



HAL
open science

A decentralized algorithm for a Mean Field Control problem of Piecewise Deterministic Markov Processes

Adrien Séguret, Thomas Le Corre, Nadia Oudjane

► **To cite this version:**

Adrien Séguret, Thomas Le Corre, Nadia Oudjane. A decentralized algorithm for a Mean Field Control problem of Piecewise Deterministic Markov Processes. *ESAIM: Probability and Statistics*, 2024, 28, pp.22-45. 10.1051/ps/2023021 . hal-03910622v2

HAL Id: hal-03910622

<https://hal.science/hal-03910622v2>

Submitted on 25 Jan 2024

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

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



Distributed under a Creative Commons Attribution 4.0 International License

A decentralized algorithm for a Mean Field Control problem of Piecewise Deterministic Markov Processes *

Adrien Séguret¹

Thomas Le Corre²

Nadia Oudjane³

Abstract

This paper provides a decentralized approach for the control of a population of N agents to minimize an aggregate cost. Each agent evolves independently according to a Piecewise Deterministic Markov dynamics controlled via unbounded jumps intensities. The N -agent high dimensional stochastic control problem is approximated by the limiting mean field control problem. A Lagrangian approach is proposed. Although the mean field control problem is not convex, it is proved to achieve zero duality gap. A stochastic version of the Uzawa algorithm is shown to converge to the primal solution. At each dual iteration of the algorithm, each agent solves its own small dimensional sub problem by means of the Dynamic Programming Principal, while the dual multiplier is updated according to the aggregate response of the agents. Finally, this algorithm is used in a numerical simulation to coordinate the charging of a large fleet of electric vehicles in order to track a target consumption profile.

1 Introduction

This paper focuses on a specific family of stochastic control problems called "large and aggregative". The problem consists in optimizing the strategies of a large number of agents, in a random environment, while the decisions of all the agents interact via the objective function because the latter depends on an "aggregative" term as the sum of the decisions of all the agents. This type of problem occurs naturally in several fields of application such as energy management [19], telecommunications [23], portfolio management [21], robot coordination [6] among others.

To deal with the curse of dimensionality in such high dimensional stochastic control problems, methods based on the value function have been developed in the convex setting. For instance, Stochastic Dual Dynamic Programming [24] relies on polyhedral approximations of the value function providing good performances for medium sized problems with a number of agents $N \leq 30$. In our case, where the number of agents is very large $N \geq 1000$, we look more specifically for decentralized solutions, allowing both to respect the privacy of each agent and to reduce the complexity of the problem. In the particular setting, where controls are bounded and strong convexity conditions are satisfied, [28] proposed an iterative stochastic algorithm providing a decentralized solution to such problems. In the same spirit as the Dual Approximate Dynamic Programming [7], this approach relies on a Lagrangian decomposition technique to obtain a decentralized solution, but it takes advantage of a mean field approximation to ensure the validity of the approach when the number of agents is sufficiently large.

This paper aims at extending the decentralized approach developed in [28] to a non-convex framework, with unbounded controls, allowing for more realistic dynamics ruling the evolution of each agent state. In particular, we are interested in the aggregative control of Piecewise Deterministic Markov Processes (PDMPs for short), where the state of each agent is modelled by a PDMP controlled by an unbounded jump intensity. PDMPs were introduced in [10] as a class of non-diffusion stochastic models, mixing random jumps and deterministic dynamics between jumps. This kind of process is used to model a wide broad of phenomena or situations, such as system reliability and maintenance [12, 32], oil production [33], biology [22, 26] etc... The approach proposed in this paper results in a decentralized algorithm, where at each iteration, each agent is intended to solve its own small dimensional optimal control problem of PDMP. This type of problem is well known and has been studied through dynamic programming and related equations [2, 9, 11, 18, 31]. Due to the large number of agents, the problem is formulated

*This research benefited from the support of the FMJH Program Gaspard Monge for optimization and operations research and their interactions with data science.

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

²Thomas Le Corre is with DI ENS, CNRS, PSL University, INRIA Paris thomas.le-corre@inria.fr

³Nadia Oudjane is with Osiris, EDF R&D, 91120 Palaiseau, France and with Finance for Energy Market Research Centre (FIME), Paris, France nadia.oudjane@edf.fr

as a Mean Field Control (MFC for short) problem. Relying on the law of large numbers, the MFC theory consists in controlling a representative agent of the population, where the interaction between the agents is approximated by the interaction of this representative agent with the probability distribution of its state or control [5]. This work is particularly motivated by demand-side management in power systems to help balancing between energy production and consumption. This problem is indeed critical today [30], due to the increasing share of uncontrollable energies (wind and solar) in the electricity generation mix, which requires to compensate the uncontrollable character of the production by controlling the demand. Controlling the sum of the consumption of each flexible consumer in order to balance the electrical system has already been investigated successfully in the specific framework of Quadratic Kullback-Leibler control problems [4]. In [29] and [20], a mean field assumption is also considered to control the charging of a large fleet of electrical vehicles (EVs for short), leading to the optimal control of partial differential equation problems.

The novelty of the present paper is to extend the Lagrangian approach proposed in [28], to a particular non-convex setting with unbounded controls by first establishing the existence of a saddle point and then ensuring the convergence of the decentralized algorithm. The first contribution lies in the originality of the proof of a saddle point existence which follows a completely different path than the one developed in [28]. In particular, we make use of the existence and of the regularity results on the solutions of an optimal control problem of a continuity equation developed in [27], to show the existence of a saddle point of our Lagrangian problem. We highlight that while bounds assumptions on the jump intensity [2, 9] or compactness assumptions to analyse the Hamilton Jacobi Bellman equation [18] are common in the optimal control of PDMPs literature, we obtain the existence of a solution without assuming any a priori bounds on the controlled jump intensity. The second contribution consists in proving that the Stochastic Uzawa Algorithm proposed in [28] is still providing a converging sequence of controls, in this new setting involving the PDMP dynamics which violates the convexity conditions originally exploited. Finally, an application to the smart charging of an electric vehicle fleet by an aggregator illustrates the performance and the interest of the approach to coordinate the charge of a large number of electric vehicles in order to track a given target power consumption profile for the whole population of electric vehicles.

The outline of this paper is as follows. In Section 2, we formulate the optimization problem, the assumptions and the main results. Section 3 presents a dual approach of this problem. The Stochastic Uzawa algorithm is presented in Section 4 and is proved to converge in this setting. Section 5 presents simulations of the coordination of power consumption of a large fleet of electrical vehicles.

In the rest of this section we will list some frequently used notations.

Notation Let \mathcal{X} be the state space defined by $\mathcal{X} := I \times [0, 1]$, where I is a finite set, of cardinality $d \in \mathbb{N}^*$. Let $[0, T]$ be a time interval with $T > 0$. The space $D([0, T], I)$ is the space of càdlàg functions from $[0, T]$ to I . The spaces $C^1(\mathcal{X})$ (resp. $C^1([0, T] \times \mathcal{X})$) denotes the sets of real-valued continuously differentiable functions defined on \mathcal{X} (resp. on $[0, T] \times \mathcal{X}$). The space $C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$ denotes the set of continuous functions from $[0, T] \times \mathcal{X}$ to \mathbb{R}_+^d and is endowed with the norm $\|f\|_\infty := \sup_{(i,t,j,z) \in I \times [0,T] \times \mathcal{X}} |f_i(t, j, z)|$. We consider the spaces

$$\begin{aligned} \mathfrak{A} &:= \{\alpha \in C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d) \mid \forall i \in I, \alpha_i(\cdot, i, \cdot) = 0\}, \\ \mathfrak{B} &:= C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d) \cap L^\infty([0, T], W^{1,\infty}(\mathcal{X}, \mathbb{R}_+^d)), \\ \mathfrak{C} &:= \{f \in C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d) \mid \partial_z f_i \in C^0([0, T] \times \mathcal{X})\}. \end{aligned} \tag{1.1}$$

where ∂_z is the space derivative w.r.t. the space variable $z \in [0, 1]$ and $W^{1,\infty}(\mathcal{X}, \mathbb{R}_+^d)$ is a Sobolev Space [3, Section 8.2], endowed with norm $\|\cdot\|_{W^{1,\infty}}$, defined by

$$\begin{aligned} W^{1,\infty}(\mathcal{X}, \mathbb{R}_+^d) &:= \{f \in L^\infty(\mathcal{X}, \mathbb{R}_+^d) \mid \partial_z f \in L^\infty(\mathcal{X}, \mathbb{R}_+^d)\} \\ \|f\|_{W^{1,\infty}} &:= \|f\|_\infty + \|\partial_z f\|_\infty, \end{aligned}$$

and $L^\infty([0, T], W^{1,\infty}(\mathcal{X}, \mathbb{R}_+^d))$ is defined as the set of functions f in $L^\infty([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$ such that $\{\|f(t, \cdot)\|_{W^{1,\infty}}\}_{t \in [0, T]}$ is bounded. The space of Borel probability measures on the space \mathcal{X} is denoted by $\mathcal{P}(\mathcal{X})$. We recall the definition of the Wasserstein distance, denoted by \mathbb{W} , on the space $\mathcal{P}(\mathcal{X})^2$:

$$\mathbb{W}(\mu, \rho) := \sup \left\{ \sum_{i \in I} \int_0^1 \varphi(i, z) (\mu(i, dz) - \rho(i, dz)) \mid \varphi \text{ is } 1 - \text{Lipschitz from } \mathcal{X} \text{ to } \mathbb{R} \right\},$$

where the discrete set is I is endowed with the usual discrete distance function (equal to 1 for different elements). The space $C^0([0, T], \mathcal{P}(\mathcal{X}))$ is the set of continuous functions from $[0, T]$ to $\mathcal{P}(\mathcal{X})$, endowed with the distance $\mathbb{W}_{\|\cdot\|_\infty}$ defined by:

$$\mathbb{W}_{\|\cdot\|_\infty}(m^1, m^2) := \sup_{t \in [0, T]} \mathbb{W}(m^1(t), m^2(t)). \quad (1.2)$$

Let \mathfrak{H} be an Hilbert space and F be a real valued function defined on \mathfrak{H} . The convex conjugate of F is denoted by F^* and is defined, for any $x \in \mathfrak{H}$, by $F^*(x) := \sup_{y \in \mathfrak{H}} \langle x, y \rangle_{\mathfrak{H}} - F(y)$.

2 Problem formulation and main results

Motivation

We consider a population of N independent and identically distributed processes, controlled by a central planner via a common feedback control α . The space of feedbacks, denoted by \mathfrak{A} , is defined in (1.1) and is endowed with the norm $\|\cdot\|_\infty$. The state of an agent, controlled by the jump intensity α , is given at time t by $X_t^\alpha = (Y_t^\alpha, Z_t^\alpha)$ with $Y_t^\alpha \in I$ and $Z_t^\alpha \in [0, 1]$. In the smart charging application in Section 5, Y^α represents the charging mode of a vehicle, specifying its discrete charging rate (e.g. $i = 0$ for not charging, $i = 1$ for charging and $i = -1$ for discharging). The continuous variable Z^α represents the state of charge (SoC for short) of an EV, that is the ratio of energy charged into the battery w.r.t. the total capacity of the battery. We say that the process X^α is a PDMP(b, α) with jump intensity α if $X^\alpha = (Y^\alpha, Z^\alpha)$, where Y^α is a jump process with values in I switching spontaneously, at jump times $\{T_k^\alpha\}_{k \in \mathbb{N}}$ given by a Poisson process with intensity α_i to switch to mode $i \in I$. The constraint $\alpha_i(\cdot, i, 0) = 0$ on the control space \mathfrak{A} in (1.1) is considered to avoid jumps of Y^α from i to i . The process Z^α follows a deterministic dynamics between two consecutive jumps $Z_t^\alpha = \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, t)$ for $t \in [T_k^\alpha, T_{k+1}^\alpha)$. For any $(\tau, t, j, z) \in [0, T] \times [\tau, T] \times \mathcal{X}$, the flow ϕ is the unique solution of the ordinary differential equation:

$$\partial_t \phi_{\tau, z}(j, t) = b(j, \phi_{\tau, z}(j, t)) \text{ with } \phi_{\tau, z}(j, \tau) = z, \quad (2.1)$$

where b is a function that is exogenous and depends on the characteristics of the battery and of the charging station. To be more specific, following the definition of a PDMP given in [13], knowing T_k^α and $X_{T_k^\alpha}^\alpha = (Y_{T_k^\alpha}^\alpha, Z_{T_k^\alpha}^\alpha)$, one obtains $(T_{k+1}^\alpha, X_{T_{k+1}^\alpha}^\alpha)$ as follows:

$$\left\{ \begin{array}{l} T_{k+1, j}^\alpha := \inf \left\{ t \geq T_k^\alpha : E_{k+1, j} < \int_{T_k^\alpha}^t \alpha_j(r, \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, r)) dr \right\} \quad \forall j \in I \\ T_{k+1}^\alpha := \min_{j \in I} T_{k+1, j}^\alpha \\ Y_{T_{k+1}^\alpha}^\alpha = \min \{ j \in I \mid T_{k+1, j}^\alpha = T_{k+1}^\alpha \} \\ Z_{T_{k+1}^\alpha}^\alpha := \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, T_{k+1}^\alpha) \\ X_{T_{k+1}^\alpha}^\alpha = (Y_{T_{k+1}^\alpha}^\alpha, Z_{T_{k+1}^\alpha}^\alpha), \end{array} \right. \quad (2.2)$$

where $\{E_{k, j}\}_{k, j}$ are independent random variables following an exponential distribution of parameter 1, independent of X_0^α . Indeed, the control process $\alpha = (\alpha_j)_{j \in I}$ is such that $\alpha_j(t, x)$ represents the intensity rate, at time $t \in [0, T]$, of jumping into mode $j \in I$ when coming from state $x \in \mathcal{X}$. We initialize $T_0 = 0$ and at time $t = 0$, the law of the couple of random variables (Y_0^α, Z_0^α) is given by $m^0 \in \mathcal{P}(\mathcal{X})$. Although the dynamics of X^α might seem specific due to the vector valued jump intensity α , the dynamics (2.1) and (2.2) define a PDMP and the link with the definition of a PDMP introduced in [10] is described in Section A.1 of the Appendix. Since the control space \mathfrak{A} is made of continuous functions and that we assume below in Assumption 1 that b is in $C^1(\mathcal{X})$, the PDMP(b, α) is well defined for any $\alpha \in \mathfrak{A}$. The cost function J^N for the N -agents problem is defined for any $\alpha \in \mathfrak{A}$ by:

$$J^N(\alpha) := \mathbb{E} \left[\int_0^T f \left(t, \frac{1}{N} \sum_{n=1}^N p(t, X_t^{n, \alpha}) \right) dt + \frac{1}{N} \sum_{n=1}^N G(\alpha, X^{n, \alpha}) \right], \quad (2.3)$$

where $\{X^{n, \alpha}\}_{n \in \{1, \dots, N\}}$ are supposed to be independent PDMP(b, α) driven by independent exponential variables. Indeed, for any $n \in \{1, \dots, N\}$, $X^{n, \alpha}$ is supposed to be controlled by the common feedback function $\alpha \in \mathfrak{A}$ only depending on the agent state X^n . Besides, the function G is common to every agent and defined as the the sum of

a running and a terminal cost, such that for any $x := (y, z) \in D([0, T], I) \times C^0([0, T], [0, 1])$, by:

$$G(\alpha, x) := \int_0^T \left[c(t, x_t) + \sum_{j \in I} L(\alpha_j(t, x_t)) \right] dt + g(x_T).$$

From a practical point of view, the function f represents a coupling cost depending on the aggregate quantity $\frac{1}{N} \sum_{n=1}^N p(t, X_t^{n, \alpha})$, and G the individual cost.

Problem formulation

In this paper, we focus on the mean field limit control problem, corresponding to the problem with an infinite population. Thus, we introduce the function J defined, for any $\alpha \in \mathfrak{A}$, by:

$$J(\alpha) := \int_0^T f(t, \mathbb{E}[p(t, X_t^\alpha)]) dt + \mathbb{E}[G(\alpha, X^\alpha)], \quad (2.4)$$

where X^α is a PDMP(b, α). This paper is dedicated to the following problem:

$$\min_{\alpha \in \mathfrak{A}} J(\alpha). \quad (2.5)$$

As the cost function (2.4) is nonlinear w.r.t. the expectation term, $\mathbb{E}[p(t, X_t^\alpha)]$, via the coupling cost f , Problem (2.5) goes beyond the scope of optimal control of PDMP. We propose to numerically solve Problem (2.5) by applying the Stochastic Uzawa Algorithm (1), that is detailed in Section 4. This algorithm is introduced in [28] and relies on the same arguments as those underlying the stochastic gradient descent algorithm [16]. The main result of this paper, stated in Theorem 2.1, ensures the convergence of the Stochastic Uzawa Algorithm 1 to the solution of Problem (2.5).

Throughout the paper, we assume the following:

General assumptions

Assumptions 1. *The vector field $b \in C^1(\mathcal{X})$ is assumed to vanish at the boundary: $b(j, 0) = b(j, 1) = 0$ for any $j \in I$.*

Assumptions 2. *The function p is in $C^1([0, T] \times \mathcal{X})$ and $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory function being lower semi-continuous (l.s.c. for short), strictly convex and twice differentiable w.r.t. the second variable with Lipschitz gradient uniformly in time. More precisely, for a.e. $t \in [0, T]$, we have $f(t, \cdot) \in C^2(\mathbb{R})$ and there exists a constant $C > 0$ independent of t , such that $\|\partial_z^2 f(t, \cdot)\|_\infty \leq C$. In addition, there exists a constant $C_f > 0$ such that, for any $(t, x) \in [0, T] \times \mathbb{R}$,*

$$\frac{x^2}{2C_f} - C_f \leq f(t, x) \leq C_f \frac{x^2}{2} + C_f. \quad (2.6)$$

Assumptions 3. *It is assumed that $c \in C^1([0, T] \times \mathcal{X})$ and $g \in C^1(\mathcal{X})$.*

Assumptions 4. *The function $L : \mathbb{R} \rightarrow \bar{\mathbb{R}}$ is defined by:*

$$L(x) := \begin{cases} l(x) & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ +\infty & \text{otherwise,} \end{cases}$$

where $l \in C^1(\mathbb{R}_+, \mathbb{R}_+)$ is an increasing strongly convex function bounded from above by a quadratic function. More explicitly, there exists $C > 0$ such that for any $x \in \mathbb{R}_+$:

$$\frac{x^2}{C} - C \leq l(x) \leq C(x^2 + 1),$$

where the first inequality is due to the strong convexity of l . We assume that l satisfies $\lim_{x \rightarrow 0} l(x) = 0$. Due to the assumptions on l , L is convex. We denote by H the convex conjugate of L , i.e. H is defined for any $x \in \mathbb{R}$ by

$$H(x) := \sup_{y \in \mathbb{R}} xy - L(y).$$

Remark 2.1. The main role of Assumption 1 is to ensure that the flow defined in (2.1) exists and takes values in $[0, 1]$. Assumption 2 ensures that $\int_0^T f(t, \mathbb{E}[p(t, X_t^\alpha)]) dt$ in (2.4) is well defined for any $\alpha \in \mathfrak{A}$. Also, the regularity assumptions w.r.t. the space variable z of the function p in Assumption 2 is crucial to show the existence of a fixed point of a function involving p in Lemma A.2. The continuity of the map $\alpha \mapsto \mathbb{E}[G(\alpha, X^\alpha)]$ is an automatic consequence of the properties of c, g and L given in Assumptions 3 and 4. By Assumption 4, H is non-decreasing, non-negative, and H' is Lipschitz continuous on \mathbb{R} .

Remark 2.2. The boundary condition on b may seem restrictive in Assumption 1. However, if $\overline{\text{supp}(m^0)} \subset I \times (0, 1)$, any function b that is such that, for any $i \in I$, $b(i, 0) \geq 0$, $b(i, 1) \leq 0$ and the sign of $b(i, \cdot)$ is constant, can be approximated by a function \tilde{b} satisfying the boundary condition. Indeed, for such a function b , there exists $\varepsilon > 0$ such that $\mathbb{P}(Z_t^\alpha \in (\varepsilon, 1 - \varepsilon)) = 1$ for any $\alpha \in \mathfrak{A}$ and any $t \in [0, T]$. Therefore, there exists an open subset containing the boundaries, that is a.s. never attained and it is possible to approximate b by a smooth function \tilde{b} , vanishing at the boundary without modifying the trajectory of $Z_t^{n, \alpha}$.

The main results of the paper are summarized in the following theorem, and provide the convergence of both the sequence of iterates and the sequence of function values generated by Algorithm 1 stated in Section 4.

Theorem 2.1. Problem 2.5 has a solution. Let $\{\alpha^k\}_{k \in \mathbb{N}}$ be a sequence in \mathfrak{A} generated by Algorithm 1 (line 5), then the following assertions hold

- (i) The sequence $\{\alpha^k\}_{k \in \mathbb{N}}$ converges a.s. to an element of $\underset{\alpha \in \mathfrak{A}}{\text{argmin}} J(\alpha)$ w.r.t. the norm $\|\cdot\|_\infty$.
- (ii) The sequence $\{J(\alpha^k)\}_{k \in \mathbb{N}}$ converges a.s. to $\min_{\alpha \in \mathfrak{A}} J(\alpha)$.

This theorem can not be obtained by a direct application of [28], which relies on additional assumptions, that are not verified in the present framework, which involves unbounded controls and PDMP dynamics. The proof of Theorem 2.1 is given in Section 4. Before, we introduce and analyze in the next section a dual problem proved to be equivalent, in some sense, to Problem 2.5.

3 Dual approach

In the same vein as [28], a Lagrangian decomposition approach is adopted to obtain a decentralized algorithm. Contrary to [28], the specific assumptions of convexity, that ensure the absence of duality gap [14, Appendix I] and the convergence of the Stochastic Uzawa Algorithm, are not satisfied here. In Sections 3 and 4, we propose new theoretical arguments allowing to demonstrate the validity of the approach proposed in [28] in this specific framework of control of PDMP.

Let F be defined for any $v \in L^2(0, T)$ by:

$$F(v) := \int_0^T f(t, v(t)) dt.$$

Owing to the properties of f in Assumption 2, one deduces that function F is strictly convex and has at least quadratic growth. Indeed, for any $v, u \in L^2(0, T)$ and $\gamma \in [0, 1]$, by the strict convexity of f , one has

$$\begin{aligned} F(\gamma v + (1 - \gamma)u) &= \int_0^T f(t, \gamma v(t) + (1 - \gamma)u(t)) dt < \int_0^T \gamma f(t, v(t)) + (1 - \gamma)f(t, u(t)) dt \\ &= \gamma F(v) + (1 - \gamma)F(u), \end{aligned}$$

and by (2.6)

$$\frac{\|v\|_2^2}{2C_f} - TC_f \leq F(v) \leq C_f \frac{\|v\|_2^2}{2} + TC_f.$$

Lemma 3.1. The problem

$$\begin{cases} \min_{\alpha \in \mathfrak{A}, v \in L^2(0, T)} \bar{J}(\alpha, v), \\ \bar{J}(\alpha, v) := F(v) + \mathbb{E}[G(\alpha, X^\alpha)], \\ \text{s.t. } \mathbb{E}[p(t, X_t^\alpha)] - v(t) = 0 \text{ a.e on } [0, T], \end{cases} \quad (3.1)$$

is equivalent to Problem (2.5).

Proof. The fact that $p \in C^1([0, T] \times \mathcal{X}, \mathbb{R})$ implies that the map $t \mapsto \mathbb{E}[p(t, X_t^\alpha)]$ is in $L^2(0, T)$. The equivalence between the two problems then follows from the definitions of J and \bar{J} . \square

Let us introduce the Lagrangian $\mathcal{L} : \mathfrak{A} \times L^2(0, T) \times L^2(0, T) \rightarrow \mathbb{R}$:

$$\mathcal{L}(\alpha, v, \lambda) := \mathcal{L}_1(\alpha, \lambda) + \mathcal{L}_2(v, \lambda), \quad (3.2)$$

where

$$\mathcal{L}_1(\alpha, \lambda) := \mathbb{E}[G(\alpha, X^\alpha)] + \int_0^T \mathbb{E}[p(t, X_t^\alpha)]\lambda(t)dt, \quad (3.3)$$

$$\mathcal{L}_2(v, \lambda) := F(v) - \int_0^T v(t)\lambda(t)dt, \quad (3.4)$$

and the associated dual function $\mathcal{W} : L^2(0, T) \rightarrow \mathbb{R}$:

$$\mathcal{W}(\lambda) := \inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda) + \inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda), \quad (3.5)$$

where $\inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ and $\inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ are each referred to as sub problem. The dual problem [14, Section 3 Chapter 3] associated to (3.1) consists of the following maximization problem:

$$\max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda). \quad (3.6)$$

The following lemma states the existence of a unique solution of the dual problem (3.6). It is derived from the assumptions on f in Assumption 2.

Lemma 3.2. *There exists a unique $\bar{\lambda} \in L^2(0, T)$ such that $\bar{\lambda} = \operatorname{argmax}_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda)$.*

Proof. By Assumption 2, one has for a.e. $t \in [0, T]$, that $f(t, \cdot) \in C^2(\mathbb{R})$ and $\|\partial_z^2 f(t, \cdot)\|_\infty \leq C$ (C being independent of t). Thus, by applying the dominated convergence theorem [3, Theorem 4.2], one deduces that for any $v_1, v_2 \in L^2(0, T)$,

$$\lim_{h \rightarrow 0} \frac{F(v_1 + hv_2) - F(v_1)}{h} = \int_0^T \partial_z f(t, v_1(t))v_2(t)dt.$$

Therefore, the map F is Gâteaux differentiable at any $v \in L^2(0, T)$ [14, Definition 5.2]. Denoting by $DF(v)$ the Gâteaux derivative of F at $v \in L^2(0, T)$, using the previous equation and that $\|\partial_z^2 f(t, \cdot)\|_\infty \leq C$ for a.e. $t \in [0, T]$, one has for any $v_1, v_2, v_3 \in L^2(0, T)$

$$\begin{aligned} |DF(v_1)(v_3) - DF(v_2)(v_3)| &= \left| \int_0^T (\partial_z f(t, v_1(t)) - \partial_z f(t, v_2(t)))v_3(t)dt \right| \\ &\leq C\|v_3\|_2\|v_1 - v_2\|_2. \end{aligned}$$

Thus, by previous inequality, F has a Lipschitz continuous gradient. Therefore, the convex conjugate F^* of F , defined by $\lambda \mapsto F^*(\lambda) := \sup_{v \in L^2(0, T)} \langle v, \lambda \rangle - F(v)$ is strongly convex (see for e.g. [28, Lemma 4.1]). By construction of the convex conjugate, F^* is l.s.c. [14, Section 4 Chapter 1], and so is the function $\lambda \mapsto -\inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$. Similarly, for any $\lambda \in L^2(0, T)$, the map $\alpha \mapsto -\inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ is convex and l.s.c.. One deduces that $-\mathcal{W}$ is strongly convex and l.s.c. on $L^2(0, T)$. Thus, Problem (3.6) has a unique solution [14, Chapter 2, Proposition 1.2]. \square

The proposition below is a key result, that enables to show the convergence of a sequence $\{\alpha^k\}_k$ generated by Algorithm 4 defined in the next section, to a solution of Problem (2.5) that belongs to the set \mathfrak{B} , defined in (1.1).

Theorem 3.1. *For any $\lambda \in L^2(0, T)$, $\operatorname{argmin}_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ is not empty. In addition, there exists a selection*

$$\lambda \mapsto \alpha[\lambda] \in \operatorname{argmin}_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda), \quad (3.7)$$

such that the map $L^2(0, T) \ni \lambda \mapsto \alpha[\lambda] \in \mathfrak{A}$ is Lipschitz continuous on any bounded subset of $L^2(0, T)$. Further, for any $\lambda \in L^0(0, T)$, $\alpha[\lambda]$ belongs to \mathfrak{B} .

Theorem 3.1 is proved in Section 3.1. This result allows to prove that the dual function \mathcal{W} is Gâteaux differentiable in $L^2(0, T)$, that is the main argument of the proof of Theorem 3.2. Theorem 3.2 gives the existence of a saddle point of the Lagrangian function \mathcal{L} . This result provides a solution of Problem (2.5).

Theorem 3.2. *There is no duality gap associated with Problem (3.6), i.e.,*

$$\max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda) = \min_{\alpha \in \mathfrak{A}, v \in L^2(0, T)} \bar{J}(\alpha, v).$$

Besides, $\operatorname{argmin}_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \bar{\lambda})$ and $\operatorname{argmin}_{v \in L^2(0, T)} \mathcal{L}_2(v, \bar{\lambda})$ are not empty and, there exists $\bar{\alpha} \in \operatorname{argmin}_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \bar{\lambda})$ and $\bar{v} \in \operatorname{argmin}_{v \in L^2(0, T)} \mathcal{L}_2(v, \bar{\lambda})$, such that $((\bar{\alpha}, \bar{v}), \bar{\lambda})$ is a saddle point of the Lagrangian \mathcal{L} and $\bar{\alpha}$ is a solution of Problem (2.5).

The proof of Theorem 3.2 is given in Section 3.2. The main argument for the proof of Theorem 3.2 is that the map $\lambda \mapsto \mathcal{W}(\lambda)$ is Gâteaux differentiable in $L^2(0, T)$. We show that $\lambda \mapsto \inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ and $\lambda \mapsto \inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ are both differentiable. The second result is more difficult to prove. Different intermediary results are needed. First, properties on the sub problem $\inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ are given in Lemma 3.3. Then, an optimal control problem of a continuity equation is introduced and proved to be equivalent to the sub problem $\inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ in Lemma 3.4. This optimal control problem is solved using results obtained in [28] and provides a solution of the sub problem $\inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ in Lemma 3.5. Theorem 3.1 is a consequence of Lemma 3.5. Finally, Lemma 3.6 shows that the map $\lambda \mapsto \inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ is Gâteaux differentiable in $L^2(0, T)$. These lemmas are stated and proved below.

The next Lemma shows that, for any $\lambda \in L^2(0, T)$, there exists a unique solution $v[\lambda]$ of the sub problem $\inf_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ and that the map $\lambda \mapsto \mathcal{L}_2(v[\lambda], \lambda)$ is Gâteaux differentiable.

Lemma 3.3.

- (i) For any $\lambda \in L^2(0, T)$, there exists a unique $v[\lambda]$ satisfying $v[\lambda] = \operatorname{argmin}_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$.
- (ii) The map $\lambda \mapsto \min_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ admits a Gâteaux derivative $\lambda \mapsto D\mathcal{L}_2(v[\lambda], \lambda)(\cdot)$ in $L^2(0, T)$, which satisfies that, for any $\mu \in L^2(0, T)$:

$$D\mathcal{L}_2(v[\lambda], \lambda)(\mu) = - \int_0^T v[\lambda](\tau) \mu(\tau) d\tau. \quad (3.8)$$

Proof. (i) By Assumption (2) the map $v \mapsto F(v)$ is strictly convex, continuous, proper on $L^2(0, T)$ and by inequality (2.6) this map is also coercive. Thus, one can deduce the existence and uniqueness of $v[\lambda] := \operatorname{argmin}_{v \in L^2(0, T)} \mathcal{L}_2(v, \lambda)$ [14, Chapter 2, Proposition 1.2].

(ii) Since the function F is proper and strictly convex, classical results from convex analysis give that its convex conjugate F^* is differentiable [17, Chapter E, Theorem 4.1.1]. Therefore, the map $\lambda \mapsto \mathcal{L}_2(v[\lambda], \lambda)$ is differentiable on $L^2(0, T)$, with derivative $\lambda \mapsto v[\lambda]$. \square

3.1 Proof of Theorem 3.1

This section is devoted to the proof of Theorem 3.1. The following optimization problem is studied for any $\lambda \in L^2(0, T)$,

$$\inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda). \quad (P_\lambda)$$

A characterization of a solution of this optimization problem is provided in Lemma 3.5. This lemma relies on the results obtained in [27] about the characterization of the solutions of a mean field control problem arising in smart charging. To exploit these results, we first introduce an optimization problem, for a given $\lambda \in L^2(0, T)$, that is equivalent to Problem (P_λ) and that fits with the framework of [27]. To do so, for a given function $\alpha \in \mathfrak{A}$, we introduce the continuity equation (CEQ_α) on $[0, T] \times \mathcal{X}$:

$$\begin{aligned} \partial_t m(t, i, z) + \partial_z (b(i, z) m(t, i, z)) &= - \sum_{j \neq i} (\alpha_j(t, i, z) m(t, i, z) - \alpha_i(t, j, z) m(t, j, z)) \quad (t, i, z) \in (0, T) \times \mathcal{X}, \\ m_i(0, z) &= m_i^0(z) \quad (i, z) \in \mathcal{X}, \end{aligned} \quad (CEQ_\alpha)$$

where $m^0 \in \mathcal{P}(\mathcal{X})$ is given. The next definition gives the characterization of weak solutions of (CEQ_α) .

Definition 3.1. For a given function $\alpha \in \mathfrak{A}$, a distribution m satisfies (CEQ_α) in the weak sense if m is in $C^0([0, T], \mathcal{P}(\mathcal{X}))$ and for any test function $\phi \in C_c^\infty([0, T] \times \mathcal{X})$, we have:

$$\begin{aligned} & \sum_{i \in I} \int_0^1 \phi_i(T, i, z) m(T, i, dz) - \phi(0, i, z) m_i^0(dz) \\ &= \int_0^T \int_0^1 \sum_{i \in I} (\partial_t \phi(t, i, z) + b(i, z) \partial_z \phi(t, i, z)) m(t, i, dz) + \sum_{j \in I, j \neq i} (\phi(t, j, z) - \phi(t, i, z)) \alpha_j(t, i, z) m(t, i, dz) dt. \end{aligned}$$

The following proposition is a classical result and is crucial to introduce a problem that is equivalent to Problem (P_λ) .

Proposition 3.1. Let $m^0 \in \mathcal{P}([0, 1] \times I)$, $\alpha \in C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$ and X^α be a PDMP(b, α). Then, the law m^α of the process X^α is the unique weak solution of (CEQ_α) in the sense of Definition 3.1.

Proof. This result is proved in [8] for controls α that are continuous in space and time independent. The extension of this result to bounded controls that are measurable in time is straightforward. \square

We introduce, for any $\lambda \in L^2(0, T)$, the objective function $\bar{\mathcal{L}}_1(\cdot, \cdot, \lambda)$, defined on $\mathfrak{A} \times C^0([0, T], \mathcal{P}(\mathcal{X}))$, by

$$\begin{aligned} \bar{\mathcal{L}}_1(\alpha, m, \lambda) := & \int_0^T \int_0^1 \sum_{i \in I} (c(t, i, z) + \lambda(t) p(t, i, z)) m^\alpha(t, i, dz) dt \\ & + \int_0^T \int_0^1 \sum_{i, j \in I} L(\alpha_j(t, i, z)) m^\alpha(t, i, dz) dt + \int_0^1 \sum_{i \in I} g(i, z) m^\alpha(T, i, dz), \end{aligned} \quad (3.9)$$

and the optimization problem

$$\begin{aligned} & \inf_{(\alpha, m) \in \mathfrak{A} \times C^0([0, T], \mathcal{P}(\mathcal{X}))} \bar{\mathcal{L}}_1(\alpha, m, \lambda) \\ & \text{where } m \text{ is a weak solution of } (CEQ_\alpha). \end{aligned} \quad (\bar{P}_\lambda)$$

The next lemma establishes the equivalence between Problem (P_λ) and Problem (\bar{P}_λ) .

Lemma 3.4. For any $\lambda \in L^2(0, T)$, Problem (P_λ) is equivalent to Problem (\bar{P}_λ) .

Proof. Let $\alpha \in \mathfrak{A}$ X^α be a PDMP(b, α) and m^α be the law of X^α . By the definitions of \mathcal{L}_1 in (3.3) and $\bar{\mathcal{L}}_1$ in (3.9), one has

$$\mathcal{L}_1(\alpha, \lambda) = \bar{\mathcal{L}}_1(\alpha, m^\alpha, \lambda). \quad (3.10)$$

Further, by Proposition 3.1, m^α is the unique element of $C^0([0, T], \mathcal{P}(\mathcal{X}))$ that is a weak solution of (CEQ_α) . Thus, (α, m^α) is admissible for Problem (\bar{P}_λ) and the problems (P_λ) and (\bar{P}_λ) are equivalent. \square

By the previous lemma, it is straightforward that if Problem (\bar{P}_λ) has a solution (α, m^α) , then α is a solution of Problem (P_λ) . For any $\lambda \in L^2(0, T)$, we define φ^λ , the unique function in \mathfrak{C} (defined in (1.1)) satisfying for any $(t, i, z) \in [0, T] \times \mathcal{X}$,

$$\begin{aligned} \varphi^\lambda(t, i, z) = & \int_t^T \sum_{j \in I, j \neq i} -H\left(\varphi^\lambda(\tau, i, \phi_{t,z}(i, \tau)) - \varphi^\lambda(\tau, j, \phi_{t,z}(i, \tau))\right) + c(\tau, i, \phi_{t,z}(i, \tau)) + p(\tau, i, \phi_{t,z}(i, \tau)) \lambda(\tau) d\tau \\ & + g(i, \phi_{t,z}(i, T)), \end{aligned} \quad (3.11)$$

where H is defined in Assumption 4 and the flow ϕ in (2.1). The existence and uniqueness of φ^λ is a proved in Lemma A.2 in Appendix A and is closely related to [27, Proposition 4.2]. The next lemma provides a solution of Problem (\bar{P}_λ) , based on φ^λ .

Lemma 3.5. For any $\lambda \in L^2(0, T)$, the control $\alpha[\lambda]$ defined for any $(t, i, j, z) \in [0, T] \times I^2 \times [0, 1]$ by

$$\alpha_j[\lambda](t, i, z) := H'(\varphi^\lambda(t, i, z) - \varphi^\lambda(t, j, z)) \quad (3.12)$$

is a solution of Problem (P_λ) .

Proof. By Assumption 4, the map H' is Lipschitz continuous from \mathbb{R} to \mathbb{R}_+ and by Lemma A.2, φ^λ is in \mathfrak{C} . Thus, $\alpha[\lambda]$ belongs $C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$ and by the uniform bound of $\{\|\partial_z \varphi^\lambda\|_\infty\}_{t \in [0, T]}$, there exists a constant $C > 0$ such that, for any $t \in [0, T]$, we have $\|\alpha[\lambda](t, \cdot)\|_\infty + \|\partial_z \alpha[\lambda](t, \cdot)\|_\infty \leq C$. Thus, the control $\alpha[\lambda]$ is in \mathfrak{B} .

By Assumption 4, one can show that $H'(0) = 0$ and therefore, that $\alpha_i[\lambda](\cdot, i, \cdot) = 0$. It comes that $\alpha[\lambda]$ belongs to \mathfrak{A} . By [27, Theorem 2.1], $(\alpha[\lambda], m^{\alpha[\lambda]})$ is a solution of Problem (\bar{P}_λ) . The conclusion follows from Lemma 3.4. \square

Remark 3.1. *The lemma 3.5 states the existence of a solution of an optimal control problem of PDMP without any assumptions on the bounds of the set \mathfrak{A} . Although the boundedness of the space of controls is a common assumption in the literature [2, 9], some recent results enable to relax partially this assumption [18]. Contrary to the literature previously cited, our existence and regularity results do not rely on a dynamic programming approach but is a direct application of the analysis of an optimal control problem of a continuity equation [27].*

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.5, for any $\lambda \in L^2(0, T)$, $\alpha[\lambda]$ defined in (3.12) is a solution of Problem (P_λ) . By the proof of Lemma 3.5, $\alpha[\lambda]$ belongs to \mathfrak{B} . Let $k > 0$ and $\lambda, \mu \in L^2(0, T)$ be such that $\|\lambda\|_2 \leq k$ and $\|\mu\|_2 \leq k$. Then, by the same arguments as in the proof of [27, Lemma 4.5], there exists a constant $K(k, T) > 0$ such that for any $(t, i, j, z) \in [0, T] \times I \times I \times [0, 1]$:

$$|\alpha_j[\lambda](t, i, s) - \alpha_j[\mu](t, i, s)| \leq K(k, T) \|p\|_\infty \|\lambda - \mu\|_2.$$

The Lipschitz continuity of the map $\lambda \mapsto \alpha[\lambda]$ follows from taking the supremum over $[0, T] \times \mathcal{X}$ of the l.h.s. of the previous inequality. \square

Remark 3.2. *Note that rewriting Problem (2.5) w.r.t. to the distribution of the states, as in Lemma 3.4 for the Problem (P_λ) , one can prove Theorem 3.2 by using a change of variable, as it has been done in [27] for a similar problem. To avoid additional notations, we have preferred to prove Theorem 3.2 by means of Theorem 3.1.*

3.2 Proof of Theorem 3.2

The main argument of the proof is that the map $\lambda \mapsto \mathcal{W}(\lambda)$ is Gâteaux differentiable in $L^2(0, T)$. In Lemma 3.3.(ii), it is proved that $\lambda \mapsto \inf_{v \in L^2(0, T)} \mathcal{L}_2(\lambda)$ admits a Gâteaux derivative in $L^2(0, T)$. The next lemma states that the map $\lambda \mapsto \inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ is Gâteaux differentiable in $L^2(0, T)$.

Lemma 3.6. *The map $\lambda \mapsto \inf_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \lambda)$ admits a Gâteaux derivative $\lambda \mapsto D\mathcal{L}_1(\alpha[\lambda], \lambda)(\cdot)$ in $L^2(0, T)$, such that, for any $\mu \in L^2(0, T)$:*

$$D\mathcal{L}_1(\alpha[\lambda], \lambda)(\mu) = \int_0^T \mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda]}) \right] \mu(\tau) d\tau.$$

Proof. For any $\lambda \in L^2(0, T)$, we recall that the definition of the Gâteaux derivative $\lambda \mapsto D\mathcal{L}_1(\alpha[\lambda], \lambda)(\cdot)$ in the direction $\mu \in L^2(0, T)$, if it exists, is:

$$D\mathcal{L}_1(\alpha[\lambda], \lambda)(\mu) := \lim_{h \rightarrow 0} \frac{\mathcal{L}_1(\alpha[\lambda + h\mu], \lambda + h\mu) - \mathcal{L}_1(\alpha[\lambda], \lambda)}{h}.$$

For any $h \in [0, T]$, we have:

$$\begin{aligned} \int_0^T \mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda + h\mu]}) \right] \mu(\tau) d\tau &= \frac{\mathcal{L}_1(\alpha[\lambda + h\mu], \lambda + h\mu) - \mathcal{L}_1(\alpha[\lambda + h\mu], \lambda)}{h} \\ &\leq \frac{\mathcal{L}_1(\alpha[\lambda + h\mu], \lambda + h\mu) - \mathcal{L}_1(\alpha[\lambda], \lambda)}{h} \\ &\leq \frac{\mathcal{L}_1(\alpha[\lambda], \lambda + h\mu) - \mathcal{L}_1(\alpha[\lambda], \lambda)}{h} \\ &= \int_0^T \mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda]}) \right] \mu(\tau) d\tau. \end{aligned} \tag{3.13}$$

Since $\lim_{h \rightarrow 0} \lambda + h\mu = \lambda$ in $L^2(0, T)$, it turns out from Theorem 3.1 that $\lim_{h \rightarrow 0} \alpha[\lambda + h\mu] = \alpha[\lambda] \in \mathfrak{B}$ w.r.t. the norm $\|\cdot\|_\infty$. Since $p \in C^1([0, T] \times \mathcal{X}, \mathbb{R})$, one deduces by Lemma A.1 in Appendix A:

$$\begin{aligned} \lim_{h \rightarrow 0} \int_0^T \mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda + h\mu]}) \right] \mu(\tau) d\tau &= \lim_{h \rightarrow 0} \int_0^T \sum_{i \in I} \int_0^1 p(\tau, i, z) m^{\alpha[\lambda + h\mu]}(\tau, i, dz) \mu(\tau) d\tau \\ &= \int_0^T \sum_{i \in I} \int_0^1 p(\tau, i, z) m^{\alpha[\lambda]}(\tau, i, dz) \mu(\tau) d\tau \\ &= \int_0^T \mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda]}) \right] \mu(\tau) d\tau. \end{aligned}$$

The conclusion follows by taking the limit $h \rightarrow 0$ in (3.13). \square

We now turn to the proof of Theorem 3.2.

Proof of Theorem 3.2. By Lemma 3.3.(ii) and Lemma 3.6, \mathcal{W} is Gâteaux differentiable in $L^2(0, T)$. For any $\lambda \in L^2(0, T)$, the Gâteaux differential of \mathcal{W} at λ is denoted by $D\mathcal{W}(\lambda)(\cdot)$ and is given in the direction $\mu \in L^2(0, T)$ by:

$$D\mathcal{W}(\lambda)(\mu) = \int_0^T \left(\mathbb{E} \left[p(\tau, X_\tau^{\alpha[\lambda]}) \right] - v[\lambda](\tau) \right) \mu(\tau) d\tau. \quad (3.14)$$

By the definition of $\bar{\lambda}$ given in Lemma 3.2, it results that $0 \in \partial(-\mathcal{W}(\bar{\lambda}))$. Since $-\mathcal{W}$ is convex and Gâteaux differentiable, its sub-differential is reduced to a singleton and is equal to $-D\mathcal{W}(\bar{\lambda})$. One has, for any $\mu \in L^2(0, T)$, by the previous equality:

$$\int_0^T \left(\mathbb{E} \left[p(\tau, X_\tau^{\alpha[\bar{\lambda}]}) \right] - v[\bar{\lambda}](\tau) \right) \mu(\tau) d\tau = \langle \partial(-\mathcal{W}(\bar{\lambda})), \mu \rangle = 0.$$

Therefore, one has for a.e. $t \in [0, T]$:

$$\mathbb{E} \left[p(\tau, X_t^{\alpha[\bar{\lambda}]}) \right] - v[\bar{\lambda}](t) = 0. \quad (3.15)$$

By the previous equality, for any $\alpha \in \mathfrak{A}$ and $v \in L^2(0, T)$, satisfying $\mathbb{E}[p(t, X_t^\alpha)] - v(t) = 0$ a.e on $[0, T]$, one has:

$$\bar{J}(\alpha[\bar{\lambda}], v[\bar{\lambda}]) = \mathcal{L}(\alpha[\bar{\lambda}], v[\bar{\lambda}], \bar{\lambda}) \leq \mathcal{L}(\alpha, v, \bar{\lambda}) = \bar{J}(\alpha, v). \quad (3.16)$$

By Lemma 3.2 and the definition of \mathcal{W} in (3.5)

$$\max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda) = \mathcal{W}(\bar{\lambda}) = \mathcal{L}(\alpha[\bar{\lambda}], v[\bar{\lambda}], \bar{\lambda}) = \bar{J}(\alpha[\bar{\lambda}], v[\bar{\lambda}]).$$

Thus, by inequality (3.16) and the previous equality, one has

$$\max_{\lambda \in L^2(0, T)} \mathcal{W}(\lambda) = \min_{\alpha \in \mathfrak{A}, v \in L^2(0, T)} \bar{J}(\alpha, v).$$

By Lemma 3.3.(i), $\operatorname{argmin}_{v \in L^2(0, T)} \mathcal{L}_2(v, \bar{\lambda})$ is not empty and by Theorem 3.1 $\operatorname{argmin}_{\alpha \in \mathfrak{A}} \mathcal{L}_1(\alpha, \bar{\lambda})$ is also not empty. Setting

$\bar{\alpha} := \alpha[\bar{\lambda}]$ and $\bar{v} := v[\bar{\lambda}]$, using (3.15) and (3.16), one has for any $(\alpha, v, \lambda) \in \mathfrak{A} \times L^2(0, T) \times L^2(0, T)$

$$\mathcal{L}(\bar{\alpha}, \bar{v}, \lambda) \leq \mathcal{L}(\bar{\alpha}, \bar{v}, \bar{\lambda}) \leq \mathcal{L}(\alpha, v, \bar{\lambda}),$$

and thus, $((\bar{\alpha}, \bar{v}), \bar{\lambda})$ is a saddle point of \mathcal{L} . Finally, by Lemma 3.1 and inequality (3.16), $\bar{\alpha}$ is a solution of Problem (2.5). \square

4 Stochastic Uzawa Algorithm

While the existence of a unique solution $\bar{\lambda}$ of the dual problem (3.6) is established in Lemma 3.2, we propose in this section an iterative algorithm converging to $\bar{\lambda}$. The convergence of the algorithm relies on the same arguments as those developed for the analysis of the gradient descent algorithm in Hilbert space [16]. This algorithm is an adaptation of the Uzawa Algorithm [1] in stochastic optimization settings. The sub gradient of the opposite of

Algorithm 1 Stochastic Uzawa

- 1: Initialization $\lambda^0 \in L^\infty(0, T)$, set $\{\rho_k\}$ and $M \in \mathbb{N}^*$
 - 2: $k \leftarrow 0$.
 - 3: **for** $k = 0, 1, \dots$ **do**
 - 4: $v^k \leftarrow v[\lambda^k]$ where $v[\lambda^k]$ is defined in Lemma 3.3.
 - 5: $\alpha^k \leftarrow \alpha[\lambda^k]$ where $\alpha[\lambda^k]$ is defined in (3.12).
 - 6: Generate M independent states realizations $(X^{1,\alpha^k}, \dots, X^{M,\alpha^k})$ independent of previous iterations simulations.
 - 7: $U^{k+1} \leftarrow v^k - \frac{1}{M} \sum_{j=1}^M p(\cdot, X^{j,\alpha^k})$.
 - 8: $\lambda^{k+1} \leftarrow \lambda^k + \rho_k U^{k+1}$.
-

the dual function \mathcal{W} is estimated by Monte Carlo simulations. Convexity assumptions introduced in [28] are not satisfied here.

We consider a sequence $\{\rho_k\}_k$ satisfying $\sum_{k=0}^{\infty} \rho_k = \infty$ and $\sum_{k=0}^{\infty} \rho_k^2 < \infty$ to ensure the convergence. Typically, consider the sequence defined by $\rho_k = \frac{a}{b+k}$ with $a > 0$ and $b > 0$ chosen empirically to accelerate the convergence.

The next Lemma shows that it is possible at any iteration k to compute $\alpha[\lambda^k]$ defined in line 5 of the Algorithm 1.

Lemma 4.1. *For any $k \in \mathbb{N}$, $\lambda^k \in L^\infty(0, T)$ a.s..*

Proof. The proof is done by induction. For $k = 0$, λ^0 is initialized as an element of $L^\infty(0, T)$. Assume now that $\lambda^k \in L^\infty(0, T)$ for an integer k . Since $p \in C^1([0, T] \times \mathcal{X}, \mathbb{R})$, one has that a.s. $\frac{1}{M} \sum_{j=1}^M p(\cdot, X^{j,\alpha^k}) \in L^\infty(0, T)$. We show now that $v[\lambda^k] \in L^\infty(0, T)$. Since f is l.s.c., it is a normal integrand. By the exchange property [25, Theorem 14.60], one has:

$$\inf_{v \in L^2(0, T)} \int_0^T f(t, v(t)) - v(t)\lambda^k(t) dt = \int_0^T \inf_{v \in \mathbb{R}} f(t, v) - v\lambda^k(t) dt.$$

and for a.e. $t \in [0, T]$:

$$v[\lambda^k](t) \in \operatorname{argmin}_{v \in \mathbb{R}} f(t, v) - v\lambda^k(t). \quad (4.1)$$

For any $t \in [0, T]$, $f^*(t, \cdot)$ denotes the convex conjugate of $f(t, \cdot)$. Since for any $t \in [0, T]$, $f(t, \cdot)$ is strictly convex and by inequality (2.6) has at most quadratic growth, $f^*(t, \cdot)$ is differentiable, with linear growth [28, Lemma A.1]. Then, there exists a constant $C > 0$, such that for any $t \in [0, T]$ one has:

$$|\nabla f^*(t, \lambda(t))| \leq C(|\lambda(t)| + 1),$$

where $\nabla f^*(t, \cdot)$ is the derivative of $z \mapsto f^*(t, z)$. By (4.1), one has, for a.e. $t \in [0, T]$, that $v[\lambda^k](t) = \nabla f^*(t, \lambda(t))$. By previous inequality, one deduces that

$$\|v[\lambda^k]\|_\infty \leq C(\|\lambda^k\|_\infty + 1). \quad (4.2)$$

The conclusion follows from the definition of λ^{k+1} at line 8 of Algorithm 1, previous inequality and the fact that a.s. $\frac{1}{M} \sum_{j=1}^M p(\cdot, X^{j,\alpha^k}) \in L^\infty(0, T)$. \square

The next Lemma provides an estimate on the norm of the gradient of \mathcal{W} , that is crucial to show the convergence of Algorithm 1.

Lemma 4.2. *There exists $C > 0$ such that the sequence $\{U^{k+1}\}_k$ generated by Algorithm 1 (line 7) satisfies, for any $k \in \mathbb{N}$,*

$$\mathbb{E}[\|U^{k+1}\|_2] \leq C(\|\lambda^k\|_2 + 1).$$

Proof. By Assumption 2, p is in $C^1([0, T] \times \mathcal{X}, \mathbb{R})$ thus, there exists $C > 0$, such that, for any $k \in \mathbb{N}$ and $j \in \{1, \dots, M\}$:

$$\mathbb{E}[\|p(\cdot, X^{j, \alpha^k})\|_2] \leq C.$$

The conclusion follows from previous inequality, inequality (4.2) in the proof of Lemma 4.1 and the definition of U^{k+1} at line 7 of Algorithm 1. \square

The convergence of Algorithm 1 is stated in the next lemma.

Lemma 4.3. *The sequence $\{\lambda^k\}_k$, generated by Algorithm 1, converges to $\bar{\lambda}$ a.s. in $L^2(0, T)$.*

Proof. Since \mathcal{W} is convex and Gateau differentiable (proof of Theorem 3.2), by Lemma 4.2, the proof is a simple adaptation of the one in [16, Theorem 3.6], as it has been done in [28]. \square

Before to give the proof of the main result of this paper, we need the following regularity result on the function J .

Lemma 4.4. *The map J defined in (2.4) is continuous on \mathfrak{B} , w.r.t. the norm $\|\cdot\|_\infty$.*

Proof. Let $\alpha \in \mathfrak{B}$, X^α be a PDMP(b, α) and m^α be the law of X^α . By the definition of m^α and of the map J in (2.4), J be rewritten as:

$$\begin{aligned} J(\alpha) = & \int_0^T f\left(t, \sum_{i \in I} \int_0^1 p(t, i, z) m^\alpha(t, i, dz)\right) dt + \int_0^T \int_0^1 \sum_{i \in I} c(t, i, z) m^\alpha(t, i, dz) dt \\ & + \int_0^T \int_0^1 \sum_{i, j \in I} L(\alpha_j(t, i, z)) m^\alpha(t, i, dz) dt + \int_0^1 \sum_{i \in I} g(i, z) m^\alpha(T, i, dz). \end{aligned}$$

By the regularity of the functions p, c, g and L in Assumptions 2-4 and the continuity of the map $\alpha \mapsto m^\alpha$ on \mathfrak{B} stated in Lemma A.1 in Appendix A, one obtains the continuity of J over \mathfrak{B} w.r.t. the norm $\|\cdot\|_\infty$. \square

We can now prove Theorem 2.1.

Proof of Theorem 2.1. The existence of a solution to Problem 2.5 is a direct consequence of Theorem 3.2. Indeed, the control $\alpha[\bar{\lambda}]$, where by Lemma 3.2 $\bar{\lambda}$ is the solution of Problem (3.6), is a solution of Problem (2.5).

(i) By Lemma 4.3, the sequence $\{\lambda^k\}_k$, generated by Algorithm 1, converges to $\bar{\lambda}$ a.s. in $L^2(0, T)$. Then, by Theorem 3.1, the sequence $\{\alpha[\lambda^k]\}_k$ converges to $\alpha[\bar{\lambda}]$ a.s. w.r.t. the norm $\|\cdot\|_\infty$. Since $\alpha[\bar{\lambda}]$ is a solution of Problem 2.5 and that by definition $\alpha^k = \alpha[\lambda^k] \in \mathfrak{B}$ for any $k \in \mathbb{N}$, the conclusion follows.

(ii) Since the sequence $\{\alpha^k\}_k$ converges to a solution of Problem 2.5 a.s. w.r.t. the norm $\|\cdot\|_\infty$ and by Lemma 4.4 the function J is continuous on \mathfrak{B} , one deduces that the sequence $\{J(\alpha^k)\}_{k \in \mathbb{N}}$ converges a.s. to $\min_{\alpha \in \mathfrak{A}} J(\alpha)$. \square

5 Simulations

This section illustrates the results with an example of smart charging. In this use case, the discrete variable taking values in I represents the charging mode and the continuous variable taking value in $[0, 1]$ represents the state of charge (SoC) of electrical vehicles relatively to the maximum energy capacity of the battery. The quantities are dimensionless.

5.1 Definition of the use Case

We consider a large fleet of EVs controlled by a central planner during their charging period $[0, T]$. The goal of the central planner is to provide ancillary services to the transmission grid by controlling the aggregate consumption profile of the fleet on the time horizon $[0, T]$. More specifically, the central planner aims at making the consumption profile of the fleet to be close to a given profile $r = (r_t)_{0 \leq t \leq T}$, supposed to be known on the whole period $[0, T]$. In our simulation we compare two situations. In the first case known as V1G, it is assumed that the vehicle batteries can only draw electricity from the grid, then the set of charging modes is $I = \{0, 1\}$, where 0 stands for idle mode and 1 for charging. In the second case known as V2G, it is assumed that the batteries can either draw or inject electricity into the grid, then the charging modes are $I = \{-1, 0, 1\}$, where -1 corresponds to injection mode. In each situation, each mode $i \in I$ is characterized by its charging rate $b(i, \cdot)$ such that

- $i = -1$ corresponds to a V2G mode with $b(-1, z) = -a$ for $z \geq 0.25$ and $b(-1, z) = -az$ for $z \leq 0.25$, with $a = \frac{2}{5}$.
- $i = 0$ corresponds to a non-charging mode with $b(0, z) = 0$,
- $i = 1$ corresponds to a charging mode with $b(1, z) = a$ for $z \leq 0.75$ and $b(1, z) = a(1 - z)$ for $z \geq 0.75$, with $a = \frac{2}{5}$.

The charging rate b of the batteries has been designed based on industrial data [15]. The state of each EV, controlled by α , is represented by a PDMP $X^\alpha = (Y^\alpha, Z^\alpha)$ with Y_t^α , the charging mode and Z_t^α , the SoC of the vehicle at time t . We assume that the power consumption of a vehicle at state $(i, z) \in \mathcal{X}$ is equal to its instantaneous charging rate $b(i, z)$. At time $t = 0$, Y_0^α equals 0 or 1 with probability 1/2, while $Z_0^\alpha = \min(1, |z_0| + 0.15)$ where $z_0 \sim \mathcal{N}(0, 0.2)$.

To ensure customer satisfaction, the final cost $g(i, z) := 30 \times (1 - e^{5(z-0.75)})^+$ penalizes vehicles with a final SoC z lower than 75%. As previously stated, high values of α are penalized through the cost $\alpha^2/2$ in order to avoid high frequencies of jumps damaging the batteries. For any $t \in [0, T]$,

$$\begin{cases} c(t, i, z) & = & 0 \\ p(t, i, z) & := & b(i, z) \\ l(a) & := & \frac{a^2}{2} \\ F(v) & := & \kappa \int_0^T (v(t) - r(t))^2 dt \end{cases}$$

with $\kappa = 1000$ in order to ensure that the overall consumption is close to the profile r .

The target profile r is a slight modification of the nominal behaviour of the fleet of EVs (without control), with the same energy consumed over the fixed time horizon $[0, T]$. The nominal behaviour corresponds to the situation where the EVs are not required to fit the target profile and seek only to satisfy their own comfort (namely $F(v) = 0$). The idea is to define a realistic target profile so that it is possible for the fleet to follow this profile while satisfying their charging needs. We consider a population of $N = 10^5$ EVs. The optimal control of Problem (2.5) is computed using Algorithm 1.

5.2 Algorithm parameters

For the implementation of the algorithm, the line 5 is computed by discretization of the Hamilton Jacobi-Bellman-Equation III.5 associated to $\alpha[\lambda]$ defined in III.6 with $N_t = 1000$ regular points of time and $N_s = 400$ regular points of SoC. The Lagrangian multipliers λ^k are obtained after $k = 1000$ iterations and with $M = 1000$ realizations and are displayed in Figure 4. The initial multiplier is $\lambda^0 = 0$ and the stepsize sequence is such that $\rho_k = \min(30, \frac{500}{1+k})$. For each multiplier λ , the associated optimal strategy $\alpha[\lambda]$ is computed.

5.3 Results

Four periods on Figure 3 are distinguished, depending on whether the target profile is above or below the nominal consumption. The reference profile is displayed in green. EVs are encouraged to consume more than the nominal consumption during periods P1 and P3, and to consume less during periods P2 and P4. First, we observe that both V1G (blue line) and V2G (red line) consumptions are close to the profile r (green line). The main difference lies in P4, where the possibility for the EVs to perform V2G allows the fleet consumption to stay closer to the target profile. As expected, the profile is better tracked with the V2G mode, due to the additional degree of liberty. On Figure 6 one can observe that V2G is used during periods P2 and P4, when the profile r is lower than the nominal consumption. Ten trajectories of SoC, generated in the V2G use case, are represented in Figure 1. The green line is an example of a PDMP switching from mode 0 to mode 1 around $t = 0.2$, from mode 1 to mode -1 around $t = 2.6$ and then from mode -1 to mode 1 around $t = 4$. The initial and final distribution of the SoC of the fleet are displayed on Figures 7 and 8. While the initial distribution is the same for both scenarios, one can observe that their final distribution are very closed, and that very few EVs have a SoC lower than 0.75% in both cases. Thus, the comfort of each agent is weakly impacted by the V2G mode.

To conclude, the interest of adding V2G is confirmed by the comparison of the optimal cost of each scenario (Figure 2). These optimal costs are decomposed into three parts: the switching cost ($\frac{1}{2} \int_0^T \sum_{j \neq I} \alpha_j(t, X_t^{i, \alpha})^2 dt$), the

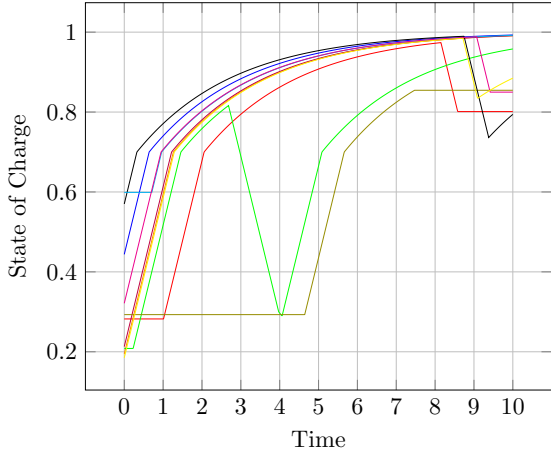


Figure 1: Representation of the SoC of 10 PDMP.

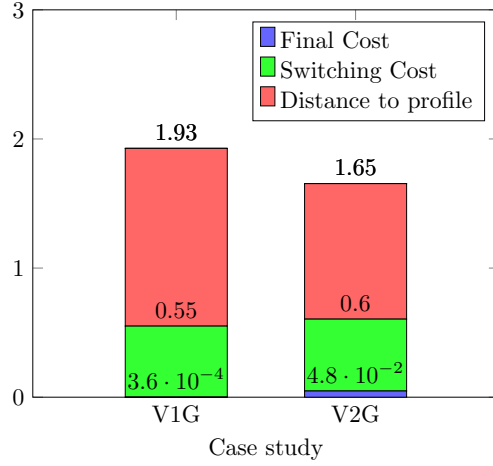


Figure 2: Decomposition of the cost function.

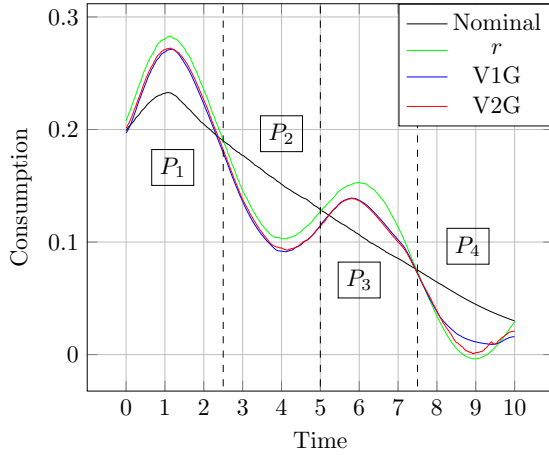


Figure 3: Consumption for V1G and V2G cases compared to the profile and nominal.

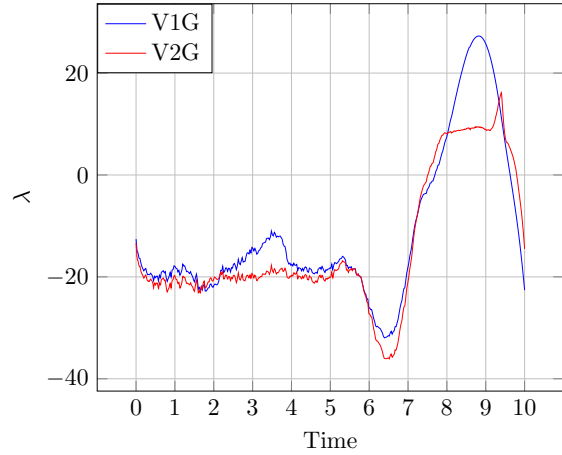


Figure 4: Optimal Lagrangian multiplier λ obtained.

final cost (g) and the distance to profile (F). The low value of the final cost (3.6×10^{-4} in the V1G use case and 4.8×10^{-2} in the V2G one) shows that, despite the tracking objective, EVs have enough battery at the end of the period, as previously stated. Moreover, this performance is achieved without significant loss in switching costs, thus preserving fleet batteries.

A Appendix

A.1 Equivalent definition for Piecewise Deterministic Markov Processes

This section aims at showing the connection between the definition of the PDMP in Section 2, that is based on [13], and the standard definition of a PDMP introduced in [10].

We consider the same settings as in Section 2, i.e. the finite time horizon $[0, T]$ with $T > 0$, the state space $\mathcal{X} := I \times [0, 1]$, where I is finite. Considering a control $\alpha \in \mathfrak{A}$, we define the measurable function $\lambda^\alpha : [0, T] \times \mathcal{X} \mapsto \mathbb{R}_+$:

$$\lambda^\alpha(t, i, z) := \sum_{j \in I} \alpha_j(t, i, z).$$

By the definition of \mathfrak{A} in (1.1), the function λ^α is non negative, continuous and bounded. Let \mathcal{M} be the Borel set of \mathcal{X} . For any $\alpha \in \mathfrak{A}$, we consider the transition measure $Q : \mathcal{M} \times [0, T] \times \mathcal{X} \mapsto [0, 1]$ defined for any

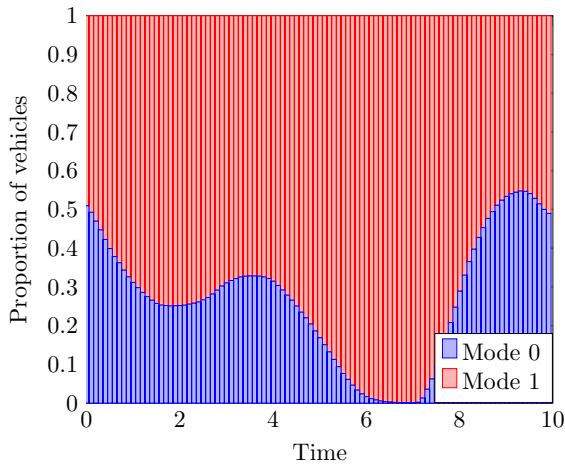


Figure 5: Evolution of the proportion of vehicles for V1G.

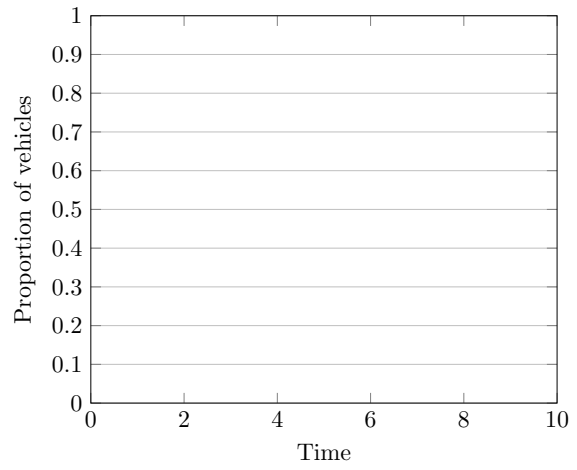


Figure 6: Evolution of the proportion of vehicles for V2G.

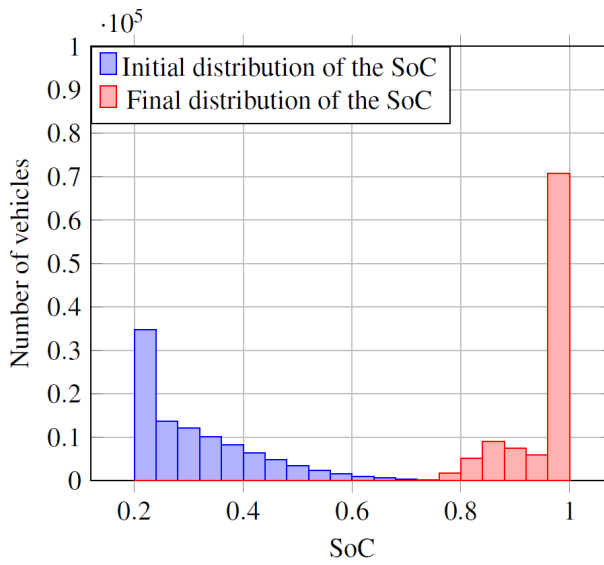


Figure 7: Initial and Final SoC in the V1G case

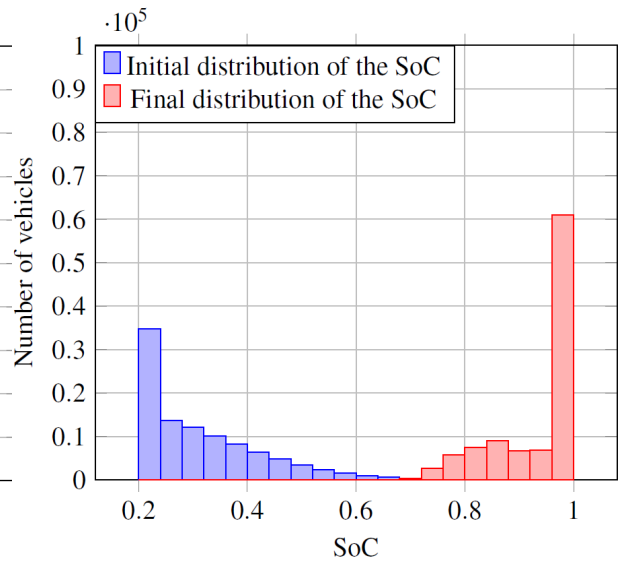


Figure 8: Initial and Final SoC in the V2G case.

$(A, t, i, z) \in \mathcal{M} \times [0, T] \times \mathcal{X}$ by

$$Q^\alpha(A; t, i, z) := \begin{cases} \frac{1}{\lambda^\alpha(t, i, z)} \sum_{j \in I} \alpha_j(t, i, z) \delta_{(j, z)}(A) & \text{if } \lambda^\alpha(t, i, z) > 0, \\ 0 & \text{else.} \end{cases} \quad (\text{A.1})$$

The PDMP(b, α) $X^\alpha = (Y^\alpha, Z^\alpha)$ defined in Section 2 by (2.2) can alternatively be described as follows [13]:

- the initial state (Y_0^α, Z_0^α) follows the law $m^0 \in \mathcal{P}(\mathcal{X})$ and we set $T_0^\alpha = 0$.
- Knowing T_k^α and $X_{T_k^\alpha}^\alpha = (Y_{T_k^\alpha}^\alpha, Z_{T_k^\alpha}^\alpha)$, the next inter-jump time Δ_{k+1}^α is a random variable, whose law \mathbb{P} is given, for any $t \in \mathbb{R}_+$ by:

$$\mathbb{P}[\Delta_{k+1}^\alpha > t] := \begin{cases} \exp\left(-\int_{T_k^\alpha}^{T_k^\alpha+t} \lambda^\alpha(\tau, Y_{T_k^\alpha}^\alpha, \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, \tau)) d\tau\right) & \text{if } t \in [0, T - T_k^\alpha] \\ 0 & \text{else,} \end{cases}$$

where the flow ϕ is the unique solution of the ordinary differential equation (2.1).

- Once Δ_{k+1}^α is determined, the next jumping time T_{k+1}^α is given by $T_{k+1}^\alpha := T_k^\alpha + \Delta_{k+1}^\alpha$.
- We consider a realization of the couple of variables $(\tilde{Y}_{k+1}^\alpha, \tilde{Z}_{k+1}^\alpha)$ with distribution $Q(\cdot; T_{k+1}^\alpha, Y_{T_k^\alpha}^\alpha, \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, T_{k+1}^\alpha))$. Note that due to the definition of Q in (A.1), we have $\tilde{Z}_{k+1}^\alpha = \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, T_{k+1}^\alpha)$ a.s..
- The trajectory of X^α for $t \in [T_k^\alpha, T_{k+1}^\alpha]$ is given by

$$X_t^\alpha = (Y_t^\alpha, Z_t^\alpha) = \begin{cases} (Y_{T_k^\alpha}^\alpha, \phi_{T_k^\alpha, Z_{T_k^\alpha}^\alpha}(Y_{T_k^\alpha}^\alpha, t)) & \text{if } t \in [T_k^\alpha, T_{k+1}^\alpha) \\ (\tilde{Y}_{k+1}^\alpha, \tilde{Z}_{k+1}^\alpha) & \text{if } t = T_{k+1}^\alpha. \end{cases}$$

A.2 Continuity property of the law of a PDMP w.r.t. the jump intensity

In this section, we provide a result stating the continuity of the map $\alpha \mapsto m^\alpha$, where m^α is the distribution of the PDMP(b, α) X^α , from \mathfrak{B} to $C^0([0, T], \mathcal{P}(\mathcal{X}))$, where \mathfrak{B} is defined in (1.1).

Lemma A.1. *The map $\alpha \mapsto m^\alpha$ is continuous from \mathfrak{B} , endowed with the norm $\|\cdot\|_\infty$, to $C^0([0, T], \mathcal{P}(\mathcal{X}))$, endowed with the distance $\mathbb{W}_{\|\cdot\|_\infty}$.*

Proof. Let $\psi \in C^1(\mathcal{X})$ and let $\varphi \in C^1([0, T] \times \mathcal{X})$ be the classical solution of the PDE:

$$\begin{aligned} \partial_\tau \varphi(\tau, i, z) + b(i, z) \partial_s \varphi(\tau, i, z) &= 0 & (\tau, i, z) \in [0, t] \times \mathcal{X} \\ \varphi(t, i, z) &= \psi(i, z) & (i, z) \in \mathcal{X}. \end{aligned} \quad (\text{A.2})$$

One has $\|\varphi\|_\infty = \|\psi\|_\infty$, and denoting by γ_ψ the Lipschitz constant of ψ , for any $t \in [0, T]$ $\varphi(t, \cdot, \cdot)$ is Lipschitz continuous with Lipschitz constant $\gamma_\psi e^{T\|b'\|_\infty}$.

Let $\alpha, \bar{\alpha} \in \mathfrak{B}$. Since X^α is a PDMP(b, α) and $X^{\bar{\alpha}}$ is a PDMP($b, \bar{\alpha}$), by Proposition 3.1, m^α and $m^{\bar{\alpha}}$ are both weak solutions of the continuity equation (CEQ_α) associated respectively with α and $\bar{\alpha}$. Then, by equality (A.2),

one has:

$$\begin{aligned}
& \int_0^1 \sum_{i \in I} \psi(i, z) (m^\alpha(t, i, dz) - m^{\bar{\alpha}}(t, i, dz)) \\
&= \sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \left(-\alpha_j(\tau, i, z) m^\alpha(\tau, i, dz) + \alpha_i(\tau, j, z) m^\alpha(\tau, j, dz) \right) d\tau \\
&\quad + \sum_{i \in I, j \neq i} \int_0^t \int_0^1 \sum_{j \in I} \varphi(i, \tau, z) \left(\bar{\alpha}_j(\tau, i, z) m^{\bar{\alpha}}(\tau, i, dz) - \bar{\alpha}_i(\tau, j, z) m^{\bar{\alpha}}(\tau, j, dz) \right) d\tau \\
&= \sum_{i \in I, j \neq i} \int_0^t \int_0^1 \sum_{j \in I} \varphi(i, \tau, z) \alpha_j(\tau, i, z) \left(m^{\bar{\alpha}}(\tau, i, dz) - m^\alpha(\tau, i, dz) \right) d\tau \\
&\quad + \sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \alpha_j(\tau, i, z) \left(m^\alpha(\tau, j, dz) - m^{\bar{\alpha}}(\tau, j, dz) \right) d\tau \\
&\quad + \sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \left(\bar{\alpha}_j(\tau, i, z) - \alpha_j(\tau, i, z) \right) m^{\bar{\alpha}}(\tau, i, dz) d\tau \\
&\quad + \sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \left(\alpha_i(\tau, j, z) - \bar{\alpha}_i(\tau, j, z) \right) m^{\bar{\alpha}}(\tau, j, dz) d\tau
\end{aligned} \tag{A.3}$$

Using that $\varphi(t, \cdot, \cdot)$ is Lipschitz continuous with Lipschitz constant $\gamma_\psi e^{T\|b'\|_\infty}$ and that $\partial_z \alpha \in L^\infty([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$, there exists $L > 0$ such that:

$$\sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \alpha_j(\tau, i, z) \left(m^{\bar{\alpha}}(\tau, i, dz) - m^\alpha(\tau, i, dz) \right) d\tau \leq L \gamma_\psi e^{T\|b'\|_\infty} \int_0^T \mathcal{W}(m^\alpha(\tau), m^{\bar{\alpha}}(\tau)) d\tau, \tag{A.4}$$

$$\sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \alpha_j(\tau, i, z) \left(m^\alpha(\tau, j, dz) - m^{\bar{\alpha}}(\tau, j, dz) \right) d\tau \leq L \gamma_\psi e^{T\|b'\|_\infty} \int_0^T \mathcal{W}(m^\alpha(\tau), m^{\bar{\alpha}}(\tau)) d\tau. \tag{A.5}$$

Since $\alpha, \bar{\alpha} \in C^0([0, T] \times \mathcal{X}, \mathbb{R}_+^d)$ and that $\|\varphi\|_\infty \leq \|\psi\|_\infty$, one has:

$$\sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \left(\bar{\alpha}_j(\tau, i, z) - \alpha_j(\tau, i, z) \right) m^{\bar{\alpha}}(\tau, i, dz) d\tau \leq |I| \|\psi\|_\infty \|\alpha - \bar{\alpha}\|_\infty, \tag{A.6}$$

$$\sum_{i \in I} \int_0^t \int_0^1 \sum_{j \in I, j \neq i} \varphi(i, \tau, z) \left(\alpha_i(\tau, j, z) - \bar{\alpha}_i(\tau, j, z) \right) m^{\bar{\alpha}}(\tau, j, dz) d\tau \leq |I| \|\psi\|_\infty \|\alpha - \bar{\alpha}\|_\infty. \tag{A.7}$$

By inequalities (A.4)-(A.7), inequality (A.3) becomes:

$$\int_0^1 \sum_{i \in I} \psi(i, z) (m^\alpha(t, i, dz) - m^{\bar{\alpha}}(t, i, dz)) \leq 2|I| \|\psi\|_\infty \|\alpha - \bar{\alpha}\|_\infty + 2L \gamma_\psi e^{T\|b'\|_\infty} \int_0^t \mathcal{W}(m^\alpha(\tau), m^{\bar{\alpha}}(\tau)) d\tau.$$

Since ψ is arbitrary and in $C^1(\mathcal{X})$, previous inequality becomes, for any $t \in [0, T]$,

$$\mathbb{W}(m^\alpha(t), m^{\bar{\alpha}}(t)) \leq 2|I|^2 \|\psi\|_\infty \|\alpha - \bar{\alpha}\|_\infty + 2Le^{T\|b'\|_\infty} \int_0^t \mathcal{W}(m^\alpha(\tau), m^{\bar{\alpha}}(\tau)) d\tau.$$

By Gronwall inequality, one deduces that there exists a constant $C > 0$ such that, for any $t \in [0, T]$,

$$\mathbb{W}(m^\alpha(t), m^{\bar{\alpha}}(t)) \leq L \|\alpha - \bar{\alpha}\|_\infty.$$

Thus, $\mathbb{W}_{\|\cdot\|_\infty}(m^\alpha, m^{\bar{\alpha}}) \leq L \|\alpha - \bar{\alpha}\|_\infty$ and the continuity of $\alpha \mapsto m^\alpha$ follows. \square

A.3 Characterization of an optimal control

This section aims at giving an existence and uniqueness result of the equation (3.11) for any $\lambda \in L^2(0, T)$.

Lemma A.2. *For any $\lambda \in L^2(0, T)$, there exists a unique $\varphi^\lambda \in \mathfrak{C}$ satisfying (3.11).*

Proof. The proof follows the path of the proof of [27, Proposition 4.1]. The main difference with the framework of the proof of [27, Proposition 4.1] is the presence of the term $\lambda(\tau)p(\tau, i, \phi_{t,z}(i, \tau))$ in (3.11). Let $C_0, C_1, \kappa, M > 0$ be constant defined below. We consider the space

$$\mathfrak{D} := \left\{ f \in \mathfrak{C} \mid \|f\|_\infty \leq M + 1 \text{ and } \|\partial_z f(t, \cdot)\|_\infty \leq C_0 e^{C_1(T-t)} \forall t \in [0, T] \right\}.$$

where \mathfrak{C} is defined in (1.1) and is endowed with the norm $\|\cdot\|_{0,1}^\kappa$ defined by:

$$\|f\|_{0,1}^\kappa := \|f\|_\infty^\kappa + \|\partial_s f\|_\infty^\kappa.$$

The space $(\mathfrak{C}, \|\cdot\|_{0,1}^\kappa)$ is a Banach space. We consider the map Γ^λ , defined on \mathfrak{C} , by

$$\begin{aligned} \Gamma^\lambda(\phi)(t, i, z) := & \int_t^T \sum_{j \in I, j \neq i} -H\left(\varphi(\tau, i, \phi_{t,z}(i, \tau)) - \varphi(\tau, j, \phi_{t,z}(i, \tau))\right) + c(\tau, i, \phi_{t,z}(i, \tau)) d\tau \\ & + \int_t^T \lambda(\tau)p(\tau, i, \phi_{t,z}(i, \tau)) d\tau + g(i, \phi_{t,z}(i, T)). \end{aligned} \quad (\text{A.8})$$

Then, by considering a suitable choice of the constants:

- $M > 0$, as in the proof of [27, Lemma 4.2] with an additional dependence on $\|\lambda\|_2 \|p\|_\infty$,
- C_0 and C_1 as in the proof of [27, Lemma 4.3] with an additional dependence on $\|\lambda\|_2 \|\partial_z p\|_\infty$,
- κ as in the proof of [27, Lemma 4.3],

one can show the existence of a fixed point of the map Γ^λ on \mathfrak{D} , following the proof of [27, Lemma 4.4]. Then, the existence and uniqueness of a solution φ^λ in \mathfrak{D} of equation (3.11) is obtained by adapting the proof of [27, Proposition 4.1]. \square

References

- [1] Kenneth Joseph Arrow, Leonid Hurwicz, Hirofumi Uzawa, and Hollis Burnley Chenery. *Studies in linear and non-linear programming*, volume 2. Stanford University Press, 1958.
- [2] Nicole Bäuerle and Ulrich Rieder. Optimal control of Piecewise Deterministic Markov Processes with finite time horizon. *Modern trends in controlled stochastic processes: theory and applications*, pages 123–143, 2010.
- [3] Haim Brezis and Haim Brézis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011.
- [4] Ana Bušić and Sean Meyn. Distributed control of thermostatically controlled loads: Kullback-Leibler optimal control in continuous time. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 7258–7265. IEEE, 2019.
- [5] René Carmona and François Delarue. Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics. *The Annals of Probability*, 43(5):2647–2700, 2015.
- [6] Guido Carnevale, Andrea Camisa, and Giuseppe Notarstefano. Distributed online aggregative optimization for dynamic multi-robot coordination. *IEEE Transactions on Automatic Control*, 2022.
- [7] Pierre Carpentier, J-Ph Chancelier, Vincent Leclère, and François Pacaud. Stochastic decomposition applied to large-scale hydro valleys management. *European Journal of Operational Research*, 270(3):1086–1098, 2018.
- [8] Christiane Coccozza-Thivent, Robert Eymard, Sophie Mercier, and Michel Roussignol. Characterization of the marginal distributions of Markov processes used in dynamic reliability. *Journal of Applied Mathematics and Stochastic Analysis*, 2006.
- [9] Oswaldo LV Costa, François Dufour, and Alexei B Piunovskiy. Constrained and unconstrained optimal discounted control of Piecewise Deterministic Markov Processes. *SIAM Journal on Control and Optimization*, 54(3):1444–1474, 2016.

- [10] Mark HA Davis. Piecewise-Deterministic Markov Processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):353–376, 1984.
- [11] Benoîte De Saporta, François Dufour, and Alizée Geeraert. Optimal strategies for impulse control of Piecewise Deterministic Markov Processes. *Automatica*, 77:219–229, 2017.
- [12] Benoîte De Saporta and Huilong Zhang. Predictive maintenance for the heated hold-up tank. *Reliability Engineering & System Safety*, 115:82–90, 2013.
- [13] Alain Durmus, Arnaud Guillin, and Pierre Monmarché. Piecewise Deterministic Markov Processes and their invariant measures. In *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, volume 57, pages 1442–1475. Institut Henri Poincaré, 2021.
- [14] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.
- [15] Fastned Support. Qu’est-ce qui détermine la vitesse de recharge? <https://support.fastned.nl/hc/fr/articles/205694717-Qu-est-ce-qui-d%C3%A9termine-la-vitesse-de-recharge->, 2020.
- [16] Caroline Geiersbach and Georg Ch Pflug. Projected stochastic gradients for convex constrained problems in Hilbert spaces. *SIAM Journal on Optimization*, 29(3):2079–2099, 2019.
- [17] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Springer Science & Business Media, 2004.
- [18] Yonghui Huang and Xianping Guo. Finite-horizon Piecewise Deterministic Markov decision processes with unbounded transition rates. *Stochastics*, 91(1):67–95, 2019.
- [19] Paulin Jacquot, Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, and Nadia Oudjane. A privacy-preserving disaggregation algorithm for non-intrusive management of flexible energy. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 890–896. IEEE, 2019.
- [20] Caroline Le Floch, Emre Can Kara, and Scott Moura. PDE modeling and control of electric vehicle fleets for ancillary services: A discrete charging case. *IEEE Transactions on Smart Grid*, 9(2):573–581, 2016.
- [21] Man-Fai Leung, Jun Wang, and Duan Li. Decentralized robust portfolio optimization based on cooperative-competitive multiagent systems. *IEEE Transactions on Cybernetics*, 2021.
- [22] Yen Ting Lin and Nicolas E Buchler. Efficient analysis of stochastic gene dynamics in the non-adiabatic regime using Piecewise Deterministic Markov Processes. *Journal of The Royal Society Interface*, 15(138):20170804, 2018.
- [23] Ya-Feng Liu and Yu-Hong Dai. On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems. *IEEE transactions on Signal Processing*, 62(3):583–596, 2013.
- [24] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1):359–375, 1991.
- [25] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [26] Ryszard Rudnicki and Marta Tyran-Kamińska. Piecewise Deterministic Markov Processes in biological models. In *Semigroups of operators-theory and applications*, pages 235–255. Springer, 2015.
- [27] Adrien Séguret. An optimal control problem for the continuity equation arising in smart charging. *Journal of Mathematical Analysis and Applications*, page 127891, 2023.
- [28] Adrien Seguret, Clémence Alasseur, J Frédéric Bonnans, Antonio De Paola, Nadia Oudjane, and Vincenzo Trovato. Decomposition of convex high dimensional aggregative stochastic control problems. *Applied Mathematics & Optimization*, 88(1):8, 2023.
- [29] Adrien Séguret, Cheng Wan, and Clémence Alasseur. A mean field control approach for smart charging with aggregate power demand constraints. In *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pages 01–05. IEEE, 2021.

- [30] Simon R Sinsel, Rhea L Riemke, and Volker H Hoffmann. Challenges and solution technologies for the integration of variable renewable energy sources — a review. *renewable energy*, 145:2271–2285, 2020.
- [31] D Verms. Optimal control of Piecewise Deterministic Markov process. *Stochastics: An International Journal of Probability and Stochastic Processes*, 14(3):165–207, 1985.
- [32] Huilong Zhang, Benoîte De Saporta, François Dufour, and Gilles Deleuze. Dynamic reliability by using simulink and stateflow. *Chemical Engineering Transactions*, 33:529–534, 2013.
- [33] Huilong Zhang, Fares Innal, François Dufour, and Yves Dutuit. Piecewise Deterministic Markov Processes based approach applied to an offshore oil production system. *Reliability Engineering & System Safety*, 126:126–134, 2014.