Wasserstein Gradient Flow for Stochastic Prediction, Filtering and Control: Theory and Algorithms

Abhishek Halder

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

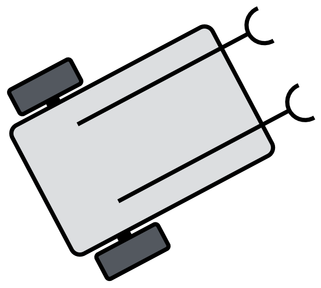
Joint work with Kenneth F. Caluya (UC Santa Cruz)
and Tryphon T. Georgiou (UC Irvine)

Overarching theme

Systems-control theory for densities

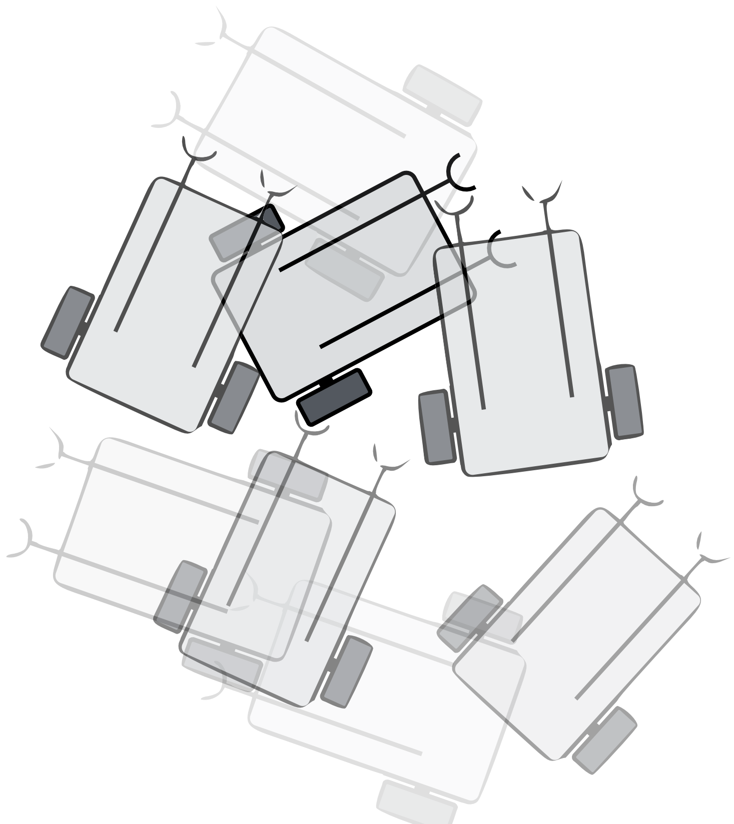
What is density?

Probability Density Fn.



$$x(t) \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

Probability Density Fn.

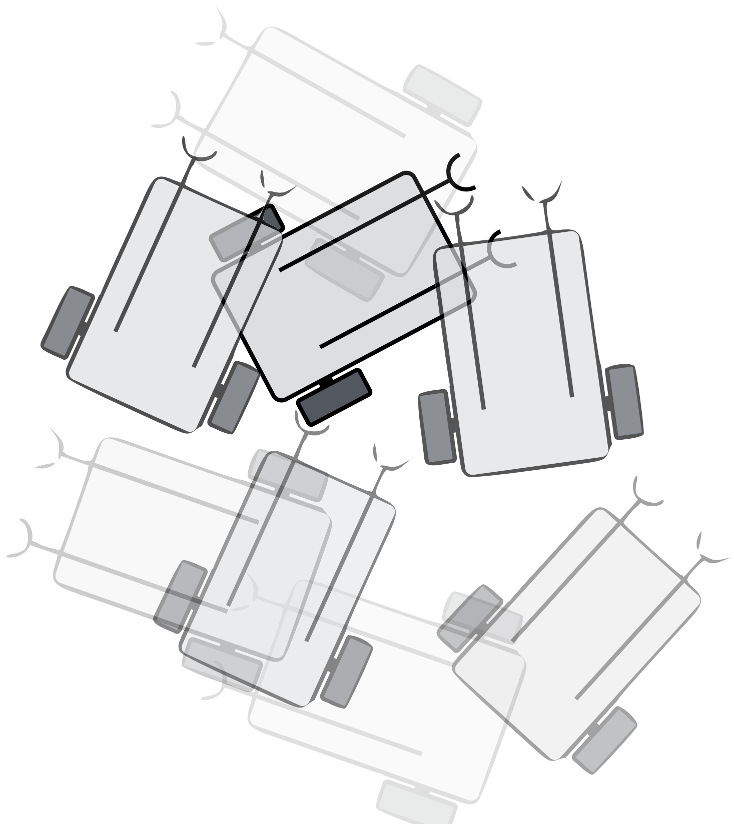


$$x(t) \in \mathcal{X} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

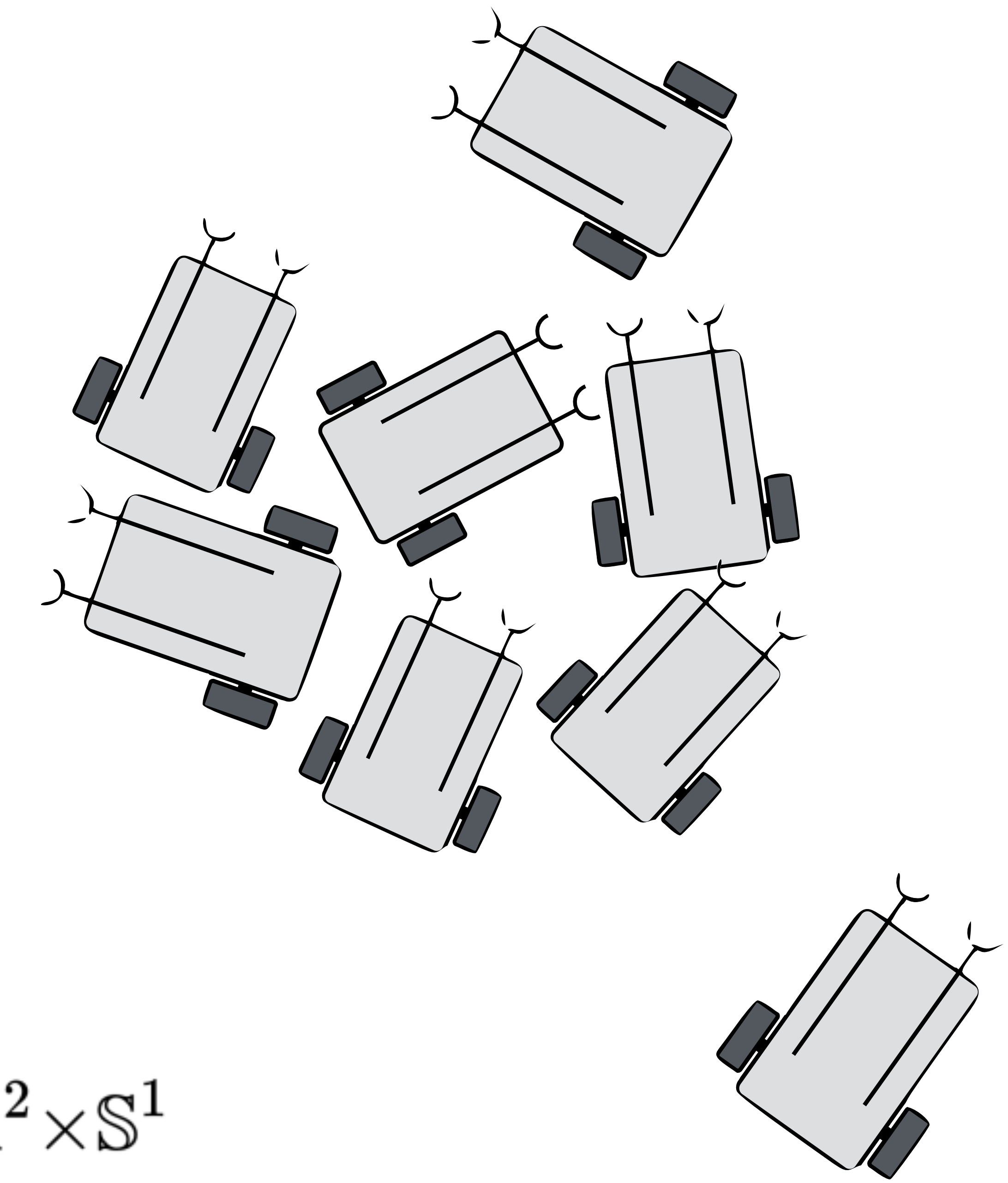
$$\rho(x, t) : \mathcal{X} \times [0, \infty) \mapsto \mathbb{R}_{\geq 0}$$

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

Probability Density Fn.



Population Density Fn.



$$x(t) \in \boldsymbol{\mathcal{X}} \equiv \mathbb{R}^2 \times \mathbb{S}^1$$

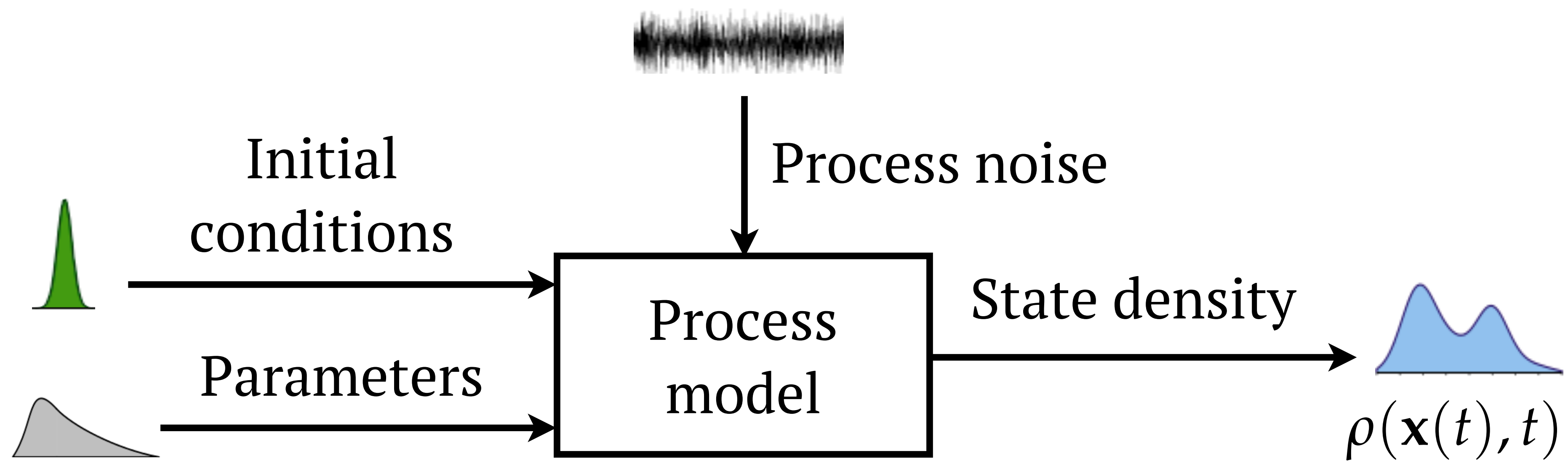
$$\rho(x,t) : \boldsymbol{\mathcal{X}} \times [0,\infty) \mapsto \mathbb{R}_{\geq 0}$$

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

Why care about densities?

Prediction Problem

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



Trajectory flow:

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

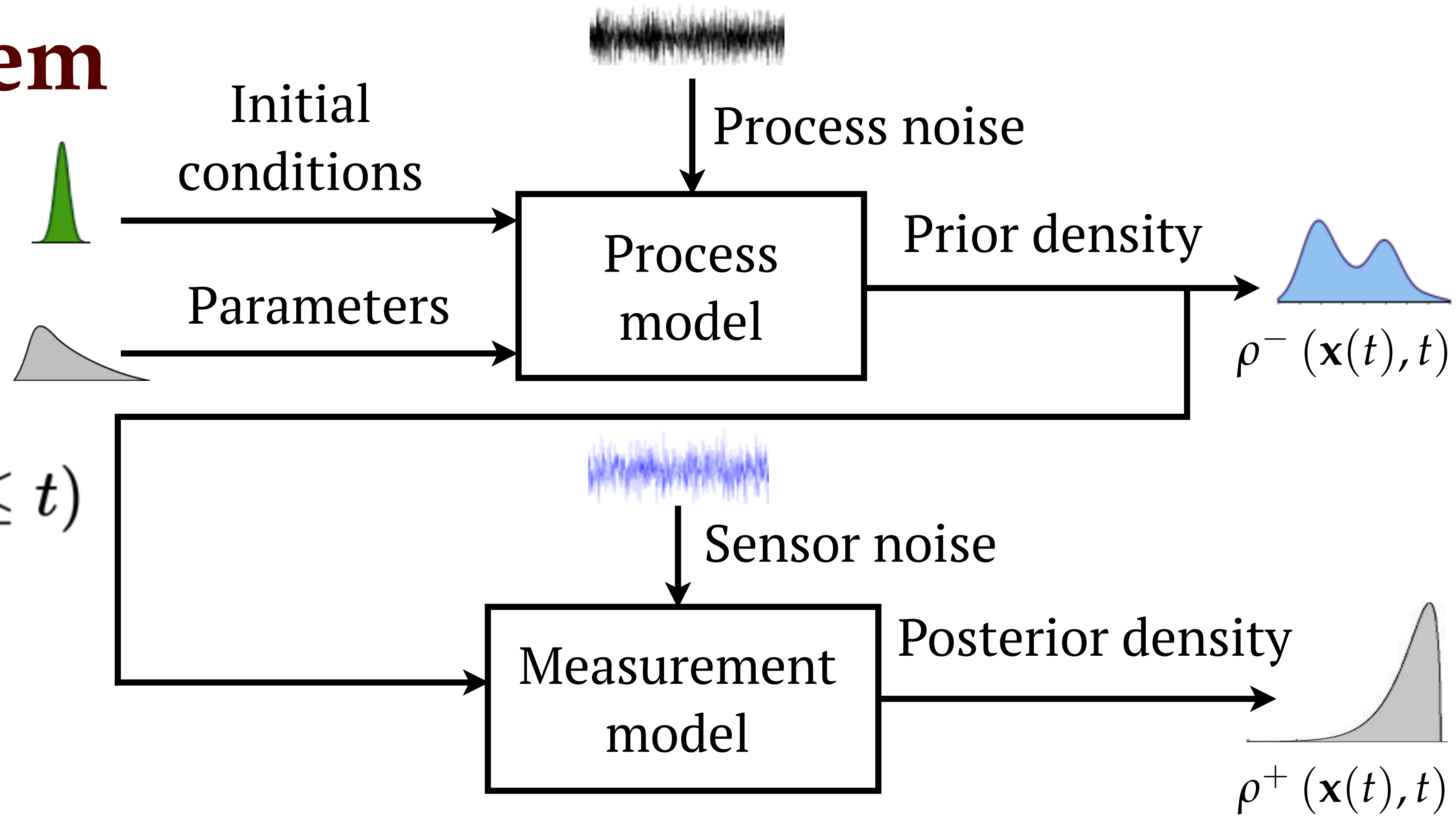
Density flow:

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

Filtering Problem

Compute conditional joint state PDF

$$\rho^+ \equiv \rho(x, t | z(s), 0 \leq s \leq t)$$



Trajectory flow:

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

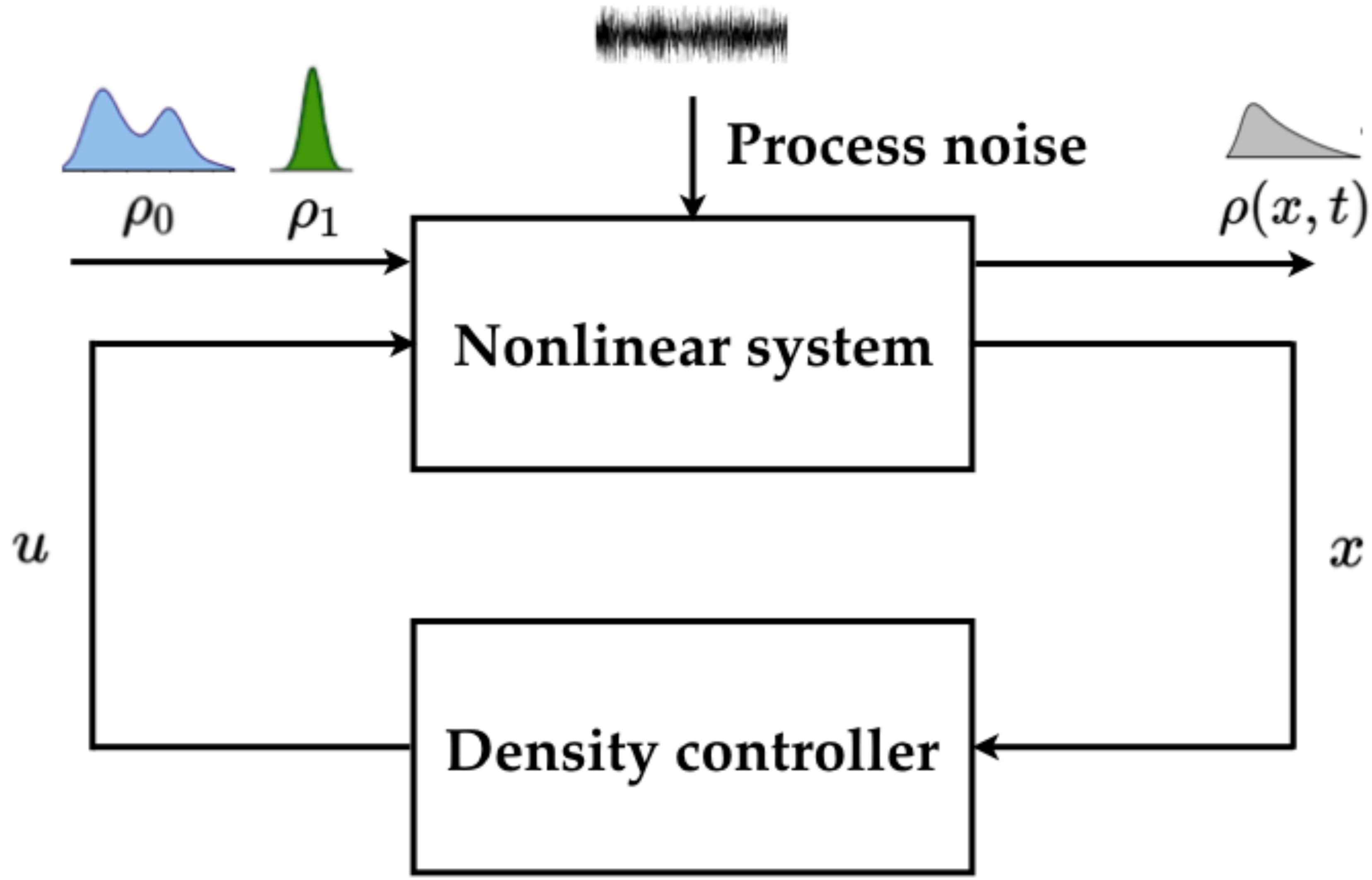
$$\mathrm{d}\mathbf{Z}(t) = \mathbf{h}(\mathbf{X}, t)\,\mathrm{d}t + \mathrm{d}\mathbf{v}(t), \qquad \mathrm{d}\mathbf{v}(t) \sim \mathcal{N}(0, \mathbf{R}\mathrm{d}t)$$

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



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

subject to

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

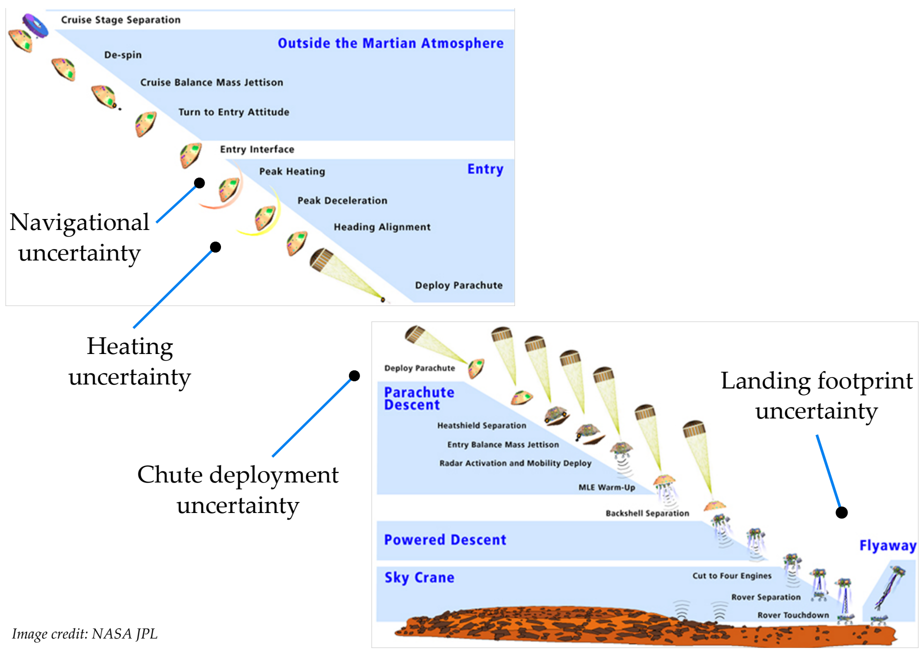
$$x(t=0) \sim \rho_0, \quad x(t=1) \sim \rho_1$$

PDFs in Mars Entry-Descent-Landing

Prediction Problem

Filtering Problem

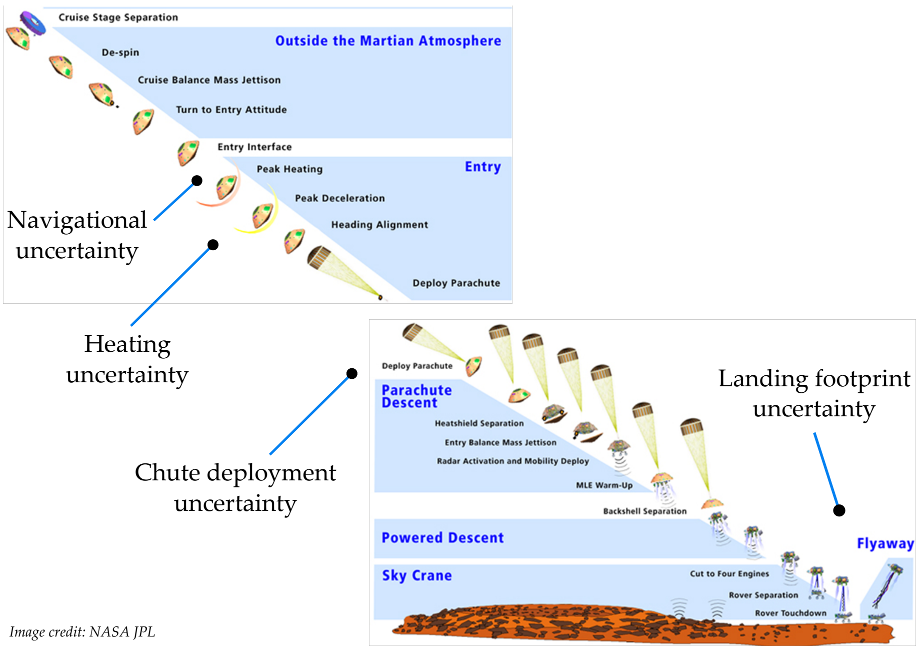
Control Problem



Predict heating rate uncertainty

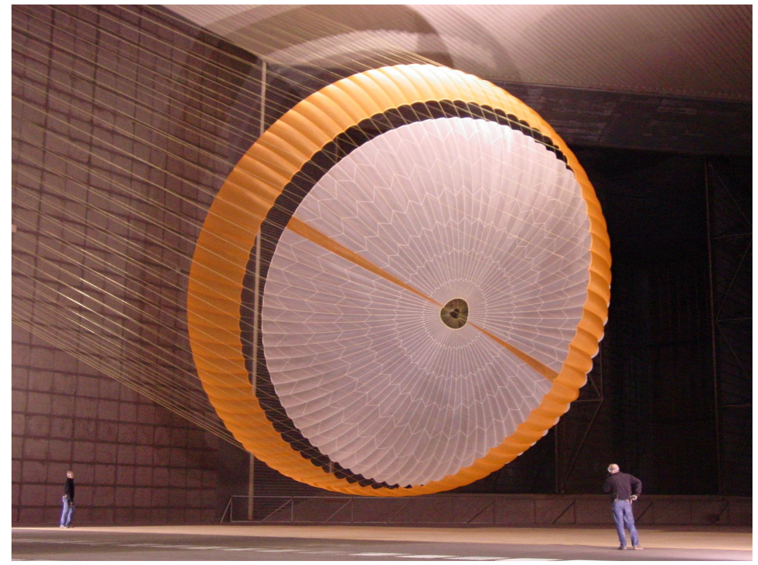
PDFs in Mars Entry-Descent-Landing

Prediction Problem

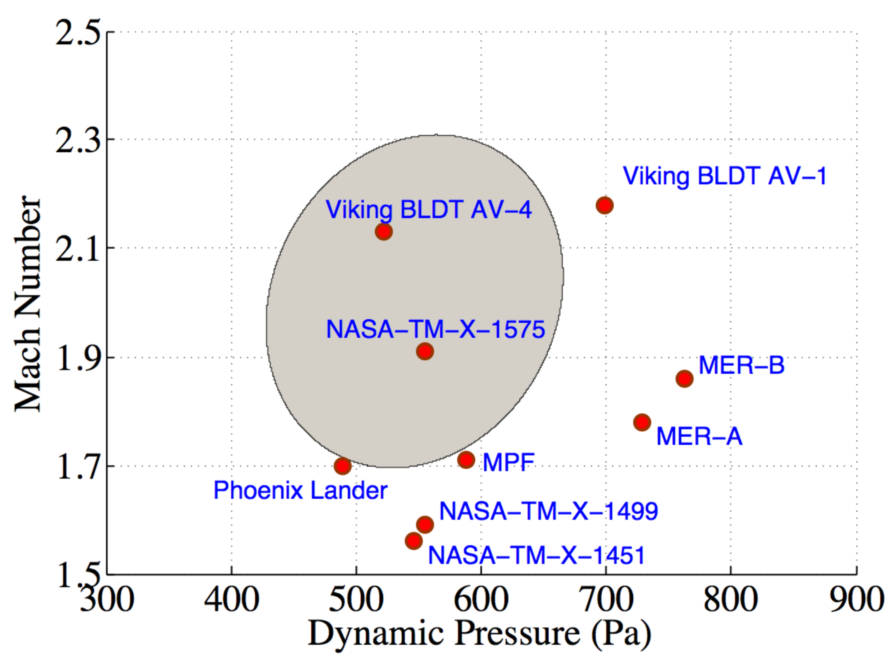


Predict heating rate uncertainty

Filtering Problem



Supersonic parachute

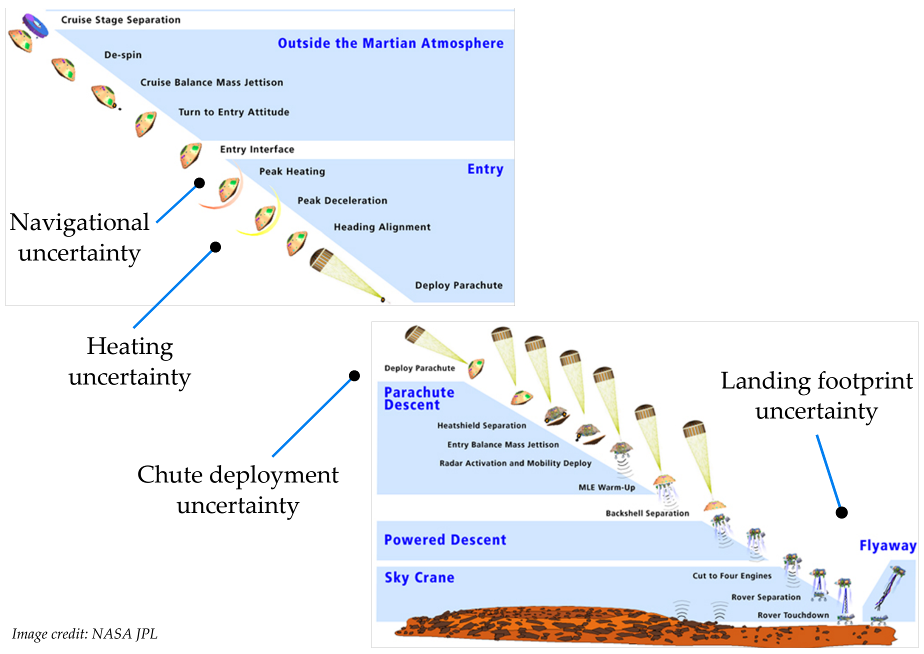


Estimate state to deploy parachute

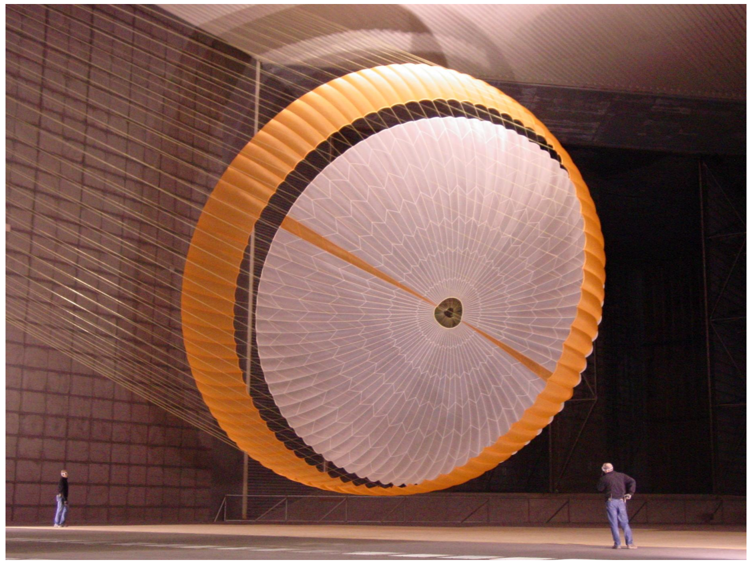
Control Problem

PDFs in Mars Entry-Descent-Landing

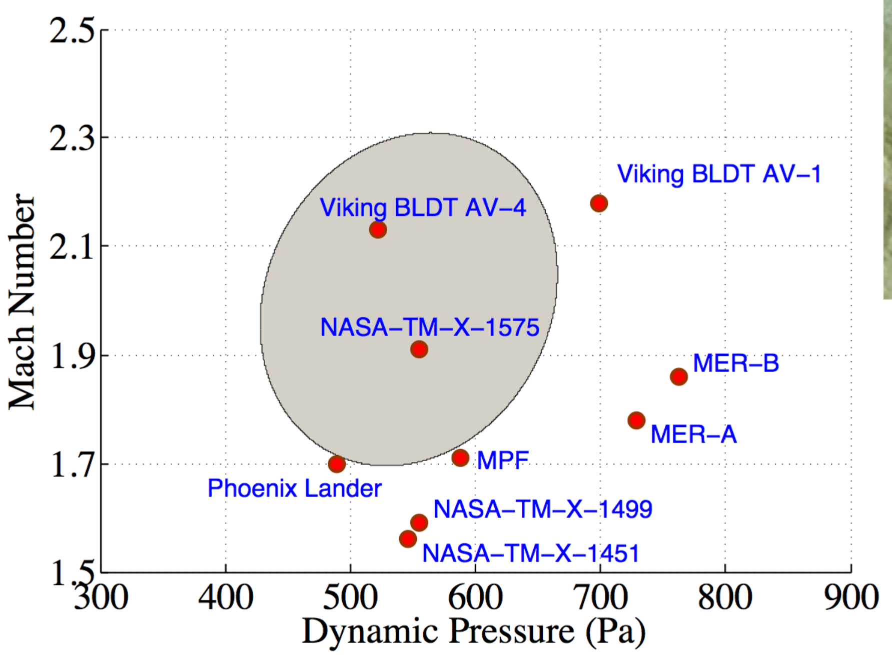
Prediction Problem



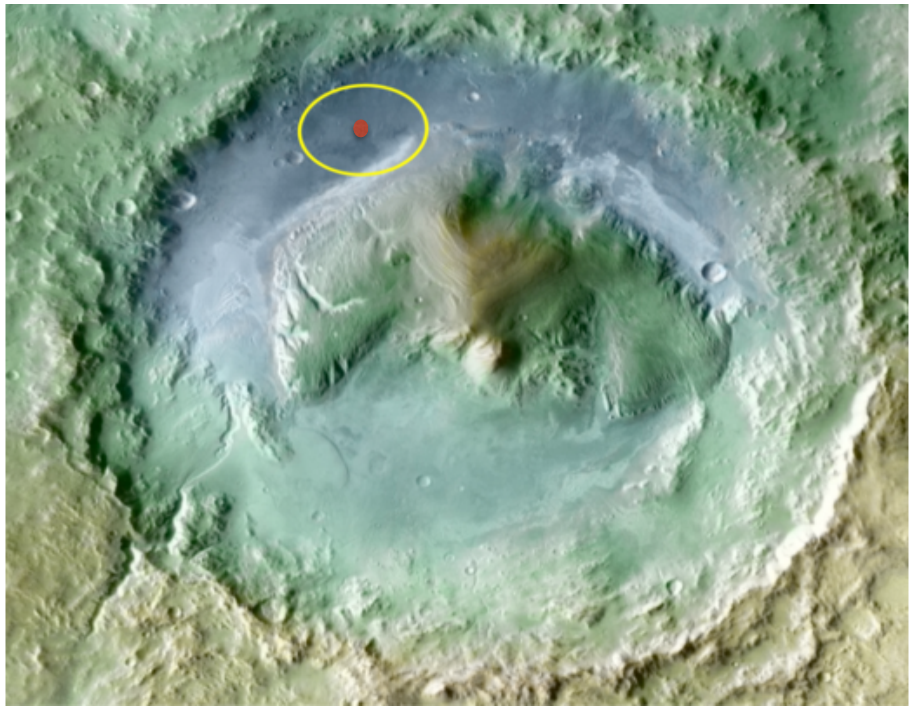
Filtering Problem



Supersonic parachute



Control Problem



Gale Crater (4.49S, 137.42E)

Predict heating rate uncertainty

Estimate state to deploy parachute

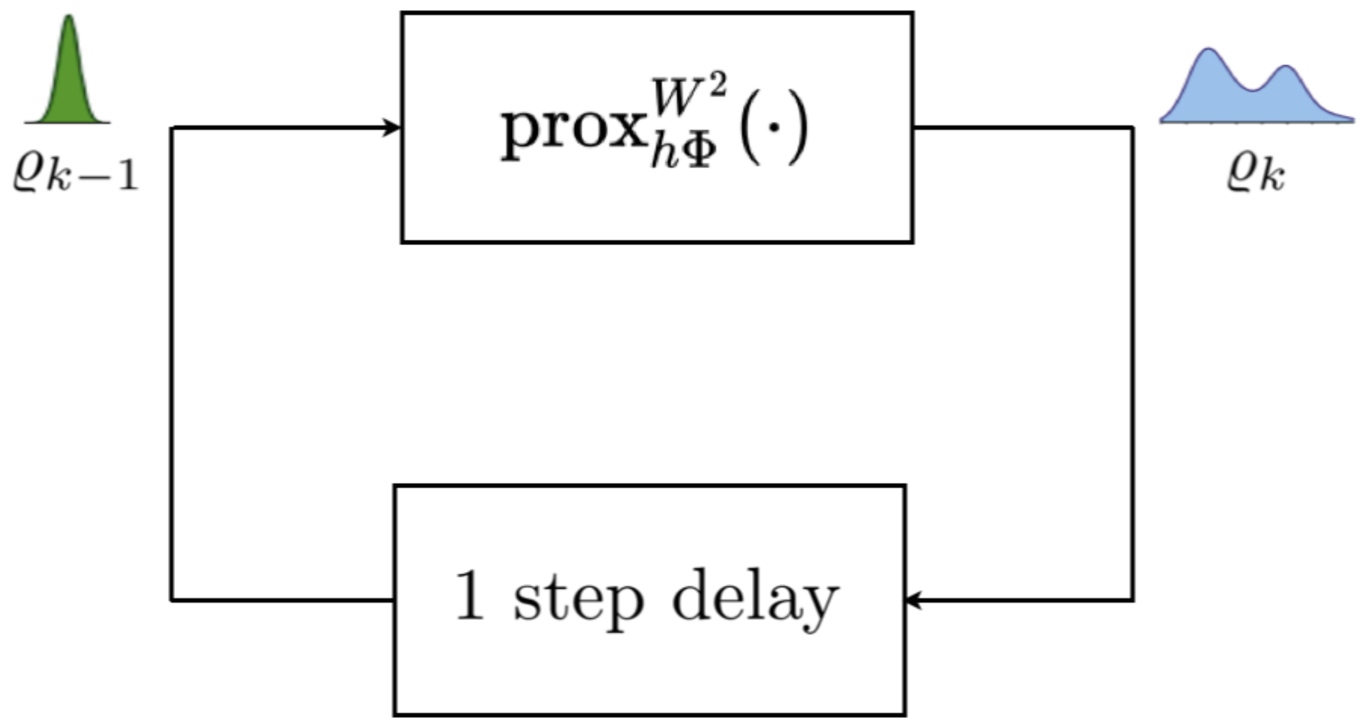
Steer state PDF to achieve desired landing footprint accuracy

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:



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

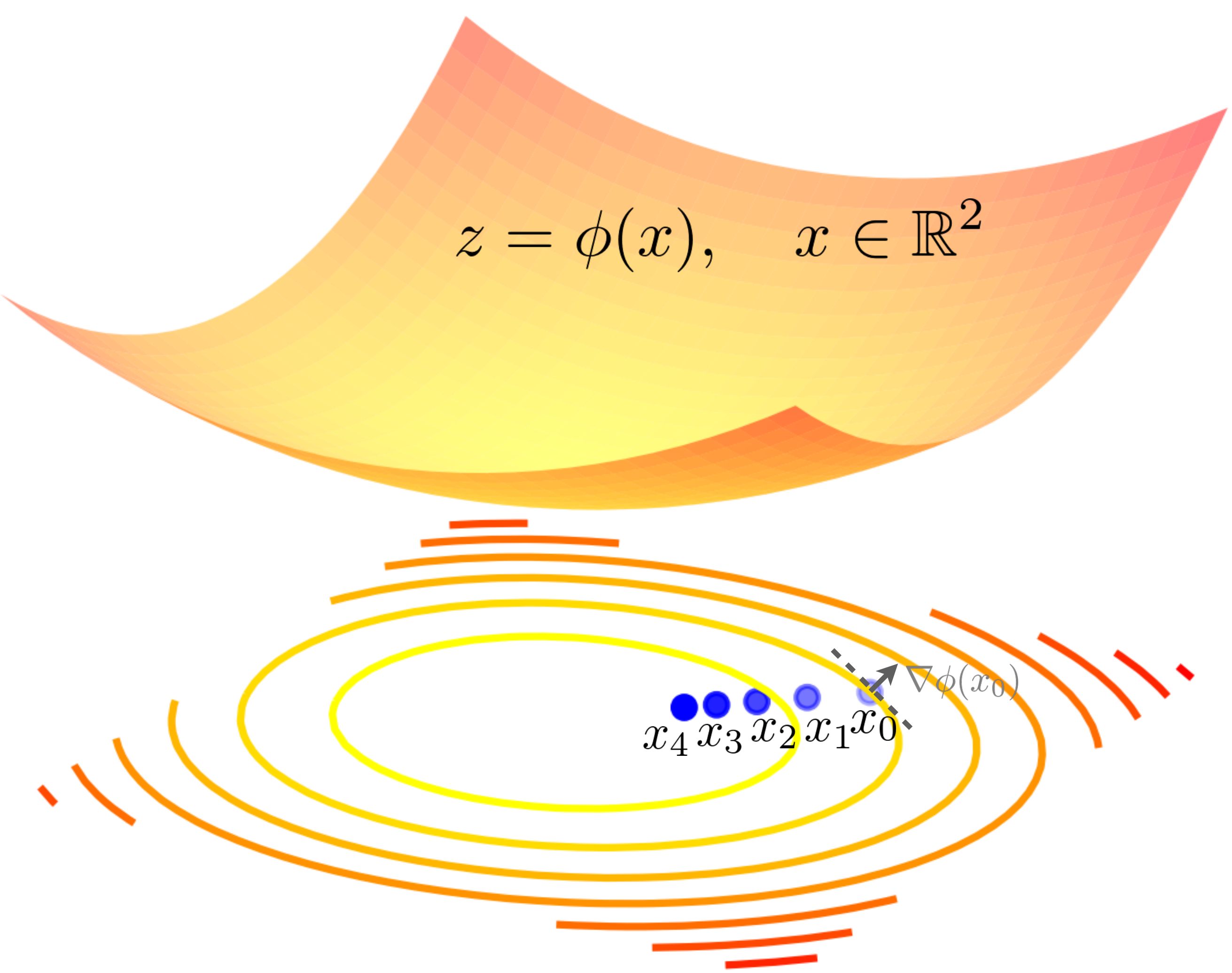
Free energy functional: $\Phi(\varrho) := \displaystyle\int_{\mathcal{X}} \psi\varrho\, \mathrm{d}x + \beta^{-1} \int_{\mathcal{X}} \varrho \log \varrho\, \mathrm{d}x$

Geometric Meaning of Gradient Flow

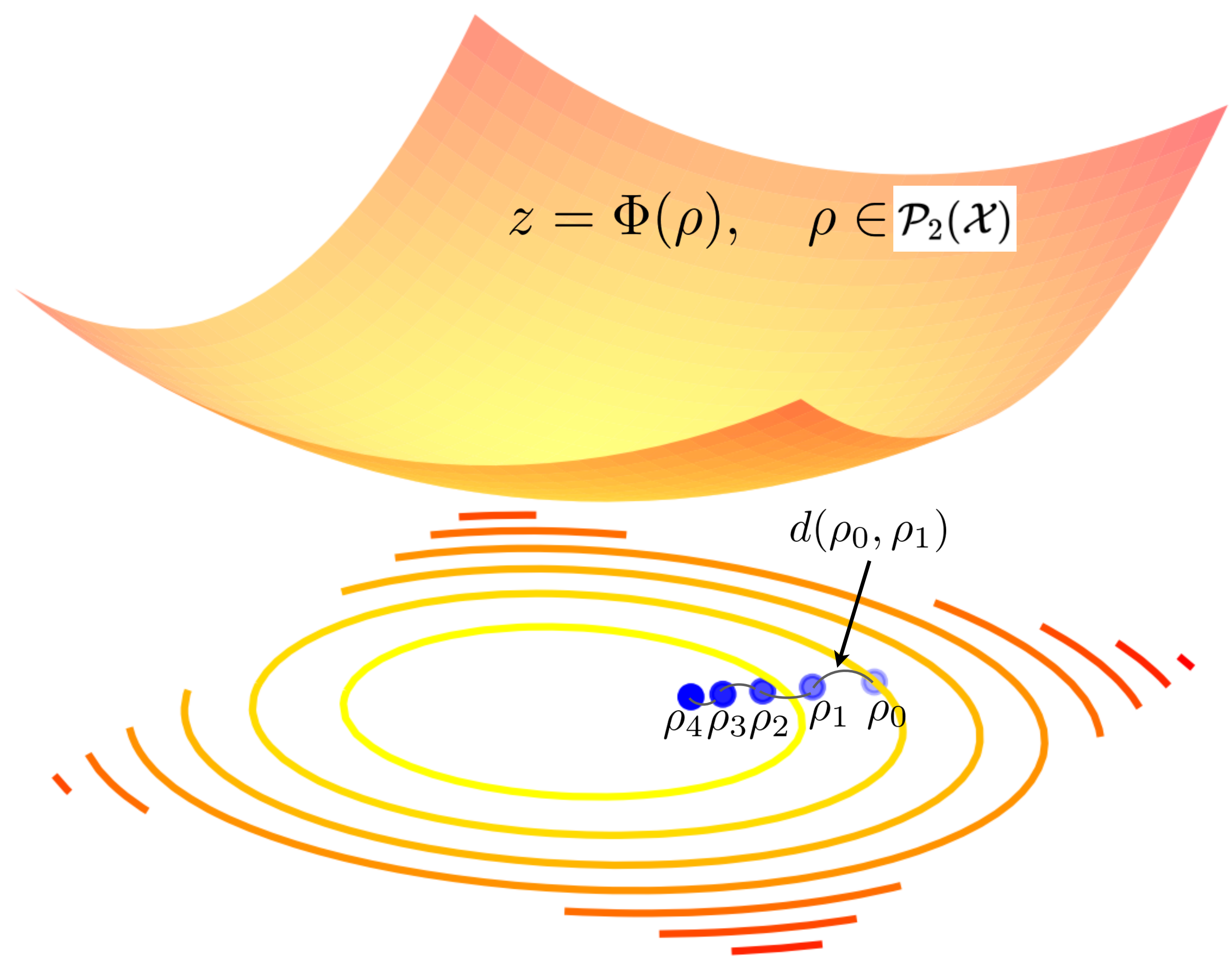
Gradient Flow in $\mathcal{X}$	Gradient Flow in $\mathcal{P}_2(\mathcal{X})$
$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$	$\frac{\partial\rho}{\partial t} = -\nabla^W\Phi(\rho), \quad \rho(\boldsymbol{x},0) = \rho_0$
Recursion: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

Geometric Meaning of Gradient Flow

Gradient Flow in $\mathcal{X}$

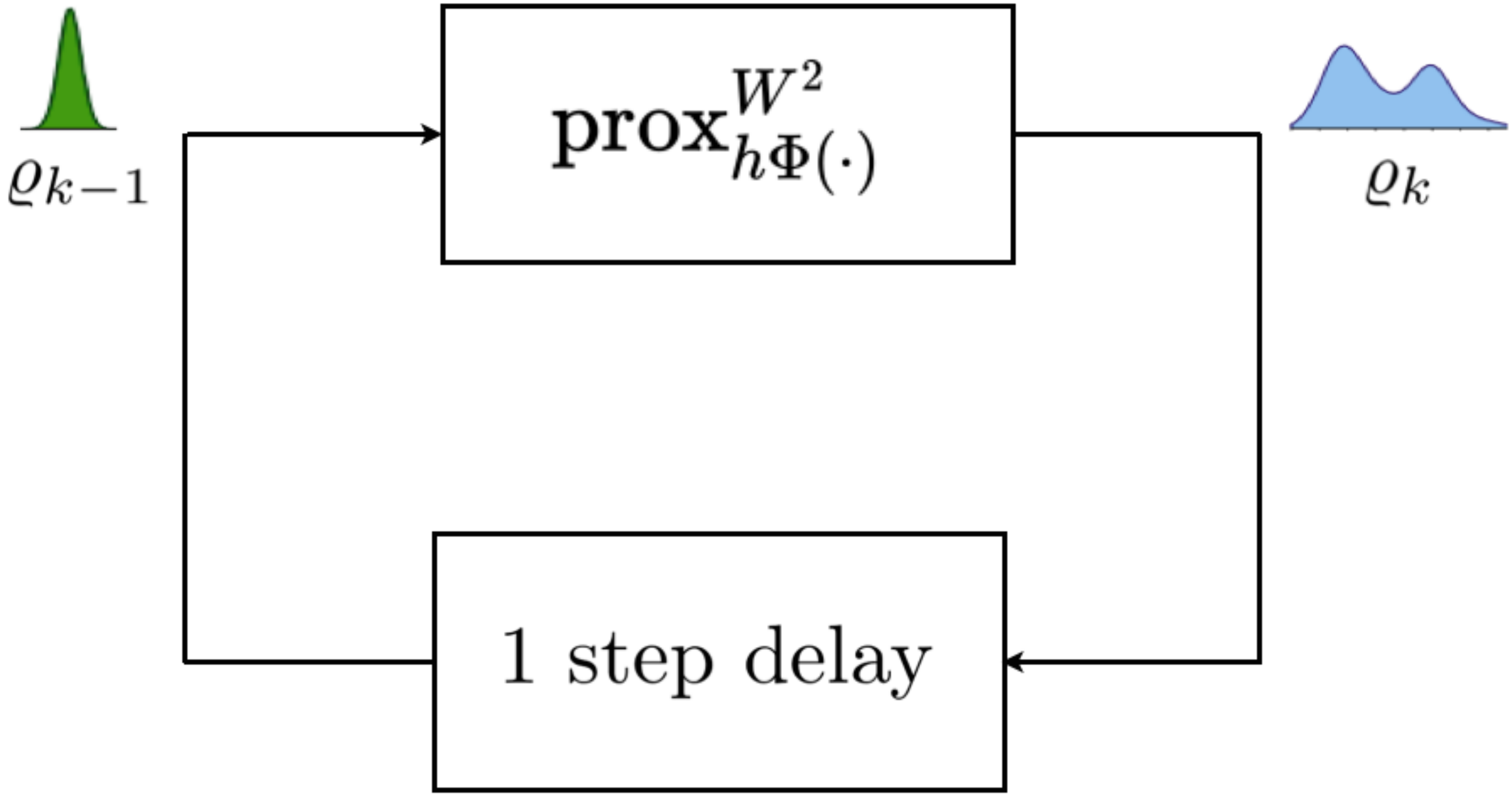


Gradient Flow in $\mathcal{P}_2(\mathcal{X})$



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**

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

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**

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

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**

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

$\Downarrow$ **Entropic Regularization**

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

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**

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

$\Downarrow$ **Entropic Regularization**

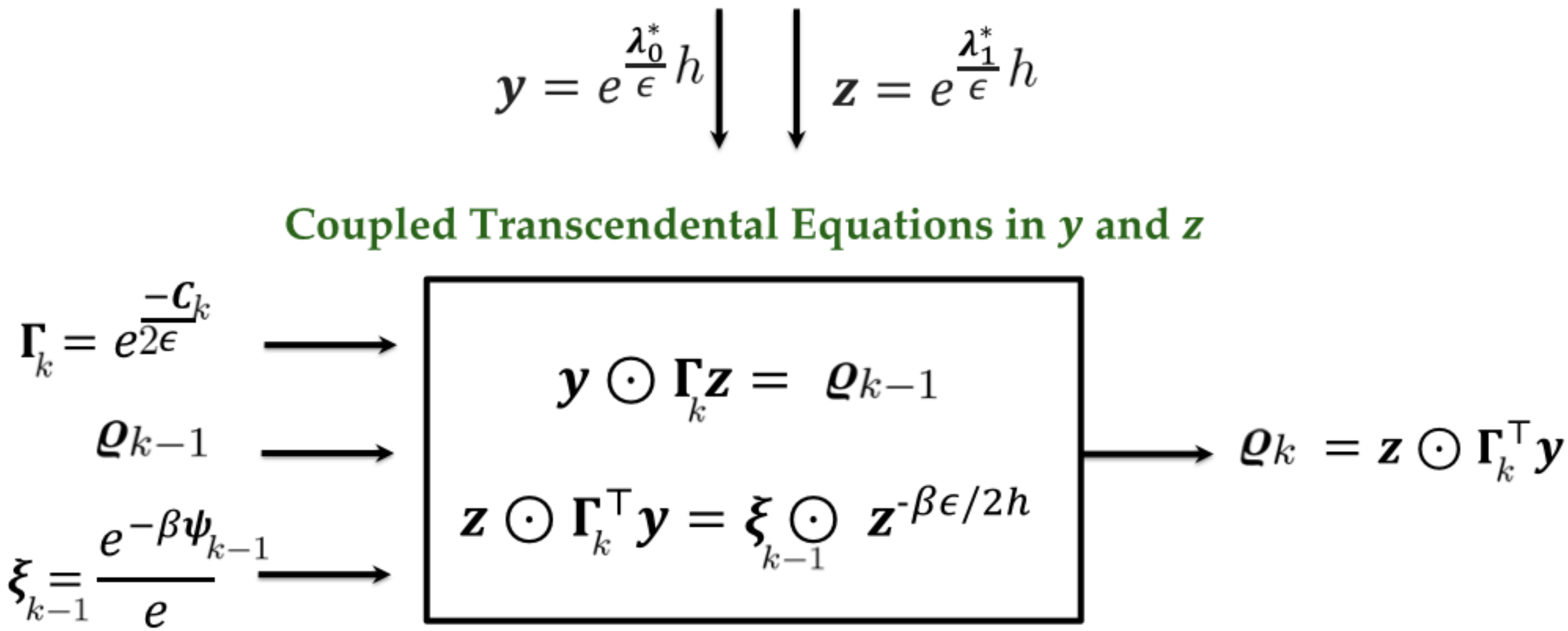
$$\boldsymbol{\varrho}_k = \arg\min_{\boldsymbol{\varrho}} \left\{ \min_{\boldsymbol{M} \in \Pi(\boldsymbol{\varrho}_{k-1}, \boldsymbol{\varrho})} \frac{1}{2} \langle \boldsymbol{C}_k, \boldsymbol{M} \rangle + \epsilon H(\boldsymbol{M}) + h \left\langle \boldsymbol{\psi}_{k-1} + \beta^{-1} \log \boldsymbol{\varrho},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

$\Updownarrow$ **Dualization**

$$\boldsymbol{\lambda}_0^{\rm opt}, \boldsymbol{\lambda}_1^{\rm opt} = \underset{\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1 \geq 0}{\arg\max} \Bigg\{ \langle \boldsymbol{\lambda}_0, \boldsymbol{\varrho}_{k-1} \rangle - F^{\star}(-\boldsymbol{\lambda}_1)$$

$$- \frac{\epsilon}{h} \left(\exp(\boldsymbol{\lambda}_0^{\top} h/\epsilon) \exp(-\boldsymbol{C}_k/2\epsilon) \exp(\boldsymbol{\lambda}_1 h/\epsilon) \right) \Bigg\}$$

Fixed Point Recursion

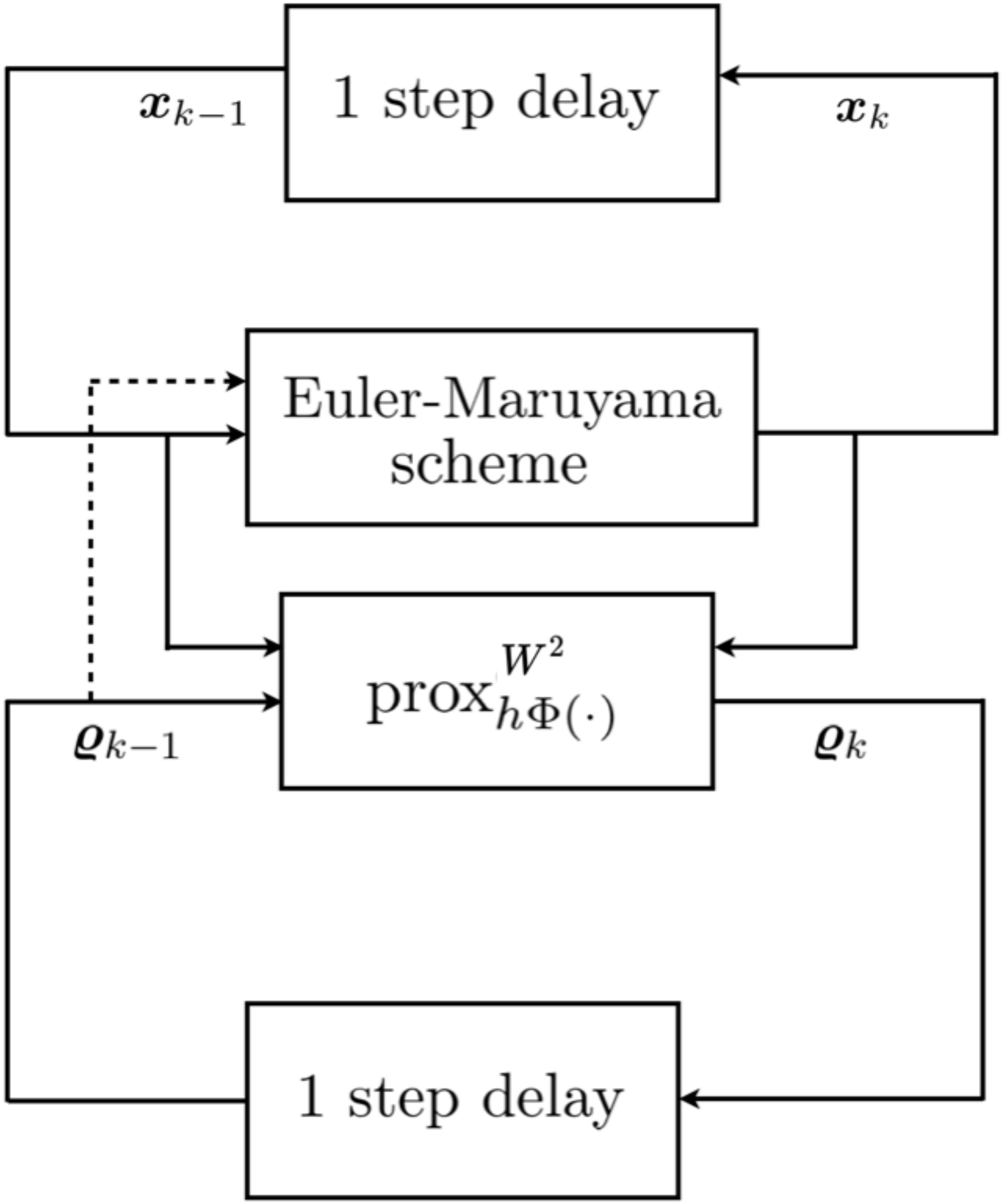


Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

$$\boldsymbol{y} \odot (\boldsymbol{\Gamma}_k \boldsymbol{z}) = \varrho_{k-1}, \quad \boldsymbol{z} \odot \left(\boldsymbol{\Gamma}_k{}^\top \boldsymbol{y}\right) = \boldsymbol{\xi}_{k-1} \odot \boldsymbol{z}^{-\frac{\beta\epsilon}{h}},$$

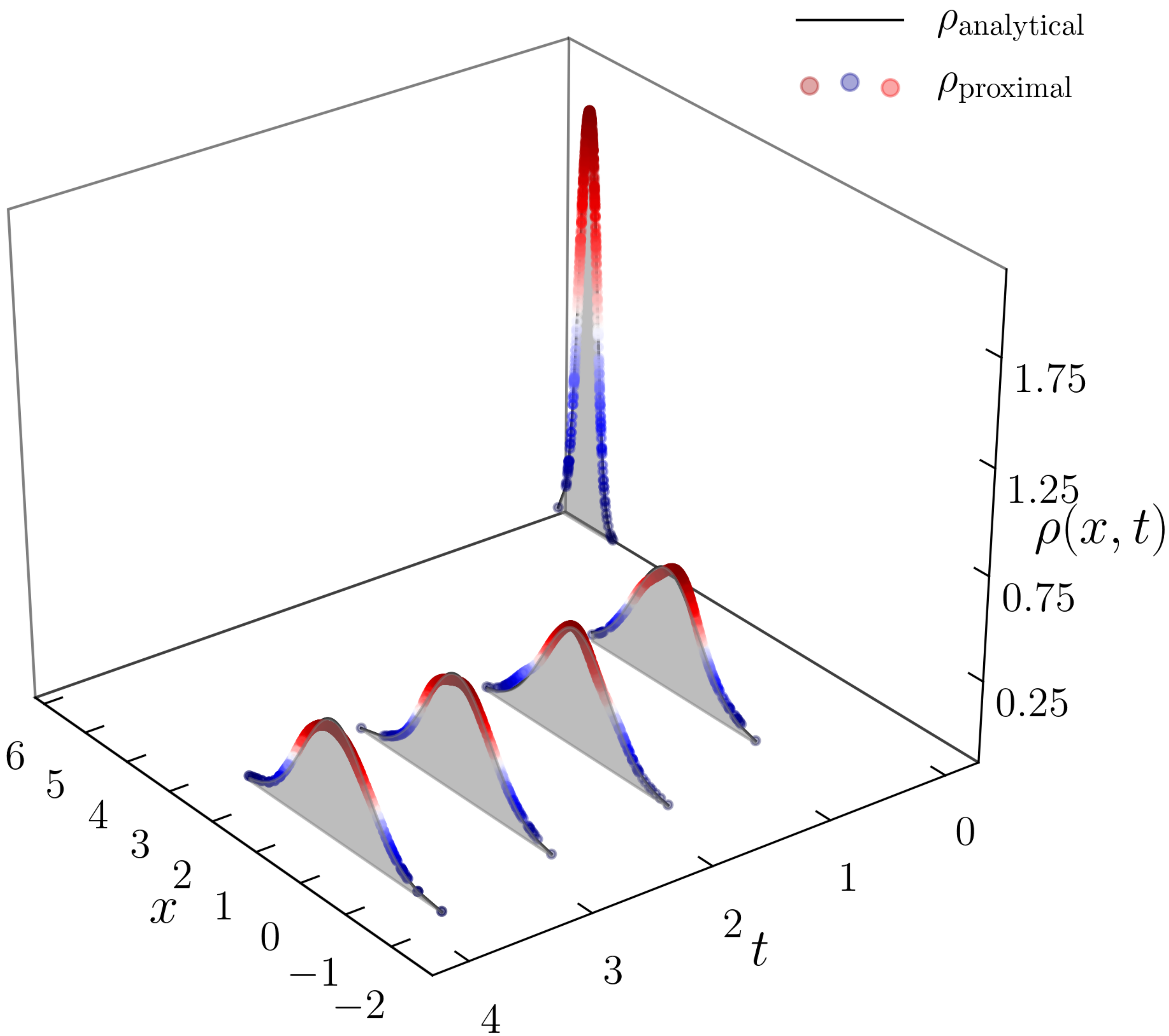
Then the solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ gives the proximal update $\varrho_k = \boldsymbol{z}^* \odot (\boldsymbol{\Gamma}_k{}^\top \boldsymbol{y}^*)$

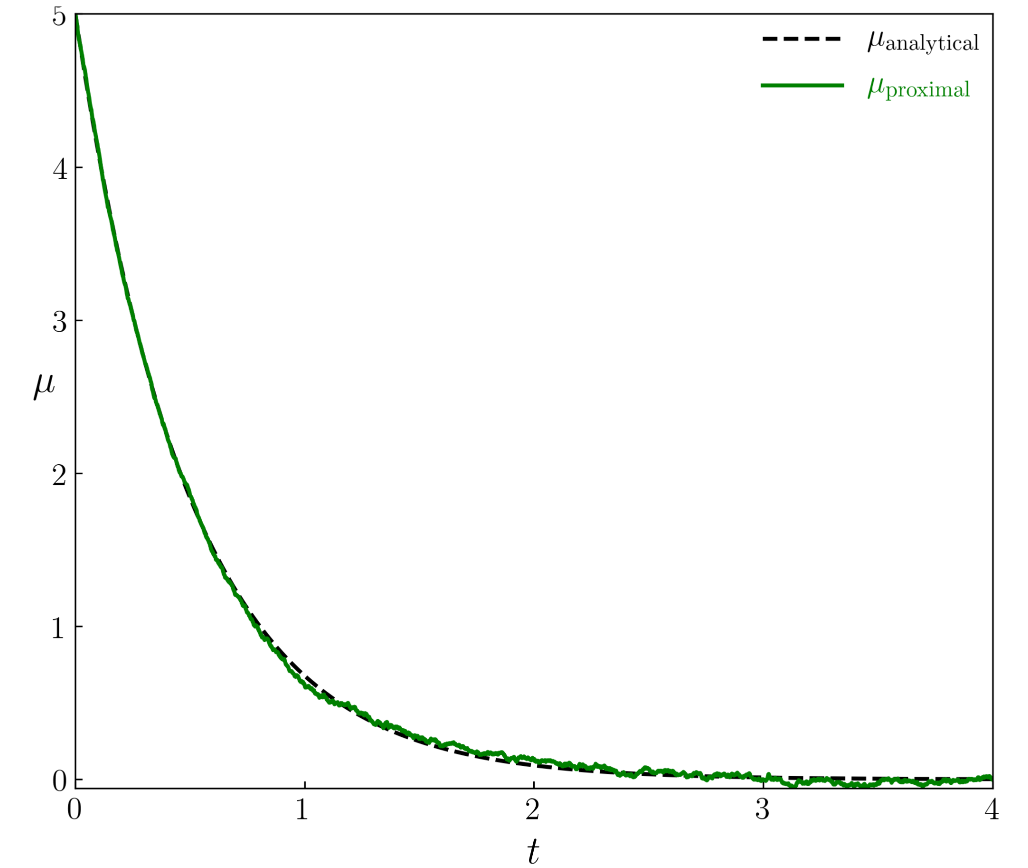
Algorithmic Setup

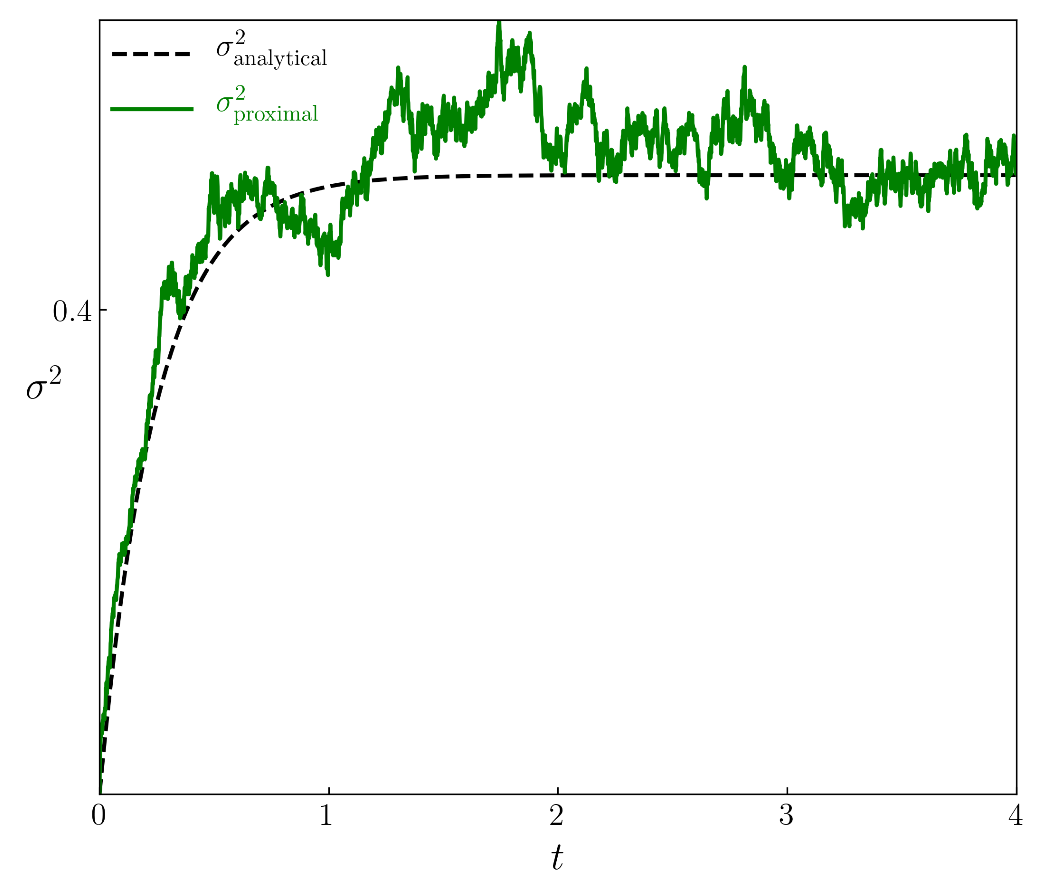


Theorem: Block co-ordinate iteration of $(\mathbf{y}, \mathbf{z})$ recursion is contractive on $\mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0}$.

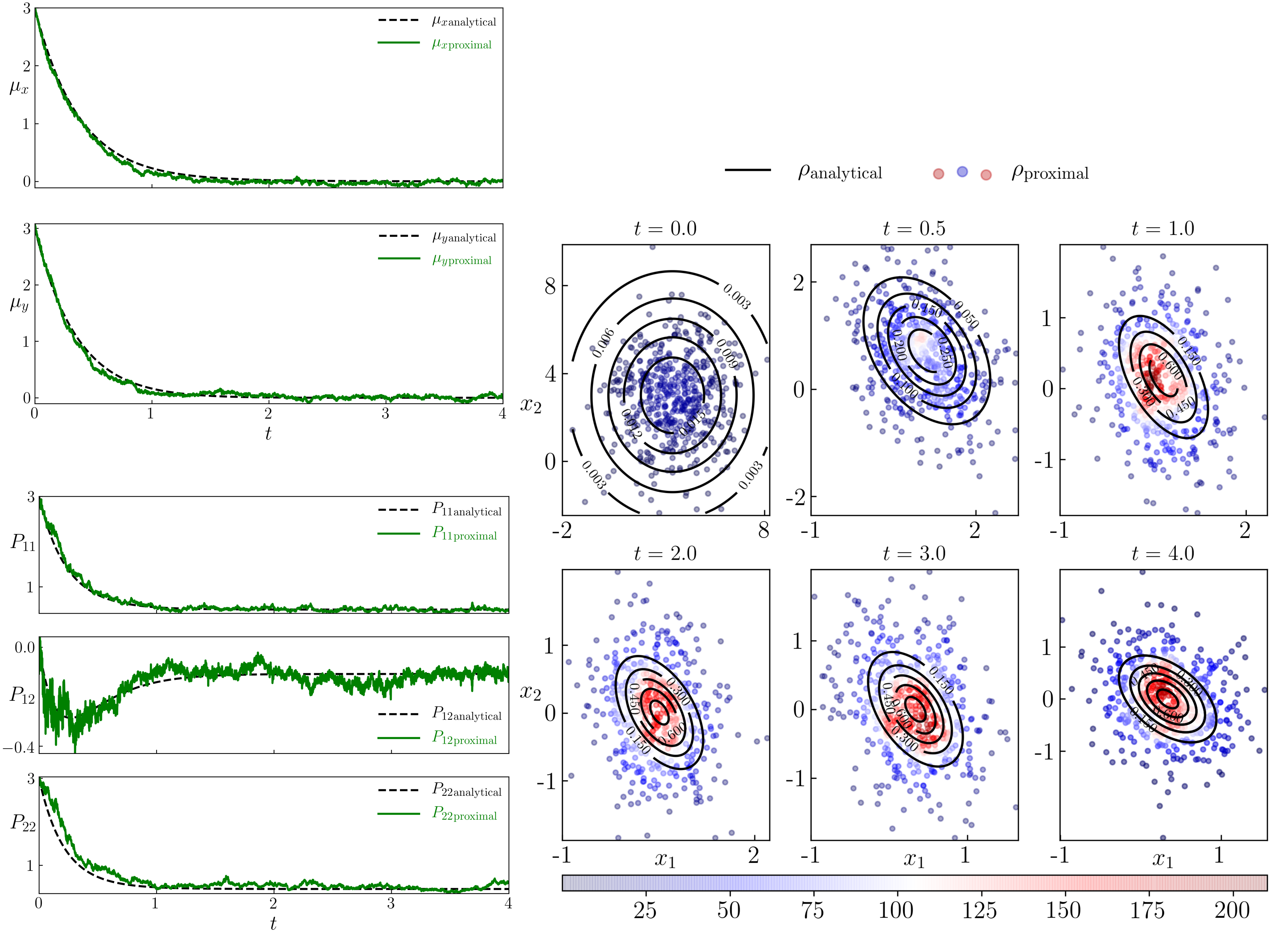
Proximal Prediction: 1D Linear Gaussian



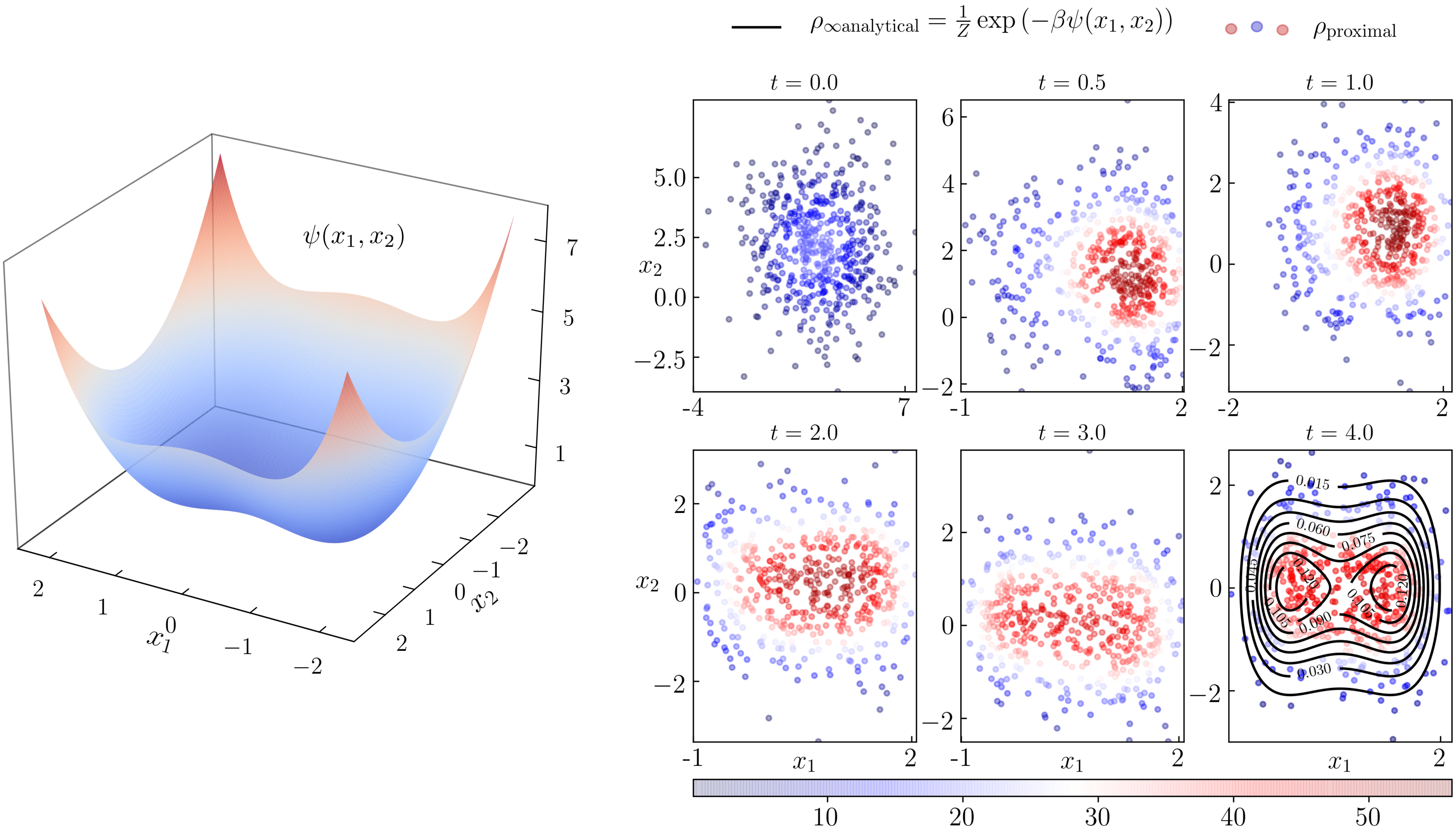




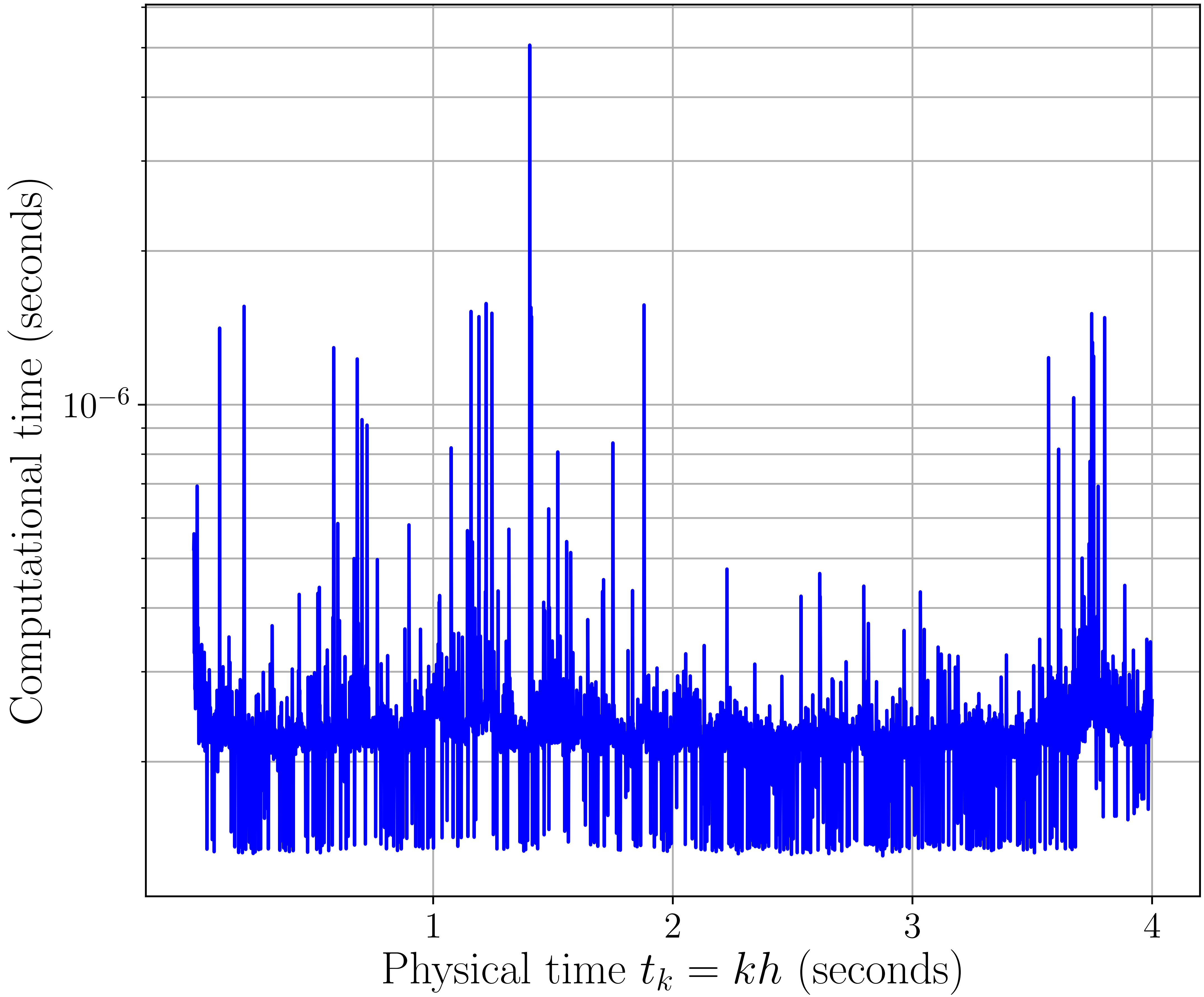
Proximal Prediction: 2D Linear Gaussian



Proximal Prediction: 2D Nonlinear Non-Gaussian



Computational Time: 2D Nonlinear Non-Gaussian



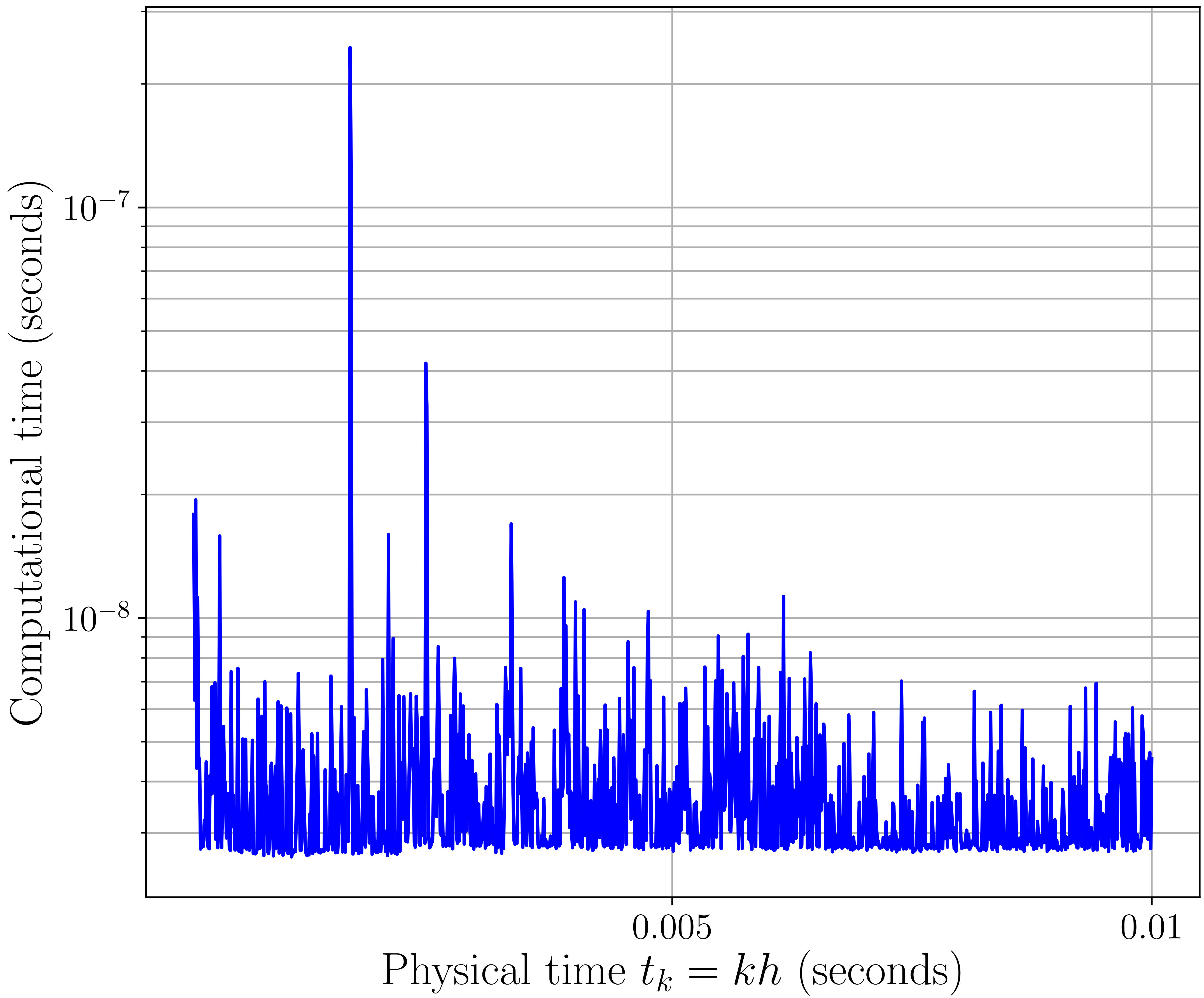
Proximal Prediction: Satellite in Geocentric Orbit

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

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

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

Computational Time: Satellite in Geocentric Orbit



Extensions: Nonlocal interactions

PDF dependent sample path dynamics:

$$\mathrm{d}\mathbf{x} = -\left(\nabla U\left(\mathbf{x}\right) + \nabla \rho * V\right)\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}\mathbf{w}$$

Mckean-Vlasov-Fokker-Planck-Kolmogorov integro PDE:

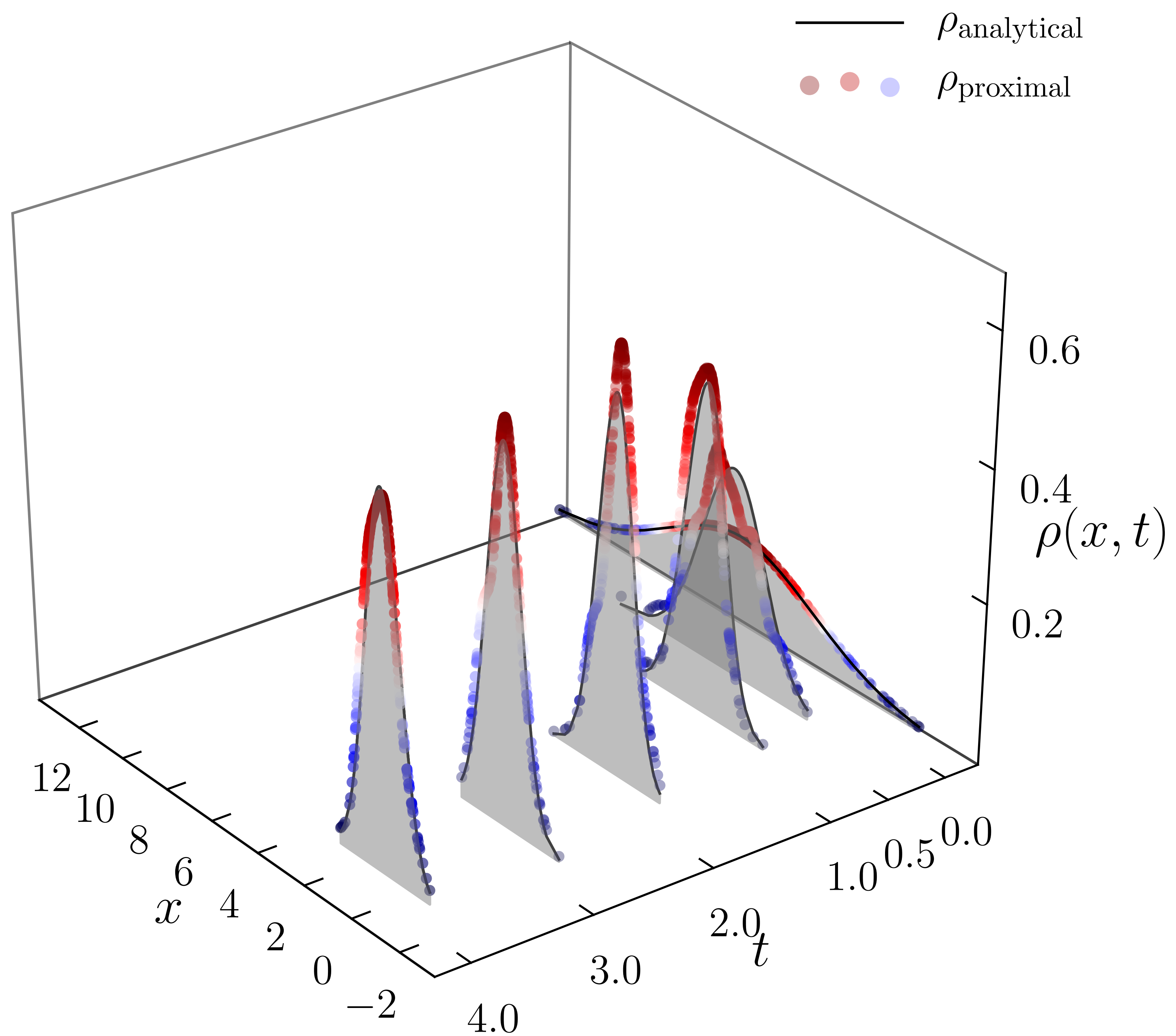
$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \left(U + \rho * V\right)\right) + \beta^{-1}\Delta\rho$$

Free energy:

$$F(\rho) := \mathbb{E}_{\rho}\left[U + \beta^{-1}\rho\log\rho + \rho * V\right]$$

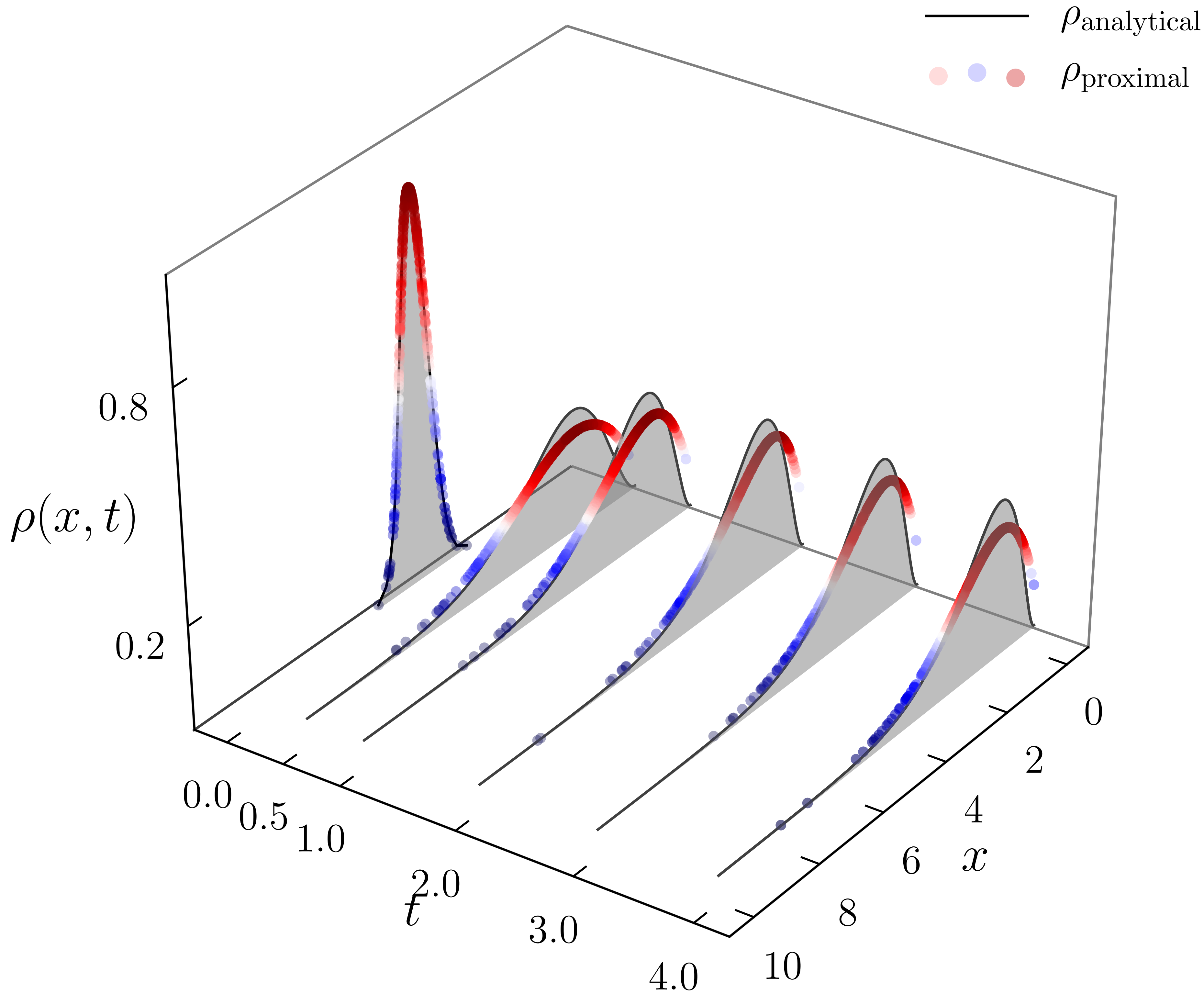
Extensions: Nonlocal interactions (contd.)

$$U(\cdot) = V(\cdot) = \| \cdot \|_2^2$$



Extensions: Multiplicative Noise

Cox-Ingersoll-Ross: $\mathrm{d}x = a(\theta - x)\,\mathrm{d}t + b\sqrt{x}\,\mathrm{d}w,\ 2a > b^2, \theta > 0$



Details on Proximal Prediction

— K.F. Caluya, and A.H., Proximal Recursion for Solving the Fokker-Planck Equation, ACC 2019.

— K.F. Caluya, and A.H., Gradient Flow Algorithms for Density Propagation in Stochastic Systems, accepted in TAC. [arXiv:1908.00533]

Git repo: github.com/kcaluya/UncertaintyPropagation

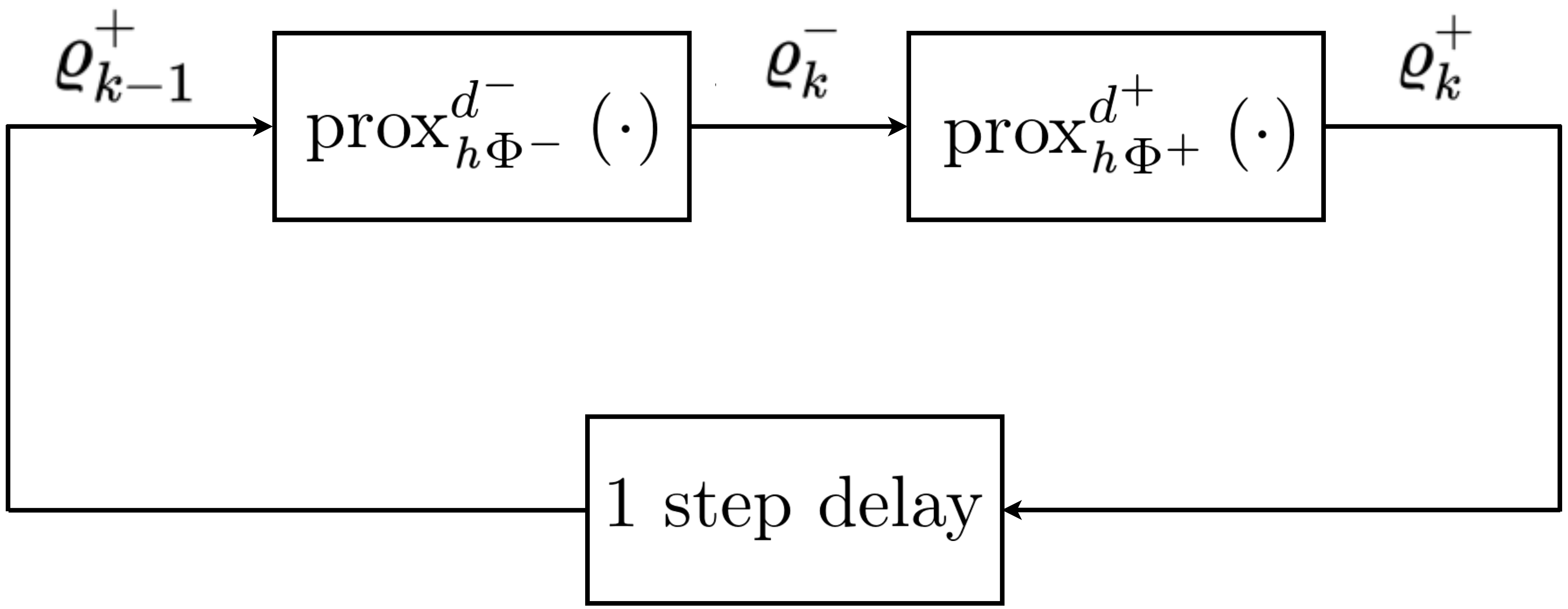
Solving filtering problem as Wasserstein gradient flow

What's New?

Main idea: Solve the Kushner-Stratonovich SPDE

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

Recursion of {deterministic ∘ stochastic} proximal operators:



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

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

For posterior: $d^{+} \equiv d_{\mathrm{FR}}^2$ 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 Kalman-Bucy Filter

Model:

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

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

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

$$\mathrm{d}\mu^+(t) = \mathbf{A}\mu^+(t)\mathrm{d}t + \underbrace{\mathbf{K}(t)}_{\mathbf{P}^+\mathbf{C}\mathbf{R}^{-1}} \left(\mathrm{d}\mathbf{z}(t) - \mathbf{C}\mu^+(t)\mathrm{d}t\right),$$

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

— 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 Wonham Filter

Model:

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

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

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

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

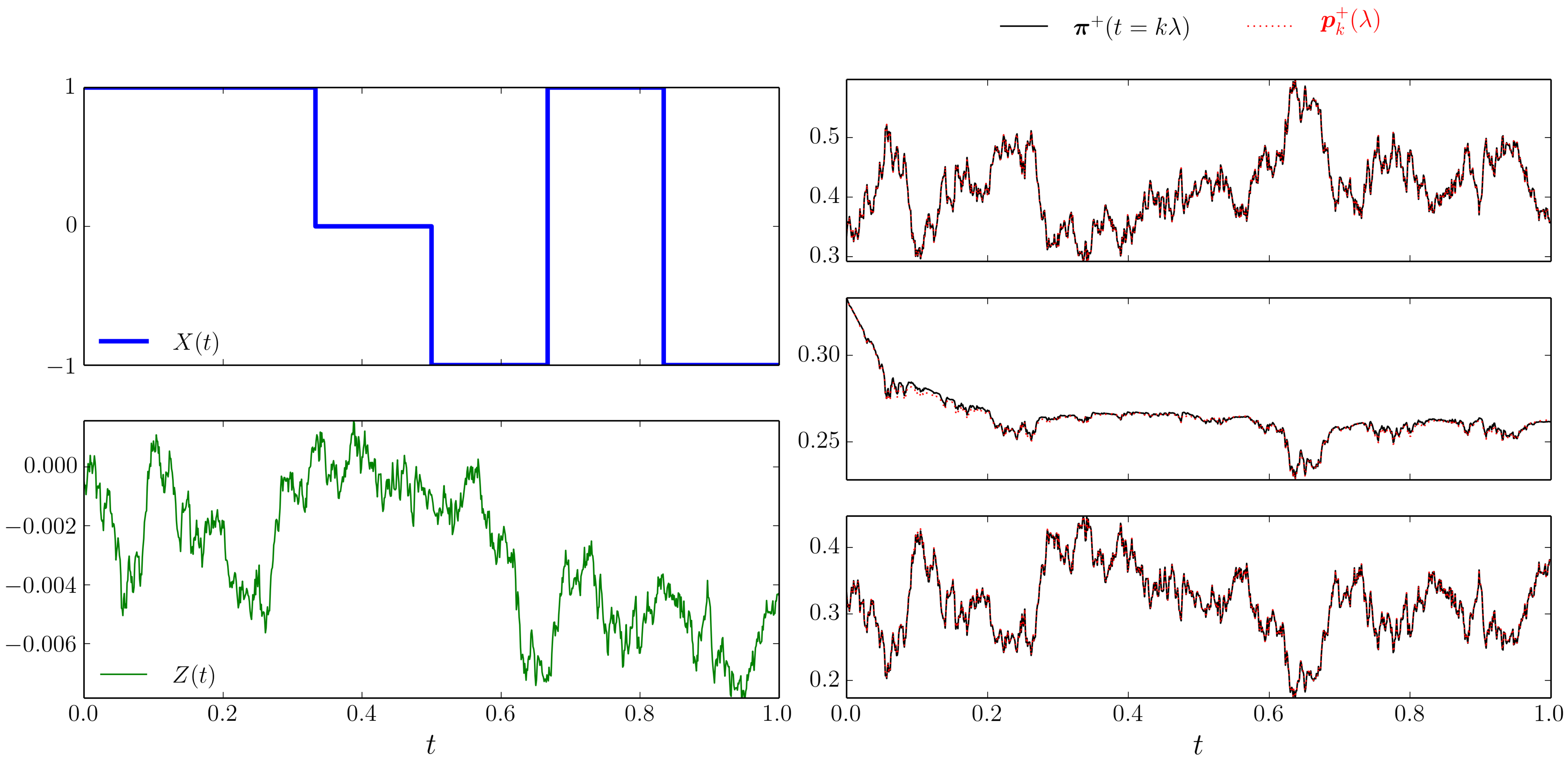
where $H := \mathrm{diag}(h(a_1), \ldots, h(a_m)), \quad \widehat{h}(t) := \sum_{i=1}^{m} h(a_i)\pi_i^+(t),$

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 Wonham Filter



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

Solving density control as Wasserstein gradient flow

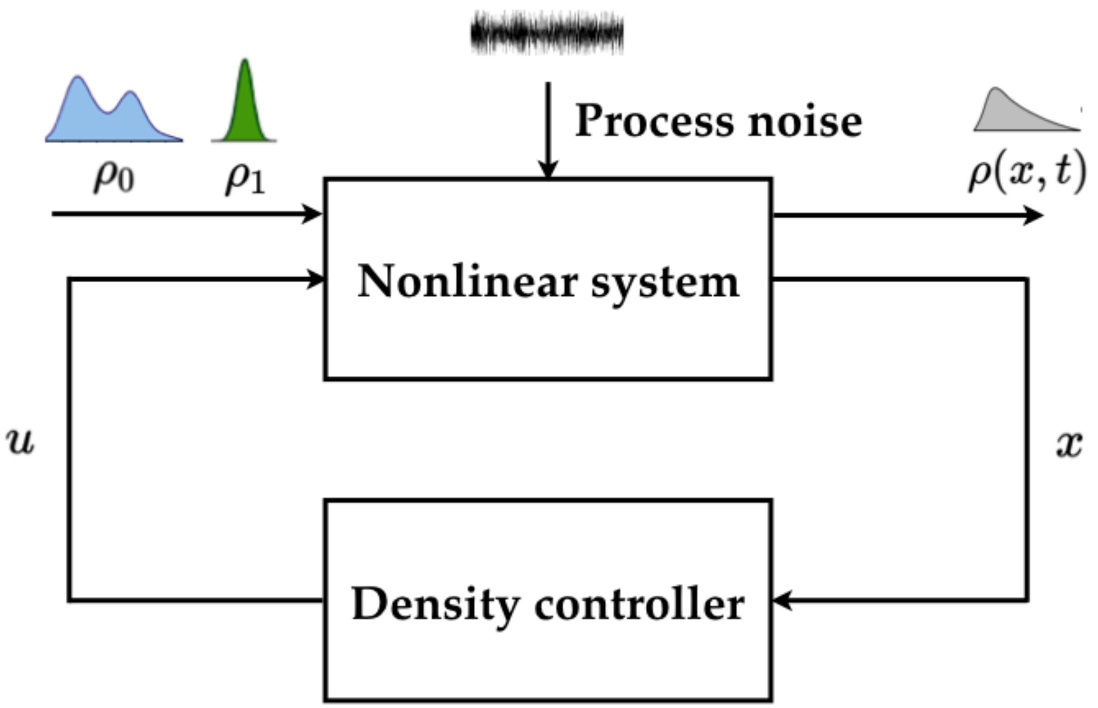
Finite Horizon Feedback Density Control

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

subject to

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

$$x(t=0)\sim\rho_0, \quad x(t=1)\sim\rho_1$$



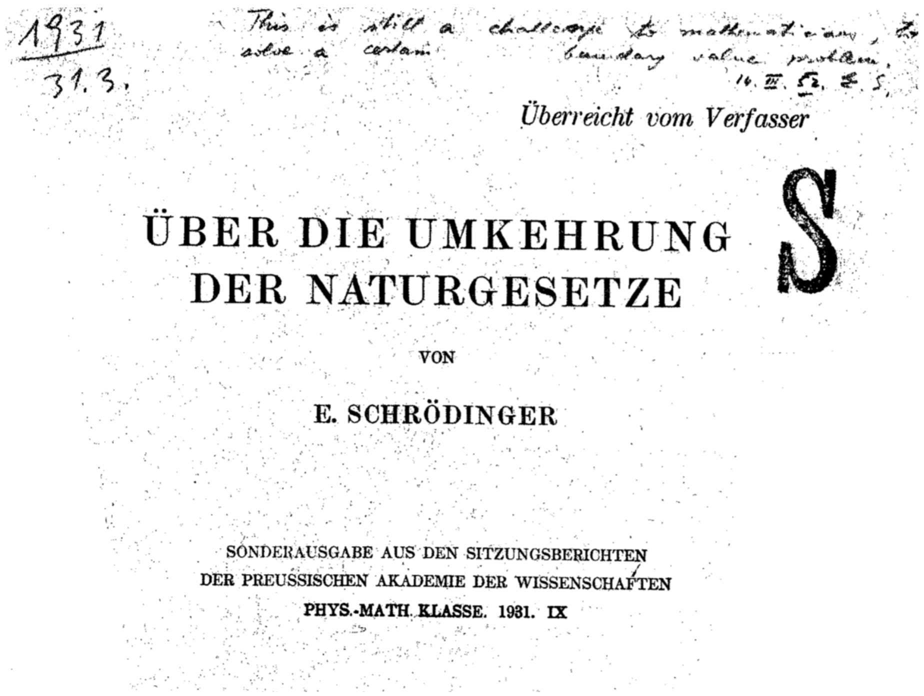
Consider e.g., $f(x,u,t) = f(x,t) + B(t)u(x,t), \quad g(x,t) = \sqrt{2\epsilon}B(t)$

System of coupled Nonlinear PDEs (FPK + HJB):	Boundary conditions:
$\frac{\partial\rho}{\partial t} + \nabla\cdot\left(\rho\left(f + B(t)^\top\nabla\psi\right)\right) = \epsilon\mathbf{1}^\top\left(D(t)\odot\mathrm{Hess}(\rho)\right)\mathbf{1}$	$\rho(x,0)=\rho_0,\ \rho(x,1)=\rho_1$
$\frac{\partial\psi}{\partial t} + \frac{1}{2}\parallel B(t)^\top\nabla\psi\parallel_2^2 + \langle\nabla\psi, f\rangle = -\epsilon\langle D(t), \mathrm{Hess}(\psi)\rangle$	**Optimal control:** $u(x,t) = B(t)^\top\nabla\psi(x,t)$

LTV + Gaussian endpoint PDFs: boundary coupled Riccati ODEs
(TAC 2017 George S. Axelby Outstanding Paper Award)

Solution via Schrödinger System

Schrödinger’s (until recently) forgotten papers:


1931
31.3.
This is still a challenge to mathematicians, to solve a certain boundary value problem, 14.III.52. E.S.

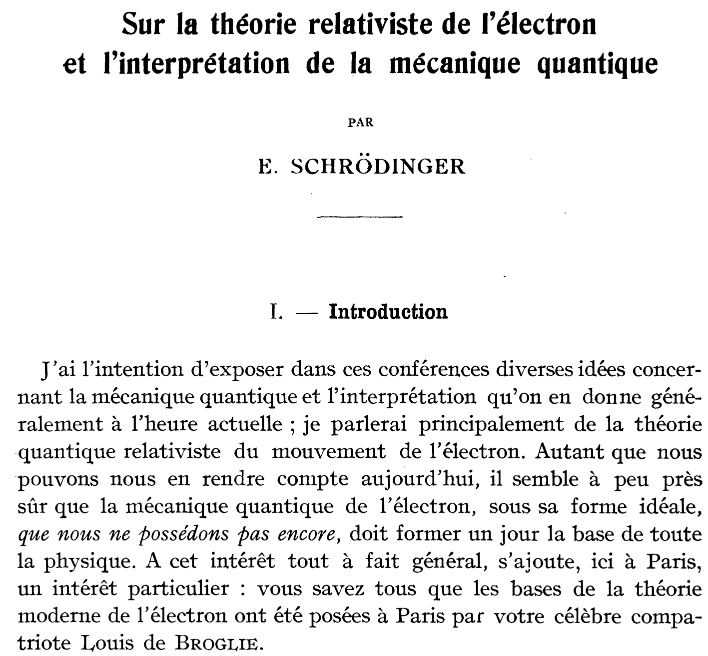
Überreicht vom Verfasser

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.

Schrödinger’s contribution: change of variable $(\rho,\psi)\mapsto(\varphi,\hat{\varphi})$

$$\varphi(x,t)=\exp\left(\frac{\psi(x,t)}{2\epsilon}\right),$$

$$\hat{\varphi}(x,t)=\rho(x,t)\exp\left(-\frac{\psi(x,t)}{2\epsilon}\right),$$

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

Optimal control: $u^*(x,t)=2\epsilon B(t)^\top\nabla\log\varphi(x,t)$

Solution via Schrödinger System

2 coupled nonlinear PDEs ⟶ boundary-coupled linear PDEs!!

Let $D(t) := B(t)B(t)^\top$

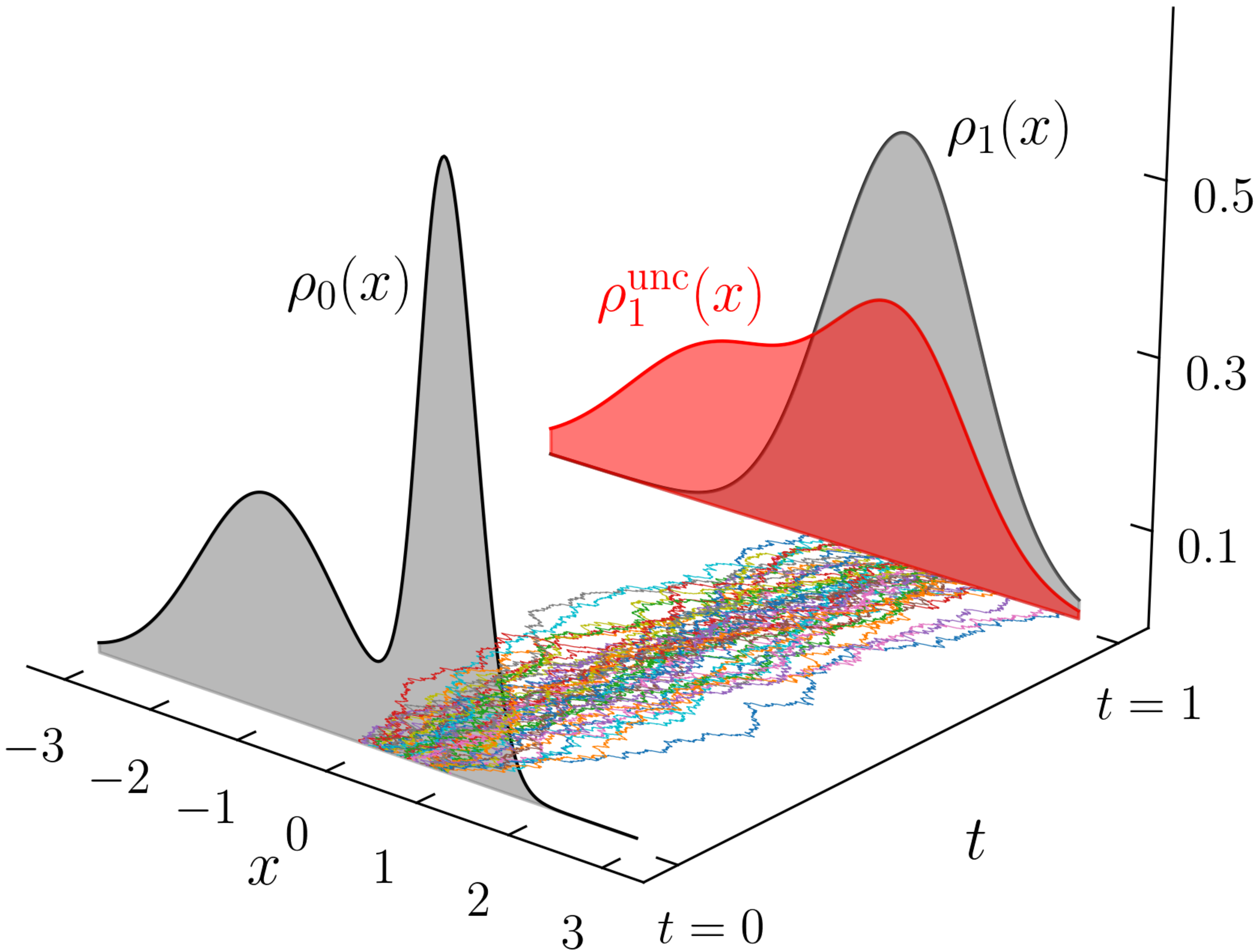
$$\frac{\partial \varphi}{\partial t} = -\langle \nabla \varphi, f \rangle - \epsilon \langle D(t), \mathrm{Hess}(\varphi) \rangle, \qquad \varphi_0 \hat{\varphi}_0 = \rho_0,$$

$$\frac{\partial \hat{\varphi}}{\partial t} = -\nabla \cdot (\hat{\varphi} f) + \epsilon \mathbf{1}^\top (D(t) \odot \mathrm{Hess}(\hat{\varphi}))\mathbf{1}, \qquad \varphi_1 \hat{\varphi}_1 = \rho_1,$$

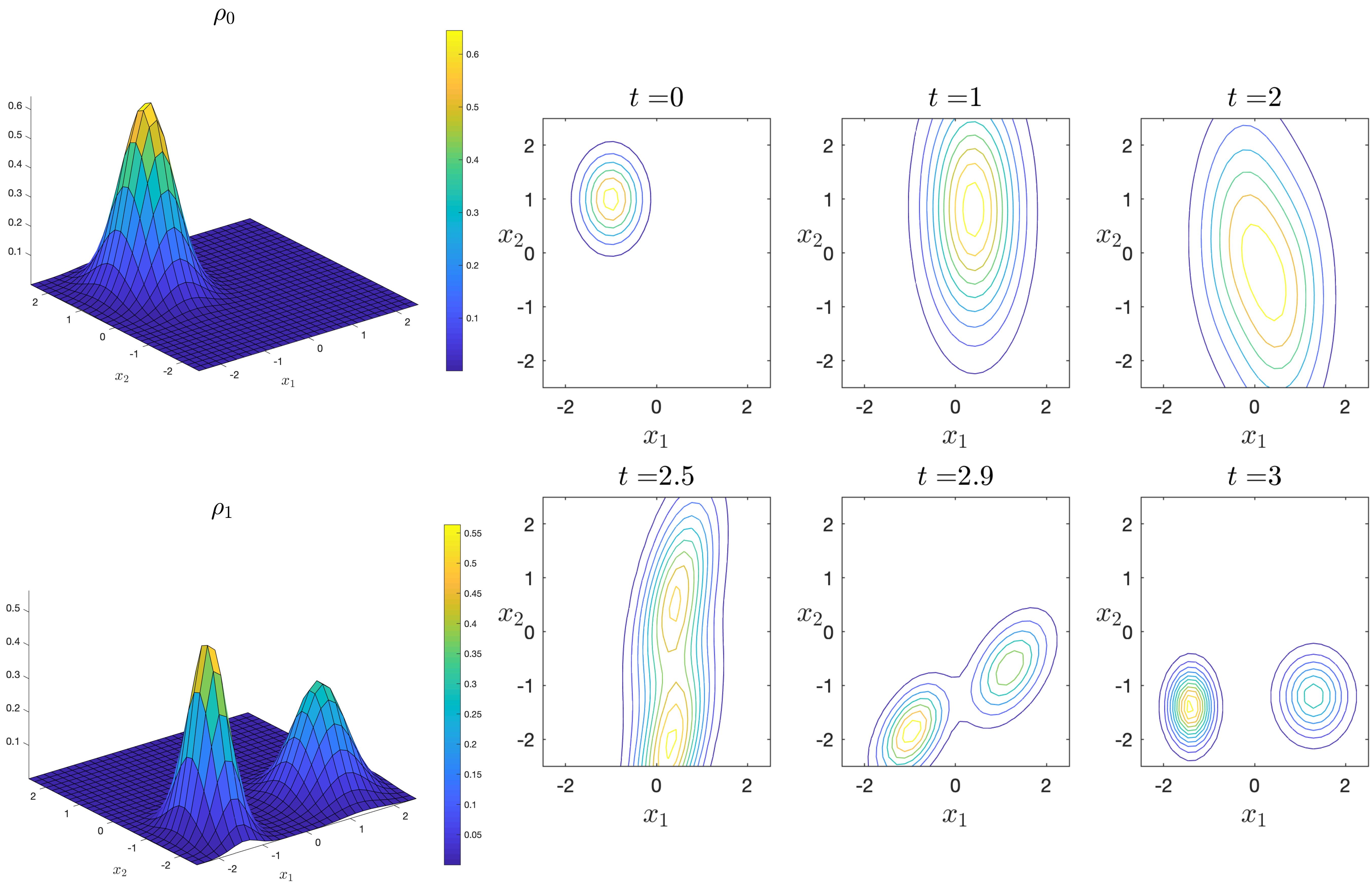
Wasserstein proximal algorithm ⟶ fixed point recursion over $(\varphi_1, \hat{\varphi}_0)$

(Contractive in Hilbert metric)

Feedback Density Control: Zero Prior Dynamics

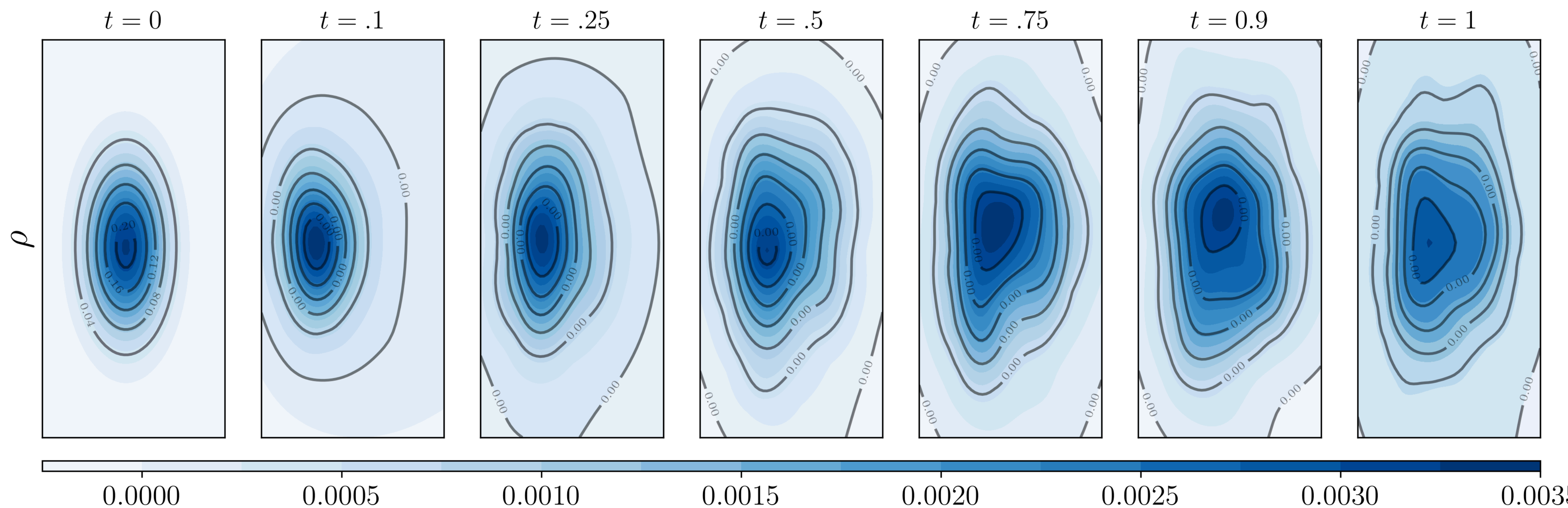


Feedback Density Control: LTI Prior Dynamics

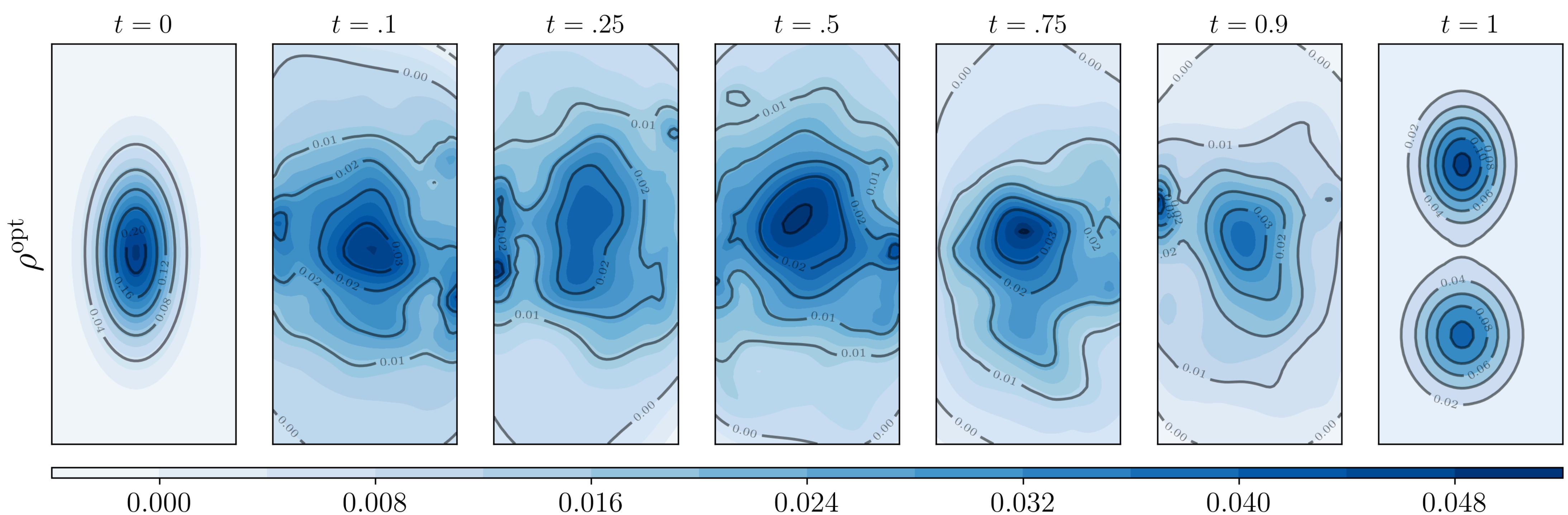


Feedback Density Control: Nonlinear Gradient Prior Dynamics

Uncontrolled joint PDF evolution:



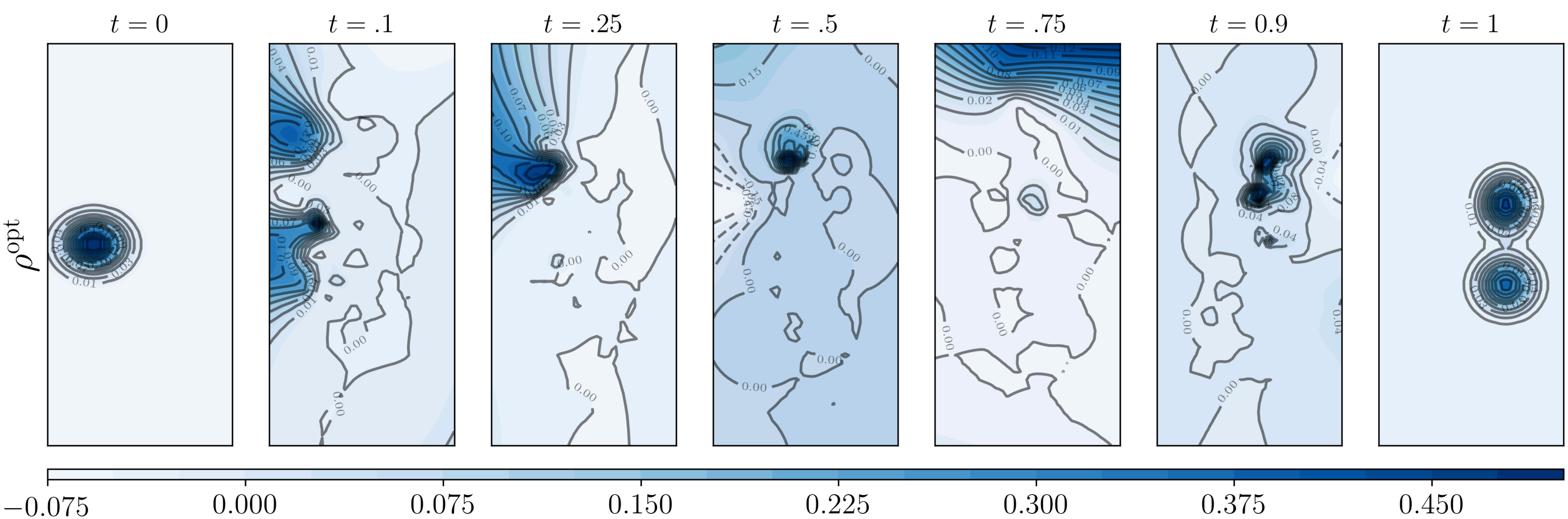
Optimal controlled joint PDF evolution:



Feedback Density Control: Nonlinear Mixed Conservative-Dissipative Prior Dynamics

$$\begin{pmatrix} \mathrm{d}\xi \\ \mathrm{d}\eta \end{pmatrix} = \Bigg\{ \underbrace{\begin{pmatrix} \eta \\ -\nabla_{\xi} V(\xi) - \kappa\eta \end{pmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{pmatrix} \mathbf{0}_{m\times m} \\ \mathbf{I}_{m\times m} \end{pmatrix}}_{\mathbf{B}} \mathbf{u} \Bigg\} \mathrm{d}t + \sqrt{2\epsilon\kappa} \underbrace{\begin{pmatrix} \mathbf{0}_{m\times m} \\ \mathbf{I}_{m\times m} \end{pmatrix}}_{\mathbf{B}} \mathrm{d}\mathbf{w}, \quad \kappa > 0,$$

Optimal controlled joint PDF evolution:



Details on Feedback Density Control for Nonlinear Systems

— K.F. Caluya, and A.H., Finite Horizon Density Control for Static State Feedback Linearizable Systems, under review in TAC. [arXiv:1904.02272]

— K.F. Caluya, and A.H., Finite Horizon Density Steering for Multi-input State Feedback Linearizable Systems, under review in ACC 2020. [arXiv:1909.12511]

— K.F. Caluya, and A.H., Wasserstein Proximal Algorithms for the Schrodinger Bridge Problem: Density Control with Nonlinear Drift, working draft.

Upcoming tutorial lectures on feedback density control:

— "Toward Feedback Control of Densities in Nonlinear Systems"
Invited lecture at "Uncertainty Synthesis Workshop", IEEE CDC 2019, Nice, France.

— "Wasserstein Gradient Flow for Filtering and Stochastic Control"
Invited lecture at SIAM UQ 2020, Munich, Germany.

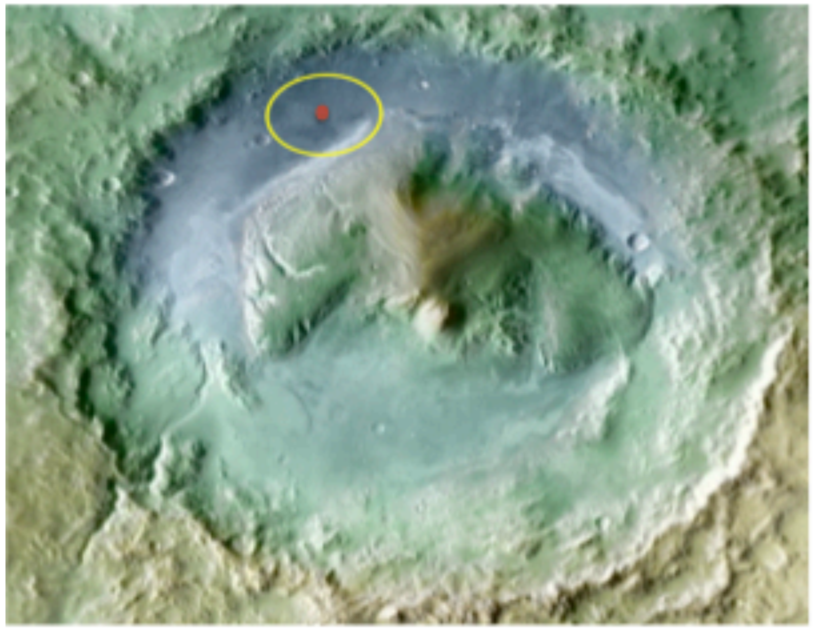
Take Home Message

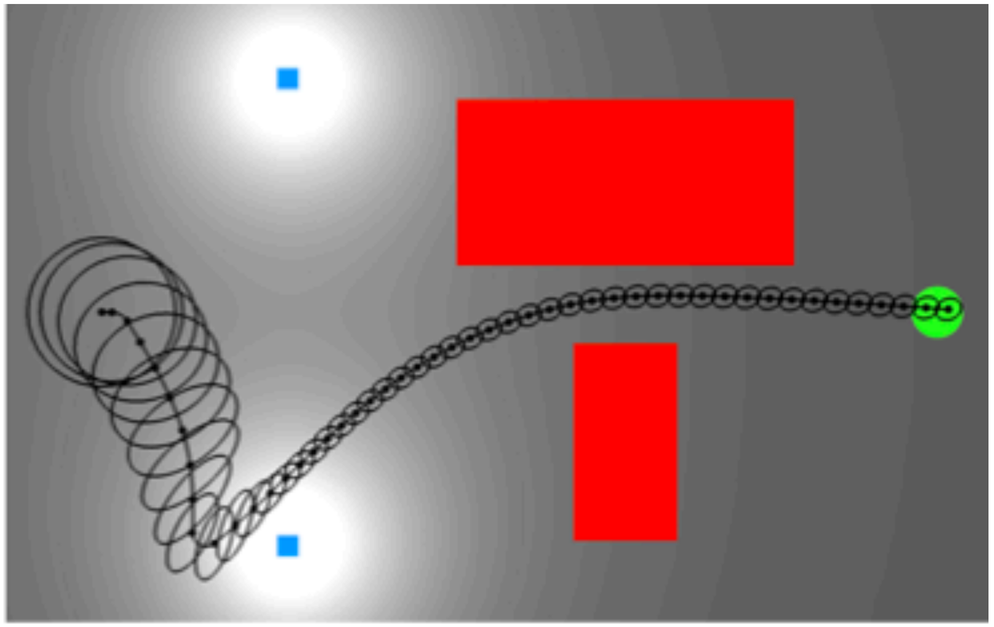
Emerging systems-control theory of densities

Three problems: prediction, filtering, control

One unifying framework: Wasserstein gradient flow

Many applications:





Thank You