ESAIM: COCV 28 (2022) 40
https://doi.org/10.1051/cocv/2022034

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

EXTENDED MCKEAN-VLASOV OPTIMAL STOCHASTIC CONTROL APPLIED TO SMART GRID MANAGEMENT[*,**]

Emmanuel Gobet[1,***] and Maxime Grangereau[2]

Abstract. We study the mathematical modeling of the energy management system of a smart grid, related to a aggregated consumer equipped with renewable energy production (PV panels *e.g.*), storage facilities (batteries), and connected to the electrical public grid. He controls the use of the storage facilities in order to diminish the random fluctuations of his residual load on the public grid, so that intermittent renewable energy is better used leading globally to a much greener carbon footprint. The optimization problem is described in terms of an extended McKean-Vlasov stochastic control problem. Using the Pontryagin principle, we characterize the optimal storage control as solution of a certain McKean-Vlasov Forward Backward Stochastic Differential Equation (possibly with jumps), for which we prove existence and uniqueness. Quasi-explicit solutions are derived when the cost functions may not be linear-quadratic, using a perturbation approach. Numerical experiments support the study.

Mathematics Subject Classification. 93E20, 49N10, 49N80.

Received January 23, 2021. Accepted April 22, 2022.

1. Introduction

The energy sector is currently facing major changes because of the raising concern about climate change, the search for energy-efficiency and the need to reduce carbon footprint. In particular, the share of renewable energy (RE for short) production has increased in most industrialized countries over the last few years, and further effort has to be done to limit the temperature increase well below 2 °C by 2100, as targeted by the 2015 Paris agreement. However, even if these renewable energies allow a huge reduction of carbon footprint during the energy production phase, they raise a major issue: the amount of energy produced is intermittent

* This work has benefited from several supports: Siebel Energy Institute (Calls for Proposals #2, 2016), ANR project CAESARS (ANR-15-CE05-0024), Association Nationale de la Recherche Technique (ANRT), Electricité De France (EDF), *Finance for Energy Market* (FiME) Lab (Institut Europlace de Finance), Chair "Stress Test, RISK Management and Financial Steering" of the Foundation Ecole Polytechnique.

** This work has been presented at the FOREWER conference Paris-June 2017, at the MCM2017 conference Montreal-July 2017, at the conference "Stochastic control, BSDEs and new developments" Roscoff-September 2017, at the conference "Advances in Stochastic Analysis for Risk Modeling" CIRM-November 2017, at the Workshop "Mean-Field Games, Energy and Environment" London-February 2018, at the conference "Advances in Modelling and Control for Power Systems of the Future" Paris-September 2018, at the conference "APS Informs" Brisbane-July 2019, at the MCM2019 conference Sydney-July 2019. The authors wish to thank the participants for their feedbacks.

Keywords and phrases: Mean-field control, energy management, numerical methods for stochastic control.

[1] Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France.

[2] Electricité de France, Palaiseau, France.

*** Corresponding author: emmanuel.gobet@polytechnique.edu

and uncertain, as a main difference with more conventional energy production units (coal/gas-fired units, or nuclear power plants).

Since the electricity production has to meet consumption at all spatial and time scales, the load balancing operations become harder in this uncertain context, this leads to higher operating costs for the whole electricity system; furthermore, it sometimes lead to ecologically catastrophic solutions such as the use of coal units to compensate the deficit of clean energy production. See [19] for an overview on how to integrate renewables in electricity markets. Therefore, a major challenge is to smooth the electricity consumption by better predicting RE production and better managing the energy system. We address the latter in the context of a consumer equipped with its own RE production (*e.g.* PV panels), and formalize the problem as a stochastic control problem of McKean-Vlasov (MKV for short) type that we solve theoretically and numerically. More specifically, we study a decentralized mechanism aimed at reducing the variability of residual consumption on the electricity network; thus, operating the network could be done at lower costs and with a lower carbon footprint. This mechanism is a setting where a consumer has to commit in advance (say $T =$ one day-ahead, to match the usual working of day-ahead markets) to a predefined load profile and then, he has to command optimally and dynamically his system according to his stochastic consumption/production. Both the optimal load profile and the optimal control are the outputs of the stochastic control problem described below. The above model is a simplified prototype of *smart grid* (as defined by the European Commission[1]): our so-called *consumer* is considered as an association of small consumers, with possibly individual RE production and individual storage facilities, that we aggregate and consider as a whole.

We take the point of view of a consumer supplied in energy by its own intermittent sources (PV panels for instance) and by the electrical public grid. We consider the situation where the non-flexible consumption and the intermittent production are exogenous and can not be predicted perfectly: a stochastic model should be used for both of them. See [4] about a recent methodology for deriving a probabilistic forecast for solar irradiance (and thus PV production). To smooth his residual consumption, the consumer can take advantage of storage facilities (for instance conventional batteries, electrical vehicle batteries, heating network, flywheel etc) which we consider as a whole. At time t, his control is denoted by u_t, the level of storage is represented by X_t^u, its net consumption on the electrical public grid is $\mathtt{P}_t^{\mathtt{grid},u}$. The (deterministic) committed profile load is the curve $(\mathtt{P}_t^{\mathtt{grid,com.}} : 0 \le t \le T)$. Optimal control of a single micro-grid has already been considered in the literature, without the optimal committed load profile. A popular yet without theoretical optimality guarantee is Model Predictive Control [26]. In discrete-time settings, Stochastic Dynamic Programming [18, 29] and Stochastic Dual Dynamic Programming [20] are popular approaches to get theoretical optimality guarantees. Long-term aging of the battery equipping a micro-grid is taken into account by two time-scales time decomposition in [12]. Continuous time optimal control problems are considered in [15] in a deterministic setting, and in [16] in a stochastic environment. By jointly optimizing with the profile $\mathtt{P}^{\mathtt{grid,com.}}$, we change the nature of the stochastic control problem, compared to these works. We shall consider general filtrations with processes possibly exhibiting jumps, to account for sudden variations of solar irradiance or consumption for instance.

In short, in a simplified setting, the optimization criterion takes the form of the following cost functional

$$\mathbb{E}\left[\int_0^T \left\{\mathtt{C}_t \mathtt{P}_t^{\mathtt{grid},u} + \frac{\mu}{2}u_t^2 + \frac{\nu}{2}\left(X_t^u - \frac{1}{2}\right)^2 + l_1\left(\mathtt{P}_t^{\mathtt{grid},u} - \mathtt{P}_t^{\mathtt{grid,com.}}\right)\right\} \mathrm{d}t + \frac{\gamma}{2}\left(X_T^u - \frac{1}{2}\right)^2\right],$$

minimized over admissible controls $(u_t)_t$. The first term in the above cost functional is the cost of buying electricity to the electrical public grid, at a price $\mathtt{C}_t$ which can be random. The second term in the cost functional accounts for a penalization of the use of the storage (*e.g.* aging cost in the case of a battery). The third and fifth terms are penalization of the deviation from the desired state of charge of the storage, which we define as $\frac{1}{2}$ by convention. The fourth term is a penalization (through a convex loss function l_1) of the deviation of the power supplied by the electrical public grid $\mathtt{P}^{\mathtt{grid},u}$ from the commitment profile $\mathtt{P}^{\mathtt{grid,com.}}$. The latter is chosen in such a way that the variability of its residual consumption on the electrical public grid is minimized in a

[1] http://www.ieadsm.org/publication/functionalities-of-smart-grid-and-smart-meters-eutf/

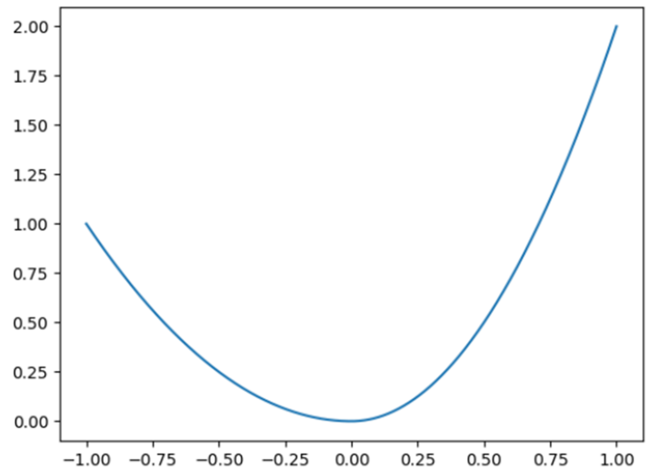


FIGURE 1. Loss function l_1 penalizing more the consumption exceedance.

consistent way. On the side of the electricity supplier on the electrical public grid, since the consumption is close to a deterministic profile chosen in advance, the operating costs are lower and the use of fossil-fueled generation units can be likely avoided. We shall highlight that presumably, good loss functions l_1 should penalize more the consumption exceedance than the consumption deficit: indeed, exceedance possibly requires the use of extra production units with high carbon footprint, this is clearly to discard as often as possible. A typical example of loss function would be:

$$l_1(x) = \alpha x^2 + \alpha_+ \max(x, 0)^2; \tag{1.1}$$

see Figure 1 for an example with $\alpha_+ = 1, \alpha = 1$. This choice is somehow related to generalized risk measures accounting for both left and right tails of the distribution, such as expectiles [6].

If $\mathtt{P}^{\mathtt{grid,com.}}$ were exogenously given, the problem would take the form of a standard stochastic control problem. In our model, it is endogenous and its optimal value at time t is obtained by solving the following stochastic optimization problem:

$$\mathtt{P}_t^{\mathtt{grid,com.}} = \arg\min_p \mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - p)\right]. \tag{1.2}$$

In general, only a numerical solution of this stochastic optimization problem is available. In addition, there is one such problem for each $t \in [0, T]$, hence the set of such problems is continuous thus uncountable. Moreover, these stochastic optimization problems depend on u, which is in turn the optimal solution of a stochastic control problem parameterized by the solutions of the stochastic optimization problems $\mathtt{P}^{\mathtt{grid,com.}}$. Hence, one could try to employ a fixed point iterative procedure, consisting in:

– Given $\mathtt{P}_t^{\mathtt{grid,com.}}$ for all $t \in [0, T]$, solving a stochastic control problem, to find and update u.
– Given u, solving the solutions of the infinite set of stochastic optimization problems (1.2) to update $\mathtt{P}^{\mathtt{grid,com.}}$,
– Then start over again until convergence is reached.

For simplification, we employ a different approach by setting:

$$\mathtt{P}_t^{\mathtt{grid,com.}} = \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right]. \tag{1.3}$$

This amounts to set $\mathtt{P}^{\mathtt{grid,com.}}$ heuristically to a consistent (though no longer optimal in general) value, depending endogenously on u, which allows to obtain a different but close optimization problem for finding good values of u and $\mathtt{P}^{\mathtt{grid,com.}}$. This choice is inspired by the quadratic case for l_1: indeed, solving (1.2) for $l_1 : x \mapsto x^2$ leads to (1.3), as the reader can easily check. Doing so, we obtain a stochastic control problem of MKV type with scalar interactions, see later. As mentioned earlier, with the choice of loss function (1.1) with $\alpha_+ > 0$, the choice $\mathtt{P}_t^{\mathtt{grid,com.}} = \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right]$ is no longer an optimal solution of (1.2) in general. However, we have the following estimation:

$$\begin{aligned}\forall p \in \mathbb{R},\ \mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - p)\right] &\geq \mathbb{E}\left[\alpha(\mathtt{P}_t^{\mathtt{grid},u} - p)^2\right] \\ &\geq \alpha\mathbb{E}\left[(\mathtt{P}_t^{\mathtt{grid},u} - \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right])^2\right] \\ &\geq \frac{\alpha}{\alpha+\alpha_+}\mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right])\right].\end{aligned}$$

In particular, we have:

$$\frac{\alpha}{\alpha+\alpha_+}\mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right])\right] \leq \inf_p \mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - p)\right] \leq \mathbb{E}\left[l_1(\mathtt{P}_t^{\mathtt{grid},u} - \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right])\right]. \tag{1.4}$$

Hence, choosing $\mathtt{P}_t^{\mathtt{grid,com.}} = \mathbb{E}\left[\mathtt{P}_t^{\mathtt{grid},u}\right]$ is a proxy for solving (1.2) with l_1 given by (1.1), which remains good as long as α_+ is small compared to α. Moreover, note that the presence of the expectation of the controlled process $\mathtt{P}^{\mathtt{grid},u}$ in the criteria arises directly from the will to control its probability distribution (by choosing $\mathtt{P}^{\mathtt{grid,com.}}$ as in (1.3)). Therefore, the underlying reason for using a McKean-Vlasov formulation here is conceptually different from another usual application of such models: the asymptotic behavior of a large number of actors interacting, see for instance [2].

Another point to stress is the need to account for jumps in the production/consumption dynamics – *i.e.* the consumption might have discontinuities as appliances/devices are switched-on/off, the power production by a solar panel might suddenly drop to zero if a cloud hides the sun. To summarize, in order to fit application needs, we shall consider non quadratic loss functions and a probabilistic setting of general filtration (allowing jumps).

We embed the previous example in a more general setting:

$$\left.\begin{array}{l}\mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t,\omega,u_t,X_t^u,\mathbb{E}\left[g(t,\omega,u_t,X_t^u)\right]\right)\mathrm{d}t + \psi\left(\omega,X_T^u,\mathbb{E}\left[k(\omega,X_T^u)\right]\right)\right] \\ \text{s.t. } X_t^u = x_0 + \int_0^t \phi(s,\omega,u_s,X_s^u)\mathrm{d}s.\end{array}\right\} \longrightarrow \min_u. \tag{1.5}$$

The functions l, g, ψ, k, ϕ depend on time, control, state variable and on the ambient randomness ω, precise assumptions are given later. Note that the control only appears in the drift of the state variable: we could also have considered a more general model $X_t^u = x + \int_0^t \phi(s,\omega,u_s,X_s^u)\mathrm{d}s + Z_t$ where Z is càdlàg semi-martingale (independent of u), but actually, this extended model is equivalent to the current one by setting $\widetilde{X}_t^u = X_t^u - Z_t$ as a new state variable and by adjusting the (already random) coefficients. Besides, note that the above dynamics for X^u is compatible with usual battery dynamics [17], like for example models of the form

$$\frac{\mathrm{d}\ \text{State of charge}}{\mathrm{d}t} = \text{constant} \cdot \text{Battery power}. \tag{1.6}$$

The problem (1.5) is of McKean-Vlasov (MKV) type since the distribution of (u, X^u) enters into the functional cost. But since this is through generalized moments *via* the functions g and k, the interactions are so-called scalar, which avoids to use the notion of derivatives with respect to probability measures, while maintaining some interesting flexibility. For a full account on control of Stochastic Differential Equations (SDE for short) of MKV type and the link with Mean Field Games, see the recent books [11] and in particular Chapter 6 of Volume I. However, in the above reference, only the distribution of SDE enters in the coefficients, not that of the control as in our setting. We refer to this more general setting as *extended* MKV stochastic optimal control.

Studies in such an extended framework are quite unusual in the literature. In [23], the general discrete case is studied. In [31] and very recently in [5], both the probability distributions of the state and control variables appear in the dynamic of the state and the cost function, but only through their first and second order moments (Linear-Quadratic problems, LQ for short). In [24], the cost functional and the dynamic depend both on the joint probability distribution of the state and control variables, but the authors consider closed-loop controls, which allows them to consider the probability distribution of the state variable only: in our setting, *we do not make any Markovian assumptions* for the characterization of the optimal control. During the preparation of this work (started in 2016), we have been aware of the recent preprint [1] which deals also with the extended MKV stochastic optimal control, with fully non-linear interaction, Markovian dynamics, in the case of a Brownian filtration.

As a difference with the previous references, we do not restrict ourselves to the LQ setting, we deal with extended MKV stochastic optimal control, without Markovian assumptions, and we do not assume that the underlying filtration is Brownian (allowing jump processes). Besides, apart "expected" results about existence/uniqueness, we provide some numerical approximations by using some perturbations analysis around the LQ case. We shall insist that MKV stochastic control is a very recent field and numerical methods are still in their infancy; see [3] for a scheme based on tree methods for solving some MKV Forward-Backward SDE (FBSDE for short) that characterize optimal stochastic controls. Our perturbation approach is different from theirs. As a consequence, we design an effective numerical scheme to address the problem raised by the optimal management of storage facilities able to reduce the variability of residual electricity consumption on the electrical public grid, in the context of uncertain production/consumption of an aggregated consumer. This presumably opens the door to a wider use of these approaches in real smart grid applications.

Now let us go into the details of mathematical/computational arguments. For characterizing the optimal control, we follow a quite standard methodology (see *e.g.* [10]), although details are quite different. This is made in three steps: necessary first order conditions, which become sufficient under additional convexity assumptions, existence of solutions to the first order equations. The derivation of the first order conditions follows the stochastic Pontryagin principle, see for instance [7, 10, 22]. This is achieved for general running and terminal cost functions. In particular, to account for jumps in the production/consumption dynamics, our mathematical analysis is performed in the context of general filtration. It gives rise to an optimality system (see Thms. 2.2 and 2.3), composed of a forward degenerate SDE and of a backward SDE (the adjoint equation), with possibly discontinuous martingale term, and an optimality condition linking the values and probability laws of the state and control variables with the adjoint variable.

In Section 2.4, we establish that this system of equations has a unique solution under some regularity conditions, an invertibility assumption and for small time horizon T (see Thm. 2.4). The condition on T is quite explicit from the proof, which makes the verification on practical examples easy. Here the proof has to be specific and restricted to small time because of non-Brownian filtration and of non-Markovian dynamics: indeed, we can not invoke neither a drift-monotony condition, as in [21], nor a non-degeneracy condition as in [13]. In Section 2.5, we discuss how the unique solution to the first order condition may or may not be the optimal solution; we provide a counter-example (Prop. 2.6), which is interesting for its own, we believe that this kind of situation is already known but we could not find an appropriate reference.

Then we show in Section 2.7 that the necessary optimality conditions established in Theorem 2.3 become sufficient if we assume some convexity conditions on the Hamiltonian and the terminal cost. We shall highlight that the usual Hamiltonian [10] (when the distribution of the control is not optimized) can not match with our framework; alternatively, we define a version in expectation (Lem. 2.9). The final optimality result is stated in Theorem 2.10.

In Section 3, we exemplify our study to the toy model presented in introduction, motivated by practical applications to smart grid management. To get a tractable and effective solution, we perform a perturbation approach around the LQ case. We establish error bounds and as an approximation, we select the expansion with the second order error terms. Numerical experiments illustrate the performance and accuracy of the method, as well the behavior on the optimally controlled system. Long and technical proofs are postponed to Section 4 in order to smooth the reading.

We list the most common notations used in all this work.

▷ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers and positive integers. The notation $|x|$ stands for the Euclidean norm of a vector x, without further reference to its dimension. For a given matrix $A \in \mathbb{R}^p \otimes \mathbb{R}^d$, $A^\top$ refers to its transpose. Its norm is that induced by the Euclidean norm, *i.e.* $|A| := \sup_{x\in\mathbb{R}^d,|x|=1} |Ax|$. Recall that $|A^\top| = |A|$. For $p \in \mathbb{N}^*$, Id_p stands for the identity matrix of size $p \times p$.

▷ *Functions, derivatives.* When a function (or a process) ψ depends on time, we write indifferently $\psi_t(z)$ or $\psi(t, z)$ for the value of ψ at time t, where z represents all other arguments of ψ.

For a smooth function $g : \mathbb{R}^q \mapsto \mathbb{R}^p$, g_x represents the Jacobian matrix of g with respect to x, *i.e.* the matrix $(\partial_{x_j} g_i)_{i,j} \in \mathbb{R}^p \otimes \mathbb{R}^q$. However, a subscript x_t refers to the value of a process x at time t (and not to a partial derivative with respect to t). We also introduce $\nabla_x f := f_x^\top$.

▷ *Probability.* To model the random uncertainty on the time interval $[0, T]$ ($T > 0$ fixed), we consider a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$, we assume that the filtration $\{\mathcal{F}_t\}_{0\le t\le T}$ is right-continuous, augmented with the $\mathbb{P}$-null sets. For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$. Denote by $\mathcal{P}$ the σ-field of predictable sets of $[0, T] \times \Omega$.

All the quantities impacted by the control u are upper-indexed by u, like Z^u for instance.

As usually, càdlàg processes stand for processes that are right continuous with left-hand limits. All the martingales are considered with their càdlàg modifications.

▷ *Spaces.* Let $k \in \mathbb{N}^*$. We define $\mathbb{L}^2([0, T], \mathbb{R}^k)$ as the Banach space of deterministic functions f on $[0, T]$ with values in $\mathbb{R}^k$ such that $\int_0^T |f_t|^2 \mathrm{d}t < +\infty$. Since the arrival space $\mathbb{R}^k$ will be unimportant, we will skip the reference to it in the notation and write the related norms as

$$\|f\|_{\mathbb{L}^2_T} := \left(\int_0^T |f(t)|^2 \mathrm{d}t\right)^{\frac{1}{2}}.$$

Let $p, q \ge 1$. The Banach space of $\mathbb{R}^k$-valued random variables X such that $\mathbb{E}[|X|^p] < +\infty$ is denoted by $\mathbb{L}^p(\Omega, \mathbb{R}^k)$, or simply $\mathbb{L}^p_\Omega$; the associated norm is

$$\|X\|_{\mathbb{L}^p_\Omega} := \mathbb{E}[|X|^p]^{\frac{1}{p}}.$$

The Banach space $\mathbb{H}^{p,q}([0, T] \times \Omega, \mathbb{R}^k)$ (resp. $\mathbb{H}^{p,q}_{\mathcal{P}}([0, T] \times \Omega, \mathbb{R}^k)$) is the set of all $\mathbb{F}$-progressively measurable (resp. $\mathbb{F}$-predictable) processes $\psi : [0, T] \times \Omega \to \mathbb{R}^k$ such that $\int_0^T \mathbb{E}[|\psi_t|^q]^{p/q} \mathrm{d}t < +\infty$. Here again we will omit the reference to $\mathbb{R}^k$, which will be clear from the context. The associated norm is

$$\|\psi\|_{\mathbb{H}^{p,q}} := \left(\int_0^T \mathbb{E}[|\psi_t|^q]^{p/q} \mathrm{d}t\right)^{\frac{1}{p}}.$$

The Banach space $\mathbb{H}^{\infty,q}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ stands for the elements of $\mathbb{H}^{p,q}\left([0, T] \times \Omega, \mathbb{R}^k\right)$ satisfying $\sup_{t\in[0,T]} \mathbb{E}[|\psi_t|^q] < +\infty$, and the related norm is

$$\|\psi\|_{\mathbb{H}^{\infty,q}([0,T]\times\Omega,\mathbb{R}^k)} := \sup_{t\in[0,T]} \mathbb{E}[|\psi_t|^q]^{\frac{1}{q}}.$$

We shall most often consider $p = q = 2$.

2. Stochastic control and MKV-FBSDEs

The aim is to analyze the control problem, about minimizing (1.5). We first discuss the smart grid setting and the class of admissible controls u; second we derive the first-order condition (Pontryagin principle) which writes as a MKV-FBSDE; third we derive sufficient conditions for the existence and uniqueness to the above; fourth in the absence of convexity conditions we provide a counter-example to optimality; last, with suitable convexity assumptions we establish that the MKV-FBSDE solution characterizes the optimal control.

2.1. Stochastic model and smart grid framework

As explained in introduction, (1.5) may describe the optimal energy management of an aggregated consumer, with storage facilities (*e.g.* battery), with his own RE production (*e.g.* building equipped with solar panel), with a connection to the electrical public grid. The management horizon T is typically short, *e.g.* 24 hours for reasons explained in introduction.

The control is made through a $\mathbb{R}^d$-valued vector process $u = (u_t : 0 \leq t \leq T)$, $d \in \mathbb{N}^*$. We consider u as a $\mathcal{F}_t$-predictable process in $\mathbb{H}^{2,2}_{\mathcal{P}}$: the intuition behind it is that decisions occurring at time t have to be made in accordance with the information available up to this time. This is coherent with the smart grid application. In particular, there has to be a slight delay between sudden events and the decisions taken by the controller, whence the predictability assumption.

The dynamics of the system are represented by a $\mathbb{R}^p$-valued state variable, denoted by X, which satisfies the following ODE

$$X^u_{t,\omega} = x_0 + \int_0^t \phi(s, \omega, u_{s,\omega}, X^u_{s,\omega})\mathrm{d}s. \tag{2.1}$$

This state variable includes information about all controlled processes in the smart grid application, like the state of charge of a battery (see (1.6) or [2]), the temperature of a water heater, as in [27] or [28], which dynamics can be modeled by first-order ordinary differential equations. Note that, by an appropriate transformation of the cost functional and a change of variables, our results remain valid if the state process X^u takes the form $X^u = (X^{u,c}, X^{nc})$ with $X^{u,c}$ having a dynamic of the form:

$$X^{u,c}_{t,\omega} = x_0 + \int_0^t \phi(s, \omega, u_{s,\omega}, X^{u,c}_{s,\omega})\mathrm{d}s$$

and X^{nc} being a general uncontrolled semi-martingale (*i.e.*, independent from u). Indeed, by an appropriate transformation of the (random) running and terminal cost, it is possible to consider only the controllable state $X^{u,c}$ as state variable. We are thus free to use very general models for exogenous stochastic processes impacting the system. For instance, in our energy-related application, such exogenous stochastic processes X^{nc} include the electricity spot price, the local electricity production by photo-voltaic panels, the inflexible electricity consumption of a household, or the impact of (exogenous) random water consumption on the temperature of hot water tank.

The cost functional is described by $\mathcal{J}(u)$, given in (1.5). In the smart grid application, Markovian-type costs would take the form, for instance, $l(t, \omega, u, x, \bar{g}) = \tilde{l}(t, Z_t(\omega), u, x, \bar{g})$ where Z would represent a multidimensional stochastic factor modeling the evolution of the exogenous uncontrolled variables (weather, consumption…), but we also allow non Markovian models. In the sequel, we omit ω when we write terms inside $\mathcal{J}(u)$ and X^u, since it is now clear that we deal with random coefficients. All in all, the optimal control problem we study is

$$\left.\begin{array}{l} \mathcal{J}(u) := \mathbb{E}\left[\int_0^T l\left(t, u_t, X^u_t, \mathbb{E}\left[g(t, u_t, X^u_t)\right]\right)\mathrm{d}t + \psi\left(X^u_T, \mathbb{E}\left[k(X^u_T)\right]\right)\right] \\ \text{s.t. } X^u_t = x_0 + \int_0^t \phi(s, u_s, X^u_s)\mathrm{d}s. \end{array}\right\} \longrightarrow \min_{u \in \mathbb{H}^{2,2}_{\mathcal{P}}}. \tag{2.2}$$

Last, we summarize the coefficients from the toy example described in page 3.

Example 2.1 (Smart grid toy example)**.** Let $\mathtt{P}^{\mathtt{load}}$ be the difference between the instantaneous consumer local consumption and his RE production: we assume this is a process in $\mathbb{H}^{2,2}([0,T]\times\Omega,\mathbb{R})$. The control $u\in\mathbb{H}^{2,2}_{\mathcal{P}}([0,T]\times\Omega,\mathbb{R})$ corresponds to the power supplied by the battery, while the state X^u corresponds to the normalized state of charge of the battery which dynamics is linear with respect to the control u, see [15]:

$$X^u_t = x_0 - \frac{1}{\mathcal{E}_{\max}}\int_0^t u_s \mathrm{d}s.$$

If $\mathtt{P}^{\mathtt{grid},u}$ is the power supplied by the electrical public grid, the power balance imposes that

$$\mathtt{P}^{\mathtt{load}}_{t-} = \mathtt{P}^{\mathtt{grid},u}_{t-} + u_t.$$

Then set $d=p=1$ and

$$\begin{aligned}
l(t,\omega,u,x,\bar{g}) &:= \mathtt{C}_{t-}(\omega)\ (\mathtt{P}^{\mathtt{load}}_{t-}(\omega)-u) + \frac{\mu_t}{2}u^2 + \frac{\nu_t}{2}(x-\frac{1}{2})^2 + l_1(\mathtt{P}^{\mathtt{load}}_{t-}(\omega)-u-\bar{g}),\\
g(t,\omega,u,x) &:= \mathtt{P}^{\mathtt{load}}_{t-}(\omega)-u,\\
\psi(\omega,x,\bar{k}) &:= \frac{\gamma}{2}(x-\frac{1}{2})^2,\\
k(\omega,x) &:= 0,\\
\phi(t,u,x) &:= -\frac{u}{\mathcal{E}_{\max}}.
\end{aligned} \tag{2.3}$$

The time-dependent coefficients μ_t and ν_t give the flexibility to include hourly effect in the management. We recall that the convex loss function l_1 may take the form (1.1). Considering the left-hand limit $t-$ in the above definitions is a technicality to fulfill the following assumptions.

2.2. Standing assumptions

From now on, we assume the following hypotheses hold. When we refer to a *constant*, we mean a *finite deterministic constant*.

(H.x) $x_0\in\mathbb{L}^2_\Omega$ and is $\mathcal{F}_0$-measurable.

(H.l) $l:(t,\omega,u,x,\bar{g})\in[0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^q\mapsto l(t,\omega,u,x,\bar{g})\in\mathbb{R}$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)$-measurable. Furthermore, $l(\cdot,\cdot,0,0,0)\in\mathbb{H}^{1,1}$, l is continuously differentiable in $(u,x,\bar{g})$ with the growth condition

$$|\nabla_u l(t,\omega,u,x,\bar{g})| + |\nabla_x l(t,\omega,u,x,\bar{g})| + |\nabla_{\bar{g}} l(t,\omega,u,x,\bar{g})| \le C\ (|u|+|x|+|\bar{g}|) + C^{(0)}_l(t,\omega)$$

for any $(t,u,x,\bar{g})\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^q$ a.s., for some constant C and some random process $C^{(0)}_l$ in $\mathbb{H}^{2,2}$.

(H.g) $g:(t,\omega,u,x)\in[0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^p\mapsto g(t,\omega,u,x)\in\mathbb{R}^q$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $g(\cdot,\cdot,0,0)\in\mathbb{H}^{2,1}$, g is continuously differentiable in (u,x) and there exist constants $C_{g,u}$ and $C_{g,x}$ such that

$$|\nabla_x g(t,\omega,u,x)|\le C_{g,x} \qquad \text{and} \qquad |\nabla_u g(t,\omega,u,x)|\le C_{g,u}$$

for any $(t,u,x)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^p$ a.s. .

(H.ψ) $\psi : (\omega, x, \bar{k}) \in \Omega \times \mathbb{R}^p \times \mathbb{R}^r \mapsto \psi(\omega, x, \bar{k}) \in \mathbb{R}$ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^r)$-measurable. Furthermore, $\psi(\cdot, 0, 0) \in \mathbb{L}^1_\Omega$, ψ is continuously differentiable in $(x, \bar{k})$ and the growth condition

$$|\nabla_x \psi(\omega, x, \bar{k})| + |\nabla_{\bar{k}} \psi(\omega, x, \bar{k})| \leq C \ (|x| + |\bar{k}|) + C^{(0)}_\psi(\omega)$$

holds for any $(x, \bar{k}) \in \mathbb{R}^p \times \mathbb{R}^r$ a.s., for some constant C and some random variable $C^{(0)}_\psi$ in $\mathbb{L}^2_\Omega$.

(H.k) $k : (\omega, x) \in \Omega \times \mathbb{R}^p \mapsto k(\omega, x) \in \mathbb{R}^r$ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $k(\cdot, 0) \in \mathbb{L}^1_\Omega$, k is continuously differentiable in x and there exists a constant $C_{k,x}$ such that

$$|\nabla_x k(\omega, x)| \leq C_{k,x}$$

holds for any $x \in \mathbb{R}^p$ a.s..

(H.ϕ) $\phi : (t, \omega, u, x) \in [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}^p \mapsto \phi(t, \omega, u, x) \in \mathbb{R}^p$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^p)$-measurable. Furthermore, $\phi(\cdot, \cdot, 0, 0) \in \mathbb{H}^{2,2}$, ϕ is continuously differentiable in (u, x) and there exist constants $C_{\phi,u}$ and $C_{\phi,x}$ such that

$$|\nabla_u \phi(t, \omega, u, x)| \leq C_{\phi,u} \qquad \text{and} \qquad |\nabla_x \phi(t, \omega, u, x)| \leq C_{\phi,x}$$

hold for any $(t, u, x) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^p$ a.s..

It is easy to check these conditions in Example 2.1.

As a consequence of **(H.ϕ)**, the dynamics of X^u in (2.1) writes as a ODE with Lipschitz-continuous stochastic coefficient: the uniqueness and existence stem from the Cauchy existence theorem for ODE, applied ω by ω. In addition, we easily show

$$|X^u_t| \leq |x_0| + \int_0^t (|\phi(s, 0, 0)| + C_{\phi,u}|u_s| + C_{\phi,x}|X^u_s|) \, \mathrm{d}s \leq C_T \left(|x_0| + \int_0^t (|\phi(s, 0, 0)| + C_{\phi,u}|u_s|) \, \mathrm{d}s \right)$$

where the second inequality comes from Gronwall's lemma. Then one directly shows that, since u and $\phi(\cdot, 0, 0)$ are in $\mathbb{H}^{2,2}$, X^u is in $\mathbb{H}^{\infty,2} \subset \mathbb{H}^{2,2}$. Then, a careful inspection of the assumptions **(H.l)**-**(H.g)**-**(H.ψ)**-**(H.k)** shows that it implies that the cost $\mathcal{J}(u)$ is finite.

2.3. Necessary condition for optimality

For admissible controls u and v, we now provide a representation of the derivative

$$\dot{\mathcal{J}}(u, v) = \partial_\varepsilon \mathcal{J}(u + \varepsilon v)|_{\varepsilon=0},$$

using an adjoint process Y^u.

Theorem 2.2 (Gâteaux derivatives). *Let $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and set $\bar{g}^u_t := \mathbb{E}\left[g(t, u_t, X^u_t)\right]$. Let $\tilde{L}^u$ be the unique solution of*

$$\tilde{L}^u_0 = \mathrm{Id}_p, \qquad \frac{\mathrm{d}\tilde{L}^u_t}{\mathrm{d}t} = \tilde{L}^u_t \nabla_x \phi(t, u_t, X^u_t).$$

Then $\tilde{L}^u$ is invertible and its inverse satisfies (see Lem. 4.1)

$$(\tilde{L}^u_0)^{-1} = \mathrm{Id}_p, \qquad \frac{\mathrm{d}(\tilde{L}^u_t)^{-1}}{\mathrm{d}t} = -\nabla_x \phi(t, u_t, X^u_t)(\tilde{L}^u_t)^{-1}.$$

Define also $L^u := ((\tilde{L}^u)^{-1})^\top$. *The following* $\mathbb{R}^p$*-valued process* Y^u *is well defined as a càdlàg process in* $\mathbb{H}^{\infty,2}$*:*

$$\begin{aligned} Y_t^u = \mathbb{E}_t & \left[(\tilde{L}_t^u)^{-1} \tilde{L}_T^u \bigg(\nabla_x \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right) + \nabla_x k(X_T^u) \mathbb{E}\left[\nabla_{\bar{k}} \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right] \bigg) \right] \\ & + \mathbb{E}_t \left[\int_t^T (\tilde{L}_t^u)^{-1} \tilde{L}_s^u \bigg(\nabla_x l(s, u_s, X_s^u, \bar{g}_s^u) + \nabla_x g(s, u_s, X_s^u) \mathbb{E}\left[\nabla_{\bar{g}} l(s, u_s, X_s^u, \bar{g}_s^u)\right] \bigg) \mathrm{d}s \right]. \end{aligned} \tag{2.4}$$

In particular, there exists a $\mathbb{R}^p$*-valued càdlàg martingale* M^u *in* $\mathbb{H}^{\infty,2}$*, vanishing at time 0, such that* (Y^u, M^u) *is the unique solution in* $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ *of the following BSDE in* (Y, M)*:*

$$\begin{aligned} -\mathrm{d}Y_t &= \Big(\nabla_x \phi(t, u_t, X_t^u) Y_t + \nabla_x l(t, u_t, X_t^u, \bar{g}_t^u) + \nabla_x g(t, u_t, X_t^u) \mathbb{E}\left[\nabla_{\bar{g}} l(t, u_t, X_t^u, \bar{g}_t^u)\right] \Big) \mathrm{d}t - \mathrm{d}M_t, \\ Y_T &= \nabla_x \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right) + \nabla_x k(X_T^u) \mathbb{E}\left[\nabla_{\bar{k}} \psi\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\right]. \end{aligned} \tag{2.5}$$

Besides, for any $u, v \in \mathbb{H}^{2,2}_{\mathcal{P}}$*, the directional derivative* $\dot{\mathcal{J}}(u, v)$ *exists and is given by*

$$\dot{\mathcal{J}}(u,v) = \mathbb{E}\left[\int_0^T \Big\{ l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u) \Big\} v_t \mathrm{d}t \right].$$

The proof is postponed to Section 4.1. At the optimal control u (whenever it exists), the above derivative $\dot{\mathcal{J}}(u, v)$ must be 0, in any direction $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$. Take for instance v given by:

$$\forall \in [0,T], \qquad v_t := l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u),$$

which ensures that $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ under our assumptions. This justifies the following statement.

Theorem 2.3 (Necessary condition for optimality)**.** *Under the notations and assumptions of Theorem 2.2, if a control* $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ *is optimal, then there exists a unique couple* $(X^u, Y^u) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ *fulfilling* (2.1) *and* (2.4) *such that*

$$l_u(t, u_t, X_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(t, u_t, X_t^u, \bar{g}_t^u)\right] g_u(t, u_t, X_t^u) + (Y_{t-}^u)^\top \phi_u(t, u_t, X_t^u) = 0 \tag{2.6}$$

holds $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$*-a.e.*

2.4. Solvability of the MKV Forward-Backward SDE

Our aim is now to provide sufficient conditions to ensure existence of solution to the system of forward-backward equations (2.1)-(2.4)-(2.6), which we call MKV-FBSDE. For this, we strengthen previous assumptions.

(H.l.2) **(H.l)** holds and there exist constants $C_{l_*,\star}$ where $*$ stands for x and $\bar{g}$, and $\star$ stands for u, x or $\bar{g}$ such that:

$$\begin{aligned} |\nabla_x l(t,\omega,u_1,x_1,\bar{g}_1) - \nabla_x l(t,\omega,u_2,x_2,\bar{g}_2)| &\le C_{l_x,u}\, |u_1-u_2| + C_{l_x,x}|x_1-x_2| + C_{l_x,\bar{g}}|\bar{g}_1-\bar{g}_2|, \\ |\nabla_{\bar{g}} l(t,\omega,u_1,x_1,\bar{g}_1) - \nabla_{\bar{g}} l(t,\omega,u_2,x_2,\bar{g}_2)| &\le C_{l_{\bar{g}},u}\, |u_1-u_2| + C_{l_{\bar{g}},x}|x_1-x_2| + C_{l_{\bar{g}},\bar{g}}|\bar{g}_1-\bar{g}_2| \end{aligned}$$

holds for any $(u_1, u_2, x_1, x_2, \bar{g}_1, \bar{g}_2) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$, $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e..

(H.g.2) **(H.g)** holds and g is affine-linear in x, of the form $g(t,u,x) = a_t^{(g)} x + b^{(g)}(t,u)$.

(H.ψ.2) **(H.ψ)** holds and there exist constants $C_{\psi_*,\star}$ where $*$ and $\star$ stand for x or $\bar{k}$ such that:

$$|\nabla_x\psi(x_1,\bar{k}_1)-\nabla_x\psi(x_2,\bar{k}_2)| \le C_{\psi_x,x}|x_1-x_2|+C_{\psi_x,\bar{k}}|\bar{k}_1-\bar{k}_2|,$$
$$|\nabla_{\bar{k}}\psi(x_1,\bar{k}_1)-\nabla_{\bar{k}}\psi(x_2,\bar{k}_2)| \le C_{\psi_{\bar{k}},x}|x_1-x_2|+C_{\psi_{\bar{k}},\bar{k}}|\bar{k}_1-\bar{k}_2|$$

holds for any $(x_1,x_2,\bar{k}_1,\bar{k}_2)\in\mathbb{R}^p\times\mathbb{R}^p\times\mathbb{R}^r\times\mathbb{R}^r$, $\mathrm{d}t\times\mathrm{d}\mathbb{P}$-a.e..

(H.k.2) **(H.k)** holds and k is affine-linear in x, of the form $k(x)=a^{(k)}x+b^{(k)}$.

(H.ϕ.2) **(H.ϕ)** holds and the dynamic of X^u is affine-linear in x, given by $\phi(t,u,x)=a_t^{(\phi)}x+b^{(\phi)}(t,u)$.

Observe again that this set of conditions is consistent with Example 2.1. We now aim at establishing the solvability of the system composed of (2.1), (2.5) and (2.6). We are going to show that this system has a unique solution for a small enough time horizon T, hence the existence and uniqueness of a solution to the optimal control problem, under the sufficient conditions of Theorem 2.10.

Theorem 2.4. *Assume* **(H.l.2)**-**(H.g.2)**-**(H.ψ.2)**-**(H.k.2)**-**(H.ϕ.2)** *hold. Assume furthermore that there exists a $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)\otimes\mathcal{B}(\mathbb{R}^q)$-measurable function $h:(t,\omega,x,y,\bar{g},\bar{\lambda})\mapsto h(t,\omega,x,y,\bar{g},\bar{\lambda})\in\mathbb{R}^d$ such that*

$$\begin{aligned}&l_u(t,u_t,X_t^u,\bar{g}_t^u)+\mathbb{E}\left[l_{\bar{g}}(t,u_t,X_t^u,\bar{g}_t^u)\right]g_u(t,u_t,X_t^u)+(Y_{t-}^u)^\top\phi_u(t,u_t,X_t^u)=0, \quad \mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e.\\ &\iff u_t=h\left(t,X_t^u,Y_{t-}^u,\bar{g}_t^u,\mathbb{E}\left[\nabla_{\bar{g}}l(t,\omega,u_t,X_t^u,\bar{g}_t^u)\right]\right), \quad \mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e..\end{aligned} \tag{2.7}$$

If h is Lipschitz continuous in $(x,y,\bar{g},\bar{\lambda})$, with Lipschitz constants denoted by $C_{h,x},C_{h,y},C_{h,\bar{g}},C_{h,\bar{\lambda}}$, and if $\Big(h(t,\omega,0,0,0,0)\Big)_{(t,\omega)\in\mathcal{P}}\in\mathbb{H}^{2,2}_{\mathcal{P}}$,

$$\Theta:\begin{cases}\mathbb{H}^{2,2}_{\mathcal{P}} & \to \mathbb{H}^{2,2}_{\mathcal{P}}\\ u & \mapsto \tilde{u}\end{cases},$$

where

$$\Theta(u)_t:=\tilde{u}_t=h\left(t,\omega,X_t^u,Y_{t-}^u,\bar{g}_t^u,\mathbb{E}\left[\nabla_{\bar{g}}l(t,u_t,X_t^u,\bar{g}_t^u)\right]\right), \quad \mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e.,$$

is well defined and Lipschitz continuous. If moreover,

$$C_{h,\bar{g}}C_{g,u}+C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u}+C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)<1, \tag{2.8}$$

then for T small enough, Θ is a contraction and therefore has a unique fixed point $u^\star$. In that case, there exists a unique $u\in\mathbb{H}^{2,2}_{\mathcal{P}}$ satisfying (2.1)-(2.5)-(2.6) and $u=u^\star$.

The proof is available in Section 4.2. Regarding the proof of a fixed point when the time interval $[0,T]$ is arbitrary large, observe that, as a difference with [21] and [13] for instance, in our setting we can rely on a monotony condition of the drifts, nor a non-degeneracy condition. This is why we shall restrict to small time condition.

Remark 2.5. If one can exhibit a $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^p)\otimes\mathcal{B}(\mathbb{R}^q)\otimes\mathcal{B}(\mathbb{R}^q)$-measurable function h such that for all $(\tilde{u},x,y,\bar{g},\bar{\lambda})\in\mathbb{R}^d\times\mathbb{R}^p\times\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^q$:

$$\begin{aligned}&\mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e.,\ l_u(t,\omega,\tilde{u},x,\bar{g})+\bar{\lambda}^\top g_u(t,\omega,\tilde{u},x)+y^\top\phi_u(t,\omega,\tilde{u},x)=0\\ &\iff \mathrm{d}\mathbb{P}\otimes\mathrm{d}t-a.e.,\ \tilde{u}=h\left(t,\omega,x,y,\bar{g},\bar{\lambda}\right),\end{aligned}$$

then (2.7) is satisfied with the same function h.

2.5. Existence and uniqueness of critical point do not necessarily imply existence of a minimum

If there exists a unique solution to the first order optimality condition (unique critical point), and under other assumptions like continuity, growth properties, it is tempting to conclude that this point is a minimum. However, this is not necessarily the case in infinite dimension. This section aims at clarifying this fact by providing an example[2] where continuity, coercivity and unique critical point are ensured, but without existence of minimum. Therefore, extra conditions are necessary to get the existence of a minimum, see later the discussion in Section 2.6.

Proposition 2.6. *Set*

$$F : \begin{cases} \mathbb{L}_1^2 := \mathbb{L}^2([0,1],\mathbb{R}) & \mapsto \mathbb{R} \\ u & \mapsto (\|u\|_{\mathbb{L}_1^2}^2 - 1)^2 + \int_0^1 t|u_t|^2 \mathrm{d}t. \end{cases}$$

Then F satisfies the following properties:

- ***Continuity:*** *F is continuous*
- ***Coercivity:*** *$F(u)$ tends to $+\infty$ when $\|u\|_{\mathbb{L}_1^2}$ tends to $+\infty$*
- ***Existence and uniqueness of critical point:*** *F is Gateaux-differentiable and has a unique critical point.*

However, F does not have a minimum.

The proof is postponed to Section 4.3. The function F defined in this example cannot be quasi-convex (and a fortiori F cannot be convex), since it would then have a minimum, as stated in the next section.

2.6. Existence of an optimal control

We now give sufficient conditions for the existence of an optimal control, *i.e.* existence of a minimizer of $\mathcal{J}$. In such a favorable case, and if the necessary optimality conditions (2.1)-(2.4)-(2.6) have a unique solution u^*, then u^* is the unique minimum of $\mathcal{J}$. We start with a general result.

Theorem 2.7. *Let E be a reflexive Banach space, let $F : E \to \mathbb{R}$ be a lower semi-continuous, quasi-convex function which satisfies the coercivity condition $\lim_{\|u\|_E \to +\infty} F(u) = +\infty$. Then F has a minimum on E.*

Proof. We adapt the arguments of Corollary 3.23, pp. 71 in [9], where the operator considered is assumed to be continuous and convex. However, the hypothesis can be relaxed to lower semi-continuity and quasi-convexity of the function F, since we only need closedness and convexity of the sub-level sets $\Gamma_\alpha^{(F)} := \{u \in E | F(u) \le \alpha\}$ for all $\alpha \in \mathbb{R}$. □

Let us add a few comments. In the finite dimensional case, any lower semi-continuous and coercive function has a minimum (since any closed and bounded set is compact). In the infinite dimensional case, the example in Section 2.5 illustrates that this may be not the case without the quasi-convexity assumption. Besides, note that without the coercivity condition, the existence of minimum may not hold, even in finite dimension (take $E = \mathbb{R}$ and $F(x) = \exp(x)$). Moreover, without the lower semi-continuity of F, the result may fail as well (take $F : (-\infty, 0] \mapsto \mathbb{R}$ defined by $F(x) = |x|\mathbf{1}_{x<0} + \mathbf{1}_{x=0}$, which is coercive and convex).

Apply the previous result with $E = \mathbb{H}^{2,2}$ and $F = \mathcal{J}$: E is an Hilbert space, thus a reflexive Banach space. The functional $\mathcal{J}$ is continuous, hence lower semi-continuous. Therefore, we have proved the following.

Corollary 2.8. *Assume that $\mathcal{J}$ defined in (2.2) is quasi-convex and that $\lim_{\|u\|_{\mathbb{H}^{2,2}} \to +\infty} \mathcal{J}(u) = +\infty$. Then the optimal control problem has a solution $u^* \in \mathbb{H}_{\mathcal{P}}^{2,2}$.*

[2]Such examples might exist in the literature, but we have not been able to find a reference for this.

2.7. Sufficient condition for optimality

Let us now give conditions under which the necessary optimality conditions are sufficient. Additionally to **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**, we assume the following conditions.

(Conv)

- The mapping $\mathcal{T}:\begin{cases}\mathbb{L}^2_\Omega & \to \mathbb{R}\\ X & \mapsto \mathbb{E}\left[\psi(X,\mathbb{E}\left[k(X)\right])\right]\end{cases}$ is convex.
- The mapping $\mathcal{I}:\begin{cases}\mathbb{H}^{2,2}_{\mathcal{P}}\times\mathbb{H}^{\infty,2} & \to \mathbb{R}\\ (\tilde{u},X) & \mapsto \int_0^T \mathbb{E}\left[l\left(t,\tilde{u}_t,X_t,\mathbb{E}\left[g\left(t,\tilde{u}_t,X_t\right)\right]\right)\right]\mathrm{d}t\end{cases}$ is convex.
- The mapping: $\phi:\begin{cases}[0,T]\times\mathbb{R}^d\times\mathbb{R}^p & \to \mathbb{R}^p\\ (t,u,X) & \mapsto \phi(t,u,X)\end{cases}$ is affine-linear in (u,X).

Lemma 2.9. *Under (**Conv**), $\mathcal{J}$ is convex. If furthermore, $\mathcal{I}$ is strictly convex in $\tilde{u}$, or $\mathcal{I}$ is strictly convex in X and ϕ_u has full column rank (which implies $p \geq d$) for almost every t in $[0,T]$, then $\mathcal{J}$ is strictly convex.*

Proof. Under the assumption on ϕ, $u \mapsto X^u$ is affine-linear. This yields the first result using the fact that a composition of an affine-linear function by a convex function is convex. If $\mathcal{I}$ is strictly convex in u then so is $\mathcal{J}$. If ϕ_u has full column rank, $u \mapsto X^u$ is an affine-linear injection and if besides $\mathcal{I}$ is strictly convex in X, we get that $\mathcal{J}$ is strictly convex. □

Let us emphasize the difference with usual stochastic maximum principle (when distributions do not enter in the cost functions). In that case, *i.e.* without the dependence w.r.t. $\mathbb{E}\left[g\left(t,\tilde{u}_t,X_t\right)\right]$ of the running cost and w.r.t. $\mathbb{E}\left[k(X)\right]$ of the terminal cost, the sufficient optimality condition is the affine-linearity in (u,X) of ϕ, the *point-wise* convexity in (u,x) of

$$(t,u,x)\mapsto l(t,u,x),$$

for any t and the *point-wise* convexity of ψ in x.

In the current MKV setting, it would be tempting to require:

$$\xi:(t,u,x)\mapsto l(t,u,x,\mathbb{E}\left[g(t,u,x)\right])$$

to be convex in $(u,X)\in\mathbb{L}^2_\Omega\times\mathbb{L}^2_\Omega$ for any t and

$$X\mapsto\psi(X,\mathbb{E}\left[k(X)\right])$$

to be convex in X in $\mathbb{L}^2_\Omega$. However, even for the simple linear-quadratic case with $d=p=q=1$, *i.e.*

$$l(t,u,x,\bar{g})=(1+\kappa)u^2-\kappa\bar{g}^2,\qquad g(t,u,x)=u,\qquad \phi(t,u,x)=u,\qquad \psi=0,$$

with parameter $\kappa>0$, this fails to be true. Indeed, denoting $\zeta(u)=\xi(t,u,x)$, we get:

$$\zeta\left(\frac{u_1+u_2}{2}\right)-\frac{1}{2}\left(\zeta(u_1)+\zeta(u_2)\right)=\frac{1}{4}\left(\kappa(\mathbb{E}\left[u_1-u_2\right])^2-(1+\kappa)(u_1-u_2)^2\right).$$

Now if u_1 is a Bernoulli random variable with parameter $\frac{1}{2}$, and $u_2=-u_1$, then on the set $\{\omega: u_1(\omega)=u_2(\omega)=0\}$ of positive probability, the above equals $\frac{\kappa}{4}>0$, which violates the convexity condition for these ω. On the contrary, $\mathbb{E}\left[\zeta(\frac{u_1+u_2}{2})-\frac{\zeta(u_1)+\zeta(u_2)}{2}\right]\leq 0$ for $\kappa\geq 0$, and it is easy to see that $\mathbb{E}\left[\zeta(u)\right]$ is convex in u, for such

κ. This discussion clarifies better why the correct convexity condition for the integrated Hamiltonian $\mathcal{I}$ or the point-wise one $\mathcal{H}$ is in expectation and not ω-wise, as stated in **(Conv)**.

We now summarize all the results for having existence and uniqueness of an optimal stochastic control. This is one of the main results of this section.

Theorem 2.10. *Assume* **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**-**(Conv)** *hold.*

– *If $\mathcal{J}$ defined in* (2.2) *satisfies the coercivity condition:*

$$\lim_{\|u\|_{\mathbb{H}^{2,2}} \to +\infty} \mathcal{J}(u) = +\infty,$$

then there exists an optimal control $u^\star \in \mathbb{H}^{2,2}_{\mathcal{P}}$, i.e. a minimum of $\mathcal{J}$ on $\mathbb{H}^{2,2}_{\mathcal{P}}$.

– *$u^\star$ is an optimal control for the problem* (2.2) *if and only if there exists $(X^\star, Y^\star) \in \mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ such that $(u^\star, X^\star, Y^\star)$ fulfills* (2.1)-(2.5)-(2.6).

– *If $\mathcal{J}$ is strictly convex, then it admits at most one minimizer.*

Proof. 1. This is a direct consequence of Theorem 2.8 and Lemma 2.9.
2. If $(u^\star, X^{u^\star}, Y^{u^\star})$ satisfies (2.1)-(2.5)-(2.6), then $\dot{\mathcal{J}}(u^\star, v) = 0$ for any $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ according to Theorem 2.2. Besides, under our assumptions, $\mathcal{J}$ is convex and therefore, for all $v \in \mathbb{H}^{2,2}_{\mathcal{P}}$ and $t \in (0, 1]$,

$$\mathcal{J}(v) - \mathcal{J}(u^\star) \geq \frac{\mathcal{J}(u^\star + t(v - u^\star)) - \mathcal{J}(u^\star)}{t}.$$

By taking the limit when $t \to 0$, we obtain $\mathcal{J}(v) - \mathcal{J}(u^\star) \geq \dot{\mathcal{J}}(u^\star, v - u^\star) = 0$, hence the optimality of $u^\star$. The direct implication $\Rightarrow$ has been established in Theorem 2.3. □

Proposition 2.11. *Assume* **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**-**(Conv)** *hold and suppose that g is affine-linear, that $l : (t, \omega, u, x, \bar{g}) \mapsto l(t, \omega, u, x, \bar{g})$ is jointly convex with respect to $(u, x, \bar{g})$ and strongly convex with respect to u, $\mathrm{d}\mathbb{P} \otimes \mathrm{d}t$-a.e. Then $\mathcal{J}$ is strongly convex, hence coercive.*

Proof. The strong convexity of $\mathcal{J}$ is immediate. Let us show that it implies its coercivity. By strong convexity of $\mathcal{J}$, there exits $\mu > 0$ such that:

$$\forall u, v \in \mathbb{H}^{2,2}_{\mathcal{P}}, \forall \alpha \in [0, 1], \qquad \mathcal{J}(\alpha u + (1 - \alpha)v) \leq \alpha \mathcal{J}(u) + (1 - \alpha)\mathcal{J}(v) - \frac{\alpha(1 - \alpha)\mu}{2} \|u - v\|^2_{\mathbb{H}^{2,2}}.$$

By taking an arbitrary $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$ with $\|u\|_{\mathbb{H}^{2,2}} > 1$, $v = 0$, $\alpha = \frac{1}{\|u\|_{\mathbb{H}^{2,2}}} \in [0, 1]$, we get:

$$\mathcal{J}(u/\|u\|_{\mathbb{H}^{2,2}}) \leq \mathcal{J}(u)/\|u\|_{\mathbb{H}^{2,2}} + (1 - 1/\|u\|_{\mathbb{H}^{2,2}})\, \mathcal{J}(0) - \mu \frac{(\|u\|_{\mathbb{H}^{2,2}} - 1)}{2}.$$

By **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**, $\mathcal{J}$ has at most quadratic growth, *i.e.*, there exits $C > 0$ such that:

$$\forall v \in \mathbb{H}^{2,2}_{\mathcal{P}}, |\mathcal{J}(v)| \leq C(1 + \|v\|_{\mathbb{H}^{2,2}}).$$

From that and the previous estimate, we get:

$$\mathcal{J}(u) \geq -2C\|u\|_{\mathbb{H}^{2,2}} - C\left(\|u\|_{\mathbb{H}^{2,2}} - 1\right) + \mu \frac{\|u\|_{\mathbb{H}^{2,2}}(\|u\|_{\mathbb{H}^{2,2}} - 1)}{2} \underset{\|u\|_{\mathbb{H}^{2,2}} \to +\infty}{\longrightarrow} +\infty,$$

hence the coercivity of $\mathcal{J}$. □

Remark 2.12. The strong convexity of l as defined in its second argument x and full column rank of ϕ_u are not sufficient assumptions to get the coercivity of $\mathcal{J}$, nor its strong convexity. Indeed, define $\mathcal{J}$ by:

$$\mathcal{J}(u) := \mathbb{E}\left[\int_0^T X_t^2 \mathrm{d}t\right]$$
$$s.t. \quad X_t = \int_0^t u_s \mathrm{d}s.$$

This problem satisfies **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**-**(Conv)**. For $0 < \varepsilon < T$, define:

$$u_t^{(\varepsilon)} = \frac{1}{\sqrt{t}} 1_{t \geq \varepsilon}.$$

Then we have $\|u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} = \sqrt{|\ln(T/\varepsilon)|} \to +\infty$ when $\varepsilon \to 0$ while $X_\cdot = \int_0^\cdot u_s^{(\varepsilon)} \mathrm{d}s$ remains bounded uniformly in ε, and hence $\mathcal{J}(u^{(\varepsilon)})$ is uniformly bounded in ε. Therefore, $\mathcal{J}$ is not coercive. As a consequence, it cannot be strongly convex either, by similar arguments as in the proof of Proposition 2.11.

3. Effective computation and approximation of battery control

3.1. Model/context

For simplicity, we assume one-dimensional processes ($p = q = r = 1$), but the results can be easily extended to any dimension, since the arguments are based on the solution of Linear-Quadratic FBSDE, which are well known (see [30]). Let us consider the following toy problem:

$$\min_{u \in \mathbb{H}^{2,2}_{\mathcal{P}}} \mathbb{E}\left[\int_0^T \left\{ \mathtt{C}_{t-} \mathtt{P}^{\mathtt{grid},u}_{t-} + \frac{\mu}{2} u_t^2 + \frac{\nu}{2}\left(X_t^u - \frac{1}{2}\right)^2 + l\left(\mathtt{P}^{\mathtt{grid},u}_{t-} - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid},u}_{t-}\right]\right)\right\} \mathrm{d}t + \frac{\gamma}{2}\left(X_T^u - \frac{1}{2}\right)^2\right]$$
$$\text{s.t.} \quad \begin{cases} X_t^u = x - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s \mathrm{d}s, \\ \mathtt{P}^{\mathtt{grid},u}_{t-} = \mathtt{P}^{\mathtt{load}}_{t-} - u_t. \end{cases}$$

This model is the same as the one presented in the introduction and has the same interpretation. We consider the following hypothesis:

(Toy)
- The parameters μ, ν, γ are deterministic and satisfy $\mu > 0$, $\nu \geq 0$, $\gamma \geq 0$.
- The mapping l is deterministic, convex, continuously differentiable with the growth condition $|l'(x)| \leq C_{l,x}(1 + |x|)$ for all x, for some constant $C_{l,x} > 0$.
- $\mathtt{P}^{\mathtt{load}} \in \mathbb{H}^{2,2}$, $C \in \mathbb{H}^{2,2}$ are $\mathbb{F}$-adapted and càdlàg.

Under assumptions **(Toy)**, **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**-**(Conv)** hold. Besides, one can show the strict convexity of $\mathcal{J}$. Then, it remains to apply Theorem 2.10 to conclude the following.

Proposition 3.1. *Under assumptions* **(Toy)**, *there exists a unique optimal control $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$. Besides, there exist unique processes $X^u \in \mathbb{H}^{\infty,2}$ and $Y^u \in \mathbb{H}^{\infty,2}$ such that (u, X^u, Y^u) satisfies the following McKean-Vlasov Forward Backward SDE:*

$$\begin{cases} X_t^u = x - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s \mathrm{d}s, \\ Y_t^u = \mathbb{E}_t\left[\int_t^T \nu(X_s^u - \frac{1}{2})\mathrm{d}s + \gamma\left(X_T^u - \frac{1}{2}\right)\right], \\ \mu u_t - \mathtt{C}_{t-} - l'\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right) + \mathbb{E}\left[l'\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)\right] = \frac{Y^u_{t-}}{\mathcal{E}_{\max}}. \end{cases} \tag{3.1}$$

Although we can derive specific results for the control problem under assumption **(Toy)** (see Props. 3.1 and 3.4), solving explicitly the system (3.1) remains difficult for general convex l. To get approximation results, we consider a specific form of l.

(ToyBis) The mapping l is given by $l(x) := \frac{\lambda}{2}x^2 + \frac{\varepsilon(\lambda+\mu)}{2}(x_+)^2$ with $\lambda \geq 0$, $|\varepsilon| < 1$.

From the application point of view, we remind that we want to penalize more consumption excess (compared to the commitment) than consumption deficit. The asymmetry parameter ε should thus be taken non-negative. Under assumptions **(Toy)** and **(ToyBis)**, the last equation in (3.1) writes:

$$(\lambda+\mu)u_t - \lambda\mathbb{E}\left[u_t\right] - \mathtt{C}_{t-} - \lambda(\mathtt{P}^{\mathtt{load}}_{t-} - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-}\right]) - \varepsilon(\lambda+\mu)\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)_+$$
$$+\,\varepsilon(\lambda+\mu)\mathbb{E}\left[\left(\mathtt{P}^{\mathtt{load}}_{t-} - u_t - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}_{t-} - u_t\right]\right)_+\right] = \frac{Y^u_{t-}}{\mathcal{E}_{\max}}.$$

We now provide a first order expansion of the solution of this problem with respect to the parameter $\varepsilon \to 0$.

3.2. Computation of first order expansion

3.2.1. Preliminary result

The computation of a first order expansion of the solution of the MKV FBSDE (3.1) will rely extensively on the following result.

Proposition 3.2. *Let a, b, c, e, f, g be deterministic real parameters with $a > 0$, $g > 0$, $b \geq 0$ and $e \geq 0$. Let $(h_t)_t$ be a stochastic process in $\mathbb{H}^{2,2}_{\mathcal{P}}$ and $x_0 \in \mathbb{L}^2(\Omega)$ be $\mathcal{F}_0$-measurable. Define:*

$$\theta_t := \begin{cases} \frac{1}{2}\left(1+e\sqrt{\frac{ag}{b}}\right)\exp(\sqrt{abg}(T-t)) + \frac{1}{2}\left(1-e\sqrt{\frac{ag}{b}}\right)\exp(\sqrt{abg}(t-T)) & \text{if } b>0,\\ eag(T-t)+1 & \text{if } b=0, \end{cases} \tag{3.2}$$

$$p_t := -\frac{\mathrm{d}\theta_t}{\mathrm{d}t}\frac{1}{ag\theta_t}, \tag{3.3}$$

$$\pi_t = \frac{1}{\theta_t}\left(f - \int_t^T (ap_s\mathbb{E}_t\left[h_s\right] - c)\theta_s \mathrm{d}s\right). \tag{3.4}$$

Define x, y and v by:

$$\begin{cases} x_t = x_0\frac{\theta_t}{\theta_0} - \int_0^t (ag\pi_s + ah_s)\frac{\theta_t}{\theta_s}\mathrm{d}s,\\ y_t = p_t x_t + \pi_t,\\ v_t = gp_t x_t + g\pi_{t-} + h_t. \end{cases} \tag{3.5}$$

Then (x, y, v) is a solution in $\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{2,2}_{\mathcal{P}}$ of the Forward-Backward system:

$$\begin{cases} x_t = x_0 - \int_0^t av_s\mathrm{d}s,\\ y_t = \mathbb{E}_t\left[\int_t^T (bx_s + c)\mathrm{d}s + ex_T + f\right],\\ v_t = gy_{t-} + h_t. \end{cases} \tag{3.6}$$

Besides, for T small enough, this solution to (3.6) is the unique one in $\mathbb{H}^{\infty,2}\times\mathbb{H}^{\infty,2}\times\mathbb{H}^{2,2}_{\mathcal{P}}$.

The proof is postponed to Section 4.4.

Remark 3.3. Uniqueness of the solution of the FBSDE (3.6) could be proved for arbitrary time horizon T, using the fact that (3.6) characterizes the solution of a (linear-quadratic) stochastic control which has a unique solution (as the associated cost function is continuous, convex and coercive [9], Corol. 3.23, pp. 71).

3.2.2. Average processes

We introduce the following notations for the average (in the sense of expectation) of the solutions of (3.1):

$$\bar{u} := \mathbb{E}[u], \qquad \bar{X} := \mathbb{E}[X^u], \qquad \bar{Y} := \mathbb{E}[Y^u], \qquad \bar{\mathtt{C}} := \mathbb{E}[\mathtt{C}].$$

By taking the expectation in (3.1), we immediately get the following simple but remarkable result: the average processes do not depend on l.

Proposition 3.4. *Under Assumption* (**Toy**), $(\bar{u}, \bar{X}, \bar{Y})$ *solves*

$$\begin{cases} \bar{X}_t = \mathbb{E}[x] - \frac{1}{\mathcal{E}_{\max}} \int_0^t \bar{u}_s \mathrm{d}s, \\ \bar{Y}_t = \int_t^T \nu(\bar{X}_s - \frac{1}{2})\mathrm{d}s + \gamma\left(\bar{X}_T - \frac{1}{2}\right), \\ \bar{u}_t = \frac{\bar{Y}_{t-}}{\mu\mathcal{E}_{\max}} + \frac{\bar{\mathtt{C}}_{t-}}{\mu}. \end{cases} \tag{3.7}$$

In particular, $(\bar{u}, \bar{X}, \bar{Y})$ *does not depend on* l.

Note that the FBSDE (3.7) is explicitly solvable, as a particular case of equation (3.6) with $x_0 := x$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = -\frac{\nu}{2}$, $e = \gamma$, $f = -\frac{\gamma}{2}$, $g = \frac{1}{\mu\mathcal{E}_{\max}}$ and $h_t = \frac{\bar{\mathtt{C}}_{t-}}{\mu}$.

3.2.3. Notations

From now on, assume that (**Toy**) and (**ToyBis**) hold. From Proposition 3.4, $(\bar{u}, \bar{X}, \bar{Y})$ does not depend on ε. We denote the processes u, X^u and Y^u by $u^{(\varepsilon)}$, $X^{(\varepsilon)}$ and $Y^{(\varepsilon)}$ respectively to insist on the dependency w.r.t. the parameter ε. $(u^{(\varepsilon)}, X^{(\varepsilon)}, Y^{(\varepsilon)})$ satisfies (3.1) with $l'(x) = \lambda x + \varepsilon(\lambda + \mu)x_+$.

For the ease of the proofs, let us define the recentered processes

$$\begin{aligned} &u^{\Delta,(\varepsilon)} := u^{(\varepsilon)} - \bar{u}, \qquad X^{\Delta,(\varepsilon)} := X^{(\varepsilon)} - \bar{X}, \qquad Y^{\Delta,(\varepsilon)} := Y^{(\varepsilon)} - \bar{Y}, \\ &\mathtt{P}^{\mathtt{load},\Delta} := \mathtt{P}^{\mathtt{load}} - \mathbb{E}\left[\mathtt{P}^{\mathtt{load}}\right], \qquad \mathtt{C}^{\Delta} := \mathtt{C} - \mathbb{E}[\mathtt{C}]. \end{aligned}$$

Then, $(u^{\Delta,(\varepsilon)}, X^{\Delta,(\varepsilon)}, Y^{\Delta,(\varepsilon)})$ satisfies:

$$\begin{cases} X_t^{\Delta,(\varepsilon)} = x - \mathbb{E}[x] - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s^{\Delta,(\varepsilon)} \mathrm{d}s, \\ Y_t^{\Delta,(\varepsilon)} = \mathbb{E}_t\left[\int_t^T \nu X_s^{\Delta,(\varepsilon)} \mathrm{d}s + \gamma X_T^{\Delta,(\varepsilon)}\right], \\ \mu u_t^{\Delta,(\varepsilon)} - \mathtt{C}_{t-}^{\Delta} - \lambda\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right) - \varepsilon(\lambda+\mu)\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+ + \varepsilon(\lambda+\mu)\mathbb{E}\left[\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(\varepsilon)}\right)_+\right] = \frac{Y_{t-}^{\Delta,(\varepsilon)}}{\mathcal{E}_{\max}}. \end{cases} \tag{3.8}$$

We now seek a first order expansion of the solution of (3.1) w.r.t. ε, as $\varepsilon \to 0$, and equivalently, as the average processes do not depend on ε (see Prop. 3.4), we will perform it for the recentered processes, by showing

$$u^{\Delta,(\varepsilon)} = u^{\Delta,(0)} + \varepsilon\dot{u} + o(\varepsilon), \qquad X^{\Delta,(\varepsilon)} = X^{\Delta,(0)} + \varepsilon\dot{X} + o(\varepsilon), \qquad Y^{\Delta,(\varepsilon)} = Y^{\Delta,(0)} + \varepsilon\dot{Y} + o(\varepsilon),$$

where $\dot{u}$, $\dot{X}$ and $\dot{Y}$ are suitable processes in $\mathbb{H}_{\mathcal{P}}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ (independent of ε) and the convergence $o(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$ holds in $\mathbb{H}^{2,2}$-norm.

Proposition 3.5. *Assume (Toy) and (ToyBis). Then $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ satisfies:*

$$\begin{cases} X_t^{\Delta,(0)} = x - \mathbb{E}[x] - \frac{1}{\mathcal{E}_{\max}} \int_0^t u_s^{\Delta,(0)} \mathrm{d}s, \\ Y_t^{\Delta,(0)} = \mathbb{E}_t \left[\int_t^T \nu X_s^{\Delta,(0)} \mathrm{d}s + \gamma X_T^{\Delta,(0)} \right], \\ u_t^{\Delta,(0)} = \frac{Y_{t-}^{\Delta,(0)}}{(\lambda+\mu)\mathcal{E}_{\max}} + \frac{\mathtt{C}_{t-}^{\Delta} + \lambda \mathtt{P}_{t-}^{\mathtt{load},\Delta}}{\lambda+\mu}. \end{cases} \tag{3.9}$$

Observe that the FBSDE (3.9) is known in a closed form, as a particular case of equation (3.6) with $x_0 := x - \mathbb{E}[x]$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = 0$, $e = \gamma$, $f = 0$, $g = \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$ and $h_t = \frac{\mathtt{C}_{t-}^{\Delta} + \lambda \mathtt{P}_{t-}^{\mathtt{load},\Delta}}{\lambda+\mu}$.

Proposition 3.6. *Assume (Toy) and (ToyBis). Define the finite differences*

$$\dot{u}^{(\varepsilon)} := \frac{u^{\Delta,(\varepsilon)} - u^{\Delta,(0)}}{\varepsilon}, \qquad \dot{X}^{(\varepsilon)} := \frac{X^{\Delta,(\varepsilon)} - X^{\Delta,(0)}}{\varepsilon}, \qquad \dot{Y}^{(\varepsilon)} := \frac{Y^{\Delta,(\varepsilon)} - Y^{\Delta,(0)}}{\varepsilon},$$

which solve

$$\begin{cases} \dot{X}_t^{(\varepsilon)} = -\frac{1}{\mathcal{E}_{\max}} \int_0^t \dot{u}_s^{(\varepsilon)} \mathrm{d}s, \\ \dot{Y}_t^{(\varepsilon)} = \mathbb{E}_t \left[\int_t^T \nu \dot{X}_s^{(\varepsilon)} \mathrm{d}s + \gamma \dot{X}_T^{(\varepsilon)} \right], \\ \dot{u}_t^{(\varepsilon)} = \frac{\dot{Y}_{t-}^{(\varepsilon)}}{(\lambda+\mu)\mathcal{E}_{\max}} + \left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(\varepsilon)} \right)_+ - \mathbb{E}\left[\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(\varepsilon)} \right)_+ \right]. \end{cases} \tag{3.10}$$

Besides, for small enough time horizon T, $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ is uniformly bounded in $\mathbb{H}_{\mathcal{P}}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ as $\varepsilon \to 0$. Define $(\dot{u}, \dot{X}, \dot{Y})$ as a solution (unique when T is small enough) to

$$\begin{cases} \dot{X}_t = -\frac{1}{\mathcal{E}_{\max}} \int_0^t \dot{u}_s \mathrm{d}s, \\ \dot{Y}_t = \mathbb{E}_t \left[\int_t^T \nu \dot{X}_s + \gamma \dot{X}_T \mathrm{d}s \right], \\ \dot{u}_t = \frac{\dot{Y}_{t-}}{(\lambda+\mu)\mathcal{E}_{\max}} + \left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(0)} \right)_+ - \mathbb{E}\left[\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(0)} \right)_+ \right]. \end{cases} \tag{3.11}$$

Then, for small enough time horizon T, the finite differences $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ are close (at order 1 in ε) to $(\dot{u}, \dot{X}, \dot{Y})$:

$$\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}_{\mathcal{P}}^{2,2}} + \|\dot{X}^{(\varepsilon)} - \dot{X}\|_{\mathbb{H}^{2,2}} + \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} = \mathcal{O}(\varepsilon).$$

The proof is postponed to Section 4.5. Note again that the FBSDE (3.11) is explicitly solvable, as a particular case of equation (3.6) with $x_0 := 0$, $a = \frac{1}{\mathcal{E}_{\max}}$, $b = \nu$, $c = 0$, $e = \gamma$, $f = 0$, $g = \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$ and $h_t = \left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(0)} \right)_+ - \mathbb{E}\left[\left(\mathtt{P}_{t-}^{\mathtt{load},\Delta} - u_t^{\Delta,(0)} \right)_+ \right]$.

Collecting all the previous results, we get the following theorem, which fully characterizes the first order expansion of the solution to the control problem.

Theorem 3.7. *Assume (Toy) and (ToyBis) hold. For small enough time horizon T, the unique solution $(u^{(\varepsilon)}, X^{(\varepsilon)}, Y^{(\varepsilon)})$ in $\mathbb{H}_{\mathcal{P}}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{H}^{2,2}$ of (3.1) can be expanded at first order w.r.t. ε (with error of second order as $\varepsilon \to 0$):*

$$u^{(\varepsilon)} = \bar{u} + u^{\Delta,(0)} + \varepsilon \dot{u} + \mathcal{O}(\varepsilon^2), \qquad X^{(\varepsilon)} = \bar{X} + X^{\Delta,(0)} + \varepsilon \dot{X} + \mathcal{O}(\varepsilon^2), \qquad Y^{(\varepsilon)} = \bar{Y} + Y^{\Delta,(0)} + \varepsilon \dot{Y} + \mathcal{O}(\varepsilon^2),$$

where errors $\mathcal{O}(\varepsilon^2)$ are measured in $\mathbb{H}^{2,2}$-norm, with $(\bar{u}, \bar{X}, \bar{Y})$ solution of (3.7), $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ solution of (3.9) and $(\dot{u}, \dot{X}, \dot{Y})$ solution of (3.11).

We shall emphasize that all terms in these expansions are solutions of FBSDEs of the form (3.6) for different input parameters (see Tab. 1) and thus they are explicitly solvable.

For other problems with more regularity (notice that $x \mapsto (x_+)^2$ is not twice continuously differentiable), the previous approach could actually be extended to a second order expansion or even higher order, but it would lead to more and more nested FBSDEs: on the mathematical side, there is no hard obstacle to derive these equations under appropriate regularity conditions. The concerns would be rather on the computational side since it would require larger and larger computational time.

Remark 3.8. Theorem 3.7 is not specific to the form of the loss function l given in **(ToyBis)** and it could be extended for any loss function l provided that it is continuously differentiable with Lipschitz-continuous derivative.

3.3. Effective simulation of first order expansion of optimal control

3.3.1. Models for random uncertainties

We assume the electricity price $\mathtt{C}$ is constant ($\bar{\mathtt{C}} = \mathtt{C}$ and $\mathtt{C}^\Delta = 0$), and we suppose $\mathtt{P}^{\mathtt{load}}$ is given by $\mathtt{P}^{\mathtt{load}} = \mathtt{P}^{\mathtt{cons}} - \mathtt{P}^{\mathtt{sun}}$, where $\mathtt{P}^{\mathtt{cons}}$ and $\mathtt{P}^{\mathtt{sun}}$ are two independent scalar SDEs[3], representing respectively the consumption and the photo-voltaic power production. For the consumption $\mathtt{P}^{\mathtt{cons}}$, we use the jump process:

$$\mathrm{d}\mathtt{P}^{\mathtt{cons}}_t = -\rho^{\mathtt{cons}}(\mathtt{P}^{\mathtt{cons}}_t - \mathtt{p}^{\mathtt{cons,ref}}_t)\mathrm{d}t + h^{\mathtt{cons}}\mathrm{d}N^{\mathtt{cons}}_t, \tag{3.12}$$

where $N^{\mathtt{cons}}$ is a compensated Poisson Process with intensity $\lambda^{\mathtt{cons}}$. Regarding the PV production, we follow [4] by setting $\mathtt{P}^{\mathtt{sun}} = \mathtt{P}^{\mathtt{sun,max}}\mathtt{X}^{\mathtt{sun}}$ where $\mathtt{P}^{\mathtt{sun,max}} : [0,T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) and $\mathtt{X}^{\mathtt{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathtt{X}^{\mathtt{sun}}_t = -\rho^{\mathtt{sun}}(\mathtt{X}^{\mathtt{sun}}_t - \mathtt{x}^{\mathtt{sun,ref}}_t)\mathrm{d}t + \sigma^{\mathtt{sun}}(\mathtt{X}^{\mathtt{sun}}_t)^{\alpha}(1 - \mathtt{X}^{\mathtt{sun}}_t)^{\beta}\mathrm{d}W_t, \tag{3.13}$$

with $\alpha, \beta \geq 1/2$. As proved in [4], there is a strong solution to the above SDE and the solution $\mathtt{X}^{\mathtt{sun}}$ takes values in $[0,1]$.

Since the drifts are affine-linear, the conditional expectation of the solution is known in closed forms (this property is intensively used in [8]):

$$\mathbb{E}_t\left[\mathtt{P}^{\mathtt{sun}}_s\right] = \left(\frac{\mathtt{P}^{\mathtt{sun}}_t}{\mathtt{P}^{\mathtt{sun,max}}_t}\exp(-\rho^{\mathtt{sun}}(s-t)) + \int_t^s \rho^{\mathtt{sun}}\mathtt{x}^{\mathtt{sun,ref}}_\tau \exp(-\rho^{\mathtt{sun}}(s-\tau))\mathrm{d}\tau\right)\mathtt{P}^{\mathtt{sun,max}}_s, \tag{3.14}$$

$$\mathbb{E}_t\left[\mathtt{P}^{\mathtt{cons}}_s\right] = \mathtt{P}^{\mathtt{cons}}_t\exp(-\rho^{\mathtt{cons}}(s-t)) + \int_t^s \rho^{\mathtt{cons}}\mathtt{p}^{\mathtt{cons,ref}}_\tau \exp(-\rho^{\mathtt{cons}}(s-\tau))\mathrm{d}\tau, \tag{3.15}$$

for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\mathtt{P}^{\mathtt{load}}_s\right]$ as required when deriving the optimal control.

3.3.2. FBSDE parameters

Proposition 3.2 is repeatedly used to solve the affine-linear FBSDEs $(\bar{u}, \bar{X}, \bar{Y})$, $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ and $(\dot{u}, \dot{X}, \dot{Y})$ arising in the first order expansion of the optimal control w.r.t. ε (see Thm. 3.7). In Algorithm 1 we give the pseudo-code of the scheme used to compute solutions of the FBSDE of the form (3.6).

[3]We consider Brownian SDEs for simplicity, but note that the current setting allows more general processes.

Algorithm 1 Sample of a path of (x, y, v), solution of (3.6)

1: **Inputs:** $x_0 \in \mathbb{L}^2_\Omega, a > 0, b \geq 0, c \in \mathbb{R}, e \geq 0, f \in \mathbb{R}, g > 0, h \in \mathbb{H}^{2,2}_{\mathcal{P}},\ N_T > 0$
2: Sample x_0 and set $X(0) \leftarrow x_0$. Set $\tau = \frac{T}{N_T}$.
3: **for** $n = 0, \dots, N_T - 1$ **do**
4: Compute the conditional expectations $(\mathbb{E}_{n\tau}[h_s])_{n\tau \leq s \leq T}$
5: Compute $\pi(n\tau)$ by numerical integration, as given in (3.4)
6: Compute $p(n\tau)$ as in (3.3)
7: $v(n\tau) \leftarrow gp(n\tau)X(n\tau) + g\pi(n\tau) + h(n\tau)$
8: $x((n+1)\tau) \leftarrow x(n\tau) - av(n\tau)\tau$
9: **end for**
10: **return** (x, y, v)

TABLE 1. Table of parameters needed to compute the expansion terms.

	$(\bar{u}, \bar{X}, \bar{Y})$	$(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$	$(\dot{u}, \dot{X}, \dot{Y})$
a	$\frac{1}{\mathcal{E}_{\max}}$		
b	ν		
c	$\frac{-\nu}{2}$	0	
e	γ		
f	$\frac{-\gamma}{2}$	0	
g	$\frac{1}{\mu \mathcal{E}_{\max}}$	$\frac{1}{(\lambda+\mu)\mathcal{E}_{\max}}$	
h_t	$\frac{\bar{\mathtt{C}}_{t-}}{\mu}$	$\frac{\mathtt{C}^{\Delta}_{t-} + \lambda \mathtt{P}^{\mathtt{load},\Delta}_{t-}}{\lambda+\mu}$	$\left(\mathtt{P}^{\mathtt{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+ - \mathbb{E}\left[\left(\mathtt{P}^{\mathtt{load},\Delta}_{t-} - u^{\Delta,(0)}_t\right)_+\right]$

In Table 1, we give the correspondence between the input parameters $(a, b, c, d, e, f, g, h_t)$ for the generic FBSDE of Proposition 3.2 and the parameters defining the 3 FBSDEs. Merged columns indicate common values of parameters. As the data involved in the system defining $(\bar{u}, \bar{X}, \bar{Y})$ is deterministic, one only needs to perform numerical integrations to compute π and therefore $(\bar{u}, \bar{X}, \bar{Y})$. For $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$ and $(\dot{u}, \dot{X}, \dot{Y})$, it becomes a bit more involved. Let us provide some details on the implementation.

- For the computation of $(u^{\Delta,(0)}, X^{\Delta,(0)}, Y^{\Delta,(0)})$, the conditional expectations $(\mathbb{E}_{n\tau}[h_s])_{n\tau \leq s \leq T}$ are given by affine-linear combinations of $\mathtt{P}^{\mathtt{cons}}_{n\tau}$ and $\mathtt{P}^{\mathtt{sun}}_{n\tau}$ with deterministic coefficients, depending on s and n, by assumption on our models for $\mathtt{P}^{\mathtt{cons}}$ and $\mathtt{P}^{\mathtt{sun}}$ (see (3.14)–(3.15)). Therefore, $\pi(n\tau)$ is also given by an affine-linear combination of $\mathtt{P}^{\mathtt{cons}}_{n\tau}$ and $\mathtt{P}^{\mathtt{sun}}_{n\tau}$ with deterministic coefficients. This allows to speed up Steps 4 and 5 in Algorithm 1.
- For the computation of $(\dot{u}, \dot{X}, \dot{Y})$, the conditional expectations $\left(\mathbb{E}_{n\tau}\left[\left(\mathtt{P}^{\mathtt{load},\Delta}_s - u^{\Delta,(0)}_s\right)_+\right]\right)_{n\tau \leq s \leq T}$ at Step 4 is estimated by Monte-Carlo methods. The procedure for doing so is given in Algorithm 2. This Step 4 has a complexity of order $\mathcal{O}((N_T - n)M_0)$, which is the most costly Step in the loop of Algorithm 1; hence sampling $(\dot{u}, \dot{X}, \dot{Y})$ has a computational cost of order $\mathcal{O}(N_T^2 M_0)$.

3.3.3. *Numerical values of parameters*

We report the values chosen for the next experiments.

Parameters for smart grid. We consider the following values for the time horizon, the size of the storage system and the initial value of its normalized state of charge.

Parameter	T	$\mathcal{E}_{\max}$	x_0
Value	24 h	200 kWh	0.5

Parameters for uncertain consumption/production. The following table gives the values of the parameters used in the modeling of the underlying exogenous stochastic processes impacting the system.

Algorithm 2 Evaluation of $\left(\mathbb{E}_{n\tau}\left[\left(\mathtt{P}_s^{\mathtt{load},\Delta} - u_s^{\Delta,(0)}\right)_+\right]\right)_{s=n\tau,\ldots,N_T\tau}$

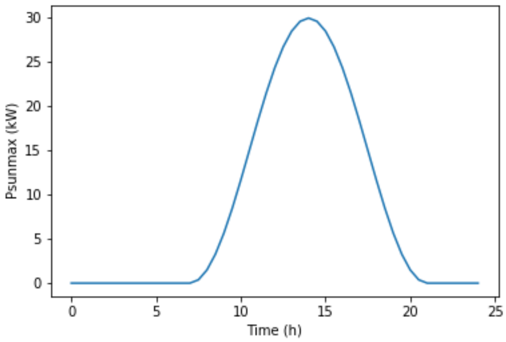
Inputs: $n < N_T$, $X_{n\tau}^{\Delta,(0)}$, $\mathtt{P}_{n\tau}^{\mathtt{sun}}$, $\mathtt{P}_{n\tau}^{\mathtt{load}}$, $M_0 > 0$
Initialization: $(R[n], R[n+1], ..., R[N_T]) \leftarrow (0, 0, \ldots, 0)$.
Compute $u^{\Delta,(0)}(n\tau)$ using a similar procedure as in Algorithm 1.
$R[n] \leftarrow \left(\mathtt{P}_{n\tau}^{\mathtt{load},\Delta} - u_{n\tau}^{\Delta,(0)}\right)_+$.
for $m = 1, \ldots, M_0$ **do**
  for $k = n+1, \ldots, N_T$ **do**
    Sample $(\mathtt{P}_{k\tau}^{\mathtt{cons}}, \mathtt{P}_{k\tau}^{\mathtt{sun}})$ conditionally to $(\mathtt{P}_{(k-1)\tau}^{\mathtt{cons}}, \mathtt{P}_{(k-1)\tau}^{\mathtt{sun}})$ using (3.12)–(3.13), independently from all other random variables simulated so far.
    Compute $u_{k\tau}^{\Delta,(0)}$ with Steps 5 to 8 of Algorithm 1 with the data of the FBSDE (3.9). Compute $X_{k\tau}^{\Delta,(0)}$.
    $R[k] \leftarrow R[k] + \frac{1}{M_0}\left(\mathtt{P}_{k\tau}^{\mathtt{load},\Delta} - u_{k\tau}^{\Delta,(0)}\right)_+$
  end for
end for
return $(R[n], R[n+1], ..., R[N_T])$

$\mathtt{P}^{\mathtt{sun}}$	$\rho^{\mathtt{sun}}$	$0.75h^{-1}$
	$\mathtt{x}_t^{\mathtt{sun,ref}}$	0.5
	$\sigma^{\mathtt{sun}}$	0.8
	α	0.8
	β	0.7
	$\mathtt{P}^{\mathtt{sun,max}}$	see Figure 2a
$\mathtt{P}^{\mathtt{cons}}$	$\rho^{\mathtt{cons}}$	$0.9h^{-1}$
	$\mathtt{p}^{\mathtt{cons,ref}}$	see Figure 2b
	$h^{\mathtt{cons}}$	5 kW
	$\lambda^{\mathtt{cons}}$	0.5 h^{-1}

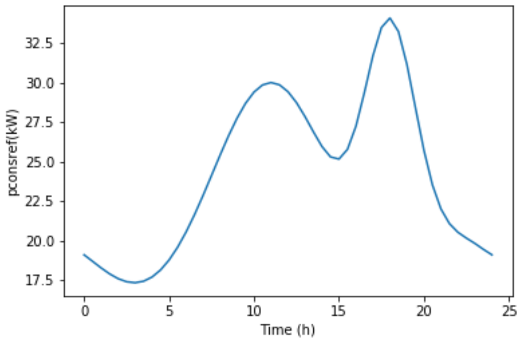
In Figure 2, we plot the time-evolution of the deterministic functions $\mathtt{P}^{\mathtt{sun,max}}$ and $\mathtt{p}^{\mathtt{cons,ref}}$, 10 independent samples of processes $\mathtt{P}^{\mathtt{sun}}$ and $\mathtt{P}^{\mathtt{cons}}$, and the time-evolution of quantiles (computed with $M_1 = 100000$ i.i.d. simulations).

Parameters of input data and optimization problem. The values of the parameters of the optimization problem are chosen such that:

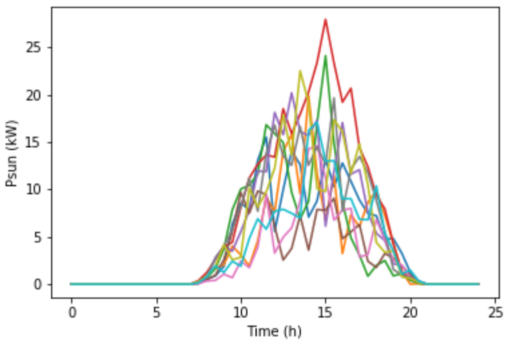
– the state of charge of the battery remains close to a reference level, which we set to 0.5,
– we observe a clear reduction of the random fluctuation of $\mathtt{P}^{\mathtt{grid}}$ on the time interval.



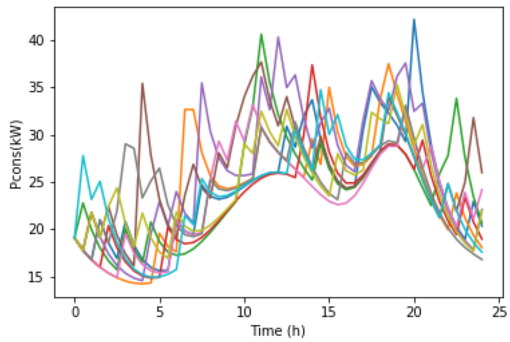
(A) Time evolution of $\mathtt{P}^{\mathtt{sun,max}}$, accounting for clear sky model



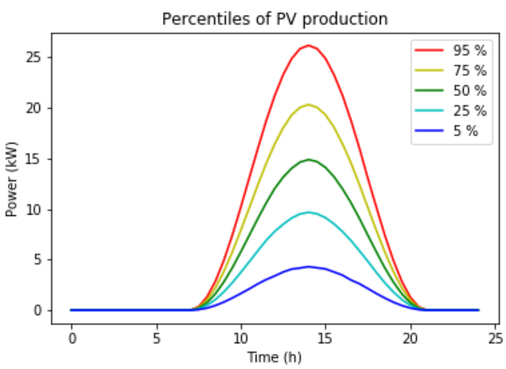
(B) Time evolution of $\mathtt{P}^{\mathtt{cons,ref}}$, accounting for intraday peaks



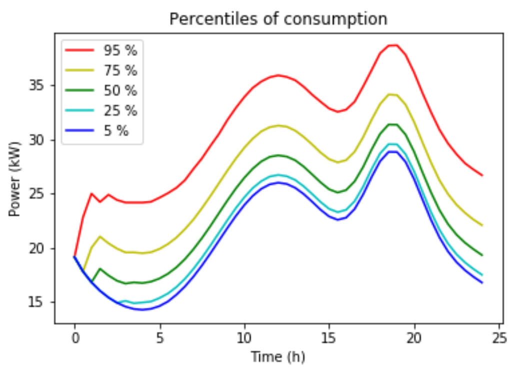
(C) Example trajectories of $\mathtt{P}^{\mathtt{sun}}$



(D) Example trajectories of $\mathtt{P}^{\mathtt{cons}}$



(E) Time evolution of quantiles of $\mathtt{P}^{\mathtt{sun}}$

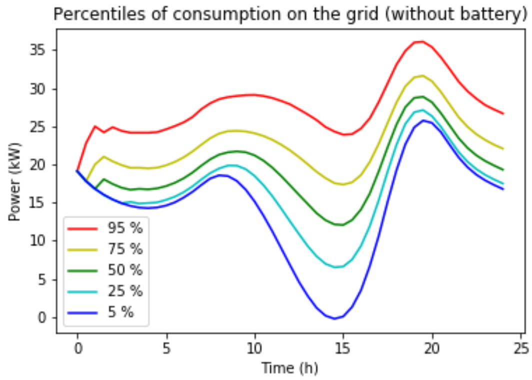


(F) Time evolution of quantiles of $\mathtt{P}^{\mathtt{cons}}$

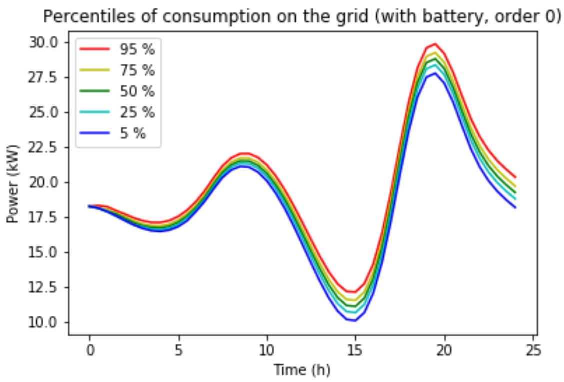
FIGURE 2. Graphical statistics of the evolution of $\mathtt{P}^{\mathtt{sun}}$ and $\mathtt{P}^{\mathtt{cons}}$.

The following table gives the values of the parameters of the cost functional of the control problem.

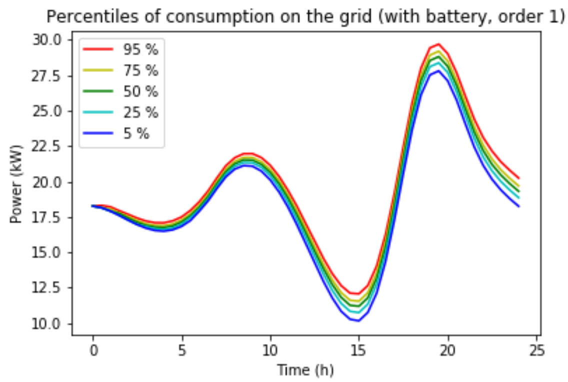
Parameter	ε	λ	μ	ν	γ	$\mathtt{C}$
Value	0.2	0.49	0.01	0.1	500	0.27
Unit	-	euros.kW^{-2}.h^{-1}	euros.kW^{-2}.h^{-1}	euros.h^{-1}	euros	euros.kW.h^{-1}



(A) Without flexibility



(B) With controlled flexibility ($\varepsilon = 0$)



(C) With controlled flexibility ($\varepsilon = 0.2$)

FIGURE 3. Quantiles of $\mathtt{P}^{\mathtt{grid}}$ as a function of time.

Time discretization. The average processes $(\bar{u}, \bar{X})$ are computed explicitly (up to numerical integration), while the recentered processes $(u^{\Delta,(0)}, X^{\Delta,(0)})$ and first order correction processes $(\dot{u}, \dot{X})$ are computed using discretization schemes (detailed in Algorithms 1 and 2) with time-step equal to $0.5h$.

Monte-Carlo simulations. To compute the first order correction, we need Monte-Carlo estimations, as explained in Algorithm 2. We choose $M_0 = 4000$. For assessing the statistical performances of the optimal control associated to a symmetric loss function ($\varepsilon = 0$), we consider $M_1 = 100000$ macro-runs. Among those M_1 trajectories, we only consider the first $M_2 = 4000$ trajectories for the computation of the first order corrections associated to $\varepsilon = 0.2$.

3.3.4. Results from the experiments

Computational time. The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. We have computed the optimal control associated to a symmetric penalization of deviations of $\mathtt{P}^{\mathtt{grid}}$ from its average ($\varepsilon = 0$) and for $M_1 = 100000$ i.i.d. simulations, which takes about 3 seconds. The computation of the first order correction when $\varepsilon = 0.2$ for $M_2 = 4000$ i.i.d. simulations takes about 80 minutes.

Reduction of fluctuations. We plot the time-evolution of quantiles (see Fig. 3) of the power supplied by the network in 3 cases: using no flexibility, with optimal control of the battery with $\varepsilon = 0$, and with the approximated

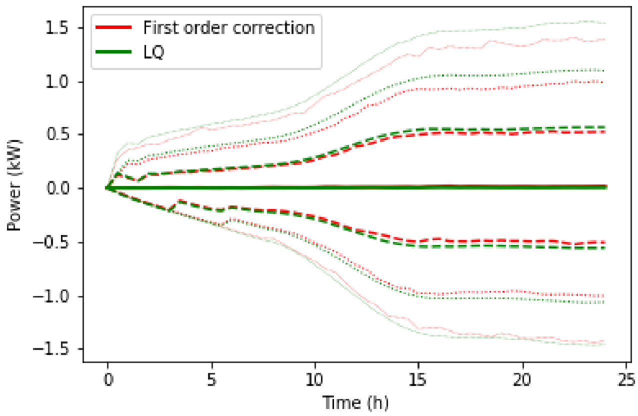
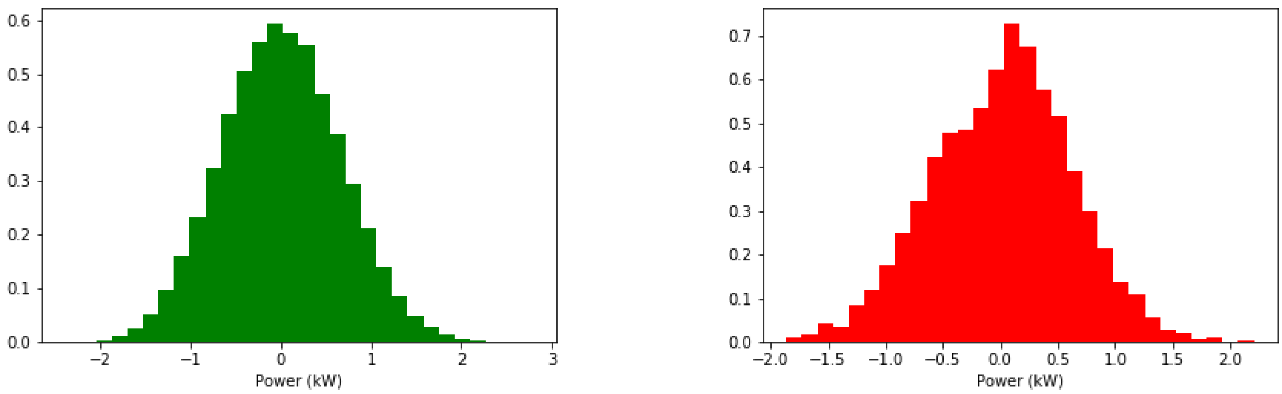


FIGURE 4. Time evolution of quantiles of deviations $\mathtt{P}^{\mathtt{grid}} - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}\right]$.



(A) Symmetric penalization ($\varepsilon = 0$)

(B) Asymmetric penalization ($\varepsilon = 0.2$)

FIGURE 5. Empirical histograms of deviations $\mathtt{P}^{\mathtt{grid}}(T) - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}(T)\right]$.

optimal control associated to $\varepsilon = 0.2$ respectively. The comparison of the first graph with the two others shows that the quantiles are much closer to each other in the case of storage use, meaning that the variability of the power supplied by the grid has been much reduced, as expected. However, the difference between the optimal control with symmetric and asymmetric loss functions is not much visible on these plots.

Impact of first order correction. Overall, the effect of the first order correction $\dot{u}$ (which has theoretically an average value of 0), is to lower the probability of large upper deviations of $\mathtt{P}^{\mathtt{grid}}$ from its expectation. This is quite visible if we plot the time-evolution of quantiles of the deviations $\mathtt{P}^{\mathtt{grid}}(t) - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}(t)\right]$ for the case $\varepsilon = 0$, in green in Figure 4, refered as "LQ" and $\varepsilon = 0.2$, in red, refered as "First Order Correction". In Figure 4, we have represented from top to bottom, the quantiles of $\mathtt{P}^{\mathtt{grid}}(t) - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}(t)\right]$ associated to levels 99%, 95%, 80%, 50%, 20%, 5% and 1%. We observe that the empirical estimations of the lower quantiles are left unchanged, while the upper quantiles 99% and 95% have been notably decreased, which was the effect sought by the choice of this loss function. To have a even more clear visualization of the change of distribution of the deviations $\mathtt{P}^{\mathtt{grid}} - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}\right]$, we have represented the empirical histograms of $\mathtt{P}^{\mathtt{grid}}(T) - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}(T)\right]$ for both cases $\varepsilon = 0$ in Figure 5a ($M_1 = 100000$ i.i.d. simulations) and $\varepsilon = 0.2$ in Figure 5b ($M_2 = 4000$ i.i.d. simulations). Observe that the impact of the first order term is to break the symmetry of the distribution around 0, and to reduce the

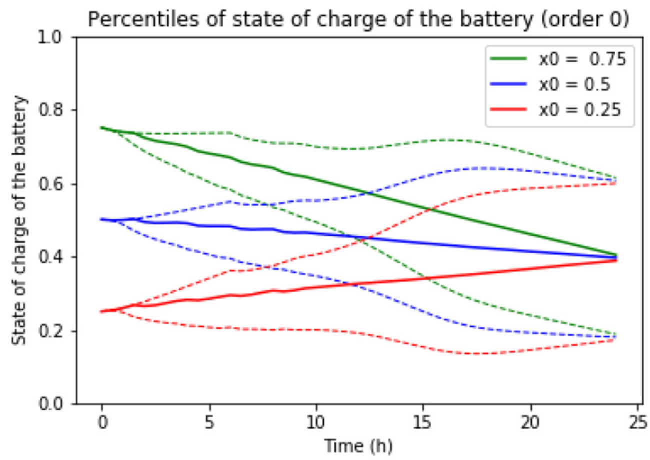


FIGURE 6. Time evolution of the empirical quantiles 95%, 50%, 5% of the state of charge of the storage system.

probability of the highest values of $\mathtt{P}^{\mathtt{grid}}(T) - \mathbb{E}\left[\mathtt{P}^{\mathtt{grid}}(T)\right]$. These results suggest that we have reached our goal of reducing the probability of high upper deviations of $\mathtt{P}^{\mathtt{grid}}$ from its average.

Distribution of state of charge of the battery. As the first order correction term has only minor impact on the distribution of the state of charge of the storage system, we only consider the case with $\varepsilon = 0$ in this paragraph. Figure 6 shows the time-evolution of the quantiles 95%, 50%, 5% of the state of charge of the battery with $\varepsilon = 0$ (computed using $M_1 = 100000$ i.i.d. simulations) for several initial conditions on the state of charge of the battery, namely $x_0 = 0.75$, $x_0 = 0.5$ and $x_0 = 0.25$. What we observe is that independently on the initial condition chosen, the state of charge remains between 0.15 and 0.75 with high probability. Besides, the terminal distribution of the state of charge is quite independent from the initial condition: the terminal values of the quantiles (levels 95%, 50%, 5%) of the state of charge are almost the same, for all initial conditions $x_0 = 0.75$, $x_0 = 0.5$ and $x_0 = 0.25$. This is presumably due to the term in the cost functional which penalizes the deviations of state of charge from a medium value (here 1/2).

Simulation-based bound on approximation error of first order expansion. Following the proof of Proposition 3.6, with our choice of parameters, we obtain an upper bound on the error in the approximation of the optimal control $u^{(\varepsilon)}$:

$$\begin{aligned}\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} &= \varepsilon\|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} \\ &\le \frac{4\varepsilon^2}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)}\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}.\end{aligned}$$

We would like to obtain a bound on the relative error committed $\frac{\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\|u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}}$. To do this, we have by triangular inequality:

$$\begin{aligned}\frac{\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\|u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}} &\le \frac{\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}}{\left|\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - \|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}\right|} \\ &\le \frac{4\varepsilon^2}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)} \frac{\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{\left|\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - \|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}\right|} \\ &\le \frac{4\varepsilon^2\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}} - 4\varepsilon^2\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}.\end{aligned}$$

In the last inequality, we used the fact that $\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$ is asymptotically small compared to $\|\bar{u} + u^{\Delta,(0)} + \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$ when ε goes to 0, as well as the previous bound on $\|u^{(\varepsilon)} - \bar{u} - u^{\Delta,(0)} - \varepsilon\dot{u}\|_{\mathbb{H}^{2,2}}$. Hence we obtain an upper bound which depends only on quantities which can be estimated by simulations in the algorithm. This is very convenient to assess the relative accuracy of our approximation. The left-hand-side in the last inequality is estimated using the $M_2 = 4000$ simulations of the first order expansion and we find a value of 0.03. In other words, the relative error is smaller than 3% when taking the first order expansion of the control instead of its true value. Note that we do not take into account errors due to the time discretization or due to residual noise in the Monte-Carlo estimations.

4. Proofs

4.1. Proof of Theorem 2.2

(a) Observe first that, in view of **(H.ϕ)** and Lemma 4.1, $\tilde{L}^u$ and $(\tilde{L}^u)^{-1}$ are uniformly bounded by a constant $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e (take $A : (t,\omega) \mapsto \nabla_x \phi(t,\omega,u_t,X_t^u)$). Therefore, and owing to **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**, the random variable inside the conditional expectation defining Y^u in (2.4) is bounded by

$$\begin{aligned}\Gamma_T :=& C_T \left(C_\psi^{(0)} + |X_T^u| + \mathbb{E}\left[|k(0)|\right] + \mathbb{E}\left[|X_T^u|\right]\right) \\ &+ C_T \int_0^T \left(C_l^{(0)}(s) + |X_s^u| + |u_s| + \mathbb{E}\left[|g(s,0,0)|\right] + \mathbb{E}\left[|u_s|\right] + \mathbb{E}\left[|X_s^u|\right] + \mathbb{E}\left[C_l^{(0)}(s)\right]\right) \mathrm{d}s\end{aligned}$$

for some constant C_T depending on the bounds in **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**. Hence by the Cauchy Schwartz inequality, for some other constant C_T:

$$\begin{aligned}\mathbb{E}\left[|\Gamma_T|^2\right] \leq& C_T \left(\mathbb{E}\left[\left(C_\psi^{(0)}\right)^2\right] + \mathbb{E}\left[|X_T^u|^2\right] + \mathbb{E}\left[|k(0)|\right]^2\right) \\ &+ C_T \int_0^T \left(\mathbb{E}\left[\left(C_l^{(0)}(s)\right)^2\right] + \mathbb{E}\left[|g(s,0,0)|\right]^2 + \mathbb{E}\left[|u_s|^2\right] + \mathbb{E}\left[|X_s^u|^2\right]\right) \mathrm{d}s.\end{aligned}$$

Note that this bound is finite and independent from t (since $C_\psi^{(0)} \in \mathbb{L}^2_\Omega$, $X^u \in \mathbb{H}^{\infty,2} \subset \mathbb{H}^{2,2}$, $k(0) \in \mathbb{L}^1$, $C_l^{(0)} \in \mathbb{H}^{2,2}$, $g(\cdot,0,0) \in \mathbb{H}^{2,1}$ and $u \in \mathbb{H}^{2,2}$). Consequently $Y^u \in \mathbb{H}^{\infty,2}$.

(b) Now observe that, by definition of Y^u,

$$N_t^u := \tilde{L}_t^u Y_t^u + \int_0^t \tilde{L}_s^u \Big(\nabla_x l(s,u_s,X_s^u,\bar{g}_s^u) + \nabla_x g(s,u_s,X_s^u)\mathbb{E}\left[\nabla_{\bar{g}} l(s,u_s,X_s^u,\bar{g}_s^u)\right]\Big) \mathrm{d}s = \mathbb{E}_t\left[N_T^u\right], \tag{4.1}$$

with N_T^u square integrable (using the same arguments as before) and therefore, N^u is a càdlàg martingale in $\mathbb{H}^{\infty,2}$. As a by-product, we obtain that Y^u is a semi-martingale, which dynamics has now to be identified.

(c) To justify that Y^u defined in (2.4) solves the BSDE (2.5) for some M^u, left-multiply both sides of (4.1) by $(\tilde{L}_t^u)^{-1}$, then apply the integration by parts formula in ([25], Cor. 2, p. 68) to $(\tilde{L}^u)^{-1}N^u$ and use the fact that $(\tilde{L}^u)^{-1}$ is continuous with finite variations. After reorganizing terms and using that N^u has countable jumps, we retrieve (2.5) with $M_t^u := \int_{0^+}^t (\tilde{L}_s^u)^{-1} \mathrm{d}N_s^u$ (which is also a càdlàg martingale in $\mathbb{H}^{\infty,2}$, see [25], Thm. 20. p. 63, Cor. 3 p. 73, Thm. 29, p. 75).

(d) We now claim that the solution (Y, M) to (2.5) is unique in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$, and thus given by (Y^u, M^u). In fact, the uniqueness is a classical result for linear BSDE, see for instance ([14], Thm. 5.1 with $p = 2$) in our context of general filtration.

(e) Let us now prove the last claim about $\dot{\mathcal{J}}(u,v)$. We first study the differentiability properties of $X^{u+\varepsilon v}$ with respect to ε. In the following computations we use different constants which we denote generically by C

(although their values may change from line to line), they do not depend on u, v, ε, they only depend on $T > 0$ and on the bounds from the assumptions **(H.x)**-**(H.g)**-**(H.l)**-**(H.k)**-**(H.ψ)**-**(H.ϕ)**. At this point of the proof, it is more convenient to work with Jacobian matrices than with gradients (as in the statement). Only at the very end are we going to make the link with Y^u and go back to the gradient notation.

Set $\theta_t^u := (t, u_t, X_t^u)$ and let $\dot{X}_t^{u,v}$ be the solution to the following linear equation

$$\dot{X}_t^{u,v} := \int_0^t \left[\phi_u(\theta_s^u)v_s + \phi_x(\theta_s^u)\dot{X}_s^{u,v}\right]\mathrm{d}s = \int_0^t (L_t^u)^{-1}L_s^u\phi_u(\theta_s^u)v_s\mathrm{d}s, \tag{4.2}$$

since it can be noticed that L^u is the unique solution of

$$L_0^u = \mathrm{Id}_p, \qquad \frac{\mathrm{d}L_t^u}{\mathrm{d}t} = -L_t^u\phi_x(\theta_t^u),$$

using Lemma 4.1 and $\nabla_x\phi = (\phi_x)^\top$. Note, whenever useful, that $\int_0^T |v_s|^2\mathrm{d}s < +\infty$ a.s. since $v \in \mathbb{H}^{2,2}$.

For $\varepsilon \neq 0$, set $\Delta X_t^{u,v,\varepsilon} := \frac{X_t^{u+\varepsilon v} - X_t^u}{\varepsilon}$ and $RX_t^{u,v,\varepsilon} := \Delta X_t^{u,v,\varepsilon} - \dot{X}_t^{u,v}$: we claim that a.s. $RX_t^{u,v,\varepsilon} \to 0$ as $\varepsilon \to 0$, *i.e.* $\dot{X}_t^{u,v}$ is the derivative of $X_t^{u+\varepsilon v}$ at $\varepsilon = 0$. To justify this, we proceed in a few steps. First the Taylor formula equality gives, for smooth φ,

$$\varphi(z, x') - \varphi(z, x) = \Big(\int_0^1 \varphi_x(z, x + \lambda(x' - x))\mathrm{d}\lambda\Big)(x' - x) := \overline{\varphi}_x(z, [x, x'])(x' - x);$$

applying that to ϕ_x and ϕ_u, we obtain

$$\begin{aligned} \Delta X_t^{u,v,\varepsilon} &= \int_0^t \Big[\overline{\phi}_x(s, u_s + \varepsilon v_s, [X_s^u, X_s^{u+\varepsilon v}])\Delta X_s^{u,v,\varepsilon} + \overline{\phi}_u(s, [u_s, u_s + \varepsilon v_s], X_s^u)v_s\Big]\mathrm{d}s \\ &= (\overline{L_t^{u,v,\varepsilon}})^{-1}\int_0^t \overline{L_s^{u,v,\varepsilon}}\ \overline{\phi}_u(s, [u_s, u_s + \varepsilon v_s], X_s^u)v_s\mathrm{d}s \end{aligned} \tag{4.3}$$

where $\overline{L^{u,v,\varepsilon}}$ is the unique solution of

$$\overline{L_0^{u,v,\varepsilon}} = \mathrm{Id}_p, \qquad \frac{\mathrm{d}\overline{L_t^{u,v,\varepsilon}}}{\mathrm{d}t} = -\overline{L_t^{u,v,\varepsilon}}\ \left(\overline{\phi}_x(t, u_t + \varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}])\right).$$

By hypothesis on ϕ_x, L^u and $(L^u)^{-1}$ are $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. uniformly bounded by a constant. Besides, $(t, \omega) \mapsto \overline{\phi}_x(t, u_t + \varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}])$ satisfies the hypothesis of Lemma 4.1 with a constant C independent from ε. Therefore $\overline{L^{u,v,\varepsilon}}$ is invertible and $L^{u,v,\varepsilon}$ and $(L^{u,v,\varepsilon})^{-1}$ are $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. uniformly bounded by a constant independent from ε. From (4.2)-(4.3) and using the fact that ϕ_u and $\overline{\phi}_u(\cdot, [u, u + \varepsilon v], X^u)$ are $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. bounded by a constant independent from ε (by hypothesis on ϕ_u), we derive

$$|\Delta X_t^{u,v,\varepsilon}| + |\dot{X}_t^{u,v}| \le C\int_0^t |v_s|\mathrm{d}s, \tag{4.4}$$

where C is a constant independent from ε. With similar arguments, we can represent the residual error $RX^{u,v,\varepsilon}$ as follows:

$$RX_t^{u,v,\varepsilon} = \int_0^t \Big(\eta_s^{u,v,\varepsilon} + \phi_x(s, u_s, X_s^u)RX_s^{u,v,\varepsilon}\Big)\mathrm{d}s = \int_0^t (L_t^u)^{-1}L_s^u\eta_s^{u,v,\varepsilon}\mathrm{d}s,$$

with L^u as before and

$$\begin{aligned}\eta_t^{u,v,\varepsilon} &:= \alpha_t^{u,v,\varepsilon}\Delta X_t^{u,v,\varepsilon} + \beta_t^{u,v,\varepsilon} v_t,\\ \alpha_t^{u,v,\varepsilon} &:= \overline{\phi}_x(t, u_t+\varepsilon v_t, [X_t^u, X_t^{u+\varepsilon v}]) - \phi_x(t, u_t, X_t^u),\\ \beta_t^{u,v,\varepsilon} &:= \overline{\phi}_u(t, [u_t, u_t+\varepsilon v_t], X_t^u) - \phi_u(t, u_t, X_t^u).\end{aligned}$$

By boundedness of L^u and $(L^u)^{-1}$, and by (4.4), we have:

$$\sup_{t\in[0,T]} |RX_t^{u,v,\varepsilon}| \le C\left(\int_0^T |\alpha_t^{u,v,\varepsilon}|\mathrm{d}t\right)\left(\int_0^T |v_t|\mathrm{d}t\right) + C\int_0^T |\beta_t^{u,v,\varepsilon}||v_t|\mathrm{d}t,$$

for some constant $C > 0$.

By continuity of the state with respect to the control and continuous differentiability of ϕ with respect to (u, x), $\alpha^{u,v,\varepsilon}$ and $\beta^{u,v,\varepsilon}$ converge point-wise to 0 when ε goes to 0 and are uniformly bounded by assumption on ϕ. Therefore $\left(\int_0^T |\alpha_t^{u,v,\varepsilon}|\mathrm{d}t\right)$ converges to 0 as ε goes to 0, by Lebesgue's domination theorem. Besides, $|\beta^{u,v,\varepsilon}||v|$ converges point-wise to 0 since $\{s \in [0, T] : |v_s| < +\infty\}$ has a full Lebesgue measure (recall that a.s. $\int_0^T |v_s|^2\mathrm{d}s < +\infty$) and is dominated by $|v|$, which is a.s. integrable. Applying again Lebesgue's domination theorem yields the convergence of $\int_0^T |\beta_t^{u,v,\varepsilon}||v_t|\mathrm{d}t$ to 0 when ε goes to 0.

Putting everything together, one gets that a.s. $\sup_{t\in[0,T]} |RX_t^{u,v,\varepsilon}| \to 0$ as $\varepsilon \to 0$, *i.e.* $\partial_\varepsilon X_t^{u+\varepsilon v}\big|_{\varepsilon=0} = \dot{X}_t^{u,v}$ a.s.

(f) We now switch to the differentiability of $\mathcal{J}(u + \varepsilon v)$ w.r.t. ε at $\varepsilon = 0$. Similar arguments as before yield

$$|\partial_\varepsilon X_t^{u+\varepsilon v}| \le C\int_0^t |v_s|\mathrm{d}s, \tag{4.5}$$

and for $\varepsilon \in [-1, 1]$,

$$|X_t^{u+\varepsilon v}| \le |X_t^u| + |\varepsilon\Delta X_t^{u,v,\varepsilon}| \le |X_t^u| + |\Delta X_t^{u,v,\varepsilon}| \le |X_t^u| + C\int_0^t |v_s|\mathrm{d}s \tag{4.6}$$

$\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e. for a constant C independent from $\varepsilon \in [-1, 1]$.

The above, combined with $X^u \in \mathbb{H}^{\infty,2}$, $u \in \mathbb{H}^{2,2}$, the smoothness of the functions l, g, ψ, k with the bounds on their derivatives (assumptions **(H.l)**-**(H.g)**-**(H.ψ)**-**(H.k)**) allow to apply the Lebesgue differentiation theorem and to obtain

$$\begin{aligned}\dot{\mathcal{J}}(u, v) = \mathbb{E}\Bigg[\int_0^T \Bigg(& l_u(\theta_t^u, \bar{g}_t^u)v_t + l_x(\theta_t^u, \bar{g}_t^u)\dot{X}_t^{u,v} + l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\mathbb{E}\left[g_u(\theta_t^u)v_t + g_x(\theta_t^u)\dot{X}_t^{u,v}\right]\Bigg)\mathrm{d}t\\ &+ \psi_x\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\dot{X}_T^{u,v} + \psi_{\bar{k}}\left(X_T^u, \mathbb{E}\left[k(X_T^u)\right]\right)\mathbb{E}\left[k_x(X_T^u)\dot{X}_T^{u,v}\right]\Bigg].\end{aligned}$$

Using Fubini's theorem and reorganizing terms, we get

$$\dot{\mathcal{J}}(u, v) = \mathbb{E}\Bigg[\int_0^T \Bigg(\Bigg\{l_u(\theta_t^u, \bar{g}_t^u) + \mathbb{E}\left[l_{\bar{g}}(\theta_t^u, \bar{g}_t^u)\right]g_u(\theta_t^u)\Bigg\}v_t$$

$$+\left\{l_x(\theta_t^u,\bar g_t^u)+\mathbb{E}\left[l_{\bar g}(\theta_t^u,\bar g_t^u)\right]g_x(\theta_t^u)\right\}\dot X_t^{u,v}\Bigg)\mathrm{d}t$$

$$+\left\{\psi_x\left(X_T^u,\mathbb{E}\left[k(X_T^u)\right]\right)+\mathbb{E}\left[\psi_{\bar k}\left(X_T^u,\mathbb{E}\left[k(X_T^u)\right]\right)\right]k_x(X_T^u)\right\}\dot X_T^{u,v}\Bigg].$$

Apply now the Itô lemma to $Y_t^u\cdot\dot X_t^{u,v}$ between $t=0$ and $t=T$, with

$$Y_T^u\cdot\dot X_T^{u,v}=\left\{\psi_x\left(X_T^u,\mathbb{E}\left[k(X_T^u)\right]\right)+\mathbb{E}\left[\psi_{\bar k}\left(X_T^u,\mathbb{E}\left[k(X_T^u)\right]\right)\right]k_x(X_T^u)\right\}\dot X_T^{u,v},\qquad Y_0^u\cdot\dot X_0^{u,v}=0,$$

note that $\dot X^{u,v}$ has finite variations, combine with (2.5) and (4.2), and take the expectation: it gives

$$\dot{\mathcal{J}}(u,v)=\mathbb{E}\left[\int_0^T\left\{l_u(\theta_t^u,\bar g_t^u)+\mathbb{E}\left[l_{\bar g}(\theta_t^u,\bar g_t^u)\right]g_u(\theta_t^u)+(Y_t^u)^\top\phi_u(\theta_t^u)\right\}v_t\mathrm{d}t\right].$$

The formula is also valid for Y_{t-}^u since the jumps of Y^u are countable. Theorem 2.2 is proved. □

4.2. Proof of Theorem 2.4

In the proof, $T\le 1$ and C denotes a generic (deterministic) constant which only depends on the bounds in the hypothesis (and not on T). For $u^{(1)}$ and $u^{(2)}$ in $\mathbb{H}^{2,2}$, if a process or variable F^u depends on u we write $F^{(1)}:=F^{u^{(1)}}$ and $F^{(2)}:=F^{u^{(2)}}$. Besides, for any function, operator or process F which depends on u, X^u,…, we write $\Delta F:=F^{(2)}-F^{(1)}$.

The proof is decomposed into several steps.

- First, notice that by our assumptions on ϕ, L^u (resp. $\tilde L^u=((L^u)^{-1})^\top$) (defined in Thm. 2.2) is independent from u, therefore, we simply write L (resp. $\tilde L$) instead. Using Lemma 4.1, L and $\tilde L$ are bounded by constants.
- Consider the application

$$\Theta^{(X)}:\begin{cases}\mathbb{H}^{2,2}_{\mathcal{P}} & \to\mathbb{H}^{\infty,2}\\ u & \mapsto X^u\end{cases}.$$

  It is well-defined, since we have already seen that $X^u\in\mathbb{H}^{\infty,2}$ whenever $u\in\mathbb{H}^{2,2}_{\mathcal{P}}$. We want to show that $\Theta^{(X)}$ is Lipschitz continuous and its Lipschitz constant is such that

$$C_{\Theta^{(X)},u}(T)=\mathcal{O}\left(\sqrt{T}\right)\quad(T\to 0).$$

  Using assumption **(H.ϕ.2)** and computations as in (4.2):

$$\Delta X_t=\int_0^t\Delta\phi_s\mathrm{d}s=\int_0^t\left(a_s^{(\phi)}\Delta X_s+b^{(\phi)}(s,u_s^{(2)})-b^{(\phi)}(s,u_s^{(1)})\right)\mathrm{d}s$$
$$=\int_0^t(L_t)^{-1}L_s(b^{(\phi)}(s,u_s^{(2)})-b^{(\phi)}(s,u_s^{(1)}))\mathrm{d}s.$$

Therefore by assumption on ϕ, we get

$$|\Delta X_t|^2 \le C\left(\int_0^T |\Delta u_s|\mathrm{d}s\right)^2 \le CT\int_0^T |\Delta u_s|^2\mathrm{d}s,$$

whence, $\|\Delta X\|^2_{\mathbb{H}^{\infty,2}} \le CT\|\Delta u\|^2_{\mathbb{H}^{2,2}}$ and the Lipschitz continuity of $\Theta^{(X)}$ as announced.

– Now consider

$$\Theta^{(Y)} : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} & \to \mathbb{H}^{\infty,2} \\ u & \mapsto Y^u \end{cases},$$

with Y^u as in (2.4). Theorem 2.2 guarantees that $\Theta^{(Y)}$ is well defined. Let us prove that it is Lipschitz continuous and its Lipschitz constant is such that

$$C_{\Theta^{(Y)},u}(T) = \mathcal{O}\left(\sqrt{T}\right) \quad (T \to 0).$$

Using the hypothesis on ϕ, g and k and the notation $\theta^u_s = (s, u_s, X^u_s)$, we get $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t - a.e.$

$$\begin{aligned}\Delta Y_t =& \mathbb{E}_t\left[\tilde{L}_t^{-1}\tilde{L}_T\left(\nabla_x\psi(X_T^{(2)}, \mathbb{E}\left[k(X_T^{(2)})\right]) - \nabla_x\psi(X_T^{(1)}, \mathbb{E}\left[k(X_T^{(1)})\right])\right)\right] \\ &+ \mathbb{E}_t\left[\tilde{L}_t^{-1}\tilde{L}_T(a^{(k)})^\top\mathbb{E}\left[\nabla_{\bar{k}}\psi(X_T^{(2)}, \mathbb{E}\left[k(X_T^{(2)})\right]) - \nabla_{\bar{k}}\psi(X_T^{(1)}, \mathbb{E}\left[k(X_T^{(1)})\right])\right]\right] \\ &+ \mathbb{E}_t\left[\int_t^T \tilde{L}_t^{-1}\tilde{L}_s\left(\nabla_x l(\theta_s^{(2)}, \mathbb{E}\left[g(\theta_s^{(2)})\right]) - \nabla_x l(\theta_s^{(1)}, \mathbb{E}\left[g(\theta_s^{(1)})\right])\right)\mathrm{d}s\right] \\ &+ \mathbb{E}_t\left[\int_t^T \tilde{L}_t^{-1}\tilde{L}_s(a_s^{(g)})^\top\mathbb{E}\left[\nabla_{\bar{g}} l(\theta_s^{(2)}, \mathbb{E}\left[g(\theta_s^{(2)})\right]) - \nabla_{\bar{g}} l(\theta_s^{(1)}, \mathbb{E}\left[g(\theta_s^{(1)})\right])\right]\mathrm{d}s\right].\end{aligned}$$

Now, owing to assumptions, the Cauchy-Schwartz inequality, the previous estimate on $\tilde{L}$, on its inverse and $\|\Delta X\|_{\mathbb{H}^{\infty,2}}$, the inequality $\|\cdot\|_{\mathbb{H}^{2,2}} \le \sqrt{T}\|\cdot\|_{\mathbb{H}^{\infty,2}}$, one gets:

$$\|\Delta Y\|^2_{\mathbb{H}^{\infty,2}} \le C\left(\mathbb{E}\left[|\Delta X_T|^2\right] + T\left\{\|\Delta u\|^2_{\mathbb{H}^{2,2}} + \|\Delta X\|^2_{\mathbb{H}^{2,2}}\right\}\right) \le CT\|\Delta u\|^2_{\mathbb{H}^{2,2}}.$$

This yields the Lipschitz continuity of $\Theta^{(Y)}$ with $C_{\Theta^{(Y)},u}(T) = \mathcal{O}\left(\sqrt{T}\right) \quad (T \to 0)$.

– Our goal is to prove that:

$$\Theta : \begin{cases} \mathbb{H}^{2,2}_{\mathcal{P}} & \to \mathbb{H}^{2,2}_{\mathcal{P}} \\ u & \mapsto \tilde{u} \end{cases},$$

with

$$\begin{aligned}&\tilde{u}_t = h\left(t, X_t^u, Y_{t-}^u, \bar{g}_t^u, \mathbb{E}\left[\nabla_{\bar{g}} l(t, u_t, X_t^u, \bar{g}_t^u)\right]\right), \\ &X^u = \Theta^{(X)}(u), \qquad Y^u = \Theta^{(Y)}(u), \qquad \bar{g}_t^u = \mathbb{E}\left[g(t, u_t, X_t^u)\right]\end{aligned}$$

is well defined, Lipschitz continuous and its Lipschitz constant satisfies:

$$C_{\Theta,u}(T) = C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right) + \mathcal{O}(T) \quad (T \to 0).$$

By construction, $\tilde{u}$ is predictable. Besides, for $u \in \mathbb{H}^{2,2}_{\mathcal{P}}$,

$$\mathbb{E}\left[\int_0^T |\tilde{u}_s|^2 \mathrm{d}s\right] \leq \mathbb{E}\Bigg[\int_0^T \Bigg(|h(s,0,0,0,0)| + (C_{h,\bar{g}} + C_{h,\bar{\lambda}} C_{l_{\bar{g}},\bar{g}})\mathbb{E}\left[|g(s,0,0)|\right] + C_{h,x}|X^u_s| + C_{h,y}|Y^u_s| + \left(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\right)\mathbb{E}\left[|u_s|\right] + \left(C_{h,\bar{g}}C_{g,x} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},x} + C_{l_{\bar{g}},\bar{g}}C_{g,x}\right)\right)\mathbb{E}\left[|X^u_s|\right]\Bigg)^2 \mathrm{d}s\Bigg].$$

Using Minkowski's inequality, this shows that the right-hand side is finite since $g(\cdot,0,0) \in \mathbb{H}^{2,1}$, $h(\cdot,0,0,0,0), X^u, Y^u$ and u are in $\mathbb{H}^{2,2}$, whence the well-posedness of Θ. Similar computations give

$$\|\Delta\tilde{u}\|_{\mathbb{H}^{2,2}} \leq \Bigg(\mathbb{E}\Bigg[\int_0^T \Bigg(C_{h,x}|\Delta X_s| + \left(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\right)\mathbb{E}\left[|\Delta u_s|\right] + C_{h,y}|\Delta Y_s| + \left(C_{h,\bar{g}}C_{g,x} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},x} + C_{l_{\bar{g}},\bar{g}}C_{g,x}\right)\right)\mathbb{E}\left[|\Delta X_s|\right]\Bigg)^2 \mathrm{d}s\Bigg]\Bigg)^{1/2}.$$

Again, from Minkowski's inequality it follows that

$$\|\Delta\tilde{u}\|_{\mathbb{H}^{2,2}} \leq \Bigg(C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right)\Bigg)\|\Delta u\|_{\mathbb{H}^{2,2}} + C\Bigg(\|\Delta X\|_{\mathbb{H}^{2,2}} + \|\Delta Y\|_{\mathbb{H}^{2,2}}\Bigg).$$

Using $\|\cdot\|_{\mathbb{H}^{2,2}} \leq \sqrt{T}\|\cdot\|_{\mathbb{H}^{\infty,2}}$ and our estimates on $\|\Delta X\|_{\mathbb{H}^{\infty,2}}$ and $\|\Delta Y\|_{\mathbb{H}^{\infty,2}}$, we obtain that Θ is Lipschitz continuous and its Lipschitz constant satisfies:

$$C_{\Theta,u}(T) \leq C_{h,\bar{g}}C_{g,u} + C_{h,\bar{\lambda}}\left(C_{l_{\bar{g}},u} + C_{l_{\bar{g}},\bar{g}}C_{g,u}\right) + \mathcal{O}(T) \quad (T \to 0).$$

- Under assumption (2.8), for T small enough, Θ is a contraction in the complete space $\mathbb{H}^{2,2}_{\mathcal{P}}$ and has therefore a unique fixed point u in $\mathbb{H}^{2,2}_{\mathcal{P}}$.
- To conclude, notice (2.1)–(2.4)–(2.6) are satisfied by (u, X^u, Y^u) with $X^u = \Theta^{(X)}(u)$ and $Y^u = \Theta^{(Y)}(u)$ if and only if u is a fixed point of Θ. □

4.3. Proof of Proposition 2.6

▷ The continuity and coercivity of F are obvious. Similar computations as in the proof of Theorem 2.2 show that F is Gateaux-differentiable and that the Gateaux-derivative of F at u in direction v is given by:

$$\dot{F}(u,v) = 4(\|u\|^2_{\mathbb{L}^2_1} - 1)\int_0^1 u_t v_t \mathrm{d}t + 2\int_0^1 t u_t v_t \mathrm{d}t := \int_0^1 \mathcal{F}(u)_t v_t \mathrm{d}t,$$

where $\mathcal{F} : \mathbb{L}^2_1 \mapsto \mathbb{L}^2_1$ is defined by $\mathcal{F}(u) : t \mapsto \big(4(\|u\|^2_{\mathbb{L}^2_1} - 1) + 2t\big)u_t$.

▷ Let us identify the critical points $u^\star \in \mathbb{L}^2_1$: for such element, we must have $\dot{F}(u^\star, \mathcal{L}(u^\star)) = \int_0^1 |\mathcal{F}(u^\star)_t|^2 \mathrm{d}t = 0$, which implies $(4(\|u^\star\|^2_{\mathbb{L}^2_1} - 1) + 2t)u^\star_t = 0$ a.e. on $[0,1]$. Clearly, it leads to $u^\star_t = 0$ a.e. on $[0,1]$ and therefore, 0 is the unique critical point of F.

▷ Let us show that $\inf_{u \in \mathbb{L}^2_1} F(u) = 0$. Since F takes values in $\mathbb{R}^+$, it is enough to exhibit a sequence $u^{(n)} \in \mathbb{L}^2_1$

s.t. $F(u^{(n)}) \to 0$ as $n \to +\infty$. Define, $\forall n \in \mathbb{N}$,

$$u^{(n)} : t \mapsto \sqrt{n+1}\mathbb{I}_{[0,\frac{1}{n+1}]}(t).$$

Then,

$$\int_0^1 |u_t^{(n)}|^2 \mathrm{d}t = 1, \qquad \int_0^1 t|u_t^{(n)}|^2 \mathrm{d}t = \int_0^{1/(n+1)} (n+1)t\mathrm{d}t = \frac{1}{2(n+1)},$$

therefore $F(u^{(n)}) = \frac{1}{2(n+1)} \to 0$, as it was sought.
$\triangleright$ Last, we prove that the minimum is not achieved. Assume the contrary with the existence of $u^\star \in \mathbb{L}_1^2$ s.t. $F(u^\star) = 0$. We must have $\|u^\star\|_{\mathbb{L}_1^2} = 1$ and $t|u_t^\star|^2 = 0$ a.e. on $[0,1]$: the second condition requires $u^\star = 0$ which is incompatible with the first condition. We are done, F does not have a minimum. ☐

4.4. Proof of Proposition 3.2

Usual results about the solution to affine-linear FBSDEs hold for Brownian filtration, see [30] for instance. Here, we consider more general filtrations, but the arguments are quite similar. For the sake of completeness, we give the proof.

The function θ is the unique solution of the following affine-linear second order ODE

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} - bag\theta_t = 0 \text{ for } t \in [0,T], \\ \theta_T = 1, \\ \frac{\mathrm{d}\theta_t}{\mathrm{d}t}|_{t=T} = -eag, \end{cases}$$

and does not vanish on $[0,T]$ according to the sign conditions on the coefficients. We can therefore define p as in (3.3). Besides, p and θ are continuous and bounded on $[0,T]$. By standard arguments, one can check that p is the unique solution of the following Riccati ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} - agp_t^2 + b = 0, \\ p_T = e. \end{cases} \tag{4.7}$$

Define the following BSDE:

$$\begin{cases} -\mathrm{d}\tilde{\pi}_t = -(agp_t\tilde{\pi}_t + ap_th_t - c)\mathrm{d}t - \mathrm{d}M_t, \\ \tilde{\pi}_T = f, \end{cases}$$

which has a unique solution in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2}$ (see [14], Thm. 5.1 with $p = 2$) in our context of general filtrations. By the integration by parts formula applied to $\tilde{\pi}_t \exp\left(\int_t^T agp_s\mathrm{d}s\right)$ (see [25], Cor. 2, p. 68), we get that $\tilde{\pi}$ is also given by:

$$\begin{aligned} \tilde{\pi}_t &= \mathbb{E}_t\left[f\exp\left(-\int_t^T agp_\tau\mathrm{d}\tau\right) - \int_t^T (ap_sh_s - c)\exp\left(-\int_t^s agp_\tau\mathrm{d}\tau\right)\mathrm{d}s\right] \\ &= \mathbb{E}_t\left[f\exp\left(\int_t^T \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right) - \int_t^T (ap_sh_s - c)\exp\left(\int_t^s \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right)\mathrm{d}s\right] \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}_t\left[f\frac{\theta_T}{\theta_t} - - \int_t^T (ap_s h_s - c)\frac{\theta_s}{\theta_t}\mathrm{d}s\right] \\
&= \frac{1}{\theta_t}\left(f - \int_t^T (ap_s\mathbb{E}_t\left[h_s\right] - c)\theta_s\mathrm{d}s\right) = \pi_t,
\end{aligned}
$$

where we used the definitions of p and π.

We deduce that the process π also has the following representation:

$$
\pi_t = \mathbb{E}_t\left[f - \int_t^T (agp_s\pi_s + ap_s h_s - c)\mathrm{d}s\right]. \tag{4.8}
$$

Since θ and p are bounded on $[0,T]$, we easily prove that $\pi \in \mathbb{H}^{\infty,2}$. From that and our assumptions on the data of the problem, it is clear that (x,y,v) as defined in (3.5) belong to $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{2,2}_{\mathcal{P}}$. In particular, v is predictable since x is continuous by construction.

We now prove that (x,y,v) defined by (3.5) solves (3.6). By definition of y and v in (3.5), we can check that:

$$
v_t = gy_{t-} + h_t.
$$

Define $\tilde{x}$ the unique solution of the following affine-linear ODE:

$$
\tilde{x}_t := x_0 - \int_0^t (agp_s\tilde{x}_s + ag\pi_s + ah_s)\mathrm{d}s. \tag{4.9}
$$

It is also given by:

$$
\begin{aligned}
\tilde{x}_t &= x_0\exp\left(-\int_0^t agp_\tau\mathrm{d}\tau\right) - \int_0^t\left\{(ag\pi_s + ah_s)\exp\left(-\int_s^t agp_\tau\mathrm{d}\tau\right)\right\}\mathrm{d}s \\
&= x_0\exp\left(\int_0^t \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right) - \int_0^t\left\{(ag\pi_s + ah_s)\exp\left(\int_s^t \frac{\mathrm{d}\theta_\tau}{\mathrm{d}t}\frac{1}{\theta_\tau}\mathrm{d}\tau\right)\right\}\mathrm{d}s \\
&= x_0\frac{\theta_t}{\theta_0} - - \int_0^t (ag\pi_s + ah_s)\frac{\theta_t}{\theta_s}\mathrm{d}s \\
&= x_t;
\end{aligned}
$$

hence, x is the unique solution of (4.9). Since π has countably many jumps and changing the Lebesgue integral is left unchanged by changing the integrand at countably many points, we then get by definition of v:

$$
x_t = x_0 - \int_0^t av_s\mathrm{d}s.
$$

It remains to show that the second equation in (3.6) is verified. Using that p is solution of (4.7), π verifies (4.8) and x is solution of (4.9), we get:

$$
\begin{aligned}
y_t &= p_t x_t + \pi_t \\
&= \mathbb{E}_t\left[p_T x_T - \int_t^T\left(\frac{\mathrm{d}p_s}{\mathrm{d}s}x_s + \frac{\mathrm{d}x_s}{\mathrm{d}s}p_s\right)\mathrm{d}s + \pi_T - \int_t^T (agp_s\pi_s + ap_s h_s - c)\mathrm{d}s\right]
\end{aligned}
$$

$$= \mathbb{E}_t \left[ex_T + f + \int_t^T \left\{ -(agp_s^2 - b)x_s + (agp_s x_s + ag\pi_s + ah_s)p_s - (agp_s\pi_s + ap_s h_s - c) \right\} \mathrm{d}s \right]$$

$$= \mathbb{E}_t \left[ex_T + f + \int_t^T (bx_s + c)\,\mathrm{d}s \right].$$

This shows that (v, x, y) is a solution of (3.6) in $\mathbb{H}^{\infty,2} \times \mathbb{H}^{\infty,2} \times \mathbb{H}^{2,2}_{\mathcal{P}}$.

The uniqueness of the solution for small time T follows from a fixed point argument as in the proof of Theorem 2.4. We do not repeat the arguments here. □

4.5. Proof of Proposition 3.6

▷ By definition of $\dot{u}^{(\varepsilon)}$, $\dot{X}^{(\varepsilon)}$ and $\dot{Y}^{(\varepsilon)}$, they are clearly solutions of (3.10). Let us turn to uniform boundedness. The first two equations yield:

$$\|\dot{X}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \le \frac{\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}, \qquad \|\dot{Y}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \le \frac{\nu T + \gamma\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}.$$

This can be easily proved following the arguments given in the proof of Theorem 2.7, details are left to the reader. From the last equation of (3.10) and the 1-Lipschitz continuity of $x \to x_+$, we get:

$$\begin{aligned} \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} &\le \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}} \|\dot{Y}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} + 2\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(\varepsilon)})_+\|_{\mathbb{H}^{2,2}} \\ &\le \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2} \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} + 2\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}} + 2\varepsilon \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}. \end{aligned}$$

When T and ε are small, such that,

$$\alpha(T) + 2\varepsilon := \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2} + 2\varepsilon < 1,$$

we obtain

$$\|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \le \frac{2\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}}{1 - \alpha(T) - 2\varepsilon},$$

whence the uniform boundedness of $\dot{u}^{(\varepsilon)}$ as $\varepsilon \to 0$, provided $\alpha(T) < 1$.

▷ Now we prove the convergence of $(\dot{u}^{(\varepsilon)}, \dot{X}^{(\varepsilon)}, \dot{Y}^{(\varepsilon)})$ to $(\dot{u}, \dot{X}, \dot{Y})$ in $\mathbb{H}^{2,2}$-norms as $\varepsilon \to 0$. Similarly as before, we have:

$$\|\dot{X}^{(\varepsilon)} - \dot{X}\|_{\mathbb{H}^{2,2}} \le \frac{\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}}, \qquad \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} \le \frac{\nu T + \gamma\sqrt{T}}{\mathcal{E}_{\max}} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}}.$$

Besides, the last equations in (3.10) and (3.11) as well as the 1-Lipschitz continuity of $x \to x_+$ give:

$$\begin{aligned} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} &\le \frac{1}{(\lambda+\mu)\mathcal{E}_{\max}} \|\dot{Y}^{(\varepsilon)} - \dot{Y}\|_{\mathbb{H}^{2,2}} + 2\|u^{\Delta,(\varepsilon)} - u^{\Delta,(0)}\|_{\mathbb{H}^{2,2}} \\ &\le \frac{\nu T + \gamma\sqrt{T}}{(\lambda+\mu)\mathcal{E}_{\max}^2} \|\dot{u}^{(\varepsilon)} - \dot{u}\|_{\mathbb{H}^{2,2}} + 2\varepsilon \|\dot{u}^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \end{aligned}$$

$$= \alpha(T)\|\dot u^{(\varepsilon)} - \dot u\|_{\mathbb{H}^{2,2}} + 2\varepsilon\|\dot u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}}.$$

For T small enough s.t. $\alpha(T) < 1$ and and for $\varepsilon < \frac{1-\alpha(T)}{2}$, we thus obtain:

$$\|\dot u^{(\varepsilon)} - \dot u\|_{\mathbb{H}^{2,2}} \le \frac{2\varepsilon}{1-\alpha(T)}\|\dot u^{(\varepsilon)}\|_{\mathbb{H}^{2,2}} \le \frac{4\varepsilon}{(1-\alpha(T))(1-\alpha(T)-2\varepsilon)}\|(\mathtt{P}^{\mathtt{load},\Delta} - u^{\Delta,(0)})_+\|_{\mathbb{H}^{2,2}}.$$

This completes the proof. □

4.6. Boundedness of solutions to linear ODE with bounded stochastic coefficient

This following result is used in the proof of Theorems 2.2 and 2.4.

Lemma 4.1. *Let $A : [0,T] \times \Omega \mapsto \mathbb{R}^p \times \mathbb{R}^p$ be a random matrix-valued process. Suppose there exists a constant C such that $|A(t,\omega)| \le C$, $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-a.e.*

Let R and L be the unique (continuous) solutions of the following linear ODEs:

$$\begin{cases} \frac{\mathrm{d}L_t}{\mathrm{d}t} = L_t A_t, \\ L_0 = \mathrm{Id}_p, \end{cases} \qquad \textit{and} \qquad \begin{cases} \frac{\mathrm{d}R_t}{\mathrm{d}t} = -A_t R_t, \\ R_0 = \mathrm{Id}_p. \end{cases}$$

Then, L and R are invertible with $L^{-1} = R$. Besides, $|L_t|$ and $|R_t|$ are uniformly bounded on $[0,T]$ by $\exp(CT)$.

Proof. A direct computation shows that

$$\mathrm{d}t \times \mathrm{d}\mathbb{P}\text{-a.e.}, \qquad \frac{\mathrm{d}(L_t R_t)}{\mathrm{d}t} = 0,$$

thus $\forall t \in [0,T]$, $L_t R_t = L_0 R_0 = \mathrm{Id}_p$. Therefore R and L are invertible with $R = L^{-1}$. Let us now turn to the uniform boundedness. Let $v \in \mathbb{R}^p$, we have

$$\frac{\mathrm{d}|L_t^\top v|^2}{\mathrm{d}t} = v^\top L_t(A_t + A_t^\top)L_t^\top v \le |A_t + A_t^\top|\,|L_t^\top v|^2 \le 2C\,|L_t^\top v|^2, \quad \mathrm{d}t \times \mathrm{d}\mathbb{P}\text{-a.e.}.$$

Therefore, by integration, $|L_t^\top v|^2 \le |v|^2 \exp(2CT)$ for $t \in [0,T]$, which yields $\sup_{0\le t\le T} |L_t^\top|^2 \le \exp(2CT)$. This proves $\exp(CT) \ge \sup_{0\le t\le T} |L_t^\top| = \sup_{0\le t\le T} |L_t|$, whence the announced bound for L. For bounding $|R|$ start from $\frac{\mathrm{d}|R_t v|^2}{\mathrm{d}t}$ and proceed similarly. □

5. Conclusion

In this work, we have identified the optimal control of storage facilities of a smart grid under uncertain consumption/production, in order to reduce the stochastic fluctuations of the residual consumption on the electrical public grid. It has been possible thanks to the resolution of a new extended McKean-Vlasov stochastic control problem, using Pontryagin principle and Forward Backward Stochastic Differential Equations. For situations where the costs are close to quadratic functions, we have derived quasi-explicit formulas for the control, using perturbation arguments.

In further works, we will consider subsequent issues like more realistic dynamics of the battery flow accounting with aging/boundary effect, sizing of the smart grid and of the storage/production capacities, risk aggregation of optimized smart grids with dependent solar productions, impact of model mis-specification on the optimal solution (risk model).

References

[1] B. Acciaio, J. Backhoff-Veraguas and R. Carmona, Extended mean field control problems: stochastic maximum principle and transport perspective. *SIAM J. Control Optim.* **57** (2019) 3666–3693.

[2] C. Alasseur, I.B. Taher and A. Matoussi, An extended mean field game for storage in smart grids. *J. Optim. Theory Appl.* **184** (2020) 644–670.

[3] A. Angiuli, C. Graves, H. Li, J.-F. Chassagneux, F. Delarue and R. Carmona, Cemracs 2017: numerical probabilistic approach to MFG. *ESAIM: PS* **65** (2019) 84–113.

[4] J. Badosa, E. Gobet, M. Grangereau and D. Kim, Day-ahead probabilistic forecast of solar irradiance: a Stochastic Differential Equation approach. Renewable Energy: Forecasting and Risk Management. Edited by P. Drobinski, M. Mougeot, D. Picard, R. Plougonven and P. Tankov, Springer Proceedings in Mathematics & Statistics (2018) 73–93.

[5] M. Basei and H. Pham, A weak martingale approach to linear-quadratic McKean–Vlasov stochastic control problems. *J. Optim. Theory Appl.* **181** (2019) 347–382.

[6] F. Bellini, B. Klar, A. Muller and E.R. Gianin, Generalized quantiles as risk measures. *Insurance: Math. Econ.* **54** (2014) 41–48.

[7] A. Bensoussan, Perturbation methods in optimal control. *Wiley/Gauthier-Villars Series in Modern Applied Mathematics*. Translated from the French by C. Tomson. John Wiley & Sons Ltd., Chichester (1988).

[8] B. Bibby, I. Skovgaard and M. Sorensen, Diffusion-type models with given marginal distribution and autocorrelation function. *Bernoulli* **11** (2005) 191–220.

[9] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations. Springer Science+Business Media (2010).

[10] R. Carmona and F. Delarue, Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.* **43** (2015) 2647–2700.

[11] R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications I-II. Springer (2018).

[12] P. Carpentier, J. Chancelier, M. De Lara and T. Rigaut, Algorithms for two-time scales stochastic optimization with applications to long term management of energy storage. HAL preprint hal-02013969 (2019).

[13] F. Delarue and G. Guatteri, Weak existence and uniqueness for forward-backward SDEs. *Stoch. Process. Appl.* **116** (2006) 1712–1742.

[14] N. El Karoui, S. Peng and M. Quenez, Backward stochastic differential equations in finance. *Math. Finance* **7** (1997) 1–71.

[15] B. Heymann, J. Bonnans, P. Martinon, F. Silva, F. Lanas and G. Jiménez-Estévez, Continuous optimal control approaches to microgrid energy management. *Energy Syst.* **9** (2018) 59–77.

[16] B. Heymann, J. Bonnans, F. Silva and G. Jimenez, A stochastic continuous time model for microgrid energy management. In *Control Conference (ECC), 2016 European*. IEEE (2016) 2084–2089.

[17] Z. Hussien, L. Cheung, M. Siam and A. Ismail, Modeling of sodium sulfur battery for power system applications. *Elektrika J. Electr. Eng.* **9** (2007) 66–72.

[18] E. Iversen, J. Morales and H. Madsen, Optimal charging of an electric vehicle using a Markov decision process. *Appl. Energy* **123** (2014) 1–12.

[19] J. Morales, A. Conejo, H. Madsen, P. Pinson and M. Zugno, Integrating renewables in electricity markets. Vol. 205 of *International Series in Operations Research & Management Science*. Springer Science + Business Media, New York (2014).

[20] F. Pacaud, P. Carpentier, J. Chancelier and M. De Lara, Stochastic optimal control of a domestic microgrid equipped with solar panel and battery. Preprint arXiv:1801.06479 (2018).

[21] E. Pardoux and S. Tang, Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Relat. Fields* **114** (1999) 123–150.

[22] S. Peng, A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* **28** (1990) 966–979.

[23] H. Pham and X. Wei, Discrete time McKean–Vlasov control problem: a dynamic programming approach. *Appl. Math. Optim.* **74** (2016) 487–506.

[24] H. Pham and X. Wei, Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: COCV* **24** (2018) 437–461.

[25] P.E. Protter, Stochastic Integration and Differential Equations. Springer Berlin Heidelberg, Berlin, Heidelberg, second edition (2003).

[26] C. Sun, F. Sun and S. Moura, Nonlinear predictive energy management of residential buildings with photovoltaics & batteries. *J. Power Sources* **325** (2016) 723–731.

[27] V. Trovato, I.M. Sanz, B. Chaudhuri and G. Strbac, Advanced control of thermostatic loads for rapid frequency response in great Britain. *IEEE Trans. Power Syst.* **32** (2016) 2106–2117.

[28] V. Trovato, S.H. Tindemans and G. Strbac, Leaky storage model for optimal multi-service allocation of thermostatic loads. *IET Gener. Trans. Distrib.* **10** (2016) 585–593.

[29] X. Wu, X. Hu, S. Moura, X. Yin and V. Pickert, Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array. *J. Power Sources* **333** (2016) 203–212.

[30] J. Yong, Linear forward-backward stochastic differential equations with random coefficients. *Prob. Theory Related Fields* **135** (2006) 53–83.

[31] J. Yong, Linear-quadratic optimal control problems for mean-field stochastic differential equations. *SIAM J. Control Optim.* **51** (2013) 2809–2838.