Computational Optimization and Applications
Comput Optim Appl
DOI 10.1007/s10589-017-9944-3

A Fokker–Planck approach to control collective motion

Souvik Roy[1] · **Mario Annunziato**[2] ·
Alfio Borzì[1] · **Christian Klingenberg**[3]

Received: 21 September 2015

Abstract A Fokker–Planck control strategy for collective motion is investigated. This strategy is formulated as the minimisation of an expectation objective with a bilinear optimal control problem governed by the Fokker–Planck equation modelling the evolution of the probability density function of the stochastic motion. Theoretical results on existence and regularity of optimal controls are provided. The resulting optimality system is discretized using an alternate-direction implicit Chang–Cooper scheme that guarantees conservativeness, positivity, L^1 stability, and second-order accuracy of the forward solution. A projected non-linear conjugate gradient scheme is used to solve the optimality system. Results of numerical experiments validate the theoretical accuracy estimates and demonstrate the efficiency of the proposed control framework.

Keywords Fokker–Planck equation · Alternate direction method · Chang–Cooper scheme · Projected gradient method · Control constrained PDE optimization

✉ Souvik Roy
souvik.roy@mathemtik.uni-wuerzburg.de

Mario Annunziato
mannunzi@unisa.it

Alfio Borzì
alfio.borzi@mathematik.uni-wuerzburg.de

Christian Klingenberg
klingen@mathematik.uni-wuerzburg.de

1 Institut für Mathematik, Universität Würzburg, Emil-Fischer-Strasse 30, 97074 Würzburg, Germany

2 Dipartimento di Matematica, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, Italy

3 Institut für Mathematik, Universität Würzburg, Emil-Fischer-Strasse 40, 97074 Würzburg, Germany

Mathematics Subject Classification 35Q84 · 35Q91 · 35Q93 · 49K20 · 49J20 · 65C20

1 Introduction

In recent years, there has been a surge of interest in the modelling and control of collective motion. Collective movements are observed in group of cells [35], colonies of bacteria, herds of animals [36], birds, and fishes [40]; see, e.g., [10] for a review on collective motion in biological systems. Moreover, collective motion appears in human behaviour as pedestrian track patterns and traffic flows [5,28], or for the interaction of dust particles including noise [18].

The successful application of differential models has motivated laboratory investigation suggesting that collective motion models could be augmented by including stochastic terms; see, e.g., [17,40]. For this reason, very recently differential models have been proposed that include noise modelled by a Wiener process; see, e.g., [8,28,35,40], where the underlying idea is that the individual random dispersal results from collision that can be described by a Brownian motion.

However, in deterministic models, an optimal control is obtained by determining a control function u that optimizes a given objective given by a cost functional J. For this purpose, many control strategies are available as, for example, the model predictive control (MPC) strategy [21] and linear–quadratic–regulator control methodologies. On the other hand, in stochastic models the state evolution $X(t)$ is random and represents an outcome of a probability space, therefore a direct insertion of $X(t)$ into the objective J results into a random variable. For this reason, in stochastic optimal control, the following average of the cost functional is considered [15]

$$J(X,u)=\mathbb{E}\left[\int_0^T L(t,X(t),u(t))\,dt+\Psi[X(T)]\right], \tag{1}$$

where L and Ψ are continuous functions which satisfy the polynomial growth conditions:

$$|L(t,X,u)|\leq C(1+|X|+|u|)^k$$
$$|\Psi(X)|\leq D(1+|X|)^k.$$

for some suitable constants C, D, k, and T is the final time. In this setting, the method of dynamic programming can be applied [15] in order to formulate the Hamilton–Jacobi–Bellman (HJB) equation for $\min_u J$ with u the control function.

However, an advantageous approach is to formulate the control problem in a deterministic framework, considering the problem from a statistical point of view. For this reason, we remark that the state of a stochastic process can be completely characterized in many cases by the shape of its statistical distribution, which is represented by the probability density function (PDF). Furthermore, we recognize that the evolution of the PDF associated to a stochastic process with Brownian noise is governed by a

Fokker–Planck (FP) equation. This is a partial differential equation of parabolic type with Cauchy data given by an initial PDF distribution. Therefore, a control methodology formulated in terms of the PDF and the use of the Fokker–Planck equation can provide an efficient control framework that can accommodate a wide class of objectives. This strategy has been investigated in [2] in the case of quadratic objectives of the PDF of the stochastic process. More recently, it has been recognized that the HJB control framework is related to the FP control strategy whenever expectation objectives are considered [3]; for a more detailed discussion on this issue see the end of Sect. 3 below.

The purpose of this work is to investigate the FP control strategy for collective motion. Such a control strategy was recently investigated for crowd motion in [34], where an objective containing a tracking-trajectory error and a terminal cost expectation functional is formulated in terms of a 'valley' potential V, such that the minimization of this functional aims at driving the motion in the region of low potential; the terminal cost is defined in a similar way. In [34], a control function dependent only on space and a H^1 cost of the control in the objective are considered. In this work, we generalize our control to be space–time dependent. Further, for the cost of the control, we consider two different functionals. In the first case, we take a H^1 cost of the control as in [34], which is usually chosen in open-loop control problems [6]. In the second case, we consider an expectation functional of the H^1 cost, which results in a control function of feedback type. In both cases, we consider the presence of box-constraints on the controls. Corresponding to these two functionals, we formulate optimal control problems governed by the FP equation related to stochastic motion and characterize the solutions to these problems as the solutions of the resulting optimality systems. Further, we analyze existence and regularity of the controls that appear in the FP bilinear control structure; see for e.g., [6].

To compute the optimal controls, we discuss the discretization of the optimality systems based on an alternate-direction implicit (ADI) scheme combined with a second-order accurate and positive preserving Chang–Cooper (CC) scheme, as in [34]. In this reference, the authors state the conservativeness, positivity, L^2 stability and second order accuracy for the ADI-CC scheme. In this work, we prove rigorously that the ADI-CC scheme is conservative, positive preserving, L^1 stable, and second-order accurate in space and time. We also discuss the extension of this scheme to the adjoint equations appearing in the optimality systems. The forward and adjoint FP equations and the optimality condition are solved with a projected gradient-based optimization procedure.

In the following section, we present the stochastic model for motion where the velocity field has the role of the control function and a Wiener process represents dispersal due to collision among individuals. In correspondence to this model, we discuss the FP equation where the drift given by the velocity field plays the role of a control coefficient in the convective term. Together with the FP equation, an initial PDF distribution is given that also represents the density distribution of the individuals at the start of the evolution. Also in the next section, we discuss two different objectives and the corresponding optimality systems with box constraints. In particular, in the case of an expectation functional, we show that the adjoint problem and the optimality condition may not depend on the PDF function (i.e., the forward problem) and result in a feedback control strategy. Section 3 is devoted to the analysis of our optimal control problems and of the corresponding optimality systems. We discuss the regularity of the control-to-state map, and prove existence of optimal controls and differentiability of the reduced objectives. In Sect. 4, we discuss a stable, second-order accurate discretization of the FP equation that is able to accommodate the inequality constraint given by the optimality condition. The challenge is to construct a scheme that is conservative and guarantees positivity for any value of the control function. We achieve this goal combining an ADI method with the CC scheme. Further, we prove stability and second-order space–time accuracy in the L^1 norm that appears to be the natural choice in the framework of FP problems. In Sect. 5, we obtain the discretized FP adjoint based on the discretize-before-optimize strategy. In Sect. 6, we discuss a projected non-linear conjugate gradient (NCG) scheme with H^1 gradient to solve our optimization problems which is an extension of NCG to optimization problems with box constraints on the controls. Section 7 is devoted to the validation of our control strategies and of our numerical analysis estimates. A section of conclusion completes the presentation of our work.

2 A Fokker–Planck control framework

We investigate a strategy for the control of the motion of an individual whose position at time t is denoted with $X(t) \in \mathbb{R}^n$, $1 \leq n \leq 3$, and its velocity field, depending on position x and time t, is given by $u = u(x,t) \in \mathbb{R}^n$. Further, we assume that the individual is subject to random collisions with other individuals, forcing our individual to a Brownian motion. This dynamics is modelled by the following continuous-time stochastic process [30]

$$\begin{aligned} dX(t) &= u(X(t), t)\, dt + \sigma\, dW(t), \\ X(t_0) &= X_0, \end{aligned} \tag{2}$$

where the state variable $X(t)$ is subject to deterministic infinitesimal increments driven by the vector valued drift function $u = (u_1, \ldots, u_n)$, and to random infinitesimal increments proportional to a multi-dimensional Wiener process $dW_t \in \mathbb{R}^n$, with stochastically independent components and $\sigma \in \mathbb{R}$ is a positive constant.

We assume that the process (2) is constrained to stay in a bounded convex domain with Lipschitz boundaries, thus $X(t) \in \Omega \subset \mathbb{R}^n$, by virtue of a reflecting barrier on $\partial\Omega$.

Now, we introduce the Fokker–Planck (FP) equation that governs the evolution of the probability density function (PDF) of the process modelled by (2). We have

$$\begin{aligned} &\partial_t f(x,t) - \frac{\sigma^2}{2} \sum_{i=1}^{n} \partial^2_{x_i x_i} f(x,t) + \sum_{i=1}^{n} \partial_{x_i} \left(u_i(x,t) f(x,t) \right) = 0 \\ &f(x,0) = f_0(x) \end{aligned} \tag{3}$$

where $f = f(x, t)$ is the PDF of the individual to be in x at time t. The function $f_0(x)$ represents the initial PDF distribution that satisfies the following

$$f_0 \geq 0, \quad \int_\Omega f_0(x)dx = 1. \tag{4}$$

The function $f_0(x)$ represents the distribution of the initial position X_0 of the process and the domain of definition of the FP problem is $Q = \Omega \times (0, T)$.

The reflecting barrier conditions assumed on the process correspond to flux zero boundary conditions for the FP equation. For this purpose, notice that (3) can be written in flux form as follows

$$\partial_t f(x, t) = \nabla \cdot F, \quad f(x, 0) = f_0(x), \tag{5}$$

where the flux F is given component-wise by

$$F_j(x, t; f) = \frac{\sigma^2}{2}\partial_{x_j} f - u_j(x, t) f. \tag{6}$$

Flux zero boundary conditions are formulated as follows

$$F \cdot \hat{n} = 0 \quad \text{on} \quad \partial\Omega \times (0, T), \tag{7}$$

where $\hat{n}$ is the unit outward normal on $\partial\Omega$.

We anticipate that the drift u represents our control function that is sought in the following admissible set

$$\begin{aligned} U_{ad} = \Big\{ u \in \left(L^2(0, T; H_0^1(\Omega)) \right)^n \mid u_a \leq u_i(x) \leq u_b, \quad i = 1, \ldots, n \\ \text{a.e. in} \quad \Omega,\ u_a, u_b \in \mathbb{R}, u_a \leq 0 \leq u_b \Big\}. \end{aligned} \tag{8}$$

We denote $U = (L^2(0, T; H_0^1(\Omega)))^n$.

Our purpose is to discuss a robust control strategy with the objective modelled by the following functional

$$\begin{aligned} J(f, u) = \alpha \int_Q V(x - x_t) f(x, t)dxdt + \beta \int_\Omega V(x - x_T) f(x, T)dx \\ + \frac{\nu}{2} \int_0^T \int_\Omega A(u(x, t))dxdt, \quad \alpha, \beta, \nu > 0, \end{aligned} \tag{9}$$

where $x_t = (x^1(t), \ldots, x^n(t))$ represents a desired trajectory, $t \in [0, T], x_T = \bar{x}(T), \alpha, \beta, \nu > 0$. The function V represents a given convex and smooth potential; see Fig. 1.

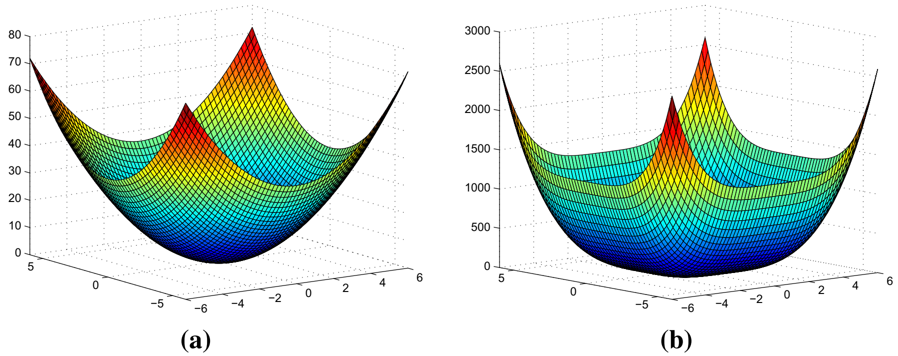


Fig. 1 Two types of potential functions V at time $t = 0$. **a** Square potential. **b** Quartic potential

We consider two choices of quantifying the cost of the control. Thus we consider $A(u)$ as follows

$$A(u(x,t)) = |u(x,t)|^2 + |\nabla u(x,t)|^2, \tag{C1}$$

$$A(u(x,t)) = (|u(x,t)|^2 + |\nabla u(x,t)|^2) f(x,t), \tag{C2}$$

where $|\cdot|$ represents the standard Euclidean norm in $\mathbb{R}^n$, ∇u is the Jacobian matrix whose entries are defined by $(\nabla u)_{ij} = \dfrac{\partial u_i}{\partial x_j}$ and $|\nabla u|$ is the Frobenius norm of ∇u. With this latter choice of $A(u)$, the cost functional J is linear in f.

Minimising J corresponds to the aim to drive the random process to follow the path of minimum potential at all times and to reach a region of low potential at the terminal time. In fact, we use $\bar{x}$ to define the minimum of V. This minimum can be interpreted as the risk-free zone for the individual.

Now we formulate the optimal control problem to find u that minimizes the objective J, given by (9), subject to the FP differential constraints (3), (4), (7), as follows

$$\min_{u \in U_{ad}} J(f,u)$$
$$\text{subject to } (3,\ 4,\ 7). \tag{10}$$

As we prove in the next section, for a given control function $u \in U_{ad}$, the solution of the Fokker–Planck model (3) and (7) is uniquely determined. We denote this dependence by $f = \Lambda(u)$ and one can prove that this mapping is differentiable. We introduce the reduced cost functional $\hat{J}$ given by

$$\hat{J}(u) = J(\Lambda(u), u). \tag{11}$$

Correspondingly, a local minimum u^* of $\hat{J}$ can be characterized by $\langle \nabla \hat{J}(u^*), v - u^* \rangle \geq 0$ for all $v \in U_{ad}$, where $\langle \cdot, \cdot \rangle$ represents the $L^2(Q)$ inner product defined as follows

$$\langle u, v \rangle = \int_0^T \int_\Omega u(x,t) v(x,t)\, dxdt = \int_0^T \int_\Omega (u(\cdot,t) v(\cdot,t))_{L^2(\Omega)}\, dt,$$

which induces the norm $\|u\|_{L^2(0,T;L^2(\Omega))}$. This local minimum can be characterized by using the following Lagrange functional

$$L(f,u,p) = J(f,u) + \langle \partial_t f - \nabla \cdot F, p\rangle, \tag{12}$$

whose stationary points establish the first-order necessary conditions to the solution of the optimal control problem (10).

For the case (C1), the first-order necessary conditions are given by

$$\begin{aligned} \partial_t f(x,t) - \frac{\sigma^2}{2}\sum_{i=1}^n \partial^2_{x_i x_i} f(x,t) + \sum_{i=1}^n \partial_{x_i}\,(u_i(x,t)f(x,t)) = 0 \\ f(x,0) = f_0(x) \quad \text{in} \quad \Omega \\ F\cdot\hat{n} = 0 \quad \text{on} \quad \partial\Omega\times(0,T), \end{aligned} \tag{13}$$

$$\begin{aligned} -\partial_t p(x,t) - \frac{\sigma^2}{2}\sum_{i=1}^n \partial^2_{x_i x_i} p(x,t) - \sum_{i=1}^n u_i(x,t)\partial_{x_i} p(x,t) + \alpha V(x-x_t) = 0 \\ p(x,T) = -\beta V(x-x_T) \quad \text{in} \quad \Omega \\ \frac{\partial p}{\partial \hat{n}} = 0 \quad \text{on} \quad \partial\Omega\times(0,T), \end{aligned} \tag{14}$$

$$\left\langle \nu u_k - \nu\Delta u_k - \frac{\partial p}{\partial x_k} f, v - u_k \right\rangle \geq 0 \quad \forall v \in U_{ad}, \quad k = 1,\ldots,n. \tag{15}$$

For the case (C2), the first-order necessary conditions are given by (13) along with

$$\begin{aligned} -\partial_t p(x,t) - \frac{\sigma^2}{2}\sum_{i=1}^n \partial^2_{x_i x_i} p(x,t) - \sum_{i=1}^n u_i(x,t)\partial_{x_i} p(x,t) \\ + \alpha V(x-x_t) + \frac{\nu}{2}(|u(x,t)|^2 + |\nabla u(x,t)|^2 = 0 \\ p(x,T) = -\beta V(x-x_T) \quad \text{in}\ \Omega \\ \frac{\partial p}{\partial \hat{n}} = 0 \quad \text{on} \quad \partial\Omega\times(0,T), \end{aligned} \tag{16}$$

$$\left\langle \nu u_k f - \nu\Delta u_k f - \frac{\partial p}{\partial x_k} f, v - u_k \right\rangle \geq 0 \quad \forall v \in U_{ad}, \quad k = 1,\ldots,n. \tag{17}$$

Notice that in the optimality system (15), the following reduced L^2 gradient components appear

$$\nabla_{u_k} \hat{J}(u)(\cdot,t) = \left[\nu u_k - \nu\Delta u_k - \frac{\partial p}{\partial x_k} f\right](\cdot,t), \quad k = 1,\ldots,n, \tag{18}$$

and in the optimality system (17), the following reduced L^2 gradient components appear

$$\nabla_{u_k} \hat{J}(u)(\cdot,t) = \left[\left(\nu u_k - \nu \Delta u_k - \frac{\partial p}{\partial x_k}\right) f\right](\cdot,t), \quad k = 1,\ldots,n, \tag{19}$$

for almost all $t \in [0,T]$, where Δ is the distributional Laplacian. Now, let us discuss the control-unconstrained case. In this case optimality requires $\nabla_{u_k} \hat{J}(u) = 0$. Because of the H^1 control costs, we have a setting that allows to include boundary conditions on the control function. By considering the derivation of the optimality system above using the Lagrange formulation [38], we find that a convenient choice is to require $u = 0$ on $\partial\Omega$. However, also homogeneous Neumann boundary conditions are appropriate. We chose homogeneous Dirichlet boundary conditions because in this case the control does not appear in the FP flux zero boundary conditions given by (7). Alternatively, we could choose Neumann boundary conditions and the control function on the boundary would result by application of the trace operator.

Notice that, assuming the last term in (18) being in $H^{-1}(\Omega)$ and because Ω is a Lipschitz convex domain, the solution of the gradient equation with homogeneous Dirichlet boundary conditions results in $u(\cdot,t) \in H_0^1(\Omega)$. However, we wish to apply a gradient-based optimization scheme where the residual of (18) is used. For this purpose, we cannot use this residual directly for updating the control, since it is in $H^{-1}(\Omega)$. Therefore, it is necessary to determine the reduced H^1 gradient. This is done based on the following fact

$$\left(\nabla \hat{J}(u)_{H^1}(\cdot,t), \varphi(\cdot)\right)_{H^1(\Omega)} = \left(\nabla \hat{J}(u)(\cdot,t), \varphi(\cdot)\right)_{L^2(\Omega)},$$

a.e., in $(0,T)$ and $\varphi \in (H_0^1(\Omega))^n$. Using the definition of the H^1 inner product and integrating by parts, we have that the H^1 gradient is obtained by solving the following boundary value problem for $t \in (0,T)$

$$-\Delta\left(\nabla_{u_k} \hat{J}(u)_{H^1}\right) + \left(\nabla_{u_k} \hat{J}(u)_{H^1}\right) = \nabla_{u_k} \hat{J}(u) \quad \text{in} \quad \Omega \tag{20}$$

$$\left(\nabla_{u_k} \hat{J}(u)_{H^1}\right) = 0 \quad \text{on} \quad \partial\Omega, \tag{21}$$

where $k = 1,\ldots,n$, $\nabla_{u_k} \hat{J}(u)_{H^1}$ denotes the kth component of $\nabla_{u_k} \hat{J}(u)_{H^1}$ and (20)–(21) is defined in the weak sense. The solution to this problem provides the appropriate gradient to be used in a gradient update of the control that includes projection to satisfy the given control constraints.

3 Theory of the Fokker–Planck optimal control problem

In this section, we discuss the existence of solutions to the optimal control problem (10) and its characterization by the optimality system (13)–(15). While we follow a 'classical' reasoning path to analyze our problems, we refer to, e.g., [32] for an interesting alternative theoretical framework.

To simplify our analysis while addressing the essential issues, we consider $\beta = 0$ in the cost functional.

Consider the following FP control problem

$$\min J(f, u), \quad \text{s.t.} \quad \mathcal{E}(f_0, u) = 0, \tag{22}$$

where the equation $\mathcal{E}(f_0, u) = 0$ denotes (13).

Our analysis of this problem starts with a discussion concerning existence of weak solutions to $\mathcal{E}(f_0, u) = 0$. In the case of (13) in a bounded domain and with reflecting boundary conditions, we refer to classical results in [9,33] for time-dependent convection–diffusion equations with Robin boundary conditions; see also the recent works [13,16]. Furthermore, to link our discussion to a more general framework, we refer to [19,20,26]. We have the following proposition.

Proposition 1 *Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$, and $u \in U_{ad} \subset U$. Then $\mathcal{E}(f_0, u) = 0$ admits a unique non-negative solution $f \in L^2(0, T; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$.*

We remark that using classical bootstrapping techniques [37], one can get the $H^2(\Omega)$ regularity in space. Furthermore, because of (5) and (7), we can prove the following theorem that states conservation of the total probability.

Proposition 2 *The FP problem (13) with (4) is conservative.*

To prove this proposition and a stability property of our FP model, we consider the L^2 scalar product with a test function $\psi \in H^1(\Omega)$. Integrating by part the diffusion operator and including the flux zero boundary conditions, we obtain the following

$$\int_\Omega \frac{\partial f}{\partial t} \psi \, dx = -\frac{\sigma^2}{2} \int_\Omega \nabla f \cdot \nabla \psi \, dx + \int_\Omega (u f) \cdot \nabla \psi \, dx. \tag{23}$$

Notice that choosing $\psi = 1$, we obtain $\int_\Omega f(x, t) dx = 1$ for all $t \in [0, T]$; this proves Proposition 2. On the other hand, by choosing $\psi = f(\cdot, t)$, we have

$$\frac{\partial}{\partial t} \|f(t)\|^2_{L^2(\Omega)} = -\sigma^2 \|\nabla f(t)\|^2_{L^2(\Omega)} + 2 \int_\Omega (u f(t)) \cdot \nabla f(t) \, dx. \tag{24}$$

Now, denote $\bar{u} = \max\{|u_a|, |u_b|\}$ and use the Cauchy inequality, $2bd \leq b^2/k + kd^2$, to estimate the last term in (24). We choose $k = \sigma^2/\bar{u}$ and obtain the following

$$\frac{\partial}{\partial t} \|f(t)\|^2_{L^2(\Omega)} \leq \frac{\bar{u}^2}{\sigma^2} \|f(t)\|^2_{L^2(\Omega)}.$$

Therefore we have

$$\|f(t)\|_{L^2(\Omega)} \leq \|f_0\|_{L^2(\Omega)} \exp\left(\frac{\bar{u}^2}{2\sigma^2} t \right). \tag{25}$$

This result provides a useful bound of the L^2 norm of the PDF.

Next, we state some further properties of the solution to (13), which we need in order to analyse our FP optimal control problem. We have the following proposition.

Proposition 3 *Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$, and $u \in U_{ad} \subset U$. Then if f is a solution to $\mathcal{E}(f_0, u) = 0$, the following inequalities hold*

$$\|f\|_{L^\infty(0,T;L^2(\Omega))} \leq c_1 \|f_0\|_{L^2(\Omega)}, \tag{26}$$

$$\|\partial_t f\|_{L^2(0,T;H^{-1}(\Omega))} \leq \left(c_2 + c_3\|u\|_{L^2(\Omega;\mathbb{R}^n)}\right) \|f_0\|_{L^2(\Omega)}, \tag{27}$$

where c_1, c_2, c_3 are positive constants. Further, if $\sigma^2 > \bar{u}$ then the following inequality holds

$$\|f\|_{L^2(0,T;H^1(\Omega))} \leq c_4 \|f_0\|_{L^2(\Omega)}, \tag{28}$$

where c_4 is a positive constant.

Proof The inequality (26) follows from (25), with

$$c_1 = \exp\left(\frac{\bar{u}^2 T}{2\sigma^2}\right).$$

For proving inequality (27), we note that

$$\|\partial_t f\|_{H^{-1}(\Omega)} = \sup_{\substack{\psi \in H_0^1(\Omega) \\ \psi \neq 0}} \frac{\langle \partial_t f, \psi\rangle_{L^2(\Omega)}}{\|\psi\|_{H_0^1(\Omega)}}.$$

From (23), using (25) we get

$$\langle \partial_t f, \psi\rangle_{L^2(\Omega)} \leq (c_2 + c_3\|u\|_{L^2(\Omega;\mathbb{R}^n)})) \|f_0\|_{L^2(\Omega)} \|\psi\|_{H_0^1(\Omega)},$$

where

$$c_2 = \frac{c_2^2 \sigma^2}{2}, \quad c_3 = c_1^2.$$

To prove (28), we first integrate (24) in $(0, T)$ to obtain

$$\|f(T)\|^2_{L^2(\Omega)} - \|f_0\|^2_{L^2(\Omega)} = -\sigma^2 \int_0^T \|\nabla f(t)\|^2_{L^2(\Omega)}\, dt \\ + 2\int_0^T \int_\Omega (uf(t)) \cdot \nabla f(t)\, dx dt.$$

Using the Cauchy inequality, we have

$$\sigma^2 \int_0^T \|\nabla f(t)\|^2_{L^2(\Omega)}\, dt \le \|f_0\|^2_{L^2(\Omega)} + \bar{u} \int_0^T \left(\|f(t)\|^2_{L^2(\Omega)} + \|\nabla f(t)\|^2_{L^2(\Omega)} \right) dt.$$

This implies

$$(\sigma^2 - \bar{u}) \int_0^T \|\nabla f(t)\|^2_{L^2(\Omega)}\, dt \le \|f_0\|^2_{L^2(\Omega)} + \bar{u} \int_0^T \|f(t)\|^2_{L^2(\Omega)}\, dt. \tag{29}$$

Adding $(\sigma^2 - \bar{u}) \int_0^T \|f(t)\|^2_{L^2(\Omega)}\, dt$ to (29) we have the following

$$\begin{aligned} &(\sigma^2 - \bar{u}) \int_0^T \left(\|f(t)\|^2_{L^2(\Omega)} + \|\nabla f(t)\|^2_{L^2(\Omega)} \right) dt \\ &\quad \le \|f_0\|^2_{L^2(\Omega)} + \sigma^2 \int_0^T \|f(t)\|^2_{L^2(\Omega)}\, dt. \end{aligned} \tag{30}$$

From (25), we have

$$\begin{aligned} \int_0^T \|f(t)\|^2_{L^2(\Omega)}\, dt &\le \|f_0\|^2_{L^2(\Omega)} \int_0^T \exp\left(\frac{\bar{u}^2}{2\sigma^2} t\right) dt \\ &= \frac{2\sigma^2}{\bar{u}^2} \left[\exp\left(\frac{\bar{u}^2}{2\sigma^2} T\right) - 1 \right] \|f_0\|^2_{L^2(\Omega)}. \end{aligned} \tag{31}$$

Therefore we obtain

$$\begin{aligned} &(\sigma^2 - \bar{u}) \int_0^T \left(\|f(t)\|^2_{L^2(\Omega)} + \|\nabla f(t)\|^2_{L^2(\Omega)} \right) dt \\ &\quad \le \frac{2\sigma^4}{\bar{u}^2} \left[\exp\left(\frac{\bar{u}^2}{2\sigma^2} T\right) - 1 + \frac{\bar{u}^2}{2\sigma^4} \right] \|f_0\|^2_{L^2(\Omega)}. \end{aligned} \tag{32}$$

This proves (28) with $c_4 = \sqrt{\frac{2\sigma^4}{\bar{u}^2(\sigma^2 - \bar{u})} \left[\exp\left(\frac{\bar{u}^2}{2\sigma^2} T\right) - 1 + \frac{\bar{u}^2}{2\sigma^4} \right]}$. □

Using the results above, we obtain that the mapping $\Lambda : U \to C([0,T]; H^1(\Omega))$, $u \to f = \Lambda(u)$ is continuous. Following the arguments given in [2,38], we can prove that this mapping is Fréchet differentiable. This is stated in the following theorem.

Proposition 4 *Let* $\mathcal{A} = -\frac{1}{2} \sum_{i=1}^n \partial^2_{x_i x_i} (a_i(x)\, \cdot)$ *and* $\mathcal{B} = -\sum_{i=1}^n \partial_{x_i}(\cdot)$ *The mapping* $\Lambda : U \to C([0,T]; H^1(\Omega))$, $u \to f = \Lambda(u)$ *is the solution to* $\mathcal{E}(f_0, u) = 0$*, is Fréchet differentiable, and* $e = \Lambda'_{u^*} \cdot h$ *satisfies the equation*

$$\dot{e} + \mathcal{A}e = \mathcal{B}(u^* e) + \mathcal{B}(h f^*), \quad e(0) = 0,$$

where $f^* = \Lambda(u^*)$ *and* $h \in U$.

In the following proposition, we discuss the properties of the cost functional.

Proposition 5 *The objective functional* (9) *is sequentially weakly lower semicontinuous (w.l.s.c.), bounded from below, coercive on* U, *and it is Fréchet differentiable.*

The proof of this proposition is straightforward, once one recalls that the PDF is a nonnegative function.

The next result states existence of an optimal control u^*.

Proposition 6 *Assume that* $f_0 \in H^1(\Omega)$ *and it satisfies* (4)*, and the objective is given by* (9)*. Then there exists a pair* $(f^*, u^*) \in C([0, T]; H^1(\Omega)) \times U_{ad}$ *such that* f^* *is a solution to* $\mathcal{E}(f_0, u^*) = 0$ *and* u^* *minimizes* J *in* U_{ad}.

Proof The proof follows standard arguments [38]. In fact, boundedness from below of J guarantees the existence of a minimising sequence (u^m). Since U is reflexive and J is sequentially w.l.s.c. and coercive in U, this sequence is bounded. Therefore it contains a weakly convergent subsequence (u^{m_l}) in $H_0^1(\Omega; \mathbb{R}^n)$, $u^{m_l} \rightharpoonup u^*$. Correspondingly, the sequence (f^{m_l}), where $f^{m_l} = \Lambda(u^{m_l})$, is bounded in $L^2(0, T; H^1(\Omega))$, while the sequence of the time derivatives, $(\partial_t f^{m_l})$, is bounded in $L^2(0, T; H^{-1}(\Omega))$. Therefore both sequences converge weakly to f^* and $\partial_t f^*$, respectively. Now, we invoke the Theorem of Aubin–Lions [25] to state strong convergence of a subsequence (f^{m_k}) in $L^2(0, T, L^2(\Omega))$. At this point, it remains to address the bilinear state-control term in the FP equation. For this purpose, notice that we need to consider the sequence $\nabla \cdot (u^{m_k} f^{m_k})$ within the weak formulation of solutions to the FP problem. Therefore we focus on $\langle \nabla \cdot (u^{m_k} f^{m_k}), \psi \rangle_{L^2(\Omega)}$ for any $\psi \in H^1(\Omega)$ (we omit the time dependence of f). Since $u^{m_k} \in H_0^1(\Omega; \mathbb{R}^n)$, we have $\langle \nabla \cdot (u^{m_k} f^{m_k}), \psi \rangle_{L^2(\Omega)} = -\langle (u^{m_k} f^{m_k}), \nabla \psi \rangle_{L^2(\Omega;\mathbb{R}^n)}$. Now, from the previous discussion, we can state weak convergence of the sequence of products $(u^{m_k} f^{m_k})$ in $L^2(0, T, L^2(\Omega; \mathbb{R}^n))$, that is, $\langle (u^{m_k} f^{m_k}), \nabla \psi \rangle_{L^2(\Omega;\mathbb{R}^n)} \to \langle (u^* f^*), \nabla \psi \rangle_{L^2(\Omega;\mathbb{R}^n)}$. With this preparation, and again using the standard argument of considering the limiting sequences in the weak formulation of solutions to the FP problem (13), it follows that $f^* = \Lambda(u^*)$, and the pair (f^*, u^*) minimizes the objective. □

Finally, the following proposition shows the differentiability of the reduced functional $\hat{J}$ defined in (11), which can be proved using similar arguments as in [38]. Notice that here Δ denotes the vector Laplacian.

Proposition 7 *The reduced functional* $\hat{J}(u)$ *is differentiable and its derivative is given by*

$$d\hat{J}(u) \cdot v = \left\langle \nu u - \nu \Delta u - f \partial_x p, v \right\rangle \quad \forall v \in U,$$

for the case (C1) *and*

$$d\hat{J}(u) \cdot v = \left\langle \nu u f - \nu \Delta u f - f \partial_x p, v \right\rangle \quad \forall v \in U,$$

for the case (C2)*, where* p *is the solution to the adjoint equation*

$$-\partial_t p(x,t) - \frac{\sigma^2}{2}\sum_{i=1}^{n}\partial^2_{x_i x_i} p - \sum_{i=1}^{n} u_i \partial_{x_i} p = -\alpha V(x - x_t),$$
$$\frac{\partial p}{\partial n} = 0 \quad on \quad \partial\Omega \times (0,T),$$

with $p(x,T) = -\beta V(x - x_T)$ *for the case* (C1)*, and*

$$-\partial_t p(x,t) - \frac{\sigma^2}{2}\sum_{i=1}^{n}\partial^2_{x_i x_i} p - \sum_{i=1}^{n} u_i \partial_{x_i} p$$
$$= -\alpha V(x - x_t) - \frac{\nu}{2}(|u(x,t)|^2 + |\nabla u(x,t)|^2),$$

with $p(x,T) = -\beta V(x - x_T)$ *for the case* (C2) *and* f *is the solution to* $\mathcal{E}(f_0, u) = 0$.

Similarly to the discussion above concerning the forward FP problem, we refer to [9,33] for results on existence and regularity of solutions to the adjoint FP problem.

At this point, we are able to elucidate the connection of the present FP open-loop approach with the feedback control strategy that results in the HJB framework. This connection is established in the case (C2), where the objective $J(f,u)$ given in (9) is linear in f and becomes an expectation cost functional. In fact, in this case the optimality condition takes the following form

$$\left\langle f \left(\nu u_k - \nu \Delta u_k - \frac{\partial p}{\partial x_k}\right), \quad v - u_k \right\rangle \geq 0 \quad \forall v \in U_{ad}, \quad k = 1,\ldots,n.$$

Now, it appears that, in the unconstrained-control case, the condition $\nu u_k - \nu\Delta u_k - \partial_{x_k} p = 0$ in Ω and $u = 0$ on $\partial\Omega$, and a.e. in $(0,T)$, is a sufficient condition for optimality. In this case, the control u is determined by this optimality condition and the adjoint equation and thus this control can be regarded as closed-loop control for our stochastic model.

We should further notice that the PDF of the stochastic process is everywhere non-negative and it cannot be zero on an open set (for any $t > 0$). Therefore the boundary value problem mentioned above constitutes also a necessary condition for optimality. Moreover, the fact that the PDF is a.e. positive allows to extend the above consideration to the constrained-control case, which then requires that the control u must satisfy an elliptic variational inequality.

4 Discretization of the optimality system

In this section, we discuss the spatial discretization to the FP and adjoint FP equations using the Chang–Cooper scheme that is a second-order accurate numerical scheme for the FP equation. We restrict ourselves to the two-dimensional case and assume that the control function is Lipschitz continuous in space with Lipschitz constant Γ independent of t, i.e.,

$$\|u(x_1, y_1, t) - u(x_2, y_2, t)\| \leq \Gamma \|(x_1, y_1) - (x_2, y_2)\|,$$
$$\forall (x_1, y_1), (x_2, y_2) \in \Omega \subset \mathbb{R}^2, t \in [0, T], \tag{33}$$

where $\|\cdot\|$ represents the Euclidean norm in $\mathbb{R}^2$ and $\Omega \equiv (-a, a) \times (-a, a)$ is a square domain. Then consider a sequence of uniform grids $\{\Omega_h\}_{h>0}$ given by

$$\Omega_h = \{(x, y) \in \mathbb{R}^2 : (x_i, y_j) = (-a + ih, -a + jh),\ (i, j) \in \{0, \dots, N_x\}^2\} \cap \Omega,$$

where N_x represents the number of cells in each direction and h is the spatial stepsize. Moreover, h is chosen such that the boundaries of Ω coincide with the grid points. Let δt be the time stepsize and N_t denotes the number of time steps. Define

$$Q_{h,\delta t} = \{(x_i, y_j, t_m) : (x_i, y_j) \in \Omega_h,\ t_m = m\delta t,\ 0 \leq m \leq N_t\}.$$

On the grid $Q_{h,\delta t}$, $f^m_{i,j}$ represents the value of the grid function in Ω_h at (x_i, y_j) and time t_m.

For space discretization, we need a second-order scheme which guarantees positivity of the PDF together with the conservation of the total probability. These are the essential features of the Chang–Cooper (CC) scheme. The first step in the formulation of the CC scheme is to consider the flux form of the FP equation (13). The divergence of the flux term, $\nabla \cdot F$, at time t_m can be discretized as follows

$$\nabla \cdot F = \frac{1}{h} \left\{ \left(F^m_{i+\frac{1}{2},j} - F^m_{i-\frac{1}{2},j} \right) + \left(F^m_{i,j+\frac{1}{2}} - F^m_{i,j-\frac{1}{2}} \right) \right\}$$

where $F^m_{i+\frac{1}{2},j}$ and $F^m_{i,j+\frac{1}{2}}$ represents the flux in the ith and jth direction respectively at the point (x_i, x_j) and is given as

$$F^m_{i+\frac{1}{2},j} = \left[-(1 - \delta^m_{i+\frac{1}{2},j}) u^1_{i+\frac{1}{2},j,m} + \frac{\sigma^2}{2h} \right] f^m_{i+1,j} - \left[\frac{\sigma^2}{2h} + \delta^m_{i+\frac{1}{2},j} u^1_{i+\frac{1}{2},j,m} \right] f^m_{i,j}, \tag{34}$$

and

$$F^m_{i,j+\frac{1}{2}} = \left[-(1 - \delta^m_{i,j+\frac{1}{2}}) u^2_{i,j+\frac{1}{2},m} + \frac{\sigma^2}{2h} \right] f^m_{i,j+1} - \left[\frac{\sigma^2}{2h} + \delta^m_{i,j+\frac{1}{2}} u^2_{i,j+\frac{1}{2},m} \right] f^m_{i,j}, \tag{35}$$

where

$$u^1_{i+\frac{1}{2},j,m} = -u_1 \left(x_{i+\frac{1}{2}}, y_j, t_m \right),$$

$$u^2_{i,j+\frac{1}{2},m} = -u_2\left(x_i, y_{j+\frac{1}{2}}, t_m\right), \tag{36}$$

and

$$\begin{aligned} \delta^m_{i+\frac{1}{2},j} &= \frac{1}{w^m_{i+\frac{1}{2},j}} - \frac{1}{\exp\left(w^m_{i+\frac{1}{2},j}\right)-1}, \quad w^m_{i+\frac{1}{2},j} = -2hu^1_{i+\frac{1}{2},j,m}/\sigma^2, \\ \delta^m_{i,j+\frac{1}{2}} &= \frac{1}{w^m_{i,j+\frac{1}{2}}} - \frac{1}{\exp\left(w^m_{i,j+\frac{1}{2}}\right)-1}, \quad w^m_{i,j+\frac{1}{2}} = -2hu^2_{i,j+\frac{1}{2},m}/\sigma^2. \end{aligned} \tag{37}$$

This scheme is discussed in [2,7,27].

4.1 An ADI-CC scheme for solving the FP equation

For time discretizations, we use the alternate-direction implicit (ADI) method [11,12, 31]. We couple the ADI scheme with the CC scheme for space discretization to solve the FP equation (13) and refer to the ADI-CC scheme. To formulate our ADI-CC scheme, we introduce an intermediate half time step $t_{m+\frac{1}{2}}$ between t_m and t_{m+1}. Thus for the FP equation (13) in 2D, the scheme can be written as follows

$$\begin{aligned} \frac{f^{m+\frac{1}{2}}_{i,j} - f^m_{i,j}}{\delta t/2} &= \frac{1}{h}\left(F^{m+\frac{1}{2}}_{i+\frac{1}{2},j} - F^{m+\frac{1}{2}}_{i-\frac{1}{2},j}\right) + \frac{1}{h}\left(F^m_{i,j+\frac{1}{2}} - F^m_{i,j-\frac{1}{2}}\right), \\ \frac{f^{m+1}_{i,j} - f^{m+\frac{1}{2}}_{i,j}}{\delta t/2} &= \frac{1}{h}\left(F^{m+\frac{1}{2}}_{i+\frac{1}{2},j} - F^{m+\frac{1}{2}}_{i-\frac{1}{2},j}\right) + \frac{1}{h}\left(F^{m+1}_{i,j+\frac{1}{2}} - F^{m+1}_{i,j-\frac{1}{2}}\right), \end{aligned} \tag{38}$$

for all $(i, j) \in \{1, \ldots, N_x - 1\}$. The flux zero boundary condition at the discrete level is given by

$$\begin{aligned} &F(i, N_x - 1/2, t_m) = 0, \;\; F(i, 1/2, t_m) = 0 \quad \forall i = 0, \ldots N_x, \\ &F(N_x - 1/2, j, t_m) = 0, \;\; F(1/2, j, t_m) = 0 \quad \forall j = 0, \ldots N_x, \end{aligned} \tag{39}$$

for all $m = 0, 1/2, 1, 3/2, \ldots, N_t$. Both equation in (38) are implicit, but only with respect to one of the two spatial dimension of the flux. The algorithm for implementing the ADI-CC scheme (38)–(39) is given below.

Algorithm 4.1 (Implementation of the ADI-CC scheme)

1. *Start with* $m = 0$.
2. *For* $j = 1, \ldots, N_x - 1$, *compute* $f^{m+\frac{1}{2}}_{i,j}$ *by solving the first equation in* (38) *for all* $i = 2, \ldots, N_x - 1$, *using the values of* $f^m_{i,j}$
3. *For* $i = 1, \ldots, N_x - 1$, *compute* $f^{m+1}_{i,j}$ *by solving the second equation in* (38) *for all* $j = 2, \ldots, N_x - 1$, *using the values of* $f^{m+\frac{1}{2}}_{i,j}$
4. *If* $m = N_t$, *stop else goto Step 2.*

4.1.1 Analysis of the ADI-CC scheme

We now study the properties of the ADI-CC scheme (38)–(39). The following lemma states the conservativeness of the ADI-CC scheme.

Lemma 4.1 *The ADI-CC scheme (38)–(39) is conservative.*

Proof To see this, we add both the equations in (38) to get

$$\frac{f_{i,j}^{m+1}-f_{i,j}^{m}}{\delta t/2}=\frac{2}{h}\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}}-F_{i-\frac{1}{2},j}^{m+\frac{1}{2}}\right)+\frac{1}{h}\left(F_{i,j+\frac{1}{2}}^{m}-F_{i,j-\frac{1}{2}}^{m}\right)+\frac{1}{h}\left(F_{i,j+\frac{1}{2}}^{m+1}-F_{i,j-\frac{1}{2}}^{m+1}\right). \tag{40}$$

Summing over all i, j, we have

$$\sum_{i,j}\frac{f_{i,j}^{m+1}-f_{i,j}^{m}}{\delta t/2}=\sum_{i,j}\left[\frac{2}{h}\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}}-F_{i-\frac{1}{2},j}^{m+\frac{1}{2}}\right)+\frac{1}{h}\left(F_{i,j+\frac{1}{2}}^{m}-F_{i,j-\frac{1}{2}}^{m}\right)+\frac{1}{h}\left(F_{i,j+\frac{1}{2}}^{m+1}-F_{i,j-\frac{1}{2}}^{m+1}\right)\right]. \tag{41}$$

The right hand side of (41) is a telescoping series. After summation we have

$$\begin{aligned}\sum_{i,j}\frac{f_{i,j}^{m+1}-f_{i,j}^{m}}{\delta t/2}&=\sum_{j}\frac{2}{h}\left(F_{N_x-\frac{1}{2},j}^{m+\frac{1}{2}}-F_{1/2,j}^{m+\frac{1}{2}}\right)+\sum_{i}\frac{1}{h}\left(F_{i,N_x-\frac{1}{2}}^{m}-F_{i,1/2}^{m}\right)\\&\quad+\sum_{i}\frac{1}{h}\left(F_{i,N_x-\frac{1}{2}}^{m+1}-F_{i,1/2}^{m+1}\right)\\&=0\ \text{(using (39))}.\end{aligned} \tag{42}$$

This gives

$$\sum_{i,j}f_{i,j}^{m+1}=\sum_{i,j}f_{i,j}^{m},\quad \forall m=0,\ldots,N_t-1, \tag{43}$$

which proves conservativeness of the ADI-CC scheme. □

To study the properties of positivity and error estimates for (38), we write the numerical method in matrix–vector form for the unknown $\widetilde{f}_j^m=(f_{1,j}^m,\ldots,f_{N_x-1,j}^m)$ to make the analysis easier. We do this for the first equation in (38). A similar analysis follows for the second equation in (38).

We define the following

$$\alpha_{i,j}^{m+\frac{1}{2}} = \frac{\sigma^2}{2h} + \delta_{i+\frac{1}{2},j}^{m+\frac{1}{2}} u^1_{i+\frac{1}{2},j,m+\frac{1}{2}} = -\frac{u^1_{i+\frac{1}{2},j,m+\frac{1}{2}}}{\left(\overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - 1\right)}, \quad 1 \leq i,\, j \leq N_x - 1,$$

$$\beta_{i,j}^{m} = \frac{\sigma^2}{2h} + \delta_{i,j+\frac{1}{2}}^{m} u^2_{i,j+\frac{1}{2},m} = -\frac{u^2_{i,j+\frac{1}{2},m}}{\left(\overline{w}_{i,j+\frac{1}{2}}^{m} - 1\right)}, \quad 1 \leq i,\, j \leq N_x - 1, \tag{44}$$

$$\alpha_{0,j}^{m+\frac{1}{2}} = 0, \quad 1 \leq j \leq N_x - 1,$$

$$\beta_{i,0}^{m} = 0, \quad 1 \leq i \leq N_x - 1,$$

where $\delta_{i+\frac{1}{2},j}^{m+\frac{1}{2}}, \delta_{i,j+\frac{1}{2}}^{m}$ are defined in (37) and $\overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}} = \exp(w_{i+\frac{1}{2},j}^{m+\frac{1}{2}})$, $\overline{w}_{i,j+\frac{1}{2}}^{m} = \exp(w_{i,j+\frac{1}{2}}^{m})$. We remark that $\alpha_{i,j}^{m+\frac{1}{2}}$, $\beta_{i,j}^{m}$ are positive.

Using (44), the first equation in (38) reads as follows

$$\mathcal{M}_j \widetilde{f}_j^{m+\frac{1}{2}} = D_{1j} \widetilde{f}_{j+1}^{m} + D_{2j} \widetilde{f}_j^{m} + D_{3j} \widetilde{f}_{j-1}^{m}, \tag{45}$$

where

$$\mathcal{M}_j = \left(I - \frac{\delta t}{2} \bar{A}_j\right) \tag{46}$$

and $\bar{A}_j$ is a tridiagonal matrix whose entries are given by

$$\begin{aligned}
(\bar{A}_j)_{i,i-1} &= \alpha_{i-1,j}^{m+\frac{1}{2}}/h, \quad 2 \leq i \leq N_x, \\
(\bar{A}_j)_{i,i} &= -\left(\alpha_{i-1,j}^{m+\frac{1}{2}} \overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\right)/h, \quad 1 \leq i \leq N_x, \\
(\bar{A}_j)_{i,i+1} &= \alpha_{i,j}^{m+\frac{1}{2}} \overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}}/h, \quad 1 \leq i \leq N_x - 1,
\end{aligned} \tag{47}$$

D_{1j}, D_{2j}, D_{3j} are diagonal matrices of order $N_x - 1$, whose ith diagonal entries are given by

$$\begin{aligned}
(D_{1j})_i &= \frac{\delta t}{2h} \beta_{ij}^{m} \overline{w}_{i,j+\frac{1}{2}}^{m}, \\
(D_{2j})_i &= 1 - \frac{\delta t}{2h}\left(\beta_{i,j-1}^{m} \overline{w}_{i,j-\frac{1}{2}}^{m} + \beta_{i,j}^{m}\right), \\
(D_{3j})_i &= \frac{\delta t}{2h} \beta_{i,j-1}^{m}.
\end{aligned} \tag{48}$$

For our forthcoming discussions, we define the logarithmic norm-1 of a matrix $\mu_1(A) = \lim_{\tau\to 0}(\|I + \tau A\|_1 - 1)/\tau$, that is $\mu_1(A) = \max_i(a_{ii} + \sum_{l\neq i}|a_{li}|)$. We also introduce the following compatible vector norms $\|f\|_1 = \sum_{i,j=0}^{n}|f_{i,j}|,\ f \in \mathbb{R}^{n^2}$ and $\|M\|_1 = \max_{j=1,\dots,n}\sum_{i=1}^{n}|M_{ij}|,\ M \in \mathbb{R}^{n^2\times n^2}$.

Remark 4.1 The matrices $\mathcal{M}_j$ defined in (46) and (47) are non singular. We recall the following theorem [23]: let A be a matrix and $\mu_1(A)$ the logarithmic norm-1, then $\mu_1(A) \le \omega$ iff $\|(I - \tau A)^{-1}\|_1 \le 1/(1-\tau\omega)$ for $\tau\omega < 1$. Since $\alpha_{ij}^{m+\frac{1}{2}} > 0$ it is $\mu_1(\bar{A}_j) = 0$, then from (46) follows that $\mathcal{M}_j^{-1}$ exists and $\|\mathcal{M}_j^{-1}\|_1 \le 1$.

Remark 4.2 We can show that $\|\mathcal{M}_j\|_1 = 1$. Similarly, as it has been shown in [27,34], if the following condition is satisfied

$$\delta t < \frac{2}{\Gamma}, \tag{49}$$

where $\Gamma > 0$ is the Lipschitz constant of the control u defined in (33), then $\mathcal{M}_j$ is an M-matrix. In fact, $\mathcal{M}_j$ is an M-matrix if $S_j = I - \text{diag}(\mathcal{M}_j)\mathcal{M}_j$ is positive and a convergent matrix [39], where $\text{diag}(\mathcal{M}_j)$ is the diagonal part of $\mathcal{M}_j$, that is from (47)

$$\text{diag}(\mathcal{M}_j)_{ii} = 1 + \frac{\delta t}{2h}\left(\alpha_{i-1,j}^{m+\frac{1}{2}}\overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\right) \quad 1 \le i \le N_x.$$

Thus S has the following non-vanishing elements

$$(S_j)_{i,i-1} = \frac{\alpha_{i-1,j}^{m+\frac{1}{2}}}{\dfrac{2h}{\delta t} + \left(\alpha_{i-1,j}^{m+\frac{1}{2}}\overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\right)},$$

$$(S_j)_{i,i+1} = \frac{\alpha_{i,j}^{m+\frac{1}{2}}\overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}}}{\dfrac{2h}{\delta t} + \left(\alpha_{i-1,j}^{m+\frac{1}{2}}\overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\right)}.$$

It is immediate to see from (44) that all these elements are non-negative. Now S is a convergent matrix if the spectral radius is less than 1. From the Gerschgorin theorem this is verified when $\sum_{\substack{j=1,\\ j\neq i}}^{N_x}|S_{ij}| < 1$, that is

$$\frac{\alpha_{i-1,j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}}}{\dfrac{2h}{\delta t} + \left(\alpha_{i-1,j}^{m+\frac{1}{2}}\overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} + \alpha_{i,j}^{m+\frac{1}{2}}\right)} < 1,$$

which is equivalent to the condition

$$\alpha_{i,j}^{m+\frac{1}{2}} \left(\overline{w}_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - 1 \right) - \alpha_{i-1,j}^{m+\frac{1}{2}} \left(\overline{w}_{i-\frac{1}{2},j}^{m+\frac{1}{2}} - 1 \right) < \frac{2h}{\delta t}.$$

By using the first equation of (44) we get $u^1_{i-\frac{1}{2},j,m+\frac{1}{2}} - u^1_{i+\frac{1}{2},j,m+\frac{1}{2}} < 2h/\delta t$. Using (33), we can show that $|u^1_{i+\frac{1}{2},j,m+\frac{1}{2}} - u^1_{i-\frac{1}{2},j,m+\frac{1}{2}}| < \Gamma h$. Thus, the condition $\Gamma h < 2h/\delta t$ implies S is a convergent matrix. Therefore, $\mathcal{M}_j^{-1}$ exists with non-negative entries under the condition $\delta t < 2/\Gamma$.

We now want to derive some numerical properties of our ADI-CC scheme. In view of this, we consider the FP equations (13) with a source term $g(x,t)$. The corresponding ADI-CC scheme is given as follows

$$\begin{aligned}
\frac{f_{i,j}^{m+\frac{1}{2}} - f_{i,j}^m}{\delta t/2} &= \frac{1}{h}\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - F_{i-\frac{1}{2},j}^{m+\frac{1}{2}} \right) + \frac{1}{h}\left(F_{i,j+\frac{1}{2}}^m - F_{i,j-\frac{1}{2}}^m \right) + g\left(x_{i,j}, t^m\right) \\
\frac{f_{i,j}^{m+1} - f_{i,j}^{m+\frac{1}{2}}}{\delta t/2} &= \frac{1}{h}\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - F_{i-\frac{1}{2},j}^{m+\frac{1}{2}} \right) + \frac{1}{h}\left(F_{i,j+\frac{1}{2}}^{m+1} - F_{i,j-\frac{1}{2}}^{m+1} \right) + g\left(x_{i,j}, t^{m+1}\right).
\end{aligned} \tag{50}$$

Let $f^m = (f^m_{1,1}, \ldots, f^m_{N_x-1,1}, \ldots, f^m_{1,2}, \ldots f^m_{N_x-1,2}, \ldots, f^m_{1,N_x-1}, \ldots, f^m_{N_x-1,N_x-1})$, then we can rewrite the first step of the ADI-CC scheme (45) or (50) as follows

$$\mathcal{M} f^{m+\frac{1}{2}} = \mathcal{D} f^m + \overline{g}^m_{i,j}, \quad m = 0, \ldots N_t - 1, \tag{51}$$

where $\overline{g} = \dfrac{\delta t}{2} g$. The matrix $\mathcal{M}$ is a block diagonal matrix with $(N_x - 1)$ blocks where the jth diagonal block is given by

$$\mathcal{M}_j = I - \frac{\delta t}{2} \bar{A}_j, \quad j = 1 \ldots N_x - 1,$$

and $\bar{A}^j$ is defined as in (47). The matrix $\mathcal{D}$ is given by

$$\begin{pmatrix}
(D_2^1)_1, \ldots, (D_2^1)_{N_x-1}, & (D_1^1)_1, \ldots (D_1^1)_{N_x-1}, & 0, \ldots & & \ldots, 0 \\
(D_3^2)_1, \ldots, (D_3^2)_{N_x-1}, & (D_2^2)_1, \ldots (D_2^2)_{N_x-1}, & (D_1^2)_1, \ldots (D_1^2)_{N_x-1}, & 0, \ldots & \ldots, 0 \\
0, \ldots\ldots\ldots\ldots\ldots 0, & (D_3^3)_1, \ldots (D_3^3)_{N_x-1}, & (D_2^3)_1, \ldots (D_2^3)_{N_x-1}, & 0, \ldots & \ldots, 0 \\
\vdots & \vdots & \vdots & \vdots & \\
0, \ldots & 0, \ldots & 0, \ldots & & (D_2^{N_x-1})_1, \ldots (D_2^{N_x-1})_{N_x-1}
\end{pmatrix}. \tag{52}$$

Remark 4.3 From (48), notice that the entries of the matrices D_{1j} and D_{3j} are non-negative. Moreover, the entries of D_{2j} are also non-negative provided that $\max_{i,j,m} \dfrac{\delta t}{2h} (\beta^m_{i,j-1} \overline{w}^m_{i,j-\frac{1}{2}} + \beta^m_{i,j}) \leq 1$. In fact, since u is bounded, by using (44), we can show that $\max_{i,j,m}(\beta^m_{i,j-1}\overline{w}^m_{i,j-\frac{1}{2}} + \beta^m_{i,j})h \le M$ with

$$M = \frac{h\underline{B}}{e^{2h\underline{B}/\sigma^2} - 1} + \frac{h\overline{B}}{1 - e^{-2h\overline{B}/\sigma^2}}, \quad \text{where}$$
$$\underline{B} = \min_{x,y,t} u_2(x,y,t), \ \overline{B} = \max_{x,y,t} u_2(x,y,t). \tag{53}$$

Thus the CFL-like condition

$$\delta t < \frac{2h^2}{M} \tag{54}$$

guarantees the non-negativity of D_{2j}. Finally, a direct calculation with (48) gives $\|\mathcal{D}\|_1 = 1$. We also note that $\lim_{h\to 0} M = \sigma^2 > 0$.

Next, we show that the ADI-CC scheme is positivity preserving. This is important because the solution of the continuous FP equation represents a probability density function which is non-negative at all times. Thus it is necessary for the numerical scheme to have the same property.

Theorem 4.1 *There exist two positive constants M and Γ such that for $\delta t < \min\{\frac{2}{\Gamma}, \frac{2h^2}{M}\}$, the ADI-CC scheme (38)–(39) applied to Eqs. (13) with (4) is positive preserving.*

Proof We note that the ADI-CC scheme is composed of two steps. We show positivity of both the steps that implies positivity of the ADI-CC scheme. The first step of the ADI-CC scheme is given by the first Eqs. (38)–(39).

Combining the statement of Remarks 4.2 and 4.3 we get the following condition of positivity of the first step of the ADI-CC scheme (38)–(39).

$$\delta t < \min\left\{\frac{2}{\Gamma}, \frac{2h^2}{M}\right\}. \tag{55}$$

In a similar way, (55) would ensure positivity of the second step of the ADI-CC scheme. Thus the ADI-CC scheme in (38) is positive under the CFL condition (55). □

We now show the stability of the ADI-CC scheme.

Theorem 4.2 *The ADI-CC scheme (38)–(39), related to Eqs. (13)–(4), is discrete L^1 stable under the CFL-like condition (55) i.e.,*

$$\|f^m\|_1 = \|f^0\|_1, \quad m = 0, \ldots, N_t - 1.$$

Proof From Lemma 4.1, we get

$$\sum_{i,j=0}^{n} f^m_{i,j} = \sum_{i,j=0}^{n} f^0_{i,j}, \quad \forall m = 1, \ldots, N_t, \tag{56}$$

Using Theorem 4.1, (56) can be written as

$$\sum_{i,j=0}^{n} |f_{i,j}^{m}| = \sum_{i,j=0}^{n} |f_{i,j}^{0}|, \quad \forall m = 1, \ldots, N_t, \tag{57}$$

which proves our result. □

Theorem 4.3 *The solution $f_{i,j}^m$ obtained from the ADI-CC scheme (38)–(39), related to Eqs. (13)–(4) with a source $g(x,t)$, under the CFL-like condition (55), satisfies the following inequality*

$$\|f^m\|_1 \leq \|f^0\|_1 + \delta t \sum_{n=0}^{m} \max\left(\|g^n\|_1, \|g^{n-1/2}\|_1\right), \quad m = 0, \ldots N_t - 1.$$

Proof Since $\mathcal{M}_j$ is invertible as shown in the Remarks 4.1 and 4.2, it follows that $\mathcal{M}$ is invertible. Thus, (51) can be written as follows

$$f^{m+\frac{1}{2}} = \mathcal{M}^{-1}\mathcal{D} f^m + \mathcal{M}^{-1}\overline{g}^m, \quad m = 0, \ldots N_t - 1. \tag{58}$$

Taking discrete L^1 norm on both sides of (58), we have

$$\|f^{m+\frac{1}{2}}\|_1 \leq \|\mathcal{M}^{-1}\|_1 \|\mathcal{D}\|_1 \|f^m\|_1 + \|\mathcal{M}^{-1}\|_1 \|\overline{g}^m\|_1, \quad m = 0, \ldots N_t - 1. \tag{59}$$

From Remarks 4.1 and 4.3, we know that both $\|\mathcal{M}^{-1}\|_1$ and $\|\mathcal{D}\|_1$ are equal to 1 under the condition (55). The same analysis can be performed for the second half time step of Eq. (50), and included in the previous inequality. Hence, by substituting the value of $\overline{g}$ with $\frac{\delta t}{2} g$ and calculating the summation, we get the required bound. □

Next, we investigate the consistency of the ADI-CC scheme (38)–(39). We define the following

$$\begin{aligned} \bar{D}_x f_{i,j}^m &= D_+ C_{i-\frac{1}{2},j}^{m+\frac{1}{2}} D_- f_i^m + D_+ B_{i-\frac{1}{2},j}^{m+\frac{1}{2}} M_\delta f_i^m, \\ \bar{D}_y f_{i,j}^m &= D_+ C_{i,j-\frac{1}{2}}^{m} D_- f_j^m + D_+ B_{i,j-\frac{1}{2}}^{m} M_\delta f_j^m, \end{aligned} \tag{60}$$

where

$$\begin{aligned} D_+ f_i &= (f_{i+1,j} - f_{i,j})/h, \\ D_- f_i &= (f_{i,j} - f_{i-1,j})/h, \\ M_\delta f_i &= (1 - \delta_{i-1}) f_{i,j} + \delta_{i-1} f_{i-1,j}, \\ D_+ f_j &= (f_{i,j+1} - f_{i,j})/h, \\ D_- f_j &= (f_{i,j} - f_{i,j-1})/h, \\ M_\delta f_j &= (1 - \delta_{j-1}) f_{i,j} + \delta_{j-1} f_{i,j-1}. \end{aligned}$$

Therefore (38), with the fluxes given in (34) and (35), can be written as follows

$$\begin{aligned}\frac{f_{i,j}^{m+\frac{1}{2}} - f_{i,j}^{m}}{\delta t/2} &= \frac{1}{h}\left(\bar{D}_x f_{i,j}^{m+\frac{1}{2}} + \bar{D}_y f_{i,j}^{m}\right)\\ \frac{f_{i,j}^{m+1} - f_{i,j}^{m+\frac{1}{2}}}{\delta t/2} &= \frac{1}{h}\left(\bar{D}_x f_{i,j}^{m+\frac{1}{2}} + \bar{D}_y f_{i,j}^{m+1}\right).\end{aligned} \tag{61}$$

Adding the two equations in (61), we get

$$\frac{f_{i,j}^{m+1} - f_{i,j}^{m}}{\delta t} = \frac{1}{2h}\left[2\bar{D}_x f_{i,j}^{m+\frac{1}{2}} + \bar{D}_y\left(f_{i,j}^{m} + f_{i,j}^{m+1}\right)\right]. \tag{62}$$

Subtracting the two equations in (61), we get

$$f_{i,j}^{m+\frac{1}{2}} = \frac{f_{i,j}^{m} + f_{i,j}^{m+1}}{2} - \frac{\delta t}{4h}\left[\bar{D}_y\left(f_{i,j}^{m+1} - f_{i,j}^{m}\right)\right]. \tag{63}$$

Substituting the value of $f_{i,j}^{m+\frac{1}{2}}$ from (63) in (62), we obtain

$$\frac{f_{i,j}^{m+1} - f_{i,j}^{m}}{\delta t} = \frac{1}{2h}\left[\left(\bar{D}_x + \bar{D}_y\right)\left(f_{i,j}^{m} + f_{i,j}^{m+1}\right)\right] - \frac{\delta t}{4h^2}\left[\bar{D}_x\bar{D}_y\left(f_{i,j}^{m+1} - f_{i,j}^{m}\right)\right]. \tag{64}$$

We define the truncation error as follows

$$\begin{aligned}\boldsymbol{\varphi}_{i,j}^{m+1} := &\frac{f(x_i, y_j, t_{m+1}) - f(x_i, y_j, t_m)}{\delta t}\\ &- \frac{1}{2h}\left[(\bar{D}_x + \bar{D}_y)(f(x_i, y_j, t_m) + f(x_i, y_j, t_{m+1})\right]\\ &+ \frac{\delta t}{4h^2}\left[\bar{D}_x\bar{D}_y(f(x_i, y_j, t_{m+1}) - f(x_i, y_j, t_m)\right].\end{aligned} \tag{65}$$

Lemma 4.2 *The truncation error* (65) *of the ADI-CC scheme* (38)–(39) *is of order* $\mathcal{O}(\delta t^2 + h^2)$ *under the CFL-like condition* (55).

Proof We can write the truncation error (65) as follows

$$\boldsymbol{\varphi}_{i,j}^{m+1} = T_1 + T_2,$$

where

$$T_1 = \frac{f(x_i, y_j, t_{m+1}) - f(x_i, y_j, t_m)}{\delta t}$$

$$-\frac{1}{2h}\left[(\bar{D}_x + \bar{D}_y)(f(x_i, y_j, t_m) + f(x_i, y_j, t_{m+1})\right],$$

and

$$T_2 = \frac{\delta t}{4h^2}\left[\bar{D}_x \bar{D}_y (f(x_i, y_j, t_{m+1}) - f(x_i, y_j, t_m)\right].$$

The term T_1 corresponds to the Crank–Nicholson method with CC discretization for space operator. Using similar arguments as in [27, Lemma 3.2 and Theorem 3.6] for $\bar{D}_x$ and $\bar{D}_y$, and by Taylor series expansion, one can show that under the CFL condition (55), the truncation error corresponding to the term T_1 is of $\mathcal{O}(\delta t^2 + h^2)$. In addition, for the term T_2, using similar arguments as in [27, Lemma 3.2 and Theorem 3.6], and Taylor series expansion, we have

$$\begin{aligned}
&\frac{1}{4h^2}\left[\bar{D}_x \bar{D}_y \left(f_{i,j}^{m+1} - f_{i,j}^m\right)\right] \\
&\quad = \delta t \left[\frac{\sigma^4}{4} \cdot \frac{\partial^5 f}{\partial t \partial x^2 \partial y^2} + \frac{\sigma^2}{2}\frac{\partial^4 (B_2 f)}{\partial t \partial x \partial y^2} + \frac{\sigma^2}{2}\frac{\partial^2}{\partial t \partial x}\left(B_1 \frac{\partial^2 f}{\partial y^2}\right) + \frac{\partial^2}{\partial t \partial x}\left(B_1 \left(\frac{\partial B_2 f}{\partial y}\right)\right)\right]_{i,j}^m \\
&\qquad + \mathcal{O}(\delta t \cdot h^2).
\end{aligned}$$

This gives

$$\frac{1}{4h^2}\left[\bar{D}_x \bar{D}_y \left(f_{i,j}^{m+1} - f_{i,j}^m\right)\right] \approx \mathcal{O}(\delta t + \delta t \cdot h^2),$$

and therefore

$$\frac{\delta t}{4h^2}\left[\bar{D}_x \bar{D}_y \left(f_{i,j}^{m+1} - f_{i,j}^m\right)\right] \approx \mathcal{O}(\delta t^2 + \delta t^2 h^2).$$

Thus, the truncation error (65) of the ADI-CC scheme (38)–(39) is at least of order $\mathcal{O}(\delta t^2 + h^2)$. □

Next, we define the global error as follows

$$e_{i,j}^m = f(x_{ij}, t^m) - f_{i,j}^m, \quad i, j = 0, \ldots, N_x, \; m = 1, \ldots, N_t.$$

Using Lemma 4.3 and Theorem 4.2, the error estimates for our ADI-CC scheme can be given as follows

Theorem 4.4 *The ADI-CC scheme* (38) *converges with an error of order* $\mathcal{O}(\delta t^2 + h^2)$ *under the CFL condition* (55) *in the discrete* L^1 *norm.*

Proof By the definition of truncation error, we have

$$\begin{aligned}
\frac{e_{i,j}^{m+1} - e_{i,j}^m}{\delta t} &= \frac{1}{2h}\left[\left(\bar{D}_x + \bar{D}_y\right)\left(e_{i,j}^m + e_{i,j}^{m+1}\right)\right] \\
&\quad - \frac{\delta t}{4h^2}\left[\bar{D}_x \bar{D}_y \left(e_{i,j}^{m+1} - e_{i,j}^m\right)\right] + \boldsymbol{\varphi}_{i,j}^{m+1}.
\end{aligned}$$

Thus, the solution error $e^m_{i,j}$ satisfies the discretized FP equation discussed above with the right hand side given by the truncation error function. Hence, using Theorem 4.3, we have

$$\|e^m\|_1 \leq \|e^0\|_1 + \delta t \sum_{n=0}^{m} \|\boldsymbol{\varphi}^n\|_1.$$

Therefore, from Lemma 4.2, the ADI-CC scheme converges with order $\mathcal{O}(\delta t^2 + h^2)$ in the discrete L^1 norm. □

5 A numerical scheme for the adjoint equations

We build the numerical scheme for the adjoint equations (14) and (16) by performing the variation on the discrete version of the Lagrangian (12), such as in the discretize-before-optimize approach [6]. Details of the derivation of the discrete adjoint equations with the boundary conditions using the discrete Lagrangian can be found in the "Appendix".

When the control cost $A(u)$ is given by (C1), the numerical scheme for the adjoint equation reads as follows

$$\begin{aligned}
\frac{p_{i,j}^{m+\frac{1}{2}} - p_{i,j}^{m+1}}{\delta t/2} &= \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+\frac{1}{2}} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+\frac{1}{2}}\right) \\
&\quad + \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+1} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+1} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+1} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+1}\right) \\
&\quad - \alpha V(x_{i,j} - x_t^{m+1}), \\
\frac{p_{i,j}^{m} - p_{i,j}^{m+\frac{1}{2}}}{\delta t/2} &= \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m}\right) \\
&\quad + \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m+\frac{1}{2}} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m+\frac{1}{2}}\right) \\
&\quad - \alpha V(x_{i,j} - x_t^{m+\frac{1}{2}}),
\end{aligned} \tag{66}$$

for all $(i, j) \in \{1, \ldots, N_x - 1\}$ and $m \in \{0, \ldots, N_t - 1\}$.

When the control cost $A(u)$ is given by (C2), the numerical scheme for the adjoint reads as follows

$$\begin{aligned}
\frac{p_{i,j}^{m+\frac{1}{2}} - p_{i,j}^{m+1}}{\delta t/2} &= \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+\frac{1}{2}} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+\frac{1}{2}}\right) \\
&\quad + \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+1} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+1} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+1} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+1}\right) \\
&\quad - \alpha V\left(x_{i,j} - x_t^{m+1}\right) \\
&\quad - \frac{\nu}{2}|u_{i,j}^{m+1}|^2 - \frac{\nu}{2}\left|\frac{u_{i+1,j}^{m+1} - u_{i,j}^{m+1}}{h}\right|^2 - \frac{\nu}{2}\left|\frac{u_{i,j-1}^{m+1} - u_{i,j}^{m+1}}{h}\right|^2,
\end{aligned} \tag{67}$$

$$\frac{p_{i,j}^{m} - p_{i,j}^{m+\frac{1}{2}}}{\delta t/2} = \frac{1}{h}(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m})$$

$$+\frac{1}{h}\left(K^{m+\frac{1}{2}}_{i-\frac{1}{2},j}p^{m+\frac{1}{2}}_{i-1,j} - R^{m+\frac{1}{2}}_{i+\frac{1}{2},j}p^{m+\frac{1}{2}}_{i,j} - K^{m+\frac{1}{2}}_{i-\frac{1}{2},j}p^{m+\frac{1}{2}}_{i,j} + R^{m+\frac{1}{2}}_{i+\frac{1}{2},j}p^{m+\frac{1}{2}}_{i+1,j}\right)$$
$$-\alpha V\left(x_{i,j} - x_t^{m+\frac{1}{2}}\right)$$
$$-\frac{\nu}{2}|u^{m+\frac{1}{2}}_{i,j}|^2 - \frac{\nu}{2}\left|\frac{u^{m+\frac{1}{2}}_{i+1,j} - u^{m+\frac{1}{2}}_{i,j}}{h}\right|^2 - \frac{\nu}{2}\left|\frac{u^{m+\frac{1}{2}}_{i,j-1} - u^{m+\frac{1}{2}}_{i,j}}{h}\right|^2, \qquad (68)$$

for all $(i, j) \in \{1, \dots, N_x - 1\}$ and $m \in \{0, \dots, N_t - 1\}$. The terms $K^m_{i,j}$ and $R^m_{i,j}$ are defined in Appendix. In both cases, we have the following terminal condition

$$p^{N_t}_{i,j} = -\beta V(x_{i,j} - x_T). \qquad (69)$$

The boundary conditions for the adjoint equations are implemented as follows

$$\begin{aligned} p^m_{N_x,j} &= p^m_{N_x-1,j} \quad \forall j = 0, \dots, N_x, \\ p^m_{i,N_x} &= p^m_{i,N_x-1} \quad \forall i = 0, \dots, N_x. \end{aligned} \qquad (70)$$

Next, we consider the discretization of the optimality conditions by direct use of the five-point finite difference Laplacian Δ^h and the first order difference operators to (15) and (17), corresponding to the control costs (C1) and (C2). Notice that in doing this, we abandon the discretize-before-optimize approach for the sake of simplifying the numerical implementation of the optimality conditions. In fact, by pursuing this approach, the structure of the resulting reduced gradient appears much more complicated.

In the case (15), we consider the following numerical evaluation of the reduced gradient

$$\begin{aligned} (\nabla \hat{J}(u))^1_{ijm} &= \nu u^1_{i,j,m} - \nu\frac{u^1_{i+1,j,m} - 2u^1_{i,j,m} + u^1_{i-1,j,m}}{h^2} - \nu\frac{u^1_{i,j+1,m} - 2u^1_{i,j,m} + u^1_{i,j-1,m}}{h^2} \\ &\quad - f(i,j,m)\cdot\frac{p(i+1,j,m) - p(i,j,m)}{h}, \\ (\nabla \hat{J}(u))^2_{ijm} &= \nu u^2_{i,j,m} - \nu\frac{u^2_{i+1,j,m} - 2u^2_{i,j,m} + u^2_{i-1,j,m}}{h^2} - \nu\frac{u^2_{i,j+1,m} - 2u^2_{i,j,m} + u^2_{i,j-1,m}}{h^2} \\ &\quad - f(i,j,m)\cdot\frac{p(i,j+1,m) - p(i,j,m)}{h}. \end{aligned} \qquad (71)$$

For the optimality condition (17), we get

$$\begin{aligned} (\nabla \hat{J}(u))^1_{ijm} &= \left(\nu u^1_{i,j,m} - \nu\frac{u^1_{i+1,j,m} - 2u^1_{i,j,m} + u^1_{i-1,j,m}}{h^2} - \nu\frac{u^1_{i,j+1,m} - 2u^1_{i,j,m} + u^1_{i,j-1,m}}{h^2}\right. \\ &\quad \left. - \cdot\frac{p(i+1,j,m) - p(i,j,m)}{h}\right) f(i,j,m), \\ (\nabla \hat{J}(u))^2_{ijm} &= \left(\nu u^2_{i,j,m} - \nu\frac{u^2_{i+1,j,m} - 2u^2_{i,j,m} + u^2_{i-1,j,m}}{h^2} - \nu\frac{u^2_{i,j+1,m} - 2u^2_{i,j,m} + u^2_{i,j-1,m}}{h^2}\right. \\ &\quad \left. - \cdot\frac{p(i,j+1,m) - p(i,j,m)}{h}\right) f(i,j,m), \end{aligned} \qquad (72)$$

where $u^m_{i,j} = (u^1_{i,j,m}, u^2_{i,j,m})$, $0 \le m \le N_t$ and $0 \le i, j \le N_x - 1$.

As discussed in Sect. 2, the grid functions $(\nabla \hat{J}(u))^1$ and $(\nabla \hat{J}(u))^2$ provide the right-hand sides of the following discretized elliptic problems

$$-\Delta^h v_k + v_k = (\nabla \hat{J}(u))^k \text{ in } \Omega_h, \quad v_k = 0 \text{ on } \partial\Omega_h,$$

where $k = 1, 2$, and v_k represents the kth component of the reduced H^1 gradient, $(\nabla \hat{J}(u)_{H^1})^k = v_k$.

6 A projected NCG optimization scheme

We solve the optimization problem (10) by implementing a projected non-linear conjugate scheme (NCG); see [29]. Such a scheme is an extension of the conjugate gradient method to constrained non-linear optimization problems. To describe this iterative method, we start with an initial guess u_0 for the control function and corresponding search direction

$$d_0 = g_0 := \nabla \hat{J}(u_0)_{H^1},$$

where $\nabla \hat{J}$ corresponds to (71) in the (C1) case, and to (72) in the (C2) case. The search directions are obtained recursively as

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \tag{73}$$

where $g_k = \nabla \hat{J}(u_k),\ k = 0, 1, \ldots$ and the parameter β_k is chosen according to the formula of Hager–Zhang [22] given by

$$\beta_k^{HG} = \frac{1}{d_k^T y_k} \left(y_k - 2 d_k \frac{\|y_k\|^2}{d_k^T y_k} \right)^T g_{k+1}, \tag{74}$$

where $y_k = g_{k+1} - g_k$ and $\| \cdot \|^2 = \langle \cdot, \cdot \rangle$. We update the value of the control u with a steepest descent scheme given as follows

$$u_{k+1} = u_k + \alpha_k\, d_k, \tag{75}$$

where k is a index of the iteration step and $\alpha_k > 0$ is a steplength obtained using a line search algorithm as in [2]. For this line search, we use the following Armijo condition of sufficient decrease of $\hat{J}$

$$\hat{J}(u_k + \alpha_k d_k) \le \hat{J}(u_k) + \delta \alpha_k \langle \nabla \hat{J}(u_k)_{H^1}, d_k \rangle_{\delta t, h}, \tag{76}$$

where $0 < \delta < 1/2$ and the scalar product $\langle u, v \rangle_{\delta t, h} = \delta t \sum_{m=0}^{N_t} \langle u, v \rangle_{H^1_h}$. For the definition of the discrete H^1_h scalar product we refer to [24].

Notice that this gradient procedure should be combined with a projection step onto U_{ad}. Therefore, we consider the following

$$u_{k+1} = P_U\left[u_k + \alpha_k\, d_k\right], \tag{77}$$

where

$$P_U\left[u\right] = \max\{u_a, \min\{u_b, u\}\}.$$

The projected NCG scheme can be summarized as follows.

Algorithm 6.1 (Projected NCG Scheme)

1. *Input: initial approx.* u_0. *Evaluate* $d_0 = -\nabla \hat{J}(u_0)_{H^1}$, *index* $k = 0$, *maximum* $k = k_{max}$, *tolerance* $= tol$.
2. *While* $(k < k_{max})$ *do*
3. *Set* $u_{k+1} = P_U\left[u_k + \alpha_k\, d_k\right]$, *where* α_k *is obtained using a line-search algorithm.*
4. *Compute* $g_{k+1} = \nabla \hat{J}(u_{k+1})_{H^1}$.
5. *Compute* β_k^{HG} *using* (74).
6. *Set* $d_{k+1} = -g_{k+1} + \beta_k^{HG} d_k$.
7. *If* $\|u_{k+1} - u_k\| < tol$, *terminate.*
8. *Set* $k = k + 1$.
9. *End while.*

The convergence of the projected NCG scheme Algorithm 6.1 is discussed in [29, Lemma 1.5, p. 235] and is stated as follows

Lemma 6.1 *Let* $\nabla \hat{J}$ *be bounded in a neighbourhood* N_{u^*} *of the optimal control* u^*, *where* $\hat{J}(u)$, *given in* (11), *is locally strictly convex, and let* $\alpha_k \geq \alpha^* > 0$ *for any* k, *where* α_k *is determined in Algorithm* 6.1. *If the sequence* $\{u_k\}$ *generated by Algorithm* 6.1 *satisfies*

$$\lim_{k\to\infty} \|u_{k+1} - u_k\| = 0,$$

in the Euclidean norm in Q, *then* $\{u_k\}$ *is a minimizing sequence in the sense that*

$$\lim_{k\to\infty} \hat{J}(u_k) = \inf\{\hat{J}(u) : u \in N_{u^*}\}.$$

The numerical optimization procedure is summarized as follows. We start with an initial guess for the control, u_k. We then solve the discrete forward FP equation using the ADI-CC scheme (38)–(39) and afterwards the associated discrete adjoint equation with the discrete schemes (66) (resp. (67)) corresponding to the control costs (C1) (resp. (C2).) Then, the gradient $\nabla_u \hat{J}(u_k)$ is evaluated using (71) and (72), which is then used in the projected NCG scheme Algorithm 6.1, iteratively, to obtain the update u_{k+1} to the discrete optimal control.

Notice that, in the case (C2), the optimization procedure above is computationally convenient but not necessary. In fact, one could decide to solve the adjoint

Table 1 Convergence of the ADI-CC scheme

N_x	N_t	Relative L^1 Error	Order
25	25	8.34e−5	–
50	50	2.01e−5	2.05
100	100	4.93e−6	2.02
200	200	1.22e−6	2.01

FP problem (backwards) together with the elliptic variational inequality of the control (at each time step). However, solving a system of a parabolic equation and an elliptic variational inequality (or even equality) corresponds to solving a nonlinear and nonsmooth evolution problem, which is very challenging with respect to the solution of the FP optimality system by well-known numerical optimization techniques.

In the next section, we validate our optimal control strategy and the corresponding numerical setup with test cases.

7 Numerical results

In this section, we present results of numerical experiments to validate our control strategies and the obtained numerical analysis estimates. Specifically, in the first part of this section, we verify second-order accuracy of our ADI-CC scheme. In the second part of this section, we consider the control of a two dimensional stochastic motion model by computing the optimal controls for the given objectives.

To demonstrate the accuracy of the ADI-CC scheme as proved in Theorem 4.4, we use the method of manufactured solutions to construct an exact solution for the FP equation (13) with a non-zero source term $g(x_1, x_2, t)$ on the right hand side. We set $u_1(x_1, x_2, t) = -x_1, u_2(x_1, x_2, t) = -x_2, \sigma = 1$. Further, we choose $T = 1$ and $\Omega = (-6, 6) \times (-6, 6)$. We assume zero-flux boundary condition. Choosing $g(x_1, x_2, t) = -1/\exp(x_1^2+x_2^2+t)$ and initial condition $f_0(x_1, x_2) = 1/\exp(x_1^2+x_2^2)$, we have the exact solution as $f_{ex} = 1/\exp(x_1^2+x_2^2+t)$. The size of the solution error is evaluated based on the following discrete L^1 norm

$$|||f|||_1 = h^2 \delta t \sum_{m=0}^{N_t} \sum_{i,j=0}^{N_x} |f_{i,j}^k|,$$

which we identify with $L^2_{\delta t}(0, T; L^1_h)$. The discrete relative L^1 error is defined as follows

$$|||f - f_{ex}|||_{1,r} = \frac{|||f - f_{ex}|||_1}{|||f_{ex}|||_1}.$$

Table 1 shows the results of experiments that evaluate the accuracy of the ADI-CC numerical scheme. We see that the resulting order of convergence is $\mathcal{O}(h^2 + \delta t^2)$.

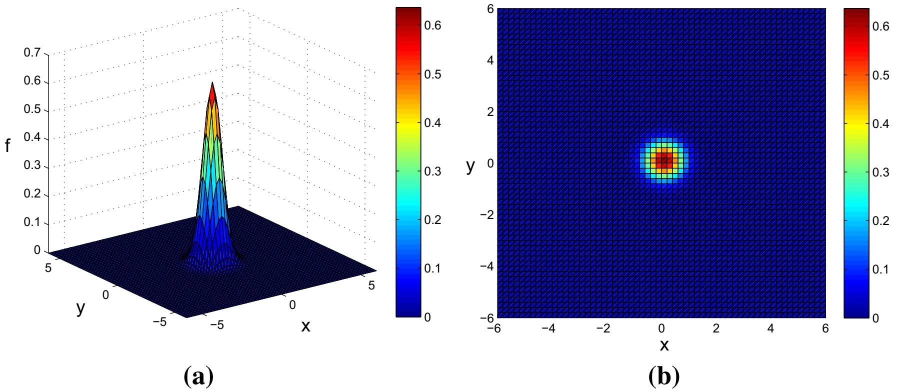


Fig. 2 Initial PDF centered at (0, 0). **a** The 3D view of the Gaussian PDF and **b** the 2D view of the same Gaussian PDF

We now discuss results concerning our control framework. We assume that the initial probability density function of the crowd is given, which is denoted with f_0. Our purpose is to determine an optimal drift that drives the crowd from a given initial distribution, along a given trajectory, to a terminal distribution at final time T. We consider two classes of trajectories: a single trajectory and a trajectory with obstacles. We implement the optimization algorithms described in Sect. 6 to solve the minimization problem (10) and thus determine the optimal control $u(x, t)$. We set the values of $\alpha = \beta = 1$ in (9).

We consider a two-dimensional stochastic process as follows

$$\begin{aligned} dX_1(t) &= u_1(X_1(t), X_2(t), t)dt + \sigma dW_1(t), \\ dX_2(t) &= u_2(X_1(t), X_2(t), t)dt + \sigma dW_2(t), \end{aligned} \tag{78}$$

where $X_1(t)$ and $X_2(t)$ represent the coordinates of the position of the individual at time t; $dW_1(t)$ and $dW_2(t)$ represent random infinitesimal increments of two mutually independent normalized Wiener processes. We take $\Omega = (-a, a) \times (-a, a)$ with $a = 6$.

We have that the drifts corresponds to u_1 and u_2, i.e., the control process, u_1 and u_2 represents the velocity of the particles. The diffusion is given as $\sigma = 1$. The initial PDF $f_0(x)$ is given by

$$f_0(x) = \hat{C} e^{-\{(x_1 - A_1)^2 - (x_2 - A_2)^2\}/0.5}, \tag{79}$$

where $(A_1, A_2) = x_t(0)$ is the starting point of the trajectory x_t, $\hat{C}$ is a normalization constant such that

$$\int_\Omega f_0(x) dx = 1$$

and $x = (x_1, x_2)$. A plot of the initial PDF is shown in Fig. 2.

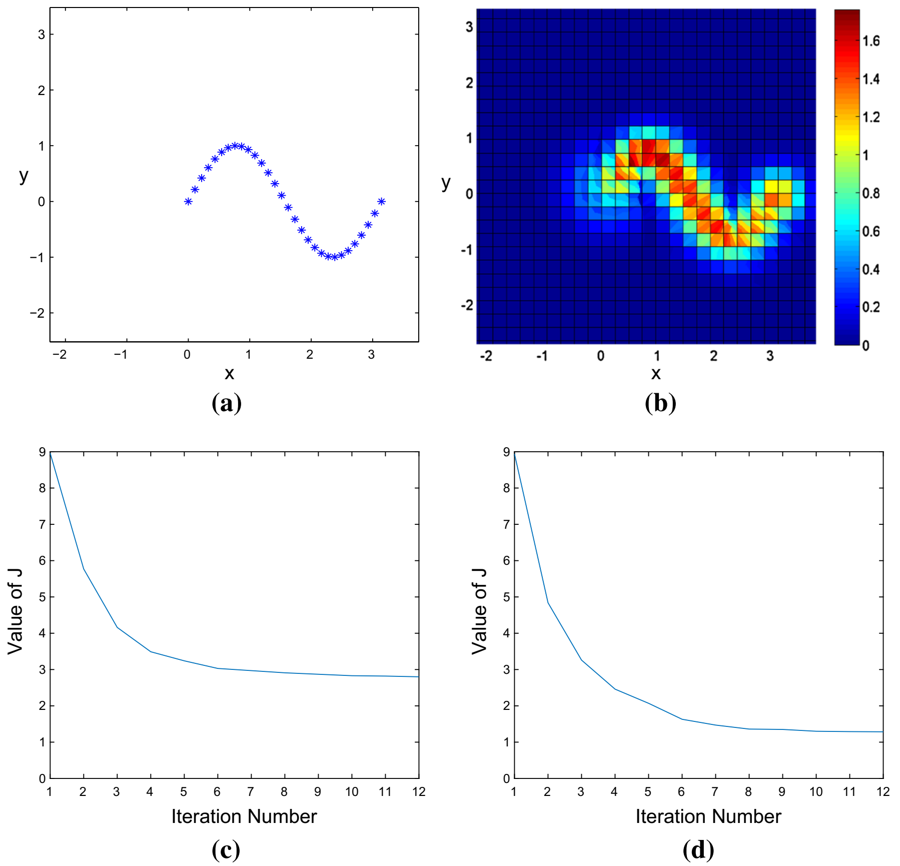


Fig. 3 Results of the numerical experiments with the controlled random process along sinusoidal trajectory. **a** Desired trajectory. **b** Resulting/controlled PDF. **c** Convergence of NCG with control cost (C1). **d** Convergence of NCG with control cost (C2)

We choose the control bounds $u_a = -5$ and $u_b = 5$. The total number of spatial grid points is $N_x = 60$ and temporal grid points is $N_t = 60$. The parameters for Algorithm 3 are: $tol = 10^{-4}$, $k_{max} = 30$. The parameters for the Armijo condition for Algorithm 4 are: $\delta = 0.1$, $k_{max} = 10$. The value of ν is taken to be 10^{-2}.

Our aim is to control the evolution of the PDF of (78) to follow a given trajectory x_t. We use the potential function V given by $V(x,t) = (x - x_t)^4$. Our desired path is given by the sinusoidal trajectory $x_t = (t, \sin(2t)),\ \ t \in [0, \pi]$. In correspondence to this setting, we solve the optimal control problem (10) to get the optimal control u.

From Fig. 3, we see that the control u drives the PDF along the desired path. The projected NCG scheme converges in 12 iterations to minimum value of $J = 2.80$ corresponding to the control cost (C1) and to the minimum value of $J = 1.29$ corresponding to the control cost (C2). The corresponding convergence history is shown in Fig. 3c, d. Figure 4 depicts the components of the control $u(x,t)$ at different times, showing that the constraints are active along the evolution.

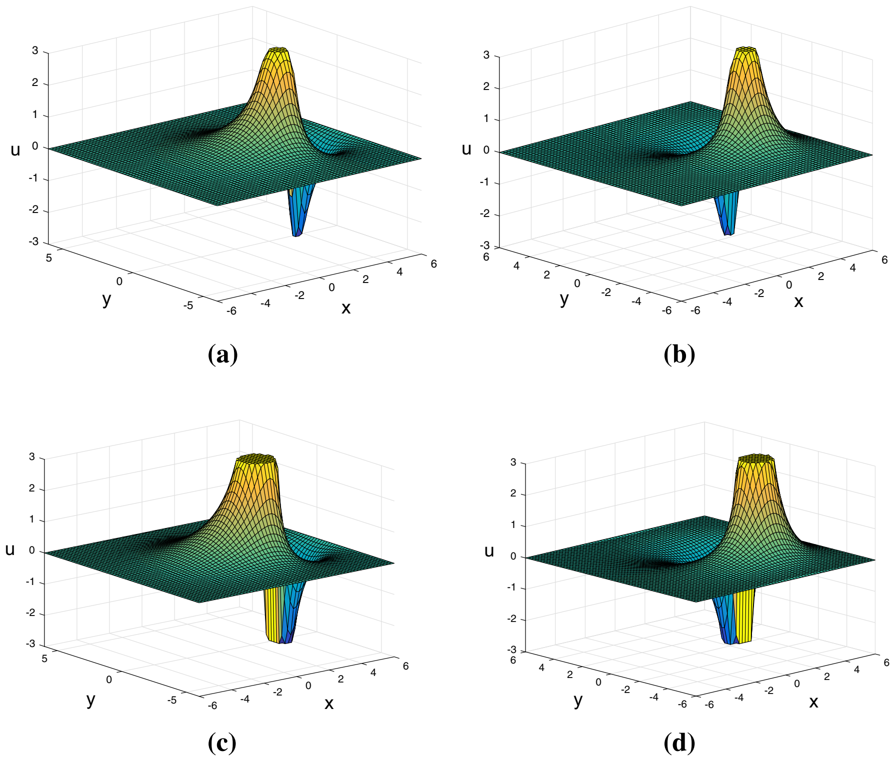
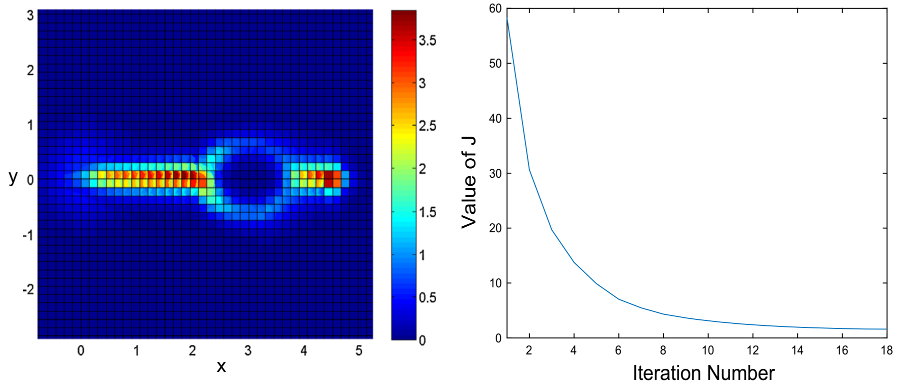


Fig. 4 Components of control u at time $t = \pi$ for the controlled random process along sinusoidal trajectory with the control costs (C1) and (C2).**a** With control cost (C1), u_1 at time $t = \pi$. **b** With control cost (C1), u_2 at time $t = \pi$. **c** With control cost (C2), u_1 at time $t = \pi$. **d** With control cost (C2), u_2 at time $t = \pi$

Next, we discuss the case of motion in the presence of an obstacle. An obstacle is represented by a high-valued concave potential function. In our framework, we use the initial PDF as given in (79) starting at the point (0, 0). The desired trajectory is given by $x_t = (1.5t, 0)$ and the potential V is given by

$$V(x,t) = \begin{cases} 100, & (x_1 - 3)^2 + x_2^2 \leq 0.2^2 \\ (x_1 - 1.5t)^2 + x_2^2, & \text{otherwise}, \end{cases} \tag{80}$$

where the obstacle is modelled by a cylinder centered at (3, 0) and radius 0.2. The time interval is chosen as $t \in [0, 2]$. In correspondence to this setting, we solve the optimal control problem (10) to get the optimal control u. Figure 5 shows that the PDF evolves along the desired path while avoiding the obstacle. The projected NCG scheme converges in 18 iterations to the minimum value of $J = 1.61$ with control cost (C1) as compared to 12 iterations in the case without obstacles. The corresponding convergence history is shown in Fig. 5b. Thus we observe faster convergence to minimum in the case we do not have obstacles.



(a) **(b)**

Fig. 5 Results of the numerical experiments with the controlled random process along a desired trajectory with an obstacle and the control cost (C1). **a** Controlled PDF. **b** Convergence of NCG

8 Conclusions

In this work, two Fokker–Planck control-constrained strategies for collective motion were presented, which resulted in an open-loop and a closed-loop control schemes. For these problems, existence and regularity of optimal control solutions were discussed, and their computation by an alternate-direction implicit Chang–Cooper scheme was illustrated. This scheme was proven to be conservative, positive preserving, L^1 stable, and second-order accurate in space and time. The discretized FP optimality system was solved with a projected non-linear conjugate gradient scheme. The effectiveness of the proposed control framework was demonstrated by considering trajectories with and without obstacles.

Acknowledgements The authors would like to gratefully acknowledge the comments by the referees which helped to improve this paper. S. Roy would like to thank A. S. Vasudeva Murthy and Praveen Chandrashekar for several fruitful discussions during the initial phases of this work. This work was supported in part by the European Union under Grant Agreement No. 304617 Marie Curie Research Training Network “Multi-ITN STRIKE—Novel Methods in Computational Finance” and the BMBF project “ROENOBIO”. S. Roy was also supported by the DAAD Passage to India Program and the AIRBUS Group Corporate Foundation Chair in Mathematics of Complex Systems established in TIFR/ICTS, Bangalore.

Appendix: Derivation of the numerical adjoint

We derive the numerical scheme for the adjoint equation (14) using the discretize-before-optimize approach. The starting point of this derivation is the Lagrangian

$$L(f, u, p) = J(f, u) + \langle \partial_t f - \nabla \cdot F, p \rangle. \tag{81}$$

In order to obtain the discrete version of the adjoint equation, we need to consider a discrete version of the Lagrange function with the ADI-CC scheme for the time–space derivatives of f. Since the ADI-CC scheme has an intermediate time step $t^{m+\frac{1}{2}}$, we define the Lagrangian on the following double grid

$$
\begin{aligned}
Q^d_{h,\delta t} &= \{(x, t_m) : x \in \Omega_h,\ t_m = m\delta t,\ 0 \le m \le N_t\} \\
&\quad \cup \{(x, t_{m+\frac{1}{2}}) : x \in \Omega_h,\ t_{m+\frac{1}{2}} = \left(m + \frac{1}{2}\right)\delta t,\ 0 \le m \le N_t - 1\}.
\end{aligned}
\tag{82}
$$

The discrete Lagrangian is given by

$$
\begin{aligned}
\hat{L}(f, u, p) &= \alpha \sum_m \sum_{i,j=1}^{N_x-1} V\left(x_{i,j} - x_t^m\right) f_{i,j}^m\, h^2 \frac{dt}{2} \\
&\quad + \alpha \sum_m \sum_{i,j=1}^{N_x-1} V\left(x_{i,j} - x_t^{m+\frac{1}{2}}\right) f_{i,j}^{m+\frac{1}{2}}\, h^2 \frac{dt}{2} \\
&\quad + \beta \sum_{i,j=1}^{N_x-1} V\left(x_{i,j} - x_t^{N_t}\right) f_{i,j}^{N_t}\, h^2 + \frac{\nu}{2} \sum_m \sum_{i,j=1}^{N_x-1} A(u_{i,j}^m)\, h^2 \frac{dt}{2} \\
&\quad + \frac{\nu}{2} \sum_m \sum_{i,j=1}^{N_x-1} A\left(u_{i,j}^{m+\frac{1}{2}}\right) h^2 \frac{dt}{2} \\
&\quad + \sum_m \sum_{i,j=1}^{N_x-1} \frac{f_{i,j}^{m+\frac{1}{2}} - f_{i,j}^m}{\delta t/2} p_{i,j}^m\, h^2 \frac{dt}{2} \\
&\quad + \sum_m \sum_{i,j=1}^{N_x-1} \frac{f_{i,j}^{m+1} - f_{i,j}^{m+\frac{1}{2}}}{\delta t/2} p_{i,j}^{m+\frac{1}{2}}\, h^2 \frac{dt}{2} \\
&\quad - \sum_m \sum_{i,j=1}^{N_x-1} \left[\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - F_{i-\frac{1}{2},j}^{m+\frac{1}{2}}\right) + \left(F_{i,j+\frac{1}{2}}^m - F_{i,j-\frac{1}{2}}^m\right)\right] p_{i,j}^m\, h \frac{dt}{2} \\
&\quad - \sum_m \sum_{i,j=1}^{N_x-1} \left[\left(F_{i+\frac{1}{2},j}^{m+\frac{1}{2}} - F_{i-\frac{1}{2},j}^{m+\frac{1}{2}}\right) + \left(F_{i,j+\frac{1}{2}}^{m+1} - F_{i,j-\frac{1}{2}}^{m+1}\right)\right] p_{i,j}^{m+\frac{1}{2}}\, h \frac{dt}{2}.
\end{aligned}
\tag{83}
$$

We write the fluxes of the FP equation (13) in the following compact form

$$
F_{i+\frac{1}{2},j}^m = K_{i+\frac{1}{2},j}^m f_{i+1,j}^m - R_{i+\frac{1}{2},j}^m f_{i,j}^m,
\tag{84}
$$

where

$$
\begin{aligned}
K_{i+\frac{1}{2},j}^m &= (1 - \delta_i) B_{i+\frac{1}{2},j}^m + \frac{\sigma^2}{h}, \\
R_{i+\frac{1}{2},j}^m &= \frac{\sigma^2}{h} - \delta_i B_{i+\frac{1}{2},j}^m.
\end{aligned}
\tag{85}
$$

Similarly, we have

$$F^{m}_{i,j+\frac{1}{2}} = K^{m}_{i,j+\frac{1}{2}} f^{m}_{i,j+1} - R^{m}_{i,j+\frac{1}{2}} f^{m}_{i,j}. \tag{86}$$

Therefore, we obtain

$$\begin{aligned}\sum_{m}\sum_{i,j=1}^{N_x-1}\left(F^{m+\frac{1}{2}}_{i+\frac{1}{2},j} - F^{m+\frac{1}{2}}_{i-\frac{1}{2},j}\right)p^{m}_{i,j} = \sum_{m}\sum_{i,j=1}^{N_x-1}\Big(&K^{m+\frac{1}{2}}_{i+\frac{1}{2},j} f^{m+\frac{1}{2}}_{i+1,j} - R^{m+\frac{1}{2}}_{i+\frac{1}{2},j} f^{m+\frac{1}{2}}_{i,j}\\ &- K^{m+\frac{1}{2}}_{i-\frac{1}{2},j} f^{m+\frac{1}{2}}_{i,j} + R^{m+\frac{1}{2}}_{i-\frac{1}{2},j} f^{m+\frac{1}{2}}_{i-1,j}\Big)p^{m}_{i,j}.\end{aligned} \tag{87}$$

Rearranging the summation on the right-hand side of (87) to collect the terms $f^{m+\frac{1}{2}}_{i,j}$ with same space index and using discrete flux zero (39), we have

$$\begin{aligned}\sum_{m}\sum_{i,j=1}^{N_x-1}\left(F^{m+\frac{1}{2}}_{i+\frac{1}{2},j} - F^{m+\frac{1}{2}}_{i-\frac{1}{2},j}\right)p^{m}_{i,j} = \sum_{m}\sum_{i,j=1}^{N_x-1}\Big(&K^{m+\frac{1}{2}}_{i-\frac{1}{2},j} p^{m}_{i-1,j} - R^{m+\frac{1}{2}}_{i+\frac{1}{2},j} p^{m}_{i,j}\\ &- K^{m+\frac{1}{2}}_{i-\frac{1}{2},j} p^{m}_{i,j} + R^{m+\frac{1}{2}}_{i+\frac{1}{2},j} p^{m}_{i+1,j}\Big) f^{m+\frac{1}{2}}_{i,j}.\end{aligned} \tag{88}$$

In a similar way, we have

$$\begin{aligned}\sum_{m}\sum_{i,j=1}^{N_x-1}\left(F^{m}_{i,j+\frac{1}{2}} - F^{m}_{i,j-\frac{1}{2}}\right)p^{m}_{i,j} = \sum_{m}\sum_{i,j=1}^{N_x-1}\Big(&K^{m}_{i,j-\frac{1}{2}} p^{m}_{i,j-1} - R^{m}_{i,j+\frac{1}{2}} p^{m}_{i,j}\\ &-K^{m}_{i,j-\frac{1}{2}} p^{m}_{i,j} + R^{m}_{i,j+\frac{1}{2}} p^{m}_{i,j+1}\Big) f^{m}_{i,j},\end{aligned} \tag{89}$$

$$\begin{aligned}\sum_{m}\sum_{i,j=1}^{N_x-1}\left(F^{m+\frac{1}{2}}_{i+\frac{1}{2},j} - F^{m+\frac{1}{2}}_{i-\frac{1}{2},j}\right)p^{m+\frac{1}{2}}_{i,j} = \sum_{m}\sum_{i,j=1}^{N_x-1}\Big(&K^{m+\frac{1}{2}}_{i-\frac{1}{2},j} p^{m+\frac{1}{2}}_{i-1,j} - R^{m+\frac{1}{2}}_{i+\frac{1}{2},j} p^{m+\frac{1}{2}}_{i,j}\\ &- K^{m+\frac{1}{2}}_{i-\frac{1}{2},j} p^{m+\frac{1}{2}}_{i,j} + R^{m+\frac{1}{2}}_{i+\frac{1}{2},j} p^{m+\frac{1}{2}}_{i+1,j}\Big) f^{m+\frac{1}{2}}_{i,j},\end{aligned} \tag{90}$$

and

$$\begin{aligned}\sum_{m}\sum_{i,j=1}^{N_x-1}\left(F^{m+1}_{i,j+\frac{1}{2}} - F^{m+1}_{i,j-\frac{1}{2}}\right)p^{m+\frac{1}{2}}_{i,j} = \sum_{m}\sum_{i,j=1}^{N_x-1}\Big(&K^{m+1}_{i,j-\frac{1}{2}} p^{m+\frac{1}{2}}_{i,j-1} - R^{m+1}_{i,j+\frac{1}{2}} p^{m+\frac{1}{2}}_{i,j}\\ &- K^{m+1}_{i,j-\frac{1}{2}} p^{m+\frac{1}{2}}_{i,j} + R^{m+1}_{i,j+\frac{1}{2}} p^{m+\frac{1}{2}}_{i,j+1}\Big) f^{m+1}_{i,j}.\end{aligned} \tag{91}$$

For our convenience, using (88)–(91) in (83), rearranging the time indices and collecting the terms $f_{i,j}^{m+\frac{1}{2}}$ and $f_{i,j}^{m+1}$, we obtain the Lagrange function in a different form as follows

$$
\begin{aligned}
&\hat{L}_1(f,u,p)\\
&= \alpha \sum_m \sum_{i,j=1}^{N_x-1} V(x_{i,j}-x_t^{m+1}) f_{i,j}^{m+1}\, h^2 \frac{dt}{2} + \alpha \sum_m \sum_{i,j=1}^{N_x-1} V(x_{i,j}-x_t^{m+\frac{1}{2}}) f_{i,j}^{m+\frac{1}{2}}\, h^2 \frac{dt}{2}\\
&+ \beta \sum_{i,j=1}^{N_x-1} V(x_{i,j}-x_t^{N_t}) f_{i,j}^{N_t}\, h^2 + \frac{\nu}{2} \sum_m \sum_{i,j=1}^{N_x-1} A\left(u_{i,j}^m\right) h^2 \frac{dt}{2} + \frac{\nu}{2} \sum_{m=0}^{N_t-1} \sum_{i,j=1}^{N_x-1} A\left(u_{i,j}^{m+\frac{1}{2}}\right) h^2 \frac{dt}{2}\\
&+ \sum_m \sum_{i,j=1}^{N_x-1} \frac{p_{i,j}^m - p_{i,j}^{m+\frac{1}{2}}}{\delta t/2} f_{i,j}^{m+\frac{1}{2}}\, h^2 \frac{dt}{2} + \sum_m \sum_{i,j=1}^{N_x-1} \frac{p_{i,j}^{m+\frac{1}{2}} - p_{i,j}^{m+1}}{\delta t/2} f_{i,j}^{m+1}\, h^2 \frac{dt}{2}\\
&- \sum_m \sum_{i,j}^{N_x-1} \left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^m - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^m - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^m + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^m\right) f_{i,j}^{m+\frac{1}{2}} h \frac{dt}{2}\\
&- \sum_m \sum_{i,j}^{N_x-1} \left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+1} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+1} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+1} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+1}\right) f_{i,j}^{m+1} h \frac{dt}{2}\\
&- \sum_m \sum_{i,j}^{N_x-1} \left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m+\frac{1}{2}} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m+\frac{1}{2}}\right) f_{i,j}^{m+\frac{1}{2}} h \frac{dt}{2}\\
&- \sum_m \sum_{i,j}^{N_x-1} \left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+\frac{1}{2}} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+\frac{1}{2}}\right) f_{i,j}^{m+1} h \frac{dt}{2}.
\end{aligned}
\tag{92}
$$

When the control cost $A(u)$ is given by (C1), taking derivative with respect to f^{m+1}, we obtain the following first integration step for the adjoint equation

$$
\begin{aligned}
\frac{p_{i,j}^{m+\frac{1}{2}} - p_{i,j}^{m+1}}{\delta t/2} &= \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+\frac{1}{2}} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+\frac{1}{2}}\right)\\
&+ \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+1} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+1} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+1} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+1}\right)\\
&- \alpha V(x_{i,j} - x_t^{m+1}).
\end{aligned}
$$

Taking derivative with respect to $f_{i,j}^{m+\frac{1}{2}}$, we obtain the following second integration step for the adjoint equation

$$
\begin{aligned}
\frac{p_{i,j}^{m} - p_{i,j}^{m+\frac{1}{2}}}{\delta t/2} &= \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m}\right)\\
&+ \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m+\frac{1}{2}} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m+\frac{1}{2}}\right)\\
&- \alpha V\left(x_{i,j} - x_t^{m+\frac{1}{2}}\right),
\end{aligned}
$$

along with the terminal condition

$$p_{i,j}^{N_t} = -\beta V(x_{i,j} - x_T).$$

When the control cost $A(u)$ is given by (C2), taking derivative with respect to f^{m+1}, we obtain the following first integration step for the adjoint equation

$$\begin{aligned}
\frac{p_{i,j}^{m+\frac{1}{2}} - p_{i,j}^{m+1}}{\delta t/2} &= \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+\frac{1}{2}} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+\frac{1}{2}} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+\frac{1}{2}}\right) \\
&+ \frac{1}{h}\left(K_{i,j-\frac{1}{2}}^{m+1} p_{i,j-1}^{m+1} - R_{i,j+\frac{1}{2}}^{m+1} p_{i,j}^{m+1} - K_{i,j-\frac{1}{2}}^{m+1} p_{i,j}^{m+1} + R_{i,j+\frac{1}{2}}^{m+1} p_{i,j+1}^{m+1}\right) \\
&- \alpha V\left(x_{i,j} - x_t^{m+1}\right) - \frac{\nu}{2}|u_{i,j}^{m+1}|^2 - \frac{\nu}{2}\left|\frac{u_{i+1,j}^{m+1} - u_{i,j}^{m+1}}{h}\right|^2 - \frac{\nu}{2}\left|\frac{u_{i,j-1}^{m+1} - u_{i,j}^{m+1}}{h}\right|^2.
\end{aligned}$$

Taking derivative with respect to $f_{i,j}^{m+\frac{1}{2}}$, we obtain the following second integration step for the adjoint equation

$$\begin{aligned}
\frac{p_{i,j}^{m} - p_{i,j}^{m+\frac{1}{2}}}{\delta t/2} &= \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m}\right) \\
&+ \frac{1}{h}\left(K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i-1,j}^{m+\frac{1}{2}} - R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} - K_{i-\frac{1}{2},j}^{m+\frac{1}{2}} p_{i,j}^{m+\frac{1}{2}} + R_{i+\frac{1}{2},j}^{m+\frac{1}{2}} p_{i+1,j}^{m+\frac{1}{2}}\right) \\
&- \alpha V\left(x_{i,j} - x_t^{m+\frac{1}{2}}\right) - \frac{\nu}{2}|u_{i,j}^{m+\frac{1}{2}}|^2 - \frac{\nu}{2}\left|\frac{u_{i+1,j}^{m+\frac{1}{2}} - u_{i,j}^{m+\frac{1}{2}}}{h}\right|^2 - \frac{\nu}{2}\left|\frac{u_{i,j-1}^{m+\frac{1}{2}} - u_{i,j}^{m+\frac{1}{2}}}{h}\right|^2.
\end{aligned}$$

References

1. Annunziato, M., Borzì, A.: Optimal control of probability density functions of stochastic processes. Math. Model. Anal. **15**, 393–407 (2010)
2. Annunziato, M., Borzì, A.: A Fokker–Planck control framework for multidimensional stochastic process. J. Comput. Appl. Math. **237**, 487–507 (2013)
3. Annunziato, M., Borzì, A., Nobile, F., Tempone, R.: On the connection between the Hamilton–Jacobi–Bellman and the Fokker–Planck control frameworks. Appl. Math. **5**, 2476–2484 (2014)
4. Aronson, D.G.: Non-negative solutions of linear parabolic equations. Ann. della Scuola Normale Superiore di Pisa - Classe di Scienze **22**(4), 607–694 (1968)
5. Bellomo, N., Bellouquid, A., Knopoff, D.: From the microscale to collective crowd dynamics. Multiscale Model. Simul. **11**(3), 943–963 (2013)
6. Borzì, A., Schulz, V.: Computational Optimization of Systems Governed by Partial Differential Equations. SIAM Book Series on Computational Science and Engineering 08. SIAM, Philadelphia, PA (2012)
7. Chang, J.S., Cooper, G.: A practical difference scheme for Fokker–Planck equations. J. Comput. Phys. **6**, 1–16 (1970)
8. Cucker, F., Mordecki, E.: Flocking in noisy environments. J. Math. Pures Appl. **89**, 278–296 (2008)
9. Dautray, R., Lions, J.-L.: Problèmes d' évolution du premier ordre en temps, (Chap 18.) Masson, Paris (1984)
10. Deutsch, A., Theraulaz, G., Vicsek, T.: Collective motion in biological systems. Interface Focus **2**(6), 689–692 (2012)
11. Douglas Jr., J.: On the numerical integration of $u_{xx} + u_{yy} = u_t$ by implicit methods. J. Soc. Ind. Appl. Math. **3**, 42–65 (1955)

12. Douglas Jr., J.: Alternating direction methods for three space variables. Numer. Math. **4**(1), 41–63 (1962)
13. Droniou, J., Vázquez, J.-L.: Noncoercive convection-diffusion elliptic problems with Neumann boundary conditions. Calc. Var. Partial Differ. Equ. **34**, 413–434 (2008)
14. Evans, L.C.: Partial Differential Equations, Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence, RI (2002)
15. Fleming, W., Soner, M.: Controlled Markov Processes and Viscosity Solutions. Springer, Berlin (2006)
16. Flotron, S., Rappaz, J.: Conservation schemes for convection-diffusion equations with Robin boundary conditions. ESAIM **47**, 1765–1781 (2013)
17. Gautrais, J., Ginelli, F., Fournier, R., Blanco, S., Soria, M., Chateé, H., Theraulaz, H.G.: Deciphering interactions in moving animal groups. PLoS Comput. Biol. **8**(9), e1002 (2012)
18. Grosskinsky, S., Klingenberg, C., Oelschläger, K.: A rigorous derivation of Smoluchowski's equation in the moderate limit. Stoch. Anal. Appl. **22**(1), 113–141 (2004)
19. Glowinski, R.: Numerical Methods for Nonlinear Variational Problems. Springer, New York (1984)
20. Glowinski, R., Lions, J.-L., He, J.: Exact and approximate controllability for distributed parameter systems. Acta Numer. **3**, 269–378 (1994)
21. Grüne, L., Pannek, J.: Nonlinear Model Predictive Control, Theory and Algorithms, Communications and Control Engineering. Springer, Berlin (2011)
22. Hager, W.W., Zhang, H.: A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. **16**(1), 170–192 (2005)
23. Hundsdorfer, W., Verwer, J.G.: Numerical Solution of Time-Dependent Advection–Diffusion–Reaction Equations. Springer Series in Computational Mathematics (2010)
24. Jovanović, B.S., Süli, E.: Analysis of Finite Difference Schemes For Linear Partial Differential Equations with Generalized Solutions. Springer Series in Computational Mathematics (2014)
25. Lions, J.-L.: Quelque Methodes de Résolution des Problemes aux Limites Non Linéaires. Dunod-Gauth. Vill, Paris (1969)
26. Lions, J.-L.: Nonhomogeneous Boundary Value Problems and Applications. Springer, Berlin (1972)
27. Mohammadi, M., Borzì, A.: Analysis of the Chang–Cooper discretization scheme for a class of Fokker–Planck equations. J. Numer. Math. **23**, 271–288 (2015)
28. Morale, D., Capasso, V., Oelschlaeger, K.: An interacting particle system modelling aggregation behavior: from individuals to populations. J. Math. Biol. **50**(1), 49–66 (2005)
29. Neittaanmaki, P., Tiba, D.: Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications. Pure and Applied Mathematics. CRC Press, London (1994)
30. Øksendal, B.K.: Stochastic Differential Equations: An Introduction with Applications. Springer, Berlin (2003)
31. Peaceman, D.W., Rachford Jr., H.H.: The numerical solution of parabolic and elliptic differential equations. J. Soc. Ind. Appl. Math. **3**(1), 28–41 (1955)
32. Petrelli, L., Kearsely, A.J.: Wasserstein metric convergence method for Fokker–Planck equations with point controls. Appl. Math. Lett. **22**(7), 1130–1135 (2009)
33. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations. Springer, Berlin (1994)
34. Roy, S., Annunziato, M., Borzì, A.: A Fokker–Planck feedback control-constrained approach for modelling crowd motion. J. Comput. Theor. Transp. **45**(6), 442–458 (2016)
35. Sepúlveda, N., Petitjean, L., Cochet, O., Grasland-Mongrain, E., Silberzan, P., Hakim, V.: Collective cell motion in an epithelial sheet can be quantitatively described by a stochastic interacting particle model. PLoS Comput. Biol. **9**(3), e1002944 (2013)
36. Strandburg-Peshkin, A., Farine, D.R., Couzin, I.D., Crofoot, M.C.: Shared decision-making drives collective movement in wild baboons. Science **348**(6241), 1358–1361 (2015)
37. Tao, T.: Nonlinear Dispersive Equations: Local and Global Analysis. American Mathematical Society, Providence, RI (2006)
38. Tröltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods and Applications. American Mathematical Society, Providence, RI (2010)
39. Varga, R.S.: Matrix Iterative Analysis. Springer Series in Computational Mathematics. Springer, Berlin (2000)
40. Zienkiewicz, A., Barton, D.A.W., Porfiri, M., di Bernardo, M.: Data-driven stochastic modelling of zebrafish locomotion. J. Math. Biol. (2014). doi:10.1007/s00285-014-0843-2