The Reliability Value of Storage in a Volatile Environment

Ali Parandehgheibi, Mardavij Roozbehani, and Asuman Ozdaglar

Abstract—**This paper examines a supply chain model with uncertain demand, supply friction, and storage. The base demand is assumed to be perfectly predictable, while deviations from the base are modeled as random shocks with stochastic arrivals. Supply is assumed to be able to track the base demand perfectly. However, due to friction, the random surge shocks cannot be tracked by supply. When storage which is assumed to be frictionless is available, it can be used to compensate for the surge in demand, otherwise, some demand must be curtailed, i.e., a blackout will occur. The problem of optimal utilization of storage is formulated as minimization of the expected discounted cost of blackouts over an infinite horizon. It is shown that when the stage cost is linear in the size of the blackout, the optimal policy is myopic in the sense that the demand shock is masked by storage up to the available level of storage. However, when the stage cost is strictly convex, it is sometimes optimal to curtail some of the demand and allow a small blackout now in favor of maintaining a higher level of reserve to avoid a large blackout in the future.**

Index Terms—**Storage, Ramp Constraints, Reliability, Probability of Large Blackouts**

I. Introduction

Supply and demand in electric power networks are subject to exogenous, impulsive, and unpredictable shocks due to generator outages, failure of transmission equipments or unexpected changes in weather conditions. On the other hand, price pressure and environmental concerns have led to a global trend in large-scale integration of renewable resources with stochastic output. This will only add to the magnitude and frequency of impulsive shocks to the supply network. We ask what is the value of storage in mitigation volatility of of supply and demand, and what are the fundamental limits that cannot be overcome by storage due to physical ramp constraints, and finally, what are the impacts of different control policies on system reliability.

In this paper we are concerned with the reliability value of storage, defined as the maximal improvement is system reliability as a function of storage capacity. Two metrics for quantifying reliability in a system are considered: one is the expected long-term discounted cost of blackouts, and the other is the probability of loss of load by a certain amount or less. By taking out other factors such as the environment, cost of energy, and cost of storage, we characterize the value of storage purely from a reliability perspective. We formulate the problem of optimal storage management as the problem of minimization of the expected long term discounted cost of blackouts, and further investigate the effects of optimal strategies with respect to the COB metric on the LOLP metric.

our work is related to these: [9] [4] [5] [6] [3] ***

Although storage can be drained instantaneously to make up for the lost power or demand shock, this myopic strategy may not always be optimal.

Notation. *Throughout the paper,* $\mathbb{I}_A$ *denotes the indicator function of a set* A.

The organization of this paper is as follows. Section II presents the elements of the model and the problem formulation. Section III includes the main results.

II. The Model

We examine an abstract model of system consisting of a single consumer, a single fully controllable supplier, a supplier with stochastic output (e.g., wind), and a storage system with finite capacity (Figure 1). These agents each represent an aggregate of several small consumers and producers.

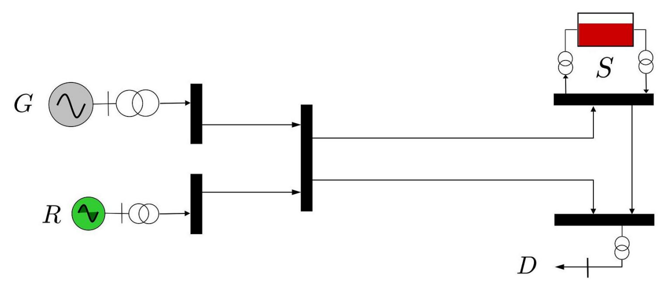

Fig. 1. Layout of the physical layer of a power supply network with conventional and renewable generation, storage, and demand.

The details of the elements of the model are as follows.

A. Supply

1) Controllable Supply: The controllable supply process is denoted by $\mathbf{G} = \{G(t) : t \geq 0\}$, where $G(t)$ is the power output at time $t \geq 0$. It is assumed that the supplier's production is subject to an upward ramp constraint, in the sense that its output cannot increase instantaneously,

$$\frac{G(t) - G(t')}{t - t'} \leq \zeta, \qquad \forall t : 0 \leq t < t'.$$

We do not assume a downward ramp constraint or a maximum capacity constraint on $G(t)$. Thus, production can shut down instantaneously, and can meet any large demand sufficiently far in the future.

This work was supported by X and Y

Ali Parandehgheibi (parandeh@mit.edu), Mardavij Roozbehani (mardavij@mit.edu), and Asuman Ozdaglar (asuman@mit.edu) are with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

2) Renewable Supply: The renewable supply process is denoted by $\mathbf{R} = \{R(t) : t \geq 0\}$. It is assumed that $\mathbf{R}$ can be modeled as a process with two components: $\mathbf{R} = \overline{\mathbf{R}} + \Delta\mathbf{R}$, where $\overline{\mathbf{R}} = \{\overline{R}(t) : t \geq 0\}$ is a deterministic process representing the predicted renewable supply, and $\Delta\mathbf{R} = \{\Delta R(t) : t \geq 0\}$ is the forecast error assumed to be a random arrival process. Thus, at any given time $t \geq 0$, the total forecast supply from the renewable and controllable generators is thus given by $G(t) + \overline{R}(t)$.

The details of the forecast error process are outlined in Section **??**.

B. Demand

The demand process is denoted by $\mathbf{D} = \{D(t) : t \geq 0\}$, where $D(t)$ is the total power demand at time t, assumed to be exogenous and inelastic. Similar to the renewable supply, $\mathbf{D}$ has two components: $\mathbf{D} = \overline{\mathbf{D}} + \Delta\mathbf{D}$, where $\overline{\mathbf{D}} = \{\overline{D}(t) : t \geq 0\}$ is the predicted demand process (deterministic), and $\Delta\mathbf{D} = \{\Delta D(t) : t \geq 0\}$ is the forecast error, again, assumed to be a random arrival process.

We refer to the event of not meeting the demand as a *blackout*, and associate to it a cost which is an increasing function of the total lost energy. Throughout the paper, $g(x)$ denotes the marginal (with respect to time) cost of losing x units of power, where $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an increasing function. The total cost associated with a blackout specified by $x(t),\ t \in T \subset \mathbb{R}_+$ where x is a right-continuous function mapping $\mathbb{R}_+$ to $\mathbb{R}_+$, is thus given by

$$(\Psi x)(T) = \int_T g(x(\tau))d\tau$$

C. Storage

The storage process is denoted by $\mathbf{S} = \{S(t) \in [0, \overline{S}] : t \geq 0\}$, where $S(t)$ is the amount of stored energy at time t, and $\overline{S} < \infty$ is the storage capacity. The storage technology is subject to an upward ramp constraint:

$$\frac{S(t) - S(t')}{t - t'} \leq r, \qquad \forall t : 0 \leq t < t'.$$

Thus, storage cannot be filled up instantaneously, though, it can be drained (to supply power) instantaneously. Let $\mathbf{U} = \{U(t) : t \geq 0\}$, be the power withdrawal process from storage. The dynamics of the storage process is then given by:

$$S(t) = S_0 + \int_0^t \mathbb{I}_{\{S<\overline{S}\}} r d\tau - \int_0^t U(\tau)\, d\tau \tag{1}$$

One of our goals in this paper is to design a causal controller $K(\mathbf{S}, \mathbf{G} + \mathbf{R} - \mathbf{D})$, such that $U(t) = K(S, G + R - D)$ maximizes the system reliability objectives.

D. The Dynamics

Definition 1. The *normalized* demand is defined as the *current demand* minus the *forecast* demand. Similarly, the *normalized* supply is defined as the *current supply* minus the *forecast* supply. The *normalized power imbalance* is the normalized demand minus the normalized supply.

$$x(t) = \Delta D(t) - \Delta R(t) \tag{2}$$

Finally, the *normalized energy imbalance* is defined as:

$$W(t) = \frac{x(t)^2}{2\zeta} \tag{3}$$

Assumption I: The forecast supply is equal to the forecast demand. That is:

$$\overline{D}(t) = G(t) + \overline{R}(t), \qquad \forall t \geq 0$$

Under Assumption I, the energy from storage will be used only to compensate for the power imbalance, since in a perfect information scenario (i.e., when $x(t) = 0$ for all $t \geq 0$), supply is always equal to demand, and storage provides no additional utility.

The power imbalance process (2) can be mapped to an energy imbalance process in the following way. Suppose that an imbalance shock arrives which causes the loss of x units of power generation. Immediately after this event, the generators ramp up their production at the maximum rate ζ. It thus takes $T_\zeta(x) = x(t)/\zeta$ units of time until the service is fully restored, which, in the absence of storage, would result in a total energy loss of $W(t) = x(t)^2/2\zeta$ units. Since storage can supply instantaneously, it is possible to compensate for this energy loss by withdrawing from the storage (subject to its availability), or choose to curtail some demand up to the size of the energy shock and incur a blackout cost.

Assumption II: The normalized energy imbalance process (3) is a compound poisson process with arrival rate Q. ** more elaboration is needed here on compound process**

The dynamics of the storage process is then given by:

$$S(t) = S_0 + \int_0^t \mathbb{I}_{\{S<\overline{S}\}} r d\tau - \int_0^t K\left(S_{\tau^-}, Y_\tau - Y_{\tau^-}\right) dN_\tau \tag{4}$$

E. Problem Formulation

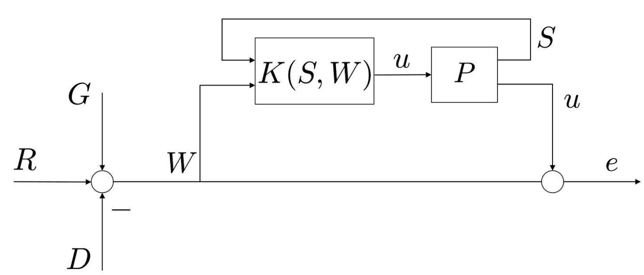

Fig. 2. The control layer of the power supply network in Figure 1.

We are now ready to formally formulate the problem. Let $J_u(s)$ denote the expected long-term discounted cost of blackouts starting from an initial state s and under control policy u,

$$J_u(s) = \mathbf{E}\left[\sum_{k=1}^{\infty} e^{-\theta\tau_k} g\left(W_k - u(s_{\tau_k -}, W_k)\right)\Big| s_0 = s\right]. \tag{5}$$

The system reliability problem can now be formulated as an infinite horizon stochastic optimal control problem as follows:

$$J_u(s) \to \min_{u\in\mathcal{U}(s)} \tag{6}$$

where the optimization problem (6) is subject to the state dynamics (4).

III. Main Results

IV. System Model and Problem Formulation

Consider a single energy storage system with capacity $\bar{s}$. The energy storage can drain power from the grid, and supports the demand.

Assume that the demand process receives positive shocks according to a compound Poisson process:

$$Y_t = \sum_{k=1}^{N_t} W_k, \tag{7}$$

where N_t is a Poisson process of rate R and W_k's denote the shock sizes and are drawn independently from distribution f_W. We assume that shock sizes are bounded by B. Given the *càdlàg* process Y_t, we may define the jump size process

$$V_t = Y_t - Y_{t^-}. \tag{8}$$

In this model, for every demand shock, the main grid increases the supply accordingly to support the demand. However, due to ramp constraint on the conventional generators, there might be a small gap until the increased demand is satisfied, which results in a partial blackout. We have the option to mask this gaps (completely or partially) using the energy storage, hence, resulting in smaller partial blackouts.

Assume that we have an energy storage of size $\bar{s}$. The storage has a ramp constraint r for charging but no ramp constraint for draining. Let s_t denote the amount of energy stored at time t. The controlled process s_t obeys a (stochastic differential equation) SDE of the following form:

$$s_t = s_0 + \int_0^t \mathbb{I}_{\{s<\bar{s}\}} r d\tau - \int_0^t u_\tau dN_\tau, \tag{9}$$

where $\mathbb{I}$ denotes the indicator function and N_t is a Poisson counter with rate R. Moreover, u_t denotes the control process. In this work, we assume that u_t is given by a stationary Markovian policy $\mu : [0,\bar{s}]\times[0,B]\to[0,B]$, that observes the storage state and demand shock size process V_t. In particular

$$u_t = \mu(s_{t^-}, V_t). \tag{10}$$

Given the shock size w and storage level s, we have the following natural constraints on the policy $\mu(s,w)$:

$$0 \le \mu(s,w) \le \min\{s,w\}, \quad \text{for all } s\in[0,\bar{s}],\ w\in[0,B].$$

We refer to such policies as *admissible*, and denote the set of all such policies by Π.

Upon arrival of a shock of size w, if $\mu(s,w) = w$, there is no service interruption (blackout) experienced by the consumers until the generators of the main grid ramp up. Otherwise, depending on the size of the blackout, $w-\mu(s,w)$, a penalty must be paid. Let $g : [0,B] \to [0,\infty)$ denote the penalty function. Throughout this work we assume that g is nondecreasing, bounded and without loss of generality $g(0) = 0$.

Let $C_\mu(s)$ denote a discounted average penalty paid over an infinite horizon. We have

$$C_\mu(s) = \mathbf{E}\left[\int_0^\infty e^{-\theta\tau} g\left(V_\tau - u_\tau\right) dN_\tau \Big| s_0 = s\right] \tag{11}$$

$$= \mathbf{E}\left[\sum_{k=1}^{\infty} e^{-\theta t_k} g\left(W_k - \mu(s_{t_k^-}, W_k)\right) \Big| s_0 = s\right], \tag{12}$$

where t_k is the k-th Poisson arrival time. The goal is to find a policy in the space of admissible policies Π, that minimize the expected penalty $C_\mu(s)$. Define the *value function* or optimal cost function as

$$C(s) = \min_{\mu\in\Pi} C_\mu(s), \quad 0\le s\le \bar{s}. \tag{13}$$

A policy $\mu^* \in \Pi$ is defined to be optimal if it achieves the value function, i.e., $C_{\mu^*}(s) = C(s)$.

A. Characterizations of the Value Function

We first provide several characterizations for the value function defined in (13) and derive specific properties for the value function which are useful in characterization of the optimal policy.

Let $J_\mu(s,w)$ be the expected long-term discounted cost under policy μ conditioned on the first jump arriving at time $t_1 = 0$, and being of size w. Here, s is the state of the system before executing the action dictated by the policy. By memoryless property of the Poisson process, we have

$$J_\mu(s,w) = g(w-\mu(s,w))$$
$$+\mathbf{E}\Big[\sum_{k=1}^{\infty} e^{-\theta t_k} g(W_k - \mu(s_{t_k^-}, W_k))\Big| s_0 = s-\mu(s,w)\Big]. \tag{14}$$

We may relate $J_\mu(s,W)$ to the total expected cost $C_\mu(s)$ defined in (11) as follows:

$$C_\mu(s) = \mathbf{E}_{(t_0,W)}\left[e^{-\theta t_0} J_\mu(\min\{s+rt_0,\bar{s}\}, W)\right], \tag{15}$$

where t_0 is an exponential random variable with mean $1/R$, and is independent of W which is drawn from distribution f_W.

From (15), it is clear that from the minimization of J_μ across all admissible policies Π, we may obtain the optimal solution to the original problem in (13). The discrete-time formulation of J_μ given by (14), facilitates deriving the Bellman equation as the necessary and sufficient optimality condition, as well as development of efficient numerical methods. For completeness, we summarize these results in the following theorem.

Theorem 1. *Define $J^*(s,w) = min_{\mu\in\Pi} J_\mu(s,w)$, where J_μ is defined in (14). If the stage cost g is bounded, and $\theta > 0$, the following fixed-point equation is both necessary and sufficient for optimality of J.*

$$J(s,w) = (TJ)(s,w) \triangleq \min_{0\le u\le \min\{s,w\}} \Big\{ g(w-u) + \mathbf{E}\Big[e^{-\theta t_0} J\big(\min\{s-u+rt_0, \bar{s}\}, W\big)\Big]\Big\} \quad (16)$$

Moreover, a stationary policy $\mu^*(s,w)$ is optimal if and only if it achieves the minimum in (16). Further, the value iteration algorithm

$$J_{k+1} = TJ_k, \quad (17)$$

converges to the optimal solution J^* for any initial condition J_0.

Proof: The result follows from establishing the contraction property of T, which is standard for discounted problems with bounded stage cost. See [2] for more details. ■

Theorem 1 allows us to develop further properties of the optimal cost function. An alternative approach for characterizing the optimal cost function is the continuous-time analysis of problem (13) which leads to Hamilton-Jacobi-Bellman (HJB) equation. In the following Theorem we present some basic properties of the optimal cost function as well as HJB equation.

Theorem 2. *Let $C(s)$ be the optimal cost function defined in (13). The following statements hold:*

(i) *$C(s)$ is nonincreasing in s. Moreover, if the stage cost function $g(\cdot)$ is increasing and $E_W[g(W)] > 0$, then $C(s)$ is strictly decreasing in s.*

(ii) *If the stage cost $g(\cdot)$ is convex, the optimal cost function $C(s)$ is also convex in s.*

(iii) *If C is continuously differentiable, then for all $0 \le s \le \bar{s}$ it obeys the following differential equation*

$$\frac{dC}{ds} = \frac{R+\theta}{r}C(s) - \frac{R}{r}\mathbf{E}_W\Big[\min_{0\le u\le s,W} g(W-u) + C(s-u)\Big], \quad (18)$$

with the boundary condition

$$\frac{dC}{ds}\Big|_{s=\bar{s}} = 0. \quad (19)$$

Moreover, the optimal policy $\mu^*(s,w)$ achieves the optimal solution of the minimization problem in (18). Further, for a given policy μ, if the expected cost function $C_\mu(s)$ is differentiable, it satisfies the following delay differential equation

$$\frac{dC}{ds} = \frac{R+\theta}{r}C(s) - \frac{R}{r}\mathbf{E}_W\Big[g(W-\mu(s,W)) + C(s-\mu(s,W))\Big], \quad (20)$$

with the boundary condition given by (19).

Proof: See Appendix. ■

The result of Theorem 2 requires continuous differentiability of the optimal cost function, which does not hold in general. However, we may establish differentiability of the value function under some mild conditions such as differentiability of the stage cost function g and the probability density function $f_W(\cdot)$ of Poisson jumps (cf. Benveniste and Scheinkman [1]). Throughout this paper, we assume that $C(s)$ is in fact continuously differentiable and we may apply the results of Theorem 2.

In the following, we present some structural properties of the optimal policy using the characterizations of the optimal cost function given by Theorems 1 to 2.

B. Characterizations of the Optimal Policy

In this part, we present several characterizations of the optimal policy for different stage cost functions $g(\cdot)$. First, we show that the myopic policy is optimal for linear stage cost functions. Then, we partially characterize the structure of optimal policy for nonlinear stage cost functions.

Theorem 3. *If the stage cost is linear, i.e., $g(x) = \beta x$, then the myopic policy*

$$\mu^*(s,w) = \min\{s,w\}, \quad (21)$$

is optimal for the problem defined in (13).

Proof: See Appendix. ■

Next, we focus on nonlinear but convex stage cost functions. For the case of nonlinear stage cost functions, the myopic policy defined in (21) is no more optimal. Intuitively, the myopic policy greedily consumes the reserve and increases the chance of a large blackout. In the linear case, the penalty for a large blackout was equivalent to the total penalty of many small blackouts. However, in the nonlinear and convex case, the penalty for a large blackout is larger than multiple small blackouts. Therefore, the optimal policy in this case tends to be more conservative in using the reserve. Nevertheless, the structure of the optimal policy in the general case shows some similarities with the myopic policy. In the following we present some characterizations of the structural properties of the optimal policy using the results from Section IV-A.

Theorem 4. *Let $\mu^*(s,w)$ be the optimal policy under Assumption 1 and strict convexity of the stage cost. $\mu^*(s,w)$ is monotonically nondecreasing in both s and w.*

Proof: See Appendix. ■

Assumption 1. Let $\bar{W} = E[W]$ be the expected jump size of the arrival process. Assume the rate of the compound Poisson process is less than the ramp constraint, i.e., $R\bar{W} \le r$.

Theorem 5. *Let μ^* denote the optimal policy in the problem (13), where the stage cost g is strictly convex. There exist a unique kernel function $\phi : [-B, \bar{s}] \to \mathbb{R}$ such that*

$$\mu^*(s,w) = \Big[w - \phi(s-w)\Big]^+, \quad \textit{for all } (s,w) \in [0,\bar{s}]\times[0,B], \quad (22)$$

where $[x]^+ = \max\{x, 0\}$*, and for all* $p \in [-B, \bar{s}]$

$$\begin{aligned} \phi(p) \quad &= \quad \textit{argmin} \quad g(x) + C(p + x) \qquad (23) \\ & \quad \textit{s.t.} \\ & \quad x \geq \max\{0, -p\} \\ & \quad x \leq \min\{B, \bar{s} - p\}. \end{aligned}$$

Moreover, under Assumption 1, we can represent the kernel function $\phi(p)$ *as follows:*

$$\phi(p) = \begin{cases} -p, & -B \leq p \leq b_0 \\ \phi^\circ(p), & b_0 \leq p \leq b_1 \\ 0, & b_1 \leq p \leq \bar{s}, \end{cases} \qquad (24)$$

where $\phi^\circ(p)$ *is the unique solution of*

$$g'(x) + C'(x + p) = 0, \qquad (25)$$

and b_0 *and* b_1 *are the break-points, where*

$$b_0 = -(g')^{(-1)}\Big(C'(0)\Big) \geq -(g')^{(-1)}\Big(\frac{R}{r}\mathbf{E}[g(W)]\Big) \geq -B, \qquad (26)$$

$$b_1 = -(C')^{(-1)}\Big(g'(0)\Big) \leq \bar{s}. \qquad (27)$$

Proof: See Appnedix. ■

Theorem 5 demonstrates a very special structure for the optimal policy. In fact, we showed that the two dimensional policy can be represented using a single dimensional kernel function. This observation allows us to significantly reduce the computational complexity of numerical methods for computing the optimal policy.

Using the results of Theorem 5, we can provide a rather complete picture of the structure of the kernel function for large range of system parameters. Figure 3 illustrates a conceptual plot of the kernel function.

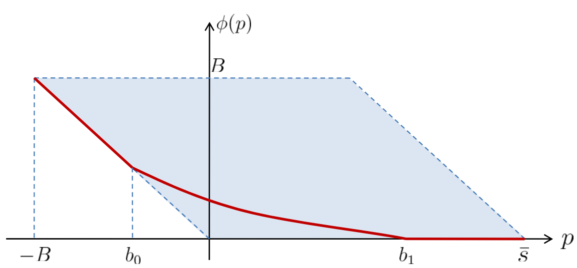

Fig. 3. Structure of the kernel function $\phi(p)$ defined in (23).

In the the following, we summarize the structure of the optimal policy using the results of Theorem 5. For any $w \in [0, B]$, there are two cases to consider. If $w \geq -b_0$, we have

$$\mu^*(s, w) = \begin{cases} s, & 0 \leq s \leq s_0(w) \\ w - \phi^\circ(s - w), & s_0(w) \leq s \leq s_1(w) \\ w, & s_1(w) \leq s \leq \bar{s}, \end{cases} \qquad (28)$$

where $s_i(w) = w + b_i$ for $i = 0, 1$. In the case where $w \leq -b_0$, we have

$$\mu^*(s, w) = \begin{cases} 0, & 0 \leq s \leq q_0(w) \\ w - \phi^\circ(s - w), & q_0(w) \leq s \leq s_1(w) \\ w, & s_1(w) \leq s \leq \bar{s}, \end{cases} \qquad (29)$$

where $s_0(w)$ is the unique solution of $\phi^\circ(s - w) = w$. Figure 4 illustrates an example of optimal policy for different ranges of s, w.

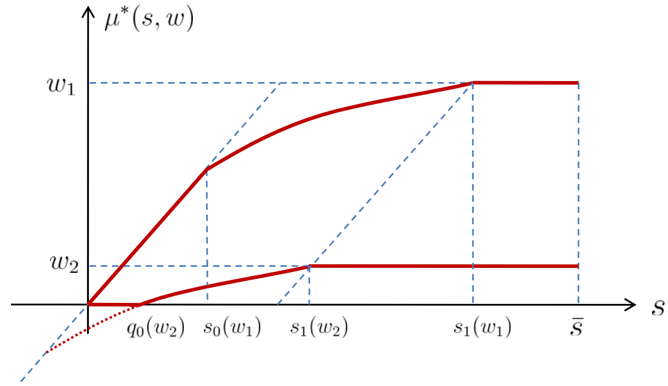

Fig. 4. Structure of the optimal policy $\mu^*(s, w)$ for a convex stage cost.

SOME EXAMPLES???

C. Statistics of Power Outages

V. Numerical Simulations

System parameters if not specified in the caption:

- $\theta = 0.1$
- $r = 1$
- $R = 0.8$
- $B = 1$
- $\bar{s} = 2$
- $g(x) = x^2$
- f_W uniform

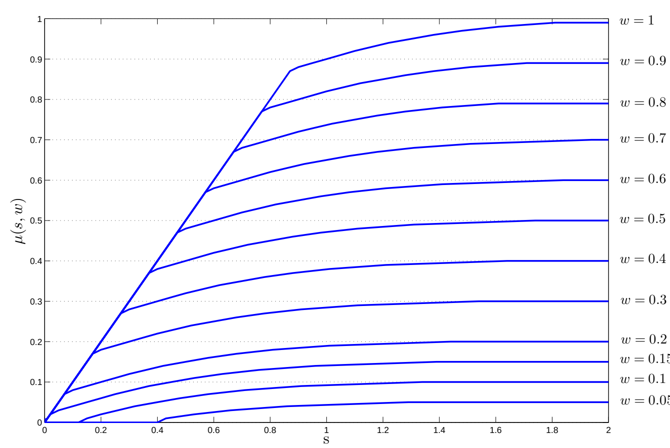

Fig. 5. Optimal policy computed by value iteration algorithm (17) for quadratic stage cost and uniform shock distribution.

$c(s; \bar{s})$ denotes the optimal cost function (13) when the storage size is given by $\bar{s}$.

VI. Conclusions

Appendix

Proof of Theorem 2: *Part (i):* The monotonicity property of the value function follows almost immediately from the

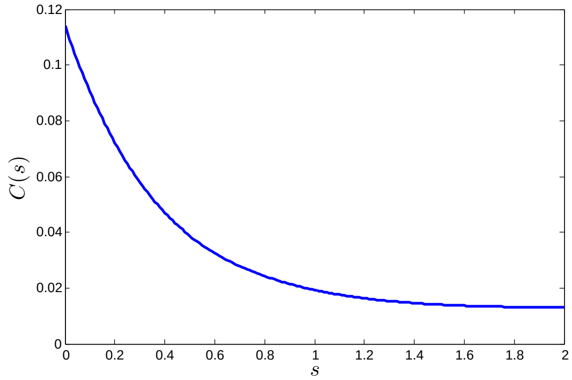

Fig. 6. Optimal cost function computed by value iteration algorithm (17) for quadratic stage cost and uniform shock distribution.

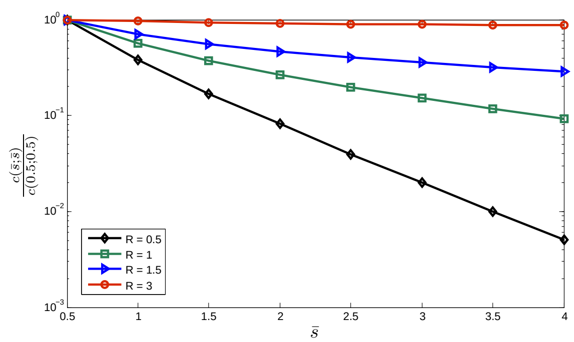

Fig. 7. Normalized blackout cost in log-scale as a function of the storage size for different arrival rates.

definition. Let $0 \leq s_1 < s_2 \leq \bar{s}$, and assume $C(s) = C_\mu(s)$ for some policy μ. Given the initial state s_1, let $u_t^{(1)}$ be the control process under policy μ. Note that for every realization ω of the compound Poisson process, the sample path $u_t^{(1)}(\omega)$ is admissible for initial condition $s_2 > s_1$. Therefore, by definitions (11) and (13), we have $C(s_2) \leq C(s_1)$.

In order to show the strict monotonicity, consider the controlled process starting from s_1. Let τ be the first arrival time such that $g(V_\tau - u_\tau^{(1)}) > 0$. From the hypothesis, we

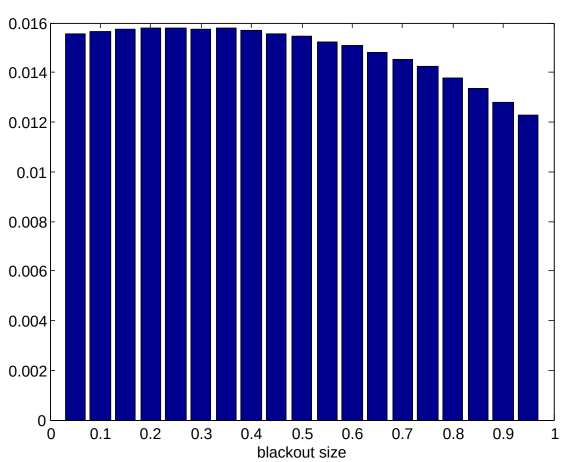

Fig. 8. Blackout distribution for myopic policy.

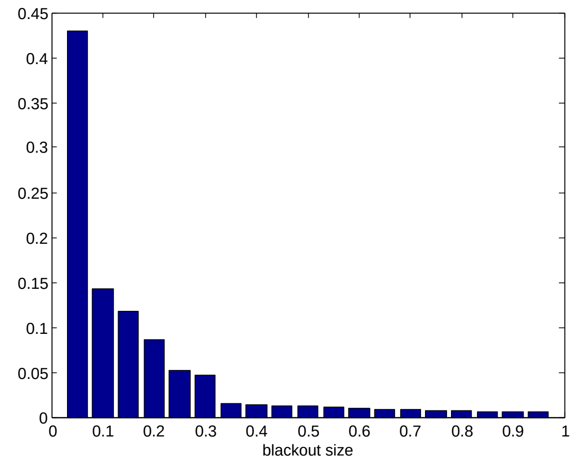

Fig. 9. Blackout distribution for optimal policy for quadratic stage cost.

have $\mathbf{P}(\tau \in [0,T]) > 0$ for some $T < \infty$. For every sample path ω, define the control process

$$u_t^{(2)}(\omega) = u_t^{(2)}(\omega) + \delta \cdot \mathbb{I}_{\{t=\tau(\omega)\}},$$

for some $\delta > 0$ such that $\delta \leq \min\{s_2 - s_1, V_{\tau(\omega)} - u_{\tau(\omega)}^{(1)}\}$.

It is clear that $u_t^{(2)}(\omega)$ is admissible for the controlled process starting from s_2. Using the definition of the expected cost function in (11), we can write

$$\begin{aligned} C(s_1) - C(s_2) &= \mathbf{E}_\omega[e^{-\theta\tau(\omega)} g(V_{\tau(\omega)} - u_{\tau(\omega)}^{(1)}) \\ &\quad - e^{-\theta\tau(\omega)} g(V_{\tau(\omega)} - u_{\tau(\omega)}^{(1)} - \delta)] \\ &\geq \mathbf{E}[\epsilon e^{-\theta\tau(\omega)}], \quad \text{for some } \epsilon > 0 \\ &\geq \epsilon e^{-\theta T} \mathbf{P}(\tau \in [0,T]) > 0, \end{aligned}$$

where the first inequality holds by strict monotonicity of g.

Part (ii): We first prove convexity of $J^*(s,w)$ defined in Theorem 1, and use it to establish convexity of $C(s)$.

In order to show convexity of $J^*(s,w)$, we need to show that the operator T defined in (16) preserves convexity. Then the claim would be immediate using the convergence of value iteration algorithm (17) to optimal cost J^*, where the initial condition is an arbitrary convex function such as $J_0 = 0$.

Next we show that the operator T preserves convexity for this particular problem. Define the objective function in (16) as $Q(s,w,u)$. We have

$$\begin{aligned} Q(s,w,u) &= g(w-u) + \mathbf{E}\Big[e^{-\theta t_0} J\big(\min\{s-u+rt_0, \bar{s}\}, W\big)\Big] \\ &= g(w-u) + \int_{\frac{\bar{s}-s+u}{r}}^{\infty} e^{-\theta t_0} \mathbf{E}[J(\bar{s}, W)] R e^{-Rt_0} dt_0 \\ &\quad + \int_0^{\frac{\bar{s}-s+u}{r}} e^{-\theta t_0} \mathbf{E}[J(s-u+rt_0, W)] R e^{-Rt_0} dt_0. \end{aligned}$$

Using the fact that J is convex, linearity of expectation and basic definition of a convex function, it is straightforward but tedious to show that $Q(s,w,u)$ is a convex function. We omit the details for brevity. Given the convexity of Q, the convexity of $(TJ)(s,w)$ is immediate, since we are minimizing a multidimensional convex function over one of its dimensions. Hence, we have established convexity of $J^*(s,w)$ in (s,w). Finally, we can express $C(s)$ in terms of $J^*(s,w)$ as in (15). This results in convexity of $C(s)$ using the above argument for proving convexity of $Q(s,w,u)$.

Part (iii): The derivation of Hamilton-Jacobi-Bellman is relatively standard. We present a proof sketch based on *principle of optimality*. For a more detailed treatment, please refer to [2], [8] and [7].

Let s_t be the state process under the optimal policy governed by the SDE in (9), and $s_0 = s < \bar{s}$. By principle of optimality, going from time 0 to time h, we have

$$C(s) = C(s_0) = \min_{u_t \in \Pi} \Big\{ \mathbf{E}\Big[\int_0^h e^{-\theta\tau} g(V_\tau - u_\tau) dN_\tau \Big] + e^{-\theta h} \mathbf{E}\big[C(s_h)\big] \Big\},$$

where the expectation is with respect to the compound Poisson process. Note that in this particular model, we assume that the control process u_t is progressively measurable with respect to the jump process.

Let N_h be the number of Poisson arrivals in $[0,h]$, where h is positive but small. There are three cases to consider: First, there are no arrivals in $[0,h]$, in which case the decision function is trivially zero, and no penalty occurs. Second, there is a single arrival in this interval. In this case, the control function u_t is reduced to a scalar decision u that is measurable w.r.t jump size W. Third, there are more than one arrivals in $[0,h]$, which occurs with probability $o(h)$. Since, the stage cost is bounded, the expected cost under this condition remains $o(h)$. Hence, for every $s < \bar{s}$,

$$\begin{aligned} C(s) &= \mathbf{E}_W \Big[\min_{0 \le u \le s, W} \Big\{ 0 \cdot \mathbf{P}(N_h = 0) \\ &\quad + g(W-u)\mathbf{P}(N_h = 1) \\ &\quad + e^{-\theta h} C(s + rdh)\mathbf{P}(N_h = 0) \\ &\quad + e^{-\theta h} C(s - u + rdh)\mathbf{P}(N_h = 1) + o(h) \Big\} \Big] \\ &= e^{-\theta h} C(s) + \mathbf{E}_W \Big[\min_{0 \le u \le s, W} \Big\{ Rh\, g(W-u) \\ &\quad + e^{-\theta h}(1 - Rh)(C(s + rdh) - C(s)) \\ &\quad + e^{-\theta h} Rh\, C(s-u) + o(h) \Big\} \Big]. \end{aligned} \tag{30}$$

Using the fact that C is differentiable on $[0,\bar{s})$, we may verify the result in (18), by dividing the above relation by h and taking the limit as h tends to zero.

The derivation of the boundary condition in (19) is similar. Note that for $s = \bar{s}$, when no Poisson arrival occurs in interval $[0,h]$, the state of the system stays at $\bar{s}$. Therefore, we can modify (30) as follows:

$$\begin{aligned} C(\bar{s}) = \mathbf{E}_W \Big[\min_{0 \le u \le \bar{s}, W} \Big\{ & 0 \cdot \mathbf{P}(N_h = 0) \\ & + g(W-u)\mathbf{P}(N_h = 1) + e^{-\theta h} C(\bar{s})\mathbf{P}(N_h = 0) \\ & + e^{-\theta h} C(\bar{s} - u)\mathbf{P}(N_h = 1) + o(h) \Big\} \Big]. \end{aligned} \tag{31}$$

Again, dividing by h and taking the limit as h goes to zero, we have

$$\frac{R+\theta}{r} C(\bar{s}) - \frac{R}{r} \mathbf{E}_W \Big[\min_{0 \le u \le \bar{s}, W} g(W-u) + C(\bar{s} - u) \Big] = 0,$$

which is the same as (19) in light of (18).

The derivation of the differential equation (20) for a particular policy is similar, noting the memoryless property of Poisson process and stationarity of the controlled state process. ■

Proof of Theorem 3: We establish optimality of μ^* by showing that it achieves an expected cost no higher than any other admissible policy.

Consider an admissible policy $\tilde{\mu}$ such that $\tilde{\mu}(s,w) < \min\{s,w\}$ for some set $(s,w) \in [0,\bar{s}] \times [0,B]$. For every sample path of the controlled process, let $\tau_1(\omega)$ be the first Poisson arrival time such that

$$\min\{s_{\tau_1^-}, V_{\tau_1}\} - \tilde{\mu}(s_{\tau_1^-}, V_{\tau_1}) = \epsilon > 0.$$

Therefore, by applying policy $\tilde{\mu}$ instead of μ^*, we pay an extra penalty of $\beta\epsilon e^{-\theta\tau_1(\omega)}$. The reward for this extra penalty is that the state process is now biased by at most ϵ, which allows us to avoid later penalties. However, since the stage cost is linear, the penalty reduction by this bias for any time $\tau_2(\omega) > \tau_1(\omega)$ is at most $\beta\epsilon e^{-\theta\tau_2(\omega)}$. Hence, for this sample path ω, the policy $\tilde{\mu}$ does worse than the myopic policy μ^* at least by

$$\beta\epsilon(e^{-\theta\tau_1(\omega)} - e^{-\theta\tau_2(\omega)}) > 0.$$

Therefore, by taking the expectation for all sample paths, the myopic policy cannot do worse than any other admissible policy. Note that this argument does not prove the uniqueness of μ^* as the optimal policy. In fact, we may construct optimal policies that are different from μ^* on a set $A \subseteq [0,\bar{s}] \times [0,B]$, where $\mathbf{P}((s_{t^-}, V_t) \in A) = 0$. ■

We delay the proof of Theorem 4 until after proof of Theorem 5. Let us start with some useful lemmas on the structure of the kernel function.

Lemma 1. *Let $\phi(p)$ be defined as in (23). We have*

1) *If $\phi(p_0) = -p_0$ for some p_0, then*

$$\phi(p) = -p, \quad \textit{for all } p \le p_0.$$

2) *If $\phi(p_0) = 0$ for some p_1, then*

$$\phi(p) = 0, \quad \textit{for all } p \ge p_1.$$

Proof: By convexity of the stage cost function and Theorem 2(ii), $\phi(p)$ is the optimal solution of a convex program. Therefore, if $\phi(p_0) = -p_0$ for some $p_0 \le 0$, we have

$$g'(-p_0) + C'(0) \ge 0.$$

Thus, by convexity of stage cost, $g(-p) \ge g(-p_0)$, for any $p \le p_0$. Therefore, by convexity of $C(\cdot)$ and $g(\cdot)$,

$$g'(x) + C'(x+p) \ge g'(-p) + C'(0) \ge 0, \text{for all } x \ge -p,$$

which immediately implies optimality of $(-p)$, for $p \le p_0$.

Similarly, for the case where $\phi(p_1) = 0$, we have

$$g'(0) + C'(p_1) \geq 0,$$

which implies

$$g'(x) + C'(x+p) \geq g'(0) + C'(p) \geq 0, \quad \text{for all } p \geq p_1,$$

hence, the objective is nondecreasing for all feasible x and $\phi(p) = 0$. ∎

Lemma 2. *Let $C(s)$ be defined as in (13), and assume that the stage cost $g(\cdot)$ is convex. Then*

$$\frac{dC}{ds}(s) \geq -\frac{R}{r}\mathbf{E}_W[g(W)], \quad 0 \leq s \leq \bar{s}. \tag{32}$$

Proof: By Theorem 2(ii), the optimal cost function $C(s)$ is convex. Hence,

$$\frac{dC}{ds}(s) \geq \frac{dC}{ds}(0).$$

On the other hand, by Theorem 2(iii), we can write

$$\begin{aligned}\frac{dC}{ds}(0) &= \frac{R+\theta}{r}C(0) \\ &- \frac{R}{r}\mathbf{E}_W\Big[\min_{u=0} g(W-0) + C(0)\Big].\end{aligned}$$

Combining the two preceding relations proves the claim. ∎

Lemma 3. *If Assumption 1 holds, then the constraint $x \leq \min\{B, \bar{s}-p\}$ in (23) is never active, i.e., $\phi(p) < \min\{B, \bar{s}-p\}$.*

Proof: We show that under Assumption 1, the slope of the objective function is always non-negative at $x = \min\{B, \bar{s}-p\}$. First, consider the case where $\bar{s} - p \leq B$. We have

$$\frac{\partial}{\partial x}\Big(g(x) + C(x+p)\Big)\Big|_{x=\bar{s}-p} = g'(\bar{s}-p) + C'(\bar{s}) \geq 0,$$

where the inequality follows from monotonicity of g and (19). For the case where $\bar{s} - p \geq B$, we employ Lemma 2 and Assumption 1 to write

$$\begin{aligned}\frac{\partial}{\partial x}\Big(g(x) + C(x+p)\Big)\Big|_{x=B} &= g'(B) + C'(B+p) \\ &\geq g'(B) - \frac{R}{r}\mathbf{E}_W[g(W)] \\ &\geq g'(B) - \frac{\mathbf{E}_W[g(W)]}{\mathbf{E}[W]} \geq 0,\end{aligned}$$

where the last inequality holds because $g(w) \leq wg'(B)$, for all $w \leq B$, which is a convexity result. ∎

Proof of Theorem 5: Under the differentiability assumptions, we may use Theorem 2 to characterize the optimal policy as

$$\begin{aligned}\mu^*(s,w) &= \text{argmin} \quad g(w-u) + C(s-u) \\ &\quad \text{s.t.} \\ &\quad 0 \leq u \leq \min\{s, w\}.\end{aligned} \tag{33}$$

Note that the optimization problem in (33) is convex, because $g(\cdot)$ and hence, $C(\cdot)$ is convex (cf. Theorem 2(ii)). Using the change of variables

$$x = w - u, \quad p = s - w,$$

we can rewrite (33) as $\mu^*(s,w) = w - x^*(p,w)$, where

$$\begin{aligned}x^*(p,w) &= \text{argmin} \quad g(x) + C(p+x) \\ &\quad \text{s.t.} \\ &\quad x \geq \max\{0, -p\} \\ &\quad x \leq w.\end{aligned} \tag{34}$$

The optimization problem in (34) depends on both parameters p and w. Since this problem is convex, we may remove the dependency on w as follows. Since $w \leq B, \bar{s} - p$, we may relax the last constraint, $x \leq w$, by replacing it with

$$x \leq \min\{B, \bar{s} - p\}$$

The optimal solution of the relaxed problem is the same as $\phi(p)$ defined in (23). If $\phi(p) < w$, then the relaxed constraint is not active, and $\phi(p)$ is also the solution of (34). Otherwise, since we have a convex problem, the constraint $x \leq w$ must be active, which uniquely identifies the optimal solution as w. Therefore, the optimal solution of the problem in (34) is given by $x^*(p,w) = \min\{\phi(p), w\}$. Combining the preceding relations, we obtain

$$\mu^*(s,w) = w - \min\{\phi(s-w), w\} = \Big[w - \phi(s-w)\Big]^+.$$

The representation in (24) is a direct consequence of Lemmas 1 and 3. Between some break-points b_0 and b_1, the optimal solution of (23) can only be an interior solution, which is given by (25). The uniqueness of $\phi^\circ(p)$ follows from strict convexity of g. Finally, by continuous differentiability of the cost function, equation (25) should hold at the break-points as well. Therefore,

$$g'(b_0) + C'(b_0 + (-b_0)) = 0, \quad g'(0) + C'(0 + b_1) = 0 =,$$

which is equivalent to the characterizations in (26) and (27). The first inequality in (26) holds by Lemma 32 and convexity of $g(\cdot)$, and the second inequality holds by Assumption 1 and applying convexity of $g(\cdot)$ again. ∎

Lemma 4. *Let $\phi(p)$ be defined as in (23), and assume that Assumption 1 holds and the stage cost $g(\cdot)$ is strictly convex. Then for all $p_1 \leq p_2$,*

$$-(p_2 - p_1) \leq \phi(p_2) - \phi(p_1) \leq 0. \tag{35}$$

Proof: We first establish the monotonicity of $\phi(p)$. Let $p_1 < p_2$. Given the structure of the kernel function in (24), there are multiple cases to consider, for most of which the claim is immediate using (24). We only present the case where $-B \leq p_1 \leq b_1$ and $b_0 \leq p_2 \leq b_1$. A necessary optimality condition at p_1 is given by

$$g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \geq 0. \tag{36}$$

Similarly, for p_2, we must have

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) = 0, \tag{37}$$

Now, assume $\phi(p_2) > \phi(p_1)$. By convexity of $C(\cdot)$ (cf. Theorem 2(ii)) and strict convexity of $g(\cdot)$, we obtain

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) > g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \geq 0,$$

which is a contradiction to (37).

For the second part of the claim, again, we should consider several cases depending on the interval to which p_1 and p_2 belong. Here, we present the case where $b_0 \leq p_1 \leq b_2$ and $b_0 \leq p_2 \leq \bar{s}$. The remaining cases are straightforward using (24). In this case, we have

$$g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) = 0, \tag{38}$$

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) \geq 0. \tag{39}$$

Combine the optimality conditions in (38) and (39) to get

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) \geq g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \tag{40}$$

Assume $\phi(p_2) < \phi(p_1)$; otherwise, the claim is trivial. By strict convexity of $g(\cdot)$, we have $g'(\phi(p_2)) < g'(\phi(p_1))$. There by (40), it is true that

$$C'(p_2 + \phi(p_2)) > C'(p_1 + \phi(p_1)). \tag{41}$$

On the other hand, assume the claim is not true, i.e., $\phi(p_2) - \phi(p_1) < -(p_2 - p_1)$. By rearranging the terms of this inequality and invoking the convexity of $C(\cdot)$, we get

$$C'(p_2 + \phi(p_2)) \leq C'(p_1 + \phi(p_1)),$$

which is in contradiction to (41). Therefore, the claim holds. ■

Proof of Theorem 4: By Theorem 5 we may represent the optimal cost in terms of the kernel function $\phi(\cdot)$. For any w and $s_1 \leq s_2$, by Lemma 4, we have

$$\phi(s_2 - w) \leq \phi(s_1 - w),$$

which implies

$$\mu^*(s_2, w) = \big[w - \phi(s_2 - w)\big]^+ \geq \big[w - \phi(s_1 - w)\big]^+ = \mu^*(s_1, w).$$

Moreover, for all s and $w_1 \leq w_2$, we can use the second part of Lemma 4 to conclude

$$\phi(s - w_1) - \phi(s - w_2) \geq -(w_2 - w_1).$$

By rearranging the terms, it follows that

$$\mu^*(s, w_2) = \big[w_2 - \phi(s - w_2)\big]^+ \geq \big[w_1 - \phi(s - w_1)\big]^+ = \mu^*(s, w_1).$$

■

REFERENCES

[1] L.M. Benveniste and J.A. Scheinkman. On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47(3), 2010.

[2] D. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1,2. Athena Scientific, 3rd edition, 2007.

[3] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman. Complex dynamics of blackouts in power transmission systems. *Chaos*, 14(3):643–652, 2004.

[4] M. Chen, I.-K. Cho, and S.P. Meyn. Reliability by design in a distributed power transmission network. *Automatica*, 42:1267–1281, August 2006. (invited).

[5] I.-K. Cho and S. P. Meyn. A dynamic newsboy model for optimal reserve management in electricity markets. Submitted for publication, SIAM J. Control and Optimization., 2009.

[6] Ian Dobson, Benjamin A. Carreras, Vickie E. Lynch, and David E. Newman. Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization. In *Chaos*, volume 17, 2007.

[7] R. J. Elliott. *Stochastic Calculus and Applications*. SpringerVerlag, 1982.

[8] W.H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2nd edition, 2005.

[9] Sean Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, New York, NY, USA, 1st edition, 2007.