On the Use of Entropy-Like Functions for Solving the Network Flow Problem with Application to Electric Power Systems

Madi Zholbaryssov, *Student Member, IEEE*, Alejandro D. Domínguez-García, *Member, IEEE*

***Abstract*— In this paper, we consider a variation of the minimum-cost network flow problem (NFP) with additional non-convex cycle constraints on nodal variables; this problem has relevance in the context of optimizing power flows in electric power networks. We propose one approach to tackle the NFP that relies on solving a convex approximation of the problem, obtained by augmenting the cost function with an entropy-like term to relax the non-convex constraints. We show that the approximation error, for which we give an upper bound, can be made small enough for practical use. An alternative approach is to solve the classical NFP, i.e., without the cycle constraints, and solve a separate optimization problem afterwards in order to recover the actual flows satisfying the cycle constraints; the solution of this separate problem maximizes the individual entropies of the cycles. Finally, we validate the practical usefulness of the theoretical results through a numerical example, where the IEEE 39–bus system is used as a test bed.**

I. Introduction

We consider a system comprised of multiple nodes that are interconnected via links (edges). Each node represents either a supplier or a consumer of a certain commodity, and links allow this commodity to flow between nodes; assuming lossless links, the difference of out-flows and in-flows at each node needs to be equal to the amount of commodity injected (positive if the corresponding node is a supplier, and negative otherwise). Both the nodal injections and the flows on each link are constrained to lie within a certain interval. In this setting, a typical objective is to minimize a certain cost function the arguments of which are the link flows and possibly the nodal commodity injections. This problem is commonly referred to as the minimum-cost network flow problem (see, e.g., [1]–[8]). In this paper, in addition to the aforementioned constraints in the classical formulation, we assume that the flow of the commodity along every link is a function of the difference of certain nodal variables associated with the link end nodes. We further assume that this relation between flows and nodal variables imposes equality constraints on the flows along every cycle in the network.

The network flow problem setting considered in this paper is of interest in electric power networks (see, e.g., [6]–[11]). Specifically, the optimal generation dispatch problem can be formulated as a network flow problem with the voltage phase angles identified as the nodal variables governing the power flows along the electrical lines. In this setting, there is a cost associated to the power provided by each generator, and the problem is to assign generator power output set-points so as to satisfy the total load while (i) respecting generation capacity limits, and thermal limits on the electrical lines; and (ii) ensuring that the resulting phase angles are such that the angle differences across the edges of every cycle add up to zero. This problem formulation shares many similarities with the standard optimal power flow problem formulation (see, e.g., [12]) in the sense that both problems are focused on finding the optimal nodal injections that need to balance the net power flow at each node. However, the main difference is that we solve the considered problem in terms of the power flows, while the optimal power flow problem is typically formulated and solved in terms of the nodal variables—voltage magnitudes and phase angles.

In the literature, the classical network flow problem is a convex optimization problem (see, e.g., [1]–[5]). The aforementioned additional cycle constraints on the flows that we consider in our network flow problem formulation are non-convex; we address this non-convexity using two approaches. In the first one, we define an entropy-like function for cycles in the network, and add this function as a penalty term to the cost function; this guarantees a satisfaction of the cycle constraints at the optimal solution. However, this penalty-based approach introduces a certain approximation of the original cost function; we show that this approximation can be made as accurate as possible by decreasing the penalty weight. The idea of relaxing the non-convex constraints using entropy-like penalty functions can also be found in several computer vision applications and combinatorial problems (see, e.g., [13], [14]), as well as in matrix scaling problems [15]. In the second approach, we solve a modified version of the original problem, where we remove the cycle constraints, and in order to obtain the actual flows satisfying the cycle constraints, we solve a separate optimization problem, the solution of which maximizes the value of the individual entropy-like functions associated to each link in a cycle.

II. Preliminaries

In this section, we first introduce a few graph-theoretic notions that are relevant to our work; the reader is referred to [16] for a more in-depth discussion on the topic. Then, we formulate the network flow problem under cycle constraints, and tailor it to electric power system applications.

A. Graph-Theoretic Notions

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote an undirected graph, where $\mathcal{V} := \{1, 2, \dots, n\}$ is the set of nodes, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, with $\{l, j\} \in \mathcal{E}$ if nodes l and j are connected. Let

M. Zholbaryssov and A. D. Domínguez-García are with the ECE Department of the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: {zholbar1,aledan}@ILLINOIS.EDU.

$\gamma_{lj} > 0$ denote some weight assigned to edge $\{l,j\} \in \mathcal{E}$, i.e., $\gamma_{lj} > 0$ if $\{l,j\} \in \mathcal{E}$, and zero, otherwise. Let $\mathcal{N}(i)$ denote the set of neighbors of a node i, i.e., $\mathcal{N}(i) = \{j \in \mathcal{V} : \{i,j\} \in \mathcal{E}\}$.

Define a one-to-one map, $\mathbb{I} : \mathcal{E} \to \mathbb{R}$, such that every e in the set $\{1, 2, \ldots, |\mathcal{E}|\}$ is arbitrarily assigned to exactly one edge $\{l,j\} \in \mathcal{E}$, i.e., $\mathbb{I}(\{l,j\}) = e$. We also arbitrarily assign an orientation to each edge in $\mathcal{E}$, and denote by (l,j) the edge oriented from l to j, i.e., (l,j) is incident to j. Let $\vec{\mathcal{E}}$ denote the set of oriented edges that result from this arbitrarily chosen orientation (note that $|\vec{\mathcal{E}}| = |\mathcal{E}| =: \ell$); then, based on this orientation, we can define a node-to-edge incidence matrix, $M \in \mathbb{R}^{n \times \ell}$, as follows. For each $e = \mathbb{I}(\{i,j\})$, $M_{ie} = 1$ and $M_{je} = -1$, if $(i,j) \in \vec{\mathcal{E}}$, and $M_{ie} = 0$ and $M_{je} = 0$, if $\{i,j\} \notin \mathcal{E}$.

Let $\mathcal{T}$ denote a spanning tree in $\mathcal{G}$; then, an edge of $\mathcal{G}$, that does not belong to $\mathcal{T}$, and the path in $\mathcal{T}$ between the vertices of this edge form a cycle, referred to as an undirected *fundamental cycle* [16, Definition 2-8]. Then, since there are $\ell - n + 1$ edges which do not belong to $\mathcal{T}$, there exist $c := \ell - n + 1$ fundamental cycles in $\mathcal{G}$.

Let $\mathcal{C}_i = (\mathcal{V}_i, \mathcal{E}_i)$, $i = 1, \ldots, c$, denote an undirected fundamental cycle in $\mathcal{G}$, formed with respect to $\mathcal{T}$, where the vertex set $\mathcal{V}_i = \{i_1, i_2, \ldots, i_{d_i}\} \subseteq \mathcal{V}$, $d_i = |\mathcal{V}_i|$, and the edge set $\mathcal{E}_i = \{\{i_1,i_2\}, \{i_2,i_3\}, \ldots, \{i_{d_i}, i_1\}\} \subseteq \mathcal{E}$. The oriented counterpart of $\mathcal{C}_i$ is denoted by $\vec{\mathcal{C}}_i$ with the orientation of the edges being induced by traversing the cycle clockwise, and the directed edge set defined as $\vec{\mathcal{E}}_i$. Additionally, we can define the entries of the fundamental cycle matrix, denoted by $N \in \mathbb{R}^{c \times \ell}$, as follows:

$$N_{ie} = \begin{cases} 1 & \text{if } (i_k, i_{k+1}) \in \vec{\mathcal{E}}_i, M_{i_k e} = 1, \\ -1 & \text{if } (i_k, i_{k+1}) \in \vec{\mathcal{E}}_i, M_{i_k e} = -1, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where $e = \mathbb{I}(\{i_k, i_{k+1}\})$; then, we have the following result (see, e.g., [16, Theorem 4-6]):

$$MN^\top = \mathbf{0}. \quad (2)$$

B. Network Flows Under Cycle Constraints

Consider a system composed of multiple nodes that are interconnected via edges, the topology of which is described by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ containing oriented fundamental cycles, $\vec{\mathcal{C}}_1, \vec{\mathcal{C}}_1, \ldots, \vec{\mathcal{C}}_c$, formed with respect to a spanning tree, $\mathcal{T}$, of $\mathcal{G}$. Let u_i denote the *injection* of a commodity into the network via node $i \in \mathcal{V}$, and let f_{lj} denote the *flow* of this commodity along edge $(l,j) \in \vec{\mathcal{E}}$; the injection u_i and the flows incident to node i are related through the flow balance equation given below

$$u_i = \sum_{(i,j) \in \vec{\mathcal{E}}} f_{ij} - \sum_{(l,i) \in \vec{\mathcal{E}}} f_{li}.$$

The relation between the nodal injections and edge flows can be compactly written as

$$u = Mf, \quad (3)$$

where $u = [u_1, \ldots, u_n]^\top$, $f = [\{f_{lj}\}_{(l,j) \in \vec{\mathcal{E}}}]^\top$. For a given u, it follows from (2) that any two solutions of (3), $f^{(1)}$ and $f^{(2)}$, can be related as follows:

$$f^{(1)} = f^{(2)} + N^\top \mu, \quad (4)$$

where $\mu \in \mathbb{R}^c$.

In the adopted network flow model, injections and flows are constrained to lie within some upper and lower bounds: $\underline{u}_i \le u_i \le \overline{u}_i$, $i \in \mathcal{V}$, and $\underline{f}_{lj} \le f_{lj} \le \overline{f}_{lj}$, $(l,j) \in \vec{\mathcal{E}}$. In addition, we assume that the edge flows depend on certain nodal variables, denoted by θ_i, $i \in \mathcal{V}$, via a functional relation of the form

$$f_{lj} = \gamma_{lj} g(\theta_l - \theta_j),$$

where $g(\cdot) : \mathcal{B} \to \mathcal{Y}$ is a continuous and odd function, i.e., $g(x) = -g(-x)$, with

$$\mathcal{B} = \{x \in \mathbb{R} : -\phi \le x \le \phi\}, \quad (5)$$

$$\mathcal{Y} = \{y \in \mathbb{R} : -g(\phi) \le y \le g(\phi)\}, \quad (6)$$

for some constant $\phi \in \mathbb{R}$. The relation between the nodal variables and the edge flows can be compactly written as

$$f = \Gamma g(M^\top \theta), \quad (7)$$

where $\Gamma \in \mathbb{R}^{\ell \times \ell}$ is a diagonal matrix with $\Gamma_{ee} = \gamma_{lj}$, $e = \mathbb{I}(\{l,j\})$, and $g(M^\top \theta) = [\{g(\theta_l - \theta_j)\}_{(l,j) \in \vec{\mathcal{E}}}]^\top \in \mathbb{R}^\ell$.

In electric power networks, u_i corresponds to the active power injection at node $i \in \mathcal{V}$, and assuming losses can be neglected, $f_{lj} = \gamma_{lj} g(\theta_l - \theta_j)$ is the flow of active power along the electrical line connecting nodes l and j, where $g(\theta_l - \theta_j) = \sin(\theta_l - \theta_j)$ for inductive lines, with θ_i being the phase angle of the voltage phasor at node i, and $\gamma_{lj} = V_l V_j B_{lj}$, with V_i being the voltage amplitude, and $-B_{lj}$ being the susceptance of the electrical line. To enforce the line capacity constraint, the voltage phase angle difference, $\theta_l - \theta_j$, across each line is limited by $\phi < \frac{\pi}{2}$. Thus, in the remainder, instead of using a generic $g(\cdot)$, we utilize $g(\cdot) = \sin(\cdot)$, because of its relevance in electric power system problems.

C. Network Flow Problem Formulation

We are interested in finding the injections and nodal variables that solve the following problem, which we refer to as P0:

$$\mathbf{P0}: \min_{u, \theta \in \mathbb{R}^n} \sum_{i \in \mathcal{V}} F_i(u_i) \quad (8)$$

$$\text{subject to } u = M\Gamma \sin(M^\top \theta), \quad (9)$$

$$\underline{u} \le u \le \overline{u}, \quad (10)$$

$$\|M^\top \theta\|_\infty \le \phi, \quad (11)$$

where $F_i(\cdot)$ is a convex, strictly increasing and continuously differentiable cost function, $\|M^\top \theta\|_\infty = \max_{(i,j) \in \vec{\mathcal{E}}} |\theta_i - \theta_j|$; this problem is hard to solve because of the non-convex constraint in (9).

Instead of directly solving P0 in terms of the injections and nodal variables, we introduce the following problem, formulated in terms of the injections and flows:

$$\mathbf{P}1: \min_{u\in\mathbb{R}^n, f\in\mathbb{R}^\ell} \sum_{i\in\mathcal{V}} F_i(u_i) \tag{12}$$
$$\text{subject to } u = Mf, \tag{13}$$
$$\underline{u} \le u \le \overline{u}, \tag{14}$$
$$\underline{f} \le f \le \overline{f}, \tag{15}$$
$$N \arcsin(\Gamma^{-1} f) = 0, \tag{16}$$

where $\arcsin(\Gamma^{-1}f) = [\{\arcsin(f_{lj}/\gamma_{lj})\}_{(l,j)\in\vec{\mathcal{E}}}]^\top \in \mathbb{R}^\ell$, $\overline{f} = [\{\overline{f}_{lj}\}_{(l,j)\in\vec{\mathcal{E}}}]^\top \in \mathbb{R}^\ell$, $\underline{f} = [\{\underline{f}_{lj}\}_{(l,j)\in\vec{\mathcal{E}}}]^\top \in \mathbb{R}^\ell$, with $\overline{f}_{lj} = -\underline{f}_{lj} = \gamma_{lj}\sin(\phi)$. The cycle constraint in (16) ensures that the voltage phase angles, that correspond to the power flows, add up to zero along each cycle, i.e.,

$$\sum_{(l,j)\in\vec{\mathcal{E}}_i} (\theta_l - \theta_j) = 0, \quad \forall \mathcal{C}_i \in \mathcal{G}.$$

Remark 1. The cycle constraint in (16) can be rewritten explicitly for each cycle $\mathcal{C}_i$ as follows:

$$\sum_{e=1}^{\ell} n_e^{(i)} \arcsin\left(\frac{f_{lj}}{\gamma_{lj}}\right) = 0, \tag{17}$$

where $n^{(i)}$ denotes the i-th column of $N^\top$, and $e = \mathbb{I}(\{l,j\})$. Based on the orientation of $\vec{\mathcal{C}}_i$, and

$$\begin{aligned} -f_{lj} &= -\gamma_{lj}\sin(\theta_l - \theta_j) \\ &= \gamma_{lj}\sin(\theta_j - \theta_l) =: f_{jl},\ (l,j)\in\vec{\mathcal{E}}, \end{aligned} \tag{18}$$

an alternative expression for (17) is given by

$$\sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{f_{lj}}{\gamma_{lj}}\right) = 0. \tag{19}$$

In the remainder, we use both (17) and (19) as deemed appropriate. ■

Let $\mathcal{Q}$ and $\mathcal{S}$ denote the constraint sets of P0 and P1, respectively, given below:

$$\mathcal{Q} := \{(u,\theta) : (u,\theta) \text{ satisfies (9) – (11)}\},$$

and

$$\mathcal{S} := \{(u,f) : (u,f) \text{ satisfies (13) – (16)}\}.$$

The following result shows that P0 and P1 are equivalent.

Lemma 1. Let $\mathcal{U}_0$ denote the set of all u's for which there exists θ such that $(u,\theta)\in\mathcal{Q}$, and let $\mathcal{U}_1$ denote the set of all u's for which there exists an f such that $(u,f)\in\mathcal{S}$. Then, $\mathcal{U}_0 \equiv \mathcal{U}_1$.

Proof. First, we show that $\mathcal{U}_0 \subseteq \mathcal{U}_1$. Suppose, $(\tilde{u},\tilde{\theta}) \in \mathcal{Q}$. Choose

$$\tilde{f}_{lj} = \gamma_{lj}\sin(\tilde{\theta}_l - \tilde{\theta}_j);$$

then, obviously, the constraints in (13) – (15) are satisfied. Also, along each $\mathcal{C}_i$, we have that

$$\sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{\tilde{f}_{lj}}{\gamma_{lj}}\right) = \sum_{(l,j)\in\vec{\mathcal{E}}_i} (\tilde{\theta}_l - \tilde{\theta}_j) = 0; \tag{20}$$

therefore, $(\tilde{u},\tilde{f}) \in \mathcal{S}$.

Now, we show that $\mathcal{U}_1 \subseteq \mathcal{U}_0$. Suppose, $(\tilde{u},\tilde{f}) \in \mathcal{S}$. We show how to construct $\tilde{\theta}$ from $\tilde{f}$. We assign any value to $\tilde{\theta}_1$, and compute, for each neighbor of node 1, $\tilde{\theta}_j = \tilde{\theta}_1 - \arcsin(\tilde{f}_{1j}/\gamma_{1j})$, $\forall j \in \mathcal{N}(1)$. Then, we perform the same computations for the neighbors of each $j \in \mathcal{N}(1)$, and continue this process. For cyclic networks, this computation continues until $\tilde{\theta}_l$ is computed from two different nodes i and k simultaneously. In order to successfully construct $\tilde{\theta}$, it is enough to show that $\tilde{\theta}_l$ as computed by node i, which is equal to $\tilde{\theta}_i - \arcsin(\tilde{f}_{il}/\gamma_{il})$, matches $\tilde{\theta}_l$ as computed by node k, which is equal to $\tilde{\theta}_k - \arcsin(\tilde{f}_{kl}/\gamma_{kl})$. Indeed, let $\mathcal{C}_1$ denote the cycle containing l, i and k, and suppose that traversing the path in the clockwise direction from i to k along $\vec{\mathcal{C}}_1$ does not go through node l, then, we have that

$$\tilde{\theta}_i - \tilde{\theta}_k = \sum_{(l,j)\in\vec{\mathcal{E}}'_1} \arcsin(\tilde{f}_{lj}/\gamma_{lj}), \tag{21}$$

where $\vec{\mathcal{E}}'_1 = \vec{\mathcal{E}}_1 \backslash \{(k,l)\cup(l,i)\}$, and, therefore,

$$\begin{aligned} &\tilde{\theta}_i - \arcsin\left(\frac{\tilde{f}_{il}}{\gamma_{il}}\right) - \tilde{\theta}_k + \arcsin\left(\frac{\tilde{f}_{kl}}{\gamma_{kl}}\right) \\ &= \sum_{(l,j)\in\vec{\mathcal{E}}'_1} \arcsin\left(\frac{\tilde{f}_{lj}}{\gamma_{lj}}\right) - \arcsin\left(\frac{\tilde{f}_{il}}{\gamma_{il}}\right) + \arcsin\left(\frac{\tilde{f}_{kl}}{\gamma_{kl}}\right) \\ &= 0, \end{aligned} \tag{22}$$

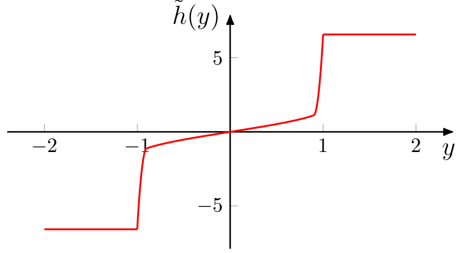
where in the last equality we used the fact that $\tilde{f}$ satisfies the cycle constraint in (16). Therefore, $\tilde{\theta}_l$ as computed by node i matches $\tilde{\theta}_l$ as computed by node k. Then, obviously, the constraints in (9) – (11) are satisfied; therefore, $\mathcal{U}_0 \equiv \mathcal{U}_1$. □

III. Entropy-Like Penalty Functions for Solving the NFP Under Cycle Constraints

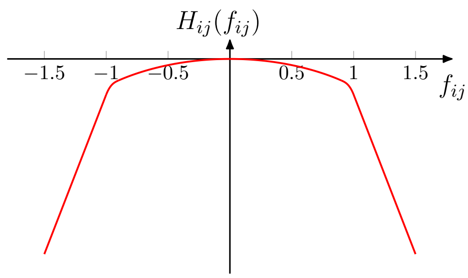
In this section, we formulate a convex approximation of the original problem P1. Although the solution of this new convex problem is an approximate solution of the original problem, reducing the penalty weight yields a better approximation.

First, we note that the set $\bigcup_{i=1}^{c} \mathcal{E}_i$ includes every edge that belongs to a cycle. Let $\breve{\mathcal{E}}$ denote the set of all edges from $\bigcup_{i=1}^{c} \mathcal{E}_i$ directed with an orientation consistent with that of the incidence matrix M, i.e. $(i,j)\in\breve{\mathcal{E}}$ only if $(i,j)\in\vec{\mathcal{E}}$. Then, we introduce the following *entropy-like function,*[1] denoted

[1]The function $H(f)$ has certain properties typical of entropy functions; for example, the negative of $H(f)$ is a convex integral, $H(f)$ is concave, and maximizing $H(f)$ allows to achieve a system-wide objective, which is, in our case, a satisfaction of the non-convex cycle constraints in (16) (to be shown later) [15], [17], [18]. This is the reason why we use the term entropy-like function to refer to $H(\cdot)$.



(a)



(b)

Fig. 1: (a) Plot of $\tilde{h}(y)$ for $\gamma_{ij} = 1$, $\epsilon = 0.1$, $\kappa = 5 \times 10^3$, and (b) corresponding entropy-like component $H_{ij}(f_{ij})$.

by $H(\cdot)$, for the cycles in the network, taking value:

$$H(f) := -\sum_{(i,j)\in\check{\mathcal{E}}} \int_0^{f_{ij}} \tilde{h}\left(\frac{x}{\gamma_{ij}}\right) dx, \tag{23}$$

where

$$\tilde{h}(y) = \begin{cases} -\phi - \kappa\epsilon^2, & y \le \underline{y}, \\ \arcsin(y) - \kappa(y + \underline{y}_\epsilon)^2, & \underline{y} \le y \le \underline{y}_\epsilon, \\ \arcsin(y), & \underline{y}_\epsilon \le y \le \overline{y}_\epsilon, \\ \arcsin(y) + \kappa(y - \overline{y}_\epsilon)^2, & \overline{y}_\epsilon \le y \le \overline{y}, \\ \phi + \kappa\epsilon^2, & \overline{y} \le y, \end{cases} \tag{24}$$

with $\overline{y} := \sin(\phi)$, $\underline{y} := -\sin(\phi)$, $\overline{y}_\epsilon := \sin(\phi) - \epsilon$, $\underline{y}_\epsilon := -\sin(\phi) + \epsilon$, for some parameters $\epsilon > 0$, and $\kappa > 0$ (to be discussed later). From (24), it can be seen that $H(f)$ penalizes the flows violating the box constraints for lines along cycles, and κ causes each entropy-like component to increase drastically when the flows are outside of the box constraints. An example of $\tilde{h}(y)$ and the entropy-like component for edge (i, j), denoted by

$$H_{ij}(f_{ij}) := -\int_0^{f_{ij}} \tilde{h}\left(\frac{x}{\gamma_{ij}}\right) dx,$$

is given in Figure 1.

In the remainder, we relax the non-convex cycle constraints and box constraints on the flows along the cycles, and instead maximize the entropy-like function $H(f)$. To be more precise, to deal with the non-convexity in P1, we add $H(f)$ as a penalty term to the cost function to replace the cycle and box constraints in (15) – (16) for flows along the cycles; this results in the following optimization problem:

$$\mathbf{P}2: \min_{u\in\mathbb{R}^n, f\in\mathbb{R}^\ell} \sum_{i\in\mathcal{V}} F_i(u_i) - \rho H(f) \tag{25}$$
$$\text{subject to } u = Mf, \tag{26}$$
$$\underline{u} \le u \le \overline{u}, \tag{27}$$
$$\underline{f}_{ij} \le f_{ij} \le \overline{f}_{ij}, \ (i,j) \in \vec{\mathcal{E}} \setminus \check{\mathcal{E}}, \tag{28}$$

where $\rho > 0$ is the penalty weight. In (28), we have the constraints on the flows along the edges, which do not belong to any cycle. The result in the next lemma establishes that P2 is convex.

Lemma 2. $H(\cdot)$ defined in (23) is concave.

Proof. To show that $H(f)$ is concave, it is enough to show that $H_{ij}(f_{ij})$ is concave, which is true since

$$\frac{\partial H_{ij}(f_{ij})}{\partial f_{ij}} = -\tilde{h}\left(\frac{f_{ij}}{\gamma_{ij}}\right) \tag{29}$$

is non-increasing in f_{ij}. □

Now, we state one of the main results of this work, for which we do not provide a proof due to space limitations.

Proposition 1. Let $(u^{(1)*}, f^{(1)*})$ and $(u^{(2)*}, f^{(2)*})$ denote the solutions of P1 and P2, respectively. If

$$\kappa = \frac{3}{\rho\epsilon^3}\Big(\sum_{i\in\mathcal{V}} F_i(\overline{u}_i) - \rho H(\overline{f} - \epsilon\gamma)\Big), \tag{30}$$

then, the minimum of P2, $(u^{(2)*}, f^{(2)*})$, satisfies the cycle constraints in (16) if

$$\underline{f}_{ij} + \epsilon\gamma_{ij} \le f_{ij}^{(2)*} \le \overline{f}_{ij} - \epsilon\gamma_{ij}, \ \forall (i,j) \in \check{\mathcal{E}},$$

and the box constraints for flows in (15) if

$$\underline{f}_{ij} + \epsilon\gamma_{ij} \le f_{ij}^{(1)*} \le \overline{f}_{ij} - \epsilon\gamma_{ij}, \ \forall (i,j) \in \check{\mathcal{E}}.$$

Proposition 1 establishes that the solution of P2 is a feasible point of the non-convex constraint set of P1, if $f^{(2)*}$ strictly satisfies the box constraints along each cycle. This solution can be made as close to the solution of P1 as possible by making the penalty weight ρ sufficiently small, as shown in the next lemma.

Lemma 3. Let $(u^{(1)*}, f^{(1)*})$ and $(u^{(2)*}, f^{(2)*})$ denote the solutions of P1 and P2, respectively. If $\underline{f} + \epsilon\gamma \le f^{(1)*} \le \overline{f} - \epsilon\gamma$, then,

$$\sum_{i\in\mathcal{V}} F_i(u_i^{(2)*}) - \sum_{i\in\mathcal{V}} F_i(u_i^{(1)*}) \le \rho(H(f^{(2)*}) - H(\overline{f} - \epsilon\gamma)). \tag{31}$$

Proof. Since $(u^{(2)*}, f^{(2)*})$ minimizes P2, we have that

$$\sum_{i\in\mathcal{V}} F_i(u_i^{(2)*}) - \rho H(f^{(2)*}) \le \sum_{i\in\mathcal{V}} F_i(u_i^{(1)*}) - \rho H(f^{(1)*}). \tag{32}$$

Since $H(\overline{f} - \epsilon\gamma) = H(\underline{f} + \epsilon\gamma)$ and $\underline{f} + \epsilon\gamma \le f^{(1)*} \le \overline{f} - \epsilon\gamma$, $H(f^{(1)*}) \ge H(\overline{f} - \epsilon\gamma)$ and (31) follows from (32) after rearranging the terms. □

IV. Entropy Maximization for Finding Feasible Flows for Cyclic Networks

In this section, due to the space limitation, we only consider networks with edge-disjoint cycles, i.e., any two cycles in the network do not have edges in common. Some recent work (see, e.g., [6]–[8]) has been done to compute the optimal injections and flows without satisfying the cycle constraints for nodal variables. Here, we assume that there is an optimal $(u^{(3)*}, f^{(3)*})$ that has already been computed by solving a modified version of P1 formed by dropping the cycle constraints. We propose a method that obtains the actual flows, denoted by f^*, corresponding to $u^{(3)*}$, which satisfy the cycle constraints in P1, assuming that there exist such flows.

More formally, let $(u^{(3)*}, f^{(3)*})$ denote an optimal solution to a modified version of P1, where the cycle constraints have been removed, referred to as P3 given below:

$$\textbf{P3}: \min_{u\in\mathbb{R}^n, f\in\mathbb{R}^\ell} \sum_{i\in\mathcal{V}} F_i(u_i) \tag{33}$$

$$\text{subject to } u = Mf, \tag{34}$$

$$\underline{u} \le u \le \overline{u}, \tag{35}$$

$$\underline{f} \le f \le \overline{f}. \tag{36}$$

It is easy to see from (4) that $f^* = f^{(3)*} + N^\top \mu^*$ for some $\mu^* \in \mathbb{R}^c$, where μ_i^* corresponds to cycle $\mathcal{C}_i$. We also note that since $(u^{(3)*}, f^{(3)*})$ is a minimum of P3, $(u^{(3)*}, f^*)$ is a minimum of P1. In the following, we show how to construct f^* from $f^{(3)*}$ by computing μ^*. The problem of obtaining μ^* will be set up as an entropy maximization problem.

We introduce the following entropy-like function, denoted by $H_i(\cdot)$, for cycle $\mathcal{C}_i$:

$$H_i(\mu_i) = -\sum_{(l,j)\in\vec{\mathcal{E}}_i} \int_0^{\mu_i} \arcsin\left(\frac{f_{lj}^{(3)*} + x}{\gamma_{lj}}\right) dx, \tag{37}$$

with the domain

$$\mathcal{D}_i = \{x \in \mathbb{R} : \underline{f}_{lj} \le f_{lj}^{(3)*} + x \le \overline{f}_{lj}, \forall (l,j) \in \vec{\mathcal{E}}_i\}.$$

In the next lemma, we show that $H_i(\mu_i)$ is strictly concave.

Lemma 4. Consider the entropy-like function $H_i(\mu_i)$ defined in (37) for cycle $\mathcal{C}_i$. Then, $H_i(\cdot)$ is strictly concave.

Proof. Since

$$\nabla H_i(\mu_i) = \frac{\partial H_i(\mu_i)}{\partial \mu_i} = -\sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{f_{lj}^{(3)*} + \mu_i}{\gamma_{lj}}\right)$$

is strictly decreasing in $\mu_i \in \mathcal{D}_i$, $H_i(\cdot)$ is strictly concave. □

For a given $(u^{(3)*}, f^{(3)*})$ that solves P3, the next proposition shows that in order to obtain μ^* so that $(u^{(3)*}, f^{(3)*} + N^\top \mu^*)$ solves P1, we need to maximize the entropy for every cycle.

Proposition 2. Consider the entropy-like function $H_i(\mu_i)$ defined in (37) for cycle $\mathcal{C}_i$. Then, $(u^{(3)*}, f^{(3)*} + N^\top \mu^*)$ is a solution of P1 if

$$\mu_i^* = \operatorname*{argmax}_{\mu_i\in\mathcal{D}_i} H_i(\mu_i), \; i = 1, \ldots, c,$$

and $\mu^* \in \text{int}(\mathcal{D})$, where $\mathcal{D}$ is the Cartesian product of the sets $\mathcal{D}_1, \ldots, \mathcal{D}_c$, and $\text{int}(\mathcal{D})$ denotes the interior of $\mathcal{D}$.

Proof. Note that $\nabla H_i(\mu_i) = 0$ is the cycle constraint for $\mathcal{C}_i$, which is satisfied at $\mu_i = \mu_i^*$, i.e.,

$$\nabla H_i(\mu_i^*) = 0, \; i = 1, \ldots, c, \tag{38}$$

if and only if $(u^{(3)*}, f^{(3)*} + N^\top \mu^*)$ is a solution of P1. Since, by Lemma 4, $H_i(\cdot)$ is concave, the first-order concavity condition in (38) is satisfied if $\mu^* \in \text{int}(\mathcal{D})$ and $H_i(\mu_i)$ achieves a maximum at μ_i^*, $\forall i$. □

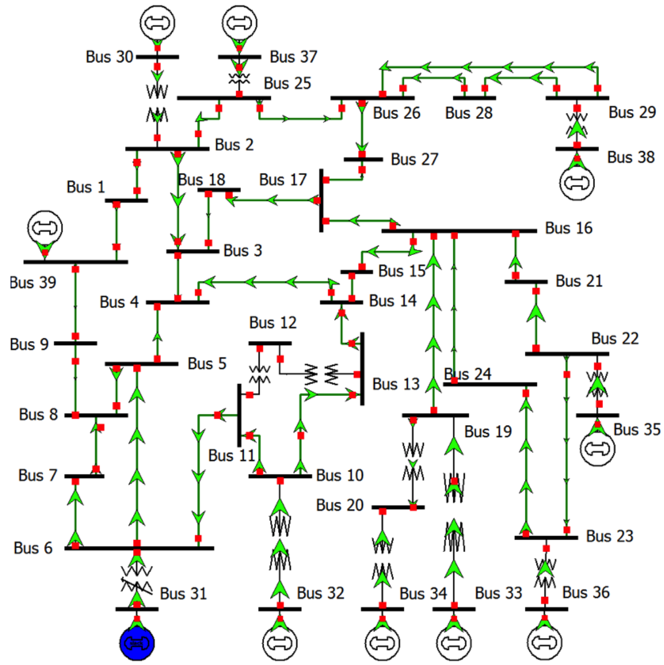


Fig. 2: IEEE 39-bus test system [20].

V. Simulations

In this section, we validate the usefulness of our theoretical results via simulations using the IEEE 39–bus test system, the topology of which is shown in Fig. 2, with all model parameters taken from [19].

A. Entropy Penalty-Based Approach

We first apply the entropy penalty-based approach from Section III and solve P2 for different values of ρ using the standard MATLAB solver *fmincon*; these values are compared with the global solution of P1, computed by solving P3 using the standard MATLAB solver *quadprog* and by applying the entropy maximization approach in Section IV to find the actual flows, which satisfy the cycle constraints, using *fmincon*.

Table I gives the values of the cost function $F(u) := \sum_{i\in\mathcal{V}} F_i(u_i)$, the parameters of which were taken from [19],

TABLE I: Simulation results for the entropy-like penalty-based approach.

ρ	$F(u)$, \$ (in pu)	ϵ, rad
Global solution	85.804	0
1	85.815	8×10^{-4}
10	86.387	1.4×10^{-4}
100	94.198	2.6×10^{-5}
1000	103.13	0

TABLE II: Simulation results for finding the feasible flows.

Global $F(u)$, \$ (in pu)	85.804
$F(u)$ from P3, \$ (in pu)	85.804
$[\epsilon_1, \epsilon_2, \epsilon_3]$, rad	$[-2.7, -1.4, 2.3] \times 10^{-5}$
$[\epsilon_4, \epsilon_5, \epsilon_6, \epsilon_7]$, rad	$[1.4, 2.5, 2.2, 9.6] \times 10^{-5}$

for different values of ρ after 150 iterations. As expected from Lemma 3, sufficiently small ρ yields a solution close to the global one. Also, to show that by solving P2 we satisfy the cycle constraints,

$$\sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{f_{lj}}{\gamma_{lj}}\right) = 0$$

for all cycles, in Table I we provide

$$\epsilon := \max_{1\le i\le c} \Big| \sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{f_{lj}}{\gamma_{lj}}\right) \Big|, \qquad (39)$$

for different values of ρ. Simulation results confirm the results in Proposition 1, that is, $\epsilon[k] \to 0$, as (u, f) converges to the optimal solution.

B. Finding Feasible Flows

Next, we applied the entropy maximization approach from Section IV to compute the actual flows for the solution of P3 to satisfy the cycle constraints. In Table II, we present the values of the cost $F(u)$ corresponding to the solution of P3 obtained using *quadprog*, after about 10 iterations, and the cycle constraint violation levels, denoted by ϵ_i and defined as

$$\epsilon_i := \sum_{(l,j)\in\vec{\mathcal{E}}_i} \arcsin\left(\frac{f_{lj}}{\gamma_{lj}}\right), \qquad (40)$$

for all seven cycles using *fmincon*, after about 30 iterations. The results in Table II demonstrate that the cycle constraint violation levels go to zero. We note that P3 can be solved faster than P2, because P3 is a quadratic program; this means that solving P3 can be preferable for large-scale problems.

VI. Concluding Remarks

In this paper, we proposed two approaches to solve the network flow problem under non-convex equality constraints on nodal variables along cycles. In the first one, we introduced an entropy-like penalty-based method to obtain a convex approximation of the original non-convex problem; reducing the penalty weight yields a better approximation of the global solution of the original problem. In the second one, we solve the modified version of the original problem, where we drop the cycle constraints, and in order to recover the actual flows satisfying the cycle constraints, we solve a separate optimization problem, the solution of which maximizes the individual entropies of the cycles.

References

[1] L. R. Ford and D. R. Fulkerson, *Flows in networks*. Princeton University Press, 1962.

[2] W. S. Jewell, "Optimal flow through networks with gains," *Operations Research*, vol. 10, no. 4, pp. 476 – 499, 1962.

[3] J. Edmonds and R. M. Karp, "Theoretical improvements in algorithmic efficiency for network flow problems," *J. Assoc. Comput. Mach.*, vol. 19, no. 2, pp. 248–264, 1972.

[4] D. P. Bertsekas, P. A. Hosein, and P. Tseng, "Relaxation methods for network flow problems with convex arc costs," *SIAM Journal on Control and Optimization*, vol. 25, no. 5, pp. 1219–1243, 1987.

[5] A. V. Goldberg, É. Tardos, and R. E. Tarjan, "Network flow algorithms," Defense Technical Information Center, Tech. Rep., 1989.

[6] S. Sojoudi and J. Lavaei, "Convexification of generalized network flow problem with application to power systems," in *Proc. IEEE Conf. Decision and Control*, Dec. 2013, pp. 7552–7559.

[7] S. T. Cady, C. N. Hadjicostis, and A. D. Domínguez-García, "Distributed frequency control of inertia-less ac microgrids," in *Proc. IEEE Conf. Decision and Control*, Dec. 2015, pp. 2018–2023.

[8] S. T. Cady, M. Zholbaryssov, A. D. Domínguez-García, and C. N. Hadjicostis, "A distributed frequency regulation architecture for islanded inertialess ac microgrids," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 6, pp. 1961–1977, Nov 2017.

[9] A. Lam, B. Zhang, and D. Tse, "Distributed algorithms for optimal power flow problem," in *Proc. IEEE Conf. Decision and Control*, 2012, pp. 430–437.

[10] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.

[11] M. Zholbaryssov and A. D. Domínguez-García, "Exploiting phase cohesiveness for frequency control of islanded inverter-based microgrids," in *Proc. IEEE Conf. Decision and Control*, Dec 2016, pp. 4214–4219.

[12] A. S. Debs, *Modern Power Systems Control and Operation*. Boston, MA: Springer US, 1988.

[13] M. Souiai, M. R. Oswald, Y. Kee, J. Kim, M. Pollefeys, and D. Cremers, "Entropy minimization for convex relaxation approaches," in *IEEE International Conference on Computer Vision*, Dec 2015, pp. 1778–1786.

[14] G. Finlayson, M. Drew, and C. Lu, "Intrinsic images by entropy minimization," *Proc. European Conference on Computer Vision*, pp. 582–595, 2004.

[15] M. Schneider, "Matrix scaling, entropy minimization, and conjugate duality (ii): The dual problem," *Mathematical Programming*, vol. 48, no. 1-3, pp. 103–124, 1990.

[16] S. Seshu and M. Reed, *Linear graphs and electrical networks*. Addison-Wesley Pub. Co., 1961.

[17] C. Léonard, "Minimization of entropy functionals," *Journal of Mathematical Analysis and Applications*, vol. 346, no. 1, pp. 183 – 204, 2008.

[18] C. R. Rao, *Convexity properties of entropy functions and analysis of diversity*, ser. Lecture Notes–Monograph Series. Hayward, CA: Institute of Mathematical Statistics, 1984, vol. 5, pp. 68–77.

[19] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[20] *IEEE 39-Bus System*, Illinois Center for a Smarter Electric Grid (ICSEG), http://icseg.iti.illinois.edu/ieee-39-bus-system/.