Improving the small-signal stability of a stochastic power system
Algorithms and mathematical analysis

Wang, Zhen; Xi, Kaihua; Cheng, Aijie; Lin, Hai Xiang; van Schuppen, Jan H.

DOI
10.1016/j.chaos.2025.116616

Publication date
2025

Document Version
Final published version

Published in
Chaos, Solitons and Fractals

Chaos, Solitons and Fractals 199 (2025) 116616

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

Improving the small-signal stability of a stochastic power system — Algorithms and mathematical analysis

Zhen Wang [a],*, Kaihua Xi [a], Aijie Cheng [a], Hai Xiang Lin [b,c], Jan H. van Schuppen [b]

[a] School of Mathematics, Shandong University, Jinan, 250100, Shandong Province, China
[b] Department of Applied Mathematics, Delft University of Technology, Delft, 2628CD, The Netherlands
[c] Institute of Environmental Sciences (CML), Leiden University, Leiden, 2333CC, The Netherlands

ARTICLE INFO

Keywords:
Fluctuations in power flows
Optimal power dispatch
Exit probability from a safe subset
Constrained nonlinear optimization
Nonconvex and nondifferentiable
Convergence analysis

ABSTRACT

Tools and analysis for improving the small-signal stability of a stochastic power system by optimal power dispatch in each short time horizon, such as five-minute intervals, are provided in this paper. An objective function which characterizes the maximal exit probability from the static stability region $(-\pi/2, \pi/2)$ across the phase-angle differences of all power lines is formulated. This objective function is proven to be Lipschitz continuous, nondifferentiable, and nonconvex, with a finite minimum defined over the region of power supply vectors. The formulas of the generalized subgradient and directional derivative of the objective function are provided, and based on these formulas, a two-step algorithm is designed to approximate a minimizer accompanied by the convergence proof: (1) using a projected generalized subgradient method to compute an effective initial vector, and (2) applying the steepest descent method to approximate a local minimizer. The algorithms have been verified using a synthesized power network, demonstrating computational validity and effectiveness in minimizing the maximal exit probability of all power lines.

1. Introduction

Motivation. The 21st-century power system differs significantly from that of the 20th century due to the gradual shift from fossil power sources to renewable energy sources, and an increase in power loads. An effect of this change is that the power supply to the transmission net will not be steady but will gradually have more and more fluctuations owing to variations caused by the weather and human factors. A serious concern is whether these fluctuations will endanger the stability of the power system.

The improvement of the small-signal stability of a transmission net of a stochastic power system, is the focus of this paper. The power system will be controlled by determining the power dispatch, the specification of the power supply during a short horizon, so that it minimizes the maximal exit probability from a safe subset of all phase-angle differences of the power network. An objective function is defined concerning the maximal exit probability over all power lines. An infimization procedure is proposed to compute the power dispatch to minimize the objective function. The algorithms are proven to converge to a local minimum.

The *objective function* is formulated as the maximum value among a finite set of secondary functions. Each secondary function is a sum of two constituent functions: the standard deviation component denoted as σ_k, exhibits non-convex characteristics, while the other absolute mean component is referred to as $|m_k|$, may lack differentiability in cases where a component of the argument becomes zero. The objective function satisfies a generic differentiability property but exhibits nondifferentiability on either an algebraic subset or a subset of lower algebraic complexity. The domain of power supply vectors denoted as P^+, takes the form of a polytope.

The Problem. The approach to the considered problem, based on periodic recomputation of an optimal set of power supply vectors, leads to an infimization problem for the associated objective function.

Literature review. Literature on stability analysis and control of power systems is reviewed to clarify the context and significance of the problem investigated in this paper. A review of literature on nonconvex and nondifferential optimization problems and relevant solution algorithms is presented to explain the mathematical techniques and methods used in this paper and to propose potential improvements to the current approach.

Stability analysis and control of a conventional power system are well known, [1], and for concepts of stability of power systems, the reader is referred to [2]. Currently, attention has been shifted to

* Corresponding author at: School of Mathematics, Shandong University, Jinan, 250100, Shandong Province, China.
E-mail addresses: Wangzhen17@mail.sdu.edu.cn, wangzhen17.sdu@vip.163.com (Z. Wang), kxi@sdu.edu.cn (K. Xi), aijie@sdu.edu.cn (A. Cheng), H.X.Lin@tudelft.nl (H.X. Lin), J.H.vanSchuppen@tudelft.nl (J.H. van Schuppen).

https://doi.org/10.1016/j.chaos.2025.116616
Received 18 January 2025; Received in revised form 11 May 2025; Accepted 17 May 2025
Available online 9 June 2025

power systems with renewable energy. See [3], the small-signal stability and transient stability of the power system are analyzed when utility-scale photovoltaic units are integrated into the power system to replace the conventional synchronous generators. The model is further enhanced by a coordinated voltage–frequency control, which can additionally provide primary frequency control, whose response to frequency deviation is faster than conventional plants in their tested power systems, [4]. Concurrently, the number of electric vehicles connected to the power grid, acting as loads or generators, is increasing, which will contribute to the reduction of the CO_2 emission. See [5] for two tools, primary frequency control and dynamic grid support, to control the transient stability of power systems when the vehicle-to-grid grows. Here in this research, the focus is on the situation where a large amount of renewable sources is integrated into the power system, which introduces uncertainty and fluctuation, necessitating the investigation on a stochastic power system.

Extensive research has been conducted on stochastic power systems, [6–8]. The focus of the investigation has been on simulations and the approximation of the probability distribution of the first exit time from the domain of attraction, [9,10]. The latter requires computing the type-one synchronous states and numerically approximating the probability distribution, which has high computational complexity, and requires analytical techniques. To the best of the author's knowledge, the approach to the corresponding optimization problem has not been found in the literature. In order to reduce the computational burden, we linearize the power system at the steady state of the power flow equation and investigate the stability of the stochastic linearized power system, which is called *small-signal* stability [2]. An objective function is then formulated, which describes the maximal exit probability from the static angle stability region $(-\pi/2, \pi/2)$ across the phase-angle differences of all power lines. Due to the constraints on the power dispatch, the problem is formulated as a constrained nonlinear optimization problem, which is nonconvex and nondifferentiable. The nondifferentiability stems from the representation of the function as the maximum of several functions, and each such function contains a term that defies differentiability at zero. The nonconvexity is due to the Lyapunov matrix equation.

Optimizing nonconvex and nondifferentiable objective functions is a challenging task. A nonconvex optimization problem is difficult because the objective function has many minimizers. A nondifferentiable optimization problem is even more challenging because the gradient, which is a fundamental tool that is often used in the iterations of optimization algorithms, is not everywhere defined. Instead, either a generalized subgradient or a generalized gradient is used to generalize the concept of a gradient to the set of nondifferentiable functions. The initial research of F. Clarke on subgradients [11,12] is well known. B. Mordukhovich and others developed optimization of nondifferentiable and nonconvex functions, [13–16]. For basic insights into semiconvexity, semicontinuity, and various properties associated with these function classes, [17] is referred to. Additionally, more comprehensive information on generalized gradients and generalized subgradients can be found in references such as [11,18,19]. Readers who are interested in optimization criteria and methodologies related to these optimization techniques, are recommended exploring [20,21].

In addition to the generalized gradient, which represents a vector class, researchers have also leveraged directional derivatives to optimize nonsmooth functions. Notably, A. Ben-Tal introduced second-order necessary and sufficient optimality criteria for four distinct types of nonsmooth minimization problems in [22]. Subsequently, the same author extended the application of both first and second directional derivatives to establish necessary and sufficient conditions for a strict local minimizer, as elaborated in [23]. Expanding on this work, A. Shapiro provided a comprehensive overview of various directional derivative concepts in [24].

In this paper, both the generalized subgradient and the directional derivative will be used to infimize our objective function. A two-step procedure is then formulated. The first step uses a projected generalized subgradient method to determine an effective initial position, and then the second step presents the steepest descent method starting from that initial position and converging to a local minimizer. Because the iterations rely on either the generalized subgradient or the first directional derivative, the proposed algorithms are first-order algorithms.

Contributions of this Paper. (1) A nonlinear objective function characterizing the maximal exit probability of all power lines, concerning the small-signal stability of stochastic power systems, and the convexity, differential analysis of the objective function per direction vector, where the formulas of directional derivatives provide a different understanding of the Jacobian matrix compared to existing literature. (2) The numerical algorithm for computing the gradient vector and the Hessian matrix of the standard deviation within the objective function, and the two-step algorithm with its convergence proof for approximating a local minimizer (3) Verification and application of the proposed algorithms to minimize the maximal exit probability of all power lines by using an academic power network as an illustration.

Paper Organization. Section 2 provides an introduction to power system fundamentals, the domain of power supply vectors, and the procedure for the calculation of the objective function. Section 3 examines its convexity, and provides the directional derivatives. Section 4 presents the algorithms: (1) the numerical algorithm mentioned above and (2) the two-step algorithm. Section 5 displays the reliability and effectiveness of the proposed algorithms to improve the small-signal stability of stochastic power systems by an academic power network example. Section 6 summarizes conclusions and outlines open research challenges.

2. The constrained nonlinear optimization problem

Notation. The following mathematical notation is used in this paper. Denote the integers, the positive integers, and the natural numbers respectively by $\mathbb{Z}$, $\mathbb{Z}_+ = \{1, 2, \ldots\}$, and $\mathbb{N} = \{0, 1, 2, \ldots\}$. For any integer $n \in \mathbb{Z}_+$ denote the finite sets $\mathbb{Z}_n = \{1, 2, \ldots, n\}$ and $\mathbb{N}_n = \{0, 1, 2, \ldots, n\}$. Denote the positive and the strictly positive real numbers, respectively by $\mathbb{R}_+$ and $\mathbb{R}_{s,+}$. Define the sign function as $\mathrm{sign}(x) = +1$ if $x > 0$, -1 if $x < 0$, and 0 if $x = 0$. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$ and it is equipped with the inner product $\langle \cdot, \cdot \rangle$, the infinity norm $\| \cdot \|_\infty$, and the Euclidean norm $\| \cdot \|_2$. For any integer $n \in \mathbb{Z}_+$, $\mathbb{R}^n_+$ and $\mathbb{R}^n_{s,+}$ denote respectively the n-fold product of $\mathbb{R}_+$ and $\mathbb{R}_{s,+}$.

The set of real matrices of size $m \times n$ is denoted by $\mathbb{R}^{m\times n}$. The matrix transpose operator is denoted by $\top$. The spectrum norm and the Frobenius norm of a matrix are denoted by $\|\cdot\|_{2,s}$ and $\|\cdot\|_F$, respectively. Denote the nth row of a matrix $A \in \mathbb{R}^{m\times n}$ by $A(n)$ and the nth column of a matrix A is denoted by A_n. Denote the set of diagonal matrices of size $n \times n$ with positive or strictly positive elements as $\mathbb{R}^{n\times n}_{+,\mathrm{diag}}$ and $\mathbb{R}^{n\times n}_{s+,\mathrm{diag}}$, respectively. A diagonal matrix with elements from the vector $v \in \mathbb{R}^n$ is represented as $\mathrm{diag}(v) \in \mathbb{R}^{n\times n}$. Unit vectors: e_k is used to represent the kth unit vector. Identity matrix: the identity matrix of size n is denoted as I_n. Column selection: For $m > n$, the nth to mth columns of the product matrix $(A\ B\ C)$ is denoted as $(A\ B\ C)([n : m])$.

Hadamard product and power: The Hadamard product and Hadamard power, representing component-wise multiplication or power, is denoted as $A \circ B$ or $A^{\circ x}$, separately. The scalar–vector product of a real number $c \in \mathbb{R}$ and a vector $v \in \mathbb{R}^n$ is denoted by $c \cdot v$.

2.1. Problem introduction

Power system stability control is managed across different time scales: primary control operates within seconds, secondary control spans 3–5 min, and tertiary control extends to 15 min. The approach to the considered problem in this paper is executed in a secondary or tertiary control framework ; hence there is used a sequence of short horizons of a few minutes, think of 3 to 15 min. In each short horizon, the power dispatch is adjusted and kept constant during this short horizon. A generally accepted sufficient condition for stability is that the phase-angle difference over any power flow remains in the safe subset $(-\pi/2,\ +\pi/2) \subset \mathbb{R}$.

Due to the fact that at every time step, the latest output of the power system is used for the computation of the future power demand [25, Subsection 3.1], the form of the approach is output feedback. The input of the power system at the next time step is a constant input with possibly a different value at each node of the power system. The vector of input values is determined as an approximation of the minimizer of the objective function. The objective function value corresponds to the maximal exit probability of all power lines from the safe subset $(-\pi/2, \pi/2)$ in the next time horizon. Details for the optimization problem will be provided below.

Because of disturbances of power sources and of power loads, which are expected to increase in the future, the phase-angle differences of power lines become random variables. The performance criterion is to minimize the maximal probability that any phase-angle difference exits from the safe subset $(-\pi/2, \pi/2)$. An objective function that describes the criterion with low computational complexity is needed. Then one obtains the objective function defined below in Definition 2.3, in which other minor approximations are used. For further details of the objective function, the reader is referred to the paper [25]. A constrained optimization problem for the power dispatch of the power system is then formulated in each short horizon. A solution to the optimization problem ensures that the maximal probability that any phase-angle difference will exit from the safe subset during a short horizon is sufficiently small. Hence, the small-signal stability of the stochastic power system is improved.

2.2. Power system model and the power supply vector

The stochastic power system employed in this paper is a network of interconnected oscillators (swing equations) driven by Brownian motion. To limit the computational complexity of determining the probability distribution of a stochastic nonlinear power system, attention is restricted to a stochastic linearized power system which is described as follows, [25, Section 2.8],[1,26], of which several other simplifications are made, e.g. the voltages are assumed to have been adjusted properly, thus their dynamics are omitted from the model,

$$dx(t) = J(\theta_s, 0)\ x(t)\ dt + K\ dv(t),\ x(0) \in G(m_{x(0)}, \sigma_{x(0)}), \tag{1}$$

$$y(t) = Cx(t), \tag{2}$$

$$x : \Omega \times T \to \mathbb{R}^{n_\mathcal{V}},\ y : \Omega \times T \to \mathbb{R}^{n_\mathcal{E}},$$

$$x(t) = \begin{bmatrix} \theta(t) \\ \omega(t) \end{bmatrix},\ K = \begin{bmatrix} 0 \\ K_2 \end{bmatrix} \in \mathbb{R}^{2n_\mathcal{V} \times n_\mathcal{V}},$$

$$J(\theta_s,\ 0) = \begin{bmatrix} 0 & I_{n_\mathcal{V}} \\ -M^{-1} BW F(\theta_s) & -M^{-1} D \end{bmatrix}, \tag{3}$$

$$C = \begin{bmatrix} B^\top & 0 \end{bmatrix} \in \mathbb{R}^{n_\mathcal{E} \times 2n_\mathcal{V}},$$

Note that this model is used to analyze the influence of the small disturbance, and give insight for the moderate disturbance which is called small-signal stability of the power system, and therefore, large disturbances, which rarely happen in reality, are not considered in this paper. For a non-Gaussian noise, one has to model the noise process as the output of a nonlinear stochastic system driven by a Brownian motion process. If there exists a probability distribution function of the state, then it satisfies a partial differential equation that can only be numerically approximated. This approach is of high complexity; one may refer to [27], and it is left for future research. The power network is modeled by a graph $G = (V,\ E)$ with the set of vertices V and the set of edges in E. There are $n_V \in \mathbb{Z}_+$ vertices and $n_E \in \mathbb{Z}_+$ edges. A line between two buses of a power network is modeled by an edge denoted by $k = (i_k,\ j_k) \in E$ which connects vertices i_k and j_k. Nodes with only power generation are indexed according to $n_1,\ n_2,\ \ldots,\ n^+$ and Nodes with only power demand are indexed by $n^+ + 1, n^+ + 2, \ldots, n_V$. The diagonal matrices of the strictly positive inertias and the positive damping coefficients are denoted by $M = \mathrm{diag}([m_1, \ldots, m_{n_V}])$ and $D = \mathrm{diag}([d_1, \ldots, d_{n_V}])$, separately. The network incidence matrix is denoted by $B \in \mathbb{R}^{n_V \times n_E}$ and the weight matrix is denoted by $W = \mathrm{diag}([w_1, \ldots, w_{n_E}]) \in \mathbb{R}^{n_E \times n_E}_{s+,\mathrm{diag}}$.

The tuple (Ω, F, P) denotes a probability space with a set Ω, a σ-algebra F, and a probability measure P. Denote by the matrix $K_2 \in \mathbb{R}^{n_V \times n_V}_{+,\mathrm{diag}}$ the standard deviation of vector-valued Brownian motion acting on the frequencies of the nodes. The Brownian motion process is denoted by $v : \Omega \times T \to \mathbb{R}^{n_V}$ which has independent increments for nonoverlapping intervals and satisfies, $v(0) = 0,\ \forall s,\ t \in T, s < t,\ v(t) - v(s) \in G(0, (t-s) \times I_{n_V})$. Furthermore, $F^{x(0)}, F^{v}_{\infty}$ are independent σ-algebras.

Denote, for a short horizon, the maximal available power supply by $p^{+,max} \in \mathbb{R}^{n^+}_+$; the prediction of the power demand by $p^- \in \mathbb{R}^{n_V - n^+}_+$; the sum of the maximal power supply by $p^{+,max}_{sum} = \sum_{i=1}^{n^+} p^{+,max}_i \in \mathbb{R}_{s,+}$; and the sum of the predicted power demand by $p^-_{sum} = \sum_{i=n^++1}^{n_V} p^-_i \in \mathbb{R}_{s,+}$. It is assumed that the sum of the maximal available power supply is larger than or equal to the sum of the predicted demand, $p^{+,max}_{sum} \geq p^-_{sum}$. Because the sum of the power supply has to equal the sum of the predicted power demand, the control vector of the power supply can be defined as $p_s = \left(p^+_1,\ \ldots,\ p^+_{n^+-1}\right)^\top \in \mathbb{R}^{n^+-1}_+$. From p_s one can compute the last element of the vector of power supplies, $p^+_{n^+} = p^-_{sum} - \sum_{i=1}^{n^+-1} p_s(i)$.

2.3. The domain of power supply vectors

Definition 2.1. Define the domain of the power supply vector $p_s \in \mathbb{R}^{n^+-1}_+$, depending on the maximal power supply $p^{+,max}$ and the sum of the predicted power demand p^-_{sum}, as the set,

$$P^+ = P^+(p^{+,max},\ p^-_{sum}) = \left\{ p_s \in \mathbb{R}^{n^+-1}_+ |\ b_1 \leq A_1\, p_s,\ p_s \leq b_2 \right\}, \tag{4}$$

$$b_1(i) = p^-_{sum} - p^{+,max}_{n^+} - \cdots - p^{+,max}_{i+1},\ b_2(i) = p^{+,max}_i,\ b_1,\ b_2 \in \mathbb{R}^{n^+-1},$$

$$A_1(i,j) = \begin{cases} 1 & \text{if } i \leq j \\ 0 & \text{if } i > j, \end{cases} \quad A_1 \in \mathbb{R}^{(n^+-1)\times(n^+-1)}.$$

The domain P^+ is set to meet the power demand and realistic constraints.

Remark 2.2. The domain P^+ is compact and convex. Moreover, it is a polytope.

2.4. The objective function

Definition 2.3. Define the objective and related functions according to,

$$\forall\ k \in \mathbb{Z}_{n_E},\ k = (i_k,\ j_k),\ f_{as} : P^+ \to \mathbb{R}^{n_E},\ \forall\ p_s \in P^+,$$

$$|m_k(p_s)| = f_{as,k}(p_s) = \arcsin\left(\left|\left(A\, p_s + b\right)_k\right|\right),\ f_{as,k} : P^+ \to \mathbb{R}_+,$$

$$f_k(p_s) = f_{as,k}(p_s) + r \cdot \sigma_k(p_s) \in \mathbb{R}_+,\ \forall\ k \in \mathbb{Z}_{n_E},\ f_k : P^+ \to \mathbb{R}_+,$$

$$f(p_s) = \|f_k(p_s)\|_\infty = \max_{k \in \mathbb{Z}_{n_E}} f_k(p_s),\ f : P^+ \to \mathbb{R}_+.$$

The objective function f corresponds to an upper bound of the maximal exit probability across all power lines associated with the parameter $r \in (0, \infty)$ according to the invariant probability distribution. The function $f_{as,k} = |m_k|$ for a power supply vector p_s stands for the absolute value of the mean of the phase-angle difference of power line k. See [28, pp. 95–97] for properties of the arcsin function. The function σ_k represents the standard deviation of the phase-angle difference of that line (standard deviation as understood in statistics). Since it is an implicit function within the objective function f, it will be also referred to as an implicit function in the following sections.

Procedure 2.4. *Computation of the objective function for a power supply vector.*

Input data. The parameters of the power system include the graph of the connected power network $G = (V,\ E)$ and the matrices $M,\ D,\ B,\ W,\ K_2$.

$$
\begin{aligned}
J_d\left(p_s\right) &= \begin{bmatrix} 0_{(n_V-1)\times(n_V-1)}, & 0_{(n_V-1)\times 1} \mid I_{n_V-1} \\ -\left(U\left(p_s\right)^\top M^{-\frac{1}{2}}\, BW(p_s)B^\top\, M^{-\frac{1}{2}}U\left(p_s\right)\right)([2:n_V]), & -U(p_s)^\top M^{-1}\, D\, U(p_s) \end{bmatrix}, \\
&\in \mathbb{R}^{(2n_V-1)\times(2n_V-1)} \\
K_d\left(p_s\right) &= \begin{bmatrix} 0_{(n_V-1)\times n_V} \\ U\left(p_s\right)^\top M^{-\frac{1}{2}}\, K_2 \end{bmatrix}, C_d\left(p_s\right) = \begin{bmatrix} \left(B^\top M^{-\frac{1}{2}}U(p_s)\right)([2:n_V]), & 0_{n_E\times n_V} \end{bmatrix}, \\
&\in \mathbb{R}^{(2n_V-1)\times n_V}, \qquad \in \mathbb{R}^{n_E\times(2n_V-1)}
\end{aligned} \tag{10}
$$

Box I.

The vectors $p^{+,max}$ and p^-. A power supply vector $p_s \in P^+$ and the vector of power demand $p^- = \begin{bmatrix} p^-_{n^++1} & \cdots & p^-_n \end{bmatrix}^\top \in \mathbb{R}^{n_V-n^+}$. Finally, the parameter of the objective function $r \in \mathbb{R}_{s+}$.

1. *Compute an orthonormal matrix U, the matrix A, and the vector b according to,*

$$
\begin{aligned}
&U^\top \Lambda^\dagger U = \left(B\, W\, B^\top\right)^\dagger, \quad U = \begin{bmatrix} U_1 & U_2 & \cdots & U_{n_E} \end{bmatrix}, \\
&A = B^\top U^\top \Lambda^\dagger \begin{bmatrix} U_1 - U_{n^+}, & U_2 - U_{n^+}, & \ldots & , U_{n^+-1} - U_{n^+} \end{bmatrix} \\
&\quad \in \mathbb{R}^{n_E\times(n^+-1)}, \\
&b = B^\top U^\top \Lambda^\dagger \begin{bmatrix} U_{n^++1} - U_{n^+}, & \ldots & , U_n - U_{n^+} \end{bmatrix} (-p^-) \in \mathbb{R}^{n_E}.
\end{aligned}
$$

2. *Solve the following Lyapunov equation for the unique matrix $Q_x \in \mathbb{R}^{2n_V\times 2n_V}$ and then compute the variance matrix $Q_y \in \mathbb{R}^{n_E\times n_E}$,*

$$0 = J\left(p_s\right)\, Q_x + Q_x\, J\left(p_s\right)^\top + K\, K^\top, \tag{5}$$

$$Q_y = C\, Q_x\, C^\top, \text{ where,} \tag{6}$$

$$J\left(p_s\right) = \begin{bmatrix} 0_{n_V\times n_V} & I_{n_V} \\ -M^{-1}\, B\, W\, F(p_s) & -M^{-1}\, D \end{bmatrix} \in \mathbb{R}^{2n_V\times 2n_V}, \tag{7}$$

$$F\left(p_s\right) = \operatorname{diag}\left(\cos\left(\arcsin\left(A\, p_s + b\right)\right)\right)\, B^\top \in \mathbb{R}^{n_E\times n_V}, \tag{8}$$

$$K = \begin{bmatrix} 0 \\ K_2 \end{bmatrix} \in \mathbb{R}^{2n_V\times n_V}, \quad C = \begin{bmatrix} B^\top & 0 \end{bmatrix} \in \mathbb{R}^{n_E\times 2n_V}. \tag{9}$$

3. *Compute the values of standard deviation σ_k and of the variance V_k for the power supply vector $p_s \in P^+$,*

$$
\begin{aligned}
&\sigma_k\left(p_s\right) = Q_y(k,k)^{\frac{1}{2}}, \ \sigma\left(p_s\right) = \begin{bmatrix} \sigma_1\left(p_s\right) & \cdots & \sigma_{n_E}\left(p_s\right) \end{bmatrix}^\top, \ \sigma : P^+ \to \mathbb{R}^{n_E}, \\
&V_k\left(p_s\right) = \sigma_k^2\left(p_s\right), \ V\left(p_s\right) = \begin{bmatrix} V_1\left(p_s\right), & \cdots, & V_{n_E}\left(p_s\right) \end{bmatrix}^\top, \ V : P^+ \to \mathbb{R}^{n_E}.
\end{aligned}
$$

4. *Compute the value of the objective function for the power supply vector $p_s \in P^+$,*

$$
\begin{aligned}
&|m_k\left(p_s\right)| = f_{as,k}\left(p_s\right) = \arcsin\left(\left|\left(A\, p_s + b\right)_k\right|\right) = \arcsin\left(\left|A(k)\, p_s + b_k\right|\right), \\
&f_k\left(p_s\right) = f_{as,k}\left(p_s\right) + r\cdot\sigma_k\left(p_s\right) \in \mathbb{R}_+, \ \forall\, k = (i_k, j_k) \in \mathbb{Z}_{n_E}, \\
&f\left(p_s\right) = \|f_k\left(p_s\right)\|_\infty = \max_{k\in\mathbb{Z}_{n_E}} f_k\left(p_s\right).
\end{aligned}
$$

Definition 2.5. Define the weight matrix function W concerning $p_s \in P^+$, as $W\left(p_s\right) = W\ \operatorname{diag}\left(\cos\left(\arcsin\left(A\, p_s + b\right)\right)\right)$. This leads to $B\, W\ F(\theta_s) = B\, W(p_s)\, B^\top$.

Remark 2.6. If the matrix $J\left(p_s\right)$ is Hurwitz, then the matrix Q_x is the unique solution of a Lyapunov matrix Eq. (6). But in fact, the system matrix $J(\theta_s, 0)$ is not Hurwitz because it has a zero eigenvalue due to the product matrix $BW\,F(p_s)$ having a zero eigenvalue. Therefore, there is a need for a reduction procedure, see [26,29]. After the deduction, the matrices in Eq. (5)), Eq. (6) become as Eq. (10) in Box I where the matrix $U\left(p_s\right)^\top M^{-\frac{1}{2}}\, BW(p_s)B^\top\, M^{-\frac{1}{2}}U\left(p_s\right)$ is diagonal, and $J_d\left(p_s\right)$ becomes Hurwitz. Therefore, Eq. (5) has a unique positive-definite solution. Additionally, the value of the implicit function σ_k for a vector p_s can be computed by,

$$\sigma_k\left(p_s\right) = \left[C_d\left(p_s\right) \int_0^\infty e^{J_d(p_s)t} K_d\left(p_s\right)\, K_d\left(p_s\right)^\top e^{J_d(p_s)^\top t} dt\ C_d\left(p_s\right)^\top\right]^{\frac{1}{2}}(k,k), \tag{11}$$

It is proven in [25] that, with respect to a condition on K, the matrix Q_y is strictly positive definite, hence that, for all $i \in \mathbb{Z}_{n_y}$, $Q_y(i,i) > 0$. Note that the values of the diagonal of the matrix K_2 have an effect on the numerical outcome of the computations, but not on the optimization procedure.

2.5. The optimization problem and the existence of a minimizer

To minimize the maximal exit probability from the safe subset $(-\pi/2, \pi/2)$ across the phase-angle differences of all power lines of a power network, with the constraints on the power supply vectors, lead to the following constrained nonlinear optimization problem.

Problem 2.7. Solve the minimization problem for the objective function, and determine a value $a \in \mathbb{R}$ and a minimizer $p_s^* \in P^+$ such that,

$$a = f\left(p_s^*\right) = \inf_{p_s\in P^+} f\left(p_s\right) = \inf_{p_s\in P^+} \max_{k\in\mathbb{Z}_{n_E}} \left[|m_k\left(p_s\right)| + r\cdot\sigma_k\left(p_s\right)\right].$$

If for a parameter $r \in (0,1)$, the value a of the optimization criterion satisfies $a < \pi/2$, then there exists a power supply vector $p^+ \in P^+$ such that the probability of exiting the safe subset is less than 2ϵ, and the new exit probability can also be computed [25, Appendix 1.4(b)].

Proposition 2.8. *The objective function f is Lipschitz continuous, and the value range of f is a closed interval, denoted as $R(f) = [a,\ c]$ for $a,\ c \in \mathbb{R}$ where $-\infty < a \leq c$.*

The details of the proof of Proposition 2.8 are omitted here, and readers interested can find it in [30, Section 3]. From Proposition 2.8, the existence of a minimizer of the objective function is clear.

3. Convexity and differentials of the objective function

3.1. Convexity of the objective function

The convexity of the objective function will be investigated first. Then its differentiability will be analyzed according to a partition of the set of power supply vectors. Readers interested in exploring optimization and convexity may study [31–34].

Lemma 3.1. *The function $f_{as,k} : P^+ \to \mathbb{R}_+, \forall\, k \in \mathbb{Z}_{n_E}$, is convex on P^+.*

The proof is by the formula of the Hessian matrix of $f_{as,k}$, see details in [30, Lemma 4.1].

Claim 3.2. *In general, each component of the standard deviation of the phase-angle difference, denoted as $\sigma_k : P^+ \to \mathbb{R}_{s,+}$, $\forall\ k \in \mathbb{Z}_{n_E}$, is not convex.*

Remark 3.3. Note that whether σ_k is convex or not cannot be conclusively determined, as it is an implicit function of p_s. To verify Claim 3.2, an algorithm to compute the gradient vector and the Hessian matrix of the function σ_k numerically by using the directional derivative method is developed. The algorithm with its analysis can be found in Section 4.2.

In our computational analysis, several cases have been examined, and the standard deviation $\sigma_k, \forall k \in \mathbb{Z}_{n_E}$ is not convex in all cases, which is due to the presence of both strictly positive and strictly negative eigenvalues in the Hessian matrix.

Consequently, the objective function f on P^+ cannot be guaranteed to be convex, which means it may have multiple local minimizers. To ensure that each local minimizer of the objective function is isolated, the following assumption is introduced.

Assumption 3.4. There does not exist an open neighborhood $\mathcal{O} \subset P^+$ for which the objective function f remains constant on this neighborhood.

3.2. A partition of the set of power supply vectors

Since the objective function is not differentiable, we turn to investigate its local differentiability. Hence, a partition of the set of power supply vectors defined in Definition 2.1 is needed. The definition of the partition is preceded by concepts of algebraic geometry.

In geometry, one describes surfaces. Ways to specify surfaces include hyperplanes described by affine functions; algebraic sets described by polynomials; and other surfaces described by functions which are not polynomials. For power systems, all three cases appear. The terminology of a surface described by a polynomial follows, [35, Chapter 1, Paragraph 2, Definition 1].

Consider a polynomial in $n \in \mathbb{Z}_+$ indeterminates of degree $d \in \mathbb{Z}_+$ as an algebraic object

$\forall\ n \in \mathbb{Z}_+,\ \forall\ d \in \mathbb{Z}_+,\ \forall\ k \in \mathbb{N}_d^n\ \left(\forall\ i \in \mathbb{Z}_n, k(i) \in \mathbb{N}_d = \{0, 1, \dots, d\}\right),$

define the monomial, $p^k = \prod_{i=1}^{n} p_i^{k(i)}$;

define the polynomial for any finite subset $\mathbb{N}_s \subset \mathbb{N}_d^n$,

$q(p) = \sum_{k \in \mathbb{N}_s} c(k)\ p^k \in \mathbb{R}[p_1,\ p_2,\ \dots,\ p_n]$, where, $\forall\ k \in \mathbb{N}_s,\ c(k) \in \mathbb{R}$;

$\forall\ Q_s \subset \mathbb{R}[p_1,\ p_2,\ \dots,\ p_n]$, a finite subset of polynomials, define,

$V(\mathbb{R}^n,\ Q_s) = \left\{p \in \mathbb{R}^n |\ \forall\ q \in Q_s,\ q(p) = 0\right\}.$

Call $V(\mathbb{R}^n,\ Q_s)$ an *algebraic set* or an *affine variety*. Call a subset $G \subseteq \mathbb{R}^n$ a *generic subset* of $\mathbb{R}^n$ if $\mathbb{R}^n \backslash G$ is an algebraic set.

This is now applied to the objective function. Define, $f_{lin,k}\left(p_s\right) = |A(k)\ p_s + b_k|,\ \forall\ k \in \mathbb{Z}_{n_E}$. Recall the notation, $A(k)$ denotes the kth row of matrix A and $A(k)\,p_s + b_k$ is an affine function of p_s, thus, a polynomial, hence $V(\mathbb{R}^{n^+-1},\ f_{as,k})$ is an algebraic set. Note that, for $x \in (-1,\ +1)$, $\arcsin(x) = 0$ if and only if $x = 0$. However, $f_k\left(p_s\right) = \left[\arcsin\left(|A(k)\,p_s + b_k|\right)\right] + r \cdot \sigma_k\left(p_s\right)$ defined in Definition 2.3 is not a polynomial because of the function arcsin. It can be proven that $\sigma_k(p_s)$ is a polynomial in terms of the components of p_s.

Definition 3.5. Consider the objective function of Definition 2.3. Distinguish the cases:

- Case 1. there exists a unique $k \in \mathbb{Z}_{n_E}$ and there exists a nonempty subset $P^+_{(k)}$ such that $P^+_{(k)} = \{p_s \in P^+ |\ f\left(p_s\right) = f_k\left(p_s\right)\}$.
- Case 1.1. there exists a unique $k \in \mathbb{Z}_{n_E}$ and there exists a nonempty subset $P^+_{(k),nz}$ such that $P^+_{(k),nz} = \{p_s \in P^+_{(k)} |\ f_{as,k}\left(p_s\right) \neq 0\}$.
- Case 1.2. there exists a unique $k \in \mathbb{Z}_{n_E}$ and there exists a nonempty subset $P^+_{(k),z}$ such that $P^+_{(k),z} = \{p_s \in P^+_{(k)} |\ f_{as,k}\left(p_s\right) = 0\}$.
- Case 2. there exist two or more $k_1,\ k_2,\ \dots,\ k_m \in \mathbb{Z}_{n_E}$ and there exists a nonempty subset $P^+_{(k_1,k_2,\dots,k_m)}$ such that $P^+_{(k_1,k_2,\dots,k_m)} = \{p_s \in P^+ |\ f\left(p_s\right) = f_{k_1}\left(p_s\right) = f_{k_2}\left(p_s\right) = \cdots = f_{k_m}\left(p_s\right)\}$.
  Denote by $I_{max}\left(p_s\right) = \{k_1,\ k_2,\ \dots,\ k_m\}$ the subset of those integers.

The subset $P^+_{(k_1,k_2)}$ may not be an algebraic set because the relation $f_{k_1}\left(p_s\right) = f_{k_2}\left(p_s\right) = \left[\arcsin\left(|A(k_2)\,p_s + b_{k_2}|\right)\right] + r \cdot \sigma_{k_2}\left(p_s\right)$ is not a polynomial in general. It is clear from the formulas that Case 1.1 is the generic case and that Case 1.2 takes place on an algebraic set.

In the generic case (Case 1.1), the objective function f is differentiable, while in Case 1.2, f is subdifferentiable due to the absolute value operator which makes the gradient vector of f at this domain different on two sides, and in Case 2, f is nondifferentiable because a gradient vector does not exist. Directional derivatives of the objective function will be discussed in the following subsections.

3.3. The first directional derivative of the objective function

Books about the directional derivatives and related concepts include [15,32,34]. The formulas of the first directional derivative of the objective function will be used in the design of the asymptotic algorithm to approximate a local minimizer, from which one can also get a different understanding of the Jacobian matrix defined below from the literature.

Definition 3.6 (*[34, Definition 3.1.3]*). Consider integers $m,\ n \in \mathbb{Z}_+$, a convex and open subset $U \subseteq X = \mathbb{R}^n$, and a function $g : U \to \mathbb{R}^m$. Assume that: (1) the function g is continuous on its domain of definition U; and (2) there does not exist an open subset $O \subseteq U$ on which the function g is constant.

One says that the function g is *directionally differentiable at* $x_s \in U$ *in the direction* $v \in \mathbb{R}^n$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that the following limit exists,

$$L(v) = \lim_{t \in \mathbb{R}_{s+},\ t \downarrow 0} \frac{g\left(x_s + t \cdot v\right) - g\left(x_s\right)}{t},\ \Leftrightarrow$$

$$0 = \lim_{t \in \mathbb{R}_{s+},\ t \downarrow 0} \frac{g\left(x_s + t \cdot v\right) - g\left(x_s\right) - t \cdot L(v)}{t};\ \text{denote,} \tag{12}$$

$$dg\left(x_s,\ v\right) = L(v) \in \mathbb{R}^m,\ \forall\ v \in \mathbb{R}^n. \tag{13}$$

Call then $dg\left(x_s,\ v\right)$ the *directional differential of* g *at* x_s *in the direction* v.

If a basis of the vector space $\mathbb{R}^n$ has been chosen to be $\left\{e_1,\ e_2,\ \dots,\ e_n\right\}$, the set of the Euclidian unit vectors, then the linear map can be represented by the Jacobian matrix,

$$J_g\left(x_s\right) = g'\left(x_s\right) = \left[L\left(e_1\right) \quad L\left(e_2\right) \quad \dots \quad L\left(e_n\right)\right] \in \mathbb{R}^{m \times n};\ \text{equivalently,}$$

$$J_g\left(x_s\right)_{i,j} = \lim_{t \in \mathbb{R}_{s+},\ t \downarrow 0} \frac{g_i\left(x_s + t \cdot e_j\right) - g_i\left(x_s\right)}{t},\ \forall\ i \in \mathbb{Z}_m,\ \forall\ j \in \mathbb{Z}_n;\ \text{then}$$

$$dg\left(x_s,\ v\right) = L(v) = J_g\left(x_s\right)\ v = g'\left(x_s\right)\ v,\ \forall\ v \in \mathbb{R}^n.$$

Call the matrix $J_g\left(x_s\right)$ the *Jacobian matrix* of g at $x_s \in U$.

In general, for Case 1.2 and Case 2, the Jacobian matrix does not exist for the objective function of Definition 2.3. Even if the Jacobian matrix for a particular direction exists then the Jacobian matrix can be different for another direction. For Case 2, the directional differential $df\left(p_s,\ v\right)$ will not be a continuous function of the direction vector v. A further formalization of a sector-wise directional derivative, is not stated in this paper.

Proposition 3.7. *Consider the objective function of Definition 2.3, and the partition of the feasible set in Definition 3.5.*

(a) *Case 1.1. For any $p_s \in P^+_{(k),nz}$, and a direction vector $v \in \mathbb{R}^{n^+-1}$, such that $p_s + v \in P^+$, the first directional derivative of the objective function exists and equals,*

$$f'\left(p_s,\ v\right) = f'_k\left(p_s, v\right) = \nabla f\left(p_s\right) v = \left(\nabla f_{as,k}\left(p_s\right) + r \cdot \nabla\sigma_k\left(p_s\right)\right) v$$
$$= \begin{cases} \left[\left(1-\left(A(k)\ p_s + b_k\right)^2\right)^{-\frac{1}{2}} \cdot A(k) + r \cdot \nabla\sigma_k\left(p_s\right)\right] v, \\ \quad \text{if } A(k)\ p_s + b_k > 0, \\ \left[-\left(1-\left(A(k)\, p_s + b_k\right)^2\right)^{-\frac{1}{2}} \cdot A(k) + r \cdot \nabla\sigma_k\left(p_s\right)\right] v, \\ \quad \text{if } A(k)\ p_s + b_k < 0. \end{cases}$$

(b) *Case 1.2. For any $p_s \in P^+_{(k),z}$, and a direction vector $v \in \mathbb{R}^{n^+-1}$, such that $p_s + v \in P^+$, $f'_k\left(p_s\right) = |A(k)\ v| + r \cdot \nabla\sigma_k\left(p_s\right)\ v$.*
Note that in this case, the Jacobian matrix depends on the direction vector v.

(c) *[15, Chapter 4, p. 157]. Case 2. For any $p_s \in P^+_{(k_1,k_2,\ldots,k_m)}$, and a direction vector $v \in \mathbb{R}^{n^+-1}$, such that $p_s + v \in P^+$, $f'\left(p_s,\ v\right) = \max_{k\in I_{max}(p_s)} f'_k\left(p_s,\ v\right)$. Note that $f'_k\left(p_s, v\right)$ takes the expression of Case 1.1 or Case 1.2, depending on whether $A(k)\, p_s + b_k$ is zero or not.*

See details of the proof in [30, Proof 8]. For Case 2 and $m = 2$, one can distinguish 9 subcases. Most subcases have two or four different expressions for the Jacobian matrices. Because Case 2 will occur less frequently than Case 1.2, the authors have decided not to include the formulas of the Jacobian matrices for each subcase.

Note that the second directional derivative and the corresponding property of the objective function are also derived; see [30, Section 4.4], which can be used for higher accuracy algorithms in future work.

4. Algorithms, the descent method, and convergence theorem

Many algorithms have been developed for nonconvex and nonsmooth optimization. T. Liu et al. introduced the successive difference-of-convex approximation method, demonstrating its global convergence for a specific class of objective functions in [36]. In a similar vein, Y. Wang proposed the ADMM algorithm for nonconvex nonsmooth optimization, which also exhibits global convergence, particularly for objective functions with equality constraints, as discussed in [37]. Another noteworthy contribution comes from B. Wen et al. who presented the proximal gradient algorithm with extrapolation and conducted a comprehensive convergence analysis, referred to as R-linear convergence in [38]. In this section, a two-step algorithm with the mathematical analysis will be provided for our objective function, which is also nonconvex and nonsmooth.

This section is organized as follows. The computation of the gradient vector and Hessian matrix of the implicit functions V_i and σ_i, $i \in \mathbb{Z}_{n_E}$ defined in Procedure 2.4, Step 3 are analyzed in Sections 4.1 and 4.2, respectively. Section 4.3, and Section 4.4 introduce an algorithm using the steepest descent method for approximating a local minimizer and the convergence analysis theorem. The computational complexity of the algorithms proposed in the previous subsections is analyzed in Section 4.5, which concerns whether they can be executed online or not, and in which an algorithm is proposed in Proposition 4.11 to compute an approximate effective vector to serve as the initial vector for the algorithm of Section 4.3.

4.1. The analysis and computation of the gradient vector and the Hessian matrix of a variance V_i, for a power supply vector $p_s \in P^+$

Since the variance V_i does not have an explicit formula, the directional derivative method inspired by [39, Theorem 6] will be used to compute the gradient vector and the Hessian matrix of that function. The formal analysis and derivation are presented in this subsection, while some concrete formulas which are less important are omitted and referred to [30].

Proposition 4.1. *The first and the second directional derivatives of the implicit function $V_i : P^+ \to \mathbb{R}_+$, $i \in \mathbb{Z}_{n_E}$, of a vector $p_s \in P^+$, can be computed according to the formulas,*

$$V_i\left(p_s + \delta\,\mu\right) = V_i^{(0)}\left(p_s\right) + V_i^{(1)}\left(p_s\right)\delta + V_i^{(2)}\left(p_s\right)\delta^2,\ \delta \in \mathbb{R},\ \mu \in \mathbb{R}^{n^+-1}, \textit{ where,}$$
$$V_i^{(1)}\left(p_s\right) = \nabla_\mu V_i\left(p_s\right) = \mu^\top \nabla V_i\left(p_s\right), \quad V_i^{(2)}\left(p_s\right) = \frac{1}{2}\,\mu^\top \nabla^2\, V_i\left(p_s\right)\ \mu.$$

Using this proposition, specifying $k,\ j \in \mathbb{Z}_{n^+-1}$, one can calculate $\nabla V_i\left(p_s\right)(k)$ by $\mu = e_k$ and $\frac{1}{2} \cdot \left(\nabla^2 V_i\left(p_s\right)(k,k) + \nabla^2 V_i\left(p_s\right)(j,j) + \nabla^2 V_i\left(p_s\right)(k,j) + + \nabla^2 V_i\left(p_s\right)(j,k)\right)$ by $\mu = e_k + e_j$.

However, after perturbing the variable from p_s to $p_s + \delta\,\mu$, one needs to differentiate each of the matrices in Eq. (11) for δ. The weight matrix function W, as defined in Definition 2.5, depends on the variable p_s and is part of the system matrix J_d in Eq. (10). Therefore, it should be investigated first. View the matrix $W(p_s + \delta\mu)$ as a matrix function with respect to δ and define $h(\delta) = W(p_s + \delta\mu)$. Its Taylor expansion equals,

$$h(\delta) = W\left(p_s + \delta\,\mu\right) = W^{(0)} + W^{(1)}\,\delta + W^{(2)}\,\delta^2, \text{ where,}$$
$$W^{(0)} = W\left(p_s\right),\ W^{(1)} = \frac{dh(\delta)}{d\delta}|_{\delta=0};\ W^{(2)} = \frac{1}{2}\,\frac{d^2h(\delta)}{d\delta^2}|_{\delta=0}.$$

In the formulas above, the matrices $W^{(1)}$ and $W^{(2)}$ need to be computed from matrices that are already available. First, we define a matrix E as $E = [U_1 - U_{n^+}, \ldots, U_{n^+-1} - U_{n^+}, U_{n^++1} - U_{n^+}, \cdots, U_{n_E} - U_{n^+}] \in \mathbb{R}^{n_V\times(n_V-1)}$, concerning the orthonormal matrix U as defined in Procedure 2.4 Step 1. Next, recall the notation $E_i \in \mathbb{R}^{n_V}$, representing the ith column of the matrix E. The formulas for the matrices $W^{(1)}$ and $W^{(2)}$ are displayed in [30, Appendix A, Eq. (19)].

Secondly, the Taylor expansions of a variance $V_i\left(p_s + \delta\,\mu\right)$, the ith row of output matrix $C_d(\delta)$ denoted as $C_{d,i}(\delta)$, and the solution of the Lyapunov equation $Q(\delta)$ with respect to δ are presented in [30, Appendix A, Eq. (20)]. Following Proposition 4.1, formulas for the first directional derivative $\nabla_\mu V_i\left(p_s\right)$ and the second directional derivative $\nabla^2_\mu\ V_i\left(p_s\right)$ are provided in [30, Appendix A, Eq. (21)]. Within [30, Appendix A, Eq. (21)], the unknown matrices or vectors are $C^{(0)}_{d,i}, C^{(1)}_{d,i}, C^{(2)}_{d,i}, Q^{(0)}, Q^{(1)}, Q^{(2)}$. Before calculating these matrices, the transformation matrix U as a function of δ, which satisfies $U(\delta)^\top M^{-\frac{1}{2}} B W(p_s+\delta\mu) B^\top M^{-\frac{1}{2}} U(\delta) = \Lambda_1$, must be analyzed. Here, the matrix Λ_1 is diagonal, and the first diagonal element of the matrix Λ_1 should be 0. If it is not zero, the columns of the matrix $U(\delta)$ need to be permuted. The Taylor expansion of the transformation matrix function U concerning δ is described as similar to the function h,

Remark 4.2. This remark pertains to the matrices $U^{(1)}$ and $U^{(2)}$, which are the first and second terms of the Taylor expansion of U concerning δ. We define a function, denoted as L, with respect to δ as $L(\delta) = M^{-\frac{1}{2}}\ B\,W(p_s + \delta\,\mu)\,B^\top\,M^{-\frac{1}{2}} : \mathbb{R} \to \mathbb{R}^{n_V\times n_V}$.

If the dimension $n^+ - 1$ of the matrix $L(\delta)$ is relatively small, say no more than 4, then one may be able to get the analytical solution of the orthonormal matrix $U(\delta)$. The values of the matrices $U^{(0)}, U^{(1)}, U^{(2)}$ then follow.

If the dimension $n^+ - 1$ of the matrix $L(\delta)$ is large, it follows from Galois' theorem that it is impossible to get the analytical expression of the orthonormal matrix $U(\delta)$. Therefore, a method to approximate the numerical value of the matrices $U^{(0)}, U^{(1)}, U^{(2)}$ needs to be formulated. It is easy to compute the numerical value of matrix $U(0)$ by the formula $U(0)^\top L(0)\,U(0) = \Lambda$, where Λ is a diagonal matrix, as the matrix $L(0)$ is already known. The authors then choose the finite difference method to approximate the first and second-order derivative matrices $U^{(1)}, U^{(2)}$. The errors between the numerical values and the real values are dependent on the step length denoted by Δ. For a step length, the reader may think of $10^{(-4)}$ as an example, we have the following approximations:

$$U^{(1)} \approx \frac{U(\Delta) - U(-\Delta)}{2\,\Delta},\ U^{(2)} \approx \frac{U(\Delta) - 2\,U(0) + U(-\Delta)}{2\,\Delta^2},$$

where $U(\Delta)$, $U(-\Delta)$ can be approximated similarly to $U(0)$.

With respect to this remark, the numerical values of the output and the input matrices $C_d^{(0)}$, $C_d^{(1)}$, $C_d^{(2)}$, $K_d^{(0)}$, $K_d^{(1)}$, $K_d^{(2)}$ can be computed by the formulas of [30, Appendix A, Eq. (22)], and the system matrices $J_d^{(0)}$, $J_d^{(1)}$, $J_d^{(2)}$ can be computed by [30, Appendix A, Eq. (23)]. These matrices will be used for the calculation of the solution matrices $Q^{(0)}$, $Q^{(1)}$, $Q^{(2)}$ of the Lyapunov equations.

Finally, the Taylor expansion of the Lyapunov equation, $Q(\delta)\, A_d(\delta)^\top + A_d(\delta)\, Q(\delta) + B_d(\delta)\, B_d(\delta)^\top = 0$ is presented in [30, Appendix A, Eq. (24)]. Solve successively the following three Lyapunov equations for the matrices $Q^{(0)}, Q^{(1)}, Q^{(2)}$,

$$0 = Q^{(0)}\, J_d^{(0)\top} + J_d^{(0)}\, Q^{(0)} + K_d^{(0)}\, K_d^{(0)\top}, \tag{14}$$

$$0 = Q^{(1)}\, J_d^{(0)\top} + J_d^{(0)}\, Q^{(1)} + \Big[Q^{(0)}\, J_d^{(1)\top} + J_d^{(1)}\, Q^{(0)} + K_d^{(0)}\, K_d^{(1)\top} + \tag{15}$$

$$+K_d^{(1)}\, K_d^{(0)\top}\Big], \tag{16}$$

$$0 = Q^{(2)}\, J_d^{(0)\top} + J_d^{(0)}\, Q^{(2)} + \Big[Q^{(0)}\, J_d^{(2)\top} + J_d^{(2)}\, Q^{(0)} + Q^{(1)}\, J_d^{(1)\top} + J_d^{(1)}\, Q^{(1)} + \tag{17}$$

$$+K_d^{(0)}\, K_d^{(2)\top} + K_d^{(1)}\, K_d^{(1)\top} + K_d^{(2)}\, K_d^{(0)\top}\Big].$$

Since the system matrix $J_d^{(0)}$ is Hurwitz, each of the above three Lyapunov equations has a unique positive definite solution.

4.2. The algorithm of computation of the gradient vector and the Hessian matrix of a standard deviation σ_i for a power supply vector $p_s \in P^+$

Section 4.1 analyzes the computation method for the gradient vector and the Hessian matrix of a variance V_i concerning a power supply vector $p_s \in P^+$. However, needed is that of a standard deviation σ_i for a vector $p_s \in P^+$ denoted as $\nabla\sigma_i(p_s)$, $\nabla^2\sigma_i(p_s)$, respectively, the formulas that should be employed are as follows,

$$\nabla\sigma_i^2 = 2\,\sigma_i\,\nabla\sigma_i \;\Rightarrow\; \nabla\sigma_i = \frac{\nabla\sigma_i^2}{2\,\sigma_i} = \frac{\nabla V_i}{(2\,\sigma_i)},$$

$$\nabla^2\sigma_i^2 = 2\,\nabla\sigma_i^\top\,\nabla\sigma_i + 2\,\sigma_i\,\nabla^2\sigma_i$$

$$\Rightarrow\; \nabla^2\sigma_i = \frac{\nabla^2\sigma_i^2 - 2\,\nabla\sigma_i^\top\,\nabla\sigma_i}{2\,\sigma_i} = \frac{H(V_i) - 2\,\nabla\sigma_i^\top\,\nabla\sigma_i}{2\,\sigma_i}.$$

Define $\{G = \nabla V_i,\ H = \nabla^2 V_i; G1 = \nabla\sigma_i,\ H1 = \nabla^2\sigma_i\}$. Our analysis leads to Algorithm 1. With this algorithm, for a given power supply vector p_s, the numerical values of $\nabla\sigma_i(p_s)$ and $\nabla^2\sigma_i(p_s)$ would be available, which will be directly used in the subsequent analysis and algorithm design.

Algorithm 1 Computation of $\nabla\sigma_i(p_s), \nabla^2\sigma_i(p_s)$.

1: **Input data:** A power supply vector $p_s \in P^+$
2: **Output data:** G_1, H_1. ▷ The gradient vector and the Hessian matrix
3: U, $\Lambda^\dagger$ ← Procedure 2.4 Step 1; E ← Section 4.1; $W^{(0)} = W(p_s)$ ← Definition 2.5. ▷ The prerequisite matrices
4: **for** $k = 1, 2, \cdots, n^+ - 1$ **do**
5: $W^{(1)}$, $W^{(2)}$, ← [30, Appendix A, Eqn. (19)]. ▷ The weight matrices
6: $U(0)$, $U^{(1)}$, $U^{(2)}$, ← (Remark 4.2) ▷ The transformation matrices
7: $J_d^{(0)}$, $J_d^{(1)}$, $J_d^{(2)}$, $K_d^{(0)}$, $K_d^{(1)}$, $K_d^{(2)}$, $C_d^{(0)}$, $C_d^{(1)}$, $C_d^{(2)}$, ← [30, Appendix A, Eqns. (22, 23)] ▷ The system, input and output matrices
8: $Q^{(0)}$, $Q^{(1)}$, $Q^{(2)}$ ← Eq. (14).
9: $G(k)$, $H(k,k)$ ← [30, Appendix A, Eqn. (21)]
10: **end for**
11: **for** $k = 1, 2, \cdots, n^+ - 2$ **do**
12: **for** $j = k+1, k+2, \cdots, n^+ - 1$ **do**
13: Repeat the steps 5-8 of the former loop.
14: $2\,H(k,j)$ ← [30, Appendix A, Eqn. (21)]$-H(k,k) - H(j,j)$.
15: **end for**
16: **end for**
17: $G_1 \leftarrow \frac{G}{2\,\sigma_i}$, $H_1 \leftarrow \frac{H - 2\,G_1^\top\,G_1}{2\,\sigma_i}$.

4.3. A steepest descent method to approximate a local minimizer of the objective function

The reader will find in this section: A description of the search for the steepest descent direction, primarily for Case 1.1, with comments provided for Cases 1.2 and 2. A description of the algorithm using the projected subgradient method to compute the steepest descent direction. A description of the line search along the steepest descent direction.

Algorithm 2 A steepest descent method.

1: **Input data:** $p_s^{(0)} \in \text{interior}(P^+)$ ▷ $p_s^{(0)}$ is an interior point.
2: **Output data:** $p_s^{(i+1)}$ ▷ An approximation of a local minimizer.
3: **while** $f\left(p_s^{(i)}\right) - f\left(p_s^{(i+1)}\right) > \epsilon$ **do**
4: Compute $v^{(i)}$ by solving $f'\left(p_s^{(i)}, v^{(i)}\right) = \min_{b_1 \le A_1\left(p_s^{(i)} + v\right),\ p_s^{(i)} + v \le b_2} f'\left(p_s^{(i)}, v\right)$, $i \in \mathbb{N}$. ▷ Apply Algorithm 3, with A_1, b_1, and b_2 as presented in Eq. (4). (Note that the steepest descent direction $v^{(i)}$ computed by Algorithm 3 is a numerical approximation of the exact value. However, in the subsequent steps of the algorithm, it is treated as the exact one.)
5: $t = 1$, $0 < \alpha < 0.5$, $0 < \beta < 1$. ▷ Set the parameters
6: **while** $f\left(p_s^{(i)} + t \cdot v^{(i)}\right) > f\left(p_s^{(i)}\right) + \alpha \cdot t \cdot f'(p_s^{(i)}, v^{(i)})$ **do**
7: $t = \beta \cdot t$. ▷ Line search
8: **end while**
9: $p_s^{(i+1)} \leftarrow p_s^{(i)} + t \cdot v^{(i)}$. ▷ Update the variable **return** to Step 4.
10: Determine whether to break or continue the loop based on the following cases,

- Case b.1: If $f'\left(p_s^{(i)}, v^{(i)}\right) = 0$ and $v^{(i)} = 0$, then $p_s^{(i)}$ is a local minimizer and breaks the loop.
- Case b.2: If $f'\left(p_s^{(i)}, v^{(i)}\right) = 0$ and $v^{(i)} \neq 0$, then for $\xi \in (0,\ 1)$ according to Definition 4.4, compute $f''\left(p_s^{(i)} + \xi \cdot v^{(i)}, v^{(i)}\right)$. (1) If $f''\left(p_s^{(i)} + \xi \cdot v^{(i)}, v^{(i)}\right) > 0$, then $p_s^{(i)}$ is a local minimizer, and breaks the loop. (2) Otherwise, if $f''\left(p_s^{(i)} + \xi \cdot v^{(i)}, v^{(i)}\right) < 0$, then $p_s^{(i)}$ is an inflection point. In this case, let $p_s^{(i+1)} = p_s^{(i)} + v^{(i)}$ and return to Step 3.
- Case b.3: If $f'(p_s^{(i)}, v^{(i)}) < 0$, continue with Step 9.

11: **end while**

Algorithm 3 A projected subgradient method.

1: **Input data:** $p_s^{(i)} \in P^+$ ▷ $p_s^{(i)}$ is the ith iteration step of the power supply vector.
2: **Output data:** $v^{(i)}$ ▷ the steepest descent direction.
3: **Denote** $f_d(v) = f'\left(p_s^{(i)}, v\right) : \mathbb{R}^{n^+-1} \to \mathbb{R}$, $\alpha_j = 1/j^\gamma$, $0 < \gamma < 1$, [40] .
4: $v_0 \leftarrow b_1 \le A_1(p_s^{(i)} + v_0)$, $p_s^{(i)} + v_0 \le b_2$. ▷ Initialization of v
5: **while** $j < L_N$, e.g. $L_N = 500$ **do**
6: $v_{j+1} \leftarrow P\left(p_s^{(i)} + v_j - \alpha_j\, g_j\right) - p_s^{(i)}$. ▷ g_j : subgradient of $f'\left(p_s^{(i)}, v\right)$ at v_j; $P : \mathbb{R}^{n^+-1} \to P^+$: Euclidean projection operator.
7: $f_{d,min}^{j+1} \leftarrow \min\left\{f_{d,min}^j,\ f_d\left(v_{j+1}\right)\right\} \Rightarrow f_{d,min}^{j+1} = \min\left\{f_d\left(v_1\right), \cdots, f_d\left(v_{j+1}\right)\right\}$, and return to the step 6.
8: **end while**

Algorithm 2 to approximate a local minimizer of the objective function is based on the steepest descent method, [33, Section 9.4]. The concept of a *subgradient* is defined that is used in the computation of the steepest descent direction.

Definition 4.3. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function. If at $x \in \mathbb{R}^n$, a vector $h \in \mathbb{R}^n$ satisfies that $g(y) \ge g(x) + h^\top(y - x), \forall\ y \in \mathbb{R}^n$, then the *subgradient of* g *at* x exists, and it is equal to the vector h, which is denoted by $\partial g(x) = h$.

The following definition will be used in the evaluation of an inflection point.

Definition 4.4. Define a function $g : \mathbb{R}_{s+} \to \mathbb{R}$ with respect to τ, by $g(\tau) = f(p_s + \tau v)$. It follows from Definition 3.6 that, $g(\tau) = g(0) + g'(0^+)\tau + g''(0^+)\tau^2/2 + o(\tau^2)$ or, $g(\tau) = g(0) + g'(0^+)\tau + g''(\xi)\tau^2/2,\ \xi \in (0,\tau)$, where,
$g'(0^+) = f'(p_s, v)$ and $g''(0^+) = f''(p_s, v)$.

The analysis of the steepest descent direction in Case 1.1

For a given vector $p_s^{(i)}$, where $p_s^{(i)} \in P^+_{(k),nz}$ and $p_s^{(i)} \in \text{intrior}(P^+)$, with $i \in \mathbb{N}$ representing the iteration step in Algorithm 2, the steepest descent direction is, $-\nabla f\left(p_s^{(i)}\right) = -\left(\nabla f_{as,k}\left(p_s^{(i)}\right) + r \cdot \nabla\sigma_k\left(p_s^{(i)}\right)\right)$. In this case, there exists a unique positive value $t_i^* > 0$ such that $p_s^{(i)} - t \cdot \nabla f\left(p_s^{(i)}\right) \notin P^+$ for $t > t_i^*$, and $p_s^{(i)} - t \cdot \nabla f\left(p_s^{(i)}\right) \in P^+$ for $t \le t_i^*$, which means that the vector $p_s^{(i)} - t_i^* \cdot \nabla f\left(p_s^{(i)}\right)$ lies on one of the facets of the polytope P^+. Furthermore, the value of t_i^* can be easily computed using the constraints of the domain P^+ defined in Definition 2.1. In conclusion, the steepest descent direction v under constraints in this case equals: $-t_i^* \cdot \left(\nabla f_{as,k}\left(p_s^{(i)}\right) + r \cdot \nabla\sigma_k\left(p_s^{(i)}\right)\right)$.

For a vector $p_s^{(i)} \in P^+_{(k),nz}$ and $p_s^{(i)}$ lies on the boundary of P^+, denoted as $p_s^{(i)} \in \partial P^+$, we encounter a scenario where the direction vector of the objective function at the vector $p_s^{(i)}$ must point into the interior of the polytope P^+. For the vector $p_s^{(i)}$, which is at the corner of the polytope P^+, the direction vector of the objective function is particularly challenging to express by using mathematical formulas. Therefore, to handle these extreme cases, a projected subgradient method, as presented in Algorithm 3 will be applied, which was first proposed in [40]. Moreover, at Step 2 of Algorithm 3 in Case 1.1, the subgradient satisfies, $g_j = \nabla f_{as,k}\left(p_s^{(i)}\right) + r \cdot \nabla\sigma_k\left(p_s^{(i)}\right)$, $\forall\, j \in \mathbb{N}$ in this case, and the more concrete expression of g_j can be found in Proposition 3.7(a). The following assumption and remark are used in Algorithm 3,

Assumption 4.5. In Algorithm 3, the minimizer satisfies $f'(p_s^*, v^*) \le 0$ due to the origin,. In the case where $f'\left(p_s^*, v^*\right) = 0$, it is assumed that there exists at most one segment that satisfies this condition. The reason is that if there exist two or more segments that satisfy this equality, then the convex combination of these line segments will also satisfy $f'\left(p_s^*, v\right) = 0$, and this case is excluded in this paper.

Remark 4.6. By solving a linear constrained quadratic program $\min_x \frac{1}{2}\|x-y\|_2^2, b_1 \le A_1\, y, y \le b_2$, the Euclidean projection of step 5 in Algorithm 3 can be programmed as an operator. In this paper, this program is solved through a dynamical system that is exponentially convergent, [41]. Note that one can also reformulate the problem as conic programming, which has very stable solvers. Additionally, readers can explore other methods like the Charged Balls Method [42, Chapter 1] for projecting a vector onto a convex set; see a review of projection algorithms in [43].

Proposition 4.7. *It has been established in [40] that the approximated sequence produced by Algorithm 3 satisfies $\lim_{j\to\infty} f_{d,min}^{j+1} = f_d^* \in \mathbb{R}$. If the iteration number j is sufficiently large, for instance, when $j > LN$, the iteration can be stopped and a v_j such that $f_d\left(v_j\right) = f_d^*$ can be specified. Note that $v^* = v_j$ may not be a unique solution for $f_d\left(v_j\right) = f_d^*$, and according to Assumption 4.5, if $f_d^* = 0$, then $v^* = 0$ or v^* lies on a line segment.*

The analysis of the steepest descent direction in the Cases 1.2 and 2

Algorithm 3 will be employed to compute the steepest descent direction in Case 1.2. Consider an iteration step v_j, the expression of the subgradient of the directional derivative function $f'\left(p_s^{(i)}, v\right)$ at v_j, denoted as g_j in Step 2 is needed. It follows Proposition 3.7, $f'\left(p_s^{(i)}, v_j\right) = f_k'\left(p_s^{(i)}, v_j\right) = |A(k)\ v_j| + r \cdot \nabla\sigma_k\left(p_s^{(i)}\right) v_j$, and thus the subgradient g_j satisfies,

$$g_j = \begin{cases} \left(\theta \cdot A(k) + (1-\theta)\cdot(-A(k)) + r\cdot\nabla\sigma_k\left(p_s^{(i)}\right)\right)^\top, & \text{if } A(k)\ v_j = 0; \\ \left(A(k) + r\cdot\nabla\sigma_k\left(p_s^{(i)}\right)\right)^\top, & \text{if } A(k)\ v_j > 0; \\ \left(-A(k) + r\cdot\nabla\sigma_k\left(p_s^{(i)}\right)\right)^\top, & \text{if } A(k)\ v_j < 0. \end{cases} \tag{18}$$

Consider Case 2 and an iteration step v_j in Algorithm 3. As indicated in Proposition 3.7, we have $f'\left(p_s^{(i)}, v_j\right) = \max_{k\in I_{max}\left(p_s^{(i)}\right)} f_k'\left(p_s^{(i)}, v_j\right)$.

To determine the subgradient g_j of the function $f'\left(p_s^{(i)}, v\right)$ at v_j, one needs to specify a $k \in I_{max}\left(p_s^{(i)}\right)$ such that $f'\left(p_s^{(i)}, v_j\right) = f_k'\left(p_s^{(i)}, v_j\right)$. Then the subgradient g_j can be determined by Eq. (18) or Case 1.1.

Proposition 4.8. *The line search step (step 6 - step 8) of Algorithm 2 can be executed in a finite number of steps.*

The proof of Proposition 4.8 is similar to [33, Chapter 9.4], and thus omitted here. The readers can also find it in [30, Proof 10]. To summarize, the algorithms proposed will be invoked as follows, which formulates the complete steepest descent algorithm.

Procedure 4.9.

1. *Input data, the power supply vector $p_s^{(0)}$ which satisfies the constraints in Definition 2.1.*
2. *While the stopping criterion of Algorithm 2 is not satisfied, then do*
   (a) *Compute the gradient $\nabla\sigma(p_s^{(k)})$ by Algorithm 1.*
   (b) *Invoke Algorithm 3 to compute the steepest descent direction $v^{(k)}$.*
   (c) *Employ Algorithm 2 to get the next iteration step $p_s^{(k+1)}$.*
3. *Output data, the optimal power dispatch $p_s^{(k+1)}$ which is a local minimizer of the objective function.*

4.4. Convergence analysis

An approximation sequence will be generated by the steps described in Procedure 4.9. The convergence analysis theorem is introduced as follows,

Theorem 4.10.

(a) *Algorithm 2 terminates after a finite number of steps.*
(b) *Algorithm 2 generates a sequence $\left\{p_s^{(k)} \in P^+, \forall\, k \in \mathbb{N}\right\}$ such that $\lim_{k\to\infty} f\left(p_s^{(k)}\right) = f^*$, where f^* is a local minimum.*

Proof. The claim (a) follows directly from Case b.1 and Case b.2 (1) of Algorithm 2.

For the claim (b), if Algorithm 2 does not break the loop according to Case 5.1 and Case 5.2 (1), then it generates a sequence of values, $\left\{f\left(p_s^{(0)}\right), f\left(p_s^{(1)}\right), \ldots \mid p_s^{(k)} \in P^+, \forall\, k \in \mathbb{N}\right\}$, which is strictly decreasing.

From Proposition 2.8, $\exists\ a \in \mathbb{R}$, such that $f\left(p_s^{(k)}\right) \ge a$, $\forall\ k \in \mathbb{N}$. Therefore, this sequence can be bounded from below. By the monotone convergence theorem [44, corollary 2.11], $\lim_{k\to\infty} f\left(p_s^{(k)}\right) = \inf_{k\in\mathbb{N}} f\left(p_s^{(k)}\right) := f^*$, where f^* is a local minimum.

The convergence rate $c \in (0,1)$ is defined as the smallest real number such that $|f\left(p_s^{(k)}\right) - f\left(p_s^{(0)}\right)| \le c^k \cdot |f\left(p_s^{(1)}\right) - f\left(p_s^{(0)}\right)|$, [33, p. 480]. The determination of the convergence rate as a function of the parameters of our problem is left for a future investigation.

4.5. Computational complexity analysis

As nested Lyapunov equations need to be solved in a loop of Algorithm 1, it is necessary to analyze the computational complexity of the proposed algorithms so as to know their adaptivity for large-scale power networks. Since the dimension of the control vector of the power supply is $n^+ - 1$, see Section 2.2. The term $n^+ - 1$ is now replaced by n, as is commonly done in complexity analysis. Here, we only consider computational complexity with respect to the problem scale, but not the accuracy of the solution.

Solving the Lyapunov equation is the most time-consuming step in Algorithm 1, so the computational complexity of Algorithm 1 is determined by this step. The computational complexity of solving an n-dimensional Lyapunov equation is $O(n^3)$. Note that the computation of the Hessian matrix is not needed in the generation of an approximation sequence. Therefore, the second loop in Algorithm 1 is not needed. However, the computation of the Hessian matrix may be needed for the refinement of a higher-accuracy algorithm, such as a second-order algorithm, in future work. To summarize, the computational complexity of Algorithm 1 is $O(n^4)$.

The line search step of Algorithm 2, i.e., step 7 of the algorithm, has a computational complexity of $O(n)$. However, the computational complexity of Algorithm 2 lies at the projection onto the feasible set e.g. step 2, which needs solving a linear constrained quadratic optimization problem, see Remark 4.6, and the computational complexity is up to $O(n^{3.5})$. Hence, the computational complexity of Algorithm 2 is $O(n^{3.5})$. The computational complexity of Algorithm 3 has a similar explanation, resulting in computational complexity of $O(n^{3.5})$.

In conclusion, the proposed algorithms, which are integrated in Procedure 4.9, have a computational complexity of $O(n^4)$.

Proposition 4.11. *The steepest descent Algorithm 2 is only locally convergent. To enhance its effectiveness, it is essential to determine an initial vector that allows the objective function to decrease significantly. Hence, a procedure is needed to compute an effective initial vector for Algorithm 2. The projected generalized subgradient method is chosen, and one can see an application of the subgradient method for solving the dual problem of a Lagrangian relaxation in [45]. Note that the algorithm for the chosen method is similar to Algorithm 3, and the difficulty lies in the formula of the generalized subgradient of our objective function. The formulas are shown in [30, Proposition 5.7], where two main theories are referred, [12, theorem 10.27], [21, corollary 4.4]. Readers interested in its convergence analysis for this problem can explore [30, Theorem 5.10]. The computational complexity of the corresponding algorithm is up to $O(n^{3.5})$.*

Note that the algorithm introduced in Proposition 4.11 will be referred to as the first-step algorithm, and the steepest descent algorithm presented in Procedure 4.9 will be referred to as the second-step algorithm.

5. Application to a dual-ring power network

In this section, an illustrative example will be provided to demonstrate the practical application of the algorithms introduced in this paper, which is a two-step procedure. First of all, the decreasing and the convergence of the sequence generated by the proposed approach will be verified. Secondly, the aim of the considered problem that the maximal exit probability from the safe subset $(-\pi/2, \pi/2)$ of all the phase-angle differences is minimized by selecting power supply vectors (or, in other words, by the optimal power dispatches) will be checked and illustrated. To show the effect of the proposed approach, the performance of a stochastic power system *with using* the control of power dispatch proposed in this paper, is compared with that *without using* the proposed control. The control law used in the case of *without using* is called the *proportional control law*, which serves as a benchmark of the proposed approach in this paper, and it is defined by the following.

For a short horizon, the *proportional control law* is defined for the power supply vector by the following procedure. Consider the vector of maximal available power supply $p^{+,max}$ and the vector of power demand p^- which satisfies that $\sum_{i=1}^{n^+} p_i^{+,max} = p_{sum}^{+,max} \geq p_{sum}^- = \sum_{j=1}^{n^-} p_j^-$. Choose $p_{sum}^+ = p_{sum}^-$, hence power supply equals power demand. Define the fraction $s = p_{sum}^+/p_{sum}^{+,max} \in (0,\ 1)$. Define then the power supply vector $p^+ \in \mathbb{R}^{n^+}$, $p_i^+ = s \times p_i^{+,max}$ for all $i \in \mathbb{Z}_{n^+}$, then $\sum_{i=1}^{n^+} p_i^+ = p_{sum}^+ = p_{sum}^-$.

Denote the exit probability distribution functions as follows, $f_{a,k} = P(\omega \in \Omega \mid \theta_{i_k,j_k}(\omega,t) < -\pi/2)$, $f_{b,k} = P(\omega \in \Omega \mid \theta_{i_k,j_k}(\omega,t) > \pi/2)$, and $\bar{p}_{out,k}$ is the upper bound of the exit probabilities of the phase-angle differences in all the power lines, computed according to the numerical solution of our proposed approach, which equals two times the larger

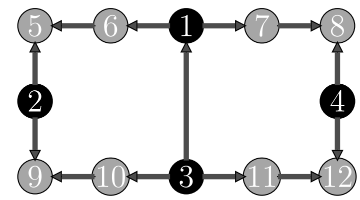

Fig. 1. A connected network with two subrings.

one between $f_{a,k}$ and $f_{b,k}$ when line k is the most vulnerable power line, [25, Appendix 1.4 (b)]. The comparison of *without using* and *with using* the proposed approach will focus on these three values,

The network topology is depicted in Fig. 1, with nodes labeled from 1 to 12. Nodes $1, 2, 3, 4$ represent power supply nodes, while nodes $5, 6, 7, 8, 9, 10, 11, 12$ are designated as power demand nodes. This network model emulates two adjacent rings of a power transmission system. Details regarding the power network's parameters and the computational results are presented in Table 1, and the iterations of the proposed algorithms are plotted in Fig. 2. All computations were conducted using MATLAB 2021b on a 64-bit MacBook equipped with an M1 chip and 8 GB of RAM.

In Figure (2.1), one can see that the sequence of the objective function values generated by the first-step algorithm stated in Proposition 4.11 continues decreasing until the 50th step, but it also decreases at the 300th step. Thus, we use the optimal point of the first iteration sequence as the initial point and do the iteration again, whose trajectory of iteration is plotted in Figure (2.2). The reason for executing the first-step algorithm twice is to compute a more effective initial point for the steepest descent algorithm. The convergence of the sequence generated by the first-step algorithm is clear from Figure (2.1) and Figure (2.2), although the convergence speed is not very rapid. From Figure (2.3), one can see that the iteration sequence converges at the 20th step, and therefore, the convergence speed of the second-step algorithm is much quicker than the first-step algorithm. In conclusion, Fig. 2 demonstrates the validity of the proposed two-step procedure in terms of computation.

Table 1.D provides results obtained using a different initial vector. A comparison between Table 1.C and Table 1.D reveals the presence of two distinct isolated local minimizers for the objective function. The exit probabilities of several critical power lines *with* and *without using* the proposed approach are displayed in Tables 2 and 3. In Table 2, the most vulnerable line is l_{4-12} and in Table 3, the line which endangers the small-signal stability is also l_{4-12}. Note in Table 2, there is a considerable reduction of $f_{b,k}$ in line l_{4-12} from $1.4923e-04$ for *without using* to $1.0863e-04$ for *with using* the proposed approach. The values of the other lines of Table 2 for $f_{b,k}$ are lower than that of the first line. Thus, the maximal exit probability is decreased from $1.4923e-04$ to $1.0863e-04$. For Table 3, the same conclusions hold. At the same time, the exit probabilities of other lines may increase or decrease; see the change in values of line l_{3-11}, line l_{7-8} from Table 2 and line l_{3-11}, l_{1-7} from Table 3, respectively. That is due to the choice of the objective function. Through the comparisons of *without using* and *with using* the proposed approach concerning the maximal exit probability of all power lines, it is clear that the small-signal stability of a stochastic power system can be improved by the proposed approach in this paper.

6. Conclusions and further research

This research provides rigorous mathematical analysis and algorithms for solving the optimization challenges related to the small-signal stability of stochastic power systems. Our investigation focuses on reducing the maximal exit probability of all power lines of the

Table 1
Several tables.

Table 1.A. Parameter specifications for the domain of feasible vectors.

$p_1^{+,max}$	$p_2^{+,max}$	$p_3^{+,max}$	$p_4^{+,max}$	$-p_5^-$	$-p_6^-$	$-p_7^-$	$-p_8^-$	$-p_9^-$	$-p_{10}^-$	$-p_{11}^-$	$-p_{12}^-$
25	25	25	25	−8	−12	−13	−7	−8	−12	−11	−9

Table 1.B. The parameter specifications of the power network are listed in this table, though not listed are that the weights of all power lines equal 20.

Node	1	2	3	4	5	6	7	8	9	10	11	12
Inertia	10	10	10	10	1	1	1	1	1	1	1	1
Damping	4	4	4	4	1	1	1	1	1	1	1	1
Noise	2	2.3	2.5	2.7	1.6	1.7	1.8	1.9	1.65	1.75	1.85	2.05

Table 1.C. The minimum of the objective function and optimal power supply vector, with $[23, 19, 24]$ as the initial power supply vector.

Minimum	p_1^+	p_2^+	p_3^+	p_4^+
1.39540	19.8388	19.8812	20.8194	19.4606

Table 1.D. The optimal objective function value and optimal power supply vector, with $[19, 19, 19]$ as the initial power supply vector.

Minimum	p_1^+	p_2^+	p_3^+	p_4^+
1.39538	14.3187	22.1401	23.1106	20.4306

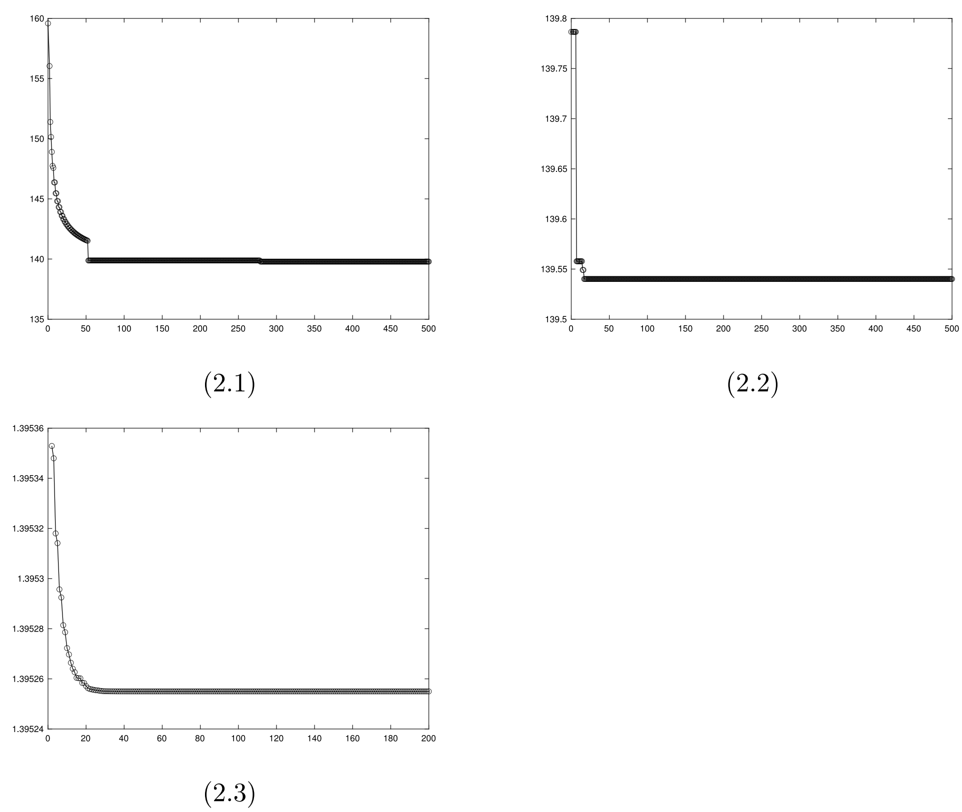

Fig. 2. Figure (2.1) depicts the initial iteration of the first-step algorithm using the projected generalized subgradient method with the starting vector set as $[23, 19, 24]$. The x-axis represents the iteration number, and the y-axis reflects the value of the objective function multiplied by a factor of 100.
Figure (2.2) shows the second iteration of the first-step algorithm using the projected generalized subgradient method. The initial vector is the optimal vector computed in (3.a). The axes are similar to (2.1).
Figure (2.3) displays the iteration of the second-step algorithm using the steepest descent method. The initial vector for this iteration is the optimal vector obtained in (2.2). Here, the x-axis corresponds to the iteration number, and the y-axis displays the objective function value.

Table 2
The tail probabilities of both *without using* and *with using* the proposed approach, according to Table 1.C.

Power line	*Without using* Exit probabilities		*With using* Exit probabilities		
$i_k - j_k$	$f_{a,k}$	$f_{b,k}$	$f_{a,k}$	$f_{b,k}$	$\bar{p}_{out,k}$
4–12	6.2453e−14	1.4923e−04	8.3976e−14	1.0863e−04	2×1.0863e−04
3–11	2.9882e−15	6.9937e−05	5.4638e−14	1.0674e−04	2×1.0717e−04
1–7	7.5828e−14	7.7720e−05	2.4307e−15	8.6967e−05	2×9.6974e−05
7–8	1.5517e−08	2.2977e−11	1.2177e−08	2.9741e−11	2×8.6042e−05

Table 3
The tail probabilities of both *without using* and *with using* the proposed approach, according to Table 1.D.

Power line	*Without using* Exit probabilities		*With using* Exit probabilities		
$i_k - j_k$	$f_{a,k}$	$f_{b,k}$	$f_{a,k}$	$f_{b,k}$	$\bar{p}_{out,k}$
4–12	6.2453e−14	1.4923e−04	8.5292e−14	1.0866e−04	2×1.0866e−04
3–11	2.9882e−15	6.9937e−05	5.4630e−14	1.0715e−04	2×1.0719e−04
1–7	7.5828e−14	7.7720e−05	7.6432e−19	7.2509e−05	2×7.7557e−05
7–8	1.5517e−08	2.2977e−11	3.8159e−08	9.1109e−12	2×8.6573e−05

stochastic linearized power system through a complex nonlinear objective function that incorporates implicit functions, maximum, and absolute value operators. The proposed algorithms have a computational complexity of $O(n^4)$. The designed algorithms can be directly utilized by electrical engineers to improve the small-signal stability of stochastic power systems.

For future research, the derivation of the convergence rates and further improvement of the computational efficiency of the proposed algorithms are important remaining issues. Refining the second-order algorithm is expected to enhance the accuracy of our methodological framework. Additionally, exploring algorithms with global convergence properties, as suggested in publications such as [36,37], presents a promising direction for our future investigations. Another promising and important trajectory is the interaction with real-world power network operations.

CRediT authorship contribution statement

Zhen Wang: Writing – original draft, Validation, Investigation, Formal analysis. **Kaihua Xi:** Supervision, Investigation, Conceptualization. **Aijie Cheng:** Supervision, Project administration. **Hai Xiang Lin:** Writing – review & editing, Supervision, Project administration. **Jan H. van Schuppen:** Writing – review & editing, Project administration, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zhen Wang reports financial support was provided by China Scholarship Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Zhen Wang has visited Department of Applied Mathematics, Delft University of Technology, Delft, The Netherlands from November 2021 till February 2024 where the results of this paper were developed, and he is grateful to the China Scholarship Council, China (No. 202106220104) for financial support for his stay at Delft University of Technology.

Data availability

The data is available in the manuscript.

References

[1] Kundur P. Power system stability and control. New York: McGraw-Hill; 1994.
[2] Kundur P, Paserba J, Ajjarapu V, Andersson G, Bose A, Canizares C, et al. Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions. IEEE Trans Power Syst 2004;19(3):1387–401.
[3] Bueno PG, Hernández JC, Ruiz-Rodriguez FJ. Stability assessment for transmission systems with large utility-scale photovoltaic units. IET Renew Power Gener 2016;10(5):584–97.
[4] Hernández JC, Bueno PG, Sanchez-Sutil F. Enhanced utility-scale photovoltaic units with frequency support functions and dynamic grid support for transmission systems. IET Renew Power Gener 2017;11(3):361–72.
[5] Hernández JC, Sanchez-Sutil F, Vidal P, Rus-Casas C. Primary frequency control and dynamic grid support for vehicle-to-grid in transmission systems. Int J Electr Power Energy Syst 2018;100:152–66.
[6] Anderson PM, Bose A. A probabilistic approach to power system stability analysis. IEEE Trans Power Appar Syst 1983;102:2430–9.
[7] De Marco CL, Bergen AR. A security measure for random load disturbances in nonlinear power system models. IEEE Trans Circuits Syst 1987;34:1546–57.
[8] Dong ZY, Zhao JH, Hill DJ. Numerical simulation for stochastic transient stability assessment. IEEE Trans Power Syst 2012;27:1741–9.
[9] Perninge M, Knazkins V, Amelin M, Soder L. Risk estimation of critical time to voltage instability induced by saddle–node bifurcation. IEEE Trans Power Syst 2010;25(3):1600–10.
[10] Wu X, Xi K, Cheng A, Zhang C, Lin HX. A critical escape probability formulation for enhancing the transient stability of power systems with system parameter design. Automatica 2025;175:112217.
[11] Clarke FH. Generalized gradients and applications. Trans Amer Math Soc 1975;205:247–62.
[12] Clarke F. Functional analysis, calculus of variations and optimal control. Graduate texts in mathematics, London: Springer; 2013.
[13] Mordukhovich B. Variational analysis and generalized differentiation – I: basic theory. Heidelberg: Springer-Verlag; 2006.
[14] Mordukhovich B. Variational analysis and generalized differentiation – II: applications. Heidelberg: Springer-Verlag; 2006.
[15] Cui Y, Pang J-S. Modern nonconvex nondifferentiable optimization. Philadelphia: SIAM; 2021.
[16] Mordukhovich B, Nam NM. An easy path to convex analysis and applications. Cham: Springer Nature; 2023.
[17] Mifflin R. Semismooth and semiconvex functions in constrained optimization. SIAM J Control Optim 1977;15(6):959–72.
[18] Rockafellar RT. Generalized directional derivatives and subgradients of nonconvex functions. Canad J Math 1980;32(2):257–80.
[19] Balder EJ. On generalized gradients and optimization. Report, Utrecht, The Netherlands: Department of Mathematics, University of Utrecht; 2009.
[20] Cánovas MJ, López MA, Mordukhovich BS, Parra J. Variational analysis in semi-infinite and infinite programming, II: Necessary optimality conditions. SIAM J Optim 2010;20(6):2788–806.
[21] Mordukhovich BS, Nghia TT. Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data. SIAM J Optim 2013;23(1):406–31.
[22] Ben-Tal A, Zowe J. Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems. Math Program 1982;24:70–91.
[23] Ben-Tal A, Zowe J. Directional derivatives in nonsmooth optimization. J Optim Theory Appl 1985;47(4):483–90.

[24] Shapiro A. On concepts of directional differentiability. J Optim Theory Appl 1990;66:477–87.
[25] Wang Z, Xi K, Cheng A, Lin HX, van Schuppen JH. Control of the power flows of a stochastic power system. 2024, Report arXiv:2311.16282v2, eess.SY, arXiv.
[26] Wang Z, Xi K, Cheng A, Lin HX, Ran AC, van Schuppen JH, et al. Synchronization of power systems under stochastic disturbances. Automatica 2023;151:110884.
[27] Kushner HJ, Dupuis P. Numerical methods for stochastic control problems in continuous time. Springer; 2001.
[28] Bourbaki N. Elements of mathematics. Functions of a real variable, Berlin: Springer; 2004.
[29] Xi K, Wang Z, Cheng A, Lin HX, van Schuppen JH, Zhang C. Synchronization of coupled phase oscillators with stochastic disturbances and the cycle space of the graph. SIAM J Appl Dyn Syst 2023;22(2):1030–52.
[30] Wang Z, Xi K, Cheng A, Lin HX, van Schuppen JH. Optimal control of a stochastic power system – algorithms and mathematical analysis. 2024, Report arXiv:2401.16879, math.OC.
[31] Rockafellar R. Convex analysis, princeton landmarks in mathematics and physics. New Jersey: Princeton University Press; 1970.
[32] Rockafellar R, Wets R. Variational analysis. 3rd ed.. Heidelberg: Springer-Verlag; 2009.
[33] Boyd SP, Vandenberghe L. Convex optimization. Cambridge: Cambridge University Press; 2004.
[34] Nesterov Y. In: Lectures on convex optimization, vol. 137, Heidelberg: Springer; 2018.
[35] Cox D, Little J, O'Shea D. Ideals, varieties, algorithms: an introduction to computational algebraic geometry and commutative algebra. Undergraduate texts in mathematics, Heidelberg: Springer; 1992.
[36] Liu T, Pong TK, Takeda A. A successive difference-of-convex approximation method for a class of nonconvex nonsmooth optimization problems. Math Program 2019;176:339–67.
[37] Wang Y, Yin W, Zeng J. Global convergence of ADMM in nonconvex nonsmooth optimization. J Sci Comput 2019;78:29–63.
[38] Wen B, Chen X, Pong TK. Linear convergence of proximal gradient algorithm with extrapolation for a class of nonconvex nonsmooth minimization problems. SIAM J Optim 2017;27(1):124–45.
[39] Poolla BK, Bolognani S, Dörfler F. Optimal placement of virtual inertia in power grids. IEEE Trans Autom Control 2017;62(12):6209–20.
[40] Boyd S, Xiao L, Mutapcic A. Subgradient methods, lecture notes of EE392o. Autumn Quarter: Stanford University; 2003.
[41] Zhang Y, Wang J. A dual neural network for convex quadratic programming subject to linear equality and inequality constraints. Phys Lett A 2002;298(4):271–8.
[42] Danilova M, Dvurechensky P, Gasnikov A, Gorbunov E, Guminov S, Kamzolov D, et al. Recent theoretical advances in non-convex optimization. In: High-dimensional optimization and probability: with a view towards data science. Cham: Springer International Publishing; 2022, p. 79–163.
[43] Bauschke HH, Borwein JM. On projection algorithms for solving convex feasibility problems. SIAM Rev 1996;38(3):367–426.
[44] Browder A. Mathematical analysis: an introduction. Heidelberg: Springer Science & Business Media; 2012.
[45] Nedić Angelia, Ozdaglar Asuman. Approximate primal solutions and rate analysis for dual subgradient methods. SIAM J Optim 2009;19(4):1757–80.