Proximal Recursion for Solving the Fokker-Planck Equation

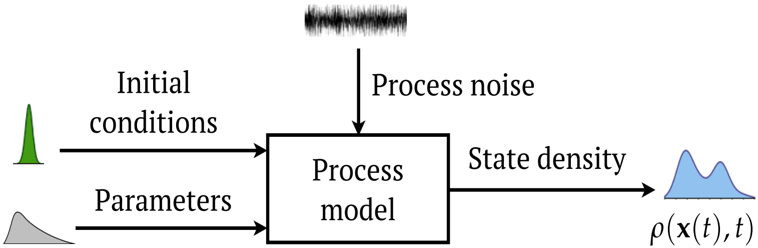
Kenneth Caluya

Joint work with Abhishek Halder

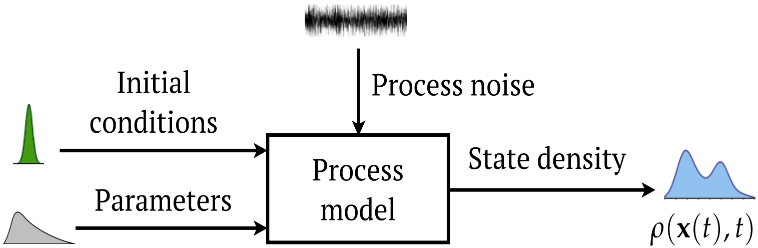
University of California, Santa Cruz

2019 American Control Conference
Philadelphia, July 12, 2019

Problem: Density Propagation



Problem: Density Propagation



State Dynamics:

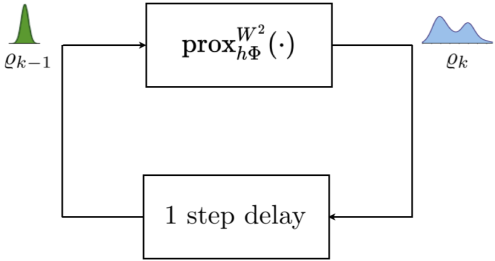
$$\mathrm{d}\boldsymbol{x} = -\nabla\psi(\boldsymbol{x}, t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}\boldsymbol{w}, \;\; \boldsymbol{x}_0 \sim \rho_0, \;\; \mathrm{d}\boldsymbol{w}(t) \sim \mathcal{N}(0, \boldsymbol{I}\mathrm{d}t), \;\; \boldsymbol{x} \in \mathbb{R}^n$$

Probability Density Function (PDF) Dynamics:

$$\frac{\partial\rho}{\partial t} = \mathcal{L}\rho := \nabla\cdot(\nabla\psi\rho) + \beta^{-1}\Delta\rho, \quad \rho(\boldsymbol{x}, 0) = \rho_0$$

What's new?

Main Idea: Solve $\dfrac{\partial \rho}{\partial t} = \mathcal{L}\rho, \ \rho(\boldsymbol{x}, 0) = \rho_0$ **as gradient flow in** $\mathcal{P}_2(\mathbb{R}^n)$



Proximal Operator: $\rho_k = \text{prox}_{h\Phi}^{W^2}(\rho_{k-1}) := \underset{\rho \in \mathcal{P}_2(\mathbb{R}^n)}{\arg\inf} \left\{ \frac{1}{2} W^2(\rho, \rho_{k-1}) + h\, \Phi(\rho) \right\}$

Optimal Transport Cost: $W^2(\rho, \rho_{k-1}) := \inf_{\pi \in \Pi(\rho, \rho_{k-1})} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\pi(\boldsymbol{x}, \boldsymbol{y})$

Free Energy Functional: $\Phi(\rho) := \int_{\mathbb{R}^n} \psi \rho \, \mathrm{d}\boldsymbol{x} + \beta^{-1} \int_{\mathbb{R}^n} \rho \log \rho \, \mathrm{d}\boldsymbol{x}$

Gradient Flow

Gradient Flow in $\mathbb{R}^n$

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

Recursion:

$$\begin{aligned}\boldsymbol{x}_k &= \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k)\\ &= \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\left\{\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_{k-1}\|_2^2 + h\varphi(\boldsymbol{x})\right\}\\ &=: \mathrm{prox}_{h\varphi}^{\|\cdot\|_2}(\boldsymbol{x}_{k-1})\end{aligned}$$

Convergence:

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

Gradient Flow

Gradient Flow in $\mathbb{R}^n$

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

Recursion:

$$\begin{aligned}\boldsymbol{x}_k &= \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k)\\ &= \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\left\{\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}_{k-1}\|_2^2 + h\varphi(\boldsymbol{x})\right\}\\ &=: \operatorname{prox}_{h\varphi}^{\|\cdot\|_2}(\boldsymbol{x}_{k-1})\end{aligned}$$

Convergence:

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

Gradient Flow in $\mathcal{P}_2(\mathbb{R}^n)$

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

Recursion:

$$\begin{aligned}\rho_k &= \rho(\cdot, t = kh)\\ &= \underset{\rho\in\mathcal{P}_2(\mathbb{R}^n)}{\arg\min}\left\{\frac{1}{2}W^2(\rho, \rho_{k-1}) + h\Phi(\rho)\right\}\\ &=: \operatorname{prox}_{h\Phi}^{W^2}(\rho_{k-1})\end{aligned}$$

Convergence:

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

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**

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

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**

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

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**

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

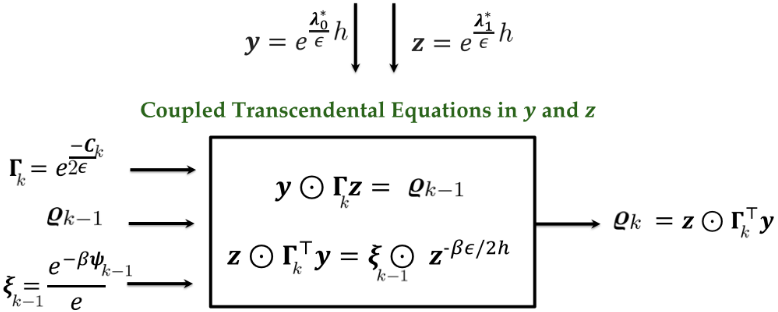
$\Downarrow$ **Entropic Regularization**

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

$\Updownarrow$ **Dualization**

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

Fixed Point Recursion

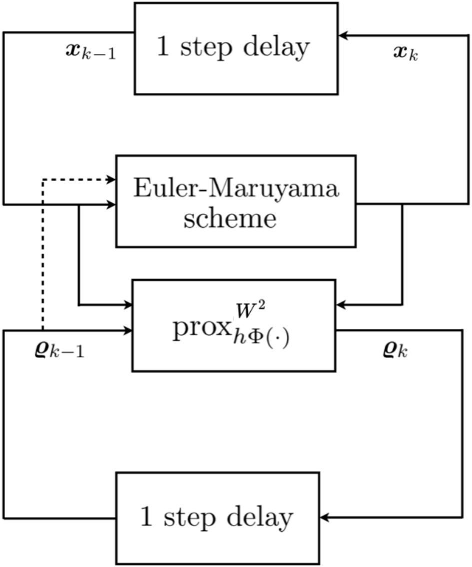


Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

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

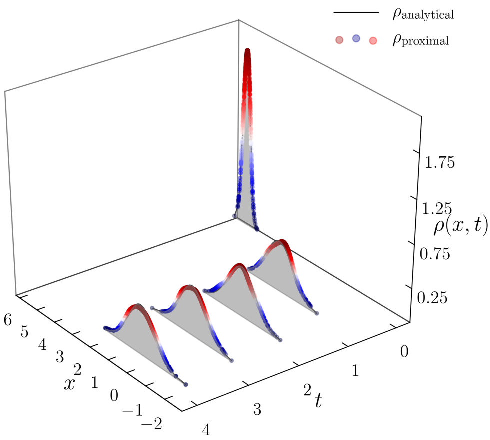
Then the solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ gives the proximal update $\varrho_k = \boldsymbol{z}^* \odot (\boldsymbol{\Gamma}_k{}^\top \boldsymbol{y}^*)$

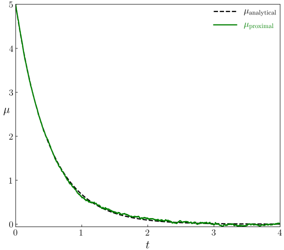
Algorithmic Setup

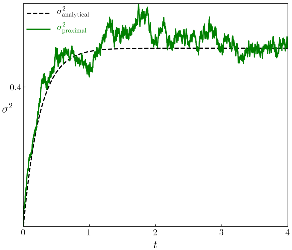


1D Linear Gaussian

$$\mathrm{d}x = -ax\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}w,\ a > 0$$

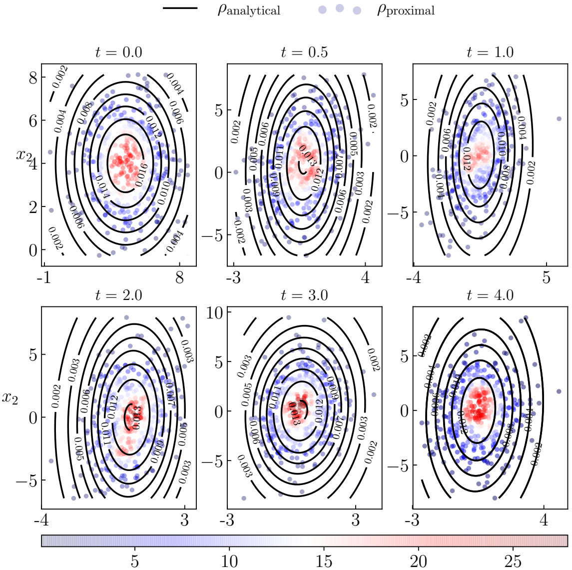






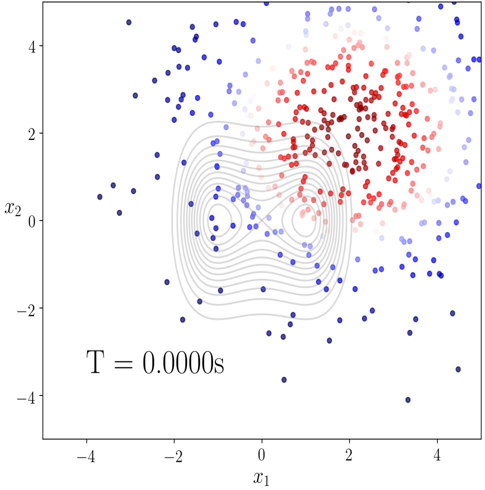
2D Linear Gaussian with Non-Gradient Drift

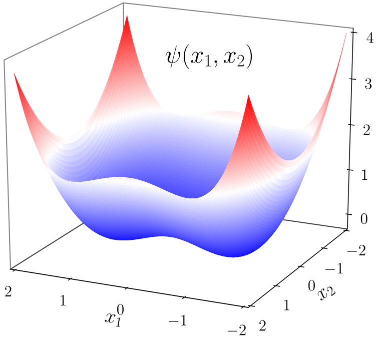
$$\mathrm{d}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}\,\mathrm{d}t + \boldsymbol{B}\,\mathrm{d}\boldsymbol{w}$$



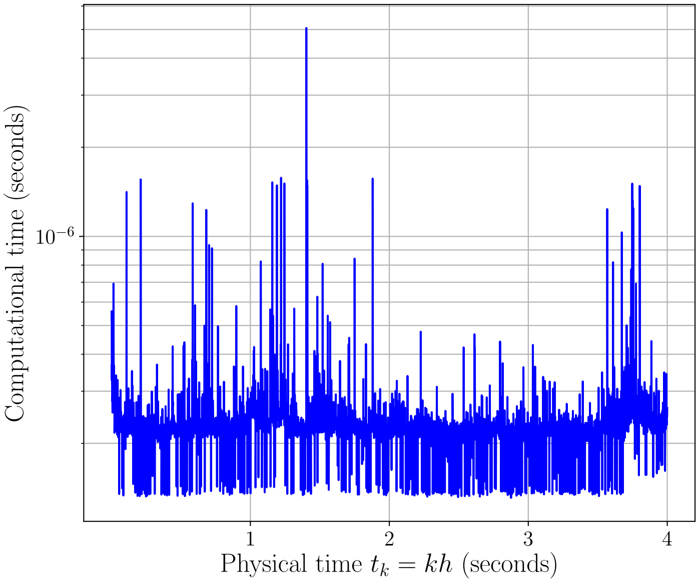
2D Nonlinear Non-Gaussian

$$\mathrm{d}\boldsymbol{x} = -\nabla\psi(\boldsymbol{x})\,\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}\boldsymbol{w},$$
$$\rho_\infty \propto \exp(-\beta\psi)$$





Computational Time for 2D Nonlinear Non-Gaussian



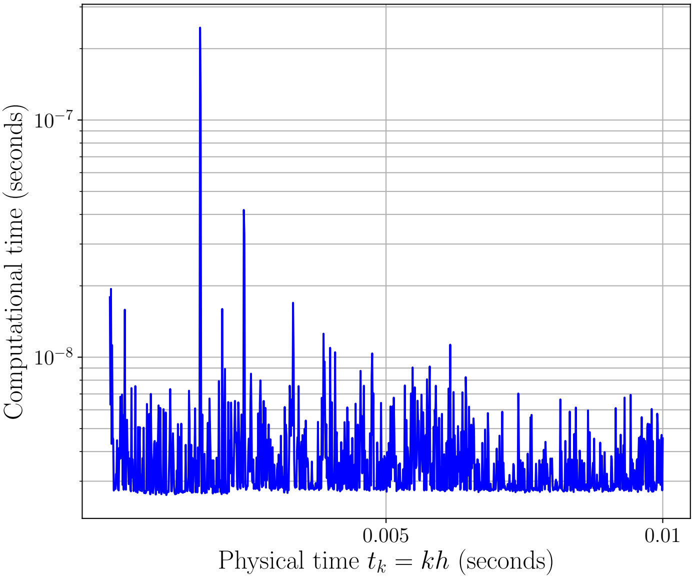
Mixed Conservative-Dissipative Drift

Relative motion of a satellite in geocentric orbit:

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

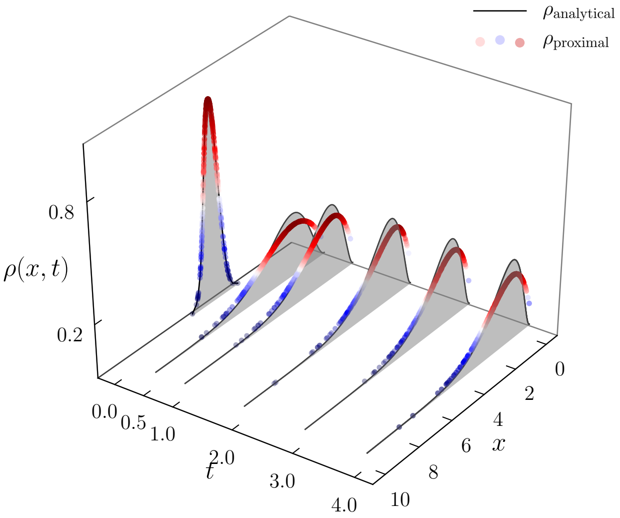
$$\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}_{\text{pert}} = \begin{pmatrix} s\theta\, c\phi & c\theta\, c\phi & -s\phi \\ s\theta\, s\phi & c\theta\, s\phi & c\phi \\ c\theta & -s\theta & 0 \end{pmatrix} \begin{pmatrix} \dfrac{k}{2r^4}\left(3(s\theta)^2 - 1\right) \\ -\dfrac{k}{r^5} s\theta\, c\theta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

Computational time for 6 State Satellite Motion



Extensions: Multiplicative Noise

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



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:

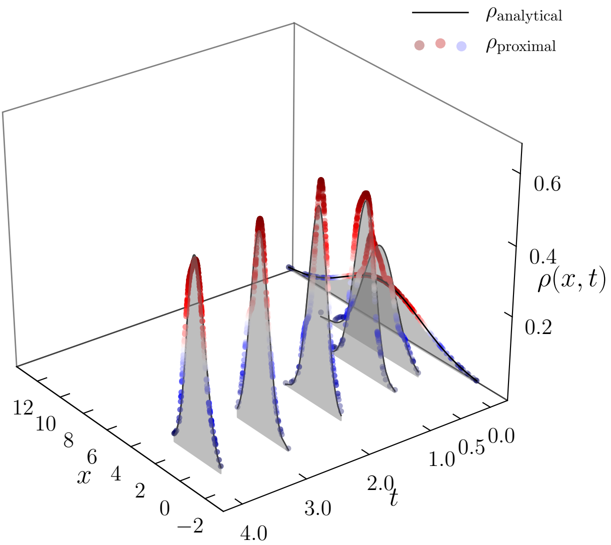
$$\frac{\partial\rho}{\partial t} = \nabla\cdot\left(\rho\nabla\left(U + \rho * V\right)\right) + \beta^{-1}\Delta\rho$$

Free energy:

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

Extensions: Nonlocal Interactions (contd.)

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



Application: Finite Time Density Steering

$$\inf_{u \in \mathcal{U}} \quad \mathbb{E}\left\{\int_0^1 \frac{1}{2}|u(\boldsymbol{x},t)|^2 \,\mathrm{d}t\right\}$$

$$\text{subject to} \quad \mathrm{d}\boldsymbol{x} = -\nabla\Psi(\boldsymbol{x},t)\,\mathrm{d}t + \; u(\boldsymbol{x},t)\,\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}\boldsymbol{w}$$

$$\boldsymbol{x}_0 \sim \rho_0 \quad \boldsymbol{x}_1 \sim \rho_1$$

Objective: Compute $u^{\mathrm{opt}}(\boldsymbol{x},t), \rho^{\mathrm{opt}}(\boldsymbol{x},t),\; u^{\mathrm{opt}}(\boldsymbol{x},t) = \nabla\lambda(\boldsymbol{x},t)$

Application: Finite Time Density Steering

$$\inf_{u \in \mathcal{U}} \quad \mathbb{E}\left\{\int_0^1 \frac{1}{2}|u(\boldsymbol{x},t)|^2 \, \mathrm{d}t\right\}$$

$$\text{subject to} \quad \mathrm{d}\boldsymbol{x} = -\nabla\Psi(\boldsymbol{x},t)\,\mathrm{d}t + u(\boldsymbol{x},t)\,\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}\boldsymbol{w}$$

$$\boldsymbol{x}_0 \sim \rho_0 \quad \boldsymbol{x}_1 \sim \rho_1$$

Objective: Compute $u^{\mathrm{opt}}(\boldsymbol{x},t), \rho^{\mathrm{opt}}(\boldsymbol{x},t)$, $u^{\mathrm{opt}}(\boldsymbol{x},t) = \nabla\lambda(\boldsymbol{x},t)$

Coupled Nonlinear PDE's

$$\frac{\partial\lambda}{\partial t} + \frac{|\nabla\lambda|^2}{2} - \nabla\lambda\cdot\nabla\Psi = -\epsilon\Delta\lambda$$

$$\frac{\partial\rho^{\mathrm{opt}}}{\partial t} + \nabla\cdot\left(\rho^{\mathrm{opt}}(\nabla\lambda - \nabla\Psi)\right) = \epsilon\Delta\rho^{\mathrm{opt}}$$

$$\rho^{\mathrm{opt}}(\boldsymbol{x},0) = \rho_0, \quad \rho^{\mathrm{opt}}(\boldsymbol{x},1) = \rho_1$$

Application: Finite Time Density Steering

$$\inf_{u\in\mathcal{U}} \quad \mathbb{E}\left\{\int_0^1 \frac{1}{2}|u(\boldsymbol{x},t)|^2\,\mathrm{d}t\right\}$$

$$\text{subject to} \quad \mathrm{d}\boldsymbol{x} = -\nabla\Psi(\boldsymbol{x},t)\,\mathrm{d}t + u(\boldsymbol{x},t)\,\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}\boldsymbol{w}$$

$$\boldsymbol{x}_0 \sim \rho_0 \quad \boldsymbol{x}_1 \sim \rho_1$$

Objective: Compute $u^{\mathrm{opt}}(\boldsymbol{x},t), \rho^{\mathrm{opt}}(\boldsymbol{x},t),\ u^{\mathrm{opt}}(\boldsymbol{x},t) = \nabla\lambda(\boldsymbol{x},t)$

Coupled Nonlinear PDE's

$$\frac{\partial\lambda}{\partial t} + \frac{|\nabla\lambda|^2}{2} - \nabla\lambda\cdot\nabla\Psi = -\epsilon\Delta\lambda$$

$$\frac{\partial\rho^{\mathrm{opt}}}{\partial t} + \nabla\cdot\left(\rho^{\mathrm{opt}}(\nabla\lambda - \nabla\Psi)\right) = \epsilon\Delta\rho^{\mathrm{opt}}$$

$$\rho^{\mathrm{opt}}(\boldsymbol{x},0) = \rho_0, \quad \rho^{\mathrm{opt}}(\boldsymbol{x},1) = \rho_1$$

Boundary Coupled Linear PDE's

$$\frac{\partial\varphi}{\partial t} = \nabla\varphi\cdot\nabla\Psi - \epsilon\Delta\varphi$$

$$\frac{\partial\hat{\varphi}}{\partial t} = \nabla\cdot(\hat{\varphi}\nabla\Psi) + \epsilon\Delta\hat{\varphi}$$

$$\rho_0 = \varphi_0\hat{\varphi}_0, \quad \rho_1 = \varphi_1\hat{\varphi}_1$$

Application: Finite Time Density Steering

$$\inf_{u \in \mathcal{U}} \quad \mathbb{E}\left\{\int_0^1 \frac{1}{2}|u(\boldsymbol{x},t)|^2 \,\mathrm{d}t\right\}$$

$$\text{subject to} \quad \mathrm{d}\boldsymbol{x} = -\nabla\Psi(\boldsymbol{x},t)\,\mathrm{d}t + u(\boldsymbol{x},t)\,\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}\boldsymbol{w}$$

$$\boldsymbol{x}_0 \sim \rho_0 \quad \boldsymbol{x}_1 \sim \rho_1$$

Objective: Compute $u^{\mathrm{opt}}(\boldsymbol{x},t), \rho^{\mathrm{opt}}(\boldsymbol{x},t),\ u^{\mathrm{opt}}(\boldsymbol{x},t) = \nabla\lambda(\boldsymbol{x},t)$

Coupled Nonlinear PDE's

$$\frac{\partial\lambda}{\partial t} + \frac{|\nabla\lambda|^2}{2} - \nabla\lambda\cdot\nabla\Psi = -\epsilon\Delta\lambda$$

$$\frac{\partial\rho^{\mathrm{opt}}}{\partial t} + \nabla\cdot\left(\rho^{\mathrm{opt}}(\nabla\lambda - \nabla\Psi)\right) = \epsilon\Delta\rho^{\mathrm{opt}}$$

$$\rho^{\mathrm{opt}}(\boldsymbol{x},0) = \rho_0, \quad \rho^{\mathrm{opt}}(\boldsymbol{x},1) = \rho_1$$

Boundary Coupled Linear PDE's

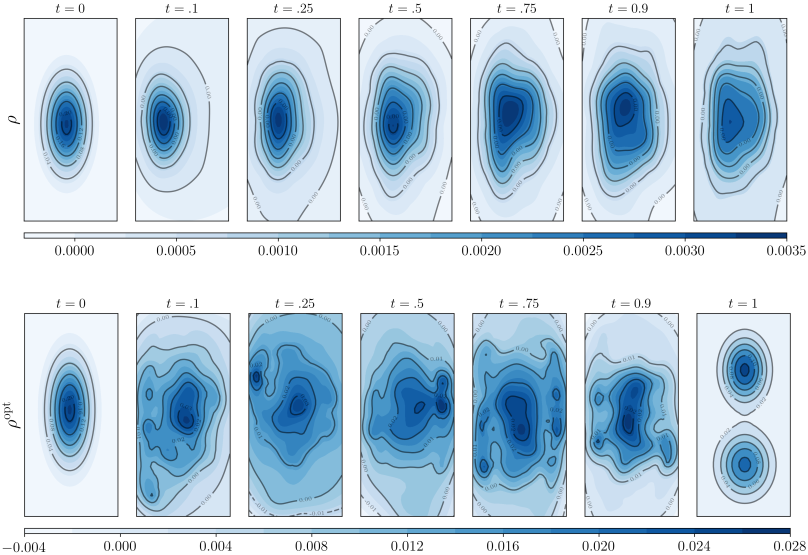
$$\frac{\partial\varphi}{\partial t} = \nabla\varphi\cdot\nabla\Psi - \epsilon\Delta\varphi$$

$$\frac{\partial\hat{\varphi}}{\partial t} = \nabla\cdot(\hat{\varphi}\nabla\Psi) + \epsilon\Delta\hat{\varphi}$$

$$\rho_0 = \varphi_0\hat{\varphi}_0, \quad \rho_1 = \varphi_1\hat{\varphi}_1$$

$$u^{\mathrm{opt}}(\boldsymbol{x},t) = 2\epsilon\nabla\varphi(\boldsymbol{x},t), \rho^{\mathrm{opt}} = \varphi(\boldsymbol{x},t)\hat{\varphi}(\boldsymbol{x},t)$$

Finite Time Density Steering: Optimal Controlled Densities



Summary

- New algorithm for density propagation
- No spatial discretization or function approximations
- Extremely fast runtime
- **Details:**
  "Gradient Flow Algorithms for Density Propagation in Stochastic Systems", *to appear in IEEE TAC*
- **Code:**
  `github.com/kcaluya/UncertaintyPropagation`

Thank You!