OPERATIONS RESEARCH
Vol. 55, No. 2, March–April 2007, pp. 234–251
ISSN 0030-364X | EISSN 1526-5463 | 07 | 5502 | 0234

informs

DOI 10.1287/opre.1060.0355

A Framework Using Two-Factor Price Lattices for Generation Asset Valuation

Chung-Li Tseng
Department of Engineering Management and Systems Engineering, University of Missouri–Rolla, Rolla, Missouri 65409, chungli@umr.edu

Kyle Y. Lin
Operations Research Department, Naval Postgraduate School, Monterey, California 93943, kylin@nps.edu

In this paper, we use a real-options framework to value a power plant. The real option to commit or decommit a generating unit may be exercised on an hourly basis to maximize expected profit while subject to intertemporal operational constraints. The option-exercising process is modeled as a multistage stochastic problem. We develop a framework for generating discrete-time price lattices for two correlated Ito processes for electricity and fuel prices. We show that the proposed framework exceeds existing approaches in both lattice feasibility and computational efficiency. We prove that this framework guarantees existence of branching probabilities at all nodes and all stages of the lattice if the correlation between the two Ito processes is no greater than $4/\sqrt{35} \approx 0.676$. With price evolution represented by a lattice, the valuation problem is solved using stochastic dynamic programming. We show how the obtained power plant value converges to the true expected value by refining the price lattice. Sensitivity analysis for the power plant value to changes of price parameters is also presented.

Subject classifications: natural resources: energy; finance: asset pricing, real options.
Area of review: Environment, Energy, and Natural Resources.
History: Received January 2002; revisions received December 2002, June 2005, December 2005; accepted March 2006.

1. Introduction

Recently there have been substantial research activities involving real-options implementations for valuation in a variety of areas (Trigeorgis 1996). Formally, real options can be defined as the options embedded in real operational processes, activities, or investment opportunities that are not financial instruments. Given the popularity of the real-options approach, the core flexibility of a power plant, such as committing or decommitting a unit, has also been valued using such a "financial" methodology in the deregulated environment (Tseng and Barz 2002). The traditional unit-commitment problem (e.g., Sheble and Fahd 1994, Tseng 1996) can be viewed as an example of optimally exercising these operational real options to achieve cost minimization in the regulated environment, although in the "cost-of-service" world, savings from efficient dispatch typically accrue to the ratepayers. With deregulation and the introduction of spot markets for both electricity output and the fuel input, generation asset owners, whether they are a utility or a merchant operator, must reassess the value of their units, accounting for the market opportunity costs. Price information must be incorporated into the unit-commitment problems to capitalize on the profitable opportunities arising in the market. In doing so, utilities and power generators not only optimize their commitment decisions, taking into account price stochastics, but also maximize the market value of their power plants over the operating period.

In Deng et al. (1998) and Hsu (1998a, b), a power plant's valuation is appraised using real-option theory. The idea of these approaches is as follows: A power plant, with its associated heat rate, converts a particular fuel to electricity. This conversion involves two commodities with different market prices. When the electricity price is high, but the fuel price is low, the power plant should run to capitalize on the positive and profitable price spread between the price for power and the unit's cost of generation. If the price spread is negative, then the optimal decision is not to run the unit. Therefore, owning a power plant can be regarded as holding a series of call options of spark spreads, defined as the electricity price less the product of the heat rate associated with the generator and the fuel price. Analytical solutions using financial option theory can date back to Margrabe (1978) and Deng et al. (1998), who specifically applied them in the context of generation asset valuation. Although using real-option theory to value a power plant is a novel approach, most implementations have overlooked the power plant's operational constraints. Without considering the operational constraints, the power plant may be overvalued.

In this paper, we assume that there are hourly spot markets for both electricity and the fuel used by the generator, and that their prices follow some Ito processes. At each hour, the power plant operator must decide whether or not to exercise the option for running the unit. The unit operation is subject to decision lead times and minimum uptime and downtime constraints, so the commitment decision must take into account intertemporal effects. This option-exercising process is modeled as a multistage stochastic problem.

In Tseng and Barz (1999, 2002), the generation asset valuation problem with physical constraints was tackled using Monte Carlo (MC) simulation. That approach, similar to backward dynamic programming (DP) steps, applies MC simulation to determine optimal decision policies starting from the last period. The process is then repeated and moved backward: Having obtained all optimal decision policies for the subsequent time periods, MC simulation is applied to the current period. The advantage of this approach is its flexibility: MC simulation can easily adopt different kinds of price processes. It can also incorporate heuristics and simulate difficult constraints, such as unit ramp constraints. The major disadvantage is the computational speed. Because in the approach in Tseng and Barz (1999, 2002) the steps involve repeated backward-moving DP recursion and forward-moving MC simulation, the computation is massive and slow.

As opposed to the MC simulation, advantages of another approach using lattices for modeling uncertainty evolution have long been recognized. These advantages include its computational efficiency and its capability of handling exercise of American-type options. In our opinion, some important mathematical properties, such as existence of the price lattice and computational complexity, have been overlooked, which limits the applicability of the lattice approach. We show that traditional approaches, such as using decoupling techniques and the approach of Hull and White (1994), all have drawbacks. In this paper, we develop a framework for generating discrete-time trinomial price lattices for two general correlated Ito processes for electricity and fuel prices. Existence and convergence properties of the price lattice are analyzed. We derive a sufficient condition for guaranteeing existence of branching probabilities at all nodes and all stages of the lattice, that may not necessarily be met by the traditional approaches. We show that the proposed framework can satisfy the sufficient condition if the correlation between the two Ito processes is no greater than $4/\sqrt{35} \approx 0.676$, which is much higher than the correlation between the electricity and natural gas prices observed in the markets (e.g., Herbert 2001). Overall, the proposed framework exceeds existing approaches in both lattice feasibility and computational complexity. With price evolution represented by a lattice, the valuation problem is solved using stochastic DP.

To approximate an Ito process via a trinomial price lattice, we first divide the time horizon into small intervals. We then, in each time interval, use a three-point discrete-distribution approximation that matches the first and the second moment of the underlying Ito process. We further show that the proposed price lattice converges to the underlying Ito processes as the length of interval tends to zero. Therefore, the estimated power plant value in the price lattice will converge to the true expected value if we endure more computational effort. Our approach not only provides a better understanding of two-factor uncertainty lattices, but also has broad applications in valuation of various financial instruments.

This paper contributes to better understanding of mathematical properties of the two-factor lattices in the following three notions: (i) lattice compactness (boundedness of the size), (ii) lattice feasibility (existence of branching probabilities for any node at any stage), and (iii) lattice convergence to the underlying continuous correlated process. We show that for a class of mean reverting processes, the lattice size can be bounded. It is further shown that a lattice may be "optimized" by adjusting its cell sizes so as to achieve the feasibility and improve the convergence.

Deng and Oren (2003) proposed another approach for modeling a two-factor lattice, which has a different branching pattern such that branches are from one node to three other nodes, like a triangular pyramid. Their lattice, however, does not have the lattice compactness property, and its size grows as the time horizon increases.

The proposed generation asset valuation framework is intended to be applicable for general situations, which may include valuation over a long life cycle and/or subject to transmission constraints. Although we present the valuation framework with hourly time increments to capture physical constraints, we will also discuss how longer-term market information, such as forward curves, can be incorporated into the framework in the online appendix. How to consider transmission constraints, which are not formulated in the proposed model, will also be discussed in the online appendix.

In numerical tests, we apply the proposed lattice framework to two correlated geometric mean-reverting processes corresponding to electricity and fuel prices. The convergence property of the joint distribution, represented by the lattice, will be presented. The valuation model is then applied to a natural gas-fueled power plant. We demonstrate that the proposed method can converge to the theoretically true power plant expected value within acceptable CPU times. Sensitivity analysis for the power plant value to changes of price parameters is also presented.

This paper is organized as follows. In §2, we model the generation asset valuation problem as a multistage stochastic problem. The framework for modeling two correlated processes by a two-factor lattice is discussed in §3. We present numerical results in §4 and conclude this paper in §5. For brevity, all mathematical proofs are placed in the appendix of this paper, which is available in the online companion at http://or.pubs.informs.org/Pages/collect.html. Some implementation issues and potential extensions, such as considering ramp constraints, long-term valuation, and network constraints are also discussed in the online appendix.

2. Short-Term Generation Asset Valuation

A power plant consumes a particular fuel and then converts the fuel into electricity. This conversion involves two

commodities with different market prices. The operational constraints of a power plant considered in this paper include decision lead times, intertemporal constraints, variable heat rate, and startup costs, all of which complicate the optimal commitment decision making, especially under price uncertainties. The interested reader is referred to Tseng and Barz (2002) for an overview of these operational constraints.

2.1. The Mathematical Model for Power Plant Operation

This section presents a mathematical model of the characteristics of power plant operations. We first introduce the following standard notation. Additional symbols will be introduced when necessary.

t: index for time (in hours) ($t = 0, \ldots, T$).

τ: unit startup time, i.e., commitment decision lead time ($\tau \geqslant 1$, $\tau \in Z$, the set of all integers).

ν: unit shutdown time, i.e., decommitment decision lead time ($\nu \geqslant 1$, $\nu \in Z$).

x_t: state variable indicating the commitment status of the unit in time period t ($x_t \in Z$, $x_t \neq 0$).

u_t: zero-one generation unit commitment decision variable to be made at state x_t in time period t.

t^{on}: the minimum number of periods the unit must remain on after it has been turned on ($t^{\text{on}} \geqslant 1$, $t^{\text{on}} \in Z$).

t^{off}: the minimum number of periods the unit must remain off after it has been turned off ($t^{\text{off}} \geqslant 1$, $t^{\text{off}} \in Z$).

t^{cold}: the minimum number of periods required to cool down the boiler of a unit after it has been turned off ($t^{\text{cold}} \geqslant t^{\text{off}}$, $t^{\text{cold}} \in Z$).

q_t: decision variable indicating the amount of power the unit is generating in time period t.

$q^{\min}$: minimum rated capacity of the unit.

$q^{\max}$: maximum rated capacity of the unit.

R: ramp rate of the unit.

P_t^E: electricity price (\$/MWh) in time period t.

P_t^F: fuel price (\$/MMBtu) in time period t.

$H(q)$: heat requirement (MMBtu/MWh) of the unit at output level q (MWh).

$S_u(x_t)$: startup cost associated with turning on the unit at state x_t in time period t.

S_d: shutdown cost associated with turning off the unit.

Next, we present major assumptions made in this paper on unit operations, A1 to A4, and price models, A5 to A6. Further justification for some assumptions will be given later when they are cited.

Assumptions

A1. *The unit operation is not subject to ramp constraints. We, however, will discuss the impact of this assumption in the online appendix.*

A2. *An online (spinning) unit can be dispatched instantaneously and optimally.*

A3. *The unit commitment decisions are made on an hourly basis.*

A4. *The operator is a price taker, i.e., he has no market power to influence market prices by his unit commitment decisions. Therefore, the electricity and fuel prices are exogenous to the unit commitment decision models. Furthermore, the operator is risk neutral.*

A5. *There are hourly spot markets for both electricity and fuel used by the generator, and their prices P_t^E and P_t^F, respectively, follow some continuous correlated diffusion processes. However, in conjunction with* A3*, the prices are taken at discrete time periods in our implementation.*

A6. *Both P_t^E and P_t^F are actual price models, instead of risk-neutral price models.*

Let Φ be the set of states for unit commitment. Φ is composed of four subsets of states: Φ_1 for the startup period, Φ_2 for the normal online period, Φ_3 for the shutdown period, and Φ_4 for the normal offline period.

$$x_t \in \Phi = \Phi_1 \cup \Phi_2 \cup \Phi_3 \cup \Phi_4, \quad \text{where} \tag{1}$$

$$\Phi_1 \equiv \{1, 2, \ldots, \tau\}, \tag{2a}$$

$$\Phi_2 \equiv \{\tau + 1, \tau + 2, \ldots, \tau + t^{\text{on}}\}, \tag{2b}$$

$$\Phi_3 \equiv \{-1, -2, \ldots, -\nu\}, \tag{2c}$$

$$\Phi_4 \equiv \{-\nu - 1, \ldots, -\nu - t^{\text{off}}, \ldots, -\nu - t^{\text{cold}}\}. \tag{2d}$$

Basically, unit startup time and shutdown time are captured in Φ_1 and Φ_3, respectively; and the unit minimum uptime and downtime are captured in Φ_2 and Φ_4, respectively. This model is consistent with the one proposed in Svoboda et al. (1997).

Decision Variables. At each state x_t at time t, a commitment decision, denoted by $u_t \in \{0, 1\}$, must be made to transition x_t to some state x_{t+1} at time $t + 1$. Note that in this paper, the commitment u_t is a decision to be made at state x_t, which is somewhat different from the traditional unit commitment formulation in which u_t and x_t have some correspondence and are determined simultaneously. This change reflects the dynamics of decision making under uncertainty. The commitment decision u_t must be made in compliance with the so-called minimum uptime/downtime constraints, that are described in the following equation:

$$u_t = \begin{cases} 1 & \text{if } 1 \leqslant x_t < \tau + t^{\text{on}}, \\ 0 & \text{if } -\nu - t^{\text{off}} < x_t \leqslant -1, \\ 0 \text{ or } 1 & \text{otherwise.} \end{cases} \tag{3}$$

Equation (3) describes that the unit must remain online if it has not been online for more than t^{on} hours, and must remain offline if it has not been offline for at least t^{off} hours. In other cases, i.e., $x_t \in \{\tau + t^{\text{on}}\} \cup \{x \in Z \mid -\nu - t^{\text{cold}} \leqslant x \leqslant -\nu - t^{\text{off}}\}$, turning the unit on or off is an option. The commitment decision u_t, once it has been made, will yield the following state transitions.

State Transition Constraints

$$x_{t+1} = \begin{cases} \min(\tau + t^{\text{on}}, x_t + 1) \\ \qquad \text{if } x_t \in \Phi_1 \cup \Phi_2 \text{ and } u_t = 1, \\ -1 \quad \text{if } x_t = \tau + t^{\text{on}} \text{ and } u_t = 0, \\ \max(-\nu - t^{\text{cold}}, x_t - 1) \\ \qquad \text{if } x_t \in \Phi_3 \cup \Phi_4 \text{ and } u_t = 0, \\ 1 \quad \text{if } x_t \leqslant -\nu - t^{\text{off}} \text{ and } u_t = 1. \end{cases} \tag{4}$$

Dispatch Constraints

If $x_t \in \Phi_1$,

$$q_t = q^{\min} x_t / \tau. \tag{5}$$

If $x_t \in \Phi_2$,

$$|q_t - q_{t-1}| \leqslant R, \tag{6}$$

$$q^{\min} \leqslant q_t \leqslant q^{\max}. \tag{7}$$

If $x_t \in \Phi_3$,

$$q_t = q^{\min}(1 + x_t/\nu). \tag{8}$$

If $x_t \in \Phi_4$,

$$q_t = 0. \tag{9}$$

Note that (5) implies that when the unit is turned on, it takes τ hours with fixed increasing rate in generation $q^{\min}/\tau$ to reach the minimum rated capacity of the unit. Therefore, $q_t = q^{\min}$ when $x_t = \tau$. Similarly, (8) implies that $q_t = 0$ when $x_t = -\nu$. That means when an online unit is turned off, its generation level has to be reduced to $q^{\min}$, then from $q^{\min}$ it takes ν hours with fixed decreasing rate in generation $q^{\min}/\nu$ to reduce to zero. To sum up, during the startup and shutdown periods, the unit continues to generate power, but not as much as in the normal online period Φ_2. To distinguish the difference, we refer to the online period the states in Φ_3 only.

The constraints specified in (6) are the so-called (operating) ramp constraints, which limit the capacity changes of a generation from one hour to the next. Because the ramp constraints can be handled by the approach to be proposed in the paper, with the state space expanded from $\{x_t\}$ to include the (discrete) generation levels $\{q_t\}$, we will not specifically include the ramp constraints in the development (A1). Certainly, the ramp constraints increase computational effort (Tseng and Barz 2002). Because the ramp constraints will not be imposed, they may be violated. We will discuss how to relieve the impact due to the violation of the ramp constraints in the online appendix.

Initial Conditions. The unit operation is subject to some initial condition on x_t at $t = 0$. This initial value is denoted by $\tilde{x}_0$.

There are costs (or profits) associated with state transition and state visit. Let $S(\cdot)$ denote the transition costs and $\pi(\cdot)$ denote the profit collected from visiting states. Transition costs S include startup cost and shutdown cost, and can be represented as

$$S(x_t, u_t) = \begin{cases} S_u(x_t) & \text{if } x_t < 0 \text{ and } u_t = 1, \\ S_d & \text{if } x_t > 0 \text{ and } u_t = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

which captures the transition cost due to the commitment decision u_t made at state x_t. The profit associated with a state x_t is

$$\pi(x_t, q_t; P_t^E, P_t^F) = \begin{cases} P_t^E q_t - H(q_t) P_t^F & \text{if } x_t > 0, \\ 0 & \text{if } x_t < 0, \end{cases} \tag{11}$$

where q_t is the generation level at time t. In this paper, it is assumed that the dispatch problem can be solved instantaneously and optimally after the prices are revealed (A2). When a unit is in the normal online period at time t, i.e., $x_t \in \Phi_2$, its optimal dispatch g_t is defined as follows:

$$g_t \equiv \arg\max\{P_t^E q_t - H(q_t) P_t^F \mid q^{\min} \leqslant q_t \leqslant q^{\max}\} \quad \forall\, x_t \in \Phi_2. \tag{12}$$

Below, the dispatch q_t in the profit function π will be replaced by its optimal value g_t.

Let $F_t(x_t; P_t^E, P_t^F)$ be the power plant value at time t and state x_t with observed electricity and fuel prices (P_t^E, P_t^F). According to A4, the operator is risk neutral and the recurrence equations can be formulated as follows:

$$\begin{aligned} &F_t(x_t; P_t^E, P_t^F) \\ &= \pi(x_t, g_t; P_t^E, P_t^F) \\ &\quad + \max_{u_t}(E_t[e^{-r} F_{t+1}(x_{t+1}; P_{t+1}^E, P_{t+1}^F)] - S(x_t, u_t)), \end{aligned} \tag{13}$$

where E_t denotes the expectation operator, the subscript t indicates the expectation is conditioned on the price information available at time t, and r is the discount rate. Equation (13) is subject to the state transition constraints (3) and (4), and the dispatch constraints (5) to (9). The boundary conditions are

$$F_T(x_T; P_T^E, P_T^F) = \pi(x_T, g_T; P_T^E, P_T^F). \tag{14}$$

The optimal value J^* representing the power plant value over the period $[0, T]$ can be obtained from the last step of the recursive relation as

$$J^* = F_0(\tilde{x}_0; P_0^E, P_0^F). \tag{15}$$

In this paper, we propose using stochastic dynamic programming (SDP) to solve (15). To do so, we need to model price uncertainties on the discrete-time space. This will be detailed in the following sections.

In (13), the expectation is taken based on actual price processes instead of risk-neutral ones, and r is a discount rate that reflects the risk preference of the decision maker. In general, the risk-neutral measure of an uncertainty is obtained by referring to some "twin portfolio" implied by existing derivative securities that has equivalent riskiness. In this paper, based on A6, we utilize actual price processes instead of risk-neutral price processes because we suppose that the model will be mostly applied to determine the impact of operational constraints over a shorter term (such as days to weeks) rather than the monthly averages upon which derivative contracts are typically based. If a much longer term (such as months or years) will be considered, one can apply a constant adjustment to the commodity price process mean to obtain the risk-neutral price process, with the magnitude of the adjustment determined from the appropriate commodity derivatives (e.g., futures). In that situation, the discount rate will be the risk-free rate.

2.2. The SDP Approach

Consider using a price tree to represent the evolution of prices (P_t^E, P_t^F), $t \in [0, T]$. We will refer to each node in the tree as a *leaf node*, or simply, a *leaf*. Suppose that all the leaf nodes belonging to each time t are arranged by some ordered index set $I(t)$. Each leaf is denoted using a superscript index by $(P_t^{E,i}, P_t^{F,i})$, $i \in I(t)$. Let $A_t(i)$ be the index set of *descendants* for a leaf $(P_t^{E,i}, P_t^{F,i})$, $i \in I(t)$, such that $A_t(i) \subset I(t+1)$. For each price branch (arc) connecting $(P_t^{E,i}, P_t^{F,i})$, $i \in I(t)$, and $(P_{t+1}^{E,j}, P_{t+1}^{F,j})$, $j \in A_t(i)$, there is an associated probability $p_t^{i,j} \geqslant 0$, indicating the chance that the transition of prices represented by this arc will occur, given the leaf $(P_t^{E,i}, P_t^{F,i})$ is visited. Therefore,

$$\sum_{j \in A_t(i)} p_t^{i,j} = 1 \quad \forall\, i \in I(t),\ \forall\, t. \tag{16}$$

With the price tree, (15) can be solved using standard SDP backward steps. For every state $x_t \in \Phi$ at each leaf node $(P_t^{E,i}, P_t^{F,i})$ in the price tree, (13) can be rewritten as

$$F_t(x_t; P_t^{E,i}, P_t^{F,i}) = \pi(x_t, g_t; P_t^{E,i}, P_t^{F,i}) + \max_{u_t} e^{-r}\left[\sum_{j \in A_t(i)} p_t^{i,j} F_{t+1}(x_{t+1}; P_{t+1}^{E,j}, P_{t+1}^{F,j}) - S(x_t, u_t)\right], \tag{17}$$

subject to state transition constraints (3) and (4) and the dispatch constraints (5) to (9).

Thus far, we have assumed that the time step of the price tree is in hours, and there is a commitment decision to make at each node and each stage in the SDP formulation (17). However, sometimes it may be desirable to have the time step of a tree shorter than an hour, say 10 minutes, to better capture price movements. To do so, we introduce a new index $k = 1, \ldots, K$ for K subintervals in every hourly time period $[t, t+1]$ such that the kth subinterval corresponds to time period $[t + (k-1)\Delta t, t + k\Delta t]$, where $\Delta t = 1/K$ (hour). Accordingly, the SDP formulation given in (17) requires some modification because now a commitment decision is not made at every price node of the tree, but only at those subintervals corresponding to the beginning of an hour, i.e., $k = 1$. To indicate the corresponding subperiod of a node, we use an additional subscript k to variables whenever necessary. Note that the state of the unit at each time t is held constant for all k, but $k = 1$. For each leaf node $(P_{t,k}^{E,i}, P_{t,k}^{F,i})$ in the price tree, the refined SDP recursive relations are summarized as follows.

If $k = 1$,

$$F_{t,k}(x_t; P_{t,k}^{E,i}, P_{t,k}^{F,i}) = \pi(x_t, g_{t,k}; P_{t,k}^{E,i}, P_{t,k}^{F,i}) + \max_{u_t} e^{-r/K}\left[\sum_{j \in A_{t,k}(i)} p_{t,k}^{i,j} F_{t,k+1}(\tilde{x}_t; P_{t,k+1}^{E,j}, P_{t,k+1}^{F,j}) - S(x_t, u_t)\right]; \tag{18a}$$

if $1 < k < K$,

$$F_{t,k}(x_t; P_{t,k}^{E,i}, P_{t,k}^{F,i}) = e^{-r/K} \sum_{j \in A_{t,k}(i)} p_{t,k}^{i,j} F_{t,k+1}(x_t; P_{t,k+1}^{E,j}, P_{t,k+1}^{F,j}); \tag{18b}$$

and if $k = K$,

$$F_{t,k}(x_t; P_{t,k}^{E,i}, P_{t,k}^{F,i}) = e^{-r/K} \sum_{j \in A_{t,k}(i)} p_{t,k}^{i,j} F_{t+1,1}(x_{t+1}; P_{t+1,1}^{E,j}, P_{t+1,1}^{F,j}). \tag{18c}$$

Note that the maximization in (18a) is subject to state transition constraints (3) and (4) with x_{t+1} replaced by $\tilde{x}_t$, and the dispatch constraints (5) to (9). That is, in (18a), $\tilde{x}_t$ is the state of the unit at time $t+1$. It is obvious that (18a) is the direct extension of (17) because a commitment decision is only considered at the beginning of each hour ($k = 1$). For the other subperiods $1 < k \leqslant K$, (18b) and (18c) only perform recursive discounting.

The SDP approach with refined subintervals proposed in this section can be viewed as a means to better approximate the discrete-time probability distribution of the next hour's price. It can be verified that within each subperiod $1 < k \leqslant K$, no profit is collected and (18b) and (18c) only perform recursive discounting. However, a refined time step may also be used to capture real-time price movements. In reality, the power plant may be compensated ex post based on some real-time prices. This situation can be handled in the proposed model by simply including profit collection in the recursive formulation, (18b) and (18c), for the subintervals.

3. Modeling Two-Factor Price Lattices

In this paper, we assume that the prices of electricity P_t^E and fuel consumed by the generator P_t^F are some smooth functions of y_1 and y_2, respectively, which are governed by the following stochastic differential equations:

$$dy_1 = \mu_1(y_1)dt + \sigma_1 dB_1 \tag{19}$$

and

$$dy_2 = \mu_2(y_2)dt + \sigma_2 dB_2, \tag{20}$$

where μ_1 and μ_2 are drift functions, σ_1 and σ_2 are constant volatilities, and B_1 and B_2 are two Wiener processes with instantaneous correlation ρ. For example, setting $P_t^E = \exp(y_1)$ and $P_t^F = \exp(y_2)$ is a common approach to avoid getting negative prices. In the online appendix we will extend the proposed valuation model to accommodate the volatilities σ_1 and σ_2 in (19) and (20) that are time dependent.

While we will develop two-factor price lattices that can be applied to general processes of the forms in (19) and (20), special attention will be paid to mean-reverting (MR) processes because they have been commonly used to represent energy price movements (e.g., Barz 1999). An MR process is defined below.

Mean-Reverting (MR) Process. The stochastic process y is an MR process if the drift function $\mu(y)$ is Lipschitz continuous and monotonically decreasing in y, and

$$\mu(y) \to \infty \quad \text{as } y \to -\infty \ \forall t, \tag{21a}$$

$$\mu(y) \to -\infty \quad \text{as } y \to \infty \ \forall t. \tag{21b}$$

The main task in this section is to represent the continuous price processes in (19) and (20) by an equivalent discrete-time price tree defined in §2.2. One property of such a general tree is that there exists a unique, simple path between every pair of distinct nodes in a tree, i.e., branches (arcs) do not recombine. This property implies that the number of branches may easily increase exponentially as the number of stages increases. Even for a simple 24-hour case in our application, the corresponding price tree would become too bushy to be manageable. Therefore, we focus on *lattices*, which are basically the same as trees except that branches are allowed to recombine. This will greatly reduce computational complexity. It can be verified that the notations for a tree defined in §2.2 are independent of the fact whether or not the branches would recombine, and hence all the SDP formulations in the same section remain valid. In the remainder of this paper, we thus use the term lattice(s) instead of tree(s).

Before we proceed to model the two-factor lattices, we review the one-factor case first.

3.1. One-Factor Trinomial Price Lattices

Consider the following Ito process:

$$dy = \mu(y)dt + \sigma dB, \tag{22}$$

where B is a Wiener process. While this process is a continuous-time model, it is desirable to derive a discrete-time model that can be used to approximate the evolution of $y(t)$. The trinomial lattice, initially proposed in Hull and White (1993) to model MR processes, can be applied to represent a general process of the form in (22). With the time horizon being divided into intervals of equal length Δt, the process can only take on values that are multiples of h, where $h = c\sigma\sqrt{\Delta t}$. Hull and White (1993) suggested the constant c to be $\sqrt{3}$ for an MR process. In §§3.3–3.5, we shall show that the choice of c may affect the feasibility and convergence properties of a price lattice.

In a trinomial lattice, a price node y branches into nodes $y+(\kappa+1)h$, $y+\kappa h$, and $y+(\kappa-1)h$ at the next stage with respective branching probabilities p_u, p_m, and p_d, where κ is an integer. The value of κ is chosen such that κh is closest to the expected drift $\mu(y)\Delta t$. That is,

$$\kappa \equiv \left\lfloor \frac{\mu(y)\Delta t}{h} + \frac{1}{2} \right\rfloor, \tag{23}$$

where $\lfloor\cdot\rfloor$ is the so-called floor function, which maps a real number to the nearest integer toward $-\infty$. The branching patterns corresponding to $\kappa = 1$, 0, -1, and -2 are depicted in Figure 1.

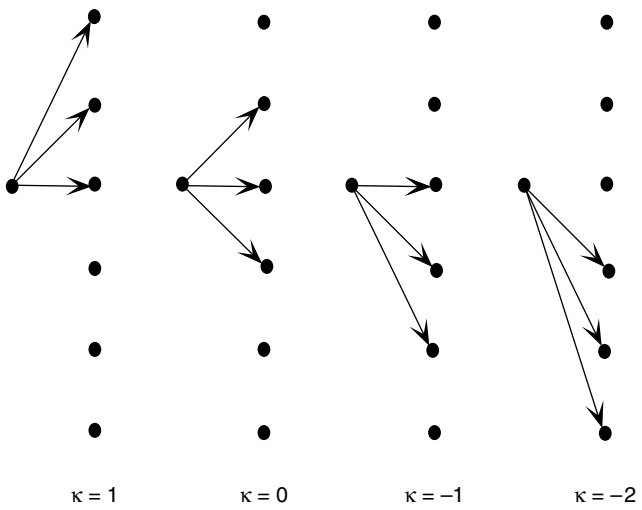
To solve the branching probabilities p_u, p_m, and p_d, we consider the following linear equations:

$$p_u(\kappa+1)h + p_m\kappa h + p_d(\kappa-1)h = \mu(y)\Delta t, \tag{24a}$$

$$p_u(\kappa+1)^2h^2 + p_m\kappa^2h^2 + p_d(\kappa-1)^2h^2 = \sigma^2\Delta t + \mu(y)^2\Delta t^2, \tag{24b}$$

$$p_u + p_m + p_d = 1. \tag{24c}$$

Figure 1. Alternative branching in a trinomial model.



Equations (24a) and (24b) intend to match the local mean and variance of the Ito process, respectively. The solution of the system of equations is

$$p_{\mathrm{u}} = \frac{1}{2}\left(\frac{1}{c^2} + \epsilon + \epsilon^2\right), \tag{25a}$$

$$p_{\mathrm{d}} = \frac{1}{2}\left(\frac{1}{c^2} - \epsilon + \epsilon^2\right), \tag{25b}$$

$$p_{\mathrm{m}} = 1 - \frac{1}{c^2} - \epsilon^2, \tag{25c}$$

where we define

$$\epsilon \equiv \frac{\mu(y)\Delta t}{h} - \kappa. \tag{26}$$

From (23), it is obvious that

$$|\epsilon| \leqslant \frac{1}{2}. \tag{27}$$

Note that ϵ is a function of y and t. Furthermore, it can be shown that the branching probabilities in (25a) to (25c) are always legitimate (between zero and one) if

$$\frac{1}{4} \leqslant \frac{1}{c^2} \leqslant \frac{3}{4} \tag{28}$$

or $2/\sqrt{3} \leqslant c \leqslant 2$.

To approximate the continuous stochastic process $\{y(t) \mid 0 \leqslant t \leqslant T\}$, consider an N-stage trinomial lattice $\{Y_n \mid n = 0, \ldots, N\}$ with the initial condition $Y_0 = y(0)$. Let $\Delta t = T/N$ and $h = c\sigma\sqrt{\Delta t}$, where c is a constant satisfying (28). Using (23) to determine the branching factor at each node, and (25a) to (25c) to compute the branching probabilities, the weak convergence theorem (Kushner 1984) guarantees that the approximation becomes exact as $N \to \infty$.

When price $\mu(y)$ satisfies the MR properties, the size of the trinomial lattice $\{Y_n \mid n = 0, \ldots, N\}$ will not grow indefinitely, i.e., the lattice size $\max_{n=0,\ldots,N}|Y_n|$ is capped by a finite number. That is, a finite interval for y exists for each time period n, say $[y_n^{\min}, y_n^{\max}]$, such that any lattice node within this interval only branches into nodes in the same interval corresponding to the following time period. Therefore, if at time 0, the initial price $y(0) \in \bigcup_{n=0}^{N}[y_n^{\min}, y_n^{\max}]$, the lattice size will be capped. This is because of the reverting property: For each n, when y approaches $y_n^{\max}$ (from below), y will revert down, and vice versa. Finally, note that even if $y(0) \notin \bigcup_{n=0}^{N}[y_n^{\min}, y_n^{\max}]$, the lattice size can still be shown to be capped in $\bigcup_{n=0}^{N}[\min(y(0), y_n^{\min}), \max(y(0), y_n^{\max})]$.

Note that although the lattice size can be capped, it does not mean that the price is necessarily bounded. It simply implies that the probability that the price is outside the bounds is negligible due to the discretization nature. Furthermore, the (discretized) price upper and lower bounds approach ∞ and $-\infty$, respectively, as $N \to \infty$. That is, the discrete distribution represented by the lattice approaches its real, underlying continuous distribution.

3.2. Two-Factor Price Lattices Using Decoupling

Consider the following linear transformation:

$$w_1 = \sigma_2 y_1 - \sigma_1 y_2, \tag{29a}$$

$$w_2 = \sigma_2 y_1 + \sigma_1 y_2. \tag{29b}$$

We have

$$\begin{aligned} dw_1 &= \sigma_2 dy_1 - \sigma_1 dy_2 \\ &= (\sigma_2\mu_1(y_1) - \sigma_1\mu_2(y_2))dt + \sigma_1\sigma_2(dB_1 - dB_2). \end{aligned} \tag{30}$$

Using the fact that

$$y_1 = \frac{w_1 + w_2}{2\sigma_2} \quad \text{and} \quad y_2 = \frac{w_2 - w_1}{2\sigma_1}, \tag{31}$$

(30) is reduced to

$$dw_1 = \hat{\mu}_1(w_1, w_2)dt + \sigma_1\sigma_2\sqrt{2(1-\rho)}dB_3, \tag{32}$$

where the drift term $\hat{\mu}_1$ is a function of w_1, w_2, and t, and $(dB_1 - dB_2)$ has been normalized and denoted in terms of another Wiener process dB_3. Similarly, we have

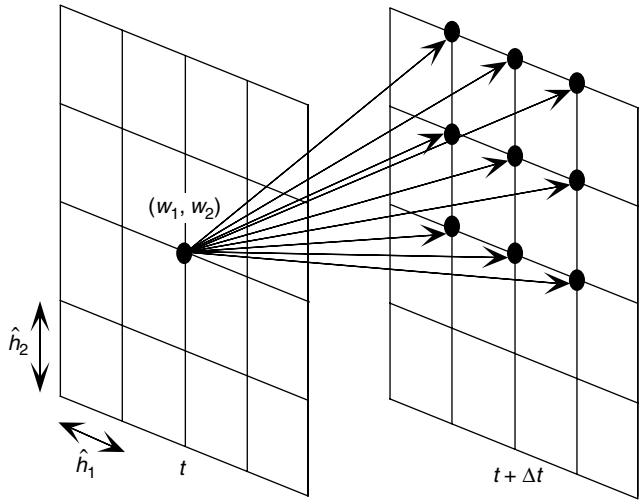
$$\begin{aligned} dw_2 &= \sigma_2 dy_1 + \sigma_1 dy_2 \\ &= (\sigma_2\mu_1(y_1) + \sigma_1\mu_2(y_2))dt + \sigma_1\sigma_2(dB_1 + dB_2) \\ &= \hat{\mu}_2(w_1, w_2)dt + \sigma_1\sigma_2\sqrt{2(1+\rho)}dB_4. \end{aligned} \tag{33}$$

It can be shown that B_3 and B_4 are uncorrelated Wiener processes. Note that we apply a hat to some notations to indicate that these notations are defined in the transformed (w_1, w_2) domain, as opposed to their counterparts defined in the (y_1, y_2) domain.

To create a two-factor price lattice, one can directly extend the one-factor model such that each two-factor price node branches into $3 \times 3 = 9$ nodes in a predetermined price lattice of the next time period. A sample branching is given in Figure 2. In this two-factor lattice, $\hat{h}_1 = c_1\sigma_1\sigma_2\sqrt{2(1-\rho)\Delta t}$ and $\hat{h}_2 = c_2\sigma_1\sigma_2\sqrt{2(1+\rho)\Delta t}$ and $2/\sqrt{3} \leqslant c_1, c_2 \leqslant 2$. Given a node (w_1, w_2), one can first use (23) to obtain a $\hat{\kappa}_1$ for w_1, then use (25a) to (25c) to determine the branching probabilities $(\hat{p}_{1\mathrm{u}}, \hat{p}_{1\mathrm{m}}, \hat{p}_{1\mathrm{d}})$, and, similarly, obtain $\hat{\kappa}_2$ and $(\hat{p}_{2\mathrm{u}}, \hat{p}_{2\mathrm{m}}, \hat{p}_{2\mathrm{d}})$. Because both B_3 and B_4 are uncorrelated, the nine combinations of the product of $(\hat{p}_{1\mathrm{u}}, \hat{p}_{1\mathrm{m}}, \hat{p}_{1\mathrm{d}})$ and $(\hat{p}_{2\mathrm{u}}, \hat{p}_{2\mathrm{m}}, \hat{p}_{2\mathrm{d}})$, together with $\hat{\kappa}_1$ and $\hat{\kappa}_2$, identify the nine nodes that are branched into in the next time period, and the corresponding branching probabilities. Note that even though B_3 and B_4 are independent, w_1 and w_2 may not be independent because of the MR components. In Figure 2, the example branching corresponds to $\hat{\kappa}_1 = 0$ and $\hat{\kappa}_2 = 1$.

At each node of the two-factor lattice in the (w_1, w_2) domain, one can obtain the corresponding (y_1, y_2) using (31), and based on this, one can value the power plant as if it were in the (y_1, y_2) domain.

Figure 2. A sample two-factor trinomial branching with $\hat{\kappa}_1 = 0$, $\hat{\kappa}_2 = 1$.



The decoupling method is generally applicable for generating two-factor lattices. However, when both y_1 and y_2 are MR, the lattice generated in the (w_1, w_2) domain will not be capped as in the one-factor case. To show this, consider the branching factor $\hat{\kappa}_1$ for w_1 such that

$$\hat{\kappa}_1 = \left\lfloor \frac{(\sigma_2\mu_1(y_1) - \sigma_1\mu_2(y_2))\Delta t}{\hat{h}_1} + \frac{1}{2} \right\rfloor. \tag{34}$$

Because both $\mu_1(y_1)$ and $\mu_2(y_2)$ are continuous and monotonically decreasing, it can be shown that at any time t, for any y_1, there exists a y_2 such that

$$\sigma_2\mu_1(y_1) = \sigma_1\mu_2(y_2), \tag{35}$$

i.e., $\hat{\kappa}_1 = 0$. Because (w_1, w_2) is merely a linear transformation of (y_1, y_2), for any w_1, there exists a w_2 such that the branching factor $\hat{\kappa}_1 = 0$ at that point. That is, for each w_1, there exists a w_2 at which the lattice is expanding. Therefore, at any given time t, there does not exist a closed and bounded region in the transformed (w_1, w_2) domain, such that any lattice node inside the region only branches into nodes in the same region corresponding to the following time period.

For any two correlated MR processes, the lattice generated using transformation will not be computationally efficient if the problem involves a large number of time periods. The number of nodes increases at least quadratically, as the number of time periods increases. For a valuation problem involving one week (168 hourly time periods), computation and memory requirement will be intense. In the next section, we will propose a new approach for generating two-factor price lattices, which is computationally efficient for generating branching probabilities, and whose lattice size is capped under the assumption of correlated MR processes.

3.3. A New Approach for Two-Factor Price Lattices

To have a two-factor price lattice capped in size for two correlated MR processes, we build the lattice in the (y_1, y_2) domain directly because the size of each one-factor lattice in the price domain can be capped. Similar to what is performed for creating a two-factor lattice in §3.2, each price node branches into $3 \times 3 = 9$ nodes in a predetermined price lattice of the next time period. Denote the two-factor N-stage trinomial lattice as $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$, which approximates the continuous stochastic processes $\{(y_1(t), y_2(t)) \mid 0 \leqslant t \leqslant T\}$ over a given time period $[0, T]$ with an initial condition $(Y_{1,0}, Y_{2,0}) = (y_1(0), y_2(0))$.

Let $\Delta t = T/N$, $h_1 = c_1\sigma_1\sqrt{\Delta t}$, and $h_2 = c_2\sigma_2\sqrt{\Delta t}$, where both c_1 and c_2 are constants satisfying (28). That is,

$$2/\sqrt{3} \leqslant c_1, c_2 \leqslant 2. \tag{36}$$

Given any node (y_1, y_2) at stage n, we first use (23) to obtain κ_1 for Y_1, then (25a) to (25c) to obtain the one-factor branching probabilities $(\tilde{p}_{1\mathrm{u}}, \tilde{p}_{1\mathrm{m}}, \tilde{p}_{1\mathrm{d}})$, and similarly obtain κ_2 and $(\tilde{p}_{2\mathrm{u}}, \tilde{p}_{2\mathrm{m}}, \tilde{p}_{2\mathrm{d}})$ for Y_2. Note that the *tilde* associated with each of these (one-factor) branching probabilities denotes the realization of (25a) to (25c) and is used to distinguish them from the variables of branching probabilities to be determined next. To simplify the notation, let $\Omega \equiv \{\mathrm{u}, \mathrm{m}, \mathrm{d}\}$ be the index set of (text) subscripts, representing upward, middle, and downward branching, respectively. Let $\kappa_{1\mathrm{u}} = \kappa_1 + 1$, $\kappa_{1\mathrm{m}} = \kappa_1$, and $\kappa_{1\mathrm{d}} = \kappa_1 - 1$, and define $\kappa_{2\mathrm{u}}$, $\kappa_{2\mathrm{m}}$, and $\kappa_{2\mathrm{d}}$ similarly. Denote the following system of linear equations (with nonnegativity constraints) by (Q) for the node (y_1, y_2):

$$(Q) \quad \sum_{j\in\Omega} p_{\mathrm{u}j} = \tilde{p}_{1\mathrm{u}}, \tag{37a}$$

$$\sum_{j\in\Omega} p_{\mathrm{m}j} = \tilde{p}_{1\mathrm{m}}, \tag{37b}$$

$$\sum_{j\in\Omega} p_{j\mathrm{u}} = \tilde{p}_{2\mathrm{u}}, \tag{37c}$$

$$\sum_{j\in\Omega} p_{j\mathrm{m}} = \tilde{p}_{2\mathrm{m}}, \tag{37d}$$

$$\sum_{i,\,j\in\Omega} p_{ij}\kappa_{1i}\kappa_{2j}h_1h_2 = \rho\sigma_1\sigma_2\Delta t + \mu_1\mu_2\Delta t^2, \tag{37e}$$

$$\sum_{i,\,j\in\Omega} p_{ij} = 1, \quad p_{ij} \geqslant 0 \quad \forall\, i, j \in \Omega. \tag{37f}$$

In (Q), (37a) and (37b) intend to match the mean and variance for $y_1(t + \Delta t)$ conditioned on the value of $y_1(t)$, respectively. Because the one-factor branching probabilities $(\tilde{p}_{1\mathrm{u}}, \tilde{p}_{1\mathrm{m}}, \tilde{p}_{1\mathrm{d}})$ have already matched the conditional mean and variance for y_1, (37a) and (37b) is an equivalent representation, as are (37c) and (37d). Note that (Q) has nine variables, but only six linear equations; hence, it has infinitely many solutions if it is feasible. Next, we show that

to guarantee the feasibility of (Q), some relation between ρ and c_1 and c_2 must be satisfied.

By rearranging terms and using the results from (25a) to (25c), (37e) can be reduced to the following equation:

$$p_{uu} + p_{dd} - p_{ud} - p_{du} = \frac{\rho}{c_1 c_2} + \left(\frac{\mu_1 \Delta t}{h_1} - \kappa_1\right)\left(\frac{\mu_2 \Delta t}{h_2} - \kappa_2\right) = \frac{\rho}{c_1 c_2} + \epsilon_1 \epsilon_2, \tag{38}$$

where $\epsilon_1 \equiv (\mu_1(y_1)\Delta t/h_1 - \kappa_1)$ and $\epsilon_2 \equiv (\mu_2(y_2)\Delta t/h_2 - \kappa_2)$, definitions directly extended from (26). Recall from (27) that $|\epsilon_1| \leqslant 1/2$ and $|\epsilon_2| \leqslant 1/2$. Now consider the following LP, denoted by (P_1):

$$(P_1) \quad \text{optimize} \quad X \equiv p_{uu} + p_{dd} - p_{ud} - p_{du} \tag{39a}$$

$$\text{s.t.} \quad \sum_{j\in\Omega} p_{ij} = \tilde{p}_{1i} \quad \forall\, i \in \Omega, \tag{39b}$$

$$\sum_{i\in\Omega} p_{ij} = \tilde{p}_{2j} \quad \forall\, j \in \Omega, \tag{39c}$$

$$p_{ij} \geqslant 0 \quad \forall\, i, j \in \Omega. \tag{39d}$$

Note that (39b) implies $\sum_{i,\, j\in\Omega} p_{ij} = 1$, and so does (39c). Let $X_{\min}$ (or $X_{\max}$) be the optimal objective value of (P_1) if it is a minimization (or maximization) problem. We have the following lemma.

LEMMA 1. *(Q) is feasible if the following inequality holds*:

$$X_{\min} \leqslant \frac{\rho}{c_1 c_2} + \epsilon_1 \epsilon_2 \leqslant X_{\max} \tag{40}$$

for all $|\epsilon_1|, |\epsilon_2| \leqslant 1/2$.

PROOF. See the online appendix.

LEMMA 2. *Pertaining to (P_1), the following statements are true*:

(i) *If* $\tilde{p}_{1u} \leqslant \tilde{p}_{2u}$ *and* $\tilde{p}_{1d} \leqslant \tilde{p}_{2d}$, $X_{\max} = \tilde{p}_{1u} + \tilde{p}_{1d}$.

(ii) *If* $\tilde{p}_{1u} > \tilde{p}_{2u}$ *and* $\tilde{p}_{1d} > \tilde{p}_{2d}$, $X_{\max} = \tilde{p}_{2u} + \tilde{p}_{2d}$.

(iii) *If* $\tilde{p}_{1u} \leqslant \tilde{p}_{2u}$ *and* $\tilde{p}_{1d} > \tilde{p}_{2d}$, $X_{\max} = \min\{\tilde{p}_{1u} + \tilde{p}_{2d}, 1 - (\tilde{p}_{1d} - \tilde{p}_{2d}) - (\tilde{p}_{2u} - \tilde{p}_{1u})\}$.

(iv) *If* $\tilde{p}_{1u} > \tilde{p}_{2u}$ *and* $\tilde{p}_{1d} \leqslant \tilde{p}_{2d}$, $X_{\max} = \min\{\tilde{p}_{2u} + \tilde{p}_{1d}, 1 - (\tilde{p}_{2d} - \tilde{p}_{1d}) - (\tilde{p}_{1u} - \tilde{p}_{2u})\}$.

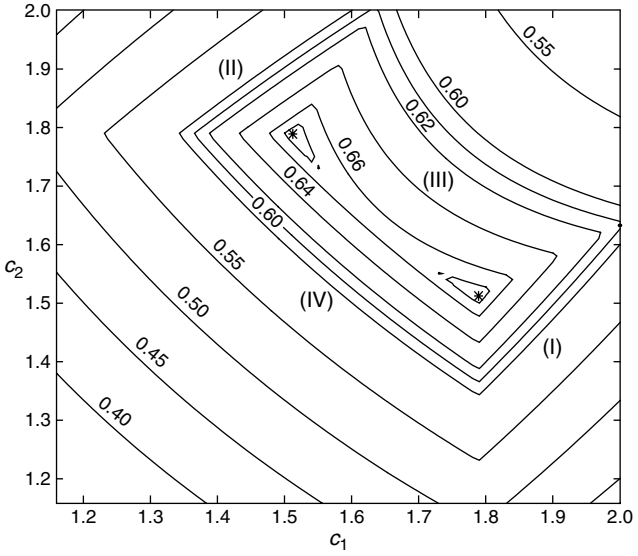
PROOF. See the online appendix.

The values of $X_{\max}$ and $X_{\min}$ in Lemmas 1 and 2 may vary from node to node. Next, we will show a (tightest) sufficient condition that guarantees existence of branching probabilities for *any* node at *any* stage. That is, this condition ensures feasibility of an entire lattice, regardless of the size of the lattice.

DEFINITION. A lattice $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$ is said to be *feasible* if (Q) is feasible at *all* price nodes at *all* stages.

For simplicity, in the sequel we refer to a lattice $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$ as one whose branching probabilities are generated by solving (Q).

Figure 3. The constant contours of $\rho^{\max}$ versus the values of c_1 and c_2.



THEOREM 3. *A lattice $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$ is feasible if*

$$|\rho| \leqslant \rho^{\max}(c_1, c_2) \equiv \min\left\{\frac{c_2}{c_1} - \frac{c_1 c_2}{16}, \frac{c_1}{c_2} - \frac{c_1 c_2}{16}, \frac{1}{2}\left(\frac{c_2}{c_1} + \frac{c_1}{c_2}\right) - \frac{c_1 c_2}{8}, \frac{c_1 c_2}{4}\right\}. \tag{41}$$

PROOF. See the online appendix.

In Figure 3, the constant contours of $\rho^{\max}$ are plotted versus the values of c_1 and c_2. The four regions, numbered from (I) to (IV) in Figure 3, correspond to the four functions in the braces of (41), in order of their appearance. Each region shows where the corresponding function prevails in the minimum operator in (41). The function $\rho^{\max}$ has two small and symmetric humps, and achieves the maximum $4/\sqrt{35} \approx 0.676$ at $(c_1, c_2) = (4/\sqrt{5}, 4/\sqrt{7})$ and $(4/\sqrt{7}, 4/\sqrt{5})$. Therefore, this two-factor trinomial lattice can be a good approximation of the continuous-time model, as long as the correlation ρ between two Wiener processes is less than 0.676, and c_1, c_2 are chosen such that $\rho^{\max}(c_1, c_2) \geqslant \rho$.

REMARK 1. Based on our tests over historical price data, the correlation between electricity and natural gas prices has constantly been observed to be between 0.2 and 0.3, which is much smaller than 0.676. Our observations are consistent with Herbert (2001), who reports a correlation of approximately 0.2. However, it should be noted that the variety of methods for accommodating seasonality, as well as the variety of data sets and methods for assessing the correlation itself (e.g., fixed window or exponential moving averages), will result in a range of observed correlation values. Therefore, the proposed framework is not only

well suited for generation asset valuation, but also for many other asset valuations involving two correlated Wiener processes.

REMARK 2. Until now, choosing the values of c_1 and c_2 has only been known to affect the "quality" of fitting probability distributions in terms of matching moments under some restricted conditions. Theorem 3 extends the knowledge of the quality measures of fitting probability distributions to include fitting feasibility, which is a prerequisite to other quality measures.

REMARK 3. If $\mu(y)$ is bounded for all y and t, a lattice can be made feasible by merely reducing $\Delta t \to 0$ for all $|\rho| < 1$. This is the approach suggested by Hull and White (1994), which, however, does not work for an MR process because its $\mu(y)$ is unbounded.

REMARK 4. It is possible that (Q) may still be feasible for some nodes in a lattice (possibly small) even when $|\rho| > \rho^{\max}$. Theorem 3 presents a sufficient condition that guarantees the feasibility of any lattice of any size. Therefore, $\rho^{\max}$ can be interpreted as the worst-case (or the most strict case) result because the theorem accommodates all possible correlated price processes of the forms in (19) and (20). On the other hand, $\rho^{\max}$ in (41) is the maximum possible ρ that can guarantee lattice feasibility. Therefore, Theorem 3 presents the tightest possible sufficient condition.

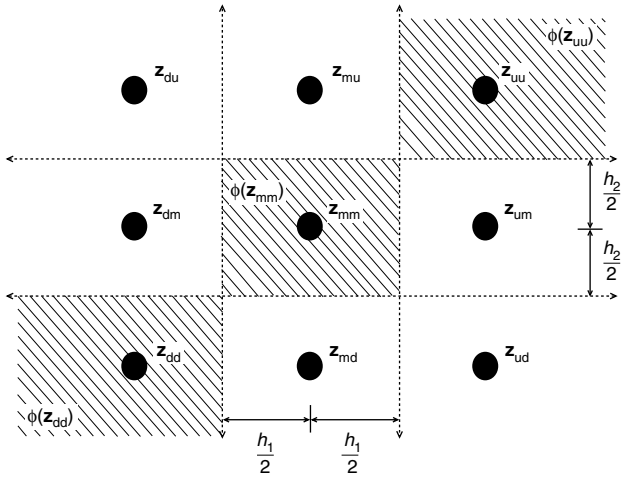
REMARK 5. Apparently, the size of the lattice generated by the proposed method is capped for two correlated underlying MR processes. Compared with the method using decoupling, which has no restriction on the correlation value, this improvement in computational complexity comes with a price of reducing $\rho^{\max}$ from 1 to 0.676. This trade-off is worthwhile for our application.

For each price node in the lattice, its branching probabilities are selected by any arbitrary feasible solution of the corresponding (Q). Two questions remain: whether the lattice converges and, if so, how to achieve a good convergence. By convergence of the lattice, we mean that the approximation of the joint distribution at time T would become exact as $N \to \infty$. The proof of the convergence is established in the online appendix.

The next issue is to measure the convergence. Consider a discrete random variable $\mathbf{Z} \in R^2$ with probability mass function $p(\mathbf{z}_l)$, where $\mathbf{z}_l \in R^2$, $l = 1, \ldots, L$, and L is the total number of the sample points. Suppose the continuous random variable that $\mathbf{Z}$ intends to approximate is $\hat{\mathbf{Z}} \in R^2$ with density function $f(\hat{\mathbf{z}})$. For each sample point $\mathbf{z}_l$, an area surrounding $\mathbf{z}_l$, denoted by $\phi(\mathbf{z}_l)$, can be assigned such that (i) all $\phi(\mathbf{z}_l)$, $l = 1, \ldots, L$, are mutually exclusive; and (ii) $\bigcup_{l=1}^{L} \phi(\mathbf{z}_l) = R^2$. Define the metric $d(p, f)$ as follows:

$$d(p, f) \equiv \sum_{l=1}^{L} \left(p(\mathbf{z}_l) - \int_{\phi(\mathbf{z}_l)} f(\mathbf{w})\, d\mathbf{w} \right)^2. \tag{42}$$

Figure 4. An example of allocating $\phi(\mathbf{z}_{ij})$.



In the proposed lattice in this paper, each node branches into nine nodes, denoted by $\mathbf{z}_{ij}$, $i, j \in \Omega$. An intuitive way for assigning surrounding area $\phi(\mathbf{z}_{ij})$ for each node $\mathbf{z}_{ij}$ is depicted in Figure 4. Two vertical lines and two horizontal lines are marked in the middle of two adjacent columns and rows, respectively. Three nodes ($\mathbf{z}_{dd}$, $\mathbf{z}_{mm}$, and $\mathbf{z}_{uu}$) and their corresponding $\phi(\mathbf{z}_{ij})$ are highlighted. Note that except for the "interior" node $\mathbf{z}_{mm}$, every other node has an unbounded $\phi(\mathbf{z}_{ij})$.

Two issues will be shown to affect the convergence: the determination of the branching probabilities for each lattice node and the lattice cell size (i.e., the values of c_1 and c_2). Note that from the proof given in the online appendix, the convergence of the lattice rests on the feasibility of (Q) at each lattice node, regardless of the selection of the solution of (Q), which typically has infinitely many feasible solutions. On the other hand, the feasibility of (Q) depends on the selection of c_1 and c_2. Intuitively, there are "good" feasible solutions of (Q) that, in combination with "good" sections of c_1 and c_2, can result in good convergence. These two issues will be discussed in the following sections.

3.4. Selecting Branching Probabilities

To determine a good, feasible solution of (Q) for good convergence, first consider a special case when y_1 and y_2 are independent (i.e., $\rho = 0$). The solution for (Q) is

$$p_{ij} = \tilde{p}_{1i}\tilde{p}_{2j} \quad \forall\, i, j \in \Omega. \tag{43}$$

Substituting (43) into (38) and applying (25a)–(25c) yields

$$\begin{aligned} &p_{\mathrm{uu}} - p_{\mathrm{ud}} - p_{\mathrm{du}} + p_{\mathrm{dd}} \\ &\quad = \tilde{p}_{1\mathrm{u}}\tilde{p}_{2\mathrm{u}} - \tilde{p}_{1\mathrm{u}}\tilde{p}_{2\mathrm{d}} - \tilde{p}_{1\mathrm{d}}\tilde{p}_{2\mathrm{u}} + \tilde{p}_{1\mathrm{d}}\tilde{p}_{2\mathrm{d}} = \epsilon_1\epsilon_2. \end{aligned} \tag{44}$$

We can now rewrite (Q) for the case of a general ρ in the following equivalent form in terms of r_{ij} denoted by $(\hat{Q})$, where

$$r_{ij} \equiv p_{ij} - \tilde{p}_{1i}\tilde{p}_{2j} \quad \forall\, i, j \in \Omega. \tag{45}$$

$$(\hat{Q}) \quad \sum_{i\in\Omega} r_{ij} = 0 \quad \forall\, j \in \Omega, \tag{46a}$$

$$\sum_{j\in\Omega} r_{ij} = 0 \quad \forall\, i \in \Omega, \tag{46b}$$

$$r_{\mathrm{uu}} - r_{\mathrm{ud}} - r_{\mathrm{du}} + r_{\mathrm{dd}} = \frac{\rho}{c_1 c_2}, \tag{46c}$$

$$0 \leqslant r_{ij} + \tilde{p}_{1i}\tilde{p}_{2j} \leqslant 1 \quad \forall\, i, j \in \Omega, \tag{46d}$$

where (46c) is obtained by subtracting (44) from (38). In $(\hat{Q})$, r_{ij} can be considered as an adjustment term added to the probability $\tilde{p}_{1i}\tilde{p}_{2j}$ from the uncorrelated case to represent the correlation.

Next, we discuss three methods for selecting a feasible solution from $(\hat{Q})$ (or equivalently from (Q)). The first two methods for selecting branching probabilities are commonly used by finance professionals. They are the method presented by Hull and White (1994) and the method of fitting higher moments. The shortcoming of these two methods will also be discussed. We will then propose our optimization-based approach.

3.4.1. Hull and White's (1994) Method. In Hull and White (1994), the authors proposed the use of the following adjustment matrix when $\rho > 0$:

$$\begin{bmatrix} r_{\mathrm{du}} & r_{\mathrm{mu}} & r_{\mathrm{uu}} \\ r_{\mathrm{dm}} & r_{\mathrm{mm}} & r_{\mathrm{um}} \\ r_{\mathrm{dd}} & r_{\mathrm{md}} & r_{\mathrm{ud}} \end{bmatrix} = \begin{bmatrix} -\delta & -4\delta & +5\delta \\ -4\delta & +8\delta & -4\delta \\ +5\delta & -4\delta & -\delta \end{bmatrix}. \tag{47}$$

It can be easily verified that this matrix satisfies (46a) and (46b). Plugging (47) into (46c) introduces a correlation equal to $12c_1c_2\delta$. By choosing $c_1 = c_2 = \sqrt{3}$, the authors suggest setting δ equal to $\rho/36$. Similarly, when $\rho < 0$,

$$\begin{bmatrix} r_{\mathrm{du}} & r_{\mathrm{mu}} & r_{\mathrm{uu}} \\ r_{\mathrm{dm}} & r_{\mathrm{mm}} & r_{\mathrm{um}} \\ r_{\mathrm{dd}} & r_{\mathrm{md}} & r_{\mathrm{ud}} \end{bmatrix} = \begin{bmatrix} +5\delta & -4\delta & -\delta \\ -4\delta & +8\delta & -4\delta \\ -\delta & -4\delta & +5\delta \end{bmatrix}. \tag{48}$$

The advantage of this approach is its computational efficiency. There is, however, a feasibility problem because such $[r_{ij}]$, $i, j \in \Omega$, obtained from (47) or (48), may violate (46d). Precisely, this approach works if

$$\rho \leqslant 12c_1c_2 \cdot \max\{\delta \,|\, 0 \leqslant r_{ij}(\delta) + \tilde{p}_{1i}\tilde{p}_{2j} \leqslant 1 \ \forall i, j \in \Omega\}, \tag{49}$$

where $r_{ij}(\delta)$ are in (47). With $\tilde{p}_{ij}$, $i, j \in \Omega$, expressed in terms of ϵ_1 and ϵ_2, and taking into account all their possible values ($|\epsilon_1|, |\epsilon_2| \leqslant 1/2$), the upper bound given in (49) is only 0.0625, if $c_1 = c_2 = \sqrt{3}$ as suggested by Hull and White (1994). The right-hand side of (49) can be maximized over the values of c_1 and c_2. The maximum is achieved when $c_1 = c_2 \approx 1.53$, with a maximized value equal to 0.201. This number is still significantly smaller than 0.676, up to which the lattice feasibility can be guaranteed.

3.4.2. Fitting Higher Moments. The reason that (Q) has infinitely many solutions is that it has more variables (nine) than equality constraints (six). One intuitive approach to deliver a unique solution is to make the system more constrained by adding three additional equations. A reasonable choice for these three constraints is to match higher moments of the joint probability density function of y_1 and y_2. While there are many alternatives for choosing the three moments, we consider those immediately higher than what we have matched. They are

$$E[(y_1(t+1) - y_1(t))^2(y_2(t+1) - y_2(t)) \,|\, y_1(t), y_2(t)] = \sum_{i,\,j\in\Omega} p_{ij}\kappa_{1i}^2\kappa_{2j}h_1^2h_2, \tag{50a}$$

$$E[(y_1(t+1) - y_1(t))(y_2(t+1) - y_2(t))^2 \,|\, y_1(t), y_2(t)] = \sum_{i,\,j\in\Omega} p_{ij}\kappa_{1i}\kappa_{2j}^2h_1h_2^2, \tag{50b}$$

$$E[(y_1(t+1) - y_1(t))^2(y_2(t+1) - y_2(t))^2 \,|\, y_1(t), y_2(t)] = \sum_{i,\,j\in\Omega} p_{ij}\kappa_{1i}^2\kappa_{2j}^2h_1^2h_2^2. \tag{50c}$$

The exact expression of these three moments can be derived by observing that $y_1(t + \Delta t)$ and $y_2(t + \Delta t)$, conditional on the values of $y_1(t)$ and $y_2(t)$, have a bivariate normal distribution. Equations (37a) to (37f), along with (50a) to (50c), form the new system of linear equations.

By assuming that both y_1 and y_2 follow some MR processes, we tested the feasibility of each given correlation value associated with a given (c_1, c_2) by simulating all possible values and combinations of (ϵ_1, ϵ_2) and (κ_1, κ_2). We found that the maximal correlation for which this approach can guarantee a feasible solution for all simulated instances is around 0.267. This number is, again, significantly lower than the theoretical upper bound 0.676.

3.4.3. Optimization-Based Approach. One natural way to identify a "good" solution from $(\hat{Q})$ is to include an objective function and turn $(\hat{Q})$ into an optimization problem. Intuitively, we seek a solution of $(\hat{Q})$ that minimizes the metric $d(p, f)$ defined in §3.3. Consider a specific node that branches into nine nodes, each of which is denoted by $\mathbf{z}_{ij}$, $i, j \in \Omega$, as shown in Figure 4. The surrounding area $\phi(\mathbf{z}_{ij})$ for each node $\mathbf{z}_{ij}$ is defined as in §3.3. Consider the following optimization problem, denoted by (QP_0):

$$(QP_0) \quad \min \sum_{i,\,j\in\Omega} (p_{ij} - \tilde{f}_{ij})^2 \quad \text{s.t. (46a) to (46d)}, \tag{51}$$

where

$$\tilde{f}_{ij} \equiv \int_{\phi(\mathbf{z}_{ij})} f(\mathbf{w})\, d\mathbf{w} \tag{52}$$

and f is the p.d.f. of the underlying processes at this given node. Because f is a bivariate normal distribution, $\tilde{f}_{ij}$

can be evaluated by approximation methods, such as Drezner (1978). However, $\tilde{f}_{ij}$ will be different from node to node, and solving the exact version of (QP_0) will not be efficient. We seek to approximate (QP_0). Define $\tilde{s}_{ij} \equiv \tilde{f}_{ij} - \tilde{p}_{1i}\tilde{p}_{2j}$ and consider the objective function of (QP_0) as follows:

$$\sum_{i,j\in\Omega}(p_{ij}-\tilde{f}_{ij})^2 = \sum_{i,j\in\Omega}(p_{ij}-\tilde{p}_{1i}\tilde{p}_{2j}+\tilde{p}_{1i}\tilde{p}_{2j}-\tilde{f}_{ij})^2 \quad (53a)$$

$$= \sum_{i,j\in\Omega}(r_{ij}-\tilde{s}_{ij})^2 \quad (53b)$$

$$= \sum_{i,j\in\Omega}(r_{ij}^2 - 2r_{ij}\tilde{s}_{ij}+\tilde{s}_{ij}^2). \quad (53c)$$

Note that $\tilde{s}_{ij}$ captures the change of the area $\phi(\mathbf{z}_{ij})$ under the p.d.f. of a bivariate normal distribution from having no correlation to having a correlation ρ. A closer look indicates that major change of the areas occurs at the four corner nodes, corresponding to $\tilde{s}_{uu}$, $\tilde{s}_{du}$, $\tilde{s}_{ud}$, and $\tilde{s}_{dd}$ (see Figure 4), and little change at the others. The observations are summarized below.

Assumptions on $\tilde{s}_{ij}$

S1. $\tilde{s}_{ij} \approx 0\ \forall ij \in \{\text{um, mu, mm, md, dm}\}$.

S2. *If* $\rho > 0$, *then* $\tilde{s}_{uu} > 0$, $\tilde{s}_{dd} > 0$, $\tilde{s}_{ud} < 0$, *and* $\tilde{s}_{du} < 0$. *Likewise, if* $\rho < 0$, *then* $\tilde{s}_{uu} < 0$, $\tilde{s}_{dd} < 0$, $\tilde{s}_{ud} > 0$, *and* $\tilde{s}_{du} > 0$.

S3. $|\tilde{s}_{ij}|$ *is roughly a constant, denoted by* $k_s > 0$ *for* $ij \in \{uu, ud, du, dd\}$.

Using the approximations S1 to S3, if $\rho > 0$, we have

$$\sum_{i,j\in\Omega}(p_{ij}-\tilde{f}_{ij})^2$$

$$\approx \sum_{i,j\in\Omega} r_{ij}^2 - 2(r_{uu}-r_{ud}-r_{du}+r_{dd})k_s + \sum_{i,j\in\Omega}\tilde{s}_{ij}^2 \quad (54a)$$

$$= \sum_{i,j\in\Omega} r_{ij}^2 - \frac{2\rho k_s}{c_1 c_2} + \sum_{i,j\in\Omega}\tilde{s}_{ij}^2. \quad (54b)$$

In (54b), (46c) is plugged, which is an equality constraint of (QP_0). If $\rho < 0$, the negative sign of the second term in (54b) is reversed. Because the last two terms of (54b) are constant, they can be left aside in the objective function. Consequently, we have the following optimization problem denoted by (QP), which is an approximate optimization problem of (QP_0):

$$(QP) \quad \min \sum_{i,j\in\Omega} r_{ij}^2 \quad \text{s.t. (46a) to (46d).} \quad (55)$$

(QP) appears to be an easier problem than (QP_0) and has the following characteristics. First, when $\rho = 0$, the optimal solution is $r_{ij} = 0\ \forall i, j \in \Omega$. That is, this objective selects the branching probabilities for the uncorrelated case $p_{ij} = \tilde{p}_{1i}\tilde{p}_{2j}\ \forall i, j \in \Omega$. Second, when $\rho \neq 0$, while there are infinitely many solutions, this problem selects the one that is closest to the probability set for the uncorrelated case in Euclidean norm. As shown above, such an efficient way—in terms of minimal deviation—for constructing a probability set that shows correlation between y_1 and y_2 also approximately minimizes the defined metric $d(p, f)$ in §3.3.

(QP) is a standard quadratic program with nine variables, subject to seven linear constraints not including the bounds of variables. Equations (46a) and (46b) describe a 3×3 matrix $[r_{ij}]$, $i, j \in \Omega$, such that the components in each column and row sum up to zero. We can substitute the matrix components in terms of only the four components at the corners as follows:

$$\begin{bmatrix} r_{du} & r_{mu} & r_{uu} \\ r_{dm} & r_{mm} & r_{um} \\ r_{dd} & r_{md} & r_{ud} \end{bmatrix} = \begin{bmatrix} r_{du} & -r_{du}-r_{uu} & r_{uu} \\ -r_{du}-r_{dd} & r_{du}+r_{dd}+r_{uu}+r_{ud} & -r_{uu}-r_{ud} \\ r_{dd} & -r_{dd}-r_{ud} & r_{ud} \end{bmatrix}.$$

Putting (46d) aside for the time being, (QP) is reduced to the following problem, which is much easier to handle:

$$\begin{aligned} (\widehat{QP}) \quad \min\ & f_Q = r_{du}^2 + (r_{du}+r_{uu})^2 + r_{uu}^2 + (r_{du}+r_{dd})^2 \\ & + (r_{du}+r_{dd}+r_{uu}+r_{ud})^2 + (r_{uu}+r_{ud})^2 \\ & + r_{dd}^2 + (r_{dd}+r_{ud})^2 + r_{ud}^2 \\ \text{s.t.}\ & r_{uu} - r_{ud} - r_{du} + r_{dd} = \frac{\rho}{c_1 c_2}. \end{aligned}$$

Define the Lagrangian $\mathcal{L}(r_{uu}, r_{ud}, r_{du}, r_{dd}, \lambda)$ as

$$\mathcal{L}(r_{uu}, r_{ud}, r_{du}, r_{dd}, \lambda) = f_Q + \lambda\left(r_{uu} - r_{ud} - r_{du} + r_{dd} - \frac{\rho}{c_1 c_2}\right), \quad (56)$$

where $\lambda > 0$ is the Lagrange multiplier of the constraint. With some algebra, the optimality condition is

$$\begin{bmatrix} \frac{\partial\mathcal{L}}{\partial r_{uu}} \\ \frac{\partial\mathcal{L}}{\partial r_{ud}} \\ \frac{\partial\mathcal{L}}{\partial r_{du}} \\ \frac{\partial\mathcal{L}}{\partial r_{dd}} \\ \frac{\partial\mathcal{L}}{\partial\lambda} \end{bmatrix} = \begin{bmatrix} 8 & 4 & 4 & 2 & 1 \\ 4 & 8 & 2 & 4 & -1 \\ 4 & 2 & 8 & 4 & -1 \\ 2 & 4 & 4 & 8 & 1 \\ 1 & -1 & -1 & 1 & 0 \end{bmatrix}$$

$$\cdot \begin{bmatrix} r_{uu} \\ r_{ud} \\ r_{du} \\ r_{dd} \\ \lambda \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \dfrac{\rho}{c_1 c_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{57}$$

At the optimal solution $(r^*_{uu}, r^*_{ud}, r^*_{du}, r^*_{dd}, \lambda^*)$, we have

$$\begin{bmatrix} r^*_{uu} \\ r^*_{ud} \\ r^*_{du} \\ r^*_{dd} \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 8 & 4 & 4 & 2 & 1 \\ 4 & 8 & 2 & 4 & -1 \\ 4 & 2 & 8 & 4 & -1 \\ 2 & 4 & 4 & 8 & 1 \\ 1 & -1 & -1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \dfrac{\rho}{c_1 c_2} \end{bmatrix}$$

$$= \frac{\rho}{4c_1 c_2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 2 \end{bmatrix}. \tag{58}$$

That is, the proposed method suggests the following adjustment matrix $[r_{ij}]$, $i, j \in \Omega$:

$$\begin{bmatrix} r_{du} & r_{mu} & r_{uu} \\ r_{dm} & r_{mm} & r_{um} \\ r_{dd} & r_{md} & r_{ud} \end{bmatrix} = \frac{\rho}{4c_1 c_2} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}. \tag{59}$$

Compared with (47) and (48), (59) shows the efficiency that we defined earlier in this section for constructing the probability set from the uncorrelated case: All entries are zeros except the four corner ones that are involved in the constraint. This result also coincides with the approximations S1 to S3.

Unfortunately, the solution in (59) may violate (46d). That means some component(s) $r_{ij} + \tilde{p}_{1i}\tilde{p}_{2j}$, $i, j \in \Omega$, may be outside of $[0, 1]$. Note that $p_{ij} = r_{ij} + \tilde{p}_{1i}\tilde{p}_{2j}$ will not be greater than one, because if it is, p_{ij} will be adjusted back to one, implying that all other branching probabilities are zero, which corresponds to a case with certainty. Therefore, we can focus on the case that p_{ij} may be negative. If p_{ij} becomes negative, it will be adjusted back to zero to satisfy (46d). This implies that the corresponding constraint will become binding, associated with a nonzero Lagrange multiplier. When this happens, the optimality condition (57) requires some minor modification. If the binding constraints (i.e., the p_{ij} that are zeros at optimality) can be known a priori, the optimal solution, along with the Lagrange multipliers, can be obtained by solving a system of linear equations. Because there are only few possible combinations for p_{ij} to be zero, a method that seeks to identify the correct binding constraints by iterations can be constructed. Because the problem is a (strictly convex) quadratic optimization subject to linear constraints, the solution is unique. Any solution sought (associated with the Lagrange multipliers) satisfying the optimality condition must be the optimal solution.

We have implemented an efficient algorithm based on such a strategy. By setting up a test bed on MATLAB for solving random instances of (QP), we have found that our algorithm takes as little as 8% of the CPU time required by MATLAB's embedded quadratic programming solver (qp.m) to solve the same problem on average.

Given a set of branching probabilities that fits the underlying processes well in the sense that the branching probability of each node matches the area of its surrounding $\phi(\cdot)$ under the corresponding probability density function, as the lattice expands over more time periods, more nodes are generated. That means that the surrounding area $\phi(\cdot)$ of each node (at a later time period) is refined and becomes smaller. The discrete branching probabilities will continue to fit the continuous counterpart well for the later period. In addition, the overall fitting generally improves as the total number of nodes increases.

3.5. Selecting the Lattice Cell Size

The size of the lattice cells can also affect the convergence of the lattice to the underlying continuous process. Using the one-factor model as an example, if c is small, ϵ in (26) tends to be small so that the middle branch can match the conditional mean better. However, if c is too small, the three-point pattern will be too narrow to capture the width of the normal distribution. Therefore, the optimal value of c that yields good convergence exists.

Hull and White (1993) suggested choosing $c = \sqrt{3}$ in the one-factor model, which can match the first five moments of a normal probability density function (Keefer 1994). The authors also suggested using $c_1 = c_2 = \sqrt{3}$ for the two-factor lattices in their subsequent paper (Hull and White 1994). Choosing $c_1 = c_2 = \sqrt{3}$ yields a $\rho^{\max}$ of 0.625, which is quite close to the maximum value 0.676. However, due to the correlation, choosing $c_1 = c_2 = \sqrt{3}$ does not necessarily match the higher moments as it does in the one-factor case. Therefore, choosing $c_1 = c_2 = \sqrt{3}$ is merely a heuristic, lacking theoretical support. Next, we propose a more general approach, such that the values of c_1 and c_2 can be customized, depending on actual problem data.

Let $p_T(Y_{1,N}, Y_{2,N})$ denote the joint probability distribution function of $(Y_{1,N}, Y_{2,N})$ at time T, and $f_T(y_1(T), y_2(T))$ the joint c.d.f. of $(y_1(T), y_2(T))$ at time T. We propose to select c_1 and c_2 by solving the following optimization problem:

$$(P_2) \quad \min_{c_1, c_2} \; d(p_T, f_T)$$
$$\text{s.t.} \quad (36) \text{ and } (41).$$

(P_2) can be solved using optimization methods that do not require derivatives (see, for example, Nelder and Mead 1965). For each N, let $(c_1^{(N)}, c_2^{(N)})$ denote the optimal solution for (P_2). Our experience shows that $(c_1^{(N)}, c_2^{(N)})$ converges quickly as N increases. In general, solving c_1 and c_2 using $N = 5$ gives satisfactory results. An example will be given in §4 for numerical results, which indicates that the optimal value of (c_1, c_2) is far from $(\sqrt{3}, \sqrt{3})$, suggested by Hull and White (1994). To obtain the exact continuous distribution f_T, note that f_T is a bivariate normal distribution, which can be uniquely identified by the means and covariance matrix of $y_1(T)$ and $y_2(T)$. It has been observed that as N increases, the moments of p_T converge much faster than the metric defined in (42), which motivated us to solve (P_2) for selecting (c_1, c_2). Note that the solution of (P_2) may vary with changes of data such as the initial prices, price process parameters, and price correlation.

4. Numerical Results

4.1. Modeling Price Uncertainty

As a practice, we assume that the prices of electricity and fuel consumed by the generator follow some correlated geometric mean reversion (GMR) processes advocated by Barz (1999). Assume that the evolutions of both P_t^E and P_t^F can be captured by the following processes:

$$d\ln(P_t^E) = -\mu_t^E(\ln(P_t^E) - m_t^E)dt + \sigma_t^E dB^E \tag{60}$$

and

$$d\ln(P_t^F) = -\mu_t^F(\ln(P_t^F) - m_t^F)dt + \sigma_t^F dB^F, \tag{61}$$

where B^E and B^F are two Wiener processes with instantaneous correlation ρ.

The GMR price models in (60) and (61) are characterized by mean reversion and lognormally distributed seasonal prices. The mean reversion parameter μ_t, in some sense, represents the storability of the commodity. For electricity, which is quite difficult to store, this parameter is large, implying little autocorrelation between today's price and tomorrow's price of the same hour. Furthermore, this parameter, in conjunction with the volatility σ_t, captures the short-term and long-term price fluctuations and characterizes the variance of the lognormal price distribution. The mean reversion parameter μ_t and the volatility σ_t may be time dependent, as indicated by a subscript t. Finally, m_t is a periodic function capturing the cyclical nature of the long-term expected prices. m_t is thus a function of the interplay between the cost of production and consumer demand for the commodity.

To obtain the parameters of the price processes, we examined the historical price data series of Nymex natural gas prices and electricity prices from the PJM (Pennsylvania, New Jersey, and Maryland) in the year 2002, taking the logarithm of these prices as our basic data series. Because there is no hourly market for natural gas, μ_t, σ_t, and m_t^F are assumed to be constant within a given day and we fit the price process to daily data. Because the model time step is hourly, however, we adjust these parameters accordingly. Specifically, because the model is a continuous-time model, we can estimate the hourly fluctuations from the daily parameters.

It may seem that by assuming that the gas prices vary over the day (A5), we obfuscate the problem. We argue that even though there is only one price for gas each day, purchases of gas at that price may not be available at every hour of the day. Besides, the prices of gas futures do vary throughout the day, and it seems reasonable to assume that spot prices likewise show such variability in actual purchases hour by hour. Furthermore, the price variability within a day may be viewed as some random opportunity cost associated with gas that reflects updates in information throughout the day. Our approach can be viewed as an approximation of the actual hourly prices for gas purchases.

For electricity, we used historical daily data from the PJM day-ahead market, taking the logarithm of these prices as our basic data series. However, to incorporate a daily price pattern, we then adjust m_t^E by overlaying the daily electricity price pattern (in terms of percentage changes). Using the method of maximum likelihood, we obtain $\mu_t^F = 6.95 \times 10^{-4}$ and $\sigma_t^F = 0.019$ independent of t. For electricity, we obtain $\mu_t^E = 0.062$, $\sigma_t^E = 0.137$ for peak hours (6 A.M. to 10 P.M.), and $\mu_t^E = 0.044$, $\sigma_t^E = 0.123$ for off-peak hours. Note that $1/\mu_t^E$ and $1/\mu_t^F$ are in units of hours and σ_t^E and σ_t^F are in units of \$/MWh and \$/MMBtu, respectively. As mentioned previously, m_t^E captures the cyclical nature of the expected electricity prices. Detailed m_t^E values are summarized in Table 1. At time $t = 0$, suppose that prices $P_0^E = 20$ (\$/MWh), and $P_0^F = 2.2$ (\$/MMBtu) are observed. We set the instantaneous correlation coefficient between the logarithms of the electricity and natural gas prices to be $\rho = 0.3$, as observed in the market. This is close to the correlation between the electricity and natural gas prices.

The two-factor lattice that represents the two processes (60) and (61) can be obtained by following the steps presented in §3.3. Letting $y_1(t) = \ln(P_t^E)$ and $y_2(t) = \ln(P_t^F)$, (60) and (61) are reduced to

$$dy_1(t) = -\mu_t^E(y_1(t) - m_t^E)dt + \sigma_t^E dB^E \tag{62}$$

Table 1. Values of hourly m_t^E.

t	m_t^E	t	m_t^E	t	m_t^E
1	2.88	9	3.27	17	3.58
2	2.74	10	3.36	18	3.63
3	2.68	11	3.45	19	3.58
4	2.65	12	3.45	20	3.51
5	2.69	13	3.46	21	3.49
6	2.85	14	3.51	22	3.34
7	3.17	15	3.53	23	3.10
8	3.23	16	3.56	24	2.96

and

$$dy_2(t) = -\mu_t^F(y_2(t) - m_t^F)dt + \sigma_t^F dB^F, \tag{63}$$

which belong to the category of (19) and (20) (except that the volatility is time dependent, which is addressed in the online appendix). The lattice is generated in the (y_1, y_2) domain. At any given node of the lattice (y_1, y_2), the profit function (11) is evaluated with respect to the corresponding price set $(P_t^E, P_t^F) = (\exp(y_1), \exp(y_2))$. It can be seen that both (62) and (63) satisfy the MR property in (21a) and (21b). Therefore, the lattice size for the GMR processes will be capped.

4.2. Convergence Analysis

Let $y_1(0) = \ln(P_0^E)$ and $y_2(0) = \ln(P_0^F)$. Using the optimization-based approach presented in §3.4.3, we generate the corresponding two-factor lattice $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$, with an initial condition $(Y_{1,0}, Y_{2,0}) = (\ln(20), \ln(2.2))$. First, we show the approximated joint distribution at time $T = 1$ by varying N in Figure 5. The three contours in Figure 5 are generated by the contour function in MATLAB, which performs some level of data interpolation and extrapolation. As can be seen in Figure 5, as N increases, more price nodes are generated at time T and a better approximation of the distribution is yielded. Intuitively, the metric $d(p_T, f_T)$ defined in §3.3 will converge to zero as N increases.

Next, we solve (P_2) defined in §3.5 using the Nelder-Mead method, which uses a sequence of polytopes (simplices) to close in and eventually locate the optimal values of c_1 and c_2. The initial polytope is located in the feasible region of (c_1, c_2) corresponding to a $\rho^{\max} \geqslant 0.3$. The optimal values of c_1 and c_2 versus N are summarized in Table 2. It can be seen in Table 2 that when $N > 5$, the minimum is achieved consistently at $(c_1, c_2) = (1.50, 1.49)$, where the best fit of the continuous density function is obtained, implying the best convergence. The corresponding least-square error of the metric $d(f, p)$ defined in §3.3 is small, indicating a good fit of the entire density function. At this point, from (41) the corresponding $\rho^{\max}$ is 0.56, which is sufficient for our application. In this test, the branching probabilities are obtained using the proposed optimization-based method in §3.4.3.

Next, we test the convergence using the other two methods for generating the branching probabilities—by Hull and White (1994) (§3.4.1) and fitting higher moments (§3.4.2)—compared with our proposed one. For Hull and White's method, $c_1 = c_2 = \sqrt{3}$ as proposed by the authors. The optimal values for $(c_1, c_2) = (1.50, 1.49)$ obtained in the previous test are adopted for the methods for fitting higher moments and using optimization. The convergence information, in terms of the least-square error of the metric, is summarized in Table 3 and is depicted in Figure 6. It can be seen that the proposed optimization-based method

Figure 5. (a) $N = 1$ (3×3 nodes); (b) $N = 3$ (7×7 nodes); (c) $N = 10$ (21×21 nodes).

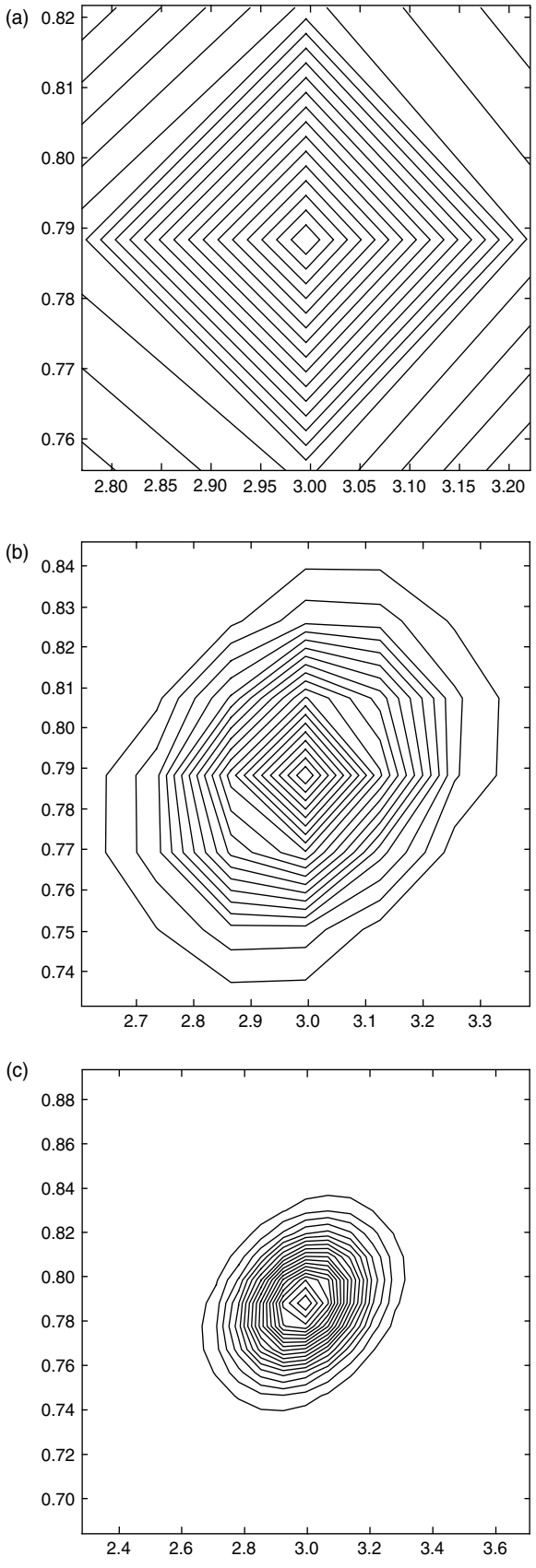


Table 2. Optimal lattice cell size (c_1, c_2) vs. N.

N	c_1	c_2	$d(p, f)$
1	1.56	1.52	0.01300
2	1.50	1.48	0.00770
3	1.51	1.50	0.00470
4	1.51	1.50	0.00300
5	1.51	1.50	0.00220
6	1.50	1.49	0.00167
7	1.50	1.49	0.00133
8	1.50	1.49	0.00109
9	1.50	1.49	0.00092
10	1.50	1.49	0.00079

has the best convergence, followed by the method of fitting higher moments, then by Hull and White's method. The convergence of the methods of optimization and higher moments is one order of magnitude better than that of Hull and White's method.

4.3. Short-Term Generation Asset Valuation

We have implemented the proposed optimization-based method for valuing a power plant using FORTRAN in a Pentium 4 PC and applied to a natural gas-fueled generating unit with the following input-output characteristics:

$$H(q_t) = 540 + 9.223q_t + 0.00234q_t^2, \tag{64}$$

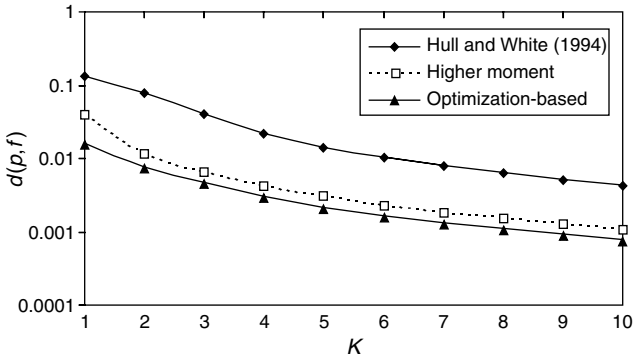
with $q^{\min} = 225$ MW, $q^{\max} = 700$ MW. We let $\tau = \nu = 2$, $t^{\text{on}} = 5$, $t^{\text{off}} = 10$ to fully capture the influence of the physical constraints. In the following numerical tests, we use $(c_1, c_2) = (2/\sqrt{3}, 2/\sqrt{3})$ for the lattice, obtained from §4.2. Because of the short duration (one day to one week) of the test problems considered here, we do not apply a discount rate.

As stated previously, as the number of stages K increases, the joint distribution approximated by the lattice converges to the exact one. Consider $T = 24$ (hours) and let $N = T \times K$, where K is the number of subintervals for each hour, the same as in §2.2. As K increases, the power plant value should converge to its true expected value. In Table 4, we show the power plant value in relationship to K, also depicted in Figure 7. It can be seen that the power plant value eventually converges to a value around \$25,050. The error due to discretization is estimated to be around 2.7% if $K = 1$ is selected. In the last column of Table 4, the CPU time increases polynomially (order ≈ 2.8) with respect to the number of subintervals K.

Table 3. Comparison of convergence $(d(p, f))$ by three different methods for obtaining branching probabilities.

N	Hull and White (1994)	Higher moments	Optimization-based
1	0.13503	0.04042	0.01612
2	0.07878	0.01185	0.00768
3	0.04018	0.00683	0.00471
4	0.02241	0.00442	0.00304
5	0.01436	0.00315	0.00218
6	0.01028	0.00240	0.00167
7	0.00788	0.00190	0.00133
8	0.00631	0.00156	0.00109
9	0.00521	0.00131	0.00092
10	0.00440	0.00112	0.00079

Figure 6. Comparison of lattice convergence of three methods (logarithmic scale).



Next, we compare the proposed lattice approach with the simulation-based approach in Tseng and Barz (2002). Using their method to solve the 24-hour case above $(K = 1)$, the CPU time required is around 13 minutes. Their method obtains \$24,566, quite close to that obtained by the lattice model (\$24,400). Apparently, the proposed method is much more computationally efficient than the simulation-based approach.

By repeatedly running the program with different parameters, we obtain the following sensitivity analysis results for the one-week case $(T = 168)$. These are: the power plant value versus T (in Table 5a and Figure 8a), the power plant value versus parallel shift of the vector of $\boldsymbol{\mu}^E \equiv \{\mu_t^E\}_{t=1}^T$ (in Table 5b and Figure 8b), the power plant value versus parallel shift of the vector of $\boldsymbol{\sigma}^E \equiv \{\sigma_t^E\}_{t=1}^T$ (in Table 5c and Figure 8c), and the power plant value versus ρ (in Table 5d).

Figure 7. The power plant value vs. K, the number of subintervals in each hour.

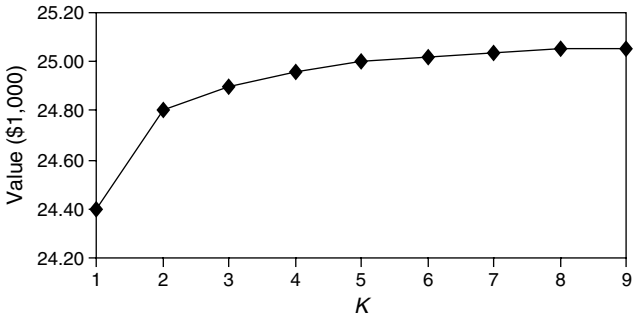


Table 4. Power plant value ($\$ \times 10^3$) vs. K, the number of subintervals in each hour.

K	Power plant value	No. of nodes at final stage	CPU time (sec.)
1	24.40	43×48	0.28
2	24.80	83×96	2.43
3	24.90	124×144	6.53
4	24.96	163×192	13.97
5	25.00	203×242	25.94
6	25.02	243×290	43.37
7	25.04	284×338	66.82
8	25.05	323×384	96.52
9	25.05	363×434	166.33

As seen in Figure 8a, the power plant value increases approximately linearly as the length of the planning horizon T increases, especially when $T \geqslant 48$. From $T =$ 24 to 48, the power plant value quadruples because this increase makes T much larger than the minimum uptime ($t^{\text{on}} = 5$) and downtime ($t^{\text{off}} = 10$) and therefore lowers their impact to the power plant value. Note that this relationship will become more concave as the discount rates increase. Table 5a also displays the number of lattice nodes in the final stage, which indicates memory requirement of the computer. In this case, P_t^F dominates the memory usage due to its much smaller reversion parameter than that of P_t^E. A higher reversion parameter implies stronger reversion. Namely, the price upper bound of the lattice becomes lower and the price lower bound becomes higher. Therefore, fewer lattice nodes need to be evaluated. In Table 5a, as T exceeds 48 (hours), the number of nodes to represent price distribution for P_t^E stops increasing because its one-factor lattice size has been capped.

From Figure 8b, it can be seen that the power plant value initially decreases and then increases as the parallel shift of $\boldsymbol{\mu}^E$ increases. To interpret the result, in general, with bigger $\boldsymbol{\mu}^E$ any price deviation from the mean does not last long; thus, there are fewer "lasting" profitable opportunities. On the other hand, the physical constraints of the unit place restrictions against the unit to react to these profitable opportunities of short durations. Both effects need to balance. This explains the initial decrease of the power plant value. When $\boldsymbol{\mu}^E$ continues to increase, the price reverts to the mean value more quickly, implying less price risk.

Table 5a. Power plant value ($\$ \times 10^3$) vs. time T.

T	24	48	72	96	120	144	168
Value	24.4	99.2	186.2	277.0	369.1	462.2	556.0
CPU (sec.)	0.28	1.80	4.73	8.61	13.68	20.18	27.69
No. of nodes at final stage	43×48	59×96	59×146	59×192	59×240	59×290	59×338

Table 5b. Power plant value ($\$ \times 10^3$) vs. parallel shift of $\boldsymbol{\mu}^E$.

Shift	−0.02	0	0.04	0.08	0.12	0.16	0.20	0.24	0.28
Value	733.6	556.0	423.0	387.0	383.6	391.8	403.0	413.9	423.8
CPU time (sec.)	45.87	27.65	16.52	12.50	10.59	9.67	8.58	8.09	7.64

Table 5c. Power plant value ($\$ \times 10^3$) vs. parallel shift of $\boldsymbol{\sigma}^E$.

Shift	−0.08	−0.04	−0.02	0	0.02	0.04	0.08
Value	253.7	378.2	461.0	556.0	663.4	785.5	1,073.8

Table 5d. Relative power plant value ($\$ \times 10^3$) vs. correlation coefficient ρ.

ρ	0	0.1	0.2	0.3	0.4	0.5
High	1,599.9	1,577.2	1,553.9	1,530.0	1,505.5	1,480.2
Medium	588.5	577.9	567.1	556.0	544.6	532.9
Low	276.7	271.7	266.6	261.3	256.0	250.5

Figure 8a. The power plant value vs. T.

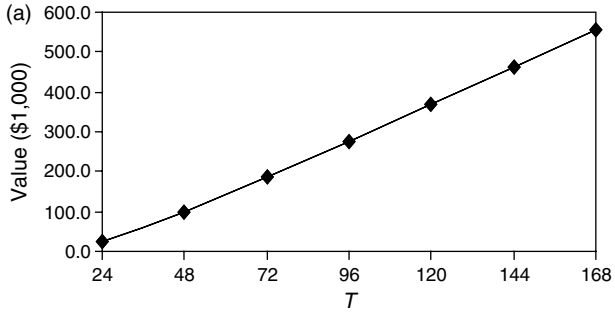


Figure 8b. The power plant value vs. parallel shift of $\boldsymbol{\mu}^E$.

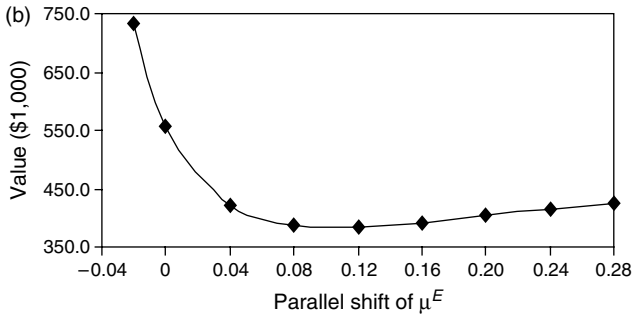
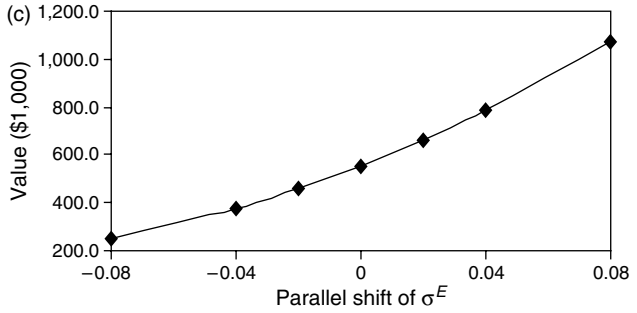


Figure 8c. The power plant value vs. parallel shift of $\boldsymbol{\sigma}^E$.



Because it becomes easier to predict prices, the power plant value increases. In this case, the power plant value reaches a minimum when the parallel shift of $\boldsymbol{\mu}^E$ is around 0.1. Also note that the CPU time decreases as the parallel shift of $\boldsymbol{\mu}^E$ decreases.

In Figure 8c, we see that the plant value increases as the parallel shift of $\boldsymbol{\sigma}^E$ increases. Moreover, the plant value is extremely sensitive to the price volatility. It can be estimated that a 1% increase (decrease) in the average hourly volatility would result in roughly a 1.4% increase (decrease) of the power plant value. This result implies that the unit is more profitable in a market place with more volatile prices.

Finally, we test the relation of the power plant value versus the correlation ρ between the electricity and fuel prices. Three cases, corresponding to high volatility, medium volatility, and low volatility, are tested. The case with medium volatility is the baseline case given in §4.1 (i.e., $\sigma_t^E = 0.137$ for peak hours, 0.123 for off-peak hours; and $\sigma_t^F = 0.019$). The high-volatility and low-volatility cases have their hourly volatility value twice and half of the hourly value in the medium-volatility case, respectively. Basically, the power plant value decreases as the value of ρ increases, meaning less-profitable opportunities with big price spreads. Their relation is approximately linear (slightly concave) in all three cases, as shown in Table 5d. Two observations can be made: (1) The power plant value in the high-volatility case is approximately 2.8 times the value in the medium-volatility case; more than the increase of the volatilities. Similarly, when the hourly volatilities are reduced by half, the power plant value is reduced by more than a half (approximately 0.47 times). This reiterates the high sensitivity of the power plant value versus the volatility, as previously discussed. (2) The power plant value in the high-volatility case is more sensitive to ρ than in the medium-volatility case, which is more sensitive than in the low-volatility case. The sensitivity (slope) increases (or decreases) by roughly the same factor of the increase (or decrease) of the volatility in the high-volatility (or low-volatility) case.

5. Conclusion

In this paper, we present a method for valuing a power plant using discrete-time price lattices. A framework that generates lattices for two correlated Ito processes is developed. Numerical tests show that the proposed model has at least two advantages: flexible modeling and efficient computation. In terms of flexible modeling, we incorporate operational constraints into the decision-making process, and the lattice framework can handle general price processes. In computation, our method provides a much more efficient approach to calculating the power plant value than the MC simulation.

6. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/ecompanion.html.

Acknowledgments

The authors express their gratitude to B. F. Hobbs, four anonymous reviewers, and Graydon Barz for several helpful comments. This research was supported by the National Science Foundation under grants 0100186, 0343011, and 0223314.

References

Barz, G. 1999. Stochastic financial models for electricity derivatives. Ph.D. dissertation, Department of Engineering—Economic Systems and Operations Research, Stanford University, Stanford, CA.

Deng, S., S. Oren. 2003. Incorporating operational characteristics and start-up costs in option-based valuation of power generation capacity. *Probab. Engrg. Inform. Sci.* **17** 155–181.

Deng, S., B. Johnson, A. Sogomonian. 1998. Exotic electricity options and the valuation of electricity generation and transmission. Presented at Chicago Risk Management Conference, Chicago, IL.

Drezner, Z. 1978. Computation of the bivariate normal integral. *Math. Comput.* **32** 277–279.

Herbert, J. 2001. The relationship between natural gas and electricity prices. The Natural Gas Supply Association (NGSA) documents. www.ngsa.org/docs/GasPrice_v_ElecPrice.pdf.

Hsu, M. 1998a. Spark spread options are hot! *Electricity J.* **11** 28–39.

Hsu, M. 1998b. Using financial options theory to value and hedge natural gas power assets. Presented at Pricing Energy in a Competitive Market Conference. Electric Power Research Institute, Washington, D.C.

Hull, J. C., A. White. 1993. One-factor interest-rate models and the valuation of interest-rate derivative securities. *J. Financial Quant. Anal.* **28** 235–254.

Hull, J. C., A. White 1994. Numerical procedures for implementing term structure models II: Two-factor models. *J. Derivatives* **2** 37–49.

Keefer, D. L. 1994. Certainty equivalent for three-point discrete-distribution approximations. *Management Sci.* **40** 760–773.

Kushner, H. J. 1984. *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. MIT Press, Cambridge, MA.

Margrabe, W. 1978. The value of an option to exchange one asset for another. *J. Finance* **33** 177–186.

Nelder, J. A., R. Mead. 1965. A simplex method for function minimization. *Comput. J.* **7** 308–313.

Sheble, G. B., G. N. Fahd. 1994. Unit commitment synopsis. *IEEE Trans. Power Systems* **9** 128–135.

Svoboda, A. J., C. L. Tseng, C. A. Li, R. B. Johnson. 1997. Short term resource scheduling with ramp constraints. *IEEE Trans. Power Systems* **12** 77–83.

Trigeorgis, L. 1996. *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. MIT Press, Cambridge, MA.

Tseng, C. L. 1996. On power system generation unit commitment problems. Ph.D. dissertation, Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA.

Tseng, C. L., G. Barz. 1999. Short-term generation asset valuation. Presented at 32nd Hawaii Internat. Conf. System Sciences, Maui, HI.

Tseng, C. L., G. Barz. 2002. Short-term generation asset valuation: A real options approach. *Oper. Res.* **50** 297–310.

OPERATIONS RESEARCH
DOI 10.1287/opre.1060.0355ec
pp. ec1–ec7

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"A Framework Using Two-Factor Price Lattices for Generation Asset Valuation" by Chung-Li Tseng and Kyle Y. Lin, *Operations Research* 2007, 55(2) 234–251.

Online Appendix

Appendix A. Proofs

PROOF OF LEMMA 1. Replacing (37e) by (38) in (Q), (P_1) is a problem that optimizes the left-hand side of (38) subject to all other constraints (37a)–(37d), (37f) in (Q). Obviously, $[X_{\min}, X_{\max}]$ provides the exact interval to bound the right-hand side of (38), $\rho/(c_1c_2)+\epsilon_1\epsilon_2$, such that (Q) is feasible. □

PROOF OF LEMMA 2. (1) In this case, because $p_{uu} \leqslant \tilde{p}_{1u}$, $p_{dd} \leqslant \tilde{p}_{1d}$, $p_{ud} \geqslant 0$, $p_{du} \geqslant 0$, it is clear that $\tilde{p}_{1u}+\tilde{p}_{1d}$ is an upper bound for X. We can conclude that $X_{\max} = \tilde{p}_{1u}+\tilde{p}_{1d}$ by presenting the following feasible solution that achieves this upper bound:

$$\begin{bmatrix} p_{du} & p_{mu} & p_{uu} \\ p_{dm} & p_{mm} & p_{um} \\ p_{dd} & p_{md} & p_{ud} \end{bmatrix} = \begin{bmatrix} 0 & \tilde{p}_{2u}-\tilde{p}_{1u} & \tilde{p}_{1u} \\ 0 & \tilde{p}_{2m} & 0 \\ \tilde{p}_{1d} & \tilde{p}_{2d}-\tilde{p}_{1d} & 0 \end{bmatrix}. \tag{EC.A1}$$

(2) Similar to (i).

(3) First, note that

$$p_{uu} \leqslant \tilde{p}_{1u}, \qquad p_{dd} \leqslant \tilde{p}_{2d}, \qquad p_{ud} \geqslant 0.$$

On the other hand,

$$p_{du} = \tilde{p}_{2u} - p_{uu} - p_{mu} \geqslant \max\{0, \tilde{p}_{2u} - \tilde{p}_{1u} - \tilde{p}_{1m}\} = (\tilde{p}_{2u} + \tilde{p}_{1d} - 1)^+,$$

where $x^+ \equiv \max(0, x)$. Hence,

$$\begin{aligned} X &= p_{uu} + p_{dd} - p_{ud} - p_{du} \\ &\leqslant \tilde{p}_{1u} + \tilde{p}_{2d} - (\tilde{p}_{2u} + \tilde{p}_{1d} - 1)^+ \\ &= \min\{\tilde{p}_{1u} + \tilde{p}_{2d}, 1 - (\tilde{p}_{1d} - \tilde{p}_{2d}) - (\tilde{p}_{2u} - \tilde{p}_{1u})\}. \end{aligned}$$

We can conclude that $X_{\max} = \min\{\tilde{p}_{1u} + \tilde{p}_{2d}, 1 - (\tilde{p}_{1d} - \tilde{p}_{2d}) - (\tilde{p}_{2u} - \tilde{p}_{1u})\}$ by presenting the following feasible solution that achieves this upper bound:

$$\begin{bmatrix} p_{du} & p_{mu} & p_{uu} \\ p_{dm} & p_{mm} & p_{um} \\ p_{dd} & p_{md} & p_{ud} \end{bmatrix} = \begin{bmatrix} (\tilde{p}_{1d} - \tilde{p}_{2d} - \tilde{p}_{2m})^+ & \tilde{p}_{2u} - \tilde{p}_{1u} - (\tilde{p}_{1d} - \tilde{p}_{2d} - \tilde{p}_{2m})^+ & \tilde{p}_{1u} \\ \min\{\tilde{p}_{2m}, \tilde{p}_{1d} - \tilde{p}_{2d}\} & (\tilde{p}_{2m} - (\tilde{p}_{1d} - \tilde{p}_{2d}))^+ & 0 \\ \tilde{p}_{2d} & 0 & 0 \end{bmatrix}. \tag{EC.A2}$$

(4) Similar to (iii). □

PROOF OF THEOREM 3. From Lemma 1, a lattice is feasible if the following inequality holds for all $|\epsilon_1|, |\epsilon_2| \leqslant 1/2$:

$$c_1c_2(X_{\min} - \epsilon_1\epsilon_2) \leqslant \rho \leqslant c_1c_2(X_{\max} - \epsilon_1\epsilon_2). \tag{EC.A3}$$

One can guarantee feasibility of the entire lattice if $\rho^{\min}(c_1, c_2) \leqslant \rho \leqslant \rho^{\max}(c_1, c_2)$, where

$$\rho^{\max}(c_1, c_2) \equiv \min_{-1/2 \leqslant \epsilon_1, \epsilon_2 \leqslant 1/2} c_1 c_2 (X_{\max} - \epsilon_1 \epsilon_2),$$

$$\rho^{\min}(c_1, c_2) \equiv \max_{-1/2 \leqslant \epsilon_1, \epsilon_2 \leqslant 1/2} c_1 c_2 (X_{\min} - \epsilon_1 \epsilon_2).$$

To determine $\rho^{\min}(c_1, c_2)$ and $\rho^{\max}(c_1, c_2)$, consider the four cases discussed in Lemma 2.

(1) If $\tilde{p}_{1u} \leqslant \tilde{p}_{2u}$ and $\tilde{p}_{1d} \leqslant \tilde{p}_{2d}$, from Lemma 2(i) and (27) for $|\epsilon_1|, |\epsilon_2| \leqslant 1/2$,

$$X_{\max} - \epsilon_1 \epsilon_2 = \tilde{p}_{1u} + \tilde{p}_{1d} - \epsilon_1 \epsilon_2 = \frac{1}{c_1^2} + \epsilon_1^2 - \epsilon_1 \epsilon_2 \geqslant \frac{1}{c_1^2} - \frac{1}{16},$$

where we use the fact that the global minimum for $\epsilon_1^2 - \epsilon_1 \epsilon_2$ subject to $|\epsilon_1|, |\epsilon_2| \leqslant 1/2$ is $-1/16$, achieved at $(\epsilon_1, \epsilon_2) = (-1/4, -1/2)$.

(2) If $\tilde{p}_{1u} > \tilde{p}_{2u}$ and $\tilde{p}_{1d} > \tilde{p}_{2d}$, similar to (i) one can obtain

$$X_{\max} - \epsilon_1 \epsilon_2 = \tilde{p}_{2u} + \tilde{p}_{2d} - \epsilon_1 \epsilon_2 = \frac{1}{c_2^2} + \epsilon_2^2 - \epsilon_1 \epsilon_2 \geqslant \frac{1}{c_2^2} - \frac{1}{16}.$$

(3) If $\tilde{p}_{1u} \leqslant \tilde{p}_{2u}$ and $\tilde{p}_{1d} > \tilde{p}_{2d}$, from Lemma 2(iii) and (27), we have

$$X_{\max} - \epsilon_1 \epsilon_2 = \min \begin{cases} \dfrac{1}{2}\left(\dfrac{1}{c_1^2} + \epsilon_1^2 + \epsilon_1\right) + \dfrac{1}{2}\left(\dfrac{1}{c_2^2} + \epsilon_2^2 - \epsilon_2\right) - \epsilon_1 \epsilon_2 \geqslant \dfrac{1}{2}\left(\dfrac{1}{c_1^2} + \dfrac{1}{c_2^2}\right) - \dfrac{1}{8}, \\ 1 + \epsilon_1 - \epsilon_2 - \epsilon_1 \epsilon_2 \geqslant \dfrac{1}{4}, \end{cases}$$

where we use the fact that $-1/8$ and $-3/4$ are the global minima for $0.5(\epsilon_1^2 + \epsilon_1 + \epsilon_2^2 - \epsilon_2) - \epsilon_1 \epsilon_2$ and $\epsilon_1 - \epsilon_2 - \epsilon_1 \epsilon_2$, respectively, subject to $|\epsilon_1|, |\epsilon_2| \leqslant 1/2$.

(4) Same as case (iii), which obtains the same lower bound for $X_{\max}$.

Summarizing all four possible cases above, we conclude that

$$\begin{aligned} \rho^{\max} &= c_1 c_2 \cdot \min\left\{\frac{1}{c_1^2} - \frac{1}{16}, \frac{1}{c_2^2} - \frac{1}{16}, \frac{1}{2}\left(\frac{1}{c_1^2} + \frac{1}{c_2^2}\right) - \frac{1}{8}, \frac{1}{4}\right\} \\ &= \min\left\{\frac{c_2}{c_1} - \frac{c_1 c_2}{16}, \frac{c_1}{c_2} - \frac{c_1 c_2}{16}, \frac{1}{2}\left(\frac{c_2}{c_1} + \frac{c_1}{c_2}\right) - \frac{c_1 c_2}{8}, \frac{c_1 c_2}{4}\right\}. \end{aligned} \tag{EC.A4}$$

Similarly, we can show that

$$\rho^{\min} = -\rho^{\max}.$$

The proof is completed. □

Appendix B. Convergence Analysis

In this section, we shall show that the proposed two-factor lattice converges to the diffusion processes that it approximates as the number of periods $N \to \infty$.

Let $\{(y_1(t), y_2(t)) \mid 0 \leqslant t \leqslant T\}$ denote the solution to (19) and (20) with initial values $(y_1(t), y_2(t)) = (y_1(0), y_2(0))$ at $t = 0$. Consider a Markov chain $\{(\bar{Y}_{1,n}, \bar{Y}_{2,n}) \mid n = 0, 1, \ldots, N\}$ that approximates $\{(y_1(t), y_2(t)) \mid 0 \leqslant t \leqslant T\}$, and let $\Delta t = T/N$. Deng and Oren (2003) in Proposition 3.1 provides a sufficient condition such that $\{(\bar{Y}_{1,n}, \bar{Y}_{2,n})\}$ converges to $\{(y_1(t), y_2(t))\}$ if the following four conditions hold:

Condition 1. The local expectations of the discrete lattice match those of the diffusion process to the dominant term:

$$\mathrm{E}[\bar{Y}_{i,n+1} \mid \bar{Y}_{i,n} = \hat{y}] = \mu_i(\hat{y})\Delta t + o(\Delta t) \quad \forall n,\ i = 1, 2. \tag{EC.B1}$$

Condition 2. The local variances and covariances of the discrete lattice match those of the diffusions processes to the dominant term. That is,

$$\mathrm{Var}(\bar{Y}_{i,n+1} \mid \bar{Y}_{i,n}) = \sigma_i^2 \Delta t + o(\Delta t) \quad \forall n,\ i = 1, 2 \tag{EC.B2}$$

and

$$\mathrm{Cov}(\bar{Y}_{1,n+1}, \bar{Y}_{2,n+1} \mid \bar{Y}_{1,n}, \bar{Y}_{2,n}) = \rho \sigma_1 \sigma_2 \Delta t + o(\Delta t) \quad \forall n. \tag{EC.B3}$$

Condition 3. There exists a function $z(\Delta t)$ such that with probability 1,

$$|\bar{Y}_{i,n+1} - \bar{Y}_{i,n}| \leqslant z(\Delta t) \quad \forall n,\ i = 1, 2. \tag{EC.B4}$$

In addition, $\lim_{\Delta t \to 0} z(\Delta t) = 0$.

CONDITION 4. The initial condition is satisfied:

$$(\bar{Y}_{1,0}, \bar{Y}_{2,0}) = (y_1(0), y_2(0)). \tag{EC.B5}$$

Next, we state the following proposition of convergence by showing that the setup of our lattice satisfies the four conditions (EC.B1)–(EC.B4).

PROPOSITION 1. *Let* $\{(y_1(t), y_2(t)) \mid 0 \leqslant t \leqslant T\}$ *denote the solution to* (19) *and* (20) *with initial values* $(y_1(t), y_2(t)) = (y_1(0), y_2(0))$ *at* $t = 0$. *The Markov chain* $\{(Y_{1,n}, Y_{2,n}) \mid n = 0, 1, \ldots, N\}$ *constructed by solving* (*Q*), (37a)–(37f) *in* §3.3, *with initial values* $(Y_{1,0}, Y_{2,0}) = (y_1(0), y_2(0))$, *converges in distribution to* $\{(y_1(t), y_2(t)) \mid 0 \leqslant t \leqslant T\}$ *as* $\Delta t \to 0$, *if* $y_i(t)$, $i = 1, 2$, *satisfies one of the following conditions*:

(1) $\mu_i(y)$ *is bounded.*

(2) $y_i(t)$ *is an MR process.*

(3) $\mu_i(y)$ *is Lipschitz continuous and there exist two constants* $\bar{y}_i > \hat{y}_i$, *such that* $\mu_i(y) < 0$ *and decreases in* y *for* $y > \bar{y}_i$, *and* $\mu_i(y) > 0$ *and decreases in* y *for* $y < \hat{y}_i$.

PROOF. To prove this proposition, we need to verify Conditions (1)–(4). First, Conditions (1) and (2) are satisfied because (*Q*) ensures that the branch probabilities of $\{(Y_{1,n}, Y_{2,n})\}$ match the local mean, local variance, and local covariance of the Markov chain to the dominant terms of those in the diffusion process $\{(y_1(t), y_2(t))\}$ by (24a), (24b), and (37e), respectively. Condition (EC.B4) is satisfied because of the initial condition $(Y_{1,0}, Y_{2,0}) = (y_1(0), y_2(0))$.

To show that Condition (3) is satisfied, we need to show that the random variable $|Y_{i,n+1} - Y_{i,n}|$ is uniformly bounded by a function that tends to zero as $\Delta t \to 0$. Because Condition (3) concerns each process individually rather than by the joint process, we omit the subscript i to simplify the notation in the following. Consider each case separately.

(1) Because $\mu(y)$ is bounded, suppose that there exists a constant M such that $|\mu(y)| < M$. From (23), we have that

$$\frac{\mu(y)\Delta t}{h} - \frac{1}{2} < \kappa \leqslant \frac{\mu(y)\Delta t}{h} + \frac{1}{2}.$$

Therefore,

$$|\kappa| \leqslant \left|\frac{\mu(y)\Delta t}{h}\right| + \frac{1}{2} < \frac{M\Delta t}{c\sigma\sqrt{\Delta t}} + \frac{1}{2} = \frac{M}{c\sigma}\sqrt{\Delta t} + \frac{1}{2}.$$

Because κ is an integer, the preceding implies that if

$$\Delta t < \left(\frac{c\sigma}{2M}\right)^2, \tag{EC.B6}$$

then $\kappa = 0$. Hence, as $\Delta t \to 0$, the random variable $|Y_{n+1} - Y_n| \leqslant h = c\sigma\sqrt{\Delta t}$, so Condition (EC.B3) is satisfied.

(2) Suppose that y is an MR process and by definition $\mu(y)$ is Lipschitz continuous. There exists a constant L such that for any $\Delta y > 0$,

$$|\mu(y + \Delta y) - \mu(y)| \leqslant L|\Delta y|. \tag{EC.B7}$$

Suppose that at some node $Y_n = \bar{y}$, its corresponding κ is equal to zero, or equivalently,

$$-\frac{1}{2} \leqslant \frac{\mu(\bar{y})\Delta t}{h} < \frac{1}{2}. \tag{EC.B8}$$

Then, in the next time period Y_{n+1} can take on three possible values: $\bar{y} + h$, $\bar{y}$, and $\bar{y} - h$.

Consider the following three cases:

(a) $Y_{n+1} = \bar{y} + h$: Because $\mu(y)$ decreases in y (by the definition of an MR process) and $\mu(y)$ is Lipschitz continuous, using (EC.B7) we have that

$$\mu(\bar{y}) - Lh \leqslant \mu(\bar{y} + h) \leqslant \mu(\bar{y}).$$

Multiplying the preceding by $\Delta t/h$ and applying (EC.B8), we then have

$$-L\Delta t - \frac{1}{2} \leqslant \frac{\mu(\bar{y} + h)\Delta t}{h} < \frac{1}{2}.$$

Therefore, if $-L\Delta t > -1$, or equivalently,

$$\Delta t < \frac{1}{L}, \tag{EC.B9}$$

the corresponding κ for $Y_{n+1} = \bar{y} + h$ is either 0 or -1.

(b) $Y_{n+1} = \bar{y}$: In this case, the corresponding κ for $Y_{n+1} = \bar{Y}$ is 0.

(c) $Y_{n+1} = \bar{y} - h$: Similar to (a). If Δt satisfies (EC.B9), the corresponding κ for $Y_{n+1} = \bar{y} - h$ is either 0 or 1.

Summarizing from these three cases, for Δt satisfying (EC.B9), if $\kappa = 0$ in one time period, then in the next time period there will be only three possible values for κ, namely, -1, 0, and 1. By defining

$$r_u(\Delta t) \equiv \inf\left\{ y \,\middle|\, \frac{\mu(y)\Delta t}{h} \leqslant -\frac{1}{2} \right\}$$

and

$$r_l(\Delta t) \equiv \sup\left\{ y \,\middle|\, \frac{\mu(y)\Delta t}{h} \geqslant \frac{1}{2} \right\},$$

we can assert that if $Y_0 = y(0) \in [r_l(\Delta t) - h, r_u(\Delta t) + h]$, then the whole process $\{Y_n \mid n = 0, 1, \ldots, N\}$ will always lie in $[r_l(\Delta t) - h, r_u(\Delta t) + h]$. Consequently, in the whole lattice there are only three possible values for κ, namely, -1, 0, and 1. Hence, for Δt satisfying (EC.B9), $|Y_{n+1} - Y_n|$ is uniformly bounded by $2h = 2c\sigma\sqrt{\Delta t}$, which tends to 0 as $\Delta t \to 0$.

To complete the proof, note that $r_u(\Delta t) \to \infty$ and $r_l(\Delta t) \to -\infty$ as $\Delta t \to 0$. Therefore, we can always find a small enough Δt to ensure that $y(0) \in [r_l(\Delta t) - h, r_u(\Delta t) + h]$.

(3) We can prove this case by combining the results from the previous two cases. Because $\mu(y)$ is continuous, the set $\{|\mu(y)| \mid \hat{y} \leqslant y \leqslant \bar{y}\}$ is compact and therefore bounded. Let $M \equiv \max\{|\mu(y)| \mid \hat{y} \leqslant y \leqslant \bar{y}\}$ and let L denote the constant so that $\mu(y)$ satisfies the Lipschitz continuity in (EC.B7). According to (EC.B6) and (EC.B9), we can conclude that for

$$\Delta t < \min\left\{ \left(\frac{c\sigma}{2M} \right)^2, \frac{1}{L} \right\},$$

if $\kappa = 0$ in one time period, in the next time period there will be only three possible values for κ, namely, -1, 0, and 1. The rest of the proof follows that in case 2. □

Proposition 1 shows that the convergence of the lattice relies on the property of the drift function. The lattice can converge for at least three types of drift function $\mu(y)$ given in the proposition, including that for an MR process. Note that $\mu(y)$ determines the branching factor κ (23). An intuitive interpretation is that if $\mu(y)$ changes too quickly over a small interval of y (e.g., not Lipschitz continuous), the movements of branches might become very large over a small time interval and cause the discrete distribution not to converge to the underlying, continuous one.

Appendix C. Relieving the Impact of the Ramp Constraints

It is well known that the SDP (lattice) approach can only handle problems with constraints that are path-independent on the network of the state space. That is, each state should be independent of the paths that lead to it. The ramp constraints (6) are path-dependent unless we expand the state space of $\{x_t\}$ to include discrete levels of the generation q_t. Recall that in the proposed lattice model at each node (y_1, y_2) at time t, there is an associated optimal generating level $g(y_1, y_2)$ (a special case with $y_1 = P_t^E$ and $y_2 = P_t^F$ is given in (12)). So, when a branch (an arc) moves from one node to another in the next time period, the difference between the two optimal generation levels associated with these two nodes determines the satisfaction of the corresponding ramp constraint. If this movement violates the ramp constraint, it means that the change of the prices over this time period (or equivalently, the time step Δt) is too big. Therefore, one may want to reduce the value of Δt to reduce the price change over each time period. It can also be expected that reducing the value of Δt can decrease the number of branches that violate the ramp constraints, although it may not be able to eliminate the violation of the ramp constraints completely. To model this effect, we present a continuous version of the ramp constraint (6) when the time step $\Delta t \leqslant 1$:

$$|q_t - q_{t-\Delta t}| \leqslant R\Delta t, \tag{EC.C10}$$

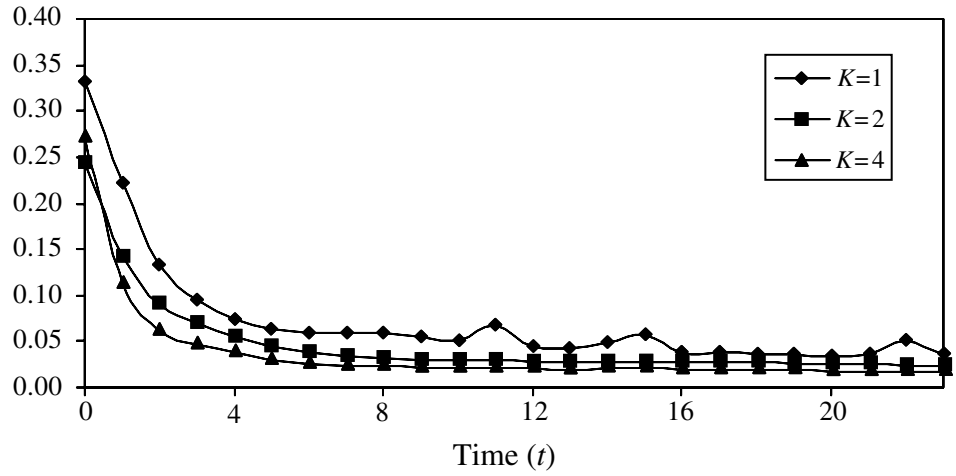
where we assume that the ramp rate is a linear function of Δt for simplicity. Given a node (y_1, y_2) at time t and a branch incident from this node, if reducing Δt can eventually satisfy the ramp constraint for this branch, that means

$$|g(y_1, y_2) - g(y_1 + (\kappa_1 + 1)h_1, y_2 + (\kappa_2 - 1)h_2)| \leqslant R\Delta t \tag{EC.C11}$$

holds for some small Δt. Note that in (EC.C11), because we try to bound the change of the generation levels from above, we consider the transition that yields the largest change in g (with high electricity price $(\kappa_1 + 1)$ and low fuel price $(\kappa_2 - 1)$). Dividing Δt on both sides of (EC.C11) and letting $\Delta t \to 0$, we have

$$\left| \frac{\partial g}{\partial y_1}\mu_1(y_1) + \frac{\partial g}{\partial y_2}\mu_2(y_2) \right| \leqslant R, \tag{EC.C12}$$

FIGURE EC.1. The percentage of branches violating the ramp constraint at each time period.



if the above partial derivatives exist at the node. (Note that g is continuous but may not be differentiable at all points.) In general, if the partial derivatives exist, $\partial g/\partial y_1 \geqslant 0$ and $\partial g/\partial y_2 \leqslant 0$. It is clear that (EC.C12) compares the increasing rate (or the total differential) of g and the constant ramp rate R. Equation (EC.C12) may be viewed as a (necessary and sufficient) condition to verify whether reducing Δt can eventually meet the ramp constraints.

It is worth noting that (EC.C12) may hold for *all* nodes under some special price processes. That means *all* the ramp constraints can be automatically satisfied by setting a Δt small enough. Equivalently, the ramp constraints can be eliminated. This, unfortunately, is not the case for the MR processes used in our numerical tests. However, numerical tests indicate that by slightly reducing the value of Δt, the majority of the branches indeed satisfy the ramp constraints. Furthermore, the impact due to the branches that violate the ramp constraints is estimated to be insignificant. Therefore, reducing Δt is an effective strategy to relieve the impact of the ramp constraints. Details of the numerical tests are presented next.

Consider the operation of the same unit used in §4.3 over a 24-hour period. Three different values of Δt are tested: $\Delta t = 1$ $(K = 1)$, $\Delta t = 0.5$ $(K = 2)$, and $\Delta t = 0.25$ $(K = 4)$. The number of branches that violate the ramp constraint are counted in each time period, and are then divided by the total number of branches in the same time period to yield a ratio, which is depicted in Figure EC.1. For example, when $\Delta t = 1$, there are nine branches incident from the initial node at $t = 0$, one-third (0.33) of them violate the ramp constraint. This ratio is reduced to 0.24 and 0.27 when $\Delta t = 0.5$ and 0.25, respectively. Note that when the value of Δt is reduced to 0.5 and 0.25 from 1, the total number of the branches also increases from 9 to 90 and 756 within the first time period $(t = 0)$, respectively. In general, the percentage of the branches that violate the ramp constraint is very low and decreases as Δt decreases, except at $t = 0$.

If the corresponding branching probability of each branch that violates the ramp constraint is also considered, the violation of the ramp constraint overall accounts for 4.4% of the branching probability when $\Delta t = 1$. This probability is reduced to 2.75% and 2.0% when Δt is reduced to 0.5 and 0.25, respectively. On the other hand, from Table 2, when $\Delta t = 1$ is reduced to 0.5 and 0.25, the plant value increases 1.6% and 2.3%, respectively. Roughly, every 1% of improvement of the branching probability that violates the ramp constraint is associated with a 1% increase in power plant value.

While indeed the ramp constraint may not be satisfied at all branches in the proposed lattice model, this test result suggests that its impact is very limited, which can be further relieved by reducing Δt.

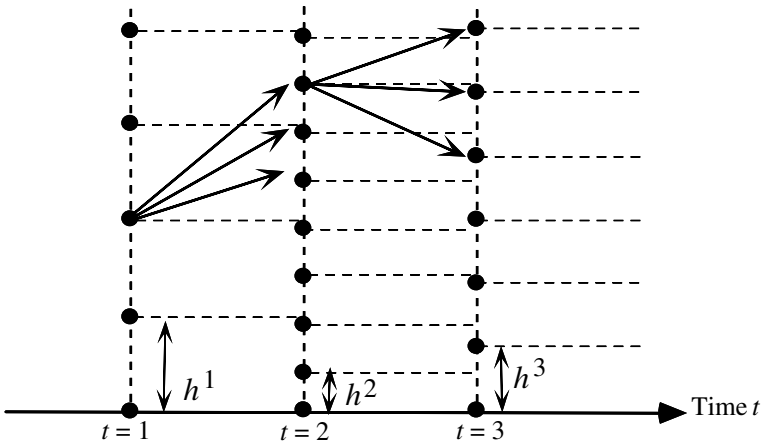
Appendix D. Extensions of the Proposed Valuation Model

This section discusses potential extensions of the proposed valuation model, including time-dependent drift and volatility, long-term valuation, and incorporation of network constraints.

EC.D1. Time-Dependent Drift and Volatility

Drift function and volatility may be time dependent. For example, the mean level of the electricity price at each hour may be different following some hourly pattern, such as peak hours and off-peaks hours. The volatility in

FIGURE EC.2. A one-factor lattice with time-dependent volatility.



the peak hours may be different from that in the off-peak hours. The lattice model developed in the preceding sections can be easily extended to accommodate time-dependent drift and volatility.

Recall the definition of branching factor κ in (23). If the drift function $\mu(y)$ is now time-dependent, say $\mu_t(y)$, one can simply replace $\mu(y)$ in (23) by $\mu_t(y)$ for determining κ. Given a lattice node y_t at time t, (23) also implies that

$$y_t + \mu(y_t)\Delta t \in \left[y_t + \left(\kappa - \tfrac{1}{2}\right)h,\, y_t + \left(\kappa + \tfrac{1}{2}\right)h\right). \tag{EC.D13}$$

Because y_t is also a lattice node at $t+1$, the branching factor κ, defined by (EC.D13), shows the relative positions of the three nodes to be branched into with respect to (w.r.t.) y_t.

Consider the following one-factor process with time-dependent drift and volatility:

$$dy = \mu_t(y)dt + \sigma_t dB. \tag{EC.D14}$$

Because the volatility determines the (discrete) price increment h of the lattice, with a time-dependent volatility σ_t the price increment may be different from hour to hour (see Figure EC.2).

Define the price increment at hour t as

$$h^t = c\sigma_t\sqrt{\Delta t}. \tag{EC.D15}$$

A superscript t is used to distinguish h^t from h_1 and h_2 previously defined in the two-factor lattice. Because now a lattice node y_t at time t may not be a lattice node at time $t+1$, it is better to change the notion of κ from a *relative* position w.r.t. y_t to an *absolute* position w.r.t. a common reference point such as zero. We thus define the time-dependent branching factor κ_{t+1} as follows:

$$\kappa_{t+1} \equiv \left\lfloor \frac{y_t + \mu_t(y_t)\Delta t}{h^{t+1}} + \frac{1}{2} \right\rfloor. \tag{EC.D16}$$

With (EC.D16), a lattice node y_t at time t now branches into three nodes $(\kappa_{t+1}-1)h^{t+1}$, $\kappa_{t+1}h^{t+1}$, and $(\kappa_{t+1}+1)h^{t+1}$ at time $t+1$. Accordingly, the definition of ϵ in (26) is modified by

$$\epsilon_{t+1} \equiv \frac{y_t + \mu_t(y_t)\Delta t}{h^{t+1}} - \kappa_{t+1} \tag{EC.D17}$$

and the following property of ϵ_t remains valid:

$$|\epsilon_t| \leqslant \tfrac{1}{2} \quad \forall\, t. \tag{EC.D18}$$

Therefore, all properties of the lattice (both one-factor and two-factor) derived previously follow.

EC.D2. Long-Term Valuation Model

There is no doubt that the generation asset valuation must account for the operational constraints, whether it be short term (weeks) or long term (years). However, because the operational constraints, such as the minimum uptime and downtime constraints, can only be modeled via hour-by-hour unit commitment (UC), it is not a trivial task to integrate a long-term valuation model with physical constraints. We believe that the proposed model can be directly applied to long-term asset valuation, say over a 10-year period. The only concern is that whether the underlying price model can capture longer term price signals (such as those provided by forward price curve). This can be handled by converting the price information observed in the forward curves to the "mean" price levels of the MR process. For example, if the MR process follows the standard Ornstein-Uhlenbeck process:

$$\mu_i(y_i) = \eta_i(m_{it} - y_i), \quad i = 1, 2, \tag{EC.D19}$$

where η_i is the speed of reversion and m_{it} is the mean level of y_i at time t, one can set the (hourly) profile of the mean level $\{m_{it}\}$ such that it reflects the forward curve. Note that a monthly forward price can be viewed as an average of all hourly prices over a month. Therefore, some scaling may be necessary to convert forward prices to hourly prices. In addition to setting the drift function of the price model, one can also set the volatility that varies from time to time over the entire planning period, which has been covered in §EC.D1. Because we have shown that the size of the proposed two-factor lattice will be capped for the MR underlying processes, the computational complexity for the version of long-term valuation will only increase linearly, as does the CPU time. Therefore, directly applying the proposed valuation to a long-term period is technically feasible.

EC.D3. Incorporating Network Constraints

In Valenzuela and Mazumdar (2003), the authors showed that with the existence of a competitive spot market, the UC problem may become decomposable if the spot market has an infinite capacity and can trade (including buying and selling) electricity instantaneously. In this situation, the UC problem can be decomposed to independent unit subproblems. An intuitive explanation is that the power imbalance of the demand constraint can be offset by the market through trading instantaneously. This coincides with our rationale to discuss single unit asset valuation in a competitive market with spot markets. When the network constraints exist, one can no longer assume that the spot market (presumably near the power plant of interest) has an infinite capacity due to transmission constraints. In this case, it makes better sense to consider the valuation of multiple generating units over various locations as a portfolio, which is beyond the scope of this paper. However, to consider the impact of the network constraints on a specific unit at a specific location, one can model the network congestion in terms of the capacity of the unit. For example, the maximal and minimal capacities of the unit may be time dependent and/or even random:

$$q_t^{\min} \leqslant q_t \leqslant q_t^{\max}. \tag{EC.D20}$$

This can be easily incorporated in the proposed lattice model.

References

See references list in the main paper.

Deng, S., S. Oren. 2003. Incorporating operational characteristics and start-up costs in option-based valuation of power generation capacity. *Probab. Engrg. Inform. Sci.* **17** 155–181.

Valenzuela, J., M. Mazumdar. 2003. Commitment of electric power generators under stochastic market prices. *Oper. Res.* **51** 880–893.