Ensemble optimal control problems governed by kinetic models

by

Alfio Borzì

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

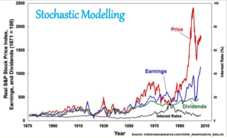
Framework

Many models in sciences and engineering aim at describing the dynamics of multiple agents subject to internal and external forces.
The manipulation of these forces allows to control the systems' dynamics in order to perform desired tasks.

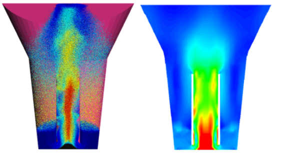
- A convenient description of the configuration of multi-agent (multi-particle) systems is achieved by means of probabilistic or material densities.
- The time evolution of these densities is governed by kinetic models.
- Optimal control theory provides the mathematical tools to formulate and solve control problems.
- Ensemble optimal control problems represent the natural framework for designing control mechanisms and objectives for systems governed by kinetic models.

Applications

collective motion, ©STIR



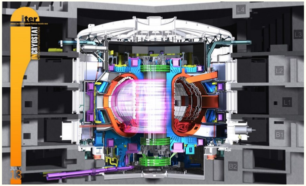
S & P stock price index



coating and mixing of powder, ©RCPE

Optimal control of stochastic and multi-particle systems to perform given tasks

space propulsion, ©SPARC



fusion reactor, ©ITER

Dynamical models

1. Deterministic models:

$$\dot{X}(t) = a(X(t), t)$$

2. Stochastic drift-diffusion-jump models:

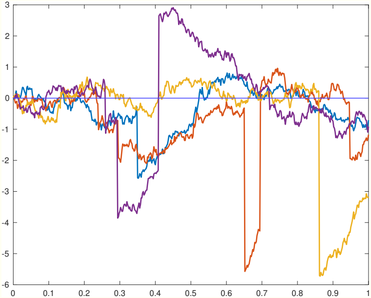
$$dX(t) = a(X(t), t)\,dt + \sigma(X(t), t)\,dW(t) + dP(t)$$

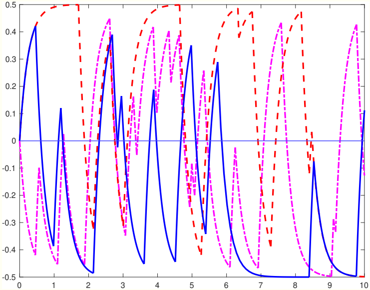
3. Piecewise-deterministic processes:

$$\dot{X}(t) = a_{\mathcal{S}(t)}(X(t), t), \quad t \in [0, \infty)$$

Most macroscopic evolution systems in, e.g., biology, climate, CFD, economics, ecology, finance, physics, etc., represent the emergent equations from the microscopic underpinnings of the models above (J. Hopfield).

Ensemble of trajectories





Left: Trajectories of a drift-diffusion-jump model: $a(x,t) = -4\,x$, $\sigma = 2$, and initial condition $X(0) = 0$.
Right: Trajectories of a PDP process: $a_1(x,t) = -4\,x + 2$, $a_2(x,t) = -4\,x - 2$, and initial condition for both states $X(0) = 0$.

In general, the initial condition $X(0) = X_0$ is given by means of a distribution function $f_0(x)$. In this case, also with deterministic models we obtain an ensemble of trajectories.

Density functions

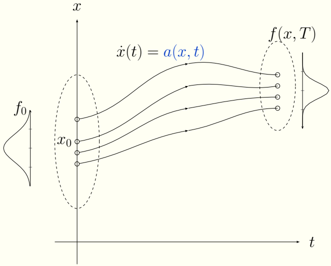
Consideration of all possible trajectories of a multi-particle system is an overwhelming task. For this reason, L. E. Boltzmann introduced the concept of material density $f(x,t)$.

In the non-interacting case, if $f_0(x)$ represents the initial density (configuration) at time $t = 0$, then the evolution of this density is modelled by

the Liouville equation

$$\partial_t f(x,t) + \operatorname{div}\left(a(x,t)\, f(x,t)\right) = 0,$$

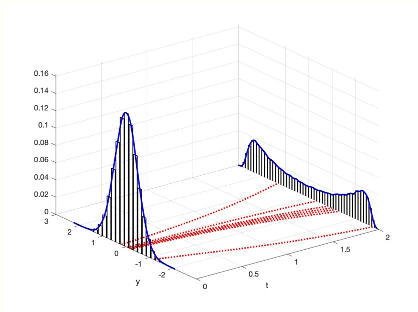
with drift a and initial condition $f(x,0) = f_0(x)$.



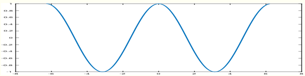
This fundamental result of statistical mechanics leads to the kinetic equations with applications in, e.g., space propulsion, electronic devices and materials, high-temperature plasma, etc..

Role of the drift

The Liouville equation $\partial_t f + \operatorname{div}\left(a(x,t)\, f\right) = 0$ is the fundamental continuity equation; the first in the hierarchy of kinetic models.



Take the dynamics
$\dot{X}(t) = \sin(X(t))$
$X(0) = X_0 \sim \mathcal{N}(\mu, \bar{\sigma}^2)$,
$\mu = 0$ and $\bar{\sigma} = 0.5$.



Notice that $a(x,t) = \sin(x) = -\frac{d}{dx}\cos(x) = -\nabla U(x)$ where $U(x) = \cos(x)$. The function U can be interpreted as a potential.

Compare with moments' equations in the case $\dot{X}(t) = [A(t)\, X(t) + b(t)]$:

$$\dot{\mu}(t) = A(t)\,\mu(t) + b(t), \qquad \dot{\Sigma}(t) = \Sigma(t)\, A(t)^\top + A(t)\,\Sigma(t).$$

A control mechanism

Optimal control applications require to identify a control mechanism in the model. We focus on time-dependent controls:

$$a(x,t;u_1,u_2) = a_0(x,t) + a_1\, u_1(t) + a_2\, u_2(t)\, x.$$

$a_0(x,t) \in \mathbb{R}^d$ smooth vector field, $a_1, a_2 \in \mathbb{R}$, $u_1(t), u_2(t) \in \mathbb{R}^d$

Moment equations: Define $m(t)$ as the mean, $v(t)$ as the variance. Choose $a(x,t;u_1,u_2) = u_1(t) + u_2(t)\, x$, and f_0 as normal Gaussian distribution.

From the Liouville equation, we obtain

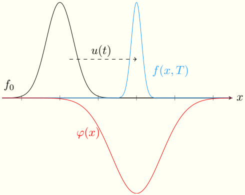
$$\begin{aligned} \dot{m}(t) &= u_1(t) + m(t)\, u_2(t), & m(0) &= m_0 \\ \dot{v}(t) &= 2\, v(t)\, u_2(t) & v(0) &= v_0. \end{aligned}$$

where

$m(t) = \int x\, f(x,t)\, dx$ and $v(t) = \int (x - m(t))^2\, f(x,t)\, dx$

Ensemble cost functional

The particles should follow a desired trajectory $x_D(t)$, $t \in [0, T]$, reach a target position x_T at $t = T$.



Ensemble control approach: define "attracting" potentials

$$\theta(x,t) = \Theta(|x - x_D(t)|)$$

$$\varphi(x) = \Phi(|x - x_T|)$$

Ensemble cost functional:

$$J(f,u) = \int_0^T \int_{\mathbb{R}^d} \theta(x,t)\, f(x,t)\, dx\, dt + \int_{\mathbb{R}^d} \varphi(x)\, f(x,T)\, dx + \kappa(u).$$

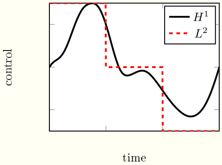
Cost of control

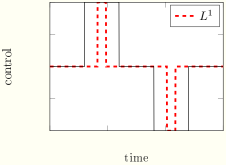
For time-dependent controls:

L^2: standard control cost.

H^1: includes time-derivative of the control (minimum attention control); turning control on at initial time and off at terminal time.

L^1: sparse controls (minimum action control)





Cost function:

$$\kappa(u) = \frac{\gamma}{2} \int_0^T \left|u(t)\right|^2 dt + \delta \int_0^T \left|u(t)\right| dt + \frac{\nu}{2} \int_0^T \left|\frac{d}{dt} u(t)\right|^2 dt$$

An ensemble control problem

A Liouville ensemble optimal control problem:

$$\min_{u\in U_{ad}} J(f,u) := \int_0^T \int_{\mathbb{R}^d} \theta(x,t)\, f(x,t)\, dx\, dt + \int_{\mathbb{R}^d} \varphi(x)\, f(x,T)\, dx$$
$$+ \frac{\gamma}{2} \int_0^T \left|u(t)\right|^2 dt + \delta \int_0^T \left|u(t)\right| dt + \frac{\nu}{2} \int_0^T \left|\frac{d}{dt}u(t)\right|^2 dt$$

subject to
$$\begin{cases} \partial_t f(x,t) + \operatorname{div}\big(a(x,t;u)\, f(x,t)\big) = 0 & \text{in } \mathbb{R}^d \times [0,T] \\ f(x,0) = f_0(x) & \text{in } \mathbb{R}^d \end{cases}$$

with the set of admissible controls

$$U_{ad} := \{u = (u_1, u_2) \in U \times U \,|\, u_a \le u(t) \le u_b,\ t \in [0,T]\},$$

$U = H^1([0,T];\mathbb{R}^d)$ or $U = L^2([0,T];\mathbb{R}^d)$.
$\gamma, \delta, \nu \ge 0, \gamma + \delta + \nu > 0, u_a < 0, u_b > 0$

Results with Liouville model

- For $u \in U_{ad}$ there exists a unique solution $f \in C\big([0,T]; H_k^m(\mathbb{R}^d)\big)$ of the Liouville initial-value problem.
- The control-to-state map G, $u \mapsto f = G(u)$ is Fréchet differentiable
- The ensemble optimal control problem admits at least one solution in U_{ad}.
- In the case $\nu = 0$, with the Lagrange multiplier q, and $\widehat{\lambda} \in \partial\left(||u||_{L^1}\right)$, the optimality system is given by

$$
\begin{aligned}
&\partial_t f + \operatorname{div}\big(a(x,t;u)\, f\big) = 0, \qquad f_{|t=0} = f_0 \\
&-\partial_t q - a(x,t;u)\cdot\nabla q = -\theta, \qquad q_{|t=T} = -\varphi \\
&\left(\gamma\, u_j^r + \widehat{\lambda}_j^r + \int_{\mathbb{R}^d} \operatorname{div}\left(\frac{\partial a}{\partial u_j^r} f\right) q\, dx\ ,\ v_j^r - u_j^r\right)_{L^2(0,T)} \geq 0 \quad v \in U_{ad}
\end{aligned}
$$

See later sections.

Further results

- Strong stability conserving Runge-Kutta method of second order in time and Kurganov-Tadmor scheme in space.
- In addition for the adjoint equation: second-order Strang splitting.
- Proved L^1 stability, second-order accuracy, positivity preserving.
- Projected semi-smooth Newton method.

See later sections.

S. Roy and A. Borzì, Numerical investigation of a class of Liouville control problems, Journal of Scientific Computing, 73 (2017), 178-202.

J. Bartsch, A. Borzì, F. Fanelli, S, Roy, A theoretical investigation of Brockett's ensemble optimal control problems Calculus of Variations and Partial Differential Equations, 58 (2019), 162.

J. Bartsch, A. Borzì, F. Fanelli, S. Roy, A numerical investigation of Brockett's ensemble optimal control problems, Numerische Mathematik, 149 (2021), 1-42.

Another ensemble control problem

In many physical systems, the density f is defined in the phase space spanned by position $x \in \Omega \subset \mathbb{R}^d$ and velocity $v \in \mathbb{R}^d$, $Q = \Omega \times \mathbb{R}^d$.

In this statistical framework, the time evolution of f can be governed by the following kinetic model (Vlasov+collision)

$$\begin{aligned} \partial_t f(x,v,t) + v \cdot \nabla_x f(x,v,t) + u(x) \cdot \nabla_v f &= C[f](v,t) \\ f|_{t=0} &= f_0, \\ f(x,v,t)|_{\partial\Omega \times \mathbb{R}^d_< \times (0,T]} &= f(x, v - 2n(n \cdot v), t) \end{aligned}$$

with specular reflection space boundaries; $\mathbb{R}^d_< := \{v \in \mathbb{R}^d \,|\, v \cdot n(x) < 0\}$, and space-dependent control field $u \in H^1_0(\Omega)$ is a force field (control), and $C[f]$ is the collision term.

We can define an ensemble optimal control problem with the functional:

$$J(f,u) \; := \int_0^T \int_Q \theta(x,v,t)\, f(x,t)\, dx\, dt \; + \int_Q \varphi(x,v)\, f(x,v,T)\, dx + \frac{\gamma}{2}\, \|u\|^2_{H^1}.$$

The Keilson-Storer model

We consider the Keilson-Storer (KS) collision model:

$$C[f](v,t) := \int f(w,t)\,A(w,v)\,dw - f(v,t)\int A(v,w)\,dw,$$

It has a gain – loss structure. We have $A(v,w) := A_0 e^{(-\beta|w-\gamma v|^2)}$ and $\gamma \in [-1,1]$, $A_0, \beta > 0$. For post-collision velocity holds $w \sim \mathcal{N}(\gamma v, (2\beta)^{-1})$.

- $\gamma \lessapprox 1$: weak collisions, Brownian motion
- $\gamma \approx 0$: strong collision, Bhatnagar-Gross-Krook (BGK) operator
- collision frequency $\frac{1}{\tau} = A_0\sqrt{\pi/\beta}$
- detailed balance: $A(w,v)\,f^{eq}(w) = A(v,w)\,f^{eq}(v)$
- equilibrium solution $f^{eq}(v)$ is the Maxwellian distribution
- A_0 and β related to the background density and temperature

Results with kinetic KS model

- Well-posedness of the control-to-state map
- Existence of optimal controls
- Monte Carlo solution of the kinetic model and of its adjoint
- Nonlinear conjugate gradient method

See later sections.

J. Bartsch, G. Nastasi, A. Borzì, Optimal Control of the Keilson-Storer Master Equation in a Monte Carlo Framework, Journal of Computational and Theoretical Transport, 50 (2021), 454-482.

J. Bartsch, A. Borzì, MOCOKI: A Monte Carlo approach for optimal control in the force of a linear kinetic model, Computer Physics Communications, 266 (2021), 108030.

Stochastic models and PDFs

Consider a continuous-time continuous-space stochastic process with $X : \mathcal{T} \times \Omega \to \mathbb{R}$, and $\text{Range}(X) = \mathbb{R}$. Since different random variables are labelled by different t in $\mathcal{T} = [0, +\infty)$, we can denote the probability density function (PDF) of the random variable X at $t \in \mathcal{T}$ with $f(\cdot, t)$.

Similarly, we denote with $f(\cdot, t_2 | v_1, t_1)$ the conditional probability density function of $X(t_2, \omega)$ given the occurrence of the value v_1 of $X(t_1, \omega)$ with $f(v_1, t_1) > 0$.

For a Markov process, we have $f(v, t) = \int_{\mathbb{R}} f(v, t|z, 0)\, f_0(z)\, dz$, and continuity implies the following identity

$$f(v, \tau|z, t) = \int_{\mathbb{R}} f(v, \tau|r, s)\, f(r, s|z, t)\, dr,$$

where $\tau > s > t$. This is the Chapman-Kolmogorov equation for the conditional PDFs.
We have $\int_{\mathbb{R}} f_0(x)\, dx = 1$, and so $\int_{\mathbb{R}} f(x, t)\, dx = 1$, $t \geq 0$.

Einstein, Smoluchowski, ...

With the Chapman-Kolmogorov equation and specification of the probability space $(\Omega, \mathcal{F}, P)$ for the random variable $X(t, \cdot)$ with t fixed, one can (re-)obtain the evolution equations for the density.

In the deterministic case, we obtain the Liouville equation.

In the case of stochastic drift-diffusion-jump processes, we have

$$\partial_t f(x,t) + \mathrm{div}(a(x,t)\, f(x,t)) = \frac{\sigma^2}{2}\, \partial^2_{xx} f(x,t) + \lambda \int_{\mathbb{R}} [f(x-y,t) - f(x,t)]\, g(y)\, dy.$$

where P_t is exp distributed in time with $\lambda e^{-\lambda \Delta t}$, λ the rate of jumps, whose amplitude is distributed according to $g = g(x)$.

Fokker-Planck, Kolmogorov, ...

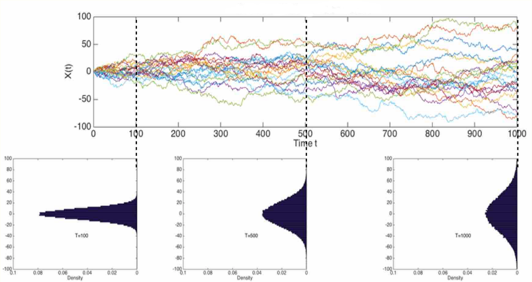
For piecewise-deterministic processes, we have

$$\partial_t f_s(x,t) + \mathrm{div}(a_s(x,t)\, f_s(x,t)) = \sum_{j=1}^{S} Q_{sj}(x)\, f_j(x,t), \qquad s = 1, \ldots, S,$$

where Q_{sj} is given by $Q_{sj} = \begin{cases} \mu_j\, q_{sj} & \text{if } j \neq s, \\ \mu_s\, (q_{ss} - 1), & \end{cases}$ corresponding to a stochastic transition probability matrix $\{q_{ij}\}$ and switching times with exponential PDF of transition events $\psi_s(t) = \mu_s e^{-\mu_s t}$.

In general, we could refer to Lévy processes and corresponding FP systems, which include Brownian motion with drift, the Poisson process, subdiffusion processes, the family of piecewise deterministic Markov processes (PDMP), Switching Diffusion Process (SDP), and Stochastic Hybrid System (SHS).

The Fokker-Planck equation



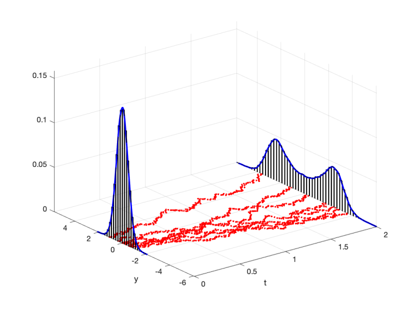
The evolution of the PDF associated to a stochastic drift-diffusion process $X(t) \in \mathbb{R}^d$ is modelled by the Fokker-Planck (FP) equation

$$\partial_t f + \sum_{i=1}^{d} \partial_{x_i} (a_i f) - \sum_{i,j=1}^{d} \partial^2_{x_i x_j} (D_{ij} f) = 0, \qquad f(0) = f_0,$$

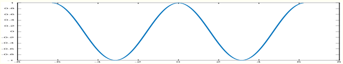
where $D = \frac{1}{2} \sigma\sigma^\top$. Boundary conditions of different type correspond to barriers for the stochastic process $X(t)$ in $\Omega \subset \mathbb{R}^d$.

The FP equation - continuity eq.

The Fokker-Planck equation $\partial_t f + \mathrm{div}\,(a(x,t)\,f) = \frac{1}{2}\sigma^2 \Delta f$ is the continuity equation of the ensemble of stochastic trajectories of a drift-diffusion process.



$dX(t) = \sin(X(t))\,dt + dW(t)$
with $X(0) = X_0 \sim \mathcal{N}(\mu,\, \bar{\sigma}^2)$,
$\mu = 0$ and $\bar{\sigma} = 0.5$.



Notice that $a(x,t) = \sin(x) = -\frac{d}{dx}\cos(x) = -\nabla U(x)$ where $U(x) = \cos(x)$.

Compare with moments' equations in the case $dX(t) = [A(t)\,X(t) + b(t)]\,dt + \sigma\,dW(t)$:

$$\dot{\mu}(t) = A(t)\,\mu(t) + b(t), \qquad \dot{\Sigma}(t) = \Sigma(t)\,A(t)^\top + A(t)\,\Sigma(t) + \sigma\,\sigma^\top$$

Motion of a pedestrian

Consider the motion of an individual, subject to random perturbation, whose (planar) position at time t is denoted with $X(t) \in \mathbb{R}^2$, and its velocity field (drift) is given by u. We have

$$dX(t) = u(X(t), t)\, dt + \sigma\, dW(t), \qquad X(t_0) = X_0,$$

where u depends on x and t.

Assume that reflecting barriers keep the pedestrian in a region $\Omega \subset \mathbb{R}^2$.

The FP equation can be written in flux form $\partial_t f = \nabla \cdot F(f)$, where

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

Reflecting barriers correspond to flux-zero boundary conditions $F \cdot n = 0$ on $\partial\Omega \times (0, T)$, where n is the unit outward normal on $\partial\Omega$.

Ensemble control problem

Consider the control of the motion of a pedestrian by the velocity field u, in order to follow a desired trajectory given by $x_D(t) = (x^1(t), x^2(t))$. We may choose $\theta(x,t) = |x - x_D(t)|^2$ (similarly for $\varphi(x)$) in the ensemble functional:

$$J(f,u) = \int_0^T \int_\Omega \left(\theta(x,t) + \frac{\nu}{2}|u(x,t)|^2\right) f(x,t)\, dx\, dt + \int_\Omega \varphi(x)\, f(x,T)\, dx.$$

In this setting, θ represents an attracting (valley). A bump θ would represent a repulsive (soft obstacle) potential.

We may require that the control belongs to the following admissible set:

$$U_{ad} = \{u \in \mathcal{U},\ u(x,t) \in K_U, \quad \text{a.e. in } Q\}.$$

FP optimality system

$$\partial_t f(x,t) + \nabla \cdot (u(x,t)\, f(x,t)) - \frac{\sigma^2}{2} \Delta f(x,t) = 0$$

$$F \cdot n = 0, \qquad f(x,0) = f_0(x)$$

$$\partial_t p(x,t) + u(x,t) \cdot \nabla p(x,t) + \frac{\nu}{2} |u(x,t)|^2 + \frac{\sigma^2}{2} \Delta p(x,t) + \theta(x,t) = 0$$

$$\partial_n p = 0, \qquad p(x,T) = \varphi(x)$$

$$\langle f\,(\nu\, u + \nabla p)\,, v - u \rangle \geq 0 \qquad v \in U_{ad}.$$

Consider this variational inequality pointwise, and notice that $f > 0$ a.e. in Q. Then, it is sufficient $(\nu\, u + \nabla p)\ (v - u) \geq 0$, which characterizes the solution to $\min_{v \in K_U} [v \cdot \nabla p + \frac{\nu}{2} |v|^2]$ at any point in Q.

By comparison, one recognizes that along the optimal solution, the adjoint equation coincides with the Hamilton-Jacobi-Bellman equation.

The HJB equation

We can consider the expected value (ensemble) functional

$$J_{t_0,x_0}(u) = \mathbb{E}[\int_{t_0}^{T} \left(\theta(X(s),s) + \frac{\nu}{2}|u(X(s),s)|^2\right) ds + \varphi(X(T)) \ \mid \ X(t_0) = x_0 \,],$$

Correspondingly, we have the optimal control $u^* = \operatorname{argmin}_{u \in \mathcal{U}} J_{t_0,x_0}(u)$, and the so-called value function

$$q(x,t) := \min_{u \in \mathcal{U}} J_{t,x}(u) = J_{t,x}(u^*),$$

which satisfies the Hamilton-Jacobi-Bellman equation

$$\partial_t q + H(x,t,\nabla q, \Delta q) = 0, \qquad q(x,T) = \varphi(x),$$

with the Hamilton-Pontryagin function

$$H(x,t,\nabla q,\Delta q) := \min_{v \in K_U}[\frac{\sigma^2}{2}\Delta q(x,t) + v \cdot \nabla q(x,t) + \theta(x,t) + \frac{\nu}{2}|v|^2].$$

Open and closed loop controls

Open-loop: $a(x,t;u) = (v(t) + w(t) \circ x)$, where $u := (v,w)$, $v,w : [0,T] \to \mathbb{R}^n$. In this case, the dependence of the control function on x is given.

$$dX(t) = (v(t) + w(t) \circ X(t))\ dt + \sigma\, dW(t),$$

where $\circ$ denotes the Hadamard product.

Closed-loop: $a(x,t;u) = u(x,t)$

$$dX(t) = u\,(X(t),t)\ dt + \sigma\, dW(t),$$

where $u : \Omega \times [0,T] \to \mathbb{R}^n$. In this case, the dependence of the control function on x has to be determined.
Potentials:

$$\theta\,(x,t) = -\frac{10^{-3}}{2\pi r^2}\,\mathrm{e}^{-\frac{|x-x_D(t)|^2}{2r^2}}, \qquad \phi\,(x) = -\frac{10^{-3}}{2\pi r^2}\,\mathrm{e}^{-\frac{|x-x_D(T)|^2}{2r^2}}.$$

$x_D(t) = (t-1,\ \sin(\pi\, t/2))$

Tracking of trajectory I

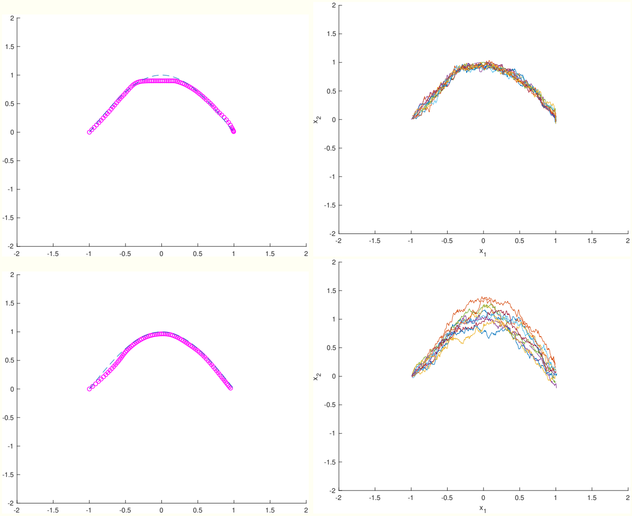


Figure Evolution of $\mathbb{E}[X(t)]$ (circles); the dashed line depicts the desired trajectory. Top: the closed-loop case; bottom: the open-loop case.

Tracking of trajectory II

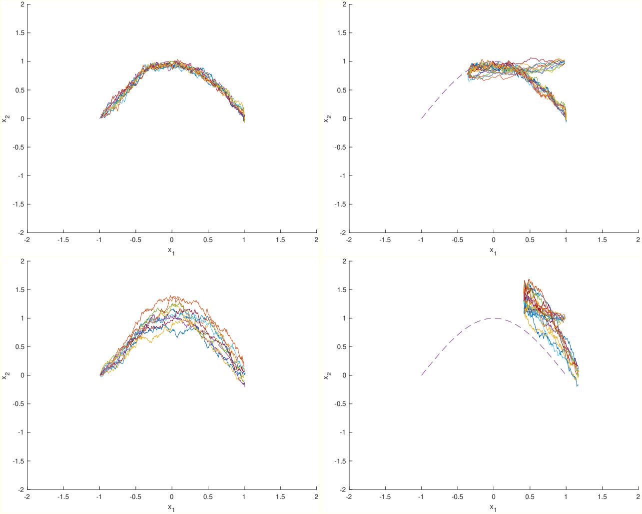


Figure Trajectories of the SDE models with the closed-loop control (top) and the open-loop control (bottom). Left: trajectories starting with $X_0 = x_0 = (-1, 0)$; right: trajectories starting at $X_0 = (1, 1)$.

Results with FP models

- Existence and uniqueness of FP models with different control mechanisms
- Existence of optimal controls
- First- and second-order analysis of FP optimality systems in the Lagrange framework
- Characterization of optimality with the Pontryagin maximum principle
- Analysis of discretization of optimality systems (Chang & Cooper, Splitting)
- Development and analysis of numerical optimization schemes (NCG, proximal methods, sequential quadratic hamiltonian method)
- Applications in biology, finance, microscopy, pedestrian motion, quantum systems

Results with FP models

M. Annunziato, A. Borzì, Optimal control of probability density functions of stochastic processes, Mathematical Modelling and Analysis, 15 (2010), 393-407.

M. Annunziato, A. Borzì, A Fokker-Planck control framework for multidimensional stochastic processes, Journal of Computational and Applied Mathematics, 237 (2013), 487-507.

M. Annunziato, A. Borzì, Fokker-Planck-based control of a two-level open quantum system, Mathematical Models and Methods in Applied Sciences (M3AS), 23 (2013), 2039-2064.

M. Annunziato, A. Borzì, Optimal control of a class of piecewise deterministic processes , European Journal of Applied Mathematics, 25 (2014), 1-25.

M. Mohammadi, A. Borzì, Analysis of the Chang-Cooper Discretization Scheme for a Class of Fokker-Planck Equations, Journal of Numerical Mathematics, 23 (2015), 271-288.

M. Annunziato, A. Borzì, M. Magdziarz, A. Weron, A fractional Fokker-Planck control framework for subdiffusion processes, Optimal Control, Applications and Methods, 37 (2016), 290-304.

S. Roy, M. Annunziato, A. Borzì, A Fokker–Planck feedback control-constrained approach for modelling crowd motion, Journal of Computational and Theoretical Transport, 45 (2016), 442-458.

B. Gaviraghi, M. Annunziato, A. Borzì, Analysis of splitting methods for solving a partial-integro differential Fokker-Planck equation, Applied Mathematics and Computation, 294 (2016), 1-17.

V. Thalhofer, M. Annunziato, A. Borzì, Stochastic modelling and control of antibiotic subtilin production, Journal of Mathematical Biology, 73 (2016), 727-749.

M. Annunziato, A. Borzì, A Fokker-Planck control framework for stochastic systems, EMS Surveys In Mathematical Sciences, 5 (2018), 65-98.

T. Breitenbach, A. Borzì, The Pontryagin maximum principle for solving Fokker-Planck optimal control problems, Computational Optimization and Applied Mathematics, 76, 499–533 (2020).

M. Annunziato, A. Borzì, A Fokker–Planck approach to the reconstruction of a cell membrane potential, SIAM Journal on Scientific Computing, 43 (2021), B623–B649.

J. Körner, A. Borzì, Second–order analysis of Fokker-Planck ensemble optimal control problems ESAIM: COCV, 28 (2022), 77.

FP Nash games and avoidance

Consider two uncoupled pedestrians, labelled with $p = 1, 2$. Each pedestrian's position is denoted with $X^{(p)}(t) \in \Omega \subset \mathbb{R}^2$:

$$\begin{cases} dX^{(p)}(t) = b^{(p)}(X^{(p)}(t), t, u^{(p)}(t))\, dt + \sigma\, dW(t) \\ X^{(p)}(0) = X_0^{(p)}, \end{cases}$$

The drift $b^{(p)}(X^{(p)}(t), t, u^{(p)}(t)) \in \mathbb{R}^2$ has the structure

$$b^{(p)}(X^{(p)}(t), t, u^{(p)}(t)) = v^{(p)}(X^{(p)}(t), t) + u^{(p)}(t)$$

where $v^{(p)}(X^{(p)}(t), t)$ represents the planned velocity (path) of the pedestrian assuming the absence of other pedestrians and of perturbations. The (open-loop) function $u^{(p)}(t)$ models the avoidance action.

A Wiener process is included to model dispersal due to, e.g., collision among individuals (crowd), and the motion is confined in a bounded region $\Omega \subset \mathbb{R}^2$ (a room).

Goals and costs

- Goals : two pedestrians wish to reach different target positions $X_T^{(p)}$, i.e. minimize $\mathbb{E}\left(V_p(X^{(p)}(T) - X_T^{(p)})\right), p = 1, 2$, where V_p is a convex potential.
- Interaction : the two pedestrians would like to avoid a situation where their distance is below an overcrowding limit, thus minimize
$Prob\left\{|X^{(2)}(t) - X^{(1)}(t)| < r\right\} = \mathbb{E}(Q_p) = \mathbb{E}(\mathbf{1}_{\{|X^{(2)}(t)-X^{(1)}(t)|<r\}})$.

$$\begin{aligned}\mathbb{E}(Q_p) &= \int_\Omega\int_\Omega \mathbf{1}_{|y-x|<\rho} f^{(1)}(x,t) f^{(2)}(y,t)\,dxdy \\ &\approx \rho\int_\Omega f^{(1)}(x,t)\, f^{(2)}(x,t)\,dx =: W(f^{(1)}, f^{(2)}),\end{aligned}$$

where $f^{(1)}$ and $f^{(2)}$ represent the probability density functions (PDFs) of the two positions, and $\rho = c\, r^D$ is used as a weighting parameter.
- We include a $H^1(0,T)$ cost of the control to guarantee a bounded continuous slow-varying control.

Avoidance as a FP game

The PDF $f^{(p)}$ associated to the stochastic process $X^{(p)}$ obeys

$$\partial_t f^{(p)}(x,t) - \frac{\sigma^2}{2}\Delta f^{(p)}(x,t) + \nabla\cdot\Big(b^{(p)}(x,t,u^{(p)}(t))\, f^{(p)}(x,t)\Big) = 0,$$

with given $f^{(p)}(x,0) = f_0^{(p)}(x)$. Reflecting barriers for the motion result in flux-zero boundary conditions for the FP equations.

To the pedestrian p is associated the following reduced payoff functional

$$J_p(u^{(1)},u^{(2)}) = \alpha \int_\Omega V_p(x - x_T^{(p)}) f^{(p)}(x,T)dx + \frac{\nu}{2}\|u^{(p)}\|^2_{H^1(0,T;\mathbb{R}^D)} + W(f^{(p)}, f^{(-p)}).$$

No player can improve payoff by unilaterally changing its strategy. The pair $(\bar{u}^1,\bar{u}^2)$ is a Nash equilibrium (NE) point if

$$(\bar{u}^1,\bar{u}^2) \quad = \quad \underset{u^1\in U^1}{\arg\min}\; J_1(u^1,\bar{u}^2) = \underset{u^2\in U^2}{\arg\min}\; J_2(\bar{u}^1,u^2)$$

A weakly coupled Nash game

The avoidance problem has the structure $J_p(u^{(1)}, u^{(2)}) = G_p(u^{(p)}) + W(u^{(1)}, u^{(2)})$ where

$$G_p(u^{(p)}) = \alpha \int_\Omega V_p(x - x_T^{(p)})\, f^{(p)}(x, T) dx + \frac{\nu}{2} \|u^{(p)}\|^2_{H^1(0,T;\mathbb{R}^D)}.$$

Now, define

$$\hat{\mathcal{J}}(u^{(1)}, u^{(2)}) = G_1(u^{(1)}) + G_2(u^{(2)}) + W(u^{(1)}, u^{(2)}).$$

Theorem

Assume that $\hat{\mathcal{J}}$ has a minimum $(\bar{u}^{(1)}, \bar{u}^{(2)})$. Then $(\bar{u}^{(1)}, \bar{u}^{(2)})$ is a Nash equilibrium of the Nash game.

This theorem states that the existence of a minimum of $\hat{\mathcal{J}}$ is a sufficient condition for a Nash equilibrium. This condition is not necessary, in the sense that there can be NE that are not minima of $\hat{\mathcal{J}}$.

An optimal control problem for avoidance

With theorem above, a solution ot our NE avoidance problem is given by the solution of the following optimal control problem: Find $\bar{u} = (\bar{u}^{(1)}, \bar{u}^{(2)})$ such that

$$\hat{\mathcal{J}}(\bar{u}) \leq \hat{\mathcal{J}}(u) \qquad \text{for all } u = (u^{(1)}, u^{(2)}) \in U^{(1)} \times U^{(2)}.$$

This is the reduced formulation of the following FP control problem.

$$\begin{gathered}
\min \hat{\mathcal{J}}(f^{(1)}, f^{(2)}, u^{(1)}, u^{(2)}) := G_1(f^{(1)}, u^{(1)}) + G_2(f^{(2)}, u^{(2)}) + W(f^{(1)}, f^{(2)}) \\
\partial_t f^{(1)}(x,t) - \frac{\sigma^2}{2} \Delta f^{(1)}(x,t) + \nabla \cdot \left(b^{(1)}(x,t,u^{(1)}(t)) \, f^{(1)}(x,t) \right) = 0, \\
f^{(1)}(x,0) = f_0^{(1)}(x), \\
\partial_t f^{(2)}(x,t) - \frac{\sigma^2}{2} \Delta f^{(2)}(x,t) + \nabla \cdot \left(b^{(2)}(x,t,u^{(2)}(t)) \, f^{(2)}(x,t) \right) = 0, \\
f^{(2)}(x,0) = f_0^{(2)}(x) \quad .
\end{gathered}$$

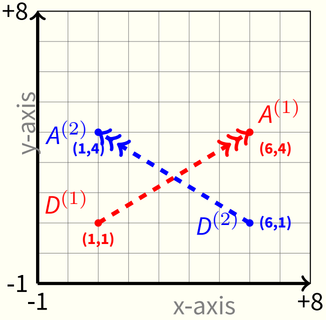
The strategies $u^{(p)} = (u_1^{(p)}, u_2^{(p)})$, $p = 1, 2$, are sought in the following admissible set

$$U^{(p)} = \{ u \in H_0^1(0,T;\mathbb{R}^2) | \, u_a \leq u_i(t) \leq u_b, \; i = 1,2 \text{ a.e. in } (0,T) \}$$

where $u_a, u_b \in \mathbb{R}, \; u_a < u_b$.

Experiment-2: Turnwald 1C-A3 – Setting

Two pedestrian are asked to reach specific targets and avoid each other [1].



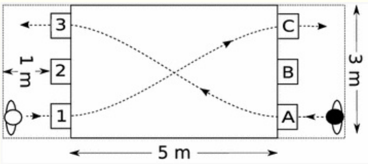


Figure Settings for the Turnwald game described in test-case Turnwald 1C-A3: (left) the computational setting ; (right) the human experiment setting.

[1] Turnwald, A., Althoff, D., Wollherr, D., Buss, M. (2016). *Understanding human avoidance behavior: interaction-aware decision making based on game theory.* International Journal of Social Robotics, 8(2), 331-351.

Experiment: Turnwald 1C-A3 – Computation

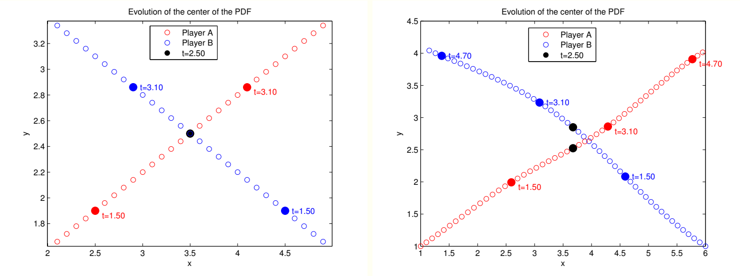


Figure Zoomed-in plots of trajectories with $\rho = 0.01, 200$ respectively. (a) motion for $\rho = 0.01$ and the two pedestrian meet at $t = 2.5$; (b) trajectories with $\rho = 200$ and avoidance.

S. Roy, A. Borzì, A. Habbal,

Pedestrian motion modelled by Fokker - Planck Nash games, Royal Society open science, 4: 170648, 2017.

Thanks a lot for your kind invitation and for your interest in my work

Additional technical details

Existence and uniqueness

Consider the weighted function spaces H^m_k: $f \in H^m_k$ if f and all its derivatives up to order m belong to $\left(L^2(\mathbb{R}^d), (1+|x|^k)\, dx\right)$.

$$\partial_t f(x,t) + \operatorname{div}\big(a(x,t;u)\, f(x,t)\big) = g(x,t), \qquad f(x,0) = f_0(x),$$

where

$$\begin{cases} g \in L^1\big([0,T]; H^m_k(\mathbb{R}^d)\big) & \text{and} \quad f_0 \in H^m_k(\mathbb{R}^d) \\ a \in L^1\big([0,T]; C^{m+1}(\mathbb{R}^d)\big) & \text{with} \quad \nabla a \in L^1\big([0,T]; C^m_b(\mathbb{R}^d)\big) \end{cases}$$

Theorem

Let $T > 0$ and $m \in \mathbb{N}$ fixed, and let a, f_0 and g satisfy our hypotheses. Then there exists a unique solution $f \in C\big([0,T]; H^m_k(\mathbb{R}^d)\big)$ of the Liouville initial-value problem. Moreover, there exists a constant $C > 0$, independent of f_0, a, g, f and T, such that the following estimate holds true for any $t \in [0,T]$:

$$\|f(t)\|_{H^m_k} \le C \left(\|f_0\|_{H^m_k} + \int_0^t \|g(\tau)\|_{H^m_k}\, d\tau \right) \exp\left(C \int_0^t \|\nabla a(\tau)\|_{C^m_b}\, d\tau \right).$$

The control-to-state map G

This map associates to a given control u a state f with a fixed initial condition f_0. Let $m \geq 2$, $k \geq 2$ and the data be given as above.

Theorem

Let δu be an arbitrary admissible variation of u. The control-to-state map G is Gâteaux differentiable at u and the Gâteaux derivative $\delta_{\delta u} G$ satisfies the Liouville problem

$$\partial_t \delta_{\delta u} G + \mathrm{div}\big(a(t,x;u)\, \delta_{\delta u} G\big) \;=\; -\; \mathrm{div}\big(\overline{a}(t,x;\delta u)\, G(u)\big)\,, \qquad \delta_{\delta u} G_{|t=0} \;=\; 0\,, \tag{1}$$

where we have defined $\overline{a}(t,x;\delta u) \;:=\; \delta u_1 + x \circ \delta u_2$.

Theorem

The map G is Fréchet differentiable from $\mathrm{int}\, U_{ad}$ into $L^\infty_T(L^2)$, and its Fréchet differential at any point $u \in \mathrm{int}\, U_{ad}$ is given by $DG(u)$. The differential $DG(u)[\delta u]$ is the unique solution to equation (1). There exists a constant $C > 0$ such that

$$\Big\| G(u+\delta u) \;-\; G(u) \;-\; DG(u)[\delta u] \Big\|_{L^\infty_T(L^2)} \;\leq\; C\; \|\delta u\|^2_{L^\infty_T}\;.$$

Existence of optimal controls

We introduce the reduced cost functional

$$J_r(u) \ := \ J\big(G(u), u\big) \ ,$$

Correspondingly, we have the reduced optimal control problem

$$\min_{u \in U_{ad}} J_r(u) \, .$$

Theorem

The ensemble optimal control problem with $\gamma \geq 0, \delta \geq 0, \ \nu > 0$ admits at least one solution $u^* \ \in \ U_{ad}$. The corresponding state $f^* \ := \ G(u^*)$ belongs to the space $C\big([0,T]; H^m_k(\mathbb{R}^d)\big)$, $(m,k) \in \mathbb{N}^2$.

Notice that J is Fréchet (sub-)differentiable over $C([0,T];L^2) \times int\, U_{ad}$, since it is linear in f and the control costs are given by (sub-)differentiable norms.

Optimality system

Consider the case $\nu = 0$, and take

$$U_{ad} = \{u \in L^2([0,T];\mathbb{R}^d) \times L^2([0,T];\mathbb{R}^d) \,|\, u_a \leq u(t) \leq u_b,\ t \in [0,T]\}$$

We introduce the Lagrange multiplier q, and $\widehat{\lambda} \in \partial\left(||u||_{L^1}\right)$.

The optimality system reads

$$\begin{aligned}
&\partial_t f \,+\, \mathrm{div}\left(a(x,t;u)\,f\right) \,=\, 0, \qquad f_{|t=0} \,=\, f_0 \\
&-\partial_t q \,-\, a(x,t;u)\cdot\nabla q \,=\, -\theta, \qquad q_{|t=T} \,=\, -\varphi \\
&\left(\gamma\, u_j^r + \widehat{\lambda}_j^r + \int_{\mathbb{R}^d} \mathrm{div}\left(\frac{\partial a}{\partial u_j^r}\, f\right) q\, dx \,,\, v_j^r - u_j^r\right)_{L^2(0,T)} \geq 0 \qquad v \in U_{ad}\,, \\
&j = 1,2\,,\ r = 1,\dots,d
\end{aligned}$$

Liouville adjoint equation

Consider the adjoint Liouville intial-value problem

$$\begin{cases} -\partial_t q - a(x,t;u)\cdot\nabla q = -\theta, & \text{in } \mathbb{R}^d \times [0,T] \\ q(x,T) = -\varphi(x), & \text{in } \mathbb{R}^d \end{cases}$$

where

$$\begin{cases} \theta \in L^1\big([0,T]; H^m_k(\mathbb{R}^d)\big) & \text{and} \quad \varphi \in H^m_k(\mathbb{R}^d) \\ a \in L^1\big([0,T]; C^{m+1}(\mathbb{R}^d)\big) & \text{with} \quad \nabla a \in L^1\big([0,T]; C^m_b(\mathbb{R}^d)\big) \end{cases}$$

Theorem

Let $T > 0$ and $m \in \mathbb{N}$ fixed, and let a, φ and θ satisfy the given assumptions. Then there exists a unique solution $q \in C\big([0,T]; H^m_k(\mathbb{R}^d)\big)$ of the initial-value problem. Moreover, there exists a constant $C > 0$, independent of φ, a, θ, q and T, such that the following estimate holds true for any $t \in [0,T]$:

$$\|q(t)\|_{H^m_k} \leq C\left(\|\varphi\|_{H^m_k} + \int_0^t \|\theta(\tau)\|_{H^m_k}\, d\tau\right) \exp\left(C\int_0^t \|\nabla a(\tau)\|_{C^m_b}\, d\tau\right).$$

Numerical approximation

Consider a computational domain $[-B, B]^2 =: \Omega$, $B \gg 0$, and a cell-centred finite-volume setting with equally spaced, non-overlapping cells; cell size $h > 0$.
Discretization in time: intervals with equal length $\Delta t > 0$
Assume that f and q have compact support of f at all times $t \in [0, T]$, $T > 0$; set homogeneous Dirichlet boundary conditions.

Liouville equation

- *strong stability conserving Runge-Kutta method of second order* in time and *Kurganov-Tadmor* scheme in space (**SSPRK2-KT**)

Adjoint equation

- additionally use second-order Strang splitting (**KTS**)

Kurganov-Tadmor approximation

Uses generalized MUSCL flux with approximation rule for f at cell-edges and has total-variation diminishing (TVD) property

Lemma The semi-discrete KT scheme is at least second-order accurate in space for smooth f, except possibly at the points of extrema of f.

Lemma The SSPRK2-KT scheme is positivity preserving and conservative, in the sense that $\sum_{i,j}^{N_x} f_{i,j}^k = \sum_{i,j}^{N_x} f_{i,j}^0, \qquad k = 1, \ldots, N_t.$

Lemma The solution $f_{i,j}^k$ with a Lipschitz continous right-hand side g obtained with the SSPRK2-KT-scheme is discrete L^1 stable in the sense that

$$\left\| f_{\cdot,\cdot}^{k+1} \right\|_{1,h} = \| f_{\cdot,\cdot}^0 \|_{1,h} + \Delta t \sum_{m=0}^{k} \| g_{\cdot,\cdot}^m \|_{1,h}, \qquad k = 0, \ldots, N_t - 1$$

provided that the CFL condition holds ($\lambda := \Delta t / h$),

$$\lambda \left(\|a_0\|_{L_T^\infty(L^\infty(\Omega))} + \left(a_1 + a_2 B\right) \max\left\{ |u^a|, |u^b| \right\} \right) \leq \frac{1}{4}.$$

Strang splitting scheme

Rewrite the adjoint equation

$$-\partial_t q(x,t) - a(x,t;u)\cdot\nabla q(x,t) = -\theta(x,t)$$

as

$$-\partial_t q(x,t) - \nabla\cdot\left(a(x,t;u(t))\,q(x,t)\right) = -\left(\nabla\cdot a(x,t;u)\right)q(x,t) - \theta(x,t).$$

We solve the Liouville part

$$-\partial_t q(x,t) - \nabla\cdot\left(a(x,t;u(t))\,q(x,t)\right) = 0$$

using the SSPRK2-KT scheme and the source and reaction term part

$$-\partial_t q(x,t) = -\left(\nabla\cdot a(x,t;u)\right)q(x,t) - \theta(x,t).$$

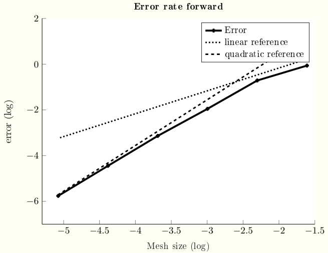
using Euler's method.

Accuracy of the SSPRK2-KT scheme

Theorem

Let $f \in C^3$ be the exact solution of the Liouville equation, with countably many extrema, and let $\left\|f^0_{\cdot,\cdot} - f_0(\cdot,\cdot)\right\|_{1,h} = \mathcal{O}(h^2)$. Under the CFL condition, the solution $f^k_{i,j}$ obtained with the SSPRK2-KT scheme is second-order accurate in the discrete L^1-norm as follows

$$\left\|f^{N_t}_{\cdot,\cdot} - f(\cdot,\cdot,T)\right\|_{1,h} \leq D(T,\Omega,\lambda)\, h^2.$$

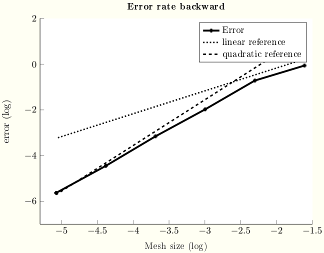


Accuracy of the KTS scheme

Theorem

Let $q \in C^3$ be the exact solution of the adjoint equation, with countably many extrema, and let $\left\| q^0_{\cdot,\cdot} + \varphi(\cdot,\cdot) \right\|_{1,h} = \mathcal{O}(h^2)$. Under the CFL condition, the solution $q^k_{i,j}$ obtained with the KTS scheme is second-order accurate in the discrete L^1-norm as follows

$$\left\| q^{N_t}_{\cdot,\cdot} - q(\cdot,\cdot,T) \right\|_{1,h} \leq E(T,\Omega,\lambda)\, h^2.$$



Projected semi-smooth Newton

Consider the generalized Jacobian at u in direction δu. It is defined as follows

$$\mathcal{J}_r(u)\,\delta u \;:=\; \delta u \,+\, \Phi,$$

where Φ is the solution to

$$\Big(-\nu\frac{d^2}{dt^2} + \gamma\Big)\Phi \;=\; -\int_{\mathbb{R}^d} \frac{\partial a}{\partial u}\hat{f}\cdot\nabla q\,dx \,+\, \int_{\mathbb{R}^d} \mathrm{div}\Big(\frac{\partial a}{\partial u}f\Big)\hat{q}\,dx,$$

$$\Phi(0) \;=\; \Phi(T) \;=\; 0\,,$$

where $\hat{f}$, $\hat{q}$ satisfy the linearised Liouville and adjoint equations.

Projection in H^1 of new iterate on U_{ad} is determined as follows

$$P_{U_{ad}}(u) = \mathrm{argmin}_{\tilde{u}\in U_{ad}}\;\frac{\gamma}{2}\,\|\tilde{u}-u\|_{L^2}^2 \,+\, \frac{\nu}{2}\left\|\frac{\mathrm{d}}{\mathrm{dt}}(\tilde{u}-u)\right\|_{L^2}^2.$$

Optimization algorithm

Algorithm 1: Projected semi-smooth Newton method

Require: $u_0, f_0, \theta, \varphi, tol > 0, n_{\max}$

Ensure: Optimal control u^* and optimal state $f^* = G(u^*)$

Set $l = 0$

while $\|u_{l+1} - u_l\| > tol$ **and** $l < n_{\max}$ **do**

    Solve Liouville equation $\partial_t f + \operatorname{div}\big(a(x,t;u)\,f\big) = 0\,, \quad f_{|t=0} = f_0$

    Solve adjoint equation $-\partial_t q - a(x,t;u)\cdot\nabla q = -\theta, \quad q_{|t=T} = -\varphi$

    Solve linearized Liouville eq. $\partial_t \hat f + \operatorname{div}(a\hat f) = -\operatorname{div}(\hat a f), \quad \hat f(x)\big|_{t=0} = 0$

    Solve linearized adjoint eq. $-\partial_t \hat q - a\cdot\nabla\hat q = \hat a\cdot\nabla q, \quad \hat q\big|_{t=T} = 0$

    Assemble Jacobian $\mathcal{J}_r(u_l)$ and the gradient $\mathcal{F}_r$

    Solve $\mathcal{J}_r(u_l)\,d_l = -\mathcal{F}_r(u_l)$ (with e.g. GMRES)

    Find stepsize σ_l (with e.g. Armijo line-search and H^1 projection)

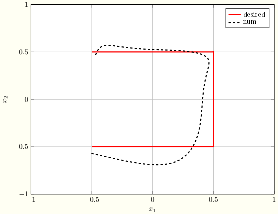
    $u_{l+1} = u_l + \sigma_l d_l$

    $l = l + 1$

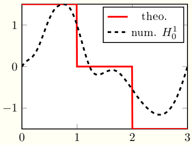
end while

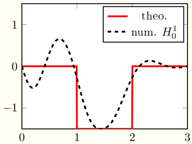
return $(f(u_l), u_l)$

Num. exp. - unimodal Gauss



mean value $\langle x \rangle$ of probability distribution function (- - -) and desired mean value x_D (——)



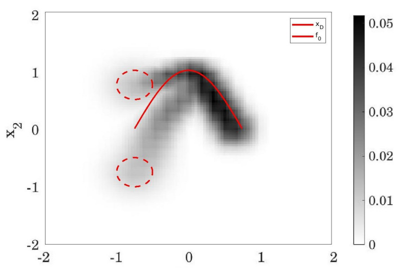


$$\langle x \rangle(t) = \int x\, f(x,t)\, dx,$$

$$f_0(x) = \frac{C_0}{2\pi\sigma^2}\left(-\frac{1}{2}\left(\frac{|x-x_0|}{\sigma}\right)^2\right),$$

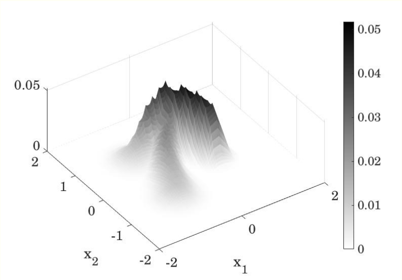
$$C_0 = 0.1, \sigma = 0.25, x_0 = (-0.5, 0.5)$$

Num. exp. - bimodal Gauss

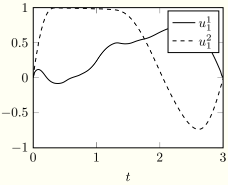


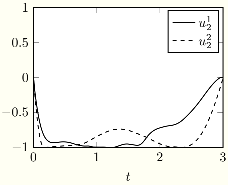
$x_D(t) = \left(-\frac{3}{4} + \frac{3t}{2T}, \sin\left(\frac{\pi t}{T}\right)\right)$

(a) top view



(b) side view





Kinetic models with collision

Keilson-Storer model

For modeling collisions, we focus on the Keilson-Storer (KS) model.

$$C[f] := \int f(w,t)\,A(w,v)\,dw \;-\; f(v,t)\int A(v,w)\,dw,$$

- linear kinetic model of colloidals suspended in a bath in thermal equilibrium
- application in, e.g., the estimation of transport coefficients, laser spectroscopy, molecular dynamics simulations, reorientation of molecules in liquid water, and quantum transport
- microscopic derivation of the KS master equation is available

The adjoint KS model has not a kinetic structure but this structure is required in order to apply Monte Carlo (MC) methods that are necessary in the regime of dilute gases, where the continuum assumption is no longer valid.

Keilson-Storer collision term

It has a gain – loss structure

$$C[f] := \int f(w,t)\,A(w,v)\,dw \;-\; f(v,t)\int A(v,w)\,dw,$$

with collision kernel $A(v,w) := A_0 e^{(-\beta|w-\gamma v|^2)}$ and $\gamma \in [-1,1]$, $A_0, \beta > 0$.
For the post-collision velocity holds $w \sim \mathcal{N}(\gamma v, (2\beta)^{-1})$.

Properties:

- $\gamma \lessapprox 1$: weak collisions, Brownian motion
- $\gamma \approx 0$: strong collision, Bhatnagar-Gross-Krook (BGK) operator
- collision frequency $\frac{1}{\tau} = A_0\sqrt{\pi/\beta}$
- detailed balance: $A(w,v)\,f^{eq}(w) = A(v,w)\,f^{eq}(v)$.
- equilibrium solution $f^{eq}(v)$ is the Maxwellian distribution
- A_0 and β related to the background density and temperature

KS optimal control problem

We consider the following Liouville KS problem

$$
\begin{aligned}
\partial_t f + v \cdot \nabla_x f + u(x) \cdot \nabla_v f &= C[f] \\
f|_{t=0} &= f_0, \\
f(x,v,t)|_{\partial\Omega \times \mathbb{R}^d_< \times (0,T]} &= f(x, v - 2n(n \cdot v), t)
\end{aligned}
$$

with specular reflection space boundaries; $\mathbb{R}^d_< := \{v \in \mathbb{R}^d \,|\, v \cdot n(x) < 0\}$. The control field u is sought in $H^1_0(\Omega)$.

Suppose a desired trajectory in the phase space $z_D(t)$, $t \in [0, T]$, and desired final configuration z_T. We choose the potentials θ and φ.

Our problem is to find $u \in H^1_0(\Omega)$ such that the following ensemble cost functional is minimised

$$
J(f,u) := \int_0^T \int_Q \theta(x,v,t) f(x,v,t)\, dx\, dv\, dt + \int_Q \varphi(x,v) f(x,v,T)\, dx\, dv + \frac{\nu}{2} \|u\|^2_{H^1}
$$

where $Q = \Omega \times \mathbb{R}^d$.

The KS adjoint equation

The Liouville KS adjoint equation is given by

$$-\partial_t q(x,v,t) - v\cdot\nabla_x q(x,v,t) - u(x)\cdot\nabla_v q(x,v,t) = \tilde{C}[q](x,v,t) - \theta(x,v,t)$$

with

$$\tilde{C}[q](x,v,t) = \int A(v,w)\, q(x,w,t)\, dw - q(x,v,t) \int A(v,w)\, dw.$$

The operator $\tilde{C}[q]$ has not a gain-loss structure, but such a structure can be partially recovered defining

- $C^*[q](v,t) = \int A^*(w,v)\, q(w,t)\, dw - q(v,t) \int A^*(v,w)\, dw,$
- $A^*(w,v) = \frac{1}{\gamma} A(v,w)$
- $\int \big(A(w,v) - A(v,w)\big)\, dw = \frac{1-\gamma}{\tau_q} =: C_0^*.$
- ‘adjoint’ mean free time $\tau_q = \gamma\, \tau$

Reformulation of the KS adjoint

We choose θ and φ as follows

$$\theta(z,t) := -\frac{C_\theta}{2\pi\sigma_\theta^2}\exp\left(-\frac{|z-z_D(t)|^2}{2\sigma_\theta^2}\right), \qquad \sigma_\theta > 0.$$

and

$$\varphi(z) := -\frac{C_\varphi}{2\pi\sigma_\varphi^2}\exp\left(-\frac{|z-z_T)|^2}{2\sigma_\varphi^2}\right), \qquad \sigma_\varphi > 0.$$

With this choice, θ and φ play the role of sources (or sinks) of particles.

The adjoint KS model is given by

$$-\partial_t q - v\cdot\nabla_x q - u(x)\cdot\nabla_v q = C^*[q] + C_0^* q - \theta, \qquad q_{|t=T} = -\varphi.$$

The forward and adjoint problems can be written introducing the free-streaming operators

$$L_u = v\cdot\nabla_x + u\cdot\nabla_v, \qquad \text{and} \qquad L_u^* = -L_u.$$

The reduced gradient in H^1

The L^2 gradient of the reduced cost functional is given by

$$\nabla J_r(u)\big|_{L^2}(x) = -\nu\,\Delta u(x) + \nu\,u(x) + \int_0^T \int_{\mathbb{R}^d} q(x,v,t)\,\nabla_v f(x,v,t)\,dv\,dt.$$

However, the update for the control needs the H^1 reduced gradient. Considering the Riesz representative of $J_r'(u)$ on different Hilbert spaces:

$$\left(\nabla J_r(u)\big|_{L^2}, \delta u\right)_{L^2} = \left(\nabla J_r(u)\big|_{H^1}, \delta u\right)_{H^1}, \qquad \delta u \in H^1.$$

Thus, the H^1 gradient is obtained as the solution to the following boundary-value problem

$$-\Delta\psi + \psi = \nabla J_r(u)\big|_{L^2}, \qquad \psi\big|_{\partial\Omega} = 0.$$

That is, $\nabla J_r(u)\big|_{H^1} = \psi$.

KS optimality system

$$
\begin{aligned}
&\partial_t f(x,v,t) + L_u f(x,v,t) = C[f](x,v,t),\\
&f(x,v,0) = f_0(x,v)\\
&f(x,v,t)|_{\partial\Omega\times\mathbb{R}^d_<} = f(x, v - 2n(n\cdot v), t)
\end{aligned}
$$

$$
\begin{aligned}
&-\partial_t q(x,v,t) + L_u^* q(x,v,t) = C^*[q](x,v,t) + C_0^*\, q(x,v,t) - \theta(x,v,t),\\
&q(x,v,T) = -\varphi(x,v)\\
&q(x,v,t)|_{\partial\Omega\times\mathbb{R}^d_>} = q(x, v - 2n(n\cdot v), t)
\end{aligned}
$$

$$
\begin{aligned}
&-\nu\,\Delta u(x) + \nu\, u(x) + \int_0^T \int_{\mathbb{R}^d} q(x,v,t)\,\nabla_v f(x,v,t)\,dv\,dt = 0\\
&u_{|\partial\Omega} = 0.
\end{aligned}
$$

Simulation scales & Numerics

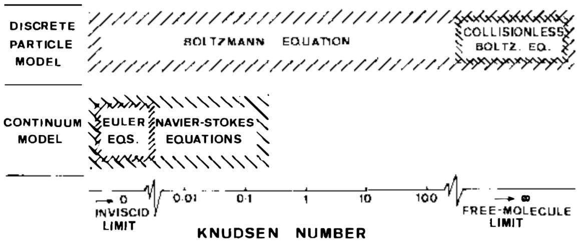


Figure 1 The Knudsen-number limits on the conventional mathematical models of neutral gas flows.

In the long term, we are concerned with methods for calibration, control, and optimization of kinetic models, especially in the mesoscopic regime where probabilistic aspects of the evolution of particles play an essential role.

This is the case in the simulation of rarefied gases with high Knudsen number (the ratio of the mean-free path and the characteristic length of the problem).

The mesoscopic setting accommodates the case where the coefficients of the model are prescribed probabilistically by some distribution functions.

Although kinetic models are partial-integro differential equations, methods developed in a deterministic context cannot always be applied and computation by Monte Carlo methods could be required.

Monte Carlo methods

Split the solution operator using the kinetic description of gases that consists of a deterministic free flight (Newton's law of motion) between two collisions

$$\dot{x} = v, \qquad \dot{v} = u(x),$$

and probabilistic collision (Keilson-Storer kernel) according to a certain collision frequency τ

$$\delta t = -\tau \log(r), \qquad w = \gamma v + \frac{p}{2\beta}.$$

where r uniform random number, p normal random number.

Particles are list of pointers F^k, Q^k storing position and velocity at time step k. A computational mesh is required only to compute the integral for the evaluation of the gradient.

KS adjoint and Monte Carlo

$$\partial_s q - v \cdot \nabla_x q \; - \; u(x) \cdot \nabla_v q \; = \; C^*[q] \; + \; C_0^* q - \theta, \qquad q_{|s=0} = -\varphi.$$

The KS adjoint model consists out of a free-streaming part, a collision part, a reaction term and a source term. We have collision frequency $(\gamma\tau)^{-1}$ and post-collision velocities $w^* \sim \mathcal{N}\left(v/\gamma, (2\beta\gamma^2)^{-1}\right)$.

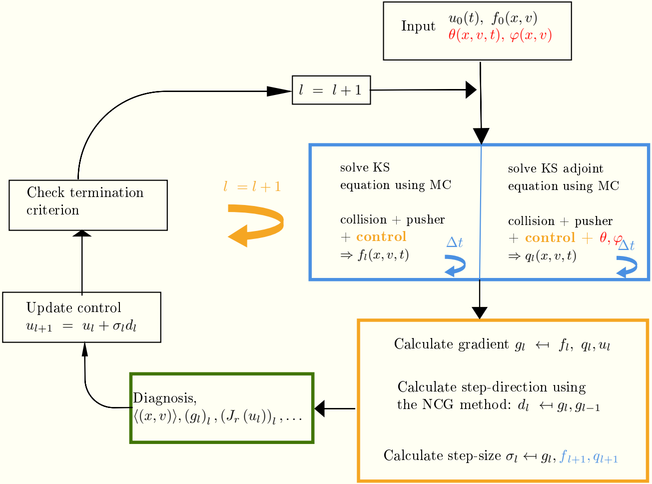
For reaction term $C_0^* q$:

- For all particles p in Q^k: Generate $r_* := \lfloor \Delta t \, C_0^* \rfloor$ particles with the velocity $Q^k[p].v$ and position $Q^k[p].x$.
- Add these particles to the existing ones in Q^k.

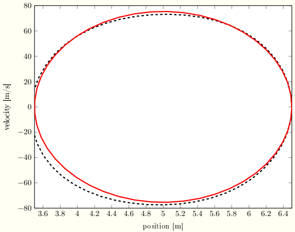
For the source term $-\theta$:

- Generate N_{frac} new particles with phase space components having the normal distribution with mean $z_D(t^k)$ and variance σ_θ^2: $v \sim \mathcal{N}\left(z_D(t^k), \sigma_\theta^2\right)$.
- Add these particles to the existing ones in Q^k

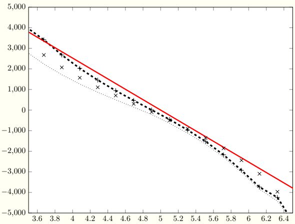
Monte Carlo algorithm



Num. exp. - Harmonic oscillator



Plot of mean values $(\langle x \rangle, \langle v \rangle)$ (- - -) and desired z_D (——).



Optimal control (- - -) and elastic force $F(x)$ of the harmonic oscillator(——). Comparison of results with number of particles N_f and with two times N_f $(\cdots\cdots)$, four times N_f $(\times\times\times)$, eight times N_f $(+++)$.

$$f_0(x, v) = \frac{1}{2\pi \cdot 0.15 \cdot 5.0} \exp\left(-\frac{1}{2}\left[\left(\frac{x-5.0}{0.15}\right)^2 + \left(\frac{v-0.0}{5.0}\right)^2\right]\right),$$
$$z_D(t) = (1.5\cos(\omega t) + 5.0, -1.5\sin(\omega t))^T, F(x) = -\omega^2(x-5)$$

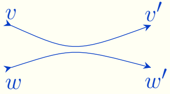
Boltzmann equation

We use the Boltzmann collision kernel (binary collision)

$$C_B[f,f](x,v,t) = \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} |v-w|c(|v-w|,\vartheta)(f_{v'}f_{w'} - f_v f_w)\,d\vartheta\,dw,$$

with

$$v' = \frac{v+w}{2} + \frac{|v-w|}{2}\vartheta, \qquad w' = \frac{v+w}{2} - \frac{|v-w|}{2}\vartheta$$



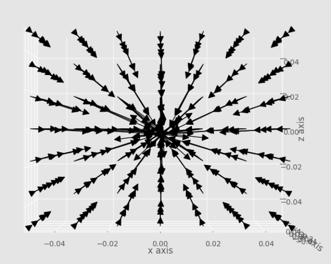
Choose the external force $u = (\mathcal{E},\mathcal{B})$

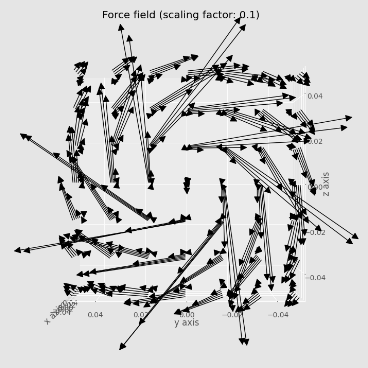
$$\mathcal{F}(x,v;u) = \Big(a_0(x) + a_1\,\mathcal{E}(x) + a_2\, v \times \mathcal{B}(x)\Big).$$

Consider the Boltzmann equation in full phase space

$$\partial_t f(x,v,t) + v\cdot\nabla_x f(x,v,t) + \mathcal{F}(x,v;u)\cdot\nabla_v f(x,v,t) = C_B[f,f](x,v,t)$$

Experiments





Control for charged particles confinement; $\mathcal{F}(x) = \mathcal{E}(x)$

Control for beam constriction; $\mathcal{F}(x, v) = v \times \mathcal{B}(x)$

Work in collaboration with SPARC Industries

Connections between models

- The Boltzmann equation for massive particles M immersed in a viscous fluid and subject to collisions with much smaller particles of mass $m << M$ is approximately linear and describes Brownian motion[2]
- A linear Boltzmann equation becomes the drift-diffusion (Fokker-Planck) equation in the small-diffusion limit[3]
- The Liouville equation results by vanishing viscosity of the Fokker-Planck equation[4]
- Collisions in the classical Boltzmann equation can be framed in the PDP framework[5]
- In phase space and with underlying hamiltonian dynamics, the Liouville equation is the Vlasov equation

---

[2] D. Montgomery, Brownian motion from Boltzmann's equation, Phys. Fluids, 14 (1971), 2088–2090.

[3] J. Keilson and J. E. Storer, On Brownian motion, Boltzmann's equation, and the Fokker-Planck equation, Quart. Appl. Math., 10 (1952), 243–253.

[4] E. Godlewski and P.-A. Raviart, Numerical Approximation of Hyperbolic Systems of Conservation Laws, Springer, 1996.

[5] S. Rjasanow and W. Wagner, Stochastic Numerics for the Boltzmann Equation, Springer, 2005.

Remarks on ensemble control

1. Open-loop and closed-loop strategies can be accommodated both in this framework.
   - Open-loop: the functional dependence on x is given while dependence on t (i.e. $u = u(t)$) must be determined;
   - Closed-loop: the functional dependence on both x and t must be determined (i.e. $u(x,t)$).
2. Also in the open-loop case a robust control can be obtained.
3. Appropriate choice of the functional dependence on x may allow to obtain a good approximation of the closed-loop control function while being easier to implement.
4. This framework can be applied in general in the context of continuity equations and density functions representations.
5. It is a rich source of challenging mathematical problems.