Bilinear Optimal Control of the Fokker-Planck Equation ⋆

Arthur Fleig * **Roberto Guglielmi** **

* *Department of Mathematics, University of Bayreuth, Germany, (e-mail: arthur.fleig@uni-bayreuth.de).*
** *Radon Institute for Computational and Applied Mathematics, Austria (e-mail: roberto.guglielmi@ricam.oeaw.ac.at)*

Abstract: The optimal tracking problem of the probability density function of a stochastic process can be expressed in term of an optimal bilinear control problem for the Fokker-Planck equation, with the control in the coefficient of the divergence term. As a function of time and space, the control needs to belong to an appropriate Banach space. We give suitable conditions to establish existence of optimal controls and the associated first order necessary optimality conditions.

Keywords: control system analysis, optimal control, bilinear control, Fokker-Planck equation, stochastic control

1. INTRODUCTION

The study of the Fokker-Planck equation (FPE), also known as Kolmogorov forward equation, has received great and increasing interest starting from the work by Kolmogoroff (1931), owing to its relation with the description of the time evolution of the Probability Density Function (PDF) of the velocity of a particle. In Blaquière (1992), the analysis of the controllability properties of the FPE has been developed in connection with quantum system and stochastic control. In recent years, it has become of main interest in mean field game theory (see Porretta (2015) for further insights on this connection).

In a similar way, our main interest in the optimal control of the Fokker-Planck equation derives from its connection with the evolution of the PDF associated with a stochastic process. Given $T > 0$, let us consider a continuous-time stochastic process described by the (Itô) stochastic differential equation

$$\begin{aligned} dX_t &= b(X_t,t)\,\mathrm{d}t + \sigma(X_t,t)\,dW_t\,, \quad t \in (0,T)\,, \\ X(t=0) &= X_0\,, \end{aligned} \tag{1}$$

where $X_0 \in \mathbb{R}^d$ is the initial condition, $d \geq 1$, $dW_t \in \mathbb{R}^m$ is an $m-$dimensional Wiener process, $m \geq 1$, $b = (b_1, \ldots, b_m)$ is a vector valued drift function, and the dispersion matrix $\sigma(X_t,t) = (\sigma_{ij}) \in \mathbb{R}^{d\times m}$ is assumed to have full rank.

Assuming for simplicity that the state variable X_t evolves in a bounded domain Ω of $\mathbb{R}^d$ with smooth boundary, we define $Q := \Omega \times (0,T)$, $\Sigma := \partial\Omega \times (0,T)$, and $a_{ij} := \sigma_{ik}\sigma_{kj}/2$, $i, j = 1, \ldots d$, where here and in the following we use the Einstein summation convention. We denote by ∂_i and ∂_t the partial derivative with respect to x_i and t, respectively, where $i = 1, \ldots, d$.

⋆ 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".

Under suitable assumptions on the coefficients b and σ, it is well known, see (Primak et al., 2004, p. 227) and (Protter, 2005, p. 297) that, given an initial distribution ρ_0, the PDF associated with the stochastic process (1) evolves according to the following FPE

$$\partial_t\rho - \partial^2_{ij}(a_{ij}\rho) + \partial_i(b_i\rho) = 0\,, \qquad \text{in } Q\,, \tag{2}$$

$$\rho(x,0) = \rho_0(x)\,, \qquad \text{in } \Omega\,. \tag{3}$$

We refer to Risken (1989) for an exhaustive theory and numerical methods for the FPE. A solution ρ to (2)-(3) shall furthermore satisfy the standard properties of a PDF, i.e.,

$$\rho(x,t) \geq 0\,, \quad (x,t) \in Q\,,$$

$$\int_\Omega \rho(x,t)\,\mathrm{d}x = 1\,, \quad t \in (0,T)\,.$$

Consider now the presence of a control function acting on (1) through the drift term b,

$$dX_t = b(X_t,t;u)\,dt + \sigma(X_t,t)\,dW_t\,, \tag{4}$$

where the control has to be chosen from a suitable class of admissible functions in a way to minimize a certain cost functional. In the non-deterministic case of (4), the state evolution X_t represents a random variable. Therefore, when dealing with stochastic optimal control, usually the average of the cost function is considered, see for example Fleming and Rishel (1975). In particular, the cost functional is usually of the form

$$J(X,u) = \mathbb{E}\left[\int_0^T L(t,X_t,u(t))dt + \psi(X_T)\right]\,,$$

for suitable running cost L and terminal cost ψ.

On the other hand, the state of a stochastic process can be characterized by the shape of its statistical distribution, which is represented by the PDF. Therefore, a control methodology defined via the PDF provides an accurate and flexible control strategy that can accommodate a wide class of objectives, cf. (Brockett, 2001, Section 4). In this direction, in Forbes et al. (2004); Jumarie (1992); Kárný (1996); Wang (1999) 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 and no average is needed.

However, in these references, stochastic methods were still adopted in order to approximate the state variable X_t of the random process. On the other hand, in Annunziato and Borzì (2010, 2013) the authors approach directly the problem of tracking the PDF associated with the stochastic process. If the control acts through the drift term as in (4), the evolution of the PDF is controlled through the advection term of equation (2). This is a rather weak action of the controller on the system, usually called of bilinear type, since the control takes action as a coefficient of the state variable. Indeed, few controllability results are known for such a kind of control system (e.g. Blaquière, 1992; Porretta, 2014). Concerning the existence of bilinear optimal control, a first result was given by Addou and Benbrik (2002) for a control function which only depends on time. Relying on this result, in Annunziato and Borzì (2010, 2013) the tracking of a PDF governed by (2) has been studied with a time dependent control function.

Notice that, in general, the space domain in (2) is $\mathbb{R}^d$ instead of 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, we might focus on the description of the evolution of the PDF in the bounded domain $\Omega \subset \mathbb{R}^d$. Assuming that the physical structure of the problem ensures the confinement of the stochastic process within Ω, it is reasonable to employ homogeneous Dirichlet boundary conditions

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

also known as absorbing boundary conditions (Primak et al., 2004, page 231) (see also Feller (1954) for a complete characterization of possible boundary conditions in dimension one).

The aim of this work is to extend the theoretical study on the existence of bilinear optimal controls of the FPE by Addou and Benbrik (2002) to the case of more general control functions, i.e., to the case of a bilinear control which depends both on time and space. In connection with our motivation from stochastic optimal control, on the one hand, a simpler controller $u = u(t)$ would be easier to implement in some applications. On the other hand, in certain situations it could be handier or even required to act on the space variable as well. In general, the richer structure of a control $u = u(x,t)$ allows to substantially improve the tracking performance of a PDF, as shown in Fleig et al. (2014). For a more detailed presentation of the results and their proofs in the current work, we refer to Fleig and Guglielmi (2015).

In the sequel, following Aronson (1968), we introduce proper assumptions on the functional framework to ensure existence of solutions to state equation of the form (2) in Section 2. Section 3 is devoted to recast the FPE in an abstract setting and to deduce useful a-priori estimates on its solution. The main result on existence of solutions to the optimal control problem is presented in Section 4, whereas in Section 5 we deduce the system of first order necessary optimality conditions that characterizes the optimal solutions. Section 6 concludes.

2. EXISTENCE OF SOLUTION TO THE FPE

In this section, we describe the functional framework that we will use to ensure the existence of solutions to

$$\partial_t y - \partial_{ij}^2 (a_{ij} y) + \partial_i (b_i(u) y) = f \qquad \text{in } Q, \qquad (5)$$

which, assuming that $a_{ij} \in C^1(\overline{Q})$ for all $i,j = 1,\dots,d$, and setting $\tilde{b}_j(u) := \partial_i a_{ij} - b_j(u)$, can be recast in the flux formulation

$$\partial_t y - \partial_j \left(a_{ij} \partial_i y + \tilde{b}_j(u) y \right) = f \qquad \text{in } Q, \qquad (6)$$

with initial and boundary conditions

$$y(x,t) = 0, \qquad (x,t) \in \Sigma, \qquad (7)$$

$$y(x,0) = y_0(x) \in L^2(\Omega), \qquad x \in \Omega, \qquad (8)$$

and associated variational formulation

$$\iint_Q f v = \iint_Q \partial_t y v - \iint_Q \left(\partial_j \left(a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \right) v$$
$$= - \iint_Q y \partial_t v - \int_\Omega y(\cdot,0) v(\cdot,0) + \iint_Q \left(a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \partial_j v$$

for any test function $v \in W_2^{1,1}(Q)$ with $v_{|\partial\Omega} = 0$ and $v(\cdot,T) = 0$, where the differentials $\mathrm{d}x$ and $\mathrm{d}t$ have been omitted for readability.

Here and in the following sections we assume the hypotheses

Assumption 1. (1) $a_{ij} \in C^1(\overline{Q})$ for all $i,j = 1,\dots,d$.

(2) $\forall \xi \in \mathbb{R}^d$ and almost all $(x,t) \in Q$:
  - (a) $a_{ij}(x,t) \xi_i \xi_j \geq \theta |\xi|^2$ for some $0 < \theta < \infty$.
  - (b) $|a_{ij}(x,t)| \leq M, i,j = 1,...,d$, for some $0 < M < \infty$.

(3) $f, \tilde{b}_j(u) \in L^q(0,T;L^\infty(\Omega)), j = 1,...,d$, with $2 < q \leq \infty$.

Under this assumption, a result by (Aronson, 1968, Thm. 1, p. 634) ensures the existence and uniqueness of (nonnegative) solutions to equation (6).

Theorem 2. (Existence of nonnegative solutions). Suppose that Assumption 1 holds and let $y_0 \in L^2(\Omega)$. Then there exists a unique $y \in L^2(0,T;H_0^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ satisfying

$$\iint_Q \left[-y \partial_t v + \left(a_{ij} \partial_i y + \tilde{b}_j(u) y \right) \partial_j v - f v \right] = \int_\Omega y_0 v(\cdot,0)$$

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 (6)-(8). Moreover, if $f \equiv 0$ and $0 \leq y_0 \leq m$ almost everywhere in Ω, then

$$0 \leq y(x,t) \leq m(1 + C_{FP} k) \quad \text{almost everywhere in } Q,$$

where

$$k := \sum_{j=1}^d \left\| \tilde{b}_j(u) \right\|_{L^q(0,T;L^\infty(\Omega))}$$

and the constant $C_{FP} > 0$ depends only on T, Ω, and the structure of the FPE.

Remark 3. If the right-hand-side in (6) is of the form $f = \operatorname{div}(F)$ with $F\colon Q \to \mathbb{R}^d$, Theorem 2 remains true assuming that $F_i \in L^2(Q)$, $i = 1,...,d$, see Aronson (1968).

The solution obtained by Theorem 2 is more regular. To this end, let us consider the Gelfand triple $V \hookrightarrow H \hookrightarrow V'$, with $H := L^2(\Omega)$, $V := H_0^1(\Omega)$, and $V' = H^{-1}(\Omega)$ the dual space of V, endowed with norms

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

respectively, where $\langle .,. \rangle_{V',V}$ represents the duality map between V and V'. We remind that

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

$\dot{y}$ denoting the time derivative of y.

Proposition 4. Under the assumptions of Theorem 2, the solution y to problem (6)-(8) belongs to $W(0,T)$, possibly after a modification on a set of measure zero.

For brevity, in the following we will refer to the space $L^p(0,T;X)$ simply by $L^p(X)$, for any $p \in [1,+\infty]$ and X Banach space.

3. A-PRIORI ESTIMATES

In this section, we deduce a-priori estimates of solutions to the FPE (5),(7),(8) with $f \in L^2(0,T;V')$. For the sake of clarity, we recast it in abstract form

$$\begin{cases} \dot{y}(t) + Ay(t) + B(u(t),y(t)) = f(t) \text{ in } V',\ t \in (0,T) \\ y(0) = y_0, \end{cases} \tag{9}$$

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

$$\langle Az, \varphi\rangle_{V',V} = \int_\Omega a_{ij}\partial_i z\, \partial_j \varphi \,\mathrm{d}x \qquad \forall z, \varphi \in V,$$

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

$$\langle B(u,y), \varphi\rangle_{V',V} = -\int_\Omega b_i(u) y\, \partial_i \varphi \,\mathrm{d}x = -\int_\Omega y b(u).\nabla\varphi \,\mathrm{d}x$$

for all $u \in L^\infty(\Omega;\mathbb{R}^d)$, $y \in H$, $\varphi \in V$. In the following, $\mathscr{E}(y_0,u,f)$ refers to (9) whenever we want to point out the data (y_0,u,f).

From this section on, we denote by M and C generic positive constants that might change from line to line, and we assume the following properties.

Assumption 5. (1) The coefficient functions $a_{ij}(x,t)$ are positive constants, i.e., $a_{ij} > 0$.

(2) The function $b\colon \mathbb{R}^{d+1} \times \mathscr{U} \to \mathbb{R}^d, (x,t;u) \mapsto b(x,t;u(x,t))$ satisfies the growth condition

$$\sum_{i=1}^{d} |b_i(x,t;u)|^2 \le M(1+|x|^2+|u(x,t)|^2) \quad \forall x \in \mathbb{R}^d, \tag{10}$$

for every $i = 1,\dots,d$, $t \in [0,T]$, and u in a suitable space $\mathscr{U}$ of admissible controls.

We assume for simplicity the coefficients a_{ij} to be constant, that is, the operator A is self-adjoint, in order to focus more specifically on the bilinear action of the control through the divergence term. However, it shall be possible to extend the analysis to the case of general diffusion coefficients satisfying Assumption 1(i)-(ii).

In this setting, $u(t) \in L^\infty(\Omega;\mathbb{R}^d)$ implies $b(t;u(t)) \in L^\infty(\Omega;\mathbb{R}^d)$, which occurs, in particular, in the case

$$b_i(x,t;u) = \gamma_i(x) + u_i(x,t), \qquad (x,t) \in Q,$$

for some $\gamma_i \in C^1(\Omega)$ and $u_i(\cdot,t) \in L^\infty(\Omega)$, $i = 1,\dots,d$. Furthermore, relation (10) ensures that

$$\|B(u,y)\|_{V'} \le M(1+\|u\|_{L^\infty(\Omega;\mathbb{R}^d)})\,\|y\|_H$$

for any $u \in L^\infty(\Omega;\mathbb{R}^d)$ and $y \in H$. Given $q > 2$, admissible controls are functions

$$u \in \mathscr{U} := L^q(0,T;L^\infty(\Omega;\mathbb{R}^d)) \subset L^2(0,T;L^\infty(\Omega;\mathbb{R}^d)),$$

for which we have

$$\|u\|_{L^2(0,T;L^\infty(\Omega;\mathbb{R}^d))} \le T^{\frac{q-2}{2q}}\,\|u\|_{\mathscr{U}}.$$

To ease the notation, we still denote by A and B the operators $A\colon L^2(V) \to L^2(V')$ and $B\colon \mathscr{U} \times L^\infty(H) \to L^q(V')$ with $1/q + 1/q' = 1$, such that, respectively,

$$Az = -\partial_j(a_{ij}\partial_i z) \qquad \forall z \in L^2(V)$$

and

$$B(u,y) = \partial_i(b_i(u)y) = \mathrm{div}(b(u)y) \quad \forall u \in \mathscr{U},\, y \in L^\infty(H).$$

Indeed, for every $u \in \mathscr{U}$ and $y \in L^\infty(H)$ we have that $\mathrm{div}(b(u)y) \in L^q(V')$ and

$$\|B(u,y)\|_{L^q(V')} = \|\mathrm{div}(b(u)y)\|_{L^q(V')} \le M(1+\|u\|_{\mathscr{U}})\,\|y\|_{L^\infty(H)}.$$

The next result gives some useful a-priori estimates on the solution to (9).

Lemma 6. Let $y_0 \in H$, $f \in L^2(V')$ and $u \in \mathscr{U}$. Then a solution y to (9) satisfies the estimates

$$\begin{aligned} \|y\|^2_{L^\infty(H)} &\le M(u)\left(\|y_0\|_H^2 + \|f\|^2_{L^2(V')}\right), \\ \|y\|^2_{L^2(V)} &\le (1+\|u\|^2_{\mathscr{U}})M(u)\left(\|y_0\|_H^2 + \|f\|^2_{L^2(V')}\right), \\ \|\dot{y}\|^2_{L^2(V')} &\le 2\,\|f\|^2_{L^2(V')} + \\ &\qquad (1+\|u\|^2_{\mathscr{U}})M(u)\left(\|y(0)\|_H^2 + \|f\|^2_{L^2(V')}\right), \end{aligned}$$

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

4. EXISTENCE OF OPTIMAL CONTROLS

In this section, we consider the minimization of a cost functional $\tilde{J}(y,u)$, where the state y is subject to equation (9) with control u and source $f \equiv 0$. We require Assumptions 1 and 5 to hold in this and the following sections.

Fixing $y_0 \in H$, we introduce the control-to-state operator $\Theta\colon \mathscr{U} \to C([0,T];H)$ such that $u \mapsto y \in C([0,T];H)$ solution of $\mathscr{E}(y_0,u,0)$. Thus, the optimization problem turns into minimizing the so-called reduced cost functional $J(u) := \tilde{J}(\Theta(u),u)$, which we assume to be bounded from below, over a suitable non-empty subset of admissible controls $\mathscr{U}_{ad}$. Without loss of generality, we assume the existence of a control $\tilde{u} \in \mathscr{U}_{ad}$ such that $J(\tilde{u}) < \infty$. In the following, we consider the usual box constraints for the space of admissible controls, i.e.,

$$\mathscr{U}_{ad} := \{u \in \mathscr{U} : u_a \le u(x,t) \le u_b \quad \text{for a.e. } (x,t) \in Q\}, \tag{11}$$

where $u_a, u_b \in \mathbb{R}^d$ and $u_a \le u_b$ is to be understood componentwise. In order to prove the main theorem we will need the following compactness result (see Aubin (1963), (Lions, 1969, Théorème 5.1, page 58) or (Simon, 1987, Corollary 4)).

Theorem 7. 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 8. Let $y_0 \in H$ and assume $b(x;u) = (\gamma_i(x) + u_i(x,t))_i$ for some $\gamma_i \in C^1(\Omega)$, $i = 1,\dots,d$. Consider the reduced cost functional $J(u) = \tilde{J}(\Theta(u),u)$, to be minimized over the controls $u \in \mathscr{U}_{ad}$. Assume that J is bounded from below and (sequentially) weakly-star lower semicontinuous. Then there exists a pair $(\bar{y},\bar{u}) \in C([0,T];H) \times \mathscr{U}_{ad}$ such that $\bar{y}$ solves $\mathscr{E}(y_0,\bar{u},0)$ and $\bar{u}$ minimizes J in $\mathscr{U}_{ad}$.

Remark 9. Requiring box constraints as in (11) might seem a too restrictive choice. However, we 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, see for example (Lions, 1971, Section 15.3, p. 237). Theorem 8 clearly also holds for any $\mathscr{U}_{ad}$ that is a bounded weakly-star closed subset of $\mathscr{U}$. However, note that in the unconstrained case $\mathscr{U}_{ad} \equiv \mathscr{U}$, asking only $J(u) \geq \lambda \|u\|_{\mathscr{U}}$ for some $\lambda > 0$ is not enough. Instead, a condition of the type $J(u) \geq \lambda \|u\|_{L^\infty(Q)}$ would allow to prove the existence of optimal controls. However, this kind of condition is not very practical in applications.

Corollary 10. Assume that $b(x,t;u) = (\gamma_i(x) + u_i(x,t))_i$ for some $\gamma_i \in C^1(\Omega)$, $i = 1,\dots,d$, with $u \in \mathscr{U}_{ad}$ as in (11), and let $y_d \in L^2(0,T;H)$, $y_\Omega \in H$, $\alpha,\beta,\lambda \geq 0$ with $\max\{\alpha,\beta\} > 0$. Then an optimal pair $(\bar{y},\bar{u}) \in C([0,T];H) \times \mathscr{U}_{ad}$ exists for the reduced cost functional $J(u)$ defined by

$$\frac{\alpha}{2}\|y - y_d\|^2_{L^2(H)} + \frac{\beta}{2}\|y(T) - y_\Omega\|^2_H + \frac{\lambda}{2}\|u\|^2_{L^2(H)}, \tag{12}$$

where $y = \Theta(u)$.

Remark 11. If one wants to use the cost functional (12) without imposing box constraints on the control, e.g., $\mathscr{U}_{ad} \equiv \mathscr{U}$, one shall require more regularity on the state y and on the control u, in order to gain the same level of compactness required in the proof of Theorem 8. Indeed, further regularity of y can be ensured by standard improved regularity results, see for example (Wloka, 1987, Theorems 27.2 and 27.5) and (Ladyzhenskaya et al., 1967, Theorem 6.1 and Remark 6.3). However, these results come at the price of requiring more regularity of the coefficients in the PDE, which, in our case, translates to more regularity of the control. In particular, one would need to require differentiability of u both in time and space, which is a feature that is scarcely ever satisfied in the numerical simulations.

Remark 12. Corollary 10 applies analogously to the case of time-independent controls in the admissible space

$$\tilde{\mathscr{U}}_{ad} := \{u \in L^\infty(\Omega) : u_a \leq u(x) \leq u_b \quad \text{for a.e. } x \in \Omega\} \tag{13}$$

for some $u_a, u_b \in \mathbb{R}^d$ such that $u_a \leq u_b$ (component-wise), and the reduced cost functional $J_2(u)$ given by

$$\frac{\alpha}{2}\|y - y_d\|^2_{L^2(H)} + \frac{\beta}{2}\|y(T) - y_\Omega\|^2_H + \frac{\lambda}{2}\|u\|^2_H,$$

where $y = \Theta(u)$.

5. ADJOINT STATE AND OPTIMALITY CONDITIONS

In this section, we consider b and B such that $b(u) = u$ and

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

respectively. This choice does not affect the generality of the problem. Indeed, for b as in Theorem 8, assuming $\max_i\{\gamma_i, \gamma_i'\}$ sufficiently small, we can include the contribution $\operatorname{div}(\gamma y)$ in the operator A, which becomes

$$A_\gamma z := Az + \operatorname{div}(\gamma z)$$

that still satisfies the assumptions required on A.

Thanks to the estimates given by Lemma 6, we deduce the following result.

Lemma 13. Let $y_0 \in H$. Then the control-to-state map Θ is differentiable in the Fréchet sense, and for every $\bar{u}$, $h \in \mathscr{U}$ the function $\Theta'(\bar{u})h$ satisfies

$$\begin{cases} \dot{z}(t) + Az(t) + B(\bar{u}(t), z(t)) = -B(h(t), \bar{y}(t)) \ \text{ in } V', \\ z(0) = 0, \end{cases} \tag{14}$$

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

Thanks to Remark 3, Theorem 2 ensures the existence of a unique weak solution of equation (14).

We introduce the operator $\tilde{B}\colon L^2(V) \to L^2(L^2(\Omega;\mathbb{R}^d))$ such that $\tilde{B}(v) = \nabla_x v$ for all $v \in L^2(V)$, where ∇_x denotes the gradient with respect to the space variable $x \in \mathbb{R}^d$. For every $u \in \mathscr{U}$, $v \in L^2(V)$, and $w \in L^\infty(H)$, we have that

$$\int_0^T \big(b(u).\tilde{B}(v), w\big)_H \,\mathrm{d}t = \iint_Q b_i(u) w\, \partial_i v \,\mathrm{d}x\mathrm{d}t = -\int_0^T \langle B(u(t), w(t)), v\rangle_{V',V} \,\mathrm{d}t$$

and the above integrals are well-defined.

In the sequel, we give the first order necessary optimality conditions for the cost functional J as in (12). We start by deducing an explicit representation formula for the derivative of J. Incidentally, let us point out that J is one of the objective functionals most commonly used in the numerical simulations, see, for example, Annunziato and Borzì (2013); Fleig et al. (2014).

Proposition 14. Let $y_d \in L^q(0,T;L^\infty(\Omega))$, $y_\Omega \in L^2(\Omega)$, and $y_0 \in L^\infty(\Omega)$. Then the functional J given by (12) is differentiable in $\mathscr{U}$ and, for all $u, h \in \mathscr{U}$,

$$\mathrm{d}J(u)h = \iint_Q h_i(t)\,[y(t)\partial_i p(t) + \lambda u_i(t)]\,\mathrm{d}x\mathrm{d}t \tag{15}$$

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

$$\begin{cases} -\dot{p}(t) + Ap(t) - b(u(t)).\tilde{B}p(t) = \alpha\,[y(t) - y_d(t)] \ \text{ in } V', \\ p(T) = \beta\,[y(T) - y_\Omega]\,. \end{cases} \tag{16}$$

Let us observe that the function $h_i\langle\partial_i p, y\rangle_{V',V}\colon (0,T) \to \mathbb{R}$ belongs to $L^1(0,T)$ for all $i = 1,\dots,d$, owing to $h_i \in L^q(L^\infty(\Omega))$ with $q > 2$, $y \in L^2(V)$ and $\partial_i p \in L^\infty(V')$. Moreover, the existence and uniqueness of solutions for equation (16) is ensured by Theorem 2. Indeed, $y_0 \in L^\infty(\Omega)$ gives $y \in L^\infty(Q)$, thus $y - y_d \in L^q(L^\infty(\Omega))$ as required by Assumption 1. Furthermore, $y(T) - y_\Omega \in L^2(\Omega)$. Therefore, by the change of variable $q(t) = p(T-t)$, $v(t) = u(T-t)$ and $f(t) = \alpha[y(T-t) - y_d(T-t)]$, equation (16) is recast in a form similar to (9) such that Theorem 2 and Proposition 4 can be applied. In addition, if $y_\Omega \in L^\infty(\Omega)$ we conclude that $p \in W(0,T) \cap L^\infty(Q)$, see (Aronson, 1968, Theorem 1, p. 634).

We note that, a priori, for every $u \in \mathscr{U}$, $\mathrm{d}J(u)$ is defined in $\mathscr{U}$. However, thanks to the representation formula (15), it may be extended to a map defined on $L^2(L^2(\Omega;\mathbb{R}^d))$.

As a consequence of Proposition 14 and the variational inequality $\mathrm{d}J(\bar{u})(u - \bar{u}) \geq 0$ for any $u \in \mathscr{U}_{ad}$ and locally optimal solution $\bar{u}$, we deduce the first order necessary optimality conditions, formulated in the next result.

Corollary 15. Let $y_d \in L^q(0,T;L^\infty(\Omega))$, $y_\Omega \in L^2(\Omega)$, and $y_0 \in L^\infty(\Omega)$. Consider the cost functional J defined by (12) with $\alpha,\beta,\gamma \geq 0$ and $\max\{\alpha,\beta\} > 0$. Then an optimal pair $(\bar{y},\bar{u}) \in C([0,T];H) \times \mathscr{U}_{ad}$ for J with corresponding adjoint state $\bar{p}$ is characterized by the following necessary conditions:

$$
\begin{aligned}
&\partial_t \bar{y} - a_{ij}\partial^2_{ij}\bar{y} + \partial_i\big(\bar{u}_i\bar{y}\big) = 0\,, && \text{in } Q\,,\\
&-\partial_t \bar{p} - a_{ij}\partial^2_{ij}\bar{p} - \bar{u}_i\partial_i\bar{p} = \alpha[\bar{y} - y_d]\,, && \text{in } Q\,,\\
&\bar{y} = \bar{p} = 0 && \text{on } \Sigma\,,\\
&\bar{y}(0) = y_0\,, \quad \bar{p}(T) = \beta[\bar{y}(T) - y_\Omega]\,, && \text{in } \Omega\,,\\
&\iint\limits_Q [\bar{y}\partial_i\bar{p} + \lambda\bar{u}_i]\,(u_i - \bar{u}_i)\,\mathrm{d}x\mathrm{d}t \geq 0 && \forall u \in \mathscr{U}_{ad}\,.
\end{aligned}
\tag{17}
$$

Remark 16. In the case of time-independent control as in Remark 12, the only modification needed in the optimality system (17) is the variational inequality, which changes to

$$
\int_\Omega \left[\int_0^T \bar{y}\partial_i\bar{p}\,\mathrm{d}t + \lambda\bar{u}_i\right](u_i - \bar{u}_i)\,\mathrm{d}x \geq 0 \quad \forall u \in \tilde{\mathscr{U}}_{ad}\,,
$$

where $\tilde{\mathscr{U}}_{ad}$ is given by (13).

6. CONCLUSION

For the controlled Fokker-Planck equation with a space-dependent control $u(x,t)$ acting on the drift term we have established theoretical results regarding the existence of optimal controls and necessary optimality conditions. Compared to just time-dependent controls $u(t)$, where the PDF can only be moved as a whole, space-dependent control allows to consider a much wider class of objectives since every single point of the PDF may be moved independently. When applying the calculated optimal control directly to the stochastic process, this results in a feedback loop, which may be interesting to a variety of applications, e.g., fluid flow, quantum control, or finance.

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.

REFERENCES

Addou, A., Benbrik, A., 2002. Existence and uniqueness of optimal control for a distributed-parameter bilinear system. J. Dynam. Control Systems 8 (2), 141–152.

Annunziato, M., Borzì, A., 2010. Optimal control of probability density functions of stochastic processes. Math. Model. Anal. 15 (4), 393–407.

Annunziato, M., Borzì, A., 2013. A Fokker-Planck control framework for multidimensional stochastic processes. J. Comput. Appl. Math. 237 (1), 487–507.

Aronson, D. G., 1968. Non-negative solutions of linear parabolic equations. Ann. Scuola Norm. Sup. Pisa (3) 22, 607–694.

Aubin, J.-P., 1963. Un théorème de compacité. C. R. Acad. Sci. Paris 256, 5042–5044.

Blaquière, A., 1992. Controllability of a Fokker-Planck equation, the Schrödinger system, and a related stochastic optimal control (revised version). Dynam. Control 2 (3), 235–253.

Brockett, R., 2001. New issues in the mathematics of control. In: Mathematics unlimited—2001 and beyond. Springer, Berlin, pp. 189–219.

Feller, W., 1954. Diffusion processes in one dimension. Trans. Amer. Math. Soc. 77, 1–31.

Fleig, A., Grüne, L., Guglielmi, R., 2014. 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. pp. 1203–1206.

Fleig, A., Guglielmi, R., 2015. Bilinear optimal control of the Fokker-Planck equation. preprint.

Fleming, W. H., Rishel, R. W., 1975. Deterministic and stochastic optimal control. Springer-Verlag, Berlin-New York.

Forbes, M. G., Forbes, J. F., Guay, M., 2004. 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.

Jumarie, G., 1992. Tracking control of nonlinear stochastic systems by using path cross-entropy and Fokker-Planck equation. Internat. J. Systems Sci. 23 (7), 1101–1114.

Kárný, M., 1996. Towards fully probabilistic control design. Automatica J. IFAC 32 (12), 1719–1722.

Kolmogoroff, A., 1931. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. 104 (1), 415–458.

Ladyzhenskaya, O., Solonnikov, V., Ural'tseva, N., 1967. Linear and quasilinear equations of parabolic type. Moskva: Izdat. 'Nauka'. 736 pp.

Lions, J.-L., 1969. Quelques méthodes de résolution des problèmes aux limites non linéaires. Dunod; Gauthier-Villars, Paris.

Lions, J.-L., 1971. 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.

Porretta, A., 2014. On the planning problem for the mean field games system. Dyn. Games Appl. 4 (2), 231–256.

Porretta, A., 2015. Weak solutions to Fokker-Planck equations and mean field games. Arch. Ration. Mech. Anal. 216 (1), 1–62.

Primak, S., Kontorovich, V., Lyandres, V., 2004. Stochastic methods and their applications to communications. John Wiley & Sons, Inc., Hoboken, NJ.

Protter, P. E., 2005. Stochastic integration and differential equations. Vol. 21 of Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin.

Risken, H., 1989. The Fokker-Planck equation, 2nd Edition. Vol. 18 of Springer Series in Synergetics. Springer-Verlag, Berlin.

Simon, J., 1987. Compact sets in the space $L^p(0,T;B)$. Ann. Mat. Pura Appl. (4) 146, 65–96.

Wang, H., 1999. Robust control of the output probability density functions for multivariable stochastic systems with guaranteed stability. IEEE Trans. Automat. Control 44 (11), 2103–2107.

Wloka, J., 1987. Partial differential equations. Cambridge University Press, Cambridge.