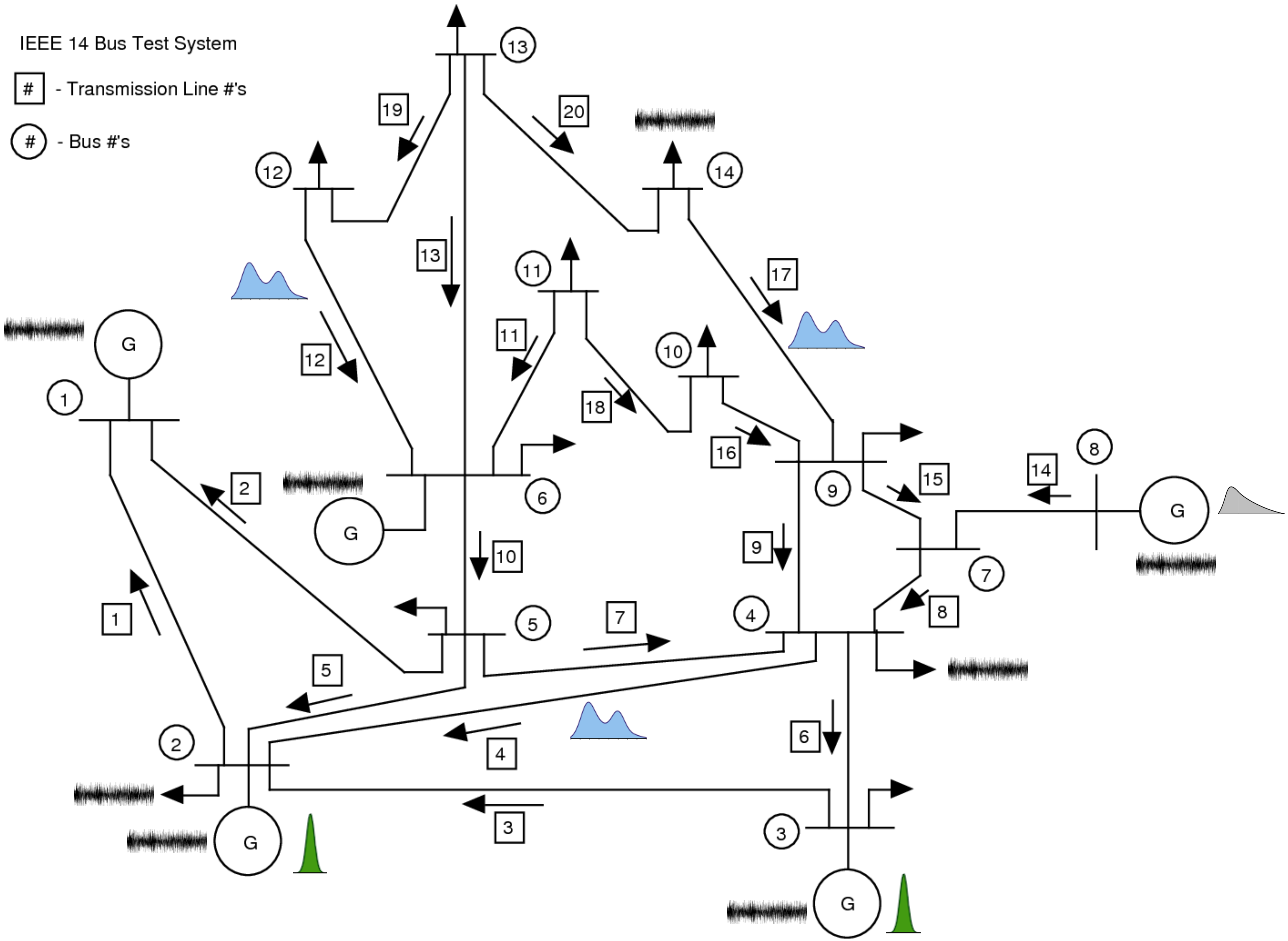
Optimal Transport Algorithms for Stochastic Uncertainty Propagation in Power Systems

Abhishek Halder

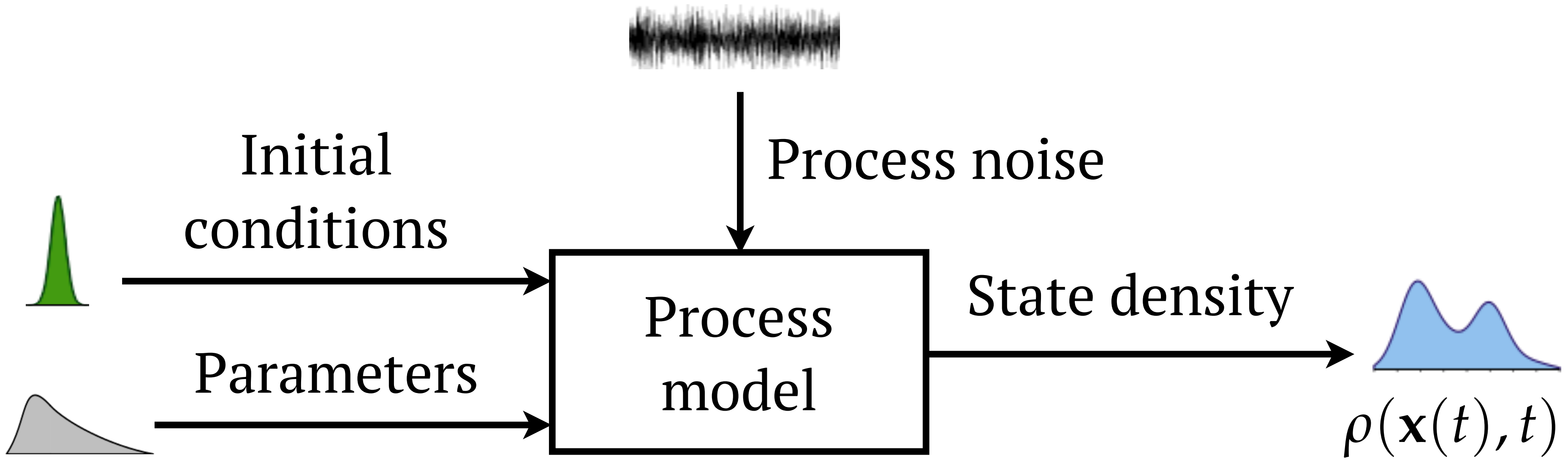
Department of Applied Mathematics
University of California, Santa Cruz
Santa Cruz, CA 95064

Acknowledgement: NSF Award 1923278

Uncertainty Propagation in Power Systems



Propagating Joint Probability Density Function



Trajectory flow:

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

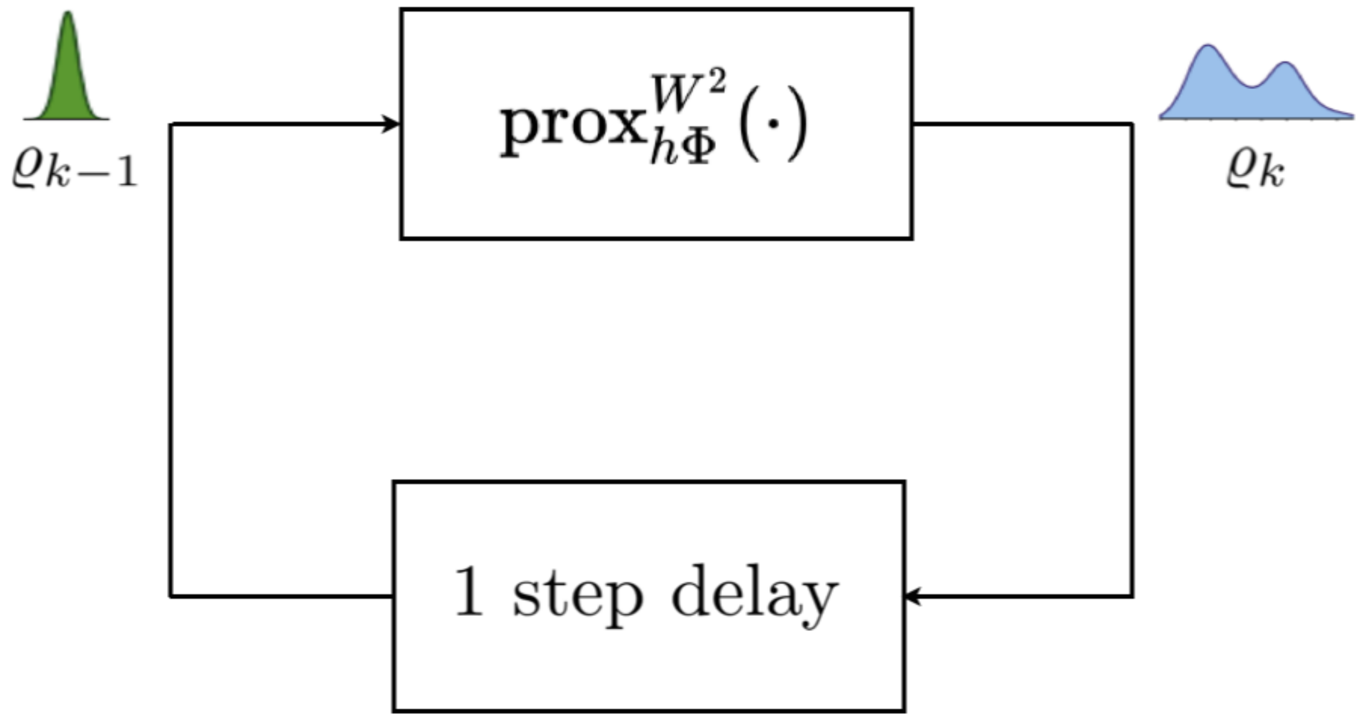
Density flow:

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

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

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

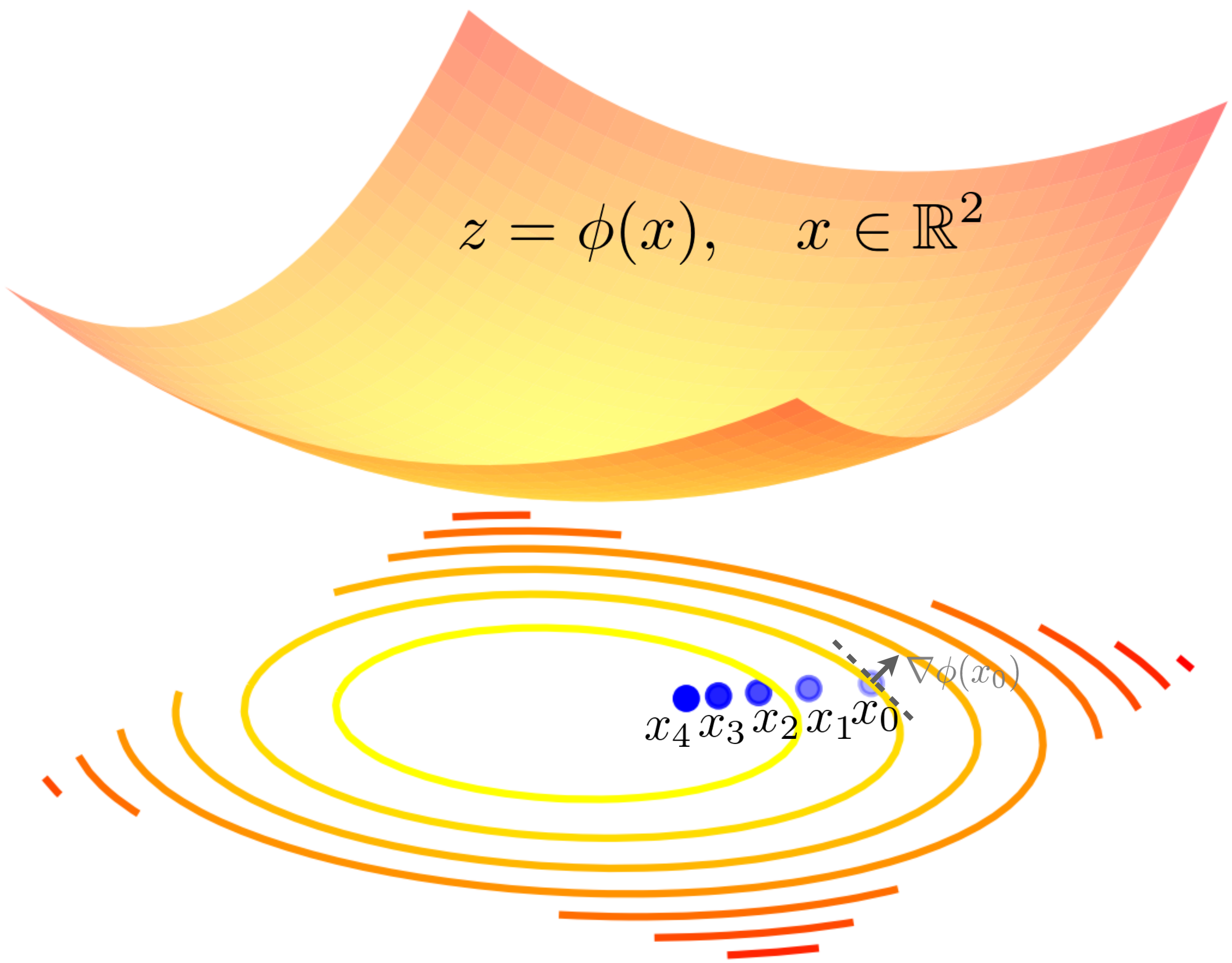
Free energy functional: $\Phi(\varrho) := \displaystyle\int_{\mathcal{X}} \psi \varrho \,\mathrm{d}x + \beta^{-1} \int_{\mathcal{X}} \varrho \log \varrho \,\mathrm{d}x$

Geometric Meaning of Gradient Flow

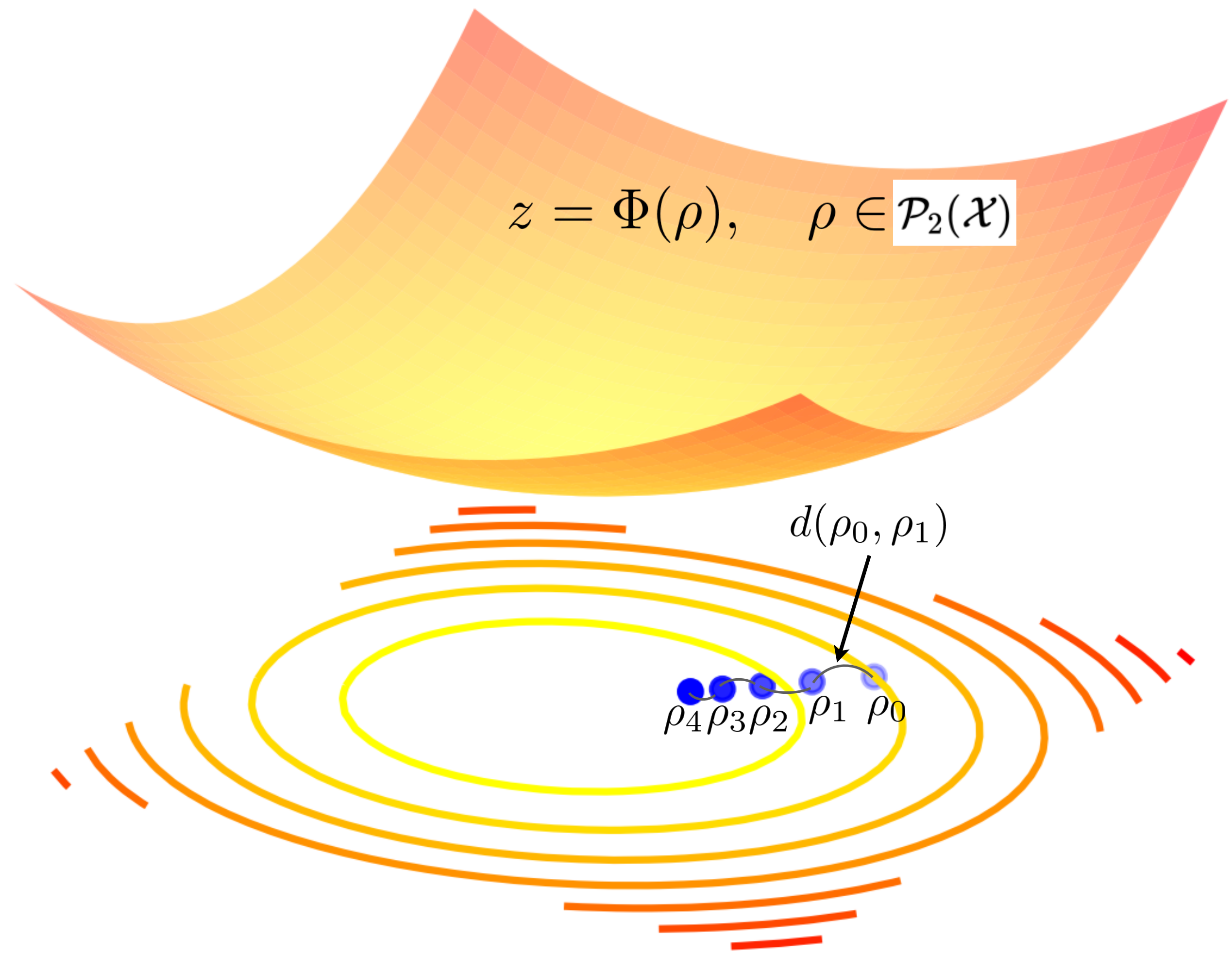
Gradient Flow in $\mathcal{X}$	**Gradient Flow in $\mathcal{P}_2(\mathcal{X})$**
$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$	$\frac{\partial\rho}{\partial t} = -\nabla^W\Phi(\rho), \quad \rho(\boldsymbol{x},0) = \rho_0$
Recursion: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

Geometric Meaning of Gradient Flow

Gradient Flow in $\mathcal{X}$

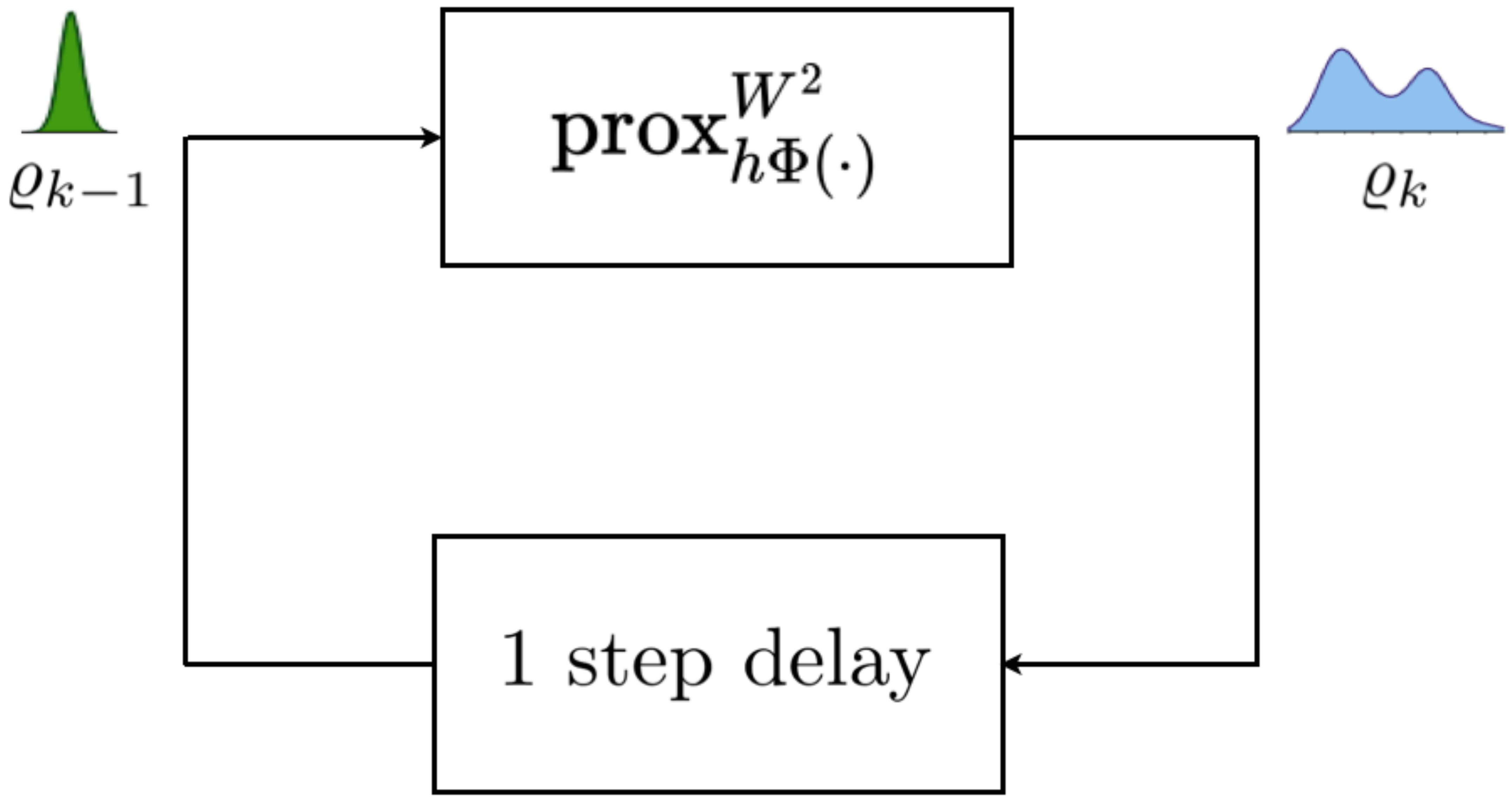


Gradient Flow in $\mathcal{P}_2(\mathcal{X})$



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**

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

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**

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

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**

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

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**

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

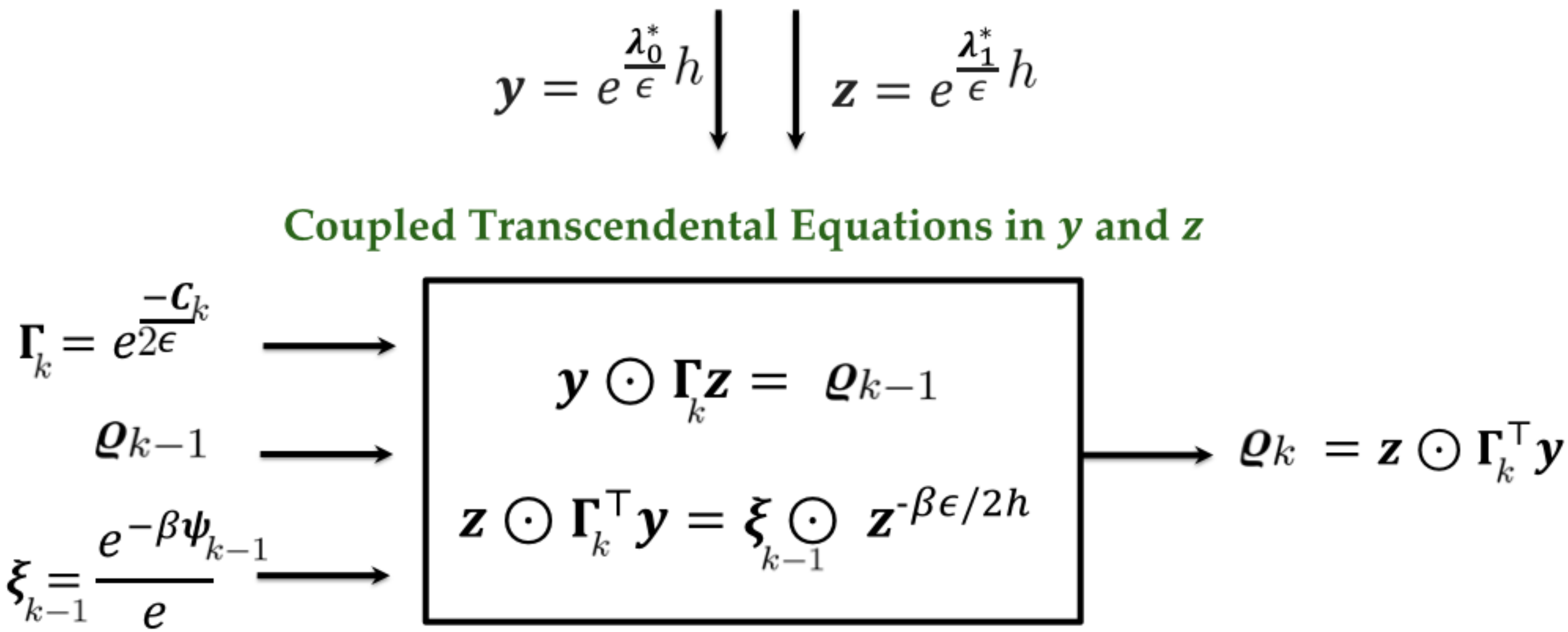
$\Downarrow$ **Entropic Regularization**

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

$\Updownarrow$ **Dualization**

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

Recursion on the Cone

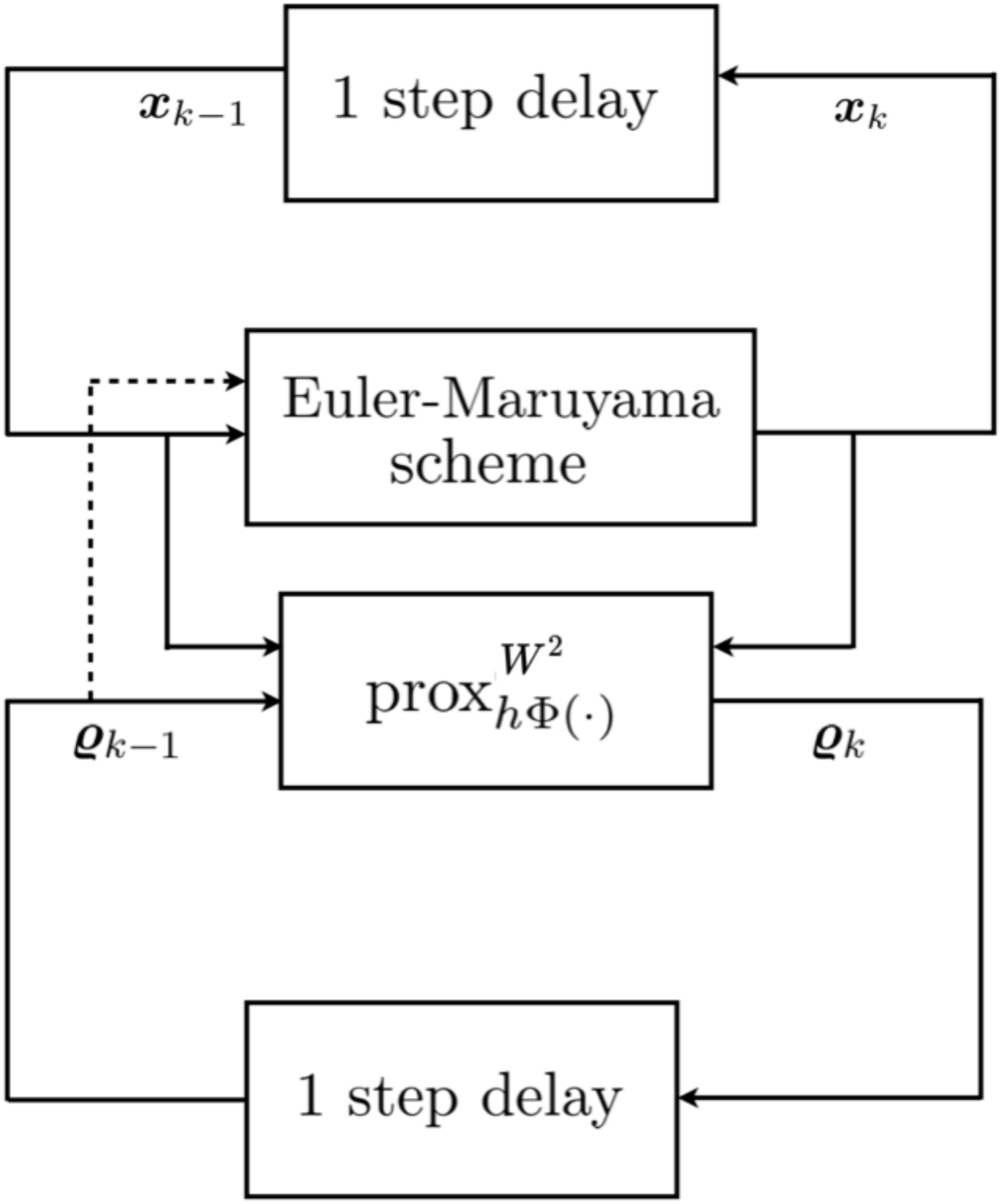


Theorem: Consider the recursion on the cone $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$

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

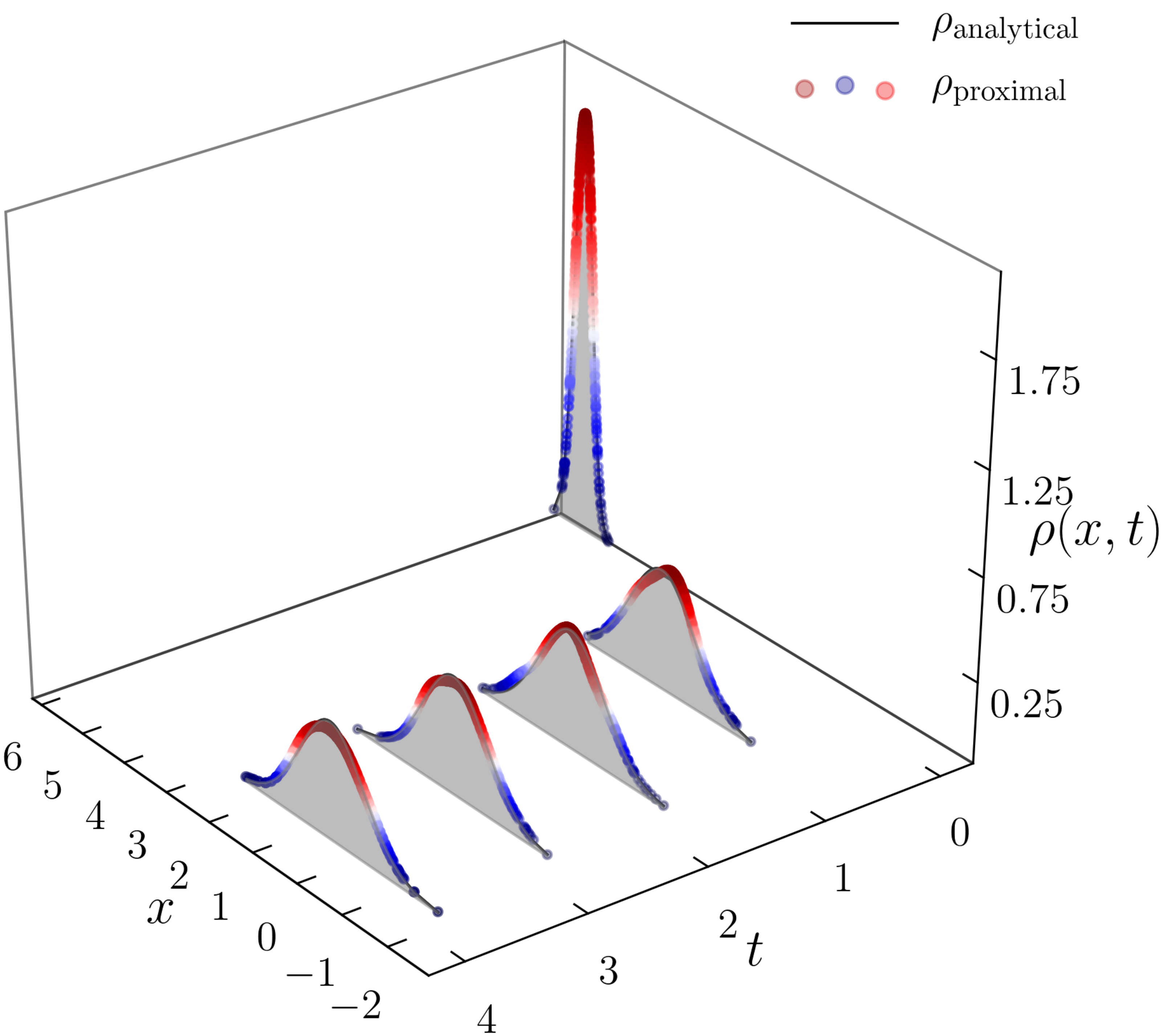
Then the solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ gives the proximal update $\varrho_k = \boldsymbol{z}^* \odot (\boldsymbol{\Gamma}_k{}^\top \boldsymbol{y}^*)$

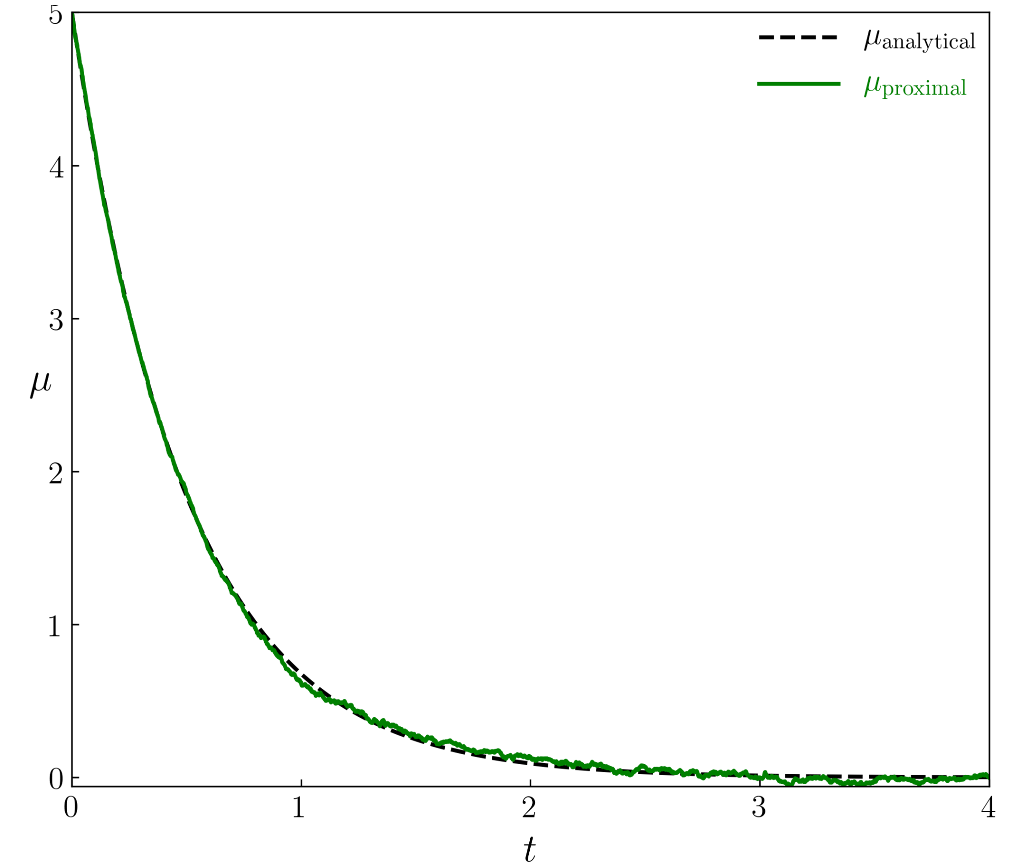
Algorithmic Setup

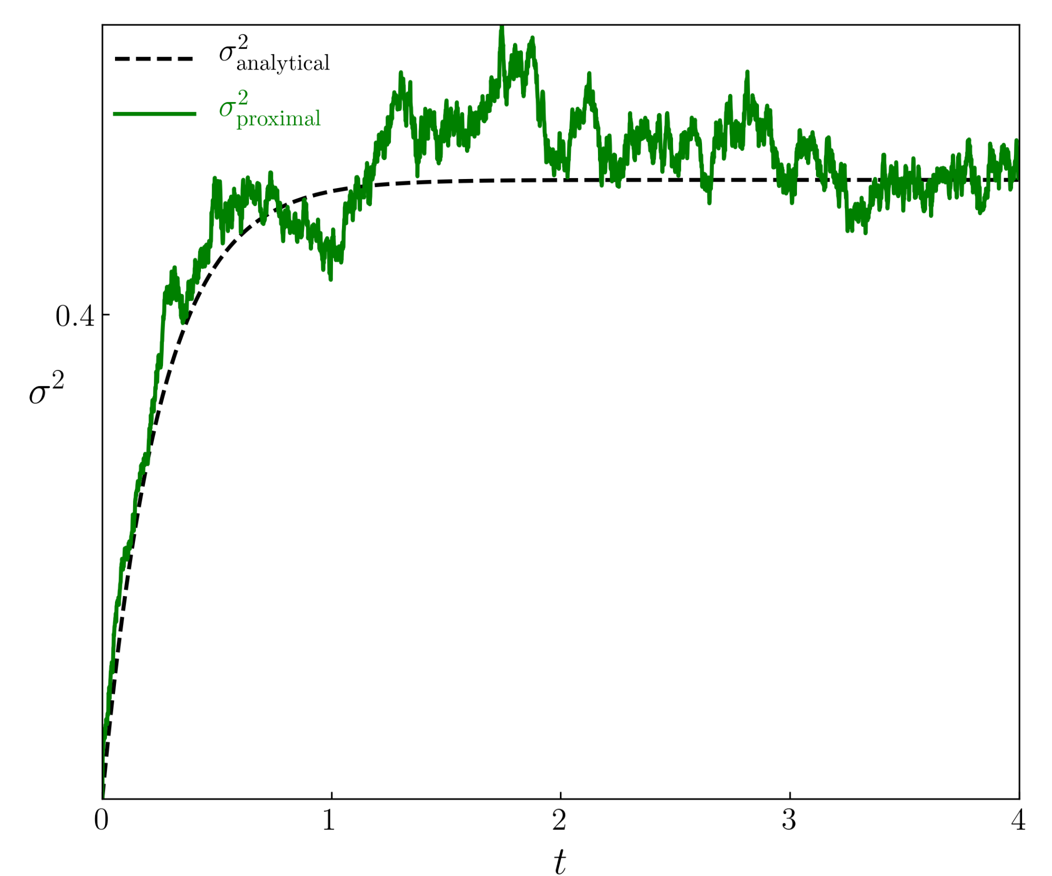


Theorem: Block co-ordinate iteration of $(\mathbf{y}, \mathbf{z})$ recursion is contractive on $\mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0}$.

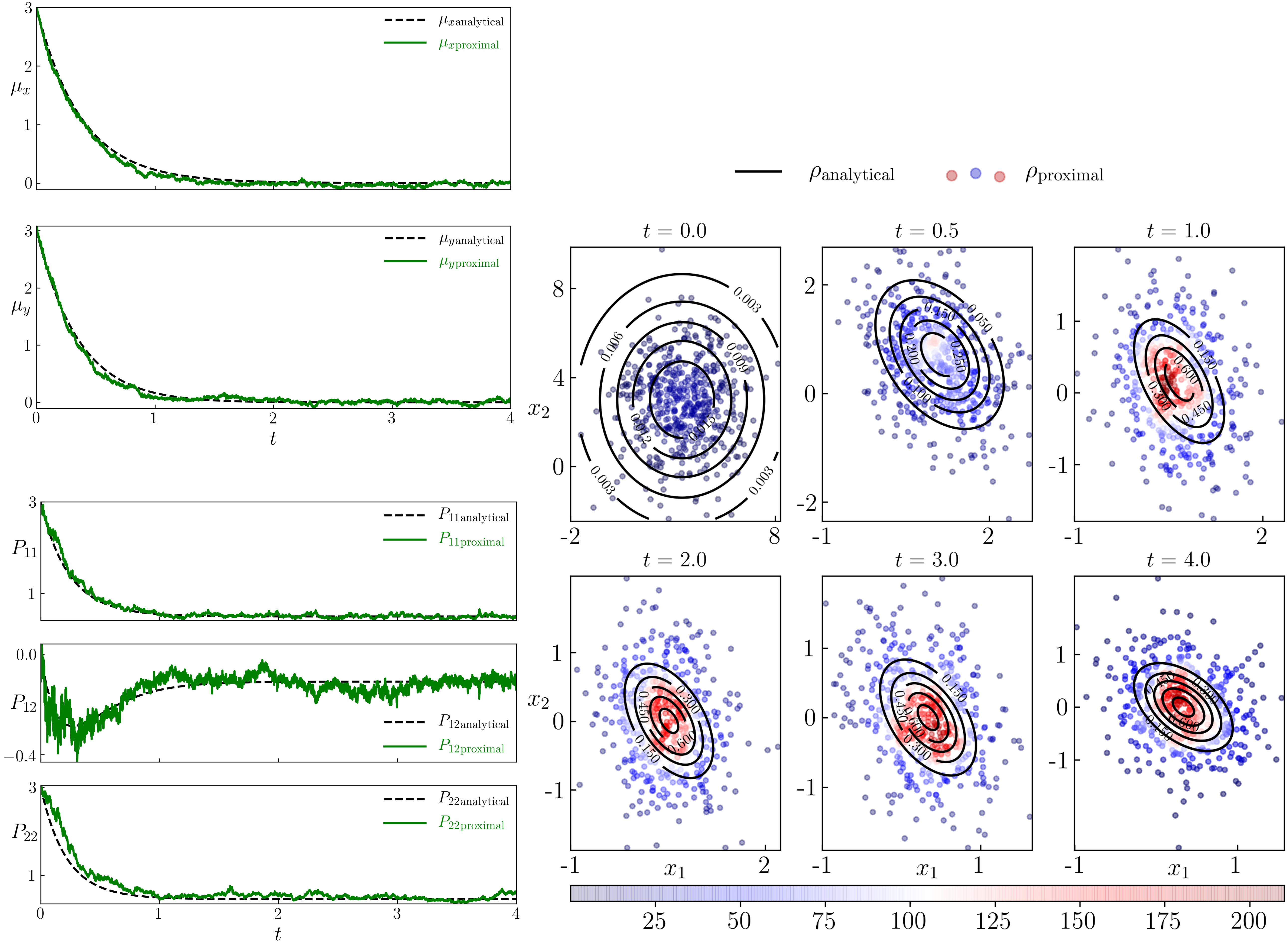
Proximal Prediction: 1D Linear Gaussian



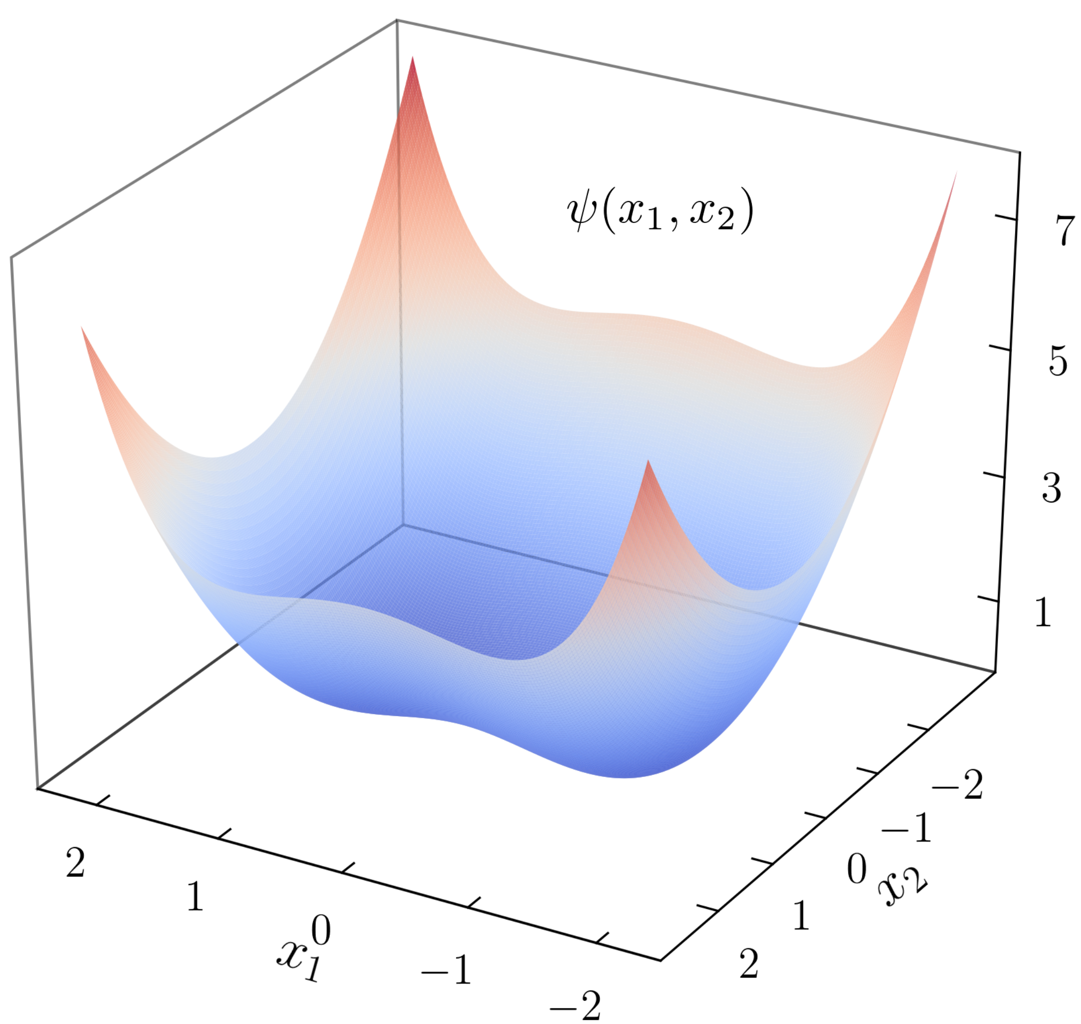


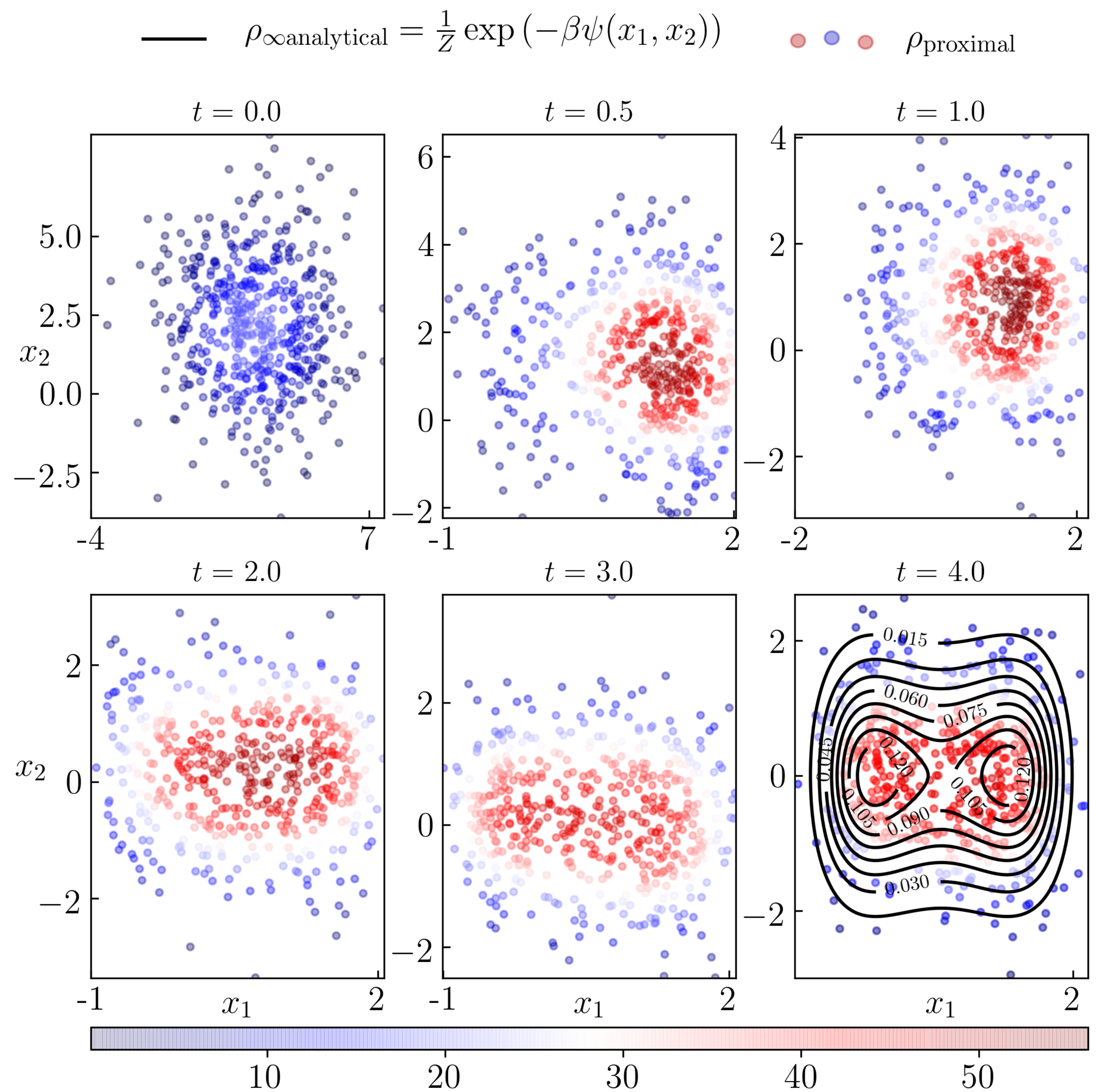


Proximal Prediction: 2D Linear Gaussian

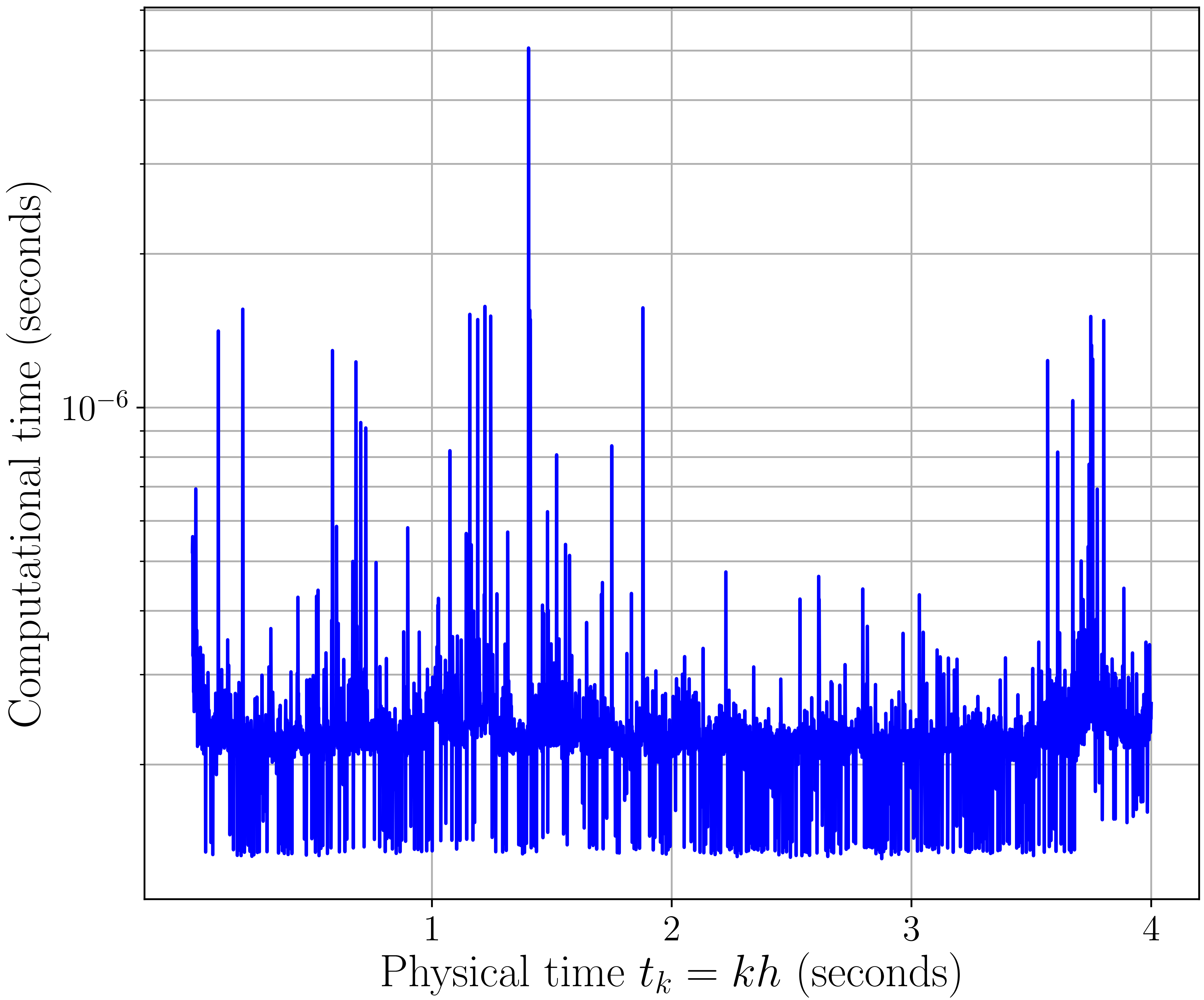


Proximal Prediction: Nonlinear Non-Gaussian





Computational Time: Nonlinear Non-Gaussian



Network Reduced Power System Model

Noisy Kuramoto (a.k.a. structure preserving power network) model

$$m_i\ddot{\theta}_i+\gamma_i\dot{\theta}_i= P_i^{\mathrm{mech}} - \sum_{j=1}^{n} k_{ij}\sin(\theta_i-\theta_j) + \sigma_i \times \text{stochastic forcing},\; i=1,\ldots,n$$

Mixed Conservative-Dissipative SDE over state variables $(\boldsymbol{\theta},\boldsymbol{\omega}) \in \mathbb{T}^n \times \mathbb{R}^n$

$$\mathrm{d}\boldsymbol{\theta} = \boldsymbol{\omega}\,\mathrm{d}t$$
$$\mathrm{d}\boldsymbol{\omega} = \left(-(\boldsymbol{\gamma}\oslash\boldsymbol{m})\odot\boldsymbol{\omega} - \nabla_{\boldsymbol{\theta}}V(\boldsymbol{\theta})\right)\mathrm{d}t \;+\; (\boldsymbol{\sigma}\oslash\boldsymbol{m})\odot\mathrm{d}\boldsymbol{w}$$

Potential function $V : \mathbb{T}^n \mapsto \mathbb{R}_{\geq 0}$

$$V(\boldsymbol{\theta}) := \sum_{i=1}^{n}\frac{1}{m_i}P_i^{\mathrm{mech}}\theta_i + \sum_{(i,j)\in\mathcal{E}}\frac{1}{m_i}k_{ij}(1-\cos(\theta_i-\theta_j))$$

Proximal Recursion for Power System Model

Consider simple case: homogeneous generators with $\sigma^2 = 2\beta^{-1}\gamma$

Lyapunov functional:

$$\Phi(\rho) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(\frac{1}{2} \parallel \boldsymbol{\omega} \parallel_2^2 + V(\boldsymbol{\theta}) \right) \rho \, \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{\omega} + \beta^{-1} \int_{\mathbb{T}^n \times \mathbb{R}^n} \rho \log \rho \, \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{\omega}$$

However, the FPK PDE is NOT a gradient descent of Φ **w.r.t.** W

Instead, do: $\varrho_k = \mathrm{prox}_{h\gamma\widetilde{\Phi}}^{\widetilde{W}}(\varrho_{k-1}), \quad k \in \mathbb{N},$

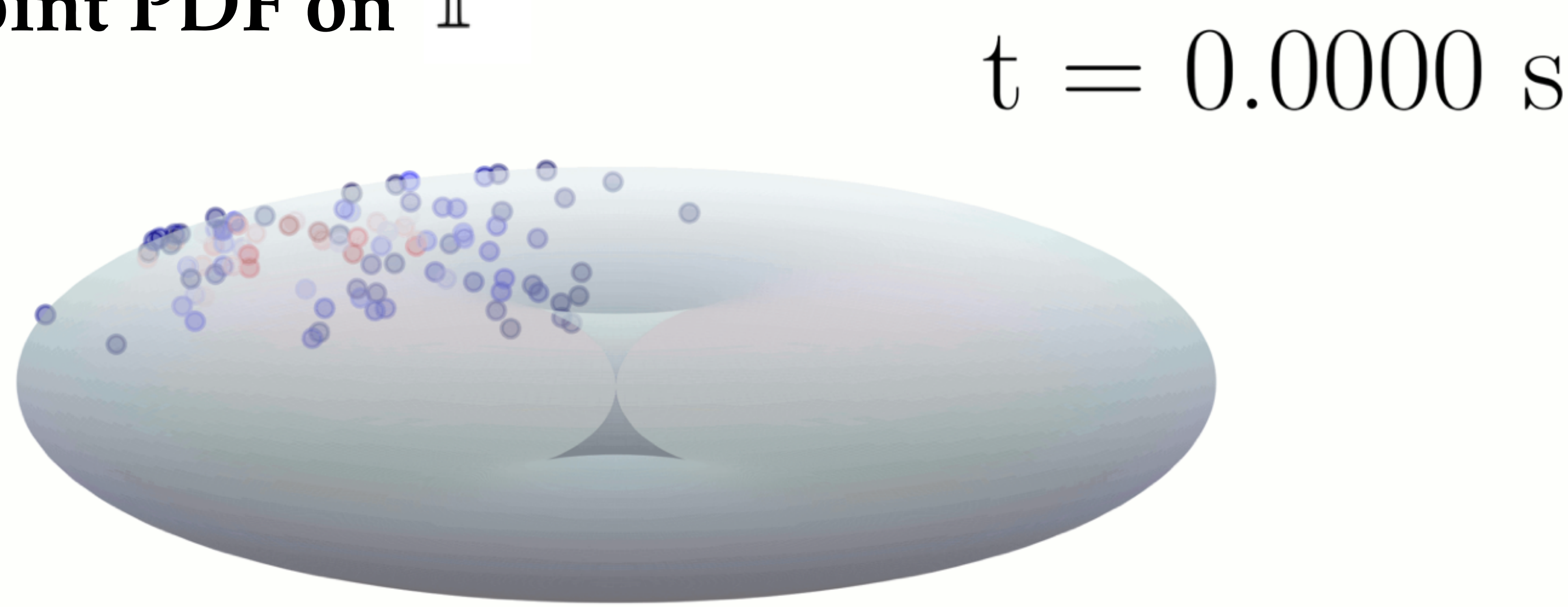
$$\widetilde{\Phi}(\rho) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \frac{1}{2} \parallel \boldsymbol{\omega} \parallel_2^2 \rho \, \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{\omega} + \beta^{-1} \int_{\mathbb{T}^n \times \mathbb{R}^n} \rho \log \rho \, \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{\omega}$$

$$\widetilde{W}^2(\varrho, \varrho_{k-1}) = \inf_{\pi \in \Pi(\varrho, \varrho_{k-1})} \int_{\mathbb{T}^{2n} \times \mathbb{R}^{2n}} s_h(\boldsymbol{\theta}, \boldsymbol{\omega}, \overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\omega}}) \, \mathrm{d}\pi(\boldsymbol{\theta}, \boldsymbol{\omega}, \overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\omega}})$$

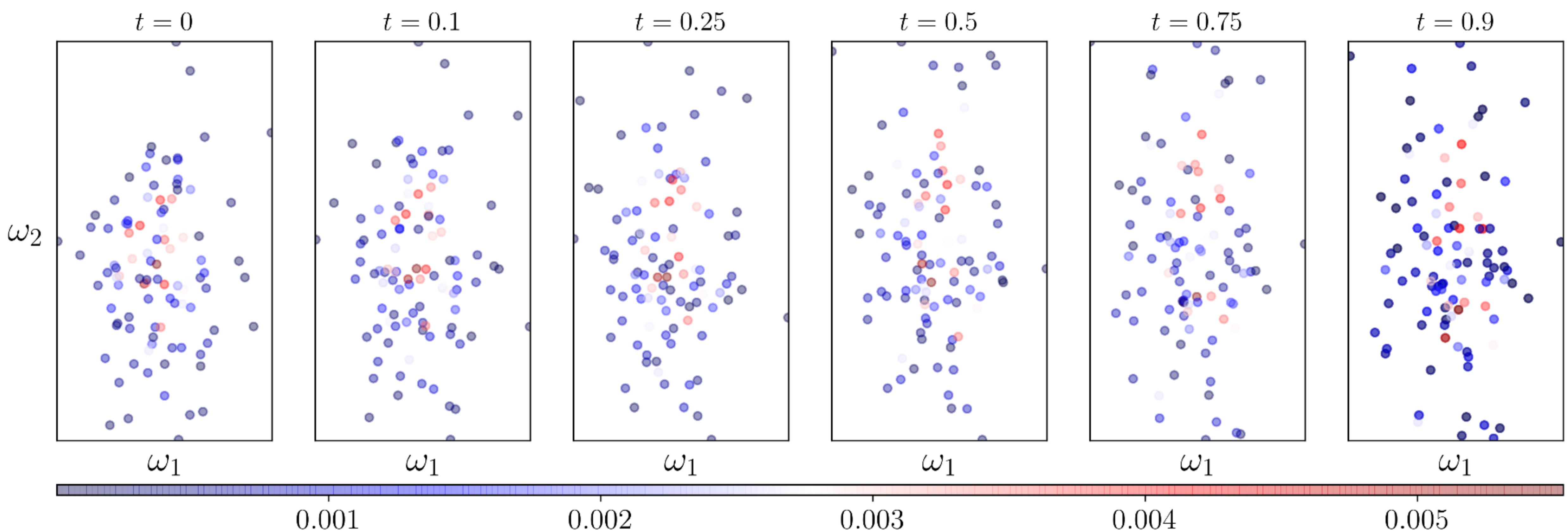
where $s_h(\boldsymbol{\theta}, \boldsymbol{\omega}, \overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\omega}}) := \left\| \overline{\boldsymbol{\omega}} - \boldsymbol{\omega} + h \nabla V(\boldsymbol{\theta}) \right\|_2^2 + 12 \left\| \frac{\overline{\boldsymbol{\theta}} - \boldsymbol{\theta}}{h} - \frac{\overline{\boldsymbol{\omega}} + \boldsymbol{\omega}}{2} \right\|_2^2$

Proximal Prediction: Power System with $n = 2$

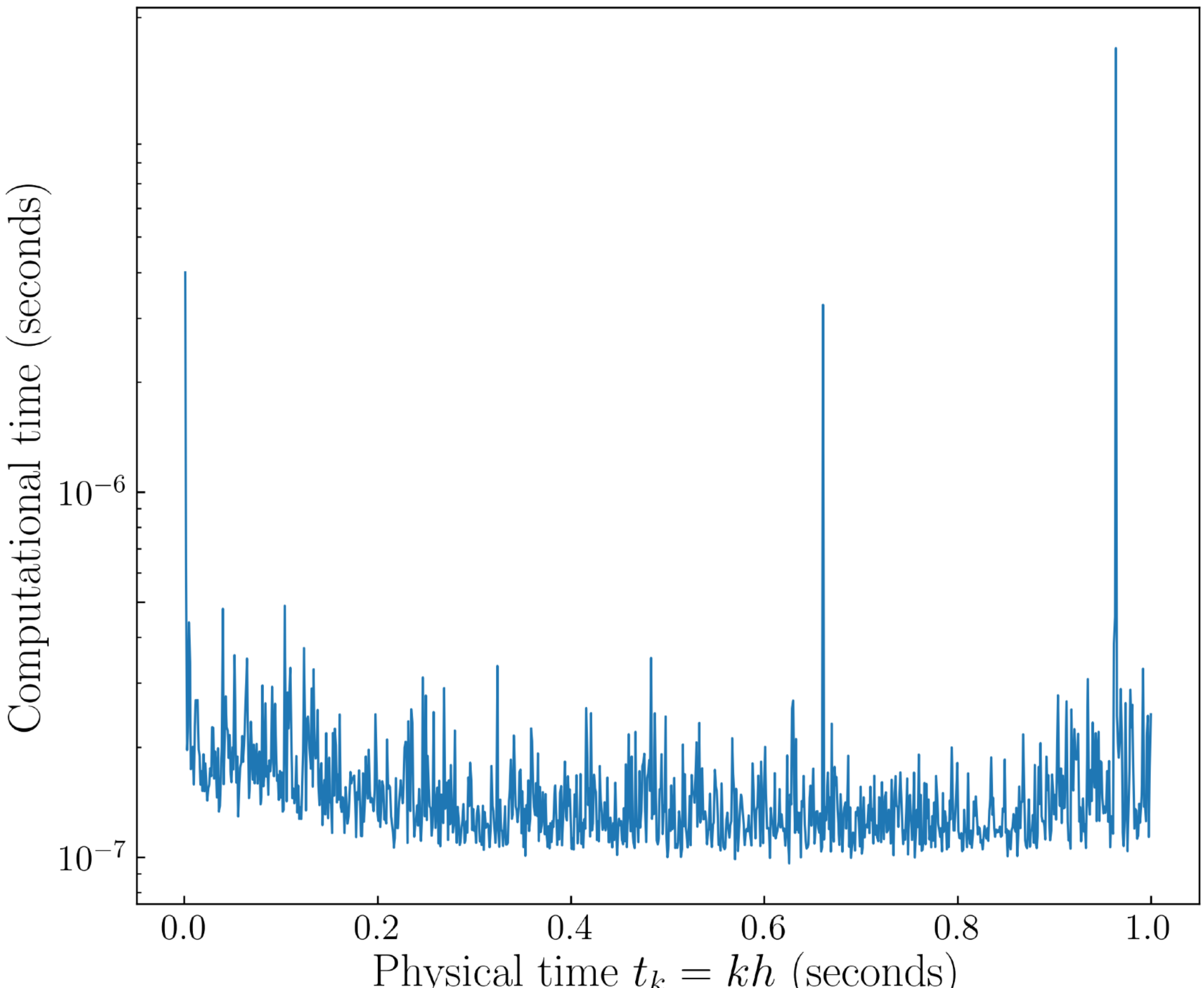
Projection of the joint PDF on $\mathbb{T}^2$



Projection of the joint PDF on $\mathbb{R}^2$



Computational Time: Power System with $n = 2$



Summary

Fast proximal recursions for PDF propagation in power systems

Ongoing

Large scale implementation: ~1000 generators in ~seconds

Address Stochastic Differential Algebraic Equations (SDAEs)

Thank You