Kalman-Wasserstein Gradient Flows

Franca Hoffmann

Computing and Mathematical Sciences
California Institute of Technology

Young Researchers Workshop: Ki-Net 2012-2019
CSCAMM, Oct 22, 2019

University of Maryland.

CliMA
CLIMATE MODELING ALLIANCE

Caltech

collaboration with:

Alfredo Garbuno-Inigo, Wuchen Li, Andrew Stuart (2019)
arXiv preprint, 1903.08866

accepted in:

SIAM Journal on Applied Dynamical Systems (SIADS)

The Big Picture

High-Level Overview

- Parameter calibration and uncertainty in complex computer models.
- Optimization approach and least squares.
- Bayesian approach and sampling.
- Ensemble Kalman Inversion (for optimization).
- Ensemble Kalman Sampling (for sampling).
- Gaussian Process Regression (for better sampling).
- Kalman-Wasserstein gradient flow structure.

Preview

Gradient Flow Structure

Inverse Problems

Numerics

The EKS and Mean Field Limits

Kalman-Wasserstein Space

Convergence to Equilibrium

Algorithmic Framework

Gradient Flow Structure

Optimization

Goal (Finite Dimensions)

Minimize $E : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^N$.

- **Dynamical Formulation**: find $\operatorname{argmin} E(x)$ by solving

$$\dot{x}(t) = -\nabla E(x(t)) \,, \qquad x(t=0) = x_0 \,.$$

Goal (Infinite Dimensions)

Minimize $E : \mathcal{P}(\Omega) \to \mathbb{R}$, where $\mathcal{P}(X) = \{\rho \in L^1(\Omega) \,|\, \int_\Omega \mathrm{d}\rho = 1\}$.

- **Dynamical Formulation**: find $\operatorname{argmin} E(\rho)$ by solving

$$\partial_t \rho = -\text{''}\nabla E(\rho)\text{''} \,, \qquad \rho(t=0) = \rho_0 \,.$$

Quadratic Wasserstein Metric $\mathcal{W}\colon \mathcal{P} \times \mathcal{P} \to \mathbb{R}$

Optimal Transport Formulation

For $\mu, \nu \in \mathcal{P}(\Omega)$,

$$W(\mu,\nu)^2 := \inf_{\gamma} \int_{\Omega\times\Omega} |x-y|^2 \mathrm{d}\gamma(x,y)$$

$$\text{subject to} \quad \gamma(A,\Omega) = \mu(A)\,, \quad \gamma(\Omega,B) = \nu(B) \quad \text{for all } A, B \subset \Omega\,.$$

Dynamic Formulation

For $\mu, \nu \in \mathcal{P}(\Omega)$,

$$W(\mu,\nu)^2 := \inf_{(\rho_t,\phi_t)} \int_0^1 \int_{\Omega} \langle \nabla\phi_t\,,\, \nabla\phi_t \rangle \, \mathrm{d}\rho_t(x)\mathrm{d}t$$

$$\text{subject to} \quad \partial_t\rho_t + \nabla\cdot(\rho_t\nabla\phi_t) = 0,\;\; \rho_0 = \mu,\;\; \rho_1 = \nu\,.$$

Wasserstein Gradient Flow

Tangent Vectors

- Gradient in Wasserstein metric:
$$\nabla_W E(\rho) = -\nabla \cdot \left(\rho \nabla \frac{\delta E(\rho)}{\delta \rho} \right) .$$

- First variation (formal definition):
$$\lim_{\varepsilon \to 0} \left(\frac{E(\rho + \varepsilon\varphi) - E(\rho)}{\varepsilon} \right) = \int_\Omega \frac{\delta E(\rho)}{\delta \rho}[x] \varphi(x) \, \mathrm{d}x \, , \qquad \forall \, \varphi \, .$$

[Ambrosio, Gigli, Savaré 2005]

Gradient Flow Structure

PDEs as GFs

Any PDE of the form

$$\partial_t \rho = \nabla \cdot (\rho \mathbf{v})$$

can be interpreted as a W-GF if there exists an energy $E : \mathcal{P}(\Omega) \to \mathbb{R}$ such that

$$\mathbf{v} = \nabla \frac{\delta E(\rho)}{\delta \rho}\,.$$

[Jordan, Kinderlehrer, Otto 1998], [Otto 2001]

Kalman-Wasserstein Space

Gradient Flow Structure

$$\partial_t \rho = \nabla \cdot \left(\rho \, \mathcal{C}(\rho) \, \nabla \frac{\delta E(\rho)}{\delta \rho} \right)$$

$$\mathcal{C}(\rho) = \int (\theta - \bar{\theta}) \otimes (\theta - \bar{\theta}) \rho(\theta, t) d\theta, \quad \bar{\theta} = \int \theta \rho(\theta, t) d\theta.$$

- Gradient flow in $\mathcal{C}(\rho)$-weighted metric.
- Suitable energy E: unique attractor is the posterior for an underlying inverse problem.
- Inspired new efficient derivative-free algorithm to solve inverse problems.

Solving Inverse Problems

Inverse Problem for Parameters

Find Parameter θ from Data y

Let $\mathsf{G} : \Theta \mapsto \mathcal{Y}$, and η be noise. Then data and parameter are related by

$$y = \mathsf{G}(\theta) + \eta, \quad \eta \sim \mathsf{N}(0, \gamma^2 I).$$

Our Setting

- Calibration and UQ for θ are both important.
- G is expensive to evaluate.
- Derivatives of G are not available.

Optimization Approach

Find Parameter θ from Data y

Let $\mathsf{G} : \Theta \mapsto \mathcal{Y}$, and η be noise. Then data and parameter are related by

$$y = \mathsf{G}(\theta) + \eta, \quad \eta \sim N(0, \gamma^2 I).$$

Mathematical Formulation

$$\theta^\star = \mathrm{argmin}_{\theta \in \Theta} \Phi_R(\theta; y),$$
$$\Phi_R(\theta; y) = \frac{1}{2\gamma^2}|y - \mathsf{G}(\theta)|^2 + \frac{1}{2}\langle \theta, \Sigma_0^{-1}\theta \rangle.$$

Algorithms: parameter θ calibration.

Bayesian Approach

- Prior: $\mathbb{P}(\theta)$; $\theta \sim \mathrm{N}(0, \Sigma_0)$
- Likelihood: $\mathbb{P}(y|\theta)$; $y - \mathrm{G}(\theta) \sim \mathrm{N}(0, \gamma^2 I)$
- Posterior: $\mathbb{P}(\theta|y)$; $\theta|y \sim ?$

Mathematical Formulation

$$\mathbb{P}(\theta|y) \propto \mathbb{P}(y|\theta) \times \mathbb{P}(\theta),$$

$$\mathbb{P}(\theta|y) \propto \exp\Big(-\frac{1}{2\gamma^2}|y - \mathrm{G}(\theta)|^2\Big) \times \exp\Big(-\frac{1}{2}\langle\theta, \Sigma_0^{-1}\theta\rangle\Big)$$

$$\propto \exp\Big(-\Phi_R(\theta; y)\Big)$$

Algorithms: parameter θ sampling.

Parameter Calibration & UQ

$$\mathbb{P}(\theta|y) \propto \exp\Big(-\Phi_R(\theta; y)\Big) \qquad \Phi_R(\theta; y) = \frac{1}{2\gamma^2}|y - \mathsf{G}(\theta)|^2 + \frac{1}{2}\langle\theta, \Sigma_0^{-1}\theta\rangle$$

Our Setting

- Calibration: maximizing $\mathbb{P}(\theta|y)$ = minimizing $\Phi_R(\theta; y)$.
- UQ: sample from $\mathbb{P}(\theta|y)$.
- G is expensive to evaluate.
- Derivatives of G are not available.

Goals

- Methods to generate approximate samples from posterior.
- Mathematical framework for analysis using gradient flow structure.

Ensemble Kalman Inversion (EKI)

Continuous Time Formulation

$$\dot{\theta}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\frac{1}{\gamma^2}\left\langle \mathsf{G}(\theta^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(\theta^{(j)}) - y \right\rangle \left(\theta^{(k)} - \bar{\theta}\right),$$

$$\bar{\theta} = \frac{1}{J}\sum_{k=1}^{J}\theta^{(k)}, \qquad \bar{\mathsf{G}} = \frac{1}{J}\sum_{k=1}^{J}\mathsf{G}\big(\theta^{(k)}\big).$$

[Evensen 1994], [Inglesias, Law, Stuart 2013], [Ernst, Sprungk, Starkloff 2015], [Schillings, Stuart 2017], [Herty, Visconti 2019]

- Sample from prior instead of posterior.
- Evolve particle ensemble: derivative-free algorithm.
- Tool for optimization: collapses to max $\mathbb{P}(\theta|y)$.

Ensemble Kalman Sampling (EKS)

$$C(\theta) = \frac{1}{J}\sum_{k=1}^{J}\Big(\theta^{(k)} - \bar{\theta}\Big) \otimes \Big(\theta^{(k)} - \bar{\theta}\Big).$$

Continuous Time Formulation

$$\dot{\theta}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\frac{1}{\gamma^2}\Big\langle \mathsf{G}(\theta^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(\theta^{(j)}) - y\Big\rangle \Big(\theta^{(k)} - \bar{\theta}\Big)$$
$$- C(\theta)\Sigma_0^{-1}\theta^{(j)} + \sqrt{2C(\theta)}\dot{W}^{(j)}$$

[Garbuno-Inigo, Li, H., Stuart 2019 (preprint)]

- Add damping term related to prior **[Chada, Stuart, Tong 2019 (preprint)]**.
- Perturb particles instead of data **[Kovachki, Stuart 2019]**.
- New noise covariance structure.

Numerics:
EKI vs EKS vs MCMC

Example: Elliptic BVP

- 1D problem for $x \in [0,1]$,

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(\theta_1)\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1,$$

  with boundary conditions $p(0) = 0$ and $p(1) = \theta_2$.
- Explicit solution is available and we define

$$G(\theta) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}.$$

- Run EKI with $J = 1000$.

Example: Elliptic BVP

- 1D problem for $x \in [0, 1]$,

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(\theta_1)\,\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1,$$

  with boundary conditions $p(0) = 0$ and $p(1) = \theta_2$.

- Explicit solution is available and we define

$$G(\theta) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}.$$

- Run EKI with $J = 1000$.

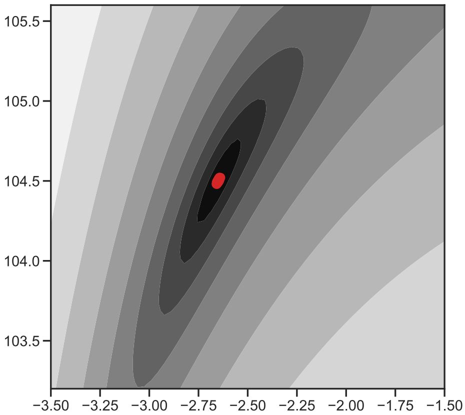


Figure: Contour plots: $\Phi(\theta) = \frac{1}{2\gamma^2}|y - G(\theta)|^2$.

Example: Elliptic BVP

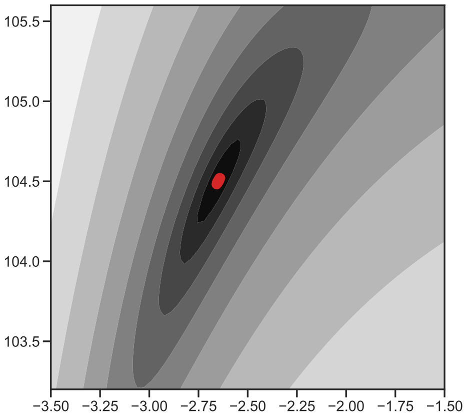
- 1D problem for $x \in [0, 1]$,

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(\theta_1)\,\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1,$$

  with boundary conditions $p(0) = 0$ and $p(1) = \theta_2$.
- Explicit solution is available and we define

$$G(\theta) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}.$$

- Run EKS with $J = 1000$.
- Compare with exact MCMC.



Example: Elliptic BVP

- 1D problem for $x \in [0,1]$,

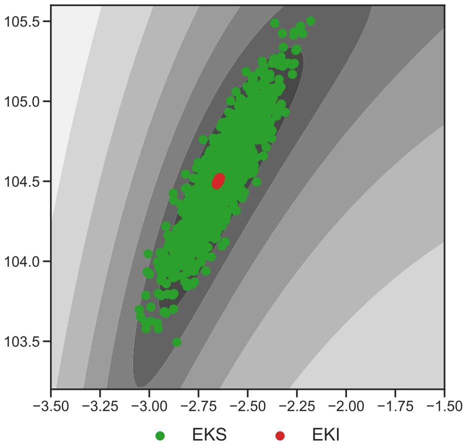
$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(\theta_1)\,\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1,$$

with boundary conditions $p(0) = 0$ and $p(1) = \theta_2$.

- Explicit solution is available and we define

$$G(\theta) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}.$$

- Run EKS with $J = 1000$.
- Compare with exact MCMC.



Example: Elliptic BVP

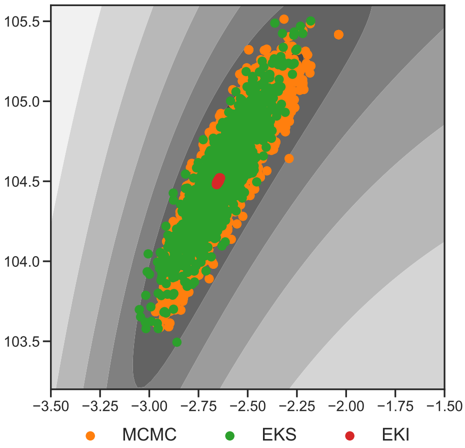
- 1D problem for $x \in [0,1]$,

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(\theta_1)\,\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1,$$

  with boundary conditions $p(0) = 0$ and $p(1) = \theta_2$.
- Explicit solution is available and we define

$$G(\theta) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}.$$

- Run EKS with $J = 1000$.
- Compare with exact MCMC.



From EKS to
Kalman-Wasserstein Gradient Flow

EKS: Approximation 1

$$\dot{\theta}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\frac{1}{\gamma^2}\left\langle \mathsf{G}(\theta^{(k)}) - \bar{\mathsf{G}}, \mathsf{G}(\theta^{(j)}) - y \right\rangle \left(\theta^{(k)} - \bar{\theta}\right) - C(\theta)\Sigma_0^{-1}\theta^{(j)} + \sqrt{2C(\theta)}\dot{W}^{(j)},$$

$$C(\theta) = \frac{1}{J}\sum_{k=1}^{J}\left(\theta^{(k)} - \bar{\theta}\right) \otimes \left(\theta^{(k)} - \bar{\theta}\right).$$

Linear Approximation

$$\left(\mathsf{G}(\theta^{(k)}) - \bar{\mathsf{G}}\right) \approx A(\theta^{(k)} - \bar{\theta}), \qquad A := D\mathsf{G}(\theta^{(j)}).$$

EKS: Approximation 1

Preconditioned Langevin Equation

$$\dot{\theta}^{(j)} = -C(\theta)\nabla\Phi_R(\theta^{(j)}) + \sqrt{2C(\theta)}\dot{W}^{(j)}.$$

$$\Phi_R(\theta) \approx \frac{1}{2\gamma^2}\left|y - A\theta\right|^2 + \frac{1}{2}\left|\Sigma_0^{-\frac{1}{2}}\theta\right|^2.$$

- If G linear, this is EKS.
- If G non-linear, expect solution to be close to EKS solution if particles close together (conjecture!).
- Preconditioning accelerates [Matthews, Leimkuhler, Weare 2018].

EKS: Approximation 2

$$\mathcal{C}(\rho) := \int (\theta - \bar{\theta}) \otimes (\theta - \bar{\theta}) \rho(\theta, t) d\theta, \qquad \bar{\theta} := \int \theta \rho(\theta, t) d\theta.$$

Mean Field Limit

$$\dot{\theta} = -\mathcal{C}(\rho) \nabla \Phi_R(\theta) + \sqrt{2\mathcal{C}(\rho)} \dot{W} .$$

Nonlinear Fokker-Planck Equation

$$\partial_t \rho = \nabla \cdot \left(\rho \, \mathcal{C}(\rho) \, \nabla \Phi_R \right) + \mathcal{C}(\rho) : D^2 \rho,$$

[Li, Din 2019 (preprint)]

Connection To Bayesian Inversion

$$\partial_t \rho = \nabla \cdot (\rho \mathcal{C}(\rho) \nabla (\Phi_R + \log \rho)) \,, \qquad E(\rho) = \int \rho (\phi_R + \log \rho) \; d\theta \,.$$

Manifold of Stationary States

$$\rho(\theta) = \delta_v(\theta) \text{ for some } v \in \mathbb{R}^d \quad \Leftrightarrow \quad \mathcal{C}(\rho) = 0 \,.$$

Steady State

Equilibrium solution to non-linear Fokker-Planck equation:

$$\rho_\infty(\theta) := \frac{e^{-\Phi_R(\theta)}}{\int e^{-\Phi_R(\theta)} \, \mathrm{d}\theta} \,.$$

This is the density of $\mathbb{P}(\theta|y)$.

Kalman-Wasserstein Space

Gradient Flow Structure

- For $\Omega \subseteq \mathbb{R}^d$ convex set, define $\mathcal{P}_+ := \{\rho \in \mathcal{P} : \rho > 0 \text{ a.e. }, \rho \in C^\infty(\Omega)\}$.

Kalman-Wasserstein Metric $\mathcal{W}_\mathcal{C} : \mathcal{P}_+ \times \mathcal{P}_+ \to \mathbb{R}$

For $\mu, \nu \in \mathcal{P}_+$,

$$W_\mathcal{C}(\mu,\nu)^2 := \inf_{(\rho_t,\phi_t)} \int_0^1 \int_\Omega \langle \nabla\phi_t \,, \mathcal{C}(\rho_t)\nabla\phi_t\rangle \; \rho_t \, \mathrm{d}x$$

$$\text{subject to} \quad \partial_t\rho_t + \nabla\cdot(\rho_t\,\mathcal{C}(\rho_t)\nabla\phi_t) = 0,\ \ \rho_0 = \mu,\ \ \rho_1 = \nu,$$

Theorem

Given a finite functional $E : \mathcal{P}_+ \to \mathbb{R}$, the gradient flow of $E(\rho)$ in $(\mathcal{P}_+, W_\mathcal{C})$ satisfies

$$\partial_t\rho = \nabla\cdot\left(\rho\,\mathcal{C}(\rho)\nabla\frac{\delta E}{\delta\rho}\right) .$$

[Garbuno-Inigo, Li, H., Stuart 2019 (preprint)]

- **Energy:** Kullback-Leibler divergence

$$E(\rho) = \int \left(\Phi_R + \ln \rho(t)\right) \rho(t) \, \mathrm{d}\theta$$
$$= \int \frac{\rho(t)}{\rho_\infty} \ln\left(\frac{\rho(t)}{\rho_\infty}\right) \rho_\infty \, \mathrm{d}\theta + \ln\left(\int e^{-\Phi_R(\theta)} \, \mathrm{d}\theta\right)$$
$$= \mathrm{KL}(\rho(t) \| \rho_\infty) + c$$

- **Euler-Lagrange condition:**

$$\frac{\delta E}{\delta \rho} = \Phi_R(\theta) + \ln \rho(\theta) = c \qquad \text{on } \mathrm{supp}\,(\rho)$$

- **Unique solution:** posterior

$$\rho_\infty(\theta) := \frac{e^{-\Phi_R(\theta)}}{\int e^{-\Phi_R(\theta)} \, \mathrm{d}\theta}.$$

- **Energy dissipation:**

$$\frac{\mathrm{d}}{\mathrm{d}t}E(\rho) = -\int \rho \left|\mathcal{C}(\rho)^{\frac{1}{2}}\nabla(\Phi_R + \ln\rho)\right|^2 \mathrm{d}\theta\,.$$

  Hence $E \searrow$ along paths until $\mathcal{C}(\rho) = 0$ or $\rho = \rho_\infty$.

- **Fisher-Information:** for any covariance matrix Λ,

$$\mathcal{I}_\Lambda(\rho\|\rho_\infty) := \int \rho \left\langle \nabla \ln\left(\frac{\rho}{\rho_\infty}\right), \Lambda\nabla \ln\left(\frac{\rho}{\rho_\infty}\right)\right\rangle \mathrm{d}\theta\,.$$

- **Kalman-Fisher information:** For $\Lambda = \mathcal{C}(\rho)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{KL}(\rho(t)\|\rho_\infty) = -\mathcal{I}_{\mathcal{C}}(\rho(t)\|\rho_\infty)\,.$$

Convergence to Equilibrium

Nonlinear forward map

Theorem (Decay to Equilibrium)

Assume $\alpha > 0$ and $\lambda > 0$ exists such that

$$\mathcal{C}(\rho(t)) \geq \alpha I_d \,, \qquad D^2\Phi_R \geq \lambda I_d \,.$$

If $\mathrm{KL}(\rho_0 \| \rho_\infty) < \infty$ then there is $c > 0$ such that

$$\|\rho(t) - \rho_\infty\|_{L^1(\mathbb{R}^d)} \leq c e^{-\alpha\lambda t} \,.$$

[Garbuno-Inigo, Li, H., Stuart 2019 (preprint)]

Proof

- $D^2\Phi_R \geq \lambda I_d$ guarantees log Sobolev inequality **[Bakry, Émery 1985]**:

$$\mathrm{KL}(\rho(t)\|\rho_\infty) \leq \frac{1}{2\lambda}\mathcal{I}_{I_d}\left(\rho(t)\|\rho_\infty\right) \qquad \forall \rho\,.$$

- $C(\rho(t)) \geq \alpha I_d$ gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{KL}(\rho(t)\|\rho_\infty) &= -\mathcal{I}_C(\rho(t)\|\rho_\infty) \\ &\leq -\alpha\mathcal{I}_{I_d}(\rho(t)|\rho_\infty) \leq -2\alpha\lambda\mathrm{KL}(\rho(t)\|\rho_\infty)\,.\end{aligned}$$

- By Csiszár-Kullback inequality:

$$\frac{1}{2}\|\rho(t)-\rho_\infty\|^2_{L^1(\mathbb{R}^d)} \leq \mathrm{KL}(\rho(t)\|\rho_\infty) \leq \mathrm{KL}(\rho_0\|\rho_\infty)e^{-2\alpha\lambda t}\,.$$

Linear forward map $G(\theta) = A\theta$

Theorem (Linear Inverse Problem).

- Closed equations for the moments.
- Mean field limt of EKS: Gaussians remain Gaussians.
- Gaussians converge exponentially fast to ρ_∞ in $L^1(\mathbb{R}^d)$ as $t \to \infty$.

[Garbuno-Inigo, Li, H., Stuart 2019 (preprint)]

Theorem (Decay to Equilibrium)

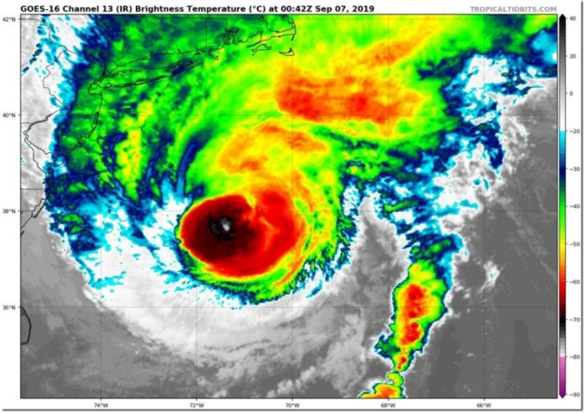
If ρ_t^1, ρ_t^2 solutions to KW-GF with initial conditions ρ_0^1, ρ_0^2 respectively, then under suitable conditions on $\rho_0^1, \rho_0^2, A, \Sigma_0, \gamma$,

$$W(\rho_t^1, \rho_t^2) \leq Ce^{-t} W(\rho_0^1, \rho_0^2)\,,$$

where C only depends on the first two moments of ρ_0^1, ρ_0^2, and onA, Σ_0.

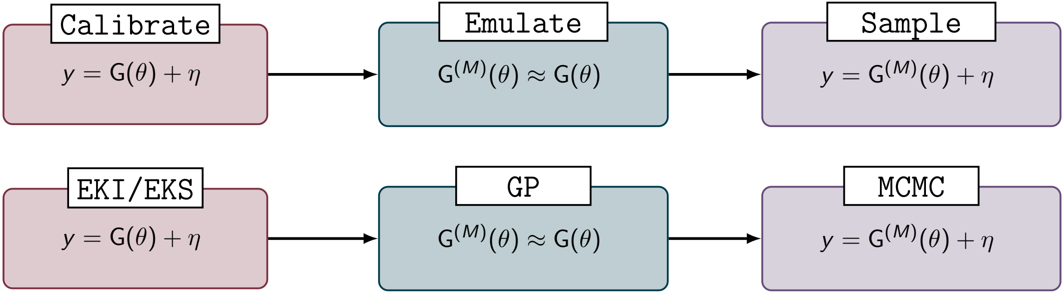
[Carrillo, Vaes 2019 (preprint)]

Calibrate - Emulate - Sample



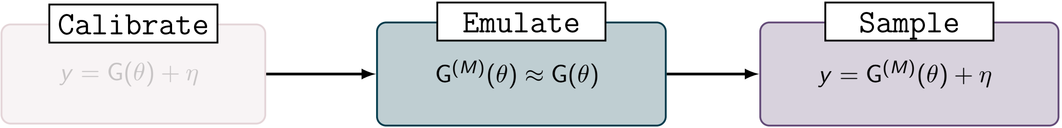
Hurricane Dorian, Sept 07, 2019. Source: tropicaltidbits.com.

Calibrate, Emulate, Sample



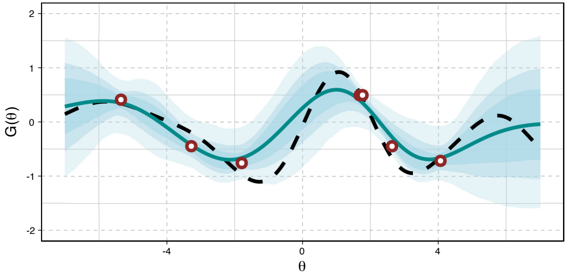
Cleary, Garbuno, Lan, Schneider, Stuart (2019)

arXiv preprint, 1911.+++++



Gaussian Process Accelerated Sampling

- From EKS we generate approximate posterior samples $\left\{\theta^{(i)}, \mathrm{G}(\theta^{(i)})\right\}_{i=1}^{J}$.
- Use parameter-output pairs to train a Gaussian Process (GP) emulator $\mathrm{G}_J(\cdot)$.



- Define $\Phi_J(\theta; y) = \frac{1}{2\gamma^2}|y - \mathrm{G}_J(\theta)|^2$. Evaluation of Φ_J is fast.
- Sample approximate posterior

$$\mathbb{P}_J(\theta|y) \propto \exp\Big(-\Phi_J(\theta; y)\Big) \times \exp\Big(-\frac{1}{2}\langle\theta, \Sigma_0^{-1}\theta\rangle\Big).$$

Conclusions

- New algorithm to **generate approximate posterior samples** for inverse problems;
- Introduced **Kalman-Wasserstein space**;
- New algorithmic framework **Calibrate-Emulate-Sample**.

Open Questions (Theory & Practice):

- Applications of the Calibrate-Emulate-Sample framework.
- How close are dynamics of EKS to KW flow for non-linear G?
- Convergence to equilibrium for non-linear G.
- Properties of the Kalman-Wasserstein space and related functional inequalities.
- General matrix $K(\rho)$: optimal rates of convergence?

References

[1] R. Kalman.
A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, **82**(1960), 35–45.

[2] G. Evensen.
Data Assimilation: The Ensemble Kalman Filter. Springer, 2006.

[3] D.S Oliver, A.C. Reynolds, N. Liu.
Inverse Theory For Petroleum Reservoir Characterization And History Matching. CUP 2008.

[4] S. Reich.
A non-parametric ensemble transform method for Bayesian inference, *SIAM J. Sci. Comput.*, 35(4), A2013-A2024, 2013.

[5] O.G. Ernst, B. Sprungk, H.-J. Starkloff.
Analysis of the ensemble and polynomial chaos Kalman Filters in Bayesian inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, (3) 823-851, 2015.

[6] M.A. Iglesias, K. Law, A.M. Stuart.
Ensemble Kalman method for inverse problems, *Inverse Problems*, 29(4), 2018.

[7] C. Matthews, J. Weare, B. Leimkuhler.
Ensemble preconditioning for Markov chain Monte Carlo simulation. *Statistics and Computing*, Springer, 28(2), 277-290, 2018.

References

[8] W. Li.
Geometry of Probability Simplex via Optimal Transport. *arXiv preprint arXiv:1803.06360*, 2018.

[9] M. Herty, G. Visconti.
Kinetic methods for inverse problems. *Kinetic & Related Models*, 12(5): 1109-1130, 2019.

[10] N.K. Chada, A.M. Stuart, X.T. Tong.
Tikhonov regularizarion within ensemble Kalman inversion. *arXiv preprint arXiv:1901.10382*, 2019.

[11] A. Garbuno-Inigo, F. Hoffmann, W. Li, A.M. Stuart.
Interacting Langevin Diffusions: Gradient Structure And Ensemble Kalman Sampler, *arXiv preprint arXiv:1903.08866*, 2019.

[12] Z. Ding, Q. Li.
Mean-field limit and numerical analysis for Ensemble Kalman Inversion: linear setting. *arXiv preprint arXiv:1908.05575*, 2019.

[13] N. Nüsken, S. Reich.
Note on Interacting Langevin Diffusions: Gradient Structure and Ensemble Kalman Sampler by Garbuno-Inigo, Hoffmann, Li and Stuart. *arXiv preprint arXiv:1908.10890*, 2019.

[14] J.A. Carrillo, U. Vaes
Wasserstein stability estimates for covariance-preconditioned FokkerPlanck equations. *arXiv preprint arXiv:1910.07555*, 2019.