



**HAL**  
open science

## Decomposition of convex high dimensional aggregative stochastic control problems

Adrien Seguret, Clémence Alasseur, Joseph Frédéric Bonnans, Antonio de Paola, Nadia Oudjane, Vincenzo Trovato

► **To cite this version:**

Adrien Seguret, Clémence Alasseur, Joseph Frédéric Bonnans, Antonio de Paola, Nadia Oudjane, et al.. Decomposition of convex high dimensional aggregative stochastic control problems. Applied Mathematics and Optimization, 2023, 88 (8), pp.35. 10.1007/s00245-023-09977-1 . hal-02917014v5

**HAL Id: hal-02917014**

**<https://inria.hal.science/hal-02917014v5>**

Submitted on 2 Mar 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Decomposition of convex high dimensional aggregative stochastic control problems\*

Adrien Seguret<sup>1</sup>      Clemence Alasseur<sup>2</sup>      J. Frédéric Bonnans<sup>3</sup>      Antonio De Paola<sup>4</sup>  
 Nadia Oudjane<sup>5</sup>      Vincenzo Trovato<sup>6</sup>

## Abstract

We consider the framework of convex high dimensional stochastic control problems, in which the controls are aggregated in the cost function. As first contribution, we introduce a modified problem, whose optimal control is under some reasonable assumptions an  $\varepsilon$ -optimal solution of the original problem. As second contribution, we present a decentralized algorithm whose convergence to the solution of the modified problem is established. Finally, we study the application of the developed tools in an engineering context, studying a coordination problem for large populations of domestic thermostatically controlled loads (TCLs).

**Keywords:** Stochastic optimization, Lagrangian decomposition, Uzawa’s algorithm, Stochastic gradient, Thermostatically controlled loads

## 1 Introduction

The present article aims at solving a high dimensional stochastic control problem  $(P_1)$  involving a large number  $n$  of agents indexed by  $i \in \{1, \dots, n\}$ , of the form:

$$(P_1) \quad \begin{cases} \min_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left[ F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i(\omega^i, \omega^{-i}) \right) + \frac{1}{n} \sum_{i=1}^n G_i(u^i(\cdot, \omega^{-i}), \omega^i) \right]. \end{cases} \quad (1.1)$$

Here the noise  $\omega := (\omega^1, \dots, \omega^n)$  belongs to  $\Omega := \prod_{i=1}^n \Omega^i$ , where  $(\Omega^i, \mathcal{F}^i, \mu^i)$  is a probability space, and  $(\Omega, \mathcal{F}, \mu)$  is the corresponding product probability space. Let  $\omega^{-i} := (\omega^1, \dots, \omega^{i-1}, \omega^{i+1}, \dots, \omega^n)$  denote an element of the space  $\Omega^{-i} := \prod_{j=1, j \neq i}^n \Omega^j$ . The associated product probability space is  $(\Omega^{-i}, \mathcal{F}^{-i}, \mu^{-i})$ , where  $\mathcal{F}^{-i} := \otimes_{j=1, j \neq i}^n \mathcal{F}^j$  and  $\mu^{-i} := \prod_{j=1, j \neq i}^n \mu^j$ . Each decision variable  $u^i$  is a random variable (i.e. is  $\mathcal{F}$ -measurable), square summable with value in a Hilbert space  $\mathbb{U}$  so that  $u := (u^1, \dots, u^n)$  belongs to  $L^2(\Omega, (\mathbb{U})^n)$ . The function  $\omega^i \mapsto u^i(\omega^i, \omega^{-i})$  is denoted by  $u^i(\cdot, \omega^{-i})$  and is a.s. (in  $\omega^{-i}$ )  $\mathcal{F}^{-i}$ -measurable and belongs to  $L^2(\Omega^i, \mathbb{U})$ . Also,  $\mathcal{U} := \prod_{i=1}^n \mathcal{U}_i$  where  $\mathcal{U}_i$  is, for  $i = 1$  to  $n$ , a closed convex subset of  $L^2(\Omega^i, \mathbb{U})$ . In the application to dynamical problems, the constraint  $u^i \in \mathcal{U}_i$  includes the constraint of adaptation of  $u^i$  to some filtration. If each  $u^i$  is a random variable of  $\omega^i$ , for  $i = 1$  to  $n$ , we say that  $u$  is a decentralized decision variable.

\*The first, second, third and fifth author thank the FiME Lab (Institut Europlace de Finance). The third author was supported by the PGM0 project “Optimal control of conservation equations”, itself supported by iCODE(IDEF Paris-Saclay) and the Hadamard Mathematics LabEx.

<sup>1</sup>Adrien Seguret is with PSL Research University, Université Paris-Dauphine, CEREMADE, Place de Lattre de Tassigny, 75016 Paris, France, and with Finance for Energy Market Research Centre (FIME), Paris, France, and with Osiris, EDF R&D, 91120 Palaiseau, France [adrien.seguret@edf.fr](mailto:adrien.seguret@edf.fr)

<sup>2</sup>Clémence Alasseur is with Osiris, EDF R&D, 91120 Palaiseau, France and with Finance for Energy Market Research Centre (FIME) [clemence.alasseur@edf.fr](mailto:clemence.alasseur@edf.fr)

<sup>3</sup>J. Frédéric Bonnans is with Disco Team, L2S, CentraleSupélec/Université Paris-Saclay and Inria-Saclay, France [Frederic.Bonnans@inria.fr](mailto:Frederic.Bonnans@inria.fr)

<sup>4</sup>Antonio De Paola is with Department of Electrical and Electronic Engineering, Imperial College London, London, UK [antonio.de-paola09@imperial.ac.uk](mailto:antonio.de-paola09@imperial.ac.uk)

<sup>5</sup>Nadia Oudjane is with Osiris, EDF R&D, 91120 Palaiseau, France and with Finance for Energy Market Research Centre (FIME), Paris, France [nadia.oudjane@edf.fr](mailto:nadia.oudjane@edf.fr)

<sup>6</sup>Vincenzo Trovato is with Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy and with Department of Electrical and Electronic Engineering, Imperial College London, London, UK [vincenzo.trovato@unitn.it](mailto:vincenzo.trovato@unitn.it)

The cost function is the sum of a *coupling term*  $F_0 : \mathbb{U} \rightarrow \mathbb{R}$ , function of the *aggregate strategies*  $\frac{1}{n} \sum_{i=1}^n u^i$ , and *local terms* functions of the local decision  $u^i$  and local noise  $\omega^i$  with  $G_i : L^2(\Omega^i, \mathbb{U}) \times \Omega^i \rightarrow \mathbb{R}$ . This framework aims at containing stochastic optimal control problems, where the states of the agents are driven by independent noises (see equations (5.5) and (5.2) developed in Section 5).

## 1.1 Motivations

This work is motivated by its potential applications to distributed coordination of large populations of small agents, with relevant real-world implications in different sectors, from communication networks to power systems. The application developed in this paper deals with the coordination of flexible electrical appliances, to support power system operation in a context of increasing penetration of renewables. Among other appliances, thermostatically controlled loads (e.g. refrigerators, air conditioners etc.) have been investigated in the last few years, for their intrinsic flexibility and potential for network support. Several papers have already assessed the potential of demand-side response actions for frequency response services of TCLs [28] and how the population recovers from significant perturbations [7]. The coordination of TCLs can be performed in a centralized way, like in [12]. However, this approach raises concerns with respect to the communication requirements and customer privacy. A common objective can be reached in a fully distributed approach, like in [32], where each TCL is able to calculate its own actions (ON/OFF switching) to pursue a common objective. This paper is related to the work of De Paola *et al.* [8], where each agent represents a flexible TCL device. In [8] a distributed solution is presented for the operation of a population of  $n = 2 \times 10^7$  refrigerators providing frequency support and load shifting. They adopt a game-theory framework, modelling the TCLs as price-responsive rational agents that schedule their energy consumption and allocate their frequency response provision in order to minimize their operational costs. The potential practical application of our work also considers a large population of TCLS which, as extension to [8], have stochastic dynamics. The proposed approach is able to minimize the overall system costs in a distributed way, with each TCL determining its optimal power consumption profile in response to price signals.

## 1.2 Related literature

The considered problem belongs to the class of stochastic control: looking for strategies minimizing the expectation of an objective function under specific constraints. One of the main approaches proposed in the literature to tackle this problem is to use random trees: this consists in replacing the almost sure constraints, induced by non-anticipativity, by a finite number of constraints, in order to get a finite set of scenarios (see [13] and [26]). Once the tree structure is built, the problem is solved by different decomposition methods such as scenario decomposition [25] or dynamic splitting [27]. The main objective of the scenario method is to reduce the problem to an approximated deterministic one. The present paper focuses on high dimensional noise problems with a large number of time steps, for which this approach is not feasible.

The idea of reducing a single high-dimensional problem to a large number of smaller problems with lower dimension has been widely studied in the deterministic case. In deterministic and stochastic problems there is the possibility of using time decomposition thanks to the Dynamic Programming Principle [2], taking advantage of the Markov property of the system. However, this method requires a specific time structure of the cost function and is not suitable for problems with high-dimensional state spaces. Under continuous linear-convex assumptions, one can deal with the curse of dimensionality by using the Stochastic Dual Dynamic Programming algorithm (SDDP) [22] to get upper and lower bounds of the value function, using polyhedral approximations. The almost-sure convergence of a broad class of SDDP algorithms has been proven [24], and complexity of the algorithm can be estimated, in the specific case of Lipschitz continuous value function [34] or by using a regularization of the value functions [33]. In [19, 21], a stopping criteria based on a dual version of SDDP, which gives a deterministic upper-bound for the primal problem, is proposed. SDDP is well-adapted for medium sized population problems ( $n \leq 30$ ), whereas it fails for problems with large populations ( $n > 1000$ ) such as the ones considered in this paper. To tackle this type of high dimensional problems, it is natural to investigate decomposition techniques in the spirit of the Dual Approximation Dynamic Programming (DADP) [11, 18]. This approach is characterized by a price decomposition of the problem, where the stochastic constraints are projected on subspaces such that the associated Lagrangian multiplier is adapted for dynamic programming. Then the optimal multiplier is estimated by implementing Uzawa's algorithm. To this end in [18], the Uzawa's algorithm, formulated in a Hilbert setting, is extended to a Banach space. DADP has been applied in different cases, such as storage management problem for electrical production in [1, 11] and hydro valley management [5]. The idea of approaching the primal and dual problems by restricting or relaxing

the set of decision variables has also been proposed in the context of stochastic programming [4, 16] to provide upper and lower bounds for the considered problem. In the proposed paper, in the same vein as DADP, we propose a price decomposition approach restricted to deterministic prices. This new approach takes advantage of the large population number in order to introduce an auxiliary problem where the coupling term is purely deterministic.

### 1.3 Contributions

The numerical difficulty of Problem  $(P_1)$  is related to the randomness of the aggregate term  $\frac{1}{n} \sum_{i=1}^n u^i$  involved in the coupling function  $F_0$ . Let us introduce the set of decentralized controls:

$$\hat{\mathcal{U}} := \prod_{i=1}^n \hat{\mathcal{U}}_i, \text{ where } \hat{\mathcal{U}}_i := \{u^i \in \mathcal{U}_i \mid u^i \text{ is } \mathcal{T}^1 \otimes \dots \otimes \mathcal{F}^i \otimes \mathcal{T}^{i+1} \otimes \dots \otimes \mathcal{T}^n \text{ - measurable}\}, \quad (1.2)$$

where  $\mathcal{T}^j$  is the trivial  $\sigma$ -field  $\{\emptyset, \Omega^j\}$ . Note that by construction, we can identify  $\hat{\mathcal{U}}_i$  with  $\mathcal{F}^i$ -measurable functions defined on  $\Omega^i$ . In addition, two decentralized controls  $u^i \in \hat{\mathcal{U}}_i$  and  $u^j \in \hat{\mathcal{U}}_j$ ,  $i \neq j$ , are independent random variables. If the control  $u$  is decentralized, since the variance of the sum of independent random variables is equal to the sum of the variances of these random variables, the aggregate term  $\frac{1}{n} \sum_{i=1}^n u^i$  can be approximated by  $\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i$  when  $n$  is large enough. Let us consider then the following approximation of Problem  $(P_1)$ :

$$(P_2) \quad \begin{cases} \min_{u \in \mathcal{U}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i \right) + \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n G_i(u^i(\cdot, \omega^{-i}), \omega^i) \right]. \end{cases} \quad (1.3)$$

A first step consists in showing that, without loss of optimality in Problem  $(P_2)$ , one can restrict the control set  $\mathcal{U}$  to  $\hat{\mathcal{U}}$ .

Theorem 2.1 states the equivalence between Problem  $(P_2)$  and its decentralized version  $(\hat{P}_2)$  defined by:

$$(\hat{P}_2) \quad \begin{cases} \min_{u \in \hat{\mathcal{U}}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i \right) + \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n G_i(u^i, \omega^i) \right]. \end{cases} \quad (1.4)$$

Through the article, the circumflex symbol  $\hat{\cdot}$  will be used to denote minimization problems w.r.t. decentralized controls. Problem  $(\hat{P}_2)$  can be written as:

$$(\hat{P}'_2) \quad \begin{cases} \min_{u \in \hat{\mathcal{U}}, v \in \mathbb{U}} \bar{J}(u, v), \\ \bar{J}(u, v) := F_0(v) + \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n G_i(u^i, \omega^i) \right], \\ \text{s.t. } g(u, v) = 0, \end{cases} \quad (1.5)$$

where  $g: \mathcal{U} \times \mathbb{U} \rightarrow \mathbb{U}$  is defined by

$$g(u, v) := \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i - v. \quad (1.6)$$

Observe that, for any  $u^i \in \hat{\mathcal{U}}_i$ ,  $G_i(u^i, \cdot)$  is independent of  $\mathcal{F}^{-i}$ . As a first contribution, this paper shows that under some convexity and regularity assumptions on  $F_0$  and  $(G_i)_{i \in \{1, \dots, n\}}$ , any solution of Problem  $(P_2)$  is an  $\varepsilon_n$ -solution of  $(P_1)$ , with  $\varepsilon_n \rightarrow 0$  when  $n \rightarrow \infty$ . In addition, we will see that an approach of price decomposition for  $(\hat{P}_2)$ , based on the formulation  $(\hat{P}'_2)$ , is tractable for dynamical problems, since the problem of minimizing the Lagrangian with deterministic dual variables can be decomposed in subproblems which are solvable by Dynamic Programming.

Since computing the dual cost of  $(\hat{P}'_2)$  is expensive, we propose *Stochastic Uzawa* and *Sampled Stochastic Uzawa* algorithms relying on the Robbins-Monroe algorithm, in the spirit of the stochastic gradient. Their convergence is established, relying on the proof provided by [10] for the convergence of the stochastic gradient in a Hilbert space. We check the effectiveness of the *Stochastic Uzawa* algorithm on a linear quadratic Gaussian framework, and we apply the *Sampled Stochastic Uzawa* algorithm to a model of power system, inspired by the work of A. De Paola *et al.* [8].

## 1.4 Assumptions

Various assumptions needed in the article are listed in this subsection.

- Assumption 1.** (i) Each set  $\mathcal{U}_i$  is bounded, i.e. there exists  $M > 0$  such that  $\mathbb{E}\|u^i\|_{\mathbb{U}}^2 \leq M^2$ , for  $i \in \{1, \dots, n\}$ .
- (ii) The function  $u^i \mapsto G_i(u^i(\cdot, \omega^{-i}), \omega^i)$  is a.s. non negative, convex and lower semi continuous (l.s.c. for short). And, for any  $u^i \in \mathcal{U}_i$ , the function  $\omega \mapsto G_i(u^i(\cdot, \omega^{-i}), \omega^i)$  is measurable.
- (iii) The function  $F_0$  is l.s.c. and proper.
- (iv) Problem  $(P_1)$  is feasible.

**Assumption 2.** The function  $F_0$  is convex.

**Assumption 3.** The function  $F_0$  is Gâteaux differentiable with  $c$ -Lipschitz derivative.

**Assumption 4.** (i) The function  $u^i \mapsto G_i(u^i, \omega^i)$  is for a.a.  $\omega^i \in \Omega^i$  strictly convex on  $\hat{\mathcal{U}}_i$ .

(ii) The function  $F_0$  has at least quadratic growth, i.e. there exist  $C_1, C_2 > 0$  such that for any  $v \in \mathbb{U}$ :

$$C_1\|v\|_{\mathbb{U}}^2 - C_2 \leq F_0(v).$$

**Remark 1.1.** By Lemma A.1 in Appendix A, if  $F_0$  satisfies Assumption 3 then,  $F_0$  has at most quadratic growth, i.e. there exists  $C > 0$  such that for any  $v \in \mathbb{U}$  one has:

$$F_0(v) \leq C(\|v\|_{\mathbb{U}}^2 + 1).$$

We denote by  $\{\rho_k\}_{k \in \mathbb{N}^*}$  the sequence of step sizes used in the *Stochastic Uzawa* and *Sampled Stochastic Uzawa* algorithms in Section 4.

**Assumption 5.** The sequence  $\{\rho_k\}_{k \in \mathbb{N}^*}$  is such that:  $\rho_k > 0$ ,  $\sum_{k=1}^{\infty} \rho_k = \infty$  and  $\sum_{k=1}^{\infty} (\rho_k)^2 < \infty$ .

Note that a sequence of the form  $\rho_k := \frac{a}{b+k}$ , with  $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}_+$ , satisfies Assumption 5.

**Assumption 6.** (i)  $F_0$  is strongly convex.

(ii) There exists  $\delta > 0$  such that for any  $i \in \{1, \dots, n\}$  and for a.a.  $\omega^i \in \Omega^i$ , the function  $\hat{\mathcal{U}}_i \ni u^i \mapsto G_i(u^i, \omega^i)$  is strongly convex with modulus of convexity greater or equal to  $\delta$ .

Obviously, Assumption 6 is stronger than Assumption 4.

## 2 Approximating the optimization problem

In this section, the link between the values of problems  $(P_1)$  and  $(P_2)$  is analyzed.

**Lemma 2.1.** Let Assumptions 1 and 2 hold. Then Problem  $(P_1)$  has a solution, i.e.  $J$  reaches its minimum over  $\mathcal{U}$ .

*Proof.* The existence of a minimum is proved by considering a minimizing sequence (which exists since  $(P_1)$  is feasible)  $\{u_k\}$  of  $J$  over  $\mathcal{U}$ . The set  $\mathcal{U}$  being bounded and weakly close, there exists a subsequence  $\{u_{k_\ell}\}$  which weakly converges to a certain  $u^* \in \mathcal{U}$ . Using Assumptions 1.(ii) and convexity of  $F_0$ , it follows that  $\liminf J(u_{k_\ell}) \geq J(u^*)$  and thus  $u^*$  is a solution of  $(P_1)$ .  $\square$

We obtain the following Corollary about  $(P_2)$ .

**Corollary 2.1.** If Assumptions 1 and 2 are satisfied, then Problem  $(P_2)$  has a solution and its value is lower or equal to the value of Problem  $(P_1)$  i.e.:

$$\inf_{u \in \mathcal{U}} \tilde{J}(u) \leq \inf_{u \in \mathcal{U}} J(u).$$

*Proof.* Assumption 1.(iv) and convexity of  $F_0$  imply that  $(P_2)$  is feasible. By using the same techniques as in the proof of Lemma 2.1, one can prove that  $(P_2)$  admits a solution. Using the convexity of  $F_0$  and Jensen's inequality, one has for any centralized control  $u \in \mathcal{U}$ :

$$F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i\right) \leq \mathbb{E}[F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right)],$$

and the conclusion follows from the definition of  $(P_1)$  in (1.1) and  $(P_2)$  in (1.3)  $\square$

We have the following key result.

**Theorem 2.1.** *If Assumption 1 is satisfied, then the decentralized Problem  $(\hat{P}_2)$  has the same value as the centralized Problem  $(P_2)$  i.e.:*

$$\inf_{u \in \hat{\mathcal{U}}} \tilde{J}(u) = \inf_{u \in \mathcal{U}} \tilde{J}(u). \quad (2.1)$$

*Proof.* Since  $\hat{\mathcal{U}} \subset \mathcal{U}$ , it is immediate that  $\inf_{u \in \mathcal{U}} \tilde{J}(u) \leq \inf_{u \in \hat{\mathcal{U}}} \tilde{J}(u)$ .

Fix  $i \in \{1, \dots, n\}$ , using the definition of conditional expectation, we define  $\tilde{u}^i \in L^2(\Omega^i, \mathbb{U})$  for any  $u^i \in \mathcal{U}_i$  by:

$$\tilde{u}^i(\omega^i) := \mathbb{E}[u^i(\omega^i, \omega^{-i}) | \omega^i] = \int_{\Omega^{-i}} u^i(\omega^i, \omega^{-i}) d\mu^{-i}(\omega^{-i}) \quad \text{for any } \omega^i \in \Omega^i.$$

Since  $G_i$  is a.s. convex w.r.t. the first variable, Jensen's inequality gives:

$$G_i(\tilde{u}^i, \omega^i) \leq \int_{\Omega^{-i}} G_i(u^i(\cdot, \omega^{-i}), \omega^i) d\mu^{-i}(\omega^{-i}) = \mathbb{E}[G_i(u^i(\cdot, \omega^{-i}), \omega^i) | \omega^i] \quad \text{a.s.} \quad (2.2)$$

On the other hand  $(u^1, \dots, u^n) \mapsto F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i\right)$  is invariant when taking the conditional expectation, thus:

$$F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i\right) = F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \tilde{u}^i\right).$$

Taking the expectation of (2.2), we have  $\inf_{u \in \hat{\mathcal{U}}} \tilde{J}(u) \leq \inf_{u \in \mathcal{U}} \tilde{J}(u)$ , and the conclusion follows.  $\square$

**Remark 2.1.** *In the applications to stochastic control problems (in discrete and continuous time) we have the constraint of having progressively measurable control policies. Since the set of progressively measurable policies is closed and convex, this enters in the above framework. In particular, the decentralized policy  $\tilde{u}^i$  constructed in the above proof is progressively measurable.*

**Remark 2.2.** *By Theorem 2.1, for any  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal solution of Problem  $(P_2)$  that is a decentralized control.*

Before stating the next result, we need to introduce Problem  $(\hat{P}_1)$ , that corresponds to the minimization of  $J$ , defined in (1.1), over the set of decentralized controls  $\hat{\mathcal{U}}$ :

$$(\hat{P}_1) \left\{ \min_{u \in \hat{\mathcal{U}}} J(u). \right.$$

**Proposition 2.1.** *If Assumption 1 is satisfied and  $F_0$  is Lipschitz continuous with constant  $\gamma$ , then any solution of  $(\hat{P}_2)$  is an  $\varepsilon$ -optimal solution of  $(\hat{P}_1)$  and, conversely, any solution of  $(\hat{P}_1)$  is an  $\varepsilon$ -optimal solution of  $(\hat{P}_2)$ , with  $\varepsilon = 2\gamma M/\sqrt{n}$ .*

*Proof.* Since  $F_0$  is Lipschitz continuous with Lipschitz constant  $\gamma$ , it holds for any  $x, y \in \mathbb{U}$ :  $|F_0(x) - F_0(y)| \leq \gamma \|x - y\|_{\mathbb{U}}$ . We set for any  $u \in \mathcal{U}$ :

$$\mathbf{u}^i := u^i - \mathbb{E} u^i. \quad (2.3)$$

Using the Lipschitz continuity of  $F_0$ , one has for any  $u \in \hat{\mathcal{U}}$ :

$$\begin{aligned} \left| \mathbb{E} \left[ F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i\right) \right] \right| &\leq \mathbb{E} \left| F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i\right) \right| \\ &\leq \frac{\gamma}{n} \mathbb{E} \left\| \sum_{i=1}^n \mathbf{u}^i \right\|_{\mathbb{U}}. \end{aligned}$$

Using the Jensen's inequality, for any  $u \in \hat{\mathcal{U}}$ , the mutual independence of the centered variables  $\mathbf{u}_i$  and  $\mathbf{u}_j$  for any  $j \neq i$  and  $\mathbb{E}\|\mathbf{u}_i\|_{\mathbb{U}}^2 \leq M^2$ , we get:

$$\frac{\gamma}{n} \mathbb{E} \left\| \sum_{i=1}^n \mathbf{u}^i \right\|_{\mathbb{U}} \leq \frac{\gamma}{n} \mathbb{E} \left[ \left\| \sum_{i=1}^n \mathbf{u}^i \right\|_{\mathbb{U}}^2 \right]^{\frac{1}{2}} \leq \frac{\gamma}{n^{\frac{1}{2}}} M. \quad (2.4)$$

Let  $\hat{u}$  denote a minimizer of  $(\hat{P}_2)$ , then using (2.4) for the first and last inequality, for any  $u \in \hat{\mathcal{U}}$  it holds:

$$J(\hat{u}) \leq \tilde{J}(\hat{u}) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq \tilde{J}(u) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq J(u) + \frac{2\gamma}{n^{\frac{1}{2}}} M. \quad (2.5)$$

Similarly, if  $u^*$  is a solution of  $(\hat{P}_1)$ , then for any  $u \in \hat{\mathcal{U}}$  one has:

$$\tilde{J}(u^*) \leq J(u^*) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq J(u) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq \tilde{J}(u) + \frac{2\gamma}{n^{\frac{1}{2}}} M. \quad (2.6)$$

□

**Theorem 2.2.** *Let Assumptions 1, 2 and 3 be satisfied. Then any solution of Problem  $(\hat{P}_2)$  is an  $\varepsilon$ -optimal solution (where  $\varepsilon = cM^2/n$ ) of Problem  $(P_1)$ .*

*Proof.* From Corollary 2.1 and Theorem 2.1, one has for any  $\hat{u} \in \hat{\mathcal{U}}$  solution of  $(\hat{P}_2)$  that:

$$\tilde{J}(\hat{u}) \leq \inf_{u \in \mathcal{U}} J(u). \quad (2.7)$$

Since  $F_0$  is convex, differentiable, with a  $c$ -Lipschitz derivative, one can derive a.s.:

$$\begin{aligned} & F_0\left(\frac{1}{n} \sum_{i=1}^n \hat{u}^i\right) - F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \hat{u}^i\right) \\ & \leq \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \hat{u}^i\right), \sum_{i=1}^n \mathbf{u}^i \rangle_{\mathbb{U}} \\ & = \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \hat{u}^i\right) - \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \hat{u}^i\right), \sum_{i=1}^n \mathbf{u}^i \rangle_{\mathbb{U}} + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \hat{u}^i\right), \sum_{i=1}^n \mathbf{u}^i \rangle_{\mathbb{U}} \\ & \leq \frac{c}{n^2} \left\| \sum_{i=1}^n \mathbf{u}^i \right\|_{\mathbb{U}}^2 + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \hat{u}^i\right), \sum_{i=1}^n \mathbf{u}^i \rangle_{\mathbb{U}}, \end{aligned} \quad (2.8)$$

where  $\mathbf{u}^i$  is defined from  $\hat{u}$  as in (2.3). From the definition of  $\mathbf{u}^i$ , one obtains:

$$\mathbb{E} \left[ \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E} \hat{u}^i\right), \sum_{i=1}^n \mathbf{u}^i \rangle_{\mathbb{U}} \right] = 0.$$

Since  $\hat{u} \in \hat{\mathcal{U}}$ , controls are mutually independent and bounded a.s. by  $M$ , one gets as in (2.4):

$$\frac{c}{n^2} \mathbb{E} \left\| \sum_{i=1}^n \mathbf{u}^i \right\|_{\mathbb{U}}^2 \leq \frac{c}{n} M^2. \quad (2.9)$$

Taking the expectation of the first and last terms of (2.8) and then incorporating (2.9) and the equality above, one obtains:

$$J(\hat{u}) - \tilde{J}(\hat{u}) \leq \frac{c}{n} M^2.$$

From previous inequality and (2.7), we get:

$$J(\hat{u}) \leq \inf_{u \in \mathcal{U}} J(u) + \frac{c}{n} M^2. \quad (2.10)$$

□

**Remark 2.3.** Observe that the value of the centralized Problem  $(P_1)$  on the l.h.s. of the inequality (2.11) below is upper bounded by the following decentralized problem on the r.h.s of this inequality i.e.

$$\inf_{u \in \mathcal{U}} J(u) \leq \inf_{u \in \hat{\mathcal{U}}} J(u). \quad (2.11)$$

Ref. [6] obtains an upper bound for the decentralized problem and a lower bound for the centralized problem. The upper bound is provided by a resource decomposition approach (with deterministic quantities) while the lower bound is obtained by a price decomposition approach with deterministic prices (see Equation (28) of [6]). Theorem 2.2 provides an upper bound for Problem  $(P_1)$  with an a priori quantification of the deviation from the optimal value which vanishes when the number of agents grows to infinity. Moreover, in Section 4 we provide an original algorithm that allows to approach the solution of the decentralized problem.

**Remark 2.4.** Let  $\hat{u}$  and  $u^*$  be respectively the optimal solutions of problems  $(\hat{P}_2)$  and  $(P_1)$ . From Jensen's inequality and by definition of  $\hat{u}$  we have:

$$-J(u^*) \leq -\tilde{J}(u^*) \leq -\tilde{J}(\hat{u}).$$

Adding  $J(\hat{u})$ , one has:

$$0 \leq J(\hat{u}) - J(u^*) \leq J(\hat{u}) - \tilde{J}(u^*) \leq J(\hat{u}) - \tilde{J}(\hat{u}). \quad (2.12)$$

Inequality (2.12) allows to compute an upper bound of the "optimality" error  $J(\hat{u}) - J(u^*)$ , by evaluating  $J(\hat{u}) - \tilde{J}(\hat{u})$ .

### 3 Dualization and Decentralization of Problem $(P_2)$

The Lagrangian function associated to the constrained optimization Problem  $(\hat{P}'_2)$ , defined in (1.5), is:  $L : \hat{\mathcal{U}} \times \mathbb{U} \times \mathbb{U} \rightarrow \mathbb{R}$  defined by:

$$L(u, v, \lambda) := \bar{J}(u, v) + \langle \lambda, \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i - v \rangle_{\mathbb{U}}. \quad (3.1)$$

The dual Problem  $(D)$  associated with  $(\hat{P}'_2)$  is:

$$(D) \quad \max_{\lambda \in \mathbb{U}} \mathcal{W}(\lambda), \quad \text{where } \mathcal{W}(\lambda) := \min_{u \in \hat{\mathcal{U}}, v \in \mathbb{U}} L(u, v, \lambda). \quad (3.2)$$

For any  $\lambda \in \mathbb{U}$ , it holds:

$$\mathcal{W}(\lambda) = -F_0^*(\lambda) + \frac{1}{n} \sum_{i=1}^n \min_{u^i \in \hat{\mathcal{U}}_i} \mathbb{E}[G_i(u^i, \omega^i)] + \langle \lambda, \mathbb{E} u^i \rangle_{\mathbb{U}}, \quad (3.3)$$

where, for any real valued function  $F$  defined on  $\mathbb{U}$ ,  $F^*$  stands for its Fenchel conjugate, defined for  $x \in \mathbb{U}$  by  $F^*(x) := \sup_{y \in \mathbb{U}} \langle x, y \rangle_{\mathbb{U}} - F(y)$ . The problem is said to be qualified if it is still feasible after a small perturbation of the constraint, in the following sense:

$$\text{There exists } \varepsilon > 0 \text{ such that } \mathcal{B}_{\mathbb{U}}(0, \varepsilon) \subset g(\hat{\mathcal{U}}, \mathbb{U}), \quad (3.4)$$

where  $\mathcal{B}_{\mathbb{U}}(0, \varepsilon)$  is the open ball of radius  $\varepsilon$  in  $\mathbb{U}$ ,  $g$  has been defined in (1.6) and  $g(\hat{\mathcal{U}}, \mathbb{U})$  is the image by  $g$  of  $\hat{\mathcal{U}} \times \mathbb{U}$ .

**Lemma 3.1.** *If Assumption 1 holds, then Problem  $(\hat{P}'_2)$  is qualified. If Assumption 2 is also satisfied, then problems  $(\hat{P}'_2)$  and  $(D)$  have the same value, the set of dual solutions  $S$  is nonempty and bounded and any primal solution  $\hat{u}$  satisfies both  $W(\hat{\lambda}) = \tilde{J}(\hat{u})$  and  $(\hat{u}, \hat{v}) \in \arg \min_{u \in \hat{\mathcal{U}}, v \in \mathbb{U}} L(\hat{\lambda}, u, v)$ , with  $\hat{\lambda} \in S$ .*

*Proof.* By Assumption 1.(iv), there exists  $\tilde{u}$  feasible for Problem  $(P_1)$ . Then using the definition of  $g$  in (1.6)

$$\mathcal{B}_{\mathbb{U}}(0, \varepsilon) \subset \mathbb{U} = g(\tilde{u}, \mathbb{U}) \subset g(\hat{\mathcal{U}}, \mathbb{U}). \quad (3.5)$$

The qualification of  $(\hat{P}'_2)$  follows. The conclusion follows by [3, Theorem 2.165].  $\square$

Since the set of admissible controls  $\hat{\mathcal{U}} = \hat{\mathcal{U}}_1 \times \dots \times \hat{\mathcal{U}}_n$  is a Cartesian product, if  $G_i$  is strictly convex with respect to its first variable, then each component  $\hat{u}^i$  of the solution  $\hat{u}$  of Problem  $(\hat{P}'_2)$ , can be uniquely determined by solving the following subproblem:

$$\hat{u}^i = \arg \min_{u^i \in \hat{\mathcal{U}}_i} \left\{ \mathbb{E} \left[ G_i(u^i, \omega^i) + \langle \hat{\lambda}, u^i \rangle_{\mathbb{U}} \right] \right\},$$

where  $\hat{\lambda} \in S$ .



**Remark 3.1.** By using the same argument as in Theorem 2.1, one can prove, for any  $\lambda \in \mathbb{U}$ :

$$\min_{u^i \in \mathcal{U}_i} \{ \mathbb{E} [G_i(u^i, \omega^i) + \langle \lambda, u^i \rangle_{\mathbb{U}}] \} = \min_{u^i \in \mathcal{U}_i} \{ \mathbb{E} [G_i(u^i(\cdot, \omega^{-i}), \omega^i) + \langle \lambda, u^i \rangle_{\mathbb{U}}] \}. \quad (3.6)$$

## 4 Stochastic Uzawa and Sampled Stochastic Uzawa algorithms

This section aims at proposing an algorithm to find a solution of the dual problem (3.2).

### 4.1 Preliminary results

Though the below result is well-known and can be found in [14] for functions defined on finite vector spaces, we adapt the proof to the Hilbert space setting for the sake of completeness.

**Lemma 4.1.** *If Assumption 1 and 2 are satisfied, then Assumption 3 holds iff  $F_0^*$  is strongly convex.*

*Proof.* (i) Let Assumption 3 holds. Since  $F_0$  is proper, convex and l.s.c.,  $F_0^*$  is l.s.c. proper. From the Lipschitz property of the gradient of  $F_0$ , it holds that  $\text{dom}(F_0) = \mathbb{U}$ .

Let  $s, \tilde{s} \in \text{dom}(F_0^*)$  such that there exist  $\lambda_s \in \partial F_0^*(s)$  and  $\lambda_{\tilde{s}} \in \partial F_0^*(\tilde{s})$ . From the differentiability, l.s.c. and convexity of  $F_0$ , it follows that:  $s = \nabla F_0(\lambda_s)$  and  $\tilde{s} = \nabla F_0(\lambda_{\tilde{s}})$ . By Assumption 3 and the extended Baillon-Haddad theorem [23, Theorem 3.1],  $\nabla F_0$  is cocoercive. In other words:

$$\begin{aligned} \langle s - \tilde{s}, \lambda_s - \lambda_{\tilde{s}} \rangle_{\mathbb{U}} &= \langle \nabla F_0(\lambda_s) - \nabla F_0(\lambda_{\tilde{s}}), \lambda_s - \lambda_{\tilde{s}} \rangle_{\mathbb{U}} \\ &\geq \frac{1}{c} \|\nabla F_0(\lambda_s) - \nabla F_0(\lambda_{\tilde{s}})\|_{\mathbb{U}}^2 \\ &= \frac{1}{c} \|s - \tilde{s}\|_{\mathbb{U}}^2, \end{aligned} \quad (4.1)$$

where  $c$  is the Lipschitz constant of  $\nabla F_0$  defined in Assumption 3. Therefore  $\partial F_0^*$  is strongly monotone, which implies the strong convexity of  $F_0^*$ .

(ii) Conversely, assume that  $F_0^*$  is proper and strongly convex. Then there exist  $\alpha, \beta > 0$  and  $\gamma \in \mathbb{U}$  such that for any  $s \in \text{dom}(F_0^*)$ :  $F_0^*(s) \geq \alpha \|s\|_{\mathbb{U}}^2 + \langle \gamma, s \rangle_{\mathbb{U}} - \beta$ , and  $F_0$  being convex, l.s.c. and proper, for any  $\lambda \in \mathbb{U}$  it holds:

$$F_0(\lambda) \leq \sup_{s \in \mathbb{U}} \langle s, \lambda - \gamma \rangle_{\mathbb{U}} - \alpha \|s\|_{\mathbb{U}}^2 + \beta = \|\lambda - \gamma\|^2 / (4\alpha) + \beta. \quad (4.2)$$

Thus,  $F_0$  is proper and uniformly upper bounded over bounded sets and therefore is locally Lipschitz. In addition, from the strong convexity of  $F_0^*$  and the convexity of  $F_0$ , for any  $\lambda \in \mathbb{U}$ ,  $\partial F_0(\lambda)$  is a singleton. Thus  $F_0$  is everywhere Gâteaux differentiable.

Let  $\lambda, \mu \in \mathbb{U}$ . Since  $F_0^*$  is strongly convex, the functions  $F_0^*(s) - \langle \lambda, s \rangle_{\mathbb{U}}$  (resp.  $F_0^*(s) - \langle \mu, s \rangle_{\mathbb{U}}$ ) has a unique minimum point  $s_\lambda$  (resp.  $s_\mu$ ), characterized by:  $\lambda \in \partial F_0^*(s_\lambda)$  and  $\mu \in \partial F_0^*(s_\mu)$ . From the strong convexity of  $F_0^*$ , the strong monotonicity of  $\partial F_0^*$  holds:  $\langle \mu - \lambda, s_\mu - s_\lambda \rangle_{\mathbb{U}} \geq \frac{1}{c} \|s_\mu - s_\lambda\|_{\mathbb{U}}^2$ , where  $c > 0$  is a constant related to the strong convexity of  $F_0^*$ . Using that  $s_\lambda = \nabla F_0(\lambda)$  and  $s_\mu = \nabla F_0(\mu)$ , it holds:

$$\langle \mu - \lambda, \nabla F_0(\mu) - \nabla F_0(\lambda) \rangle_{L^2(0,T)} \geq \frac{1}{c} \|\nabla F_0(\mu) - \nabla F_0(\lambda)\|_{L^2(0,T)}^2, \quad (4.3)$$

meaning that  $\nabla F_0$  is cocoercive. Applying the Cauchy-Schwarz inequality to the left hand side of the previous inequality, the Lipschitz property of  $\nabla F_0$  follows.  $\square$

**Lemma 4.2.** *If Assumptions 1, 2 and 3 hold, then  $\mathcal{W}$  is strongly concave.*

*Proof.* For any  $\lambda \in \mathbb{U}$ , the expression of  $\mathcal{W}(\lambda)$  is given by (3.3), where for any  $i \in \{1, \dots, n\}$ ,  $\lambda \mapsto \inf_{u^i \in \mathcal{U}_i} \mathbb{E} G_i(u^i, \omega^i) + \langle \lambda, \mathbb{E} u^i \rangle_{\mathbb{U}}$  is concave and, from Lemma 4.1,  $-F_0^*$  is strongly concave. Since the sum of a concave function and of a strongly concave function is strongly concave, the result follows.  $\square$

### 4.2 Analysis of the algorithms

Assumptions 1, 2 and 4 are supposed to hold throughout Section 4.2. For all  $i \in \{1, \dots, n\}$ , and  $\lambda \in \mathbb{U}$ , we define the optimal control  $u^i(\lambda)$ :

$$u^i(\lambda) := \arg \min_{u^i \in \mathcal{U}_i} \{ \mathbb{E} [G_i(u^i, \omega^i) + \langle \lambda, u^i \rangle_{\mathbb{U}}] \}, \quad (4.4)$$

which is well defined since  $u^i \rightarrow \mathbb{E} G_i(u^i, \omega^i)$  is strictly convex.

For any  $\lambda \in \mathbb{U}$ , the subset  $V(\lambda)$  is defined by:

$$V(\lambda) := \arg \min_{v \in \mathbb{U}} \{F_0(v) - \langle \lambda, v \rangle_{\mathbb{U}}\}. \quad (4.5)$$

Since  $F_0$  is convex and has at least quadratic growth,  $V(\lambda)$  is a non empty subset of  $\mathbb{U}$  and is reduced to a singleton if  $F_0$  is strictly convex. For any  $\lambda \in \mathbb{U}$ , we denote by  $v(\lambda)$  an element of  $V(\lambda)$ , and for any  $v(\lambda) \in V(\lambda)$ , one has  $v(\lambda) \in \partial F_0^*(\lambda)$ .

Uzawa's algorithm seems particularly fitting for this problem. However, at each dual iteration  $k$  and any  $i \in \{1, \dots, n\}$ , one would have to compute the quantities  $\mathbb{E}[u^i(\lambda^k)]$  for the update of  $\lambda^{k+1}$ , which is hard in practice. Therefore two algorithms are proposed where, at each iteration  $k$ ,  $\lambda^{k+1}$  is updated thanks to a realization of  $u^i(\lambda^k)$ .

We introduce the function  $f : \mathbb{U} \rightarrow \mathbb{U}$  defined by:

$$f(\lambda) := g(u(\lambda), v(\lambda)) = \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i(\lambda) - v(\lambda). \quad (4.6)$$

By Assumption 4.(ii),  $F_0$  has at least quadratic growth, whence  $F_0^*$  has at most quadratic growth. Indeed, using the definition of the Fenchel's conjugate, the fact that  $F_0$  has at least quadratic growth, and the Cauchy-Schwarz inequality, there exist  $\bar{C}_1, \bar{C}_2 > 0$  such that for any  $\lambda \in \mathbb{U}$ :

$$F_0^*(\lambda) \leq \sup_{\mu \in \mathbb{U}} \|\mu\|_{\mathbb{U}} \|\lambda\|_{\mathbb{U}} - \bar{C}_1 \|\mu\|_{\mathbb{U}}^2 + \bar{C}_2 = \frac{\|\lambda\|_{\mathbb{U}}^2}{2\bar{C}_1} + \bar{C}_2.$$

Then using Lemma A.1 in Appendix A and that  $v(\lambda) \in \partial F_0^*(\lambda)$ , there exists  $C > 0$  such that for any  $\lambda \in \mathbb{U}$ :

$$\|v(\lambda)\|_{\mathbb{U}} \leq C(\|\lambda\|_{\mathbb{U}} + 1). \quad (4.7)$$

Using the definition of  $\mathcal{U}$ , one has  $\frac{1}{n} \sum_{i=1}^n \|\mathbb{E} u^i(\lambda)\|_{\mathbb{U}} \leq M$  for any  $\lambda \in \mathbb{U}$ . Therefore, from the definition of  $f$  in (4.6), there exist  $M_1, M_2 > 0$  such that for any  $\lambda \in \mathbb{U}$  one has:

$$\|f(\lambda)\|_{\mathbb{U}}^2 \leq M_1 + M_2 \|\lambda\|_{\mathbb{U}}^2. \quad (4.8)$$

For any  $\lambda \in \mathbb{U}$ , we denote by  $\partial(-\mathcal{W}(\lambda))$  the subgradient of  $-\mathcal{W}$  at  $\lambda$ . Therefore, for any  $\lambda \in \mathbb{U}$ :

$$\partial(-\mathcal{W}(\lambda)) \ni -f(\lambda). \quad (4.9)$$

The iterative algorithm, proposed as an approximation scheme for  $\lambda^* \in \arg \max_{\lambda} \mathcal{W}(\lambda)$ , is summarized in the *Stochastic Uzawa Algorithm 1*.

---

**Algorithm 1** Stochastic Uzawa

---

- 1: Initialization  $\lambda^0 \in \mathbb{U}$ , set  $\{\rho_k\}_{k \in \mathbb{N}^*}$  satisfying Assumption 5.
  - 2:  $k \leftarrow 0$ .
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:  $v^k \leftarrow v(\lambda^k)$  where  $v(\lambda^k) \in V(\lambda^k)$ , this set being defined in (4.5).
  - 5:  $u^{i,k} \leftarrow u^i(\lambda^k)$  where  $u^i(\lambda^k)$  is defined in (4.4) for any  $i \in \{1, \dots, n\}$ .
  - 6: Generate  $n$  independent noises  $(\omega^{1,k+1}, \dots, \omega^{n,k+1})$ , independent also of  $\{\omega^{i,p} : 1 \leq i \leq n, p \leq k\}$ .
  - 7: Compute the associated control realization  $(u^{1,k}(\omega^{1,k+1}), \dots, u^{n,k}(\omega^{n,k+1}))$ .
  - 8:  $Y^{k+1} \leftarrow \frac{1}{n} \sum_{i=1}^n u^{i,k}(\omega^{i,k+1}) - v^k$ .
  - 9:  $\lambda^{k+1} \leftarrow \lambda^k + \rho_k Y^{k+1}$ .
- 

At any dual iteration  $k$  of Algorithm 1,  $Y^{k+1}$  is an estimator of  $\mathbb{E}[\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i,k+1}) - v(\lambda^k)]$ . An alternative approach, proposed in the *Sampled Stochastic Uzawa Algorithm 2*, consists in performing less simulations

---

**Algorithm 2** Sampled Stochastic Uzawa

- 
- 1: Initialization of  $m$  a positive integer and  $\tilde{\lambda}^0 \in \mathbb{U}$ , set  $\{\rho_k\}_{k \in \mathbb{N}^*}$  satisfying Assumption 5.
  - 2:  $k \leftarrow 0$ .
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:    $v^k \leftarrow v(\tilde{\lambda}^k)$  where  $v(\tilde{\lambda}^k) \in V(\tilde{\lambda}^k)$ , this set being defined in (4.5).
  - 5:   Generate  $m$  i.i.d. discrete random variables  $I_1^k, \dots, I_m^k$  uniformly in  $\{1, \dots, n\}$ .
  - 6:    $u^{I_j^k, k} \leftarrow u^{I_j^k}(\tilde{\lambda}^k)$  where  $u^{I_j^k}(\tilde{\lambda}^k)$  is defined in (4.4) for any  $j \in \{1, \dots, m\}$ .
  - 7:   Generate  $m$  independent noises  $(\omega^{1, k+1}, \dots, \omega^{m, k+1})$ , independent also of  $\{\omega^{i, p} : 1 \leq i \leq m, p \leq k\}$ .
  - 8:   Compute the associated control realization  $(u^{I_1^k, k}(\omega^{1, k+1}), \dots, u^{I_m^k, k}(\omega^{m, k+1}))$ .
  - 9:    $\tilde{Y}^{k+1} \leftarrow \frac{1}{m} \sum_{j=1}^m u^{I_j^k, k}(\omega^{I_j^k, k+1}) - v^k$
  - 10:    $\tilde{\lambda}^{k+1} \leftarrow \tilde{\lambda}^k + \rho_k \tilde{Y}^{k+1}$ .
- 

at each iteration, by taking  $m < n$ , at the risk of performing more dual iterations, to estimate the quantity  $\mathbb{E}[\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i, k+1}) - v(\lambda^k)]$ .

The complexity of the *Sampled Stochastic Uzawa* Algorithm 2 is proportional to  $m \times K$ , where  $K$  is the total number of dual iterations and  $m$  the number of simulations performed at each iteration. The error  $\mathbb{E} \|\lambda^{k+1} - \lambda^*\|_{\mathbb{U}}^2$  for  $\lambda^* \in S$  (we recall that  $S$  is defined by  $S := \arg \max_{\lambda \in \mathbb{U}} \mathcal{W}(\lambda)$  and from Lemma 3.1  $S$  is non empty) is the sum of the square of the bias (which only depends on  $K$  and not on  $m$ ) and the variance (which both depends on  $K$  and  $m$ ). Therefore, this algorithm enables a bias variance trade-off for a given complexity. Similarly, for a given error, it enables to optimize the complexity of the algorithm.

The following result establishes the convergence of the *Stochastic Uzawa* Algorithm 1:

**Lemma 4.3.** *Let Assumptions 1, 2, 4 and 5 hold and let  $\{\lambda_k\}_k$  be a sequence of multipliers generated by Algorithm 1. Then:*

- (i)  $\{\|\lambda^k - \lambda\|_{\mathbb{U}}^2\}$  converges a.s., for all  $\lambda \in S$ .
- (ii)  $\mathcal{W}(\lambda^k) \xrightarrow[k \rightarrow \infty]{} \max_{\lambda \in \mathbb{U}} \mathcal{W}(\lambda)$  a.s.
- (iii)  $\{\lambda^k\}$  weakly converges to some  $\bar{\lambda} \in S$  in  $\mathbb{U}$  a.s.
- (iv) If Assumption 3 holds, then a.s.  $\{\lambda^k\}$  converges to  $\bar{\lambda}$  in  $\mathbb{U}$ , with  $S := \{\bar{\lambda}\}$ .

The proof follows from [10, Theorem 3.6]. The cited reference (changing minimization in maximization) is interested in the maximization of a function  $\mathcal{W}$  of the specific form  $\mathcal{W}(\lambda) = \mathbb{E} \mathbf{W}(\lambda, \omega)$ , where  $\mathbf{W}(\cdot, \omega)$  is concave a.s. in  $\omega$ . However, in our setting we cannot in general exhibit such a representation for the dual function  $\mathcal{W}$ , defined in (3.2). Using the definition of  $u(\lambda)$  in (4.4) and  $v(\lambda)$  in (4.5), we have  $\mathcal{W}(\lambda) = \mathbb{E} \mathbf{W}(\lambda, \omega)$ , where

$$\lambda \mapsto \mathbf{W}(\lambda, \omega) := F_0(v(\lambda)) + \frac{1}{n} \sum_{i=1}^n G_i(u^i(\lambda), \omega^i) + \langle \lambda, \frac{1}{n} \sum_{i=1}^n u^i(\lambda) - v(\lambda) \rangle_{\mathbb{U}}$$

Note that  $\mathbf{W}(\cdot, \omega)$  is not a concave function of  $\lambda$  for a.a.  $\omega \in \Omega$ . Although our setting does not enter in the framework considered in [10], the proof of Lemma 4.3 follows from an obvious adaptation of the one in [10, Theorem 3.6]. It is enough to provide the first steps of the proof.

*Proof of Lemma 4.3.* First consider point (i). Let  $\lambda \in S$ . For any  $k$ ,  $\mathcal{G}_{k+1}$  is the filtration defined by:

$$\mathcal{G}_{k+1} := \sigma(\{\omega^{i, p}\} : 1 \leq i \leq n, p \leq k+1). \quad (4.10)$$

Using the definition of  $Y^{k+1} \in \mathbb{U}$  line 8 in the *Stochastic Uzawa* Algorithm 1, we have:

$$\begin{aligned} \|\lambda^{k+1} - \lambda\|_{\mathbb{U}}^2 &= \|\lambda^k + \rho_k Y^{k+1} - \lambda\|_{\mathbb{U}}^2 \\ &= \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \langle \lambda^k - \lambda, Y^{k+1} \rangle_{\mathbb{U}} + (\rho_k)^2 \|Y^{k+1}\|_{\mathbb{U}}^2. \end{aligned} \quad (4.11)$$

Since  $Y^{k+1}$  is independent from  $\mathcal{G}_k$ , it follows that:

$$\mathbb{E}[\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k] = \mathbb{E} \left\| \frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i,k+1}) - v(\lambda^k) \right\|_{\mathbb{U}}^2. \quad (4.12)$$

Using previous equality and the inequality (4.8), one can easily show that there exists  $M_3, M_4 > 0$  such that, for any  $k \in \mathbb{N}$ , one has:

$$\mathbb{E}[\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k] \leq M_1 + M_2 \|\lambda^k\|_{\mathbb{U}}^2 \leq M_3 + M_4 \|\lambda^k - \lambda\|_{\mathbb{U}}^2 \quad (4.13)$$

Since  $\lambda^k$  is  $\mathcal{G}_k$ -measurable and that  $\mathbb{E}[Y^{k+1} | \mathcal{G}_k] = f(\lambda^k)$ , we have that:

$$\begin{aligned} \mathbb{E}[\|\lambda^{k+1} - \lambda\|_{\mathbb{U}}^2 | \mathcal{G}_k] &= \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \mathbb{E}[\langle \lambda^k - \lambda, Y^{k+1} \rangle_{\mathbb{U}} | \mathcal{G}_k] + (\rho_k)^2 \mathbb{E}[\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k] \\ &\leq \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \langle \lambda^k - \lambda, f(\lambda^k) \rangle_{\mathbb{U}} + (\rho_k)^2 (M_3 + M_4 \|\lambda^k - \lambda\|_{\mathbb{U}}^2) \\ &\leq \|\lambda^k - \lambda\|_{\mathbb{U}}^2 (1 + M_4 \rho_k^2) + (\rho_k)^2 M_3 - 2\rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)). \end{aligned} \quad (4.14)$$

In the last inequality, we used the concavity of  $\mathcal{W}$  and (4.9). The rest of the proof follows [10, Theorem 3.6].  $\square$

Recalling the definition of  $\bar{J}(u, v)$  in (1.5) and of  $\bar{\lambda}$  in Lemma 4.3.(iii), we define  $\bar{u}$ :

$$\bar{u} := \arg \min_{u \in \hat{\mathcal{U}}} \left\{ \mathbb{E} \left[ \sum_{i=1}^n G_i(u^i, \omega^i) + \langle \bar{\lambda}, u^i \rangle_{\mathbb{U}} \right] \right\}. \quad (4.15)$$

Under Assumption 4,  $G_i$  is strictly convex w.r.t. the first variable, and then  $\bar{u}$  is well defined. If  $F_0$  is strictly convex, then  $V(\bar{\lambda})$  is a singleton and we can write:

$$\bar{v} := \arg \min_{v \in \mathbb{U}} \{F_0(v) + \langle \bar{\lambda}, v \rangle_{\mathbb{U}}\}. \quad (4.16)$$

**Remark 4.1.** *If  $F_0$  is convex, by Lemma 3.1, there is no duality gap associated to the Lagrangian  $L$  defined in (3.1). Further, if  $F_0$  is strictly convex, then  $(\bar{u}, \bar{v}, \bar{\lambda})$  is the unique saddle point associated to the Lagrangian  $L$ . Indeed, by Assumption 3,  $\bar{\lambda}$  is the unique solution of the dual problem (D), by Assumption 4.(i),  $\bar{u}$  is unique and by strict convexity of  $F_0$ ,  $\bar{v}$  is also the unique minimizer in the right hand side of (4.16).*

**Theorem 4.1.** *Let the Assumptions 1, 2, 3, 4 and 5 hold, then we have:*

- (i)  $\{u(\lambda^k)\}$  weakly converges a.s. to  $\bar{u}$ .

Furthermore, if  $F_0$  is strictly convex, then from Remark 4.1,  $\bar{u}$  is the unique minimizer of  $\tilde{J}$  in  $\hat{\mathcal{U}}$  and:

- (ii)  $\tilde{J}(u(\lambda^k)) \xrightarrow[k \rightarrow \infty]{} \tilde{J}(\bar{u})$  a.s.
- (iii)  $\limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \inf_{u \in \hat{\mathcal{U}}} J(u) + 2\varepsilon$  a.s. where  $\varepsilon = cM^2/n$ .

*Proof.* Proof of point (i). By Lemma 4.3.(iv), the sequence  $\{\lambda^k\}$  is bounded in  $\mathbb{U}$ . Thus, using inequality (4.7) one deduces that  $\{v(\lambda^k)\}$  is also bounded in  $\mathbb{U}$ . Since the sequence  $\{(u(\lambda^k), v(\lambda^k))\}$  is bounded in  $\hat{\mathcal{U}} \times \mathbb{U}$ , there exists a weakly convergent subsequence  $\{(u(\lambda^{\theta_k}), v(\lambda^{\theta_k}))\}$  such that:

$$(u(\lambda^{\theta_k}), v(\lambda^{\theta_k})) \rightharpoonup (u^\theta, v^\theta) \in \hat{\mathcal{U}} \times \mathbb{U}. \quad (4.17)$$

Using the definition of  $\lambda \mapsto u(\lambda)$  in (4.4), it holds for any  $k > 0$ :

$$\mathbb{E} [G_i(\bar{u}^i, \omega^i) + \langle \lambda^{\theta_k}, \bar{u}^i \rangle_{\mathbb{U}}] \geq \mathbb{E} [G_i(u^i(\lambda^{\theta_k}), \omega^i) + \langle \lambda^{\theta_k}, u^i(\lambda^{\theta_k}) \rangle_{\mathbb{U}}]. \quad (4.18)$$

Using that  $u^i \mapsto G_i(u^i, \omega^i)$  is a.s. w.l.s.c. on  $\hat{\mathcal{U}}_i$  and the a.s. convergence of  $\{\lambda^k\}$ , resulting from Lemma 4.3.(iv), we have from (4.18) when  $k \rightarrow \infty$ :

$$\mathbb{E} [G_i(\bar{u}^i, \omega^i) + \langle \bar{\lambda}, \bar{u}^i \rangle_{\mathbb{U}}] \geq \mathbb{E} [G_i(u^{i,\theta}, \omega^i) + \langle \bar{\lambda}, u^{i,\theta} \rangle_{\mathbb{U}}]. \quad (4.19)$$

Since  $\bar{u}$  is unique, it follows  $u^\theta = \bar{u}$  and (4.19) is an equality. Using that every weakly convergent subsequence of  $\{u(\lambda^k)\}$  has the same weak limit  $\bar{u}$ , (i) is deduced.

Proof of point (ii).

From point (i) and (4.19), it follows for any  $i \in \{1, \dots, n\}$ :

$$\lim_{k \rightarrow \infty} \mathbb{E} G_i(u^i(\lambda^k), \omega^i) = \mathbb{E} G_i(\bar{u}^i, \omega^i). \quad (4.20)$$

Using (4.17), the w.l.s.c. of  $F_0$ , equation (4.16), and applying the same previous argument to  $\{v(\lambda^k)\}$ , it holds that:

$$\lim_{k \rightarrow \infty} F_0(v(\lambda^k)) - \langle \lambda^k, v(\lambda^k) \rangle_{\mathbb{U}} = F_0(\bar{v}) - \langle \bar{\lambda}, \bar{v} \rangle_{\mathbb{U}}, \quad (4.21)$$

and  $v(\lambda^k) \xrightarrow[k \rightarrow \infty]{} \bar{v}$ .

From the two previous equalities and the a.s. convergence of  $\{\lambda^k\}$ , it follows:

$$\lim_{k \rightarrow \infty} F_0(v(\lambda^k)) = F_0(\bar{v}). \quad (4.22)$$

Using that  $(\bar{u}, \bar{v}, \bar{\lambda})$  is a saddle point, it follows:

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \bar{u}^i = \bar{v}. \quad (4.23)$$

From (4.22) and (4.23), it holds:

$$\lim_{k \rightarrow \infty} F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i(\lambda^k) \right) = F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} \bar{u}^i \right). \quad (4.24)$$

Then adding (4.20) and (4.24):  $\lim_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) = \tilde{J}(\bar{u})$ .

Proof of point (iii). From point (ii), inequality (2.10) and Theorem 2.2, it holds:

$$\limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \limsup_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) + \varepsilon = \inf_{u \in \mathcal{U}} \tilde{J}(u) + \varepsilon \leq \inf_{u \in \mathcal{U}} J(u) + 2\varepsilon. \quad (4.25)$$

□

**Lemma 4.4.** *Let Assumptions 1 and 6.(i) hold, then the function  $\lambda \mapsto v(\lambda)$  is Lipschitz on  $\mathbb{U}$ .*

*Proof.* From the definition of  $v$  in (4.5), we have for any  $\lambda \in \mathbb{U}$  that  $\lambda \in \partial F_0(v(\lambda))$ . Thus, for any  $\lambda, \mu \in \mathbb{U}$ , we have from the strong convexity of  $F_0$ :

$$\begin{cases} F_0(v(\mu)) & \geq F_0(v(\lambda)) + \langle \lambda, v(\mu) - v(\lambda) \rangle_{\mathbb{U}} + \alpha \|v(\mu) - v(\lambda)\|_{\mathbb{U}}^2 \\ F_0(v(\lambda)) & \geq F_0(v(\mu)) + \langle \mu, v(\lambda) - v(\mu) \rangle_{\mathbb{U}} + \alpha \|v(\lambda) - v(\mu)\|_{\mathbb{U}}^2. \end{cases} \quad (4.26)$$

Adding the two previous inequalities, after simplifications, we get:

$$\langle \lambda - \mu, v(\lambda) - v(\mu) \rangle_{\mathbb{U}} \geq 2\alpha \|v(\lambda) - v(\mu)\|_{\mathbb{U}}^2. \quad (4.27)$$

Applying the Cauchy-Schwarz inequality and simplifying by  $\|v(\lambda) - v(\mu)\|_{\mathbb{U}}$ , we get the desired Lipschitz inequality. □

**Lemma 4.5.** *If Assumptions 1 and 6.(ii) hold, the function  $\lambda \mapsto u(\lambda)$  is Lipschitz on  $\mathbb{U}$ .*

*Proof.* The proof is similar to the proof of Lemma 4.4. □

**Theorem 4.2.** *Let Assumptions 1, 3, 5, and 6 hold, then:  $u(\lambda^k) \xrightarrow[k \rightarrow \infty]{} u(\bar{\lambda})$  a.s.*

*Proof.* The convergence follows from the Lipschitz property of  $\lambda \mapsto u(\lambda)$  (as a result of Assumption 6) associated with the a.s. convergence of  $\{\lambda^k\}$ . □

**Remark 4.2.** *Note that Lemma 4.3 and Theorems 4.1 and 4.2 still hold when replacing  $\lambda^k$  by  $\check{\lambda}^k$  and  $Y^k$  by  $\check{Y}^k$  (as defined respectively in line 9 and 10 of the Sampled Stochastic Uzawa Algorithm 2). This can be proved by the same argument, using that  $\check{Y}^k$  is bounded a.s. and  $\mathbb{E}[\check{Y}^k | \check{\mathcal{G}}_k] = f(\check{\lambda}^k)$  for any  $k$ , where:*

$$\check{\mathcal{G}}_k = \sigma \left( \{W^{\ell,p}\} : 1 \leq \ell \leq m, p \leq k \right) \vee \sigma \left( \{I_\ell^p\} : 1 \leq \ell \leq m, p \leq k \right), \quad (4.28)$$

with  $W^{\ell,p}$  and  $I_\ell^p$  defined respectively in line 7 and 5 of the Sampled Stochastic Uzawa Algorithm 2.

**Remark 4.3.** We wish to emphasize that the proposed Algorithm 2 is particularly suitable for practical distributed implementations in agent-based scenarios. In these cases, the quantity  $\lambda$  can be interpreted as a common price signal that is broadcast to the independent agents, which in turn compute independently their optimal solution  $u(\lambda)$  on the basis of their local parameters (step 5).

To illustrate the results, we consider in the next section stochastic control problems in both continuous and discrete time settings.

## 5 Application to stochastic control

### 5.1 Continuous time setting

Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete filtered probability space on which  $W = (W^i)_{i=1, \dots, n}$  is a  $n \times d$ -dimensional Brownian motion such that, for any  $t \in [0, T]$  and  $i \in \{1, \dots, n\}$ ,  $W_t^i$  takes value in  $\mathbb{R}^d$  and generates the filtration  $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ . In the considered notation,  $\mathbb{P}$  stands for the Wiener measure associated with this filtration and  $\mathbb{F}$  for the augmented filtration by all  $\mathbb{P}$ -null sets. The following notations are used:

$$\mathbb{X} := \{\varphi : \Omega \rightarrow \mathcal{C}([0, T], \mathbb{R}^d) \mid \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted, } \|\varphi\|_{\infty, 2} := \mathbb{E} \sup_{\substack{1 \leq k \leq d \\ s \in [0, T]}} |\varphi_k(s)|^2 < \infty\},$$

$$\mathbb{U} := L^2([0, T], \mathbb{R}^p) := \{\varphi : [0, T] \rightarrow \mathbb{R}^p \mid \int_0^T \sum_{k=1}^p |\varphi_k(t)|^2 dt < \infty\},$$

For any  $i \in \{1, \dots, n\}$ , the feasible set of controls is defined by:

$$\begin{aligned} \mathcal{U}_i := \{ & v : \Omega \times [0, T] \rightarrow \mathbb{R}, v(\cdot) \text{ is } \mathbb{F}\text{-prog. measurable,} \\ & v(\omega) \in \mathbb{U} \text{ and } v_t(\omega) \in [-M_i, M_i]^p, \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega\}, \end{aligned} \quad (5.1)$$

and we set  $M := \max_{i \in \{1, \dots, n\}} M_i$ , where  $M_i > 0$ .

Each local agent  $i = 1, \dots, n$  is supposed to control its state variable through the control process  $u^i \in \mathcal{U}_i$  and is subject to independent uncertainties. More specifically, the state process of each agent,  $X^{i, u^i} = (X_t^{i, u^i})_{t \in [0, T]}$ , for  $i = 1, \dots, n$  takes values in  $\mathbb{R}^d$  and follows the dynamics for  $i \in \{1, \dots, n\}$ :

$$\begin{cases} dX_t^{i, u^i} &= \mu_i(t, u_t^i(\cdot, W^i), X_t^{i, u^i}) dt + \sigma_i(t, X_t^{i, u^i}) dW_t^i, \text{ for } t \in [0, T], \\ X_0^{i, u^i} &= x_0^i \in \mathbb{R}^d; \end{cases} \quad (5.2)$$

We assume that, for any  $i$ , there exist five functions  $\alpha_i \in L^\infty([0, T], \mathbb{R}^{d \times p})$ ,  $\beta_i, \theta_i \in L^\infty([0, T], \mathbb{R}^{d \times d})$ ,  $\gamma_i \in L^\infty([0, T], \mathbb{R}^d)$  and  $\xi_i \in L^\infty([0, T], \mathbb{R}^{d \times d \times d})$  such that, for any  $(t, \nu, x) \in [0, T] \times [-M, M]^p \times \mathbb{R}^d$ :

$$\mu_i(t, \nu, x) = \alpha_i(t)\nu + \beta_i(t)x + \gamma_i(t) \text{ and } \sigma_i(x, t) = \xi_i(t)x + \theta_i(t). \quad (5.3)$$

Without loss of generality, the initial states  $x_0^i$  are supposed to be deterministic. The process  $X^{i, u^i}$  is  $\mathbb{F}$ -progressively measurable. For all  $i$ ,  $\mathcal{F}^i$  stands for the natural filtration of the Brownian motion  $W^i$ .

#### 5.1.1 On the well-posedness of $(P_1)$

In this section, we discuss some conditions under which Problem  $(P_1)$  is well-posed.

**Lemma 5.1.** *Let  $i \in \{1, \dots, n\}$  and  $v \in \mathcal{U}_i$  be a control process. The map  $v \mapsto X^{i, v}$  is linear continuous from  $\mathcal{U}_i$  to  $\mathbb{X}$  and there exists a unique process  $X^{i, v} \in \mathbb{X}$  satisfying (5.2) (in the strong sense) such that, for any  $p \in [1, \infty)$ :*

$$\mathbb{E} \left[ \sup_{\substack{0 \leq t \leq T \\ 1 \leq k \leq d}} |X_{k, t}^{i, v}|^p \right] < C(r, T, x_0, K) < \infty. \quad (5.4)$$

*Proof.* The proof for the existence and uniqueness of a solution of (5.2) relies on [20, Theorem 3.6, Chapter 2]. The inequality is a result of [20, Theorem 4.4, Chapter 2].  $\square$

Let  $F_0 : \mathbb{U} \rightarrow \mathbb{R}$  be a proper, convex and lower semi continuous function, satisfying Assumptions 3 and 4.(ii). For any  $i \in \{1, \dots, n\}$ , we assume that there exists  $F_i$  such that the local cost  $G_i$  is of the form:

$$u^i \mapsto G_i(u^i(\cdot, \omega^{-i}), \omega^i) = F_i(u^i(\omega^i, \omega^{-i}), X^{i, u^i}(\omega^i)), \quad (5.5)$$

where  $F_i : \mathbb{U} \times \mathcal{C}([0, T] \times \mathbb{R}^d) \rightarrow \mathbb{R}_+$  is a proper and lower semi continuous function. Additional assumptions are formulated below.

**Assumption 7.** For any  $i \in \{1, \dots, n\}$ :

- (i)  $F_i$  is jointly convex w.r.t. both variables and strictly convex w.r.t. the first variable.
- (ii) there exists a positive integer  $r$  such that  $F_i$  has  $r$ -polynomial growth, i.e there exists  $K > 0$  such that for any  $x^i \in \mathcal{C}([0, T], \mathbb{R}^d)$  and  $u^i \in \mathbb{U}$ :  $|F_i(u^i, x^i)| \leq K(1 + \sup_{\substack{0 \leq t \leq T \\ 0 \leq k \leq n}} |x_{k,t}^i|^r)$ .

Note that if Assumption 7.(i) is satisfied, using the definition of  $F_0$  and  $F_i$  above in this section, Assumptions 1 to 4 hold.

**Remark 5.1.** It is worth highlighting the following aspects regarding Assumption 7:

1. Assumption 7.(i) is satisfied if there exist  $g_i : L^2((0, T), \mathbb{R}^p) \rightarrow \mathbb{R}$  strictly convex and  $h_i : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  convex, such that  $F_i(v, X) = g_i(v) + h_i(X)$ .
2. Observe that Assumption 7 satisfies Assumptions 1.(ii) and 4.(i)

From now on, Assumption 7 is in force. Now the optimization problems  $(P_1^c)$  and  $(\hat{P}_2^c)$  can be clearly defined:

$$(P_1^c) \begin{cases} \inf_{u \in \mathcal{U}} J^c(u) \\ J^c(u) := \mathbb{E} \left[ F_0\left(\frac{1}{n} \sum_{i=1}^n u^i(\omega)\right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i(\omega), X^{i, u^i}(\omega^i)) \right], \end{cases} \quad (5.6)$$

and

$$(\hat{P}_2^c) \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}^c(u) \\ \tilde{J}^c(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i \right) + \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n F_i(u^i(\omega), X^{i, u^i}(\omega^i)) \right], \end{cases} \quad (5.7)$$

Using the results of Section 2, we can state the following corollary:

**Corollary 5.1.** (i) Problems  $(P_1^c)$  and  $(\hat{P}_2^c)$  admit both a unique solution.

(ii) Any optimal solution of Problem  $(\hat{P}_2^c)$  is an  $\varepsilon$ -optimal solution, where  $\varepsilon = cM^2/n$ , of Problem  $(P_1^c)$ .

*Proof.* The proof of point (i) is a specific case of Lemma 2.1. Similarly, point (ii) is a particular case of Theorem 2.2.  $\square$

**Remark 5.2.** A practical example of this type of stochastic optimization problem is illustrated in Section 7, which considers the interactions between a large population of price-responsive self-interested domestic appliances and a central system operator which has to meet the prescribed levels of demand at minimum generation costs.

## 5.2 Discrete time setting

The main results of the paper are instantiated to the discrete time setting in this subsection. The following notations are used.

- Let  $n \in \mathbb{N}^*$  be the number of agents,  $d, p \in \mathbb{N}^*$  the dimension respectively of their state and control variables at any time step, and  $T \in \mathbb{N}^*$  the finite time horizon.
- For any matrix  $M$ , its transpose is denoted by  $M^\top$ .

- We consider a global noise process as a sequence of independent random variables  $(W_1, \dots, W_T)$ , where for any  $t \in \{1, \dots, T\}$ ,  $W_t$  is a vector of  $d$ -dimensional independent random variables, with finite variance, defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ :  $W_t := (W_t^1, \dots, W_t^d)$ , with  $W_t^i \in \mathbb{R}^d$ . For any  $i \in \{1, \dots, n\}$  and  $t \in \{1, \dots, T\}$  we define  $\mathcal{F}_t^i := \sigma(W_1^i, \dots, W_t^i)$  and  $\mathcal{F}_t := \otimes_{i=1}^n \mathcal{F}_t^i$ .
- The space  $\mathbb{X}$  is defined by:

$$\mathbb{X} := \{x = (x_0, \dots, x_T) \mid \forall t \in \{0, \dots, T\}, \mathbb{R}^d \ni x_t \text{ is } \mathcal{F}_t \text{-measurable and } \mathbb{E}\|x_t\|_2^2 < \infty\}. \quad (5.8)$$

- For any  $i \in \{1, \dots, n\}$ , we define the space of control  $\mathcal{U}^i$  of agent  $i$  by:

$$\mathcal{U}^i := \{u^i = (u_0^i, \dots, u_{T-1}^i) \mid \forall t \in \{0, \dots, T-1\}, \mathbb{R}^p \ni u_t^i \text{ is } \mathcal{F}_t \text{-measurable and } u_k^i(\omega) \in [-M, M]^p \text{ } \mathbb{P}\text{-a.s.}\}, \quad (5.9)$$

where  $M > 0$ . We finally set  $\mathcal{U} := \prod_{i=1}^n \mathcal{U}^i$ .

- For any  $i \in \{1, \dots, n\}$ ,  $X^{i, u^i} := (X_0^{i, u^i}, \dots, X_T^{i, u^i}) \in \mathbb{X}$  is the state trajectory of agent  $i$  controlled by  $u^i \in \mathcal{U}^i$ . We have the following dynamics:

$$\begin{cases} X_{t+1}^{i, u^i} &= A^i X_t^{i, u^i} + B^i u_t^i + W_{t+1}^i, \quad \text{for } t \in \{0, \dots, T-1\}, \\ X_0^{i, u^i} &= x_0 \in \mathbb{R}^d, \end{cases} \quad (5.10)$$

where  $A^i \in \mathbb{R}^{d \times d}$  and  $B^i \in \mathbb{R}^{d \times p}$ .

Let  $F_0 : \mathbb{R}^{p \times T} \rightarrow \bar{\mathbb{R}}$  be proper, lower semi continuous, convex and satisfy Assumptions 3 and 4.(ii). Similarly to the previous subsection, we assume that, for any  $i$ , there exists a function  $F_i : \mathbb{R}^{p \times T} \times \mathbb{R}^{d \times T} \rightarrow \mathbb{R}$  such that  $G_i$  and  $F_i$  satisfy (5.5), and  $F_i$  satisfies Assumption 7 for an integer  $r$  such that  $\mathbb{E}\|W_t\|^r$  is finite for any  $t \in \{1, \dots, T\}$ .

Now, for any  $n \in \mathbb{T}^*$ , the optimization problems  $(P_1^d)$  and  $(\hat{P}_2^d)$  can be clearly defined:

$$(P_1^d) \begin{cases} \inf_{u \in \mathcal{U}} J^d(u) \\ J^d(u) := \mathbb{E} \left[ F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i, X^{i, u^i}) \right], \end{cases} \quad (5.11)$$

and

$$(\hat{P}_2^d) \begin{cases} \inf_{u \in \hat{\mathcal{U}}} \tilde{J}^d(u) \\ \tilde{J}^d(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i \right) + \frac{1}{n} \mathbb{E} \left[ \sum_{i=1}^n F_i(u^i, X^{i, u^i}) \right]. \end{cases} \quad (5.12)$$

In the same spirit as in the previous subsection, we have the following results, which will be useful for the next section.

**Corollary 5.2.** (i) Problems  $(P_1^d)$  and  $(\hat{P}_2^d)$  admit both a unique solution.

(ii) Any optimal solution of Problem  $(\hat{P}_2^d)$  is an  $\varepsilon$ -optimal solution, where  $\varepsilon = cM^2/n$ , of Problem  $(P_1^d)$ .

*Proof.* The proof of point (i) is analogous to the one of Lemma 2.1. Similarly, proof of point (ii) is analogous to the one of Theorem 2.2.  $\square$

One can implement the *Stochastic Uzawa* (Algo 1) and the *Sampled Stochastic Uzawa* (Algo 2) in this discrete time setting with Lemma 4.3 and Theorems 4.1 and 4.2 still ensuring the algorithm convergence.



## 6 A numerical example: the LQG (Linear Quadratic Gaussian) problem

This section aims at illustrating numerically the convergence of the *Stochastic Uzawa* (Algo 1) on a simple example. The speed of convergence of the algorithm is evaluated according to the number of dual iterations and of agents. A linear quadratic formulation is considered, with  $n$  agents in a discrete setting Problem  $(\hat{P}_2^{LQG})$ . We use the notations of Section 5.2.

This framework constitutes a simple test case, since the (deterministic) Uzawa's algorithm can be performed, and one can compare the resulting multiplier estimate with the one provided by the *Stochastic Uzawa* algorithm. All the assumptions required for the convergence of the *Stochastic Uzawa* (Algo 1) are satisfied for Problem  $(\hat{P}_2^{LQG})$ . Moreover, the optimal solutions to the local problems (line 5 of Algorithm 1) can be resolved analytically.

Problem  $(\hat{P}_2^{LQG})$  is similar to  $(\hat{P}_2^d)$  defined in (5.12) but, in this specific case, the function  $F_0$  is a quadratic function of the aggregate strategies of the agents

$$F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i \right) := \frac{\nu}{2} \sum_{t=0}^T \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u_t^i - r_t \right)^2, \quad (6.1)$$

where  $\nu > 0$  and  $\{r_t\}$  is a deterministic target sequence. Similarly, the cost term  $F_i$  of the individual agents is expressed as a quadratic function of their state  $X^{i,u^i}$  and control  $u^i$

$$F_i(u^i, X^{i,u^i}) := \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2, \quad (6.2)$$

where  $q_i > 0$  and  $d_i > 0$  for any  $i \in \{1, \dots, n\}$ . Defining the matrices  $D = \text{diag}(d_1, \dots, d_n)$ ,  $Q = \text{diag}(q_1, \dots, q_n)$  and  $D^f = \text{diag}(d_1^f, \dots, d_n^f)$ , we get:

$$\sum_{i=1}^n F_i(u^i, X^{i,u^i}) = \frac{1}{2} \left( \sum_{t=0}^T X_t^{u^\top} D X_t^u + u_t^\top Q u_t \right) + \frac{1}{2} X_T^{u^\top} D^f X_T^u, \quad (6.3)$$

where, for any  $t \in \{0, \dots, T\}$ ,  $X_t^u := (X_t^{1,u^1}, \dots, X_t^{n,u^n}) \in \mathbb{R}^n$  is the controlled state vector of all the agents. Now the optimization Problem  $(\hat{P}_2^{LQG})$  is clearly defined.

To find the optimal multiplier and control of  $(\hat{P}_2^{LQG})$ , the *Stochastic Uzawa* Algorithm 1 is applied. In this specific case, the lines 4 and 6 take respectively the following form at any dual iteration  $k$ :

$$u^i(\lambda^k) := \arg \min_{u^i \in \hat{U}^i} \left\{ \mathbb{E} \left[ \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 + \lambda_t^k u_t^i \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2 \right] \right\}, \quad (6.4)$$

$$v(\lambda^k) := \arg \min_{v \in \mathbb{R}^T} \left\{ \left( \sum_{t=0}^T \nu (v_t - r_t)^2 - \lambda_t^k v_t \right) \right\}. \quad (6.5)$$

The optimization problem (6.4) solved by each local agent also falls within the LQG framework. One can solve these problems using the results of [30]. The resolution via Riccati equations of (6.4) shows that  $u^i(\lambda^k)$  is a linear function of the state  $X^{i,u^i}$  and of the price  $\lambda^k$ . Therefore, in this specific example, one can explicitly compute  $\mathbb{E}[u_t^i(\lambda^k) | \mathcal{G}_k]$  for any  $t$ , with  $\mathcal{G}_k$  as defined in (4.10).

Within this described framework, it is possible to implement the (deterministic) Uzawa's algorithm and use it as a reference to evaluate the performances of the *Stochastic Uzawa* algorithm.

Different population sizes  $n$  are considered, with  $n$  ranging between 1 and  $10^4$ . Similarly, the algorithm is stopped after different numbers of dual iteration  $k$ , ranging between 10 and  $10^4$ . In order to evaluate the bias and variance of the *Stochastic Uzawa* algorithm, this has been performed over  $J = 1000$  runs.

It is possible to define a Problem  $(\hat{P}_2^{LQG})$  and a dual Problem  $(D^{LQG})$  from Problem  $(\hat{P}_2^{LQG})$  following the same approach presented in (1.5) and (3.2) for the definition of  $(\hat{P}_2^d)$  and  $(D)$ , respectively, from  $(\hat{P}_2)$ . It can be shown that, for any  $n$ , there exists a unique optimal multiplier  $\bar{\lambda}^n$ , solution of  $(D^{LQG})$ . For any  $n$ , the quantity  $\lambda^{k,n,j}$  denotes the dual price computed during the  $j^{\text{th}}$  simulations ( $j = 1, \dots, J$ ) of the *Stochastic Uzawa* algorithm, after  $k$  dual iterations.

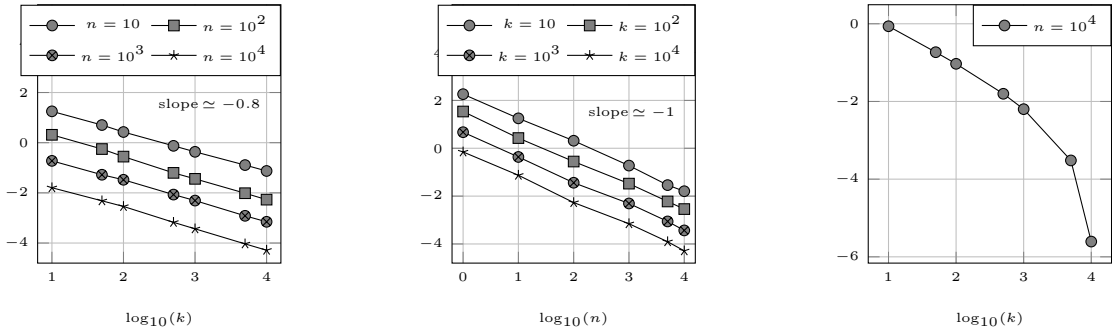


Figure 1: Variance term  $\log_{10}(v_{k,n})$ , expressed as a function of  $k$ , for different number of agents  $n$ . Figure 2: Variance term  $\log_{10}(v_{k,n})$ , expressed as a function of  $n$ , for different number of iterations  $k$ . Figure 3: Bias term  $\log_{10}(\|b_{k,n}\|_2^2)$ , expressed as a function of  $k$ , given the number of agents  $n = 10^4$ .

For any  $n$ , the deterministic multiplier  $\bar{\lambda}^n$  is obtained by applying Uzawa's algorithm, after  $10^4$  dual iterations. To this end, we applied the *Stochastic Uzawa Algorithm* 1, where we ignored the line 8 and we replaced the update of  $\lambda^k$  line 9 by:  $\bar{\lambda}^{k+1} \leftarrow \bar{\lambda}^k + \rho_k \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i(\bar{\lambda}^k) - v(\bar{\lambda}^k) \right)$ .

At each dual iteration  $k$ , the computation of  $\mathbb{E} u^i(\lambda^k)$  is straightforward in this specific case, since  $u^i(\lambda^k)$  is a linear function of  $X^{i,u^i}$  and  $\lambda^k$ , as explained in the previous subsection.

The multipliers  $\lambda^{k,n,j}$  and  $\bar{\lambda}^n$ , obtained by applying the *Stochastic Uzawa* and *Uzawa algorithms*, respectively, are now compared. For any  $k$  and  $n$ , let  $b_{k,n}$ ,  $v_{k,n}$  and  $\ell_{k,n}$  denote an estimation of the bias, the variance and the  $L^2$  norm of the error, respectively, as computed via Monte Carlo method with  $J$  simulations. For any  $k$  and  $n$ , these quantities are defined as follows:

$$b_{k,n} := \frac{1}{J} \sum_{j=1}^J \lambda^{k,n,j} - \bar{\lambda}^n,$$

$$v_{k,n} := \frac{1}{J} \sum_{j=1}^J \|\lambda^{k,n,j} - \bar{\lambda}^n - b_{k,n}\|_2^2,$$

$$\ell_{k,n} := v_{k,n} + \|b_{k,n}\|_2^2.$$

Since numerical simulations are based on finite dimensional approximations, it is relevant to compare the empirical convergence rates, shown in Figures 1-3, with the associated theoretical asymptotic rates presented in the literature for a finite dimensional setting.

In Figure 1, we observe a behavior in  $1/k^\alpha$  (with  $\alpha \simeq 0.8$ ) of the variance  $v_{k,n}$  w.r.t. the number of iterations  $k$ . This rate of convergence is consistent with [9, Theorem 2.2.12, Chapter 2], where the best asymptotic convergence rate for the Robbins-Monro algorithm is proved to be of the order of  $1/k$  (for the quadratic error).

In Figure 2 we observe a behavior in  $1/n^\beta$  (with  $\beta \simeq 1$ ) of the variance  $v_{k,n}$  w.r.t. the number of agents  $n$ . This is expected, following [9, Theorem 2.2.12, Chapter 2] and the observation that the variance of  $Y^{k+1}$  is of the order of  $1/n$  for any iteration  $k$ .

Finally, in Figure 3, we note that the bias  $\|b_{k,n}\|_2^2$  decreases faster than  $1/k$  w.r.t. the number of iterations  $k$ . Thus, for a large number of iterations ( $k > 0$ ), the dominant term impacting the error  $\ell_{k,n}$  is the variance  $v_{k,n}$ .

## 7 Price-based coordination of a large population of thermostatically controlled loads

The goal of this section is to demonstrate the applicability of the presented approach for the coordination of TCLs in the context of flexible power systems. In particular, the problem analyses the daily operation of a power system with a large penetration of price-responsive TCLs, adopting a modelling framework similar to [8]. Two distinct elements are considered: i) a system operator, which must schedule a portfolio of generation assets in order to satisfy the

energy demand at a minimum cost, and ii) a population of price-responsive TCLs that individually determine their ON/OFF power profile in response to price, with the objective of minimizing their operating cost while fulfilling users' requirements. Note that the operations of the two elements are interconnected, since the aggregate power consumption of the TCLs will modify the system-level demand profile that needs to be accommodated by the system operator.

## 7.1 Formulation of the problem

In the considered problem, the function  $F_0$  represents the minimized power production cost and corresponds to the resolution of an Unit Commitment (UC) problem. The UC determines generation scheduling decisions (in terms of energy production and frequency response (FR) provision) in order to minimize the short term operating cost of the system while matching generation and demand. The demand quantity is the sum of an inflexible deterministic term (denoted for any time instant  $t \in [0, T]$  by  $\bar{D}(t)$ ) and of a stochastic component  $n \times U_{TCL}(t)$ , i.e. the product of the population size  $n$  and the average demand profile  $U_{TCL}(t)$  of the TCL population.

For simplicity, a Quadratic Programming (QP) formulation in a discrete time setting is adopted for the UC problem. The central planner disposes of  $Z$  generation technologies (gas, nuclear, wind) and schedules their production and allocates response by slot of 30 min every day. For any  $j \in \{1, \dots, Z\}$  and  $\ell \in \{1, \dots, 48\}$ , the quantities  $H_j(t_\ell)$ ,  $G_j(t_\ell)$  and  $R_j(t_\ell)$  denote respectively the commitment, the power production and the frequency-response from unit  $j$  during the time interval  $[t_\ell, t_{\ell+1}]$  (all expressed in MWh). The associated vectors are denoted by  $H(t_\ell) = [H_1(t_\ell), \dots, H_Z(t_\ell)]$ ,  $G(t_\ell) = [G_1(t_\ell), \dots, G_Z(t_\ell)]$  and  $R(t_\ell) = [R_1(t_\ell), \dots, R_Z(t_\ell)]$ .

The cost sustained at time  $t_\ell$  by unit  $j$  is linear with respect to the commitment  $H_j(t_\ell)$  and quadratic with respect to generation  $G_j(t_\ell)$  and can be expressed as  $c_{1,j}H_j(t_\ell)G_j^{Max}(t_\ell) + c_{2,j}G_j(t_\ell) + c_{3,j}G_j(t_\ell)^2$ . In this cost expression,  $G_j^{Max}$  denotes the production limit allocated by each generation technology,  $c_{1,j}$  [ $\text{£}/\text{MWh}$ ] is the no-load cost term, whereas  $c_{2,j}$  [ $\text{£}/\text{MWh}$ ] and  $c_{3,j}$  [ $\text{£}/\text{MW}^2\text{h}$ ] are the production cost coefficients of the generation technology  $j$ . The optimization of  $F_0$  must satisfy the following constraints for all  $\ell \in \{1, \dots, 48\}$  and  $j \in \{1, \dots, Z\}$ :

$$\sum_{j=1}^Z G_j(t_\ell) - \int_{t_\ell}^{t_{\ell+1}} (\bar{D}(t) + nU_{TCL}(t))dt = 0, \quad (7.1)$$

$$0 \leq H_j(t_\ell) \leq 1, \quad (7.2)$$

$$R_j(t_\ell) - r_j H_j(t_\ell) G_j^{max}(t_\ell) \leq 0, \quad (7.3)$$

$$R_j(t_\ell) - s_j (H_j(t_\ell) G_j^{max}(t_\ell) - G_j(t_\ell)) \leq 0, \quad (7.4)$$

$$\Delta G_L - \Lambda (\bar{D}(t_\ell) + n(\bar{U}_{TCL}(t_\ell) - \bar{R}_{TCL}(t_\ell))) \Delta f_{qss}^{max} - \hat{R}(t_\ell) \leq 0, \quad (7.5)$$

$$2\Delta G_L t_{ref} t_d - t_{ref}^2 \hat{R}(t_\ell) - 4\Delta f_{ref} t_d \hat{H}(t_\ell) \leq 0, \quad (7.6)$$

$$\bar{q}(t) - \hat{H}(t) \hat{R}(t) \leq 0 \quad (7.7)$$

$$\mu r_j H_j(t_\ell) G_j^{max}(t_\ell) - G_j(t_\ell) \leq 0, \quad (7.8)$$

where (7.1) equals production and aggregated demand (i.e. the system inelastic demand  $\bar{D}$  and the TCL flexible demand  $nU_{TCL}$ ). The quantities  $\hat{R}$  and  $\hat{H}$  denote the total reserve and inertia of the system, respectively, and are defined for any  $\ell \in \{1, \dots, 48\}$  as:

$$\hat{R}(t_\ell) = \sum_{j=1}^Z R_j(t_\ell) + nR_{TCL}(t_\ell),$$

$$\hat{H}(t_\ell) = \sum_{j=1}^Z \frac{h_j H_j(t_\ell) G_j^{max} - h_L \Delta G_L}{f_0}.$$

In (7.2) it is supposed that, for any generation technology  $j$ , the capacity of the single power plant is significantly smaller than the total installed capacity. As a result, it is reasonable to consider the continuous relaxation of the UC problem by assuming  $H_j(t_\ell) \in [0, 1]$ .

The amount of response allocated by each generation technology is limited by the headroom  $r_j H_j(t_\ell) G_j^{max}(t_\ell)$  in (7.3) and by the slope  $s_j$  linking the FR with the dispatch level (7.4). Constraints (7.5) to (7.8) deal with frequency response provision and  $R_{TCL}$  (the mean of FR allocated by TCLs). They guarantee secure frequency

deviations following sudden generation loss  $\Delta G_L$ . Inequality (7.5) allocates enough FR (with delivery time  $t_d$ ) such that the quasi-steady-state frequency remains above  $\Delta f_{qss}^{max}$ , with  $\Lambda$  accounting for the damping effect introduced by the loads [17]. The constraint (7.7) imposes the maximum tolerable frequency deviation  $\Delta f_{nad}$ , following the formulation and methodology presented in [29] and [31]. The rate of change of frequency is taken into account in (7.6) where at  $t_{rcf}$  the frequency deviation remains above  $\Delta f_{ref}$ . Constraint (7.8) prevents trivial unrealistic solutions that may arise in the proposed formulation, such as high values of committed generation  $H_j(t_\ell)$  in correspondence with low (even zero) generation dispatch  $G_j(t_\ell)$ . The reader can refer to [8] for more details on the UC problem.

The solution  $F_0$  of the UC problem can be defined by the following optimization problem:

$$F_0(U_{TCL}, R_{TCL}) := \min_{H, G, R} \sum_{\ell=1}^{48} \sum_{j=1}^Z c_{1,j} H_j(t_\ell) G_j^{max}(t_\ell) + c_{2,j} G_j(t_\ell) + c_{3,j} G_j(t_\ell)^2, \quad (7.9)$$

subject to equations (7.1)-(7.8).

Note that the formulation of the present problem does not fulfill all the assumptions presented in Section 4. In particular, the function  $F_0$  is not strictly convex, as instead supposed in Theorem 4.1.(ii).(iii). Nevertheless, the numerical simulations of Section 7.2 shows that the proposed approach is still able to achieve convergence.

Regarding the modelling of the individual price-responsive TCLs, each TCL  $i \in \{1, \dots, n\}$  is characterized at any time  $t \in [0, T]$  by its temperature state  $X_t^{i, u^i}$  [ $^\circ C$ ] and by its power consumption control  $u_t^i$  [ $W$ ]. The thermal dynamic  $X_t^{i, u^i}$  of a single TCL  $i$  is given by:

$$\begin{cases} dX_t^{i, u^i} &= -\frac{1}{\gamma_i} (X_t^{i, u^i} - X_{OFF}^i + \zeta_i u_t^i) dt + \sigma_i dW_t^i, \quad \text{for } t \in [0, T], \\ X_{0, u^i}^i &= x_0^i \in \mathbb{R}, \end{cases} \quad (7.10)$$

where:

- $\gamma_i$  is its thermal time constant [ $s$ ].
- $X_{OFF}^i$  is the ambient temperature [ $^\circ C$ ].
- $\zeta_i$  is the heat exchange parameter [ $^\circ C/W$ ].
- $\sigma_i$  is a positive constant [ $(^\circ C)s^{\frac{1}{2}}$ ],
- $W^i$  is a Brownian Motion [ $s^{\frac{1}{2}}$ ], independent from  $W^j$  for any  $j \neq i$ .

For any  $i \in \{1, \dots, n\}$ , the set of control  $\mathcal{U}_i$  is defined by:

$$\mathcal{U}_i := \left\{ v : \Omega \times [0, T] \rightarrow \mathbb{R}, v(\cdot) \text{ is } \mathbb{F} - \text{prog. measurable}, \right. \\ \left. v(\omega) \in \mathbb{U} \text{ and } v_t(\omega) \in \{0, P_{ON, i}\}, \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega \right\}, \quad (7.11)$$

The TCLs dynamics in (7.10) have been derived according to [15], with the addition of the stochastic term  $\sigma_i dW_t^i$  to account for the influence of the environment (opening/closing of the fridge, environment temperature, etc.) on the evolution of the TCL temperature.

By combining the objective functions of the systems, the system operator has to solve the following optimization problem:

$$(P_1^{TCL}) \begin{cases} \inf_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left[ F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i, \frac{1}{n} \sum_{i=1}^n r_i(u^i, X^{i, u^i}) \right) \right] \\ + \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i, u^i}) ds + \gamma_i (X_T^{i, u^i} - \bar{X}^i)^2 \right], \end{cases} \quad (7.12)$$

The term  $r_i(u^i, X^{i, u^i})$  denotes the maximum amount of FR allocated by the TCL  $i$  at time  $s$  and can be expressed as:

$$r_i(u^i, X^{i, u^i})(s) := u_s^i \frac{X_s^{i, u^i} - X_{min}^i}{X_{max}^i - X_{min}^i}. \quad (7.13)$$

The discomfort term of the single TCL  $i$  at time  $s$  is denoted by  $f_i(u_s^i, X_s^{i,u^i})$ , which takes the following expression:

$$f_i(u_s^i, X_s^{i,u^i}) := \alpha_i (X_s^{i,u^i} - \bar{X}^i)^2 + \beta_i ((X_{\min}^i - X_s^{i,u^i})_+^2 + (X_s^{i,u^i} - X_{\max}^i)_+^2), \quad (7.14)$$

where:

- $\alpha_i (X_s^{i,u^i} - \bar{X}^i)^2$  is a discomfort term penalizing temperature deviations from some comfort target  $\bar{X}$  [ $^\circ\text{C}$ ], considering  $\alpha_i$  [ $\text{£}/h(^\circ\text{C})^2$ ] as a discomfort term parameter.
- $\beta_i ((X_s^{i,u^i} - X_{\min}^i)_+^2 + (X_{\max}^i - X_s^{i,u^i})_+^2)$  is a penalization term meant to maintain the temperature within the interval  $[X_{\min}^i, X_{\max}^i]$ , considering the cost parameter  $\beta_i$  [ $\text{£}/s(^\circ\text{C})^2$ ] and the maximum function  $(a)_+ = \max(0, a)$ .
- $\gamma_i (X_T^{i,u^i} - \bar{X}^i)^2$  is a terminal cost term meant to impose soft periodic constraints by quadratically penalizing the deviations of the final temperature state  $X_T^{i,u^i}$  with respect to the initial temperature value  $\bar{X}^i$ , considering the cost parameter  $\gamma$  [ $\text{£}/s(^\circ\text{C})^2$ ].

Note that the control set  $\mathcal{U}$  is not convex. We can mention a possible relaxation of the problem by taking the control in the interval  $[0, P_{ON,i}]$ .

In order to solve  $(P_1^{TCL})$ , the modified Problem  $(P_2^{TCL})$  is studied:

$$(P_2^{TCL}) \quad \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} u^i, \frac{1}{n} \sum_{i=1}^n \mathbb{E} r_i(u^i, X^{i,u^i}) \right) \\ + \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i,u^i}) ds + \gamma_i (X_T^{i,u^i} - \bar{X}^i)^2 \right]. \end{cases} \quad (7.15)$$

## 7.2 Decentralized implementation

The *Sampled Stochastic Uzawa* Algorithm 2 is applied to solve  $(P_2^{TCL})$ , with  $m = 317$  simulations per iteration. At each iteration  $k$ , the lines 4 and 6 of Algorithm 2 correspond to the solution of a deterministic UC problem and of an Hamilton Jacobi Bellman (HJB) equation, respectively. The time steps  $\Delta t = 7.6$  s and temperature steps  $\Delta T = 0.15^\circ\text{C}$  are chosen for the discretization of the HJB equation. Let us note that, at line 6, each TCL solves its own local problem on the basis of the received price signal  $\lambda^k = (p^k, \rho^k)$ :

$$\inf_{u^i \in \mathcal{U}_i} \int_0^T f_i(u_s^i, X_s^{i,u^i}) + u_s^i p_s^k - r_i(u^i, X^{i,u^i})(s) \rho_s^k ds, \quad (7.16)$$

where  $f_i(u_s^i, X_s^{i,u^i})$  is a discomfort term defined in (7.14),  $u_s^i p_s^k$  can be interpreted as consumption cost and  $r_i(u^i, X^{i,u^i})(s) \rho_s^k$  as fee awarded for FR provision. This implementation has a practical sense: each TCL uses local information and the received price signals to schedule its power consumption on the time interval  $[0, T]$ , with the objective of minimizing its overall costs. It follows that, with the proposed approach, it is possible to optimize the total system costs in  $(P_1^{TCL})$  in a distributed manner, with each TCL acting independently and pursuing its own cost minimization.

## 7.3 Results

In the proposed case study, the considered generation technologies available in the system are nuclear, combined cycle gas turbines (CCGT), open cycle gas turbines (OCGT) and wind. The characteristics and parameters of the UC in this simulation are the same as in [8].

It is assumed that the population of TCLs corresponds to  $n = 2 \times 10^7$  fridges with built-in freeze compartment that operate in the system according to the proposed price-based control scheme. For any TCL  $i$ , we set the consumption parameter  $P_{ON,i} = 180\text{W}$ . The values of the TCL dynamic parameters  $\gamma_i$  and  $X_{OFF}^i$  of (7.10) are equal to the ones considered in [8]. The initial temperatures of the TCLs are selected randomly according to a uniform probability distribution, considering temperature values between  $-21^\circ\text{C}$  and  $-14^\circ\text{C}$ . For any TCL  $i$ , the parameters of the individual cost function  $f_i$ , defined in (7.14), are:  $\alpha_i = 0.2 \times 10^{-4}$   $\text{£}/s(^\circ\text{C})^2$ ,  $\beta_i = 50 \text{£}/s(^\circ\text{C})^2$ ,  $\bar{X}^i = -17.5^\circ\text{C}$  and  $X_{\max} = -14^\circ\text{C}$ ,  $X_{\min} = -21^\circ\text{C}$ . The parameter  $\beta_i$  is intentionally taken very large to ensure

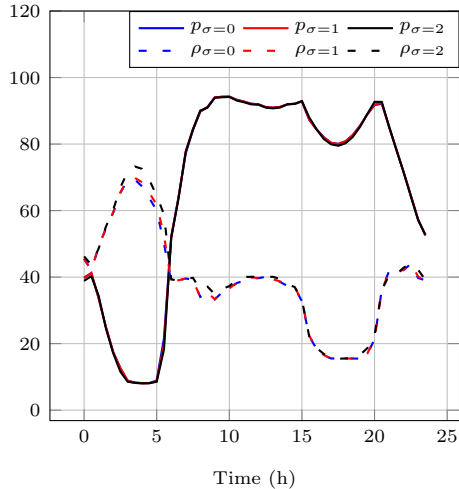
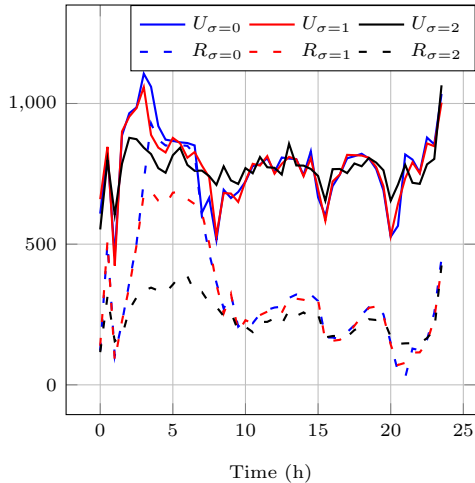


Figure 4: Total power consumption  $U$  and allocated response  $R$  (MW) of the TCLs after 75 algorithm iterations. Figure 5: Electricity price  $p$  and response availability price  $\rho$  (£/MWh) after 75 algorithm iterations.

that the TCL temperature remains within the interval  $[X_{max}^i, X_{min}^i]$ . Note that the individual problems solved by the TCLs are distinct than the ones in [8] (different terms and parameters).

Simulations are performed for different volatility values  $\sigma_i := 0, 1, 2$  (all the TCLs have the same volatility in the simulations), with  $\sigma_i$  defined as in (7.10). The *Sampled Stochastic Uzawa* Algorithm is stopped after 75 iterations.

The resulting profiles of total power consumption  $U = nU_{TCL}$  and total allocated response  $R = nR_{TCL}$  by the TCLs population are reported in Figure 4, while the resulting electricity prices  $p$  and response availability prices  $\rho$  are shown in Figure 5. As observed in [8], the total consumption  $U$  is higher when the electricity price  $p$  is lower. Conversely, the total allocated response  $R$  is higher when the FR remuneration price  $\rho$  is also higher. This can be observed in particular during the first hours of the day, between 0 and 6 h. The power consumption  $U$  exhibits smaller oscillations during the rest of the day, as the internal temperature of the TCLs is maintained within feasible levels. Although the prices do not seem to be particularly sensitive with respect to the volatility parameter  $\sigma$ , the power consumption  $U$  and frequency response  $R$  are highly correlated to the volatility of the TCLs temperature.

The TCLs impact on system commitment decisions and consequent energy/FR dispatch levels is also analyzed and displayed in Figures 6 and 7. In this analysis, the “flexibility scenario”, obtained with the proposed optimization strategy and considering flexible price-responsive TCLs, is compared to a “business-as-usual” scenario where the TCL do not respond to external price signals and do not perform any optimization of their costs. In the “business-as-usual”, we impose  $R_{TCL}(t) = 0$  and we assume that the TCLs operate exclusively according to their internal temperature  $X^{i,u^i}$ . They switch ON ( $u^i(t) = P_{ON,i}$ ) when they reach their maximum feasible temperature  $X_{max}^i$  and they switch back OFF again ( $u^i(t) = 0$ ) when they reach the minimum temperature  $X_{min}^i$ . In Figure 6, we can clearly observe that TCL’s flexibility allows to increase the contribution of wind to the energy balance of the system while decreasing the contribution of CCGT both in energy and frequency response. In the “business-as-usual” scenario, without frequency support by the TCL, the optimal solution envisages a further curtailment of wind output in favor of an increase in CCGT generation, as wind does not provide any FR. As expected, the influence of the TCL on the system is larger when the temperature volatility is lower.

A comparison of the system costs (i.e. UC solution) between the “flexibility scenario” (FS) and the “Business-as-usual” (BAU) framework is provided in Table 1. As expected, costs are lower in the FS, as the flexibility of the TCLs positively supports system operation, allowing to replace gas generation from OCGT and CCGT plants with cheaper wind energy. The reduction is higher (about 1.9%) for  $\sigma = 0$  with respect to the cases with  $\sigma = 1$  or  $\sigma = 2$  (about 1.6% and 1.2%, respectively). This confirms previous indications that TCLs tend to be more flexible when the volatility of their internal temperature is lower.

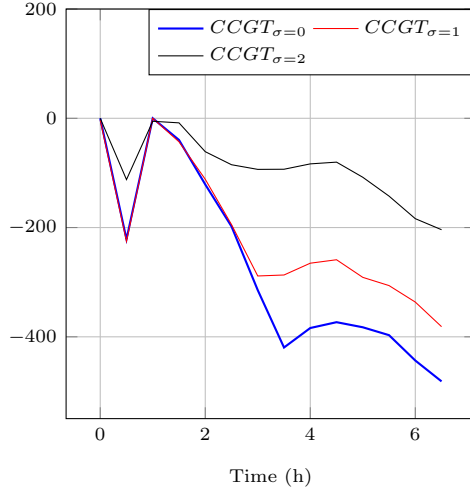
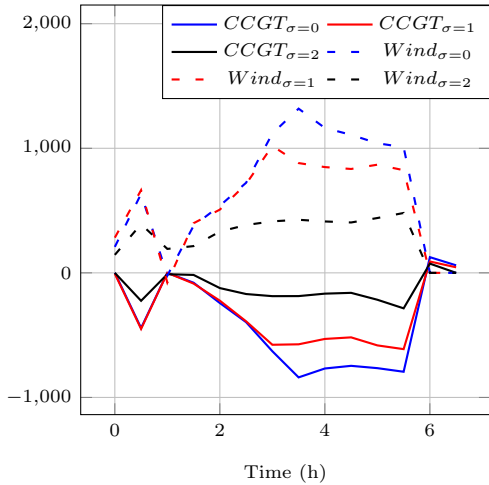


Figure 6: Deviation of generation profiles (MW) from the “business-as-usual” scenario during the first hours of the day, considering three different values of temperature volatility  $\sigma$ .

Figure 7: Deviation of Frequency Response (MW) allocated by CCGT technology with respect to the “business-as-usual” scenario during the first hours of the day, considering three different values of temperature volatility  $\sigma$ .

	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
BAU	$2.770 \times 10^7$	$2.770 \times 10^7$	$2.772 \times 10^7$
FS	$2.719 \times 10^7$	$2.725 \times 10^7$	$2.740 \times 10^7$

Table 1: Minimized system costs in (£)

## 8 Conclusions

Randomness and high dimensionality usually make the resolution of an optimization problem quite difficult. However, in the specific case of convex aggregative control problems, we have shown that, under independent noise assumptions, one can take advantage of the high dimension to approximate accurately the original Problem ( $P_1$ ) by a decentralized Problem ( $\hat{P}_2$ ), whose numerical resolution is more tractable. We highlight the fact that the approximation error is of order  $\frac{1}{n}$ , where  $n$  is the number of agents. The extension of this approach to stochastic control problems with common noise or to non convex problems may be challenging but interesting topics for further work.

## A Appendix

**Lemma A.1.** *Let  $H$  be a Hilbert space and  $f : H \rightarrow \mathbb{R}$  be l.s.c. and convex. The function  $f$  has at most quadratic growth if and only if its subgradient has linear growth.*

*Proof.* Let the subgradient have linear growth, that is,  $\|q\|_H \leq c_1(1 + \|x\|_H)$  whenever  $x \in H$  and  $q \in \partial f(x)$ . Then  $f(x) \leq f(0) + \langle q, x \rangle_H \leq f(0) + \|q\|_H \|x\|_H \leq c_2(1 + \|x\|_H^2)$ , so that  $f$  has at most quadratic growth.

Conversely, let  $f$  have at most quadratic growth. Since  $f$  is convex, one has for all  $x \in H$  and  $q_0 \in \partial f(0)$ :

$$f(x) \geq f(0) + \langle q_0, x \rangle_H \geq -c_3(1 + \|x\|_H^2),$$

where  $c_3 > 0$  depends only on  $f(0)$  and  $q_0$ . Then, using the growth assumption on  $f$  and the inequality above, one gets for all  $x, y \in H$  and  $q \in \partial f(x)$ :

$$c_4(1 + \|y\|_H^2) \geq f(y) \geq f(x) + \langle q, y - x \rangle_H \geq -c_3(1 + \|x\|_H^2) + \langle q, y - x \rangle_H.$$

Take  $y = x + \alpha q$ , with  $\alpha \in (0, 1)$ , we get

$$2c_4(1 + \|x\|_H^2 + \alpha^2 \|q\|_H^2) \geq -c_3(1 + \|x\|_H^2) + \alpha \|q\|_H^2$$

so that  $(\alpha - 2c_4\alpha^2)\|q\|_H^2 \leq (2c_4 + c_3)(1 + \|x\|_H^2)$ . Take  $\alpha = 1/(4c_4)$ , then  $\alpha - 2c_4\alpha^2 = 1/(8c_4) > 0$  and then

$$\|q\|_H^2 \leq 8c_4(2c_4 + c_3)(1 + \alpha\|x\|_H^2)$$

and the conclusion follows. □

## References

- [1] Kengy Barty, Pierre Carpentier, and Pierre Girardeau, *Decomposition of large-scale stochastic optimal control problems*, RAIRO-Operations Research **44** (2010), no. 3, 167–183.
- [2] Dimitri Bertsekas and Steven E Shreve, *Stochastic optimal control: the discrete-time case*, vol. 5, Athena Scientific, 1996.
- [3] J Frédéric Bonnans and Alexander Shapiro, *Perturbation analysis of optimization problems*, Springer Science & Business Media, 2013.
- [4] David B Brown, James E Smith, and Peng Sun, *Information relaxations and duality in stochastic dynamic programs*, Operations research **58** (2010), no. 4-part-1, 785–801.
- [5] Pierre Carpentier, J-Ph Chancelier, Vincent Leclère, and François Pacaud, *Stochastic decomposition applied to large-scale hydro valleys management*, European Journal of Operational Research **270** (2018), no. 3, 1086–1098.
- [6] Pierre Carpentier, Jean-Philippe Chancelier, Michel De Lara, and François Pacaud, *Mixed spatial and temporal decompositions for large-scale multistage stochastic optimization problems*, Journal of Optimization Theory and Applications **186** (2020), no. 3, 985–1005.
- [7] Michael Chertkov and Vladimir Chernyak, *Ensemble of thermostatically controlled loads: Statistical physics approach*, Scientific reports **7** (2017), no. 1, 8673.
- [8] Antonio De Paola, Vincenzo Trovato, David Angeli, and G. Strbac, *A mean field game approach for distributed control of thermostatic loads acting in simultaneous energy-frequency response markets*, IEEE Transactions on Smart Grid **PP** (2019), 1–1.
- [9] Marie Duflo, *Random iterative models*, vol. 34, Springer Science & Business Media, 2013.
- [10] Caroline Geiersbach and Georg Ch Pflug, *Projected stochastic gradients for convex constrained problems in Hilbert spaces*, SIAM Journal on Optimization **29** (2019), no. 3, 2079–2099.
- [11] Pierre Girardeau, *Solving large-scale dynamic stochastic optimization problems*, PhD thesis, Université Paris-Est, December 2010.
- [12] He Hao, Borhan M Sanandaji, Kameshwar Poolla, and Tyrone L Vincent, *Aggregate flexibility of thermostatically controlled loads*, IEEE Transactions on Power Systems **30** (2014), no. 1, 189–198.
- [13] Julia L Higle and Suvrajeet Sen, *Stochastic decomposition: a statistical method for large scale stochastic linear programming*, vol. 8, Springer Science & Business Media, 2013.
- [14] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal, *Fundamentals of convex analysis*, Springer Science & Business Media, 2004.
- [15] Arman C Kizilkale, Rabih Salhab, and Roland P Malhamé, *An integral control formulation of mean field game based large scale coordination of loads in smart grids*, Automatica **100** (2019), 312–322.
- [16] Daniel Kuhn, Wolfram Wiesemann, and Angelos Georghiou, *Primal and dual linear decision rules in stochastic and robust optimization*, Mathematical Programming **130** (2011), no. 1, 177–209.
- [17] Prabha Kundur, Neal J Balu, and Mark G Lauby, *Power system stability and control*, vol. 7, McGraw-hill New York, 1994.
- [18] Vincent Leclere, *Contributions to decomposition methods in stochastic optimization*, PhD thesis, Université Paris-Est, June 2014.



- [19] Vincent Leclere, Pierre Carpentier, Jean-Philippe Chancelier, Arnaud Lenoir, and François Pacaud, *Exact converging bounds for stochastic dual dynamic programming via fenchel duality*, SIAM Journal on Optimization **30** (2020), no. 2, 1223–1250.
- [20] Xuerong Mao, *Stochastic differential equations and applications*, Elsevier, 2007.
- [21] François Pacaud, *Decentralized optimization for energy efficiency under stochasticity*, PhD thesis, Université Paris-Est, October 2018.
- [22] Mario VF Pereira and Leontina MVG Pinto, *Multi-stage stochastic optimization applied to energy planning*, Mathematical programming **52** (1991), no. 1-3, 359–375.
- [23] Pedro Pérez-Aros and Emilio Vilches, *An enhanced Baillon-Haddad Theorem for convex functions defined on convex sets*, Applied Mathematics & Optimization **83** (2021), no. 3, 2241–2252.
- [24] Andrew B Philpott and Ziming Guan, *On the convergence of stochastic dual dynamic programming and related methods*, Operations Research Letters **36** (2008), no. 4, 450–455.
- [25] R Tyrrell Rockafellar and Roger J-B Wets, *Scenarios and policy aggregation in optimization under uncertainty*, Mathematics of operations research **16** (1991), no. 1, 119–147.
- [26] Andrzej Ruszczyński and Alexander Shapiro, *Stochastic programming models*, Handbooks in operations research and management science **10** (2003), 1–64.
- [27] David H Salinger, *A splitting algorithm for multistage stochastic programming with application to hydropower scheduling*, University of Washington, 1997.
- [28] Joe A Short, David G Infield, and Leon L Freris, *Stabilization of grid frequency through dynamic demand control*, IEEE Transactions on power systems **22** (2007), no. 3, 1284–1293.
- [29] Fei Teng, Vincenzo Trovato, and Goran Strbac, *Stochastic scheduling with inertia-dependent fast frequency response requirements*, IEEE Transactions on Power Systems **31** (2015), no. 2, 1557–1566.
- [30] Emanuel Todorov, *Optimal control theory*, Bayesian brain: probabilistic approaches to neural coding (2006), 269–298.
- [31] Vincenzo Trovato, Agnès Bialecki, and Anes Dallagi, *Unit commitment with inertia-dependent and multispeed allocation of frequency response services*, IEEE Transactions on Power Systems **34** (2018), no. 2, 1537–1548.
- [32] Vincenzo Trovato, Simon H Tindemans, and Goran Strbac, *Leaky storage model for optimal multi-service allocation of thermostatic loads*, IET Generation, Transmission & Distribution **10** (2016), no. 3, 585–593.
- [33] Shixuan Zhang and Xu Andy Sun. Stochastic dual dynamic programming for multistage stochastic mixed-integer nonlinear optimization. *Mathematical Programming*, 196(1):935–985, 2022.
- [34] Guanghui Lan. Complexity of stochastic dual dynamic programming. *Mathematical Programming*, pages 1–38, 2020.