

Optimal Control of the Fokker-Planck Equation with Space-Dependent Controls

A. Fleig^{*} and R. Guglielmi[†]

19.04.2017

Abstract

This paper is devoted to the analysis of a bilinear optimal control problem subject to the Fokker-Planck equation. The control function depends on time and space and acts as a coefficient of the advection term. For this reason, suitable integrability properties of the control function are required to ensure well-posedness of the state equation. Under these low regularity assumptions and for a general class of objective functionals, we prove the existence of optimal controls. Moreover, for common quadratic cost functionals of tracking and terminal type, we derive the system of first order necessary optimality conditions.

Keywords Bilinear control · Fokker-Planck equation · Optimal control theory · Optimization in Banach spaces · Probability density function · Stochastic process

Mathematics Subject Classification (2010) 35Q84 · 35Q93 · 49J20 · 49K20

1 Introduction

Initiated by Kolmogorov's work [20], the study of the Fokker-Planck (FP) equation, also known as Kolmogorov forward equation, has received great and increasing attention, since, for a large class of stochastic processes, it describes the evolution of the associated probability density function (PDF). Using the FP equation has proven to be a viable approach in several physical, chemical, and biological applications that involve noise. A large amount of literature has been developed on the FP equation in connection with transition PDFs that are associated to stochastic processes; see, for example, [15, 17]. In recent years, the well-posedness of the FP equation under low regularity assumptions on the coefficients has been studied in connection with existence, uniqueness and stability of martingale solutions

^{*}Chair of Applied Mathematics, Department of Mathematics, University of Bayreuth, 95440 Bayreuth, Germany, email: arthur.fleig@uni-bayreuth.de

[†]Radon Institut for Computational and Applied Mathematics - RICAM, 4040 Linz, Austria, email: roberto.guglielmi@ricam.oeaw.ac.at

to the related stochastic differential equation [22, 11]. Furthermore, control properties of the FP equation have become of main interest in mean field game theory; see [26].

Our focus is on the optimal control of the FP equation. It stems from a statistical approach, which allows to recast an optimal control problem subject to an Itô stochastic differential equation into a deterministic optimization problem subject to a FP equation. The idea behind this approach is that the state of a stochastic process can be characterized by the associated PDF. Therefore, controlling the PDF yields an accurate and flexible control strategy, which can accommodate a wide class of objectives; see also [8, Section 4]. In this direction, in [14, 18, 19, 33], PDF-control schemes were proposed, where the cost functional depends on the PDF of the stochastic state variable. In this way, a deterministic objective results. In particular, no average over all possible states of the stochastic process appears in the objective functional, which is usually the case in stochastic optimal control problems; see, e.g., [13]. Still, in [14, 18, 19, 33], stochastic methods were adopted in order to approximate the state variable of the random process.

To the contrary, in [2, 3] the authors approach the problem of tracking the PDF associated with the stochastic process directly. If the control acts through the drift term, then the evolution of the PDF is controlled through the advection term of the FP equation. This is a rather weak action of the controller on the system, usually called of bilinear type, since the control appears as a coefficient in the state equation. Indeed, only few controllability results are known for such kind of control systems, for instance in connection with quantum control systems and stochastic control [7] or in relation to the planning problem for the mean field games system [25]. Concerning the existence of bilinear optimal controls for a parabolic system of fourth order, a first result was given in [1], with a control function that only depends on time. This has been used in [3] in order to show existence of optimal controls for a FP equation with constant or time-dependent control functions. In this setting, however, due to the absence of space-dependent controls, there is no mechanism to cope with the diffusion term in the FP equation. Hence, unsurprisingly, acting on the space variable substantially improves tracking performance, as demonstrated in the numerical simulations in [12].

The aim of this work is to extend the theoretical study on the existence of bilinear optimal controls of the FP equation by [3] to the case of more general control functions, which depend on both time and space. We do not require any differentiability property of the control, which is in accordance with the simulations in [12]. For this reason, a careful analysis of the well-posedness of the FP equation is required. Indeed, suitable integrability assumptions are needed on the coefficient of the advection term in order to give meaning to the weak formulation of the equation. For this purpose, we use the functional framework proposed in the works of Aronson [4] and Aronson-Serrin [5]. In this setting, the advection coefficient belongs to a Bochner space that prevents us from choosing the set of square-integrable functions as the space of controls. As a result, the optimization problem is defined on a Banach space, a setting often considered whenever the state variable is subject to a nonlinear partial differential equation (PDE); see, for example, [9, 29]. Let us observe that, in recent works [22, 26], the well-posedness of the FP equation has been established even for drift coefficients that are square-integrable in time and space, in the context of renormalized

solutions. These papers could describe the right framework for studying the optimal control problem of the FP equation in a Hilbert setting.

The remainder of this paper is organized as follows. In Section 2, we formulate our optimal control problem and state general assumptions. In Section 3, we ensure the existence and uniqueness of (non-negative) solutions to the state equation. Section 4 is devoted to recast the FP equation in an abstract setting and to deduce a-priori estimates of its solution. These are used to prove our main result (Theorem 7 and Corollary 9) on existence of solutions to the considered optimal control problem for a general class of cost functionals. In Section 6, we deduce the system of first order necessary optimality conditions for common quadratic cost functionals. Section 7 concludes.

2 Problem Formulation and Assumptions

Given a final time $T > 0$, let us consider a controlled continuous-time stochastic process described by the (Itô) stochastic differential equation

$$\begin{aligned} dX_t &= b(X_t, t; u(X_t, t)) dt + \sigma(X_t, t) dW_t, \quad t \in]0, T[, \\ X(t=0) &= X_0, \end{aligned} \tag{1}$$

where $X_0 \in \mathbb{R}^d$ is the initial condition, $d \geq 1$, $W_t \in \mathbb{R}^m$ is an m -dimensional Wiener process, $m \geq 1$, $b = (b_1, \dots, b_d)$ is a vector valued drift function, and the dispersion matrix $\sigma(X_t, t) \in \mathbb{R}^{d \times m}$ is assumed to have full rank. The control $u(X_t, t)$, acting on (1) through the drift term b , has to be chosen from a suitable class of admissible functions to minimize a certain cost functional.

Assuming for simplicity that the state variable X_t evolves in a bounded domain Ω of \mathbb{R}^d with C^1 boundary $\partial\Omega$, we set the notations $Q := \Omega \times]0, T[$, $\Sigma := \partial\Omega \times]0, T[$, and $a_{ij} = \sum_{k=1}^d \sigma_{ik} \sigma_{kj} / 2$, $i, j = 1, \dots, d$. The matrix $(a_{ij})_{i,j}$ may depend on space $x = (x_1, \dots, x_d) \in \Omega$ and time $t \in]0, T[$ and is symmetric positive definite. We will denote by ∂_i and ∂_t the partial derivative with respect to x_i and t , respectively, where $i = 1, \dots, d$.

Under suitable assumptions on the coefficients b and σ , it is well known [27, p. 227], [28, p. 297] that, given an initial PDF ρ_0 , the evolution of the PDF associated with the stochastic process (1) satisfies the following FP equation

$$\partial_t \rho - \sum_{i,j=1}^d \partial_{ij}^2 (a_{ij} \rho) + \sum_{i=1}^d \partial_i (b_i(u) \rho) = 0 \quad \text{in } Q, \tag{2}$$

$$\rho(\cdot, 0) = \rho_0(\cdot) \quad \text{in } \Omega, \tag{3}$$

where the arguments (x, t) are omitted here and in the following, whenever clear from the context. Similarly, we use the notation $b_i(u)$ and $b_i(t; u(t))$ in order to stress the action of the control u through the coefficient b_i and to underline the time dependence, respectively, omitting the other arguments. We refer to [30] for an overview on classical methods for solving the Fokker-Planck equation and several applications regarding the description of

transitions of a system from a macroscopic point of view. A solution ρ to (2)-(3) shall furthermore satisfy the standard properties of a PDF, i.e.,

$$\rho(x, t) \geq 0, \quad (x, t) \in Q \quad \text{and} \quad \int_{\Omega} \rho(x, t) \, dx = 1, \quad t \in]0, T[. \quad (4)$$

Note that, in general, the FP equation evolves in the space domain \mathbb{R}^d rather than in Ω . However, if localized SDEs are under consideration or if the objective is to keep the PDF within a given compact set of Ω and the probability to find X_t outside of Ω is negligible, then we might focus on the description of the evolution of the PDF in the bounded domain $\Omega \subset \mathbb{R}^d$, employing homogeneous Dirichlet boundary conditions as in [3],

$$\rho(x, t) = 0 \quad \text{on } \Sigma, \quad (5)$$

also known as absorbing boundary conditions [27, p. 231]. For a complete characterization of possible boundary conditions in one space dimension, see the work of Feller [10]. In the multidimensional case, usually either absorbing boundary conditions or reflecting boundary conditions $n \cdot j(x, t) = 0$ on Σ are adopted, where j denotes the probability flux¹ and n the unit normal vector to the surface $\partial\Omega$. See also [17, Chapter 5] for a comparison between the Gihman-Skorohod [16] and the Feller classification of boundary conditions. Note that in general, the conservation of mass property in (4) only holds for reflecting boundary conditions.

With this in mind, we can formulate the optimal control problem. The aim is to control the stochastic process via the FP equation in an optimal way, i.e., to minimize some state- and control-dependent cost functional \tilde{J} , subject to the FP equation (2) with initial and boundary conditions, (3) and (5), over some set of admissible controls \mathcal{U}_{ad} . The absorbing boundary conditions (5), also used in [2, 3], are not only easier to analyze, they also provide the most suitable conditions in certain situations. For instance, when considering the Shiryaev stochastic diffusion on a bounded domain rather than on $[0, +\infty)$, a particle hitting the boundary shall leave the domain (by being absorbed) instead of being reflected back.

Furthermore, we do not demand the non-negativity and conservation of mass properties (4) explicitly, for the following reasons. As will be shown in Section 3, the former holds automatically if the initial state is non-negative. Regarding the latter, on the one hand, in the above Shiryaev example, the loss of the conservation of mass property is pertinent to the model. On the other hand, numerical results in [2, 3, 12] indicate that requiring this property can, at least in practice, often be circumvented by choosing a large enough domain Ω . However, under these conditions, the state is not necessarily a PDF and hence, will be denoted by y instead of ρ from now on.

To summarize, we consider the following optimal control problem:

$$\begin{aligned} \min_{u \in \mathcal{U}_{ad}, y} \tilde{J}(y, u) \quad \text{s.t.:} \quad & \partial_t y - \sum_{i,j=1}^d \partial_{ij}^2 (a_{ij} y) + \sum_{i=1}^d \partial_i (b_i(u) y) = 0 \quad \text{in } Q, \\ & y(\cdot, 0) = y_0(\cdot) \quad \text{in } \Omega, \\ & y = 0 \quad \text{on } \Sigma, \end{aligned} \quad (P)$$

¹The probability flux describes the flow of probability in terms of probability per unit time per unit area.

where

$$\mathcal{U}_{ad} := \{u \in \mathcal{U} : u_a \leq u(x, t) \leq u_b \text{ for almost all } (x, t) \in Q\}, \quad (6)$$

with $u_a, u_b \in \mathbb{R}^d$ and $u_a \leq u_b$ component-wise. The space of controls

$$\mathcal{U} := L^q(0, T; L^\infty(\Omega; \mathbb{R}^d)) \subset L^2(0, T; L^\infty(\Omega; \mathbb{R}^d)) \quad (7)$$

with $2 < q \leq \infty$ is motivated by the integrability requirements in [4] to ensure well-posedness of the state equation; see Section 3. Unless stated otherwise, we will use the above spaces \mathcal{U}_{ad} and \mathcal{U} throughout the paper, together with the following hypotheses.

Assumption 1

1. $\forall i, j = 1, \dots, d : a_{ij} \in C^1(\overline{\Omega})$.
2. $\exists \theta > 0$ such that $\forall \xi \in \mathbb{R}^d$ and almost all $x \in \Omega : \sum_{i,j=1}^d a_{ij}(x) \xi_i \xi_j \geq \theta |\xi|^2$.
3. The function $b : \mathbb{R}^{d+1} \times \mathcal{U} \rightarrow \mathbb{R}^d, (x, t; u) \mapsto b(x, t; u)$ satisfies the growth condition

$$\sum_{i=1}^d |b_i(x, t; u)|^2 \leq M(1 + |u(x, t)|^2) \quad \forall x \in \mathbb{R}^d, \quad (8)$$

for every $i = 1, \dots, d, t \in [0, T], u \in \mathcal{U}$, and some constant $M > 0$.

For simplicity, we assume the coefficients a_{ij} to be independent of time in order to cope with an autonomous operator. In Sections 5 and 6, Assumption 1(3) is replaced by the following, stronger requirement:

Assumption 2

$\exists \gamma_i \in L^\infty(\Omega) : b_i(x, t; u) = \gamma_i(x) + u_i(x, t), i = 1, \dots, d$.

The fact that b is affine on u is exploited in particular in the proofs of Theorem 7 and Lemma 11, in order to prove the existence of optimal solutions and the differentiability of the control-to-state operator, which will be introduced in Section 5.

3 Well-Posedness of the Fokker–Planck Equation

In this section, we ensure the well-posedness of the FP equation in (P), where we add a source term $f : Q \rightarrow \mathbb{R}$ on the right-hand-side, which will be of use for the well-posedness of the adjoint equation in Section 6.

Setting $\tilde{b}_j(u) := \sum_{i=1}^d \partial_i a_{ij} - b_j(u)$, we can recast the FP equation in flux formulation

$$\partial_t y - \sum_{j=1}^d \partial_j \left(\sum_{i=1}^d a_{ij} \partial_i y + \tilde{b}_j(u) y \right) = f \quad \text{in } Q.$$

Together with the initial and boundary conditions in (P), we have the associated weak formulation

$$\begin{aligned} \iint_Q f v \, dx dt &= \iint_Q \partial_t y v \, dx dt - \iint_Q \left(\sum_{j=1}^d \partial_j \left(\sum_{i=1}^d a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \right) v \, dx dt \\ &= - \iint_Q y \partial_t v \, dx dt - \int_{\Omega} y(\cdot, 0) v(\cdot, 0) \, dx \\ &\quad + \iint_Q \sum_{j=1}^d \left(\sum_{i=1}^d a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \partial_j v \, dx dt \end{aligned}$$

for test functions $v \in W_2^{1,1}(Q)$ with $v|_{\partial\Omega} = 0$ and $v(\cdot, T) = 0$.

We make use of this in the following theorem, which is a special case of [4, Thm. 1, p. 634] and guarantees the existence and uniqueness of (non-negative) solutions.

Theorem 3

Let $y_0 \in L^2(\Omega)$. Additionally, let $f \in L^q(0, T; L^\infty(\Omega))$ or $f = \operatorname{div}(\tilde{f})$ for some $\tilde{f}: Q \rightarrow \mathbb{R}^d$ with $\tilde{f}_j \in L^2(Q)$, $j = 1, \dots, d$. Then, there exists a unique $y \in L^2(0, T; H_0^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$ satisfying

$$\iint_Q -y \partial_t v + \sum_{j=1}^d \left(\sum_{i=1}^d a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \partial_j v - f v \, dx dt = \int_{\Omega} y_0 v(\cdot, 0) \, dx \quad (9)$$

for every $v \in W_2^{1,1}(Q)$ with $v|_{\partial\Omega} = 0$ and $v(\cdot, T) = 0$, i.e., y is the unique weak solution of the Fokker-Planck initial boundary value problem defined in (P), including a right-hand-side f in the FP equation. Moreover, if $f \equiv 0$ and $0 \leq y_0 \leq m$ almost everywhere in Ω for some $m > 0$, then y is bounded and $y(x, t) \geq 0$ almost everywhere in Q .

Note that, due to the choice of \mathcal{U} in (7) and Assumption 1(3), $\tilde{b}_j(u)$ belongs to $L^q(0, T; L^\infty(\Omega))$ for $j = 1, \dots, d$, as required in [4].

The solution obtained by Theorem 3 is more regular; indeed, it belongs to the $W(0, T)$ space. We remind that

$$W(0, T) := \{y \in L^2(0, T; V) : \dot{y} \in L^2(0, T; V')\} \subset C([0, T]; H),$$

where \dot{y} denotes the time derivative of y and $H := L^2(\Omega)$, $V := H_0^1(\Omega)$, and $V' := H^{-1}(\Omega)$, the dual space of V , endowed with norms

$$\|y\|_H^2 := \int_{\Omega} y^2 \, dx, \quad \|y\|_V^2 := \int_{\Omega} |\nabla y|^2 \, dx, \quad \|L\|_{V'} := \sup_{y \in V, \|y\|_V=1} |\langle L, y \rangle_{V', V}|,$$

respectively, from the Gelfand triple $V \hookrightarrow H \hookrightarrow V'$. We denote by $|\cdot|$ the Euclidean norm and by $\langle \cdot, \cdot \rangle_{V', V}$ the duality map between V and V' . This notation and these spaces are used throughout the paper.

Proposition 4

Under the assumptions of Theorem 3, the solution y in Theorem 3 belongs to $W(0, T)$, possibly after a modification on a set of measure zero.

Proof. The proof is analogous to the one of [32, Theorem 3.12], the only change being a different functional F . The idea is to show that F belongs to $L^2(0, T; V')$ and to rewrite the variational formulation of the PDE in terms of F to show that $\dot{y} = F$ in the sense of vector-valued distributions. In our case, for any fixed t , the linear functional is given by $F(t): V \rightarrow \mathbb{R}$,

$$v \mapsto - \sum_{j=1}^d \left(\sum_{i=1}^d a_{ij} \partial_i y(t) + \tilde{b}_j(t; u(t)) y(t), \partial_j v \right)_H + (f(t), v)_H.$$

We first assume $f \in L^q(0, T; L^\infty(\Omega))$ with $2 < q \leq \infty$. $F(t)$ is bounded and thus continuous for all $t \in]0, T[$:

$$\begin{aligned} |F(t)v| &= \left| - \sum_{j=1}^d \int_{\Omega} \left(\sum_{i=1}^d a_{ij} \partial_i y(t) + \tilde{b}_j(t; u(t)) y(t) \right) \partial_j v \, dx + \int_{\Omega} f(t) v \, dx \right| \\ &\leq \sum_{j=1}^d \int_{\Omega} \sum_{i=1}^d |a_{ij}| |\partial_i y(t)| |\partial_j v| \, dx + \int_{\Omega} |f(t)| |v| \, dx \\ &\quad + \sum_{j=1}^d \int_{\Omega} |\tilde{b}_j(t; u(t))| |y(t)| |\partial_j v| \, dx \\ &\leq \underbrace{\sum_{i,j=1}^d \|a_{ij}\|_{L^\infty(\Omega)}}_{=:C} \|y(t)\|_V \|v\|_V + c_\Omega \|f(t)\|_H \|v\|_V \\ &\quad + \sum_{j=1}^d \|\tilde{b}_j(t; u(t))\|_{L^\infty(\Omega)} \|y(t)\|_H \|v\|_V, \end{aligned}$$

where c_Ω is such that $\|v\|_H \leq c_\Omega \|v\|_V$ for any $v \in V = H_0^1(\Omega)$. Therefore,

$$\|F(t)\|_{V'} \leq C \|y(t)\|_V + \sum_{j=1}^d \|\tilde{b}_j(t; u(t))\|_{L^\infty(\Omega)} \|y(t)\|_H + c_\Omega \|f(t)\|_H. \quad (10)$$

Since $\|y(t)\|_V \in L^2(0, T)$, $\|\tilde{b}_j(t; u(t))\|_{L^\infty(\Omega)} \in L^q(0, T)$, $\|y(t)\|_H \in L^\infty(0, T)$, and $\|f(t)\|_H \in L^q(0, T)$ with $q > 2$, the right-hand-side of (10) belongs to $L^2(0, T)$, i.e., $F \in L^2(0, T; V')$. The remaining steps are the same as in the proof of [32, Theorem 3.12].

If $f = \operatorname{div}(\tilde{f})$, the spatial derivatives are transferred to v , which results in a very similar calculation and, in particular, also in $F \in L^2(0, T; V')$. \square \square

Furthermore, note that we have $\int_{\Omega} y_0 v \, dx = \lim_{t \rightarrow 0} \int_{\Omega} y(t) v \, dx = \int_{\Omega} y(0) v \, dx$ for all $v \in V$, where the first equality follows from (9) and the second holds because $W(0, T) \subset C([0, T]; H)$. Consequently, $y(0) = y_0$ in Ω .

4 A-priori Estimates

The purpose of this section is to deduce a-priori estimates of solutions to the Fokker-Planck initial boundary value problem given in (P), including a right-hand-side $f \in L^2(0, T; V')$ in the FP equation. For the sake of clarity, we recast it in abstract form

$$\begin{cases} \dot{y}(t) + Ay(t) + B(u(t), y(t)) = f(t) & \text{in } V', \, t \in]0, T[, \\ y(0) = y_0, \end{cases} \quad (11)$$

where $y_0 \in H$, $A: V \rightarrow V'$ is a linear and continuous operator such that

$$\langle Az, \varphi \rangle_{V', V} := \int_{\Omega} \sum_{i,j=1}^d \partial_i(a_{ij}z) \partial_j \varphi \, dx \quad \forall \varphi \in V,$$

and the operator $B: L^\infty(\Omega; \mathbb{R}^d) \times H \rightarrow V'$ is defined by

$$\langle B(u, y), \varphi \rangle_{V', V} := - \int_{\Omega} \sum_{i=1}^d b_i(u) y \partial_i \varphi \, dx = - \int_{\Omega} y b(u) \cdot \nabla \varphi \, dx \quad \forall \varphi \in V.$$

In the following, $\mathcal{E}(y_0, u, f)$ refers to (11), whenever we want to point out the data (y_0, u, f) . To ease the notation, we will still denote by A and B the two operators $A: L^2(0, T; V) \rightarrow L^2(0, T; V')$ and $B: \mathcal{U} \times L^\infty(0, T; H) \rightarrow L^q(0, T; V')$ such that for all $\varphi \in L^2(0, T; V)$, we have $Az = - \sum_{i,j=1}^d \partial_{ij}^2(a_{ij}z)$ and

$$\int_0^T \langle Az(t), \varphi(t) \rangle_{V', V} \, dt = \iint_Q \sum_{i,j=1}^d \partial_i(a_{ij}z) \partial_j \varphi \, dx dt, \quad (12)$$

and $B(u, y) = \sum_{i=1}^d \partial_i(b_i(u)y) = \operatorname{div}(b(u)y)$ such that

$$\int_0^T \langle B(u(t), y(t)), \varphi(t) \rangle_{V', V} \, dt = - \iint_Q \sum_{i=1}^d b_i(u) y \partial_i \varphi \, dx dt \quad (13)$$

for all $\varphi \in L^{q'}(0, T; V)$ with $1/q + 1/q' = 1$. Indeed, thanks to Assumption 1(3), we have $\operatorname{div}(b(u)y) \in L^q(0, T; V')$ and

$$\|B(u, y)\|_{L^q(0, T; V')} = \|\operatorname{div}(b(u)y)\|_{L^q(0, T; V')} \leq M(1 + \|u\|_{\mathcal{U}}) \|y\|_{L^\infty(0, T; H)}.$$

Note that the integral on the r.h.s. in (12) is not symmetric in z and φ , owing to the fact that A is not self-adjoint. The bilinear form $a:]0, T[\times V \times V \rightarrow \mathbb{R}$ associated with the FP equation is defined by

$$\begin{aligned} a(t, \psi, \varphi) &:= \int_{\Omega} \left(\sum_{i,j=1}^d \partial_i(a_{ij}\psi) \partial_j \varphi - \sum_{i=1}^d b_i(t; u(t)) \psi \partial_i \varphi \right) dx \\ &= \int_{\Omega} \left(\sum_{i,j=1}^d a_{ij} \partial_i \psi \partial_j \varphi + \sum_{j=1}^d \tilde{b}_j(t, u(t)) \psi \partial_j \varphi \right) dx. \end{aligned}$$

Thanks to the uniform ellipticity of A and Young's inequality, for every $\varepsilon > 0$, $t \in]0, T[$, and every $\varphi \in V$, we have that

$$\begin{aligned} \theta \int_{\Omega} |\nabla \varphi|^2 dx &\leq \int_{\Omega} \sum_{i,j=1}^d a_{ij} \partial_i \varphi \partial_j \varphi dx = a(t, \varphi, \varphi) - \int_{\Omega} \sum_{j=1}^d \tilde{b}_j(t; u(t)) \varphi \partial_j \varphi dx \\ &\leq a(t, \varphi, \varphi) + \|\tilde{b}(t; u(t))\|_{L^\infty(\Omega; \mathbb{R}^d)} \left(\varepsilon \int_{\Omega} |\nabla \varphi|^2 dx + \frac{1}{4\varepsilon} \int_{\Omega} |\varphi|^2 dx \right). \end{aligned}$$

Thus, with $\varepsilon = 3\theta/(4\|\tilde{b}(t; u(t))\|_{L^\infty(\Omega; \mathbb{R}^d)})$, we conclude

$$\frac{\theta}{4} \|\varphi\|_V^2 \leq a(t, \varphi, \varphi) + C_1(t) \|\varphi\|_H^2, \quad (14)$$

where

$$C_1(t) := \|\tilde{b}(t; u(t))\|_{L^\infty(\Omega; \mathbb{R}^d)}^2 / (3\theta). \quad (15)$$

We now derive some a-priori estimates on the solution of (11). We will need them in the following sections. From this section on, we denote by M and C generic, positive constants that might change from line to line.

Lemma 5

Let $y_0 \in H$, $f \in L^2(0, T; V')$ and $u \in \mathcal{U}$. Then, a solution y of (11) satisfies the estimates

$$\|y\|_{L^\infty(0, T; H)}^2 \leq M(u) \left(\|y(0)\|_H^2 + \|f\|_{L^2(0, T; V')}^2 \right), \quad (16)$$

$$\|y\|_{L^2(0, T; V)}^2 \leq (1 + \|u\|_{\mathcal{U}}^2) M(u) \left(\|y(0)\|_H^2 + \|f\|_{L^2(0, T; V')}^2 \right), \quad (17)$$

$$\|\dot{y}\|_{L^2(0, T; V')}^2 \leq (1 + \|u\|_{\mathcal{U}}^2) M(u) \left(\|y(0)\|_H^2 + \|f\|_{L^2(0, T; V')}^2 \right) + 2 \|f\|_{L^2(0, T; V')}^2, \quad (18)$$

where $M(u) := Ce^{c(1+\|u\|_{\mathcal{U}}^2)}$ for some positive constants c , C .

Proof. Let y a solution of (11) and $t \in]0, T[$. Multiplying (11) by $y(t)$, we get

$$\frac{1}{2} \frac{d}{dt} (\|y(t)\|_H^2) + a(t, y(t), y(t)) = \langle f(t), y(t) \rangle_{V', V}, \quad t \in]0, T[,$$

and thus

$$\begin{aligned} \frac{d}{dt} (\|y(t)\|_H^2) + \frac{\theta}{2} \|y(t)\|_V^2 &\leq \frac{d}{dt} (\|y(t)\|_H^2) + 2a(t, y(t), y(t)) + 2C_1(t) \|y(t)\|_H^2 \\ &= 2\langle f(t), y(t) \rangle_{V', V} + 2C_1(t) \|y(t)\|_H^2 \\ &\leq 2\varepsilon \|y(t)\|_V^2 + \frac{1}{2\varepsilon} \|f(t)\|_{V'}^2 + 2C_1(t) \|y(t)\|_H^2. \end{aligned}$$

Fixing $\varepsilon = \theta/8$, we deduce the relation

$$\frac{d}{dt} (\|y(t)\|_H^2) + \frac{\theta}{4} \|y(t)\|_V^2 \leq \frac{4}{\theta} \|f(t)\|_{V'}^2 + 2C_1(t) \|y(t)\|_H^2. \quad (19)$$

Applying Gronwall's inequality, we have that

$$\|y(t)\|_H^2 \leq e^{\int_0^t 2C_1(\tau) d\tau} \left(\|y(0)\|_H^2 + \frac{4}{\theta} \int_0^t \|f(\tau)\|_{V'}^2 d\tau \right).$$

For $u \in \mathcal{U}$, the inequality

$$\|u\|_{L^2(0,T;L^\infty(\Omega;\mathbb{R}^d))} \leq T^{\frac{q-2}{2q}} \|u\|_{\mathcal{U}} \quad (20)$$

holds. With $C_1(t)$ from (15) and due to Assumption 1(3) and (20), we deduce that $\int_0^T 2C_1(t) dt \leq M(1)$ and thus

$$\|y\|_{L^\infty(0,T;H)}^2 \leq C e^{c(1+\|u\|_{\mathcal{U}}^2)} \left(\|y(0)\|_H^2 + \|f\|_{L^2(0,T;V')}^2 \right).$$

Moreover, integrating (19) over $]0, T[$, we conclude that

$$\begin{aligned} \|y\|_{L^2(0,T;V)}^2 &\leq C \left(\|y(0)\|_H^2 + \|f\|_{L^2(0,T;V')}^2 \right) + C(1 + \|u\|_{\mathcal{U}}^2) \|y\|_{L^\infty(0,T;H)}^2 \\ &\leq C(1 + \|u\|_{\mathcal{U}}^2) e^{c(1+\|u\|_{\mathcal{U}}^2)} \left(\|y(0)\|_H^2 + \|f\|_{L^2(0,T;V')}^2 \right). \end{aligned}$$

We remind that C might change from line to line. Finally, multiplying (11) by $\varphi \in L^2(0, T; V)$ and integrating over $]0, T[$ yields

$$\begin{aligned} \left| \int_0^T \langle \dot{y}(t), \varphi(t) \rangle_{V', V} dt \right| &\leq \|y\|_{L^\infty(0,T;H)} \|u\|_{L^2(0,T;L^\infty(\Omega;\mathbb{R}^d))} \|\varphi\|_{L^2(0,T;V)} \\ &\quad + C_\alpha \|y\|_{L^2(0,T;V)} \|\varphi\|_{L^2(0,T;V)} + \|f\|_{L^2(0,T;V')} \|\varphi\|_{L^2(0,T;V)}, \end{aligned}$$

where $C_\alpha > 0$ is such that $\|A\xi\|_{V'} \leq C_\alpha \|\xi\|_V$ for all $\xi \in V$. Thanks to (20),

$$\|\dot{y}\|_{L^2(0,T;V')} \leq C_\alpha \|y\|_{L^2(0,T;V)} + C \|y\|_{L^\infty(0,T;H)} \|u\|_{\mathcal{U}} + \|f\|_{L^2(0,T;V')}.$$

Using twice the estimate $(a+b)^2 \leq 2a^2 + 2b^2$, we derive (18) by the estimates on $\|y\|_{L^\infty(0,T;H)}$ and $\|y\|_{L^2(0,T;V)}$. \square \square

5 Existence of Optimal Controls

This section contains our main result: the existence of optimal controls for (P), with \mathcal{U}_{ad} and \mathcal{U} as in (6) and (7). Fixing $y_0 \in H$, we introduce the control-to-state operator $\Theta: \mathcal{U} \rightarrow C([0, T]; H)$ such that $u \mapsto y \in C([0, T]; H)$ is a solution of $\mathcal{E}(y_0, u, 0)$. Thus, the optimization problem turns into minimizing the so-called reduced cost functional $J: \mathcal{U} \rightarrow \mathbb{R}$ such that $J(u) := \tilde{J}(\Theta(u), u)$, which we assume to be bounded from below, over the non-empty subset of admissible controls $\mathcal{U}_{ad} \subset \mathcal{U}$. In order to prove the main theorem, we will need the following compactness result (see [6], [23, Théorème 5.1, p. 58] or [31]).

Theorem 6

Let I be an open and bounded interval of \mathbb{R} , and let X, Y, Z be three Banach spaces, with dense and continuous inclusions $Y \hookrightarrow X \hookrightarrow Z$, the first one being compact. Then, for every $p \in [1, +\infty[$ and $r > 1$, we have the compact inclusions

$$L^p(I; Y) \cap W^{1,1}(I; Z) \hookrightarrow L^p(I; X)$$

and

$$L^\infty(I; Y) \cap W^{1,r}(I; Z) \hookrightarrow C(\bar{I}; X).$$

Theorem 7

Let $y_0 \in H$. Consider the minimization of the reduced cost functional $J(u) = \tilde{J}(\Theta(u), u)$ over \mathcal{U}_{ad} . Assume that J is bounded from below and (sequentially) weakly-star lower semi-continuous. Then, there exists a pair $(\bar{y}, \bar{u}) \in C([0, T]; H) \times \mathcal{U}_{ad}$ such that \bar{y} solves $\mathcal{E}(y_0, \bar{u}, 0)$ and \bar{u} minimizes J in \mathcal{U}_{ad} .

Proof. Let $(u_n)_{n \geq 1}$ be a minimizing sequence, i.e., $J(u_n) \rightarrow I := \inf_{u \in \mathcal{U}_{ad}} J(u)$ as $n \rightarrow \infty$. Since $(u_n)_{n \geq 1} \subset \mathcal{U}_{ad}$, we have $\|u_n\|_{\mathcal{U}} \leq c \|u_n\|_{L^\infty(Q)} \leq C$ for some constants $c, C > 0$ and any $n \geq 1$. Moreover, the pair (u_n, y_n) satisfies the state equation

$$\dot{y}_n(t) + Ay_n(t) + B(u_n(t), y_n(t)) = 0, \quad y_n(0) = y_0. \quad (21)$$

The a-priori estimates of Lemma 5 ensure that there exists a positive constant, still denoted by C , such that

$$\|y_n\|_{L^\infty(0,T;H)}, \quad \|y_n\|_{L^2(0,T;V)}, \quad \|\dot{y}_n\|_{L^2(0,T;V')} \leq C,$$

and so we deduce that

$$\begin{aligned} \|Ay_n\|_{L^2(0,T;V')} &\leq C_\alpha \|y_n\|_{L^2(0,T;V)} \leq C, \\ \|B(u_n, y_n)\|_{L^2(0,T;V')} &\leq c \|B(u_n, y_n)\|_{L^q(0,T;V')} \\ &\leq M(1 + \|u_n\|_{\mathcal{U}}) \|y_n\|_{L^\infty(0,T;H)} \leq C, \end{aligned}$$

where we remind that the constant $C_\alpha > 0$, which appears in the proof of Lemma 5, is such that $\|A\xi\|_{V'} \leq C_\alpha \|\xi\|_V$ for all $\xi \in V$. Thus, there exist subsequences (still indexed with the subscript n) such that

$$\begin{aligned} u_n &\overset{*}{\rightharpoonup} \bar{u} && \text{weakly-star in } \mathcal{U}, \\ y_n &\overset{*}{\rightharpoonup} \bar{y} && \text{weakly-star in } L^\infty(0, T; H), \\ y_n &\rightharpoonup \bar{y} && \text{weakly in } L^2(0, T; V), \\ \dot{y}_n &\rightharpoonup \dot{\bar{y}} && \text{weakly in } L^2(0, T; V'), \\ Ay_n &\rightharpoonup \chi && \text{weakly in } L^2(0, T; V'), \\ B(u_n, y_n) &\rightharpoonup \Lambda && \text{weakly in } L^2(0, T; V'). \end{aligned}$$

Since the Banach-Alaoglu theorem ensures that \mathcal{U}_{ad} is weakly-star closed, we deduce that $\bar{u} \in \mathcal{U}_{ad}$. We now want to pass to the limit in the state equation (21). First of all, we observe that $\psi = \dot{\bar{y}}$, thanks to the convergence in the $\sigma(\mathcal{D}(0, T; V), \mathcal{D}'(0, T; V'))$ topology. Thus, $\bar{y} \in W(0, T) \subset C([0, T]; H)$. Moreover, since the operator $A: L^2(0, T; V) \rightarrow L^2(0, T; V')$ is strongly continuous, and therefore weakly continuous, too, we deduce that $A\bar{y} = \chi$. Finally, we claim that $B(\bar{u}, \bar{y}) = \Lambda$, which, because of the bilinear action of the control, is the most difficult part of the proof. Note that, thanks to the first relation in Theorem 6 with $Y := V$, $X := H$, and $Z := V'$, the embedding $W(0, T) \subset L^2(0, T; H)$ is compact. Thus, $(y_n)_n$ admits a subsequence strongly convergent to \bar{y} in $L^2(0, T; H)$. Therefore, for every $\varphi \in L^2(0, T; V)$,

$$\begin{aligned} &\int_0^T \langle B(\bar{u}(t), \bar{y}(t)) - \Lambda(t), \varphi(t) \rangle_{V', V} dt \\ &= - \iint_Q \bar{y} b(\bar{u}) \cdot \nabla \varphi \, dx dt - \lim_{n \rightarrow \infty} \int_0^T \langle B(u_n(t), y_n(t)), \varphi(t) \rangle_{V', V} dt \\ &= - \iint_Q \bar{y} b(\bar{u}) \cdot \nabla \varphi \, dx dt + \lim_{n \rightarrow \infty} \iint_Q y_n b(u_n) \cdot \nabla \varphi \, dx dt \\ &= - \lim_{n \rightarrow \infty} \iint_Q (\bar{y} b(\bar{u}) - y_n b(u_n)) \cdot \nabla \varphi \, dx dt \\ &= - \lim_{n \rightarrow \infty} \iint_Q \bar{y} (b(\bar{u}) - b(u_n)) \cdot \nabla \varphi \, dx dt - \lim_{n \rightarrow \infty} \iint_Q (\bar{y} - y_n) b(u_n) \cdot \nabla \varphi \, dx dt, \end{aligned}$$

where ∇ denotes the gradient with respect to the space variable $x \in \mathbb{R}^d$. We observe that $\bar{y} \in L^\infty(0, T; H)$ and $\partial_i \varphi \in L^2(0, T; H)$ for all $i = 1, \dots, d$, thus $\bar{y} \partial_i \varphi \in L^2(0, T; L^1(\Omega)) \subset L^{q'}(0, T; L^1(\Omega))$ with q' such that $1/q + 1/q' = 1$ and $L^q(0, T; L^\infty(\Omega)) = [L^{q'}(0, T; L^1(\Omega))]^*$, since the Lebesgue measure is σ -finite. We remind that b is affine on u ; see Assumption 2. Therefore, $b(\bar{u}) - b(u_n) = (\bar{u}_i - u_{n,i})_{i=1, \dots, d}$. Now $u_n \overset{*}{\rightharpoonup} \bar{u}$ weakly-star in \mathcal{U} ensures that the first integral goes to 0 as $n \rightarrow +\infty$. Furthermore, since the sequence $(b(u_n))_n$ is uniformly bounded and $y_n \rightarrow \bar{y}$ strongly in $L^2(0, T; H)$,

$$\left| \iint_Q (\bar{y} - y_n) b(u_n) \cdot \nabla \varphi \, dx dt \right| \leq C \|\bar{y} - y_n\|_{L^2(0, T; H)} \|\varphi\|_{L^2(0, T; V)} \rightarrow 0$$

as $n \rightarrow +\infty$. Additionally, we observe that $\bar{y}(0) = y_0$, hence

$$\dot{\bar{y}}(t) + A\bar{y}(t) + B(\bar{u}(t), \bar{y}(t)) = 0, \quad \bar{y}(0) = y_0.$$

Finally, owing to the weakly-star lower semicontinuity of J , we conclude that

$$J(\bar{u}) \leq \liminf_{n \rightarrow \infty} J(u_n) = I.$$

Thus, (\bar{y}, \bar{u}) is an optimal pair for the considered optimal control problem. \square \square

Theorem 7 clearly also holds for any $\mathcal{U}_{ad} \subset \mathcal{U}$ bounded and weakly-star closed. However, observe that in the unconstrained case $\mathcal{U}_{ad} \equiv \mathcal{U}$, asking only $J(u) \geq \lambda \|u\|_{\mathcal{U}}$ for some $\lambda > 0$ is not enough. Instead, one can modify the proof straightforwardly by requiring $J(u) \geq \lambda \|u\|_{L^\infty(Q; \mathbb{R}^d)}$, which is not very practical. An alternative might be to require more regularity on the state y and on the control u , in order to gain the same level of compactness required to deduce that $B(\bar{u}, \bar{y}) = \Lambda$. Indeed, further regularity of y can be ensured by standard improved regularity results; see, for example, [34, Theorems 27.2 and 27.5] and [21, Theorem 6.1 and Remark 6.3]. However, these results require more regularity of the coefficients in the PDE and hence, on the control. In particular, one would need differentiability of u both in time and space. In comparison, requiring box constraints as in (6) seems to be a less restrictive choice. Note that, in case of bilinear action of the control into the system, even box constraints might not suffice to ensure the existence of optimal controls in general; see, for example, [24, Section 15.3, p. 237].

Remark 8

We have shown in the previous proof that the control-to-state map $\Theta: \mathcal{U}_{ad} \subset \mathcal{U} \rightarrow C([0, T]; H) \subset L^2(0, T; H)$, $u \mapsto \Theta(u) = y \in L^2(0, T; H)$, where y solves $\mathcal{E}(y_0, u, 0)$, is sequentially continuous from \mathcal{U}_{ad} (with the weak-star topology induced by \mathcal{U}) to $L^2(0, T; H)$ (with the strong topology).

Corollary 9

Let $y_d \in L^2(0, T; H)$, $y_\Omega \in H$, and $\alpha, \beta, \lambda \geq 0$. Consider the final time observation operator $S_T: W(0, T) \rightarrow H$ such that $y \mapsto y(T)$. Then, an optimal pair $(\bar{y}, \bar{u}) \in C([0, T]; H) \times \mathcal{U}_{ad}$ exists for the reduced cost functional

$$J(u) := \frac{\alpha}{2} \|\Theta(u) - y_d\|_{L^2(Q)}^2 + \frac{\beta}{2} \|S_T \Theta(u) - y_\Omega\|_{L^2(\Omega)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q; \mathbb{R}^d)}^2. \quad (22)$$

Proof. The cost functional (22) is bounded from below by zero. Moreover, it is weakly lower semicontinuous in $L^2(Q; \mathbb{R}^d)$. This is due to Remark 8 and the fact that the embedding $W(0, T) \subset C([0, T]; H)$ is continuous, that the operator S_T is linear and continuous, and that the norm functionals $\|\cdot\|_H^2$ and $\|\cdot\|_{L^2(Q; \mathbb{R}^d)}^2$ are weakly lower semicontinuous on H and $L^2(Q; \mathbb{R}^d)$, respectively. Moreover, a minimizing sequence $(u_n)_{n \geq 1}$ in \mathcal{U}_{ad} converging to I is uniformly bounded both in \mathcal{U} and in $L^2(Q; \mathbb{R}^d)$. Since the weak-star convergence in \mathcal{U} implies the weak convergence in $L^2(Q; \mathbb{R}^d)$, we do not need to require weakly-star lower semicontinuity of J . Therefore, we can conclude the existence of an optimal pair $(\bar{y}, \bar{u}) \in C([0, T]; H) \times \mathcal{U}_{ad}$. \square \square

Remark 10

Corollary 9 applies analogously to the case of time-independent controls u in the admissible space

$$\tilde{\mathcal{U}}_{ad} := \{u \in L^\infty(\Omega; \mathbb{R}^d) : u_a \leq u(x) \leq u_b \text{ for almost every } x \in \Omega\} \quad (23)$$

for the reduced cost functional

$$J_2(u) := \frac{\alpha}{2} \|\Theta(u) - y_d\|_{L^2(Q)}^2 + \frac{\beta}{2} \|\Theta(u)(T) - y_\Omega\|_{L^2(\Omega)}^2 + \frac{\lambda}{2} \|u\|_{L^2(\Omega; \mathbb{R}^d)}^2.$$

6 Adjoint State and Optimality Conditions

For the optimal control problem (P), modified by using the reduced cost functionals considered in Corollary 9 and Remark 10, we derive the first order necessary optimality conditions in this section. Incidentally, let us point out that these quadratic objective functionals are commonly used in theory and practice; see, for example, [3, 12, 32]. We start by denoting the operator

$$D(u, y) := B(u - \gamma, y) = \operatorname{div}(uy) \quad \forall u \in \mathcal{U}, y \in L^\infty(0, T; H),$$

and proving the differentiability of the control-to-state operator.

Lemma 11

Let $y_0 \in H$. The control-to-state operator Θ from Section 5 is Fréchet-differentiable and, for every \bar{u} , $h \in \mathcal{U}$, the function $\Theta'(\bar{u})h$ satisfies

$$\begin{cases} \dot{z}(t) + Az(t) + B(\bar{u}(t), z(t)) = -D(h(t), \bar{y}(t)) & \text{in } V', \quad t \in]0, T[, \\ z(0) = 0, \end{cases} \quad (24)$$

where $\bar{y} = \Theta(\bar{u})$.

Note that Theorem 3 ensures the existence of a unique weak solution of (24).

Proof. Thanks to Assumption 2, the map $L: \mathcal{U} \rightarrow C([0, T]; H)$, such that $h \mapsto z \in C([0, T]; H)$ is a solution of (24), is linear. Moreover, L is continuous; indeed, the estimate (16) yields

$$\begin{aligned} \|z\|_{L^\infty(0, T; H)}^2 &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|D(h, \bar{y})\|_{L^2(0, T; V')}^2 \\ &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|\bar{y}\|_{L^\infty(0, T; H)}^2 \|h\|_{\mathcal{U}}^2 \leq C \|h\|_{\mathcal{U}}^2. \end{aligned}$$

Let us now introduce $y_h := \Theta(\bar{u} + h)$, the solution of $\mathcal{E}(y_0, \bar{u} + h, 0)$, and set $y := y_h - \bar{y}$. Thus, y satisfies

$$\begin{cases} \dot{y}(t) + Ay(t) + B(\bar{u}(t), y(t)) = -D(h(t), y_h(t)) & \text{in } V', \quad t \in]0, T[, \\ y(0) = 0. \end{cases}$$

Moreover, $D(h, y_h) \in L^q(0, T; V') \subset L^2(0, T; V')$, and (16) ensures

$$\begin{aligned} \|y\|_{L^\infty(0, T; H)}^2 &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|D(h, y_h)\|_{L^2(0, T; V')}^2 \\ &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|y_h\|_{L^\infty(0, T; H)}^2 \|h\|_{\mathcal{U}}^2, \end{aligned}$$

with $\|y_h\|_{L^\infty(0, T; H)}^2 \leq C e^{c(1+\|\bar{u}+h\|_{\mathcal{U}}^2)} \|y_0\|_H^2$, which is locally bounded in h . Finally, $w := y - z$ is a solution of $\mathcal{E}(0, \bar{u}, -D(h(t), y(t)))$ and satisfies

$$\begin{aligned} \|w\|_{L^\infty(0, T; H)}^2 &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|D(h, y)\|_{L^2(0, T; V')}^2 \\ &\leq C e^{c(1+\|\bar{u}\|_{\mathcal{U}}^2)} \|y\|_{L^\infty(0, T; H)}^2 \|h\|_{\mathcal{U}}^2, \end{aligned}$$

that is,

$$\|\Theta(\bar{u} + h) - \Theta(\bar{u}) - z\|_{L^\infty(0, T; H)}^2 \leq C \|y_0\|_H^2 e^{c(1+\|\bar{u}+h\|_{\mathcal{U}}^2)} \|h\|_{\mathcal{U}}^4.$$

Therefore, Θ is Fréchet differentiable and, for all $\bar{u}, h \in \mathcal{U}$, the operator $\Theta'(\bar{u}): \mathcal{U} \rightarrow C([0, T]; H)$ is defined by $\Theta'(\bar{u})h := z$, where z solves (24). \square \square

Next, we introduce the two operators $A^*: L^2(0, T; V) \rightarrow L^2(0, T; V')$ and $\tilde{B}: L^2(0, T; V) \rightarrow L^2(0, T; L^2(\Omega; \mathbb{R}^d))$ such that $A^*z := -\sum_{i,j=1}^d a_{ij} \partial_{ij}^2 z$ and $\tilde{B}(z) := \nabla z$, respectively, where ∇ denotes the gradient with respect to $x \in \mathbb{R}^d$. Observe that, for every $v, \varphi \in L^2(0, T; V)$,

$$\int_0^T \langle A^*v(t), \varphi(t) \rangle_{V', V} dt = \int_0^T \langle A\varphi(t), v(t) \rangle_{V', V} dt$$

and for every $u \in \mathcal{U}, v \in L^2(0, T; V)$ and $w \in L^\infty(0, T; H)$,

$$\begin{aligned} \int_0^T (b(u) \cdot \tilde{B}(v), w)_H dt &= \iint_Q \sum_{i=1}^d b_i(u) w \partial_i v dx dt \\ &= - \int_0^T \langle B(u(t), w(t)), v(t) \rangle_{V', V} dt, \end{aligned} \quad (25)$$

and the above integrals are well-defined.

With this in mind, we can provide an explicit representation formula for the derivative of J as in Corollary 9.

Proposition 12

Consider J of the form (22) with $y_d \in L^q(0, T; L^\infty(\Omega))$ and $y_\Omega \in H$. Let $y_0 \in L^\infty(\Omega)$. Then, J is differentiable in \mathcal{U} and, for all $u, h \in \mathcal{U}$,

$$dJ(u)h = \sum_{i=1}^d \iint_Q h_i [y \partial_i p + \lambda u_i] dx dt \quad (26)$$

holds, where $y \in W(0, T) \cap L^\infty(Q)$ is the solution of $\mathcal{E}(y_0, u, 0)$ and $p \in W(0, T)$ is the solution of the adjoint equation

$$\begin{cases} -\dot{p}(t) + A^*p(t) - b(u(t)) \cdot \tilde{B}p(t) = \alpha [y(t) - y_d(t)] & \text{in } V', \ t \in]0, T[, \\ p(T) = \beta [y(T) - y_\Omega] . \end{cases} \quad (27)$$

Observe that, for all $i = 1, \dots, d$, the function $h_i \langle \partial_i p, y \rangle_{V', V} :]0, T[\rightarrow \mathbb{R}$ belongs to $L^1(0, T)$, owing to $h_i \in L^q(0, T; L^\infty(\Omega))$ with $q > 2$, $y \in L^2(0, T; V)$ and $\partial_i p \in L^\infty(0, T; V')$. Existence and uniqueness of solutions for (27) are ensured by Aronson [4]. Indeed, $y_0 \in L^\infty(\Omega)$ implies $y \in L^\infty(Q)$, and therefore $y - y_d \in L^q(0, T; L^\infty(\Omega))$, as required by Theorem 3. Moreover, we have $y(T) - y_\Omega \in L^2(\Omega)$. By the change of variable $q(t) = p(T - t)$, $v(t) = u(T - t)$ and $f(t) = \alpha [y(T - t) - y_d(T - t)]$, (27) is recast in a form such that the same results as in Theorem 3 and Proposition 4 can be applied, following the remark in [4, page 621] concerning the adjoint operator.

Proof. Thanks to Lemma 11, the functional J is differentiable in \mathcal{U} . Let z solution of (24). We set $z = \Theta'(u)h \in C([0, T]; H)$ and derive

$$\begin{aligned} dJ(u)h &= \langle z, \alpha[y - y_d] \rangle_{L^2(0, T; H)} + \langle z(T), \beta[y(T) - y_\Omega] \rangle_H \\ &\quad + \lambda \langle h, u \rangle_{L^2(0, T; L^2(\Omega; \mathbb{R}^d))} \end{aligned}$$

for all $u, h \in \mathcal{U}$, where y is the solution of the state equation $\mathcal{E}(y_0, u, 0)$. We now exploit the adjoint state p in order to figure out the dependence of $dJ(u)h$ on h . Indeed, owing to relations (25) and (27), we have that

$$\begin{aligned} &\int_0^T \langle z(t), \alpha[y(t) - y_d(t)] \rangle_H dt \\ &= \int_0^T \langle -\dot{p}(t) + A^*p(t) - b(u(t)) \cdot \tilde{B}p(t), z(t) \rangle_{V', V} dt \\ &= \int_0^T \langle \dot{z}(t) + Az(t) + B(u(t), z(t)), p(t) \rangle_{V', V} dt - \langle z(T), p(T) \rangle_H + \langle z(0), p(0) \rangle_H \\ &= -\langle z(T), p(T) \rangle_H - \int_0^T \langle D(h(t), y(t)), p(t) \rangle_{V', V} dt \\ &= -\langle z(T), p(T) \rangle_H + \iint_Q y(t)h(t) \cdot \nabla p(t) \, dx dt . \end{aligned}$$

Since $\langle z(T), \beta[y(T) - y_\Omega] \rangle_H = \langle z(T), p(T) \rangle_H$, we conclude

$$dJ(u)h = \iint_Q y h \cdot \nabla p \, dx dt + \lambda \sum_{i=1}^d \langle h_i, u_i \rangle_{L^2(Q)} = \sum_{i=1}^d \iint_Q h_i [y \partial_i p + \lambda u_i] \, dx dt ,$$

which is the assertion. □

A priori, $dJ(u)$ is defined only in \mathcal{U} for every $u \in \mathcal{U}$. However, thanks to the representation formula (26), it admits an extension operator that is well defined on $L^2(0, T; L^2(\Omega; \mathbb{R}^d))$. With this, Proposition 12, and the variational inequality $dJ(\bar{u})(u - \bar{u}) \geq 0$, which holds for any $u \in \mathcal{U}_{ad}$ and locally optimal solution \bar{u} , we deduce the system of first order necessary optimality conditions. Note that, since the control-to-state operator is nonlinear, the reduced cost functional is non-convex even for standard quadratic costs like (22). In particular, there may be controls that are not optimal, not even locally, and nevertheless satisfy the necessary optimality conditions. Yet, this system plays an important role in the development of efficient numerical methods for the FP optimal control; see [2, 3]. Moreover, the simulations in [2, 3, 12] suggest that these conditions are viable to be used in practice.

Corollary 13

Let $y_0 \in L^\infty(\Omega)$, $y_d \in L^q(0, T; L^\infty(\Omega))$, and $y_\Omega \in H$. Consider the cost functional J defined by (22) with $\alpha, \beta, \lambda \geq 0$. An optimal pair $(\bar{y}, \bar{u}) \in C([0, T]; H) \times \mathcal{U}_{ad}$ for J with corresponding adjoint state \bar{p} satisfies the following necessary conditions:

$$\begin{aligned}
& \partial_t \bar{y} - \sum_{i,j=1}^d \partial_{ij}^2 (a_{ij} \bar{y}) + \sum_{i=1}^d \partial_i ((\gamma_i + \bar{u}_i) \bar{y}) = 0 & \text{in } Q, \\
& -\partial_t \bar{p} - \sum_{i,j=1}^d a_{ij} \partial_{ij}^2 \bar{p} - \sum_{i=1}^d (\gamma_i + \bar{u}_i) \partial_i \bar{p} = \alpha [\bar{y} - y_d] & \text{in } Q, \\
& \bar{y} = \bar{p} = 0 & \text{on } \Sigma, \\
& \bar{y}(\cdot, 0) = y_0(\cdot), \quad \bar{p}(\cdot, T) = \beta [\bar{y}(\cdot, T) - y_\Omega(\cdot)] & \text{in } \Omega, \\
& \iint_Q [\bar{y} \partial_i \bar{p} + \lambda \bar{u}_i] (u_i - \bar{u}_i) dx dt \geq 0 & \forall u \in \mathcal{U}_{ad}, i = 1, \dots, d.
\end{aligned} \tag{28}$$

Proof. The necessary optimality conditions (28) are derived by combining the state equation $\mathcal{E}(y_0, \bar{u}, 0)$ for \bar{y} , (27) for the adjoint \bar{p} , and the variational inequality $dJ(\bar{u})(u - \bar{u}) \geq 0$ for all $u \in \mathcal{U}_{ad}$ and locally optimal \bar{u} . Thanks to (12), (13), and (25), which define the operators A , B , and \tilde{B} , respectively, we deduce the desired system. \square \square

Following [32, Section 2.8], we can derive pointwise conditions for the variational inequality in (28). Indeed, if $\lambda = 0$, it follows for all $i = 1, \dots, d$ and almost all $(x, t) \in Q$ that,

$$\bar{u}_i(x, t) = \begin{cases} u_{a_i}, & \text{if } \bar{y}(x, t) \partial_i \bar{p}(x, t) > 0, \\ u_{b_i}, & \text{if } \bar{y}(x, t) \partial_i \bar{p}(x, t) < 0, \end{cases}$$

and no value can be assigned if $\bar{y}(x, t) \partial_i \bar{p}(x, t) = 0$. If $\lambda > 0$, then we get the standard projection formula for almost all $(x, t) \in Q$:

$$\bar{u}_i(x, t) = \mathbb{P}_{[u_{a_i}, u_{b_i}]} \left\{ -\frac{1}{\lambda} \bar{y}(x, t) \partial_i \bar{p}(x, t) \right\}.$$

In case of time-independent controls considered in Remark 10, the only modification needed in the optimality system (28) is the variational inequality, which, for $\tilde{\mathcal{U}}_{ad}$ given by (23), changes to

$$\int_{\Omega} \left[\int_0^T \bar{y} \partial_i \bar{p} \, dt + \lambda \bar{u}_i \right] (u_i - \bar{u}_i) \, dx \geq 0 \quad \forall u \in \tilde{\mathcal{U}}_{ad}, i = 1, \dots, d.$$

7 Conclusions

In this paper, we considered a bilinear optimal control problem subject to the Fokker-Planck equation with a time- and space-dependent control. We proved existence of optimal controls associated with a non-negative state solution and derived the first order necessary optimality conditions rigorously, thereby extending the results of [3]. Since the problem is non-convex due to the nonlinear control-to-state operator, the task to find sufficient conditions is still open. Another open question is the uniqueness of the optimal control.

Acknowledgements

The authors wish to express their gratitude to Lars Grüne for suggesting them this very interesting subject and for many helpful comments. They would also like to thank Alfio Borzì for very helpful discussions and the referees for their valuable comments that helped to improve the manuscript. This work was partially supported by the EU under the 7th Framework Program, Marie Curie Initial Training Network FP7-PEOPLE-2010-ITN SADCO, GA 264735-SADCO, by the DFG project Model Predictive Control for the Fokker-Planck Equation, GR 1569/15-1, and by the INdAM through the GNAMPA Research Project 2015 "Analisi e controllo di equazioni a derivate parziali nonlineari". Most of the results proved in this paper have been announced in a less general setting in a proceedings of the 2nd IFAC Workshop on Control of Systems Governed by Partial Differential Equations.

References

- [1] A. Addou and A. Benbrik. Existence and uniqueness of optimal control for a distributed-parameter bilinear system. *J. Dynam. Control Systems*, 8(2):141–152, 2002.
- [2] M. Annunziato and A. Borzì. Optimal control of probability density functions of stochastic processes. *Math. Model. Anal.*, 15(4):393–407, 2010.
- [3] M. Annunziato and A. Borzì. A Fokker-Planck control framework for multidimensional stochastic processes. *J. Comput. Appl. Math.*, 237(1):487–507, 2013.
- [4] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 22:607–694, 1968.

- [5] D. G. Aronson and James Serrin. Local behavior of solutions of quasilinear parabolic equations. *Arch. Rational Mech. Anal.*, 25:81–122, 1967.
- [6] Jean-Pierre Aubin. Un théorème de compacité. *C. R. Acad. Sci. Paris*, 256:5042–5044, 1963.
- [7] A. Blaquière. Controllability of a Fokker-Planck equation, the Schrödinger system, and a related stochastic optimal control (revised version). *Dynam. Control*, 2(3):235–253, 1992.
- [8] Roger Brockett. New issues in the mathematics of control. In *Mathematics unlimited—2001 and beyond*, pages 189–219. Springer, Berlin, 2001.
- [9] Eduardo Casas. Pontryagin’s principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.*, 35(4):1297–1327, 1997.
- [10] William Feller. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.*, 77:1–31, 1954.
- [11] Alessio Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
- [12] Arthur Fleig, Lars Grüne, and Roberto Guglielmi. Some results on model predictive control for the Fokker-Planck equation. In *MTNS 2014 : 21st International Symposium on Mathematical Theory of Networks and Systems, July 7-11, 2014*, University of Groningen, The Netherlands, pages 1203–1206, 2014.
- [13] Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control*. Springer-Verlag, Berlin-New York, 1975. Applications of Mathematics, No. 1.
- [14] Michael G. Forbes, J. Fraser Forbes, and Martin Guay. Regulating discrete-time stochastic systems: focusing on the probability density function. *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms*, 11(1-2):81–100, 2004.
- [15] Crispin Gardiner. *Stochastic methods*. Springer Series in Synergetics. Springer-Verlag, Berlin, fourth edition, 2009. A handbook for the natural and social sciences.
- [16] Ī. Ī. Gihman and A. V. Skorohod. *Stochastic differential equations*. Springer-Verlag, New York-Heidelberg, 1972. Translated from the Russian by Kenneth Wickwire, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 72.
- [17] Werner Horsthemke and René Lefever. *Noise-induced transitions*, volume 15 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1984. Theory and applications in physics, chemistry, and biology.
- [18] Guy Jumarie. Tracking control of nonlinear stochastic systems by using path cross-entropy and Fokker-Planck equation. *Internat. J. Systems Sci.*, 23(7):1101–1114, 1992.

- [19] Miroslav Kárný. Towards fully probabilistic control design. *Automatica J. IFAC*, 32(12):1719–1722, 1996.
- [20] A. Kolmogoroff. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.*, 104(1):415–458, 1931.
- [21] O.A. Ladyzhenskaya, V.A. Solonnikov, and N.N. Ural'tseva. *Linear and quasilinear equations of parabolic type*. Moskva: Izdat. 'Nauka'. 736 pp., 1967.
- [22] C. Le Bris and P.-L. Lions. Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. *Comm. Partial Differential Equations*, 33(7-9):1272–1317, 2008.
- [23] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris, 1969.
- [24] J.-L. Lions. *Optimal control of systems governed by partial differential equations*. Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York-Berlin, 1971.
- [25] Alessio Porretta. On the planning problem for the mean field games system. *Dyn. Games Appl.*, 4(2):231–256, 2014.
- [26] Alessio Porretta. Weak solutions to Fokker-Planck equations and mean field games. *Arch. Ration. Mech. Anal.*, 216(1):1–62, 2015.
- [27] Serguei Primak, Valeri Kontorovich, and Vladimir Lyandres. *Stochastic methods and their applications to communications*. John Wiley & Sons, Inc., Hoboken, NJ, 2004.
- [28] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005.
- [29] J. P. Raymond and H. Zidani. Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Appl. Math. Optim.*, 39(2):143–177, 1999.
- [30] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, second edition, 1989.
- [31] Jacques Simon. Compact sets in the space $L^p(0, T; B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.
- [32] Fredi Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010.
- [33] Hong Wang. Robust control of the output probability density functions for multi-variable stochastic systems with guaranteed stability. *IEEE Trans. Automat. Control*, 44(11):2103–2107, 1999.
- [34] J. Wloka. *Partial differential equations*. Cambridge University Press, Cambridge, 1987.