Generalized Gradient Flows for Stochastic Prediction, Estimation, Learning and Control

Abhishek Halder

Department of Applied Mathematics
University of California, Santa Cruz
Santa Cruz, CA 95064

Joint work with students and collaborators








Aerospace Engineering Department Seminar, Iowa State University
November 17, 2022

Topic of this talk

Control theory and algorithms for measures/distributions and densities

Intuition

measure a.k.a. distribution = mass

mass = density $\times$ volume

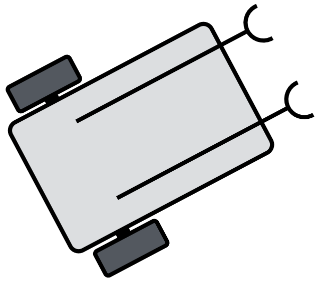
conservation of mass:

$$\int \mathrm{d}(\text{mass}) = \int \text{density} \times \mathrm{d}(\text{volume})$$

$$= \text{constant, say } 1$$

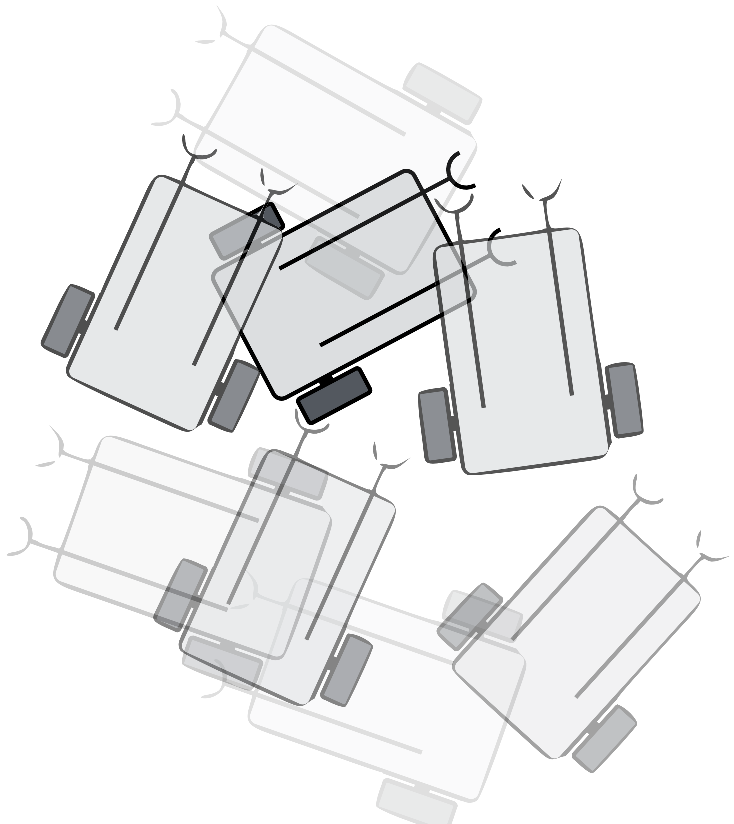
But what do we mean by density?

Probability Density Fn.



$$\boldsymbol{x}(t) = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

Probability Density Fn.

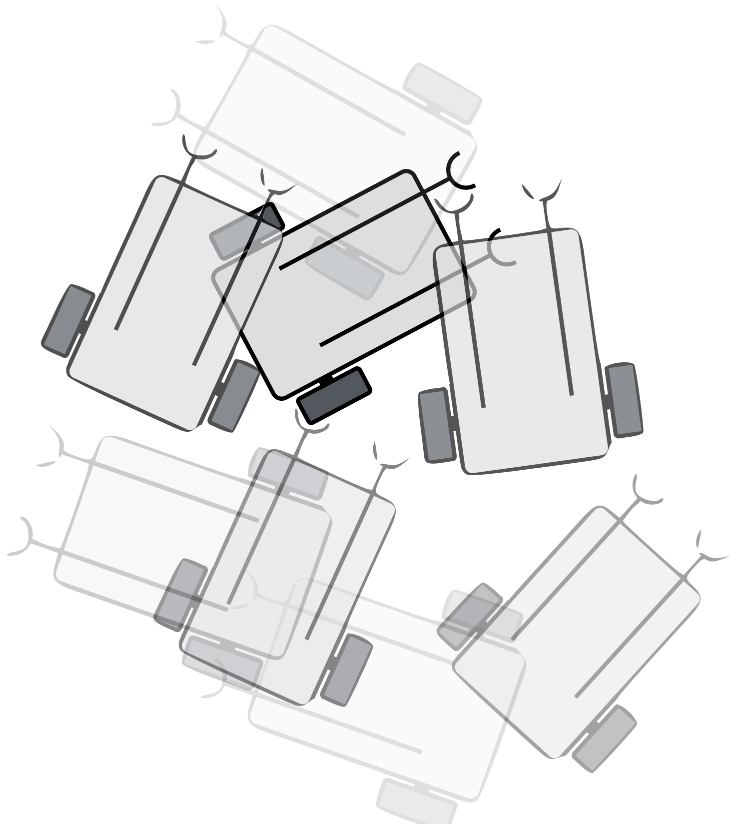


$$\boldsymbol{x}(t) = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

$$\rho\left(\boldsymbol{x}, t\right) : \mathcal{X} \times [0, \infty) \mapsto \mathbb{R}_{\geq 0}$$

$$\int_{\mathcal{X}} \mathrm{d}\mu = \int_{\mathcal{X}} \rho \, \mathrm{d}\boldsymbol{x} = 1 \quad \text{for all } t \in [0, \infty)$$

Probability Density Fn.



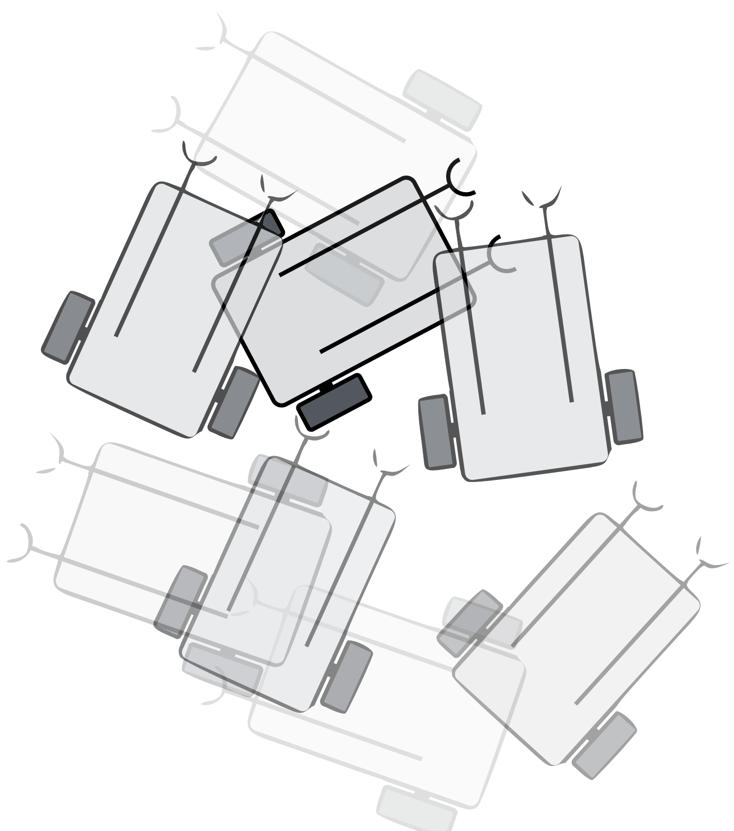
$$\boldsymbol{x}(t) = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

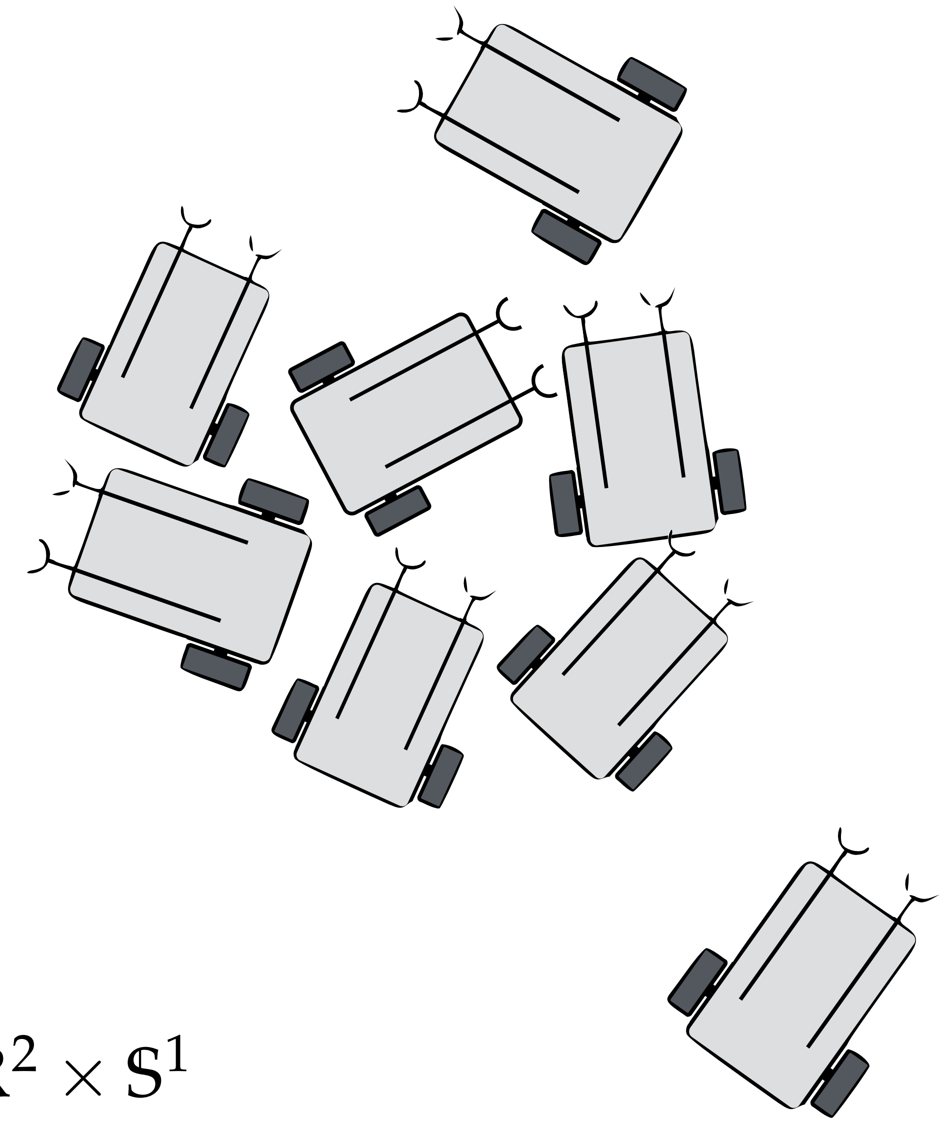
$$\rho\left(\boldsymbol{x}, t\right) : \mathcal{X} \times [0, \infty) \mapsto \mathbb{R}_{\geq 0}$$

probability measure → μ; probability density function → ρ

$$\int_{\mathcal{X}} \mathrm{d}\mu = \int_{\mathcal{X}} \rho \, \mathrm{d}\boldsymbol{x} = 1 \quad \text{for all } t \in [0, \infty)$$

Probability Density Fn. Population Density Fn.





$$\boldsymbol{x}(t) = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

$$\rho\left(\boldsymbol{x}, t\right) : \mathcal{X} \times [0, \infty) \mapsto \mathbb{R}_{\geq 0}$$

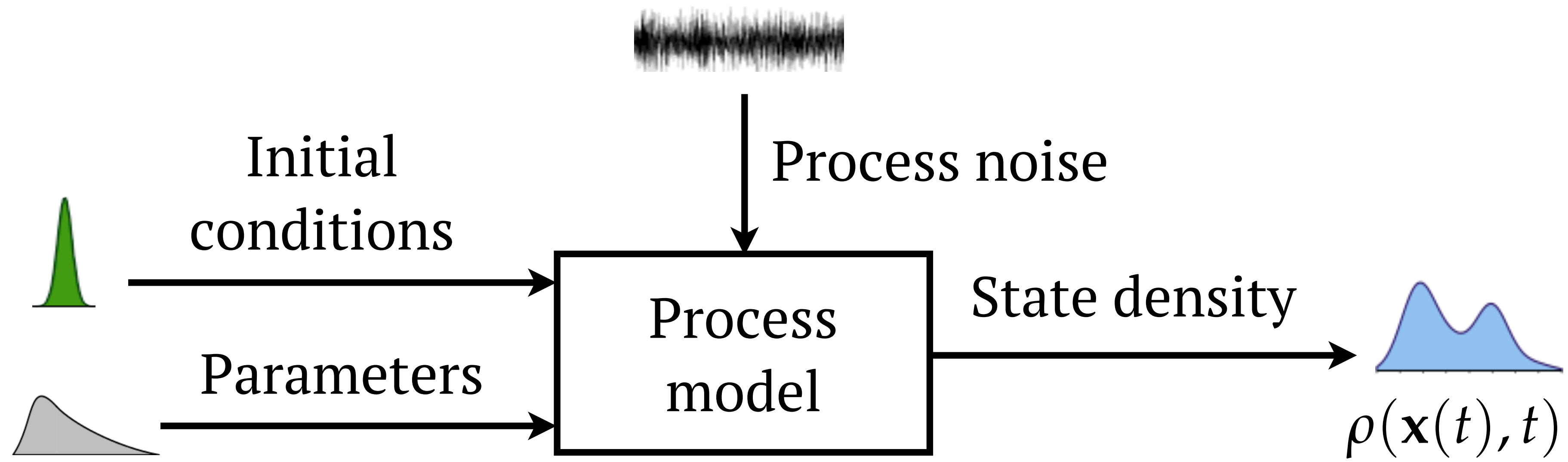
population measure population density function

$$\int_{\mathcal{X}} \mathrm{d}\mu = \int_{\mathcal{X}} \rho \, \mathrm{d}\boldsymbol{x} = 1 \quad \text{for all } t \in [0, \infty)$$

Why care about densities?

Prediction Problem

Compute
joint state PDF
$\rho\left(x, t\right)$



Trajectory flow:

$$\mathrm{d}\mathbf{X}(t) = \mathbf{f}(\mathbf{X}, t)\,\mathrm{d}t + \mathbf{g}(\mathbf{X}, t)\,\mathrm{d}\mathbf{w}(t), \quad \mathrm{d}\mathbf{w}(t) \sim \mathcal{N}(0, \mathbf{Q}\mathrm{d}t)$$

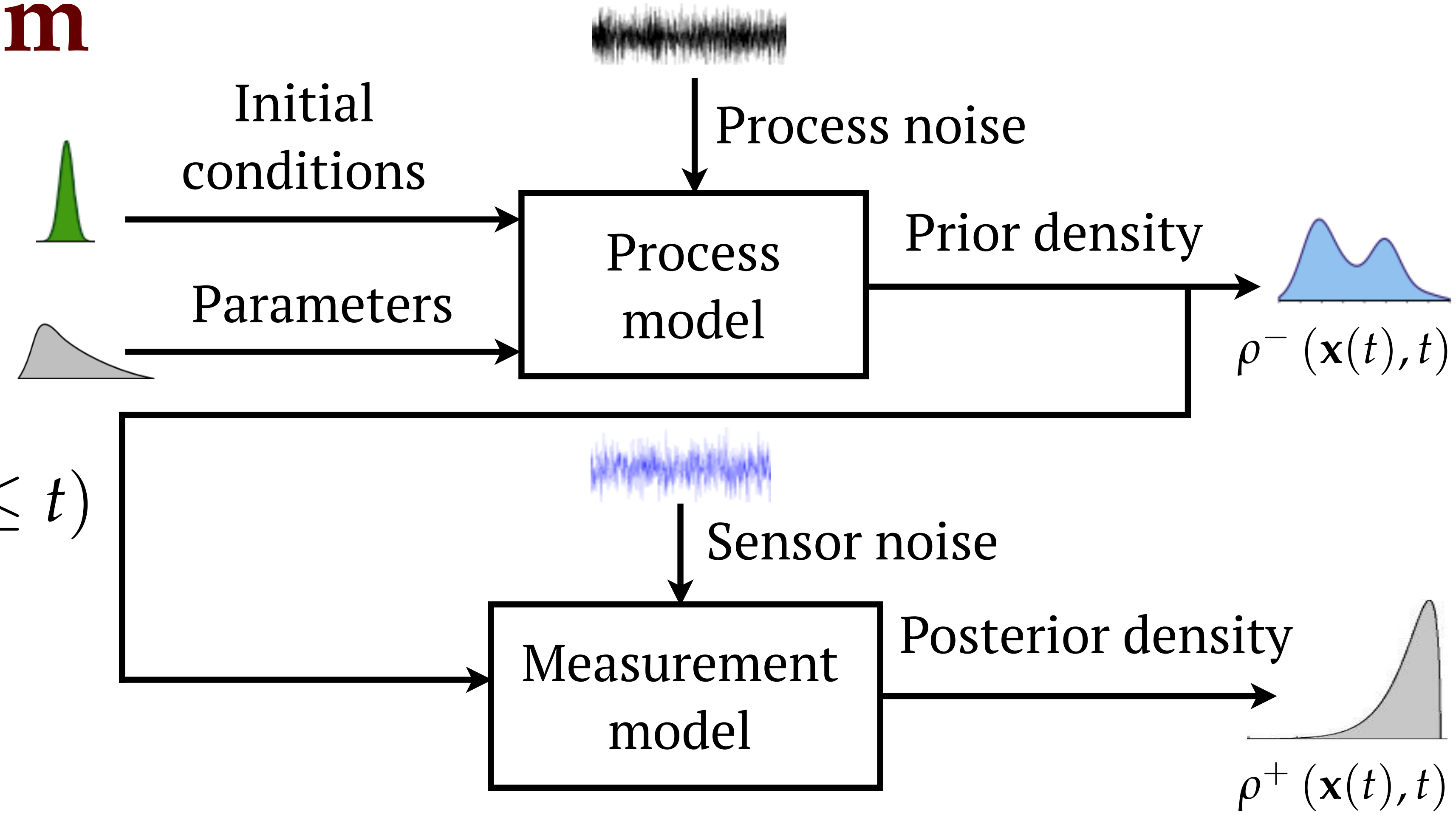
Density flow:

$$\frac{\partial \rho}{\partial t} = \mathcal{L}_{\mathrm{FP}}(\rho) := -\nabla \cdot (\rho \mathbf{f}) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} \left(\left(\mathbf{g}\mathbf{Q}\mathbf{g}^{\top} \right)_{ij} \rho \right)$$

Filtering Problem

Compute conditional joint state PDF

$$\rho^{+} := \rho\left(x, t \mid z(s), 0 \leq s \leq t\right)$$



Trajectory flow:

$$\mathrm{d}\mathbf{X}(t) = \mathbf{f}(\mathbf{X}, t)\,\mathrm{d}t + \mathbf{g}(\mathbf{X}, t)\,\mathrm{d}\mathbf{w}(t), \quad \mathrm{d}\mathbf{w}(t) \sim \mathcal{N}(0, \mathbf{Q}\mathrm{d}t)$$

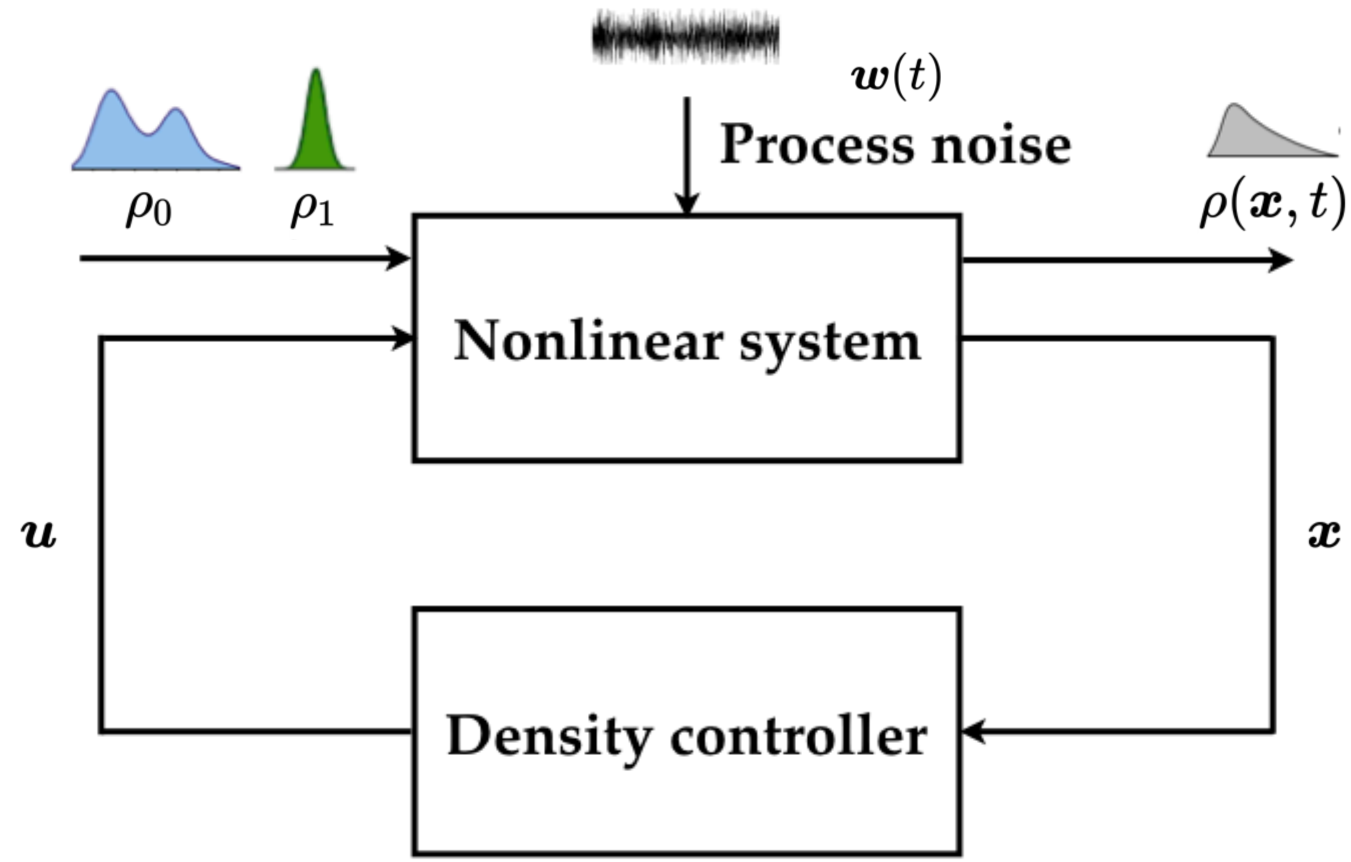
$$\mathrm{d}\mathbf{Z}(t) = \mathbf{h}(\mathbf{X}, t)\,\mathrm{d}t + \mathrm{d}\mathbf{v}(t), \quad \mathrm{d}\mathbf{v}(t) \sim \mathcal{N}(0, \mathbf{R}\mathrm{d}t)$$

Density flow:

$$\mathrm{d}\rho^{+} = \left[\mathcal{L}_{\mathrm{FP}}\mathrm{d}t + \left(\mathbf{h}(\mathbf{x}, t) - \mathbb{E}_{\rho^{+}}\{\mathbf{h}(\mathbf{x}, t)\}\right)^{\top}\mathbf{R}^{-1}\left(\mathrm{d}\mathbf{z}(t) - \mathbb{E}_{\rho^{+}}\{\mathbf{h}(\mathbf{x}, t)\}\mathrm{d}t\right)\right]\rho^{+}$$

Control Problem

Steer joint state PDF via feedback control over finite time horizon



$$\underset{u \in \mathcal{U}}{\text{minimize}} \quad \mathbb{E}\left[\int_0^1 \|\boldsymbol{u}\|_2^2 \,\mathrm{d}t\right]$$

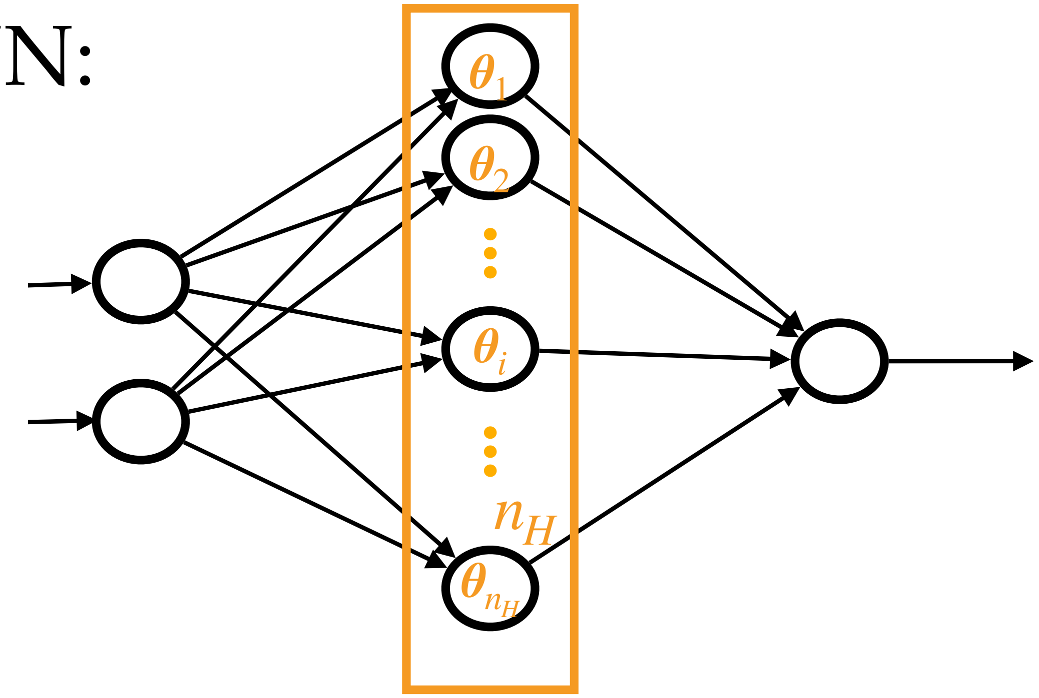
subject to

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{f}\left(\boldsymbol{x}, \boldsymbol{u}, t\right)\,\mathrm{d}t + \boldsymbol{g}\left(\boldsymbol{x}, t\right)\,\mathrm{d}\boldsymbol{w},$$

$$\boldsymbol{x}\left(t = 0\right) \sim \rho_0, \quad \boldsymbol{x}\left(t = 1\right) \sim \rho_1$$

Mean Field Neural Network Learning Problem

Infinite width limit of fully connected NN:



Mei, Montanari and Nguyen, *Proceedings of the National Academy of Sciences*, 2018

Chizat and Bach, *NeurIPS*, 2018

Rotskoff and Vanden-Eijnden, *NeurIPS*, 2018

Sirignano and Spiliopoulos, *Stochastic Processes and their Applications*, 2020

Mean field learning problem: $$\inf_{\rho \in \mathcal{P}_2(\mathbb{R}^p)} R\left(\int \Phi(\boldsymbol{x}, \boldsymbol{\theta})\rho(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\right)$$

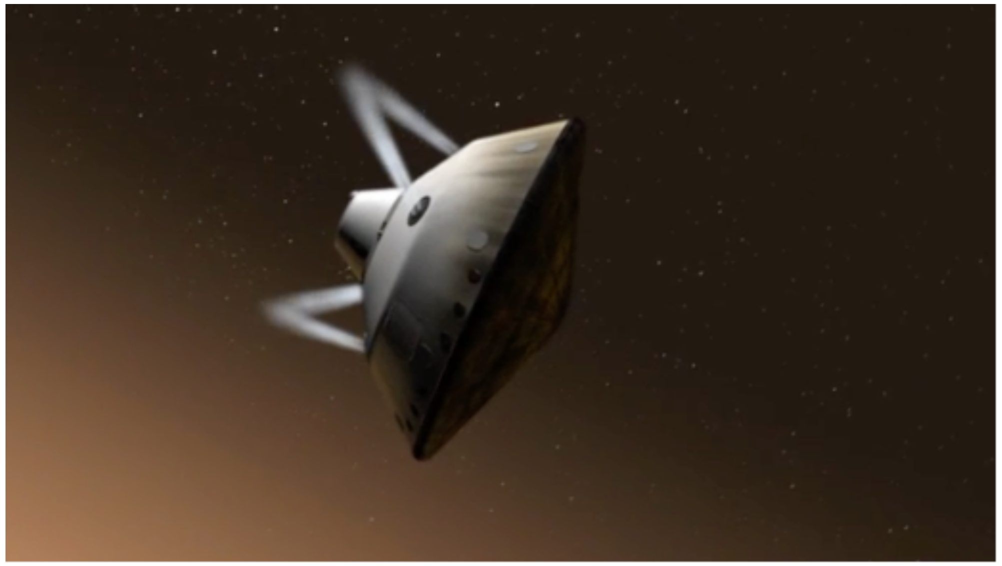
manifold of PDFs supported on $\mathbb{R}^p$ with finite second moments

PDF dynamics:

$$\frac{\partial \rho}{\partial t} = -\nabla^W R\left(\int \Phi\rho\right) = \nabla \cdot \left(\rho \nabla \frac{\delta}{\delta\rho} R\left(\int \Phi\rho\right)\right)$$

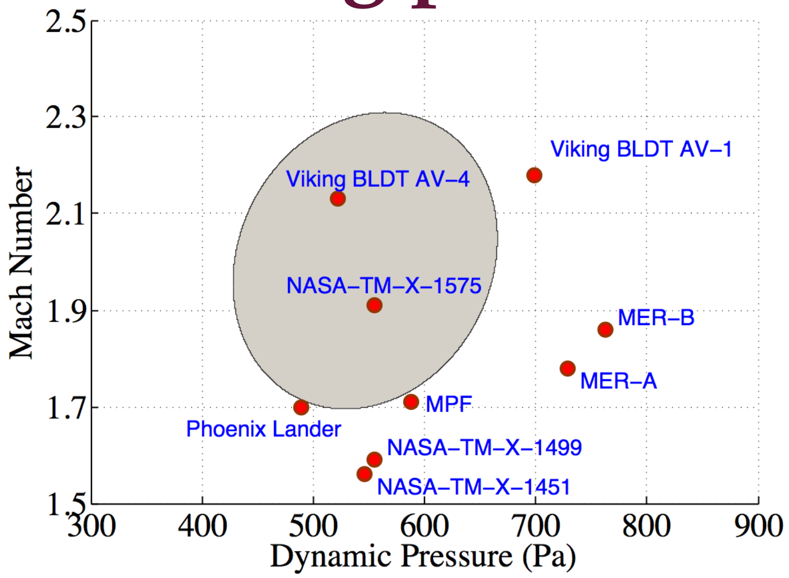
PDFs in Mars Entry-Descent-Landing

Prediction problem



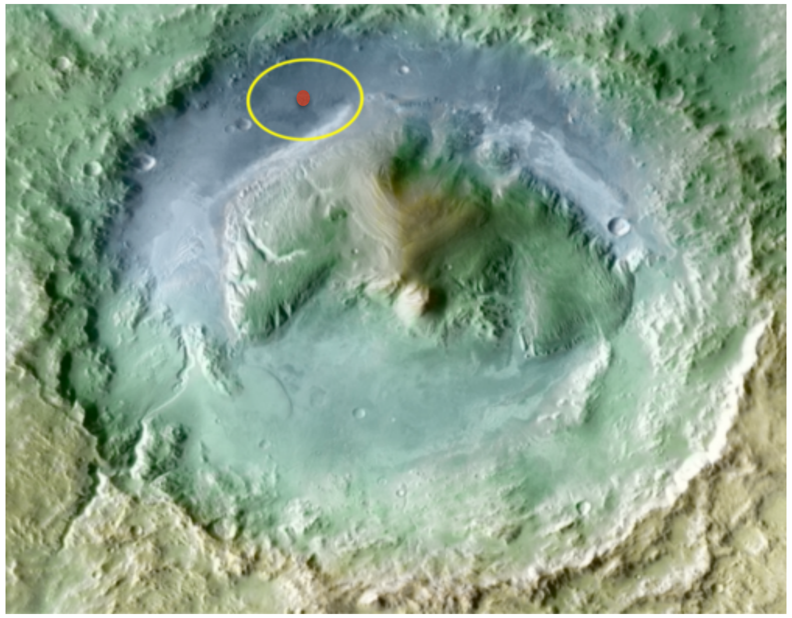
Predict heating rate uncertainty

Filtering problem



Estimate state to deploy parachute

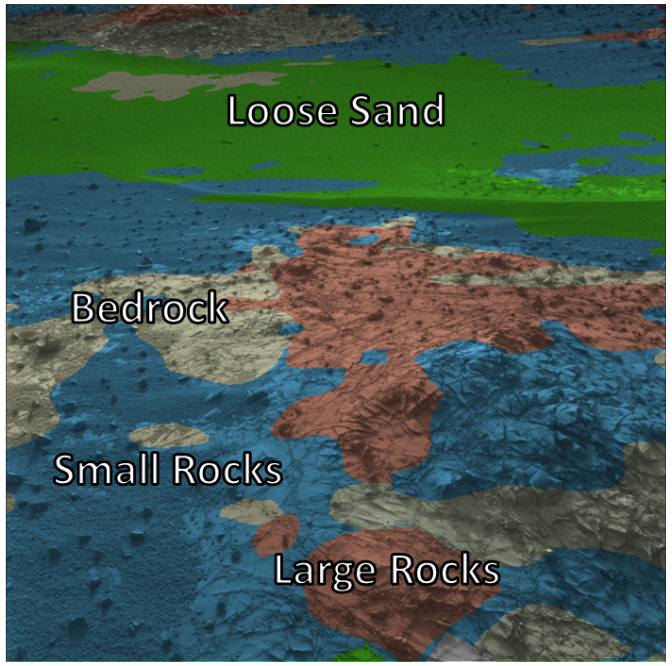
Control problem



Gale Crater (4.49S, 137.42E)

Steer state PDF to achieve desired landing footprint accuracy

Learning problem



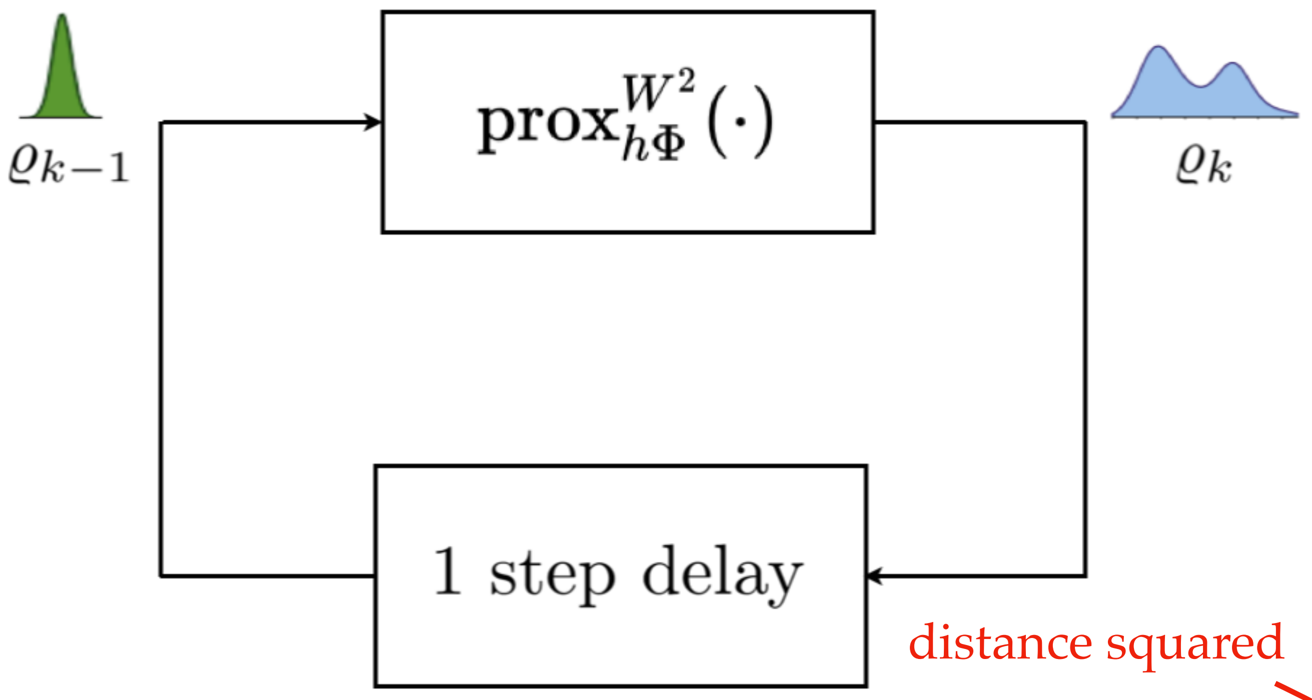
Learn surface feature from data

Solving prediction problem as Wasserstein gradient flow

What's New?

Main idea: Solve $\dfrac{\partial \rho}{\partial t} = \mathcal{L}_{\mathrm{FP}}\rho, \;\; \rho(x, t=0) = \rho_0$ **as gradient flow in** $\mathcal{P}_2(\mathcal{X})$

Infinite dimensional variational recursion:



distance squared — time-step — energy-like functional

Proximal operator: $$\varrho_k = \mathrm{prox}_{h\Phi}^{W^2}(\varrho_{k-1}) := \underset{\varrho \in \mathcal{P}_2(\mathcal{X})}{\arg\inf} \left\{ \frac{1}{2} W^2(\varrho, \varrho_{k-1}) + h\Phi(\varrho) \right\}$$

Optimal transport cost: $$W^2(\varrho, \varrho_{k-1}) := \inf_{\pi \in \Pi(\varrho, \varrho_{k-1})} \int_{\mathcal{X} \times \mathcal{X}} c(x, y)\, \mathrm{d}\pi(x, y)$$

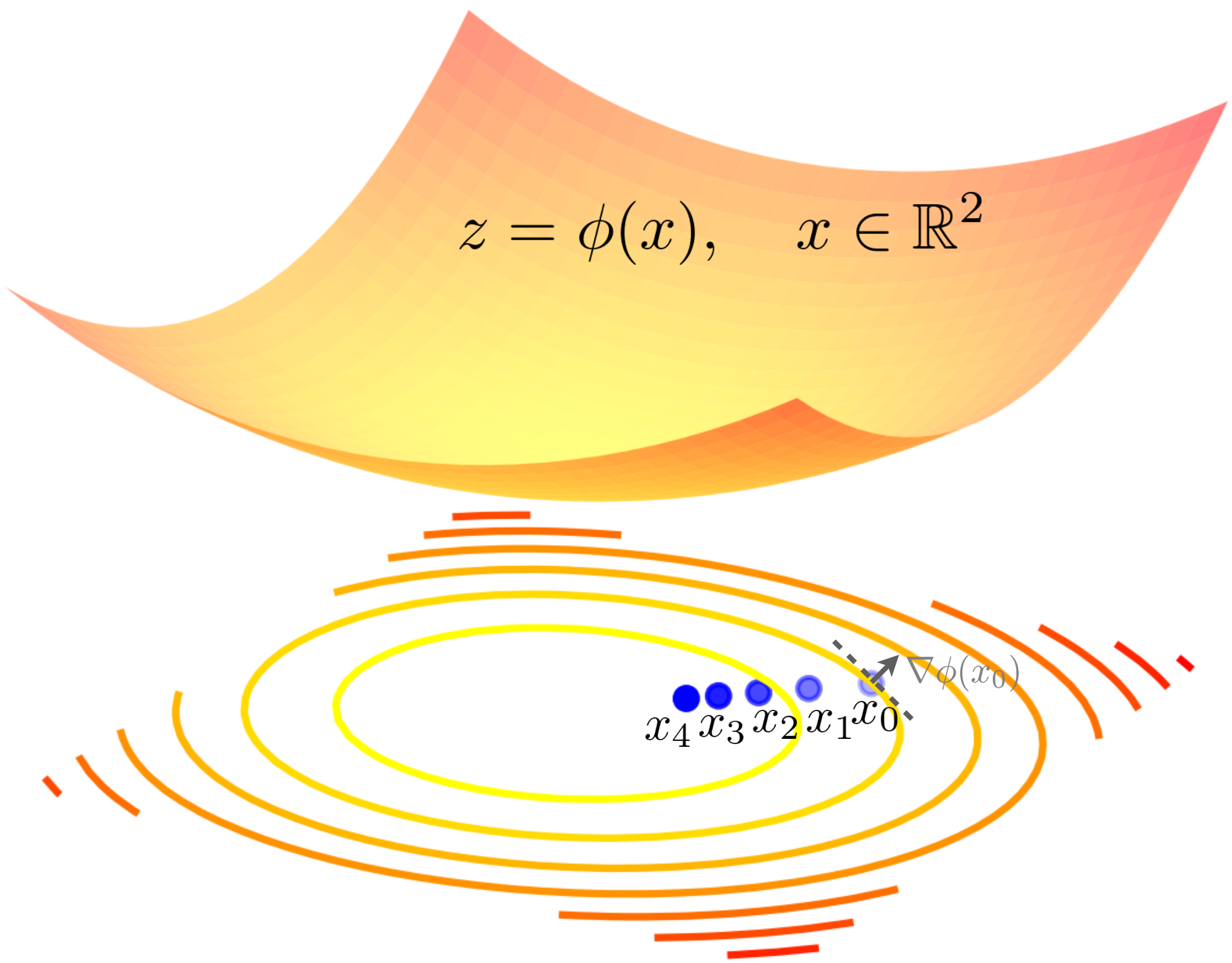
Free energy functional: $$\Phi(\varrho) := \int_{\mathcal{X}} \psi \varrho \, \mathrm{d}x + \beta^{-1} \int_{\mathcal{X}} \varrho \log \varrho \, \mathrm{d}x$$

Geometric Meaning of Gradient Flow

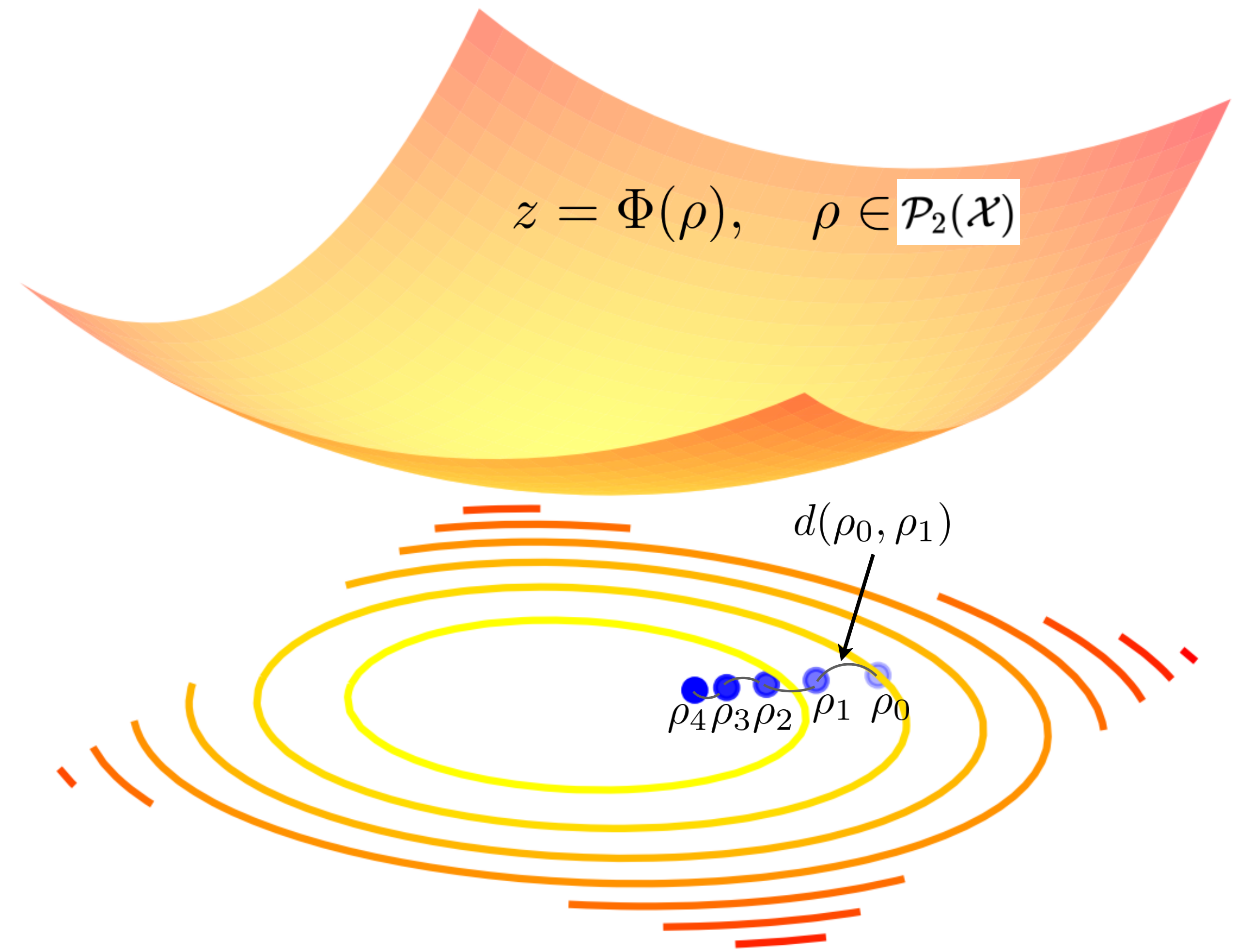
Gradient Flow in $\mathcal{X}$	Gradient Flow in $\mathcal{P}_2(\mathcal{X})$
$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$	$\frac{\partial\rho}{\partial t} = -\nabla^W\Phi(\rho), \quad \rho(\boldsymbol{x},0) = \rho_0$
Recursion: $\boldsymbol{x}_k = \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k) = \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\min}\left\{\frac{1}{2}\lVert\boldsymbol{x}-\boldsymbol{x}_{k-1}\rVert_2^2 + h\varphi(\boldsymbol{x})\right\} =: \mathrm{prox}_{h\varphi}^{\lVert\cdot\rVert_2}(\boldsymbol{x}_{k-1})$	**Recursion:** $\rho_k = \rho(\cdot, t = kh) = \underset{\rho\in\mathcal{P}_2(\mathcal{X})}{\arg\min}\left\{\frac{1}{2}W^2(\rho,\rho_{k-1}) + h\Phi(\rho)\right\} =: \mathrm{prox}_{h\Phi}^{W^2}(\rho_{k-1})$
Convergence: $\boldsymbol{x}_k \to \boldsymbol{x}(t = kh) \quad \text{as} \quad h\downarrow 0$	**Convergence:** $\rho_k \to \rho(\cdot, t = kh) \quad \text{as} \quad h\downarrow 0$
φ as Lyapunov function: $\frac{\mathrm{d}}{\mathrm{d}t}\varphi = -\lVert\nabla\varphi\rVert_2^2 \;\leq\; 0$	**Φ as Lyapunov functional:** $\frac{\mathrm{d}}{\mathrm{d}t}\Phi = -\mathbb{E}_\rho\left[\left\lVert\nabla\frac{\delta\Phi}{\delta\rho}\right\rVert_2^2\right] \;\leq\; 0$

Geometric Meaning of Gradient Flow

Gradient Flow in $\mathcal{X}$

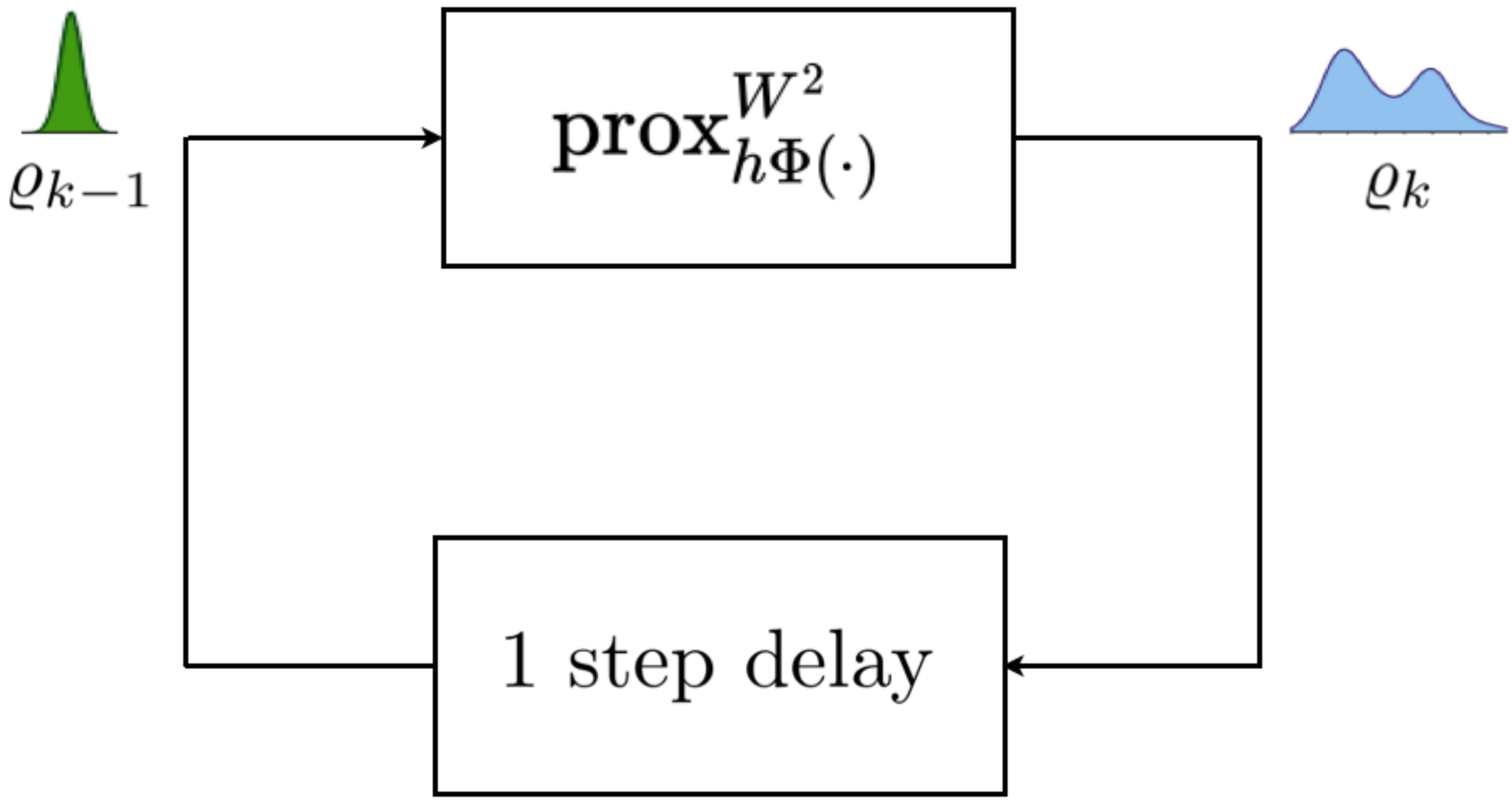


Gradient Flow in $\mathcal{P}_2(\mathcal{X})$



Algorithm: Gradient Ascent on the Dual Space

Uncertainty propagation via point clouds



No spatial discretization or function approximation

Algorithm: Gradient Ascent on the Dual Space

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \psi \rho) + \beta^{-1} \Delta \rho$$

$\Updownarrow$ **Proximal Recursion**

$$\rho_k = \rho(\boldsymbol{x}, t = kh) = \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$$

$\Downarrow$ **Discrete Primal Formulation**

$$\boldsymbol{\varrho}_k = \underset{\boldsymbol{\varrho}}{\arg\min} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + h \, \langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \rangle \right\}$$

$\Downarrow$ **Entropic Regularization**

$$\boldsymbol{\varrho}_k = \underset{\boldsymbol{\varrho}}{\arg\min} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + \epsilon H(\boldsymbol{M}) + h \, \langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho}, \boldsymbol{\varrho} \rangle \right\}$$

$\Updownarrow$ **Dualization**

$$\boldsymbol{\lambda}_0^{\text{opt}}, \boldsymbol{\lambda}_1^{\text{opt}} = \underset{\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1 \geq 0}{\arg\max} \Big\{ \langle \boldsymbol{\lambda}_0, \boldsymbol{\varrho}_{k-1} \rangle - F^\star(-\boldsymbol{\lambda}_1) - \frac{\epsilon}{h} \left(\exp(\boldsymbol{\lambda}_0^\top h/\epsilon) \exp(-\boldsymbol{C}_k/2\epsilon) \exp(\boldsymbol{\lambda}_1 h/\epsilon) \right) \Big\}$$

Recursion on the Cone

$$y = e^{\frac{\lambda_0^*}{\epsilon}h} \downarrow \qquad \downarrow z = e^{\frac{\lambda_1^*}{\epsilon}h}$$

Coupled Transcendental Equations in y and z

$$\Gamma_k = e^{\frac{-C_k}{2\epsilon}} \longrightarrow, \qquad \varrho_{k-1} \longrightarrow, \qquad \xi_{k-1} = \frac{e^{-\beta\psi_{k-1}}}{e} \longrightarrow$$

$$\boxed{\begin{aligned} y \odot \Gamma_k z &= \varrho_{k-1} \\ z \odot \Gamma_k^\top y &= \xi_{k-1} \odot z^{-\beta\epsilon/2h} \end{aligned}} \longrightarrow \varrho_k = z \odot \Gamma_k^\top y$$

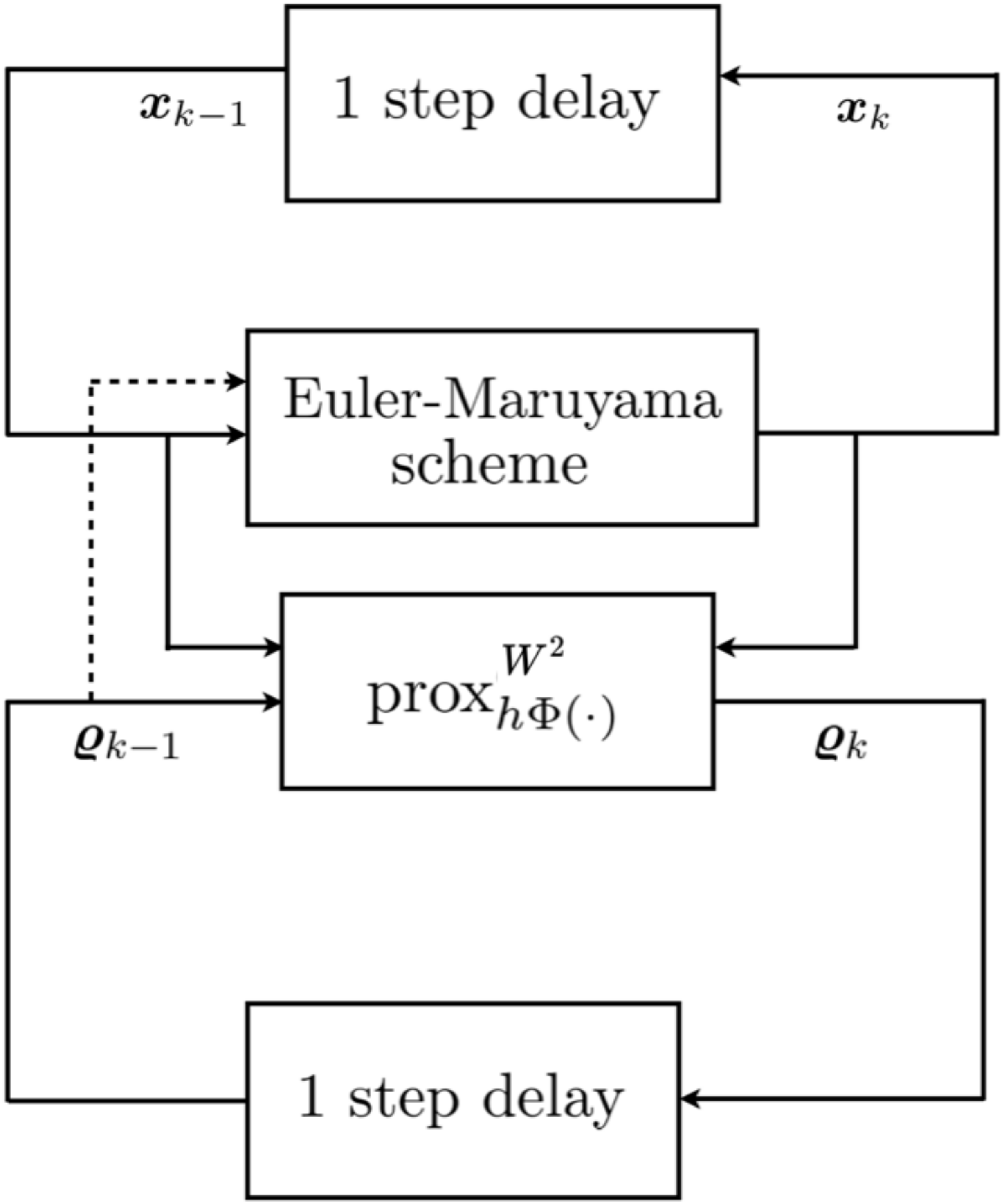
Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

$$y \odot (\Gamma_k z) = \varrho_{k-1}, \quad z \odot \left(\Gamma_k{}^\top y\right) = \xi_{k-1} \odot z^{-\frac{\beta\epsilon}{h}},$$

Then the solution (y^*, z^*) gives the proximal update $\varrho_k = z^* \odot (\Gamma_k{}^\top y^*)$

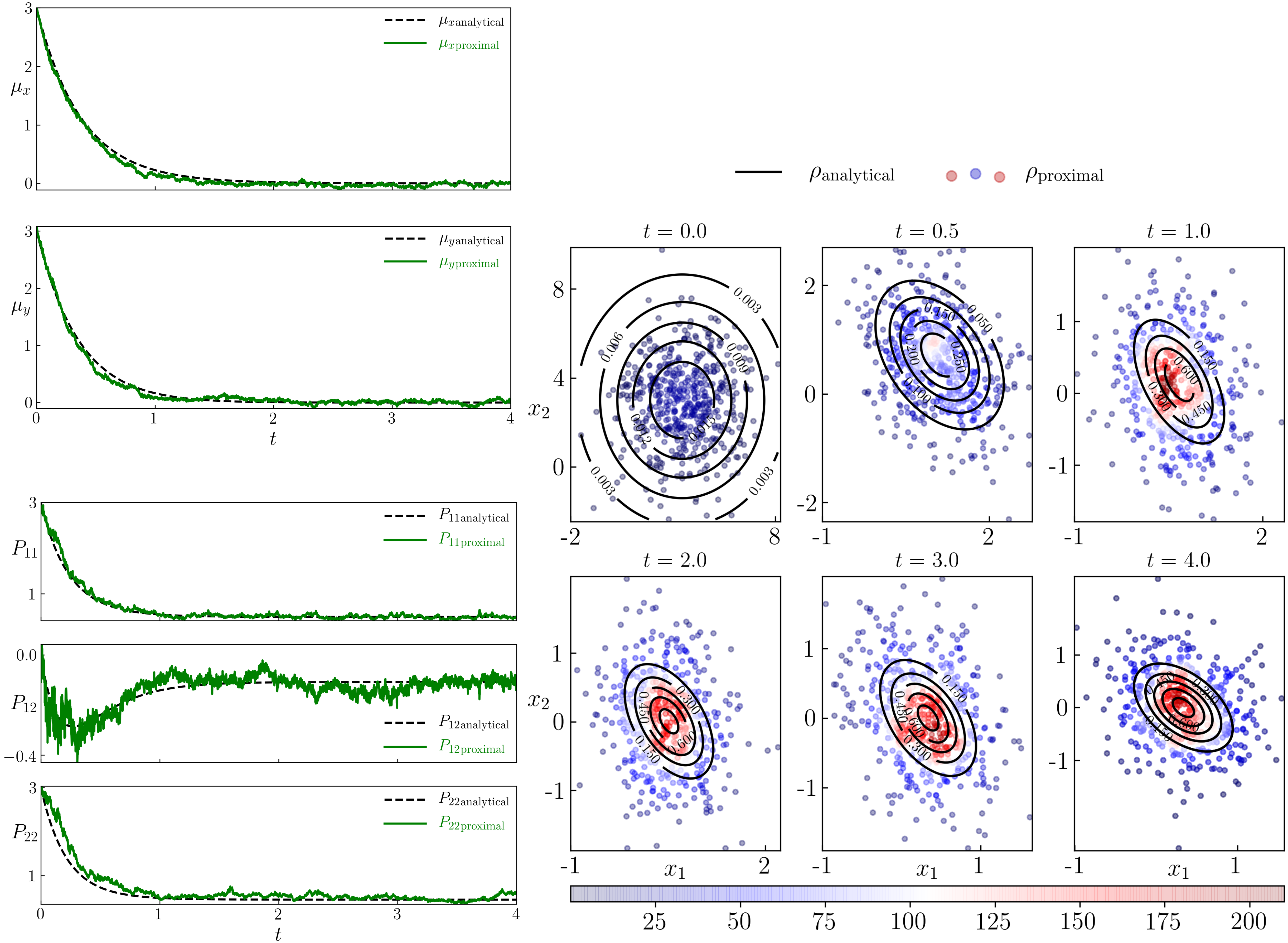
K.F. Caluya and A.H., Gradient flow algorithms for density propagation in stochastic systems, *IEEE TAC* 2019.

Algorithmic Setup

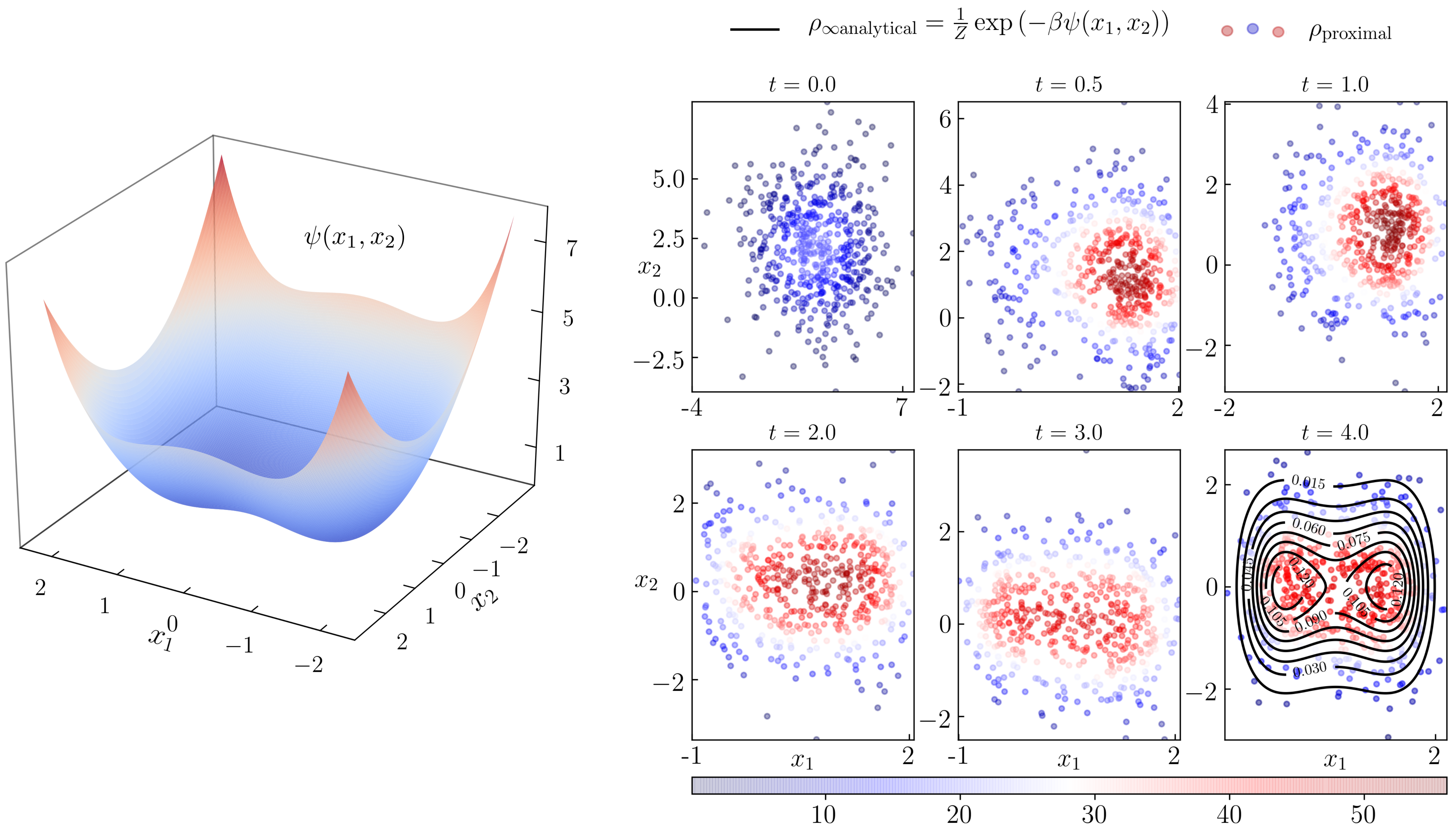


Theorem: Block co-ordinate iteration of $(\mathbf{y}, \mathbf{z})$ recursion is contractive on $\mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0}$.

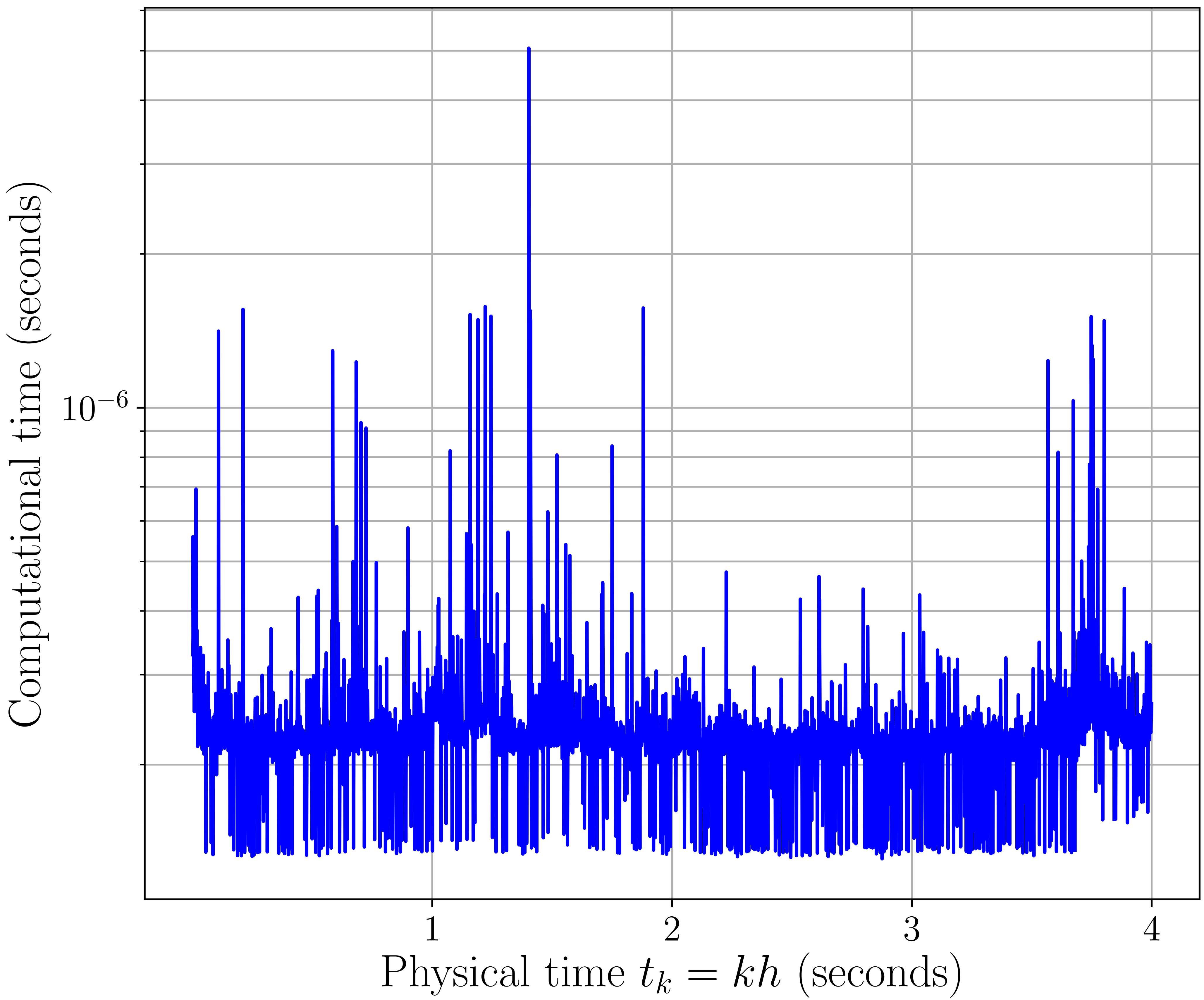
Proximal Prediction: 2D Linear Gaussian



Proximal Prediction: Nonlinear Non-Gaussian



Computational Time: Nonlinear Non-Gaussian



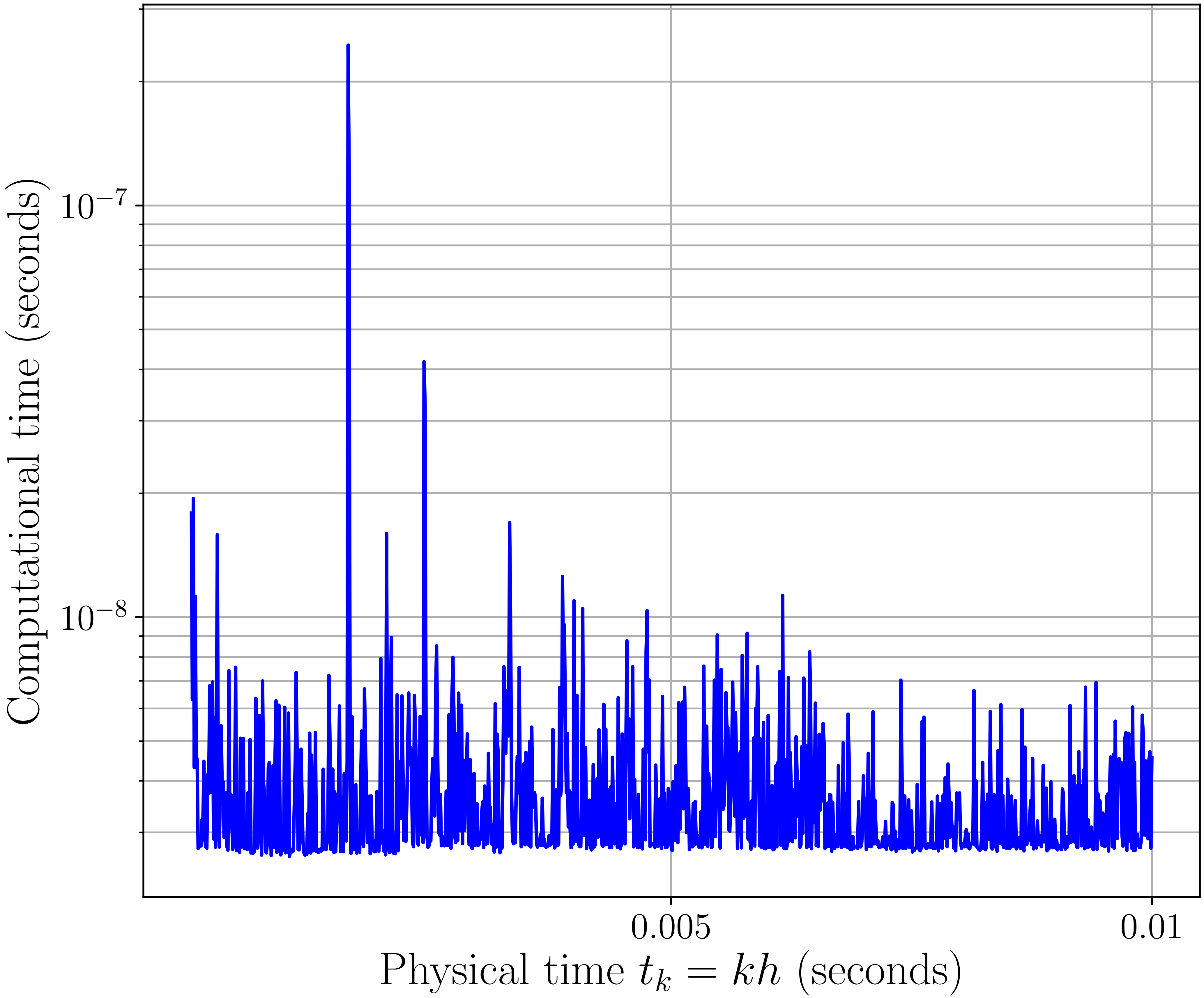
Proximal Prediction: Satellite in Geocentric Orbit

Here, $\mathcal{X} \equiv \mathbb{R}^6$

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \\ \mathrm{d}v_x \\ \mathrm{d}v_y \\ \mathrm{d}v_z \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ v_z \\ -\dfrac{\mu x}{r^3} + (f_x)_{\text{pert}} - \gamma v_x \\ -\dfrac{\mu y}{r^3} + (f_y)_{\text{pert}} - \gamma v_y \\ -\dfrac{\mu z}{r^3} + (f_z)_{\text{pert}} - \gamma v_z \end{pmatrix} \mathrm{d}t + \sqrt{2\beta^{-1}\gamma} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathrm{d}w_1 \\ \mathrm{d}w_2 \\ \mathrm{d}w_3 \end{pmatrix},$$

$$\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}_{\text{pert}} = \begin{pmatrix} s\theta\, c\phi & c\theta\, c\phi & -s\phi \\ s\theta\, s\phi & c\theta\, s\phi & c\phi \\ c\theta & -s\theta & 0 \end{pmatrix} \begin{pmatrix} \dfrac{k}{2r^4}\left(3(s\theta)^2 - 1\right) \\ -\dfrac{k}{r^5} s\theta\, c\theta \\ 0 \end{pmatrix}, k := 3J_2 R_{\mathrm{E}}^2, \mu = \text{constant}$$

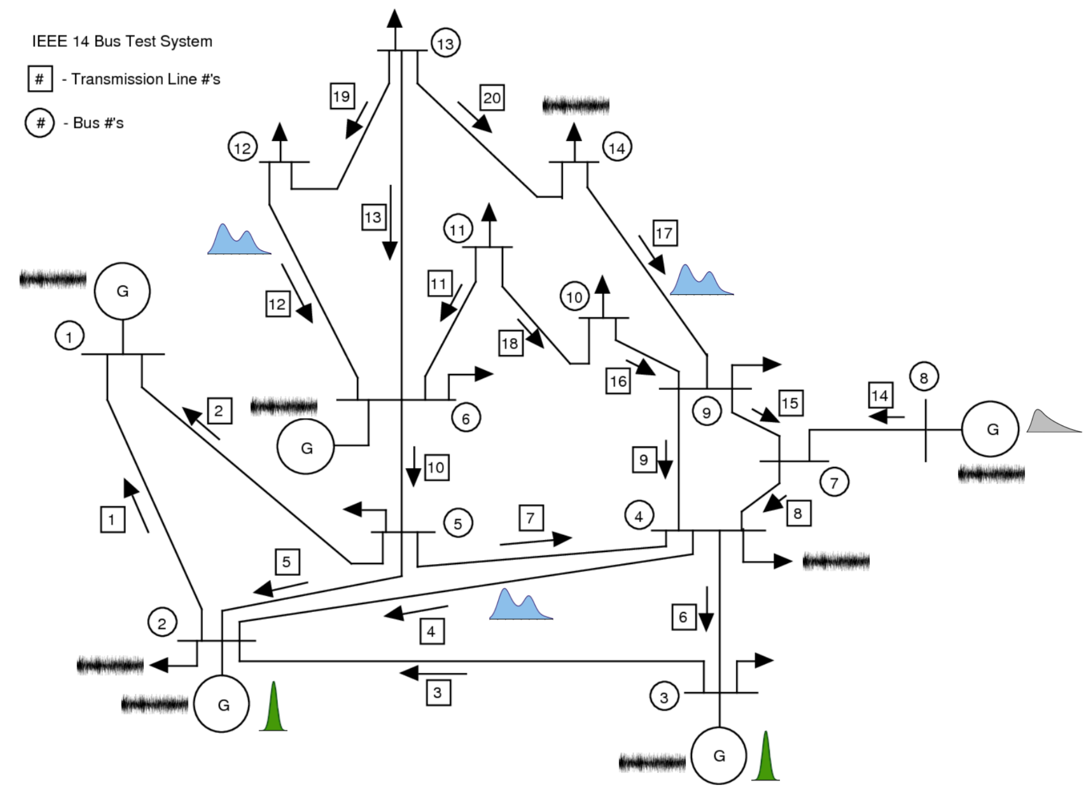
Computational Time: Satellite in Geocentric Orbit



Extensions and Applications

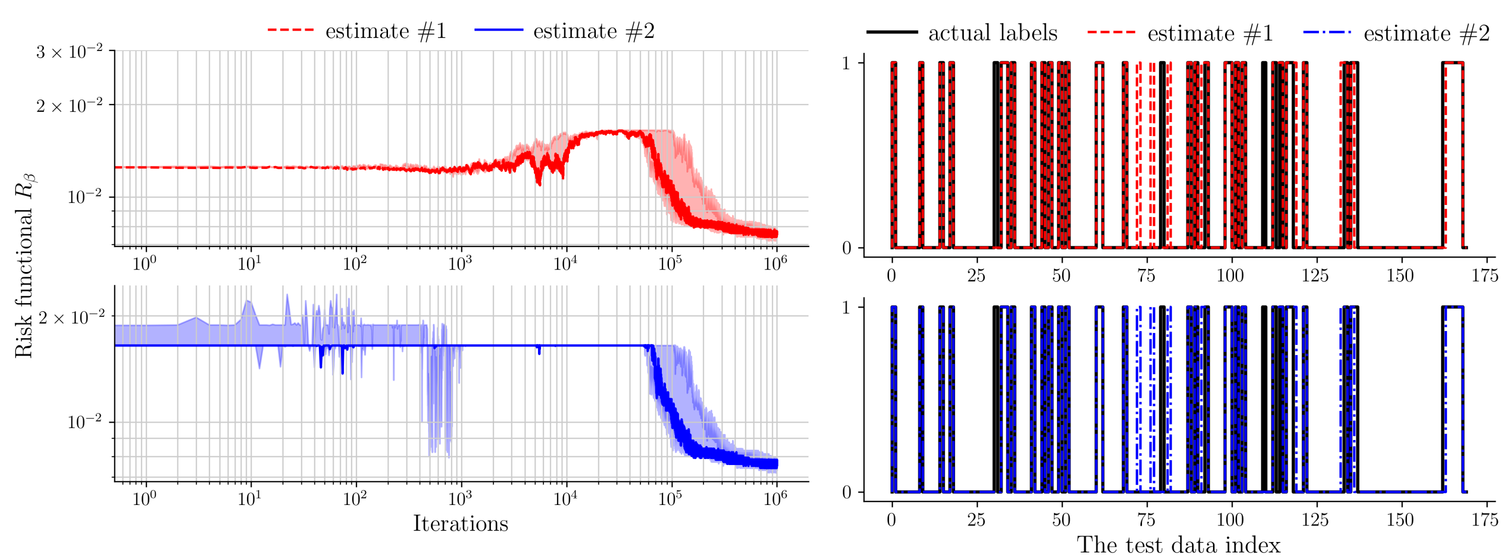
Networked nonlinear power system dynamics with $O(100)$ states

A.H., K.F. Caluya, P. Ojaghi, and X. Geng, Stochastic uncertainty propagation in power system dynamics with measure-valued proximal recursions, *IEEE Transactions on Power Systems*, 2022.



Mean field learning in NN

A.M.H. Teter, I. Nodozi, and A.H., Proximal mean field learning in shallow neural networks, *arXiv:2210:13879*, 2022.



Case study: Wisconsin Breast Cancer Diagnostic (WBCD) Data Set

GPU: Jetson TX2 NVIDIA Pascal GPU 256 CUDA cores, 64 bit NVIDIA Denver + ARM Cortex A57 CPUs (≈ 2 hrs runtime)

Classification accuracy for the WBDC dataset

β	Estimate #1	Estimate #2
0.03	91.17%	92.35%
0.05	92.94%	92.94%
0.07	78.23%	92.94%

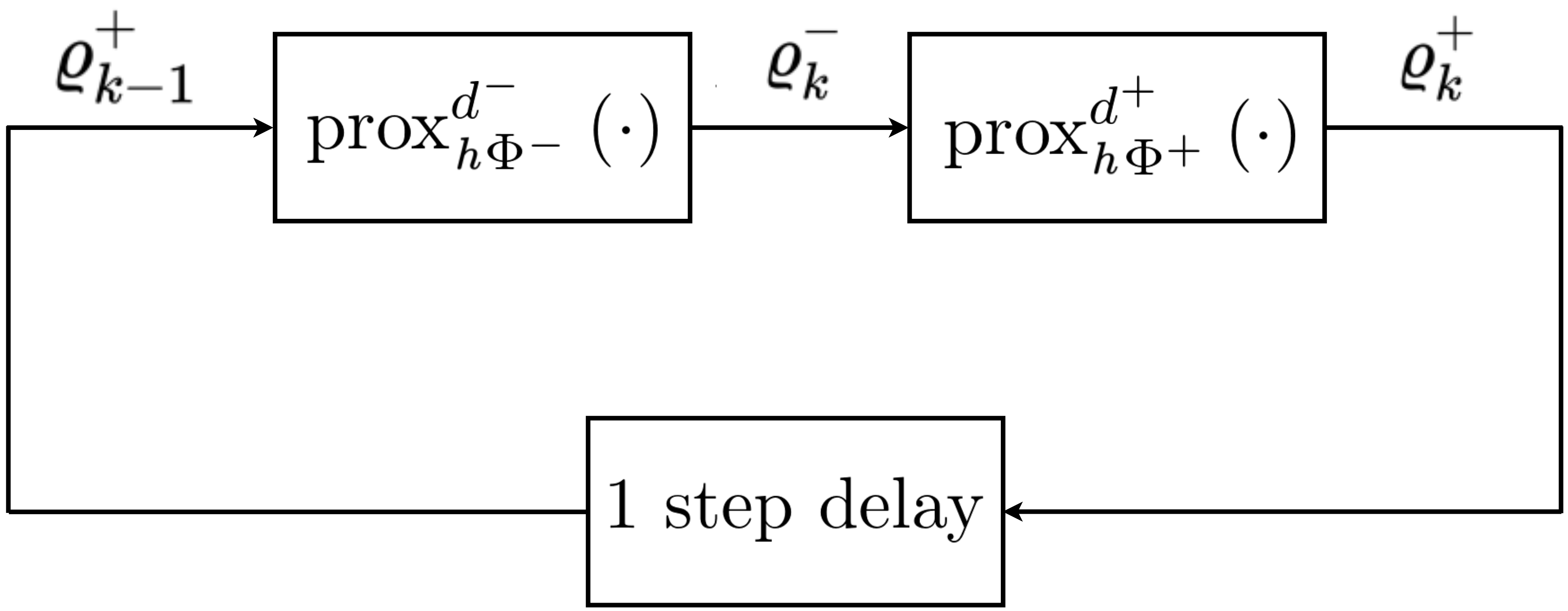
Solving filtering as generalized gradient flow

What's New?

Main idea: Solve the Kushner-Stratonovich SPDE

$$\mathrm{d}\rho^{+} = \left[\mathcal{L}_{\mathrm{FP}}\mathrm{d}t + \mathcal{L}\left(\mathrm{d}z, \mathrm{d}t, \rho^{+}\right)\right]\rho^{+}, \;\; \rho(x, t=0) = \rho_0 \text{ as gradient flow in } \mathcal{P}_2(\mathcal{X})$$

Recursion of {deterministic ∘ stochastic} proximal operators:



Convergence: $\varrho_k^{+}(h) \rightarrow \rho^{+}(x, t = kh) \quad \text{as} \quad h \downarrow 0$

For prior, as before: $d^{-} \equiv W^2, \quad \Phi^{-} \equiv \mathbb{E}_{\varrho}\left[\psi + \beta^{-1}\log\varrho\right]$

For posterior: $d^{+} \equiv d_{\mathrm{FR}}^2 \text{ or } D_{\mathrm{KL}}, \quad \Phi^{+} \equiv \dfrac{1}{2}\mathbb{E}_{\varrho^{+}}\left[(y_k - h(x))^{\top}R^{-1}(y_k - h(x))\right]$

Explicit Recovery of the Kalman-Bucy Filter

Model:

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{B}d\mathbf{w}(t), \quad d\mathbf{w}(t) \sim \mathcal{N}(0, \mathbf{Q}dt)$$

$$d\mathbf{z}(t) = \mathbf{C}\mathbf{x}(t)dt + d\mathbf{v}(t), \quad d\mathbf{v}(t) \sim \mathcal{N}(0, \mathbf{R}dt)$$

Given $\mathbf{x}(0) \sim \mathcal{N}(\mu_0, \mathbf{P}_0)$, want to recover:

$$\mathbf{P}^{+}\mathbf{C}\mathbf{R}^{-1}$$

$$d\mu^{+}(t) = \mathbf{A}\mu^{+}(t)dt + \mathbf{K}(t)\left(d\mathbf{z}(t) - \mathbf{C}\mu^{+}(t)dt\right),$$

$$\dot{\mathbf{P}}^{+}(t) = \mathbf{A}\mathbf{P}^{+}(t) + \mathbf{P}^{+}(t)\mathbf{A}^{\top} + \mathbf{B}\mathbf{Q}\mathbf{B}^{\top} - \mathbf{K}(t)\mathbf{R}\mathbf{K}(t)^{\top}.$$

A.H. and T.T. Georgiou, Gradient Flows in Uncertainty Propagation and Filtering of Linear Gaussian Systems, *CDC 2017*.

A.H. and T.T. Georgiou, Gradient Flows in Filtering and Fisher-Rao Geometry, *ACC 2018*.

Explicit Recovery of the Wonham Filter

Model:

$$x(t) \sim \text{Markov}(Q),$$
$$\mathrm{d}z(t) = h(x(t))\,\mathrm{d}t \;+\; \sigma_v(t)\mathrm{d}v(t)$$

J.SIAM Control
Ser. A, Vol. 2, No. 3
Printed in U.S.A., 1965

SOME APPLICATIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS TO OPTIMAL NONLINEAR FILTERING*

W. M. WONHAM†

State space: $\Omega := \{a_1, \ldots, a_m\}$

Posterior $\pi^+(t) := \{\pi_1^+(t), \ldots, \pi_m^+(t)\}$ **solves the nonlinear SDE:**

$$\mathrm{d}\pi^+(t) = \pi^+(t)Q\,\mathrm{d}t \;+\; \frac{1}{(\sigma_v(t))^2}\pi^+(t)\Big(H - \widehat{h}(t)I\Big)\Big(\mathrm{d}z(t) - \widehat{h}(t)\mathrm{d}t\Big),$$

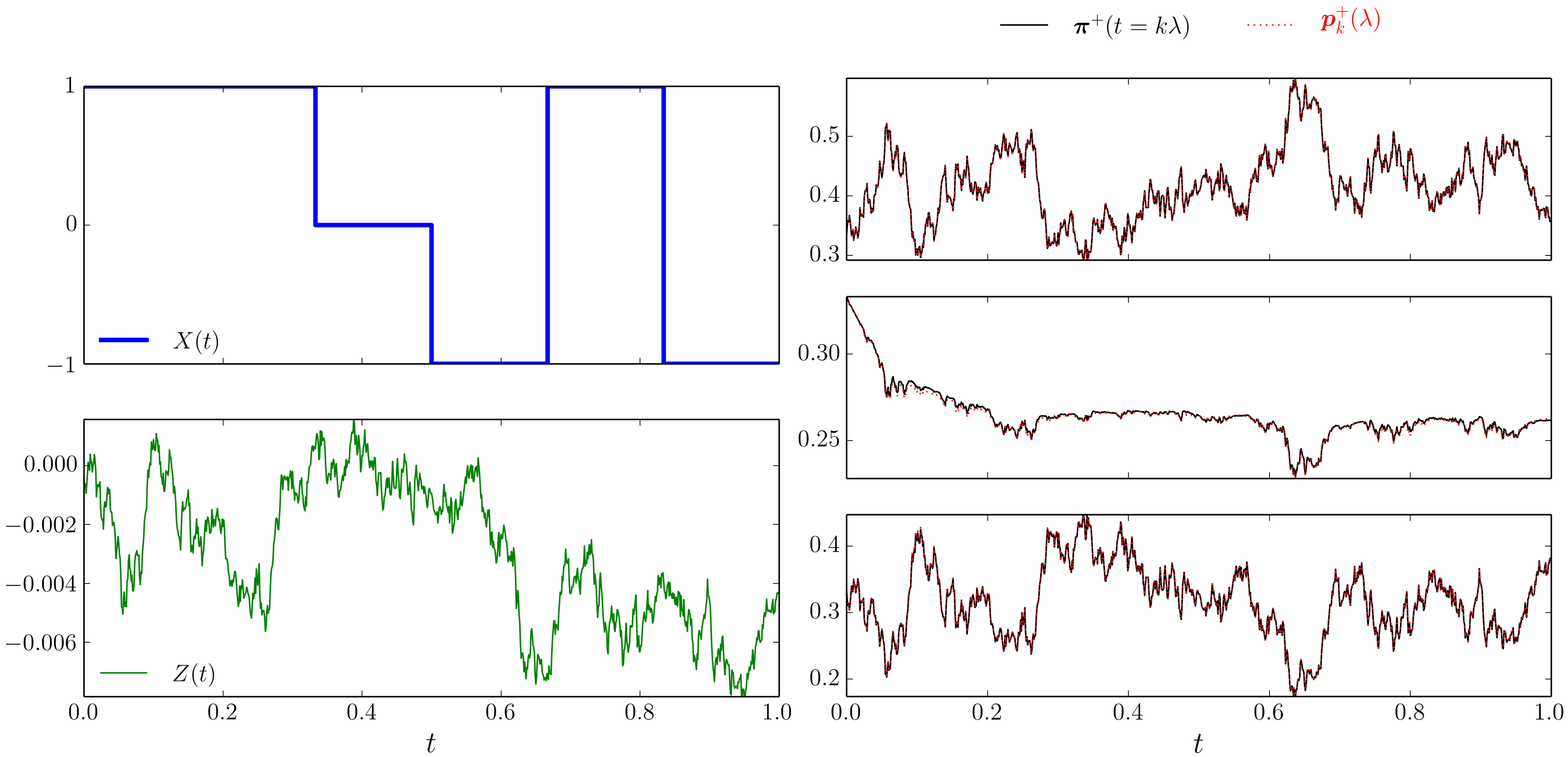
where $H := \text{diag}(h(a_1), \ldots, h(a_m)), \quad \widehat{h}(t) := \sum_{i=1}^{m} h(a_i)\pi_i^+(t),$

Initial condition: $\pi^+(t=0) = \pi_0,$

By defn. $\pi^+(t) = \mathbb{P}(x(t) = a_i \mid z(s), 0 \leq s \leq t)$

A.H. and T.T. Georgiou, Proximal Recursion for the Wonham Filter, *CDC 2019*.

Numerical Results for the Wonham Filter

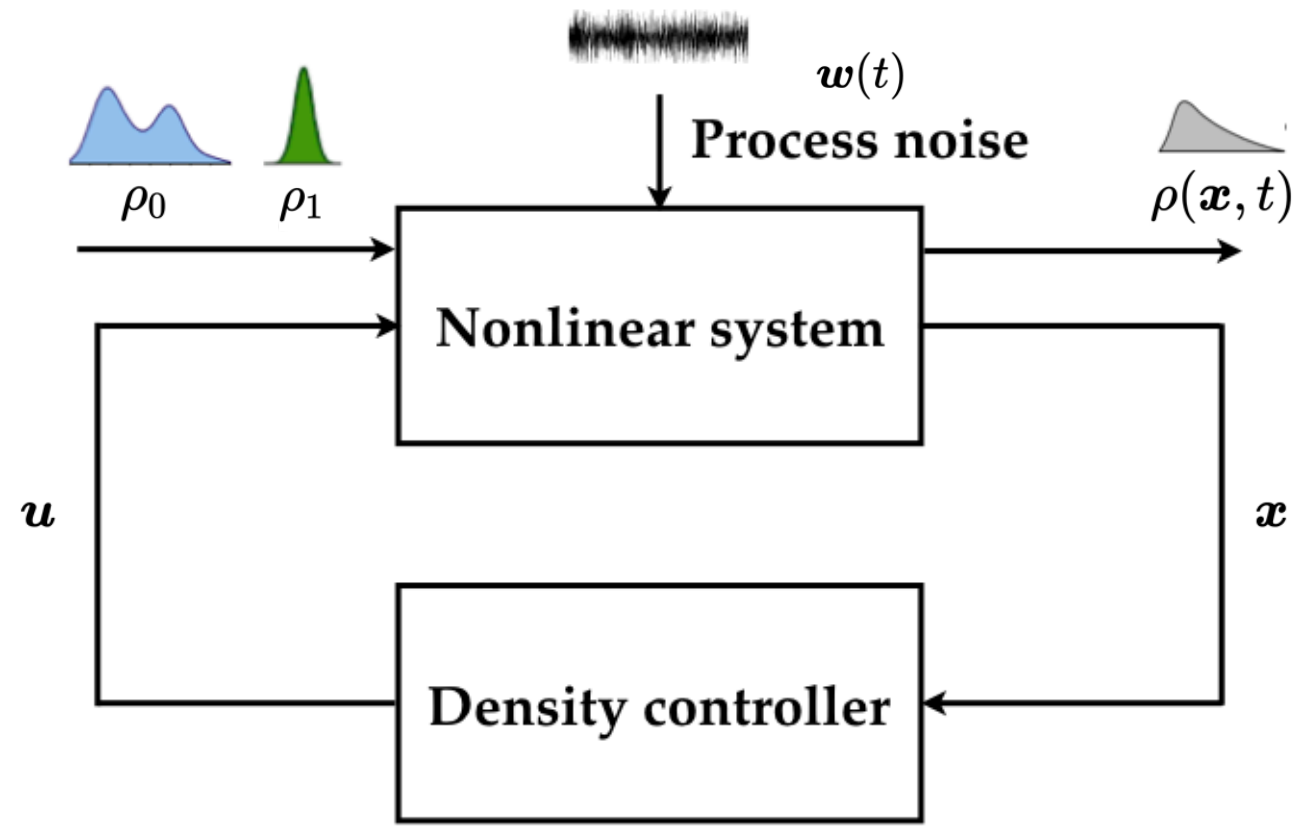


A.H. and T.T. Georgiou, Proximal Recursion for the Wonham Filter, *CDC 2019*.

Solving density control as generalized gradient flow

State Feedback Density Steering

Steer joint state PDF via feedback control over finite time horizon



Common scenario: $\boldsymbol{G} \equiv \boldsymbol{B}$

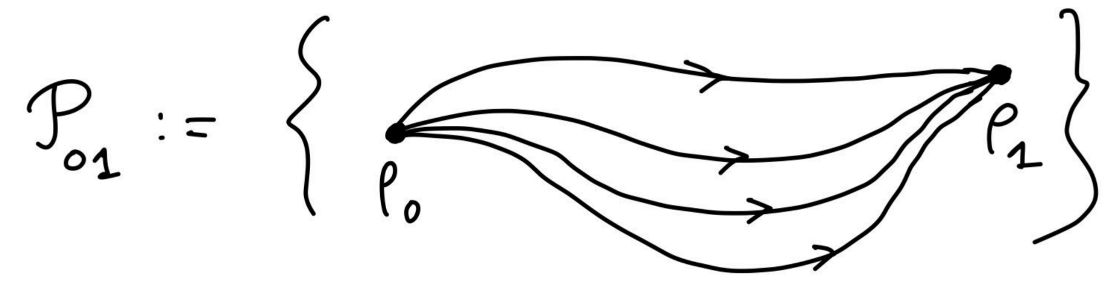
$$\underset{u \in \mathcal{U}}{\text{minimize}} \quad \mathbb{E}\left[\int_0^1 \left(\frac{1}{2}\|\boldsymbol{u}(t, \boldsymbol{x}_t^u)\|_2^2 + q(t, \boldsymbol{x}_t^u)\right) \, \mathrm{d}t\right]$$

subject to

$$\mathrm{d}\boldsymbol{x}_t^u = \{\boldsymbol{f}(t, \boldsymbol{x}_t^u) + \boldsymbol{B}(t, \boldsymbol{x}_t^u)\boldsymbol{u}\}\mathrm{d}t + \sqrt{2}\boldsymbol{G}(t, \boldsymbol{x}_t^u)\mathrm{d}\boldsymbol{w}_t$$

$$\boldsymbol{x}_0^u := \boldsymbol{x}_t^u(t=0) \sim \rho_0, \quad \boldsymbol{x}_1^u := \boldsymbol{x}_t^u(t=1) \sim \rho_1$$

Optimal Control Problem over PDFs



Diffusion tensor: $\boldsymbol{D} := \boldsymbol{G}\boldsymbol{G}^\top$

Hessian operator w.r.t. state: Hess

$$\inf_{(\rho,\boldsymbol{u})\in\mathcal{P}_{01}\times\mathcal{U}} \quad \int_{\mathbb{R}^n}\int_0^1 \left(\frac{1}{2}\|\boldsymbol{u}(t,\boldsymbol{x}_t^{\boldsymbol{u}})\|_2^2 + q(t,\boldsymbol{x}_t^{\boldsymbol{u}})\right)\rho(t,\boldsymbol{x}_t^{\boldsymbol{u}})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}_t^{\boldsymbol{u}}$$

subject to

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\left(\left(\boldsymbol{f}+\boldsymbol{B}\boldsymbol{u}\right)\rho\right) = \langle \mathrm{Hess}, \boldsymbol{D}\rho\rangle$$

$$\rho(t=0,\boldsymbol{x}_0^u) = \rho_0, \quad \rho(t=1,\boldsymbol{x}_1^u) = \rho_1$$

Necessary Conditions of Optimality (Assuming $G \equiv B$)

Coupled nonlinear PDEs + linear boundary conditions

Controlled Fokker-Planck or Kolmogorov's forward PDE

$$\frac{\partial \rho^{\text{opt}}}{\partial t} + \nabla \cdot \left(\left(\boldsymbol{f} + \boldsymbol{D} \nabla \psi \right) \rho^{\text{opt}} \right) = \langle \text{Hess}, \boldsymbol{D} \rho \rangle$$

Hamilton-Jacobi-Bellman-like PDE

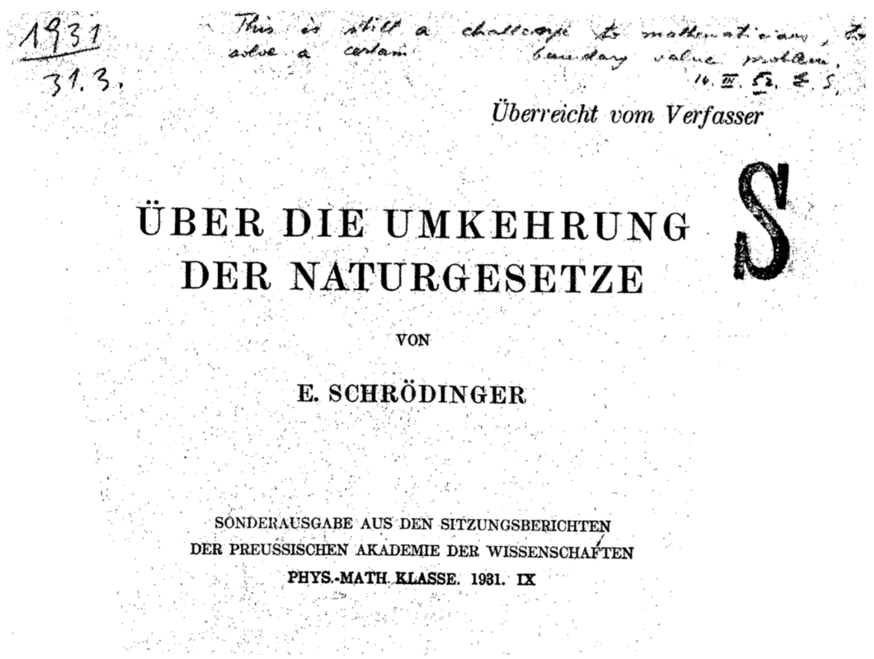
$$\frac{\partial \psi}{\partial t} + \langle \nabla \psi, \boldsymbol{f} \rangle + \langle \boldsymbol{D}, \text{Hess}(\psi) \rangle + \frac{1}{2} \langle \nabla \psi, \boldsymbol{D} \nabla \psi \rangle = q$$

Boundary conditions:

$$\rho^{\text{opt}}(\cdot, t=0) = \rho_0, \quad \rho^{\text{opt}}(\cdot, t=1) = \rho_1$$

Optimal control: $\boldsymbol{u}^{\text{opt}} = \boldsymbol{B}^{\top} \nabla \psi$

Feedback Synthesis via the Schrödinger System



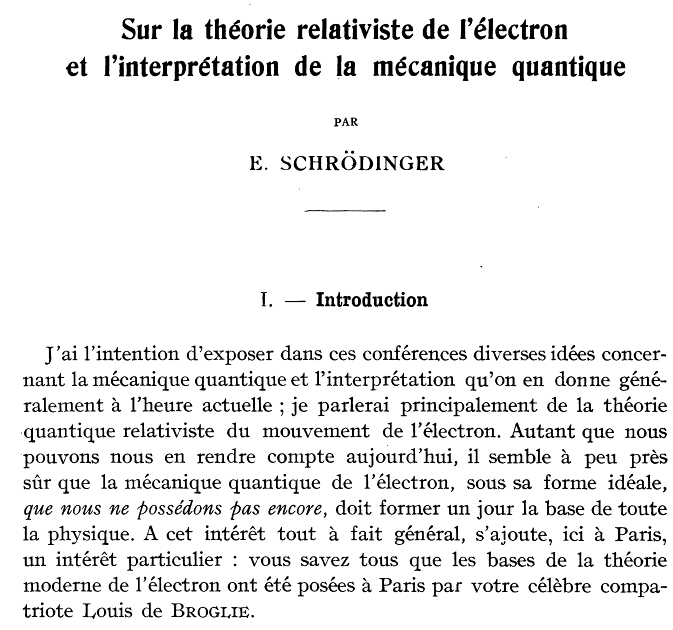
Überreicht vom Verfasser

ÜBER DIE UMKEHRUNG DER NATURGESETZE

VON

E. SCHRÖDINGER

SONDERAUSGABE AUS DEN SITZUNGSBERICHTEN DER PREUSSISCHEN AKADEMIE DER WISSENSCHAFTEN PHYS.-MATH. KLASSE. 1931. IX



Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique

PAR

E. SCHRÖDINGER

I. — **Introduction**

J'ai l'intention d'exposer dans ces conférences diverses idées concernant la mécanique quantique et l'interprétation qu'on en donne généralement à l'heure actuelle ; je parlerai principalement de la théorie quantique relativiste du mouvement de l'électron. Autant que nous pouvons nous en rendre compte aujourd'hui, il semble à peu près sûr que la mécanique quantique de l'électron, sous sa forme idéale, *que nous ne possédons pas encore*, doit former un jour la base de toute la physique. A cet intérêt tout à fait général, s'ajoute, ici à Paris, un intérêt particulier : vous savez tous que les bases de la théorie moderne de l'électron ont été posées à Paris par votre célèbre compatriote Louis de Broglie.

Hopf-Cole a.k.a. Fleming's logarithmic transform:

$$\left(\rho^{\text{opt}}, \psi\right) \mapsto \left(\widehat{\varphi}, \varphi\right)$$ — **Schrödinger factors**

$$\widehat{\varphi}(x,t) = \rho^{\text{opt}}(x,t)\exp\left(-\psi\left(x,t\right)\right)$$

$$\varphi(x,t) = \exp\left(\psi\left(x,t\right)\right) \qquad \text{for all} \qquad (x,t) \in \mathbb{R}^n \times [0,1]$$

Feedback Synthesis via the Schrödinger System

2 coupled nonlinear PDEs $\longrightarrow$ boundary-coupled linear PDEs!!

Uncontrolled forward-backward Kolmogorov PDEs:

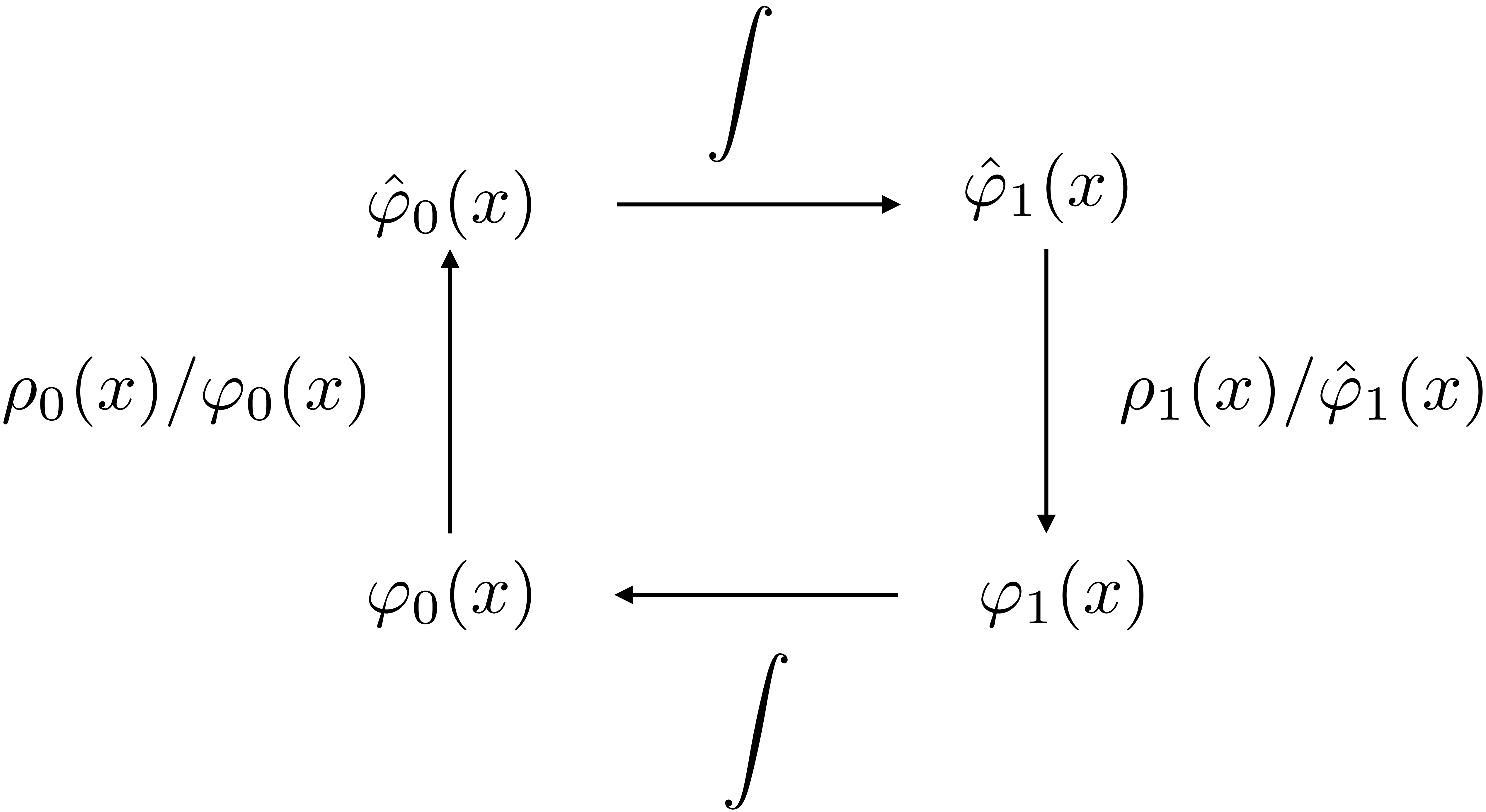
$$\frac{\partial \widehat{\varphi}}{\partial t} = -\nabla \cdot (\widehat{\varphi} \boldsymbol{f}) + \langle \text{Hess}, \boldsymbol{D} \widehat{\varphi} \rangle - q \widehat{\varphi}, \qquad \widehat{\varphi}_0 \varphi_0 = \rho_0,$$

$$\frac{\partial \varphi}{\partial t} = -\langle \nabla \varphi, \boldsymbol{f} \rangle - \langle \text{Hess}(\varphi), \boldsymbol{D} \rangle + q \varphi, \qquad \widehat{\varphi}_1 \varphi_1 = \rho_1,$$

Optimal controlled joint state PDF: $\rho^{\text{opt}}(\boldsymbol{x}, t) = \widehat{\varphi}(\boldsymbol{x}, t) \varphi(\boldsymbol{x}, t)$

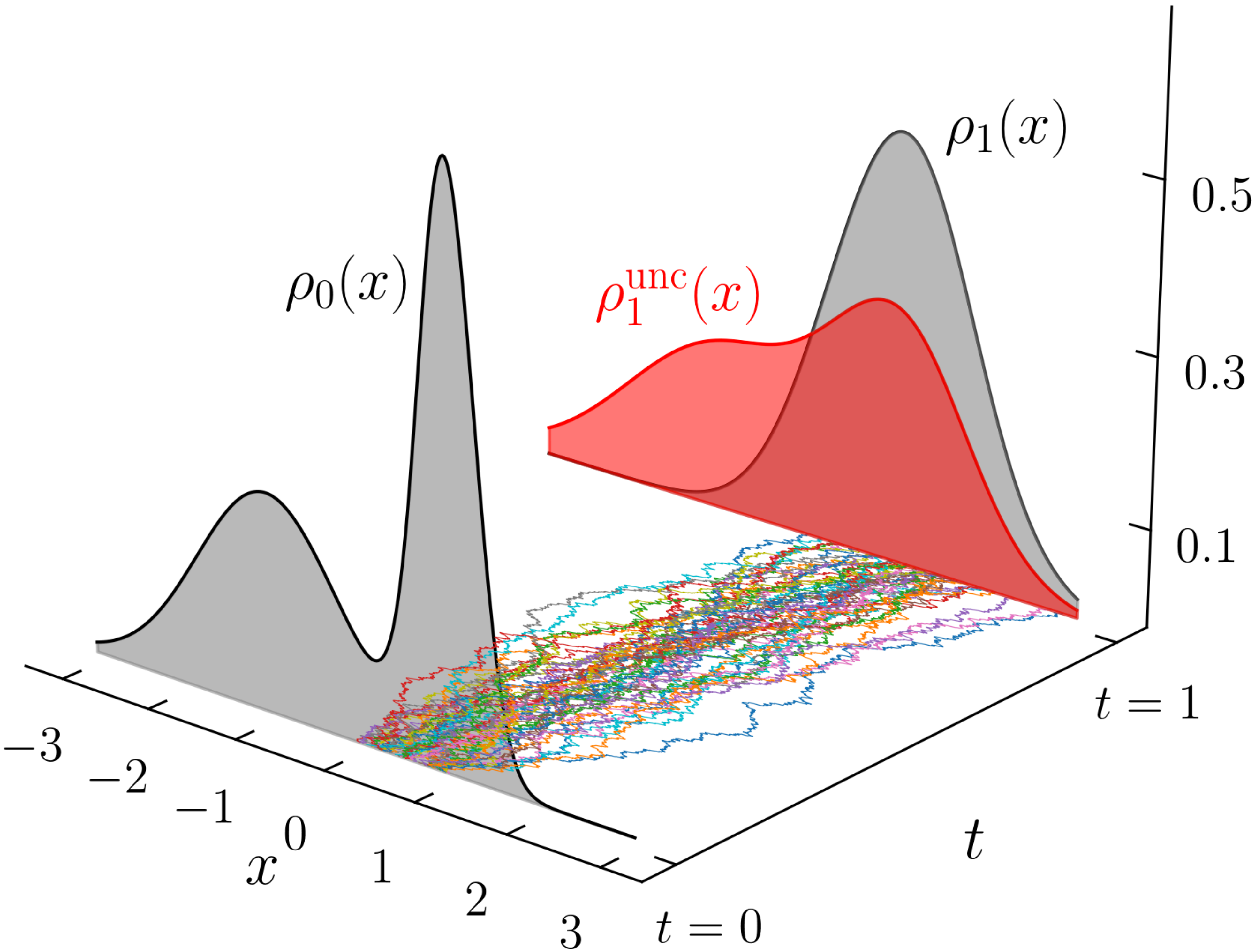
Optimal control: $\boldsymbol{u}^{\text{opt}}(\boldsymbol{x}, t) = 2\boldsymbol{B}^{\top} \nabla_{\boldsymbol{x}} \log \varphi(\boldsymbol{x}, t)$

Fixed Point Recursion over $(\hat{\varphi}_0, \varphi_1)$



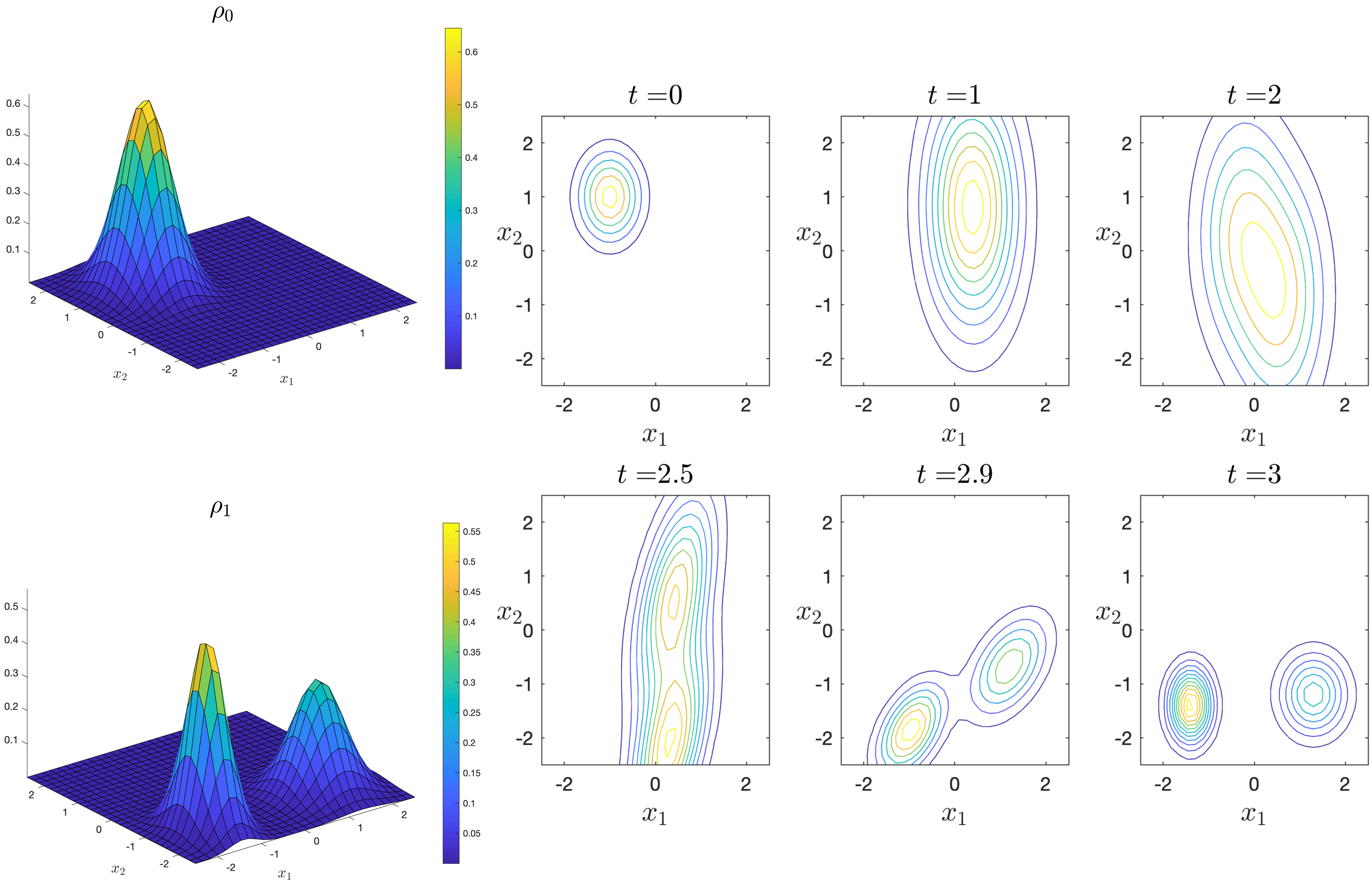
This recursion is contractive in the Hilbert metric!!

Feedback Density Control: $f \equiv 0, B = G \equiv I, q \equiv 0$



Zero prior dynamics

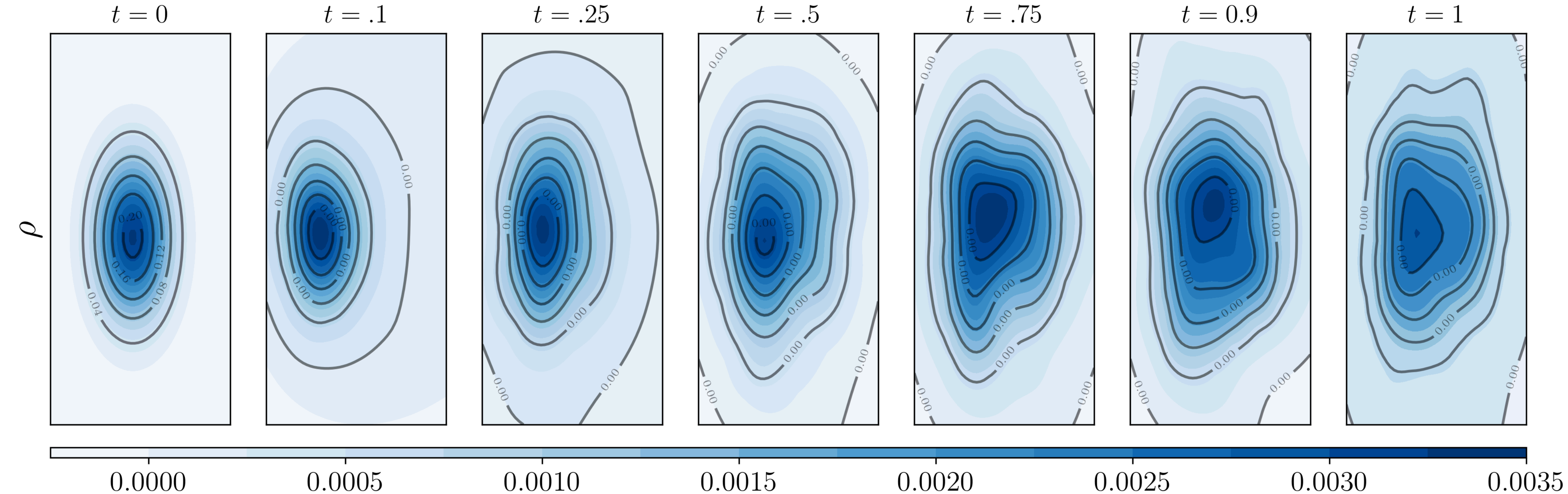
Feedback Density Control: $f \equiv Ax, B = G, q \equiv 0$



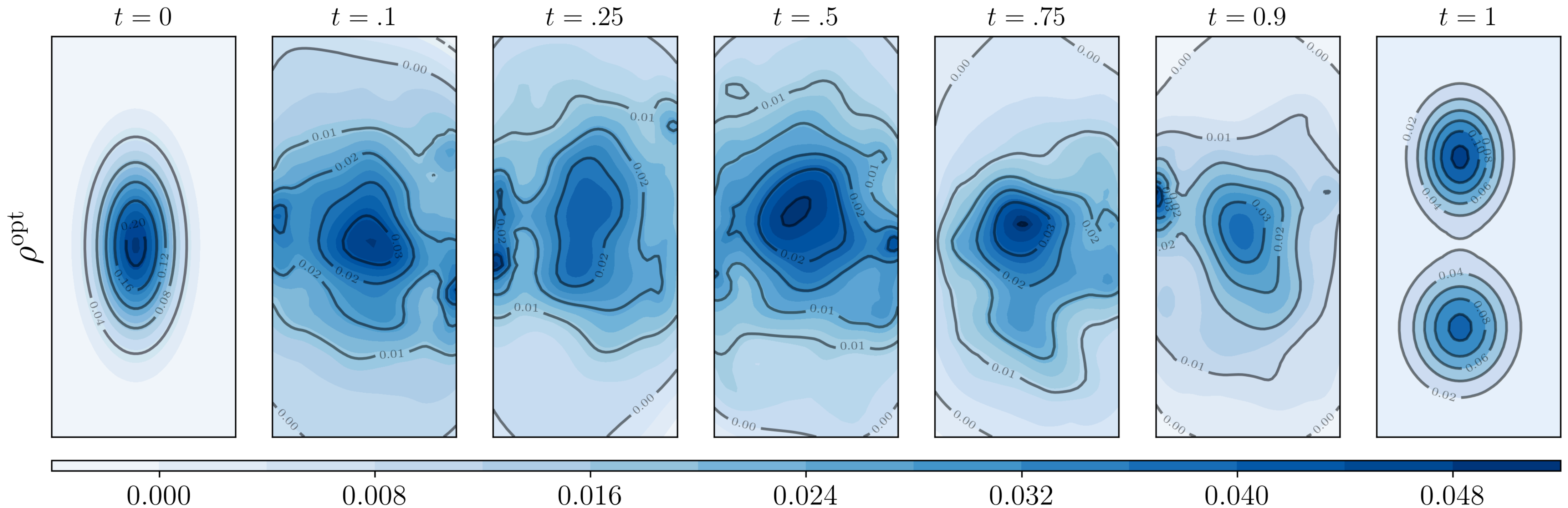
Linear prior dynamics

Feedback Density Control: Nonlinear Grad. Drift

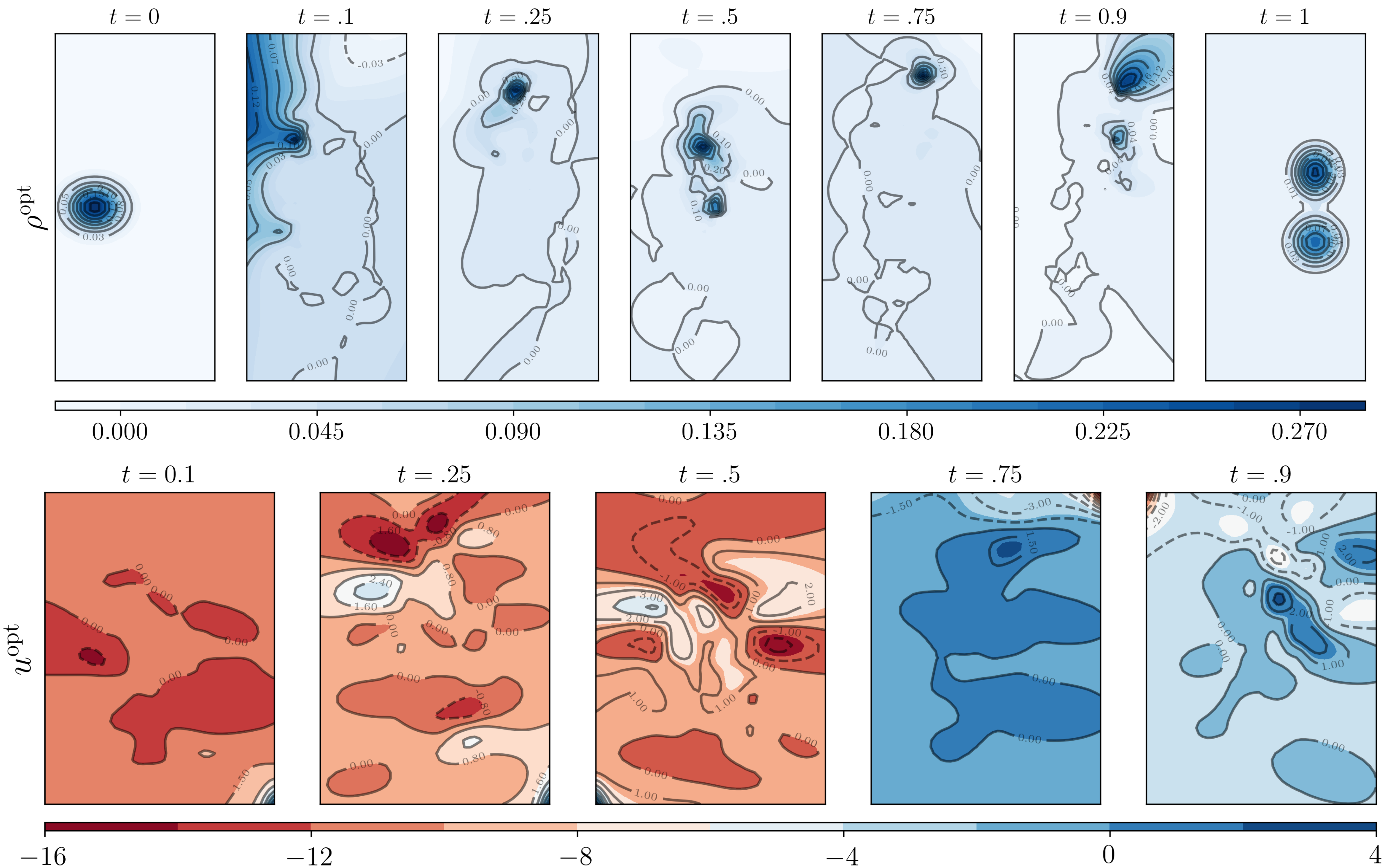
Uncontrolled joint PDF evolution:



Optimal controlled joint PDF evolution:

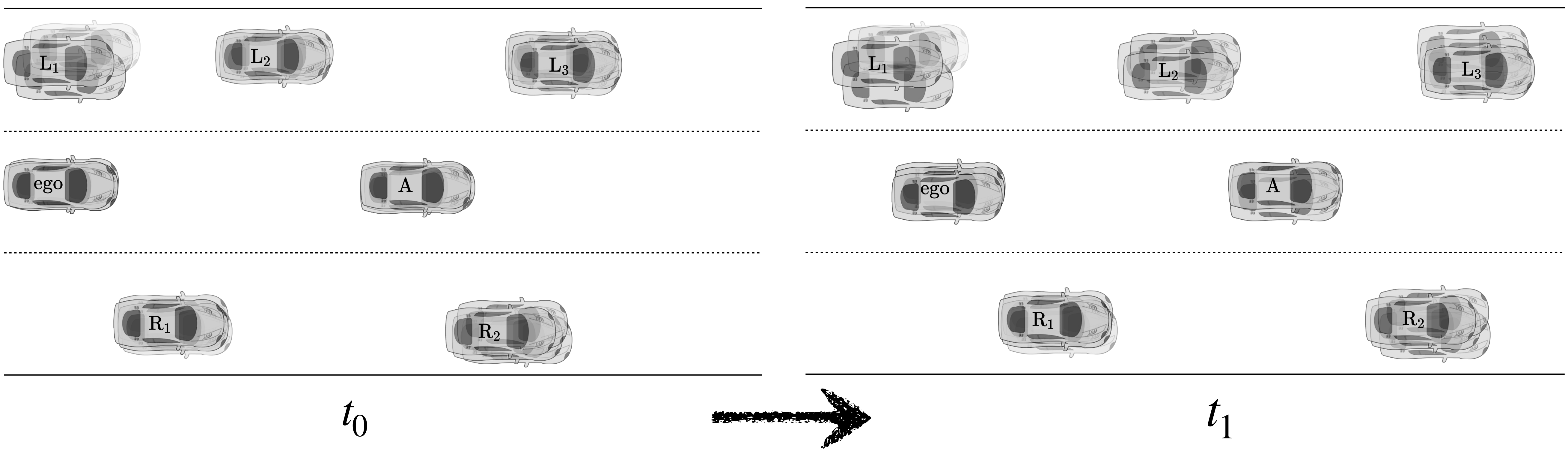


Feedback Density Control: Mixed Conservative-Dissipative Drift

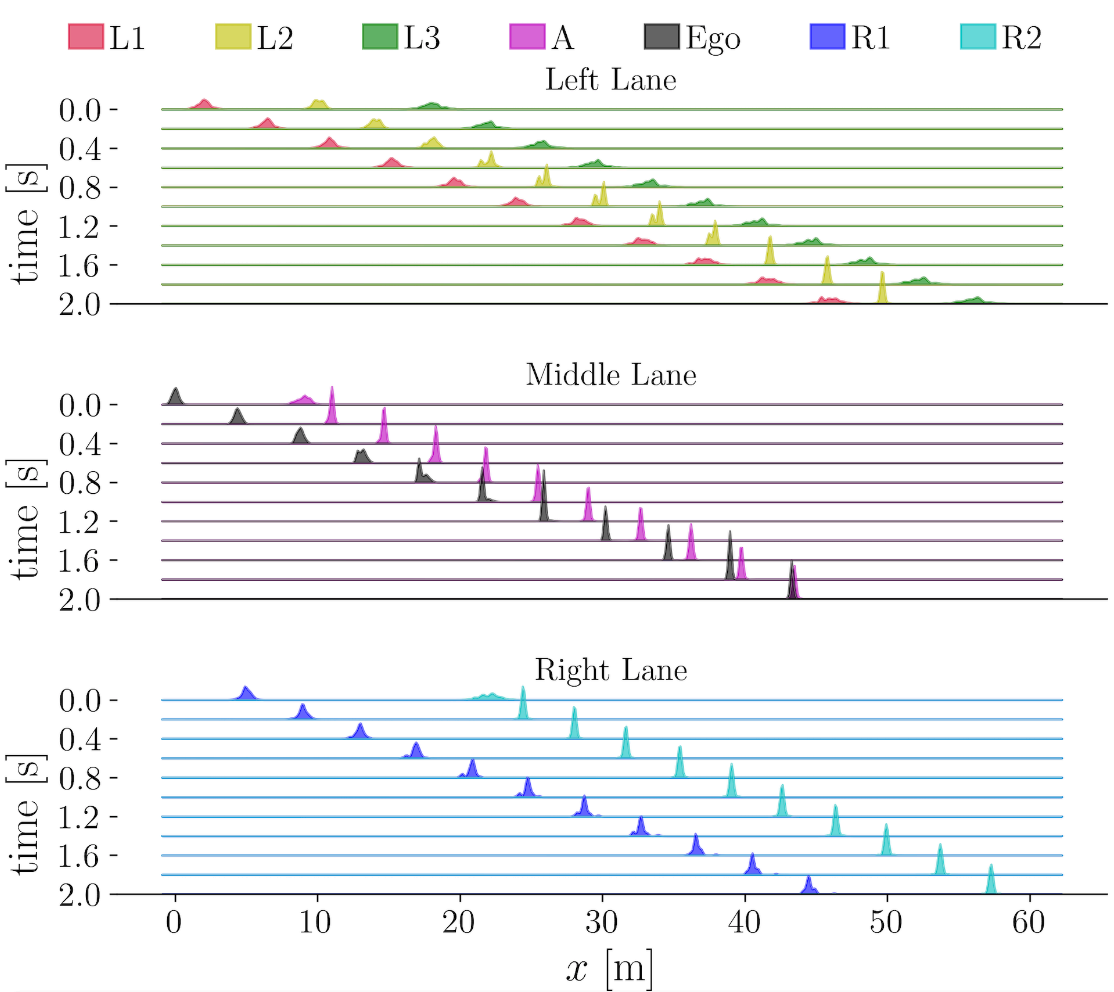


K.F. Caluya and A.H., Wasserstein proximal algorithms for the Schrödinger bridge problem: density control with nonlinear drift, *IEEE TAC* 2021.

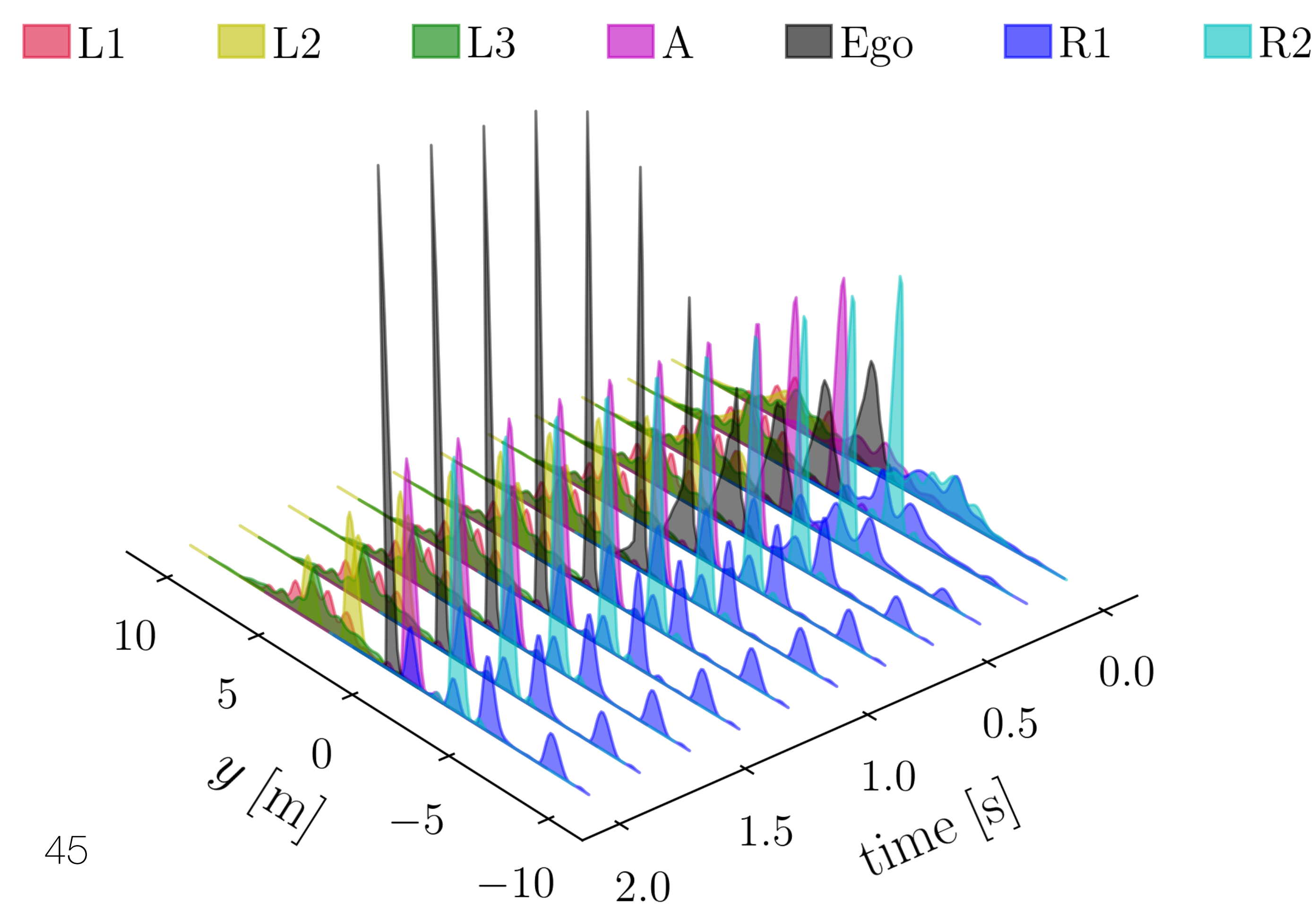
Application: Multi-lane Automated Driving



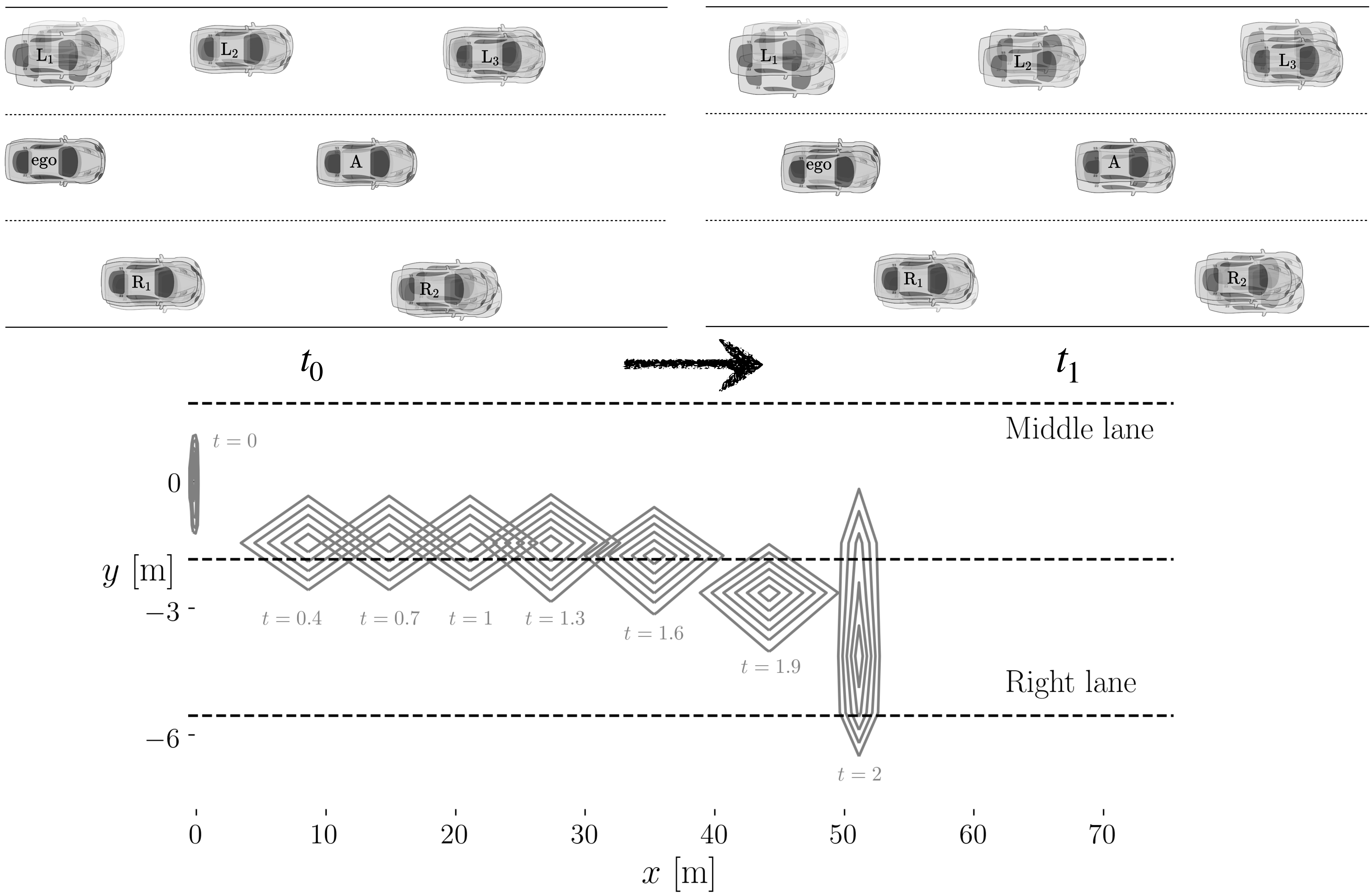
x marginals



y marginals



Application: Multi-lane Automated Driving



Extensions and Applications

Multi-lane automated driving

S. Haddad, A.H., and B. Singh, Density-based stochastic reachability computation for occupancy prediction in automated driving, *IEEE Transactions on Control Systems Technology*, 2022.

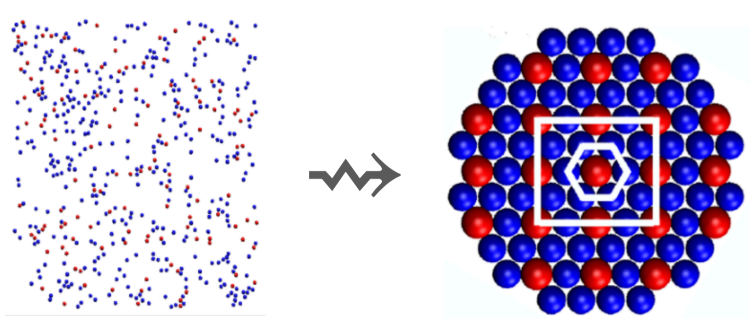
S. Haddad, K.F. Caluya, A.H., and B. Singh, Prediction and optimal feedback steering of probability density functions for safe automated driving, *IEEE Control Systems Letters*, 2021.

Hard path constraints, non-standard nonlinear drift and diffusion

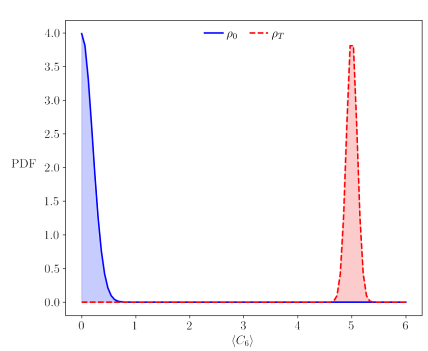
K.F. Caluya, and A.H., Reflected Schrödinger bridge: density control with path constraints, *ACC* 2021.

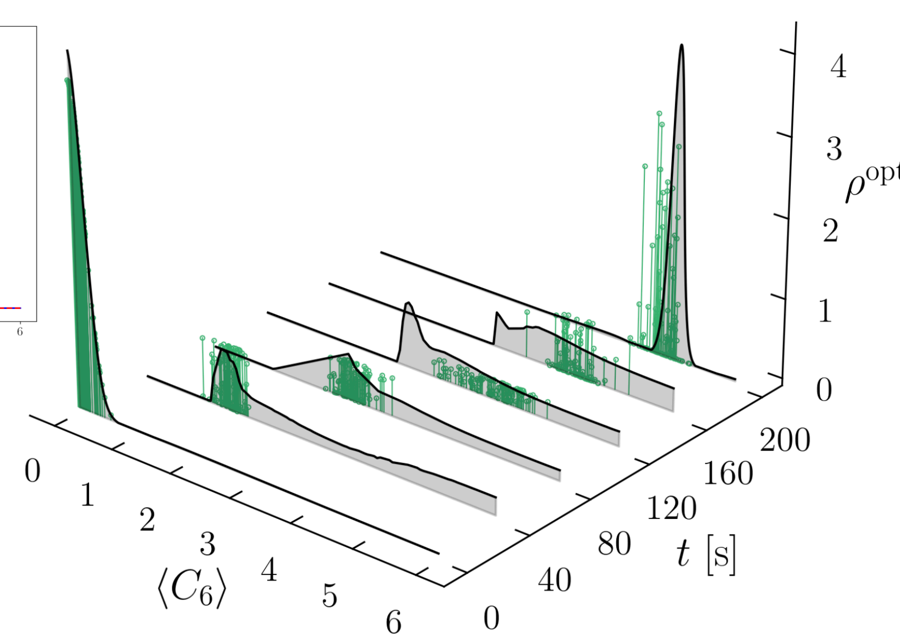
I. Nodozi, and A.H., Schrödinger meets Kuramato via Feynman-Kac: minimum effort distribution steering for noisy nonuniform Kuramoto oscillators, *CDC 2022*.

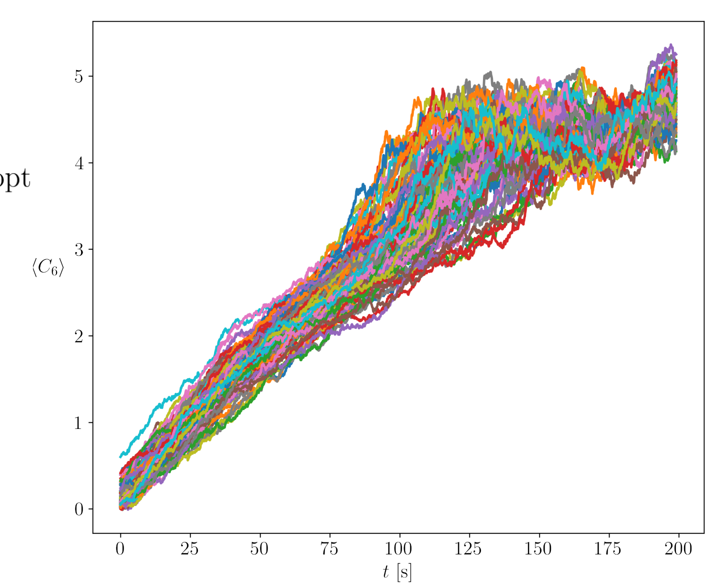
Controlled colloidal self-assembly



Dispersed particles **Ordered structure**



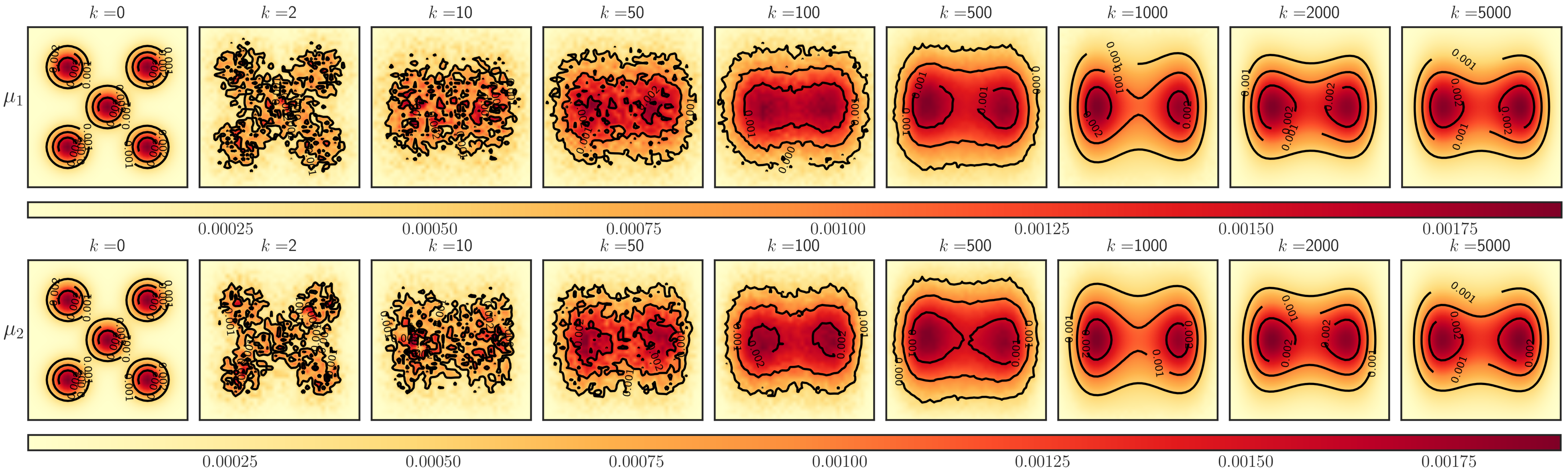




arXiv: 2208.09182, 2022

Ongoing Efforts

Distributed algorithms

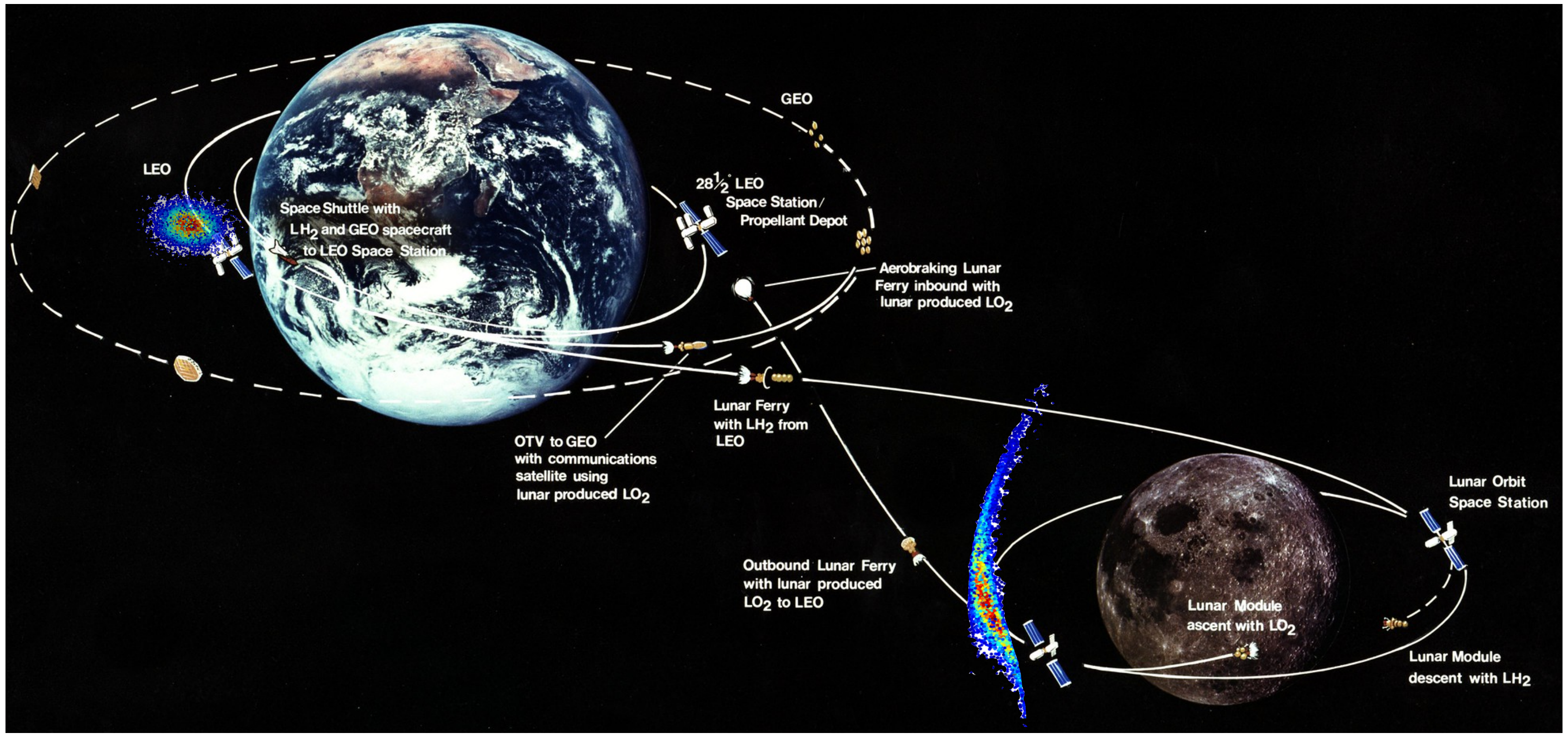


Stochastic guidance of spacecraft

Provably correct guidance law synthesis

Nonparametric statistical computation

Orbit transfer case studies



Take Home Message

Thank You

Support: NSF 1923278, 2112755, 2111688 Ford CITRIS PEOPLE AND ROBOTS