

# Inverse Equilibrium Analysis of Oligopolistic Electricity Markets

Simon Risanger, *Student Member, IEEE*,  
Stein-Erik Fleten, and Steven A. Gabriel, *Senior Member, IEEE*

**Abstract**—Inverse equilibrium modeling fits parameters of an equilibrium model to observations. This allows investigation of whether market structures fit observed outcomes and it has predictive power. We introduce a methodology that leverages relaxed stationarity conditions from Karush-Kuhn-Tucker conditions to set up inverse equilibrium problems. This facilitates reframing of existing equilibrium approaches on power systems into inverse equilibrium programs. We illustrate the methodology on network-constrained and unconstrained Nash-Cournot games between price-making power generators. The inverse equilibrium problems in this paper reformulate into linear programming problems that are flexible and interpretable. Still, inverse equilibrium modeling provides generally inconsistent estimation and econometric approaches are better for this purpose.

**Index Terms**—Inverse equilibrium, inverse optimization, equilibrium modeling, electricity markets.

## I. INTRODUCTION

**D**ESPITE the liberalization of electricity markets, features such as a limited amount of large producers, high investment costs, and transmission constraints may cause price-making behavior, barriers of entry, and reduce access to markets. As a result, the markets are vulnerable to abuse of market power. Equilibrium models, which represent these oligopolistic tendencies, are therefore widely used to study electricity markets [1].

When we study actual energy markets, it is generally easy to observe the equilibrium outcomes, such as prices and flows. The theoretical development in inverse equilibrium modeling [2], [3] leverages this fact. The framework expands the theory of inverse optimization [4], which fit parameters of an optimization problem given observations of decision variables. As a result, we

can use actual data to analyze markets and participant behavior to a greater extent.

Recent literature shows an increased interest from the power systems community in inverse optimization. Applications include the investigation of price response of consumers [5], [6], estimation of offer prices from rival producers [7], and investigation of the parameters of transmission constraints in electricity markets based on locational marginal prices [8]. Relevant work on inverse equilibrium models include [9] and [10], which use the variational inequality approach of [3] to estimate bid curves of competing firms that employ strategic bidding.

Expanding the literature cited above, we show how to use a Karush-Kuhn-Tucker (KKT) representation [11] to formulate inverse equilibrium models. This allows existing equilibrium models from KKT formulations to be rearranged into inverse problems. Although [2] also considers inverse nonlinear complementarity problems, their approach requires initial estimates of parameters. Our methodology follows the idea of [3] and [11], where they minimize relaxed optimality conditions. As a result, we can apply our observations directly and solve the inverse equilibrium problem as an optimization problem.

Considering the rich history of equilibrium modeling in the power system community, it is natural to assume that inverse equilibrium modeling can be a valuable tool. While this is true to some extent, the approach also has limitations. The goal of this paper is to highlight both strengths and weaknesses of inverse equilibrium to modelers who consider using this method. Our contributions are the following:

- We develop a method to fit objective function coefficients of participants in a power system by inverting an equilibrium model from KKT conditions.
- We explain how inverse equilibria relate to similar concepts in econometrics and machine learning.
- We invert a Nash-Cournot game of transmission-constrained and unconstrained electricity markets.
- We use examples to illustrate how inverse equilibrium fits models and describe its performance in the presence of noise.
- We discuss performance, implementation, and challenges of inverse equilibrium models, as illustrated

Manuscript received April 25, 2019; revised September 19, 2019 and January 24, 2020; accepted April 26, 2020.

(Corresponding author: Simon Risanger)

S. Risanger and S.-E. Fleten are with the Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway (e-mail: simon.risanger@ntnu.no).

S. A. Gabriel is with the Department of Mechanical Engineering and the Applied Mathematics, Statistics, and Scientific Computation Program, University of Maryland, College Park, Maryland, USA, and the Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway.

by our examples.

The remainder of this paper is as follows: Section II outlines how inverse equilibrium modeling relates to econometrics and machine learning. Section III provides an introduction to equilibrium models from KKT conditions and explain how to utilize stationarity conditions to invert the problem. We apply the method on relevant examples in Section IV. Section V addresses implementation challenges, while Section VI concludes the paper.

## II. RELATIONSHIP TO ECONOMETRICS AND MACHINE LEARNING

At first glance, inverse equilibrium modeling may seem like another addition to the literature on structural econometrics [12]. Several econometric studies exist on electricity markets, especially intending to expose market power (see e.g. [13], [14] and [15]). However, the major difference is that inverse equilibrium modeling is a completely data-driven method. As a result, we make no assumptions on the distribution of our observations. Rather, we try to fit an equilibrium model of a market structure and see whether it fits the data well or not. Econometric estimation, on the other hand, assumes that there is an underlying population, which our sample data should reasonably represent, and tries to estimate true parameters of the population. This gives greater explanatory power than inverse equilibrium modeling. The cost, however, is careful data collection and estimator formulations. For instance, the estimators require that data comply with certain attributes, a traditionally prominent example is the Gauss-Markov assumptions, to enjoy statistical properties like unbiasedness and consistency. Although state-of-the-art econometrics have non-parametric estimation methods and approaches to handle challenges such as heteroskedasticity, serial correlation and endogeneity, the field nevertheless require a set of assumptions on the data in order to infer from it. Estimations from inverse equilibrium modeling, which do not require these assumptions, do thereby not share these properties. We illustrate this by example in Section IV. For an example of structural estimation in power systems see [16], for an overview on econometric methods, see e.g. [17] or [18]. In addition, [3, Appendix 2] discusses the relationship between inverse equilibrium modeling and structural estimation, while [19] suggest poor accuracy from estimation by first-order conditions within a conjectural variations framework of an oligopolistic electricity market.

Inverse equilibrium modeling relates more to a machine learning philosophy, which values prediction over explanation, than econometrics. However, inverse equilibrium modeling adds more structure than a pure machine learning predictor. Most notably, inverse equilibrium modeling has a strong prior. We believe that a

certain equilibrium market structure is the basis for the observations and want to see whether or not this is correct. If an inverse model is a good fit to the data, we can insert the fitted parameters in the original problem to obtain good predictive power [3]. Although this is a nice feature, we limit this paper to only consider formulating and solving inverse equilibrium problems, and refer the interested reader to [10] and [20] that use inverse optimization for prediction.

From the discussion, we see that inverse equilibrium complements existing econometrics and machine learning methods. We emphasize that inverse equilibrium modeling is generally an inconsistent estimator. Even if we get interpretable fitted parameters, such as costs or willingness-to-pay, we cannot conclude with confidence that they represent those of an underlying market. They are merely a good fit. If the goal of a study is to estimate true market parameters, econometric approaches should be used. That being said, inverse equilibrium modeling has several advantages:

- We require no assumptions on the input data.
- Inverse equilibrium modeling is flexible, and one can easily add or remove constraints and alter the problem.
- The problem often rearranges into a tractable linear programming problem.
- One can obtain estimates for other values than objective function coefficients, for instance coefficients of transmission constraints as shown in [8].
- By using the KKT approach of this paper, it is simple to invert mixed complementarity models.
- Inverse equilibrium models have more structure than pure machine learning predictors, which increases interpretability.

## III. INVERSE EQUILIBRIUM MODELING

### A. Equilibrium models

We consider a set of decision-makers,  $\mathcal{P} = \{1, \dots, |\mathcal{P}|\}$ , where each player  $p \in \mathcal{P}$  has an optimization problem illustrated by (1). Functions  $f_p$ ,  $g_{pi}$ , and  $h_{pj}$  may be different or similar for the different decision-makers. Moreover,  $\theta$ ,  $\phi$  and  $\psi$  denote the parameters of the respective functions. Notice that the objective (1a) is dependent on  $x_{-p} = (x_k)_{k \in \mathcal{P} \setminus p}$ , which denotes the decisions of the other players, in addition to its own decision variable vector  $x_p$ . The problem can be restricted by inequality constraints  $i \in \mathcal{I}$  and equality constraints  $j \in \mathcal{J}$ . Because restrictions (1b) and (1c) do not depend on  $x_{-p}$ , they are internal constraints for player  $p$ . Finally, we note that  $\lambda_{pi}$  and  $\nu_{pj}$  represent the

dual variables of constraints (1b) and (1c), respectively.

$$\min_{x_p} f_p(x_p, x_{-p}; \theta_p, \theta_{-p}) \quad (1a)$$

$$\text{s.t.} \quad g_{pi}(x_p; \phi_p) \leq 0, \quad (\lambda_{pi}) \quad i \in \mathcal{I} \quad (1b)$$

$$h_{pj}(x_p; \psi_p) = 0, \quad (\nu_{pj}) \quad j \in \mathcal{J} \quad (1c)$$

The decision-makers cannot optimize their own problem without considering the responses of the other players. Solving all  $p \in \mathcal{P}$  problems simultaneously leads to an equilibrium problem. Both variational inequalities (VIs) and mixed complementarity problems (MCPs) are paradigms to model the simultaneous solution of these player-specific problems. VIs are based on considering the variational principle related to non-negative directional derivatives for feasible directions (to minimization problems). MCPs rely on the KKT conditions and involve both primal and dual variables, which has a modeling advantage in some cases [1]. We only consider MCPs in the remainder of this paper.

We assume that problem (1) for all  $p \in \mathcal{P}$  satisfies a constraint qualification that makes the KKT conditions necessary. The KKT conditions are sufficient, for example, when  $f_p$  is convex (concave for a maximization problem) while  $g_{pi}$  and  $h_{pj}$  are affine. A solution that satisfies the KKT conditions (when these conditions are sufficient) is thus an optimal solution of (1). Likewise, a solution that simultaneously satisfies the KKT conditions for all  $p \in \mathcal{P}$ , as shown in (2), is an equilibrium solution.

$$\nabla_{x_p} f_p(x_p, x_{-p}; \theta_p, \theta_{-p}) + \sum_{i \in \mathcal{I}} \lambda_{pi} \nabla_{x_p} g_{pi}(x_p; \phi_p) \quad (2a)$$

$$+ \sum_{j \in \mathcal{J}} \nu_{pj} \nabla_{x_p} h_{pj}(x_p; \psi_p) = 0, \quad p \in \mathcal{P}$$

$$g_{pi}(x_p; \phi_p) \leq 0, \quad i \in \mathcal{I}, p \in \mathcal{P} \quad (2b)$$

$$h_{pj}(x_p; \psi_p) = 0, \quad j \in \mathcal{J}, p \in \mathcal{P} \quad (2c)$$

$$\lambda_{pi} \geq 0, \quad i \in \mathcal{I}, p \in \mathcal{P} \quad (2d)$$

$$\lambda_{pi} g_{pi}(x_p; \phi_p) = 0, \quad i \in \mathcal{I}, p \in \mathcal{P} \quad (2e)$$

## B. Inverse equilibrium models

Problem (2) assumes that parameters,  $\theta$ ,  $\phi$  and  $\psi$ , are fixed and seeks a solution satisfying all the conditions. By contrast, inverse equilibrium modeling is the reverse-engineering direction to this. Namely, given an equilibrium solution, it seeks to find the parameters  $\theta$ ,  $\phi$  and  $\psi$  that best fit the observed solution. The equilibrium outcomes, represented by the decision variables  $x_1, \dots, x_{|\mathcal{P}|}$  become fixed observations, and thus parameters,  $\tilde{x}_1, \dots, \tilde{x}_{|\mathcal{P}|}$ , in the inverse problem. Our method is similar to [11], which applies KKT relaxations to convex optimization problems.

We allow stationarity conditions (2a) to be relaxed, while constraints (2b) to (2e) must hold. A deviation

from (2a) results in near-equilibrium solutions, but outcomes are still feasible when (2b) to (2e) hold. We can thus relax the stationarity condition by deviation  $\epsilon_p$ , as shown in (3), to create a near-equilibrium solution. This allows the inverse model to consider observations that are not necessarily optimal strategies for its assumed model. Note that the deviations are not independent because we relax the stationarity condition, which includes decision variables of the other problems.

$$\begin{aligned} \nabla_{x_p} f_p(x_p, x_{-p}; \theta_p, \theta_{-p}) - \sum_{i \in \mathcal{I}} \lambda_{pi} \nabla_{x_p} g_{pi}(x_p; \phi_p) \\ - \sum_{j \in \mathcal{J}} \nu_{pj} \nabla_{x_p} h_{pj}(x_p; \psi_p) = \epsilon_p \end{aligned} \quad (3)$$

We assume that observations come from rational players, and thus are optimal decisions in the actual market. The inverse equilibrium problem (4) therefore seeks to minimize the vector norm of these deviations,  $\|\epsilon\|$  where  $\epsilon = \{\epsilon_p : p \in \mathcal{P}\}$ . This fits the parameters in a manner where the observations are as optimal as possible for the assumed model. Recall that observations  $\tilde{x}_1, \dots, \tilde{x}_{|\mathcal{P}|}$  are parameters in the inverse problem. The dual variables,  $\lambda_{ki}$  and  $\nu_{kj}$ , become parameters if they are observable. A notable example is prices, which are dual variables of market-clearing constraints and observable at the power exchange. If unobservable, the dual variables continue to be decision variables, which we assume for the remainder of the paper. The parameters we want to fit, for instance cost coefficients, slopes or intercepts of inverse demand functions, also become decision variables.

$$\min_{\epsilon, \lambda, \nu, \theta, \psi, \phi} \|\epsilon\| \quad (4a)$$

$$\begin{aligned} \text{s.t.} \quad \nabla_{x_p} f_p(\tilde{x}_p, \tilde{x}_{-p}; \theta_p, \theta_{-p}) \\ - \sum_{i \in \mathcal{I}} \lambda_{pi} \nabla_{x_p} g_{pi}(\tilde{x}_p; \phi_p) \\ - \sum_{j \in \mathcal{J}} \nu_{pj} \nabla_{x_p} h_{pj}(\tilde{x}_p; \psi_p) = \epsilon_p, \quad p \in \mathcal{P} \end{aligned} \quad (4b)$$

Constraints (2b) to (2e)

Depending on the number of variables that are observable and how many parameters we try to fit, there may be several optimal solutions for (4). With respect to interpretability, we want the solution space as small as possible. We can achieve this by adding constraints, getting observations for variables and fitting fewer parameters. Several different observations also increase the probability of having marginal observations, i.e. observations that reveals some limit of the variables. This reduces scale invariance, which is the situation where the fitted parameters has a range of optimal solutions.

We therefore introduce  $h \in \mathcal{H} = \{1, \dots, |\mathcal{H}|\}$  as index for different observations. For instance, the electricity market outcomes for multiple hours or days. We

introduce observations  $\tilde{x}_{1h}, \dots, \tilde{x}_{|\mathcal{P}|h}$  into the inverse equilibrium problem and minimize the deviation at each observation,  $\epsilon_{ph}$ , constrained to (4b) and (2b) to (2e) for all observations.

The inverse equilibrium problem has several convenient computational properties compared to ordinary equilibrium problems. Complementarity constraints of equilibrium problems are non-convex, and thus computationally challenging for large instances. When decision variables become fixed observations, they cease being variables. If an observed variable is part of a bilinear term, the term becomes linear. If one wants to fit parameters in an inequality constraint, i.e.  $\phi$ , complementarity conditions can arise because we multiply  $\phi$  with the dual variable  $\lambda$  in constraint (2e). However, this is not an issue if we do not need to estimate  $\phi$  or if we have observations of its corresponding dual variable  $\tilde{\lambda}$ .

Objective function (4a) minimizes the distance from the objective and can be represented by any norm. An  $L_1$ -norm (the sum of absolute values) or  $L_\infty$ -norm (the single largest magnitude in a vector) linearizes the inverse equilibrium objective. For the examples in Section IV, we use the  $L_1$ -norm. If the constraints are affine, then (4) becomes a linear programming problem. Consequently, we are able to solve much larger instances of inverse equilibrium problems than equilibrium problems.

### C. Pre-process data to reduce problem size

Although we can solve the inverse equilibrium problem in its original form (4), pre-processing data reduces problem size and decreases the risk of numerical complications. Take for instance restriction (2e):

$$\lambda_{pi} g_{pi}(x_p; \phi) = 0.$$

Given an observation  $\tilde{x}_p$  and we know  $\phi$ , then we know the value of  $g_{pi}(\tilde{x}_p; \phi)$ , which now becomes a parameter in the problem. If  $g_{pi}(\tilde{x}_p; \phi) = 0$ , we can omit restriction (2e), because we know it is satisfied. Similarly, if  $g_{pi}(\tilde{x}_p; \phi) \neq 0$ , we can set  $\lambda_{pi} = 0$  instead of the numerically more complicated (2e). In addition, non-negativity constraint (2d) becomes redundant.

## IV. ILLUSTRATIVE EXAMPLES OF INVERSE EQUILIBRIUM MODELS

To illustrate the computational aspects of solving inverse equilibrium problems, we introduce two Nash-Cournot games where strategic generators use market power to maximize profits. Throughout the section, we use the PATH solver [21] in GAMS to solve the equilibrium problems, while we implement the inverse equilibrium problems, which become linear programming problems, in the Pyomo package for Python and solve with the Gurobi solver.

### A. Generic Nash-Cournot game

1) *Model formulation:* First we consider a generic Nash-Cournot game between  $p \in \mathcal{P}$  price-making generators with finite capacity. They supply a price-sensitive load without any transmission constraints. Generation is denoted  $x_p$ , and has a marginal cost  $c_p$ , as described by optimization problem (5). Each generator tries to maximize its profits, given by objective function (5a). A linear inverse demand function with slope  $a \geq 0$  and intercept  $b \geq 0$  determines the price. We include  $\xi$  as a demand shock that increases or decreases the demand intercept. In actual application, there is significant uncertainty regarding  $\xi$ . We include it merely to generate different observations for the case study. A generator cannot exceed its maximum generation capacity  $Q_p^{max}$ , as enforced by (5b), and generation is non-negative. Finally,  $\mu_p$  denotes the dual variable of the maximum generation restriction.

$$\max_{x_p} \quad -c_p x_p + \left( b + \xi - a \sum_{k \in \mathcal{P}} x_k \right) x_p \quad (5a)$$

$$\text{s.t.} \quad x_p \leq Q_p^{max} \quad (\mu_p) \quad (5b)$$

$$x_p \geq 0 \quad (5c)$$

We formulate the KKT conditions of (5) as described in Section III-A. The objective (5a) is concave and constraints (5b) and (5c) are affine, so the KKT conditions (6) are necessary and sufficient to represent a global optimum of (5). The market equilibrium is the set of  $x_1, \dots, x_{|\mathcal{P}|}$  that satisfy (6) for all players, where the perp operator  $\perp$  signifies that the product of the constraints on both sides of the operator must equal zero.

$$0 \leq c_p - b - \xi + a \left( x_p + \sum_{k \in \mathcal{P}} x_k \right) + \mu_p \quad (6a)$$

$$\perp \quad x_p \geq 0$$

$$0 \leq -x_p + Q_p^{max} \quad \perp \quad \mu_p \geq 0 \quad (6b)$$

We apply the option to deviate by  $\epsilon_h$  from the stationarity condition (6a), as explained in Section III-B, and use several observations  $h \in \mathcal{H}$ . Each observation differs by realizations of the demand shock  $\xi_h$ . Equation set (7) becomes the inverse equilibrium problem, where the objective function (7a) is to minimize the distance to an equilibrium point considering all observations.

$$\min_{c,a,b,\mu,\epsilon} \|\epsilon\| \quad (7a)$$

$$\text{s.t.} \quad \left( c_p - b - \xi_h + a \left( x_{ph} + \sum_{k \in \mathcal{P}} x_{kh} \right) \right. \quad (7b)$$

$$\left. + \mu_{ph} + \epsilon_{ph} \right) x_{ph} = 0, \quad p \in \mathcal{P}, h \in \mathcal{H}$$

$$0 \leq c_p - b - \xi_h + a \left( x_{ph} + \sum_{k \in \mathcal{P}} x_{kh} \right) \quad (7c)$$

$$+ \mu_{ph} + \epsilon_{ph}, \quad p \in \mathcal{P}, h \in \mathcal{H}$$

$$(-x_{ph} + Q_p^{max}) \mu_{ph} = 0, \quad p \in \mathcal{P}, h \in \mathcal{H} \quad (7d)$$

$$\mu_{ph} \geq 0, \quad p \in \mathcal{P}, h \in \mathcal{H} \quad (7e)$$

2) *Illustrative case study:* To illustrate the inverse Nash-Cournot game, we create a case study where we consider three price-making electricity generators. All have a maximum production of  $Q_p^{max} = 5000 MWh$  and their marginal costs are  $c_1 = 50.0 \text{ €/MWh}$  and  $c_2 = c_3 = 60.0 \text{ €/MWh}$ . Their collective consumers are represented by a linear inverse demand function with slope  $a = 0.01 \text{ €/MWh}^2$  and intercept  $b = 200.0 \text{ €/MWh}$ . We insert these values into the equilibrium problem (6) and solve. The Nash-Cournot equilibrium is  $x_1 = 4250 MWh$  and  $x_2 = x_3 = 3250 MWh$  when the demand shock  $\xi = 0$ .

We solve equilibrium problem (6) a hundred times to produce observations  $\tilde{x}_{1h}$ ,  $\tilde{x}_{2h}$ , and  $\tilde{x}_{3h}$ . Each observation has a different demand shock  $\xi_h$  selected at random from a normal distribution with mean of 0 and standard deviation  $20 \text{ €/MWh}$ . We thus have  $|\mathcal{H}| = 100$  different observations.

The inverse generic Nash-Cournot game (7) takes observations  $\tilde{x}_{1h}$ ,  $\tilde{x}_{2h}$ , and  $\tilde{x}_{3h}$  as parameters and solves for  $c_1, c_2, c_3, a, b, \mu$ , and  $\epsilon$ . We assume that the demand shocks  $\xi_h$  are known and thus parameters as well. Note that this is not a realistic assumption, but prevents noise in the example, which is a topic we consider in Section IV-A4.

The objective value of (7a) becomes  $8 \cdot 10^{-5}$ , so sufficiently small to indicate that the model fits the data. Slope  $a$  is correctly fitted to  $0.01 \text{ €/MWh}^2$ , but some deviation occurs for  $b = 150.0 \text{ €/MWh}$ ,  $c_1 = 0.0$ , and  $c_2 = c_3 = 10.0 \text{ €/MWh}$ . All deviations are fitted exactly  $50.0 \text{ €/MWh}$  less than the original value, so we have a case of scale invariance. Whenever we are dealing with a market, we can use price observations  $\tilde{\lambda}_h$ . We introduce the relationship that the inverse demand function determines price, as shown in (8), as a scaling constraint.

$$\tilde{\lambda}_h = b + \tilde{\xi}_h - a \sum_{k \in \mathcal{P}} \tilde{x}_{kh}, \quad h \in \mathcal{H} \quad (8)$$

When we include (8) to the inverse problem (7), we obtain the same objective value, but parameters fit exactly

to the true value. Hence, we show that if data coincide with the inverse equilibrium model, it fits perfectly.

3) *Fit inverse equilibrium models to other market structures:* The inverse equilibrium approach fits data to models. To illustrate, we fit data from a competitive equilibria to the inverse Cournot model (7). We use 100 observations from when a social planner coordinates all decisions. Table I outlines the results.

TABLE I  
RESULTS OF FITTING PERFECT COMPETITION DATA TO INVERSE COURNOT MODEL.

	True	Without (8)	With (8)
Deviation, $\epsilon$ [€/MWh]	0	208.1	1113.2
Intercept [€/MWh], $b$	200.0	143.6	200.0
Slope, $a$ [€/MWh <sup>2</sup> ]	0.01	0.0067	0.01
Cost gen. 1, $c_1$ [€/MWh]	50.0	0.0	0.0
Cost gen. 2, $c_2$ [€/MWh]	60.0	20.3	17.6
Cost gen. 3, $c_3$ [€/MWh]	60.0	20.3	14.2

In contrast to the previous example, we observe a non-zero deviation. The inverse model does not manage to fit parameters such that the observations become an equilibrium of (7). In other words, the players deviate from their optimal Cournot strategy and a Cournot model is not a good representation of the data.

Table I also shows that the price relationship (8) increases the deviation  $\epsilon$  and thus changes the solution space. It is therefore no longer a scaling constraint. We also note that the fitted parameters do not resemble the true parameters. This example illustrates the strength of inverse equilibrium modeling to test different market structures. It also emphasizes caution towards considering the fitted parameters as true estimations.

4) *Performance under noise:* In general, we cannot prove that inverse equilibrium modeling, as inverse optimization in its canonical form, yield consistent estimators. That is, as the number of observations increase, the fitted parameters will not converge to a true value. If the goal is to estimate parameters, consistency is an important feature. For this reason, we cannot recommend inverse equilibrium as an estimator.

To display the caveat of using inverse equilibrium as an estimator, we solve the generic Nash-Cournot game for  $|\mathcal{H}| = 10, 100, 500$ , and 1000 observations with a known random demand shock. We then add a normally distributed noise with mean 0 and standard deviation  $200 MWh$  to the output of Generator 3. If production with noise exceeds its production limit, we simply set it to  $Q_3^{max}$ .

A consistent estimator would be able to reduce the noise as the number of observations increase and converge to the true value. Table II shows that this is not the case for the inverse equilibrium model. In fact, the fitted parameters show no significant trend and adhere to the randomness of the noise. The total deviation  $\epsilon$  in

Table II shows a steady increase because it gets more terms that deviate. Theoretically, we can observe this from objective (7a) of the inverse equilibrium model. We only minimize the deviation from optimum and have no noise correcting term. With a noise correcting term, the problem becomes non-convex (see [22]) and thus computationally hard to solve.

TABLE II  
PERFORMANCE OF INVERSE COURNOT MODEL WHEN GENERATOR 3 HAS NOISE THAT FOLLOW DISTRIBUTION  $\mathcal{N}(0, 200MWh)$ .

Observations, $ \mathcal{H} $	10	100	500	1000
$\varepsilon$ [€/MWh]	4.85	61.88	294.55	611.33
$b$ [€/MWh]	198.88	200.78	199.21	198.78
$a$ [€/MWh]	0.0099	0.010	0.0099	0.0098
$c_1$ [€/MWh]	50.64	50.33	50.89	51.57
$c_2$ [€/MWh]	60.43	60.38	60.81	61.47
$c_3$ [€/MWh]	60.44	60.87	60.68	61.64

### B. Nash-Cournot equilibrium in power systems

1) *Model formulation:* To illustrate the inverse equilibrium method for power systems, we use the model formulation of [23] that neglects the presence of arbitrageurs. See [23] for the assumptions that provide a unique equilibrium solution. We want to fit demand and supply function parameters to observations. The inverse demand function (9) sets the price  $\lambda_i$  at a particular bus  $i \in \mathcal{N}$ , where  $\mathcal{N}$  is the set of nodes, with respect to total quantity  $q_i$ , slope  $a_i$  and intercept  $b_i$ . Equation (10) denotes the linear marginal cost for a producer  $p$ , where  $x_p$  is its generation.

$$f_i^{-1}(q_i) = \lambda_i(q_i) = b_i - a_i q_i \quad (9)$$

$$MC_p(x_p) = d_p + c_p x_p \quad (10)$$

A profit-maximizing producer  $p$  decides its sales to a particular node  $s_{pi}$  and its generation  $x_p$  according to problem (11). The objective (11a) is to maximize profits, given by the difference between revenue and cost. The cost of using the transmission network,  $w_i$ , is a parameter in problem (11), but we define it later as the dual variable of the market-clearing condition (15). Constraint (11b) enforces a maximum limit on  $x_p$ , while restriction (11c) ensures that sales are equal to generation.

$$\max_{s_{pi}, x_p} \sum_{i \in \mathcal{N}} (b_i - a_i \sum_{k \in \mathcal{P}} s_{ki} - w_i) s_{pi} \quad (11a)$$

$$\text{s.t. } x_p - Q_p^{max} \leq 0, \quad (\alpha_p) \quad (11b)$$

$$\sum_{i \in \mathcal{N}} s_{pi} - x_p = 0, \quad (\beta_p) \quad (11c)$$

$$s_{pi} \geq 0, \quad x_p \geq 0 \quad (11d)$$

The KKT conditions of the producer problem (11) become (12). Notation  $p(i)$  denotes the mapping from producer  $p$  to node  $i$ , i.e. the location of the generator.

$$0 \leq -b_i + a_i(s_{pi} + \sum_{k \in \mathcal{P}} s_{ki}) + w_i + \beta_p \perp s_{pi} \geq 0, \quad i \in \mathcal{N} \quad (12a)$$

$$0 \leq d_p + 2c_p x_p - w_{p(i)} + \alpha_p - \beta_p \perp x_p \geq 0 \quad (12b)$$

$$0 \leq -x_p + Q_p^{max} \perp \alpha_p \geq 0 \quad (12c)$$

$$\sum_{i \in \mathcal{N}} s_{pi} - x_p = 0, \quad \beta_p \in \mathbb{R} \quad (12d)$$

A system operator oversees energy flow while maximizing revenue from grid use, as shown in problem (13), where  $y_i$  is net energy injection at node  $i$ . Constraints (13b) and (13c) guarantee flows within the minimum and maximum limits of line  $l \in \mathcal{L}$ , where  $\mathcal{L}$  is the set of lines. A PTDF matrix determines the flows in the system, where element  $PTDF_{li}$  gives the ratio of flow on line  $l$  caused by power injections at node  $i$ . Although the system operator has an optimization problem, the net injection  $y_i$  is in fact determined by sales and production by the producers, as we show later in the market-clearing condition (15). Consequently, the system operator does not act strategically.

$$\max_{y_i} \sum_{i \in \mathcal{N}} w_i y_i \quad (13a)$$

$$\text{s.t. } -F_l^{cap} - \sum_{i \in \mathcal{N}} PTDF_{li} y_i \leq 0, \quad (\gamma_l^-) \quad l \in \mathcal{L} \quad (13b)$$

$$\sum_{i \in \mathcal{N}} PTDF_{li} y_i - F_l^{cap} \leq 0, \quad (\gamma_l^+) \quad l \in \mathcal{L} \quad (13c)$$

The KKT conditions of the system operator problem (13) are (14):

$$w_i + \sum_{l \in \mathcal{L}} PTDF_{li} (\gamma_l^- - \gamma_l^+) = 0, \quad y_i \in \mathbb{R} \quad i \in \mathcal{N} \quad (14a)$$

$$0 \leq F_l^{cap} + \sum_{i \in \mathcal{N}} PTDF_{li} y_i \perp \gamma_l^- \geq 0, \quad l \in \mathcal{L} \quad (14b)$$

$$0 \leq F_l^{cap} - \sum_{i \in \mathcal{N}} PTDF_{li} y_i \perp \gamma_l^+ \geq 0, \quad l \in \mathcal{L} \quad (14c)$$

Finally, the market-clearing condition (15) states that the net injection for each node must be equal to the difference between sales to the node and its internal production.

$$\sum_{p \in \mathcal{P}} s_{pi} - x_{p(i)} = y_i, \quad w_i \in \mathbb{R} \quad i \in \mathcal{N} \quad (15)$$

Both the producer and system operator problems are concave with affine constraints, so the KKT conditions are necessary and sufficient to represent the global

optimum. The equilibrium problem is to find the set of variables that satisfy (12) for all the players, (14), and (15).

We invert the equilibrium problem to (16) for multiple observations  $h \in \mathcal{H}$  according to the method of Section III-B. Because the producer problem has two decision variables, sales,  $s_{pi}$ , and production,  $x_p$ , it has two stationarity conditions. Consequently, we introduce two sets of deviation variables,  $\epsilon_{pih}^s$  and  $\epsilon_{ph}^x$ , for  $s_{pih}$  and  $x_{ph}$ , respectively. In the example, we weigh the deviations equally.

$$\min_{a,b,c,d,\alpha,\beta,\gamma^-, \gamma^+, w, \epsilon} \|\epsilon\| \quad (16a)$$

$$\text{s.t.} \quad (16b)$$

$$\left( -b_i + a_i(s_{pih} + \sum_{k \in \mathcal{P}} s_{kih}) + w_i \right. \quad (16c)$$

$$\left. + \beta_{ph} + \epsilon_{pih}^s \right) s_{pih} = 0, \quad p \in \mathcal{P}, i \in \mathcal{N}, h \in \mathcal{H}$$

$$0 \leq -b_i + a_i(s_{pih} + \sum_{k \in \mathcal{P}} s_{kih}) + w_i \quad (16d)$$

$$+ \beta_{ph} + \epsilon_{pih}^s, \quad \forall p \in \mathcal{P}, i \in \mathcal{N}, h \in \mathcal{H}$$

$$\left( d_p + 2c_p x_{ph} - w_{p(i)} + \alpha_{ph} \right. \quad (16e)$$

$$\left. - \beta_{ph} + \epsilon_{ph}^x \right) x_{ph} = 0, \quad p \in \mathcal{P}, h \in \mathcal{H}$$

$$0 \leq d_p + 2c_p x_{ph} - w_{p(i)} + \alpha_{ph} - \beta_{ph} + \epsilon_{ph}^x, \quad (16f)$$

$$p \in \mathcal{P}, h \in \mathcal{H}$$

$$(-x_{ph} + Q_p^{max}) \alpha_{ph} = 0, \quad p \in \mathcal{P}, h \in \mathcal{H} \quad (16g)$$

$$w_{ih} + \sum_{l \in \mathcal{L}} PTDF_{l,i} (\gamma_{lh}^- - \gamma_{lh}^+) = 0, \quad i \in \mathcal{N}, h \in \mathcal{H} \quad (16h)$$

$$(F_l^{cap} + \sum_{i \in \mathcal{N}} PTDF_{l,i} y_{ih}) \gamma_{lh}^- = 0, \quad l \in \mathcal{L}, h \in \mathcal{H} \quad (16i)$$

$$(F_l^{cap} - \sum_{i \in \mathcal{N}} PTDF_{l,i} y_{ih}) \gamma_{lh}^+ = 0, \quad l \in \mathcal{L}, h \in \mathcal{H} \quad (16j)$$

$$\alpha_{ph} \geq 0, \quad p \in \mathcal{P}, h \in \mathcal{H} \quad (16k)$$

$$\gamma_{lh}^-, \gamma_{lh}^+ \geq 0, \quad l \in \mathcal{L}, h \in \mathcal{H} \quad (16l)$$

$$y_i, w_i \in \mathbb{R}, i \in \mathcal{N}, \quad \beta_p \in \mathbb{R}, p \in \mathcal{P} \quad (16m)$$

2) *Illustrative case study:* As a case study, we consider the the 6-bus system from [24], as shown in Figure 1. Network flows behave according to the PTDF matrix represented in Table III where we define bus 1 as the hub. The line from bus 1 to 6 has a capacity of 200MW, bus 2 to 5 has 250MW, while the rest are sufficiently high not to limit any flows. Buses 1, 2, and 4 contain price-making producers, while buses 3, 5, and 6 are price-taking consumers. Table IV outlines the intercept and slope of both producer marginal cost and inverse demand.

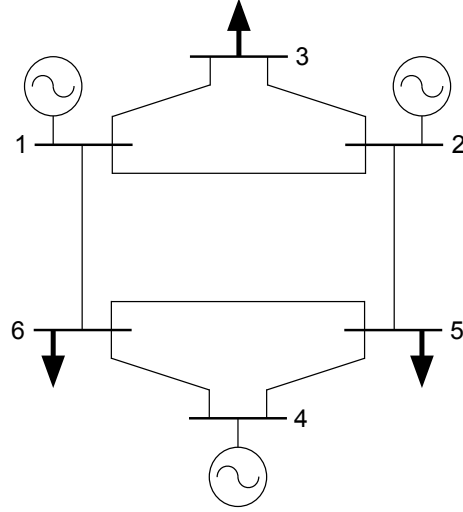


Fig. 1. Illustration of the 6 bus network from [24] used for equilibrium in power systems example.

TABLE III  
PTDF MATRIX OF 6-BUS EXAMPLE.

Line/bus	1	2	3	4	5	6
(1,2)	0	-0.583	-0.292	-0.292	-0.333	-0.25
(1,3)	0	-0.292	-0.646	-0.146	-0.167	-0.125
(1,6)	0	-0.125	-0.063	-0.563	-0.5	-0.625
(2,3)	0	0.292	-0.354	0.146	0.167	0.125
(2,5)	0	0.125	0.063	-0.438	-0.5	-0.375
(4,5)	0	-0.042	-0.021	0.479	-0.167	0.125
(4,6)	0	0.042	0.021	0.521	0.167	-0.125
(5,6)	0	0.083	0.042	0.042	0.333	-0.25

TABLE IV  
FITTED PARAMETERS FROM 6-BUS EXAMPLE. INTERCEPTS IN €/MWh AND SLOPES IN €/MWh<sup>2</sup>.

Bus	Fitted intercept	Fitted slope	True intercept	True slope
1	10.0	0.05	10.0	0.05
2	15.0	0.05	15.0	0.05
3	37.5	0.05	37.5	0.05
4	42.5	0.25	42.5	0.025
5	75.0	0.1	75.0	0.1
6	80.0	0.1	80.0	0.1

We obtain observations by solving the KKT conditions (12) for all the players, (14), and (15) as an equilibrium problem using the input data of the 6-bus example. To get different observations we apply both supply and demand shocks. We assume that all producers have fossil fuel generators with equal emission per unit energy and must pay a carbon price,  $\lambda^{CO_2}$ , for their emissions. We select carbon prices at random from a normal distribution with mean of 10€/MWh and standard deviation 2€/MWh. The carbon price becomes an additional term in the marginal cost,  $MC_p(x_p)$  from (10), of the producers. The demand shock  $\xi_h$  comes from a

normal distribution with mean 0 and standard deviation  $2\text{€}/MWh$ . We assume a sufficiently high generation limit,  $Q_p^{max} = 1000MWh$ , as to not be binding for any of the observations.

We select 100 random carbon prices and demand shocks before solving the equilibrium model to generate observations. The objective value of (16a) becomes slightly above zero at 0.021. Thus we can conclude that the Nash-Cournot model fits the data. Moreover, we see from Table IV that the fitted parameters coincide with the actual variable. In contrast to the example in Section IV-A we have no scale invariance. The data provides sufficient marginal observations to scale the fitted parameters correctly.

## V. COMMENTS ON IMPLEMENTATION

As demonstrated in the examples of Section IV, existing equilibrium models can easily be recast as inverse equilibrium models. Although the fitted parameters in our examples provide good estimates of objective function parameters, we emphasize that the data was generated in a controlled environment. In real applications, the data will be noisy and the results more challenging to interpret. Inverse equilibrium modeling tries to fit a hypothesis, i.e. equilibrium structure, to data. A benefit of this approach is that the inverse equilibrium models are interpretable. While this limits generalization, it enables the modeler to use domain knowledge.

Data from real-world applications are subject to noise. Inverse equilibrium models are unlikely to enjoy as small deviations as our examples. This is expected, as it only shows that an equilibrium structure does not perfectly fit the data. An interesting feature is that we can try different equilibrium set-ups and observe what structure has the least deviation, and thus is the best fit for the data. Note that there may be several reasons for deviations; the model structure may not adhere to the observations, the observations can be noisy or there may be underlying dynamics or costs unobserved by the modeler.

The KKT approach in this paper benefits from the close relationship to existing MCP models applied to power systems. Consequently, the deviations are measured in costs per variable unit, which is less intuitive to interpret than just costs. The VI approach [3], on the other hand, measures deviations in the unit of the objective function. However, this requires a VI representation of the equilibrium problem.

As discussed in Section II, it is important to be cautious when investigating the fitted parameters. They are not representative of characteristics of an underlying population as in econometrics, they are merely the best fit to the data. Estimation of underlying market parameters is an important task for market monitors. For this purpose we recommend consistent estimators established

in the econometric literature. If the reader is interested to try inverse equilibrium approaches in an estimation direction, we refer to [25], [26] and [22], which consider inverse optimization with noisy observations.

Inverse equilibrium modeling is a general approach that can be applied to any equilibrium problem. In this paper we use Cournot models because they are familiar to the power system modeling community. An alternative approach are conjectural variations models (see e.g. [27], [28] and [29]), which are more general. A challenge with inverting for instance the model in [27], is that even if the KKT conditions of the problem are necessary and sufficient, the inverse problem becomes non-convex in parameters. Hence, to make the inverse equilibrium problem convex, we need observations on a parameter in the bilinear term. For more information on estimation of conjectural variations models in power systems we refer to [30].

## VI. CONCLUSION

Inverse equilibrium modeling is a data-driven method that fit parameters of an equilibrium model in order to minimize the deviation from an observation. This paper shows how to use Karush-Kuhn-Tucker (KKT) conditions to invert equilibrium problems. As shown in two applications, a constrained and an unconstrained Nash-Cournot game between power producers, this only requires a small deviation from the original equilibrium problem. Our methodology is thus easy to apply on existing equilibrium models applied to power systems, where working with KKT conditions is prominent. Inverse equilibrium models as shown in this paper can transform into linear programming problems. The method can investigate if data fit a model structure and it has predictive power. However, its estimation is generally inconsistent and econometric approaches are better for this purpose.

## ACKNOWLEDGMENT

The authors would like to thank Paolo Pisciella for support with equilibrium model formulation and GAMS implementation. Gratitude is also extended to four anonymous reviewers for their feedback.

## REFERENCES

- [1] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity Modeling in Energy Markets*. Springer New York, 2013.
- [2] J.-Z. Zhang, J.-B. Jian, and C.-M. Tang, "Inverse problems and solution methods for a class of nonlinear complementarity problems," *Computational Optimization and Applications*, vol. 49, no. 2, pp. 271–297, Jun 2011.
- [3] D. Bertsimas, V. Gupta, and I. C. Paschalidis, "Data-driven estimation in equilibrium using inverse optimization," *Mathematical Programming*, vol. 153, no. 2, pp. 595–633, 11 2015.
- [4] R. K. Ahuja and J. B. Orlin, "Inverse optimization," *Operations Research*, vol. 49, no. 5, pp. 771–783, 10 2001.



- [5] J. Saez-Gallego, J. M. Morales, M. Zugno, and H. Madsen, "A data-driven bidding model for a cluster of price-responsive consumers of electricity," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5001–5011, 11 2016.
- [6] J. Saez-Gallego and J. M. Morales, "Short-term forecasting of price-responsive loads using inverse optimization," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4805–4814, Sep. 2018.
- [7] C. Ruiz, A. J. Conejo, and D. J. Bertsimas, "Revealing rival marginal offer prices via inverse optimization," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3056–3064, 8 2013.
- [8] J. R. Birge, A. Hortaçsu, and J. M. Pavlin, "Inverse optimization for the recovery of market structure from market outcomes: An application to the MISO electricity market," *Operations Research*, vol. 65, no. 4, pp. 837–855, 8 2017.
- [9] R. Chen, I. C. Paschalidis, and M. C. Caramanis, "Strategic equilibrium bidding for electricity suppliers in a day-ahead market using inverse optimization," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 12 2017, pp. 220–225.
- [10] R. Chen, I. C. Paschalidis, M. C. Caramanis, and P. Andrianeis, "Learning from past bids to participate strategically in day-ahead electricity markets," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5794–5806, 2019.
- [11] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *2011 IEEE International Symposium on Intelligent Control*. IEEE, 9 2011, pp. 613–619.
- [12] J. Rust, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, vol. 55, no. 5, pp. 999–1033, 1987.
- [13] S. Borenstein, J. B. Bushnell, and F. A. Wolak, "Measuring market inefficiencies in California's restructured wholesale electricity market," *The American Economic Review*, vol. 92, no. 5, pp. 1376–1405, 2002.
- [14] C. D. Wolfram, "Measuring duopoly power in the British electricity spot market," *American Economic Review*, vol. 89, no. 4, pp. 805–826, September 1999.
- [15] F. A. Wolak, "Identification and estimation of cost functions using observed bid data: An application to electricity markets," no. 8191, March 2001.
- [16] S.-E. Fleten, E. Haugom, A. Pichler, and C. J. Ullrich, "Structural estimation of switching costs for peaking power plants," *European Journal of Operational Research*, 2019.
- [17] D. Ackerberg, L. Benkard, S. Berry, and A. Pakes, "Econometric tools for analyzing market outcomes," in *The Handbook of Econometrics*, J. Heckman and E. Leamer, Eds., 2007, vol. 6A, pp. 4171–4276.
- [18] P. C. Reiss and F. A. Wolak, "Structural econometric modeling: Rationales and examples from industrial organization," in *Handbook of Econometrics*. Elsevier, 1 2007, vol. 6, ch. 64, pp. 4277–4415.
- [19] D.-W. Kim and C. R. Knittel, "Biases in Static Oligopoly Models? Evidence from the California Electricity Market," *Journal of Industrial Economics*, vol. 54, no. 4, pp. 451–470, 12 2006.
- [20] J. Saez-Gallego and J. M. Morales, "Short-term forecasting of price-responsive loads using inverse optimization," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4805–4814, 9 2018.
- [21] S. P. Dirkse and M. C. Ferris, "The path solver: a nonmonotone stabilization scheme for mixed complementarity problems," *Optimization Methods and Software*, vol. 5, no. 2, pp. 123–156, 1995.
- [22] A. Aswani, Z.-J. M. Shen, and A. Siddiq, "Inverse optimization with noisy data," *Operations Research*, vol. 66, no. 3, pp. 870–892, 6 2018.
- [23] B. Hobbs, "Linear complementarity models of Nash-Cournot competition in bilateral and POOLCO power markets," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 194–202, 5 2001.
- [24] H.-P. Chao and S. C. Peck, "Reliability management in competitive electricity markets," *Journal of Regulatory Economics*, vol. 14, no. 2, pp. 189–200, 1998.
- [25] A. Aswani, "Statistics with set-valued functions: applications to inverse approximate optimization," *Mathematical Programming*, Mar 2018.
- [26] J. Thai and A. M. Bayen, "Imputing a variational inequality function or a convex objective function: A robust approach," *Journal of Mathematical Analysis and Applications*, vol. 457, no. 2, pp. 1675–1695, 1 2018.
- [27] C. Day, B. Hobbs, and Jong-Shi Pang, "Oligopolistic competition in power networks: a conjectured supply function approach," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 597–607, 8 2002.
- [28] J. Liu, T. Lie, and K. Lo, "An empirical method of dynamic oligopoly behavior analysis in electricity markets," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 499–506, 5 2006.
- [29] J. Lagarto, J. de Sousa, and A. Martins, "Application of a conjectural variations model to analyze the competitive behavior in the Iberian electricity market," in *2011 8th International Conference on the European Energy Market (EEM)*. IEEE, 5 2011, pp. 857–862.
- [30] A. García-Alcalde, M. Ventosa, M. Rivier, A. Ramos, and G. Relación, "Fitting electricity market models: A conjectural variations approach," in *14th PSCC*, 2002.



**Simon Risanger** is a PhD student at the Department of Industrial Economics and Technology Management, NTNU. He holds an M. Sc. in Energy and Environmental Engineering from the same university. His research interests are the application of mathematical programming towards power system challenges and corresponding analysis.



challenges to managing the uncertainty of energy prices and other related factors.

**Stein-Erik Fleten** is professor of Operations Research in the Department of Industrial Economics and Technology Management, NTNU. His research interests lie in the domains of energy analytics and stochastic programming. These interests concern development and implementation of financial engineering models and methods for engineering-economic analysis of investment in, and operations of energy businesses. He is particularly interested in applications where there are



development with applications in networked industries such as energy and transportation.

**Steven A. Gabriel** is a Full Professor of Operations Research in the Department of Mechanical Engineering and in the Applied Mathematics, Statistics, and Scientific Computation Program at the University of Maryland (College Park). In addition he is an adjunct Professor in the Department of Industrial Economics and Technology Management at NTNU. His interests and expertise are in optimization, equilibrium, and game theory modeling and related algorithm development