Polynomial-Time Algorithms for Setting Tight Big-M Coefficients in Transmission Expansion Planning with Disconnected Buses

Behnam Jabbari Marand[1*] and Adolfo R. Escobedo[2]

[1,2]Industrial & Systems Engineering, North Carolina State University, 915 Partners Way, Raleigh, 27606, NC, US.

*Corresponding author(s). E-mail(s): Bjabbar@ncsu.edu;
Contributing authors: Arescobedo@ncsu.edu;

Abstract

The increasing penetration of renewable energy into power systems necessitates the development of effective methodologies to integrate initially disconnected generation sources into the grid. This paper introduces the Longest Shortest-Path-Connection (LSPC) algorithm, a graph-based method to enhance the mixed-integer linear programming disjunctive formulation of Transmission Expansion Planning (TEP) using valid inequalities (VIs). Traditional approaches for determining big-M coefficients in disconnected TEP networks typically rely on solving the computationally intensive Longest Path Problem (LPP). In contrast, LSPC circumvents these limitations by efficiently identifying relevant power flow paths between disconnected buses within the expansion network. We demonstrate that the VIs generated from these identified paths dominate those derived from LPP-based methods and other existing approaches.

Keywords: Transmission expansion planning, Mixed-integer linear programming, Valid inequalities, Renewable energy sources integration

1 Introduction

The Transmission Expansion Planning (TEP) problem entails adding transmission lines within and between systems to accommodate future demand growth at the lowest possible cost [1]. The strategic importance of TEP in power systems cannot be overstated, given its long-term implications for system operations. The evolving landscape, characterized by renewable energy integration, large-scale generation projects, and market integration, has significantly elevated the complexity of TEP problem, demanding more effective solution methods [2].

TEP problem is often addressed via a linear approximation model, namely Direct Current Optimal Power Flow (DC-OPF), which is widely applied in power system optimization (e.g. [3],[4],[5]). This linearization is obtained by assuming uniform voltage magnitudes, minimal angle differences, and disregarding reactive power, considering the low conductance of transmission lines [6]. These simplifications provide an effective trade-off between simplicity and accuracy, making them suitable for TEP, where operational considerations are less critical due to the long-term planning horizon and extensive power transmission distances [7].

The introduction of discrete decisions to DC-OPF transforms it from a linear program (LP) into a mixed-integer linear program (MILP), which is generally intractable (i.e., NP-hard [8]). Its complexity explains the current emphasis on metaheuristics and other approximate methods ([9],[10],[11]), even though these approaches do not provide formal guarantees of the solution quality. Hybrid methods for DC-TEP that couple heuristics with branch-and-bound algorithms (e.g., [12],[13],[14]) have been explored. However, despite having theoretical guarantees, their high computational cost limits them to smaller problems, hindering their practical scalability. Benders' decomposition is another popular method for solving DC-TEP that guarantees optimal solutions ([15],[16],[17],[18]), but it suffers from slow convergence in large-scale problems.

Cutting planes offer another exact approach for accelerating the solution process of DC-TEP. They have been extensively investigated for various power systems optimization problems involving discrete decisions, such as DC OPF-based Optimal Transmission Switching (DC-OTS) and Unit Commitment (DC-UC), with their computational benefits widely corroborated ([19],[20],[21],[22]). However, cutting plane methods for DC-TEP have received relatively less attention. Tsamasphyrou [23] and Binato [15] conducted pioneering research on the use of cutting planes to enhance the MILP disjunctive formulation of DC-TEP. However, the cutting planes they developed prove to be either too loose or limited in scope to translate into significant computational improvements. Skolfield et al. [24] provide the most in-depth study to date, deriving a broader set of cutting planes that incorporate a wide range of variables to further strengthen the problem formulation. Nonetheless, real-world transmission expansion situations, particularly those involving the integration of new buses into the existing network, remain unexplored. To bridge this gap, this paper proposes an efficient graph-based methodology for deriving new classes of cutting planes, leveraging the structural properties of the DC-TEP disjunctive formulation.

The rest of the paper is organized as follows. Section 2 presents the DC-TEP formulation and introduces background concepts. Section 3 motivates the featured methodology, denoted as the longest shortest-path-connection algorithm, to tackle situations that require new-bus integration. Section 4 provides a detailed description of this algorithm, along with proofs of correctness and complexity.

2 Modeling framework and background

This section provides the notation and underlying mathematical model that serves as the basis for the methodology developed in this work. In power systems, *buses* (i.e., nodes) represent connection points for various electrical components (e.g., power plants, substations) and are linked by *corridors*, which are transmission pathways between buses. This work distinguishes between *established corridors*, which connect buses solely through existing transmission lines, and *expansion corridors*, which incorporate candidate lines for potential network expansion. For simplicity, it is assumed that each corridor can accommodate only a single existing line or a candidate line.

2.1 Notation overview

Sets

$n \in \mathcal{B}$	Buses (i.e., nodes)
$(i,j) \in \Omega^0$	Established corridors: corridors containing only an established line
$(i,j) \in \Omega^1$	Expansion corridors: corridors containing only a candidate line

Parameters

c_{ij}	Cost of installing a line in corridor $(i,j) \in \Omega^1$
c_n	Cost per unit of power generation at bus n
$\overline{g}_n$	Upper limit of power generation at bus n
d_n	Active power demand at bus n
$\overline{\theta}_{ij}$	Limit on the angle difference between buses i and j
$\overline{P}_{ij}$	Maximum capacity of (candidate) line within corridor $(i,j) \in \Omega^1$
$\overline{P}^0_{ij}$	Maximum capacity of (existing) line within corridor $(i,j) \in \Omega^0$
x_{ij}	Reactance of line in corridor (i,j)
σ	A scaling factor for aligning generation costs with transmission investment costs

Variables

P^0_{ij}	Active power transmitted through the existing line in corridor (i,j)
P_{ij}	Active power transmitted through the candidate line in corridor (i,j)
g_n	Active power produced by the generator at bus n
θ_n	Voltage angle at bus n
y_{ij}	$\begin{cases} 1 & \text{If a candidate line within corridor } (i,j) \text{ is purchased} \\ 0 & \text{Otherwise} \end{cases}$

2.2 MILP disjunctive formulation of DC-TEP

This paper employs the MILP disjunctive model of DC-TEP [25] with a single investment period. This formulation is as follows:

$$\min \sum_{(i,j)\in\Omega^1} c_{ij} y_{ij} + \sum_{n\in\mathcal{B}} \sigma c_n g_n \tag{1a}$$

$$\sum_{(i,n)\in\Omega^0} P^0_{in} + \sum_{(i,n)\in\Omega^1} P_{in} - \sum_{(n,i)\in\Omega^0} P^0_{ni} - \sum_{(n,i)\in\Omega^1} P_{ni} + g_n = d_n \quad \forall n \in \mathcal{B} \tag{1b}$$

$$-\overline{P}^0_{ij} \le P^0_{ij} \le \overline{P}^0_{ij} \quad \forall (i,j) \in \Omega^0 \tag{1c}$$

$$-\overline{P}_{ij} y_{ij} \le P_{ij} \le \overline{P}_{ij} y_{ij} \quad \forall (i,j) \in \Omega^1 \tag{1d}$$

$$x_{ij}P^0_{ij} = (\theta_i - \theta_j) \qquad \forall (i,j) \in \Omega^0 \qquad (1e)$$

$$-\overline{\theta}_{ij}(1-y_{ij}) \leq x_{ij}P_{ij} - (\theta_i - \theta_j) \leq \overline{\theta}_{ij}(1-y_{ij}) \qquad \forall (i,j) \in \Omega^1 \qquad (1f)$$

$$g_n \leq \overline{g}_n \qquad \forall n \in \mathcal{B} \qquad (1g)$$

$$y_{ij} \in \{0,1\} \qquad \forall (i,j) \in \Omega^1 \qquad (1h)$$

$$g_n \geq 0, \theta_n \text{ unr.} \qquad \forall n \in \mathcal{B} \qquad (1i)$$

$$P^0_{ij}, P_{ij} \text{ unr.} \qquad \forall (i,j) \in \Omega^0 \cup \Omega^1 \qquad (1j)$$

Objective function (1a) minimizes the total cost of adding new lines and power generation; generation costs are scaled by a factor of σ to ensure comparability to investment costs. Constraint (1b) enforces Kirchoff's Current Law (KCL), also known as the flow balance equations. They ensure that at each bus, the inflow and generation are together equal to the summed outflow and demand. Constraints (1c) and (1d) are capacity limits for existing and candidate lines, respectively. Constraints (1e) and (1f) enforce Kirchhoff's Voltage Law (KVL) for existing and candidate lines, respectively, by equating the product of line reactance and power flow to the corresponding bus angle difference within a corridor. Here, the susceptance parameter b_{ij} from the standard DCOPF formulation is replaced with the reactance $x_{ij} = \frac{-1}{b_{ij}}$ to simplify the notation in the formulations. Constraint (1f) employs a sufficiently large disjunctive parameter, i.e., $\overline{\theta}_{ij}$, to guarantee inequality redundancy for unconstructed corridors. Previous studies use a big-M parameter for the disjunctive coefficient; this work uses the parameter $\overline{\theta}_{ij}$ to generalize its application to all bus pairs (the standard TEP formulation defines this parameter only for adjacent buses). The remaining constraints specify the domain of values for the decision variables.

2.3 Review of TEP formulation improvement approaches

In constraint (1f), the disjunctive parameter $\overline{\theta}_{ij}$ imposes an upper limit on the angle difference between buses connected by expansion corridor $(i,j) \in \Omega^1$. The selection of $\overline{\theta}_{ij}$ significantly contributes to the quality of the problem formulation. These coefficients must be sufficiently large not to cut off any integer-feasible solution. However, they should be kept as small as possible to provide tighter LP relaxations, thereby expediting solution times (and mitigating numerical issues [26]). To elaborate, consider a pair of buses i and j, succinctly represented as $[i,j]$, that are connected by an established corridor. A valid upper bound on the angle difference between the pair can be derived by incorporating the line capacity constraint (1c) into the KVL constraint (1e) as:

$$|\theta_i - \theta_j| \leq x_{ij}\overline{P}^0_{ij}, \qquad (2)$$

from the fact that $|P^0_{ij}| \leq \overline{P}^0_{ij}$. Henceforth, the *capacity-reactance product* of a line (i,j) (e.g., see the right-hand side of (2), $x_{ij}\overline{P}^0_{ij}$) is abbreviated as CR_{ij}.

The above bound is valid when a transmission line is constructed between buses i and j, but it may not apply to corridors not already connected via an established corridor. Nonetheless, it is possible to utilize (2) to establish bounds on buses that can be connected via expansion lines. A simple bound was derived by Tsamasphyrou et al. [23], who calculate a single bound γ that is applicable to any pair of adjacent buses in Ω^1 and is given by:

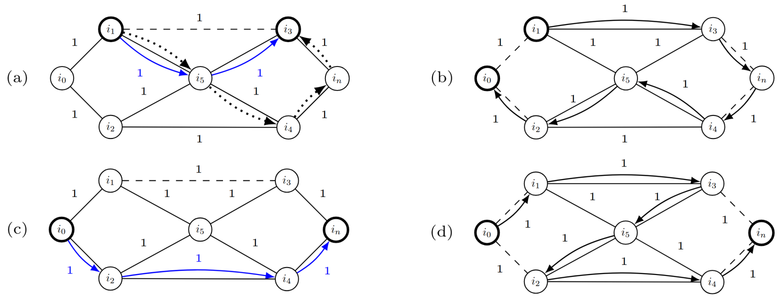

Fig. 1: Identifying relevant paths between bus pairs based on their connectivity in G^0 and G: (a) Bus pair $[i_1, i_3]$ is adjacent in G and connected in G^0; (b) Bus pair $[i_0, i_1]$ is adjacent in G but disconnected in G^0; (c) Bus pair $[i_0, i_n]$ is non-adjacent in G yet connected in G^0; (d) Bus pair $[i_0, i_n]$ is non-adjacent in G and disconnected in G^0.

$$|\theta_i - \theta_j| \leq \sum_{(k,l)\in\Omega^0\cup\Omega^1} CR_{kl} = \gamma, \quad \forall (i,j) \in \Omega^1. \tag{3}$$

Upper bound γ effectively represents a worst-case scenario where any flow traveling between i and j would traverse all the corridors ($\Omega^0 \cup \Omega^1$) in the network. It is obtained by sequentially applying the angle difference inequality (2) to each corridor and summing the results.

Upper bound γ is straightforward to compute, but its magnitude becomes excessive even for very small instances, meaning it does not provide real computational advantages. To derive tighter upper bounds, it is necessary to restrict attention to more relevant power flows between bus pairs. This entails identifying and analyzing only the relevant paths that power flow can take between buses i and j, which may either be an *established path*, denoted by ρ^0_{ij}, composed of existing lines, or a *potential path*, denoted by ρ_{ij}, consisting of expansion corridors with or without existing corridors. The general approach involves efficiently identifying such paths between bus pairs, and calculating their lengths — by summing the lines' capacity-reactance products (see 2) — to refine $\overline{\theta}_{ij}$. To help explain the applicability of this approach, let $G = (\mathcal{B}, \Omega^0 \cup \Omega^1)$ be the *expansion network* associated with the inclusion of potential investment decisions in TEP. Additionally, it is necessary to define the *initial network* $G^0 = (\mathcal{B}, \Omega^0)$, which consists solely of existing lines.

Figure 1 illustrates different network structures and expansion situations, to motivate the applicability and limitations of existing approaches for identifying relevant paths between bus pairs and deriving their angle difference bounds. In the figure, the network G is composed of both established corridors, depicted by solid edges, and expansion corridors, shown by dashed edges, with edge weights denoting the capacity-reactance product of the lines. When buses i and j within expansion corridor (i, j) are connected via an established path ρ^0_{ij} in G^0, a valid upper bound on their voltage angle difference is obtained by traversing the path. That is, starting from one endpoint of the expansion corridor, i, and following path ρ^0_{ij} to the opposite endpoint j, summing the inequalities (2) creates a telescoping effect on the left-hand side, resulting in the angle difference $|\theta_i - \theta_j|$. Simultaneously, the right-hand side accumulates the capacity-reactance products (CR_{ij}) of the traversed lines.

Example 1. *Figure 1a depicts expansion corridor* (i_1, i_3) *and established path* $\rho^0_{i_1 i_3} := \langle (i_1, i_5), (i_5, i_4), (i_4, i_n), (i_n, i_3) \rangle$. *Summing the angle difference inequalities along* $\rho^0_{i_1 i_3}$, *represented by the dotted arrows, yields the angle difference upper bound for* $[i_1, i_3]$*:*

$$|\sum_{(i,j)\in\rho^0_{i_1 i_3}} (\theta_i - \theta_j)| \leq \underbrace{|\theta_{i_1} - \theta_{i_5}|}_{\leq CR_{i_1 i_5}} + \underbrace{|\theta_{i_5} - \theta_{i_4}|}_{\leq CR_{i_5 i_4}} + \underbrace{|\theta_{i_4} - \theta_{i_n}|}_{\leq CR_{i_4 i_n}} + \underbrace{|\theta_{i_n} - \theta_{i_3}|}_{\leq CR_{i_n i_3}} \tag{4a}$$

$$\Rightarrow |\theta_{i_1} - \theta_{i_3}| \leq CR_{i_1 i_5} + CR_{i_5 i_4} + CR_{i_4 i_n} + CR_{i_n i_3} = 1 + 1 + 1 + 1 = 4 \tag{4b}$$

The presence of multiple established paths connecting i and j results in multiple inequalities similar to (4b). Binato [15] proposes solving the Shortest Path Problem (SPP) within G^0 to identify the tightest bound on $\overline{\theta}_{ij}$. In Example 1, the shortest established path $\underline{\rho}^0_{i_1 i_3} := \langle (i_1, i_5), (i_5, i_3) \rangle$ (indicated by blue solid arrows) yields an upper bound of 2 on the angle difference for buses i_1 and i_3, dominating the previously determined bound.

Skolfield et al. [24] extended the application of this approach, which was previously limited to adjacent buses (i.e., those with a single expansion corridor between them), to establish angle difference bounds, $\overline{\theta}_{ij}$, for any $i, j \in \mathcal{B}$, connected within G^0 (i.e., buses that can reach each other via existing lines). The authors derive a valid upper bound $\overline{\theta}_{ij}$ by solving SPP across all simple established paths, where the associated capacity-reactance product sum is denoted as $\underline{CR(\rho^0_{ij})} := \min_r \{CR(\rho^0_{ij_r})\}$. Figure 1c illustrates non-adjacent buses i_0 and i_n connected within G^0, where the shortest path between them yields an upper bound. Considering all bus pairs enables the derivation of many additional angle difference bounds beyond those in Binato's work, thereby enhancing the DC-TEP formulation.

The cutting planes derived by Binato and Skolfield et al. require established paths connecting buses in G^0 to efficiently establish angle difference upper bounds. However, such paths may not exist in certain network configurations and expansion decisions, for example, when the expansion of transmission grids involves integrating new buses. Figures 1b and 1d showcase two relevant situations, which can be motivated by the need to connect renewable energy sources. For instance, the integration of wind farms, as exemplified by ERCOT's CREZ initiative to deliver remote wind power to high-demand regions in Texas ([27],[28]), requires the addition of new buses (e.g., buses i_0 and i_n in Figure 1d). The rise of distributed generation (e.g., rooftop solar panels [29]) and the expansion of transmission grids through line branching are additional drivers for adding new buses in TEP. While angle difference bounds analogous to those obtained for connected bus pairs can be derived for disconnected pairs, obtaining the respective coefficients proves impractical. This is because any single potential path between the disconnected bus pairs in G could be relevant, depending on the solution to the problem, and thus solving the SPP is inapplicable in these cases.

3 Motivating the longest shortest-path-connection algorithm

It is inefficient to use the aforementioned power flow-path based approaches to derive angle difference upper bounds for bus pairs that are disconnected in the initial network. The computational difficulties arise from the combinatorial nature of the expansion decisions — that is, not knowing a priori whether these decisions will affect the connectivity of currently disconnected bus pairs — leading to numerous potential connection scenarios. Indeed, Binato proposes employing a total enumeration approach, meaning evaluating all potential paths in

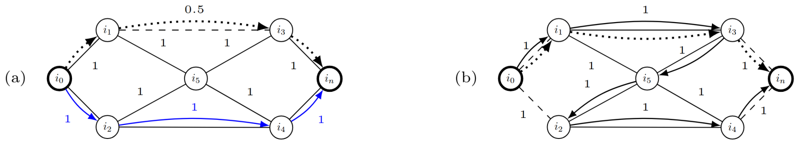

Fig. 2: Tightening SPP- and LPP-based angle difference bounds using path-based VIs: (a) Bus pair $[i_0, i_n]$ is connected in G^0 and a potential path (dotted arrows) shorter than the SPP-based path exists; (b) Bus pair $[i_0, i_n]$ is disconnected in G^0 and a potential path (dotted arrows) shorter than the LPP-based path exists.

G connecting the disconnected bus pair and selecting the longest one to ensure no feasible solutions are excluded ([15],[24],[30]). In contrast, Skolfield et al. [24] introduce a more sophisticated approach that leverages parallel paths between the bus pair to further tighten the angle difference bound from what is possible with total enumeration. However, both approaches still necessitate solving the Longest (Simple) Path Problem (LPP), which is NP-hard [31]. Apart from this computational difficulty, the lengths of LPP-based paths often exceed reasonable lengths of plausible power flows between the disconnected bus pair due to the meshed structure and the size of transmission grids, leading to overly conservative angle difference bounds. Such aspects hinder the effectiveness of both Binato's total enumeration approach and Skolfield's bound-tightening technique.

Next, we explain how LPP can often be circumvented in practical situations where, aside from the new disconnected buses, the initial network is well-connected and can be leveraged to reduce the relevant connection possibilities. To motivate this insight, it is necessary to formally define valid inequalities (VIs). For an integer programming problem, written succinctly as $\min\{cx : x \in X\}$, with feasible region $\mathcal{S} = \{x \in \mathbb{Z}^n : Ax \leq b\}$, an inequality $\pi x \leq \pi_0$ is valid if it holds for all feasible solutions $x \in \mathcal{S}$ [32]. A VI is deemed effective if it reduces the feasible region of the relaxed problem represented by polyhedral set $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$. For instance, Skolfield et al. introduce *path-based* VIs to tighten the angle difference bounds derived from SPP or LPP. For potential paths $\rho_{i_0 i_n}$ parallel to the original SPP- or LPP-based path connecting i_0 and i_n, these VIs can be written as

$$|\theta_{i_0} - \theta_{i_n}| \leq CR(\rho_{i_0 i_n}) + (1 - y_{\rho_{i_0 i_n}})(\overline{\theta}_{i_0 i_n} - CR(\rho_{i_0 i_n})), \tag{5}$$

where $y_{\rho_{i_0 i_n}} \in \{0, 1\}$. Without loss of generality, path-based VI (5) is presented in a simplified form from Skolfield's exposition, associating a binary variable with a path; in contrast, the original path-based VI enumerates all corridors along the path, with one binary variable for each expansion corridor. In (5), binary variable $y_{\rho_{i_0 i_n}}$ equals 1 if the potential path $\rho_{i_0 i_n}$ is constructed, and 0 otherwise. When $y_{\rho_{i_0 i_n}} = 1$, the tighter upper bound $CR(\rho_{i_0 i_n})$ is enforced. If $y_{\rho_{i_0 i_n}} = 0$, the initial upper bound $\overline{\theta}_{i_0 i_n}$ is maintained, which equals $\underline{CR(\rho^0_{i_0 i_n})}$ when i_0 and i_n are connected in G^0, and $\overline{CR(\rho_{i_0 i_n})}$ (representing the longest path length) otherwise. Figures 2a and 2b show example candidate paths for deriving path-based VIs (indicated by the dotted arrows) when a bus pair is connected (left subfigure) and disconnected in G^0 (right subfigure).

As a first contribution, we add a condition for ensuring the validity of the path-based inequalities, namely,

$$\{\rho_{i_0 i_n} \in \mathcal{C}^G_{i_0 i_n} \mid CR(\rho_{i_0 i_n}) < \overline{\theta}_{i_0 i_n}\}, \tag{6}$$

where $\mathcal{C}^G_{i_0 i_n}$ represents the set of all paths connecting i_0 to i_n in the network G. Violating this condition, which is missing from the original expression by Skolfield et al., could lead to (5) generating invalid inequalities. Proposition 4 in Appendix 5 presents the complete expression of inequality (5) (i.e., using expansion corridor variables) and demonstrates that it is valid only for potential paths that meet condition (6). In addition, Proposition 5 in Appendix 5 demonstrates that for potential paths $\rho_{i_0 i_n}$ within the *angle difference effective domain* — defined as $|\theta_{i_0} - \theta_{i_n}| \geq CR(\rho_{i_0 i_n})$ (representing the tightest achievable bound through $\rho_{i_0 i_n}$) — the path-based VIs constructed using an LPP-based initial bound are dominated by those initialized with the bounds developed in Section 4. This proposition indicates that the effectiveness of path-based VIs diminishes as the initial angle difference bound $\overline{\theta}_{i_0 i_n}$ grows, with excessively large bounds failing to tighten the LP relaxation. This underscores the significance of identifying the shortest relevant paths between initially disconnected bus pairs. Indeed, LPP is unnecessary and ineffective when the existing network between two disconnected buses has a high degree of connectivity. In such cases, shorter connections can often be identified. In practice, very few buses are disconnected; hence, the number of potential corridor combinations for connecting them is small. This is justified by the fact that, in large-scale power systems, bus degrees follow an exponential distribution with a mean of 2 [33], indicating a low number of incident lines.

Example 2. *Figure 2b illustrates the longest path between buses i_0 and i_n, namely $\overline{\rho}_{i_0 i_n} := \langle (i_0, i_1), (i_1, i_3), (i_3, i_5), (i_5, i_2), (i_2, i_4), (i_4, i_n) \rangle$, which is represented by solid arrows and has a length of 6. However, assuming the construction of corridors (i_0, i_1) and (i_4, i_n), a shorter and more relevant path with a length of 4 can be identified: $\rho_{i_0 i_n} := \langle (i_0, i_1), (i_1, i_5), (i_5, i_4), (i_4, i_n) \rangle$. Both paths traverse the terminal corridors (i_0, i_1) and (i_4, i_n), but the latter employs a shorter path through the existing grid to connect the intermediate buses i_1 and i_4.*

Building upon these observations, we introduce the Longest Shortest-Path-Connection (LSPC) algorithm for deriving tight angle difference upper bounds for new-bus integration situations. For a given disconnected pair, the algorithm circumvents the prohibitive computational effort needed to solve LPP by evaluating connections between the disconnected buses and their neighboring buses. These neighboring buses are then linked through the shortest path within G^0. By merging the connections with the neighbors and the SPP-based sub-paths between them, complete paths between the original buses are formed. The longest of these shortest-path-connections is then selected to establish a valid angle difference bound.

Before proceeding, it is crucial to distinguish the proposed concept from the *diameter* of graph $G(V, E)$, denoted as $diam(G) := \max_{u,v \in V} d(u, v)$ [34]. In words, the diameter represents the maximum shortest path distance between any two nodes in a static graph. This metric fails to guarantee a feasible bound due to the variable nature of the network in expansion planning. When applied to the expansion graph G, the diameter may result in overly conservative estimates by considering the bus pair with the longest shortest path among all pairs in the network. While it evaluates shortest paths between all buses, which may or may not be constructed, the LSPC algorithm focuses only on the part of the graph that is relevant to the specific disconnected bus pair.

4 The longest shortest-path-connection algorithm

This section introduces the LSPC algorithm, a novel graph-based approach for tightening the formulation of the DC-TEP problem. The algorithm has two phases, which are applicable to different expansion situations. The results from the first phase are used to derive additional bounds for the second phase.

4.1 LSPC Phase I: Bridging-bus integration

The LSPC algorithm establishes an upper bound on the angle difference between disconnected buses in minimally disconnected networks, where a small number of investments in candidate lines would integrate the new buses into the existing network. Phase I focuses on the simplest case, common to many real-world situations, where a new bus is separated from the existing grid by a single expansion corridor.

To proceed, let $N_G(i)$ denote the set of buses adjacent to bus i in G. Phase I applies to disconnected bus pair $[i, j]$, where, for all neighboring pairs $[n_i, n_j]$ with $n_i \in N_G(i)$ and $n_j \in N_G(j)$, if a path exists between n_i and n_j in the expansion network (i.e., $\mathcal{C}^{G}_{n_i n_j} \neq \emptyset$), a corresponding path also exists in the initial network (i.e., $\mathcal{C}^{G^0}_{n_i n_j} \neq \emptyset$). The existence of these paths ensures that all potential paths between i and j can be realized by connecting i to n_i and j to n_j. In such cases, a relevant path between i and j that extends the connection between n_i and n_j is determined by first identifying the shortest path between n_i and n_j in G^0 and then adding the lengths of the corresponding terminal corridors, (i, n_i) and (n_j, j), to the length of the latter sub-path.

Phase I iterates over all connected neighboring bus pairs $[n_i, n_j]$ in G to construct all relevant paths between buses i and j. Nevertheless, given the typically low degree of buses in a power grid, this usually yields a small number of options. The longest among them is extracted to establish $\overline{\theta}_{ij}$. To proceed, we introduce two concepts:

Definition 1. *A node-isolated graph of G with respect to node i, excludes all edges connecting node i to its neighbors, that is,* $\overline{G}_i := (\mathcal{B}, \{\Omega^0 \cup \Omega^1\} \backslash \{(i, N_G(i))\})$.

Using a *node-isolated* graph ensures that only the neighbors of i and j with a potential path uniting them (i.e., $\mathcal{C}^{\overline{G}_i \cap \overline{G}_j}_{n_i n_j} \neq \emptyset$) are considered in identifying candidate paths between i and j. Absent this, if one pair of neighbors is connected by a path in G, other neighboring pairs may be incorrectly deemed to be connected by the same path, reaching it through the corridors $(i, N_G(i))$ and $(N_G(j), j)$.

Definition 2. *The reachability set* $\mathcal{C}^N_{ij}$ *for buses i and j includes* $[n_i, n_j]$*, whenever* n_i *and* n_j *are connected in* $\overline{G}_i \cap \overline{G}_j$*, and it includes* $[i, j]$ *if* $(i, j) \in \Omega^1$*; in mathematical notation, its elements are given by*

$$\{[n_i, n_j] | n_i \in N_G(i), n_j \in N_G(j), n_i \neq j, n_j \neq i, \mathcal{C}^{\overline{G}_i \cap \overline{G}_j}_{n_i n_j} \neq \emptyset\} \cup \{[i, j] | (i, j) \in \Omega^1\}. \quad (7)$$

The length of the path ρ_{ij} corresponding to each element $[\eta_i, \eta_j] \in \mathcal{C}^N_{ij}$, where $\eta_i \in \{i, n_i\}$, yields a lower bound for $\overline{\theta}_{ij}$ (since the path may or may not be chosen in the TEP solution). Specifically, the set $\mathcal{C}^N_{ij}$ includes two types of elements, distinguished by the number of corridors along the corresponding path. For each element $[\eta_i, \eta_j]$, the lower bounds are given by:

Fig. 3: Examples of situations where the reachability set should be refined: (a) In $\mathcal{C}^N_{i_0 i_1}$, $\{[i_2, i_3], [i_3, i_3]\}$ are replaced by $[i_0, i_3]$, as $\{(i_0, i_2), (i_0, i_3)\} \in \Omega^0$; (b) In $\mathcal{C}^N_{i_0 i_6}$, $[i_4, i_4]$, $[i_4, i_5]$, and $[i_2, i_4]$ are replaced by $[i_0, i_4]$, since $\{(i_0, i_2), (i_0, i_4)\} \in \Omega^0$, and since $[i_0, i_4]$ is connected in G^0, $[i_0, i_4]$ also replaces $[i_1, i_4]$.

i. Adjacent: The path consists of a single expansion corridor, specifically $\rho_{ij_{[i,j]}} := \langle (i, j) \rangle$. Therefore, $\overline{\theta}_{ij} \geq CR(\rho_{ij_{[i,j]}}) = CR_{ij}$.
ii. Non-adjacent (connection via extended path): For a neighboring pair $[n_i, n_j]$, an SPP is solved to obtain $\overline{\theta}_{n_i n_j} = \underline{CR(\rho^0_{n_i n_j})}$. Subsequently, the terminal-corridor lengths CR_{in_i} and $CR_{n_j j}$ are added to obtain the length of the full potential path $CR(\rho_{ij_{[n_i,n_j]}})$, resulting in the lower bound

$$\overline{\theta}_{ij} \geq CR(\rho_{ij_{[n_i,n_j]}}) = x_{in_i} \overline{P}_{in_i} + \underline{CR(\rho^0_{n_i n_j})} + x_{n_j j} \overline{P}_{n_j j}. \tag{8}$$

Example 3. *Consider the disconnected bus pair i_0 and i_1 in Figure 3a. Given that $(i_0, i_1) \in \Omega^1$, case (i) is applicable to the element $[i_0, i_1] \in \mathcal{C}^N_{i_0 i_1}$, resulting in lower bound $\overline{\theta}_{i_0 i_1} \geq CR(\rho_{i_0 i_1{}_{[i_0,i_1]}}) = CR_{i_0 i_1} = 1$. Additionally, for the neighbor pair $[i_4, i_5] \in \mathcal{C}^N_{i_0 i_1}$, which corresponds to case (ii), the algorithm extends the shortest path between i_4 and i_5 to create a full path that connects i_0 to i_1, giving the lower bound*

$$\overline{\theta}_{i_0 i_1} \geq CR(\rho_{i_0 i_1{}_{[i_4,i_5]}}) = x_{i_0 i_4} \overline{P}_{i_0 i_4} + \underline{CR(\rho^0_{i_4 i_5})} + x_{i_5 i_1} \overline{P}_{i_5 i_1} = 1 + 1 + 1 = 3.$$

When bus i from a disconnected pair $[i, j]$ is connected to the initial grid (e.g., bus i_0 in Figures 3a and 3b), it is necessary to factor the existing connections between bus i and G^0 to ensure the inclusion of all potential paths and to expand the applicability of Phase I.

Example 4. *In Figure 3b, if the elements $[i_4, i_4]$, $[i_4, i_5]$, $[i_2, i_4]$, and $[i_2, i_5]$ are included in $\mathcal{C}^N_{i_0 i_6}$, the shortest potential path, namely $\rho_{i_0 i_6(1)} := \langle (i_0, i_2), (i_2, i_4), (i_4, i_6) \rangle$, is excluded from consideration. Instead, the longer paths $\rho_{i_0 i_6(2)} := \langle (i_0, i_4), (i_4, i_6) \rangle$, $\rho_{i_0 i_6(3)} := \langle (i_0, i_4), (i_4, i_2), (i_2, i_5), (i_5, i_6) \rangle$, and $\rho_{i_0 i_6(4)} := \langle (i_0, i_2), (i_2, i_5), (i_5, i_6) \rangle$ are obtained from case (ii) of Definition 2. In effect, enumerating all ordered neighboring pairs $[n_{i_0}, n_{i_6}]$ connected in G to construct $\mathcal{C}^N_{i_0 i_6}$ impedes obtaining the tightest value for $\overline{\theta}_{i_0 i_6}$. Additionally, including $[i_1, i_4]$ in $\mathcal{C}^N_{i_0 i_6}$ requires an established path between i_1 and i_4. However, since i_4 is already reachable from i_0, requiring connections from any $n_{i_0} \in N_G(i_0)$ to i_4 becomes unnecessary and restrictive. In summary, listing all pairs $[n_{i_0}, n_{i_6}]$ may limit the applicability of Phase I based on the existing connectivity of i_0.*

Definition 3. *For a disconnected bus pair where one bus is connected to the current network, the reachability set is refined by replacing $[n_i, n_j]$ with $[i, n_j]$ if $(i, n_i) \in \Omega^0$ or $\mathcal{C}^{G^0}_{in_j} \neq \emptyset$, and with $[n_i, j]$ if $(n_j, j) \in \Omega^0$ or $\mathcal{C}^{G^0}_{n_i j} \neq \emptyset$.*

To simplify the upcoming lemma and proof, connections of the form $[i, n_j]$ and $[n_i, j]$ are jointly represented as $[i, n_j]$ since G and G^0 are undirected (i.e., $[i, j]$ is equivalent to $[j, i]$).

Lemma 1. *Refining the reachability set $\mathcal{C}_{ij}^N$ by replacing $[n_i, n_j]$ with $[i, n_j]$ when $(i, n_i) \in \Omega^0$ or $\mathcal{C}_{in_j}^{G^0} \neq \emptyset$ ensures that all necessary connections are included to capture the relevant paths between buses i and j through n_j.*

Proof. Pairs $[n_i, n_j]$ are replaced with $[i, n_j]$ in two cases:

Case 1: $(i, n_i) \in \Omega^0$. Let $N^0(i) \subseteq N_G(i)$ denote the neighbors of bus i with $(i, n_i) \in \Omega^0$. It is correct to replace $[n_i, n_j]$ with $[i, n_j]$ since the SPP accounts for traversing all neighbors $n_i \in N^0(i)$, when determining the shortest existing path from i to n_j. When bus i and at least two of its neighbors, say n_i^1 and n_i^2, form a cycle of established corridors $c := \langle i, n_i^1, n_i^2, i \rangle$, there is a path with at least two corridors connecting each $n_i \in N^0(i)$ to i by traversing through other neighbors in $N^0(i)$. In other words, this refinement ensures that the reachability set captures all paths that pass through multiple neighbors. As a result, all sub-paths in the refined reachability set are no more than one corridor away from the original buses.

Case 2: $\mathcal{C}_{in_j}^{G^0} \neq \emptyset$ (i.e., bus i has a path to n_j in G^0). Assume that $n_i' \in N^0(i)$, and therefore the neighbor pair $[n_i', n_j] \in \mathcal{C}_{ij}^N$ is replaced with $[i, n_j]$ due to case 1. We demonstrate that the potential path through $[n_i, n_j] \in \mathcal{C}_{ij}^N$, where $\{(i, n_i), (n_j, j)\} \in \Omega^1$, should be excluded from $\mathcal{C}_{ij}^N$. Constructing the full path through $[n_i, n_j]$ requires adding lines across both corridors (i, n_i) and (n_j, j), whereas constructing the path corresponding to $[i, n_j]$ only necessitates building the corridor (n_j, j). More specifically, the construction of both lines is redundant and restrictive in this case: when the path for $[n_i, n_j]$ is constructed, the path for $[i, n_j]$ is automatically formed, whereas connections where only (n_j, j) is constructed are incorrectly discarded (since a path between i and j is still formed). $\square$

This refinement leads to a third type of elements in $\mathcal{C}_{ij}^N$:

iii. <u>Non-adjacent (existing connection via neighbor)</u>: For connections of the form $[i, n_j]$, $\overline{\theta}_{in_j}$ is obtained as $\underline{CR(\rho_{in_j}^0)}$ by solving the SPP for $[i, n_j]$. Then, the terminal corridor length $CR_{n_j j}$ is added to create the candidate path $\rho_{ij_{[i,n_j]}}$, providing the corresponding lower bound:
$$\overline{\theta}_{ij} \geq CR(\rho_{ij_{[i,n_j]}}) = \underline{CR(\rho_{in_j}^0)} + x_{n_j j}\overline{P}_{n_j j}. \tag{9}$$

Example 5. *In Example 3, the reachability set $\mathcal{C}_{i_0 i_1}^N = \{[i_0, i_1], [i_2, i_3], [i_3, i_3], [i_4, i_5]\}$ is refined to $\mathcal{C}_{i_0 i_1}^N = \{[i_0, i_1], [i_0, i_3], [i_4, i_5]\}$ by replacing $[i_2, i_3]$ and $[i_3, i_3]$ with $[i_0, i_3]$, as $\{(i_0, i_2), (i_0, i_3)\} \in \Omega^0$. The lower bound for $\overline{\theta}_{i_0 i_1}$ corresponding to the element $[i_0, i_3]$ is established as*
$$\overline{\theta}_{i_0 i_1} \geq CR(\rho_{i_0 i_{1[i_0,i_3]}}) = \underline{CR(\rho_{i_0 i_3}^0)} + x_{i_3 i_1}\overline{P}_{i_3 i_1} = 1 + 1 = 2.$$

The following theorem formally establishes the tightest angle difference bound for an initially disconnected bus pair by computing the lengths of candidate paths associated with the three types of elements of the reachability set.

Theorem 1. *For an initially disconnected bus pair $[i, j]$, if all pairs in $\mathcal{C}_{ij}^N \backslash \{[i, j]\}$ are connected in G^0, the tightest bound on the angle difference is determined by the longest shortest path within refined $\mathcal{C}_{ij}^N$ (considering cases i, ii, iii) as:*

Algorithm 1 Longest Shortest-Path-Connection Algorithm (LSPC) - Phase I

Inputs: $G, G^0, [i,j], \gamma$
Output: A tighter upper bound on the angle difference between buses i and j
$\overline{\theta}_{ij} \leftarrow 0$ ▷ Initialize $\overline{\theta}_{ij}$
Construct $\mathcal{C}_{ij}^N$ ▷ Construct the reachability set by applying Definition 2
if $\exists[n_i,n_j] \in \mathcal{C}_{ij}^N \mid (i,n_i) \in \Omega^0$ **or** $\mathcal{C}_{in_j}^{G^0} \neq \emptyset$ **then**
  Refine $\mathcal{C}_{ij}^N$ ▷ Refine the reachability set by applying Definition 3
end if
if $\mathcal{C}_{[\eta_i,\eta_j]}^{G^0} \neq \emptyset, \forall\, [\eta_i,\eta_j] \in \mathcal{C}_{ij}^N \backslash \{[i,j]\}$ **then** ▷ Verify Phase I's applicability to $[i,j]$
  for all $[\eta_i,\eta_j] \in \mathcal{C}_{ij}^N$ **do** ▷ Compute $CR(\rho_{ij_{[\eta_i,\eta_j]}})$ using cases (i),(ii), and (iii)
    $\overline{\theta}_{ij} \leftarrow \max(\overline{\theta}_{ij}, CR(\rho_{ij_{[\eta_i,\eta_j]}}))$
  end for
else
  $\overline{\theta}_{ij} \leftarrow \gamma$ ▷ Phase I fails
end if

$$|\theta_i - \theta_j| \leq \max_{[\eta_i,\eta_j] \in \mathcal{C}_{ij}^N} \{CR(\rho_{ij_{[\eta_i,\eta_j]}})\}; \tag{10}$$

where, the length of each path $\rho_{ij_{[\eta_i,\eta_j]}}$ *is computed as:*

$$CR(\rho_{ij_{[\eta_i,\eta_j]}}) = x_{ij}\overline{P}_{ij}, \qquad \forall\, [\eta_i,\eta_j] = [i,j] \tag{11a}$$

$$CR(\rho_{ij_{[\eta_i,\eta_j]}}) = \underline{CR(\rho_{in_j}^0)} + x_{n_jj}\overline{P}_{n_jj}, \qquad \forall\, [\eta_i,\eta_j] \in \{[i,n_j]\} \tag{11b}$$

$$CR(\rho_{ij_{[\eta_i,\eta_j]}}) = x_{in_i}\overline{P}_{in_i} + \underline{CR(\rho_{n_in_j}^0)} + x_{jn_j}\overline{P}_{jn_j}, \qquad \forall\, [\eta_i,\eta_j] \in \{[n_i,n_j]\} \tag{11c}$$

Proof. To ensure the reachability set $\mathcal{C}_{ij}^N$ captures all relevant paths between buses i and j, all ordered neighboring pairs $[n_i,n_j]$ connected in G, as well as the pair $[i,j]$ when $(i,j) \in \Omega^1$, are enumerated by applying (7). To extend the applicability of the LSPC algorithm and derive the tightest upper bound, $\mathcal{C}_{ij}^N$ is refined following Definition 3, as established by Lemma 1. Assuming all $[\eta_i,\eta_j] \in \mathcal{C}_{ij}^N \setminus \{[i,j]\}$ are connected in G^0, the length of the candidate path corresponding to $[\eta_i,\eta_j] \in \mathcal{C}_{ij}^N$ is calculated as follows:

Case i: For $[i,j] \in \mathcal{C}_{ij}^N$, if a line is established along the expansion corridor (i,j), the inequality

$$|\theta_i - \theta_j| \leq CR(\rho_{ij_{[i,j]}}) = CR_{ij} \tag{12}$$

follows from (2).

Case ii: For $[n_i,n_j] \in \mathcal{C}_{ij}^N$, we employ the SPP to determine their angle difference bounds $\overline{\theta}_{n_in_j}$. By extending the terminal expansion corridors (with lengths CR_{in_i} and CR_{n_jj}) on both sides of each SPP-based sub-path, a potential path between i and j is formed. Assuming the construction of both (i,n_i) and (n_j,j), the path is added to the network, yielding the angle difference VI

$$|\theta_i - \theta_j| \leq \underbrace{|\theta_i - \theta_{n_i}|}_{\leq CR_{in_i}} + \underbrace{|\theta_{n_i} - \theta_{n_j}|}_{\leq \underline{CR(\rho^0_{n_i n_j})}} + \underbrace{|\theta_{n_j} - \theta_j|}_{\leq CR_{n_j j}} \tag{13a}$$

$$\Rightarrow |\theta_i - \theta_j| \leq CR_{in_i} + \underline{CR(\rho^0_{n_i n_j})} + CR_{n_j j} = CR(\rho_{ij_{[n_i,n_j]}}). \tag{13b}$$

Case iii: For the candidate paths represented by $[i, n_j]$, constructing the corridor (n_j, j) extends the SPP-based sub-path from i to n_j and is required to guarantee that

$$|\theta_i - \theta_j| \leq \underbrace{|\theta_i - \theta_{n_j}|}_{\leq \underline{CR(\rho^0_{in_j})}} + \underbrace{|\theta_{n_j} - \theta_j|}_{\leq CR_{n_j j}} \tag{14a}$$

$$\Rightarrow |\theta_i - \theta_j| \leq \underline{CR(\rho^0_{in_j})} + CR_{n_j j} = CR(\rho_{ij_{[i,n_j]}}). \tag{14b}$$

Since the inequalities (12),(13), and (14) are derived from prospective paths not yet included in the current network, it is essential to ensure that $\overline{\theta}_{ij} \geq \underset{[\eta_i,\eta_j]\in\mathcal{C}^N_{ij}}{\cup} CR(\rho_{ij_{[\eta_i,\eta_j]}})$. To maintain the validity and tightness of $\overline{\theta}_{ij}$, it must be at least as large as the maximum lower bound, thereby establishing inequality (10).

To show that (10) is the tightest achievable angel difference VI for $[i, j]$, it is important to note that a smaller $\overline{\theta}_{ij}$ would require a shorter path for at least one pair in $\mathcal{C}^N_{ij}$. However, this is impossible, as the SPP is solved to connect all pairs in the reachability set and form sub-paths. Subsequently, sub-paths are merged with at most one candidate line on each side to form full paths, as supported by case 2 in Lemma 1, leaving no room for further improvement. □

Proposition 2. *Assuming $n = |\mathcal{B}|$ and $P(deg(i) > K) \sim \exp(-0.5K)$ for all $i \in \mathcal{B}$ and $K \in \mathbb{N}$, the worst-case time complexity of determining $\overline{\theta}_{ij}$ with LSPC Phase I is $\mathcal{O}(n^2)$.*

Proof. In the worst-case, buses i and j are connected to all other buses in the network through an expansion corridor, meaning that $|N_G(i) \setminus \{j\}| = |N_G(j) \setminus \{i\}| = n - 2$. With all neighboring buses fully connected in G^0, the cardinality of the set $\mathcal{C}^N_{ij}$ reaches its maximum value of $(n-2)^2 + 1$. Dijkstra's algorithm is applied to each of $(n-2)$ neighbors to find the shortest path to all $n-3$ other neighboring buses. The set $\mathcal{C}^N_{ij}$ contains $(n-2)(n-3)$ elements, to which case (ii) applies giving a total of $2(n-2)(n-3)$ additions. Furthermore, there are $(n-2)$ pairs in $\mathcal{C}^N_{ij}$ to which case (iii) applies, each requiring a single addition operation. Afterwards, maximizing the obtained bounds requires $(n-2)^2 + 1$ comparisons. Given the quadratic time complexity of Dijkstra's algorithm for both additions and comparisons, the worst-case time complexity of Phase I is $\mathcal{O}(n^3)$. However, based on the assumed bus-degree distribution, the number of invocations of Dijkstra's algorithm is decreased to a constant, reducing the overall complexity to $\mathcal{O}(n^2)$. □

4.2 Phase II: Tree-structured bus integration

Other new-bus integration situations of interest may involve a tree-like structure of new buses, such as wind farms and intermediate substations, or interconnected components, such as another isolated existing network, to be added to the grid.

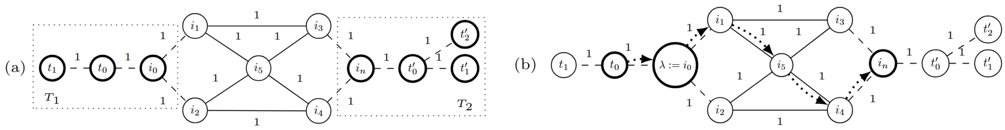

Fig. 4: Integration of new buses from network-expansion trees into the grid: (a) Buses within trees T_1 and T_2 to be integrated into the network; (b) Phase II combines the Phase I-based sub-paths $\rho_{t_0 i_0} := \langle (t_0, i_0) \rangle$ and $\rho_{i_0 i_n} := \langle (i_0, i_1), (i_1, i_5), (i_5, i_4), (i_4, i_n) \rangle$ to link t_0 to i_n via the intermediary bus i_0.

Definition 4. *A network-expansion tree $T(\mathcal{B}_T, \Omega_T^1)$ is defined as an acyclic, connected subgraph of G, rooted at a bus that is separated from G^0 by a single expansion corridor.*

LSPC Phase II leverages the fact that any path to a new bus in a network-expansion tree must pass through a certain intermediary neighboring bus. Formally, Phase II applies to disconnected buses i and j, for which relevant paths to an *intermediary bus* λ, with corresponding bounds $\overline{\theta}_{i\lambda}$ and $\overline{\theta}_{\lambda j}$, can be identified through Phase I.

Example 6. *In Figure 4a, T_1 and T_2 exemplify network-expansion trees, where the buses in these trees illustrate the applicability of Phase II. In this network, when power flows into bus i_n from i_3 or i_4, the only pathway for the power to reach bus t_1' is through bus t_0'. As a result, once a path to i_n is built, it can be easily extended along the tree structure rooted at i_n to reach additional new buses located deeper in the tree.*
Consider now the buses t_0 and t_0'. Once power flows into bus i_0 from t_0, the sub-network between i_0 and i_n can be bypassed through the Phase I-based path $\rho_{i_0 i_n}$ (the dotted arrows originating from i_0) to reach t_0'. Specifically, by merging this sub-path with two Phase I-based sub-paths, $\rho_{t_0 i_0} := \langle (t_0, i_0) \rangle$ and $\rho_{i_n t_0'} := \langle (i_n, t_0') \rangle$, a relevant path connecting t_0 and t_0', denoted as $\rho_{t_0 t_0'} := \langle (t_0, i_0), \rho_{i_0 i_n}, (i_n, t_0') \rangle$, is obtained.

From the set of all buses reachable from both i and j, the intermediate bus λ that provides the shortest complete path, formed by combining $\rho_{i\lambda}$ and $\rho_{\lambda j}$, should be selected to avoid *non-simple* paths, i.e., traversing a bus more than once, and thereby prevent loose upper bounds. This implies that Phase II can be applied iteratively, utilizing previously established connections to construct paths to new buses located deeper into network-expansion trees.

Example 7. *Consider the bus pair $[t_0, i_n]$ in Figure 4b. Although Phase I is not directly applicable — as no established path connects $i_0 \in N_G(t_0)$ to buses in $N_G(i_n)$ — it can be applied to another bus in the tree, namely i_0, to establish bounds $\overline{\theta}_{t_0 i_0} = 1$ and $\overline{\theta}_{i_0 i_n} = 4$. Using $\lambda = i_0$ as the intermediate bus and merging these paths, the angle difference bound associated with the resulting complete path is determined as:*

$$|\theta_{t_0} - \theta_{i_n}| \leq \underbrace{|\theta_{t_0} - \theta_{i_0}|}_{\leq \overline{\theta}_{t_0 i_0} = 1} + \underbrace{|\theta_{i_0} - \theta_{i_n}|}_{\leq \overline{\theta}_{i_0 i_n} = 4} \leq 5 \tag{15}$$

The path obtained between t_0 and i_n can be merged with the path $\rho_{i_n t_0'}$ in the subsequent iteration of Phase II, thereby building a connection between t_0 to t_0'.

Algorithm 2 Longest Shortest-Path-Connection Algorithm (LSPC) - Phase II

Inputs: $\{i, j\} \in \mathcal{B}, \gamma, \overline{\theta}_{ij}$ ▷ SPP or Phase I bounds (or the γ bound, if they fail)
Outputs: A tighter bound on the angle difference between buses i and j
for all $\lambda \in \mathcal{B} \backslash \{i, j\}$ **do**
  if $\overline{\theta}_{i\lambda} < \gamma$ **and** $\overline{\theta}_{\lambda j} < \gamma$ **then**
    $\overline{\theta}_{ij} \leftarrow \min(\overline{\theta}_{ij}, \overline{\theta}_{i\lambda} + \overline{\theta}_{\lambda j})$
  end if
end for

Proposition 3. *Consider buses i and j that are disconnected in the initial network, with at least one of them belonging to a network-expansion tree. If bounds $\overline{\theta}_{i\lambda}$ and $\overline{\theta}_{\lambda j}$ can be determined for at least one $\lambda \in \mathcal{B} \backslash \{i, j\}$, then $\overline{\theta}_{ij}$ can be obtained as:*

$$|\theta_i - \theta_j| \leq \min_{\lambda \in \{1,\ldots,|\mathcal{B}|\} \backslash \{i,j\}, (\overline{\theta}_{i\lambda} < \gamma \, \wedge \, \overline{\theta}_{\lambda j} < \gamma)} \{\overline{\theta}_{i\lambda} + \overline{\theta}_{\lambda j}\}. \tag{16}$$

Proof. With the angle difference bounds $\overline{\theta}_{i\lambda}$ and $\overline{\theta}_{\lambda j}$ obtained from LSPC Phase I, a complete path between i and j can be formed by linking the paths connecting them to λ. Traversing this full path creates a telescoping effect on the left side, with the sum of the lengths of the sub-paths, $\overline{\theta}_{i\lambda} + \overline{\theta}_{\lambda j}$, yielding an upper bound on the overall angle difference:

$$|\theta_i - \theta_j| \leq \underbrace{|\theta_i - \theta_\lambda|}_{\leq \overline{\theta}_{i\lambda}} + \underbrace{|\theta_\lambda - \theta_j|}_{\leq \overline{\theta}_{\lambda j}} \leq \overline{\theta}_{i\lambda} + \overline{\theta}_{\lambda j}. \tag{17}$$

The presence of multiple intermediary buses λ results in multiple upper bounds. The smallest is selected to yield the strongest valid upper bound from these options, yielding inequality (16). □

5 Conclusion

This paper introduces the Longest Shortest-Path-Connection (LSPC) algorithm to determine tight upper bounds on voltage angle differences between disconnected bus pairs within transmission expansion planning networks. LSPC is a polynomial-time algorithm that overcomes the practical drawbacks of the longest path problem approach, which is the only existing option for generating valid upper bounds for new-bus integration scenarios. The paper demonstrates that path-based VIs initialized with LSPC dominate those derived from LPP or any other larger initial bound. In future work, we will focus on evaluating the computational benefits of the proposed inequalities to solve large-scale power systems expansion planning problems.

Appendix A

We define ρ_{ij} as a candidate path in G connecting buses i and j that includes at least one expansion corridor. Let $N_e(\rho_{ij})$ be the count of such expansion corridors along ρ_{ij}. The complete form of the path-based VIs (see (5)) for potential paths ρ_{ij} is provided in the following proposition. We also demonstrate that when $CR(\rho_{ij})$ exceeds the initial angle difference bound $\overline{\theta}_{ij}$, it may exclude integer-feasible solutions.

Proposition 4. *The following expression, which provides the complete form of the path-based inequalities, is valid only for potential paths $\rho_{ij} \in \mathcal{C}^G_{ij}$ with $CR(\rho_{ij}) < \overline{\theta}_{ij}$:*

$$|\theta_i - \theta_j| \leq CR\left(\rho_{ij}\right) + \left(\overline{\theta}_{ij} - CR\left(\rho_{ij}\right)\right) \left(N_e\left(\rho_{ij}\right) - \sum_{(k,l)\in\rho_{ij}} \mathbb{I}_{kl} y_{kl} \right). \tag{18}$$

Here, the indicator function $\mathbb{I}_{kl}$ takes a value of 1 if $(k,l) \in \rho_{ij} \cap \Omega^1$ and 0 otherwise.

Proof. We will show that absent the condition $CR(\rho_{ij}) < \overline{\theta}_{ij}$, inequalities (18) may eliminate integer-feasible solutions. For succinctness, define the variable $\nu := \left(N_e\left(\rho_{ij}\right) - \sum_{(k,l)\in\rho_{ij}} \mathbb{I}_{kl} y_{kl}\right)$ and use it to reformulate inequality (18) as

$$|\theta_i - \theta_j| \leq (1-\nu)\, CR(\rho_{ij}) + \nu\, \overline{\theta}_{ij}. \tag{19}$$

When $CR(\rho_{ij}) > \overline{\theta}_{ij}$ and $\nu > 1$ (i.e., more than one unbuilt expansion corridor exists along ρ_{ij}), the right-hand side of (19) becomes smaller than the initial upper bound $\overline{\theta}_{ij}$, despite no path shorter than $\overline{\theta}_{ij}$ having been constructed. Consequently, inequality (19) becomes invalid. □

Appendix B

Let $LSPC_{ij}$ represent the angle difference bound obtained using the LSPC algorithm for buses i and j. Proposition 5 evaluates the LSPC-based bound against the LPP-based bound to compare the tightness of the resulting path-based VIs associated with their respective bounds.

Proposition 5. *Let ρ_{ij} be a candidate path connecting buses i and j in G, defining the effective domain for their angle difference as $|\theta_i - \theta_j| \geq CR(\rho_{ij})$. The path-based VI associated with ρ_{ij}, constructed using $LSPC_{ij}$, dominates those derived with LPP as well as other initial bounds greater than $LSPC_{ij}$.*

Proof. Consider the path-based VIs derived using $LSPC_{ij}$ and $\overline{CR(\rho_{ij})}$ as the initial bound $\overline{\theta}_{ij}$, respectively:

$$\sum_{(k,l)\in\rho_{ij}}^{N_e} \mathbb{I}_{kl} y_{kl} \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{LSPC_{ij} - CR(\rho_{ij})} \tag{20}$$

$$\sum_{(k,l)\in\rho_{ij}}^{N_e} \mathbb{I}_{kl} y_{kl} \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{\overline{CR(\rho_{ij})} - CR(\rho_{ij})} \tag{21}$$

These inequalities can be represented as $\pi x \leq \pi_0^1$ and $\pi x \leq \pi_0^2$. We need to demonstrate that π_0^1, i.e., the right-hand side of (20), is smaller than π_0^2, i.e., the right-hand side of (21) within the effective domain, namely, whenever $|\theta_i - \theta_j| \geq CR(\rho_{ij})$. From the definition of the longest path, we have that $LSPC_{ij} \leq \overline{CR(\rho_{ij})}$, with equality indicating the worst-case scenario in which all simple paths between i and j have identical lengths. Excluding the case of equality, two distinct cases arise for the resulting path-based VIs to be comparable:

Case 1: $CR(\rho_{ij}) = LSPC_{ij} < \overline{CR(\rho_{ij})}$. The path ρ_{ij} offers no further improvement to the LSPC-based bound, and a path-based VI can only be derived setting $\overline{\theta}_{ij} = \overline{CR(\rho_{ij})}$. The VI eliminates solutions from the relaxed problem's space when $|\theta_i - \theta_j| \geq CR(\rho_{ij})$, but this removal is redundant, as the VI $|\theta_i - \theta_j| \leq LSPC_{ij} = CR(\rho_{ij})$ dominates it. Furthermore, the VI becomes ineffective when $|\theta_i - \theta_j| \leq CR(\rho_{ij})$, as it is dominated by the trivial *corridor enumeration* VI, formulated as $\sum_{(k,l)\in\rho_{ij}}^{N_e} \mathbb{I}_{kl} y_{kl} \leq N_e(\rho_{ij})$. This is because

$$N_e(\rho_{ij}) \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{\overline{CR(\rho_{ij})} - CR(\rho_{ij})} \tag{22}$$

(the numerator is guaranteed to be non-negative, and the denominator is positive in this case). Therefore, (21) does not provide a tighter VI than the LSPC-based and the corridor enumeration VI.

Case 2: $CR(\rho_{ij}) < LSPC_{ij} < \overline{CR(\rho_{ij})}$. Both initial bounds $LSPC_{ij}$ and $\overline{CR(\rho_{ij})}$, can be used to derive the path-based VIs (20) and (21), respectively. When $|\theta_i - \theta_j| \leq CR(\rho_{ij})$, both VIs are dominated by $\sum_{(k,l)\in\rho_{ij}} \mathbb{I}_{kl} y_{kl} \leq N_e(\rho_{ij})$, that is,

$$N_e(\rho_{ij}) \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{\overline{CR(\rho_{ij})} - CR(\rho_{ij})} \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{LSPC_{ij} - CR(\rho_{ij})}, \tag{23}$$

since the fractions have an identical numerator, and $\overline{CR(\rho_{ij})} - CR(\rho_{ij}) > LSPC_{ij} - CR(\rho_{ij}) > 0$. Conversely, when $|\theta_i - \theta_j| \geq CR(\rho_{ij})$, the LSPC-based VI dominates the other two, as a larger fraction term is subtracted from $N_e(\rho_{ij})$:

$$N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{LSPC_{ij} - CR(\rho_{ij})} \leq N_e(\rho_{ij}) + \frac{CR(\rho_{ij}) - |\theta_i - \theta_j|}{\overline{CR(\rho_{ij})} - CR(\rho_{ij})} \leq N_e(\rho_{ij}). \tag{24}$$

This demonstrates that when $|\theta_i - \theta_j| \geq CR(\rho_{ij})$, the LSPC-based VI provides a tighter bound than that derived using $\overline{CR(\rho_{ij})}$ or any path longer than $LSPC_{ij}$.

Consequently, within the effective domain, the LPP-based VIs are either dominated by LSPC-based angle difference VIs (case 1) or the path-based VIs with LSPC-based initial bounds (case 2). □

References

[1] Garver, L.L.: Transmission network estimation using linear programming. IEEE Transactions on power apparatus and systems (7), 1688–1697 (1970)

[2] Lumbreras, S., Ramos, A.: The new challenges to transmission expansion planning. survey of recent practice and literature review. Electric Power Systems Research **134**, 19–29 (2016)

[3] Pan, X., Zhao, T., Chen, M., Zhang, S.: Deepopf: A deep neural network approach for security-constrained dc optimal power flow. IEEE Transactions on Power Systems **36**(3), 1725–1735 (2020)

[4] Kargarian, A., Mohammadi, J., Guo, J., Chakrabarti, S., Barati, M., Hug, G., Kar, S., Baldick, R.: Toward distributed/decentralized dc optimal power flow implementation in future electric power systems. IEEE Transactions on Smart Grid **9**(4), 2574–2594 (2016)

[5] Minot, A., Lu, Y.M., Li, N.: A parallel primal-dual interior-point method for dc optimal power flow. In: 2016 Power Systems Computation Conference (PSCC), pp. 1–7 (2016). IEEE

[6] Hörsch, J., Ronellenfitsch, H., Witthaut, D., Brown, T.: Linear optimal power flow using cycle flows. Electric Power Systems Research **158**, 126–135 (2018)

[7] Lumbreras, S., Ramos, A., Sánchez, P.: Automatic selection of candidate investments for transmission expansion planning. International Journal of Electrical Power & Energy Systems **59**, 130–140 (2014)

[8] Oertel, D., Ravi, R.: Complexity of transmission network expansion planning: Np-hardness of connected networks and minlp evaluation. Energy systems **5**(1), 179–207 (2014)

[9] Gallego, R., Monticelli, A., Romero, R.: Transmision system expansion planning by an extended genetic algorithm. IEE Proceedings-Generation, Transmission and Distribution **145**(3), 329–335 (1998)

[10] Oliveira, E.J., Da Silva, I., Pereira, J.L.R., Carneiro, S.: Transmission system expansion planning using a sigmoid function to handle integer investment variables. IEEE Transactions on Power Systems **20**(3), 1616–1621 (2005)

[11] Abdi, H., Moradi, M., Lumbreras, S.: Metaheuristics and transmission expansion planning: A comparative case study. Energies **14**(12), 3618 (2021)

[12] Sahraei-Ardakani, M., Korad, A., Hedman, K.W., Lipka, P., Oren, S.: Performance of ac and dc based transmission switching heuristics on a large-scale polish system. In: 2014 IEEE PES General Meeting— Conference & Exposition, pp. 1–5 (2014). IEEE

[13] Gopalakrishnan, A., Raghunathan, A.U., Nikovski, D., Biegler, L.T.: Global optimization of optimal power flow using a branch & bound algorithm. In: 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 609–616 (2012). IEEE

[14] Sousa, A.S., Asada, E.N.: A heuristic method based on the branch and cut algorithm to the transmission system expansion planning problem. In: 2011 IEEE Power and Energy Society General Meeting, pp. 1–6 (2011). IEEE

[15] Binato, S., Pereira, M.V.F., Granville, S.: A new benders decomposition approach to solve power transmission network design problems. IEEE Transactions on Power Systems **16**(2), 235–240 (2001)

[16] Mohammadi, J., Hug, G., Kar, S.: A benders decomposition approach to corrective security constrained opf with power flow control devices. In: 2013 IEEE Power & Energy Society General Meeting, pp. 1–5 (2013). IEEE

[17] Haffner, S., Monticelli, A., Garcia, A., Mantovani, J., Romero, R.: Branch and bound algorithm for transmission system expansion planning using a transportation model. IEE Proceedings-Generation, Transmission and Distribution **147**(3), 149–156 (2000)

[18] Mégel, O., Andersson, G., Mathieu, J.L.: Reducing the computational effort of stochastic multi-period dc optimal power flow with storage. In: 2016 Power Systems Computation Conference (PSCC), pp. 1–7 (2016). IEEE

[19] Dey, S.S., Kocuk, B., Redder, N.: Node-based valid inequalities for the optimal transmission switching problem. Discrete Optimization **43**, 100683 (2022)

[20] Kocuk, B., Jeon, H., Dey, S.S., Linderoth, J., Luedtke, J., Sun, X.A.: A cycle-based formulation and valid inequalities for dc power transmission problems with switching. Operations Research **64**(4), 922–938 (2016)

[21] Lorca, A., Sun, X.A., Litvinov, E., Zheng, T.: Multistage adaptive robust optimization for the unit commitment problem. Operations Research **64**(1), 32–51 (2016)

[22] Hedman, K.W., Ferris, M.C., O'Neill, R.P., Fisher, E.B., Oren, S.S.: Co-optimization of generation unit commitment and transmission switching with n-1 reliability. IEEE Transactions on Power Systems **25**(2), 1052–1063 (2010)

[23] Tsamasphyrou, P., Renaud, A., Carpentier, P.: Transmission network planning under uncertainty with benders decomposition. In: Optimization: Proceedings of the 9th Belgian-French-German Conference on Optimization Namur, September 7–11, 1998, pp. 457–472 (2000). Springer

[24] Skolfield, J.K., Escobar, L.M., Escobedo, A.R.: Derivation and generation of path-based valid inequalities for transmission expansion planning. Annals of Operations Research **312**(2), 1031–1049 (2022)

[25] Villasana, R.V.: Transmission Network Planning Using Linear and Linear Mixed Integer Programming. Rensselaer Polytechnic Institute, United States (1984)

[26] Rahmani, M.: Study of new mathematical models for transmission expansion planning problem. PhD thesis, Ilha Solteira UNESP São Paulo, Brazil (2013)

[27] Kirby, B.: Evaluating transmission costs and wind benefits in texas: Examining the ercot crez transmission study. The Wind Coalition and Electric Transmission Texas, LLC,

Texas PUC Docket (33672) (2007)

[28] Du, P.: Renewable integration at ercot. In: Renewable Energy Integration for Bulk Power Systems: ERCOT and the Texas Interconnection, pp. 1–26. Springer, Switzerland (2023)

[29] Azmy, A.M., Erlich, I.: Impact of distributed generation on the stability of electrical power system. In: IEEE Power Engineering Society General Meeting, 2005, pp. 1056–1063 (2005). IEEE

[30] Moulin, L.S., Poss, M., Sagastizábal, C.: Transmission expansion planning with re-design. Energy systems **1**, 113–139 (2010)

[31] Schrijver, A., *et al.*: Combinatorial Optimization: Polyhedra and Efficiency vol. 24. Springer, Berlin (2003)

[32] Wolsey, L.A., Nemhauser, G.L.: Integer and Combinatorial Optimization. John Wiley & Sons, Hoboken, NJ, USA (2014)

[33] Albert, R., Albert, I., Nakarado, G.L.: Structural vulnerability of the north american power grid. Physical review E **69**(2), 025103 (2004)

[34] Merris, R.: Graph Theory. John Wiley & Sons, New York, NY, USA (2011)