Modeling and Online Control of Generalized Energy Storage Networks

Junjie Qin, Yinlam Chow, Jiyan Yang
Computational and Mathematical Engineering
Stanford University
Stanford, CA, 94305
{jqin, ychow, jiyan}@stanford.edu

Ram Rajagopal
Civil and Environmental Engineering
Stanford University
Stanford, CA, 94305
ramr@stanford.edu

ABSTRACT

The integration of intermittent and volatile renewable energy resources requires increased flexibility in the operation of the electric grid. Storage, broadly speaking, provides the flexibility of shifting energy over time; network, on the other hand, provides the flexibility of shifting energy over geographical locations. The optimal control of general storage networks in uncertain environments is an important open problem. The key challenge is that, even in small networks, the corresponding constrained stochastic control problems with continuous spaces suffer from curses of dimensionality, and are intractable in general settings. For large networks, no efficient algorithm is known to give optimal or near-optimal performance.

This paper provides an efficient and provably near-optimal algorithm to solve this problem in a very general setting. We study the optimal control of generalized storage networks, *i.e.*, electric networks connected to distributed generalized storages. Here generalized storage is a unifying dynamic model for many components of the grid that provide the functionality of shifting energy over time, ranging from standard energy storage devices to deferrable or thermostatically controlled loads. An online algorithm is devised for the corresponding constrained stochastic control problem based on the theory of Lyapunov optimization. We prove that the algorithm is near-optimal, and construct a semidefinite program to minimize the sub-optimality bound. The resulting bound is a constant that depends only on the parameters of the storage network and cost functions, and is independent of uncertainty realizations. Numerical examples are given to demonstrate the effectiveness of the algorithm.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Modeling techniques, Performance attributes; G.4 [**Mathematical Software**]: Algorithm design and analysis

General Terms

Algorithms, Design, Theory

Keywords

Energy systems, energy storage, stochastic control, online optimization

1. INTRODUCTION

To ensure a sustainable energy future, deep penetration of renewable energy generation is essential. Renewable energy resources, such as wind and solar, are intrinsically variable. Uncertainties associated with these intermittent and volatile resources pose a significant challenge to their integration into the existing grid infrastructure [1]. More flexibility, especially in shifting energy supply and/or demand across time and network, are desired to cope with the increased uncertainties.

Energy storage provides the functionality of shifting energy across time. A vast array of technologies, such as batteries, flywheels, pumped-hydro, and compressed air energy storages, are available for such a purpose [2, 3]. Furthermore, flexible or controllable demand provides another ubiquitous source of storage. Deferrable loads – including many thermal loads, loads of internet data-centers and loads corresponding to charging electric vehicles (EVs) over certain time interval [4] – can be interpreted as *storage of demand* [5]. Other controllable loads which can possibly be shifted to an earlier or later time, such as thermostatically controlled loads (TCLs), may be modeled and controlled as a storage with negative lower bound and positive upper bound on the storage level [6]. These forms of storage enable intertemporal shifting of excess energy supply and/or demand, and significantly reduce the reserve requirement and thus system costs.

On the other hand, shifting energy across a network, *i.e.*, moving excess energy supply to meet unfulfilled demand among different geographical locations with transmission or distribution lines, can achieve similar effects in reducing the reserve requirement for the system. Thus in practice, it is natural to consider these two effects together. Yet, it remains mathematically challenging to formulate a sound and tractable problem that accounts for these effects in electric grid operations. Specifically, due to the power flow and network constraints, control variables in connected buses are coupled. Due to the storage constraints, control variables in different time periods are coupled as well. On top of that, uncertainties associated with stochastic generation and de-

mand dramatically complicate the problem, because of the large number of recourse stages and the need to account for all probable realizations.

Two categories of approaches have been proposed in the literature. The first category is based on exploiting structures of specific problem instances, usually using dynamic programming. These structural results are valuable in providing insights about the system, and often lead to analytical solution of these problem instances. However, such approaches rely heavily on specific assumptions of the type of storage, the form of the cost function, and the distribution of uncertain parameters. Generalizing results to other specifications and more complex settings is usually difficult, and consequently this approach is mostly used to analyze single storage systems. For instance, analytical solutions to optimal storage arbitrage with stochastic price have been derived in [7] without storage ramping constraints, and in [8] with ramping constraints. Problems of using energy storage to minimize energy imbalance are studied in various contexts; see [9, 10] for reducing reserve energy requirements in power system dispatch, [11, 12] for operating storage co-located with a wind farm, [13, 14] for operating storage co-located with end-user demands, and [15] for storage with demand response.

The other category is to use heuristic algorithms, such as Model Predictive Control (MPC)[16] and look-ahead policies [17], to identify sub-optimal storage control rules. Usually based on deterministic (convex) optimization, these approaches can be easily applied to general networks. The major drawback is that these approaches usually do not have any performance guarantee. Consequently, it lacks theoretical justification for implementing them in real systems. Examples of this category can be found in [16] and references therein.

This work aims to bring together the best of both worlds, *i.e.*, to design online deterministic optimizations that solve the stochastic control problem with provable guarantees. It contributes to the existing literature in the following ways. First, we formalize the notion of *generalized storage* as a dynamic model that captures a variety of power system components which provide the functionality of storage. Second, we formulate the problem of storage network operation as a stochastic control problem with general cost functions, and provide examples of applications that can be encapsulated by such a formulation. Third, we devise an online algorithm for the problem based on the theory of Lyapunov optimization [1], and prove guarantees for its performance in terms of a bound of its sub-optimality. We also show that the bound is independent of the realizations of the uncertain parameters. The bound is useful not only in assessing the worst-case performance of our algorithm, but also in evaluating the performance of other sub-optimal algorithms when the optimal costs are hard to obtain. It can also be used to estimate the maximum cost reduction that can be achieved by *any* storage operation, thus provides understanding for the limit of a certain storage system. To the best of our knowledge, this is the first algorithm with provable guarantees for the storage operation problem with general electric networks.

Our methodology is closely related to that of [14], where the focus is on solving the problem of operating an idealized

---

[1] Although closely related to the classical Lyapunov theory for stability, the theory and techniques of Lyapunov optimization are relatively recent. See [18] for more details.

energy storage (with no energy dissipation over time, and no charging/discharging conversion loss) at data-centers. Our objective is to provide an algorithm to operate generalized storage network in a wide range of different settings. This requires an extended or a new analysis in the following aspects. From the modeling perspective, in order to capture applications such as deferrable loads and TCLs, we do not assume storage level is non-negative, instead, we only assume each storage is feasible (see Assumption 2.1 for more details). Furthermore, modeling the dissipation of energy over time leads to a new sub-optimality bound; the bound in [14] becomes a special case of our bound when the dissipation factor (or storage efficiency) is one. A semidefinite program is constructed to decide parameters of the algorithm in order to minimize the sub-optimality bound. Finally, the aspect of power network appears to be completely new.

The rest of the paper is organized as follows. Section 2 formulates the problem of operating a generalized storage network under uncertainty. Section 3 gives the online algorithm and states the performance guarantee. Section 4 analyzes the single bus case in detail with a generalized storage, and Section 5 provides a summary of results for general storage networks. Numerical examples are then given in Section 6. Section 7 concludes the paper.

2. PROBLEM FORMULATION

2.1 Generalized Storage Models

We start by defining a generalized storage model for each fixed bus of the electric network. A diagram is shown in Figure 1. Such a model may be used for a single bus system by setting the network inflow to be zero, or as a component of an electric network as discussed in Section 2.3. We work with a slotted time model, where t is used as the index for an arbitrary time period. Given that the actual length of each time interval is constant, this allows for simple conversion from power units (*e.g.*, MW) to energy units (*e.g.*, MWh) and vice versa. Thus we assume all quantities under consideration in this paper are in energy units, albeit many power system quantities are conventionally specified in power units.
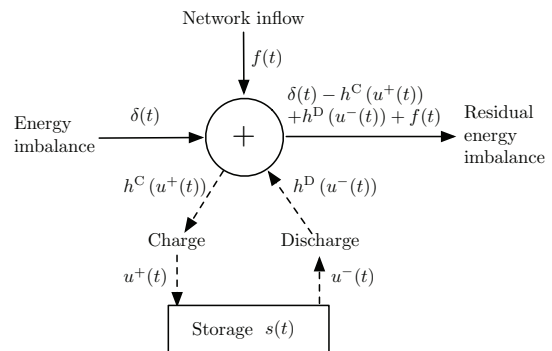

Figure 1: Diagram of a single-bus storage system.

For the bus under consideration and time period t, the local *energy imbalance* $\delta(t)$ is defined to be the difference between the local generation and demand. Both the local generation and demand can be stochastic, and therefore $\delta(t)$ is stochastic in general. The bus may be connected to other

parts of the network, whose net energy inflow is denoted by $f(t)$. The bus is also connected to a *generalized storage*, which is specified by the following elements:

- The *storage level* or State of Charge (SoC) $s(t)$ summarizes the status of the storage at time period t. If $s(t) \geq 0$, it represents the amount of energy in storage; if $s(t) \leq 0$, $-s(t)$ can represent the amount of currently deferred (and not fulfilled) demand. It satisfies $s(t) \in [S^{\min}, S^{\max}]$, where $S^{\max}$ is the storage capacity, and $S^{\min}$ is the minimum allowed storage level.

- The *storage operation* $u(t)$ summarizes the charging (when $u(t) \geq 0$) and discharging (when $u(t) \leq 0$) operations of the storage. It satisfies charging and discharging ramping constraints, *i.e.*, $u(t) \in [U^{\min}, U^{\max}]$, where $U^{\min} \leq 0$ whose magnitude is the maximum discharge within each time period, and $U^{\max} \geq 0$ is the maximum charge within each time period. We also use $u^+(t) = \max(u(t), 0)$ and $u^-(t) = \max(-u(t), 0)$ to denote the charging and discharging operations respectively.

- The *storage conversion function* h maps the storage operation $u(t)$ into its effect on the bus. In particular, it is composed of two linear functions, namely the *charging conversion function* h^C, and the *discharging conversion function* h^D, such that the quantity $h^C(u^+(t))$ is the amount of energy that is withdrawn from the bus due to $u^+(t)$ amount of charge, and $h^D(u^-(t))$ is the amount of energy that is injected into the bus due to $u^-(t)$ amount of discharge, whence

$$h(u(t)) = h^D(u^-(t)) - h^C(u^+(t))$$

is the net energy injection into the bus.

- The *storage dynamics* is then

$$s(t+1) = \lambda s(t) + u(t), \qquad (1)$$

where $\lambda \in (0, 1]$ is the *storage efficiency* which models the loss over time even if there is no storage operation.

The storage parameters satisfy the following consistency conditions.

Assumption 2.1 (Feasibility). *Starting from any feasible storage level, there exists a feasible storage operation such that the storage level in the next time period is feasible, that is*

1. $\lambda S^{\min} + U^{\max} \geq S^{\min}$,

2. $\lambda S^{\max} + U^{\min} \leq S^{\max}$.

The *residual energy imbalance*, after accounting for the network inflow and storage operation, is then given by:

$$\delta(t) + h(u(t)) + f(t) = \delta(t) - h^C(u^+(t)) + h^D(u^-(t)) + f(t). \qquad (2)$$

We give a few examples of generalized storage models as follows.

Example 2.2 (Storage of energy). *Storage of energy can be modeled as a generalized storage with $S^{\max} \geq S^{\min} \geq 0$. Here $U^{\min}$ and $U^{\max}$ correspond to the power rating of the storage, up to a multiple of the length of each time period. By setting $h^C(u^+(t)) = (1/\mu^C)u^+(t)$, and $h^D(u^-(t)) =*

$\mu^D u^-(t)$, *one models the energy loss during charging and discharging operations. Here $\mu^C \in (0, 1]$ is the charging efficiency; $\mu^D \in (0, 1]$ is the discharging efficiency; and the round-trip efficiency of the energy storage is $\mu^C \mu^D$. For instance, based on the information from [19], a NaS (sodium sulfur) battery can be modeled with parameters:*

$$(S^{\min}, S^{\max}, U^{\min}, U^{\max}, \mu^C, \mu^D, \lambda) = (0MWh, 100MWh, -10MW \times 1h, 10MW \times 1h, 0.85, 0.85, 0.97),$$

and a CAES (compressed air energy storage) can be modeled with parameters:

$$(S^{\min}, S^{\max}, U^{\min}, U^{\max}, \mu^C, \mu^D, \lambda) = (0MWh, 3000MWh, -300MW \times 1h, 300MW \times 1h, 0.85, 0.85, 1.00).$$

Example 2.3 (Storage of demand). *Pre-emptive deferrable loads may be modeled as storage of demand, with $-s(t)$ corresponding to the accumulated deferred (but not yet fulfilled) load up to time t, and with $u(t)$ corresponding to the amount of load to defer/fulfill in time period t. We have $S^{\min} \leq S^{\max} \leq 0$ in this case. Storage of demand differs from storage of energy in the sense that it has to be discharged before charging is allowed. The conversion function can usually be set to $h(u(t)) = u(t)$, and generally $\lambda = 1$ in deferrable load related applications.*

Example 2.4 (Generalized battery models). *It is shown recently that an aggregation of TCLs may be modeled as a generalized battery [6]. A discrete time version of such a model can be cast into our framework by setting $S^{\min} = -S^{\max}$ and $S^{\max} \geq 0$. Other storage parameters can be set properly according to Definition 1 of [6], and we have $\lambda \leq 1$ to model energy dissipation.*

We consider the following *stochastic piecewise linear cost function* for each fixed bus

$$g(t) = \sum_{\ell=1}^{L} p(t,\ell) \Big(\alpha^\delta(\ell)\delta(t) - \alpha^C(\ell)h^C(u^+(t)) \qquad (3)$$
$$+ \alpha^D(\ell)h^D(u^-(t)) + \alpha^F(\ell)f(t) + \alpha^{\text{Const}}(t,\ell) \Big)^+,$$

where the parameter $p(t,\ell)$ is in general stochastic, and follows a prescribed probability law, and $\alpha^\delta(\ell)$, $\alpha^C(\ell)$, $\alpha^D(\ell)$, $\alpha^F(\ell)$ and $\alpha^{\text{Const}}(t,\ell)$ are constants, for each $\ell = 1, \ldots, L$ and t. This cost function serves as a generalization of positive (and/or negative) part cost function of the residual energy imbalance, and it encapsulates many applications of storage as shown in Section 2.2. Our analysis applies to a more general class of cost functions; see Appendix A for more details.

2.2 Applications in Single Bus Systems

The storage operation problem on a single bus system ($f(t) = 0$) can be posed as an infinite horizon average cost stochastic control problem as follows:

$$\text{minimize} \quad \lim_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} g(t) \right] \qquad (4a)$$

$$\text{subject to} \quad s(t+1) = \lambda s(t) + u(t), \qquad (4b)$$

$$S^{\min} \leq s(t) \leq S^{\max}, \qquad (4c)$$

$$U^{\min} \leq u(t) \leq U^{\max}, \qquad (4d)$$

where we aim to find a control policy that maps the state $s(t)$ to storage operation $u(t)$, minimizes the expected average cost and satisfies all constraints for each time period t. Here, the initial state $s(1) \in [S^{\min}, S^{\max}]$ is given.

Combining some specific cases of the generalized storage model given in Examples 2.2-2.4 with properly defined cost functions leads to possible problem instances of optimal control of storage under uncertainty. Here we provide examples that are considered in the literature.

Example 2.5 (Balancing). *Storage may be used to minimize residual energy imbalance given some stochastic $\{\delta(t) : t \geq 1\}$ process. Typical cost functions penalize the positive and negative residual energy imbalance differently, and may have different penalties at different time periods . (For example, to model the different consequences of load shedding at different times of the day.) The problem of optimal storage control for such a purpose can be modeled by problem (4) with the cost function*

$$
\begin{aligned}
g(t) =& q^+(t)\left(\delta(t) - h^{\mathrm{C}}(u^+(t)) + h^{\mathrm{D}}(u^-(t))\right)^+ \\
& + q^-(t)\left(\delta(t) - h^{\mathrm{C}}(u^+(t)) + h^{\mathrm{D}}(u^-(t))\right)^- \\
=& q^+(t)\left(\delta(t) - h^{\mathrm{C}}(u^+(t)) + h^{\mathrm{D}}(u^-(t))\right)^+ \\
& + q^-(t)\left(-\delta(t) + h^{\mathrm{C}}(u^+(t)) - h^{\mathrm{D}}(u^-(t))\right)^+ ,
\end{aligned}
$$

where $q^+(t)$ and $q^-(t)$ are the penalties[2] for each unit of positive and negative residual energy imbalance at time period t, respectively.

Example 2.6 (Arbitrage). *Given that the locational marginal prices $\{p^{\mathrm{LMP}}(t) : t \geq 1\}$ are stochastic, a storage may be used to exploit arbitrage opportunities in electricity markets. The problem of maximizing the expected arbitrage profit using storage operations can be cast as an instance of (4), with the cost function (i.e., negative profit) given by:*

$$
\begin{aligned}
g(t) =& p^{\mathrm{LMP}}(t)(h^{\mathrm{C}}(u^+(t)) - h^{\mathrm{D}}(u^-(t))) \\
=& p^{\mathrm{LMP}}(t)\left(h^{\mathrm{C}}(u^+(t)) - h^{\mathrm{D}}(u^-(t))\right)^+ \\
& - p^{\mathrm{LMP}}(t)\left(-h^{\mathrm{C}}(u^+(t)) + h^{\mathrm{D}}(u^-(t))\right)^+ .
\end{aligned}
$$

Example 2.7 (Storage co-located with a stochastic generation or demand). *Applications of this type may be cast into our framework using $\{\delta(t) : t \geq 1\}$ to model the stochastic generation or demand process, and $\{p(t,\ell) : t \geq 1\}$ to model the stochastic prices. A possible cost function is*

$$
g(t) = p^{\mathrm{LMP}}(t)\left(-\delta(t) + h^{\mathrm{C}}(u^+(t)) - h^{\mathrm{D}}(u^-(t))\right)^+ ,
$$

where the residual energy is curtailed with no cost/benefit, and the residual demand is supplied via buying energy from the market at stochastic price $p^{\mathrm{LMP}}(t)$.

2.3 Network Models

The electric network can be modeled as a directed graph $G(V, E)$. Let $n = |V|$, $m = |E|$, and E^{R} be the edge set with all edges reversed. We use the notation $e \sim v$ to indicate that $e \in \{(v', v) \in E \cup E^{\mathrm{R}} : v' \in V\}$. Each edge models

a transmission (or distribution) line, and is associated with some power flow. Assuming the power system is operated in steady state, and the power flow is approximately a constant over each time period t, the energy flow through the line can be obtained by multiplying the power flow by the length of each time period and is denoted by $f_e(t)$ for $e \in E$, with the direction of the edge indicating the positive direction of the flow.[3] The flow vector $\mathbf{f}(t) \in \mathbb{R}^m$ satisfies power flow constraints, which can be compactly summarized by the following set of linear constraints using the classical DC power flow approximations to AC power flow equations [20]:

$$
\mathbf{f}(t) \in \mathcal{F}, \quad \mathcal{F} = \{\mathbf{f} \in \mathbb{R}^m : -\mathbf{F}^{\max} \leq \mathbf{f} \leq \mathbf{F}^{\max}, K\mathbf{f} = 0\}, \tag{5}
$$

where $\mathbf{F}^{\max} \in \mathbb{R}^m$ is the vector of the line capacities of the network, and $K \in \mathbb{R}^{(m-n+1) \times m}$ is a matrix summarizing the Kirchhoff's voltage law. The construction of this K matrix from network topology and line parameters can be found in [21]. Note that additional network constraints may be included in the definition of the set $\mathcal{F}$.

Each node models a bus in the electric network. On bus $v \in V$, a set of variables as described in Section 2.1 is defined, with a subscript v attached to each of the bus variables, and the network inflow is replaced by network flows to the bus from incident lines. The cost for bus v and time period t is then given by

$$
\begin{aligned}
g_v(t) = \sum_{\ell=1}^{L_v} p_v(t,\ell) \Big(& \alpha_v^{\delta}(\ell)\delta_v(t) - \alpha_v^{\mathrm{C}}(\ell)h_v^{\mathrm{C}}(u_v^+(t)) \\
& + \alpha_v^{\mathrm{D}}(\ell)h_v^{\mathrm{D}}(u^-(t)) + \alpha_v^{\mathrm{F}}(\ell)\sum_{e \sim v} f_e(t) + \alpha_v^{\mathrm{Const}}(t,\ell) \Big)^+ ,
\end{aligned} \tag{6}
$$

and the networked storage stochastic control problem is defined as follows:

$$
\begin{aligned}
\text{minimize} \quad & \lim_{T \to \infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^T \sum_{v \in V} g_v(t)\right] & \text{(7a)} \\
\text{subject to} \quad & s_v(t+1) = \lambda_v s_v(t) + u_v(t), & \text{(7b)} \\
& S_v^{\min} \leq s_v(t) \leq S_v^{\max}, & \text{(7c)} \\
& U_v^{\min} \leq u_v(t) \leq U_v^{\max}, & \text{(7d)} \\
& \mathbf{f}(t) \in \mathcal{F}. & \text{(7e)}
\end{aligned}
$$

In this problem, we aim to find a control policy that maps the state $\mathbf{s}(t)$ to storage operation $\mathbf{u}(t)$ and network flow $\mathbf{f}(t)$, and minimizes the expected average cost objective function(7a), such that constraints (7b)-(7d) hold for each t and v, and (7e) holds for each t.

3. ONLINE ALGORITHM AND PERFORMANCE GUARANTEES

This paper provides an online algorithm for solving (7) with provable performance guarantees. Here we give a preview of the algorithm (Algorithm 1) and its sub-optimality bound (Theorem 3.3). The performance theorem will hold under the following additional technical assumptions.

---

[2]These penalties are usually prescribed deterministic sequences [9].

[3]To lighten the notation, for each $e = (v_1, v_2) \in E$, we also define $f_{e'}(t) = -f_e(t)$ for $e' = (v_2, v_1)$, and therefore for all $v \in V$, the net inflow $\sum_{e \sim v} f_e(t) = \sum_{e=(v',v) \in E} f_e(t) + \sum_{e=(v',v) \in E^{\mathrm{R}}} f_e(t) = \sum_{e=(v',v) \in E} f_e(t) - \sum_{e=(v,v') \in E} f_e(t)$.

Algorithm 1 Online Lyapunov Optimization for Storage Network Control

Input: (i) Storage specifications ($S_v^{\min}$, $S_v^{\max}$, $U_v^{\min}$, $U_v^{\max}$, h_v^{C}, h_v^{D}, λ_v), (ii) cost parameters in $g_v(t)$ (including an upper bound and a lower bound on the sub-derivative of $g_v(t)$ with respect to $u_v(t)$, denoted by $\overline{D}g_v(t)$ and $\underline{D}g_v(t)$, and excluding any information about stochastic parameters $\delta_v(t)$ and $p_v(t,\ell)$, $\ell = 1,\ldots,L_v$) for each bus $v \in V$, and (iii) network parameters K and $\mathbf{F}^{\max}$.

Offline-Phase: Determine algorithmic parameters Γ_v, for each bus $v \in V$, and W by solving semidefinite program (24).

Online-Phase:

    for each time period t **do**

        Each bus $v \in V$ observes realizations of stochastic parameters $\delta_v(t)$ and $p_v(t,\ell)$, $\ell = 1,\ldots,L_v$.

        Solve the following deterministic optimization for storage operation $\mathbf{u}(t)$ and network flow $\mathbf{f}(t)$:

$$\text{minimize} \sum_{v \in V} \lambda_v(s_v(t) + \Gamma_v)u_v(t) + Wg_v(t) \qquad (8\text{a})$$

$$\text{subject to } \mathbf{U}^{\min} \le \mathbf{u}(t) \le \mathbf{U}^{\max}, \qquad (8\text{b})$$

$$\mathbf{f}(t) \in \mathcal{F}. \qquad (8\text{c})$$

    end for

Assumption 3.1. *For each bus $v \in V$, the range of storage control is smaller than the effective capacity of the storage, i.e., $U_v^{\max} - U_v^{\min} < S_v^{\max} - S_v^{\min}$.*

Since the bounds for storage control $U_v^{\max}$ and $U_v^{\min}$ are the product of the power rating of storage (in unit MW for example) and the length of each time period, this assumption holds for most systems as long as the length of each time period is made small enough. For instance, this assumption is satisfied for both energy storage examples in Example 2.2.

We make the following assumption on the stochastic parameters of the system.

Assumption 3.2. *Let the collection of stochastic parameters be $\boldsymbol{\theta}(t) = \{\delta_v(t), p_v(t,\ell), \ell = 1,\ldots,L_v, v \in V\}$. Then one of the following two assumptions is in force:*

1. *$\boldsymbol{\theta}(t)$ is independent and identically distributed (i.i.d.) across time t, and is supported on a compact set.*

2. *$\boldsymbol{\theta}(t)$ is some deterministic function of the system stochastic state $\omega(t)$, which is supported on an (arbitrarily large) finite set Ω, and follows an ergodic Markov Chain.*

The first assumption is used to give a simple proof of the performance theorem and provide insights on the construction of our algorithm. The second alternative assumption intends to generalize the performance bounds to non-i.i.d. cases. The additional assumptions such as that $\omega(t)$ lies in a finite set are introduced to reduce the required technicality in view of the page limit. The same result can be obtained in a more general setting where $\{\boldsymbol{\theta}(t) : t \ge 1\}$ follows a renewal process. The performance bounds in these setting are the same, up to a multiple of the mean recurrence time of the stochastic process under consideration.

Under Assumptions 3.1 and 3.2, the following performance guarantee holds.

Theorem 3.3. *The control actions $(\mathbf{u}(t), \mathbf{f}(t))$ generated by Algorithm 1 are feasible for (7) and sub-optimal, whose sub-optimality[4] is bounded by a constant that depends only on the parameters of the storages and cost functions, and is independent of realizations of the stochastic parameters.*

The precise expressions of the sub-optimality bounds for the single bus case and general network case are given in Section 4 and Section 5, respectively, both under the i.i.d. assumption (Assumption 3.2.1). The bounds for settings with the Markov assumption (Assumption 3.2.2) are given in [21].

Remark 3.4 (Convexity). *Our result holds without assuming that $g_v(t)$ is convex in $u_v(t)$, $v \in V$. However, we do assume the online optimization (8) can be solved efficiently, and in all numerical examples we work with convex cost functions.*

4. ANALYSIS FOR SINGLE BUS SYSTEMS

To demonstrate the proof ideas without unfolding all technicalities, we prove Theorem 3.3 for a single bus system under the following simplifying assumptions.

Assumption 4.1. *We assume in this section:*

- *the imbalance process $\{\delta(t) : t \ge 1\}$ is independent and identically distributed (i.i.d.) across t and is supported on a compact interval $[\delta^{\min}, \delta^{\max}]$;*

- *for each $\ell = 1,\ldots,L$, the process $\{p(t,\ell) : t \ge 1\}$ is i.i.d. across t and is supported on a compact interval $[p^{\min}(\ell), p^{\max}(\ell)]$.*

Define

$$\bar{u} \triangleq \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} u(t)\right], \quad \bar{s} \triangleq \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} s(t)\right].$$

Note that for $s(1) \in [S^{\min}, S^{\max}]$,

$$\bar{u} = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} s(t+1) - \lambda s(t)\right] = (1-\lambda)\bar{s}.$$

As $s(t) \in [S^{\min}, S^{\max}]$ for all $t \ge 0$, the above expression implies

$$(1-\lambda)S^{\min} \le \bar{u} \le (1-\lambda)S^{\max}.$$

Problem (4) can be equivalently written as follows

$$\textbf{P1: } \text{minimize} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} g(t)\right] \qquad (9\text{a})$$

$$\text{subject to } s(t+1) = \lambda s(t) + u(t), \qquad (9\text{b})$$

$$S^{\min} - \lambda s(t) \le u(t) \le S^{\max} - \lambda s(t), \quad (9\text{c})$$

$$U^{\min} \le u(t) \le U^{\max}, \qquad (9\text{d})$$

$$(1-\lambda)S^{\min} \le \bar{u} \le (1-\lambda)S^{\max}, \qquad (9\text{e})$$

where bounds on $s(t)$ are replaced by (9c), and (9e) is added without loss of optimality.

The proof procedure is depicted in the diagram shown in Figure 2, where we use $J_{\mathrm{P1}}(u)$ to denote the objective value

---

[4]Here the sub-optimality is defined as the difference between the objective value of (7) with $(\mathbf{u}(t), \mathbf{f}(t))$ generated by Algorithm 1 and the optimal cost of (7).

of **P1** with storage operation sequence u (as an abbreviation of $\{u(t) : t \geq 1\}$), $u^\star(\mathbf{P1})$ to denote an optimal sequence of storage operation for **P1**, $J_{\text{P1}}^\star \triangleq J_{\text{P1}}(u^\star(\mathbf{P1}))$, and we define similar quantities for **P2** and **P3**. Here **P2** is an auxilliary problem we construct to bridge the infinite horizon storage control problem **P1** to online Lyapunov optimization problems **P3** (8) (or (15) for single storage case). It has the following form

$$\mathbf{P2:} \quad \text{minimize} \quad \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} g(t)\right] \qquad (10\text{a})$$

$$\text{subject to } U^{\min} \leq u(t) \leq U^{\max}, \qquad (10\text{b})$$

$$(1-\lambda)S^{\min} \leq \bar{u} \leq (1-\lambda)S^{\max}. \qquad (10\text{c})$$

Notice that it has the same objective as **P1**, and evidently it is a relaxation of **P1**. This implies that $u^\star(\mathbf{P2})$ may not be feasible for **P1**, and

$$J_{\text{P2}}^\star = J_{\text{P1}}(u^\star(\mathbf{P2})) \leq J_{\text{P1}}^\star. \qquad (11)$$

The reason for the removal of state-dependent constraints (9c) (and hence (9b) as the sequence $\{s(t) : t \geq 1\}$ becomes irrelevant to the optimization of $\{u(t) : t \geq 1\}$) in **P2** is that the state-independent problem **P2** has easy-to-characterize optimal stationary control policies. In particular, from the theory of stochastic network optimization [18], the following result holds.

Lemma 4.2 (Optimality of Stationary Disturbance-Only Policies). *Under Assumption 4.1 there exists a stationary disturbance-only[5] policy $u^{\text{stat}}(t)$, satisfying (10b) and (10c), and providing the following guarantees for all t:*

$$(1-\lambda)S^{\min} \leq \mathbb{E}[u^{\text{stat}}(t)] \leq (1-\lambda)S^{\max}, \qquad (12)$$

$$\mathbb{E}[g(t)|u(t) = u^{\text{stat}}(t)] = J_{\text{P2}}^\star, \qquad (13)$$

where the expectation is taken over the randomization of $\delta(t)$, $p(t,\ell)$, $\ell = 1, \ldots, L$, and (possibly) $u^{\text{stat}}(t)$.
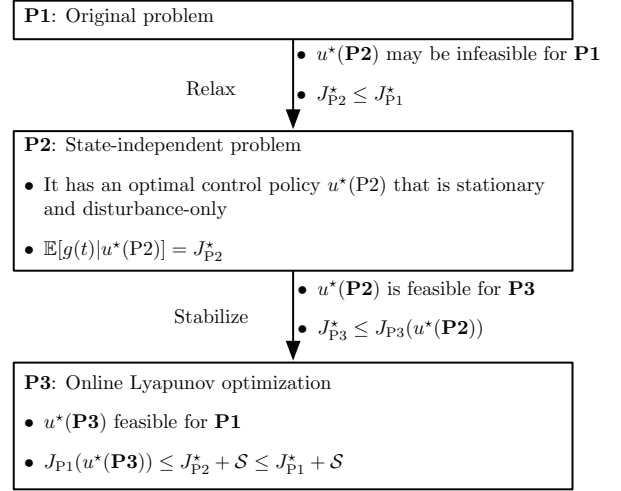
Equation (13) not only assures the storage operation induced by the stationary disturbance-only policy achieves the optimal cost, but also guarantees that the expected stage-wise cost is a constant across time t and equal to the optimal time average cost. This fact will later be exploited in order to establish the performance guarantee of our online algorithm. By the merits of this Lemma, in the sequel, we overload $u^\star(\mathbf{P2})$ to denote the storage operation sequence obtained from an optimal stationary disturbance-only policy.

An issue with $u^\star(\mathbf{P2})$ for the original problem is that it may not be feasible for **P1**. To have the $\{s(t) : t \geq 1\}$ sequence induced by the storage operation sequence lie in the interval $[S^{\min}, S^{\max}]$, we construct a virtual queue related to $s(t)$ and use techniques from Lyapunov optimization to "stabilize" such a queue. Let the queueing state be a shifted version of the storage level:

$$\tilde{s}(t) = s(t) + \Gamma, \qquad (14)$$

where the shift constant Γ will be specified later. We wish to minimize the stage-wise cost $g(t)$ and at the same time to maintain the queueing state close to zero. This motivates us to consider solving the following optimization online (*i.e.*,

---

[5]The policy is a pure function (possibly randomized) of the current disturbances $\delta(t)$ and $p(t,\ell)$, $\ell = 1, \ldots, L$.

P1: Original problem
- $u^\star(\mathbf{P2})$ may be infeasible for **P1**
- $J_{\text{P2}}^\star \leq J_{\text{P1}}^\star$

Relax

P2: State-independent problem
- It has an optimal control policy $u^\star(\text{P2})$ that is stationary and disturbance-only
- $\mathbb{E}[g(t)|u^\star(\text{P2})] = J_{\text{P2}}^\star$

Stabilize
- $u^\star(\mathbf{P2})$ is feasible for **P3**
- $J_{\text{P3}}^\star \leq J_{\text{P3}}(u^\star(\mathbf{P2}))$

P3: Online Lyapunov optimization
- $u^\star(\mathbf{P3})$ feasible for **P1**
- $J_{\text{P1}}(u^\star(\mathbf{P3})) \leq J_{\text{P2}}^\star + \mathcal{S} \leq J_{\text{P1}}^\star + \mathcal{S}$

Figure 2: An illustration of the proof procedure as relations between problems considered in Section 4. Here $\mathcal{S}$ denotes the sub-optimality bound.

at the beginning of each time period t after the realizations of stochastic parameters $p(t,\ell)$, $\ell = 1, \ldots, L$, and $\delta(t)$ have been observed)

$$\mathbf{P3:} \quad \text{minimize } \lambda\tilde{s}(t)u(t) + Wg(t) \qquad (15\text{a})$$

$$\text{subject to } U^{\min} \leq u(t) \leq U^{\max}, \qquad (15\text{b})$$

where the optimization variable is $u(t)$, the stochastic parameters in $g(t)$ are replaced with their observed realizations, and $W > 0$ is a weight parameter. Note that the objective here is a weighted combination of the stage-wise cost and a linear term of $u(t)$, whose coefficient is positive when $s(t)$ is large, and negative when $s(t)$ is small. We use the notation $u^{\text{ol}}(t)$ for the solution to **P3** at time period t, $u^\star(\mathbf{P3})$ for the sequence $\{u^{\text{ol}}(t) : t \geq 1\}$, $J_{\text{P3},t}(u(t))$ for the objective function of **P3** at time period t, and $J_{\text{P3},t}^\star$ for the corresponding optimal cost. In the rest of this section, we give conditions for parameters Γ and W such that solving **P3** online will result in a feasible $\{s(t) : t \geq 1\}$ sequence (Section 4.1), characterize the sub-optimality of $u^\star(\mathbf{P3})$ as a function of Γ and W and state the semidefinite program for identifying the optimal Γ and W pair (Section 4.2).

4.1 Feasibility

We start with a structural result for the online optimization problem **P3**. It follows from Lemma A.1 which is proved for general cost functions in Appendix A.

Lemma 4.3. *At each time period t, the solution to **P3**, $u^{\text{ol}}(t)$, satisfies*

1. $u^{\text{ol}}(t) = U^{\min}$ whenever $\lambda\tilde{s}(t) \geq -W\underline{D}g$,

2. $u^{\text{ol}}(t) = U^{\max}$ whenever $\lambda\tilde{s}(t) \leq -W\overline{D}g$,

where

$$\underline{D}g \triangleq \inf \left\{ \xi \in \partial_{u(t)}g(t) \middle| \begin{array}{l} u(t) \in [U^{\min}, U^{\max}], \\ p(t,\ell) \in [p^{\min}(\ell), p^{\max}(\ell)], \forall\ell, \\ \delta(t) \in [\delta^{\min}, \delta^{\max}], \\ \mathbf{f}(t) \in \mathcal{F}, \ t \geq 1 \end{array} \right\},$$

and

$$\overline{D}g \triangleq \sup \left\{ \xi \in \partial_{u(t)}g(t) \left| \begin{array}{l} u(t) \in [U^{\min}, U^{\max}], \\ p(t,\ell) \in [p^{\min}(\ell), p^{\max}(\ell)], \forall \ell, \\ \delta(t) \in [\delta^{\min}, \delta^{\max}], \\ \mathbf{f}(t) \in \mathcal{F}, \ t \geq 1 \end{array} \right. \right\}$$

are the greatest lower bound and the least upper bound of the sub-derivatives of $g(t)$, respectively.

Remark 4.4 (Evaluation of $\overline{D}g$, $\underline{D}g$)**.** *Any finite lower bound and upper bound for the sub-derivative of the cost $g(t)$ can be used as $\underline{D}g$ and $\overline{D}g$, respectively. Here we use the greatest lower bound and least upper bound to provide the tightest performance bounds. For cases with simple cost functions, e.g., for idealized storage with $L = 1$, $\underline{D}g$ and $\overline{D}g$ can be easily obtained from $p^{\min}(\ell)$, $p^{\max}(\ell)$, and constants in the cost function $g(t)$ (such as $\alpha^{\mathrm{C}}(\ell)$ and $\alpha^{\mathrm{D}}(\ell))$. In cases where $g(t)$ is differentiable with respect to $u(t)$, $\overline{D}g$ and $\underline{D}g$ may be obtained by solving a simple optimization problem.*

This allows us to construct the following sufficient condition that will assure the feasibility of the $\{s(t) : t \geq 1\}$ sequence induced by $u^\star(\mathbf{P3})$.

Theorem 4.5 (Feasibility)**.** *Suppose the initial storage level satisfies $s(1) \in [S^{\min}, S^{\max}]$, then the storage level sequence $\{s(t) : t \geq 1\}$ induced by the sequence of storage operation $u^\star(\mathbf{P3})$ is feasible with respect to storage level constraints, i.e., $s(t) \in [S^{\min}, S^{\max}]$ for all t, provided that*

$$\Gamma^{\min} \leq \Gamma \leq \Gamma^{\max}, \tag{16}$$

$$0 < W \leq W^{\max}, \tag{17}$$

where

$$\Gamma^{\min} = \frac{1}{\lambda}\left[-W\underline{D}g + (U^{\max} - (1-\lambda)S^{\max})^+\right] - S^{\max}, \tag{18}$$

$$\Gamma^{\max} = \frac{1}{\lambda}\left[-W\overline{D}g - \left((1-\lambda)S^{\min} - U^{\min}\right)^+\right] - S^{\min}, \tag{19}$$

and

$$W^{\max} = \frac{1}{\overline{D}g - \underline{D}g}\left[\lambda(S^{\max} - S^{\min}) - \left((1-\lambda)S^{\min} - U^{\min}\right)^+\right.$$
$$\left. - (U^{\max} - (1-\lambda)S^{\max})^+\right]. \tag{20}$$

Proof. The result is proved by induction, where Lemma 4.3 is used to partially characterize the $u^{\mathrm{ol}}(t)$ sequence. See Appendix B for more details. $\qquad\square$

4.2 Performance

In the previous result, we have established that $u^\star(\mathbf{P3})$ is feasible for **P1** as long as parameters Γ and W satisfy (16) and (17). In the next theorem, we characterize the sub-optimality of $u^\star(\mathbf{P3})$ for fixed Γ and W.

Theorem 4.6 (Performance)**.** *The sub-optimality of storage operation $u^\star(\mathbf{P3})$ is bounded by $M(\Gamma)/W$, that is*

$$J_{\mathrm{P1}}^\star \leq J_{\mathrm{P1}}(u^\star(\mathbf{P3})) \leq J_{\mathrm{P1}}^\star + M(\Gamma)/W, \tag{21}$$

where

$$M(\Gamma) = M^u(\Gamma) + \lambda(1-\lambda)M^s(\Gamma),$$

$$M^u(\Gamma) = \frac{1}{2}\max\left(\left(U^{\min} + (1-\lambda)\Gamma\right)^2, (U^{\max} + (1-\lambda)\Gamma)^2\right),$$

$$M^s(\Gamma) = \max\left(\left(S^{\min} + \Gamma\right)^2, (S^{\max} + \Gamma)^2\right).$$

Proof. A quadratic Lyapunov function is constructed. The relation between the Lyapunov drift and the objective of **P3** is exploited, which in turn relates to the objective of **P2** and so **P1**. Appendix B contains the whole proof. $\qquad\square$

The theorem above guarantees that the worst-case cost (among different uncertainty realizations) of our online algorithm is bounded above by $J_{\mathrm{P1}}^\star + M(\Gamma)/W$. The sub-optimality bound $M(\Gamma)/W$ reduces to a much simpler form if $\lambda = 1$.

Remark 4.7 (Sub-Optimality Bound, $\lambda = 1$)**.** *For a storage with $\lambda = 1$, we have*

$$M \triangleq M(\Gamma) = (1/2)\max((U^{\min})^2, (U^{\max})^2),$$

and the online algorithm is no worse than M/W sub-optimal. In this case, one would optimize the performance by setting

$$W = W^{\max} = \frac{(S^{\max} - S^{\min}) - (U^{\max} - U^{\min})}{\overline{D}g - \underline{D}g},$$

and the corresponding interval $[\Gamma^{\min}, \Gamma^{\max}]$ turns out to be a singleton, where

$$\Gamma^{\min} = \Gamma^{\max} = -\frac{\overline{D}g(S^{\max} - U^{\max}) + \underline{D}g(U^{\min} - S^{\min})}{\overline{D}g - \underline{D}g}.$$

Let $S^{\max} - S^{\min} = \rho(U^{\max} - U^{\min})$. Suppose $|U^{\max}| = |U^{\min}|$. For efficient storage ($\lambda = 1$), the sub-optimality bound is

$$\frac{M}{W} = \frac{(1/2)(\overline{D}g - \underline{D}g)(U^{\max})^2}{(S^{\max} - S^{\min}) - (U^{\max} - U^{\min})} = \frac{\overline{D}g - \underline{D}g}{4(\rho - 1)}U^{\max}.$$

For fixed $U^{\max}$, as storage capacity increases, i.e., $\rho \to \infty$, the sub-optimality $(M/W) \to 0$. If $U^{\max}$ and $S^{\max}$ increases with their ratio ρ fixed, the bound increases linearly with $U^{\max}$.

The remaining case $\lambda \in (0,1)$ requires solving an optimization program to identify the bound-minimizing parameter pair (Γ, W). In the next result, we state a semidefinite program to find $(\Gamma^\star, W^\star)$ that solves the following parameter optimization program

P3-PO: minimize $\quad M(\Gamma)/W$

$\qquad$ subject to $\Gamma^{\min} \leq \Gamma \leq \Gamma^{\max}, \ 0 < W \leq W^{\max}$.

In the current form, this program appears to be non-convex. The next result reformulates **P3-PO** into a semidefinite program. Note that $\Gamma^{\min}$ and $\Gamma^{\max}$ are linear functions of W as defined in (18) and (19).

Lemma 4.8 (Semidefinite Reformulation of **P3-PO**)**.** *Let symmetric positive definite matrices $X^{\min,u}$, $X^{\max,u}$, $X^{\min,s}$, and $X^{\max,s}$ be defined as follows*

$$X^{(\cdot),u} = \begin{bmatrix} \eta^u & U^{(\cdot)} + (1-\lambda)\Gamma \\ * & 2W \end{bmatrix}, \ X^{(\cdot),s} = \begin{bmatrix} \eta^s & S^{(\cdot)} + \Gamma \\ * & W \end{bmatrix},$$

*where $(\cdot)$ can be either $\max$ or $\min$, and η^u and η^s are auxiliary variables. Then **P3-PO** can be solved via the following semidefinite program*

$$\text{minimize} \quad \eta^u + \lambda(1-\lambda)\eta^s \tag{23a}$$

$$\text{subject to } \Gamma^{\min} \leq \Gamma \leq \Gamma^{\max}, \ 0 < W \leq W^{\max}, \tag{23b}$$

$$X^{\min,u}, X^{\max,u}, X^{\min,s}, X^{\max,s} \succeq 0. \tag{23c}$$

Proof. The result follows from Schur complement. See Appendix B for details. □

We close this section by discussing several implications of the performance theorem.

Remark 4.9 (Optimality at the Fast-Acting Limit). *Let the length of each time period be Δt. At the limit $\Delta t \to 0$, the online algorithm is optimal. Indeed, as discussed in Section 2, both $|U^{\min}|$ and $|U^{\max}|$ are linear in Δt, such that $|U^{\max}| \to 0$ and $|U^{\min}| \to 0$ as $\Delta t \to 0$. Meanwhile, $\lambda \to 1$ as $\Delta t \to 0$. So by Remark 4.7, it is easy to verify that the sub-optimality M/W converges to zero as $\Delta t \to 0$.*

Remark 4.10 (Operational Value of Storage and Percentage Cost Savings). *Operational Value of Storage (VoS) is broadly defined as the savings in the long term system cost due to storage operation. Such an index is usually calculated by assuming storage is operated optimally. In stochastic environments, the optimal system cost with storage operation is hard to obtain in general settings. In our notations, let u^{NS} denote the sequence $\{u(t) : u(t) = 0, t \geq 1\}$ which corresponds to no storage operation. Then*

$$\mathrm{VoS} = J_{\mathrm{P1}}(u^{\mathrm{NS}}) - J_{\mathrm{P1}}^{\star},$$

and it can be estimated by the interval

$$\left[J_{\mathrm{P1}}(u^{\mathrm{NS}}) - J_{\mathrm{P1}}(u^{\star}(\mathbf{P3})),\ J_{\mathrm{P1}}(u^{\mathrm{NS}}) - J_{\mathrm{P1}}(u^{\star}(\mathbf{P3})) + \frac{M}{W} \right].$$

Additionally, for a storage operation sequence u, the percentage cost savings due to storage can then be defined by $(J_{\mathrm{P1}}(u^{\mathrm{NS}}) - J_{\mathrm{P1}}(u))/J_{\mathrm{P1}}(u^{\mathrm{NS}})$. An upper bound of this for any storage control policy can be obtained via $(J_{\mathrm{P1}}(u^{\mathrm{NS}}) - J_{\mathrm{P1}}(u^{\star}(\mathbf{P3})) + M/W)/J_{\mathrm{P1}}(u^{\mathrm{NS}})$, which to an extent summarizes the limit of a storage system in providing cost reduction.

5. RESULTS FOR NETWORKED SYSTEMS

For completeness, we provide a summary of results for networks with more than one buses and storages; detailed discussions and proofs can be found in [21]. For each bus $v \in V$, we use X_v to denote the corresponding variable X defined for single bus systems in Section 4, and $\mathbf{X}$ to denote the vector $\{X_v\}_{v \in V}$. The online optimization, which corresponds to **P3** in the single bus systems, is as stated in (8). Suppose that Assumption 4.1 is in force for each bus $v \in V$. The following results hold for the control sequence $(\mathbf{u}^{\star}(\mathbf{P3}), \mathbf{f}^{\star}(\mathbf{P3}))$ obtained from Algorithm 1.

Theorem 5.1 (Feasibility). *Suppose the initial storage level satisfies $s_v(1) \in [S_v^{\min}, S_v^{\max}]$, for all $v \in V$, then the storage level sequence $\{\mathbf{s}(t) : t \geq 1\}$ induced by the sequence of storage operation $(\mathbf{u}^{\star}(\mathbf{P3}))$ is feasible with respect to storage level constraints, i.e., $s_v(t) \in [S_v^{\min}, S_v^{\max}]$ for all t and $v \in V$, provided that*

$$\Gamma_v^{\min} \leq \Gamma_v \leq \Gamma_v^{\max},\ 0 < W \leq W_v^{\max},$$

for all $v \in V$.

Theorem 5.2 (Performance). *The sub-optimality of control sequence $(\mathbf{u}^{\star}(\mathbf{P3}), \mathbf{f}^{\star}(\mathbf{P3}))$ is bounded by $M(\mathbf{\Gamma})/W$, that is*

$$J_{\mathrm{P1}}^{\star} \leq J_{\mathrm{P1}}(\mathbf{u}^{\star}(\mathbf{P3}), \mathbf{f}^{\star}(\mathbf{P3})) \leq J_{\mathrm{P1}}^{\star} + \frac{M(\mathbf{\Gamma})}{W},$$

where

$$M(\mathbf{\Gamma}) = M^u(\mathbf{\Gamma}) + M^s(\mathbf{\Gamma}),$$

$$M^u(\mathbf{\Gamma}) = \sum_{v \in V} M_v^u(\Gamma_v),$$

and

$$M^s(\mathbf{\Gamma}) = \sum_{v \in V} \lambda_v (1 - \lambda_v) M_v^s(\Gamma_v).$$

The semidefinite program for minimizing the bound is as follows.

Lemma 5.3 (Semidefinite Optimization of $M(\mathbf{\Gamma})/W$). *Let symmetric positive definite matrices $X_v^{\min,u}$, $X_v^{\max,u}$, $X_v^{\min,s}$, and $X_v^{\max,s}$, $v \in V$, be defined as follows*

$$X_v^{(\cdot),u} = \begin{bmatrix} \eta_v^u & U_v^{(\cdot)} + (1 - \lambda_v)\Gamma_v \\ * & 2W \end{bmatrix}, \quad X_v^{(\cdot),s} = \begin{bmatrix} \eta_v^s & S_v^{(\cdot)} + \Gamma_v \\ * & W \end{bmatrix},$$

where $(\cdot)$ can be either $\max$ or $\min$, and η_v^u and η_v^s are auxiliary variables, for any $v \in V$. Then the sub-optimality bound $M(\mathbf{\Gamma})/W$ can be optimized by solving the following semidefinite program

$$\text{minimize} \quad \sum_{v \in V} \eta_v^u + \lambda_v(1 - \lambda_v)\eta_v^s \tag{24a}$$

$$\text{subject to} \quad \Gamma_v^{\min} \leq \Gamma_v \leq \Gamma_v^{\max}, \tag{24b}$$

$$0 < W \leq W_v^{\max}, \tag{24c}$$

$$X_v^{\min,u}, X_v^{\max,u}, X_v^{\min,s}, X_v^{\max,s} \succeq 0, \tag{24d}$$

where constraints (24b)-(24d) hold for all $v \in V$.

6. NUMERICAL EXPERIMENTS

6.1 Single Storage Example

We first test our algorithm in a simple setting where the analytical solution for the optimal control policy is available, so that the algorithm performance can be compared against the true optimal costs. We consider the problem of using a single energy storage to minimize the energy imbalance as studied in [9], where it is shown that greedy storage operation is optimal if $\lambda = 1$ and if the following cost is considered

$$g(t) = |\delta(t) - (1/\mu^{\mathrm{C}})u^+(t) + \mu^{\mathrm{D}}u^-(t)|.$$

As in [9], we specify storage parameters in per unit, and $S^{\min} = 0$. Let $\mu^{\mathrm{C}} = \mu^{\mathrm{D}} = 1$ so that the parameterization of storage operation here is equivalent to that of [9]. We assume each time period represents an hour, and $-U^{\min} = U^{\max} = (1/10)S^{\max}$. In order to evaluate the performance, we simulate the $\delta(t)$ process by drawing i.i.d. samples from zero-mean Laplace distribution, with standard deviation $\sigma_\delta = 0.149$ per unit obtained from NREL data [9]. The time horizon for the simulation is chosen to be $T = 1000$. Figure 3(a) depicts the performance of the our algorithm and the optimal cost $J_{\mathrm{P1}}^{\star}$ obtained from the greedy policy, where it is shown that the algorithm performance is near-optimal, and better than what the (worst-case) sub-optimality bound predicts. [6]

---

[6] By an abuse of notation, in this section, we use $J_{\mathrm{P1}}^{\star}$ and J_{P1} to denote the results from simulation, which are estimates of the true expectations.

A slight modification of the cost function would render a problem which does not have an analytical solution. Consider the setting where only unsatisfied demand is penalized with a higher penalty during the day (7 am to 7 pm):

$$g(t) = \begin{cases} 3\left(\delta(t) - (u^+(t)/\mu^{\mathrm{C}}) + \mu^{\mathrm{D}}u^-(t)\right)^-, & t \in \mathcal{T}^{\mathrm{Day}}, \\ \left(\delta(t) - (u^+(t)/\mu^{\mathrm{C}}) + \mu^{\mathrm{D}}u^-(t)\right)^-, & \text{otherwise,} \end{cases} \quad (25)$$

where $\mathcal{T}^{\mathrm{Day}} = \{t \geq 1 : 7 \leq t \bmod 24 < 19\}$. We run the same set of tests above, with the modification that now $\mu^{\mathrm{C}} = \mu^{\mathrm{D}} = 0.95$. Note that the greedy policy is only a sub-optimal heuristic for this case. Figure 3(b) shows our online algorithm performs significantly better than the greedy algorithm. The costs of our algorithm together with the lower bounds give a narrow envelope for the optimal average cost $J_{\mathrm{P1}}^\star$ in this setting, which can be used to evaluate the performance of other sub-optimal algorithms numerically. In both experiments, we also plot the costs of predictive/nominal storage control, whose solution can be shown to be $u(t) = 0$ for all t. Consequently, the costs of such operation rule are the same as the system costs when there is no storage.
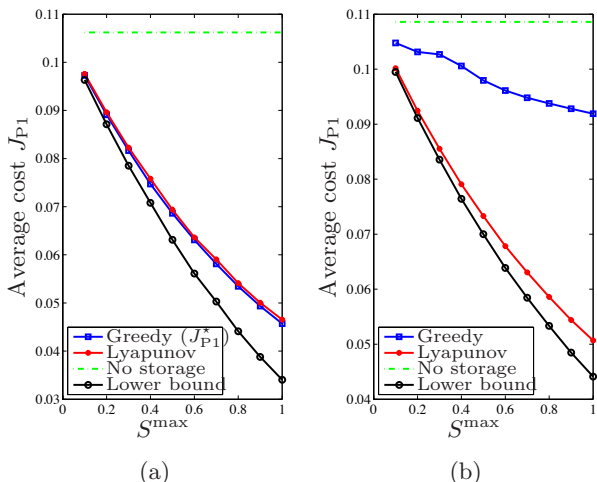
Figure 3: Performance of algorithms in a single bus network. Note that, different from the setup in Remark 4.7, we scale $U^{\min}$, $U^{\max}$ together with $S^{\max}$ in this and the following numerical examples.

Figure 4 translates the results in Figure 3 into percentage cost savings due to storage operation, which are computed following the discussion in Remark 4.10. Using the cost of our online algorithm and the theoretical sub-optimality bound, we obtain an upper bound of percentage cost reduction of energy storage for *any control policy* (see black curve in each panel of Figure 4). It indicates the systemic limit of using storage to provide cost reduction, and is useful for system design considerations especially when the optimal cost cannot be calculated efficiently.

6.2 Storage Network Example

We consider a setting similar to that in the single storage numerical example, in which now distributed storages are coordinated to minimize the power imbalance over a tree network with N buses. We assume the storage network is homogeneous, *i.e.*, the storage installed on each bus of the
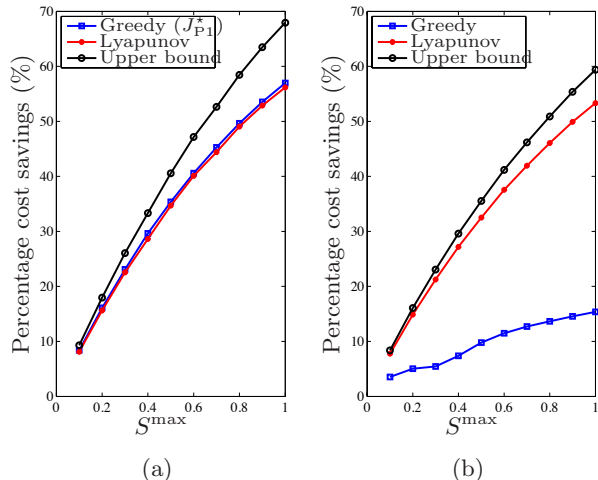
Figure 4: Percentage cost savings of a single storage operated for balancing.

network has the same specifications and the same cost functions. Two cases with different cost functions are considered. In the first case, time homogeneous costs of the form

$$g_v^{\mathrm{H}}(t) = \left(\delta_v(t) - (1/\mu_v^{\mathrm{C}})u_v^+(t) + \mu_v^{\mathrm{D}}u_v^-(t) + \sum_{e \sim v} f_e(t)\right)^-, \quad (26)$$

are considered, where $\mu_v^{\mathrm{C}} = \mu_v^{\mathrm{D}} = 0.95$, and $\delta_v(t)$ is i.i.d. following the same distribution as in the single storage example. In the second case, each bus has a cost function similar to (25):

$$g_v(t) = \begin{cases} 3g_v^{\mathrm{H}}(t), & t \in \mathcal{T}^{\mathrm{Day}}, \\ g_v^{\mathrm{H}}(t), & \text{otherwise,} \end{cases}$$

with $g_v^{\mathrm{H}}(t)$ as defined in (26). We consider non-idealized storages which are operated frequently such that $\lambda_v = 0.999$ for all $v \in V$. As in the single storage example, we fix $-U_v^{\min} = U_v^{\max} = (1/10)S_v^{\max}$. The storages are connected in a star network, with $N = 5$ and $F_e^{\max} = \sigma_\delta$ for each line $e \in E$. The time horizon for the simulation is chosen to be $T = 1000$. Figure 5 shows the percentage cost savings, where it is demonstrated that the online algorithm performs consistently superior to the greedy heuristic, and leads to percentage cost savings values that are close to the derived upper bound. Therefore near-optimal performance is achieved by our algorithm in both cases.

7. CONCLUSIONS AND FUTURE WORK

This work is motivated by the fundamental question of how to *optimally* shift energy over space and time to achieve uncertainty reduction and to facilitate renewable integration. To this end, we consider the problem of optimal control of generalized storage networks under uncertainty. The notion of generalized storage is proposed as a dynamic model to capture many forms of storage conveniently. An online control strategy is then proposed to solve the corresponding constrained stochastic control problem, whose performance is analyzed in detail. We prove that the algorithm is near optimal, and its sub-optimality can be bounded by a constant that only depends on the parameters of the storage network
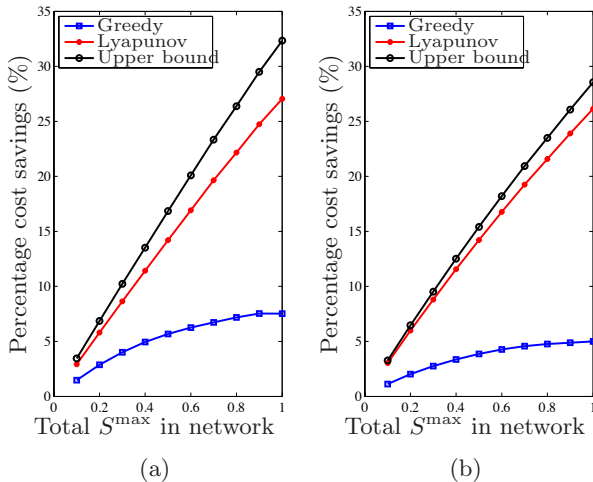
Figure 5: Percentage cost savings of a storage network operated for balancing.

and cost functions, and is independent of the realizations of uncertainties.

Although we have provided analysis for a relatively general setting (see [21] for more details), potential improvements can be achieved in many directions. (i) Our formulation starts by minimizing the long run *expected average* cost, and lands on an online algorithm that has *robust* performance guarantees in the form of a sub-optimality bound that holds for all uncertainty realizations. Relaxing such requirements may result in approaches that trade *risk* with *performance*. Better performance guarantees (in terms of smaller sub-optimality) may hold with large probability (instead of with probability one), which leads to, in a sense, probably approximately correct (PAC) algorithms[22]. (ii) While our online control solves deterministic optimization respecting network constraints, the sub-optimality bound is derived independent of network topology and network parameters such as line capacities. Utilizing such information may lead to a tighter performance bound or a more informed choice of algorithmic parameters. (iii) It can be an advantage or a disadvantage that our online algorithm does not use any statistical information about the uncertain parameters, depending on whether such information is readily available. Observe that our approach in fact can be generalized immediately to settings with additional *same-stage variables* which do not affect the (temporal) states of the system. Incorporating statistical information and forecast updates may improve the performance of the algorithm, and make the algorithm applicable to other settings where *lookahead variables* (such as wind farm contract level for the next stage) are considered together with storage operation. (iv) While the focus of this paper is on energy networks, the algorithm may be applied to other networks since our analysis does not rely on properties of the given constraints on network flow. This also implies that a more accurate AC power model can be used in this study as long as the online optimization can be solved efficiently. Recent advances in tight convex relaxation of AC optimal power flow [23] can be utilized for such purposes. (v) This paper provides a procedure to convert the hard stochastic problem to a sequence of easy deterministic problems which fit into

today's grid operational paradigms (especially for transmission grids operated by centralized system operators). For the future, the integration of distributed energy resources would require a decentralized solution to these problems. We note that many methods have been developed for distributed/decentralized *deterministic* optimization (*cf.* [24]); incorporating these methods for solving the online problems is an important future direction.

Acknowledgments

8. REFERENCES

[1] National Renewable Energy Laboratory. Western Wind and Solar Integration Study, 2010.

[2] National Renewable Energy Laboratory. The Role of Energy Storage with Renewable Electricity Generation, 2010.

[3] D. Lindley. Smart Grids: The Energy Storage Problem. *Nature*, 463(7277):18, 2010.

[4] B. Daryanian and R. E. Bohn. Sizing of Electric Thermal Storage under Real Time Pricing. *IEEE Transactions on Power Systems*, 8(1):35–43, 1993.

[5] G. O'Brien and R. Rajagopal. A Method for Automatically Scheduling Notified Deferrable Loads. In *Proc. of American Control Conference*, pages 5080–5085, 2013.

[6] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent. Aggregate Flexibility of Thermostatically Controlled Loads. *IEEE Transactions on Power Systems*, submitted.

[7] J. Qin, R. Sevlian, D. Varodayan, and R. Rajagopal. Optimal Electric Energy Storage Operation. In *Proc. of IEEE Power and Energy Society General Meeting*, pages 1–6, 2012.

[8] A. Faghih, M. Roozbehani, and M. A. Dahleh. On the Economic Value and Price-Responsiveness of Ramp-Constrained Storage. *ArXiv e-prints*, 2012.

[9] H. I. Su and A. El Gamal. Modeling and Analysis of the Role of Energy Storage for Renewable Integration: Power Balancing. *IEEE Transactions on Power Systems*, 28(4):4109–4117, 2013.

[10] J. Qin, H. I. Su, and R. Rajagopal. Storage in Risk Limiting Dispatch: Control and Approximation. In *Proc. of American Control Conference (ACC)*, pages 4202–4208, 2013.

[11] E. Bitar, R. Rajagopal, P. Khargonekar, and K. Poolla. The Role of Co-Located Storage for Wind Power Producers in Conventional Electricity Markets. In *Proc. of American Control Conference (ACC)*, pages 3886–3891, 2011.

[12] J. H. Kim and W. B. Powell. Optimal Energy Commitments with Storage and Intermittent Supply. *Operations Research*, 59(6):1347–1360, 2011.

[13] P. M. van de Ven, N. Hegde, L. Massoulie, and T. Salonidis. Optimal Control of End-User Energy Storage. *ArXiv e-prints*, 2012.

[14] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. Optimal Power Cost Management Using Stored Energy in Data Centers. In *Proc. of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '11, pages 221–232, 2011.

[15] L. Huang, J. Walrand, and K. Ramchandran. Optimal Demand Response with Energy Storage Management. In *Proc. of IEEE Third International Conference on Smart Grid Communications*, pages 61–66, 2012.

[16] L. Xie, Y. Gu, A. Eskandari, and M. Ehsani. Fast MPC-Based Coordination of Wind Power and Battery Energy Storage Systems. *Journal of Energy Engineering*, 138(2):43–53, 2012.

[17] National Renewable Energy Laboratory. The Value of Energy Storage for Grid Applications, 2013.

[18] M. J. Neely. Stochastic Network Optimization with Application to Communication and Queueing Systems. *Synthesis Lectures on Communication Networks*, 3(1):1–211, 2010.

[19] P. Mokrian and M. Stephen. A Stochastic Programming Framework for the Valuation of Electricity Storage. In *Proc. of 26th USAEE/IAEE North American Conference*, pages 24–27, 2006.

[20] B. Stott, J. Jardim, and O. Alsac. DC Power Flow Revisited. *IEEE Transactions on Power Systems*, 2009.

[21] J. Qin, Y. Chow, J. Yang, and R. Rajagopal. Modeling and Online Control of Generalized Energy Storage Networks. Technical report, 2014. Available at `http://www.stanford.edu/~jqin/pdf/QCYR2014report.pdf`.

[22] A. Strehl, L. Li, E. Wiewiora, J. Langford, and M. Littman. PAC Model-Free Reinforcement Learning. In *Proc. of the 23rd International Conference on Machine Learning*, pages 881–888. ACM, 2006.

[23] B. Zhang and D. Tse. Geometry of Feasible Injection Region of Power Networks. *IEEE Transactions on Power Systems*, 28(2):788–797, 2013.

[24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

APPENDIX

A. STRUCTURAL PROPERTIES OF THE ON-LINE OPTIMIZATION

We consider replacing the $g(t)$ defined in (3) with an extended real-valued function

$$g(t) = \phi_t(u(t), \boldsymbol{\kappa}(t)), \qquad (27)$$

where $\boldsymbol{\kappa}(t)$ is a vector of auxiliary parameters that captures both stochastic parameters and deterministic parameters, and it is supported on a compact set C_K. Observe that this would make our analysis applicable to general cost functions. Similar to discussions in Section 4, we are interested in solving the following optimization in each time period t for $u(t)$

$$\text{minimize} \quad \lambda \tilde{s}(t)u(t) + W\phi_t(u(t), \boldsymbol{\kappa}(t)) \qquad (28a)$$

$$\text{subject to} \quad U^{\min} \leq u(t) \leq U^{\max}, \qquad (28b)$$

after observing the realization of $\boldsymbol{\kappa}(t)$.

Lemma A.1 (Structural Properties of Single Bus Online Optimization). *For an extended real valued function $\phi_t(u(t), \boldsymbol{\kappa}(t))$, let $\hat{\phi}^{\boldsymbol{\kappa}}(u) \triangleq \phi_t(u, \boldsymbol{\kappa})$, where $\boldsymbol{\kappa}$ is equal to the observed value of $\boldsymbol{\kappa}(t)$. The following statements hold, regardless of the realizations of $\boldsymbol{\kappa}(t)$.*

1. *if $\lambda \tilde{s}(t) + W\underline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u) \geq 0$, then $u^{\mathrm{ol}}(t) = U^{\min}$;*

2. *if $\lambda \tilde{s}(t) + W\overline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u) \leq 0$, then $u^{\mathrm{ol}}(t) = U^{\max}$.*

Here,

$$\underline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u) \triangleq \inf \left\{ \xi \in \partial\hat{\phi}^{\boldsymbol{\kappa}}(u) \,\middle|\, \begin{matrix} u \in [U^{\min}, U^{\max}], \\ \boldsymbol{\kappa} \in C_K \end{matrix} \right\},$$

$$\overline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u) \triangleq \sup \left\{ \xi \in \partial\hat{\phi}^{\boldsymbol{\kappa}}(u) \,\middle|\, \begin{matrix} u \in [U^{\min}, U^{\max}], \\ \boldsymbol{\kappa} \in C_K \end{matrix} \right\}.$$

and $\partial\hat{\phi}^{\boldsymbol{\kappa}}(u)$ is the sub-differential of $\phi^{\boldsymbol{\kappa}}(u)$[7].

Proof. To show the set of sufficient conditions for $u^{\mathrm{ol}}(t)$ takes $U^{\max}$ (or $U^{\min}$), notice that the condition

$$\lambda \tilde{s}(t) \leq -W\overline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u)$$

implies $\partial J_t^{\boldsymbol{\kappa}, \tilde{s}}(u) \subseteq (-\infty, 0]$ for any $u(t)$, $\boldsymbol{\kappa}(t)$, $\tilde{s}(t)$. Thus, for every given $u \in [U^{\min}, U^{\max}]$, if β is a constant such that

$$J_t^{\boldsymbol{\kappa}, \tilde{s}}(v) - J_t^{\boldsymbol{\kappa}, \tilde{s}}(u) \geq \beta \cdot (v - u), \ \forall v \in [U^{\min}, U^{\max}],$$

then the sub-differential condition implies that $\beta \leq 0$. Now, by substituting $u = U^{\max}$ in the above expression, one obtains $\beta \cdot (v - u) \geq 0$ and $J_t^{\boldsymbol{\kappa}, \tilde{s}}(v) \geq J_t^{\boldsymbol{\kappa}, \tilde{s}}(U^{\max})$, for all $v \in [U^{\min}, U^{\max}]$. Therefore, one concludes that $u(t) = U^{\max}$ attains an optimal solution in problem (28). Similarly, the condition

$$\lambda \tilde{s}(t) \geq -W\underline{D}\hat{\phi}^{\boldsymbol{\kappa}}(u)$$

implies $\partial J_t^{\boldsymbol{\kappa}, \tilde{s}}(u(t)) \subseteq [0, \infty)$. Based on analogous arguments, one concludes that $u(t) = U^{\min}$ attains an optimal solution in problem (28). □

B. PROOF OF SINGLE BUS RESULTS

Proof of Theorem 4.5 We first validate that the intervals of Γ and W are non-empty. Note that from Assumption 3.1, $W^{\max} > 0$, thus it remains to show $\Gamma^{\max} \geq \Gamma^{\min}$. Based on (20), $W \geq 0$, and $\overline{D}g \geq \underline{D}g$, one obtains

$$W(\overline{D}g - \underline{D}g) \leq \lambda(S^{\max} - S^{\min}) - (U^{\max} - (1-\lambda)S^{\max})^+ - ((1-\lambda)S^{\min} - U^{\min})^+.$$

Re-arranging terms results in

$$\left[-W\underline{D}g + (U^{\max} - (1-\lambda)S^{\max})^+ \right] - \lambda S^{\max}$$
$$\leq \left[-W\overline{D}g - \left((1-\lambda)S^{\min} - U^{\min}\right)^+ \right] - \lambda S^{\min}$$

---

[7] We say β is a sub-derivative of an extended real-valued function ϕ at $u_0 \in \mathbb{R}$ if $\beta \cdot (u - u_0) \leq \phi(u) - \phi(u_0)$, for any $u \in \mathbb{R}$, and denote the set of all such β, namely the sub-differential of ϕ at u_0, by $\partial\phi(u_0)$.

which further implies $\Gamma^{\max} \geq \Gamma^{\min}$.

We proceed to show that

$$S^{\min} \leq s(t) \leq S^{\max}, \qquad (29)$$

for $t = 1, 2, \ldots$, when $u^\star(\mathbf{P3})$ is implemented. The base case holds by assumption. Let the inductive hypothesis be that (29) holds at time t. The storage level at $t+1$ is then $s(t+1) = \lambda s(t) + u^{\mathrm{ol}}(t)$. We show (29) holds at $t+1$ by considering the following three cases.

Case 1. $-W\underline{D}g \leq \lambda\tilde{s}(t) \leq \lambda(S^{\max} + \Gamma)$.

First, it is easy to verify that the above interval for $\lambda\tilde{s}(t)$ is non-empty using (18) and $\Gamma \geq \Gamma^{\min}$. Next, based on Lemma 4.3, one obtains $u^{\mathrm{ol}}(t) = U^{\min} \leq 0$ in this case. Therefore

$$s(t+1) = \lambda s(t) + U^{\min} \leq \lambda S^{\max} + U^{\min} \leq S^{\max},$$

where the last inequality follows from Assumption 2.1. On the other hand,

$$
\begin{aligned}
s(t+1) = \lambda s(t) + U^{\min} &\geq -W\underline{D}g - \lambda\Gamma + U^{\min} \\
&\geq -W\underline{D}g - \lambda\Gamma^{\max} + U^{\min} \\
&\geq \left((1-\lambda)S^{\min} - U^{\min}\right)^+ + \lambda S^{\min} + U^{\min} \geq S^{\min},
\end{aligned}
$$

where the third line used $\overline{D}g \geq \underline{D}g$.

Case 2. $\lambda(S^{\min} + \Gamma) \leq \lambda\tilde{s}(t) \leq -W\overline{D}g$.

The above interval for $\lambda\tilde{s}(t)$ is non-empty by (19) and $\Gamma \leq \Gamma^{\max}$. Lemma 4.3 implies $u^{\mathrm{ol}}(t) = U^{\max} \geq 0$ in this case. Therefore, using Assumption 2.1,

$$s(t+1) = \lambda s(t) + U^{\max} \geq \lambda S^{\min} + U^{\max} \geq S^{\min}.$$

On the other hand,

$$
\begin{aligned}
s(t+1) = \lambda s(t) + U^{\max} &\leq -W\overline{D}g - \lambda\Gamma + U^{\max} \\
&\leq -W\overline{D}g - \lambda\Gamma^{\min} + U^{\max} \\
&\leq -(U^{\max} - (1-\lambda)S^{\max})^+ + \lambda S^{\max} + U^{\max} \leq S^{\max},
\end{aligned}
$$

where the third line again is by $\overline{D}g \geq \underline{D}g$.

Case 3. $-W\overline{D}g < \lambda\tilde{s}(t) < -W\underline{D}g$.

By $U^{\min} \leq u^{\mathrm{ol}}(t) \leq U^{\max}$, one obtains

$$
\begin{aligned}
s(t+1) = \lambda s(t) + u^{\mathrm{ol}}(t) &\leq \lambda s(t) + U^{\max} \\
&< -W\underline{D}g - \lambda\Gamma + U^{\max} \\
&\leq -W\underline{D}g - \lambda\Gamma^{\min} + U^{\max} \\
&\leq -(U^{\max} - (1-\lambda)S^{\max})^+ + \lambda S^{\max} + U^{\max} \leq S^{\max}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
s(t+1) = \lambda s(t) + u^{\mathrm{ol}}(t) &\geq \lambda s(t) + U^{\min} \\
&> -W\overline{D}g - \lambda\Gamma + U^{\min} \\
&\geq -W\overline{D}g - \lambda\Gamma^{\max} + U^{\max} \\
&\geq \left((1-\lambda)S^{\min} - U^{\min}\right)^+ + \lambda S^{\min} + U^{\min} \geq S^{\min}.
\end{aligned}
$$

Combining these three cases, and by mathematical induction, we conclude (29) holds for all $t = 1, 2, \ldots$.

Proof of Theorem 4.6 Consider a quadratic Lyapunov function $L(s) = s^2/2$. Let the corresponding Lyapunov drift be

$$\Delta(\tilde{s}(t)) = \mathbb{E}\left[L(\tilde{s}(t+1)) - L(\tilde{s}(t))|\tilde{s}(t)\right].$$

Recall that $\tilde{s}(t+1) = s(t+1) + \Gamma = \lambda\tilde{s}(t) + u(t) + (1-\lambda)\Gamma$, and so

$$
\begin{aligned}
\Delta(\tilde{s}(t)) =\ & \mathbb{E}\big[(1/2)(u(t) + (1-\lambda)\Gamma)^2 - (1/2)(1-\lambda^2)\tilde{s}(t)^2 \\
& + \lambda\tilde{s}(t)u(t) + \lambda(1-\lambda)\tilde{s}(t)\Gamma|\tilde{s}(t)\big] \\
\leq\ & M^u(\Gamma) - (1/2)(1-\lambda^2)\tilde{s}(t)^2 \\
& + \mathbb{E}\big[\lambda\tilde{s}(t)u(t) + \lambda(1-\lambda)\tilde{s}(t)\Gamma|\tilde{s}(t)\big] \\
\leq\ & M^u(\Gamma) + \mathbb{E}\left[\lambda\tilde{s}(t)(u(t) + (1-\lambda)\Gamma)|\tilde{s}(t)\right]. \quad (30)
\end{aligned}
$$

It follows that, with arbitrary storage operation $u(t)$,

$$
\begin{aligned}
&\Delta(\tilde{s}(t)) + W\mathbb{E}[g(t)|\tilde{s}(t)] \qquad\qquad (31) \\
&\leq M^u(\Gamma) + \lambda(1-\lambda)\tilde{s}(t)\Gamma + \mathbb{E}[J_{\mathrm{P3},t}(u(t))|\tilde{s}(t)],
\end{aligned}
$$

where it is clear that minimizing the right hand side of the above inequality over $u(t)$ is equivalent to minimizing the objective of **P3**. Given that $u^{\mathrm{stat}}(t)$, the disturbance-only stationary policy of **P2** described in Lemma 4.2, is feasible for **P3**, the above inequality implies

$$
\begin{aligned}
&\Delta(\tilde{s}(t)) + W\mathbb{E}[g(t)|\tilde{s}(t), u(t) = u^{\mathrm{ol}}(t)] \qquad (32) \\
&\leq M^u(\Gamma) + \lambda(1-\lambda)\tilde{s}(t)\Gamma + \mathbb{E}\big[J_{\mathrm{P3},t}^\star|\tilde{s}(t)\big] \\
&\leq M^u(\Gamma) + \lambda(1-\lambda)\tilde{s}(t)\Gamma + \mathbb{E}\big[J_{\mathrm{P3},t}(u^{\mathrm{stat}}(t))|\tilde{s}(t)\big] \\
&\overset{(a)}{=} M^u(\Gamma) + \lambda\tilde{s}(t)\mathbb{E}\left[u^{\mathrm{stat}}(t) + (1-\lambda)\Gamma\right] + W\mathbb{E}[g(t)|u^{\mathrm{stat}}(t)] \\
&\overset{(b)}{\leq} M(\Gamma) + W\mathbb{E}[g(t)|u^{\mathrm{stat}}(t)] \overset{(c)}{\leq} M(\Gamma) + W J_{\mathrm{P1}}^\star.
\end{aligned}
$$

Here (a) uses the fact that $u^{\mathrm{stat}}(t)$ is induced by a disturbance-only stationary policy; (b) follows from inequalities $|\tilde{s}(t)| \leq \left(\max\left((S^{\max} + \Gamma)^2, (S^{\min} + \Gamma)^2\right)\right)^{1/2}$ and $\left|\mathbb{E}\left[u^{\mathrm{stat}}(t)\right] + (1-\lambda)\Gamma\right| \leq (1-\lambda)(\max((S^{\max} + \Gamma)^2, (S^{\min} + \Gamma)^2))^{1/2}$; and (c) used $\mathbb{E}[g(t)|u^{\mathrm{stat}}(t)] = J_{\mathrm{P2}}^\star$ in Lemma 4.2 and $J_{\mathrm{P2}}^\star \leq J_{\mathrm{P1}}^\star$. Taking expectation over $\tilde{s}(t)$ on both sides gives

$$
\begin{aligned}
&\mathbb{E}\left[L(\tilde{s}(t+1)) - L(\tilde{s}(t))\right] + W\mathbb{E}\left[g(t)|u(t) = u^{\mathrm{ol}}(t)\right] \\
&\leq M(\Gamma) + W J_{\mathrm{P1}}^\star. \qquad\qquad (33)
\end{aligned}
$$

Summing expression (33) over t from 1 to T, dividing both sides by WT, and taking the limit $T \to \infty$, we obtain the performance bound in expression (21).

Proof of Lemma 4.8 Based on the following re-parametrizations:

$$\eta^u = M^u(\Gamma)/W, \eta^s = M^s(\Gamma)/W,$$

(since $W > 0$) one can easily show that problem **P3-PO** has a same solution as the following optimization problem:

$$
\begin{aligned}
\text{minimize}\quad & \eta^u + \lambda(1-\lambda)\eta^s \\
\text{subject to}\quad & \Gamma^{\min} \leq \Gamma \leq \Gamma^{\max}, 0 < W \leq W^{\max}, \\
& 2\eta^u W \geq \left(U^{\min} + (1-\lambda)\Gamma\right)^2, \\
& 2\eta^u W \geq \left(U^{\max} + (1-\lambda)\Gamma\right)^2, \\
& \eta^s W \geq \left(S^{\min} + \Gamma\right)^2, \eta^s W \geq (S^{\max} + \Gamma)^2.
\end{aligned}
$$

The proof is completed by applying Schur complement on the last four constraints of the above optimization.