

Discrete Shortest Paths in Optimal Power Flow Feasible Regions

Daniel Turizo, *Member, IEEE*, Diego Cifuentes, Anton Leykin, and Daniel K. Molzahn, *Senior Member, IEEE*

Abstract—Optimal power flow (OPF) is a critical optimization problem for power systems to operate at points where cost or operational objectives are optimized. Due to the non-convexity of the set of feasible OPF operating points, it is non-trivial to transition the power system from its current operating point to the optimal one without violating constraints. On top of that, practical considerations dictate that the transition should be achieved using a small number of small-magnitude control actions. To solve this problem, this paper proposes an algorithm for computing a transition path by framing it as a shortest path problem. This problem is formulated in terms of a discretized piece-wise linear path, where the number of pieces is fixed a priori in order to limit the number of control actions. This formulation yields a nonlinear optimization problem (NLP) with a block tridiagonal structure, which we leverage by utilizing a specialized interior point method. An initial feasible path for our method is generated by solving a sequence of relaxations which are then tightened in a homotopy-like procedure. Numerical experiments illustrate the effectiveness of the algorithm.

Index Terms—Optimal power flow, shortest path, nonlinear optimization, interior point method

I. INTRODUCTION

THE optimal power flow (OPF) is arguably the most important problem in steady state power system operation. OPF is an optimization problem that seeks to minimize an objective (usually operation cost) subject to the power flow equations governing the power system behavior and the engineering and technical constraints associated with physical operation of the system and its components [1]. A complete formulation of the OPF problem, called Alternating Current OPF (ACOPF), is a nonconvex problem with nonlinear equality constraints and hundreds to thousands of variables.

After solving an ACOPF problem, the operator must determine how to transition the system from the current operating point to the resulting optimal point. An OPF solution provides values that controllable variables must take in order to minimize the objective function. Such variables may be manipulated physically by, for example, controlling a flood-gate in a hydro plant or the boiler in a thermal plant. As such, the transition process between values of the controllable variables must be performed in terms of a sequence of few simple control actions, as the physical implementation limits the complexity of the execution. Furthermore, the transition

between states should respect the system constraints in the same way that the optimal solution does.

The problem of state transitioning in terms of few simple actions is not trivial, but some approaches have been explored in the literature. Some authors have used linear OPF approximations to tractably generate the transition as a sequence of corrective actions involving a small subset of the controllable variables. References [2] and [3] construct a mixed-integer linear program (MILP) as an approximation to the ACOPF, while also adding hard constraints on the amount of controllable variables modified. Reference [4] applies sparse techniques based on high-dimensional statistics to the DCOPF formulation to generate sparse solutions with respect to a base state. These approaches, while tractable, rely on linear approximations to the original problem, and so they do not guarantee that constraints are not violated during the transition. Moreover, these linear approximations improve tractability at the expense of ignoring the non-convex and possibly non-connected geometry of the feasible space [5], [6]. In light of these drawbacks, [7] and [8] extend previous formulations to consider the full ACOPF, obtaining a mixed-integer nonlinear program (MINLP). These papers approximate the binary constraints in the MINLP using barrier functions, obtaining a continuous nonlinear program (NLP). This new approximation represents the original feasible set more accurately, yet still does not guarantee feasibility during the transition.

The issue of guaranteeing feasibility during the transition process has been tackled by recent work in [9] and [10]. Reference [9] proposes a method for iteratively generating a sequence of convex restrictions (i.e., convex inner approximations) for the ACOPF feasible set. The sequence of sets are pairwise connected, and at some point the method generates a convex restriction containing the optimal operating point. The output of the method is a finite sequence sequence of operating points which define a piece-wise linear path connecting the current operating point and the optimal operating point. This path is guaranteed to be feasible, as it is contained in a chain of connected convex restrictions containing both operating points. Reference [10] proposes an algorithm for iteratively generating a sequence of feasible operating points using sensitivity information and a Newton iteration. The transition is constructed using each point in the sequence. The main drawback of approaches like those of [9] and [10] is that there is not control over the number of intermediate operating points generated during the iteration process. That is, while these methods output a finite sequence of intermediate transition points, the length of the sequence can be arbitrarily large.

An important issue that, to the authors knowledge, has

Daniel Turizo and Daniel K. Molzahn are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, {djturizo,molzahn}@gatech.edu. Support from NSF contract #2023140.

Diego Cifuentes is with the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, dfc3@gatech.edu.

Anton Leykin is with the School of Mathematics, Georgia Institute of Technology, leykin@math.gatech.edu. Support from NSF DMS award #2001267.

not been studied in the literature regards the *amplitude*. By amplitude we mean the size of the change each variable undertakes during a control action (or equivalently, the distance between states before and after the control action takes place). Even if the transition can be done using a few control actions involving few variables, large amplitudes for these actions can be detrimental. For example, large amplitude control actions in battery energy storage systems can increase the depth-of-discharge, thus increasing battery degradation [11]. Ideally, the best transition path would be the straight line joining the current and optimal operating points since this path represents a single control action with the minimal possible amplitude. If the constraints are violated by the straight line, the transition path should be modified to avoid constraint violations, thus increasing the number and amplitude of control actions.

This paper addresses two of the issues of operating point transition: the number and amplitude of control actions. We formulate the problem of minimizing the amplitude of control actions as a *shortest path* problem that seeks the shortest path joining the current and optimal operating points inside the feasible space. To this end, we propose an algorithm that computes a piece-wise linear approximation of this shortest path as a discretized path defined in terms of a chosen number of intermediate operating points. We formulate the shortest path problem as an NLP where the objective function is the path length and the optimization variables are the coordinates of the intermediate operating points, subject to the ACOF constraints. The NLP is solved using a feasible interior point method coupled with an homotopy procedure to generate an initial feasible path. When the interior point method is applied to our formulation, the matrices involved show a block tridiagonal structure. We show how to exploit this structure to reduce the interior point method's complexity, so that each iteration scales linearly with the number of intermediate operating points. We thus obtain a scalable algorithm that minimizes the amplitude of control actions and enables specifying the number of intermediate points. Numerical experiments on multiple test cases of varying sizes show the algorithm's effectiveness in finding a discretized shortest path for a specific number of points.

The rest of the paper is organized as follows. Section II describes the formulation of the ACOF problem and the corresponding shortest path problem. Section III elaborates on the implementation of a feasible interior point method that leverages the special structure of the shortest path problem. Section IV provides a description of the complete algorithm, including a homotopy procedure for generating an initial feasible path required to execute the interior point algorithm. Section V illustrates the numerical experiments we performed. Section VI discusses conclusion and future work.

II. SHORTEST PATH OPF PROBLEM FORMULATION

We consider an arbitrary power system with two different operating points of interest. We wish to connect these points through a continuous path such that every point in the path is a feasible operating condition with respect to the OPF constraints. For a power system with n buses, let $x \in \mathbb{R}^{2n}$ denote the real and imaginary parts of the voltage phasors

for all buses, i.e., the state vector of the power system. Let $u \in \mathbb{R}^{2g}$ denote the vector of controlled variables¹, where $g \leq n$ is the number of generators. In particular, we denote the points we want to connect by u_0 and $u_1 \neq u_0$. The relationship between x and u is given by the power flow equations:

$$f(x, u) = [f_1(x, u), \dots, f_{2n}(x, u)]^T = 0 \in \mathbb{R}^{2n},$$

$$f_k(x, u) = \begin{cases} \frac{1}{2}x^T H_k x + r_k^T x + c_k - u_k, & k \leq 2g, \\ \frac{1}{2}x^T H_k x + r_k^T x + c_k, & k > 2g \end{cases} \quad (1)$$

for appropriate symmetric matrices $H_k \in \mathbb{R}^{2n \times 2n}$ (which correspond to the Y_k matrices in [12]) and vectors $r_k \in \mathbb{R}^{2n}$. Matrices H_k are highly structured: they have at most two non-zero rows and columns, and they have at most rank four.

The OPF feasible set consists of all pairs (u, x) satisfying the power flow equations and the OPF constraints g_i and h_i (like voltage limits, line flow limits, etc.):

$$g_i(u) \leq 0, \quad i \in \mathcal{U}, \quad (2a)$$

$$h_i(x) \leq 0, \quad i \in \mathcal{X}, \quad (2b)$$

for appropriate disjoint index sets \mathcal{U}, \mathcal{X} . We assume that all OPF constraints inequalities depend on either u ($i \in \mathcal{U}$) or x ($i \in \mathcal{X}$), but not both.² The vector x corresponds to the state vector associated with u that satisfies (1). The existence of such x is not trivial, for some values u there exists multiple solutions or possibly none [13]. From the implicit function theorem [14], we can specify a branch of the mapping to define a continuous and injective function φ from u to x in a neighborhood of u , as long as the Jacobian of (1) with respect to x is non-singular in said neighborhood (see Fig. 1). We can use this information to restrict ourselves to a single branch of the mapping. Consider the pair (u_0, x_0) where u_0 is the starting operating point and x_0 is the solution of (1) associated with u_0 for the branch we are interested in. Let $J(x) = \partial f(x, u) / \partial x$ denote the Jacobian of the power flow equations with respect to the state vector x (the Jacobian with respect to x is independent of u). If we assume that $J(x_0)$ is non-singular, then there exists a continuous and injective function $\varphi(u)$ defined by the branch of (1) satisfying $\varphi(u_0) = x_0$. We impose the additional constraint $u \in \mathcal{F}$ where \mathcal{F} is defined as

$$\mathcal{F} = \{u \in \mathbb{R}^{2g} : J(\varphi(u)) \text{ is not singular}\}, \quad (3a)$$

$$\mathcal{F} = \{u \in \mathbb{R}^{2g} : -|\det J(\varphi(u))| < 0\}. \quad (3b)$$

To use this formulation, we require some assumptions:

- **Assumption 1:** The Jacobian $J(x_0)$ is non-singular.
- **Assumption 2:** The function $\varphi(u)$ can be computed.
- **Assumption 3:** Both u_0 and u_1 belong to the same connected component of \mathcal{F} .

¹Usually the controlled variables of OPF problem are the voltage magnitude and active power outputs of each generator. Other type of controlled variables are valid, as long as they fit within the proposed framework.

²In the standard OPF problem the entries of u are the generator voltage magnitudes and active power of PV buses. As such, $g_{\mathcal{U}}$ contains the voltage limits of generator buses and the active power limits of PV buses. On the other hand, $g_{\mathcal{X}}$ contains the voltage and active power limits of remaining buses, reactive power limits, line flow limits, and angle difference constraints.

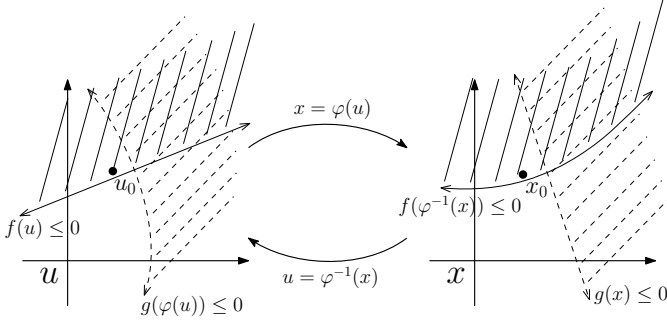


Fig. 1. Variables u and x in a neighborhood of u_0 and x_0 are related by the power flow mapping φ . Feasible sets generated by inequalities in x can be mapped back to feasible sets in u and vice-versa. As the power flow mapping φ is nonlinear, the geometry of the mapped feasible sets will be altered.

In particular, u_0 and u_1 may be in different connected components if their associated states x_0 and x_1 belong to different branches of (1).

Under the previous assumptions, we can define the functions g_i for all $i \in \mathcal{X}$ as

$$g_i(u) = h_i(\varphi(u)). \quad (4)$$

This way all constraints depend only on u now, so we no longer need to consider the state vector x as an optimization variable. In most cases, x is multiple times larger than u , so there is a significant computational gain in reducing the dimension of the optimization problem. This gain does not come for free though, as the power flow equations need to be solved to find $\varphi(u)$, and derivative computations become more involved due to the implicit function φ . Moreover, we have to make sure that φ is indeed defined at a given vector u . We achieve this by introducing an additional constraint representing the power flow feasible set. For some singleton index set \mathcal{P} disjoint from \mathcal{U}, \mathcal{X} , we define

$$g_i(u) = -|\det J(\varphi(u))|, \quad i \in \mathcal{P}. \quad (5)$$

Define $\mathcal{I} = \mathcal{U} \cup \mathcal{X} \cup \mathcal{P}$. The interior of the power flow feasible set ($g_i, i \in \mathcal{P}$) and the OPF constraints' feasible set ($g_i, i \in \mathcal{U} \cup \mathcal{X}$) is given by all points $u \in \mathbb{R}^{2g}$ such that

$$g_i(u) < 0, \quad i \in \mathcal{I}. \quad (6)$$

For interior point methods, the distinction between $<$ and \leq is inconsequential, as the numerical solution always lies in the interior of the feasible set.

A. Optimal Control Problem

Finding a path between two points in a set is a classical optimal control problem. If we seek the *shortest* path, we then obtain an optimization problem. We define a continuation parameter $t \in [0, 1]$ and the decision vector $u(t) \in C[0, 1]$, where $C[0, 1]$ denotes the set of continuous functions defined on the interval $[0, 1]$. The shortest path problem is

$$\begin{aligned} \inf_u \quad & \int_0^1 (u'^T(t)u'(t))^{1/2} dt \\ \text{s.t.} \quad & u(0) = u_0, \quad u(1) = u_1, \\ & g_i(u(t)) < 0, \quad \forall t \in [0, 1], \quad i \in \mathcal{I}. \end{aligned} \quad (7)$$

This is a calculus of variations problem with constraints. The objective function may not be differentiable at some points (due to the square root). Moreover, problem (7) is naturally ill-defined, as even in the unconstrained case there are infinite gradient maps that yield a straight line between u_0 and u_1 . These issues can be avoided by requiring the gradient map to have constant norm (constant “speed” of transition along the path), which also simplifies the objective function. To illustrate this, assume that the path has constant norm, i.e. $\|u'(t)\| = \zeta > 0$ for all $t \in [0, 1]$, then the objective function becomes

$$\int_0^1 (u'^T(t)u'(t))^{1/2} dt = \int_0^1 \|u'(t)\| dt = \int_0^1 \zeta dt = \zeta, \quad (8)$$

so ζ not only denotes the “speed” of a particle traversing the path but also the “time” it takes for the particle to go from u_0 to u_1 . This formulation yields the following eikonal equation problem in terms of the arclength ζ (see [15], [16]):

$$\begin{aligned} \inf_{u, \zeta} \quad & \zeta \\ \text{s.t.} \quad & u(0) = u_0, \quad u(1) = u_1, \\ & \|u'(t)\| = \zeta, \quad \forall t \in [0, 1], \\ & g_i(u(t)) < 0, \quad i \in \mathcal{I}. \end{aligned} \quad (9)$$

Any numerical approach to solving this problem must honor the feasible set constraints in (6), as there does not exist a state vector x associated with any $u \notin \mathcal{F}$. Also, there is no trivial feasible starting path available in general. To circumvent this issue, we next propose a discretized version of the problem.

B. Piece-wise Linear Path Approximation

We restrict the search space from $C[0, 1]$ to the space of piece-wise linear paths $PL[0, 1]$.³ More specifically, we will consider the space of piece-wise linear paths with $K + 1$ pieces, $PL_{K+1}[0, 1]$. Let the characteristic (sometimes called indicator) function $\chi_E(t)$ be defined as

$$\chi_E(t) = \begin{cases} 1 & t \in E \\ 0 & t \notin E \end{cases} \quad (10)$$

We consider a piece-wise linear path $p(t) \in PL_{K+1}[0, 1]$ defined by its $K + 2$ points $\{p_k\}_{k=0}^{K+1}$ and parameters $\{t_k\}_{k=0}^{K+1}$:

$$\begin{aligned} p(t) &= p_0 \chi_{\{0\}}(t) + \sum_{k=1}^{K+1} c_k(t) \chi_{(t_{k-1}, t_k]}(t), \\ \text{with } c_k(t) &= p_k - 1 + (p_k - p_{k-1}) \frac{t - t_{k-1}}{t_k - t_{k-1}}. \end{aligned} \quad (11)$$

The parameter values t_k satisfy

$$t_0 = 0 < t_1 < \dots < t_K < t_{K+1} = 1. \quad (12)$$

We also set $p_0 = u_0$ and $p_{K+1} = u_1$ to satisfy the endpoint constraints. We want to compute the path $p(t)$ that minimizes the objective function. Note that, for fixed values $\{t_k\}_{k=0}^{K+1}$, $p(t) \in PL_{K+1}[0, 1]$ can be identified with $\{p_k\}_{k=0}^{K+1}$. Thus, the control problem reduces to computing the points $\{p_k\}_{k=0}^{K+1}$ that minimize the objective (recall that $p_0 = u_0$ and $p_{K+1} = u_1$):

$$\begin{aligned} \inf_{p_1, \dots, p_K, \zeta} \quad & \zeta \\ \text{s.t.} \quad & \|p'(t)\| = \zeta, \quad \forall t \in [0, 1], \\ & g_i(u(t)) < 0, \quad i \in \mathcal{I}. \end{aligned} \quad (13)$$

³Note that $PL[0, 1]$ is dense in $C[0, 1]$ with respect to the uniform norm, as the Schauder system of $C[0, 1]$ is composed of piece-wise linear functions [17].

We concatenate the points $\{p_k\}_{k=1}^K$ into a single vector $p = [p_1^T, \dots, p_K^T]^T \in \mathbb{R}^{2gK}$. Replacing (11) in (13) yields

$$\begin{aligned} \inf_{p, \zeta} \quad & \zeta \\ \text{s.t.} \quad & c_k(p, t) = p_{k-1} + (p_k - p_{k-1}) \frac{t - t_{k-1}}{t_k - t_{k-1}}, \\ & \|c'_k(p, \tau)\| = \zeta, \quad \forall \tau \in (t_{k-1}, t_k], \\ & g_j(c_k(p, \tau)) < 0, \quad j \in \mathcal{I}, \\ & \forall t \in [0, 1], \quad k = 1, \dots, K+1. \end{aligned} \quad (14)$$

The optimization problem is now finite dimensional, yet the constraints are still infinite dimensional. For the purpose of tractability, we will relax the constraints by only enforcing them at the corner points p_k . This means that the path may violate constraints in between corner points. However, if needed, we can add more discretization points to mitigate this issue. As each piece of the path is linear, the infinite-dimensional constant speed constraint is equivalent to the finite dimensional constraint that enforces the slopes of each piece of the path to be equal in norm. Also note that the constant speed constraint implies that $\zeta \geq 0$, so minimizing ζ is equivalent to minimizing ζ^2 . These changes yield the following problem:

$$\begin{aligned} \inf_{p, \zeta} \quad & \zeta^2 \\ \text{s.t.} \quad & \frac{\|p_k - p_{k-1}\|}{t_k - t_{k-1}} = \zeta, \quad k = 1, \dots, K+1, \\ & g_j(p_i) < 0, \quad j \in \mathcal{I}. \end{aligned} \quad (15)$$

The norm constraints are nonlinear inequalities, and hence are non-convex. Any solution method for this problem should be able to at least converge to a local optimum, even in the presence of non-convexities. To this end, we will reformulate the problem in a way that is advantageous for the numerical method we will use. Define $w_k = (t_k - t_{k-1})^{-2}/(K+1) > 0$ for $k = 1, \dots, K+1$. Then (15) is equivalent to

$$\inf_p \quad \sum_{k=1}^{K+1} w_k \|p_k - p_{k-1}\|^2 \quad (16a)$$

$$\text{s.t.} \quad w_{i+1} \|p_{i+1} - p_i\|^2 = w_i \|p_i - p_{i-1}\|^2, \quad i = 1, \dots, K, \quad (16b)$$

$$g_j(p_i) < 0, \quad j \in \mathcal{I}. \quad (16c)$$

In this formulation, the Hessian of the objective function is positive definite, which will prove useful for the interior point iteration described in the next section.

III. LOG-BARRIER NEWTON METHOD IMPLEMENTATION

The discretized shortest path problem in (16) has a tridiagonal structure which is not leveraged by standard interior point solvers. For this reason, we developed an specialized interior point implementation that makes use of the problem structure to reduce the computational complexity of solving the problem. This section is dedicated to explaining in detail the core iterative process behind this specialized solver.

A. Interior Point Iteration

The shortest path problem has a parallel structure since the constraints do not depend on the full decision vector p , but only on their associated point p_i . The only source of coupling between points comes from the objective and the equality constraints, which both have simple block-tridiagonal structures that can be exploited by a specialized interior point iteration. The log-barrier formulation that is central to the interior point method embeds the inequality constraints into the objective and then numerically solves the first-order Karush-Kuhn-Tucker (KKT) equations. We define the index set $\mathcal{E} = \{1, \dots, K\}$ and the functions c_i as

$$c_i(p) = w_i \|p_i - p_{i-1}\|^2 - w_{i+1} \|p_{i+1} - p_i\|^2, \quad i \in \mathcal{E}, \quad (17)$$

where p_i is defined as usual. We also define the objective as

$$\phi(p) = \sum_{k=1}^{K+1} w_k \|p_k - p_{k-1}\|^2. \quad (18)$$

We reformulate (16) using a small barrier parameter $\mu > 0$:

$$\begin{aligned} \inf_{p, s} \quad & \phi(p) - \mu \sum_{i=1}^K \sum_{j \in \mathcal{I}} \ln[(s)_{|\mathcal{I}|(i-1)+j}] \\ \text{s.t.} \quad & g_j(p_i) + (s)_{|\mathcal{I}|(i-1)+j} = 0, \quad j \in \mathcal{I}, \quad i = 1, \dots, K, \\ & c_j(p) = 0, \quad j \in \mathcal{E}, \end{aligned} \quad (19)$$

where s is a vector of size $K|\mathcal{I}|$ and $(s)_k$ denotes the k -th entry of s . Note that this formulation is only equivalent to (16) when all the entries of s are strictly positive. However, such constraints are unnecessary, as the logarithmic terms act as a barrier preventing the entries of s from becoming non-positive. Define $g_{\mathcal{I}}(p_i)$ as the vector of inequality constraints (evaluated at a particular point on the path), $c_{\mathcal{E}}(p)$ as the vector of equality constraints, and D_p as the Jacobian operator (with respect to p). Let $y \in \mathbb{R}^{|\mathcal{E}|}$ and $z \in \mathbb{R}^{K|\mathcal{I}|}$ be vectors of Lagrange multipliers of the equality and inequality constraints, respectively. Specifically, we write

$$z^T = [z_1^T, \dots, z_K^T], \quad (20)$$

where $z_i \in \mathbb{R}^{|\mathcal{I}|}$ is the vector of Lagrange multipliers associated with inequality constraints evaluated at p_i , for $i = 1, \dots, K$. The stationarity condition, split for the derivatives with respect to p and s , is

$$0 = \nabla_p \phi(p) + [D_p c_{\mathcal{E}}(p)]^T y + \sum_{i=1}^K [D_{p_i} g_{\mathcal{I}}(p_i)]^T z_i, \quad (21a)$$

$$0 = -(\mu \vec{1}) \oslash s + z, \quad (21b)$$

where \oslash denotes element-wise division, $\vec{1}$ is a vector of ones, and $\nabla_p(D_p)$ denotes the gradient (Jacobian) with respect to p . More explicitly, the gradient term is

$$\nabla_p \phi(p) = \left[\frac{\partial \phi(p)}{\partial p_i} \right]_{i=1}^K, \quad (22)$$

where the notation $[\cdot]_{i=1}^K$ indicates vertical concatenation of scalars/vectors/matrices indexed by i , along the ordered set $1, \dots, K$. In the same fashion, we can write the Jacobian as

$$D_p c_{\mathcal{E}}(p) = [\nabla_p^T c_i(p)]_{i \in \mathcal{E}}. \quad (23)$$

Define the vectors d_k as

$$d_k = p_k - p_{k-1}, \quad k = 1, \dots, K+1, \quad (24)$$

then we have that

$$\frac{\partial \phi(p)}{\partial p_i^T} = 2w_i d_i - 2w_{i+1} d_{i+1}, \quad i = 1, \dots, K. \quad (25)$$

We also notice that

$$D_p g_{\mathcal{I}}(p_i) = [D_{p_1} g_{\mathcal{I}}(p_i), \dots, D_{p_K} g_{\mathcal{I}}(p_i)], \quad (26a)$$

$$D_p g_{\mathcal{I}}(p_i) = [0_{|\mathcal{I}| \times 2g(i-1)}, D_{p_i} g_{\mathcal{I}}(p_i), 0_{|\mathcal{I}| \times 2g(K-i)}], \quad (26b)$$

so the stationarity condition can be expressed as

$$0 = \nabla_p \phi(p) + [D_p c_{\mathcal{E}}(p)]^T y + ([D_p g_{\mathcal{I}}(p_i)]_{i=1}^K)^T z, \quad (27a)$$

$$0 = -(\mu \vec{1}) \odot s + z. \quad (27b)$$

The stationarity condition combined with the equality constraints define a set of nonlinear equations that can be solved numerically to find a KKT point. The Lagrangian of the problem, excluding the barrier terms, is

$$L(p, s, y, z) = \phi(p) + y^T c_{\mathcal{E}}(p) + z^T ([g_{\mathcal{I}}(p_i)]_{i=1}^K + s), \quad (28)$$

so the first-order KKT conditions can be written as

$$0 = \nabla_p L(p, s, y, z), \quad (29a)$$

$$0 = s \odot z - \mu \vec{1}, \quad (29b)$$

$$0 = c_{\mathcal{E}}(p), \quad (29c)$$

$$0 = g_{\mathcal{I}}(p_i) + (s)_{(|\mathcal{I}|(i-1)+1):|\mathcal{I}|i}, \quad i = 1, \dots, K, \quad (29d)$$

where \odot denotes element-wise multiplication. Notice that (29b) implies that the entries of s and z must have the same sign, so z must also have positive entries. For brevity, we will rename some vectors and matrices as

$$\begin{aligned} D_{\mathcal{E}} &= D_p c_{\mathcal{E}}, & D_{\mathcal{I}} &= [D_p g_{\mathcal{I}}(p_i)]_{i=1}^K, \\ \Sigma &= \text{diag}(z \odot s), & c_{\mathcal{I}}(p) &= [g_{\mathcal{I}}(p_i)]_{i=1}^K. \end{aligned} \quad (30)$$

Applying Newton's method, and omitting dependencies for brevity, we obtain the following update equation:

$$\begin{bmatrix} \nabla_{pp}^2 L & 0 & D_{\mathcal{E}}^T & D_{\mathcal{I}}^T \\ 0 & \Sigma & 0 & I \\ D_{\mathcal{E}} & 0 & 0 & 0 \\ D_{\mathcal{I}} & I & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta s \\ \Delta y \\ \Delta z \end{bmatrix} = - \begin{bmatrix} \nabla_p L \\ z - (\mu \vec{1}) \odot s \\ c_{\mathcal{E}} \\ c_{\mathcal{I}}(p) + s \end{bmatrix}, \quad (31)$$

where the second row block has been left-multiplied by $\text{diag}(s)^{-1}$ to make the matrix symmetric. The rank of the Newton matrix depends directly on $\nabla_{pp}^2 L$ and $D_{\mathcal{E}}$. More specifically, if both $\nabla_{pp}^2 L$ and $D_{\mathcal{E}}$ are full rank, then the Newton matrix is invertible. We will first provide conditions under which $D_{\mathcal{E}}$ has full rank. Recall that

$$D_{\mathcal{E}} = D_p c_{\mathcal{E}} = [D_{p_1} c_{\mathcal{E}}, \dots, D_{p_K} c_{\mathcal{E}}], \quad (32a)$$

$$D_{\mathcal{E}} = \left[\frac{\partial c_j}{\partial p_i^T} \right]_{j \in \mathcal{E}}, \quad i = 1, \dots, K, \quad (32b)$$

$$\frac{\partial c_j}{\partial p_i^T} = \begin{cases} -2w_j d_j^T, & i = j-1 \\ 2w_j d_j^T + 2w_{j+1} d_{j+1}^T, & i = j \\ -2w_{j+1} d_{j+1}^T, & i = j+1 \\ 0, & \text{else} \end{cases}, \quad (32c)$$

so $D_{\mathcal{E}}$ is a $K \times 2gK$ block tridiagonal matrix, with blocks of size $1 \times 2g$. Next, we prove the claim.

Theorem 1. For any $j = 1, \dots, K+1$ define

$$b_j(p) = w_j d_j \neq 0, \quad (33)$$

and assume that $b_j(p) \neq 0$ for all $j = 1, \dots, K+1$ (this is true if and only if $d_j \neq 0$). Let q_j be

$$q_j(p) = b_j / b_j^T b_j, \quad \forall j = 1, \dots, K+1. \quad (34)$$

If $\sum_{j=1}^{K+1} q_j(p) \neq 0$, then $D_{\mathcal{E}}$ is full rank.

Proof. See Appendix D. \square

From this result, we can guarantee that $D_{\mathcal{E}}$ is full rank as long as we prevent any d_j from becoming 0. We also need to safeguard the algorithm against cases where $\sum_{j=1}^{K+1} q_k = 0$. This condition is a vector generalization of the condition of impedance loops not adding to zero in order to guarantee the invertibility of the admittance matrix in transmission systems (see [18]). We propose a simple step rejection procedure as a safeguard; this is detailed in Appendix B.

The Lagrangian Hessian, $\nabla_{pp}^2 L$, may not be invertible, but it is a very structured matrix. In Appendix A, we show that $\nabla_{pp}^2 L$, $\nabla_{pp}^2 \phi$ and $\nabla_{pp}^2 (y^T c_{\mathcal{E}})$ are symmetric and block tridiagonal, and $\nabla_{pp}^2 (L - \phi - y^T c_{\mathcal{E}})$ is block diagonal (with block sizes $2g \times 2g$ for all matrices). Invertibility and other issues (like indefiniteness) can be easily corrected by leveraging the block structure of $\nabla_{pp}^2 L$ and its components, as shown in the next subsection.

B. Reduced Newton Step

The Newton step computation requires solving the linear system (31), which has size $2K(g + |\mathcal{I}| + 1)$, so a matrix factorization requires $O(K^3(g + |\mathcal{I}|)^3)$ operations. The Jacobians $D_p g_j$ and the Hessians $\nabla_{pp}^2 g_j$ are usually dense for constraints in the state vector x . Therefore, the Newton matrix is relatively sparse, but with dense blocks. The total number of non-zero entries is of the order of $O(Kg(g + |\mathcal{I}|))$. This means that solving the linear system (31) with an iterative method would require $O(K^2g(g + |\mathcal{I}|)^2)$ operations. We will show that we can reduce the operation cost even further by reducing the size of the linear system. Specifically, we will show that the Newton step can be written in terms of a block tridiagonal matrix, which reduces the cost of computing the solution to $O(Kg^3)$. The first and second derivatives of the OPF constraints can be computed in $O(K(g^3 + n^3) + |\mathcal{I}|n^2)$ time complexity, so the overall time complexity remains linear in K . Refer to Appendix C for a detailed description on computing the OPF constraints' derivatives. We start by removing Δs from (31) by substituting the second row block into the fourth one (recall that $z, s > 0$, so $\Sigma \succ 0$), yielding

$$\begin{bmatrix} \nabla_{pp}^2 L & D_{\mathcal{E}}^T & D_{\mathcal{I}}^T \\ D_{\mathcal{E}} & 0 & 0 \\ D_{\mathcal{I}} & 0 & -\Sigma^{-1} \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta y \\ \Delta z \end{bmatrix} = - \begin{bmatrix} \nabla_p L \\ c_{\mathcal{E}} \\ c_{\mathcal{I}}(p) + (\mu \vec{1}) \odot z \end{bmatrix}. \quad (35)$$

Substituting the third row block in the first removes Δz :

$$\begin{bmatrix} \nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}} & D_{\mathcal{E}}^T \\ D_{\mathcal{E}} & 0 \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta y \end{bmatrix} =$$

$$- \left[\begin{array}{c} \nabla_p L + D_{\mathcal{I}}^T (\Sigma c_{\mathcal{I}}(p) + (\mu \vec{1}) \odot s) \\ c_{\mathcal{E}} \end{array} \right]. \quad (36)$$

The removed steps can be recovered as

$$\Delta z = \Sigma(D_{\mathcal{I}} \Delta p + c_{\mathcal{I}}(p) + (\mu \vec{1}) \odot z), \quad (37a)$$

$$\Delta s = \Sigma^{-1}((\mu \vec{1}) \odot s - z - \Delta z). \quad (37b)$$

To ensure that the Newton step in the primal variables, Δp , yields a descent direction, we require $\nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}}$ to be positive definite in the tangent space of the equality constraints. More formally, let Z be a null space matrix of $D_{\mathcal{E}}$, then we require that $Z^T (\nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}}) Z \succ 0$. The simplest way to satisfy the condition is to modify $\nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}}$ to make it positive definite. To this end, notice that

$$\nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}} = \nabla_{pp}^2 \phi + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}} + \nabla_{pp}^2 (L - \phi). \quad (38)$$

We already proved that $\nabla_{pp}^2 \phi \succ 0$, so any source of indefiniteness must come from the Lagrangian terms of the inequalities, $L - \phi$. We know that $\nabla_{pp}^2 L$ and $\nabla_{pp}^2 \phi$ are block tridiagonal, so $\nabla_{pp}^2 (L - \phi)$ is block tridiagonal as well. We modify the Hessian by adding a matrix of the form

$$S = \begin{bmatrix} \delta_1 I & & \\ & \ddots & \\ & & \delta_K I \end{bmatrix}, \quad \mathbb{R} \ni \delta_i \geq 0, \quad i = 1, \dots, K. \quad (39)$$

A strategy for selecting values of δ_i is discussed in Appendix B1. Define ψ and the block diagonal matrix Γ as

$$\psi = \phi + y^T c_{\mathcal{E}}, \quad (40a)$$

$$\Gamma = D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}} + \nabla_{pp}^2 (L - \psi) + S = \begin{bmatrix} \Gamma_1 & & \\ & \ddots & \\ & & \Gamma_K \end{bmatrix}. \quad (40b)$$

Then the modified linear system is

$$\begin{bmatrix} \nabla_{pp}^2 \psi + \Gamma & D_{\mathcal{E}}^T \\ D_{\mathcal{E}} & 0 \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta y \end{bmatrix} = - \left[\begin{array}{c} \nabla_p L + D_{\mathcal{I}}^T (\Sigma c_{\mathcal{I}}(p) + (\mu \vec{1}) \odot s) \\ c_{\mathcal{E}} \end{array} \right]. \quad (41)$$

Note that Γ is dense and block diagonal. Both $\nabla_{pp}^2 \phi$ and $D_{\mathcal{E}}$ are block tridiagonal. We can permute the rows and columns of the reduced Newton matrix to get a block tridiagonal matrix. Denote the permutation by the matrix P . We then have

$$P \begin{bmatrix} \nabla_{pp}^2 \psi + \Gamma & D_{\mathcal{E}}^T \\ D_{\mathcal{E}} & 0 \end{bmatrix} P^T P \begin{bmatrix} \Delta p \\ \Delta y \end{bmatrix} = - P \left[\begin{array}{c} \nabla_p L + D_{\mathcal{I}}^T (\Sigma c_{\mathcal{I}}(p) + (\mu \vec{1}) \odot s) \\ c_{\mathcal{E}} \end{array} \right], \quad (42)$$

where P is such that

$$P \begin{bmatrix} \Delta p \\ \Delta y \end{bmatrix} = \begin{bmatrix} (\Delta p)_{(2g(i-1)+1):2gi} \\ (\Delta y)_i \end{bmatrix}_{i=1}^K. \quad (43)$$

Define the following matrices:

$$\Phi = \begin{bmatrix} \Phi_1 & & \\ & \ddots & \\ & & \Phi_{K+1} \end{bmatrix}, \quad (44a)$$

$$\Phi_k = 2w_k \begin{bmatrix} (1 + (y)_k - (y)_{k-1})I & d_k \\ d_k^T & 0 \end{bmatrix}. \quad (44b)$$

We then have that

$$P \begin{bmatrix} \nabla_{pp}^2 \psi + \Gamma & D_{\mathcal{E}}^T \\ D_{\mathcal{E}} & 0 \end{bmatrix} P^T = \begin{bmatrix} \begin{bmatrix} \Gamma_1 & 0 \\ 0 & 0_{1 \times 1} \end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix} \Gamma_K & 0 \\ 0 & 0_{1 \times 1} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \Phi_1 + \Phi_2 & -\Phi_2 & & \\ -\Phi_2 & \ddots & \ddots & \\ & \ddots & \ddots & -\Phi_K \\ & & -\Phi_K & \Phi_K + \Phi_{K+1} \end{bmatrix}. \quad (45)$$

With this permutation, we obtain a $(2g+1)K \times (2g+1)K$ block tridiagonal system with blocks of size $(2g+1) \times (2g+1)$. Thus, computing the solution costs $O(Kg^3)$, as previously claimed.

C. Newton Iteration Algorithm

Thus far, we have detailed a procedure for computing the Newton step in an interior point iteration for solving (16). However, a robust implementation must also incorporate safeguards for issues related to strong non-linearity, indefiniteness, strict positivity of dual variables, and scale disparity between primal and dual variables. We discuss these issues and their solutions in Appendix B. Once a complete Newton iteration for the interior point method is implemented, we can solve the barrier problem for a fixed barrier parameter μ , as long as we are provided an initial feasible path. Pseudo-code of the procedure given an initial feasible path p is described in Appendix B4.

IV. INITIAL FEASIBLE PATH GENERATION

The last missing part of the full algorithm is a procedure for generating an initial feasible path. In the unconstrained case, the straight line connecting u_0 to u_1 is a feasible path (and, in fact, the shortest one). To include the effect of constraints, we introduce an homotopy-like procedure: we start with a relaxed version of the problem where the straight line is feasible and then we solve increasingly tighter relaxations until the original problem is recovered. A way to interpret this procedure is to consider the constraints as continuously pushing and deforming the straight line until a curved feasible path is obtained. If the original problem is infeasible (u_0 and u_1 lie in different connected components of the feasible region), then at some point of the homotopy some constraints will try to cut the path to get each piece to a different connected component. If the path's corners are too close, such a transformation of the path would violate the constant speed constraint (16b) and the homotopy would fail (see Fig. 3).

We next formally describe the path generation procedure. First, we notice that the power flow feasibility constraint ($g_i, i \in \mathcal{P}$, see (5)) is a special case as it is not differentiable on its boundary. This means that there exists no differentiable relaxation of it. Nevertheless, the power flow feasible region (i.e., the set of power injections for which a power flow solution exists) is typically much larger than the OPF constraints'

feasible region, so we can thus assume that the straight line does not violate the power flow feasibility constraint:

- **Assumption 4:** The straight line joining u_0 and u_1 is contained in the power flow feasibility set \mathcal{F} .

Under Assumption 4, we do not need to include the power flow feasibility constraint in the homotopy process. The homotopy procedure for addressing the remaining constraints is relatively simple. Assume that the user provides a path spacing $\{t_k\}_{k=0}^{K+1}$ satisfying (12). Let p be the current candidate path. At the start of the procedure, p is a straight line, so its corners are

$$p_k = u_0 + t_k(u_1 - u_0), \quad i = 0, \dots, K+1. \quad (46)$$

Next we compute the vector of relaxation parameters v , as the vector of maximum violations of each constraint across all path corners multiplied by a margin $\beta > 1$:

$$(v)_j = \beta \cdot \max_{i=1, \dots, K} (\max \{g_j(p_i), 0\}), \quad j \in \mathcal{I}. \quad (47)$$

Note that the power flow feasibility constraint is never positive, so its corresponding entry on v is always 0. The vector of relaxed constraints, g_v , is

$$g_v(u) = g_{\mathcal{I}}(u) - v. \quad (48)$$

Clearly the path p is contained in the relaxed feasible set defined by g_v . More formally:

$$g_v(p_i) < 0, \quad i = 1, \dots, K. \quad (49)$$

If we choose β close to (but still greater than) 1, then the boundary of each violated constraint's relaxation will be very close to some corner of p . We leverage this situation by calling the interior point solver, whose iterations will naturally push the path towards the interior of the (relaxed) feasible region. By using a large barrier parameter μ_{hi} , we can obtain a new path that will not be close to any boundary of the relaxed constraint vector g_v , allowing us to reduce the relaxation parameters (the entries of v). Thus, we just need to recompute v and repeat this process until v is close enough to 0, indicating that the corner points of the path satisfy the original (non-relaxed) constraints. If this process stagnates for any reason (entries of v stop decreasing), we report failure under suspicion that a feasible path may not exist (see Fig. 3). Pseudo-code of the complete shortest path algorithm, including the generation of a feasible path, is given by Algorithm 1. Upon finding a feasible path, we compute the *shortest* path by calling the interior point solver with a small barrier parameter μ .

Some OPF cases have inequalities that are so close that they roughly behave like equalities, making the feasible region nearly a lower-dimensional manifold with no interior. In such cases, the interior point algorithm may present convergence difficulties or even fail completely. As a safeguard against these issues, the last solver call uses the relaxed constraints g_v with a small relaxation vector $v = \epsilon_{ls}\bar{1}$. This slightly increases the size of the feasible region's interior, so that the solver has enough "space" in the feasible set to move the candidate path towards the solution. We may also have situations where the relative decrease of the violations is less than β , so v ends up increasing slightly after each iteration. This issue is prevented by taking the minimum (entry-wise) between v and the previous iteration vector v^- (see Step 7 of Algorithm 1).

Algorithm 1 Shortest Path Algorithm (Outer Loop)

```

1: procedure SHORTESTPATH( $f, g_{\mathcal{I}}, \{t_k\}_{k=0}^{K+1}, \beta, \mu_{hi}, \epsilon_{st},$ 
    $\mu_{lo}, \epsilon_{tol}, \text{iter}_{\max}, \tau, \gamma, \eta, \epsilon_{ls}, \rho_{\max}$ )
2:   compute  $p$  from (46) and compute  $v$  from (47)
3:   while  $\|v\|_{\infty} > \epsilon_{ls}$  do
4:     compute  $g_v$  from (48) and assign  $p^- \leftarrow p$ 
5:      $p, s \leftarrow \text{BARRIER\_SOLVE}(f, g_v, p^-, \mu_{hi}, \dots)$ 
6:     assign  $v^- \leftarrow v$  and compute  $v$  from (47)
7:      $v \leftarrow \min\{v, v^-\}$ 
8:     if  $\|v - v^-\|_{\infty} \leq \epsilon_{st}$  and  $\|v\|_{\infty} > \epsilon_{ls}$  then
9:       report failure and break
10:    if  $\|v\|_{\infty} \leq \epsilon_{ls}$  then
11:      assign  $v \leftarrow \epsilon_{ls}\bar{1}$  and compute  $g_v$  from (48)
12:       $p, s \leftarrow \text{BARRIER\_SOLVE}(f, g_v, p, \mu_{lo}, \dots)$ 
13:    return  $p, \|v\|_{\infty}$ 

```

V. NUMERICAL EXPERIMENTS

This section describes experiments performed to assess the performance of the proposed algorithm. We provide a public implementation of the algorithm, illustrative examples, and experiments on power systems of different scales.

A. Implementation

We developed a Julia code that implements the shortest path algorithm. The code is publicly available at the following page: github.com/djturizo/Shortest-Path-OPF

All experiments were run using Julia 1.10 on a Windows 11 PC with 32GB of RAM and an AMD Ryzen™ PRO 7840U CPU. Unless specified otherwise, we used the following parameters:

$$\begin{aligned}
K &= 19, & t_k &= 0.05 \cdot k, & k &= 0, \dots, K+1, \\
\beta &= 1.01, & \mu_{hi} &= 10^{-1}, & \epsilon_{st} &= 10^{-3}, \\
\mu_{lo} &= 10^{-6}, & \epsilon_{st} &= 10^{-3}, & \text{iter}_{\max} &= 100, \\
\tau &= 0.99, & \gamma &= 0.5, & \eta &= 10^{-4}, \\
\epsilon_{ls} &= 10^{-6}, & \rho_{\max} &= 100.0.
\end{aligned}$$

The power flow equations were solved using the Newton-Raphson method with a tolerance of 10^{-8} and a limit of 20 iterations (see Step 9 of Algorithm 2). The shortest path algorithm uses a network model with one generator per node at most and rectangular coordinates for the voltage phasors, in order to have quadratic power flow equations and constraints (except for line flow constraints). Some test cases have multiple generators in a single node, but it is possible to compute a single equivalent generator. Angle difference constraints can be written as quadratic inequalities whenever the corresponding angle limit lies in the interval $(-\pi/2, \pi/2)$ (see [19]), which is the case in practice.

During the execution of the experiments, we noticed that evaluating the power flow feasibility constraint (5) took a significant portion of the execution time, but it was never active. This is consistent with the expectation that the boundary of the power flow feasibility constraint is significantly larger than that of all other constraints, so the feasible set ends up being determined by the standard OPF constraints. This means that

the power flow feasibility constraint has no effect at all on the results of the shortest path algorithm (and we confirmed this on the experiments). We thus ignored this constraint in our experiments to increase the execution speed of the algorithm.

B. Example: Two Variants of the 9-Bus Case

To illustrate how the algorithm works in different situations we used the 9-bus OPF case of MATPOWER [20]. The system has three generators, at nodes 1 to 3, with node 1 being the slack node. The control variables are the voltage magnitudes of the generators (V_1, V_2, V_3) and the active power of non-slack generators (P_{G2}, P_{G3}). We consider two variants of the 9-bus case obtained by modifying the system parameters. The first one, called variant 1 from now on, is modified to introduce an obstacle in the feasible region. First we set the generator voltage magnitudes to be 1 p.u. ($V_1 = V_2 = V_3 = 1$). The control vector in the subspace is chosen as $u = [P_{G2}, P_{G3}]^T$. We generate the obstacle by setting the lower reactive power limit of the generator at bus 3 to -2 MVA ($Q_{G3\min} = -0.02$). For the endpoints we choose $u_0 = [0.5, 0.5]^T$ and $u_1 = [1.5, 1.3]^T$.

We executed the shortest path algorithm, obtaining the results illustrated in Fig. 2. The feasible region is colored in green, and the relaxations generated by the algorithm are colored in red hues. Later iterations have smaller constraint violations, which lead to tighter relaxations, represented with darker shades of red. The shortest path is computed for each relaxation. Paths corresponding to tighter relaxations are colored with lighter shades of blue for contrast. The figure shows the continuous deformation of the path as it moves away from the boundary. After multiple iterations of this process, the algorithm obtains a feasible path, and then the final iteration tightens the candidate path while preserving feasibility.

We next consider another modification, called variant 2 from now on, where no feasible path exists. For this variant we used the 9-bus OPF case of MATPOWER [20], modified as in [6]. We also fix the generator voltage magnitudes to the following p.u. values: $V_1 = 0.920, V_2 = 0.935, V_3 = 0.943$. The control vector in the subspace is chosen as $u = [P_{G3}, P_{G2}]^T$. The endpoints, are chosen to be from different connected regions. Namely, we chose $u_0 = [0.12, 0.16]^T$ and $u_1 = [1.57, 0.24]^T$.

We executed the shortest path algorithm, obtaining the results illustrated in Fig. 3. The figure shows how tighter relaxations become narrower around the center in an attempt to eventually break into two components. As a result, the candidate path ends up “choked” in this narrow passage, which attempts stretch the path, separating the corner points into two distant clusters. Such a deformation would violate the constant speed constraints that require the corner points to preserve the relative distance between them. As a result, the algorithm is unable to reduce the constraint violations any further, and it appropriately reports failure to find a feasible path.

As a last experiment for this case, we modified the value of μ_{lo} to observe its effect on the computation of the shortest path from a given feasible path (step 12 of Algorithm 1). For this experiment, we consider the variant 1 of the 9-bus case and we solve the shortest path problem for multiple values of μ_{lo} in the range $[10^{-11}, 10^{-1}]$. For each value of μ_{lo} , we

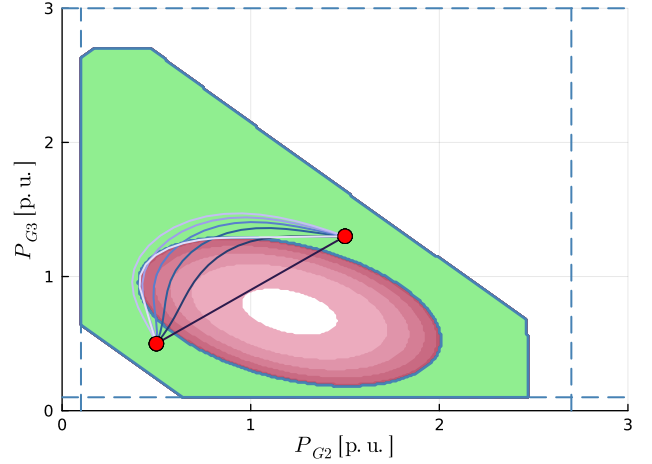


Fig. 2. Variant 1 of the 9-bus case. The straight line path is not feasible, but the algorithm deforms the path to achieve feasibility.

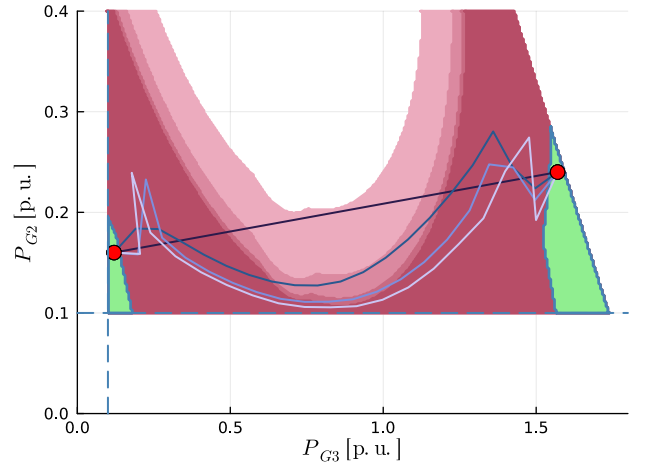


Fig. 3. Variant 2 of the 9-bus case. The endpoints are disconnected, so the algorithm fails to find a feasible path.

compute the length of the shortest path found as the percentage increase over the path length of the unconstrained solution (i.e., the straight line joining the endpoints). As shown in Fig. 4, the feasible path generated by the homotopy process may be significantly larger than the shortest path, warranting the last optimization process that is executed with a lower barrier parameter. For small values of μ_{lo} , observe that the solution does not change until μ_{lo} becomes small enough that the non-linearity of the barrier function introduces numerical artifacts. This means that the value of μ_{lo} must be chosen as small as possible with risking numerical issues.

C. Multiple scale OPF cases

For this experiment, we used multiple OPF benchmark test cases from the Power Grid Library PGLib [21]. We selected nine cases of different sizes, ranging from 14 to 118 buses. Since these cases have high-dimensional feasible spaces that are hard to visualize, selecting non-trivial endpoints (where the straight line is not feasible) is not always straightforward. We therefore follow the heuristic presented in [9]: namely,

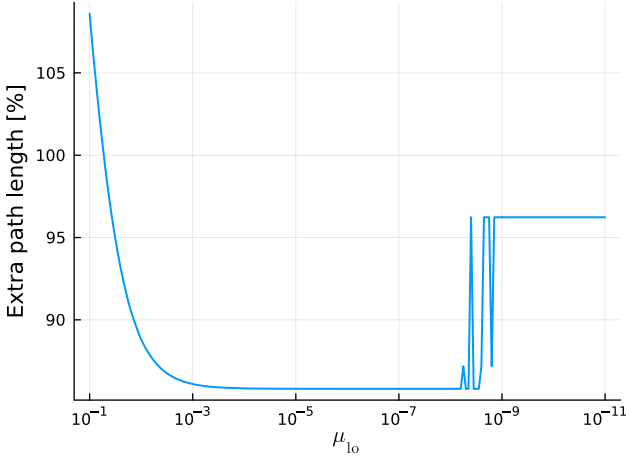


Fig. 4. Variant 1 of the 9-bus case. For smaller barrier parameters, the path length decreases until stabilizing at the shortest path. However, very small barrier parameters introduce numerical artifacts.

we selected the endpoints as the solution of the minimum loss problem and the OPF solution. For each test case, we computed the maximum constraint violation over the path points in the starting straight line (i.e., before running the algorithm) and over the final path resulting from running the algorithm, ignoring the endpoints (because they are fixed and not modified by the algorithm). If the maximum constraint violation after running the algorithm is negative, then the final path found is feasible and the algorithm has thus identified a shortest path (in a local sense, at least).

We also computed solution and objective function metrics as follows: let $p(t)$ be the piece-wise linear shortest path approximation (as defined in (11)) resulting from the algorithm, and let $L(t)$ be the straight line path associated with the endpoints. With both p and L parameterized by arclength, we computed the relative difference between the paths as

$$\text{path-diff\%} = \frac{\int_0^1 \|p(t) - L(t)\| dt}{\int_0^1 \|L(t)\| dt} \times 100\%.$$

Similarly, we computed the relative objective function increase, or gap, with respect to the value at the straight line:

$$\text{obj-fun-gap\%} = \frac{\int_0^1 \|p(t)\| dt - \int_0^1 \|L(t)\| dt}{\int_0^1 \|L(t)\| dt} \times 100\%.$$

The results are reported in Table I. The algorithm succeeded in finding a locally shortest path on all test cases. In many test cases, the straight line is slightly infeasible, with one notable exception being the 60-bus case where the straight line has violations as large as is 2.22 p.u. In the 14- and 30-bus cases, the straight line is feasible, but very close to the boundary of infeasibility. In such cases, the continuous nature of the barrier function will make the algorithm deform the straight line to move it further away from the boundary. This is an unintended but desirable behavior of the algorithm, as it introduces a safety margin around the path. The magnitude of this safety margin depends on the sensitivity of the constraints around the boundary. For example, in the 30-bus case, a change no larger than $\sim 10^{-8}$ p.u. per constraint is required to satisfy

the algorithm tolerance, yet this change introduces relative a difference of $\sim 4\%$ between the shortest path and the straight line. On the other hand, in the 60-bus case, changes as large as 2.22 p.u. are needed to correct the constraint violations, yet the relative difference between the shortest path and the straight line is only $\sim 1\%$.

We also executed the algorithm on eight test cases selected from [22], which were crafted specifically to be challenging for OPF solvers. The results for this second batch of test cases are shown in Table II. As expected, these test cases proved to be more challenging, as in three of the eight cases the algorithm failed to find a feasible path. We remark that it is possible that the endpoints of those cases are not connected, but the algorithm may also fail even if a feasible path exists (after all, this is a non-convex optimization problem). For the five remaining cases the straight line was already feasible in two of them, and for the other three the algorithm succeeded in generating a locally shortest path. We observed that the relative path differences are usually larger than in the PGLib test cases, and we suspect this tendency is due to more pronounced non-convexities resulting from the fact that these test cases have been engineered to challenge OPF solvers.

VI. CONCLUSIONS

In this paper, we developed an algorithm for computing a discretized shortest path from an initial feasible operating point to an optimal one (or between any two feasible points in general), thus minimizing the amplitude of the control actions required to transition from one point to another. The discretized shortest path is represented as a sequence of intermediate feasible points, the number and relative spacing of which can be specified a priori. The algorithm computes the intermediate points by solving a nonlinear optimization problem via a specialized interior point method, provided an initial feasible path is given. By leveraging the nature of barrier functions in interior point methods, an initial feasible path is found by solving a sequence of relaxed, but increasingly tighter relaxations of the shortest path problem, where in the initial relaxation the straight line joining the endpoints is feasible. The resulting sequence of shortest paths converges to a feasible path of the original problem in a finite number of iterations. The interior point solver for the algorithm was modified to exploit the special block tridiagonal structure of the shortest path problem. Multiple numerical experiments show that the proposed algorithm can effectively compute a shortest path for a specified number of intermediate points.

The algorithm we developed tackles the issues of amount and amplitude of control actions in the problem of transitioning between operating points. One issue not considered in this work is feasibility of the path in the continuous sense. While our algorithm provides a sequence of intermediate points that are guaranteed to be feasible, the lines joining them may cross the boundary of the feasible set. An avenue of future work consists on extending the current algorithm with a methodology to provide mathematical guarantees that the line pieces comprising the discrete path are entirely contained in the feasible set.

TABLE I
RESULTS OF RUNNING THE SHORTEST PATH ALGORITHM ON PGLIB TEST CASES

Test case	n	g	Max. con. bef. [p.u.]	Exec. time [s]	Found path?	Max. con. aft. [p.u.]	Path diff.	Obj. fun. gap
case9 (variant 1)	9	2	2.79E-2	4.9	Yes	-6.92E-8	85.8%	34.8%
case14_ieee	14	5	-9.92E-7	2.4	Yes	-1.00E-6	0.88%	0.01%
case24_ieee_rts	24	11	9.92E-4	12.7	Yes	-1.00E-6	0.24%	0.00%
case30_ieee	30	6	-9.91E-7	2.5	Yes	-1.00E-6	4.03%	0.12%
case39_epri	39	10	9.66E-2	17.9	Yes	-5.81E-5	0.72%	0.00%
case57_ieee	57	7	2.50E-3	4.8	Yes	-1.00E-6	0.18%	0.00%
case60_c	60	23	2.22E+0	57.3	Yes	-1.00E-6	1.08%	0.01%
case73_ieee_rts	73	33	9.59E-4	75.7	Yes	-1.00E-6	0.18%	0.00%
case118_ieee	118	54	2.42E-2	605.8	Yes	-1.00E-6	0.23%	0.00%

TABLE II
RESULTS OF RUNNING THE SHORTEST PATH ALGORITHM ON THE TEST CASES OF [22]

Test case	n	g	Max. con. bef. [p.u.]	Exec. time [s]	Found path?	Max. con. aft. [p.u.]	Path diff.	Obj. fun. gap
nmwc3acyclic_connected	3	2	-1.84E-2	2.2	Yes	-1.84E-2	0.58%	0.00%
nmwc3acyclic_disconnected	3	2	1.96E-5	3.8	Yes	-2.77E-4	11.6%	0.99%
nmwc3cyclic	3	2	9.28E-3	2.3	No	8.76E-3	—	—
nmwc4	4	2	-1.09E-4	2.2	Yes	-7.58E-4	10.7%	0.92%
nmwc5	5	2	2.34E-2	14.2	Yes	-1.91E-4	2.18%	0.03%
nmwc14	14	5	9.26E-4	3.7	No	5.57E-4	—	—
nmwc24	24	11	3.71E-3	43.8	Yes	-9.97E-7	0.80%	0.00%
nmwc57	57	7	2.89E-3	10.6	No	2.37E-3	—	—

REFERENCES

- [1] M. L. Crow, *Computational Methods for Electric Power Systems*, 3rd ed. CRC Press, 2015.
- [2] F. Capitanescu and L. Wehenkel, "Optimal power flow computations with a limited number of controls allowed to move," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 586–587, 2010.
- [3] —, "Redispatching active and reactive powers using a limited number of control actions," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1221–1230, 2011.
- [4] D. T. Phan and X. A. Sun, "Minimal impact corrective actions in security-constrained optimal power flow via sparsity regularization," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1947–1956, 2015.
- [5] D. Lee, H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Convex restriction of power flow feasibility sets," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1235–1245, 2019.
- [6] D. K. Molzahn, "Computing the feasible spaces of optimal power flow problems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4752–4763, 2017.
- [7] F. Capitanescu, "Suppressing ineffective control actions in optimal power flow problems," *IET Generation, Transmission & Distribution*, vol. 14, no. 13, pp. 2520–2527, 2020. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-gtd.2019.1783>
- [8] I.-I. Avramidis, G. Cheimonidis, and P. Georgilakis, "Ineffective control actions in opf problems: Identification, suppression and security aspects," *Electric Power Systems Research*, vol. 212, p. 108228, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378779622004369>
- [9] D. Lee, K. Turitsyn, D. K. Molzahn, and L. A. Roald, "Feasible Path Identification in Optimal Power Flow with Sequential Convex Restriction," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3648–3659, September 2020.
- [10] R. Martins Barros, G. Guimarães Lage, and R. de Andrade Lira Rabêlo, "Sequencing paths of optimal control adjustments determined by the optimal reactive dispatch via lagrange multiplier sensitivity analysis," *European Journal of Operational Research*, vol. 301, no. 1, pp. 373–385, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221721009310>
- [11] J.-O. Lee and Y.-S. Kim, "Novel battery degradation cost formulation for optimal scheduling of battery energy storage systems," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107795, 2022.
- [12] B. Ghaddar, J. Marecek, and M. Mevissen, "Optimal power flow as a polynomial optimization problem," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 539–546, 2016.
- [13] C. J. Tavora and O. J. M. Smith, "Equilibrium analysis of power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 3, pp. 1131–1137, 1972.
- [14] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, 1st ed. Academic Press, Inc., 1970.
- [15] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, 1st ed. Springer, 1997.
- [16] Z. Clawson, A. Chacon, and A. Vladimirovsky, "Causal domain restriction for eikonal equations," *SIAM Journal on Scientific Computing*, vol. 36, no. 5, pp. A2478–A2505, 2014.
- [17] C. Heil, *A Basis Theory Primer*, 1st ed. Springer, 2011.
- [18] D. Turizo and D. K. Molzahn, "Invertibility conditions for the admittance matrices of balanced power systems," *IEEE Transactions on Power Systems*, vol. 38, no. 4, pp. 3841–3853, 2023.
- [19] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "The qc relaxation: A theoretical and computational study on optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3008–3018, 2016.
- [20] R. D. Zimmerman and C. E. Murillo-Sánchez, "Matpower user's manual," 2020. [Online]. Available: <https://matpower.org/docs/MATPOWER-manual-7.1.pdf>
- [21] IEEE PES Task Force on Benchmarks for Validation of Emerging Power System Algorithms, "The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms," *arXiv:1908.02788v2*, Jan. 2021.
- [22] M. R. Narimani, D. K. Molzahn, D. Wu, and M. L. Crow, "Empirical Investigation of Non-Convexities in Optimal Power Flow Problems," *American Control Conference (ACC)*, June 2018.
- [23] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000, vol. 71.
- [24] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Review*, vol. 44, no. 4, pp. 525–597, 2002.
- [25] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, Mar 2006.
- [26] J. R. Bunch and L. Kaufman, "Some stable methods for calculating inertia and solving symmetric linear systems," *Mathematics of computation*, vol. 31, no. 137, pp. 163–179, 1977.
- [27] C. Ashcraft, R. G. Grimes, and J. G. Lewis, "Accurate symmetric indefinite linear equation solvers," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 513–561, 1998.
- [28] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2013.
- [29] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. Springer, 2006.
- [30] R. H. Byrd, J. C. Gilbert, and J. Nocedal, "A trust region method based

on interior point techniques for nonlinear programming,” *Mathematical Programming*, vol. 89, no. 1, pp. 149–185, Nov 2000.

- [31] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed. Wiley, 2007.
- [32] D. Kulkarni, D. Schmidt, and S.-K. Tsui, “Eigenvalues of tridiagonal pseudo-Toeplitz matrices,” *Linear Algebra and its Applications*, vol. 297, no. 1, pp. 63–80, 1999.

APPENDICES

Appendix A: Structure of the Lagrangian Hessian

The Lagrangian of the Hessian, $\nabla_{pp}^2 L$, has components associated to the objective function and the constraints. Each of these components has a specific matrix structure the we analyze next. First we compute $\nabla_{pp}^2 L$ using the independence of s on p :

$$\nabla_{pp}^2 L = \nabla_{pp}^2 \phi + \nabla_{pp}^2 (y^T c_{\mathcal{E}}) + \nabla_{pp}^2 (z^T ([g_{\mathcal{I}}(p_i)]_{i=1}^K + s)), \quad (50a)$$

$$\nabla_{pp}^2 L = \nabla_{pp}^2 \phi + \sum_{j \in \mathcal{E}} (y)_j \nabla_{pp}^2 c_j + \sum_{i=1}^K \sum_{j \in \mathcal{I}} (z_i)_j \nabla_{pp}^2 g_j(p_i). \quad (50b)$$

The corresponding Hessian for any function f is given by

$$\nabla_{pp}^2 f = \begin{bmatrix} \frac{\partial f}{\partial p_1 \partial p_1^T} & \cdots & \frac{\partial f}{\partial p_1 \partial p_K^T} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial p_K \partial p_1^T} & \cdots & \frac{\partial f}{\partial p_K \partial p_K^T} \end{bmatrix}. \quad (51)$$

First we analyze the Hessian of the objective function, $\nabla_{pp}^2 \phi$. Let $I \in \mathbb{R}^{2g \times 2g}$ be the identity matrix. The derivatives of ϕ are

$$\frac{\partial^2 \phi}{\partial p_i \partial p_i^T} = 2(w_i + w_{i+1})I, \quad i = 1, \dots, K, \quad (52a)$$

$$\frac{\partial^2 \phi}{\partial p_{i-1} \partial p_i^T} = \frac{\partial^2 \phi}{\partial p_i \partial p_{i-1}} = -2w_i I, \quad i = 2, \dots, K, \quad (52b)$$

$$\frac{\partial^2 \phi}{\partial p_i \partial p_j} = 0, \quad |i - j| > 1, \quad (52c)$$

so $\nabla_{pp}^2 \phi$ is a constant, symmetric, and block tridiagonal matrix with symmetric blocks of size $2g \times 2g$. We can write $\nabla_{pp}^2 \phi$ as

$$\nabla_{pp}^2 \phi = 2Y \otimes I, \quad (53a)$$

$$Y = \begin{bmatrix} w_1 + w_2 & -w_2 & & & \\ -w_2 & \ddots & \ddots & & \\ & \ddots & \ddots & -w_K & \\ & & -w_K & w_K + w_{K+1} & \end{bmatrix}. \quad (53b)$$

The matrix Y can be seen as the admittance matrix of a single loop circuit with line positive resistances given by $w_k, k = 2, \dots, K$, and two shunts corresponding to w_1 and w_{K+1} . This admittance matrix is guaranteed to be invertible [18]. In particular, Y can be factored as in [18] to show that Y is positive definite. Moreover, the eigenvalues of the Kronecker product correspond to all pairwise products between eigenvalues of the two factors (see exercise 7.8.11 (b) of [23]), so $\nabla_{pp}^2 \phi$ is positive definite.

Next we consider the equality term $\nabla_{pp}^2 \phi + \nabla_{pp}^2 (y^T c_{\mathcal{E}})$. We compute the derivative terms of $c_j, j \in \mathcal{E}$ as

$$\frac{\partial^2 c_j}{\partial p_j \partial p_j^T} = 2(w_j - w_{j+1})I, \quad (54a)$$

$$\frac{\partial^2 c_j}{\partial p_{j-1} \partial p_{j-1}^T} = 2w_j I, \quad (54b)$$

$$\frac{\partial^2 c_j}{\partial p_{j+1} \partial p_{j+1}^T} = -2w_{j+1} I, \quad (54c)$$

$$\frac{\partial^2 c_j}{\partial p_j \partial p_{j-1}^T} = \frac{\partial^2 c_j}{\partial p_{j-1} \partial p_j^T} = -2w_j I, \quad (54d)$$

$$\frac{\partial^2 c_j}{\partial p_j \partial p_{j+1}^T} = \frac{\partial^2 c_j}{\partial p_{j+1} \partial p_j^T} = 2w_{j+1} I, \quad (54e)$$

$$\frac{\partial^2 c_j}{\partial p_i \partial p_k} = 0, \quad \text{any other case.} \quad (54f)$$

With a slight abuse of notation, we define

$$w_0 = (y)_0 = (y)_{K+1} = 0 \in \mathbb{R}. \quad (55)$$

Notice that

$$\frac{\partial^2 (y^T c_{\mathcal{E}})}{\partial p_i \partial p_j^T} = \sum_{j \in \mathcal{E}} (y)_j \frac{\partial^2 c_j}{\partial p_i \partial p_j^T}. \quad (56)$$

Therefore, we can write, for $j = 1, \dots, K$, that

$$\frac{\partial^2 (y^T c_{\mathcal{E}})}{\partial p_{j-1} \partial p_j^T} = \frac{\partial^2 (y^T c_{\mathcal{E}})}{\partial p_j \partial p_{j-1}} = -2w_j ((y)_j - (y)_{j-1})I, \quad (57a)$$

$$\frac{\partial^2 (y^T c_{\mathcal{E}})}{\partial p_j \partial p_j^T} = 2[w_j ((y)_j - (y)_{j-1}) + w_{j+1} ((y)_{j+1} - (y)_j)]I, \quad (57b)$$

$$\frac{\partial^2 (y^T c_{\mathcal{E}})}{\partial p_i \partial p_j} = 0, \quad |i - j| > 1, \quad (57c)$$

so the matrix $\nabla_{pp}^2 (y^T c_{\mathcal{E}})$ is symmetric and block tridiagonal (with symmetric blocks of size $2g \times 2g$).

Lastly we consider the inequality term of the Lagrangian Hessian, $\nabla_{pp}^2 (z^T ([g_{\mathcal{I}}(p_i)]_{i=1}^K + s))$. As every inequality constraint depends only on one specific p_i , it is easy to see that the inequality term is block diagonal, with symmetric blocks of size $2g \times 2g$. This implies that the Lagrangian Hessian, $\nabla_{pp}^2 L$, is symmetric and block tridiagonal (with symmetric blocks of size $2g \times 2g$).

Appendix B: Implementation Details of the Newton Iteration

1) Indefiniteness Correction:

The computation of the Newton step is performed by solving (36). We also require Δp to be a descent direction, which happens if and only if $Z^T (\nabla_{pp}^2 L + D_{\mathcal{I}}^T \Sigma D_{\mathcal{I}}) Z \succ 0$, where Z is a null space matrix of $D_{\mathcal{E}}$. This condition is equivalent to the inertia of the Newton matrix of (36) being equal to $(2gK, K, 0)$ [24]. The standard way to enforce this condition (as is done in IPOPT, for example [25]) is to factor the Newton matrix using the Bunch-Kaufman algorithm (see [26], [27]) and then verify the inertia condition. If the inertia is not correct, then a positive diagonal perturbation is added to $\nabla_{pp}^2 L$ and the Newton matrix is refactored. The process

is repeated using increasingly larger perturbations until the inertia condition is satisfied [25]. This approach is disadvantageous in our problem setting because generic factorization algorithms destroy the block-diagonal pattern of the matrix. Also, the factorization algorithm may be used multiple times, with each call being computationally expensive. Instead, we adopted a simpler heuristic that guarantees a descent direction in a single factorization, at the cost of additional line search evaluations (see the next subsection for details on the line search procedure).

As indicated in (39), we consider a block-wise perturbation. Recall that the only source of indefiniteness comes from $\nabla_{pp}^2(L - \phi)$, which is block tridiagonal:

$$\nabla_{pp}^2(L - \phi) = \nabla_{pp}^2(y^T c_{\mathcal{E}}) + \nabla_{pp}^2(z^T ([g_{\mathcal{I}}(p_i)]_{i=1}^K + s)), \quad (58a)$$

$$\nabla_{pp}^2(L - \phi) = \nabla_{pp}^2(y^T c_{\mathcal{E}}) + \begin{bmatrix} \nabla_{p_1 p_1}^2(z_1^T g_{\mathcal{I}}(p_1)) & & \\ & \ddots & \\ & & \nabla_{p_K p_K}^2(z_K^T g_{\mathcal{I}}(p_K)) \end{bmatrix}. \quad (58b)$$

We take the perturbation sizes as

$$\delta_i = l_{\mathcal{E}} + \|\nabla_{p_i p_i}^2(z_i^T g_{\mathcal{I}}(p_i))\|_F, \quad i = 1, \dots, K, \quad (59)$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $l_{\mathcal{E}} \geq 0$ is an upper bound for the magnitude of the largest negative eigenvalue of $\nabla_{pp}^2(y^T c_{\mathcal{E}})$, i.e., $-l_{\mathcal{E}}$ is a lower bound of the minimum eigenvalue of $\nabla_{pp}^2(y^T c_{\mathcal{E}})$. Notice that the Frobenius norm is never smaller than the induced 2-norm of a matrix (see exercise 5.6.P23 of [28]), so this choice of δ_i guarantees that

$$\nabla_{p_i p_i}^2(z_i^T g_{\mathcal{I}}(p_i)) + (\delta_i - l_{\mathcal{E}})I \succeq 0, \quad i = 1, \dots, K, \quad (60)$$

and thus $\nabla_{pp}^2(L - \phi) + S \succeq 0$. We still need to discuss how to choose $l_{\mathcal{E}}$. The next theorem shows that we can find an appropriate $l_{\mathcal{E}}$ with little computational effort.

Theorem 2. *Let the entries of $\nabla_{pp}^2(y^T c_{\mathcal{E}})$ be given by (57), and choose*

$$l_{\mathcal{E}} = -4 \left(1 + \cos \left(\frac{\pi}{K+1} \right) \right) \cdot \min \left\{ 0, \min_{j=1, \dots, K+1} [w_j((y)_j - (y)_{j-1})] \right\}, \quad (61)$$

then

$$\lambda_{\min}(\nabla_{pp}^2(y^T c_{\mathcal{E}})) \geq -l_{\mathcal{E}}. \quad (62)$$

Proof. See Appendix E. \square

The time cost of computing S is $O(Kg^2)$, so the time cost of the Newton step remains asymptotically unchanged. The indefiniteness condition is not applied at every iteration, but only when necessary according to the line search.

2) Positivity of Dual Variables and Line Search:

The KKT conditions of the interior point formulation require the entries of vectors z and s to have the same sign (see (29b)). We require that $s > 0$, as otherwise the barrier terms are undefined. Therefore, we also require that $z > 0$. The Newton step may yield an update Δz that makes some of the entries of $z + \Delta z$ non-positive. To prevent this situation, we scale Δz using the fraction-to-boundary rule (see [29]) to ensure that the updated vector remains in the positive orthant. Another source of difficulties for the Newton iteration is the high nonlinearity that may arise in the interior point formulation. In such cases, the Newton linearization is not a good approximation, except for small step lengths. Thus, if the Newton step does not yield a good enough decrease of the objective function, we should take a smaller step. We achieve this by using a backtracking line search over the Armijo condition of an appropriate merit function [30]. In summary, the new iterate variables are computed as follows:

$$\alpha_z^{\max} = \max \{ \alpha \in [0, 1] : z + \alpha \Delta z \geq (1 - \tau)z \}, \quad (63a)$$

$$\alpha_z^{\max} = \min \{ 1, \min \{ -\tau(z)_i / (\Delta z)_i : (\Delta z)_i < 0 \} \}, \quad (63b)$$

$$p^+ = p + \gamma^M \Delta p, \quad (63c)$$

$$y^+ = y + \gamma^M \alpha_z^{\max} \Delta y, \quad (63d)$$

$$z^+ = z + \gamma^M \alpha_z^{\max} \Delta z, \quad (63e)$$

$$s^+ = -c_{\mathcal{I}}(p^+), \quad (63f)$$

for parameters $\tau, \gamma \in (0, 1)$. We remark that (63a) is the fraction-to-boundary rule (see [30]) and (63b) is an equivalent definition that is easier to implement computationally. M is the smallest non-negative integer satisfying the line search conditions. More formally, define the merit function as

$$\psi(p) = \psi_o(p) + \psi_c(p), \quad \text{where} \quad (64a)$$

$$\psi_o(p) = \phi(p) - \mu \sum_{i=1}^K \sum_{j \in \mathcal{I}} \ln[\max\{-g_j(p_i), 0\}], \quad (64b)$$

$$\psi_c(p) = \nu \|c_{\mathcal{E}}(p)\|_1, \quad (64c)$$

for some parameter $\nu \geq 0$, where we use the convention that $\ln 0 = -\infty$. Then the line search conditions are

$$\psi(p + \gamma^M \Delta p) \leq \psi(p) + \eta [(\nabla_p \psi_o(p))^T \gamma^M \Delta p + \psi_c(p + \gamma^M \Delta p) - \psi_c(p)], \quad (65a)$$

$$\frac{\min_{j=1, \dots, K+1} \|b_j(p + \gamma^M \Delta p)\|_{\infty}}{\max_{j=1, \dots, K+1} \|b_j(p + \gamma^M \Delta p)\|_{\infty}} > \epsilon_{\text{ls}}, \quad (65b)$$

$$\frac{(K+1)^{-1} \left\| \sum_{j=1}^{K+1} q_j(p + \gamma^M \Delta p) \right\|_{\infty}}{\max_{j=1, \dots, K+1} \|q_j(p + \gamma^M \Delta p)\|_{\infty}} > \epsilon_{\text{ls}}, \quad (65c)$$

for some parameters $\eta, \epsilon_{\text{ls}} \in (0, 1)$. Equation (65a) is the Armijo condition [30]. Equations (65b) and (65c) are the necessary conditions for Theorem 1, evaluated at the candidate path $p + \gamma^M \Delta p$. For this paper, we chose $\tau = 0.99$, $\gamma = 0.5$, $\nu = 0$ and $\eta = 10^{-4}$. These values were adapted from typical values used in solvers like IPOPT, with the exception of ν . We remark that the equality constraints corresponding to the spacing between points are used to eliminate an ill-conditioning inherent to the problem formulation. As such, slight violations of these constraints are not problematic, so it

is not strictly necessary not penalize them. Furthermore, our numerical tests showed that the algorithm is unable to make any progress when the equality constraints are penalized, so we chose $\nu = 0$. M is computed by trial and error starting with $M = 0$ and increasing its value by 1 until all conditions are satisfied (in the extended real sense) or

$$\gamma^M \leq \epsilon_{\text{ls}}, \quad (66)$$

for some parameter $\epsilon_{\text{ls}} \in (0, 1)$. For this paper, we chose $\epsilon_{\text{ls}} = 10^{-6}$. It may happen that the Newton direction may not be productive at all, in which case $M \rightarrow \infty$. In such cases, the violation of (66) signals the need for a safer step direction, so the indefiniteness condition is applied and the step is recomputed. If (66) is violated again after the indefiniteness correction, then the algorithm reports failure, returns the current solution, and terminates.

3) Stopping Criterion:

The Newton iteration seeks primal and dual variables that solve the first-order KKT equations, but we are only interested in the values of the primal variables. In some situations, the Newton iteration may converge in the primal variables, but not the dual variables (when the constraint qualifications are “almost” not satisfied, for example). To avoid this problem, we follow the approach of IPOPT (see [25]) to define an error metric that is scaled with respect to the dual variables:

$$\rho_d = \max \left\{ \rho_{\max}, \frac{\|y\|_1 + \|z\|_1}{|\mathcal{E}| + K|\mathcal{I}|} \right\} / \rho_{\max}, \quad (67a)$$

$$\rho_c = \max \left\{ \rho_{\max}, \frac{\|z\|_1}{K|\mathcal{I}|} \right\} / \rho_{\max}, \quad (67b)$$

$$E_\mu(p, s, y, z) = \max \left\{ \frac{\|\nabla_p L(p, s, y, z)\|_\infty}{\rho_d}, \frac{\|s \circ z - \mu \vec{1}\|_\infty}{\rho_c}, \frac{\|c_\mathcal{E}(p)\|_\infty}{\rho_c} \right\}, \quad (67c)$$

for some parameter $\rho_{\max} > 0$. For this work, we chose $\rho_{\max} = 100$. The stopping criterion for the Newton iteration is

$$E_\mu(p, s, y, z) \leq \epsilon_{\text{tol}}, \quad (68)$$

where ϵ_{tol} is the error tolerance specified by the user. This criterion provides the advantage of being robust against problems where the primal variables converge but the dual variables diverge (e.g., the solution does not satisfy the KKT conditions).

4) Newton Iteration Algorithm:

We have described all the necessary tools to implement a complete Newton iteration for the interior point method. This way we can solve the barrier problem for a fixed barrier parameter μ , as long as we are provided an initial feasible path. Pseudo-code of the procedure given an initial feasible path p is presented in Algorithm 2. The idea is to iteratively compute Newton steps until the error criterion is satisfied (success) or a maximum number of iterations (denoted as iter_{\max}) is reached (failure). For the starting values of the remaining variables, we choose $y = 0$, $s = -c_\mathcal{I}(p)$, and $z = \mu \odot s$. For a small enough μ , the barrier problem’s solution will be a good

enough approximation to the solution of the original problem. We remark that the parameter ν should be kept at 0, so it is not considered an input. On the other hand, Algorithm 2 requires the user to specify the vectors of power flow equations f and OPF constraints $g_\mathcal{I}$, which together completely characterize the power system and the OPF feasible region.

Algorithm 2 Solution of Barrier Problem (Inner Loop)

```

1: procedure BARRIERSOLVE( $f, g_\mathcal{I}, p, \mu, \epsilon_{\text{tol}}, \text{iter}_{\max}, \tau,$ 
    $\gamma, \eta, \epsilon_{\text{ls}}, \rho_{\max}$ )
2:    $s \leftarrow -c_\mathcal{I}(p)$ 
3:    $y \leftarrow 0, z \leftarrow \mu \vec{1} \odot s, \text{iter} \leftarrow 0$ 
4:   compute  $\nabla_{pp}^2 \phi$  from (53)
5:   for  $\text{iter} = 1, \dots, \text{iter}_{\max}$  do
6:     didCorrection  $\leftarrow$  False
7:     compute  $\nabla_p L$  and  $E_\mu(p, s, y, z)$ 
8:     if  $E_\mu(p, s, y, z) \leq \epsilon_{\text{tol}}$  then break
9:     compute  $D_\mathcal{E}, D_\mathcal{I}, \Sigma, c_\mathcal{I}(p)$  from (30)
10:    compute  $\nabla_{pp}^2 (L - \phi)$ 
11:     $\delta_i \leftarrow 0$  for  $i = 1, \dots, K$ 
12:    compute  $S$  from (39) and compute  $\Gamma$  from (40b)
13:    compute  $\Delta p, \Delta y$  by solving (42) with a block
      tridiagonal routine
14:    compute  $\Delta z$  from (37a)
15:     $\gamma^M \leftarrow$  perform line search until (65) is satisfied
      or (66) is violated
16:    if  $\gamma^M \leq \epsilon_{\text{ls}}$  and didCorrection then
17:      report “failed after inertia correction”
18:      break
19:    else if  $\gamma^M \leq \epsilon_{\text{ls}}$  then
20:      compute  $l_\mathcal{E}$  from (61)
21:      compute  $\delta_i$ , for  $i = 1, \dots, K$ , from (59)
22:      didCorrection  $\leftarrow$  True
23:      go to 12:
24:      compute  $p^+, y^+, z^+, s^+$  from (63)
25:       $(p, y, z, s) \leftarrow (p^+, y^+, z^+, s^+)$ 
26:  return  $p, s$ 

```

Appendix C: OPF Derivatives Computation

The OPF constraints may depend on either the control vector u or the state vector x . The value of x depends directly on u , so the Implicit Function Theorem is required to compute gradients and Hessians of these constraints. We first need to compute the first- and second-order derivatives of x with respect to u . From the power flow equations, we have that $f(x, u) = 0$, and thus

$$0 = \frac{df}{du}, \quad (69a)$$

$$0 = \frac{\partial f}{\partial x^T} \frac{dx}{du^T} + \frac{\partial f}{\partial u^T}, \quad (69b)$$

$$\frac{dx}{du^T} = - \left(\frac{\partial f}{\partial x^T} \right)^{-1} \frac{\partial f}{\partial u^T}, \quad (69c)$$

as long as the Jacobian is invertible. For the second-order derivatives, we have, for $m = 1, \dots, 2g$, that

$$0 = \frac{d^2 f}{d(u)_m du^T}, \quad (70a)$$

$$0 = \frac{d}{d(u)_m} \left(\frac{\partial f}{\partial x^T} \frac{dx}{du^T} + \frac{\partial f}{\partial u^T} \right), \quad (70b)$$

$$0 = \frac{\partial^2 f}{d(u)_m \partial x^T} \frac{dx}{du^T} + \frac{\partial f}{\partial x^T} \frac{d^2 x}{d(u)_m du^T} + \frac{\partial^2 f}{d(u)_m \partial u^T}, \quad (70c)$$

$$0 = \left(\sum_{k=1}^{2n} \frac{\partial^2 f}{\partial(x)_k \partial x^T} \frac{d(x)_k}{d(u)_m} + \frac{\partial^2 f}{\partial(u)_m \partial x^T} \right) \frac{dx}{du^T} + \frac{\partial f}{\partial x^T} \frac{d^2 x}{d(u)_m du^T} + \sum_{k=1}^{2n} \frac{\partial^2 f}{\partial(x)_k \partial u^T} \frac{d(x)_k}{d(u)_m} + \frac{\partial^2 f}{\partial(u)_m \partial u^T}. \quad (70d)$$

For the power flow equations, we have, for $k = 1, \dots, 2n$ and $m = 1, \dots, 2g$, that

$$\frac{\partial^2 f}{d(u)_m \partial x^T} = 0, \quad \frac{\partial^2 f}{\partial(x)_k \partial u^T} = 0, \quad \frac{\partial^2 f}{\partial(u)_m \partial u^T} = 0, \quad (71)$$

and

$$\left(\frac{dx}{du^T} \right)_{km} = \frac{d(x)_k}{d(u)_m}, \quad (72)$$

therefore

$$\frac{\partial^2 f}{d(u)_m \partial x^T} = \sum_{k=1}^{2n} \frac{\partial^2 f}{\partial(x)_k \partial x^T} \cdot \left(\frac{dx}{du^T} \right)_{km}. \quad (73)$$

Replacing:

$$0 = \left[\sum_{k=1}^{2n} \frac{\partial^2 f}{\partial(x)_k \partial x^T} \cdot \left(\frac{dx}{du^T} \right)_{km} \right] \frac{dx}{du^T} + \frac{\partial f}{\partial x^T} \frac{d^2 x}{d(u)_m du^T}, \quad (74a)$$

$$\frac{d^2 x}{d(u)_m du^T} = - \left(\frac{\partial f}{\partial x^T} \right)^{-1} \left[\sum_{k=1}^{2n} \frac{\partial^2 f}{\partial(x)_k \partial x^T} \cdot \left(\frac{dx}{du^T} \right)_{km} \right] \frac{dx}{du^T}, \quad (74b)$$

where the second-order partial derivative term is constant for the power flow equations. More specifically:

$$\frac{\partial^2 f}{\partial(x)_k \partial x^T} = J_k. \quad (75)$$

This comes from the fact that in rectangular coordinates there exists constant matrices J_0, J_1, \dots, J_{2n} such that the power flow Jacobian at x , denoted $J(x)$, can be written as

$$J(x) = J_0 + \sum_{k=1}^{2n} J_k \cdot (x)_k. \quad (76)$$

Replacing in (74b) we get that

$$\frac{d^2 x}{d(u)_m du^T} = - \left(\frac{\partial f}{\partial x^T} \right)^{-1} \left[\sum_{k=1}^{2n} J_k \cdot \left(\frac{dx}{du^T} \right)_{km} \right] \frac{dx}{du^T}. \quad (77)$$

The optimal power constraints can be divided in two types. The first type corresponds to constraints of the form $g(u) \leq 0$ that do not depend directly on x (that is, $g = g_i, i \in \mathcal{U}$). Gradients and Hessians are computed as usual in this case. The second type corresponds to constraints of the form $g(x) \leq 0$

that do not depend directly on u (that is, $g = g_i, i \in \mathcal{X} \cup \mathcal{P}$). For the second type, we have that

$$\frac{dg}{du^T} = \frac{dg}{dx^T} \frac{dx}{du^T}, \quad (78)$$

and

$$\frac{dg}{d(u)_m du^T} = \frac{d}{d(u)_m} \left(\frac{dg}{dx^T} \frac{dx}{du^T} \right), \quad (79a)$$

$$\frac{dg}{d(u)_m du^T} = \frac{dg}{d(u)_m dx^T} \frac{dx}{du^T} + \frac{dg}{dx^T} \frac{d^2 x}{d(u)_m du^T}, \quad (79b)$$

$$\frac{dg}{d(u)_m du^T} = \frac{dx^T}{d(u)_m} \frac{d^2 g}{dx dx^T} \frac{dx}{du^T} + \frac{dg}{dx^T} \frac{d^2 x}{d(u)_m du^T}, \quad (79c)$$

$$\frac{dg}{dud u^T} = \left(\frac{dx}{du^T} \right)^T \frac{d^2 g}{dx dx^T} \frac{dx}{du^T} + \left[\frac{dg}{dx^T} \frac{d^2 x}{d(u)_m du^T} \right]_{m=1}^{2g}, \quad (79d)$$

$$\frac{dg}{dud u^T} = \left(\frac{dx}{du^T} \right)^T \frac{d^2 g}{dx dx^T} \frac{dx}{du^T} - \left[\left(\sum_{k=1}^{2n} \frac{dg}{dx^T} \cdot \left(\frac{\partial f}{\partial x^T} \right)^{-1} J_k \cdot \left(\frac{dx}{du^T} \right)_{km} \right) \frac{dx}{du^T} \right]_{m=1}^{2g}, \quad (79e)$$

$$\frac{dg}{dud u^T} = \left(\frac{dx}{du^T} \right)^T \frac{d^2 g}{dx dx^T} \frac{dx}{du^T} - \left[\sum_{k=1}^{2n} \left(\frac{dg}{dx^T} \cdot \left(\frac{\partial f}{\partial x^T} \right)^{-1} J_k \right) \left(\frac{dx}{du^T} \right)_{km} \right]_{m=1}^{2g} \frac{dx}{du^T}, \quad (79f)$$

$$\frac{dg}{dud u^T} = \left(\frac{dx}{du^T} \right)^T \frac{d^2 g}{dx dx^T} \frac{dx}{du^T} - \left(\frac{dx}{du^T} \right)^T \left[\frac{dg}{dx^T} \cdot \left(\frac{\partial f}{\partial x^T} \right)^{-1} J_k \right]_{k=1}^{2n} \frac{dx}{du^T}, \quad (79g)$$

$$\frac{dg}{dud u^T} = \left(\frac{dx}{du^T} \right)^T \left(\frac{d^2 g}{dx dx^T} - \left[\frac{dg}{dx^T} \cdot \left(\frac{\partial f}{\partial x^T} \right)^{-1} J_k \right]_{k=1}^{2n} \right) \frac{dx}{du^T}, \quad (79h)$$

There is one special constraint: the power flow feasibility set. This constraint can be expressed as

$$g(x) = -|\det J(x)|. \quad (80)$$

We assume that the Jacobian is non-singular, so $\det J(x) \neq 0$ and thus the absolute value can be differentiated, yielding:

$$\frac{dg}{d(x)_k} = -\text{sign}(\det J) \frac{d}{d(x)_k} (\det J). \quad (81)$$

Using Jacobi's formula for the derivative of the determinant (see [31]) and the fact that the Jacobian is invertible in the feasible region, we get that

$$\frac{dg}{d(x)_k} = -\text{sign}(\det J) \text{tr} \left[(\det J) J^{-1} \frac{dJ}{d(x)_k} \right], \quad (82a)$$

$$\frac{dg}{d(x)_k} = -|\det J| \text{tr} (J^{-1} J_k). \quad (82b)$$

For the Hessian, we have that

$$\frac{dg}{d(x)_m d(x)_k} = \frac{d}{d(x)_m} [-|\det J| \operatorname{tr}(J^{-1} J_k)], \quad (83a)$$

$$\begin{aligned} \frac{dg}{d(x)_m d(x)_k} &= \frac{d}{d(x)_m} (-|\det J|) \operatorname{tr}(J^{-1} J_k) \\ &\quad - |\det J| \frac{d}{d(x)_m} \operatorname{tr}(J^{-1} J_k), \end{aligned} \quad (83b)$$

$$\frac{dg}{d(x)_m d(x)_k} = \frac{dg}{d(x)_m} \operatorname{tr}(J^{-1} J_k) - |\det J| \operatorname{tr}\left(\frac{dJ^{-1}}{d(x)_m} J_k\right), \quad (83c)$$

$$\begin{aligned} \frac{dg}{d(x)_m d(x)_k} &= \frac{dg}{d(x)_m} \operatorname{tr}(J^{-1} J_k) \\ &\quad + |\det J| \operatorname{tr}(J^{-1} J_m J^{-1} J_k). \end{aligned} \quad (83d)$$

Now the Jacobian of the inequality constraints is simply

$$D_{p_i} g_{\mathcal{I}}(p_i) = \left[\frac{g_j(p_i)}{du^T} \right]_{j \in \mathcal{I}}. \quad (84)$$

The Hessian of the Lagrangian term of the inequality constraints is⁴

$$\nabla_{pp}^2(z^T([g_{\mathcal{I}}(p_i)]_{i=1}^K + s)) = \nabla_{pp}^2(z^T[g_{\mathcal{I}}(p_i)]_{i=1}^K), \quad (85a)$$

$$\nabla_{pp}^2(z^T([g_{\mathcal{I}}(p_i)]_{i=1}^K + s)) = \operatorname{diag}\left([\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{I}}(p_i)]_{i=1}^K\right). \quad (85b)$$

We decompose $g_{\mathcal{I}}$ as

$$g_{\mathcal{I}}(p_i) = g_{\mathcal{U}}(p_i) + g_{\mathcal{X} \cup \mathcal{P}}(p_i), \quad (86)$$

where

$$(g_{\mathcal{U}}(p_i))_j = \begin{cases} g_j(p_i), & j \in \mathcal{U} \\ 0, & \text{else} \end{cases}, \quad (87a)$$

$$(g_{\mathcal{X} \cup \mathcal{P}}(p_i))_j = \begin{cases} g_j(p_i), & j \in \mathcal{X} \cup \mathcal{P} \\ 0, & \text{else} \end{cases}. \quad (87b)$$

Now we can write

$$\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{I}}(p_i) = \nabla_{p_i p_i}^2 z_i^T g_{\mathcal{U}}(p_i) + \nabla_{p_i p_i}^2 z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i). \quad (88)$$

The first term, $\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{U}}(p_i)$, can be computed trivially, as $g_{\mathcal{U}}(p_i)$ is an explicit function of u . For the second term, we have that

$$\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i) = \frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dud u^T}, \quad (89a)$$

$$\begin{aligned} \nabla_{p_i p_i}^2 z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i) &= \left(\frac{dx}{du^T} \right)^T \frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dxdx^T} \frac{dx}{du^T} \\ &\quad + \left[\frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dx^T} \frac{d^2 x}{d(u)_m du^T} \right]_{m=1}^{2g}, \end{aligned} \quad (89b)$$

⁴We generalize the diag operator for matrices of suitable size as follows: let B_k be a $n \times n$ matrix for all $k = 1, \dots, K$ and define the $(nK) \times n$ matrix M as $M = [B_k]_{k=1}^K$, then

$$\operatorname{diag}(M) = \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_K \end{bmatrix} \in \mathbb{F}^{(nK) \times (nK)}$$

for an appropriate field \mathbb{F} .

$$\frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dxdx^T} = \sum_{j \in \mathcal{X} \cup \mathcal{P}} (z_i)_j \frac{dg_j(p_i)}{dxdx^T}, \quad (89c)$$

$$\frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dx^T} = \sum_{j \in \mathcal{X} \cup \mathcal{P}} (z_i)_j \frac{dg_j(p_i)}{dx^T}. \quad (89d)$$

The computation of $\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i)$ involves multiple matrix-matrix products, so a closer inspection is warranted in order to develop an efficient implementation. To this end, we first compute $d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))/(dxdx^T)$ and $d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))/dx^T$ as usual. Next, we compute intermediate variables in the order presented next:

$$\theta_{1,i} = \frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dx^T} \left(\frac{\partial f}{\partial x^T} \right)^{-1}, \quad (90a)$$

$$\theta_{2,ik} = \theta_{1,i} J_k, \quad k = 1, \dots, 2n, \quad (90b)$$

$$\Theta_{3,i} = \frac{d(z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i))}{dxdx^T} - [\theta_{2,ik}]_{k=1}^{2n}. \quad (90c)$$

Lastly, we compute $\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{X} \cup \mathcal{P}}(p_i)$ as

$$\nabla_{p_i p_i}^2 z_i^T g_{\mathcal{I}}(p_i) = \left(\frac{dx}{du^T} \right)^T \Theta_{3,i} \frac{dx}{du^T}. \quad (91)$$

Appendix D: Proof of Theorem 1

Proof. For brevity, we omit the dependence of $b_j(p)$ and $q_j(p)$ on p . Notice that $D_{\mathcal{E}}$ can be written as

$$D_{\mathcal{E}} = \begin{bmatrix} b_1^T + b_2^T & -b_2^T & & & \\ -b_2^T & \ddots & \ddots & & \\ & \ddots & \ddots & -b_K^T & \\ & & -b_K^T & b_K^T + b_{K+1}^T & \end{bmatrix}. \quad (92)$$

Define, for any $k \in \mathbb{N}$, the following matrices:

$$L_k = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \ddots & \ddots & \\ 1 & \dots & 1 & 1 \end{bmatrix} \in \mathbb{R}^{k \times k}, \quad (93a)$$

$$U_k = \begin{bmatrix} 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix} \in \mathbb{R}^{k \times k}, \quad (93b)$$

and, for $k = 2, \dots, K$, define

$$D_k = \operatorname{diag}([q_j^T]_{j \in \{1, \dots, K+1\} \setminus \{k\}})^T \in \mathbb{R}^{2gK \times K}, \quad (94a)$$

$$M_k = \begin{bmatrix} L_{k-1} \otimes I_{2g} & 0 \\ 0 & L_{K-k+1}^T \otimes I_{2g} \end{bmatrix} \in \mathbb{R}^{2gK \times 2gK}, \quad (94b)$$

where I_k denotes the $k \times k$ identity matrix. Lastly, define

$$Q = ([q_j^T]_{j=1}^{K+1})^T, \quad (95)$$

and let $e_k \in \mathbb{R}^k$ be the last (rightmost) column of the $k \times k$ identity matrix. Then, for any $k = 2, \dots, K$, we have by direct computation that

$$(D_p c_{\mathcal{E}}) \cdot (M_k D_k) =$$

$$\begin{bmatrix} U_{k-1} + e_{k-1}(b_k^T(Q)_{1:k-1,:}) & -e_{k-1}(b_k^T(Q)_{k+1:K+1,:}) \\ -e_{k-1}(b_k^T(Q)_{1:k-1,:}) & U_{K-k+1}^T + e_{k-1}(b_k^T(Q)_{k+1:K+1,:}) \end{bmatrix} \quad (96)$$

We next eliminate the off-diagonal entries of U_{k-1} and U_{K-k+1}^T using elementary column operations. Thus, there exists an invertible matrix $C \in \mathbb{R}^{K \times K}$ such that

$$(D_p c_{\mathcal{E}}) \cdot (M_k D_k) C = \begin{bmatrix} I_{k-2} & & & \\ \boxtimes & 1 + \sum_{j=1}^{k-1} b_k^T q_j & -\sum_{j=k+1}^{K+1} b_k^T q_j & \boxtimes \\ \boxtimes & -\sum_{j=1}^{k-1} b_k^T q_j & 1 + \sum_{j=k+1}^{K+1} b_k^T q_j & \boxtimes \\ & & & I_{K-k} \end{bmatrix}, \quad (97)$$

where the symbol \boxtimes denotes blocks with possibly non-zero, but unimportant entries. We can eliminate the \boxtimes entries using elementary row operations, so there exists an invertible matrix $R \in \mathbb{R}^{K \times K}$ such that

$$A_k = R(D_p c_{\mathcal{E}}) \cdot (M_k D_k) C = \begin{bmatrix} I_{k-2} & & & \\ & 1 + \sum_{j=1}^{k-1} b_k^T q_j & -\sum_{j=k+1}^{K+1} b_k^T q_j & \\ & -\sum_{j=1}^{k-1} b_k^T q_j & 1 + \sum_{j=k+1}^{K+1} b_k^T q_j & \\ & & & I_{K-k} \end{bmatrix}, \quad (98)$$

where we named the final matrix as A_k for convenience. The determinant of A_k can be readily computed as

$$\det(A_k) = \left(1 + \sum_{j=1}^{k-1} b_k^T q_j\right) \left(1 + \sum_{j=k+1}^{K+1} b_k^T q_j\right) - \left(\sum_{j=1}^{k-1} b_k^T q_j\right) \left(\sum_{j=k+1}^{K+1} b_k^T q_j\right), \quad (99a)$$

$$\det(A_k) = 1 + b_k^T \left(\sum_{j=1}^{k-1} q_j + \sum_{j=k+1}^{K+1} q_j\right). \quad (99b)$$

Notice that $b_k^T q_k = 1$, and hence

$$\det(A_k) = b_k^T \sum_{j=1}^{K+1} q_j = b_k^T Q \vec{1}, \quad (100)$$

where $\vec{1}$ denotes a vector with all entries equal to 1. The previous argument (with small modifications) still holds for $k = 1$ and $k = K + 1$, so (100) holds for $k = 1, \dots, K + 1$. Assume for contradiction that $D_{\mathcal{E}}$ is rank deficient, then from properties of the rank (see [28]), we have that

$$\text{rank}(A_k) \leq \text{rank}(D_{\mathcal{E}}) < K, \quad (101)$$

so A_k is singular and therefore

$$q_k^T Q \vec{1} = (b_k^T b_k)^{-1} b_k^T Q \vec{1} = (b_k^T b_k)^{-1} \det(A_k) = 0, \quad (102)$$

for $k = 1, \dots, K + 1$. In matrix-vector form we have that

$$Q^T Q \vec{1} = 0, \quad (103)$$

so

$$\left(\sum_{j=1}^{K+1} q_j\right)^2 = (Q \vec{1})^T (Q \vec{1}) = \vec{1}^T (Q^T Q \vec{1}) = 0, \quad (104)$$

which implies that $\sum_{j=1}^{K+1} q_j = 0$, and so we have a contradiction. \square

Appendix E: Proof of Theorem 2

Proof. For brevity, we define the scalars a_j as

$$a_j = \begin{cases} 0, & j = 0 \\ 2w_j((y)_j - (y)_{j-1}) & j \in \{1, \dots, K+1\} \end{cases}, \quad (105)$$

and λ^- as

$$\lambda^- = \min_{k=0, \dots, K+1} a_k \leq 0, \quad (106)$$

so $l_{\mathcal{E}}$ becomes

$$l_{\mathcal{E}} = -2\lambda^- \left(1 + \cos\left(\frac{\pi}{K+1}\right)\right), \quad (107)$$

and we can write $\nabla_{pp}^2(y^T c_{\mathcal{E}})$ as

$$\nabla_{pp}^2(y^T c_{\mathcal{E}}) = \begin{bmatrix} (a_1 + a_2)I & -a_2I & & \\ -a_2I & \ddots & \ddots & \\ & \ddots & \ddots & -a_K I \\ & & -a_K I & (a_K + a_{K+1})I \end{bmatrix}. \quad (108)$$

Define the following matrices:

$$D = \begin{bmatrix} a_1 I & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{K+1} I \end{bmatrix} \in \mathbb{R}^{2g(K+1) \times 2g(K+1)}, \quad (109a)$$

$$A = \begin{bmatrix} I & 0 & \cdots & 0 \\ -I & I & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -I & I \\ 0 & \cdots & 0 & 0 & -I \end{bmatrix} \in \mathbb{R}^{2g(K+1) \times 2gK}. \quad (109b)$$

We can then factor $\nabla_{pp}^2(y^T c_{\mathcal{E}})$ as

$$\nabla_{pp}^2(y^T c_{\mathcal{E}}) = A^T D A. \quad (110)$$

Let the following matrix be:

$$D_1 = \begin{bmatrix} (a_1 - \lambda^-)I & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (a_{K+1} - \lambda^-)I \end{bmatrix}, \quad (111)$$

then clearly

$$\nabla_{pp}^2(y^T c_{\mathcal{E}}) = \lambda^- A^T A + A^T D_1 A. \quad (112)$$

The matrix D_1 is block diagonal with positive semidefinite blocks, so $D_1 \succeq 0$. This means that $A^T D_1 A \succeq 0$ and thus, from the concavity of the smallest eigenvalue (see [31]), we have that

$$\begin{aligned} \lambda_{\min}(\nabla_{pp}^2(y^T c_{\mathcal{E}})) &\geq \lambda_{\min}(\lambda^- A^T A) + \lambda_{\min}(A^T D_1 A) \\ &\geq \lambda_{\min}(\lambda^- A^T A). \end{aligned} \quad (113)$$

Denote the Kronecker product by \otimes , then

$$A^T A = \begin{bmatrix} 2I & -I & 0 & \cdots & 0 \\ -I & 2I & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2I & -I \\ 0 & \cdots & 0 & -I & 2I \end{bmatrix} = Y \otimes I, \quad (114)$$

where

$$Y = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix} \in \mathbb{R}^{K \times K}. \quad (115)$$

Denote the eigenvalues of Y by ξ_k for $k = 1, \dots, K$, then (see [32])

$$\xi_k = 2 - 2 \cos \left(\frac{k\pi}{K+1} \right). \quad (116)$$

The eigenvalues of the Kronecker product correspond to the product of each matrix eigenvalues, so the eigenvalues of $Y \otimes I$ correspond to ξ_k , each repeated K times [23]. As $\lambda^- \leq 0$ we get that

$$\lambda_{\min} (\nabla_{pp}^2 (y^T c_{\mathcal{E}})) \geq \lambda^- \cdot \lambda_{\max} (A^T A), \quad (117a)$$

$$\lambda_{\min} (\nabla_{pp}^2 (y^T c_{\mathcal{E}})) \geq \lambda^- \cdot \lambda_{\max} (Y \otimes I), \quad (117b)$$

$$\lambda_{\min} (\nabla_{pp}^2 (y^T c_{\mathcal{E}})) \geq \lambda^- \cdot \lambda_{\max} (Y), \quad (117c)$$

$$\lambda_{\min} (\nabla_{pp}^2 (y^T c_{\mathcal{E}})) \geq \lambda^- \left(2 - 2 \cos \left(\frac{K\pi}{K+1} \right) \right), \quad (117d)$$

$$\lambda_{\min} (\nabla_{pp}^2 (y^T c_{\mathcal{E}})) \geq 2\lambda^- \left(1 + \cos \left(\frac{\pi}{K+1} \right) \right) = -l_{\mathcal{E}}, \quad (117e)$$

and the claim follows. \square