ONLINE SUPPLEMENT: OPTIMAL HOUR–AHEAD BIDDING IN THE REAL–TIME ELECTRICITY MARKET WITH BATTERY STORAGE USING APPROXIMATE DYNAMIC PROGRAMMING

DANIEL R. JIANG* AND WARREN B. POWELL*

1. Proofs

Proposition 1. *For an initial resource level $r \in \mathcal{R}$, a vector of intra–hour prices $P \in \mathbb{R}^M$, a bid $b = (b^-, b^+) \in \mathcal{B}$, and a subinterval m, the resource transition function $g_m^R\big(r, q(P,b)\big)$ is nondecreasing in r, b^-, and b^+.*

Proof. Since

$$g_{m+1}^R(R_t, q_s) = \big[\min\{g_m^R(R_t, q_s) - e_m^\intercal q_s, R_{\max}\}\big]^+,$$

it is clear that the transition from g_m^R to g_{m+1}^R is nondecreasing in the value of g_m and nonincreasing in the value of $e_m^\intercal q_s$. Thus, a simple induction argument shows that for $r_1, r_2 \in \mathcal{R}$ and $q_1, q_2 \in \{-1, 0, 1\}^M$ where $r_1 \leq r_2$ and $q_1 \leq q_2$,

$$g_M^R(r_1, q_2) \leq g_M^R(r_2, q_1).$$

The result follows from the fact that $q(P, b)$ is nonincreasing in b. □

Proposition 2. *For an initial $l \in \mathcal{L}$, a vector of intra–hour prices $P \in \mathbb{R}^M$, a bid $b = (b^-, b^+) \in \mathcal{B}$, and a subinterval m, the transition function $g_m^L\big(l, d(P,b)\big)$ is nondecreasing in l, b^-, and b^+.*

Proof. The transition

$$g_{m+1}^L(L_t, d_s) = \big[g_m^L(L_t, d_s) - e_m^\intercal d_s\big]^+$$

is nondecreasing in g_m^L and nonincreasing in $e_m^\intercal d_s$. Like in Proposition 1, induction shows that for $l_1, l_2 \in \mathcal{L}$ and $d_1, d_2 \in \{0, 1\}^M$ where $l_1 \leq l_2$ and $d_1 \leq d_2$,

$$g_M^L(l_1, d_2) \leq g_M^L(l_2, d_1).$$

The result follows from the fact that $d(P, b)$ is nonincreasing in b. □

Proposition 3. *The contribution function $C_{t,t+2}(S_t, b_t)$, with $S_t = (R_t, L_t, b_{t-1}, P_t^S)$ is nondecreasing in R_t, L_t, b_{t-1}^-, and b_{t-1}^+.*

Proof. First, we argue that the revenue function $C(r, l, P, b)$ is nondecreasing in r and l. From their respective definitions, we can see that γ_m and U_m are both nondecreasing in their first arguments. These arguments can be written in terms of r and l through the transition functions g_m^R and g_m^L. Applying Proposition 1 and Proposition 2, we can confirm that $C(r, l, P, b)$ is nondecreasing in r and l. By its definition,

$$C_{t,t+2}(S_t, b_t) = \mathbf{E}\Big[C\big(g^R(R_t, P_{(t,t+1]}, b_{t-1}), g^L(L_t, P_{(t,t+1]}, b_{t-1}), P_{(t+1,t+2]}, b_t\big) | S_t\Big].$$

Again, applying Proposition 1 and Proposition 2 (for $m = M$), we see that the term inside the expectation is nondecreasing in R_t, b_{t-1}^-, and b_{t-1}^+ (composition of nondecreasing functions) for any outcome of $P_{(t,t+1]}$ and $P_{(t+1,t+2]}$. Thus, the expectation itself is nondecreasing. □

*Department of Operations Research and Financial Engineering, Princeton University

Proposition 4. *The optimal value function* $V_t^*(S_t)$*, with* $S_t = (R_t, L_t, b_{t-1}, P_t^S)$ *is nondecreasing in* R_t*,* L_t*,* b_{t-1}^-*, and* b_{t-1}^+*.*

Proof. Define the function $V_t^b(S_t, b_t) = \mathbf{E}(V_{t+1}^*(S_{t+1})|S_t, b_t)$, often called the *post–decision* value function (see Powell [2011]). Thus, we can rewrite the optimality equation as:

$$\begin{aligned} V_t^*(S_t) &= \max_{b_t \in \mathcal{B}} \big\{ C_{t,t+2}(S_t, b_t) + V_t^b(S_t, b_t) \big\} \text{ for } t = 0, 1, 2, \ldots, T-1, \\ V_T^*(S_T) &= C_{\text{term}}(S_T). \end{aligned} \tag{1}$$

The proof is by backward induction on t. The base case is $t = T$ and since $V_T^*(\cdot)$ satisfies monotonicity for any state $s \in \mathcal{S}$ by definition. Notice that the state transition function satisfies the following property. Suppose we have a fixed action b_t and two states $S_t = (R_t, L_t, b_{t-1}, P_t^S)$ and $S_t' = (R_t', L_t', b_{t-1}', P_t^S)$ where $(R_t, L_t, b_{t-1}) \le (R_t', L_t', b_{t-1}')$. Then, for any realization of the intra–hour prices $P_{(t,t+1]}$ (by Propositions 1 and 2),

$$\begin{aligned} S_{t+1} &= (R_{t+1}, L_{t+1}, b_t, P_{t+1}^S) = S^M(S_t, b_t, P_{(t,t+1]}), \\ S_{t+1}' &= (R_{t+1}', L_{t+1}', b_t, P_{t+1}^S) = S^M(S_t', b_t, P_{(t,t+1]}), \end{aligned}$$

with $R_{t+1} \le R_{t+1}'$ and $L_{t+1} \le L_{t+1}'$, implying that $S_{t+1} \le S_{t+1}'$. This means that the transition function satisfies a specialized nondecreasing property. Using this and supposing that $V_{t+1}^*(\cdot)$ satisfies the statement of the proposition (induction hypothesis), it is clear that $V_t^b(S_t, b_t)$ is nondecreasing in R_t, L_t, and b_{t-1}. Now, by the previous proposition, we see that the term inside the maximum of (1) is nondecreasing in R_t, L_t, and b_{t-1} for any action b_t. Hence, we can take the pointwise maximum and retain monotonicity; the inductive step is complete. □

Proposition 5. *The post–decision value function* $V_t^b(S_t^b)$*, with* $S_t^b = (R_t, L_t, b_{t-1}, b_t, P_t^S)$ *is nondecreasing in* R_t*,* L_t*,* b_{t-1}^-*,* b_{t-1}^+*,* b_t^-*, and* b_t^+*.*

Proof. Previously in the proof of Proposition 4, we argued that $V_t^b(S_t^b)$ is monotone in R_t, L_t, b_{t-1}^-, and b_{t-1}^+. To see the monotonicity in b_t^- and b_t^+, first fix an outcome of $P_{(t,t+1]}$ and $b_t, b_t' \in \mathcal{B}$, with $b_t \le b_t'$. Observe that if we let $(R_{t+1}, L_{t+1}, b_t, P_{t+1}^S) = S^M(S_t, b_t, P_{(t,t+1]})$, then $(R_{t+1}, L_{t+1}, b_t', P_{t+1}^S) = S^M(S_t, b_t', P_{(t,t+1]})$, with only the bid dimensions changed. Therefore,

$$S^M(S_t, b_t, P_{(t,t+1]}) \le S^M(S_t, b_t', P_{(t,t+1]}).$$

Thus, by Proposition 4, for a fixed S_t, any outcome of the price process $P_{(t,t+1]}$, and $b_t \le b_t'$,

$$V_{t+1}^*\big(S^M(S_t, b_t, P_{(t,t+1]})\big) \le V_{t+1}^*\big(S^M(S_t, b_t', P_{(t,t+1]})\big).$$

Hence, after taking expectations, we get the desired result: $V_t^b(S_t, b_t) \le V_t^b(S_t, b_t')$. □

Lemma 1. *Define deterministic bounding sequences* L_t^k *and* U_t^k *in the following way. Let* $U^0 = V^* + V_{\max} \cdot e$ *and* $L^0 = V^* - V_{\max} \cdot e$*, where* e *is a vector of ones. In addition,* $U^{k+1} = (U^k + HU^k)/2$ *and* $L^{k+1} = (L^k + HL^k)/2$*. Then, for each* $s \in \mathcal{S}^b$ *and* $t \le T-1$*,*

$$\begin{aligned} L_t^k(s) &\to V_t^b(s), \\ U_t^k(s) &\to V_t^b(s), \end{aligned}$$

where the limit is in k*.*

Proof. We first show that H satisfies the following properties:

(i) $V \leq V' \Longrightarrow HV \leq HV'$.
(ii) V^* is a unique fixed point of H, i.e., $HV^* = V^*$.
(iii) $HV - \eta e \leq H(V - \eta e) \leq H(V + \eta e) \leq HV + \eta e$, for $\eta > 0$.

Statement (i) is trivial and follows directly from the monotonicity of the max and expectation operators. Statement (ii) follows from the fact that the finite horizon dynamic program exhibits a unique optimal value function (and thus, post–decision value function as well) determined by the backward recursive Bellman equations. Statement (iii) is easy to see directly from the definition of H. Now, applying Lemma 4.6 of Bertsekas and Tsitsiklis [1996] gives us the desired limit result. □

Lemma 2. *U^k and L^k both satisfy the monotonicity property: for each t, k, and $s_1, s_2 \in \mathcal{S}^b$ such that $s_1 \preccurlyeq^b s_2$,*

$$\begin{aligned} U_t^k(s_1) &\leq U_t^k(s_2), \\ L_t^k(s_1) &\leq L_t^k(s_2). \end{aligned} \tag{2}$$

Proof. To show this, first note that given a fixed $t \leq T-2$ and *any* vector $Y \in \mathbb{R}^{|\mathcal{S}^b|}$ (defined over the post–decision state space) that satisfies the monotonicity property, it is true that the vector $h_t Y$, whose component at $s \in \mathcal{S}^b$ is defined using the post–decision Bellman recursion,

$$(h_t Y)(s) = \mathbf{E}\Big[\max_{b_{t+1} \in \mathcal{B}} \big\{C_{t+1,t+3}(S_{t+1}, b_{t+1}) + Y(S_{t+1}^b)\big\} | S_t^b = s\Big],$$

also obeys the monotonicity property. We point out that there is a small difference between the operator H and h_t in that H operates on vectors of dimension $T \cdot |\mathcal{S}^b|$. To verify monotonicity, $s_1, s_2 \in \mathcal{S}^b$ such that $s_1 \preccurlyeq^b s_2$. For a fixed sample path of prices P, let $S_{t+1}(s_1, P)$ and $S_{t+1}(s_2, P)$ be the respective downstream pre–decision states. Applying Propositions 1 and 2, we have that $S_{t+1}(s_1, P) \preccurlyeq^b S_{t+1}(s_2, P)$. For any fixed $b_{t+1} \in \mathcal{B}$, we apply the monotonicity of the contribution function $C_{t+1,t+3}$ (Proposition 3) and the monotonicity of Y to see that

$$C_{t+1,t+3}(S_{t+1}(s_1, P), b_{t+1}) + Y((S_{t+1}(s_1, P), b_{t+1})) \tag{3}$$
$$\leq C_{t+1,t+3}(S_{t+1}(s_2, P), b_{t+1}) + Y((S_{t+1}(s_2, P), b_{t+1})), \tag{4}$$

which confirms that $(h_t Y)(s_1) \leq (h_t Y)(s_2)$. When $t = T-1$, we set $(h_t Y)(s) = \mathbf{E}[C_{\text{term}}(S_{t+1}) | S_t^b = s]$ and the same monotonicity result holds.

Now, we can easily proceed by induction on k, noting that U^0 and L^0 satisfy monotonicity for each t. Assuming that U^k satisfies monotonicity, we can argue that U^{k+1} does as well; we first note that for any t, by the definition of U^{k+1},

$$U_t^{k+1} = \frac{U_t^k + \big(HU^k\big)_t}{2} = \frac{U_t^k + \big(h_t U_{t+1}^k\big)}{2}.$$

By the induction hypothesis and the property of h_t proved above, it is clear that U_t^{k+1} also satisfies monotonicity and the proof is complete. □

References

D. P. Bertsekas and J. N. Tsitsiklis. *Neuro–Dynamic Programming*. Athena Scientific, Belmont, MA, 1996.

W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley, 2nd edition, 2011.