DECOMPOSITION OF HIGH DIMENSIONAL AGGREGATIVE STOCHASTIC CONTROL PROBLEMS *

ADRIEN SEGURET†, CLÉMENCE ALASSEUR‡, J. FRÉDÉRIC BONNANS§, ANTONIO DE PAOLA¶, NADIA OUDJANE‖, AND VINCENZO TROVATO#

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

Key words. Stochastic optimization, Lagrangian decomposition, Uzawa's algorithm, stochastic gradient.

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

1. Introduction.

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

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

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

1.1. Motivations.

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

---

* Submitted to the editors September 23, 2020

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

†OSIRIS department, EDF Lab, Paris-Saclay, France(adrien.seguret@edf.fr).

‡OSIRIS department, EDF Lab, Paris-Saclay, France(clemence.alasseur@edf.fr).

§CMAP, Ecole polytechnique, 91128 Palaiseau, France, and Inria-Saclay(frederic.bonnans@inria.fr).

¶Center for Sustainable Power Distribution University of Bath, Bath, UK(adp50@bath.ac.uk).

‖OSIRIS department, EDF Lab, Paris-Saclay, France(nadia.oudjane@edf.fr).

#Department of Electrical and Electronic Engineering, Imperial College London, London, UK(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 algorithm. To this end in [12], the Uzawa's algorithm, formulated in a Hilbert setting, is extended to a Banach space. DADP has been applied in different cases, such as storage management problem for electrical production in [7, chapter 4] and hydro valley management [2]. In the proposed paper, in the same vein as DADP we propose a price decomposition approach restricted to deterministic prices. This new approach takes advantage of the large population number in order to introduce an auxiliary problem where the coupling term is purely deterministic.

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

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

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

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

The following notations are used:

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

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

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

and we set $M := \max_{i\in\{1,\dots,n\}} M_i$, where $M_i > 0$. The set of admissible controls is $\mathcal{U} := \mathcal{U}_1 \times \ldots \times \mathcal{U}_n$, whose elements are denoted by $u := (u^1,\dots,u^n)$.

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

$\mathbb{R}$ and follows the dynamics

(2.2)
$$\begin{cases} dX_t^{i,u^i} & = & \mu_i(t,u_t^i,X_t^{i,u^i})dt + \sigma_i(t,X_t^{i,u^i})dW_t^i, \text{ for } t\in[0,T], \text{ for } i\in\{1,\dots,n\} \\ X_{0,u^i}^i & = & x_0^i \in \mathbb{R}. \end{cases}$$

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

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

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

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

(2.3)
$$\begin{aligned} &|\mu_i(t,\nu,x)-\mu_i(t,\nu,y)| + |\sigma_i(t,\nu,x)-\sigma_i(t,\nu,y)| \le K_i\,|x-y|, \\ &|\mu_i(t,\nu,x)| + |\sigma_i(t,\nu,x)| \le K_i\,(1+|x|), \end{aligned} \qquad \text{for any } x,y\in\mathbb{R}.$$

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

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

Proof. The proof for the existence and uniqueness of a solution of (2.2) relies on [13, Theorem 3.6, Chapter 2]. The inequality is a result of [13, Theorem 4.4, Chapter 2]. ∎

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

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

Define $G_i : L^2(0,T)\times\mathcal{C}([0,T])\to\bar{\mathbb{R}}$ by $G_i(z,\omega) = F_i(z,X^{i,z}(\omega))$. Additional assumptions are formulated below.

Assumption 2.3. For any $i\in\{1,\dots,n\}$:
(i) G_i is strictly convex w.r.t. the first variable.
(ii) there exists a positive integer p such that F_i has p-polynomial growth, i.e there exists $K>0$ such that for any $x^i\in\mathcal{C}([0,T],\mathbb{R})$ and $u^i\in L^2(0,T)$: $|F_i(u^i,x^i)|\le K(1+\sup_{0\le t\le T}|x_t^i|^p)$.

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

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

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

Then Assumption 2.3.(i) is satisfied if:

(i) g_i is strictly convex and h_i convex.

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

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) 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:

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

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

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

Remark 2.5. If for any $i \in \{1,\dots,n\}$, Assumption 2.3.(i) holds, then G_i is w.l.s.c. w.r.t. the first variable. Indeed, G_i as a function of the first variable being convex, finite valued and bounded on bounded subsets of $L^2(0,T)$ (from the polynomial growth of F_i and the inequality (2.4)),thus G_i is continuous w.r.t. the first variable.

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

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

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

Proof. Clearly the control $\hat{u} \in \mathcal{U}$ defined in (2.5) is feasible. The existence and uniqueness of a minimum is proved by considering a minimizing sequence $\{u_k\}$ of J over $\mathcal{U}$. The set $\mathcal{U}$ being bounded and weakly closed, there exists a sub-sequence $\{u_{k_\ell}\}$ which weakly converges to a certain $u^* \in \mathcal{U}$. Using Assumptions 2.3.(i)(ii) and convexity of F_0, it follows that $\liminf J(u_{k_\ell}) \geq J(u^*)$ and thus u^* is a solution of (P_1). The uniqueness is due to the strict convexity of G_i w.r.t. the first variable. ☐

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

3. Approximating the optimization problem. In this section, the link between problems (P_1) and (P_2) is analyzed.

Assumption 3.1. Problem (P_2) admits a unique solution.

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

We have the following key result.

THEOREM 3.2. *Under Assumption 3.1, $\tilde{J}$ reaches its minimum over $\mathcal{U}$ at a unique 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 $\tilde{u}^j$ are mutually independent.*

Proof. Fix $i \in \{1,\dots,n\}$, since G_i is proper, convex and l.s.c. w.r.t. the first variable, using Jensen's inequality we get:

$$\mathbb{E}\big(G_i(u^i, W^i)|W^i\big) \geq G_i\big(\mathbb{E}(u^i\,|W^i), W^i\big). \tag{3.1}$$

On the other hand $(u^1,\dots,u^n)\mapsto F_0(\frac{1}{n}\sum_{i=1}^n \mathbb{E}(u^i))$ is invariant when taking the conditional expectation, thus $F_0\left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}(u^i)\right) = F_0\left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}\big(\mathbb{E}(u^i|W^i)\big)\right)$. From Assumption 3.1, we know that there exists a solution u^* to (P_2).

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, $\tilde u^i := \mathbb{E}(u^{*i}|W^i)$ is $\mathcal{F}_i$-adapted. Using the definition of u^* and (3.1), one can derive that $\inf_{u\in\mathcal{U}} \tilde J(u) = \tilde J(u^*) \geq \tilde J(\tilde u)$. ◻

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

$$\hat{\mathcal{U}} := \{u\in\mathcal{U}\,|\,u^i \text{ is } \mathcal{F}^i-\text{adapted for all } i\in\{1,\dots,n\}\} \tag{3.2}$$

From Theorem 3.2, if Assumption 3.1 holds, then:

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

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

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

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

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

Using the Jensen and Hölder inequalities, $(\mathbb{E}|Y|)\leq(\mathbb{E}|Y|^2)^{\frac12}$, the fact that for any $j\neq i$, u_i and u_j are mutually independent, and that u^i is bounded by M, we have $\forall u\in\hat{\mathcal{U}}$:

$$\begin{aligned}
|\mathbb{E}\left(F_0(\frac{1}{n}\sum_{i=1}^n u^i)) - F_0(\frac{1}{n}\sum_{i=1}^n \mathbb{E}(u^i)\right)| &\leq \mathbb{E}\left(|F_0(\frac{1}{n}\sum_{i=1}^n u^i)) - F_0(\frac{1}{n}\sum_{i=1}^n \mathbb{E}(u^i))|\right)\\
&\leq \frac{\gamma}{n}\mathbb{E}(\|\sum_{i=1}^n \hat u^i\|_{L^2(0,T)})\\
&\leq \frac{\gamma}{n}\mathbb{E}(\|\sum_{i=1}^n \hat u^i\|^2_{L^2(0,T)})^{\frac12}\\
&= \frac{\gamma}{n}(\int_0^T \mathrm{Var}(\sum_{i=1}^n u^i_t)dt)^{\frac12}\\
&\leq \frac{\gamma}{n^{\frac12}}M\sqrt{T}.
\end{aligned} \tag{3.5}$$

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:

$$J(\tilde u^*) \leq \tilde J(\tilde u^*) + \frac{\gamma}{n^{\frac12}}M\sqrt{T} \leq \tilde J(u) + \frac{\gamma}{n^{\frac12}}M\sqrt{T} \leq J(u) + \frac{2\gamma}{n^{\frac12}}M\sqrt{T}. \tag{3.6}$$ ◻

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

THEOREM 3.6. *Suppose F_0 is convex, then the following ε-optimality results hold:*
(i) *For any $u \in \mathcal{U}$, $\tilde{J}(u) \leq J(u)$.*
(ii) *Suppose Assumption 3.5 holds, then any optimal solution of problem (P_2) is an ε-optimal solution (where $\varepsilon = 2cTM^2/n$) of problem (P_1).*

Proof. Proof of point (i).
By Jensen's inequality, we have that:

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

which gives the result.

Proof of point (ii).
Since F_0 is convex, differentiable, with a c-Lipschitz differential, one can derive for any $u \in \hat{\mathcal{U}}$ and a.s.:

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

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

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

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

$$\frac{c}{n^2}\,\mathbb{E}\big(\|\sum_{i=1}^{n}\hat{u}^i\,\|^2_{L^2(0,T)}\big) = \frac{c}{n^2}\int_0^T\sum_{i=1}^{n}\mathrm{Var}(u_t^i)dt \leq \frac{c}{n}TM^2. \tag{3.9}$$

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

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

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

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

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

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

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

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

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

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

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

The optimization problem (P_2) is equivalent to:

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

where $g(u,v) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(u^i) - v$ and $\mathcal{V} := \{\nu \in L^2(0,T)|; |v(t)| \leq 2M \, \forall t \in [0,T]\}$.
The *Lagrangian* function associated with the constrained optimization problem (P_3) is: $L : \mathcal{U} \times L^2(0,T) \times L^2(0,T) \to \bar{\mathbb{R}}$ defined by:

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

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

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

For any $\lambda \in L^2(0,T)$, it holds:

$$\mathcal{W}(\lambda) = -F_0^*(\lambda) + \frac{1}{n}\sum_{i=1}^{n} \inf_{u^i \in \mathcal{U}_i} \mathbb{E}(G_i(u^i, W^i)) + \langle \lambda, \mathbb{E}(u^i)\rangle_{L^2(0,T)}, \tag{4.4}$$

where $F_0^*(\lambda) := \sup_{v \in \mathcal{V}} \langle \lambda, v\rangle_{L^2(0,T)} - F_0(v)$.

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

$$\text{There exists } \varepsilon > 0 \text{ such that } \mathcal{B}_{L^2(0,T)}(0,\varepsilon) \subset g(\mathcal{U},\mathcal{V}), \tag{4.5}$$

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

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

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

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

Since the set of admissible controls $\mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_n$ is a Cartesian product and by strict convexity of G_i w.r.t. the first variable, each component u^{*i} can be uniquely determined by solving the following sub problem:

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

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

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

$$\begin{aligned} &\min_{u^i \in \hat{\mathcal{U}}_i} \left\{ \mathbb{E}\left(F_i(u^i, X^{i,u^i}) + \langle \lambda^*, u^i \rangle_{L^2(0,T)} \right) \right\} \\ &= \min_{u^i \in \mathcal{U}_i} \left\{ \mathbb{E}\left(F_i(u^i, X^{i,u^i}) + \langle \lambda^*, u^i \rangle_{L^2(0,T)} \right) \right\}. \end{aligned} \tag{4.6}$$

5. ***Stochastic Uzawa*** and ***Sampled Stochastic Uzawa*** algorithms.

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

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

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

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

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

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

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

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

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

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

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

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

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

Let $s, \tilde{s} \in \mathrm{dom}(F_0^*)$ such that there exist $\lambda_s \in \partial F_0^*(s)$ and $\mu_{\tilde{s}} \in \partial F_0^*(\tilde{s})$. From the differentiability, l.s.c. and convexity of F_0, it follows that: $s = \nabla F_0(\lambda_s)$ and $\tilde{s} = \nabla F_0(\mu_{\tilde{s}})$. By Assumption 3.5 and the extended Baillon-Haddad theorem [16, Theorem 3.1], ∇F_0 is cocoervice. In other words:

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

Therefore ∂F_0^* is strongly monotone, which implies the strong convexity of F_0^*.
(ii) Conversely, assume that F_0^* is proper and strongly convex. Then there exist $\alpha, \beta > 0$ such that for any $s \in \mathrm{dom}(F_0^*)$: $F_0^*(s) \geq \alpha \|s\|^2_{L^2(0,T)} - \beta$, and F_0 being convex, l.s.c. and proper, for any $\lambda \in L^2(0,T)$ it holds:

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

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

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

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

meaning that ∇F_0 is cocoercive. Applying the Cauchy–Schwarz inequality to the left hand side of the previous inequality, the Lipschitz property of ∇F_0 follows. ∎

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

Proof. For any $\lambda \in L^2(0,T)$, the expression of $\mathcal{W}(\lambda)$ is given by 4.4, where for any $i \in \{1, \ldots, n\}$, $\lambda \mapsto \inf_{u^i \in \mathcal{U}_i} \mathbb{E}(G_i(u^i, W^i)) + \langle \lambda, E(u)^i \rangle_{L^2(0,T)}$ is concave and from Lemma 5.1 $-F_0^*$ is strongly concave. Since the sum of a concave function and of a strongly concave function is strongly concave, the result follows. ∎

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

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

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

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

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

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

The iterative algorithm, proposed as an approximation scheme for $\lambda^* \in \arg\max_{\lambda} \mathcal{W}(\lambda)$, 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 $\{W^{i,p} : 1 \le i \le n, p \le k\}$.
7: Compute the associated state realizations $(X^{1,u^1(\lambda^k)}, \dots, X^{n,u^n(\lambda^k)})$.
8: $Y^{k+1} \leftarrow \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} u^i(\lambda^k)(W^{i,k+1}) - v(\lambda^k)$.
9: $\lambda^{k+1} \leftarrow \lambda^k + \rho_k Y^{k+1}$.

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

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

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

Some assumptions on the step size are introduced.

Assumption 5.4. The sequence $(\rho_k)_k$ is such that: $\rho_k > 0$, $\sum_{k=1}^{\infty} \rho_k = \infty$ and $\sum_{k=1}^{\infty} (\rho_k)^2 < \infty$.

Note that a sequence of the form $\rho_k := \dfrac{a}{b+k}$, with $(a,b) \in \mathbb{R}_+^* \times \mathbb{R}_+$, satisfies Assumption 5.4.

Algorithm 5.2 Sampled Stochastic Uzawa

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

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

The following result establishes the convergence of the *Stochastic Uzawa* Algorithm 5.1:

THEOREM 5.5. *Let Assumption 5.4 hold, then:*
(i) $\{\|\lambda^k - \lambda\|^2_{L^2(0,T)}\}$ *converges a.s., for all* $\lambda \in S$.
(ii) $\mathcal{W}(\lambda^k) \xrightarrow[k\to\infty]{} \max_{\lambda \in L^2(0,T)} \mathcal{W}(\lambda)$ *a.s.*
(iii) $\{\lambda^k\}$ *weakly converges to some* $\bar{\lambda} \in S$ *in* $L^2(0,T)$ *a.s.*
(iv) *If Assumption 3.5 holds, then a.s.* $\{\lambda^k\}$ *converges to* $\bar{\lambda}$ *in* $L^2(0,T)$*, with* $S := \{\bar{\lambda}\}$.

Though the proof is similar to [6, Theorem 3.6], the current framework is different from the one of that reference, and for the convenience of the reader we provide the proof.
We first state two lemmas.

LEMMA 5.6 (Robbins-Siegmund). *Let* $\{\mathcal{G}_k\}$ *be an increasing sequence of* σ*-algebra and* d_k*,* a_k*,* b_k *and* c_k *be nonnegative random variables adapted to* $\mathcal{G}_k$*. Assume that:* $\mathbb{E}(d_{k+1}|\mathcal{G}_k) \le d_k(1 + a_k) + b_k - c_k$ *and* $\sum_{k=1}^{\infty} a_k < \infty$ *a.s.,* $\sum_{k=1}^{\infty} b_k < \infty$ *a.s. Then with probability one,* $\{d_k\}$ *is convergent and it holds that* $\sum_{k=1}^{\infty} c_k < \infty$.

Proof. See [5], Theorem 1.3.12. ∎

LEMMA 5.7. *Let* $\{\alpha_k\}$ *be a nonnegative deterministic sequence and* $\{\beta_k\}$ *a nonnegative random sequence adapted to* $\{\mathcal{G}_k\}$*. Assume that* $\sum_{k=1}^{\infty} \alpha_k = \infty$ *a.s. and* $\mathbb{E}(\sum_{k=1}^{\infty} \alpha_k \beta_k) < \infty$ *a.s. Moreover assume that* $\beta_k - \mathbb{E}(\beta_{k+1}|\mathcal{G}_k) \le c\alpha_k$ *a.s. for all*

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

Proof. See [6], Proposition 3.2. ∎

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

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

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

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

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

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

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

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

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

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

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 nonnegative; c_k is nonnegative since $\lambda \in S$. By Lemma 5.6, the sequence $\{\|\lambda^k - \lambda\|^2_{L^2(0,T)}\}$ converges a.s. Now we show point (ii) thanks to Lemma 5.7.

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 both side of (5.12), we get, using the deterministic version of Lemma 5.6 that: $\mathbb{E}\left(\sum_{k=1}^{\infty} \rho_k(\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k))\right) < \infty$. By concavity of $\mathcal{W}$ and the Cauchy-Schwarz inequality, we have:

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

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

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

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

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

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

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 expectation on both sides, noticing that $\mathbb{1}_{\tau_M > k}$ is $\mathcal{G}_k$-measurable, and considering (5.14) and (5.16), we get $\beta_k - \mathbb{E}(\beta_{k+1}|\mathcal{G}_k) \leq \rho_k (M')^2$.
By Lemma 5.7, on the set $B_M := \{\tau_M = \infty\}$, β_k converges to 0 and coincides with $\mathcal{W}(\lambda) - \mathcal{W}(\lambda^k)$. Since $\|\lambda^k - \lambda\|$ converges a.s, $\|\lambda^k\|$ is bounded in probability and therefore the probability of the set B_M ca be made arbitrarily close to 1 by choosing 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 almost surely.

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

$$\mathcal{W}(\bar{\lambda}) \geq \lim_{m\to\infty} \mathcal{W}(\lambda^{k_m}) = \mathcal{W}(\tilde{\lambda}). \tag{5.17}$$

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

$$\|\lambda^{k_m} - \lambda_2\|^2 = \|\lambda^{k_m} - \lambda_1\|^2 + \|\lambda_1 - \lambda_2\|^2 + 2\langle \lambda^{k_m} - \lambda_1, \lambda_1 - \lambda_2\rangle, \tag{5.18}$$

$$\|\lambda^{k_l} - \lambda_1\|^2 = \|\lambda^{k_1} - \lambda_2\|^2 + \|\lambda_2 - \lambda_1\|^2 + 2\langle \lambda^{k_l} - \lambda_2, \lambda_2 - \lambda_1\rangle, \tag{5.19}$$

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

$$\lim_{m\to\infty} \|\lambda^{k_m} - \lambda_2\|^2 - \|\lambda^{k_m} - \lambda_1\|^2 = \|\lambda_2 - \lambda_1\|^2, \tag{5.20}$$

$$\lim_{l\to\infty} \|\lambda^{k_l} - \lambda_1\|^2 - \|\lambda^{k_l} - \lambda_2\|^2 = \|\lambda_2 - \lambda_1\|^2. \tag{5.21}$$

By a.s. convergence of the sequence $\{\|\lambda^k - \lambda\|^2\}$ for all $\lambda \in S$, the limit of each subsequence is equal to the limit of the entire sequence with probability one, so $\lim\limits_{m\to\infty} \|\lambda^{k_m} - \lambda_1\|^2 = \lim\limits_{k\to\infty} \|\lambda^k - \lambda_1\|^2 =: l_1$ and similarly $\lim\limits_{m\to\infty} \|\lambda^{k_m} - \lambda_2\|^2 = \lim\limits_{k\to\infty} \|\lambda^k - \lambda_2\|^2 =: l_2$. Therefore (5.20) and (5.21) imply $l_2 - l_1 = \|\lambda_1 - \lambda_2\|^2 = l_1 - l_2$, meaning $\|\lambda_1 - \lambda_2\|^2 = 0$ and thus the weak limits coincide. Therefore $\{\lambda^k\}$ is weakly convergent to a unique limit with probability one.

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

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

Since $-\langle f(\bar{\lambda}), \lambda^k - \bar{\lambda}\rangle_{L^2(0,T)} > 0$, by optimality of $\bar{\lambda}$, $\lim\limits_{k\to\infty} \mathcal{W}(\bar{\lambda}) - \mathcal{W}(\lambda^k) = 0$ a.s. implies $\lim\limits_{k\to\infty} \|\lambda^k - \bar{\lambda}\|_{L^2(0,T)} = 0$ a.s. ∎

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

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

If F_0 is strictly convex, then we define:

$$\bar{v} := \operatorname*{arg\,min}_{v\in\mathcal{V}} \left\{ F_0(v) + \langle \bar{\lambda}, v \rangle_{L^2(0,T)} \right\}. \tag{5.24}$$

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 the unique minimizer of $\tilde{J}$ in $\mathcal{U}$.

THEOREM 5.8. *Let the Assumptions 3.5 and 5.4 hold, then we have:*
(i) $\{u(\lambda^k)\}$ *weakly converges a.s. to* $\bar{u}$.
If F_0 *is strictly convex, then:*
(ii) $\tilde{J}(u(\lambda^k)) \xrightarrow[k\to\infty]{} \tilde{J}(\bar{u})$ *a.s.*
(iii) $\limsup_{k\to\infty} J(u(\lambda^k)) \leq \inf_{u\in\mathcal{U}} J(u) + 2\,\varepsilon$ *a.s. where* ε *is defined in Theorem 3.6.(ii).*

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

$$(u(\lambda^{\theta_k}), v(\lambda^{\theta_k})) \underset{k\to\infty}{\rightharpoonup} (u^\theta, v^\theta) \in \mathcal{U} \times \mathcal{V}. \tag{5.25}$$

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

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

Using that $z \mapsto F_i(z, X^{i,z}(\omega))) = G_i(z,\omega)$, with $z \in L^2(0,T)$, is w.l.s.c. for any $\omega \in C(0,T)$ (see Remark 2.5) and the a.s. convergence of $\{\lambda^k\}$, resulting from Theorem 5.5.(iv), we have from (5.26) when $k \to \infty$:

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

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

Proof of point (ii).

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

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

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

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

and $v(\lambda^k) \underset{k\to\infty}{\rightharpoonup} \bar{v}$.

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

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

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

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

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

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

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

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

$$\limsup_{k\to\infty} J(u(\lambda^k)) \le \limsup_{k\to\infty} \tilde{J}(u(\lambda^k)) + \varepsilon = \inf_{u\in\mathcal{U}} \tilde{J}(u) + \varepsilon \le \inf_{u\in\mathcal{U}} J(u) + 2\varepsilon\,, \tag{5.33}$$

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

Assumption 5.9. (i) F_0 is strongly convex.
(ii) For any $i \in \{1, \dots, n\}$ and $\omega \in C(0,T)$, the function $u^i \mapsto \mathbb{E}(G_i(u^i, \omega))$ is strongly convex.

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

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

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

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

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

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

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

Proof. The proof is similar to the proof of Lemma 5.10. ∎

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

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

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

$$\check{\mathcal{G}}_k = \sigma\left(\{W^{I_\ell^p,p}\} : 1 \le \ell \le m,\, p \le k\}\right) \vee \sigma\left(\{I_\ell^p\} : 1 \le \ell \le m,\, p \le k\}\right), \tag{5.36}$$

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

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

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

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

The following notations are used:

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

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

  where A and B are diagonal matrices, C is a positive diagonal matrix of size n. The global noise process is a sequence of independent random variables $(W_1, \dots, W_T)$, where for any $t \in \{1, \dots, T\}$, W_t is a vector of centered, reduced and independent Gaussian variables, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$: $W_t := (W_t^1, \dots, W_t^n)$. For any $i \in \{1, \dots, n\}$ and $t \in \{1, \dots, T\}$ we define $\mathcal{F}_t^i := \sigma(W_1^i, \dots, W_t^i)$.
- 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 with: $U_t^i := \{\alpha : \Omega \mapsto \mathbb{R} \,|\, \alpha \,\text{is}\, \mathcal{F}_t^i - \text{measurable and } \alpha(\omega) \in [-M, M] \quad \mathbb{P}\text{-a.s.}\}$, where $M > 0$. We finally set $\mathcal{U} := \prod_{i=1}^n \mathcal{U}^i$.

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

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

and

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

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

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

Proof. The proof of point (i) is the same as for the Lemma 2.6. Similarly, point (ii) is obtained by using the same proof of Theorem 3.6.(ii). ∎

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

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

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

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

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

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

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

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

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

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

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

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

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

The optimization problem (6.4) solved by each local agent is also in the LQG framework. One can solve these problems using the results of [23]. The resolution via Riccati equations of (6.4) shows that $u^i(\lambda^k)$ is a linear function of the state X^{i,u^i} and of the price λ^k. Therefore, in this specific example, for any t one can explicitly 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 (deterministic) Uzawa's algorithm as a reference to evaluate the performances of the *Stochastic Uzawa* algorithm.

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

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

For any n, the deterministic multiplier $\bar{\lambda}^n$ is obtained by applying Uzawa's algorithm, after 10^4 dual iterations. To this end, we applied the *Stochastic Uzawa* Algorithm 5.1 where we ignored the line 8 and we replaced the update of λ^k line 9 by: $\bar{\lambda}^{k+1} \leftarrow \bar{\lambda}^k + \rho_k(\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(u^i(\bar{\lambda}^k)) - v(\bar{\lambda}^k))$.

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

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

For any k and n, $b_{k,n}$, $v_{k,n}$ and $\ell_{k,n}$ denotes respectively an estimation of the bias, the variance and the L2 norm of the error, via Monte Carlo method with J simulations. Thus we have for any k and n: $b_{k,n} = \frac{1}{J}\sum_{j=1}^{J} \lambda^{k,n,j} - \bar{\lambda}^n$, $v_{k,n} = \frac{1}{J}\sum_{j=1}^{J} \|\lambda^{k,n,j} - \bar{\lambda}^n - b_{k,n}\|_2^2$, $\ell_{k,n} = v_{k,n} + \|b_{k,n}\|_2^2$

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

On Figure 6.0.2 we observe a behavior in $1/n^\beta$ (with $\beta \simeq 1$) of the variance $v_{k,n}$ w.r.t. the number of agents n. This is expected, see [5, Theorem 2.2.12, Chapter 2] and observing that the variance of Y^{k+1} is of order $1/n$ for any iteration k.

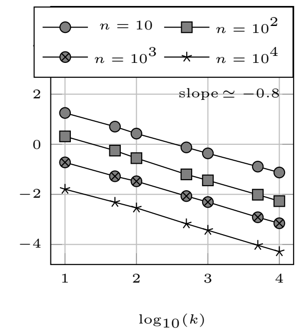

Figure 6.0.1: $\log_{10}(v_{k,n})$ function of k, given the number of agents n

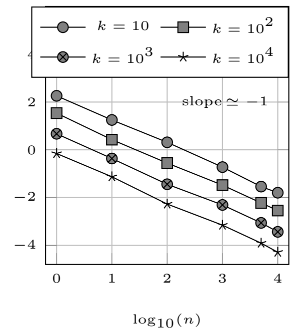

Figure 6.0.2: $\log_{10}(v_{k,n})$ function of n, given the number of agents k

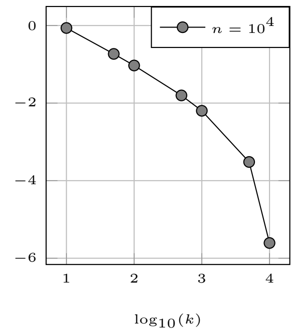

Figure 6.0.3: $\log_{10}(\|b_{k,n}\|_2^2)$ function of k, given the number of agents $n = 10^4$

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

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

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

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

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

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

$$0 \le H_j(t_\ell) \le 1, \tag{7.2}$$

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

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

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

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

$$\bar{q}(t) - \hat{H}(_\ell)\hat{R}(_\ell) \le 0 \tag{7.7}$$

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

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

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

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

The amount of response allocated by each generation technology is limited by the headroom $r_j H_j(t_\ell) G_j^{max}(t_\ell)$ in (7.3) and the slope s_j linking the FR with the dispatch level (7.4). Constraints (7.5) to (7.8) deal with frequency response provision and R_{TCL} (the mean of FR allocated by TCLs). They guaranty secure frequency deviations following sudden generation loss ΔG_L. Inequality (7.5) allocates enough FR (with delivery time t_d) such that the quasi-steady-state frequency remains above Δf_{qss}^{max}, with Λ accounting for the damping effect introduced by the loads [11]. Finally (7.7) constraints the maximum tolerable frequency deviation Δf_{nad}, following the formulation and methodology presented in [22] and [24]. The rate of change of frequency is taken into account in (7.6) where at t_{rcf} the frequency deviation remains above Δf_{ref}. Constraint (7.8) prevents trivial unrealistic solutions that may arise in the proposed formulation, such as high values of committed generation $H_j(t_\ell)$ in correspondence with low (even zero) generation dispatch $G_j(t_\ell)$. The reader can refer to [4] for more details on the UC problem.

The solution of the UC problem, corresponding to the function F_0, can be described by the following optimization problem:

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

subject to equations (7.1)-(7.8).

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

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

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

where:

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

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

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

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

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

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

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

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

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

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

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

  where:
  - $\alpha_i(X_s^{i,u^i} - \bar{X}^i)^2$ is a discomfort term penalizing temperature deviation from some comfort target $\bar{X}$ $[°C]$, with α_i a discomfort term parameter $[£\,/h(°C)^2]$.
  - $\beta_i((X_s^{i,u^i} - X_{\min}^i)_+^2 + (X_{\max}^i - X_s^{i,u^i})_+^2)$ is a penalization term to keep the temperature in the interval $[X_{\min}^i, X_{\max}^i]$, with β_i a target term parameter $[£/s(°C)^2]$ and for any $x \in \mathbb{R}$, $(a)_+ = \max(0, a)$.
- $\gamma_i(X_T^{i,u^i} - \bar{X}_i)^2$ is a terminal cost imposing periodic constraints, with γ a target term parameter $[£/s(°C)^2]$.

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

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

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

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

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

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

7.3. Results. The generation technologies available in the system are nuclear, combined cycle gas turbines (CCGT), open cycle gas turbines (OCGT) and wind.

The characteristics and parameters of the UC in this simulation are the same as in [4].

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

Simulations are performed for different values of volatility $\sigma_i := 0, 1, 2$ (all the TCLs have the same volatility in the simulations), where σ_i is defined in (7.10). The *Sampled Stochastic Uzawa* Algorithm is stopped after 75 iterations or when the relative variation $2\|\lambda^{k+1} - \lambda^k\|_2^2/\|\lambda^{k+1} + \lambda^k\|_2^2$ between two successive prices λ^k and λ^{k+1} is less than 10^{-4}.

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

The system costs (i.e. UC solution) obtained with the flexibility scenario (FS)

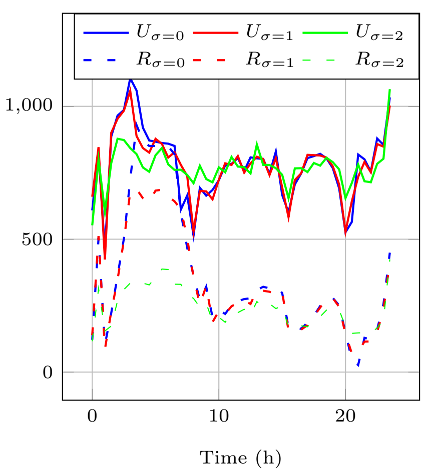

Figure 7.3.1: Total power consumption U and allocated response R (MW) of TCLs after 75 iterations of the algorithm.

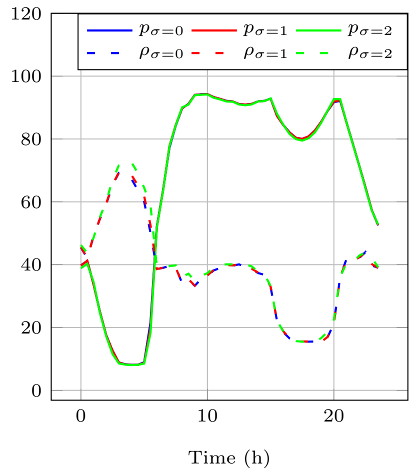

Figure 7.3.2: Electricity price p and response availability price ρ (£/MWh) after 75 iterations of the algorithm.

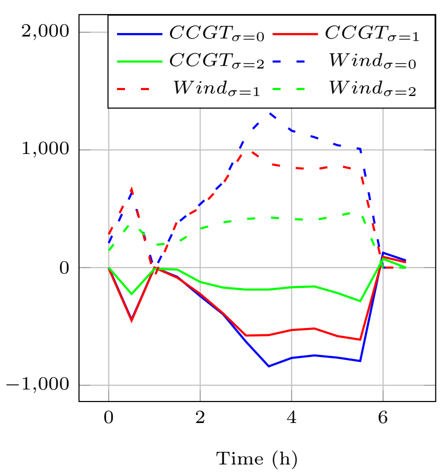

Figure 7.3.3: Deviation of generation profiles (MW) from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

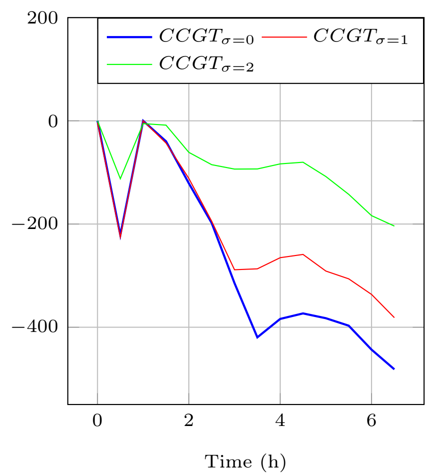

Figure 7.3.4: Deviation of Frequency Response (MW) allocated by CCGT from the "no-flexibility scenario" for three different "flexibility scenario" corresponding to three temperature volatilities.

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

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

Table 1: Minimized system costs in (£)

REFERENCES

[1] Dimitri P Bertsekas and Steven Shreve. *Stochastic optimal control: the discrete-time case.* 2004.

[2] Pierre Carpentier, J-Ph Chancelier, Vincent Leclère, and François Pacaud. Stochastic decomposition applied to large-scale hydro valleys management. *European Journal of Operational Research*, 270(3):1086–1098, 2018.

[3] Michael Chertkov and Vladimir Chernyak. Ensemble of thermostatically controlled loads: Statistical physics approach. *Scientific reports*, 7(1):8673, 2017.

[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.

[5] Marie Duflo. *Random iterative models*, volume 34. Springer Science & Business Media, 2013.

[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.

[7] Pierre Girardeau. *Solving large-scale dynamic stochastic optimization problems.* PhD thesis, Université Paris-Est, December 2010.

[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.

[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.

[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.

[11] Prabha Kundur, Neal J Balu, and Mark G Lauby. *Power system stability and control*, volume 7. McGraw-hill New York, 1994.

[12] Vincent Leclere. *Contributions to decomposition methods in stochastic optimization.* PhD thesis, Université Paris-Est, June 2014.

[13] Xuerong Mao. *Stochastic differential equations and applications.* Elsevier, 2007.

[14] François Pacaud. *Decentralized optimization for energy efficiency under stochasticity.* PhD thesis, Université Paris-Est, October 2018.

[15] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1-3):359–375, 1991.

[16] Pedro Pérez-Aros and Emilio Vilches. An enhanced Baillon-Haddad theorem for convex functions on convex sets. *arXiv preprint arXiv:1904.04885*, 2019.

[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.

[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.

[19] Andrzej Ruszczyński and Alexander Shapiro. Stochastic programming models. *Handbooks in operations research and management science*, 10:1–64, 2003.

[20] David H Salinger. *A splitting algorithm for multistage stochastic programming with application to hydropower scheduling.* PhD thesis, 1997.

[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.

[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.

[23] Emanuel Todorov. Optimal control theory. *Bayesian brain: probabilistic approaches to neural coding*, pages 269–298, 2006.

[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.

[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.