

Universidade Federal de Juiz de Fora
Faculdade de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica

Lucas Meirelles Pires Deotti

**Contribuições para a Resolução do Problema de Extração dos Parâmetros do Modelo
Fotovoltaico de Um Diodo**

Juiz de Fora

2026

Lucas Meirelles Pires Deotti

Contribuições para a Resolução do Problema de Extração dos Parâmetros do Modelo Fotovoltaico de Um Diodo

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora, na área de concentração em Sistemas de Energia Elétrica, como requisito parcial para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Ivo Chaves da Silva Junior, D.Sc.

Juiz de Fora

2026

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF com os dados fornecidos pelo(a) autor(a)

Deotti, Lucas Meirelles Pires.

Contribuições para a Resolução do Problema de Extração dos Parâmetros do Modelo Fotovoltaico de Um Diodo / Lucas Meirelles Pires Deotti. – 2026.
203 f. : il.

Orientador: Ivo Chaves da Silva Junior, D.Sc.

Tese (Doutorado) – Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Engenharia Elétrica, 2026.

1. Geração Solar Fotovoltaica. 2. Modelo de Um Diodo. 3. Estimação de Parâmetros. 4. Função W de Lambert. 5. Método de Ponto Exterior. I. Silva Junior, Ivo Chaves, orient. II. Título.

Lucas Meirelles Pires Deotti

Contribuições para a Resolução do Problema de Extração dos Parâmetros do Modelo Fotovoltaico de Um Diodo

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica

Aprovada em 16 de março de 2026.

BANCA EXAMINADORA

Prof. Dr. Ivo Chaves da Silva Junior - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Edmarcio Antonio Belati

Universidade Federal do ABC

Prof. Dr. Fabiano Perin Gasparin

Universidade Federal do Rio Grande do Sul

Prof. Dr. João Alberto Passos Filho

Universidade Federal de Juiz de Fora

Prof. Dr. Bruno Henriques Dias

Universidade Federal de Juiz de Fora

Juiz de Fora, 29/01/2026.



Documento assinado eletronicamente por **Ivo Chaves da Silva Junior, Professor(a)**, em 16/03/2026, às 11:19, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Joao Alberto Passos Filho, Professor(a)**, em 16/03/2026, às 11:27, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Edmarcio Antonio Belati, Usuário Externo**, em 16/03/2026, às 11:49, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bruno Henriques Dias, Professor(a)**, em 16/03/2026, às 13:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fabiano Perin Gasparin, Usuário Externo**, em 16/03/2026, às 13:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2847088** e o código CRC **8582282B**.

À memória do estimado Professor José Luiz Rezende Pereira.

AGRADECIMENTOS

A Deus, por ter dado o sopro inicial e por sempre nos agraciar com Sua luz.

Ao Prof. José Luiz Rezende Pereira (*in memoriam*), por apontar as direções primárias para o desenvolvimento desta pesquisa; por sua dedicação admirável; e por seus conhecimentos e vivências generosamente compartilhados.

Ao Prof. Ivo Chaves da Silva Junior, por ter mantido as portas da Universidade Federal de Juiz de Fora abertas para mim; por sua orientação efetiva e por sua amizade atenciosa.

Aos meus pais, Roseli Meirelles Pires e Antônio Carlos Silveira Deotti, por todo amor e suporte. À minha irmã, Juliane Meirelles Pires Deotti, pelo amparo e incentivo. Ao meu afilhado, Rafael Deotti Freitas, por nos trazer tantas alegrias. À Sandra da Silva (Dadá), pelo carinho imensurável. À Rose Mary Silveira Diotte (*in memoriam*), minha eterna Tia Ni. A todos os meus tios e tias, primos e primas, por serem meu refúgio de diversão e proteção.

À minha noiva, Maria Carmela Kneip Lopes Fonseca, por seu amor e cumplicidade, pela força e pelas felicidades proporcionadas. À sua família, pelo grande acolhimento e apreço.

Aos meus amigos de infância, de colégio e de universidade, por todo apoio e pelos momentos de descontração. Aos colegas do Núcleo de Pesquisa em Sistemas de Potência, pela amistosa convivência. Ao Matheus Alberto de Souza, por todo o companheirismo. À Wanessa de Oliveira Guedes, pela valiosa parceria nos tempos difíceis da pandemia.

Ao corpo docente do Programa de Pós-Graduação em Engenharia Elétrica e da Faculdade de Engenharia, por contribuir para a minha formação. Ao Prof. Flavio Vanderson Gomes, por ter me apresentado à energia solar fotovoltaica. Ao Prof. Bruno Henriques Dias, por ter me encorajado a realizar pesquisas variadas. Ao Prof. Igor Delgado de Melo, por me apresentar o método de ponto interior. Ao Prof. André Luis Marques Marcato, pelas contribuições dadas à publicação de trabalhos. Ao Prof. Tiago Soares, pela disponibilidade e ensinamentos.

Aos colegas do Operador Nacional do Sistema Elétrico, em especial da Gerência de Metodologias e Modelos Energéticos, pelas experiências profissionais e cotidianas únicas. À Tatiana Frade Gonçalves Mundstock, por toda compreensão e confiança. Ao Carlos Alberto de Araujo Junior, pelos aprendizados e colaborações. Ao Paulo Sergio de Castro Nascimento, pelas conversas enriquecedoras e pelo incentivo constante.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, ao Instituto Nacional de Ciência e Tecnologia em Energia Elétrica e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais, por fornecerem os recursos necessários ao meu doutoramento.

A todos que, de algum modo, contribuíram para a consolidação da minha jornada acadêmica e àqueles que, em algum instante, expressaram votos de sucesso.

“All models are wrong, but some are useful.”

George E. P. Box (1919-2013)

RESUMO

O modelo de um diodo constitui a base físico-matemática prevalecente para a representação do comportamento elétrico de geradores fotovoltaicos, sustentando aplicações de simulação de desempenho e de caracterização de seus componentes. Seu emprego prático, contudo, requer o ajuste dos cinco parâmetros de sua equação característica. Em princípio, a forma mais consistente de realizar essa tarefa é extrair tais parâmetros a partir de uma curva corrente-tensão ($I-V$) do gerador em estudo, de modo que a equação característica reproduza da maneira mais fidedigna possível. Essa etapa conduz a um problema de otimização desafiador, marcado por não linearidade, transcendência e não convexidade, além de sensibilidade numérica e forte dependência das escolhas de formulação e da abordagem de cálculo adotada. A partir de uma revisão bibliográfica abrangente e de uma análise crítica dos tratamentos existentes, constatou-se que a abordagem formal do modelo — baseada na reescrita explícita da equação característica via função W de Lambert — combinada a uma formulação com restrições de canalização coerentes constitui a alternativa mais adequada e promissora para a obtenção de soluções de alta qualidade. Esse diagnóstico motivou a proposição de um arcabouço metodológico, desenvolvido em duas frentes complementares. A primeira dedicou-se ao cálculo da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo, culminando em um *método preditor-corretor autoiniciado*, capaz de retornar soluções com acurácia na ordem da precisão de máquina e baixo custo computacional. A segunda frente concentrou-se em uma iniciativa, ao que consta, sem precedentes na literatura: o desenvolvimento de um método determinístico de busca local, sob medida para a resolução direta do problema. Em consequência, concebeu-se o “*tailored exterior-point method*” (TEPM), cuja arquitetura combina distintos mecanismos para lidar com as dificuldades estruturais e numéricas intrínsecas ao problema, tais como barreira logarítmica modificada, regularização inercial e escalonamento simétrico de matrizes. Por meio de estudos de caso difundidos na literatura, verificou-se que ambos os métodos propostos são competitivos, destacando-se frente às alternativas de estado da arte em termos de eficácia e robustez. Adicionalmente, em uma aplicação envolvendo a resolução de mais de um milhão de problemas de extração a partir de curvas $I-V$, esses desempenhos foram corroborados: o elevado potencial de acurácia do TEPM foi evidenciado, ao passo que o método preditor-corretor autoiniciado mostrou-se pertinente ao proporcionar tempos de simulação reduzidos, compatíveis com as exigências da aplicação. Em síntese, as contribuições originais desta tese consistem sobretudo na proposição de dois métodos dedicados — o método preditor-corretor autoiniciado e o TEPM —, os quais constituem instrumentos para a resolução efetiva, reproduzível e bem fundamentada do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$, com valor direto para a comunidade científica e para o setor fotovoltaico.

Palavras-chave: Geração Solar Fotovoltaica. Modelo de Um Diodo. Estimação de Parâmetros. Função W de Lambert. Método de Ponto Exterior.

ABSTRACT

The single diode model is the leading physical–mathematical framework for representing the electrical behavior of photovoltaic generators, supporting both performance simulation and component characterization. However, practical application requires adjusting the five parameters of its characteristic equation. The most consistent approach involves extracting these parameters from a current–voltage (I – V) curve of the generator under study, ensuring the characteristic equation accurately mirrors the observed behavior. This process presents a challenging optimization problem, characterized by nonlinearity, transcendence, nonconvexity, numerical sensitivity, and a strong dependence on the chosen formulation and calculation approach. Through a comprehensive literature review and critical analysis of existing treatments, it was established that a formal approach to the model—based on the explicit rewriting of the characteristic equation using the Lambert W function—combined with a formulation incorporating coherent boundary constraints, constitutes the most appropriate and promising alternative for obtaining high-quality solutions. This insight led to the proposal of an integrated methodological framework developed along two complementary lines. The first focuses on efficient computation of the Lambert W function within the physically relevant domain of the single-diode model, culminating in a *self-started predictor–corrector method* capable of delivering solutions with machine-precision accuracy and low computational cost. The second line focused on an initiative that, to the best of the author’s knowledge, has no precedent in the literature: the proposal of a deterministic local search method specifically designed for the direct solution of the problem. Consequently, the *tailored exterior-point method* (TEPM) was conceived, whose architecture combines mechanisms specifically designed to address the structural and numerical challenges, such as a modified logarithmic barrier, inertial regularization, and symmetric matrix scaling. Through case studies widely reported in the literature, both proposed methods were shown to be highly competitive, standing out among state-of-the-art alternatives in terms of efficacy and robustness. Moreover, in a large-scale application involving the solution of more than one million parameter extraction problems from I – V curves, these performances were further corroborated: the high accuracy potential of the TEPM was evidenced, while the self-started predictor–corrector method proved suitable by providing reduced simulation times compatible with the application’s requirements. In summary, the original contributions of this thesis consist primarily in the proposal of two dedicated methods—the self-started predictor–corrector method and the TEPM—which constitute concrete tools for the effective, reproducible, and technically grounded solution of the single diode model parameter extraction problem from an I – V curve, with direct value for both the scientific community and the photovoltaic industry.

Keywords: Solar Photovoltaic Generation. Single Diode Model. Parameter Estimation. Lambert W Function. Exterior-Point Method.

LISTA DE ILUSTRAÇÕES

Figura 1 – Circuito elétrico equivalente do modelo de um diodo. Alguns termos secundários, necessários para a sua análise, encontram-se representados em cor acinzentada.	37
Figura 2 – (a) Módulo fotovoltaico composto por $N_{cel} = 36$ células conectadas em série; (b) representação do seu circuito elétrico equivalente via modelo de um diodo ($N_p = 1$ e $N_s = 36$).	40
Figura 3 – (a) Arranjo fotovoltaico formado por $N_p = 3$ strings de $N_{mod} = 3$ módulos, sendo que cada um destes é composto por $N_{cel} = 36$ células; (b) representação do seu circuito elétrico equivalente via modelo de um diodo ($N_p = 3$ e $N_s = 108$).	41
Figura 4 – Influência dos parâmetros do modelo de um diodo na simulação da curva $I-V$ do gerador fotovoltaico: (a) curva real típica; (b)-(f) efeito da variação isolada de I_{ph} , I_0 , n , R_{sh} e R_s	43
Figura 5 – Curva $I-V$ da célula fotovoltaica de tecnologia de silício cristalino RTC France.	47
Figura 6 – Esquema geral de resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico em estudo.	51
Figura 7 – Fluxograma de aplicação do método de Newton para o cálculo da equação característica do modelo de um diodo.	54
Figura 8 – Reprodução da curva $I-V$ da célula de silício RTC France, considerando as diferentes soluções ótimas alcançadas pelo nDEBCO para os cinco parâmetros do modelo de um diodo.	82
Figura 9 – Comportamento da função $f(w) = w \exp(w)$ no eixo real.	86
Figura 10 – Comportamento da função W de Lambert no eixo real, com destaque para os seus ramos analíticos.	88
Figura 11 – Magnitude do argumento z da função W de Lambert presente na equação característica do modelo de um diodo, em relação à tensão nos terminais de um gerador fotovoltaico.	91
Figura 12 – Comparação da função W de Lambert com outras funções regulares, considerando o domínio fisicamente relevante ao modelo de um diodo.	111
Figura 13 – Relação entre b^2 e $4ac$ verificada para a equação quadrática associada ao passo corretor no escopo do modelo de um diodo.	115
Figura 14 – Perfil do erro relativo absoluto (ARE) verificado em cada etapa do método proposto.	117
Figura 15 – Boxplot dos valores de ARE verificados nas soluções retornadas pelo método proposto.	120
Figura 16 – Tempos médios totais de simulação referentes aos cálculos requeridos da função W de Lambert.	121

Figura 17 – Comparação entre as curvas $I-V$ geradas em cada estudo de caso pela função lambertw do MATLAB e pelo método proposto.	124
Figura 18 – Comparação dos valores de ARE verificados para os métodos considerados na curva $I-V$ de cada estudo de caso.	125
Figura 19 – Fluxograma básico do método de ponto exterior.	137
Figura 20 – Fluxograma de otimização das variáveis primais do método de ponto exterior desenvolvido.	143
Figura 21 – Representação da lógica subjacente à condição de Armijo.	150
Figura 22 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo xSi12922, em Cocoa - FL.	176
Figura 23 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo mSi460A8, em Eugene - OR.	177
Figura 24 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo CIGS39013, em Golden - CO.	178

LISTA DE TABELAS

Tabela 1	– Relações de equivalência tipicamente adotadas para compactação da equação característica geral do modelo de um diodo.	40
Tabela 2	– Principais características das abordagens de cálculo do modelo de um diodo no escopo da função objetivo do problema de extração dos parâmetros. . . .	56
Tabela 3	– Visão geral dos métodos determinísticos já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$	60
Tabela 4	– Visão geral dos procedimentos de inicialização propostos para a aplicação de métodos determinísticos locais.	61
Tabela 5	– Visão geral dos métodos metaheurísticos já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$	66
Tabela 6	– Visão geral dos métodos conjugados já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$	71
Tabela 7	– Soluções ótimas alcançadas por diferentes métodos de otimização para o problema de extração dos parâmetros da célula RTC France, considerando a abordagem simplificada do modelo de um diodo. Note que estas estão segregadas de acordo com a classe correspondente do método.	75
Tabela 8	– Soluções ótimas alcançadas por diferentes métodos de otimização para o problema de extração dos parâmetros da célula RTC France, considerando a abordagem formal do modelo de um diodo. Note que estas estão segregadas de acordo com a classe correspondente do método.	78
Tabela 9	– Estimativas iniciais consideradas na aplicação do LMA e do TRR na resolução do problema de extração dos parâmetros da célula RTC France.	79
Tabela 10	– Valores de RMSE relativos a cada abordagem do modelo de um diodo, considerando as respectivas soluções alcançadas pelo nDEBCO.	81
Tabela 11	– Parâmetros característicos dos estudos de caso.	119
Tabela 12	– Métricas estatísticas dos valores de ARE verificados para o método proposto e para os métodos recursivos de segunda e quarta ordem.	122
Tabela 13	– Comparação dos tempos médios totais de simulação nos cálculos da função W de Lambert.	122
Tabela 14	– Valores de RMSE, dados em ampères, verificados para os métodos proposto e correlatos.	123
Tabela 15	– Comparação dos tempos médios totais de simulação registrados nos cálculos requeridos da equação característica do modelo de um diodo.	126
Tabela 16	– Parâmetros de controle do método de ponto exterior proposto.	156

Tabela 17 – Pares medidos de tensão e corrente das curvas características de cada estudo de caso.	160
Tabela 18 – Limites estabelecidos para os parâmetros do modelo de um diodo em cada estudo de caso.	160
Tabela 19 – Estimativas iniciais consideradas para cada estudo de caso.	160
Tabela 20 – Soluções retornadas pelos métodos avaliados no caso da célula RTC France.	162
Tabela 21 – Soluções retornadas pelos métodos avaliados no caso do módulo Photowatt-PWP201.	162
Tabela 22 – Indicadores de eficiência computacional verificados para cada método avaliado nos estudos de caso considerados.	164
Tabela 23 – RMSE médio e tempo médio de simulação associados às estimativas iniciais fornecidas pelo procedimento de inicialização para os problemas de extração dos parâmetros das curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.	173
Tabela 24 – Comparação do RMSE médio, dado em ampères, obtido pelos métodos avaliados na resolução dos problemas de extração dos parâmetros associados às curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.	175
Tabela 25 – Comparação do tempo de simulação médio, dado em segundos, obtido pelos métodos avaliados na resolução dos problemas de extração dos parâmetros associados às curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.	180

LISTA DE ABREVIATURAS E SIGLAS

ABC	Colônia de abelhas artificiais
ABC-TRR	Método de região de confiança reflexiva conjugado com a colônia de abelhas artificiais
ABSO	Otimização por enxame de abelhas artificiais
ARE	Erro relativo absoluto
BA	Algoritmo de morcego
BBO	Otimização baseada em biogeografia
BFA	Algoritmo de forrageamento bacteriano
BHCS	Algoritmo de busca de cuco heterogêneo baseado em biogeografia
BM	Método de barreira
BMO	Otimizador de acasalamento de pássaros
CdTe	Tecnologia fotovoltaica baseada em material de telureto de cádmio
CIGS	Tecnologia fotovoltaica baseada em material de seleneto de cobre, índio e gálio
CO	Estado norte-americano Colorado
COA	Algoritmo de coiote
CPMPSO	Algoritmo de otimização por enxame de partículas baseado em mutação por perturbação classificada
CS	Busca de cuco
CSO	Otimização por enxame de gatos
DE	Evolução diferencial
DE/WOA	Evolução diferencial híbrida com algoritmo de otimização de baleia
EABOA	Algoritmo de otimização de borboleta adaptativa aprimorado
EHA-NMS	Método simplex de Nelder–Mead com a colônia de abelhas artificiais e a estratégia de águia de três estágios
ELBA	Algoritmo de morcego aprimorado com voo de Lévy

ELPSO	Otimização por enxame de partículas com líder aprimorado
EPM	Método de ponto exterior
FE	Número de avaliações da função objetivo
FEm	Número máximo de avaliações da função objetivo
FFO	Otimizador de fertilidade das terras agrícolas
FL	Estado norte-americano da Flórida
FPA	Algoritmo de polinização de flores
GA	Algoritmo genético
GB	Gigabyte
GCPSO	Otimização por enxame de partículas com convergência garantida
GHz	Gigahertz
GOBL	Mecanismo de aprendizado baseado em oposição generalizada
GOFPANM	Método simplex de Nelder–Mead conjugado com o algoritmo de polinização de flores e o mecanismo de aprendizado baseado em oposição generalizada
GWO	Otimizador de lobo cinzento
GWOCs	Algoritmo híbrido baseado no otimizador de lobo cinzento e busca de cuco
HCS	Busca de cuco heterogênea
HISA	Algoritmo de busca interior hibridizada
HIT	Tecnologia fotovoltaica baseada em material de heterojunção intrínseca com camada fina
HS	Busca harmônica
IABBA	Algoritmo “ <i>branch-and-bound</i> ” baseado em aritmética intervalar
IEC	Comissão Eletrotécnica Internacional
IEEE	Instituto de Engenheiros Eletricistas e Eletrônicos
IGHS	Busca harmônica global inovadora
IMPA	Algoritmo de predadores marinhos melhorado
IPM	Método de ponto interior

ISA	Algoritmo de busca interior
IT	Número de iterações
ITLBO	Otimização baseada no ensino-aprendizagem melhorada
ITm	Número máximo de iterações
ImSMA	Algoritmo de bolor limoso melhorado
KKT	Condições de Karush–Kuhn–Tucker
LICQ	Qualificação de restrições por independência linear
LMA	Algoritmo de Levenberg-Marquardt
LMSA	Algoritmo de Levenberg-Marquardt conjugado com o recozimento simulado
MADE	Método simplex de Nelder–Mead conjugado com a evolução diferencial adaptativa baseada no histórico de sucesso e a estratégia de eliminação baseada em ranking
MCSWOA	Algoritmo de otimização de baleias com cruzamento assistido de estratégias de busca modificadas e operador de seleção
MFO	Otimização por atração de mariposa
MP	Ponto de máxima potência
MPSO	Versão modificada da otimização por enxame de partículas
NM-MPSO	Método simplex de Nelder–Mead conjugado com a versão modificada da otimização por enxame de partículas
NMMLP	Método de Newton modificado com um parâmetro de Levenberg
NMS	Método simplex de Nelder–Mead
NMSOLMFO	Método simplex de Nelder–Mead conjugado com a otimização por atração de mariposa e a estratégia de aprendizagem ortogonal
NREL	Laboratório Nacional de Energias Renováveis
OC	Circuito aberto
OLS	Estratégia de aprendizagem ortogonal
OR	Estado norte-americano do Oregon
PGJAYA	JAYA guiado pelo desempenho

PSO	Otimização por enxame de partículas
RAM	Memória de acesso aleatório
RD	Quantidade de aplicações repetidas
RES	Estratégia de eliminação baseada em ranking
RMSE	Raiz do erro quadrático médio
Rcr-IJADE	Evolução diferencial adaptativa reparada
SA	Recozimento simulado
SC	Curto-circuito
SDO	Otimização baseada em oferta-demanda
SFS	Otimização de busca fractal estocástica
SHADE	Evolução diferencial adaptativa baseada no histórico de sucesso
SQP	Programação sequencial quadrática
SSE	Soma dos erros quadráticos
STC	Condição padrão de teste
TC	Tempo médio de simulação para alcançar a solução final
TEPM	Método de ponto exterior sob medida
TES	Estratégia de águia de três estágios
TLABC	Algoritmo de colônia de abelhas artificiais baseado em ensino-aprendizagem
TLBO	Otimização baseada em ensino-aprendizagem
TRR	Método de região de confiança reflexiva
TS	Tempo (médio ou específico) de simulação de uma aplicação
TT	Tempo de simulação do conjunto de aplicações
WDO	Otimização orientada pelo vento
WOA	Algoritmo de otimização de baleias
aSi	Tecnologia fotovoltaica baseada em material de silício amorfo
mSi	Tecnologia fotovoltaica baseada em material de silício multicristalino

nDEBCO	Novo algoritmo de evolução diferencial híbrida e inteligência de colônia de abelhas artificiais
xSi	Tecnologia fotovoltaica baseada em material de silício monocristalino

LISTA DE SÍMBOLOS

Conjuntos:

\mathbb{C}	Conjunto dos números complexos
\mathbb{N}	Conjunto dos números naturais
\mathbb{R}	Conjunto dos números reais
\mathbb{Z}	Conjunto dos números inteiros

Escalares:

G_{sh}	Condutância de derivação (<i>shunt</i>)
$H_{j,l}$	Elemento (j,l) de uma matriz Hessiana
I	Corrente elétrica nos terminais do gerador fotovoltaico
I_0	Corrente de saturação reversa
$I_{c,j}$	Corrente elétrica calculada/modelada no j -ésimo ponto da curva $I-V$
I_{cel}	Corrente elétrica nos terminais da célula fotovoltaica
I_d	Corrente no diodo
$I_{m,j}$	Corrente elétrica medida no j -ésimo ponto da curva $I-V$
I_{ph}	Fotocorrente
I_{sh}	Corrente na ramificação de derivação (<i>shunt</i>)
N	Número de pares de medições que descreve uma curva $I-V$
N_{cel}	Número de células fotovoltaicas
N_{mod}	Número de módulos fotovoltaicos
N_p	Número de <i>strings</i> fotovoltaicas interconectadas em paralelo
N_s	Número de células fotovoltaicas interconectadas em série
R_s	Resistência série
R_{sh}	Resistência paralelo
T_c	Temperatura de operação da célula fotovoltaica ou gerador fotovoltaico
V	Tensão elétrica nos terminais de um gerador fotovoltaico

V_{cel}	Tensão elétrica nos terminais da célula fotovoltaica
$V_{m,j}$	Tensão elétrica medida no j -ésimo ponto da curva $I-V$
Ψ	Classe de funções duas vezes continuamente diferenciáveis usada no princípio do reescalonamento não linear
α	Comprimento do passo associado à direção de busca
β	Fator multiplicativo da estratégia de “ <i>backtracking</i> ”
δ	Multiplicador de Lagrange associado a uma restrição do problema transformado
ε	Tolerância numérica
γ	Parâmetro de regularização
κ	Número de condicionamento para a inversão
λ	Multiplicador de Lagrange associado às restrições do problema original
\mathcal{F}	Função objetivo de um problema de otimização
\mathcal{L}	Função Lagrangiana associada a um problema de otimização
\mathcal{L}_{epm}	Função Lagrangiana associada ao problema de otimização transformado do método de ponto exterior
\mathcal{N}	Função unidimensional auxiliar do modelo de um diodo
\mathcal{R}	Função residual associada à função W de Lambert
W	Função W de Lambert
W_{-1}	Ramo inferior da função W de Lambert
W_0	Ramo principal da função W de Lambert
W_c	Aproximação da função W de Lambert no passo corretor
W_i	Aproximação da função W de Lambert no passo inicial
W_p	Aproximação da função W de Lambert no passo preditor
μ	Parâmetro de barreira associado ao método de ponto exterior
ϕ	Variável de entrada genérica da função decorrente da equação característica do modelo de um diodo

φ	Erro entre uma aproximação e o valor exato
ψ	Função pertencente a Ψ aplicada às restrições de desigualdade no reescalonamento não linear
σ	Fator de escalonamento associado ao reescalonamento não linear
τ	Parâmetro auxiliar da regra da fração até a fronteira
v	Fração mínima do decréscimo esperado a ser satisfeita
f	Função matemática qualquer
j	Subscrito de indexação (índice j)
k	Constante de Boltzmann ($1,38065 \times 10^{-23}$ J/K)
lb	Limite inferior de uma variável de decisão
n	Fator de idealidade
q	Valor absoluto da carga típica de um elétron ($1,60218 \times 10^{-19}$ C)
t	Contador auxiliar de iteração
ub	Limite superior de uma variável de decisão
w	Variável auxiliar usada no âmbito da função W de Lambert
z	Argumento da função W de Lambert
Vetores:	
δ	Vetor de multiplicadores de Lagrange associados às restrições do problema transformado (variáveis duais)
λ	Vetor de multiplicadores de Lagrange associados ao problema original
$\mathbf{0}$	Vetor nulo
$\tilde{\mathbf{p}}$	Vetor direção de busca no espaço escalonado
\mathbf{b}	Vetor de termos independentes (quando aplicável)
\mathbf{c}	Vetor de restrições
\mathbf{e}	Vetor de uns
\mathbf{g}	Vetor gradiente da função objetivo
\mathbf{lb}	Vetor de limites inferiores

p	Vetor direção de busca
ub	Vetor de limites superiores
v	Vetor auxiliar de avaliação de restrições
x	Vetor de variáveis de decisão (variáveis primais)

Matrizes:

A	Matriz simétrica associada à formulação linearizada
C	Matriz auxiliar na formulação transformada
D	Matriz diagonal de escalonamento
H	Matriz Hessiana da função objetivo
I	Matriz identidade
J	Matriz Jacobiana

Operadores:

\approx	Operador de aproximação
\geq	Operador de desigualdade “maior ou igual”
\leq	Operador de desigualdade “menor ou igual”
\mapsto	Operador de mapeamento
$\nabla_{\mathbf{x}}$	Operador gradiente com respeito a \mathbf{x}
$\nabla_{\mathbf{xx}}$	Operador Hessiano com respeito a \mathbf{x}
\odot	Operador de multiplicação elemento a elemento
Im	Operador de parte imaginária
diag	Operador de matriz diagonal
T	Operador de transposição

Sub/sobrescritos:

max	Subscrito que indica o valor máximo associado à grandeza em análise
min	Subscrito que indica o valor mínimo associado à grandeza em análise

SUMÁRIO

1 – INTRODUÇÃO	25
1.1 CONTEXTUALIZAÇÃO	25
1.2 DELIMITAÇÃO DO TEMA	26
1.3 MOTIVAÇÃO DA PESQUISA	29
1.4 JUSTIFICATIVA ESTRATÉGICA	31
1.5 OBJETIVOS DO TRABALHO	34
1.6 PUBLICAÇÕES DECORRENTES	35
1.7 ESTRUTURA DO DOCUMENTO	35
2 – REVISÃO BIBLIOGRÁFICA	37
2.1 MODELO DE UM DIODO	37
2.1.1 Formulação básica	38
2.1.2 Escalonamento estrutural	39
2.1.3 Parâmetros característicos	42
2.1.4 Aspectos matemáticos	44
2.2 PROBLEMA DE EXTRAÇÃO DOS PARÂMETROS	45
2.2.1 Formulação geral	46
2.2.2 Classificação matemática	49
2.2.3 Processo de resolução	50
2.3 ABORDAGENS DE CÁLCULO DA FUNÇÃO OBJETIVO	52
2.3.1 Abordagem simplificada	52
2.3.2 Abordagem formal	53
2.3.3 Comentários adicionais	55
2.4 MÉTODOS DE OTIMIZAÇÃO DISPONÍVEIS	57
2.4.1 Métodos determinísticos	57
2.4.2 Métodos metaheurísticos	64
2.4.3 Métodos conjugados	70
2.4.4 Observações complementares	72
2.5 ANÁLISE COMPREENSIVA	74
2.5.1 Resultados verificados na abordagem simplificada	74
2.5.2 Resultados verificados na abordagem formal	77
2.5.3 Discussão	80
2.6 CONCLUSÕES PARCIAIS	82

3 – FORMALIZAÇÃO DO PROBLEMA DE EXTRAÇÃO DOS PARÂMETROS	85
3.1 ABORDAGEM DO MODELO DE UM DIODO VIA FUNÇÃO W DE LAMBERT	85
3.1.1 Função W de Lambert	85
3.1.2 Equação característica explícita	89
3.1.3 Domínio fisicamente relevante	90
3.2 ANÁLISE DO PROBLEMA DE OTIMIZAÇÃO	92
3.2.1 Formulação adotada	92
3.2.2 Função Lagrangiana	94
3.2.3 Condições de otimalidade	94
3.2.4 Avaliação de convexidade	97
3.3 QUESTÕES NUMÉRICAS	98
3.3.1 Escala das variáveis	99
3.3.2 Condicionamento	100
3.4 CONCLUSÕES PARCIAIS	102
4 – MÉTODO PROPOSTO PARA O CÁLCULO DA FUNÇÃO W DE LAMBERT	104
4.1 PROSPECÇÃO METODOLÓGICA	104
4.1.1 Abordagens de cálculo numérico	104
4.1.2 Método híbrido	107
4.1.3 Método numérico	107
4.1.4 Método explícito	108
4.1.5 Análise compreensiva	109
4.2 MÉTODO PROPOSTO	110
4.2.1 Passo inicial	111
4.2.2 Passo preditor	112
4.2.3 Passo corretor	113
4.2.4 Consolidação matemática	115
4.2.5 Implementação computacional	117
4.3 RESULTADOS EXPERIMENTAIS	118
4.3.1 Avaliação de efetividade	120
4.3.2 Avaliação de pertinência	121
4.3.3 Avaliação de competitividade	122
4.4 CONCLUSÕES PARCIAIS	126
5 – MÉTODO PROPOSTO PARA A RESOLUÇÃO DO PROBLEMA	128
5.1 PROSPECÇÃO METODOLÓGICA	128
5.1.1 Avaliação preliminar	128
5.1.2 Métodos de barreira	130
5.1.3 Princípio do reescalonamento não linear	132

5.2	MÉTODO PROPOSTO	135
5.2.1	Problema transformado	135
5.2.2	Estrutura lógica	136
5.2.3	Método de Newton escalonado e regularizado	142
5.2.4	Racionalização da arquitetura	152
5.2.5	Consolidação do algoritmo	154
5.2.6	Implementação computacional	157
5.3	RESULTADOS EXPERIMENTAIS	159
5.3.1	Avaliação de acurácia	161
5.3.2	Avaliação de eficiência	163
5.4	CONCLUSÕES PARCIAIS	164
6	– APLICAÇÃO PRÁTICA	166
6.1	CARACTERIZAÇÃO DO ESTUDO	166
6.1.1	Dados utilizados	166
6.1.2	Formulação do problema	167
6.1.3	Recurso computacional	168
6.1.4	Metodologia de avaliação	168
6.2	PROCEDIMENTOS AUXILIARES	169
6.2.1	Restrições de canalização	169
6.2.2	Procedimento de inicialização	171
6.2.3	Validação preliminar	172
6.3	SÍNTESE DOS RESULTADOS	173
6.3.1	Níveis de acurácia	174
6.3.2	Qualidade das soluções	175
6.3.3	Quesitos de eficiência	179
6.4	CONCLUSÕES PARCIAIS	180
7	– CONSIDERAÇÕES FINAIS	182
7.1	CONCLUSÕES	182
7.2	PERSPECTIVAS DE CONTINUIDADE DA PESQUISA	186
	REFERÊNCIAS BIBLIOGRÁFICAS	188
	APÊNDICE A – FORMA EXPLÍCITA DO MODELO DE UM DIODO	201

1 INTRODUÇÃO

A tese de doutorado registrada neste documento encontra-se inserida em um campo interdisciplinar de pesquisa, centrado na modelagem do desempenho de geradores fotovoltaicos operando em regime permanente, com ênfase nos métodos numéricos requeridos para a sua implementação prática.

Neste capítulo introdutório, apresenta-se o contexto no qual se desenvolve esta tese e delimita-se o seu tema de pesquisa. Em seguida, explicitam-se as motivações e justificativas que fundamentam a sua realização. À luz do exposto, formalizam-se os objetivos definidos para a condução das investigações e para a consequente produção de contribuições originais. Adicionalmente, listam-se as publicações já alcançadas em decorrência dos resultados obtidos. Ao final, descreve-se sucintamente a estrutura capitular que compõe o restante deste documento.

1.1 CONTEXTUALIZAÇÃO

As constantes evoluções tecnológicas, o crescimento da cooperação internacional e o agravamento de diferentes questões político-ambientais transformaram a geração de energia elétrica a partir de fontes renováveis, tais como hídrica, eólica, solar e biomassa, em uma solução padrão para o desenvolvimento sustentável da sociedade contemporânea [1, 2]. Dentre essas fontes e as suas respectivas tecnologias de conversão, a chamada *geração solar fotovoltaica* é aquela que vem se apresentando como a mais promissora [2–5].

Em linhas gerais, essa tecnologia é materializada por *sistemas fotovoltaicos conectados à rede elétrica* [2, 3]. Esses sistemas são essencialmente constituídos por um gerador fotovoltaico e um inversor [6]. O primeiro é tipicamente formado por um arranjo de módulos fotovoltaicos interconectados entre si [7]. Cada módulo é, por sua vez, composto por um certo número de células fotovoltaicas idênticas, as quais constituem as unidades responsáveis por converter a energia eletromagnética contida na radiação proveniente do Sol em energia elétrica [6]. Todavia, como tal conversão se traduz sob a forma de corrente elétrica contínua, cumpre ao inversor condicioná-la eletronicamente à forma de corrente elétrica alternada, com as mesmas características fundamentais de frequência e formato de onda daquela estabelecida na rede elétrica local [8, 9]. Dado o seu caráter de interface, o inversor também assume as responsabilidades de otimizar, supervisionar e garantir a operação segura do sistema [6–9].

A ampla disponibilidade do recurso solar, aliada à modularidade inerente dos sistemas fotovoltaicos conectados à rede elétrica e à sua operação livre de emissões de poluentes, torna tal solução excepcionalmente atraente tanto para aplicações residenciais urbanas quanto para projetos em escala de utilidade em áreas remotas [3, 10, 11]. Ademais, o fato de esses sistemas virem apresentando um custo relativo cada vez mais competitivo, bem como o de constituir a alternativa de geração renovável que mais cria oportunidades de emprego, representa um diferencial estratégico frente aos tomadores de decisão e aos agentes políticos [5, 12].

Em decorrência desses fatores, verifica-se que, nos últimos anos, a expansão anual da capacidade instalada de sistemas fotovoltaicos no mundo vem superando expressivamente a de qualquer outra tecnologia de geração, incluindo aquelas baseadas em combustíveis fósseis e energia nuclear [1]. Além disso, há uma forte expectativa de que tal protagonismo se mantenha ao menos pelo restante da década [1–5].

Apesar desse cenário favorável, há uma série de questões que ainda precisam ser superadas para que a geração solar fotovoltaica sustente essa trajetória de crescimento e amplie sua participação na matriz elétrica mundial. Por um lado, identificam-se barreiras regulatórias e entraves comerciais que têm prejudicado a viabilidade de investimento em novos sistemas fotovoltaicos conectados à rede elétrica em diferentes países [4, 13, 14]. Por outro, constata-se que o próprio aumento da penetração desses sistemas nas redes de transmissão e distribuição de energia elétrica tem aumentado sobremaneira as suas complexidades operacionais [11, 15, 16].

Sob uma perspectiva técnica, essa conjuntura indica a proliferação de oportunidades e desafios no setor de energia. Ainda assim, essas facetas somente poderão ser gerenciadas de forma eficaz se houver profissionais qualificados, infraestrutura adequada e ferramentas computacionais especializadas para a condução das atividades de engenharia [3, 10, 12, 17].

1.2 DELIMITAÇÃO DO TEMA

Dado o presente contexto, a modelagem do desempenho de sistemas fotovoltaicos conectados à rede elétrica torna-se crucial para a análise técnica e econômica de novas unidades a serem instaladas, bem como daquelas que já se encontram em operação [10, 17–19]. Em aplicações mais rigorosas de engenharia, essa tarefa costuma ser conduzida por meio de uma cadeia de modelos intercambiáveis, capaz de estimar, tanto em quantidade quanto em tempo, a produção de energia elétrica do sistema [10, 18, 20–22].

Nessa cadeia, o modelo destinado a representar o desempenho do gerador fotovoltaico destaca-se como o elemento mais importante [10, 23]. Em linhas gerais, esse modelo busca quantificar a fração da irradiação solar incidente que é efetivamente convertida em energia elétrica, considerando a influência de diversos fatores, sobretudo a temperatura de operação das células fotovoltaicas que o compõem [23, 24]. Logo, os resultados fornecidos por este configuram estimativas da máxima produção que se pode esperar do sistema em análise [7, 10, 18]. Adicionalmente, tais resultados podem servir como referência para a avaliação da integridade física e da conformidade operacional dos próprios componentes do gerador [25–31].

Na literatura, diversos modelos de desempenho de geradores fotovoltaicos encontram-se disponíveis [23, 24, 32]. Em meio a estes, o chamado *modelo de um diodo* apresenta-se como o mais prevalente [23]. A sua adoção em obras consagradas, como [6], [7] e [33], reflete a sua consistência conceitual. Ademais, o extenso conjunto de trabalhos publicados ao longo das últimas décadas ao seu respeito evidencia a sua relevância no meio acadêmico [34, 35].

Paralelamente, a implementação do modelo de um diodo no cerne de ferramentas computacionais de uso difundido na indústria fotovoltaica corrobora o seu valor prático no meio profissional. Nesse sentido, destaca-se a sua exploração primária no PVSYSY, *software* comercial de projeto e simulação de sistemas fotovoltaicos, frequentemente tido como referência em nível mundial [20, 36, 37]. Vale também mencionar a sua incorporação no SYSTEM ADVISOR MODEL (SAM), *software* gratuito voltado à simulação técnico-econômica de projetos de geração conectados à rede, mantido pela renomada instituição norte-americana *National Renewable Energy Laboratory* (NREL) [22, 38].

A rigor, o modelo de um diodo é classificado como um modelo físico, o qual busca descrever os principais atributos e processos internos que definem o comportamento elétrico de uma célula fotovoltaica operando em regime permanente, por meio do estabelecimento de um circuito elétrico equivalente [6, 10, 33]. Embora a concepção dessa estrutura tenha sido orientada por aspectos associados às primeiras gerações de células de silício cristalino, as experiências acumuladas na literatura, bem como na indústria, mostram que esta permanece efetiva na representação de dispositivos mais modernos, com construções físicas mais elaboradas e materiais diversificados, sendo adaptações pontuais necessárias apenas em casos específicos [20, 23, 33, 39].

Além disso, é essencial frisar que, como a célula fotovoltaica constitui a unidade elementar de módulos fotovoltaicos e, conseqüentemente, de seus arranjos, o modelo de um diodo pode ser propriamente empregado para representar geradores de qualquer porte [23, 40]. A partir da análise do circuito resultante, é possível obter uma equação que permite estimar a intensidade da corrente elétrica (I) produzida pelo gerador em estudo como função da tensão (V) simultaneamente verificada em seus terminais [10, 23]. Em termos estritos, é essa equação que formaliza o modelo de um diodo e possibilita a sua aplicação em estudos e simulações [20, 23, 34, 35].

Apesar de todas essas qualidades positivas, a implementação prática do modelo de um diodo não é, de fato, uma tarefa simples. Inegociavelmente, existem duas questões associadas à sua equação característica que precisam ser devidamente tratadas. A primeira diz respeito à sua natureza transcendente [41, 42]. Em outras palavras, o termo I aparece implicitamente em tal equação, de modo que $I = f(V, I)$ [41]. Dessa forma, faz-se mandatário adotar alguma abordagem matemática que viabilize o seu cálculo [43, 44].

A segunda questão concerne à sua estrutura paramétrica. Em princípio, existem cinco parâmetros nessa equação que precisam ser propriamente determinados, a saber: *fotocorrente*, *corrente de saturação reversa*, *fator de idealidade*, *resistência paralelo* e *resistência série* [6, 23, 33]. Em conjunto, esses parâmetros constituem o mecanismo pelo qual o modelo de um diodo se adapta às particularidades construtivas e operacionais de um gerador fotovoltaico e, por conseguinte, se torna capaz de representar o seu comportamento elétrico [23].

Assim, a forma mais precisa de determinar os parâmetros do modelo de um diodo consiste em *extrair* os seus valores diretamente de uma *curva corrente-tensão* ($I-V$) levantada experimentalmente nos terminais do gerador fotovoltaico [34, 45]. A depender do escopo do

estudo e dos recursos disponíveis, tais curvas podem ser registradas em laboratório, com auxílio de simuladores solares, ou em campo, por meio de traçadores de curva $I-V$ [10, 46–51]. Em geral, adotam-se cuidados para assegurar que o gerador se encontre submetido a condições tão homogêneas quanto possível de irradiância solar incidente e temperatura de operação [6, 10, 23, 50]. Como resultado, obtém-se uma curva $I-V$ que descreve o comportamento elétrico do gerador sob uma condição específica de operação [10, 50, 51]. Ou, para ser mais exato, obtém-se um conjunto de dezenas a milhares de pares discretos de medição (V, I) , distribuídos ao longo dessa curva [10, 50, 51].

A imposição simultânea desses pares de medições (V, I) na equação característica do modelo de um diodo conduz naturalmente a um sistema sobredeterminado de equações, o qual é usualmente tratado na forma de um problema de otimização [34, 35]. Embora esse problema tenha apenas cinco variáveis de decisão — isto é, os cinco parâmetros do modelo de um diodo — e seja comparável ao bem conhecido problema dos mínimos quadrados, a sua natureza multidimensional, contínua, restrita, não linear e não convexa torna a sua resolução particularmente difícil [52, 53]. Alguns autores chegam a apontar, inclusive, que esse problema pode ser interpretado como um legítimo *benchmark* para a avaliação do desempenho de métodos de otimização [53].

Cumprir observar que as soluções obtidas por meio desse procedimento de extração não possuem caráter universal [23, 54]. De fato, o conjunto de medições (V, I) disponível reflete o desempenho real do gerador em um dado instante, condicionado ao seu nível de degradação física, a eventuais falhas localizadas e a não idealidades construtivas [55–58]. Além disso, esses pontos operativos registrados decorrem da resposta do gerador, em tal estado físico, a uma combinação bem definida de irradiância solar incidente sobre as superfícies coletoras das células que o compõem e de suas respectivas temperaturas de operação [54, 59]. Logo, os parâmetros estimados a partir desse conjunto tornam o modelo de um diodo apto a reproduzir o comportamento elétrico do gerador fotovoltaico primordialmente nas mesmas condições de operação em que as medições foram propriamente realizadas [23].

Em diversos estudos, essa singularidade dos parâmetros do modelo de um diodo é interpretada como uma espécie de assinatura elétrica do gerador fotovoltaico em um dado instante de tempo e sob uma condição específica de operação [25–27, 31, 60]. A título de exemplo, em campanhas voltadas à avaliação dos impactos da degradação física do gerador sobre o seu desempenho, é prática comum registrar curvas $I-V$ de forma sistemática ao longo do tempo [28, 29, 60]. Para cada curva registrada, podem então ser extraídos os valores dos cinco parâmetros do modelo de um diodo [26, 28]. Desse modo, a comparação temporal desses valores — com destaque para aqueles associados às resistências em série e em paralelo — permite identificar tendências compatíveis com fenômenos de degradação e alterações estruturais dos dispositivos de conversão [27–29, 31]. Tais conhecimentos, por sua vez, oferecem subsídios para o diagnóstico do estado do gerador e para a tomada de decisão em manutenção e operação [28, 31].

A lógica descrita no exemplo anterior também se aplica a iniciativas de monitoramento operacional e diagnóstico de falhas em tempo real, nas quais curvas $I-V$ podem ser registradas de forma recorrente e os parâmetros do modelo de um diodo podem ser acompanhados como indicadores elétricos do comportamento esperado e de seus desvios [30,31]. Não obstante, é pertinente notar que a realização da extração desses parâmetros ao longo de séries extensas de curvas $I-V$ impõe requisitos severos de eficiência computacional [29, 30]. Caso contrário, o esforço de processamento associado à resolução reiterada do problema de otimização correspondente pode tornar esses estudos impraticáveis [29].

Em complemento, é pertinente mencionar que a singularidade dos parâmetros do modelo de um diodo não necessariamente impõe uma limitação ao seu emprego em um escopo mais amplo. Uma vez conhecidos os seus valores para uma condição de referência, torna-se possível, em certos casos, recorrer às chamadas equações de translação [23]. Em linhas gerais, essas equações permitem estimar os valores que seriam assumidos por cada parâmetro sob condições arbitrárias, tomando justamente como base aqueles previamente obtidos a partir de uma curva $I-V$ levantada em condição de operação conhecida [34, 59, 61, 62]. Portanto, se o problema de extração dos parâmetros for resolvido com êxito, o modelo de um diodo torna-se capaz de emular o comportamento elétrico do gerador fotovoltaico em diferentes condições de irradiância solar incidente e temperatura de operação que sejam de interesse [20, 36, 61].

Diante dessas considerações, depreende-se que o *problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ do gerador fotovoltaico em estudo* constitui um tema de pesquisa genuíno e de relevância preponderante para a indústria fotovoltaica.

1.3 MOTIVAÇÃO DA PESQUISA

Seja pela sua relevância para a indústria fotovoltaica, seja pela sua natureza matematicamente desafiadora, fato é que, nos últimos anos, inúmeros métodos de otimização têm sido propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico [34, 35]. Em meio a estes, identificam-se *métodos determinísticos*, tanto baseados em estratégias de *busca local* quanto em estratégias de *busca global* [52–54, 63–68]. Contudo, verifica-se que os *métodos metaheurísticos* são os mais explorados por pesquisadores interessados no tema [43, 69–103]. Há ainda uma terceira classe que vem ganhando notoriedade: a dos *métodos conjugados*. Em particular, esses últimos decorrem de diferentes arranjos de métodos determinísticos e metaheurísticos [104–111].

Em linhas gerais, o processo de validação desses métodos de otimização envolve a resolução de problemas de extração dos parâmetros¹ bem difundidos na literatura, tais como aqueles originalmente apresentados em [63], [99] e [112]. A ideia central por trás dessa prática é verificar em que medida a solução retornada por um método proposto permite que o modelo

¹ Por brevidade, o problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ será, em geral, referido daqui em diante somente como “problema de extração dos parâmetros”.

de um diodo reproduza a curva $I-V$ respectiva ao problema [85]. Se constatado que esse nível de acurácia é igual ou superior àqueles previamente reportados para métodos de estado da arte, pode-se então concluir que o método proposto cumpre efetivamente o seu objetivo [35, 85].

Essa prática de validação é constatada na grande maioria dos trabalhos disponíveis sobre o tema — veja [67] ou [85], por exemplo. Consequentemente, verifica-se que métodos determinísticos, metaheurísticos e conjugados, mesmo em versões mais recentes e avançadas, tendem formalmente a apresentar níveis similares de acurácia quanto à resolução do problema de extração dos parâmetros [35, 43, 52, 65, 66, 85, 109, 110]. O mesmo, porém, não pode ser afirmado sobre as suas respectivas eficiências. A esse respeito, verifica-se que os métodos determinísticos de busca local — ou, simplesmente, *métodos determinísticos locais* — costumam apresentar os melhores indicadores: em geral, demandam um número reduzido de iterações e dezenas ou centenas de avaliações da função objetivo para alcançar a solução ótima do problema [53, 54, 64–67]; enquanto métodos conjugados podem demandar até milhares de iterações e avaliações [106–110]; e métodos metaheurísticos, até dezenas de milhares [35, 43, 84, 85, 89, 91, 93, 94, 101, 103]. Ademais, enquanto métodos de busca local tipicamente despendem no máximo alguns segundos para atingir tal solução, métodos determinísticos de busca global podem despende horas [52, 54].

Além do elevado nível de eficiência, os métodos determinísticos locais apresentam atributos que os tornam particularmente adequados a aplicações que exigem alto nível de confiabilidade [45]. Mais especificamente, o fato de esses métodos retornarem sempre a mesma solução ótima para uma dada estimativa inicial faz com que seus resultados sejam, em princípio, reproduzíveis [113]. Por outro lado, a natureza probabilística dos métodos metaheurísticos e conjugados dificulta que se ofereça, na prática, o mesmo nível de transparência [114]. Adicionalmente, diversos métodos determinísticos locais oferecem, ainda que localmente, garantias teóricas de otimalidade das soluções retornadas [113]. Quando comparados a métodos metaheurísticos e à maior parte dos métodos conjugados, os quais não dispõem de qualquer garantia intrínseca, esses métodos se destacam como as alternativas conceitualmente mais consistentes [114].

Não obstante, subsiste uma questão estrutural que carece de maior debate. Conforme destacado anteriormente, a equação característica do modelo de um diodo apresenta a peculiaridade de ser transcendente em relação à variável que representa a corrente produzida pelo gerador fotovoltaico [41]. Essa variável, por sua vez, é precisamente aquela manipulada no âmbito da função objetivo do problema de extração dos parâmetros [43]. Assim, torna-se compulsória, para a resolução do problema, a adoção de alguma abordagem matemática que viabilize o cálculo da equação característica do modelo de um diodo [43, 43, 52, 63, 67, 85, 101, 103, 110].

Ao analisar a vasta quantidade de trabalhos associados à proposição de métodos de otimização para a resolução do problema de extração dos parâmetros, verifica-se que duas abordagens são recorrentemente adotadas. A mais comum, ora denominada *abordagem simplificada*, busca descaracterizar a natureza transcendente da equação característica por meio da introdução de

aproximações razoáveis [43, 52, 63, 85, 103, 110]. A outra, ora denominada *abordagem formal*, procura lidar estritamente com essa natureza por meio de técnicas apropriadas, tais como o *método de Newton* e a *função W de Lambert* [43, 67, 101].

De modo geral, observa-se que a escolha entre essas abordagens para o cálculo da equação característica tende a seguir a orientação estabelecida pelas principais referências adotadas pelos pesquisadores na condução de seus estudos [67, 85]. Ainda assim, o trabalho apresentado em [43] distingue-se por considerar, de forma pioneira, ambas as abordagens. Notadamente, os resultados reportados nesse trabalho revelaram que as soluções ótimas obtidas a partir de cada abordagem, para um mesmo problema de extração dos parâmetros, são distintas [43]. Com efeito, essa situação indica que a dinâmica imposta à função objetivo pela abordagem simplificada pode diferir significativamente daquela imposta sob a abordagem formal, de modo que, na prática, cada alternativa conduz a uma formulação de otimização efetivamente distinta. Ademais, tais resultados evidenciaram que a adoção de uma ou outra abordagem influencia o desempenho de diferentes métodos de otimização, tanto em termos do número de avaliações da função objetivo quanto do tempo computacional [43].

Essas constatações sinalizam, portanto, a necessidade de que estudos dedicados à proposição e à validação de métodos de otimização para a resolução do problema de extração dos parâmetros passem a tratar a escolha da abordagem de cálculo da equação característica do modelo de um diodo de forma mais criteriosa, atentando para seus possíveis impactos sobre os resultados obtidos e as conclusões consolidadas.

1.4 JUSTIFICATIVA ESTRATÉGICA

Em termos estritos, observa-se que os trabalhos contemporâneos que tratam da resolução do problema de extração dos parâmetros por meio de métodos determinísticos locais são bastante limitados, tanto em quantidade quanto em escopo.

Se, por um lado, encontram-se na literatura dezenas de estudos regulares dedicados a métodos metaheurísticos, por outro, identificam-se poucas referências consistentes que envolvem métodos determinísticos locais, a saber: [53], [54], [64], [65], [66] e [67]. Além disso, esse pequeno conjunto de trabalhos segue uma estrutura bastante similar. Inicialmente, propõe-se um método original para a resolução aproximada do problema de extração dos parâmetros [53, 54, 64–67]. Em seguida, a solução aproximada retornada por esse método é submetida a um processo recorrentemente chamado de “refinamento” [53, 54, 64–67]. Esse processo, entretanto, consiste em utilizar tal solução como estimativa inicial para a aplicação de um método determinístico local, disponibilizado por ferramentas computacionais dedicadas — isto é, *solvers* padrão — na resolução do próprio problema de extração dos parâmetros [53, 54, 64–67].

Dado esse panorama, depreende-se que esse conjunto de trabalhos converge, em última instância, para a proposição de procedimentos de inicialização voltados à aplicação de métodos

determinísticos locais, e não para o desenvolvimento de um método especificamente concebido para a resolução direta do problema de extração dos parâmetros. Mais precisamente, suas contribuições avançam na etapa de inicialização, mas não no enfrentamento sistemático das dificuldades matemáticas intrínsecas ao problema.

É essencial ponderar que essa lógica de investigação centrada na etapa de inicialização é, até certo ponto, justificável. Em princípio, a qualidade da estimativa inicial considerada na aplicação de um método determinístico local constitui um fator crítico para a resolução efetiva de um problema de extração dos parâmetros, sobretudo em virtude de sua natureza não convexa [113, 115]. Dessa forma, é não apenas razoável, mas necessário que se direcionem esforços à elaboração de procedimentos de inicialização qualificados.

Adicionalmente, é importante reconhecer que o desenvolvimento de um novo método determinístico local, ou mesmo a implementação criteriosa de uma alternativa semelhante já existente, não é uma tarefa trivial [113]. Diferentemente dos métodos metaheurísticos, os métodos determinísticos locais possuem um arcabouço matemático denso e um grau significativamente menor de liberdade para customização [113, 114, 116]. Não bastasse isso, esses métodos frequentemente se deparam com questões numéricas sensíveis [113, 117–119]. Assim, o tratamento adequado de métodos determinísticos locais exige do pesquisador um conhecimento aprofundado tanto das características matemáticas do problema quanto das nuances do próprio método.

Não obstante, o interesse limitado da comunidade científica em investigar a aplicação de métodos determinísticos locais na resolução direta do problema de extração dos parâmetros — ou, em outras palavras, em aprimorar meticulosamente o próprio processo de refinamento — tem acarretado lacunas substanciais na literatura e perpetuado certas inconsistências.

Sob a perspectiva de otimização, ressalta-se o fato de que o referido conjunto de trabalhos recorreu a métodos determinísticos locais disponibilizados em funções especiais do *software* MATLAB [120, 121]. Mais precisamente, [53] emprega o *algoritmo de Levenberg–Marquardt* (“*Levenberg–Marquardt algorithm*”, LMA), enquanto [54], [64], [65], [66] e [67] utilizam o *algoritmo de região de confiança reflexiva* (“*trust-region-reflective*”, TRR). Ou seja, apenas dois métodos determinísticos locais foram recentemente testados para a resolução do problema de extração dos parâmetros, em contraste com dezenas de métodos metaheurísticos e diversos métodos conjugados. Diante desse quadro, um questionamento que naturalmente emerge é se não haveria outro método determinístico local capaz de oferecer desempenho superior — seja em termos de acurácia, seja em algum quesito de eficiência — a esses dois até então formalmente avaliados na literatura.

Já sob a perspectiva da modelagem, destaca-se o fato de que esse conjunto de trabalhos considerou uma formulação irrestrita do problema de extração dos parâmetros, desconsiderando as suas genuínas restrições de canalização [53, 54, 64–67]. Uma possível explicação para essa prática reside nas limitações intrínsecas dos métodos empregados. Por exemplo, o LMA foi originalmente desenvolvido para a resolução de problemas de otimização irrestritos [113]. Ademais,

a versão do MATLAB disponível à época da publicação de [53] ainda não contemplava uma variante desse método munida de mecanismos adicionais para lidar com restrições de canalização, como ocorre em versões mais recentes [120]. Outra explicação possível decorre da crença de que as estimativas iniciais retornadas pelos procedimentos de inicialização propostos estariam suficientemente próximas da solução ótima global, de modo que o método convergiria para esta de forma inequívoca. Todavia, é fundamental ponderar que, na ausência de restrições, qualquer método de otimização está sujeito a convergir para uma solução que pode ser rigorosamente infactível, dada a natureza não convexa do problema [113, 114, 116]. Logo, uma dúvida que imediatamente surge é se não seria mais adequado — ou mesmo mais seguro — considerar explicitamente tais restrições na formulação do problema de extração dos parâmetros.

Por fim, sob a perspectiva estrutural, chama a atenção o fato de que o mesmo conjunto de trabalhos adotou uma abordagem formal para o cálculo da equação característica do modelo de um diodo, baseada na função W de Lambert [53, 54, 64–67]. Embora se trate de uma alternativa moderna e comprovadamente efetiva, a avaliação do valor dessa função especial constitui, por si só, um tópico em aberto de pesquisa [122, 123]. Nesse cenário, verifica-se que os métodos numéricos dedicados a esse propósito variam de forma significativa em termos de acurácia e de eficiência [41, 42]. Tal variabilidade é particularmente crítica no âmbito do problema de extração dos parâmetros, pois pode gerar desvios na função objetivo, condicionar o custo por avaliação da mesma e, em casos adversos, afetar a confiabilidade do processo de otimização [44]. Ainda assim, nota-se que essa questão é apenas marginalmente abordada nos trabalhos referidos, a ponto de muitos não explicitarem nem justificarem a alternativa metodológica adotada para a computação da função W de Lambert. Diante dessa conjuntura, uma indagação que se impõe é em que medida a escolha do método dedicado à computação da função W de Lambert influencia o desempenho de métodos determinísticos locais e condiciona a própria resolução do problema de extração dos parâmetros.

A comunhão dessas considerações indica que seria não apenas pertinente, mas também inovador, investigar de forma aprofundada a aplicação de um método determinístico local na resolução do problema de extração dos parâmetros, incorporando explicitamente as restrições de canalização e adotando a abordagem formal do modelo de um diodo via função W de Lambert.

Nesse sentido, a literatura clássica [113] e as consolidações mais recentes apresentadas em [116] e [124] apontam que *métodos de programação sequencial quadrática* (“*sequential quadratic programming*”, SQP) e *métodos de barreira* (“*barrier methods*”, BM) representam o estado da arte em otimização não linear restrita. Estes últimos, em especial, apresentam afinidade pronunciada com problemas de otimização cujas restrições são expressas por desigualdades, como é o caso das restrições de canalização no problema de extração dos parâmetros [113, 116]. Além disso, observa-se que um número expressivo de variantes bem-sucedidas de BMs já foi proposto para a resolução de problemas de otimização não convexos, com destaque para: [125], [126], [127], [128], [129] e [130]. Portanto, uma investigação mais focada no emprego de um

BM para a resolução do problema de extração dos parâmetros mostra-se uma alternativa ao mesmo tempo adequada e promissora.

Em síntese, a literatura recente evidencia uma lacuna clara: embora a etapa de inicialização tenha recebido atenção considerável, ainda não há uma proposta consolidada de método determinístico local que seja (i) dedicado à resolução do problema de extração dos parâmetros, (ii) trate, de forma apropriada, das restrições de canalização e (iii) incorpore a abordagem formal do modelo de um diodo via função W de Lambert, compreendendo claramente as implicações dessa escolha. A principal inovação desta tese consiste em propor e demonstrar, de maneira sistemática, um BM concebido especificamente para esse contexto, acompanhado de implementação criteriosa e validação abrangente, oferecendo, assim, uma alternativa determinística local mais acurada, robusta e tecnicamente justificável para a extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico em estudo.

1.5 OBJETIVOS DO TRABALHO

À luz do exposto, estabelece-se como objetivo geral da pesquisa a proposição de um novo método determinístico local para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico em estudo, tendo em vista a formulação restrita do problema e a abordagem formal da equação característica do modelo via função W de Lambert.

Para o cumprimento desse objetivo geral, definem-se os seguintes objetivos específicos:

1. Analisar o problema de extração dos parâmetros como um problema de otimização, caracterizando sua natureza matemática e suas dificuldades intrínsecas, bem como os impactos decorrentes da adoção da abordagem formal do modelo de um diodo via função W de Lambert na sua estrutura.
2. Examinar as particularidades da função W de Lambert e as estratégias numéricas disponíveis na literatura específica, a fim de propor um método eficaz e eficiente para o seu cálculo no escopo da equação característica do modelo de um diodo.
3. Desenvolver um método de barreira que permita a resolução acurada e robusta do problema de extração dos parâmetros, considerando as suas genuínas restrições de canalização.
4. Indicar procedimentos que permitam formular adequadamente as restrições de canalização do problema de extração dos parâmetros, bem como gerar estimativas iniciais com a qualidade necessária para a aplicação efetiva do método de barreira desenvolvido.
5. Validar esse arcabouço metodológico concebido por meio da resolução de problemas de extração dos parâmetros bem difundidos na literatura e em estudos de caso associados a demandas práticas da indústria fotovoltaica.

1.6 PUBLICAÇÕES DECORRENTES

Alguns dos estudos conduzidos com vistas ao cumprimento dos objetivos definidos para esta pesquisa já foram devidamente sintetizados e publicados em periódicos internacionais indexados. Em particular, destacam-se os trabalhos listados abaixo:

- DEOTTI, L. M. P.; JUNIOR, I. C. da S. A self-started predictor–corrector method for calculating the Lambert W function within the scope of the photovoltaic single diode model. *Solar Energy*, Elsevier, v. 276, p. 112681, 2024.
- DEOTTI, L. M. P.; JUNIOR, I. C. da S. A survey on the parameter extraction problem of the photovoltaic single diode model from a current–voltage curve. *Solar Energy*, Elsevier, v. 263, p. 111930, 2023.
- DEOTTI, L. M. P.; PEREIRA, J. L. R.; JUNIOR, I. C. da S. Parameter extraction of photovoltaic models using an enhanced Lévy flight bat algorithm. *Energy Conversion and Management*, Elsevier, v. 221, p. 113114, 2020.

Adicionalmente, encontra-se em fase final de elaboração um artigo intitulado “*Parameter extraction of the photovoltaic single diode model using a tailored exterior-point method*”. À luz de seu escopo e das contribuições obtidas, pretende-se submetê-lo a um periódico de reconhecida relevância na área, como *Energy Conversion and Management* ou *Solar Energy*.

Além dessas publicações em periódicos e do manuscrito em preparação, parte dos resultados desta pesquisa também foi divulgada em evento científico nacional. Mais especificamente, o trabalho indicado a seguir foi apresentado oralmente no *LIV Simpósio Brasileiro de Pesquisa Operacional (SBPO 2022)*, realizado entre os dias 8 e 11 de novembro de 2022, na cidade de Juiz de Fora, Minas Gerais:

- DEOTTI, L. M. P.; JUNIOR, I. C. da S. Avaliação dos impactos produzidos pelas principais abordagens de cálculo do modelo fotovoltaico de um diodo no seu problema de determinação dos parâmetros. In: *LIV Simpósio Brasileiro de Pesquisa Operacional*. [S.l.: s.n.], 2022.

1.7 ESTRUTURA DO DOCUMENTO

Este documento está organizado em seis capítulos, seguidos de um capítulo de conclusões e de um apêndice técnico, estruturados de modo a conduzir o leitor desde a identificação da lacuna científica até a consolidação e validação do arcabouço metodológico proposto.

O Capítulo 2 apresenta um levantamento bibliográfico e uma revisão crítica do problema de extração dos parâmetros do modelo de um diodo a partir de curvas $I-V$ de um gerador fotovoltaico. A princípio, são discutidas a equação característica, a formulação típica do problema

e as principais abordagens de cálculo adotadas na literatura, distinguindo-se explicitamente as vertentes simplificada e formal. À luz dessa distinção, são documentados dezenas de trabalhos representativos segundo diferentes classes de métodos de otimização, bem como segundo a abordagem de cálculo empregada em seus testes de validação, permitindo uma análise comparativa do estado da arte em termos de acurácia, eficiência e maturidade tecnológica. Por fim, as constatações obtidas evidenciam lacunas metodológicas que fundamentam os desenvolvimentos apresentados nos capítulos subsequentes.

O Capítulo 3 formaliza o problema de extração dos parâmetros sob a abordagem formal via função W de Lambert, adotando uma formulação restrita fisicamente consistente. Em seguida, examinam-se as propriedades matemáticas do problema resultante, com destaque para a sua natureza não convexa, e investigam-se aspectos numéricos decisivos para a sua resolução, em particular as discrepâncias de escala e o potencial mal condicionamento das matrizes envolvidas. Com isso, estabelecem-se as bases conceituais que orientam o desenvolvimento de métodos dedicados nos capítulos subsequentes.

O Capítulo 4 dedica-se ao estudo do cálculo da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo. A partir de uma prospecção metodológica das alternativas existentes, é proposto um *método preditor–corretor autoiniciado*, concebido para combinar elevada acurácia com baixo custo computacional, viabilizando a adoção sistemática da abordagem formal do modelo de um diodo no problema de extração dos parâmetros.

O Capítulo 5 apresenta a principal contribuição metodológica da tese: o desenvolvimento do *tailored exterior-point method* (TEPM), um método determinístico local estruturado sob o paradigma de métodos de barreira e concebido sob medida para a formulação formal e restrita estabelecida. Nesse sentido, são detalhados seus fundamentos, sua arquitetura algorítmica e sua validação frente a alternativas consolidadas.

O Capítulo 6 consolida o arcabouço metodológico proposto por meio de sua aplicação em um estudo massivo envolvendo mais de um milhão de curvas I - V , representativo de cenários de interesse prático. Os resultados obtidos confirmam a robustez, a acurácia e a viabilidade prática das estratégias desenvolvidas.

O Capítulo 7 sintetiza os principais resultados alcançados, estabelece a correspondência entre os objetivos delineados e as contribuições efetivamente obtidas, e discute as implicações científicas e tecnológicas do arcabouço proposto, além de indicar perspectivas de continuidade da pesquisa.

Por fim, o Apêndice A apresenta desenvolvimentos matemáticos elementares associados à formulação via função W de Lambert, oferecendo suporte técnico adicional às análises conduzidas ao longo do texto.

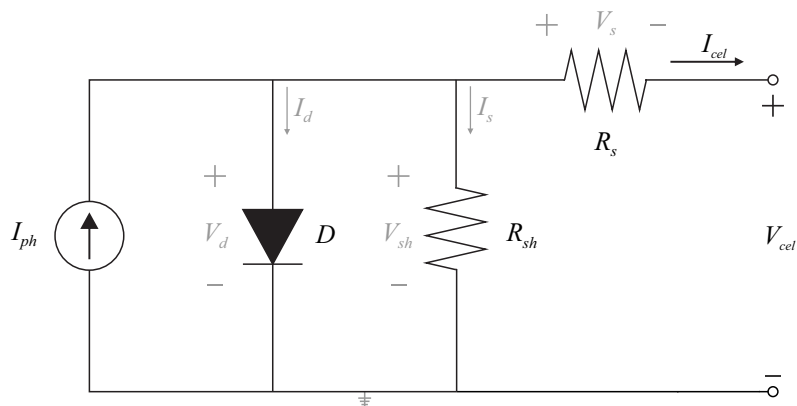
2 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta um levantamento bibliográfico acerca do tema de pesquisa desta tese. Inicialmente, a Seção 2.1 introduz o modelo de um diodo e desenvolve a sua equação característica. Em seguida, a Seção 2.2 descreve a formulação típica do problema de extração dos parâmetros dessa equação e esclarece questões fundamentais relacionadas à sua resolução. Uma vez construída essa base, a Seção 2.3 detalha as abordagens mais comumente adotadas para viabilizar o cálculo da equação característica no escopo da função objetivo do problema. A Seção 2.4, por sua vez, documenta alguns dos mais notáveis métodos de otimização já propostos para a sua resolução, organizando-os conforme as suas classes e de acordo com a abordagem de cálculo da equação característica considerada em seus respectivos testes de validação. Na sequência, a Seção 2.5 realiza uma análise compreensiva do desempenho desses métodos, bem como dos impactos produzidos por cada abordagem de cálculo na solução final do problema. Por fim, a Seção 2.6 apresenta uma síntese do conteúdo exposto ao longo do capítulo e explicita as principais lacunas identificadas na literatura sobre o tema de pesquisa.

2.1 MODELO DE UM DIODO

O chamado *modelo de um diodo* é caracterizado pelo circuito elétrico equivalente apresentado na Figura 1 [23]. Conforme se observa, esse circuito é composto por uma fonte de corrente ideal em antiparalelo com um diodo (D), um resistor paralelo e um resistor série. Cada um desses componentes representa um aspecto fundamental que define o comportamento elétrico de uma célula fotovoltaica operando em regime permanente e, em última instância, o seu desempenho. Discussões mais detalhadas a respeito podem ser encontradas em [6], [7] ou [131].

Figura 1 – Circuito elétrico equivalente do modelo de um diodo. Alguns termos secundários, necessários para a sua análise, encontram-se representados em cor acinzentada.



A equação que descreve o comportamento elétrico nos terminais desse circuito e, conseqüentemente, caracteriza o modelo de um diodo é desenvolvida a seguir. Posteriormente, introduz-se a sua generalização para representar geradores fotovoltaicos de maior escala, com ênfase nos aspectos mais relevantes de seus parâmetros físicos e de sua estrutura matemática.

2.1.1 Formulação básica

Ao aplicar a Lei de Kirchhoff das Correntes na análise (em corrente contínua) do circuito apresentado na Figura 1, encontra-se que a corrente elétrica nos terminais de uma célula fotovoltaica, I_{cel} , é dada por (2.1). Nessa equação, I_{ph} representa a *fotocorrente*, ou seja, a corrente elétrica produzida pela célula fotovoltaica a partir da absorção de uma parcela da radiação eletromagnética contida na irradiância solar incidente sobre a sua superfície [6]. Já I_d representa a corrente elétrica que flui através do diodo ou, mais especificamente, da junção pn que caracteriza a célula fotovoltaica em estudo. Por fim, I_{sh} representa a corrente elétrica que flui através do resistor paralelo.

$$I_{cel} = I_{ph} - I_d - I_{sh} \quad (2.1)$$

Em particular, a corrente elétrica que flui através da junção pn da célula fotovoltaica pode ser descrita pela bem estabelecida equação de Shockley, expressa em (2.2). De maneira mais específica, I_0 representa a *corrente de saturação reversa*, isto é, a corrente que flui em polarização reversa devido à geração térmica de portadores de carga (elétrons e lacunas) na junção pn [131]. Por sua vez, n representa o *fator de idealidade*, um índice adimensional representativo do quão ideal é o transporte de portadores de carga na junção pn [131]. Complementarmente, V_d corresponde à tensão elétrica estabelecida sobre a junção pn , enquanto T_c exprime a sua *temperatura de operação* ou, para fins práticos, a temperatura de operação da própria célula fotovoltaica (em kelvin) [131]. Os outros dois termos, k e q , são, respectivamente, a constante de Boltzmann ($1,38065 \times 10^{-23}$ J/K) e o valor absoluto da carga do elétron ($1,60218 \times 10^{-19}$ C).

$$I_d = I_0 \left[\exp \left(\frac{qV_d}{nkT_c} \right) - 1 \right] \quad (2.2)$$

A tensão elétrica estabelecida sobre a junção pn pode ser obtida a partir da aplicação da Lei de Kirchhoff das Tensões ao circuito e da Lei de Ohm ao resistor série. Dessa forma, chega-se às relações apresentadas em (2.3), em que V_{cel} representa a tensão elétrica entre os terminais da célula fotovoltaica, V_s corresponde à queda de tensão sobre o resistor série e R_s é a *resistência série* que representa as perdas ôhmicas internas da célula fotovoltaica [33].

$$V_{cel} = V_d - V_s = V_d - I_{cel}R_s \quad \therefore \quad V_d = V_{cel} + I_{cel}R_s \quad (2.3)$$

Uma vez que a tensão elétrica estabelecida sobre o resistor paralelo é a mesma que se verifica na junção pn , conclui-se que $V_{sh} = V_d$ e, portanto, esta pode ser igualmente descrita por (2.3). Assim, ao se aplicar a Lei de Ohm ao resistor paralelo, obtêm-se as relações reunidas em (2.4), em que R_{sh} denota a *resistência paralelo* que representa as perdas internas da célula fotovoltaica associadas a correntes de fuga.

$$I_{sh} = \frac{V_{sh}}{R_{sh}} = \frac{V_d}{R_{sh}} \quad \therefore \quad I_{sh} = \frac{V_{cel} + I_{cel}R_s}{R_{sh}} \quad (2.4)$$

Ao incorporar (2.2), (2.3) e (2.4) em (2.1), encontra-se então que a corrente elétrica nos terminais de uma célula fotovoltaica em estudo é mais propriamente expressa por (2.5).

$$I_{cel} = I_{ph} - I_0 \left[\exp \left(\frac{q(V_{cel} + I_{cel}R_s)}{nkT_c} \right) - 1 \right] - \frac{V_{cel} + I_{cel}R_s}{R_{sh}} \quad (2.5)$$

Por explicitar os aspectos físicos que definem o comportamento elétrico de uma célula fotovoltaica, (2.5) é essencialmente tida como a *equação característica* do modelo de um diodo.

2.1.2 Escalonamento estrutural

A célula fotovoltaica é o dispositivo que, de forma elementar, compõe um módulo fotovoltaico e, conseqüentemente, os seus arranjos [6]. Sendo assim, o modelo de um diodo pode ser empregado de forma apropriada para representar um gerador de qualquer escala [23].

Para a maioria dos estudos, é suficiente considerar que o gerador fotovoltaico em questão apresenta aspectos construtivos próximos do ideal e se encontra sujeito a condições de operação homogêneas [23, 40]. Nesse sentido, considera-se tipicamente que todas as células que o compõem são idênticas e estão sujeitas às mesmas condições ambientais [23]. Além disso, assume-se que essas células estão interconectadas segundo uma estrutura matricial $N_s \times N_p$, tal como ilustrado nas Figuras 2 e 3 [23, 40]. Nessas circunstâncias, a corrente elétrica nos terminais de um gerador fotovoltaico genérico pode ser expressa por (2.6), sendo $I = N_p \cdot I_{cel}$ e $V = N_s \cdot V_{cel}$ [40].

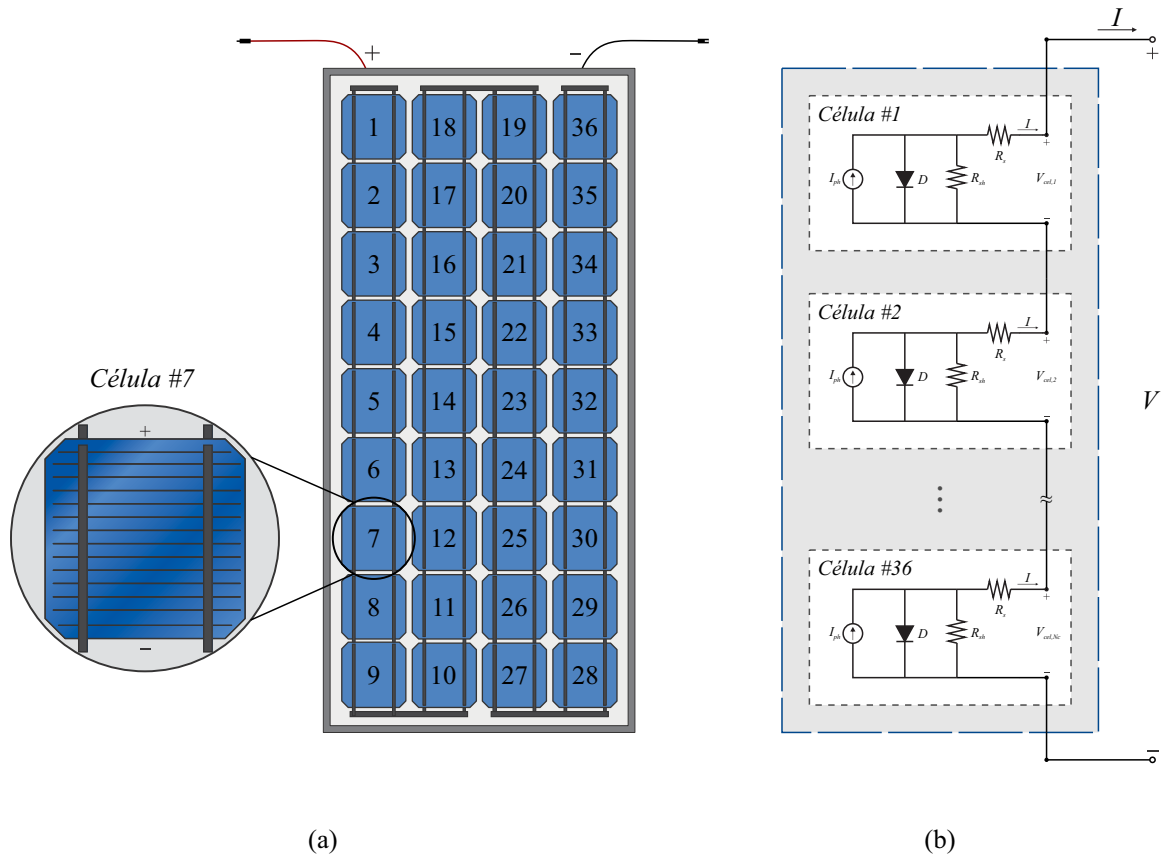
$$I = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V + I(N_s/N_p)R_s)}{N_s nkT_c} \right) - 1 \right] - \frac{V + I(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} \quad (2.6)$$

Dessa forma, para representar uma célula fotovoltaica por meio de (2.6), basta definir $N_p = N_s = 1$ (um), de modo que tal expressão se torna exatamente igual a (2.5) [40]. Já para representar um módulo fotovoltaico, deve-se definir $N_p = 1$ (um) e N_s igual ao número N_{cel} de células conectadas em série que o compõem, isto é, $N_s = N_{cel}$ [40]. Em relação aos arranjos fotovoltaicos, é necessário lembrar que estes são regularmente formados por uma ou mais fileiras, isto é, *strings*, igualmente constituídas por um número N_{mod} de módulos fotovoltaicos semelhantes interconectados em série [7, 23, 40]. Logo, para representar um arranjo, deve-se definir N_p igual ao número de *strings* que o formam e N_s igual ao produto do número de células que compõem uma unidade do módulo fotovoltaico considerado pelo número de módulos que constituem regularmente uma *string*, isto é, $N_s = N_{cel} \cdot N_{mod}$ [40]. Observe que essas definições de N_p e N_s estão explicitadas nas legendas das Figuras 2 e 3 para fins de esclarecimento.

Cumprе mencionar que essa equação característica geral do modelo de um diodo pode ser reescrita de forma mais compacta, interpretando alguns de seus termos que se interrelacionam como *parâmetros concentrados* [40, 61]. Ao aplicar as equivalências apresentadas na Tabela 1 sobre (2.6), por exemplo, obtém-se (2.7) [40, 61].

$$I = I'_{ph} - I'_0 \left[\exp \left(\frac{V + IR'_s}{a} \right) - 1 \right] - \frac{V + IR'_s}{R'_{sh}} \quad (2.7)$$

Figura 2 – (a) Módulo fotovoltaico composto por $N_{cel} = 36$ células conectadas em série; (b) representação do seu circuito elétrico equivalente via modelo de um diodo ($N_p = 1$ e $N_s = 36$).



Fonte: a ilustração (a) foi adaptada de [6].

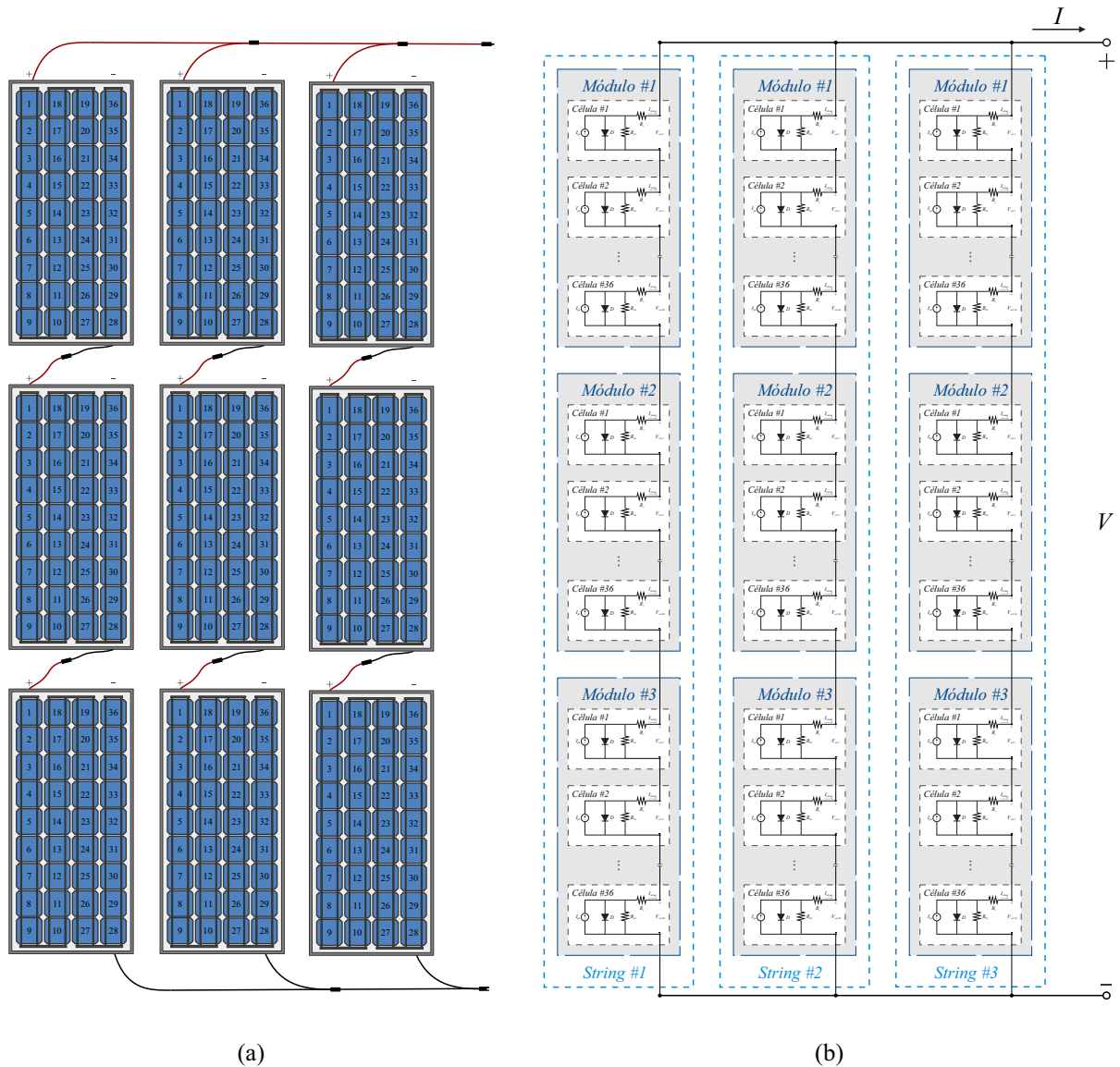
Tabela 1 – Relações de equivalência tipicamente adotadas para compactação da equação característica geral do modelo de um diodo.

Termos concentrados	Equivalência
I'_{ph}	$N_p I_{ph}$
I'_0	$N_p I_0$
a	$(N_s n k T_c) / q$
R'_s	$(N_s / N_p) R_s$
R'_{sh}	$(N_s / N_p) R_{sh}$

Como se observa, (2.7) preserva a estrutura matemática da equação característica do modelo de um diodo, sendo, contudo, mais enxuta do que a formulação geral dada em (2.6). Uma particularidade interessante, que muitas vezes se mostra conveniente, é que (2.7) dispensa o conhecimento explícito de N_p e N_s [42]. Em contrapartida, justamente por ser insensível a esses termos de conexão série–paralelo, essa formulação oferece um menor grau de detalhamento, o que pode não ser desejável em certos tipos de estudos, especialmente aqueles que dependem de análises de sensibilidade em relação à configuração física do gerador [6, 7, 23].

Diante dessas considerações, verifica-se que a escolha entre (2.6) e (2.7) como formulação primária da equação característica do modelo de um diodo varia amplamente nos trabalhos da literatura [6, 23, 40, 42, 61, 85, 131]. Em muitos casos, essa decisão decorre dos dados de que

Figura 3 – (a) Arranjo fotovoltaico formado por $N_p = 3$ strings de $N_{mod} = 3$ módulos, sendo que cada um destes é composto por $N_{cel} = 36$ células; (b) representação do seu circuito elétrico equivalente via modelo de um diodo ($N_p = 3$ e $N_s = 108$).



Fonte: a ilustração (a) foi adaptada de [6].

se dispõe para o gerador fotovoltaico em estudo, de restrições práticas associadas ao nível de detalhamento, do tratamento das incertezas experimentais ou, ainda, da conveniência algébrica de cada forma. De todo modo, ambas as expressões são amplamente aceitas e empregadas na literatura como pontos de partida equivalentes para a modelagem e a análise do gerador.

Posto isto, faz-se necessário registrar que, ao longo deste documento, todos os desenvolvimentos matemáticos decorrentes do modelo de um diodo serão conduzidos tomando (2.6) como sua equação característica. Embora mais onerosa, entende-se que tal formulação é mais fiel ao propósito físico do modelo e, por isso, capaz de oferecer um valor analítico mais imediato às aplicações consideradas. Ademais, reforça-se que quaisquer expressões daí deduzidas podem ser reescritas sem o conhecimento explícito de N_p e N_s , recorrendo às relações providas na Tabela 1.

2.1.3 Parâmetros característicos

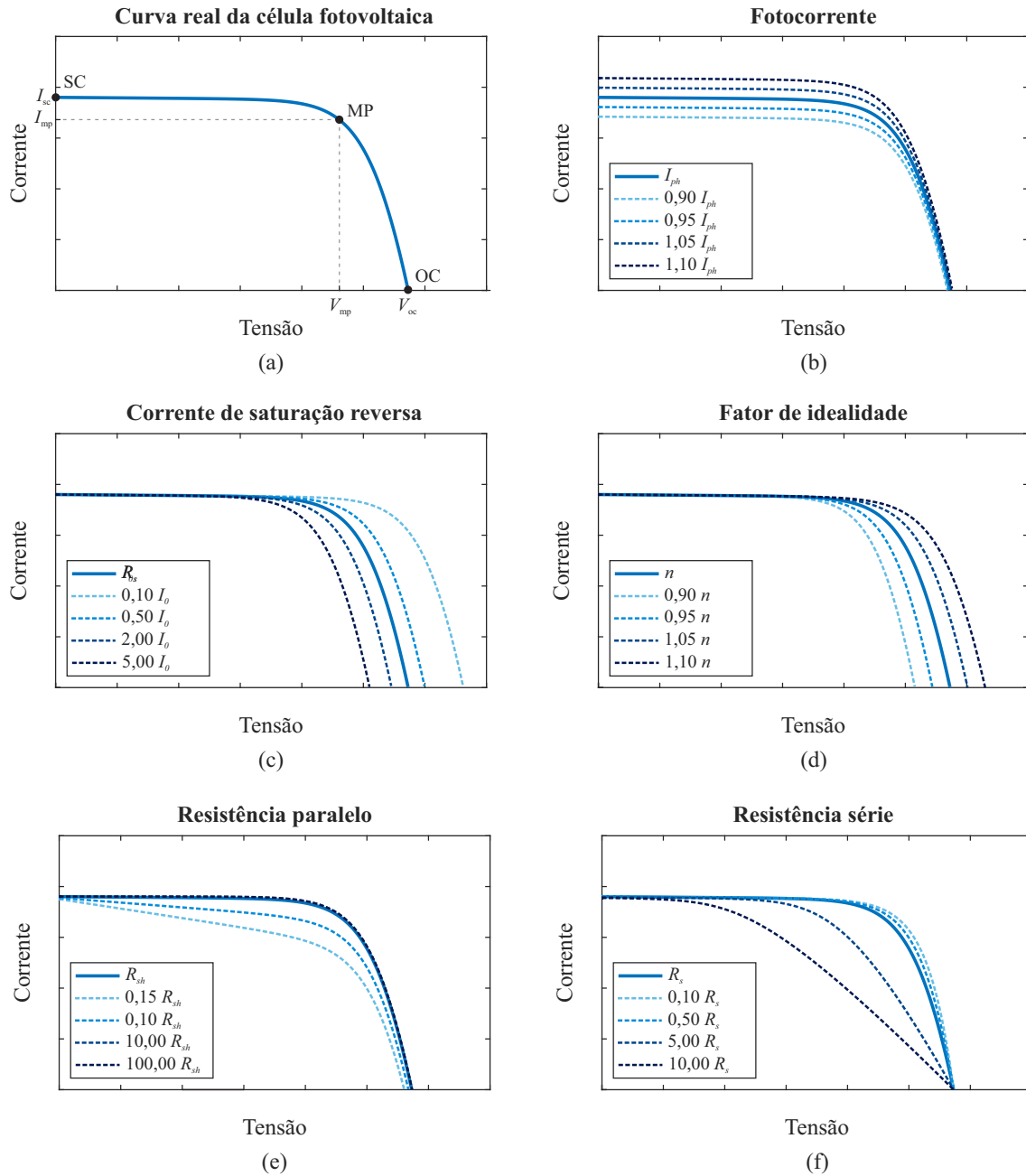
Ao analisar os termos que constituem a equação característica do modelo de um diodo, dada por (2.6), observa-se que: (i) k e q são constantes físicas, cujos valores são essencialmente fixos; (ii) N_p e N_s são dados tipicamente conhecidos do gerador fotovoltaico em estudo; (iii) I e V são grandezas elétricas que podem ser prontamente medidas em seus terminais com o auxílio de instrumentos como multímetros e traçadores de *curva corrente-tensão* (I - V); e (iv) T_c é uma grandeza térmica que pode ser adequadamente estimada por meio de termopares e câmeras infravermelhas [10, 48–50, 132–136]. Assim, restam apenas cinco termos efetivamente desconhecidos nessa equação ou, em outras palavras, cinco *parâmetros característicos*: I_{ph} , I_0 , n , R_{sh} e R_s [23].

Em conjunto, esses cinco parâmetros constituem o mecanismo pelo qual o modelo de um diodo se ajusta às particularidades construtivas e operacionais de um gerador fotovoltaico e, por conseguinte, se torna capaz de representar matematicamente o seu comportamento elétrico [23]. Para fins de esclarecimento, apresenta-se na Figura 4 um panorama das contribuições individuais de cada parâmetro nesse processo.

Em particular, o gráfico (a) da Figura 4 apresenta o formato típico da curva I - V de uma célula fotovoltaica padrão — e, por extensão, de um gerador fotovoltaico. Complementarmente, destacam-se nessa curva três pontos notáveis de operação elétrica: o ponto de *curto-circuito* (“*short-circuit*”, SC), o ponto de *circuito aberto* (“*open circuit*”, OC) e o ponto de *máxima potência* (“*maximum power*”, MP) [6, 7, 23, 131]. Note-se que, no ponto de curto-circuito, a corrente é máxima e a tensão é nula, configurando o par $(0, I_{sc})$; enquanto, no ponto de circuito aberto, ocorre o inverso, com tensão máxima e corrente nula, configurando o par $(V_{oc}, 0)$. Em ambos os casos, portanto, a potência transferida do módulo fotovoltaico para uma carga ou outro componente eventualmente acoplado a seus terminais é nula. Por sua vez, o ponto de máxima potência caracteriza a situação oposta, na qual essa transferência é efetivamente máxima. Esse ponto está associado a valores específicos de tensão e corrente, de modo que o par (V_{mp}, I_{mp}) tipicamente se localiza no “joelho” da curva I - V [6, 7, 23, 131].

Em princípio, essa curva I - V pode ser reproduzida de forma fiel pela equação característica do modelo de um diodo, desde que se ajustem adequadamente os valores de seus cinco parâmetros. Posto isso, os demais gráficos da Figura 4 ilustram os efeitos provocados pela variação isolada de cada parâmetro, mantendo-se os demais fixos em valores ideais. Conforme se observa no gráfico (b), I_{ph} atua principalmente na translação vertical da curva simulada, guardando estreita relação com a corrente de curto-circuito (I_{sc}). Já nos gráficos (c) e (d), verifica-se que I_0 e n afetam predominantemente a posição horizontal da curva simulada, estando ambos fortemente associados à tensão de circuito aberto (V_{oc}). Por sua vez, o gráfico (e) indica que R_{sh} influencia sobretudo a inclinação da curva na região próxima ao curto-circuito, ao passo que o gráfico (f) evidencia que R_s afeta a inclinação da curva na região próxima ao circuito aberto. Assim, a combinação apropriada dos valores desses parâmetros permite que a equação

Figura 4 – Influência dos parâmetros do modelo de um diodo na simulação da curva $I-V$ do gerador fotovoltaico: (a) curva real típica; (b)-(f) efeito da variação isolada de I_{ph} , I_0 , n , R_{sh} e R_s .



Fonte: curva real da célula fotovoltaica baseada nas medições reportadas em [63].

característica do modelo de um diodo assimile, com boa precisão, os aspectos observados na curva $I-V$ real de um gerador fotovoltaico em estudo, a ponto de se tornar apta a reproduzi-la de forma fidedigna em análises de engenharia [20, 23, 26, 31, 61].

Convém destacar, entretanto, que esses parâmetros geralmente assumem valores dentro de faixas e ordens de grandeza bem conhecidas. Mais expressamente, I_{ph} tende naturalmente a ser próximo do valor da corrente de curto-circuito da célula fotovoltaica que compõe o gerador na condição de operação considerada [23, 59]. Quanto a I_0 , a literatura indica que tal parâmetro tradicionalmente assume valores em ordens de grandeza reduzidas, tais como

alguns microampères ou nanoampères [23, 137]. A respeito de n , tem-se que o seu valor é fundamentalmente próximo de 1 (um) para células fotovoltaicas cujo transporte interno de portadores de carga alcança comportamento ideal, próximo de 2 (dois) quando esse transporte ocorre de maneira mais problemática e ainda mais elevado em situações excepcionais [131, 138, 139]. No que se refere a R_{sh} , esta pode assumir valores da ordem de dezenas a milhares de ohms, dependendo das características construtivas das células que compõem o gerador fotovoltaico e da condição de operação considerada [23, 54]. Já R_s dificilmente atinge a ordem de uma unidade de ohm [23, 54].

2.1.4 Aspectos matemáticos

O processo de adaptação do modelo de um diodo ao gerador fotovoltaico em estudo e, posteriormente, seu emprego em análises de engenharia dependem, em última instância, da possibilidade de calcular de forma consistente as soluções associadas à sua equação característica [20, 23, 26, 31, 61]. Nesse sentido, compreender os principais aspectos matemáticos de (2.6) constitui um requisito crucial para antecipar dificuldades de cálculo e, sobretudo, para embasar a escolha entre tratamentos numéricos e analíticos adequados.

Ao se inspecionar (2.6), observa-se que sua estrutura combina termos algébricos e um termo exponencial, caracterizando uma relação *não linear* entre corrente e tensão. Tal característica é compatível com a morfologia típica da curva $I-V$ observada em geradores fotovoltaicos, conforme ilustrado na Figura 4. Além disso, ao se examinar a participação de cada parâmetro, nota-se que algumas dependências são relativamente diretas, ao passo que outras introduzem não linearidades essenciais. Por exemplo, I_{ph} aparece apenas no termo aditivo $N_p I_{ph}$, de modo que a equação é afim em I_{ph} . De maneira análoga, mantidos fixos I , V , n , R_s , R_{sh} e T_c , I_0 atua como fator multiplicativo do colchete que contém o termo exponencial, resultando em uma dependência linear, no sentido estrito da expressão. Em contrapartida, n figura no denominador do argumento do termo exponencial, de modo que variações em seu valor alteram a contribuição desse termo de forma intrinsecamente não linear. Já R_{sh} aparece no denominador da parcela racional, implicando uma dependência propriamente não linear — embora esta se torne linear caso o problema seja reparametrizado em termos da condutância $G_{sh} = 1/R_{sh}$.

Por sua vez, R_s merece atenção especial, pois participa de dois mecanismos distintos. Por um lado, aparece no numerador do argumento do termo exponencial, influenciando diretamente sua contribuição; por outro, figura na parcela racional associada a R_{sh} . Esse duplo papel acopla os componentes exponencial e algébrico em (2.6), o que é determinante tanto para o comportamento global do modelo quanto para a dificuldade de manipulação de sua equação característica.

De fato, é pertinente notar que tanto R_s quanto I e V aparecem simultaneamente dentro e fora do termo exponencial em (2.6), o que inviabiliza o isolamento de qualquer um desses termos por meio de operações algébricas elementares. Essa condição implica que a equação característica do modelo de um diodo é *transcendente* com respeito a tais termos. Em consequência, sempre

que um desses termos assume o papel de incógnita em um problema, torna-se necessário recorrer a abordagens matemáticas específicas para viabilizar a obtenção da solução desejada [43, 44].

A despeito disso, sob a perspectiva computacional, (2.6) apresenta uma propriedade favorável: para valores fisicamente admissíveis de I_{ph} , I_0 , n , R_{sh} , R_s e dos demais termos envolvidos, trata-se de uma expressão composta por operações algébricas elementares e pela função exponencial e , portanto, continuamente diferenciável nas variáveis e parâmetros de interesse. Essa *suavidade* de (2.6) fornece, portanto, uma base propícia para a aplicação de métodos que exploram derivadas, seja no cálculo numérico de suas soluções, seja na resolução de problemas de otimização formulados a partir do modelo de um diodo [44, 53, 54, 64–67, 95–97, 99–101].

2.2 PROBLEMA DE EXTRAÇÃO DOS PARÂMETROS

A determinação dos parâmetros do modelo de um diodo pode ser realizada por diferentes procedimentos, cuja escolha depende essencialmente do conjunto de informações disponível sobre o gerador fotovoltaico em estudo — ou, de modo mais determinante, do tipo e da quantidade de dados de que se dispõe a respeito de uma de suas curvas $I-V$.

Por exemplo, em estudos voltados ao projeto de novos sistemas fotovoltaicos conectados à rede elétrica, o gerador fotovoltaico ainda está em fase de especificação e, por isso, é comum que as únicas informações acessíveis sejam aquelas fornecidas na ficha técnica dos módulos cotados para a sua composição [10, 20, 22, 36, 61, 140]. Apesar de valiosas, tais informações tipicamente se restringem aos três pontos notáveis (SC, OC e MP) de uma curva $I-V$ obtida em laboratório sob condição padrão de teste (“*standard test condition*”, STC)¹. Nessas circunstâncias, a determinação dos parâmetros do modelo de um diodo não pode ser realizada apenas pela imposição direta desses poucos pontos em sua equação característica, sendo necessário empregar recursos auxiliares que permitam fechar o problema, seja por meio de um sistema determinado de equações heterogêneas, seja por formulações independentes para a estimação direta de certos parâmetros [39, 43, 49, 61, 63, 85, 140, 142–145].

Em contraste, em estudos que permitem o ensaio experimental do gerador fotovoltaico — como campanhas voltadas à avaliação de impactos de degradação física e ao diagnóstico de falhas em tempo real —, ou quando se dispõe, por qualquer via, de um conjunto suficientemente extenso e bem distribuído de pontos de uma de suas curvas $I-V$, a situação é substancialmente distinta [25–31]. Em particular, a imposição simultânea dos diversos pares de medições conhecidos na equação característica do modelo de um diodo conduz naturalmente a um sistema sobredeterminado, usualmente tratado na forma de um problema de otimização [34, 35, 113, 116].

¹ Conforme estabelecido pela norma IEC 61215:2005, as condições padrão de teste são definidas por um nível de irradiância de 1.000 W/m^2 , conforme o espectro global padrão de massa de ar 1,5, incidente de maneira normal à superfície coletora do dispositivo fotovoltaico em avaliação, estando este a uma temperatura uniforme de 25°C [141].

Conseqüentemente, os métodos de determinação dos parâmetros aplicados nesse âmbito se concentram em técnicas de otimização e em procedimentos numéricos desenvolvidos especificamente para possibilitar a eficaz *extração* dos seus valores a partir da curva $I-V$ [35, 52, 68, 85].

Embora os métodos orientados apenas pelos pontos notáveis da curva $I-V$ se destaquem pela simplicidade e, em muitos casos, pelo reduzido custo computacional, tais abordagens apresentam vulnerabilidades intrínsecas que limitam sua adoção quando se dispõe de um conjunto mais extenso e representativo de medições ao longo da curva. Mais precisamente, as abordagens que recorrem a sistemas de equações heterogêneas tendem a admitir múltiplas soluções [139, 140]. Em outras palavras, diferentes combinações de parâmetros podem reproduzir satisfatoriamente os pontos SC, OC e MP, mas resultar em comportamentos significativamente distintos ao longo do restante da curva $I-V$ [139, 140]. Por outro lado, as abordagens baseadas em formulações independentes tendem a produzir soluções pouco acuradas ou mesmo fisicamente inadmissíveis, quando aplicadas a curvas fora do STC, além de, frequentemente, se mostrarem numericamente instáveis [39]. Assim, sempre que uma curva $I-V$ estiver efetivamente disponível e não houver limitações adicionais impeditivas, a prática mais consistente reside em extrair os valores dos parâmetros do modelo de um diodo diretamente do conjunto completo de pontos disponíveis, de modo a maximizar a aderência do modelo ao comportamento elétrico conhecido do gerador fotovoltaico em estudo [34, 45].

Posto isso, o problema de otimização tipicamente formulado para permitir a extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ do gerador fotovoltaico em estudo é detalhadamente discutido, a seguir.

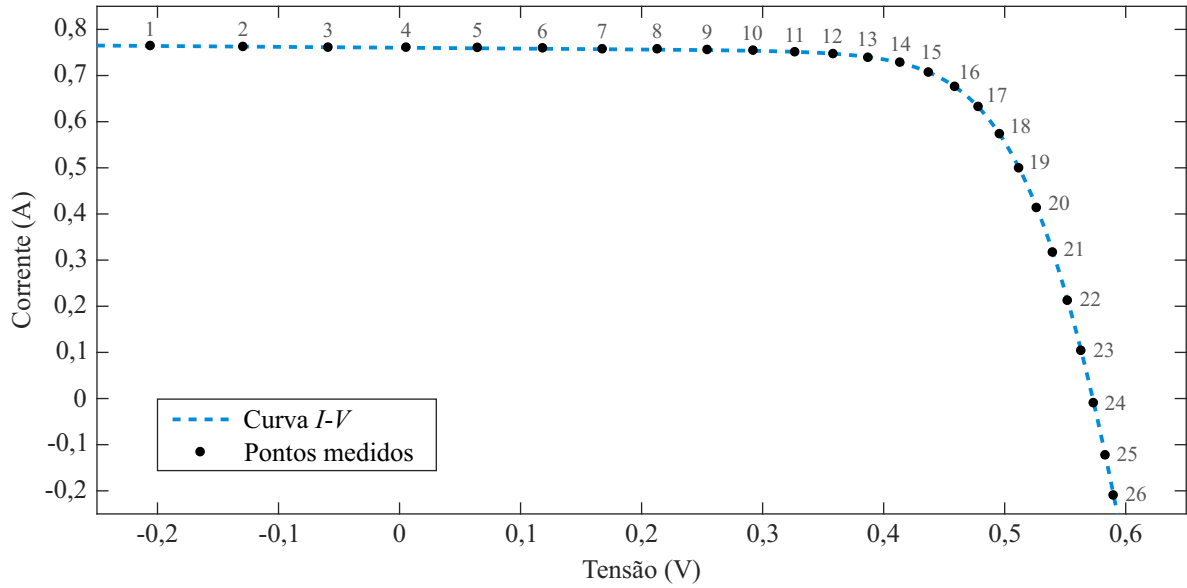
2.2.1 Formulação geral

Conforme ilustrado na Figura 5, a curva $I-V$ de um gerador fotovoltaico é descrita, na prática, por um conjunto de N pares de medições discretas [63]. Em aplicações usuais, N pode variar de dezenas a milhares; logo, assume-se que em todo caso satisfaz $N > 5$ [49–51, 63].

Assim, considere inicialmente uma amostra $j \in \{1, \dots, N\}$ desse conjunto de medições. Ao substituir $V_{m,j}$ na equação característica do modelo de um diodo, dada por (2.6), obtém-se (2.8). Nessa expressão, os parâmetros I_{ph} , I_0 , n , R_{sh} e R_s podem então ser ajustados de modo que a corrente calculada pelo modelo, aqui denotada por $I_{c,j}$, se aproxime de $I_{m,j}$.

$$I_{c,j} = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q (V_{m,j} + I_{c,j} (N_s / N_p) R_s)}{N_s n k T_c} \right) - 1 \right] - \frac{V_{m,j} + I_{c,j} (N_s / N_p) R_s}{(N_s / N_p) R_{sh}} \quad (2.8)$$

É essencial notar, entretanto, que as soluções encontradas para esses cinco parâmetros tendo em vista somente uma amostra j podem não ser apropriadas para outras amostras do mesmo conjunto. Logo, esse processo de ajuste deve ser realizado considerando simultaneamente todos os N pares de medições, o que naturalmente conduz a resolução do sistema de equações homogêneas apresentado em (2.9).

Figura 5 – Curva $I-V$ da célula fotovoltaica de tecnologia de silício cristalino RTC France.

Fonte: medições reportadas em [63].

$$\left\{ \begin{array}{l}
 I_{m,1} - I_{c,1} = I_{m,1} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,1} + I_{c,1}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,1} + I_{c,1}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0 \\
 I_{m,2} - I_{c,2} = I_{m,2} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,2} + I_{c,2}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,2} + I_{c,2}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0 \\
 I_{m,3} - I_{c,3} = I_{m,3} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,3} + I_{c,3}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,3} + I_{c,3}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0 \\
 I_{m,4} - I_{c,4} = I_{m,4} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,4} + I_{c,4}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,4} + I_{c,4}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0 \\
 I_{m,5} - I_{c,5} = I_{m,5} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,5} + I_{c,5}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,5} + I_{c,5}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0 \\
 \vdots \\
 I_{m,j} - I_{c,j} = I_{m,N} - N_p I_{ph} + N_p I_0 \left[\exp\left(\frac{q(V_{m,N} + I_{c,N}(N_s/N_p)R_s)}{N_s n k T_c}\right) - 1 \right] + \frac{V_{m,N} + I_{c,N}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} = 0
 \end{array} \right. \quad (2.9)$$

O fato de (2.9) constituir um sistema sobredeterminado, isto é, possuir mais equações ($N > 5$) do que incógnitas (I_{ph} , I_0 , n , R_{sh} , R_s), indica que há mais restrições do que graus de liberdade [113, 116]. Nessas condições, pequenas inconsistências entre as amostras, decorrentes de ruídos de medição e das próprias limitações do modelo, podem impedir que todas as equações sejam satisfeitas simultaneamente, tornando pouco realista, na prática, a busca por uma solução exata [23, 49, 146]. Por essa razão, é comum reinterpretar esse processo como um problema de otimização, no qual se busca o conjunto de valores dos parâmetros que minimiza um erro global entre as correntes medidas e as correntes pelo modelo de um diodo [113, 116].

No contexto desse problema, os cinco parâmetros do modelo de um diodo são tomados como as *variáveis de decisão*, conforme designado em (2.10). Vale esclarecer que “ \top ” representa o operador de transposição vetorial. Além disso, como esses parâmetros podem assumir, sob uma perspectiva estritamente matemática, qualquer valor contínuo no domínio real (\mathbb{R}), é adequado estabelecer $\mathbf{x} \in \mathbb{R}^5$.

$$\mathbf{x} := [I_{ph}, I_0, n, R_{sh}, R_s]^\top \quad (2.10)$$

A *função objetivo* $\mathcal{F} : \mathbb{R}^5 \mapsto \mathbb{R}$, por sua vez, é tipicamente definida por uma métrica de erro [35, 113]. Dentre as mais tradicionais, destacam-se a *soma dos erros quadráticos* (“*sum of squared errors*”, SSE) e a *raiz do erro quadrático médio* (“*root mean square error*”, RMSE).

No primeiro caso, a expressão matemática da função objetivo é definida por (2.11), em que $I_{c,j}$ é calculado a partir de (2.8), como função de $V_{m,j}$, \mathbf{x} e de um eventual conjunto de parâmetros auxiliares, aqui denotado genericamente por ϕ .

$$\text{SSE}(\mathbf{x}) = \sum_{j=1}^N [I_{m,j} - I_{c,j}(V_{m,j}, \mathbf{x}, \phi)]^2 \quad (2.11)$$

Ao analisar (2.11), conclui-se que a SSE assume apenas valores não negativos. Dessa forma, quanto mais próximo de zero for essa métrica, mais adequados são os valores estabelecidos em \mathbf{x} para os cinco parâmetros e maior é o nível de acurácia dos resultados então retornados pelo modelo de um diodo [35].

No segundo caso, a expressão matemática da função objetivo é dada por (2.12). Assim como a métrica de erro anterior, a RMSE também assume apenas valores não negativos. Logo, quanto mais próximo de zero esta for, mais adequados são os valores estabelecidos em \mathbf{x} para os cinco parâmetros e menor é a incerteza associada aos resultados então retornados pelo modelo de um diodo [147].

$$\text{RMSE}(\mathbf{x}) = \sqrt{\frac{\sum_{j=1}^N [I_{m,j} - I_{c,j}(V_{m,j}, \mathbf{x}, \phi)]^2}{N}} \quad (2.12)$$

Ao comparar (2.11) e (2.12), observa-se que o SSE constitui o núcleo avaliativo do RMSE. Em termos equivalentes, $\text{RMSE} = \sqrt{\text{SSE}/N}$, sendo RMSE uma transformação monótona de SSE. Conseqüentemente, ambas as métricas de erro são capazes de conduzir o processo de otimização para uma mesma solução ótima [52]. Todavia, é pertinente notar que a adoção do SSE tende a ser mais eficiente computacionalmente, pois o seu cálculo requer a execução de um número menor de operações algébricas [52].

Complementarmente, observa-se que *restrições de canalização* são frequentemente introduzidas na formulação do problema de otimização em questão [43, 52, 85, 103, 110]. O propósito por trás dessas restrições é duplo. Em primeiro lugar, busca-se garantir que os sinais matemáticos predefinidos para os cinco parâmetros do modelo de um diodo em (2.8) sejam respeitados, uma vez que estes decorrem essencialmente da análise do circuito elétrico

equivalente apresentado na Figura 1. Desse modo, valores negativos não podem ser admitidos para as variáveis de decisão, sob pena de violar seus significados físicos ou sentidos vetoriais. Em segundo lugar, procura-se assegurar que a solução ótima do problema mantenha compromisso com as faixas de valores tidas como consistentes para os parâmetros do modelo de um diodo — veja a Subseção 2.1.3.

Atento a isso, os valores permitidos para as variáveis de decisão podem ser estabelecidos considerando os limites inferiores e superiores designados, respectivamente, em (2.13) e (2.14). Note que o subscrito “min” presente em \mathbf{lb} indica o valor mínimo permitido para a respectiva variável de decisão, enquanto o subscrito “max” presente em \mathbf{ub} indica o valor máximo permitido.

$$\mathbf{lb} := [I_{ph,\min}, I_{0,\min}, n_{\min}, R_{sh,\min}, R_{s,\min}]^T \quad (2.13)$$

$$\mathbf{ub} := [I_{ph,\max}, I_{0,\max}, n_{\max}, R_{sh,\max}, R_{s,\max}]^T \quad (2.14)$$

Por conveniência, as restrições de canalização podem ser escritas conforme (2.15), em que $\mathbf{lb} \in \mathbb{R}^5$ e $\mathbf{ub} \in \mathbb{R}^5$. Ressalva-se, entretanto, que as desigualdades devem ser interpretadas componente a componente.

$$\mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub} \quad (2.15)$$

À luz dessas considerações, o problema de otimização em questão pode ser enunciado, em sua estrutura mais abrangente, conforme (2.16), em que $\mathcal{F}(\mathbf{x})$ é definida por (2.11) ou (2.12).

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^5}{\text{minimizar}} \quad \mathcal{F}(\mathbf{x}) \\ & \text{sujeito a:} \quad \mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub} \end{aligned} \quad (2.16)$$

Esse problema oferece, portanto, um procedimento sistemático para determinar I_{ph} , I_0 , n , R_{sh} e R_s a partir do conjunto de N pares de medições que descrevem uma curva I - V do gerador fotovoltaico em estudo. Através da sua resolução, os valores ótimos dos parâmetros são efetivamente *extraídos* dessas medições, permitindo ao modelo de um diodo reproduzir os pontos experimentalmente observados e, assim, representar a curva I - V com a máxima fidelidade possível. Por esse motivo, tal problema é frequentemente denominado na literatura como o *problema de extração dos parâmetros* do modelo de um diodo [85].

2.2.2 Classificação matemática

Para escolher o método mais apropriado para resolver um dado problema de otimização, é necessário classificá-lo e compreender de que modo seus atributos influenciam a eficácia e a adequação das alternativas disponíveis [116]. Essas atividades são importantes, pois não existe um método de resolução que seja eficiente, ou mesmo apropriado, para todos os tipos de problemas [116]. Nesse sentido, problemas de otimização costumam ser classificados com base em dois aspectos principais: a sua formulação e as propriedades das funções envolvidas [116].

No que diz respeito à formulação, o conteúdo apresentado na subseção anterior estabelece que o problema de extração dos parâmetros é *multidimensional*, possuindo cinco variáveis de decisão *contínuas* e *reais*. Ademais, (2.16) evidencia que se trata de um problema de otimização com *única função objetivo* e que pode ser declarado tanto na forma *irrestrita* quanto na forma *restrita*, a depender da adoção de restrições de canalização. Já no que se refere às propriedades da função objetivo e das restrições, a análise é mais delicada, devendo uma série de atributos ser particularmente examinada.

O primeiro desses atributos é a linearidade. Nesse sentido, as restrições dadas em (2.15) são, por natureza, funções afins em relação a cada variável de decisão, sendo lineares em \mathbf{x} sob a perspectiva de programação matemática. Em contrapartida, o mesmo não pode ser afirmado acerca da função objetivo. Observa-se que a equação característica do modelo de um diodo, na forma expressa em (2.8), é, por si só, fortemente não linear; afinal, as variáveis de decisão do problema aparecem correlacionadas por meio de operações de multiplicação, divisão e exponenciação. Além disso, tanto o SSE quanto o RMSE envolvem os erros associados a (2.8) em operações quadráticas, conforme (2.11) e (2.12). Assim, conclui-se que a função objetivo e, conseqüentemente, o problema de extração dos parâmetros são *não lineares*.

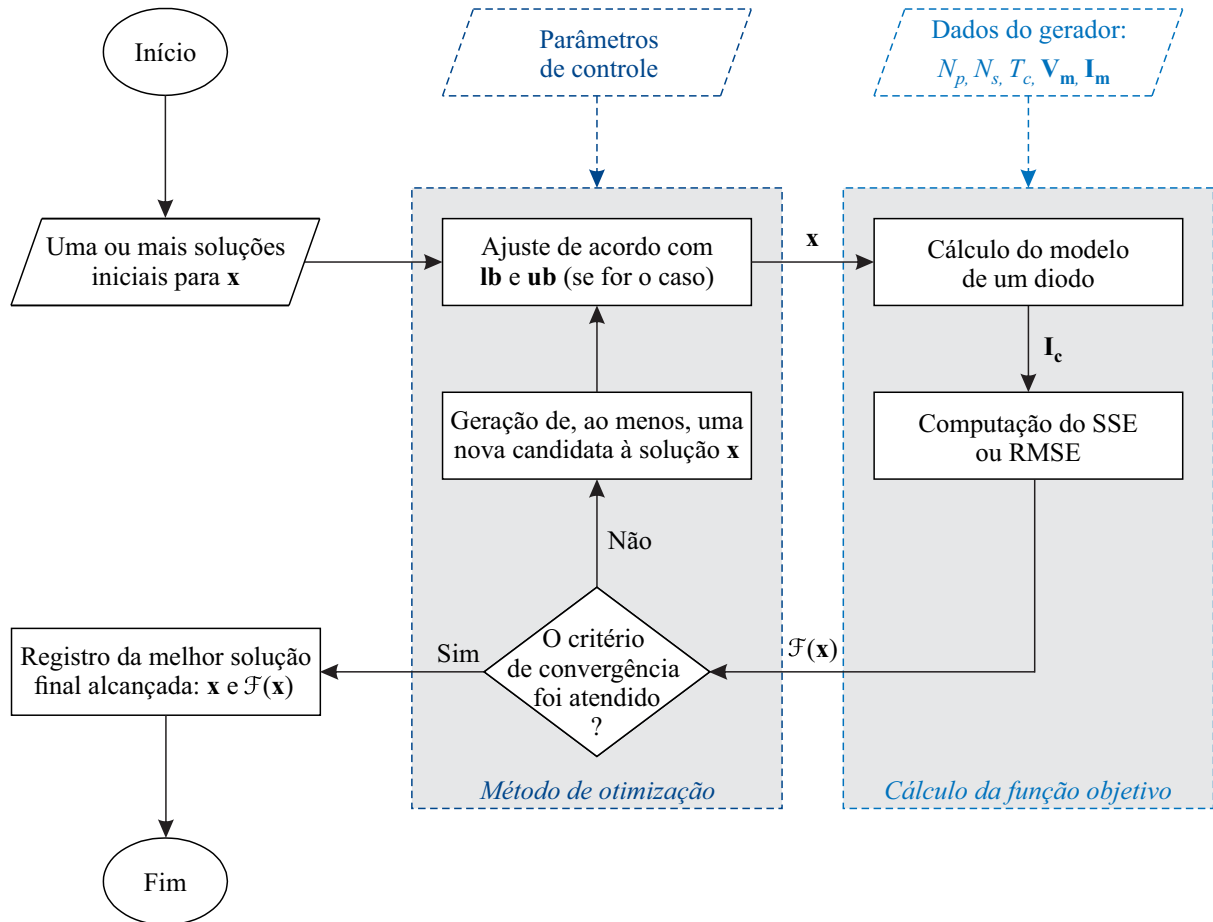
O segundo atributo é a suavidade, isto é, o número de vezes que cada função é continuamente diferenciável [116]. Como mencionado, as restrições dadas em (2.15) são afins e, portanto, continuamente diferenciáveis em relação a cada variável de decisão. Para a função objetivo, a análise é mais complexa. A diferenciação de SSE e RMSE está vinculada à diferenciação de (2.8). Contudo, conforme destacado na Subseção 2.1.4, a equação característica do modelo de um diodo apresenta derivadas parciais de primeira e segunda ordem bem definidas. Logo, o problema de extração dos parâmetros pode ser efetivamente considerado *suave*.

O terceiro e último atributo é a convexidade. Vale recordar que uma função é convexa quando todo segmento de reta que liga dois pontos arbitrários de seu gráfico permanece acima ou sobre o próprio gráfico, sem jamais interceptá-lo [116]. Como consequência, funções convexas são necessariamente unimodais (embora a recíproca não seja verdadeira), isto é, seus gráficos apresentam descidas que conduzem a um único mínimo [116]. Em contraste, funções não convexas podem ser multimodais, exibindo regiões de descida e subida que conduzem a múltiplos mínimos [116]. Posto isto, diversos trabalhos apontam, de maneira consistente, que o problema de extração dos parâmetros apresenta, em geral, natureza *não convexa* [35, 52, 67].

2.2.3 Processo de resolução

Em linhas gerais, o processo de resolução do problema de extração dos parâmetros segue o fluxograma apresentado na Figura 6. Ressalta-se que $\mathbf{V}_m = [V_{m,1}, V_{m,2}, \dots, V_{m,N}]$ e $\mathbf{I}_m = [I_{m,1}, I_{m,2}, \dots, I_{m,N}]$ são vetores auxiliares que sintetizam os pares de medições disponíveis da curva $I-V$ do gerador fotovoltaico em estudo. Nessa mesma linha, $\mathbf{I}_c = [I_{c,1}, I_{c,2}, \dots, I_{c,N}]$ reúne os respectivos resultados retornados pela equação característica do modelo de um diodo.

Figura 6 – Esquema geral de resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico em estudo.



Ao analisar cada uma de suas etapas, constata-se que duas questões fundamentais estão envolvidas nesse processo. A questão mais evidente é: “*qual método de otimização deve ser aplicado para encontrar a solução ótima requerida?*”. À luz da classificação realizada na subseção anterior, essa escolha deve respeitar um critério objetivo de aplicabilidade: métodos aptos a lidar com problemas de otimização multidimensionais, suaves, não lineares e não convexos são, em princípio, admissíveis; caso se adotem restrições de canalização, requer-se, adicionalmente, que o método seja capaz de tratar restrições de desigualdade.

Vale lembrar que métodos de otimização são normalmente caracterizados por algoritmos iterativos [113, 148]. A partir de uma ou mais estimativas de soluções iniciais, esses algoritmos geram uma sequência de novas candidatas à solução, em geral associadas a valores progressivamente menores para a função objetivo [113, 148]. A estratégia adotada para aprimorar tais candidatas a cada iteração e a forma como lidam com diferentes tipos de restrições são o que, essencialmente, diferencia um método de outro [113, 148]. Além disso, distintos critérios de convergência podem ser empregados, a depender dos recursos matemáticos incorporados [52, 85, 110, 113, 148].

A outra questão envolvida no processo de resolução do problema de extração dos parâmetros, e que muitas vezes passa despercebida, diz respeito ao cálculo da função objetivo. Mais expressamente: “*como calcular os valores de $I_{c,j}$ requeridos para a computação de (2.11) ou (2.12)?*”. Assim como destacado na Subseção 2.1.4, a natureza transcendente da equação característica do modelo de um diodo impede que $I_{c,j}$ seja determinado de forma direta a partir de (2.8). Portanto, é necessário adotar alguma simplificação ou tratamento matemático específico que viabilize o seu cálculo.

Uma vez elucidadas essas duas questões, torna-se claro que não há como aplicar qualquer método de otimização à resolução do problema de extração dos parâmetros sem que se estabeleça previamente um procedimento para o cálculo da equação característica do modelo de um diodo no escopo da função objetivo. Por essa razão, as principais abordagens matemáticas adotadas para superar essa questão são detalhadas na próxima seção.

2.3 ABORDAGENS DE CÁLCULO DA FUNÇÃO OBJETIVO

Na literatura, observa-se que há duas abordagens matemáticas comumente adotadas pelos pesquisadores para viabilizar o cálculo do modelo de um diodo no escopo da função objetivo do problema de extração dos parâmetros. No presente documento, a primeira é chamada de *abordagem simplificada* do modelo de um diodo, enquanto a segunda é chamada de *abordagem formal*. As características que definem cada uma dessas abordagens e justificam tais denominações são apresentadas a seguir.

2.3.1 Abordagem simplificada

A *abordagem simplificada* do modelo de um diodo procura contornar a natureza transcendente de sua equação característica ao definir um valor prático para os termos $I_{c,j}$ presentes no lado direito de (2.8). Nesse sentido, o único valor conhecido antecipadamente que se mostra uma opção razoável é o próprio valor de $I_{m,j}$. Assim, a equação característica do modelo de um diodo é expressa, nessa abordagem, conforme (2.17).

$$I_{c,j} = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V_{m,j} + I_{m,j}(N_s/N_p)R_s)}{N_s n k T_c} \right) - 1 \right] - \frac{V_{m,j} + I_{m,j}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} \quad (2.17)$$

É interessante notar que, de certa forma, a abordagem simplificada pressupõe um cenário ideal de emprego do modelo de um diodo. Em outras palavras, considera-se que os pares $(V_{m,j}, I_{m,j})$ que descrevem a curva $I-V$ do gerador fotovoltaico em estudo apresentam ruídos ínfimos e incertezas de medição e registro desprezíveis. Por sua vez, assume-se que o modelo de um diodo é capaz de representar perfeitamente o comportamento elétrico desse gerador. Nessas circunstâncias, quando uma candidata à solução do problema corresponde aos valores ótimos dos cinco parâmetros do modelo de um diodo, as correntes calculadas a partir da equação característica são exatamente iguais às correntes medidas. Logo, a relação dada em (2.17) é

satisfeita. Não obstante, vale observar que, se a candidata à solução não corresponder aos valores ótimos dos cinco parâmetros, ou se alguma das condições de idealidade mencionadas não for efetivamente verificada na prática, os resultados obtidos por meio de (2.17) serão apenas aproximados.

2.3.2 Abordagem formal

A *abordagem formal* do modelo de um diodo procura lidar com a natureza transcendente de sua equação característica por meio de tratamentos matemáticos consolidados. Uma reunião das alternativas de tratamento disponíveis é oportunamente apresentada em [44]. Todavia, observa-se que duas destas se destacam como os preferidos dos pesquisadores: o *método de Newton* e a *função W de Lambert*. Os conceitos por trás desses tratamentos, bem como as suas condições de uso, são objetivamente apresentados a seguir.

2.3.2.1 Resolução implícita via método de Newton

Uma vez designado o devido valor de $V_{m,j}$ e fixados os valores dos cinco parâmetros do modelo de um diodo no vetor \mathbf{x} , a única variável que permanece incógnita em (2.8) é $I_{c,j}$. O tratamento mais tradicional para lidar com uma expressão transcendente dessa forma consiste em reinterpretá-la como um problema de determinação de raízes [119, 149]. Para tanto, transcreve-se (2.8) na função auxiliar (\mathcal{N}) apresentada em (2.18), de modo que a solução procurada é dada precisamente pelo valor de $I_{c,j}$ que satisfaz *implicitamente* $\mathcal{N}(I_{c,j}) = 0$.

$$\mathcal{N}(I_{c,j}) := N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V_{m,j} + I_{c,j}(N_s/N_p)R_s)}{N_s n k T_c} \right) - 1 \right] - \frac{V_{m,j} + I_{c,j}(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} - I_{c,j} \quad (2.18)$$

A técnica mais imediata para resolver esse problema é o bem estabelecido *método de Newton-Raphson* ou, simplesmente, *método de Newton* [95–97, 99–101]. Em termos simples, tal método consiste em um algoritmo de solução de raízes que pode ser aplicado a qualquer função contínua e diferenciável de primeira ordem [149]. A sua ideia central é aproximar linearmente o comportamento da função na vizinhança de uma suposta raiz, por meio da expansão em série de Taylor truncada em seus dois primeiros termos [149]. Para o presente contexto, tal lógica se traduz em (2.19a), em que \mathcal{N}' indica a derivada de primeira ordem de \mathcal{N} e é definida por (2.19b).

$$I_{c,j}^{t+1} = I_{c,j}^t - \frac{\mathcal{N}(I_{c,j}^t)}{\mathcal{N}'(I_{c,j}^t)} \quad (2.19a)$$

$$\mathcal{N}'(I_{c,j}^t) = \frac{d\mathcal{N}(I_{c,j}^t)}{dI_{c,j}^t} = -\frac{R_s}{R_{sh}} - \frac{I_0 R_s q \exp \left(\frac{q(V_{m,j} + I_{c,j}^t(N_s/N_p)R_s)}{N_s n k T_c} \right)}{n k T_c} - 1 \quad (2.19b)$$

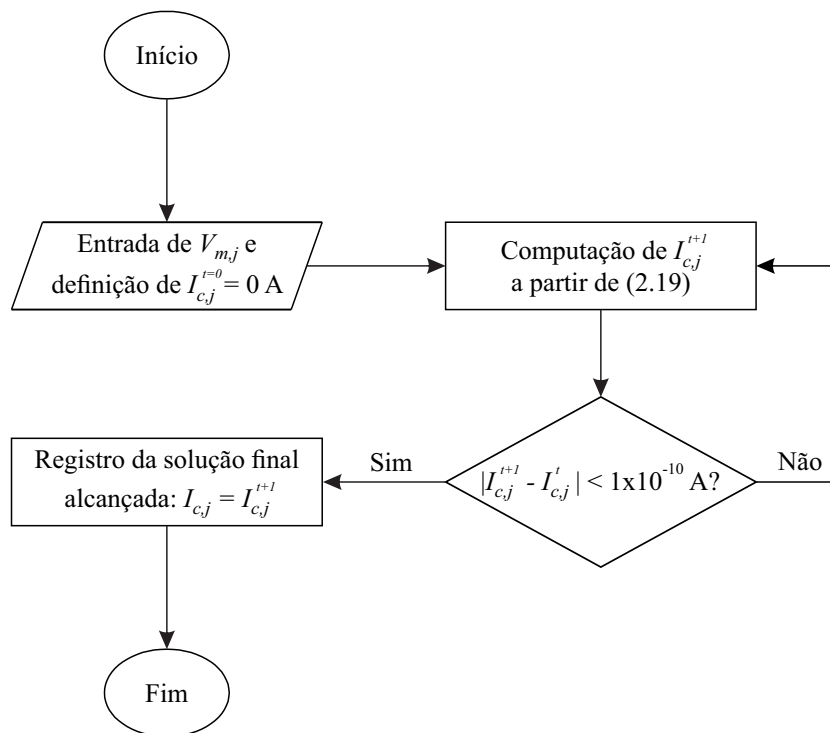
Ao calcular (2.19) recursivamente, é possível obter uma solução cada vez mais acurada para $I_{c,j}$. Observe, inclusive, que o sobrescrito t indica o contador de iteração. Em condições

regulares, esse procedimento apresenta convergência quadrática, isto é, o número de algarismos corretos da solução aproximada em $I_{c,j}^{t+1}$ tende a dobrar a cada iteração [149].

Em geral, os trabalhos encontrados na literatura adotam como estimativa inicial $I_{c,j}^{t=0} = 0$ [95, 101]. Além disso, consideram que o procedimento iterativo deve ser interrompido quando a solução obtida em uma dada iteração não se difere mais significativamente daquela obtida na iteração anterior [95–97, 99–101]. Esse critério de convergência é então definido por uma margem de tolerância, a qual é frequentemente dada por 1×10^{-10} A [97, 100, 101].

Diante dessas considerações, apresenta-se na Figura 7 o fluxograma de aplicação do método de Newton para o cálculo da equação característica do modelo de um diodo.

Figura 7 – Fluxograma de aplicação do método de Newton para o cálculo da equação característica do modelo de um diodo.



2.3.2.2 Resolução explícita via função W de Lambert

A rigor, a *função W de Lambert* é definida como o inverso multivalorado da função dada por $f(w) = w \exp(w)$, em que a variável w assume qualquer número complexo (\mathbb{C}) [122]. Apesar desse conceito intrincado, essa função especial satisfaz a identidade fundamental mostrada em (2.20), de modo que z é uma imagem de $f(w)$ e $W(z) \equiv w$ [123].

$$W(z) \exp(W(z)) = z \quad (2.20)$$

Assim, o principal interesse por trás da função W de Lambert é que esta permite determinar *explicitamente* o valor da variável w que resolve a transcendência $w \exp(w) = z$, conhecendo-se apenas o valor de z [123]. Consequentemente, qualquer equação que possa ser reescrita dessa

forma passa a admitir solução fechada por meio da função W de Lambert [150]. Conforme observado há algumas décadas, este é justamente o caso da equação característica do modelo de um diodo [151]. Após realizar uma série de manipulações estritamente algébricas², esta pode ser reescrita conforme (2.21).

$$I_{c,j} = \frac{N_p R_{sh} (I_{ph} + I_0) - (N_p / N_s) V_{m,j}}{R_{sh} + R_s} - \frac{N_p n k T_c}{q R_s} W(z) \quad (2.21a)$$

$$z = \frac{q I_0 R_{sh} R_s}{n k T_c (R_{sh} + R_s)} \exp\left(\frac{q R_{sh} [N_s R_s (I_{ph} + I_0) + V_{m,j}]}{N_s n k T_c (R_{sh} + R_s)}\right) \quad (2.21b)$$

2.3.3 Comentários adicionais

Uma vez apresentadas as abordagens comumente adotadas para o cálculo do modelo de um diodo no escopo da função objetivo do problema de extração dos parâmetros, torna-se pertinente elucidar e discutir algumas de suas particularidades. Para auxiliar esse processo, sintetiza-se, na Tabela 2, as suas principais características.

No que concerne à abordagem simplificada, é fundamental perceber que (2.17) define $I_{c,j} = f(V_{m,j}, \mathbf{x}, \phi)$, em que $\phi := I_{m,j}$. Logo, essa abordagem permite a pronta determinação de $I_{c,j}$ por meio apenas de operações algébricas elementares. Por conseguinte, a sua implementação computacional é de fácil execução em qualquer ambiente de programação científica. Além disso, é relevante observar que esta permite a realização de uma série de estudos tanto de caráter numérico como de caráter analítico, o que é especialmente interessante a métodos de otimização que se baseiam nas derivadas da função objetivo [113].

Por apresentar esses benefícios, verifica-se que a abordagem simplificada é a mais amplamente adotada pelos pesquisadores [43, 52, 63, 68–89, 91–94, 103–110, 152]. Não obstante, é essencial ter em mente que esta resulta de uma idealização da qualidade de medição da curva I - V , bem como do próprio potencial representativo do modelo de um diodo. Dessa forma, as soluções obtidas a partir de (2.17) tendem, por construção, a ser menos acuradas.

A abordagem formal, por sua vez, apresenta como grande diferencial justamente o alto nível de acurácia das soluções produzidas para $I_{c,j}$. Esse benefício, entretanto, pode vir acompanhado de diferentes custos computacionais, a depender do tratamento matemático adotado. Por exemplo, a resolução implícita da equação característica do modelo de um diodo via método de Newton representa uma alternativa segura, uma vez que este é bem estabelecido na literatura técnica geral [149]. Todavia, o fato de se tratar de um método iterativo que, muitas vezes, exige implementação manual é um fator que pode ser considerado inconveniente por alguns profissionais [95–97, 99–101]. Além disso, por se tratar de um método numérico, este não suporta a realização de estudos analíticos da função objetivo, o que restringe, de certa forma, o seu uso em estudos teóricos mais densos.

² Essas manipulações são contextualizadas no Capítulo 3 e desenvolvidas passo a passo no Apêndice A.

Tabela 2 – Principais características das abordagens de cálculo do modelo de um diodo no escopo da função objetivo do problema de extração dos parâmetros.

Característica	Abordagem simplificada	Abordagem formal	
		Método de Newton	Função W de Lambert
Complexidade teórica	Baixa	Média	Alta
Representação de $I_{c,j}$	$I_{c,j} = f(V_{m,j}, \mathbf{x}, \phi)$, em que $\phi := I_{m,j}$	$\mathcal{N}(I_{c,j}) = 0$	$I_{c,j} = f(V_{m,j}, \mathbf{x}, \phi)$, em que $\phi := \emptyset$
Determinação de $I_{c,j}$	Direta	Iterativa	Tipicamente iterativa
Principais equações	(2.17)	(2.18) e (2.19)	(2.21)
Possibilidades de estudos	Numérica e analítica	Numérica	Numérica e analítica
Vantagens	- Fácil resolução - Fácil implementação computacional	- Acurácia das soluções - Bem estabelecido entre os pesquisadores	- Acurácia das soluções - Implementação facilitada por <i>solvers</i> dedicados
Desvantagens	- Soluções relativamente menos acuradas	- Implementação computacional manual - Não suporta estudos analíticos	- Alto custo computacional dos <i>solvers</i> dedicados - Pouco difundido no meio acadêmico e profissional
Ocorrências na literatura	[43, 52, 63, 68–89, 91–94, 103–110, 152]	[95–97, 99–101]	[43, 53, 54, 64–66, 102]

Em contrapartida, a resolução explícita da equação característica do modelo de um diodo via função W de Lambert surge como uma opção intermediária. Ao analisar (2.21), é possível notar que tal expressão define $I_{c,j} = f(V_{m,j}, \mathbf{x}, \phi)$, em que $\phi := \emptyset$ (entrada vazia). Assim, à primeira vista, a sua implementação computacional é tão fácil de executar quanto aquela da abordagem simplificada do modelo de um diodo. Do mesmo modo, esta também permite a realização de estudos numéricos e analíticos da função objetivo, uma vez que a função W de Lambert é, sob certas condições, continuamente diferenciável [123]. Entretanto, é crucial notar que a pronta determinação de $I_{c,j}$ somente é viável se o valor da função W de Lambert para o argumento z for previamente conhecido. Logo, a questão envolvida nessa abordagem formal do modelo de um diodo não é mais “*como calcular os valores de $I_{c,j}$ requeridos para a computação de (2.11) ou (2.12)?*”, mas sim “*como calcular o valor da função W de Lambert requerido por (2.21a) no ponto definido por (2.21b)?*”.

A esse respeito, cabe destacar que o cálculo da função W de Lambert é um tema de pesquisa ainda hoje em aberto [42, 153]. Como (2.20) é transcendente com respeito a $W(z)$, sua solução pode ser determinada implicitamente por meio do método de Newton, seguindo uma lógica análoga àquela introduzida anteriormente nesta seção [153–155]. Contudo, outros métodos numéricos mais avançados podem oferecer desempenho superior a um custo computacional não tão expressivo nesse caso, a destacar o *método de Halley* [122]. Há ainda uma série de tratamentos alternativos e métodos emergentes que podem ser adotados para esse fim, sendo alguns destes desenvolvidos para escopos genéricos, enquanto outros são oportunamente voltados ao escopo do próprio modelo de um diodo [41, 42, 153–155].

Paralelamente, é possível encontrar ferramentas computacionais ou, em outras palavras, *solvers* dedicados ao cálculo da função W de Lambert em diferentes ambientes de programação científica e bibliotecas computacionais amplamente difundidas, como a rotina correspondente no *software* MAPLE e as funções homônimas `lambertw` oferecidas pelo *software* MATLAB e pelo pacote *SciPy* da linguagem Python [122, 156, 157]. Por se basearem no método de Halley, essas funções demonstram ser capazes de retornar soluções bastante efetivas [41, 44, 122]. No entanto, as suas respectivas eficiências são um ponto que carece de certa atenção. Alguns estudos publicados recentemente apontam, por exemplo, que o código computacional de alto nível da referida função do MATLAB torna o seu tempo de simulação criticamente longo [41, 44].

2.4 MÉTODOS DE OTIMIZAÇÃO DISPONÍVEIS

Conforme introduzido no Capítulo 1, inúmeros métodos de otimização já foram propostos para a resolução do problema de extração dos parâmetros. Em geral, tais métodos são classificados em *determinísticos* e *metaheurísticos*, a depender dos recursos matemáticos incorporados em suas respectivas estratégias de busca da solução ótima. Nos últimos anos, contudo, novos métodos passaram a ser desenvolvidos a partir da combinação de dois ou mais métodos originalmente pertencentes a uma dessas duas classes. Tais métodos encarados, no presente documento, como integrantes de uma terceira classe de métodos de otimização, a qual foi denominada de *conjugados*.

Posto isso, as principais características que definem cada uma dessas classes são apresentadas a seguir. Além disso, alguns dos métodos mais notáveis de cada uma destas são documentados, considerando o seu contexto de aplicação na resolução do problema de extração dos parâmetros.

2.4.1 Métodos determinísticos

Um *método determinístico* pode ser definido como um rigoroso procedimento iterativo [148]. A partir de uma solução inicialmente estimada para o problema, este é capaz de gerar uma sequência de candidatas à solução que se destacam por serem cada vez mais acuradas [113]. Esse processo geralmente termina quando nenhum progresso adicional é obtido ou quando uma solução com um nível mínimo desejado de acurácia é alcançada [63, 120]. As estratégias adotadas para melhorar a qualidade de uma candidata à solução normalmente envolvem funções bem definidas e parâmetros de controle constantes [113, 115]. Assim, sempre que aplicado a partir de uma mesma estimativa inicial, um método determinístico seguirá sempre o mesmo caminho no espaço de soluções do problema e alcançará a mesma solução final [113, 148]. Embora apresente esse nível excepcional de reprodutibilidade, essa propriedade, em geral, não se estende à robustez ou à eficiência.

A maioria dos métodos determinísticos disponíveis na literatura se caracteriza por estratégias de *busca local* da solução ótima [114]. Isto é, tais métodos exploram algumas

informações a respeito da função objetivo e das eventuais restrições do problema em um ponto do seu espaço de soluções, o qual é representado pela candidata à solução em análise, para então encontrar uma direção de ajuste apropriada para os valores das variáveis de decisão que a definem [113, 115]. Nos métodos tidos como sendo de *busca direta*, essas informações são essencialmente referentes ao valor da própria função objetivo [113]. Já nos *métodos baseados em gradiente*, essas informações estão relacionadas às derivadas da função objetivo e, possivelmente, das restrições existentes [113, 120]. Se, por um lado, os métodos de busca direta são computacionalmente mais simples, por outro, os métodos baseados em gradiente costumam ser mais eficientes [116]. Em outras palavras, esses últimos conseguem alcançar a solução ótima de certo problema requerendo um número menor de avaliações da sua respectiva função objetivo [116]. Além disso, o inerente uso das referidas derivadas permite que tais métodos adotem critérios mais rígidos para otimização e, conseqüentemente, forneçam garantias teóricas mais consistentes sobre a otimalidade local das suas soluções finais [113, 116].

Não obstante, é importante perceber que em problemas de natureza não-convexa, tal como é o problema de extração dos parâmetros, esses métodos determinísticos de busca local ou, simplesmente, *métodos determinísticos locais*, estão sujeitos a ficar presos na região de um mínimo local, caso a estimativa inicial considerada para a sua aplicação não esteja suficientemente próxima do mínimo global [113, 114]. Portanto, não há garantia, a princípio, de que a solução final alcançada por um método determinístico local seja a solução ótima global para um dado problema [114].

Em contrapartida, alguns métodos determinísticos baseados em estratégias de *busca global* da solução ótima também podem ser encontrados na literatura [158–160]. Em geral, essas estratégias costumam ser executadas em um plano superior, seguindo uma estrutura enumerativa de “*branch-and-bound*” [160, 161]. Nesse sentido, o espaço de soluções do problema é dividido ou “ramificado” em sub-regiões [161]. Para cada uma destas, são estabelecidos limites de valores para a função objetivo do problema [161]. O limite superior é definido resolvendo um subproblema de otimização similar ao problema original, porém, estando este restrito à sub-região então selecionada [161]. Para realizar essa tarefa, um método determinístico local é usualmente aplicado [159, 161]. Já o limite inferior pode ser definido aplicando uma técnica de relaxação convexa ao problema original, estando este também restrito à sub-região selecionada [160, 161]. Graças à natureza convexa desse subproblema de otimização resultante, a solução final obtida pelo método determinístico local aplicado em sua resolução será sempre uma solução ótima global [113, 114]. Se esses limites superior e inferior encontrados convergirem para um mesmo valor dentro de uma margem de tolerância, então a solução ótima global da sub-região foi encontrada [161]. Caso contrário, esta sub-região é dividida em novas sub-regiões menores, as quais serão posteriormente examinadas [161]. Nesse processo, qualquer sub-região que apresente um valor de limite inferior maior do que o melhor limite superior até então encontrado, considerando todas as sub-regiões já examinadas, é prontamente “podada” [160, 161]. Assim, a árvore figurativa de candidatas à solução vai sendo ramificada e podada, à medida que as

sub-regiões do espaço de soluções do problema vão sendo examinadas. Ao final, é idealmente esperado que reste somente o nó referente à sub-região que contém o mínimo global do problema original.

Essa estrutura enumerativa de *branch-and-bound* consegue oferecer, portanto, garantias teóricas de que a solução final alcançada por um *método determinístico global* é efetivamente a solução ótima global de um dado problema de otimização, mesmo este sendo de natureza não-convexa [158–161]. Todavia, como esta requer a resolução de inúmeros subproblemas de otimização, um significativo esforço computacional é exigido [52, 68, 162]. Além disso, não há garantia de que este irá convergir em um número tratável de iterações ou em um tempo razoável de simulação [68, 161].

Feitas essas considerações, os métodos determinísticos que se tem ciência de terem sido consistentemente testados na resolução do problema de extração dos parâmetros são listados na Tabela 3. Destaca-se que estes se encontram segregados de acordo com a abordagem de cálculo adotada para o modelo nos respectivos testes de validação.

Como pode ser notado, apenas algumas iniciativas isoladas para resolver o problema de extração dos parâmetros considerando a abordagem simplificada do modelo de um diodo se destacam na literatura. A primeira diz respeito ao *método de Newton modificado com um parâmetro de Levenberg* (“*Newton’s method modified with a Levenberg parameter*”, NMMLP), um método determinístico local baseado em gradiente proposto ainda em meados da década de 1980 [63]. A segunda concerne ao algoritmo *branch-and-bound* baseado em aritmética intervalar (“*interval arithmetic-based branch-and-bound algorithm*”, IABBA), um método determinístico global que admite a integração de métodos determinísticos locais baseados em gradiente e que se encontra implementado em uma biblioteca de código aberto em C++, denominada *ibexopt* [68].

Em contraste, verifica-se que a resolução do problema de extração dos parâmetros considerando a abordagem formal do modelo de um diodo tem atraído maior atenção dos pesquisadores. Nos trabalhos encontrados, nota-se que o uso da função W de Lambert é predominante [53, 54, 64–67]. Além disso, destaca-se o fato de todos estes terem recorrido às funções `lsqnonlin` ou `lsqcurvefit` do MATLAB [53, 54, 64–67]. Essas funções são *solvers* dedicados a problemas de mínimos quadrados não lineares, os quais diferem, essencialmente, quanto à forma de declaração dos dados de entrada [120, 121]. Assim, ambas possuem implementados em seu núcleo os seguintes métodos determinísticos locais baseados em gradiente: *algoritmo de Levenberg-Marquardt* (“*Levenberg-Marquardt algorithm*”, LMA) e o *método de região de confiança reflexiva* (“*trust region reflective method*”, TRR) [120, 121].

Essa ampla assistência buscada no MATLAB talvez se justifique pelo fato de esses trabalhos terem como objetivo principal o desenvolvimento de um *método de resolução aproximada* do problema de extração dos parâmetros, e não de um método determinístico local para a sua pronta resolução. De modo mais estrito, observa-se que os métodos de resolução aproximada propostos em [53], [54], [64], [65], [66] e [67] são capazes de retornar soluções de

Tabela 3 – Visão geral dos métodos determinísticos já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$.

Trabalho		Método determinístico			Formulação do problema		
Ref.	Ano	Acrônimo	Busca	Implementação (<i>solver</i>)	\mathcal{F}	lb/ub	Abordagem/Técnica
[63]	1986	NMMLP	Local	Basic	SSE	Não	Simplificada
[68]	2020	IABBA	Global	Minibex/C++ (<i>ibexopt</i>)	SSE	Sim	Simplificada
[52]	2020	IABBA	Global	Julia/C++ (<i>ibexopt</i>)	SSE	Sim	Simplificada
[53]	2014	LMA	Local	MATLAB (<i>lsqnonlin</i>)	SSE	Não	Função W de Lambert
[54]	2016	TRR ^a	Local	MATLAB (<i>lsqnonlin</i>)	SSE	Não	Função W de Lambert
[64]	2018	TRR ^a	Local	MATLAB (<i>lsqcurvefit</i>)	SSE	Não	Função W de Lambert
[65]	2020	TRR ^a	Local	MATLAB (<i>lsqnonlin</i>)	SSE	Não	Função W de Lambert
[66]	2020	TRR ^a	Local	MATLAB (<i>lsqnonlin</i>)	SSE	Não	Função W de Lambert
[67]	2021	TRR	Local	MATLAB (<i>lsqnonlin</i>)	SSE	Não	Função W de Lambert

^aInformações obtidas por meio de contato direto com os autores correspondentes dos trabalhos.

elevada qualidade para o referido problema, isto é, soluções que se aproximam da solução ótima global correspondente. Cientes disso, os autores desses métodos sugerem que um “refinamento” seja aplicado sobre essas soluções aproximadas para que, dessa forma, se obtenha a solução ótima global. Todavia, esse refinamento consiste em utilizar essas soluções aproximadas como estimativas iniciais para a aplicação de um método determinístico local na resolução do problema. Logo, tais autores parecem ter recorrido aos *solvers* anteriormente mencionados do MATLAB, para evitar maiores complicações em seus trabalhos.

É essencial ainda notar que os métodos de resolução aproximada propostos nos referidos trabalhos podem ser encarados como legítimos *procedimentos de inicialização*. Nesse sentido, é importante esclarecer que cada um desses procedimentos foi respectivamente concebido, segundo uma lógica distinta de formulação e resolução. Uma visão geral dessas características é devidamente apresentada mais à frente, no Item 2.4.1.1.

Antes, porém, é pertinente comentar que os trabalhos apresentados em [64], [65] e [66] tratam o modelo de um diodo de forma normalizada. Em [65] e [66], essa normalização é feita em *por unidade* — ou, simplesmente, *pu* — para as correntes, tensões e resistências que compõem, essencialmente, a equação característica do modelo de um diodo. Por sua vez, em [64], tal normalização é feita em relação a um valor unitário. De acordo com os autores desse último trabalho, o tratamento normalizado se justifica pela diferença de ordem de grandeza entre as variáveis de decisão [64]. Assim, a principal vantagem oferecida pela sua adoção reside na possível redução da complexidade computacional do problema de extração dos parâmetros [65, 113].

2.4.1.1 Procedimentos de inicialização

A Tabela 4 apresenta um panorama dos procedimentos de inicialização (ou métodos de resolução aproximada) propostos nos trabalhos encontrados que aplicam métodos determinísticos locais na resolução do problema de extração dos parâmetros.

Tabela 4 – Visão geral dos procedimentos de inicialização propostos para a aplicação de métodos determinísticos locais.

Trabalho	Núcleo de resolução	Principais aspectos
[63]	Etapas alternadas não lineares/lineares	<ul style="list-style-type: none"> - Baseado na abordagem simplificada do modelo de um diodo. - Não requer conhecimento de nenhum ponto chave da curva $I-V$. - Não apresenta fórmulas para definir a estimativa inicial de n e R_s.
[53]	Problema de otimização bidimensional	<ul style="list-style-type: none"> - Baseado em ambas as abordagens do modelo de um diodo. - Requer, ao menos, o conhecimento de SC e OC da curva $I-V$. - Recorre ao LMA implementado na função <code>lsqnonlin</code> do MATLAB. - Adota $n = 1$ como estimativa inicial e apresenta uma formulação para inicializar R_s como função de n.
[54]	Decomposição do tipo Benders generalizada	<ul style="list-style-type: none"> - Baseado em ambas as abordagens do modelo de um diodo. - Recorre ao TRR implementado na função <code>lsqnonlin</code> do MATLAB. - Não requer conhecimento de pontos chave da curva $I-V$. - Adota $n = 1$ como estimativa inicial e apresenta uma formulação para inicializar R_s a partir da curva $I-V$.
[64]	Problema de mínimos quadrados lineares em dois estágios	<ul style="list-style-type: none"> - Baseado na abordagem simplificada do modelo de um diodo. - Não requer conhecimento de pontos chave da curva $I-V$. - Não requer que nenhuma variável de decisão seja previamente inicializada.
[65]	Problema de otimização tridimensional	<ul style="list-style-type: none"> - Baseado em ambas as abordagens do modelo de um diodo, dadas em pu. - Requer conhecimento do ponto de MP da curva $I-V$. - Recorre ao TRR implementado na função <code>lsqnonlin</code> e à função <code>fsolve</code> (com método não declarado) do MATLAB. - Apresenta uma formulação para inicializar n, R_{sh} e R_s a partir da curva $I-V$.
[66]	Sistema de equações não lineares	<ul style="list-style-type: none"> - Baseado na abordagem simplificada do modelo de um diodo, dada em pu. - Não requer conhecimento de pontos chave da curva $I-V$. - Adota $n = 1,5$ como estimativa inicial e apresenta formulações para inicializar R_{sh} e R_s a partir da curva $I-V$.
[67]	Problema de mínimos quadrados não lineares separável	<ul style="list-style-type: none"> - Baseado em ambas as abordagens do modelo de um diodo. - Não requer conhecimento de pontos chave da curva $I-V$. - Recorre ao método NMS implementado na função <code>fminsearch</code> do MATLAB. - Adota $n = 2$ como estimativa inicial e apresenta uma formulação para inicializar R_s a partir da curva $I-V$.

A maior parte desses procedimentos de inicialização é caracterizada por um ou mais subproblemas de otimização que apresentam um espaço de soluções reduzido (menos de cinco dimensões) e podem ser resolvidos com maior facilidade do que o problema original de extração dos parâmetros [53, 54, 63–65, 67]. Para a concepção desses subproblemas, os autores costumam explorar as relações de linearidade e não linearidade existentes entre os cinco parâmetros na equação característica do modelo de um diodo. Por exemplo, a resistência paralela pode ser escrita em termos de condutância ($G_{sh} = 1/R_{sh}$) em (2.8), resultando em (2.22). Assim, se valores forem diretamente atribuídos para n e R_s , tal equação passa a assumir caráter linear, o que é bastante atrativo para a formulação de um subproblema.

$$I_{c,j} = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V_{m,j} + I_{c,j}(N_s/N_p)R_s)}{N_s n k T_c} \right) - 1 \right] - (N_p/N_s) G_{sh} (V_{m,j} + I_{c,j}(N_s/N_p)R_s) \quad (2.22)$$

Esse entendimento foi inicialmente explorado no pioneiro trabalho apresentado em [63]. Com o objetivo de viabilizar a convergência adequada do NMMLP, o autor desenvolveu um

procedimento de inicialização baseado em *etapas alternadas não lineares/lineares*. Nesse procedimento, as variáveis de decisão n e R_s precisam ser convenientemente inicializadas. Com esses valores fixados, I_{ph} , I_0 e G_{sh} são estimados por regressão linear [63]. Em seguida, os resultados obtidos para essas três variáveis de decisão são fixados na função objetivo do problema de extração dos parâmetros, a qual é posteriormente minimizada em relação a n e R_s mediante aplicação do método de Newton–Raphson [63]. Esses passos são repetidos com os novos valores de n e R_s até que a diferença relativa entre dois valores consecutivos da função objetivo (calculados após cada regressão linear) se torne inferior a uma tolerância pré-estabelecida [63]. Embora esse procedimento de inicialização tenha se mostrado útil, é pertinente destacar que nenhuma formulação foi explicitamente apresentada para o cálculo inicial de n e R_s [63].

Recentemente, o autor de [67] propôs um procedimento de inicialização baseado em um *problema de mínimos quadrados não lineares separável*, o qual guarda estreita semelhança com aquele descrito anteriormente. Em particular, tal procedimento faz uso da abordagem formal do modelo de um diodo e recomenda o emprego do método de busca direta denominado *algoritmo simplex de Nelder–Mead* (“*Nelder–Mead simplex algorithm*”, NMS), disponível na função `fminsearch` do MATLAB, para resolver o subproblema não linear associado a n e R_s [67, 163]. Um método de penalização também é introduzido para evitar soluções com $n < 1$ e $R_s < 0$ [67]. Para formalizar a aplicação desse procedimento, o autor sugere adotar $n = 2$ como estimativa inicial e apresenta uma formulação para determinar o valor inicial de R_s a partir dos pontos extremos da curva I – V (independentemente de estarem associados a SC e OC) [67].

O procedimento de inicialização apresentado em [54] segue a mesma linha dos anteriores; contudo, os subproblemas de otimização são formulados com base em uma chamada “aproximação educada” e por meio de uma *decomposição do tipo Benders generalizada*. Assim, as soluções obtidas no subproblema primário para n e R_s , o qual se baseia na abordagem formal do modelo de um diodo, são utilizadas como dados de entrada na resolução do subproblema secundário, que inclui as variáveis de decisão I_{ph} , I_0 e G_{sh} e se baseia na abordagem simplificada do modelo de um diodo [54]. Para resolver esses subproblemas, os autores recorreram à função `lsqnonlin` do MATLAB [54]. Além disso, sugeriram adotar $n = 1$ como estimativa inicial e apresentaram uma formulação para determinar o valor inicial de R_s a partir dos pontos extremos da curva I – V [54].

Outro procedimento de inicialização baseado em (2.22) foi desenvolvido em [53]. Fazendo uso de expressões analíticas associadas à curva I – V , os autores conceberam duas formas distintas de escrever I_{ph} , I_0 e G_{sh} como funções de n e R_s [53]. A primeira depende, em particular, das informações de SC, OC e MP; ao passo que a segunda depende de SC, OC e de qualquer outro ponto da curva I – V [53]. Apesar dessa dependência dos pontos operacionais notáveis da curva I – V , os subproblemas de otimização formulados a partir dessas duas formas apresentam a vantagem de envolver apenas duas variáveis de decisão e de possuir natureza convexa [53]. Portanto, podem ser resolvidos com facilidade por um método determinístico local (ou mesmo

por um método metaheurístico, conforme mostrado em [164]). Em conformidade com isso, os autores recomendaram o uso da função `lsqnonlin` do MATLAB para resolver o subproblema considerado [53]. Além disso, propuseram adotar $n = 1$ como estimativa inicial e apresentaram uma formulação para determinar o valor inicial de R_s , dependente apenas de n [53].

O autor de [65] propôs um procedimento de inicialização que, de certo modo, se relaciona ao anterior. Em síntese, tal procedimento também faz uso de expressões analíticas e do ponto de MP da curva $I-V$ [65]. Entretanto, por não trabalhar com a condutância paralela, o subproblema de otimização formulado passa a contar com três variáveis de decisão: n , R_{sh} e R_s [65]. É pertinente destacar que a equação característica do modelo de um diodo e, conseqüentemente, as variáveis de decisão desse subproblema são tratadas em *pu* [65]. Dessa forma, o autor sugere o uso da função `lsqnonlin` do MATLAB para resolver tal subproblema [65]. Ademais, para determinar o valor inicial de n , o autor propõe a resolução de um sistema de equações não lineares por meio da função `fsolve` do MATLAB [65, 165]. Já para os valores iniciais de R_s e R_{sh} , são apresentadas formulações diretas baseadas nos pontos extremos da curva $I-V$ [65].

Por sua vez, o procedimento de inicialização apresentado em [64] segue uma lógica mais particular. A partir da análise do formato típico da curva $I-V$ e de manipulações realizadas na equação característica do modelo de um diodo, os autores obtiveram uma expressão matemática alternativa com duas partes bem definidas: uma linear e outra formada por termos exponenciais [64]. Como esta última pode ser linearizada de forma direta mediante uma transformação logarítmica, o subproblema de otimização resultante é então resolvido pela aplicação de um *método de mínimos quadrados lineares em dois estágios* [64]. Cabe notar que a primeira etapa desse método está associada à parte linear e, em última instância, relaciona-se às variáveis de decisão I_{ph} , I_0 , R_{sh} e R_s [64]. A etapa subsequente está associada à parte exponencial linearizada e depende das variáveis de decisão I_0 , n , R_{sh} e R_s [64]. É oportuno destacar que esse método não requer qualquer estimativa inicial, o que é um aspecto bastante atrativo.

Finalmente, o procedimento de inicialização desenvolvido em [66] pode ser considerado o mais singular. Diferentemente dos demais procedimentos descritos anteriormente, caracterizados essencialmente pela resolução iterativa de um ou mais subproblemas de otimização de dimensão reduzida (menos de cinco), tal procedimento consiste na resolução sequencial de um sistema de equações não lineares [66]. Mais especificamente, adota-se $n = 1,5$ como estimativa inicial; propõem-se formulações para determinar os valores iniciais de R_{sh} e R_s a partir da curva $I-V$; e, com o auxílio de expressões analíticas, escrevem-se I_{ph} e I_0 como funções de n , R_{sh} e R_s [66]. É pertinente ressaltar que essas expressões matemáticas são fornecidas em *pu* [66]. Além disso, é interessante mencionar que tais expressões requerem apenas informações sobre os dois pontos extremos da curva $I-V$ como dados de entrada [66].

A simplicidade desse último procedimento de inicialização pode sugerir, em um primeiro momento, que se trata de uma alternativa ideal ou mesmo da mais recomendável para determinar as estimativas iniciais dos cinco parâmetros do modelo de um diodo. Não obstante, é necessário

adotar certa cautela diante desse julgamento. Por exemplo, os procedimentos de inicialização desenvolvidos em [54], [64] e [67] foram testados com sucesso no repositório do NREL, o qual contém mais de um milhão de curvas $I-V$ [49]. Ademais, os procedimentos de inicialização propostos em [53] e [64] encontram-se implementados em uma ferramenta *online* de grande destaque, o que sugere uma elevada confiabilidade em sua aplicação prática — veja [166]. O procedimento de inicialização proposto em [66], por outro lado, foi validado apenas com estudos de caso de referência. Portanto, testes adicionais precisam ser realizados para que se possa chegar a uma opinião mais conclusiva acerca do seu real potencial.

2.4.2 Métodos metaheurísticos

A maioria dos métodos metaheurísticos disponíveis é inspirada na natureza, isto é, foi desenvolvida com base na abstração matemática de algum fenômeno natural [114]. A razão por trás dessa lógica reside no fato de a natureza ter evoluído ao longo de milhões de anos e ter encontrado soluções viáveis para quase todos os problemas que enfrentou [115]. Dependendo do fenômeno natural adotado como inspiração, a estratégia de busca de um método metaheurístico pode ser *baseada em trajetória* ou *baseada em população* [148]. A primeira se concentra em modificar e melhorar uma única candidata à solução, enquanto a segunda se concentra em manter e aprimorar um conjunto de candidatas à solução [148].

O desempenho dessas tarefas depende diretamente de como o comportamento do fenômeno natural é matematicamente modelado. Essa modelagem é, por sua vez, geralmente estruturada em um algoritmo simples, formado por equações intuitivas e independentes do problema, isto é, que não consideram as derivadas da função objetivo, mas apenas os seus valores em certos pontos [114, 148]. No entanto, a composição ou aplicação dessas equações está quase sempre condicionada a um número (pseudo)aleatório que, em última análise, representa a aleatoriedade intrínseca do fenômeno natural [148]. Esses números são usualmente sorteados a partir de distribuições probabilísticas, como a distribuição Normal ou a de Lévy, cada qual oferecendo diferentes potenciais de randomização ao método [148]. Dessa forma, a implementação computacional dos métodos metaheurísticos é regularmente realizada com a ajuda de ferramentas dedicadas ao sorteio de números aleatórios, como a função `rand` do MATLAB [148, 167].

Apesar dessas particularidades, a estratégia de busca de qualquer método metaheurístico é tipicamente caracterizada por *processos de intensificação e diversificação* [114, 115, 148]. O primeiro busca tornar uma candidata à solução mais acurada, explorando a proximidade da melhor candidata à solução até então encontrada [148]. Em contrapartida, o segundo visa gerar uma candidata à solução significativamente diferente de suas versões anteriores ou de outras candidatas concorrentes, explorando o espaço de soluções do problema em uma escala global [148]. Logo, esse processo evita que candidatas à solução fiquem presas na região de um mínimo local do espaço de soluções do problema e, ao mesmo tempo, aumenta a diversidade das candidatas à solução geradas durante as iterações [115].

Uma vez que os métodos metaheurísticos são normalmente aplicados considerando, como critério de convergência, um número máximo de avaliações da função objetivo ou de iterações, o equilíbrio entre esses processos de intensificação e diversificação desempenha um papel fundamental em seus respectivos sucessos [85, 148]. Além disso, graças à coexistência desses dois processos, os métodos metaheurísticos são entendidos como adequados para otimização global [115]. É crucial notar, entretanto, que não há garantias teóricas de que a solução final alcançada por qualquer método metaheurístico seja, em princípio, a solução ótima global do problema em questão [148].

Portanto, um *método metaheurístico* pode ser definido como um procedimento iterativo que usa certos *trade-offs* de busca local e randomização para gerar uma sequência de candidatas à solução cada vez mais acuradas para determinado problema [148]. Consequentemente, se for aplicado várias vezes na resolução desse problema, considerando uma mesma estimativa inicial ou um mesmo conjunto de estimativas, este não necessariamente seguirá o mesmo caminho no espaço de soluções e, possivelmente, nem mesmo alcançará a mesma solução final [114]. Contudo, se a sua estratégia de busca for bem projetada, pode se mostrar robusto e estável o suficiente para obter soluções finais próximas e de boa qualidade em todas as suas aplicações [85, 114, 148]. Por essa razão, os testes de validação de um método metaheurístico geralmente envolvem uma análise estatística das soluções finais alcançadas em várias rodadas de aplicação na resolução de um mesmo problema de interesse — veja, por exemplo, [43] ou [85].

Devido à inerente liberdade de criação e ao crescente interesse da comunidade científica nos métodos metaheurísticos, verifica-se que um número impressionante desses métodos tem sido sistematicamente testado, ou mesmo propriamente desenvolvido, para a resolução do problema de extração dos parâmetros. Alguns dos métodos metaheurísticos mais notáveis da literatura são listados na Tabela 5.

A partir da Tabela 5, é possível constatar que a formulação do problema de extração dos parâmetros considerando a abordagem simplificada do modelo de um diodo é predominante no âmbito dos métodos metaheurísticos. De certa forma, a facilidade de implementação dessa abordagem combina bem com a simplicidade de implementação computacional desses métodos. Além disso, em comparação com a abordagem formal, o seu menor custo computacional é uma característica particularmente atraente, uma vez que os métodos metaheurísticos geralmente realizam milhares de avaliações da função objetivo durante a resolução de um problema — veja, por exemplo, [43] ou [85].

Assim, observa-se que diversos métodos metaheurísticos foram propostos para a resolução do problema em questão, especialmente no decorrer da última década. Dentre as diversas alternativas encontradas na literatura, destacam-se: *algoritmo genético* (“*genetic algorithm*”, GA) [69], *otimização por enxame de partículas* (“*particle swarm optimization*”, PSO) [69], *evolução diferencial* (“*differential evolution*”, DE) [70], *recozimento simulado* (“*simulated annealing*”, SA) [71], *otimização por enxame de abelhas artificiais* (“*artificial bee swarm*

Tabela 5 – Visão geral dos métodos metaheurísticos já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$.

Trabalho		Método metaheurístico			Formulação do problema		
Ref.	Ano	Acrônimo	Tipo	Implementação	\mathcal{F}	lb/ub	Abordagem/Técnica
[69]	2009	GA	Básico	MATLAB	RMSE	—	Simplificada
[69]	2009	PSO	Básico	MATLAB	RMSE	—	Simplificada
[70]	2010	DE	Básico	MATLAB	Outra	—	Simplificada
[71]	2012	SA	Básico	—	Outra	—	Simplificada
[81]	2012	IGHS	Modificado	MATLAB	RMSE	Sim	Simplificada
[72]	2012	ABSO	Básico	MATLAB	RMSE	Sim	Simplificada
[73]	2013	BMO	Básico	MATLAB	RMSE	Sim	Simplificada
[74]	2013	CS	Básico	MATLAB	RMSE	Sim	Simplificada
[82]	2013	Rcr-IJADE	Modificado	C++	RMSE	Sim	Simplificada
[75]	2014	BFA	Básico	—	RMSE	Sim	Simplificada
[75]	2014	ABC	Básico	—	RMSE	Sim	Simplificada
[76]	2014	TLBO	Básico	LabVIEW	Outra	Sim	Simplificada
[77]	2016	CSO	Básico	—	RMSE	Sim	Simplificada
[78]	2017	WDO	Básico	MATLAB	RMSE	Sim	Simplificada
[79]	2017	WOA	Básico	MATLAB	RMSE	—	Simplificada
[92]	2018	TLABC	Híbrido	—	RMSE	Sim	Simplificada
[91]	2018	DE/WOA	Híbrido	MATLAB	RMSE	Sim	Simplificada
[80]	2019	SDO	Básico	MATLAB	RMSE	Sim	Simplificada
[83]	2019	PGJAYA	Modificado	MATLAB	RMSE	Sim	Simplificada
[84]	2019	ITLBO	Modificado	MATLAB	RMSE	Sim	Simplificada
[86]	2019	CPMPSO	Modificado	MATLAB	RMSE	Sim	Simplificada
[87]	2019	MCSWOA	Modificado	MATLAB	RMSE	Sim	Simplificada
[93]	2019	BHCS	Híbrido	—	RMSE	Sim	Simplificada
[85]	2020	ELBA	Modificado	MATLAB	RMSE	Sim	Simplificada
[94]	2020	GWOCs	Híbrido	MATLAB	RMSE	Sim	Simplificada
[88]	2021	IMPA	Modificado	Java	RMSE	Sim	Simplificada
[89]	2021	EABOA	Modificado	MATLAB	RMSE	Sim	Simplificada
[43]	2021	nDEBCO	Híbrido	MATLAB	RMSE	Sim	Simplificada
[90]	2022	SDGBO	Modificado	MATLAB	RMSE	Sim	Simplificada
[98]	2022	SGDE	Híbrido	MATLAB	RMSE	Sim	Simplificada
[99]	2018	ELPSO	Modificado	—	RMSE	Sim	Método de Newton
[100]	2018	GCPSO	Modificado	MATLAB	RMSE	Sim	Método de Newton
[101]	2019	HISA	Híbrido	MATLAB	RMSE	Sim	Método de Newton
[95]	2020	COA	Básico	MATLAB	RMSE	Sim	Método de Newton
[96]	2020	FFO	Básico	MATLAB	RMSE	Sim	Método de Newton
[102]	2021	ImSMA	Modificado	MATLAB	RMSE	Sim	Função W de Lambert
[97]	2021	SFS	Básico	—	RMSE	Sim	Método de Newton
[43]	2021	nDEBCO	Híbrido	MATLAB	RMSE	Sim	Função W de Lambert

optimization”, ABSO) [72], *otimizador de acasalamento de pássaros* (“*bird mating optimizer*”, BMO) [73], *busca de cuco* (“*cuckoo search*”, CS) [74], *algoritmo de forrageamento bacteriano* (“*bacteria foraging algorithm*”, BFA) [75], *colônia de abelhas artificiais* (“*artificial bee colony*”, ABC) [75], *otimização baseada em ensino-aprendizagem* (“*teaching-learning-based optimization*”, TLBO) [76], *otimização por enxame de gatos* (“*cat swarm optimization*”, CSO) [77], *otimização orientada pelo vento* (“*wind driven optimization*”, WDO) [78], *algoritmo de otimização de baleias* (“*whale optimization algorithm*”, WOA) [79] e *otimização baseada em oferta-demanda* (“*supply-demand-based optimization*”, SDO) [80]. É interessante comentar que, com exceção do SA, todos esses métodos metaheurísticos apresentam uma estratégia de busca da solução ótima baseada em população.

Embora a maioria desses métodos tenha obtido resultados satisfatórios nos testes de validação realizados, inúmeras questões relacionadas aos seus desempenhos começaram a surgir

com o passar do tempo, colocando em dúvida o amplo sucesso de sua aplicação na resolução do problema de extração dos parâmetros. Alguns métodos metaheurísticos, como o PSO, DE e WDO, apresentam uma característica de *convergência prematura*, isto é, tendem a concentrar a busca em uma região específica sem ter explorado suficientemente o espaço de soluções do problema [35, 168]. Portanto, se tal região estiver associada a uma solução ótima local, esses métodos possivelmente ficarão presos nesta, não atingindo a região do mínimo global do espaço de soluções do problema. Por outro lado, certos métodos metaheurísticos, como o SA, TLBO e WOA, tendem a ter uma *convergência lenta*, isto é, requerem muitas iterações para começarem a intensificar propriamente suas buscas em uma região onde o mínimo global do espaço de soluções do problema poderia potencialmente estar [35, 168]. Essa característica compromete, sobretudo, a eficiência desses métodos.

Além disso, a maioria dos métodos metaheurísticos tem seu desempenho afetado, em menor ou maior grau, pelo ajuste de seus parâmetros de controle [152]. Vale ressaltar que esses parâmetros frequentemente definem o equilíbrio entre os processos de intensificação e diversificação de um método metaheurístico [148]. Em vista disso, o trabalho apresentado em [152] se dedicou a avaliar o desempenho de GA, PSO, DE, CS, BFA e ABC na resolução de alguns problemas de extração dos parâmetros do modelo de um diodo, considerando diferentes configurações para seus respectivos parâmetros de controle [152]. Os resultados obtidos revelaram que o PSO é o método mais sensível ao ajuste de parâmetros dentre os avaliados, enquanto o CS se mostrou o menos sensível, além de mais acurado e robusto [152].

A fim de mitigar esses inconvenientes, observa-se que diversos pesquisadores começaram a desenvolver e testar versões modificadas desses métodos metaheurísticos na resolução do problema de extração dos parâmetros. O trabalho apresentado em [81] foi um dos primeiros encontrados a seguir essa linha. Neste, é demonstrado que o método metaheurístico modificado denominado *busca harmônica global inovadora* (“*innovative global harmony search*”, IGHS) é capaz de retornar resultados mais promissores do que o método metaheurístico de *busca harmônica* (“*harmony search*”, HS) em sua versão básica e em outra versão modificada [81]. Em seguida, [82] propôs o chamado *evolução diferencial adaptativa reparada* (“*repaired adaptive differential evolution*”, Rcr-IJADE) e demonstrou que a sua acurácia é superior à verificada para o IGHS e outros cinco métodos metaheurísticos [82]. Curiosamente, o estudo comparativo apresentado em [103] avaliou o desempenho desse último método metaheurístico modificado frente ao DE em sua versão básica e outras nove versões modificadas bem estabelecidas [103]. A partir dos resultados encontrados, os autores concluíram que o Rcr-IJADE é o mais acurado e robusto dentre todos os avaliados [103].

Em outra linha de pesquisa, o trabalho apresentado em [83] propôs o método metaheurístico modificado denominado *JAYA guiado pelo desempenho* (“*performance-guided JAYA*”, PGJAYA) e demonstrou que o seu desempenho é superior aos observados para o JAYA em sua versão básica e para uma de suas variantes [83]. Por sua vez, [84] propôs a *otimização baseada*

no ensino-aprendizagem melhorada (“*improved teaching-learning-based optimization*”, ITLBO) e demonstrou que a sua acurácia e o seu tempo de simulação são melhores que os verificados para o TLBO básico e três de suas variantes [84]. De modo análogo, [86] propôs o *algoritmo de otimização por enxame de partículas baseado em mutação por perturbação classificada* (“*classified perturbation mutation based particle swarm optimization algorithm*”, CPMP SO) e mostrou que tal método metaheurístico modificado supera o PSO básico e duas de suas variantes [86]. Paralelamente, [87] propôs o *algoritmo de otimização de baleias com cruzamento assistido de estratégias de busca modificadas e operador de seleção* (“*modified search strategies assisted crossover whale optimization algorithm with selection operator*”, MCSWOA) e demonstrou que este é competitivo em relação ao WOA básico e a outras onze variantes avançadas deste [87].

Em contraste, o autor do presente documento propôs em [85] o chamado *algoritmo de morcego aprimorado com voo de Lévy* (“*enhanced Lévy flight bat algorithm*”, ELBA) e demonstrou que a sua acurácia, eficiência e tempo de simulação são significativamente superiores àqueles verificados para o *algoritmo de morcego* (“*bat algorithm*”, BA) em sua versão básica e outras duas versões modificadas bem consolidadas [85]. Além disso, nesse trabalho foi demonstrado que o ELBA é tão robusto quanto o Rcr-IJADE, tão rápido quanto o ITLBO e tão ou mais acurado que Rcr-IJADE, PGJAYA, ITLBO, CPMP SO e MCSWOA [85].

Mais recentemente, outros métodos metaheurísticos modificados foram propostos, apresentando resultados promissores, como o *algoritmo de predadores marinhos melhorado* (“*improved marine predators algorithm*”, IMPA) [88], o *algoritmo de otimização de borboleta adaptativa aprimorado* (“*enhanced adaptive butterfly optimization algorithm*”, EABOA) [89] e o *método de otimização baseado em gradiente diferencial cosseno-seno* (“*sine cosine differential gradient-based optimization method*”, SDGBO) [90].

Ao invés de modificar um único método metaheurístico básico para obter uma técnica mais confiável de resolução do problema em questão, alguns pesquisadores optaram por hibridizar características potencialmente complementares de dois ou mais desses métodos. Nesse sentido, em [91], o processo de diversificação do DE foi combinado com o processo de intensificação do WOA para formar o denominado *evolução diferencial híbrida com algoritmo de otimização de baleia* (“*hybrid differential evolution with whale optimization algorithm*”, DE/WOA) [91]. Já em [92], os autores associaram as notórias capacidades de intensificação do TLBO com a boa habilidade de exploração do ABC para criar o *algoritmo de colônia de abelhas artificiais baseado em ensino-aprendizagem* (“*teaching-learning-based artificial bee colony algorithm*”, TLABC) [92]. Posteriormente, em [93], os incisivos operadores de busca local da *otimização baseada em biogeografia* (BBO) foram combinados com os notáveis recursos de diversificação de um método metaheurístico modificado chamado *busca de cuco heterogênea* (“*heterogeneous cuckoo search*”, HCS) para compor o *algoritmo de busca de cuco heterogêneo baseado em biogeografia* (“*biogeography-based heterogeneous cuckoo search algorithm*”, BHCS) [93]. Alternativamente, os autores de [94] aproveitaram a grande capacidade de intensificação do *otimizador de lobo*

cinzento (“*grey wolf optimizer*”, GWO) em conjunto com a eficiente capacidade de exploração do CS para construir um *algoritmo híbrido baseado no otimizador de lobo cinzento e busca de cuco* (“*hybrid algorithm based on gray wolf optimizer and cuckoo search*”, GWOCS) [94]. Mais recentemente, os autores de [98] propuseram o *método de evolução diferencial assistido por grupo social* (“*social group assisted differential evolution*”, SGDE), um método híbrido que também tem se mostrado competitivo em relação a outros métodos de última geração [98]. Em geral, todos esses métodos metaheurísticos híbridos demonstraram em seus testes de validação um desempenho superior ao de seus métodos originários e de variantes correlatas [91–94, 98].

No que se refere à formulação do problema de extração dos parâmetros considerando a abordagem formal do modelo de um diodo, nota-se que, embora seja menos popular do que a abordagem simplificada, um número significativo de métodos metaheurísticos também já foi validado nesse contexto, especialmente nos últimos anos. Nessa formulação, não há um consenso acerca da técnica mais adequada para o cálculo da equação característica do modelo, isto é, se este deve ser realizado via método de Newton ou via função W de Lambert.

Dentre os métodos metaheurísticos básicos já testados considerando o método de Newton para o cálculo do modelo de um diodo, chamam a atenção o *algoritmo de coiote* (“*coyote optimization algorithm*”, COA) [95], o *otimizador de fertilidade das terras agrícolas* (“*farmland fertility optimizer*”, FFO) [96] e a *otimização de busca fractal estocástica* (“*stochastic fractal search optimization*”, SFS) [97]. Em relação aos métodos modificados, duas variantes de PSO se destacam: a *otimização por enxame de partículas com líder aprimorado* (“*enhanced leader particle swarm optimization*”, ELPSO) [99] e a *otimização por enxame de partículas com convergência garantida* (“*guaranteed convergence particle swarm optimization*”, GCPSO) [100]. No que concerne aos métodos híbridos, o *algoritmo de busca interior hibridizada* (“*hybridized interior search algorithm*”, HISA) apresenta resultados interessantes, integrando operadores de DE ao *algoritmo de busca interior* (“*interior search algorithm*”, ISA) [101].

Entre os métodos metaheurísticos já propostos para a resolução do problema em questão considerando a função W de Lambert, sobressaem-se o método metaheurístico modificado conhecido como *algoritmo de bolor limoso melhorado* (“*improved slime mould algorithm*”, ImSMA) [102] e o método metaheurístico híbrido denominado *novo algoritmo de evolução diferencial híbrida e inteligência de colônia de abelhas artificiais* (“*novel hybrid differential evolution and artificial bee colony intelligence*”, nDEBCO) [43]. É importante notar que esse último também foi testado na resolução do problema de extração dos parâmetros considerando a abordagem simplificada do modelo de um diodo [43]. Inclusive, o nDEBCO é o único método (entre determinísticos, metaheurísticos e conjugados) encontrado na literatura que foi originalmente validado tendo em vista as duas abordagens mais comuns de cálculo do modelo de um diodo no escopo da função objetivo do problema de extração dos parâmetros.

2.4.3 Métodos conjugados

Diante do conteúdo até então exposto, é possível perceber que tanto os métodos determinísticos quanto os métodos metaheurísticos apresentam pontos fortes e fracos. Por exemplo, os métodos determinísticos locais podem, em princípio, alcançar de forma bastante eficiente a solução ótima global de um dado problema, mas apenas se forem inicializados suficientemente próximos da região do mínimo global do espaço de soluções. Já os métodos metaheurísticos podem atingir tal solução independentemente de sua inicialização; contudo, à custa de uma quantidade significativamente mais onerosa de avaliações da função objetivo e de iterações e, em geral, sem garantias formais de otimalidade global. Ademais, os desempenhos dos métodos pertencentes a ambas as classes são afetados, em menor ou maior grau, pelo ajuste de seus respectivos parâmetros de controle [113, 120, 121, 148, 152].

Cientes disso, alguns pesquisadores passaram a combinar características complementares desses métodos ou a explorar de forma dirigida as suas potencialidades específicas, com o intuito de formar um novo método de otimização, em geral mais eficiente. No presente documento, tais técnicas são genericamente classificadas como *métodos conjugados*.

Conforme listado na Tabela 6, um número razoável de métodos conjugados já foi proposto para a resolução do problema de extração dos parâmetros. Vale destacar, entretanto, que todos foram validados considerando a abordagem simplificada da equação característica do modelo.

Nesse contexto, observa-se que um dos primeiros métodos conjugados foi o LMSA, apresentado em [104]. Em linhas gerais, seus autores utilizaram o método metaheurístico básico SA para definir, a cada iteração, o valor de um parâmetro de controle do método determinístico local LMA [104]. A fim de viabilizar a efetiva aplicação do LMSA, eles também propuseram um procedimento de inicialização semelhante ao desenvolvido em [64] (veja a Tabela 4) [104].

Não obstante, a maioria dos métodos conjugados propostos após o LMSA adotou um paradigma de concepção exatamente oposto. Em outras palavras, esses métodos buscaram essencialmente usar um método determinístico local para melhorar o processo de intensificação de certo método metaheurístico. Nesse sentido, o método determinístico local de busca direta NMS desponta como o favorito entre os pesquisadores. A implementação relativamente simples do seu algoritmo, aliada ao fato de não exigir o cálculo de derivadas da função objetivo, facilita sobremaneira a sua integração com métodos metaheurísticos que, por sua vez, compartilham essas mesmas características [113].

Dessa forma, os autores do trabalho apresentado em [105] incorporaram o NMS a uma versão modificada do PSO (*“modified particle swarm optimization”*, MPSO) e demonstraram que o método conjugado resultante, referido como NM-MPSO, apresenta desempenho promissor [105]. Já em [106], os autores utilizaram uma *estratégia de águia de três estágios* (*“three-stage eagle strategy”*, TES) para hibridizar o ABC com o NMS. Os resultados reportados mostram que o método conjugado resultante, denominado EHA-NMS, é mais eficiente do que

Tabela 6 – Visão geral dos métodos conjugados já propostos para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$.

Trabalho		Método conjugado					Formulação do problema ^a	
Ref.	Ano	Acrônimo	Determinístico	+	Metaheurístico	Implementação	\mathcal{F}	lb/ub
[104]	2014	LMSA	LMA	+	SA	MATLAB	SSE	Não
[105]	2016	NM-MPSO	NMS	+	MPSO	MATLAB	RMSE	—
[106]	2016	EHA-NMS	NMS	+	ABC+TES	MATLAB	RMSE	Sim
[107]	2017	GOFPANM	NMS	+	FPA+GOBL	MATLAB	RMSE	Sim
[108]	2018	ABC-TRR	TRR	+	ABC	MATLAB	SSE/RMSE	Sim
[109]	2019	MADE	NMS	+	SHADE+RES	MATLAB	RMSE	Sim
[110]	2020	NMSOLMFO	NMS	+	MFO+OLS	MATLAB	RMSE	Sim
[111]	2022	ISNMWOA	NMS	+	WOA+ISSM	MATLAB	RMSE	Sim

^aTodos os trabalhos consideraram a abordagem simplificada do modelo de um diodo.

outros métodos metaheurísticos, incluindo ABSO, CSO e Rcr-IJADE [106]. Posteriormente, em [107], foi proposta a integração do *algoritmo de polinização de flores* (“*flower pollination algorithm*”, FPA) com o NMS e com um *mecanismo de aprendizado baseado em oposição generalizada* (“*generalized opposition-based learning mechanism*”, GOBL), visando melhorar a capacidade de escapar da região de uma solução ótima local [107]. Os autores demonstraram de forma abrangente que o método conjugado resultante, referido como GOFPANM, apresenta desempenho superior ao de outras possíveis variantes desses dois métodos metaheurísticos, bem como ao de outros métodos da mesma classe [107].

Na mesma linha, em [109] foi proposta a introdução do NMS em uma variante do DE conhecida como *evolução diferencial adaptativa baseada no histórico de sucesso* (“*success-history based adaptive differential evolution*”, SHADE), além da adoção de uma *estratégia de eliminação baseada em ranking* (“*ranking-based elimination strategy*”, RES) para reforçar o procedimento de diversificação [109]. O método conjugado resultante, referido como MADE, mostrou-se um pouco mais acurado que o SHADE e mais eficiente que o GOFPANM [109]. Mais recentemente, em [110], foi proposto o uso conjunto do NMS e de uma *estratégia de aprendizagem ortogonal* (“*orthogonal learning strategy*”, OLS) para potencializar os recursos de busca local e exploração da chamada *otimização por atração de mariposa* (“*moth flame optimization*”, MFO) [110]. Os resultados apresentados indicam que o método conjugado resultante, denominado NMSOLMFO, supera o MFO básico e outras de suas variantes metaheurísticas [110]. Adicionalmente, o trabalho apresentado em [111] propôs o método conjugado ISNMWOA, no qual o NMS é associado ao *algoritmo de otimização de baleias* (“*whale optimization algorithm*”, WOA) e a um *mecanismo de busca por compartilhamento de informação* (“*information sharing search mechanism*”, ISSM). Os resultados reportados sugerem que o ISNMWOA é competitivo em relação a outros métodos de última geração baseados em NMS [111].

Por fim, é relevante comentar que um paradigma distinto de concepção é verificado no caso do método conjugado denominado ABC-TRR, proposto em [108]. Nesse caso, a solução final encontrada pelo método metaheurístico básico ABC é utilizada como estimativa inicial

para a aplicação do método determinístico local TRR [108]. Assim, trata-se de um acoplamento sequencial entre um método metaheurístico e um método determinístico, em que o primeiro pode ser interpretado como um procedimento de inicialização do segundo.

2.4.4 Observações complementares

Ao analisar as Tabelas 3, 5 e 6, constata-se que a formulação do problema de extração dos parâmetros é, em alguma medida, influenciada pelo método de otimização previsto para ser aplicado na sua resolução.

No que se refere à função objetivo, nota-se que a SSE, dado por (2.11), é a métrica de erro mais comumente adotada quando o método previsto para a resolução do problema é determinístico. É importante perceber, entretanto, que essa preferência é parcialmente induzida pelos próprios *solvers* do MATLAB, os quais adotam tal métrica de erro como função objetivo padrão [120, 121]. Em contrapartida, observa-se que a RMSE, dado por (2.12), é a métrica de erro mais frequentemente utilizada quando o método previsto é metaheurístico ou conjugado. Se, por um lado, essa métrica de erro é computacionalmente mais onerosa do que a SSE, por outro, fornece informações fundamentais para estudos que envolvem análise de risco, isto é, fornece uma medida da incerteza associada aos resultados retornados pelo modelo de um diodo [147]. Dessa forma, acredita-se que muitos pesquisadores tenham preferência pela RMSE porque tal métrica oferece, de imediato, um resultado potencialmente mais proveitoso.

A respeito das restrições de canalização, observa-se que estas são geralmente adotadas quando o método previsto é caracterizado por uma estratégia de busca global da solução ótima. Em particular, quando se trata de um método determinístico global, de um método metaheurístico ou de um método conjugado em sua forma convencional, isto é, à exceção do LMSA. Vale ressaltar que essas restrições são essenciais para delimitar o espaço de soluções ótima do problema. Além disso, auxiliam no controle do sorteio das estimativas iniciais a serem consideradas na aplicação dessas duas últimas classes de métodos [85, 110]. Portanto, as restrições de canalização são, em essência, indispensáveis à aplicação dos métodos em questão.

Em contrapartida, verifica-se que as restrições de canalização costumam ser desconsideradas quando o método previsto é, em última instância, caracterizado por uma estratégia de busca local. Mais especificamente, quando se trata de um método determinístico local ou de um método conjugado não convencional, tal como o LMSA. Uma possível explicação para essa prática reside nas limitações intrínsecas dos métodos determinísticos envolvidos. Por exemplo, o método de Newton-Raphson e o LMA foram originalmente desenvolvidos para resolver problemas de otimização irrestritos [113, 120, 121, 149]. Consequentemente, modificações precisariam ser adequadamente executadas em seus algoritmos para que passassem a lidar de forma apropriada com as restrições de canalização. Outra possível explicação relaciona-se à própria necessidade de inicializações de alta qualidade exigida por esses métodos. Como as estimativas iniciais consideradas para suas aplicações tendem a ser asseguradamente próximas do mínimo global

do espaço de soluções do problema, esses métodos tendem a convergir diretamente para essa solução, tornando, em muitos casos, menos relevantes os limites estabelecidos pelas restrições de canalização. Ainda assim, é pertinente ressaltar que, na ausência dessas restrições, qualquer método de otimização se encontra sujeito a convergir para uma solução final que pode ser, a rigor, infactível. Logo, o descarte das restrições de canalização não é uma prática recomendável.

Ademais, é relevante chamar a atenção para o fato de o MATLAB se apresentar como o ambiente de programação preferencial dos pesquisadores da área. Além de ser amplamente difundido no meio acadêmico, seus *solvers* dedicados facilitam sobremaneira o desenvolvimento de estudos, tal como foi diversas vezes destacado ao longo do conteúdo anteriormente exposto [120, 121, 156, 163, 165, 167].

Por fim, é pertinente registrar que, no decorrer da atividade de revisão dos trabalhos citados ao longo desta seção, foi possível notar a ocorrência sistemática de alguns equívocos no que diz respeito à forma de validação de métodos de otimização propostos para a resolução do problema de extração dos parâmetros. Nesse contexto, é prática comum comparar as soluções alcançadas por um novo método, em estudos de caso proeminentes, com aquelas previamente reportadas para outros métodos de estado da arte. Todavia, observou-se que essa comparação tem sido muitas vezes realizada sem levar em consideração sob qual abordagem de cálculo do modelo de um diodo essas soluções foram originalmente obtidas.

Em particular, nos trabalhos apresentados em [95], [96], [97], [99], [100], [101] e [102], os métodos metaheurísticos propostos foram aplicados à resolução de conhecidos problemas de extração dos parâmetros considerando a abordagem formal do modelo de um diodo (veja a Tabela 5). Não obstante, as soluções finais alcançadas por esses métodos foram diretamente comparadas com as soluções reportadas para outros métodos avançados, sendo que alguns, quando não todos estes, haviam sido aplicados considerando a abordagem simplificada do modelo de um diodo. Conforme discutido na Seção 2.3, as soluções verificadas para um mesmo problema de extração dos parâmetros considerando a abordagem formal do modelo de um diodo tendem, por construção, a ser mais acuradas do que aquelas obtidas considerando a sua abordagem simplificada. Logo, os autores desses trabalhos se valeram de comparações conceitualmente equivocadas para afirmar que os métodos metaheurísticos que estavam propondo seriam capazes de alcançar soluções finais mais acuradas para os problemas avaliados do que aquelas disponíveis até o momento na literatura.

Infelizmente, falhas análogas também foram identificadas em alguns trabalhos de revisão contemporâneos. Por exemplo, tanto em [35] como em [169] são apresentadas comparações das soluções finais alcançadas por vários métodos de otimização para um problema de extração dos parâmetros tradicional na literatura, sem qualquer ponderação com relação às abordagens de cálculo do modelo de um diodo que foram consideradas na aplicação desses métodos. Tais comparações indiscriminadas, expostas em periódicos internacionais indexados e de grande credibilidade, podem induzir pesquisadores da área a conclusões precipitadas ou mesmo errôneas

a respeito do desempenho relativo dos métodos de otimização já propostos para a resolução do problema em questão e, de forma preocupante, prejudicar a condução de novas pesquisas.

2.5 ANÁLISE COMPREENSIVA

A fim de estabelecer uma base comum para discussão, considere o tradicional estudo de caso da célula fotovoltaica de tecnologia de silício cristalino ($N_p = N_s = 1$), denominada *RTC France*. Uma curva $I-V$ medida com esse dispositivo operando a 33°C ($T_c = 306,15\text{ K}$) foi reportada, há cerca de quatro décadas, no trabalho apresentado em [63]. Conforme mostrado na Figura 5, tal curva é descrita por $N = 26$ pares de medições. É importante observar que alguns desses pares se encontram em regiões operativas incomuns ou indesejáveis na prática, definidas por valores negativos de tensão e de corrente elétrica. Não obstante, essas medições oferecem informações adicionais sobre o comportamento elétrico do gerador fotovoltaico em estudo [6,23]. Por essa razão, todos os pares de medições que descrevem a curva em questão são, em geral, considerados no escopo do problema de extração dos parâmetros da célula RTC France.

No que se refere à formulação desse problema, é pertinente comentar que as restrições de canalização, quando adotadas, são frequentemente definidas por $\mathbf{lb} = [0, 0, 1, 0, 0]$ e $\mathbf{ub} = [1, 1, 2, 100, 0,5]$, com os limites de I_{ph} dados em ampères (A), os limites de I_0 dados em microampères (μA) e os limites de R_{sh} e R_s dados, respectivamente, em ohms (Ω) [85]. Ao considerar os valores tipicamente assumidos pelos cinco parâmetros do modelo de um diodo, mencionados na Subseção 2.1.3, constata-se que alguns desses limites mínimos e máximos se mostram substancialmente relaxados. Um exemplo evidente ocorre com I_{ph} : um valor próximo de 0 para esse parâmetro é inconcebível, pois a corrente de curto-circuito indicada pela curva $I-V$ disponível para a célula RTC France é de aproximadamente 0,7 A — veja a Figura 5. Apesar disso, observa-se na prática que tais limites oferecem a flexibilidade necessária para que um método de otimização aplicado ao problema caminhe oportunamente no espaço de soluções assim delimitado e convirja para uma solução ótima.

Posto isso, as soluções finais de maior qualidade formalmente alcançadas pelos diferentes métodos de otimização apresentados na seção anterior, para o problema de extração dos parâmetros da célula RTC France, são analisadas na sequência. Destaca-se, entretanto, que essa análise é conduzida distinguindo explicitamente as diferentes abordagens de cálculo do modelo de um diodo.

2.5.1 Resultados verificados na abordagem simplificada

As soluções reportadas para os cinco parâmetros do modelo de um diodo, considerando a abordagem simplificada de sua equação característica, são sintetizadas na Tabela 7. A qualidade dessas soluções é indicada em termos de RMSE, uma vez que essa é a métrica de erro mais comumente adotada como função objetivo — veja a Seção 2.4. Além disso, registram-se observações acerca das condições de convergência sob as quais essas soluções foram obtidas.

Tabela 7 – Soluções ótimas alcançadas por diferentes métodos de otimização para o problema de extração dos parâmetros da célula RTC France, considerando a abordagem simplificada do modelo de um diodo. Note que estas estão segregadas de acordo com a classe correspondente do método.

Método	I_{ph} (A)	I_0 (μ A)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)	Observações ^d
NMMLP [63]	0,7608	0,3223	1,4837	53,7634	0,0364	9,6964E-03	
IABBA [68]	0,760771090719	0,323459337407	1,48204583019	53,7981707126	0,0363725593061	9,8603E-04	
IABBA [52]	0,760779120136	0,322873926858	1,48113747635	53,7009537057	0,0363792207867	9,860250417458982E-04	TS = 13.547s
IGHS [81]	0,76077	0,34351	1,48740	53,2845	0,03613	9,9306E-04	ITm = 5.000
ABSO [72]	0,76080	0,30623	1,47583	52,2903	0,03659	9,9124E-04	ITm = 5.000
BMO [73]	0,76077	0,32479	1,48173	53,8716	0,03636	9,8608E-04	ITm = 5.000
CS [74]	0,7608	0,323	1,4812	53,7185	0,0364	1,0E-03	ITm = 5.000
ReI-JADE [82]	0,760776	0,323021	1,481184	53,718526	0,036377	9,8602E-04	FEm = 10.000, RD = 100
BFA [75]	0,7602	0,8000	1,6951	50,8691	0,0325	2,9E-02	ITm = 10.000, RD = 35
ABC [75]	0,7608	0,3251	1,4817	53,6433	0,0364	9,862E-04	ITm = 10.000, RD = 35
CSO [77]	0,76078	0,3230	1,48118	53,7185	0,03638	9,8602E-04	ITm = 500, RD = 50
TLABC [92]	0,76078	0,32302	1,48118	53,71636	0,03638	9,86022E-04	FEm = 50.000, RD = 30
DE/WOA [91]	0,760776	0,323021	1,481184	53,718524	0,036377	9,860219E-04	FEm = 50.000, RD = 50
SDO [80]	0,7608	0,3230	1,4812	53,7185	0,0364	9,8602E-04	FEm = 50.000, RD = 50
PGJAYA [83]	0,7608	0,3230	1,4812	53,7185	0,0364	9,8602E-04	FEm = 50.000, RD = 30
ITLBO [84]	0,7608	0,3230	1,4812	53,7185	0,0364	9,8602E-04	FEm = 50.000, RD = 30, TT = 5,95s
CPMPSO [86]	0,760776	0,323021	1,481184	53,71852	0,036377	9,860219E-04	FEm = 50.000, RD = 30
MCSWOA [87]	0,7608	0,3230	1,4812	53,7185	0,0364	9,8602E-04	FEm = 50.000, RD = 50
BHCS [93]	0,76078	0,32302	1,48118	53,71852	0,03638	9,86022E-04	FEm = 50.000, RD = 30
ELBA [85]	0,760776	0,323021	1,481185	53,718523	0,036377	9,860219E-04	FEm = 50.000, RD = 30, TS = 0,407s, TT = 12,23s
GWOCs [94]	0,760773	0,32192	1,4808	53,6320	0,03639	9,8607E-04	FEm = 50.000, RD = 30, TS = 3,89s
IMPA [88]	0,760775530386165	0,3230208166104389	1,4811835921250962	53,71852391990669	0,03637709258093378	9,860218778914944E-04	FEm = 20.000, RD = 30, TS = 3,89s
EABOA [89]	0,760771077	0,322929	1,481153457	53,76600144	0,036379593	9,8602E-04	FEm = 50.000, RD = 30, TT = 74,409s
nDEBCO [43]	0,760775530434227	0,323020774222916	1,48118068229836	53,7185206085772	0,0363770931030067	9,86021877891545E-04	FE = 2.050, FEm = 50.000, RD = 30, TC = 0,134s
NM-MPSO [105]	0,76078	0,32306	1,48120	53,7222	0,03638	9,8602E-04	ITm = 5.000
EHA-NMS [106]	0,760776	0,323021	1,481184	53,718521	0,036377	9,8602E-04	FEm = 5.000, RD = 1.000
GOPANM [107]	0,7607755	0,3230208	1,4811836	53,7185203	0,0363771	9,860219E-04	FEm = 10.000, RD = 30
ABC-TRR [108]	0,760776	0,323021	1,481184	53,718521	0,036377	9,8602E-04	FEm = 1.000, RD = 1.000
MADE [109]	0,7608	0,3230	1,4812	53,7185	0,0364	9,8602E-04	FEm = 5.000, RD = 30, TS = 0,1267s
NMSOLMFO [110]	0,760775529	0,323021	1,481183605	53,71853135	0,036377092	9,8602E-04	FEm = 20.000, RD = 30

^dLegenda: FE = número de avaliações da função objetivo, FEm = número máximo de avaliações da função objetivo, IT = número de iterações, ITm = tempo de iterações, RD = quantidade de aplicações repetidas, TC = tempo médio de simulação para alcançar a solução final, TS = tempo (médio) de simulação de uma aplicação, TT = tempo de simulação do conjunto de aplicações.

A respeito dessas soluções reportadas, é pertinente destacar que:

- A provável solução ótima global para cada parâmetro do modelo de um diodo é dada — considerando os seis dígitos mais significativos — por $I_{ph} = 0,760776$ A, $I_0 = 0,323021$ μ A, $n = 1,481185$, $R_{sh} = 53,718523$ Ω e $R_s = 0,036377$ Ω , o que corresponde a um valor de RMSE próximo de $9,860219 \times 10^{-4}$ A.
- As soluções alcançadas pelo IABBA apresentam valor de RMSE cerca de uma ordem de grandeza inferior à verificada para a solução obtida pelo NMMLP. Embora seja tentador atribuir essa diferença exclusivamente ao fato de o primeiro ser um método determinístico global e o segundo um método determinístico local (veja a Tabela 3), tal conclusão não pode ser feita com segurança. O NMMLP foi aplicado em meados da década de 1980, enquanto o IABBA foi aplicado apenas em 2020; logo, há uma distância temporal e tecnológica significativa entre ambos, que inevitavelmente influencia essa comparação.
- Por oferecer garantias teóricas de otimalidade, apresentar efetividade impecável e ter sido aplicado sob condições de teste equivalentes e por meio do mesmo *solver* [52, 68], seria razoável esperar que as duas aplicações do IABBA convergissem para uma mesma solução final. No entanto, observa-se que nenhuma destas alcançou sequer a provável solução ótima global do problema. Essa situação revela uma certa incipiência tecnológica associada a esse método.
- Os métodos metaheurísticos inicialmente aplicados à resolução do problema, em sua maioria métodos básicos, foram capazes de alcançar soluções com diferentes níveis de qualidade. Em contraste, os métodos metaheurísticos aplicados mais recentemente, caracterizados predominantemente por versões modificadas e híbridas (veja a Tabela 5), convergiram para a provável solução ótima global do problema. Essa evolução indica um ganho real de maturidade dessa classe de métodos de otimização e corrobora a pertinência das modificações e hibridizações introduzidas a partir de algoritmos básicos.
- Todos os métodos conjugados convergiram para a provável solução ótima global. Considerando que esses métodos começaram a ser desenvolvidos apenas nos últimos anos e reúnem, de forma mais equilibrada, recursos eficientes de busca local e de exploração global do espaço de soluções, esse nível de desempenho e acurácia já era, em certa medida, esperado.

No que concerne à eficiência com a qual cada método de otimização alcançou as respectivas soluções reportadas, é necessário adotar certo cuidado ao estabelecer qualquer comparação direta. Conforme destacado nas observações apresentadas na Tabela 7, muitos desses métodos foram aplicados sob critérios de convergência distintos. Além disso, todos foram possivelmente executados a partir de diferentes estimativas iniciais. Somam-se a isso questões relacionadas à forma de implementação computacional, à versão do *software* utilizada nas simulações e às

especificações da máquina empregada, as quais podem influenciar de maneira expressiva o tempo de simulação. Apesar dessas ressalvas, as referidas observações revelam algumas peculiaridades e tendências das diferentes classes de métodos de otimização, a saber:

- Os métodos determinísticos globais demandam, em geral, tempos de simulação demasiadamente longos para convergir para uma solução final. Por exemplo, o IABBA, mesmo sendo aplicado por meio de um *solver*, que provavelmente segue boas práticas de implementação computacional e otimização de recursos, demandou algumas horas para retornar uma solução subótima. Em contraste, o método metaheurístico modificado ELBA, implementado pelo autor do presente documento, mostrou-se capaz de retornar a provável solução ótima do problema em um tempo médio de simulação inferior a meio segundo [85]. Situação semelhante é verificada para o método conjugado MADE, o qual demonstrou ser capaz de alcançar essa mesma solução em uma fração ainda menor de segundo [109].
- Os métodos metaheurísticos são regularmente aplicados considerando, como critério de convergência, um número máximo da ordem de 50.000 avaliações da função objetivo. Já os métodos conjugados costumam ser aplicados com limites entre 1.000 e 20.000 avaliações da função objetivo. Essa diferença, somada às comparações de tempo apresentadas no item anterior, ajuda a corroborar o maior nível de eficiência usualmente observado para os métodos conjugados em relação aos métodos metaheurísticos.
- A maioria dos pesquisadores costuma realizar em torno de 30 aplicações do método metaheurístico ou conjugado proposto, de modo a viabilizar análises estatísticas de seu nível de eficiência, estabilidade e robustez. Embora os detalhes dessas análises transcendam o escopo deste capítulo, registra-se que tais procedimentos são didaticamente conduzidos no trabalho apresentado em [85], o qual é assinado pelo mesmo autor do presente documento.

2.5.2 Resultados verificados na abordagem formal

As soluções reportadas para os cinco parâmetros do modelo de um diodo, considerando a abordagem formal de sua equação característica, são sintetizadas na Tabela 8. Tal como feito anteriormente, a qualidade dessas soluções é indicada em termos de RMSE, sendo apresentadas também algumas observações acerca das condições de convergência sob as quais foram obtidas.

Sobre essas soluções reportadas, é pertinente destacar que:

- Como as técnicas de cálculo aplicadas no âmbito da abordagem formal do modelo de um diodo são, em última instância, baseadas em métodos numéricos, as soluções reportadas exibem pequenos ruídos nas casas decimais mais avançadas. Ainda assim, é possível declarar que a provável solução ótima global para cada parâmetro do modelo de um diodo se aproxima de $I_{ph} = 0,760788$ A, $I_0 = 0,310685$ μ A, $n = 1,477265$, $R_{sh} = 52,889791$ Ω e $R_s = 0,036547$ Ω , sugerindo um valor de RMSE próximo de $7,730063 \times 10^{-4}$ A.

Tabela 8 – Soluções ótimas alcançadas por diferentes métodos de otimização para o problema de extração dos parâmetros da célula RTC France, considerando a abordagem formal do modelo de um diodo. Note que estas estão segregadas de acordo com a classe correspondente do método.

Método	I_{ph} (A)	I_0 (μ A)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)	Observações ^d
LMA [53]	0,7607884	0,3102482	1,4769641	52,859056	0,03655304	7,7301E-04	FE = 138, IT = 16
TRR [54]	0,760788	0,3106847	1,477105	52,890468	0,036547	7,7301E-04	FE = 36, IT = 8
TRR [64]	0,76078797	0,31068485	1,4771052	52,889804	0,036546942	7,7300627E-04	FE = 24, IT = 3
TRR [65]	0,7607880	0,3106884	1,4771063	52,8900296	0,0365469	7,7300627E-04	FE = 18, IT = 2
TRR [66]	0,760787954	0,310698	1,477109318	52,89061647	0,036546759	7,7300628E-04	FE = 48, IT = 7
ELPSO [99]	0,760788	0,310677	1,475256	52,889336	0,036547	7,7301E-04	FE _m = 151.500, RD = 30
GCPPO [100]	0,76078796	0,310685316	1,47726802	52,88987895	0,03654694	7,730063E-04	IT _m = 10.000, RD = 100, TS = 98s
HISA [101]	0,7607879660	0,3106845918	1,47726778476	52,8897942635	0,03654694543	7,730062689941E-04	IT _m = 1.000, RD = 50
COA [95]	0,76076929153	0,3083945801266	1,47654776591	52,82666150326	0,03655460766	7,75470161606E-04	IT _m = 1.000, RD = 30
FFO [96]	0,7608	0,3223	1,5215	53,8989	0,0364	7,75184E-04	IT _m = 1.500, RD = 100
ImSMA [102]	0,76079	0,31068	1,51770	52,89000	0,03655	7,73006E-04	IT _m = 500, RD = 100, TS = 3,05
SFS [97]	0,7609	0,3167	1,47918	53,2805	0,03648	7,931E-04	IT _m = 400
nDEBCO [43]	0,760787966601641	0,310684633139038	1,47726490925928	52,8897913461548	0,0365469447815297	7,73006268994108E-04	FE = 2.050, FE _m = 50.000, RD = 30, TS = 0,755s

^dLegenda: FE = número de avaliações da função objetivo, FE_m = número máximo de avaliações da função objetivo, IT = número de iterações, IT_m = número máximo de iterações, RD = quantidade de aplicações repetidas, TC = tempo médio de simulação para alcançar a solução final, TS = tempo (médio) de simulação de uma aplicação.

Tabela 9 – Estimativas iniciais consideradas na aplicação do LMA e do TRR na resolução do problema de extração dos parâmetros da célula RTC France.

Método	I_{ph} (A)	I_0 (μ A)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)
LMA [53]	0,761060	0,290125	1,470097	49,973561	0,036800	8,8437E-4
TRR [54]	0,760788	0,3106845	1,477105	52,890463	0,036547	7,7301E-4
TRR [64]	0,76074014	0,31285196	1,4777295	55,907380	0,036615485	7,9439241E-4
TRR [65]	0,7607191	0,3319572	1,4837043	54,2878608	0,0361310	8,0737695E-4
TRR [67]	0,76078734	0,31075161	1,4778224	52,909893	0,036547082	7,730072975E-4

- Os métodos determinísticos locais alcançaram soluções finais com nível de qualidade comparável ao da provável solução ótima. Observa-se, contudo, que os valores dos parâmetros retornados pelo LMA se afastam de forma mais expressiva daqueles obtidos pelas diferentes aplicações do TRR. Uma possível justificativa para essa diferença reside nas próprias potencialidades de otimização de cada um desses métodos determinísticos locais; outra está nas estimativas iniciais consideradas em suas respectivas aplicações. Como mostrado na Tabela 9, as discrepâncias entre essas estimativas iniciais não são desprezíveis, ainda que todas já apresentem valores de RMSE bastante reduzidos. De modo análogo, as diferenças observadas nas casas decimais mais avançadas das soluções alcançadas por cada aplicação do TRR podem ser atribuídas às diferentes inicializações e aos tratamentos adicionais eventualmente atribuídos ao modelo de um diodo, como normalizações em pu ou em valor unitário.
- Os métodos metaheurísticos conseguiram alcançar soluções finais com nível de qualidade muito próximo ao da provável solução ótima. Todavia, verifica-se que as soluções obtidas pelos métodos metaheurísticos básicos são ligeiramente inferiores àquelas retornadas pelos métodos metaheurísticos modificados e híbridos (veja a Tabela 5). Essa constatação reforça, mais uma vez, a relevância do desenvolvimento dessas versões avançadas.
- Como era de se esperar, a técnica utilizada para resolver a abordagem formal do modelo de um diodo não impacta a solução ótima do problema em análise. Um exemplo claro disso pode ser obtido comparando as soluções resultantes dos métodos metaheurísticos híbridos HISA e nDEBCO (veja novamente a Tabela 5). O primeiro recorreu ao método de Newton, enquanto o segundo utilizou a função W de Lambert. Independentemente dessa diferença, ambos convergiram para soluções finais muito semelhantes, em plena concordância com a provável solução ótima do problema.

Com relação às observações apresentadas na Tabela 8, é relevante chamar a atenção para os seguintes fatos:

- Os métodos determinísticos locais se mostram capazes de alcançar uma solução final para o problema demandando um número muito menor de avaliações da função objetivo e de iterações, em comparação com os métodos metaheurísticos. Por exemplo, na sua aplicação

menos eficiente, o TRR demandou 48 avaliações da função objetivo para retornar uma solução com qualidade comparável à da provável solução ótima [66]. Já o método metaheurístico híbrido nDEBCO necessitou de cerca de 2.150 avaliações da função objetivo, dentre as 50.000 permitidas pelo seu critério de convergência, para alcançar uma solução de qualidade semelhante [43]. Há de se ponderar, no entanto, que o primeiro foi aplicado considerando uma estimativa inicial com valor de RMSE bastante reduzido, como indicado na Tabela 9, ao passo que o último foi aplicado considerando estimativas iniciais de menor qualidade, conforme revelado em [43].

- A aplicação do TRR reportada em [54] apresenta um dos piores indicadores de eficiência, mesmo tendo sido realizada a partir da melhor estimativa inicial conhecida (veja a Tabela 9). Em contrapartida, a aplicação do TRR reportada em [65] exhibe os melhores indicadores de eficiência, embora tenha sido conduzida a partir da estimativa inicial de pior qualidade entre as consideradas (veja novamente a Tabela 9). Uma possível justificativa para essa situação, à primeira vista inesperada, reside no tratamento em *pu* atribuído ao modelo de um diodo no âmbito desse último trabalho (veja a Subseção 2.4.1). Outra possibilidade é que os autores desses estudos tenham adotado configurações distintas para um ou mais parâmetros de controle do TRR, já que tais valores não foram explicitados em nenhuma das publicações.
- Os métodos metaheurísticos são frequentemente aplicados considerando, como critério de convergência, um número máximo de iterações, sem que haja, entretanto, consenso sobre qual deva ser esse limite. Vale lembrar que, no caso da abordagem simplificada, observou-se que esse critério costuma corresponder a aproximadamente 50.000 avaliações da função objetivo (veja a Tabela 7).

2.5.3 Discussão

Ao comparar os resultados verificados para os diferentes métodos de otimização, considerando a abordagem simplificada do modelo de um diodo, com aqueles obtidos sob a abordagem formal, constata-se que estes apontam para soluções ótimas distintas para o problema de extração dos parâmetros da célula RTC France.

A explicação mais plausível para essa situação reside no fato de os problemas de otimização resultantes de cada uma dessas abordagens não serem equivalentes. Uma maneira simples de confirmar essa hipótese é aplicar a solução ótima apontada em determinada abordagem do modelo de um diodo diretamente na formulação correspondente à outra abordagem. Nesse sentido, considere as soluções apresentadas nas Tabelas 7 e 8 para o método metaheurístico híbrido nDEBCO. Além de esse ser o único método de otimização encontrado que foi originalmente validado considerando ambas as abordagens, destaca-se que este alcançou, em ambos os casos, soluções finais compatíveis com as prováveis soluções ótimas globais verificadas, em conjunto com os demais métodos, para o respectivo problema.

Tabela 10 – Valores de RMSE relativos a cada abordagem do modelo de um diodo, considerando as respectivas soluções alcançadas pelo nDEBCO.

Solução verificada	Função objetivo (A)	
	Abordagem simplificada	Abordagem formal
Abordagem simplificada	9,860219E-04	7,754348E-04
Abordagem formal	9,892442E-04	7,730063E-04

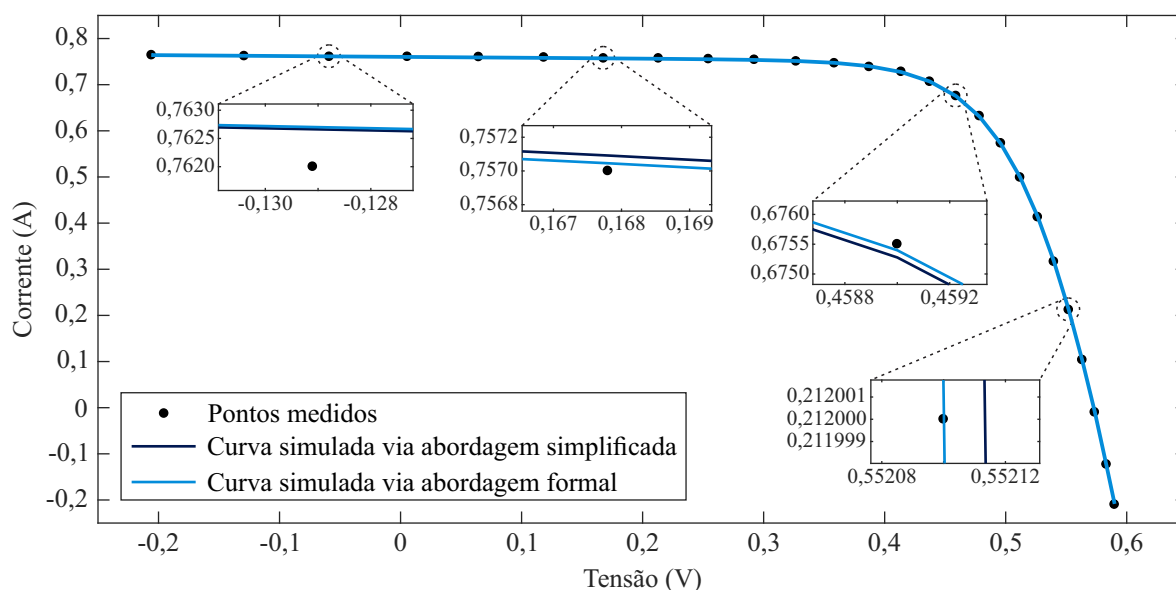
Conforme indicado na Tabela 10, ao se substituir a solução alcançada pelo nDEBCO considerando a abordagem formal na formulação da abordagem simplificada, isto é, em (2.17), obtém-se que a função objetivo do problema é dada por $RMSE = 9,892442 \times 10^{-4}$ A. Esse valor é superior ao verificado para a solução alcançada por esse mesmo método considerando a própria abordagem simplificada, qual seja, $RMSE = 9,860219 \times 10^{-4}$ A. Constata-se, portanto, que a solução ótima apontada para o problema de otimização resultante da abordagem formal é subótima para o problema de otimização resultante da abordagem simplificada.

Por sua vez, ao se substituir a solução alcançada pelo nDEBCO considerando a abordagem simplificada em uma das formulações que definem a abordagem formal, isto é, (2.19) no caso do método de Newton ou (2.21) no caso da função W de Lambert, encontra-se que a função objetivo do problema é dada por $RMSE = 7,754348 \times 10^{-4}$ A. Novamente, observa-se que esse valor é maior do que o verificado para a solução alcançada por esse mesmo método considerando a própria abordagem formal, qual seja, $RMSE = 7,730063 \times 10^{-4}$ A. Logo, constata-se que a solução ótima apontada para o problema de otimização resultante da abordagem simplificada é subótima para aquele resultante da abordagem formal.

Essas constatações ratificam, portanto, que os problemas de otimização resultantes de cada abordagem de cálculo do modelo de um diodo são, de fato, distintos. Posto isso, surge imediatamente a seguinte questão: “*qual dessas abordagens do modelo de um diodo é a mais adequada e, conseqüentemente, quais soluções verificadas para os seus cinco parâmetros são as mais apropriadas?*”

Conforme ilustrado na Figura 8, as soluções alcançadas pelo nDEBCO em ambas as abordagens permitem que o modelo de um diodo reproduza de forma satisfatória a curva $I-V$ da célula RTC France. Entretanto, tal como destacado, a curva simulada considerando a abordagem formal representa de maneira mais fiel os pontos operativos que descrevem essa curva experimental. Justamente por esse motivo se verifica que o valor de RMSE correspondente, isto é, $RMSE = 7,730063 \times 10^{-4}$ A, é inferior àquele obtido considerando a abordagem simplificada, ou seja, $RMSE = 9,860219 \times 10^{-4}$ A. Além disso, é fundamental lembrar que, sob a perspectiva teórica, as soluções verificadas na abordagem formal para os cinco parâmetros tendem a ser mais apropriadas, pois não resultam de idealizações sobre a qualidade da curva $I-V$ em questão ou sobre o potencial representativo do modelo de um diodo, tal como ocorre em algum grau no caso da abordagem simplificada. Portanto, a abordagem formal do modelo de um diodo é aquela capaz de conduzir às soluções mais adequadas para o problema em questão.

Figura 8 – Reprodução da curva $I-V$ da célula de silício RTC France, considerando as diferentes soluções ótimas alcançadas pelo nDEBCO para os cinco parâmetros do modelo de um diodo.



Não obstante, é pertinente fazer uma ponderação: a julgar pelo desempenho do nDEBCO, a abordagem formal do modelo de um diodo via função W de Lambert apresenta um custo computacional significativamente mais elevado do que a abordagem simplificada. Mais especificamente, verifica-se, nos resultados apresentados na Tabela 8, que esse método demandou 2.150 avaliações da função objetivo e 0,799 segundos para alcançar uma solução final comparável à provável solução ótima global do problema, considerando a referida abordagem formal do modelo de um diodo [43]. Já na abordagem simplificada, observa-se, a partir dos resultados indicados na Tabela 7, que esse mesmo método demandou 2.050 avaliações da função objetivo e apenas 0,134 segundos para alcançar a provável solução ótima global [43]. A razão entre o tempo de simulação e o número de avaliações da função objetivo mostra-se, portanto, expressivamente maior no caso da abordagem formal em comparação com o da abordagem simplificada, evidenciando o impacto adicional do custo de cálculo associado à função W de Lambert.

2.6 CONCLUSÕES PARCIAIS

Este capítulo revisou de forma criteriosa o problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico. Nesse sentido, a equação característica do modelo de um diodo foi propriamente desenvolvida, sendo analisados os principais aspectos associados à sua estrutura, bem como aos seus cinco parâmetros. De maneira igualmente fundamental, a formulação típica do problema de extração desses parâmetros foi sistematizada, explicitando-se a construção da função objetivo a partir dos pontos que descrevem a curva $I-V$ e o papel das restrições de canalização na obtenção de soluções fisicamente admissíveis. Uma vez estabelecida essa base, foram elucidadas as principais abordagens de cálculo empregadas para viabilizar a avaliação da equação característica do modelo

de um diodo no escopo da função objetivo. Adicionalmente, cerca de cinquenta métodos de otimização propostos na literatura para a sua resolução foram documentados, tendo em vista as suas classes e a abordagem de cálculo adotada em seus respectivos testes de validação. Por fim, as melhores soluções reportadas por esses métodos para um estudo de caso amplamente reconhecido foram devidamente comparadas. Essa comparação permitiu caracterizar o estado da arte desses métodos em termos de acurácia, eficiência e maturidade tecnológica. Ademais, evidenciou que os problemas de otimização resultantes de cada abordagem de cálculo do modelo de um diodo são, de fato, distintos.

À luz das discussões desenvolvidas ao longo desta revisão, diversas lacunas na literatura e oportunidades de pesquisa foram sendo progressivamente identificadas. De modo preocupante, verificou-se que a falta de clareza quanto à coexistência de diferentes abordagens de cálculo do modelo de um diodo, bem como às suas implicações no problema de extração dos parâmetros, tem conduzido, em alguns estudos, a comparações metodologicamente inconsistentes. Esse quadro indica a necessidade de um reexame crítico de parte dos resultados reportados. Além disso, outras oportunidades relevantes incluem a:

- Proposição de metodologias para determinar limites inferiores e superiores admissíveis para os parâmetros do modelo de um diodo, com vistas a subsidiar formulações mais robustas e aplicações mais eficientes de diferentes métodos de otimização;
- Reavaliação da aplicação de métodos determinísticos globais em estudos de caso amplamente difundidos na literatura, uma vez que as soluções por estes obtidas podem servir como referência sólida para a validação de outras alternativas a serem propostas para a resolução do problema de extração dos parâmetros;
- Condução de estudos comparativos mais abrangentes sobre os procedimentos de inicialização já propostos para a aplicação de métodos determinísticos locais, visando à análise sistemática de seus respectivos desempenhos.

Não obstante, dentre as lacunas e oportunidades identificadas, destacam-se aquelas que se alinham diretamente aos objetivos específicos desta tese. Para fins de síntese e delimitação do escopo, tais pontos são formalizados a seguir:

1. Uma vez verificado que a abordagem formal do modelo de um diodo é aquela capaz de conduzir às soluções mais adequadas para o problema de extração dos parâmetros, torna-se pertinente investigar de forma mais aprofundada a sua aplicabilidade, em particular no que se refere ao emprego da função W de Lambert. Em contraponto à resolução implícita da equação característica via método de Newton, essa alternativa mostra-se promissora por viabilizar a realização de estudos não apenas numéricos, mas também analíticos. Assim, torna-se pertinente avaliar tanto os aspectos subjacentes da função W de Lambert quanto o

desempenho dos principais métodos de cálculo do seu valor no contexto da abordagem formal do modelo de um diodo.

2. Diante de sua consistência conceitual, de seu potencial de eficiência e de sua reprodutibilidade intrínseca, torna-se pertinente investigar a aplicação de um método determinístico local, baseado em gradiente, que incorpore explicitamente restrições de canalização na resolução do problema de extração dos parâmetros. A imposição dessas restrições tende a conferir maior segurança ao processo iterativo e a reduzir a dependência em relação à qualidade da estimativa inicial. Ademais, à luz de tendências observadas em outras classes de métodos, mostra-se oportuno considerar o desenvolvimento de uma alternativa determinística local concebida especificamente para esse propósito, uma vez que não foi encontrado, na literatura, um trabalho com tal escopo.

Por fim, ressalta-se que parte significativa da revisão sistematizada neste capítulo foi previamente publicada em [170], o que corrobora a materialidade e a relevância do conteúdo aqui registrado, bem como das constatações deste decorrentes.

3 FORMALIZAÇÃO DO PROBLEMA DE EXTRAÇÃO DOS PARÂMETROS

Este capítulo estabelece as bases matemáticas que sustentam a formulação adotada para o problema de extração dos parâmetros no âmbito desta pesquisa, partindo da premissa de que se deve empregar a abordagem formal do modelo de um diodo via função W de Lambert. Nesse sentido, a Seção 3.1 apresenta as propriedades fundamentais dessa função especial e delimita o seu domínio fisicamente relevante. Em seguida, a Seção 3.2 declara propriamente a formulação do problema e analisa suas propriedades matemáticas. Na sequência, a Seção 3.3 discute questões de natureza numérica que podem influenciar de maneira decisiva a sua resolução. Ao final, a Seção 3.4 sintetiza as implicações dessas análises para os desenvolvimentos metodológicos apresentados nos capítulos subsequentes.

3.1 ABORDAGEM DO MODELO DE UM DIODO VIA FUNÇÃO W DE LAMBERT

A expressão decorrente da abordagem da equação característica do modelo de um diodo via função W de Lambert foi anteriormente exposta como (2.21) no âmbito do Capítulo 2. Contudo, em razão do caráter sintético deste, a sua apresentação foi feita de maneira concisa. Isto é, não se detalhou o procedimento algébrico empregado em sua dedução, nem se discutiram as nuances associadas à avaliação da função W de Lambert. Posto isto, esta seção se dedica a aprofundar esses aspectos, com vistas à construção de uma compreensão crítica e operacional da abordagem, necessária à formalização do problema de extração e aos desenvolvimentos metodológicos apresentados nos capítulos subsequentes.

3.1.1 Função W de Lambert

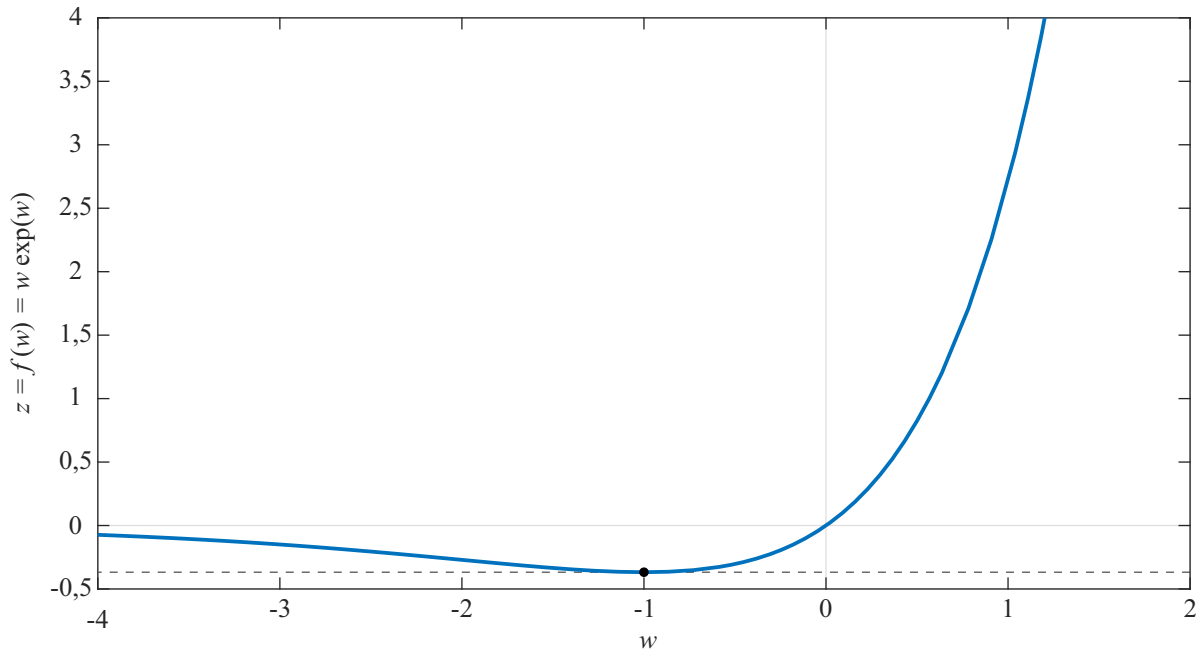
A função W de Lambert é definida como o inverso multivalorado da função f apresentada em (3.1), na qual a variável w pode assumir qualquer valor no plano complexo (\mathbb{C}) [122].

$$f(w) := w \exp(w) \quad (3.1)$$

Embora essa definição seja matematicamente correta, esta não é, por si só, muito elucidativa quanto ao significado da função nem quanto ao potencial de aplicação desta. Para tornar essas interpretações mais transparentes, ilustra-se na Figura 9 o perfil de (3.1) nas proximidades da origem no eixo real.

Como pode ser observado, a curva descritiva apresenta um ponto de mínimo em $w = -1$, no qual $f(-1) = -1/\exp(1)$. De fato, a função f é estritamente decrescente no intervalo $(-\infty, -1]$ e estritamente crescente em $[-1, \infty)$, o que caracteriza esse ponto como o mínimo global [122, 123]. Essa condição implica que f não é uma função injetora [171]. Em particular, para uma imagem z maior que $-1/\exp(1)$, a equação proveniente de (3.1), explicitada em (3.2), admite duas soluções distintas em certos trechos do domínio real.

$$w \exp(w) = z \quad (3.2)$$

Figura 9 – Comportamento da função $f(w) = w \exp(w)$ no eixo real.

Os possíveis valores de w associados univocamente à imagem z em certo intervalo do domínio no qual f é estritamente monótona podem, então, ser recuperados aplicando-se a função inversa de f em ambos os lados de (3.2). Dessa forma, obtém-se a relação apresentada em (3.3), onde f^{-1} denota a inversa de f restrita ao intervalo considerado, por simplicidade.

$$f^{-1}(z) = f^{-1}(w \exp(w)) = f^{-1}(f(w)) = w \quad (3.3)$$

Conforme estabelecido inicialmente, é exatamente essa operação de inversão que caracteriza a função W de Lambert. Assim, a definição desta função pode ser alternativamente expressa por meio de (3.4), para o intervalo considerado.

$$W(z) := w \quad (3.4)$$

Ao substituir (3.4) em (3.2), conclui-se que o valor retornado por $W(z)$ deve satisfazer a *identidade fundamental* apresentada em (3.5). Essa identidade é frequentemente tratada na literatura específica como a *equação definidora* da função W de Lambert, uma vez que esta serve de base para o desenvolvimento de uma parcela expressiva de seus atributos analíticos e numéricos [122, 123, 150, 153, 155, 172, 173].

$$W(z) \exp(W(z)) = z \quad (3.5)$$

É importante ainda destacar que (3.5) evidencia o principal interesse por trás da própria função W de Lambert: tal função permite determinar explicitamente o valor da variável w que resolve a transcendência $w \exp(w) = z$, conhecendo-se apenas o valor de z [123]. Consequentemente, qualquer equação que possa ser reescrita na forma de “variável multiplicada pela própria exponencial” passa a admitir solução fechada por meio da função W de Lambert [150].

O emprego prático dessa função especial, entretanto, demanda cautela, em virtude de sua natureza multivalorada. Em princípio, estruturas matemáticas com esse comportamento não constituem funções no sentido estrito, pois deixam de associar cada elemento do domínio a uma única imagem. Para lidar analiticamente com tal fenômeno, recorre-se usualmente ao conceito de *ramos* [123]. Em termos gerais, um ramo é um subconjunto tão amplo quanto possível do domínio no qual a função multivalorada se torna univocamente definida, contínua e infinitamente diferenciável — ou, em linguagem mais técnica, uma função de classe C^∞ [123]. A avaliação exclusiva de cada ramo corresponde, assim, à seleção consistente de uma das possíveis soluções da função original [123].

No caso da função W de Lambert, as ramificações são especialmente ricas: existem infinitos ramos no plano complexo, usualmente denotados por $W_m(z)$, onde $m \in \mathbb{Z}$ [123]. Contudo, ao restringir a análise ao eixo real — cenário suficiente para a maior parte das aplicações e, em particular, para o modelo de um diodo — persistem apenas dois desses ramos [122]. A fim de elucidar as particularidades correspondentes, ilustra-se na Figura 10 o perfil da função W de Lambert nas proximidades da origem no eixo real.

O gráfico apresentado na Figura 10 evidencia um ponto crítico em $z = -1/\exp(1)$, no qual $W(z) = -1$ e a curva apresenta uma tangente vertical. Em outras palavras, a inclinação desta função cresce sem limite à medida que z se aproxima desse valor. Esse aspecto singular delimita precisamente a fronteira entre as regiões em que (3.5) admite, ou não, soluções reais. De fato, (i) se $z < -1/\exp(1)$, nenhuma solução real existe; (ii) se $z = -1/\exp(1)$, há uma única solução real; (iii) se $-1/\exp(1) < z < 0$, existem duas soluções reais, ambas negativas; e (iv) se $z \geq 0$, existe exatamente uma solução real, necessariamente não negativa.

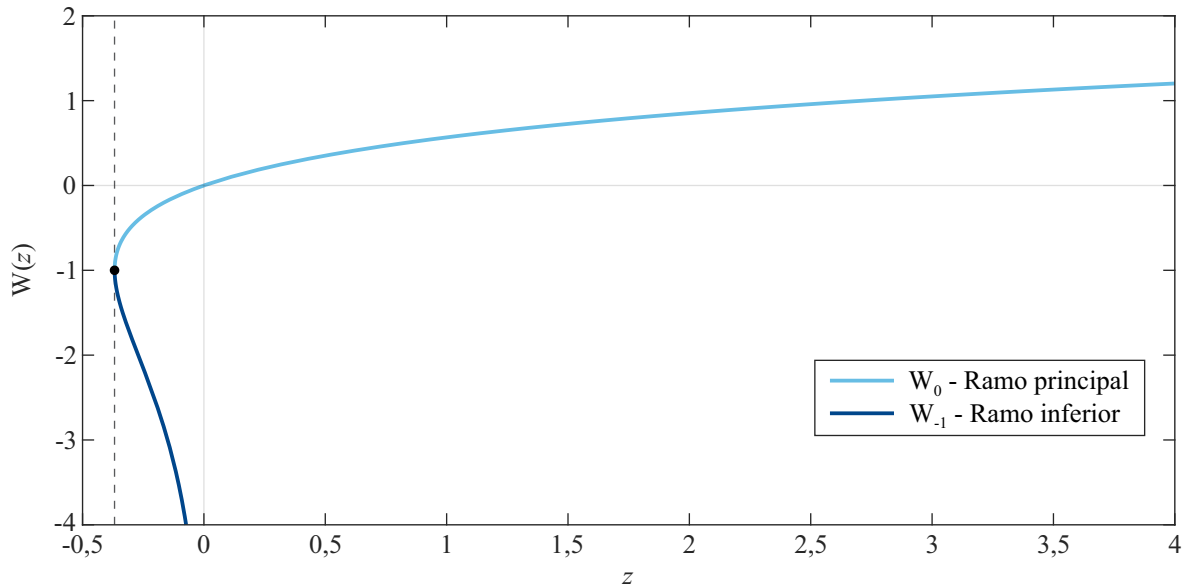
Esses comportamentos observados na curva descritiva da função W de Lambert conduzem intrinsecamente ao estabelecimento de dois ramos reais. O *ramo inferior*, denotado por W_{-1} , é restrito ao intervalo $z \in [-\exp(-1), 0)$ e satisfaz $W_{-1}(z) \leq -1$ [153]. Já o *ramo superior*, denotado por W_0 , é definido para $z \in [-\exp(-1), \infty)$ e satisfaz $W_0(z) \geq -1$ [153]. Vale mencionar que este último é mais frequentemente referido na literatura específica como o *ramo principal* [122, 123, 150, 153, 155, 172, 173].

Como cada um desses ramos é analítico no interior do respectivo domínio, as suas propriedades qualitativas podem ser examinadas por meio das derivadas associadas. Por exemplo, a derivada de primeira ordem do ramo principal é dada por (3.6). Nesse caso, os sinais de $W_0(z)$ e de z coincidem em todo o domínio real, e o denominador permanece estritamente positivo. Assim, conclui-se que $W'_0(z) > 0$ e, portanto, o ramo principal é *estritamente crescente* em todo o seu domínio [123].

$$W'_0(z) = \frac{dW_0(z)}{dz} = \frac{W_0(z)}{z(1 + W_0(z))} \quad (3.6)$$

Importa esclarecer que essa análise da derivada de primeira ordem do ramo principal envolve uma questão sensível. Como se observa na Figura 10, o ramo principal satisfaz $W_0(0) =$

Figura 10 – Comportamento da função W de Lambert no eixo real, com destaque para os seus ramos analíticos.



0, resultado que também se verifica prontamente por meio da equação definidora apresentada em (3.5). À primeira vista, essa condição poderia sugerir a existência de uma singularidade em $z = 0$, uma vez que (3.6) contém explicitamente o termo $1/z$. Contudo, esse comportamento é apenas aparente: a forma alternativa apresentada em (3.7), obtida manipulando-se (3.6) propriamente em conjunto com (3.5), indica que $W'_0(0) = 1$ [123].

$$W'_0(z) = \frac{1}{\exp(W_0(z)) (1 + W_0(z))} \quad (3.7)$$

A derivada de segunda ordem do ramo principal, por sua vez, é dada por (3.8). Como os fatores quadráticos são sempre positivos e $1 + W_0(z) > 0$ em todo o domínio, o sinal da expressão é determinado pelo termo $-(W_0(z) + 2)$, que é sempre negativo. Logo, $W''_0(z) < 0$ e o ramo principal é *estritamente côncavo* [123].

$$W''_0(z) = \frac{d^2 W_0(z)}{dz^2} = -\frac{W_0(z)^2 (W_0(z) + 2)}{z^2 (1 + W_0(z))^3} \quad (3.8)$$

Ao realizar uma análise análoga para o ramo inferior, verifica-se que $W_{-1}(z)$ é *estritamente decrescente* em todo o seu domínio [123]. Além disso, esse ramo apresenta um ponto de inflexão em $z = -2/\exp(2)$, sendo *convexo* no intervalo $[-\exp(-1), -2/\exp(2))$ e *côncavo* em $(-2/\exp(2), 0)$ [123]. Esse comportamento pode ser percebido, ainda que sutilmente, na Figura 10.

3.1.2 Equação característica explícita

A equação característica do modelo de um diodo foi devidamente apresentada na Seção 2.1 do capítulo anterior. Por conveniência, recupera-se abaixo sua expressão geral.

$$I = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V + I(N_s/N_p)R_s)}{N_s n k T_c} \right) - 1 \right] - \frac{V + I(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} \quad (2.6)$$

Ao analisar (2.6) tomando o termo I — correspondente à corrente elétrica produzida pelo gerador fotovoltaico — como única incógnita, observa-se que essa equação contém uma transcendência do tipo “variável somada a uma exponencial que, por sua vez, contém a própria variável”. Por outro lado, a função W de Lambert fornece soluções fechadas para expressões da forma “variável multiplicada pela sua própria exponencial”, conforme discutido na Subseção 3.1.1. Assim, torna-se necessário efetuar uma reformulação algébrica substancial de (2.6), de modo a conduzi-la à estrutura matemática apresentada em (2.21).

Embora as manipulações algébricas necessárias para conduzir (2.6) à forma canônica requerida pela função W de Lambert sejam raramente apresentadas de modo explícito na literatura, a compreensão do encadeamento dessas etapas pode ser pertinente para apreender com maior clareza as nuances da abordagem. Todavia, como tal desenvolvimento implica uma carga matemática considerável, opta-se por apresentá-lo integralmente no Apêndice A, a fim de não onerar ainda mais este capítulo, que já é denso por natureza. Assim, nesta subseção, registram-se apenas as expressões finais obtidas.

Posto isto, a equação característica do modelo de um diodo pode ser transcrita via função W de Lambert conforme (3.9) — expressão esta equivalente àquela apresentada em (2.21).

$$I = \frac{N_p R_{sh}(I_{ph} + I_0) - (N_p/N_s)V}{R_{sh} + R_s} - \frac{N_p n k T_c}{q R_s} W(z) \quad (3.9a)$$

$$z = \frac{q I_0 R_{sh} R_s}{n k T_c (R_{sh} + R_s)} \exp \left(\frac{q R_{sh} [N_s R_s (I_{ph} + I_0) + V]}{N_s n k T_c (R_{sh} + R_s)} \right) \quad (3.9b)$$

A título de complementariedade, é pertinente registrar que a forma compacta da equação característica do modelo de um diodo apresentada na Seção 2.1 do capítulo anterior também pode ser transcrita explicitamente via função W de Lambert. Para tanto, pode-se aplicar a (2.7) o procedimento descrito no Apêndice A ou, de forma mais direta, empregar as relações apresentadas na Tabela 1 em (3.9). Em ambos os casos, obtém-se a expressão (3.10) [42].

$$I = \frac{1}{R'_{sh} + R'_s} \left(R'_{sh} (I'_{ph} + I'_0) - V \right) - \frac{a}{R'_s} W_0(z) \quad (3.10a)$$

$$z = \frac{I'_0 R'_{sh} R'_s}{a (R'_{sh} + R'_s)} \exp \left(\frac{R'_{sh}}{a (R'_{sh} + R'_s)} \left(R'_s (I'_{ph} + I'_0) + V \right) \right) \quad (3.10b)$$

É pertinente notar que (3.9) — e, equivalentemente (3.10) — é obtida exclusivamente a partir de manipulações algébricas aplicadas à equação característica do modelo de um diodo.

Durante esse desenvolvimento, não é necessário discutir qual ramo analítico da função W de Lambert está sendo implicitamente assumido e, por conseguinte, a expressão final não inclui qualquer indicação evidente a esse respeito. A identificação do ramo analítico apropriado requer uma análise contextualizada de (3.9), a qual é apresentada na sequência.

3.1.3 Domínio fisicamente relevante

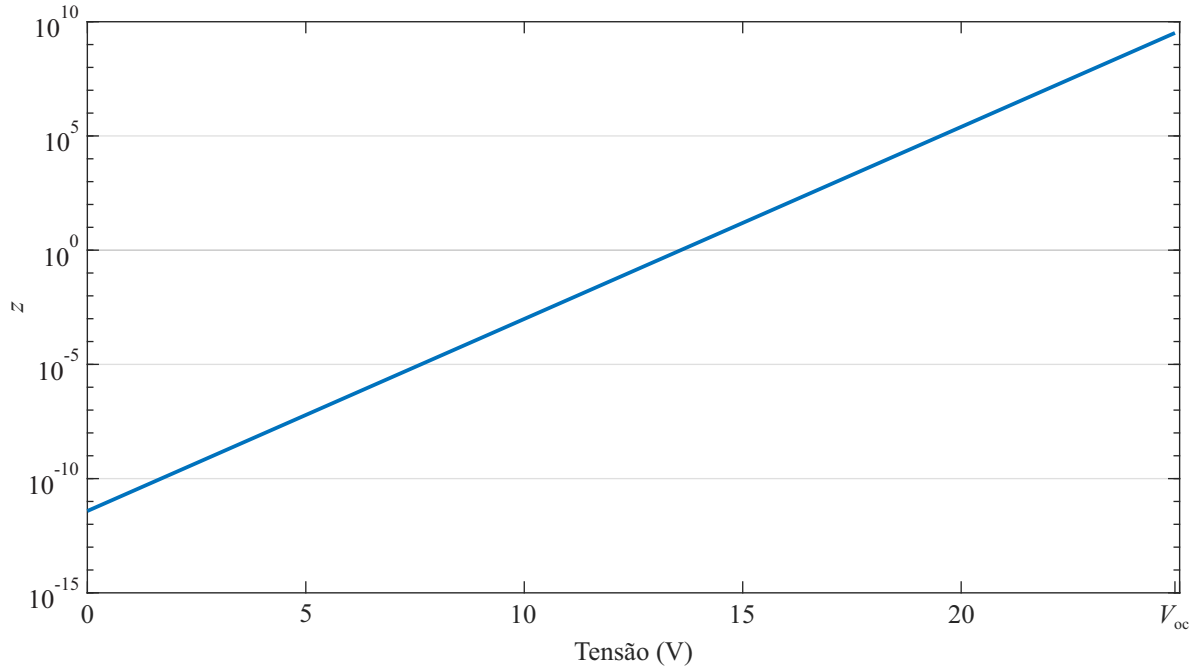
Para determinar em qual ramo analítico a função W de Lambert presente em (3.9) deve ser avaliada, é crucial recordar que o modelo de um diodo possui natureza fundamentalmente física. Conforme apresentado na Seção 2.1 do capítulo anterior, a sua equação característica foi concebida de modo a preservar os sinais esperados das grandezas envolvidas quando o gerador fotovoltaico opera em condições normais. A transcrição dessa equação via função W de Lambert não altera esse aspecto: por construção, essa reformulação apenas reorganiza algebricamente os termos de (2.6), transferindo a transcendência para a função W de Lambert, sem modificar o conteúdo físico do modelo. Além disso, é importante observar que todas as grandezas envolvidas no modelo de um diodo pertencem ao domínio real. Assim, mesmo antes de se examinar a estrutura de (3.9), já se pode afirmar que qualquer interpretação fisicamente consistente dessa equação deve recorrer a um dos dois ramos reais da função W de Lambert.

Posto isto, a escolha entre o ramo principal ou o ramo inferior pode ser adequadamente estabelecida a partir da análise direta do argumento z , definido por (3.9b) e empregado na função W de Lambert presente em (3.9a). Nesse sentido, observa-se que tal argumento é composto por um prefator não negativo multiplicado por um termo exponencial estritamente positivo. De fato, as constantes físicas q e k são estritamente positivas, assim como N_p e N_s são números naturais (\mathbb{N}). Ademais, no contexto físico do gerador fotovoltaico, T_c , I_0 , n , R_{sh} e R_s são termos, a rigor, positivos. Já I_{ph} é não negativa, sendo nula apenas na ausência de irradiância solar incidente. Cabe ainda notar que a função exponencial presente em (3.9b) permanece estritamente positiva para qualquer argumento real, inclusive quando a tensão nos terminais do gerador é negativa. Assim, conclui-se que $z \geq 0$, sendo $z = 0$ apenas em casos-limite ou idealizações nas quais o prefator se anula.

Outro atributo pertinente diz respeito à ordem de grandeza do argumento em questão. Conforme também discutido na Seção 2.1, R_{sh} tende a assumir valores da ordem de milhares de ohms, enquanto R_s se aproxima de zero. Nessas circunstâncias, a contribuição dominante para a variação do argumento z em (3.9b) é fornecida pela tensão [42]. Para valores reduzidos de V , z tende a se aproximar de zero; já na vizinhança da tensão de circuito aberto (V_{oc}), o termo exponencial presente em (3.9b) pode elevar o valor de z a ordens de grandeza extremas, assim como ilustrado na Figura 11. Portanto, o domínio fisicamente relevante do argumento z se estende propriamente por todo *semi-eixo real não negativo*.

Uma vez que $z \in [0, \infty)$ e o ramo inferior é definido apenas para argumentos no intervalo $[-1/e, 0)$, conclui-se que a função W de Lambert presente em (3.9) deve necessariamente ser

Figura 11 – Magnitude do argumento z da função W de Lambert presente na equação característica do modelo de um diodo, em relação à tensão nos terminais de um gerador fotovoltaico.



avaliada no ramo principal, conforme explicitado em (3.11).

$$I = \frac{N_p R_{sh} (I_{ph} + I_0) - (N_p / N_s) V}{R_{sh} + R_s} - \frac{N_p n k T_c}{q R_s} W_0(z) \quad (3.11a)$$

$$z = \frac{q I_0 R_{sh} R_s}{n k T_c (R_{sh} + R_s)} \exp\left(\frac{q R_{sh} [N_s R_s (I_{ph} + I_0) + V]}{N_s n k T_c (R_{sh} + R_s)}\right) \quad (3.11b)$$

Adicionalmente, é interessante notar que a adoção do ramo principal também é coerente com as propriedades físicas da curva $I-V$ de um gerador fotovoltaico. Como ilustrado na Figura 5, essa curva apresenta um perfil monotonicamente decrescente com respeito à tensão, sob condições normais de operação. A expressão dada em (3.11a), por sua vez, preserva essa particularidade. O primeiro termo do lado direito de (3.11a) é uma função afim de V , com coeficiente angular negativo. O segundo termo contém $W_0(z)$ — que, conforme demonstrado anteriormente, é estritamente crescente no semi-eixo real não negativo — multiplicada por um coeficiente negativo. Uma vez que o argumento z definido em (3.9b) cresce rapidamente com a tensão, especialmente para V na vizinhança de V_{oc} , a parcela dependente de $W_0(z)$ reforça a tendência de diminuição da corrente à medida que V aumenta. Sob parâmetros fisicamente razoáveis, ambas as contribuições atuam em consonância, assegurando a monotonicidade decrescente da curva $I-V$.

Em síntese, tanto a análise do argumento z definido em (3.9b) quanto a inspeção das propriedades estruturais de (3.9a) conduzem à mesma conclusão: no escopo do modelo de um diodo, a função W de Lambert deve ser avaliada no seu ramo principal, em particular no semi-eixo real não negativo.

3.2 ANÁLISE DO PROBLEMA DE OTIMIZAÇÃO

A adoção da abordagem formal via função W de Lambert não apenas fornece uma expressão fechada para a corrente elétrica descrita pelo modelo de um diodo, como também estabelece a estrutura analítica que será incorporada à função objetivo do problema de extração dos parâmetros. Desse modo, as propriedades diferenciais dessa função especial e as restrições intrinsecamente associadas ao domínio fisicamente relevante passam a integrar, de maneira indissociável, a própria natureza matemática do problema.

Atento a essas implicações, esta seção apresenta a formulação do problema de extração dos parâmetros efetivamente adotada nesta pesquisa e, em seguida, analisa suas propriedades sob a ótica da otimização.

3.2.1 Formulação adotada

Conforme explicitado na Subseção 2.2.1 do capítulo anterior, as variáveis de decisão do problema de extração dos parâmetros são propriamente definidas pelos cinco parâmetros do modelo de um diodo. Recordar-se, portanto, que essas variáveis são designadas de acordo com (2.10), onde “ \top ” representa o operador de transposição vetorial e $\mathbf{x} \in \mathbb{R}^5$.

$$\mathbf{x} := [I_{ph}, I_0, n, R_{sh}, R_s]^\top \quad (2.10)$$

A função objetivo, por sua vez, é tradicionalmente definida pela soma dos erros quadráticos ou pela raiz do erro quadrático médio, isto é, SSE ou RMSE. Embora ambas conduzam o processo iterativo de otimização à mesma solução ótima, a adoção da SSE oferece um ganho de eficiência, pois tanto a sua avaliação quanto o cálculo de suas derivadas são mais simples. Além disso, como revelado no levantamento bibliográfico documentado na Subseção 2.4.1, tal métrica é aquela comumente empregada quando se considera a aplicação de métodos determinísticos locais para a resolução do problema de extração dos parâmetros — o que se alinha aos objetivos específicos da pesquisa. Logo, adota-se a SSE como função objetivo na presente formulação, sendo $\mathcal{F} : \mathbb{R}^5 \mapsto \mathbb{R}$ expressa por (3.12).

$$\mathcal{F}(\mathbf{x}) := \sum_{j=1}^N [I_{m,j} - I_{c,j}(V_{m,j}, \mathbf{x})]^2 \quad (3.12)$$

A respeito de seus termos, ressalta-se que $N \in \mathbb{N}$ representa o número total de pares de medições $(V_{m,j}, I_{m,j})$ que descrevem a curva I - V disponível do gerador fotovoltaico em estudo e, em todo caso, satisfaz $N > 5$. Já $I_{c,j}$ representa, no caso, o valor de corrente retornado pela equação característica explícita do modelo de um diodo. Em consonância com as discussões realizadas na seção anterior, esta é então dada por (3.13), em que $W_0(z)$ representa o ramo principal da função W de Lambert.

$$I_{c,j}(V_{m,j}, \mathbf{x}) = \frac{N_p R_{sh} (I_{ph} + I_0) - (N_p / N_s) V_{m,j}}{R_{sh} + R_s} - \frac{N_p n k T_c}{q R_s} W_0(z_{c,j}) \quad (3.13a)$$

$$z_{c,j} = \frac{qI_0R_{sh}R_s}{nkT_c(R_{sh} + R_s)} \exp\left(\frac{qR_{sh}[N_sR_s(I_{ph} + I_0) + V_{m,j}]}{N_snkT_c(R_{sh} + R_s)}\right) \quad (3.13b)$$

Para finalizar a presente formulação, é necessário deliberar sobre a incorporação das restrições de canalização. O levantamento bibliográfico apresentado na Subseção 2.4.1 evidenciou que tais restrições têm sido, até o momento, sistematicamente desconsideradas quando se trata da aplicação de métodos determinísticos locais. Todavia, à luz do conteúdo exposto no capítulo anterior e até o momento neste, depreende-se que tal prática não é recomendada.

Em primeiro lugar, porque somente na presença de limites inferiores e superiores explícitos é possível assegurar que a solução final retornada por um método de otimização qualquer permaneça fisicamente coerente, bem como consistente com as respectivas faixas de valores esperados para os parâmetros do modelo de um diodo (ver discussão na Subseção 2.2.1).

Em segundo lugar, porque apenas com a imposição adequada desses limites é possível garantir a validade da abordagem formal do modelo de um diodo via função W de Lambert no âmbito do próprio problema. Como pode ser inferido das discussões realizadas na seção anterior, caso se permita que os parâmetros do modelo de um diodo assumam valores negativos, abre-se a possibilidade de que $z_{c,j}$ também se torne negativo e, em casos mais críticos, satisfaça $z_{c,j} < -\exp(-1)$. Nessa situação, $W_0(z_{c,j})$ deixa de ser real e passa a assumir valores no plano complexo. Dessa forma, as restrições de canalização tornam-se necessárias para manter os valores calculados de $I_{c,j}$ bem definidos em \mathbb{R} ao longo do processo iterativo de resolução do problema.

Posto isto, adotam-se as restrições de canalização na presente formulação, com o intuito de preservar a natureza fundamental do modelo de um diodo, bem como do problema de extração dos parâmetros. Para fins de padronização, os limites inferiores $\mathbf{lb} \in \mathbb{R}^5$ e superiores $\mathbf{ub} \in \mathbb{R}^5$ são introduzidos na forma usual de desigualdades do tipo “maior ou igual a”, conforme expresso em (3.14), onde $\mathbf{0}$ representa um vetor nulo (isto é, com todos os elementos iguais a zero), de dimensões compatíveis. Com isso, torna-se possível reunir as condições explicitadas no capítulo anterior em (2.13) e (2.14), em um único conjunto de restrições de desigualdade, dado pela função vetorial $\mathbf{c} : \mathbb{R}^5 \mapsto \mathbb{R}^{10}$.

$$\mathbf{c}(\mathbf{x}) := \begin{bmatrix} \mathbf{x} - \mathbf{lb} \\ \mathbf{ub} - \mathbf{x} \end{bmatrix} \geq \mathbf{0} \quad (3.14)$$

Em síntese, a formulação adotada para o problema de extração dos parâmetros é enunciada em (3.15), onde $\mathcal{F}(\mathbf{x})$ é definida por (3.12), com $I_{c,j}$ dado por (3.13), enquanto $\mathbf{c}(\mathbf{x})$ é definida por (3.14).

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^5}{\text{minimizar}} && \mathcal{F}(\mathbf{x}) \\ & \text{sujeito a:} && \mathbf{c}(\mathbf{x}) \geq \mathbf{0} \end{aligned} \quad (3.15)$$

À luz das avaliações conduzidas na Subseção 2.2.2 e das propriedades da função W de Lambert apresentadas na Subseção 3.1.1, observa-se que (3.15) enquadra o problema de extração

dos parâmetros como um problema de otimização multidimensional restrito, de natureza não linear e, em geral, não convexa. Ademais, sob condições regulares, a função objetivo de (3.15) é suave, isto é, de classe C^∞ , no interior do espaço de soluções viáveis — em particular, a corrente calculada $I_{c,j}$ é continuamente diferenciável sempre que $z_{c,j} > -1/\exp(1)$.

3.2.2 Função Lagrangiana

Em termos gerais, a *função Lagrangiana* associa a um problema de otimização restrito uma formulação irrestrita, ao incorporar as restrições diretamente na função objetivo por meio dos *multiplicadores de Lagrange* [116]. Embora tal construção não elimine a necessidade de impor a viabilidade, esta fornece uma ferramenta útil para a análise do problema, pois permite tratá-lo com base em técnicas clássicas de cálculo diferencial, mantendo consistência com as soluções ótimas [113].

No caso do problema de extração dos parâmetros, a função Lagrangiana é construída designando-se multiplicadores às restrições de canalização impostas às variáveis de decisão [116]. Para tanto, definem-se multiplicadores distintos para as restrições relativas aos limites inferiores e superiores, conforme (3.16) e (3.17). Note-se que $\boldsymbol{\lambda}_l \in \mathbb{R}^5$ está vinculado às restrições de limite inferior, isto é, $\mathbf{x} - \mathbf{lb} \geq \mathbf{0}$. De maneira análoga, $\boldsymbol{\lambda}_u \in \mathbb{R}^5$ está vinculado às restrições de limite superior, isto é, $\mathbf{ub} - \mathbf{x} \geq \mathbf{0}$.

$$\boldsymbol{\lambda}_l := [\lambda_{l,I_{ph}}, \lambda_{l,I_0}, \lambda_{l,n}, \lambda_{l,R_{sh}}, \lambda_{l,R_s}]^\top \quad (3.16)$$

$$\boldsymbol{\lambda}_u := [\lambda_{u,I_{ph}}, \lambda_{u,I_0}, \lambda_{u,n}, \lambda_{u,R_{sh}}, \lambda_{u,R_s}]^\top \quad (3.17)$$

Uma vez que as restrições de canalização do problema de extração dos parâmetros foram reunidas em (3.14), torna-se conveniente congregiar também os seus respectivos multiplicadores em um único vetor. Assim, define-se $\boldsymbol{\lambda} \in \mathbb{R}^{10}$ como a concatenação de $\boldsymbol{\lambda}_l$ e $\boldsymbol{\lambda}_u$, conforme (3.18).

$$\boldsymbol{\lambda} := \begin{bmatrix} \boldsymbol{\lambda}_l \\ \boldsymbol{\lambda}_u \end{bmatrix} \quad (3.18)$$

Dessa forma, a função Lagrangiana $\mathcal{L} : \mathbb{R}^5 \times \mathbb{R}^{10} \mapsto \mathbb{R}$ associada ao problema de extração dos parâmetros, formulado em (3.15), é formalmente definida por (3.19) — veja [113] para mais detalhes.

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) := \mathcal{F}(\mathbf{x}) - \boldsymbol{\lambda}^\top \mathbf{c}(\mathbf{x}) \quad (3.19)$$

Cabe mencionar que, no contexto da função Lagrangiana, as variáveis de decisão \mathbf{x} são usualmente denominadas *variáveis primais*, ao passo que os multiplicadores $\boldsymbol{\lambda}$ são denominados *variáveis duais* [113].

3.2.3 Condições de otimalidade

O fato de a formulação adotada para o problema de extração dos parâmetros ser regularmente suave abre caminho para uma série de análises avançadas [113, 116]. Em particular,

o comportamento da função objetivo nas vizinhanças de um ponto do espaço de soluções delimitado pelas restrições pode ser aproximado por meio de “modelos locais” obtidos a partir de expansões em série de Taylor [116]. Como consequência, podem-se estabelecer condições para avaliar se o ponto ora definido por uma candidata à solução constitui, ao menos, um mínimo local [113]. Dentre estas, destacam-se as *condições necessárias de primeira ordem*, também conhecidas como *condições de Karush–Kuhn–Tucker (KKT)* [113, 116].

As condições de KKT para o problema de extração dos parâmetros, formulado conforme (3.15), são expressas de acordo com (3.20), em que “ \odot ” representa o operador de multiplicação elemento a elemento entre os termos correspondentes.

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{g}(\mathbf{x}) - \mathbf{J}(\mathbf{x})^\top \boldsymbol{\lambda} = \mathbf{0} \quad (3.20a)$$

$$\mathbf{c}(\mathbf{x}) \geq \mathbf{0} \quad (3.20b)$$

$$\boldsymbol{\lambda} \odot \mathbf{c}(\mathbf{x}) = \mathbf{0} \quad (3.20c)$$

$$\boldsymbol{\lambda} \geq \mathbf{0} \quad (3.20d)$$

É pertinente esclarecer que (3.20a) corresponde às derivadas parciais de primeira ordem da função Lagrangiana dada por (3.19) com respeito às variáveis primais. Nesse contexto, $\mathbf{g} : \mathbb{R}^5 \mapsto \mathbb{R}^5$ denota o *vetor gradiente da função objetivo*, ou seja, o vetor com as derivadas parciais de primeira ordem de (3.12) com respeito às variáveis primais. Logo, este é determinado de acordo com (3.21).

$$\mathbf{g}(\mathbf{x}) := \nabla\mathcal{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial\mathcal{F}(\mathbf{x})}{\partial I_{ph}} \\ \frac{\partial\mathcal{F}(\mathbf{x})}{\partial I_\theta} \\ \frac{\partial\mathcal{F}(\mathbf{x})}{\partial n} \\ \frac{\partial\mathcal{F}(\mathbf{x})}{\partial R_{sh}} \\ \frac{\partial\mathcal{F}(\mathbf{x})}{\partial R_s} \end{bmatrix} \quad (3.21)$$

Por sua vez, $\mathbf{J} : \mathbb{R}^5 \mapsto \mathbb{R}^{10 \times 5}$ denota a *matriz Jacobiana das restrições*. A rigor, essa matriz é obtida ao se reunirem os vetores gradiente de cada restrição, isto é, $\mathbf{J}(\mathbf{x}) := \nabla\mathbf{c}(\mathbf{x})$ [116]. Entretanto, como (3.14) é uma função afim de \mathbf{x} , a matriz Jacobiana é constante e assume a forma apresentada em (3.22), em que $\mathbf{I} \in \mathbb{R}^{5 \times 5}$ denota a matriz identidade.

$$\mathbf{J}(\mathbf{x}) \equiv [\mathbf{I}, -\mathbf{I}]^\top = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \quad (3.22)$$

À luz da Teoria do Cálculo Diferencial, (3.20a) estabelece as *condições de estacionariedade* para a candidata à solução [116]. Em outras palavras, para que o ponto definido pela candidata à solução possa ser considerado um mínimo, não deve existir, em sua vizinhança, nenhuma direção de descida.

Adicionalmente, (3.20b) estabelece as *condições de viabilidade primal*. Isto é, para que o ponto definido pela candidata à solução possa ser considerado válido, este deve pertencer ao espaço de soluções viáveis do problema de extração dos parâmetros.

Por fim, (3.20c) estabelece as chamadas *condições de complementaridade*, enquanto (3.20d) estabelece as *condições de não negatividade das variáveis duais*. Na prática, essas condições conectam cada restrição de desigualdade ao seu respectivo multiplicador de Lagrange. Mais propriamente, para cada componente c_j de \mathbf{c} e o multiplicador associado λ_j de $\boldsymbol{\lambda}$, tem-se que, se a restrição estiver *inativa*, isto é, $c_j(\mathbf{x}) > 0$, então necessariamente $\lambda_j = 0$. Em contraste, se a restrição estiver *ativa*, isto é, $c_j(\mathbf{x}) = 0$, então $\lambda_j \geq 0$, sendo comum $\lambda_j > 0$ quando essa restrição efetivamente influencia a solução [113, 116]. Dessa forma, apenas as restrições ativas impactam em (3.20a) e influenciam a solução ótima, de modo que os seus próprios multiplicadores representam uma medida da sensibilidade da função objetivo a variações no respectivo valor-limite imposto [113, 116]. Cumpre notar, entretanto, que as condições de complementaridade introduzem componentes de natureza combinatória no problema, elevando a sua complexidade computacional [113, 116].

Assim, as condições de KKT expressas em (3.20) estabelecem que a candidata à solução descrita em \mathbf{x} e, de forma complementar, em $\boldsymbol{\lambda}$ deve necessariamente constituir um ponto estacionário da função Lagrangiana no espaço de soluções viáveis do problema de extração dos parâmetros, para que possa ser considerada, ao menos, localmente ótima.

Por completude, ressalta-se que a validade das condições de KKT pressupõe que o conjunto de restrições ativas do problema atenda à *qualificação de independência linear* (“*linear independence constraint qualification*”, LICQ) [113, 116]. Em termos simples, tal condição requer que os gradientes das restrições ativas no ponto viável sob análise sejam linearmente independentes [113]. No entanto, no caso do problema de extração dos parâmetros, a LICQ se verifica de forma natural. Como o problema compreende apenas restrições de canalização, os vetores gradiente correspondentes são constantes e coincidem (a menos de sinal) com os vetores canônicos associados a cada variável de decisão, conforme refletido na estrutura de $\mathbf{J}(\mathbf{x})$ em (3.22). Ademais, para limites estritos ($\mathbf{lb} < \mathbf{ub}$), se uma restrição associada à j -ésima variável x_j estiver ativa, então ocorre necessariamente apenas uma das possibilidades: ou $x_j - lb_j = 0$ (isto é, $x_j = lb_j$) ou $ub_j - x_j = 0$ (isto é, $x_j = ub_j$), mas nunca ambas simultaneamente. Logo, o conjunto de gradientes das restrições ativas é composto por vetores canônicos distintos, os quais são linearmente independentes.

Vale ainda ponderar que as condições de KKT constituem condições de otimalidade necessárias, porém não suficientes, para assegurar que um ponto candidato corresponda a um

mínimo local [113]. Com efeito, um ponto estacionário no conjunto de soluções viáveis do problema pode representar um mínimo, um máximo ou mesmo um ponto de sela [116]. A discriminação entre esses casos requer a avaliação de informações de ordem superior associadas à função Lagrangiana, o que conduz ao estabelecimento de condições de otimalidade adicionais [113, 116]. Maiores detalhes a respeito podem ser encontrados em [113] e [116].

3.2.4 Avaliação de convexidade

Assim como introduzido na subseção anterior, a função Lagrangiana constitui um instrumento importante para a avaliação do comportamento local do problema. Conseqüentemente, esta pode ser explorada para investigar a existência de regiões convexas e não convexas no espaço de soluções viáveis [113]. Nesse sentido, a *matriz Hessiana da função Lagrangiana* assume papel central, por sintetizar informações de curvatura associadas à vizinhança de um ponto sob análise e, assim, subsidiar uma apreciação mais precisa da geometria local do problema [116, 117].

Mais precisamente, a matriz Hessiana da função Lagrangiana é tipicamente caracterizada como uma matriz simétrica composta pelas derivadas parciais de segunda ordem dessa função com respeito às variáveis primais, isto é, $\mathbf{H}_{\mathcal{L}}(\mathbf{x}, \boldsymbol{\lambda}) := \nabla_{\mathbf{xx}}^2 \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ [113, 116]. Tendo isso em vista, o comportamento na vizinhança de um ponto viável $(\mathbf{x}, \boldsymbol{\lambda})$ pode ser considerado *localmente convexo* se essa matriz assumir uma forma *semidefinida positiva*: $\mathbf{H}_{\mathcal{L}} \succeq \mathbf{0}$ [113, 117]. Conforme preveem os conceitos de Álgebra Linear, esse tipo de matriz possui todos os seus autovalores positivos ou nulos, sendo todos os elementos da sua diagonal principal não negativos [117]. Por sua vez, a Teoria do Cálculo Diferencial fornece que os autovalores de uma matriz Hessiana quantificam os eixos principais de curvatura da vizinhança do ponto sob análise [116]. Assim, se existir ao menos um autovalor negativo, então há ao menos uma direção de curvatura negativa, o que indica comportamento não convexo naquela região do espaço de soluções viáveis do problema [113].

No caso do problema de extração dos parâmetros, a matriz Hessiana da função Lagrangiana dada em (3.19) é formalmente definida por (3.23). Nesse contexto, $\mathbf{H} : \mathbb{R}^5 \mapsto \mathbb{R}^{5 \times 5}$ denota a *matriz Hessiana da função objetivo*, ou seja, a matriz simétrica composta pelas derivadas parciais de segunda ordem de (3.12) com respeito às variáveis primais. Logo, esta é determinada de acordo com (3.24).

$$\mathbf{H}_{\mathcal{L}}(\mathbf{x}, \boldsymbol{\lambda}) := \nabla_{\mathbf{xx}}^2 \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{H}(\mathbf{x}) - \sum_{j=1}^{10} \lambda_j \nabla_{\mathbf{xx}}^2 c_j(\mathbf{x}) \quad (3.23)$$

$$\mathbf{H}(\mathbf{x}) := \nabla_{\mathbf{xx}}^2 \mathcal{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_{ph}^2} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_{ph} \partial I_0} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_{ph} \partial n} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_{ph} \partial R_{sh}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_{ph} \partial R_s} \\ \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_0 \partial I_{ph}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_0^2} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_0 \partial n} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_0 \partial R_{sh}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial I_0 \partial R_s} \\ \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial n \partial I_{ph}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial n \partial I_0} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial n^2} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial n \partial R_{sh}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial n \partial R_s} \\ \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_{sh} \partial I_{ph}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_{sh} \partial I_0} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_{sh} \partial n} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_{sh}^2} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_{sh} \partial R_s} \\ \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_s \partial I_{ph}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_s \partial I_0} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_s \partial n} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_s \partial R_{sh}} & \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial R_s^2} \end{bmatrix} \quad (3.24)$$

Não obstante, como as restrições em (3.14) são funções afins de \mathbf{x} , todas as derivadas parciais de segunda ordem dessas restrições com respeito às variáveis primais são nulas. Mais especificamente, $\nabla_{\mathbf{xx}}^2 c_j(\mathbf{x}) = \mathbf{0}$ para todo $j \in \{1, \dots, 10\}$. Em consequência, a matriz Hessiana da função Lagrangiana independe dos multiplicadores de Lagrange e coincide com a matriz Hessiana da função objetivo, tal como indicado em (3.25).

$$\mathbf{H}_{\mathcal{L}}(\mathbf{x}, \boldsymbol{\lambda}) \equiv \mathbf{H}(\mathbf{x}) \quad (3.25)$$

Assim, a avaliação da convexidade na vizinhança de um ponto pertencente ao espaço de soluções viáveis do problema de extração dos parâmetros pode ser prontamente realizada a partir da matriz Hessiana da função objetivo. Note-se, inclusive, que essa matriz é de ordem cinco e, portanto, possui apenas um conjunto de cinco autovalores, o que torna o cálculo destes uma tarefa pouco onerosa.

A título de exemplo, considere o tradicional estudo de caso da célula fotovoltaica RTC France, descrito na Subseção 2.5. Ao introduzir uma perturbação de 0,01 A no parâmetro I_{ph} da sua solução ótima, chega-se ao ponto viável \mathbf{x}^t definido por $I_{ph} = 0,770788$ A, $I_0 = 0,310685$ μ A, $n = 1,477265$, $R_{sh} = 53,889791$ Ω e $R_s = 0,0365467$ Ω . Ao montar a matriz Hessiana da função objetivo nesse ponto, encontra-se (3.26).

$$\mathbf{H}(\mathbf{x}^t) = \begin{bmatrix} 4,135 \times 10^1 & -1,450 \times 10^7 & 4,310 \times 10^1 & 4,375 \times 10^{-3} & -2,612 \times 10^1 \\ -1,450 \times 10^7 & 2,562 \times 10^{13} & -7,703 \times 10^7 & -2,855 \times 10^3 & 1,269 \times 10^7 \\ 4,310 \times 10^1 & -7,703 \times 10^7 & 2,355 \times 10^2 & 8,520 \times 10^{-3} & -3,441 \times 10^1 \\ 4,375 \times 10^{-3} & -2,855 \times 10^3 & 8,520 \times 10^{-3} & -9,245 \times 10^{-7} & -4,618 \times 10^{-3} \\ -2,612 \times 10^1 & 1,269 \times 10^7 & -3,441 \times 10^1 & -4,618 \times 10^{-3} & 9,363 \times 10^1 \end{bmatrix} \quad (3.26)$$

Como se pode notar, a diagonal principal dessa matriz apresenta um elemento negativo, dado por $-9,245 \times 10^{-7}$, o que já implica que $\mathbf{H}(\mathbf{x}^t) \not\equiv \mathbf{0}$. De fato, ao se calcularem os autovalores dessa matriz com o auxílio da função `eig` do MATLAB (versão 2016b), verifica-se que estes são dados por: $-6,535 \times 10^{-4}$; $3,747$; $2,720 \times 10^1$; $9,346 \times 10^1$; e $2,562 \times 10^{13}$. Portanto, esses resultados indicam que o ponto \mathbf{x}^t em análise pertence a uma região do espaço de soluções viáveis na qual coexistem direções de curvatura positiva e negativa, substanciando a natureza não convexa do problema em questão.

3.3 QUESTÕES NUMÉRICAS

Além de apresentar uma natureza intrincada, o problema de extração dos parâmetros também está sujeito a questões numéricas que podem tornar sua resolução particularmente difícil. A seguir, discutem-se duas das principais.

3.3.1 Escala das variáveis

Conforme explicitado na Subseção 2.1.3 do capítulo anterior, I_0 tipicamente assume valores da ordem de 10^{-6} ou inferiores, ao passo que R_{sh} pode atingir valores da ordem de 10^5 . Essa discrepância, em conjunto com as diferentes posições que tais parâmetros ocupam na equação característica do modelo de um diodo, ora dada por (3.13), implica que pequenas variações em seus respectivos valores produzam impactos expressivamente distintos no valor da função objetivo. Consequentemente, o problema de extração dos parâmetros se caracteriza, em essência, como um problema de otimização *mal escalonado*.

É relevante comentar que tal característica pode prejudicar de forma significativa o desempenho de métodos de otimização de diferentes classes, pois torna mais difícil a obtenção de direções de busca para a solução ótima que sejam efetivas no espaço de soluções viáveis do problema [113, 116]. Como resultado, o processo iterativo tende a demandar maior esforço computacional e pode se tornar menos robusto, afetando a eficiência e, em casos mais extremos, até mesmo a sua convergência [85]. Portanto, trata-se de um aspecto que deve ser cuidadosamente considerado durante a seleção e a implementação de um método de otimização para a resolução do problema de extração dos parâmetros.

A literatura específica aponta que o mal escalonamento de um problema de otimização pode ser mitigado ainda na fase de modelagem, mediante a aplicação de técnicas de escalonamento das variáveis de decisão [113, 116, 117]. Nesse sentido, é pertinente exaltar os tratamentos normalizados aplicados às variáveis de decisão do problema de extração dos parâmetros em trabalhos correlatos, a saber, [64], [65] e [66]. Não obstante, faz-se necessário chamar a atenção para a eventual dificuldade de se estabelecer valores de referência adequados para esses tratamentos em certos casos. Em particular, os parâmetros I_0 e R_s podem variar ao longo de uma ampla faixa de ordens de grandeza [23, 43, 54, 65, 131, 139]. Assim, a escolha de um valor de referência para a normalização que seja satisfatório para tais parâmetros ao longo de todo o processo iterativo de otimização, no qual estes podem variar entre diferentes magnitudes, pode se mostrar desafiadora.

De maneira alternativa, ou mesmo complementar, existem métodos de otimização que dispõem de recursos matemáticos e computacionais voltados a lidar com o mal escalonamento do problema [113]. A título de exemplo, é possível citar alternativas que foram propriamente desenvolvidas para a resolução do problema de extração dos parâmetros com atenção a essa questão, como o método determinístico local NMMLP e o método metaheurístico modificado ELBA [63, 85]. Assim, a escolha de métodos de otimização dotados dessas características tende a ser mais confortável e promissora para a resolução do problema.

3.3.2 Condicionamento

Além das questões associadas à escala das variáveis de decisão, existe outra dificuldade numérica que pode complicar a resolução do problema de extração dos parâmetros: o condicionamento das matrizes Hessianas neste envolvidas.

Diversos métodos de otimização, em especial os determinísticos locais baseados em gradiente, determinam direções de busca para a solução ótima por meio da resolução de sistemas lineares cujas matrizes de coeficientes são fundamentalmente estabelecidas pela matriz Hessiana da função objetivo ou pela matriz Hessiana da função Lagrangiana [113, 116]. Nessa circunstância, o bom condicionamento dessas matrizes torna-se determinante para a estabilidade do processo iterativo, uma vez que influencia diretamente a confiabilidade das operações algébricas subjacentes e, por conseguinte, a qualidade das novas candidatas à solução geradas a cada iteração [113].

Para fins de clareza, é relevante enfatizar que o condicionamento de matrizes simétricas, tais como as matrizes Hessianas, é usualmente avaliado por meio do *número de condicionamento para a inversão* (κ). Esse número é caracterizado por $\kappa \in [1, \infty)$ e pode ser calculado, de forma geral, conforme (3.27), em que \mathbf{A} representa a matriz simétrica em questão e $\|\cdot\|$ denota uma norma matricial qualquer [113, 149, 174, 175]. De modo mais prático, κ também pode ser obtido por meio de rotinas dedicadas, como a função `cond` disponível no MATLAB [174].

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \quad (3.27)$$

Os conceitos por trás do número de condicionamento são discutidos em detalhe nos trabalhos apresentados em [119], [149], [175] e [176]. Por ora, cabe mencionar que tal métrica atua como um estimador dos impactos produzidos por erros de arredondamento que surgem durante o processo de inversão de uma matriz simétrica [119, 149, 175, 176].

A título de exemplo, considere que o número de condicionamento de uma matriz simétrica qualquer seja $\kappa(\mathbf{A}) = 1 \times 10^4$. Na prática, é possível esperar uma perda de precisão da ordem de $\log_{10} \kappa(\mathbf{A})$, isto é, de 4 dígitos significativos, ao se inverter essa matriz. O grau de severidade dessa perda de precisão depende, por sua vez, do padrão adotado para a realização dos processos numéricos. No caso do padrão do *Institute of Electrical and Electronics Engineers* (IEEE) de ponto flutuante em precisão dupla — comum em ambientes de programação técnico-científica como o MATLAB —, a *precisão de máquina*¹ é $\varepsilon_{\text{maq}} \approx 2,2 \times 10^{-16}$, de modo que cada número é representado por cerca de 16 dígitos significativos [119, 177]. Logo, se a matriz em questão fosse invertida segundo esse padrão, os valores atribuídos à sua inversa apresentariam apenas os seus 12 primeiros dígitos significativos com acurácia. Portanto, quanto mais próximo de 1 (um) for o número de condicionamento de uma matriz simétrica, mais bem condicionada esta

¹ O termo “precisão de máquina” pode ser entendido como o limite superior do erro relativo decorrente de arredondamentos na aritmética em ponto flutuante [119].

pode ser considerada, dado que maior tende a ser a confiabilidade dos resultados obtidos nas operações numéricas realizadas envolvendo a mesma [117–119, 149].

No que concerne às matrizes Hessianas da função objetivo e da função Lagrangiana associada a um problema de otimização, a obra apresentada em [118] indica que um número de condicionamento da ordem de até 10^5 pode ser considerado moderado, no âmbito do padrão IEEE de ponto flutuante em precisão dupla. Por outro lado, um número de condicionamento da ordem de até 10^9 pode ser considerado elevado, e da ordem de até 10^{14} , extremamente elevado [118]. Caso o número de condicionamento ultrapasse de forma expressiva a ordem de 10^{14} , a inversão da matriz Hessiana pode ser considerada inviável, uma vez que o seu resultado não seria fidedigno [118].

No caso do problema de extração dos parâmetros, observa-se que há uma escassez de informações na literatura que tratem explicitamente do condicionamento das matrizes Hessianas envolvidas e de seus impactos sobre a resolução numérica do problema. Uma hipótese plausível para essa lacuna reside na popularidade e na forma de aplicação dos métodos de otimização empregados em trabalhos correlatos. Conforme revelado pelo levantamento bibliográfico documentado na Seção 2.4, a maioria desses trabalhos privilegia métodos metaheurísticos e conjugados, os quais, de forma fundamental, independem da matriz Hessiana da função objetivo e dispensam a construção da função Lagrangiana associada ao problema. Por outro lado, os poucos trabalhos que envolvem métodos determinísticos os utilizam apenas por meio de *solvers* padrão. Por conseguinte, seus autores não avaliam nem reportam explicitamente o condicionamento dessas matrizes, uma vez que tal informação permanece encapsulada nessas rotinas.

Apesar disso, a experiência acumulada ao longo desta pesquisa indica que o condicionamento das matrizes Hessianas associadas à função objetivo do problema de extração dos parâmetros e, por consequência, à função Lagrangiana associada tende a ser crítico. Uma evidência disso pode ser prontamente obtida a partir do exemplo numérico conduzido na seção anterior, com base no tradicional estudo de caso da célula fotovoltaica RTC France. Ao se calcular o número de condicionamento da matriz Hessiana da função objetivo descrita em (3.26) com o auxílio da função `cond` do MATLAB (versão 2016b), obtém-se $\kappa(\mathbf{H}(\mathbf{x}^t)) = 1,700 \times 10^{19}$ — um indicativo inequívoco de mal condicionamento severo.

É crucial notar que esse *mal condicionamento* impõe um desafio de grande magnitude à aplicação, sobretudo, de métodos determinísticos locais baseados em gradientes que exploram informação de segunda ordem. Reforça-se que valores elevados do número de condicionamento tendem a amplificar erros de arredondamento e a reduzir a quantidade de dígitos significativos efetivamente preservados nas operações numéricas realizadas a cada iteração [119, 149, 175, 176]. Como consequência, as candidatas à solução geradas no âmbito desses métodos podem se tornar progressivamente menos confiáveis, comprometendo a robustez de seus processos iterativos e podendo culminar em degradação acentuada de seus desempenhos ou mesmo em falhas de convergência, na ausência de salvaguardas numéricas adequadas [113].

3.4 CONCLUSÕES PARCIAIS

Este capítulo teve por objetivo estabelecer, de forma rigorosa, as bases matemáticas que sustentam a formulação adotada para o problema de extração dos parâmetros no âmbito desta pesquisa. Nesse sentido, foram discutidas as propriedades analíticas da função W de Lambert e delimitado o seu domínio fisicamente relevante na abordagem formal do modelo de um diodo. Em seguida, formalizou-se o problema de otimização correspondente e analisaram-se sua natureza estrutural, condições de otimalidade e aspectos de convexidade. Por fim, examinaram-se questões numéricas associadas à sua resolução, com destaque para a escala das variáveis e para o condicionamento das matrizes que permeiam métodos determinísticos locais.

À luz das análises conduzidas, ressaltam-se os seguintes pontos:

- No escopo do modelo de um diodo, a avaliação da função W de Lambert recai sobre o seu ramo principal, sendo o domínio fisicamente relevante restrito ao semi-eixo real não negativo. Assim, o emprego da forma explícita da equação característica dada em (3.13a) implica restrições ao argumento definido em (3.13b) e, em última instância, aos próprios parâmetros do modelo. Isso porque, caso tais parâmetros assumam valores irregulares e conduzam a $z_{c,j} < -\exp(-1)$, $W_0(z_{c,j})$ presente em (3.13a) passa a assumir valores no plano complexo e, por conseguinte, a corrente calculada deixa de ser real, inviabilizando a sua interpretação física.
- As restrições de canalização introduzidas em (3.14) desempenham, portanto, um papel operacional indispensável na formulação (3.15): além de delimitar um conjunto de parâmetros fisicamente coerente, estas atuam como salvaguardas para manter o argumento $z_{c,j}$ no domínio real admissível do ramo principal, preservando a avaliação de $W_0(z_{c,j})$ e garantindo que $I_{c,j}$ permaneça bem definida em \mathbb{R} ao longo do processo iterativo.
- A formulação adotada para o problema de extração dos parâmetros, enunciada em (3.15), incorpora essas considerações e caracteriza-se como um problema de otimização multidimensional restrito, de natureza não linear, (regularmente) suave e, em geral, não convexo.
- As distintas ordens de grandeza assumidas pelos parâmetros do modelo de um diodo, em conjunto com as suas disposições na equação característica explícita, tornam o problema intrinsecamente mal escalonado, afetando a sensibilidade da função objetivo e a eficiência das direções de busca computadas por métodos de otimização.
- O condicionamento das matrizes Hessianas associadas à função objetivo e à função Lagrangiana do problema de extração dos parâmetros pode assumir valores criticamente elevados, caracterizando cenários de mal condicionamento severo que comprometem a estabilidade numérica de métodos de otimização baseados em derivadas de segunda ordem.

Esses pontos evidenciam que a resolução efetiva do problema de extração dos parâmetros requer alternativas metodológicas capazes de lidar simultaneamente com restrições estruturais do domínio, não convexidade, mal escalonamento e mal condicionamento numérico. Nesse contexto, os capítulos seguintes dedicam-se à proposição de métodos específicos: primeiramente, um método voltado ao cálculo eficaz da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo e, posteriormente, um método de otimização propriamente concebido para enfrentar as particularidades do problema aqui evidenciadas.

4 MÉTODO PROPOSTO PARA O CÁLCULO DA FUNÇÃO W DE LAMBERT

Este capítulo documenta o método proposto para o cálculo da função W de Lambert no escopo da equação característica explícita do modelo de um diodo. De maneira fundamental, a Seção 4.1 apresenta as principais alternativas disponíveis na literatura específica para esse propósito e discute seus potenciais e limitações. Já a Seção 4.2 descreve o método oportunamente desenvolvido para viabilizar o cálculo da função W de Lambert no domínio fisicamente relevante de maneira eficaz e eficiente. A Seção 4.3, por sua vez, reporta os testes realizados para validação do método proposto e sintetiza os resultados obtidos frente às alternativas de estado da arte. Por fim, a Seção 4.4 sumariza o conteúdo exposto ao longo do capítulo, explicitando a originalidade e a contribuição técnica do método proposto no atendimento ao propósito estabelecido.

4.1 PROSPECÇÃO METODOLÓGICA

Conforme estabelecido no Capítulo 3, a formulação explícita da equação característica do modelo de um diodo apresentada em (3.9) — ou, alternativamente, em (3.10) — recai sobre o ramo principal da função W de Lambert, sendo o domínio fisicamente relevante para sua avaliação restrito ao semi-eixo real não negativo. A seguir, apresenta-se um panorama acerca dos principais métodos disponíveis para o cálculo numérico da função W de Lambert nesse escopo.

4.1.1 Abordagens de cálculo numérico

Em princípio, o cálculo do valor da função W de Lambert em um de seus ramos reais consiste em determinar a solução de sua equação definidora [122, 123, 153–155]. Ressalta-se que tal equação foi introduzida na Subseção 3.1.1 do capítulo anterior e é, por conveniência, reexibida abaixo.

$$W(z) \exp(W(z)) = z \quad (3.5)$$

Entretanto, a natureza transcendental dessa equação impede que a solução desejada seja obtida por meio de operações elementares. Mais especificamente, o fato de a incógnita $W(z)$ surgir simultaneamente dentro e fora do termo exponencial em (3.5) inviabiliza qualquer tentativa de isolá-la recorrendo apenas a manipulações algébricas [122, 123].

Na literatura específica, é possível encontrar diferentes tratamentos matemáticos para contornar essa limitação [41, 42, 123, 153, 178]. As alternativas de uso mais amplo caracterizam-se, em geral, pela aplicação de métodos iterativos, os quais constroem aproximações sucessivas para $W(z)$ a partir de expansões locais, aproximações assintóticas ou rearranjos específicos da própria equação definidora e de outras equações funcionais associadas [122, 154, 178].

Embora algumas abordagens desenvolvidas nessa linha apresentem particularidades interessantes — a exemplo do esquema recursivo de ordem superior proposto em [154] e das iterações baseadas em identidades logarítmicas discutidas em [155] — há uma estratégia que se

consolidou historicamente como a mais proeminente e difundida. Em essência, esta consiste em reinterpretar a equação definidora como um problema de determinação de raízes. Para tanto, reescreve-se (3.5) na forma apresentada em (4.1), adotando-se a variável w por convenção. Assim, a solução procurada para $W(z)$ é dada pelo valor de w que satisfaz implicitamente $\mathcal{R}(w) = 0$ no ramo real de interesse.

$$\mathcal{R}(w) := w \exp(w) - z \quad (4.1)$$

A técnica mais imediata para resolver esse problema é o *método de Newton–Raphson* ou, simplesmente, *método de Newton* [122]. Conforme comentado no Capítulo 2, trata-se de um algoritmo recursivo de solução de raízes fundamentado na expansão em série de Taylor truncada em seus dois primeiros termos, isto é, em uma aproximação linear da função na vizinhança da suposta raiz [149]. A sua atratividade decorre sobretudo do baixo custo computacional por iteração, uma vez que a expressão resultante exige apenas a avaliação de $\mathcal{R}(w)$ e de sua derivada de primeira ordem [122]. Ademais, sob condições regulares, o método apresenta convergência quadrática, de modo que o número de algoritmos corretos da solução tende a aproximadamente dobrar a cada iteração [149].

No caso específico de (4.1), entretanto, essa economia de custo por iteração não desempenha papel decisivo. Isso ocorre porque todas as derivadas do produto $w \exp(w)$ admitem forma fechada particularmente simples. Mais precisamente, a derivada de ordem m é dada por $(w + m) \exp(w)$, sendo obtida essencialmente ao custo de uma única multiplicação após o cálculo — já necessário — do termo $\exp(w)$ [122]. Assim, o acréscimo de custo associado ao uso de derivadas de ordem superior torna-se marginal nesse cenário, abrindo espaço para métodos iterativos de ordem mais elevada, potencialmente mais eficientes.

Posto isso, o *método de Halley* emerge como uma técnica particularmente eficaz [122]. Assim como o método de Newton, trata-se de um algoritmo de solução de raízes construído a partir da expansão em série de Taylor, porém truncada em seus três primeiros termos e posteriormente reorganizada de modo a produzir uma forma racional numericamente mais estável [179]. Como consequência, esse método fornece uma aproximação de caráter hiperbólico na vizinhança da raiz, por meio de uma expressão que exige a avaliação não apenas de $\mathcal{R}(w)$ e de sua derivada de primeira ordem, mas também de sua derivada de segunda ordem [179]. Quando aplicado a (4.1), tal expressão assume a forma apresentada em (4.2), em que $w^{(t)}$ denota a t -ésima aproximação gerada pelo processo iterativo. Em condições regulares, a execução recursiva dessa equação exibe convergência cúbica, fazendo com que o número de algoritmos corretos da solução tenda a aproximadamente triplicar a cada iteração [155, 179].

$$w^{(t+1)} = w^{(t)} - \frac{w^{(t)} \exp(w^{(t)}) - z}{\exp(w^{(t)}) (w^{(t)} + 1) - \frac{(w^{(t)} + 2) (w^{(t)} \exp(w^{(t)}) - z)}{2(w^{(t)} + 1)}} \quad (4.2)$$

A obra seminal apresentada em [122] reportou uma análise compreensiva do desempenho do método de Newton, do método de Halley e do método recursivo de ordem superior proposto em [154]. Os resultados evidenciaram que o método de Halley oferece desempenho mais atrativo ao longo de uma ampla faixa de argumentos, ramos e requisitos de precisão numérica [122]. Essa constatação, aliada às indicações registradas na própria obra para a construção de estimativas iniciais, exerceu papel decisivo na consolidação desse método como a técnica padrão para o cálculo da função W de Lambert em seus ramos reais [123].

Ademais, a difusão prática dessa escolha é reforçada pelo fato de que variantes do método de Halley encontram-se implementadas no núcleo de diversos *solvers* disponibilizados em ambientes de programação científica e bibliotecas computacionais amplamente difundidas, como a rotina correspondente no *software* MAPLE e as funções homônimas `lambertw` oferecidas pelo *software* MATLAB e pelo pacote *SciPy* da linguagem Python [122, 156, 157].

Embora esses *solvers* venham desempenhando um papel preponderante na popularização da função W de Lambert e, tipicamente, se mostrem efetivos para os mais diversos usos, alguns trabalhos publicados recentemente têm chamado atenção para o fato de que suas rotinas nem sempre fornecem o nível de eficiência computacional requerido em determinadas circunstâncias de emprego do modelo de um diodo [41, 42, 44]. A própria função `lambertw` do MATLAB, embora reconhecida pela sua elevada acurácia, pode apresentar tempos de simulação criticamente longos [41, 42, 44]. Assim, o cálculo reiterado da função W de Lambert com esses recursos revela-se um ponto sensível, sobretudo no âmbito de problemas de extração dos parâmetros¹, nos quais de dezenas a milhões de avaliações da equação característica do modelo de um diodo podem ser regularmente exigidas.

Atentos a essa situação, alguns grupos de pesquisadores da área de modelagem fotovoltaica passaram a desenvolver métodos voltados ao cálculo eficiente da função W de Lambert no escopo do modelo de um diodo. Nesse contexto, duas oportunidades de aprimoramento mostraram-se particularmente promissoras. Em primeiro lugar, como o domínio fisicamente relevante à equação característica restringe-se a $z \in [0, \infty)$, a estratégia de cálculo pode ser especializada para esse intervalo, explorando propriedades locais e evitando a necessidade de contemplar outras regiões do ramo principal ou mesmo o ramo inferior. Em segundo lugar, a sobrecarga computacional associada a *solvers* de uso geral pode ser reduzida por meio de formulações e implementações mais diretas, com lógica mais uniforme e custo mais previsível, aspecto especialmente relevante em aplicações que demandam avaliações massivas.

Diante dessa conjuntura, apresenta-se, a seguir, um panorama desses métodos especializados ao modelo de um diodo, acompanhado de uma análise de seus potenciais e limitações.

¹ Note que cada avaliação da função objetivo do problema implica no cálculo de N avaliações da equação característica do modelo de um diodo e, portanto, de N avaliações da função W de Lambert, sendo cada uma referente a um ponto específico da curva $I-V$ em questão.

4.1.2 Método híbrido

Um dos primeiros métodos desenvolvidos especificamente para o cálculo da função W de Lambert no escopo do modelo de um diodo foi introduzido em [180]. Posteriormente batizado como *método híbrido* por um de seus autores, esse método combina duas expansões em série da função W de Lambert em seu ramo principal [41]. A primeira, dada por (4.3a), consiste nos seis primeiros termos da expansão em série divulgada na obra seminal apresentada em [122] e é aplicada para $z \in [0, 9)$. A segunda, dada por (4.4), consiste nos sete primeiros termos da expansão assintótica e é aplicada para $z \in [9, \infty)$.

$$W_0(z) = r + \frac{r}{1+r}d + \frac{r}{2(1+r)^3}d^2 - \frac{r(2r-1)}{6(1+r)^5}d^3 \quad (4.3a)$$

$$+ \frac{r(6r^2 - 8r + 1)}{24(1+r)^7}d^4 - \frac{r(24r^3 - 58r^2 + 22r - 1)}{120(1+r)^9}d^5 \quad (4.3b)$$

$$r = z \exp(-1) \quad (4.3c)$$

$$d = 1 - z \exp(-1) \quad (4.3d)$$

$$W_0(z) = L_1 - L_2 + \frac{L_2}{L_1} + \frac{L_2(-2 + L_2)}{2L_1^2} \quad (4.4a)$$

$$+ \frac{L_2(6 - 9L_2 + 2L_2^2)}{6L_1^3} + \frac{L_2(-12 + 36L_2 - 22L_2^2 + 3L_2^3)}{12L_1^4} \quad (4.4b)$$

$$+ \frac{L_2(60 - 300L_2 + 350L_2^2 - 125L_2^3 + 12L_2^4)}{60L_1^5} \quad (4.4c)$$

$$L_1 = \ln(z) \quad (4.4d)$$

$$L_2 = \ln(\ln(z)) \quad (4.4e)$$

Os resultados reportados em [41] e [44] indicam que o método híbrido é significativamente mais rápido que a função `lambertw` do MATLAB. Além disso, esses resultados indicam que as suas aproximações permitem que o modelo de um diodo retorne valores de corrente, por meio de (3.9), com incertezas da ordem de 10^{-5} A.

4.1.3 Método numérico

Mais recentemente, o trabalho apresentado em [42] propôs uma interessante estratégia para o cálculo da função W de Lambert no escopo do modelo de um diodo. Trabalhando com diferentes parametrizações e equações funcionais, seus autores recomendaram o uso dos conjuntos de expressões apresentados em (4.5) e (4.6). Nesse sentido, o valor da variável auxiliar u em (4.5a) é obtido como a raiz que satisfaz (4.5b). De forma análoga, o valor da variável auxiliar w em (4.6a) é obtido como a raiz que satisfaz (4.6b). Os autores sugerem, então, a utilização do método de Newton (leia-se Newton–Raphson) para determinar essas raízes,

considerando (4.5c) e (4.6c) como estimativas iniciais. A escolha entre um ou outro conjunto segue uma lógica condicional: (4.5) é aplicado para $z \in [0, \exp(1)]$, enquanto (4.6) é aplicado para $z \in (\exp(1), \infty)$.

$$W_0(z) = zu \quad (4.5a)$$

$$u - (\exp(-z))^u = 0 \quad (4.5b)$$

$$u^0 = \frac{1 + \ln(1 - \ln(\exp(-z)))}{1 - 2 \ln(\exp(-z))} \quad (4.5c)$$

$$W_0(z) = w \quad (4.6a)$$

$$\ln(w) + w - \ln(z) = 0 \quad (4.6b)$$

$$w^0 = \ln(z) - \ln(\ln(z)) + \frac{\ln(\ln(z))}{\ln(z)} + \frac{\ln(\ln(z))(-2 + \ln(\ln(z)))}{2(\ln(z))^2} \quad (4.6c)$$

Ao ser aplicado com um limite suficientemente elevado de iterações internas do método de Newton, os resultados reportados para esse chamado *método numérico* indicam que as suas soluções permitem que o modelo de um diodo retorne valores de corrente, por meio de (3.9), com incertezas da ordem de 10^{-16} A [42]. Além disso, esses resultados indicam que tais soluções podem ser obtidas em um tempo de simulação cerca de três ordens de grandeza inferior ao requerido pela função `lambertw` do MATLAB [42].

4.1.4 Método explícito

Com base na mesma estratégia do método numérico, os autores do trabalho apresentado em [42] propuseram uma alternativa simplificada para o cálculo da função W de Lambert no escopo do modelo de um diodo, a qual foi apelidada de *método explícito*. A principal diferença reside no fato de que esse método não permite qualquer iteração interna do método de Newton. Assim, uma vez selecionada a formulação adequada em função do valor do argumento z , calcula-se uma estimativa inicial de acordo com (4.5c) ou (4.6c) e executa-se uma única vez a equação resultante do método de Newton associada a (4.5b) ou (4.6b). Vale ressaltar que tais equações possuem uma forma relativamente simples, conforme mostram (4.7) e (4.8), respectivamente.

$$u^1 = u^0 - \frac{u^0 - (\exp(-z))^{u^0}}{1 + z (\exp(-z))^{u^0}} \quad (4.7)$$

$$w^1 = w^0 - \frac{\ln(w^0) + w^0 - \ln(z)}{1 + 1/w^0} \quad (4.8)$$

Assim como era de se esperar, os resultados reportados em [42] confirmam que o método explícito é mais rápido que o método numérico, porém ao custo de fornecer soluções de menor qualidade. Em linhas gerais, tais soluções permitem que o modelo de um diodo retorne valores de corrente, por meio de (3.9), com incertezas comparáveis ou ligeiramente inferiores às aquelas do método híbrido [42].

4.1.5 Análise compreensiva

Do ponto de vista prático, os três métodos apresentados acima para o cálculo da função W de Lambert no escopo do modelo de um diodo possuem vantagens e limitações próprias. À primeira vista, os métodos híbrido e explícito se destacam por serem não iterativos e, portanto, por apresentarem custo computacional fixo. Contudo, o nível de acurácia das soluções retornadas por esses dois métodos é um aspecto que merece atenção. Considerando o padrão IEEE de ponto flutuante em precisão dupla, a ordem da precisão de máquina é $\epsilon_{\text{maq}} \approx 2,2 \times 10^{-16}$. Assim, as soluções fornecidas por esses dois métodos podem ser consideradas de qualidade intermediária e, inclusive, revelar-se inadequadas para determinados estudos. Por exemplo, os resultados apresentados em [44] demonstram que as soluções encontradas para o tradicional problema da célula RTC France via método híbrido são significativamente inferiores às aquelas obtidas pela função correspondente do MATLAB e, conseqüentemente, ineficazes.

Por sua vez, o método numérico se destaca pelo potencial de alcançar soluções tão acuradas que a incerteza resultante se encontra da ordem da própria precisão de máquina. Apesar disso, os resultados reportados em [42] deixam claro que, em geral, mais de uma iteração do método de Newton é necessária para a obtenção dessas soluções; caso contrário, o método explícito também retornaria aproximações com esse nível elevado de qualidade. Dessa forma, o método numérico apresenta um custo computacional variável, que tende a ser superior àquele verificado para os métodos não iterativos.

Adicionalmente, é importante observar que tais métodos compartilham um ponto sensível: todos dedicam expressões específicas a diferentes subintervalos do domínio $z \in [0, \infty)$. Sob a perspectiva de implementação, essa lógica condicional representa um inconveniente, sobretudo porque agrega complexidade ao código e à eventual necessidade de validação de resultados. Há ainda, como agravante, o fato de que cada ramificação conduz a blocos matemáticos substancialmente distintos — ora polinomiais, ora apoiados em formulações com logaritmos ou exponenciais —, implicando regimes de cálculo heterogêneos e, por conseqüência, um tempo de processamento menos uniforme.

Essas constatações conduzem à ideia de que há espaço para o desenvolvimento de novos métodos de cálculo da função W de Lambert no escopo do modelo de um diodo, desde que apresentem um melhor compromisso entre acurácia e custo computacional. Não obstante, é crucial notar que a concepção de um método desse tipo a partir das lógicas empregadas naqueles já existentes não se revela promissora. Por exemplo, aprimorar a acurácia do método híbrido inevitavelmente exigiria expandir as séries utilizadas, considerando termos adicionais ou até mesmo recorrer a novas expansões em certos subintervalos, o que aumentaria simultaneamente a complexidade do método e o seu custo computacional. De modo semelhante, melhorar o desempenho dos métodos explícito e numérico implicaria substituir o método de Newton por algoritmos de solução de raízes de ordem superior, mais onerosos, ou lidar com o desafio de desenvolver estimativas iniciais mais acuradas para cada um dos seus conjuntos de expressões.

Além disso, em qualquer desses cenários, a natureza condicional intrínseca a tais métodos permaneceria inalterada. Sendo assim, a concepção desse novo método exige a exploração de novos conceitos e estruturas matemáticas, capazes de superar as limitações dessas alternativas de estado da arte.

4.2 MÉTODO PROPOSTO

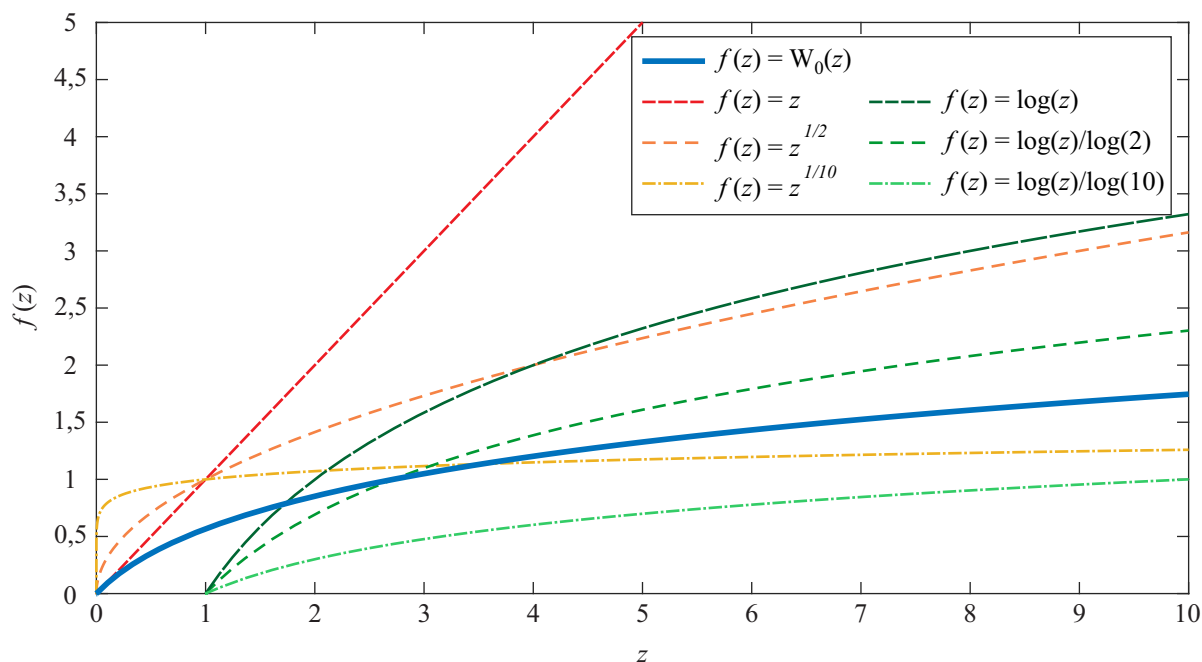
A análise conduzida na Seção 4.1 revelou que há espaço para o desenvolvimento de um novo método de cálculo da função W de Lambert no escopo do modelo de um diodo. Nesse contexto, tal método deve ser capaz de: (i) retornar soluções cuja acurácia e, conseqüentemente, incerteza introduzida no modelo sejam da ordem da precisão de máquina; (ii) operar com o menor custo computacional possível, sendo este preferencialmente fixo; e (iii) evitar, tanto quanto viável, o uso de estruturas condicionais em sua formulação.

Posto isso, conduziu-se uma investigação abrangente a fim de assimilar a experiência acumulada na literatura e identificar uma estrutura matemática ainda formalmente inexplorada, mas adequada ao desenvolvimento do método pretendido. Como resultado, verificou-se que a lógica comum aos *métodos preditores–corretores de resolução de equações não lineares* poderia servir de base para a construção de um procedimento que atendesse às características mencionadas [181, 182].

Em linhas gerais, um método preditor–corretor é concebido a partir da associação de dois métodos já existentes para o mesmo propósito, de modo a obter uma técnica potencialmente mais eficiente [119, 181, 182]. No âmbito da resolução de equações não lineares, esse tipo de método geralmente parte de uma estimativa inicial arbitrária e realiza dois passos encadeados para aprimorar a sua acurácia [181, 182]. No primeiro passo, obtém-se uma solução aproximada para a equação em estudo por meio da expressão matemática de um dos métodos selecionados. No segundo passo, corrige-se parte das casas decimais dessa solução aproximada por meio da expressão pertencente ao outro método selecionado. Caso a solução obtida no segundo passo ainda não apresente o nível de qualidade desejado, esta passa a servir como nova estimativa inicial para o primeiro passo, e repete-se sucessivamente o procedimento descrito até que o nível de acurácia requerido seja atingido.

À luz dessas considerações, um método preditor–corretor foi propriamente desenvolvido para calcular a função W de Lambert no domínio fisicamente relevante ao modelo de um diodo, isto é, no semi-eixo real não negativo. Todavia, a fim de garantir que sua aplicação seja autocontida e alcance soluções com o nível de acurácia desejado de forma não iterativa, uma equação específica foi adotada para determinar as estimativas iniciais necessárias. Os detalhes desse *método preditor–corretor autoiniciado* — ou, na nomenclatura internacional adotada para fins de publicação, “*self-started predictor–corrector method*” — são apresentados a seguir. Ao final, suas principais equações são sintetizadas e a sua implementação computacional é devidamente elucidada.

Figura 12 – Comparação da função W de Lambert com outras funções regulares, considerando o domínio fisicamente relevante ao modelo de um diodo.



4.2.1 Passo inicial

Encontrar uma estimativa, ainda que grosseira, do valor da função W de Lambert no semi-eixo real não negativo não é uma tarefa trivial. Como ilustrado na Figura 12, esse intervalo exhibe um comportamento não linear peculiar.

Para lidar com esse desafio, os métodos apresentados na Seção 4.1, assim como a maioria daqueles propostos para calcular a função W de Lambert em escopos mais amplos, adotam estratégias do tipo “*divide et impera*” [42, 122, 153, 154, 173, 180]. Em termos simples, tal estratégia consiste em particionar o intervalo de interesse do ramo principal em alguns subintervalos e, então, atribuir a cada um destes uma expressão matemática representativa. Embora algumas dessas estratégias exibam níveis de acurácia surpreendentes (veja, por exemplo, [173]), a complexidade das equações envolvidas e, sobretudo, a sua natureza condicional tornam a implementação computacional custosa.

Por outro lado, é possível encontrar na literatura específica formulações não condicionais construídas para representar razoavelmente o comportamento da função W de Lambert no ramo principal [178]. Um exemplo particularmente interessante é introduzido no trabalho apresentado em [155]. Segundo os seus autores, a expressão dada em (4.9) reproduz os dois primeiros termos da expansão em série de potências da função W de Lambert em torno da origem ($z = 0$), bem como os dois primeiros termos de uma expansão logarítmica assintótica para valores elevados do argumento z [155].

$$W_i = \ln\left(1 + \frac{z}{1 + 0,5 \ln(1 + z)}\right) \quad (4.9)$$

Em consequência, o estudo reportado em [178] revela que tal equação fornece soluções aproximadas para a função W de Lambert com erro relativo limitado superiormente a aproximadamente $3,53 \times 10^{-2}$ para $z \in [0, \infty)$. Embora esse nível de acurácia seja relativamente modesto, verifica-se que os seus resultados são adequados para uso como estimativas iniciais [155]. Assim, considerando sua baixa complexidade e boa aderência ao comportamento geral da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo, a expressão (4.9) foi escolhida como ponto de partida para o método proposto. Vale notar, inclusive, que o subscrito “ i ” foi atribuído a W_i em alusão à ideia de “inicial”.

4.2.2 Passo preditor

Uma alternativa para aprimorar a estimativa inicial dada por (4.9) consiste em explorar diretamente a própria estrutura da função W de Lambert [154, 155]. Para tornar essa ideia mais clara, considere que a estimativa inicial W_i difere da solução exata $W_0(z)$ por um erro denotado por φ_i . Em termos matemáticos, essa relação é expressa de acordo com (4.10).

$$W_0(z) = W_i + \varphi_i \quad (4.10)$$

Ao substituir (4.10) na equação definidora da função W de Lambert, apresentada em (3.5), e realizar algumas manipulações algébricas, obtém-se que o erro φ_i deve satisfazer a relação apresentada em (4.11). Se essa equação pudesse ser prontamente satisfeita, a estimativa inicial W_i poderia ser imediatamente corrigida. Contudo, a sua natureza é transcendente, pois φ_i aparece simultaneamente dentro e fora do termo logarítmico à direita e, portanto, não há como isolá-lo recorrendo apenas a operações elementares.

$$\varphi_i = \ln\left(\frac{z}{W_i}\right) - W_i - \ln\left(1 + \frac{\varphi_i}{W_i}\right) \quad (4.11)$$

Para contornar essa dificuldade, o trabalho apresentado em [155] sugere expandir o termo logarítmico da última parcela de (4.11) em série de Taylor. Considerando apenas o primeiro termo dessa expansão, encontra-se a aproximação indicada em (4.12).

$$\ln\left(1 + \frac{\varphi_i}{W_i}\right) \approx \frac{\varphi_i}{W_i} \quad (4.12)$$

Ao substituir (4.12) em (4.11) e realizar novas manipulações, conclui-se que o erro associado à estimativa inicial W_i pode ser aproximadamente determinado por (4.13).

$$\varphi_i \approx \frac{W_i}{1 + W_i} \left[\ln\left(\frac{z}{W_i}\right) - W_i \right] \quad (4.13)$$

Assim, ao combinar (4.10) e (4.13), produz-se (4.14). Essa expressão constitui a base do método recursivo proposto em [155] para calcular a função W de Lambert no ramo principal, tendo em vista um escopo genérico.

$$W_p = \frac{W_i}{1 + W_i} \left[1 + \ln\left(\frac{z}{W_i}\right) \right] \quad (4.14)$$

Ainda de acordo com [155], a relação (4.14) é capaz de realizar uma correção de segunda ordem; isto é, o número de casas decimais corretas da estimativa inicial tende a aproximadamente dobrar após sua aplicação. Corroborando esse potencial, o estudo apresentado em [178] indica que a combinação entre (4.9) e (4.14) pode gerar soluções aproximadas para a função W de Lambert com erro relativo limitado superiormente a cerca de $2,16 \times 10^{-4}$ para $z \in [0, \infty)$. Embora tal nível de acurácia ainda seja reduzido, destaca-se novamente a baixa complexidade das equações envolvidas. Por essa razão, (4.14) foi escolhida para compor o *passo predictor* do método proposto — daí o subscrito “ p ” em W_p .

4.2.3 Passo corretor

O passo corretor deve ser executado imediatamente após o passo predictor. Assim, a relação apresentada em (4.11) assume a forma mostrada em (4.15), na qual φ_p denota o erro associado à diferença entre a solução aproximada W_p e a solução exata $W_0(z)$.

$$\varphi_p = \ln\left(\frac{z}{W_p}\right) - W_p - \ln\left(1 + \frac{\varphi_p}{W_p}\right) \quad (4.15)$$

Conforme discutido anteriormente, a expansão em série de Taylor do último termo logarítmico poderia ser utilizada para construir uma equação corretiva de segunda ordem, análoga àquela que originou o passo predictor. Entretanto, mostra-se mais vantajoso, neste ponto, empregar formulações ligeiramente mais elaboradas, porém capazes de produzir correções substancialmente mais precisas. Uma alternativa particularmente eficiente consiste em expandir o referido termo logarítmico não em série de Taylor, mas em fração contínua [154]. Considerando tal expansão truncada em seus dois primeiros termos, obtém-se (4.16).

$$\ln\left(1 + \frac{\varphi_p}{W_p}\right) \approx \frac{\varphi_p W_p + \frac{1}{6} \varphi_p^2}{W_p^2 + \frac{2}{3} \varphi_p W_p} \quad (4.16)$$

Substituindo (4.16) em (4.15) e realizando manipulações algébricas, chega-se à equação quadrática dada em (4.17)². Para fins de clareza, introduz-se o parâmetro auxiliar y , definido em (4.17b), o qual simplifica a visualização dos termos envolvidos.

$$\underbrace{\left(\frac{2}{3} W_p + \frac{1}{6}\right)}_a \varphi_p^2 + \underbrace{\left(W_p^2 + W_p - \frac{2}{3} y W_p\right)}_b \varphi_p + \underbrace{(-y W_p^2)}_c = 0 \quad (4.17a)$$

$$y = \ln\left(\frac{z}{W_p}\right) - W_p \quad (4.17b)$$

² Ressalta-se que o termo a assinalado nessa equação denota um coeficiente do polinômio de segundo grau em φ_p , não devendo ser confundido com o parâmetro concentrado do modelo de um diodo, listado na Tabela 1 e aplicado em (3.10).

Em um olhar mais atento, nota-se que (4.17a) é um polinômio de segundo grau em φ_p . Os seus coeficientes característicos, usualmente denotados por a , b e c , estão assinalados em (4.17a) e explicitados em (4.18).

$$a = \frac{2}{3}W_p + \frac{1}{6} \quad (4.18a)$$

$$b = W_p^2 + W_p - \frac{2}{3}yW_p \quad (4.18b)$$

$$c = -yW_p^2 \quad (4.18c)$$

Em princípio, o erro φ_p é determinado por uma das duas raízes de (4.17). Essas raízes são dadas pela fórmula quadrática tradicional, reproduzida em (4.19), na qual o operador “ \pm ” representa os sinais de mais e de menos e permite uma compactação da expressão das duas soluções possíveis [183].

$$\varphi_p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (4.19)$$

Do ponto de vista analítico, (4.19) caracteriza completamente o conjunto de soluções de (4.17). Contudo, em computação numérica, sua aplicação direta pode ser problemática [183]. Em particular, quando $b^2 \gg 4ac$, observa-se que $\sqrt{b^2 - 4ac}$ torna-se numericamente muito próximo de $|b|$. Nessas circunstâncias, a raiz associada ao sinal positivo em (4.19) envolve a subtração entre dois números quase iguais, o que configura uma situação clássica para o surgimento do fenômeno de *cancelamento subtrativo* (ou *cancelamento catastrófico*), o qual é responsável pela perda de dígitos significativos da solução [119]. Com efeito, a Figura 13 evidencia que tal condição inconveniente é primordialmente verificada no escopo do modelo de um diodo.

Para contornar esse problema, diversos trabalhos recomendam substituir a fórmula quadrática tradicional pela chamada *fórmula citardauq* [119, 183]. Curiosamente, essa denominação origina-se da escrita do termo em inglês “quadratic” em ordem inversa [183]. No presente caso, tal fórmula assume a forma indicada em (4.20).

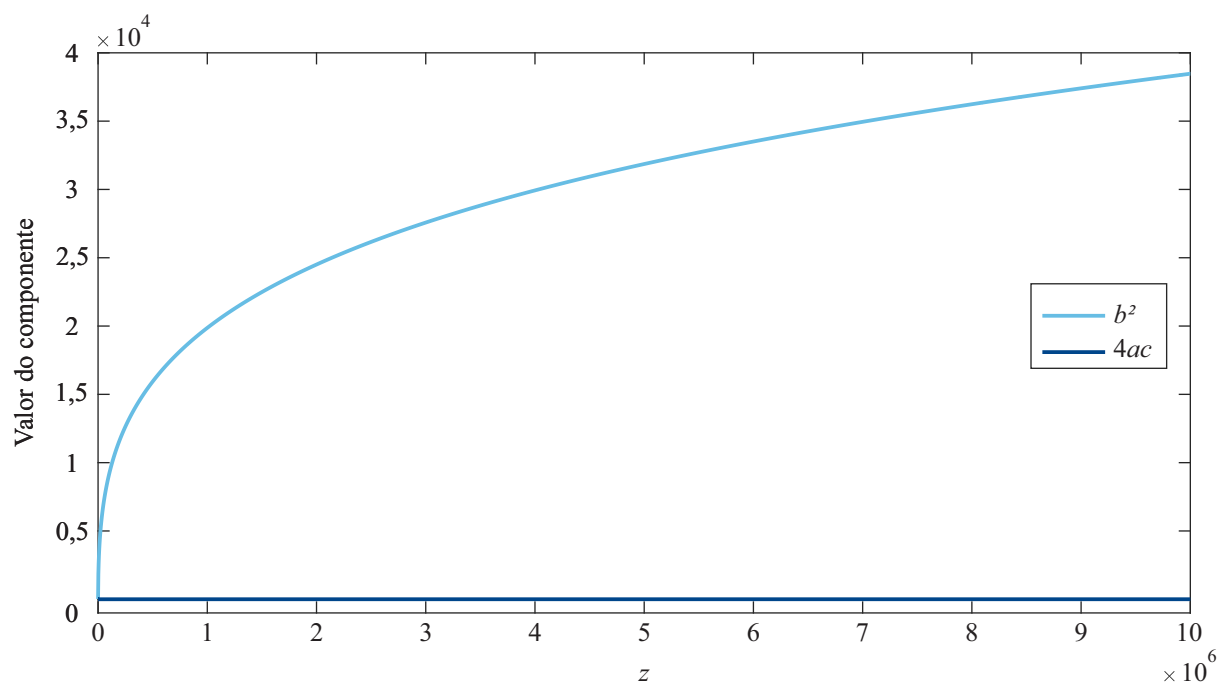
$$\varphi_p = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}} \quad (4.20)$$

Ao resolver (4.17) por meio da raiz associada ao sinal positivo em (4.20), ressignificando os coeficientes conforme (4.18) e realizando algumas simplificações algébricas, obtém-se a expressão apresentada em (4.21). A escolha dessa raiz se justifica pelo fato de que esta tende a 0 quando $y \rightarrow 0$, comportamento compatível com o esperado para o erro φ_p [154]. Em contraste, a raiz associada ao sinal negativo pode crescer em módulo de forma descontrolada, acarretando instabilidade numérica e, portanto, não deve ser empregada.

$$\varphi_p = \frac{2yW_p}{(1 + W_p - \frac{2}{3}y) + \sqrt{(1 + W_p + \frac{2}{3}y)^2 - 2y}} \quad (4.21)$$

A partir de (4.21), é possível finalmente escrever a equação corretiva apresentada em (4.22). Cabe mencionar que essa expressão constitui o núcleo do método recursivo desenvolvido

Figura 13 – Relação entre b^2 e $4ac$ verificada para a equação quadrática associada ao passo corretor no escopo do modelo de um diodo.



em [154], ainda na década de 1970, para o cálculo da função W de Lambert no ramo principal, tendo em vista um escopo genérico.

$$W_c = W_p + \frac{2yW_p}{(1 + W_p - \frac{2}{3}y) + \sqrt{(1 + W_p + \frac{2}{3}y)^2 - 2y}} \quad (4.22)$$

Embora mais complexa do que a estrutura de (4.14), (4.22) é capaz de conduzir uma correção de quarta ordem [173]. Na prática, isso significa que o número de casas decimais corretas da solução aproximada obtida no passo preditor tende a ser praticamente quadruplicado após a sua aplicação. Considerando que tal solução aproximada apresenta erro relativo limitado superiormente a algo em torno de $2,16 \times 10^{-4}$ para $z \in [0, \infty)$, essa correção de quarta ordem permite, então, que W_c atinja um erro relativo da ordem de 10^{-16} no mesmo intervalo. Por essa razão, (4.22) foi escolhida para compor o *passo corretor* do método proposto — daí o subscrito “c” em W_c .

4.2.4 Consolidação matemática

À luz do exposto, o método proposto pode ser sintetizado no conjunto de expressões apresentado em (4.23). Ressalta-se que (4.23a), (4.23b), (4.23c) e (4.23d) devem ser avaliadas sequencialmente, uma vez que cada passo depende diretamente do resultado do passo precedente. Não há, entretanto, impedimento técnico para condensar tais expressões em formas mais compactas — inclusive em uma única equação — caso isso seja do interesse da parte responsável pela sua aplicação. A apresentação segmentada é adotada aqui apenas para fins de clareza expositiva.

$$W_i = \ln\left(1 + \frac{z}{1 + 0,5 \ln(1 + z)}\right) \quad (4.23a)$$

$$W_p = \frac{W_i}{1 + W_i} \left[1 + \ln\left(\frac{z}{W_i}\right)\right] \quad (4.23b)$$

$$y = \ln\left(\frac{z}{W_p}\right) - W_p \quad (4.23c)$$

$$W_c = W_p + \frac{2yW_p}{(1 + W_p - \frac{2}{3}y) + \sqrt{(1 + W_p + \frac{2}{3}y)^2 - 2y}} \quad (4.23d)$$

Conforme destacado anteriormente, (4.23) é capaz de retornar, por construção, valores para a função W de Lambert com erro relativo em torno da ordem de 10^{-16} em todo o semi-eixo real não negativo. Considerando que, no padrão IEEE de ponto flutuante em precisão dupla $\epsilon_{\text{maq}} \approx 2,2 \times 10^{-16}$, conclui-se que qualquer solução indicada por W_c , já na primeira execução da sequência “passo inicial \rightarrow passo preditor \rightarrow passo corretor”, tende a coincidir com $W_0(z)$ até o limite imposto pela própria aritmética. Assim, não se faz necessária qualquer iteração entre os passos preditor e corretor, nem o estabelecimento de ramificações lógicas acompanhadas de conjuntos de expressões complementares.

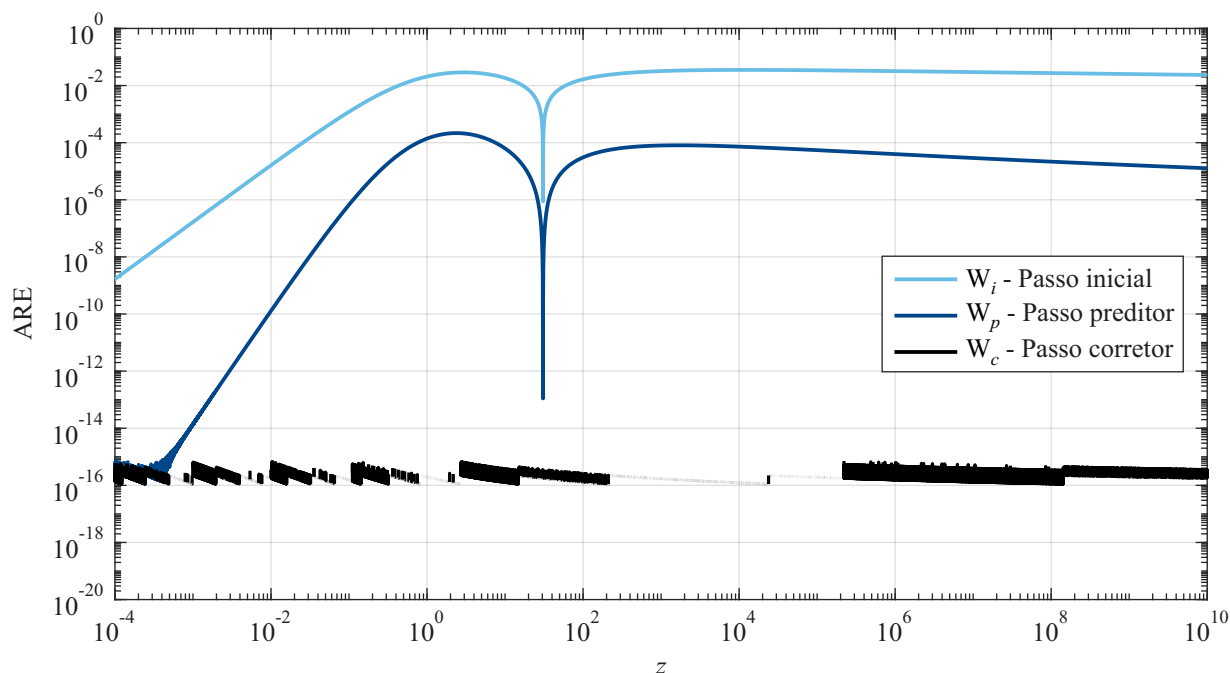
A fim de ilustrar esse desempenho, a Figura 14 apresenta o perfil do *erro relativo absoluto* (“*absolute relative error*”, ARE) associado a cada passo do método proposto, considerando valores reais não negativos do argumento z . Essa métrica é definida em (4.24), em que $W_0(z)_{\text{calc}}$ ³ denota a solução calculada em um passo específico do método proposto — isto é, W_i , W_p ou W_c — e $W_0(z)_{\text{ref}}$ corresponde ao valor de referência usualmente fornecido por um *solver* bem estabelecido, em particular a função `lambertw` do MATLAB [41, 42, 44].

$$\text{ARE}(z) = \left| \frac{W_0(z)_{\text{calc}} - W_0(z)_{\text{ref}}}{W_0(z)_{\text{ref}}} \right| \quad (4.24)$$

A análise da Figura 14 revela que os valores de ARE obtidos nos passos inicial e preditor estão em plena consonância com os limites de erro documentados na literatura específica. Ademais, observa-se que os valores de ARE associados ao passo corretor permanecem sistematicamente abaixo da ordem de 10^{-15} , evidenciando de modo inequívoco a eficácia da sua correção de quarta ordem. Em consequência, confirma-se que W_c é numericamente compatível com $W_0(z)$ até o limite de precisão de máquina, de modo que o método proposto retorna, na prática, os efetivos valores da função $W_0(z)$ no semi-eixo real não negativo.

³ Esclarece-se que o subscrito “*calc*” foi adotado aqui, em dissonância com o subscrito “*c*” empregado nas formulações de métricas de erro anteriores, para evitar confusão com o subscrito do passo corretor.

Figura 14 – Perfil do erro relativo absoluto (ARE) verificado em cada etapa do método proposto.



4.2.5 Implementação computacional

A rigor, a implementação computacional do método proposto constitui uma tarefa pouco onerosa, pois envolve a declaração de apenas quatro equações encadeadas, as quais apresentam complexidade moderada. Não obstante, fazem-se necessárias algumas medidas auxiliares em prol de sua robustez.

O arcabouço analítico que fundamenta o método proposto estabelece que este é, em sua formulação teórica, inteiramente não condicional, uma vez que a sequência (4.23a)–(4.23d) é válida e aplicável em todo o semi-eixo real não negativo. Ao se transpor esse conjunto de expressões para o plano computacional, entretanto, emerge uma questão relevante: ainda que o procedimento executório seja conceitualmente livre de ramificações lógicas, a sua implementação numérica requer ponderação específica para o caso em que o argumento da função W de Lambert coincida com a origem do eixo real, sob pena de se incorrer em indeterminações algébricas e dificuldades numéricas.

Para fins de clareza, considere $z = 0$. O passo inicial do método, definido por (4.23a), retorna imediatamente $W_i = 0$, o qual já representa corretamente o valor exato de $W_0(0)$. Já no âmbito do passo preditor, verifica-se em (4.23b) que $z/W_i(z) \rightarrow 1$ quando $z \rightarrow 0^+$, o que conduz a $\ln(1) = 0$ e, portanto, a $W_p = 0$. Todavia, essa regularidade se manifesta apenas no contexto analítico. Na prática, a substituição direta dos valores de z e W_i em (4.23b) produz o termo $\ln(0/0)$, o qual é algebricamente indefinido.

A constatação dessa questão, estritamente associada ao argumento $z = 0$, justifica, sob a perspectiva computacional, a introdução de uma verificação pontual para esse caso, retornando-se

prontamente $W_0(0) = 0$. Cumpre observar, contudo, que tal condicional não altera a natureza do método: trata-se apenas da codificação explícita de uma condição de fronteira imposta pela forma algébrica empregada na sua concepção, e não de uma regra seletiva entre conjuntos de expressões concorrentes. Assim, diferentemente do observado nos métodos especializados discutidos no item 4.1.5, esse ajuste não configura alternância estrutural entre blocos matemáticos distintos nem introduz caminhos de cálculo alternativos, permanecendo, portanto, neutro em termos de custo computacional e de arquitetura lógica.

À luz dessas considerações, apresenta-se no Algoritmo 1 o pseudocódigo de implementação computacional do método proposto, ora nomeado como *método preditor–corretor autoiniciado*. Complementarmente, elucida-se no Algoritmo 2 a sua aplicação no escopo do modelo de um diodo.

Algoritmo 1 Cálculo da função W de Lambert via método preditor–corretor autoiniciado.

Entrada: z

Saída: $W_0(z)$

- 1: **Se** $z = 0$ **então**
 - 2: **Retorna** $W_0(0) = 0$
 - 3: **Caso contrário**
 - 4: Calcular W_i usando (4.23a)
 - 5: Calcular W_p usando (4.23b)
 - 6: Calcular y usando (4.23c)
 - 7: Calcular W_c usando (4.23d)
 - 8: **Retorna** $W_0(z) = W_c$
 - 9: **Fim**
-

Algoritmo 2 Cálculo da equação característica do modelo de um diodo via função W de Lambert.

Entradas: $k, q, N_p, N_s, T_c, V, I_{ph}, I_0, n, R_{sh}, R_s$

Saída: I

- 1: Calcular z via (3.9b)
 - 2: Calcular $W_0(z)$ conforme Algoritmo 1
 - 3: Calcular I via (3.9a)
 - 4: **Retorna** I
-

4.3 RESULTADOS EXPERIMENTAIS

Em consonância com as práticas adotadas em trabalhos correlatos, os seis estudos de caso recentemente introduzidos por [42] foram tomados como base para a validação do método proposto. Cada um destes é caracterizado por uma curva $I-V$, gerada sinteticamente por meio da abordagem formal da equação característica do modelo de um diodo via função W de Lambert. É pertinente destacar que os valores assumidos pelo argumento z em seus âmbitos variam entre as ordens de grandeza de 10^{-12} e 10^{35} [42]. Esses valores, portanto, abrangem uma faixa ampla e representativa do semi-eixo real não negativo, que é precisamente o domínio fisicamente relevante ao escopo do modelo de um diodo.

Tabela 11 – Parâmetros característicos dos estudos de caso.

Parâmetro	Unidade	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
I'_{ph}	A	15,8800	1,0320	3,6540	0,5780	0,7610	4,8020
I'_0	A	7,4400E-10	2,5130E-6	3,9990E-21	1,3400E-10	3,1070E-7	4,0160E-7
a	V	14,6700	1,3000	0,5160	0,0118	0,0390	0,0370
R'_{sh}	Ω	425,2000	744,7140	2329,0000	612,0000	52,8900	1167,0000
R'_s	Ω	2,0400	1,2390	2,6900	0,0127	0,0370	0,5906
V_{oc}	V	348,1000	16,7750	24,8930	0,2620	0,5730	0,6030

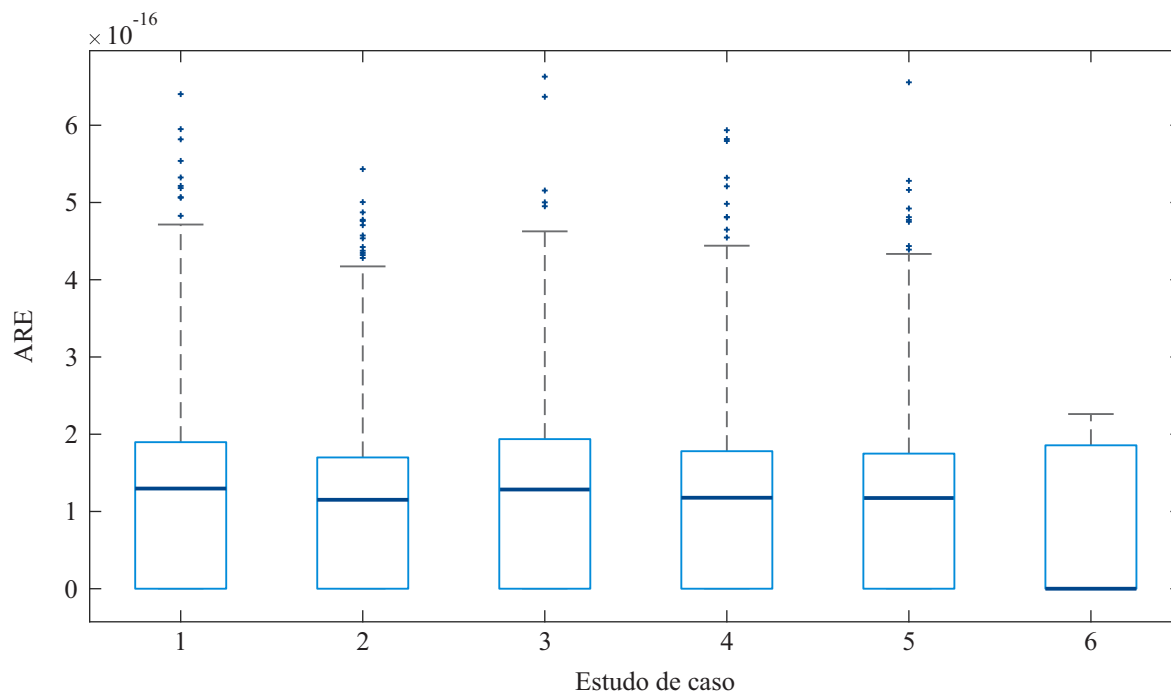
Por razões que lhes são próprias, os autores do referido trabalho optaram por empregar a expressão compacta da equação característica do modelo de um diodo — e não aquela geral que vem sendo adotada ao longo desta pesquisa. Ressalta-se que essa expressão foi explicitada em (3.10), no âmbito do capítulo anterior. Dessa forma, a curva I – V de cada estudo de caso é determinada a partir dos dados apresentados na Tabela 11 [42].

Cabe esclarecer que no processo de geração dessas curvas I – V , os valores correspondentes dos parâmetros concentrados I'_{ph} , I'_0 , a , R'_{sh} e R'_s são substituídos em (3.10), sendo um valor de I é calculado para cada valor de V considerado entre zero e a tensão de circuito aberto (V_{oc}). Esse procedimento de cálculo, por sua vez, segue exatamente a lógica exaltada no Algoritmo 2. Assim, a curva I – V de cada estudo de caso foi inicialmente gerada aplicando-se o método proposto para o cálculo da função W de Lambert, conforme o Algoritmo 1. Para fins de comparação de desempenho, essas mesmas curvas foram posteriormente obtidas recorrendo a outros métodos correlatos, os quais são especificados oportunamente nas subseções seguintes.

Todos os métodos considerados foram implementados no MATLAB, versão 2016b. Consequentemente, a função `lambertw` disponibilizada por esse *software* foi incluída como alternativa de referência nos estudos [156]. Conforme discutido ao longo deste capítulo, tal *solver* retorna soluções acuradas para a função W de Lambert no ramo principal, aplicando essencialmente o método de Halley [156]. É importante também mencionar que a função `linspace` disponibilizada pelo mesmo *software* foi usada para especificar $N = 1000$ pontos igualmente espaçados entre 0 e V_{oc} para cada curva I – V em questão [184]. As simulações foram todas executadas em um computador equipado com processador Intel Core i7-4510U @ 2/2,60 GHz, 8 GB de memória RAM, sob o sistema operacional Windows 10 de 64 bits.

Os resultados obtidos foram avaliados em termos de erro relativo absoluto, raiz do erro quadrático médio e tempo médio total de simulação. Em particular, o erro relativo absoluto, ou simplesmente ARE, foi calculado para cada solução retornada por cada método implementado, conforme (4.24). Já a raiz do erro quadrático médio, ou RMSE, foi calculada de acordo com a expressão (2.12), apresentada no Capítulo 2, porém tomando-se como referência as soluções fornecidas pela função `lambertw` do MATLAB, em vez de medições. Vale mencionar que a sua adoção, em complemento ao ARE, se deve ao fato de esta ser a principal figura de mérito utilizada em estudos relacionados ao desempenho do modelo de um diodo. Por fim, o tempo médio total de simulação foi obtido aplicando-se cada método de cálculo da função W de

Figura 15 – Boxplot dos valores de ARE verificados nas soluções retornadas pelo método proposto.



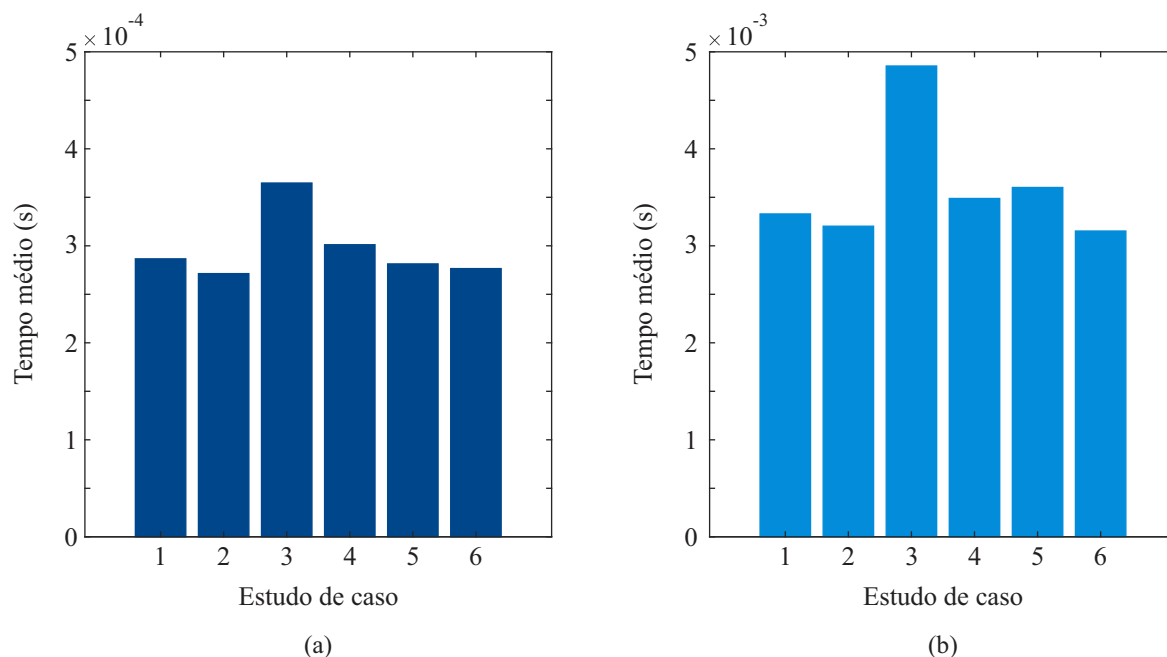
Lambert 100 vezes no âmbito de cada estudo de caso, tal como praticado em [44]. As soluções retornadas pelos métodos e, conseqüentemente, pelo modelo de um diodo, foram idênticas em todas essas repetidas execuções, uma vez que decorrem de expressões determinísticas e não envolvem qualquer componente de natureza probabilística. Em contrapartida, os tempos de execução variaram ligeiramente em razão de instabilidades inerentes ao computador utilizado. Logo, a média desses tempos foi registrada como métrica mais consistente para avaliação. A seguir, apresenta-se uma síntese dessas análises.

4.3.1 Avaliação de efetividade

A Figura 15 apresenta uma visão estatística dos valores de ARE verificados para as soluções encontradas pelo método proposto em cada estudo de caso. A sua análise revela que todos esses valores situam-se na almejada ordem de 10^{-16} . Observa-se ainda que o terceiro quartil permanece abaixo de 2×10^{-16} em todos os casos, o que significa que ao menos 75% das soluções retornadas pelo método proposto resultaram em erros inferiores à própria precisão de máquina ($\varepsilon_{\text{maq}} \approx 2,2 \times 10^{-16}$). Além disso, no caso 6, a mediana praticamente coincide com o valor mínimo observado, evidenciando uma robustez particularmente destacada.

Adicionalmente, a Figura 16(a) mostra o tempo médio total gasto pelo método proposto para calcular todos os valores requeridos da função W de Lambert em cada estudo de caso. Nota-se que os tempos envolvidos são bastante reduzidos. Considerando que cada caso compreende 1000 cálculos dessa função, infere-se que o método proposto necessitou de aproximadamente 10^{-7} segundos para determinar um único valor de $W_0(z)$.

Figura 16 – Tempos médios totais de simulação referentes aos cálculos requeridos da função W de Lambert.



De forma complementar, a Figura 16(b) apresenta os tempo médios totais despendidos para realizar todos os cálculos da equação característica do modelo de um diodo utilizando as soluções retornadas pelo método proposto. Como esperado, esses tempos também apresentam ordem de grandeza pouco expressiva.

Esses resultados evidenciam que o método proposto cumpre efetivamente o seu objetivo e, portanto, retorna soluções com nível de acurácia compatível com a precisão de máquina e com custo computacional reduzido, no intervalo do ramo principal da função W de Lambert relevante ao modelo de um diodo.

4.3.2 Avaliação de pertinência

Uma vez constatado que o método proposto cumpre o seu objetivo, torna-se necessário avaliar a qualidade de seu desempenho em relação aos métodos recursivos selecionados para a sua concepção. Caso contrário, sua materialidade pode ser colocada em dúvida.

Nesse sentido, a Tabela 12 apresenta as principais métricas estatísticas dos valores de ARE verificados para as soluções retornadas pelo método proposto, bem como para: (i) o método de segunda ordem apresentado em [155], cuja expressão coincide com o passo predictor do método proposto, isto é, (4.23b); e (ii) o método de quarta ordem concebido em [154], cujo equacionamento tem o mesmo cerne do passo corretor, isto é, (4.23c) e (4.23d). Ressalta-se que, para estabelecer uma comparação justa, as estimativas iniciais exigidas por esses métodos recursivos foram fornecidas pelo passo inicial do método proposto, isto é, (4.23a). Os melhores resultados comparativos encontram-se destacados em **negrito**.

Tabela 12 – Métricas estatísticas dos valores de ARE verificados para o método proposto e para os métodos recursivos de segunda e quarta ordem.

Caso	Alternativa	Mínimo	Mediana	Média	Máximo	Desvio-padrão
1	Método de segunda ordem	0,0000E+00	1,4775E-16	1,5055E-16	6,7588E-16	1,3323E-16
	Método de quarta ordem	0,0000E+00	1,2806E-16	1,2080E-16	6,0963E-16	1,2078E-16
	Método proposto	0,0000E+00	1,2969E-16	1,2201E-16	6,4029E-16	1,2243E-16
2	Método de segunda ordem	0,0000E+00	1,3150E-16	1,2484E-16	6,3678E-16	1,2603E-16
	Método de quarta ordem	0,0000E+00	1,1469E-16	1,0327E-16	5,9449E-16	1,1890E-16
	Método proposto	0,0000E+00	1,1513E-16	1,0075E-16	5,4323E-16	1,1322E-16
3	Método de segunda ordem	0,0000E+00	1,3798E-16	1,3842E-16	6,6292E-16	1,3391E-16
	Método de quarta ordem	0,0000E+00	1,3094E-16	1,2054E-16	6,6292E-16	1,2025E-16
	Método proposto	0,0000E+00	1,2842E-16	1,1792E-16	6,6292E-16	1,1994E-16
4	Método de segunda ordem	0,0000E+00	1,4316E-16	1,4175E-16	6,4520E-16	1,3176E-16
	Método de quarta ordem	0,0000E+00	1,2220E-16	1,1311E-16	5,9359E-16	1,2184E-16
	Método proposto	0,0000E+00	1,1777E-16	1,0746E-16	5,9359E-16	1,2037E-16
5	Método de segunda ordem	0,0000E+00	1,3760E-16	1,3456E-16	8,1948E-16	1,2944E-16
	Método de quarta ordem	0,0000E+00	1,2151E-16	1,0782E-16	6,5558E-16	1,1437E-16
	Método proposto	0,0000E+00	1,1742E-16	1,0302E-16	6,5558E-16	1,1216E-16
6	Método de segunda ordem	0,0000E+00	0,0000E+00	5,0508E-17	4,4442E-16	8,3650E-17
	Método de quarta ordem	0,0000E+00	0,0000E+00	5,8279E-17	2,3048E-16	8,6816E-17
	Método proposto	0,0000E+00	0,0000E+00	6,1373E-17	2,2613E-16	8,8500E-17

Tabela 13 – Comparação dos tempos médios totais de simulação nos cálculos da função W de Lambert.

Alternativa	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Método de segunda ordem	2,0759	2,2258	2,1787	2,6684	2,1408	2,5291
Método de quarta ordem	2,2588	2,4139	2,4038	2,4657	2,2808	2,5302
Método proposto	1	1	1	1	1	1

Os resultados apresentados na Tabela 12 mostram que os valores de ARE do método proposto são essencialmente equivalentes àqueles verificados para os métodos recursivos. Notadamente, este preserva o elevado nível de acurácia, robustez e estabilidade inerente aos métodos de segunda e quarta ordem.

Por sua vez, a Tabela 13 apresenta uma comparação qualitativa do tempo médio total de simulação gasto pelo método proposto para calcular todos os valores da função W de Lambert requeridos em cada estudo de caso, frente aos tempos registrados para os métodos recursivos de segunda e quarta ordem. Mais precisamente, o tempo médio total verificado para cada alternativa foi normalizado em relação ao tempo correspondente do método proposto (veja a Figura 16(a)).

Os resultados indicados na Tabela 13 revelam que o método proposto é, no mínimo, duas vezes mais rápido que os métodos recursivos de segunda e quarta ordem. Assim, sua estrutura “preditor–corretor” confirma-se não apenas mais eficiente, mas também adequada ao propósito de aprimorar a avaliação numérica da função W de Lambert no escopo do modelo de um diodo.

4.3.3 Avaliação de competitividade

Diante do êxito de desenvolvimento do método proposto, torna-se interessante avaliar a sua competitividade frente a outras alternativas destinadas ao mesmo propósito. Nesse contexto,

Tabela 14 – Valores de RMSE, dados em ampères, verificados para os métodos proposto e correlatos.

Alternativa	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Método híbrido	7,0380E-04	1,1447E-05	1,8836E-05	7,7203E-06	1,0772E-05	3,9930E-11
Método explícito	4,8794E-04	1,0502E-05	9,0955E-06	2,1375E-07	8,2638E-07	8,3943E-16
Método numérico	6,9822E-16	3,3307E-17	3,1512E-16	1,8073E-17	2,3288E-17	5,3989E-16
Método proposto	6,7056E-16	2,9997E-17	2,8375E-16	1,5799E-17	2,0396E-17	4,6555E-16

destacam-se os métodos especializados discutidos na Seção 4.1, a saber: o método híbrido formalizado em [41] e os métodos explícito e numérico apresentados em [42].

A Tabela 14 apresenta os valores de RMSE verificados para o método proposto em cada estudo de caso, bem como para os métodos híbrido, explícito e numérico. Vale frisar que esses resultados decorrem da aplicação da equação característica do modelo de um diodo com as soluções respectivamente retornadas para a função W de Lambert por cada um desses métodos.

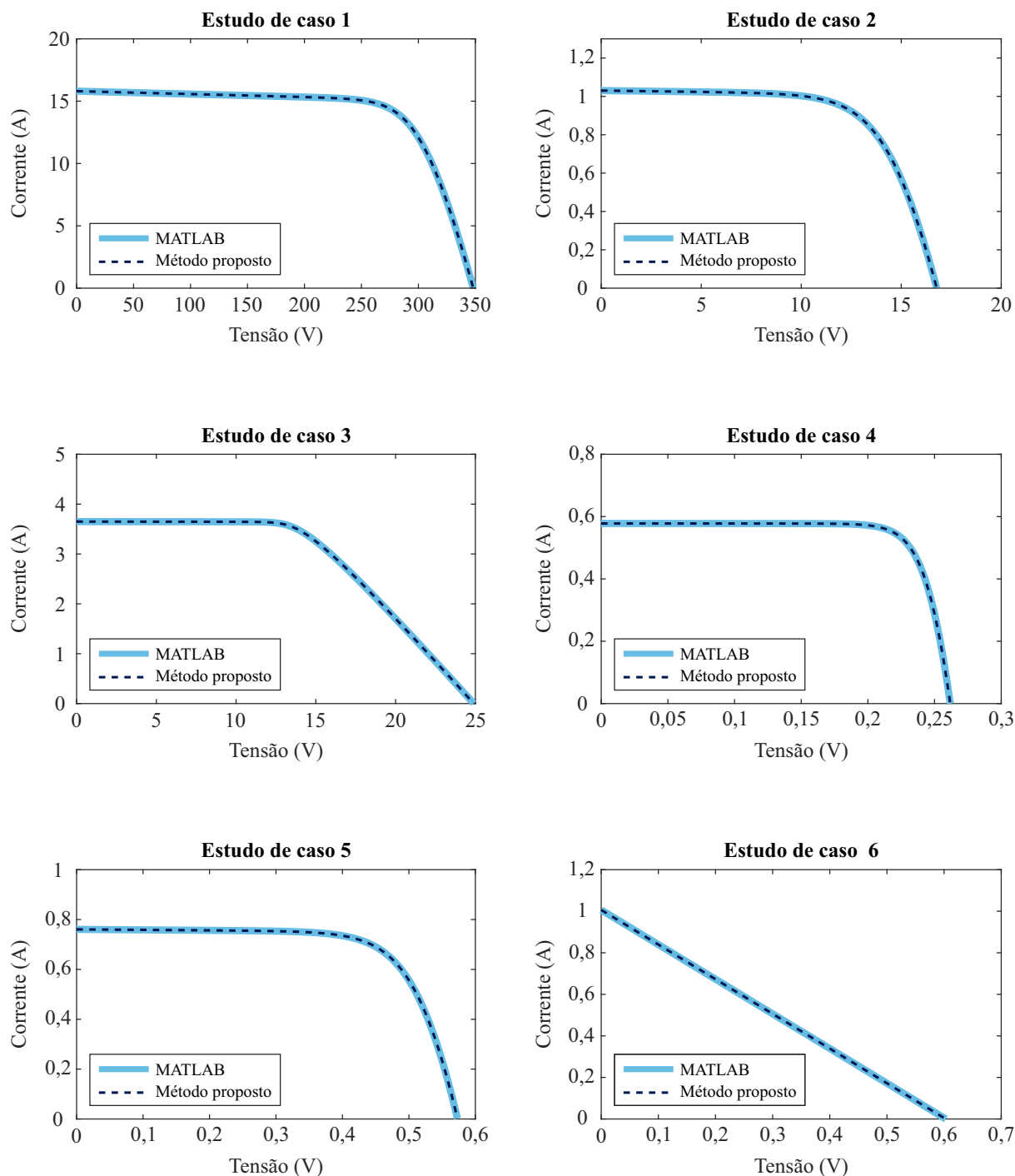
Os resultados da Tabela 14 revelam que os valores de RMSE verificados para o método proposto são iguais ou inferiores à ordem de grandeza da precisão de máquina em todos os estudos de caso. Consequentemente, observa-se na Figura 17 que as curvas $I-V$ geradas a partir de suas soluções coincidem com aquelas obtidas por meio da função `lambertw` do MATLAB. Esses resultados corroboram, portanto, que o método proposto pode ser confiavelmente aplicado no escopo do modelo de um diodo.

No que diz respeito aos demais métodos, os valores de RMSE apresentados na Tabela 14 indicam que o método proposto supera expressivamente os métodos não iterativos. Por exemplo, os valores de RMSE do método híbrido não atingem a ordem de 10^{-12} , e os do método explícito ultrapassam essa ordem de grandeza em apenas um estudo de caso. Ademais, verifica-se que as soluções retornadas pelo método proposto são comparáveis àquelas obtidas pelo método numérico, o qual se destoa por ser iterativo.

Adicionalmente, a Figura 18 apresenta os perfis de ARE verificados para as soluções retornadas pelos métodos considerados ao longo dos pontos da curva $I-V$ de cada estudo de caso. Nota-se que os valores de ARE associados ao método híbrido são significativamente superiores aos dos demais métodos. Em relação ao método explícito, observa-se que, nos casos 1–5, seus valores de ARE exibem oscilações que culminam no aumento de algumas ordens de grandeza a partir de determinados valores de tensão e , e, consequentemente, de argumento z . No caso 6, entretanto, seus valores de ARE aproximam-se daqueles verificados para os métodos numérico e proposto. Sobre esses dois últimos, constata-se que seus valores de ARE permanecem consistentemente em ordem de grandeza compatível com a precisão de máquina. Desse modo, esses resultados substanciam as conclusões inferidas com base nos valores de RMSE apresentados na Tabela 14.

Por sua vez, a Tabela 15 apresenta uma comparação qualitativa entre os tempos médios totais de simulação gastos pelo método proposto para realizar todos os cálculos requeridos da equação característica do modelo de um diodo e aqueles registrados para os métodos correlatos e

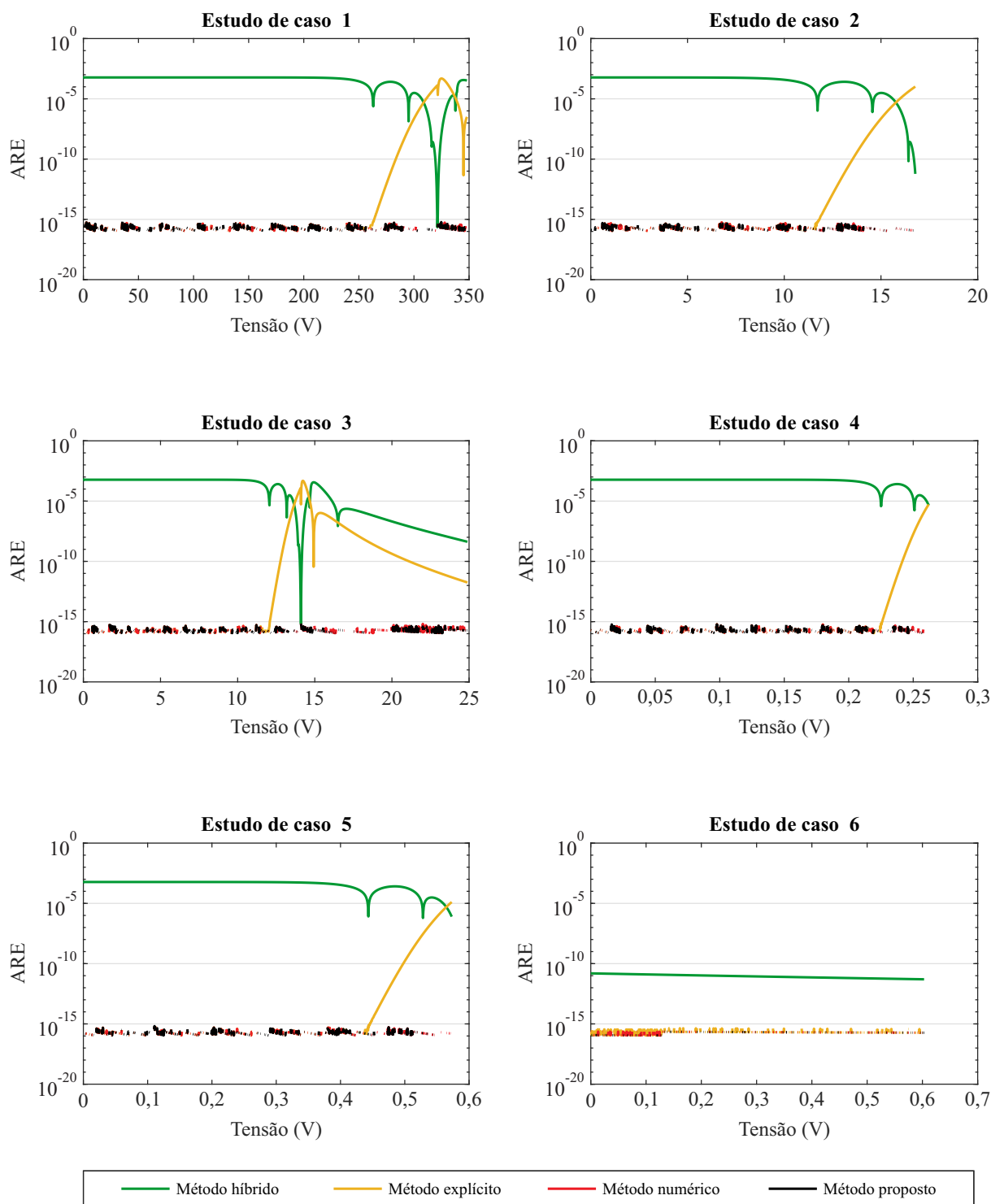
Figura 17 – Comparação entre as curvas $I-V$ geradas em cada estudo de caso pela função `lambertw` do MATLAB e pelo método proposto.



para a função `lambertw` do MATLAB. Conforme adotado anteriormente, os tempos médios totais foram normalizados em relação ao tempo do método proposto (veja a Figura 16(b)).

Os resultados indicados na Tabela 15 mostram que o método proposto é até cerca de duas vezes mais rápido que o método numérico e tende a apresentar tempos de simulação semelhantes aos registrados para os métodos explícito e híbrido, podendo, contudo, ser aproximadamente 40% inferior a estes. Além disso, verifica-se que o método proposto é cerca de três ordens de grandeza mais rápido que a função `lambertw` do MATLAB.

Figura 18 – Comparação dos valores de ARE verificados para os métodos considerados na curva $I-V$ de cada estudo de caso.



Diante dessas constatações, é possível afirmar que o método proposto não apenas demonstra competitividade, mas configura uma alternativa particularmente atrativa para o cálculo da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo. Em primeiro lugar, porque apresenta acurácia equivalente àquela da função `lambertw` do MATLAB, mas com eficiência significativamente superior. Em segundo lugar, porque se mostra consistentemente capaz de superar o desempenho de outros métodos concebidos para o mesmo propósito.

Tabela 15 – Comparação dos tempos médios totais de simulação registrados nos cálculos requeridos da equação característica do modelo de um diodo.

Alternativa	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
MATLAB (<code>lambertw</code>)	6021,1989	6668,1898	4090,1743	5461,3339	5948,2197	7022,3943
Método híbrido	1,3247	1,3835	1,2393	1,7322	1,2252	1,3628
Método explícito	1,0515	1,1123	0,9915	1,1319	1,0796	1,0069
Método numérico	1,2382	1,4121	1,2985	2,0333	1,1757	1,3144
Método proposto	1	1	1	1	1	1

4.4 CONCLUSÕES PARCIAIS

Este capítulo apresentou uma prospecção metodológica acerca das principais alternativas disponíveis para o cálculo da função W de Lambert no domínio relevante ao modelo de um diodo. Para além de métodos de uso geral, foi evidenciada uma vertente recente dedicada ao desenvolvimento de métodos especializados a esse escopo, motivada pelo fato de que *solvers* consagrados — como a função `lambertw` do MATLAB — embora reconhecidos pela elevada acurácia, podem impor tempos de simulação proibitivos em aplicações intensivas, como ocorre no problema de extração dos parâmetros.

Nesse contexto, foram discutidos três métodos especializados recorrentes na literatura e foram explicitados seus principais potenciais e limitações: enquanto os métodos híbrido e explícito tendem a reduzir o tempo de cálculo à custa de soluções de menor qualidade, o método numérico pode atingir soluções mais acuradas, porém com menor eficiência. Adicionalmente, foi observado como inconveniente estrutural desses métodos a dependência de lógica condicional, que conduz a conjuntos distintos de expressões e a custos de processamento menos uniformes.

À luz dessas constatações, foi então concebido o chamado *método preditor–corretor autoiniciado*. Tomando como base as experiências acumuladas na literatura, a sua formulação foi construída a partir de uma aproximação inicial de baixa complexidade para a função W de Lambert, a qual é aprimorada por meio de um encadeamento sequencial de passos preditor e corretor. Como resultado, foi obtido um procedimento não iterativo e não condicional, caracterizado por uma única passagem por quatro equações, capaz de retornar soluções de alta qualidade com baixo custo computacional. Em complemento, sua implementação e sua aplicação direta na abordagem formal do modelo de um diodo foram sistematizadas por meio do Algoritmo 1 e do Algoritmo 2.

A validação desse método proposto foi realizada com base em seis estudos de caso compatíveis com a literatura recente. Os resultados obtidos evidenciaram que o procedimento desenvolvido é capaz de retornar valores para a função W de Lambert com erro relativo da ordem da precisão de máquina no domínio fisicamente relevante ao modelo de um diodo, a um custo computacional reduzido. Consequentemente, as curvas I – V geradas a partir da equação característica do modelo mostraram-se coerentes e representativas do comportamento esperado, confirmando a pertinência da abordagem adotada.

Além disso, foi constatado que tal método é não apenas competitivo, mas também particularmente atraente frente a alternativas de estado da arte: por um lado, apresentou acurácia equivalente à da função `lambertw` do MATLAB, porém com tempos de execução cerca de três ordens de grandeza inferiores; por outro, superou de forma consistente o desempenho dos métodos híbrido, numérico e explícito, os quais foram concebidos para o mesmo propósito.

Diante do exposto, o *método preditor–corretor autoiniciado* consolida-se como uma contribuição original desta tese, tanto pela inovação associada à sua formulação matemática quanto pelo compromisso alcançado entre acurácia, robustez e eficiência computacional no domínio fisicamente relevante do modelo de um diodo. Em decorrência dessas características, é estabelecido um suporte técnico direto para a adoção sistemática da abordagem formal do modelo de um diodo via função W de Lambert no âmbito do problema de extração dos parâmetros. Registra-se, ainda, que a pertinência e a materialidade técnica dessa contribuição são corroboradas por sua publicação em [185].

5 MÉTODO PROPOSTO PARA A RESOLUÇÃO DO PROBLEMA

Este capítulo documenta o método determinístico local proposto para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico. Inicialmente, a Seção 5.1 sistematiza a prospecção metodológica realizada e o encadeamento de critérios que conduziu à adoção do paradigma de método de barreira considerado mais promissor. Na sequência, a Seção 5.2 formaliza a arquitetura do método proposto para a resolução acurada e robusta do problema, descrevendo sua estrutura lógica, os mecanismos incorporados para tratar as dificuldades intrínsecas do problema, bem como detalhes de implementação computacional. A Seção 5.3, por sua vez, reporta os testes realizados para validação do método proposto e sintetiza os resultados obtidos frente às alternativas do estado da arte. Por fim, a Seção 5.4 sumariza o conteúdo exposto ao longo do capítulo, explicitando a originalidade e a contribuição técnica do método proposto no atendimento ao propósito delineado.

5.1 PROSPECÇÃO METODOLÓGICA

Conforme estabelecido no Capítulo 3, o problema de extração dos parâmetros, formulado conforme (3.15), constitui um problema de otimização multidimensional restrito, de natureza não linear, (regularmente) suave e, em geral, não convexo. Ademais, as discussões realizadas deixaram claro que se trata de um problema mal escalonado e mal condicionado. Apesar de reunir atributos tão intrincados, é possível identificar na literatura distintos métodos de otimização capazes de resolvê-lo. Nesse contexto, esta seção sintetiza as avaliações metodológicas conduzidas acerca das alternativas determinísticas de busca local disponíveis, culminando na identificação do paradigma considerado mais promissor para embasar a proposição de um novo método de resolução.

5.1.1 Avaliação preliminar

Assim como introduzido na Subseção 2.4.1 do Capítulo 2, a maioria dos métodos determinísticos disponíveis na literatura se caracteriza por estratégias de busca local da solução ótima. Dentre estes, os métodos baseados em gradiente — isto é, aqueles que exploram informações relacionadas às derivadas da função objetivo e das restrições existentes para produzir novas candidatas à solução — figuram como os mais eficientes e teoricamente rigorosos.

Na literatura específica, é possível identificar duas categorias de métodos determinísticos locais baseados em gradiente que se mostram elegíveis para a resolução do problema de extração dos parâmetros. A primeira diz respeito àqueles desenvolvidos ou implementados de modo mais particular para a resolução de problemas de mínimos quadrados [113]. Em termos gerais, esses métodos buscam gerar novas candidatas à solução observando essencialmente o comportamento analítico da função objetivo [186, 187]. Apenas quando necessário, ou quando explicitamente previsto, estes aplicam correções numéricas para considerar o efeito das restrições de canali-

zação [186, 187]. Entre os representantes mais notórios dessa categoria, mencionam-se os já apresentados *algoritmo de Levenberg–Marquardt* e *método de região de confiança reflexiva*, isto é, o LMA e o TRR [186, 187].

A segunda categoria concerne a métodos concebidos com um escopo mais amplo, isto é, voltados à resolução de problemas não lineares restritos, independentemente da forma de construção da função objetivo [113]. Em linhas gerais, esses métodos se pautam na *função Lagrangiana* associada ao problema em questão ou em alguma versão devidamente transformada desta [113, 116, 124]. Em última instância, as condições de primeira ordem associadas a essa função materializam parte das condições de KKT do problema original (ver Subseção 3.2.3 do Capítulo 3) [113, 116, 124]. Dessa forma, estratégias baseadas no *método de Newton* costumam ser internamente empregadas para resolver um sistema de equações particular que, de algum modo, incorpora essas condições [113, 116, 124]. Nessa lógica, as novas candidatas à solução são então geradas considerando intrinsecamente o efeito das restrições do problema [113]. Como representantes dessa categoria, mencionam-se os outrora apresentados *método de programação quadrática sequencial* e *método de barreira*, isto é, SQP e BM [113, 116, 124].

A esta altura, torna-se pertinente chamar a atenção para alguns fatos. Conforme revelado pelo levantamento bibliográfico documentado na Subseção 2.4.1, tanto o LMA quanto o TRR já foram testados no âmbito do problema de extração dos parâmetros. Embora isso tenha sido feito desconsiderando a existência das restrições de canalização do problema — o que, conforme discutido no Capítulo 3, constitui uma escolha profundamente sensível e não recomendável —, a análise apresentada na Seção 2.5 revela que estes apresentaram relativo sucesso. Em contrapartida, nenhum método pertencente à segunda categoria foi, até onde se tem conhecimento, publicamente avaliado para a resolução desse problema. Consequentemente, não há comprovação formal, ou mesmo indícios, quanto à efetividade e ao nível de eficiência que ao menos um de seus representantes poderia apresentar quando aplicado a esse fim. Assim, torna-se evidente que a adoção de um SQP ou de um BM como base para a proposição de um novo método determinístico local para a resolução do problema de extração dos parâmetros, levando-se em conta a existência de suas restrições de canalização, constitui uma alternativa que oferece o maior potencial de inovação e de contribuição para o enriquecimento do tema.

A escolha entre um SQP e um BM, por sua vez, não é tão trivial. Em princípio, ambos configuram o estado da arte de métodos determinísticos locais aplicáveis à otimização não linear restrita [113, 116]. Ainda assim, uma análise mais aprofundada de suas características revela que o BM apresenta afinidade particularmente pronunciada com problemas de otimização cujas restrições se caracterizam por desigualdades, como é o caso das restrições de canalização do problema de extração dos parâmetros [113, 116]. Ademais, observa-se que um número excepcionalmente elevado de variantes desse método já foi proposto com sucesso para a resolução de problemas de otimização com natureza não-convexa, a saber: [125], [126], [127], [128], [129] e [130].

Dada essa conjuntura, a investigação acerca da aplicação de um BM para a resolução do problema de extração dos parâmetros, considerando a existência de suas restrições de canalização, materializa-se como uma alternativa não apenas apropriada, mas também promissora.

5.1.2 Métodos de barreira

A ideia por trás dos métodos de barreira ou, simplesmente, BMs, reside na aplicação de uma função especial às restrições de desigualdade do problema em questão [113, 124]. Em princípio, essa função apresenta propriedades capazes de forçar numericamente o atendimento dessas restrições, fornecendo assim um mecanismo para a construção de um procedimento de otimização [113, 124]. Não obstante, diferentes características e formas de emprego podem ser consideradas para essa função, dando origem a paradigmas distintos e, conseqüentemente, a linhagens heterogêneas de BMs [113, 116, 124, 188].

O paradigma mais tradicional de BMs remonta aos *métodos de penalidade*, desenvolvidos há aproximadamente cinco décadas [113, 116]. Atualmente, o representante mais refinado e amplamente reconhecido dessa linhagem é o chamado *método de ponto interior* (“*interior-point method*”, IPM) [113, 116]. Para se ter uma perspectiva, os métodos de otimização implementados no núcleo de *solvers* proeminentes, como LOQO, IPOPT e a função `fmincon` do MATLAB, são variantes respectivamente personalizadas do IPM [128–130, 189].

Para resolver um problema de otimização como o problema de extração dos parâmetros, o IPM moderno exige inicialmente a conversão de suas restrições de desigualdade em restrições de igualdade [113]. Tipicamente, essa tarefa é realizada por meio da simples introdução de *variáveis de folga* no problema [113]. Contudo, essas variáveis somente são efetivas quando assumem valores não negativos [113, 116]. Assim, em vez de se considerar novas restrições de desigualdade para impor essa condição, essas variáveis são convenientemente inseridas em funções logarítmicas e incorporadas à função objetivo, tal como *penalidades* [113]. Como resultado, um problema dado como (3.15) assume a forma expressa em (5.1), em que o escalar real $\mu \in [0, \infty)$ representa o chamado *parâmetro barreira*; $\mathbf{s} \in \mathbb{R}^{10}$ corresponde ao vetor de variáveis de folga; e $\mathbf{e} \in \mathbb{R}^{10}$ é um vetor auxiliar cujos elementos são iguais a um, isto é, $\mathbf{e} = [1, \dots, 1]^\top$.

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^5, \mathbf{s} \in \mathbb{R}^{10}}{\text{minimizar}} && \mathcal{F}(\mathbf{x}) - \mu \mathbf{e}^\top \ln(\mathbf{s}) \\ & \text{sujeito a} && \mathbf{c}(\mathbf{x}) - \mathbf{s} = \mathbf{0} \end{aligned} \tag{5.1}$$

Vale recordar que, no domínio real, a função logarítmica não admite argumentos negativos e, portanto, atua como uma *barreira* que propriamente emula o papel desempenhado por uma restrição de desigualdade. Além disso, essa característica garante que, dada uma candidata à solução localizada internamente no espaço de soluções viáveis do problema original, todas as demais candidatas subsequentemente geradas durante o processo iterativo permanecerão estritamente em seu *interior*.

Apesar de seu sucesso, o IPM apresenta algumas vulnerabilidades [113, 116, 124, 190]. Uma destas concerne à limitação intrínseca de sua barreira logarítmica [190]. Para elucidar essa questão, considere um problema cujo mínimo se localiza exatamente sobre uma das fronteiras do espaço de soluções viáveis, isto é, em que algum componente s_j de \mathbf{s} é nulo. Como $-\ln(s_j) \rightarrow \infty$ quando $s_j \rightarrow 0$, o IPM pode apenas se aproximar da solução ótima desse problema, e fazê-lo com elevado nível de acurácia tende a ser numericamente desafiador [190].

Outra vulnerabilidade diz respeito à sensibilidade à escala de sua barreira logarítmica [113, 116]. Em princípio, $\ln(s_j) < 0$ para $0 < s_j < 1$ e $\ln(s_j) > 0$ para $s_j > 1$. Logo, a depender dos valores assumidos pela componente s_j , é possível que, de uma iteração para outra, o termo de barreira deixe de representar uma penalização da função objetivo e passe, inadvertidamente, a atuar como um incentivo, comprometendo, em última instância, a eficácia do IPM.

Ao se trazer essas questões para o contexto do problema de extração dos parâmetros, algumas preocupações emergem. Como as suas restrições de desigualdade são essencialmente restrições de canalização, qualquer variável de folga introduzida pelo IPM será diretamente comparada a uma determinada variável de decisão e , conseqüentemente, ao respectivo parâmetro do modelo de um diodo. Dessa forma, não é irrazoável supor que, em um dado problema, o valor ótimo de um ou mais destes parâmetros efetivamente se situe em seu limite mínimo ou máximo admissível, sobretudo quando tal limite tiver sido rigorosamente definido. Resultados como aqueles apresentados em [54], inclusive, indicam que essa é uma situação possivelmente comum em relação ao parâmetro n . Adicionalmente, observa-se em trabalhos correlatos que os intervalos de valores admissíveis para alguns parâmetros, como I_{ph} e R_s , frequentemente permitem a transição entre valores decimais e unitários [43, 100, 101]. Portanto, essas situações lançam dúvidas quanto à estrita aplicabilidade e à robustez do IPM na resolução do problema de extração dos parâmetros.

É pertinente reconhecer que essas vulnerabilidades identificadas na barreira logarítmica clássica do IPM podem, em certos cenários, ser contornadas pela parte interessada na resolução do problema de extração dos parâmetros. Por exemplo, pode-se relaxar os limites das restrições de canalização ou adotar uma modelagem normalizada para o problema, como proposto em [64], [65] e [66]. No entanto, é relevante ter em mente que não há garantias de que tais iniciativas sejam efetivas ou mesmo convenientes. Cumpre ainda notar que, ao se propor um novo método para a resolução do problema de extração dos parâmetros, o ideal é que este seja capaz de lidar, de forma nativa, com as particularidades envolvidas, eliminando qualquer esforço adicional por parte do responsável pela sua aplicação.

Diante dessas considerações, outros paradigmas de BMs, embora menos difundidos, passam a figurar como alternativas potenciais [124, 188]. Dentre estes, o denominado *princípio do reescalonamento não linear* destaca-se por sua maturidade teórica e versatilidade [124]. Embora seu escopo seja amplo, é precisamente na formulação de uma *barreira logarítmica modificada* que esse paradigma encontra seu principal expoente [124–127, 191–194]. Uma das

grandes virtudes dessa barreira reside no fato de que esta preserva as melhores propriedades da barreira logarítmica clássica do IPM, ao mesmo tempo em que se encontra livre de suas vulnerabilidades [125, 191]. Conseqüentemente, os BMs desenvolvidos com base nesta compartilham similaridades estruturais e conceituais com o IPM, mas dispensam algumas de suas fragilidades fundamentais [124–126, 195]. Posto isto, uma explicação mais profunda acerca do princípio do reescalonamento não linear é apresentada na sequência.

5.1.3 Princípio do reescalonamento não linear

O princípio do reescalonamento não linear consiste em transformar um problema de otimização com restrições de desigualdade em um problema equivalente, porém dotado de características numéricas mais favoráveis [124, 195]. A rigor, essa tarefa é conduzida com o propósito explícito de utilizar a função Lagrangiana do problema transformado como instrumento para o desenvolvimento de métodos de otimização e para a realização de análises teóricas [195].

Assim como o próprio nome sugere, a transformação em questão é obtida pela aplicação de uma função não linear às restrições de desigualdade do problema original. Essa função é tipicamente escolhida a partir de uma classe Ψ de funções duas vezes continuamente diferenciáveis no intervalo (t_0, t_1) , com $-\infty \leq t_0 \leq 0 \leq t_1 \leq +\infty$ [195]. Em particular, funções $\psi \in \Psi$ satisfazem as propriedades listadas a seguir, para todo $t \in (t_0, t_1)$ ¹ [195]:

1. $\psi(0) = 0$;
2. $\psi'(t) > 0$;
3. $\psi'(0) = 1$;
4. $\psi''(t) < 0$;
5. existe $a > 0$ tal que $\psi(t) \leq -at^2$, para $t \leq 0$;
6. (a) $\psi'(t) \leq b_0 t^{-1}$; (b) $-\psi''(t) \leq b_1 t^{-2}$, para $t > 0$, com $b_0 > 0$ e $b_1 > 0$.

Em termos mais didáticos, é útil interpretar o argumento t como uma medida de *margem* associada a uma restrição: valores $t \geq 0$ representam satisfação (com possível folga), enquanto valores $t < 0$ representam violação. Nesse contexto, a propriedade (1) garante que a fronteira $t = 0$ não introduz qualquer viés artificial, uma vez que a transformação preserva o valor nulo na exata satisfação da restrição. Em seguida, a propriedade (2) assegura monotonicidade estrita, de modo que margens maiores resultem em valores maiores de $\psi(t)$, preservando a ordenação natural imposta pela restrição original. Por sua vez, a propriedade (3) normaliza a escala local

¹ Ressalta-se que o termo t utilizado nesta descrição das propriedades representa apenas uma variável de análise, não devendo ser confundido com o elemento dedicado à contagem de iterações. Apesar da inconveniência dessa coincidência, optou-se por prosseguir com tal denotação, em alinhamento com a representação típica da literatura.

da transformação, pois impõe que, nas proximidades da fronteira, pequenas variações de t sejam refletidas de maneira aproximadamente linear, isto é, sem amplificações ou atenuações artificiais. Complementarmente, a propriedade (4) estabelece concavidade estrita, promovendo um comportamento suave e controlado para $t > 0$ e contribuindo para a estabilidade numérica do problema transformado. Já a propriedade (5) é aquela que garante robustez no regime inviável: para $t \leq 0$, a desigualdade $\psi(t) \leq -at^2$ impõe uma penalização pelo menos quadrática, fazendo com que violações mais severas sejam desencorajadas de forma progressivamente mais intensa. Por fim, a propriedade (6) controla o comportamento assintótico no regime viável com folga: ao impor limites superiores para $\psi'(t)$ e para $-\psi''(t)$ quando $t > 0$, esta evita que a função cresça ou se deforme de maneira excessiva, mitigando efeitos de escalonamento e preservando a estabilidade numérica quando as restrições são amplamente satisfeitas.

Dessa forma, um problema dado como (3.15), assume a forma expressa em (5.2), em que $\boldsymbol{\psi}(\cdot)$ denota a aplicação componente a componente de uma função ψ às restrições de desigualdade. Note que, nessa formulação, não se faz necessário incorporar explicitamente um termo de barreira à função objetivo do problema original, como ocorre, por exemplo, em (5.1) no âmbito do IPM [125]. Em conjunto, as propriedades (1)–(6) fornecem um arcabouço que tende a ser suficiente para que o reescalamento promova uma convergência eficiente e estável no processo de otimização do problema transformado, rumo a uma solução consistente com a do problema original [124, 195].

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^3}{\text{minimizar}} && \mathcal{F}(\mathbf{x}) \\ & \text{sujeito a} && \boldsymbol{\psi}(\mathbf{c}(\mathbf{x})) \geq \mathbf{0} \end{aligned} \quad (5.2)$$

Alguns exemplos de funções ψ podem ser encontrados na *transformação exponencial* apresentada em [196], na *transformação hiperbólica* proposta por [191] e na *transformação logarítmica* introduzida também por [191]. Embora cada uma dessas funções possua seus próprios méritos, a última se destaca como a mais trabalhada na literatura [124–127, 192–194]. Sendo assim, (5.3) expressa uma de suas formas quando aplicada a uma componente c_j de \mathbf{c} , onde $\mu \in (0, \infty)$ denota o parâmetro barreira [126, 127, 188, 194].

$$\psi(c_j(\mathbf{x})) = \mu \ln \left(1 + \frac{c_j(\mathbf{x})}{\mu} \right) \quad (5.3)$$

À primeira vista, o fato de que a desigualdade $c_j(\mathbf{x}) \geq 0$ pode ser reformulada em termos de $\psi(c_j(\mathbf{x}))$ pode não parecer imediato. Entretanto, o seu processo de construção torna essa relação clara, pois a transformação $\psi(\cdot)$ é estritamente monótona no domínio de interesse, de modo que $c_j(\mathbf{x}) \geq 0$ e $\psi(c_j(\mathbf{x})) \geq 0$ descrevem o mesmo conjunto viável [124, 126, 195]. Nesse sentido, considere uma restrição de desigualdade dada por $c_j(\mathbf{x}) \geq 0$. Ao se adicionar o parâmetro barreira μ a ambos os lados, obtém-se (5.4).

$$\mu + c_j(\mathbf{x}) \geq \mu \quad (5.4)$$

Em seguida, ao se dividir ambos os lados de (5.4) por μ e, então, elevar ambos os lados à potência μ , chega-se a (5.5).

$$\left[1 + \frac{c_j(\mathbf{x})}{\mu}\right]^\mu \geq [1]^\mu \quad (5.5)$$

Por fim, ao se aplicar a função logarítmica natural a ambos os lados de (5.5) e realizar simplificações algébricas, encontra-se (5.6).

$$\mu \ln\left(1 + \frac{c_j(\mathbf{x})}{\mu}\right) \geq 0 \quad (5.6)$$

Desse modo, conclui-se que a desigualdade original $c_j(\mathbf{x}) \geq 0$ pode ser substituída, de maneira equivalente, pela condição $\psi(c_j(\mathbf{x})) \geq 0$, preservando-se o conjunto viável e explicitando-se o papel do parâmetro barreira μ na construção dessa reformulação.

Ainda sobre (5.3), é possível notar que sua expressão se assemelha à barreira logarítmica clássica do IPM, apresentada em (5.1). Por essa razão, esta é frequentemente denominada *barreira logarítmica modificada* [124–126, 188, 191, 192, 194]. Não obstante, é crucial perceber que tal barreira lida perfeitamente com situações em que $c(\mathbf{x}) = 0$, pois, conforme estabelece a propriedade (1) listada anteriormente, $\psi(c_j(\mathbf{x})) = \psi(0) = 0$. Ademais, para $c_j(\mathbf{x}) \geq 0$, independentemente de $0 < c_j(\mathbf{x}) < 1$ ou de $c_j(\mathbf{x}) > 1$, esta sempre retorna um valor não negativo. Portanto, a barreira logarítmica modificada é imune à limitação intrínseca e à sensibilidade de escala anteriormente identificadas na barreira logarítmica clássica do IPM.

Cabe ainda observar que a barreira logarítmica modificada permite certa relaxação das restrições de desigualdade. Mais especificamente, (5.3) torna aceitável qualquer $c_j(\mathbf{x}) \geq -\mu$. Essa situação pode ser devidamente constatada ao se dividir ambos os lados de (5.6) por μ e, na sequência, aplicar a função exponencial a ambos os lados da expressão obtida, simplificando-se, por fim, a expressão resultante. Desse modo, para um eventual $\mu > 0$, torna-se possível tolerar, de maneira criteriosa, situações em que $c_j(\mathbf{x}) < 0$. Em outras palavras, o processo de otimização pode, eventualmente, aventurar-se no *exterior* do espaço de soluções viáveis do problema original. Dada essa particularidade, os métodos desenvolvidos com base na barreira logarítmica modificada são frequentemente chamados de *métodos de ponto exterior* (“*exterior-point methods*”, EPM) [124, 125, 127].

À luz do exposto, é razoável concluir que a adoção de um EPM como base para a proposição de um novo método determinístico local para a resolução do problema de extração dos parâmetros constitui uma alternativa metodologicamente adequada. Em especial, as propriedades associadas à barreira logarítmica modificada sugerem maior robustez e eficiência no tratamento das restrições de canalização inerentes ao problema, sem incorrer nas fragilidades observadas em BMs mais tradicionais, como o IPM. Além disso, vale ressaltar que a literatura reporta aplicações bem-sucedidas de variantes de EPMs em problemas de otimização não convexos, o que reforça a plausibilidade dessa escolha [124–127].

5.2 MÉTODO PROPOSTO

A prospecção metodológica sintetizada na seção anterior culminou no EPM como base adequada para a proposição de um novo método determinístico local voltado à resolução do problema de extração dos parâmetros. A partir dessa diretiva, conduziu-se uma série de investigações conceituais e experimentos numéricos com o objetivo de se alcançar um algoritmo capaz de cumprir tal propósito com os níveis de acurácia requeridos pelo contexto de aplicação. Como resultado, desenvolveu-se uma alternativa original, na qual cada procedimento matemático foi especificamente delineado para lidar, de modo eficaz, com as nuances da natureza do problema e, ao mesmo tempo, assegurar um processo de otimização equilibrado e robusto.

Inicialmente, esta seção apresenta os quesitos fundamentais que caracterizam esse método propriamente desenvolvido, tais como a sua conceituação, estrutura lógica e formulações matemáticas associadas. Uma vez feito isso, racionaliza-se a sua arquitetura, consolida-se o seu algoritmo principal e discutem-se aspectos relevantes de sua implementação prática.

5.2.1 Problema transformado

Para resolver o problema de extração dos parâmetros, o EPM requer apenas o reescalonamento não linear de suas restrições de canalização. Ao realizar esse procedimento por meio da barreira logarítmica modificada, apresentada em (5.3), o problema enunciado em (3.15) transforma-se em (5.7). Ressalta-se que $\mathbf{e} \in \mathbb{R}^{10}$ denota um vetor auxiliar de uns, isto é, $\mathbf{e} = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^\top$.

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^5}{\text{minimizar}} && \mathcal{F}(\mathbf{x}) \\ & \text{sujeito a} && \mu \ln \left(\mathbf{e} + \frac{\mathbf{c}(\mathbf{x})}{\mu} \right) \geq \mathbf{0} \end{aligned} \quad (5.7)$$

Em consonância com o exposto na Subseção 3.2.2 do Capítulo 3, a função Lagrangiana associada a esse problema transformado pode ser construída designando-se multiplicadores de Lagrange às restrições reescaladas. No presente contexto, tais multiplicadores são definidos conforme (5.8). Note-se que $\boldsymbol{\delta}_l \in \mathbb{R}^5$ está vinculado às restrições de limite inferior, ao passo que $\boldsymbol{\delta}_u \in \mathbb{R}^5$ está vinculado às restrições de limite superior. Ademais, por conveniência, firma-se $\boldsymbol{\delta} \in \mathbb{R}^{10}$ como a concatenação desses dois últimos.

$$\boldsymbol{\delta}_l := [\delta_{l,I_{ph}}, \delta_{l,I_0}, \delta_{l,n}, \delta_{l,R_{sh}}, \delta_{l,R_s}] \quad (5.8a)$$

$$\boldsymbol{\delta}_u := [\delta_{u,I_{ph}}, \delta_{u,I_0}, \delta_{u,n}, \delta_{u,R_{sh}}, \delta_{u,R_s}] \quad (5.8b)$$

$$\boldsymbol{\delta} := [\boldsymbol{\delta}_l, \boldsymbol{\delta}_u]^\top \quad (5.8c)$$

Uma vez definidos esses multiplicadores, a função Lagrangiana $\mathcal{L}_{\text{epm}} : \mathbb{R}^5 \times \mathbb{R}^{10} \mapsto \mathbb{R}$ associada ao problema transformado assume a forma expressa em (5.9).

$$\mathcal{L}_{\text{epm}}(\mathbf{x}, \boldsymbol{\delta}; \mu) := \mathcal{F}(\mathbf{x}) - \boldsymbol{\delta}^\top \left[\mu \ln \left(\mathbf{e} + \frac{\mathbf{c}(\mathbf{x})}{\mu} \right) \right] \quad (5.9)$$

Cumprido destacar que a barreira logarítmica modificada presente em (5.9) estabelece intrinsecamente o subespaço primal viável do problema transformado, sendo este formalmente dado por: $\{\mathbf{x} \in \mathbb{R}^5 \mid \mathbf{c}(\mathbf{x}) \geq -\mu \mathbf{e}\}$ [124, 126]. Em complemento, vale memorar que o subespaço dual viável é, em princípio, dado por: $\{\boldsymbol{\delta} \in \mathbb{R}^{10} \mid \boldsymbol{\delta} \geq \mathbf{0}\}$ [113, 124].

5.2.2 Estrutura lógica

Uma vez que a barreira logarítmica modificada incorporada à função Lagrangiana define, por si só, o subespaço primal de soluções viáveis, a expressão (5.9) pode ser convenientemente interpretada como uma reformulação irrestrita de (5.7) [124, 195]. Nessa perspectiva, a sua minimização direta, desde que acompanhada de salvaguardas teóricas e numéricas, conduziria à solução ótima do problema transformado e, conseqüentemente, do problema original [124].

De fato, observa-se na literatura específica que esse raciocínio é exitosamente explorado na construção de uma parcela significativa dos algoritmos de EPM [124–127, 191, 194, 195]. Todavia, nota-se que, na abordagem mais tradicional, tais algoritmos se concentram na minimização de (5.9) apenas com respeito às variáveis primais, tomando as variáveis duais como elementos condutores do processo de otimização, assim como o parâmetro de barreira [124–127, 191].

A fim de tornar essa lógica explícita, apresenta-se na Figura 19 o fluxograma do algoritmo padrão do EPM. Embora conciso, verifica-se em trabalhos correlatos que há variações significativas no entendimento acerca de como cada uma de suas etapas deve ser conduzida [124–127, 191, 194, 195]. Dessa forma, descrevem-se a seguir os detalhes associados a cada uma destas, exaltando-se as construções matemáticas entendidas como sendo as mais apropriadas ao presente contexto, bem como introduzindo-se algumas formulações oportunamente concebidas no âmbito do EPM desenvolvido.

5.2.2.1 Otimização das variáveis primais

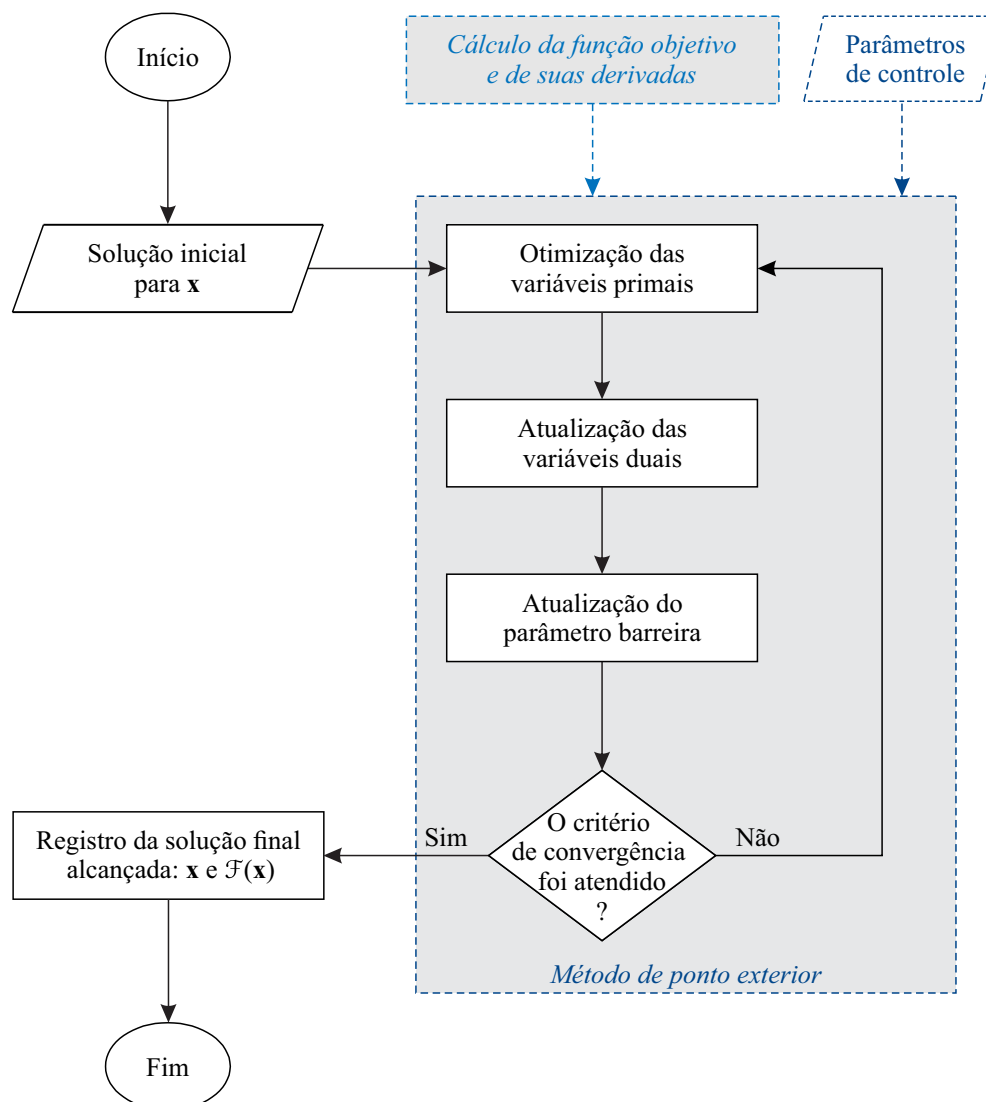
A princípio, considere que o EPM dispõe de uma candidata à solução não ótima, representada por $(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$. Ao longo de uma iteração t , o seu algoritmo padrão busca então gerar uma nova candidata mais acurada, a começar pelo ajuste das variáveis primais desta última conhecida. A rigor, tal tarefa é conduzida com base em (5.10) [124, 125, 127, 195].

$$\mathbf{x}^{t+1} = \arg \min \left\{ \mathcal{L}_{\text{epm}}(\mathbf{x}, \boldsymbol{\delta}^t; \mu^t) \mid \mathbf{x} \in \mathbb{R}^5 \right\} \quad (5.10)$$

É essencial notar que (5.10) induz um subproblema de otimização irrestrito, no qual \mathbf{x}^{t+1} representa as variáveis de decisão, enquanto $\boldsymbol{\delta}^t$ e μ^t são tratados como constantes [113, 124]. Em termos estritos, as condições de otimalidade necessárias de primeira ordem que se aplicam ao mesmo são as condições de estacionariedade do problema transformado, as quais são expressas de acordo com (5.11) [124, 126].

$$\nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^t; \mu^t) = \mathbf{g}(\mathbf{x}^{t+1}) - \mathbf{J}(\mathbf{x}^{t+1})^\top \frac{\mu^t \boldsymbol{\delta}^t}{\mu^t \mathbf{e} + \mathbf{c}(\mathbf{x}^{t+1})} = \mathbf{0} \quad (5.11)$$

Figura 19 – Fluxograma básico do método de ponto exterior.



Ressalta-se que $\mathbf{g} : \mathbb{R}^5 \mapsto \mathbb{R}^5$ denota o vetor gradiente da função objetivo, sendo dado conforme (3.12). Por sua vez, $\mathbf{J} : \mathbb{R}^5 \mapsto \mathbb{R}^{10 \times 5}$ denota a matriz Jacobiana das restrições, a qual assume a forma apresentada em (3.22).

Não obstante, como $\boldsymbol{\delta}^t$ e $\boldsymbol{\mu}^t$ tendem a assumir valores distantes do ótimo, sobretudo nas iterações iniciais, em geral não é vantajoso despender recursos na obtenção de uma solução exata para esse subproblema [113, 117, 124]. A experiência acumulada na literatura mostra que é suficiente encontrar uma candidata \mathbf{x}^{t+1} que, segundo algum critério de mérito devidamente estabelecido, seja superior àquela corrente dada em \mathbf{x}^t , garantindo o progresso consistente do processo iterativo do EPM [113, 124, 126, 195].

Os métodos empregados para esse propósito são tipicamente implantados sob um ar-

cabouço de globalização² [113, 116, 124]. Dentre as opções convencionais, a mais simples e amplamente adotada é conhecida como *busca em linha* [113, 116, 117, 124]. Como indicado em (5.12), o ajuste das variáveis primais é realizado, em seu âmbito, pela definição de uma *direção de busca*, estabelecida em $\mathbf{p} \in \mathbb{R}^5$, e de um *comprimento de passo* a ser tomado nessa direção, denotado pelo escalar real $\alpha \in (0, 1]$ [116].

$$\mathbf{x}^{t+1} = \mathbf{x}^t + \alpha \mathbf{p} \quad (5.12)$$

Uma vez que o subproblema em questão é de minimização, a direção de busca deve ser preferencialmente de descida, isto é, deve apontar para regiões do subespaço primal nas quais seja possível encontrar valores de \mathbf{x} associados a reduções efetivas da função Lagrangiana associada ao problema transformado [113, 124]. Todavia, ainda que essa direção seja apropriada, o ajuste das variáveis primais somente será eficaz se o comprimento de passo for determinado de maneira propícia. Em termos gerais, passos demasiadamente longos podem provocar a má exploração da direção de busca e elevar o valor da função Lagrangiana, enquanto passos excessivamente curtos podem resultar em progresso marginal e convergência impraticavelmente lenta [113, 116, 117]. Portanto, a efetividade da estratégia de busca em linha aplicada ao ajuste das variáveis primais depende do equilíbrio e do alinhamento conceitual entre os mecanismos aplicados na determinação desses dois termos [113, 116, 117, 124, 197, 198].

Nesse contexto, observa-se que diferentes abordagens do *método de Newton com busca em linha* figuram entre as construções matemáticas mais bem consolidadas na literatura específica [113, 116, 117, 124]. Em particular, o potencial de convergência quadrática e a elevada robustez frente a discrepâncias de escala nas variáveis envolvidas tornam essas abordagens especialmente atrativas [113, 116, 124]. É imperativo ponderar, entretanto, que tais abordagens nem sempre se mostram plenamente satisfatórias, uma vez que limitações de caráter algébrico e geométrico inerentes ao método de Newton podem colocar em xeque a sua aplicabilidade [113, 117, 199]. Ainda assim, a própria literatura indica que essas fragilidades podem ser mitigadas por mecanismos internos de escalonamento, regularização e globalização adequadamente concebidos, preservando-se a eficiência do método em cenários não ideais [113, 117, 199].

À luz dessas considerações e ciente da natureza intrincada do problema de extração dos parâmetros, propõe-se, no âmbito do EPM desenvolvido, um procedimento autoral para o ajuste das variáveis primais via (5.12). Como recurso primário, este emprega uma versão escalonada e regularizada do método de Newton, a qual foi particularmente concebida para operar de forma acurada e resiliente mesmo na presença de não convexidade local e de mal condicionamento numérico. A fim de não sobrecarregar a presente exposição da estrutura lógica do EPM, esclarece-se que os detalhes técnicos associados a esse procedimento serão apresentados na seção subsequente 5.2.3.

² Ressalta-se que o termo “globalização” refere-se, neste contexto, à ideia de tornar o procedimento robusto o bastante para convergir a um ponto de mínimo local mesmo quando iniciado a partir de candidatas de qualidade arbitrária [116].

5.2.2.2 Atualização das variáveis duais

Na literatura específica, observa-se haver um consenso de que a atualização das variáveis duais deve ser realizada por meio de uma equação especializada, a qual é dada, no presente contexto, por (5.13) [124–127, 191, 194, 195].

$$\boldsymbol{\delta}^{t+1}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^t; \mu^t) = \boldsymbol{\delta}^t \frac{\mu^t}{\mu^t \mathbf{e} + \mathbf{c}(\mathbf{x}^{t+1})} \quad (5.13)$$

A escolha dessa equação mostra-se interessante não apenas por sua baixa complexidade computacional, mas também pelo impacto que promove [124–127, 191, 194, 195]. Mais expressamente, ao se incorporar (5.13) no âmbito das condições de estacionariedade do problema transformado, apresentadas em (5.11), obtém-se (5.14).

$$\nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \mu^t) = \mathbf{g}(\mathbf{x}^{t+1}) - \mathbf{J}(\mathbf{x}^{t+1})^\top \boldsymbol{\delta}^{t+1}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^t; \mu^t) = \mathbf{0} \quad (5.14)$$

Ao analisar (5.14), constata-se que esta coincide estruturalmente com as condições de estacionariedade do problema original, enunciadas em (3.20a), no Capítulo 3. Portanto, (5.13) é capaz de induzir o processo iterativo à condição idealizada em (5.15), reforçando a compatibilidade da solução retornada pelo EPM com o problema original.

$$\nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \mu^t) \equiv \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0} \quad (5.15)$$

É pertinente mencionar que, em certas variantes do EPM, (5.13) é oportunamente manipulada e associada a (5.11), a fim de formar uma espécie de *sistema primal-dual* [124, 125, 195]. Essa abordagem — bastante comum no âmbito de métodos correlatos, como o IPM — propicia o cálculo de \mathbf{x}^{t+1} e $\boldsymbol{\delta}^{t+1}$ de modo simultâneo [113, 195]. Todavia, variantes como aquelas introduzidas nos trabalhos apresentados em [191], [126] e [194] demonstram que a atualização das variáveis primais e duais de modo desacoplado e sequencial também é efetiva. Como essa última abordagem dispensa o aumento tanto do escopo quanto dos esforços computacionais associados ao procedimento de otimização essencialmente requerido para o ajuste das variáveis primais, esta se mostra, a princípio, a alternativa mais atrativa.

Não obstante, entende-se que a atualização das variáveis duais de maneira segregada das variáveis primais exige, à nível algorítmico, certas ponderações acerca do emprego de (5.13). Em particular, o fato de $\boldsymbol{\delta}^{t+1}$ estar diretamente relacionada a $\boldsymbol{\delta}^t$ nessa equação abre espaço para a situação crítica evidenciada em (5.16). Ou seja, se algum componente δ_j de $\boldsymbol{\delta}$ assumir valor nulo em uma dada iteração, tal variável dual permanecerá indefinidamente presa a esse valor. Consequentemente, se a restrição c_j associada a δ_j se tornar ativa, isto é, $c_j(\mathbf{x}) = 0$, tal variável dual não será mais capaz de refletir a sua possível influência sobre a solução ótima do problema.

$$\delta_j^t = 0 \quad \therefore \quad \delta_j^{t+1} = 0 \quad \implies \quad \delta_j^{t+2} = \delta_j^{t+3} = \dots = 0 \quad (5.16)$$

Um fato que chama à atenção é que nenhum dos trabalhos citados como referências desta abordagem segregada de atualização das variáveis duais apresenta correções para essa questão

ou sequer a discute. Sendo assim, propõe-se, no âmbito do EPM desenvolvido, o procedimento autoral sintetizado em (5.17). Como pode ser observado, se $\delta_j^t = 0$ e a restrição associada estiver ativa na iteração corrente, então δ_j^{t+1} recebe um valor padrão não nulo, definido por $\delta_{\text{aux}} \in (0, \infty)$. Caso contrário, δ_j^{t+1} recebe o valor propriamente provido por (5.13).

$$\delta_j^{t+1} \leftarrow \begin{cases} \delta_{\text{aux}}, & \text{se } c_j(\mathbf{x}^{t+1}) = 0 \text{ e } \delta_j^t = 0 \\ \delta_j^t \frac{\mu^t}{\mu^t + c_j(\mathbf{x}^{t+1})}, & \text{caso contrário} \end{cases} \quad (5.17)$$

Cabe esclarecer que a referência a $c_j(\mathbf{x}^{t+1}) = 0$ foi adotada como condição, ao invés de $c_j(\mathbf{x}^{t+1}) = -\mu^t$ pois, como explicitado anteriormente, δ_j tende a se alinhar ao valor de λ_j que satisfaz a condição de estacionariedade do problema original — onde $c_j(\mathbf{x}^{t+1}) \geq 0$. Ademais, vale destacar que a perturbação imposta por δ_{aux} constitui um meio simples e suficiente de desvincular, na iteração corrente, a variável dual em questão do valor nulo, permitindo que esta assuma valores adequados nas iterações subseqüentes.

5.2.2.3 Atualização do parâmetro barreira

Em contraste com a etapa anterior, observa-se que não há, na literatura específica, um consenso acerca das formulações que devem ser efetivamente adotadas para realizar a atualização do parâmetro de barreira [124–127, 191, 194, 195]. Uma possível explicação para tal dispersão reside no fato de esse parâmetro desempenhar um papel preponderante no balanceamento da convergência do EPM e, por conseguinte, requerer valores ajustados às particularidades do algoritmo implementado [125–127, 194, 195]. Esse caráter sensível é ainda acentuado pelo fato de o parâmetro de barreira constituir um elemento estrutural comum a diversas linhagens de BMs, o que amplia substancialmente o repertório de regras e equações disponíveis [113, 124, 200].

Ainda assim, dentre as diferentes estratégias encontradas em trabalhos correlatos, chamam à atenção aquelas que se baseiam em alguma medida de complementaridade entre as restrições do problema e suas variáveis duais associadas [113, 200]. Em particular, tais estratégias buscam conferir sensibilidade à necessidade de ajuste do parâmetro de barreira na iteração corrente. No contexto mais amplo, esse tipo de estratégia adaptativa é reconhecido pela robustez que tende a oferecer [113, 200]. Ademais, os resultados apresentados em [125] e [126] corroboram a sua efetividade no âmbito do EPM.

Nesse sentido, uma das medidas de complementaridade mais populares é a expressa em (5.18) [126, 200]. Nessa equação, m_c representa o número de restrições do problema (sendo, portanto, igual a 10 no presente caso), enquanto $\sigma \in (0, 1]$ representa o chamado *parâmetro de centragem*.

$$\mu^{t+1} = \sigma \frac{\mathbf{c}(\mathbf{x}^{t+1})^\top \boldsymbol{\delta}^{t+1}}{m_c} \quad (5.18)$$

Uma das formulações para definição do parâmetro de centragem que reconhecidamente apresenta maior êxito prático é aquela implementada, de forma pioneira, no IPM do *solver*

LOQO [200]. Conforme expresso em (5.19), esta se baseia no desvio do menor produto de complementaridade $c_j(\mathbf{x}^{t+1}) \delta_j^{t+1}$ em relação à média, em que c_j denota o j -ésimo componente de \mathbf{c} e, analogamente, δ_j o j -ésimo componente de $\boldsymbol{\delta}^{t+1}$ [200].

$$\sigma = 0,1 \min \left(0,05 \frac{1-\varepsilon}{\varepsilon}, 2 \right)^3 \quad (5.19a)$$

$$\varepsilon = \frac{\min_j \left\{ c_j(\mathbf{x}^{t+1}) \delta_j^{t+1} \right\}}{\mathbf{c}(\mathbf{x}^{t+1})^\top \boldsymbol{\delta}^{t+1} / m_c} \quad (5.19b)$$

Em (5.19b), quando o parâmetro auxiliar $\varepsilon \in (0, 1]$ se aproxima de 1 (um), todos os produtos individuais se tornam próximos de sua média e, então, o parâmetro de barreira é reduzido de maneira mais agressiva. Vale notar, porém, que (5.19a) sempre retorna $\sigma \leq 0,8$. Desse modo, ainda que o valor do parâmetro de barreira possa aumentar de uma iteração para a subsequente, este nunca será escolhido acima do valor corrente de complementaridade, dado por $\mathbf{c}(\mathbf{x}^{t+1})^\top \boldsymbol{\delta}^{t+1} / m_c$. Portanto, essa construção controla de maneira equilibrada a sua evolução, promovendo reduções mais expressivas quando a complementaridade se encontra suficientemente equilibrada, o que contribui para maior robustez numérica, ao mesmo tempo em que permite aumentos moderados quando necessário para influenciar positivamente a convergência do algoritmo [113].

Não obstante, é crucial notar que o arcabouço formado por (5.18), (5.19a) e (5.19b), por ter tido seu desenvolvimento orientado ao IPM, permite a ocorrência de uma situação especialmente crítica ao EPM: $\mu^{t+1} = 0$. Mais expressamente, quando o parâmetro de barreira assume valor nulo, (5.3) colapsa para $0 \cdot \ln(\infty)$. Além disso, na eventualidade de haver alguma restrição ativa, isto é, $c_j(\mathbf{x}^{t+1}) = 0$, as derivadas da barreira logarítmica modificada associada passarão a ser dominadas pela razão $1/0$, causando uma série de problemas numéricos — repare em (5.11).

Para contornar essa questão, propõe-se, no âmbito do EPM desenvolvido, que a medida de complementariedade seja calculada conforme (5.20).

$$\mu^{t+1} = \max \left(\sigma \frac{\mathbf{c}(\mathbf{x}^{t+1})^\top \boldsymbol{\delta}^{t+1}}{m_c}, \mu_{\min} \right) \quad (5.20)$$

Em termos práticos, a definição de $\mu \in (\mu_{\min}, 1]$, com $0 < \mu_{\min} < \varepsilon_{\text{bar}}$, onde $\varepsilon_{\text{bar}} \in (0, 1)$ denota uma tolerância numérica (igual ou inferior àquelas dos critérios de convergência que serão apresentadas a seguir), garante que o parâmetro de barreira permaneça em uma faixa numericamente adequada, preservando a estabilidade e as propriedades de convergência do processo iterativo.

5.2.2.4 Critérios de convergência

Uma vez concluído o ajuste otimizado das variáveis primais mediante (5.12), a atualização das variáveis duais segundo (5.17) e a atualização do parâmetro de barreira conforme (5.19)

e (5.20), o EPM passa a dispor de uma nova candidata à solução, dada por $(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \boldsymbol{\mu}^{t+1})$. Resta então verificar se essa candidata atende, de modo satisfatório, aos requisitos técnicos do problema em questão. Para tanto, observa-se que dois critérios avaliativos são tipicamente estabelecidos: um *critério de viabilidade*, baseado na violação das restrições do problema original; e um *critério de otimalidade*, baseado nas condições de estacionariedade do problema transformado [124, 127, 194, 195].

No âmbito do EPM desenvolvido, adota-se como medida de viabilidade a norma infinita do conjunto de restrições do problema original que se encontram eventualmente violadas. Assim, o critério de viabilidade é dado por (5.21a), em que $\varepsilon_{\text{viab}} \in (0, 1)$ denota uma tolerância numérica associada [194]. Ressalta-se que $\mathbf{v} \in \mathbb{R}^{10}$ representa um vetor auxiliar, cujos componentes v_j se relacionam com os componentes c_j de \mathbf{c} segundo a regra estabelecida em (5.21b) [201].

$$\|\mathbf{v}(\mathbf{x}^{t+1})\|_{\infty} \leq \varepsilon_{\text{viab}} \quad (5.21a)$$

$$v_j(\mathbf{x}^{t+1}) = \min(c_j(\mathbf{x}^{t+1}), 0) \quad (5.21b)$$

Adicionalmente, adota-se como medida de otimalidade a norma infinita do resíduo das condições de estacionariedade do problema transformado. Assim, o critério de otimalidade é expresso por (5.22), em que $\varepsilon_{\text{otim}} \in (0, 1)$ denota uma tolerância numérica associada [124, 127, 195].

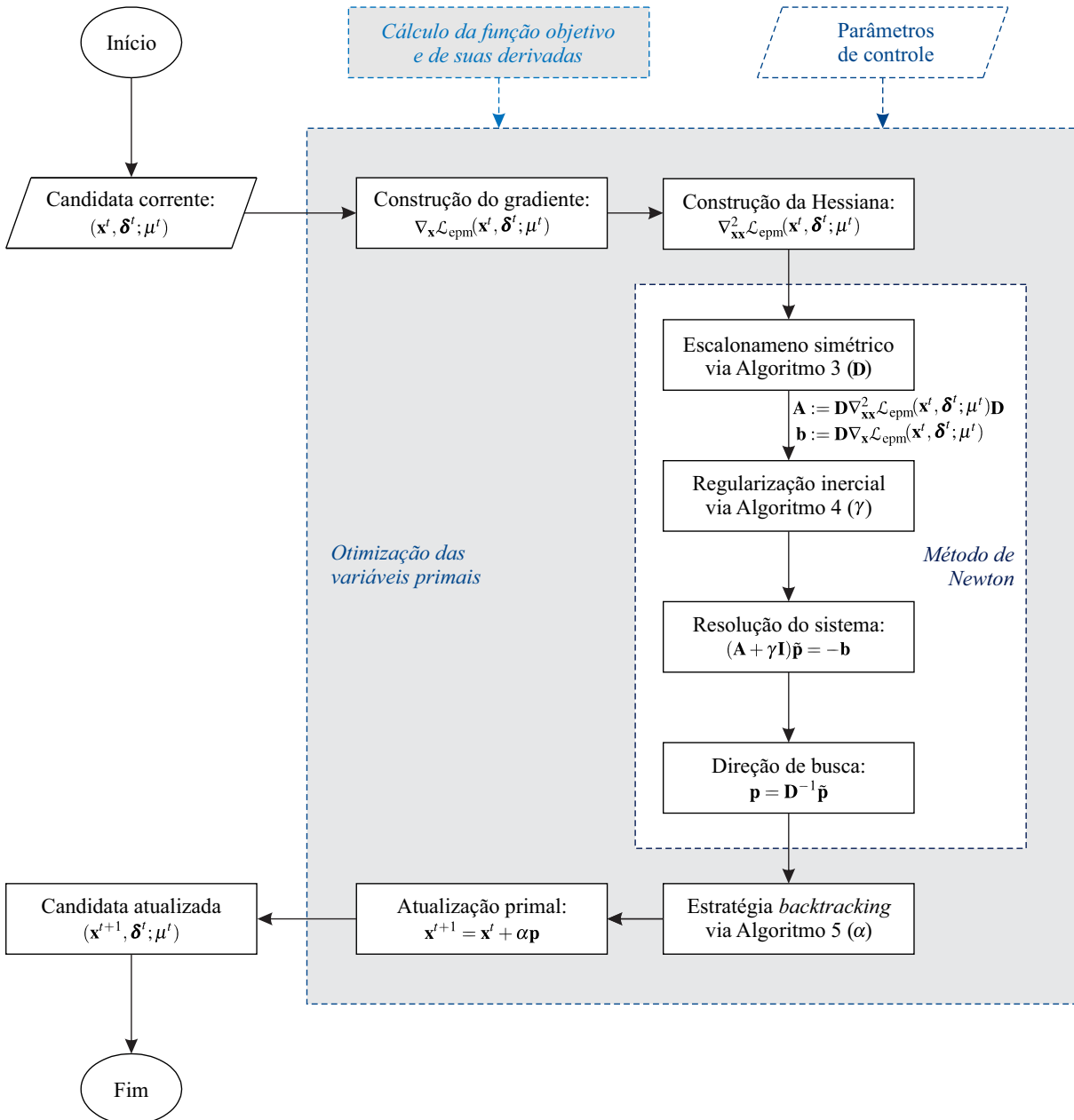
$$\left\| \nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \boldsymbol{\mu}^{t+1}) \right\|_{\infty} \leq \varepsilon_{\text{otim}} \quad (5.22)$$

Dessa forma, caso $(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \boldsymbol{\mu}^{t+1})$ satisfaça simultaneamente (5.21) e (5.22), esta pode ser declarada, na prática, como uma solução ótima local para o problema em questão, sendo o processo iterativo prontamente encerrado. Caso contrário, uma nova candidata deve ser buscada, tomando-se essa última obtida como base e reexecutando-se os procedimentos correspondentes às etapas anteriores ao longo de uma nova iteração.

5.2.3 Método de Newton escalonado e regularizado

Conforme introduzido na subseção anterior, o ajuste otimizado das variáveis primais é realizado, no âmbito do EPM desenvolvido, por meio de um arcabouço de busca em linha. Em particular, a direção de busca \mathbf{p} é obtida a partir de uma versão escalonada e regularizada do método de Newton aqui concebida. Por sua vez, o comprimento de passo α a ser tomado ao longo dessa direção é determinado seguindo as melhores práticas reportadas na literatura associada. Assim, as grandezas \mathbf{p} e α requeridas por (5.12) são determinadas de modo a produzir uma candidata à solução compatível com o problema transformado e capaz de promover progresso efetivo no processo iterativo. Na Figura 20, apresenta-se um panorama do encadeamento lógico que estrutura essa etapa de ajuste, bem como dos principais recursos incorporados à formulação do método Newton proposta. Na sequência, tais elementos são descritos em detalhe.

Figura 20 – Fluxograma de otimização das variáveis primais do método de ponto exterior desenvolvido.



5.2.3.1 Direção de busca

Ao analisar a estrutura das condições de estacionariedade do problema transformado, apresentadas em (5.11), constata-se que estas assumem a forma de um sistema de equações não lineares. Assim, ao se aplicar o *método de Newton–Raphson multivariado* na resolução desse sistema, encontra-se que a direção de busca \mathbf{p} pode ser diretamente determinada a partir da resolução do sistema linear estabelecido por (5.23) [116, 124].

$$\nabla_{xx}^2 \mathcal{L}_{epm}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \mathbf{p} = -\nabla_x \mathcal{L}_{epm}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \quad (5.23)$$

Note-se que o termo do lado direito de (5.23) é definido pelas derivadas parciais de

primeira ordem da função Lagrangiana do problema transformado ou, em outras palavras, pelo gradiente de (5.9) com respeito às variáveis primais. Dessa forma, este é dado por (5.24).

$$\nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) = \mathbf{g}(\mathbf{x}^t) - \mathbf{J}(\mathbf{x}^t)^\top \frac{\mu^t \boldsymbol{\delta}^t}{\mu^t \mathbf{e} + \mathbf{c}(\mathbf{x}^t)} \quad (5.24)$$

Por sua vez, o termo do lado esquerdo de (5.23) é definido pelas derivadas parciais de segunda ordem da função Lagrangiana do problema transformado, isto é, pela matriz Hessiana de (5.9) com respeito às variáveis primais. Em termos práticos, essa matriz pode ser expressa conforme (5.25).

$$\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) = \mathbf{H}(\mathbf{x}^t) + \mathbf{c}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \quad (5.25a)$$

$$\mathbf{c}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) = \mathbf{J}(\mathbf{x}^t)^\top \left[\text{diag} \left(\frac{\mu^t \boldsymbol{\delta}^t}{(\mu^t \mathbf{e} + \mathbf{c}(\mathbf{x}^t))^2} \right) \right] \mathbf{J}(\mathbf{x}^t) \quad (5.25b)$$

Ressalta-se que $\mathbf{H} : \mathbb{R}^5 \mapsto \mathbb{R}^{5 \times 5}$ denota a matriz Hessiana da função objetivo, sendo dada conforme (3.12). Já $\mathbf{c} : \mathbb{R}^5 \times \mathbb{R}^{10} \mapsto \mathbb{R}^{5 \times 5}$ é uma matriz auxiliar, definida de forma conveniente para facilitar as análises subsequentes. Em particular, o operador “diag(·)” apresentado em (5.25b) denota a matriz diagonal cujos elementos na diagonal principal são justamente os componentes do vetor resultante do seu argumento. Assim, como os termos desse argumento são, a rigor, não negativos e a estrutura regular da Jacobiana das restrições é aquela indicada em (3.22), conclui-se que \mathbf{c} é diagonal, com elementos não negativos e, portanto, semidefinida positiva (ou definida positiva, quando tais elementos são estritamente positivos).

À luz dessa explanação, cumpre chamar à atenção para uma particularidade dessa abordagem de cálculo da direção de busca. Em princípio, a direção \mathbf{p} é definida pelo método de Newton–Raphson multivariado, o qual constitui uma generalização do homônimo univariado introduzido na Seção 2.3.2. Sob a perspectiva teórica, tal direção então decorre de uma aproximação linear do comportamento da função vetorial que reflete o sistema de equações em questão, na vizinhança do ponto definido pela candidata à solução corrente [113, 116, 149]. No presente caso, porém, esse sistema corresponde às condições de estacionariedade do problema transformado, isto é, a equações já expressas em termos de derivadas de primeira ordem. Consequentemente, a linearização empregada pelo método de Newton–Raphson equivale, de forma natural, à construção de um modelo quadrático local para a função Lagrangiana associada, permitindo uma representação mais qualificada do seu comportamento na vizinhança do ponto $(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$ [113, 116].

5.2.3.2 Aspectos críticos

Embora o procedimento de determinação da direção de busca \mathbf{p} por meio da resolução do sistema linear estabelecido por (5.23) pareça simples, na prática há uma série de questões algébricas e geométricas que carecem de atenção especial [113, 116, 117, 124]. Mais precisamente,

existem três atributos da matriz Hessiana definida em (5.25) que precisam ser observadas: singularidade, condicionamento e autovalores [113].

No que diz respeito à singularidade, o tópico central é a própria viabilidade de resolução do sistema linear estabelecido por (5.23). Como $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ representa a matriz de coeficientes desse sistema, o cálculo da direção \mathbf{p} somente é possível se esta for invertível ou, mais expressamente, não singular [113, 202]. Vale lembrar que uma matriz quadrada, tal como a Hessiana, é singular quando o seu determinante é nulo e, conseqüentemente, não admite inversa [113, 202]. Analogamente, tal condição implica que suas linhas (ou colunas) são linearmente dependentes e que o seu posto é estritamente menor do que a sua ordem, de modo que o sistema linear associado pode não admitir solução única [113, 202].

Já no que concerne o condicionamento, a preocupação reside na confiabilidade do resultado obtido a partir da resolução de (5.23). Conforme discutido na Subseção 3.3.2, o número de algarismos significativos corretos em cálculos computacionais que envolvem a inversa de uma matriz simétrica, tal como a Hessiana, tende a ser maior à medida que a magnitude do número de condicionamento da própria matriz for menor. Dessa forma, ainda que $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ seja não singular — assegurando a viabilidade de resolução do sistema linear —, a direção \mathbf{p} encontrada somente será acurada se tal matriz for relativamente bem condicionada [113, 117].

Por fim, no que se refere aos autovalores, a atenção se volta para a efetividade da direção de busca resultante de (5.23). Como explicado na Subseção 3.2.4, os autovalores da matriz Hessiana quantificam as curvaturas principais na vizinhança do ponto sob análise e, portanto, revelam as possíveis direções que \mathbf{p} pode assumir. Em particular, a existência de ao menos um autovalor negativo implica uma direção de curvatura negativa, evidenciando a não convexidade daquela região do espaço de soluções viáveis e colocando em xeque a efetividade do inerente modelo quadrático [113]. Assim, mesmo que $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ seja não singular e relativamente bem condicionada — assegurando a resolução acurada do sistema linear —, a direção \mathbf{p} encontrada somente será pertinente se representar uma direção de descida, ou seja, rumo a um mínimo local. Para tanto, é desejável que a matriz Hessiana em questão não apresente qualquer autovalor negativo e, idealmente, $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}} \succ \mathbf{0}$, pois, nesse caso, a direção \mathbf{p} será garantidamente de descida [113, 191].

Ao trazer essas questões para o contexto do problema de extração dos parâmetros, um cenário bastante sensível se configura. A princípio, a ocorrência de singularidade em $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ não é esperada. Uma das justificativas para tanto é a diversificação genuína dos elementos de \mathbf{H} , determinada conforme (3.24), a qual sugere que suas linhas (ou colunas) sejam linearmente independentes. Outra justificativa decorre da contribuição diagonal oferecida por \mathbf{c} em (5.25a), a qual tende a reforçar essa hipótese de independência linear. Não obstante, não há como garantir que tal situação não se verifique na prática. Ademais, os efeitos numéricos associados à elevada disparidade de magnitudes das variáveis envolvidas — e, por conseqüência, dos elementos de \mathbf{H} e \mathbf{c} — podem induzir condições de quase-singularidade ao longo do processo iterativo.

Algoritmo 3 Escalonamento simétrico da Hessiana, considerando norma infinita.

Entradas: $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \boldsymbol{\mu}^t)$, $\varepsilon_{\text{esc}, \text{min}}$, $t_{\text{esc}, \text{max}}$
Saídas: \mathbf{D} , t_{esc} , ε_{esc}

- 1: Estabelecer $\mathbf{U} \leftarrow \nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \boldsymbol{\mu}^t)$
 - 2: Estabelecer $\mathbf{D} \leftarrow \mathbf{I}$
 - 3: Calcular $\mathbf{r}_0 \leftarrow \max(|\mathbf{U}|, \text{por linha})$
 - 4: Identificar $\mathcal{J}_0 \leftarrow \{j : (\mathbf{r}_0)_j = 0\}$
 - 5: Estabelecer $\varepsilon_{\text{esc}} \leftarrow +\infty$ e $t_{\text{esc}} \leftarrow 0$
 - 6: **Enquanto** $t_{\text{esc}} < t_{\text{esc}, \text{max}}$ e $\varepsilon_{\text{esc}} > \varepsilon_{\text{esc}, \text{min}}$ **faça**
 - 7: $t_{\text{esc}} \leftarrow t_{\text{esc}} + 1$
 - 8: Calcular $\mathbf{r} \leftarrow \max(|\mathbf{U}|, \text{por linha})$
 - 9: Para $j \in \mathcal{J}_0$, impor $r_j \leftarrow 1$
 - 10: Calcular $\boldsymbol{\sigma} \leftarrow \sqrt{\mathbf{r}}$
 - 11: Atualizar $\mathbf{D} \leftarrow \mathbf{D} \text{diag}(\boldsymbol{\sigma})^{-1}$
 - 12: Atualizar $\mathbf{U} \leftarrow \text{diag}(\boldsymbol{\sigma})^{-1} \mathbf{U} \text{diag}(\boldsymbol{\sigma})^{-1}$
 - 13: Calcular $\mathbf{r}_{\text{nov}} \leftarrow \max(|\mathbf{U}|, \text{por linha})$
 - 14: Para $j \in \mathcal{J}_0$, impor $(r_{\text{nov}})_j \leftarrow 1$
 - 15: Estabelecer $\varepsilon_{\text{esc}} \leftarrow \max_j |(r_{\text{nov}})_j - 1|$
 - 16: **Fim**
 - 17: **Retorna** \mathbf{D} , t_{esc} , ε_{esc}
-

Por outro lado, o mal condicionamento e a presença de autovalores negativos em $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ são não apenas esperados, mas também prováveis. Conforme demonstrado na Subseção 3.3.2, o número de condicionamento de \mathbf{H} tende a ser severamente elevado, característica que naturalmente acomete o condicionamento de $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$. Além disso, como evidenciado na Subseção 3.2.4, é possível que $\mathbf{H} \not\preceq \mathbf{0}$. Dessa forma, ainda que, por construção, $\mathbf{c} \succeq \mathbf{0}$ (ou $\mathbf{c} \succ \mathbf{0}$), não há como garantir que a soma dessas duas matrizes em (5.25a) resulte em uma que contenha apenas autovalores positivos e nulos.

Diante desse cenário, propõe-se o emprego de duas salvaguardas numéricas, com o objetivo de tornar estável e eficaz a computação da direção de Newton com base em (5.23). A seguir, os detalhes de cada uma dessas salvaguardas são devidamente apresentados.

5.2.3.3 Escalonamento simétrico

A primeira salvaguarda proposta se pauta na técnica de *escalonamento de matrizes* introduzida no trabalho apresentado em [199]. Em termos simples, trata-se de um algoritmo iterativo que busca equilibrar as normas das linhas e colunas de uma dada matriz em torno de 1 (um), por meio de pré-multiplicação por uma matriz diagonal e de pós-multiplicação por outra matriz diagonal [199]. Quando a matriz a ser escalonada é simétrica, tal como é o caso da Hessiana, essa matriz diagonais são idênticas e, portanto, essa técnica lida efetivamente com uma única *matriz diagonal de escalonamento* \mathbf{D} [199]. Para fins de clareza, apresenta-se no Algoritmo 3 o pseudocódigo dessa técnica, quando aplicada à matriz Hessiana definida por (5.25). Note-se que, além da matriz Hessiana a ser escalonada, este recebe dois parâmetros

adicionais de entrada: uma tolerância para o critério de aceitação, dada por $\varepsilon_{\text{esc},\min} \in (0, 1)$, e um número máximo de iterações do escalonamento, dado por $t_{\text{esc},\max} \in (1, \infty)$, os quais controlam, respectivamente, o nível de balanceamento requerido e o custo computacional do procedimento [199].

Uma das justificativas para a adoção dessa técnica em específico é que o escalonamento imposto por \mathbf{D} em (5.26a), origina uma *matriz Hessiana escalonada* \mathbf{A} que tende comprovadamente a apresentar um número de condicionamento expressivamente reduzido em comparação com $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$ [199]. Todavia, para que essa melhoria de condicionamento se traduza em maior robustez no cálculo da direção de Newton, faz-se necessário aplicar o escalonamento de forma consistente a todo o sistema linear estabelecido em (5.23). Nesse sentido, define-se complementarmente o *vetor gradiente escalonado* \mathbf{b} de acordo com (5.26b) e a *direção de busca escalonada* $\tilde{\mathbf{p}}$ como (5.26c).

$$\mathbf{A} := \mathbf{D} \nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \mathbf{D} \quad (5.26a)$$

$$\mathbf{b} := \mathbf{D} \nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \quad (5.26b)$$

$$\tilde{\mathbf{p}} := \mathbf{D} \mathbf{p} \quad (5.26c)$$

Outra justificativa para a adoção dessa técnica é que o escalonamento imposto por \mathbf{D} em (5.26a) preserva a estrutura simétrica de $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$, de modo que \mathbf{A} permanece simétrica [199]. Vale destacar que esse não é um aspecto compartilhado por outras técnicas de escalonamento de matrizes e, portanto, constitui um diferencial da sua abordagem [199]. No presente contexto, essa manutenção da simetria é particularmente conveniente porque, além de resguardar a natureza do subproblema de otimização em questão, viabiliza o emprego de rotinas de fatoração dedicadas e reconhecidamente estáveis, tais como Cholesky ou LDL^T , na resolução do sistema linear escalonado expresso em (5.27) [113, 117].

$$\mathbf{A} \tilde{\mathbf{p}} = -\mathbf{b} \quad (5.27)$$

Não obstante, é importante ressaltar que a técnica adotada não altera, de maneira substancial, as propriedades espectrais relevantes da matriz Hessiana original. Mais expressamente, por se tratar de uma transformação por congruência, \mathbf{A} herda a *inércia* de $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}$, isto é, as respectivas quantidades de autovalores positivos, negativos e nulos [113, 202]. Assim, embora seja central para a melhoria do condicionamento e, conseqüentemente, para a estabilidade numérica do cálculo da direção \mathbf{p} , essa técnica não é capaz, nem se destina, a tratar situações em que $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}} \not\approx \mathbf{0}$. Essa circunstância justifica o emprego da segunda salvaguarda numérica.

5.2.3.4 Regularização inercial

A segunda salvaguarda proposta se pauta em uma técnica de *regularização inercial*, a qual é amplamente discutida na literatura, sob diferentes denominações, incluindo modificação da

Algoritmo 4 Regularização inercial via verificação espectral no espaço escalonado.

Entradas: \mathbf{A} , γ_0 , β_{reg} , γ_{max}
Saída: γ

- 1: Estabelecer $\gamma \leftarrow 0$
 - 2: Calcular os autovalores de $(\mathbf{A} + \gamma\mathbf{I})$
 - 3: **Enquanto** existir ao menos um autovalor negativo em $(\mathbf{A} + \gamma\mathbf{I})$ **faça**
 - 4: **Se** $\gamma = 0$ **então**
 - 5: Estabelecer $\gamma \leftarrow \gamma_0$
 - 6: **Caso contrário**
 - 7: Atualizar $\gamma \leftarrow \beta_{\text{reg}} \cdot \gamma$
 - 8: **Fim**
 - 9: **Se** $\gamma > \gamma_{\text{max}}$ **então**
 - 10: **Interromper** (*falha da regularização inercial*)
 - 11: **Retorna** γ
 - 12: **Fim**
 - 13: Calcular os autovalores de $(\mathbf{A} + \gamma\mathbf{I})$
 - 14: **Fim**
 - 15: **Retorna** γ
-

Hessiana, correção de inércia, regularização diagonal ou regularização de Levenberg-Marquardt [113, 117, 124, 129]. De forma sucinta, essa técnica altera a natureza indefinida da matriz Hessiana pela adição de uma parcela diagonal, construída a partir da multiplicação de um *parâmetro de regularização*, dado pelo escalar real $\gamma \in [0, \infty)$, por uma matriz identidade de dimensões compatíveis [113, 129]. De modo prático, o valor desse parâmetro pode ser determinado por meio de uma rotina iterativa simples, cujo objetivo é assegurar que a matriz resultante seja, ao menos, semidefinida positiva [113, 129].

No presente contexto, é conveniente observar que a matriz Hessiana escalonada, definida em (5.26a), tende a apresentar elementos com magnitudes mais equilibradas e reduzidas do que aqueles da matriz Hessiana original, dada por (5.25). Assim, a determinação de γ diretamente sobre \mathbf{A} tende a ser mais previsível, sem prejudicar sua eficácia. Atento a isso, apresenta-se no Algoritmo 4 o pseudocódigo adaptado da técnica de regularização inercial, onde γ é inicializado em zero e, se necessário, aumentado a partir de um valor inicial $\gamma_0 \in (0, \infty)$ por atualizações multiplicativas via $\beta_{\text{reg}} \in (1, \infty)$, até que a matriz resultante de $\mathbf{A} + \gamma\mathbf{I}$ seja ao menos semidefinida positiva.

Embora $\gamma \neq 0$ implique uma distorção controlada da natureza do subproblema, é importante notar que sua introdução permite encontrar uma direção de descida que, mesmo não sendo necessariamente ótima, é suficiente para propiciar o progresso do processo iterativo do EPM [113, 129]. Ademais, essa perturbação acarreta efeitos algébricos que são desejáveis. Um destes é a mitigação de uma eventual incidência de singularidade ou quase-singularidade em \mathbf{A} , reforçando a viabilidade do sistema linear regularizado [113, 129]. Outro é que, ao afastar os autovalores de zero, esta tende a contribuir marginalmente para a melhoria do condicionamento de \mathbf{A} e, conseqüentemente, para a estabilidade numérica do referido sistema [117, 202].

Note-se, contudo, que a introdução de um valor máximo admissível para o parâmetro de regularização, denotado por γ_{\max} , constitui uma salvaguarda numérica importante. Embora o aumento progressivo de γ seja, em geral, eficaz para assegurar a semidefinitude positiva de $\mathbf{A} + \gamma\mathbf{I}$, valores demasiadamente elevados podem induzir distorções desnecessárias no subproblema regularizado, fazendo com que o termo diagonal passe a dominar a matriz e reduzindo a influência das informações de segunda ordem contidas em \mathbf{A} . Assim, a imposição de γ_{\max} evita a escalada excessiva do parâmetro e limita a regularização ao regime em que esta permanece efetivamente “controlada”, prevenindo ajustes exagerados que tenderiam a produzir direções pouco informativas do ponto de vista do modelo local.

Em posse de um valor adequado para o parâmetro de regularização, a direção de busca pode ser efetivamente determinada, no espaço escalonado, a partir da resolução do sistema linear regularizado estabelecido em (5.28).

$$(\mathbf{A} + \gamma\mathbf{I})\tilde{\mathbf{p}} = -\mathbf{b} \quad (5.28)$$

5.2.3.5 Encadeamento executório

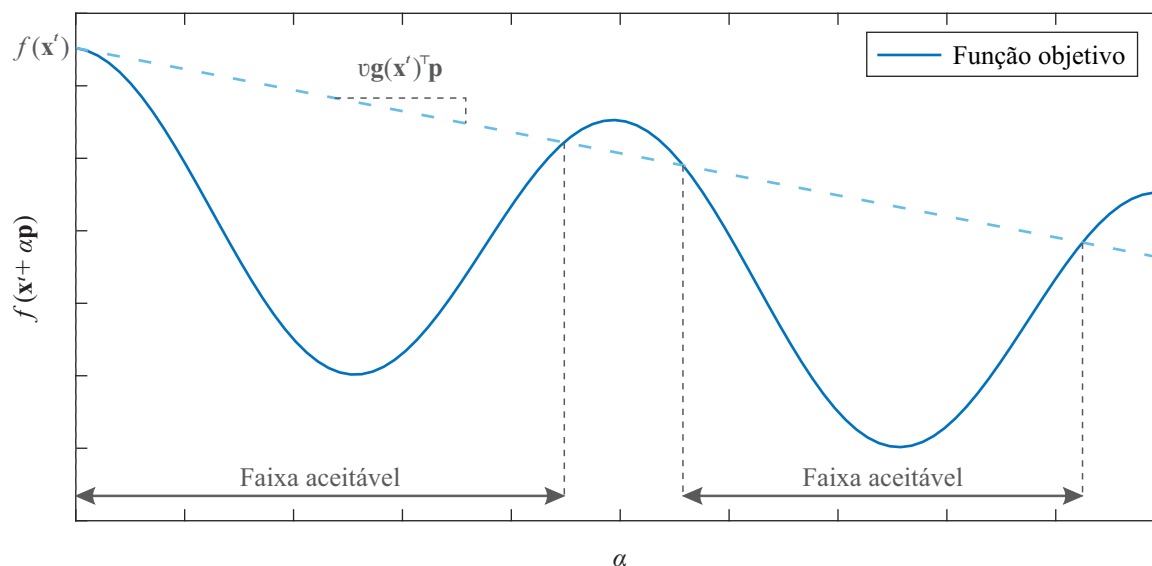
O conteúdo anteriormente exposto evidencia que a determinação da direção de busca \mathbf{p} no âmbito da versão escalonada e regularizada concebida do Método de Newton possui fundamentação sólida e coerente, porém envolve um conjunto não trivial de rotinas e formulações matemáticas. Com intuito de tornar mais claro o seu encadeamento executório, explicita-se a seguir o passo a passo a ser adotado para a sua consolidação:

1. Construir o vetor gradiente da função Lagrangiana, de acordo com (5.24);
2. Construir a matriz Hessiana da função Lagrangiana, de acordo com (5.25);
3. Obter a matriz de escalonamento por meio do Algoritmo 3;
4. Escalonar o vetor gradiente construído, conforme (5.26b);
5. Escalonar a matriz Hessiana construída da função Lagrangiana, conforme (5.26a);
6. Determinar um valor adequado do parâmetro de regularização através do Algoritmo 4;
7. Resolver o sistema linear regularizado, formalizado em (5.28), de modo a obter $\tilde{\mathbf{p}}$;
8. Mapear a solução $\tilde{\mathbf{p}}$ para o espaço original por meio de (5.29), a fim de determinar \mathbf{p} .

$$\mathbf{p} = \mathbf{D}^{-1}\tilde{\mathbf{p}} \quad (5.29)$$

Uma vez realizado esse passo a passo e, assim, determinada a direção de busca \mathbf{p} , resta definir o comprimento de passo α a ser tomado ao longo dessa direção. A seguir, apresenta-se a estratégia adotada para esse propósito.

Figura 21 – Representação da lógica subjacente à condição de Armijo.



5.2.3.6 Comprimento do passo

A rigor, o máximo benefício associado à direção de busca \mathbf{p} pode ser obtido por meio da resolução exata do subproblema de otimização unidimensional enunciado em (5.30) [113]. Entretanto, observa-se que o cumprimento dessa tarefa normalmente implica um elevado custo computacional [113, 117].

$$\underset{\alpha \in \mathbb{R}_+}{\text{minimizar}} \quad f(\mathbf{x}^t + \alpha \mathbf{p}) \quad (5.30)$$

Tendo em vista o compromisso com efetividade e eficiência, algoritmos de uso corrente costumam então recorrer a uma estratégia denominada, em língua inglesa, *backtracking* [113]. Nesta, o comprimento de passo é progressivamente reduzido até que se observe um decréscimo suficiente da função objetivo [113]. Uma das métricas de avaliação mais amplamente empregadas para esse propósito é a chamada *condição de Armijo*, segundo a qual um dado valor de α é aceito se satisfizer (5.31) [113, 116].

$$\mathcal{F}(\mathbf{x}^t + \alpha \mathbf{p}) \leq \mathcal{F}(\mathbf{x}^t) + v \alpha \mathbf{g}(\mathbf{x}^t)^\top \mathbf{p} \quad (5.31)$$

Conforme ilustrado na Figura 21, o produto $\alpha \mathbf{g}(\mathbf{x}^t)^\top \mathbf{p}$, com $\mathbf{g}(\mathbf{x}^t)^\top \mathbf{p} < \mathbf{0}$ representa o decréscimo esperado da função objetivo segundo sua aproximação linear local [116]. Já o fator multiplicativo, dado pelo escalar real $v \in (0, 1)$, garante que (5.31) seja satisfeita quando ao menos uma pequena fração dessa diminuição esperada for efetivamente alcançada [116]. Logo, qualquer valor de α que satisfaça a condição de Armijo assegura progresso do processo iterativo no sentido de minimizar o valor correspondente da função objetivo.

Não o bastante, dada a natureza restrita do problema transformado, não é recomendável que a aceitação de α se baseie exclusivamente em critérios de otimalidade [113]. Em particular, é pertinente certificar se $\mathbf{x}^t + \alpha \mathbf{p}$ reflete, ao menos, uma candidata à solução “menos inviável” do

que aquela corrente dada por \mathbf{x}^t [113, 201]. Nesse sentido, exige-se complementarmente que α seja aprovado no *filtro de viabilidade*, dado pela condição apresentada em (5.32) [201].

$$\|\mathbf{v}(\mathbf{x}^t + \alpha\mathbf{p})\|_\infty \leq \|\mathbf{v}(\mathbf{x}^t)\|_\infty \quad (5.32a)$$

$$v_j(\mathbf{x}^t + \alpha\mathbf{p}) = \min(\mathbf{x}^t + \alpha\mathbf{p}, -\boldsymbol{\mu}^t) \quad (5.32b)$$

Cabe ressaltar que o vetor auxiliar $\mathbf{v} \in \mathbb{R}^{10}$ é definido, no âmbito desse filtro, mediante a regra estabelecida em (5.32b), a qual é distinta daquela adotada em (5.21b) no critério de convergência. Isso se deve ao fato de que, neste ponto, tal vetor está sendo empregado de forma orientada ao problema transformado, e não à avaliação em face do problema original.

Assim, apresenta-se no Algoritmo 5 o pseudocódigo da estratégia de *backtracking*, considerando esses critérios de otimalidade e viabilidade. Em consonância com as práticas recomendadas na literatura específica, o valor inicial de α é definido como o maior comprimento de passo admissível do ponto de vista da viabilidade primária, denotado por α_{max} . Tal valor é usualmente obtido por meio da chamada *regra da fração até a fronteira*³, a qual assegura que a candidata $\mathbf{x}^t + \alpha\mathbf{p}$ permaneça estritamente no interior da região viável relaxada, evitando violações prematuras das restrições que definem o subespaço primal. Uma das possíveis formas de expressão dessa regra é apresentada em (5.33a), onde τ é um parâmetro auxiliar dado por (5.33b), considerando um valor real mínimo admissível $\tau_{min} \in (0, 1)$ [129].

$$\alpha_{max} = \max\{\alpha \in (0, 1) : \mathbf{x}^t + \alpha\mathbf{p} \geq (1 - \tau)\mathbf{x}^t\} \quad (5.33a)$$

$$\tau = \max\{\tau_{min}, 1 - \mu^t\} \quad (5.33b)$$

Vale ressaltar que α_{max} não corresponde, por si só, a um comprimento de passo automaticamente aceitável, mas sim a um limite superior estrutural a partir do qual se inicia o processo de busca em linha. Dessa forma, caso o valor corrente de α_{max} não satisfaça simultaneamente a condição de Armijo (5.31) e o filtro de viabilidade (5.32), procede-se à sua redução por meio de um fator multiplicativo, dado pelo escalar real $\beta_{back} \in (0, 1)$. Esse processo é repetido iterativamente até que ambas as condições sejam atendidas ou, em última instância, até que α atinja um valor mínimo preestabelecido.

Note-se que a introdução de um comprimento de passo mínimo, denotado por α_{min} , constitui uma salvaguarda numérica importante, impedindo reduções excessivas que conduzam a deslocamentos desprezíveis ao longo da direção de busca [113, 129]. Valores inferiores a esse limiar são, em geral, tidos como inefetivos na prática, pois tendem a não produzir progresso mensurável nem na redução da função objetivo nem na melhoria do grau de viabilidade. Nessas circunstâncias, o procedimento de busca em linha é interrompido, sinalizando a necessidade de reavaliar a direção de busca \mathbf{p} e, conseqüentemente, a falha do presente método de Newton com busca em linha.

³ Uma tradução livre da denominação em língua inglesa: “*fraction-to-the-boundary rule*”.

Algoritmo 5 Determinação do comprimento do passo via estratégia de *backtracking*.

Entradas: $\mathcal{F}(\cdot)$, $\mathbf{g}(\cdot)$, $\mathbf{c}(\cdot)$, \mathbf{x}^t , \mathbf{p} , μ^t , v , β , τ_{\min} , α_{\min}

Saída: α

- 1: Calcular α_{\max} via (5.33)
 - 2: Estabelecer $\alpha \leftarrow \alpha_{\max}$
 - 3: **Enquanto** (5.31) e (5.32) não forem satisfeitas **faça**
 - 4: Atualizar $\alpha \leftarrow \beta_{\text{back}} \cdot \alpha$
 - 5: **Se** $\alpha < \alpha_{\min}$ **então**
 - 6: **Interromper** (*falha da busca em linha*)
 - 7: **Retorna** α
 - 8: **Fim**
 - 9: **Fim**
 - 10: **Retorna** α
-

5.2.4 Racionalização da arquitetura

Com a apresentação do EPM desenvolvido concluída, convém agora explicitar como as suas escolhas de projeto se articulam com o diagnóstico estabelecido no Capítulo 3, de modo a evidenciar o caráter sob medida de sua concepção.

Como discutido ao longo do referido capítulo, a formulação do problema de extração dos parâmetros impõe dificuldades que se manifestam diretamente no subespaço primal, isto é, no domínio onde a solução é construída. Nesse sentido, o EPM desenvolvido foi delineado para evoluir nesse subespaço por meio da minimização irrestrita de $\mathcal{L}_{\text{epm}}(\mathbf{x}, \boldsymbol{\delta}^t; \mu^t)$, conforme (5.10), enquanto as variáveis duais e o parâmetro de barreira atuam como recursos auxiliares de condução e balanceamento. Assim, em vez de contornar tais dificuldades por tratamentos externos, o método as enfrenta diretamente por meio de mecanismos incorporados à sua estrutura. A seguir, destacam-se essas inter-relações de forma organizada.

(i) Restrições de canalização e admissibilidade físico-matemática

A primeira dificuldade identificada reside na necessidade de respeitar restrições de canalização que assegurem, simultaneamente, a admissibilidade física dos parâmetros do modelo de um diodo e a consistência matemática da abordagem formal baseada na função W de Lambert, devido às condições de domínio dessa função (ver Seções 3.1 e 3.2).

No EPM desenvolvido, essa exigência não foi tratada como um ajuste posterior, mas incorporada à própria base da formulação por meio da barreira logarítmica modificada. Ao conduzir ao problema transformado (5.7), essa barreira redefine o espaço de soluções viáveis efetivamente explorado pelo algoritmo, fazendo com que o processo iterativo evolua em uma região primal relaxada $\{\mathbf{c}(\mathbf{x}) \geq -\mu \mathbf{e}\}$, cuja extensão é governada pelo ajuste adaptativo do parâmetro de barreira.

Ao ampliar temporária e controladamente esse conjunto de soluções admissíveis, a relaxação favorece a exploração do espaço de busca sobretudo nas iterações iniciais. Isso é

particularmente relevante quando as candidatas à solução correntes ainda se encontram distantes do regime ótimo, pois reduz a tendência a passos excessivamente conservadores induzidos por um conjunto factível demasiadamente estreito. À medida que μ é atualizado, essa margem é reduzida de forma sistemática, reforçando progressivamente a imposição das restrições do problema original. Por essa razão, a barreira logarítmica modificada constitui o alicerce estrutural do EPM desenvolvido: esta define, desde a origem, o domínio primal de operação e o mecanismo de reforço progressivo da viabilidade em direção ao regime do problema original.

(ii) *Não convexidade e qualidade das soluções*

A segunda dificuldade decorre da natureza não convexa do problema (ver Seção 3.2). Em termos práticos, isso implica que o procedimento de otimização pode ser conduzido a regiões nas quais a matriz Hessiana $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$ apresenta autovalores negativos, de modo que a direção calculada a partir das informações locais por esta fornecidas pode deixar de constituir uma direção de descida e, conseqüentemente, não produzir progresso efetivo ou favorecer a estabilização prematura em soluções de qualidade inferior.

Para lidar com esse cenário, introduziu-se a regularização inercial, cujo propósito é assegurar que o sistema linear empregado no cálculo da direção de busca permaneça coerente com o objetivo de minimização em cada iteração. Ao adicionar de maneira controlada o termo $\gamma \mathbf{I}$, promove-se uma correção mínima e controlada quando a curvatura local se mostra indefinida, viabilizando uma direção de busca compatível com descida.

Assim, a regularização não substitui a informação de segunda ordem disponível, mas a ajusta no estritamente necessário para que possa ser utilizada de maneira efetiva no avanço do processo iterativo. Dessa forma, a regularização inercial atua como salvaguarda para que a não convexidade não impeça o progresso em direção a soluções de melhor qualidade.

(iii) *Discrepância de escala e estabilidade numérica*

A terceira dificuldade está associada à discrepância de escala entre variáveis e ao mal condicionamento das matrizes envolvidas (ver Seção 3.3). Em particular, no âmbito do método de Newton, tal discrepância de escala tende a se manifestar na própria matriz Hessiana $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$, a qual pode apresentar elementos de magnitudes bastante distintas, impactando desfavoravelmente o seu condicionamento e tornando sensível a resolução dos sistemas lineares empregados no cálculo da direção de busca.

A fim de mitigar esse efeito, incorporou-se o escalonamento simétrico da matriz Hessiana $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$, cujo propósito é equilibrar as magnitudes das linhas e colunas antes da resolução do sistema linear correspondente. Desse modo, reduz-se o número de condicionamento e reforça-se a robustez numérica dos cálculos, preservando-se a estrutura simétrica da matriz.

Assim, esses três mecanismos compõem uma arquitetura metodológica coerente com o diagnóstico estabelecido no Capítulo 3: a barreira logarítmica modificada estrutura e condiciona o domínio primal desde a origem em função das restrições de canalização; a regularização inercial impede que a não convexidade comprometa o avanço do processo iterativo; e o escalonamento simétrico reforça a robustez numérica frente às discrepâncias de escala entre variáveis e ao consequente mal condicionamento das matrizes envolvidas.

Por fim, cumpre registrar que, em decorrência de sua concepção sob medida, o EPM desenvolvido passa a ser denominado *tailored exterior-point method* (TEPM). Registre-se que se optou por formalizar o método apenas sob a nomenclatura em língua inglesa, em consonância com a prática internacional observada na literatura específica; nesse contexto, a sigla TEPM será adotada como forma de referência predominante ao longo do texto, à semelhança do que se verifica em trabalhos correlatos.

5.2.5 Consolidação do algoritmo

Tendo sido explicitadas as bases conceituais e as inter-relações que fundamentam o método doravante denominado TEPM, apresenta-se no Algoritmo 6 a sua consolidação formal. O pseudocódigo sintetiza, de forma estruturada, o encadeamento lógico das etapas responsáveis pelo ajuste das variáveis primais, pela atualização das variáveis duais e do parâmetro de barreira, incorporando explicitamente as salvaguardas numéricas delineadas ao longo desta seção.

Inicialmente, nas linhas 1–2, são estabelecidos o contador de iterações e a candidata inicial à solução, definida por $(\mathbf{x}^0, \boldsymbol{\delta}^0; \mu^0)$. Em seguida, o laço principal (linha 3) conduz o processo iterativo até que seja atingido o *número máximo de iterações*, dado por $t_{\max} \in (0, \infty)$, o qual é introduzido como critério adicional de parada, conforme prática consolidada no campo de métodos de otimização não linear [120, 189].

As linhas 4–12 concentram o procedimento proposto para o ajuste das variáveis primais via método de Newton escalonado e regularizado com busca em linha. Especificamente, as linhas 4–5 constroem o gradiente e a Hessiana da função Lagrangiana transformada; a linha 6 aplica o escalonamento simétrico; as linhas 7–9 realizam a regularização inercial e a resolução do sistema linear correspondente; a linha 10 recupera a direção no espaço original; e as linhas 11–12 determinam o comprimento de passo via *backtracking* e atualizam as variáveis primais.

As linhas 13–15 implementam a atualização das variáveis duais, incorporando a correção proposta para evitar o aprisionamento de multiplicadores em valor nulo. Já as linhas 16–17 tratam da atualização adaptativa do parâmetro de barreira com imposição de piso numérico, assegurando estabilidade mesmo em cenários de complementaridade muito reduzida.

Por fim, as linhas 18–23 verificam simultaneamente os critérios de viabilidade e otimalidade. Caso ambos sejam satisfeitos, o processo iterativo é interrompido por convergência e a solução corrente é retornada. Caso contrário, antes de efetivamente iniciar uma nova iteração,

Algoritmo 6 Método de ponto exterior proposto sob medida.

Entradas: $\mathbf{x}^0, \boldsymbol{\delta}^0, \mu^0, \varepsilon_{\text{otim}}, \varepsilon_{\text{viab}}, \varepsilon_{\text{fobj}}, \varepsilon_{\text{varx}}, t_{\text{max}}, \varepsilon_{\text{esc,min}}, t_{\text{esc,max}}, \gamma_0, \gamma_{\text{max}}, \beta_{\text{reg}}, \alpha_{\text{min}}, \nu, \tau_{\text{min}}, \beta_{\text{back}}, \delta_{\text{aux}}, \mu_{\text{min}},$

Saídas: $(\mathbf{x}^*, \boldsymbol{\delta}^*; \mu^*)$

- 1: Estabelecer $t \leftarrow 0$
 - 2: Estabelecer $(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t) \leftarrow (\mathbf{x}^0, \boldsymbol{\delta}^0; \mu^0)$
 - 3: **Enquanto** $t < t_{\text{max}}$ **faça**
 - 4: Calcular $\nabla_{\mathbf{x}} \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$ via (5.24)
 - 5: Calcular $\nabla_{\mathbf{xx}}^2 \mathcal{L}_{\text{epm}}(\mathbf{x}^t, \boldsymbol{\delta}^t; \mu^t)$ via (5.25)
 - 6: Obter \mathbf{D} via Algoritmo 3
 - 7: Calcular \mathbf{A} e \mathbf{b} via (5.26)
 - 8: Determinar γ via Algoritmo 4
 - 9: Resolver (5.28) para obter $\tilde{\mathbf{p}}$
 - 10: Recuperar \mathbf{p} via (5.29)
 - 11: Determinar α via Algoritmo 5
 - 12: Atualizar $\mathbf{x}^{t+1} \leftarrow \mathbf{x}^t + \alpha \mathbf{p}$ via (5.12)
 - 13: **Para** $j = 1, \dots, m_c$ **faça**
 - 14: Atualizar δ_j^{t+1} via (5.17)
 - 15: **Fim**
 - 16: Calcular σ via (5.19)
 - 17: Atualizar μ^{t+1} via (5.20)
 - 18: Verificar (5.21)
 - 19: Verificar (5.22)
 - 20: **Se** (5.21a) e (5.22) forem satisfeitas **então**
 - 21: **Interromper** (*convergência alcançada*)
 - 22: **Retorna** $(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \mu^{t+1})$
 - 23: **Fim**
 - 24: Verificar (5.34)
 - 25: Verificar (5.35)
 - 26: **Se** (5.34) e (5.35) forem satisfeitas **então**
 - 27: **Interromper** (*estagnação do processo*)
 - 28: **Retorna** $(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \mu^{t+1})$
 - 29: **Fim**
 - 30: Estabelecer $t \leftarrow t + 1$
 - 31: **Fim**
 - 32: **Retorna** $(\mathbf{x}^{t+1}, \boldsymbol{\delta}^{t+1}; \mu^{t+1})$
-

procede-se a uma checagem de continuidade do processo iterativo, conforme as linhas 24–29. Como é usual em implementações práticas de métodos de otimização, tal checagem é realizada por meio de *critérios de estagnação* associados à variação das variáveis primais e ao decréscimo da função objetivo. Esses critérios são frequentemente estabelecidos a partir das condições mostradas em (5.34) e (5.35), nas quais $\varepsilon_{\text{var}} \in (0, 1)$ e $\varepsilon_{\text{obj}} \in (0, 1)$ denotam, respectivamente, as tolerâncias numéricas associadas.

$$\|\mathbf{x}^{t+1} - \mathbf{x}^t\|_{\infty} \leq \varepsilon_{\text{var}} \quad (5.34)$$

$$|\mathcal{F}(\mathbf{x}^{t+1}) - \mathcal{F}(\mathbf{x}^t)| \leq \varepsilon_{\text{obj}} \quad (5.35)$$

Tabela 16 – Parâmetros de controle do método de ponto exterior proposto.

Escopo	Parâmetro	Valor padrão
Critério de convergência	ϵ_{otim}	1×10^{-18}
	ϵ_{viab}	1×10^{-12}
Critério de estagnação	ϵ_{obj}	1×10^{-18}
	ϵ_{var}	1×10^{-18}
Critério de parada	t_{max}	1×10^3
	f_{max}	110×10^3
Escalonamento simétrico	$\epsilon_{\text{esc,min}}$	1×10^{-12}
	$t_{\text{esc,max}}$	125
Regularização inercial	γ_0	1×10^{-12}
	γ_{max}	1×10^{40}
	β_{reg}	8
Procedimento <i>backtracking</i>	α_{min}	1×10^{-40}
	ν	1×10^{-6}
	τ_{min}	0,9995
	β_{back}	0,5
Perturbação da variável dual	δ_{aux}	0,1
Limite inferior do parâmetro barreira	μ_{min}	1×10^{-12}
Cálculo do gradiente via passo complexo	h_{min}	1×10^{-30}
	h_{rel}	1×10^{-20}

Se esses critérios de estagnação indicarem ausência de avanço significativo, o laço principal é encerrado por um critério prático de parada e a solução corrente é retornada. Apenas no caso em que nem convergência nem estagnação sejam detectadas é que o contador iterativo é incrementado (linha 30), dando início a uma nova iteração.

Cumpra destacar que, além do número máximo de iterações t_{max} , na implementação prática também é usual controlar o *número total de avaliações da função objetivo*, o qual é dado por $f_{\text{max}} \in (0, \infty)$ [120, 189]. Tal mecanismo constitui uma salvaguarda adicional contra custos computacionais excessivos. Não obstante, esse contador não foi explicitamente representado no pseudocódigo, uma vez que depende diretamente da estratégia adotada para o cálculo das derivadas envolvidas, devendo ser entendido como subjacente à implementação do método.

No que concerne às condições iniciais das variáveis duais e do parâmetro de barreira, diversos testes preliminares foram conduzidos com o objetivo de estabelecer estimativas iniciais adequadas. Os melhores resultados foram obtidos ao se adotar valores iniciais iguais a 0, 1 para essas grandezas, assegurando simultaneamente coerência com o subespaço primal relaxado e condições iniciais compatíveis com o princípio de complementaridade do método.

Adicionalmente, uma série extensa de experimentos numéricos foi realizada para calibrar os parâmetros de controle do algoritmo — tais como tolerâncias, fatores multiplicativos e limites máximos admissíveis. Os valores que se mostraram mais eficazes são sintetizados na Tabela 16. Ressalta-se, contudo, que tais parâmetros não possuem caráter universal, podendo demandar ajustes finos conforme o contexto de aplicação e as características do problema tratado.

5.2.6 Implementação computacional

Uma vez consolidado o algoritmo do TEPM e esclarecidos todos os seus recursos e parâmetros de controle, sua implementação computacional pode ser conduzida de forma estruturada. Contudo, há um aspecto que transcende a lógica estrita do método e que se revela determinante para o seu desempenho prático: o cálculo das derivadas de primeira e de segunda ordem da função objetivo ao longo do processo iterativo (veja a Figura 20). Em particular, como o TEPM se apoia em informações diferenciais para compor direções de busca no subespaço primal, a qualidade dessas derivadas influencia diretamente a robustez do procedimento e a estabilidade numérica das iterações.

Em princípio, as derivadas requeridas poderiam ser obtidas por diferenciação simbólica, diferenciação automática ou diferenciação numérica [113, 116]. No contexto desta pesquisa, a diferenciação simbólica mostrou-se pouco efetiva no ambiente do MATLAB, versão 2016b, tanto por impor sobrecarga computacional quanto por apresentar inconsistências numéricas em frequentes situações. Já a diferenciação automática, embora reconhecidamente atrativa por fornecer derivadas acuradas, não se mostrou viável na prática, seja por limitações de suporte nativo na versão disponível do MATLAB, seja pela necessidade de tratamentos adicionais requeridos à função W de Lambert.

Diante disso, adotou-se uma abordagem numérica para o cálculo das derivadas, em linha com a prática consolidada em *solvers* acadêmicos e comerciais. Ressalta-se, entretanto, que “diferenciação numérica” compreende uma família de estratégias, cujas propriedades de acurácia, estabilidade e custo podem variar significativamente [113, 116, 119]. No âmbito desta pesquisa, foram realizados testes preliminares abrangentes com alternativas puramente baseadas em diferenças finitas, bem como com aproximações que evitam a matriz Hessiana explícita, a exemplo das estratégias quasi-Newton, e com aproximações típicas de métodos de mínimos quadrados, como variações do tipo Gauss–Newton [113, 116]. Essas alternativas, embora atrativas do ponto de vista de custo, mostraram fragilidades em termos de estabilidade e robustez quando aplicadas ao problema de extração dos parâmetros na formulação considerada.

Como resultado, a estratégia que apresentou melhor desempenho global foi uma metodologia híbrida: emprega-se o *método do passo complexo* para o cálculo do gradiente e um esquema de *diferenças finitas centradas* aplicado ao gradiente para aproximar a Hessiana. Nesse arranjo, cada componente g_j do vetor gradiente da função objetivo, dada de acordo com (3.21), é diretamente aproximado pelo método do passo complexo, conforme (5.36) [203, 204]. Destaca-se que x_j é o j -ésimo componente do vetor de variáveis de decisão, podendo ser respectivo a I_{ph} , I_0 , n , R_{sh} , ou R_s . Por sua vez, \mathbf{v}_j define uma base canônica em j , ou seja, um vetor majoritariamente nulo onde somente o componente j é dado por 1 (um). Ressalta-se ainda que o operador “ $\text{Im}(\cdot)$ ” retorna a parte imaginária do seu argumento.

$$g_j(\mathbf{x}) := \frac{\partial \mathcal{F}(\mathbf{x})}{\partial x_j} \approx \frac{\text{Im}\left(\mathcal{F}(\mathbf{x} + ih_j^{cs} \mathbf{v}_j)\right)}{h_j^{cs}} \quad (5.36)$$

Embora o *passo de perturbação* dado pelo escalar real $h_j^{cs} > 0$ em (5.36) possa, a rigor, ser tomado extremamente pequeno sem induzir cancelamento subtrativo, optou-se por atrelar o seu valor à escala da variável perturbada. Essa escolha visa lidar de forma robusta com diferentes ordens de magnitude entre as variáveis de decisão do problema de extração dos parâmetros e, ao mesmo tempo, evitar que a perturbação se torne tão pequena a ponto de produzir variações numericamente inócuas na avaliação da função, em razão de limitações inerentes à aritmética de ponto flutuante [176, 203, 204]. Sendo assim, o valor do passo é determinado dinamicamente conforme (5.37), em que o escalar real $h_{rel} \in (0, 1)$ estabelece um fator relativo à escala da variável perturbada e $h_{min} < h_{rel}$ impõe um valor mínimo admissível (valores adotados indicados na Tabela 16).

$$h_j^{cs} = \max\{h_{min}, h_{rel} \min\{1, |x_j|\}\} \quad (5.37)$$

Por outro lado, cada componente $H_{j,l}$ da matriz Hessiana da função objetivo, dada de acordo com (3.24), é aproximado por um esquema de diferenças finitas centradas aplicado ao gradiente, conforme (5.38) [113, 119]. Nessa expressão, cada g_j denota o j -ésimo componente do gradiente, previamente computado por meio de (5.36), enquanto \mathbf{v}_j representa a base canônica associada ao índice j .

$$H_{j,l}(\mathbf{x}) := \frac{\partial^2 \mathcal{F}(\mathbf{x})}{\partial x_j \partial x_l} \approx \frac{g_j(\mathbf{x} + h_j^{df} \mathbf{v}_j) - g_j(\mathbf{x} - h_j^{df} \mathbf{v}_j)}{2h_j^{df}} \quad (5.38)$$

Diferentemente do método do passo complexo, o passo de perturbação dado pelo escalar real $h_j^{df} \in (0, 1)$, se tomado extremamente pequeno, pode induzir cancelamento subtrativo em (5.38). Dessa forma, adota-se um passo compatível com o compromisso entre erro de truncamento e erro de arredondamento, intrínseco ao método de diferenças finitas, definindo-se o seu valor dinamicamente por meio de (5.39) [113, 119, 176], em que ϵ_{maq} denota a precisão de máquina.

$$h_j^{df} = \sqrt[4]{\epsilon_{maq}} \min\{1, |x_j|\} \quad (5.39)$$

Essa metodologia híbrida privilegia elevada acurácia no cálculo das derivadas de primeira ordem, ao mesmo tempo em que preserva simplicidade operacional para a obtenção de derivadas de segunda ordem completas. Não obstante, é importante reconhecer que essa não é a alternativa mais eficiente, sobretudo no cálculo da matriz Hessiana: como a aproximação por diferenças finitas centradas é construída a partir de avaliações adicionais do gradiente, o custo por iteração tende a ser significativo, exigindo um número elevado de avaliações da função objetivo em uma única iteração do TEPM.

No algoritmo consolidado, esse efeito se reflete diretamente na necessidade de adotar um limite relativamente alto para o número total de avaliações permitidas da função objetivo

(f_{\max}), conforme indicado na Tabela 16. Ainda assim, como a estratégia híbrida se mostrou mais eficaz em termos de robustez, esta foi adotada como configuração inicial de referência. A investigação de alternativas mais eficientes permanece, assim, como uma oportunidade natural de aprimoramento em desenvolvimentos futuros.

5.3 RESULTADOS EXPERIMENTAIS

Em consonância com a prática regularmente adotada na literatura, dois estudos de caso tradicionais foram tomados como base para a validação do método proposto. Em ambos, o modelo de um diodo é empregado para representar o comportamento elétrico de geradores fotovoltaicos distintos, exigindo-se a extração dos seus cinco parâmetros a partir das respectivas curvas $I-V$ disponíveis. Em particular, esses estudos de caso são caracterizados como descrito a seguir, sendo seus dados reportados na Tabela 17 [63].

1. *Célula fotovoltaica RTC France* — dispositivo de silício cristalino ($N_p = N_s = 1$), cuja curva $I-V$ é caracterizada por $N = 26$ pares de medição obtidos sob temperatura de operação de 33°C ($T_c = 306,15\text{ K}$) [63].
2. *Módulo fotovoltaico Photowatt-PWP201* — dispositivo composto por uma fileira de 36 células de silício policristalino conectadas em série ($N_p = 1$ e $N_s = 36$), cuja curva $I-V$ é caracterizada por $N = 26$ pares de medição obtidos sob temperatura de operação de 45°C ($T_c = 318,15\text{ K}$) [63].

Ressalta-se que o primeiro estudo de caso é o mesmo abordado no âmbito da análise compreensiva conduzida no Capítulo 2 e, portanto, amplamente difundido na presente área de pesquisa. O segundo, por sua vez, foi reportado no mesmo trabalho seminal que introduz o primeiro, o que lhe confere notoriedade comparável.

O problema de extração dos parâmetros associado a cada um desses estudos de caso foi formulado de acordo com (3.15). A função objetivo, por sua vez, foi definida a partir da abordagem formal do modelo de um diodo via função W de Lambert, conforme (3.13). Nesse sentido, o *método preditor-corretor autoiniciado* proposto no capítulo anterior foi empregado para realizar todos os cálculos requeridos da função W de Lambert. Ademais, registra-se que todos os pontos que caracterizam as curvas $I-V$ correspondentes a cada estudo de caso foram considerados na montagem da função objetivo.

Quanto aos limites inferiores e superiores exigidos para a composição das restrições de cada problema, foram considerados os valores indicados na Tabela 18, em concordância com aqueles adotados por trabalhos correlatos. Como explicitado no Capítulo 2, embora tais intervalos possam ser interpretados como relaxados, estes ao menos se mostram funcionais.

Visando limitar o escopo desses testes de validação, foram consideradas nas avaliações realizadas apenas estimativas iniciais já reportadas formalmente para cada problema de extração

Tabela 17 – Pares medidos de tensão e corrente das curvas características de cada estudo de caso.

Medição	Célula RTC France		Photowatt-PWP201	
	Tensão (V)	Corrente (A)	Tensão (V)	Corrente (A)
1	-0,2057	0,7640	-1,9426	1,0345
2	-0,1291	0,7620	0,1248	1,0315
3	-0,0588	0,7605	1,8093	1,0300
4	0,0057	0,7605	3,3511	1,0260
5	0,0646	0,7600	4,7622	1,0220
6	0,1185	0,7590	6,0538	1,0180
7	0,1678	0,7570	7,2364	1,0155
8	0,2132	0,7570	8,3189	1,0140
9	0,2545	0,7555	9,3097	1,0100
10	0,2924	0,7540	10,2163	1,0035
11	0,3269	0,7505	11,0449	0,9880
12	0,3585	0,7465	11,8018	0,9630
13	0,3873	0,7385	12,4929	0,9255
14	0,4137	0,7280	13,1231	0,8725
15	0,4373	0,7065	13,6983	0,8075
16	0,4590	0,6755	14,2221	0,7265
17	0,4784	0,6320	14,6995	0,6345
18	0,4960	0,5730	15,1346	0,5345
19	0,5119	0,4990	15,5311	0,4275
20	0,5265	0,4130	15,8929	0,3185
21	0,5398	0,3165	16,2229	0,2085
22	0,5521	0,2120	16,5241	0,1010
23	0,5633	0,1035	16,7987	-0,0080
24	0,5736	-0,0100	17,0499	-0,1110
25	0,5833	-0,1230	17,2793	-0,2090
26	0,5900	-0,2100	17,4885	-0,3030

Tabela 18 – Limites estabelecidos para os parâmetros do modelo de um diodo em cada estudo de caso.

Parâmetro	Unidade	RTC France		Photowatt-PWP201	
		lb	ub	lb	ub
I_{ph}	A	0	1	0	2
I_0	μA	0	1	0	50
n	—	1	2	1	50
R_{sh}	Ω	0	100	0	2000
R_s	Ω	0	0,5	0	2

Tabela 19 – Estimativas iniciais consideradas para cada estudo de caso.

Item	Referência	I_{ph} (A)	I_0 (μA)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)
1	[53]	0,761060	0,290125	1,470097	49,973561	0,036800	8,8437E-4
2	[54]	0,760788	0,3106845	1,477105	52,890463	0,036547	7,7301E-4
3	[64]	0,76074014	0,31285196	1,4777295	55,907380	0,036615485	7,9439241E-4
4	[65]	0,7607191	0,3319572	1,4837043	54,2878608	0,0361310	8,0737695E-4
5	[67]	0,76078734	0,31075161	1,4778224	52,909893	0,036547082	7,730072975E-4
1	[53]	1,033537	2,825571	1,329426	689,321	1,224053	2,1547E-03
2	[54]	1,032345	2,515158	1,317240	747,943156	1,238972	2,04680E-03
3	[64]	1,0335685	2,2709763	1,3069558	687,87337	1,2599674	2,1722792E-03
4	[65]	1,0335104	2,6257110	1,3216167	683,5180201	1,2246947	2,1726166E-03
5	[67]	1,0323418	2,5162508	1,3179039	748,46159	1,2389511	2,0467844E-03

dos parâmetros e que se mostraram adequadas à aplicação de métodos determinísticos locais. Para fins de registro, essas inicializações são explicitadas na Tabela 19. Cabe esclarecer que, no caso da célula RTC France, as soluções iniciais coincidem com aquelas previamente reportadas na Tabela 9, sendo reapresentadas aqui apenas por conveniência.

Dessa forma, o TEPM foi aplicado à resolução dos problemas de extração dos parâmetros associados a esses dois estudos de caso de maneira repetida, considerando-se, em cada execução, uma das respectivas soluções iniciais indicadas na Tabela 19. As implementações computacionais foram desenvolvidas em ambiente MATLAB, versão 2016b, e todas as simulações foram executadas em um computador equipado com processador Intel Core Ultra 7 155H @ 3,8 GHz, 32 GB de memória RAM e sistema operacional Windows 10 (64 bits).

Para fins de comparação, também foram empregados na resolução desses estudos de caso métodos determinísticos locais disponibilizados nativamente pelo MATLAB. Em particular, considerou-se o IPM implementado na função `fmincon`, por constituir o principal representante dos métodos de barreira, bem como o TRR disponibilizado na função `lsqnonlin`, reconhecido como alternativa de estado da arte para esse problema, conforme evidenciado no Capítulo 2 [187, 189]. Esclarece-se que o LMA, igualmente disponível em `lsqnonlin`, não foi considerado, pois, na versão utilizada do MATLAB, sua implementação não oferece suporte direto a restrições de canalização, sendo destinada apenas a problemas irrestritos, o que inviabilizaria uma comparação metodologicamente consistente.

Cumprir registrar que, sob as configurações padrão desses *solvers*, os resultados inicialmente obtidos, tanto para o IPM quanto para o TRR, mostraram-se insatisfatórios, tendo sido observados, inclusive, comportamentos divergentes em determinadas execuções. Diante desse cenário, procedeu-se a uma calibração sistemática de suas opções internas, com o objetivo de adequá-las às especificidades inerentes ao problema de extração dos parâmetros. Nesse processo, um dos aspectos centrais investigados foi a metodologia de cálculo das derivadas, verificando-se melhor desempenho com a adoção de diferenças finitas centradas. Adicionalmente, constatou-se ganho de estabilidade ao se fornecer explicitamente a solução inicial como valor típico das variáveis de decisão, o que contribuiu para maior consistência numérica ao longo das iterações.

Além disso, salienta-se que o TEPM compartilha critérios de convergência, de estagnação e de parada semelhantes àqueles adotados internamente nos *solvers* do MATLAB. Assim, com o intuito de tornar a comparação dos resultados razoavelmente justa, tanto as execuções do TEPM quanto as dos demais métodos considerados foram conduzidas com parâmetros de controle harmonizados, tomando como referência os valores apresentados na Tabela 16. A esse respeito, ressalta-se que as tolerâncias dos critérios de convergência foram estabelecidas de forma rigorosa, ao passo que as tolerâncias dos critérios de parada foram mantidas suficientemente permissivas, de modo a evitar interrupções prematuras e assegurar que cada método pudesse buscar a melhor solução para cada problema. A seguir, apresenta-se uma síntese dos resultados obtidos e das análises decorrentes.

5.3.1 Avaliação de acurácia

As soluções encontradas por cada método na resolução de cada problema de extração dos parâmetros, considerando-se as diferentes alternativas de inicialização previamente estabelecidas,

Tabela 20 – Soluções retornadas pelos métodos avaliados no caso da célula RTC France.

Inicialização	Método	I_{ph} (A)	I_0 (μ A)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)
1	TRR	0,760830	0,293963	1,471757	51,403097	0,036779	7,780958E-04
	IPM	0,760788	0,310715	1,477279	52,891746	0,036547	7,730063E-04
	TEPM	0,760788	0,310688	1,477270	52,890024	0,036547	7,730063E-04
2	TRR	0,760788	0,310682	1,477269	52,889604	0,036547	7,730063E-04
	IPM	0,760788	0,310698	1,477274	52,890465	0,036547	7,730063E-04
	TEPM	0,760788	0,310689	1,477271	52,890039	0,036547	7,730063E-04
3	TRR	0,760787	0,311745	1,477611	52,965203	0,036532	7,730250E-04
	IPM	0,760788	0,310715	1,477279	52,891747	0,036547	7,730063E-04
	TEPM	0,760788	0,310689	1,477271	52,890076	0,036547	7,730063E-04
4	TRR	0,760770	0,326784	1,482361	54,016191	0,036326	7,772445E-04
	IPM	0,760788	0,310716	1,477279	52,891753	0,036547	7,730063E-04
	TEPM	0,760788	0,310690	1,477271	52,890130	0,036547	7,730063E-04
5	TRR	0,760788	0,310702	1,477275	52,891063	0,036547	7,730063E-04
	IPM	0,760787	0,310795	1,477305	52,909898	0,036546	7,730067E-04
	TEPM	0,760788	0,310690	1,477271	52,890132	0,036547	7,730063E-04

Tabela 21 – Soluções retornadas pelos métodos avaliados no caso do módulo Photowatt-PWP201.

Inicialização	Método	I_{ph} (A)	I_0 (μ A)	n	R_{sh} (Ω)	R_s (Ω)	RMSE (A)
1	TRR	1,032130	2,721092	1,325434	21,747134	0,034181	2,049035E-03
	IPM	1,032345	2,503679	1,316915	20,827777	0,034451	2,040002E-03
	TEPM	1,032358	2,496626	1,316629	20,786879	0,034460	2,039992E-03
2	TRR	1,032358	2,496694	1,316634	20,786249	0,034459	2,040002E-03
	IPM	1,032355	2,498053	1,316687	20,795115	0,034458	2,039993E-03
	TEPM	1,032358	2,496626	1,316629	20,786879	0,034460	2,039992E-03
3	TRR	1,032560	2,307613	1,308705	19,977871	0,034714	2,047064E-03
	IPM	1,032345	2,503678	1,316914	20,827771	0,034451	2,040002E-03
	TEPM	1,032358	2,496626	1,316629	20,786879	0,034460	2,039992E-03
4	TRR	1,032244	2,606502	1,321012	21,254698	0,034321	2,042098E-03
	IPM	1,032355	2,498053	1,316687	20,795117	0,034458	2,039993E-03
	TEPM	1,032358	2,496626	1,316629	20,786879	0,034460	2,039992E-03
5	TRR	1,032357	2,496975	1,316649	20,787206	0,034459	2,040053E-03
	IPM	1,032355	2,498053	1,316687	20,795116	0,034458	2,039993E-03
	TEPM	1,032358	2,496626	1,316629	20,786878	0,034460	2,039992E-03

são apresentadas nas Tabelas 20 e 21. Vale destacar que o nível de acurácia dessas soluções foi mensurado por meio da raiz do erro quadrático médio (RMSE), métrica tradicionalmente adotada em estudos relacionados ao desempenho do modelo de um diodo, conforme discutido no Capítulo 2. Mais precisamente, o seu cálculo foi realizado de acordo com a expressão apresentada em (2.12).

Ao analisar as soluções obtidas pelo TEPM para os parâmetros do modelo de um diodo em cada estudo de caso, observa-se que estas permanecem mutuamente compatíveis mesmo diante de diferentes estimativas iniciais. Esse comportamento indica baixa sensibilidade à escolha do ponto inicial e, conseqüentemente, um nível interessante de robustez do método proposto.

Na Tabela 20, em particular, verifica-se que as soluções retornadas pelo TEPM para o estudo de caso da célula RTC France mostraram-se idênticas àquelas identificadas como ótimas globais no âmbito da análise compreensiva conduzida no Capítulo 2. Esses resultados sugerem,

portanto, que o TEPM é capaz de retornar, de forma consistente, soluções de alta qualidade para o problema de extração dos parâmetros, mesmo sob distintas condições iniciais.

Em contraste, ao examinar os resultados obtidos por meio do IPM e TRR, verifica-se um comportamento menos uniforme. Para ambos os estudos de caso, as soluções retornadas apresentam maior dispersão entre si quando se variam as estimativas iniciais, o que evidencia maior sensibilidade ao ponto de partida e, por consequência, menor robustez. Ademais, observa-se que tais métodos nem sempre convergiram para as melhores soluções identificadas, retornando, em algumas execuções, soluções de qualidade inferior às aquelas fornecidas pelo TEPM, conforme refletido pelos valores de RMSE reportados.

Em síntese, os resultados obtidos corroboram que o TEPM cumpre o propósito de resolver o problema de extração dos parâmetros de maneira acurada e robusta. Além disso, as diferenças observadas em relação ao desempenho do IPM reforçam a pertinência do desenvolvimento de um BM especializado, ao passo que os ganhos verificados frente ao TRR evidenciam o seu potencial de competitividade.

5.3.2 Avaliação de eficiência

Os indicadores de eficiência computacional considerados neste estudo foram o número de iterações (IT), o número de avaliações da função objetivo (FE) e o tempo médio total de simulação (TC) requerido até a convergência. Em particular, o tempo médio total foi estimado a partir de 100 execuções independentes de cada método, na resolução de cada problema e respectiva inicialização, com o intuito de mitigar efeitos de variabilidade inerentes ao ambiente computacional utilizado. Os resultados correspondentes são apresentados na Tabela 22.

Ao analisar esses indicadores, observa-se inicialmente que o TEPM apresentou, em termos gerais, os maiores valores para o número de avaliações da função objetivo. Em diversos casos, o método demandou milhares de avaliações, ao passo que o IPM e o TRR atingiram soluções de qualidade comparável com algumas dezenas ou centenas. Naturalmente, esse maior esforço de avaliação refletiu-se diretamente no tempo médio total de simulação, elevando o custo computacional associado ao TEPM em relação aos demais métodos. Cabe notar, contudo, que tal comportamento era, em certa medida, esperado, em função da metodologia de cálculo das derivadas adotada na implementação do TEPM, em particular no que se refere ao tratamento e à atualização da matriz Hessiana.

Por outro lado, nota-se que o TEPM tipicamente requer um número razoável de iterações, o qual é consistentemente inferior ao observado para o IPM. Isso sugere que parte do ônus computacional está concentrada no custo por iteração, e não necessariamente na necessidade de longas sequências iterativas. Assim, ao se empregar uma metodologia de cálculo de derivadas mais eficiente, há margem concreta para ganhos de eficiência, o que pode ampliar a competitividade do método nesse quesito.

Tabela 22 – Indicadores de eficiência computacional verificados para cada método avaliado nos estudos de caso considerados.

Inicialização	Método	RTC France			Photowatt-PWP201		
		IT	FE	TC (s)	IT	FE	TC (s)
1	TRR	3	44	0,000031	4	55	0,000033
	IPM	53	639	0,000373	88	1039	0,000586
	TEPM	11	1359	0,019093	18	2186	0,027681
2	TRR	2	33	0,000027	2	33	0,000026
	IPM	29	400	0,000305	78	934	0,000590
	TEPM	10	1280	0,016433	13	1606	0,020365
3	TRR	2	33	0,000026	3	44	0,000028
	IPM	41	514	0,000334	85	1008	0,000600
	TEPM	8	1023	0,013071	12	1529	0,019145
4	TRR	3	44	0,000029	2	33	0,000025
	IPM	50	614	0,000420	96	1140	0,000679
	TEPM	13	1593	0,019870	17	2041	0,025354
5	TRR	2	33	0,000025	2	33	0,000026
	IPM	34	483	0,000406	85	1015	0,000622
	TEPM	13	1578	0,020058	14	1748	0,021894

Por fim, destaca-se que o TRR apresentou, de modo geral, os melhores indicadores de eficiência dentre os métodos avaliados. Ainda assim, como discutido na subseção anterior, tal desempenho não se traduziu, em todas as execuções, na obtenção das melhores soluções conhecidas.

5.4 CONCLUSÕES PARCIAIS

Este capítulo apresentou o método determinístico local proposto para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico. A revisão bibliográfica conduzida no Capítulo 2 indicou que métodos de escopo amplo, voltados à resolução de problemas não lineares restritos de estrutura genérica, ainda não haviam sido aplicados de forma sistemática a esse problema e que, dentre as alternativas dessa categoria, os BMs mostravam-se particularmente alinhados à sua natureza. Uma investigação foi então realizada a respeito desse grupo, da qual emergiu o IPM como candidato proeminente. Todavia, uma análise de suas fundamentações teóricas evidenciou fragilidades potenciais frente às características específicas do problema, o que motivou a exploração de paradigmas alternativos dentro da mesma classe. Esse percurso culminou na adoção do EPM como base mais adequada e promissora para o desenvolvimento da abordagem proposta.

Na sequência, a forma de aplicação e a estrutura lógica do EPM foram estabelecidas em consonância com os aspectos conceituais e operacionais verificados como usuais na literatura específica. Contudo, em face das dificuldades mapeadas no Capítulo 3, foram realizadas deliberações para que sua utilização fosse eficaz no escopo do problema de extração dos parâmetros. Em primeiro lugar, a barreira logarítmica modificada, característica desse paradigma, lida diretamente com implicações trazidas pelas restrições do problema, estruturando e condicionando

o domínio primal desde a origem. Em segundo lugar, o mecanismo de regularização inercial incorporado mitiga os efeitos da natureza não convexa e preserva o avanço do processo iterativo. Por fim, o recurso de escalonamento simétrico introduzido reforça a robustez numérica frente às discrepâncias de escala entre variáveis e ao mal condicionamento das matrizes envolvidas. O método resultante foi, portanto, concebido sob medida para esse problema, sendo denominado *tailored exterior-point method* (TEPM).

A validação foi conduzida em dois estudos de caso clássicos da literatura, adotando-se a formulação via função W de Lambert e empregando-se o método preditor–corretor autoiniciado desenvolvido no capítulo anterior. As soluções retornadas, sob diferentes inicializações, mostraram-se consistentes com aquelas identificadas como ótimas globais na literatura, evidenciando a capacidade do TEPM de retornar, de forma robusta, soluções de alta qualidade para o problema de extração dos parâmetros. O desempenho superior frente ao IPM reforçou a pertinência do desenvolvimento de um BM concebido sob medida para esse problema. Ademais, o maior nível de acurácia verificado em relação ao TRR evidenciou seu potencial de competitividade. Em termos de eficiência, observou-se que o maior tempo computacional em relação aos demais métodos considerados esteve associado sobretudo ao custo por iteração imposto pelo cálculo numérico das derivadas. Ainda assim, o número de iterações requerido não se mostrou demasiado, indicando margem concreta para ganhos adicionais de eficiência mediante estratégias de derivadas mais econômicas.

Diante do exposto, o TEPM consolida-se como uma contribuição original desta tese, tanto pela arquitetura delineada a partir do diagnóstico estabelecido no Capítulo 3 quanto pelo compromisso alcançado entre acurácia, robustez e viabilidade prática. Adicionalmente, os resultados obtidos sustentam seu posicionamento como alternativa técnica competitiva frente a métodos consolidados da literatura, reforçando a pertinência e o potencial de aplicabilidade da abordagem desenvolvida.

Registra-se, ainda, que o conteúdo técnico aqui apresentado encontra-se em fase final de consolidação em um manuscrito, intitulado “*Parameter extraction of the photovoltaic single diode model using a tailored exterior-point method*”, conforme antecipado no Capítulo 1.

6 APLICAÇÃO PRÁTICA

Este capítulo apresenta uma aplicação em larga escala dos métodos desenvolvidos nos capítulos anteriores. Nesse contexto, a Seção 6.1 descreve os principais aspectos do estudo realizado, incluindo os dados utilizados, a formulação dos problemas de extração dos parâmetros e o esquema computacional empregado em suas resoluções. A Seção 6.2, por sua vez, detalha os procedimentos adotados para o estabelecimento dos limites inferiores e superiores dos parâmetros do modelo de um diodo em cada problema, bem como o método utilizado para obter as estimativas iniciais requeridas à resolução. Na Seção 6.3, apresentam-se, de forma sintética, os resultados obtidos no âmbito do estudo, discutindo-se as suas implicações. Por fim, a Seção 6.4 sumariza o conteúdo exposto ao longo do capítulo, explicitando as conclusões alcançadas acerca do arcabouço metodológico proposto ao longo desta tese.

6.1 CARACTERIZAÇÃO DO ESTUDO

Nos capítulos anteriores, foram apresentados: *(i)* o método preditor–corretor autoiniciado proposto para o cálculo eficaz e eficiente da função W de Lambert; e *(ii)* o método de ponto exterior concebido sob medida para a resolução acurada e robusta do problema de extração dos parâmetros, o qual atende por TEPM. Em consonância com as práticas adotadas na literatura, ambos os métodos foram validados por meio de estudos de caso representativos, de modo que os resultados obtidos evidenciaram que cumprem os objetivos a que se propõem e se mostram competitivos frente a alternativas de estado da arte.

Não obstante, embora tais testes sejam fundamentais para aferir a efetividade, a pertinência e o desempenho potencial dos métodos propostos, o vislumbre de seu emprego em situações cotidianas demanda avaliações adicionais em cenários mais desafiadores e mais próximos de condições práticas de aplicação. Essa ampliação do escopo de testes é necessária para produzir evidências mais conclusivas e, sobretudo, para subsidiar de maneira objetiva a adoção desses métodos na resolução de problemas de extração dos parâmetros nos diferentes contextos do mundo real.

Nesse sentido, apresenta-se neste capítulo um estudo em larga escala, no qual a aplicação dos métodos propostos é submetida a um conjunto massivo de problemas de extração dos parâmetros, sob condições numéricas variadas e elevada carga computacional. Em particular, nesta seção são apresentados os principais aspectos estruturantes desse estudo.

6.1.1 Dados utilizados

Os dados utilizados no presente estudo advêm de um repositório público disponibilizado pelo NREL, concebido especificamente para apoiar a validação e o desenvolvimento de modelos de desempenho de geradores fotovoltaicos [49].

Esse repositório foi estabelecido há pouco mais de uma década e, em linhas gerais, contempla medições de campo para 33 módulos fotovoltaicos, representativos de diferentes tecnologias comerciais disponíveis à época, incluindo: silício monocristalino (xSi), silício multicristalino (mSi), telureto de cádmio (CdTe), seleneto de cobre, índio e gálio (CIGS), bem como módulos de heterojunção (HIT) e variantes baseadas em silício amorfo (aSi) [49].

A campanha de aquisição dessas medições foi conduzida em localidades com condições climáticas contrastantes no território continental dos Estados Unidos da América: Cocoa, no estado da Flórida (FL), com clima subtropical; Eugene, no estado do Oregon (OR), com clima oceânico temperado; e Golden, no estado do Colorado (CO), com clima semiárido [49]. Nessas localidades, as medições foram realizadas ao ar livre, em estrutura fixa do tipo “*open-rack*”, com registro sistemático de variáveis meteorológicas, elétricas e térmicas por períodos de aproximadamente um ano em cada localidade [49]. Ademais, procedimentos de garantia de qualidade foram aplicados no pós-processamento a fim de remover registros potencialmente não confiáveis; por essa razão, as séries disponibilizadas não são serialmente completas [49].

Os registros operacionais consolidados para o conjunto de módulos compõem um acervo de exatamente 1.025.599 curvas $I-V$, correspondentes a condições amplamente variadas de irradiância solar incidente e temperatura de operação [49]. Ademais, cada curva $I-V$ é descrita por um conjunto de pares de medições de tamanho variável, compreendendo, no caso mais reduzido, 150 pontos e, no mais amplo, 295 pontos [49].

Cumprir ponderar que os módulos fotovoltaicos e os equipamentos utilizados na concepção desse repositório correspondem, em grande medida, a gerações tecnológicas anteriores. Ainda assim, trata-se, até o presente, de um dos acervos de dados públicos mais valiosos e amplamente disponíveis para a condução de estudos relacionados à modelagem de geradores fotovoltaicos [49]. Dentre os diversos contextos em que esse repositório pode ser explorado, inclui-se o problema de extração dos parâmetros. De fato, alguns dos trabalhos correlatos que empregam métodos determinísticos locais adotam esse repositório como base para avaliações mais abrangentes, a saber [54], [64] e [67]. Portanto, a escolha desse acervo para a caracterização do estudo em larga escala aqui delineado mostra-se apropriada.

6.1.2 Formulação do problema

Para cada curva $I-V$ disponibilizada no referido repositório, formulou-se um problema particular de extração dos parâmetros. À luz das discussões conduzidas ao longo dos capítulos anteriores, esses problemas foram estabelecidos de acordo com (3.15).

Mais especificamente, a função objetivo foi definida com base na abordagem formal do modelo de um diodo via função W de Lambert, conforme (3.13). Nesse contexto, o método preditor-corretor autoiniciado foi empregado para realizar os cálculos requeridos dessa função especial, seguindo o Algoritmo 2. Ademais, registra-se que todos os pontos que caracterizam cada curva $I-V$ foram considerados na construção da função objetivo do problema correspondente.

No que concerne às restrições de canalização, ressalta-se que os limites inferiores e superiores empregados em sua composição foram estabelecidos por meio de uma formulação especificamente desenvolvida para esse fim. Uma vez que tal formulação integra, em caráter subsidiário, o arcabouço metodológico proposto nesta tese, ela é apresentada em destaque na seção subsequente.

Nessa mesma linha, as estimativas iniciais requeridas para a resolução de cada problema de extração dos parâmetros formulado foram adaptadas de um trabalho correlato, sendo maiores detalhes também apresentados na seção subsequente.

6.1.3 Recurso computacional

As implementações computacionais necessárias ao processamento dos dados do repositório, à declaração dos problemas de extração dos parâmetros, à sua resolução e ao posterior tratamento dos resultados foram desenvolvidas em ambiente MATLAB, versão 2016b. Todas as simulações foram executadas em um computador equipado com processador Intel Core Ultra 7 155H @ 3,8 GHz, 32 GB de memória RAM e sistema operacional Windows 10 (64 bits). Embora o processamento em paralelo seja possível e, em aplicações práticas, recomendável diante do volume de problemas, tal estratégia não foi considerada no âmbito do presente estudo; conseqüentemente, cada instância foi resolvida de forma sequencial.

6.1.4 Metodologia de avaliação

O TEPM foi implementado conforme o Algoritmo 6, adotando-se os parâmetros de controle indicados na Tabela 16. Além disso, seguiu-se a estratégia híbrida de cálculo de derivadas discutida na Seção 5.2 do capítulo anterior: o vetor gradiente da função objetivo do problema foi internamente computado pelo método do passo complexo, enquanto a sua matriz Hessiana foi aproximada por diferenças finitas centradas.

Uma vez que o MATLAB foi adotado como ambiente de implementação e execução do estudo, seus *solvers* nativos foram também explorados com o objetivo de viabilizar avaliações comparativas. Nesse sentido, além do TEPM, consideraram-se métodos determinísticos locais disponibilizados nativamente pelo *software*. Em particular, empregou-se o IPM implementado na função `fmincon`, por constituir o principal representante da classe de métodos de barreira, bem como o TRR disponibilizado na função `lsqnonlin`, reconhecido como alternativa de estado da arte para o problema de extração dos parâmetros, conforme o levantamento bibliográfico realizado no Capítulo 2 [187, 189]. Ressalta-se que tais *solvers* foram configurados com parâmetros de controle harmonizados, em linha com os critérios e escolhas discutidos na Seção 5.3 do capítulo anterior, de modo a assegurar comparabilidade entre os desempenhos observados.

Com vistas a possibilitar uma análise abrangente, foram armazenados, para cada instância do problema, os dados extraídos do repositório (curva $I-V$ e variáveis associadas), os limites inferiores e superiores estabelecidos, as estimativas iniciais consideradas e as soluções retor-

nadas por cada método. Complementarmente, registraram-se métricas essenciais de avaliação, incluindo o erro quadrático médio (RMSE), calculado conforme a expressão (2.12), introduzida no Capítulo 2.

6.2 PROCEDIMENTOS AUXILIARES

Conforme estabelecido, os desenvolvimentos apresentados ao longo desta tese concentraram-se na concepção e na análise de métodos numéricos voltados ao cálculo da função W de Lambert e à resolução do problema de extração dos parâmetros do modelo de um diodo. No entanto, a aplicação sistemática desse arcabouço — sobretudo em estudos de larga escala — requer a definição criteriosa de condições estruturais do problema, as quais influenciam diretamente a viabilidade numérica, a robustez do processo iterativo e a interpretabilidade física das soluções obtidas.

À luz das discussões conduzidas no Capítulo 3, a formulação do problema de extração dos parâmetros envolve restrições explícitas e implícitas que não podem ser negligenciadas. Em particular, impõe-se o estabelecimento de limites admissíveis para os parâmetros do modelo de um diodo, de modo a evitar soluções fisicamente inconsistentes e, simultaneamente, prevenir extrapolações para regiões que não pertencem ao domínio real associado ao ramo principal da função W de Lambert. Na ausência dessas salvaguardas, podem ocorrer inviabilidades numéricas ou estimativas destituídas de significado físico, comprometendo análises subsequentes, as quais são fundamentais, por exemplo, em estudos práticos sobre o impacto da degradação do gerador em seu desempenho ou na identificação de faltas em seu monitoramento operacional [25–31].

Adicionalmente, o emprego de métodos determinísticos locais requer o fornecimento de estimativas iniciais de qualidade razoável. A escolha dessas estimativas influencia decisivamente a convergência a soluções de interesse, sobretudo diante da possível presença de mínimos locais em situações práticas. Assim, a etapa de inicialização deve ser tratada como componente estratégico do processo, com impacto direto tanto na robustez quanto na eficiência computacional.

Posto isto, apresentam-se a seguir as formulações adotadas no presente estudo para a definição das restrições de canalização e para a obtenção de estimativas iniciais, as quais servem como referência prática para aplicações subsequentes dos métodos propostos. Ressalta-se que, em conjunto, tais procedimentos consolidam condições adequadas para a aplicação sistemática do arcabouço metodológico desenvolvido nesta tese na resolução de problemas de extração dos parâmetros envolvendo diferentes tecnologias fotovoltaicas e variadas condições de operação.

6.2.1 Restrições de canalização

Conforme discutido no Capítulo 2, os trabalhos correlatos que empregam de maneira consistente métodos determinísticos locais para a resolução do problema de extração dos parâmetros tipicamente não incorporam restrições de canalização na formulação do problema. Em

consequência, não há referências diretas das quais possam ser extraídas formulações operacionais para o estabelecimento de limites mínimos e máximos para os parâmetros do modelo de um diodo.

Nos trabalhos que abordam métodos determinísticos globais, metaheurísticos e conjugados, por outro lado, restrições de canalização são usualmente adotadas. Todavia, como tais métodos são, em geral, avaliados apenas no contexto de estudos de caso difundidos na literatura, observa-se que a definição dos limites inferior e superior tende a ser conduzida de maneira arbitrária. Em particular, no conjunto de trabalhos revisados no âmbito desta pesquisa, não foi identificada uma metodologia sistemática, generalizável e operacionalmente orientada para o estabelecimento desses limites em bases de dados extensas e heterogêneas.

Em um escopo mais amplo, apenas um trabalho foi efetivamente identificado propondo formulações práticas e bem fundamentadas para a determinação de limites admissíveis para os parâmetros do modelo de um diodo, a saber [139]. Não obstante, as expressões propostas em [139] fundamentam-se em propriedades geométricas específicas de curvas $I-V$ observadas sob STC ou, sobretudo, sob níveis mais elevados de irradiância solar incidente. Em testes preliminares realizados no contexto do presente estudo, verificou-se que tais expressões não fornecem limites adequados para uma parcela relevante das curvas $I-V$, em especial para aquelas associadas a condições de operação mais adversas, como regimes de baixa irradiância, nos quais a forma da curva apresenta um caráter monotonicamente decrescente mais acentuado, comprometendo a robustez das relações geométricas exploradas.

Diante desse diagnóstico, realizou-se uma investigação complementar, abrangendo diferentes vertentes da literatura sobre o modelo de um diodo, com o objetivo de identificar valores usuais e aproximações consagradas de sua equação característica que pudessem ser explorados na definição de limites admissíveis para os cinco parâmetros. A partir desse levantamento, propôs-se a formulação apresentada em (6.1) para a determinação dos limites inferiores, bem como em (6.2) para a determinação dos limites superiores.

$$I_{ph,\min} := \frac{I_{sc}^2}{I_{ph,\max}} \quad (6.1a)$$

$$I_{0,\min} := 0 \quad (6.1b)$$

$$n_{\min} := 10^{-30} \quad (6.1c)$$

$$R_{sh,\min} := \max \left\{ \frac{V_{xx}}{I_{sc} - I_{xx}} - R_{s,\max}, 10^{-30} \right\} \quad (6.1d)$$

$$R_{s,\min} := 0 \quad (6.1e)$$

$$I_{ph,\max} := \max \left\{ I_{sc} \left(1 + \frac{R_{s,\max}}{R_{sh,\min}} \right), 1,5 I_{sc} \right\} \quad (6.2a)$$

$$I_{0,\max} := I_{ph,\max} \exp \left(- \frac{q V_{oc}}{n_{\max} N_s k T_k} \right) \quad (6.2b)$$

$$n_{\max} := 5 \quad (6.2c)$$

$$R_{sh,\max} := 10^6 \quad (6.2d)$$

$$R_{s,\max} := \frac{V_{oc} - V_x}{I_x} \quad (6.2e)$$

Cabe esclarecer que as grandezas V_x e I_x correspondem a um ponto intermediário da curva situado à esquerda do ponto de máxima potência, enquanto V_{xx} e I_{xx} representam um ponto intermediário situado à direita desse ponto. Tais pontos são empregados com o propósito de capturar informações geométricas adicionais da curva experimental, permitindo inferências aproximadas sobre a inclinação local e, conseqüentemente, sobre limites plausíveis para R_s e R_{sh} (ver discussão na Seção 2.1). Além disso, destaca-se que a escolha de $n_{\max} = 5$ fundamenta-se em valores típicos reportados na literatura [139]. Já as demais constantes numéricas empregadas, como $R_{sh,\max} = 10^6$ e o piso 10^{-30} , foram adotadas com o objetivo de evitar degenerescências numéricas e preservar a viabilidade computacional do problema.

Por fim, ressalta-se que tal formulação não tem por objetivo constituir uma caracterização teórica exaustiva, tampouco estabelecer limites ideais sob qualquer perspectiva. Seu propósito central é definir intervalos que, na prática, sejam suficientemente informativos para evitar extrapolações numéricas indevidas, ao mesmo tempo em que permaneçam aderentes à realidade de diferentes tecnologias fotovoltaicas e a distintas condições de operação. Ademais, a construção de uma metodologia de canalização mais diretamente ancorada no comportamento físico — ou mesmo estatístico — dos parâmetros do modelo de um diodo constitui uma oportunidade natural de investigação futura, conforme já delineado no Capítulo 2.

6.2.2 Procedimento de inicialização

Conforme reiterado ao longo desta tese, trabalhos correlatos que empregam de maneira consistente métodos determinísticos locais para a resolução do problema de extração dos parâmetros frequentemente se concentram na proposição de procedimentos de inicialização. Diante desse acúmulo de contribuições, torna-se natural aproveitar uma dessas alternativas no presente estudo.

Nesse sentido, realizou-se uma análise preliminar das alternativas de inicialização disponíveis. A partir dessa análise, selecionou-se o procedimento proposto em [54]. Conforme exposto no Item 2.4.1.1 do Capítulo 2, tal procedimento fundamenta-se na resolução de dois subproblemas de otimização formulados a partir da chamada “aproximação educada”, sendo organizados por meio de uma *decomposição do tipo Benders generalizada*. Em linhas gerais, essa alternativa destacou-se pela consistência de sua fundamentação teórica e pela qualidade das estimativas iniciais retornadas, inclusive quando aplicada no âmbito do mesmo repositório de curvas I - V adotado no presente estudo, o que a torna particularmente atraente.

Adicionalmente, ressalta-se que o procedimento selecionado não foi empregado de forma estritamente idêntica à proposta original. Para assegurar compatibilidade entre as estimativas

iniciais e as restrições de canalização definidas em (6.1) e (6.2), o subproblema primário do procedimento foi resolvido já incorporando os limites estabelecidos para os parâmetros n e R_s . Considerando que o subproblema subsequente é resolvido de forma diretamente condicionada às estimativas obtidas na etapa primária, entende-se que a imposição de restrições apenas nesta é suficiente para induzir estimativas factíveis para os cinco parâmetros do modelo de um diodo.

Por fim, para fins de registro, salienta-se que outras alternativas de inicialização podem ser conceitualizadas tendo em vista o conteúdo documentado no Capítulo 2 e avaliadas em desenvolvimentos futuros. Por exemplo, à luz das discussões apresentadas na Seção 2.2, pode-se vislumbrar procedimentos dedicados ao uso exclusivo dos três pontos notáveis de uma curva $I-V$ (SC, OC e MP), os quais podem se revelar competitivos em regimes operacionais específicos. De modo complementar, a partir do levantamento apresentado na Subseção 2.4.3, torna-se igualmente possível conceber um arranjo no qual um método metaheurístico seja empregado de forma restrita à determinação de estimativas iniciais. Nesse caso, sob uma perspectiva mais ampla, obtém-se um método conjugado envolvendo o método determinístico local considerado, com estrutura análoga à do ABC-TRR proposto em [108], potencialmente resultando em ganhos adicionais de desempenho.

6.2.3 Validação preliminar

Neste ponto, torna-se pertinente evidenciar que os procedimentos adotados para o estabelecimento dos limites necessários às restrições de canalização e para a obtenção das estimativas iniciais mostram-se mutuamente compatíveis.

Para esse fim, apresenta-se na Tabela 23 o RMSE médio, denotado por $\overline{\text{RMSE}}$, verificado a partir das soluções retornadas pelo procedimento de inicialização para os problemas de extração dos parâmetros associados às curvas $I-V$ disponíveis de cada um dos 33 módulos fotovoltaicos do repositório, bem como o tempo médio de simulação correspondente.

Os resultados indicam que as estimativas iniciais produzidas pelo procedimento são efetivas. Ainda que este tenha sido aplicado considerando apenas o subproblema primário sob restrições (isto é, com n_{\min} , n_{\max} , $R_{s,\min}$ e $R_{s,\max}$ explicitamente impostos), verificou-se que as estimativas obtidas para cada parâmetro permaneceram compatíveis com os limites definidos em (6.1) e (6.2). Ademais, os valores de $\overline{\text{RMSE}}$ observados situam-se na ordem de grandeza de 10^{-2} A e 10^{-3} A, o que corrobora a boa qualidade dessas estimativas. De modo complementar, os tempos médios reportados indicam um custo computacional compatível com a aplicação em larga escala considerada.

Diante dessas constatações, depreende-se que os limites estabelecidos para as restrições de canalização pela formulação proposta são razoáveis e que o procedimento de inicialização selecionado se integra de forma consistente a esses limites, consolidando condições propícias para a aplicação sistemática do TEPM, bem como dos demais métodos considerados no âmbito do presente estudo. Além disso, a adoção de um procedimento de inicialização com esse nível

Tabela 23 – RMSE médio e tempo médio de simulação associados às estimativas iniciais fornecidas pelo procedimento de inicialização para os problemas de extração dos parâmetros das curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.

Módulo fotovoltaico	Local	Quantidade de curvas	RMSE (A)	TC (s)
CIGS39017	Cocoa - FL	34775	1,662662E-02	0,009520
CIGS8-001	Cocoa - FL	38939	5,036485E-03	0,007653
CdTe75638	Cocoa - FL	39080	5,322195E-03	0,008950
HIT05667	Cocoa - FL	38377	1,634777E-02	0,012909
aSiMicro03036	Cocoa - FL	39037	3,551475E-03	0,005442
aSiTandem72-46	Cocoa - FL	39186	5,872585E-03	0,006881
aSiTriple28324	Cocoa - FL	38485	9,964700E-03	0,008994
mSi0166	Cocoa - FL	36765	3,343358E-03	0,009274
mSi0188	Cocoa - FL	39102	3,401927E-03	0,009539
mSi460A8	Cocoa - FL	38929	1,361536E-02	0,013524
xSi12922	Cocoa - FL	38989	1,564775E-02	0,013241
CIGS39017	Eugene - OR	42674	1,148901E-02	0,008440
CIGS8-001	Eugene - OR	43146	4,032289E-03	0,007379
CdTe75638	Eugene - OR	42248	3,512698E-03	0,007966
HIT05667	Eugene - OR	43271	1,483331E-02	0,009367
aSiMicro03036	Eugene - OR	43343	4,024919E-03	0,004582
aSiTandem72-46	Eugene - OR	43266	5,510800E-03	0,005902
aSiTriple28324	Eugene - OR	42705	1,154204E-02	0,009629
mSi0166	Eugene - OR	43268	2,804299E-03	0,007311
mSi0188	Eugene - OR	43127	2,752894E-03	0,007424
mSi460A8	Eugene - OR	43115	9,089378E-03	0,010690
xSi12922	Eugene - OR	43185	1,017804E-02	0,010364
CIGS1-001	Golden - CO	12011	5,229501E-03	0,007801
CIGS39013	Golden - CO	11437	1,427030E-02	0,008852
CdTe75669	Golden - CO	11953	5,276116E-03	0,008141
HIT05662	Golden - CO	11876	1,727636E-02	0,013094
aSiMicro03038	Golden - CO	12148	4,518076E-03	0,004696
aSiTandem90-31	Golden - CO	12070	8,881059E-03	0,005336
aSiTriple28325	Golden - CO	11445	1,721319E-02	0,007730
mSi0247	Golden - CO	11912	4,158850E-03	0,008142
mSi0251	Golden - CO	11887	3,991583E-03	0,008387
mSi460BB	Golden - CO	11919	2,218687E-02	0,010037
xSi11246	Golden - CO	11929	1,451449E-02	0,012531

de desempenho torna os testes subsequentes mais informativos: caso o TEPM, ou algum dos demais métodos considerados, não retorne soluções acuradas mesmo quando inicializado a partir de estimativas de elevada qualidade, obtém-se um indicativo objetivo de limitação do método avaliado nas condições consideradas.

6.3 SÍNTESE DOS RESULTADOS

Uma vez caracterizado o estudo em larga escala, estabelecidos os procedimentos de canalização e de inicialização e definidos os critérios de avaliação, apresentam-se nesta seção uma síntese dos principais resultados obtidos. Salienta-se que, em respeito ao escopo delimitado para a presente pesquisa, a exposição é conduzida de maneira objetiva, com foco na verificação da qualidade das soluções retornadas e do desempenho computacional observado. Uma exploração mais abrangente do comportamento dos parâmetros estimados — incluindo interpretações físico-operacionais e análises estatísticas aprofundadas — é deliberadamente reservada para momento oportuno, podendo constituir desdobramento natural em trabalhos futuros.

6.3.1 Níveis de acurácia

A Tabela 24 apresenta o $\overline{\text{RMSE}}$ verificado a partir das soluções retornadas pelo TEPM para os problemas de extração dos parâmetros associados às curvas $I-V$ disponíveis para cada um dos 33 módulos fotovoltaicos do repositório.

A princípio, chama atenção o fato de os valores de $\overline{\text{RMSE}}$ obtidos com o TEPM situarem-se tipicamente na ordem de 10^{-3} A, atingindo, em casos, patamares da ordem de 10^{-4} A, em linha com os estudos de caso reportados no capítulo anterior. Esse resultado evidencia elevada qualidade média das soluções retornadas. Ao comparar esses resultados com aqueles apresentados na Tabela 23, observa-se ainda uma redução sistemática do $\overline{\text{RMSE}}$ para todos os módulos fotovoltaicos avaliados, chegando, em alguns casos, a uma diminuição de até 7 vezes (por exemplo, nos módulos identificados como CdTe75638 e xSi11246). Em conjunto, esse comportamento, aliado aos baixos valores médios observados, corrobora a robustez e o elevado nível de acurácia do TEPM.

Para fins de comparação, a Tabela 24 apresenta adicionalmente o $\overline{\text{RMSE}}$ verificado a partir das soluções retornadas pelo IPM, assim como pelo TRR. Embora esses métodos também tenham se mostrado capazes de retornar soluções de boa qualidade, observa-se que os ganhos obtidos foram, em geral, mais modestos. Em particular, o IPM apresentou, de forma recorrente, os maiores valores de $\overline{\text{RMSE}}$. O TRR, por sua vez, apresentou soluções de qualidade intermediária e, em 3 dos 33 casos avaliados, foi capaz de retornar os menores valores médios de $\overline{\text{RMSE}}$.

Diante dessas evidências, constata-se que o TEPM cumpre efetivamente o objetivo proposto. Para 30 dos 33 módulos fotovoltaicos, esse método foi capaz de retornar, em média, as soluções mais acuradas e fisicamente coerentes. Mesmo nos 3 casos restantes, as soluções obtidas permaneceram consistentes e exibiram qualidade superior àquela associada às estimativas iniciais, indicando progresso sistemático. Ademais, o desempenho particularmente superior em relação ao IPM corrobora a pertinência do desenvolvimento de um BM concebido sob medida para o problema de extração dos parâmetros.

Por fim, ressalta-se que os 3 casos em que o TEPM não apresentou o menor $\overline{\text{RMSE}}$ foram examinados em maior profundidade. Verificou-se que o principal fator limitante esteve associado ao atingimento do número máximo de avaliações permitido para a função objetivo e, portanto, a uma interrupção forçada do processo iterativo. Esse efeito relaciona-se diretamente à abordagem híbrida de cálculo de derivadas adotada na implementação, a qual eleva o custo por iteração em termos de avaliações da função objetivo. Assim, uma vez identificada e incorporada uma alternativa igualmente eficaz, porém menos custosa sob esse quesito, há margem concreta para ganhos adicionais de acurácia nessas situações específicas, permitindo que o TEPM explore de forma mais plena o seu potencial.

Tabela 24 – Comparação do RMSE médio, dado em ampères, obtido pelos métodos avaliados na resolução dos problemas de extração dos parâmetros associados às curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.

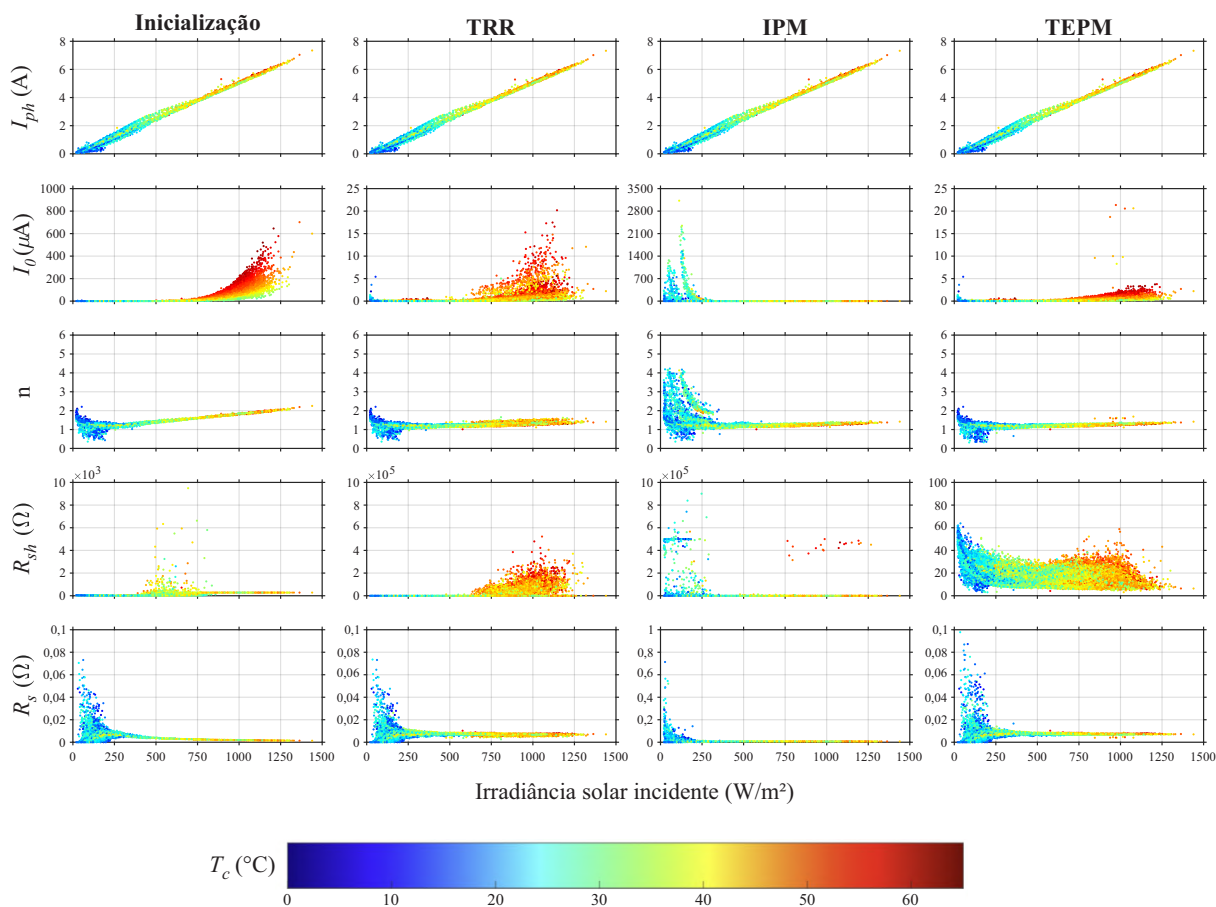
Módulo fotovoltaico	Local	Quantidade de curvas	TRR	IPM	TEPM
CIGS39017	Cocoa - FL	34775	5,420051E-03	8,335627E-03	5,371362E-03
CIGS8-001	Cocoa - FL	38939	3,579049E-03	3,803606E-03	3,578680E-03
CdTe75638	Cocoa - FL	39080	7,946890E-04	1,825751E-03	7,822160E-04
HIT05667	Cocoa - FL	38377	4,579217E-03	6,630222E-03	4,311710E-03
aSiMicro03036	Cocoa - FL	39037	6,144664E-04	8,619191E-04	5,845638E-04
aSiTandem72-46	Cocoa - FL	39186	1,593915E-03	2,261719E-03	1,577351E-03
aSiTriple28324	Cocoa - FL	38485	9,585731E-03	9,589757E-03	9,894231E-03
mSi0166	Cocoa - FL	36765	2,720184E-03	4,014597E-03	2,718203E-03
mSi0188	Cocoa - FL	39102	2,153264E-03	3,469423E-03	2,150735E-03
mSi460A8	Cocoa - FL	38929	3,150983E-03	4,112932E-03	2,935741E-03
xSi12922	Cocoa - FL	38989	2,828562E-03	3,998026E-03	2,427803E-03
CIGS39017	Eugene - OR	42674	4,434618E-03	6,497308E-03	4,426041E-03
CIGS8-001	Eugene - OR	43146	3,655623E-03	3,873163E-03	3,654326E-03
CdTe75638	Eugene - OR	42248	6,819476E-04	1,385201E-03	6,737977E-04
HIT05667	Eugene - OR	43271	6,461188E-03	1,015942E-02	5,869976E-03
aSiMicro03036	Eugene - OR	43343	8,216838E-04	1,192868E-03	8,165496E-04
aSiTandem72-46	Eugene - OR	43266	1,788900E-03	2,092909E-03	1,735555E-03
aSiTriple28324	Eugene - OR	42705	9,449317E-03	9,462510E-03	1,066707E-02
mSi0166	Eugene - OR	43268	2,467689E-03	3,849023E-03	2,465062E-03
mSi0188	Eugene - OR	43127	1,996680E-03	3,470812E-03	1,991892E-03
mSi460A8	Eugene - OR	43115	2,326791E-03	3,903555E-03	2,177211E-03
xSi12922	Eugene - OR	43185	2,145388E-03	4,020838E-03	2,082523E-03
CIGS1-001	Golden - CO	12011	2,411880E-03	3,174868E-03	2,411237E-03
CIGS39013	Golden - CO	11437	4,848525E-03	6,924763E-03	4,743306E-03
CdTe75669	Golden - CO	11953	1,013974E-03	1,689448E-03	1,006375E-03
HIT05662	Golden - CO	11876	4,116152E-03	5,313480E-03	3,435626E-03
aSiMicro03038	Golden - CO	12148	8,000006E-04	9,619525E-04	7,744550E-04
aSiTandem90-31	Golden - CO	12070	2,748180E-03	3,247387E-03	2,706461E-03
aSiTriple28325	Golden - CO	11445	1,336590E-02	1,342850E-02	1,502735E-02
mSi0247	Golden - CO	11912	2,919584E-03	4,101623E-03	2,916311E-03
mSi0251	Golden - CO	11887	2,280876E-03	3,565312E-03	2,276486E-03
mSi460BB	Golden - CO	11919	8,567894E-03	9,713709E-03	8,461150E-03
xSi11246	Golden - CO	11929	2,517612E-03	3,996284E-03	2,168275E-03

6.3.2 Qualidade das soluções

Com o intuito de complementar a análise quantitativa apresentada, a Figura 22 ilustra o comportamento dos parâmetros do modelo de um diodo, indicado tanto pelas estimativas fornecidas pelo procedimento de inicialização adotado quanto pelas soluções retornadas por cada método determinístico local considerado, para o caso do módulo fotovoltaico xSi12922, alocado em Cocoa - FL. Note que esse comportamento é representado em função da irradiância solar incidente, com a temperatura de operação do módulo sendo indicada por meio da escala de cores. Além disso, ressalta-se que as escalas adotadas no eixo das ordenadas para um mesmo parâmetro não são necessariamente idênticas entre os diferentes gráficos, tendo estas sido ajustadas de forma a favorecer a visualização coerente dos respectivos resultados apresentados.

Ao analisar a Figura 22, verifica-se que as estimativas fornecidas pelo procedimento de inicialização, embora apresentem qualidade numérica razoável em termos de RMSE, capturam apenas de forma limitada a natureza física subjacente ao modelo de um diodo. Em particular, observa-se que para curvas $I-V$ correspondentes a condições de irradiância mais elevadas, tais

Figura 22 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo xSi12922, em Cocoa - FL.

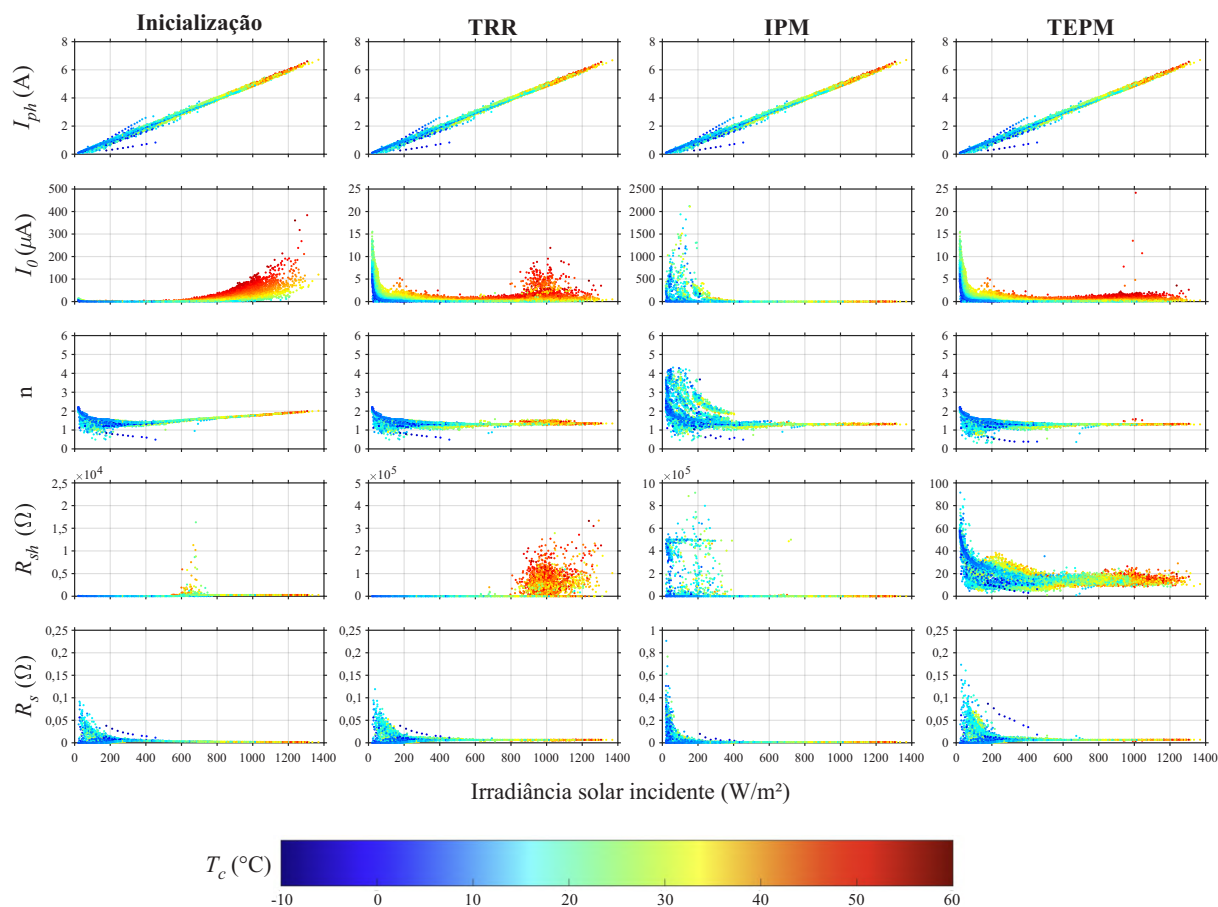


estimativas parecem atribuir ao parâmetro I_0 valores demasiadamente elevados. Adicionalmente, nota-se que o parâmetro n apresenta evolução pouco compatível com aquela evidenciada nas soluções retornadas pelos métodos determinísticos locais. Não bastasse isso, percebe-se um agrupamento de valores despropositados para R_{sh} em faixas intermediárias de irradiância.

Por outro lado, ao se examinar as soluções retornadas pelos métodos determinísticos locais, verifica-se que estas tendem, em maior ou menor grau, a capturar as dinâmicas comportamentais dos parâmetros do modelo de um diodo. Ainda assim, irregularidades significativas podem ser identificadas. No caso do TRR, observam-se “explosões” recorrentes nos valores do parâmetro R_{sh} para curvas $I-V$ correspondentes a condições de irradiância mais elevadas. Já no caso do IPM, notam-se valores demasiadamente elevados para os parâmetros I_0 , n e R_{sh} , sobretudo em níveis de irradiância mais baixos. Em contrapartida, no caso do TEPM constata-se que os cinco parâmetros exibem comportamento relativamente homogêneo, com transições suaves e fisicamente mais plausíveis ao longo de toda a faixa de operação.

Cabe ponderar que as constatações realizadas a partir da Figura 22 não constituem situações isoladas, mas sim manifestações de um quadro que se repete, em alguma medida, nos casos de outros módulos fotovoltaicos, independentemente de suas tecnologias e contextos

Figura 23 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo mSi460A8, em Eugene - OR.

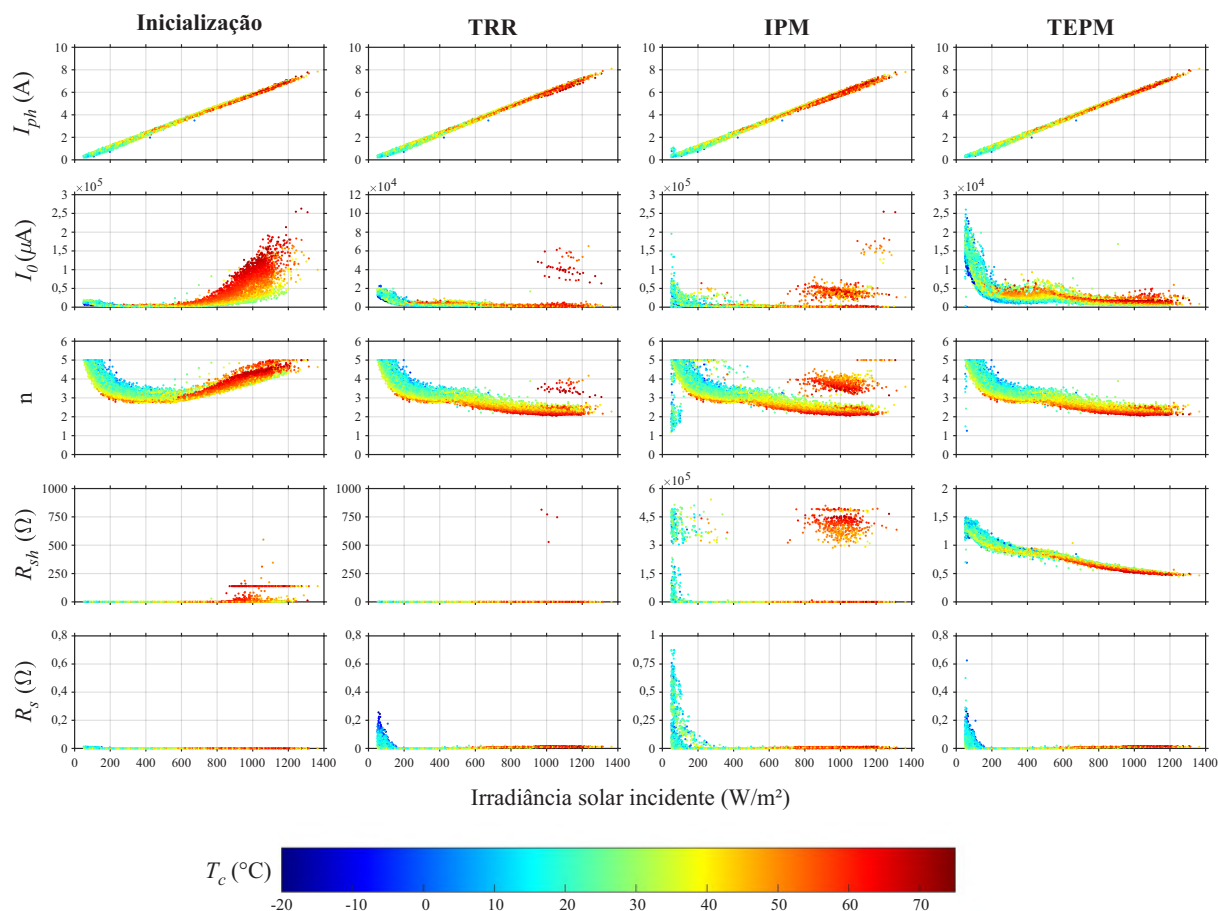


operacionais. A Figura 23, por exemplo, apresenta resultados análogos para o módulo mSi460A8, alocado em Eugene - OR, evidenciando novamente limitações nas estimativas fornecidas pelo procedimento de inicialização, bem como a ocorrência de distorções nas soluções retornadas por parte do TRR e do IPM.

Indo além, existem casos com manifestações de inconformidade ainda mais preocupantes. Por exemplo, a Figura 24 ilustra o comportamento dos parâmetros do modelo de um diodo verificado para o módulo CIGS39013, alocado em Golden - CO. Ao analisar as estimativas fornecidas pelo procedimento de inicialização, observa-se que, no caso dos parâmetros I_0 e n , estas apresentam tendência oposta àquela apontada, de maneira consonante, pelas soluções dos métodos determinísticos locais. Esse comportamento ratifica que, embora as estimativas fornecidas pelo procedimento de inicialização sejam factíveis e numericamente razoáveis, estas não devem ser empregadas diretamente em análises práticas, devendo ser compreendidas apenas como pontos de partida para um refinamento posterior e necessário.

Ainda sobre a Figura 24, verifica-se que inconformidades permanecem presentes em meio às soluções dos métodos determinísticos locais. No caso do IPM, observa-se convergência

Figura 24 – Comportamento dos parâmetros do modelo de um diodo em função da condição de operação verificado para o módulo CIGS39013, em Golden - CO.



sistemática para soluções associadas a valores excessivamente elevados de I_0 , n e R_{sh} , sobretudo em níveis de irradiância elevados. Uma situação semelhante, embora menos pronunciada, é notada no caso do TRR. Em contraste, as soluções obtidas a partir do TEPM preservam comportamento consistente ao longo de toda a faixa de irradiância. Em particular, nota-se ausência de descontinuidades abruptas e agrupamentos artificiais em quaisquer faixas de operação.

Cumprе salientar que essas distorções frequentemente verificadas nas soluções retornadas pelo TRR e pelo IPM são particularmente críticas, pois aplicações desses métodos restritas às faixas de irradiância mais sensíveis poderiam conduzir a interpretações equivocadas acerca do estado físico do módulo fotovoltaico em questão. Em estudos de degradação, por exemplo, valores artificialmente elevados de I_0 e R_{sh} podem sugerir deteriorações inexistentes, distorcendo conclusões técnicas relevantes. Ciente disso, a maior estabilidade estrutural constatada para o TEPM reforça sua robustez não apenas sob o ponto de vista numérico, mas também sob a perspectiva interpretativa, aspecto fundamental em aplicações voltadas ao monitoramento e à análise de desempenho de geradores fotovoltaicos.

6.3.3 Quesitos de eficiência

A Tabela 25 apresenta os tempos médios de simulação verificados para os problemas de extração dos parâmetros associados às curvas $I-V$ disponíveis para cada um dos 33 módulos fotovoltaicos do repositório.

No que se refere a esse quesito, observa-se que o TEPM não apresentou desempenho tão proeminente quanto aquele verificado na análise de qualidade das soluções. Em comparação com o IPM e com o TRR, o TEPM foi o método mais veloz em apenas 2 dos 33 casos analisados. Nos demais, o TRR evidenciou eficiência mais pronunciada, destacando-se como o método mais rápido na maior parte dos casos. O IPM, por sua vez, apresentou novamente desempenho relativamente inferior, situando-se recorrentemente entre os maiores tempos médios de simulação.

Não obstante, esses resultados também revelam aspectos relevantes do TEPM. Em primeiro lugar, mesmo tendo sido implementado com uma abordagem de cálculo de derivadas notadamente custosa do ponto de vista computacional, o TEPM mostrou-se capaz de retornar soluções em intervalos de tempo reduzidos, inclusive superando o IPM em diversos casos. Assim, uma vez identificada e incorporada uma alternativa igualmente eficaz, porém mais eficiente, para o cálculo das derivadas requeridas, há margem concreta para ganhos adicionais de eficiência.

Ainda assim, ressalta-se que os tempos médios observados para o TEPM situaram-se na ordem de décimos e centésimos de segundo, permanecendo alinhados às ordens de grandeza verificadas para os demais métodos e compatíveis com a condução de análises em larga escala. Sob essa perspectiva, constata-se que o TEPM, tal como implementado no presente estudo, já apresenta um nível de eficiência adequado a um contexto de aplicação de caráter intensivo.

Cabe ainda destacar que, embora as questões subjacentes ao cálculo da função W de Lambert tenham permanecido em segundo plano na análise aqui conduzida, os tempos médios reportados para todos os métodos considerados, em ordens tão reduzidas, somente foram viabilizados em razão do emprego do método preditor–corretor autoiniciado no âmbito das avaliações da função objetivo. Conforme evidenciado no Capítulo 4, a avaliação por meio da função `lambertw` do MATLAB pode demandar até 3 ordens de grandeza a mais no tempo de execução, em comparação com o método proposto. Para ilustrar a implicação dessa diferença, considere o tempo médio agregado observado para o TEPM ao longo do presente estudo, de 0,123672 segundos por problema. Nesse cenário, a resolução de 1.025.599 problemas de extração dos parâmetros (associados a curvas $I-V$ descritas por, no mínimo, 150 pontos) corresponde a aproximadamente 1,47 dias de processamento serial agregado. Sob a hipótese conservadora de um custo até 10^3 vezes superior ao se empregar `lambertw`, o mesmo experimento demandaria cerca de 1.468 dias, isto é, aproximadamente 4,02 anos. Conclui-se, portanto, que a eficiência do cálculo da função W de Lambert constitui elemento determinante para a viabilidade prática de estudos em larga escala como o considerado neste capítulo, corroborando também a pertinência do desenvolvimento do método preditor–corretor autoiniciado.

Tabela 25 – Comparação do tempo de simulação médio, dado em segundos, obtido pelos métodos avaliados na resolução dos problemas de extração dos parâmetros associados às curvas $I-V$ dos 33 módulos fotovoltaicos do repositório.

Módulo fotovoltaico	Local	Quantidade de curvas	TRR	IPM	TEPM
CIGS39017	Cocoa - FL	34775	0,100819	0,186719	0,131973
CIGS8-001	Cocoa - FL	38939	0,038497	0,105427	0,064855
CdTe75638	Cocoa - FL	39080	0,061482	0,201465	0,200575
HIT05667	Cocoa - FL	38377	0,074454	0,374565	0,276538
aSiMicro03036	Cocoa - FL	39037	0,008378	0,146803	0,088065
aSiTandem72-46	Cocoa - FL	39186	0,006458	0,168744	0,068858
aSiTriple28324	Cocoa - FL	38485	0,004163	0,063043	0,043472
mSi0166	Cocoa - FL	36765	0,029533	0,109745	0,050251
mSi0188	Cocoa - FL	39102	0,034266	0,112382	0,062665
mSi460A8	Cocoa - FL	38929	0,052897	0,138841	0,189577
xSi12922	Cocoa - FL	38989	0,059359	0,206436	0,215889
CIGS39017	Eugene - OR	42674	0,149768	0,193384	0,106076
CIGS8-001	Eugene - OR	43146	0,049384	0,123049	0,057968
CdTe75638	Eugene - OR	42248	0,045302	0,196491	0,140401
HIT05667	Eugene - OR	43271	0,057062	0,421277	0,262567
aSiMicro03036	Eugene - OR	43343	0,005335	0,110743	0,098132
aSiTandem72-46	Eugene - OR	43266	0,004922	0,107279	0,081378
aSiTriple28324	Eugene - OR	42705	0,004847	0,056051	0,068251
mSi0166	Eugene - OR	43268	0,022908	0,113954	0,033110
mSi0188	Eugene - OR	43127	0,024754	0,112034	0,041959
mSi460A8	Eugene - OR	43115	0,037035	0,125656	0,128650
xSi12922	Eugene - OR	43185	0,043943	0,220193	0,144116
CIGS1-001	Golden - CO	12011	0,040611	0,084445	0,085004
CIGS39013	Golden - CO	11437	0,138988	0,141564	0,114285
CdTe75669	Golden - CO	11953	0,056751	0,171438	0,186122
HIT05662	Golden - CO	11876	0,072342	0,345905	0,259211
aSiMicro03038	Golden - CO	12148	0,006438	0,107999	0,087586
aSiTandem90-31	Golden - CO	12070	0,005648	0,094981	0,099646
aSiTriple28325	Golden - CO	11445	0,004294	0,044030	0,100830
mSi0247	Golden - CO	11912	0,031443	0,099763	0,064010
mSi0251	Golden - CO	11887	0,031818	0,096920	0,074111
mSi460BB	Golden - CO	11919	0,080611	0,216161	0,274184
xSi11246	Golden - CO	11929	0,046983	0,171502	0,180861

6.4 CONCLUSÕES PARCIAIS

Este capítulo apresentou uma aplicação em larga escala do arcabouço metodológico desenvolvido ao longo desta tese, com ênfase na avaliação de seu desempenho em um cenário numericamente exigente e representativo de condições práticas. Para tanto, empregou-se um repositório público de medições em campo, composto por mais de um milhão de curvas $I-V$ associadas a diferentes tecnologias fotovoltaicas e a condições amplamente variadas de irradiância solar incidente e temperatura de operação. Para cada curva, formulou-se um problema particular de extração dos parâmetros, definindo-se a função objetivo sob a abordagem formal do modelo de um diodo via função W de Lambert e adotando-se, de forma sistemática, procedimentos dedicados para a definição das restrições de canalização e para a obtenção de estimativas iniciais razoáveis. Os problemas de otimização assim concebidos foram então resolvidos pelo TEPM, bem como por alternativas determinísticas locais avançadas, possibilitando uma comparação direta sob critérios uniformes de avaliação e com registro estruturado das saídas e métricas essenciais.

Os resultados obtidos evidenciaram, em termos médios, elevada qualidade das soluções retornadas pelo TEPM ao longo do conjunto massivo de instâncias, com valores de $\overline{\text{RMSE}}$ compatíveis com aqueles verificados nos estudos de caso representativos, e com redução sistemática do erro em relação às estimativas iniciais para todos os módulos avaliados. Adicionalmente, a análise qualitativa do comportamento dos parâmetros identificados em função das condições operacionais indicou maior regularidade e coerência física das soluções retornadas pelo TEPM quando comparadas às alternativas consideradas. Ainda que o TRR tenha apresentado, em geral, maior eficiência computacional, constatou-se que o TEPM mantém tempos médios de simulação em patamares compatíveis com análises intensivas, preservando níveis elevados de robustez e acurácia, especialmente quando comparado ao IPM. Por fim, a análise agregada reforçou que a eficiência do cálculo da função W de Lambert é determinante para a viabilidade prática de experimentos dessa natureza, corroborando a pertinência do método preditor–corretor autoiniciado como componente essencial do arcabouço proposto.

Em síntese, os resultados aqui apresentados corroboram os aspectos já verificados experimentalmente nos capítulos anteriores, notadamente no que concerne à robustez numérica e à elevada acurácia das soluções. Em conjunto, tais evidências confirmam que o TEPM, em comunhão com o método preditor–corretor autoiniciado e com os procedimentos auxiliares adotados, mostra-se adequado para aplicações em larga escala envolvendo a resolução massiva do problema de extração dos parâmetros do modelo de um diodo a partir de curvas $I-V$. Registra-se, ademais, que parcela substantiva dos resultados consolidados neste capítulo integra o manuscrito em preparação dedicado à apresentação do TEPM, intitulado “*Parameter extraction of the photovoltaic single diode model using a tailored exterior-point method*”, conforme antecipado no Capítulo 1. Dessa forma, as evidências aqui reportadas não apenas consolidam as contribuições originais desta tese, mas também subsidiam a sua disseminação científica no âmbito da literatura especializada.

7 CONSIDERAÇÕES FINAIS

7.1 CONCLUSÕES

A extração dos parâmetros do modelo de um diodo a partir de curvas $I-V$ constitui um problema central na modelagem de geradores fotovoltaicos, com implicações diretas para a simulação de desempenho, a caracterização de dispositivos e a interpretação física de seus mecanismos internos. Apesar de sua aparente simplicidade, trata-se de um problema matematicamente desafiador, marcado por não linearidade, transcendência e não convexidade, além de sensibilidade numérica e forte dependência de escolhas de formulação. Nesse contexto, esta tese investigou criticamente os tratamentos existentes e propôs um arcabouço metodológico integrado, fundamentado na abordagem formal da equação característica do modelo via função W de Lambert e em uma formulação restrita fisicamente consistente, visando viabilizar a obtenção reprodutível de soluções de alta qualidade com desempenho computacional compatível com aplicações práticas, inclusive em cenários intensivos.

No Capítulo 2, foi conduzido um levantamento bibliográfico e uma revisão crítica do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ de um gerador fotovoltaico. Nesse sentido, a equação característica foi desenvolvida e discutida acerca de sua estrutura e de seus cinco parâmetros. De maneira igualmente fundamental, sistematizou-se a formulação típica do problema de extração, explicitando-se a construção da função objetivo a partir dos pontos da curva $I-V$ e o papel das restrições de canalização na obtenção de soluções fisicamente admissíveis. Uma vez estabelecida essa base, foram distinguidas as principais abordagens empregadas para contornar a natureza transcendente da equação característica no escopo da função objetivo, identificando-se duas vertentes: uma abordagem simplificada e uma abordagem formal. À luz dessa distinção, cerca de cinquenta métodos de otimização propostos na literatura para a sua resolução foram documentados, considerando-se não apenas as suas classes, mas também a abordagem de cálculo adotada em seus respectivos testes de validação. As melhores soluções reportadas por esses métodos para um estudo de caso amplamente difundido foram comparadas, permitindo caracterizar o estado da arte em termos de acurácia, eficiência e maturidade tecnológica, bem como evidenciar que os problemas de otimização induzidos por cada abordagem de cálculo do modelo de um diodo são, de fato, distintos.

A análise compreensiva do Capítulo 2 evidenciou que a coexistência dessas abordagens tem, em diversos casos, sido tratada de forma imprecisa na literatura, conduzindo a comparações metodologicamente inconsistentes. Nesse contexto, verificou-se que a abordagem formal é conceitualmente mais apropriada e tende a produzir soluções de melhor qualidade, o que impõe a escolha do tratamento matemático para efetivá-la. Em particular, a reescrita explícita via função W de Lambert mostrou-se alternativa atraente e em ascensão, por viabilizar investigações tanto analíticas quanto numéricas com maior liberdade do que tratamentos implícitos, como o método de Newton. Não obstante, constatou-se que sua adoção ainda esbarra em limitações de eficiência

associadas ao cálculo da própria função W de Lambert, configurando tópico de pesquisa em aberto. Adicionalmente, observou-se que métodos determinísticos locais tendem a ser mais eficientes do que alternativas metaheurísticas e conjugadas. Notou-se, contudo, que os trabalhos que exploram tais métodos concentram-se em formulações irrestritas e direcionam esforços sobretudo a procedimentos de inicialização, em detrimento de métodos dedicados à resolução direta do problema restrito. Essa desconsideração das restrições é especialmente preocupante diante da fundamentação física dos parâmetros e da natureza não convexa do problema, pois pode conduzir a soluções sem significância conceitual e comprometer interpretações de interesse.

Diante dessas considerações, o Capítulo 3 formalizou o problema de extração sob a abordagem explícita da equação característica do modelo de um diodo via função W de Lambert. A análise das propriedades dessa função mostrou que sua aplicabilidade exige cuidadosa consideração do domínio, uma vez que determinadas regiões podem conduzir a valores complexos, incompatíveis com a natureza real do modelo. Essa constatação reforçou a necessidade de uma formulação restrita que assegure avaliações em regiões do domínio que preservem a admissibilidade física. Estabeleceu-se, assim, uma formulação restrita do problema de extração dos parâmetros como plataforma de pesquisa. No âmbito dessa formulação, a análise dos principais aspectos matemáticos evidenciou a natureza não convexa do problema e a investigação numérica indicou que este é inerentemente mal escalonado e pode apresentar mal condicionamento, características que influenciam de maneira decisiva a robustez e a eficiência de métodos determinísticos locais.

A partir desse diagnóstico, consolidaram-se duas linhas complementares de investigação. A primeira, desenvolvida no Capítulo 4, dedicou-se ao estudo de métodos de cálculo da função W de Lambert no domínio fisicamente relevante ao modelo de um diodo. A segunda, apresentada no Capítulo 5, direcionou-se ao desenvolvimento de um método determinístico local dedicado, coerente com as propriedades estruturais e numéricas evidenciadas para o problema formalizado.

No Capítulo 4, uma prospecção metodológica revelou uma vertente recente de desenvolvimento na área de modelagem fotovoltaica voltada à proposição de métodos especializados para o cálculo da função W de Lambert nesse escopo. A revisão dos métodos disponíveis explicitou potenciais, limitações e oportunidades concretas de aprimoramento. Aproveitando-se essa janela, concebeu-se o *método preditor–corretor autoiniciado* (“*self-started predictor–corrector method*”), caracterizado por um procedimento não iterativo e essencialmente não condicional, dado por uma única passagem por quatro equações explícitas de complexidade moderada. Os resultados demonstraram desempenho particularmente atraente frente ao estado da arte, com acurácia equivalente à da função `lambertw` do MATLAB e custo computacional substancialmente inferior, além de alta competitividade em relação a métodos especializados concebidos para o mesmo propósito. Em decorrência, estabeleceu-se suporte técnico direto para a adoção sistemática da abordagem formal via função W de Lambert no problema de extração dos parâmetros.

No Capítulo 5, foi desenvolvida a principal contribuição metodológica desta tese: a concepção de um método determinístico local dedicado à resolução do problema de extração dos parâmetros sob a abordagem formal via função W de Lambert e a formulação restrita previamente estabelecida. Com base na revisão do Capítulo 2, evidenciou-se que métodos de escopo amplo para problemas não lineares restritos ainda não haviam sido aplicados sistematicamente a esse problema e que, dentre essas alternativas, os métodos de barreira (BM) são particularmente alinhados à sua natureza. A investigação conduzida nesse grupo culminou na adoção do paradigma de método de ponto exterior (EPM) como base para o desenvolvimento do método proposto. Sua estrutura lógica foi estabelecida em consonância com a literatura específica e munida de mecanismos voltados a lidar com as dificuldades mapeadas no Capítulo 3. Em síntese, a barreira logarítmica modificada trata implicações diretas das restrições; a regularização inercial mitiga efeitos da não convexidade; e o escalonamento simétrico reforça a robustez numérica frente a discrepâncias de escala e ao mal condicionamento. O método resultante, concebido sob medida, foi denominado *tailored exterior-point method* (TEPM). Os resultados obtidos em estudos de caso clássicos evidenciaram que o TEPM é capaz de retornar, de forma robusta, as melhores soluções conhecidas na literatura, apresentando desempenho competitivo frente às alternativas determinísticas locais de estado da arte. Ademais, o desempenho relativamente superior ao do IPM confirmou a pertinência do desenvolvimento de um método de barreira concebido sob medida para esse problema.

No Capítulo 6, o arcabouço metodológico foi submetido a uma aplicação em larga escala, envolvendo mais de um milhão de curvas $I-V$, em um cenário numericamente exigente e representativo de condições práticas. Neste âmbito, foram explicitados procedimentos para formulação das restrições de canalização e inicialização, os quais, embora não tenham sido objeto de investigação aprofundada, foram indicados de maneira consistente. Os resultados confirmaram, em ambiente massivo, as constatações previamente obtidas: o TEPM manteve desempenho mais atrativo frente ao IPM e apresentou o nível de acurácia mais elevado dentre todos os métodos determinísticos locais de estado da arte considerados. Ainda que haja espaço para aprimoramentos em termos de eficiência, os tempos computacionais permaneceram compatíveis, em ordem de grandeza, com os demais métodos analisados. Além disso, cumpre destacar que os tempos reduzidos observados somente foram viabilizados pela utilização do método preditor–corretor autoiniciado para os requeridos cálculos da função W de Lambert; sem tal recurso, o estudo massivo dificilmente poderia ter sido conduzido em tempo hábil ou com a mesma qualidade de soluções.

À luz dessas análises, conclui-se que os objetivos específicos definidos no Capítulo 1 foram plenamente atendidos. A partir da sólida revisão bibliográfica realizada (Capítulo 2), o problema foi formalizado sob a abordagem formal do modelo de um diodo via função W de Lambert e devidamente analisado sob aspectos matemáticos e numéricos, evidenciando as suas dificuldades de resolução (Capítulo 3); a função W de Lambert foi estudada, e um método foi proposto para permitir o seu cálculo eficaz e eficiente no domínio fisicamente admissível ao

modelo de um diodo (Capítulo 4); um método de barreira especializado foi desenvolvido para resolver, de forma acurada e robusta, o problema de extração dos parâmetros considerando as suas restrições de canalização (Capítulo 5); e procedimentos para a formulação das restrições de canalização e para a inicialização foram adequadamente propostos, tendo o arcabouço metodológico sido criteriosamente validado tanto em estudos de caso clássicos quanto em uma aplicação massiva, representativa de um cenário realista (Capítulos 4, 5 e 6).

Por conseguinte, o objetivo geral foi efetivamente alcançado, consolidando-se a proposição de uma nova alternativa de método determinístico local para a resolução do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$, à luz da formulação restrita do problema e da abordagem formal baseada na função W de Lambert aplicada à equação característica do modelo.

Dessa forma, encerra-se o arco lógico estabelecido desde a identificação das lacunas metodológicas até a validação prática do arcabouço proposto. Ao articular uma formulação matemática rigorosa e métodos numéricos sob medida, esta tese oferece como contribuições originais:

- **Delimitação conceitual e operacional das vertentes de cálculo do modelo de um diodo.** Estabeleceu-se uma distinção objetiva entre as abordagens simplificada e formal, discutindo-se as suas implicações diretas sobre a natureza do problema de otimização induzido. Essa sistematização fornece base conceitual mais clara para análises comparativas e mitiga inconsistências metodológicas observadas na literatura.
- **Método dedicado ao cálculo da função W de Lambert no domínio fisicamente relevante.** Concebeu-se o método preditor–corretor autoiniciado para a avaliação da função W de Lambert no escopo do modelo de um diodo, combinando acurácia na ordem da precisão de máquina com custo computacional reduzido. Esse resultado viabiliza o emprego sistemático da abordagem formal em cenários intensivos.
- **Método determinístico local de barreira dedicado à resolução do problema.** Desenvolveu-se o TEPM para a resolução acurada e robusta do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$, incorporando mecanismos voltados a lidar diretamente com restrições, não convexidade e sensibilidade numérica. O método é capaz de retornar, de forma reproduzível, soluções de elevada qualidade e de sustentar aplicações massivas.

Em termos de impacto, essas contribuições oferecem valor direto tanto para a comunidade científica quanto para a indústria fotovoltaica, reafirmando a centralidade do problema de extração dos parâmetros do modelo de um diodo a partir de uma curva $I-V$ do gerador fotovoltaico em estudo e fornecendo instrumentos concretos para a sua resolução de maneira efetiva e tecnicamente fundamentada.

7.2 PERSPECTIVAS DE CONTINUIDADE DA PESQUISA

Diante da relevância da temática investigada e das constatações registradas ao longo deste documento, vislumbram-se oportunidades concretas de aprofundamento e expansão, tanto no âmbito do próprio arcabouço proposto quanto em tópicos correlatos. Destacam-se, nesse contexto, as seguintes frentes de continuidade:

- **Disponibilização em repositório público** – estruturar e publicar o arcabouço metodológico desenvolvido em repositório aberto, tanto na linguagem original (MATLAB) quanto em linguagens de ampla adoção recente, como PYTHON e JULIA, de modo a facilitar sua apropriação pela comunidade científica e pela indústria fotovoltaica.
- **Cálculo de derivadas no TEPM** – investigar estratégias de cálculo de derivadas que preservem a robustez do esquema híbrido atualmente empregado (gradiente via passo complexo e Hessiana via diferenças finitas centradas), reduzindo custo computacional e ampliando a eficiência global do método.
- **Aprimoramentos estruturais motivados pela não convexidade** – aprofundar a análise dos efeitos da não convexidade do problema de extração dos parâmetros formulado sob a abordagem formal do modelo de um diodo via função W de Lambert, visando subsidiar aprimoramentos na regularização inercial e, eventualmente, em outros elementos estruturais do TEPM.
- **Uso de base massiva de resultados para refinamento de componentes do arcabouço** – analisar os parâmetros do modelo de um diodo extraídos das mais de um milhão de curvas $I-V$ do repositório do NREL, caracterizando dependências com irradiância e temperatura, e explorando esse entendimento para aprimorar metodologias de limites admissíveis, procedimentos de inicialização e equações de translação.
- **Metodologias para determinação de limites admissíveis** – propor estratégias sistemáticas para determinar limites inferiores e superiores fisicamente admissíveis para os parâmetros do modelo de um diodo, com vistas a subsidiar formulações restritas do problema de extração dos parâmetros mais robustas e aplicações mais eficientes de diferentes métodos de otimização.
- **Procedimentos de inicialização e combinações metodológicas** – conduzir estudos comparativos abrangentes sobre procedimentos de inicialização já propostos para métodos determinísticos locais, avaliando sistematicamente seus impactos na robustez e na qualidade das soluções obtidas para os parâmetros do modelo de um diodo; adicionalmente, investigar estratégias alternativas fundamentadas em métodos destinados ao uso estrito de pontos notáveis da curva $I-V$, bem como a combinação estruturada de métodos correlatos (incluindo metaheurísticas reconhecidamente eficazes), com vistas à concepção de abordagens híbridas.

- **Reavaliação de métodos determinísticos globais como referência** – reavaliar a aplicação de métodos determinísticos globais no âmbito do problema de extração dos parâmetros, com ênfase em estudos de caso amplamente difundidos, de modo a estabelecer referências sólidas para a validação comparativa de métodos de outras classes e tipos.
- **Reexame crítico de comparações na literatura** – diante da constatação de que a coexistência de abordagens de cálculo do modelo de um diodo tem sido tratada de forma imprecisa em parte da literatura, resultando em comparações metodologicamente inconsistentes, conduzir um reexame crítico de resultados reportados sob essas premissas.
- **Impactos técnico-econômicos de estimativas paramétricas não acuradas** – investigar como erros na estimação dos parâmetros do modelo de um diodo podem repercutir sobre atividades de monitoramento, diagnóstico e manutenção de geradores fotovoltaicos, induzindo interpretações equivocadas sobre degradação, falhas e desempenho, bem como implicações econômicas associadas a intervenções desnecessárias, à postergação de ações corretivas e a perdas de receita por subdesempenho não identificado.
- **Extensão para modelos físicos correlatos** – avaliar a aplicabilidade do arcabouço desenvolvido, originalmente orientado ao modelo de um diodo, no escopo de modelos fisicamente mais elaborados, com destaque para o modelo de dois diodos.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] IRENA. World energy transitions outlook: 1.5° c pathway. *International Renewable Energy Agency, Abu Dhabi*, 2024.
- [2] IEA et al. The progress energy report: Tracking sdg 7. *World Bank, Washington, DC*, 2024.
- [3] MASSON, G.; L'EPINE, M. de; KAIZUKA, I. Trends in photovoltaic applications 2023. *Photovoltaic Power Systems Programme, International Energy Agency*, 2014.
- [4] REN24. Renewables 2024. global status report. *REN21 Secretariat*, 2024.
- [5] IRENA. Renewable power generation costs in 2023. *International Renewable Energy Agency, Abu Dhabi*, 2024.
- [6] MERTENS, K. *Photovoltaics: fundamentals, technology, and practice*. [S.l.]: John Wiley & Sons, 2019.
- [7] HÄBERLIN, H. *Photovoltaics: system design and practice*. [S.l.]: John Wiley & Sons, 2012.
- [8] TEODORESCU, R.; LISERRE, M.; RODRIGUEZ, P. *Grid converters for photovoltaic and wind power systems*. [S.l.]: John Wiley & Sons, 2011.
- [9] YANG, Y. et al. *Advances in grid-connected photovoltaic power conversion systems*. [S.l.]: Woodhead Publishing, 2018.
- [10] PEARSALL, N. *The performance of photovoltaic (PV) systems: modelling, measurement and assessment*. [S.l.]: Woodhead Publishing, 2016.
- [11] DEOTTI, L. et al. Technical and economic analysis of battery storage for residential solar photovoltaic systems in the brazilian regulatory context. *Energies, Multidisciplinary Digital Publishing Institute*, v. 13, n. 24, p. 6517, 2020.
- [12] IRENA; ILO. Renewable energy and jobs: Annual review 2024. *International Renewable Energy Agency, Abu Dhabi, and International Labour Organization, Geneva*, 2024.
- [13] GARLET, T. B. et al. Paths and barriers to the diffusion of distributed generation of photovoltaic energy in southern brazil. *Renewable and Sustainable Energy Reviews, Elsevier*, v. 111, p. 157–169, 2019.
- [14] PREHODA, E.; PEARCE, J. M.; SCHELLY, C. Policies to overcome barriers for renewable energy distributed generation: A case study of utility structure and regulatory regimes in michigan. *Energies, MDPI*, v. 12, n. 4, p. 674, 2019.
- [15] SEGUIN, R. et al. *High-penetration PV integration handbook for distribution engineers*. [S.l.], 2016.
- [16] OLOWU, T. O. et al. Future challenges and mitigation methods for high photovoltaic penetration: A survey. *Energies, MDPI*, v. 11, n. 7, p. 1782, 2018.
- [17] PARRA, I. De la et al. Pv performance modelling: A review in the light of quality assurance for large pv plants. *Renewable and Sustainable Energy Reviews, Elsevier*, v. 78, p. 780–797, 2017.

- [18] STEIN, J. S. Energy prediction and system modeling. *Photovoltaic Solar Energy*, Wiley Online Library, p. 564, 2017.
- [19] DEOTTI, L. M. P.; GOMES, F. V. Análise de viabilidade técnica e econômica de um sistema de microgeração solar fotovoltaica residencial com autoconsumo remoto. In: *VII Congresso Brasileiro de Energia Solar*. [S.l.: s.n.], 2018.
- [20] MERMOUD, A.; WITTMER, B. Pvsyst user's manual. *Switzerland*. Disponível em: <<http://files.pvsyst.com/help/>> [Acessado em 10 de junho de 2019], 2019.
- [21] VALENTIN, G. Pv*sol expert help documentation. *Valentin Software GmbH*. Disponível em: <<https://help.valentin-software.com/pvsol/calculation/>> [Acessado em 10 de junho de 2019], 2019.
- [22] GILMAN, P. et al. Sam photovoltaic model technical reference update. *NREL: Golden, CO, USA*, 2018.
- [23] PETRONE, G. et al. *Photovoltaic sources modeling*. [S.l.]: Wiley Online Library, 2017.
- [24] DEOTTI, L. et al. Empirical models applied to distributed energy resources—an analysis in the light of regulatory aspects. *Energies*, Multidisciplinary Digital Publishing Institute, v. 14, n. 2, p. 326, 2021.
- [25] LUO, W. et al. Potential-induced degradation in photovoltaic modules: a critical review. *Energy & environmental science*, Royal Society of Chemistry, v. 10, n. 1, p. 43–68, 2017.
- [26] HRELJA, N. et al. Single diode model applied to pv module aging. In: *35th European Photovoltaic Solar Energy Conference and Exhibition*. [S.l.: s.n.], 2018. p. 1290–1293.
- [27] KALLIOJÄRVI, H.; LAPPALAINEN, K. Identification of photovoltaic module ageing from symmetrically and asymmetrically cut current–voltage curves. *Solar Energy Materials and Solar Cells*, Elsevier, v. 296, p. 114081, 2026.
- [28] PILIOUGINE, M. et al. Parameters extraction of single diode model for degraded photovoltaic modules. *Renewable Energy*, Elsevier, v. 164, p. 674–686, 2021.
- [29] KALLIOJÄRVI, H.; LAPPALAINEN, K. Methods for online identification of photovoltaic module ageing by series resistance from measured current–voltage curves. *Energy Reports*, Elsevier, v. 13, p. 1558–1570, 2025.
- [30] LAZZARETTI, A. E. et al. A monitoring system for online fault detection and classification in photovoltaic plants. *Sensors*, MDPI, v. 20, n. 17, p. 4688, 2020.
- [31] LIU, Y.; WU, Y. Fault diagnosis of photovoltaic modules: A review. *Solar Energy*, Elsevier, v. 293, p. 113489, 2025.
- [32] DUPRÉ, O.; VAILLON, R.; GREEN, M. A. Thermal behavior of photovoltaic devices. *Physics and engineering*, Springer, v. 10, p. 978–3, 2017.
- [33] GREEN, M. A. Solar cells: operating principles, technology, and system applications. *Englewood Cliffs*, 1982.
- [34] JORDEHI, A. R. Parameter estimation of solar photovoltaic (pv) cells: A review. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 61, p. 354–371, 2016.

- [35] LI, S.; GONG, W.; GU, Q. A comprehensive survey on meta-heuristic algorithms for parameter extraction of photovoltaic models. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 141, p. 110828, 2021.
- [36] DEVILLE, L. et al. Parameter translation for photovoltaic single-diode models. *IEEE Journal of Photovoltaics*, IEEE, 2025.
- [37] INTELMARKETRESEARCH. *Solar Photovoltaic Design Software Market Growth Analysis, Dynamics, Key Players and Innovations, Outlook and Forecast 2026-2034*. [S.l.], 2026.
- [38] DOBOS, A. P.; MACALPINE, S. M. Procedure for applying iec-61853 test data to a single diode model. In: IEEE. *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*. [S.l.], 2014. p. 2846–2849.
- [39] BATZELIS, E. Non-iterative methods for the extraction of the single-diode model parameters of photovoltaic modules: A review and comparative assessment. *Energies*, MDPI, v. 12, n. 3, p. 358, 2019.
- [40] TIAN, H. et al. A cell-to-module-to-array detailed model for photovoltaic panels. *Solar energy*, Elsevier, v. 86, n. 9, p. 2695–2706, 2012.
- [41] BATZELIS, E. I. et al. Computation of the lambert w function in photovoltaic modeling. *Lecture Notes in Electrical Engineering; JB Metzler: Berlin, Germany*, p. 583–595, 2020.
- [42] TOLEDO, F. J. et al. Quick and accurate strategy for calculating the solutions of the photovoltaic single-diode model equation. *IEEE Journal of Photovoltaics*, IEEE, v. 12, n. 2, p. 493–500, 2022.
- [43] HACHANA, O. et al. Photovoltaic mono and bifacial module/string electrical model parameters identification and validation based on a new differential evolution bee colony optimizer. *Energy Conversion and Management*, Elsevier, v. 248, p. 114667, 2021.
- [44] MATHEW, L. E.; PANCHAL, A. K. A complete numerical investigation on implicit and explicit pv single-diode-models using i-and v-approaches. *IEEE Journal of Photovoltaics*, IEEE, v. 11, n. 3, p. 827–837, 2021.
- [45] HANSEN, C. *Parameter estimation for single diode models of photovoltaic modules*. [S.l.], 2015.
- [46] ESEN, V.; SAĞLAM, Ş.; ORAL, B. Light sources of solar simulators for photovoltaic devices: A review. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 77, p. 1240–1250, 2017.
- [47] CORTÉS-SEVERINO, R. et al. Optimal design and experimental test of a solar simulator for solar photovoltaic modules. *Energy Science & Engineering*, Wiley Online Library, v. 9, n. 12, p. 2514–2528, 2021.
- [48] MORALES-ARAGONÉS, J. I. et al. A review of i–v tracers for photovoltaic modules: topologies and challenges. *Electronics*, MDPI, v. 10, n. 11, p. 1283, 2021.
- [49] MARION, B. et al. *Data for validating models for pv module performance*. [S.l.], 2021.
- [50] BOYD, M. Performance data from the nist photovoltaic arrays and weather station. *Journal of Research of the National Institute of Standards and Technology*, v. 122, p. 1, 2017.

- [51] BRAID, J. *Photovoltaic Module Current-Voltage and Electroluminescence Image Data (PV-IV-EL)*. [S.l.], 2025.
- [52] GAO, S. et al. Rethinking solar photovoltaic parameter estimation: global optimality analysis and a simple efficient differential evolution method. *arXiv preprint arXiv:2011.12114*, 2020.
- [53] LAUDANI, A.; FULGINEI, F. R.; SALVINI, A. High performing extraction procedure for the one-diode model of a photovoltaic panel from experimental i–v curves by using reduced forms. *Solar Energy*, Elsevier, v. 103, p. 316–326, 2014.
- [54] CÁRDENAS, A. A. et al. Experimental parameter extraction in the single-diode photovoltaic model via a reduced-space search. *IEEE Transactions on Industrial Electronics*, IEEE, v. 64, n. 2, p. 1468–1476, 2016.
- [55] AYANG, A. et al. Faults diagnosis and monitoring of a single diode photovoltaic module based on estimated parameters. In: IEEE. *2018 IEEE Electrical Power and Energy Conference (EPEC)*. [S.l.], 2018. p. 1–6.
- [56] ZHU, Y.; XIAO, W. A comprehensive review of topologies for photovoltaic i–v curve tracer. *Solar Energy*, Elsevier, v. 196, p. 346–357, 2020.
- [57] XIAO, K. et al. Study on performance degradation and damage modes of thin-film photovoltaic cell subjected to particle impact. *Scientific Reports*, Nature Publishing Group UK London, v. 11, n. 1, p. 782, 2021.
- [58] REN, B. et al. Fault diagnosis of photovoltaic systems based on global iv features and similarity learning. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 171, p. 110987, 2025.
- [59] RUSCHEL, C. S.; GASPARIN, F. P.; KRENZINGER, A. Experimental analysis of the single diode model parameters dependence on irradiance and temperature. *Solar Energy*, Elsevier, v. 217, p. 134–144, 2021.
- [60] CHEN, X. et al. Open data sets for assessing photovoltaic system reliability. *Applied Energy*, Elsevier, v. 395, p. 126132, 2025.
- [61] SOTO, W. D.; KLEIN, S. A.; BECKMAN, W. A. Improvement and validation of a model for photovoltaic array performance. *Solar energy*, Elsevier, v. 80, n. 1, p. 78–88, 2006.
- [62] IBRAHIM, H.; ANANI, N. Variations of pv module parameters with irradiance and temperature. *Energy Procedia*, Elsevier, v. 134, p. 276–285, 2017.
- [63] EASWARAKHANTHAN, T. et al. Nonlinear minimization algorithm for determining the solar cell parameters with microcomputers. *International journal of solar energy*, Taylor & Francis, v. 4, n. 1, p. 1–12, 1986.
- [64] TOLEDO, F. J.; BLANES, J. M.; GALIANO, V. Two-step linear least-squares method for photovoltaic single-diode model parameters extraction. *IEEE Transactions on Industrial Electronics*, IEEE, v. 65, n. 8, p. 6301–6308, 2018.
- [65] PANCHAL, A. K. I–v data operated high-quality photovoltaic solution through per-unit single-diode model. *IEEE Journal of Photovoltaics*, IEEE, v. 10, n. 4, p. 1175–1184, 2020.

- [66] PANCHAL, A. K. A per-unit-single-diode-model parameter extraction algorithm: A high-quality solution without reduced-dimensions search. *Solar Energy*, Elsevier, v. 207, p. 1070–1077, 2020.
- [67] XU, J. Separable nonlinear least squares search of parameter values in photovoltaic models. *IEEE Journal of Photovoltaics*, IEEE, v. 12, n. 1, p. 372–380, 2021.
- [68] CHENOUCARD, R.; EL-SEHIEMY, R. A. An interval branch and bound global optimization algorithm for parameter estimation of three photovoltaic models. *Energy Conversion and Management*, Elsevier, v. 205, p. 112400, 2020.
- [69] YE, M.; WANG, X.; XU, Y. Parameter extraction of solar cells using particle swarm optimization. *Journal of Applied Physics*, American Institute of Physics, v. 105, n. 9, p. 094502, 2009.
- [70] COSTA, W. T. da et al. Identification of photovoltaic model parameters by differential evolution. In: IEEE. *2010 IEEE International Conference on Industrial Technology*. [S.l.], 2010. p. 931–936.
- [71] EL-NAGGAR, K. M. et al. Simulated annealing algorithm for photovoltaic parameters identification. *Solar Energy*, Elsevier, v. 86, n. 1, p. 266–274, 2012.
- [72] ASKARZADEH, A.; REZAZADEH, A. Artificial bee swarm optimization algorithm for parameters identification of solar cell models. *Applied Energy*, Elsevier, v. 102, p. 943–949, 2013.
- [73] ASKARZADEH, A.; REZAZADEH, A. Extraction of maximum power point in solar cells using bird mating optimizer-based parameters identification approach. *Solar energy*, Elsevier, v. 90, p. 123–133, 2013.
- [74] MA, J. et al. Parameter estimation of photovoltaic models via cuckoo search. *Journal of applied mathematics*, Hindawi, v. 2013, 2013.
- [75] OLIVA, D.; CUEVAS, E.; PAJARES, G. Parameter identification of solar cells using artificial bee colony optimization. *Energy*, Elsevier, v. 72, p. 93–102, 2014.
- [76] PATEL, S. J.; PANCHAL, A. K.; KHERAJ, V. Extraction of solar cell parameters from a single current–voltage characteristic using teaching learning based optimization algorithm. *Applied Energy*, Elsevier, v. 119, p. 384–393, 2014.
- [77] GUO, L. et al. Parameter identification and sensitivity analysis of solar cell models with cat swarm optimization algorithm. *Energy conversion and management*, Elsevier, v. 108, p. 520–528, 2016.
- [78] DERICK, M. et al. An improved optimization technique for estimation of solar photovoltaic parameters. *Solar Energy*, Elsevier, v. 157, p. 116–124, 2017.
- [79] ELAZAB, O. S. et al. Parameters estimation of single-and multiple-diode photovoltaic model using whale optimisation algorithm. *IET Renewable Power Generation*, Wiley Online Library, v. 12, n. 15, p. 1755–1761, 2018.
- [80] XIONG, G. et al. Application of supply-demand-based optimization for parameter extraction of solar photovoltaic models. *Complexity*, Hindawi, v. 2019, 2019.

- [81] ASKARZADEH, A.; REZAZADEH, A. An innovative global harmony search algorithm for parameter identification of a pem fuel cell model. *IEEE Transactions on industrial electronics*, IEEE, v. 59, n. 9, p. 3473–3480, 2011.
- [82] GONG, W.; CAI, Z. Parameter extraction of solar cell models using repaired adaptive differential evolution. *Solar Energy*, Elsevier, v. 94, p. 209–220, 2013.
- [83] YU, K. et al. A performance-guided jaya algorithm for parameters identification of photovoltaic cell and module. *Applied energy*, Elsevier, v. 237, p. 241–257, 2019.
- [84] LI, S. et al. Parameter extraction of photovoltaic models using an improved teaching-learning-based optimization. *Energy Conversion and Management*, Elsevier, v. 186, p. 293–305, 2019.
- [85] DEOTTI, L. M. P.; PEREIRA, J. L. R.; JÚNIOR, I. C. da S. Parameter extraction of photovoltaic models using an enhanced lévy flight bat algorithm. *Energy Conversion and Management*, Elsevier, v. 221, p. 113114, 2020.
- [86] LIANG, J. et al. Classified perturbation mutation based particle swarm optimization algorithm for parameters extraction of photovoltaic models. *Energy Conversion and Management*, Elsevier, v. 203, p. 112138, 2020.
- [87] XIONG, G. et al. Modified search strategies assisted crossover whale optimization algorithm with selection operator for parameter extraction of solar photovoltaic models. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 11, n. 23, p. 2795, 2019.
- [88] ABDEL-BASSET, M. et al. Parameter estimation of photovoltaic models using an improved marine predators algorithm. *Energy Conversion and Management*, Elsevier, v. 227, p. 113491, 2021.
- [89] LONG, W. et al. Parameters identification of photovoltaic models by using an enhanced adaptive butterfly optimization algorithm. *Energy*, Elsevier, v. 229, p. 120750, 2021.
- [90] YU, S. et al. Parameter identification of photovoltaic models using a sine cosine differential gradient based optimizer. *IET Renewable Power Generation*, Wiley Online Library, 2022.
- [91] XIONG, G. et al. Parameter extraction of solar photovoltaic models by means of a hybrid differential evolution with whale optimization algorithm. *Solar Energy*, Elsevier, v. 176, p. 742–761, 2018.
- [92] CHEN, X. et al. Teaching–learning–based artificial bee colony for solar photovoltaic parameter estimation. *Applied energy*, Elsevier, v. 212, p. 1578–1588, 2018.
- [93] CHEN, X.; YU, K. Hybridizing cuckoo search algorithm with biogeography-based optimization for estimating photovoltaic model parameters. *Solar Energy*, Elsevier, v. 180, p. 192–206, 2019.
- [94] LONG, W. et al. A new hybrid algorithm based on grey wolf optimizer and cuckoo search for parameter extraction of solar photovoltaic models. *Energy Conversion and Management*, Elsevier, v. 203, p. 112243, 2020.
- [95] DIAB, A. A. Z. et al. Coyote optimization algorithm for parameters estimation of various models of solar cells and pv modules. *Ieee Access*, IEEE, v. 8, p. 111102–111140, 2020.

- [96] AGWA, A. M.; EL-FERGANY, A. A.; MAKSOUD, H. A. Electrical characterization of photovoltaic modules using farmland fertility optimizer. *Energy Conversion and Management*, Elsevier, v. 217, p. 112990, 2020.
- [97] REZK, H. et al. A robust parameter estimation approach based on stochastic fractal search optimization algorithm applied to solar pv parameters. *Energy Reports*, Elsevier, v. 7, p. 620–640, 2021.
- [98] REDDY, A. K. V. K.; NARAYANA, K. V. L. Investigation of a social group assisted differential evolution for the optimal pv parameter extraction of standard and modified diode models. *Energy Conversion and Management*, Elsevier, v. 268, p. 115955, 2022.
- [99] JORDEHI, A. R. Enhanced leader particle swarm optimisation (elpso): an efficient algorithm for parameter estimation of photovoltaic (pv) cells and modules. *Solar Energy*, Elsevier, v. 159, p. 78–87, 2018.
- [100] NUNES, H. et al. A new high performance method for determining the parameters of pv cells and modules based on guaranteed convergence particle swarm optimization. *Applied energy*, Elsevier, v. 211, p. 774–791, 2018.
- [101] KLER, D. et al. A novel approach to parameter estimation of photovoltaic systems using hybridized optimizer. *Energy Conversion and Management*, Elsevier, v. 187, p. 486–511, 2019.
- [102] EL-FERGANY, A. *Parameters identification of PV model using improved slime mould optimizer and Lambert W-function*. *Energy Reports*. 7, 875–887. 2021.
- [103] YANG, X.; GONG, W.; WANG, L. Comparative study on parameter extraction of photovoltaic models via differential evolution. *Energy Conversion and Management*, Elsevier, v. 201, p. 112113, 2019.
- [104] DKHICHI, F. et al. Parameter identification of solar cell model using levenberg–marquardt algorithm combined with simulated annealing. *Solar Energy*, Elsevier, v. 110, p. 781–788, 2014.
- [105] HAMID, N. F. A.; RAHIM, N. A.; SELVARAJ, J. Solar cell parameters identification using hybrid nelder-mead and modified particle swarm optimization. *Journal of Renewable and Sustainable Energy*, AIP Publishing LLC, v. 8, n. 1, p. 015502, 2016.
- [106] CHEN, Z. et al. Parameters identification of photovoltaic models using hybrid adaptive nelder-mead simplex algorithm based on eagle strategy. *Applied Energy*, Elsevier, v. 182, p. 47–57, 2016.
- [107] XU, S.; WANG, Y. Parameter estimation of photovoltaic modules using a hybrid flower pollination algorithm. *Energy Conversion and Management*, Elsevier, v. 144, p. 53–68, 2017.
- [108] WU, L. et al. Parameter extraction of photovoltaic models from measured iv characteristics curves using a hybrid trust-region reflective algorithm. *Applied Energy*, Elsevier, v. 232, p. 36–53, 2018.
- [109] LI, S. et al. Parameter estimation of photovoltaic models with memetic adaptive differential evolution. *Solar Energy*, Elsevier, v. 190, p. 465–474, 2019.

- [110] ZHANG, H. et al. Orthogonal nelder-mead moth flame method for parameters identification of photovoltaic modules. *Energy Conversion and Management*, Elsevier, v. 211, p. 112764, 2020.
- [111] PENG, L. et al. Information sharing search boosted whale optimizer with nelder-mead simplex for parameter estimation of photovoltaic models. *Energy Conversion and Management*, Elsevier, v. 270, p. 116246, 2022.
- [112] TONG, N. T.; PORA, W. A parameter extraction technique exploiting intrinsic properties of solar cells. *Applied energy*, Elsevier, v. 176, p. 104–115, 2016.
- [113] NOCEDAL, J.; WRIGHT, S. *Numerical optimization*. [S.l.]: Springer Science & Business Media, 2006.
- [114] YANG, X.-S. *Nature-inspired algorithms and applied optimization*. [S.l.]: Springer, 2018.
- [115] YANG, X.-S. *Optimization techniques and applications with examples*. [S.l.]: John Wiley & Sons, 2018.
- [116] MARTINS, J. R.; NING, A. *Engineering design optimization*. [S.l.]: Cambridge University Press, 2021.
- [117] BECK, A. *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. [S.l.]: SIAM, 2014.
- [118] EDGAR, T. F. et al. *Optimization of chemical processes*. [S.l.: s.n.], 2001.
- [119] SAUER, T. *Numerical analysis*. [S.l.]: Pearson, 2017.
- [120] MATLAB. *lsqnonlin: solve nonlinear least-squares problems*. Natick, Massachusetts: The MathWorks Inc., 2021.
- [121] MATLAB. *lsqcurvefit: solve nonlinear curve-fitting problems in least-squares sense*. Natick, Massachusetts: The MathWorks Inc., 2021.
- [122] CORLESS, R. M. et al. On the lambertw function. *Advances in Computational mathematics*, Springer, v. 5, n. 1, p. 329–359, 1996.
- [123] MEZO, I. *The Lambert W function: its generalizations and applications*. [S.l.]: Chapman and Hall/CRC, 2022.
- [124] POLYAK, R. A. et al. *Introduction to continuous optimization*. [S.l.]: Springer, 2021. v. 172.
- [125] GRIVA, I. Numerical experiments with an interior-exterior point method for nonlinear programming. *Computational Optimization and Applications*, Springer, v. 29, p. 173–195, 2004.
- [126] SOUSA, V. D.; BAPTISTA, E. C.; COSTA, G. da. Loss minimization by the predictor–corrector modified barrier approach. *Electric Power Systems Research*, Elsevier, v. 79, n. 5, p. 803–808, 2009.
- [127] PINHEIRO, R. B. N. et al. Interior–exterior point method with global convergence strategy for solving the reactive optimal power flow problem. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 66, p. 235–246, 2015.

- [128] VANDERBEI, R. J.; SHANNO, D. F. An interior-point algorithm for nonconvex nonlinear programming. *Computational Optimization and Applications*, Springer, v. 13, p. 231–252, 1999.
- [129] WÄCHTER, A.; BIEGLER, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, Springer, v. 106, p. 25–57, 2006.
- [130] WALTZ, R. A. et al. An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical programming*, Springer, v. 107, n. 3, p. 391–408, 2006.
- [131] HONSBURG, C.; BOWDEN, S. Pv education. *ORG. Disponível em: <<http://pveducation.org/pvcdrom/properties-of-sunlight/sun-position-calculator/>> [Acessado em 05 de maio de 2022]*, 2022.
- [132] YORDANOV, G. H. Characterization and analysis of photovoltaic modules and the solar resource based on in-situ measurements in southern norway. *Norwegian University of Science and Technology, Department of Electric Power Engineering, Trondheim*, 2012.
- [133] MADETI, S. R.; SINGH, S. Monitoring system for photovoltaic plants: A review. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 67, p. 1180–1207, 2017.
- [134] VUGT, R. v. Quality assurance in solar with the use of I-V curves. In: *Eternal Sun Whitepaper*. [S.l.]: Eternal Sun, 2022.
- [135] MARTÍNEZ, M. A.; ANDÚJAR, J. M.; ENRIQUE, J. M. Temperature measurement in pv facilities on a per-panel scale. *Sensors*, MDPI, v. 14, n. 8, p. 13308–13323, 2014.
- [136] MIHAYLOV, B. et al. Uncertainty estimation of temperature coefficient measurements of pv modules. *IEEE Journal of Photovoltaics*, IEEE, v. 6, n. 6, p. 1554–1563, 2016.
- [137] BOYLESTAD, R. L.; NASHELKY, L. *Dispositivos Eletrônicos e teoria de circuitos*. [S.l.]: São Paulo: Pearson, 2013.
- [138] BREITENSTEIN, O. et al. The origin of ideality factors $n > 2$ of shunts and surfaces in the dark iv curves of si solar cells. In: WIP MUNICH. *Proceedings of the 21st European photovoltaic solar energy conference*. [S.l.], 2006. p. 625–628.
- [139] TOLEDO, F. et al. In-depth analysis of single-diode model parameters from manufacturer's datasheet. *Renewable Energy*, Elsevier, v. 163, p. 1370–1384, 2021.
- [140] LAUDANI, A.; FULGINEI, F. R.; SALVINI, A. Identification of the one-diode model for photovoltaic modules from datasheet values. *Solar energy*, Elsevier, v. 108, p. 432–446, 2014.
- [141] PHOTOVOLTAIC, C. S. T. Modules—design qualification and type approval. *IEC*, v. 1215, p. 2005–05, 2005.
- [142] SERA, D.; TEODORESCU, R.; RODRIGUEZ, P. Pv panel model based on datasheet values. In: IEEE. *2007 IEEE international symposium on industrial electronics*. [S.l.], 2007. p. 2392–2396.
- [143] VILLALVA, M. G.; GAZOLI, J. R.; FILHO, E. R. Comprehensive approach to modeling and simulation of photovoltaic arrays. *IEEE Transactions on power electronics*, IEEE, v. 24, n. 5, p. 1198–1208, 2009.

- [144] BRANO, V. L.; CIULLA, G. An efficient analytical approach for obtaining a five parameters model of photovoltaic modules using only reference data. *Applied Energy*, Elsevier, v. 111, p. 894–903, 2013.
- [145] DOBOS, A. P.; MACALPINE, S. M. Procedure for applying iec-61853 test data to a single diode model. In: IEEE. *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*. [S.l.], 2014. p. 2846–2849.
- [146] LI, B. et al. A robust i–v curve correction procedure for degraded photovoltaic modules. *Renewable Energy*, Elsevier, v. 224, p. 120108, 2024.
- [147] RICHTER, M. et al. Best practice guide on uncertainty in pv modelling. *Public report Performance Plus WP2 Deliverable D*, v. 2, 2015.
- [148] YANG, X.-S. *Nature-inspired optimization algorithms*. [S.l.]: Elsevier, 2014.
- [149] CHAPRA, S. C.; CANALE, R. P. *Numerical methods for engineers*. [S.l.]: Mcgraw-hill New York, 2021.
- [150] EDWARDS, S. Extension of algebraic solutions using the lambert w function. *arXiv preprint arXiv:1902.08910*, 2019.
- [151] JAIN, A.; KAPOOR, A. Exact analytical solutions of the parameters of real solar cells using lambert w-function. *Solar Energy Materials and Solar Cells*, Elsevier, v. 81, n. 2, p. 269–277, 2004.
- [152] MA, J. et al. Comparative performance on photovoltaic model parameter identification via bio-inspired algorithms. *Solar Energy*, Elsevier, v. 132, p. 606–616, 2016.
- [153] LÓCZI, L. Explicit and recursive estimates of the lambert w function. *arXiv preprint arXiv:2008.06122*, 2020.
- [154] FRITSCH, F. N.; SHAFER, R.; CROWLEY, W. Solution of the transcendental equation $wew = x$. *Communications of the ACM*, ACM New York, NY, USA, v. 16, n. 2, p. 123–124, 1973.
- [155] IACONO, R.; BOYD, J. P. New approximations to the principal real-valued branch of the lambert w-function. *Advances in Computational Mathematics*, Springer, v. 43, n. 6, p. 1403–1436, 2017.
- [156] MATLAB. *lambertw: Lambert W function*. Natick, Massachusetts: The MathWorks Inc., 2021.
- [157] VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020.
- [158] FLOUDAS, C. A.; GOUNARIS, C. E. A review of recent advances in global optimization. *Journal of Global Optimization*, Springer, v. 45, n. 1, p. 3–38, 2009.
- [159] FLOUDAS, C. A.; PARDALOS, P. M. *Frontiers in global optimization*. [S.l.]: Springer Science & Business Media, 2013. v. 74.
- [160] BUTENKO, S.; PARDALOS, P. M. Chapter 13: Deterministic global optimization. In: *Advances and Trends in Optimization with Engineering Applications*. [S.l.]: SIAM, 2017. p. 163–173.

- [161] ZHUANG, E.; YUE, D.; YOU, F. Spatial branch and bound method. 2015.
- [162] WÄCHTER, A. Chapter 17: Nonlinear optimization algorithms. In: *Advances and Trends in Optimization with Engineering Applications*. [S.l.]: SIAM, 2017. p. 221–235.
- [163] MATLAB. *fminsearch: search for local minimum of unconstrained multivariable function using derivative-free method*. Natick, Massachusetts: The MathWorks Inc., 2023.
- [164] FRANCO, R. A. P.; FILHO, G. L.; VIEIRA, F. H. T. Firefly algorithm applied to the estimation of the parameters of a photovoltaic panel model. *Advances in Nature-Inspired Computing and Applications*, Springer, p. 101–134, 2019.
- [165] MATLAB. *fsolve: solve system of nonlinear equations*. Natick, Massachusetts: The MathWorks Inc., 2023.
- [166] TOLEDO, F. J. et al. Photovoltaic modelling. *www.pvmodel.umh.es*, 2023.
- [167] MATLAB. *rand: uniformly distributed random numbers*. Natick, Massachusetts: The MathWorks Inc., 2021.
- [168] OLIVA, D. et al. A review on meta-heuristics methods for estimating parameters of solar cells. *Journal of Power Sources*, Elsevier, p. 126683, 2019.
- [169] OULCAID, M. et al. Parameter extraction of photovoltaic cell and module: Analysis and discussion of various combinations and test cases. *Sustainable Energy Technologies and Assessments*, Elsevier, v. 40, p. 100736, 2020.
- [170] DEOTTI, L. M. P.; JUNIOR, I. C. da S. A survey on the parameter extraction problem of the photovoltaic single diode model from a current–voltage curve. *Solar Energy*, Elsevier, v. 263, p. 111930, 2023.
- [171] STEWART, J. *Cálculo: Volume 1 - Tradução da sétima edição norte-americana*. [S.l.]: Cengage Learning, 2013.
- [172] HAYES, B. Why w? *American Scientist*, Sigma Xi, The Scientific Research Society, v. 93, p. 104–108, 2005.
- [173] VEBERIČ, D. Lambert w function for applications in physics. *Computer Physics Communications*, Elsevier, v. 183, n. 12, p. 2622–2628, 2012.
- [174] MATLAB. *cond: condition number for inversion*. Natick, Massachusetts: The MathWorks Inc., 2022.
- [175] MOLER, C. What is the condition number of a matrix? *MathWorks Blogs*. Disponível em: <<https://blogs.mathworks.com/cleve/2017/07/17/what-is-the-condition-number-of-a-matrix/>> [Acessado em 02 de novembro de 2022], 2017.
- [176] HIGHAM, N. J. *Accuracy and stability of numerical algorithms*. [S.l.]: SIAM, 2002.
- [177] MATLAB. *Floating-Point numbers*. Natick, Massachusetts: The MathWorks Inc., 2026.
- [178] HOWARD, R. M. Analytical approximations for the principal branch of the lambert w function. *European Journal of Mathematical Analysis*, v. 2, p. 14–14, 2022.

- [179] SCAVO, T.; THOO, J. On the geometry of halley's method. *The American mathematical monthly*, Taylor & Francis, v. 102, n. 5, p. 417–426, 1995.
- [180] BATZELIS, E. I.; ROUTSOLIAS, I. A.; PAPATHANASSIOU, S. A. An explicit PV string model based on the Lambert W function and simplified mpp expressions for operation under partial shading. *IEEE Transactions on Sustainable Energy*, IEEE, v. 5, n. 1, p. 301–312, 2013.
- [181] NOOR, M. A.; KHAN, W. A.; HUSSAIN, A. A new modified halley method without second derivatives for nonlinear equation. *Applied mathematics and computation*, Elsevier, v. 189, n. 2, p. 1268–1273, 2007.
- [182] SOLAIMAN, O. S.; HASHIM, I. Two new efficient sixth order iterative methods for solving nonlinear equations. *Journal of King Saud University-Science*, Elsevier, v. 31, n. 4, p. 701–705, 2019.
- [183] GOUALARD, F. The ins and outs of solving quadratic equations with floating-point arithmetic. *Authorea Preprints*, Authorea, 2023.
- [184] MATLAB. *linspace: generate linearly spaced vector*. Natick, Massachusetts: The MathWorks Inc., 2026.
- [185] DEOTTI, L. M. P.; JUNIOR, I. C. da S. A self-started predictor–corrector method for calculating the lambert w function within the scope of the photovoltaic single diode model. *Solar Energy*, Elsevier, v. 276, p. 112681, 2024.
- [186] MATLAB. Least-squares (model fitting) algorithms. The MathWorks Inc. *Disponível em: <<https://www.mathworks.com/help/optim/ug/least-squares-model-fitting-algorithms.html>> [Acessado em 23 de novembro de 2022]*, 2022.
- [187] MATLAB. Constrained nonlinear optimization algorithms. The MathWorks Inc. *Disponível em: <<https://www.mathworks.com/help/optim/ug/constrained-nonlinear-optimization-algorithms.html>> [Acessado em 23 de novembro de 2022]*, 2022.
- [188] CONN, A.; GOULD, N.; TOINT, P. A globally convergent lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds. *Mathematics of computation*, v. 66, n. 217, p. 261–288, 1997.
- [189] MATLAB. *fmincon: find minimum of constrained nonlinear multivariable function*. Natick, Massachusetts: The MathWorks Inc., 2022.
- [190] OLIVEIRA, E. J. et al. An optimal power flow based on safety barrier interior point method. *International Journal of Electrical power & energy systems*, Elsevier, v. 64, p. 977–985, 2015.
- [191] POLYAK, R. Modified barrier functions (theory and methods). *Mathematical programming*, Springer, v. 54, p. 177–222, 1992.
- [192] NASH, S. G.; POLYAK, R.; SOFER, A. A numerical comparison of barrier and modified barrier methods for large-scale bound-constrained optimization. *Large scale optimization: State of the Art*, Springer, p. 319–338, 1994.
- [193] BREITFELD, M. G.; SHANNO, D. F. Computational experience with penalty-barrier methods for nonlinear programming. *Annals of Operations Research*, Springer, v. 62, n. 1, p. 439–463, 1996.

- [194] SOUSA, V. D.; BAPTISTA, E. C.; COSTA, G. D. Optimal reactive power flow via the modified barrier lagrangian function approach. *Electric power systems research*, Elsevier, v. 84, n. 1, p. 159–164, 2012.
- [195] GRIVA, I.; POLYAK, R. A. Primal–dual methods for nonlinear constrained optimization. *Wiley Encyclopedia of Operations Research and Management Science*, Wiley Online Library, 2010.
- [196] KORT, B.; BERTSEKAS, D. Multiplier methods for convex programming. In: IEEE. *1973 IEEE Conference on Decision and Control including the 12th Symposium on Adaptive Processes*. [S.l.], 1973. p. 428–432.
- [197] GOULD, N. I. M. et al. Exploiting negative curvature directions in linesearch methods for unconstrained optimization. *Optimization Methods and Software*, v. 14, n. 1–2, p. 75–98, 2000.
- [198] MCCORMICK, G. P. A modification of armijo’s step-size rule for negative curvature. *Mathematical Programming*, v. 13, n. 1, p. 111–115, 1977.
- [199] KNIGHT, P. A.; RUIZ, D.; UÇAR, B. A symmetry preserving algorithm for matrix scaling. *SIAM journal on Matrix Analysis and Applications*, SIAM, v. 35, n. 3, p. 931–955, 2014.
- [200] NOCEDAL, J.; WÄCHTER, A.; WALTZ, R. A. Adaptive barrier update strategies for nonlinear interior methods. *SIAM Journal on Optimization*, SIAM, v. 19, n. 4, p. 1674–1693, 2009.
- [201] BENSON, H. Y.; VANDERBEI, R. J.; SHANNO, D. F. Interior-point methods for nonconvex nonlinear programming: Filter methods and merit functions. *Computational Optimization and Applications*, Springer, v. 23, n. 2, p. 257–272, 2002.
- [202] GOLUB, G. H.; LOAN, C. F. V. *Matrix computations*. [S.l.]: JHU press, 2013.
- [203] SQUIRE, W.; TRAPP, G. Using complex variables to estimate derivatives of real functions. *SIAM review*, SIAM, v. 40, n. 1, p. 110–112, 1998.
- [204] MARTINS, J. R.; STURDZA, P.; ALONSO, J. J. The complex-step derivative approximation. *ACM Transactions on Mathematical Software (TOMS)*, ACM New York, NY, USA, v. 29, n. 3, p. 245–262, 2003.

APÊNDICE A – FORMA EXPLÍCITA DO MODELO DE UM DIODO

Este apêndice apresenta, de maneira didática, a derivação algébrica que conduz da equação característica do modelo de um diodo, na forma implícita dada por (2.6), à sua forma explícita via função W de Lambert, conforme (2.21).

O objetivo é explicitar o encadeamento das manipulações necessárias para condicionar a transcendência do problema à estrutura canônica introduzida em (3.2), na qual a função W de Lambert pode ser aplicada diretamente, isto é, conforme (3.4). Por conveniência, essas duas expressões são retomadas a seguir.

$$w \exp(w) = z \quad (3.2)$$

$$W(z) := w \quad (3.4)$$

Assim, toma-se como ponto de partida a equação (2.6), transcrita abaixo.

$$I = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V + I(N_s/N_p)R_s)}{N_s nkT_c} \right) - 1 \right] - \frac{V + I(N_s/N_p)R_s}{(N_s/N_p)R_{sh}} \quad (2.6)$$

O primeiro passo consiste em separar, na última parcela do lado direito de (2.6), o termo racional que contém I . Ao transferir essa contribuição para o lado esquerdo, cancelar fatores comuns e reorganizar os coeficientes multiplicativos e racionais, obtém-se (A.1).

$$I \left(\frac{R_{sh} + R_s}{R_{sh}} \right) = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V + I(N_s/N_p)R_s)}{N_s nkT_c} \right) - 1 \right] - \frac{V}{(N_s/N_p)R_{sh}} \quad (A.1)$$

Em seguida, divide-se ambos os lados de (A.1) pelo fator $(R_{sh} + R_s)/R_{sh}$, obtendo-se (A.2).

$$I = \frac{N_p I_{ph} R_{sh}}{R_{sh} + R_s} - \frac{N_p I_0 R_{sh}}{R_{sh} + R_s} \left[\exp \left(\frac{q(V + I(N_s/N_p)R_s)}{N_s nkT_c} \right) - 1 \right] - \frac{V}{(N_s/N_p)(R_{sh} + R_s)} \quad (A.2)$$

A expressão acima ainda reúne, em um mesmo termo, a dependência exponencial em V e em I . Para evidenciar a contribuição de I no expoente, expõem-se os termos entre colchetes e decompõe-se a exponencial associada. Com isso, (A.2) assume a forma apresentada em (A.3).

$$I = \frac{N_p R_{sh}(I_{ph} + I_0) - (N_p/N_s)V}{R_{sh} + R_s} - \frac{N_p I_0 R_{sh}}{R_{sh} + R_s} \exp \left(\frac{qV}{N_s nkT_c} \right) \exp \left(\frac{qIR_s}{N_p nkT_c} \right) \quad (A.3)$$

Neste ponto, é útil introduzir parâmetros auxiliares que simplificam a notação e tornam mais transparente a estrutura matemática do problema. Recorrendo aos parâmetros A , B e C definidos em (A.4), (A.3) pode ser reexpressa na forma compacta mostrada em (A.5).

$$A = \frac{N_p R_{sh}(I_{ph} + I_0) - (N_p/N_s)V}{R_{sh} + R_s} \quad (A.4a)$$

$$B = \frac{N_p I_0 R_{sh}}{R_{sh} + R_s} \exp \left(\frac{qV}{N_s nkT_c} \right) \quad (A.4b)$$

$$C = \frac{qR_s}{N_p nkT_c} \quad (A.4c)$$

$$I = A - B \exp(CI) \quad (\text{A.5})$$

A seguir, isola-se o termo exponencial em (A.5) e multiplica-se ambos os lados por $-C \exp(C(A - I))$. Essa escolha não é arbitrária: esta permite obter exatamente o produto de uma variável por sua própria exponencial. Assim, obtém-se (A.6), na qual a transcendência já se manifesta na forma requerida.

$$\underbrace{C(A - I)}_w \exp(\underbrace{C(A - I)}_w) = \underbrace{BC \exp(AC)}_z \quad (\text{A.6})$$

Conforme assinalado, a estrutura matemática de (A.6) é análoga àquela da expressão originária da função W de Lambert apresentada em (3.2), desde que se adotem as identificações $w \equiv C(A - I)$ e $z \equiv BC \exp(AC)$. Logo, ao aplicar a definição dada em (3.4), chega-se a (A.7).

$$C(A - I) = W(BC \exp(AC)) \quad (\text{A.7})$$

A equação (A.7) já fornece a relação implícita necessária para obter I . Ao isolar a corrente, obtém-se imediatamente que a equação característica do modelo de um diodo pode ser expressa de forma *explícita* conforme (A.8).

$$I = A - \frac{1}{C} W(BC \exp(AC)) \quad (\text{A.8})$$

Resta, então, explicitar o argumento da função W de Lambert em termos dos parâmetros do modelo. Para isso, avaliam-se separadamente os produtos BC e AC de acordo com (A.4), obtendo-se as expressões intermediárias mostradas em (A.9) e (A.10).

$$\begin{aligned} BC &= \left[\frac{N_p I_0 R_{sh}}{R_{sh} + R_s} \exp\left(\frac{qV}{N_s nkT_c}\right) \right] \left[\frac{qR_s}{N_p nkT_c} \right] \\ &= \frac{N_p I_0 R_{sh}}{R_{sh} + R_s} \frac{qR_s}{N_p nkT_c} \exp\left(\frac{qV}{N_s nkT_c}\right) \\ &= \frac{qI_0 R_{sh} R_s}{nkT_c (R_{sh} + R_s)} \exp\left(\frac{qV}{N_s nkT_c}\right) \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} AC &= \left[\frac{N_p R_{sh} (I_{ph} + I_0) - (N_p/N_s)V}{R_{sh} + R_s} \right] \left[\frac{qR_s}{N_p nkT_c} \right] \\ &= \frac{qR_s}{nkT_c (R_{sh} + R_s)} \left[R_{sh} (I_{ph} + I_0) - \frac{V}{N_s} \right] \\ &= \frac{qR_s}{N_s nkT_c (R_{sh} + R_s)} [N_s R_{sh} (I_{ph} + I_0) - V] \end{aligned} \quad (\text{A.10})$$

A partir de (A.9) e (A.10), obtém-se que o argumento da função W de Lambert presente em (A.8) é definido por (A.11).

$$BC \exp(AC) = \frac{qI_0 R_{sh} R_s}{nkT_c (R_{sh} + R_s)} \exp\left(\frac{qR_{sh} [N_s R_s (I_{ph} + I_0) + V]}{N_s nkT_c (R_{sh} + R_s)}\right) \quad (\text{A.11})$$

À luz de (A.4), (A.8) e (A.11), chega-se, por fim, à expressão (A.12), que transcreve a equação característica do modelo de um diodo via função W de Lambert e é equivalente a (2.21).

$$I = \frac{N_p R_{sh} (I_{ph} + I_0) - (N_p / N_s) V}{R_{sh} + R_s} - \frac{N_p n k T_c}{q R_s} W(z) \quad (\text{A.12a})$$

$$z = \frac{q I_0 R_{sh} R_s}{n k T_c (R_{sh} + R_s)} \exp\left(\frac{q R_{sh} [N_s R_s (I_{ph} + I_0) + V]}{N_s n k T_c (R_{sh} + R_s)}\right) \quad (\text{A.12b})$$