

A Quantization Procedure for Nonlinear Pricing with an Application to Electricity Markets

Quentin Jacquet, Wim van Ackooij, Clémence Alasseur and Stéphane Gaubert

Abstract—We consider a revenue maximization model, in which a company aims at designing a menu of contracts, given a population of customers. A standard approach consists in constructing an incentive-compatible continuum of contracts, i.e., a menu composed of an infinite number of contracts, where each contract is especially adapted to an infinitesimal customer, taking his type into account. Nonetheless, in many applications, the company is constrained to offering a limited number of contracts. We show that this question reduces to an optimal quantization problem, similar to McEneaney’s *pruning problem* that appeared in the max-plus based numerical methods in optimal control. We develop a new quantization algorithm, which, given an initial menu of contracts, iteratively prunes the less important contracts, to construct an implementable menu of the desired cardinality, while minimizing the revenue loss. We apply this algorithm to solve a pricing problem with price-elastic demand, originating from the electricity retail market. Numerical results show an improved performance by comparison with earlier pruning algorithms.

I. INTRODUCTION

A. Motivation from electricity markets

Electricity retail markets are now open to competition in most countries, and providers are free to design a *menu* of offers/contracts in addition to regulated alternatives (fixed prices), so that each consumer can select among the vast jungle of offers the one which maximizes his utility. In this paper, the choice of a contract is based on the minimization of the invoice (rational choice theory, see e.g. [1]), and we suppose that each customer can adjust his consumption to the electricity prices (*price elasticity*). This phenomenon is highlighted by the current spike in energy prices: consumers are likely to make huge consumption reduction efforts in order to save money.

A key problem for electricity providers is to design an optimal menu of offers, maximizing their revenue, under a restriction on the “size” of the menu (number of contracts). In fact, from an optimization point of view, proposing more contracts increases the revenue, as it allows one to adjust the menu to the individual preferences of the different types of customers. However, in practice, it is essential to restrict the number of contracts, in order to make the commercial offer more visible to agents, easier to understand, and also to keep an implementable menu for the company.

Q. Jacquet, W. van Ackooij and C. Alasseur are with EDF R&D Saclay, Palaiseau, France {quentin.jacquet, wim.van-ackooij, clemence.alasseur}@edf.fr

Q. Jacquet and S. Gaubert are with INRIA, CMAP, Ecole Polytechnique, IP Paris, CNRS, Palaiseau, France stephane.gaubert@inria.fr

B. The optimal nonlinear pricing problem

We consider more generally the revenue maximization problem faced by a seller, called *principal*, or *leader* in the setting of Stackelberg games [2]. This problem has been addressed by the theory of mechanism design [3] through the question of nonlinear pricing. The so-called *monopolist problem* is among the most studied ones: in this approach, the population is represented as a continuum of buyers (called *agents* or *followers*), and a contract can be specifically designed for each agent (continuum menu). In the seminal paper [4], Rochet and Choné study the monopolist problem by introducing a dual approach. In some specific cases (linear-quadratic setting and specific agents distribution), analytic solutions can be found in one [5] or many dimensions [6], via reformulation as welfare maximization using virtual valuation technique. Extending the framework of Rochet and Choné to decomposable variational problem under convexity requirement, Carlier [7] addresses the question of the existence and uniqueness of a solution, and proposes an iterative algorithm. In the specific case \mathbb{R}^2 , Mirebeau [8] introduces a more efficient method using an adaptive mesh based on stencils. The infinite-size menu is therefore characterized by a value-function satisfying the incentive-compatibility conditions as with the full-participation condition, the latter supposing that contracting with the whole population is optimal. Bergemann, Yeh and Zhang recently considered the question of the optimal quantization of a menu [9].

C. Contributions

Our main contribution is the development of new *quantization* algorithms which, given the infinite-size menu, aim at finding the best n -contracts approximation that maximizes the revenue. This 2-step strategy bypasses the combinatorial difficulty tackled in bilevel pricing – see e.g. [10, 11] – where formulations directly embed customer choices over the n contracts, becoming rapidly untractable for large size of menu. We show that the quantization problem is equivalent to the *pruning problem*, which arose, following McEneaney [12], in the development of the max-plus based curse-of-dimensionality attenuation methods in numerical optimal control, see [13, 14, 15], and [16] for an application. In these methods, the value function of an optimal control problem is represented as a supremum of “basis functions”, and one looks for a sparse representation – with a prescribed number of basis functions. In the present application, the basis functions are linear functions, representing contracts. We develop a *greedy descent* algorithm which iteratively removes the less “important” contracts. We consider different

importance measures, taking into account the L_1 and L_∞ approximation errors previously considered in the study of the pruning problem, and also a specific measure of the loss of revenue, see Algorithms 1 and 2. An essential feature of these algorithms is the low incremental cost per iteration, with an update rule requiring only *local* computations – in a “small neighborhood” of the active set of a basis function. To do so, we exploit discrete geometry techniques, by associating to a basis decomposition a polyhedral complex, which is updated dynamically.

To apply this algorithm to the optimal design of a menu in the electricity retail market, we generalize the framework of [7] to allow for a nondecomposable (still convex) cost. Indeed, the revenue of the provider depends on the supply cost, supposed to be an increasing function of the *global* consumption, see e.g. ([17, 18]). In this extended setting, we prove the existence and uniqueness of the solution for the infinite-size menu (Proposition 2.2). The solving of this problem is then tackled by a direct method (discretization of the variational problem). We also take into account the elastic behavior of the customers, who adapt their consumption according to prices, even on an annual scale (we focus here on year-based contracts). We show that, after a change of variables, the addition of a uniform elasticity actually reduces to the previous model (Proposition 4.1). Numerical tests, on a realistic instance (arising from the French electricity market), illustrate the efficiency of our approach both in terms of revenue gain and of computational time, see Figure 3. Our algorithm also allows one to estimate the minimal admissible number of contracts, given a target of acceptable revenue loss by comparison with the infinite-size case.

D. Related works

In the nonlinear pricing context, the restriction to a finite number of offers has been regarded only recently. In [9], the authors analyze the loss of revenue induced by this restriction, exhibiting upper bounds of order $1/n^{2/d}$, where d is the dimension and n the maximal number of contracts. A similar asymptotic error rate arose in a different setting of quantization theory, see e.g. [14]. Moreover, in the linear-quadratic setting of [9], the extreme distributions realizing the worst revenue loss satisfies separability conditions à la Armstrong [6], leading to an explicit expression for the optimal quantization. We do not satisfy these requirements here, as we tackle a broader class of variational problem, hence the need of efficient methods to solve the pricing problems with a finite number of contracts. In [19], a discretization is obtained by writing the utility function as a supremum of finitely many affine functions, and so the solution they obtain can be viewed as a n -contracts menu. However, the scheme also discretizes the population (with the same size as the contracts). In the present application, this is not desirable, since the size of the population has to be much larger than the size of the menu.

The present algorithms should be compared with the pruning methods to compute a sparse representation of a function as maximum of a prescribed number of basis

functions. The pruning problem was shown in [14] to be a continuous version of the facility location problem, a hard combinatorial optimization problem. The pruning algorithms developed in [13, 14] rely on a notion of importance metric, measuring the contribution of each basis function to the approximation error. A basic algorithm in [13, 14] performs a single pass which keeps only the n basis functions with the highest importance metric, the latter being evaluated either by solving a convex programming problem or in approximate way, after a discretization of the state space. A *greedy ascent* algorithm is also implemented in [14], adding incrementally functions by decreasing order of importance. In contrast, the present algorithm does not require a discretization of the state space. Moreover, the use of fast (local) updates of the importance measure allows us to perform a greedy descent starting from the complete family of basis functions, and removing at each stage the less important one. This leads to improved performances on our application case.

The paper is organized as follows: in Section II, we define the nonlinear pricing problem, adapted to our application case, and encompassing the monopolist framework. In Section III, we approximate the continuum menu by a finite set of contracts, and present refined pruning algorithms with local update. Then, in Section IV, we specify the problem encountered in electricity markets, and show how it boils down to the general case of Section II. Finally, we numerically study the effectiveness of our approach in Section V.

II. NONLINEAR PRICING WITH COUPLING COSTS

A. Notation

For two vectors x and y of \mathbb{R}^d , we denote by $\langle x, y \rangle$ the scalar product and $x \odot y$ the entrywise product. Moreover, for a discrete set S , we denote by $|S|$ the cardinality of S .

B. Generalized monopolist problem

Let us consider a heterogeneous population, where each agent in the population is defined by a d -dimensional vector of characteristics $x \in X$. We suppose that $X \subset \mathbb{R}_{>0}^d$ is a compact polyhedral domain. An agent of type x will derive a utility $\langle x, \alpha \odot q_k \rangle - p_k$ from consuming a good k with quality $q_k \in \mathbb{R}_{>0}^d$ and price $p_k \in \mathbb{R}_{>0}$. The vector $\alpha \in (\mathbb{R}^*)^d$ is an exogeneous data, viewed as a varying perception of the quality. The agents are distributed according to ρ satisfying $\int_X \rho(x) dx = 1$.

Let us consider a *monopolist* (principal) who designs a contract menu represented by a pair of functions $x \mapsto (p(x), q(x)) \in P \times Q$. For each agent x , these functions indicate respectively the price and the quality that the agent is supposed to prefer. Here, P and Q are compact subsets of $\mathbb{R}_{>0}$ and $\mathbb{R}_{>0}^d$. To ensure that the contract $(p(x), q(x))$ really satisfies agent of type x (i.e., is optimal for him), an additional constraint on the shape of the function, called *incentive-compatibility* condition is required: denoting by $u(x) := \langle x, \alpha \odot q(x) \rangle - p(x)$ the utility function for the menu designed by the monopolist,

$$u(y) - u(x) \geq \langle y - x, \alpha \odot q(x) \rangle, \forall x, y \in X. \quad (1)$$

Let U_x be the set of admissible values of u for type x :

$$U_x := \{ \langle x, \alpha \odot q \rangle - p \mid (p, q) \in P \times Q \} .$$

Each set U_x is compact by compactness of P and Q .

Proposition 2.1 ([20]): Let $q(\cdot)$ be defined on X , with values in Q . There exists a function $p : X \rightarrow P$ such that $u(\cdot)$ satisfies (1) if and only if

- (i) $u(x) \in U_x$, for $x \in X$,
- (ii) u is convex on X ,
- (iii) $\nabla u(x) = \alpha \odot q(x)$ for a.e. $x \in X$.

The aim of the monopolist is then to maximize a revenue function, defined as

$$J(u, q) := \int_X L(x, u(x), q(x)) dx - C \left(\int_X M(x, q(x)) dx \right), \quad (2)$$

In (2), the cost function C takes as input data aggregated on the whole domain X . Such coupling cost naturally appears in some applications, for instance in electricity retail market, see Section IV.

Assumption 2.1: The integrand L is linear in u and q . Moreover, the integrand M is strictly convex in q , and C is increasing and strictly convex.

In addition to the incentive-compatibility condition, the utility must be greater than a reservation utility:

$$u(x) \geq R(x) . \quad (3)$$

The problem solved by the monopolist is then

$$\max_{u, q} \left\{ J(u, q) \mid \begin{array}{l} u, q \text{ satisfy (1), (3)} \\ (u(x), q(x)) \in U_x \times Q \text{ for } x \in X \end{array} \right\} \quad (4)$$

Theorem 2.2: Under Assumption 2.1, Problem (4) has a unique optimal solution.

The proof of Theorem 2.2 is given in Section VII-A. This result should be compared with [21], where the (decomposable) criteria is defined by an integrand that must satisfy coercivity condition, which entails that a minimizing sequence (u_n) must be bounded in the $W^{1,1}$ Sobolev norm. Here, J is not necessarily coercive. Instead, the compactness argument directly comes with assumptions on P and Q .

C. Resolution of the infinite-size case

As an extension of the monopolist problem, Problem (4) can be solved to optimality through a discretization scheme. In [19], the authors proved the convergence of the discretized problem to the continuous one, which can be extended to nondecomposable cost. Efficient numerical methods have been proposed in [7] and [8]. Let us define a regular grid Σ of X . Each of the methods provides a solution $\{(\hat{p}_i, \hat{q}_i)\}_{i \in \Sigma}$, inducing a convex utility function \hat{u}_Σ that can be represented as the supremum of affine functions, with the notation:

$$\hat{u}_S(x) = \max_{i \in S} \hat{u}_i(x), \quad S \subseteq \Sigma, \quad (5)$$

where $\hat{u}_i : x \in \mathbb{R}^d \mapsto \langle \alpha \odot \hat{q}_i, x \rangle - \hat{p}_i$. In the context of max-plus methods [12, 22], the functions \hat{u}_i are called *basis functions* and can be more general than affine functions, but we focus here on this specific case, as this naturally appears in the model (affine contracts).

III. PRUNING PROCEDURES

A. Pruning method for max-plus basis decomposition

Let us now suppose that the monopolist has a maximal number of n contracts he can design. Given the discretized infinite-size solution u_Σ , the question can be recast as the following combinatorial problem:

$$\min_{S \subseteq \Sigma} \{ d(\hat{u}_S, \hat{u}_\Sigma) \text{ s.t. } |S| \leq n \} , \quad (6)$$

where the function $d(\cdot)$ can be either

- (i) the L_∞ norm $d_\infty(u, v) = \|u - v\|_{L_\infty(X)}$,
- (ii) the L_1 norm $d_1(u, v) = \|u - v\|_{L_1(X)}$,
- (iii) and the J -based criterion $d_J(u, v) = J(v, \alpha^{-1} \odot \nabla v) - J(u, \alpha^{-1} \odot \nabla u)$.

The third case corresponds to the maximization of the function J , where $\alpha \odot q = \nabla u$ thanks to Proposition 2.1.

Theorem 3.1 ([14]): Let $X \subseteq \mathbb{R}^d$ and $v : X \rightarrow \mathbb{R}$ strongly convex of class \mathcal{C}^2 . Then, both L_1 and L_∞ approximation errors are $\Omega\left(\frac{1}{n^{2/d}}\right)$ as $n \rightarrow \infty$.

Theorem 3.1 exhibits an error rate identical to the complexity bound proved in [9] in a different setting.

We define the *importance metric* of basis function i as

$$\nu(S, i) = d(\hat{u}_{S \setminus \{i\}}, \hat{u}_S) . \quad (7)$$

This corresponds to an incremental version of the criteria (6). For the L_∞ and L_1 case, if $\nu(S, i) = 0$, then the i -th basis function does not contribute to the max-sum. Otherwise, if $\nu(S, i) > 0$, then it expresses the maximal difference between the shape of \hat{u}_S with and without \hat{u}_i , depending on the criterion. For the criterion d_J , it expresses the loss of revenue for the principal when contract i is removed.

B. Specific case: minimizing L_∞ error

For a L_∞ approximation error, the importance metric (7) can be computed by solving a linear program, see [14]:

$$\begin{array}{ll} \max_{x \in X, \nu} & \nu \\ \text{s.t.} & \forall j \in S \setminus \{i\}, \quad \hat{u}_i(x) - \hat{u}_j(x) \geq \nu \quad (\lambda_{ij}) \end{array} \quad (P_i^S)$$

In (P_i^S) , we denote by $(\lambda_{ij})_j$ the dual variable associated with each constraint. The set of saturated constraints is then characterized by the positive variables λ_{ij} .

Algorithm 1 Pruning for L_∞ importance metric

```

Require:  $n$  ▷ Desired number of contracts
1:  $S \leftarrow \Sigma$  ▷ Indices of kept contracts
2:  $I \leftarrow \Sigma$  ▷ Indices of problems to re-compute
3: for  $t = 1 : |\Sigma| - n$  do
4:   for  $i \in I$  do
5:      $\nu_i, \lambda_i \leftarrow$  solution of  $(P_i^S)$ 
6:      $J_i \leftarrow \{j \in S \setminus \{i\} \mid \lambda_{ij} > 0\}$ 
7:   end for
8:    $r \leftarrow \arg \min_{i \in S} \nu_i$  ▷ Contract to remove
9:    $S \leftarrow S \setminus \{r\}$ 
10:   $I \leftarrow \{i \in S \mid r \in J_i\}$ 
11: end for
12: return  $S$ 

```

Algorithm 1 describes a *greedy descent* procedure: we start from the complete set of contracts S , and iteratively remove the less important contract exploiting a fast local update of the importance metric. Compared with [14], the importance metric is computed *exactly*, i.e., without discretization of the space X . Moreover, we take advantage of the linearity of the basis functions \hat{u}_i to exploit the optimal dual variables λ_{ij} in the linear program (P_i^S) :

Proposition 3.2 (Local update): Let λ_{ij} be the optimal dual variables in (P_i^S) for a contract $i \in S$. Then, the importance metric of i stays *unchanged* when we remove a contract $j \in S$ s.t. $\lambda_{ij} = 0$, i.e., $\nu(S \setminus \{j\}, i) = \nu(S, i)$.

Proposition 3.2 ensures the correctness of Algorithm 1, where we only re-compute at each iteration the values ν_i for a very small subset of Σ . This leads to a huge gain in computation time, see Section V.

C. L_1 and J -based approximation error

Contrary to the L_∞ case, the computation exploits the geometric structure. Indeed, the representation of the function \hat{u}_S as a maximum of basis functions' $\hat{u}_j, j \in S$ induces a polyhedral complex, in which every function \hat{u}_i determines a polyhedral cell C_i , consisting of the types $x \in X$ such that $\hat{u}_S(x) = \hat{u}_i(x)$. Removing a basis function \hat{u}_i from the supremum $\hat{u}_S = \sup_{j \in S} \hat{u}_j$ yields a local modification of the latter supremum, concentrated on a neighborhood of the cell C_i . Hence, we will need to compute at each iteration the neighbors of each contract cell C_i with $i \in S$. This idea may be compared with the notion of Delaunay triangulation associated to a Voronoï diagram [23]. During the algorithm, we keep in memory two sets: J_i represents the neighboring cells of cell i , and V_i is the vertex representation of cell i . Two routines are used for both the L_1 and J -based criterion:

- ◊ $\text{VREP}(S, i)$ returns the V-representation (representation by vertices) of the polyhedral cell C_i induced by contract i for a given set S , taking as input the H-representation (representation by half-spaces) $\{x \in X \mid \hat{u}_i(x) \geq \hat{u}_j(x), \forall j \in S\}$ of the cell i . This is done using the revised *reverse search* algorithm implemented in the library `lrs`, see [24].
- ◊ $\text{UPDATENEIGHBORS}((V_S)_{i \in I})$ updates the neighbors of each cell $i \in I$ knowing the vertex representation.

Proposition 3.3 (Local update): The importance metric of a contract $i \in S$ stays *unchanged* when we remove a contract j which is not in the neighborhood of i , i.e., $\nu(S \setminus \{j\}, i) = \nu(S, i)$ for $j \in S \setminus J_i$.

Proposition 3.3 ensures the correctness of Algo. 2, where we only re-compute vertex representations for a small subset of contracts (corresponding to the neighboring cells of the lastly removed contract, see line 8 of the algorithm). This local update is illustrated in Figure 1. The update of the importance metric in line 11 differs between the L_1 and J -based cases, and is described in Algos. 3a–3b. In Algo. 3a, the integral that appears in the computation of ν_i can be evaluated analytically using Green's formula, as it integrates a linear form over a polytope, see Appendix VII-B. In Algo. 3b, δ_L can be computed in the same way. For M_0 and

δ_M , this generally involves the integration of the function $x \mapsto M(x, \hat{q}_i)$. In the present application, this function is linear, and so the direct integration is possible, see (14)–(15).

Algorithm 2 Pruning with local update (for L_1 and J -based)

```

Require:  $n$  ▷ Desired number of contracts
1: for  $i \in \Sigma$  do
2:    $V_i \leftarrow \text{VREP}(\Sigma, i)$  ▷ Vertex representation
3: end for
4:  $S \leftarrow \Sigma$  ▷ Indices of kept contracts
5:  $I \leftarrow \Sigma$  ▷ Indices of problems to re-compute
6: for  $t = 1 : |\Sigma| - n$  do
7:    $(J_i)_{i \in I} \leftarrow \text{UPDATENEIGHBORS}((V_i)_{i \in I})$ 
8:   for  $i \in I, j \in J_i$  do
9:      $F_{j,-i} \leftarrow \text{VREP}(S \setminus \{i\}, j)$  ▷ Future cells
10:  end for
11:   $\nu \leftarrow \text{UPDATEIMP METRIC}(I, (V_i)_{i \in S}, (F_{j,-i})_{j \in J_i, i \in S})$ 
12:   $r \leftarrow \arg \min_{i \in S} \nu_i$  ▷ Contract to remove
13:   $S \leftarrow S \setminus \{r\}$ 
14:  for  $j \in J_r$  do
15:     $V_j \leftarrow F_{j,-r}$  ▷ Update vertex representation
16:  end for
17:   $I \leftarrow J_r$ 
18: end for
19: return  $S$ 

```

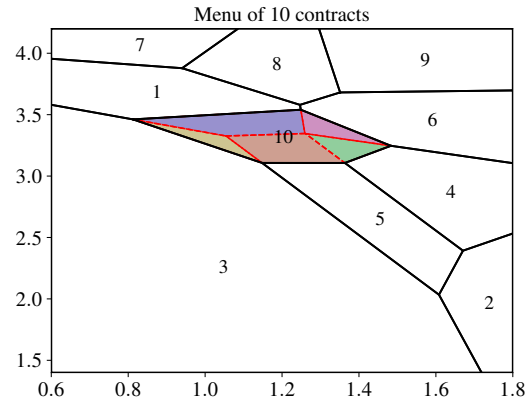


Fig. 1: Evaluation of contract by dividing into subregions ($d = 2$). The green polyhedron corresponds to $F_{4,-10} \cap V_{10}$.

Algorithm 3a UPDATEIMP METRIC (L_1 error)

```

Require:  $I, (V_i)_{i \in S}, (F_{j,-i})_{i \in I, j \in J_i}$ 
1: for  $i \in I$  do ▷ Update metric on cells
2:    $\nu_i \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} (\hat{u}_i(x) - \hat{u}_j(x)) dx$ 
3: end for
4: return  $\nu$ 

```

Algorithm 3b UPDATEIMP METRIC (J -based error)

```

Require:  $I, (V_i)_{i \in S}, (F_{j,-i})_{i \in I, j \in J_i}$ 
1:  $M_0 \leftarrow \sum_{i \in S} \iint_{V_i} M(x, \hat{q}_i) dx$ 
2: for  $i \in S$  do ▷ Update metric on cells
3:    $\delta_L \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} L(x, \hat{u}_i(x), \hat{q}_i) - L(x, \hat{u}_j(x), \hat{q}_j) dx$ 
4:    $\delta_M \leftarrow \sum_{j \in J_i} \iint_{F_{j,-i} \cap V_i} M(x, \hat{q}_j) - M(x, \hat{q}_i) dx$ 
5:    $\nu_i \leftarrow \delta_L - C(M_0) + C(M_0 + \delta_M)$ 
6: end for

```

Proposition 3.4 (Critical steps): Let m be the maximum number of neighbors of a polyhedral cell during the execution of the algorithm (for all t and i , $|J_i| \leq m$). Then,

- ◊ The number of linear programs (P_i^S) solved in Algo. 1 is in $O(m|\Sigma|)$,
- ◊ The number of computations of a vertex representation of a polyhedral cell (calls to $\text{VREP}(S, i)$ in Algo. 2 / reverse search) is in $O(m^2|\Sigma|)$.

By comparison with Proposition 3.4, a naïve implementation (full recomputation of the importance metric at each step) of the two algorithms would respectively lead to a number of critical steps in $O(|\Sigma|^2)$ and $O(m|\Sigma|^2)$. Indeed, each linear program P_i^S can be solved in polynomial time (by an interior point method), and reverse search has an incremental running time of $O(|\Sigma|d)$ per vertex if the input is nondegenerate, see [24].

IV. APPLICATION TO ELECTRICITY MARKETS

A. Price elasticity

Let us consider a provider holding several contracts, each of them defined by a fixed price component $p \in \mathbb{R}$ (in €), and d variable price components $z \in \mathbb{R}^d$ (in €/kWh). In France, the contracts often take into account $d = 2$ time periods, with different prices for Peak / Off-peak consumptions. Moreover, the price coefficients (p, z) of each contract are supposed to belong to a non-empty polytope $P \times Z \subset \mathbb{R}^{d+1}$:

Assumption 4.1: Let p^-, p^+ be in $\mathbb{R}_{>0}$ and z^-, z^+ be in $\mathbb{R}_{>0}^d$. Then, $P = [p^-, p^+]$, and the polytope Z is of the following form:

$$Z := \{z^- \leq z \leq z^+ \mid z_{i_1} \leq \kappa_{i_1, i_2} z_{i_2} \text{ for } i_1 \leq_{\mathcal{P}} i_2\} ,$$

where \mathcal{P} is a *partially ordered set* (poset) of $\{1, \dots, d\}$, and $\leq_{\mathcal{P}}$ the ordering relation, and $\kappa_{i_1, i_2} > 0$. When $\kappa \equiv 1$, $z^- \equiv 0$ and $z^+ \equiv 1$, Z is known as an *order polytope* [25]. Assumption 4.1 is natural for the electricity pricing problem: the price can be freely determined within a box (bounds), as long as some inequalities between peak price coefficients and off-peak price coefficients are fulfilled.

We suppose that each individual in the (infinite-size) population is characterized by a *reference* consumption vector $\tilde{x} \in X \subset \mathbb{R}_{>0}^d$. Here, supposing a continuum of agents is justified since we consider in the application case the population of a whole country. We suppose that the consumption is *elastic* to prices, i.e., a consumer can deviate from its reference consumption \tilde{x} . In addition, we suppose that electricity elasticity can be captured into a utility-based framework, see e.g. [26] for the properties that the utility must satisfy. Here, we focus on isoelastic utilities:

Assumption 4.2 (Isoelastic utility function): For a reference consumption \tilde{x} , the utility of consuming an amount of energy $x \in \mathbb{R}_{>0}^d$ is depicted through a *Constant Relative Risk Aversion* (CRRA, [27, 28]) or isoelastic utility:

$$\mathcal{U}_{\tilde{x}} : x \in \mathbb{R}_{>0}^d \mapsto \frac{1}{\eta} \sum_{i=1}^d \beta_{\tilde{x}i} (x_i)^\eta, \quad \eta \in (-\infty, 0) \cup (0, 1] , \quad (8)$$

where $\beta_{\tilde{x}} \in \mathbb{R}_{>0}^d$ expresses the intensity of energy needs. The coefficient η is called the *risk aversion* coefficient.

In this context, this elasticity measure depicts the easiness of a customer to adopt another energy source to fulfill his needs. In [28], the authors model the electric elasticity by this kind of utility function, and separate the case $\eta < 0$ and $\eta \in (0, 1]$. The first regime ($\eta < 0$) will model a household consumption: the satisfaction coming from consuming energy saturates to a maximum utility, and a zero consumption is prohibited. In contrast, the second regime ($\eta \in (0, 1]$) will represent the high flexibility of the industrial sector, which can adapt more easily its consumption according to price. We refer to [29] and references therein for empirical studies on the intensity of the elasticity coefficient η .

For a contract defined by price coefficients $(p, z) \in \mathbb{R} \times \mathbb{R}^d$, a consumer \tilde{x} will optimize his consumption in order to maximize the *welfare function*, obtained by subtracting the electricity cost to (8):

$$\mathcal{U}_{\tilde{x}}^* : (p, z) \in \mathbb{R} \times \mathbb{R}^d \mapsto \max_{x \in \mathbb{R}_{>0}^d} \{\mathcal{U}_{\tilde{x}}(x) - \langle x, z \rangle\} - p . \quad (9)$$

We denote by $\mathcal{U}_{\tilde{x}}^*$ the welfare function as the maximization term in (9) corresponds to a Fenchel-Legendre transform up to a change of sign. As a consequence, $\mathcal{U}_{\tilde{x}}^*$ is convex and nonincreasing. We now make the following assumption to fix the value of β :

Assumption 4.3: The reference consumption $\tilde{x} \in \mathbb{R}^d$ is obtained for reference prices $\tilde{p} \in \mathbb{R}$ and $\tilde{z} \in \mathbb{R}^d$.

Under Assumption 4.3, the optimal consumption of customer \tilde{x} on period $i \in \{1, \dots, d\}$, denoted $\mathcal{E}_{\tilde{x}i}$, is given by

$$\mathcal{E}_{\tilde{x}i}(z) = \tilde{x}_i (z_i / \tilde{z}_i)^{\frac{-1}{1-\eta}} \geq 0 , \quad (10)$$

and the welfare function is given by

$$\mathcal{U}_{\tilde{x}}^*(p, z) = \left(\frac{1}{\eta} - 1\right) \sum_{i=1}^d \tilde{x}_i \tilde{z}_i (z_i / \tilde{z}_i)^{\frac{-\eta}{1-\eta}} - p . \quad (11)$$

Equations (10) and (11) are obtained from the first order optimality condition (zero derivative) for (9) ($\beta_{\tilde{x}i} = \tilde{z}_i (\tilde{x}_i)^{1-\eta}$).

B. Infinite-size menu of offers

In this section, we relax the assumption of a finite number of contracts, by supposing that the provider is able to define *as many offers as* consumers. Therefore, the infinite-size menu of offers can be represented by two functions $p : X \rightarrow \mathbb{R}$ and $z : X \rightarrow \mathbb{R}^d$, representing respectively the fixed price component and the variable price components. Let us define the (weighted) invoice of a consumer as

$$\mathcal{L}_{\tilde{x}} : (p, z) \in \mathbb{R} \times \mathbb{R}^d \mapsto (p + \langle \mathcal{E}_{\tilde{x}}(z), z \rangle) \rho(\tilde{x}) , \quad (12)$$

where $\rho(\tilde{x}) \geq 0$ represents the density of customers with reference consumption \tilde{x} . The provider's revenue maximization problem is then

$$\max_{p, z} \mathcal{J}^1(p, z) - \mathcal{J}^2(z) \quad (13a)$$

$$\text{s.t. } \mathcal{U}_x^*(p(x), z(x)) \geq \mathcal{U}_x^*(p(y), z(y)), \quad \forall x, y \in X \quad (13b)$$

$$\mathcal{U}_x^*(p(x), z(x)) \geq R(x), \quad \forall x \in X \quad (13c)$$

$$p(x) \in P, \quad z(x) \in Z \quad (13d)$$

where $\mathcal{J}^1(p, z) = \int_X \mathcal{L}_x(p(x), z(x)) dx$ and $\mathcal{J}^2(z) = C \left(\int_X \sum_{i=1}^d \mathcal{E}_{x_i}(z(x)) \rho(x) dx \right)$.

Equations (13b) and (13c) are respectively the *incentive-compatibility condition* and *participation constraint*. Taking C as a strictly convex increasing function of the global consumption is often considered in the literature. In particular, this cost function is often modeled as a piecewise linear function, see e.g. [18], or as a quadratic function, see e.g. [17]. In fact, the marginal cost to supply electricity is not constant and increases with the consumption. The convexity of the reservation utility is also a classical assumption, as this reservation utility should be a supremum over the utilities of alternative offers (each of them being linear function of the reference consumption).

Let us make the following change of variables:

$$q_i := (z_i / \tilde{z}_i)^{\frac{-\eta}{1-\eta}}.$$

Then, the consumption on period $i \in \{1, \dots, d\}$ is a convex function of q_i , expressed as $\mathcal{E}_{\tilde{x}_i}(q_i) = \tilde{x}_i [q_i]^{\frac{1}{\eta}}$, and both the utility and the weighted invoice now read as linear functions of p and q : defining $\alpha = (\eta^{-1} - 1)\tilde{z}$,

$$\begin{aligned} u(x) &:= \langle x, \alpha \odot q(x) \rangle - p(x), \\ L(x, u(x), q(x)) &:= \left(\frac{1}{\eta} \langle x, \tilde{z} \odot q(x) \rangle - u(x) \right) \rho(x), \end{aligned} \quad (14)$$

Theorem 4.1: Under Assumption 4.1, the provider's revenue maximization problem (13) is equivalent to a monopolist problem of the form (4) with

$$M(x, q(x)) := \rho(x) \sum_{i=1}^d x_i [q_i(x)]^{\frac{1}{\eta}} \quad (15)$$

and, if $\eta < 0$,

$$Q = \left\{ q \in \mathbb{R}^d \left| \begin{array}{l} (z^- / \tilde{z})^{\frac{-\eta}{1-\eta}} \leq q \leq (z^+ / \tilde{z})^{\frac{-\eta}{1-\eta}} \\ q_{i_1} \leq \left(\kappa_{i_1, i_2} \frac{\tilde{z}_{i_2}}{\tilde{z}_{i_1}} \right)^{\frac{-\eta}{1-\eta}} q_{i_2} \text{ for } i_1 \leq_{\mathcal{P}} i_2 \end{array} \right. \right\},$$

otherwise,

$$Q = \left\{ q \in \mathbb{R}^d \left| \begin{array}{l} (z^+ / \tilde{z})^{\frac{-\eta}{1-\eta}} \leq q \leq (z^- / \tilde{z})^{\frac{-\eta}{1-\eta}} \\ q_{i_1} \geq \left(\kappa_{i_1, i_2} \frac{\tilde{z}_{i_2}}{\tilde{z}_{i_1}} \right)^{\frac{-\eta}{1-\eta}} q_{i_2} \text{ for } i_1 \leq_{\mathcal{P}} i_2 \end{array} \right. \right\}$$

Proof: Owing to assumption on the set Q and the strict monotonicity of $z \mapsto z^{\frac{-\eta}{1-\eta}}$ (increasing for $\eta < 0$ and decreasing for $\eta > 0$), one can explicitly derive the form of Q . The rest of the formulation is immediate. ■

V. NUMERICAL RESULTS

A. Instance

The numerical results were obtained on a laptop i7-1065G7 CPU@1.30GHz. We provide in Table I the values of the parameters used in the application. In particular, we consider reference prices (\hat{p}, \hat{z}) corresponding to French regulated prices, and reference consumption spread around the average annual French consumption per household ($\mathcal{E}_{\text{moy}} = 4\text{MWh}$). The cost function is taken as a quadratic function, scaled so that the marginal cost $C'(\mathcal{E}_{\text{moy}}) = 0.08\text{€/kWh}$.

In comparison, the production cost is estimated in France around 0.05€/kWh for nuclear plants¹ and up to 0.09€/kWh for wind energy².

η	-0.1
\hat{p}	140€
\hat{z}	(0.174, 0.19)€/kWh
$C(\cdot)$	$0.01(\cdot)^2$
(p^-, p^+)	(0, 500)€
(q_1^-, q_1^+)	(0.05, 0.5)€/kWh
(q_2^-, q_2^+)	(0.05, 0.5)€/kWh
ρ	Uniform([0.6, 1.8] × [1.4, 4.2])
$R(\cdot)$	linear function (one regulated contract)

TABLE I: Instance used in the numerical results.

We display in Figure 2 the infinite-size menu and the quantized solution for two different sizes of menu (25 contracts and 10 contracts). In each cell \mathcal{C}_i , the contract i brings to customers of reference consumption $x \in \mathcal{C}_i$ the maximal utility given the quantized menu, i.e., $\hat{u}_S(x) = \hat{u}_i(x)$ for $x \in \mathcal{C}_i$. We observe that there is a region/cell (light gray region) where the monopolist reproduces the alternative option (of utility R). On the other side, for high consumption (peak or off-peak), the monopolist manages to design contracts that provide strictly higher utility than the regulated offer, and at the same time, procure to the monopolist a higher revenue.

B. Comparison of pruning objectives

In the upper graph of Figure 3, the three pruning objectives studied in the paper (L_∞ , L_1 and J -based) are compared with the 1-step approach of [13, 14]. The approach consists in sorting the importance metrics for all $i \in \Sigma$, and directly taking the n contracts with highest importance metric (here we consider the J -based importance metric). We display the relative objective loss, defined as $1 - J_t / J_{\text{ref}}$, where J_t is the objective for a menu of size t and J_{ref} the objective obtained with the infinite-size menu. Note that removing a contract can induce a violation of the full-participation constraint ($u \geq R$). Therefore, in order to recover a feasible solution at each iteration, we lift up the solution with the simple rule $u \leftarrow u + \max\{\max_{x \in X} \{R(x) - u(x)\}, 0\}$.

On this example, the pruning procedure of Algo. 2 (greedy descent) leads to a significant loss reduction, whatever the criterion, compared with the 1-step approach. As expected, we observe that the J -based pruning has the smallest relative loss in the objective, as we minimize the error at each iteration of the process. In contrast, the L_∞ -norm does not capture sufficiently well the behavior of the objective function J , and has larger objective loss, even for a large number of contracts.

We also depicted the cumulated time along the iterations in the lower graph of Figure 3 (we do not display the time for the 1-step procedure, as it is very fast, in less than 0.5s). For comparison, we add the cumulative time of a “naïve” J -based pruning, recomputing at each iteration the importance metric of each cell (global update). On this

¹CRE (2022), *Délibération n° 2022-45*

²ADEME (2016), *Coûts des énergies renouvelables en France*

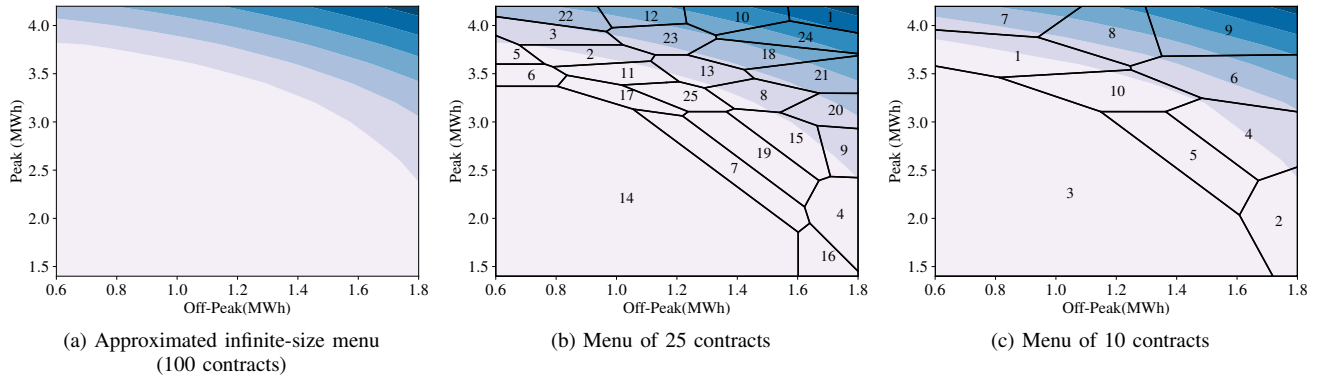


Fig. 2: L_1 -norm pruning for the electricity market case.

The normalized utility $u - R$ is depicted with colormap (light gray corresponds to the zero value and blue to high value).

example, we observe that the computational time is already reduced by a factor almost 3 (this factor would be greater in higher dimension, as the neighborhood would be larger). As expected, the L_∞ criterion is the fastest, owing to the fast local update rule exploiting the sparsity of optimal Lagrange multipliers (Algorithm 1), and the J -based and L_1 -norm criteria have similar computational time, as they use the same algorithmic architecture, see Algo. 2. In terms of loss minimization, the J -based pruning shows a loss of revenue reduced by a factor of around 2 by comparison with other methods. This approach allows us to determine the minimum number of contracts given an admissible revenue loss: e.g., Figure 3 shows that, with a J -based quantization, a menu of 10 contracts suffices to limit the revenue loss to 4%.

supply costs depend on the global consumption. We have developed a quantization procedure, allowing to maximize the revenue of a provider, given a cardinality constraint on the set of contracts. This relies on refined pruning procedures, inspired by the max-plus basis methods in numerical optimal control. In particular, we exploited the local nature of the pruning process, in order to reduce the computational time. Thus, this leads to a new class of applications for methods originally developed in optimal control, and this also improves the complexity of a key ingredient of these methods.

A strong parallel with vector quantization can be made, see e.g. [30]. In this context, a different quantization problem is addressed by Lloyd's procedures, *ibid.*. Whether these ideas can be adapted to the quantization of the maximum of affine functions with revenue criterion is left for further work.

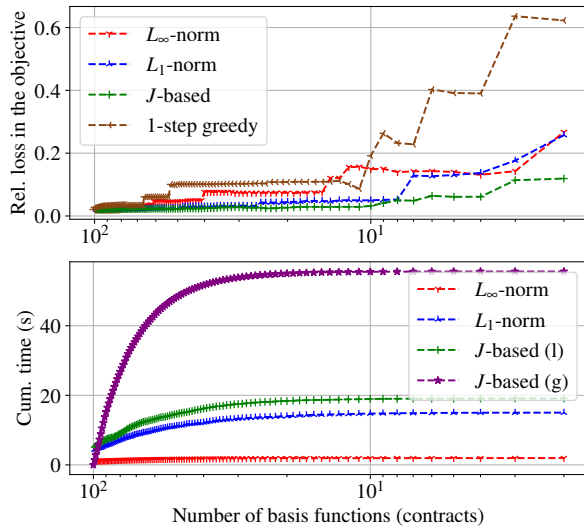


Fig. 3: Comparison of error bounds for the three types of pruning objective. The upper graph shows the loss of optimality induced by a reduced number of contracts. The lower graph depicts cumulative time along the iterations. (g) stands for global update while (l) stands for local update.

VI. CONCLUSION

We have addressed a nonlinear pricing problem incorporating coupling costs. This arises in electricity markets, where

REFERENCES

- [1] J. Scott. "Rational choice theory". In: *Understanding contemporary society: Theories of the present* 129 (2000), pp. 126–138.
- [2] M. Simaan and J. B. Cruz. "On the Stackelberg strategy in nonzero-sum games". In: *Journal of Optimization Theory and Applications* 11.5 (1973), pp. 533–555.
- [3] T. Börgers, D. Krämer, and R. Strausz. *An Introduction to the Theory of Mechanism Design*. Oxford University Press, New York, May 2015.
- [4] J.-C. Rochet and P. Choné. "Ironing, sweeping, and multi-dimensional screening". In: *Econometrica* (1998), pp. 783–826.
- [5] M. Mussa and S. Rosen. "Monopoly and product quality". In: *Journal of Economic theory* 18.2 (1978), pp. 301–317.
- [6] M. Armstrong. "Multiproduct Nonlinear Pricing". In: *Econometrica* 64.1 (Jan. 1996), p. 51.
- [7] G. Carlier and X. Dupuis. "An iterated projection approach to variational problems under generalized convexity constraints". In: *Applied Mathematics and Optimization* 76.3 (2017), pp. 565–592.
- [8] J.-M. Mirebeau. "Adaptive, anisotropic and hierarchical cones of discrete convex functions". In: *Numerische Mathematik* 132.4 (2016), pp. 807–853.
- [9] D. Bergemann, E. Yeh, and J. Zhang. "Nonlinear pricing with finite information". In: *Games and Economic Behavior* 130 (Nov. 2021), pp. 62–84.
- [10] M. Labbé, P. Marcotte, and G. Savard. "A bilevel model of taxation and its application to optimal highway pricing". In: *Management science* 44 (1998), pp. 1608–1622.

- [11] E. Baldwin and P. Klemperer. “Understanding Preferences: “Demand Types”, and the Existence of Equilibrium With Indivisibilities”. In: *Econometrica* 87.3 (2019), pp. 867–932.
- [12] W. M. McEneaney. “A Curse-of-Dimensionality-Free Numerical Method for Solution of Certain HJB PDEs”. In: *SIAM Journal on Control and Optimization* 46.4 (Jan. 2007).
- [13] W. M. McEneaney, A. Deshpande, and S. Gaubert. “Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs”. In: *2008 American Control Conference*. IEEE, June 2008.
- [14] S. Gaubert, W. McEneaney, and Z. Qu. “Curse of dimensionality reduction in max-plus based approximation methods: Theoretical estimates and improved pruning algorithms”. In: *IEEE Conference on Decision and Control and European Control Conference*. IEEE, Dec. 2011.
- [15] S. Gaubert, Z. Qu, and S. Sridharan. “Bundle-based pruning in the max-plus curse of dimensionality free method”. In: *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems July 7-11, 2014, Groningen, The Netherlands*. 2014, pp. 166–172.
- [16] W. M. McEneaney and P. M. Dower. “The Principle of Least Action and Fundamental Solutions of Mass-Spring and N-Body Two-Point Boundary Value Problems”. In: *SIAM Journal on Control and Optimization* 53.5 (2015).
- [17] W. van Ackooij, I. D. Lopez, A. Frangioni, F. Lalacandra, and M. Tahanan. “Large-scale unit commitment under uncertainty: an updated literature survey”. In: *Annals of Operations Research* 271.1 (Sept. 2018), pp. 11–85.
- [18] E. Alekseeva, L. Brotcorne, S. Lepaul, and A. Montmeat. “A bilevel approach to optimize electricity prices”. In: *Yugoslav Journal of Operations Research* (2019).
- [19] I. Ekeland and S. Moreno-Bromberg. “An algorithm for computing solutions of variational problems with global convexity constraints”. In: *Numerische Mathematik* 115.1 (Oct. 2009), pp. 45–69.
- [20] J.-C. Rochet. “A necessary and sufficient condition for rationalizability in a quasi-linear context”. In: *Journal of Mathematical Economics* 16.2 (Jan. 1987), pp. 191–200.
- [21] G. Carlier. “A general existence result for the principal-agent problem with adverse selection”. In: *Journal of Mathematical Economics* 35.1 (Feb. 2001), pp. 129–150.
- [22] M. Akian, S. Gaubert, and A. Lakhoua. “The Max-Plus Finite Element Method for Solving Deterministic Optimal Control Problems: Basic Properties and Convergence Analysis”. In: *SIAM Journal on Control and Optimization* 47.2 (Jan. 2008), pp. 817–848.
- [23] S. Fortune. “Voronoi Diagrams and Delaunay Triangulations”. In: *Lecture Notes Series on Computing*. World Scientific, Jan. 1995, pp. 225–265.
- [24] D. Avis. “A Revised Implementation of the Reverse Search Vertex Enumeration Algorithm”. In: *Polytopes — Combinatorics and Computation*. Ed. by G. Kalai and G. M. Ziegler. Basel: Birkhäuser Basel, 2000, pp. 177–198.
- [25] R. P. Stanley. “Two poset polytopes”. In: *Discrete & Computational Geometry* 1.1 (Mar. 1986), pp. 9–23.
- [26] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong. “Advanced Demand Side Management for the Future Smart Grid Using Mechanism Design”. In: *IEEE Transactions on Smart Grid* 3.3 (2012), pp. 1170–1180.
- [27] R. S. Pindyck. “Uncertain outcomes and climate change policy”. In: *Journal of Environmental Economics and Management* 63.3 (May 2012), pp. 289–303.
- [28] C. Alasseur, I. Ekeland, R. Élie, N. H. Santibáñez, and D. Possamaï. “An Adverse Selection Approach to Power Pricing”. In: *SIAM Journal on Control and Optimization* 58.2 (Jan. 2020), pp. 686–713.
- [29] A. Niromandfam and S. P. Choboghloo. *Modeling electricity demand, welfare function and elasticity of electricity demand based on the customers risk aversion behavior*. 2020.
- [30] G. Pagès. “Introduction to vector quantization and its applications for numerics”. In: *ESAIM: Proceedings and Surveys* 48 (Jan. 2015). Ed. by N. Champagnat, T. Lelièvre, and A. Nouy, pp. 29–79.
- [31] I. Ekeland and R. Temam. *Convex analysis and variational problems*. SIAM, 1999.

VII. APPENDIX

A. Proof of Proposition 2.2

Using Proposition 2.1, for any solution, $\nabla u(x) = \alpha \odot q(x)$ for a.e. $x \in X$. We then directly study the existence and uniqueness in $(u, \nabla u)$.

Existence. Let $H^1(X)$ be the Sobolev space associated with X . We define

$$\mathcal{K} = \left\{ u \left| \begin{array}{l} u \text{ convex and } u \geq R \\ u(x) \in U_x, \alpha^{-1} \odot \nabla u(x) \in Q, \forall x \in X \end{array} \right. \right\}.$$

The set \mathcal{K} is a closed, convex and bounded subset of $H^1(X)$ (it is bounded since X is bounded and $\|u\|_{L^\infty}$ and $\|\nabla u\|_{L^\infty}$ are bounded too; it is convex since R is convex).

Besides, J is concave (Assumption 2.1). Moreover, as $Q \subset \mathbb{R}_{>0}^d$, there exist $a, b > 0$ such that for any $x \in X$, $a \leq \|\nabla u(x)\| \leq b$. Therefore, there exists $c \in \mathbb{R}_+$ such that $|J(u, \nabla u)| \leq c$, and as a consequence, J is continuous on \mathcal{K} , see [31, Chapter 1, Proposition 2.5].

Using the fact that $H^1(X)$ is reflexive and [31, Chapter 2, Proposition 1.2], Problem (4) admits at least one solution.

Uniqueness. (Same arguments as in [4]) Let now consider two distinct solutions u_1 and u_2 . Then, if $\nabla u_1 \neq \nabla u_2$ on a measurable subset, any function $tu_1 + (1-t)u_2$ is valid and gives a strictly better solution than u_1 and u_2 (due to strict convexity of the cost function $u \mapsto C(\int_X M(x, \nabla u(x))dx)$ and linearity of L). Therefore, $u_1 - u_2$ is a constant function. By linearity of L , the objective value obtained with u_1 and u_2 differ by the same constant. This contradicts the optimality of the two solutions u_1 and u_2 .

B. Fast metric updates using Green’s formula

The next proposition allows us to implement efficiently the local updates of the importance metric performed in 3a–3b.

Proposition 7.1: Let P a 2D-polytope describes by its vertices $(x_i, y_i) \in \mathbb{R}^2$ (counter-clockwise ordered). Then for any $a, b, c \in \mathbb{R}$,

$$\iint_P (ax + by + c) dx dy = \sum_{i=1}^N \left[\int_{y_i}^{y_{i+1}} b(q_i + \frac{1}{\tau_i} y) y dy - \int_{x_i}^{x_{i+1}} (ax + c)(p_i + \tau_i x) dx \right],$$

with $\tau_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$, $p_i := y_i - \tau_i x_i$ and $q_i := x_i - \frac{1}{\tau_i} y_i$.

Proof: The application of the Green formula gives :

$$\iint_P (ax + by + c) dx dy = \oint_{\mathcal{C}_P} (bxy) dy - (ax + c)y dx,$$

where \mathcal{C}_P is the contour of the polytope P . We then decompose on each edges, and use the change of variable $x = q + y/\tau$ in the first integral and $y = x + \tau x$ in the second one. ■