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)

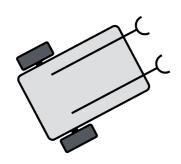


Overarching theme

Systems-control theory for densities

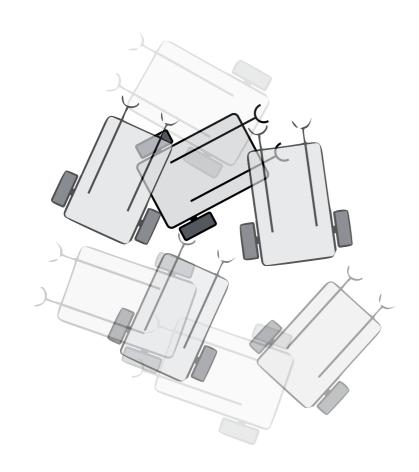
What is density?

Probability Density Fn.



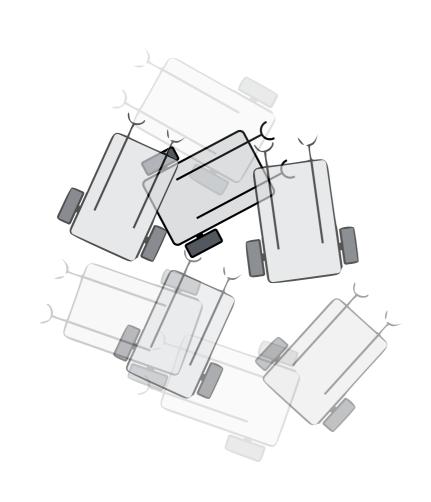
$$x(t) \in \mathcal{X} \equiv \ \mathbb{R}^2{ imes}\mathbb{S}^1$$

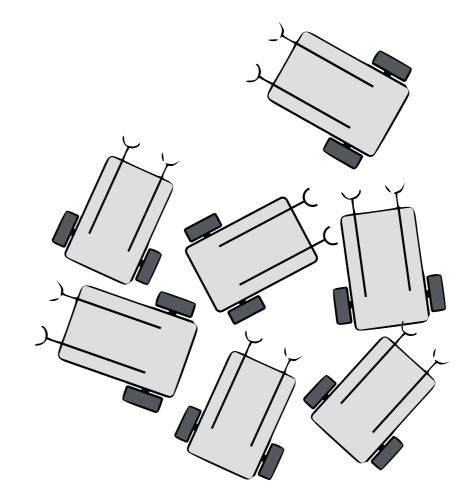
Probability Density Fn.



$$x(t) \in \mathcal{X} \equiv \ \mathbb{R}^2 imes \mathbb{S}^1$$
 $ho(x,t): \mathcal{X} imes [0,\infty) \mapsto \mathbb{R}_{\geq 0}$ $\int_{\mathcal{X}}
ho \, \mathrm{d}x = 1 \quad ext{for all} \quad t \in [0,\infty)$

Probability Density Fn. Population Density Fn.



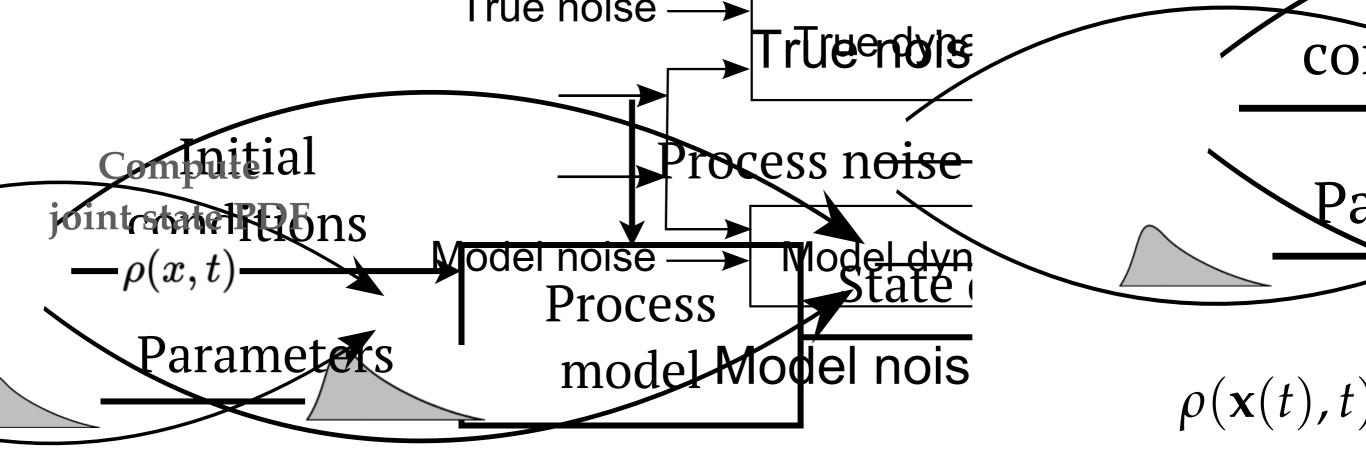


$$x(t) \in \mathcal{X} \equiv \ \mathbb{R}^2{ imes}\mathbb{S}^1$$

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

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

Why care about densities?

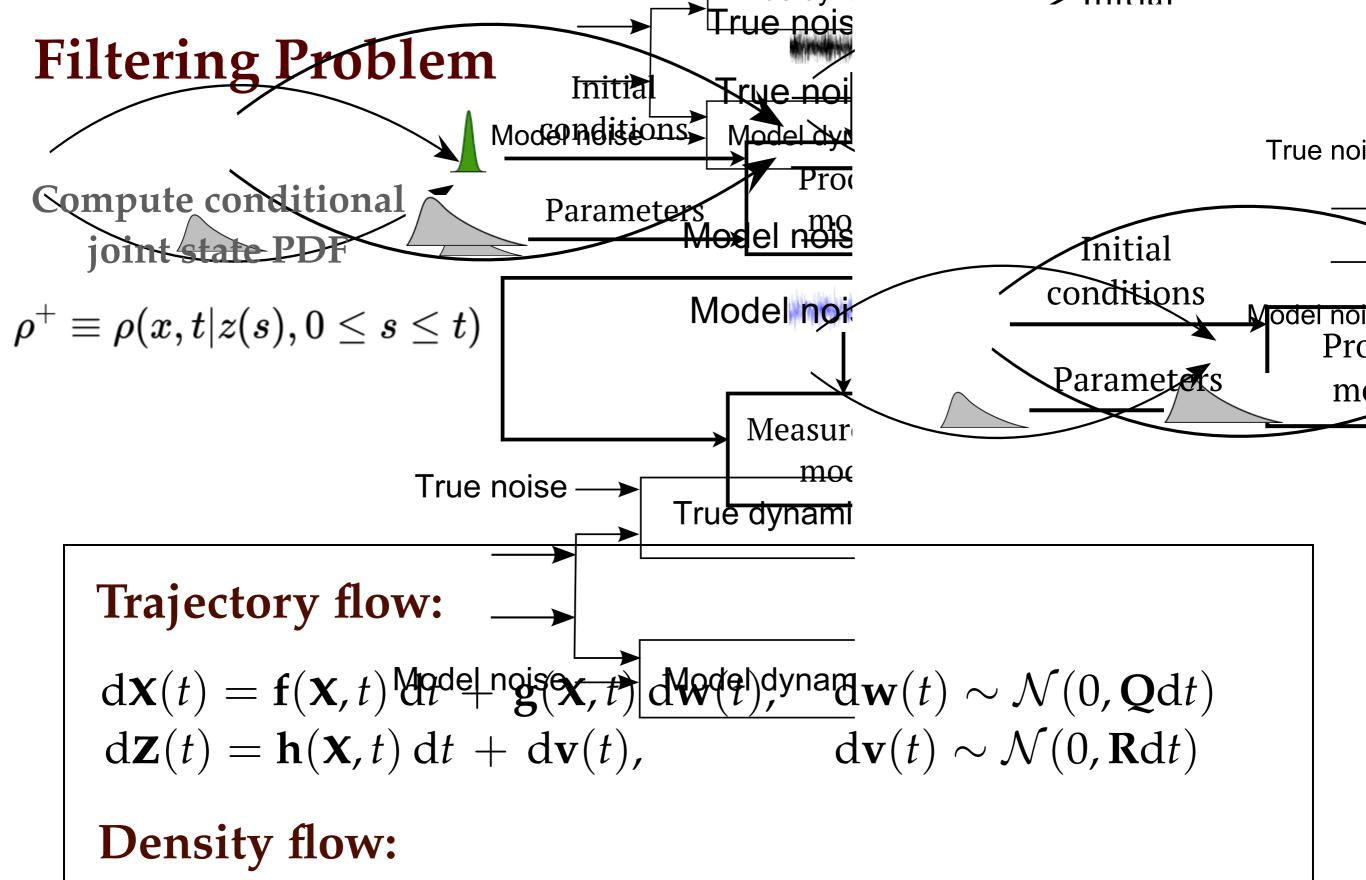


Trajectory flow:

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

Density flow:

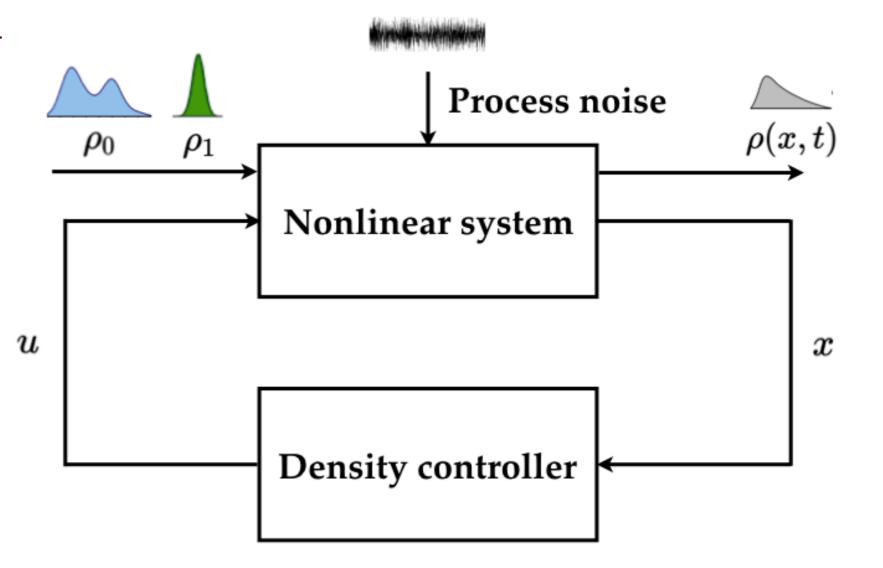
$$\frac{\partial \rho}{\partial t} = \mathcal{L}_{\text{FP}}(\rho) := -\nabla \cdot (\rho \mathbf{f}) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} \left(\left(\mathbf{g} \mathbf{Q} \mathbf{g}^{\top} \right)_{ij} \rho \right)$$



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

Control Problem

Steer joint state PDF via feedback control



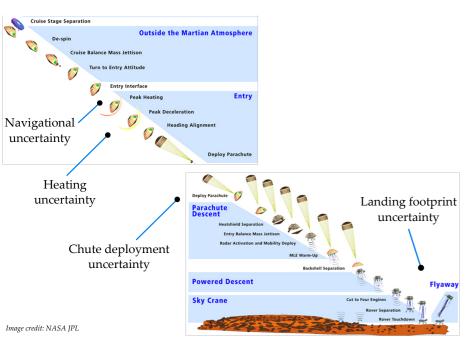
$$egin{align} ext{minimize} & \mathbb{E}igg[\int_0^1 \|u\|_2^2 \, \mathrm{d}tigg] \ ext{subject to} \ & \mathrm{d}x = f(x,u,t) \, \mathrm{d}t \, + \, g(x,t) \, \mathrm{d}w, \ & x(t=0) \sim
ho_0, \quad x(t=1) \sim
ho_1 \ \end{aligned}$$

PDFs in Mars Entry-Descent-Landing

Prediction Problem

Filtering Problem

Control Problem





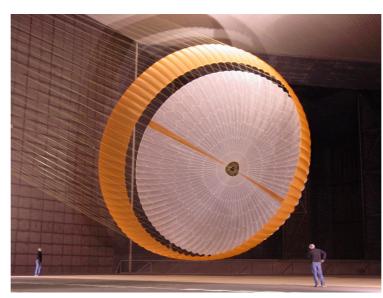
Predict heating rate uncertainty

PDFs in Mars Entry-Descent-Landing

Prediction Problem

Cruise Stage Separation De-spin De-spin Turn to Entry Attitude Entry Interface Peak Heating Entry Peak Deceleration Heading Alignment Uncertainty Deploy Parachute Descent Entry Peak Heating Uncertainty Deploy Parachute Descent Entry Deploy Parachute Descent Desc

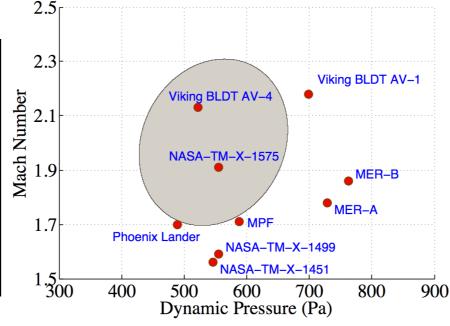
Filtering Problem



Control Problem

Supersonic parachute





Predict heating rate uncertainty

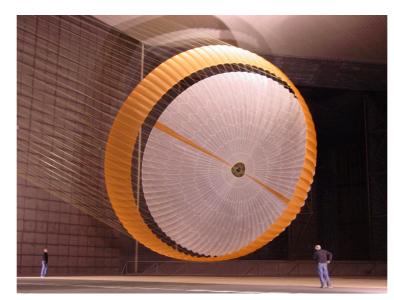
Estimate state to deploy parachute

PDFs in Mars Entry-Descent-Landing

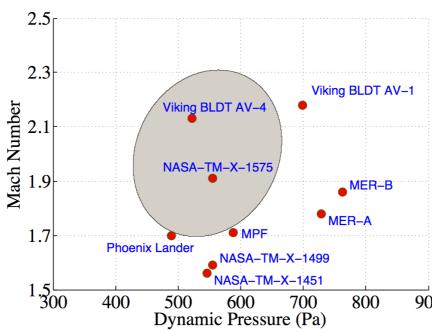
Prediction Problem

Navigational uncertainty Heating uncertainty Landing footprint uncertainty Chute deployment uncertainty Powered Descent Image credit: NASA [PI

Filtering Problem



Supersonic parachute



deployment (shaped region).

Predict heating rate uncertainty

was minimizedightals 6 Avs. and additional superysonicas predsthear

Smart AM to bear that whint it had was more in the later that the many

Estimate state to deploy parachute

Steer state PDF to achieve desired landing footprint accuracy

measure either of these quantities. Therefore, all missions have had to rely on proxy measurements of other states in or-

Control Problem

higher Mach numbers result in increased aerothermal heating of parachute structure, which can reduce material strength; and (3) at Mach numbers above Mach 1.5, DGB parachutes exhibit an instability, known as areal oscillations, which result in multiple partial collapses and violent re-inflations. The chief concern with high Mach number deployments, for parachute deployments in regions where the heating is not a driving factor, is therefore, the increased exposure to area oscillations.

The Viking parachute system was qualified to deploy between Mach 1.4 and 2.1, and a dynamic pressure between 250 and 700 Pa [1]. However, Mach 2.1 is not a hard limit for successfully operating DBG parachutes at Mars and there is very little flight test data above Mach 2.1 with which to quantify the amount of increased EDL system risk. Figure 3 shows the relevant flight tests and flight experience in the region of the planned MSh payashure deployte While magashure experts agree that higher Mach numbers result in a higher probability of failure, they have different opinions on where the limit should be blaced. For example, Gillis [5] 1918 proposed Waltupp per bound and Mach 2 for parachute aerodynamic decelerators at Mars. However, Cruz [3] places the upper Mach number range somewhere between two and three .03 E -2.25

 $4.49^{\circ}S$ | $137.42^{\circ}E$ Gale Crater presents a challenge for EDL system designer rswalde Crater 23.80 \$ 320.73 designer t then weigh the system performance gains and ris

social designation of the state altitudes and Mach numbers, against a very real, but not well

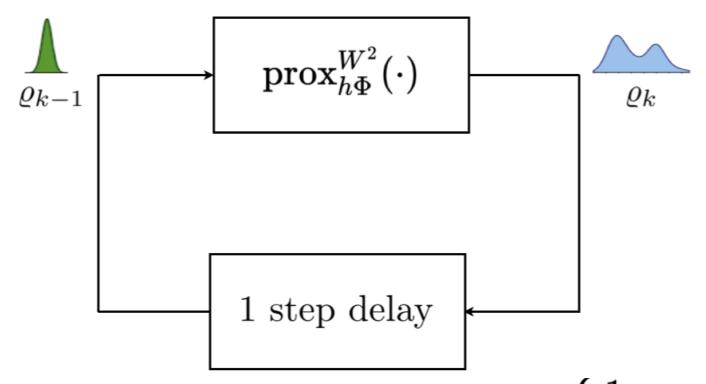
quantified, probability of parachute failure. It is clear that deploying a DGB at Mach 2.5 or 3.0 represents a significant pose of proposing and selecting possible landing sites. While here are the risk ever an initiation at What 2.9. However, it is many crites hwere initially horaposed at the first of by assumerkshapsttheautonnearfiths 4th Landing Site Workshop in 12008 was a flist mats for de candidate trite at reister langeable en la material de la company de la comp 900 four then all stites a VE berns wall despectately have object the high less still exaction at win45, that resolvan vehidhrise significantity ibeVowltherabioude capability of the system (estimated to be somewhere around

Solving prediction problem as Wasserstein gradient flow

What's New?

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

Infinite dimensional variational recursion:



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

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

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 $\mathcal{P}_2(\mathcal{X})$

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = -\nabla \varphi(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0$$

$$\frac{\partial \rho}{\partial t} = -\nabla^W \Phi(\rho), \quad \rho(\mathbf{x}, 0) = \rho_0$$

Recursion:

$$\mathbf{x}_k = \mathbf{x}_{k-1} - h \nabla \varphi(\mathbf{x}_k)$$

$$= \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{arg min}} \left\{ \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}_{k-1}\|_{2}^{2} + h\varphi(\boldsymbol{x}) \right\}$$

$$=: \operatorname{prox}_{h\varphi}^{\|\cdot\|_2}(\boldsymbol{x}_{k-1})$$

Recursion:

$$\rho_k = \rho(\cdot, t = kh)$$

$$\rho \in \mathcal{P}_2(\mathcal{X})$$

$$= \underset{\rho \in \mathcal{P}_2(\mathcal{X})}{\operatorname{arg min}} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h \Phi(\rho) \right\}$$

$$=: \operatorname{prox}_{h\Phi}^{W^2}(\rho_{k-1})$$

Convergence:

$$\mathbf{x}_k \to \mathbf{x}(t = kh)$$
 as $h \downarrow 0$

$$\rho_k \to \rho(\cdot, t = kh)$$
 as $h \downarrow 0$

Convergence:

φ as Lyapunov function:

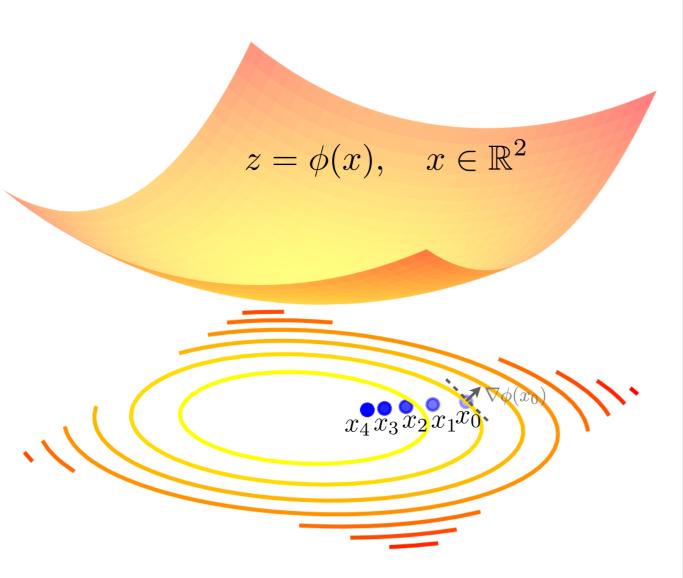
$$rac{\mathrm{d}}{\mathrm{d}t}arphi = -\parallel
abla arphi \parallel_2^2 \ \le \ 0$$

Φ as Lyapunov functional:

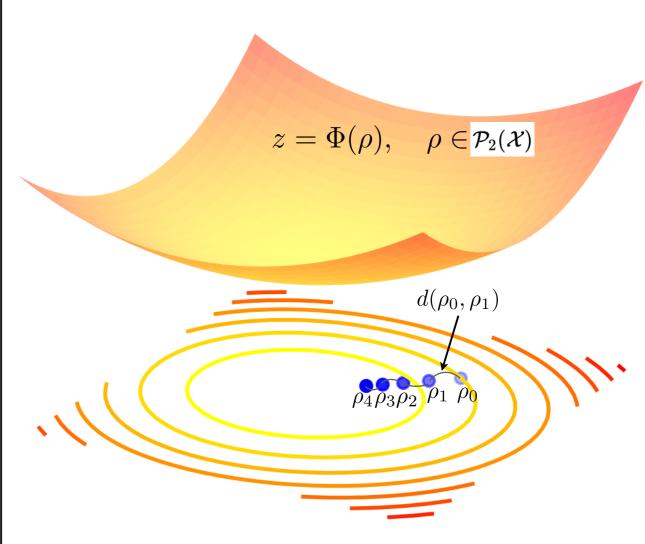
$$rac{\mathrm{d}}{\mathrm{d}t}\Phi = -\mathbb{E}_
hoigg[igg\|
ablarac{\delta\Phi}{\delta
ho}igg\|_2^2igg] ~\leq ~0$$

Geometric Meaning of Gradient Flow

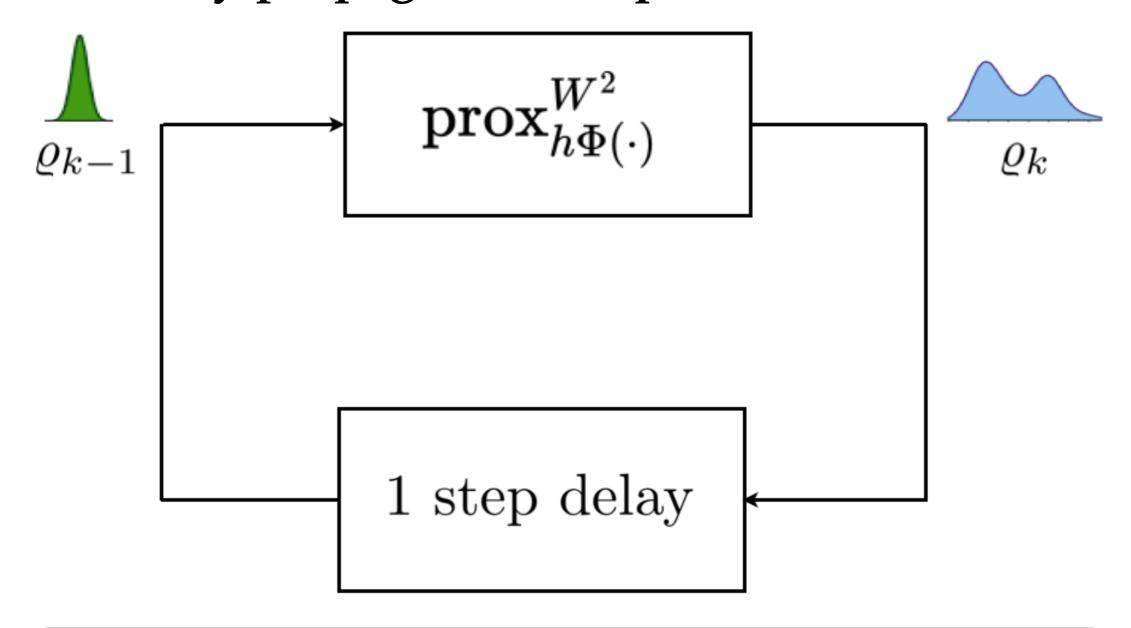
Gradient Flow in ${\mathcal X}$



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



Uncertainty propagation via point clouds



No spatial discretization or function approximation

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

$$\updownarrow \quad \text{Proximal Recursion}$$

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

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

$$\updownarrow \quad \text{Proximal Recursion}$$

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

Discrete Primal Formulation

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

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

$$\updownarrow \quad \text{Proximal Recursion}$$

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

Discrete Primal Formulation

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

Entropic Regularization

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

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

↑ Proximal Recursion

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

Discrete Primal Formulation

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

Entropic Regularization

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

$$\uparrow \quad \text{Dualization}$$

$$m{\lambda}_0^{ ext{opt}}, m{\lambda}_1^{ ext{opt}} = rg \max_{m{\lambda}_0, m{\lambda}_1 \geq 0} \left\{ \langle m{\lambda}_0, m{arrho}_{k-1}
angle - m{F}^{\star}(-m{\lambda}_1) - m{\epsilon} \left(\exp(m{\lambda}_0^{\top} h/\epsilon) \exp(-m{C}_k/2\epsilon) \exp(m{\lambda}_1 h/\epsilon) \right) \right\}$$

Fixed Point Recursion

$$\mathbf{y} = e^{\frac{\lambda_0^*}{\epsilon}h} \int \mathbf{z} = e^{\frac{\lambda_1^*}{\epsilon}h}$$

Coupled Transcendental Equations in y and z

$$\Gamma_{k} = e^{\frac{-C_{k}}{2\epsilon}} \longrightarrow y \odot \Gamma_{k} z = \varrho_{k-1}$$

$$\varrho_{k-1} \longrightarrow \varrho_{k} = z \odot \Gamma_{k}^{\mathsf{T}} y$$

$$\xi_{k-1} = e^{-\beta \psi_{k-1}} \longrightarrow z \odot \Gamma_{k}^{\mathsf{T}} y = \xi_{k-1} \odot z^{-\beta \epsilon/2h}$$

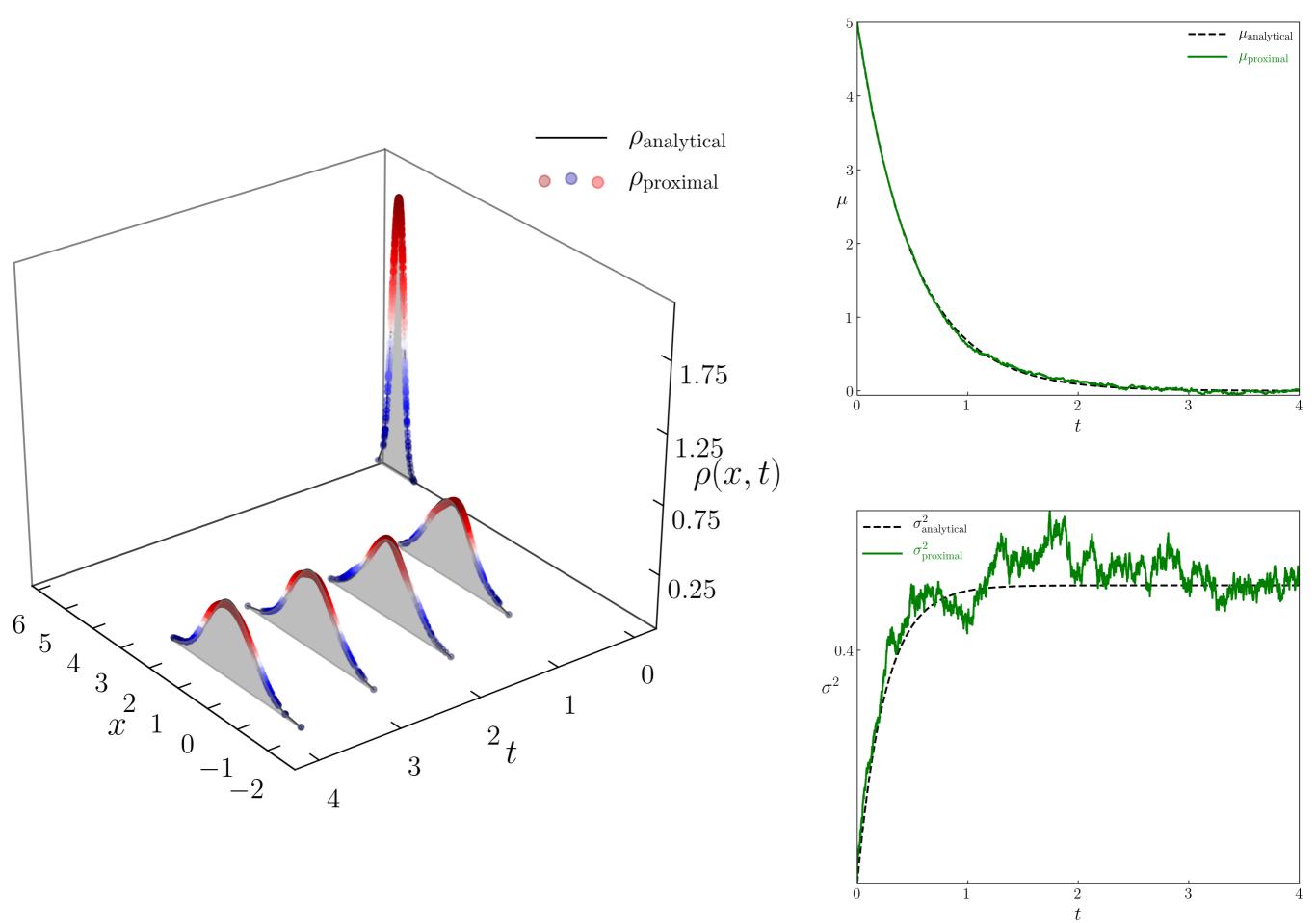
Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

$$m{y}\odot(m{\Gamma}_km{z})=m{arrho}_{k-1},\quad m{z}\odotig(m{\Gamma}_k^{\ \ }m{y}ig)=m{\xi}_{k-1}\odotm{z}^{-rac{eta\epsilon}{h}},$$

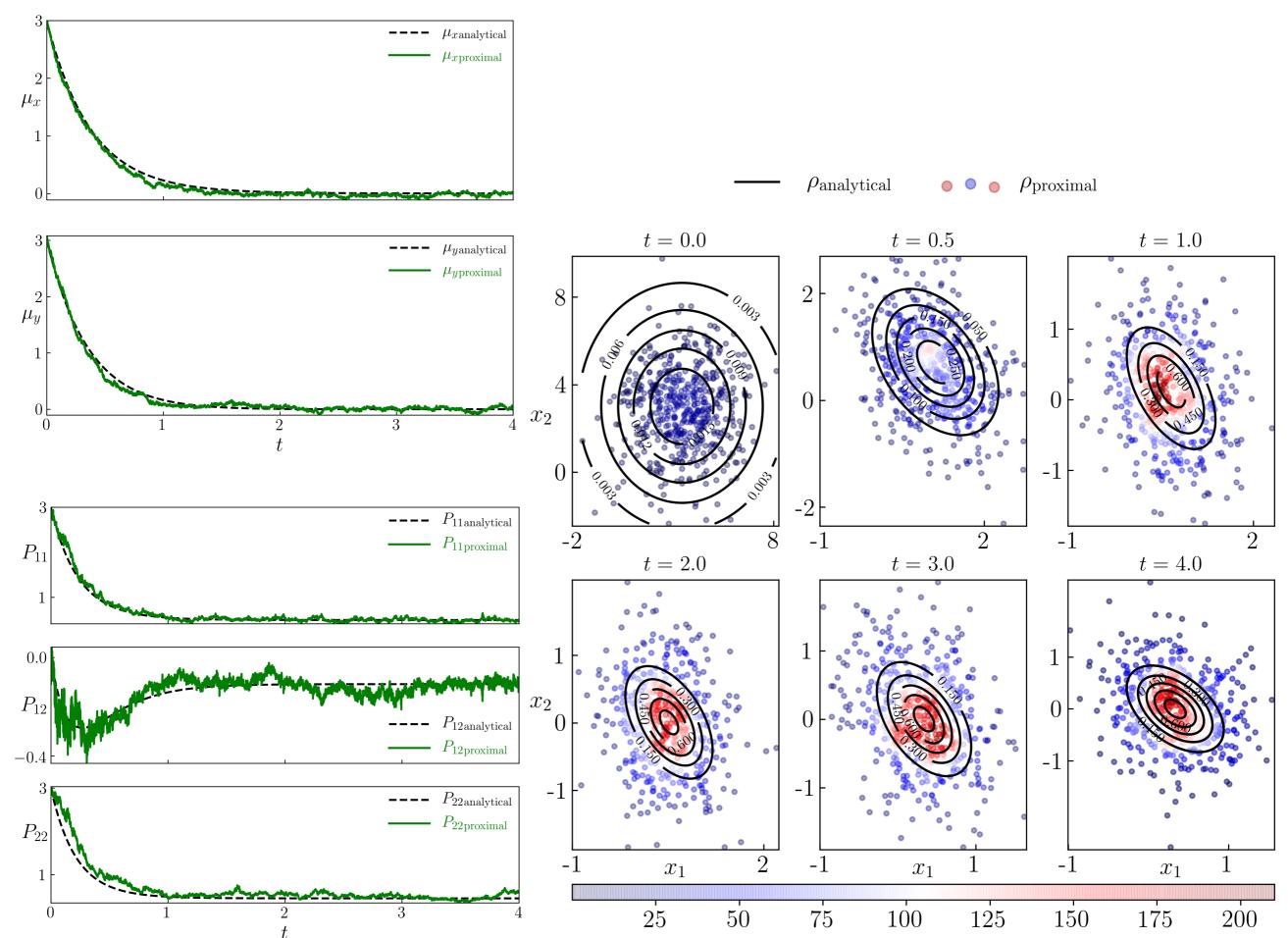
Then the solution (\pmb{y}^*, \pmb{z}^*) gives the proximal update $\pmb{\varrho}_k = \pmb{z}^* \odot (\pmb{\Gamma}_k^{\ \ } \pmb{y}^*)$

Definition 1: The Kullback-leiplentheir as a variational recursion 1: The Kullback-leiplentheir respects by $\mathbf{x} = \mathbf{x} + \mathbf{x} +$ ptimization literatures the imagning line rature, the trapping ay_{k-1} $\mapsto x_k$, given by $D_{\mathrm{KI}} (\mathrm{d}\pi_1 \parallel \mathrm{d}\pi_2) := \int_{\mathrm{KO}} \mathrm{d}x \log \frac{\rho_1(x)}{\rho_2(x)} \mathrm{d}x$, $\rho_1(x) = \int_{\mathrm{KO}} \mathrm{d}x \log \frac{\rho_1(x)}{\rho_2(x)} \mathrm{d}x$. $(x_{k-1}) := rg \min_{x} \sum_{k=1}^{n} ||x_{k-x}||_{T} ||x_{k-x}$ However, (11) is not all nyearier, sind it is the inherrisy sin the "proximal of second or the "proximal presented of and the second of the orange in equality the triangle in nerated by the proximat necession the proximal recursion Definition 2: The 2-Waskeritiem in the between the $x_k = \text{prox}_{h\varphi}^{\|\cdot\|}(x_{k-1}), \quad x_k = \text{prox}_{h\varphi}^{\|\cdot\|}(x_{k-1}), \quad x_{k-1}, \quad x$ $(\varrho_{k-1}) := \underset{\varrho \in \mathscr{D}_2}{\arg\inf} \frac{1}{2} d^2(\varrho, \varrho_{k-1}) := \underset{\varrho \in \mathscr{D}_2}{\arg\inf} \frac{1}{2} d^2(\varrho, \varrho_{k-1}) + \underset{\varrho \in \mathscr{D}_2}{h} \Phi(\varrho),$ $W(\pi_1,\pi_2) :=$ $W(\pi_1, \pi_2) :=$ $\int_{\mathbf{y}} \left(\inf_{\mathbf{y} \in \mathbf{y}} \left(\frac{\inf_{\mathbf{y}, \pi_{\mathbf{0}}} \int_{\mathbf{y}} \left(\mathbf{y}, \mathbf{y} \right) \right) \mathbf{z}^{\frac{1}{2}} - \mathbf{y} \right)$ where $\Pi(\pi_1, \pi_2)$ denotes the co 3) converges to the flow of the PDF (x) i.e. the kh as the step-size where if (π_1, π_2) denotes the sollection of all appoints satisfies $\varrho_k(x)_h \rightarrow 0$ we t_{also} while asathe step size dimensional case, he e also 1 he Han **Theorem:** Block co-ordinate iteration of (y, z) recur- $\frac{\mathrm{d}}{\mathrm{d}t}\varphi =$ +20foW6 sion is contractive on $\mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0}$. nattoWp(2 plies arphis a me the particular the pa

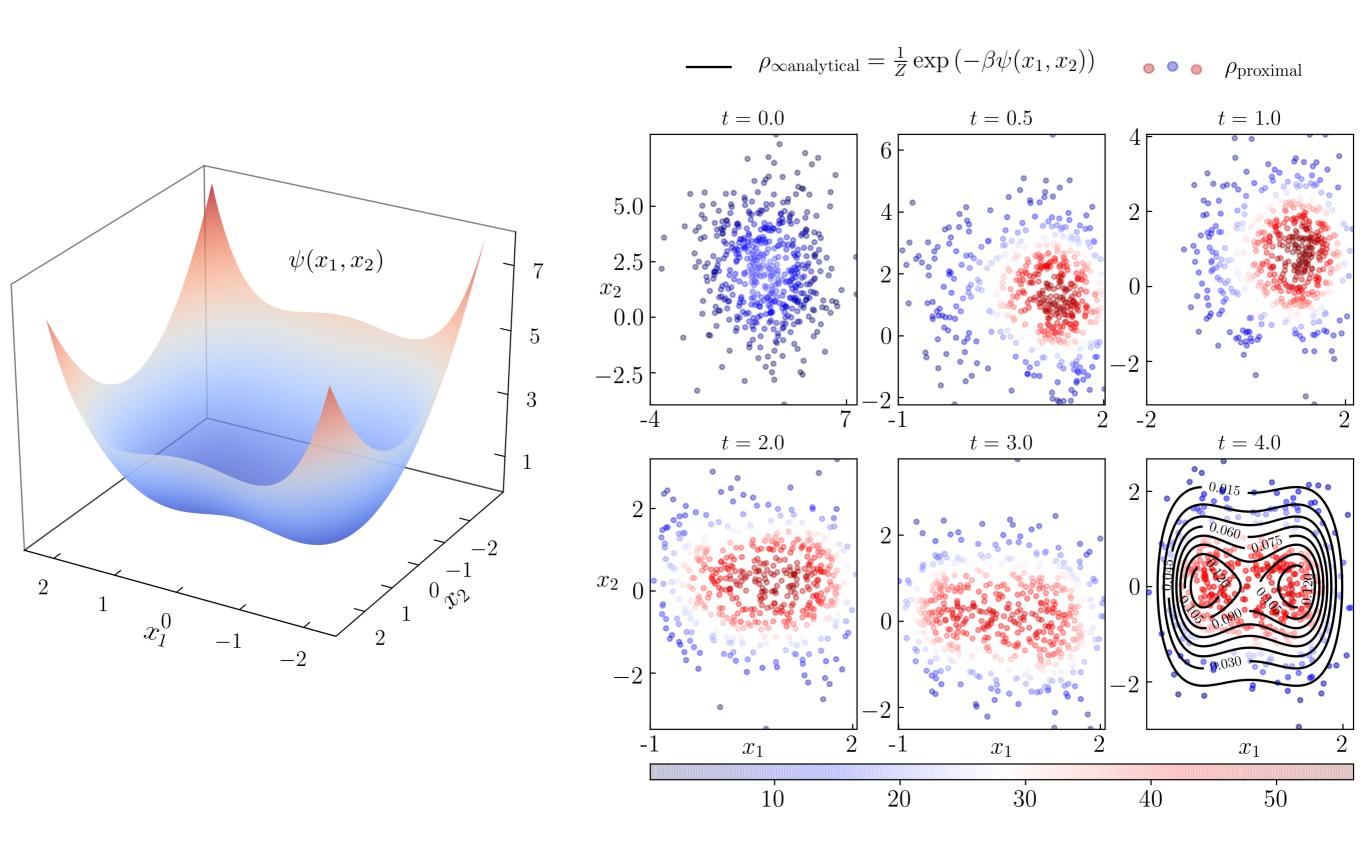
Proximal Prediction: 1D Linear Gaussian



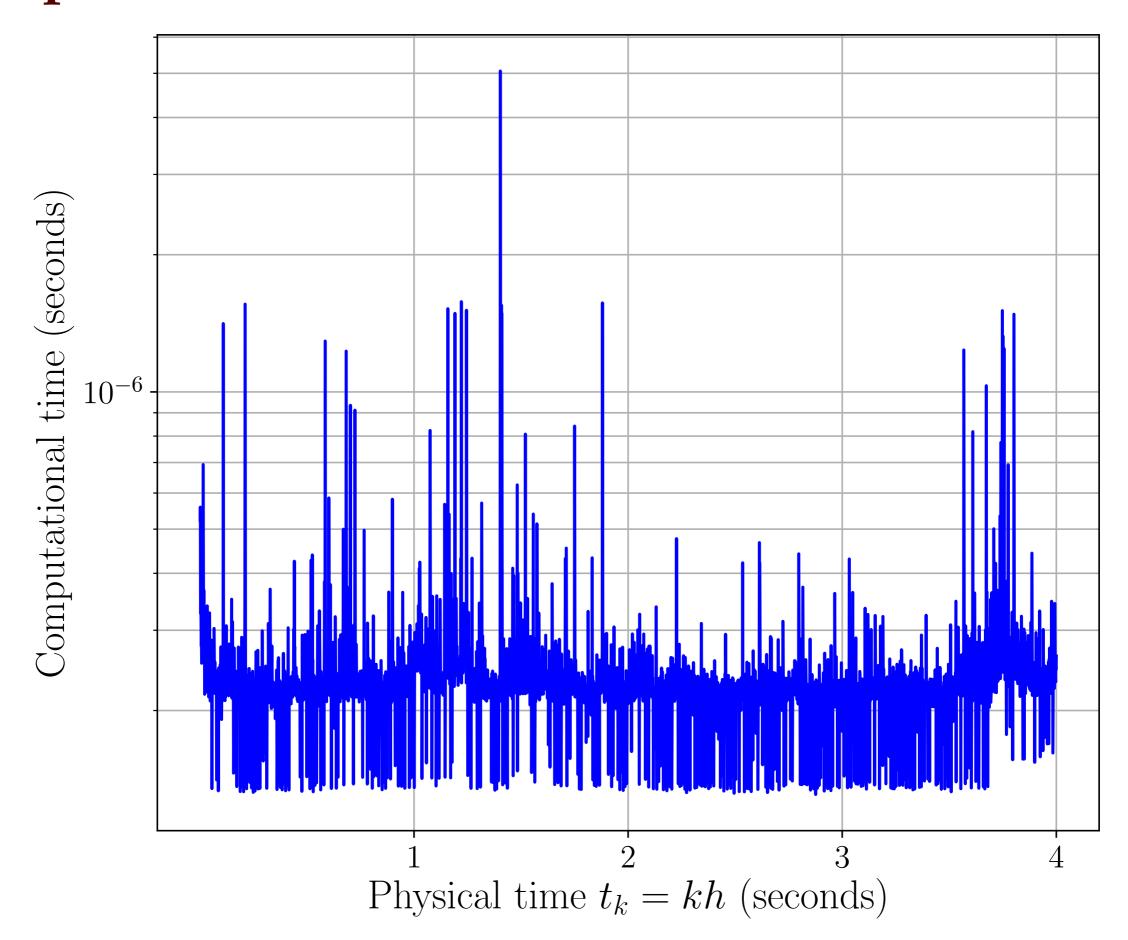
Proximal Prediction: 2D Linear Gaussian



Proximal Prediction: 2D Nonlinear Non-Gaussian



Computational Time: 2D Nonlinear Non-Gaussian



Proximal Prediction: Satellite in Geocentric Orbit

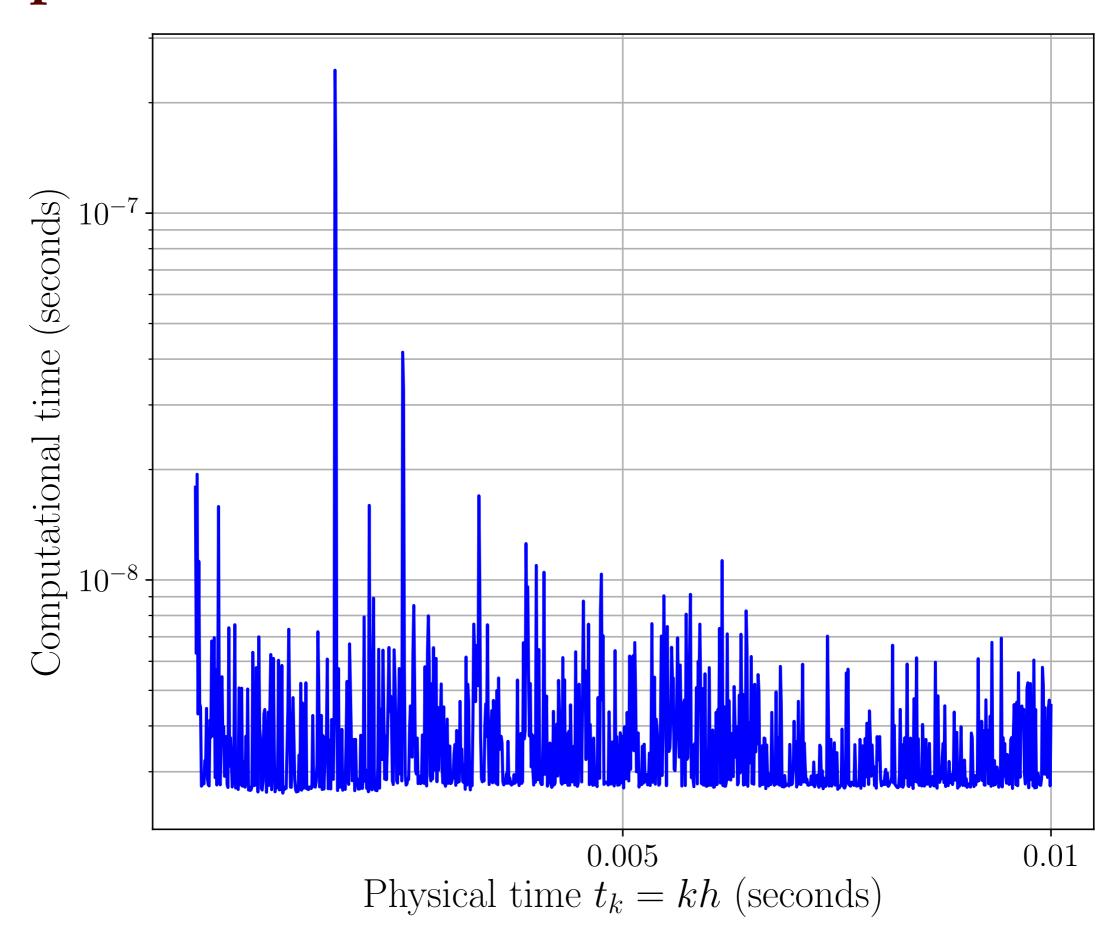
Here, $\mathcal{X} \equiv \mathbb{R}^6$

$$\frac{dx}{dy} \\
\frac{dy}{dz} \\
\frac{dv_{x}}{dv_{y}} \\
\frac{dv_{y}}{dv_{z}} = \begin{pmatrix} v_{x} \\ v_{y} \\ -\frac{\mu x}{r^{3}} + (f_{x})_{pert} - \gamma v_{x} \\ -\frac{\mu y}{r^{3}} + (f_{y})_{pert} - \gamma v_{y} \\ -\frac{\mu z}{r^{3}} + (f_{z})_{pert} - \gamma v_{z} \end{pmatrix} dt + \sqrt{2\beta^{-1}\gamma} \begin{pmatrix} 0 \\ 0 \\ 0 \\ dw_{1} \\ dw_{2} \\ dw_{3} \end{pmatrix},$$

$$\frac{dr}{dr} = \begin{pmatrix} s\theta & s\phi & s\theta & s\phi & -s\phi \end{pmatrix} \begin{pmatrix} \frac{k}{r^{4}} (3(s\theta)^{2} - 1) \\ \frac{k}{r^{4}} (3(s\theta)^{2} - 1) \end{pmatrix}$$

$$\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}_{\mathsf{pert}} = \begin{pmatrix} s\theta \ c\phi & c\theta \ c\phi & -s\phi \\ s\theta \ s\phi & c\theta \ s\phi & c\phi \\ c\theta & -s\theta & 0 \end{pmatrix} \begin{pmatrix} \frac{k}{2r^4} \left(3(s\theta)^2 - 1\right) \\ -\frac{k}{r^5}s\theta \ c\theta \\ 0 \end{pmatrix}, k := 3J_2R_{\mathrm{E}}^2, \mu = \mathsf{constant}$$

Computational Time: Satellite in Geocentric Orbit



Extensions: Nonlocal interactions

PDF dependent sample path dynamics:

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

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

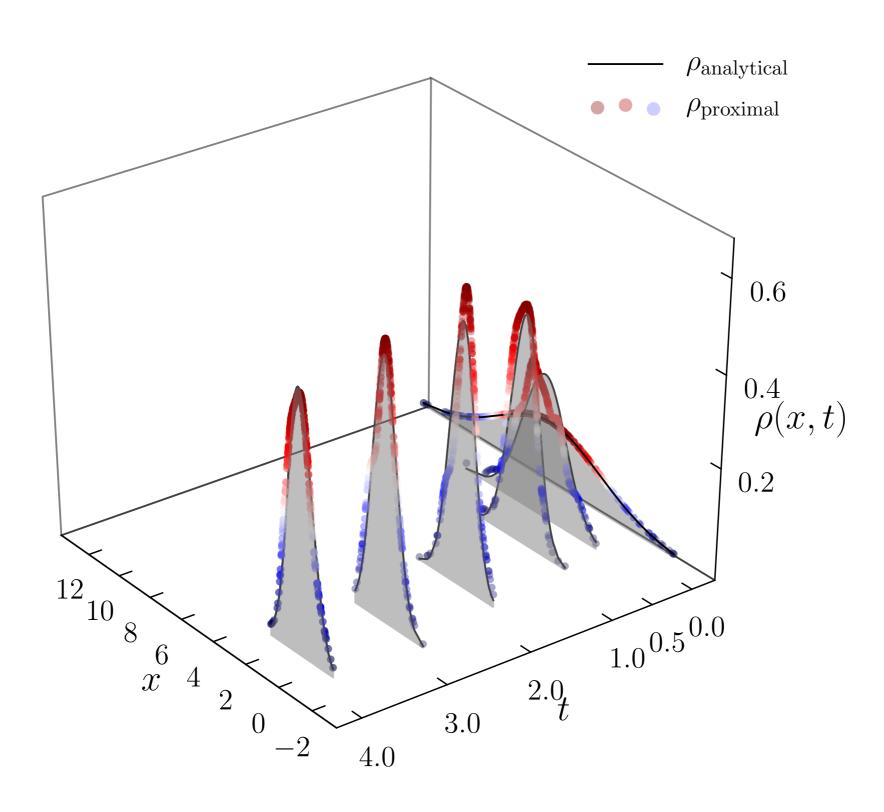
$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla (U + \rho * V)) + \beta^{-1} \Delta \rho$$

Free energy:

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

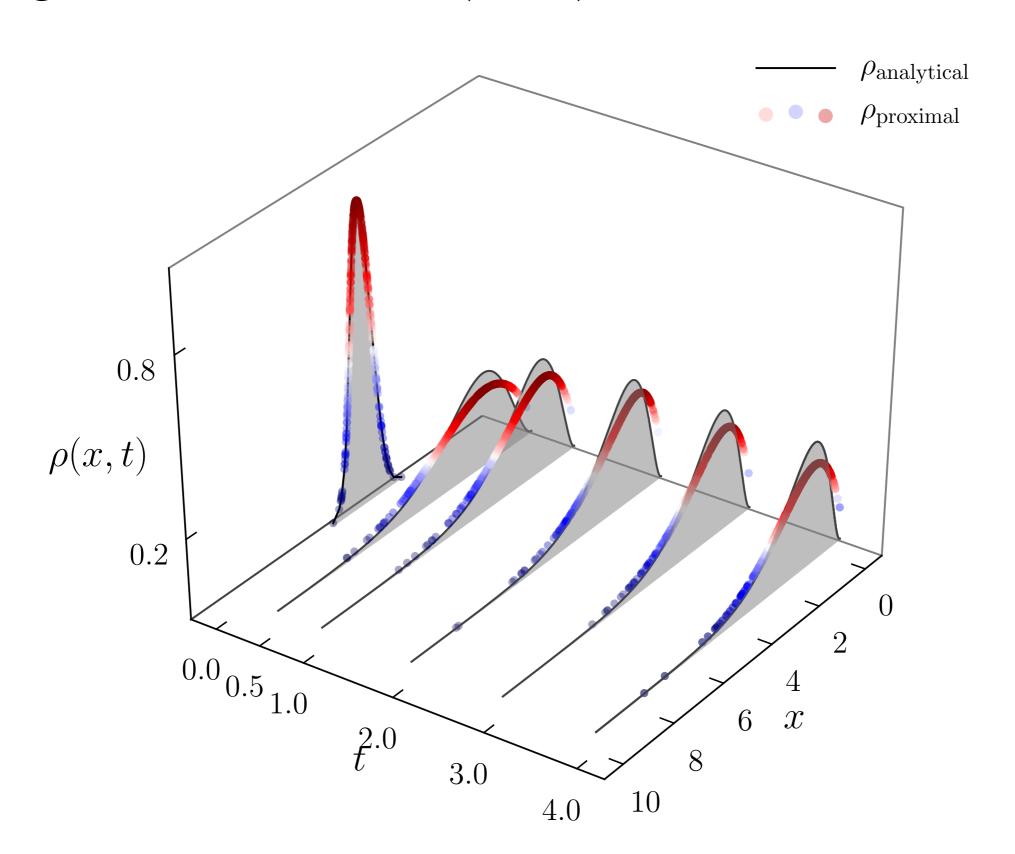
Extensions: Nonlocal interactions (contd.)

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



Extensions: Multiplicative Noise

Cox-Ingersoll-Ross: $dx = a(\theta - x) dt + b\sqrt{x} dw$, $2a > b^2$, $\theta > 0$



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

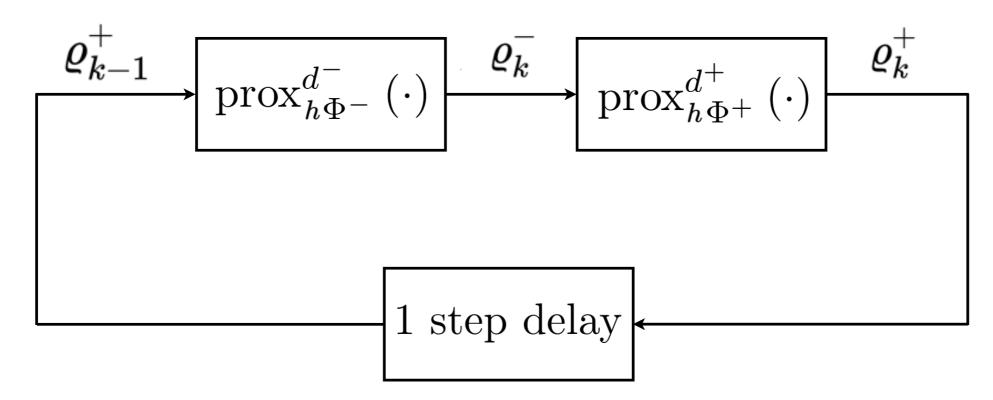
Solving filtering problem as Wasserstein gradient flow

What's New?

Main idea: Solve the Kushner-Stratonovich SPDE

$$\mathrm{d}
ho^+ = ig[\mathcal{L}_{\mathrm{FP}}\mathrm{d}t + \mathcal{L}ig(\mathrm{d}z,\mathrm{d}t,
ho^+ig)ig]
ho^+,\;
ho(x,t=0) =
ho_0 ext{ as gradient flow in } \mathcal{P}_2(\mathcal{X})$$

Recursion of {deterministic o stochastic} proximal operators:



Convergence: $\varrho_k^+(h) o
ho^+(x,t=kh)$ as $h\downarrow 0$

For prior, as before: $d^- \equiv W^2, \quad \Phi^- \equiv \ \mathbb{E}_{arrho} ig[\psi + eta^{-1} \log arrhoig]$

For posterior: $d^+ \equiv d_{ ext{FR}}^2 ext{ or } D_{ ext{KL}}, \quad \Phi^+ \equiv \; rac{1}{2} \mathbb{E}_{arrho^+} \Big[\left(y_k - h(x)
ight)^ op R^{-1} (y_k - h(x)) \Big]$

Explicit Recovery of 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:

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

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

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

- 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:

$$egin{aligned} x(t) &\sim \operatorname{Markov}(Q), \ \operatorname{d}\!z(t) &= h(x(t)) \ \operatorname{d}\!t \, + \, \sigma_v(t) \operatorname{d}\!v(t) \end{aligned}$$

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

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

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

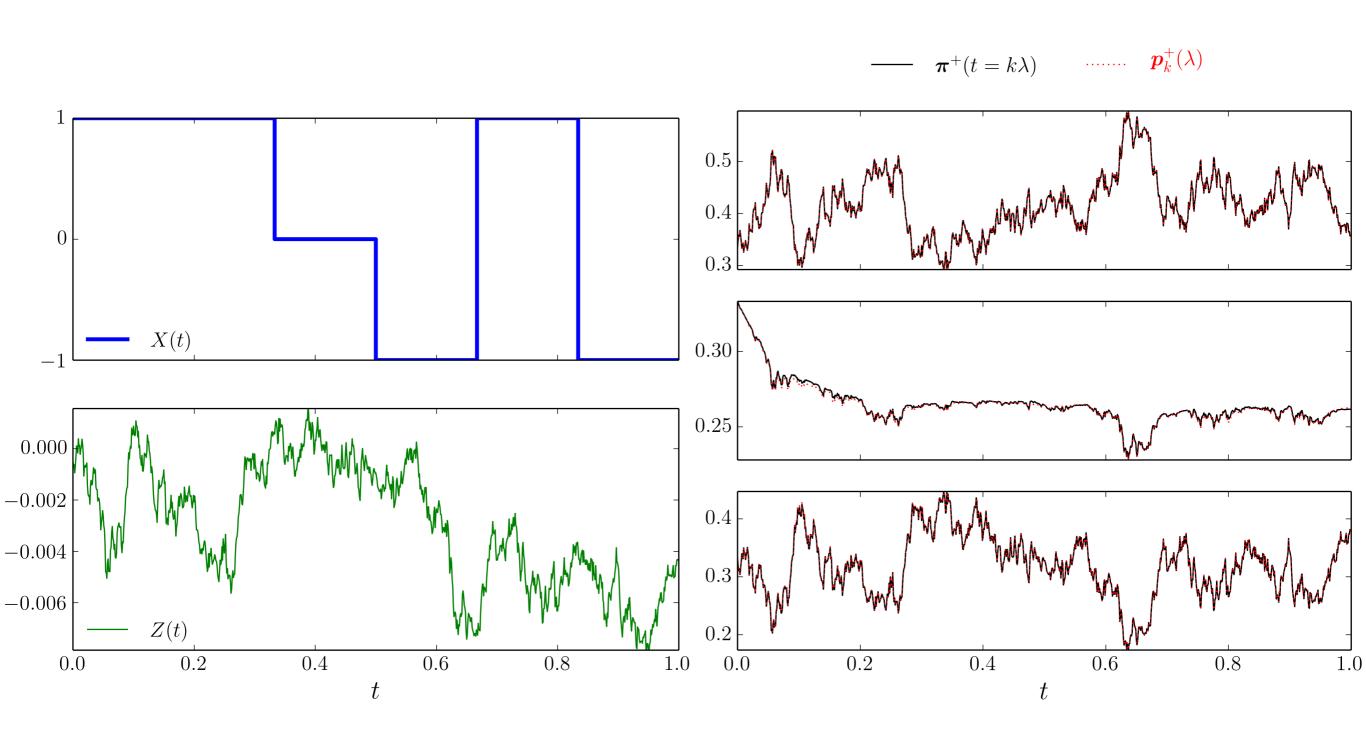
where
$$H:=\operatorname{diag}(h(a_1),\ldots,h(a_m)), \quad \widehat{h}(t):=\sum_{i=1}^m h(a_i)\pi_i^+(t),$$

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

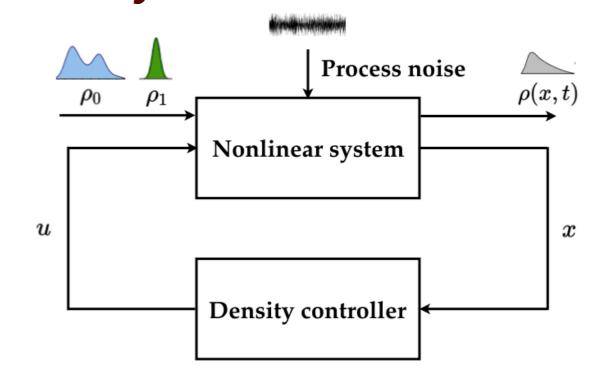


— A.H. and T.T. Georgiou, Proximal Recursion for the Wonham Filter, CDC 2019.

Solving density control as Wasserstein gradient flow

Finite Horizon Feedback Density Control

$$egin{aligned} & ext{minimize} \ \mathbb{E}igg[\int_0^1 \|u\|_2^2 \ \mathrm{d}tigg] \ & ext{subject to} \ & ext{d}x = f(x,u,t) \ \mathrm{d}t \, + \, g(x,t) \ \mathrm{d}w, \ & x(t=0) \sim
ho_0, \quad x(t=1) \sim
ho_1 \end{aligned}$$



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

$$g(x,t) = \sqrt{2\epsilon}B(t)$$

System of coupled Nonlinear PDEs (FPK + HJB):

$$rac{\partial
ho}{\partial t} +
abla \cdot \left(
ho ig(f + B(t)^{\! op}
abla \psi ig)
ight) = \epsilon \mathbf{1}^{\! op} (D(t) \odot \operatorname{Hess}(
ho)) \mathbf{1}$$

$$rac{\partial \psi}{\partial t} + rac{1}{2} \parallel B(t)^{\! op}
abla \psi \parallel_2^2 + \langle
abla \psi, f
angle = -\epsilon \langle D(t), \operatorname{Hess}(\psi)
angle \ .$$

Boundary conditions:

$$ho(x,0)=
ho_0,\;
ho(x,1)=
ho_1$$

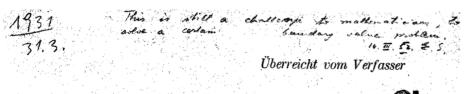
Optimal control:

$$u(x,t) = B(t)^{\! op}
abla \psi(x,t)$$

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:



ÜBER DIE UMKEHRUNG DER NATURGESETZE



VON

E. SCHRÖDINGER

SONDERAUSGABE AUS DEN SITZUNGSBERICHTEN
DER PREUSSISCHEN AKADEMIE DER WISSENSCHAFTEN
PHYS.-MATH. KLASSE, 1981. IX

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

PAI

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 $(\rho, \psi) \mapsto (\varphi, \hat{\varphi})$

$$arphi(x,t) = \expigg(rac{\psi(x,t)}{2\epsilon}igg),$$

$$\hat{arphi}(x,t) =
ho(x,t) \expigg(-rac{\psi(x,t)}{2\epsilon}igg),$$

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

Optimal control: $u^*(x,t) = 2\epsilon B(t)^{\! op}
abla \log arphi(x,t)$

Solution via Schrödinger System

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

Let
$$D(t) := B(t)B(t)^{\! op}$$

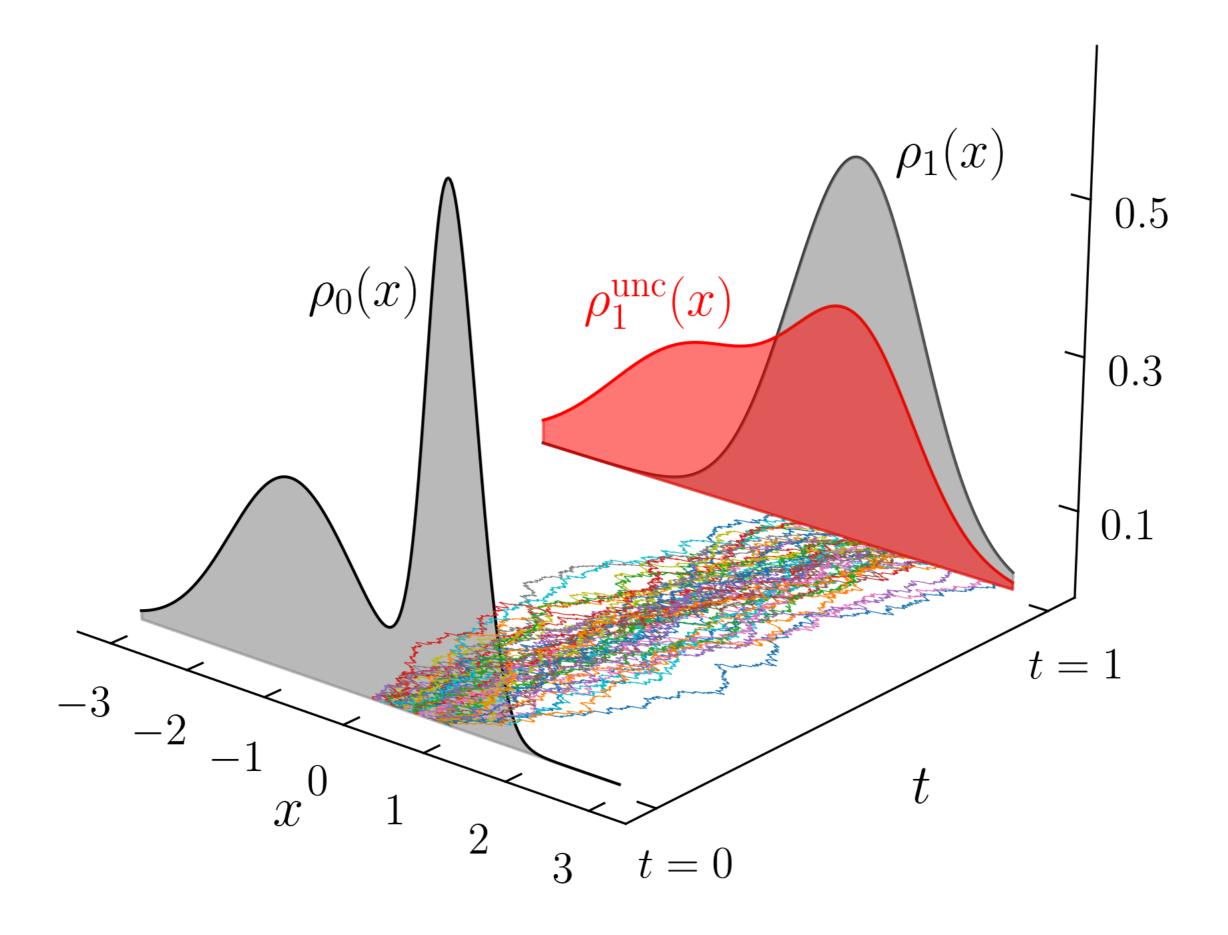
$$rac{\partial arphi}{\partial t} = -\langle
abla arphi, f
angle - \epsilon \langle D(t), \operatorname{Hess}(arphi)
angle, \qquad arphi_0 \hat{arphi}_0 =
ho_0,$$

$$rac{\partial \hat{arphi}}{\partial t} = -
abla \cdot (\hat{arphi} f) + \epsilon \mathbf{1} \cdot (D(t) \odot \operatorname{Hess}(\hat{arphi})) \mathbf{1}, \qquad arphi_1 \hat{arphi}_1 =
ho_1,$$

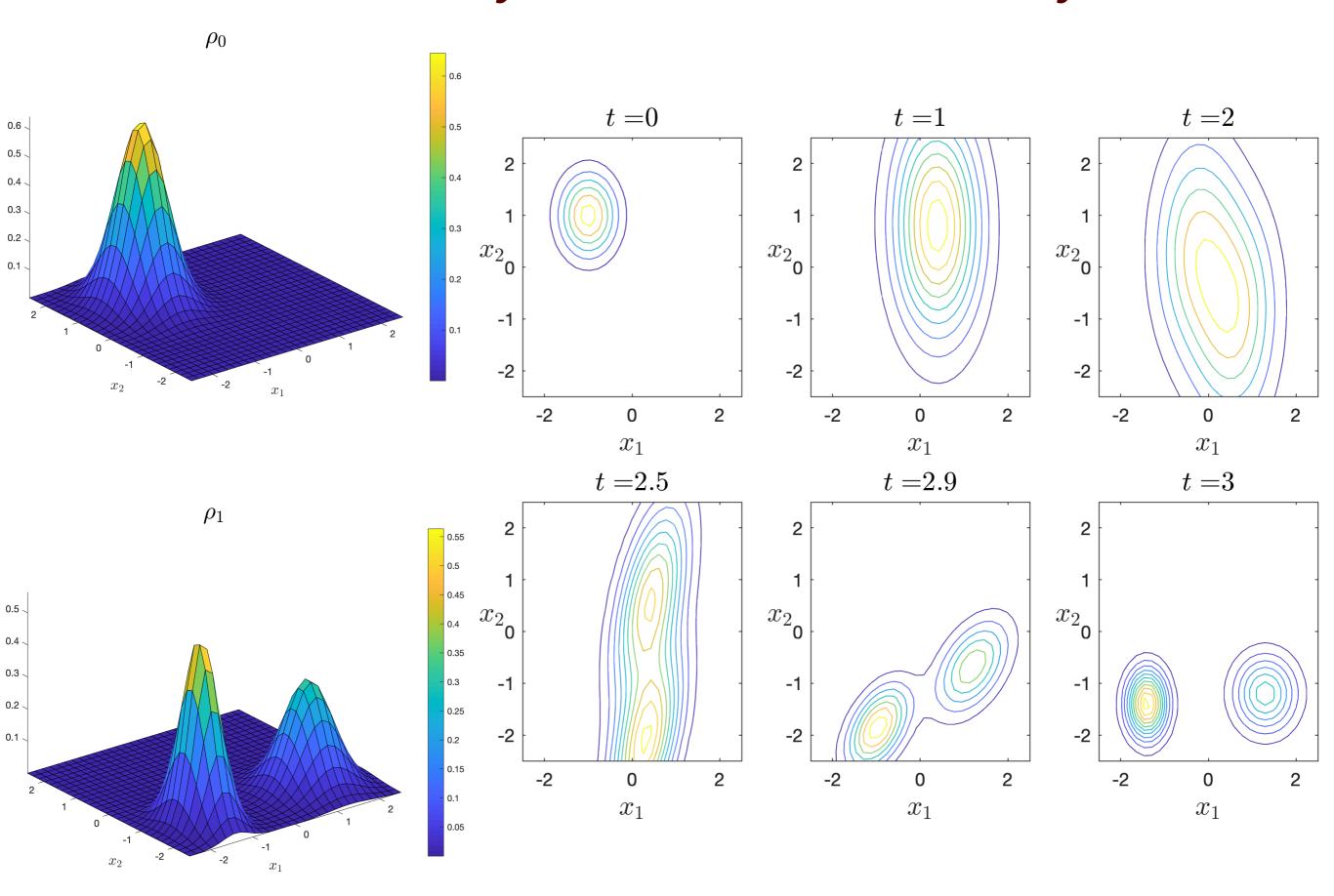
Wasserstein proximal algorithm \longrightarrow fixed point recursion over $(\varphi_1, \hat{\varphi}_0)$

(Contractive in Hilbert metric)

Feedback Density Control: Zero Prior Dynamics

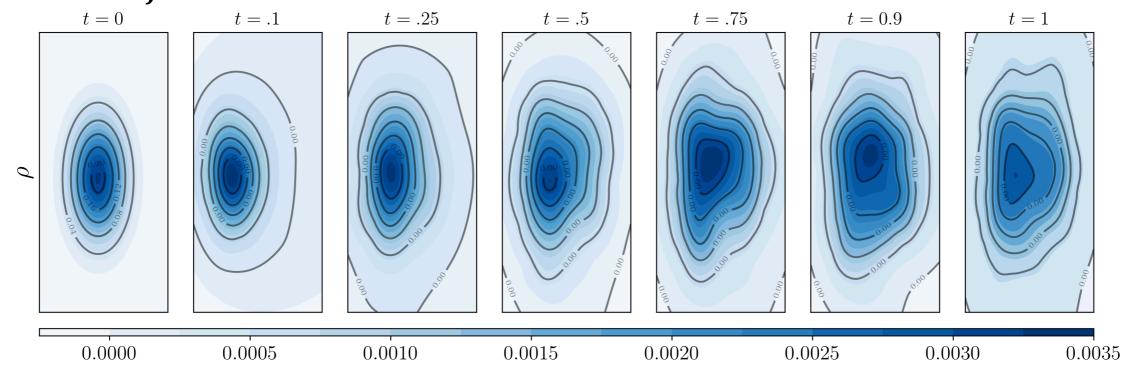


Feedback Density Control: LTI Prior Dynamics

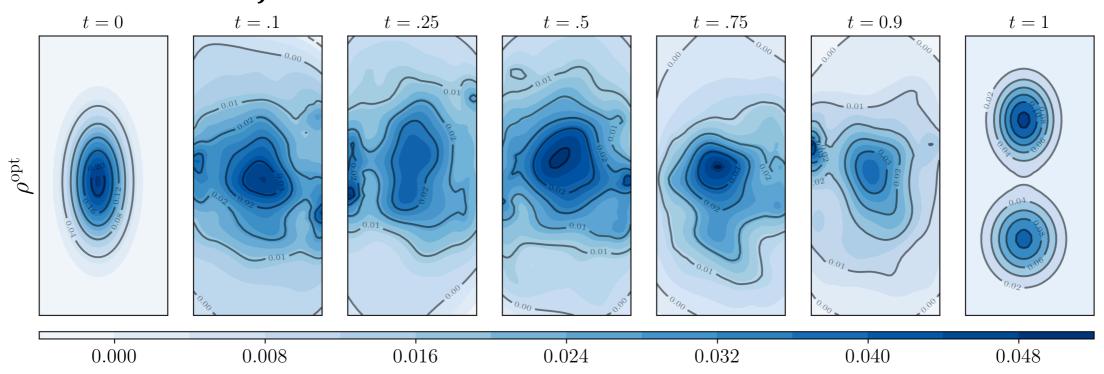


Feedback Density Control: Nonlinear Gradient Prior Dynamics

Uncontrolled joint PDF evolution:



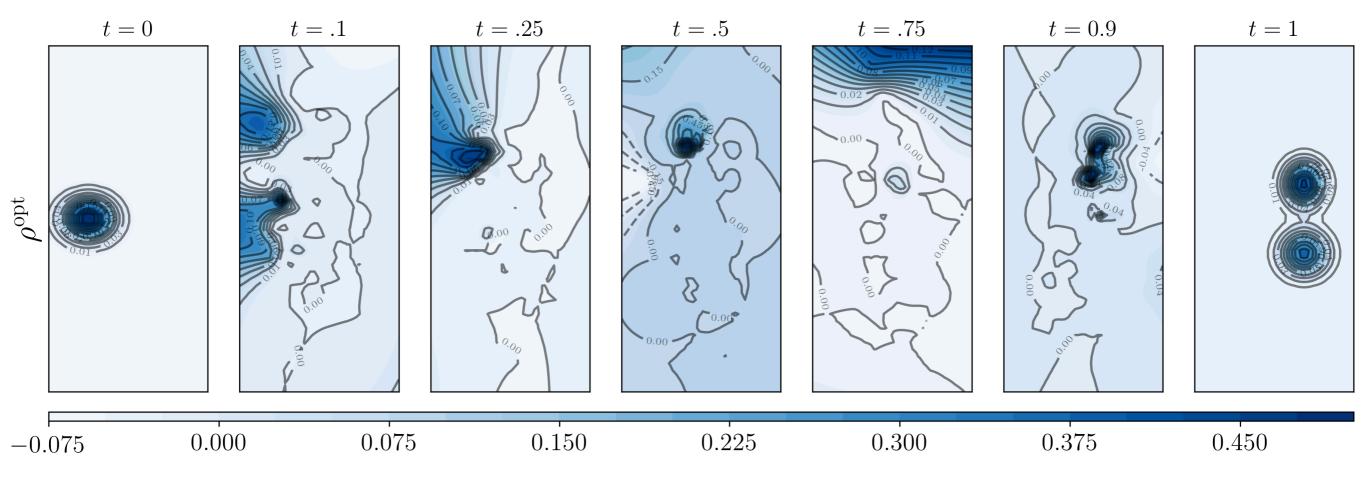
Optimal controlled joint PDF evolution:



Feedback Density Control: Nonlinear Mixed Conservative-Dissipative Prior Dynamics

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

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.

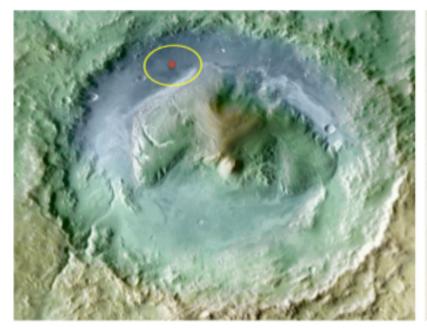
Take Home Message

Emerging systems-control theory of densities

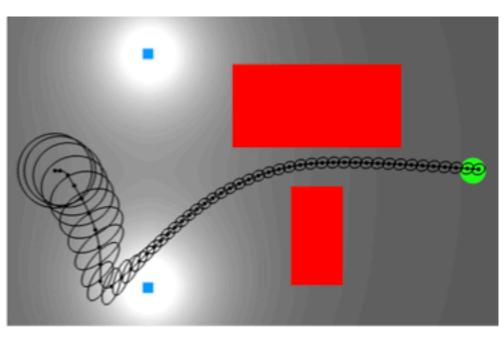
Three problems: prediction, filtering, control

One unifying framework: Wasserstein gradient flow

Many applications:







Thank You