Addressing supply-side risk in uncertain power markets: stochastic Nash models, scalable algorithms and error analysis

Aswin Kannan [a] , Uday V. Shanbhag [b] & Harrison M. Kim [b]

[a] Argonne National Laboratory, Mathematics and Computer Science
Division, Urbana, IL, 61801, USA

[b] Department of Industrial and Enterprise Systems Engineering,
University of Illinois at Urbana-Champaign, Urbana, IL, USA

Available online: 17 May 2012

PLEASE SCROLL DOWN FOR ARTICLE

Addressing supply-side risk in uncertain power markets: stochastic Nash models, scalable algorithms and error analysis

Aswin Kannan[a], Uday V. Shanbhag[b]*, and Harrison M. Kim[b]

*[a]Argonne National Laboratory, Mathematics and Computer Science Division, Urbana, IL 61801, USA;
[b]Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

Increasing penetration of volatile wind-based generation into the fuel mix is leading to growing supply-side volatility. As a consequence, the reliability of the power grid continues to be a source of much concern, particularly since the impact of supply-side risk exposure, arising from aggressive bidding,[1] is not felt by risk-seeking generation firms; instead, the system operator is largely responsible for managing shortfalls in the real-time market. We propose an alternate design in which the cost of such risk is transferred to firms responsible for imposing such risk. The resulting strategic problem can be cast as a two-period generalized stochastic Nash game with shared strategy sets. A subset of equilibria is given by a solution to a related stochastic variational inequality, that is shown to be both monotone and solvable. Computing solutions of this variational problem is challenging since the size of the problem grows with the cardinality of the sample space, network size and the number of participating firms. Consequently, direct schemes are inadvisable for most practical problems. Instead, we present a distributed regularized primal–dual scheme and a dual projection scheme where both primal and dual iterates are computed separately. Rates of convergence estimates are provided and error bounds are developed for inexact extensions of the dual scheme. Unlike projection schemes for deterministic problems, here the projection step requires the solution of a possibly massive stochastic programme. By utilizing cutting plane methods, we ensure that the complexity of the projection scheme scales slowly with the size of the sample space. We conclude with a study of a 53-node electricity network that allows for deriving insights regarding market design and operation, particularly for accommodating firms with uncertain generation assets.

Keywords: variational inequalities; stochastic programming; Nash games; projected gradient schemes; cutting plane methods

1. Introduction

As electricity markets gravitate towards regimes where intermittent renewables such as wind power are an integral part of a firm's generation mix, the associated reliability of power markets in the real-time setting assumes increasing relevance. Currently, most markets operate a day-ahead and a real-time market. Of these, the day-ahead market is a financially binding market that provides generators with an operating schedule prior to the real-time settlement. The real-time

market provides an economically optimal schedule that accords with changes in load, generation and system conditions after the settlement of the day-ahead market.

A majority of market designs employs financial penalties, referred to as deviation penalties, in an effort to reduce deviations from day-ahead bids. These penalties, in part, are expected to aid in covering the cost of real-time shortfall when generators do indeed fail to generate at contracted levels. But, generators are currently *not* charged for imposing risk. In fact, regardless of whether such deviations occur, the system operator often has to hold reserves in an effort to maintain grid reliability. But, the cost of holding reserves remains unpriced by the market and the independent system operator (ISO) is responsible for holding reserves. As wind power makes increasingly deep inroads, the bidding behaviour of wind-based generators may impose significant risk onto the market. Importantly, this risk will continue to be borne by the system operator.

In this paper, we introduce a risk-based framework in which firms are penalized in proportion with their risk exposure; in effect, firms are rendered *risk-averse*. We see these penalties as contributing towards the cost of holding reserves by the ISO. In fact, these penalties are not driven by events in the real-time market but are derived entirely on the basis of the generator bids and the associated distributions of their generation assets in the real-time market.

In an era of growing incentives for renewables, one may naturally question the role of introducing risk-based penalties. Our research is motivated by the observation that few policy instruments exist for addressing reliability. In fact, such concerns are exacerbated when wind-power firms make (risky) bids that are defined by a high likelihood of shortfall in the real-time market. In current designs, the onus of this risk is borne by the system operator while we present a framework in which such firms are charged a risk premium, in accordance with a chosen risk measure. Importantly, such premia are not uniformly imposed on all wind-based bids; specifically, if bids that do not introduce undue risk, then little or no penalty is imposed. In short, we believe that such risk measures provide an additional policy instrument that copes with aggressive bidding.

1.1 *A motivating example*

Consider the bidding by a wind-based generator in the day-ahead market. When firms change their real-time generation levels from their day-ahead schedule, a deviation cost is imposed. This cost takes on a variety of forms and is often a convex increasing function of the deviation. This charge contributes to the cost faced by the ISO in addressing shortfalls that may emerge in the real-time.[2] Figure 1 shows the deviation costs arising from low and high bids in the day-ahead market. If the actual output of the generator in the real-time market is at least as much as the day-ahead bid, then no penalty is imposed.
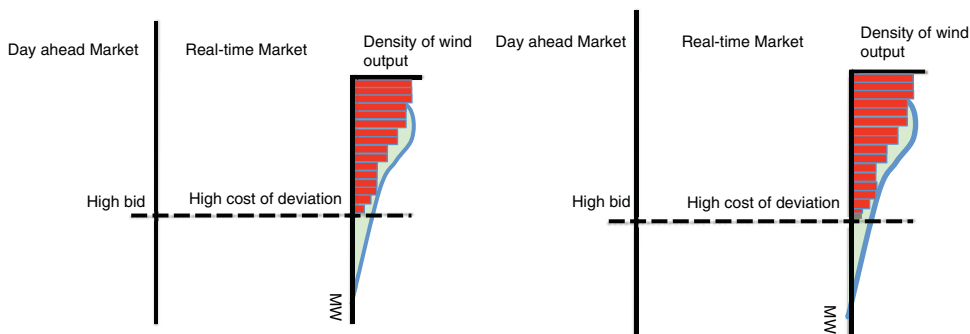

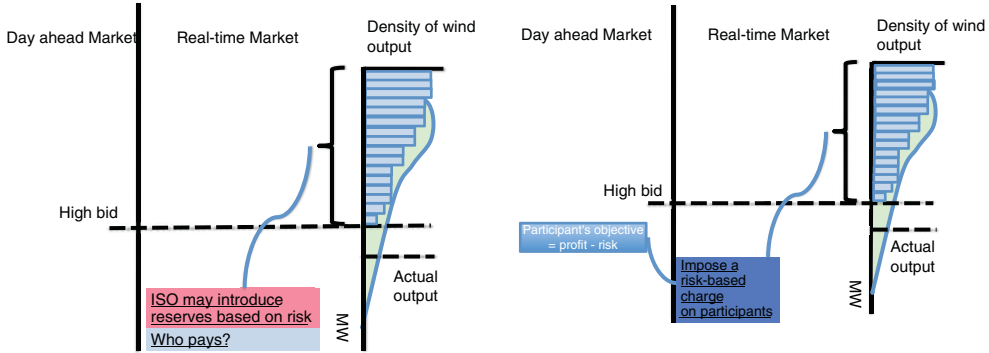Figure 1. Deviation costs associated with low and high day-ahead bids.

Figure 2.   Risk exposure and risk-based penalties.

However, as Figure 2 shows, a high bid may impose a significant risk and the ISO has to buy reserves to ensure that grid reliability is maintained. If in the real-time, the generator is dispatched at the day-ahead level, then the ISO is not remunerated for this risk exposure. An alternative lies in imposing a penalty on the generators based on this exposure, as seen in Figure 2. In fact, this penalty may be computed in an ex ante fashion allowing the system operator to determine the risk exposure ahead of clearing the real-time market.

1.2   *Outline and contributions*

Crucial to answering such questions is the development of a new generation of game-theoretic models that can contend with the uncertainty and risk, in the context of sequential electricity markets. In the past, deterministic variants have proved to be useful in analysing a range of questions in the design and operation of markets, both in a single-settlement framework [4,29,30,45] and in a two-settlement framework [9,31,32,63].[3] Yet, past work provides little from the standpoint of characterizing and computing equilibria, particularly in settings complicated by risk and uncertainty. The current paper is fuelled by natural questions arising from the resulting two-period risk-averse stochastic Nash games: (a) *Characterization of equilibria:* Can one *characterize* equilibria in such games; and (b) *Scalable computational schemes:* can such equilibria be computed via *efficient scalable and convergent* schemes?

This range of questions falls at the interstices of stochastic programming and continuous-strategy Nash games. Of these, the former is a subclass of mathematical programming first discussed by Dantzig [15] and Beale [5] and allows for both adaptive [7,33,54,60] (such as models allowing for recourse actions in the second-period, contingent on first-period decisions) and anticipative models [7,11,55] (such as chance-constrained models that impose a probabilistic or reliability constraint on the underlying optimization model), amongst others. Game theory [22,49] has its roots in the work by Von Neumann and Morgenstern [47] while the Nash-equilibrium solution concept was forwarded by Nash in 1950 [46].

Our focus is on *N*-person risk-averse stochastic Nash games over continuous strategy sets and is inspired by settings where agents make simultaneous bids in the first (such as a forward market) period followed by *recourse* bids in the second (such as a real-time market) period. We use a conditional value-at-risk (CVaR) measure [51] to capture the risk associated with bidding with assets whose availability is uncertain in the real-time market. The class of games under consideration depart from canonical models in at least two ways: first, the strategy sets of the players are coupled, implying that the agents are competing in a generalized Nash game [17,25]; second, each player solves a two-period recourse-based risk-averse optimization problem. Several challenges

are encountered in addressing the characterization and computational questions fuelling this paper. When considering the characterization of equilibria, in the realm of continuous strategy games, a common avenue relies on the analysis of the sufficient equilibrium conditions, namely a variational inequality or a complementarity problem, arising from the game. Unfortunately, in the current setting, this approach is fraught with several difficulties. First, the strategy sets across agents are coupled when one works within a regime of a networked electricity market, implying that the equilibrium conditions lead to a quasi-variational inequality [10,50], generally a less tractable object. Second, given that risk-averse agents employ CVaR measures, the resulting objectives are possibly non-smooth and the resulting variational inequality has a multivalued mapping. Third, in general, neither are the mappings of the resulting variational inequalities strongly monotone nor are the strategy sets compact. In short, a direct conclusion regarding existence and uniqueness of equilibria is unavailable.

When considering the development of scalable computational schemes, the solution of the resulting complementarity problems in practical settings is constrained by several issues. While a direct application of a solver such as PATH [21] or KNITRO [8] is often the best choice for solving such problems, it is unlikely that the computational effort will scale well with growth in problem size. Consequently, the solution of truly large-scale instances via direct schemes becomes increasingly difficult and suggests the construction of distributed schemes. Motivated by these challenges, the present work makes the following contributions.

1.2.1 *Analysis of equilibria*

When agents are faced by a forward market and a subsequent uncertain real-time market, we employ a two-period stochastic Nash model. Since, the strategy sets are coupled across firms, a generalized variant of the Nash solution concept is employed. In general, the presence of a CVaR measure implies that the agent objectives are non-smooth. By a suitable reformulation, an equilibrium to the game is given by a single-valued variational inequality [20]. We show that even in the absence of a standard risk-neutrality assumption, the resulting variational inequalities are monotone and admit compact non-empty solution sets.

1.2.2 *Convergent scalable schemes with error bounds*

We present two distributed projection-based cutting-plane schemes for computing equilibria, the first a single timescale (primal–dual) method while the second is a two timescale (dual) scheme. For the dual scheme, we develop estimates of the convergence rate and extend the analysis to contend with practical implementations. In particular, we analyse the error associated with bounded complexity implementations where the underlying primal scheme is run for a finite number of steps. The ability of the scheme to contend with the size arising from the uncertainty rests on being able to solve the projection problems effectively. By observing that these problems are two-period stochastic convex programs with complete recourse, we employ a cutting-plane method whose effort grows linearly with the cardinality of the sample space. Numerical results suggest that the overall scheme scales well with problem size.

1.2.3 *Insights for market design and operation*

A numerical implementation on a 53-node model of the Belgian network provides numerous insights for market design. For instance, we observe that higher levels of risk-aversion lead to lower participation in forward markets by firms with wind generation assets while higher levels of wind penetration results in greater participation in forward markets.

The paper is organized into five sections. In Section 2, we introduce the stochastic two-settlement electricity market model and define the related shared-constraint games and associated variational conditions. In Section 3, we analyse the properties of equilibria arising in such games. A novel hybrid distributed scheme that combines projection methods with cutting-plane algorithms is presented in Section 4. In Section 5, we obtain insights through a two-settlement networked electricity market model via a risk-based generalized Nash game. The paper concludes with some brief remarks in Section 6.

2. A two-settlement electricity market model

While extant research has laid the foundation for drawing insights pertaining to agent behaviour in power markets [29,30,34,63], these models and the consequent solution concepts are inadequate at least from two standpoints: (1) First, the majority of the past effort has presented a largely deterministic viewpoint, barring [57,63], and has ignored the native uncertainty in real-time markets arising in generation costs and availability; (2) Second, much of the past research on bidding in two-period markets assumes fully rational agents and leads to highly intractable problems, in general. For instance, in [63], firms participating in the forward market compete subject to equilibrium in the spot market. This in itself is not a shortcoming, but the resulting agent problems are given by mathematical programmes with equilibrium constraints (MPECs) [43], a class of ill-posed non-convex nonlinear programs. Little existence theory exists for the resulting games, called multi-leader multi-follower games [41], barring results in either conjectured settings [57] or under rather strong assumptions [2,59]. Further, even when equilibria are known to exist, there are no known convergent algorithms for computing these equilibria. Both shortcomings become even more pronounced when one considers the introduction of risk measures.

The present work is principally motivated by analysing a class of game-theoretic models that can overcome some of the shortcomings described in (1) and (2). We address (1) through a stochastic game-theoretic framework in which agents have heterogeneous risk preferences and employ a conditional VaR metric to capture the risk of capacity shortfall. This can be viewed as an adapted open-loop game, first studied in an optimization setting by Haurie and Moresino [26] and Haurie and Zaccour [27] for modelling multistage decision-making problems. This avenue alleviates some of the challenges articulated in (2), namely from the standpoint of characterizing and computing equilibria. In particular, we consider a simpler question of agents making simultaneous bids in the forward market and the recourse-based bids in the spot-market. Two interpretations of the resulting game can be given: (i) *Economics:* it can be viewed as a bounded-rationality simplification of the fully rational game in which firms compete in Nash with respect to the ISO, rather than assuming a leadership role, a model studied by Hobbs [29], amongst others; (ii) *Mathematical programming:* it can also be viewed as a Nash game played at the forward market by agents solving two-period stochastic programs. In particular, agents play a game in the first period and for every scenario in the second period, where recourse decisions may be taken.

Using the model suggested by Yao *et al.* [63] as a basis, we now describe some features of our framework that are common to the models we introduce in the sections to follow. The notation of the model is summarized in Table A1. Suppose the uncertainty in the second-stage is captured by the random vector ξ, and $\xi : \Omega \longrightarrow \mathbb{R}^{\tilde{n}}$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is finite in cardinality. Throughout this paper, we qualify an uncertain parameter or decision by using a subscript ω. Consider a market in which g firms compete in an electricity network where inflow/outflow decisions are managed by the ISO. Let $\mathcal{N}$ and $\mathcal{N}_j$, respectively, denote the set of nodes in this network and the set at which firm j owns generation facilities where $j = 1, \ldots, g$. Two-settlement markets are constructed around a sequence of clearings, in which the first settlement specifies the forward price while the second is a consequence of physical

transactions and determines the real-time price. We denote by x_{ij} the forward position at node i corresponding to firm j while the corresponding physical generation in scenario ω is denoted by y_{ij}^{ω}. Further, the forward and real-time prices (in scenario ω) at node i are denoted by p_i^0 and p_i^{ω}, respectively. The ISO manages injections and outflows at all nodes, where the inflow at the ith node under scenario ω is denoted by r_i^{ω} where $i \in \mathcal{N}$. Note that a positive (negative) value of r_i^{ω} marks an inflow (outflow).

2.1 *Pricing and risk mechanisms*

Given positive scalars $(a_i^{\omega}, b_i^{\omega})$, we define the nodal spot prices at scenario ω as an affine function of nodal consumption at that node, given by the total generation by all firms at node i modified by the ISO's injection, denoted by r_i^{ω}.

$$p_i^{\omega}(\bar{y}_i^{\omega}, r_i^{\omega}) \triangleq a_i^{\omega} - b_i^{\omega}(\bar{y}_i^{\omega} + r_i^{\omega}) \quad \forall i \in \mathcal{N}, \tag{1}$$

where $\bar{y}_i^{\omega} = \sum_{j \in \mathcal{J}} y_{ij}^{\omega}$. Traditional models have imposed an arbitrage-free assumption that required that the nodal forward prices were given by nodal expected spot prices. In practice, forward prices are a consequence of a market clearing and need not necessarily match expected spot prices, as discussed by Kamat and Oren [34]. Alternate models of forward pricing [34] point to the non-storability of electricity as being one reason for why the no-arbitrage condition may not hold. In accordance with Kamat and Oren [34], in one of our models, we employ a Cournot-based price function in the forward market. Specifically, p_i^0 the forward price at node i is given by

$$p_i^0(\bar{x}_i) \triangleq a_i^0 - b_i^0(\bar{x}_i), \tag{2}$$

where $\bar{x}_i = \sum_{j \in \mathcal{J}_i} x_{ij}$ and a_i^0 and b_i^0 are positive scalars for all $i \in \mathcal{N}$. Furthermore, during scenario ω, we denote the cost of generation of firm j at node i by $\zeta_{ij}^{\omega}(y_{ij}^{\omega})$ and the positive and negative deviations from the forward position x_{ij} by u_{ij}^{ω} and v_{ij}^{ω} respectively. The generation in the real-time market by firm j at node i is denoted by y_{ij}^{ω} and is defined by

$$y_{ij}^{\omega} = x_{ij} + u_{ij}^{\omega} - v_{ij}^{\omega}.$$

Shortfall in real-time generation capacity is penalized through a deviation cost, implying that the total cost of negative deviation arising for capacity shortfalls provides an estimate of the reliability of the market. While grid reliability generally has a more expansive definition, the 'likelihood of generation shortfall in the real-time market' is taken as a proxy for market reliability. For instance, if generators make low forward bids, then the likelihood of real-time shortfall is correspondingly lower. Unfortunately, such a measure of reliability is available upon the settlement of the real-time market and is, in effect, an ex post measure. In this subsection, we present a modified model that replaces deviation costs with a risk measure that incorporates the losses associated with a shortfall in real-time generation. Such a modification has several benefits. First, it allows firms to compete with heterogeneous risk preferences where the risk corresponds to the losses associated with capacity shortfall in the real-time market. Second, the risk measure provides an ex ante measure of market reliability.

Current market models discourage deviations from forward positions through the imposition of convex costs on deviations. As a consequence, the firms minimize their expected revenue less their expected cost of generation and deviation. For instance, if $X(\omega; y)$ represents the random loss under realization ω, given forward decision y, then an expected-value approach would consider the metric $\min_{y \in Y} \mathbb{E}[X(\omega; y)]$. However, such a model focuses on the *average* and does not consider the possibility that levels of real-time capacity may result in massive deviation costs. In effect, the expected-value approach does not allow for capturing risk-averseness.

Classical approaches to modelling risk preferences require the use of expected utility theory leading to agents maximizing their expected utility. In particular, if $\mathbf{u} : \mathbb{R} \to \mathbb{R}$ is a concave utility function, then a risk-averse firm would maximize $\mathbb{E}[\mathbf{u}(X(\omega; y))]$. Unfortunately, eliciting the utility functions of the agents remains rather challenging and often arbitrarily selected utility functions lead to solutions that are difficult to interpret. More recently, an approach for addressing risk aversion is through the use of risk measures. Recently, the VaR measure has gained popularity in the financial industry and is defined as

$$\text{VaR}_\alpha(X; y) \triangleq H_X^{-1}(1 - \beta),$$

where $H_X(x; y) \triangleq \mathbb{P}(X(\omega; y) \le x)$. Unfortunately, the VaR measure does not satisfy the properties of coherence [3] or convexity. Finally, the VaR measure ignores losses beyond the $\text{VaR}_\beta(X)$ level and consequently these can be arbitrarily large. The conditional value-at-risk or CVaR measure is coherent, convex and does consider the expectation of the losses beyond the VaR level and is defined as

$$\text{CVaR}_\tau(X(\omega; y)) \triangleq \min_{m \in \mathbb{R}} \left(m + \frac{1}{1 - \tau} \mathbb{E}(X(\omega; y) - m)^+ \right), \tag{3}$$

where $w^+ \triangleq \max(w, 0)$ and $\tau \in (0, 1)$ denotes the confidence level [51]. In the past, CVaR measures, and more generally coherent risk measures, have been employed in the context of risk management in a power setting [14] as well as an inventory control context [1]. An analogous measure for capturing the risk of a shortfall faced by firm j's bid at node j as follows:

$$\text{CVaR}_{\tau_j}(\varrho(\text{cap}_{ij}^\omega; x_{ij})) \triangleq \min_{m_{ij}} \left(m_{ij} + \frac{1}{1 - \tau_j} \mathbb{E}(\varrho_{ij}(\text{cap}_{ij}^\omega; x_{ij}) - m_{ij})^+ \right),$$

where $\varrho(\text{cap}_{ij}^\omega, x_{ij})$ denotes the loss function given a bid x_{ij}, cap_{ij}^ω denotes the capacity of firm i's plant at node j in the second period, and τ_j refers to the confidence level imposed by firm j. An instance of such a loss function is given by

$$\varrho(\text{cap}_{ij}^\omega, x_{ij}) = \chi(x_{ij} - \text{cap}_{ij}^\omega)^+,$$

where $\chi > 0$. Such a loss function is positive if the bid exceeds the random availability. As a consequence, when x_{ij} is set closer to the right tail of the distribution cap_{ij}^ω, then this loss function is positive with a high probability; correspondingly, if x_{ij} is low, then this loss function is zero with a high probability.

For purposes of constructing a compact notation, the strategy of the jth firm is denoted by z_j and is defined as

$$z_j \triangleq \begin{pmatrix} x_{\bullet j} \\ m_{\bullet j} \\ y_{\bullet j}^\bullet \\ u_{\bullet j}^\bullet \\ v_{\bullet j}^\bullet \end{pmatrix}, \quad w_{\bullet j} \triangleq \begin{pmatrix} w_{1j} \\ \vdots \\ w_{Nj} \end{pmatrix} \quad \text{and} \quad w_{\bullet j}^\bullet \triangleq \begin{pmatrix} w_{\bullet j}^1 \\ \vdots \\ w_{\bullet j}^{|\Omega|} \end{pmatrix}.$$

For $j = 1, \ldots, g$, the resulting profit functions of firm j are given by the summation of nodal expected profits less risk of shortfall:

$$\pi_j(z_j; z_{-j}) \triangleq \sum_{i \in \mathcal{N}} \underbrace{p_i^0 x_{ij} + \mathbb{E}(p_{ij}^\omega(y_{ij}^\omega - x_{ij}) - \zeta_{ij}^\omega(y_{ij}^\omega))}_{\text{Mean profit}} - \kappa_j \sum_{i \in \mathcal{N}} \underbrace{\text{CVaR}_{\tau_j}(\varrho(\text{cap}_{ij}^\omega; x_{ij}))}_{\text{Shortfall risk}}, \tag{4}$$

where κ_j represents the risk-aversion parameter of agent j. If $\mathcal{J}_i$ denotes the set of firms that have generation resources at node i, then the jth firm's strategy set is given by $\mathbf{Z}_j \cap \mathcal{D}_j(z_{-j})$ where

$$\mathbf{Z}_j \triangleq \left\{ z_j : \begin{cases} y_{ij}^\omega = x_{ij} + u_{ij}^\omega - v_{ij}^\omega \\ y_{ij}^\omega \leq \mathrm{cap}_{ij}^\omega \\ x_{ij}, u_{ij}^\omega, v_{ij}^\omega, y_{ij}^\omega \geq 0, \end{cases} \forall i \in \mathcal{N}, \ \forall \omega \in \Omega \right\}$$

and

$$\mathcal{D}_j(z_{-j}) \triangleq \left\{ z_j : \left\{ \sum_{j \in \mathcal{J}_i} y_{ij}^\omega + r_i^\omega \geq 0 \right\}, \ \forall i \in \mathcal{N}, \ \forall \omega \in \Omega \right\},$$

respectively.

In the definition of $\mathbf{Z}_j$, the first set of constraints relate real-time generation to the forward positions through the deviation levels while the second set of constraints impose a bound on real-time generation based on available capacity. The set-valued mapping $\mathcal{D}_j(z_{-j})$ is defined by a set of algebraic constraints which specify that the net outflow at every node is non-negative.

Next, we define z_{g+1} which is subsequently employed in constructing the ISO's problem:

$$z_{g+1} \triangleq r_\bullet^\bullet, \quad r_\bullet^\bullet \triangleq \begin{pmatrix} r_1^\bullet \\ \vdots \\ r_N^\bullet \end{pmatrix}, \quad r_i^\bullet \triangleq \begin{pmatrix} r_i^1 \\ \vdots \\ r_i^{|\Omega|} \end{pmatrix}.$$

If $\bar{\mathcal{N}}$ represents the set of nodes in the network less the slack node, then the ISO's strategy set is given by $\mathbf{Z}_{g+1} \cap \mathcal{D}_{g+1}(z_{-(g+1)})$ where

$$\mathbf{Z}_{g+1} \triangleq \left\{ z_{g+1} : \left\{ \begin{array}{c} \displaystyle\sum_{i \in \mathcal{N}} r_i^\omega = 0 \\ \displaystyle\sum_{i \in \bar{\mathcal{N}}} Q_{l,i} r_i^\omega \leq K_l^\omega, \sum_{i \in \bar{\mathcal{N}}} Q_{l,i} r_i^\omega \geq -K_l^\omega, \end{array} \right\} \forall l \in \mathcal{L}, \ \forall \omega \in \Omega \right\}$$

and

$$\mathcal{D}_{g+1}(z_{-(g+1)}) \triangleq \left\{ r_i^\omega : \left\{ \sum_{j \in \mathcal{J}_i} y_{ij}^\omega + r_i^\omega \geq 0 \right\}, \ \forall i \in \mathcal{N}, \ \forall \omega \in \Omega \right\},$$

respectively.

Note that in $\mathbf{Z}_{g+1}$, the first set of constraints are the power balance requirements, while the second and third represent the transmission capacity constraints[4]. The ISO's objective, taken as the social welfare [34,62], is given by the expected spot-market revenue (the area under the price-quantity curve) less generation cost or

$$\pi_{g+1}(z_{g+1}; z_{-(g+1)}) \triangleq \sum_{i \in \mathcal{N}} \mathbb{E}\left(\int_0^{\sum_{j \in \mathcal{J}} y_{ij}^\omega + r_i^\omega} p(\tau) \, d\tau - \sum_{j \in \mathcal{J}} \zeta_{ij}^\omega(y_{ij}^\omega) \right).$$

Note that since the ISO does not directly control real-time generation (denoted by y), the second term may be dropped from the objective.

It is worth remarking how the recourse model relates to the functioning of the real-time market. Standard designs necessitate that every firm provides a single bid in the real-time market, prior to the realization of the uncertainty. The recourse-based approach requires that the probability space is known to all firms and the ISO and every firm provides a set of 'recourse' bids, each of

which corresponds to one realization of uncertainty; In effect, each firm would submit $|\Omega|$ real-time bids. Furthermore, the ISO has no forward decision but does provide a set of recourse-based injections/withdrawals at a nodal level.

2.2 *Generalized Nash games and variational equilibria*

The resulting parameterized optimization problem faced by the jth firm is given by and the associated generalized Nash game is given by the following.

DEFINITION 2.1 *Consider a generalized Nash game in which the jth firm solves*

$$
\begin{aligned}
\text{Ag}(z_{-(j)})\text{maximize} \quad & \pi_j(z_j; z_{-j}) \\
\text{subject to} \quad & z_j \in \mathbf{Z}_j \cap \mathcal{D}_j(z_{-j}),
\end{aligned}
$$

where $j = 1, \ldots, g+1$. Then, the associated generalized Nash equilibrium is given by a tuple $\{z_j^*\}_{j=1}^{g+1}$ where z_j^* solves the problem $\text{Ag}(z_{-j}^*)$ for all $j \in \mathcal{A}$ or $z_j^* \in \text{SOL}(\text{Ag}(z_{-j}^*))$ for $j = 1, \ldots, g+1$.

The classical Nash solution concept does not allow for an interaction in the strategy sets. Yet in our setting, we observe that the strategy sets are indeed coupled, leading to a generalized Nash game. In general, under suitable convexity and differentiability assumptions, the resulting equilibrium conditions of the shared-constraint Nash game are given by a quasi-variational inequality, an extension of the variational inequality [24,53]. Recent work by Facchinei *et al.* [20] has shown that if the strategy sets are coupled through a shared constraint, an equilibrium of the game is given by the solution of an appropriately defined scalar variational inequality. This holds in our setting where the firms and the ISO are coupled through

$$
\sum_{j \in \mathcal{J}_i} y_{ij}^\omega + r_i^\omega \geq 0 \quad \forall i \in \mathcal{N} \ \forall \omega \in \Omega.
$$

The analysis of generalized Nash equilibrium problems with a set of convex shared constraints has been studied recently in [17,19,20]. Consider a mapping $\mathbf{F}$ and a set $\mathbf{Z}$ given by[5]

$$
\mathbf{F}(z) \triangleq (-\nabla_{z_j} \pi_j(z))_{j=1}^{g+1}, \quad \mathbf{Z} \triangleq \left(\prod_{j=1}^{g+1} \mathbf{Z}_j \right) \cap \mathcal{D}, \tag{5}
$$

where

$$
\mathcal{D} \triangleq \left\{ z : \sum_{j \in \mathcal{J}_i} y_{ij}^\omega + r_i^\omega \geq 0, \ \forall i \in \mathcal{N}, \ \forall \omega \in \Omega \right\}.
$$

Note that $z \in \mathbb{R}^M, \mathbf{Z} \subseteq \mathbb{R}^M$ and $F : \mathbb{R}^M \longrightarrow \mathbb{R}^M$. Then the key result in [20] proves that the solvability of $\text{VI}(\mathbf{Z}, \mathbf{F})$ suffices for ensuring that the original shared-constraint game admits an equilibrium.[6] Recall that $\text{VI}(\mathbf{Z}, \mathbf{F})$ is defined as the problem of finding a vector $z^* \in \mathbf{Z}$ such that

$$
\mathbf{F}(z^*)^{\mathrm{T}}(z - z^*) \geq 0, \quad \forall z \in \mathbf{Z}.
$$

The equilibrium corresponding to a solution of this variational problem is referred to as the normalized equilibrium [53] or the variational equilibrium (VE) [20] and its relationship to the shared-constraint game is given by the following.

THEOREM 2.2   *Suppose the objective function $\pi_j(z_j; z_{-j})$ is concave and differentiable in z_j for all z_{-j} for all $j \in \mathcal{A}$ and $\mathcal{D}$, $\mathbf{Z}_1, \ldots, \mathbf{Z}_{g+1}$ are closed and convex sets. Then every solution to $\mathrm{VI}(\mathbf{Z}, \mathbf{F})$ is a solution to the shared-constraint game given by Definition 2.1.*

A similar result is available when π_j can only be shown to be continuous and concave for all $j \in \mathcal{A}$; specifically every solution to an appropriately defined variational inequality with a multivalued mapping is a solution to the game [19].

3. Existence and uniqueness of equilibria

The analysis of the VE rests on the properties of the variational inequality, denoted by $\mathrm{VI}(\mathbf{Z}, \mathbf{F})$. When $\mathbf{Z}$ is closed and convex and $\mathbf{F}$ is continuous, compactness of $\mathbf{Z}$ suffices for existence [18]. Similarly, uniqueness follows if $\mathbf{F}$ is strongly monotone over $\mathbf{Z}$, which requires that there exists a $\nu > 0$ such that

$$(\mathbf{F}(x) - \mathbf{F}(y))^{\mathrm{T}}(x - y) \geq \nu \|x - y\|^2 \quad \forall x, y \in \mathbf{Z}.$$

Unfortunately, in the current setting, neither compactness of $\mathbf{Z}$ nor strong monotonicity of $\mathbf{F}$ holds. These complications motivate a deeper analysis of $\mathrm{VI}(\mathbf{Z}, \mathbf{F})$ and represent the core of this section. Note that the analysis of variational inequalities enjoys a long history and an expansive discussion of these topics may be found in [18, Chapter 2,3]. We make the following assumptions on costs and prices and invoke them when necessary.

ASSUMPTION 3.1 (A1)   *The cost of generation ζ_{ij}^{ω} is a convex twice-continuously differentiable function of y_{ij}^{ω} for all $i \in \mathcal{N}, j \in \mathcal{J}$ and for all $\omega \in \Omega$.*
(A2)   *The nodal spot-market price is defined by the affine price function (1) for all $i \in \mathcal{N}$ and for all $\omega \in \Omega$.*

Invoking the definition of the conditional VaR, we can reformulate the non-smooth firm problem as a smooth convex program by the addition of a set of convex constraints, each corresponding to one realization of uncertainty. Effectively, agent j's parameterized problem is given by

$$\mathrm{Ag}(z_{-j}) \text{ maximize } \sum_{i \in G} \left(\pi_{ij}(x_{ij}) + \sum_{\omega \in \Omega} \rho^{\omega} \pi_{ij}^{\omega}(y_{ij}^{\omega}; r_i^{\omega}) - \kappa_j \left(m_{ij} + \sum_{\omega \in \Omega} \rho^{\omega} \frac{s_{ij}^{\omega}}{1 - \tau_j} \right) \right)$$

$$\text{subject to } \begin{cases} y_{ij}^{\omega} = x_{ij} + u_{ij}^{\omega} - v_{ij}^{\omega} \\ y_{ij}^{\omega} \leq \mathrm{cap}_{ij}^{\omega} \\ s_{ij}^{\omega} \geq \varrho_{ij}(x_{ij}, \mathrm{cap}_{ij}^{\omega}) - m_{ij} \\ \sum_{j \in \mathcal{J}_i} y_{ij}^{\omega} + r_i^{\omega} \geq 0 \\ x_{ij}, u_{ij}^{\omega}, v_{ij}^{\omega}, y_{ij}^{\omega}, s_{ij}^{\omega} \geq 0 \end{cases}, \forall i \in \mathcal{N} \quad \forall \omega \in \Omega.$$

Based on the redefinition of the agent problems, in this subsection, the mapping $\mathbf{F}$ is appropriately redefined to include the gradients of s_{ij}^{ω} and m_{ij}. Similarly, $\mathbf{Z}_j$ is extended to account for s_{ij}^{ω} and m_{ij}. The characterization of equilibria to the game (Definition 2.1) requires the following assumption on the loss function as well as a relationship between the slopes of the real-time and forward-market price functions.

ASSUMPTION 3.2 (A3) *The loss function $\varrho_{ij}(x_{ij}, cap_{ij}^\omega)$ is convex and increasing in x_{ij} and $\mathbb{E}b_i^\omega \leq 4b_i^0$ for all $i \in \mathcal{N}$.*

The above assumption ensures the convexity of the problem and allows for showing that the game admits an equilibrium. We begin by proving an intermediate result that shows that the objective function is convex under a mild assumption on the slopes of the price functions.

LEMMA 3.3 *Suppose (A1)–(A3) hold. Then the objective functions of the firms and the ISO are concave.*

Proof It suffices to prove the convexity of the expectation term of every agent's objective (w.r.t minimization), given by $\eta_{ij}(x_{ij}, y_{ij}; y_{i,-j})$, defined as

$$\eta_{ij}(x_{ij}, y_{ij}; x_{i,-j}, y_{i,-j}) = -\left(a_i^0 - b_i^0 \sum_{j \in \mathcal{J}} x_{ij}\right) x_{ij}$$

$$- \sum_{\omega \in \Omega} \rho^\omega \left(a_i^\omega - b_i^\omega \left(\sum_{j \in \mathcal{J}} y_{ij}^\omega + r_i^\omega\right)\right)(y_{ij}^\omega - x_{ij}).$$

The gradient and Hessian of this function are given by

$$\nabla \eta_{ij} = \begin{pmatrix} b_i^0 x_{ij} + b_i^0 \sum_{j \in \mathcal{J}} x_{ij} - a_i^0 + \sum_{\omega \in \Omega} \rho^\omega a_i^\omega - \sum_{\omega \in \Omega} \rho^\omega b_i^\omega \left(\sum_{j \in \mathcal{J}} y_{ij}^\omega + r_i^\omega\right) \\ \rho^\omega \left(-a_i^1 + b_i^1 \left(y_{ij}^1 + \sum_{j \in \mathcal{J}} y_{ij}^1\right) + b_i^1 r_i^1 - b_i^1 x_{ij}\right) \\ \vdots \\ \rho^n \left(-a_i^n + b_i^n \left(y_{ij}^n + \sum_{j \in \mathcal{J}} y_{ij}^n\right) + b_i^n r_i^n - b_i^n x_{ij}\right) \end{pmatrix}$$

and

$$\nabla^2 \eta_{ij} = \begin{pmatrix} 2b_i^0 & -\rho^1 b_i^1 & \cdots & -\rho^n b_i^n \\ -\rho^1 b_i^1 & 2\rho^1 b_i^1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\rho^n b_i^n & 0 & \cdots & 2\rho^n b_i^n \end{pmatrix},$$

respectively. Let s be an arbitrary non-zero vector. Then by adding and subtracting terms, we have

$$s^T \nabla^2 \eta_{ij} s = 2b_i^0 s_1^2 - 2s_1 \sum_{\omega=1}^n \rho^\omega b_i^\omega s_{\omega+1} + 2 \sum_{\omega=1}^n \rho^\omega b_i^\omega s_{\omega+1}^2$$

$$= \left(2b_i^0 - \sum_{\omega=1}^n \rho^\omega \frac{b_i^\omega}{2}\right) s_1^2 + \sum_{\omega=1}^n \rho^\omega \frac{b_i^\omega}{2} s_1^2 - 2s_1 \sum_{\omega=1}^n \rho^\omega b_i^\omega s_{\omega+1} + 2 \sum_{\omega=1}^n \rho^\omega b_i^\omega s_{\omega+1}^2$$

$$= \left(2b_i^0 - \sum_{\omega=1}^n \rho^\omega \frac{b_i^\omega}{2}\right) s_1^2 + \sum_{\omega=1}^n \rho^\omega b_i^\omega \left(\frac{s_1}{\sqrt{2}} - \sqrt{2}s_{\omega+1}\right)^2.$$

By assumption $\mathbb{E}(b_i^\omega) \leq 4b_i^0$ implying that $s^\mathrm{T}\nabla^2\eta_{ij}s > 0$ for all non-zero s and $\eta_{ij}(x_{ij}, y_{ij}; y_{i,-j})$ is a strictly convex function in x_{ij} and y_{ij} for all $x_{i,-j}$ and $y_{i,-j}$. The convexity of π_j in z_j follows from recalling that the generation costs and the CVaR measure are known to be convex. ∎

Having proved the convexity of the objectives, the variational inequality specifying the VE is necessary and sufficient. To claim the existence of a VE, we employ the following sufficiency condition [18]:

THEOREM 3.4   *Let* $\mathbf{Z}$ *be closed and convex and* $\mathbf{F}$ *be a continuous mapping. If there exists a vector* $z^\mathrm{ref} \in \mathbf{Z}$ *such that*

$$\liminf_{z\in\mathbf{Z},\|z\|\longrightarrow\infty} \boldsymbol{F}(z)^\mathrm{T}(z - z^\mathrm{ref}) > 0,$$

then the $\mathrm{VI}(\mathbf{Z},\boldsymbol{F})$ *has a non-empty compact solution set.*

Our main existence result proves that the variational problem associated with the game does indeed satisfy these conditions.

PROPOSITION 3.5 (Existence of a Nash equilibrium)   *Consider the stochastic Nash game specified by Definition* 2.1 *and let assumptions* (A1)–(A3) *hold. Then the game admits a non-empty compact set of equilibria.*

Proof   Based on Theorem 2.2, it suffices to prove the existence of a solution to $\mathrm{VI}(\mathbf{Z},\boldsymbol{F})$. By Theorem 3.4, this variational inequality is solvable if there exists a $z^\mathrm{ref} \in \mathbf{Z}$ such that the expression in Theorem 3.4 holds. If we set $(s_{ij}^\omega)^\mathrm{ref} \triangleq \varrho_{ij}(0, \mathrm{cap}_{ij}^\omega)$, and $x^\mathrm{ref}, y^\mathrm{ref}, r^\mathrm{ref}, m^\mathrm{ref} \triangleq 0$, then $z^\mathrm{ref} \in \mathbf{Z}$. It suffices to show that the coercivity result holds with u and v dropped from the formulation (since there exist no deviation penalties). By our choice of z^ref, the term $\boldsymbol{F}(z)^\mathrm{T}(z)$ can be written as

$$
\begin{aligned}
\boldsymbol{F}(z)^\mathrm{T}(z) &= \underbrace{\sum_{\omega\in\Omega}\sum_{i\in\mathcal{N}}\rho^\omega\left(-a_i^\omega + b_i^\omega\left(\sum_{j\in\mathcal{J}}y_{ij}^\omega + r_i^\omega\right)\right)r_i^\omega}_{(\boldsymbol{F}_r(z))^\mathrm{T}(r)} \\
&+ \underbrace{\sum_{j\in\mathcal{J}}\kappa_j\sum_{i\in\mathcal{N}_j}\sum_{\omega\in\Omega}\rho^\omega\left(\frac{s_{ij}^\omega - (s_{ij}^\omega)^\mathrm{ref}}{1-\tau} + m_{ij}\right)}_{(\boldsymbol{F}_{s_j}(z))^\mathrm{T}(s_j)+(\boldsymbol{F}_{m_j}(z))^\mathrm{T}(m_j)} \\
&+ \underbrace{\sum_{j\in\mathcal{J}}\sum_{\omega\in\Omega}\sum_{i\in\mathcal{N}_j}\rho^\omega\left(-a_i^\omega + \frac{\partial\zeta_{ij}^\omega}{\partial y_{ij}^\omega} + b_i^\omega y_{ij}^\omega + b_i^\omega\left(\sum_{j\in\mathcal{J}}y_{ij}^\omega + r_i^\omega\right) - b_i^\omega x_{ij}\right)y_{ij}^\omega}_{\boldsymbol{F}_{y_j}(z)^\mathrm{T}(y_j)} \\
&+ \underbrace{\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{N}_j}\left(b_i^0 x_{ij} + b_i^0\sum_{j\in\mathcal{J}}x_{ij} - a_i^0 + \sum_{\omega\in\Omega}\rho^\omega a_i^\omega - \sum_{\omega\in\Omega}\rho^\omega b_i^\omega\left(\sum_{j\in\mathcal{J}}y_{ij}^\omega + r_i^\omega\right)\right)x_{ij}}_{\boldsymbol{F}_{x_j}(z)^\mathrm{T}(x_j)}
\end{aligned}
$$

$$
= \underbrace{\sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} \sum_{i \in \mathcal{N}_j} \rho^\omega \left(-a_i^\omega + \frac{\partial \zeta_{ij}^\omega}{\partial y_{ij}^\omega} + b_i^\omega y_{ij}^\omega + b_i^\omega \left(\sum_{j \in \mathcal{J}} y_{ij}^\omega + r_i^\omega \right) \right) y_{ij}^\omega}_{\text{term 1}}
$$

$$
+ \underbrace{\sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{N}_j} \left(b_i^0 x_{ij} + b_i^0 \sum_{j \in \mathcal{J}} x_{ij} - a_i^0 + \sum_{\omega \in \Omega} \rho^\omega a_i^\omega - \sum_{\omega \in \Omega} \rho^\omega b_i^\omega \left(\sum_{j \in \mathcal{J}} 2 y_{ij}^\omega + r_i^\omega \right) \right) x_{ij}}_{\text{term 2}}
$$

$$
+ \underbrace{\sum_{\omega \in \Omega} \sum_{i \in \mathcal{N}} \rho^\omega \left(-a_i^\omega + b_i^\omega \left(\sum_{j \in \mathcal{J}} y_{ij}^\omega + r_i^\omega \right) \right) r_i^\omega}_{\text{term 3}}
$$

$$
+ \underbrace{\sum_{j \in \mathcal{J}} \kappa_j \sum_{i \in \mathcal{N}_j} \sum_{\omega \in \Omega} \rho^\omega \left(\frac{s_{ij}^\omega - (s_{ij}^\omega)^{\text{ref}}}{1 - \tau} + m_{ij} \right)}_{\text{term 4}}. \tag{6}
$$

From the boundedness of y, the definition of the shared constraints and from $\sum_{i \in \mathcal{N}} r_i^\omega = 0, \forall \omega \in \Omega$, the boundedness of r follows. Therefore, we may conclude that term 1 is bounded and for any sequence $\{z_k\}$, such that $\|z_k\| \to \infty$, it follows that one of the sequences $\{\|x_k\|\}$, $\{\|s_k\|\}$ and $\{|m_k|\}$ is tending to $+\infty$.

Case 1: Suppose the forward generation bid x^k tends to infinity implying that term 2 tends to $+\infty$ at a quadratic rate.

Case 2: Suppose either (or both) s^k or $|m^k|$ tend to $+\infty$. $s^k \in \mathbf{Z}$, $m^k \in \mathbf{Z}$, $s^k \geq 0$ and $s^k + m^k$ is bounded from below. If, m^k tends to $-\infty$, then s^k tends to $+\infty$. Hence, term 4 grows to $+\infty$. If m^k or s^k tend to $+\infty$, then term 4 tends to $+\infty$. [7]

Case 3: Suppose x^k tends to $+\infty$ and any combination of s^k, and $|m^k|$ tends to $+\infty$. If m^k alone tends to $-\infty$, then term 4 tends to $-\infty$ and term 2 tends to $+\infty$ at a quadratic rate. Consequently, the entire sum tends to $+\infty$. If s^k tends to $+\infty$, m^k reduces to $-\infty$ and x^k tends to $+\infty$ then Cases 1 and 3 can be used in conjunction. The other possibilities lead to immediate results of the sequence tending to $+\infty$.

Consider any sequence $\{z^k\} \in \mathbf{Z}$ such that $\lim_{k \to \infty} \|z^k\| = \infty$. Since none of the terms tend to $-\infty$ and at least one of the terms tend to ∞, it follows that

$$
\liminf_{z \in \mathbf{Z}, \|z\| \longrightarrow \infty} \mathbf{F}(z)^{\mathrm{T}}(z) = \infty.
$$

This completes the proof. ∎

A uniqueness result rests on being able to show that the mapping is strictly monotone. However, in the current setting, the mapping arising from the risk-based game can only be shown to be monotone, as the next result shows. This requires showing $\nabla \mathbf{F}$, given by

$$
\nabla \mathbf{F}(z) = \begin{pmatrix} \nabla_1 \mathbf{F}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \nabla_2 \mathbf{F}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \nabla_N \mathbf{F}_N \end{pmatrix},
$$

is positive semidefinite for all z, where $\nabla_i F_i$ represents the gradient mapping (in the order $x_i, y_i, r_i, s_i, m_i, u_i$ and v_i) with respect to the nodal variables corresponding to node i. If e and

I denote the column of ones and the identity matrix, respectively, then the matrix $\nabla F_i, \forall i \in G$ is given by

$$
\nabla \mathbf{F}_i = \begin{pmatrix}
P_i^0 & P_i^1 & \dots & P_i^n & H_i & 0 \\
R_i^1 & S_i^1 & \dots & 0 & F_i^1 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
R_i^n & 0 & \dots & S_i^n & F_i^n & 0 \\
0 & T_i^1 & \dots & T_i^n & K_i & 0 \\
0 & 0 & \dots & \dots & 0 & V_i
\end{pmatrix},
$$

where $P_i^\omega = -\rho^\omega b_i^\omega e\, e^{\mathrm T}$, $R_i^\omega = -\rho^\omega I$, $S_i^\omega = \rho^\omega b_i^\omega (I + e\, e^{\mathrm T}) + \mathrm{diag}(d_{i1}^\omega \ \dots \ d_{ig}^\omega) \forall \omega \in \Omega$, $P_i^0 = b_i^0 (I + e\, e^{\mathrm T})$, $H_i = e(\rho^1 b_i^1 \ \dots \ \rho^n b_i^n)$, $K_i = \mathrm{diag}(\rho^1 b_i^1 \ \dots \ \rho^n b_i^n)$ and $V_i = \mathbf{0}$. Furthermore, $F_i^1, \dots, F_i^n$ and $T_i^1, \dots, T_i^n$ are defined as

$$
F_i^1 = \begin{pmatrix} \rho^1 b_i^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ \rho^1 b_i^1 & \dots & 0 \end{pmatrix}, \dots, \quad
F_i^n = \begin{pmatrix} 0 & \dots & \rho^n b_i^n \\ \vdots & \ddots & \vdots \\ 0 & \dots & \rho^n b_i^n \end{pmatrix},
$$

$$
T_i^1 = \begin{pmatrix} \rho^1 b_i^1 & \dots & \rho^1 b_i^1 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \dots, \quad
T_i^n = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ \rho^n b_i^n & \dots & \rho^n b_i^n \end{pmatrix}.
$$

Note that V_i refers to the zero matrix representing the second-order derivatives with respect to s, m, u and v. Next, we show that the game in question admits a unique ϵ-Nash equilibrium, given by a solution to $\mathrm{VI}(K, \mathbf{F}_\epsilon)$ where $\mathbf{F}_\epsilon = F + \epsilon \mathbf{I}$ and $\mathbf{I}$ is the identity mapping.

PROPOSITION 3.6 (Uniqueness of regularized variational equilibrium) *Consider the stochastic Nash game given by Definition* 2.1 *and let* (A1)–(A3) *hold. Then the resulting mapping* $\mathbf{F}(z)$ *is monotone over* $\mathbf{Z}$ *and the stochastic Nash game admits a unique* ϵ-*Nash equilibrium.*

Proof Since, the matrix V_i is a zero matrix, it suffices to show that principal submatrix of ∇F_i, without the last row and column corresponding to V_i, is positive semidefinite. If $\hat{\mathbf{F}}$ represents this mapping in the reduced space, it suffices to show that for all $s \neq 0$ we have $s^{\mathrm T} \nabla \hat{\mathbf{F}}_i s > 0$ where $s^{\mathrm T} \nabla \hat{\mathbf{F}}_i s$ is given by

$$
s^{\mathrm T} \nabla \hat{\mathbf{F}}_i s = b_i^0 \sum_{k=1}^{g} s_k^2 + b_i^0 \left(\sum_{k=1}^{g} s_k \right)^2 - \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \sum_{k=1}^{g} s_k s_{\omega g+k}
$$

$$
- \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \sum_{k=1}^{g} s_k \sum_{k=1}^{g} s_{\omega g+k} + \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(\sum_{k=1}^{g} s_{\omega g+k}^2 + \left(\sum_{k=1}^{g} s_{\omega g+k} \right)^2 \right)
$$

$$
+ \sum_{\omega=1}^{n} \rho^\omega \left(\sum_{k=1}^{g} d_{ik}^\omega s_{\omega g+k}^2 \right)
$$

$$
+ \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \sum_{k=1}^{g} s_{\omega g+k} \right) - \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \sum_{k=1}^{g} s_k \right)
$$

$$
+ \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \sum_{k=1}^{g} s_{\omega g+k} \right) + \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \right)^2.
$$

Adding and subtracting terms, the right-hand side is given by

$$
\begin{aligned}
s^{\mathrm{T}} \nabla \hat{\mathbf{F}}_i s = {} & \left(b_i^0 - \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \right) \sum_{k=1}^{g} s_k^2 + \left(b_i^0 - \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \right) \left(\sum_{k=1}^{g} s_k \right)^2 - \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \sum_{k=1}^{g} s_k s_{\omega g+k} \\
& + \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \sum_{k=1}^{g} s_k^2 + \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \left(\sum_{k=1}^{g} s_k \right)^2 - \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \sum_{k=1}^{g} s_k \sum_{k=1}^{g} s_{\omega g+k} \\
& + \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(\sum_{k=1}^{g} s_{\omega g+k}^2 + \left(\sum_{k=1}^{g} s_{\omega g+k} \right)^2 \right) + \sum_{\omega=1}^{n} \rho^\omega \left(\sum_{k=1}^{g} d_{ik}^\omega s_{\omega g+k}^2 \right) \\
& + 2 \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \sum_{k=1}^{g} s_{\omega g+k} \right) - \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \sum_{k=1}^{g} s_k \right) \\
& + \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} \right)^2 .
\end{aligned}
$$

On rearranging, $s^{\mathrm{T}} \nabla \hat{\mathbf{F}}_i s$ is given by

$$
\begin{aligned}
& \left(b_i^0 - \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \right) \sum_{k=1}^{g} s_k^2 + \left(b_i^0 - \sum_{\omega=1}^{n} \frac{\rho^\omega b_i^\omega}{4} \right) \left(\sum_{k=1}^{g} s_k \right)^2 + \sum_{\omega=1}^{n} \left(\rho^\omega b_i^\omega \sum_{k=1}^{g} \left(\frac{s_k}{2} - s_{\omega g+k} \right)^2 \right) \\
& + \sum_{\omega=1}^{n} \rho^\omega \left(\sum_{k=1}^{g} d_{ik}^\omega s_{\omega g+k}^2 \right) + \sum_{\omega=1}^{n} \rho^\omega b_i^\omega \left(s_{(n+1)g+\omega} + \sum_{k=1}^{g} s_{\omega g+k} - \sum_{k=1}^{g} \frac{s_k}{2} \right)^2 .
\end{aligned}
$$

Since $4b_i^0 \geq \mathbb{E} b_i^\omega$ holds by assumption, it follows that $s^{\mathrm{T}} \nabla \hat{\mathbf{F}}_i s \geq 0$ for all $i \in G$ implying that ∇F_i is also positive semidefinite for all $i \in G$. The gradient mapping for all $i \in G^c$ is given by a mapping with all zeros except for the block K_i that is positive semi-definite. Since the gradient mappings corresponding to the load nodes are positive semidefinite, the positive semidefiniteness of the entire gradient mapping $\nabla \mathbf{F}$ follows. Consequently, $\mathbf{F}$ is a monotone mapping and its regularization, namely $\mathbf{F}_\epsilon = \mathbf{F} + \epsilon \mathbf{I}$, is a strongly monotone mapping. It follows that a unique solution to $\mathrm{VI}(\mathbf{Z}, \mathbf{F}_\epsilon)$ exists, allowing us to conclude a unique ϵ-Nash equilibrium exists. ∎

Note that since $\mathbf{F}$ is a continuous monotone mapping, if a solution to $\mathrm{VI}(\mathbf{Z}, \mathbf{F})$ is shown to be locally unique, then global uniqueness of the Nash equilibrium follows [18, Theorem 3.6.6]. Local uniqueness of the VE is not immediate and requires a closer examination of ∇F and remains a focus of future work.

4. A scalable cutting-plane projection scheme

The game-theoretic problem introduced in Section 2 leads to monotone stochastic variational problems. While much exists for solving monotone variational inequalities [18,35], unfortunately most schemes can neither be implemented in a distributed setting (since the constraint sets are coupled) nor do they possess the scalability required to address this class of problems since our class of problems can be arbitrarily large in terms of the number of agents, the size of the network and the cardinality of the sample space.

It should be remarked that there have been relatively few attempts to examine the class of generalized Nash games while even fewer have considered their stochastic counterparts. Fukushima and Pang [50] suggested a sequential penalization approach for solving such problems while a review of approaches can be found in [17]. Of note is a recent approach that uses a Nikaido–Isoda (NI) function by Von Heusinger and Kanzow [28] and a relaxation algorithm using the NI function by Krawczyk and Uryasev [38]. In [16,61], sample-average approximation schemes are suggested but neither the coupled nature of their strategy sets nor their semi-infinite nature can be accommodated. Note that it may be possible to employ their approach on the complementarity problem that emerges from this setting. An alternate scheme that relies on matrix splitting techniques is suggested in [57].

Accordingly, our focus is on developing convergent algorithms with suitable error bounds, for addressing this class of problems. We place an emphasis on the construction of distributed schemes that scale with the cardinality of the sample space, namely $|\Omega|$, the number of agents $|\mathcal{J}|$ and the size of the network $|\mathcal{N}|$. To address these needs, we develop a *distributed projection-based method* that employs a *cutting-plane method* for solving the agent-specific projection problems.

In Section 4.1, we describe a dual and a primal–dual projection method for the solution of shared-constraint stochastic Nash games. At the heart of these schemes is a projection step which in general leads to a massive stochastic convex program. In Section 4.2, we employ a cutting-plane method for the solution of such problems that scales with $|\Omega|$. Rate estimates and error bounds, particularly for inexact generalizations, are presented for the projection schemes in Section 4.3. Finally, in Section 4.4, while examining the numerical behavior of the schemes, we observe that the schemes display the desired scalability properties and the inexact generalizations prove to have significant benefits.

4.1 *Distributed primal–dual and dual projection methods*

We begin with an introduction to projection(or gradient)-based methods for monotone variational inequalities. Given a variational inequality VI($\mathbf{Z}, \mathbf{F}$) where $\mathbf{F}$ is a strongly monotone Lipschitz continuous mapping over $\mathbf{Z}$ with monotonicity constant η and Lipschitz constant L, then given an x^0 one may construct a sequence $\{x^k\}$ as follows:

$$x^{k+1} := \Pi_{\mathbf{Z}}(x^k - \gamma \mathbf{F}(x^k)), \quad k \geq 0, \tag{7}$$

where $\Pi_{\mathbf{Z}}(y)$ is the projection of y on $\mathbf{Z}$. If $\gamma < 2\eta/L^2$, then $x^k \longrightarrow x^*$ as $k \longrightarrow \infty$, where x^* is a solution of VI($\mathbf{Z}, \mathbf{F}$). This gradient-based framework may be extended to a multi-player game-theoretic regime, as shown next.

Consider an N-player deterministic Nash game in which the jth agent solves the parameterized convex optimization problem given by

$$
\boxed{
\begin{array}{ll}
\text{A}(z_{-j}) & \text{maximize} \quad \pi_j(z_j; z_{-j}) \\
& \text{subject to} \quad z_j \in \mathbf{Z}_j,
\end{array}
}
$$

where $\pi_j(z_j; z_{-j})$ is a convex differentiable function of z_j for all z_{-j} and $\mathbf{Z}_j$ is a closed and convex set. Then a standard distributed projection scheme is given by

$$z_j^{k+1} := \Pi_{\mathbf{Z}_j}(z_j^k + \gamma \nabla \pi_j(z_j^k; z_{-j}^k)) \quad \text{for all } j = 1, \dots, N,$$

where γ is a fixed steplength. Yet, the convergence of such a scheme relies on two properties: First, the gradient mapping given by $F(z)$ needs to satisfy strict monotonicity, strong monotonicity or

co-coercivity property [18] over a set $\mathbf{Z}$ where $F(z)$ and $\mathbf{Z}$ are defined as

$$F(z) := -(\nabla_{z_i}\pi_i)_{i=1}^{g+1} \quad \text{and} \quad \mathbf{Z} \triangleq \prod_{j=1}^{g+1} \mathbf{Z}_j.$$

Second, the strategy sets across agents cannot be coupled. In our setting, neither assumption holds and therefore a direct application of the aforementioned approach cannot be employed. Note that this approach will be referred to as a *primal* approach, since it does not involve relaxing the shared constraints. If one does proceed with such a relaxation, then a scheme that updates both primal and dual variables (associated with the shared constraints) is considered.[8]

Instead, we observe that the shared-constraint game can be cast as a monotone complementarity problem in the primal–dual space. By solving a sequence of regularized (and therefore strongly monotone) complementarity problems through a Tikhonov regularization scheme [18], we obtain a solution to the original problem. Note that the monotonicity of the mapping in the primal–dual space suffices for the Tikhonov trajectory to converge to the solution of the original problem [18, Chapter 12]. This avenue allows us to leverage fixed steplength projection schemes for the solution of each regularized complementarity problem. Importantly, each subproblem can be massive, with a size proportional to $|\Omega| \times |\mathcal{J}| \times |\mathcal{N}|$, and a direct solution of such problems is only possible in modest settings. To cope with such a challenge, we develop a distributed framework that relies on decomposition methods that scale well with all three sources of complexity.

We now proceed to describe the distributed projection framework. If the Lagrange multipliers corresponding to the shared constraint, denoted by $d(z) \geq 0$, are denoted by λ, then it follows that (z^*, λ^*) is an equilibrium of shared-constraint Nash game if and only if (z^*, λ^*) is a solution of set of coupled fixed-point problems:

$$z = \Pi_{\mathbf{Z}}(z - \gamma\mathbf{F}_z(z, \lambda)), \tag{8}$$

$$\lambda = \Pi_{\mathbb{R}_m^+}(\lambda - \gamma\mathbf{F}_\lambda(z, \lambda)), \tag{9}$$

where

$$\mathbf{F}_z(z, \lambda) = \begin{pmatrix} -\nabla_{z_1}\pi_1 - \nabla_{z_1}d(z)^{\mathrm{T}}\lambda \\ \vdots \\ -\nabla_{z_N}\pi_N - \nabla_{z_N}d(z)^{\mathrm{T}}\lambda \end{pmatrix} \quad \text{and} \quad \mathbf{F}_\lambda(z, \lambda) = d(z). \tag{10}$$

The fixed-point representations motivate a *primal–dual* method that requires constructing a primal and dual method on the same timescale (primal and dual steps taken simultaneously) with a fixed steplength γ_{pd} in a regularized setting. Specifically, this entails the following set of regularized primal and dual steps for $k \geq 0$:

$$z_j^{k+1} = \Pi_{\mathbf{Z}_j}(z_j^k - \gamma_{\mathrm{pd}}(\mathbf{F}_z(z_j^k; z_{-j}^k, \lambda^k) + \epsilon^\ell z_j^k)) \quad \text{for all } j, \tag{11}$$

$$\lambda^{k+1} = \Pi_{\mathbb{R}_m^+}(\lambda^k - \gamma_{\mathrm{pd}}(\mathbf{F}_\lambda(z^k, \lambda^k) + \epsilon^\ell \lambda^k)), \tag{12}$$

where ϵ^ℓ is the regularization parameter at the ℓth iteration of the outer Tikhonov scheme. In the regularized primal–dual approach, the steplength γ_{pd} has to be chosen in accordance with the monotonicity and Lipschitz constant of the appropriate mappings in *both* the primal and dual spaces (see Section 4.3 for more details). In effect, if the mappings in one of the spaces has a large Lipschitz constant (or alternately a low monotonicity constant), the progress of the entire algorithm may be hampered.

A dual method for solving the monotone complementarity problem does not tie the primal and dual steplengths together and can be employed instead. This requires that for every update in the

dual space, an *exact* primal solution is required. In particular, for $k \geq 0$, this leads to a set of iterations given by

$$z_j^k = \Pi_{\mathbf{Z}_j}(z_j^k - \gamma_d(\mathbf{F}_z(z_j^k; z_{-j}^k, \lambda^k) + \epsilon^\ell z_j^k)) \quad \text{for all } j \tag{13}$$

$$\lambda^{k+1} = \Pi_{\mathbb{R}_m^+}(\lambda^k - \gamma_p(\mathbf{F}_\lambda(z^k, \lambda^k) + \epsilon^\ell \lambda^k)), \tag{14}$$

where γ_p and γ_d are the primal and dual steplengths, respectively. The termination of the inner scheme occurs when the error in the fixed-point problem falls within a threshold and is ensured by the following for the primal–dual scheme

$$\left\| \left(\frac{\|z^{k+1} - z^k\|}{1 + \|z^k\|}, \frac{\|\lambda^{k+1} - \lambda^k\|}{1 + \|\lambda^k\|} \right) \right\| \leq \epsilon^{\text{inner}}, \tag{15}$$

and the dual scheme

$$\frac{\|\lambda^{k+1} - \lambda^k\|}{1 + \|\lambda^k\|} \leq \epsilon^{\text{inner}}. \tag{16}$$

The exact solution of such a problem may prove difficult, suggesting instead that we may need to employ inexact or approximate solutions. Expectedly, this would lead to errors that require quantification. This analysis is provided, along with suitable convergence results, in Section 4.3. We conclude this subsection with an algorithm statement for the projection-based schemes.

Algorithm 1 Distributed primal–dual and dual projection methods

---

    Initialize $k = 0, \ell = 0$

    Choose constants $\epsilon^0, \epsilon^{inner}, \epsilon^{outer} > 0$ and γ_{pd}, γ_p and γ_d, initial solution (z^0, λ^0), $\bar{\gamma} < 1$

3: **while** $\epsilon^\ell > \epsilon^{outer}$ **do**

      while conditions (15) or (16) are not satisfied **do**

        Let λ^{k+1} be given by (12) (Primal-dual) or (14) (Dual)

6:       Let z^{k+1} be given by (11) (Primal-dual) or the solution of (13) (Dual)

        $k := k + 1$

        Update regularization $\epsilon^{\ell+1} := \bar{\gamma}\epsilon^\ell$

9:       $\ell := \ell + 1$

      end while

    end while

4.2 *A scalable cutting-plane method for the projection problem*

In the projection schemes presented in the earlier section, the solution of the primal projection step, as denoted by (11) and (13), requires the solution of a large convex program of size $O(|\Omega|)$. This is generally only possible via direct solvers for modest sample spaces and in this subsection, we discuss how one may solve such problems in a scalable fashion for arbitrarily large sample spaces.

In the current setting, $\mathbf{Z}_j$ is a polyhedral set implying that the projection problem is a quadratic program (QP) and, given that the problem originates from a projection problem, this QP is, in fact, strongly convex. In the past, QPs have been solved by a variety of schemes, such as interior-point methods, active-set methods and others [48]. All of these schemes are necessarily direct approaches in that they make no obvious effort to utilize the structure of the problem. However, in this instance, the problems belong to a class of *recourse-based stochastic quadratic programs* [7].

The key computational challenge in solving recourse-based stochastic optimization problems lies in ensuring that scenario-specific second-stage problems are solved in parallel, effectively allowing for a scalable method. In 1969, based on a decomposition scheme suggested by Benders [6], Van-Slyke and Wets [60] stated a cutting-plane method for the solution of recourse-based stochastic linear programs (LPs) that allows for precisely such a parallelization. While much has been done on the solution of stochastic LPs (cf. [7,33]), stochastic convex programming has been less studied in general [52]. Parallel schemes for the solution of stochastic QPs via splitting and projection methods were discussed by Womersley and Chen [13] while extensions to the L-shaped cutting-plane method have been suggested by Zakeri *et al.* [64]. More recently, Kulkarni and Shanbhag [40] and Kulkarni *et al*. [39] developed an inexact-cut and a trust-region L-shaped method for solving stochastic QPs that was subsequently employed as a QP solver within a more general sequential quadratic programming method for solving non-convex stochastic NLPs [39,40]. We employ a similar L-shaped scheme for solving the stochastic quadratic program arising from the projection problem.

Computing the projection in the primal space (13) and (11), requires solving a stochastic program given by

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}(\hat{x}_j - \bar{z}_j^k)^{\text{T}}(\hat{x}_j - \bar{z}_j^k) \\
\text{subject to} & \hat{x}_j \in \mathbf{Z}_j,
\end{array}
$$

where

$$
\bar{z}_j^k = (z_j^k - \gamma F_{z_j}(z_j; z_{-j}^k, \lambda^k)), \quad \hat{x}_j = \begin{pmatrix} \hat{x}_j \\ (\hat{y}_j^\omega)_{\omega \in \Omega} \end{pmatrix}, \quad \hat{x}_j = \begin{pmatrix} x_j \\ m_j \end{pmatrix},
$$

$$
\hat{y}_j^\omega = \begin{pmatrix} u_j^\omega \\ v_j^\omega \\ y_j^\omega \\ s_j^\omega \end{pmatrix}, \quad \forall j \in \mathcal{J}, \hat{x}_{g+1} = (0), \quad \hat{y}_{g+1}^\omega = \begin{pmatrix} r_1^\omega \\ \vdots \\ r_N^\omega \end{pmatrix}.
$$

In settings where the loss function in the risk measure is affine (or in the risk-neutral deviation cost setting), the projection problem reduces to a stochastic quadratic program given by

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}\hat{x}_j^{\text{T}}\hat{x}_j + \hat{x}_j^{\text{T}}\bar{x}_j + \sum_{\omega \in \Omega}\left(\frac{1}{2}(\hat{y}_j^\omega)^{\text{T}}(\hat{y}_j^\omega) - (\hat{y}_j^\omega)^{\text{T}}\bar{y}_j^\omega\right) \\
\text{subject to} & (\hat{x}_j, \hat{y}_j^\omega) \in \mathbf{Z}_j = \left\{(\hat{x}_j, \hat{y}_j^\omega) : \begin{cases} A_1\hat{x}_j + A_2\hat{y}_j^\omega = \hat{b}_j^\omega \\ A_3\hat{y}_j^\omega \leq \text{cap}_{ij}^\omega, \\ A_4\hat{x}_j \geq 0, \\ \hat{y}^\omega \geq 0, \end{cases} \right\} \forall \omega \in \Omega \right\},
\end{array}
$$

where A_1, A_2, A_3 and A_4 are defined appropriately. As Ω grows in cardinality, a direct solution of the quadratic program becomes challenging. Instead, we pursue a stochastic programming avenue by noting that the constraint structure allows one to cast the problem as a recourse-based stochastic program. Specifically, we have

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}\hat{x}_j^{\text{T}}\hat{x}_j + \hat{x}_j^{\text{T}}\bar{x}_j + \mathcal{Q}(\hat{x}_j) \\
\text{subject to} & A_4\hat{x}_j \geq 0,
\end{array}
$$

where $\mathcal{Q}(\hat{x}_j)$, the cost of recourse is given by $\mathcal{Q}(\hat{x}_j) = \mathbb{E}Q(\hat{x}_j; \omega)$ and $Q(\hat{x}_j; \omega)$ is the optimal value of the scenario-specific quadratic program:

$$
\boxed{
\begin{array}{ll}
\mathrm{Sub}(\hat{x}_j; \omega) & \text{minimize} \quad (\frac{1}{2}(\hat{y}_j^\omega)^{\mathrm{T}}(\hat{y}_j^\omega) - (\hat{y}_j^\omega)^{\mathrm{T}}\bar{y}_j^\omega) \\
& \text{subject to} \quad y_j^\omega \in \mathbf{Y}_j^\omega(\hat{x}_j)
\end{array}
}
$$

and

$$
\mathbf{Y}^\omega(\hat{x}_j) = \left\{ \hat{y}_j^\omega : \left\{ \begin{array}{c} A_2\hat{y}_j^\omega = \hat{b}_j^\omega - A_1\hat{x}_j \\ A_3\hat{y}_j^\omega \leq \mathrm{cap}_{ij}^\omega \\ \hat{y}^\omega \geq 0 \end{array} \right\} \right\}.
$$

It should be emphasized that, in general, a first-stage decision $\hat{x}$ might render the $\mathbf{Y}^\omega(\hat{x})$ empty. However, in this particular case, the non-negative deviation levels u^ω and v^ω can be made arbitrarily large to ensure that the second-stage problem is *always feasible* and the resulting problem is said to possess complete recourse.

The L-shaped method for the solution of stochastic QPs requires solving a sequence of increasingly constrained (QPs) (called the master problem) where the additional constraints, termed as cuts, arise from the solution of the set of scenario-specific second-stage problems. The master problem is given by

$$
\boxed{
\begin{array}{ll}
\mathrm{Master}_k \underset{\hat{x}_j, \theta_j}{\text{minimize}} & \frac{1}{2}\hat{x}_j^{\mathrm{T}}\hat{x}_j + \hat{x}_j^{\mathrm{T}}\bar{x}_j + \theta_j \\
\text{subject to} & \begin{array}{rcl} A_4\hat{x}_j & \geq & 0_j \\ \theta - G_{j,i}^{\mathrm{T}}\hat{x}_j & \geq & g_{j,i}, \quad i = 1, \ldots, k, \end{array}
\end{array}
}
$$

where $(G_{j,i}, g_{j,i})$ are the coefficients of the ith (see [56] for more details) defined as

$$
G_{j,i} \triangleq -\sum_{\omega \in \Omega} A_1^{\mathrm{T}}\pi^\omega \quad \text{and} \quad g_{j,i} \triangleq \sum_{\omega \in \Omega} (\pi^\omega)^{\mathrm{T}}\hat{b}_j^\omega - \frac{1}{2}\sum_{\omega \in \Omega}(\hat{y}^\omega)^{\mathrm{T}}\hat{y}^\omega,
$$

where π^ω represents the vector of dual variables corresponding to the sub problem (scenario ω) and I represents the identity matrix. Note that the ith cut associated with the jth agent requires the solution of $\mathrm{Sub}(\hat{x}_j^i)$. It is worth reiterating that the complexity arising from a massive sample space is addressed by decomposing what is a potentially massive QP into a set of $|\Omega|$ smaller QPs. In the L-shaped method, the termination is contingent on the lower bound L_j^k and upper bound U_j^k, where L_j^k and U_j^k are, respectively, defined as

$$
L_j^k \equiv \frac{1}{2}(\hat{x}_j^k)^{\mathrm{T}}(\hat{x}_j^k) + (\hat{x}_j^k)^{\mathrm{T}}\bar{x}_j^k + \theta_j^k \quad \text{and} \quad U_j^k \equiv \min\{U_j^{k-1}, \frac{1}{2}(\hat{x}_j^k)^{\mathrm{T}}(\hat{x}_j^k) + (\hat{x}_j^k)^{\mathrm{T}}\bar{x}_j^k + \mathcal{Q}(\hat{x}_j^k)\}.
$$

Notice that $\{L_j^k\}$ is a monotonically increasing sequence while $\{U_j^k\}$ is a monotonically decreasing sequence. Algorithm 2 provides a formal statement of the L-shaped method [56] and its convergence is easily proved and can be found in [7,54].

4.3 *Convergence and error analysis of projection methods*

Convergence of projection schemes is reliant on the underlying mappings satisfying a strict or strong monotonicity property. The absence of such a property may be addressed through a

---

Algorithm 2 L-shaped method.

Initialize $k = 1, j \in \mathcal{J}, U_j^k = \infty, L_j^k = -\infty$
Choose $\epsilon_1, \tau, u > 1$
3: **while** $|U_j^k - L_j^k| > \tau$ **do**
    Solve (Master$_k$) to get $(\hat{x}_j^k, \theta_j^k)$
    Update lower bound L_j^k
6:    Solve Sub$(\hat{x}_j^k; \omega)$ for all $\omega \in \Omega$
    Construct (G_I^k, g_I^k)
    Update upper bound U_j^k and add optimality cut (G_I^k, g_I^k) to (Master$_k$)
9:    $k = k + 1$
end while

Tikhonov-based regularization scheme [18]. Each iterate of the Tikhonov scheme may be solved efficiently and in this subsection, we provide the convergence theory for the suggested dual and primal–dual schemes for solving precisely such problems. In this section, we present three sets of results. First, our convergence statements require a precise specification of the Lipschitz and monotonicity constants of the relevant mapping and represents our first result. Second, we present a convergence result for the dual scheme in a regularized setting and further equip this result with rate estimates. The *exact* form of the dual scheme requires exact primal iterates for a given dual solution. In a regime where a bound on the primal strategy sets is assumed to be available, we relax this requirement in constructing an *inexact dual method* and allow for *inexact* primal solutions. The third set of results focus on developing error bounds for the inexact dual scheme in this setting along with suitable bounds on the primal suboptimality and primal infeasibility.

Before proving the Lipschitzian and monotonicity properties of $\mathbf{F}^\epsilon$, we consider the polyhedral shared constraint denoted by $Bz \geq 0$ and provide a precise relationship between $\|B\|$ and the problem size, where z is specified as follows:

$$z = \begin{pmatrix} \bar{p}_1 \\ \vdots \\ p_{N_g} \\ \bar{p}_0 \end{pmatrix}, \quad \bar{p}_i = \begin{pmatrix} \bar{p}_i^1 \\ \vdots \\ \bar{p}_i^n \end{pmatrix}, \quad \bar{p}_i^\omega = \begin{pmatrix} y_{i1}^\omega \\ \vdots \\ y_{iJ}^\omega \\ r_i^\omega \end{pmatrix} \quad \forall i \in G \;\; \forall \omega \in \Omega,$$

and $\bar{p}_0$ represents the other components of the vector z, not indicated above. Consequently, the matrix B is defined as

$$B = \begin{pmatrix} B_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & B_{N_g} & 0 \end{pmatrix} \quad \text{where } B_i = \begin{pmatrix} B_i^1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & B_i^n \end{pmatrix},$$

$$B_i^\omega = \begin{pmatrix} 1 & \ldots & 1 \end{pmatrix} \quad \forall i \in G \; \forall \omega \in \Omega. \tag{17}$$

The following result gives a bound on $\|B\|$, that is subsequently employed in our rate analysis.

LEMMA 4.1 *Consider the matrix B defined in (17). If N_f and N_g are the total number of players in the game and the number of generating nodes, respectively, then $\|B\|_2 \leq \sqrt{N_f N_g n}$.*

Proof Recall that $\|B\|_2 \le \|B\|_F$, where $\|B\|_F$ represents the Froebenius norm of the matrix (see [23]). When B is given by (17), then

$$\|B\|_F = \sqrt{\sum_{\omega \in \Omega} \sum_{i \in \mathcal{N}_g} (\|\mathcal{J}\| + 1)} = \sqrt{N_f N_g n}.$$

$\blacksquare$

By recalling the definitions of $\mathbf{F}_z$ and $\mathbf{F}_\lambda$ in (10), we further define $\mathbf{F}_z^\epsilon, \mathbf{F}_\lambda^\epsilon, \mathbf{F}_f^\epsilon$ and $\mathbf{F}_d$ as $\mathbf{F}_z^\epsilon := \mathbf{F}_z + \epsilon z, \mathbf{F}_\lambda^\epsilon := \mathbf{F}_\lambda + \epsilon \lambda$ and

$$\mathbf{F}_f^\epsilon := \begin{pmatrix} \nabla_{z_1} \pi_1 + \epsilon z_1 \\ \vdots \\ \nabla_{z_{g+1}} \pi_{g+1} + \epsilon z_{g+1} \end{pmatrix}, \quad \mathbf{F}_d := \begin{pmatrix} \nabla_{z_1} d^{\mathrm{T}} \lambda \\ \vdots \\ \nabla_{z_{g+1}} d^{\mathrm{T}} \lambda \end{pmatrix}, \quad \mathbf{F}_z^\epsilon := \mathbf{F}_f^\epsilon - \mathbf{F}_d. \tag{18}$$

Furthermore, we define $z, z_i, l_i^\omega, l_i\, u_i, v_i, s_i, m_i$ and x_i as follows:

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_{g+1} \end{pmatrix}, \quad z_i = \begin{pmatrix} l_i \\ u_i \\ v_i \\ s_i \\ m_i \\ x_i \end{pmatrix}, \quad l_i^\omega = \begin{pmatrix} y_{i1}^\omega \\ \vdots \\ y_{ig}^\omega \\ r_i^\omega \end{pmatrix}, \quad l_i = \begin{pmatrix} l_i^1 \\ \vdots \\ l_i^n \end{pmatrix},$$

$$u_i = \begin{pmatrix} u_{i1}^1 \\ \vdots \\ u_{ig}^n \end{pmatrix}, \quad v_i = \begin{pmatrix} v_{i1}^1 \\ \vdots \\ v_{ig}^n \end{pmatrix}, \quad s_i = \begin{pmatrix} s_{i1}^1 \\ \vdots \\ s_{ig}^n \end{pmatrix}, \quad m_i = \begin{pmatrix} m_{i1} \\ \vdots \\ m_{ig} \end{pmatrix}, \quad x_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{ig} \end{pmatrix}.$$

Using these definitions, the Lipschitz continuity and strong monotonicity constants of $\mathbf{F}_\epsilon$ can be derived.

LEMMA 4.2 *Consider the mapping $\mathbf{F}_\epsilon(z, \lambda)$, defined in (18), arising from the Nash game. Suppose assumptions (A1)–(A2), (A3) hold and suppose the cost functions ζ_{ij}^ω are Lipschitz continuous with constants $L_\zeta^{ij,\omega}$, for all $i \in \mathcal{N}, j \in \mathcal{J}$ and for all $\omega \in \Omega$. Then this mapping is Lipschitz continuous and strongly monotone with constants L and ϵ, respectively, where*

$$L \triangleq (M + \|B\| + \epsilon), \quad M \triangleq \max_{i \in G} (2N_f^2 (b_i^0 + \mathbb{E}(b_i^\omega + \bar{L}_\zeta^{i,\omega}))),$$

and $\|B\| \le \sqrt{N_f N_g n}$.

Proof We first derive the Lipschitz constant for $\mathbf{F}_\epsilon$. This requires analysing each of the three terms.

$$\|\mathbf{F}^\epsilon(z^1, \lambda^1) - \mathbf{F}^\epsilon(z^2, \lambda^2)\| = \left\| \begin{pmatrix} \mathbf{F}_f^\epsilon(z^1, \lambda^1) - \mathbf{F}_f^\epsilon(z^2, \lambda^2) + \mathbf{F}_d(z^1, \lambda^1) - \mathbf{F}_d(z^2, \lambda^2) \\ \mathbf{F}_\lambda(z^1, \lambda^1) - \mathbf{F}_\lambda(z^2, \lambda^2) \end{pmatrix} \right\| \tag{19}$$

$$\leq \underbrace{\|\mathbf{F}_f^\epsilon(z^1, \lambda^1) - \mathbf{F}_f^\epsilon(z^2, \lambda^2)\|}_{\text{Term 1}} + \underbrace{\|\mathbf{F}_d(z^1, \lambda^1) - \mathbf{F}_d(z^2, \lambda^2)\|}_{\text{Term 2}}$$

$$+ \underbrace{\|\mathbf{F}_\lambda^\epsilon(z^1, \lambda^1) - \mathbf{F}_\lambda^\epsilon(z^2, \lambda^2)\|}_{\text{Term 3}}. \tag{20}$$

We bound each of the three terms as follows:

Term 1: Given two vectors z^1 and z^2, we may decompose $\mathbf{F}$ into $H + B$ allowing term 1 to be expressed as

$$\mathbf{F}_i(z^1) - \mathbf{F}_i(z^2) = \begin{pmatrix} F_i^l(z^1) - F_i^l(z^2) \\ F_i^s(z^1) - F_i^s(z^2) \\ F_i^m(z^1) - F_i^m(z^2) \\ F_i^x(z^1) - F_i^x(z^2) \end{pmatrix} = \underbrace{\begin{pmatrix} H_i^l(z^1) - H_i^l(z^2) \\ H_i^s(z^1) - H_i^s(z^2) \\ H_i^m(z^1) - H_i^m(z^2) \\ H_i^x(z^1) - H_i^x(z^2) \end{pmatrix}}_{\text{term 4}} + \underbrace{\begin{pmatrix} B_i^l(z^1) - B_i^l(z^2) \\ B_i^s(z^1) - B_i^s(z^2) \\ B_i^m(z^1) - B_i^m(z^2) \\ B_i^x(z^1) - B_i^x(z^2) \end{pmatrix}}_{\text{term 5}}.$$

Terms in l and x are non-zero in the specification of term 4 and the first of these for $i \in G$ is bounded as shown below.

$$(H_i^l(z^1) - H_i^l(z^2))_\omega = \begin{pmatrix} 2\rho^\omega b_i^\omega(y_{i1,\omega}^1 - y_{i1,\omega}^2) + (\zeta_{i1}^\omega(y_{i1,\omega}^1) - \zeta_{i1}^\omega(y_{i1,\omega}^2)) \\ \vdots \\ 2\rho^\omega b_i^\omega(y_{ig,\omega}^1 - y_{ig,\omega}^2) + (\zeta_{ig}^\omega(y_{ig,\omega}^1) - \zeta_{ig}^\omega(y_{ig,\omega}^2)) \\ \rho^\omega b_i^\omega(r_{i,\omega}^1 - r_{i,\omega}^2) \end{pmatrix}$$

$$\leq \begin{pmatrix} 2\rho^\omega b_i^\omega(y_{i1,\omega}^1 - y_{i1,\omega}^2) \\ \vdots \\ 2\rho^\omega b_i^\omega(y_{ig,\omega}^1 - y_{ig,\omega}^2) \\ \rho^\omega b_i^\omega(r_{i,\omega}^1 - r_{i,\omega}^2) \end{pmatrix} + \begin{pmatrix} \rho^\omega(\zeta_{i1}^\omega(y_{i1,\omega}^1) - \zeta_{i1}^\omega(y_{i1,\omega}^2)) \\ \vdots \\ \rho^\omega(\zeta_{ig}^\omega(y_{ig,\omega}^1) - \zeta_{ig}^\omega(y_{ig,\omega}^2)) \end{pmatrix}$$

$$\leq \begin{pmatrix} \rho^\omega(2b_i^\omega + L_\zeta^{i1,\omega})(y_{i1,\omega}^1 - y_{i1,\omega}^2) \\ \vdots \\ \rho^\omega(2b_i^\omega + L_\zeta^{ig,\omega})(y_{ig,\omega}^1 - y_{ig,\omega}^2) \\ \rho^\omega b_i^\omega(r_{i,\omega}^1 - r_{i,\omega}^2) \end{pmatrix} \leq M_i^{l,\omega} \begin{pmatrix} (y_{i1,\omega}^1 - y_{i1,\omega}^2) \\ \vdots \\ (y_{ig,\omega}^1 - y_{ig,\omega}^2) \\ (r_{i,\omega}^1 - r_{i,\omega}^2) \end{pmatrix},$$

where $M_i^{l,\omega} = \rho^\omega(2b_i^\omega + \max_{j \in \mathcal{J}} L_\zeta^{ij,\omega})$. Similarly, for $i \in G$, the other non-zero term in term 4 is bounded as follows:

$$H_i^x(z^1) - H_i^x(z^2) = M_i^x \begin{pmatrix} (x_{i1}^1 - x_{i1}^2) \\ \vdots \\ (x_{ig}^1 - x_{ig}^2) \end{pmatrix},$$

where $M_i^x = 2b_i^0$. By noting that when $i \in G^c$, $M_i^{l,\omega} = \rho^\omega b_i^\omega$, the Lipschitz constant for term 4, denoted by M_4, is given by

$$M_4 = \max_{i \in G \cup G^c} \left(\sum_{\omega \in \Omega} M_i^{l,\omega} + M_i^x \right) \leq \max_{i \in G}(2(\mathbb{E}b_i^\omega + b_i^0) + \mathbb{E}\bar{L}_\zeta^{i,\omega}), \quad \bar{L}_\zeta^{i,\omega} = \max_{j \in \mathcal{J}} L_\zeta^{ij,\omega}.$$

Similarly, for $i \in G$, the norms of the two non-zero terms in term 5, may be bounded through the use of the triangle inequality in the following fashion:

$$\|(B_i^l(z^1) - B_i^l(z^2))_\omega\|$$

$$= \left\| \begin{pmatrix} \rho^\omega \left(b_i^\omega \left(\sum_{j \in \mathcal{J}, j \neq 1} (y_{ij,\omega}^1 - y_{ij,\omega}^2) \right) + b_i^\omega(r_{i,\omega}^1 - r_{i,\omega}^2) - b_i^\omega(x_{i1}^1 - x_{i1}^2) \right) \\ \vdots \\ \rho^\omega \left(b_i^\omega \left(\sum_{j \in \mathcal{J}, j \neq g} (y_{ij,\omega}^1 - y_{ij,\omega}^2) \right) + b_i^\omega(r_{i,\omega}^1 - r_{i,\omega}^2) - b_i^\omega(x_{iJ}^1 - x_{iJ}^2) \right) \\ \rho^\omega \left(b_i^\omega \left(\sum_{j \in \mathcal{J}} (y_{ij,\omega}^1 - y_{ij,\omega}^2) \right) \right) \end{pmatrix} \right\|$$

$$\leq \bar{M}_i^{l,\omega} \|z^1 - z^2\|,$$

$\bar{M}_i^{l,\omega} = \rho^\omega b_i^\omega (1 + (g+1)^2)$ and

$$\|B_i^x(z^1) - B_i^x(z^2)\| = \left\| \begin{pmatrix} b_i^0 \sum_{j \in \mathcal{J}, j \neq 1} (x_{ij}^1 - x_{ij}^2) - \sum_{\omega \in \Omega} \rho^\omega b_i^\omega \left(\sum_{j \in \mathcal{J}} (y_{ij,\omega}^1 - y_{ij,\omega}^2) + r_{i,\omega}^1 - r_{i,\omega}^2 \right) \\ \vdots \\ b_i^0 \sum_{j \in \mathcal{J}, j \neq g} (x_{ij}^1 - x_{ij}^2) - \sum_{\omega \in \Omega} \rho^\omega b_i^\omega \left(\sum_{j \in \mathcal{J}} (y_{ij,\omega}^1 - y_{ij,\omega}^2) + r_{i,\omega}^1 - r_{i,\omega}^2 \right) \end{pmatrix} \right\|$$

$$\leq \bar{M}_i^x \|z^1 - z^2\|,$$

where $\bar{M}_i^x = g^2(b_i^0 + \mathbb{E}b_i^\omega)$. The corresponding constant for $i \in G^c$ is seen to be zero allowing us to define M_5, the Lipschitz constant for term 5, by

$$M_5 = \max_{i \in G \cup G^c} \left(\sum_{\omega \in \Omega} \bar{M}_i^{l,\omega} + \bar{M}_i^x \right) = \max_{i \in G} (2(g+1)^2(b_i^0 + \mathbb{E}b_i^\omega)).$$

If $N_f = (g+1)$, then the overall Lipschitz constant for term 1 is given by

$$M \triangleq \max_{i \in G} (2N_f^2(b_i^0 + \mathbb{E}(b_i^\omega + \bar{L}_\zeta^{i,\omega}))).$$

Term 2: Term 2 may be bounded as

$$\mathbf{F}_d(z^1, \lambda^1) - \mathbf{F}_d(z^2, \lambda^2)\| = \|\nabla d(z^1)^{\mathrm{T}}\lambda_1 - \nabla d(z^2)^{\mathrm{T}}\lambda^2\|$$

$$\leq \|\nabla d(z^1)^{\mathrm{T}}\lambda_1 - \nabla d(z^2)^{\mathrm{T}}\lambda^1\| + \|\nabla d(z_2)^{\mathrm{T}}(\lambda^1 - \lambda^2)\|$$

$$\leq \|\nabla d(z^1) - \nabla d(z^2)\|\|\lambda^1\| + \|\nabla d(z^2)\|\|\lambda^1 - \lambda^2\|,$$

where the inequalities follow from the application of the triangle inequality and the Cauchy–Schwartz inequality. Furthermore, $\nabla d(z)$ is a constant since $d(z)$ is a polyhedral constraint given

by $d(z) = Bz$ implying that $\|\nabla d(z^1) - \nabla d(z^2)\| = 0$, allowing us to conclude that

$$\|\mathbf{F}_d(z^1, \lambda^1) - \mathbf{F}_d(z^2, \lambda^2)\| \leq \|B\|\|\lambda^1 - \lambda^2\|.$$

Term 3: Term 3 may be bounded by recalling that $d(z)$ is polyhedral, allowing us to proceed as follows:

$$\|\mathbf{F}_\lambda(z_1, \lambda_1) - \mathbf{F}_\lambda(z_2, \lambda_2)\| \leq \|d(z^1) - d(z^2)\| + \epsilon\|\lambda^1 - \lambda^2\|$$
$$\leq \|B\|\|z^1 - z^2\| + \epsilon\|\lambda^1 - \lambda^2\|,$$

where the inequalities follow again from the triangle inequality, the Cauchy–Schwartz inequality and the functional form of $d(z)$. It follows that the Lipschitz constant for the overall mapping is given by $(M + \|B\| + \epsilon)$.

The strong monotonicity of the mapping $\mathbf{F}^\epsilon$ with monotonicity constant ϵ can be deduced by noting that $\nabla \mathbf{F}^\epsilon$, given by

$$\nabla \mathbf{F}^\epsilon = \begin{pmatrix} \nabla_z \mathbf{F}_z + \epsilon\mathbf{I} & -\nabla d^{\mathrm{T}} \\ \nabla d & \epsilon\mathbf{I} \end{pmatrix},$$

is positive definite since $\nabla_z \mathbf{F}_z$ is positive semidefinite for all z. ∎

4.3.1 *Primal–dual scheme*

When the mapping $\mathbf{F}_\epsilon(z, \lambda)$ is Lipschitz continuous and strongly monotone, the convergence of the primal–dual scheme can be claimed. Note that weaker conditions such as strict monotonicity can also be used to be guarantee convergence while mere monotonicity requires alternate schemes (such as two-step methods) (see [18, Chapter 12]).

PROPOSITION 4.3 (Convergence of primal–dual scheme [18]) *Consider the primal–dual scheme given by (11) and (12). Suppose that assumptions* (A1)–(A2), (A3) *hold. If the steplength $\gamma^{pd} \leq 2\epsilon/L^2$, then the sequence $\{(z^k, \lambda^k)\}$ converges to $(z^*_\epsilon, \lambda^*_\epsilon)$, an ϵ-Nash equilibrium of the game.*

4.3.2 *Exact and inexact dual schemes*

In this subsection, we consider the dual scheme both in its exact and inexact forms. While a proof for the convergence of the original dual scheme is provided in [35], we present a different argument in a regularized setting. Crucial to this result is the supporting requirement on co-coercivity of $d(z(\lambda))$. We provide a proof that uses the mapping $\mathbf{F}^\epsilon_z, \mathbf{F}^\epsilon_f$ and $\mathbf{F}_d$ as defined in (20), adapted from a result in [35]. It must be emphasized that the inexact dual has been studied recently by the second author in a multiuser optimization setting [36,37] and our results, while couched in a stochastic game-theoretic setting, are closely related. Yet, given that they have never been proved for equilibrium problems, we see the results here being of relevance. Furthermore, the polyhedral nature of $d(z)$ simplifies some of the proofs are often simpler and allows for somewhat different yet more refined bounds that relate the error directly to *problem size*.

LEMMA 4.4 *Consider the function $d(z(\lambda))$ where $z(\lambda)$ is a solution to the primal problem (8). Then $d(z(\lambda)) \equiv Bz$ is co-coercive with constant η_{cc} or*

$$(\lambda_2 - \lambda_1)^{\mathrm{T}}(d(z(\lambda_1)) - d(z(\lambda_2))) \geq \eta_{cc}\|d(z(\lambda_2)) - d(z(\lambda_1))\|^2 \quad \text{for all } \lambda_1, \lambda_2 \in \mathbb{R}^m_+,$$

where $\eta_{cc} = \epsilon/(N_f N_g n)$. Furthermore, we have

$$\|z(\lambda_1) - z(\lambda_2)\| \leq \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda_1 - \lambda_2\| \quad \text{for all } \lambda_1, \lambda_2 \in \mathbb{R}^m_+. \tag{21}$$

Proof Let $z_1 \equiv z(\lambda_1)$ and $z_2 \equiv z(\lambda_2)$ represent solutions to $\text{VI}(\mathbf{Z}, \mathbf{F}_z^\epsilon(z; \lambda_1))$ and $\text{VI}(\mathbf{Z}, \mathbf{F}_z^\epsilon(z; \lambda_2))$, respectively. Then, we have

$$(z_2 - z_1)^{\text{T}} \mathbf{F}_z^\epsilon(z_1, \lambda_1) \geq 0 \quad \text{and} \quad (z_1 - z_2)^{\text{T}} \mathbf{F}_z^\epsilon(z_2, \lambda_2) \geq 0.$$

By recalling from (10), we have that

$$(z_2 - z_1)^{\text{T}}(\mathbf{F}_d(z_1, \lambda_1) - \mathbf{F}_d(z_2, \lambda_2)) \geq (z_2 - z_1)^{\text{T}}(\mathbf{F}_f^\epsilon(z_2, \lambda_2) - \mathbf{F}_d^\epsilon(z_1, \lambda_1))$$
$$\geq \epsilon \|z_2 - z_1\|^2, \tag{22}$$

where the second inequality follows from the strong monotonicity of $\mathbf{F}_f^\epsilon$ with constant ϵ. It follows from the definition of $d(z)$ that

$$(z_2 - z_1)^{\text{T}}(\mathbf{F}_d(z_1, \lambda_1) - \mathbf{F}_d(z_2, \lambda_2)) = (z_2 - z_1)^{\text{T}}(-B^{\text{T}}\lambda_1 + B^{\text{T}}\lambda_2)$$
$$= (Bz_2 - Bz_1)^{\text{T}}(-\lambda_1 + \lambda_2) \geq \epsilon \|z_2 - z_1\|^2$$
$$\geq \frac{\epsilon}{\|B\|^2} \|d(z_1) - d(z_2)\|^2,$$

where the last two inequalities follow from (22) and the Lipschitz continuity of $d(z)$ with constant $\|B\|$. Finally by applying the Cauchy–Schwartz inequality to the first inequality above, the second result (21) may be obtained as follows:

$$\|z_2 - z_1\|^2 \leq \frac{1}{\epsilon}(d(z_2) - d(z_1))^{\text{T}}(\lambda_2 - \lambda_1) \leq \frac{1}{\epsilon}\|B\|\|z_2 - z_1\|\|\lambda_2 - \lambda_1\|$$

giving us

$$\|z_2 - z_1\| \leq \frac{\|B\|}{\epsilon}\|\lambda_2 - \lambda_1\| \leq \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda_2 - \lambda_1\|. \qquad \blacksquare$$

Using the co-coercivity of $d(z(\lambda))$, the convergence of the iterates constructed from regularized dual scheme can be shown to converge to λ_ϵ^*, a dual solution to the regularized problem.

PROPOSITION 4.5 (Convergence of exact dual scheme) *Consider the dual scheme given by (13) and (14). If $d(z(\lambda))$ is co-coercive with constant $\eta_{cc} = \epsilon/N_f N_g n$ and γ_d satisfies*

$$\gamma_d < \frac{2\epsilon}{2\epsilon^2 + N_f N_g n}, \tag{23}$$

then $\|\lambda^{k+1} - \lambda_\epsilon^*\| \leq q_d^k \|\lambda^0 - \lambda_\epsilon^*\|$ where $q_d = (1 - \gamma_d \epsilon)$.

Proof By invoking the definition of λ^{k+1}, noting that λ^* is a fixed-point of (9) and the non-expansivity of the Euclidean projector, we have

$$\|\lambda^{k+1} - \lambda_\epsilon^*\| = \|\Pi_{\mathbb{R}_m^+}(\lambda^k - \gamma_d d(z^k) - \gamma_d \epsilon \lambda^k) - \lambda_\epsilon^*\|$$
$$= \|\Pi_{\mathbb{R}_m^+}(\lambda^k - \gamma_d d(z^k) - \gamma_d \epsilon \lambda^k) - \Pi_{\mathbb{R}_m^+}(\lambda_\epsilon^* - \gamma_d d(z_\epsilon^*) - \gamma_d \epsilon \lambda_\epsilon^*)\|$$
$$\leq \|(\lambda^k - \gamma_d d(z^k) - \gamma_d \epsilon \lambda^k) - (\lambda_\epsilon^* - \gamma_d d(z_\epsilon^*) - \gamma_d \epsilon \lambda^*)\|$$
$$= \|(1 - \gamma_d \epsilon)(\lambda^k - \lambda_\epsilon^*) - \gamma_d(d(z^k) - d(z_\epsilon^*))\|.$$

Then, by expanding the square of the expression on the right-hand side and by leveraging the co-coercivity of $d(\lambda(z))$ with respect to z, we have the following inequality:

$$\|\lambda^{k+1} - \lambda_\epsilon^*\|^2 \leq (1 - \gamma_d\epsilon)^2 \|\lambda^k - \lambda_\epsilon^*\|^2$$
$$+ (\gamma_d)^2 \|d(z^k) - d(z_\epsilon^*)\|^2 - 2\gamma_d(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z^k) - d(z_\epsilon^*))$$
$$\leq (1 - \gamma_d\epsilon)^2 \|\lambda^k - \lambda_\epsilon^*\|^2 + (\gamma_d^2 - 2\gamma_d\eta_{cc}(1 - \gamma_d\epsilon))\|d(z^k) - d(z_\epsilon^*)\|^2,$$

where the second inequality follows from the co-coercivity of $d(z(\lambda))$ with a constant η_{cc}. Convergence of the scheme follows if γ_d is chosen in accordance with

$$\gamma_d < \min\left\{\frac{1}{\epsilon}, \frac{2\eta_{cc}}{1 + 2\eta_{cc}\epsilon}\right\} \quad \text{where } \eta_{cc} = \frac{\epsilon}{N_f N_g n}.$$

But we have

$$\frac{2\eta_{cc}}{1 + 2\eta_{cc}\epsilon} = \frac{1}{N_f N_g n/2\epsilon + \epsilon} < \frac{1}{\epsilon} \text{implying that } \gamma_d < \frac{2\epsilon}{2\epsilon^2 + N_f N_g n}. \qquad \blacksquare$$

The convergence of λ^k to λ_ϵ^* allows for deriving similar statements for z^k and the infeasibility, namely $\max(0, d(z^k))$.

LEMMA 4.6 *Consider the dual scheme given by (13) and (14) and suppose $d(z(\lambda))$ is co-coercive with constant $\epsilon/\|B\|^2$. Then, for any $k \geq 0$, we have*

$$\|z^k - z^*\| \leq \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\| \quad \text{and} \quad \max(0, -d(z^k)) \leq \frac{N_f N_g n}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\|.$$

Proof A bound on the suboptimality may be directly obtained from Lemma 4.4. The infeasibility in the constraint $d(z) \geq 0$, namely $\max(0, -d(z))$, is bounded as shown through the following sequence of relationships, that use the Cauchy–Schwartz inequality and the bound on the suboptimality of z^k:

$$\max(0, -d(z^k)) \leq -Bz^k = -B(z^k + z_\epsilon^* - z_\epsilon^*)$$
$$\leq B(z_\epsilon^* - z^k)$$
$$\leq \|B\|\|z_\epsilon^* - z^k\| \leq \frac{N_f N_g n}{\epsilon}\|\lambda_\epsilon^* - \lambda^k\|. \qquad \blacksquare$$

A shortcoming of the dual scheme is the need for exact primal solutions for every dual solution. Since this requires iteratively solving a fixed-point problem, it can prove to be an inordinately expensive component of the algorithm. Our intent is in constructing a *bounded complexity* variant that requires that only K iterations of the primal scheme be made for a given value of the dual iterates. This is given by

$$z_j^{t+1} = \Pi_{\mathbf{Z}_j}(z_j^t - \gamma_d(\mathbf{F}_z(z_j^t; z_{-j}^t, \lambda^k) + \epsilon^\ell z_j^t)) \quad \text{for all } j, t = 0, \ldots, K-1. \tag{24}$$

However, in obtaining error bounds, we require that the primal strategy sets be bounded. It is worth remarking that in general this bound may be difficult to obtain in closed-form but we assume that such a bound is available for purposes of this analysis. In the current setting, one avenue for deriving such a bound would be through imposing a bound on forward positions. In the remainder of this section, we assume that $\|z\| \leq M_z$ throughout the remainder of this section. Finally, the strong

monotonicity of the primal problem implies that $\|z^t - z^*\| \le q_p^{t/2}\|z^0 - z^*\|$, where $q_p = 2\epsilon/M^2 < 1$ where M is the Lipschitz constant of the primal mapping $\mathbf{F}_f(z)$, as specified in Lemma 4.2.

PROPOSITION 4.7 (Error bounds for inexact-dual scheme) *Consider the inexact dual scheme given by (24) and (14). If $d(z(\lambda))$ is co-coercive with constant $\epsilon/\|B\|^2$, $\|z\| \le M_z$ and γ_d satisfies*

$$\gamma_d < \frac{2\epsilon}{2\epsilon^2 + N_f N_g n},$$

then we have

$$\|\lambda^k - \lambda_\epsilon^*\| \le q_d^k \|\lambda^0 - \lambda_\epsilon^*\|^k$$
$$+ \left(\frac{1 - q_d^k}{1 - q_d}\right)\left(\left(\frac{2}{\epsilon^2} + 4\right)(N_f N_g n)^{1/2} q_p^{K/2} M_z^2 (1 + (N_f N_g n)^{1/2} q_p^{K/2})\right).$$

Proof As earlier, the definition of λ^{k+1} and the fixed-point property of λ_ϵ^*, we have the following inequality:

$$\|\lambda^{k+1} - \lambda_\epsilon^*\| = \|\Pi_{\mathbb{R}_m^+}(\lambda^k - \gamma_d(d(z_K^k) + \epsilon_k\lambda^k)) - \Pi_{\mathbb{R}_m^+}(\lambda_\epsilon^* - \gamma_d(d(z^*) + \epsilon_k\lambda_\epsilon^*))\|$$
$$\le \|(\lambda^k - \gamma_d(d(z_K^k) + \epsilon_k\lambda^k)) - (\lambda_\epsilon^* - \gamma_d(d(z^*) + \epsilon_k\lambda_\epsilon^*))\|.$$

By adding and subtracting terms and by using the triangle inequality, the right-hand side can be shown to be

$$\|(\lambda^k - \gamma_d(d(z_K^k) + \epsilon\lambda^k)) - (\lambda^* - \gamma_d(d(z^*) + \epsilon_k\lambda_\epsilon^*))\|^2$$
$$= \|(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*) - \gamma_d(d(z_K^k) - d(z_\epsilon^*))\|^2$$
$$= (1 - \gamma_d\epsilon)^2\|\lambda^k - \lambda_\epsilon^*\|^2 + \gamma_d^2 \underbrace{\|d(z_K^k) - d(z_\epsilon^*)\|^2}_{\text{term 1}} \underbrace{-2\gamma_d(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z_K^k) - d(z_\epsilon^*))}_{\text{term 2}}.$$

By noting that $d(z_k)$ is given by $Bz_k \ge 0$ for some matrix B, it follows that term 1 can be bounded by

$$\|d(z_K^k) - d(z_\epsilon^*)\|^2 \le \|d(z_K^k) - d(z^k)\|^2 + \|d(z^k) - d(z_\epsilon^*)\|^2$$
$$+ 2\|d(z^k) - d(z_K^k)\|\|d(z^k) - d(z_\epsilon^*)\|.$$

Furthermore, by using the co-coercivity of $d(x(\lambda))$, term 2 may be bounded in the following fashion:

$$-2\gamma_d(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z_K^k) - d(z_\epsilon^*))$$
$$= -2\gamma_d(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z^k) - d(z_\epsilon^*)) - 2\gamma_d(1 - \gamma_d\epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z_K^k) - d(z^k))$$
$$\le -2\gamma_d(1 - \gamma_d\epsilon)\frac{\epsilon}{\|B\|^2}\|d(z^k) - d(z_\epsilon^*)\|^2 + \gamma_d^2\|\lambda^k - \lambda_\epsilon^*\|^2 + (1 - \gamma_d\epsilon)^2\|d(z_K^k) - d(z^k)\|^2.$$

Using the bounds on terms 1 and 2, we have the following:

$$(1 - \gamma_d \epsilon)^2 \|\lambda^k - \lambda_\epsilon^*\|^2 + \gamma_d^2 \|d(z_K^k) - d(z_\epsilon^*)\|^2 - 2\gamma_d (1 - \gamma_d \epsilon)(\lambda^k - \lambda_\epsilon^*)^{\mathrm{T}}(d(z_K^k) - d(z_\epsilon^*))$$

$$\leq (1 - \gamma_d \epsilon)^2 \|\lambda^k - \lambda_\epsilon^*\|^2 + \gamma_d^2 \left(\|d(z_K^k) - d(z^k)\|^2 + \|d(z^k) - d(z_\epsilon^*)\|^2 \right)$$

$$+ \gamma_d^2 \left(2\|d(z^k) - d(z_K^k)\| \|d(z^k) - d(z_\epsilon^*)\| \right) - 2\gamma_d (1 - \gamma_d \epsilon) \frac{\epsilon}{\|B\|^2} \|d(z^k) - d(z_\epsilon^*)\|^2$$

$$+ \gamma_d^2 \|\lambda^k - \lambda_\epsilon^*\|^2 + (1 - \gamma_d \epsilon)^2 \|d(z_K^k) - d(z^k)\|^2$$

$$= \underbrace{\left((1 - \gamma_d \epsilon)^2 + \gamma_d^2 \right) \|\lambda^k - \lambda_\epsilon^*\|^2 + \left(\gamma_d^2 - 2\gamma_d \frac{\epsilon}{\|B\|^2}(1 - \gamma_d \epsilon) \right) \|d(z_k) - d(z_\epsilon^*)\|^2}_{\text{term 3}}$$

$$+ \underbrace{(\gamma_d^2 + (1 - \gamma_d \epsilon)^2)\|d(z_K^k) - d(z^k)\|^2 + 2\gamma_d^2 \|d(z^k) - d(z_K^k)\| \|d(z^k) - d(z_\epsilon^*)\|}_{\text{term 4}}.$$

If γ_d is chosen in accordance with

$$((1 - \gamma_d \epsilon)^2 + \gamma_d^2) < 1, \quad \gamma_d < \frac{1 + \epsilon^2}{2\epsilon}$$
$$(\gamma_d^2 - 2\gamma_d \eta_{cc}(1 - \gamma_d \epsilon)) < 0, \quad \gamma_d < \frac{2\eta_{cc}}{1 + 2\eta_{cc}\epsilon}, \qquad \implies \gamma_d < \min \left(\frac{1 + \epsilon^2}{2\epsilon}, \frac{2\eta_{cc}}{1 + 2\eta_{cc}\epsilon} \right).$$

then term 3 would lead to a contraction. However, it can be seen that

$$\frac{2\eta_{cc}}{1 + 2\eta_{cc}\epsilon} = \frac{1}{N_f N_g n/2\epsilon + \epsilon} < \frac{1}{2\epsilon} < \frac{1 + \epsilon^2}{2\epsilon},$$

if $N_f N_g n/2\epsilon > \epsilon$ or $N_f N_g n > 2\epsilon^2$. It suffices that

$$\gamma_d < \frac{2\epsilon}{2\epsilon^2 + N_f N_g n}.$$

Note that the error arising from term 4 may be bounded by recalling that $d(z) = Bz$ is a Lipschitz continuous mapping implying that

$$(\gamma_d^2 + (1 - \gamma_d \epsilon)^2)\|d(z_K^k) - d(z^k)\|^2 + 2\gamma_d^2 \|d(z^k) - d(z_K^k)\| \|d(z^k) - d(z_\epsilon^*)\|$$

$$\leq (\gamma_d^2 + (1 - \gamma_d \epsilon)^2)\|B\|^2 \|z_K^k - z^k\|^2 + 2\gamma_d^2 \|B\| \|z^k - z_K^k\| M_z.$$

Then by observing that $\|z^k - z_K^k\| \leq \|z^k - z_0^k\| q_p^{K/2} \leq M_z q_p^{K/2}$, where the first inequality follows from geometric convergence of the sequence $\{z_K^k\}$ to z^k as $K \longrightarrow \infty$ and the second follows from the boundedness of the primal space with bound M_z. It follows that

$$(\gamma_d^2 + (1 - \gamma_d \epsilon)^2)\|B\|^2 \|z_K^k - z^k\|^2 + 2\gamma_d^2 \|B\| \|z^k - z_K^k\| M_z$$

$$\leq (\gamma_d^2 + (1 - \gamma_d \epsilon)^2)\|B\| q_p^K M_z^2 + 2\gamma_d^2 \|B\|^2 q_p^{K/2} M_z^2.$$

Finally, by observing that $(\gamma_d^2 + (1 - \gamma_d\epsilon)^2) \leq (\gamma_d^2 + (1 + \gamma_d\epsilon)^2)$ which is further bounded by $(1/\epsilon^2 + 4)$ and $\gamma_d^2 \leq 1/\epsilon^2$, we have

$$(\gamma_d^2 + (1 - \gamma_d\epsilon)^2)\|B\|q_p^K M_z^2 + 2\gamma_d^2\|B\|^2 q_p^{K/2} M_z^2$$

$$\leq \left(\frac{1}{\epsilon^2} + 4\right)\|B\|q_p^K M_z^2 + \frac{2}{\epsilon^2}\|B\|q_p^{K/2} M_z^2$$

$$\leq \left(\frac{2}{\epsilon^2} + 4\right)\|B\|q_p^{K/2} M_z^2 (1 + \|B\|q_p^{K/2})$$

$$\leq \left(\frac{2}{\epsilon^2} + 4\right)(N_f N_g n)^{1/2} q_p^{K/2} M_z^2 (1 + (N_f N_g n)^{1/2} q_p^{K/2}).$$

Then given a starting point λ^0, we have

$$\|\lambda^k - \lambda_\epsilon^*\| \leq q_d^k \|\lambda^0 - \lambda_\epsilon^*\|^k$$

$$+ \underbrace{\left(\frac{1 - q_d^k}{1 - q_d}\right)\left(\left(\frac{2}{\epsilon^2} + 4\right)(N_f N_g n)^{1/2} q_p^{K/2} M_z^2 (1 + (N_f N_g n)^{1/2} q_p^{K/2})\right)}_{\text{Error from inexact solution of primal}}. \qquad \blacksquare$$

It can be seen that the error term arising from inexact primal solutions converges to zero as $K \longrightarrow \infty$. We conclude this section with a bound on the suboptimality of z^k and infeasibility associated with $d(z^k)$ if the dual scheme terminates prematurely.

LEMMA 4.8   *Consider the inexact dual scheme given by (24) and (14). If $d(z(\lambda))$ is co-coercive with constant $\epsilon/\|B\|^2$, $\|z\| \leq M_z$ and γ_d satisfies*

$$\gamma_d < \frac{2\epsilon}{2\epsilon^2 + N_f N_g n}.$$

Then, for any non-negative integers $k, K \geq 0$, we have

$$\|z_K^k - z_\epsilon^*\| \leq q_p^{K/2} M_z + \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\|,$$

$$\max(0, -d(z_K^k)) \leq \sqrt{N_f N_g n}\left(q_p^{K/2} M_z + \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\|\right).$$

Proof   The first result follows easily by using the triangle inequality and employing the earlier result.

$$\|z_K^k - z_\epsilon^*\| \leq \|z_K^k - z^k\| + \|z^k - z_\epsilon^*\|$$

$$\leq q_p^{K/2} M_z + \frac{\sqrt{N_f N_g n}}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\|.$$

Similarly, the bound on the infeasibility at a point z_K^k is provided by adding and subtracting $d(z_\epsilon^*)$, applying the triangle and Cauchy–Schwartz inequality:

$$\max(0, -d(z_K^k)) \leq -d(z_K^k) = -d(z_K^k) + d(z_\epsilon^*) - d(z_\epsilon^*)$$

$$\leq -B(z_K^k - z_\epsilon^*) \leq \|B\|\|z_K^k - z^k\| + \|B\|\|z^k - z_\epsilon^*\|$$

$$\leq (N_f N_g n)^{1/2} q_p^{K/2} M_z + \frac{N_f N_g n}{\epsilon}\|\lambda^k - \lambda_\epsilon^*\|. \qquad \blacksquare$$

4.3.2.1 Summary of mathematical findings: We conclude this subsection with a review of the main results of the findings from Sections 3 and 4:

(1) *Existence and ϵ-uniqueness*: In Section 3, the focus lies on developing statements of existence (Proposition 3.5) and ϵ-uniqueness (Proposition 3.6), the latter being a consequence of the monotonicity of the associated mapping.
(2) *Cutting-plane projection scheme*: In Section 4, we consider a projection-based scheme that is naturally distributed across firms. However, each projection step reduces to stochastic quadratic programme that direct solvers find challenging to solve since its size grows with $|\Omega|$. Instead, a cutting-plane scheme is presented that scales slowly in complexity with $|\Omega|$.
(3) *Error analysis*: Finally, in Section 4, we examine three variants of a standard projection scheme, a primal–dual scheme, an exact dual scheme and an inexact dual scheme in the context of a regularized problem. While convergence of the first of these follows immediately, Proposition 4.5 proves the convergence of the exact dual scheme while error bounds are provided in Lemma 4.6. However, computing the exact dual scheme relies on an exact solution of the primal solution, often a computationally burdensome requirement. One may choose to use a bounded complexity variant where a fixed number of gradient steps is employed for computing an inexact primal solution. In this context, Proposition 4.7 and Lemma 4.8 derive analogous error bounds that are directly tied to the number of gradient steps in the primal.

4.4 *Numerical performance*

In this section, we examine the performance of our hybrid projection-based cutting-plane scheme with a focus on several questions. First, we consider whether the scheme scales with $|\Omega|$, $|\mathcal{J}|$ and $|\mathcal{N}|$. Second, we examine the relative performance of the primal–dual versus the dual scheme. Finally, we examine the benefits arising from inexact solutions of the primal problem.

We confine our discussion to the game and examine the behaviour of the scheme on a regularized game with $\epsilon = 1 \times 10^{-3}$. In our computational results, we define the loss function to be of the form: $\rho_{ij}^\omega = \chi (x_{ij} - \text{cap}_{ij}^\omega)^+$, where $\chi = 0.5$. Therefore, in addition to the earlier set of constraints we have another constraint stating that $s_{ij}^\omega \geq -m_{ij}$. Furthermore, we maintain χ to be the same across all agents. The risk aversion parameters are assumed to be 0.5 for all the agents unless specified otherwise. The nodal demand function intercepts were taken to be 150 and 200 for the spot and forward markets, respectively, across all nodes while the slopes of the spot-market price functions are specified to be normally distributed as per $N(1, 0.02)$ in the forward and spot markets. The algorithm was implemented in Matlab 7.0 on a Linux OS with a processor with a clockspeed of 2.39 GHz and a memory of 16 GB.

4.4.1 *Scalability*

The algorithm is implemented in a distributed fashion with each agent solving his projection problem independently. As a consequence, we expect that the effort should scale with the number of agents. When the number of firms is raised from 2 to 11, the variation of serial and parallel times are shown in Figure 3. Note that the parallel time is computed assuming that there are as many processors as there are agents. The variation in the number of overall projection steps with increase in the number of firms is also shown in Figure 3. The projection scheme is terminated when $\epsilon^{\text{inner}} = 5 \times 10^{-3}$. Both graphs show that the effort, both in terms of CPU time and projection steps, grows slowly with the number of firms.

If an analogous question is studied when the number of generating nodes is varied, we observe similar results, as shown in Figure 4. Note that the nodal problems decompose by the firm level implying that large networks, while computationally expensive, will not lead to rapid growth in
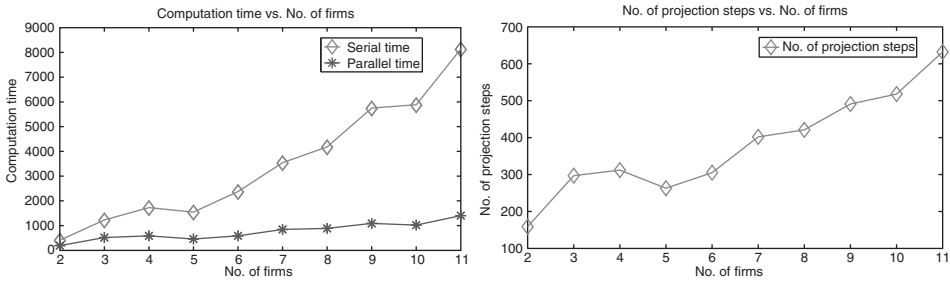
Figure 3.   Scalability of effort with number of firms.
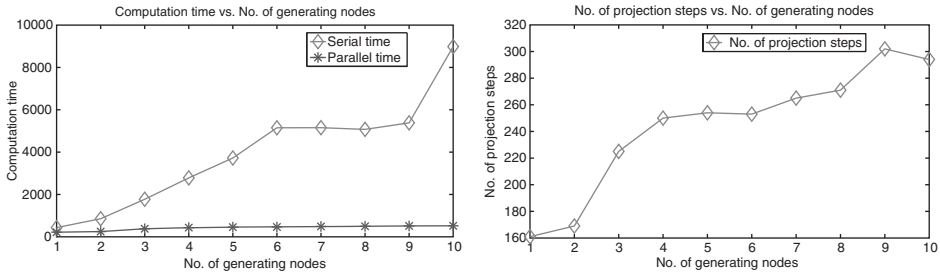
Figure 4.   Scalability of effort with number of generating nodes.
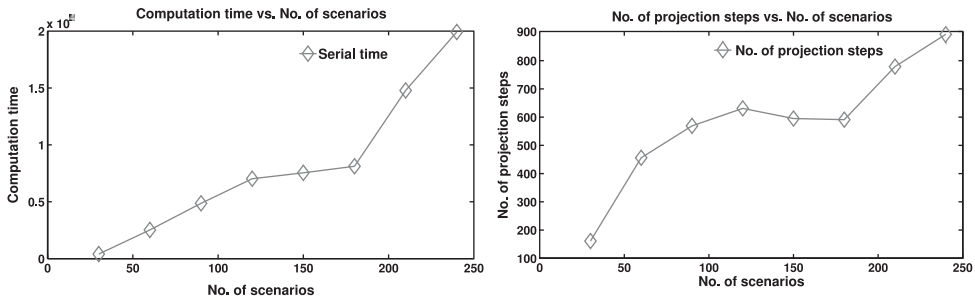
Figure 5.   Scalability of effort with sample size.

effort. Instead, such settings will necessitate the solution of a larger number of separable nodal problems.

Perhaps the most challenging source of complexity arises from the scenario-based approach for capturing uncertainty. This leads to arbitrarily large projection problems which are addressed through a cutting-plane method. If the number of scenarios is increased from 30 to 240, then the variation of serial times is as seen in Figure 5. Additionally, the variation in the number of overall projection steps is also shown in Figure 5. Finally, it is observed that the effort grows slowly with an increase in the size of the sample space, suggesting that the decomposition schemes are indeed scalable.

It is worth noting that direct solvers are poorly suited for addressing such problems. Consider the application of PATH to this class of problems. Table 1 reports the time taken by the PATH solver for seventeen problems [9]. A case with three generating firms and five generating nodes was considered and the number of scenarios were varied from $n = 10$ to $n = 200$. The forward and spot intercepts were taken to be 1400 and 1200 respectively. The capacities were taken to follow

Table 1. PATH solver
(CPU time).

n	Time (s)
10	0.23
20	0.73
30	2.13
40	3.76
50	10.98
60	15.48
70	21.00
80	33.26
90	39.21
100	49.23
110	64.51
120	69.99
130	83.68
140	99.02
150	129.52
160	138.64
170	*
180	*

$N(400, 2)$ for all generators across all nodes. The linear and quadratic generation costs were taken to be $N(12, 1)$ and $N(0.3, 0.01)$ respectively for the first two generators. The third generator was assumed to have no costs of generation. The linear and quadratic penalties for deviation were taken to be 0.1 for all generators across all nodes. Finally, failure of PATH is denoted by * and indicates that PATH does not return a solution. We observe that beyond 160 scenarios, PATH appears to have difficulty loading the problem into memory and provides further incentive for the development of decomposition-based schemes.

4.4.2 *Comparison between primal-dual and dual schemes*

A two firm problem, under the setting of one generating node was taken as a case study to compare the primal-dual and inexact dual schemes. The primal and dual step lengths were taken to be 2 for all the cases. Different instances of the above problem were solved by varying the demand and generation capacities. Instances 1 to 6 represent increasing values of (a, a^0) from $(150, 200)$ to $(400, 450)$ respectively in steps of 50. Generation capacities were correspondingly increased from $(N(100, 0.5))$ to $(N(162.7, 0.5))$ in steps of 12.7. The above set of problems was solved for ten, fifteen, twenty and twenty-five scenarios. Table 2 shows the number of iterations and the time taken to solve each problem by means of the primal dual and inexact dual methods. In the case of inexact dual methods, we show results for one, five and nine inexact primal steps. It can be seen that the primal-dual schemes tend to be more efficient than dual scheme while fewer inner primal steps are generally advisable in the context of dual schemes.

4.4.3 *CVaR measures under general sample spaces*

In this paper, we assume that the sample space is finite and comprises of the scenarios provided. In the numerical results, these scenarios are generated from prescribed distributions. In such a regime, the CVaR can be exactly computed. If, however, one considers a general measure space, the obtained solutions to the scenario-based problem are merely estimators. Sample-average approximation techniques have been studied extensively with two explicit goals: (i) Almost-sure convergence of the associated estimators to the true object and (ii) rate of convergence

Table 2. Comparison: Primal dual and inexact dual algorithms.

		Primal–dual		Dual-1			Dual-5			Dual-9		
n	Inst.	Steps	Time (s)	Dual steps	Total	Time (s)	Dual steps	Total	Time (s)	Dual steps	Total	Time (s)
10	1	53	49.34	53	53	50.50	14	70	79.80	9	81	98.18
	2	50	45.12	50	50	46.18	12	60	53.07	9	81	81.54
	3	50	47.72	50	50	48.64	12	60	46.94	8	72	57.00
	4	49	55.92	49	49	57.06	12	60	58.00	9	81	89.13
	5	50	54.36	50	50	55.56	14	70	81.56	10	90	110.53
	6	50	48.76	50	50	49.71	12	60	53.60	8	72	63.84
15	1	81	88.24	81	81	89.95	19	95	105.49	12	108	129.69
	2	76	101.36	76	76	103.00	18	90	121.50	11	99	127.95
	3	74	95.94	74	74	97.91	18	90	110.55	13	117	146.83
	4	74	103.84	74	74	105.47	18	90	122.76	11	99	127.25
	5	75	120.41	75	75	122.85	18	90	133.81	12	108	171.43
	6	75	122.29	75	75	124.70	18	90	149.65	11	99	149.65
20	1	108	180.32	108	108	183.22	25	125	207.85	16	144	249.98
	2	99	202.71	99	99	206.86	24	120	252.98	15	135	282.00
	3	97	173.93	97	97	176.48	24	120	220.43	14	126	209.15
	4	99	191.78	99	99	194.93	24	120	208.75	14	126	208.99
	5	98	183.08	98	98	186.31	24	120	212.35	14	126	202.91
	6	100	191.86	100	100	195.46	25	125	227.13	17	153	244.62
25	1	134	328.33	134	134	337.08	32	160	391.47	19	171	420.41
	2	122	289.25	122	122	294.83	29	145	337.35	18	162	388.56
	3	122	302.28	122	122	310.46	29	145	331.56	22	198	488.53
	4	120	289.69	120	120	294.15	29	145	343.48	21	189	381.28
	5	121	275.68	121	121	280.89	29	145	307.61	19	171	364.32
	6	121	272.19	121	121	274.31	29	145	309.52	19	171	366.11

analysis of the associated estimators. The presence of the CVaR measure makes the sample-average problem non-smooth and complicates the convergence analysis of the estimators. Some of these questions have been examined in the context of optimizing CVaR measures [12,44] but less so in game-theoretic regimes.

5. Insights for market design and operation

In this section, we provide some insights for market design and operations by examining the strategic behaviour of agents in the setting of a 53-node network, referred to as the Belgian grid and shown in Figure A1 in the appendix. This network has provided the basis for prior studies [62,63] and the impedances and capacities along all the transmission lines are listed in Table A2. We assume that nodes 7, 9, 10, 11, 14, 22 and 24 house generation facilities. We assume that the generation mix at each of these nodes is identical and is specified by Table 3 (total across all nodes is shown). Note that in practice, *truncated normal distributions* are employed throughout; specifically, the left tail of the distribution is truncated beyond 1. Here, the generation capacities and costs are assumed to be normally distributed across 30 scenarios ($n = 30$). Demand at all the nodes is articulated through affine functions. In the forward-clearing model, the intercepts in the forward and spot markets are taken to be fixed at 1500 at all nodes while the slopes in the spot market are assumed to vary normally with a mean of 1 and a standard deviation of 0.02. The parameter τ_j is taken to be 0.9 for all the firms and $\chi = 40$.

Our intent lies in ascertaining the relationship of a variety of parameters, such as risk-aversion, uncertainty and demand levels, on market outcomes such as forward market participation and penetration levels of wind resources. The complementarity problems are solved via KNITRO [8].[10] The detailed formulation of the complementarity problems can be found in the appendix.

Table 3. Generator details.

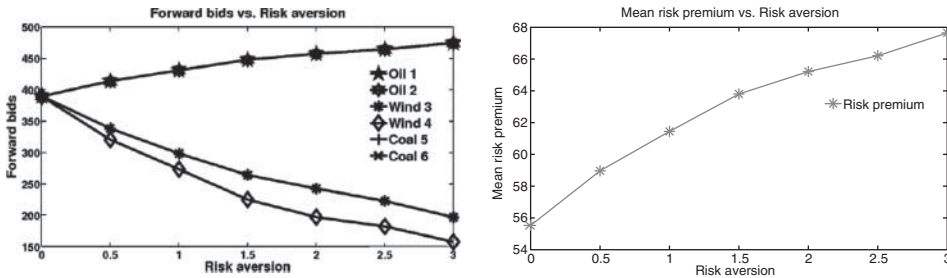| Generator type	Capacity	Linear cost	Quadratic cost
Oil 1	$N(2000, 10)$	$N(10, 1)$	$N(0.3, 0.01)$
Oil 2	$N(2000, 10)$	$N(10, 1)$	$N(0.3, 0.01)$
Wind 3	$N(650, 270)$	$N(0, 0)$	$N(0, 0)$
Wind 4	$N(730, 320)$	$N(0, 0)$	$N(0, 0)$
Coal 5	$N(1400, 10)$	$N(12, 1)$	$N(0.25, 0.01)$
Coal 6	$N(1400, 10)$	$N(12, 1)$	$N(0.25, 0.01)$

Figure 6. Impacts of increasing risk aversion.

5.1 *Risk aversion*

In this setting, we vary risk aversion parameter κ_i for all the firms from 0 to 3 in steps of 0.5. This can be viewed as an effect of higher risk-based penalties imposed by the system operator. As shown in Figure 6, we find that the forward bids drop for the wind generators and increase for the coal and oil generators. This behaviour suggests that as generators become risk-averse, firms with a larger number of wind-based assets tend to be conservative in forward market bidding. This is primarily because firms with uncertain generation face much higher risk of shortfall. As they are penalized higher amounts for exposing the market to such risk, firms tend to bid lower, reducing their risk exposure. This is manifested through lower participation in the forward market by wind-based generators. In a prisoner's dilemma-type effect, generators exposed to less risk tend to increase their positions in the forward market. Figure 6 also shows that the excess of forward price over expected spot price (risk premium) increases with risk aversion. This is expected as total forward participation reduces, thereby raising forward prices, and leading to higher premiums.

5.2 *Uncertainty in generation capacity*

Under the assumption that firms are assumed to have a constant risk aversion (fixed at 1 for all firms), we examine the relationship between uncertainty in capacity and risk exposure and level of forward participation. While coal and oil generators are expected to be close to deterministic in their availability, we assume that wind generators are faced with far greater uncertainty. In our numerical experiments, we vary the standard deviation of the wind generators (Wind 3 and Wind 4) from 10 to 885 in steps of 175. Expectedly, the risk exposure increases as the variability in wind assets grows (Figure 7). Moreover, while the general belief would be that participation in the forward markets would aid in hedging spot-market uncertainty, when risk-based penalties are introduced, we observe that wind-based generators are less inclined to participate. It should be emphasized that the deviation costs tend to have a similar impact on behaviour. Note that drops in
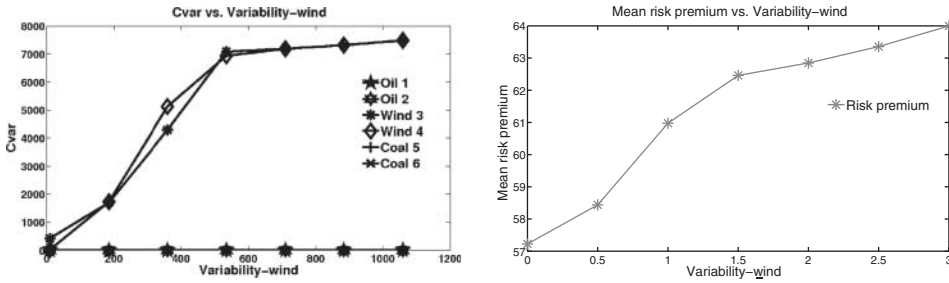
Figure 7. Impact of increasing uncertainty in wind-based capacity.

Table 4. System risk, nodal forward and spot prices with varying uncertainty.

Uncertainty θ	System risk (total CVaR)	Node 7			Node 22		
		p_i^0	$\mathbb{E}p_i^\omega$	$p_i^0 - \mathbb{E}p_i^\omega$	p_i^0	$\mathbb{E}p_i^\omega$	$p_i^0 - \mathbb{E}p_i^\omega$
2^0	11381.38	1381.17	1412.83	−31.67	1427.27	1415.06	12.21
2^{-1}	5263.32	1374.64	1409.92	−35.28	1424.73	1412.35	12.38
2^{-2}	0.00	1366.43	1408.52	−42.09	1423.28	1411.29	11.99
2^{-3}	0.00	1361.19	1407.87	−46.68	1422.73	1410.85	11.68
2^{-4}	0.00	1358.55	1407.55	−49.00	1422.16	1410.64	11.52
2^{-5}	0.00	1357.23	1407.38	−50.15	1421.97	1410.53	11.44
2^{-6}	0.00	1356.57	1407.30	−50.73	1421.87	1410.48	11.40
2^{-7}	0.00	1356.24	1407.26	−51.02	1421.83	1410.45	11.38
2^{-8}	0.00	1356.08	1407.24	−51.16	1421.80	1410.44	11.37
2^{-9}	0.00	1356.00	1407.23	−51.24	1421.79	1410.43	11.36

forward market participation lead to higher prices in the forward market with respect to the spot and are captured by an increase in risk-premium with higher uncertainty in wind assets.

5.3 *System uncertainty*

In addition to uncertain availability, a natural question is how system uncertainty affects nodal prices in such settings. We consider a setting with four generators with seven generating nodes. The total generation capacity for each generator is given by $\mathcal{N}(1500, 400\theta)$. The linear and quadratic costs of generation are given by $\mathcal{N}(12, 2\theta)$ and $\mathcal{N}(0.3, 0.03\theta)$, respectively, where θ is the scaling of the variance. The risk aversion parameters κ and χ are assumed to be 1 and 0.5, respectively, for all agents and 10 scenarios are employed. The forward and spot intercepts are fixed to be 1500 across all nodes and for all scenarios while the forward and spot slopes are taken to be 1 across all nodes and all scenarios. Table 4 reports the spot prices, forward prices, and the risk premium as θ is driven to zero; in effect, we consider how prices change from a *stochastic* setting to an almost *deterministic* setting. Several observations may be made. A key observation that can be made is that when firms compete with uncertain assets, they encounter a risk penalty; however, as the level of uncertainty falls, firms make decisions that lead to little or no risk since they have complete foresight on the future.

5.4 *Specification of forward price functions*

A crucial question is how the choice of forward price function influences the results. In no-arbitrage models, this problem does not appear since the forward price function is not explicitly

Table 5. Relationship of forward participation and risk premiums to forward price functions.

Intercepts	Node 7		Node 10		Node 11	
	Total bids	$p_i^0 - \mathbb{E}p_i^\omega$	Total bids	$p_i^0 - \mathbb{E}p_i^\omega$	Total bids	$p_i^0 - \mathbb{E}p_i^\omega$
450	0	−621.64	0	−689.73	0	−424.67
600	0	−471.64	0	−539.73	0	−274.67
750	0	−321.64	0	−389.73	0	−124.67
900	0	−171.64	0	−239.74	22.20	3.53
1050	0	−21.64	0	−89.73	149.35	28.71
1200	107.48	19.62	52.73	8.52	275.17	55.20
1350	229.01	46.68	178.06	121.77	403.47	35.55
1500	351.73	72.74	303.77	62.19	534.44	100.68
1650	478.82	93.97	434.17	84.24	665.41	122.11

defined. In our market clearing models, we expect that our assumption on forward price function have significant impact on the results that emerge. Yet, it appears that for sufficiently low forward price intercepts, there is no forward market participation since the revenues garnered through participation are not sufficient. However, beyond a certain level, forward market participation becomes positive. Therefore, while the precise level of the forward market intercept is not as relevant, if the prices are set too low (a consequence of low intercepts), then this adversely affects bidding in this market.

In our experiments, we fix the spot intercepts, slopes and forward slopes and vary the forward intercepts from 150 to 1800 in steps of 150. We find that there are no forward bids till a particular threshold of the forward intercept. Beyond this level, the forward bids and the premium increases as the forward intercept increases. Table 5 shows the variation of the forward bids and premium across nodes 7, 10 and 11. We find that when there is no risk premium, there is no forward participation (when the expected spot prices are greater than the forward prices) and vice versa. When the risk premium is positive, there is an incentive for bidding in the forward market.

5.5 *Increasing penetration of wind*

As the role of renewables in the nation's fuel mix grows, a question that remains is whether forward markets will continue to attract participation. We investigate this question by increasing the mean of the capacity of the wind generators from 300 to 2050 in steps of 350 and also raise the standard deviations in availability from 150 to 1025 in steps of 175. We observe that for a fixed level of risk aversion, the forward bids of all the firms increase with increasing wind power penetration. This is in response to the volatility in the spot market with wind power penetration (Figure 8). It is also observed that with increasing wind power penetration, there is a significant
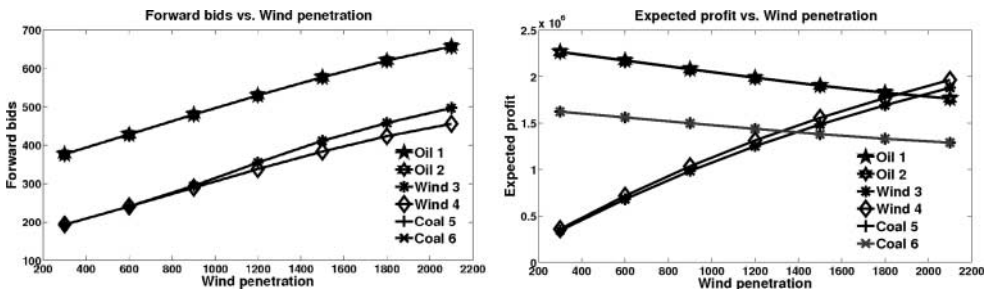

Figure 8. Increasing penetration-wind.

increase in profits of wind generators at the expense of the profits of firms with no wind assets as shown.

5.6 *Introduction of ramping constraints*

The current model may be extended to include ramping constraints. In one such extension, firms may day-ahead bids for a set of T consecutive hours. The second stage problem requires taking recourse over T consecutive period, coupled by linear ramping constraints. The associated variational problem is slightly more intricate but analogous avenues may be employed for providing existence statements. From an algorithmic standpoint, it can be seen that these constraints destroy the block diagonal property of the Jacobian of the variational mapping (as associated with the second-stage problems). The current algorithm can be applied directly since the structure of the projection problem is maintained as a stochastic quadratic program, albeit a more intricate one. In fact, the second-stage problem of the two-stage stochastic QP is a T-period problem which can itself be decomposed.

6. **Summary**

This paper is motivated by the need to manage risk exposure in the face of growing supply-side uncertainty in power markets, as a consequence of increasing penetration of wind power. Current market designs do not have a mechanism for managing this risk, and risk seeking generators are not held responsible for risk exposure; instead, this responsibility is borne by the system operator.

In this paper, we consider an uncertain two-period stochastic game where firms are charged a risk-based penalty. The resulting problem is a generalized stochastic Nash game where agents can be viewed as risk-averse and make first-period and second-period recourse decisions. By observing that the coupling between the strategy sets is through a set of shared constraints, a subset of equilibria to the original game are given by the solution to an appropriately defined variational inequality.

Risk-averseness in the agent problems is captured through a conditional value-at-risk (CVaR) measure that leads to non-smoothness. In fact, when these agent-specific measures are independent of competitive interactions, the related smooth games are shown to lead to monotone variational inequalities that are shown to admit solutions.

The monotonicity of the mapping in the variational problem allows for the use of regularized distributed projection schemes, both in a single time-scale (primal-dual) setting and a two time-scale (dual) setting. Rate of convergence estimates are provided for the dual scheme when the primal solution is computed exactly. A bounded complexity extension that allows for inexact computations of primal solution is also studied and leads to the provision of error bounds for the primal solution, dual solution and the infeasibility. The scalability of the projection scheme with $|\Omega|$, the cardinality of the sample space is contingent on effective solution of the projection step. In fact, we observe that this step essentially requires the solution of a strongly convex stochastic program and can be solved through a cutting-plane method that scales well with the cardinality of the sample space. Numerical results support that the scheme scales well with the size of the network, the number of firms and the size of the sample space.

The paper concludes with a discussion of insights for market design and operation by applying the model to a 53-node network drawn from the Belgian grid. Through this model, we observe that higher levels of risk-aversion lead to lower participation in the forward markets by agents with uncertain assets. Furthermore, higher levels of uncertainty in generation capacity leads to lower levels of forward participation. When forward price intercepts are sufficiently high, firms have incentives to participate in the forward market leading to a positive premium.

Acknowledgements

Notes

1. If generators make bids in the day-ahead market that are characterized by a higher likelihood of shortfall in the real-time market than a prescribed threshold, then such bids are referred to as 'aggressive.'
2. It should be noted that over generation may also lead to reliability concerns. A direct extension of such penalties to accommodate both under and over-generation may discourage firms from deviating from day-ahead bids. Incorporating such concerns remains a focus of future research.
3. A single settlement market model refers to one where a single clearing is analysed while a two-settlement market considers two market clearings (such as day-ahead and real-time markets) and, in some instances, models the second clearing as uncertain.
4. Network constraints are modelled by means of distribution factors or in other words, a DC approximation (linearization) of Kirchhoff's laws. The details on computing various power distribution factors are discussed in [42]. In our work, we employ one of the distribution factors, namely the Injection Shift Factor (ISF). If Q refers to the power distribution factor matrix, then $Q_{l,i}$ (lth row and ith column of Q) refers to the power flowing in line l due to unit injection of power at node i. In the use of ISF, one of the nodes in the network is assumed to be a slack node. By this assumption, injection or withdrawal of power at a slack node does not induce flow on any line in the network.
5. In general, firms cannot generate at all nodes but, for notational ease, we assume that all firms can generate at all nodes. This is overlaid by a set of additional constraints that reflect whether these firms *can* indeed generate at such nodes. Specifically, $y_{ij}^\omega, u_{ij}^\omega, v_{ij}^\omega, s_{ij}^\omega, m_{ij}, x_{ij} \equiv 0, \forall i \in G^c, \forall \omega \in \Omega, \forall j \in \mathcal{J}$. Refer to table A1 for notational details. Firms do not house generation facilities at all nodes. Therefore this also holds $\forall i \in \mathcal{J}_i^c$. These are introduced in the set $\mathbf{Z}_j$ appropriately and excluded from the formulation to ease the notation.
6. Note that in some settings, the variational equilibrium does not suffice. For instance, the question of interconnectedness across multiple grids has been recently studied in detail by Smeers and his coauthors [58] and the focus therein lies on analysing the quasi-variational inequalities associated with the generalized Nash game. In such settings, the Lagrange multiplier (interpreted as prices) associated with the 'shared' constraint need not be shared and such equilibria cannot be captured by a VE and one needs to instead focus on the quasi-variational inequality and its solution set.
7. Here s^{ref} is finite because *cap* is finite and $\varrho_{ij}(0, \text{cap}_{ij}^\omega)$ is also finite.
8. Note that Π denotes the projection, $\prod$ denotes the cartesian product and π denotes the player objectives.
9. The implementation was done on Matlab 7.11.0.584 (R 2011b) on a Linux OS with a processor with a clockspeed of 2.651 GHz and a memory of 8 GB.
10. Note that the PATH solver would have proved to be a better choice but was not available to us on the Tomlab environment on Linux.

References

[1] S. Ahmed, U. Çakmak, and A. Shapiro, *Coherent risk measures in inventory problems*, Eur. J. Oper. Res. 182 (2007), pp. 226–238.
[2] B. Allaz, *Oligopoly, uncertainty and strategic forward transactions*, J. Ind. Organ. 10 (1992), pp. 297–308.
[3] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath, *Coherent measures of risk*, in *Risk Management: Value at Risk and Beyond (Cambridge, 1998)*, Cambridge University Press, Cambridge, 2002, pp. 145–175.
[4] R. Baldick, *Electricity market equilibrium models: The effect of parametrization*, IEEE Trans. Power Syst. 17 (2002), pp. 1170–1176.
[5] E.M.L. Beale, *On minimizing a convex function subject to linear inequalities*, J. Roy. Statist. Soc. Ser. B. 17 (1955), pp. 173–184; discussion, 194–203, (Symposium on linear programming.).
[6] J.F. Benders, *Partitioning procedures for solving mixed-variables programming problems*, Numer. Math. 4 (1962), pp. 238–252.
[7] J.R. Birge and F. Louveaux, *Introduction to Stochastic Programming: Springer Series in Operations Research*, Springer, Berlin, 1997.
[8] R.H. Byrd, M.E. Hribar, and J. Nocedal, *An interior point algorithm for large-scale nonlinear programming*, SIAM J. Optim. 9 (1999), pp. 877–900 (electronic), dedicated to John E. Dennis, Jr., on his 60th birthday.
[9] J. Cardell, C. Hitt, and W. Hogan, *Market power and strategic interaction in electricity networks*, Resource Energy Econ. 19 (1997), pp. 109–137.
[10] D. Chan and J.S. Pang, *The generalized quasivariational inequality problem*, Math. Oper. Res. 7 (1982), pp. 211–222.
[11] A. Charnes and W.W. Cooper, *Chance-constrained programming*, Manage. Sci. 6 (1959/1960), pp. 73–79.

[12] X. Chen, R.J.B. Wets, and Y. Zhang, *Stochastic variational inequalities: Residual minimization smoothing/sample average approximations*, Optimization On-Line (2011), www.optimization-online.org/DB_HTML/2011/02/2930.html.

[13] X. Chen and R.S. Womersley, *Random test problems and parallel methods for quadratic programs and quadratic stochastic programs*, Optim. Methods Softw. 13 (2000), pp. 275–306.

[14] A.J. Conejo, F.J. Nogales, and J.M. Arroyo, *Risk-constrained self-scheduling of a thermal power producer*, IEEE Trans. Power Systems 19 (2004), pp. 1569–1574.

[15] G.B. Dantzig, *Linear programming under uncertainty*, Manage. Sci. 1 (1955), pp. 197–206.

[16] V. DeMiguel and H. Xu, *A stochastic multiple leader stackelberg model: Analysis, computation, and application*, Oper. Res. 57 (2009), pp. 1220–1235.

[17] F. Facchinei and C. Kanzow, *Generalized Nash equilibrium problems*, 4OR 5 (2007), pp. 173–210.

[18] F. Facchinei and J.S. Pang, *Finite Dimensional Variational Inequalities and Complementarity Problems: Vols. I and II*, Springer, New York, 2003.

[19] F. Facchinei and J.S. Pang, *Nash-Equilibria: The Variational Approach*, Convex Optimization in Signal Processing and Communications, Cambridge University Press, Cambridge, UK, 2009, pp. 443–493.

[20] F. Facchinei, A. Fischer, and V. Piccialli, *On generalized Nash games and variational inequalities*, Oper. Res. Lett. 35 (2007), pp. 159–164.

[21] M.C. Ferris and T.S. Munson, *Complementarity problems in GAMS and the PATH solver*, J. Econ. Dynam. Control 24 (2000), pp. 165–188.

[22] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, MA, 1991.

[23] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Johns Hopkins Series in the Mathematical Sciences, Vol. 3, Johns Hopkins University Press, Baltimore, MD, 1983, pp. xvi+476.

[24] P.T. Harker, *A variational inequality approach for the determination of oligopolistic market equilibrium*, Math. Programming 30 (1984), pp. 105–111.

[25] P.T. Harker, *Generalized Nash games and quasi-variational inequalities*, Eur. J. Oper. Res. 54 (1991), pp. 81–94.

[26] A. Haurie and F. Moresino, *Computation of S-adapted equilibria in piecewise deterministic games via stochastic programming methods*, in *Advances in Dynamic Games and Applications (Maastricht, 1998)*, Annals of International Society for Dynamic Games Vol. 6, Birkhäuser Boston, Boston, MA, 2001, pp. 225–252.

[27] A. Haurie and G. Zaccour, *S-adapted equilibrium in games played over event trees: an overview*, in *Advances in Dynamic Games*, Annals of International Society for Dynamic Games Vol. 7, Birkhäuser Boston, Boston, MA, 2005, pp. 417–444.

[28] A. Heusingervon and C. Kanzow, *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions*, Comput. Optim. Appl. 43 (2009), pp. 353–377.

[29] B.F. Hobbs, *Linear complementarity models of Nash–Cournot competition in bilateral and poolco power markets*, IEEE Trans. Power Syst. 16 (2001), pp. 194–202.

[30] B.F. Hobbs and J.S. Pang, *Nash-Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints*, Oper. Res. 55 (2007), pp. 113–127.

[31] B.F. Hobbs, C.B. Metzler, and J.S. Pang, *Strategic gaming analysis for electric power systems: An MPEC approach*, IEEE Trans. Power Syst. 15 (2000), pp. 638–645.

[32] X. Hu and D. Ralph, *Using EPECs to model bilevel games in restructured electricity markets with locational prices*, Oper. Res. 55 (2007), pp. 809–827.

[33] G. Infanger, *Planning under Uncertainty*, Boyd and Fraser Publishing Co., New York, 1994.

[34] R. Kamat and S. Oren, *Two-settlement systems for electricity markets under network uncertainty and market power*, J. Regulatory Econ. 25 (2004), pp. 5–37.

[35] I.V. Konnov, *Equilibrium Models and Variational Inequalities*, Elsevier, Amsterdam, 2007.

[36] J. Koshal, A. Nedić, and U.V. Shanbhag, *Distributed multiuser optimization: Algorithms and error analysis*, Proceedings of the 49th IEEE Conference on Decision and Control, CDC 2010, December 15–17, 2010, Atlanta, Georgia, USA, pp. 4372–4377.

[37] J. Koshal, A. Nedić, and U.V. Shanbhag, *Multiuser optimization: Distributed algorithms and error analysis*, SIAM J. Optim. 21 (2011), pp. 1046–1081.

[38] J. Krawczyk and S. Uryasev, *Relaxation algorithms to find nash equilibria with economic applications*, Environ. Model. Assess. 5 (2000), pp. 63–73.

[39] A. Kulkarni and U.V. Shanbhag, *Recourse-based stochastic nonlinear programming: Properties and Benders-SQP algorithms*, Comput. Optim. Appl. 51 (2012), pp. 77–123.

[40] A. Kulkarni, A. Rossi, J. Alameda, and U.V. Shanbhag, *A grid computing framework for quadratic programming under uncertainty*, in *Proceedings of the TeraGrid*, June 4–8, Madison, Wisconsin, 2007.

[41] S. Leyffer and T. Munson, *Solving multi-leader-common-follower games*, Optim. Meth. Softw. 25 (2010), pp. 601–623.

[42] M. Liu and G. Gross, *Role of distribution factors in congestion revenue rights applications*, IEEE Trans. Power Syst. 19 (2004), pp. 802–810.

[43] Z.Q. Luo, J.S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.

[44] F. Meng, J. Sun, and M. Goh, *A smoothing sample average approximation method for stochastic optimization problems with cvar risk measure*, Comput. Optim. Appl. 50 (2011), pp. 379–401.

[45] C. Metzler, B.F. Hobbs, and J.S. Pang, *Nash-Cournot equilibria in power markets on a linearized DC network with arbitrage: Formulations and properties*, Networks and Spatial Theory 3 (2003), pp. 123–150.

[46] J.F. Nash, *Equilibrium points in N-person games*, Proc. Natl Acad. Sci. USA 36 (1950), pp. 48–49.

[47] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, anniversary ed., Princeton University Press, Princeton, NJ, 2007, pp. xxxii+739, with an introduction by Harold W. Kuhn and an afterword by Ariel Rubinstein.

[48] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer, New York, 1999, pp. xxii+636.

[49] M.J. Osborne and A. Rubinstein, *A Course in Game Theory*, MIT Press, Cambridge, MA, 1994, pp. xvi+352.

[50] J.S. Pang and M. Fukushima, *Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games*, Comput. Manage. Sci. 2 (2005), pp. 21–56.

[51] R. Rockfellar and S. Uryasev, *Optimization of conditional value-at-risk*, J. Risk 2 (2000), pp. 21–41.

[52] R.T. Rockafellar and R.J.B. Wets, *Stochastic convex programming: Kuhn–Tucker conditions*, J. Math. Econ. 2 (1975), pp. 349–370.

[53] J.B. Rosen, *Existence and uniqueness of equilibrium points for concave n-person games*, Econometrica 33 (1965), pp. 520–534.

[54] A. Ruszczynski, *Decomposition methods*, in *Handbook in Operations Research and Management Science* Vol. 10, Elsevier Science, Amsterdam, 2003, pp. 141–212.

[55] A. Ruszczynski and A. Shapiro, *Introduction*, in *Handbook in Operations Research and Management Science* Vol. 10, Elsevier Science, Amsterdam, 2003, pp. 1–64.

[56] U.V. Shanbhag, *Decomposition and sampling methods for stochastic equilibrium problems*, Ph.D. thesis, Department of Management Science and Engineering (Operations Research), Stanford University, 2006.

[57] U.V. Shanbhag, G. Infanger, and P.W. Glynn, *A complementarity framework for forward contracting under uncertainty*, Oper. Res. 59 (2011), pp. 810–834.

[58] Y. Smeers, G. Oggioni, E. Allevi, and S. Schaible, *Generalized Nash equilibrium and market coupling in the European power system*, Technical Report on Electricity Policy Research Group, University of Cambridge, EPRG1016, 2010.

[59] C.L. Su, *Analysis on the forward market equilibrium model*, Oper. Res. Lett. 35 (2007), pp. 74–82.

[60] R.M. Van Slyke and R. Wets, *L-shaped linear programs with applications to optimal control and stochastic programming*, SIAM J. Appl. Math. 17 (1969), pp. 638–663.

[61] H. Xu and D. Zhang, *Stochastic Nash equilibrium problems: sample average approximation and applications*, Optimization-online (2009), http://www.optimization-online.org/DB_HTML/2009/05/2299.html.

[62] J. Yao, S.S. Oren, and I. Adler, *Two-settlement electricity markets with price caps and cournot generation firms*, Eur. J. Oper. Res. 181 (2007), pp. 1279–1296.

[63] J. Yao, I. Adler, and S.S. Oren, *Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network*, Oper. Res. 56 (2008), pp. 34–47.

[64] G. Zakeri, A.B. Philpott, and D.M. Ryan, *Inexact cuts in Benders decomposition*, SIAM J. Optim. 10 (2000), pp. 643–657.

Appendix 1. Notation and Network details

A.1 *The complementarity problem*

Under the assumptions of regularity, the variational inequality can be written in the form of a complementarity problem. The solution to the complementarity problem is the same as that of the VI and in turn a solution for the original GNP. For computation, we solve the complementarity problem to obtain the solution of the VI and in turn a solution to the original GNP. For our computation, we assume the loss function to be linear. Let us assign the multipliers α_{ij}^{ω} and β_{ij}^{ω} to equality and capacity constraints, respectively, for the firms' problems. Let γ_{ij}^{ω} and δ_{ij}^{ω} refer to the constraints with respect to s_{ij}^{ω} (firms' problems). Let, μ^{ω} and $\sigma_l^{\omega}, \eta_l^{\omega}$ be the multipliers assigned to the power balance/equality and transmission constraints of the independent system operator. Let ϕ_i^{ω} represent the multiplier for the shared constraint. Then, the complementarity problem is given by (note that indexing is omitted for purposes of brevity):

$$0 \leq x_{ij} \perp b_i^0 x_{ij} + b_i^0 \sum_{j \in \mathcal{J}} x_{ij} - a_i^0 + \sum_{\omega \in \Omega} \rho^{\omega} a_i^{\omega} - \sum_{\omega \in \Omega} \rho^{\omega} b_i^{\omega} \left(\sum_{j \in \mathcal{J}} y_{ij}^{\omega} + r_i^{\omega} \right)$$

$$- \sum_{\omega \in \Omega} \alpha_{ij}^{\omega} + \chi \sum_{\omega \in \Omega} \delta_{ij}^{\omega} \geq 0,$$

$$0 \le y_{ij}^{\omega} \perp \rho^{\omega}\left(-a_i^{\omega} + c_i^{\omega} + (b_i^{\omega} + d_{ij}^{\omega})y_{ij}^{\omega} + b_i^{\omega}\left(\sum_{j\in\mathcal{J}} y_{ij}^{\omega}\right) + b_i^{\omega}r_i^{\omega} - b_i^{\omega}x_{ij}\right)$$

$$+ \alpha_{ij}^{\omega} + \beta_{ij}^{\omega} - \phi_i^{\omega} \ge 0,$$

$$0 \le u_{ij}^{\omega} \perp -\alpha_{ij}^{\omega} \ge 0,$$

$$0 \le v_{ij}^{\omega} \perp \alpha_{ij}^{\omega} \ge 0,$$

$$0 \le s_{ij}^{\omega} \perp \frac{\kappa_j \rho^{\omega}}{1-\tau} - \gamma_{ij}^{\omega} - \delta_{ij}^{\omega} \ge 0,$$

$$\text{free} \perp \kappa_j - \sum_{j\in\mathcal{J}} \gamma_{ij}^{\omega} - \sum_{j\in\mathcal{J}} \delta_{ij}^{\omega} = 0,$$

$$0 \le \beta_{ij}^{\omega} \perp \text{cap}_{ij}^{\omega} - y_{ij}^{\omega} \ge 0,$$

$$0 \le \gamma_{ij}^{\omega} \perp s_{ij}^{\omega} + m_{ij} \ge 0,$$

$$0 \le \delta_{ij}^{\omega} \perp s_{ij}^{\omega} + m_{ij} - \chi(x_{ij} - \text{cap}_{ij}^{\omega}) \ge 0,$$

$$\text{free} \perp y_{ij}^{\omega} - x_{ij} + u_{ij}^{\omega} - v_{ij}^{\omega} = 0,$$

$$0 \le \phi_i^{\omega} \perp \sum_{j\in\mathcal{J}} y_{ij}^{\omega} + r_i^{\omega} \ge 0,$$

$$\text{free} \perp -\rho^{\omega}a_i^{\omega} + \rho^{\omega}b_i^{\omega}\left(\sum_{j\in\mathcal{J}} y_{ij}^{\omega} + r_i^{\omega}\right) + \mu^{\omega} + \sum_{l\in\mathcal{L}} Q_{l,i}(\sigma_l^{\omega} - \eta_l^{\omega}) - \phi_i^{\omega} = 0, \quad i \in G,$$

$$r_i^{\omega} \perp -\rho^{\omega}a_i^{\omega} + \rho^{\omega}b_i^{\omega}r_i^{\omega} + \mu^{\omega} + \sum_{l\in\mathcal{L}} Q_{l,i}(\sigma_l^{\omega} - \eta_l^{\omega}) \ge 0, \quad i \in (G^c - \{51\}),$$

$$r_i^{\omega} \perp -\rho^{\omega}a_i^{\omega} + \rho^{\omega}b_i^{\omega}r_i^{\omega} + \mu^{\omega} \ge 0, \quad \text{slack node-51},$$

$$\text{free} \perp \sum_{i\in\mathcal{N}} r_i^{\omega} = 0,$$

$$\sigma_l^{\omega} \perp K_l^{\omega} - \sum_{i\in\mathcal{N}} Q_{l,i}r_i^{\omega} \ge 0,$$

$$\eta_l^{\omega} \perp K_l^{\omega} + \sum_{i\in\mathcal{N}} Q_{l,i}r_i^{\omega} \ge 0.$$

Table A1. Notation.

x_{ij}	Forward decision of generation from firm j at node i
$u_{ij}^{\omega}, v_{ij}^{\omega}$	Positive and negative deviations, respectively, at scenario ω from firm j at node i
$y_{ij}^{\omega}, \mathrm{cap}_{ij}^{\omega}$	Total spot generation decision and total generation capacity at scenario ω for firm j at node i
r_i^{ω}	ISO's spot decision at scenario ω at node i
n, Ω, ρ^{ω}	Number of scenarios, set of all scenarios and probability of scenario ω
p_i^{ω}	Nodal demand function or price at scenario ω at node i
$c_{ij}^{\omega}, d_{ij}^{\omega}$	Coefficient of linear and quadratic terms in the cost function at scenario ω for firm j at node i
f_p, f_n	Penalty functions for positive and negative deviations
N_g, N	Number of generating nodes and total nodes in the network
a_i^0, b_i^0	Intercept and Slope respectively at node i in the forward market
$a_i^{\omega}, b_i^{\omega}$	Intercept and Slope respectively at node i at scenario ω
$g + 1$	Number of agents including g firms and the ISO – $(g + 1)$th agent
$Q_{l,i}$	Power flowing across line l due to unit injection/withdrawal of power at node i
$\kappa_j, \forall j \in \mathcal{J}$	Risk factor or risk aversion parameter for firm j
$\mathcal{N}_j, \mathcal{N}_j^c$	Set of all generating nodes and non-generating nodes for firm j respectively
$\mathcal{J}_i$	Set of all generating firms at node i
m_{ij}	Value at risk for firm j at node i
s_{ij}^{ω}	Conditional value at risk for firm j at node i at scenario ω
$\mathcal{L}, \mathcal{N}$	Set of all transmission lines and set of all nodes respectively
G, G^c	Set of all generating nodes and load nodes respectively
$\mathcal{J}, \mathcal{A}$	Set of all generating firms and set of all agents (firms and the ISO), respectively

Table A2. Network details.

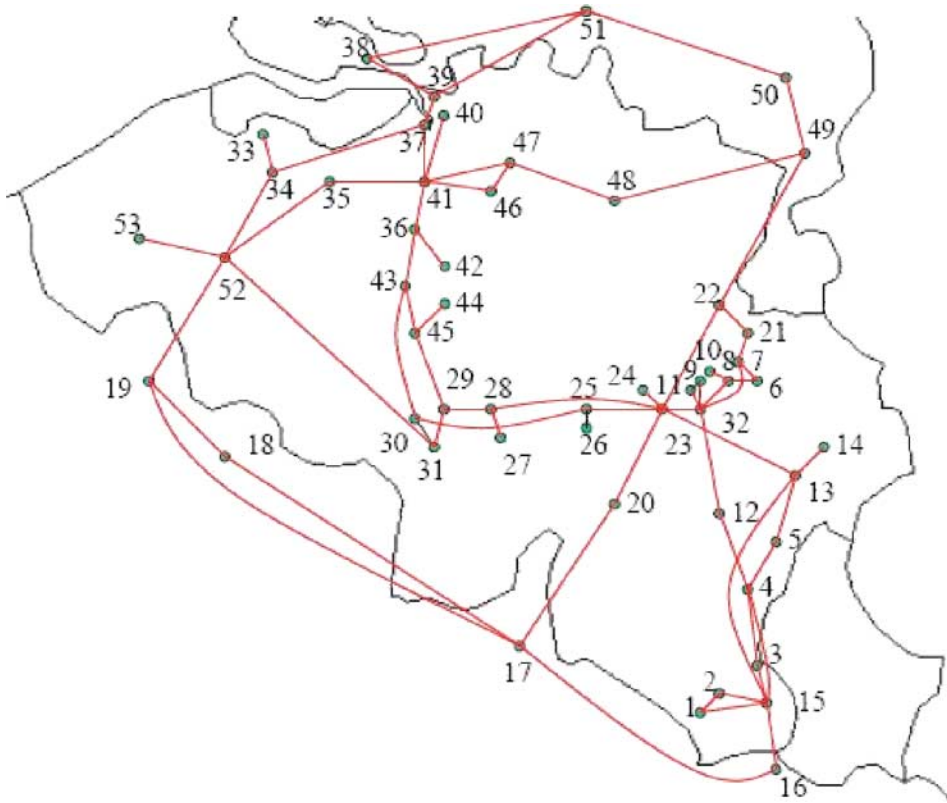| Line	Imp. (Ohm)	Cap.(MW)	Line	Imp. (Ohm)	Cap.(MW)	Line	Imp. (Ohm)	Cap.(MW)
1–2	23,716	345	16–17	2,633	5,154	34–37	7,048	1,350
1–15	6,269	345	17–18	4,236	1,715	34–52	12,234	1,350
2–15	8,534	345	17–19	1,939	5,140	35–41	14,204	1,350
3–4	5,339	240	17–20	8,071	1,179	35–52	9,026	1,420
3–15	11,686	240	18–19	1,465	13,170	36–41	15,777	2,770
4–5	6,994	510	19–52	11,321	1,179	36–42	11,186	2,840
4–12	5,887	405	20–23	13,165	1,316	36–43	15,408	2,770
4–15	3,644	240	21–22	47,621	1,420	37–39	66,471	1,420
5–13	6,462	510	22–23	11,391	1,350	37–41	21,295	1,350
6–7	23,987	300	22–49	9,138	1,350	38–39	10,931	1,650
6–8	9,138	400	23–24	41,559	5,540	38–51	17,168	946
7–21	14,885	541	23–25	16,982	1,420	39–51	8,596	1,650
7–32	5,963	410	23–28	8,610	1,350	40–41	11,113	2,770
8–9	45,360	400	23–32	33,255	1,350	41–46	11,509	2,840
8–10	26,541	800	25–26	134,987	1,420	41–47	13,797	1,420
8–32	11,467	400	25–30	11,991	1,420	43–45	34,468	1,350
9–11	20,157	410	27–28	64,753	1,420	44–45	47,128	1,420
9–32	10,012	375	28–29	38,569	1,350	46–47	34,441	1,420
11–32	18,398	375	29–31	284,443	1,350	47–48	14,942	1,420
12–32	4,567	405	29–45	14,534	1,350	48–49	6,998	1,420
13–14	121,410	2,700	30–31	269,973	1,420	49–50	5,943	3,784
13–15	5,094	790	30–43	10,268	1,420	50–51	2,746	5,676
13–23	5,481	2,770	31–52	1,453	400	52–53	1,279	2,840
15–16	8,839	400	33–34	40,429	1,420			

Figure A1.   The Belgian grid.