# Structure Preserving Discretizations of Wasserstein Gradient Flows
# Part IV of IV: Discretization on a grid

**Daniel Matthes** (TU Munich)

"A school on gradient flows"
ICTS Bengaluru 7.–11.9.2026

# Where do we need to go today?

**Question**

Can we discretize a $\mathbb{W}$-gradient flow in space replacing $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$\rho^{(h)} = \sum_{x \in \Gamma} \rho_x^{(h)} \, \delta_x \qquad \text{with parameters} \quad \rho_x^{(h)} \in \mathbb{R}_{\geq 0},$$

with a discrete point set $\Gamma \subset \mathbb{R}^d$?

# Where do we need to go today?

**Question**

Can we discretize a $\mathbb{W}$-gradient flow in space replacing $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$\rho^{(h)} = \sum_{x\in\Gamma} \rho_x^{(h)}\,\delta_x \qquad \text{with parameters} \quad \rho_x^{(h)} \in \mathbb{R}_{\geq 0},$$

with a discrete point set $\Gamma \subset \mathbb{R}^d$?

**Answer:** Yes.

# Where do we need to go today?

**Question**

Can we discretize a $\mathbb{W}$-gradient flow in space replacing $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$\rho^{(h)} = \sum_{x \in \Gamma} \rho_x^{(h)} \, \delta_x \qquad \text{with parameters} \quad \rho_x^{(h)} \in \mathbb{R}_{\geq 0},$$

with a discrete point set $\Gamma \subset \mathbb{R}^d$?

**Answer:** Yes. But it's much more complicated than one might think.

# Where do we need to go today?

**Question**

Can we discretize a $\mathbb{W}$-gradient flow in space replacing $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$\rho^{(h)} = \sum_{x \in \Gamma} \rho_x^{(h)} \, \delta_x \qquad \text{with parameters} \quad \rho_x^{(h)} \in \mathbb{R}_{\geq 0},$$

with a discrete point set $\Gamma \subset \mathbb{R}^d$?

**Answer:** Yes. But it's much more complicated than one might think.

*One* difficulty is how to make sense of $\int h(\rho^{(h)}) \, \mathrm{d}x$ etc.
But there is an even more fundamental one.

# Who said fundamental difficulty?

Here's a no-go example by Filippo.

# No-go example (1/3)

Consider the following linear transport equation in 1D:

$$\partial_t \rho = \partial_x (x\, \rho) \quad \text{is the } \mathbb{W}\text{-gradient flow for} \quad \boldsymbol{E}(\rho) = \int_{\mathbb{R}} \frac{x^2}{2} \rho \, \mathrm{d}x.$$

# No-go example (1/3)

Consider the following linear transport equation in 1D:

$$\partial_t \rho = \partial_x (x\, \rho) \quad \text{is the } \mathbb{W}\text{-gradient flow for} \quad \boldsymbol{E}(\rho) = \int_{\mathbb{R}} \frac{x^2}{2} \rho \, \mathrm{d}x.$$

The initial value problem with $\rho_0 = \delta_{\bar{x}}$ has the explicit solution

$$\rho_t = \delta_{x_t} \quad \text{with} \quad x_t = e^{-t} \bar{x}.$$

# No-go example (1/3)

Consider the following linear transport equation in 1D:

$$\partial_t \rho = \partial_x (x\,\rho) \quad \text{is the } \mathbb{W}\text{-gradient flow for} \quad \boldsymbol{E}(\rho) = \int_{\mathbb{R}} \frac{x^2}{2} \rho \, \mathrm{d}x.$$

The initial value problem with $\rho_0 = \delta_{\bar{x}}$ has the explicit solution

$$\rho_t = \delta_{x_t} \quad \text{with} \quad x_t = e^{-t}\bar{x}.$$

The JKO approximation

$$\rho_\tau^n = \operatorname*{argmin}_{\rho} \left( \frac{1}{2\tau} \mathbb{W}(\rho, \rho_\tau^{n-1})^2 + \boldsymbol{E}(\rho) \right)$$

produces the iterates $\rho_\tau^n = \delta_{x_\tau^n}$ with

$$x_\tau^n = \operatorname*{argmin}_{x \in \mathbb{R}} \left( \frac{(x - x_\tau^{n-1})^2}{2\tau} + \frac{x^2}{2} \right),$$

that is $x_\tau^n = (1+\tau)^{-n}\bar{x}$ — which yields a fairly good approximation.

# No-go example (2/3)

Let's try with $\rho^{(h)} = \sum_{x\in\Gamma} \rho_x^{(h)}\, \delta_x$ with $\Gamma \subset \mathbb{R}$ such that $\bar{x} \in \Gamma$ and

$$h := \inf \left\{ |x - x'| \;\middle|\; x, x' \in \Gamma,\, x \neq x' \right\} > 0.$$

# No-go example (2/3)

Let's try with $\rho^{(h)} = \sum_{x\in\Gamma} \rho_x^{(h)}\,\delta_x$ with $\Gamma \subset \mathbb{R}$ such that $\bar{x} \in \Gamma$ and

$$h := \inf\left\{|x - x'| \;\middle|\; x, x' \in \Gamma,\, x \neq x'\right\} > 0.$$

The first JKO step functional

$$\frac{1}{2\tau}\mathbb{W}(\rho^{(h)}, \delta_{\bar{x}})^2 + \boldsymbol{E}(\rho^{(h)}) = \sum_{x\in\Gamma}\left(\frac{(x-\bar{x})^2}{2\tau} + \frac{x^2}{2}\right)\rho_x^{(h)}$$

is minimized by $\delta_{x^*}$ where

$$x^* = \operatorname*{argmin}_{x\in\Gamma}\left(\frac{(x-\bar{x})^2}{2\tau} + \frac{x^2}{2}\right).$$

# No-go example (2/3)

Let's try with $\rho^{(h)} = \sum_{x\in\Gamma} \rho_x^{(h)}\,\delta_x$ with $\Gamma \subset \mathbb{R}$ such that $\bar{x} \in \Gamma$ and

$$h := \inf\left\{|x - x'| \;\middle|\; x, x' \in \Gamma,\, x \neq x'\right\} > 0.$$

The first JKO step functional

$$\frac{1}{2\tau}\mathbb{W}(\rho^{(h)}, \delta_{\bar{x}})^2 + \boldsymbol{E}(\rho^{(h)}) = \sum_{x\in\Gamma}\left(\frac{(x-\bar{x})^2}{2\tau} + \frac{x^2}{2}\right)\rho_x^{(h)}$$

is minimized by $\delta_{x^*}$ where

$$x^* = \operatorname*{argmin}_{x\in\Gamma}\left(\frac{(x-\bar{x})^2}{2\tau} + \frac{x^2}{2}\right).$$

If $\tau > 0$ is so small that $2\tau\bar{x} < (1+\tau)h$, then $x^* = \bar{x}$.
The evolution is frustrated.

# No-go example (3/3)

Still, good news:

**Theorem ([Hraivoronska&Santambrogio'25])**

*Under the inverse CFL condition* $h/\tau \to 0$*,*
*this discretization converges as* $h, \tau \to 0$*.*
*It even works with (very) nonlinear diffusion.*

# No-go example (3/3)

Still, good news:

**Theorem** ( `[Hraivoronska&Santambrogio'25]` )

*Under the inverse CFL condition* $h/\tau \to 0$*,*
*this discretization converges as* $h, \tau \to 0$*.*
*It even works with (very) nonlinear diffusion.*

**In any case:** The scheme is no good for semi-discretization in space only.

## So that's the end for today? Goodbye!

Fortunately, there is a workaround by Alex (and several others).

# Explained in two-point space (1/3)

Consider the problem on a two-point space $\Gamma = \{0, h\}$.
Parametrize $\rho^{(h)} \in \mathcal{P}_2(\Gamma)$ by $r \in [0,1]$ via

$$\rho^{(h)} = \mu[r] = r\delta_h + (1-r)\delta_0, \quad \text{thus} \quad \boldsymbol{E}(\mu[r]) = \frac{h^2}{2} r.$$

Then

$$\mathbb{W}(\mu[\hat{r}], \mu[\check{r}])^2 = h^2 |\hat{r} - \check{r}|,$$

which is

- quadratic in the fixed distance $h$,
- one-homogeneous in the parameters $\hat{r}$ and $\check{r}$.

# Explained in two-point space (1/3)

Consider the problem on a two-point space $\Gamma = \{0, h\}$.
Parametrize $\rho^{(h)} \in \mathcal{P}_2(\Gamma)$ by $r \in [0,1]$ via

$$\rho^{(h)} = \mu[r] = r\delta_h + (1-r)\delta_0, \quad \text{thus} \quad \boldsymbol{E}(\mu[r]) = \frac{h^2}{2} r.$$

Then

$$\mathbb{W}(\mu[\hat{r}], \mu[\check{r}])^2 = h^2 |\hat{r} - \check{r}|,$$

which is

- quadratic in the fixed distance $h$,
- one-homogeneous in the parameters $\hat{r}$ and $\check{r}$.

The respective JKO step looks strange:

$$\frac{\mathbb{W}(\mu[r], \mu[r_\tau^{n-1}])^2}{2\tau} + \boldsymbol{E}(\mu[r]) = \frac{h^2}{2\tau}\left|r_\tau^{n-1} - r\right| + \frac{h^2}{2} r \to \min .$$

Its minimizer is either $r = 0$ (for $\tau > 1$) or $r = r_\tau^{n-1}$ (for $\tau < 1$).

# Explained in two-point space (2/3)

Replace $\mathbb{W}^2$ by another "inertia" $\mathbf{d}^2$ that is nonlinear w.r.t. $\hat{r}$, $\check{r}$.

**Example:** Try with relative entropy

$$\mathbf{d}(\mu[\hat{r}], \mu[\check{r}])^2 = h^2 \hat{r}\, \psi\left(\frac{\check{r}}{\hat{r}}\right), \quad \psi(z) = z \log z - z + 1.$$

The respective JKO step reads

$$\frac{\mathbf{d}(\mu[r], \mu[r_\tau^{n-1}])^2}{2\tau} + \boldsymbol{E}(\mu[r]) \cong \frac{h^2}{2\tau}\left(r - r_\tau^{n-1} \log r\right) + \frac{h^2}{2} r \to \min,$$

which leads to the iteration

$$\frac{r_\tau^n - r_\tau^{n-1}}{\tau} = -r_\tau^n$$

and thus $r_\tau^n = (1+\tau)^{-n} r_\tau^0$,

# Explained in two-point space (2/3)

Replace $\mathbb{W}^2$ by another "inertia" $\mathbf{d}^2$ that is nonlinear w.r.t. $\hat{r}$, $\check{r}$.

**Example:** Try with relative entropy

$$\mathbf{d}(\mu[\hat{r}], \mu[\check{r}])^2 = h^2 \hat{r}\, \psi\left(\frac{\check{r}}{\hat{r}}\right), \quad \psi(z) = z \log z - z + 1.$$

The respective JKO step reads

$$\frac{\mathbf{d}(\mu[r], \mu[r_\tau^{n-1}])^2}{2\tau} + \boldsymbol{E}(\mu[r]) \cong \frac{h^2}{2\tau}\big(r - r_\tau^{n-1} \log r\big) + \frac{h^2}{2} r \to \min,$$

which leads to the iteration

$$\frac{r_\tau^n - r_\tau^{n-1}}{\tau} = -r_\tau^n$$

and thus $r_\tau^n = (1+\tau)^{-n} r_\tau^0$, which yields a fairly good approximation, and has limit $r(t) = e^{-t} r(0)$ for $\tau \to 0$.

# Explained in two-point space (3/3)

**Lemma**

*The JKO iterates $(\rho_\tau^n)$ satisfy*

$$\boldsymbol{E}(\rho_\tau^n) \leq (1+\tau)^{-1}\boldsymbol{E}(\rho_\tau^{n-1}),$$
$$\mathbf{d}(\rho_\tau^n, \rho_\infty)^2 \leq (1+\tau)^{-1}\mathbf{d}(\rho_\tau^{n-1}, \rho_\infty)^2.$$

*Moreover, the semi-discrete evolution is contractive in the sense that for two JKO approximations $(\hat{r}_\tau^n)$ and $(\check{r}_\tau^n)$:*

$$\mathbf{d}\big(\hat{r}_\tau^n, \check{r}_\tau^n\big)^2 \leq (1+\tau)^{-1}\mathbf{d}\big(\hat{r}_\tau^{n-1}, \check{r}_\tau^{n-1}\big)^2.$$

**Proof:** Homogeneity $\mathbf{d}(ca, cb)^2 = c\mathbf{d}(a,b)^2$ for all $a, b \in [0,1]$ and $c > 0$.

# Explained in two-point space (3/3)

**Lemma**

*The JKO iterates $(\rho_\tau^n)$ satisfy*

$$\boldsymbol{E}(\rho_\tau^n) \leq (1+\tau)^{-1}\boldsymbol{E}(\rho_\tau^{n-1}),$$
$$\mathbf{d}(\rho_\tau^n, \rho_\infty)^2 \leq (1+\tau)^{-1}\mathbf{d}(\rho_\tau^{n-1}, \rho_\infty)^2.$$

*Moreover, the semi-discrete evolution is contractive in the sense that for two JKO approximations $(\hat{r}_\tau^n)$ and $(\check{r}_\tau^n)$:*

$$\mathbf{d}\left(\hat{r}_\tau^n, \check{r}_\tau^n\right)^2 \leq (1+\tau)^{-1}\mathbf{d}\left(\hat{r}_\tau^{n-1}, \check{r}_\tau^{n-1}\right)^2.$$

**Proof:** Homogeneity $\mathbf{d}(ca, cb)^2 = c\mathbf{d}(a,b)^2$ for all $a, b \in [0,1]$ and $c > 0$.

**New angle of attack:**

discrete energy $+$ discrete metric $\Longrightarrow$ contractive discrete PDE,

is replaced by

discrete energy $+$ discrete PDE $\Longrightarrow$ discrete metric with contractivity.

# ODE analogy on contractivity (1/3)

**General:** on manifold $\mathcal{M}$ with metric tensor $\langle \cdot, \cdot \rangle_x$ at each $x \in \mathcal{M}$, define the cotangent-to-tangent map $\mathbb{K}_x : \mathrm{T}_x^\star\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ by

$$\langle v, \mathbb{K}_x \xi \rangle_x = \xi[v] \quad \text{for all } v \in \mathrm{T}_x\mathcal{M},\ \xi \in \mathrm{T}_x^\star\mathcal{M}.$$

The gradient flow of $\boldsymbol{e} : \mathcal{M} \to \mathbb{R}$ is

$$\dot{x} \cong -\mathbb{K}_x \mathrm{D}_x \boldsymbol{e}.$$

# ODE analogy on contractivity (1/3)

**General:** on manifold $\mathcal{M}$ with metric tensor $\langle\cdot,\cdot\rangle_x$ at each $x \in \mathcal{M}$, define the cotangent-to-tangent map $\mathbb{K}_x : \mathrm{T}_x^\star\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ by

$$\langle v, \mathbb{K}_x \xi\rangle_x = \xi[v] \quad \text{for all } v \in \mathrm{T}_x\mathcal{M},\ \xi \in \mathrm{T}_x^\star\mathcal{M}.$$

The gradient flow of $\boldsymbol{e} : \mathcal{M} \to \mathbb{R}$ is

$$\dot{x} \cong -\mathbb{K}_x \mathrm{D}_x \boldsymbol{e}.$$

**Specifically:** on $\mathcal{M} = \mathbb{R}^n$, define metric tensor by

$$\langle v, w\rangle_x = v^T \mathbb{G}_x w$$

with an $x$-dependent symmetric positive definite matrix $x \mapsto \mathbb{G}_x \in \mathbb{R}^{n\times n}$. Then

$$\mathbb{K}_x \cong \mathbb{G}_x^{-1}.$$

# ODE analogy on contractivity (2/3)

Consider a family $(\gamma^t)_{t\geq 0}$ of curves $(\gamma^t_s)_{s\in[0,1]}$
that "follow a vector field" $F : \mathcal{M} \to \mathrm{T}\mathcal{M}$ e.g. $F = -\mathbb{K}\mathrm{D}\boldsymbol{e}$,

$$\partial_t \gamma^t_s = F(\gamma^t_s).$$

# ODE analogy on contractivity (2/3)

Consider a family $(\gamma^t)_{t\geq 0}$ of curves $(\gamma^t_s)_{s\in[0,1]}$
that "follow a vector field" $F : \mathcal{M} \to \mathrm{T}\mathcal{M}$ e.g. $F = -\mathbb{K}\mathrm{D}\boldsymbol{e}$,

$$\partial_t \gamma^t_s = F(\gamma^t_s).$$

The kinetic energy of $(\gamma^t_s)_{s\in[0,1]}$

$$a(\gamma^t) = \int_0^1 (\partial_s \gamma)^T \mathbb{G}_\gamma (\partial_s \gamma)\,\mathrm{d}s$$

# ODE analogy on contractivity (2/3)

Consider a family $(\gamma^t)_{t\geq 0}$ of curves $(\gamma^t_s)_{s\in[0,1]}$
that "follow a vector field" $F : \mathcal{M} \to \mathrm{T}\mathcal{M}$ e.g. $F = -\mathbb{K}\mathrm{D}\boldsymbol{e}$,

$$\partial_t \gamma^t_s = F(\gamma^t_s).$$

The kinetic energy of $(\gamma^t_s)_{s\in[0,1]}$ evolves as follows:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} a(\gamma^t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 (\partial_s\gamma)^T \mathbb{G}_\gamma (\partial_s \gamma)\,\mathrm{d}s \\ &= -2\int_0^1 (\partial_s\gamma)^T \mathbb{G}_\gamma M_\gamma \mathbb{G}_\gamma (\partial_s\gamma)\,\mathrm{d}s,\end{aligned}$$

with

$$M_x = -\mathrm{D}_x F[\mathbb{K}_x(\cdot)] + \frac{1}{2}\mathrm{D}_x \mathbb{K}[F].$$

# ODE analogy on contractivity (2/3)

Consider a family $(\gamma^t)_{t\geq 0}$ of curves $(\gamma_s^t)_{s\in[0,1]}$
that "follow a vector field" $F : \mathcal{M} \to \mathrm{T}\mathcal{M}$ e.g. $F = -\mathbb{K}\mathrm{D}\boldsymbol{e}$,

$$\partial_t \gamma_s^t = F(\gamma_s^t).$$

The kinetic energy of $(\gamma_s^t)_{s\in[0,1]}$ evolves as follows:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} a(\gamma^t) &= \frac{\mathrm{d}}{\mathrm{d}t} \int_0^1 (\partial_s \gamma)^T \mathbb{G}_\gamma (\partial_s \gamma)\,\mathrm{d}s \\ &= -2 \int_0^1 (\partial_s \gamma)^T \mathbb{G}_\gamma M_\gamma \mathbb{G}_\gamma (\partial_s \gamma)\,\mathrm{d}s,\end{aligned}$$

with

$$M_x = -\mathrm{D}_x F[\mathbb{K}_x(\cdot)] + \frac{1}{2}\mathrm{D}_x \mathbb{K}[F].$$

**Lemma**

*If $M_x \geq \lambda \mathbb{K}_x$ with some $\lambda$ for all $x \in \mathcal{M}$, then $F$ is $\lambda$-contractive:*

$$a(\gamma^t) \leq e^{-2\lambda t} a(\gamma^0).$$

# ODE analogy on contractivity (3/3)

$$M_x = -\mathrm{D}_x F[\mathbb{K}_x(\cdot)] + \frac{1}{2}\mathrm{D}_x\mathbb{K}[F] \geq \lambda\mathbb{K}_x$$

looks ugly, but is just an inequality between matrices.

# ODE analogy on contractivity (3/3)

$$M_x = -\mathrm{D}_x F[\mathbb{K}_x(\cdot)] + \frac{1}{2}\mathrm{D}_x\mathbb{K}[F] \geq \lambda\mathbb{K}_x$$

looks ugly, but is just an inequality between matrices.

**Fundamental idea behind Eulerian calculus**

If $F = -\mathbb{K}\mathrm{D}\boldsymbol{e}$ is a gradient vector field,
then $M \geq \lambda\mathbb{K}$ implies that $\boldsymbol{e}$ is $\lambda$-geodesically convex
(and subsequently exponential control of energy, dissipation, speed, ...).

**Proof:** With some painful algebra,

$$-\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{D}_x\boldsymbol{e}^T\mathbb{K}_x\mathrm{D}_x\boldsymbol{e}\big) = 2\mathrm{D}_x\boldsymbol{e}^T M_x\mathrm{D}_x\boldsymbol{e} \geq 2\lambda\big(\mathrm{D}_x\boldsymbol{e}^T\mathbb{K}_x\mathrm{D}_x\boldsymbol{e}\big).$$

From here, everything else follows by Taylor expansion and integration along trajectories.

## Are we still doing anything discrete today?

This is applicable to Fokker-Planck.

# Factoring the heat equation (1/2)

**Our plan:**

1. You're good at dualities? Pretend for the moment you're not!
2. Forget how $\mathbb{W}$-gradient flows had been defined.
3. Try to guess $\mathbb{K}$ from the linear Fokker-Planck equation.
4. Show that the flow is $1$-contractive.
5. Then repeat, but discretely.

# Factoring the heat equation (2/2)

To write the heat equation on $\mathbb{R}$,

$$\partial_t \rho = \partial_{xx} \rho$$

in the form $\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}$ with the entropy

$$\boldsymbol{H}(\rho) = \int \rho \log \rho \, \mathrm{d}x \qquad .$$

# Factoring the heat equation (2/2)

To write the heat equation on $\mathbb{R}$,

$$\partial_t \rho = \partial_{xx} \rho = \partial_x \big( \rho \, \partial_x \log \rho \big)$$

in the form $\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}$ with the entropy

$$\boldsymbol{H}(\rho) = \int \rho \log \rho \, \mathrm{d}x \quad \text{with} \quad \frac{\delta \boldsymbol{H}}{\delta \rho} = 1 + \log \rho.$$

# Factoring the heat equation (2/2)

To write the heat equation on $\mathbb{R}$,

$$\partial_t \rho = \partial_{xx} \rho = \partial_x \big( \rho \, \partial_x \log \rho \big)$$

in the form $\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}$ with the entropy

$$\boldsymbol{H}(\rho) = \int \rho \log \rho \, \mathrm{d}x \quad \text{with} \quad \frac{\delta \boldsymbol{H}}{\delta \rho} = 1 + \log \rho.$$

we choose the (symmetric, positive, ...) cotangent-to-tangent map

$$\mathbb{K}_\rho \xi = -\partial_x \big( \rho \, \partial_x \xi \big).$$

# Factoring the heat equation (2/2)

To write the heat equation on $\mathbb{R}$,

$$\partial_t \rho = \partial_{xx}\rho = \partial_x\big(\rho\,\partial_x \log\rho\big)$$

in the form $\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}$ with the entropy

$$\boldsymbol{H}(\rho) = \int \rho \log\rho \,\mathrm{d}x \quad \text{with} \quad \frac{\delta \boldsymbol{H}}{\delta\rho} = 1 + \log\rho.$$

we choose the (symmetric, positive, ...) cotangent-to-tangent map

$$\mathbb{K}_\rho \xi = -\partial_x\big(\rho\,\partial_x \xi\big).$$

**Lemma**

*The flow is 0-contractive, i.e.,*

$$0 \le \int_{\mathbb{R}} \xi\, M_\rho \xi \,\mathrm{d}x = \int_{\mathbb{R}} \xi \Big( -\partial_{xx}\big(\mathbb{K}_\rho \xi\big) + \frac{1}{2}\mathrm{D}_\rho \mathbb{K}[\partial_{xx}\rho]\xi \Big)\,\mathrm{d}x.$$

# Some discretion, please!

- $h\mathbb{Z} = \{\dots, -h, 0, h, 2h, \dots\}$ is the one-dimensional uniform grid with mesh width $h > 0$
- $h\mathbb{Z}^* = \{\dots, -\frac{1}{2}h, \frac{1}{2}h, \frac{3}{2}h, \dots\}$ is the dual grid
- for a function $u : h\mathbb{Z} \to \mathbb{R}$, define its gradient $\nabla^{(h)} u : h\mathbb{Z}^* \to \mathbb{R}$ by

$$\nabla_y^{(h)} u = \frac{u_{y+h/2} - u_{y-h/2}}{h}.$$

- for a flux $J : h\mathbb{Z}^* \to \mathbb{R}$, define its divergence ${\nabla^{(h)}}^* \omega : h\mathbb{Z} \to \mathbb{R}$ by

$$\nabla_x^{(h)^*} \omega = \frac{\omega_{x+h/2} - \omega_{x-h/2}}{h}.$$

- the Laplace operator $\Delta^{(h)}$ is

$$\Delta_x^{(h)} u = \nabla_x^{(h)^*}(\nabla_u^{(h)}) = \frac{u_{x+h} - 2u_x + u_{x-h}}{h^2}.$$

# A complicated writing of the discrete heat equation (1/2)

The simplest spatial discretization of the heat equation on $h\mathbb{Z}$ is

$$\dot{\rho}^{(h)} = \Delta^{(h)} \rho^{(h)}.$$

# A complicated writing of the discrete heat equation (1/2)

The simplest spatial discretization of the heat equation on $h\mathbb{Z}$ is

$$\dot\rho^{(h)} = \Delta^{(h)}\rho^{(h)}.$$

We want to write $\dot\rho^{(h)} = -\mathbb{K}_\rho^{(h)}\mathrm{D}_\rho\mathbf{H}^{(h)}$ with the discretized entropy

$$\mathbf{H}^{(h)}(\rho^{(h)}) = h\sum_{x\in h\mathbb{Z}} \rho_x^{(h)}\log\rho_x^{(h)} \qquad .$$

# A complicated writing of the discrete heat equation (1/2)

The simplest spatial discretization of the heat equation on $h\mathbb{Z}$ is

$$\dot\rho^{(h)} = \Delta^{(h)}\rho^{(h)}.$$

We want to write $\dot\rho^{(h)} = -\mathbb{K}_\rho^{(h)} \mathrm{D}_\rho \mathbf{H}^{(h)}$ with the discretized entropy

$$\mathbf{H}^{(h)}(\rho^{(h)}) = h \sum_{x\in h\mathbb{Z}} \rho_x^{(h)} \log \rho_x^{(h)} \quad \text{with} \quad \mathrm{D}_\rho \boldsymbol{H}^{(h)} = 1 + \log \rho^{(h)}.$$

# A complicated writing of the discrete heat equation (1/2)

The simplest spatial discretization of the heat equation on $h\mathbb{Z}$ is

$$\dot{\rho}^{(h)} = \Delta^{(h)} \rho^{(h)}.$$

We want to write $\dot{\rho}^{(h)} = -\mathbb{K}_{\rho}^{(h)} \mathrm{D}_{\rho} \mathbf{H}^{(h)}$ with the discretized entropy

$$\mathbf{H}^{(h)}(\rho^{(h)}) = h \sum_{x \in h\mathbb{Z}} \rho_x^{(h)} \log \rho_x^{(h)} \quad \text{with} \quad \mathrm{D}_{\rho} \boldsymbol{H}^{(h)} = 1 + \log \rho^{(h)}.$$

Write

$$\Delta^{(h)} \rho^{(h)} = {\nabla^{(h)}}^{*} \nabla^{(h)} \rho^{(h)} = {\nabla^{(h)}}^{*} \left( \frac{\nabla^{(h)} \rho}{\nabla^{(h)} \log \rho^{(h)}} \nabla^{(h)} \log \rho^{(h)} \right).$$

# A complicated writing of the discrete heat equation (1/2)

The simplest spatial discretization of the heat equation on $h\mathbb{Z}$ is

$$\dot{\rho}^{(h)} = \Delta^{(h)}\rho^{(h)}.$$

We want to write $\dot{\rho}^{(h)} = -\mathbb{K}_\rho^{(h)}\mathrm{D}_\rho\mathbf{H}^{(h)}$ with the discretized entropy

$$\mathbf{H}^{(h)}(\rho^{(h)}) = h\sum_{x\in h\mathbb{Z}}\rho_x^{(h)}\log\rho_x^{(h)} \quad \text{with} \quad \mathrm{D}_\rho\boldsymbol{H}^{(h)} = 1+\log\rho^{(h)}.$$

Write

$$\Delta^{(h)}\rho^{(h)} = \nabla^{(h)*}\nabla^{(h)}\rho^{(h)} = \nabla^{(h)*}\left(\frac{\nabla^{(h)}\rho}{\nabla^{(h)}\log\rho^{(h)}}\nabla^{(h)}\log\rho^{(h)}\right).$$

Thus we have the desired form for the (symmetric, positive, ...) operator

$$\mathbb{K}_\rho^{(h)}\xi = -\nabla^{(h)*}\big(\Lambda(\rho^{(h)})\nabla^{(h)}\xi\big) \quad \text{w/} \quad \Lambda_y(\rho^{(h)}) = \frac{\rho_{y+h/2}^{(h)} - \rho_{y-h/2}^{(h)}}{\log\rho_{y+h/2}^{(h)} - \log\rho_{y-h/2}^{(h)}}.$$

# A complicated writing of the discrete heat equation (2/2)

Remarkably, this weird choice is a good one.

**Theorem** ([Maas'11,Mielke'13])

*The discrete heat equation is $0$-contractive, i.e., $M \geq 0$.*

# A complicated writing of the discrete heat equation (2/2)

Remarkably, this weird choice is a good one.

**Theorem** ([Maas'11,Mielke'13])

*The discrete heat equation is 0-contractive, i.e., $M \geq 0$.*

**Proof:** mainly uses two ingredients:

- $M_\rho$ is a tridiagonal matrix with nice algebraic properties, and
- $\Lambda$ has a lot of nice algebraic properties.

# Now what for linear Fokker-Planck? (1/3)

**Return to continuous.** For $V$ such that $V'' \geq \lambda > 0$ and $\rho_\infty := e^{-V} \in \mathcal{P}_2(\mathbb{R})$, consider the relative entropy

$$\boldsymbol{H}_V(\rho) = \int \rho \log \rho \, \mathrm{d}x + \int \rho V \, \mathrm{d}x \qquad .$$

# Now what for linear Fokker-Planck? (1/3)

**Return to continuous.** For $V$ such that $V'' \geq \lambda > 0$ and $\rho_\infty := e^{-V} \in \mathcal{P}_2(\mathbb{R})$, consider the relative entropy

$$\boldsymbol{H}_V(\rho) = \int \rho \log \rho \, \mathrm{d}x + \int \rho V \, \mathrm{d}x = \int \rho \log \frac{\rho}{\rho_\infty} \, \mathrm{d}x.$$

# Now what for linear Fokker-Planck? (1/3)

**Return to continuous.** For $V$ such that $V'' \geq \lambda > 0$ and $\rho_\infty := e^{-V} \in \mathcal{P}_2(\mathbb{R})$, consider the relative entropy

$$\boldsymbol{H}_V(\rho) = \int \rho \log \rho \, \mathrm{d}x + \int \rho V \, \mathrm{d}x = \int \rho \log \frac{\rho}{\rho_\infty} \, \mathrm{d}x.$$

With the same $\mathbb{K}$ as before, we obtain the linear Fokker-Planck equation

$$\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}_V \qquad .$$

# Now what for linear Fokker-Planck? (1/3)

**Return to continuous.** For $V$ such that $V'' \geq \lambda > 0$ and $\rho_\infty := e^{-V} \in \mathcal{P}_2(\mathbb{R})$, consider the relative entropy

$$\boldsymbol{H}_V(\rho) = \int \rho \log \rho \,\mathrm{d}x + \int \rho V \,\mathrm{d}x = \int \rho \log \frac{\rho}{\rho_\infty} \,\mathrm{d}x.$$

With the same $\mathbb{K}$ as before, we obtain the linear Fokker-Planck equation

$$\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}_V = \partial_x \left( \rho_\infty \frac{\rho}{\rho_\infty} \partial_x \log \frac{\rho}{\rho_\infty} \right) = \partial_x \left( \rho_\infty \, \partial_x \left[ \frac{\rho}{\rho_\infty} \right] \right).$$

# Now what for linear Fokker-Planck? (1/3)

**Return to continuous.** For $V$ such that $V'' \geq \lambda > 0$ and $\rho_\infty := e^{-V} \in \mathcal{P}_2(\mathbb{R})$, consider the relative entropy

$$\boldsymbol{H}_V(\rho) = \int \rho \log \rho \,\mathrm{d}x + \int \rho V \,\mathrm{d}x = \int \rho \log \frac{\rho}{\rho_\infty} \,\mathrm{d}x.$$

With the same $\mathbb{K}$ as before, we obtain the linear Fokker-Planck equation

$$\partial_t \rho = -\mathbb{K}_\rho \mathrm{D}_\rho \boldsymbol{H}_V = \partial_x \left( \rho_\infty \frac{\rho}{\rho_\infty} \partial_x \log \frac{\rho}{\rho_\infty} \right) = \partial_x \left( \rho_\infty \, \partial_x \left[ \frac{\rho}{\rho_\infty} \right] \right).$$

**Lemma**

*The flow is $\lambda$-contractive.*

# Now what for linear Fokker-Planck? (2/3)

Choose a discrete steady state $\rho_\infty^{(h)} \in \mathcal{P}_2(h\mathbb{Z})$ by

$$\rho_{\infty,x}^{(h)} = \fint_{x-h/2}^{x+h/2} \rho_\infty(x')\,\mathrm{d}x'$$

and accordingly

$$\mathbf{H}_V^{(h)}(\rho^{(h)}) = h\sum_x \rho^{(h)} \log \frac{\rho^{(h)}}{\rho_\infty^{(h)}} \quad \text{with} \quad \mathrm{D}_\rho \mathbf{H}_V^{(h)} \cong 1 + \log \frac{\rho^{(h)}}{\rho_\infty^{(h)}}.$$

Further, introduce a balanced inhomogeneity:

$$\mathbb{K}_\rho^{(h)}\xi = \nabla^{(h)^*}\left[\omega\,\Lambda\left(\frac{\rho^{(h)}}{\rho_\infty^{(h)}}\right)\nabla^{(h)}\xi\right] \quad \text{w/} \quad \omega_y = \sqrt{\rho_{\infty,y+h/2}^{(h)}\rho_{\infty,y-h/2}^{(h)}}.$$

Then the gradient flow is:

$$\dot\rho_x^{(h)} = \nabla_x^{(h)^*}\left(\omega\,\nabla^{(h)}\left[\frac{\rho}{\rho_\infty}\right]\right).$$

# Now what for linear Fokker-Planck? (3/3)

**Theorem ([Mielke'13])**

*If $V'' \geq \lambda > 0$, then there is some $\lambda_h \xrightarrow{h \searrow 0} \lambda$ s.t.*

- $\rho_\infty^{(h)}$ *is the unique minimizer of* $\mathbf{H}_V^{(h)}$
- *the entropy decays exponentially:*

$$\mathbf{H}_V^{(h)}(\rho_t^{(h)}) - \mathbf{H}_V^{(h)}(\rho_\infty^{(h)}) \leq e^{-2\lambda_h t}\big(\mathbf{H}_V^{(h)}(\rho_0^{(h)}) - \mathbf{H}_V^{(h)}(\rho_\infty^{(h)})\big),$$

  *and so do velocity, dissipation, and distance to equilibrium.*
- *the flow is $\lambda_h$-contractive:*

$$\mathbb{W}_V^{(h)}\big(\rho^{(\hat{h})}{}_t, \rho^{(\check{h})}{}_t\big)^2 \leq e^{-2\lambda_h t}\mathbb{W}_V^{(h)}\big(\rho^{(\hat{h})}{}_0, \rho^{(\check{h})}{}_0\big)^2.$$

# What next? Same for NONlinear Fokker-Planck?

Try same idea for quadratic porous medium equation

$$\partial_t \rho = \Delta \rho^2,$$

which is the $\mathbb{W}$-gradient flow of

$$\boldsymbol{H}(\rho) = \int \rho^2 \, \mathrm{d}x.$$

# What next? Same for NONlinear Fokker-Planck?

Try same idea for quadratic porous medium equation

$$\partial_t \rho = \Delta \rho^2,$$

which is the $\mathbb{W}$-gradient flow of

$$\boldsymbol{H}(\rho) = \int \rho^2 \, \mathrm{d}x.$$

Discretizing this as

$$\dot{\rho}^{(h)} = \Delta\big(\rho^{(h)}\big)^2 \quad \text{and} \quad \mathbf{H}^{(h)}(\rho^{(h)}) = h \sum_x \left(\rho^{(h)}\right)^2$$

and consequently

$$\mathbb{K}_x^{(h)}[\xi] = \nabla^{(h)^*}\left(A(\rho^{(h)})\nabla^{(h)}\xi\right) \quad \text{w/} \quad A_y(\rho^{(h)}) = \frac{\rho_{y+h/2}^{(h)} + \rho_{y-h/2}^{(h)}}{2}.$$

# What next? Same for NONlinear Fokker-Planck?

Try same idea for quadratic porous medium equation

$$\partial_t \rho = \Delta \rho^2,$$

which is the $\mathbb{W}$-gradient flow of

$$\boldsymbol{H}(\rho) = \int \rho^2 \, \mathrm{d}x.$$

Discretizing this as

$$\dot{\rho}^{(h)} = \Delta \big(\rho^{(h)}\big)^2 \quad \text{and} \quad \mathbf{H}^{(h)}(\rho^{(h)}) = h \sum_x \left(\rho^{(h)}\right)^2$$

and consequently

$$\mathbb{K}_x^{(h)}[\xi] = \nabla^{(h)^*} \left( A(\rho^{(h)}) \nabla^{(h)} \xi \right) \quad \text{w/} \quad A_y(\rho^{(h)}) = \frac{\rho^{(h)}_{y+h/2} + \rho^{(h)}_{y-h/2}}{2}.$$

**Theorem ( [Erbar&Maas'12] )**

*This discretization has no convexity at all.*

# And . . . what about DLSS?

Residual convexity survives!

# Yes, you have seen this before

**Recall:** the rescaled DLSS equation

$$\partial_t \rho = -2\partial_x \left( \rho \, \partial_x \frac{\partial_{xx}\sqrt{\rho}}{\sqrt{\rho}} \right) + \partial_x (x \, \rho)$$

is the $\mathbb{W}$-gradient flow of the relative Fisher information

$$\boldsymbol{F}'(\rho) = 2 \int_{\mathbb{R}} \left(\partial_x \sqrt{\rho}\right)^2 \mathrm{d}x + \int_{\mathbb{R}} \frac{x^2}{2} \rho \, \mathrm{d}x,$$

which is **totally not** displacement convex, but still,

**Theorem**

$$\boldsymbol{H}'(\rho_t) - \boldsymbol{H}'(\rho_\infty) \leq e^{-2t} \big( \boldsymbol{H}'(\rho_0) - \boldsymbol{H}'(\rho_\infty) \big),$$
$$\boldsymbol{F}'(\rho_t) - \boldsymbol{F}'(\rho_\infty) \leq e^{-2t} \big( \boldsymbol{F}'(\rho_0) - \boldsymbol{F}'(\rho_\infty) \big).$$

**Key:** $\boldsymbol{F}'(\rho) + 1 = \frac{1}{2} \left| \partial \boldsymbol{H}'(\rho) \right|^2$ with 1-displacement convex entropy

$$\boldsymbol{H}'(\rho) = \int_{\mathbb{R}} \rho \log \rho \, \mathrm{d}x + \int_{\mathbb{R}} \frac{x^2}{2} \rho \, \mathrm{d}x.$$

# Recycling is a great thing

With the Alex-style discretizations

$$\dot\rho^{(h)} = \nabla^{(h)*}\left(\omega\,\nabla^{(h)}\left[\frac{\rho^{(h)}}{\rho^{(h)}_\infty}\right]\right) \quad \& \quad \mathbf{H}^{(h)'} = h\sum_{x\in\delta\mathbb{Z}}\rho^{(h)}_x\log\left(\frac{\rho^{(h)}_x}{\rho^{(h)}_{\infty,x}}\right),$$

introduce a discretized relative Fisher information by

$$\mathbf{F}^{(h)'} = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{H}^{(h)'} = h\sum_{y\in(h\mathbb{Z})^*}\omega_y\nabla^{(h)}_y\log\left(\frac{\rho^{(h)}}{\rho^{(h)}_\infty}\right)\nabla^{(h)}_y\left(\frac{\rho^{(h)}}{\rho^{(h)}_\infty}\right).$$

The gradient flow of $\mathbf{F}^{(h)'}$ with respect to $\mathbb{K}$ looks horrible, but:

**Theorem ( [Maas+M'16] )**

*There is some* $\lambda_h \xrightarrow{h\searrow 0} 1$ *s.t.*

$$\mathbf{H}^{(h)'}(\rho^{(h)}_t) - \mathbf{H}^{(h)'}(\rho^{(h)}_\infty) \le e^{-2\lambda_h t}\big(\mathbf{H}^{(h)'}(\rho_0) - \mathbf{H}^{(h)'}(\rho^{(h)}_\infty)\big),$$
$$\mathbf{F}^{(h)'}(\rho^{(h)}_t) - \mathbf{F}^{(h)'}(\rho^{(h)}_\infty) \le e^{-2\lambda_h t}\big(\mathbf{F}^{(h)'}(\rho_0) - \mathbf{F}^{(h)'}(\rho^{(h)}_\infty)\big).$$

# That's all for this course

Thank you for following!