## On the Distributed Energy Storage Investment and Operations
## Online Appendices

## A. Proofs

**Proof of Lemma 1:** (i) We prove by induction. The statement holds for $V_{T+1}(\cdot) = 0$. For a given $t \in \mathcal{T}$, suppose $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})$ is decreasing and convex in $(\mathbf{s}_{t+1}, \mathbf{S})$ for any $\mathbf{d}_{t+1}$.

The constraint $q(\mathbf{s}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) \geq 0$ in (6) defines a non-convex feasible set, which is difficult for analysis. We consider relaxing (6) and show that this relaxation does not affect the value function:

$$V_t(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S}) = \min_{\mathbf{s}_{t+1} \in \mathcal{A}} f_t(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S}), \tag{A.1}$$

where $f_t(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S}) \stackrel{\text{def}}{=} c\big([q(\mathbf{s}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) - r_t]^+\big) + \gamma \mathbb{E}_t\big[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})\big]$.

To prove (A.1), suppose there exists $\widehat{\mathbf{s}}_{t+1}$ such that $q(\widehat{\mathbf{s}}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) < 0$. Because $q(\mathbf{S} - \mathbf{s}_t, \mathbf{d}_t) \geq 0$ and $q(\cdot, \cdot)$ is a continuous function, we can apply the intermediate value theorem and find $\widetilde{\mathbf{s}}_{t+1}$ such that $\widehat{\mathbf{s}}_{t+1} \leq \widetilde{\mathbf{s}}_{t+1} \leq \mathbf{S}$ and $q(\widetilde{\mathbf{s}}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) = 0$. The objective value at $\widetilde{\mathbf{s}}_{t+1}$ is lower than at $\widehat{\mathbf{s}}_{t+1}$ because

$$\begin{aligned} f_t(\widehat{\mathbf{s}}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S}) &= c(0) + \gamma \mathbb{E}_t\big[V_{t+1}(\widehat{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})\big] \\ &\geq c(0) + \gamma \mathbb{E}_t\big[V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})\big] = f_t(\widetilde{\mathbf{s}}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S}), \end{aligned}$$

where the inequality follows from the induction hypothesis that $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})$ decreases in $\mathbf{s}_{t+1}$. Therefore, in (A.1), when minimizing $f_t(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ over $\mathbf{s}_{t+1} \in \mathcal{A}$, we can restrict our attention to the states satisfying $q(\mathbf{s}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) \geq 0$, which is equivalent to the original problem (5)-(6).

To prove the monotonicity of the value function, for any given $(\mathbf{s}_t, \mathbf{d}_t, \mathbf{S})$, let $\mathbf{s}^*_{t+1}$ be an optimal decision found by (A.1). For any $(\widetilde{\mathbf{s}}_t, \widetilde{\mathbf{S}}) \geq (\mathbf{s}_t, \mathbf{S})$, we have

$$V_t(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S}) = f_t(\mathbf{s}^*_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S}) \geq f_t(\mathbf{s}^*_{t+1}, \widetilde{\mathbf{s}}_t, \mathbf{d}_t; \widetilde{\mathbf{S}}) \geq V_t(\widetilde{\mathbf{s}}_t, \mathbf{d}_t; \widetilde{\mathbf{S}}),$$

where the first inequality follows from $c\big([q(\Delta\mathbf{s}, \mathbf{d}_t) - r_t]^+\big)$ increasing in $\Delta\mathbf{s}$ and the induction hypothesis that $V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1}; \mathbf{S}) \geq V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1}; \widetilde{\mathbf{S}})$. Thus, $V_t(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ decreases in $(\mathbf{s}_t, \mathbf{S})$.

To prove the convexity of $V_t(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ in $(\mathbf{s}_t, \mathbf{S})$, note that $c\big([q(\Delta\mathbf{s}, \mathbf{d}_t) - r_t]^+\big)$ is convex in $\Delta\mathbf{s}$ due to the composition of convex increasing functions, and $\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}; \mathbf{S})\big]$ is convex in $(\mathbf{s}_{t+1}, \mathbf{S})$ by the induction hypothesis. Therefore, $f_t(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ is jointly convex in $(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{S})$ on a closed convex set: $\{(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{S}) : \mathbf{0} \leq \mathbf{s}_{t+1} \leq \mathbf{S}, \mathbf{0} \leq \mathbf{s}_t \leq \mathbf{S}, \mathbf{0} \leq \mathbf{S} \leq \overline{\mathbf{S}}\}$, where $\overline{\mathbf{S}}$ is some large vector. Using the theorem on convexity preservation under minimization from Heyman and Sobel (1984, p. 525), we conclude that $V_t(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ as minimized in (A.1) is convex in $(\mathbf{s}_t, \mathbf{S})$.

(ii) Part (i) implies that $V_1(\mathbf{S}, \mathbf{d}_1; \mathbf{S})$ is decreasing and convex in $\mathbf{S}$. Thus, $V(\mathbf{S}) = \mathbb{E}V_1(\mathbf{S}, \mathbf{d}_1; \mathbf{S})$ is decreasing and convex in $\mathbf{S}$. ∎

**Proof of Lemma 2:** (i) We prove by induction. The statement holds for $V_{T+1}(\cdot) = 0$. Suppose the statement holds in period $t+1$. For period $t$, we consider two states $(\mathbf{s}_t, \mathbf{d}_t)$ and $(\widetilde{\mathbf{s}}_t, \mathbf{d}_t)$ that satisfy the conditions in part (i). Let $\mathbf{s}^*_{t+1}$ be the optimal decision for state $(\mathbf{s}_t, \mathbf{d}_t)$. Denote $\Delta \mathbf{s}^*_t = \mathbf{s}^*_{t+1} - \mathbf{s}_t$ and $q^*_t = q(\Delta \mathbf{s}^*_t, \mathbf{d}_t) - r_t$. We construct a feasible decision for state $(\widetilde{\mathbf{s}}_t, \mathbf{d}_t)$. Consider two cases:

*Case 1:* $\widetilde{\mathbf{s}}_t + \Delta \mathbf{s}^*_t \in \mathcal{A}$. In this case, a feasible decision for state $(\widetilde{\mathbf{s}}_t, \mathbf{d}_t)$ is to produce $[q^*_t]^+$ and change inventory to $\widetilde{\mathbf{s}}_{t+1} = \widetilde{\mathbf{s}}_t + \Delta \mathbf{s}^*_t$. Then, $\widetilde{\mathbf{s}}_{t+1}$ and $\mathbf{s}^*_{t+1}$ are related as follows: $s^*_{j,t+1} = \widetilde{s}_{j,t+1} - \delta$, $s^*_{k,t+1} = \widetilde{s}_{k,t+1} + \beta^2 \delta$, and $s^*_{i,t+1} = \widetilde{s}_{i,t+1}$ for all $i \neq j, k$. Using the induction hypothesis, we have $V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}) \leq V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1})$ for any $\mathbf{d}_{t+1}$, which leads to

$$V_t(\widetilde{\mathbf{s}}_t, \mathbf{d}_t) \leq c([q^*_t]^+) + \gamma \mathbb{E}_t\big[V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1})\big] \leq c([q^*_t]^+) + \gamma \mathbb{E}_t\big[V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1})\big] = V_t(\mathbf{s}_t, \mathbf{d}_t).$$

*Case 2:* $\widetilde{\mathbf{s}}_t + \Delta \mathbf{s}^*_t \notin \mathcal{A}$, i.e., $s^*_{j,t+1} + \delta > S_j$ or $s^*_{k,t+1} - \beta^2 \delta < 0$ or both inequalities hold. Let $\widetilde{\delta} \equiv \min\left\{S_j - s^*_{j,t+1},\, s^*_{k,t+1}/\beta^2\right\}$. By definition, $\widetilde{\delta} \in [0, \delta)$. For state $(\widetilde{\mathbf{s}}_t, \mathbf{d}_t)$, construct a feasible inventory decision $\widetilde{\mathbf{s}}_{t+1} \in \mathcal{A}$ satisfying $\widetilde{s}_{j,t+1} = s^*_{j,t+1} + \widetilde{\delta}$, $\widetilde{s}_{k,t+1} = s^*_{k,t+1} - \beta^2 \widetilde{\delta}$, and $\widetilde{s}_{i,t+1} = s^*_{i,t+1}$ for all $i \neq j, k$. Then, the induction hypothesis implies $V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}) \leq V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1})$ for any $\mathbf{d}_{t+1}$. Define $\Delta \widetilde{\mathbf{s}}_t = \widetilde{\mathbf{s}}_{t+1} - \widetilde{\mathbf{s}}_t$ and $\widetilde{q}_t = q(\Delta \widetilde{\mathbf{s}}_t, \mathbf{d}_t) - r_t$. If we can show $\widetilde{q}_t \leq q^*_t$, we have the intended result:

$$V_t(\widetilde{\mathbf{s}}_t, \mathbf{d}_t) \leq c([\widetilde{q}_t]^+) + \gamma \mathbb{E}_t\big[V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1})\big] \leq c([q^*_t]^+) + \gamma \mathbb{E}_t\big[V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1})\big] = V_t(\mathbf{s}_t, \mathbf{d}_t), \quad \text{(A.2)}$$

where we used the relation $V_t(\mathbf{s}_t, \mathbf{d}_t) = \min_{\mathbf{s}_{t+1} \in \mathcal{A}} f_t(\mathbf{s}_{t+1}, \mathbf{s}_t, \mathbf{d}_t)$ given in (A.1).

The rest of the proof shows $\widetilde{q}_t \leq q^*_t$. The choice of $\widetilde{\delta}$ gives $\widetilde{s}_{j,t+1} = S_j$ or $\widetilde{s}_{k,t+1} = 0$, which implies

$$\Delta \widetilde{s}_{j,t} = \widetilde{s}_{j,t+1} - \widetilde{s}_{j,t} \geq 0 \quad \text{or} \quad \Delta \widetilde{s}_{k,t} = \widetilde{s}_{k,t+1} - \widetilde{s}_{k,t} \leq 0. \quad \text{(A.3)}$$

Let $\varepsilon = \delta - \widetilde{\delta} > 0$. Then, by definitions, we have $\Delta s^*_{j,t} = \Delta \widetilde{s}_{j,t} + \varepsilon$, $\Delta s^*_{k,t} = \Delta \widetilde{s}_{k,t} - \beta^2 \varepsilon$, and $\Delta s^*_{0,t} = \Delta \widetilde{s}_{0,t}$. Using the definition in (4), we have

$$\begin{aligned} q^*_t - \widetilde{q}_t &= \psi_\beta\big(d_{j,t} + \psi_\alpha(\Delta s^*_{j,t})\big) - \psi_\beta\big(d_{j,t} + \psi_\alpha(\Delta \widetilde{s}_{j,t})\big) - \big[\psi_\beta\big(d_{k,t} + \psi_\alpha(\Delta \widetilde{s}_{k,t})\big) - \psi_\beta\big(d_{k,t} + \psi_\alpha(\Delta s^*_{k,t})\big)\big] \\ &\geq \beta\big[\psi_\alpha(\Delta \widetilde{s}_{j,t} + \varepsilon) - \psi_\alpha(\Delta \widetilde{s}_{j,t})\big] - \beta^{-1}\big[\psi_\alpha(\Delta \widetilde{s}_{k,t}) - \psi_\alpha(\Delta \widetilde{s}_{k,t} - \beta^2 \varepsilon)\big] \equiv \Gamma, \end{aligned} \quad \text{(A.4)}$$

where the inequality is because $\psi_\beta(u)$ increases in $u$ with a slope of either $\beta$ or $\beta^{-1}$. Now consider the cases under the two conditions derived in (A.3):

- If $\Delta \widetilde{s}_{j,t} \geq 0$, then $\Gamma = \beta \alpha^{-1} \varepsilon - \beta^{-1}\big[\psi_\alpha(\Delta \widetilde{s}_{k,t}) - \psi_\alpha(\Delta \widetilde{s}_{k,t} - \beta^2 \varepsilon)\big] \geq \beta \alpha^{-1} \varepsilon - \beta^{-1} \alpha^{-1} \beta^2 \varepsilon = 0.$
- If $\Delta \widetilde{s}_{k,t} \leq 0$, then $\Gamma = \beta\big[\psi_\alpha(\Delta \widetilde{s}_{j,t} + \varepsilon) - \psi_\alpha(\Delta \widetilde{s}_{j,t})\big] - \beta^{-1} \alpha \beta^2 \varepsilon \geq \beta \alpha \varepsilon - \beta \alpha \varepsilon = 0.$

Hence, $\widetilde{q}_t \leq q^*_t$ and the result in (A.2) holds.

(ii) For the case of $k = 0$, the proof follows the same lines as in part (i), except that $s_{0,t}$ exceeds $\widetilde{s}_{0,t}$ by $\beta \delta$ instead of $\beta^2 \delta$. The case of $j = 0$ can be proved similarly. ∎

**Proof of Lemma 3:** In (5)-(6), for a given $\mathbf{S}$, we express an optimal decision rule in period $t$ as $\mathbf{s}^*_{t+1}(\mathbf{s}_t, \mathbf{d}_t; \mathbf{S})$ and abbreviate it as $\mathbf{s}^*_{t+1}$. Let $\{\mathbf{s}^*_{t+1} : t \in \mathcal{T}\}$ denote an optimal policy, and let

$u^*_{i,t} = d_{i,t} + \psi_\alpha(\Delta s^*_{i,t})$ be the corresponding energy flow according to (3).

(i) Under the centralized investment $\mathbf{S}^c = \big(S_0 + \beta^{-1}\sum_{i\in\mathcal{L}} S_i,\ 0,\ \ldots,\ 0\big)$, we construct a policy $\{\widetilde{\mathbf{s}}_{t+1} = (\widetilde{s}_{0,t+1},\, 0, \ldots, 0) : t \in \mathcal{T}\}$ such that

$$\widetilde{s}_{0,t+1} = s^*_{0,t+1} + \beta^{-1} g_{t+1}, \qquad \forall\, t \in \mathcal{T}, \tag{A.5}$$

$$g_1 = \sum_{i\in\mathcal{L}} S_i, \tag{A.6}$$

$$g_{t+1} = \min\Big\{ \sum_{i\in\mathcal{L}} S_i,\ \ g_t + \big(\sum_{i\in\mathcal{L}} \Delta s^*_{i,t}\big) + \tfrac{1-\beta^2}{\alpha} \sum_{i\in\mathcal{L}} (-u^*_{i,t})^+ \Big\}. \tag{A.7}$$

The definition in (A.7) implies $g_{t+1} \in \big[\sum_{i\in\mathcal{L}} s^*_{i,t+1}, \sum_{i\in\mathcal{L}} S_i\big]$ for all $t \in \mathcal{T}$.[8] Hence, $0 \le \widetilde{s}_{0,t+1} \le S_0 + \beta^{-1}\sum_{i\in\mathcal{L}} S_i$, that is, the constructed policy $\{\widetilde{\mathbf{s}}_{t+1} : t \in \mathcal{T}\}$ is feasible under investment $\mathbf{S}^c$.[9] Next, we prove $q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t) \ \le\ q(\Delta\mathbf{s}^*_t, \mathbf{d}_t)$. Consider two cases:

1) Case of $u^*_{i,t} \ge 0$ for all $i \in \mathcal{L}$. In this case, $\psi_\beta(u^*_{i,t}) = \beta^{-1} u^*_{i,t}$. Then,

$$\begin{aligned} q(\Delta\mathbf{s}^*_t, \mathbf{d}_t) &= \psi_\alpha(\Delta s^*_{0,t}) + \sum_{i\in\mathcal{L}} \beta^{-1}(d_{i,t} + \psi_\alpha(\Delta s^*_{i,t})) = \psi_\alpha(\Delta s^*_{0,t}) + \sum_{i\in\mathcal{L}} \big[\beta^{-1} d_{i,t} + \psi_\alpha(\beta^{-1}\Delta s^*_{i,t})\big] \\ &\ge \psi_\alpha\big(\Delta s^*_{0,t} + \sum_{i\in\mathcal{L}} \beta^{-1}\Delta s^*_{i,t}\big) + \beta^{-1}\sum_{i\in\mathcal{L}} d_{i,t} \\ &\ge \psi_\alpha(\Delta s^*_{0,t} + \beta^{-1}\Delta g_t) + \beta^{-1}\sum_{i\in\mathcal{L}} d_{i,t} \ = \ \psi_\alpha(\Delta\widetilde{s}_{0,t}) + \beta^{-1}\sum_{i\in\mathcal{L}} d_{i,t} \ = \ q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t), \end{aligned}$$

where the first inequality utilizes the subadditivity of $\psi_\alpha(\cdot)$, i.e., $\psi_\alpha(x) + \psi_\alpha(y) \ge \psi_\alpha(x+y)$, and the second inequality is because (A.7) with $u^*_{i,t} \ge 0$ implies that $\Delta g_t \equiv g_{t+1} - g_t \le \sum_{i\in\mathcal{L}} \Delta s^*_{i,t}$.

2) Case of $u^*_{j,t} < 0$ for $j \in \mathcal{L}^- \subset \mathcal{L}$, i.e., energy is sent from the nodes in set $\mathcal{L}^-$ to other nodes. This immediately implies that $\Delta\mathbf{s}^*_t \le 0$ because Lemma 2 states that energy should not be released from one node only to store it in another node. (Formally, if $u_{j,t} < 0$, $\Delta s_{j,t} < 0$, and $\Delta s_{k,t} > 0$ for some $j, k \in \mathcal{L}$, then an alternative decision $\Delta\widetilde{s}_{j,t} = \Delta s_{j,t} + \delta$ and $\Delta\widetilde{s}_{k,t} = \Delta s_{k,t} - \beta^2\delta$ with $\delta = \min\{-\Delta s_{j,t}, \Delta s_{k,t}/\beta^2\}$ would lead to the same production but a lower expected future cost according to Lemma 2.) These conditions imply that the last two terms in (A.7) are

$$\begin{aligned} &\big(\sum_{i\in\mathcal{L}} \Delta s^*_{i,t}\big) - \tfrac{1-\beta^2}{\alpha} \sum_{j\in\mathcal{L}^-} (d_{j,t} + \alpha\Delta s^*_{j,t}) \\ &= \big(\sum_{k\in\mathcal{L}\setminus\mathcal{L}^-} \Delta s^*_{k,t}\big) + \beta^2\big(\sum_{j\in\mathcal{L}^-} \Delta s^*_{j,t}\big) - \tfrac{1-\beta^2}{\alpha} \sum_{j\in\mathcal{L}^-} d_{j,t} < 0. \end{aligned}$$

Thus, $\Delta g_t = g_{t+1} - g_t = \big(\sum_{k\in\mathcal{L}\setminus\mathcal{L}^-} \Delta s^*_{k,t}\big) + \beta^2\big(\sum_{j\in\mathcal{L}^-} \Delta s^*_{j,t}\big) - \tfrac{1-\beta^2}{\alpha} \sum_{j\in\mathcal{L}^-} d_{j,t} < 0.$

[8] We show $g_{t+1} \ge \sum_{i\in\mathcal{L}} s^*_{i,t+1}$ by induction. This is true in (A.6). Suppose $g_t \ge \sum_{i\in\mathcal{L}} s^*_{i,t}$ for some $t < T$. Then, $g_t + \sum_{i\in\mathcal{L}} \Delta s^*_{i,t} \ge \sum_{i\in\mathcal{L}} (s^*_{i,t} + \Delta s^*_{i,t}) = \sum_{i\in\mathcal{L}} s^*_{i,t+1}$. Since the last term in (A.7) is non-negative, we have $g_{t+1} \ge \sum_{i\in\mathcal{L}} s^*_{i,t+1}$.

[9] We do not require $\widetilde{\mathbf{s}}_t$ to satisfy constraint (6); see the proof of Lemma 1 for constraint relaxation.

Then,

$$q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t) = \alpha(\Delta s^*_{0,t} + \beta^{-1}\Delta g_t) + \beta^{-1}\sum_{i\in\mathcal{L}} d_{i,t}$$
$$= \alpha\Delta s^*_{0,t} + \beta\sum_{j\in\mathcal{L}^-}(\alpha\Delta s^*_{j,t} + d_{j,t}) + \beta^{-1}\sum_{k\in\mathcal{L}\setminus\mathcal{L}^-}(\alpha\Delta s^*_{k,t} + d_{k,t}) \;=\; q(\Delta\mathbf{s}^*_t, \mathbf{d}_t).$$

Note that $u^*_{j,t} < 0$ for all $j \in \mathcal{L}$ is not possible because reverse flows on all lines are suboptimal by Lemma 2. Therefore, in all cases, we have $q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t) \le q(\Delta\mathbf{s}^*_t, \mathbf{d}_t)$, implying that the policy $\{\widetilde{\mathbf{s}}_{t+1} : t \in \mathcal{T}\}$ achieves an operating cost no higher than $V(\mathbf{S})$. Therefore, $V(\mathbf{S}^c) \le V(\mathbf{S})$.

(ii) Under $\mathbf{S}^l = (0,\, S_1 + \beta S_0,\, \ldots,\, S_n + \beta S_0)$, we construct a policy $\{\widehat{\mathbf{s}}_{t+1} : t \in \mathcal{T}\}$:

$$\widehat{s}_{0,t+1} = 0, \qquad \widehat{s}_{j,t+1} = s^*_{j,t+1} + \beta g_{j,t+1}, \quad j \in \mathcal{L}, \qquad \forall\, t \in \mathcal{T},$$
$$g_{j,1} = S_0, \qquad j \in \mathcal{L},$$
$$\Delta g_{j,t} = g_{j,t+1} - g_{j,t} = \begin{cases} \max\left\{\Delta s^*_{0,t} - \sum_{i\in\mathcal{L},\, i<j}\Delta g_{i,t},\ -u^{*+}_{j,t}/(\alpha\beta)\right\}, & \text{if } \Delta s^*_{0,t} < 0, \\ \min\left\{\Delta s^*_{0,t} - \sum_{i\in\mathcal{L},\, i<j}\Delta g_{i,t},\ S_0 - g_{j,t}\right\}, & \text{if } \Delta s^*_{0,t} \ge 0. \end{cases}$$

Using techniques similar to part (i), we can prove $V(\mathbf{S}^l) \le V(\mathbf{S})$. ∎

The proof of Proposition 1 requires some properties of the optimal operating policy and the value function when $d^{\min}_j > 0$, as stated in the following lemma.

**Lemma A.1** *Suppose $d^{\min}_j > 0$ for a given demand node $j \in \mathcal{L}$. If storage investment $\mathbf{S}$ satisfies $\alpha S_j \le d^{\min}_j$, then,*

*(i) There exists an optimal policy satisfying $\Delta s^*_{0,t} \cdot \Delta s^*_{j,t} \ge 0$ for all $t \in \mathcal{T}$;*

*(ii) In period $t$, suppose $\mathbf{s}_t$, $\widetilde{\mathbf{s}}_t \in \mathcal{A}$ satisfy $\widetilde{s}_{0,t} = s_{0,t} - \delta$ and $\widetilde{s}_{j,t} = s_{j,t} + \beta\delta$ for some $\delta > 0$, and $\widetilde{s}_{i,t} = s_{i,t}$ for all $i \in \mathcal{L}$, $i \ne j$, then $V_t(\mathbf{s}_t, \mathbf{d}_t) = V_t(\widetilde{\mathbf{s}}_t, \mathbf{d}_t)$ for any $\mathbf{d}_t$.*

**Proof of Lemma A.1:** The condition $\alpha S_j \le d^{\min}_j$ means that there is never surplus energy that can be sent from node $j$ back to node 0.

We prove by induction. Suppose part (ii) holds for period $t+1$ (it clearly holds for period $T+1$). In period $t$, we consider any given state $(\mathbf{s}, \mathbf{d})$ and any decision $\mathbf{s}_{t+1}$ with inventory change $\Delta\mathbf{s} \equiv \mathbf{s}_{t+1} - \mathbf{s}$ satisfying $\Delta s_0 > 0$ and $\Delta s_j < 0$. Set $\delta = \min\{\Delta s_0,\, -\beta^{-1}\Delta s_j\} > 0$. We now show that a strictly better decision is $\widehat{\mathbf{s}}_{t+1}$ with $\widehat{s}_{0,t+1} = s_{0,t+1} - \delta$, $\widehat{s}_{j,t+1} = s_{j,t+1} + \beta\delta$, and $\widehat{s}_{i,t+1} = s_{i,t+1}$ for $i \ne j$. This new decision satisfies $\Delta\widehat{s}_0 = \Delta s_0 - \delta \ge 0$, $\Delta\widehat{s}_j = \Delta s_j + \beta\delta \le 0$, and $\Delta\widehat{s}_0 \cdot \Delta\widehat{s}_j = 0$. To show the superiority of $\widehat{\mathbf{s}}_{t+1}$, note that $V_{t+1}(\widehat{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}) = V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ by the induction hypothesis and

$$q(\Delta\widehat{\mathbf{s}}, \mathbf{d}) - q(\Delta\mathbf{s}, \mathbf{d}) = \beta^{-1}(d_j + \alpha\Delta\widehat{s}_j) + \alpha^{-1}\Delta\widehat{s}_0 - \beta^{-1}(d_j + \alpha\Delta s_j) - \alpha^{-1}\Delta s_0 = \alpha\delta - \alpha^{-1}\delta < 0.$$

Similarly, any decision $\mathbf{s}_{t+1}$ with $\Delta s_0 < 0$ and $\Delta s_j > 0$ is also suboptimal. The optimal decision must satisfy $\Delta s_0 \cdot \Delta s_j \geq 0$. Thus, part (i) holds for period $t$. We next prove part (ii) for period $t$.

Consider states $(\mathbf{s}, \mathbf{d})$ and $(\widetilde{\mathbf{s}}, \mathbf{d})$ in period $t$, with $\widetilde{s}_0 = s_0 - \delta$, $\widetilde{s}_j = s_j + \beta\delta$ for some $\delta > 0$, and $\widetilde{s}_i = s_i$ for all $i \in \mathcal{L}$, $i \neq j$. Lemma 2 implies that $V_t(\mathbf{s}, \mathbf{d}) \leq V_t(\widetilde{\mathbf{s}}, \mathbf{d})$. Thus, we only need to show $V_t(\widetilde{\mathbf{s}}, \mathbf{d}) \leq V_t(\mathbf{s}, \mathbf{d})$. Let $\mathbf{s}^*_{t+1}$ be the optimal decision for $(\mathbf{s}, \mathbf{d})$ and denote $\Delta\mathbf{s}^* = \mathbf{s}^*_{t+1} - \mathbf{s}$. For state $(\widetilde{\mathbf{s}}, \mathbf{d})$, we construct a decision $\widetilde{\mathbf{s}}_{t+1}$ satisfying $\widetilde{s}_{0,t+1} = s^*_{0,t+1} - \widetilde{\delta}$, $\widetilde{s}_{j,t+1} = s^*_{j,t+1} + \beta\widetilde{\delta}$, with $\widetilde{\delta} = \min\left\{\delta,\, s^*_{0,t+1},\, \beta^{-1}(S_j - s^*_{j,t+1})\right\}$, and $\widetilde{s}_{i,t+1} = s^*_{i,t+1}$ for all $i \in \mathcal{L}$, $i \neq j$. We next show that $\widetilde{\mathbf{s}}_{t+1}$ for $(\widetilde{\mathbf{s}}, \mathbf{d})$ gives the same production cost as $\mathbf{s}^*_{t+1}$ for $(\mathbf{s}, \mathbf{d})$. Let $\Delta\widetilde{\mathbf{s}} = \widetilde{\mathbf{s}}_{t+1} - \widetilde{\mathbf{s}} = \Delta\mathbf{s}^* - (-\varepsilon, 0, \ldots, 0, \beta\varepsilon, 0, \ldots, 0)$, where $\varepsilon = \delta - \widetilde{\delta}$. Consider two cases:

- Case 1: $\Delta s^*_0 \geq 0$ and $\Delta s^*_j \geq 0$. We have $s^*_{0,t+1} \geq s_0 = \widetilde{s}_0 + \delta \geq \delta$. Thus, either $\widetilde{\delta} = \delta$ or $\widetilde{\delta} = \beta^{-1}(S_j - s^*_{j,t+1})$. In either case, we can verify that $\Delta\widetilde{s}_j \geq 0$ and $\Delta\widetilde{s}_0 \geq 0$. Hence,

$$\begin{aligned} q(\Delta\widetilde{\mathbf{s}}, \mathbf{d}) - q(\Delta\mathbf{s}^*, \mathbf{d}) &= \beta^{-1}(d_j + \alpha^{-1}\Delta\widetilde{s}_j) + \alpha^{-1}\Delta\widetilde{s}_0 - \beta^{-1}(d_j + \alpha^{-1}\Delta s^*_j) - \alpha^{-1}\Delta s^*_0 \\ &= -\beta^{-1}\alpha^{-1}\beta\varepsilon + \alpha^{-1}\varepsilon \;=\; 0. \end{aligned} \tag{A.8}$$

- Case 2: $\Delta s^*_0 \leq 0$ and $\Delta s^*_j \leq 0$. Using similar logic, we can show $\Delta\widetilde{s}_0 \leq 0$, $\Delta\widetilde{s}_j \leq 0$, and $q(\Delta\widetilde{\mathbf{s}}, \mathbf{d}) = q(\Delta\mathbf{s}^*, \mathbf{d})$.

These are the only cases we need to consider, as indicated by part (i).

By the induction hypothesis, $V_{t+1}(\mathbf{s}^*_{t+1}, \mathbf{d}_{t+1}) = V_{t+1}(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1})$. Therefore, $\widetilde{\mathbf{s}}_{t+1}$ for $(\widetilde{\mathbf{s}}, \mathbf{d})$ gives the same production cost and same future expected cost as $\mathbf{s}^*_{t+1}$ for $(\mathbf{s}, \mathbf{d})$, implying that $V_t(\widetilde{\mathbf{s}}, \mathbf{d}) \leq V_t(\mathbf{s}, \mathbf{d})$. This, together with Lemma 2, implies that $V_t(\mathbf{s}, \mathbf{d}) = V_t(\widetilde{\mathbf{s}}, \mathbf{d})$. ∎

**Proof of Proposition 1:** (i) Under investment $\mathbf{S}$, let $\{\mathbf{s}^*_{t+1} : t \in \mathcal{T}\}$ be an optimal policy satisfying $\Delta s^*_{0,t} \cdot \Delta s^*_{j,t} \geq 0$, which follows from Lemma A.1(i). Under investment $\widetilde{\mathbf{S}}$ given in Proposition 1(i), we define $\delta_t = \min\{\delta,\, s^*_{0,t}\}$ and construct a policy $\{\widetilde{\mathbf{s}}_{t+1} : t \in \mathcal{T}\}$ such that $\widetilde{s}_{0,t} = s^*_{0,t} - \delta_t$, $\widetilde{s}_{j,t} = s^*_{j,t} + \beta\delta_t$, and $\widetilde{s}_{i,t} = s^*_{i,t}$ for $i \in \mathcal{L}$ and $i \neq j$. The policy $\{\widetilde{\mathbf{s}}_{t+1} : t \in \mathcal{T}\}$ is feasible under $\widetilde{\mathbf{S}}$ because by construction $0 \leq \widetilde{s}_{j,t} \leq s^*_{j,t} + \beta\delta \leq S_j + \beta\delta = \widetilde{S}_j$ and $\widetilde{s}_{0,t} = \max\{s^*_{0,t} - \delta, 0\} \in [0, \widetilde{S}_0]$.

We next show that the two policies yields the same production levels. Lemma A.1(i) suggests that we only need to consider optimal policies with $\Delta s^*_{0,t} \cdot \Delta s^*_{j,t} \geq 0$. If $\Delta s^*_{j,t} \geq 0$ and $\Delta s^*_{0,t} \geq 0$, we have $\delta_{t+1} - \delta_t \in [0, \Delta s^*_{0,t}]$, which implies $\Delta\widetilde{s}_{j,t} = \Delta s^*_{j,t} + \beta(\delta_{t+1} - \delta_t) \geq 0$ and $\Delta\widetilde{s}_{0,t} = \Delta s^*_{0,t} - (\delta_{t+1} - \delta_t) \geq 0$. Then, following exactly the same logic as in (A.8), $q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t) = q(\Delta\mathbf{s}^*_t, \mathbf{d}_t)$. If $\Delta s^*_{j,t} \leq 0$ and $\Delta s^*_{0,t} \leq 0$, similar logic applies. Therefore, $q(\Delta\widetilde{\mathbf{s}}_t, \mathbf{d}_t) = q(\Delta\mathbf{s}^*_t, \mathbf{d}_t)$ for all $t \in \mathcal{T}$, and consequently the total production costs are the same for both policies. Because the constructed policy is feasible, we have $V(\widetilde{\mathbf{S}}) \leq V(\mathbf{S})$. The opposite inequality $V(\widetilde{\mathbf{S}}) \geq V(\mathbf{S})$ can be proved similarly.

(ii) Replacing $\delta$ units of storage capacity at demand node $k$ with $\delta$ units at node $j$ can be considered as replacing it first with $\beta^{-1}\delta$ units of central storage and then replacing $\beta^{-1}\delta$ units of central storage with $\delta$ units of local storage at node $j$. Then, we have

$$\begin{aligned} V(S_0,\ldots,S_j,\ldots,S_k,\ldots) &\ge V(S_0+\beta^{-1}\delta,\ldots,S_j,\ldots,S_k-\delta,\ldots) \\ &= V(S_0,\ldots,S_j+\delta,\ldots,S_k-\delta,\ldots), \end{aligned}$$

where the inequality follows from a generalization of Lemma 3(i) that replacing local storage by central storage and increasing its size by a factor of $\beta^{-1}$ reduces production cost (proof of Lemma 3(i) can be generalized), and the equality follows directly from part (i). ∎

**Proof of Proposition 2:** (i) Let $C(\mathbf{S}) \equiv p\,|\mathbf{S}| + V(\mathbf{S})$, which is convex in $\mathbf{S}$ (Lemma 1). Thus, it suffices to show that $\mathbf{S}^{l*}$ achieves a local minimum. Let $\widetilde{\mathbf{S}} \overset{\text{def}}{=} \mathbf{S}^{l*} + \boldsymbol{\delta}$, where $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_n)$ satisfies $-\mathbf{S}^{l*} \le \boldsymbol{\delta} < \frac{1}{2}(\alpha^{-1}d_j^{\min} - S_j^{l*})\mathbf{1}$. We aim to show $C(\mathbf{S}^{l*}) \le C(\widetilde{\mathbf{S}})$.

Note that $\delta_0 \in \big[0, \frac{1}{2}(\alpha^{-1}d_j^{\min} - S_j^{l*})\big)$. Define another localized investment $\widehat{\mathbf{S}}$ such that $\widehat{S}_0 = \widetilde{S}_0 - \delta_0 = 0$, $\widehat{S}_j = \widetilde{S}_j + \beta\delta_0$, and $\widehat{S}_i = \widetilde{S}_i$ for $i \in \mathcal{L}, i \ne j$. By definition, $\widehat{S}_j = S_j^{l*} + \delta_j + \beta\delta_0 < \alpha^{-1}d_j^{\min}$. Then, we have

$$\begin{aligned} C(\widetilde{\mathbf{S}}) - C(\mathbf{S}^{l*}) &= V(\widetilde{\mathbf{S}}) - V(\mathbf{S}^{l*}) + p\big(\delta_0 + \sum_{i\in\mathcal{L}} \delta_i\big) \\ &\ge V(\widehat{\mathbf{S}}) - V(\mathbf{S}^{l*}) + p\big(\beta\delta_0 + \sum_{i\in\mathcal{L}} \delta_i\big) \;=\; C(\widehat{\mathbf{S}}) - C(\mathbf{S}^{l*}) \;\ge\; 0, \end{aligned}$$

where the first inequality follows from Proposition 1(i) and $\delta_0 \ge \beta\delta_0$, and the last inequality follows from optimality of $\mathbf{S}^{l*}$ for the constrained investment problem (9). This proves the optimality of $\mathbf{S}^{l*}$. Furthermore, if $\delta_0$ is set to be positive, then $\delta_0 > \beta\delta_0$ and the first inequality holds strictly, which implies that an investment with $S_0 > 0$ is strictly dominated by $\mathbf{S}^{l*}$.

(ii) The statement in the proposition clearly holds when $d_i^{\min} = 0$ for all $i \in \mathcal{L}$. We only need to prove the case when $d_j^{\min} > 0$ for some $j \in \mathcal{L}$. We prove by contradiction. Let the optimal investment be $\mathbf{S}^*$ with $S_0^* > 0$, and suppose $S_j^* < \alpha^{-1}d_j^{\min}$. Define $\widetilde{\mathbf{S}}$ such that $\widetilde{S}_0 = S_0^* - \delta$ and $\widetilde{S}_j = S_j^* + \beta\delta$, where $\delta = \min\{S_0^*, (\alpha^{-1}d_j^{\min} - S_j^*)/2\}$. Note that $\widetilde{S}_j < \alpha^{-1}d_i^{\min}$. Then, by Proposition 1(i), we have $V(\mathbf{S}^*) = V(\widetilde{\mathbf{S}})$. Because $|\mathbf{S}^*| > |\widetilde{\mathbf{S}}|$, we have $C(\mathbf{S}^*) > C(\widetilde{\mathbf{S}})$, contradicting to the optimality of $\mathbf{S}^*$. ∎

**Proof of Proposition 3:** If $d_i^{\min} = 0$ for all $i \in \mathcal{L}$, the statements clearly hold. We focus on the case where at least one minimum demand is positive.

(i) Consider a given capacity allocation satisfying $S_0 > 0$. Since $\alpha|\mathbf{S}| \le \sum_{i\in\mathcal{L}} d_i^{\min}$, there exists some $j \in \mathcal{L}$ such that $\alpha S_j < d_j^{\min}$. Then, according to Proposition 1(i), we can achieve the same production cost by reducing $S_0$ by $\delta$ and increasing $S_j$ by $\beta\delta$ (thus leaving an extra $(1-\beta)\delta$ units of capacity), where $\delta = \min\{S_0, (\alpha^{-1}d_j^{\min} - S_j)/\beta\}$. Allocating the extra $(1-\beta)\delta$ units of capacity to any node will decrease the production cost (Lemma 1). Therefore, we can limit our attention to localized capacity allocations without losing optimality.

Consider any localized capacity allocation satisfying $\alpha S_k > d_k^{\min}$ for some $k \in \mathcal{L}$. Since $\alpha|\mathbf{S}| \leq \sum_{i\in\mathcal{L}} d_i^{\min}$, there exists some $j \in \mathcal{L}$ such that $\alpha S_j < d_j^{\min}$. Then, Proposition 1(ii) shows that we can improve the allocation by shifting some capacity from $k$ to $j$. Therefore, we do not lose optimality by restricting our attention to all the allocations satisfying $\alpha S_i \leq d_i^{\min}$ for all $i \in \mathcal{L}$.

If $\alpha|\mathbf{S}| = \sum_{i\in\mathcal{L}} d_i^{\min}$, then $\alpha S_i = d_i^{\min}$ for all $i \in \mathcal{L}$ is the unique allocation that satisfies $\alpha S_i \leq d_i^{\min}$ for all $i \in \mathcal{L}$ and is thus optimal.

If $\alpha|\mathbf{S}| < \sum_{i\in\mathcal{L}} d_i^{\min}$, consider two allocations $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ that satisfy $\widetilde{S}_j = S_j + \delta \leq \alpha^{-1} d_j^{\min}$ and $S_k = \widetilde{S}_k + \delta \leq \alpha^{-1} d_k^{\min}$ for some $\delta > 0$, and $\widetilde{S}_i = S_i \leq d_i^{\min}$ for $i \in \mathcal{L}$, $i \neq j, k$. Then, Proposition 1(ii) implies that $V(\widetilde{\mathbf{S}}) \leq V(\mathbf{S})$ and $V(\mathbf{S}) \leq V(\widetilde{\mathbf{S}})$, leading to $V(\widetilde{\mathbf{S}}) = V(\mathbf{S})$. Therefore, reallocating $\delta$ units of capacity from node $j$ to $k$ does not change the production cost. Using a finite number of such reallocation, we can convert any allocation to any other allocation with the same $|\mathbf{S}|$ and $\alpha S_i \leq d_i^{\min}$ for all $i \in \mathcal{L}$. Therefore, all allocations satisfying $\alpha S_i \leq d_i^{\min}$ for all $i \in \mathcal{L}$ are optimal.

Now, we show that the optimal production cost is equal to the production cost under a centralized investment $\big(\beta^{-1}|\mathbf{S}|, 0, \ldots, 0\big)$. Applying Lemma 1 repeatedly, we can replace $\beta^{-1} S_i$ units of central capacity by $S_i$ units of capacity at node $i$, where $\alpha S_i \leq d_i^{\min}$ for all $i \in \mathcal{L}$. These replacements maintain the same production cost. Therefore, the optimal production cost is $V\big(\beta^{-1}|\mathbf{S}|, 0, \ldots, 0\big)$.

(ii) If $\alpha|\mathbf{S}| > \sum_{i\in\mathcal{L}} d_i^{\min}$ and $\mathbf{S}$ satisfies $\alpha S_j < d_j^{\min}$ for some $j \in \mathcal{L}$, then $\mathbf{S}$ must also satisfy either $S_0 > 0$ or $\alpha S_k > d_k^{\min}$ for some $k \in \mathcal{L}$ or both. Then, Proposition 1(i) and (ii) provide ways to improve on $\mathbf{S}$, until the condition $\alpha S_j < d_j^{\min}$ no longer holds. Hence, there exists an optimal capacity allocation satisfying $\alpha S_i^* \geq d_i^{\min}$ for all $i \in \mathcal{L}$. ∎

**Proof of Proposition 4:** For any given $\mathbf{S} \geq 0$ and the associated optimal policy $\{\mathbf{s}_t^* : t \in \mathcal{T}\}$, we construct a new system with node 0 and a single demand node. The storage size and operations at node 0 remain the same as in the original system. The single demand node combines the demand and storage of all $n$ nodes in the original system: demand is $d_{Lt} = \sum\limits_{i\in\mathcal{L}} d_{i,t}$, storage size is $S_L = \sum\limits_{i\in\mathcal{L}} S_i$, and a feasible operating policy is $s_{0,t} = s_{0,t}^*$, $s_{Lt} = \sum\limits_{i\in\mathcal{L}} s_{i,t}^*$, $t \in \mathcal{T}$. Let $C(\mathbf{S}) \equiv p\,|\mathbf{S}| + V(\mathbf{S})$ denote the total cost under investment $\mathbf{S}$ in the original system, and let $\widetilde{C}(S_0, S_L)$ denote the total cost under $(S_0, S_L)$ for the new system. The subadditivity of $\psi_\alpha$ and $\psi_\beta$ implies

$$\psi_\beta\big(d_{Lt} + \psi_\alpha(\Delta s_{Lt})\big) \leq \psi_\beta\big(d_{Lt} + \sum_{i\in\mathcal{L}} \psi_\alpha(\Delta s_{i,t}^*)\big) \leq \sum_{i\in\mathcal{L}} \psi_\beta\big(d_{i,t} + \psi_\alpha(\Delta s_{i,t}^*)\big), \quad t \in \mathcal{T},$$

which in turn implies that the new system produces no more than the original system. Thus,

$$\widetilde{C}\big(S_0, \sum_{i\in\mathcal{L}} S_i\big) \leq C(S_0, S_1, \ldots, S_n). \tag{A.9}$$

Furthermore, (A.9) holds with equality if $d_{i,t} = k_i\, d_{1,t}$ and $S_i = k_i\, S_1$, for all $i \in \mathcal{L}$. This can be shown by using the optimal policy for the new system to construct a feasible policy for the original system that yields the same production cost. The construction maintains the local storage levels at the ratios $k_i$. In other words, under $d_{i,t} = k_i\, d_{1,t}$, we have

$$\widetilde{C}(S_0, S_L) = C\Big(S_0, \frac{k_1 S_L}{\sum_{i\in\mathcal{L}} k_i}, \cdots, \frac{k_n S_L}{\sum_{i\in\mathcal{L}} k_i}\Big). \tag{A.10}$$

Following the reasoning after Proposition 4 in the paper, a localized investment is optimal for the new system with only one demand node. Denote the optimal localized investment as $S_L^*$. Then, under $d_{i,t} = k_i\, d_{1,t}$, we have

$$C\Big(0, \frac{k_1 S_L^*}{\sum_{i\in\mathcal{L}} k_i}, \cdots, \frac{k_n S_L^*}{\sum_{i\in\mathcal{L}} k_i}\Big) = \widetilde{C}(0, S_L^*) \leq \widetilde{C}\big(S_0, \sum_{i\in\mathcal{L}} S_i\big) \leq C(\mathbf{S}), \tag{A.11}$$

Because $\mathbf{S}$ is arbitrary, we conclude from (A.11) that the localized investment is optimal. ∎

**Proof of Proposition 5:** Consider any $p_1$ and $p_2$ with $p_1 < p_2$. The optimality of $\mathbf{S}^*(p_1)$ suggests $p_1\,|\mathbf{S}^*(p_1)| + V(\mathbf{S}^*(p_1)) \leq p_1\,|\mathbf{S}^*(p_2)| + V(\mathbf{S}^*(p_2))$. Similarly, $p_2\,|\mathbf{S}^*(p_2)| + V(\mathbf{S}^*(p_2)) \leq p_2\,|\mathbf{S}^*(p_1)| + V(\mathbf{S}^*(p_1))$. Combining these two inequalities, we have

$$p_1(|\mathbf{S}^*(p_1)| - |\mathbf{S}^*(p_2)|) \leq V(\mathbf{S}^*(p_2)) - V(\mathbf{S}^*(p_1)) \leq p_2(|\mathbf{S}^*(p_1)| - |\mathbf{S}^*(p_2)|),$$

which implies $(p_1 - p_2)(|\mathbf{S}^*(p_1)| - |\mathbf{S}^*(p_2)|) \leq 0$. Because $p_1 < p_2$, we have $|\mathbf{S}^*(p_1)| \geq |\mathbf{S}^*(p_2)|$. ∎

**Proof of Lemma 4:** Symmetry in demand nodes ($S_1 = S_2 = \cdots = S_n$ and (17)) implies that the cost function is also symmetric: $\mathbb{E}_t\big[V_{t+1}\big(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}\big)\big] = \mathbb{E}_t\big[V_{t+1}\big(\widetilde{\mathbf{s}}_{t+1}, \mathbf{d}_{t+1}\big)\big]$, where $\widetilde{s}_{0,t+1} = s_{0,t+1}$ and $(\widetilde{s}_{i,t+1}, i \in \mathcal{L})$ is any permutation of $(s_{i,t+1}, i \in \mathcal{L})$.

Suppose $\mathbf{s}_1^{\mathbf{b}}$ is a minimizer for (14) (thus feasible for (18)), but it is not a minimizer for (18). Then, there exists $\mathbf{s}_2^{\mathbf{b}}$ that is feasible and achieves a lower objective in (18). Such $\mathbf{s}_2^{\mathbf{b}}$ can be obtained by decreasing $s_{1,p}^{\mathbf{b}}$ and increasing $s_{1,q}^{\mathbf{b}}$ for some $p, q \in \mathcal{L}$, i.e., $s_{2,p}^{\mathbf{b}} = s_{1,p}^{\mathbf{b}} - \varepsilon$, $s_{2,q}^{\mathbf{b}} = s_{1,q}^{\mathbf{b}} + \varepsilon$, and $s_{2,i}^{\mathbf{b}} = s_{1,i}^{\mathbf{b}}$ for $i \in \mathcal{L}$ and $i \neq p, q$, and $0 < \varepsilon < (s_{1,p}^{\mathbf{b}} - s_{1,q}^{\mathbf{b}})/2$.

Next, swap $s_{1,p}^{\mathbf{b}}$ and $s_{1,q}^{\mathbf{b}}$ and define the new vector as $\widetilde{\mathbf{s}_1^{\mathbf{b}}}$. Similarly, swap $s_{2,p}^{\mathbf{b}}$ and $s_{2,q}^{\mathbf{b}}$ and define the new vector as $\widetilde{\mathbf{s}_2^{\mathbf{b}}}$. Notice that $\mathbf{s}_2^{\mathbf{b}}$ and $\widetilde{\mathbf{s}_2^{\mathbf{b}}}$ each are convex combination of $\mathbf{s}_1^{\mathbf{b}}$ and $\widetilde{\mathbf{s}_1^{\mathbf{b}}}$. Furthermore, $\mathbf{s}_2^{\mathbf{b}} + \widetilde{\mathbf{s}_2^{\mathbf{b}}} = \mathbf{s}_1^{\mathbf{b}} + \widetilde{\mathbf{s}_1^{\mathbf{b}}}$. Therefore, we have

$$\begin{aligned}
\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_1^{\mathbf{b}}, \mathbf{d}_{t+1})\big] &= \frac{1}{2}\Big(\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_1^{\mathbf{b}}, \mathbf{d}_{t+1})\big] + \mathbb{E}_t\big[V_{t+1}(\widetilde{\mathbf{s}_1^{\mathbf{b}}}, \mathbf{d}_{t+1})\big]\Big) \\
&\geq \frac{1}{2}\Big(\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_2^{\mathbf{b}}, \mathbf{d}_{t+1})\big] + \mathbb{E}_t\big[V_{t+1}(\widetilde{\mathbf{s}_2^{\mathbf{b}}}, \mathbf{d}_{t+1})\big]\Big) \\
&= \mathbb{E}_t\big[V_{t+1}(\mathbf{s}_2^{\mathbf{b}}, \mathbf{d}_{t+1})\big].
\end{aligned}$$

Because $\mathbf{s}_1^{\mathbf{b}}$ is a minimizer for (14), the above inequality must hold with equality, i.e., $\mathbf{s}_2^{\mathbf{b}}$ is also a minimizer for (14) and achieves a lower objective value in (18). Continuing this procedure, we can identify a constrained balanced inventory that is also a minimizer for (18).

Finally, if allocating the total local inventory $z$ equally across all demand nodes is feasible, i.e.,

$\mathbf{x} \le \big(x_0, z/n, \ldots, z/n\big) \le \mathbf{y}$, then by the symmetry and convexity of the objective function, such an equal allocation minimizes the expected cost. ∎

**Proofs of Propositions 6 and 7: Overview and Preliminaries**

Proposition 6 is for the case of $\alpha \le \beta$; Proposition 7 is for the case of $\alpha > \beta$, stated in Appendix E. These two propositions provide structures of the optimal solution to (12). Because $q(\mathbf{s}_{t+1} - \mathbf{s}_t, \mathbf{d}_t)$ increases in $\mathbf{s}_{t+1}$ and $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ decreases in $\mathbf{s}_{t+1}$, the problem in (12) with equality constraint (13) is equivalent to the following problem with an inequality constraint:

$$W_t(q_t, \mathbf{s}_t, \mathbf{d}_t) = \min_{\mathbf{s}_{t+1} \in \mathcal{A}} \big\{\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})\big] : q(\mathbf{s}_{t+1} - \mathbf{s}_t, \mathbf{d}_t) \le q_t\big\}, \tag{A.12}$$

where we minimize a convex function over a convex set. Thus, to show that a solution is optimal, we only need to prove that it achieves a local minimum in (12).

Using the definition from (4), the set of feasible $\mathbf{s}_{t+1}$ for (12) is the iso-production hypersurface

$$\mathcal{A}(q) = \Big\{\mathbf{s}_{t+1} \in \mathcal{A} : \psi_\alpha(s_{0,t+1} - s_{0,t}) + \sum_{i \in \mathcal{L}} \psi_\beta\big(d_{i,t} + \psi_\alpha(s_{i,t+1} - s_{i,t})\big) = q\Big\}, \tag{A.13}$$

which separates $\mathcal{A}$ into two parts (production $< q$ and $> q$). Note that $\mathcal{A}(q_t)$ is a piecewise linear hypersurface in $\mathcal{A}$, because $\psi_\alpha(s_{0,t+1} - s_{0,t})$ is piecewise linear in $s_{0,t+1}$ with slopes $\alpha$ and $\alpha^{-1}$ (slope changes at $s_{0,t+1} = s_{0,t}$), and $\psi_\beta\big(d_{i,t} + \psi_\alpha(s_{i,t+1} - s_{i,t})\big)$ is piecewise linear in $s_{i,t+1}$ with slopes $\alpha\beta$, $\alpha\beta^{-1}$, and $\alpha^{-1}\beta^{-1}$ (slope changes at $s_{i,t+1} = s_{i,t} - d_{i,t}/\alpha$ and $s_{i,t+1} = s_{i,t}$; if $s_{i,t} - d_{i,t}/\alpha \le 0$, the segment with slope $\alpha\beta$ does not exist). For ease of exposition, we refer to any linear hypersurface of $\mathcal{A}(q)$ as a face (which has $n$ dimensions), and the intersection of any two adjacent faces as an edge (which has $n-1$ dimensions)

To prove local minimum, we show that the objective value $\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})\big]$ in (12) increases as $\mathbf{s}_{t+1}$ deviates from the prescribed $\mathbf{s}^*_{t+1}$ (or the set containing $\mathbf{s}^*_{t+1}$). We prove this using two steps:

*Step 1.* Find all faces of $\mathcal{A}(q_t)$ that intersect the prescribed $\mathbf{s}^*_{t+1}$ (or the set containing $\mathbf{s}^*_{t+1}$), and identify the edges formed by these faces.

*Step 2.* Prove that the objective value $\mathbb{E}_t\big[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})\big]$ increases when $\mathbf{s}_{t+1}$ moves away from $\mathbf{s}^*_{t+1}$ (or the set containing $\mathbf{s}^*_{t+1}$) in the directions parallel to any of the edges identified in Step 1. (We in fact prove a stronger result that $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ increases for any realization of $\mathbf{d}_{t+1}$.)

Steps 1 and 2 prove local minmum because from $\mathbf{s}^*_{t+1}$ (or the set containing $\mathbf{s}^*_{t+1}$) we can reach any point in any neighboring face identified in Step 1 by taking at most $n$ moves parallel to the edges of the faces; each move increases the objective value, as shown in Step 2.

Instead of repeating Step 1 for every case, we first identify all possible faces and edges of $\mathcal{A}(q_t)$. To find faces, we can consider $\mathcal{A}(q_t)$ as consisting of two parts: a part with $s_{0,t+1} > s_{0,t}$ (store at node 0) and another part with $s_{0,t+1} < s_{0,t}$ (release at node 0). The boundary of the two parts, which contains the edges formed by faces from both parts, corresponds no change in storage level at node 0.

Let $k$ index the faces of $\mathcal{A}(q_t)$. Face $k$ satisfies (A.13), which can be expressed as a linear equation

$$\mathbf{a}_k \cdot \mathbf{s}_{t+1} \equiv \sum_{i \in \mathcal{L} \cup \{0\}} a_{k,i} s_{i,t+1} = b_k, \qquad \text{for } \mathbf{s}_{t+1} \in \text{face } k,$$

where $a_{k,0}$ is either $\alpha$ or $\alpha^{-1}$, while $a_{k,i}$, $i \in \mathcal{L}$, takes three possible values: $\alpha\beta$, $\alpha\beta^{-1}$, or $\alpha^{-1}\beta^{-1}$. These values are exactly the slopes discussed after (A.13).

For the part of $\mathcal{A}(q_t)$ with $s_{0,t+1} > s_{0,t}$, we have $a_{k,0} = \alpha^{-1}$, while $a_{k,i}$, $i \in \mathcal{L}$, have $3^n$ combinations. Thus, this part of $\mathcal{A}(q_t)$ has up to $3^n$ faces (the actual number of faces depends on $q_t$). For $n = 2$, the contours of the 9 possible faces are shown in Figure A.1(a), labeled from 0 to 8. The other part of $\mathcal{A}(q_t)$ with $s_{0,t+1} < s_{0,t}$ consists of faces with $a_{k,0} = \alpha$. These faces are shown in Figure A.1(b) and labeled from $0'$ to $8'$. Note that $\mathcal{A}(q_t)$ cannot contain the lower-left area because $s_{0,t+1} < s_{0,t}$ implies that $\sum_{i \in \mathcal{L}} \psi_\beta\big(d_{i,t} + \psi_\alpha(s_{i,t+1} - s_{i,t})\big) > 0$ due to (A.13).

Let $\{\mathbf{e}_{ij}^{(m)} : m = 1, \ldots, n-1\}$ denote a basis for the $(n-1)$-dimensional vector space parallel to the edge formed by faces $i$ and $j$. Because all coefficients $a_{k,i} > 0$, we can always choose the basis such that $\mathbf{e}_{ij}^{(m)}$ contains exactly two non-zero elements, with one being $-1$ and the other belongs to $(0, 1]$. For $n = 2$ and $\alpha \le \beta$, these basis are shown as vectors in Figure A.1; we omit index $m$ because each edge has only one dimension.

We next prove a lemma on how the value function changes along the directions of these basis.

**Lemma A.2** *(In this lemma, '. . . ' represents omitted zeros.) For any* $\mathbf{d}_{t+1}$*, we have*

*(i)* $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ *increases as* $\mathbf{s}_{t+1}$ *moves along* $\mathbf{e}_{ij}^{(m)} = (\ldots, -1, \ldots, \beta^2, \ldots)$ *or* $(\ldots, \beta^2, \ldots, -1, \ldots)$.

*(ii)* $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ *increases as* $\mathbf{s}_{t+1}$ *moves along* $\mathbf{e}_{ij}^{(m)} = (\beta, \ldots, -1, \ldots)$ *or* $(-1, \ldots, \beta, \ldots)$.

*(iii)* $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ *increases as* $\mathbf{s}_{t+1}$ *moves along* $\mathbf{e}_{ij}^{(m)} = (\alpha^2\beta, \ldots, -1, \ldots)$ *or* $(-1, \ldots, \alpha^2\beta, \ldots)$ *or* $(\ldots, -1, \ldots \alpha^2\beta^2, \ldots)$ *or* $(\ldots, \alpha^2\beta^2, \ldots, -1, \ldots)$.

*(iv) If* $\alpha \le \beta$*, then* $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ *increases as* $\mathbf{s}_{t+1}$ *moves along* $\mathbf{e}_{ij}^{(m)} = (\alpha^2/\beta, \ldots, -1, \ldots)$ *or* $(\ldots, -1, \ldots \alpha^2, \ldots)$ *or* $(\ldots, \alpha^2, \ldots, -1, \ldots)$.

**Proof of Lemma A.2:** Parts (i) and (ii) follow directly from Lemma 2(i) and (ii), respectively.

For part (iii), suppose $\mathbf{s}_{t+1}$ moves in the direction $(\alpha^2\beta, \ldots, -1, \ldots)$ by a small amount $\delta > 0$. Then,

$$V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}) \le V_{t+1}(\mathbf{s}_{t+1} + (\beta\delta, \ldots, -\delta, \ldots), \mathbf{d}_{t+1}) \le V_{t+1}(\mathbf{s}_{t+1} + (\alpha^2\beta\delta, \ldots, -\delta, \ldots), \mathbf{d}_{t+1}),$$

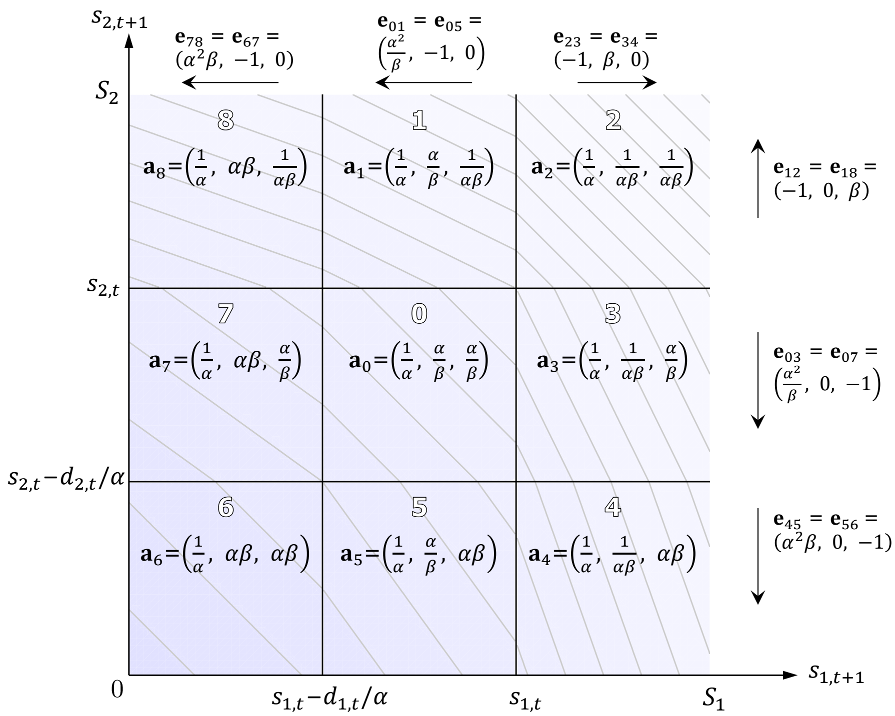
where the first inequality is due to Lemma 2(ii) and the second inequality follows from $\alpha \le 1$ and the monotonicity in Lemma 1. The proof for all other directions in part (iii) are similar.

For part (iv), suppose $\mathbf{s}_{t+1}$ moves in the direction $(\alpha^2/\beta, \ldots, -1, \ldots)$ by a small amount $\delta > 0$.

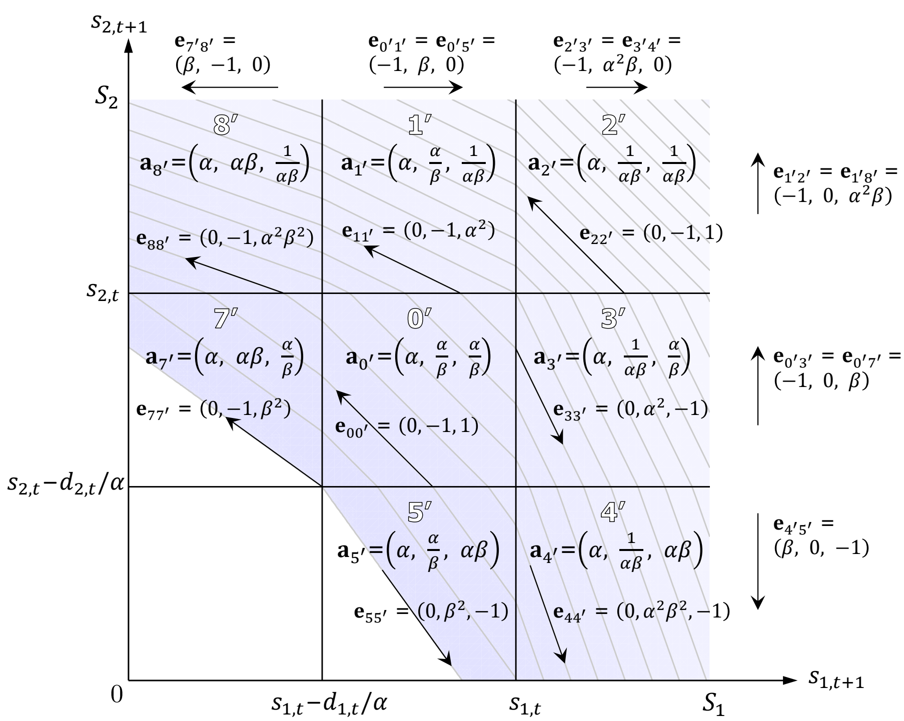
Figure A.1: Contours of $\mathcal{A}(q_t)$, faces, and edges: $n = 2$

This figure illustrates the iso-production surface $\psi_\alpha(s_{0,t+1} - s_{0,t}) + \sum_{i=1,2} \psi_\beta\big(d_{i,t} + \psi_\alpha(s_{i,t+1} - s_{i,t})\big) = q$. Each contour line represents a fixed $s_{0,t+1}$ level; the lower-left contour line has the highest $s_{0,t+1}$ level. If $s_{i,t} - d_{i,t}/\alpha \leq 0$, then the faces between 0 and $s_{i,t} - d_{i,t}/\alpha$ do not exist.

(a) $s_{0,t+1} > s_{0,t}$: store energy at node 0




(b) $s_{0,t+1} < s_{0,t}$: release energy from node 0



Then,

$$V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1}) \le V_{t+1}(\mathbf{s}_{t+1} + (\beta\delta, \dots, -\delta, \dots), \mathbf{d}_{t+1}) \le V_{t+1}(\mathbf{s}_{t+1} + (\alpha^2/\beta \cdot \delta, \dots, -\delta, \dots), \mathbf{d}_{t+1}),$$

where the second inequality follows from $\alpha \le \beta$ (so that $\alpha^2/\beta < \beta$) and the monotonicity in Lemma 1. The proof for all other directions in part (iv) are similar. ∎

**Proof of Proposition 6:** For $q_t \in (q_t^o, \overline{q}_t)$, per proposition, $\mathbf{s}_{t+1}^* = \mathbf{s}_t + \big(\alpha(q_t - q_t^o), 0, \dots, 0\big)$, i.e., serve all demand using $q_t^o$ and store remaining $q_t - q_t^o$ at node 0; no operations at the local storage. Possible directions of deviation from this prescribed solution includes $(-1, \dots, \beta, \dots)$, i.e., move some energy from central storage to a local storage, $(\alpha^2/\beta, \dots, -1, \dots)$, i.e., use some local storage and store more at the central, and $(0, \dots, -1, \dots \alpha^2, \dots)$, i.e., use some local storage and store at another demand node. Lemma A.2 asserts that $V_{t+1}(\mathbf{s}_t, \mathbf{d}_t)$ increases along these directions, which ensures the local optimality of $\mathbf{s}_{t+1}^*$.

For $q_t \in (\underline{q}_t, q_t^o)$, per proposition, $\mathbf{s}_{t+1}^* = \mathbf{s}^{\mathbf{b}}(\underline{\mathbf{s}}_t, \mathbf{s}_t, z, \mathbf{d}_t)$. From the definition of $\mathbf{s}^{\mathbf{b}}(\underline{\mathbf{s}}_t, \mathbf{s}_t, z, \mathbf{d}_t)$ in (14), $\mathbf{s}_{t+1}^*$ minimizes $\mathbb{E}_t[V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})]$ within a set $X = \big\{\mathbf{s}_{t+1} : s_{0,t+1} = s_{0,t}, s_{i,t+1} \in \big[(s_{i,t} - d_{i,t}/\alpha)^+, s_{i,t}\big], \sum_{i \in \mathcal{L}} s_{i,t+1} = z\big\}$. (For $n = 2$, the set $X$ is the intersection of faces 0 and 0′.) We only need to show that $V_{t+1}(\mathbf{s}_{t+1}, \mathbf{d}_{t+1})$ increases as $\mathbf{s}_{t+1}$ moves away from $X$. Possible directions of deviation includes $(-1, \dots, \beta, \dots)$, i.e., use some energy from central storage instead of local storage to serve demand, $(0, \dots, -1, \dots, \beta^2, \dots)$, i.e., use remote storage, $(\alpha^2/\beta, \dots, -1, \dots)$, i.e., store energy at the central node instead of sending to leaves. Lemma A.2 ensures that $V_{t+1}(\mathbf{s}_t, \mathbf{d}_t)$ increases along these directions.

For $q = \underline{q}_t$, per proposition, $\mathbf{s}_{t+1}^* = \underline{\mathbf{s}}_t$, i.e., the set $X$ shrinks to a point, which leads to additional directions of deviation: $(\alpha^2\beta, \dots, -1, \dots)$, i.e., release energy from a local storage and store it at the central. Similarly, for $q = q_t^o$, the set $X$ shrinks to $\mathbf{s}_{t+1}^* = \mathbf{s}_t$, which leads to additional directions: $(-1, \dots, \alpha^2\beta, \dots)$, i.e., release energy from the central and store it at a demand node. Lemma A.2 ensures that $V_{t+1}(\mathbf{s}_t, \mathbf{d}_t)$ increases along these directions.

Proofs for the other cases of Proposition 6 are parallel. ∎

## B. Performance of the Heuristic Method

To ensure that the heuristic method described in Section 6.2 performs well in our context, we consider a simpler demand setting and compute the "optimal" cost of the stochastic dynamic program with discretization, which is then compared with the cost found by the heuristic method. In this simpler setting, there are $n = 4$ demand nodes. At each node $i$, during the even-numbered periods $d_{i,2t} = 0$, while during the odd-numbered periods $d_{i,2t-1}$ takes three possible values: 0, 15, and 30, with probability 0.4, 0.4, and 0.2, respectively, independent across nodes and time. Although simple, the demands exhibit both predictable and unpredictable variabilities, reflecting the characteristics of energy demand.

Table 1: Estimated long-run discounted production cost $V(\mathbf{S})$

(a) $S_0 = 18$, various $S_i$

| $S_i$ | $\lvert\mathbf{S}\rvert$ | Stochastic method | Heuristic method | Dif. |
|---|---|---|---|---|
| 0 | 18 | 104725.5 | 104341.3 | -0.37% |
| 1 | 22 | 100219.6 | 99886.0 | -0.33% |
| 2 | 26 | 97231.8 | 97152.9 | -0.08% |
| 3 | 30 | 94828.9 | 94957.4 | 0.14% |
| 4 | 34 | 92782.1 | 93138.5 | 0.38% |
| 5 | 38 | 91032.0 | 91622.1 | 0.65% |

(b) $S_i = 5$, $i = 1, 2, 3, 4$, various $S_0$

| $S_0$ | $\lvert\mathbf{S}\rvert$ | Stochastic method | Heuristic method | Dif. |
|---|---|---|---|---|
| 0 | 20 | 102552.4 | 102231.4 | -0.31% |
| 4 | 24 | 98849.9 | 98673.9 | -0.18% |
| 8 | 28 | 96048.7 | 96109.7 | 0.06% |
| 12 | 32 | 93794.8 | 94087.7 | 0.31% |
| 16 | 36 | 91876.6 | 92383.5 | 0.55% |
| 20 | 40 | 90266.7 | 90919.9 | 0.72% |

Table 1 shows the expected long-run production cost estimated by solving the stochastic dynamic program in (5) and by the heuristic method. Because solving the dynamic program requires state space discretization and value function interpolation, the estimated cost is slightly higher than the exact optimal cost. Under the heuristic method, the reported cost is an average over 100 sample paths, and thus can be slightly below or above the exact optimal cost. Overall, the differences between the cost estimates by the two policies are very small.

## C. Nonlinear Line Losses

In this appendix, we conduct a robustness test using a quadratic line loss function, which captures the theoretical relationship between resistive heating (also known as Joule heating) and electric current.

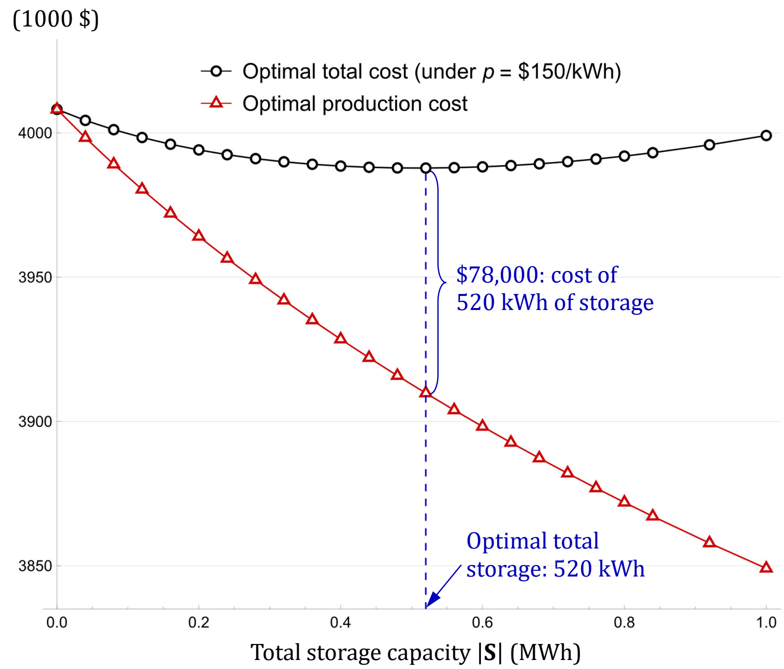
Because our numerical optimization procedure described in Section 6.2 is most efficient when all constraints are linear, we use a piecewise linear function to approximate the quadratic loss function. We choose the quadratic loss function to be $0.04q^2$, where the coefficient 0.04 is chosen so that the total energy loss is about the same as the total energy loss under the linear loss model in the paper.

We repeated the numerical analysis under the piecewise linear loss function. The results are shown in Figure A.2, which can be directly compared to Figure 9 in the paper. We studied 24 different total storage capacity levels in this robustness test, ranging from 0 to 1 MWh. (In Figure 9 in the paper, the range is larger, but the optimal total storage capacity is below 1 MWh.) Computational time for these 24 different levels is about 400 hours.

As shown in Figure A.2, the optimal storage capacity is 520 kWh, with 280 kWh for the industrial facilities and 240 kWh for the residential areas. The key difference from the analysis in the paper is that more storage is placed at the residential areas under the nonlinear loss model. (To compare, under the linear loss model in the paper, the optimal storage consists of 360 kWh for industrial sites

Figure A.2: Optimal cost under given total storage size

Note: For each given total storage capacity (horizontal axis), we optimize storage capacity allocations across the 5-node network to minimize the production cost. Then, we minimize the total cost (sum of production cost and storage investment cost) by choosing the best total storage capacity, which is 520 kWh in this example. This consists of 280 kWh of storage located at the industrial facilities and 240 kWh of storage located at the residential areas.




and only 80 kWh for residential areas.) This is because the residential demand has a higher peak than industrial demand (see Figure 8 in the paper), and quadratic line losses provide an opportunity for the storage to reduce losses by shaving the peak energy transmission.

Despite the differences, the results under nonlinear losses are consistent with the results in the paper in that a) no storage capacity is allocated to the central node, i.e., optimal storage investment still favors distributed storage, and b) the optimal storage investment still places more storage at the industrial demand, due to the minimum demand effect detailed in the paper.

## D. Impact of Central Solar Energy Generation on Storage Investment

In this appendix, we examine the system illustrated in Figure 8 of the paper, where solar power generation is incorporated at the central node. We analyze two scenarios of solar generation. In the first scenario, the amount of solar power is such that the average net demand (total demand minus solar) remains positive for most of the time. In the second scenario, the amount of solar power is doubled, leading to a considerable amount of excess solar power at times, such as during sunny weekends.

An essential part of the analysis is modeling the renewable energy generation process. We adopt the approach in Peng et al. (2023) and model the potential solar generation $r_t$ as $r_t = \overline{r}_t \frac{e^{h_t}}{1+e^{h_t}}$. The deterministic process $\overline{r}_t$ models the diurnal variations of the maximum solar power under a clear sky; the stochastic multiplier $s_t \equiv \frac{e^{h_t}}{1+e^{h_t}} \in (0,1)$ models the intermittency of solar power ($s_t$ is close to 1 under a clear sky; $s_t$ is close to 0 under thick clouds). The inverse relation, $h_t = \ln \frac{s_t}{1-s_t}$, translates $s_t \in (0,1)$ into $h_t \in (-\infty, \infty)$, which we model as a mean-reverting process. Following Peng et al. (2023), we use data from Florida Power and Light to estimate the model parameters.

Figure A.3: Total cost reduction by given total storage capacity

For each solar penetration setting, we compute the total production cost without storage first. Then, for any given total storage capacity, we optimize the storage allocations across the nodes to maximize the production cost reduction. The cost reduction relative to the no storage case is plotted. The peak cost reduction correpond to the optimal storage investment.

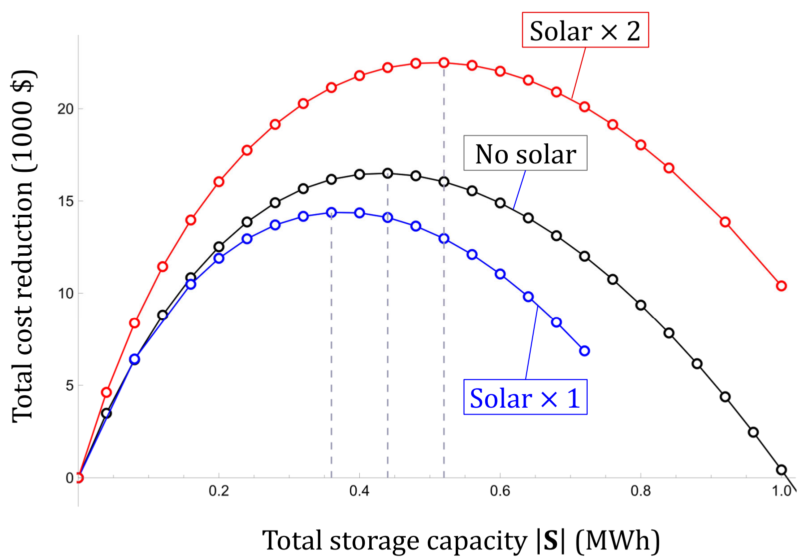



Figure A.3 illustrates the total cost reduction achieved by investing in storage assets and operating them optimally. The "No solar" curve is obtained by subtracting the optimal total cost in Figure 9 of the paper from the cost without storage; the "Solar×1" and "Solar×2" curves correspond to the two scenarios of solar power. The optimal storage investment decisions are:

- No solar: 440 kWh of storage in total, of which 360 kWh is for the two industrial facilities and 80 kWh is for residential areas, and no investment at central node;
- Solar×1: 360 kWh of storage in total, with 320 kWh at industrial sites and 40 kWh at residential areas, and no investment at central node;
- Solar×2: 520 kWh of storage in total, with 360 kWh at industrial sites and 160 kWh at residential areas, and no investment at central node.

In both solar scenarios, investing in centralized storage does not emerge as an optimal decision, despite solar power being generated at the central node. This aligns with the paper's main numerical insight that distributed storage investment is generally preferred. Furthermore, the minimum demand of 100 kWh at each industrial facility continues to drive the allocation of storage investment towards the industrial sites, which is in line with the minimum demand effect found in the paper.

It is evident that increasing the amount of solar energy is beneficial for the entire system. However, the results in Figure A.3 also reveal that both the value of storage and the total storage investment are non-monotone in the amount of solar power. Specifically, the total storage capacity decreases from 440 kWh to 360 kWh, and then increases to 520 kWh as solar power increases. This non-monotonic relationship can be attributed to two factors. First, solar power reduces the peak load on the system, thereby decreasing the need for energy storage. Second, when there is excess solar power, storage capacity helps reducing the waste of the excess energy. These results are consistent with the main findings of Peng et al. (2023).

## E. Optimal Storage Operations for the Case of $\alpha > \beta$

To confine this online appendix to 16 pages as per journal requirement, the analysis of this case has been made available at https://ssrn.com/abstract=4480308.