Linear Complementarity Models of Nash–Cournot Competition in Bilateral and POOLCO Power Markets

Benjamin F. Hobbs, Member, IEEE

Abstract—Two Cournot models of imperfect competition among electricity producers are formulated as mixed linear complementarity problems (LCPs), and a simple example is presented to illustrate their application. The two models simulate bilateral markets. The models include a congestion pricing scheme for transmission, but other transmission pricing approaches can also be represented in this framework. The two models differ from each other in that one has no arbitrage between nodes of the network, while in the other model, arbitrages erase any noncost based differences in price. The latter bilateral model turns out to be equivalent to a Cournot model of a POOLCO. The models differ from other Cournot market models in that they include both of Kirchhoff's laws via a DC approximation; can include arbitrages; possess unique solutions; and are readily solved by efficient LCP algorithms. The key assumption that permits their formulation as LCPs is that each producer naively assumes that its output will not affect transmission prices.

Index Terms—Complementarity, Cournot, electricity competition, electricity generation, market models, strategic pricing.

I. INTRODUCTION

R ESTRUCTURED power markets take a wide variety of forms. Their design and structure differ in many important ways, such as how transmission is priced, and whether generators sell to a central auction (POOLCO) or bilaterally to customers [1]. These and other differences affect the nature and outcome of competition among power producers; therefore, market models should reflect those differences.

A wide range of models are proposed for simulating the interaction of competing generation companies who price strategically [2]–[4]. Such models can be used to identify how market power might be wielded in restructured power markets and the impacts of proposed mergers. This paper presents two specific models, that like some previous models [5]–[7], adopt a Nash–Cournot game theoretic framework [8]–[14] and represent transmission constraints by a linearized DC network. Unlike previous models, however, the formulations presented here account for arbitrage, readily lend themselves to computation even for large markets with hundreds of transmission interfaces and nodes, and guarantee the existence of unique price equilibria. With very few exceptions [7], other applications of Cournot models are to small networks of 3–7 busses and potentially have either no price equilibria or many equilibria [15].

The two models presented here are based on the following market assumptions. They address a bilateral market in which imperfectly competitive generators purchase transmission services from an ISO who prices scarce transmission capacity in order to ration it efficiently. In the first model, there is no arbitraging between different locations in the network; this allows noncost based price differences to arise, so that generators can raise prices where competition is weak or demand is inelastic while competing more intensely elsewhere. In the second model, there are arbitrages/marketers who eliminate price differences between locations that are unjustified by cost [16]. This is shown below to be equivalent to a POOLCO-based system using locational marginal pricing.

In terms of strategies, each generating company in both models plays a Nash game in quantities sold. This is equivalent to each generation company assuming that other firms will not alter their outputs—a Nash–Cournot game. In addition, each generator naively assumes that its outputs will not significantly affect transmission prices. In game theoretic terms, this a Bertrand game with respect to transmission. This belief about transmission diverges from the Nash–Cournot models in [5], [6]. In the latter models, sophisticated producers recognize transmission limits and correctly predict the effect of their decisions on the transmission prices. However, such models are not numerically tractable for large systems. In contrast, the Bertrand assumption does not permit simulation of some strategies for manipulating transmission; however, the resulting models are solvable for realistically large systems. Thus, the use here of a Bertrand game for transmission is a compromise between the objectives of:

- realism in representation of strategic behavior,
- realism in representation of physical constraints, and
- computability.

Given the above market and strategy assumptions, both models calculate a market equilibrium for generation and transmission. A market equilibrium is defined as a set of prices, generator outputs, transmission flows, and consumption that simultaneously satisfy each market participant’s first order conditions for maximization of its profit while clearing the market (supply = demand). A solution satisfying those conditions possesses the property that no participant will want to alter its decisions unilaterally: a Nash equilibrium. Smeers [17] concludes his survey of gas and electric market models by

Manuscript received May 24, 1999; revised May 26, 2000. This work was supported by the National Science Foundation under Grant ECS9696014, and the FERC Office of Economic Policy.

The author is with the Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD 21218 (e-mail: bhobbs@jhu.edu).

Publisher Item Identifier S 0885-8950(01)03781-6.
arguing that explicit statement and solution of equilibrium conditions is a promising theoretical and computational approach to modeling strategic behavior.

Such equilibrium conditions are obtained here by deriving first order and market clearing conditions and solving them simultaneously. The first order (Kuhn–Kareesh–Tucker/KKT) conditions for a constrained optimization problem \( \text{MAX} \ F(x, y) \) subject to \( G(x, y) = 0 \) and \( H(x, y) \leq 0 \), \( x \geq 0 \), are:

\[
\begin{align*}
& x; \partial F / \partial x - \lambda \partial G / \partial x - \mu \partial H / \partial x \leq 0; \quad x \geq 0; \\
& y; \partial F / \partial y - \lambda \partial G / \partial y - \mu \partial H / \partial y = 0 \\
& \lambda; G(x, y) = 0 \\
& \mu; H(x, y) \leq 0; \quad \mu \geq 0; \quad \mu H(x, y) = 0
\end{align*}
\]

where \( \lambda \) and \( \mu \) are the dual variables for constraints \( G \) and \( H \), respectively. The equations associated with the nonnegative variables \( x \) and \( y \) are called complementarity conditions. The models of this paper are created by combining the KKT conditions for all the market participants and then adding equality conditions to represent clearing of the market. The resulting problem involves both equality and complementarity conditions, and is termed a mixed complementarity problem (MCP). A general MCP is defined as follows: find vector \( z \) given functions \( f(z) \) and \( g(z) \) such that \( z \geq 0 \), \( f(z) \geq 0 \), \( f(z)^T z = 0 \), and \( g(z) = 0 \), where \( z \), \( f(z) \), and \( g(z) \) are vectors. If \( f(z) \) and \( g(z) \) are affine, then this is a mixed linear complementarity problem, or a mixed LCP.

The models of this paper are mixed LCPs as a result of using linear demand functions and marginal generation costs. However, the models can be generalized to the nonlinear case, yielding nonlinear complementarity problems, NCPs. Direct solution of the market equilibrium conditions by complementarity methods has important computational advantages. Mixed LCPs involving thousands of variables and complementarity conditions can be solved using available LCP software, such as implementations of Lemke’s algorithm [18] and the MILES and PATH solvers within GAMS [19]. This permits application of strategic market models to large systems with thousands of power plants and hundreds of constrained transmission interfaces.

In the next section, previous models that are based upon explicit statement and solution of equilibrium conditions are reviewed. Then, the assumptions of the two models are summarized. A presentation of the two models follows, along with illustrative results for a simple three-bus system.

II. RELATED MODELS

A number of power market models are proposed in the literature that calculate price equilibria in two steps: 1) formulation of a set of conditions that directly state the market equilibrium conditions of profit maximization and market clearing; and 2) numerical solution of those equations. These conditions are most commonly phrased as a MCP.

In an early paper, Schmalensee and Golub [20] calculate a Cournot equilibrium in each of 170 US market areas, considering power producers who own 871 generating plants. In their model, the equilibria are calculated for each area separately without considering interactions with other areas, and transmission is represented assuming tariffs and losses are proportional to distance. “Pseudosupply functions” are derived for privately owned sellers based upon the first-order conditions for profit maximization by a Cournot producer; i.e., a producer who implicitly assumes that other producers will not alter their sales. These conditions are solved for all firms at once—but only for one market at a time—and then the well-known Herfindahl concentration index is calculated for each market area.

Later work based on equilibrium conditions attempts to make market simulations more realistic by considering all markets simultaneously while recognizing transmission capacity limitations. Several power market models of this type in which producers behave competitively rather than strategically have appeared. One is PMDAM [21] which iteratively adjusts prices at nodes in the network for all periods until the market clears; very large systems, such as the entire western US, have been solved in this way. Another is Qi and Harker [22]. They solve a NCP in which power producers in eastern North America compete to supply power in an aggregate 12 link—9 node transportation network in which Kirchhoff’s voltage law is not imposed. The model is nonlinear because demand curves were assumed to be of the constant elasticity form \( (P = aQ^b) \). Existence of an equilibrium is proven. Finally, Boucher and Smeers [23] derive equilibrium conditions for several variations of POOLCO and bilateral markets under the assumptions of perfect competition and efficient rationing of transmission capacity. They prove that the alternatives yield the same market prices and efficiency, confirming earlier assertions by others e.g., [24]. Their models impose general linear constraints upon transmission flows, which include DC load flow models and nomograms or other reliability constraints as special cases. The result of Section IV-B, below, that POOLCO and bilateral (with arbitrage) markets yield the same equilibrium for Cournot generators can be viewed as an generalization of their results to imperfect competition.

The remainder of the models reviewed in this section, as well as the models proposed in this paper, consider the possibility of strategic behavior by power producers. Jing-Yuan and Smeers [11] present a model that directly solves the equilibrium conditions for a bilateral power market on a radial network. Each Cournot producer assumes that other producers will not change their output, sales, or the flows they induce on the network, while recognizing the presence of transmission constraints. Transmission tariffs can be set by a regulated body to recover grid costs using either distance-dependent or postage-stamp fees. The authors point out that such a system can have multiple equilibria with widely diverging effects on the profits and outputs of individual firms. That is, their MCP generally has multiple solutions.

However, Jing-Yuan and Smeers [11] use a variational inequality (VI) solution approach to solve for one of the possible equilibria. They prove that a VI solution exists and is unique, even if, as Stoft [15] shows, there are actually multiple Cournot equilibria. In contrast, the mixed LCP models defined in this paper and [7] yield unique solutions, implying that the market equilibria are unique. This is made possible by making the simplifying Bertrand assumption that each generator does not anticipate how its actions will affect congestion and transmission.
prices. One of the reasons for this result is that in order to represent the effect of generation decisions on transmission prices, it is necessary to embed the KKT conditions for the grid operator’s optimal power flow problem within each generator’s optimization problem [4]–[6]; the resulting “mathematical program with equilibrium constraints” is highly nonconvex and could have multiple local optima. In contrast, the Bertrand transmission assumption means that the generator’s model is much simpler. It includes just linear terms for transmission costs in the objective function.

The multiple solutions of the formulation in [11] underlie much of the debate between Oren [5] and Stoft [15] about i) the realism of this type of Cournot model and ii) its apparent implication that Cournot generators will eliminate congestion in order to force transmission congestion charges to zero, even if there are very many producers. As Stoft [25] points out, this lack of uniqueness occurs in formulations such as [11] because of the absence of markets for transmission capacity. However, in our models and those in [7], creation of a competitive market for transmission services yields a unique price equilibrium. Basically, the eliminated equilibria are solutions in which the marginal valuations by different firms of the same transmission capacity diverge.

Smeers and Jing-Yuan [7] present a model equivalent to that of Section IV-A below in which there are markets for energy and transmission capacity, generators adopt Cournot strategies in the energy market, transmission capacity is rationed efficiently, power flows over a linearized DC network, and no arbitragers exist to erase noncost-based differences in energy prices between different locations. This model, like [11], also includes the possibility of generation capacity expansion. Smeers and Jing-Yuan prove existence and uniqueness of a market solution, which they obtain by VI methods. Presently, the model is being applied to the EU power system, and includes thousands of variables and equilibrium conditions. The models of Section IV-B below can be regarded as an extension of this approach to POOLCO and arbitrageder bilateral power markets.

Two proposed market models explicitly include the first order conditions for Cournot producers for intertemporal power production decisions while omitting transmission constraints. Bushnell [10] considers how Cournot producers would allocate hydropower over time. An iterative price-adjustment approach similar to PMDAM [21] is used. He finds that Cournot producers produce less on peak (thus raising prices at that time) and more off-peak compared to producers who behave competitively.

The second such model, the dynamic model of Ramos et al. [12], instead focuses on unit commitment over a 24 hour planning period. Their model is the only Cournot equilibrium model that includes integer variables, which represent commitment decisions. Firms are Cournot producers. Their profit-maximizing behavior is cleverly captured in the model by a constraint that the marginal revenue MR earned by a generating unit in a given hour must be at least equal to its marginal running cost MC if the unit is committed, while MC can exceed MR for uncommitted units.

Their formulation is correct for a single hour commitment problem; however, it can cause difficulties for multiperiod problems. In real power systems, prices and thus MR often fall below MC for committed units during low load periods because its owner decides it is more profitable to keep the unit committed rather than to turn it off and then later ramp it back up. However, this difficulty can be handled by defining additional integer variables that would relax the MR ≤ MC constraint when plant output is at its minimum run level. Alternatively, the objective function of the model could be modified in the manner proposed at the end of Section IV-A, below. A second difficulty with the model is that the added realism represented by integer variables unfortunately implies that existence or uniqueness of Cournot equilibria may be impossible to prove. Indeed, the infamous “duality gap” of unit commitment models can mean that, in general, no market clearing equilibrium exists even for perfect competition (nonstrategic) models of this type.

A model that focuses on the role of imperfectly competitive arbitragers is Smeers and Jing-Yuan [14]. Their model assumes that generators behave competitively, but that each arbitrageur assumes that rival arbitragers will not alter the amounts they buy and sell and the resulting DC power flows—-a Cournot assumption. In contrast, the model of Section IV-B, below, represents strategic generators, while low barriers to entry imply that arbitragers behave competitively. But like the models of this paper, [14] assumes that transmission rights are traded in a market and that market participants believe that they cannot affect the prices of those rights. Numerical examples in [14] indicate that prices are likely to converge to competitive levels as the number of arbitragers grows.

III. MODEL ASSUMPTIONS

Each producer $f$ owns power generating facilities $h = 1, \ldots, H(f, i)$ located at nodes $i$ of the network. The indices $i$ and $j$ designate nodes. $C_{fih}$ is the per MWh cost of power generation $x_{fih}$, in MW. The capacity of a generator is $X_{fih}$, MW.

Consumers at a node $i$ consume $q_i$, MW, which is price responsive. In order to use the LCP framework, we assume linear demand functions $p_i(q_i) = P_{io} - (P_{io}/Q_{io})q_i$, $$/MWh, with $P_{io}$ $$/MWh and $Q_{io}$, MW being the price and quantity intercepts, respectively. Nonlinear demand functions (as in [22]) would yield a NCP, which generally are more difficult to solve. It is assumed that this is a bilateral market, in which $s_{fj}$ MW is the quantity sold by producer $f$ to consumers at node $j$. Assuming market clearing and no arbitrage, $\Sigma_f s_{fj} = q_j$. If there is arbitrage, then $\Sigma_f s_{fj} + a_j = q_j$, where $a_j$, MW is the net amount of power sold by arbitragers to node $j$. The generators determine the level of sales to each node, and then request transmission service from the grid. An energy balance is imposed on each firm: $\Sigma_{i,h} x_{fih} = \Sigma_{j} s_{fj}$.

There exist a variety of transmission pricing policies that could be simulated. In these models, it is assumed that transmission is priced using a congestion pricing scheme [26], or its functional equivalent, a Chao–Peck [27] market for interface capacity. The owner of the grid charges a congestion-based wheeling fee $W_i$, $$/MWh for transmitting power from an arbitrary hub node to node $i$. For simplicity, it is assumed that there is neither generation nor consumption at the hub. Because
of the linearity of the DC network [26], all generation and sales can be modeled as being routed through the hub node. A firm pays $-W_i$ to get power to the hub from a generator at $i$ and then pays $+W_j$ to convey power for sale from the hub to customers at $j$. Thus, the total cost of transmitting power from a generator at $i$ to the point of sale at $j$ is $-W_i + W_j$.

The total transmission service that the grid provides for power transferred from the hub to a node $j$ is defined as $y_j$ MW, which may be negative. Consistent with the linear DC approximation [26], flows through interfaces $k$ are modeled using power transmission distribution factors; i.e., the net MW flow through $k$ is $\sum_i P_{TDF,k} y_i$. The lower and upper bounds on real power flows through an interface $k$ are $-T_{k-}$ and $T_{k+}$. We assume no losses and that congestion is only basis for pricing. However, more general assumptions can be accommodated in a LCP e.g., see [26]. For instance, zonal pricing or transmission path pricing can be imposed. However, in order to respect physical constraints, it would be necessary to assume that the grid operator in addition imposes an efficient nonprice mechanism to relieve congestion similar to the UK constrained-on and constrained-off approach.

The owner of the grid is assumed to ration limited interface capacity to maximize the value of the transmission services $y_j$, as expressed by generators’ willingness to pay. This behavior can be shown to be equivalent to having the grid choose values of $y_j$ to maximize its revenue $\sum_i W_i y_i$ as if the $W_i$ are fixed, while respecting interface constraints. It is also equivalent to a competitive market for transmission rights in which generators do not exercise market power [7], [25], [27].

A final assumption concerns arbitragers. As in [7], they are presumed to be absent in the first model; consequently, the differences between prices at nodes $i$ and $j$ can diverge from the cost $-W_i + W_j$ of transmitting power from $i$ to $j$. This is termed “spatial price discrimination.” A firm can then optimize sales to each node without worrying about how those sales will affect its sales or prices at other nodes. The no-arbitrage model is conceptually related to the continuous spatial competition model of Greenhut and Greenhut [28]. There, separate Cournot equilibria are calculated for each demand point among spatially separated producers (but unlike [20], these equilibria are calculated simultaneously). As a result, spatial price differences do not necessarily reflect transport cost differentials. Hashimoto [29] and Kolstad and Abbey [30] implemented that general approach for network-based coal market models. Schmalensee and Golub’s [20] Cournot analysis of power markets implicitly embodies the assumption that price discrimination can persist over space. But these analyses did not consider link capacity limits or networks that adhere to Kirchhoff’s voltage law, unlike [7].

The second model presented below recognizes that arbitrage will occur. Arbitragers are also explicitly considered in Qui and Harker’s [22] perfect competition model and Jing-Yuan and Smears’ [14] model of strategic power marketers. The arbitrage model assumes that arbitragers are price takers, and will sell power from $i$ to $j$ as long as $p_i(y_i) - W_i + W_j < p_j(y_j)$. Thus, in equilibrium, $p_i(y_i) - W_i = p_j(y_j) - W_j = p_H$, the hub price. Generators are assumed to recognize that this will occur, so that if firm $f$ shrinks its sales at node $j$, then the resulting increase in $j$’s price will create arbitrage opportunities; when arbitragers exploit these opportunities, prices and quantities demanded will be affected at other nodes. This yields different behavior and price equilibria than the nonarbitraged model, as Section V’s application shows.

IV. THE MODELS

A. No-Arbitrage Model

We first present the producers’ and grid owner’s optimization problems; combining their KKT conditions with the market clearing condition then yields a mixed LCP.

1) Producers: The producer’s equilibrium conditions result from the KKT conditions for the following quadratic program. This model states that producer $f$ chooses generation $x_{fih}$ and sales $s_{fj}$ in order to maximize profit ($$/hr), equal to revenue minus transmission and generation costs:

$$\text{MAX} \sum_f [(P_{f0} - (P_{f0}/Q_{f0})(\sum y_i s_{fj})) - W_j s_{fj}]$$

subject to:

$$x_{fih} \leq X_{fih} \quad (\rho_{fih}) \quad \forall \text{generators } i, h$$

$$\sum_j s_{fj} = \sum_i h x_{fih} \quad (\theta_f) \quad \forall s_{fj}, x_{fih} \geq 0$$

$$\rho_{fih} \text{ is the dual multiplier for the generator capacity constraint, while } \theta_f \text{ is the dual for the energy balance, interpretable as } f’s \text{ marginal cost at the hub. Note that all } s_{fj} \text{ for } g \neq f \text{ are assumed fixed—the Nash–Cournot assumption.}$$

Of course, the above model of dispatch is simplistic. However, the mixed LCP can accommodate more realistic assumptions, such as increasing marginal costs, minimum run levels, fuel choice options, emissions allowances and tonnage limits, and energy storage. Generation capacity expansion can also be modeled, if capacity is represented as a continuous variable with no scale economies, as in [7], [11].

The KKT conditions for generator $f$’s problem are:

- For $s_{fj}, \forall j$:
  $$[(P_{f0} - (P_{f0}/Q_{f0}) (2 s_{fj} + \sum y_i s_{fj})) - W_j] - \theta_f \leq 0;$$
  $$s_{fj} \geq 0;$$
  $$\sum_j s_{fj} = \sum_i h x_{fih};$$
  $$\{[(P_{f0} - (P_{f0}/Q_{f0}) (2 s_{fj} + \sum y_i s_{fj})) - W_j] - \theta_f \} = 0$$

(f1)

- For $x_{fih}, \forall i, h$:
  $$- (C_{fih} - W_i) - \rho_{fih} + \theta_f \leq 0;$$
  $$x_{fih} \geq 0;$$
  $$\{-(C_{fih} - W_i) - \rho_{fih} + \theta_f \} = 0$$

(f2)

- For $\rho_{fih}, \forall i, h$:
  $$x_{fih} \leq X_{fih};$$
  $$\rho_{fih} \geq 0;$$
  $$\rho_{fih} (x_{fih} - X_{fih}) = 0$$

(f3)

- For $\theta_f$:
  $$\sum_j s_{fj} = \sum_i h x_{fih},$$

(f4)

2) Grid Owner: The grid’s equilibrium conditions result from the KKT conditions of the following LP. The grid chooses
y; to maximize its profit from bilateral transactions, adopting the naive Nash–Bertrand assumption that it cannot affect the fees it gets for providing transmission:

\[
\text{MAX } \sum_i W_i \quad \text{s.t.}
\]

\[
-\sum_k PTDF_{ik} y \leq T_{h-k} \quad (\lambda_{h-k}) \quad \forall \text{ interfaces } k,
\]

\[
\sum_k PTDF_{ik} y \leq T_{k-h} \quad (\lambda_{k-h}) \quad \forall k.
\]

The \( \lambda \)s are the duals associated with the interface constraints. The KKT conditions defining the optimal solution are:

- For \( y_i \), \( \forall i \):

\[
W_i + \sum_k PTDF_{ik} (\lambda_{h-k} - \lambda_{k-h}) = 0 \quad (G1)
\]

- For \( \lambda_{h-k}, \forall k \):

\[
-\sum_i PTDF_{ik} y_i \leq T_{h-k} - \lambda_{h-k} \geq 0; \quad \lambda_{k-h} (\sum_k PTDF_{ik} y_k + \lambda_{k-h}) = 0 \quad (G2)
\]

- For \( \lambda_{k-h}, \forall k \):

\[
\sum_i PTDF_{ik} y_i \leq T_{k-h}; \quad \lambda_{k-h} \geq 0; \quad \lambda_{h-k} (\sum_k PTDF_{ik} y_k - T_{h-k}) = 0. \quad (G3)
\]

3) Market Clearing: The total transmission service demanded by generators from the hub to any \( i \) must equal the transmission service the grid provides between those nodes:

\[
\sum_j s_{ij} = \sum_k h x_{fih} = y_i \quad \forall i. \quad (MC1)
\]

Solution Approach: Gathering the producers’ and grid’s KKT conditions and the market clearing equations results in a (perhaps very large) set of conditions \([f1]–[f4], \forall f; (G1)–(G3); \) and \( (MC1) \). The resulting problem is a mixed LCP. Solving these equations simultaneously for the following primal and dual variables produces an equilibrium to the no-arbitrage market game: \([s_{ij}, x_{fih}, y_i, W_i, \rho_{fih}, \theta_f, \lambda_{h-k}, \lambda_{k-h}] \). Note that the transmission prices \( W_i \) are variables in the LCP, not fixed parameters—even though each generator’s optimization problem \((f1)–(f4)\) and the grid’s problem \((G1)–(G3)\) naively presume they are fixed. The LCP algorithm solves for the \( W_i \) that clear the market for transmission services.

Finally, note that the number of conditions (complementarity conditions and equality constraints) equals the number of variables. This can be seen by considering the fact that one set of KKT conditions results for each of the variables \( x_{fih}, y_i, W_i, \rho_{fih}, \theta_f, \lambda_{h-k}, \lambda_{k-h} \). While there is one market clearing equation for each \( W_i \), this “squareness” condition is needed for mixed LCP algorithms to find a solution. In contrast, the NCP in [11] has more variables than conditions, implying existence of multiple equilibria.

There are two general advantages to phrasing the problem as a mixed LCP. First, using theoretical results [31], it is possible to determine if this system of equations satisfies certain sufficient conditions for existence and uniqueness of the solution. Second, efficient mixed LCP solvers, such as those in GAMS, can be used.

Hashimoto [29] points out that a Cournot equilibrium on a transportation network can be calculated by solving a single QP under two conditions:

- supply and demand functions are linear; and
- transportation costs are proportional to flows on the network, and there are no flow limits.

Below, an analogous QP is formulated in which the network is instead governed by both of Kirchhoff’s laws and there are line flow limits. The KKT conditions for this QP are precisely the same as \([f1]–[f4], (G1)–(G3), (MC1)\). Consequently, this QP can be used to derive the no-arbitrage equilibrium. Further, as long as the feasible region is nonempty, a solution will exist; moreover, it can be shown that the concavity of the objective function implies that the resulting market equilibrium prices and profits for each firm are unique [32]. However, the plant outputs \( x_{fih} \) might not be unique. For example, two plants at the same bus \( i \) owned by the same firm \( f \) might have identical marginal costs, and so there may be alternative dispatches that yield the same cost and outputs for \( f \).

The QP is as follows. Choose \( s_{ij}, x_{fih}, \) and \( y_i \) to solve:

\[
\text{MAX } \sum_j [P_{jo^2} s_{ij} - P_{jo} (2Q_{jo}) (\sum_j s_{ij})]^2 \quad - P_{jo} (2Q_{jo}) (\sum_j s_{ij}^2) = \sum_i h C_{fih} x_{fih} \quad \forall f, i, h
\]

s.t.

\[
x_{fih} \geq X_{fih} \quad (\rho_{fih}) \quad \forall f, i, h
\]

\[
\sum_i h x_{fih} = \sum_i h x_{fih} \quad (\theta_f) \quad \forall f
\]

\[
-\sum_i PTDF_{ik} y_i \leq T_{h-k} - \lambda_{h-k} \geq 0; \quad \lambda_{k-h} (\sum_k PTDF_{ik} y_k + \lambda_{k-h}) = 0 \quad (G2)
\]

\[
\sum_i PTDF_{ik} y_i \leq T_{k-h}; \quad \lambda_{k-h} \geq 0; \quad \lambda_{h-k} (\sum_k PTDF_{ik} y_k - T_{h-k}) = 0. \quad (G3)
\]

Note that the dual variables for the market clearing constraint are the transmission fees \( W_i \). This QP can be solved by standard nonlinear optimizers or specialized QP codes [18].

There is one important difference between the QPs objective function and the social welfare objective used in perfect competition models. The difference is the addition of the term \( P_{jo} (2Q_{jo}) (\sum_j s_{ij}^2) \). Adding this term ensures that KKT conditions correctly calculate marginal revenue for \( f \), accounting for how expansion of sales would depress prices.

Two other observations are worth making about the no-arbitrage model. First, the difficulty in the model of [12] with low price periods can be corrected if the above QP objective is substituted for their linearized social welfare objective, and their equilibrium constraint \( (\text{MR} \geq \text{MC}) \) is dropped. Second, the no-arbitrage model yields a value of sales of each firm at each node. That makes it possible to calculate equivalent Herfindahl indices at each \( i \).

B. Arbitrage/POOLCO Model

In this model, generators recognize that marketers/arbitragers will buy and resell power where price differences exceed the cost of transmission. It is assumed here that there are many arbitragers and that they behave competitively, so it is unnecessary to develop individual models for each arbitrager, unlike [14]. Instead, the equilibrium condition that price differences reflect transport cost differences can be directly imposed upon the other
players, as reflected in the following models for the generators and grid owner.

1) Producers: The producer’s model can be formulated as a generalization of the no-arbitrage case, incorporating variables representing arbitrage transactions along with two additional set of constraints. This generalization is implemented as follows:

i) Substitute \( \Sigma g a_g j + a_j \) for \( \Sigma g s g j \) as the quantity consumed in the inverse demand function at each node \( j \), and

ii) Add two arbitrage constraints: \( p_j = p_H + W_i \) for all nodes \( j \) (where \( p_H \) is the price at the hub bus, which is treated as a decision variable in \( f \)’s model) and \( \Sigma a_j = 0 \). The first constraint forces differences in node prices to equal transmission cost, while the second compels arbitragers to be neither net producers nor consumers.

The resulting producer model is as follows:

\[
\text{MAX } \Sigma j [(P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j)) - W_j] - \Sigma_i, h (C_{fih} - W_i)x_{fih} \\
\text{s.t.:} \]

\[
\begin{align*}
x_{fih} & \leq X_{fih} \\
\Sigma_i, h x_{fih} & = \Sigma_j s_{fj} \\
P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j) & = p_H + W_j \\
\Sigma_j a_j & = 0 \\
\forall s_{fj}, x_{fih} & \geq 0.
\end{align*}
\]

Note, once again, that the \( s_{gj} \) for \( g \neq f \) are not decision variables in this model. The primal decision variables are \( f \)’s generation and sales, the arbitrage transactions, and the hub price. Note too that \( a_j \) and \( p_H \) are indexed by firm \( f \); this indicates that the firm views them as being affected by its actions. However, in a market equilibrium, these variables will be equal for all firms. New dual variables \( \alpha_{fj} \) and \( \beta_{fj} \) are introduced for the two arbitrage constraints.

Because of the no-discrimination constraint \( p_j = p_H + W_j \); this model is equivalent to a POOLCO model in which each generator sells power to the grid at the prevailing price at its bus, and not to individual customers at other nodes. That is, in the presence of arbitrage, Cournot competition in a bilateral market yields the same equilibrium as Cournot competition among generators in a POOLCO.

This assertion can be proven by starting with any optimal solution to the above problem. Say that sales \( s_{fj} \) in such a solution are being made to a node in an amount different from \( f \)’s generation at that node \( \Sigma h x_{fjh} \). By increasing \( a_j \) by an amount \( s_{fj} - \Sigma h x_{fjh} \), and decreasing \( s_{fj} \) by the same amount, and si-
multaneously decreasing \( a_j \) and increasing \( s_{fj} \) by that same amount for some other node \( i \), sales and generation can be made equal at \( j \). Such a change would not affect the firm’s revenues net of its transmission costs because:

- the price received for sales at \( j \) net of \( W_j \) is the same as the price at \( i \) net of \( W_i \)—i.e., both equal \( p_H \); and
- the firm’s total sales are unchanged.

Further, such a change will not alter the firm’s generation costs, because the \( x_{fih} \) variables are unaffected. Thus, profit is the same. Finally, the new solution is feasible because generation and delivered prices are unchanged, and the arbitrage energy balance remains satisfied. Therefore, this new solution must also be optimal. Because \( \Sigma i, h x_{fih} = \Sigma j s_{fj} \), it must be possible to arrange a set of such adjustments such that \( \Sigma j x_{fj} = s_{fj}, \forall j \). This solution is equivalent to each generator selling its output at its bus at the bus’s prevailing price.

This result means that generator \( f \)’s model can be simplified by eliminating the \( s_{fj} \) variables as follows:

\[
\text{MAX } \Sigma j [(P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j)) - W_j] \\
\text{s.t.:} \]

\[
\begin{align*}
x_{fih} & \leq X_{fih} \\
\Sigma_i, h x_{fih} & = \Sigma_j s_{fj} \\
P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j) & = p_H + W_j \\
\Sigma a_j & = 0 \\
\forall s_{fj}, x_{fih} & \geq 0.
\end{align*}
\]

The KKT conditions for this model are as follows:

- For \( x_{fih}, \forall i, h, j \):

  \[
  x_{fih} \leq X_{fih}, \quad (\rho_{fih}) \quad \forall i, h \\
  \Sigma_i, h x_{fih} = \Sigma_j s_{fj}, \quad (\theta_f) \\
  P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j) = p_H + W_j, \quad (\alpha_{fj}) \quad \forall j \\
  \Sigma_j a_j = 0, \quad (\beta_f) \\
  x_{fih} \geq 0, \quad \forall i, h.
  \]

Note, once again, that the \( s_{gj} \) for \( g \neq f \) are not decision variables in this model. The primal decision variables are \( f \)’s generation and sales, the arbitrage transactions, and the hub price. Note too that \( a_j \) and \( p_H \) are indexed by firm \( f \); this indicates that the firm views them as being affected by its actions. However, in a market equilibrium, these variables will be equal for all firms. New dual variables \( \alpha_{fj} \) and \( \beta_{fj} \) are introduced for the two arbitrage constraints.

Because of the no-discrimination constraint \( p_j = p_H + W_j \); this model is equivalent to a POOLCO model in which each generator sells power to the grid at the prevailing price at its bus, and not to individual customers at other nodes. That is, in the presence of arbitrage, Cournot competition in a bilateral market yields the same equilibrium as Cournot competition among generators in a POOLCO.

This assertion can be proven by starting with any optimal solution to the above problem. Say that sales \( s_{fj} \) in such a solution are being made to a node in an amount different from \( f \)’s generation at that node \( \Sigma h x_{fjh} \). By increasing \( a_j \) by an amount \( s_{fj} - \Sigma h x_{fjh} \), and decreasing \( s_{fj} \) by the same amount, and si-
multaneously decreasing \( a_j \) and increasing \( s_{fj} \) by that same amount for some other node \( i \), sales and generation can be made equal at \( j \). Such a change would not affect the firm’s revenues net of its transmission costs because:

- the price received for sales at \( j \) net of \( W_j \) is the same as the price at \( i \) net of \( W_i \)—i.e., both equal \( p_H \); and
- the firm’s total sales are unchanged.

Further, such a change will not alter the firm’s generation costs, because the \( x_{fih} \) variables are unaffected. Thus, profit is the same. Finally, the new solution is feasible because generation and delivered prices are unchanged, and the arbitrage energy balance remains satisfied. Therefore, this new solution must also be optimal. Because \( \Sigma i, h x_{fih} = \Sigma j s_{fj} \), it must be possible to arrange a set of such adjustments such that \( \Sigma j x_{fj} = s_{fj}, \forall j \). This solution is equivalent to each generator selling its output at its bus at the bus’s prevailing price.

This result means that generator \( f \)’s model can be simplified by eliminating the \( s_{fj} \) variables as follows:

\[
\text{MAX } \Sigma j [(P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j)) - W_j] \\
\text{s.t.:} \]

\[
\begin{align*}
x_{fih} & \leq X_{fih} \\
\Sigma_i, h x_{fih} & = \Sigma_j s_{fj} \\
P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j) & = p_H + W_j \\
\Sigma a_j & = 0 \\
\forall s_{fj}, x_{fih} & \geq 0.
\end{align*}
\]

The KKT conditions for this model are as follows:

- For \( x_{fih}, \forall i, h, j \):

  \[
  x_{fih} \leq X_{fih}, \quad (\rho_{fih}) \quad \forall i, h \\
  \Sigma_i, h x_{fih} = \Sigma_j s_{fj}, \quad (\theta_f) \\
  P_{j0} - (P_{j0}/Q_{j0})(\Sigma g s g j + a_j) = p_H + W_j, \quad (\alpha_{fj}) \quad \forall j \\
  \Sigma a_j = 0, \quad (\beta_f) \\
  x_{fih} \geq 0, \quad \forall i, h.
  \]

Two implicit assumptions of the above model are that all nodes (except the hub) have a demand curve, and that the equilibrium price never exceeds its choke price \( P_0 \) at any \( i \). More general assumptions can also be handled by a LCP.

2) Grid Owner: The grid model corresponding to the full producer model is identical to the no-arbitrage case.

3) Market Clearing: For an arbitrage model solution to represent an equilibrium, the following balances must be maintained between the transmission services provided by the grid and the services demanded by the arbitragers:

\[
y_i = a_{fj}, \quad \forall f, i. \quad (MC1')
\]
Furthermore, the hub prices assumed by the firms must be consistent, resulting in $F - 1$ of the following equations, where $F$ is number of firms:

$$P_{Hf} = P_{H1} \quad \forall f > 1.$$  

(MC2')

4) Solution Approach: The arbitrated market equilibrium problem can be solved by gathering the producers' KKT conditions ($f^1_f - f^0_f$, $\forall f$); the grid owner's KKT conditions (G1)–(G3); and the market clearing conditions (MC1'), (MC2').

The resulting mixed LCP can be solved for the equilibrium values of $\dot{q}$.

But this system cannot be solved directly, since there are more conditions than variables. This occurs because there are more market clearing conditions (MC1'), (MC2') than transmission prices $W_f$. There are $IF + F - 1$ such conditions, but only $I$ prices, where $I$ is the number of nodes other than the hub. Thus, there are $IF + F - 1 - I$ equations too many. However, it turns out that precisely this number of conditions are redundant. In particular:

- $(F - 1)I$ equations ($f^3_f$) can be dropped [because those that correspond to $f > 1$ are equivalent to those for $f = 1$, given (MC1'), (MC2')] and
- $F - 1$ equations ($f^0_f$) can be deleted [as (MC1') implies that the $(f^0_f)$ for $f > 1$ are redundant to $(f^0_f)$ for $f = 1$].

Omission of those equations yields the required "square" MCP system, with the number of conditions equaling the number of variables.

By formulating an equivalent $QP$, Metzler [32] shows that this MCP has a solution, and that the prices and profits resulting from that solution are unique. She also formulates alternative models that yield equivalent solutions, including one in which arbitrage constraints and variables are removed from each producer $f$'s model. There, she defines a separate set of KKT conditions which derives from a single arbitrager's profit maximization problem. These conditions force uneconomic price differences among locations to disappear in equilibrium.

V. EXAMPLE

This simple example illustrates the application of the above models, and is designed to permit verification by the reader. There are three busses, $i = 1, 2, 3$, each of which has customers. However, generation occurs only at busses $i = 1, 2$. Each pair of busses is interconnected by a single transmission line; all three lines have equal impedances. The demand functions are $s^1$, $s^2$, and $s^3$/MWh. These functions imply that demand is more elastic at the demand-only node (bus 3). There are two producers, each with one generator. Firm 1's generator is sited at $i = 1$, while 2's is at $i = 2$. Both generators have unlimited capacity, and a constant marginal cost: $15$/MWh for firm 1, and $20$/MWh for firm 2. The only transmission cost arises from congestion.

Two different transmission systems are considered below: one without congestion—infinite transmission capacity—and one with congestion on a single interface between busses 1 and 2 ($\dot{q} = 40 - 0.08\dot{q}_1$, $i = 1, 2$ and $\dot{q}_2 = 32 - 0.05\dot{q}_1$/MWh). These functions imply that demand is more elastic at the demand-only node (bus 3). There are two producers $f = 1, 2$, each with one generator. Firm 1's generator is sited at $i = 1$, while 2's is at $i = 2$. Both generators have unlimited capacity, and a constant marginal cost: $15$/MWh for firm 1, and $20$/MWh for firm 2. The only transmission cost arises from congestion.

Two different transmission systems are considered below: one without congestion—infinite transmission capacity—and one with congestion on a single interface ($k = 1$) between busses 1 and 2 ($\dot{q}_{1+} = \dot{q}_{1-} = 25$ MW). For each case, Tables I and II show the results for each of three types of competition:

### TABLE I

<table>
<thead>
<tr>
<th>Case</th>
<th>Labor, Prices, and Profits, Three Bus Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delivered Prices, $/MWh</td>
</tr>
<tr>
<td>$T_2 = \infty$ (No limit): Perfect Competition</td>
<td>$p_1, p_2, p_3$</td>
</tr>
<tr>
<td>Cournot, No Arbitrage</td>
<td>25</td>
</tr>
<tr>
<td>Cournot, Arbitrage</td>
<td>23.8</td>
</tr>
<tr>
<td>$T_2 = 25$ MW: Perfect Competition</td>
<td>15</td>
</tr>
<tr>
<td>Cournot, No Arbitrage</td>
<td>24.1</td>
</tr>
<tr>
<td>Cournot, Arbitrage</td>
<td>22.5</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Case</th>
<th>Generation, Sales, and Transmission for Three Bus Example (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quantities demanded</td>
</tr>
<tr>
<td></td>
<td>$q_1, q_2, q_3$</td>
</tr>
<tr>
<td>$T_2 = \infty$ (No limit): Perfect Competition</td>
<td>312.5</td>
</tr>
<tr>
<td>Cournot, No Arbitrage</td>
<td>187.5</td>
</tr>
<tr>
<td>Cournot, Arbitrage</td>
<td>202</td>
</tr>
<tr>
<td>$T_2 = 25$ MW: Perfect Competition</td>
<td>312.5</td>
</tr>
<tr>
<td>Cournot, No Arbitrage</td>
<td>199.1</td>
</tr>
<tr>
<td>Cournot, Arbitrage</td>
<td>218.2</td>
</tr>
</tbody>
</table>
perfect competition, Cournot competition/no-arbitrage, and Cournot competition/arbitrage. Perfect competition is simulated by a model that maximizes social welfare, defined as the sum of generator profit, grid owner profit, and consumer surplus. This solution is equivalent to marginal cost pricing by generators.

An obvious difference among the solutions is the effect of imperfect competition. Perfect competition yields much lower prices and higher welfare than the Cournot solutions. Without transmission constraints, the cheapest firm \( f = 1 \) serves the entire demand at all nodes under perfect competition. But under imperfect competition, prices climb enough to allow the more costly firm \( f = 2 \) to enter. When the 25 MW interface constraint is imposed, firm 2 generates even under perfect competition, as firm 1 cannot ship enough power to meet all of \( i = 2 \) and \( 3 \)'s demand.

Another difference between the solutions is the impact of arbitrage. Without arbitrage, Cournot prices can differ significantly among nodes even in the absence of transmission constraints. Further, these price differences may bear little relationship to the costs of transmission (as reflected in the \( W_i \)). In particular, because node 3's demand is relatively elastic, its no-arbitrage prices are lower than the other nodes', even though power flows to that node from nodes 1 and 2. Thus, power seems to flow the wrong way, from high priced nodes to node 3's. This solution is lower, rather than higher, than that at 1. This is because the node 3’s elasticity motivates generators to cut delivered prices there, and there are no marketers to arbitrage away the ensuing noncost based price difference.

The ability to arbitrage alters those price differences. Arbitrage aligns bus prices \( p_i \) with transmission costs \( W_i \); thus, when there are no transmission limits, price differences are completely eliminated. Note also that the arbitrage solutions yield higher welfare than the no-arbitrage cases because arbitrage eliminates unjustified price differences. This, however, is not necessarily a general result [33].

In [34], results are shown for a two node system, including a counter-intuitive outcome that reduced transmission capacity increases welfare in some cases. This occurs there because imperfect competition reverses the direction of flows relative to perfect competition, and tighter limits lessen these inefficient flows.

VI. CONCLUSION

Nash–Cournot models are popular although not necessarily realistic methods for modeling strategic interactions in power markets. However, previously proposed Nash–Cournot models either ignore the grid or represent it as a simple radial network, or they pose computational difficulties for large networks, such as nonexistence or nonuniqueness of equilibria. The models proposed here are able to compute imperfectly competitive equilibria for networks including hundreds or even thousands of control areas or buses and similarly large numbers of interfaces. An immediate task is the application to such systems in the context of merger evaluation and other market power studies. We have obtained initial results for the UK and Eastern Interconnection [35], [36].

Another question that should be addressed is: how do the model results compare to solutions from other proposed market models? Examples include Cournot models in which producers correctly anticipate how changing output affects congestion [5], [6], [11] or the Cournot and supply function equilibria models surveyed in [2], [3]. Such models may yield more realistic equilibria if the behavior they represent is more representative of how producers behave. However, most of those models are more difficult to compute than the models presented here and have existence and uniqueness problems [4], [15].

Finally, versions of the models including other transmission pricing systems, including zonal and FERC Order 888-type pricing, should be formulated and implemented. As Harvey and Hogan suggest [37], alternative transmission pricing schemes could have important implications for the exercise and effects of market power.

ACKNOWLEDGMENT

U. Helman, C. Day, W. Stewart, R. O’Neill, M. Rothkopf, Y. Smeers and the reviewers provided helpful comments. Special thanks are due to C. Metzler and J.-S. Pang whose encouragement, criticisms and proofs have been particularly valuable. Opinions and errors are the responsibility of the author.

REFERENCES


[29] B. F. Hobbs is Professor of Geography and Environmental Engineering and of Mathematical Sciences (Joint) at The Johns Hopkins University, and an Adjunct Professor of Electrical Engineering and Computer Science at Case Western Reserve University. He consults with FERC and the USDOE Energy Information Agency. He earned a Ph.D. in environmental systems engineering from Cornell University in 1983. He was with Brookhaven National Laboratory from 1977–1979 and Oak Ridge National Laboratory 1982–1984. From 1984 to 1995, he was on the Faculty of the Departments of Systems Engineering and Civil Engineering at CWRU. Among his research interests are simulation of deregulated power markets and ecosystem management under risk and multiple objectives.