Stochastic Systems

Optimization of Stochastic Lossy Transport Networks and Applications to Power Grids

Alessandro Zocca, Bert Zwart

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Optimization of Stochastic Lossy Transport Networks and Applications to Power Grids

Alessandro Zocca,[a] Bert Zwart[b]

[a] Department of Mathematics, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, Netherlands; [b] Centrum Wiskunde and Informatica, 1098 XG Amsterdam, Netherlands
Contact: a.zocca@vu.nl, https://orcid.org/0000-0001-6585-4785 (AZ); bert.zwart@cwi.nl, https://orcid.org/0000-0001-9336-0096 (BZ)

Abstract. Motivated by developments in renewable energy and smart grids, we formulate a stylized mathematical model of a transport network with stochastic load fluctuations. Using an affine control rule, we explore the trade-off between the number of controllable resources in a lossy transport network and the performance gain they yield in terms of expected power losses. Our results are explicit and reveal the interaction between the level of flexibility, the intrinsic load uncertainty, and the network structure.

1. Introduction

A transport network is an abstract model describing a structure in which some commodity is transferred from the "source" nodes of the network to the "sink" nodes according to a specified routing that is determined by some external principle or design (see Daganzo 1997, Whittle 2007). Examples of transport networks are road networks, railways, pipes, and power grids.

In this work, we focus in particular on *lossy* transport networks where a fraction of the transported good is inevitably lost, having in mind as primary application power systems in which part of the transported electricity is lost because of heat dissipation in the transmission lines.

The main question that we want to address in the present paper is how these transportation losses can be reduced in the scenario in which we have no direct control on the routing but some of the nodes of the network have controllable loads. This is the case for power systems in which the line flows are determined by physical laws but at the same time feature an increasing number of controllable energy resources, like energy storage devices, smart buildings and appliances, and electric vehicles.

In the present work, we consider a probabilistic model to describe the stochastic fluctuations of the load in (a subset of) the nodes of the network. This is instrumental to model the stochastic load fluctuations because of power demand uncertainty and intermittent generation by renewable energy sources. The current power grids were originally built around conventional power generation systems; therefore, they are not equipped to cope with this massive amount of uncertainty, especially in power supply. Managing this uncertainty on such a large scale with existing methods will soon become crucial: in the next decades, power grids will have to become more flexible and robust to reduce the likelihood of contingencies and blackouts, whose social and economic impact is enormous.

As mentioned earlier, next to the increasing renewable penetration, there is another powerful trend that is driving this pervasive evolution of the power system: the advent of distributed energy resources. At a high level, all these resources can be seen as "virtual storage/batteries," in the sense that they can dynamically reduce their power consumptions and even inject electricity in the power grid when necessary (see Taylor and Mathieu 2015, Mathias et al. 2017, Barooah 2019). Even if at the present stage these resources are not fully incorporated, they have a huge potential: if their penetration increases and we can actively and optimally control them, then they can make power grids more flexible and at the same time effectively mitigate the volatile nature of renewable power generation and allow a higher share of renewable energy sources.

The controllable loads that we consider in this work should be seen as an abstraction of more concrete examples, such as (i) actual energy storage that is neither full nor empty; (ii) distribution grids with ample flexible and/or deferrable load (Meyn et al. 2013, 2015; Mathias et al. 2016, 2017); or (iii) conventional generators that can provide balancing services.

The stochastic network model considered in this paper aims to understand the potential that these controllable resources can have in mitigating the load uncertainty and in particular the transportation losses due to the stochastic load fluctuations in the network nodes. Specifically, we consider a network with random sources and sinks modeled by an undirected connected weighted graph G consisting of n nodes and m edges to which are associated nonnegative weights $\beta \in \mathbb{R}_+^m$ and investigate how much could the *average* total loss be reduced by operating optimally the subset $B \subseteq V$ of nodes with controllable loads.

The metric we consider here to quantify the transportation losses due to stochastic fluctuations is a quadratic function of the load profile that has been introduced in Johnson and Chertkov (2010) as proxy for the total power losses in AC power grids, as we will review in more detail in the next section. Such a metric generalizes the notion of *total effective resistance* $R_{\text{tot}}(G)$ of the graph G, also known as *Kirchhoff index*. This is a key quantity that measures how well connected and robust a network is and for this reason has been extensively studied and rediscovered in various contexts, such as complex network analysis (Ellens et al. 2011) and theoretical chemistry (for an overview, see Zhou and Trinajstić 2008 and references therein).

In this paper, we address the question of how to optimally operate distributed energy resources to minimize power losses in power systems. Because of imperfect sensing and delayed communication, a real-time perfect coordination between network nodes based on the realized fluctuations is unfeasible. We thus consider a static decentralized policy that amounts to an affine control rule for the controllable loads, which is inspired by the Automatic Generation Control mechanism, currently used in power systems (Wood et al. 2014). This control mechanism prescribes that in case of a power imbalance, all the generators have to either increase or decrease their power generation *proportionally to their participation factors*, which are static nonnegative scalars set beforehand based on economic principles.

In the present work, we envision a scenario where not just a few big conventional generators but several other distributed energy resources could also join this effort to balance power fluctuations. We thus formulate a constrained optimization problem to find the optimal static *load-sharing factors* for the controllable loads. We prove that the solution is unique and gives a closed-form expression for such an optimal control from which reflects the interplay between the network structure, the location of the controllable nodes, and the correlation structure of the load fluctuations.

In particular, our result shows how the correlation structure between load fluctuations affects the optimal operations of the network resources; this is extremely relevant in power grids with geographically close wind or solar farms, whose power outputs are obviously highly correlated. These insights are derived without making Gaussian assumptions on the distribution of the fluctuations.

We then use this explicit solution to explore the trade-off between the number of controllable resources available in a network and the performance gain they yield, quantified as transportation loss reduction. The analysis builds on and extends that of Zocca and Zwart (2016), which considered the scenario with only two controllable loads. In particular, the authors show therein that the expected total loss due to fluctuations can be reduced by 25% in a line network by adding one controllable storage device. In the present paper, we extended that insight by showing that for large graphs the total expected loss can be reduced on average by a factor $(1 + 1/k)/2$ in the scenario where k controllable loads are available. The key insight is larger values of k yield diminishing improvements in terms of total expected loss.

This suggests that, even if power grids are becoming locally more robust, transportation losses can be reduced by up to 50% and the number of controllable loads or balancing services k quantify how close one can get to this reduction. Though our model is stylized, it provides a simple quantitative estimate on the value of balancing services in a scenario where each node in the network is self-sufficient on average.

We remark that our stylized model is "static," in the sense the optimal control we derive does rebalance the total power mismatch in the network in any scenario but does not depend on the *realized* load fluctuations. Indeed, it depends only on the network structure, on the location of the controllable loads, and on their average covariance structure of the load fluctuations. For this reason, our model is not specific for a precise timescale and provides insights into the value of balancing services both in real-time operations as well as long-term planning.

Our work provides a new mathematical approach that can be of help for the design of future power grids and of control schemes for distributed energy resources. In this respect, it complements a large body of literature on optimal topology design for power grids (Ghosh et al. 2008, Johnson and Chertkov 2010,

Deka et al. 2017) and on optimal control of multiple controllable devices and/or generators where often the designed participation factors are also affine in the stochastic load fluctuations (see, e.g., Kanoria et al. 2011; Sjodin et al. 2012; Apostolopoulou et al. 2014; Bienstock et al. 2016; Lubin et al. 2016; Roald et al. 2016a, b; Sundar et al. 2016; Chertkov and Dvorkin 2017; Guggilam et al. 2017a; Guggilam et al. 2017b; Matke et al. 2017). Optimal policies for storage management, especially aiming at the mitigation of the uncertainties in wind generation, have been explored in Bejan et al. (2012), Gast et al. (2012, 2014), and Van De Ven et al. (2013), where, however, the physical network is not modeled explicitly. Optimal storage placement can increase network reliability, as shown in Bhaumik et al. (2016) by simulation techniques. Storage can also be used for arbitrage, exploiting temporal price differences (Cruise et al. 2014, Cruise and Zachary 2018) and the impact of storage on energy markets has been studied in Gast et al. (2013) and Cruise et al. (2018).

The rest of the paper is organized as follows. We provide a detailed model description in Section 2. In Section 3, we investigate the optimal load sharing factors in several scenarios. These results are applied in Section 4 where a scaling law is presented for large networks. A more general optimization problem, which takes into account economic factors and/or further limitations of the controllable loads, is presented in Section 5. In Section 6, we report several numerical experiments and, lastly, Section 7 concludes.

2. Model Description

In this paper, we model a lossy transport network as a weighted graph (G, w), where the graph G is a simple undirected graph $G = (V, E)$ with $|V| = n$ nodes labelled as $V = \{1, \ldots, n\}$, and $|E| = m$ edges and $\boldsymbol{\beta} \in \mathbb{R}^m_+$ is the collection of edge weights. In the context of power grids, the nodes of G are often referred to as *buses* and the edges as *transmission lines* and the quantity $\beta_{i,j}$ is the *susceptance* of the transmission line $\ell = (v, w) \in E$ connecting buses i and j. Missing edges can be thought as edges with zero weight.

The *weighted Laplacian matrix* of the graph G, often referred to also as the *susceptance matrix* of G in the case of a power system, is the matrix $L \in \mathbb{R}^{n \times n}$ defined for every $i, j = 1, \ldots, n$ as

$$L_{i,j} := \begin{cases} \sum_{k \neq i} \beta_{i,k} & \text{if } i = j, \\ -\beta_{i,j} & \text{if } i \neq j. \end{cases}$$

It is well known that L is a real symmetric positive semidefinite matrix. By construction, all the rows of L sum up to zero and thus the matrix L is singular. Under the assumption that G is a connected graph, the eigenvalue zero has multiplicity one and the corresponding eigenvector is equal to the vector with all unit entries, which we denote by $\mathbf{1} \in \mathbb{R}^n$.

Let $L^+ \in \mathbb{R}^{n \times n}$ be the *Moore-Penrose pseudoinverse* of the weighted Laplacian matrix L. Using the eigenspace structure of L, the pseudoinverse L^+ can be defined as

$$L^+ := \left(L + \frac{\mathbf{1}\mathbf{1}^T}{n} \right)^{-1} - \frac{\mathbf{1}\mathbf{1}^T}{n}.$$

The matrix L^+ is also real, symmetric, and positive semidefinite. For the proof of these properties of the matrices L and L^+ and for further spectral properties of graphs, we refer the reader to Van Mieghem (2011) and Ranjan et al. (2014).

Denote by $\mathbf{p} \in \mathbb{R}^n$ the *load profile* at the network nodes, where $\mathbf{p}_i$ is the load at node i for every $i = 1, \ldots, n$. In the context of power grids, the i-th entry of the vector $\mathbf{p}$ models the power generated (if $\mathbf{p}_i > 0$) or consumed (if $\mathbf{p}_i < 0$) at node i. We say that a load profile $\mathbf{p} \in \mathbb{R}^n$ is *balanced* if $\mathbf{1}^T \mathbf{p} = 0$.

Given a network with a balanced load profile $\mathbf{p} \in \mathbb{R}^n$, we define its *total loss* $\mathcal{H} = \mathcal{H}(\mathbf{p})$ as

$$\mathcal{H} := \frac{1}{2} \mathbf{p}^T L^+ \mathbf{p}.$$

The scalar quantity $\mathcal{H}$ is a quadratic form of the load profile vector $\mathbf{p}$ and, as such, is always nonnegative, thanks to the fact that L^+ is a positive semidefinite matrix.

The total aggregated loss $\mathcal{H}$ will be central in our analysis, being the most natural choice for an efficiency metric in lossy transport networks such as power systems. More specifically, $\mathcal{H}$ has been shown by Johnson and Chertkov (2010) to be a good approximation for *total power loss* in AC power grids, as we will briefly outline.

AC current flows are described in terms of complex amplitudes and lines with complex impedances. In power transmission networks operating in stationary conditions (i) all voltages are mostly constants and

(ii) we can ignore reactive power flows and line conductances because they are an order of magnitude smaller than the active power flows and line susceptances, respectively. Under these two assumptions, it is reasonable to consider the linearized version of the AC Kirchhoff equations for the power flows, that is, the so-called *DC-approximation*, called in this way because it resembles in structure the equations describing a resistive network with DC flows (cf. Wood et al. 2014, chapter 6). If $\mathbf{p}$ denotes the vector of active power injections in the network nodes, by keeping only the leading term of such a DC-approximation, the power flows $f \in \mathbb{R}^m$ on the network lines are given by

$$f = \Lambda L^+ \mathbf{p}, \tag{1}$$

where L is the weighted Laplacian matrix derived looking only the imaginary part of the network admittance matrix and $\Lambda \in \mathbb{R}^{m \times n}$ is the corresponding weighted edge-vertex incidence matrix

$$\Lambda_{\ell,i} = \begin{cases} \beta_\ell & \text{if } \ell = (i,j), \\ -\beta_\ell & \text{if } \ell = (j,i), \\ 0 & \text{otherwise.} \end{cases}$$

Because the power loss on each line ℓ is proportional to $\beta_\ell^{-1} f_\ell^2$, the aggregated power loss is (up a constant factor capturing the conductance-to-admittance ratio of the lines) equal to

$$\mathcal{H} = \frac{1}{2} \sum_{\ell \in E} \beta_\ell^{-1} f_\ell^2 = \frac{1}{2} f^T \text{diag}(\beta_1, \ldots, \beta_m)^{-1} f = \frac{1}{2} \mathbf{p}^T L^+ \Lambda^T \text{diag}(\beta_1, \ldots, \beta_m) \Lambda L^+ \mathbf{p} = \frac{1}{2} \mathbf{p}^T L^+ \mathbf{p},$$

where we used the fact that $L = \Lambda^T \text{diag}(\beta_1, \ldots, \beta_m)^{-1} \Lambda$ and $L^+ L L^+ = L^+$. Therefore, modulo a trivial rescaling, the quantity $\mathcal{H}$ is an approximation for the total power losses calculated using only the leading order DC-approximation of the AC network flows.

The quantity $\mathcal{H}$ has also been considered as a scalar measure of the *network tension* in Lai and Low (2013) and Guo et al. (2017), where it is shown to be monotone along a cascade failure, as long as the network remains connected. Lastly, as we will show later, $\mathcal{H}$ can also be seen as a generalized total effective resistance in the sense that it quantifies how robust the network G is against a stochastic load profile with a predefined covariance structure.

2.1. Stochastic Fluctuations and Load-Sharing Factors

In this work, we are particularly interested in a transport network with a *stochastic* load profile, which means that we take $\mathbf{p}$ to be a multivariate random variable.

More precisely, we assume $\mathbf{p}$ to be of the form $\mathbf{p} = \mu + \omega$, where $\mu \in \mathbb{R}^n$ is the nominal load profile in the network and ω is a n-dimensional random vector modeling the fluctuations. We henceforth assume that $\mathbb{E}\omega = \mathbf{0}$ and that ω has finite second moment. We further denote by $\Sigma \in \mathbb{R}^{n \times n}$ the covariance matrix of the load fluctuations, namely, $\Sigma_{i,j} = \text{cov}(\omega_i, \omega_j) = \mathbb{E}[\omega_i \omega_j] < \infty$. Nodes in which there are no stochastic load fluctuations can be modeled by setting all the entries equal to zero in the corresponding row and the column of the matrix Σ. We denote by $S \subseteq V$ the subset of nodes with stochastic load fluctuations, and we will henceforth assume that $|S| \geq 1$.

We further assume that the nominal load profile is balanced on average, namely, $\mathbf{1}^T \mu = 0$; this assumption is reasonable, as after every 5–15 minutes, the setpoint μ is adjusted by solving the so-called Optimal Power Flow (OPF) problem, which specifically constraints the net total power to be equal to zero. This assumption, however, does not guarantee that every stochastic realization of the load profile is balanced: indeed, the total mismatch $\mathbf{1}^T \mathbf{p}$ in the network is a random variable, which can be expressed as the net sum of fluctuations, namely,

$$\mathbf{1}^T \mathbf{p} = \mathbf{1}^T(\mathbf{p} - \mu) = \mathbf{1}^T \omega = \sum_{i=1}^n \omega_i.$$

Let $\sigma^2 := \text{Var}(\sum_{i=1}^n \omega_i)$ be the variance of the sum of the load fluctuations, which also rewrites as

$$\sigma^2 = \sum_{i,j=1}^n \sum_{i,j} = \mathbf{1}^T \Sigma \mathbf{1} = \text{tr}\left(\Sigma \mathbf{1} \mathbf{1}^T\right). \tag{2}$$

Because we assumed that there is at least one node with stochastic load fluctuations, we have $\sigma^2 > 0$.

In order to cope with the stochastic fluctuations and, in particular, to keep the network balanced, we assume that load at each node is *controllable*: for every $i = 1, \ldots, n$, node i can deal with (either generate or store) a controllable fraction $\alpha_i \in \mathbb{R}$ of the realized total mismatch $\mathbf{1}^T(\mathbf{p} - \boldsymbol{\mu}) = \sum_{i=1}^{n} \omega_i$. In other words, we assume that although using the *load-sharing factors* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, the *net* load profile $\mathbf{p}(\boldsymbol{\alpha})$ is given by

$$\mathbf{p}(\boldsymbol{\alpha}) = \left(p_1 - \alpha_1 \sum_{i=1}^{n} \omega_i, \ldots, p_n - \alpha_n \sum_{i=1}^{n} \omega_i \right) = \left(\mu_1 + \omega_1 - \alpha_1 \sum_{i=1}^{n} \omega_i, \ldots, \mu_n + \omega_n - \alpha_n \sum_{i=1}^{n} \omega_i \right).$$

For any $j = 1, \ldots, n$, the term $\alpha_j \sum_{i=1}^{n} \omega_i$ corresponds to the power generated or stored in the corresponding controllable load when using an affine control responsive to stochastic load fluctuations. We can rewrite the net power load profile $\mathbf{p}(\boldsymbol{\alpha})$ in vector form as

$$\mathbf{p}(\boldsymbol{\alpha}) = C_{\boldsymbol{\alpha}} \mathbf{p} = C_{\boldsymbol{\alpha}}(\boldsymbol{\mu} + \boldsymbol{\omega}) = \boldsymbol{\mu} + C_{\boldsymbol{\alpha}} \boldsymbol{\omega}, \tag{3}$$

where $C_{\boldsymbol{\alpha}} \in \mathbb{R}^{n \times n}$ is the matrix defined as

$$C_{\boldsymbol{\alpha}} := I - \boldsymbol{\alpha} \mathbf{1}^T = \begin{pmatrix} 1 - \alpha_1 & -\alpha_1 & \ldots & & -\alpha_1 \\ -\alpha_2 & 1 - \alpha_2 & -\alpha_2 & \ldots & -\alpha_2 \\ & & \ddots & \ddots & \ddots \\ \vdots & & & & \vdots \\ -\alpha_n & & \ldots & -\alpha_n & 1 - \alpha_n \end{pmatrix}.$$

The last equality in (3) follows from the fact that $\mathbf{1}^T \boldsymbol{\mu} = 0$, because $C_{\boldsymbol{\alpha}} \boldsymbol{\mu} = (I - \boldsymbol{\alpha} \mathbf{1}^T) \boldsymbol{\mu} = \boldsymbol{\mu} - 0 \cdot \boldsymbol{\alpha} = \boldsymbol{\mu}$. When the load-sharing factors $\boldsymbol{\alpha}$ are used, the total mismatch in the network is equal to

$$\mathbf{1}^T \mathbf{p}(\boldsymbol{\alpha}) = \mathbf{1}^T C_{\boldsymbol{\alpha}}(\boldsymbol{\mu} + \boldsymbol{\omega}) = \mathbf{1}^T (I - \boldsymbol{\alpha} \mathbf{1}^T) \boldsymbol{\omega} = (\mathbf{1}^T - (\mathbf{1}^T \boldsymbol{\alpha}) \mathbf{1}^T) \boldsymbol{\omega}.$$

From this expression, it is immediate to derive the condition on $\boldsymbol{\alpha}$ that guarantees that the net load profile $\mathbf{p}(\boldsymbol{\alpha})$ is balanced for *any* realization $\boldsymbol{\omega}$, which is stated in the next lemma.

Lemma 1 (Stochastic Load Profile Balance Condition). *Consider a network with balanced nominal load profile $\mathbf{1}^T \boldsymbol{\mu} = 0$ and controllable loads using load-sharing factors $\boldsymbol{\alpha} \in \mathbb{R}^n$. The stochastic load profile $\mathbf{p}(\boldsymbol{\alpha})$ is balanced for every realization of the fluctuations if and only if*

$$\mathbf{1}^T \boldsymbol{\alpha} = \sum_{i=1}^{n} \alpha_i = 1. \tag{4}$$

2.1.1. Model Discussion. Before presenting our results in the next section, we briefly discuss here the motivation behind the assumptions on the distribution of $\boldsymbol{\omega}$ and the motivation for considering an affine static control policy for the fluctuations.

First of all, the random variable $\sum_{j=1}^{n} \omega_j$, which describes the net power mismatch, can have any sign. We thus tacitly assume that our controllable nodes can respond to both positive and negative fluctuations, meaning that they can both store energy in excess or provide energy if the network needs it. For the sake of generality, we impose no further constraints on the random variable $\sum_{j=1}^{n} \omega_j$; thus, the power absorbed/injected by controllable node i, that is, $\alpha_i \sum_{j=1}^{n} \omega_j$, could theoretically take very large (positive and negative) values. This may seem a dubious and unrealistic assumption, but it is not for various reasons, which we now outline. Firstly, in normal operating conditions and on short timescales, the quantity $\sum_{j=1}^{n} \omega_j$ is typically "small." Indeed, our model is particularly relevant for a prompt response on short timescales, in which *large unforeseen power fluctuations are highly unlikely*. If an extremely large net power fluctuation occurs, it would be physically impossible for controllable loads and storage devices to resolve this imbalance by themselves; the network operator needs to take specific ad hoc emergency actions anyway. In particular, the real-time energy market and a consequent Optimal Power Flow routine on the 5–15 minutes scale will allow for a new safe setpoint $\boldsymbol{\mu}$ and for energy reserves to be used/bought. Secondly, because a detailed statistical modeling of demand fluctuations or renewable power production is not within the scope of this paper, we consider a general multidimensional distribution for the stochastic fluctuations $\omega_1, \ldots, \omega_n$. It is, however, very reasonable to assume that the support of these random variables is bounded, because, for instance, renewable energy sources cannot produce a negative amount of power nor produce beyond a specific power rating (above which automatic protective relay mechanisms are automatically triggered causing so-called power curtailment).

Lastly, a controllable load that *at a specific moment* does not have the two-directional flexibility (i.e., that cannot both store and release energy) can temporarily "go offline"; we can account for it by adding the constraint $\alpha_i = 0$ for the corresponding network node.

An affine control modeled using the load-sharing factors $\alpha \in \mathbb{R}^n$ is clearly a simplification, especially when the controllable loads model energy storage. Indeed, it does not incorporate many details, such as possible ramping constraints or the current state of charge. In our model, such details are omitted on purpose to have a mathematically tractable optimization problem and to better identify the interplay between load uncertainty and storage operations. A further simplification we make is that we allow to choose α without invoking line limits, as we also do this for mathematical tractability; but the extension we present in Section 5 accounts for these limits and several others.

We remark that these assumptions on the stochastic fluctuations ω and on the affine control of the deviations from the nominal setpoints are fairly common in power systems operations (see, e.g., Bienstock et al. 2014; Roald et al. 2016a, b).

2.2. Expected Total Loss: Definition and Properties

For any $1 \le k \le n$, we can model a network where exactly k nodes have controllable loads by imposing that the remaining $n - k$ nodes have load-sharing factors equal to zero, so that for any such node $\mathbf{p}_i = \mathbf{p}_i(\alpha)$. Using the net load profile $\mathbf{p}(\alpha)$, the total loss rewrites as

$$\mathcal{H}(\alpha) = \frac{1}{2}\mathbf{p}(\alpha)^T L^+ \mathbf{p}(\alpha);$$

therefore, $\{\mathcal{H}(\alpha)\}_{\alpha \in \mathbb{R}^n}$ is a family of random variables parametrized by the control α. Being a quadratic form and being L^+, a positive semidefinite matrix, it immediately follows that $\mathcal{H}(\alpha)$ is a nonnegative random variable for any $\alpha \in \mathbb{R}^n$.

The next proposition, which is proved in Appendix A, shows how, leveraging the properties of the matrices L^+ and C_α, the expected total loss $\mathbb{E}\mathcal{H}(\alpha)$ rewrites as the sum of two contributions, one stochastic and one deterministic, which is not affected by the control α. Furthermore, it rewrites the expected total loss as a quadratic function of the load-sharing vector α.

Proposition 1. *Consider a network with a balanced nominal load profile μ and zero-mean stochastic fluctuations ω with covariance matrix Σ. Then, the expected total loss using the control α is given by*

$$\mathbb{E}\mathcal{H}(\alpha) = \mathbb{E}\mathcal{H}_s(\alpha) + \frac{1}{2}\mu^T L^+ \mu, \tag{5}$$

where $\mathbb{E}\mathcal{H}_s(\alpha)$ is the expected total loss due to the stochastic fluctuations. Furthermore, $\mathbb{E}\mathcal{H}_s(\alpha) \ge 0$ for every $\alpha \in \mathbb{R}^n$, and the following identity holds:

$$\mathbb{E}\mathcal{H}_s(\alpha) = \frac{\sigma^2}{2}\left(\alpha^T L^+ \alpha\right) - \mathbf{1}^T \sum L^+ \alpha + \frac{1}{2}\mathrm{tr}\left(\sum L^+\right). \tag{6}$$

The first important remark is that the expected total loss $\mathbb{E}\mathcal{H}(\alpha)$ is a quadratic form in the vector $\alpha \in \mathbb{R}^n$ because it can be rewritten as

$$\mathbb{E}\mathcal{H}(\alpha) = \frac{\sigma^2}{2}\alpha^T A \alpha - b^T \alpha + c, \tag{7}$$

where $A = L^+$ is a positive semidefinite matrix, $\sigma^2 > 0$, $b = L^+ \sum \mathbf{1} \in \mathbb{R}^n$, and $c = \mathrm{tr}(\sum L^+)/2 + \frac{1}{2}\mu^T L^+ \mu \in \mathbb{R}^+$. Furthermore, we can already conclude that the nominal load profile μ has no impact on the optimal control α, because it appears only in the constant term c.

Before investigating what is the optimal load-sharing vector for a given network and covariance structure of the noise, we argue here why the quantity $\mathbb{E}\mathcal{H}(\alpha)$ can be seen as a generalized notion of effective resistance. In order to do so, we will first recall some classical definitions.

The *effective resistance* $R_{i,j}$ between a pair of nodes i and j of the network G is defined as the electrical resistance measured across nodes i and j when we look at G as the electrical network in which resistors with conductance $\beta_1^{-1}, \ldots, \beta_m^{-1}$ are placed at the corresponding network edges. Equivalently,

$$R_{i,j} := \left(\mathbf{e}_i - \mathbf{e}_j\right)^T L^+ \left(\mathbf{e}_i - \mathbf{e}_j\right) = L_{i,i}^+ + L_{j,j}^+ - 2L_{i,j}^+,$$

where $\mathbf{e}_i$ denotes the vector with a one in the i-th coordinate and zeros elsewhere. The *total effective resistance* of a graph G is then defined as

$$R_{\text{tot}}(G) := \frac{1}{2}\sum_{i,j=1}^{n} R_{i,j}.$$

The same quantity is also known as the *Kirchhoff index* when the network G is such that all the edge weights are equal to one. As mentioned in the introduction, the total effective resistance has been used in various contexts (Zhou and Trinajstić 2008, Ellens et al. 2011) to quantify how well connected a given network is. In the context of electrical networks, the total effective resistance $R_{\text{tot}}(G)$ is shown in Ghosh et al. (2008) to be proportional to the average power dissipated in a resistive DC network (G, w) when random independent and identically distributed (i.i.d.) currents with zero mean and unit variance are injected at the nodes.

We can look at the network $(G, w, \boldsymbol{\alpha})$ with controllable loads introduced earlier as a flexible transport network, where the load-sharing factors $\alpha_1, \ldots, \alpha_n$ can be tuned to respond optimally to specific stochastic load fluctuations. In this respect, we claim that the quantity $\mathbb{E}\mathcal{H}_s(\boldsymbol{\alpha})$ can be seen as a (rescaled) *generalized* total effective resistance that measures how "robust" the network $(G, w, \boldsymbol{\alpha})$ is against stochastic load fluctuations with covariance structure Σ. To further corroborate this claim, we now show that the expected total loss reduces to the classical total effective resistance $R_{\text{tot}}(G)$ in the special case where the load-sharing factors are all equal, that is, $\alpha_i = 1/n$ for every $i = 1, \ldots, n$, and the stochastic load fluctuations are i.i.d. random variables with zero mean and unit variance, that is, $\Sigma = I$. Using (6), the expected total loss due to fluctuations re-writes as

$$\mathbb{E}\mathcal{H}_s\left(\frac{1}{n}\mathbf{1}\right) = \frac{n}{2n^2}\left(\mathbf{1}^T L^+ \mathbf{1}\right) - \mathbf{1}^T L^+ \mathbf{1} + \frac{1}{2}\text{tr}(L^+) = \frac{1}{2}\text{tr}(L^+) = \frac{R_{\text{tot}}(G)}{2n}, \tag{8}$$

where, in the last step, we use the well-known identity $R_{\text{tot}}(G) = n \cdot \text{tr}(L^+)$ proved by Klein and Randić (1993) that relates the total effective resistance of a graph with its spectrum.

Rayleigh's monotonicity principle (Doyle and Snell 2000) states that the pairwise effective resistance $R_{i,j}$ is a nonincreasing function of the edge weights and, as a consequence, the total effective resistance $R_{\text{tot}}(G)$ also is. The following proposition, proved in Appendix A, shows that a similar property also holds for $\mathcal{H}(\boldsymbol{\alpha})$: regardless of the load-sharing vector $\boldsymbol{\alpha}$, the total power loss does not increase when edges are added or weights are increased.

Proposition 2 (Monotonicity of Total Power Loss). *Let G be a weighted connected graph and let G' the graph obtained from G by increasing the weight of edge $e = (i, j)$ by $\beta > 0$ or by adding the edge $e = (i, j)$ with weight $\beta > 0$. For any load-sharing vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ and any realization of the stochastic load fluctuations $\boldsymbol{\omega}$, the following inequality holds*

$$\mathcal{H}^{G'}(\boldsymbol{\alpha}) \leq \mathcal{H}^{G}(\boldsymbol{\alpha}),$$

and, in particular, $\mathbb{E}\mathcal{H}^{G'}(\boldsymbol{\alpha}) \leq \mathbb{E}\mathcal{H}^{G}(\boldsymbol{\alpha})$.

3. Optimal Load-Sharing Control
In this section, we consider the problem of minimizing the expected total loss $\mathbb{E}\mathcal{H}(\boldsymbol{\alpha})$ given a network G and a stochastic load covariance structure Σ.

Let $B \subseteq V$ be the subset of k nodes with controllable loads. We focus first on the scenario in which not all the nodes have controllable loads and thus assume $1 \leq k < n$. Besides the constraint (4), we further add $n - k$ constraints on the optimal load-sharing vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ to account for the absence of controllable loads in the nodes in $V \setminus B$, obtaining the following constrained optimization problem in $\mathbb{R}^n$:

$$
\begin{aligned}
\min_{\boldsymbol{\alpha} \in \mathbb{R}^n} \quad & \mathbb{E}\mathcal{H}_s(\boldsymbol{\alpha}) \\
\text{s.t.} \quad & \mathbf{1}^T \boldsymbol{\alpha} = 1, \\
& \alpha_v = 0, \quad \forall v \in V \setminus B.
\end{aligned}
\tag{9}
$$

Note we consider only the expected total loss due to stochastic load fluctuations in the objective function, because in view of (5) it differs only by a constant from the expected total loss.

We henceforth assume that the k nodes with the controllable loads are those with labels $1, \ldots, k$, that is, $B = \{1, \ldots, k\}$. We can make this assumption without loss of generality, as it amounts to relabeling the network nodes. If this is the case, the rows and columns of matrices L, L^+, and Σ and the entries of the vector μ are also rearranged accordingly.

Let $P_B \in \{0,1\}^{n \times k}$ be the binary matrix that maps any k-dimensional vector $\widetilde{\alpha} \in \mathbb{R}^k$ to the n-dimensional vector $P_B \widetilde{\alpha} = (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_k, 0, \ldots, 0) \in \mathbb{R}^n$. Such a matrix can be defined component-wise as $(P_B)_{i,j} := \delta_{\{i=j\}} \delta_{\{i \le k\}}$, for $i = 1, \ldots, n$ and $j = 1, \ldots, k$, and has the following structure:

$$P_B = \begin{pmatrix} I_k \\ \mathbb{O} \end{pmatrix},$$

where I_k is the $k \times k$ identity matrix and $\mathbb{O} \in \mathbb{R}^{n-k \times k}$ is a matrix with all entries equal to zero.

In our first main result, which is proved in Appendix B, we present a closed-form expression for the optimal load-sharing factors of k controllables.

Theorem 1 (Optimal Load-Sharing Between $k < n$ Controllables)**.** *Consider a network with balanced nominal load profile $\mathbf{1}^T \mu = 0$ in which the nodes in $B = \{1, \ldots, k\}$ have controllable loads. The solution of the optimization problem* (9) *is the load-sharing vector* $\alpha^* = (\widetilde{\alpha}^* \, \mathbf{0})$, *with*

$$\widetilde{\alpha}^* = \frac{1}{t_B} \left(L_B^+ \right)^{-1} \mathbf{1} + \left(I_k - \frac{1}{t_B} \left(L_B^+ \right)^{-1} \mathbf{1} \mathbf{1}^T \right) \left(L_B^+ \right)^{-1} P_B^T L^+ \frac{\Sigma \mathbf{1}}{\sigma^2} \in \mathbb{R}^k, \tag{10}$$

where L_B^+ is the $k \times k$ principal submatrix of L^+, that is, $L_B^+ := P_B^T L^+ P_B$, and $t_B := \mathbf{1}^T (L_B^+)^{-1} \mathbf{1} > 0$.

If all the nodes with stochastic load fluctuations have controllable loads, that is, $S \subseteq B$, then

$$\widetilde{\alpha}^* = \frac{\Sigma \mathbf{1}}{\sigma^2}, \tag{11}$$

and, in particular, $\widetilde{\alpha}_v^* = 0$ for every $v \in B \setminus S$.

The involved expression (10) for the optimal load-sharing factors reflects the interplay that exists between the network structure, the location of the controllable loads B, and the correlation structure of the load fluctuations in determining the losses.

In the special case where all the nodes with stochastic load fluctuations have controllable loads, there is a nice interpretation for the optimal load-sharing factors: indeed α_i^* is proportional to how much the stochastic fluctuations of node i contribute in relative terms to the variance of the total mismatch, because

$$\alpha_i^* = \frac{1}{\sigma^2} e_i^T \Sigma \mathbf{1} = \frac{1}{\sigma^2} \sum_{j=1}^{n} \Sigma_{i,j} = \frac{\mathrm{Var}(\omega_i) + \sum_{j \ne i} \mathrm{Cov}(\omega_i, \omega_j)}{\mathrm{Var}\left(\sum_{i=1}^{n} \omega_i \right)}$$

$$= \frac{\mathrm{Var}(\omega_i) + \sum_{j \ne i} \mathrm{Cov}(\omega_i, \omega_j)}{\sum_{k=1}^{n} \left(\mathrm{Var}(\omega_k) + \sum_{j \ne k} \mathrm{Cov}(\omega_k, \omega_j) \right)}. \tag{12}$$

We further remark that in the case of i.i.d. stochastic fluctuations in all the nodes, the second term in (10) vanishes, because the vector $\Sigma \mathbf{1}$ lies in the null space of L^+ (being a multiple of $\mathbf{1}$); therefore, the optimal control is equal to

$$\widetilde{\alpha}^* = \frac{(L_B^+)^{-1} \mathbf{1}}{\mathbf{1}^T (L_B^+)^{-1} \mathbf{1}}.$$

In particular, it does not depend on the variance of the load fluctuations but only on the network structure and on the location B of the controllable loads, both encoded in the matrix L_B^+.

3.1. Full Controllability

In this subsection, we focus on the special case where the load is controllable in every node, that is, $B = V$, which is not covered in Theorem 1. Indeed, the proof method does not work in this scenario because of the noninvertibility of the graph Laplacian L and for this reason is treated separately here.

The problem of minimizing the expected loss $\mathbb{E}\mathcal{H}(\alpha)$ when all the nodes have controllable loads can be written as an optimization problem on $\mathbb{R}^n$ with a single constraint, namely,

$$\min_{\alpha \in \mathbb{R}^n} \quad \mathbb{E}\mathcal{H}(\alpha)$$
$$\text{s.t.} \quad \mathbf{1}^T \alpha = 1. \tag{13}$$

As mentioned earlier, thanks to Proposition 1, we can immediately conclude that the optimal load-sharing vector does not depend on the vector μ, which appears only in the constant term in the equality (5) for the expected loss.

The next theorem, whose proof is presented in Appendix C, derives an analytical expression for the optimal solution of this optimization problem.

Theorem 2 (Optimal Load-Sharing Between *n* Controllables). *Consider a network G with n nodes and a balanced nominal load profile* $\mathbf{1}^T \mu = 0$. *The solution of the optimization problem* (13) *is the load-sharing vector* α^* *given by*

$$\alpha^* = \frac{\Sigma \mathbf{1}}{\sigma^2}.$$

The highlight of this result is that in the scenario where all nodes have controllable loads, the optimal control α^* does not depend on the graph structure but only on the covariance structure of the fluctuations. The same interpretation as in the special case $S \subseteq B$ of Theorem 1 holds here for the optimal load-sharing factors: α_i^* is proportional to how much the stochastic fluctuations of node i contribute in relative terms to the variance of the total mismatch (see (12)). In particular, when a node i does not have stochastic load fluctuations, then it is optimal not to use the controllable load in that node, because $\alpha_i^* = 0$ in view of the fact that the i-th row of Σ is identically zero.

It immediately follows from Theorem 2 that when the stochastic load fluctuations are independent and identically distributed, the optimal load-sharing factors are all equal, namely,

$$\alpha^* = \frac{1}{n}\mathbf{1}.$$

Furthermore, the expected total loss due to stochastic fluctuations when using the optimal load-sharing factors rewrites as

$$\mathbb{E}\mathcal{H}_s(\alpha^*) = \frac{1}{2}\left(\text{tr}\left(\textstyle\sum L^+\right) - \frac{1}{\sigma^2}\mathbf{1}^T \textstyle\sum L^+ \textstyle\sum \mathbf{1} \right).$$

In this special case, the fact that $\mathbb{E}\mathcal{H}_s(\alpha^*) \geq 0$ can equivalently be proved as follows:

$$\mathbf{1}^T \textstyle\sum L^+ \textstyle\sum \mathbf{1} = \text{tr}\left(\textstyle\sum L^+ \textstyle\sum J\right) \leq \sqrt{\text{tr}\left(\left(\textstyle\sum L^+\right)^2\right) \text{tr}\left(\left(\textstyle\sum \mathbf{1}\mathbf{1}^T\right)^2\right)}$$

$$\leq \sqrt{\text{tr}\left(\textstyle\sum L^+\right)^2 \text{tr}\left(\textstyle\sum \mathbf{1}\mathbf{1}^T\right)^2} = \text{tr}\left(\textstyle\sum L^+\right)\text{tr}\left(\textstyle\sum \mathbf{1}\mathbf{1}^T\right) = \sigma^2 \text{tr}\left(\textstyle\sum L^+\right),$$

where both the inequalities leverage in a crucial way that all the matrices Σ, $\mathbf{1}\mathbf{1}^T$, and L^+ are positive semidefinite. The first inequality follows from the fact that, in the space of positive semidefinite matrices, trace is a proper inner-product and thus obeys the Cauchy-Schwarz inequality,

$$\text{tr}(AB) \leq \sqrt{\text{tr}(A^2)\,\text{tr}(B^2)} \qquad \forall\, A, B \text{ positive semi-definite matrices.}$$

The second inequality follows from the fact that $\text{tr}(A^2) \leq \text{tr}(A)^2$ for any positive semidefinite matrix A, obtained by applying the Cauchy-Schwarz inequality with $A = B$.

4. Scaling Properties of the Expected Total Loss

In this section, we explore the relation between the expected total loss and the number of controllable loads. Even if the intuition suggests that the expected total loss should be a decreasing function in the number of controllable loads, this fact may not be true in general, as the total loss depends both on the location of the controllable loads as well as on the load-sharing factors. For instance, in Section 6, we present a counterexample of a network in which by adding the one controllable load and readjusting the load-share factors to be all equal, the expected total power loss increases.

To get rid of these heterogeneities and obtain a more transparent result for the impact of the number of controllable loads, we calculate the expected total loss for a fixed number of controllable loads *averaging* on all their possible locations and assuming they share the load equally. Consider an integer $1 \le k \le n$ and denote by $\mathcal{B}_k \subset \mathbb{R}^n$ the collection of load-sharing vectors with exactly k nonzero identical entries (and thus equal to $1/k$, in view of (4)), namely,

$$\mathcal{B}_k := \left\{ \frac{1}{k} \sum_{i \in B} e_i \ : \ B \subseteq V, \ |B| = k \right\},$$

where $e_i \in \mathbb{R}^n$ is the vector with the i-th entry equal to one and zero elsewhere.

Let $\mathcal{H}_k$ denote the expected total loss due to stochastic load fluctuations averaging over all their possible locations of k controllable nodes share the load equally, that is,

$$\mathcal{H}_k := \frac{1}{|\mathcal{B}_k|} \sum_{\alpha \in \mathcal{B}_k} \mathbb{E}\mathcal{H}_s(\alpha).$$

This average consists of $|\mathcal{B}_k| = \binom{n}{k}$ terms, one for each possible arrangements of k controllable nodes in a network with n nodes. The following theorem states an explicit expression for $\mathcal{H}_k$ that makes the dependence on the number of controllable node k very explicit, showing in particular that $\mathcal{H}_k$ is, up to a constant, proportional to $1/k$. The proof of the theorem is presented in Appendix D.

Theorem 3 (Average Total Loss with k Controllable Loads). *Consider a network G of n nodes with balanced nominal profile load, that is, $\mathbf{1}^T \boldsymbol{\mu} = 0$. Then,*

$$\mathcal{H}_k = C_1 + \frac{C_2}{k},$$

where $C_1, C_2 \in \mathbb{R}$ are two constants that do not depend on k given by

$$C_1 = C_1(G) := \frac{1}{2} \mathrm{tr}\left(\sum L^+\right) - \sigma^2 \frac{\mathrm{tr}(L^+)}{2n(n-1)} \quad and \quad C_2 = C_2(G) := \sigma^2 \frac{\mathrm{tr}(L^+)}{2(n-1)}.$$

Both the constants $C_1(G)$ and $C_2(G)$ depend on the graph structure via L^+, on its size n, and on the covariance matrix Σ. We remark that the constant C_2 is always strictly positive, as $\mathrm{tr}(L^+) > 0$ (L^+ being a positive semidefinite matrix) and $\sigma^2 > 0$, in view of (2).

We are interested in understanding how the expected total loss scales for large graphs. Assume we can take a sequence of graphs $\{G_n\}_{n \in \mathbb{N}}$ of growing size, $|V_n| = n$, and of covariance matrices $\{\Sigma_n\}_{n \in \mathbb{N}}$ with total variance $\sigma_n^2 = \mathbf{1}^T \Sigma_n \mathbf{1}$, so that the limit

$$\gamma := \lim_{n \to \infty} \frac{(n-1)\,\mathrm{tr}(\Sigma_n L_n^+)}{\sigma_n^2\,\mathrm{tr}(L_n^+)}$$

exists. Note that the fact that inequality $\mathcal{H}_n \ge 0$ holds for every $n \in \mathbb{N}$ (see Proposition 1) guarantees that $\gamma \ge 0$. Under these assumptions, Theorem 1 readily implies that as the graph size n grows large, the relative gain of having k controllable loads with respect to a single one scales as

$$\lim_{n \to \infty} \frac{\mathcal{H}_k}{\mathcal{H}_1} = \left(\frac{1}{1 + \gamma^{-1}}\right) + \left(\frac{1}{1 + \gamma}\right)\frac{1}{k}.$$

In the scenario where the load fluctuations are independent and identically distributed, *regardless of the graph structure*, this asymptotic scaling reads

$$\lim_{n \to \infty} \frac{\mathcal{H}_k}{\mathcal{H}_1} = \frac{1}{2} + \frac{1}{2k}. \tag{14}$$

as

$$\gamma = \lim_{n \to \infty} \frac{(n-1)\,\mathrm{tr}(\Sigma_n L_n^+)}{\sigma_n^2\,\mathrm{tr}(L_n^+)} = \lim_{n \to \infty} \frac{(n-1)\,\sigma_n^2\,\mathrm{tr}(L_n^+)}{n\,\sigma_n^2\,\mathrm{tr}(L_n^+)} = \lim_{n \to \infty} \frac{n-1}{n} = 1.$$

In Subsection 6.6, we show numerically that the way the expected total loss scales on average with the number of controllable loads as stated in Theorem 3 is pretty accurate in general, even without averaging over all possible locations of the controllable loads.

5. Extensions
5.1. Generalized Model
The optimization problems (9) and (13) considered in Section 3 focus on the network efficiency but do not prevent the excessive usage of specific controllable loads nor account for economic factors. In this section, we present a more general optimization problem that addresses both these issues and that further includes additional constraints typical of the OPF problem, which network operators use to set the operational set-points of power systems every 5–15 minutes.

Let P be a positive definite $n \times n$ matrix, $R \in \mathbb{R}^{n \times n}$, and $q \in \mathbb{R}^n$. Introducing a nonnegative real number $\xi \geq 0$, we can consider the following generalized optimization problem:

$$\min_{\alpha \in \mathbb{R}^n} \quad \mathbb{E}\mathcal{H}(\alpha) + \xi\, \alpha^T P \alpha, \tag{15}$$

$$\text{s.t.} \quad \mathbf{1}^T \alpha = 1, \tag{16}$$

$$\alpha_v = 0, \qquad \forall\, v \in V \setminus B, \tag{17}$$

$$R\alpha \leq q, \tag{18}$$

$$|f_\ell(\alpha)| \leq f_\ell^{\max}, \quad \forall\, \ell \in E. \tag{19}$$

This optimization problem differs from that in (9) in two ways, namely, its objective function (15) has an extra term and it has two new sets of constraints (18) and (19).

The term $\alpha^T P \alpha$ can be seen as a *penalty* or *cost* term and the scalar $\xi \geq 0$ weights its relative importance with respect to the average total loss $\mathbb{E}\mathcal{H}(\alpha)$. More specifically, by taking P to be a diagonal matrix with nonnegative entries, the additional term $\alpha^T P \alpha = \sum_{i=1}^n P_{i,i}\alpha_i^2$ in the objective function penalizes excessive usage (i.e., large load-sharing factor) of the controllable loads with large coefficients $P_{i,i}$. This means that we can model less flexible or more costly controllable loads by tuning the corresponding terms $P_{i,i} > 0$ accordingly.

The additional constraint (18) can be used to set any desired upper and/or lower bounds on the load-sharing factor of any particular node. In particular, we could restrict the load-sharing factors for some of the nodes $v \in B$ to be nonnegative, $\alpha_v \geq 0$, or their absolute values $|\alpha_v| < \varepsilon_v$ with ad hoc constants $\varepsilon_v > 0$.

Lastly, the constraint (19) captures the physical limits of the lines, by imposing the power flow $f_\ell(\alpha)$ on any given network line $\ell \in E$ stays below the corresponding *capacity* $f_\ell^{\max} > 0$ of that line. As outlined in Section 2, using the linear approximation (1), the power flows $f(\alpha) = \Lambda L^+ p(\alpha)$ are linear in the load profile $p(\alpha)$ and, ultimately, linear also in the load-sharing factors α.

Clearly (15) is still a convex optimization problem: all the constraints are still linear in α and the objective function rewrites as a quadratic form with the matrix $L^+ + \xi P$ appearing in the leading term that is positive definite in view of the assumption made on the matrix P and Proposition 1. It is impossible, however, to derive the optimal control α^* in closed form for the generalized problem (15) and thus we cannot explore its structure as we did in the previous sections, but we present some numerical results in next section.

Nonetheless, we argue that the original problem (9) still gives valuable insight as the line capacity limits appearing in (19) are often redundant. Indeed, the OPF problem automatically sets a nominal setpoint μ so that all the nominal power flows f are within their limits and the optimal control α^* should decrease the largest line flows (in absolute value) $f(\alpha)$ even further, as it is intuitive from the expression $\mathcal{H}(\alpha) = \frac{1}{2}\sum_{\ell \in E} \beta_\ell^{-1} f_\ell^2(\alpha)$ (cf. Section 2).

Moreover, even if we set a specific load-sharing factor α_v to be nonnegative or smaller than ε_v with constraint (18), the controllable load in that node should still be capable of both storing and to outputting a possibly very large amount of power, as we outlined at the end of Subsection 2.1. Alternatively, a chance-constraint for the stochastic effort $\alpha_i \sum_j \omega_j$ of node i could be included in the optimization problem like it has been done in (Bienstock et al. 2014). However, additional assumptions on the multivariate distribution of ω must be made to rewrite the constraint so that the resulting optimization problem is still convex. Other chance-constraints for line limits could be obtained using concentration inequalities as in (Nesti et al. 2017) or using decay rates and large deviation theory as in (Nesti et al. 2018).

We conclude by deriving the optimal control α^* when considering the optimization problem with augmented objective function (15) in the special case where all the nodes have controllable loads, that is, $B = V$, but ignoring the constraints (18) and (19). In this scenario, the optimal solution can still be calculated in closed form as

$$\alpha^* = \left(\frac{\sigma^2}{2}L^+ + \xi P\right)^{-1}\left(I + \frac{1}{t}\mathbf{1}\mathbf{1}^T\left(\frac{\sigma^2}{2}L^+ + \xi P\right)^{-1}\right)L^+\sum\mathbf{1} + \frac{1}{t}\left(\frac{\sigma^2}{2}L^+ + \xi P\right)^{-1}\mathbf{1},$$

with $t := \mathbf{1}^T(\frac{\sigma^2}{2}L^+ + \xi P)^{-1}\mathbf{1} > 0$. This expression can be obtained in a similar way of that in Theorem 1, leveraging the fact that the matrix $L^+ + \xi P$ is positive definite.

5.2. Variance of the Aggregated Total Loss

It could be of interest choosing the load-sharing factors $\alpha_1, \ldots, \alpha_n$ not only to minimize the expected aggregated total loss $\mathbb{E}\mathcal{H}(\alpha)$ but also its *variance* $\mathrm{Var}(\mathcal{H}(\alpha))$. Aiming to possibly include this in our optimization, we need a closed-form expression for the variance of a quadratic form of a random vector. Unfortunately, this is known only in the case in which the random vector has a multivariate Gaussian distribution. Under this additional assumption, the variance of the aggregated total loss can be calculated explicitly as follows:

$$\mathrm{Var}(\mathcal{H}(\alpha)) = \mathrm{Var}(\mathcal{H}_s(\alpha)) = \mathrm{Var}\left(\frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega + \mu^T L^+ C_\alpha \omega\right)$$

$$= \frac{1}{4}\mathrm{Var}(\omega^T C_\alpha^T L^+ C_\alpha \omega) + \mathrm{Var}(\mu^T L^+ C_\alpha \omega) + \mathrm{Cov}(\omega^T C_\alpha^T L^+ C_\alpha \omega, \mu^T L^+ C_\alpha \omega)$$

$$= \frac{1}{2}\mathrm{tr}\left(\left(C_\alpha^T L^+ C_\alpha \sum\right)^2\right) + \mu^T L^+ C_\alpha \sum C_\alpha^T L^+ \mu.$$

The last step follows from standard results for quadratic forms of multivariate Gaussian random variables (Rencher and Schaalje 2008, theorems 5.2c and 5.2d) and the fact that $\mathbb{E}\omega = \mathbf{0}$.

Following similar steps to those in the proof of Proposition 1 in Appendix A, the expression for $\mathrm{Var}(\mathcal{H}(\alpha))$ can be further rewritten as a polynomial of degree four in the load-sharing factors $\alpha_1, \ldots, \alpha_n$. This new term could potentially be included in the objective function of the optimization problem (15), but the resulting problem is not guaranteed to still be convex in general.

6. Numerical Examples

In this section, we present some numerical results. First of all, in Subsection 6.1, we compare the performance of our static optimal load-sharing factors α^* with that of idealized real-time load-sharing factors $\alpha(\omega)$ that dynamically respond to fluctuations. Then, in Subsections 6.2 and 6.3, we study the impact of the covariance matrix and of the relative position of the nodes affected by fluctuations and those housing controllable loads. We present two examples to illustrate that (i) optimal load-sharing factors have opposite signs in Subsection 6.4 and (ii) expected total loss is not always decreasing in the number of controllable loads in Subsection 6.5. Lastly, in Subsection 6.6, we corroborate the accuracy of the scaling limit derived in Section 4.

6.1. Real-Time Adaptive Load-Sharing Factors

We briefly consider here the scenario in which it is possible to (i) have exact knowledge about the realized fluctuations ω and (ii) have instantaneous coordination between controllable loads. In this setting, a centralized controller could dynamically change the load-sharing factors depending on the realization of the noise ω, trying to minimize the total loss. In this case, the total aggregated power loss $\widetilde{\mathcal{H}}(\alpha, \omega) = \frac{1}{2}(\mu - C_\alpha\omega)^T L^+(\mu - C_\alpha\omega)$ is not a random variable but simply a quadratic form in α. Thus, for every realization of the noise ω, we can find the optimal load-sharing factors $\alpha^*(\omega) \in \mathbb{R}^n$ that solve

$$\begin{aligned} \min_{\alpha \in \mathbb{R}^n} \quad & \widetilde{\mathcal{H}}_s(\alpha, \omega) \\ \text{s.t.} \quad & \mathbf{1}^T\alpha = 1, \\ & \alpha_v = 0, \quad \forall v \in V \setminus B. \end{aligned} \tag{16}$$

Figure 1. The IEEE 14-Bus Test Network with Three Controllable Loads $B = \{2, 6, 9\}$ (Depicted as Squares) and All 14 Nodes Affected by Stochastic Fluctuations, That Is, $S = V$, with Correlations Captured by the Matrix Σ

The topology of the IEEE 14-bus test network

Scatter plot of the correlation matrix Σ
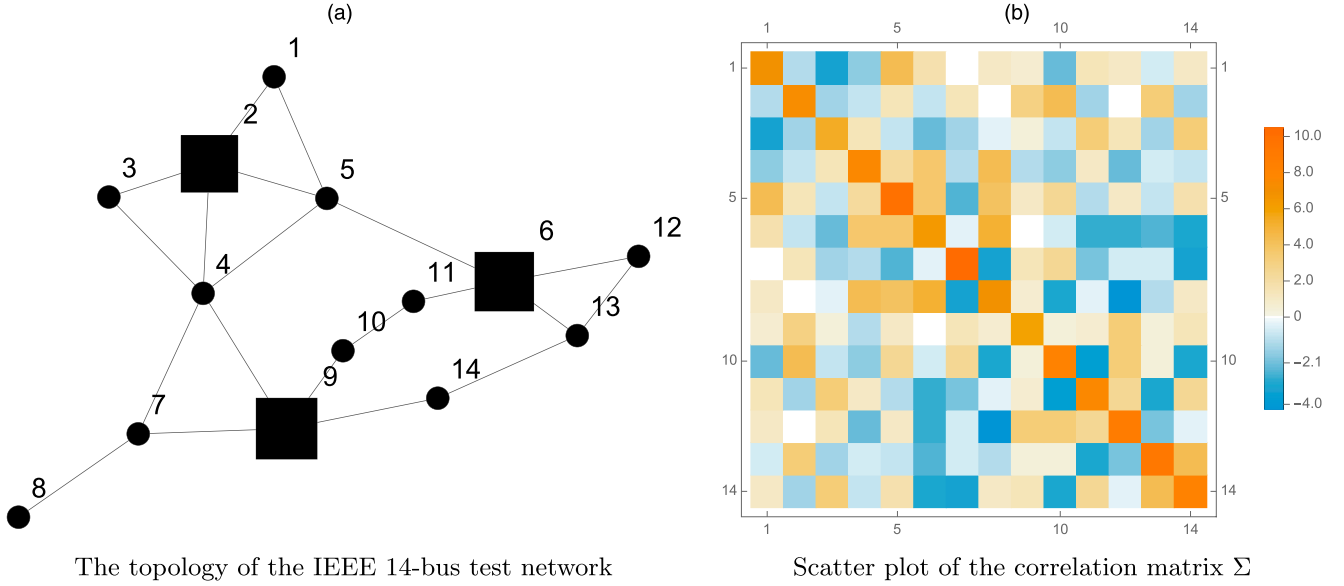
From our previous analysis, it is easy to see that this is still a convex optimization problem in α with a quadratic objective function and linear equality constraints. Its solution can be obtained from the corresponding Karush-Kuhn-Tucker conditions

$$\begin{cases} -(\mathbf{1}^T\boldsymbol{\omega})L^+(\boldsymbol{\mu} + \boldsymbol{\omega} - (\mathbf{1}^T\boldsymbol{\omega})\boldsymbol{\alpha}) + \pi\mathbf{1} + \boldsymbol{\nu} = \mathbf{0} \\ \mathbf{1}^T\boldsymbol{\alpha} = 1, \\ \alpha_v = 0, \quad \forall v \in V \setminus B, \end{cases}$$

where $\pi \in \mathbb{R}$ is the dual variable corresponding to the constraint $\mathbf{1}^T\boldsymbol{\alpha} = 1$ and $\boldsymbol{\nu} \in \mathbb{R}^n$ is the vector whose nonzero entries are equal to the dual variables of the constraints $\alpha_v = 0$, that is, $\nu_v = \nu \in \mathbb{R}$ if $v \in V \setminus B$ and zero otherwise. In the special case $B = V$, the vector $\boldsymbol{\nu}$ is equal to $\mathbf{0}$; and the system of equations above can be solved explicitly leveraging the spectral structure of L^+ to obtain

$$\alpha_i = (\mathbf{1}^T\boldsymbol{\omega})^{-1}(\mu_i + \omega_i) = \frac{\mu_i + \omega_i}{\sum_{j=1}^n \omega_j}.$$

This result corresponds to the unrealistic scenario in which the response of the controllable loads is such that the net load profile $p_i(\boldsymbol{\alpha})$ becomes zero in every node (and no power flows in the lines).

Figure 2. Traces of the Optimal Dynamic Load-Sharing Factors for the Three Controllable Loads $B = \{2, 6, 9\}$ in the IEEE 14-Bus Test Network of Figure 1 Over 100 Realizations of the Fluctuations $\boldsymbol{\omega}$

$\widetilde{\alpha}_2^*(\boldsymbol{\omega})$

$\widetilde{\alpha}_6^*(\boldsymbol{\omega})$

$\widetilde{\alpha}_9^*(\boldsymbol{\omega})$
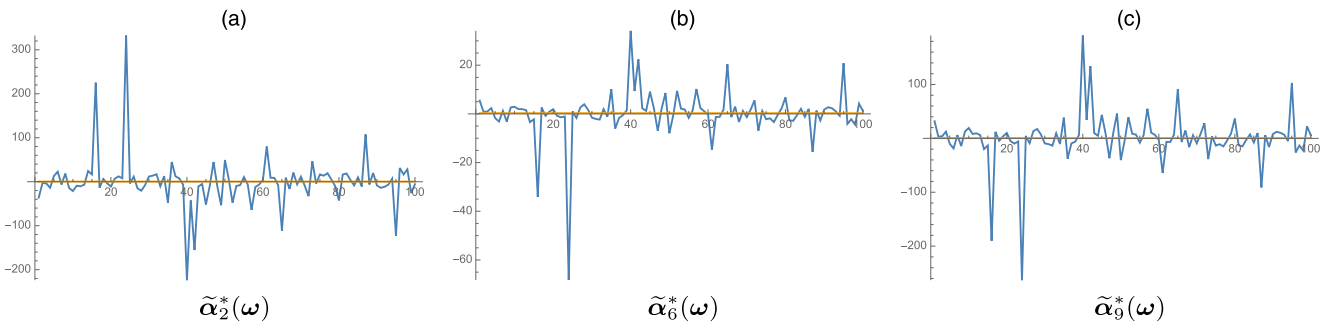
We thus focus the case where $B \subsetneq V$ and compare numerically the performance of the optimal static load-sharing factors α^* against the dynamically changing $\alpha^*(\omega)$ for the same realizations of the fluctuations ω. We consider the IEEE 14-bus test network, a standard test case used for simulations in the power grid literature (Zimmerman et al. 2011), with correlated noise (see Figure 1).

The resulting static optimal control for this network and correlation structure is $\widetilde{\alpha}^* = (\widetilde{\alpha}_2^*, \widetilde{\alpha}_6^*, \widetilde{\alpha}_9^*) = (0.3510, 0.2805, 0.3685)$. We sample 100 realizations of ω from a multivariate Gaussian distribution with mean $\mathbf{0}$ and correlation matrix Σ; for each of them, we calculate the total aggregated loss using $\widetilde{\alpha}^*$ and finding the optimal dynamic load-sharing factors $\widetilde{\alpha}^*(\omega)$. For these realizations, the average total loss using the dynamical control is equal to 18,250.8, whereas it results in a 20% higher average total loss, namely, 21,985.1, when using the static control. However, as shown in Figure 2, the dynamic control $\widetilde{\alpha}^*(\omega)$ takes incredibly large (both positive and negative) values, which are highly unrealistic to be implemented in practical settings.

6.2. Impact of Covariance Structure

The next two figures illustrate how the covariance structure influence the optimal control for a small network of 14 nodes, the IEEE 14-bus test network. Figure 3(a) covers the case in which all the nodes have stochastic load fluctuations, that is, $S = V$, whereas in Figure 3(b) we present a scenario where $S \subsetneq V$, where only the nodes in black are affected by fluctuations. In both figures, the nodes with controllable loads are $B = \{2, 6, 9\}$ and are drawn as squares.

For these two networks, we compare the optimal static control α^* for the three controllable loads in the various cases in which we vary the covariance structure of the noise (see Figure 4). The correlation matrix Σ_c in Figure 5(c) has been generated at random; the correlation matrix Σ_b used in case (b) is obtained using the diagonal entries of Σ_c; and that in case (a) is equal to $\Sigma_a = \sigma^2 I_{14}$, with $\sigma^2 = \mathbf{1}^T \Sigma_c \mathbf{1} = \mathbf{1}^T \Sigma_b \mathbf{1} = 255.75$. In this way, all the three correlation matrices have been rescaled so that their total variance is equal in all three cases, namely, $\sigma_a^2 = \sigma_b^2 = \sigma_c^2$. The correlation matrices for the second network are displayed in Figure 6; they have been obtained analogously; but only a subset of $|S| = 9$ nodes is affected by stochastic fluctuations, and the other ones have corresponding rows and columns equal to zero.

Figure 3. The IEEE 14-Bus Test Network with Three Controllable Loads $B = \{2, 6, 9\}$ (Depicted as Squares)
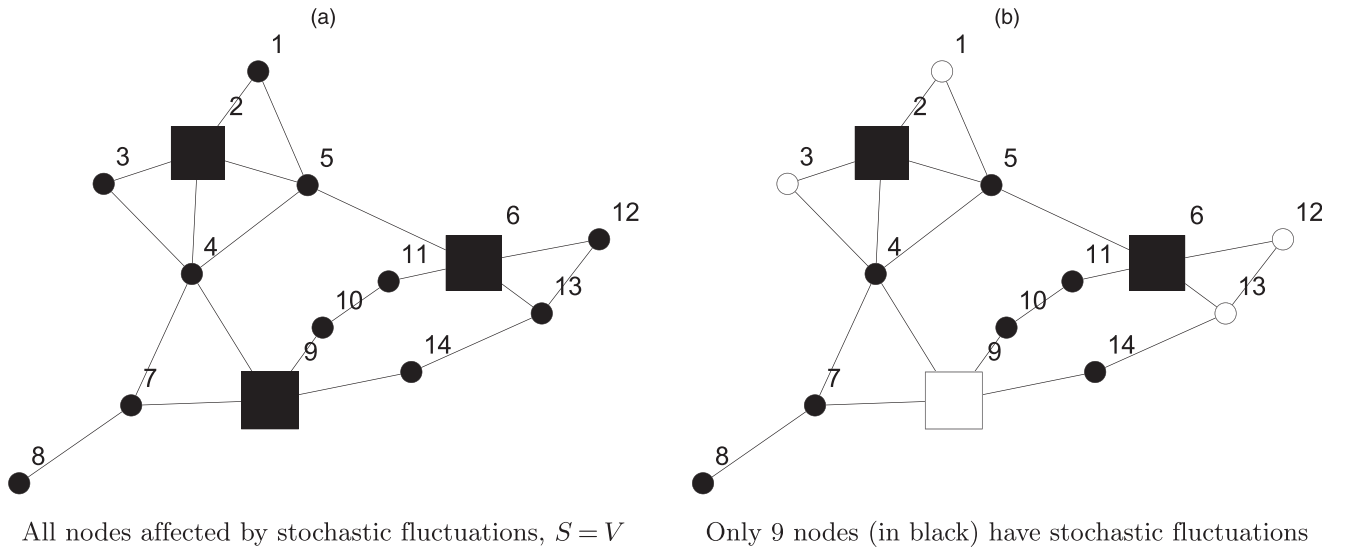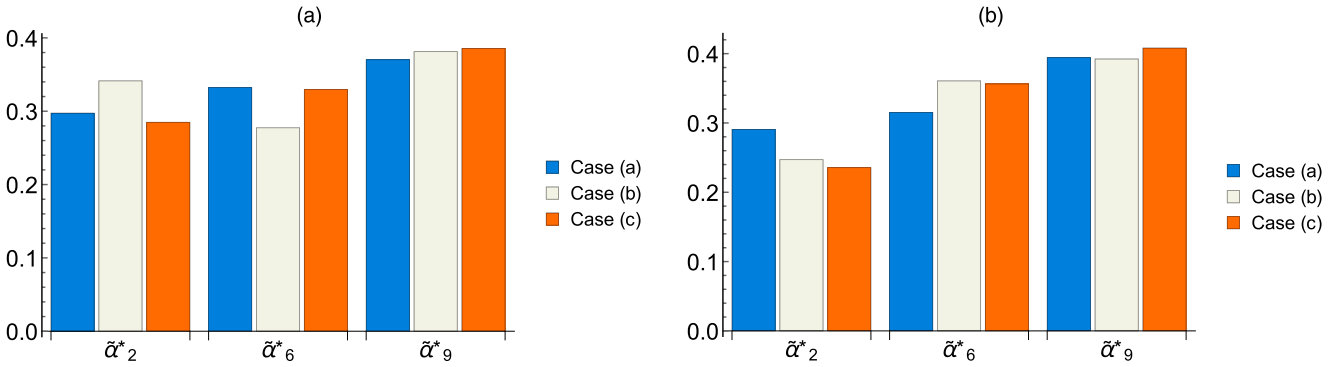

All nodes affected by stochastic fluctuations, $S = V$ Only 9 nodes (in black) have stochastic fluctuations

Figure 4. Values of the Optimal Load-Sharing Factors $\widetilde{\alpha}^*$ with Different Noise Correlation Structures

Values of $\boldsymbol{\alpha}^*$ for the three cases described in Figure 5   Values of $\widetilde{\boldsymbol{\alpha}}^*$ for the three cases described in Figure 6

6.3. Relative Position of Controllable Loads and Stochastic Nodes

As suggested by Theorem 1, the relative position of the nodes affected by fluctuations and that of the controllable loads play a crucial role in terms of the achievable total power loss, as we will illustrate in the following example on a ring network. Figure 7 visualizes the optimal load-sharing factors in the scenario where there are $|B| = 6$ controllable nodes and $|S| = 4$ nodes affected by fluctuations. The subset of nodes affected by fluctuation is fixed, $S = \{1, 4, 7, 10\}$, as well as the covariance matrix; but the location of the controllable loads, that is, the subset B, changes.

Figure 7(a) presents the scenario in which $S \subset B$ (i.e., all the stochastic nodes have controllable loads) and the optimal control $\boldsymbol{\alpha}^*$ for all other nodes in $B \setminus S = \{5, 11\}$ are equal to zero, as prescribed by Theorem 1. In the other two cases, in Figure 7, (b) and (c), we picked two different subsets B of controllable loads such that $S \not\subset B$ and $S \subset B^c$, respectively. The corresponding values of the expected loss are higher in these cases than in case (a) and suggest that it may be optimal to place the controllable loads in the nodes affected by stochastic fluctuations. Lastly, note in Figure 7(b) that the load-sharing factors for node one and seven are much larger than the other nodes in B, because these two nodes are affected by stochastic fluctuations but the remaining four are not.

Figure 5. Scatter Plots of the Correlation Matrices in Three Different Scenarios and Resulting Optimal Load-Sharing Factors for the Three Controllable Loads in $B = \{2, 6, 9\}$ for the Network in Figure 3(a)

Σ_{a} (i.i.d. load fluctuations) Σ_{b} (independent but not identi- Σ_{c} (correlated load fluctuations)

$\widetilde{\boldsymbol{\alpha}}^* = (0.2974, 0.3324, 0.3702)$ cally distributed load fluctuations) $\widetilde{\boldsymbol{\alpha}}^* = (0.2849, 0.3294, 0.3857)$

$\widetilde{\boldsymbol{\alpha}}^* = (0.3414, 0.2774, 0.3812)$

Figure 6. Scatter Plots of the Correlation Matrices in Three Different Scenarios and Resulting Optimal Load-Sharing Factors for the Three Controllable Loads in $B = \{2, 6, 9\}$ for the Network in Figure 3(b)
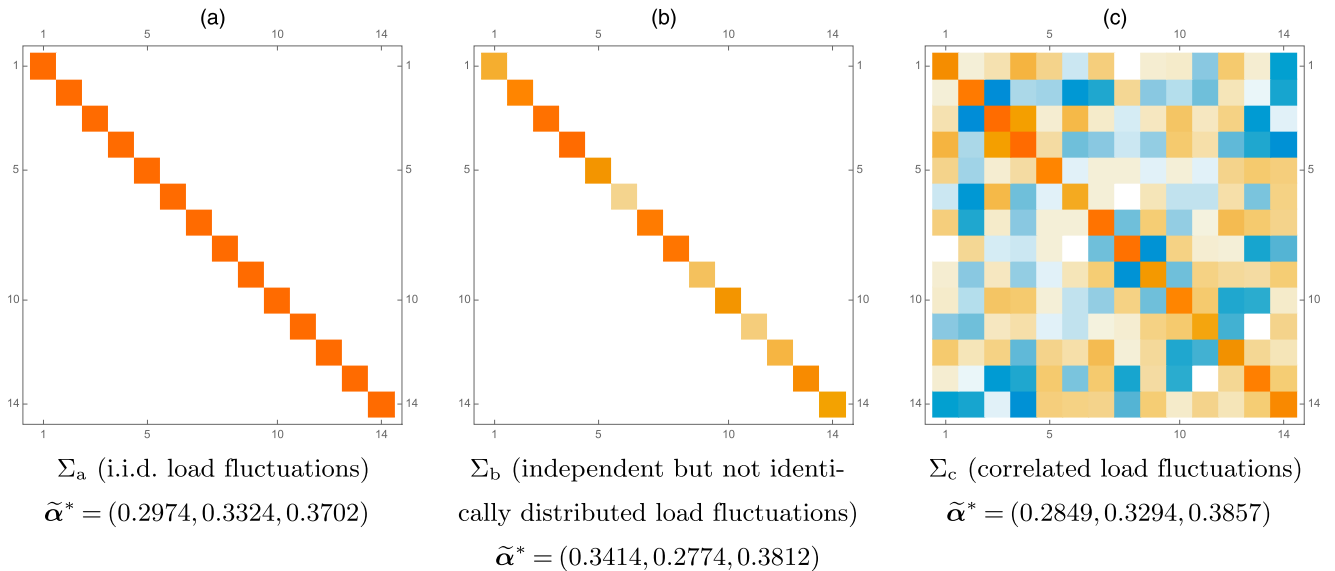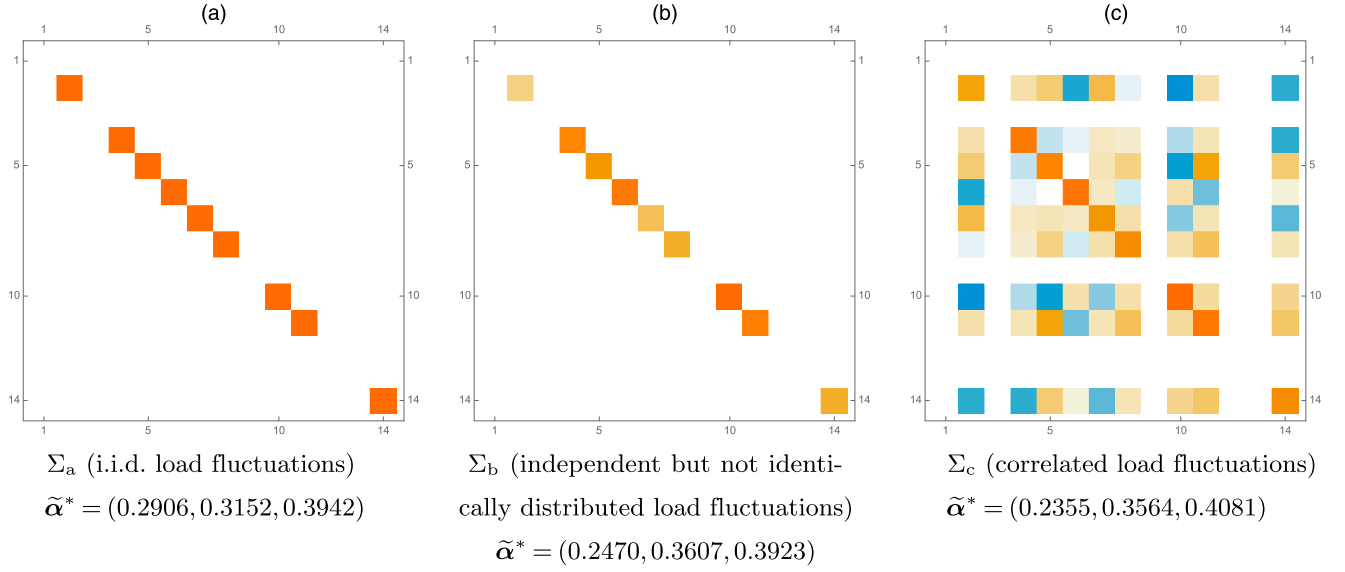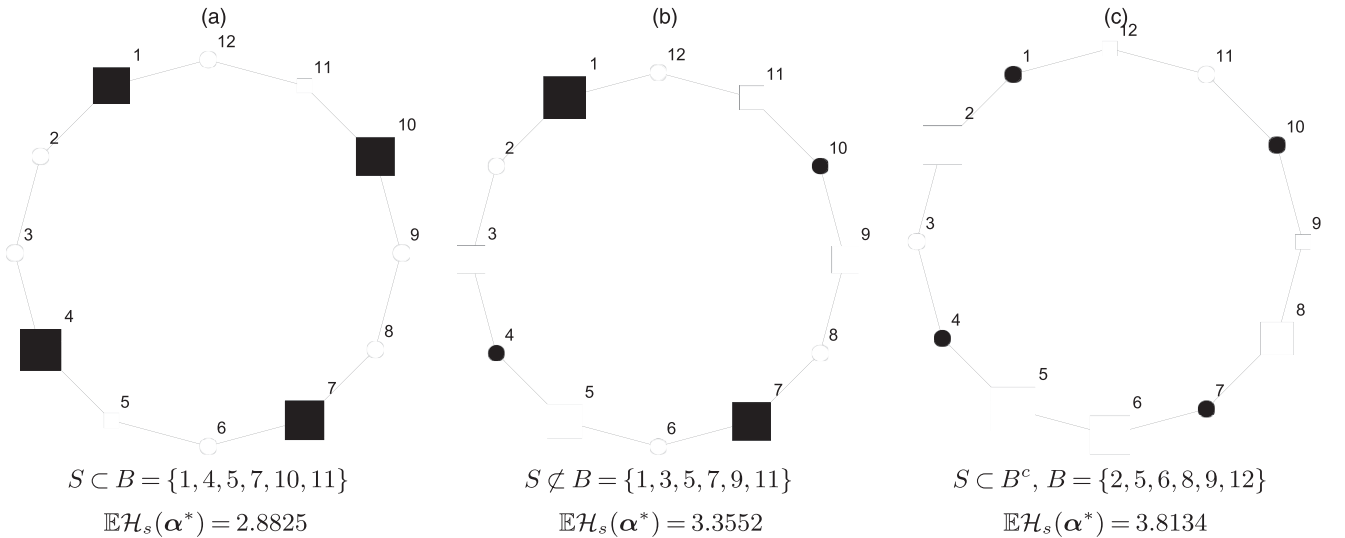

Σ_a (i.i.d. load fluctuations)

$\widetilde{\alpha}^* = (0.2906, 0.3152, 0.3942)$

Σ_b (independent but not identically distributed load fluctuations)

$\widetilde{\alpha}^* = (0.2470, 0.3607, 0.3923)$

Σ_c (correlated load fluctuations)

$\widetilde{\alpha}^* = (0.2355, 0.3564, 0.4081)$

Figure 7. Different Locations of Controllable Loads B (Depicted as Squares with Area Proportional to the Optimal Load-Sharing Factors α^*) and Corresponding Expected Total Loss

$S \subset B = \{1, 4, 5, 7, 10, 11\}$

$\mathbb{E}\mathcal{H}_s(\alpha^*) = 2.8825$

$S \not\subset B = \{1, 3, 5, 7, 9, 11\}$

$\mathbb{E}\mathcal{H}_s(\alpha^*) = 3.3552$

$S \subset B^c, B = \{2, 5, 6, 8, 9, 12\}$

$\mathbb{E}\mathcal{H}_s(\alpha^*) = 3.8134$

6.4. Negative Load-Sharing Factors

The fact that a load-sharing factor is nonnegative means that the corresponding node absorbs part of the power excess (if $\sum_{i=1}^n \omega_i > 0$) and balance out shortages (if $\sum_{i=1}^n \omega_i < 0$). In most of the related work in primary response mechanisms and automatic generation controls for power grids, the participation factors (that our load-sharing factors generalize) are in fact taken to be nonnegative, that is, $\alpha_v \geq 0$ for all $v \in B$. This assumption tacitly implies that all the controllable generators and storage have "coordinated" actions, that is, they either all increase or all decrease their power output.

   In our formulation of the optimization problems (9) and (13), we do not make such an assumption and load-sharing factors can also be negative, as long as the condition (4) is met. This is crucial, as for certain covariance structures of the load fluctuations (especially when there are strong negative correlations), it is optimal to have negative load-sharing factors in some nodes: we illustrate this fact for a small network in Figure 8.

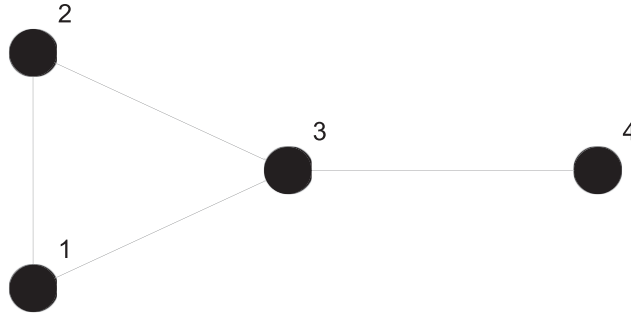**Figure 8.** A Small Network Modeled by a Graph with $n = 4$ Nodes and $m = 4$ Edges with Unit Weights

Table 1. Sets of Controllable Loads for the Network in Figure 8 and Corresponding Optimal Load-Sharing Factors

B	$\boldsymbol{\alpha}^*(B)$
$\{1, 4\}$	$(6/5, 0, 0, -1/5)$
$\{1, 2, 4\}$	$(2/3, 2/3, 0, -1/3)$
$\{1, 2, 3, 4\}$	$(1/2, 1/2, 1/2, -1/2)$

Table 1 lists the optimal load-sharing factors $\boldsymbol{\alpha}^*$ corresponding to a different set of controllable loads in the network in Figure 8 where the load fluctuations covariance structure is assumed to be

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 & -0.5 \\ 0 & 1 & 0 & -0.5 \\ 0 & 0 & 1 & -0.5 \\ -0.5 & -0.5 & -0.5 & 1 \end{pmatrix}.$$

The best way for the controllable loads to respond to the negative correlations that the load fluctuations have in this network is having the controllable load in node four taking actions "mirroring" those of the other three nodes, in the sense that $\alpha_4^* < 0$, whereas the load-sharing factors of the other nodes in B are always positive.

6.5. Nonmonotonicity of Expected Power Loss When Adding Controllable Loads

An extra controllable load always reduces the expected total power loss if the corresponding optimal load-sharing vector $\boldsymbol{\alpha}^*$ is selected, because it corresponds to removing one constraint in the optimization problem (9). However, if the chosen load-sharing factors of the augmented subset of controllable loads are *not* the optimal ones, adding an extra controllable load does not necessarily reduce the expected total power loss. We illustrate this fact with an example in which the control is always assumed to be equal-share between the controllable loads in B, that is, $\widetilde{\boldsymbol{\alpha}} = \frac{1}{|B|}\mathbf{1}$. Consider the network given in Figure 9 and assume that the stochastic loads are i.i.d. with unit variance.

Table 2 lists the expected total power losses for some subsets B of controllable loads and compares them with those for some augmented subset $B \cup \{4\}$. It is evident that in every one of these case, adding an

Figure 9. A Small Network Modeled by a Graph with $n = 7$ Nodes and $m = 9$ Edges with Unit Weights
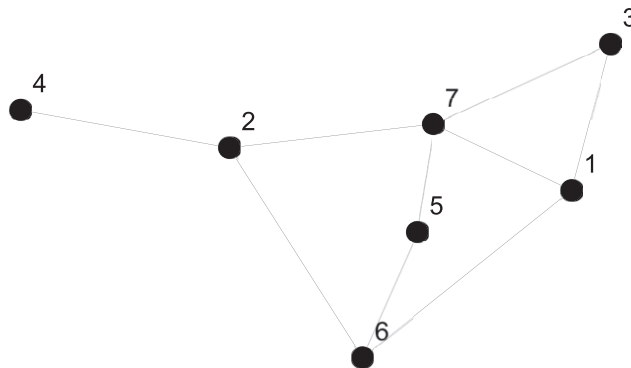
Table 2. Expected Total Power Losses for the Network in Figure 9 Assuming Equal Load-Sharing Factors for Some Subsets B of Controllable Loads and Then for the Subsets Augmented with an Extra Node, Namely, $B \cup \{4\}$

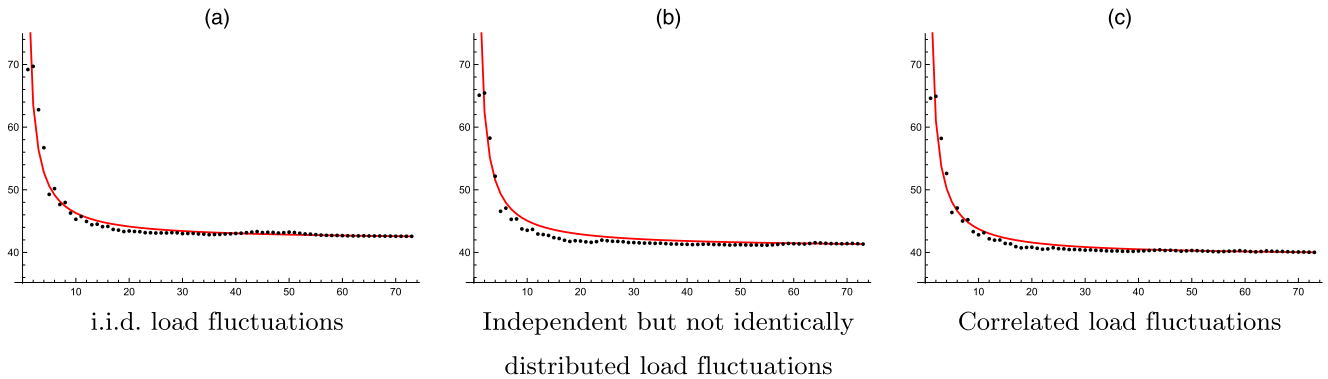| B	$\mathbb{E}\mathcal{H}(B)$	$\mathbb{E}\mathcal{H}(B \cup \{4\})$
$\{2\}$	2.6875	3.0625
$\{2,5\}$	2.0625	2.1319
$\{2,6,7\}$	1.7986	1.8438

additional controllable load in node four without optimally readjusting the load-sharing factors results in a higher expected total loss.

6.6. Empirical Evidence of the Scaling Law

The scaling law derived in Section 4, despite having been obtained by averaging over the possible locations of the controllable loads, still gives precious insight about how the average total loss decreases with the number of controllable loads. Aiming to corroborate this fact, we consider the IEEE RTS 96-bus test network (Zimmerman et al. 2011) and track the average total loss while adding one by one controllable loads in random locations and assuming equal share among them.

As illustrated by Figure 10, the theoretical scaling for the expected total loss scales with the number of controllable loads stated in Theorem 3, while averaging all possible locations, is also in fact very accurate for a single instance where new controllable locations are randomly added.

Figure 10. Theoretical Scaling (in Red) of the Expected Total Loss as Predicted by Theorem 3 vs. Empirical Total Loss (in Black) While Adding Controllable Loads One by One in Random Locations

(a)	(b)	(c)
i.i.d. load fluctuations	Independent but not identically distributed load fluctuations	Correlated load fluctuations

7. Concluding Remarks

In this paper, we consider a stochastic lossy transport network in which some nodes have controllable loads and derive a closed-form expression for the optimal control when aiming to minimize the average total loss. The model is inspired by power systems where distributed energy resources can be used as virtual storage to mitigate the fluctuations in the power generated by renewable energy sources and in power demand. Our analysis unveils the complex interplay between the network structure, the location of the controllable loads, and the covariance structure of the power fluctuations and gives insight into how much the average total loss can be reduced by adding a given number of controllable loads to the network.

We derived explicit optimal load-sharing factors for controllable loads in various scenarios. Our analysis, even if it uses a stylized mathematical model, suggests that the optimal displacement and operations of distributed energy resources must account for the possible correlations of the power fluctuations. For this reason, it complements the recent efforts in the electrical engineering community in upgrading the existing models for power grids to account both for the intrinsic volatility of renewable energy generation and storage capabilities (see, e.g., Kanoria et al. 2011; Lin and Bitar 2016a, b, 2017; Lorca et al. 2016; Lorca and Sun 2017).

Lastly, we presented a more general optimization problem that can be instrumental to explore numerically the trade-off between the best operations for the network and the corresponding cost or penalties for the excessive usage of the controllable nodes. This is particularly relevant for the design of primary response

mechanisms and automatic generation controls for power grids (Apostolopoulou et al. 2014, Lubin et al. 2016, Roald et al. 2016a, Sundar et al. 2016, Chertkov and Dvorkin 2017, Guggilam et al. 2017a, b).

Appendix A. Properties of the Total Power Loss

Proof of Proposition 1. Using the expression for C_α given in (3), we can rewrite $\mathcal{H}(\alpha)$ as

$$\mathcal{H}(\alpha) = \frac{1}{2}\left(\mu + C_\alpha \omega\right)^T L^+ \left(\mu + C_\alpha \omega\right) = \frac{1}{2}\mu^T L^+ \mu + \frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega + \mu^T L^+ C_\alpha \omega.$$

Note that one of the terms on the right-hand side, namely, $\frac{1}{2}\mu^T L^+ \mu$, is not random and does not depend on the control α. Define the random variable

$$\mathcal{H}_s(\alpha) := \mathcal{H}(\alpha) - \frac{1}{2}\mu^T L^+ \mu = \frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega + \mu^T L^+ C_\alpha \omega, \tag{A.1}$$

which describes precisely the contribution of the stochastic fluctuations to the transportation losses. From the fact that L^+ is a positive semidefinite matrix it follows that

$$\frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega = \frac{1}{2}(C_\alpha \omega)^T L^+ (C_\alpha \omega) \geq 0 \quad \forall\, \omega \in \mathbb{R}^n. \tag{A.2}$$

Combining (A.1) and (A.2), and the fact that $\mathbb{E}(\mu^T L^+ C_\alpha \omega) = \mu L^+ C_\alpha \mathbb{E}\omega = 0$, yields that

$$\mathbb{E}\mathcal{H}_s(\alpha) = \mathbb{E}\left(\frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega\right) \geq 0.$$

Applying a classic result for quadratic forms of a random vector (see, e.g., Mathai and Provost 1992, corollary 3.2b.1), we derive

$$\mathbb{E}\left(\frac{1}{2}\omega^T C_\alpha^T L^+ C_\alpha \omega\right) = \frac{1}{2}\,\text{tr}\left(C_\alpha^T L^+ C_\alpha \sum\right). \tag{A.3}$$

Because $\mathbb{E}\omega = 0$, it follows that $\mathbb{E}(\mu^T L^+ C_\alpha \omega) = \mu^T L^+ C_\alpha \mathbb{E}\omega = 0$; thus, identity (A.1) can be rewritten as

$$\mathbb{E}\mathcal{H}_s(\alpha) = \frac{1}{2}\,\text{tr}\left(C_\alpha^T L^+ C_\alpha \sum\right).$$

We now derive identity (6). Recall the following well-known properties of the trace of a matrix:
i. The trace is invariant under cyclic permutations, that is, for any $r \in \mathbb{N}$,

$$\text{tr}(A_1 \ldots A_r) = \text{tr}(A_2 \ldots A_r A_1) = \ldots = \text{tr}(A_r A_1 \ldots A_{r-1}).$$

ii. The trace of a matrix and of its transpose coincide, that is, $\text{tr}(A) = \text{tr}(A^T)$.
iii. The trace of the outer product of two vectors is their inner product, namely,

$$\text{tr}\left(v w^T\right) = \text{tr}(v \otimes w) = v^T w.$$

First note that we can rewrite

$$C_\alpha^T L^+ C_\alpha = \left(I - \alpha \mathbf{1}^T\right)^T L^+ \left(I - \alpha \mathbf{1}^T\right) = \left(I - \mathbf{1}\alpha^T\right)\left(L^+ - L^+ \alpha \mathbf{1}^T\right)$$
$$= L^+ - L^+ \alpha \mathbf{1}^T - \mathbf{1}\alpha^T L^+ + \mathbf{1}\alpha^T L^+ \alpha \mathbf{1}^T. \tag{A.4}$$

The aforementioned properties of the trace yield

$$\text{tr}\left(\sum L^+ \alpha \mathbf{1}^T\right) \overset{(ii)}{=} \text{tr}\left(\left(\sum L^+ \alpha \mathbf{1}^T\right)^T\right) = \text{tr}\left(\mathbf{1}\alpha^T L^+ \sum\right) \overset{(i)}{=} \text{tr}\left(\sum \mathbf{1}\alpha^T L^+\right), \tag{A.5}$$

and

$$\text{tr}\left(\sum \mathbf{1}\alpha^T L^+\right) = \text{tr}\left(\sum \mathbf{1}(L^+ \alpha)^T\right) = \text{tr}\left(\left(\sum \mathbf{1}\right) \otimes (L^+ \alpha)\right) \overset{(iii)}{=} \left(\sum \mathbf{1}\right)^T (L^+ \alpha) \overset{(iii)}{=} \mathbf{1}^T \sum L^+ \alpha. \tag{A.6}$$

By combining all these equalities and exploiting the linearity of the trace operator, we obtain

$$\operatorname{tr}\left(\sum C_\alpha^T L^+ C_\alpha\right) \overset{(20)}{=} \operatorname{tr}\left(\sum L^+\right) - \operatorname{tr}\left(\sum L^+ \boldsymbol{\alpha} \mathbf{1}^T\right) - \operatorname{tr}\left(\sum \mathbf{1} \boldsymbol{\alpha}^T L^+\right) + \operatorname{tr}\left(\sum \mathbf{1} \boldsymbol{\alpha}^T L^+ \boldsymbol{\alpha} \mathbf{1}^T\right)$$

$$\overset{(21)}{=} \operatorname{tr}\left(\sum L^+\right) - 2 \cdot \operatorname{tr}\left(\sum \mathbf{1} \boldsymbol{\alpha}^T L^+\right) + \left(\boldsymbol{\alpha}^T L^+ \boldsymbol{\alpha}\right) \cdot \operatorname{tr}\left(\sum \mathbf{1} \mathbf{1}^T\right)$$

$$\overset{(22)}{=} \operatorname{tr}\left(\sum L^+\right) - 2 \cdot \left(\mathbf{1}^T \sum L^+ \boldsymbol{\alpha}\right) + \sigma^2 \left(\boldsymbol{\alpha}^T L^+ \boldsymbol{\alpha}\right),$$

where we also used the fact that $\boldsymbol{\alpha}^T L^+ \boldsymbol{\alpha}$ is a scalar in the second step and identity (2) in the third step. $\quad\square$

Proof of Proposition 2. Assume that $e = (i, j) \in (V \times V)$ is the edge with weight $\beta > 0$ that has been added to G or whose edge weight has been increased by $\beta > 0$ and let $\boldsymbol{m}_e = (\boldsymbol{e}_i - \boldsymbol{e}_j) \in \mathbb{R}^n$ be the corresponding nonweighted incidence vector. In both cases, the Laplacian matrix of the newly obtained graph G' can be written as

$$L_{G'} = L_G + \beta \, \boldsymbol{m}_e \boldsymbol{m}_e^T$$

and, using the generalized version of the Sherman-Morrison formula in (Meyer 1973), we get

$$L_{G'}^+ = L_G^+ - \frac{1}{\beta^{-1} + \boldsymbol{m}_e^T L_G^+ \boldsymbol{m}_e} L_G^+ \boldsymbol{m}_e \boldsymbol{m}_e^T L_G^+.$$

We can thus rewrite the total loss corresponding to any net load profile $\mathbf{p}(\boldsymbol{\alpha})$ as

$$\mathcal{H}^{G'}(\boldsymbol{\alpha}) = \frac{1}{2} \mathbf{p}(\boldsymbol{\alpha})^T L_{G'} \mathbf{p}(\boldsymbol{\alpha})$$

$$= \frac{1}{2} \mathbf{p}(\boldsymbol{\alpha})^T L_G^+ \mathbf{p}(\boldsymbol{\alpha}) - \frac{1}{2\left(\beta^{-1} + \boldsymbol{m}_e^T L_G^+ \boldsymbol{m}_e\right)} \mathbf{p}(\boldsymbol{\alpha})^T L_G^+ \boldsymbol{m}_e \boldsymbol{m}_e^T L_G^+ \mathbf{p}(\boldsymbol{\alpha})$$

$$= \mathcal{H}^G(\boldsymbol{\alpha}) - \frac{\left(\boldsymbol{m}_e^T L_G^+ \mathbf{p}(\boldsymbol{\alpha})\right)^2}{2\left(\beta^{-1} + \boldsymbol{m}_e^T L_G^+ \boldsymbol{m}_e\right)},$$

and conclude by noticing that $\boldsymbol{m}_e^T L_G^+ \boldsymbol{m}_e \geq 0$ and $(\boldsymbol{m}_e^T L_G^+ \mathbf{p}(\boldsymbol{\alpha}))^2 \geq 0$. $\quad\square$

Appendix B. Proof of Theorem 1

In this proof, we use the so-called block matrix inversion formula, which is stated in the next lemma.

Lemma B.1 (Block Matrix Inversion Formula). Consider a matrix with the block structure $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$. If both A and $D - CA^{-1}B$ are nonsingular matrices, then

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} - A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}.$$

For any vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ that is such that $\boldsymbol{\alpha}_i = 0$ for every node $i \in V \setminus B$, there exists a unique k-dimensional vector $\widetilde{\boldsymbol{\alpha}} \in \mathbb{R}^k$ such that $\boldsymbol{\alpha} = P_B \widetilde{\boldsymbol{\alpha}}$. Using this correspondence and the fact that the nominal load profile is balanced, that is, $\mathbf{1}^T \boldsymbol{\mu} = 0$, we can rewrite

$$\mathbb{E} \mathcal{H}_s(\boldsymbol{\alpha}) = \mathbb{E} \mathcal{H}(P_B \widetilde{\boldsymbol{\alpha}}) \overset{(5)}{=} \frac{\sigma^2}{2} \left(\widetilde{\boldsymbol{\alpha}}^T P_B^T L^+ P_B \widetilde{\boldsymbol{\alpha}}\right) - \mathbf{1}^T \sum L^+ P_B \widetilde{\boldsymbol{\alpha}} + \frac{1}{2} \operatorname{tr}\left(\sum L^+\right).$$

Therefore, the n-dimensional optimization problem (9) rewrites as a k-dimensional optimization problem with a single constraint, namely,

$$\min_{\widetilde{\boldsymbol{\alpha}} \in \mathbb{R}^k} \quad \frac{\sigma^2}{2} \left(\widetilde{\boldsymbol{\alpha}}^T P_B^T L^+ P_B \widetilde{\boldsymbol{\alpha}}\right) - \mathbf{1}^T \sum L^+ P_B \widetilde{\boldsymbol{\alpha}}$$

$$\text{s.t.} \quad \mathbf{1}^T \widetilde{\boldsymbol{\alpha}} = 1. \tag{B.1}$$

The matrix $L_B^+ := P_B^T L^+ P_B \in \mathbb{R}^{k \times k}$ is positive definite, as for any vector $v \in R^k$, $v \neq \mathbf{0}$,

$$v^T L_B^+ v = v^T P_B^T L^+ P_B v = (v \quad \mathbf{0}) L^+ \begin{pmatrix} v \\ \mathbf{0} \end{pmatrix} > 0,$$

where the last inequality follows from the fact that the vector $(v\ \ 0)$ is not a multiple of the vector $\mathbf{1}$ and thus does not lie in the null space of L^+. The optimization problem in (B.1) has then a unique solution, because the corresponding Hessian is positive definite.

Let $\gamma \in \mathbb{R}$ be the Lagrange multiplier γ associated with the unique equality constraint of the optimization problem (B.1). The associated Lagrangian associated is

$$\mathcal{L}(\widetilde{\alpha}, \gamma) = \frac{\sigma^2}{2} \widetilde{\alpha}^T \left(P_B^T L^+ P_B \right) \widetilde{\alpha} - \left(P_B^T L^+ \sum \mathbf{1} \right)^T \widetilde{\alpha} - \gamma \left(\mathbf{1}^T \widetilde{\alpha} - 1 \right).$$

Setting $b := P_B^T L^+ \Sigma \mathbf{1} \in \mathbb{R}^k$, the optimality conditions read

$$\begin{cases} \sigma^2 L_B^+ \widetilde{\alpha} - \gamma \mathbf{1} = b, \\ \mathbf{1}^T \widetilde{\alpha} = 1, \end{cases}$$

or, equivalently, in matrix form

$$\begin{pmatrix} \sigma^2 L_B^+ & -\mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\alpha} \\ \gamma \end{pmatrix} = \begin{pmatrix} b \\ 1 \end{pmatrix}. \tag{B.2}$$

Being positive definite, L_B^+ is invertible and its inverse is also positive definite, which means that $t_B := \mathbf{1}^T (L_B^+)^{-1} \mathbf{1} > 0$. In view of the fact that $\sigma^2 t_B^{-1} \neq 0$ and L_B^+ are invertible, we can use the block matrix inversion formula given in Lemma B.1 to obtain

$$\begin{pmatrix} \sigma^2 L_B^+ & -\mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\sigma^2} (L_B^+)^{-1} \left(I - \frac{1}{t_B} \mathbf{1}\mathbf{1}^T (L_B^+)^{-1} \right) & \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} \\ \frac{1}{t_B} \mathbf{1}^T (L_B^+)^{-1} & \frac{\sigma^2}{t_B} \end{pmatrix}.$$

The solution of the linear system (B.2) then reads

$$\begin{pmatrix} \widetilde{\alpha} \\ \gamma \end{pmatrix} = \begin{pmatrix} \sigma^2 L_B^+ & -\mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} b \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} (L_B^+)^{-1} \left(I - \frac{1}{t_B} \mathbf{1}\mathbf{1}^T (L_B^+)^{-1} \right) & \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} \\ \frac{1}{t_B} \mathbf{1}^T (L_B^+)^{-1} & \frac{\sigma^2}{t_B} \end{pmatrix} \begin{pmatrix} b \\ 1 \end{pmatrix};$$

thus, the optimal load-sharing vector $\widetilde{\alpha} \in \mathbb{R}^k$ is given by

$$\begin{aligned} \widetilde{\alpha}^* &= \frac{(L_B^+)^{-1}}{\sigma^2} \left(I - \frac{1}{t_B} \mathbf{1}\mathbf{1}^T (L_B^+)^{-1} \right) b + \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} \\ &= \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} + (L_B^+)^{-1} \left(I - \frac{1}{t_B} \mathbf{1}\mathbf{1}^T (L_B^+)^{-1} \right) \frac{b}{\sigma^2} \\ &= \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} + \left((L_B^+)^{-1} - \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1}\mathbf{1}^T (L_B^+)^{-1} \right) P_B^T L^+ \frac{\Sigma \mathbf{1}}{\sigma^2} \\ &= \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1} + \left(I - \frac{1}{t_B} (L_B^+)^{-1} \mathbf{1}\mathbf{1}^T \right) (L_B^+)^{-1} P_B^T L^+ \frac{\Sigma \mathbf{1}}{\sigma^2}. \end{aligned}$$

We now focus on the special case where $S \subseteq B$ and prove identity (11). Rewrite L^+ as a block matrix

$$\begin{pmatrix} L_B^+ & L_C^+ \\ (L_C^+)^T & L_{B^c}^+ \end{pmatrix},$$

with $L_B^+ \in \mathbb{R}^{k \times k}$, $L_C^+ \in \mathbb{R}^{k \times n-k}$, and $L_{B^c}^+ \in \mathbb{R}^{n-k \times n-k}$. Note that L_B^+ and $L_{B^c}^+$ are symmetric matrices, because L^+ is symmetric. This is consistent with the former definition of L_B^+, because

$$P_B^T L^+ P_B = (I_k | \mathbb{O}) L_+ (I_k | \mathbb{O})^T = (I_k | \mathbb{O})^T \begin{pmatrix} L_B^+ & L_C^+ \\ (L_C^+)^T & L_{B^c}^+ \end{pmatrix} (I_k | \mathbb{O}) = L_B^+.$$

We start by noticing that

$$\left(L_B^+\right)^{-1} P_B^T L^+ = \left(L_B^+\right)^{-1}(I_k \mid \mathbb{O})\begin{pmatrix} L_B^+ & L_C^+ \\ (L_C^+)^T & L_{B^c}^+ \end{pmatrix} = \left(L_B^+\right)^{-1}\left(L_B^+ \mid L_C^+\right) = \left(I_k \mid \left(L_B^+\right)^{-1} L_C^+\right) \in \mathbb{R}^{k \times n}.$$

From the assumption $S \subseteq B$ it follows that the covariance matrix can be rewritten as

$$\sum = \begin{pmatrix} \sum_B & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix},$$

where $\sum_B \in \mathbb{R}^{k \times k}$ is itself a covariance matrix (hence, a symmetric positive semidefinite matrix). Trivially, $\mathbf{1}^T(\sum_B \mid \mathbb{O})\mathbf{1} = \text{tr}(\sum_B \mathbf{1}\mathbf{1}^T) = \text{tr}(\sum \mathbf{1}\mathbf{1}^T) = \sigma^2$. Furthermore,

$$\left(L_B^+\right)^{-1} P_B^T L^+ \sum = \left(L_B^+\right)^{-1} P_B^T L^+ \begin{pmatrix} \sum_B & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix} = \left(I_k \mid \left(L_B^+\right)^{-1} L_C^+\right)\begin{pmatrix} \sum_B & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix} = \left(\sum_B \mid \mathbb{O}\right).$$

The optimal control $\widetilde{\alpha}^*$ then rewrites as

$$\begin{aligned}
\widetilde{\alpha}^* &= \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1} + \left(I_k - \frac{\left(L_B^+\right)^{-1}\mathbf{1}\mathbf{1}^T}{t_B}\right)\left(L_B^+\right)^{-1} P_B^T L^+ \frac{\sum\mathbf{1}}{\sigma^2} \\
&= \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1} + \left(I_k - \frac{\left(L_B^+\right)^{-1}\mathbf{1}\mathbf{1}^T}{t_B}\right)\left(\sum_B \mid \mathbb{O}\right)\frac{\mathbf{1}}{\sigma^2} \\
&= \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1} + \left(\sum_B \mid \mathbb{O}\right)\frac{\mathbf{1}}{\sigma^2} - \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1}\frac{\mathbf{1}^T(\sum_B \mid \mathbb{O})\mathbf{1}}{\sigma^2} \\
&= \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1} + \frac{\sum\mathbf{1}}{\sigma^2} - \frac{\left(L_B^+\right)^{-1}}{t_B}\mathbf{1} \\
&= \frac{\sum\mathbf{1}}{\sigma^2}. \quad \square
\end{aligned}$$

Appendix C. Proof of Theorem 2

We already rewrote in (7) the objective function of the optimization problem (13) as a quadratic form in α,

$$\mathbb{E}\mathcal{H}_s(\alpha) = \frac{\sigma^2}{2}\alpha^T A\alpha - b^T \alpha + c,$$

where $A = L^+$, $b = L^+ \sum \mathbf{1} \in \mathbb{R}^n$, and $c = \text{tr}(\sum L^+)/2 \in \mathbb{R}^+$. Note that, for the purpose of solving the optimization problem (13), we can ignore the constant term c.

Denote by $0 = \lambda_1 < \lambda_2 < \ldots < \lambda_n$ the eigenvalues of the weighted Laplacian matrix L and let $v_1, v_2, \ldots, v_n$ be the corresponding orthonormal basis of eigenvectors. Consider the representation of the vector $\alpha \in \mathbb{R}^n$ in this basis, namely,

$$\alpha = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n, \tag{C.1}$$

with $a_1, \ldots, a_n \in \mathbb{R}$. In view of the fact that the rows of L sum up to zero, it immediately follows that $v_1 = \frac{1}{\sqrt{n}}\mathbf{1}$. From the constraint (4), that is, $\mathbf{1}^T \alpha = 1$, it immediately follows that $a_1 = 1/\sqrt{n}$. Indeed,

$$a_1 \cdot \sqrt{n} = a_1 \cdot \frac{\mathbf{1}^T \mathbf{1}}{\sqrt{n}} = a_1 \mathbf{1}^T v_1 = a_1 \mathbf{1}^T v_1 + a_2 v_1^T v_2 + \ldots + a_n v_1^T v_n = \mathbf{1}^T(a_1 v_1 + a_2 v_2 + \ldots + a_n v_n) = \mathbf{1}^T \alpha = 1.$$

Defining the real coefficients $\kappa_2, \ldots, \kappa_n$ as

$$\kappa_i := \left\langle \sum \mathbf{1}, v_i \right\rangle = \mathbf{1}^T \sum v_i, \qquad i = 2, \ldots, n,$$

and using the representation (C.1), the two terms of the quadratic form above rewrites as

$$\alpha^T L^+ \alpha = (a_1 v_1 + a_2 v_2 + \ldots + a_n v_n)^T L^+ (a_1 v_1 + a_2 v_2 + \ldots + a_n v_n) = a_1^2 v_1^T L^+ v_1 + \sum_{i=2}^n a_i^2 v_i^T L^+ v_i = \sum_{i=2}^n \frac{a_i^2}{\lambda_i},$$

and

$$\boldsymbol{b}^T \boldsymbol{\alpha} = \mathbf{1}^T \sum L^+ \boldsymbol{\alpha} = \mathbf{1}^T \sum L^+ (a_1 \boldsymbol{v}_1 + a_2 \boldsymbol{v}_2 + \ldots + a_n \boldsymbol{v}_n) = \sum_{i=2}^{n} \frac{a_i}{\lambda_i} \mathbf{1}^T \sum \boldsymbol{v}_i = \sum_{i=2}^{n} a_i \frac{\kappa_i}{\lambda_i},$$

where we used twice the fact that $L^+ \boldsymbol{v}_1 = \mathbf{0}$. The objective function thus can be rewritten as

$$\mathbb{E}\mathcal{H}(\boldsymbol{\alpha}) = \frac{\sigma^2}{2} \boldsymbol{\alpha}^T A \boldsymbol{\alpha} - \boldsymbol{b}^T \boldsymbol{\alpha} = \frac{\sigma^2}{2} \sum_{i=2}^{n} \frac{a_i^2}{\lambda_i} - \sum_{i=2}^{n} a_i \frac{\kappa_i}{\lambda_i} =: g(a_2, \ldots, a_n). \tag{C.2}$$

The optimization problem (13) is therefore equivalent to a unconstrained minimization problem in $n-1$ variables, $a_2, \ldots, a_n$, which can be expressed using the newly introduced function $g : \mathbb{R}^{n-1} \to \mathbb{R}$. The gradient of g can be calculated as $\nabla g(a_2, \ldots, a_n) = (\frac{\sigma^2}{\lambda_i} a_i - \frac{\kappa_i}{\lambda_i})_{i=2,\ldots,n}$, and the Hessian is the diagonal matrix $H(g) = \sigma^2 \cdot \text{diag}(\lambda_2^{-1}, \ldots, \lambda_n^{-1})$. The Hessian $H(g)$ is constant as it does not depend on $a_2, \ldots, a_n$ and is a positive definite matrix, because all its diagonal terms are positive, in view of the fact that $\lambda_i > 0$ for $i = 2, \ldots, n$ and that $\sigma^2 > 0$. The problem is then strictly convex, and any stationary point satisfying $\nabla g(a_2, \ldots, a_n) = \mathbf{0}$ would then be a minimum for the function g. Solving the optimality condition $\nabla g(a_2, \ldots, a_n) = \mathbf{0}$ yields

$$a_i^* = \frac{\kappa_i}{\sigma^2}, \qquad i = 2, \ldots, n.$$

Consequently, the optimal load-sharing factor vector $\boldsymbol{\alpha}^*$ is unique and is given by

$$\boldsymbol{\alpha}^* = \frac{1}{\sqrt{n}} \boldsymbol{v}_1 + \frac{1}{\sigma^2} \sum_{i=2}^{n} \kappa_i \boldsymbol{v}_i = \frac{1}{n} \mathbf{1} + \frac{1}{\sigma^2} \sum_{i=2}^{n} \kappa_i \boldsymbol{v}_i.$$

Lastly, setting $\boldsymbol{v}_1 := \frac{1}{\sqrt{n}} \mathbf{1}$ and $\kappa_1 := \langle \Sigma \mathbf{1}, \boldsymbol{v}_1 \rangle = \mathbf{1}^T \Sigma \boldsymbol{v}_1 = \mathbf{1}^T \Sigma \frac{1}{\sqrt{n}} \mathbf{1}$, the vector $\boldsymbol{\alpha}^*$ rewrites as

$$\boldsymbol{\alpha}^* = \frac{\kappa_1}{\sigma^2} \frac{1}{\sqrt{n}} \mathbf{1} + \frac{1}{\sigma^2} \sum_{i=2}^{n} \kappa_i \boldsymbol{v}_i = \frac{1}{\sigma^2} \left(\sum_{i=1}^{n} \kappa_i \boldsymbol{v}_i \right) = \frac{1}{\sigma^2} \left(\sum_{i=1}^{n} \langle \Sigma \mathbf{1}, \boldsymbol{v}_i \rangle \boldsymbol{v}_i \right) = \frac{\Sigma \mathbf{1}}{\sigma^2}. \quad \square$$

Appendix D. Proof of Theorem 3

The starting point of the proof is two identities that leverage the properties of the pseudoinverse L^+ of the graph Laplacian. Firstly,

$$\sum_{B \subseteq V : |B|=k} L^+ \left(\sum_{i \in B} \boldsymbol{e}_i \right) = L^+ \left(\sum_{B \subseteq V : |B|=k} \sum_{i \in B} \boldsymbol{e}_i \right) = L^+ \left(\sum_{i=1}^{n} \sum_{B \subseteq V : |B|=k} \boldsymbol{e}_i \mathbb{1}_{\{i \in B\}} \right)$$

$$= L^+ \left(\sum_{i=1}^{n} \boldsymbol{e}_i \sum_{B \subseteq V : |B|=k} \mathbb{1}_{\{i \in B\}} \right) = L^+ \left(\sum_{i=1}^{n} \boldsymbol{e}_i \binom{n-1}{k-1} \right)$$

$$= \binom{n-1}{k-1} L^+ \left(\sum_{i=1}^{n} \boldsymbol{e}_i \right) = \binom{n-1}{k-1} L^+ \mathbf{1} = \mathbf{0}, \tag{D.1}$$

where we use the fact that in a graph with n nodes, each node belongs to exactly $\binom{n-1}{k-1}$ subsets of k nodes. We further claim that

$$\sum_{B \subseteq V : |B|=k} \left(\sum_{i \in B} \boldsymbol{e}_i \right)^T L^+ \left(\sum_{i \in B} \boldsymbol{e}_i \right) = \binom{n-2}{k-1} \text{tr}(L^+), \tag{D.2}$$

with the convention that $\binom{n-2}{n-1} = 0$. Because

$$\left(\sum_{i \in B} \boldsymbol{e}_i \right)^T L^+ \left(\sum_{i \in B} \boldsymbol{e}_i \right) = \sum_{i \in B} \boldsymbol{e}_i^T L^+ \boldsymbol{e}_i + \sum_{i,j \in B, i \neq j} \boldsymbol{e}_i^T L^+ \boldsymbol{e}_j,$$

we can rewrite the left-hand side of (D.2) as

$$\sum_{B \subseteq V : |B|=k} \left(\sum_{i \in B} \boldsymbol{e}_i \right)^T L^+ \left(\sum_{i \in B} \boldsymbol{e}_i \right) = \sum_{B \subseteq V : |B|=k} \left(\sum_{i \in B} \boldsymbol{e}_i^T L^+ \boldsymbol{e}_i \right) + \sum_{B \subseteq V : |B|=k} \left(\sum_{i,j \in B, i \neq j} \boldsymbol{e}_i^T L^+ \boldsymbol{e}_j \right). \tag{D.3}$$

The first term on the right-hand side of (D.3) can be rewritten as

$$\sum_{B \subseteq V : |B|=k} \left(\sum_{i \in B} e_i^T L^+ e_i \right) = \sum_{i=1}^n e_i^T L^+ e_i \left(\sum_{B \subseteq V : |B|=k} \mathbb{1}_{\{i \in B\}} \right) = \sum_{i=1}^n \binom{n-1}{k-1} e_i^T L^+ e_i = \binom{n-1}{k-1} \sum_{i=1}^n e_i^T L^+ e_i$$

$$= \binom{n-1}{k-1} \sum_{i=1}^n L_{i,i}^+ = \binom{n-1}{k-1} \mathrm{tr}(L^+),$$

whereas the second term on the right-hand side of (D.3) is equal to

$$\sum_{B \subseteq V : |B|=k} \left(\sum_{i,j \in B, i \neq j} e_i^T L^+ e_j \right) = \sum_{i \neq j} e_i^T L^+ e_j \left(\sum_{B \subseteq V : |B|=k} \mathbb{1}_{\{i \in B, j \in B\}} \right) = \sum_{i=1}^n \binom{n-2}{k-2} e_i^T L^+ e_j$$

$$= \binom{n-2}{k-2} \sum_{i \neq j} e_i^T L^+ e_j = -\binom{n-2}{k-2} \mathrm{tr}(L^+).$$

In the last step, we used the fact that $\sum_{i \neq j} e_i^T L^+ e_j = -\mathrm{tr}(L^+)$, which immediately follows from

$$\sum_{i \neq j} e_i^T L^+ e_j + \mathrm{tr}(L^+) = \sum_{i \neq j} e_i^T L^+ e_j + \sum_i L_{i,i}^+ = \sum_{i \neq j} e_i^T L^+ e_j + \sum_i e_i^T L^+ e_i = \left(\sum_{k=1}^n e_k \right)^T L^+ \left(\sum_{k=1}^n e_k \right)$$

$$= \mathbf{1}^T L^+ \mathbf{1} = 0.$$

Hence, (D.3) rewrites as

$$\sum_{B \subseteq V : |B|=k} \left(\sum_{i \in B} e_i \right)^T L^+ \left(\sum_{i \in B} e_i \right) = \left(\binom{n-1}{k-1} - \binom{n-2}{k-2} \right) \mathrm{tr}(L^+) = \binom{n-2}{k-1} \mathrm{tr}(L^+),$$

which concludes the proof of identity (D.2).

Each load-sharing vector $\alpha \in \mathcal{B}_k$ can be written as $\alpha = \frac{1}{k} \sum_{i \in B} e_i$, for some $B \subseteq V$, $|B| = k$. By Proposition 1, the expected total loss due to stochastic fluctuations when using this load-sharing vector is given by

$$\mathbb{E} \mathcal{H}_s(\alpha) = \frac{\sigma^2}{2} \left(\frac{1}{k} \sum_{i \in B} e_i \right)^T L^+ \left(\frac{1}{k} \sum_{i \in B} e_i \right) - \mathbf{1}^T \sum L^+ \left(\frac{1}{k} \sum_{i \in B} e_i \right) + \frac{1}{2} \mathrm{tr}(\sum L^+).$$

Therefore,

$$\mathcal{H}_k = \frac{1}{2} \mathrm{tr}(\sum L^+) + \frac{1}{|\mathcal{B}_k|} \sum_{B \subseteq V : |B|=k} \left[\frac{\sigma^2}{2} \left(\frac{1}{k} \sum_{i \in B} e_i \right)^T L^+ \left(\frac{1}{k} \sum_{i \in B} e_i \right) - \mathbf{1}^T \sum L^+ \left(\frac{1}{k} \sum_{i \in B} e_i \right) \right].$$

Using (D.1) and (D.2), we get

$$\mathcal{H}_k = \frac{1}{2} \mathrm{tr}(\sum L^+) + \sigma^2 \frac{\mathrm{tr}(L^+)}{2} \frac{\binom{n-2}{k-1}}{k^2 \binom{n}{k}} = \frac{1}{2} \mathrm{tr}(\sum L^+) + \sigma^2 \frac{\mathrm{tr}(L^+)}{2} \frac{n-k}{k \cdot n(n-1)}$$

$$= \frac{1}{2} \mathrm{tr}(\sum L^+) + \sigma^2 \frac{\mathrm{tr}(L^+)}{2n(n-1)} \frac{1 - \frac{k}{n}}{\frac{k}{n}} = \frac{1}{2} \mathrm{tr}(\sum L^+) + \sigma^2 \frac{\mathrm{tr}(L^+)}{2(n-1)} \left(\frac{1}{k} - \frac{1}{n} \right). \quad \square$$

References

Apostolopoulou D, Sauer P, Dominguez-Garcia A (2014) Automatic generation control and its implementation in real time. *47th Hawaii Internat. Conf. System Sci.* (IEEE, Piscataway, NJ), 2444–2452.

Barooah P (2019) *Virtual Energy Storage from Flexible Loads: Distributed Control with QoS Constraints* (Springer International Publishing, Cham, Switzerland).

Bejan A, Gibbens R, Kelly F (2012) Statistical aspects of storage systems modelling in energy networks. *46th Annual Conf. Inform. Sci. Systems (CISS)* (IEEE, Piscataway, NJ), 1–6.

Bhaumik D, Crommelin D, Zwart B (2016) A computational method for optimizing storage placement to maximize power network reliability. Roeder TM, Frazier PI, Szechtman R, Zhou E, eds. *Proc. 2016 Winter Simulation Conf.* (IEEE, Piscataway, NJ), 883–894.

Bienstock D, Chertkov M, Harnett S (2014) Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Rev.* 56(3):461–495.

Bienstock D, Matke C, Munoz G, Yang S (2016) Robust linear control of nonconvex battery operation in transmission systems. Preprint, submitted October 29, https://arxiv.org/abs/1610.09432.

Chertkov M, Dvorkin Y (2017) Chance constrained optimal power flow with primary frequency response. *IEEE 56th Annual Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 4484–4489.

Cruise J, Zachary S (2018) The optimal control of storage for arbitrage and buffering, with energy applications. Mougeot M, Picard D, Tankov P, Plougonven R, Drobinski P eds. *Renewable Energy: Forecasting and Risk Management* (Springer International Publishing, Cham, Switzerland), 209–227.

Cruise JR, Flatley L, Zachary S (2018) Impact of storage competition on energy markets. *Eur. J. Oper. Res.* 269(3):998–1012.

Cruise JR, Gibbens RJ, Zachary S (2014) Optimal control of storage for arbitrage, with applications to energy systems. *48th Annual Conf. Inform. Sci. Systems* (IEEE, Piscataway, NJ), 1–6.

Daganzo CF (1997) *Fundamentals of Transportation and Traffic Operations* (Emerald Group Publishing Limited, Somerville, MA).

Deka D, Nagarajan H, Backhaus S (2017) Optimal topology design for disturbance minimization in power grids. *2017 Amer. Control Conf. (ACC)* (IEEE, Piscataway, NJ), 2719–2724.

Doyle PG, Snell JL (2000) *Random Walks and Electric Networks*. (The Mathmatical Association of America, Inc., Washington D.C.)

Ellens W, Spieksma F, Van Mieghem P, Jamakovic A, Kooij R (2011) Effective graph resistance. *Linear Algebra Appl.* 435(10):2491–2506.

Gast N, Tomozei DC, Le Boudec JY (2012) Optimal storage policies with wind forecast uncertainties. *ACM SIGMETRICS Performance Evaluation Rev.* 40(3):28–32.

Gast N, Tomozei D, Le Boudec J (2014) Optimal generation and storage scheduling in the presence of renewable forecast uncertainties. *IEEE Trans. Smart Grid* 5(3):1328–1339.

Gast N, Le Boudec JY, Proutière A, Tomozei DC (2013) Impact of storage on the efficiency and prices in real-time electricity markets. *Proc. 4th Internat. Conf. Future Energy Systems—e-Energy '13* (ACM, New York), 15–26.

Ghosh A, Boyd S, Saberi A (2008) Minimizing effective resistance of a graph. *SIAM Rev.* 50(1):37–66.

Guggilam S, Zhao C, Dall'Anese E, Chen Y, Dhople S (2017a) Primary frequency response with aggregated DERs. *2017 Amer. Control Conf. (ACC)* (IEEE, Piscataway, NJ), 3386–3393.

Guggilam SS, Zhao C, Dall'Anese E, Chen YC, Dhople SV (2017b) Engineering inertial and primary-frequency response for distributed energy resources. *2017 IEEE 56th Annual Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 5112–5118.

Guo L, Liang C, Low SH (2017) Monotonicity properties and spectral characterization of power redistribution in cascading failures. *55th Annual Allerton Conf. Commun. Control Comput. (Allerton)* (IEEE, Piscataway, NJ), 918–925.

Johnson J, Chertkov M (2010) A majorization-minimization approach to design of power transmission networks. *49th IEEE Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 3996–4003.

Kanoria Y, Montanari A, Tse D, Zhang B (2011) Distributed storage for intermittent energy sources: Control design and performance limits. *49th Annual Allerton Conf. Commun. Control, Comput.* (IEEE, Piscataway, NJ), 1310–1317.

Klein DJ, Randić M (1993) Resistance distance. *J. Math. Chemistry* 12(1):81–95.

Lai C, Low S (2013) The redistribution of power flow in cascading failures. *51st Annual Allerton Conf. Comm., Control, Comput. (Allerton)* (IEEE, Piscataway, NJ), 1037–1044.

Lin W, Bitar E (2016a) Decentralized control of distributed energy resources in radial distribution systems. *2016 IEEE Internat. Conf. Smart Grid Commun. (SmartGridComm)* (IEEE, Piscataway, NJ), 296–301.

Lin W, Bitar E (2016b) Performance bounds for robust decentralized control. *2016 Amer. Control Conf. (ACC)* (IEEE, Piscataway, NJ), 4323–4330.

Lin W, Bitar E (2017) Decentralized stochastic control of distributed energy resources. *IEEE Trans. Power Systems* 33(1):888–900.

Lorca A, Sun XA (2017) Multistage robust unit commitment with dynamic uncertainty sets and energy storage. *IEEE Trans. Power Systems* 32(3):1678–1688.

Lorca Á, Sun XA, Litvinov E, Zheng T (2016) Multistage adaptive robust optimization for the unit commitment problem. *Oper. Res.* 64(1):32–51.

Lubin M, Dvorkin Y, Backhaus S (2016) A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Trans. Power Systems* 31(5):3840–3849.

Mathai A, Provost S (1992) *Quadratic Forms in Random Variables: Theory and Applications*, Statistics: Textbooks and Monographs, vol. 126 (CRC Press, Boca Raton, FL.)

Mathias J, Bušić A, Meyn S (2017) Demand dispatch with heterogeneous intelligent loads. Bui T, ed. *Proc. 50th Annual Hawaii Internat. Conf. System Sci. (HICSS)*, 3138–3147.

Mathias J, Kaddah R, Bušić A, Meyn S (2016) Smart fridge/dumb grid? Demand dispatch for the power grid of 2020. Bui TX, Sprague RH, eds. *Proc. 49th Hawaii Internat. Conf. System Sci. (HICSS)* (IEEE, Piscataway, NJ), 2498–2507.

Matke C, Bienstock D, Muñoz G, Yang S, Kleinhans D, Sager S (2017) Robust optimization of power network operation: Storage devices and the role of forecast errors in renewable energies. Cherifi H, Gaito S, Quattrociocchi W, Sala A, eds. *Complex Networks & Their Applications V*, Studies in Computational Intelligence, vol. 693 (Springer, Cham, Switzerland), 809–820.

Meyer CD Jr (1973) Generalized inversion of modified matrices. *SIAM J. Appl. Math.* 24(3):315–323.

Meyn S, Barooah P, Bušić A, Ehren J (2013) Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in Florida. *52nd IEEE Conf. Decision Control* (IEEE, Piscataway, NJ), 6946–6953.

Meyn S, Barooah P, Bušić A, Chen Y, Ehren J (2015) Ancillary service to the grid using intelligent deferrable loads. *IEEE Trans. Automatic Control* 60(11):2847–2862.

Nesti T, Zocca A, Zwart B (2017) Line failure probability bounds for power grids. *2017 IEEE Power Engrg. Soc. General Meeting* (IEEE, Piscataway, NJ), 1–5.

Nesti T, Zocca A, Zwart B (2018) Emergent failures and cascades in power grids: A statistical physics perspective. *Phys. Rev. Lett.* 120(25):258301.

Ranjan G, Zhang ZL, Boley D (2014) Incremental computation of pseudo-inverse of Laplacian. Zhang Z, Wu L, Xu W, Du DZ, eds. *Combinatorial Optim. Appl.*, Lecture Notes in Computer Science, vol. 8881 (Springer International Publishing, Cham, Switzerland), 729–749.

Rencher AC, Schaalje GB (2008) *Linear Models in Statistics*, 2nd ed. (John Wiley & Sons, Hoboken, NJ).

Roald L, Misra S, Krause T, Andersson G (2016b) Corrective control to handle forecast uncertainty: A chance constrained optimal power flow. *IEEE Trans. Power Systems* 32(2):1626–1637.

Roald L, Andersson G, Misra S, Chertkov M, Backhaus S (2016a) Optimal power flow with wind power control and limited expected risk of overloads. *2016 Power Systems Comput. Conf. (PSCC)*, (IEEE, Piscataway, NJ), 1–7.

Sjodin E, Gayme D, Topcu U (2012) Risk-mitigated optimal power flow for wind powered grids. *2012 Amer. Control Conf. (ACC)* (IEEE, Piscataway, NJ), 4431–4437.

Sundar K, Nagarajan H, Lubin M, Roald L, Misra S, Bent R, Bienstock D (2016) Unit commitment with N-1 security and wind uncertainty. *2016 Power Systems Comput. Conf. (PSCC)* (IEEE, Piscataway, NJ), 1–7.

Taylor JA, Mathieu JL (2015) Uncertainty in demand response—Identification, estimation, and learning. Aleman DM, Thiele AC, eds. *The Operations Research Revolution*, TutORials in Operations Research (INFORMS, Catonsville, MD), 56–70.

Van De Ven P, Hegde N, Massoulié L, Salonidis T (2013) Optimal control of end-user energy storage. *IEEE Trans. Smart Grid* 4(2):789–797.

Van Mieghem P (2011) *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, UK).

Whittle P (2007) *Networks: Optimisation and Evolution* (Cambridge University Press, Cambridge, UK).

Wood A, Wollenberg B, Sheble G (2014) *Power Generation, Operation, and Control*, 3rd ed. (John Wiley & Sons, Hoboken, NJ).

Zhou B, Trinajstić N (2008) A note on Kirchhoff index. *Chemical Phys. Lett.* 455(1–3):120–123.

Zimmerman RD, Murillo-Sánchez CE, Thomas RJ (2011) Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Systems* 26(1):12–19.

Zocca A, Zwart B (2016) Minimizing heat loss in DC networks using batteries. *54th Annual Allerton Conf. Commun., Control, Comput.* (IEEE, Piscataway, NJ), 1306–1313.