



**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. 2020. hal-02917014v3

**HAL Id: hal-02917014**

**<https://inria.hal.science/hal-02917014v3>**

Preprint submitted on 17 Sep 2020 (v3), 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   ADRIEN SEGURET<sup>†</sup>, CLÉMENCE ALASSEUR<sup>‡</sup>, J. FRÉDÉRIC BONNANS<sup>§</sup>, ANTONIO  
4   DE PAOLA<sup>¶</sup>, NADIA OUDJANE<sup>||</sup>, AND VINCENZO TROVATO<sup>#</sup>

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

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

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

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

17  (1.1)            $(P_1) \quad \begin{cases} \inf_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right). \end{cases}$

18  The dynamics of the state of each agent  $X^{i,u^i}$  is driven by independent Brownian motions  
19   $W^i$  (no common noise) so that potential interactions between agents dynamics  
20  is only due to the non anticipative controls  $u^i$  supposed to be progressively measur-  
21  able w.r.t. to the Brownian noise  $W = (W^i)_{i \in \{1, \dots, n\}}$ . We emphasize, the specific  
22  structure of that problem whose cost function is the sum of, on one side, additively  
23  separable terms  $F_i$  between agents and a coupling term  $F_0$ , function of the *aggregate*  
24  strategies  $\frac{1}{n} \sum_{i=1}^n u^i$ .

25  **1.1. Motivations.** This work is motivated by its potential applications for large-  
26  scale coordination of flexible appliances, to support power system operation in a con-  
27  text of increasing penetration of renewables. One type of appliances that has been  
28  consistently investigated in the last few years, for its intrinsic flexibility and potential

---

\* Submitted to the editors September 17, 2020

**Funding:** The first, second, third and fifth author thank the FiME Lab (Institut Europlace de Finance). The third author was supported by the PGM0 project “Optimal control of conservation equations”, itself supported by iCODE(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)).

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

<sup>§</sup>CMAP, Ecole polytechnique, 91128 Palaiseau, France, and Inria-Saclay([frederic.bonnans@inria.fr](mailto:frederic.bonnans@inria.fr)).

<sup>¶</sup>Center for Sustainable Power Distribution University of Bath, Bath, UK([adp50@bath.ac.uk](mailto:adp50@bath.ac.uk)).

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

<sup>#</sup>Department of Electrical and Electronic Engineering, Imperial College London, London, UK([v.trovato10@imperial.ac.uk](mailto:v.trovato10@imperial.ac.uk)).

for network support, includes thermostatically controlled loads (TCLs) such as refrigerators or air conditioners. Several papers have already investigated the potential of dynamic demand control and frequency response services of TCLs [21] and how the population recovers from significant perturbations [3]. The coordination of TCLs can be performed in a centralized way, like in [8]. However this approach raises challenging problems in terms of communication requirements and customer privacy. A common objective can be reached in a fully distributed approach, like in [25], where each TCL is able to calculate its own actions (ON/OFF switching) to pursue a common objective. This paper is related to the work of De Paola *et al.* [4], where each agent represents a flexible TCL device. In [4] a distributed solution is presented for the operation of a population of  $n = 2 \times 10^7$  refrigerators providing frequency support and load shifting. They adopt a game-theory framework, modelling the TCLs as price-responsive rational agents that schedule their energy consumption and allocate their frequency response provision in order to minimize their operational costs. The potential practical application of our work also considers a large population of TCLs which, contrarily to [4], have stochastic dynamics. The proposed approach is able to minimize the overall system costs in a distributed way, with each TCL determining its optimal power consumption profile in response to price signals.

**1.2. Related literature.** The considered problem belongs to the class of stochastic control: looking for strategies minimizing the expectation of an objective function under specific constraints. One of the main approaches proposed in the literature to tackle this problem is to use random trees: this consists in replacing the almost sure constraints, induced by non-anticipativity, by a finite number of constraints to get a finite set of scenarios (see. [9] and [19]). Once the tree structure is built, the problem is solved by different decomposition methods such as scenario decomposition [18] or dynamic splitting [20]. The main objective of the scenario method is reducing the problem to an approximated deterministic one. The paper focuses on high dimensional noise problems with large number of time steps, for which this approach is not feasible. The idea of reducing a single high dimensional problem to a large number with low dimension has been widely studied in the deterministic case. In deterministic and stochastic problems a possibility is to use time decomposition thanks to the Dynamic Programming Principle [1] taking advantage of Markov property of the system. However, this method requires a specific time structure of the cost function and fails when applied to problems for which the state space dimension is greater than five. One can deal with the curse of dimensionality, under continuous linear-convex assumptions, by using the Stochastic Dual Dynamic Programming algorithm (SDDP) [15] 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 [17], there is no guarantee on the speed of the convergence and there is no good stopping test. In [14], 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 [7], [12]). 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

78 algorithm. To this end in [12], the Uzawa's algorithm, formulated in a Hilbert set-  
 79 ting, is extended to a Banach space. DADP has been applied in different cases, such  
 80 as storage management problem for electrical production in [7, chapter 4] and hydro  
 81 valley management [2]. In the proposed paper, in the same vein as DADP we propose  
 82 a price decomposition approach restricted to deterministic prices. This new approach  
 83 takes advantage of the large population number in order to introduce an auxiliary  
 84 problem where the coupling term is purely deterministic.

85 **1.3. Contributions.** We consider the following approximation of problem  $(P_1)$ :

$$86 \quad (1.2) \quad (P_2) \quad \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) + \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right). \end{cases}$$

87 As a first contribution, this paper shows that under some convexity and regularity  
 88 assumptions on  $F_0$  and  $(F_i)_{i \in \{1, \dots, n\}}$ , any solution of problem  $(P_2)$  is an  $\varepsilon_n$ -solution  
 89 of  $(P_1)$ , with  $\varepsilon_n \rightarrow 0$  when  $n \rightarrow \infty$ . In addition, an approach of price decomposition  
 90 for  $(P_2)$  is easier than for  $(P_1)$ , since the Lagrange multiplier is deterministic for  $(P_2)$ ,  
 91 whereas it is stochastic for  $(P_1)$ . Since computing the dual cost of  $(P_2)$  is expensive,  
 92 we propose *Stochastic Uzawa* and *Sampled Stochastic Uzawa* algorithms relying on  
 93 Robbins Monroe algorithm in the spirit of the stochastic gradient. Its convergence  
 94 is established. We check the effectiveness of the *Stochastic Uzawa* algorithm on a  
 95 linear quadratic Gaussian framework, and we apply the *Sampled Stochastic Uzawa*  
 96 algorithm to a model of power system, inspired by the work of A. De Paola *et al.* [4].

97 **2. General framework.** Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete filtered probability  
 98 space on which  $W = (W^i)_{i=1, \dots, n}$  is a  $n$ -dimensional Brownian motion, such that  
 99 for any  $t \in [0, T]$  and  $i \in \{1, \dots, n\}$ ,  $W_t^i$  takes value in  $\mathbb{R}$ , and generates the filtration  
 100  $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ .  $\mathbb{P}$  stands for the Wiener measure associated with this filtration and  
 101  $\mathbb{F}$  for the augmented filtration by all  $\mathbb{P}$ -null sets.

102 The following notations are used:

$$\begin{aligned} \mathbb{X} &:= \{ \varphi : \Omega \rightarrow \mathcal{C}([0, T], \mathbb{R}) \mid \varphi(\cdot) \text{ is } \mathbb{F} \text{-adapted, } \|\varphi\|_{\infty, 2} := \mathbb{E} \left( \sup_{s \in [0, T]} |\varphi(s)|^2 \right)^{\frac{1}{2}} < \infty \}, \\ 103 \quad L^2(0, T) &:= \{ \varphi : [0, T] \rightarrow \mathbb{R} \mid \int_0^T |\varphi(t)|^2 dt < \infty \}, \\ \mathbb{U} &:= \{ \varphi : [0, T] \times \Omega \rightarrow \mathbb{R} \mid \varphi(\cdot) \text{ is } \mathbb{F} \text{-prog. measurable, } \mathbb{E} \int_0^T |\varphi(t)|^2 dt < \infty \}, \end{aligned}$$

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

$$105 \quad (2.1) \quad \mathcal{U}_i := \{ v \in \mathbb{U} \text{ and } v_t(\omega) \in [-M_i, M_i], \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega \},$$

106 and we set  $M := \max_{i \in \{1, \dots, n\}} M_i$ . The set of admissible controls is  $\mathcal{U} := \mathcal{U}_1 \times \dots \times \mathcal{U}_n$ ,

107 whose elements are denoted by  $u := (u^1, \dots, u^n)$ .

108 Each local agent  $i = 1, \dots, n$  is supposed to control its state variable through the  
 109 control process  $u^i \in \mathcal{U}_i$  and suffers from independent uncertainties. More specifically,  
 110 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

111  $\mathbb{R}$  and follows the dynamics

(2.2)

$$112 \quad \begin{cases} dX_t^{i,u^i} &= \mu_i(t, u_t^i, X_t^{i,u^i})dt + \sigma_i(t, X_t^{i,u^i})dW_t^i, \text{ for } t \in [0, T], \text{ for } i \in \{1, \dots, n\} \\ X_{0,u^i}^i &= x_0^i \in \mathbb{R}. \end{cases}$$

113 Without loss of generality, the initial states  $x_0^i$  are supposed to be deterministic.

114 The process  $X^i$  is  $\mathbb{F}$ -progressively measurable. For all  $i$ ,  $\mathcal{F}^i$  stands for the natural  
115 filtration of the Brownian motion  $W^i$ .

116 **2.1. On the well-posedness of  $(P_1)$ .** In this section, the assumptions needed  
117 for  $(P_1)$  to be well posed are studied.

118 *Assumption 2.1.* For any  $i \in \{1, \dots, n\}$ , the functions  $\mu_i$  and  $\sigma_i$  are continuous  
119 w.r.t  $(u, x)$  uniformly in  $t$ . In addition there exists  $K_i > 0$  such that, for any  $t \in [0, T]$   
120 and  $\nu \in [-M, M]$ :

(2.3)

$$121 \quad \begin{aligned} |\mu_i(t, \nu, x) - \mu_i(t, \nu, y)| + |\sigma_i(t, \nu, x) - \sigma_i(t, \nu, y)| &\leq K_i |x - y|, \\ |\mu_i(t, \nu, x)| + |\sigma_i(t, \nu, x)| &\leq K_i (1 + |x|), \end{aligned} \quad \text{for any } x, y \in \mathbb{R}.$$

122 **LEMMA 2.2.** *Let  $i \in \{1, \dots, n\}$  and  $v \in \mathcal{U}_i$  be a control process. If Assumption*  
123 *2.1 holds, then there exists a unique process  $X^{i,v} \in \mathbb{X}$  satisfying (2.2) (in the strong*  
124 *sense) such that for any  $p \in [1, \infty)$ :*

$$125 \quad (2.4) \quad \mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^{i,v}|^p \right) < C(p, T, x_0, K) < \infty .$$

126 *Proof.* The proof for the existence and uniqueness of a solution of (2.2) relies on  
127 [13, Theorem 3.6, Chapter 2]. The inequality is a result of [13, Theorem 4.4, Chapter  
128 2].  $\square$

129 Let  $F_0 : L^2(0, T) \rightarrow \bar{\mathbb{R}}$  and  $F_i : L^2(0, T) \times \mathcal{C}([0, T], \bar{\mathbb{R}}) \rightarrow \mathbb{R}$  be proper and lower  
130 semi continuous functions, and there exists  $\hat{u} \in \mathcal{U}$  such that:

$$131 \quad (2.5) \quad \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n \hat{u}^i \right) \right) < \infty .$$

132 Define  $G_i : L^2(0, T) \rightarrow \bar{\mathbb{R}}$  by  $G_i(u^i) = F_i(u^i, X^{i,u^i})$  for any  $u^i \in \mathcal{U}_i$ . Additional  
133 assumptions are formulated below.

134 *Assumption 2.3.* For any  $i \in \{1, \dots, n\}$ :

135 (i)  $u^i \mapsto \mathbb{E}(G_i(u^i))$  is strictly convex.

136 (ii) there exists a positive integer  $p$  such that  $F_i$  has  $p$ -polynomial growth, i.e  
137 there exists  $K > 0$  such that for any  $x^i \in \mathcal{C}([0, T], \mathbb{R})$  and  $u^i \in L^2(0, T)$ :  
138  $|F_i(u^i, x^i)| \leq K(1 + \sup_{0 \leq t \leq T} |x_t^i|^p)$ .

139 Assumption 2.3.(1) holds in different cases, like in the example below.

140 *Example 2.4.* For any  $i \in \{1, \dots, n\}$ , there exists  $g_i : L^2(0, T) \rightarrow \mathbb{R}$  and  $h_i : \mathcal{C}[0, T] \rightarrow \mathbb{R}$   
141 such that for any  $(v, X) \in L^2(0, T) \times \mathcal{C}[0, T]$ ,  $F_i(v, X) = g_i(v) + h_i(X)$   
142 and there exists five  $L^\infty([0, T])$  scalar functions  $\alpha_i, \beta_i, \gamma_i, \xi_i$  and  $\theta_i$  such that for any  
143  $(t, \nu, x) \in [0, T] \times [-M, M] \times \mathbb{R}$ :

$$144 \quad (2.6) \quad \mu_i(t, \nu, x) = \alpha_i(t)\nu + \beta_i(t)x + \gamma_i(t) \text{ and } \sigma_i(x, t) = \xi_i(t)x + \theta_i(t).$$

145 Then Assumption 2.3.(i) is satisfied if:

146 (i)  $g_i$  is strictly convex and  $h_i$  convex.

147 (ii) for a.e.  $t \in [0, T]$ ,  $\alpha(t) \neq 0$ ,  $g_i$  is convex and  $h_i$  strictly convex.

148 Indeed, for any  $i \in \{1, \dots, n\}$ ,  $u, v \in \mathbb{U}$ ,  $\delta \in [0, 1]$  and  $t \in [0, 1]$ , it holds from (2.6)

149 that  $X_t^{i, \delta u + (1-\delta)v} = \delta X_t^{i, u} + (1-\delta)X_t^{i, v}$ . If point (i) holds, then:

$$150 \quad (2.7) \quad h_i(X^{i, \delta u + (1-\delta)v}) \leq \delta h_i(X^{i, u}) + (1-\delta)h_i(X^{i, v}).$$

151 Assumption 2.3.(i) follows from (2.7) and strict convexity of  $g_i$ .

152 Similarly, if point (ii) holds, then the inequality in (2.7) is strict, and Assumption  
153 2.3.(i) follows using also the convexity of  $g_i$ .

154 *Remark 2.5.* If for any  $i \in \{1, \dots, n\}$ , Assumption 2.3.(i) holds, then  $G_i$  is w.l.s.c.  
155 on  $\mathcal{U}_i$ . Indeed,  $G_i$  being convex, finite valued and bounded on bounded subsets  
156 of  $L^2(0, T)$  (from the polynomial growth of  $F_i$  and the inequality (2.4)), thus  $G_i$  is  
157 continuous on  $\mathcal{U}_i$ .

158 From now on, Assumptions 2.1 and 2.3 are in force in the sequel.

159 The following lemma ensures the well-posedness of  $(P_1)$ .

160 LEMMA 2.6. *Suppose that  $F_0$  is convex. Then  $J$  reaches its minimum over  $\mathcal{U}$  at*  
161 *a unique point.*

162 *Proof.* Clearly the control  $\hat{u} \in \mathcal{U}$  defined in (2.5) is feasible. The existence and  
163 uniqueness of a minimum is proved by considering a minimizing sequence  $\{u_k\}$  of  
164  $J$  over  $\mathcal{U}$ . The set  $\mathcal{U}$  being bounded and weakly closed, there exists a sub-sequence  
165  $\{u_{k_\ell}\}$  which weakly converges to a certain  $u^* \in \mathcal{U}$ . Using Assumptions 2.3.(i)(ii) and  
166 convexity of  $F_0$ , it follows that  $\liminf J(u_{k_\ell}) \geq J(u^*)$  and thus  $u^*$  is a solution of  
167  $(P_1)$ . The uniqueness is due to the strict convexity of  $u^i \mapsto \mathbb{E}(F_i(u^i, X^{i, u^i}))$ .  $\square$

168 *Remark 2.7.* This kind of stochastic optimization problem is illustrated in Section  
169 7 with a problem of coordination of a large population of domestic appliances, where  
170 a system operator has to meet the demand while producing at low cost. The state  $X_t^i$   
171 can represent for instance the temperature or the battery level of the agent  $i$  at time  $t$ ,  
172 and  $u_t^i$  its proper power generation or consumption.  $F_0$  can be assimilated to the cost  
173 function to satisfy the demand, and for any  $i$ ,  $F_i$  to the cost function connected to the  
174 proper functioning of the TCLs (characterized by individual cost function, comfort  
175 constraints, etc...).

176 **3. Approximating the optimization problem.** In this section, the link be-  
177 tween problems  $(P_1)$  and  $(P_2)$  is analyzed.

178 *Assumption 3.1.* Problem  $(P_2)$  admits a unique solution.

179 Notice that by using the same techniques as for Lemma 2.6, one can prove that the  
180 above assumption is satisfied when  $F_0$  is convex.

181 We have the following key result.

182 THEOREM 3.2. *Under Assumption 3.1,  $\tilde{J}$  reaches its minimum over  $\mathcal{U}$  at a unique*  
183 *point,  $\tilde{u} \in \mathcal{U}$ , such that for any  $i$ ,  $\tilde{u}^i$  is  $\mathcal{F}^i$ -adapted and thus for any  $j \neq i$ ,  $\tilde{u}^i$  and  $\tilde{u}^j$*   
184 *are mutually independent.*

185 *Proof.* Fix  $i \in \{1, \dots, n\}$ , using the convexity and l.s.c properties of  
186  $u^i \mapsto \mathbb{E}(G_i(u^i))$  and Jensen's inequality, we get:

$$187 \quad (3.1) \quad \mathbb{E}(G_i(u^i) | W^i) \geq G_i(\mathbb{E}(u^i | W^i)).$$

188 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-  
 189 ditional expectation, thus  $F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i)\right) = F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(\mathbb{E}(u^i|W^i))\right)$ . From As-  
 190 sumption 3.2, we know that there exists a solution  $u^*$  to  $(P_2)$ .

191 We set  $\tilde{u} := (\mathbb{E}(u^{*1}|W^1), \mathbb{E}(u^{*2}|W^2), \dots, \mathbb{E}(u^{*n}|W^n))$ . For any  $i$ ,  $\tilde{u}^i :=$   
 192  $\mathbb{E}(u^{*i}|W^i)$  is  $\mathcal{F}_i$ -adapted. Using the definition of  $u^*$  and (3.1), one can derive that  
 193  $\inf_{u \in \mathcal{U}} \tilde{J}(u) = \tilde{J}(u^*) \geq \tilde{J}(\tilde{u})$ .  $\square$

194 Let  $\hat{\mathcal{U}}$  be a subset of  $\mathcal{U}$  associated to decentralized controls, in the sense that:

$$195 \quad (3.2) \quad \hat{\mathcal{U}} := \{u \in \mathcal{U} \mid u^i \text{ is } \mathcal{F}^i\text{-adapted for all } i \in \{1, \dots, n\}\}$$

196 From Theorem 3.2, if Assumption 3.1 holds, then:

$$197 \quad (3.3) \quad \min_{u \in \hat{\mathcal{U}}} \tilde{J}(u) = \min_{u \in \mathcal{U}} \tilde{J}(u).$$

198

199 *Remark 3.3.* If Assumption 3.1 isn't satisfied, we can prove by same arguments  
 200 that for any  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal solution such that the individual controls  
 201 are mutually independent.

202 **LEMMA 3.4.** *If  $F_0$  is Lipschitz with constant  $\gamma$ , then an optimal solution in  $\hat{\mathcal{U}}$  of*  
 203 *problem  $(P_2)$  is an  $\varepsilon$ -optimal solution in  $\hat{\mathcal{U}}$  of problem  $(P_1)$ , with  $\varepsilon = 2\gamma M \sqrt{T/n}$ .*

204 *Proof.* Indeed, there exists a number  $\gamma$  such that  $\gamma > 0$  and for all  $x, y \in H_1$  we  
 205 have  $|F_0(x) - F_0(y)| < \gamma \|x - y\|_{H_1}$ . We set for any  $u \in \mathcal{U}$ :

$$206 \quad (3.4) \quad \hat{u}^i := u^i - \mathbb{E}(u^i).$$

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

$$210 \quad (3.5) \quad \begin{aligned} \left| \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) - F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) \right) \right| &\leq \mathbb{E} \left( \left| F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) - F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) \right| \right) \\ &\leq \frac{\gamma}{n} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{L^2(0,T)} \right) \\ &\leq \frac{\gamma}{n} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{L^2(0,T)}^2 \right)^{\frac{1}{2}} \\ &= \frac{\gamma}{n} \left( \int_0^T \text{Var} \left( \sum_{i=1}^n u_t^i \right) dt \right)^{\frac{1}{2}} \\ &\leq \frac{\gamma}{n^{\frac{1}{2}}} M \sqrt{T}. \end{aligned}$$

211 Let  $\tilde{u}^*$  denote a minimizer of  $\tilde{J}$  on  $\hat{\mathcal{U}}$ , then using (3.5) for any  $u \in \hat{\mathcal{U}}$  it holds:

$$212 \quad (3.6) \quad J(\tilde{u}^*) \leq \tilde{J}(\tilde{u}^*) + \frac{\gamma}{n^{\frac{1}{2}}} M \sqrt{T} \leq \tilde{J}(u) + \frac{\gamma}{n^{\frac{1}{2}}} M \sqrt{T} \leq J(u) + \frac{2\gamma}{n^{\frac{1}{2}}} M \sqrt{T}. \quad \square$$

213 *Assumption 3.5.*  $F_0$  is Gâteaux differentiable with  $c$ -Lipschitz derivative.

214 **THEOREM 3.6.** *Suppose  $F_0$  is convex, then the following  $\varepsilon$ -optimality results hold:*

215 (i) *For any  $u \in \mathcal{U}$ ,  $\tilde{J}(u) \leq J(u)$ .*

216 (ii) *Suppose Assumption 3.5 holds, then any optimal solution of problem  $(P_2)$  is*  
 217 *an  $\varepsilon$ -optimal solution (where  $\varepsilon = 2cTM^2/n$ ) of problem  $(P_1)$ .*

218 *Proof.* Proof of point (i).

219 By Jensen's inequality, we have that:

$$220 \quad (3.7) \quad F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i)\right) \leq \mathbb{E}\left(F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right)\right), \forall u \in \mathcal{U},$$

221 which gives the result.

222 Proof of point (ii).

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

$$\begin{aligned} & F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right) \\ & \leq \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right), \sum_{i=1}^n \hat{u}^i \rangle_{L^2(0,T)} \\ 225 \quad (3.8) \quad & = \frac{1}{n} \langle (\nabla F_0\left(\frac{1}{n} \sum_{i=1}^n u^i\right) - \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right)), \sum_{i=1}^n \hat{u}^i \rangle_{L^2(0,T)} \\ & \quad + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i)\right), \sum_{i=1}^n \hat{u}^i \rangle_{L^2(0,T)} \\ & \leq \frac{c}{n^2} \left\| \sum_{i=1}^n \hat{u}^i \right\|_{L^2(0,T)}^2 + \frac{1}{n} \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i)\right), \sum_{i=1}^n \hat{u}^i \rangle_{L^2(0,T)}, \end{aligned}$$

where  $\hat{u}^i$  is defined in (3.4). Taking the expectation of (3.8),

$$\mathbb{E} \left( \langle \nabla F_0\left(\frac{1}{n} \sum_{i=1}^n \mathbb{E}[u^i]\right), \sum_{i=1}^n \hat{u}^i \rangle_{L^2(0,T)} \right) = 0$$

226 , and using the mutual independence of the controls and their boundedness we get as  
 227 in (3.5):

$$228 \quad (3.9) \quad \frac{c}{n^2} \mathbb{E} \left( \left\| \sum_{i=1}^n \hat{u}^i \right\|_{L^2(0,T)}^2 \right) = \frac{c}{n^2} \int_0^T \sum_{i=1}^n \text{Var}(u_i^t) dt \leq \frac{c}{n} TM^2.$$

229 Let  $\tilde{u}^*$  denote a minimizer of  $\tilde{J}$  on  $\hat{\mathcal{U}}$ , then using (3.3), (3.9) and (3.7), for any  $u' \in \mathcal{U}$   
 230 we have:

$$231 \quad (3.10) \quad J(\tilde{u}^*) \leq \tilde{J}(\tilde{u}^*) + \frac{c}{n} TM^2 \leq \tilde{J}(u') + \frac{c}{n} TM^2 \leq J(u') + \frac{2c}{n} TM^2.$$

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

234 **PROPOSITION 3.7.** *If  $F_0$  is convex, then we have the following inequalities:*

$$235 \quad (3.11) \quad J(\tilde{u}) - \tilde{J}(\tilde{u}) \geq J(\tilde{u}) - J(u^*) \geq 0,$$

236 where  $\tilde{u}$  and  $u^*$  are respectively the optimal controls of problems  $(P_2)$  and  $(P_1)$ .



237 *Proof.* From Jensen inequality and by definition of  $\tilde{u}$  we have:

$$238 \quad (3.12) \quad J(u^*) \geq \tilde{J}(u^*) \geq \tilde{J}(\tilde{u}),$$

239 therefore from the two previous inequalities and adding  $J(\tilde{u})$  we get (3.12).  $\square$

240 *Remark 3.8.* An approximation scheme to compute  $\tilde{u}$  is provided in Section 5.  
 241 The practical interest of inequality (3.11) is that one can compute an upper bound  
 242 for the error  $J(\tilde{u}) - J(u^*)$ , that can be automatically derived from this approximation.

243 **4. Dualization and Decentralization of problem  $(P_2)$ .** From now on, in  
 244 addition to Assumptions 2.1 and 2.3, the assumption that  $F_0$  is convex is in force in  
 245 the sequel. The problem  $(P_2)$  defined in (1.2) is dualized in order to decouple the  
 246 controls in this problem.

247

248 The optimization problem  $(P_2)$  is equivalent to:

$$249 \quad (4.1) \quad (P_3) \quad \begin{cases} \min_{u \in \mathcal{U}, v \in \mathcal{V}} \bar{J}(u, v), \\ \bar{J}(u, v) := F_0(v) + \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n F_i(u^i, X^{i, u^i}) \right), \\ \text{s.t. } g(u, v) = 0, \end{cases}$$

250 where  $g(u, v) := \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) - v$  and  $\mathcal{V} := \{v \in L^2(0, T); |v(t)| \leq 2M \forall t \in [0, T]\}$ .

251 The *Lagrangian* function associated with the constrained optimization problem  $(P_3)$   
 252 is:  $L : \mathcal{U} \times L^2(0, T) \times L^2(0, T) \rightarrow \bar{\mathbb{R}}$  defined by:

$$253 \quad (4.2) \quad L(u, v, \lambda) := \bar{J}(u, v) + \langle \lambda, \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) - v \rangle_{L^2(0, T)}.$$

254 The dual problem  $(D)$  associated with  $(P_3)$  is:

$$255 \quad (4.3) \quad (D) \quad \max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda), \quad \text{where } \mathcal{W}(\lambda) := \min_{u \in \mathcal{U}, v \in \mathcal{V}} L(u, v, \lambda).$$

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

$$258 \quad (4.4) \quad \text{There exists } \varepsilon > 0 \text{ such that } \mathcal{B}_{L^2(0, T)}(0, \varepsilon) \subset g(\mathcal{U}, \mathcal{V}),$$

259 where  $\mathcal{B}_{L^2(0, T)}(0, \varepsilon)$  is the open ball of radius  $\varepsilon$  in  $L^2(0, T)$  and  $g$  has been defined in  
 260 (4.1).

261 **LEMMA 4.1.** *Problem  $(P_3)$  is qualified.*

262 *Proof.* Choose  $\varepsilon := M$ . Then:  $\mathcal{B}_{L^2(0, T)}(0, \varepsilon) \subset \overline{\mathcal{B}_{L^2(0, T)}(0, 2M)} = g(0, \mathcal{V}) \subset$   
 263  $g(\mathcal{U}, \mathcal{V})$ , where  $g(0, \mathcal{V})$  and  $g(\mathcal{U}, \mathcal{V})$  are respectively the image by  $g$  of  $\{0\} \times \mathcal{V}$  and  
 264  $\mathcal{U} \times \mathcal{V}$ . The conclusion follows.  $\square$

265 By Assumption 2.3, Lemma 4.1 and the convexity of  $F_0$ , the strong duality holds:  
 266  $\mathcal{W}(\lambda^*) = \tilde{J}(u^*)$ , where  $\lambda^* \in \arg \max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda)$  and  $u^* \in \arg \min_{u \in \mathcal{U}, v \in \mathcal{V}} L(\lambda^*, u, v)$ .

Since the set of admissible controls  $\mathcal{U} = \mathcal{U}_1 \times \dots \times \mathcal{U}_n$  is a Cartesian product and by strict convexity of  $u^i \mapsto \mathbb{E}(F_i(u^i, X^{i,u^i}))$ , each component  $u^{*i}$  can be uniquely determined by solving the following sub problem:

$$u^{*i} = \arg \min_{u^i \in \mathcal{U}_i} \left\{ \mathbb{E} \left( F_i(u^i, X^{i,u^i}) + \langle \lambda^*, u^i \rangle_{L^2(0,T)} \right) \right\},$$

267 where  $\hat{\mathcal{U}}_i := \{u \in \mathcal{U}_i \mid u^i \text{ is } \mathcal{F}^i\text{-adapted}\}$ .

268 *Remark 4.2.* By using the same argument as in Theorem 3.2, one can prove:

$$(4.5) \quad \begin{aligned} & \min_{u^i \in \hat{\mathcal{U}}_i} \left\{ \mathbb{E} \left( F_i(u^i, X^{i,u^i}) + \langle \lambda^*, u^i \rangle_{L^2(0,T)} \right) \right\} \\ & = \min_{u^i \in \mathcal{U}_i} \left\{ \mathbb{E} \left( F_i(u^i, X^{i,u^i}) + \langle \lambda^*, u^i \rangle_{L^2(0,T)} \right) \right\}. \end{aligned}$$

## 270 5. Stochastic Uzawa and Sampled Stochastic Uzawa algorithms.

271 **5.1. Continuous time setting.** We recall that Assumptions 2.1 and 2.3 are in  
272 force, as well as convexity of  $F_0$ .

273 This section aims at proposing an algorithm to find a solution of the dual problem  
274 (4.3).

275 For all  $i \in \{1, \dots, n\}$ , and  $\lambda \in L^2(0, T)$ , we define the optimal control  $u^i(\lambda)$ :

$$(5.1) \quad u^i(\lambda) := \arg \min_{u^i \in \hat{\mathcal{U}}_i} \left\{ \mathbb{E} \left( F_i(u^i, X^{i,u^i}) + \langle \lambda, u^i \rangle_{L^2(0,T)} \right) \right\},$$

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

278 For any  $\lambda \in L^2(0, T)$ , the subset  $V(\lambda)$  is defined by:

$$(5.2) \quad V(\lambda) := \arg \min_{v \in \mathcal{V}} \{F_0(v) - \langle \lambda, v \rangle_{L^2(0,T)}\}.$$

280 Since  $F_0$  is convex and  $\mathcal{V}$  is bounded,  $V(\lambda)$  is a non empty subset of  $\mathcal{V}$  and is reduced  
281 to a singleton if  $F_0$  is strictly convex.

282 For any  $\lambda \in L^2(0, T)$ , a function  $v(\lambda)$ , which is a selection of  $V(\lambda)$ , is associated.

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

288 For any real valued function  $F$  defined on  $L^2(0, T)$ ,  $F^*$  stands for its Fenchel  
289 conjugate.

290 **LEMMA 5.1.** *Assumption 3.5 holds iff  $F_0^*$  is proper and strongly convex.*

291 *Proof.* (i) Let Assumption 3.5 hold. Since  $F_0$  is proper, convex and l.s.c.,  $F_0^*$   
292 is l.s.c. proper. From the Lipschitz property of the gradient of  $F_0$ , it holds that  
293  $\text{dom}(F_0) = L^2(0, T)$ .

294 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  
295 the differentiability, l.s.c. and convexity of  $F_0$ , it follows that:  $s = \nabla F_0(\lambda_s)$  and  
296  $\tilde{s} = \nabla F_0(\mu_{\tilde{s}})$ . By Assumption 3.5 and the extended Baillon-Haddad theorem [16,

297 Theorem 3.1],  $\nabla F_0$  is cocoervice. In other words:

$$\begin{aligned}
(5.3) \quad \langle s - \tilde{s}, \lambda_s - \mu_{\tilde{s}} \rangle_{L^2(0,T)} &= \langle \nabla F_0(\lambda_s) - \nabla F_0(\mu_{\tilde{s}}), \lambda_s - \mu_{\tilde{s}} \rangle_{L^2(0,T)} \\
&\geq \frac{1}{c} \|\nabla F_0(\lambda_s) - \nabla F_0(\mu_{\tilde{s}})\|_{L^2(0,T)}^2 \\
&= \frac{1}{c} \|s - \tilde{s}\|_{L^2(0,T)}^2.
\end{aligned}$$

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

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

$$(5.4) \quad F_0(\lambda) \leq \sup_{s \in L^2(0,T)} \langle s, \lambda \rangle_{L^2(0,T)} - \alpha \|s\|_{L^2(0,T)}^2 + \beta = \|\lambda\|^2 / \alpha + \beta.$$

304 Thus  $F_0$  is proper and uniformly upper bounded over bounded sets and therefore is  
305 locally Lipschitz. In addition, from the strong convexity of  $F_0^*$  and the convexity  
306 of  $F_0$ , for any  $\lambda \in L^2(0, T)$ ,  $\partial F_0(\lambda)$  is a singleton. Thus  $F_0$  is everywhere Gâteaux  
307 differentiable.

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

$$(5.5) \quad \langle \mu - \lambda, \nabla F_0(\mu) - \nabla F_0(\lambda) \rangle_{L^2(0,T)} \geq \frac{1}{c} \|\nabla F_0(\mu) - \nabla F_0(\lambda)\|_{L^2(0,T)}^2,$$

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

317 LEMMA 5.2. *If Assumption 3.5 holds, then  $\mathcal{W}$  is strongly concave.*

318 *Proof.* For any  $\lambda \in L^2(0, T)$ , we have:  $\mathcal{W}(\lambda) = -(F_0^*(\lambda) + \frac{1}{n} \sum_{i=1}^n F_i^*(\lambda))$ , where  
319 for any  $i \in \{1, \dots, n\}$ ,  $F_i^*$  is convex and from Lemma 5.1  $F_0^*$  is strongly convex.  $\square$

320 We introduce the function  $f : L^2(0, T) \rightarrow L^2(0, T)$  where for any  $\lambda \in L^2(0, T)$ :

$$(5.6) \quad f(\lambda) := \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i(\lambda)) - v(\lambda).$$

322 From the boundedness of  $\mathcal{U}$  and  $\mathcal{V}$ , it easily follows that there exists a finite positive  
323 real  $M_1$  such that for any  $\lambda \in L^2(0, T)$ :

$$(5.7) \quad \|f(\lambda)\|_{L^2(0,T)}^2 \leq M_1.$$

325 For any  $\lambda \in L^2(0, T)$ , we denote by  $\partial(-\mathcal{W}(\lambda))$  the subgradient of  $-\mathcal{W}$  at  $\lambda$ . Therefore  
326 for any  $\lambda \in L^2(0, T)$ :

$$(5.8) \quad \partial(-\mathcal{W}(\lambda)) \ni -f(\lambda).$$

328 The iterative algorithm, proposed as an approximation scheme for  $\lambda^* \in \arg \max_{\lambda} \mathcal{W}(\lambda)$ ,  
329 is summarized in the *Stochastic Uzawa Algorithm 5.1*.

---

**Algorithm 5.1** Stochastic Uzawa
 

---

- 1: Initialization  $\lambda^0 \in L^2(0, T)$ , set  $\{\rho_k\}$  satisfying Assumption 5.4.
  - 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 (5.2).
  - 5:      $u^{i,k} \leftarrow u^i(\lambda^k)$  where  $u^i(\lambda^k)$  is defined in (5.1) for any  $i \in \{1, \dots, n\}$ .
  - 6:     Generate  $n$  independent realizations of Brownian motions  $(W^{1,k+1}, \dots, W^{n,k+1})$ , independent also with  $\{W^{i,p} : 1 \leq i \leq n, p \leq k\}$ .
  - 7:     Compute the associated state realizations  $(X^{1,u^1(\lambda^k)}, \dots, X^{n,u^n(\lambda^k)})$ .
  - 8:      $Y^{k+1} \leftarrow \frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(W^{i,k+1}) - v(\lambda^k)$ .
  - 9:      $\lambda^{k+1} \leftarrow \lambda^k + \rho_k Y^{k+1}$ .
- 

330     *Remark 5.3.* For the purpose of notation,  $u^i(\lambda^k)(W^{i,k+1})$  in (8) corresponds to  
 331 the realization of  $u^i(\lambda^k)$  resulting from a realization of the Brownian  $W^{i,k+1}$ .

332     At any dual iteration  $k$  of Algorithm 5.1,  $Y^{k+1}$  is an estimator of  
 333  $\mathbb{E}(\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(W^{i,k+1}) - v(\lambda^k))$ . Therefore an alternative approach proposed in the  
 334 *Sampled Stochastic Uzawa* Algorithm 5.2 consists in performing less simulations at  
 335 each iteration, by taking  $m < n$ , at the risk of performing more dual iterations, to  
 estimate the quantity  $\mathbb{E}(\frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(W^{i,k+1}) - v(\lambda^k))$ .

---

**Algorithm 5.2** Sampled Stochastic Uzawa
 

---

- 1: Initialization of  $m$  a positive integer and  $\check{\lambda}^0 \in L^2(0, T)$ , set  $\{\rho_k\}$  satisfying Assumption 5.4.
  - 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 (5.2).
  - 5:     Generate  $m$  i.i.d discrete random variables  $I_1^k, \dots, I_m^k$  uniformly in  $\{1, \dots, n\}$ .
  - 6:      $u^{I_j^k,k} \leftarrow u^{I_j^k}(\check{\lambda}^k)$  where  $u^{I_j^k}(\check{\lambda}^k)$  is defined in (5.1) for any  $j \in \{1, \dots, m\}$ .
  - 7:     Generate  $m$  independent realizations of Brownian motions  $(W^{I_1^k,k+1}, \dots, W^{I_m^k,k+1})$ , independent also with  $\{W^{i,p} : 1 \leq i \leq m, p \leq k\}$ .
  - 8:     Compute the associated state realizations  $(X^{I_1^k,u^{I_1^k}(\check{\lambda}^k)}, \dots, X^{I_m^k,u^{I_m^k}(\check{\lambda}^k)})$ .
  - 9:      $\check{Y}^{k+1} \leftarrow \frac{1}{m} \sum_{j=1}^m u^{I_j^k}(\check{\lambda}^k)(W^{I_j^k,k+1}) - v(\check{\lambda}^k)$
  - 10:      $\check{\lambda}^{k+1} \leftarrow \check{\lambda}^k + \rho_k \check{Y}^{k+1}$ .
- 

336     The complexity of the *Stochastic Uzawa* Algorithm 5.2 is proportional to  $m \times K$ ,  
 337 where  $K$  is the total number of dual iterations and  $m$  the number of simulations  
 338 performed at each iteration. The error  $\mathbb{E}(\|\lambda^{k+1} - \lambda^*\|^2)$  for  $\lambda^* \in S$  is the sum of  
 339 the square of the bias (which only depends on  $K$  and not on  $m$ ) and the variance  
 340 (which both depends on  $K$  and  $m$ ). Therefore this algorithm enables a bias variance  
 341 trade-off for a given complexity. Similarly for a given error it enables to optimize the  
 342 complexity of the algorithm.  
 343

344 Some assumptions on the step size are introduced.

345 *Assumption 5.4.* The sequence  $(\rho_k)_k$  is such that:  $\rho_k > 0$ ,  $\sum_{k=1}^{\infty} \rho_k = \infty$  and

346 
$$\sum_{k=1}^{\infty} (\rho_k)^2 < \infty.$$

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

348 *Assumption 5.4.*

349 Let us denote  $S := \arg \max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda)$ , where  $S$  is nonempty because of the strong

350 convexity of  $\mathcal{W}$ .

351 The following result establishes the convergence of the *Stochastic Uzawa Algo-*  
352 *rithm 5.1:*

353 **THEOREM 5.5.** *Let Assumption 5.4 hold, then:*

354 (i)  $\{\|\lambda^k - \lambda\|_{L^2(0, T)}^2\}$  converges a.s., for all  $\lambda \in S$ .

355 (ii)  $\mathcal{W}(\lambda^k) \xrightarrow[k \rightarrow \infty]{} \max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda)$  a.s.

356 (iii)  $\{\lambda^k\}$  weakly converges to some  $\bar{\lambda} \in S$  in  $L^2(0, T)$  a.s.

357 (iv) If *Assumption 3.5* holds, then a.s.  $\{\lambda^k\}$  converges to  $\bar{\lambda}$  in  $L^2(0, T)$ , with  
358  $S := \{\bar{\lambda}\}$ .

359 Though the proof is similar to [6, Theorem 3.6], the current framework is different  
360 from the one of that reference, and for the convenience of the reader we provide the  
361 proof.

362 We first state two lemmas.

363 **LEMMA 5.6** (Robbins-Siegmund). *Let  $\{\mathcal{G}_k\}$  be an increasing sequence of  $\sigma$ -*  
364 *algebra and  $d_k, a_k, b_k$  and  $c_k$  be nonnegative random variables adapted to  $\mathcal{G}_k$ . Assume*  
365 *that:  $\mathbb{E}(d_{k+1}|\mathcal{G}_k) \leq d_k(1 + a_k) + b_k - c_k$  and  $\sum_{k=1}^{\infty} a_k < \infty$  a.s.,  $\sum_{k=1}^{\infty} b_k < \infty$  a.s. Then*

366 *with probability one,  $\{d_k\}$  is convergent and it holds that  $\sum_{k=1}^{\infty} c_k < \infty$ .*

367 *Proof.* See [5], Theorem 1.3.12. □

368 **LEMMA 5.7.** *Let  $\{\alpha_k\}$  be a nonnegative deterministic sequence and  $\{\beta_k\}$  a non-*  
369 *negative random sequence adapted to  $\{\mathcal{G}_k\}$ . Assume that  $\sum_{k=1}^{\infty} \alpha_k = \infty$  a.s. and*

370  $\mathbb{E}(\sum_{k=1}^{\infty} \alpha_k \beta_k) < \infty$  a.s. *Moreover assume that  $\beta_k - \mathbb{E}(\beta_{k+1}|\mathcal{G}_k) \leq c\alpha_k$  a.s. for all*

371  $k$  and some  $c > 0$ . *Then  $\beta_k \xrightarrow{a.s.} 0$ .*

372 *Proof.* See [6], Proposition 3.2. □

373 *Proof of Theorem 5.5.* First consider point (i). Let  $\lambda \in S$ . For any  $k$ ,  $\mathcal{G}_{k+1}$  is the  
374 filtration defined by:

375 (5.9) 
$$\mathcal{G}_{k+1} := \sigma(\{W^{i,p} : 1 \leq i \leq n, p \leq k+1\}).$$

376 Using the definition of  $Y^{k+1} \in L^2(0, T)$  line 8 in the *Stochastic Uzawa Algorithm 5.1*,

377 we have:

$$\begin{aligned}
 & \|\lambda^{k+1} - \lambda\|_{L^2(0,T)}^2 = \|\lambda^k + \rho_k Y^{k+1} - \lambda\|_{L^2(0,T)}^2 \\
 378 \quad (5.10) \quad & = \|\lambda^k - \lambda\|_{L^2(0,T)}^2 + 2\rho_k \langle \lambda^k - \lambda, Y^{k+1} \rangle_{L^2(0,T)} \\
 & \quad + (\rho_k)^2 \|Y^{k+1}\|_{L^2(0,T)}^2.
 \end{aligned}$$

379 Since  $Y^{k+1}$  is independent from  $\mathcal{G}^k$ , and using (5.7), it follows that:

$$380 \quad (5.11) \quad \mathbb{E}(\|Y^{k+1}\|_{L^2(0,T)}^2 | \mathcal{G}_k) = \mathbb{E} \left( \left\| \frac{1}{n} \sum_{i=1}^n u^i(\lambda^k)(W^{i,k+1}) - v(\lambda^k) \right\|_{L^2(0,T)}^2 \right) \leq M_1$$

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

$$\begin{aligned}
 & \mathbb{E}[\|\lambda^{k+1} - \lambda\|_{L^2(0,T)}^2 | \mathcal{G}_k] \\
 382 \quad (5.12) \quad & = \|\lambda^k - \lambda\|_{L^2(0,T)}^2 + 2\rho_k \mathbb{E}(\langle \lambda^k - \lambda, Y^{k+1} \rangle | \mathcal{G}_k) + (\rho_k)^2 \mathbb{E}[\|Y^{k+1}\|_{L^2(0,T)}^2 | \mathcal{G}_k] \\
 & \leq \|\lambda^k - \lambda\|_{L^2(0,T)}^2 + 2\rho_k \langle \lambda^k - \lambda, f(\lambda^k) \rangle + (\rho_k)^2 M_1 \\
 & \leq \|\lambda^k - \lambda\|_{L^2(0,T)}^2 + (\rho_k)^2 M_1 - 2\rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)).
 \end{aligned}$$

383 In the last inequality, we used the concavity of  $\mathcal{W}$  and (5.8). We set:

$$384 \quad (5.13) \quad a_k = 0, b_k = (\rho_k)^2 M_1, c_k = 2\rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)),$$

385 We have that  $\sum_{k=1}^{\infty} a_k < \infty$  a.s. and  $\sum_{k=1}^{\infty} b_k < \infty$  a.s. Clearly,  $a_k$  and  $b_k$  are nonnega-  
 386 tive;  $c_k$  is nonnegative since  $\lambda \in S$ . By Lemma 5.6, the sequence  $\{\|\lambda^k - \lambda\|_{L^2(0,T)}^2\}$   
 387 converges a.s. Now we show point (ii) thanks to Lemma 5.7.

388 By Lemma 5.6:  $\sum_{k=1}^{\infty} \rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)) < \infty$  a.s. Taking the expected value in  
 389 both side of (5.12), we get, using the deterministic version of Lemma 5.6 that:  
 390  $\mathbb{E} \left( \sum_{k=1}^{\infty} \rho_k (\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)) \right) < \infty$ . By concavity of  $\mathcal{W}$  and the Cauchy-Schwarz in-  
 391 equality, we have:

$$392 \quad (5.14) \quad \mathcal{W}(\lambda^{k+1}) - \mathcal{W}(\lambda^k) \leq \langle f(\lambda^k), \lambda^{k+1} - \lambda^k \rangle \leq \rho_k \|f(\lambda^k)\| \|Y^{k+1}\|.$$

393 Let  $\tau_M$  be the stopping time  $\tau_M := \inf\{k : \|\lambda^k\| > M\}$  for  $M \in \mathbb{N}$ .

394 The sequence  $\{\beta_k\}$  is defined by:

$$395 \quad (5.15) \quad \beta_k := \begin{cases} \mathcal{W}(\lambda) - \mathcal{W}(\lambda^k) & \text{if } \{\tau_M > k\}, \\ \mathcal{W}(\lambda) - \mathcal{W}(\lambda^{\tau_M}) & \text{otherwise, with } \beta_{\tau_M+k} = \beta_{\tau_M}, k \geq 1. \end{cases}$$

396 Notice that if  $\|\lambda^k\| \leq M$ , there exists by (5.11) and (5.7)  $M' > 0$  such that

$$397 \quad (5.16) \quad \|f(\lambda^k)\| \mathbb{E}(\|Y^{k+1}\| | \mathcal{G}_k) \leq (M')^2.$$

398 Now:  $\beta_k - \beta_{k+1} = \mathbb{1}_{\tau_M > k} (\mathcal{W}(\lambda^{k+1}) - \mathcal{W}(\lambda^k))$ , and therefore by taking the conditional  
 399 expectation on both sides, noticing that  $\mathbb{1}_{\tau_M > k}$  is  $\mathcal{G}_k$ -measurable, and considering

400 (5.14) and (5.16), we get  $\beta_k - \mathbb{E}(\beta_{k+1}|\mathcal{G}_k) \leq \rho_k(M')^2$ .

401 By Lemma 5.7, on the set  $B_M := \{\tau_M = \infty\}$ ,  $\beta_k$  converges to 0 and coincides with  
 402  $\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)$ . Since  $\|\lambda^k - \lambda\|$  converges a.s.,  $\|\lambda^k\|$  is bounded in probability and  
 403 therefore the probability of the set  $B_M$  can be made arbitrarily close to 1 by choosing  
 404  $M$  large. Since  $\mathbb{P}(\cup_{M=1}^{\infty} B_M) = 1$ , we may infer that  $\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)$  converges to 0  
 405 almost surely.

406

407 For point (ii), since  $\{\|\lambda^k - \lambda\|^2\}$  converges a.s. for all  $\lambda \in S$ , it is bounded  
 408 in probability, so the sequence  $\{\lambda^k\}$  generated by the algorithm has a.s. a weak  
 409 accumulation point  $\bar{\lambda}$  (the point  $\bar{\lambda}$  is random in general). Let  $\{\lambda^{k_m}\}$  such that  $\lambda^{k_m} \rightharpoonup$   
 410  $\bar{\lambda}$ . Since  $\mathcal{W}$  is concave and upper semi-continuous, it is weakly upper semi-continuous,

$$411 \quad (5.17) \quad \mathcal{W}(\bar{\lambda}) \geq \lim_{m \rightarrow \infty} \mathcal{W}(\lambda^{k_m}) = \mathcal{W}(\bar{\lambda}).$$

412 In particular,  $\bar{\lambda} \in S$ . To show uniqueness, let  $\lambda_1, \lambda_2 \in S$  be two distinct weak limits  
 413 of  $\{\lambda^k\}$ , i.e.  $\lambda^{k_m} \rightharpoonup \lambda_1$  and  $\lambda^{k_l} \rightharpoonup \lambda_2$ . Then

$$414 \quad (5.18) \quad \|\lambda^{k_m} - \lambda_2\|^2 = \|\lambda^{k_m} - \lambda_1\|^2 + \|\lambda_1 - \lambda_2\|^2 + 2\langle \lambda^{k_m} - \lambda_1, \lambda_1 - \lambda_2 \rangle,$$

415

$$416 \quad (5.19) \quad \|\lambda^{k_l} - \lambda_1\|^2 = \|\lambda^{k_l} - \lambda_2\|^2 + \|\lambda_2 - \lambda_1\|^2 + 2\langle \lambda^{k_l} - \lambda_2, \lambda_2 - \lambda_1 \rangle,$$

417 so by weak convergence of each subsequence, (5.18) and (5.19) are combined to obtain

$$418 \quad (5.20) \quad \lim_{m \rightarrow \infty} \|\lambda^{k_m} - \lambda_2\|^2 - \|\lambda^{k_m} - \lambda_1\|^2 = \|\lambda_2 - \lambda_1\|^2,$$

419

$$420 \quad (5.21) \quad \lim_{l \rightarrow \infty} \|\lambda^{k_l} - \lambda_1\|^2 - \|\lambda^{k_l} - \lambda_2\|^2 = \|\lambda_2 - \lambda_1\|^2.$$

421 By a.s. convergence of the sequence  $\{\|\lambda^k - \lambda\|^2\}$  for all  $\lambda \in S$ , the limit of each  
 422 subsequence is equal to the limit of the entire sequence with probability one, so

$$423 \quad \lim_{m \rightarrow \infty} \|\lambda^{k_m} - \lambda_1\|^2 = \lim_{k \rightarrow \infty} \|\lambda^k - \lambda_1\|^2 =: l_1 \text{ and similarly } \lim_{m \rightarrow \infty} \|\lambda^{k_m} - \lambda_2\|^2 =$$

$$424 \quad \lim_{k \rightarrow \infty} \|\lambda^k - \lambda_2\|^2 =: l_2. \text{ Therefore (5.20) and (5.21) imply } l_2 - l_1 = \|\lambda_1 - \lambda_2\|^2 = l_1 - l_2,$$

425 meaning  $\|\lambda_1 - \lambda_2\|^2 = 0$  and thus the weak limits coincide. Therefore  $\{\lambda^k\}$  is weakly  
 426 convergent to a unique limit with probability one.

427 Finally, the last statement can now be proved. By strong convexity,  $-\mathcal{W}$  has an  
 428 unique minimum  $\bar{\lambda}$ , so  $S = \{\bar{\lambda}\}$ . By strong convexity, there exists a  $\mu > 0$  such that

$$429 \quad (5.22) \quad \mathcal{W}(\bar{\lambda}) - \mathcal{W}(\lambda^k) \geq -\langle f(\bar{\lambda}), \lambda^k - \bar{\lambda} \rangle_{L^2(0,T)} + \frac{\mu}{2} \|\lambda^k - \bar{\lambda}\|_{L^2(0,T)}^2.$$

430 Since  $-\langle f(\bar{\lambda}), \lambda^k - \bar{\lambda} \rangle_{L^2(0,T)} > 0$ , by optimality of  $\bar{\lambda}$ ,  $\lim_{k \rightarrow \infty} \mathcal{W}(\bar{\lambda}) - \mathcal{W}(\lambda^k) = 0$  a.s.

431 implies  $\lim_{k \rightarrow \infty} \|\lambda^k - \bar{\lambda}\|_{L^2(0,T)} = 0$  a.s.  $\square$

432 We recall the definition of  $\bar{J}(u, v)$  in (4.1) and we define  $\bar{u}$ :

$$433 \quad (5.23) \quad \bar{u} := \arg \min_{u \in \mathcal{U}} \left\{ \mathbb{E} \left( \sum_{i=1}^n F_i(u^i, X^{i,u^i}) + \langle \bar{\lambda}, u^i \rangle_{L^2(0,T)} \right) \right\},$$

434 If  $F_0$  is strictly convex, then we define:

$$435 \quad (5.24) \quad \bar{v} := \arg \min_{v \in \mathcal{V}} \{F_0(v) + \langle \bar{\lambda}, v \rangle_{L^2(0,T)}\}.$$

436 If Assumption 3.5 holds and  $F_0$  is strictly convex,  $(\bar{u}, \bar{v}, \bar{\lambda})$  is a saddle point and  $\bar{u}$  is  
 437 the unique minimizer of  $\bar{J}$  in  $\mathcal{U}$ .

438 THEOREM 5.8. *Let the Assumptions 3.5 and 5.4 hold, then we have:*

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

440 *If  $F_0$  is strictly convex, then:*

441 (ii)  $\tilde{J}(u(\lambda^k)) \xrightarrow[k \rightarrow \infty]{} \tilde{J}(\bar{u})$  a.s.

442 (iii)  $\limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \inf_{u \in \mathcal{U}} J(u) + 2\varepsilon$  a.s. where  $\varepsilon$  is defined in Theorem 3.6.(ii).

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

445 (5.25) 
$$(u(\lambda^{\theta_k}), v(\lambda^{\theta_k})) \xrightarrow[k \rightarrow \infty]{} (u^\theta, v^\theta) \in \mathcal{U} \times \mathcal{V}.$$

446 Using the definition of  $\lambda \mapsto u(\lambda)$  in (5.1), it holds for any  $k > 0$ :

447 (5.26) 
$$\begin{aligned} & \mathbb{E} \left( F_i(\bar{u}^i, X^{i, \bar{u}^i}) + \langle \lambda^{\theta_k}, \bar{u}^i \rangle_{L^2(0, T)} \right) \\ & \geq \mathbb{E} \left( F_i(u^i(\lambda^{\theta_k}), X^{i, u^i(\lambda^{\theta_k})}) + \langle \lambda^{\theta_k}, u^i(\lambda^{\theta_k}) \rangle_{L^2(0, T)} \right). \end{aligned}$$

448 Using that  $u^i \mapsto F_i(u^i(\lambda^k), X^{i, u^i(\lambda^k)})$  is w.l.s.c. (see Remark 2.5) and the a.s. con-  
 449 vergence of  $\{\lambda^k\}$ , resulting from Theorem 5.5.(iv), we have from (5.26) when  $k \rightarrow \infty$   
 450 :

451 (5.27) 
$$\mathbb{E} \left( F_i(\bar{u}^i, X^{i, \bar{u}^i}) + \langle \bar{\lambda}, \bar{u}^i \rangle_{L^2(0, T)} \right) \geq \mathbb{E} \left( F_i(u^{i, \theta}, X^{i, u^{i, \theta}}) + \langle \bar{\lambda}, u^{i, \theta} \rangle_{L^2(0, T)} \right).$$

452 Since  $\bar{u}$  is uniquely defined (see (5.23)), it follows  $u^\theta = \bar{u}$  and (5.27) is an equality.  
 453 Using that every weakly convergent sub sequence of  $\{u(\lambda^k)\}$  has the same weak limit  
 454  $\bar{u}$ , (i) is deduced.

455 *Proof of point (ii).*

456 From point (i) and (5.27), it follows for any  $i \in \{1, \dots, n\}$ :

457 (5.28) 
$$\lim_{k \rightarrow \infty} \mathbb{E} \left( F_i(u^i(\lambda^k), X^{i, u^i(\lambda^k)}) \right) = \mathbb{E} \left( F_i(\bar{u}^i, X^{i, \bar{u}^i}) \right).$$

458 Using 5.25, the w.l.s.c of  $F_0$ , equation (5.24), and applying the same previous argument  
 459 to  $\{v(\lambda^{\theta_k})\}$ , it holds that:

460 (5.29) 
$$\lim_{k \rightarrow \infty} F_0(v(\lambda^k)) - \langle \lambda^k, v(\lambda^k) \rangle_{L^2(0, T)} = F_0(\bar{v}) - \langle \bar{\lambda}, \bar{v} \rangle_{L^2(0, T)},$$

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

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

463 (5.30) 
$$\lim_{k \rightarrow \infty} F_0(v(\lambda^k)) = F_0(\bar{v}).$$

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

465 (5.31) 
$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}(\bar{u}^i) = \bar{v}.$$

466 From (5.30) and (5.31), it holds:

467 (5.32) 
$$\lim_{k \rightarrow \infty} F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i(\lambda^k)) \right) = F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\bar{u}^i) \right).$$



468 Then adding (5.28) and (5.32):  $\lim_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) = \tilde{J}(\bar{u})$ .

469 Proof of point (iii). From point (ii), inequality (3.10) and Theorem 3.6.(ii), it  
470 holds:

$$471 \quad (5.33) \quad \limsup_{k \rightarrow \infty} J(u(\lambda^k)) \leq \limsup_{k \rightarrow \infty} \tilde{J}(u(\lambda^k)) + \varepsilon = \inf_{u \in \mathcal{U}} \tilde{J}(u) + \varepsilon \leq \inf_{u \in \mathcal{U}} J(u) + 2\varepsilon,$$

472 where  $\varepsilon = 2cTM^2/n$ . The conclusion follows.  $\square$

473 *Assumption 5.9.* (i)  $F_0$  is strongly convex.

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

475 **LEMMA 5.10.** *Let Assumption 5.9.(i) hold, then the function  $\lambda \mapsto v(\lambda)$  is Lips-*  
476 *chitz on  $L^2(0, T)$ .*

477 *Proof.* From the definition of  $v$  in (5.2), we have for any  $\lambda \in L^2(0, T)$ :  $\lambda \in$   
478  $\partial F_0(v(\lambda))$ . Thus for any  $\lambda, \mu \in L^2(0, T)$ , we have from the strong convexity of  $F_0$ :

$$479 \quad (5.34) \quad \begin{cases} F_0(v(\mu)) & \geq F_0(v(\lambda)) + \langle \lambda, v(\mu) - v(\lambda) \rangle_{L^2(0, T)} + \alpha \|v(\mu) - v(\lambda)\|_{L^2(0, T)}^2 \\ F_0(v(\lambda)) & \geq F_0(v(\mu)) + \langle \mu, v(\lambda) - v(\mu) \rangle_{L^2(0, T)} + \alpha \|v(\lambda) - v(\mu)\|_{L^2(0, T)}^2. \end{cases}$$

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

$$481 \quad (5.35) \quad \langle \lambda - \mu, v(\lambda) - v(\mu) \rangle_{L^2(0, T)} \geq 2\alpha \|v(\lambda) - v(\mu)\|_{L^2(0, T)}^2.$$

482 Applying Cauchy-Schwarz inequality and simplifying by  $\|v(\lambda) - v(\mu)\|_{L^2(0, T)}$ , we get  
483 the desired Lipschitz inequality.  $\square$

484 **LEMMA 5.11.** *Let Assumption 5.9.(ii) hold, thus the function  $\lambda \mapsto u(\lambda)$  is Lips-*  
485 *chitz on  $L^2(0, T)$ .*

486 *Proof.* The proof is similar to the proof of Lemma 5.10.  $\square$

487 **THEOREM 5.12.** *Let the Assumption 3.5, 5.4, and 5.9 hold, then:  $u(\lambda^k) \xrightarrow[k \rightarrow \infty]{} u(\bar{\lambda})$*   
488 *a.s.*

489 *Proof.* The convergence follows from the Lipschitz property of  $\lambda \mapsto u(\lambda)$  (as a  
490 result of assumption 5.9) associated with the a.s. convergence of  $\{\lambda^k\}$ .  $\square$

491 *Remark 5.13.* Note that Theorems 5.5, 5.8 and 5.12 still hold when replacing  $\lambda^k$   
492 by  $\check{\lambda}^k$  and  $Y^k$  by  $\check{Y}^k$  (defined resp. line 9 and 10 in the *Sampled Stochastic Uzawa*  
493 *Algorithm 5.2*). This can be proved by same argument, using that  $\check{Y}^k$  is bounded a.s.  
494 and  $\mathbb{E}(\check{Y}^k | \check{\mathcal{G}}_k) = f(\check{\lambda}^k)$  for any  $k$ , where:

$$495 \quad (5.36) \quad \check{G}_k = \sigma \left( \{W_{\ell}^{I_{\ell}^p, p}\} : 1 \leq \ell \leq m, p \leq k \right) \vee \sigma \left( \{I_{\ell}^p\} : 1 \leq \ell \leq m, p \leq k \right),$$

496 with  $W_{\ell}^{I_{\ell}^p, p}$  and  $I_{\ell}^k$  defined respectively at lines 7 and 5 of the *Sampled Stochastic*  
497 *Uzawa Algorithm 5.2*.

498 *Remark 5.14.* From a practical point of view, this algorithm can be implemented  
499 in a decentralized way, where the system operator sends the signal  $\lambda$ , which can  
500 be assimilated to a price, to the domestic appliances, which compute their optimal  
501 solution  $u(\lambda)$ , depending on their local parameters.

502 In (2.2), the states and controls of the agents are described in a continuous time  
503 setting with finite horizon. However all the previous results are easy to extend if  
504 we consider a discrete time setting with finite horizon, the proofs using the same  
505 arguments as in continuous time setting.

506 **5.2. Extension to the discrete time setting.** The results of the previous  
 507 sections are extended to the discrete time setting in this subsection.

508 The following notations are used:

- 509 • Let  $n \in \mathbb{N}^*$  be the number of agents and  $T \in \mathbb{N}^*$  the finite time horizon.
- 510 • For any matrix  $M$ ,  $M^\top$  denotes its transpose
- 511 • For any  $i \in \{1, \dots, n\}$ ,  $X^{i,u^i} := (x_0^i, \dots, x_T^i) \in \mathbb{R}^T$  is the state trajectory  
 512 of agent  $i$  controlled by  $u^i := (u_0^i, \dots, u_{T-1}^i) \in \mathbb{R}^T$ . Similarly, for any  $t \in$   
 513  $\{0, \dots, T\}$   $X_t^u := (x_t^1, \dots, x_t^n) \in \mathbb{R}^n$  is the state vector of all the agents  
 514 controlled by  $u_j := (u_t^1, \dots, u_t^n) \in \mathbb{R}^n$ . We have the following dynamics:

$$515 \quad (5.37) \quad \begin{cases} X_{t+1}^u &= AX_t^u + Bu_t + CW_{t+1}, \quad \text{for } t \in \{0, \dots, T-1\}, \\ X_0^u &= x_0 \in \mathbb{R}^n, \end{cases}$$

516 where  $A$  and  $B$  are diagonal matrices,  $C$  is a positive diagonal matrix of size  
 517  $n$ . The global noise process is a sequence of independent random variables  
 518  $(W_1, \dots, W_T)$ , where for any  $t \in \{1, \dots, T\}$ ,  $W_t$  is a vector of centered, re-  
 519 duced and independent Gaussian variables, defined on the probability space  
 520  $(\Omega, \mathcal{F}, \mathbb{P})$ :  $W_t := (W_t^1, \dots, W_t^n)$ . For any  $i \in \{1, \dots, n\}$  and  $t \in \{1, \dots, T\}$  we  
 521 define  $\mathcal{F}_t^i := \sigma(W_1^i, \dots, W_t^i)$ .

- 522 • For any  $i \in \{1, \dots, n\}$ , we define  $\mathcal{U}^i := \prod_{t=0}^{T-1} U_t^i$  the control space of agent  $i$   
 523 with:  $U_t^i := \{\alpha : \Omega \mapsto \mathbb{R} \mid \alpha \text{ is } \mathcal{F}_t^i\text{-measurable and } \alpha(\omega) \in [-M, M] \text{ } \mathbb{P}\text{-a.s.}\}$ ,  
 524 where  $M > 0$ . We finally set  $\mathcal{U} := \prod_{i=1}^n \mathcal{U}^i$ .

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

$$527 \quad (5.38) \quad (P_1^d) \quad \begin{cases} \inf_{u \in \mathcal{U}} J^d(u) \\ J^d(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i \right) + \frac{1}{n} \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right), \end{cases}$$

528 and

$$529 \quad (5.39) \quad (P_2^d) \quad \begin{cases} \inf_{u \in \mathcal{U}} \tilde{J}^d(u) \\ \tilde{J}^d(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) + \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n F_i(u^i, X^{i,u^i}) \right), \end{cases}$$

530 where  $F_0 : \mathbb{R}^T \rightarrow \bar{\mathbb{R}}$  and  $F_i : \mathbb{R}^T \times \mathbb{R}^T \rightarrow \mathbb{R}$  are proper and lower semi continuous, and  
 531 there exists  $\hat{u} \in \mathcal{U}$  such that:  $\mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n \hat{u}^i \right) \right) < \infty$ . In addition we suppose that  
 532  $F_0$  is convex and differentiable with  $c$ -Lipschitz derivative and for any  $i \in \{1, \dots, n\}$ ,  
 533  $u^i \mapsto \mathbb{E}(F_i(u^i, X^{i,u^i}))$  is strictly convex.

534 **COROLLARY 5.15.** (i) Problems  $(P_1^d)$  and  $(P_2^d)$  admit both a unique solution.  
 535 (ii) Any optimal solution of problem  $(P_2^d)$  is an  $\varepsilon$ -optimal solution, where  $\varepsilon =$   
 536  $2cNM^2/n$ , of problem  $(P_1^d)$ .

537 *Proof.* The proof of point (i) is the same as for the Lemma 2.6. Similarly, point  
 538 (ii) is obtained by using the same proof of Theorem 3.6.(ii).  $\square$

539 By adapting the *Stochastic Uzawa* (Algo 5.1) and the *Sampled Stochastic Uzawa*  
 540 (Algo 5.2) to this discrete time setting, one can obtain similar results to Theorems  
 541 5.5, 5.8 and 5.12.

542 **6. A numerical example: the LQG (Linear Quadratic Gaussian) prob-**  
 543 **lem.** This sections aims at illustrating numerically the convergence of the *Stochastic*  
 544 *Uzawa* (Algo 5.1) on a simple example. The algorithm speed of convergence is stud-  
 545 ied, depending on the number of dual iterations and of agents. A linear quadratic  
 546 formulation is considered, with  $n$  agents in a discrete setting problem ( $P_2^{LQG}$ ). We  
 547 use the notations of Section 5.2.

548 This framework constitutes a simple test case, since the (deterministic) Uzawa's  
 549 algorithm can be performed, and one can compare the resulting multiplier estimate  
 550 with the one provided by the *Stochastic Uzawa* algorithm. Besides all the assump-  
 551 tions required for the convergence of the *Stochastic Uzawa* (Algo 5.1) are satisfied  
 552 for problem ( $P_2^{LQG}$ ). In addition the local problems (line 5 of this algorithm) can be  
 553 resolved analytically.

554 Problem ( $P_2^{LQG}$ ) is similar to ( $P_2^d$ ) defined in (5.39), but in this specific case, the  
 555 function  $F_0$  is a quadratic function of the aggregate strategies of the agents

$$556 \quad (6.1) \quad F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i) \right) := \frac{\nu}{2} \sum_{t=0}^T \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u_t^i) - r_t \right)^2,$$

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

$$559 \quad (6.2) \quad F_i(u^i, X^{i,u^i}) := \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2,$$

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

$$562 \quad (6.3) \quad \sum_{i=1}^n F_i(u^i, X^{i,u^i}) = \frac{1}{2} \left( \sum_{t=0}^T X_t^{u^\top} D X_t^u + u_t^\top Q u_t \right) + \frac{1}{2} X_T^{u^\top} D^f X_T^u.$$

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

564 To find the optimal multiplier and control of ( $P_2^{LQG}$ ), the *Stochastic Uzawa* Al-  
 565 gorithm 5.1 is applied where in this specific case the lines 4 and 6 take respectively  
 566 the following form at any dual iteration  $k$ :

$$567 \quad (6.4) \quad u^i(\lambda^k) := \arg \min_{u^i \in \tilde{U}^i} \left\{ \mathbb{E} \left( \frac{1}{2} \left( \sum_{t=0}^T d_i (X_t^{i,u^i})^2 + q_i (u_t^i)^2 + \lambda_t^k u_t^i \right) + \frac{d_i^f}{2} (X_T^{i,u^i})^2 \right) \right\},$$

568

$$569 \quad (6.5) \quad v(\lambda^k) := \arg \min_{v \in \mathbb{R}^T} \left\{ \left( \sum_{t=0}^T \nu (v_t - r_t)^2 - \lambda_t^k v_t \right) \right\}.$$

570 The optimization problem (6.4) solved by each local agent is also in the LQG frame-  
 571 work. One can solve these problems using the results of [23]. The resolution via

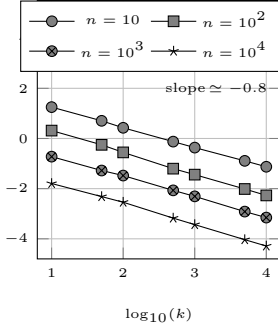


Figure 6.0.1:  $\log_{10}(v_{k,n})$  function of  $k$ , given the number of agents  $n$

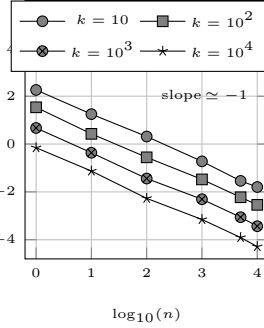


Figure 6.0.2:  $\log_{10}(v_{k,n})$  function of  $n$ , given the number of agents  $k$

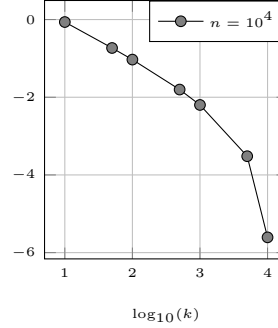


Figure 6.0.3:  $\log_{10}(\|b_{k,n}\|_2^2)$  function of  $k$ , given the number of agents  $n = 10^4$

572 Riccati equations of (6.4) shows that  $u^i(\lambda^k)$  is a linear function of the state  $X^{i,u^i}$   
 573 and of the price  $\lambda^k$ . Therefore, in this specific example, for any  $t$  one can explicitly  
 574 compute  $\mathbb{E}(u_t^i(\lambda^k)|\mathcal{G}_k)$ , where  $\mathcal{G}_k$  is defined in (5.9). It allows us to implement the  
 575 (deterministic) Uzawa's algorithm as a reference to evaluate the performances of the  
 576 *Stochastic Uzawa* algorithm.

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

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

585 For any  $n$ , the deterministic multiplier  $\bar{\lambda}^n$  is obtained by applying Uzawa's  
 586 algorithm, after  $10^4$  dual iterations. To this end, we applied the *Stochastic Uzawa*  
 587 Algorithm 5.1 where we ignored the line 8 and we replaced the update of  $\lambda^k$  line 9

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

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

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

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

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

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

597 On Figure 6.0.1, we observe a behavior in  $1/k^\alpha$  (with  $\alpha \simeq 0.8$ ) of the variance  
 598  $v_{k,n}$  w.r.t. the number of iterations  $k$ . This rate of convergence is consistent with [5,  
 599 Theorem 2.2.12, Chapter 2] for Robbins Monro algorithm where the convergence is  
 600 proved to be of order at most in  $1/k$ .

601 On Figure 6.0.2 we observe a behavior in  $1/n^\beta$  (with  $\beta \simeq 1$ ) of the variance  $v_{k,n}$

602 w.r.t. the number of agents  $n$ . This is expected, see [5, Theorem 2.2.12, Chapter 2]  
 603 and observing that the variance of  $Y^{k+1}$  is of order  $1/n$  for any iteration  $k$ .

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

607 **7. Price-based coordination of a large population of thermostatically**  
 608 **controlled loads.** The goal of this section is to demonstrate the applicability of the  
 609 presented approach for the coordination of thermostatic loads in a smart grid context.  
 610 The problem analyses the daily operation of a power system with a large penetration  
 611 of price-responsive demand, adopting a modelling framework similar to [4]. Two dis-  
 612 tinct elements are considered: i) a system operator, that must schedule a portfolio  
 613 of generation assets in order to satisfy the energy demand at a minimum cost, and  
 614 ii) a population of price-responsive loads (TCLs) that individually determine their  
 615 ON/OFF power consumption profile in response to energy prices with the objective  
 616 of minimizing their operating cost while fulfilling users' requirements. Note that the  
 617 operations of the two elements are interconnected, since the aggregate power consump-  
 618 tion of the TCLs will modify the demand profile that needs to be accommodated by  
 619 the system operator.

620 **7.1. Formulation of the problem.** In the considered problem, the function  $F_0$   
 621 represents the minimized power production cost and corresponds to the resolution of  
 622 an Unit Commitment (UC) problem. The UC determines generation scheduling deci-  
 623 sions (in terms of energy production and frequency response (FR) provision) in order  
 624 to minimize the short term operating cost of the system while matching generation  
 625 and demand. The latter is the sum of an inflexible deterministic component (denoted  
 626 for any instant  $t \in [0, T]$  by  $\bar{D}(t)$ ) and of a stochastic part, which corresponds to the  
 627 total TCL demand profile  $nU_{TCL}(t)$ .

628 For simplicity, a Quadratic Programming (QP) formulation in a discrete time  
 629 setting is adopted for the UC problem. The central planner disposes of  $Z$  genera-  
 630 tion technologies (gas, nuclear, wind) and schedules their production and allocated  
 631 response by slot of 30 min every day. For any  $j \in \{1, \dots, Z\}$  and  $\ell \in \{1, \dots, 48\}$ ,  
 632  $H_j(t_\ell)$ ,  $G_j(t_\ell)$  and  $R_j(t_\ell)$  are respectively the commitment, the power production  
 633 and response [MWh] from unit  $j$  during the time interval  $[t_\ell, t_{\ell+1}]$ . The associated  
 634 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  
 635  $R(t_\ell) = [R_1(t_\ell), \dots, R_Z(t_\ell)]$ .

636 The cost sustained at time  $t_\ell$  by unit  $j$  is linear with respect to the commit-  
 637 ment  $H_j(t_\ell)$  and quadratic with respect to generation  $G_j(t_\ell)$  and can be expressed  
 638 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-  
 639 tion allocated by each generation technology,  $c_{1,j}$  [€/MWh] as no-load cost and  $c_{2,j}$   
 640 [€/MWh] and  $c_{3,j}$  [€/MW<sup>2</sup>h] as production cost of the generation technology  $j$ . The  
 641 optimization of  $F_0$  must satisfy the following constraints for all  $\ell \in \{1, \dots, 48\}$  and  
 642  $\ell \in \{1, \dots, 48\}$ :

$$643 \quad (7.1) \quad \sum_{j=1}^Z G_j(t_\ell) - \int_{t_\ell}^{t_{\ell+1}} (\bar{D}(t) + nU_{TCL}(t))dt = 0,$$

644

$$645 \quad (7.2) \quad 0 \leq H_j(t_\ell) \leq 1,$$

646

$$647 \quad (7.3) \quad R_j(t_\ell) - r_j H_j(t_\ell) G_j^{max}(t_\ell) \leq 0,$$

648

$$649 \quad (7.4) \quad R_j(t_\ell) - s_j(H_j(t_\ell)G_j^{max}(t_\ell) - G_j(t_\ell)) \leq 0,$$

650

$$651 \quad (7.5) \quad \Delta G_L - \Lambda (\bar{D}(t_\ell) + n(\bar{U}_{TCL}(t_\ell) - \bar{R}_{TCL}(t_\ell))) \Delta f_{qss}^{max} - \hat{R}(t_\ell) \leq 0,$$

652

$$653 \quad (7.6) \quad 2\Delta G_L t_{ref} t_d - t_{ref}^2 \hat{R}(t_\ell) - 4\Delta f_{ref} t_d \hat{H}(t_\ell) \leq 0,$$

654

$$655 \quad (7.7) \quad \bar{q}(t) - \hat{H}(t) \hat{R}(t) \leq 0$$

656

$$657 \quad (7.8) \quad \mu r_j H_j(t_\ell) G_j^{max}(t_\ell) - G_j(t_\ell) \leq 0,$$

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

$$662 \quad \hat{R}(t_\ell) = \sum_{j=1}^Z R_j(t_\ell) + nR_{TCL}(t_\ell) \quad \text{and} \quad \hat{H}(t_\ell) = \sum_{j=1}^Z \frac{h_j H_j(t_\ell) G_j^{max} - h_L \Delta G_L}{f_0}.$$

663 Assuming that for any generic generation technology  $j$ , the size of single plants  
 664 included in  $j$  is quite smaller than the aggregate installed capacity of  $j$ , inequality  
 665 (7.2) sets that commitment decisions can be extended to the fleet and expressed by  
 666 continuous variables  $H_j(t_\ell) \in [0, 1]$ .

667 The amount of response allocated by each generation technology is limited by  
 668 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  
 669 dispatch level (7.4). Constraints (7.5) to (7.8) deal with frequency response provision  
 670 and  $R_{TCL}$  (the mean of FR allocated by TCLs). They guaranty secure frequency  
 671 deviations following sudden generation loss  $\Delta G_L$ . Inequality (7.5) allocates enough  
 672 FR (with delivery time  $t_d$ ) such that the quasi-steady-state frequency remains above  
 673  $\Delta f_{qss}^{max}$ , with  $\Lambda$  accounting for the damping effect introduced by the loads [11]. Fi-  
 674 nally (7.7) constraints the maximum tolerable frequency deviation  $\Delta f_{nad}$ , following  
 675 the formulation and methodology presented in [22] and [24]. The rate of change of  
 676 frequency is taken into account in (7.6) where at  $t_{ref}$  the frequency deviation remains  
 677 above  $\Delta f_{ref}$ . Constraint (7.8) prevents trivial unrealistic solutions that may arise  
 678 in the proposed formulation, such as high values of committed generation  $H_j(t_\ell)$  in  
 679 correspondence with low (even zero) generation dispatch  $G_j(t_\ell)$ . The reader can refer  
 680 to [4] for more details on the UC problem.

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

$$683 \quad (7.9) \quad F_0(U_{TCL}, R_{TCL}) := \min_{H, G, R} \sum_{\ell=1}^{48} \sum_{j=1}^Z c_{1,j} H_j(t_\ell) G_j^{max}(t_\ell) + c_{2,j} G_j(t_\ell) + c_{3,j} G_j(t_\ell)^2,$$

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

685 Note that the formulation of the present problem does not fulfill all the assumption  
 686 presented in Sections 2 and 5. In particular, the function  $F_0$  is not strictly convex, as  
 687 instead supposed in Theorem 5.8.(ii).(iii). Nevertheless, the numerical simulations of  
 688 Section 7.2 shows that the proposed approach is still able to achieve convergence.

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

$$693 \quad (7.10) \quad \begin{cases} dX_t^{i,u^i} &= -\frac{1}{\gamma_i}(X_t^{i,u^i} - X_{OFF}^i + \zeta_i u_t^i)dt + \sigma_i dW_t^i, \quad \text{for } t \in [0, T], \\ X_{0,u^i}^i &= x_0^i \in \mathbb{R}, \end{cases}$$

694 where:

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

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

$$701 \quad (7.11) \quad \mathcal{U}_i := \{\nu \in H_i \text{ and } \nu_t(\omega) \in \{0, P_{ON,i}\} \text{ for a.a. } (t, \omega) \in [0, T] \times \Omega^i\}.$$

702 The TCLs dynamics in (7.10) have been derived according to [10], with the addition  
 703 of the stochastic term  $\sigma_i dW_t^i$  to account for the influence of the environment (open-  
 704 ing/closing of the fridge, environment temperature etc) on the evolution of the TCL  
 705 temperature.

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

$$708 \quad (P_1^{TCL}) \quad \begin{cases} \inf_{u \in \mathcal{U}} J(u) \\ J(u) := \mathbb{E} \left( F_0 \left( \frac{1}{n} \sum_{i=1}^n u^i, \frac{1}{n} \sum_{i=1}^n r_i(u^i, X^{i,u^i}) \right) \right) \\ \quad + \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i,u^i}) ds + \gamma_i (X_T^{i,u^i} - \bar{X}^i)^2 \right), \end{cases}$$

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

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

$$711 \quad (7.13) \quad r_i(u^i, X^{i,u^i})(s) := u_s^i \frac{X_s^{i,u^i} - X_{min}^i}{X_{max}^i - X_{min}^i}.$$

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

$$713 \quad (7.14) \quad f_i(u_s^i, X_s^{i,u^i}) := \alpha_i (X_s^{i,u^i} - \bar{X}^i)^2 + \beta_i ((X_{min}^i - X_s^{i,u^i})_+^2 + (X_s^{i,u^i} - X_{max}^i)_+^2),$$

714 where:

- 715 -  $\alpha_i (X_s^{i,u^i} - \bar{X}^i)^2$  is a discomfort term penalizing temperature deviation  
 716 from some comfort target  $\bar{X}$  [ $^\circ C$ ], with  $\alpha_i$  a discomfort term parameter  
 717 [ $\mathcal{L}/h(^\circ C)^2$ ].
- 718 -  $\beta_i ((X_s^{i,u^i} - X_{min}^i)_+^2 + (X_{max}^i - X_s^{i,u^i})_+^2)$  is a penalization term to keep  
 719 the temperature in the interval  $[X_{min}^i, X_{max}^i]$ , with  $\beta_i$  a target term pa-  
 720 rameter [ $\mathcal{L}/s(^\circ C)^2$ ] and for any  $x \in \mathbb{R}$ ,  $(a)_+ = \max(0, a)$ .



721 •  $\gamma_i(X_T^{i,u^i} - \bar{X}_i)^2$  is a terminal cost imposing periodic constraints, with  $\gamma$  a  
 722 target term parameter [ $\text{£}/\text{s}(\text{°C})^2$ ].

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

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

(7.15)

$$726 \quad (P_2^{TCL}) \quad \left\{ \begin{array}{l} \inf_{u \in \mathcal{U}} \tilde{J}(u) \\ \tilde{J}(u) := F_0 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}(u^i), \frac{1}{n} \sum_{i=1}^n \mathbb{E}(r_i(u^i, X^{i,u^i})) \right) \\ \quad + \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^n \int_0^T f_i(u_s^i, X_s^{i,u^i}) ds + \gamma_i (X_T^{i,u^i} - \bar{X}^i)^2 \right). \end{array} \right.$$

727 **7.2. Decentralized implementation.** The *Sampled Stochastic Uzawa* Algo-  
 728 rithm 5.2 is applied to solve ( $P_2^{TCL}$ ), with  $m = 317$  simulations per iteration. At each  
 729 iteration  $k$ , the lines 4 and 6 correspond respectively to the solution of a deterministic  
 730 UC problem and of an Hamilton Jacobi Bellman (HJB) equation. The time steps  
 731  $\Delta t = 7.6$  s and temperature steps  $\Delta T = 0.15$ °C are chosen for the discretization of  
 732 the HJB equation. Let us note that at line 6, each TCL solves its own local problem  
 733 on the basis of the received price signal  $\lambda^k = (p^k, \rho^k)$ :

$$734 \quad (7.16) \quad \inf_{u^i \in \mathcal{U}_i} \int_0^T f_i(u_s^i, X_s^{i,u^i}) + u_s^i p_s^k - r_i(u^i, X^{i,u^i})(s) \rho_s^k ds,$$

735 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  
 736 as consumption cost and  $r_i(u^i, X^{i,u^i})(s) \rho_s^k$  as fee awarded for FR provision. This  
 737 implementation has a practical sense: each TCL uses local information and a price  
 738 that is communicated to them to schedule its power consumption on the time interval  
 739  $[0, T]$ . It follows that, with the proposed approach, it is possible to optimize the overall  
 740 system costs in ( $P_1^{TCL}$ ) in a distributed manner, with each TCL acting independently  
 741 and pursuing the minimization of its own costs.

742 **7.3. Results.** The generation technologies available in the system are nuclear,  
 743 combined cycle gas turbines (CCGT), open cycle gas turbines (OCGT) and wind.  
 744 The characteristics and parameters of the UC in this simulation are the same as in  
 745 [4].

746 It is assumed that a population of  $n = 2 \times 10^7$  fridges with built-in freeze compart-  
 747 ment operates in the system according to the proposed price-based control scheme.  
 748 For any agent  $i$  we set the consumption parameter  $P_{ON,i} = 180W$ . The values of the  
 749 TCL dynamic parameters  $\gamma_i$  and  $X_{OFF}^i$  of (7.10) are equal to the ones taken in [4].  
 750 Note that it is possible to take a population of heterogeneous TCLs with different  
 751 parameter values. The initial temperature are picked randomly uniformly between  
 752  $-21$ °C and  $-14$ °C. For any agent  $i$ , the parameters of the individual cost function  $f_i$ ,  
 753 defined in (7.14), are:  $\alpha_i = 0.2 \times 10^{-4}$   $\text{£}/\text{s}(\text{°C})^2$ ,  $\beta_i = 50 \text{£}/\text{s}(\text{°C})^2$ ,  $\bar{X}^i = -17.5$ °C and  
 754  $X_{max} = -14$ °C,  $X_{min} = -21$ °C. The parameter  $\beta_i$  is taken intentionally very large  
 755 to make the temperature stay in the interval  $[X_{max}^i, X_{min}^i]$ . Note that the individual  
 756 problems solved by the TCLs are distinct than the ones in [4] (different terms and  
 757 parameters).

758 Simulations are performed for different values of volatility  $\sigma_i := 0, 1, 2$  (all the  
 759 TCLs have the same volatility in the simulations), where  $\sigma_i$  is defined in (7.10).



760 The *Sampled Stochastic Uzawa* Algorithm is stopped after 75 iterations or when the  
 761 relative variation  $2\|\lambda^{k+1} - \lambda^k\|_2^2 / \|\lambda^{k+1} + \lambda^k\|_2^2$  between two successive prices  $\lambda^k$  and  
 762  $\lambda^{k+1}$  is less than  $10^{-4}$ .

763 The resulting profile of total power consumption  $nU_{TCL}$  and total allocated re-  
 764 sponse  $nR_{TCL}$  by the TCLs population are reported on figure 7.3.1. in three "flexibil-  
 765 ity scenario" each corresponding to a case where TCL flexibility is enabled with three  
 766 different volatilities  $\sigma = 0$ ;  $\sigma = 1$  and  $\sigma = 2$ . The electricity prices  $p$  and response  
 767 availability prices  $\rho$  are shown in Figure 7.3.2. As observed in [4], the total con-  
 768 sumption  $nU_{TCL}$  is higher when the price  $p$  is lower and inversely the total allocated  
 769 response  $nR_{TCL}$  is higher when the price signal  $\rho$  is also higher. This can be observed  
 770 during the first hours of the day, between 0 and 6h. The power  $U_{TCL}$  then oscillates  
 771 during the day in order to maintain feasible levels of the internal temperature of the  
 772 TCLs. Though the prices seem not to be sensitive to the values taken by  $\sigma$ , the  
 773 average consumption  $U_{TCL}$  and response  $R_{TCL}$  are highly correlated to the volatility  
 774 of the temperature of the TCLs. The less noisy their temperature are, the more price  
 775 sensitive and flexible their consumption profiles are. The TCLs impact on system  
 776 commitment decisions and consequent energy/FR dispatch levels is also analyzed and  
 777 displayed in Figure 7.3.3 and 7.3.4. The production and reserve in the "flexibility sce-  
 778 nario" minus the production and reserve in the "no-flexibility scenario" are plotted,  
 779 for different volatilities  $\sigma$ . In the no-flexibility scenario we impose  $R_{TCL}(t) = 0$  and  
 780 we consider that the TCLs operate exclusively according to their internal tempera-  
 781 ture  $X^{i,u^i}$ . They switch ON ( $u^i(t) = P_{ON,i}$ ) when they reach their maximum feasible  
 782 temperature  $X_{max}^i$  and they switch back OFF again ( $u^i(t) = 0$ ) when they reach the  
 783 minimum temperature  $X_{min}^i$ . In figure 7.3.3, we can clearly observe that TCL's flexi-  
 784 bility allows to increase the contribution of wind generation (reducing curtailment) to  
 785 the energy balance of the system while decreasing the contribution of CCGT both in  
 786 energy and frequency response. Without TCL support, the optimal solution envisages  
 787 a further curtailment of wind output in favor of an increase in CCGT generation, as  
 788 wind does not provide FR. As expected, the influence of the TCL on the system is  
 789 larger when the temperature volatility is lower.

790  
 791

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

802

#### REFERENCES

- 803 [1] Dimitri P Bertsekas and Steven Shreve. *Stochastic optimal control: the discrete-time case*.  
 804 2004.  
 805 [2] Pierre Carpentier, J-Ph Chancelier, Vincent Leclère, and François Pacaud. Stochastic decom-  
 806 position applied to large-scale hydro valleys management. *European Journal of Operational*  
 807 *Research*, 270(3):1086–1098, 2018.

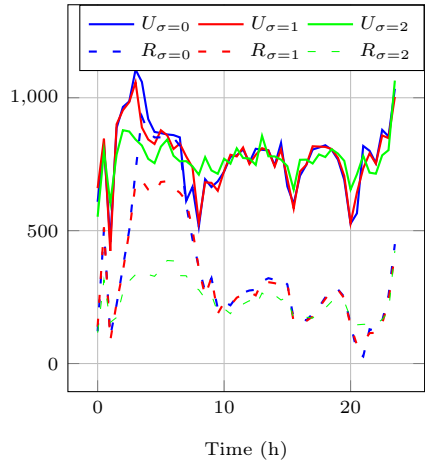


Figure 7.3.1: Total power consumption  $U$  and allocated response  $R$  (MW) of TCLs after 75 iterations of the algorithm.

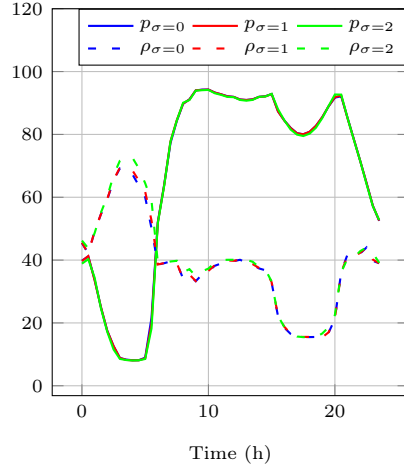


Figure 7.3.2: Electricity price  $p$  and response availability price  $\rho$  (£/MWh) after 75 iterations of the algorithm.

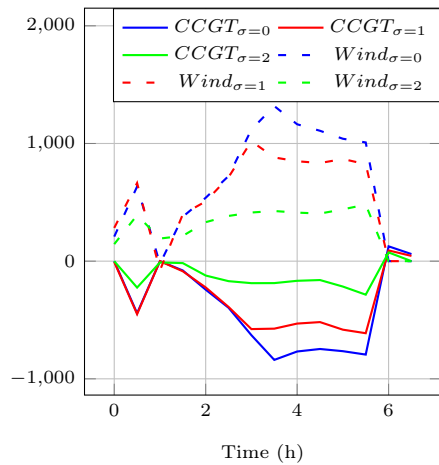


Figure 7.3.3: Deviation of generation profiles (MW) from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

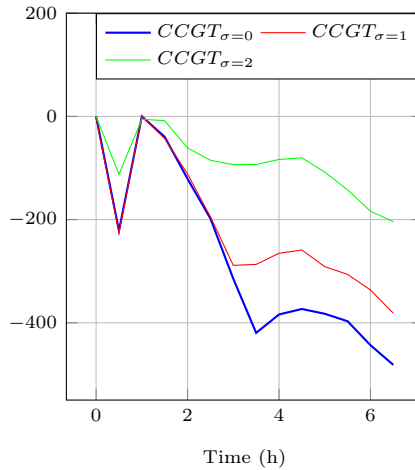


Figure 7.3.4: Deviation of Frequency Response (MW) allocated by CCGT from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

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

- 808 [3] Michael Chertkov and Vladimir Chernyak. Ensemble of thermostatically controlled loads: Statistical physics approach. *Scientific reports*, 7(1):8673, 2017.
- 809 [4] Antonio De Paola, Vincenzo Trovato, David Angeli, and G. Strbac. A mean field game approach for distributed control of thermostatic loads acting in simultaneous energy-frequency response markets. *IEEE Transactions on Smart Grid*, PP:1–1, 01 2019.
- 810 [5] Marie Duflo. *Random iterative models*, volume 34. Springer Science & Business Media, 2013.
- 811 [6] Caroline Geiersbach and Georg Ch Pflug. Projected stochastic gradients for convex constrained problems in Hilbert spaces. *SIAM Journal on Optimization*, 29(3):2079–2099, 2019.
- 812 [7] Pierre Girardeau. *Solving large-scale dynamic stochastic optimization problems*. PhD thesis, Université Paris-Est, December 2010.
- 813 [8] He Hao, Borhan M Sanandaji, Kameshwar Poolla, and Tyrone L Vincent. Aggregate flexibility of thermostatically controlled loads. *IEEE Transactions on Power Systems*, 30(1):189–198, 2014.
- 814 [9] Julia L Higle and Suvrajeet Sen. *Stochastic decomposition: a statistical method for large scale stochastic linear programming*, volume 8. Springer Science & Business Media, 2013.
- 815 [10] Arman C Kizilkale, Rabih Salhab, and Roland P Malhamé. An integral control formulation of mean field game based large scale coordination of loads in smart grids. *Automatica*, 100:312–322, 2019.
- 816 [11] Prabha Kundur, Neal J Balu, and Mark G Lauby. *Power system stability and control*, volume 7. McGraw-hill New York, 1994.
- 817 [12] Vincent Leclere. *Contributions to decomposition methods in stochastic optimization*. PhD thesis, Université Paris-Est, June 2014.
- 818 [13] Xuerong Mao. *Stochastic differential equations and applications*. Elsevier, 2007.
- 819 [14] François Pacaud. *Decentralized optimization for energy efficiency under stochasticity*. PhD thesis, Université Paris-Est, October 2018.
- 820 [15] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1-3):359–375, 1991.
- 821 [16] Pedro Pérez-Aros and Emilio Vilches. An enhanced Baillon-Haddad theorem for convex functions on convex sets. *arXiv preprint arXiv:1904.04885*, 2019.
- 822 [17] Andrew B Philpott and Ziming Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.
- 823 [18] R Tyrrell Rockafellar and Roger J-B Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of operations research*, 16(1):119–147, 1991.
- 824 [19] Andrzej Ruszczyński and Alexander Shapiro. Stochastic programming models. *Handbooks in operations research and management science*, 10:1–64, 2003.
- 825 [20] David H Salinger. *A splitting algorithm for multistage stochastic programming with application to hydropower scheduling*. PhD thesis, 1997.
- 826 [21] Joe A Short, David G Infield, and Leon L Freris. Stabilization of grid frequency through dynamic demand control. *IEEE Transactions on power systems*, 22(3):1284–1293, 2007.
- 827 [22] Fei Teng, Vincenzo Trovato, and Goran Strbac. Stochastic scheduling with inertia-dependent fast frequency response requirements. *IEEE Transactions on Power Systems*, 31(2):1557–1566, 2015.
- 828 [23] Emanuel Todorov. Optimal control theory. *Bayesian brain: probabilistic approaches to neural coding*, pages 269–298, 2006.
- 829 [24] Vincenzo Trovato, Agnès Bialecki, and Anes Dallagi. Unit commitment with inertia-dependent and multispeed allocation of frequency response services. *IEEE Transactions on Power Systems*, 34(2):1537–1548, 2018.
- 830 [25] Vincenzo Trovato, Simon H Tindemans, and Goran Strbac. Leaky storage model for optimal multi-service allocation of thermostatic loads. *IET Generation, Transmission & Distribution*, 10(3):585–593, 2016.