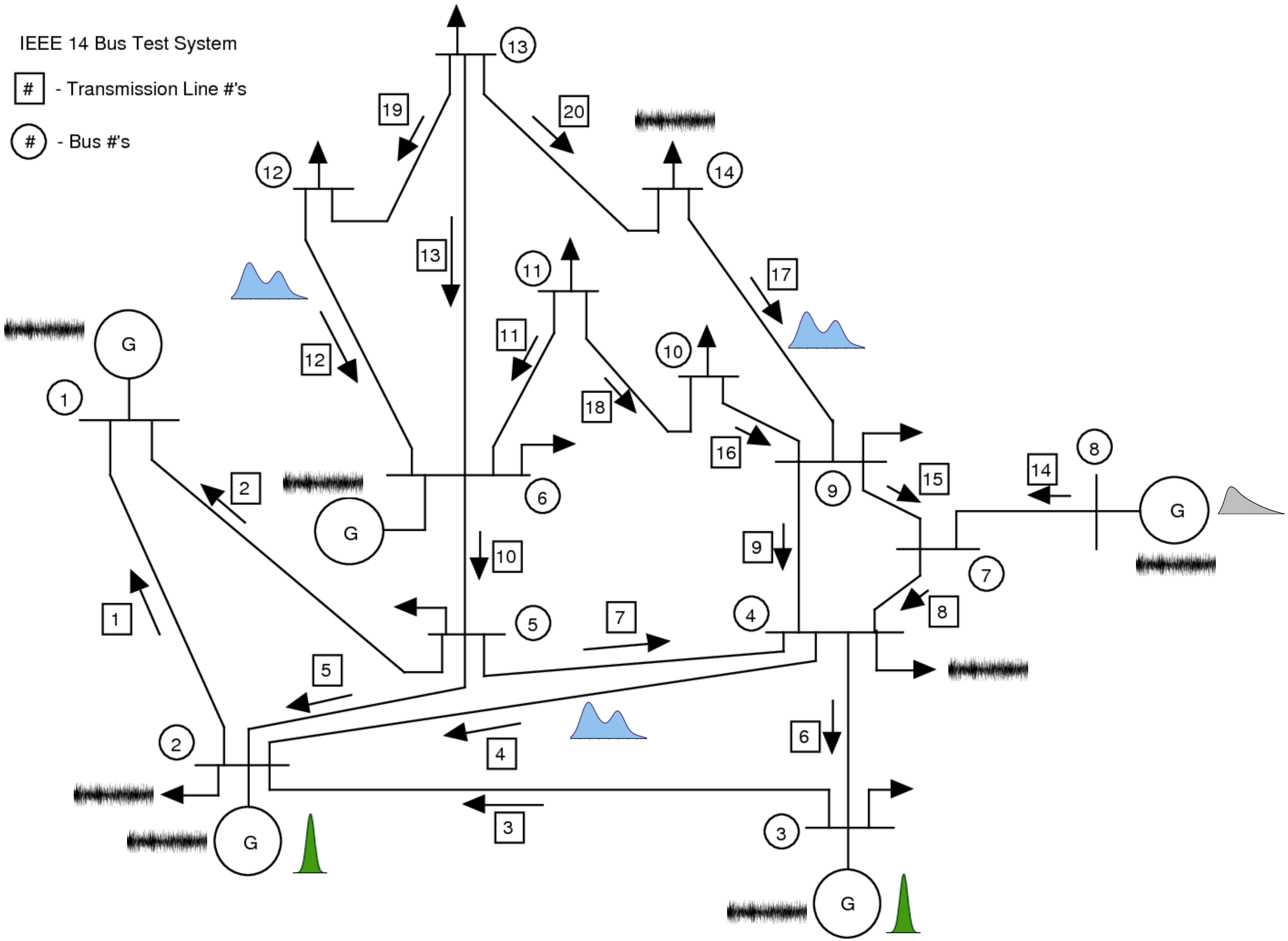
Stochastic Uncertainty Propagation in Power System Dynamics using Measure-valued Proximal Recursions

Abhishek Halder

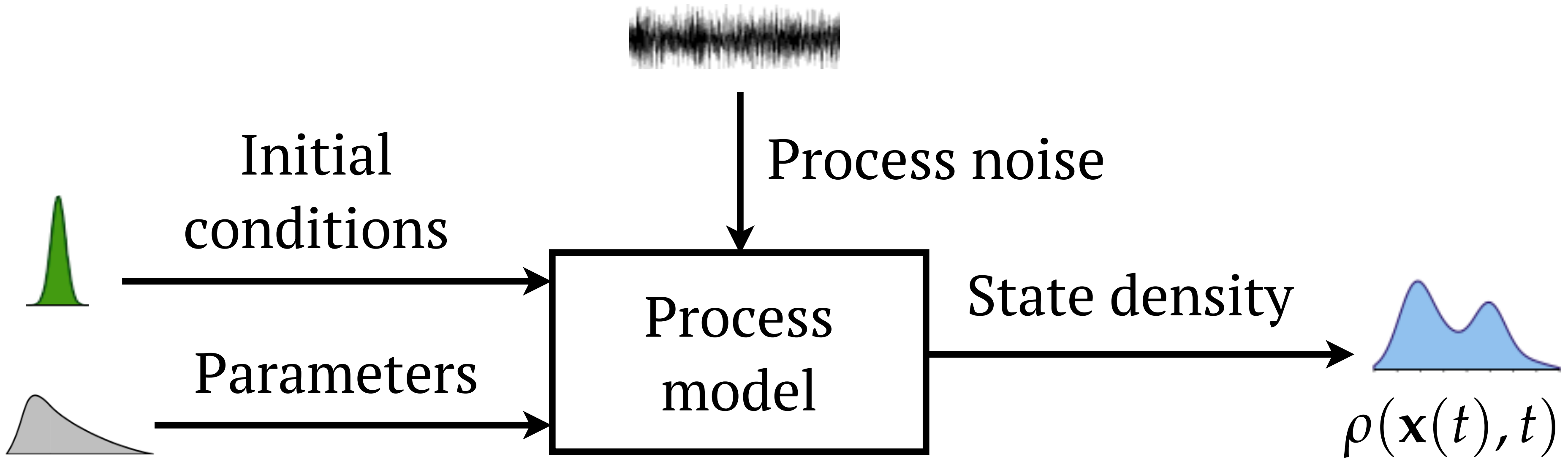
Department of Applied Mathematics
University of California, Santa Cruz
Santa Cruz, CA 95064

Joint work with K.F. Caluya, P. Ojaghi, X. Geng

Uncertainty propagation in power systems



Propagating joint PDFs



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:

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

Long history of methods for solving such PDE

But all need grid

— Monte Carlo methods: p.w. constant approximation

— classical finite element, finite volume methods

— tensor decomposition techniques

— Kernel and other density estimation algorithms

— Recast as NN regression with suitable loss, then do SGD

Can we do gridless computation?

Evolve time-varying weighted point clouds?

We want the transient joint PDF computation be

Non-parametric

— nonlinearity in dynamics destroys Gaussianity

— don't *a priori* know the dimensionality of sufficient statistics

— cannot rigorously justify statistical approximations (exponential family, mixture of Gaussians) for the transient

— don't want to make dynamical approximations (linear, Taylor series)

We want the transient joint PDF computation be

Non-parametric

— nonlinearity in dynamics destroys Gaussianity

— don't *a priori* know the dimensionality of sufficient statistics

— cannot rigorously justify statistical approximations (exponential family, mixture of Gaussians) for the transient

— don't want to make dynamical approximations (linear, Taylor series)

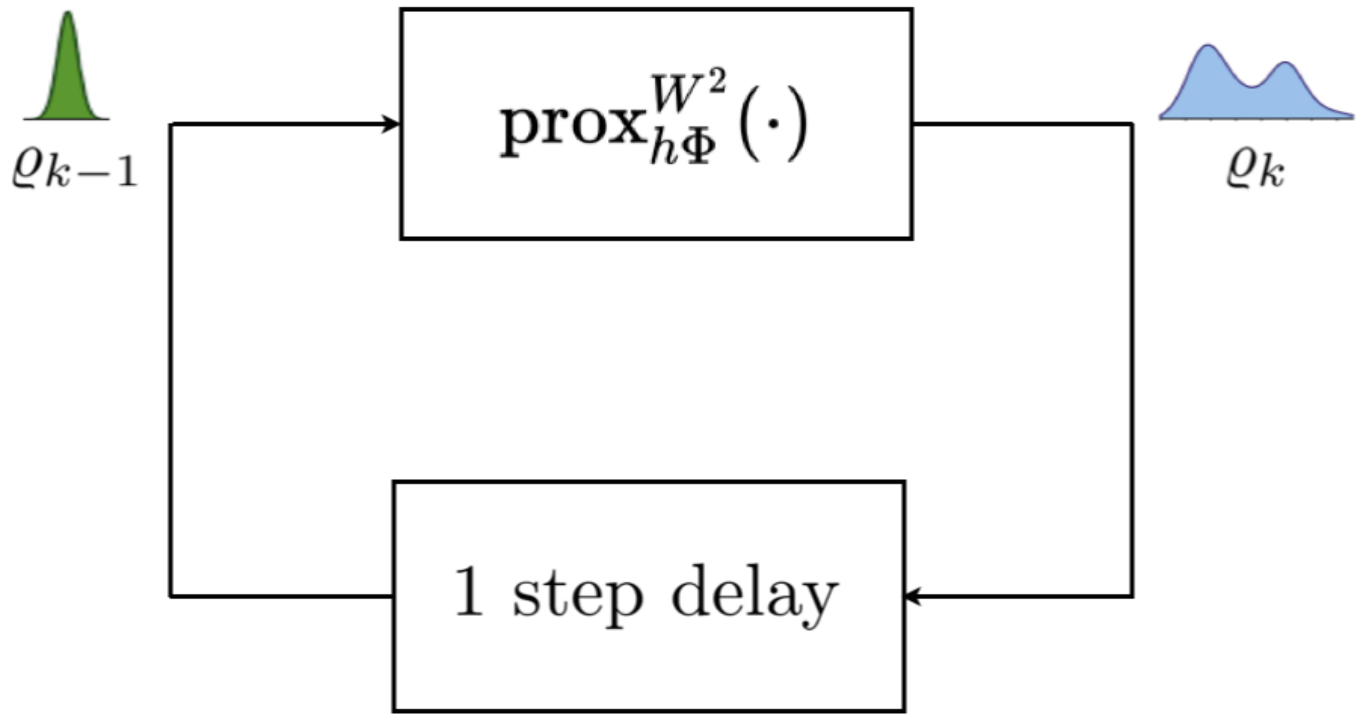
Embrace and make use of

~~Fight~~ exact nonlinearity

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

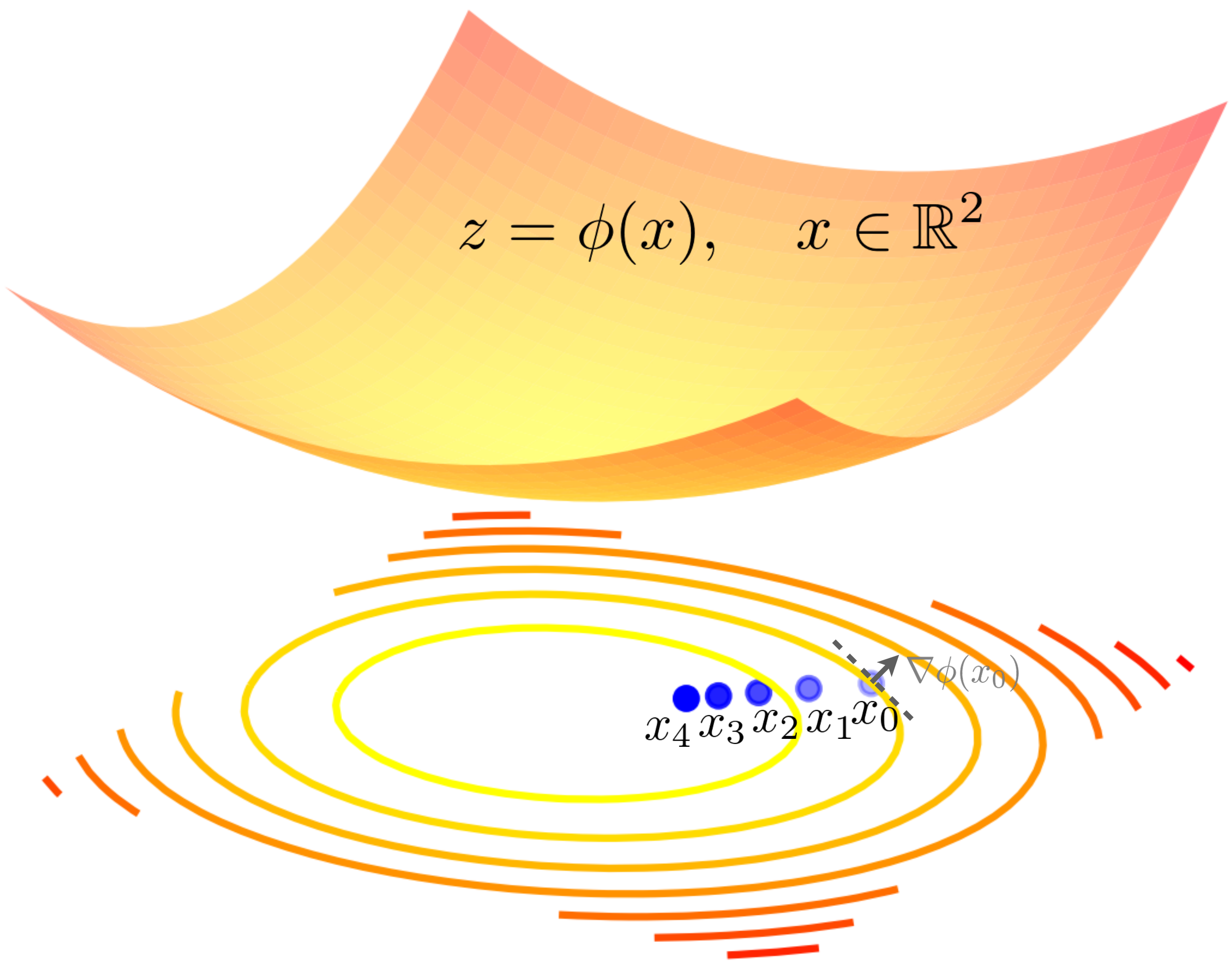
Free energy functional: $\Phi(\varrho) := \displaystyle\int_{\mathcal{X}} \psi\varrho \,\mathrm{d}x + \beta^{-1} \int_{\mathcal{X}} \varrho \log \varrho \,\mathrm{d}x$

Generalized gradient flow

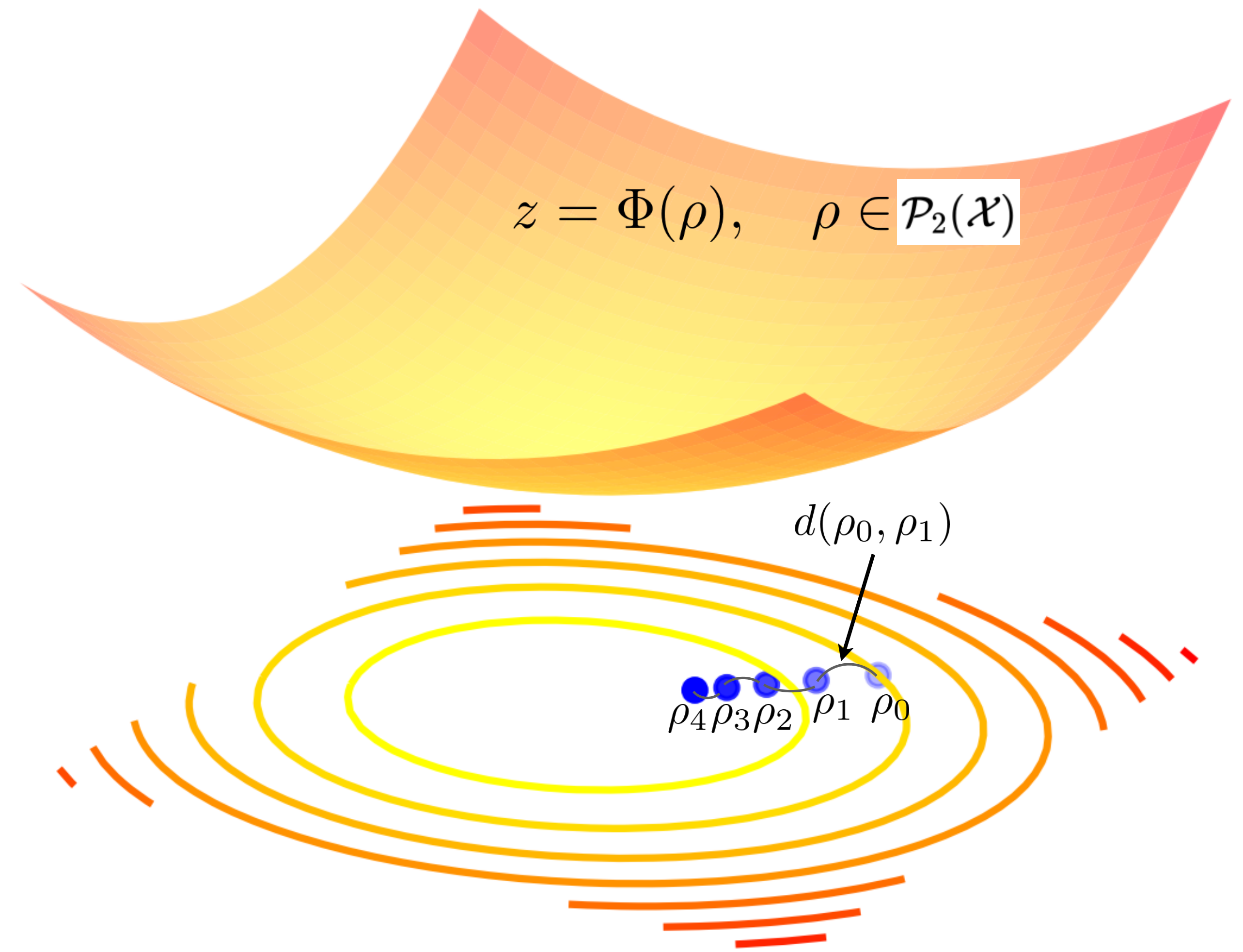
Gradient Flow in $\mathcal{X}$	**Gradient Flow in $\mathcal{P}_2(\mathcal{X})$**
$\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -\nabla\varphi(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$	$\dfrac{\partial\rho}{\partial t} = -\nabla^W \Phi(\rho), \quad \rho(\boldsymbol{x},0) = \rho_0$
Recursion:	**Recursion:**
$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} - h\nabla\varphi(\boldsymbol{x}_k)$ $= \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\min}\left\{\frac{1}{2}\Vert\boldsymbol{x}-\boldsymbol{x}_{k-1}\Vert_2^2 + h\varphi(\boldsymbol{x})\right\}$ $=: \mathrm{prox}_{h\varphi}^{\Vert\cdot\Vert_2}(\boldsymbol{x}_{k-1})$	$\rho_k = \rho(\cdot, t = kh)$ $= \underset{\rho\in\mathcal{P}_2(\mathcal{X})}{\arg\min}\left\{\frac{1}{2}W^2(\rho,\rho_{k-1}) + h\Phi(\rho)\right\}$ $=: \mathrm{prox}_{h\Phi}^{W^2}(\rho_{k-1})$
Convergence:	**Convergence:**
$\boldsymbol{x}_k \to \boldsymbol{x}(t = kh) \quad \text{as} \quad h \downarrow 0$	$\rho_k \to \rho(\cdot, t = kh) \quad \text{as} \quad h \downarrow 0$
φ as Lyapunov function:	**Φ as Lyapunov functional:**
$\dfrac{\mathrm{d}}{\mathrm{d}t}\varphi = -\Vert\nabla\varphi\Vert_2^2 \leq 0$	$\dfrac{\mathrm{d}}{\mathrm{d}t}\Phi = -\mathbb{E}_\rho\left[\left\Vert\nabla\dfrac{\delta\Phi}{\delta\rho}\right\Vert_2^2\right] \leq 0$

Generalized gradient flow

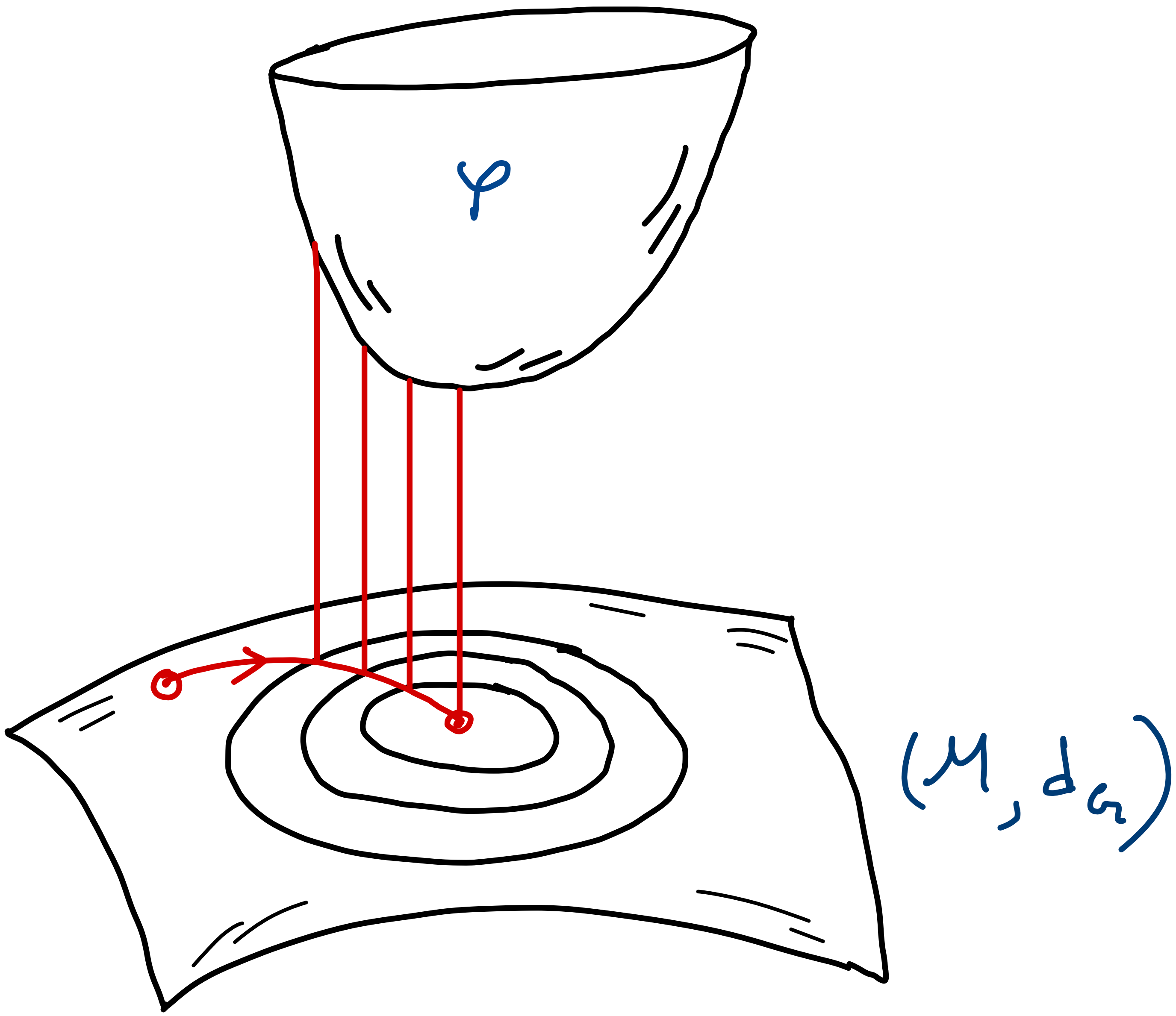
Gradient Flow in $\mathcal{X}$

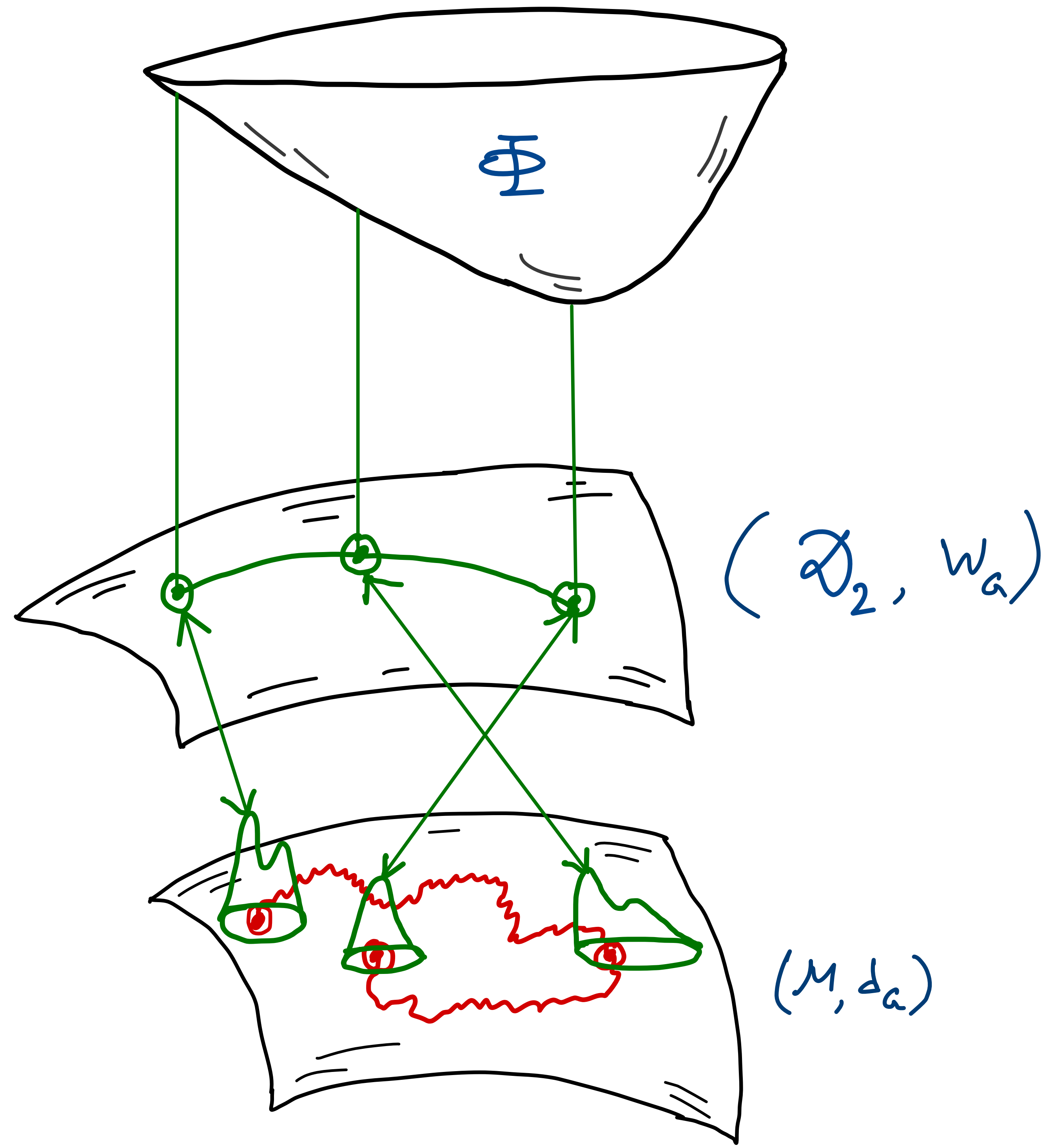


Gradient Flow in $\mathcal{P}_2(\mathcal{X})$



Generalized gradient flow with base manifold

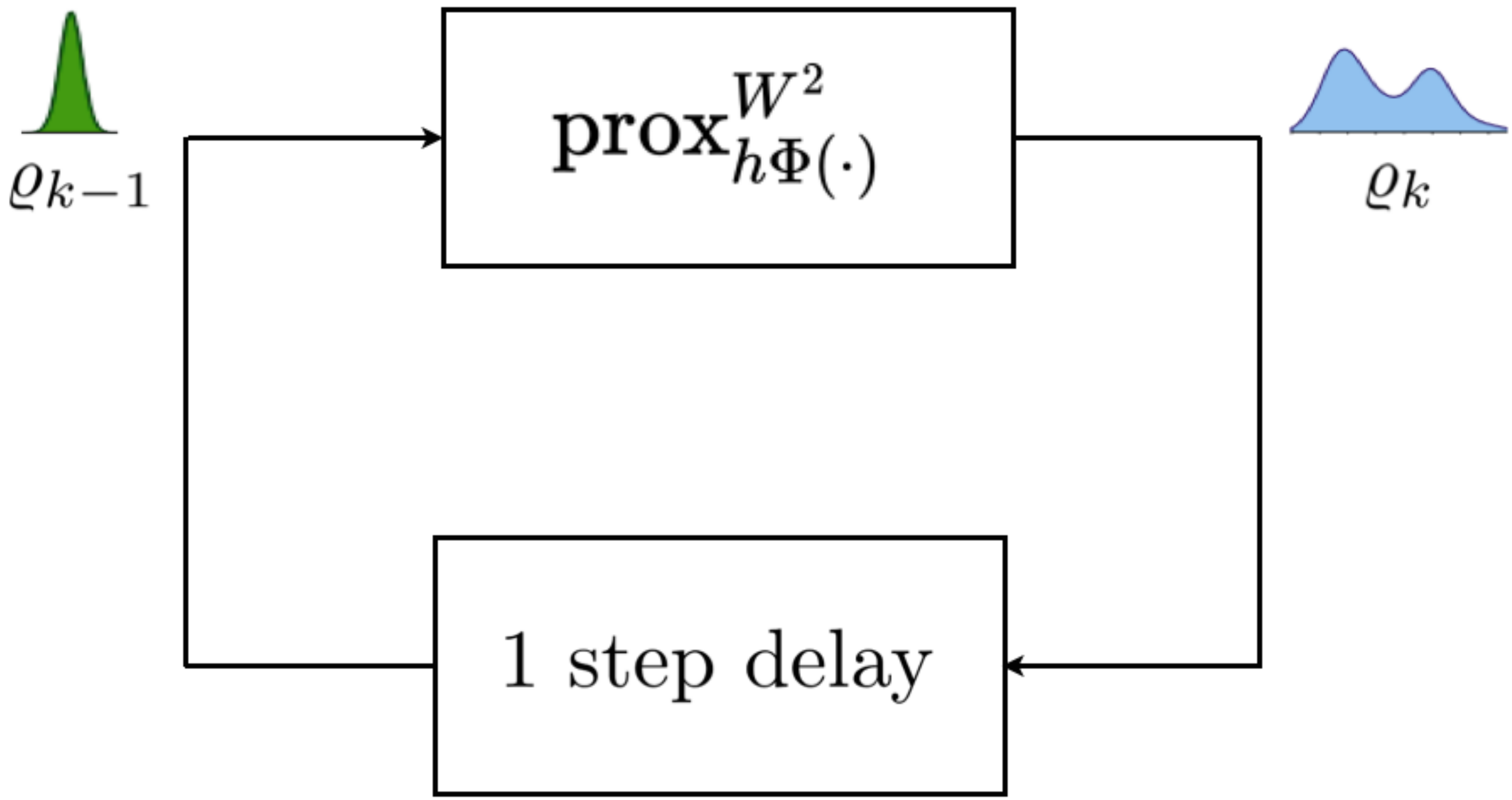




Design pair (W_G, Φ)

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**

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

Algorithm: gradient ascent on the dual space

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

$$\Updownarrow \quad \textbf{Proximal Recursion}$$

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

$$\Downarrow \quad \textbf{Discrete Primal Formulation}$$

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

$$\Downarrow \quad \textbf{Entropic Regularization}$$

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

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**

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

$\Downarrow$ **Entropic Regularization**

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

$\Updownarrow$ **Dualization**

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

Recursion on the cone

$$\boldsymbol{y} = e^{\frac{\boldsymbol{\lambda}_0^*}{\epsilon}h} \Big\downarrow \qquad \Big\downarrow \boldsymbol{z} = e^{\frac{\boldsymbol{\lambda}_1^*}{\epsilon}h}$$

Coupled Transcendental Equations in *y* and *z*

$$\boldsymbol{\Gamma}_k = e^{\frac{-\boldsymbol{C}_k}{2\epsilon}} \longrightarrow \qquad \varrho_{k-1} \longrightarrow \qquad \boldsymbol{\xi}_{k-1} = \frac{e^{-\beta\boldsymbol{\psi}_{k-1}}}{e} \longrightarrow$$

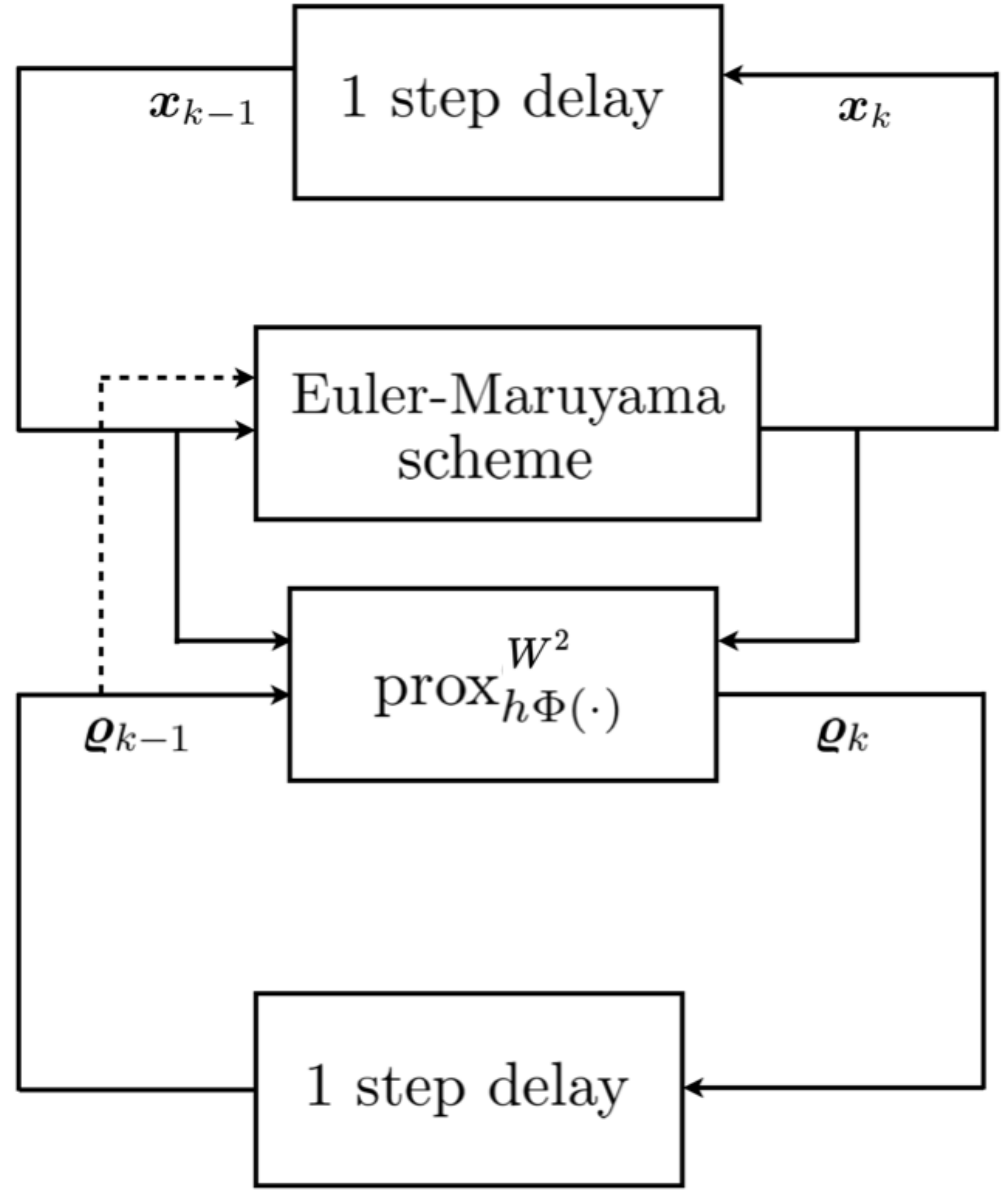
$$\boxed{\begin{aligned} \boldsymbol{y} \odot \boldsymbol{\Gamma}_k \boldsymbol{z} &= \varrho_{k-1} \\ \boldsymbol{z} \odot \boldsymbol{\Gamma}_k^\top \boldsymbol{y} &= \boldsymbol{\xi}_{k-1} \odot \boldsymbol{z}^{-\beta\epsilon/2h} \end{aligned}} \longrightarrow \varrho_k = \boldsymbol{z} \odot \boldsymbol{\Gamma}_k^\top \boldsymbol{y}$$

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

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

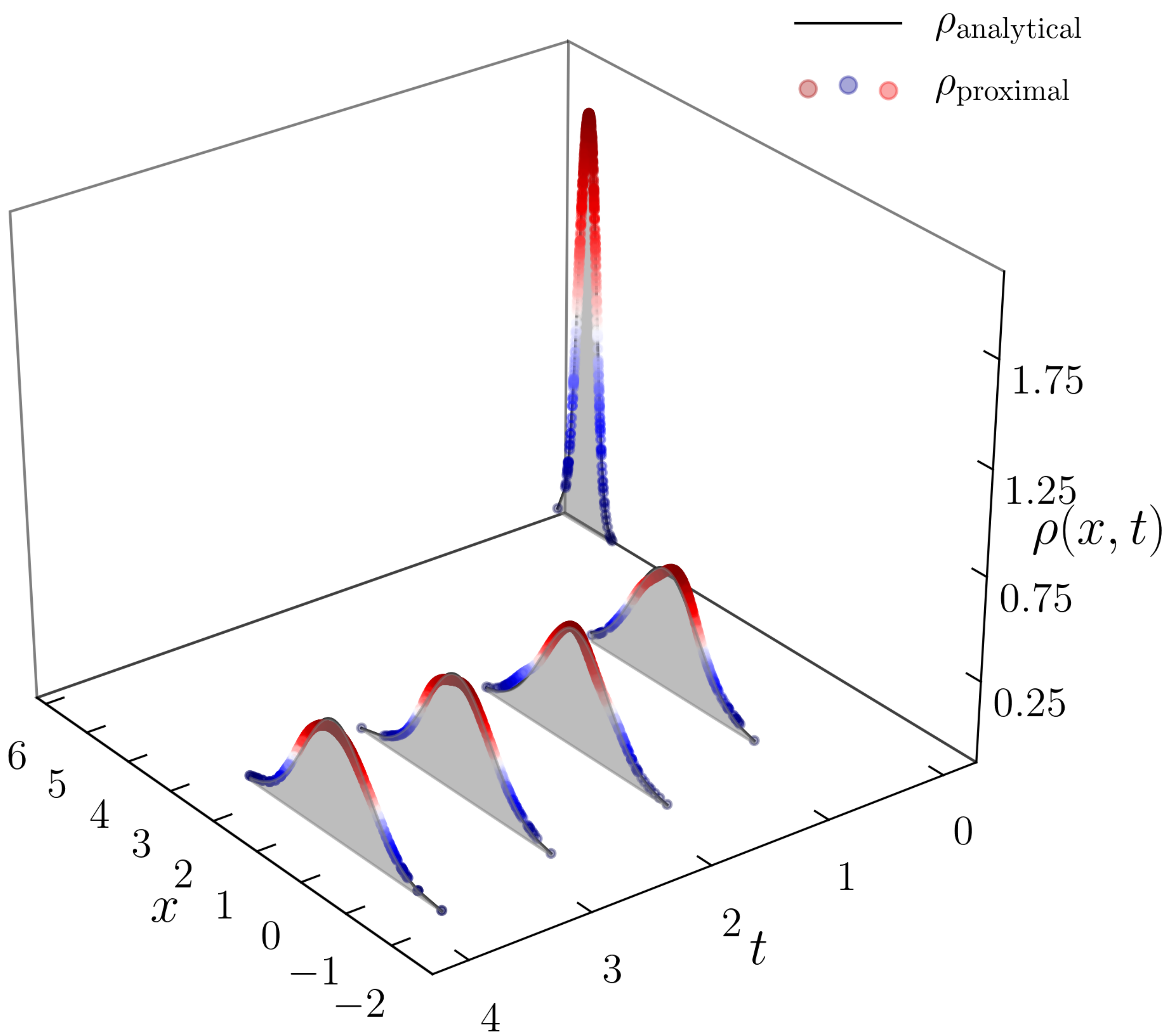
Then the solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ gives the proximal update $\varrho_k = \boldsymbol{z}^* \odot (\boldsymbol{\Gamma}_k{}^\top \boldsymbol{y}^*)$

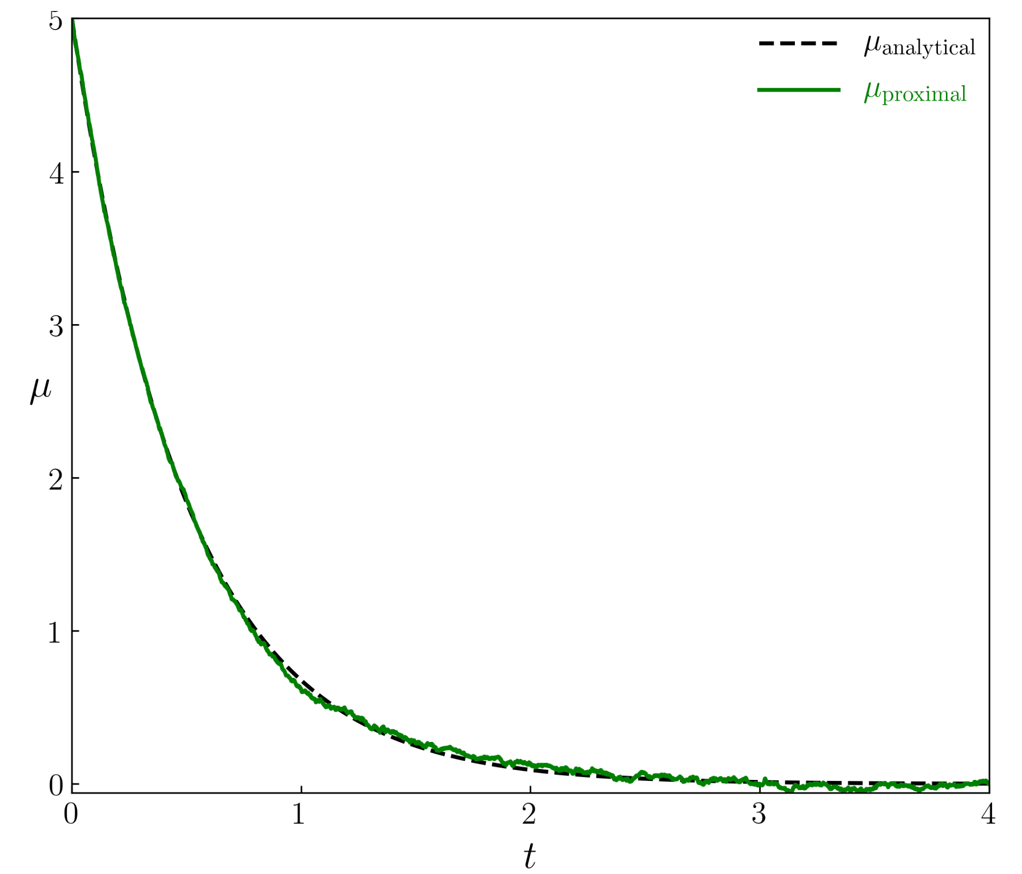
Algorithmic setup

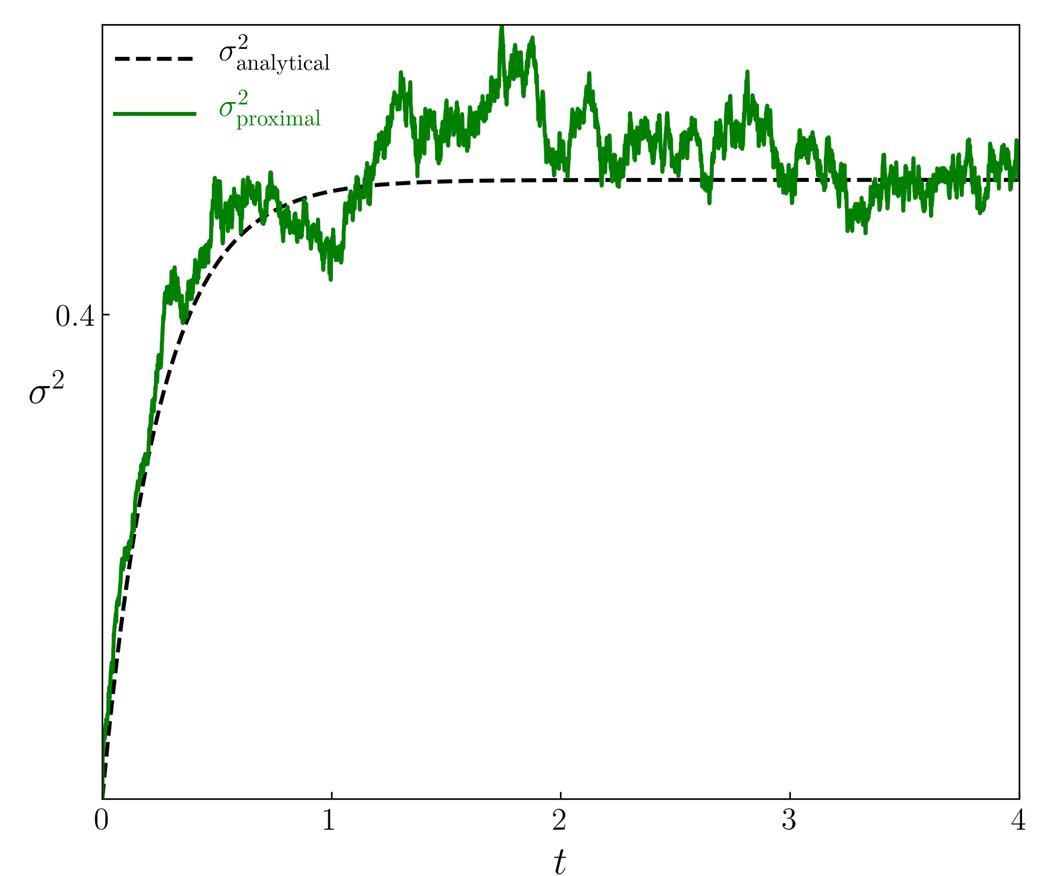


Theorem: Block co-ordinate iteration of $(\mathbf{y}, \mathbf{z})$ recursion is contractive on $\mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0}$.

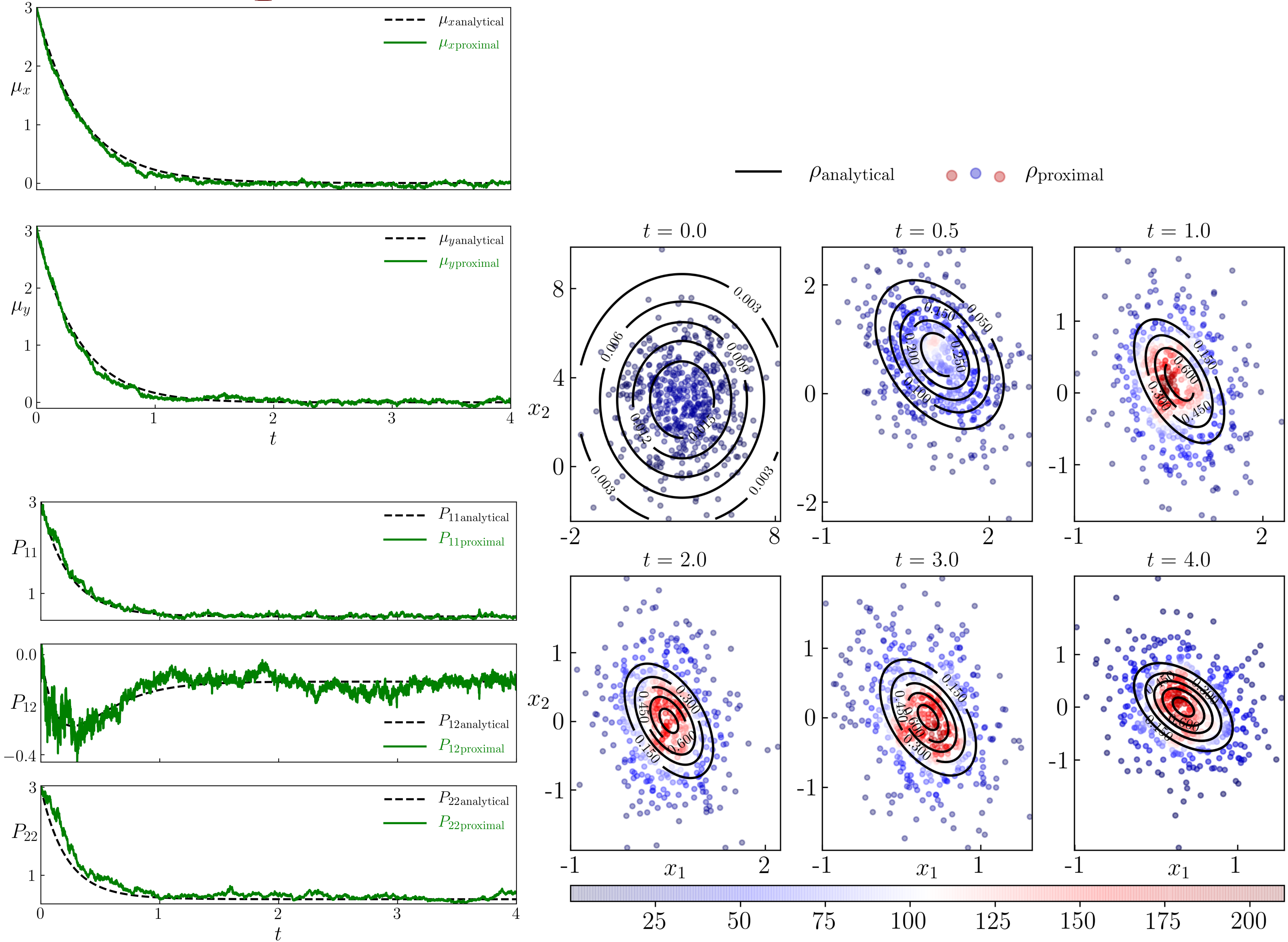
Proximal prediction: 1D linear Gaussian



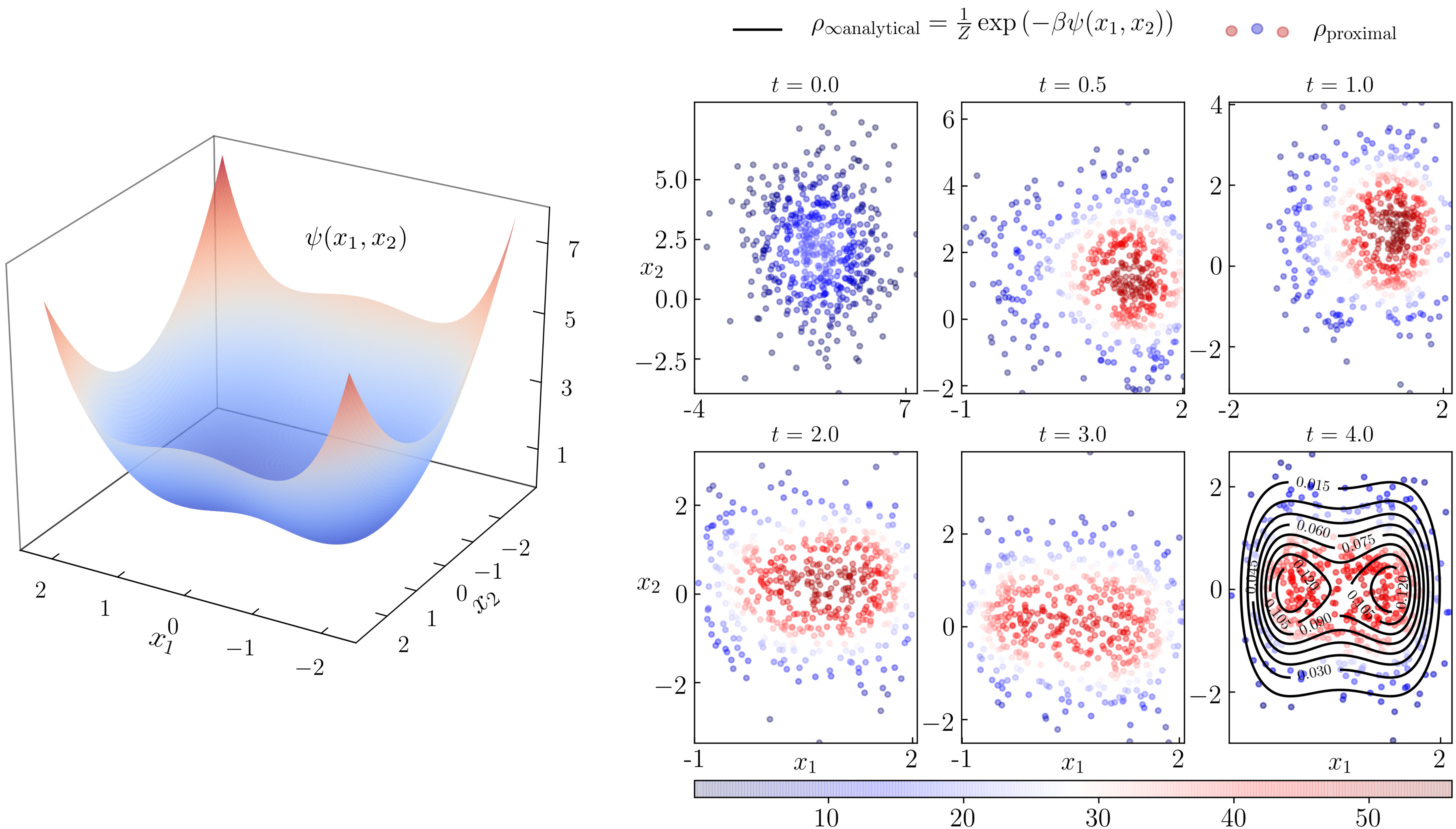




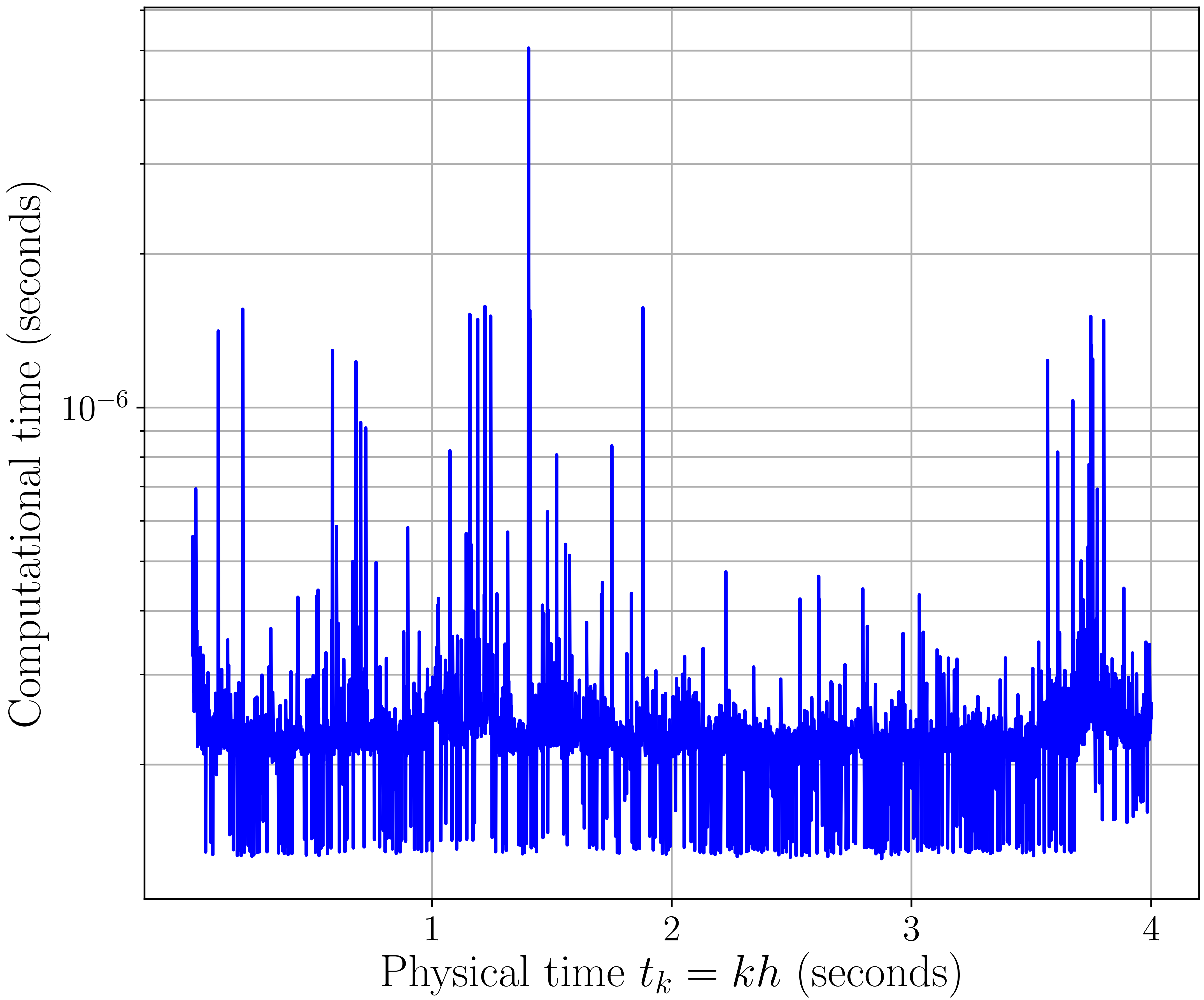
Proximal prediction: 2D linear Gaussian



Proximal prediction: nonlinear non-Gaussian



Computational time: nonlinear non-Gaussian



Details and more numerical case studies

K.F. Caluya, and A. H., Gradient flow algorithms for density propagation in stochastic systems , *IEEE Trans. Automatic Control*, 65(10), pp. 3991–4004, 2019.

K.F. Caluya, and A. H., Proximal recursion for solving the Fokker-Planck equation, *Proc. American Control Conference*, pp. 4098–4103, 2019.

Network reduced power system model

Noisy nonuniform Kuramoto model

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

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

$$\begin{pmatrix} \mathrm{d}\boldsymbol{\theta} \\ \mathrm{d}\boldsymbol{\omega} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\omega} \\ -\boldsymbol{M}^{-1}\nabla_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) - \boldsymbol{M}^{-1}\boldsymbol{\Gamma}\boldsymbol{\omega} \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \boldsymbol{0}_{n\times n} \\ \boldsymbol{M}^{-1}\boldsymbol{\Sigma} \end{pmatrix} \mathrm{d}\boldsymbol{w}$$

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

$$V(\boldsymbol{\theta}) := -\sum_{i=1}^{n} P_i\theta_i + \sum_{i<j} k_{ij}(1 - \cos(\theta_i - \theta_j - \varphi_{ij}))$$

Network reduced power system model

Kron reduced admittance matrix

$$\boldsymbol{Y} = [Y_{ij}]_{i,j=1}^{n} \in \mathbb{C}^{n\times n}, \quad Y_{ij} = Y_{ji}$$

Internal voltage and current for generator i

$$E_i, I_i \in \mathbb{C}$$

Parameters in the dynamics

$$P_i = P_i^{\mathrm{mech}} - P_i^{\mathrm{load}} - |E_i|^2 \Re(Y_{ii}) + \Re(E_i \cdot I_i^*),$$

$$\varphi_{ij} = \begin{cases} -\arctan\left(\frac{\Re(Y_{ij})}{\Im(Y_{ij})}\right), & \text{if } i \neq j, \\ 0, & \text{otherwise}, \end{cases}$$

$$k_{ij} = \begin{cases} |E_i||E_j||Y_{ij}|, & \text{if } i \neq j, \\ 0, & \text{otherwise}, \end{cases}$$

Not straightforward to apply the prox recursion

Kinetic Fokker-Planck PDE

$$\frac{\partial \rho}{\partial t} = -\langle \boldsymbol{\omega}, \nabla_\theta \rho \rangle + \nabla_{\boldsymbol{\omega}} \cdot \left(\rho \left(\boldsymbol{M}^{-1} \boldsymbol{\Gamma} \boldsymbol{\omega} + \boldsymbol{M}^{-1} \nabla_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) + \frac{1}{2} \boldsymbol{M}^{-1} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^\top \boldsymbol{M}^{-1} \nabla_{\boldsymbol{\omega}} \log \rho \right) \right)$$

Einstein relation does not hold ...

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top = \beta^{-1}\left(\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^\top\right) \quad \text{for some } \beta > 0$$

Consider change of variable $\begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{pmatrix} \mapsto \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} := \boldsymbol{\Psi} \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{pmatrix}$

where invertible linear map $\boldsymbol{\Psi} := \boldsymbol{I}_2 \otimes \left(\boldsymbol{M}\boldsymbol{\Sigma}^{-1}\right)$

From anisotropic to isotropic diffusion

Transformed SDE

$$\begin{pmatrix} \mathrm{d}\boldsymbol{\xi} \\ \mathrm{d}\boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\eta} \\ -\boldsymbol{\Upsilon}\nabla_{\boldsymbol{\xi}} U(\boldsymbol{\xi}) - \nabla_{\boldsymbol{\eta}} F(\boldsymbol{\eta}) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \mathbf{0}_{n\times n} \\ \boldsymbol{I}_n \end{pmatrix} \mathrm{d}\boldsymbol{w}$$

Set up

$$\begin{aligned} \tilde{\varrho}_k &= \mathrm{prox}_{h\widetilde{\Phi}}^{\widetilde{W}}(\tilde{\varrho}_{k-1}) \\ &\equiv \underset{\tilde{\rho}\in\mathcal{P}_2}{\arg\inf} \frac{1}{2}\widetilde{W}^2(\tilde{\varrho}, \tilde{\varrho}_{k-1}) + h\widetilde{\Phi}(\tilde{\rho}), \qquad \tilde{\varrho}_0 := \tilde{\rho}_0, \end{aligned}$$

where

$$\widetilde{\Phi}(\tilde{\rho}) := \int_{\mathbb{T}^n\times\mathbb{R}^n} \left(F(\boldsymbol{\eta}) + \frac{1}{2}\log\tilde{\rho} \right) \tilde{\rho}\,\mathrm{d}\boldsymbol{\xi}\mathrm{d}\boldsymbol{\eta}$$

From anisotropic to isotropic diffusion

Consistency guarantee

$$\tilde{\varrho}_k(\boldsymbol{\xi},\boldsymbol{\eta}) \xrightarrow{h\downarrow 0} \tilde{\varrho}(t=kh,\boldsymbol{\xi},\boldsymbol{\eta}) \text{ in } L^1(\mathbb{T}^n \times \mathbb{R}^n)$$

Evolve weighted point cloud $\left\{\tilde{\boldsymbol{x}}_k^i,\, \tilde{\varrho}_k^i\right\}_{i=1}^N$

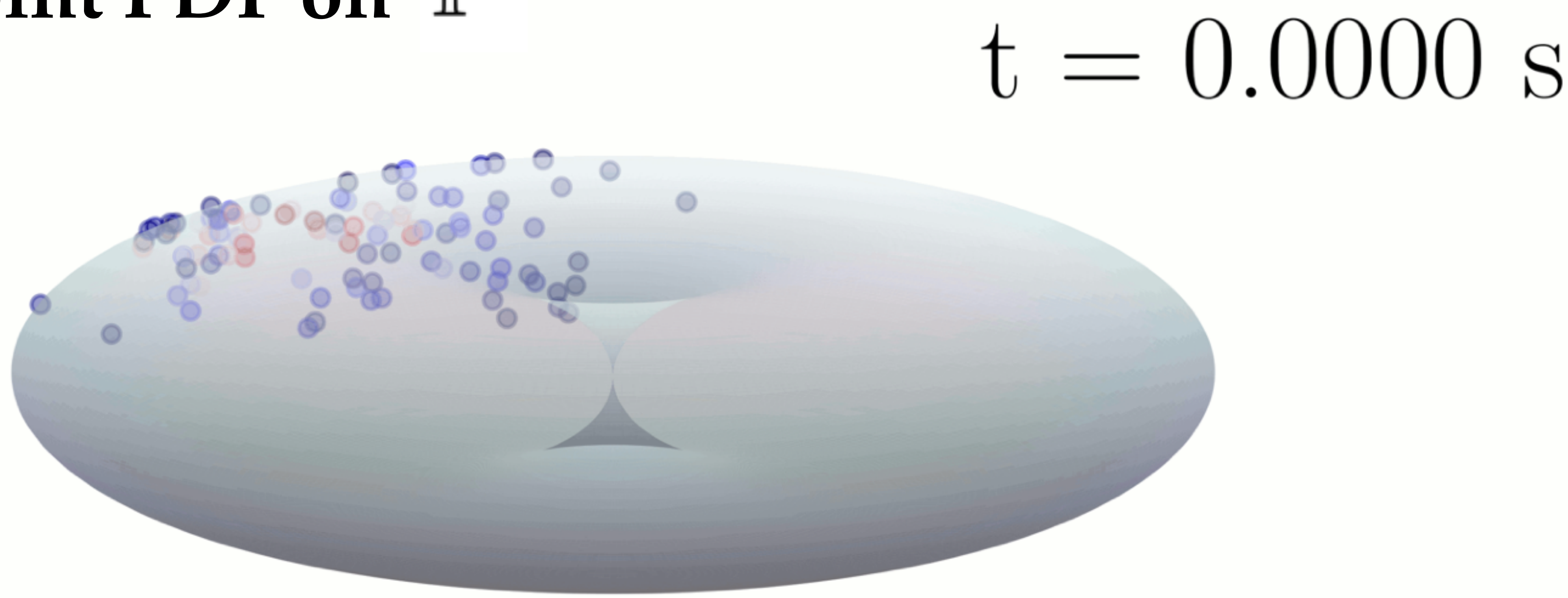
where

$$\tilde{\boldsymbol{x}}_k^i := \left(\boldsymbol{\xi}_k^i, \boldsymbol{\eta}_k^i\right)^\top, \quad i=1,\ldots,N, \quad \boldsymbol{k} \in \mathbb{N}$$

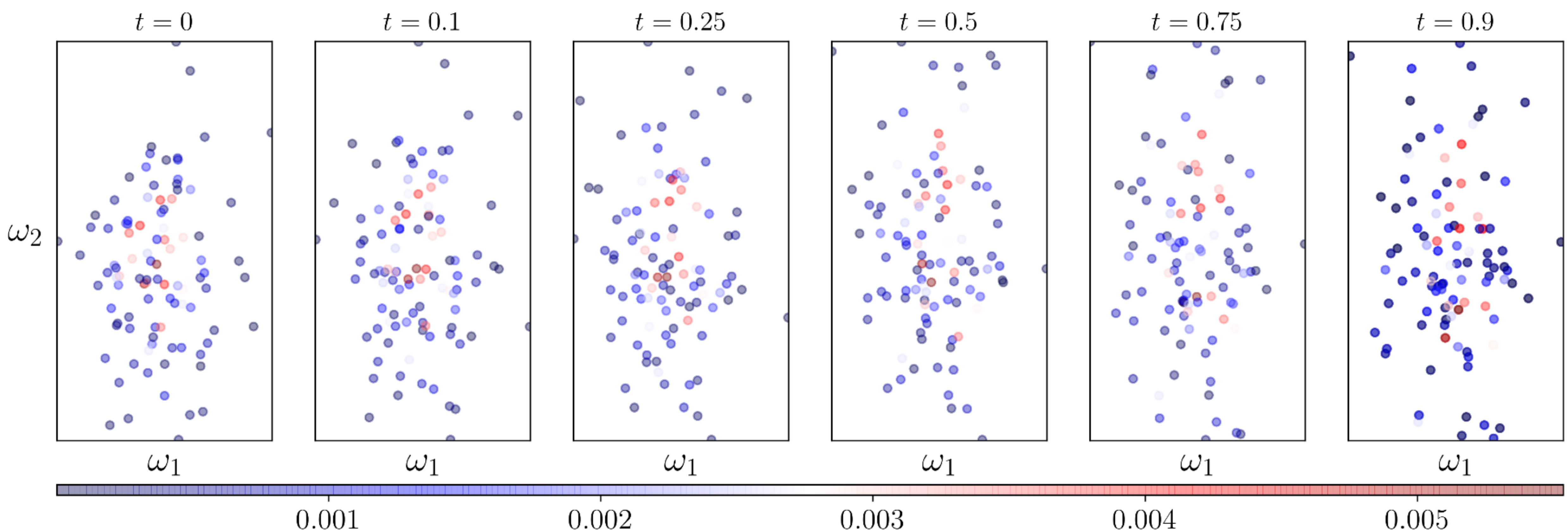
and then push forward via inverse map to come back to $(\boldsymbol{\theta},\boldsymbol{\omega})$ coordinate

Proximal Prediction: Power System with $n = 2$

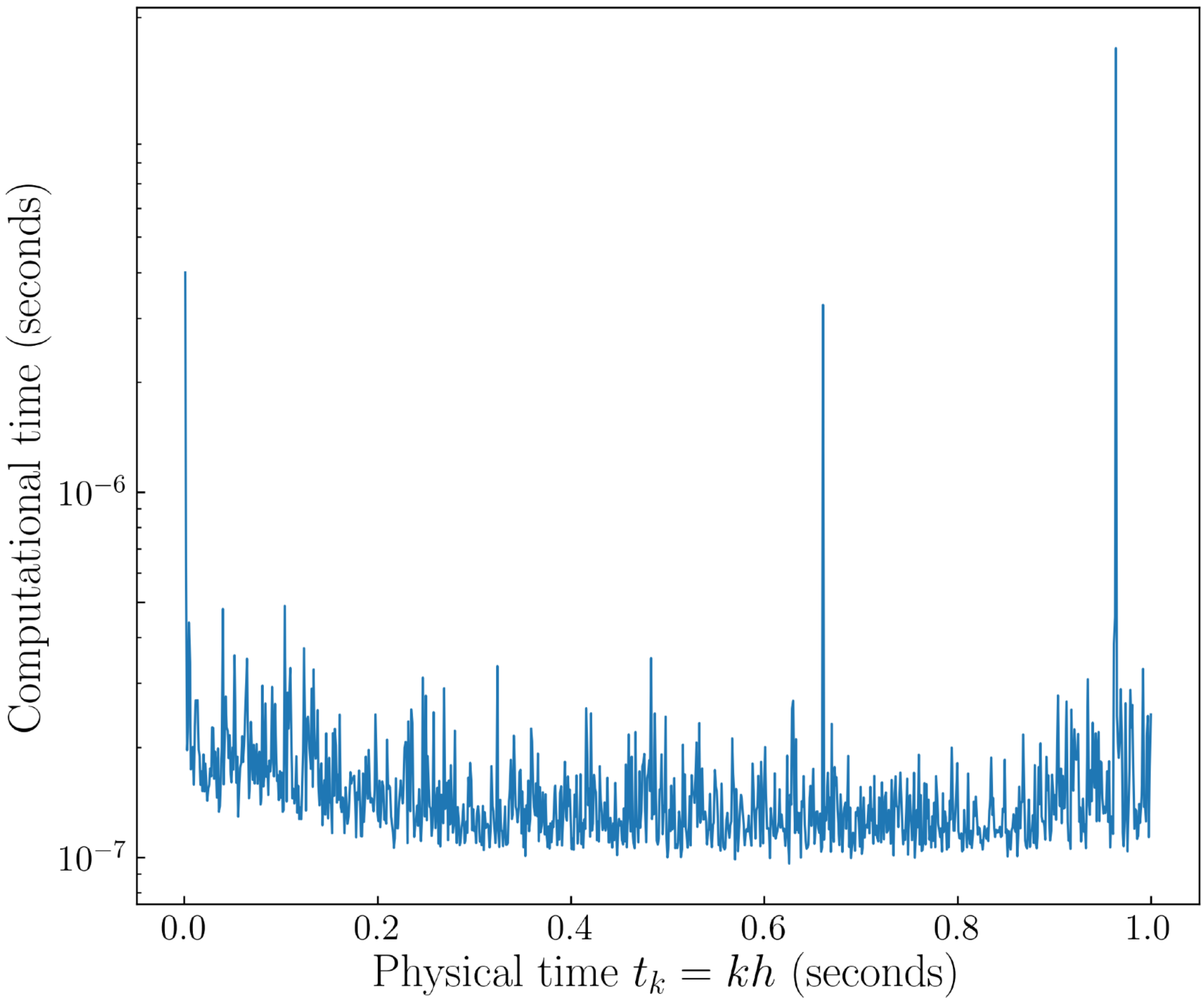
Projection of the joint PDF on $\mathbb{T}^2$



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



Computational Time: Power System with $n = 2$



IEEE 14 bus case study

Parameters from MATPOWER and ANDES

Nominal case (Case I): $P_i^{\text{mech}}, P_i^{\text{load}}$ from steady state power flow

Post-contingency case (Case II): Line 13 fails at $t = 0$

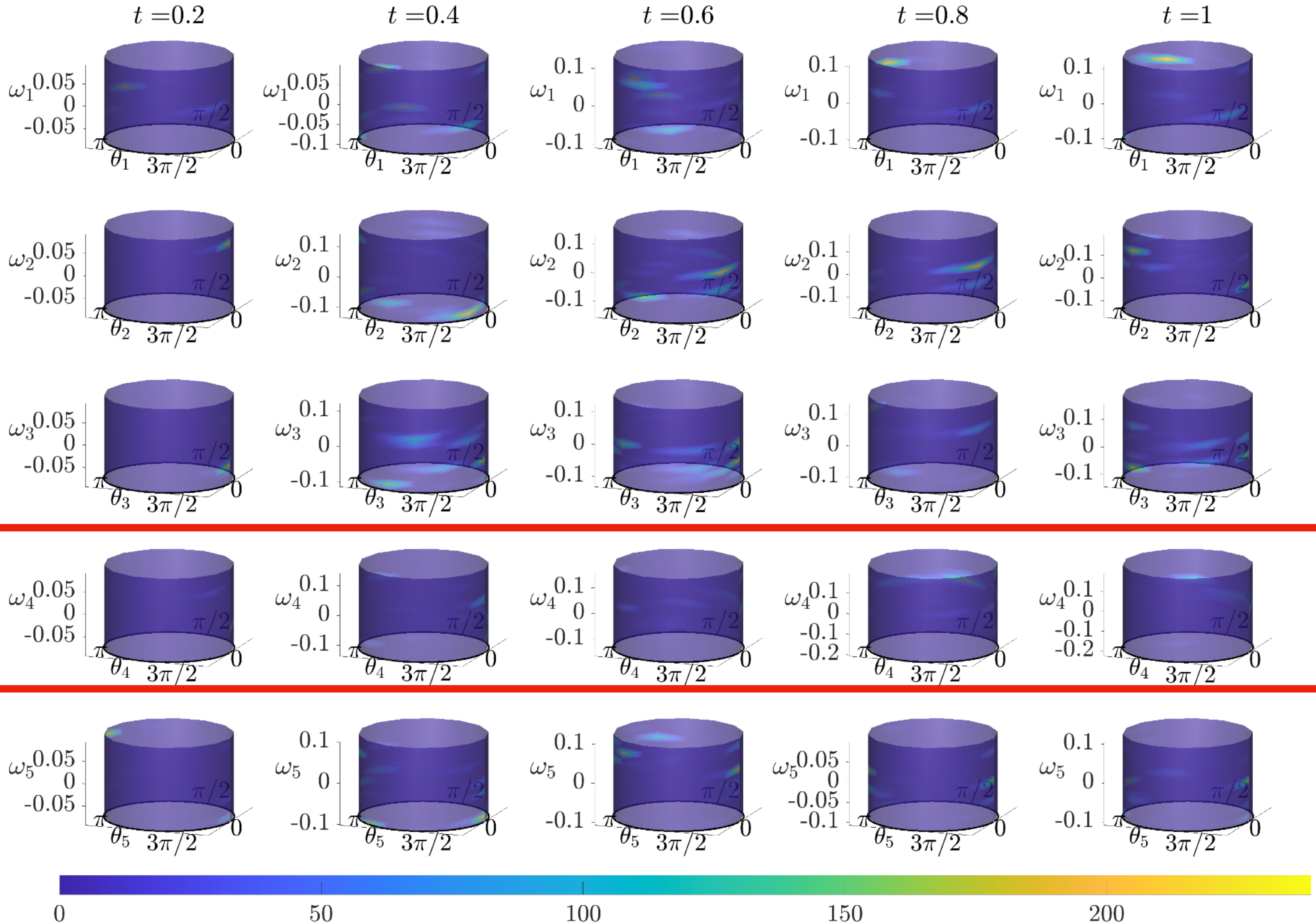
State space: $\mathbb{T}^5 \times \mathbb{R}^5$

Can account time-varying $P_i^{\text{mech}}, P_i^{\text{load}}$

Initial $\mathbb{T}^n$ marginal as product von Mises with mean angles from the steady state AC power flow

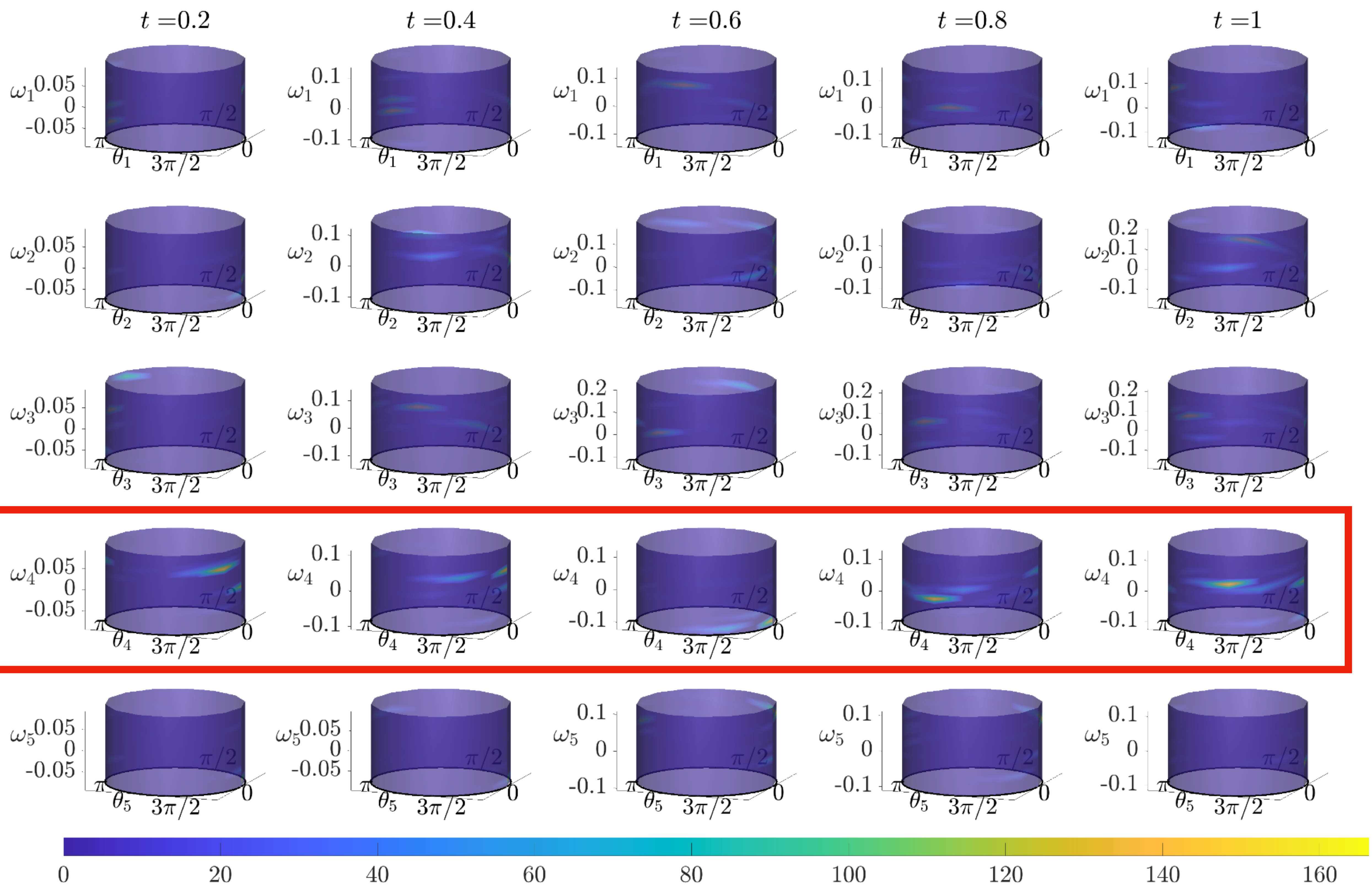
IEEE 14 bus case study: bivariate (θ, ω) marginals

Nominal case:

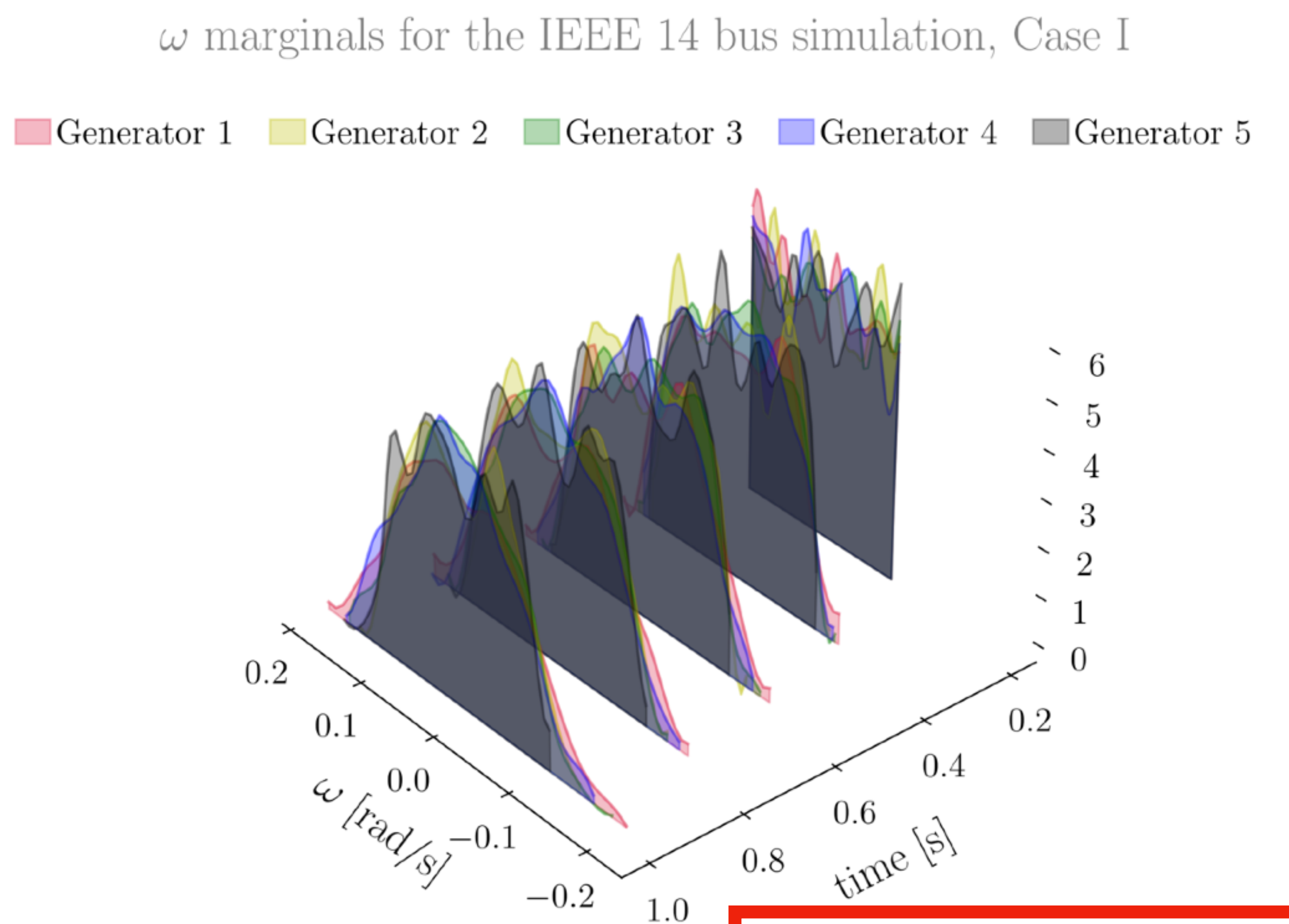


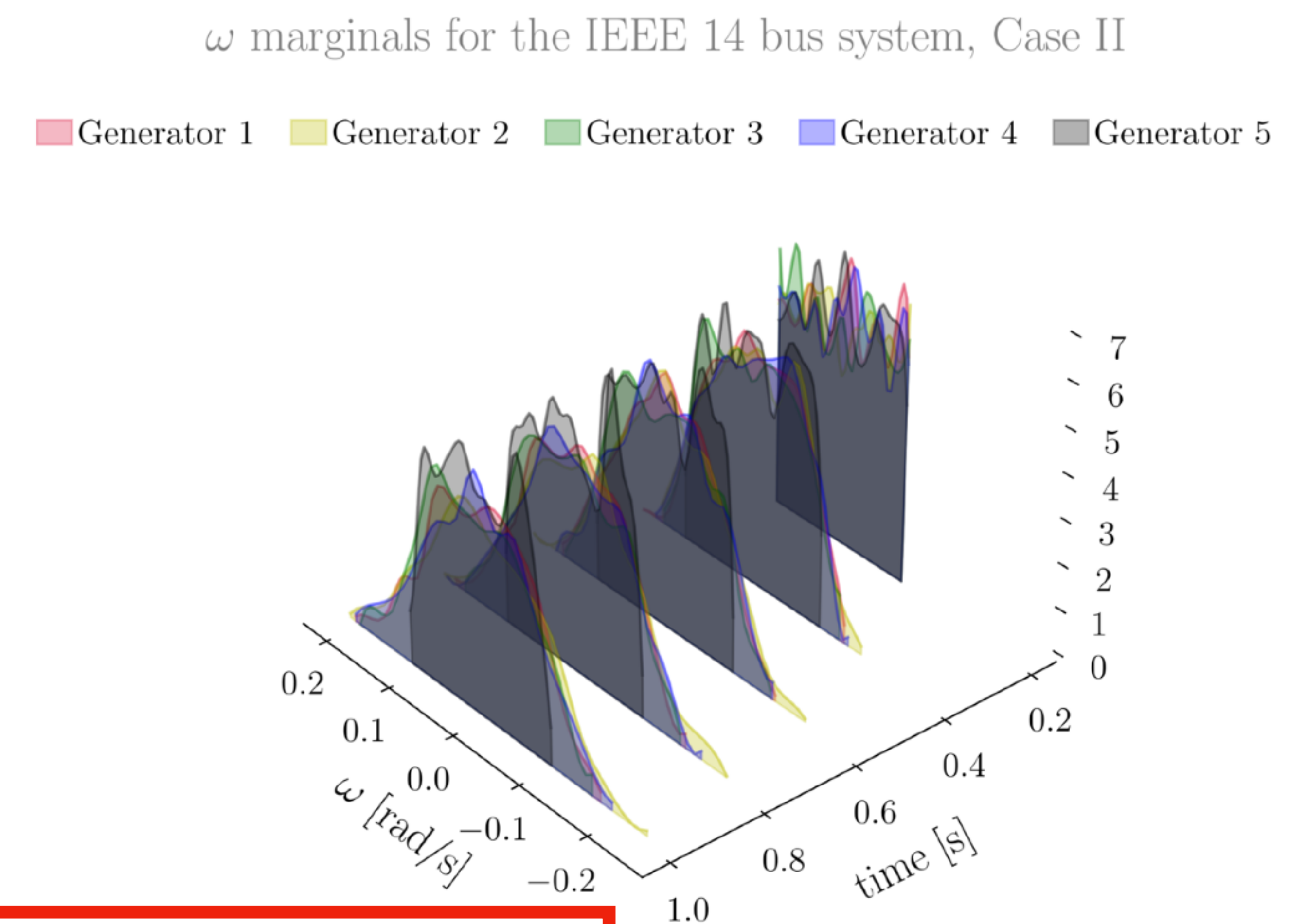
IEEE 14 bus case study: bivariate (θ, ω) marginals

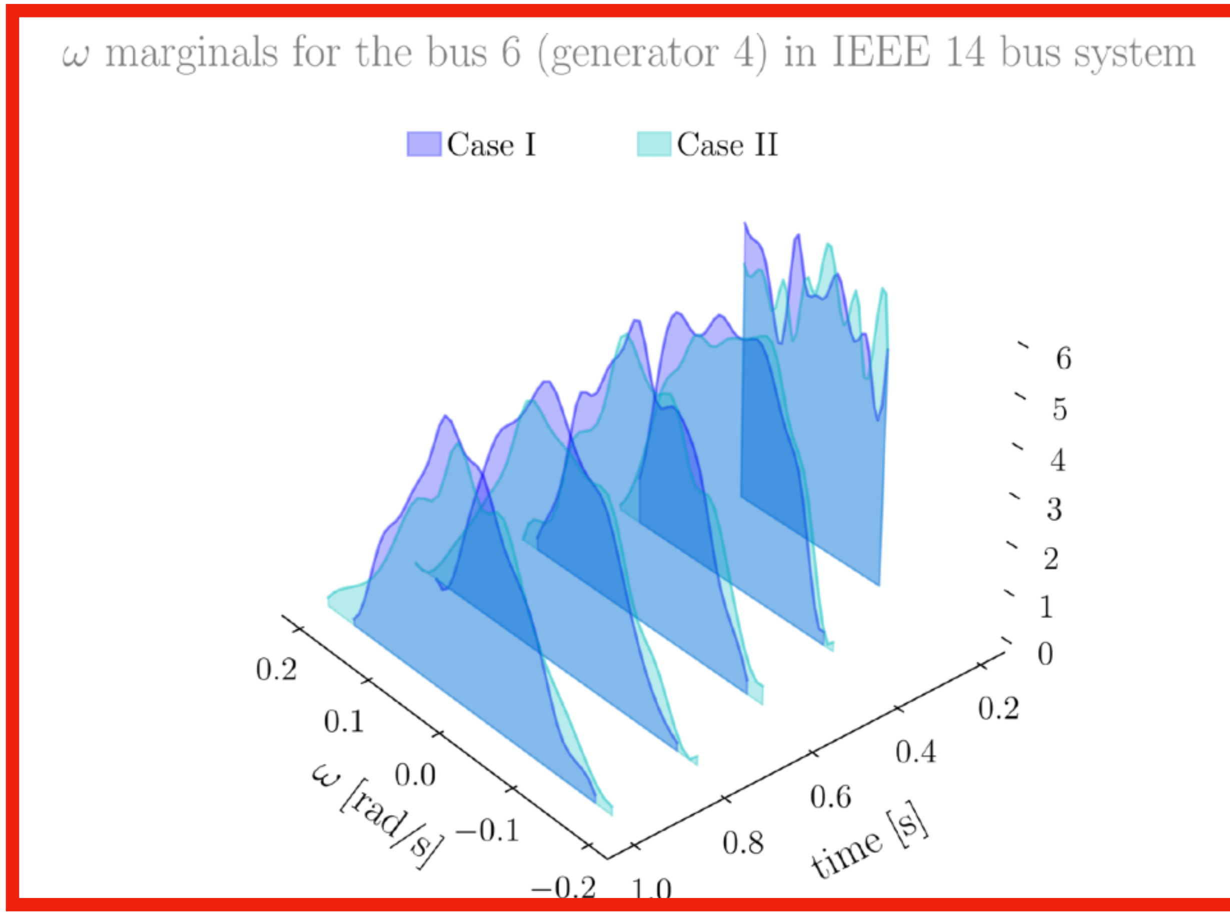
Post-contingency case:



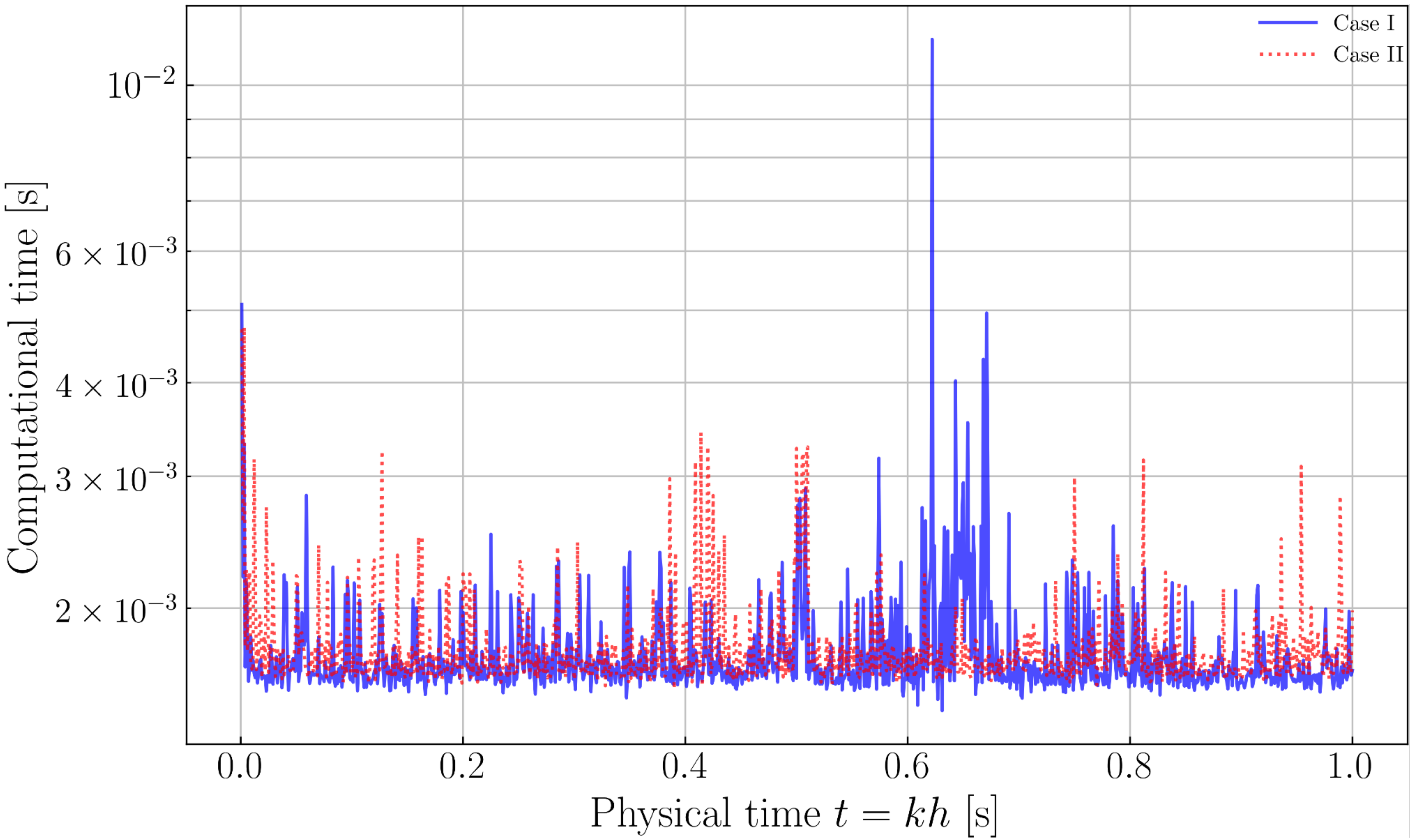
IEEE 14 bus case study: univariate ω marginals



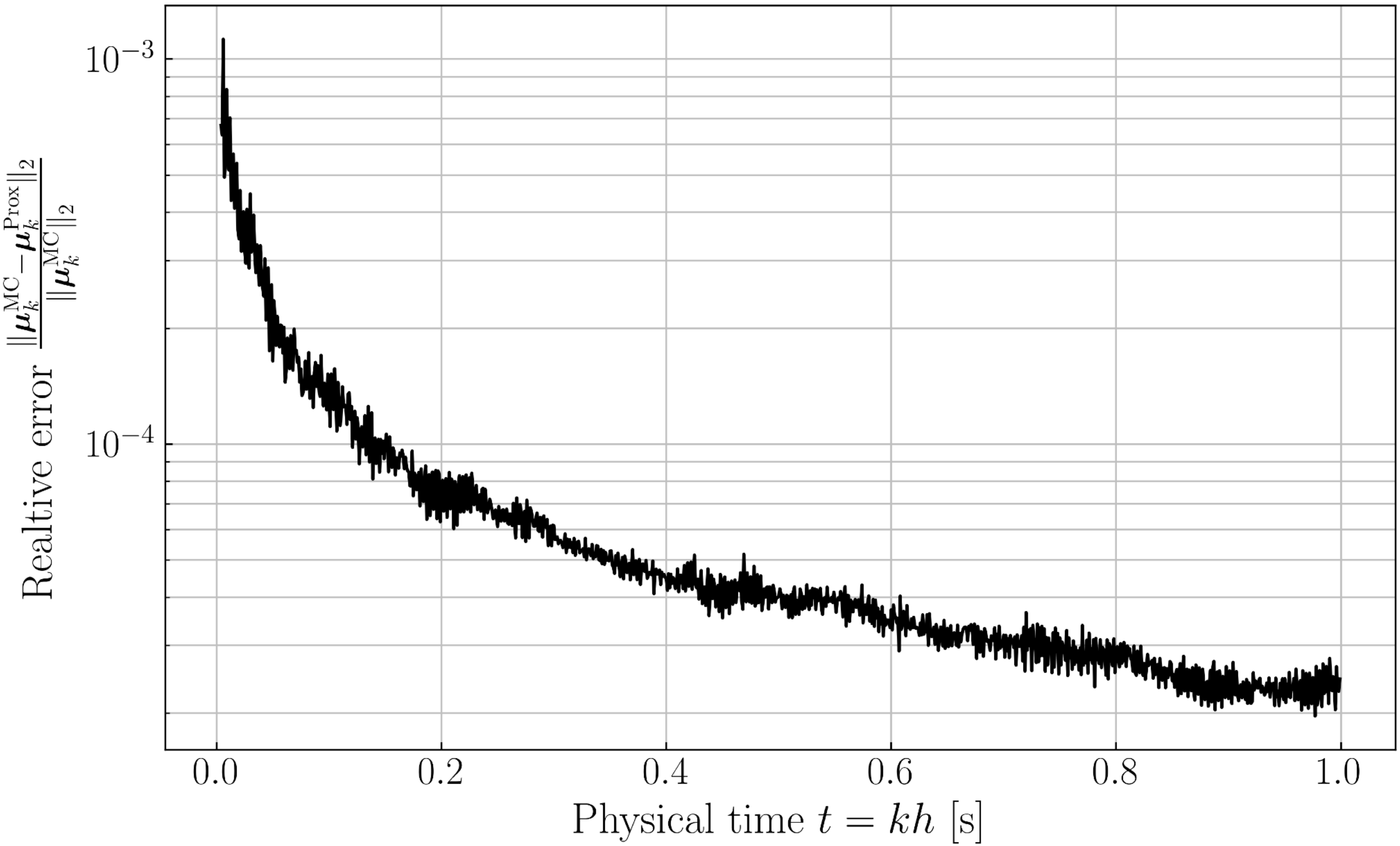




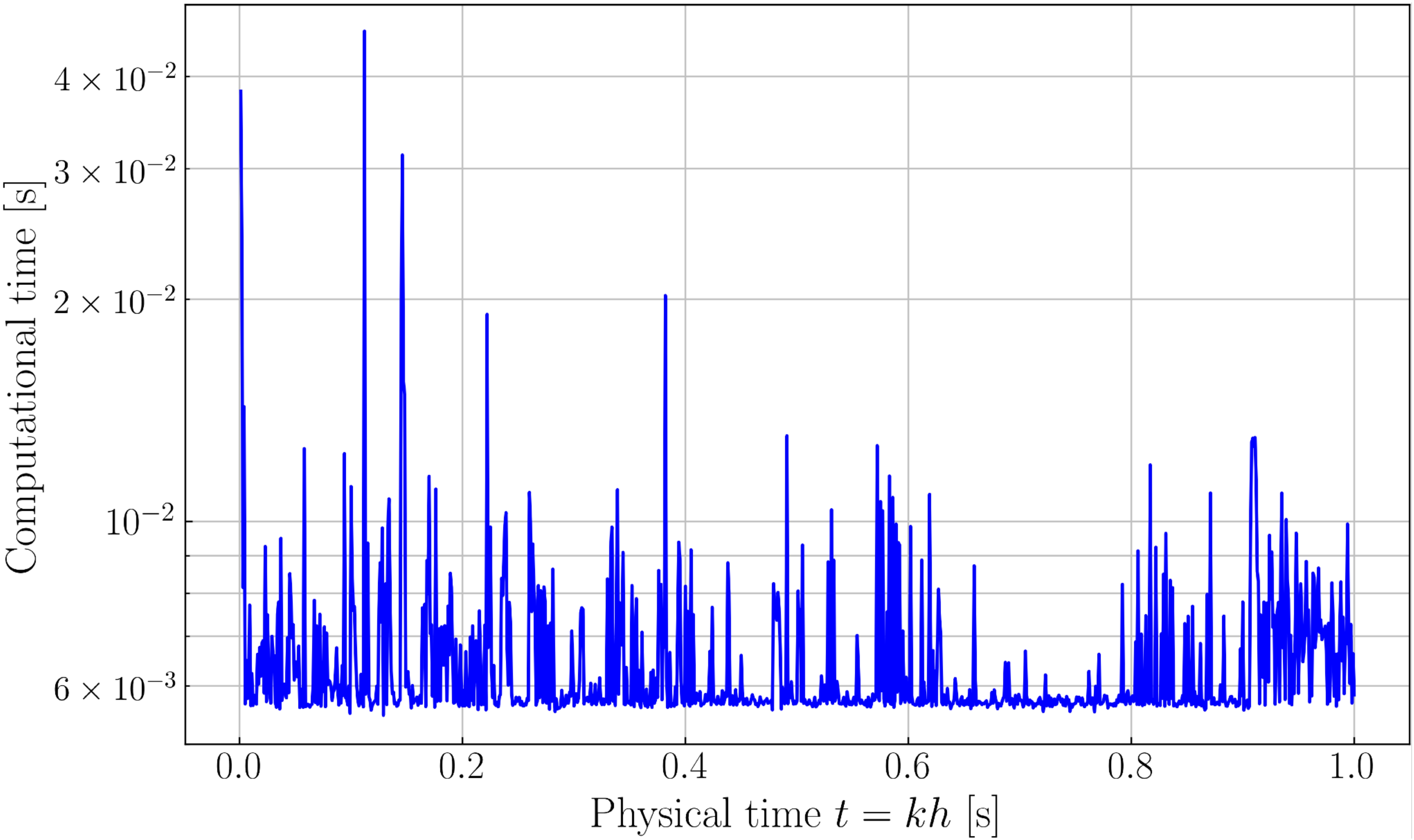
IEEE 14 bus case study: computational time



Synthetic system: $n = 50$: accuracy



Synthetic system: $n = 50$: computational time



Summary

Fast proximal recursions for PDF propagation in power systems

Ongoing work

Large scale implementation: ~1000 generators in ~seconds

Address Stochastic Differential Algebraic Equations (SDAEs)

More complex models

Control of distributional uncertainties

Thank You

Support: