SIAM J. OPTIM.
Vol. xx, No. x, pp. x–x

REGULARIZED EQUILIBRIUM PROBLEMS WITH EQUILIBRIUM CONSTRAINTS WITH APPLICATION TO ENERGY MARKETS *

JUAN PABLO LUNA †, CLAUDIA SAGASTIZÁBAL ‡, JULIA FILIBERTI §, STEVE A. GABRIEL ¶, AND MIKHAIL V. SOLODOV‖

Abstract. Equilibrium problems with equilibrium constraints are appropriate modeling formulations in a number of important areas, such as energy markets, transportation planning, and logistics. These models often correspond to bilevel games, in which certain dual variables, representing the equilibrium price, play a fundamental role. We consider multi-leader single-follower equilibrium problems having a linear program in the lower level. Because in this setting the lower-level response to the leaders' decisions may not be unique, the game formulation becomes ill-posed. We resolve possible ambiguities by considering a sequence of bilevel equilibrium problems, endowed with a special regularization term. We prove convergence of the approximating scheme. Our technique proves useful numerically, over several instances related to energy markets. When using PATH to solve the corresponding mixed-complementarity formulations, we exhibit that, in the given context, the regularization approach computes a genuine equilibrium price almost always, while without regularization the outcome is quite the opposite.

Key words. equilibrium problems with equilibrium constraints; multi-leader single-follower games, dual-primal regularization; energy markets.

AMS subject classifications. 90C33, 91A10, 49M05, 49M15

1. Introduction.

Multi-agent hierarchical decision-making can be modelled by equilibrium problems with equilibrium constraints (EPEC). This is a suitable setting for competitive environments, with sales and purchases that involve bids whose payment depends on balancing supply and demand. The price that clears the market is an important output in such transactions. For the energy sector, in particular, the EPEC corresponds to a multi-leader single-follower bilevel game. In the upper level, agents define their bids of price and generation, subject to operational constraints. The leaders aim to maximize their individual profit. The social planner problem in the lower level, shared by all the leaders, is solved by the independent system operator (ISO). The ISO collects bids from all the leaders and defines the optimal dispatch and its price. This primal-dual pair solves a linear program that minimizes social cost over a set of system-wide operational requirements. Multipliers associated to balancing constraints, ensuring that the dispatched generation meets the demand, provide the clearing price.

Our study is motivated by pitfalls that arise when prices are determined as in the energy example. Efficient market mechanisms should aim to maximize the total welfare of producers and consumers, respectively represented by the leaders and the follower. However, the efficacy of Lagrange multipliers for pricing purposes has been questioned in the literature; [14]. It was observed that Lagrange multipliers of multi-leader single-follower games sometimes result in overpricing and, hence, in lesser social welfare. Remark 2.1 explains this issue in more detail.

We therefore consider noncooperative bilevel games involving several leaders and a single follower, with a focus on the dual output of the process. Each leader's objective function depends on the product of the follower's primal and dual solution to a linear complementarity problem that is parameterized by all the leaders' decisions.

The considered setting, with only one follower player whose optimization problem is a linear

* The authors acknowledge the support of TACEMM, the Transatlantic Consortium of Energy Markets and Modeling. Research of the second author is partly funded by CEPID CeMEAI, CNPq Grant 306089/2019-0, and PRONEX–Optimization. The second author also would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme "The Mathematics of Energy Systems", supported by EPSRC grant number EP/R014604/1. The fifth author is supported by CNPq Grant 303913/2019-3, by FAPERJ Grant E-26/202.540/2019, and by PRONEX–Optimization.

† COPPE, Universidade Federal do Rio de Janeiro, Brazil (jpluna@po.coppe.ufrj.br).

‡ IMECC, Unicamp, Brazil (sagastiz@unicamp.br).

§ University of Maryland, College Park, Maryland, USA (julia.filiberti@gmail.com).

¶ University of Maryland, College Park, Maryland, USA; and Norwegian University of Science and Technology, Trondheim, Norway (sgabriel@umd.edu).

‖ IMPA – Instituto Nacional de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil (solodov@impa.br).

program, is theoretically and numerically challenging. One reason is that the lower-level problem is shared by all the leaders. Another point is certain ambiguity in the interaction between the two levels. This issue, already present in bilevel programming, refers to situations in which the response of the lower level is not unique. To eliminate the ambiguity, many works assume the lower objective function to be strictly convex, as in [17]; see also the final comments in [2, Sec. 2]. While this assumption makes the lower-level response unique regarding the primal variable (the dispatch in the energy example), it still leaves pending the ambiguity for the dual variables (the important clearing prices). A similar difficulty regarding nonuniqueness of prices is pointed out with simple stochastic games with risk-averse leaders in [10]. Finally, the assumption of strict convexity is not always realistic. For large systems (like Brazil's energy market, as one example), it is sound to formulate the lower-level problem as a linear program, a model which precludes any hope of strict convexity. We discuss this point and relate with the existing literature in § 2.3 below.

Ill-posedness is a drawback inherent to EPEC models involving multi-agent hierarchical decision making. Yet, bilevel multi-leader single-follower games represent the market behavior well in many applications. In order to mitigate the model's negative aspects, we introduce a sequence of approximating EPECs, depending on a certain regularization parameter that is driven to zero.

Rather than forcing strict convexity, so that the follower gives a unique primal output, we allow the lower level to maintain its linear programming problem structure. The scheme penalizes the dual of the follower's problem by means of a fairly general class of regularizing functions, that includes ℓ_r-norms. When $r = 1$ or $+\infty$, the regularized lower-level problem remains a linear program, as in the original EPEC. To each value of the regularization parameter corresponds an EPEC whose regularized lower-level problem yields a specific primal-dual value in the (possibly set-valued) follower's response. With our approach, social welfare is maximized when the regularization parameter goes to zero. This can be seen in Theorems 4.3 and 3.3 below, where it is shown that, in the limit, the regularized lower level delivers a solution to the original EPEC whose dual component has minimal norm.

The rest of this work is organized as follows. In Section 2 we start with some notation and introduce a simple EPEC model for the energy market that is used throughout to motivate and illustrate the development. Section 3 is devoted to our regularization scheme. The methodology is explained first in Section 3.1 for the simple energy EPEC, using the dual formulation. Section 3.2 focuses on general EPECs, formed by several leaders solving bilevel programs with a single follower solving a linear programming problem in the lower level. The dual regularization approach is reformulated for a general EPEC from a primal point of view also in Section 3.2. Section 4 discusses various properties of the regularized energy problem. Section 5 extends the regularization technique to a multi-market model, which amounts to the follower solving a variational inequality instead of a linear program; see Section 5.1. Finally, to illustrate the type of numerical difficulties that must be addressed when solving EPECs, Section 5.2 reports on several experiments with the PATH solver [6] using GAMS, for models with one and two markets. The benchmark provides insights on how our approach helps in obtaining solutions to the complementarity system that are equilibria for the EPEC. In the last section, some concluding remarks are provided.

2. On energy markets.

We start with a particular EPEC model that corresponds to the Brazilian energy system described in [3]. In this stylized static version, there is no uncertainty neither dynamic relations. The model will be used to illustrate the proposed methodology.

2.1. The setting.

Suppose the market is composed by N leaders who bid on energy for one time period. For the ith agent, $i \in \{1, \ldots, N\}$, the *bid* $0 \leq (p_i, g_i) \in \mathbb{R}^2$ consists, respectively, of a selling price and the amount of energy that the agent is willing to generate for this price. The unit cost of generation is φ_i, and upper bounds for the bid are denoted by $(p_i^{\max}, g_i^{\max}) > 0$.

The ISO is the single follower, who receives all the bids, that is $(p, g) \in \mathbb{R}^{2N}$, where $p = (p_1, p_2, \ldots, p_N)$ and $g = (g_1, g_2, \ldots, g_N)$. Taking into account the bids (p, g) as parameters, and considering the social cost and system-wide constraints related to demand and network, the ISO decides the generation *dispatch* $l = (l_1, l_2, \ldots, l_N) \in \mathbb{R}^N$ and the *market price* $P \in \mathbb{R}$, to be paid for each unit of energy (the same price for all the agents, pricing is uniform).

Regarding the lower-level output l, it is clearly not possible for the ISO to dispatch amounts larger than the quantity bids ($l_i \leq g_i$ for all i). Partial dispatch is possible ($l_k < g_k$ for some $k \in \{1, \ldots, N\}$), for example to clear the market. The market regulator determines when to partially dispatch an agent and how to treat tie-breaks.

Regarding the price P, its value is given by a function defined by the market rules. An ISO that prefers not to over-pay would set

$$P(p, g, l) := \max \{p_j : l_j > 0 \,, j = 1, \ldots, N\} \,. \tag{2.1}$$

In this case, the clearing price is the highest value of the bids, among all the prices bid by agents who were dispatched by the ISO. The dependence of the price function $P(\cdot)$ on the generation bid g is indirect, through the dispatch. A more accurate writing (notationally more heavy) would be $P = P(l(p, g))$. Finally, notice that the rule (2.1) is implicit, it depends on how the leaders and the follower interact, until an equilibrium is attained and the market is cleared.

Figure 1 represents a market with two leaders. The arrows in the diagram display the output of one level that enters as a parameter in the optimization problem of the other level.

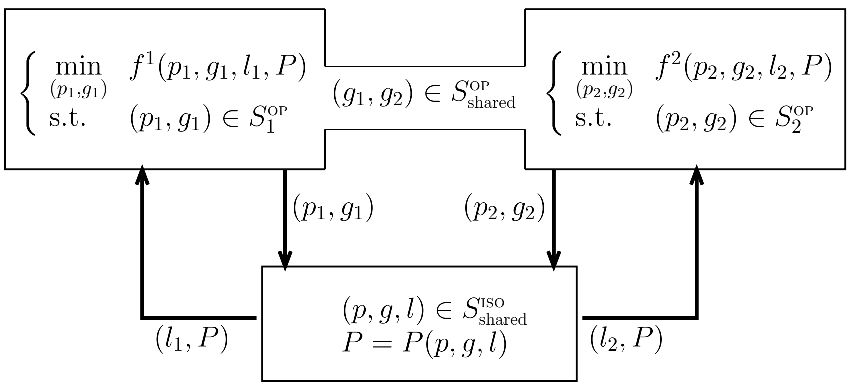

Fig. 1: Market with two agents and the ISO. In the upper level, the leaders decide their bids on price and generation, taking into account operational constraints (sets S_i^{OP} for $i = 1, 2$). Some operational constraints can be shared (set $S^{\text{OP}}_{\text{SHARED}}$). The leaders' objective function is parameterized by the dispatch and market price, which is the follower's output. The decision on dispatch and price in the lower level is done by the ISO, taking into account system constraints (set $S^{\text{ISO}}_{\text{shared}}$), and having the bids as parameters.

In Figure 1, the ith agent determines bids by solving a bilevel minimization problem, using an objective function that represents a disutility, such as the negative of profit,

$$f^i(p_i, g_i, l_i, P) = \varphi_i l_i - l_i P \,,$$

or some convex function to hedge against downside profit risk. The feasible region is defined by three sets, S_i^{OP}, $S^{\text{OP}}_{\text{shared}}$ and $S^{\text{ISO}}_{\text{shared}}$, with different structure. Operational constraints, depending on the technology employed to generate energy, are included in the first set, S_i^{OP}, that is endogenous, specific to each agent. For example,

$$S_i^{\text{OP}} := \{(p, g) : \varphi_i \leq p \leq p_i^{\max} \,, 0 \leq g \leq g_i^{\max}\} \,. \tag{2.2}$$

The lower bounds for price and generation are the marginal cost of generation and zero, respectively. Other choices are possible. In models with temporal dynamics, ramping and warm-up/shutdown requirements constrain the generation. In order to be operational at full capacity at peak times (when prices are higher), at times preceding the peak leaders bid low prices, to encourage the ISO to dispatch their utility. This is achieved by setting the price lower bound below cost.

The second feasible set, $S^{\text{OP}}_{\text{shared}}$, includes operational constraints that are *shared* by several leaders. This situation arises for a group of agents generating hydropower from a set of cascaded reservoirs. Like the first set of constraints, this second set is explicitly described by equality and inequality constraints involving components of the generation vector g (e.g., stream-flow balance constraint, exchange limits, etc.).

The third set, $S^{\text{ISO}}_{\text{shared}}$, is shared by all the agents, leaders and follower. Market-clearing and network system-wide constraints are part of this set, that we write abstractly as follows:

$$S^{\text{ISO}}_{\text{shared}} := \{(p,g,l) : l \in D(p,g)\} \ , \tag{2.3}$$

where the *dispatch multifunction* $D : \mathbb{R}^N \times \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ returns, for a given bid (p,g), a set $D(p,g)$ of dispatched quantities. In a manner similar to the pricing rule (2.1), the dispatch multifunction is available only implicitly, as it is the outcome of the ISO solving an optimization problem parameterized by the leaders' decisions. Typically, the ISO lower-level problem minimizes the consumers' expenses or maximizes social welfare, as in problem (2.5) below.

Remark 2.1 (Social cost and clearing prices from Lagrange multipliers). Even though very intuitive, (2.1) turns out to be difficult to implement. Because the rule is not explicit, to solve the EPEC in practice, the function $P(\cdot)$ therein is replaced by the Lagrange multiplier corresponding to the market clearing constraint in $S^{\text{ISO}}_{\text{shared}}$, denoted below by π. In many situations, however, such replacement provides only an upper bound. When this happens, the EPEC price at equilibrium is larger than the right-hand side in (2.1). While it is reasonable for dispatched generators to be remunerated at least the price they bid, as stated by (2.1), it is not sound for the ISO to settle the price at values higher than the largest bidding price. From the leaders' side, higher prices are better, but not from the side of the consumers. If social welfare is reduced, the ISO is not fulfilling its role and the market is said to be inefficient.

Overpricing is not uncommon in multi-leader single-follower games. In the disutility, the term $l_i P$ is the agent's remuneration and the Lagrange multiplier replacement yields

$$f^i(p_i, g_i, l_i, \pi) = \varphi_i l_i - l_i \pi \, . \tag{2.4}$$

Since the upper-level problems maximize profit (or minimize disutility), in situations in which multipliers are not unique, the leaders will always favor larger values of π, making the price replacement an overestimation of the rule (2.1). □

In order to deal with the undesirable overpricing, we regularize the dual of the follower problem. By this token, when the regularization parameter goes to zero, the Lagrange multiplier in the limit will be the price signal with minimum norm, that is, corresponding to the right-hand side in (2.1). We explain the technique first for an illustrative formulation representing the ISO decision process.

2.2. A simple ISO problem. Having the bidding prices p and quantities g as parameters, the ISO minimizes the total expense, $p^\top l$, in a manner that satisfies the demand $d \in \mathbb{R}$:

$$\left\{ \begin{array}{lll} \min\limits_{l} & p^\top l & \\ \text{s.t.} & 0 \leq l & \\ & l \leq g & (\lambda) \\ & \sum\limits_{j=1}^{N} l_j = d & (\pi) \, . \end{array} \right. \tag{2.5}$$

The right-most variables in parentheses, λ and π, denote the Lagrange multipliers associated to the constraints on the left of them. The meaning of these multipliers (the dual variables) as *shadow prices* is well-known. An economic interpretation of the optimal multiplier λ^* as *marginal rent* will be discussed in Section 3. The optimal multiplier π^* is the *marginal price*, representing the infinitesimal change in the expense arising from an infinitesimal change in the demand.

The simple formulation (2.5) has optimal primal and dual solutions that can be expressed in a closed form. To do so, the statement and its proof below use the following notation:

– the agent whose bid completes the demand is called *marginal* and is denoted by index $j_{\mathtt{mg}}$:

$$\sum_{k<\mathtt{mg}} g_{j_k} < d \leq \sum_{k\leq \mathtt{mg}} g_{j_k}\,.$$

– The N-dimensional vector with all components equal to 1 is denoted by $\mathbf{1}$.
– For two vectors of matching dimensions, $u \perp v$ means that $u^\top v = 0$.
– For a scalar c, we write $[c]^+ = \max\{0, c\}$.

Proposition 2.2 (Solution to ISO problem (2.5)). *Re-ordering the indices if necessary, suppose in problem (2.5) the bidding prices satisfy the relation $p_{j_1} \leq p_{j_2} \leq \cdots \leq p_{j_N}$, and set $p_{j_{N+1}} := \infty$. The following holds:*

(i) The dispatch defined by

$$l^*_j = \begin{cases} g_j & \text{if } j \in \{j_k : 1 \leq k < \mathtt{mg}\} \\ d - \sum\limits_{k<\mathtt{mg}} l^*_{j_k} = d - \sum\limits_{k<\mathtt{mg}} g_{j_k} & \text{if } j = j_{\mathtt{mg}} \\ 0 & \text{otherwise,} \end{cases}$$

solves (2.5) and, hence, $l^* \in D(p, g)$, the set-valued mapping from (2.3).

(ii) The optimal Lagrange multipliers associated to the demand constraint in (2.5) are

$$\Pi^* = \begin{cases} \{p_{j_{\mathtt{mg}}}\} & \text{if } l^*_{j_{\mathtt{mg}}} < g_{j_{\mathtt{mg}}} \\ \left[p_{j_{\mathtt{mg}}}, \min\limits_{k>\mathtt{mg}, g_{j_k}>0} p_{j_k}\right] & \text{if } l^*_{j_{\mathtt{mg}}} = g_{j_{\mathtt{mg}}}, \text{ and } j_{\mathtt{mg}} < N \\ \left[p_{j_{\mathtt{mg}}}, +\infty\right) & \text{if } l^*_{j_{\mathtt{mg}}} = g_{j_{\mathtt{mg}}}, \text{ and } j_{\mathtt{mg}} = N\,. \end{cases}$$

possibly non-unique.

(iii) For each π^* in the set-valued mapping Π^*, the vector λ^{π^*} with components

$$\lambda^{\pi^*}_j = [\pi^* - p_j]^+ \text{ for } 1, \ldots, N$$

is an optimal multiplier for the capacity constraint in (2.5).

Proof. The vector l^* is clearly feasible for (2.5) and $\lambda^{\pi^*} \geq 0$. Also, we have that

$$\lambda^{\pi^*}_j = \begin{cases} \pi^* - p_j & \text{if } j \in \{j_k : 1 \leq k \leq \mathtt{mg}\} \\ 0 & \text{otherwise} \end{cases}$$

and if $l^*_{j_{\mathtt{mg}}} < g_{j_{\mathtt{mg}}}$, then $\pi^* = p_{j_{\mathtt{mg}}}$ and so $\lambda^{\pi^*}_{j_{\mathtt{mg}}} = 0$. Taking into account all these relations and the definition of l^*, yields the complementarity relation

$$0 \leq g - l^* \perp \lambda^{\pi^*} \geq 0.$$

On the other hand, note that

$$p + \lambda^{\pi^*} - \pi^*\mathbf{1} = \left([\pi^* - p_j]^+ - (\pi^* - p_j)\right)_{j=1}^N \geq 0$$

and, since $l^*_{j_k} = 0$ for $k > \mathtt{mg}$, we have also that

$$0 \leq l^* \perp p + \lambda^{\pi^*} - \pi^*\mathbf{1} \geq 0\,,$$

which shows that the primal-dual point $(l^*, \pi^*, \lambda^{\pi^*})$ is optimal for (2.5). Finally, taking another pair of multipliers $(\bar\pi, \bar\lambda)$, since $l^*_{j_{\mathtt{mg}}} > 0$, from $l^*_{j_{\mathtt{mg}}}(p_{j_{\mathtt{mg}}} + \bar\lambda_{j_{\mathtt{mg}}} - \bar\pi) = 0$, we have that $0 \leq \bar\lambda_{j_{\mathtt{mg}}} \leq \bar\pi - p_{j_{\mathtt{mg}}}$. Also, when $l^*_{j_{\mathtt{mg}}} < g_{j_{\mathtt{mg}}}$ it holds that $\bar\lambda_{j_{\mathtt{mg}}} = 0$ which implies $p_{j_{\mathtt{mg}}} = \bar\pi$. In case $l^*_{j_{\mathtt{mg}}} = g_{j_{\mathtt{mg}}}$, and $k > \mathtt{mg}$, with $g_{j_k} > 0$. Since $l^*_{j_k} = 0$ and $(g_{j_k} - l^*_{j_k})\bar\lambda_{j_k} = 0$ we have that $\bar\lambda_{j_k} = 0$ and using that $p_{j_k} + \bar\lambda_{j_k} - \bar\pi \geq 0$ yields $p_{j_k} \geq \bar\pi$, as stated. ∎

Remark 2.3 (On the lack of uniqueness for the dispatch). By Proposition 2.2, whenever several agents bid the same price, and this happens to be the highest dispatched price, the equilibrium price will be the same, but not the dispatch, because $p_{j_{\mathtt{mg}}}$ is associated with different generation offers $g_{j_{\mathtt{mg},1}}, g_{j_{\mathtt{mg},2}}, \ldots$. The dispatch function (2.3) is $D^1(p,g) = \{l^* \text{ solving (2.5)}\}$, a set-valued mapping (that is, there is more than one minimizer l^*). Incidentally, this is the reason for writing (2.3) as an inclusion. This situation creates an indifference set for the ISO regarding the optimal dispatch, any distribution of the marginal dispatch $d - \sum_{k<\mathtt{mg}} g_{j_k}$ among the marginal agents yields the same output from the ISO point of view. The simple model (2.3) has no transmission constraints, so all the different $g_{j_{\mathtt{mg},l}}$ are the same when it comes to satisfaction of the demand. If there is a network, indifference sets appear when two agents inject power in the same bus and bid the same marginal price. Since the ISO problem is part of an EPEC, we observed that this phenomenon of indifference leads to numerical difficulties in the solution process.

To get a unique primal solution from the ISO, the authors of [1] impose a so-called *equity property* which would amount to include in (2.5) constraints of the form

$$p_i = p_j \Longrightarrow l_i = l_j \text{ for } i,j = 1,\ldots,N\,.$$

In the market considered by [1], the leaders' problems have no operational constraints, their bid is only on prices. The equity constraint can be considered in this case, because the ISO decides the dispatch without taking into account the generation capacity of the leaders. The same technical artifact is not applicable in our setting, with bids of price and generation. □

The equilibrium problem of interest is the EPEC that results from simultaneously considering all the agents' problems. Suppose for simplicity that there are no coupling operational constraints ($S^{\text{OP}}_{\text{shared}}$ is void). If the ISO behavior is given by (2.3) and (2.5), we consider the following EPEC:

$$\begin{array}{c}\text{Find an}\\ \text{equilibrium,}\\ \text{solving}\\ \text{for } i=1,\ldots,N\\ \text{the } \textbf{bilevel}\\ \text{problems}\end{array}\quad \left\{\begin{array}{ll}\min\limits_{g_i,p_i,l_i} & f^i(p_i,g_i,l_i,\pi)\\ \text{s.t.} & 0 \le g_i \le g_i^{\max}\\ & \varphi_i \le p_i \le p_i^{\max}\\ & l \in \arg\min(S^{\text{ISO}}_{\text{shared}}) \;=\; \left\{\begin{array}{lll}\min\limits_{l} & p^\top l & \\ \text{s.t.} & 0 \le l & \\ & l \le g & (\lambda)\\ & \sum\limits_{j=1}^N l_j = d & (\pi)\end{array}\right\}\end{array}\right. .\tag{2.6}$$

To handle the difficult implicit constraint $S^{\text{ISO}}_{\text{shared}}$, we write the dual of the follower problem (2.5):

$$\left\{\begin{array}{ll}\max\limits_{\pi,\lambda} & \pi d - \lambda^\top g\\ \text{s.t.} & \pi - \lambda_j \le p_j,\ j=1,\ldots,N\\ & \lambda \ge 0\,,\end{array}\right.\tag{2.7}$$

and replace the problem by its optimality conditions. By primal-dual feasibility and strong duality,

$$\begin{array}{c}\text{Find an}\\ \text{equilibrium,}\\ \text{solving}\\ \text{for } i=1,\ldots,N,\\ \text{the } \textbf{bilinear}\\ \text{problems}\end{array}\quad \left\{\begin{array}{ll}\min\limits_{g_i,p_i,l_i,\pi,\lambda} & f^i(p_i,g_i,l_i,\pi)\\ \text{s.t.} & 0 \le g_i \le g_i^{\max}\\ & \varphi_i \le p_i \le p_i^{\max}\\ & 0 \le l \le g\\ & \sum\limits_{j=1}^N l_j = d\\ & \pi - \lambda_j \le p_j,\ j=1,\ldots,N\\ & \lambda \ge 0\\ & p^\top l = \pi d - \lambda^\top g\,.\end{array}\right.\tag{2.8}$$

This equivalent formulation: eliminates the bilevel setting, at the expense of adding bilinear terms in the constraints, that are not simple to tackle, but have the merit of being explicit.

2.3. On related models and solution methods. Research on equilibrium problems is very rich. We mention mostly work on energy markets, without any pretension of thoroughness. Initial mathematical programming models [11] evolved to more sophisticated equilibrium problems, cast as Nash games, complementarity formulations, or EPECs [16], [9], [17]. The last work is a multi-leader single-follower model that assumes strict convexity in the lower level. This assumption appears frequently in the literature for equilibrium problems with many leaders and one follower. For optimality conditions, it can be traced back to [29]. Later it was assumed by [7] and [12]. Additional works having a strictly convex follower are listed in Section 2 of [2].

Multi-leader multi-follower and two-stage non-cooperative games are studied in [20] and [26], respectively. Multi-agent hierarchical games under uncertainty are the topic of [4]. The taxonomy of bilevel convex games therein considers a unifying formulation that again requires strict convexity in the case of a single follower (cf. the lower-level problem in [4, 2.3(a)]).

As explained in the introduction, strict convexity is not a suitable assumption for the follower in large hydro-dominated markets (like Brazil's, as one example). A linear programming lower-level problem is more realistic in this case, see [3]. In the work [1], some bids of price can be linear, as in the objective function in (2.5). But, as explained in Remark 2.3, the equity property used as a replacement of strict convexity is not applicable in our setting.

Regarding computational schemes, we can mention [8] for a bilevel approach with potentially discrete variables, [30] for stochastic variational inequalities, and [24] and [23] for risk-neutral and risk-averse stochastic games and relations with complementarity formulations. Finally, for a discussion on non-unique EPEC multipliers we refer to [22]; see also the analysis of consistent Lagrange multipliers in [20].

3. Regularizing the lower-level problem. We modify the dual lower-level problem (2.7) with a particular regularization term, that is afterwards interpreted in the initial primal problem. This justifies the naming "dual-primal" for the proposed approach.

For the energy model the procedure corresponds to adopting the viewpoint of a *dual* ISO, that could be thought as being more concerned with prices than with dispatch (a dispatch-oriented ISO would directly solve the primal problem). If the clearing price is the demand multiplier, then any dispatched agent that bids below the market price has a positive λ_j in (2.7) that represents an inframarginal rent. The wording *rent*, or surplus, refers to an amount that is received without any effort (revenue, by contrast, involves some work of the agent, to generate energy that may be dispatched). The rent λ_j is positive only when $p_j < \pi$, that is, when the agent gets paid more than the bid. Generators typically rely upon such rent to cover fixed costs. There is overpricing when the marginal agent has a positive marginal rent ($\lambda_{j_{\mathtt{mg}}} > 0$). The phenomenon is avoided with our regularization scheme, thanks to the penalization term, that gives preference to lower values of the marginal rent.

3.1. Regularized dual ISO problem of the energy EPEC. By Proposition 2.2, all the dispatched agents except the marginal ones are dispatched at their bidding level $l_i^* = g_i$, so for those agents it holds that $\lambda_i g_i = \lambda_i l_i^*$. If for a marginal agent the dispatch satisfies $l_{j_{\mathtt{mg}}}^* < g_{j_{\mathtt{mg}}}$, the rent will be zero, because $\pi^* = p_{j_{\mathtt{mg}}}$. Moreover, since in the dual problem (2.7) the objective function is $\pi d - \lambda^\top g$, the dual ISO ends in fact maximizing the overall payment to the agents net from any rent. On the other hand, the unfavorable situation $l_{j_{\mathtt{mg}}}^* = g_{j_{\mathtt{mg}}}$ leads to a positive rent for the marginal agent and an overall increase in the rent of all the other dispatched agents.

In order to control the rent, instead of solving (2.7), we define a *regularized* problem for the dual ISO that discourages large values of marginal rent. This is done by penalizing λ through a convex function $h(\lambda)$ satisfying $h(\lambda) > 0$ for $\lambda > 0$, and $h(0) = 0$.

Given a penalty parameter $\beta \geq 0$, the new dual ISO problem is

$$(3.1)_\beta \qquad \begin{cases} \max\limits_{\pi,\lambda} & \pi d - \lambda^\top g - \beta h(\lambda) \\ \text{s.t.} & \pi - \lambda_j \leq p_j\,, \text{ for } j = 1, \dots, N \\ & \lambda_j \geq 0\,, \text{ for } j = 1, \dots, N\,. \end{cases}$$

Notice that when taking $\beta = 0$, problem $(3.1)_0$ recovers the original dual linear program (2.7).

Regarding assumptions on the regularizing function, they will be stated as needed. For the moment we just review briefly two classical concepts from Convex Analysis: the *subdifferential* and the *Fenchel conjugate*; more detailed explanations can be found in [15].

Given a convex function $h : \mathbb{R}^N \to \mathbb{R}$ and $x \in \mathbb{R}^N$, the subdifferential of h at x is the set

$$\partial h(x) = \{s : h(y) \geq h(x) + s^\top (y - x) \quad \forall\, y\}\,. \tag{3.2}$$

The elements of $\partial h(x)$ are called subgradients of h at x.

The Fenchel conjugate $h^* : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty, \infty\}$ of h is defined by

$$h^*(s) = \sup_{x \in \mathbb{R}^N} \{x^\top s - h(x)\}. \tag{3.3}$$

It is well-known that the subgradient generalizes the concept of gradient, when the convex function h is not differentiable (if h is differentiable, the set is a singleton − the usual derivative). On the other hand, the Fenchel conjugate h^* is a convex function with an interesting economic interpretation. Suppose $h(x)$ represents the cost of production of a good x that can be sold at a price s. Then, the Fenchel conjugate at s defined in (3.3) represents the optimal profit that can be achieved for the price s by choosing the quantity x to be produced and sold. Finally, when h represents a cost, the subdifferential of h at x is the set of all the prices that ensure an optimal profit when x is the production level.

These concepts are used to express the regularized problem in a primal form, by writing the dual problem to $(3.1)_\beta$. We now explain the process for a multi-leader single-follower setting that extends the simple energy problem to a more general EPEC given in (3.4) below.

3.2. Dual and primal regularized formulations for general EPECs. The regularization scheme ($(3.1)_\beta$), stated for the particular EPEC instance represented by the energy problem, is applicable in a more general setting. We adopt a compact notation, in which each agent i has variables x^i, and minimizes a function f^i that depends on the follower's primal and dual decisions, denoted by y and μ, respectively. Operational constraints in the upper level are denoted by $G^i(x^i) \leq 0$. Gathering all the agents decisions in the vector x, the lower-level problem solved by the follower is an abstract linear program, parameterized by x. Accordingly, consider the general EPEC that results from agents $i = 1, \dots, N$ solving the following problem:

$$\left\{\begin{array}{ll} \min\limits_{x^i} & f^i(x^i, y, \mu) \\ \text{s.t.} & G^i(x^i) \leq 0 \\ & y \text{ solves} \quad \left\{\begin{array}{ll} \min\limits_{y} & c(x)^\top y \\ \text{s.t.} & y \geq 0 \\ & By = b(x) \quad (\mu)\,. \end{array}\right. \end{array}\right. \tag{3.4}$$

To see that the energy problem (2.6) is a particular instance of (3.4), it suffices to make the following identifications for the upper- and lower-level variables:

$$x^i = (p_i, g_i)\,, \quad \text{and} \quad y = (\ell, w) \text{ for a slack variable } w \geq 0.$$

Identification for the ith agent's objective and constraints functions are straightforward, while the follower problem amounts to taking

$$c(x) = (p, 0)\,, \quad B = \begin{bmatrix} -I_N & -I_N \\ \mathbf{1}^\top & 0 \end{bmatrix}, \quad \text{and} \quad b(x) = \begin{pmatrix} -g \\ d \end{pmatrix}.$$

Note in passing that, thanks to the affine constraint $By = b(x)$, the general setting (3.4) allows to handle an ISO dealing with network constraints, modeled in a DC formulation.

In order to regularize only some of the multipliers (as with the marginal rent in the ISO problem), we make use of a row-matrix P to project the dual variable μ onto some of its components (for the stylized problem, P has null components except for those indices corresponding to λ, where the entry is 1, so that $P\mu = \lambda$). As a result, to the linear program

$$\begin{cases} \min\limits_{y\geq 0} & c^\top y \\ s.t. & By = b\,, \end{cases}$$

we associate the regularized dual

$$\text{(3.5)} \qquad \begin{cases} \max\limits_{\mu} & b^\top \mu - \beta h(P\mu) \\ s.t. & B^\top \mu \leq c\,. \end{cases}$$

The dual of this dual problem brings back the formulation to primal setting.

Proposition 3.1 (Regularized primal problem). *Consider the linear program*

$$\begin{cases} \min\limits_{y\geq 0} & c^\top y \\ \text{s.t.} & By = b\,. \end{cases}$$

Given a convex function h satisfying $h(\lambda) \geq 0$ and $h(0) = 0$, the primal problem associated with the regularized dual (3.5) *is*

$$\text{(3.6)} \qquad \begin{cases} \min\limits_{y\geq 0, s} & c^\top y + \beta h^*(s) \\ \text{s.t.} & By + \beta P^\top s = b\,. \end{cases}$$

Proof. The Lagrangian $L(\mu, y) = -b^\top \mu + \beta h(P\mu) + y^\top (B^\top \mu - c)$ gives for (3.5) the optimality conditions

$$\text{(3.7)} \qquad \exists \Big(\mu, y, s \in \partial h(P\mu)\Big) \text{ s.t. } \begin{cases} 0 = -b + \beta P^\top s + By \\ 0 = y^\top (B^\top \mu - c)\,, \qquad y \geq 0\,, B^\top \mu \leq c\,. \end{cases}$$

Using the equivalence $s \in \partial h(x) \iff h(x) + h^*(s) = x^\top s$ for the objective function in (3.5), we then obtain that

$$\begin{aligned} -b^\top \mu + \beta h(P\mu) &= -b^\top \mu + \beta s^\top (P\mu) - \beta h^*(s) \\ &= (-b + \beta P^\top s)^\top \mu - \beta h^*(s) \\ &= -(By)^\top \mu - \beta h^*(s) \\ &= -c^\top y - \beta h^*(s)\,, \end{aligned}$$

where the third equality is by the first relation in (3.7), and the last is by the complementarity relation in (3.7). Taking into account the change in sign, this completes the proof that (3.6) is the primal problem that corresponds to (3.5). ∎

Tikhonov's regularization amounts to taking $h(\cdot) := \frac{1}{2}\|\cdot\|_2^2$, a self-dual function ($h^* = h$). The next example shows the interest of considering more general functions. With a suitable function h, the corresponding regularized primal and dual problems keep the linear structure of the original problem. By contrast, with the Tikhonov approach, the regularized problems become quadratic.

Example 3.2 (Regularizing with a norm). Problem (3.5) remains a linear program when the h is a polyhedral norm. In the family of ℓ_r-norms, this is the case with $r = 1$ and $r = +\infty$:

$$\|\lambda\|_1 := \sum_{j=1}^{N} |\lambda_j| \quad \text{and} \quad \|\lambda\|_\infty := \max_{\{j=1,\ldots,N\}} |\lambda_j|\,.$$

In particular, by definition of conjugate function,

$$\text{if} \quad h(\lambda) = \|\lambda\|_\infty \quad \text{then} \quad h^*(s) = \begin{cases} 0 & \text{if } \|s\|_1 \leq 1 \\ +\infty & \text{otherwise,} \end{cases}$$

and the regularized problem is akin to penalizing infeasibility of the constraints, as (3.5) becomes

$$\text{(3.8)} \qquad \begin{cases} \min\limits_{y\geq 0, s} & c^\top y + \beta \|s\|_1 \\ s.t. & By + \beta P^\top s = b\,, \end{cases}$$

Explicit formulations for the regularized primal problems can be derived when the function h is a norm. Specifically, to each norm $h(\cdot) = \|\cdot\|$ we associate a dual norm, defined by

$$h_D(x) = \max_{h(y)=1} x^\top y\,.$$

Using again the conjugate definition,

$$h^*(x) = \begin{cases} 0 & \text{if } h_D(x) \leq 1 \\ +\infty & \text{otherwise,} \end{cases}$$

the regularized primal problem (3.6) can be expressed as

$$\text{(3.9)} \qquad \begin{cases} \min\limits_{y\geq 0, s} & c^\top y \\ s.t. & By + \beta P^\top s = b \\ & h_D(s) \leq 1\,. \end{cases}$$

The dual norm of the ℓ_∞-norm is the ℓ_1-norm. Since, in addition,

$$s \geq 0 \text{ and } \|s\|_1 \leq 1 \iff s \geq 0 \text{ and } \sum_{j=1}^N s_j \leq 1\,,$$

the EPEC problem (3.8) is equivalent to

$$\text{(3.10)} \qquad \begin{cases} \min\limits_{y\geq 0, s} & c^\top y \\ s.t. & By + \beta P^\top s = b \\ & \sum_{j=1}^N s_j \leq 1\,. \end{cases}$$

□

In our numerical experience we observed that, rather than solving the single problem (3.4), it is preferable to solve a sequence of regularized EPECs, decreasing the values of the regularization parameter. In this approach, a run is warm-started taking as initial values for the upper- and lower-level variables the output of the run with the previous β. This technique has the effect of stabilizing the computed solutions, in the sense that the output does not depend on the starting point. Otherwise, different starting points provided to PATH to solve the single problem (3.4), often result in different output. Our next result exhibits that the regularized output in Proposition 3.1 is continuous in β and always converges to a special price, specified below.

Theorem 3.3 (Asymptotic properties of regularized multipliers). *For a convex function $h(\lambda)$ satisfying $h(\lambda) > 0$ for $\lambda > 0$ and $h(0) = 0$, let h^* denote its conjugate. Consider the family of regularized EPECs*

$$\text{for decreasing } \beta > 0, \quad \begin{cases} \min\limits_{x^i} & f^i(x^i, y, \mu) \\ \text{s.t.} & g^i(x^i) \leq 0 \\ & y \text{ solves} \quad \begin{cases} \min\limits_{y,s} & c(x)^\top y + \beta h^*(s) \\ \text{s.t.} & y \geq 0\,, \\ & By + \beta P^\top s = b(x) \quad (\mu)\,. \end{cases} \end{cases}$$

Let $\mu(\beta)$ be a dual solution to the (regularized) lower- level problem above. Then the following statements hold, regarding the convergence to a solution of the lower-level problem in the general EPEC (3.4).

1. *As $\beta \to 0$, the sequences $\{h(P\mu(\beta))\}$ and $\{b^\top \mu(\beta)\}$ are convergent. Furthermore, for any dual solution μ^* of*

$$\left\{\begin{array}{lll} \min\limits_{y} & c(x)^\top y & \\ \text{s.t.} & y \geq 0 & \\ & By - b(x) = 0 & (\mu)\,, \end{array}\right. \tag{3.11}$$

it holds that $\lim_{\beta\to 0} h(P\mu(\beta)) \leq h(P\mu^*)$ and $\lim_{\beta \to 0} b^\top \mu(\beta) = b^\top \mu^*$.

2. *Every accumulation point $\hat{\mu}$ of $\{\mu(\beta)\}$ is a dual solution to (3.11) satisfying*

$$h(P\hat{\mu}) = \min\{h(P\mu^*) : \mu^* \textit{dual solution to } (3.11)\} \tag{3.12}$$

Proof. We drop the dependence on x in what follows. Recall that (3.5) is a dual of the regularized lower- level problem (3.6). Taking $\beta = 0$, (3.5) also gives the dual problem to (3.11).

Let μ^* be any dual solution in (3.11), fixed for now. The fact that μ^* is a dual solution in (3.11) implies that it is feasible in (3.5). Hence,

$$b^\top \mu^* - \beta h(P\mu^*) \leq b^\top \mu(\beta) - \beta h(P\mu(\beta)).$$

On the other hand, by the optimality of μ^* in (3.5) for $\beta = 0$ and the fact that $\mu(\beta)$ is feasible for it (since it solves (3.5) for $\beta > 0$), it holds that

$$b^\top \mu(\beta) \leq b^\top \mu^*\,.$$

Combining the last two relations, we obtain that for all $\beta > 0$,

$$b^\top \mu^* - \beta h(P\mu^*) \leq b^\top \mu(\beta) - \beta h(P\mu(\beta)) \leq b^\top \mu^* - \beta h(P\mu(\beta)). \tag{3.13}$$

Then, from the first and last terms of (3.13), we conclude that

$$0 \leq h(P\mu(\beta)) \leq h(P\mu^*),$$

which shows that the sequence $\{h(P\mu(\beta))\}$ is bounded. Then, taking the limit as $\beta \to 0$ in (3.13),

$$\lim_{\beta\to 0} b^\top \mu(\beta) = b^\top \mu^*.$$

For showing convergence of $\{h(P\mu(\beta))\}$ as $\beta \to 0$, consider any subsequences $\beta_k \to 0$, and $\{h(P\mu(\beta_k)\} \to h(P\mu(\beta))$, passing onto subsequence if necessary (recall that $\{h(P\mu(\beta))\}$ had been proven to be bounded). For any $\beta > 0$, by the optimality of $\mu(\beta)$ in the corresponding problem, and as explained in more detail in related considerations above, we have that

$$b^\top \mu(\beta_k) - \beta h(P\mu(\beta_k)) \leq b^\top \mu(\beta) - \beta h(P\mu(\beta)) \leq b^\top \mu^* - \beta h(P\mu(\beta)).$$

Then, taking the limit in k, we obtain that

$$b^\top \mu^* - \beta \lim_k h(P\mu(\beta_k)) \leq b^\top \mu^* - \beta h(P\mu(\beta)),$$

which implies that

$$h(P\mu(\beta)) \leq \liminf_{\beta\to 0} h(P\mu(\beta)).$$

Thus,

$$\limsup_{\beta\to 0} h(P\mu(\beta)) \leq \liminf_{\beta\to 0} h(P\mu(\beta)) \leq h(P\mu^*),$$

which shows the convergence of $\{h(P\mu(\beta))\}$.

The second item follows directly from the results established above. ∎

The aforementioned "special" price is the one satisfying (3.12). If the regularizing function h is a norm, this confirms the assertions in the introduction: the limit of the regularized prices is a price for the original EPEC that has minimal norm.

Notice that existence of accumulation points of $\{\mu(\beta)\}$ is not discussed in Theorem 3.3. This is clearly a model-dependent issue. For example, the boundedness of $\{h(P\mu(\beta))\}$ and $\{b^\top\mu(\beta)\}$ combined with some additional assumptions, like $b > 0$, might imply the boundedness of $\{\mu(\beta)\}$. Also, whenever B has full rank, the dual feasible set of (3.11) is bounded and thus so is $\{\mu(\beta)\}$. Such is the case for the energy problem, considered in detail in the next section.

Finally, it is important to keep in mind that existence of lower-level regularized multipliers and their limit does not grant the existence of equilibria for the EPEC (3.4). Determining if the regularization is beneficial for existence properties remains an open question, beyond the scope of this work. For general EPECs, this is a very challenging problem. Barring the clean result in [21], existence statements for hierarchical problems are rare in the literature. For more information, we refer the reader to the analysis initiated in [27] and extended in [28]; see also [5].

4. Illustration for the energy market. We next specialize some results to the energy market model in Section 2. In particular, notice that Theorem 3.3(ii) implies that every accumulation point $\hat\mu$ of $\{\mu(\beta)\}$ has the minimal projection $P\hat\mu$ among all dual solutions to (3.11). In Theorem 4.3 below we show that the output of the sequence of regularized EPECs provides the minimal price in the limit.

4.1. Dual regularized ISO problems. Example 3.2 shows the regularized scheme when penalizing with a norm. However, this is not the only choice that might be useful, and keeping a general function h gives better insight into the properties and consequences of dealing with a regularization term. For instance, in Theorem 3.3, item (i), we showed that the sequence $\{h(P\mu(\beta))\}$ is convergent. Under additional assumptions, like h being strictly convex and Assumption 4.1 below, it is possible to show that actually the projected sequence $\{P\mu(\beta)\}$ is convergent as well (Theorem 4.3(i)).

The following properties, which hold for any ℓ_r-norm $h(x) = \|x\|_r$ with $1 \le r \le \infty$, are sufficient in our setting.

Assumption 4.1 (Conditions on the penalizing function).
The convex function $h : \mathbb{R}^N \to \mathbb{R}$ satisfies $h \ge 0$ with $h(0) = 0$ and the conditions:
1. *If $0 \le x \le y$, then $h(x) \le h(y)$.*
2. *If $0 \le x \le y$ and $x_j < y_j$ whenever $y_j > 0$, then $h(x) < h(y)$.*
3. *For any $M \in \mathbb{R}$, the level set $\{x : h(x) \le M\}$ is bounded.* □

In what follows, the positive-part function of a scalar s is defined by

$$[s]^+ := \max(s, 0)\,.$$

Lemma 4.2 (Properties of the regularized dual ISO problem). *Let $(\pi(\beta), \lambda(\beta))$ be any solution to $(3.1)_\beta$, where the function h satisfies Assumption 4.1. Then the following hold.*

(i) The marginal price is non-negative, $\pi(\beta) \ge 0$, and the marginal rent defined as

$$\lambda^\pi_j(\beta) := [\pi(\beta) - p_j]^+ \text{ for } j = 1, \dots, N\,, \tag{4.1}$$

satisfies $\lambda^\pi(\beta) \le \lambda(\beta)$.

(ii) The pair $(\pi(\beta), \lambda^\pi(\beta))$ is also a solution to problem $(3.1)_\beta$.

(iii) Given any two solutions $(\pi^1(\beta), \lambda^{1,\pi}(\beta))$ and $(\pi^2(\beta), \lambda^{2,\pi}(\beta))$ to problem $(3.1)_\beta$,

$$\pi^1(\beta) \le \pi^2(\beta) \iff \lambda^{1,\pi}(\beta) \le \lambda^{2,\pi}(\beta).$$

If, in addition, $\pi^1(\beta) < \pi^2(\beta)$ and $\lambda^{2,\pi}(\beta) \ne 0$, then $h(\lambda^{1,\pi}(\beta)) < h(\lambda^{2,\pi}(\beta))$.

Proof. Given a solution $(\pi(\beta), \lambda(\beta))$ of $(3.1)_\beta$, the fact that $\pi(\beta) \ge 0$ is clear. Also, it is easy to see that the pair $(\pi(\beta), \lambda^\pi(\beta))$ is feasible in $(3.1)_\beta$ and that, as stated in item (i), $\lambda^\pi(\beta) \le \lambda(\beta)$.

To show item (ii), notice that by Assumption (4.1) we have that $h(\lambda^\pi(\beta)) \le h(\lambda(\beta))$; and since $g \ge 0$ we have $g^\top\lambda^\pi(\beta) \le g^\top\lambda(\beta)$. Thus $\pi(\beta)d - g^\top\lambda(\beta) - \beta h(\lambda(\beta)) \le \pi(\beta)d - g^\top\lambda^\pi(\beta) - \beta h(\lambda^\pi(\beta))$, which shows that $(\pi(\beta), \lambda^\pi(\beta))$ is also a solution.

For item (iii), first note that $\pi^1(\beta) \le \pi^2(\beta)$, combined with the facts that $[\cdot]^+$ is monotonically non-decreasing and (4.1), implies that $\lambda^{1,\pi}(\beta) \le \lambda^{2,\pi}(\beta)$. The converse statement assumes that $\lambda^{1,\pi}(\beta) \le \lambda^{2,\pi}(\beta)$. Now, let us consider the following two cases. Suppose first that $\lambda^{2,\pi}(\beta) = 0$. Then $\lambda^{1,\pi}(\beta) = 0$ and, by $(3.1)_\beta$, this forces $\pi^1(\beta) = \min_j\{p_j\} = \pi^2(\beta)$.
In the second case, when $\lambda^{2,\pi}(\beta) \ne 0$, from (4.1), there exists some index j such that

$$\pi^2(\beta) - p_j = [\pi^2(\beta) - p_j]^+ = \lambda_j^{2,\pi}(\beta) \ge \lambda_j^{1,\pi}(\beta) = [\pi^1(\beta) - p_j]^+ \ge \pi^1(\beta) - p_j\,,$$

from which the desired relation follows. Finally, assuming $\pi^1(\beta) < \pi^2(\beta)$, we have that $\lambda^{1,\pi}(\beta) \le \lambda^{2,\pi}(\beta) \ne 0$, and for all j,

$$\pi^2(\beta) - p_j > \pi^1(\beta) - p_j\,.$$

As a result, for all the components j for which $\lambda_j^{2,\pi}(\beta) > 0$, we have that

$$\lambda_j^{2,\pi}(\beta) = \pi^2(\beta) - p_j > \max\{\pi^1(\beta) - p_j, 0\} = \lambda_j^{1,\pi}(\beta)\,,$$

and, by Assumption (4.1), we have that $h(\lambda^{2,\pi}(\beta)) > h(\lambda^{1,\pi}(\beta))$, which concludes the proof. $\square$

Lemma 4.2 is useful when considering convergence of a sequence of approximations, as the regularization parameter tends to zero. In (iii) and (iv) below we show that the approach converges to the price with minimal norm. The statement in item (iv), in particular, states that the limit price will always be the price bid by the marginal agent.

Theorem 4.3 (Asymptotic behavior of regularized dual ISO problems). *Consider any sequence of solutions to $(3.1)_\beta$ $\{(\pi(\beta), \lambda^\pi(\beta))\}$, parameterized by β. Under the assumptions in Lemma 4.2, the following hold.*

(i) As $\beta \to 0$, the sequence $\{(\pi(\beta), \lambda^\pi(\beta))\}$ converges to a point $(\bar\pi, \lambda^{\bar\pi})$.

(ii) The limit point $(\bar\pi, \lambda^{\bar\pi})$ solves problem $(3.1)_0$, that is, the original dual ISO problem (2.7).

(iii) For any other solution to (2.7), say (π^0, λ^0), it holds that

$$\bar\pi \le \pi^0 \quad \text{and} \quad \lambda^{\bar\pi} \le \lambda^0.$$

(iv) The limit price $\bar\pi$ coincides with the marginal price $p_{j_{\rm mg}}$ in Proposition 2.2.

Proof. For item (i), we start by proving that the sequence is bounded, and then show that all its accumulation points are the same.

Applying Theorem 3.3 to (2.6), with $\mu(\beta) = (\pi(\beta), \lambda^\pi(\beta))$ and $P\mu(\beta) = \lambda^\pi(\beta)$, we have that $\{h(P\mu(\beta))\} = \{h(\lambda^\pi(\beta))\}$ is bounded. By Assumption 4.1, it follows that $\{\lambda^\pi(\beta)\}$ is also bounded. Then, combining (4.1), the constraints in $(3.1)_\beta$, and the fact that $\pi(\beta) \ge 0$, we have that the sequence $\{\pi(\beta)\}$, is bounded as well.

Consider any accumulation point $(\pi^{\rm acc}, \lambda^{\rm acc})$ of $\{(\pi(\beta), \lambda^\pi(\beta))\}$. From Theorem 3.3 we have that $(\pi^{\rm acc}, \lambda^{\rm acc})$ is also a solution to $(3.1)_0$ and $h(\lambda^{\rm acc}) \le h(\lambda^{\pi^0})$.

Suppose, for contradiction purposes, that $\pi^{\rm acc} > \pi^0$. Lemma 4.2 ensures that $\lambda^{\rm acc} \ge \lambda^{\pi^0} \ge 0$, which implies $h(\lambda^{\pi^0}) \le h(\lambda^{\rm acc})$. Therefore, $h(\lambda^{\pi^0}) = h(\lambda^{\rm acc})$.

If $\lambda^{\rm acc} = 0$ then $\lambda^{\rm acc} = \lambda^{\pi^0} = 0$. By Lemma 4.2, we have that $\pi^{\rm acc} = \pi^0$, which contradicts our assumption. On the other hand, assuming $\lambda^{\rm acc} \ne 0$, implies, using Lemma 4.2(ii), that $h(\lambda^{\pi^0}) < h(\lambda^{\rm acc})$, which also contradicts the assumption $\pi^{\rm acc} > \pi^0$. Thus, $\pi^{\rm acc} \le \pi^0$, which also yields $\lambda^{\rm acc} \le \lambda^{\pi^0} \le \lambda^0$.

Considering any other accumulation point $(\hat\pi^{\rm acc}, \hat\lambda^{\rm acc})$ of $\Big\{(\pi(\beta), \lambda^\pi(\beta))\Big\}$, since $(\hat\pi^{\rm acc}, \hat\lambda^{\rm acc})$ solves (2.7), we have

$$\pi^{\rm acc} \le \hat\pi^{\rm acc} \quad \text{and} \quad \lambda^{\rm acc} \le \hat\lambda^{\rm acc}.$$

By a similar argument we can show that

$$\hat\pi^{\rm acc} \le \pi^{\rm acc} \quad \text{and} \quad \hat\lambda^{\rm acc} \le \lambda^{\rm acc}\,.$$

Hence, $(\pi^{\rm acc}, \lambda^{\rm acc}) = (\hat\pi^{\rm acc}, \hat\lambda^{\rm acc})$. We have therefore established that all accumulation points of the bounded sequence $\Big\{(\pi(\beta), \lambda^\pi(\beta))\Big\}$ coincide, i.e., the sequence converges. And since $\lambda^{\rm acc} = \lambda^{\pi^{\rm acc}}$, we have that the limit point can be written as $(\bar\pi, \lambda^{\bar\pi})$. This concludes item (i). Then, items (ii) and (iii) follow in a straightforward way.

To see the final item (iv), recall from Proposition 2.2 that $p_{j_{\texttt{mg}}} \le \bar\pi$ and since $(p_{j_{\texttt{mg}}}, \lambda_{p_{j_{\texttt{mg}}}})$ solves (2.7), from item (iii), this means that $\bar\pi \le p_{j_{\texttt{mg}}}$. This concludes the proof. ∎

Recall from Proposition 2.2(iii) that when the marginal agent is dispatched up to the bid, that is when $l^*_{j_{\texttt{mg}}} = g_{j_{\texttt{mg}}}$, the equilibrium price lies in the interval $\Pi^* = \left[p_{j_{\texttt{mg}}}, \min_{k>\texttt{mg}, g_{j_k}>0} p_{j_k}\right]$. A remarkable feature of Theorem 4.3(iv) is that the property $\bar\pi = p_{j_{\texttt{mg}}}$ holds *independently* of the dispatch. This ensures that, as announced, in the limit the multipliers$\pi(\beta)$ provide the smallest possible value for the price, among all the (infinite) choices in the multiplier set Π^*.

4.2. Back to the primal ISO problem. As mentioned, when considering successive EPECs with diminishing regularization parameters, in the numerical experiments we take a small value for β and use the corresponding ISO problem $(3.1)_\beta$ as an approximation for (2.7). In order to guide the choice of the bound for the regularization parameter, and determine its impact (or interference) in the bidding process, it is useful to consider the following regularized primal ISO problem, that results from (3.6), using the identifications given for the energy EPEC after (3.4):

$$(4.2)_\beta \qquad \begin{cases} \min\limits_{l,w} & l^\top p + \beta h^*\Big(\dfrac{l+w-g}{\beta}\Big) \\ \text{s.t.} & \mathbf{1}^\top l = d \\ & l\,, w \ge 0\,. \end{cases}$$

In the original ISO problem, the marginal rent is the multiplier associated to the capacity constraint $l \le g$, that is no longer explicit in $(4.2)_\beta$. The meaning of the variable λ in the regularized problem depends on the penalizing function. The choice $h(\cdot) = \frac{1}{2}\|\cdot\|^2$, which by definition of the conjugate implies that $h^* = h$, gives in $(4.2)_\beta$ the objective function

$$l^\top p + \beta h^*\Big(\frac{l+w-g}{\beta}\Big) = l^\top p + \frac{1}{2\beta}\|l+w-g\|^2\,.$$

This Tikhonov regularization scheme amounts to a quadratic penalization of the capacity constraint.

If the penalizing function is the ℓ_∞-norm, then (3.10) yields

$$(4.3) \qquad \begin{cases} \min\limits_{l\,,s} & l^\top p \\ \text{s.t.} & \mathbf{1}^\top l = d \\ & l \le g + \beta s \\ & \mathbf{1}^\top s \le 1 \\ & l, s \ge 0\,. \end{cases}$$

Since now the capacity constraint appears explicitly, the corresponding optimal multiplier λ_β plays the role of a genuine marginal rent. This primal format reveals the regularized primal ISO as disposing of a *generation reserve* equal to βs. Indeed, the capacity constraint $l_j - \beta s_j \le g_j$ results in values $l_j > g_j$ that can be optimal, seemingly allowing the ISO dispatch the jth agent *beyond* the bid. Of course, this is not possible. Rather, this situation, that leads to positive values for an optimal $s^*_j(\beta)$, is to be understood as the ISO having access to an additional source of energy, out of the market – a battery perhaps. With such reserve, the ISO can complete the dispatch and keep controlled both the price and the rent. Corollary 4.8, given at the end of this section, summarizes all the results for the ℓ_∞-regularization and provides the theoretical background for our numerical assessment. By (3.10), a similar interpretation holds for any penalty in the family of norms, as in Example 3.2.

Solutions to the original problem (2.5), given in Proposition 2.2, can be related to solutions to the regularized primal ISO problem $(4.2)_\beta$ if the penalty verifies the following relation.

Assumption 4.4 (Additional condition on the penalty). For any $x \in \mathbb{R}^N$, it holds that

$$h(x) = h(\mathrm{abs}(x)) \text{ where } \mathrm{abs}(x) := (|x_1|, |x_2|, \ldots, |x_N|)\,. \qquad \square$$

Once again, Assumption 4.4 is satisfied by any ℓ_r-norm with $1 \le r \le \infty$. We state several technical properties related with the new assumption.

Lemma 4.5 (Consequences of Assumption 4.4). *The following holds for a function h satisfying Assumption 4.4.*

(i) The conjugate h^* satisfies Assumption 4.4.

(ii) For any $\lambda \in \mathbb{R}^N$ and a subgradient $s \in \partial h(\lambda)$,

$$\lambda \ge 0 \textit{ with } \lambda \ne 0 \Longrightarrow s_j \ge 0 \textit{ for any component } j \textit{ for which } \lambda_j > 0.$$

(iii) Suppose, in addition, that h is a norm whose dual norm h_D has the property that, for any $0 \le x \le y$,

$$\exists j \in \{1, \ldots, N\} \textit{ such that } x_j < y_j \implies h_D(x) < h_D(y)\,.$$

Then $s \in \partial h(\lambda)$ in item (ii) is such that $s_j = 0$ whenever $\lambda_j = 0$.

Proof. To show (i), denote $\mathrm{sign}(x) := (\mathrm{sign}(x_1), \mathrm{sign}(x_2), \ldots, \mathrm{sign}(x_N))$. Let $\circ$ represent the Hadamard product between vectors, that is

$$\mathrm{sign}(x) \circ y := (\mathrm{sign}(x_1)y_1, \mathrm{sign}(x_2)y_2, \ldots, \mathrm{sign}(x_N)y_N)\,.$$

Then the identity below holds for any $x, y \in \mathbb{R}^N$, showing the assertion:

$$\begin{aligned} x^\top y - h(y) &= \mathrm{abs}(x)^\top(\mathrm{sign}(x) \circ y) - h(y) \\ &= \mathrm{abs}(x)^\top(\mathrm{sign}(x) \circ y) - h(\mathrm{sign}(x) \circ y)\,. \end{aligned}$$

By definition of conjugate function, the subgradient $s \in \partial h(\lambda)$ solves the problem

$$h^{**}(\lambda) = \sup_z \{\lambda^\top z - h^*(z)\}\,.$$

Since $h = h^{**}$ by convexity of h, this means that

$$h(\lambda) = \lambda^\top s - h^*(s) \ge \lambda^\top z - h^*(z) \quad \text{for all } z \in \mathbb{R}^N\,. \tag{4.4}$$

To show item (ii), we proceed by contradiction and suppose there exists j such that $\lambda_j > 0$ and $s_j < 0$. Then we have that $\lambda^\top \mathrm{abs}(s) > \lambda^\top s$ and since $h^*(s) = h^*(\mathrm{abs}(s))$ by item (i),

$$\lambda^\top \mathrm{abs}(s) - h^*(\mathrm{abs}(s)) > \lambda^\top s - h^*(s),$$

which contradicts (4.4). Thus $s_j \ge 0$ whenever $\lambda_j > 0$, as stated.

Continuing with item (iii), when h is a norm, as explained in Example 3.2, having $\lambda \ne 0$ implies that $h_D(s) = 1$. This gives in (4.4) the following:

$$h(\lambda) = \lambda^\top s \ge \lambda^\top z \quad \text{for all } z \text{ with } h_D(z) = 1\,. \tag{4.5}$$

Suppose again by contradiction that for some component j_0 it holds that $s_{j_0} \ne 0$ and $\lambda_{j_0} = 0$. Defining $\hat{s}$ by

$$\hat{s}_j = \begin{cases} 0 & \text{if } j = j_0 \\ |s_j| & \text{otherwise,} \end{cases}$$

we have that $\hat{s} \ne 0$, $0 \le \hat{s} \le \mathrm{abs}(s)$ and $\hat{s}_{j_0} < \mathrm{abs}(s)_{j_0}$. Then

$$0 < h_D(\hat{s}) < h_D(\mathrm{abs}(s)) = h_D(s) = 1.$$

This contradicts (4.5), because $\dfrac{\lambda^\top \hat{s}}{h_D(\hat{s})} > \lambda^\top \hat{s} = \lambda^\top \mathrm{abs}(s) \ge \lambda^\top s$. This completes the proof. $\square$

Thanks to Lemma 4.5, we now characterize the dispatch of the regularized ISO.

Proposition 4.6 (Solution to regularized ISO problem $(4.2)_\beta$). *Suppose h in $(3.1)_\beta$ satisfies Assumptions 4.1 and 4.4. Let $(\pi(\beta), \lambda^\pi(\beta))$ be the marginal pair from Lemma 4.2(i), solving the regularized dual ISO problem $(3.1)_\beta$. The following holds for $(l(\beta), w(\beta))$, a solution to the regularized primal problem $(4.2)_\beta$.*

(i) $s = \dfrac{l(\beta) + w(\beta) - g}{\beta} \in \partial h(\lambda^\pi(\beta))\,.$

(ii) For $j = 1, 2, \ldots, N$,

$$\begin{array}{lllll} \text{if } p_j < \pi(\beta), & \text{then} & s_j \geq 0\,, & l_j(\beta) = g_j + \beta s_j\,, & w_j = 0 \\ \text{if } p_j > \pi(\beta), & \text{then} & & l_j(\beta) = 0 & w_j = g_j + \beta s_j\,. \end{array}$$

(iii) If, in addition h is a norm and $\lambda^\pi(\beta) \neq 0$, then s has length 1 in the dual norm. Furthermore, if the dual norm h_D satisfies the conditions in Lemma 4.5(iii), then the statement in (ii) also includes that $s_j = 0$ whenever $p_j \geq \pi(\beta)$.

Proof. The first item is straightforward from the optimality conditions of $(4.2)_\beta$ and

$$l_\pi(\beta) = g - w^\pi(\beta) + \beta s\,.$$

Consider $j = 1, 2, \ldots, N$. If $p_j < \pi(\beta)$, then $\lambda_j^\pi(\beta) = \pi(\beta) - p_j > 0$. The complementarity condition between λ and w implies that $w_j^\pi = 0$ and, hence, $l_j(\beta) = g_j + \beta s_j$.

For the case $p_j > \pi(\beta)$, we have that $\pi(\beta) - p_j < 0 \leq \lambda_j^\pi(\beta)$. Again, the complementarity condition between this constraint and the Lagrange multiplier l implies that $l_j(\beta) = 0$ and $w_j(\beta) = g_j + \beta s_j$, showing item (ii).

Finally, item (iii) follows from Lemma 4.5, that yields $h_D(s) = 1$. ∎

It is worth noting that the property required for the norm in Lemma 4.5(iii) is satisfied by the ℓ_∞-norm, but *not by the ℓ_1-norm*. This is the reason why in our numerical results the regularized EPECs are defined using the former option.

Like in Theorem 4.3, we consider convergence of a sequence of approximations as the parameter β tends to zero, now from the primal point of view.

Theorem 4.7 (Behavior of regularized primal ISO problems). *Given the marginal price $p_{j_{\text{mg}}}$ from Proposition 2.2, consider the index-sets*

$$J^- = \left\{j : p_j < p_{j_{\text{mg}}}\right\} \quad \text{and} \quad J^+ = \left\{j : p_j > p_{j_{\text{mg}}}\right\}.$$

Let $\{(l(\beta), w(\beta), s(\beta))\}$ be any sequence parameterized by $\beta > 0$, where $(l(\beta), w(\beta))$ solves $(4.2)_\beta$ and $s(\beta) = \frac{l(\beta)+w(\beta)-g}{\beta}$. Under the assumptions in Proposition 4.6, the following holds.

(i) The sequence $\{(l(\beta), w(\beta), s(\beta))\}$ is bounded.

(ii) Any accumulation point l^{acc} of $\{l(\beta)\}$ solves the primal ISO problem (2.5).

(iii) There exists $M > 0$ such that for $\beta > 0$ sufficiently small,

$$\begin{array}{lllll} j \in J^- & \Longrightarrow & l_j(\beta) = g_j + \beta s_j(\beta)\,, & w_j(\beta) = 0\,, & |l_j(\beta) - g_j| \leq M\beta \\ j \in J^+ & \Longrightarrow & l_j(\beta) = 0\,, & w_j(\beta) = g_j + \beta s_j(\beta)\,, & |w_j(\beta) - g_j| \leq M\beta\,. \end{array}$$

(iv) If, in addition h is a norm satisfying the conditions in Lemma 4.5(iii), then the statement above can be refined by taking $M = 1$, which implies that $w_j(\beta) = g_j$ for $j \in J^+$.

Furthermore, the sequences $\{l_j(\beta)\}$ and $\{w_j(\beta)\}$ converge for any $j \in J^- \cup J^+$.

Proof. Consider a sequence of dual solutions $\Big\{(\pi(\beta), \lambda^\pi(\beta))\Big\}$, shown to be convergent in Proposition 4.6. Then $p_{j_{\text{mg}}} = \lim_{\beta\to 0} \pi(\beta)$, with $s(\beta) \in \partial h(\lambda^\pi(\beta))$. Since the sequence $\{\lambda^\pi(\beta)\}$ is bounded and h is convex, we have that the family of subdifferentials $\{\partial h(\lambda^\pi(\beta))\}$ is uniformly bounded, and so, the sequence $\{s(\beta)\}$ is bounded: there exists $M > 0$ such that

$$\|s(\beta)\| \leq M.$$

On the other hand, it is clear that $l(\beta)$ is feasible for problem (2.5) and since the feasible set of this problem is bounded, we have that the sequence $\{l(\beta)\}$ is bounded. Then, from $w(\beta) = g - l(\beta) + \beta s(\beta)$, we have that the sequence $\{w(\beta)\}$ is also bounded. Item (i) is established.

In order to prove item (ii), note that $h^*(s(\beta)) = s(\beta)^\top \lambda^\pi(\beta) - h(\lambda^\pi(\beta))$ implies that the sequence $\{h^*(s(\beta))\}$ is bounded, and hence, $\beta h^*(s(\beta)) \to 0$, as $\beta \to 0$. Now, considering the strong duality condition for $(4.2)_\beta$ – $(3.1)_\beta$

$$\pi(\beta)d - \lambda^\pi(\beta)^\top g - \beta h(\lambda^\pi(\beta)) = p^\top l(\beta) + \beta h^*(s(\beta))$$

and passing to the limit as $\beta \to 0$, taking a convergent subsequence if necessary, we have that

$$p_{j_{\text{mg}}} d - \lambda^\top_{p_{j_{\text{mg}}}} g = p^\top l^{\text{acc}}.$$

This shows that the strong duality condition also holds for (2.5) – (2.7), and since l^{acc} is feasible for (2.5), item (ii) follows.

Finally, note that, letting $p^- = \max_{j \in J^-} p_j$ and $p^+ = \min_{j \in J^+} p_j$, we have that

$$p^- < p_{j_{\text{mg}}} < p^+ .$$

Therefore, for $\beta > 0$ small enough,

$$p^- < \pi(\beta) < p^+ .$$

The final statement is straightforward from Proposition 4.6, Lemma 4.5, and the fact that $0 \leq s_j(\beta) \leq h_D(s_j(\beta) = 1$, for $j \in J^-$. ∎

Most of the items in the theorem above are of asymptotic nature. A remarkable exception is item (iii), that characterizes the optimal dispatch for all small β. The characterization does not involve the marginal agents because, similarly to the situation pointed out in Remark 2.3 for the original primal problem ($\beta = 0$), there is an ambiguity created by the ISO's indifference that arises when more than one agent bids the same marginal price.

We conclude our theoretical analysis for the energy EPEC by gathering the results specific for the ℓ_∞-norm, which is polyhedral and satisfies not only Assumptions 4.1 and 4.4, but also the condition given in Lemma 4.5(iii). The results stated below are useful in the numerical section, to assess when the output computed by PATH is a true equilibrium.

Corollary 4.8 (Summary of theory for the energy EPEC with ℓ_∞-regularization). *With the notation and assumptions in Proposition 2.2, Theorems 4.3 and 4.7, consider the regularized dual ISO problem obtained with $h = \|\cdot\|_\infty$. The sequence $\{(\pi(\beta), \lambda^\pi(\beta))\}$ of solutions to the dual version $(3.1)_\beta$ satisfies the following:*

$$\left.\begin{array}{l} \lim_{\beta\to 0} \pi(\beta) = p_{j_{\text{mg}}} \\ \lim_{\beta\to 0} \|\lambda^\pi(\beta)\|_\infty = \max(p_{j_{\text{mg}}} - p_{j_1}, 0) \end{array}\right\} \text{ provide minimal-norm solutions to (2.7).}$$

In addition, for any $\beta > 0$ sufficiently small, the pair $(l(\beta), w(\beta))$ solving the primal version (4.3) is such that

$$l_j(\beta) = \begin{cases} g_j & \text{if } p_j > p_{\min}, j \in J^- \\ 0 & \text{if } j \in J^+, \end{cases}$$

where $p_{\min} = \min_j p_j$. Finally, the marginal dispatch completes the demand, following an arbitrary distribution among the marginal agents, if there is more than one bidding $p_{j_{\text{mg}}}$, as noted in Remark 2.3. □

Proof. The statements follow from our previous results. The only exception concerns the value of $l_j(\beta)$ for indices j such that $p_j > p_{\min}$ and $j \in J^-$. For such j-indices, Theorem 4.7(iii) states that $l_j(\beta) = g_j + \beta s_j$. Then, from (4.3), optimality arguments imply that $s_j = 0$ for any j such that $p_j > p_{\min}$. ∎

Remark 4.9. In the proofs above, it is easy to see that $\sum s_j = 1$. Then, in case that there exists only one index j such that $p_j > p_{\min}$, this would force $s_j = 1$, which in turn implies that $l_j(\beta) = g_j + \beta$. When the index is not unique, there is no unique solution to (4.3), yet we can always choose one j that assigns a dispatch $g_j + \beta$ to one of the least expensive units. $\square$

5. Putting the method in perspective. Our formulation of problem (3.11), allows the ISO to include network constraints when defining the dispatch. Such would be the case in the US, where day-ahead markets rely on the ISO solving very detailed security-constrained models. It is explained in [13] that such feature originated in a pre-existent integrated structure of Regional Transmission Operators. Europe focused instead on a single market implementation supported by power exchanges (for such a market, (3.11) may not include transmission constraints). Clearly, the level of detail introduced in the lower-level problem depends on the market configuration. The abstract setting (3.11) encompasses various market formats, including future ones, with market bids and clearing systems adapted to new agents dealing with storage and renewable intermittent sources of energy. We refer to [13] for a thorough discussion on the topic.

We extend our approach to a more complex configuration, with agents placing bids in several markets, handled by separate ISOs, but connected by some transmission line. Many numerical tests to assess the interest of the proposal in the single and multi-market setting are given afterwards.

5.1. Bidding to more than one market. Suppose that, instead of (3.11), in the lower level there are different ISO optimization problems and a constraint coupling all the markets through a transmission line. The resulting EPEC, which has in the lower level a variational inequality and no longer one optimization problem, can still be tackled following our approach developed above.

Suppose there are K markets, handled by separate ISOs. Then, for $k \in K$, ISO_k determines the market price μ_k, to be paid for the k-th market demand d_k, as well as the dispatch, y_k. If the transmission line has capacity κ, then

$$\sum_{k=1}^{K} T_k y_k \leq \kappa \quad (\eta)$$

is the constraint coupling the markets. The value of the dual variable η represents the unit fee to be paid for transporting energy through the transmission line.

In this setting, instead of (3.4), we have the following multi-market EPEC

$$(5.1) \qquad \begin{cases} \min\limits_{x^i} & f^i(x^i, y, \mu, \eta) \\ \text{s.t.} & G^i(x^i) \leq 0 \\ & \text{for } k = 1, \dots, K \quad y_k \text{ solves} \quad \begin{cases} \min\limits_{y_k} & c_k(x)^\top y_k \\ \text{s.t.} & y_k \geq 0 \\ & B_k y_k = b_k(x) \quad (\mu_k) \end{cases} \\ & \text{and } \sum\limits_{k=1}^{K} T_k y_k \leq \kappa \quad (\eta)\,. \end{cases}$$

The corresponding family of regularized EPECs is obtained by replacing the K ISO problems by their regularized formulations, as in (3.6).

The upper-level variables and functions in (5.1) are defined to represent the multi-market setting. As an illustration, consider the situation depicted by the diagram in Figure 2. There are two markets, with respective demands d_1 and d_2, so $K = 2$. The N generating companies are distributed into two sets, I_1 and I_2, gathering agents in each market. Bids can be placed in both markets, noting that if, for instance, agent $i \in I_1$ bids to market 2, the exported energy goes through the transmission line, and incurs an additional expense, depending on η. The generation g_i is distributed into two parts, a fraction $\theta_i \in [0, 1]$ of g_i will be sold locally, in the market where the agent is located. The remaining generation $(1 - \theta_i)g_i$ is offered to the other market, to be exported through the transmission line, of capacity κ.

Agent i's decision variables are $x^i := (g_i, \theta_i, p_i, q_i)$, given that quantities and prices are:

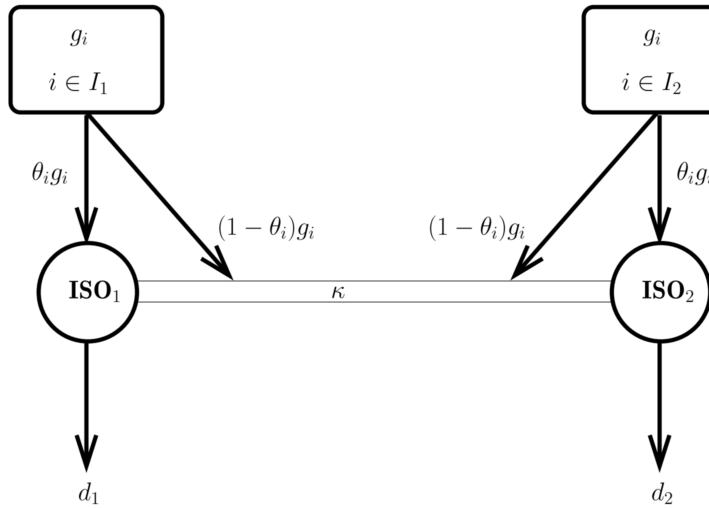

Fig. 2: Two markets with one transmission line. For $\theta_i \in [0,1]$, agent i in market 1 bids locally the generation $\theta_i g_i$ at price p_i, and offers to export to market 2 the fraction $(1-\theta_i)g_i$ at price q_i. Exchanges between markets are limited by the capacity κ of the transmission line.

- in its local market, bids $\theta_i g_i$ at price p_i
- to the other market, offers to export $(1-\theta_i)g_i$ at price q_i through the transmission line.

Accordingly, when for example $i \in I_1$, the objective function in the upper level is

$$f^i(x^i, y, \mu) = (\pi_1 - \varphi_1)l_1 + (\pi_2 - \varphi_1 - \eta)l_2 \,.$$

With respect to the formulation with only one market, the only difference is that in the lower level now there is a variational inequality, representing the separate operation of the two markets and including the capacity constraint of the transmission line,

$$\sum_{k \in \{1,2\}} T_k y_k \le \kappa \,.$$

In the multi-market problem, the lower level has decision variables $\Big((y_k, \mu_k)_{k=1}^2, \eta\Big)$, where the multiplier η of the capacity constraint defines the charge that agents have to pay for each unit of energy transmitted through the line.

5.2. Numerical assessment. In the single-market EPEC model (3.4), the leaders disutilities approximate the price function $P(g,p,l)$ with the demand multiplier. In the lower level, the regularized ISO solves problem (4.3), corresponding to taking the ℓ_∞-norm in the dual.

There are a number of ways to tackle the non-convexity that appears in EPECs. With our proposal, the EPEC can be handled computationally, for example by the PATH solver [6]. A direct solution of (3.4) leads to a computationally challenging problem with severely non-convex/disjunctive complementarity constraints, [25], [18, Chapter 7.3]. It is important to keep in mind that, while being a stationary point for the complementarity system residual, the solution provided by PATH may not be an equilibrium (having a bilinear upper objective and bilevel formulation makes the resulting mixed complementarity problem non-monotone). One of our conclusions is that the output is very sensitive to the initial input. That said, experimentation and the resulting appropriate tuning of the regularization parameter, implementing the mechanism of warm starts for decreasing values of β, leads to useful, optimal/equilibrium solutions. The results below are meant to illustrate these conclusions for the given problem.

5.2.1. Benchmark information and values at equilibrium. The family of regularized EPECs considers decreasing values of the parameter,

$$\beta = (0.2 - 0.02j)d\,, \text{for } j = 0, 1, \ldots, 10\,, \quad \text{and a given demand } d\,, \tag{5.2}$$

which amounts to the ISO having access to a "battery" that covers 20%, 18%,. . . , 2%, and 0% of the market's demand. The results provided by PATH for $\beta = 0$ (after the sequence of regularized EPECs and warm starts with $\beta > 0$) are then compared to the output obtained by PATH for the original EPEC (3.4), without regularization. Both approaches (without and with regularization) use the same starting point, randomly taken in the generators' bid feasible sets.

Thanks to the results shown for the stylized model in Proposition 2.2 and Corollary 4.8, we can make a thorough assessment of the output and gauge its quality. In particular we show empirically, over thousands of starting points in several experiments, that the output of PATH for the regularized EPECs succeeds in finding genuine equilibria much more often than the direct approach.

Given as input the number of players N, the marginal cost as well as the maximum bidding price and generation increase with the index of the player:

$$\text{for } j = 1, \ldots, N \text{ we set }\,, \quad \varphi_j = \frac{j}{2}\,, \quad p_j^{\max} = 2\varphi_j = j\,, \quad g_j^{\max} = j\,.$$

The index of the marginal agent $j_{\mathtt{mg}} \leq N$ is also an input, and we consider the following two values for the demand:

$$D := \sum_{j=1}^{j_{\mathtt{mg}-1}} g_j^{\max} + 0.5 g_{j_{\mathtt{mg}}}^{\max} \quad \text{and} \quad \tilde{D} := \sum_{j=1}^{j_{\mathtt{mg}-1}} g_j^{\max} + 1.0 g_{j_{\mathtt{mg}}}^{\max}\,.$$

To understand the consequences of this setting, recall from Proposition 2.2 that generators bidding a price cheaper than the marginal one are dispatched by the ISO at their maximal capacity. The value chosen for the demand D absorbs all the generation capacity of inframarginal agents, but not that of the marginal agent. With $\tilde{D}$, by contrast, the marginal agent is dispatched up to its bid. Since by Proposition 2.2, in the latter case the price is not unique, we expect the runs with demand $\tilde{D}$ to be more challenging.

At least for sufficiently small β, by Corollary 4.8 and Proposition 2.2. the ISO values at equilibrium are

$$\begin{array}{lll} \text{dispatch} & l_{j_k}(\beta) = \begin{cases} g_{j_k}^{\max} & \text{for } k = 1, \ldots, \mathtt{mg} - 1 \\ 0.5(\text{or } 1) g_{j_{\mathtt{mg}}}^{\max} & \text{for } k = \mathtt{mg} \text{ (if the demand is } D \text{ or } \tilde{D}) \\ 0 & \text{for } k = \mathtt{mg} + 1, \ldots, N\,, \text{and} \end{cases} \\ \text{price} & \pi(\beta) \in [j_{\mathtt{mg}}/2, j_{\mathtt{mg}+1}/2]\,, \text{ the interval of bidding prices of agent } j_{\mathtt{mg}}. \end{array} \tag{5.3}$$

The result is also valid for the original EPEC (3.4). Accordingly, both without and with regularization, runs providing a nonzero dispatch for an agent with index larger than the marginal one, or with a price too large cannot correspond to an equilibrium.

We coded the model in GAMS and PATH to directly solve the EPEC (3.4). Thanks to the Extended Mathematical Programming (EMP) extension available in GAMS, instead of manually writing down the complementarity system, replicating variables to make the system square, it suffices to write the model at a high level, indicating which variables are owned by the agents and which ones by the ISO.

The EMP framework vastly facilitates a direct formulation of multiple optimization problems with equilibrium constraints such as (3.4). For full details on the different features of the extension, we refer to [19]. Here we just give an overview for (3.4). We consider generating companies act as optimization agents in the upper level, with objective function `OBJ(i)` representing f^i, and variable `X(pn,i)` representing x^i, where `pn` refers to name of the component (price p_i or generation g_i in the vector x^i). Similarly, in the lower level, the so-called "equilibrium agent" solves a variational inequality derived from the optimality conditions of (3.11), defined in GAMS as `defVI`, on variables `ZZ(dualv, i)`, representing the dispatch $DIS(i)$ of each agent (l_i), their marginal rent (λ_i) and the market price (π). After defining the corresponding functions and variables in GAMS, the EMP code section is simply

```
equilibrium
implicit  DISPATCH(i),PRICE,defDISPATCH(i),defPRICE
min OBJ(i) s.t. X(pn,i),defOBJ(i)
vi defVI(dualv,i),ZZ(dualv,i)
```

where the `implicit` line contains lower-level variables that appear in the upper level problem.

All the experiments were performed on a notebook running under Ubuntu 18.04.4 LTS, with i7 CPU 1.90GHzx8 cores and 31.3GiB of memory. The parameter settings for PATH is the default provided by GAMS.

5.2.2. Single-market results. We start considering a setting where $N = 3$ agents bid to one market and set the marginal agent to be $j_{\mathtt{mg}} = 2$. Even for this very simple market instance, to obtain a genuine equilibrium when solving a complementarity system with PATH is very delicate. The proposed regularization scheme is beneficial, even though it also fails sometimes.

As the output is very dependent on the initial point, the experiment repeats the runs with 2500 different starting bids $\left(p_j^0, g_j^0\right)$, randomly generated in $[\varphi_j, p_j^{\max}] \times [0, g_j^{\max}]$. Each run takes a starting point and first calls PATH to solve the original EPEC (3.4). Then, from the same starting point, PATH is called eleven times, for decreasing values of β, as in (5.2), with $d \in \{D, \tilde{D}\}$. Barring the first run ($\beta = 0.2d$), the regularization procedure warm-starts each run using as initial point the output of the run with the previous value of β. Altogether, without and with regularization, obtaining the output for the 2500 starting points involves running PATH 30000 times ($30000 = 2500 \times 12$ values of β). In general, PATH runs were very fast, taking less than 15 minutes to complete the full experiment.

As already mentioned, not all solutions obtained by PATH are an equilibrium. As shown in (5.3), the largest value acceptable for the equilibrium price is $j_{\mathtt{mg}+1}/2$. Accordingly, a run is declared a failure if an agent j with $j > \mathtt{mg}$ is dispatched, or if the computed price is larger than $1.05 j_{\mathtt{mg}+1}/2$. A run can also be declared a failure by PATH itself, if a local solution was not found. Over 5000 runs, the decisions over the non-dispatched agents was always correct with both approaches. By contrast, the original (3.4) miscalculated the equilibrium price more than half of the runs, while with regularization the output was incorrect about 25% of the runs. Failures occurred mostly with the demand set to $\tilde{D}$ (80% and 25% of the runs without and with regularization, respectively).

Figure 3 shows the dispatch and generation bids computed in mean for each value of β, including the output without regularization. The graph confirms (5.3): the ISO dispatches agent 1 at maximum capacity, and completes the demand with generation from agent 2.

In Figure 3 notice that, as β decreases, the regularized dispatch of agent 2 progressively increases to the optimal value ($l_2^* = d - 1$). The "battery" provided by the regularization scheme accounts for the difference. The impact of the regularization in the decision making process is perceptible in Figure 4, with the mean composition of the dispatch, for each value of β.

Other than failing less often, there is not a noticeable difference in the dispatch computed with and without regularization. The beneficial effect of the proposed approach is perceptible on the dual variables, where it acts as a stabilizing mechanism. A graphical illustration of this phenomenon is given by Figure 5, with the values of the prices output by PATH with both approaches. The points correspond to 10000 runs, repeating the experiment 4 times, considering the maximum bidding price of the marginal agent is 1.25 or 1.5, and varying the demand in $\{D, \tilde{D}\}$. Recall from (5.3) that the marginal agent, $j_{\mathtt{mg}} = 2$, could bid any price between 1 and its maximum price. The green points in the plots indicate that PATH yielded prices in the allowed range. Such was the case most of the times with the regularization, but not with the direct solution approach.

To illustrate the role of the marginal agent in the price determination, we consider the setting in Figure 5 having $p_2^{\max} = 1.25$ and demand $D = 2$. Since agent 1 is dispatched at value 1, agent 2

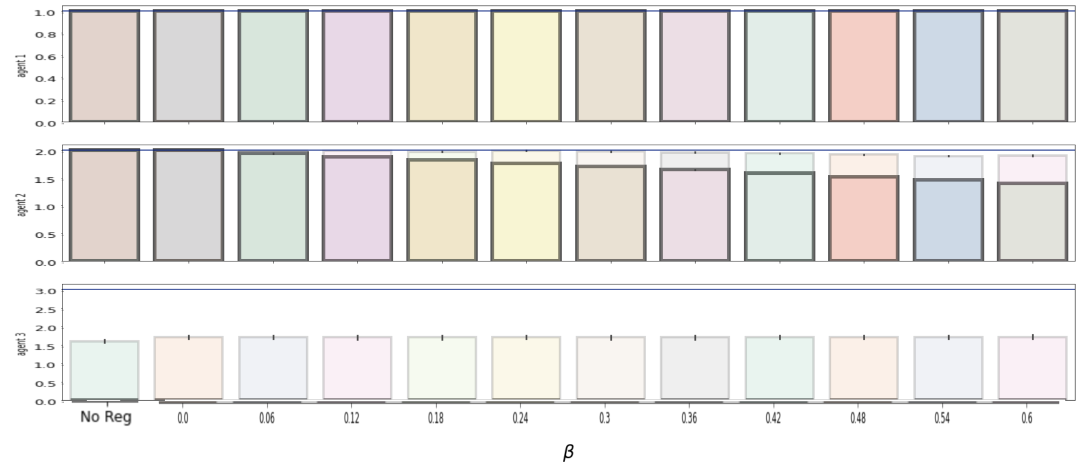

Fig. 3: Mean dispatch per agent, computed by PATH for different values of β when $d = \tilde{D} = 3$. Shaded background areas illustrate the generation bids of the agents and the left-most column corresponds to the runs without regularization. The behavior with $d = D = 2$ was similar.

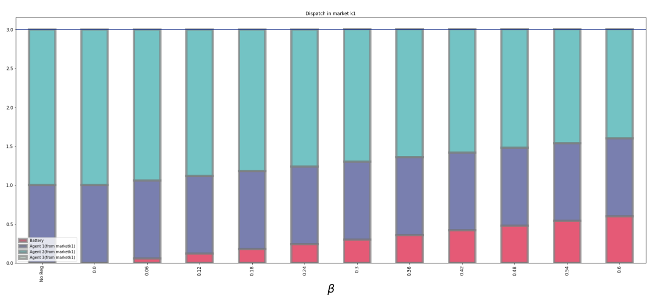

Fig. 4: Mean dispatch composition when $d = \tilde{D} = 3$. Agent 1 in dark blue is dispatched to the maximum and agent 2 in light blue completes the demand (the "battery" is displayed in red). The left-most column corresponds to the runs without regularization, followed for increasing values of β (the behavior with $d = D = 2$ was similar).

covers the residual demand, equal to 1 and the equilibrium price should be in the interval $[1, 1.25]$. We repeated 2500 runs, varying the maximum capacity of the marginal agent

$$g_2^{\max} = 2 - 0.1j \quad \text{for} \quad j \in \{0, 1, \dots, 10\}\,.$$

The statistic for the equilibrium price is reported in Figure 6. There is a sharp increase in the price computed without regularization when the capacity coincides with the residual demand.

We made a final set of experiments for one market, varying the number of agents between 3 and 30, selecting randomly the marginal agent, and with the two demand values, D and $\tilde{D}$. Table 1 informs the corresponding failures, noting that each market configuration was run with both approaches using 2500 starting points.

The conclusion is similar to the case with $N = 3$ agents. More precisely, PATH appears to be very sensitive to the starting point, and the regularization increases the number of runs that can provide an equilibrium (because at the very least the computed price is within the bidding interval of the marginal agent, see (5.3)).

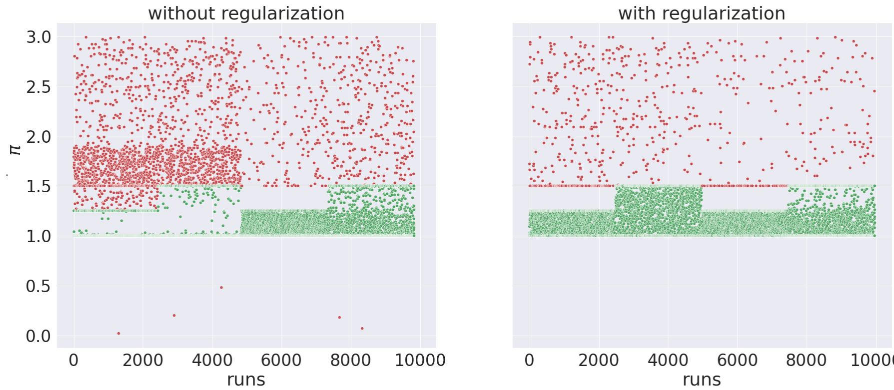

Fig. 5: Prices found with both approaches (only runs with $\pi^* \leq 3$ are displayed, for better visibility). Each graph reports the results of four experiments, running for 2500 different starting points a problem with $(p_2^{\max}, d) \in \{(1.25, \tilde{D}), (1.50, \tilde{D}), (1.25, D), (1.50, D)\}$. Points in red cannot correspond to an equilibrium. We notice much more green points on the right graph, indicating the benefits of the regularization approach.

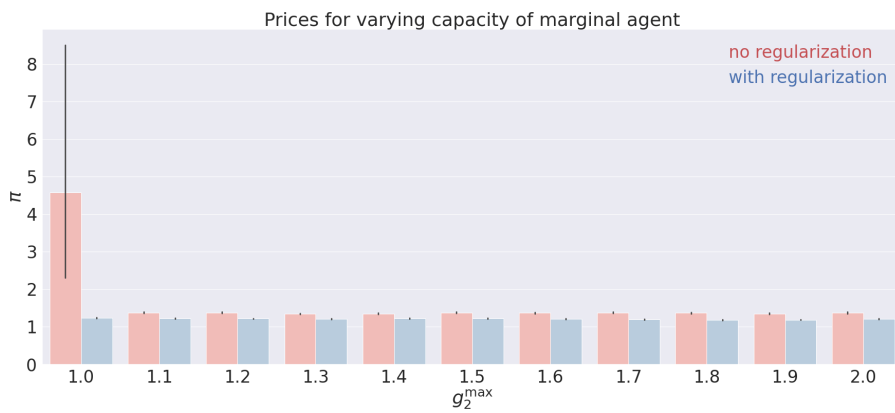

Fig. 6: Prices found with both approaches, as the marginal agent maximum generation drops down to the value of the residual demand. Without regularization, equilibrium prices are always larger than without regularization. Notice the significant increase when $g_2^{\max} = 1$: without regularization, the computed price is on average larger than 4, and can reach values equal to 8, while with regularization the equilibrium price stays at its correct value, 1.25.

5.2.3. Two markets. In order to assess the multi-market EPEC (5.1) we consider the same three agents, but now distributed in two markets, as follows

$$\text{in Figure 2,} \quad \begin{array}{l} I_1 = \{1\} \\ I_2 = \{2, 3\}, \end{array} \quad \text{so agent 1 is in market 1, while agents 2 and 3 are in market 2}\,.$$

Both markets face a demand $d_k = 2$. In market 1, as agent 1 generating capacity is 1, agents 2 and 3 must export to market 1 the missing energy. Also, since the total demand is 4 and agents 1 and 2 can generate 3 in total, agent 3 is always dispatched. We expect the equilibrium price not to exceed the $p_3^{\max} = 3$. The transmission line has a capacity varying in $\kappa = \{1.5, 3, 6\}$. In particular, $\kappa = 6$ is sufficiently large to render free the energy transmission between markets.

We repeated the same procedure, comparing for 2500 starting points the output of a direct solution of (5.1) with the regularization approach, solving the regularized EPEC for decreasing values of β. The total number of PATH solves was 90000 ($= 3 \times 2500 \times 12$), out of which PATH

Table 1: Type of output obtained with PATH when $N \in [3, 30]$ (over 40000 runs)

`PATH fails`		`Too small price`		`Too large price`	
original	regularized	original	regularized	original	regularized
224	175	3	2	15372	10202

failed in 7780 runs, most of them when solving directly (5.1), without regularization (6672 times). The regularization technique proves again beneficial, as illustrated by Figures 7 and 8, with the prices π and η computed in the experiment.

Fig. 7: Prices with both approaches (top and bottom) for markets 1 and 2 (left and right). Each graph reports the results of three experiments, running for 2500 different starting points a problem with $\kappa \in \{1.5, 3, 6\}$. Points in red cannot correspond to an equilibrium. The different length between the top and bottom graphs measures the larger number of failures in the top, without regularization, most of them occur when the line is congested ($\kappa = 1.5$). With regularization, there are much more green points, confirming once more the benefits of the regularization approach.

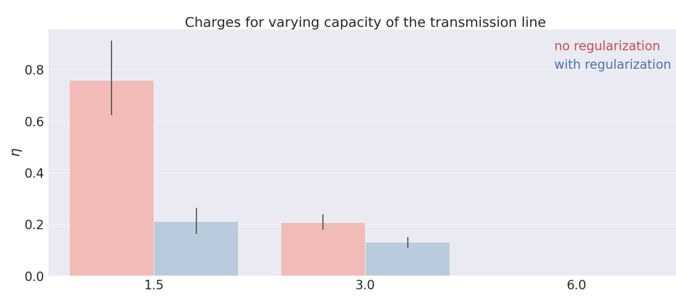

Fig. 8: Statistics for the transmission charge η, when the capacity line is $\kappa \in \{1.5, 3, 6\}$. Without regularization, the fee is larger and exhibits more variability (the length of the vertical black line in each bar), particularly when κ is small. As expected, for the larger value of κ, the transmission constraint is inactive and the value of η is zero with both approaches.

Concluding Remarks. We presented theoretical analysis pointing out some downsides of multi-leader single follower models, including potential overpricing. In situations in which the multiplier is not unique, the proposed regularization drives the computed multipliers to a clearing price with minimal norm.

An interesting topic of future research, raised by a reviewer, refers to a market with agents also bidding on the demand. It is not clear to which extent demand elasticity impacts on the leaders' decisions and price formation.

The regularized problem, which remains a linear programming problem if a polyhedral norm is employed, has interesting economic interpretations derived from analyzing both its dual and primal versions. Namely, the regularization process can be seen as endowing the ISO with a small reserve that allows to control the marginal rent of the dispatched agents. The reserve can be thought of as being available out of the market, or simply being incorporated in the corrections that modify the generation when operating the system in real time.

Our theoretical analysis is complemented with a thorough numerical assessment. The experiments, designed to shed a light on the numerical difficulty inherent to solving EPECs, show that the regularization scheme behaves as a stabilizing device that helps guiding the process towards an output that is usually an equilibrium, even if the mixed complementarity formulation resulting from EPEC is not monotone.

Acknowledgments. The authors are grateful to the Associate Editor and two reviewers, whose thorough reading and detailed comments were very beneficial for this work.

References.

[1] D. Aussel, P. Bendotti, and M. Pištěk. "Nash equilibrium in a pay-as-bid electricity market Part 2 - best response of a producer". In: *Optimization* 66.6 (2017), pp. 1027–1053.

[2] D. Aussel and A. Svensson. "A Short State of the Art on Multi-Leader-Follower Games". In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Cham: Springer International Publishing, 2020, pp. 53–76.

[3] M. Cruz, E. Finardi, V. de Matos, and J. P. Luna. "Strategic bidding for price-maker producers in predominantly hydroelectric systems". In: *Electric Power Systems Research* 140 (2016), pp. 435–444.

[4] S. Cui and U. Shanbhag. *On the computation of equilibria in monotone and potential stochastic hierarchical games*. Tech. rep. https://arxiv.org/pdf/2104.07860.pdf. 2021.

[5] V. DeMiguel and H. Xu. "A Stochastic Multiple-Leader Stackelberg Model: Analysis, Computation, and Application". In: *Operations Research* 57.5 (Oct. 2009), pp. 1220–1235.

[6] S. Dirkse and M. Ferris. "The PATH solver : A nonmonotone stabilization scheme for mixed complementarity problems". In: *Optimizations Methods and Software* 5 (1995), pp. 123–156.

[7] J. F. Escobar and A. Jofré. "Monopolistic competition in electricity networks with resistance losses". In: *Economic Theory* 44.1 (2010), pp. 101–121.

[8] S. A. Gabriel and F. U. Leuthold. "Solving discretely-constrained MPEC problems with applications in electric power markets". In: *Energy Economics* 32.1 (2010), pp. 3–14.

[9] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz. *Complementarity Modeling in Energy Markets*. Springer Publishing Company, Incorporated, 2012.

[10] H. Gérard, V. Leclère, and A. Philpott. "On risk averse competitive equilibrium". In: *Oper. Res. Lett.* 46.1 (2018), pp. 19–26.

[11] H. J. Greenberg and F. H. Murphy. "Computing Market Equilibria with Price Regulations Using Mathematical Programming". In: *Oper. Res.* 33.5 (Oct. 1985), pp. 935–954.

[12] R. Henrion, J. Outrata, and T. Surowiec. "Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market". In: *ESAIM: COCV* 18.2 (2012), pp. 295–317.

[13] I. Herrero, P. Rodilla, and C. Batlle. "Evolving Bidding Formats and Pricing Schemes in USA and Europe Day-Ahead Electricity Markets". In: *Energies* 13.19 (2020).

[14] B. Heymann and A. Jofré. "Mechanism Design and Auctions for Electricity Networks". In: *Forum for Interdisciplinary Mathematics*. Springer Singapore, 2017, pp. 65–81.

[15] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals*. Vol. 305. Springer science & business media, 2013.

[16] B. Hobbs, C. Metzler, and J.-S. Pang. "Strategic gaming analysis for electric power systems: an MPEC approach". In: *Power Systems, IEEE Transactions on* 15.2 (2000), pp. 638–645.

[17] X. Hu and D. Ralph. "Using EPECs to Model Bilevel Games in Restructured Electricity Markets with Locational Prices". In: *Operations Research* 55.5 (2007), pp. 809–827.

[18] A. Izmailov and M. Solodov. *Newton-type methods for optimization and variational problems.* Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2014.

[19] Y. Kim and M. C. Ferris. "Solving equilibrium problems using extended mathematical programming". In: *Math. Program. Comput.* 11.3 (2019), pp. 457–501.

[20] A. A. Kulkarni and U. V. Shanbhag. "A Shared-Constraint Approach to Multi-Leader Multi-Follower Games". In: *Set-Valued and Variational Analysis* 22.4 (2014), pp. 691–720.

[21] A. A. Kulkarni and U. V. Shanbhag. "An Existence Result for Hierarchical Stackelberg v/s Stackelberg Games". In: *IEEE Transactions on Automatic Control* 60.12 (Dec. 2015), pp. 3379–3384.

[22] S. Leyffer and T. Munson. "Solving multi-leader–common-follower games". In: *Optimization Methods and Software* 25.4 (Aug. 2010), pp. 601–623.

[23] J. P. Luna, C. Sagastizábal, and M. Solodov. "An approximation scheme for a class of risk-averse stochastic equilibrium problems". In: *Mathematical Programming* 157.2 (2016), pp. 451–481.

[24] J. P. Luna, C. Sagastizábal, and M. Solodov. "Complementarity and game-theoretical models for equilibria in energy markets: deterministic and risk-averse formulations". In: *International Series in Operations Research and Management Science*. Ed. by R. Kovacevic, G. Pflug, and M. Vespucci. Vol. 199. Springer, 2013, pp. 231–258.

[25] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints.* Cambridge University Press, 1996.

[26] J.-S. Pang, S. Sen, and U. V. Shanbhag. "Two-stage non-cooperative games with risk-averse players". In: *Mathematical Programming* 165.1 (2017), pp. 235–290.

[27] H. D. Sherali. "A Multiple Leader Stackelberg Model and Analysis". In: *Operations Research* 32.2 (Apr. 1984), pp. 390–404.

[28] C.-L. Su. "Analysis on the forward market equilibrium model". In: *Operations Research Letters* 35.1 (Jan. 2007), pp. 74–82.

[29] J. J. Ye and X. Y. Ye. "Necessary Optimality Conditions for Optimization Problems with Variational Inequality Constraints". In: *Mathematics of Operations Research* 22.4 (1997), pp. 977–997.

[30] F. Yousefian, A. Nedić, and U. V. Shanbhag. "On Stochastic Mirror-prox Algorithms for Stochastic Cartesian Variational". In: *Set-Valued and Variational Analysis* 26.4 (2018), pp. 789–819.