Online Appendices

Appendix A: Proofs

Proof of Lemma 1: Clearly, formulation (2) is a special case of formulation (1). To show the opposite direction, we transform formulation (1) into (2), whose coefficients and states are represented by hat symbols. Let $\hat{n} = np$, $\hat{m} = m + p - 1$, and divide stage i into p sub-stages $\{i_1, \dots, i_p\}$. At the start of stage i, let $\hat{\mathbf{x}}_{i_1} \equiv (\mathbf{x}_i, \mathbf{y}_{i_1}) \in \mathbb{R}^{m+p-1}$ and $\mathbf{y}_{i_1} = \mathbf{0}$. The multivariate action $\mathbf{u}_i$ can be decomposed as follows. For $j = 1, \dots, p-1$, let $\hat{u}_{i_j} = \mathbf{u}_i(j)$ (jth component of $\mathbf{u}_i$) and $\hat{\mathbf{x}}_{i_{j+1}} = (\mathbf{x}_i, \mathbf{y}_{i_{j+1}})$, where $\mathbf{y}_{i_{j+1}}(k) = \mathbf{y}_{i_j}(k)$ for $k \neq j$ and $\mathbf{y}_{i_{j+1}}(j) = \hat{u}_{i_j}$. In other words, we record the action $\mathbf{u}$ by the auxiliary state vector $\mathbf{y}$. Meanwhile, let $\hat{S}_{i_j} = S_i$ and $\hat{C} = 0$. For $j = p$, let $\hat{u}_{i_p} = \mathbf{u}_i(p)$ and $\hat{A}_{i_p}$, $\hat{B}_{i_p}$, and $\hat{C}_{i_p}$ be the same as A_i, B_i, and C_i, except that $\mathbf{y}_{i_p} = \mathbf{u}_i(1:p-1)$ is now the state rather than the action. Hence, stage i_p of the new formulation is equivalent to stage i, the formulation (1). In terms of the constraint, $F'_{S_i,i}\mathbf{x}_i \leq D'_{S_i,i}\mathbf{u}_i \leq G'_{S_i,i}\mathbf{x}_i$ is considered only at sub-stage $i_{j'}$, where $\mathbf{u}_i(i_{j'})$ is the last component appearing in the constraint, i.e., $D_{S_i,i}(i_{j'}) \neq 0$, while $D_{S_i,i}(i_{j'+1}) = \cdots = D_{S_i,i}(i_p) = 0$. This naturally defines a constraint at sub-stage $i_{j'}$, as the first $j'-1$ components of $\mathbf{u}_i$ are recorded by the endogenous state $\mathbf{y}_{i_{j'}}$. It is easy to see that the transformed problem, in the form of (2), is equivalent to the original problem. □

LEMMA 2. *If* (4) *holds, then* $V_i(\mathbf{x}_i, S_i, u_i)$ *is strictly convex in* u_i *for any* $(\mathbf{x}_i, S_i) \in \mathbb{R}^m \times \{s_1, \dots, s_l\}$ *and* $i = 0, \dots, n-1$.

Proof of Lemma 2: We show that for any $\lambda \in (0,1)$ and two feasible actions $u_i^{(1)}$ and $u_i^{(2)}$, $\lambda V_i(\mathbf{x}_i, S_i, u_i^{(1)}) + (1-\lambda)V_i(\mathbf{x}_i, S_i, u_i^{(2)}) \geq V_i(\mathbf{x}_i, S_i, \lambda u_i^{(1)} + (1-\lambda)u_i^{(2)})$. For any feasible policy $\{u_t\}_{t=i}^n$, the realized cost-to-go is convex by the condition. Because expectation is simply a weighted sum of all realizations, the expected cost-to-go is also convex. Let $\{u_t^{(1)}\}_{t=i+1}^n$ and $\{u_t^{(2)}\}_{t=i+1}^n$ be the corresponding policies that minimize the cost starting from $i+1$ after $u_i^{(1)}$ and $u_i^{(2)}$ are taken at i. We can define a policy $\{\lambda u_t^{(1)} + (1-\lambda)u_t^{(2)}\}_{t=i}^n$. Clearly, it is a feasible and non-anticipating policy. Its expected cost-to-go is less than $\lambda V_i(\mathbf{x}_i, S_i, u_i^{(1)}) + (1-\lambda)V_i(\mathbf{x}_i, S_i, u_i^{(2)})$ by the convexity shown above. Because it is a member of all feasible policies with $u_i = \lambda u_i^{(1)} + (1-\lambda)u_i^{(2)}$, the expected cost is greater than equal to $V_i(\mathbf{x}_i, S_i, \lambda u_i^{(1)} + (1-\lambda)u_i^{(2)})$. This completes the proof. □

Proof of Theorem 1: Now, we prove Theorem 1 by backward induction, naturally leading to the partitioning algorithm. For $i = n$, it is obvious that in the interior of the region $(G_{S_n,n})'\mathbf{x}_n \leq (F_{S_n,n})'\mathbf{x}_n$, the problem is infeasible. We rearrange the objective function of the last stage into $a_{S_n,n}u_n^2 + (b_{S_n,n})'\mathbf{x}_n u_n + \mathbf{x}'_n c_{S_n,n}\mathbf{x}_n$, where $a \in \mathbb{R}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^{m \times m}$. By Lemma 2, $a_{S_n,n} > 0$, and the unconstrained minimizer is $\tilde{u}_n^*(\mathbf{x}_n, S_n) = -(b_{S_n,n})'\mathbf{x}_n / 2a_{S_n,n}$, a linear function of the state $\mathbf{x}_n$. For each $k = 1, \dots, l$ and $S_n = s_k$, we partition the feasible state space by the linear boundaries $\tilde{u}_n^* \leq (F_{S_n,n})'\mathbf{x}_n$, $(F_{S_n,n})'\mathbf{x}_n \leq \tilde{u}_n^* \leq (G_{S_n,n})'\mathbf{x}_n$ and $(G_{S_n,n})'\mathbf{x}_n \leq \tilde{u}_n^*$. In each region, we let $u_n^* = (F_{S_n,n})'\mathbf{x}_n$, $u_n^* = \tilde{u}_n^*$ and $u_n^* = (G_{S_n,n})'\mathbf{x}_n$, respectively, and compute the quadratic value function. For differentiability, it is sufficient to check the boundaries, e.g., $(F_{S_n,n} + b_{S_n,n}/2a_{S_n,n})'\mathbf{x}_n = 0$. One can easily verify that the value functions of the two neighboring regions differ by $a_{S_n,n}\mathbf{x}'_n LL'\mathbf{x}_n$, where $L = F_{S_n,n} + b_{S_n,n}/2a_{S_n,n}$. Therefore, they have equal gradients at the boundary, and J_n is differentiable if $(G_{S_n,n} - F_{S_n,n})'\mathbf{x}_n > 0$.

Suppose that the result holds for $i+1$. At time i, we define a refinement $\{\mathscr{P}_r^{i+1}\}_{r=1}^{n_{i+1}}$ of all partitions $\{\mathscr{P}_r^{k,i+1}\}_{r=1}^{n_{k,i+1}}$, $k=1,\ldots,l$, i.e., two points belong to the same region in $\{\mathscr{P}_r^{i+1}\}_{r=1}^{n_{i+1}}$ if and only if they belong to the same region of all l partitions. For $S_i=s_j$, in the feasible region $(F_{s_j,i}-G_{s_j,i})'\mathbf{x}_i\le 0$, consider the set of $(\mathbf{x}_i,u_i)$ that transit into the region $\mathscr{P}_r^{i+1}$ contained in $\mathscr{P}_{r_k}^{k,i+1}$ for $k=1,\ldots,l$. We require $\mathscr{P}_{r_k}^{k,i+1}$ to be a feasible region if $P_{jk}^{(i)}>0$. The unconstrained minimizer $\tilde{u}_i^*$ is given by a scalar quadratic program:

$$\begin{aligned}\tilde{u}_i^* &= \arg\min_{u_i\in\mathbb{R}}\left\{\begin{pmatrix}u_i\\ \mathbf{x}_i\end{pmatrix}' C_{s_j,i}\begin{pmatrix}u_i\\ \mathbf{x}_i\end{pmatrix}+\sum_{k=1}^{l}P_{jk}^{(i)}\mathbf{x}_{i+1}'N_{r_k}^{k,i+1}\mathbf{x}_{i+1}\right\}\\ &= \arg\min_{u_i\in\mathbb{R}}\left\{a_{s_j,i}u_i^2+(b_{s_j,i})'\mathbf{x}_iu_i+\mathbf{x}_i'c_{s_j,i}\mathbf{x}_i\right\},\end{aligned}$$

when we substitute in $\mathbf{x}_{i+1}=A_{s_j,i}\mathbf{x}_i+B_{s_j,i}u_i$. Because a is positive, as implied by Lemma 2, $\tilde{u}_i^*=-(b_{s_j,i})'\mathbf{x}_i/2a_{s_j,i}$ is a linear function of $\mathbf{x}_i$. Because of the differentiability of J_{i+1} and, thus, $V_i(\cdot,s_j,\cdot)$, the minimum of the Bellman equation is attained if and only if one of the following occurs: (1) $\tilde{u}_i^*\le(F_{s_j,i})'\mathbf{x}_i$ and $u_i^*=(F_{s_j,i})'\mathbf{x}_i$; (2) $(F_{s_j,i})'\mathbf{x}_i\le\tilde{u}_i^*\le(G_{s_j,i})'\mathbf{x}_i$ and $u_i^*=\tilde{u}_i^*$; or (3) $(G_{s_j,i})'\mathbf{x}_i\le\tilde{u}_i^*$ and $u_i^*=(G_{s_j,i})'\mathbf{x}_i$. Each corresponds to a region of the state space bounded by the following:

$$\begin{aligned}&(1)\ \left(A_{s_j,i}+B_{s_j,i}(F_{s_j,i})'\right)\mathbf{x}_i\in\mathscr{P}_{r_k}^{k,i+1},\ -(b_{s_j,i})'\mathbf{x}_i/2a_{s_j,i}\le(F_{s_j,i})'\mathbf{x}_i,\ (F_{s_j,i}-G_{s_j,i})'\mathbf{x}_i\le 0;\\ &(2)\ \left(A_{s_j,i}-B_{s_j,i}(b_{s_j,i})'/2a_{s_j,i}\right)\mathbf{x}_i\in\mathscr{P}_{r_k}^{k,i+1},\ (F_{s_j,i})'\mathbf{x}_i\le-(b_{s_j,i})'\mathbf{x}_i/2a_{s_j,i}\le(G_{s_j,i})'\mathbf{x}_i;\\ &(3)\ \left(A_{s_j,i}+B_{s_j,i}(G_{s_j,i})'\right)\mathbf{x}_i\in\mathscr{P}_{r_k}^{k,i+1},\ (G_{s_j,i})'\mathbf{x}_i\le-(b_{s_j,i})'\mathbf{x}_i/2a_{s_j,i},(F_{s_j,i}-G_{s_j,i})'\mathbf{x}_i\le 0;\end{aligned}$$

for all $k=1,\ldots,l$ such that $P_{jk}^{(i)}>0$. By the induction hypothesis, $\mathscr{P}_{r_k}^{k,i+1}$ is a polyhedron. Hence, the boundaries above are linear in $\mathbf{x}_i$. We substitute in u_i^*, which is linear in $\mathbf{x}_i$, into the Bellman equation and compute the value function, which is quadratic in $\mathbf{x}_i$.

We then show that the regions defined above for all possible $(r_1,\ldots,r_l)$ are disjoint and form a partition of $\mathbb{R}^m$. If the intersection of two regions has interior points, then for any such interior point, there exist $u_i^1\ne u_i^2$, being a local minimizer and bringing $\mathbf{x}_i$ to $\mathbf{x}_{i+1}^1$ and $\mathbf{x}_{i+1}^2$ in different regions at $i+1$ for at least one $S_{i+1}=s_k$. However, this possibility is ruled out by Lemma 2 because of the strict convexity of V_i in u_i. Moreover, the points that are not covered in any region listed above must be infeasible.

Finally, we show that $J_i(\cdot,s_j)$ is differentiable in the interior of the feasible set of the state space for all $j=1,\ldots,l$. If a boundary $L'\mathbf{x}_i=0$ is of the type $(F_{S_i,i}+b_{S_i,i}/2a_{S_i,i})'\mathbf{x}_i=0$ or $(G_{S_i,i}+b_{S_i,i}/2a_{S_i,i})'\mathbf{x}_i=0$, the differentiability can be shown, similar to the case $i=n$. If the boundary is inherited from $i+1$, then because of the differentiability of J_{i+1}, the value function is differentiable at the boundary. Suppose that $J_i(\cdot,s_j)$ is non-differentiable at some $\mathbf{x}_i$. Then, we can always find a boundary that contains $\mathbf{x}_i$, and the gradients of the value function of the two neighboring regions are not equal at the boundary. This cannot occur because the boundary must be of either or both types. Therefore, we have completed the inductive step. $\square$

Proof of Proposition 1: Note that the Bellman equation is in the following form:

$$\begin{aligned}J_n(x_n,d_n,Q_n)&=\left(d_n+\frac{x_n}{2q_n}\right)x_n\\ J_i(x_i,d_i,Q_i)&=\min_{0\le u_i\le x_i}\left\{\left(d_i+\frac{u_i}{2Q_i}\right)u_i+\mathbb{E}_i\left[J_{i+1}\left(x_i-u_i,e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right),Q_{i+1}\right)\right]\right\}\\ u_n^*(x_n,d_n,Q_n)&=x_n\\ u_i^*(x_i,d_i,Q_i)&=\arg\min_{0\le u_i\le x_i}\left\{\left(d_i+\frac{u_i}{2Q_i}\right)u_i+\mathbb{E}_i\left[J_{i+1}\left(x_i-u_i,e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right),Q_{i+1}\right)\right]\right\},\end{aligned}$$

where $\mathbb{E}_i[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_i]$. Similarly, for (7), we have

$$\begin{aligned}
\bar{J}_n(x_n,d_n,Q_n,M_n) &= \left(M_n + d_n + \frac{x_n}{2Q_n}\right) x_n \\
\bar{J}_i(x_i,d_i,Q_i,M_i) &= \min_{0\le u_i\le x_i} \left\{ \left(M_i + d_i + \frac{u_i}{2Q_i}\right) u_i + \mathbb{E}_i\left[\bar{J}_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}, M_{i+1}\right)\right]\right\} \\
\bar{u}_n^*(x_n,d_n,Q_n,M_n) &= x_n \\
\bar{u}_i^*(x_i,d_i,Q_i,M_i) &= \arg\min_{0\le u_i\le x_i} \left\{ \left(M_i + d_i + \frac{u_i}{2Q_i}\right) u_i + \mathbb{E}_i\left[\bar{J}_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}, M_{i+1}\right)\right]\right\}.
\end{aligned}$$

For part (i) and (ii), we use backward induction to show that

$$\begin{aligned}
\bar{J}_i(x_i,d_i,Q_i,M_i) &= J_i(x_i,d_i,Q_i) + x_i M_i \\
\bar{u}_i^*(x_i,d_i,Q_i,M_i) &= u_i^*(x_i,d_i,Q_i).
\end{aligned}$$

At time t_n, we have

$$\begin{aligned}
\bar{J}_n\left(x_n,d_n,Q_n,M_n\right) &= \left(d_n + \frac{x_n}{2Q_n}\right) x_n + x_n M_n = J_n(x_n,d_n,Q_n) + x_n M_n \\
\bar{x}_n^*(x_n,d_n,Q_n,M_n) &= x_n = x_n^*(x_n,d_n,Q_n),
\end{aligned}$$

and the result holds. Suppose that at stage $i+1$, we have

$$\begin{aligned}
\bar{J}_{i+1}\left(x_{i+1},d_{i+1},Q_{i+1},M_{i+1}\right) &= J_{i+1}(x_{i+1},d_{i+1},Q_{i+1}) + x_{i+1}M_{i+1} \\
\bar{u}_{i+1}^*(x_{i+1},d_{i+1},Q_{i+1},M_{i+1}) &= u_{i+1}^*(x_{i+1},d_{i+1},Q_{i+1}).
\end{aligned}$$

Then, according to the Bellman equation, we have

$$\begin{aligned}
\bar{J}_i(x_i,d_i,Q_i,M_i) &= \min_{0\le u_i\le x_i} \left\{ \left(M_i + d_i + \frac{u_i}{2Q_i}\right) u_i + \mathbb{E}_i\left[\bar{J}_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}, M_{i+1}\right)\right]\right\} \\
&= \min_{0\le u_i\le x_i} \left\{ \left(d_i + \frac{u_i}{2Q_i}\right) u_i + u_i M_i \right. \\
&\quad \left. + \mathbb{E}_i\left[J_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}\right)\right] + \mathbb{E}_i[(x_i - u_i)M_{i+1}]\right\} \\
&= \min_{0\le u_i\le x_i} \left\{ \left(d_i + \frac{u_i}{2Q_i}\right) u_i + u_i M_i \right. \\
&\quad \left. + \mathbb{E}_i\left[J_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}\right)\right] + (x_i - u_i)M_i\right\} \\
&= \min_{0\le u_i\le x_i} \left\{ \left(d_i + \frac{u_i}{2Q_i}\right) u_i + \mathbb{E}_i\left[J_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}\right)\right]\right\} + x_i M_i \\
&= J(x_i,d_i,Q_i) + x_i M_i,
\end{aligned}$$

because M_i is a martingale. Given that the optimization problem of u_i does not involve M_i, we can deduce that u_i^* and $\bar{u}_i^*$ must be equal. Thus, the result is proved by induction.

For part (iii), we first show that $J_i(x_i,d_i,Q_i)$ is an increasing function of d_i by backward induction: It is straightforward for t_n; the inductive step is given by the Bellman equation

$$J_i(x_i,d_i,Q_i) = \min_{0\le u_i\le x_i} \left\{ \left(d_i + \frac{u_i}{2Q_i}\right) u_i + \mathbb{E}_i\left[J_{i+1}\left(x_i - u_i, e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{Q_i}\right), Q_{i+1}\right)\right]\right\}.$$

Note that the expectation is increasing in d_i by induction. Thus, J_i increases in d_i. Next, in the above Bellman equation, ρ_i appears only in d_{i+1}. As J is increasing in d, it must be decreasing in ρ.

For part (iv), we use backward induction to prove a stronger result: $J_i(x_i, d_i, Q_i') \le J_i(x_i, d_i, Q_i)$ for $Q_i' \ge Q_i$. Note that at time t_n, $J_n(x_n, d_n, Q_n) = (d_n + x_n/2Q_n)\, x_n$ is a decreasing function of Q_n. If this is true for $i+1$, then for i, we have

$$\begin{aligned} J_i(x_i,d_i,Q_i) &= \min_{0\le u_i\le x_i}\left\{\left(d_i+\frac{u_i}{2Q_i}\right)u_i+\sum_{j=1}^{l}\mathbb{P}(Q_{i+1}=s_j|Q_i)J_{i+1}\left(x_i-u_i, e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right), Q_{i+1}\right)\right\} \\ &\ge \min_{0\le u_i\le x_i}\left\{\left(d_i+\frac{u_i}{2Q_i'}\right)u_i+\sum_{j=1}^{l}\mathbb{P}(Q_{i+1}=s_j|Q_i)J_{i+1}\left(x_i-u_i, e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right), Q_{i+1}\right)\right\}. \end{aligned}$$

By induction, J_{i+1} is a decreasing function of Q_{i+1}. Thus,

$$\begin{aligned} &\sum_{j=1}^{l}\left(\mathbb{P}\left(Q_{i+1}=s_j|Q_i\right)-\mathbb{P}\left(Q_{i+1}=s_j|Q_i'\right)\right)J_{i+1}\left(x_i-u_i, e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right), Q_{i+1}\right) \\ \le &\sum_{j=1}^{l}\left(\mathbb{P}\left(Q_{i+1}=s_j|Q_i\right)-\mathbb{P}\left(Q_{i+1}=s_j|Q_i'\right)\right)J_{i+1}\left(x_i-u_i, e^{-\rho_i\Delta t}\left(d_i+\frac{u_i}{Q_i}\right), a\right)=0. \end{aligned}$$

Therefore, we have shown that $J_i(x_i, d_i, Q_i) \ge J_i(x_i, d_i, Q_i')$ for i and have thus completed the proof. □

Proof of Proposition 2: Note that the realized cost can be expressed as follows:

$$\begin{aligned} \sum_{i=0}^{n}u_i\left(d_i+\frac{u_i}{2Q_i}\right) &= \sum_{i=0}^{n}\left(\exp\left(-\sum_{j=0}^{i-1}\rho_j\Delta t\right)d_0+\sum_{j=0}^{i-1}\exp\left(-\sum_{k=j}^{i-1}\rho_k\Delta t\right)\frac{u_j}{Q_j}+\frac{u_i}{Q_i}\right)u_i \\ &= \sum_{i=0}^{n}\frac{u_i^2}{2Q_i}+\sum_{i<j}^{n}\exp\left(-\sum_{k=i}^{j-1}\rho_k\Delta t\right)\frac{u_iu_j}{Q_i}+d_0\left(\sum_{i=0}^{n}\exp\left(-\sum_{j=0}^{i-1}\rho_j\Delta t\right)u_i\right). \end{aligned}$$

If we index the rows and columns from 0, then the symmetric quadratic matrix $H \in \mathbb{R}^{(n+1)\times(n+1)}$ is in the form $H_{ii} = 1/2Q_i$ and $H_{ij} = \exp(-\sum_{k=i}^{j-1}\rho_k\Delta t)/2Q_i$ for $0 \le i < j \le n$. If $Q_i > \exp(-2\rho_i\Delta t)Q_{i+1}$ for $i = 0, 1, \ldots, n-1$, then we can perform the Cholesky decomposition of H: Let L be a lower triangular matrix with

$$L_{00}=\sqrt{\frac{1}{2Q_0}},\quad L_{ii}=\sqrt{\frac{1}{2Q_i}-e^{-2\rho_{i-1}\Delta t}\frac{1}{2Q_{i-1}}},\quad L_{ij}=e^{-\rho_{i-1}\Delta t}L_{i-1,j},$$

for $0 \le j < i \le n$. It is easy to confirm that $LL' = S$. Because $L_{ii} > 0$, H is positive definite. Therefore, if $s_1 > \exp(-2\rho_i\Delta t)\, s_l$, then the realized cost is always convex in $\{u_i\}_{i=0}^n$. The same argument holds for the realized cost starting from stage i for $i = 0, \ldots, n-1$. Thus, using Lemma 2, we have completed the proof. □

Proof of Proposition 3: Part (i) is straightforward because the set of all admissible deterministic policies is a subset of Θ. For part (ii), let $\{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\}$ be a deterministic vector valued in the space $\{s_1, \ldots, s_l\}^{n+1}$. We construct a sequence of Markov chains $\{Q_i\}_{0\le i\le n}^{(k)}$ that converge to $\{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\}$, with volatility $\sigma_i^{(k)}$. The initial exogenous states $Q_0^{(k)}$ are equal to $\tilde{q}_0$, and their parameters satisfy the following:

$$\alpha = \min_{0\le i\le n-1}\min_{s_j\ne\tilde{q}_{i+1}}\left\{(s_j-\tilde{q}_i-\theta(\mu_i-\tilde{q}_i)\Delta t)^2-(\tilde{q}_{i+1}-\tilde{q}_i-\theta(\mu_i-\tilde{q}_i)\Delta t)^2\right\}>0.$$

The condition guarantees that the deterministic chain $\{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\}$ is the only likely sample path in $\{Q_i\}^{(k)}_{0\le i\le n}$. Let $C^{(k)}_{det}$, $C^{(k)}_{sto}$ be the optimal costs of the deterministic formulation and the MDP formulation when the market depth follows $\{Q_i\}^{(k)}_{0\le i\le n}$. We will show that if $\sigma^{(k)} \to 0$ as $k \to \infty$, then $\lim_{k\to\infty}\left|C^{(k)}_{det} - C^{(k)}_{sto}\right| = 0$.

First, we compute the likelihood of $\{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\}$:

$$
\begin{aligned}
&\mathbb{P}\left(\{Q^{(k)}_0, Q^{(k)}_1, \ldots, Q^{(k)}_n\} = \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\}\right)\\
=&\mathbb{P}\left(Q^{(k)}_1 = \tilde{q}_1 | Q^{(k)}_0 = \tilde{q}_0\right)\mathbb{P}\left(Q^{(k)}_2 = \tilde{q}_2 | Q^{(k)}_1 = \tilde{q}_1\right)\cdots\mathbb{P}\left(Q^{(k)}_n = \tilde{q}_n | Q^{(k)}_{n-1} = \tilde{q}_{n-1}\right)\\
=&\prod_{i=0}^{n-1}\frac{\exp\left(-(\tilde{q}_{i+1} - \tilde{q}_i - \theta(\mu_i - \tilde{q}_i)\Delta t)^2/2(\sigma^{(k)}\tilde{q}_i)^2\Delta t\right)}{\sum_{j=1}^{l}\exp\left(-(s_j - \tilde{q}_i - \theta(\mu_i - \tilde{q}_i)\Delta t)^2/2(\sigma^{(k)}\tilde{q}_i)^2\Delta t\right)}\\
=&\prod_{i=0}^{n-1}\frac{1}{\sum_{j=1}^{l}\exp\left(-\frac{(s_j - \tilde{q}_i - \theta(\mu_i - \tilde{q}_i)\Delta t)^2 - (\tilde{q}_{i+1} - \tilde{q}_i - \theta(\mu_i - \tilde{q}_i)\Delta t)^2}{2(\sigma^{(k)}\tilde{q}_i)^2\Delta t}\right)}\\
\ge&\left(\frac{1}{1 + (l-1)\exp\left(-\alpha/2(\sigma^{(k)}\tilde{q}_i)^2\Delta t\right)}\right)^n\\
=&1 - O\left(\exp\left(-\frac{\alpha}{2(\sigma^{(k)}\max\{s_j\})^2\Delta t}\right)\right),
\end{aligned}
$$

as $k \to \infty$. Therefore, the Markov chain $\left\{Q^{(k)}_i\right\}$ converges to the deterministic chain in probability.

Next, consider the deterministic optimization problem of $\{x_i\}$:

$$
\begin{aligned}
\min \quad & \sum_{i=0}^{n} u_i\left(d_i + \frac{u_i}{2\tilde{q}_i}\right)\\
\text{s.t.} \quad & \sum_{i=0}^{n} u_i = X,\ u_i \ge 0 \quad i = 0, 1, \ldots, n\\
& d_{i+1} = e^{-\rho_i\Delta t}\left(d_i + \frac{u_i}{\tilde{q}_i}\right)
\end{aligned}
$$

The objective function is a quadratic function of u_i, and the constraints form a closed set. Hence, its minimum can be achieved. Denote the optimal cost and an optimal solution as $\tilde{C}$ and $\{\tilde{u}_i\}$. Let $\left\{x^{(k)}_i\right\}$ be any policy of Θ. We use the notation $\left\{x^{(k)}_i(\omega)\right\}$ to emphasize the stochastic nature of the policy. Now, we have

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=0}^{n} x^{(k)}_i(\omega)\left(d_i(\omega) + \frac{x^{(k)}_i(\omega)}{2Q^{(k)}_i(\omega)}\right)\right]\\
=&\mathbb{E}\left[\sum_{i=0}^{n} x^{(k)}_i(\omega)\left(d_i(\omega) + \frac{x^{(k)}_i(\omega)}{2Q^{(k)}_i(\omega)}\right)1(\{Q^{(k)}_0, Q^{(k)}_1, \ldots, Q^{(k)}_n\} = \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right]\\
&+\mathbb{E}\left[\sum_{i=0}^{n} x^{(k)}_i(\omega)\left(d_i(\omega) + \frac{x^{(k)}_i(\omega)}{2Q^{(k)}_i(\omega)}\right)1(\{Q^{(k)}_0, Q^{(k)}_1, \ldots, Q^{(k)}_n\} \ne \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right]\\
\ge&\mathbb{E}\left[\sum_{i=0}^{n} x^{(k)}_i(\omega)\left(d_i(\omega) + \frac{x^{(k)}_i(\omega)}{2\tilde{q}^{(k)}_i}\right)1(\{q^{(k)}_0, Q^{(k)}_1, \ldots, Q^{(k)}_n\} = \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right]\\
\ge&\tilde{C}\mathbb{E}\left[1(\{Q^{(k)}_0, Q^{(k)}_1, \ldots, Q^{(k)}_n\} = \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right]\\
\ge&\tilde{C}\left(1 - O\left(\exp\left(-\frac{\alpha}{2(\sigma^{(k)}\max\{s_j\})^2\Delta t}\right)\right)\right).
\end{aligned}
$$

Therefore, the optimal policy in Θ can achieve no better optimum than $\tilde{C}$ when k goes to infinity. However, if we let $u_i^{(k)}(\omega) = \tilde{u}_i$, which is a deterministic policy and a special member of Θ, then the objective function is

$$
\begin{aligned}
\mathbb{E}\left[\sum_{i=0}^{n} \tilde{x}_i \left(d_i(\omega) + \frac{\tilde{x}_i}{2Q_i^{(k)}(\omega)}\right)\right] &= \mathbb{E}\left[\sum_{i=0}^{n} \tilde{x}_i \left(d_i(\omega) + \frac{\tilde{x}_i}{2Q_i^{(k)}(\omega)}\right) 1(\{Q_0^{(k)}, Q_1^{(k)}, \ldots, Q_n^{(k)}\} = \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right] \\
&\quad + \mathbb{E}\left[\sum_{i=0}^{n} \tilde{x}_i \left(d_i(\omega) + \frac{\tilde{x}_i}{2Q_i^{(k)}(\omega)}\right) 1(\{Q_0^{(k)}, Q_1^{(k)}, \ldots, Q_n^{(k)}\} \neq \{\tilde{q}_0, \tilde{q}_1, \ldots, \tilde{q}_n\})\right] \\
&\leq \tilde{C} + O\left(\exp\left(-\frac{\alpha}{2(\sigma^{(k)} \max\{s_j\})^2 \Delta t}\right)\right)
\end{aligned}
$$

The second term in the last inequality is due to the boundedness of all variables, x, d and q. Thus, we have shown that $\lim_{k\to\infty} C_{sto}^{(k)} = \tilde{C}$.

Because the policy $\{\tilde{x}_i\}_{i=0}^n$ is deterministic, the same argument holds for $C_{det}^{(k)}$, and we can obtain $\lim_{k\to\infty} C_{det}^{(k)} = \tilde{C}$. Hence, we have proven that the upper bound converges to the optimal value of the original problem. $\square$

The Upper Bound for the Complexity of Both Applications (Section 2.4): To show that $(2l)^{n+1}$ is an upper bound for the LOB application, note that the terminal stage has two regions in a partition. Because $m = 2$, the refinement of l partitions, each having k regions, generates lk regions. For each such region, wait/buy gives two options. Therefore, backward induction multiplies the complexity by $2l$ at each stage and leads to the upper bound.

To show the bound for the complexity of renewable electricity management, note that the constraint is present only at the terminal stage when there are two regions. Thus, only the refinement of l partitions creates new regions, which multiplies the complexity by l at each stage. In the case of binomial noise, we have $m = 3$ and $l = 4$, leading to the bound $2l^n$. $\square$

Appendix B: Details of the Numerical Examples in the Paper

When random noise is unobserved: Without loss of generality, consider the following formulation:

$$
\begin{aligned}
\min_{\{u_i\}_{i=0}^n} \quad & \mathbb{E}\left[\sum_{i=0}^{n} \begin{pmatrix} u_i \\ \mathbf{x}_i \end{pmatrix}' C_i \begin{pmatrix} u_i \\ \mathbf{x}_i \end{pmatrix}\right] \\
\text{s.t.} \quad & F_i'\mathbf{x}_i \leq u_i \leq G_i'\mathbf{x}_i,\ i = 0, \ldots, n \\
& \mathbf{x}_{i+1} = A_i\mathbf{x}_i + B_i u_i + \epsilon_i,\ i = 0, \ldots, n-1,
\end{aligned} \tag{10}
$$

where ϵ_i is a Markov chain, and only its distribution conditional on ϵ_{i-1} (not its realization) is known at stage i. This fits into the usual setting of random noise. We next show how (10) can be transformed into (2). Consider the following formulation:

$$
\begin{aligned}
\min_{\{u_i\}_{i=0}^{2n}} \quad & \mathbb{E}\left[\sum_{i=0}^{n} \begin{pmatrix} u_{2i} \\ \mathbf{x}_{2i} \end{pmatrix}' C_i \begin{pmatrix} u_{2i} \\ \mathbf{x}_{2i} \end{pmatrix}\right] \\
\text{s.t.} \quad & F_{i/2}'\mathbf{x}_i \leq u_i \leq G_{i/2}'\mathbf{x}_i,\ i = 0, 2, \ldots, 2n \\
& 0 \leq u_i \leq 0,\ i = 1, 3, \ldots, 2n-1 \\
& \mathbf{x}_{i+1} = A_{i/2}\mathbf{x}_i + B_{i/2}u_i,\ i = 0, 2, \ldots, 2n-2 \\
& \mathbf{x}_{i+1} = \mathbf{x}_i + \epsilon_{(i-1)/2},\ i = 1, 3, \ldots, 2n-1,
\end{aligned} \tag{11}
$$

where $\epsilon_{(i-1)/2}$ is observed at i if i is odd, as in formulation (2). In formulation (11), we divide each decision stage i in (10) into a pre-decision stage $2i$, when the decision is made, and a post-decision stage $2i+1$, when the randomness is realized. It is not difficult to show that (11) is equivalent to (10): Their costs and dynamics are the same considering the even stages of (11); in terms of the information structure, when decision u_{2i} is made in (11), the decision-maker observes ϵ_{i-1} and knows only the distribution of ϵ_i. Therefore, the setting of unobserved random noises can be incorporated.

The Matrix Formulation of Example 1: The convexity condition Assumption 1 is clearly satisfied, since the stagewise cost is convex. The matrix $A \in \mathbb{R}^{11\times 11}$ does not depend on i and S_i. $A(1,1) = A(2,2) = A(j,j+1) = 1$ for $j = 2,\dots,10$. For matrix $B \in \mathbb{R}^{11}$, $B_{S_i,i}(2) = 1$ if $S_i = 1$ and $B_{S_i,i}(11) = 1$ if $S_i = 9$. The matrix $C \in \mathbb{R}^{12\times 12}$ is independent of i and S_i with $C(1,1) = C(2,2) = C(3,3) = 1$ and $C(3,2) = C(2,3) = -1$. The matrix $F = \mathbf{0}$, and there is no upper bound. All unstated components of the matrices are zero. To incorporate the terminal cost $c(\mathbf{x}_3) = \sum_{i=3}^{12}(x_i - 1)^2$, we can use the terminal value function $V_3(\mathbf{x}) = c(\mathbf{x}_3 = \mathbf{x}) = (x-1)^2 + (x+d_1-2)^2 + \cdots + (x+d_1+\cdots+d_9-10)^2$, which is quadratic in the endogenous state.

The MDP Formulation of the Electricity Management Problem: We use the pre- and post-decision formulation for unobserved random noise. Let the Markov chain S_i be i.i.d. draws of $(\epsilon^{(1)}, \epsilon^{(2)})$, with four states $s_1 = (1,1)$, $s_2 = (1,-1)$, $s_3 = (-1,1)$ and $s_4 = (-1,-1)$, from distribution $\mathbb{P}(s_1) = \mathbb{P}(s_4) = (1+\rho)/4$ and $\mathbb{P}(s_2) = \mathbb{P}(s_3) = (1-\rho)/4$. Let $\mathbf{x} = (x, p, d, \mathbf{1})$, where $\mathbf{1}$ represents a state variable that always equals 1. For $i = 0, 2, 4, \dots, 2n-2$,

$$A_{s_j,i} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \mu\Delta t \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; B_{s_j,i} = \begin{pmatrix} 1 \\ \nu \\ 0 \\ 0 \end{pmatrix}, \; C_{s_j,i} = \begin{pmatrix} \gamma & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and there is no constraint (no boundaries associated with F or G are generated in the partition). For $i = 1, 3, \dots, 2n-1$,

$$A_{s_j,i} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \sigma_p s_j(1)\sqrt{\Delta t} \\ 0 & 0 & 1 & \sigma_d s_j(2)\sqrt{\Delta t} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; B_{s_j,i} = 0^4, \; C_{s_j,i} = 0^{5\times 5}, \; F_{s_j,i} = 0^4; G_{s_j,i} = 0^4.$$

For $i = n$,

$$C_{s_j,i} = \begin{pmatrix} \beta+\eta & \eta & 0 & -\eta & 0 \\ \eta & \eta & 0 & -\eta & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\eta & -\eta & 0 & \eta & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and the constraint is $\hat{u} \geq 0$. Then, our algorithm can be applied to solve the optimal trading/production problem for the power producer. To reduce complexity, one can also replace the state d and x with their difference $d - x$.

Although we divide a stage into pre- and post-decision stages, the computational cost is significantly less than $(m, 2n, l)$ because ϵ_i is white noise: at even stages in (11), there is no dependence on the previous information, and no further partitions are generated. In fact, the complexity is still (m, n, l).

Random Coefficient Matrices in the Randomized Computational Experiments (Section 2.4.1): A is the identity matrix of size m for all (s_j, i); the components of $B_{s_j,i}$ are drawn independently and uniformly from $[-1,1]$; for C, we first generate a matrix $L \in \mathbb{R}^{(m+1)\times(m+1)}$ whose components are independent and uniformly distributed in $[-1,1]$, and then, let $C = L^T L$; F is a zero vector of size m; the components of G are drawn independently and uniformly from $[0,1]$; and the components in the transition probability matrix are independent and uniformly distributed in $[0,1]$ and then normalized by their row sum. Note that for B, C and G, we generate l sets of coefficient matrices (one for each exogenous state). The coefficients are assumed to be constant over time. We use special forms of A and F; otherwise, the problem is likely to be infeasible in the whole state space even for a small n.

The Random Experiments for the Applications (Section 2.4.1): For the LOB application, for given m, n and l, we randomly generate $\rho \sim U(0,6)$, $T \sim U(0,5)$, $x_0 \sim U(0,10000)$, $\sigma \sim U(0,3)$ and $\theta \sim U(0,5)$, where $U(a,b)$ is a random variable uniformly distributed within $[a,b]$. In the electricity application, $m = 3$ ((d,x) collapsing to a single state $d - x$, price p and the constant state $\mathbf{1}$) and $l = 4$. We randomly generate $T \sim U(1,24)$, $\sigma_p \sim U(0,1)$, $\sigma_d \sim U(0,1000)$, $\beta \sim U(0,0.2)$, $\eta \sim U(0,100)$ and $\rho \sim U(0,0.8)$ and let $\gamma = 0.2$, $\nu = 4 \times 10^{-5}$.