



**HAL**  
open science

## Decomposition of High Dimensional Aggregative Stochastic Control Problems

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

► **To cite this version:**

Adrien Seguret, Clémence Alasseur, J Frédéric Bonnans, Antonio de Paola, Nadia Oudjane, et al..  
Decomposition of High Dimensional Aggregative Stochastic Control Problems. 2021. hal-02917014v4

**HAL Id: hal-02917014**

**<https://inria.hal.science/hal-02917014v4>**

Preprint submitted on 20 Oct 2021 (v4), last revised 2 Mar 2023 (v5)

**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.

1           **DECOMPOSITION OF HIGH DIMENSIONAL AGGREGATIVE**  
2           **STOCHASTIC CONTROL PROBLEMS**

3           \*

4           ADRIEN SEGURET<sup>†</sup>, CLÉMENTCE ALASSEUR<sup>‡</sup>, J. FRÉDÉRIC BONNANS<sup>§</sup>, ANTONIO  
5           DE PAOLA<sup>¶</sup>, NADIA OUDJANE<sup>||</sup>, AND VINCENZO TROVATO<sup>#</sup>

6           **Abstract.** We consider the framework of high dimensional stochastic control problem, in which  
7           the controls are aggregated in the cost function. As first contribution we introduce a modified  
8           problem, whose optimal control is under some reasonable assumptions an  $\varepsilon$ -optimal solution of the  
9           original problem. As second contribution, we present a decentralized algorithm whose convergence  
10          to the solution of the modified problem is established. Finally, we study the application to a problem  
11          of coordination of energy production and consumption of domestic appliances.

12          **Key words.** Stochastic optimization, Lagrangian decomposition, Uzawa’s algorithm, stochastic  
13          gradient.

14          **AMS subject classifications.** 93E20,65K10, 90C25, 90C39, 90C15.

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

18          (1.1)    ( $P_1$ )    
$$\begin{cases} \text{Min}_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i(\omega^i, \omega^{-i}) \right) + \frac{1}{n} \sum_{i=1}^n G_i(u^i(\cdot, \omega^{-i}), \omega^i) \right). \end{cases}$$

19          Here the noise  $\omega := (\omega^1, \dots, \omega^n)$  belongs to  $\Omega := \prod_{i=1}^n \Omega^i$ ,  $(\Omega^i, \mathcal{F}^i, \mu^i)$  is a proba-  
20          bility space, and  $(\Omega, \mathcal{F}, \mu)$  is the corresponding product probability space. Let  $\omega^{-i} :=$   
21           $(\omega^1, \dots, \omega^{i-1}, \omega^{i+1}, \dots, \omega^n)$  denotes an element of the space  $\Omega^{-i} := \prod_{j=1, j \neq i}^n \Omega^j$ . The  
22          associated product probability space is  $(\Omega^{-i}, \mathcal{F}^{-i}, \mu^{-i})$ , where  $\mathcal{F}^{-i} := \otimes_{j=1, j \neq i}^n \mathcal{F}^j$   
23          and  $\mu^{-i} := \prod_{j=1, j \neq i}^n \mu^j$ . Each decision variable  $u^i$  is a random variable (i.e. is  
24           $\mathcal{F}$ -measurable), square summable with value in a Hilbert space  $\mathbb{U}$  so that  $u :=$   
25           $(u^1, \dots, u^n)$  belongs to  $L^2(\Omega, (\mathbb{U})^n)$ . The function  $\omega^i \mapsto u^i(\omega^i, \omega^{-i})$  is denoted by  
26           $u^i(\cdot, \omega^{-i})$  and is a.s. (in  $\omega^{-i}$ )  $\mathcal{F}^i$ -measurable and belongs to  $L^2(\Omega^i, \mathbb{U})$ . Also,  $\mathcal{U} :=$   
27           $\prod_{i=1}^n \mathcal{U}_i$  where  $\mathcal{U}_i$  is, for  $i = 1$  to  $n$ , a closed convex subset of  $L^2(\Omega^i, \mathbb{U})$ . (In the ap-  
28          plication to dynamical problems, the constraint  $u^i \in \mathcal{U}_i$  includes the constraint of

\* Submitted to the editors September 23, 2020

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

<sup>†</sup>OSIRIS department, EDF Lab, Paris-Saclay, France ([adrien.seguret@edf.fr](mailto:adrien.seguret@edf.fr)) and PSL Research University, Universite Paris-Dauphine, CEREMADE, Paris, France.

<sup>‡</sup>OSIRIS department, EDF Lab, Paris-Saclay, France ([clemence.alasseur@edf.fr](mailto:clemence.alasseur@edf.fr)).

<sup>§</sup>Disco Team, L2S, CentraleSupélec/Université Paris-Saclay and Inria-Saclay ([Fred-eric.Bonnans@inria.fr](mailto:Fred-eric.Bonnans@inria.fr)).

<sup>¶</sup>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>||</sup>OSIRIS department, EDF Lab, Paris-Saclay, France ([nadia.oudjane@edf.fr](mailto:nadia.oudjane@edf.fr)).

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

29 adaptation of  $u^i$  to some filtration.) If each  $u^i$  is a random variable of  $\omega^i$ , for  $i = 1$   
 30 to  $n$ , we say that  $u$  is a decentralized decision variable.

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

36 **1.1. Motivations.** This work is motivated by its potential applications for large-  
 37 scale coordination of flexible appliances, to support power system operation in a con-  
 38 text of increasing penetration of renewables. One type of appliances that has been  
 39 consistently investigated in the last few years, for its intrinsic flexibility and potential  
 40 for network support, includes thermostatically controlled loads (TCLs) such as re-  
 41 frigerators or air conditioners. Several papers have already investigated the potential  
 42 of dynamic demand control and frequency response services of TCLs [22] and how  
 43 the population recovers from significant perturbations [4]. The coordination of TCLs  
 44 can be performed in a centralized way, like in [9]. However this approach raises chal-  
 45 lenging problems in terms of communication requirements and customer privacy. A  
 46 common objective can be reached in a fully distributed approach, like in [26], where  
 47 each TCL is able to calculate its own actions (ON/OFF switching) to pursue a com-  
 48 mon objective. This paper is related to the work of De Paola *et al.* [5], where each  
 49 agent represents a flexible TCL device. In [5] a distributed solution is presented for  
 50 the operation of a population of  $n = 2 \times 10^7$  refrigerators providing frequency sup-  
 51 port and load shifting. They adopt a game-theory framework, modelling the TCLs as  
 52 price-responsive rational agents that schedule their energy consumption and allocate  
 53 their frequency response provision in order to minimize their operational costs. The  
 54 potential practical application of our work also considers a large population of TCLS  
 55 which, contrarily to [5], have stochastic dynamics. The proposed approach is able to  
 56 minimize the overall system costs in a distributed way, with each TCL determining  
 57 its optimal power consumption profile in response to price signals.

58 **1.2. Related literature.** The considered problem belongs to the class of  
 59 stochastic control: looking for strategies minimizing the expectation of an objective  
 60 function under specific constraints. One of the main approaches proposed in the litera-  
 61 ture to tackle this problem is to use random trees: this consists in replacing the almost  
 62 sure constraints, induced by non-anticipativity, by a finite number of constraints to  
 63 get a finite set of scenarios (see. [10] and [20]). Once the tree structure is built, the  
 64 problem is solved by different decomposition methods such as scenario decomposition  
 65 [19] or dynamic splitting [21]. The main objective of the scenario method is reducing  
 66 the problem to an approximated deterministic one. The paper focuses on high dimen-  
 67 sional noise problems with large number of time steps, for which this approach is not  
 68 feasible. The idea of reducing a single high dimensional problem to a large number  
 69 with low dimension has been widely studied in the deterministic case. In determinis-  
 70 tic and stochastic problems a possibility is to use time decomposition thanks to the  
 71 Dynamic Programming Principle [1] taking advantage of Markov property of the sys-  
 72 tem. However, this method requires a specific time structure of the cost function and  
 73 fails when applied to problems for which the state space dimension is greater than  
 74 five. One can deal with the curse of dimensionality, under continuous linear-convex  
 75 assumptions, by using the Stochastic Dual Dynamic Programming algorithm (SDDP)

[16] to get upper and lower bounds of the value function, using polyhedral approximations. Though the almost-sure convergence of a broad class of SDDP algorithms has been proved [18], there is no guarantee on the speed of the convergence and there is no good stopping test. In [15], 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 magnitude similar to one of the present paper ( $n > 1000$ ). It is natural for this type of high dimensional problem to investigate decomposition techniques in the spirit of the Dual Approximation Dynamic Programming (DADP). DADP has been developed in PhD theses (see [8], [13]). 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 [13], 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 [8, chapter 4] and hydro valley management [2]. 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.** We consider the following approximation of problem ( $P_1$ ):

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

Let  $\hat{\mathcal{U}}$  be the set of decentralized controls, defined by:

$$(1.3) \quad \hat{\mathcal{U}} := \prod_{i=1}^n \hat{\mathcal{U}}_i, \text{ where } \hat{\mathcal{U}}_i := \{u^i \in \mathcal{U}_i \mid u^i \text{ is } \mathcal{F}^i - \text{measurable}\}.$$

The decentralized version of problem ( $P_2$ ) (i.e.  $\tilde{J}$  is optimized over the set  $\hat{\mathcal{U}}$ ) can be written as:

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

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

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

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  $\varepsilon_n \rightarrow 0$

106 when  $n \rightarrow \infty$ . In addition, an approach of price decomposition for  $(P_2)$ , based on the  
 107 formulation  $(P'_2)$ , is easier than for  $(P_1)$ , since the Lagrange multiplier is deterministic  
 108 for  $(P'_2)$ , whereas it is stochastic for  $(P_1)$ . Since computing the dual cost of  $(P_2)$  is  
 109 expensive, we propose *Stochastic Uzawa* and *Sampled Stochastic Uzawa* algorithms  
 110 relying on Robbins Monroe algorithm in the spirit of the stochastic gradient. Its  
 111 convergence is established, relying on the proof provided by [7] for the convergence of  
 112 the stochastic gradient in a Hilbert space. We check the effectiveness of the *Stochastic*  
 113 *Uzawa* algorithm on a linear quadratic Gaussian framework, and we apply the *Sampled*  
 114 *Stochastic Uzawa* algorithm to a model of power system, inspired by the work of A.  
 115 De Paola *et al.* [5].

116 **2. Approximating the optimization problem.** In this section, the link be-  
 117 tween the values of problems  $(P_1)$  and  $(P_2)$  is analyzed.

118 *Assumption 2.1.* (i) Each set  $\mathcal{U}_i$  is bounded, i.e. there exists  $M > 0$  such  
 119 that  $\mathbb{E}\|u_i\|_{\mathbb{U}_i}^2 \leq M^2$ , for  $i \in \{1, \dots, n\}$ .

120 (ii) The function  $u^i \mapsto G_i(u^i(\cdot, \omega^{-i}), \omega^i)$  is a.s. non negative, convex and l.s.c.

121 (iii) Problem  $(P_1)$  is feasible.

122 From now on, Assumption 2.1 is supposed to hold.

LEMMA 2.2. *Suppose that  $F_0$  is proper, l.s.c. convex. Then Problem  $(P_1)$  has a*  
 123 *solution, i.e.  $J$  reaches its minimum over  $\mathcal{U}$ .*

124 *Proof.* The existence and uniqueness of a minimum is proved by considering a  
 125 minimizing sequence (which exists since  $(P_1)$  is feasible)  $\{u_k\}$  of  $J$  over  $\mathcal{U}$ . The set  
 126  $\mathcal{U}$  being bounded and weakly closed, there exists a subsequence  $\{u_{k_\ell}\}$  which weakly  
 127 converges to a certain  $u^* \in \mathcal{U}$ . Using Assumptions 2.1.(ii) and convexity of  $F_0$ , it  
 128 follows that  $\liminf J(u_{k_\ell}) \geq J(u^*)$  and thus  $u^*$  is a solution of  $(P_1)$ .  $\square$

129 We have the following key result.

130 THEOREM 2.3. *The decentralized problem in the l.h.s. of the following equality*  
 131 *has the same value as the centralized problem in the r.h.s. equality i.e.*

$$132 \quad (2.1) \quad \inf_{u \in \hat{\mathcal{U}}} \tilde{J}(u) = \inf_{u \in \mathcal{U}} \tilde{J}(u).$$

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

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

$$136 \quad \tilde{u}^i(\omega^i) := \mathbb{E}[u^i(\omega^i, \omega^{-i}) | \omega^i] = \int_{\Omega^{-i}} u^i(\omega^i, \omega^{-i}) d\mu^{-i}(\omega^{-i}) \quad \text{for any } \omega^i \in \Omega^i.$$

137 Since  $G_i$  is a.s. convex w.r.t. the first variable, the Jensen inequality gives:

$$138 \quad G_i(\tilde{u}^i, \omega^i) \leq \int_{\Omega^{-i}} G_i(u^i(\cdot, \omega^{-i}), \omega^i) d\mu^{-i}(\omega^{-i}) = \mathbb{E}[G_i(u^i(\cdot, \omega^{-i}), \omega^i) | \omega^i] \quad \text{a.s.}$$

139 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 con-  
 140 ditional expectation, thus:

$$141 \quad F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i)\right) = F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(\tilde{u}^i)\right).$$

142 Taking the expectation of (2), we have  $\inf_{u \in \hat{\mathcal{U}}} \tilde{J}(u) \leq \inf_{u \in \mathcal{U}} \tilde{J}(u)$ , and the conclusion  
 143 follows.  $\square$

144 *Remark 2.4.* In the applications to stochastic control problems (in discrete and  
 145 continuous time) we have the constraint of having progressively measurable control  
 146 policies. Since the set of progressively measurable policies is closed and convex, this  
 147 enters in the above framework. In particular, the decentralized policy  $\tilde{u}^i$  constructed  
 148 in the above proof is progressively measurable.

149 *Remark 2.5.* By Theorem 2.3, for any  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal solution of  
 150 problem  $(P_2)$  that is a decentralized control.

151 **PROPOSITION 2.6.** *If  $F_0$  is Lipschitz with constant  $\gamma$ , then an optimal solution in*  
 152  *$\hat{\mathcal{U}}$  of problem  $(P_2)$  (resp.  $(P_1)$ ) is an  $\varepsilon$ -optimal solution in  $\hat{\mathcal{U}}$  of problem  $(P_1)$  (resp.*  
 153  *$(P_2)$ ), with  $\varepsilon = \gamma M / \sqrt{n}$ .*

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

$$156 \quad (2.2) \quad \hat{u}^i := u^i - \mathbb{E}(u^i).$$

157 Using the Jensen and Hölder inequalities,  $(\mathbb{E}|Y|) \leq (\mathbb{E}|Y|^2)^{\frac{1}{2}}$ , the fact that for any  
 158  $j \neq i$ ,  $u_i$  and  $u_j$  are mutually independent, and that  $\|u_i\|_{\mathbb{U}}$  is bounded a.s. by  $M$ , we  
 159 have  $\forall u \in \hat{\mathcal{U}}$ :

$$160 \quad (2.3) \quad \left| \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) - F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) \right) \right| \leq \mathbb{E} \left| F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) - F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) \right|$$

$$\leq \frac{\gamma}{n} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{\mathbb{U}} \right)$$

$$\leq \frac{\gamma}{n} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{\mathbb{U}}^2 \right)^{\frac{1}{2}}$$

$$\leq \frac{\gamma}{n^{\frac{1}{2}}} M.$$

161 Let  $\tilde{u}^*$  denote a minimizer of  $\tilde{J}$  on  $\hat{\mathcal{U}}$ , then using (2.3) for any  $u \in \hat{\mathcal{U}}$  it holds:

$$162 \quad (2.4) \quad J(\tilde{u}^*) \leq \tilde{J}(\tilde{u}^*) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq \tilde{J}(u) + \frac{\gamma}{n^{\frac{1}{2}}} M \leq J(u) + \frac{\gamma}{n^{\frac{1}{2}}} M. \quad \square$$

163 If  $F_0$  is convex, using Jensen inequality we have for any centralized control  $u \in \mathcal{U}$ :

$$164 \quad (2.5) \quad F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) \leq \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) \right).$$

165 Assumption 2.1.(iii) and convexity of  $F_0$  implies that  $(P_2)$  is feasible. By using the  
 166 same techniques as for Lemma 2.2, one can prove that  $(P_2)$  admits a solution and  
 167 from (2.5) that  $\min_{u \in \hat{\mathcal{U}}} \tilde{J}(u) \leq \min_{u \in \mathcal{U}} J(u)$ , when  $F_0$  is convex.

168 *Assumption 2.7.*  $F_0$  is Gâteaux differentiable with  $c$ -Lipschitz derivative.

169 **THEOREM 2.8.** *Suppose  $F_0$  is convex and Assumption 2.7 holds, then any decen-*  
 170 *tralized optimal solution of problem  $(P_2)$  is an  $\varepsilon$ -optimal solution (where  $\varepsilon = cM^2/n$ )*  
 171 *of problem  $(P_1)$ .*

172 *Remark 2.9.* Observe that the centralized problem  $(P_1)$  on the l.h.s. of the below  
 173 inequality is bounded by the following decentralized problem on the r.h.s of this  
 174 inequality i.e.

$$175 \quad \inf_{u \in \mathcal{U}} J(u) \leq \inf_{u \in \hat{\mathcal{U}}} J(u).$$

176 The article by [3] proposes an upper bound for the decentralized problem and a lower  
 177 bound for the centralized problem. The upper bound is provided by a resource decom-  
 178 position approach (with deterministic quantities) while the lower bound is provided  
 179 by a price decomposition approach with deterministic prices (see equation (28) of [3]).  
 180 Theorem 2.8 provides an upper bound for problem  $(P_1)$  with an a priori quantification  
 181 of the deviation from the optimal value which vanishes when the number of agents  
 182 grows to infinity. Moreover, in Section 4 we provide an original algorithm that allows  
 183 to approach the solution of the decentralized problem.

184 *Proof.* Since  $F_0$  is convex, differentiable, with a  $c$ -Lipschitz differential, one can  
 185 derive for any  $u \in \hat{\mathcal{U}}$  and a.s.:

$$\begin{aligned} & F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right) \\ & \leq \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right), \sum_{i=1}^n \hat{u}^i \rangle_{\mathbb{U}} \\ 186 \quad (2.6) \quad & = \frac{1}{n} \langle (\nabla F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right)), \sum_{i=1}^n \hat{u}^i \rangle_{\mathbb{U}} \\ & \quad + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right), \sum_{i=1}^n \hat{u}^i \rangle_{\mathbb{U}} \\ & \leq \frac{c}{n^2} \left\| \sum_{i=1}^n \hat{u}^i \right\|_{\mathbb{U}}^2 + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right), \sum_{i=1}^n \hat{u}^i \rangle_{\mathbb{U}}, \end{aligned}$$

where  $\hat{u}^i$  is defined in (2.2). Taking the expectation of (2.6),

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

187 and using the mutual independence of the controls and their boundedness we get as  
 188 in (2.3):

$$189 \quad (2.7) \quad \frac{c}{n^2} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{\mathbb{U}}^2 \right) \leq \frac{c}{n} M^2.$$

190 Let  $\tilde{u}^*$  denote a minimizer of  $\tilde{J}$  on  $\hat{\mathcal{U}}$ , then using (2.1), (2.7) and (2.5), for any  $u \in \mathcal{U}$   
 191 we have:

$$192 \quad (2.8) \quad J(\tilde{u}^*) \leq \tilde{J}(\tilde{u}^*) + \frac{c}{n} M^2 \leq \tilde{J}(u) + \frac{c}{n} M^2 \leq J(u) + \frac{c}{n} M^2.$$

193 Thus for  $\varepsilon = cM^2/n$ ,  $\tilde{u}^*$  constitutes an  $\varepsilon$ -optimal solution to the stochastic control  
 194 problem  $(P_1)$ .  $\square$

195 *Remark 2.10.* Let  $\tilde{u}^*$  and  $u^*$  be respectively the optimal controls of problems  $(P_2)$   
 196 and  $(P_1)$ . From Jensen inequality and by definition of  $\tilde{u}^*$  we have:

$$197 \quad J(u^*) \geq \tilde{J}(u^*) \geq \tilde{J}(\tilde{u}^*).$$

198 Adding  $J(\tilde{u}^*)$ , one has:

$$199 \quad (2.9) \quad J(\tilde{u}^*) - \tilde{J}(\tilde{u}^*) \geq J(\tilde{u}^*) - J(u^*) \geq 0.$$

200 An approximation scheme to compute  $\tilde{u}^*$  is provided in Section 4. The practical  
 201 interest of inequality (2.9) is that one can compute an upper bound for the error  
 202  $J(\tilde{u}^*) - J(u^*)$ , that can be automatically derived from this approximation.

203 **3. Dualization and Decentralization of problem  $(P_2)$ .** From now on, the  
 204 assumption that  $F_0$  is convex is in force in the sequel. The *Lagrangian* function  
 205 associated to the constrained optimization problem  $(P'_2)$ , defined in (1.4), is:  $L : \hat{\mathcal{U}} \times \mathbb{U} \times \mathbb{U} \rightarrow \bar{\mathbb{R}}$  defined by:

$$207 \quad (3.1) \quad L(u, v, \lambda) := \bar{J}(u, v) + \langle \lambda, \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) - v \rangle_{\mathbb{U}}.$$

208 The dual problem  $(D)$  associated with  $(P'_2)$  is:

$$209 \quad (3.2) \quad (D) \quad \max_{\lambda \in \mathbb{U}} \mathcal{W}(\lambda), \quad \text{where } \mathcal{W}(\lambda) := \text{Min}_{u \in \hat{\mathcal{U}}, v \in \mathbb{U}} L(u, v, \lambda).$$

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

$$211 \quad (3.3) \quad \mathcal{W}(\lambda) = -F_0^*(\lambda) + \frac{1}{n} \sum_{i=1}^n \text{Min}_{u^i \in \hat{\mathcal{U}}_i} \mathbb{E}(G_i(u^i, \omega^i)) + \langle \lambda, \mathbb{E}(u^i) \rangle_{\mathbb{U}},$$

212 where  $F_0^*(\lambda) := \sup_{v \in \mathbb{U}} \langle \lambda, v \rangle_{\mathbb{U}} - F_0(v)$ .

213 The problem is said to be qualified if it is still feasible after a small perturbation  
 214 of the constraint, in the following sense:

$$215 \quad (3.4) \quad \text{There exists } \varepsilon > 0 \text{ such that } \mathcal{B}_{\mathbb{U}}(0, \varepsilon) \subset g(\hat{\mathcal{U}}, \mathbb{U}),$$

216 where  $\mathcal{B}_{\mathbb{U}}(0, \varepsilon)$  is the open ball of radius  $\varepsilon$  in  $\mathbb{U}$ ,  $g$  has been defined in (1.4) and  $g(\hat{\mathcal{U}}, \mathbb{U})$   
 217 is the image by  $g$  of  $\hat{\mathcal{U}} \times \mathbb{U}$ .

218 **LEMMA 3.1.** *Problem  $(P'_2)$  is qualified.*

219 *Proof.* By Assumption 2.1.(iii), there exists  $\hat{u}$  feasible for problem  $(P_1)$ . Then

$$220 \quad (3.5) \quad \mathcal{B}_{\mathbb{U}}(0, \varepsilon) \subset \mathbb{U} = g(\hat{u}, \mathbb{U}) \subset g(\hat{\mathcal{U}}, \mathbb{U}).$$

221 The conclusion follows.  $\square$

By Assumption 5.2, Lemma 3.1 and the convexity of  $F_0$ , the strong duality holds. Let us denote the set of solutions of the dual problem by  $S$ . Since the primal problem is qualified, the primal and dual values are equal, and the set of dual solutions  $S$  is nonempty and bounded. In addition, taking  $\lambda^* \in S$ , any primal solution  $u^*$  satisfies both  $W(\lambda^*) = \tilde{J}(u^*)$  and  $(u^*, v^*) \in \arg \min_{u \in \hat{\mathcal{U}}, v \in \mathbb{U}} L(\lambda^*, u, v)$ . 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 the first variable, then each component  $u^{*i}$  can be uniquely determined by solving the following sub problem:

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

222 *Remark 3.2.* By using the same argument as in Theorem 2.3, one can prove:

$$\begin{aligned} 223 \quad (3.6) \quad & \min_{u^i \in \hat{\mathcal{U}}_i} \{ \mathbb{E} (G_i(u^i, \omega^i) + \langle \lambda^*, u^i \rangle_{\mathbb{U}}) \} \\ & = \min_{u^i \in \mathcal{U}_i} \{ \mathbb{E} (G_i(u^i(\cdot, \omega^{-i}), \omega^i) + \langle \lambda^*, u^i \rangle_{\mathbb{U}}) \}. \end{aligned}$$

224 **4. Stochastic Uzawa and Sampled Stochastic Uzawa algorithms.** We  
225 recall that Assumption 2.1 is in force, as well as convexity of  $F_0$ .

226 This section aims at proposing an algorithm to find a solution of the dual problem  
227 (3.2).

228 *Assumption 4.1.* (i) The function  $u^i \mapsto G_i(u^i, \omega^i)$  is for a.a.  $\omega^i \in \Omega^i$  strictly  
229 convex on  $\hat{\mathcal{U}}_i$ .

230 (ii) The function  $F_0$  has quadratic growth.

231 For all  $i \in \{1, \dots, n\}$ , and  $\lambda \in \mathbb{U}$ , we define the optimal control  $u^i(\lambda)$ :

$$232 \quad (4.1) \quad u^i(\lambda) := \arg \min_{u^i \in \hat{\mathcal{U}}_i} \{ \mathbb{E} (G_i(u^i, \omega^i) + \langle \lambda, u^i \rangle_{\mathbb{U}}) \},$$

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

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

$$235 \quad (4.2) \quad V(\lambda) := \arg \min_{v \in \mathbb{U}} \{ F_0(v) - \langle \lambda, v \rangle_{\mathbb{U}} \}.$$

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

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

244 For any real valued function  $F$  defined on  $\mathbb{U}$ ,  $F^*$  stands for its Fenchel conjugate.

245 **LEMMA 4.2.** *Assumption 2.7 holds iff  $F_0^*$  is proper and strongly convex.*

246 *Proof.* (i) Let Assumption 2.7 hold. Since  $F_0$  is proper, convex and l.s.c.,  $F_0^*$   
247 is l.s.c. proper. From the Lipschitz property of the gradient of  $F_0$ , it holds that  
248  $\text{dom}(F_0) = \mathbb{U}$ .

249 Let  $s, \tilde{s} \in \text{dom}(F_0^*)$  such that there exist  $\lambda_s \in \partial F_0^*(s)$  and  $\mu_{\tilde{s}} \in \partial F_0^*(\tilde{s})$ . From  
250 the differentiability, l.s.c. and convexity of  $F_0$ , it follows that:  $s = \nabla F_0(\lambda_s)$  and  
251  $\tilde{s} = \nabla F_0(\mu_{\tilde{s}})$ . By Assumption 2.7 and the extended Baillon-Haddad theorem [17,

252 Theorem 3.1],  $\nabla F_0$  is cocoercive. In other words:

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

254 Therefore  $\partial F_0^*$  is strongly monotone, which implies the strong convexity of  $F_0^*$ .

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

$$(4.4) \quad F_0(\lambda) \leq \sup_{s \in \mathbb{U}} \langle s, \lambda - \gamma \rangle_{\mathbb{U}} - \alpha \|s\|_{\mathbb{U}}^2 + \beta = \|\lambda - \gamma\|^2 / (4\alpha) + \beta.$$

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

262 Let  $\lambda, \mu \in \mathbb{U}$ . Since  $F_0^*$  is strongly convex, the functions  $F_0^*(s) - \langle \lambda, s \rangle_{\mathbb{U}}$  (resp.  
 263  $F_0^*(s) - \langle \mu, s \rangle_{\mathbb{U}}$ ) has a unique minimum point  $s_\lambda$  (resp.  $s_\mu$ ), characterized by:  $\lambda \in$   
 264  $\partial F_0^*(s_\lambda)$  and  $\mu \in \partial F_0^*(s_\mu)$ . From the strong convexity of  $F_0^*$ , the strong mono-  
 265 tonicity 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  
 266 related to the strong convexity of  $F_0^*$ . Using that  $s_\lambda = \nabla F_0(\lambda)$  and  $s_\mu = \nabla F_0(\mu)$ , it  
 267 holds:

$$(4.5) \quad \langle \mu - \lambda, \nabla F_0(\mu) - \nabla F_0(\lambda) \rangle_{L^2(0,T)} \geq \frac{1}{c} \|\nabla F_0(\mu) - \nabla F_0(\lambda)\|_{L^2(0,T)}^2,$$

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

271 LEMMA 4.3. *If Assumption 2.7 holds, then  $\mathcal{W}$  is strongly concave.*

272 *Proof.* For any  $\lambda \in \mathbb{U}$ , the expression of  $\mathcal{W}(\lambda)$  is given by 3.3, where for any  
 273  $i \in \{1, \dots, n\}$ ,  $\lambda \mapsto \inf_{u^i \in \tilde{\mathcal{U}}_i} \mathbb{E}(G_i(u^i, \omega^i)) + \langle \lambda, E(u^i) \rangle_{\mathbb{U}}$  is concave and from Lemma  
 274 4.2  $-F_0^*$  is strongly concave. Since the sum of a concave function and of a strongly  
 275 concave function is strongly concave, the result follows.  $\square$

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

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

278 Since  $F_0$  has at least quadratic growth, one deduces that  $F_0^*$  has at most quadratic  
 279 growth. Using the boundedness of  $\mathcal{U}$  and Lemma A.1 in Appendix A, there exist  
 280  $M_1, M_2 > 0$  such that for any  $\lambda \in \mathbb{U}$  one has:

$$(4.7) \quad \|f(\lambda)\|_{\mathbb{U}}^2 \leq M_1 + M_2 \|\lambda\|_{\mathbb{U}}^2.$$

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

$$(4.8) \quad \partial(-\mathcal{W}(\lambda)) \ni -f(\lambda).$$

285 The iterative algorithm, proposed as an approximation scheme for  $\lambda^* \in \arg \max_{\lambda} \mathcal{W}(\lambda)$ ,  
 286 is summarized in the *Stochastic Uzawa* Algorithm 4.1. Some assumptions on the step  
 287 size are introduced.

288 *Assumption 4.4.* The sequence  $(\rho_k)_k$  is such that:  $\rho_k > 0$ ,  $\sum_{k=1}^{\infty} \rho_k = \infty$  and  
 289  $\sum_{k=1}^{\infty} (\rho_k)^2 < \infty$ .

290 Note that a sequence of the form  $\rho_k := \frac{a}{b+k}$ , with  $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}_+$ , satisfies  
 Assumption 4.4.

---

**Algorithm 4.1** Stochastic Uzawa
 

---

- 1: Initialization  $\lambda^0 \in \mathbb{U}$ , set  $\{\rho_k\}$  satisfying Assumption 4.4.
  - 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 (4.2).
  - 5:    $u^{i,k} \leftarrow u^i(\lambda^k)$  where  $u^i(\lambda^k)$  is defined in (4.1) 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  $(u^1(\lambda^k)(\omega^{1,k+1}), \dots, u^n(\lambda^k)(\omega^{n,k+1}))$ .
  - 8:    $Y^{k+1} \leftarrow \frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i,k+1}) - v(\lambda^k)$ .
  - 9:    $\lambda^{k+1} \leftarrow \lambda^k + \rho_k Y^{k+1}$ .
- 

291 At any dual iteration  $k$  of Algorithm 4.1,  $Y^{k+1}$  is an estimator of  
 292  $\mathbb{E}(\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i,k+1}) - v(\lambda^k))$ . Therefore an alternative approach proposed in the  
 293 *Sampled Stochastic Uzawa* Algorithm 4.2 consists in performing less simulations at  
 294 each iteration, by taking  $m < n$ , at the risk of performing more dual iterations, to  
 295 estimate the quantity  $\mathbb{E}(\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(\omega^{i,k+1}) - v(\lambda^k))$ .

297 The complexity of the *Stochastic Uzawa* Algorithm 4.2 is proportional to  $m \times K$ ,  
 298 where  $K$  is the total number of dual iterations and  $m$  the number of simulations  
 299 performed at each iteration. The error  $\mathbb{E}(\|\lambda^{k+1} - \lambda^*\|_{\mathbb{U}}^2)$  for  $\lambda^* \in S$  is the sum of  
 300 the square of the bias (which only depends on  $K$  and not on  $m$ ) and the variance  
 301 (which both depends on  $K$  and  $m$ ). Therefore this algorithm enables a bias variance  
 302 trade-off for a given complexity. Similarly for a given error it enables to optimize the  
 303 complexity of the algorithm.

304 We recall that  $S$  is defined by  $S := \arg \max_{\lambda \in \mathbb{U}} \mathcal{W}(\lambda)$  and that  $S$  is non empty due  
 305 to strong duality. The following result establishes the convergence of the *Stochastic*  
 306 *Uzawa* Algorithm 4.1:

---

**Algorithm 4.2** Sampled Stochastic Uzawa
 

---

- 1: Initialization of  $m$  a positive integer and  $\check{\lambda}^0 \in \mathbb{U}$ , set  $\{\rho_k\}$  satisfying Assumption 4.4.
  - 2:  $k \leftarrow 0$ .
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:      $v^k \leftarrow v(\check{\lambda}^k)$  where  $v(\check{\lambda}^k) \in V(\check{\lambda}^k)$ , this set being defined in (4.2).
  - 5:     Generate  $m$  i.i.d. discrete random variables  $I_1^k, \dots, I_m^k$  uniformly in  $\{1, \dots, n\}$ .
  - 6:      $u_j^{I_j^k, k} \leftarrow u_j^{I_j^k}(\check{\lambda}^k)$  where  $u_j^{I_j^k}(\check{\lambda}^k)$  is defined in (4.1) for any  $j \in \{1, \dots, m\}$ .
  - 7:     Generate  $m$  independent noises  $(\omega^{1, k+1}, \dots, W^{m, k+1})$ , independent also of  $\{\omega^{i, p} : 1 \leq i \leq m, p \leq k\}$ .
  - 8:     Compute the associated control realization  $(u^{I_1^k}(\check{\lambda}^k)(\omega^{1, k+1}), \dots, u^{I_m^k}(\check{\lambda}^k)(\omega^{m, k+1}))$ .
  - 9:      $\check{Y}^{k+1} \leftarrow \frac{1}{m} \sum_{j=1}^m u_j^{I_j^k}(\check{\lambda}^k)(\omega^{I_j^k, k+1}) - v(\check{\lambda}^k)$
  - 10:     $\check{\lambda}^{k+1} \leftarrow \check{\lambda}^k + \rho_k \check{Y}^{k+1}$ .
- 

LEMMA 4.5. *Let Assumption 4.4 hold, 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 2.7 holds, then a.s.  $\{\lambda^k\}$  converges to  $\bar{\lambda}$  in  $\mathbb{U}$ , with  $S := \{\bar{\lambda}\}$ .

The proof follows [7, Theorem 3.6]. That reference considers (changing minimization in maximization) the framework of maximization a function  $\mathcal{W}(\lambda) = \mathbb{E}(\mathbf{W}(\lambda, \omega))$  where  $\mathbf{W}(\cdot, \omega)$  is a.s. concave. Although our setting does not enter in this framework, due to the minimization of the Lagrangian w.r.t. the variable  $u$ , the proof of Lemma 4.5 follows from an obvious adaptation of the one in [7, Theorem 3.6]. It is enough to provide the first steps of the proof.

*Proof of Lemma 4.5.* First consider point (i). Let  $\lambda \in S$ . For any  $k$ ,  $\mathcal{G}_{k+1}$  is the filtration defined by:

$$(4.9) \quad \mathcal{G}_{k+1} := \sigma(\{W^{i, p}\} : 1 \leq i \leq n, p \leq k+1).$$

Using the definition of  $Y^{k+1} \in \mathbb{U}$  line 8 in the *Stochastic Uzawa* Algorithm 4.1, we have:

$$(4.10) \quad \begin{aligned} \|\lambda^{k+1} - \lambda\|_{\mathbb{U}}^2 &= \|\lambda^k + \rho_k Y^{k+1} - \lambda\|_{\mathbb{U}}^2 \\ &= \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \langle \lambda^k - \lambda, Y^{k+1} \rangle_{\mathbb{U}} \\ &\quad + (\rho_k)^2 \|Y^{k+1}\|_{\mathbb{U}}^2. \end{aligned}$$

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

$$(4.11) \quad \mathbb{E}(\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k) = \mathbb{E}\left(\left\|\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(W^{i, k+1}) - v(\lambda^k)\right\|_{\mathbb{U}}^2\right).$$

Using previous equality and the inequality (4.7), one can easily show that there exists  $M_3, M_4 > 0$  such that for any  $k \in \mathbb{N}$  one has:

$$(4.12) \quad \mathbb{E}(\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k) \leq M_1 + M_2 \|\lambda^k\|_{\mathbb{U}}^2 \leq M_3 + M_4 \|\lambda^k - \lambda\|_{\mathbb{U}}^2$$

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

$$\begin{aligned}
& \mathbb{E}[\|\lambda^{k+1} - \lambda\|_{\mathbb{U}}^2 | \mathcal{G}_k] \\
&= \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \mathbb{E}(\langle \lambda^k - \lambda, Y^{k+1} \rangle_{\mathbb{U}} | \mathcal{G}_k) + (\rho_k)^2 \mathbb{E}[\|Y^{k+1}\|_{\mathbb{U}}^2 | \mathcal{G}_k] \\
326 \quad (4.13) \quad &\leq \|\lambda^k - \lambda\|_{\mathbb{U}}^2 + 2\rho_k \langle \lambda^k - \lambda, f(\lambda^k) \rangle_{\mathbb{U}} + (\rho_k)^2 (M_3 + M_4 \|\lambda^k - \lambda\|_{\mathbb{U}}^2) \\
&\leq \|\lambda^k - \lambda\|_{\mathbb{U}}^2 (1 + M_4 \rho_k^2) + (\rho_k)^2 M_3 - 2\rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)).
\end{aligned}$$

327 In the last inequality, we used the concavity of  $\mathcal{W}$  and (4.8). The rest of the proof  
328 follows [7, Theorem 3.6].  $\square$

329 Recalling the definition of  $\bar{J}(u, v)$  in (1.4), we define  $\bar{u}$ :

$$330 \quad (4.14) \quad \bar{u} := \arg \min_{u \in \mathcal{U}} \left\{ \mathbb{E} \left( \sum_{i=1}^n G_i(u^i, \omega^i) + \langle \bar{\lambda}, u^i \rangle_{\mathbb{U}} \right) \right\}.$$

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

$$333 \quad (4.15) \quad \bar{v} := \arg \min_{v \in \mathbb{U}} \{F_0(v) + \langle \bar{\lambda}, v \rangle_{\mathbb{U}}\}.$$

334

335 **THEOREM 4.6.** *Let the Assumptions 2.7 and 4.4 hold, then we have:*

- 336 (i)  $\{u(\lambda^k)\}$  weakly converges a.s. to  $\bar{u}$ .  
337 If furthermore  $F_0$  is strictly convex, then  $(\bar{u}, \bar{v}, \bar{\lambda})$  is the unique saddle point  $\mathcal{L}$ , there-  
338 fore  $\bar{u}$  is the unique minimizer of  $\bar{J}$  in  $\mathcal{U}$  and:  
339 (ii)  $\bar{J}(u(\lambda^k)) \xrightarrow[k \rightarrow \infty]{} \bar{J}(\bar{u})$  a.s.  
340 (iii)  $\limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \inf_{u \in \mathcal{U}} J(u) + 2\varepsilon$  a.s. where  $\varepsilon = cM^2/n$ .

341 *Proof.* Proof of point (i). Since the sequence  $\{(u(\lambda^k), v(\lambda^k))\}$  is bounded in  $\mathbb{U} \times$   
342  $L^2(0, T)$ , there exists a weakly convergent subsequence  $\{(u(\lambda^{\theta_k}), v(\lambda^{\theta_k}))\}$  such that:

$$343 \quad (4.16) \quad (u(\lambda^{\theta_k}), v(\lambda^{\theta_k})) \rightharpoonup (u^\theta, v^\theta) \in \mathcal{U} \times \mathbb{U}.$$

344 Using the definition of  $\lambda \mapsto u(\lambda)$  in (4.1), it holds for any  $k > 0$ :

$$\begin{aligned}
345 \quad (4.17) \quad & \mathbb{E} (G_i(\bar{u}^i, \omega^i) + \langle \lambda^{\theta_k}, \bar{u}^i \rangle_{\mathbb{U}}) \\
& \geq \mathbb{E} (G_i(u^i(\lambda^{\theta_k}), \omega^i) + \langle \lambda^{\theta_k}, u^i(\lambda^{\theta_k}) \rangle_{\mathbb{U}}).
\end{aligned}$$

346 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  $\{\lambda^k\}$ ,  
347 resulting from Lemma 4.5.(iv), we have from (4.17) when  $k \rightarrow \infty$  :

$$348 \quad (4.18) \quad \mathbb{E} (G_i(\bar{u}^i, \omega^i) + \langle \bar{\lambda}, \bar{u}^i \rangle_{\mathbb{U}}) \geq \mathbb{E} (G_i(u^{i,\theta}, \omega^i) + \langle \bar{\lambda}, u^{i,\theta} \rangle_{\mathbb{U}}).$$

349 Since  $\bar{u}$  is unique, it follows  $u^\theta = \bar{u}$  and (4.18) is an equality. Using that every weakly  
350 convergent subsequence of  $\{u(\lambda^k)\}$  has the same weak limit  $\bar{u}$ , (i) is deduced.

351 *Proof of point (ii).*

352 From point (i) and (4.18), it follows for any  $i \in \{1, \dots, n\}$ :

$$353 \quad (4.19) \quad \lim_{k \rightarrow \infty} \mathbb{E} (G_i(u^i(\lambda^k), \omega^i)) = \mathbb{E} (G_i(\bar{u}^i, \omega^i)).$$

354 Using 4.16, the w.l.s.c. of  $F_0$ , equation (4.15), and applying the same previous argu-  
 355 ment to  $\{v(\lambda^{\theta^k})\}$ , it holds that:

$$356 \quad (4.20) \quad \lim_{k \rightarrow \infty} F_0(v(\lambda^k)) - \langle \lambda^k, v(\lambda^k) \rangle_{\mathbb{U}} = F_0(\bar{v}) - \langle \bar{\lambda}, \bar{v} \rangle_{\mathbb{U}},$$

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

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

$$359 \quad (4.21) \quad \lim_{k \rightarrow \infty} F_0(v(\lambda^k)) = F_0(\bar{v}).$$

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

$$361 \quad (4.22) \quad \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\bar{u}^i) = \bar{v}.$$

362 From (4.21) and (4.22), it holds:

$$363 \quad (4.23) \quad \lim_{k \rightarrow \infty} F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i(\lambda^k)) \right) = F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\bar{u}^i) \right).$$

364 Then adding (4.19) and (4.23):  $\lim_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) = \tilde{J}(\bar{u})$ .

365 Proof of point (iii). From point (ii), inequality (2.8) and Theorem 2.8, it holds:

$$366 \quad (4.24) \quad \limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \limsup_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) + \varepsilon = \inf_{u \in \mathcal{U}} \tilde{J}(u) + \varepsilon \leq \inf_{u \in \mathcal{U}} J(u) + 2\varepsilon. \quad \square$$

367 *Assumption 4.7.* (i)  $F_0$  is strongly convex.

368 (ii) For any  $i \in \{1, \dots, n\}$  and  $\omega \in \Omega$ , the function  $u^i \mapsto \mathbb{E}(G_i(u^i, \omega))$  is strongly  
 369 convex.

370 LEMMA 4.8. *Let Assumption 4.7.(i) hold, then the function  $\lambda \mapsto v(\lambda)$  is Lipschitz on  $\mathbb{U}$ .*

371 *Proof.* From the definition of  $v$  in (4.2), we have for any  $\lambda \in \mathbb{U}$ :  $\lambda \in \partial F_0(v(\lambda))$ .

372 Thus for any  $\lambda, \mu \in \mathbb{U}$ , we have from the strong convexity of  $F_0$ :

$$373 \quad (4.25) \quad \begin{cases} F_0(v(\mu)) & \geq F_0(v(\lambda)) + \langle \lambda, v(\mu) - v(\lambda) \rangle_{\mathbb{U}} + \alpha \|v(\mu) - v(\lambda)\|_{\mathbb{U}}^2 \\ F_0(v(\lambda)) & \geq F_0(v(\mu)) + \langle \mu, v(\lambda) - v(\mu) \rangle_{\mathbb{U}} + \alpha \|v(\lambda) - v(\mu)\|_{\mathbb{U}}^2. \end{cases}$$

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

$$375 \quad (4.26) \quad \langle \lambda - \mu, v(\lambda) - v(\mu) \rangle_{\mathbb{U}} \geq 2\alpha \|v(\lambda) - v(\mu)\|_{\mathbb{U}}^2.$$

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

378 LEMMA 4.9. *Let Assumption 4.7.(ii) hold, thus the function  $\lambda \mapsto u(\lambda)$  is Lipschitz on  $\mathbb{U}$ .*

379 *Proof.* The proof is similar to the proof of Lemma 4.8.  $\square$

380 THEOREM 4.10. *Let the Assumption 2.7, 4.4, and 4.7 hold, then:  $u(\lambda^k) \xrightarrow[k \rightarrow \infty]{} u(\bar{\lambda})$  a.s.*

381  $u(\bar{\lambda})$  a.s.

382 *Proof.* The convergence follows from the Lipschitz property of  $\lambda \mapsto u(\lambda)$  (as a  
383 result of assumption 4.7) associated with the a.s. convergence of  $\{\lambda^k\}$ .  $\square$

384 *Remark 4.11.* Note that Lemma 4.5 and Theorems 4.6 and 4.10 still hold when  
385 replacing  $\lambda^k$  by  $\check{\lambda}^k$  and  $Y^k$  by  $\check{Y}^k$  (defined resp. line 9 and 10 in the *Sampled Stochastic*  
386 *Uzawa* Algorithm 4.2). This can be proved by same argument, using that  $\check{Y}^k$  is  
387 bounded a.s. and  $\mathbb{E}(\check{Y}^k | \check{\mathcal{G}}_k) = f(\check{\lambda}^k)$  for any  $k$ , where:

$$388 \quad (4.27) \quad \check{\mathcal{G}}_k = \sigma \left( \{W^{\ell,p}\} : 1 \leq \ell \leq m, p \leq k \right) \vee \sigma \left( \{I_\ell^p\} : 1 \leq \ell \leq m, p \leq k \right),$$

389 with  $W^{\ell,p}$  and  $I_\ell^k$  defined respectively at lines 7 and 5 of the *Sampled Stochastic*  
390 *Uzawa* Algorithm 4.2.

391 *Remark 4.12.* From a practical point of view, this algorithm can be implemented  
392 in a decentralized way, where the system operator sends the signal  $\lambda$ , which can  
393 be assimilated to a price, to the domestic appliances, which compute their optimal  
394 solution  $u(\lambda)$ , depending on their local parameters.

395 In (5.2), the states and controls of the agents are described in a general framework.  
396 To illustrate the results, we consider in the next section stochastic control problems  
397 in both continuous and discrete time settings.

## 398 5. Application to stochastic control.

399 **5.1. Continuous time setting.** Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete filtered proba-  
400 bility space on which  $W = (W^i)_{i=1,\dots,n}$  is a  $n \times d$ -dimensional Brownian motion,  
401 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  
402 the filtration  $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ .  $\mathbb{P}$  stands for the Wiener measure associated with this  
403 filtration and  $\mathbb{F}$  for the augmented filtration by all  $\mathbb{P}$ -null sets. The following notations  
404 are used:

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

$$405 \quad \mathbb{U} := L^2((0, T), \mathbb{R}^p) := \{ \varphi : [0, T] \rightarrow \mathbb{R}^p \mid \int_0^T \sum_{k=1}^p |\varphi_k(t)|^2 dt < \infty \},$$

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

$$407 \quad (5.1) \quad \mathcal{U}_i := \{ v : \Omega \times [0, T] \rightarrow \mathbb{R}, v(\cdot) \text{ is } \mathbb{F}\text{-prog. measurable,} \\ v(\omega) \in \mathbb{U} \text{ and } v_t(\omega) \in [-M_i, M_i]^p, \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega \},$$

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

409 Each local agent  $i = 1, \dots, n$  is supposed to control its state variable through the  
410 control process  $u^i \in \mathcal{U}_i$  and suffers from independent uncertainties. More specifically,  
411 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  
412  $\mathbb{R}^d$  and follows the dynamics for  $i \in \{1, \dots, n\}$ :

$$413 \quad (5.2) \quad \begin{cases} dX_t^{i, u^i} &= \mu_i(t, u_t^i(\cdot), W^i), X_t^{i, u^i}) dt + \sigma_i(t, X_t^{i, u^i}) dW_t^i, \text{ for } t \in [0, T], \\ X_0^{i, u^i} &= x_0^i \in \mathbb{R}^d; \end{cases}$$

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

$$417 \quad (5.3) \quad \mu_i(t, \nu, x) = \alpha_i(t)\nu + \beta_i(t)x + \gamma_i(t) \quad \text{and} \quad \sigma_i(x, t) = \xi_i(t)x + \theta_i(t).$$

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

421 **5.1.1. On the well-posedness of  $(P_1)$ .** In this section, the assumptions needed  
 422 for  $(P_1)$  to be well posed are studied.

LEMMA 5.1. *Let  $i \in \{1, \dots, n\}$  and  $v \in \mathcal{U}_i$  be a control process. The map  $v^i \mapsto X^{i, v}$  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  $p \in [1, \infty)$ :*

$$423 \quad (5.4) \quad \mathbb{E} \left( \sup_{\substack{0 \leq t \leq T \\ 1 \leq k \leq d}} |X_{k,t}^{i, v}|^r \right) < C(r, T, x_0, K) < \infty .$$

424 *Proof.* The proof for the existence and uniqueness of a solution of (5.2) relies on  
 425 [14, Theorem 3.6, Chapter 2]. The inequality is a result of [14, Theorem 4.4, Chapter  
 426 2].  $\square$

427 Let  $F_0 : \mathbb{U} \rightarrow \mathbb{R}$  be proper, convex and lower semi continuous function, satisfying  
 428 Assumptions 2.7 and 4.1.(ii). For any  $i \in \{1, \dots, n\}$ , we assume that there exists  $F_i$   
 429 such that the local cost  $G_i$  is of the form:

$$430 \quad (5.5) \quad u^i \mapsto G_i(u^i(\cdot, \omega^{-i}), \omega^i) = F_i(u^i(\omega^i, \omega^{-i}), X^{i, u^i}(\omega^i)),$$

431 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.  
 432 Additional assumptions are formulated below.

433 *Assumption 5.2.* For any  $i \in \{1, \dots, n\}$ :

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

439 *Remark 5.3.* 1. Assumption 5.2.(i) is satisfied if there exist  
 440  $g_i : L^2((0, T), \mathbb{R}^p) \rightarrow \mathbb{R}$  strictly convex and  $h_i : \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  convex,  
 441 such that  $F_i(v, X) = g_i(v) + h_i(X)$ .

442 2. Observe that Assumption 5.2 satisfies Assumptions 2.1.(ii) and 4.1.(i)

443 From now on, Assumption 5.2 is in force in the sequel. Now the optimization  
 444 problems  $(P_1^c)$  and  $(P_2^c)$  can be clearly defined:

$$445 \quad (5.6) \quad (P_1^c) \quad \begin{cases} \inf_{u \in \mathcal{U}} J^c(u) \\ J^c(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i(\omega) \right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i(\omega), X^{i, u^i}(\omega^i)) \right), \end{cases}$$



446 and

$$447 \quad (5.7) \quad (P_2^c) \quad \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}^c(u) \\ \tilde{J}^c(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) + \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n F_i(u^i(\omega), X^{i,u^i}(\omega^i)) \right), \end{cases}$$

448 Using the results of Section 2, we can state the following Corollary.

449 **COROLLARY 5.4.** (i) Problems  $(P_1^c)$  and  $(P_2^c)$  admit both a unique solution.  
 450 (ii) Any optimal solution of problem  $(P_2^c)$  is an  $\varepsilon$ -optimal solution, where  $\varepsilon =$   
 451  $cM^2/n$ , of problem  $(P_1^c)$ .

452 *Proof.* The proof of point (i) is a specific case of Lemma 2.2. Similarly, point (ii)  
 453 is a particular case of Theorem 2.8.  $\square$

454 *Remark 5.5.* This kind of stochastic optimization problem is illustrated in Section  
 455 7 with a problem of coordination of a large population of domestic appliances, where  
 456 a system operator has to meet the demand while producing at low cost.

457 **5.2. Discrete time setting.** The main results of the paper are instantiated to  
 458 the discrete time setting in this subsection. The following notations are used:

- 459 • Let  $n \in \mathbb{N}^*$  be the number of agents,  $d, p \in \mathbb{N}^*$  the dimension respectively of  
 460 their state and control variables at any time step, and  $T \in \mathbb{N}^*$  the finite time  
 461 horizon.
- 462 • For any matrix  $M$ ,  $M^\top$  denotes its transpose
- 463 • We consider a global noise process as a sequence of independent random  
 464 variables  $(W_1, \dots, W_T)$ , where for any  $t \in \{1, \dots, T\}$ ,  $W_t$  is a vector of  $d$ -  
 465 dimensional centered, reduced and independent Gaussian variables, defined  
 466 on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ :  $W_t := (W_t^1, \dots, W_t^n)$ , with  $W_t^i \in \mathbb{R}^d$ . For  
 467 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  
 468  $\mathcal{F}_t := \otimes_{i=1}^n \mathcal{F}_t^i$ .
- 469 • The space  $\mathbb{X}$  is defined by:

$$470 \quad (5.8) \quad \mathbb{X} := \{x = (x_0, \dots, x_T) \mid \forall k \in \{0, \dots, T\}, \mathbb{R}^d \ni x_k \text{ is} \\ \mathcal{F}_k \text{-measurable and } \mathbb{E}\|x_k\|_2^2 < \infty\}.$$

- 471 • For any  $i \in \{1, \dots, n\}$ ,  $X^{i,u^i} := (x_0^i, \dots, x_T^i) \in \mathbb{X}$  is the state trajectory  
 472 of agent  $i$  controlled by  $u^i := (u_0^i, \dots, u_{T-1}^i) \in \mathbb{R}^{p \times T}$ . Similarly, for any  
 473  $t \in \{0, \dots, T\}$   $X_t^u := (x_t^1, \dots, x_t^n) \in \mathbb{R}^{d \times n}$  is the state vector of all the agents  
 474 controlled by  $u_j := (u_t^1, \dots, u_t^n) \in \mathbb{R}^{p \times n}$ . We have the following dynamics:

$$475 \quad (5.9) \quad \begin{cases} X_{t+1}^{i,u^i} &= A^i X_t^{i,u^i} + B^i u_t^i + C^i W_{t+1}^i, \quad \text{for } t \in \{0, \dots, T-1\}, \\ X_0^{i,u^i} &= x_0 \in \mathbb{R}^d, \end{cases}$$

476 where  $A \in \mathbb{R}^{d \times d}$ ,  $B \in \mathbb{R}^{d \times p}$  and  $C \in \mathbb{R}_+^{d \times d}$ .

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

$$478 \quad (5.10) \quad \mathcal{U}^i := \{u = (u_0, \dots, u_{T-1}) \mid \forall k \in \{0, \dots, T-1\}, \mathbb{R}^p \ni u_k \text{ is} \\ \mathcal{F}_k \text{-measurable and } u_k(\omega) \in [-M, M]^p \text{ } \mathbb{P}\text{-a.s.}\},$$

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

480 Let  $F_0 : \mathbb{R}^{p \times T} \rightarrow \bar{\mathbb{R}}$  be proper, lower semi continuous, convex and satisfy Assumptions  
 481 2.7 and 4.1.(ii). Similarly to the previous subsection, we assume that there exists for  
 482 any  $i$  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  $F_i$   
 483 satisfies Assumption 5.2.(i).

484 Now for any  $n \in \mathbb{T}^*$  the optimization problems  $(P_1^d)$  and  $(P_2^d)$  can be clearly  
 485 defined:

$$486 \quad (5.11) \quad (P_1^d) \begin{cases} \inf_{u \in \mathcal{U}} J^d(u) \\ J^d(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right), \end{cases}$$

487 and

$$488 \quad (5.12) \quad (P_2^d) \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}^d(u) \\ \tilde{J}^d(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) + \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right). \end{cases}$$

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

491 **COROLLARY 5.6.** (i) Problems  $(P_1^d)$  and  $(P_2^d)$  admit both a unique solution.  
 492 (ii) Any optimal solution of problem  $(P_2^d)$  is an  $\varepsilon$ -optimal solution, where  $\varepsilon =$   
 493  $cM^2/n$ , of problem  $(P_1^d)$ .

494 *Proof.* The proof of point (i) is analogous to the one of Lemma 2.2. Similarly,  
 495 proof of point (ii) is analogous to the one of Theorem 2.8.  $\square$

496 One can implement the *Stochastic Uzawa* (Algo 4.1) and the *Sampled Stochastic*  
 497 *Uzawa* (Algo 4.2) in this discrete time setting with Lemma 4.5 and Theorems 4.6  
 498 and 4.10 still ensuring the algorithm convergence.

499 **6. A numerical example: the LQG (Linear Quadratic Gaussian) prob-**  
 500 **lem.** This sections aims at illustrating numerically the convergence of the *Stochastic*  
 501 *Uzawa* (Algo 4.1) on a simple example. The algorithm speed of convergence is stud-  
 502 ied, depending on the number of dual iterations and of agents. A linear quadratic  
 503 formulation is considered, with  $n$  agents in a discrete setting problem  $(P_2^{LQG})$ . We  
 504 use the notations of Section 5.2.

505 This framework constitutes a simple test case, since the (deterministic) Uzawa's  
 506 algorithm can be performed, and one can compare the resulting multiplier estimate  
 507 with the one provided by the *Stochastic Uzawa* algorithm. Besides all the assump-  
 508 tions required for the convergence of the *Stochastic Uzawa* (Algo 4.1) are satisfied  
 509 for problem  $(P_2^{LQG})$ . In addition the local problems (line 5 of this algorithm) can be  
 510 resolved analytically.

511 Problem  $(P_2^{LQG})$  is similar to  $(P_2^d)$  defined in (5.12), but in this specific case, the  
 512 function  $F_0$  is a quadratic function of the aggregate strategies of the agents

$$513 \quad (6.1) \quad F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) := \frac{\nu}{2} \sum_{t=0}^T \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u_t^i) - r_t \right)^2,$$

514 where  $\nu > 0$ ,  $\{r_t\}$  is a deterministic target sequence. Similarly, the cost functions  $F_i$   
515 of the agents is expressed in a quadratic form of its state  $X^{i,u^i}$  and control  $u^i$ .

$$516 \quad (6.2) \quad F_i(u^i, X^{i,u^i}) := \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2,$$

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

$$519 \quad (6.3) \quad \sum_{i=1}^n F_i(u^i, X^{i,u^i}) = \frac{1}{2} \left( \sum_{t=0}^T X_t^{u^\top} D X_t^u + u_t^\top Q u_t \right) + \frac{1}{2} X_T^{u^\top} D^f X_T^u.$$

520 Now the optimization problem  $(P_2^{LQG})$  is clearly defined.

521 To find the optimal multiplier and control of  $(P_2^{LQG})$ , the *Stochastic Uzawa* Al-  
522 gorithm 4.1 is applied where in this specific case the lines 4 and 6 take respectively  
523 the following form at any dual iteration  $k$ :

$$524 \quad (6.4) \quad u^i(\lambda^k) := \arg \min_{u^i \in \bar{U}^i} \left\{ \mathbb{E} \left( \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 + \lambda_t^k u_t^i \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2 \right) \right\},$$

525

$$526 \quad (6.5) \quad v(\lambda^k) := \arg \min_{v \in \mathbb{R}^T} \left\{ \left( \sum_{t=0}^T \nu (v_t - r_t)^2 - \lambda_t^k v_t \right) \right\}.$$

527 The optimization problem (6.4) solved by each local agent is also in the LQG frame-  
528 work. One can solve these problems using the results of [24]. The resolution via  
529 Riccati equations of (6.4) shows that  $u^i(\lambda^k)$  is a linear function of the state  $X^{i,u^i}$   
530 and of the price  $\lambda^k$ . Therefore, in this specific example, for any  $t$  one can explicitly  
531 compute  $\mathbb{E}(u_t^i(\lambda^k) | \mathcal{G}_k)$ , where  $\mathcal{G}_k$  is defined in (4.9). It allows us to implement the  
532 (deterministic) Uzawa's algorithm as a reference to evaluate the performances of the  
533 *Stochastic Uzawa* algorithm.

534 Different population sizes  $n$  are considered, with  $n$  ranging between 1 and  $10^4$ .  
535 Similarly the algorithm is stopped for different numbers of dual iteration  $k$ , ranging  
536 between 1 and  $10^4$ . In order to evaluate the bias and variance of the *Stochastic Uzawa*  
537 algorithm, we have performed  $J = 1000$  runs of the *Stochastic Uzawa* algorithm.

538 For any  $n$ , given the strong convexity of the dual function associated with  $(P_2^{LQG})$ ,  
539 there exists a unique optimal multiplier  $\bar{\lambda}^n$ . For any  $n$ ,  $\lambda^{k,n,j}$  denotes the dual price  
540 computed during the  $j^{\text{th}}$  simulations ( $j = 1, \dots, J$ ) of the *Stochastic Uzawa* algorithm,  
541 after  $k$  dual iterations.

542 For any  $n$ , the deterministic multiplier  $\bar{\lambda}^n$  is obtained by applying Uzawa's al-  
543 gorithm, after  $10^4$  dual iterations. To this end, we applied the *Stochastic Uzawa*  
544 Algorithm 4.1 where we ignored the line 8 and we replaced the update of  $\lambda^k$  line 9

545 by:  $\bar{\lambda}^{k+1} \leftarrow \bar{\lambda}^k + \rho_k \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i(\bar{\lambda}^k)) - v(\bar{\lambda}^k) \right)$ .

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

548 The following results compare the multipliers  $\lambda^{k,n,j}$  and  $\bar{\lambda}^n$ , obtained respectively  
549 by applying the *Stochastic Uzawa* and Uzawa algorithms.

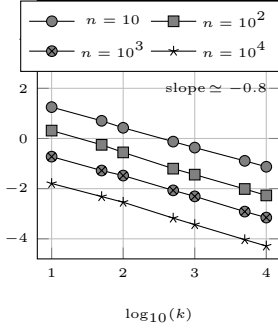


Figure 6.0.1:  $\log_{10}(v_{k,n})$  function of  $k$ , given the number of agents  $n$

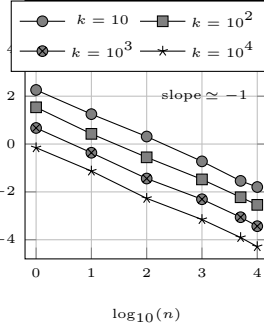


Figure 6.0.2:  $\log_{10}(v_{k,n})$  function of  $n$ , given the number of agents  $k$

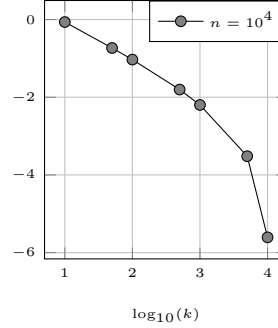


Figure 6.0.3:  $\log_{10}(\|b_{k,n}\|_2^2)$  function of  $k$ , given the number of agents  $n = 10^4$

550 For any  $k$  and  $n$ ,  $b_{k,n}$ ,  $v_{k,n}$  and  $\ell_{k,n}$  denotes respectively an estimation of the bias,  
 551 the variance and the L2 norm of the error, via Monte Carlo method with  $J$  simulations.

552 Thus we have for any  $k$  and  $n$ :  $b_{k,n} = \frac{1}{J} \sum_{j=1}^J \lambda^{k,n,j} - \bar{\lambda}^n$ ,  $v_{k,n} = \frac{1}{J} \sum_{j=1}^J \|\lambda^{k,n,j} - \bar{\lambda}^n -$

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

554 On Figure 6.0.1, we observe a behavior in  $1/k^\alpha$  (with  $\alpha \simeq 0.8$ ) of the variance  
 555  $v_{k,n}$  w.r.t. the number of iterations  $k$ . This rate of convergence is consistent with [6,  
 556 Theorem 2.2.12, Chapter 2] for Robbins Monro algorithm where the convergence is  
 557 proved to be of order at most in  $1/k$ .

558 On Figure 6.0.2 we observe a behavior in  $1/n^\beta$  (with  $\beta \simeq 1$ ) of the variance  $v_{k,n}$   
 559 w.r.t. the number of agents  $n$ . This is expected, see [6, Theorem 2.2.12, Chapter 2]  
 560 and observing that the variance of  $Y^{k+1}$  is of order  $1/n$  for any iteration  $k$ .

561 On Figure 6.0.3 we observe a faster behavior than  $1/k$  of the bias  $\|b_{k,n}\|_2^2$  w.r.t. the  
 562 number of iterations  $k$ . Thus for a large number of iterations ( $k > 0$ ), the dominant  
 563 term impacting the error  $\ell_{k,n}$  is the variance  $v_{k,n}$ .

564 **7. Price-based coordination of a large population of thermostatically**  
 565 **controlled loads.** The goal of this section is to demonstrate the applicability of the  
 566 presented approach for the coordination of thermostatic loads in a smart grid context.  
 567 The problem analyses the daily operation of a power system with a large penetration  
 568 of price-responsive demand, adopting a modelling framework similar to [5]. Two distinct  
 569 elements are considered: i) a system operator, that must schedule a portfolio  
 570 of generation assets in order to satisfy the energy demand at a minimum cost, and  
 571 ii) a population of price-responsive loads (TCLs) that individually determine their  
 572 ON/OFF power consumption profile in response to energy prices with the objective  
 573 of minimizing their operating cost while fulfilling users' requirements. Note that the  
 574 operations of the two elements are interconnected, since the aggregate power consumption  
 575 of the TCLs will modify the demand profile that needs to be accommodated by  
 576 the system operator.

577 **7.1. Formulation of the problem.** In the considered problem, the function  $F_0$   
 578 represents the minimized power production cost and corresponds to the resolution of  
 579 an Unit Commitment (UC) problem. The UC determines generation scheduling decisions  
 580 (in terms of energy production and frequency response (FR) provision) in order

581 to minimize the short term operating cost of the system while matching generation  
 582 and demand. The latter is the sum of an inflexible deterministic component (denoted  
 583 for any instant  $t \in [0, T]$  by  $\bar{D}(t)$ ) and of a stochastic part, which corresponds to the  
 584 total TCL demand profile  $nU_{TCL}(t)$ .

585 For simplicity, a Quadratic Programming (QP) formulation in a discrete time  
 586 setting is adopted for the UC problem. The central planner disposes of  $Z$  genera-  
 587 tion technologies (gas, nuclear, wind) and schedules their production and allocated  
 588 response by slot of 30 min every day. For any  $j \in \{1, \dots, Z\}$  and  $\ell \in \{1, \dots, 48\}$ ,  
 589  $H_j(t_\ell)$ ,  $G_j(t_\ell)$  and  $R_j(t_\ell)$  are respectively the commitment, the power production  
 590 and response [MWh] from unit  $j$  during the time interval  $[t_\ell, t_{\ell+1}]$ . The associated  
 591 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  
 592  $R(t_\ell) = [R_1(t_\ell), \dots, R_Z(t_\ell)]$ .

593 The cost sustained at time  $t_\ell$  by unit  $j$  is linear with respect to the commit-  
 594 ment  $H_j(t_\ell)$  and quadratic with respect to generation  $G_j(t_\ell)$  and can be expressed  
 595 as  $c_{1,j}H_j(t_\ell)G_j^{Max}(t_\ell) + c_{2,j}G_j(t_\ell) + c_{3,j}G_j(t_\ell)^2$ , with  $G_j^{Max}$  as the limit of produc-  
 596 tion allocated by each generation technology,  $c_{1,j}$  [€/MWh] as no-load cost and  $c_{2,j}$   
 597 [€/MWh] and  $c_{3,j}$  [€/MW<sup>2</sup>h] as production cost of the generation technology  $j$ . The  
 598 optimization of  $F_0$  must satisfy the following constraints for all  $\ell \in \{1, \dots, 48\}$  and  
 599  $\ell \in \{1, \dots, 48\}$ :

$$600 \quad (7.1) \quad \sum_{j=1}^Z G_j(t_\ell) - \int_{t_\ell}^{t_{\ell+1}} (\bar{D}(t) + nU_{TCL}(t))dt = 0,$$

601

$$602 \quad (7.2) \quad 0 \leq H_j(t_\ell) \leq 1,$$

603

$$604 \quad (7.3) \quad R_j(t_\ell) - r_j H_j(t_\ell) G_j^{max}(t_\ell) \leq 0,$$

605

$$606 \quad (7.4) \quad R_j(t_\ell) - s_j (H_j(t_\ell) G_j^{max}(t_\ell) - G_j(t_\ell)) \leq 0,$$

607

$$608 \quad (7.5) \quad \Delta G_L - \Lambda (\bar{D}(t_\ell) + n(\bar{U}_{TCL}(t_\ell) - \bar{R}_{TCL}(t_\ell))) \Delta f_{qss}^{max} - \hat{R}(t_\ell) \leq 0,$$

609

$$610 \quad (7.6) \quad 2\Delta G_L t_{ref} t_d - t_{ref}^2 \hat{R}(t_\ell) - 4\Delta f_{ref} t_d \hat{H}(t_\ell) \leq 0,$$

611

$$612 \quad (7.7) \quad \bar{q}(t) - \hat{H}(t) \hat{R}(t) \leq 0$$

613

$$614 \quad (7.8) \quad \mu r_j H_j(t_\ell) G_j^{max}(t_\ell) - G_j(t_\ell) \leq 0,$$

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

$$619 \quad \hat{R}(t_\ell) = \sum_{j=1}^Z R_j(t_\ell) + nR_{TCL}(t_\ell) \quad \text{and} \quad \hat{H}(t_\ell) = \sum_{j=1}^Z \frac{h_j H_j(t_\ell) G_j^{max} - h_L \Delta G_L}{f_0}.$$

620

621 Assuming that for any generic generation technology  $j$ , the size of single plants  
 621 included in  $j$  is quite smaller than the aggregate installed capacity of  $j$ , inequality

622 (7.2) sets that commitment decisions can be extended to the fleet and expressed by  
 623 continuous variables  $H_j(t_\ell) \in [0, 1]$ .

624 The amount of response allocated by each generation technology is limited by  
 625 the headroom  $r_j H_j(t_\ell) G_j^{max}(t_\ell)$  in (7.3) and the slope  $s_j$  linking the FR with the  
 626 dispatch level (7.4). Constraints (7.5) to (7.8) deal with frequency response provision  
 627 and  $R_{TCL}$  (the mean of FR allocated by TCLs). They guaranty secure frequency  
 628 deviations following sudden generation loss  $\Delta G_L$ . Inequality (7.5) allocates enough  
 629 FR (with delivery time  $t_d$ ) such that the quasi-steady-state frequency remains above  
 630  $\Delta f_{qss}^{max}$ , with  $\Lambda$  accounting for the damping effect introduced by the loads [12]. Fi-  
 631 nally (7.7) constraints the maximum tolerable frequency deviation  $\Delta f_{nad}$ , following  
 632 the formulation and methodology presented in [23] and [25]. The rate of change of  
 633 frequency is taken into account in (7.6) where at  $t_{rcf}$  the frequency deviation remains  
 634 above  $\Delta f_{ref}$ . Constraint (7.8) prevents trivial unrealistic solutions that may arise  
 635 in the proposed formulation, such as high values of committed generation  $H_j(t_\ell)$  in  
 636 correspondence with low (even zero) generation dispatch  $G_j(t_\ell)$ . The reader can refer  
 637 to [5] for more details on the UC problem.

638 The solution of the UC problem, corresponding to the function  $F_0$ , can be de-  
 639 scribed by the following optimization problem:

$$640 \quad (7.9) \quad F_0(U_{TCL}, R_{TCL}) := \min_{H, G, R} \sum_{\ell=1}^{48} \sum_{j=1}^Z c_{1,j} H_j(t_\ell) G_j^{max}(t_\ell) + c_{2,j} G_j(t_\ell) + c_{3,j} G_j(t_\ell)^2,$$

641 subject to equations (7.1)-(7.8).

642 Note that the formulation of the present problem does not fulfill all the assumption  
 643 presented in Section 4. In particular, the function  $F_0$  is not strictly convex, as instead  
 644 supposed in Theorem 4.6.(ii).(iii). Nevertheless, the numerical simulations of Section  
 645 7.2 shows that the proposed approach is still able to achieve convergence.

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

$$650 \quad (7.10) \quad \begin{cases} dX_t^{i, u^i} &= -\frac{1}{\gamma_i} (X_t^{i, u^i} - X_{OFF}^i + \zeta_i u_t^i) dt + \sigma_i dW_t^i, \quad \text{for } t \in [0, T], \\ X_{0, u^i}^i &= x_0^i \in \mathbb{R}, \end{cases}$$

651 where:

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

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

$$658 \quad (7.11) \quad \mathcal{U}_i := \{v : \Omega \times [0, T] \rightarrow \mathbb{R}, v(\cdot) \text{ is } \mathbb{F} - \text{prog. measurable}, \\ v(\omega) \in \mathbb{U} \text{ and } v_t(\omega) \in \{0, P_{ON, i}\}, \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega\},$$

659 The TCLs dynamics in (7.10) have been derived according to [11], with the addition  
 660 of the stochastic term  $\sigma_i dW_t^i$  to account for the influence of the environment (open-

661 ing/closing of the fridge, environment temperature etc) on the evolution of the TCL  
662 temperature.

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

(7.12)

$$665 \quad (P_1^{TCL}) \quad \begin{cases} \inf_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i, \frac{1}{n} \sum_{i=1}^n r_i(u^i, X^{i,u^i}) \right) \right) \\ \quad + \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i,u^i}) ds + \gamma_i (X_T^{i,u^i} - \bar{X}^i)^2 \right), \end{cases}$$

666 where, for any  $i \in \{1, \dots, n\}$  and any  $s \in [0, T]$ :

667 •  $r_i(u^i, X^{i,u^i})(s)$  is the maximum amount of FR allocated by TCL  $i$  at time  $s$ :

$$668 \quad (7.13) \quad r_i(u^i, X^{i,u^i})(s) := u_s^i \frac{X_s^{i,u^i} - X_{min}^i}{X_{max}^i - X_{min}^i}.$$

669 •  $f_i(u_s^i, X_s^{i,u^i})$  is the individual discomfort term of the TCL  $i$  at time  $s$ :

$$670 \quad (7.14) \quad f_i(u_s^i, X_s^{i,u^i}) := \alpha_i (X_s^{i,u^i} - \bar{X}^i)^2 + \beta_i ((X_{min}^i - X_s^{i,u^i})_+^2 + (X_s^{i,u^i} - X_{max}^i)_+^2),$$

671 where:

672 –  $\alpha_i (X_s^{i,u^i} - \bar{X}^i)^2$  is a discomfort term penalizing temperature deviation  
673 from some comfort target  $\bar{X}$  [ $^{\circ}C$ ], with  $\alpha_i$  a discomfort term parameter  
674 [ $\pounds/h(^{\circ}C)^2$ ].

675 –  $\beta_i ((X_s^{i,u^i} - X_{min}^i)_+^2 + (X_{max}^i - X_s^{i,u^i})_+^2)$  is a penalization term to keep  
676 the temperature in the interval  $[X_{min}^i, X_{max}^i]$ , with  $\beta_i$  a target term pa-  
677 rameter [ $\pounds/s(^{\circ}C)^2$ ] and for any  $x \in \mathbb{R}$ ,  $(a)_+ = \max(0, a)$ .

678 •  $\gamma_i (X_T^{i,u^i} - \bar{X}^i)^2$  is a terminal cost imposing periodic constraints, with  $\gamma$  a  
679 target term parameter [ $\pounds/s(^{\circ}C)^2$ ].

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

682 The modified problem  $(P_2^{TCL})$  is studied to solve  $(P_1^{TCL})$ .

(7.15)

$$683 \quad (P_2^{TCL}) \quad \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i), \frac{1}{n} \sum_{i=1}^n \mathbb{E}(r_i(u^i, X^{i,u^i})) \right) \\ \quad + \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i,u^i}) ds + \gamma_i (X_T^{i,u^i} - \bar{X}^i)^2 \right). \end{cases}$$

684 **7.2. Decentralized implementation.** The *Sampled Stochastic Uzawa* Algo-  
685 rithm 4.2 is applied to solve  $(P_2^{TCL})$ , with  $m = 317$  simulations per iteration. At each  
686 iteration  $k$ , the lines 4 and 6 correspond respectively to the solution of a deterministic  
687 UC problem and of an Hamilton Jacobi Bellman (HJB) equation. The time steps  
688  $\Delta t = 7.6$  s and temperature steps  $\Delta T = 0.15^{\circ}C$  are chosen for the discretization of  
689 the HJB equation. Let us note that at line 6, each TCL solves its own local problem

690 on the basis of the received price signal  $\lambda^k = (p^k, \rho^k)$ :

$$691 \quad (7.16) \quad \inf_{u^i \in \mathcal{U}_i} \int_0^T f_i(u_s^i, X_s^{i,u^i}) + u_s^i p_s^k - r_i(u^i, X^{i,u^i})(s) \rho_s^k ds,$$

692 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  
 693 as consumption cost and  $r_i(u^i, X^{i,u^i})(s) \rho_s^k$  as fee awarded for FR provision. This  
 694 implementation has a practical sense: each TCL uses local information and a price  
 695 that is communicated to them to schedule its power consumption on the time interval  
 696  $[0, T]$ . It follows that, with the proposed approach, it is possible to optimize the overall  
 697 system costs in  $(P_1^{TCL})$  in a distributed manner, with each TCL acting independently  
 698 and pursuing the minimization of its own costs.

699 **7.3. Results.** The generation technologies available in the system are nuclear,  
 700 combined cycle gas turbines (CCGT), open cycle gas turbines (OCGT) and wind.  
 701 The characteristics and parameters of the UC in this simulation are the same as in  
 702 [5].

703 It is assumed that a population of  $n = 2 \times 10^7$  fridges with built-in freeze compart-  
 704 ment operates in the system according to the proposed price-based control scheme.  
 705 For any agent  $i$  we set the consumption parameter  $P_{ON,i} = 180W$ . The values of the  
 706 TCL dynamic parameters  $\gamma_i$  and  $X_{OFF}^i$  of (7.10) are equal to the ones taken in [5].  
 707 Note that it is possible to take a population of heterogeneous TCLs with different  
 708 parameter values. The initial temperature are picked randomly uniformly between  
 709  $-21^\circ C$  and  $-14^\circ C$ . For any agent  $i$ , the parameters of the individual cost function  
 710  $f_i$ , defined in (7.14), are:  $\alpha_i = 0.2 \times 10^{-4} \text{ £/s}(\text{ }^\circ C)^2$ ,  $\beta_i = 50 \text{ £/s}(\text{ }^\circ C)^2$ ,  $\bar{X}^i = -17.5^\circ C$   
 711 and  $X_{max} = -14^\circ C$ ,  $X_{min} = -21^\circ C$ . The parameter  $\beta_i$  is taken intentionally very  
 712 large to make the temperature stay in the interval  $[X_{max}^i, X_{min}^i]$ . Note that the indi-  
 713 vidual problems solved by the TCLs are distinct than the ones in [5] (different terms  
 714 and parameters).

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

718 The resulting profile of total power consumption  $nU_{TCL}$  and total allocated re-  
 719 sponse  $nR_{TCL}$  by the TCLs population are reported on figure 7.3.1. in three "flexibil-  
 720 ity scenario" each corresponding to a case where TCL flexibility is enabled with three  
 721 different volatilities  $\sigma = 0$ ;  $\sigma = 1$  and  $\sigma = 2$ . The electricity prices  $p$  and response  
 722 availability prices  $\rho$  are shown in Figure 7.3.2. As observed in [5], the total con-  
 723 sumption  $nU_{TCL}$  is higher when the price  $p$  is lower and inversely the total allocated  
 724 response  $nR_{TCL}$  is higher when the price signal  $\rho$  is also higher. This can be observed  
 725 during the first hours of the day, between 0 and 6h. The power  $U_{TCL}$  then oscillates  
 726 during the day in order to maintain feasible levels of the internal temperature of the  
 727 TCLs. Though the prices seem not to be sensitive to the values taken by  $\sigma$ , the  
 728 average consumption  $U_{TCL}$  and response  $R_{TCL}$  are highly correlated to the volatility  
 729 of the temperature of the TCLs. The less noisy their temperature are, the more price  
 730 sensitive and flexible their consumption profiles are. The TCLs impact on system  
 731 commitment decisions and consequent energy/FR dispatch levels is also analyzed and  
 732 displayed in Figure 7.3.3 and 7.3.4. The production and reserve in the "flexibility sce-  
 733 nario" minus the production and reserve in the "no-flexibility scenario" are plotted,  
 734 for different volatilities  $\sigma$ . In the no-flexibility scenario we impose  $R_{TCL}(t) = 0$  and  
 735 we consider that the TCLs operate exclusively according to their internal tempera-  
 736 ture  $X^{i,u^i}$ . They switch ON ( $u^i(t) = P_{ON,i}$ ) when they reach their maximum feasible



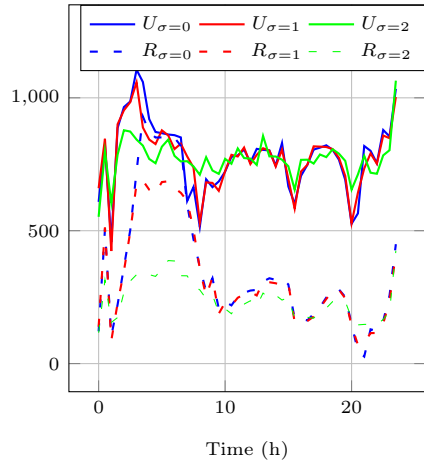


Figure 7.3.1: Total power consumption  $U$  and allocated response  $R$  (MW) of TCLs after 75 iterations of the algorithm.

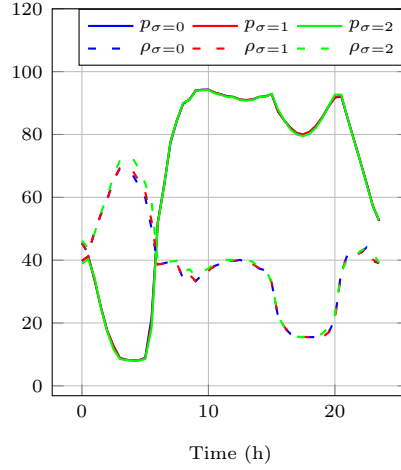


Figure 7.3.2: Electricity price  $p$  and response availability price  $\rho$  (£/MWh) after 75 iterations of the algorithm.

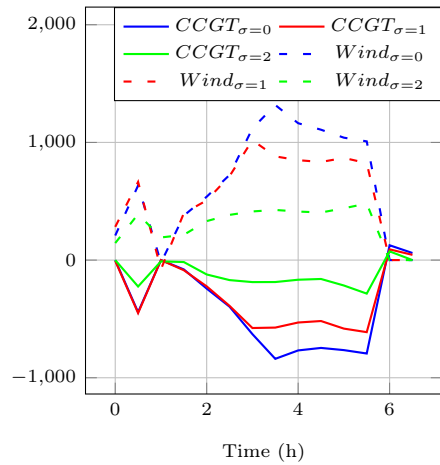


Figure 7.3.3: Deviation of generation profiles (MW) from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

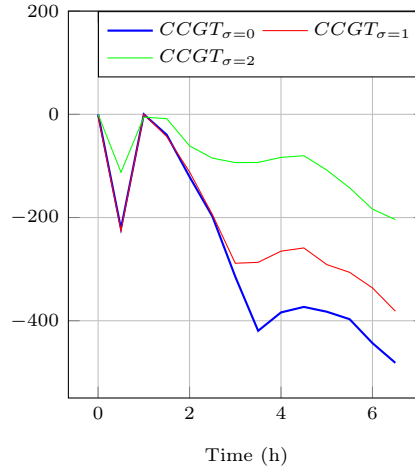


Figure 7.3.4: Deviation of Frequency Response (MW) allocated by CCGT from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

737 temperature  $X_{max}^i$  and they switch back OFF again ( $u^i(t) = 0$ ) when they reach the  
 738 minimum temperature  $X_{min}^i$ . In figure 7.3.3, we can clearly observe that TCL's flexi-  
 739 bility allows to increase the contribution of wind generation (reducing curtailment) to  
 740 the energy balance of the system while decreasing the contribution of CCGT both in  
 741 energy and frequency response. Without TCL support, the optimal solution envisages  
 742 a further curtailment of wind output in favor of an increase in CCGT generation, as  
 743 wind does not provide FR. As expected, the influence of the TCL on the system is  
 744 larger when the temperature volatility is lower.

745

	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
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 (£)

746

747

748

749

750

751

752

753

754

755

756

The system costs (i.e. UC solution) obtained with the flexibility scenario (FS) are now compared with the Business-as-usual (BAU) framework ones (the TCLs do not exploit their flexibility and they operate exclusively according to their internal temperature as previously explained) in Tab. 1. As expected the costs are lower in the CF where TCLs participate in reducing the system generation costs. The reduction is higher for  $\sigma = 0$ , where the reduction is about 1.9%, than for  $\sigma = 1$  or  $\sigma = 2$ , where the the reduction is respectively about 1.6% and 1.2%. This relies on the tendency of the TCLs to be more flexible when their volatility is low. The reduction observed in the CF scenario is due to the smaller use of OCGT and CCGT generation technologies for the benefit of wind.

757

### Appendix A. Appendix.

758

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

759

760

761

762

763

*Proof.* We suppose for all  $x \in H$  and  $q \in \partial f(x)$  that we have  $\|q\|_H \leq C(1 + \|x\|_H)$ . For all  $q \in \partial f(x)$ , we have that  $f(0) \geq f(x) - \langle q, x \rangle_H$ , so that  $f(x) \leq f(0) + \|q\|_H \|x\|_H$ , and  $f(x) \leq C(1 + \|x\|_H^2)$ . So the subquadratic growth property holds if  $\partial f$  has linear growth, at points where  $\partial f(x)$  is non empty. Since the subdifferential is nonempty in the interior of the domain, the subquadratic growth property holds everywhere.

Conversely let the subquadratic growth property holds. Then for all  $x, y \in H$  and  $q \in \partial f(x)$ :

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

Take  $y = x + \alpha q$ , we get

$$C(1 + \|x\|_H^2 + \alpha^2 \|q\|_H^2) \geq -C(1 + \|x\|_H^2) + \alpha \|q\|_H^2.$$

We deduce that:

$$C(2 + 2\|x\|_H^2 + \alpha^2 \|q\|_H^2) \geq \alpha \|q\|_H^2.$$

764

765

766

Suppose  $\|x_k\|_H$  tends to infinity, if we can take  $q_k$  in  $\partial f(x^k)$  and  $\|x_k\|_H / \|q_k\|_H$  converging to 0, then there exists  $\alpha = \alpha_k$  converging to 0, such that  $2 + 2\|x_k\|_H^2 + \alpha_k^2 \|q_k\|_H^2 \leq 2\alpha_k^2 \|q_k\|_H^2$  so that  $1 \leq 2C\alpha_k$ , and this gives a contradiction.  $\square$

767

### REFERENCES

768

769

770

771

772

- [1] Dimitri P Bertsekas and Steven Shreve. *Stochastic optimal control: the discrete-time case*. 2004.
- [2] 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(3):1086–1098, 2018.

- 773 [3] Pierre Carpentier, Jean-Philippe Chancelier, Michel De Lara, and François Pacaud. Mixed spatial  
774 and temporal decompositions for large-scale multistage stochastic optimization problems. *Journal of Optimization Theory and Applications*, 186(3):985–1005, 2020.  
775
- 776 [4] Michael Chertkov and Vladimir Chernyak. Ensemble of thermostatically controlled loads: Statistical physics approach. *Scientific reports*, 7(1):8673, 2017.  
777
- 778 [5] Antonio De Paola, Vincenzo Trovato, David Angeli, and G. Strbac. A mean field game approach  
779 for distributed control of thermostatic loads acting in simultaneous energy-frequency response markets. *IEEE Transactions on Smart Grid*, PP:1–1, 01 2019.  
780
- 781 [6] Marie Dufflo. *Random iterative models*, volume 34. Springer Science & Business Media, 2013.  
782
- 783 [7] Caroline Geiersbach and Georg Ch Pflug. Projected stochastic gradients for convex constrained  
784 problems in Hilbert spaces. *SIAM Journal on Optimization*, 29(3):2079–2099, 2019.  
785
- 786 [8] Pierre Girardeau. *Solving large-scale dynamic stochastic optimization problems*. PhD thesis,  
787 Université Paris-Est, December 2010.  
788
- 789 [10] Julia L Higle and Suvrajeet Sen. *Stochastic decomposition: a statistical method for large scale  
790 stochastic linear programming*, volume 8. Springer Science & Business Media, 2013.
- 791 [11] Arman C Kizilkale, Rabih Salhab, and Roland P Malhamé. An integral control formulation  
792 of mean field game based large scale coordination of loads in smart grids. *Automatica*,  
793 100:312–322, 2019.
- 794 [12] Prabha Kundur, Neal J Balu, and Mark G Lauby. *Power system stability and control*, volume 7.  
795 McGraw-hill New York, 1994.
- 796 [13] Vincent Leclere. *Contributions to decomposition methods in stochastic optimization*. PhD  
797 thesis, Université Paris-Est, June 2014.
- 798 [14] Xuerong Mao. *Stochastic differential equations and applications*. Elsevier, 2007.
- 799 [15] François Pacaud. *Decentralized optimization for energy efficiency under stochasticity*. PhD  
800 thesis, Université Paris-Est, October 2018.
- 801 [16] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to  
802 energy planning. *Mathematical programming*, 52(1-3):359–375, 1991.
- 803 [17] Pedro Pérez-Aros and Emilio Vilches. An enhanced Baillon-Haddad theorem for convex functions  
804 on convex sets. *arXiv preprint arXiv:1904.04885*, 2019.
- 805 [18] Andrew B Philpott and Ziming Guan. On the convergence of stochastic dual dynamic programming  
806 and related methods. *Operations Research Letters*, 36(4):450–455, 2008.
- 807 [19] R Tyrrell Rockafellar and Roger J-B Wets. Scenarios and policy aggregation in optimization  
808 under uncertainty. *Mathematics of operations research*, 16(1):119–147, 1991.
- 809 [20] Andrzej Ruszczyński and Alexander Shapiro. Stochastic programming models. *Handbooks in  
810 operations research and management science*, 10:1–64, 2003.
- 811 [21] David H Salinger. *A splitting algorithm for multistage stochastic programming with application  
812 to hydropower scheduling*. PhD thesis, 1997.
- 813 [22] Joe A Short, David G Infield, and Leon L Freris. Stabilization of grid frequency through  
814 dynamic demand control. *IEEE Transactions on power systems*, 22(3):1284–1293, 2007.
- 815 [23] Fei Teng, Vincenzo Trovato, and Goran Strbac. Stochastic scheduling with inertia-dependent  
816 fast frequency response requirements. *IEEE Transactions on Power Systems*, 31(2):1557–  
817 1566, 2015.
- 818 [24] Emanuel Todorov. Optimal control theory. *Bayesian brain: probabilistic approaches to neural  
819 coding*, pages 269–298, 2006.
- 820 [25] Vincenzo Trovato, Agnès Bialecki, and Anes Dallagi. Unit commitment with inertia-dependent  
821 and multispeed allocation of frequency response services. *IEEE Transactions on Power  
822 Systems*, 34(2):1537–1548, 2018.
- 823 [26] Vincenzo Trovato, Simon H Tindemans, and Goran Strbac. Leaky storage model for optimal  
824 multi-service allocation of thermostatic loads. *IET Generation, Transmission & Distribution*,  
825 10(3):585–593, 2016.