Technische Universität Berlin

Telegrapher's Generative Model via Kac Flows

Gabriele Steidl

TU Berlin

Joint work with

Richard Duong, Jannis Chemseddine and Peter Friz

Berlin Mathematics Research Center MATH+

Funded under Germany's Excellence Strategy by DFG Deutsche Forschungsgemeinschaft

Outline

1. Basic Idea of Generative Flows
2. Kac Model and Telegrapher's Equation in 1D
3. Multidimensional Generalizations of the Model
4. Velocity Field of the Multidimensional Kac Flow
5. Numerical Examples

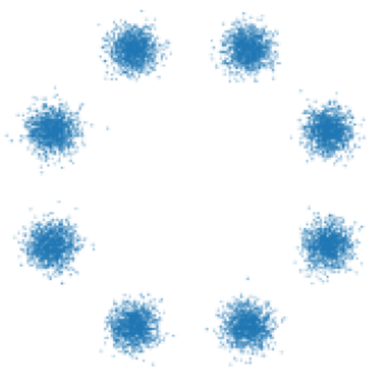
1. Basic Idea of Generative Flows

Aim: Sampling from a probability distribution μ_X (target distribution) given:

- some samples (oft many = data hungry)
- probability density function $p_X(x) = \exp(-\psi(x))/C$ (Boltzmann density) with or without normalizing factor C

It is only easy to sample from

- 1d distributions (cdf)
- $d \gg 1$ uniform or Gaussian "friendly" distributions μ_Z (latent distribution)



$d = 2$

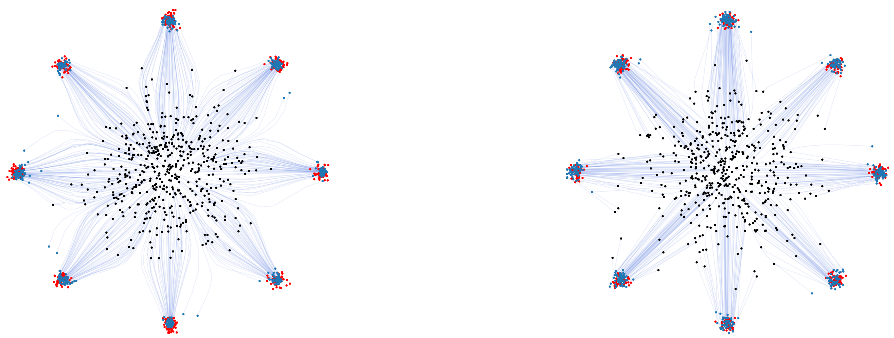
$d = 28^2 = 784$

Flow Based Generative Models

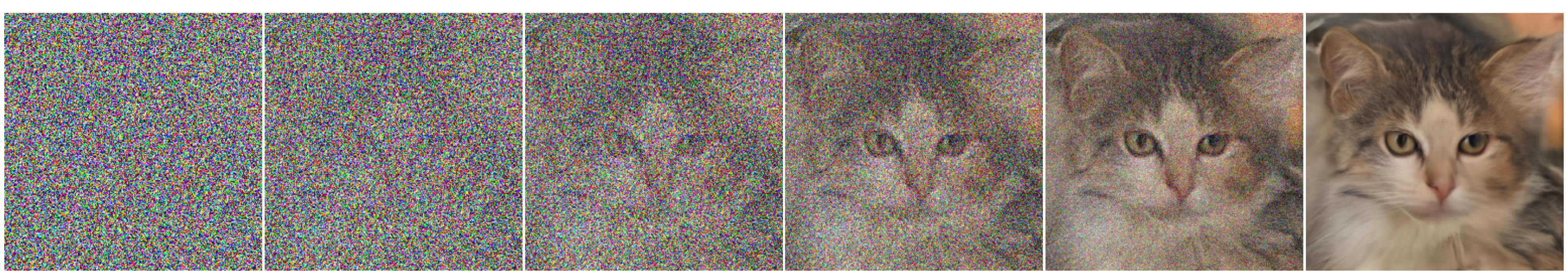
Idea: Learn a curve from μ_Z to μ_X or reverse

Basic methods:

- score-based diffusion models (Sohl-Dickstein 2015; Song/Ermon et al. 2019)
- flow matching (Lipman et al. 2022/23, Liu et al. 2022/23)



Trajectories of points via flow matching from different couplings, $d = 2$



Single trajectory for a flow matching model trained on cat images, $d = 256^2$

Refs: Wald/St.: Flow Matching: Markov kernels, transport plans and stochastic processes, 2025

Theorem (Characterization of absolutely continuous curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$)

$\mu_t : I \to \mathcal{P}_2(\mathbb{R}^d)$ absolutely continuous, here $AC^2(I; \mathcal{P}_2(\mathbb{R}^d))$, if and only if there exists a Borel measurable vector field $v : I \times \mathbb{R}^d \to \mathbb{R}^d$ such that

i) $\|v_t\|_{L_2(\mu_t)} \in L_2(I)$,

ii) curve-velocity pair (μ_t, v_t) satisfies the continuity equation (CE)

$$\partial_t \mu_t + \nabla_x \cdot (\mu_t v_t) = 0$$

Theorem Let $\mu_t : I \to \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve with associated velocity field v_t such that for every compact Borel set $B \subset \mathbb{R}^d$ it holds

$$\int_I \sup_{x \in B} \|v_t(x)\| + \mathrm{Lip}(v_t, B) \;\, \mathrm{d}t < \infty. \qquad (*)$$

Then there exists a solution $\phi : I \times \mathbb{R}^d \to \mathbb{R}^d$ of the ODE

$$\partial_t \phi(t, x) = v_t(\phi(t, x)), \quad \phi(0, x) = x$$

and $\phi(t, \cdot)_\sharp \mu_0 = \mu_0 \circ \phi(t, \cdot)^{-1} = \mu_t$.

If v_t is such that $\phi(1,\cdot)_\sharp\mu_0 = \mu_1$, then use your favorite ODE solver to transport samples from $\mu_0 = \mu_X$ to samples of $\mu_1 = \mu_Z$.

All you need is v_t !

Task: Learn/approximate $(v_t)_t$ by neural network $(v_t^\theta)_t$?

Loss function for learning:

$$\mathcal{L}(\theta) := \mathbb{E}_{(t,x)\sim\mathcal{U}(0,T)\times\mu_t}\left[\left\|v^\theta(t,x) - v(t,x)\right\|^2\right],$$

not accessible, but make it conditional so that up to a constant

$$\mathcal{L}(\theta) = \mathbb{E}_{t\sim\mathcal{U}(0,T),\, x_0\sim\mu_0,\, x\sim\mu_t(\cdot|x_0)}\left[\left\|v^\theta(t,x) - v(t,x|x_0)\right\|^2\right]$$

where in general μ_t starting in p_0 has density

$$p_t(x) = \int p_t(x|y)p_0(y)\,\mathrm{d}(y), \qquad v_t(x) = \frac{1}{p_t(x)}\int p_t(x|y)\,v_t(x|y)p_0(y)\,\mathrm{d}(y)$$

Motivation

Exploding vector fields in existing models: $\|v_t\|_{L_2(\mu_t)} \in L_2([0,1])$ neglected

1. Flow matching from couplings α of $\mu_1 = \mu_Z$ and $\mu_1 = \mu_X$:

$$\mu_t := e_{t,\sharp}\alpha \quad \text{with} \quad e_t := (1-t)x + ty$$

If $\mu_0 = \mathcal{N}(0, I_d)$, $\mu_1 = \delta_0$ and $\alpha = \mu_0 \times \mu_1$, then CE is fulfilled for

$$p_t(x) = \left(2\pi(1-t)^2\right)^{-\frac{d}{2}} \mathrm{e}^{-\frac{\|x\|^2}{2(1-t)^2}}, \qquad t \in [0,1)$$

$$v_t(x) = \nabla\left(\frac{1-t}{t}\log p_t(x) + \frac{1}{2}\|x\|^2\right) = -\frac{x}{1-t}$$

(*) not fulfilled due to singularity at $t = 1$.

But no issue, since ODE has solution $\phi_t(x) = (1-t)x$ so that $v_t(\phi_t(x)) = -x$ and

$$\|v_t\|^2_{L_2(\mu_t)} = \int_{\mathbb{R}^d} \|v_t(\phi_t(x))\|^2 \,\mathrm{d}\mu_0(x) = \int_{\mathbb{R}^d} \|x\|^2 \,\mathrm{d}\mu_0(x) = d$$

2. Diffusion flows: standard diffusion equation

$$\partial_t p_t = \frac{\sigma^2}{2} \nabla \cdot (p_t \nabla \log p_t) = \frac{\sigma^2}{2} \Delta p_t, \quad t \in (0, 1], \quad \lim_{t \downarrow 0} p_t = \delta_0,$$

starting in δ_0 has solution

$$p_t(x) = (2\pi\sigma^2 t)^{-\frac{d}{2}} \mathrm{e}^{-\frac{\|x\|^2}{2\sigma^2 t}}, \quad t \in (0, \infty).$$

Reparameterizing $\tilde{p}_t = p_{1-t}$, $t \in I = [0, 1)$, the reverse flow reads as

$$\tilde{p}_t(x) = \big(2\pi(1-t)\big)^{-\frac{d}{2}} \mathrm{e}^{-\frac{\|x\|^2}{2(1-t)}}, \quad \tilde{p}_0 = \mathcal{N}(0, I_d),$$

$$\tilde{v}_t(x) = -\frac{x}{2(1-t)}, \qquad \|\tilde{v}_t\|^2_{L_2(\tilde{p}_t)} = \frac{d}{4(1-t)}$$

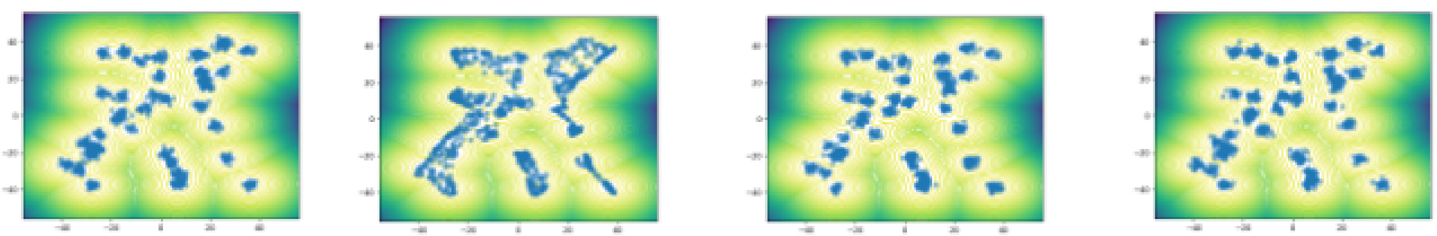
- instability issues concerning the learning of $\tilde{v}_t$ are caused by explosion at times close to the target, and need to be avoided by e.g. *time truncations*, see Kim et al. ICLR 2022
- similar with drifts, see Pidstrigach, NeuRips 2022

Further Examples

3. Flows from linear interpolation of Boltzmann energies: $p_t = e^{-f_t}/Z_t$

$$f_t = (1-t)f_0 + tf_1 \quad \Rightarrow \quad p_t = p_0^{1-t} p_1^t$$

Refs: Mate/Fleuret, TMLR 2023, Maurais/Marzuok ICML 2024, Chemsedine, Wald, Duong, St. ICLR 2025, Berner, Richter et al. 2024, ICLR 2025)



ground truth | linear | learned | gradient flow

4. Poisson flows: on $\mathbb{R}^d$ augmented by an additional dimension z:

$$-\Delta\phi(x,z) = p_{\text{data}}(x)\delta_0(z), \quad x \in \mathbb{R}^d, z \in \mathbb{R}.$$

z-dependent velocity field, which explodes for $z \to 0$

Refs: Xu, Liu, Tegmark, Jakkola, NeuRips 2022

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

In General

It appears to be unphysical that particles in diffusion can travel with unbounded velocities, contradicting Einstein's principles of relativity.

On the other hand

Diffusion (mathematical) can be seen as the limit of the damped wave equation with infinite damping and infinite propagation speed.

Let us consider the damped wave equation!

Refs: Liu, Lu, Xu, Jaakkola, Tengmark, GenPhys: from physical processes to generative models, 2023

Outline

Particle starts at $x_0 = 0$ to the right and moves with step Δt and velocity $c > 0$ in either positive or negative direction, where

- same direction with probability $1 - a\Delta t$
- reverse direction with the probability $a\Delta t$ $\quad\rightarrow$ random variable $\varepsilon_k \in \{\pm 1\}$

Particle displacement S_n after n steps (random variable):

$$S_n = c\Delta t(1 + \varepsilon_1 + \varepsilon_1\varepsilon_2 + \ldots + \varepsilon_1 \ldots \varepsilon_n) = c\Delta t \sum_{k=0}^{n} (-1)^{N_k}$$

where $N_k \sim B(k, a\Delta t)$ number of reversals up to step k

- Continuous limit for $n \to \infty$ such that $na\Delta t \to at$, then $N_n \to \mathrm{Poi}(at)$

Number of reversals $N(t)$ up to time t is given by a

homogeneous Poisson point process with rate a:

i) $N(0) = 0$;

ii) Independent increments of $N(t)$

iii) $N(t) - N(s) \sim \mathrm{Poi}\big(a(t-s)\big)$ for all $0 \le s < t$.

Particle displacement:

$$S(t) = c \underbrace{\int_0^t (-1)^{N(s)} \, \mathrm{d}s}_{\tau_t \text{ random time}}$$

Kac process starting in 0: $\quad K(t) := \mathrm{B}_{\frac{1}{2}} S(t), \qquad \mathrm{B}_{\frac{1}{2}} \sim B(1, \tfrac{1}{2})$

Kac process starting in X_0: $\quad X_t := X_0 + K(t)$

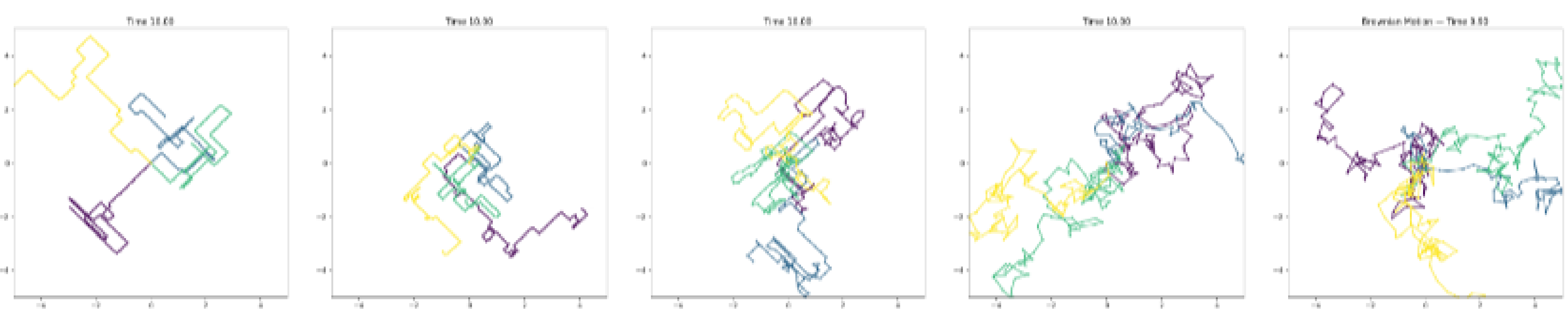


Figure 1: Paths of the componentwise Kac walk in 2D, simulated until time $T = 10$ with damping/velocity parameters $(a, c) = (1,1), (2,1), (4,2), (25,5)$, and a standard Brownian motion (right).

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

Damped wave equation (telegrapher's equation):

$$\begin{aligned} \partial_{tt}u(t,x) + 2a\,\partial_t u(t,x) &= c^2\Delta_x u(t,x) \\ u(0,x) &= f_0(x) \\ \partial_t u(0,x) &= 0 \end{aligned}$$

- $a > 0$ damping coefficient
- $c > 0$ velocity of the wave front

Distributional solution (f_0 distribution):

$$u(t,x) = \frac{1}{2}e^{-at}\Big(f_0(x+ct) + f_0(x-ct) + \beta ct\int_{x-ct}^{x+ct}\frac{I_1(\beta r_t(x-y))}{r_t(x-y)}f_0(y)\,\mathrm{d}y$$
$$+\,\beta\int_{x-ct}^{x+ct} I_0(\beta r_t(x-y))\,f_0(y)\,\mathrm{d}y\Big),$$

where I_k is k-th modified Bessel function of first kind and

$$r_t(x) := \sqrt{c^2t^2 - x^2}, \quad \beta := \frac{a}{c},$$

- Special case $f_0 = \delta_0$:

$$u(t,x) = \frac{1}{2}e^{-at}\big(\delta_0(x+ct) + \delta_0(x-ct)\big) + \tilde{u}(t,x),$$

where

$$\tilde{u}(t,x) := \frac{1}{2}e^{-at}\Big(\beta ct\frac{I_1(\beta r_t(x))}{r_t(x)} + \beta I_0(\beta r_t(x))\Big)1_{[-ct,ct]}(x)$$

- Associated probability measure μ_t:

$$\mu_t(A) = \frac{e^{-at}}{2}\left(\delta_{-ct}(A) + \delta_{ct}(A)\right) + \int_A \tilde{u}(t,x)\,\mathrm{d}x \qquad \forall A \in \mathcal{B}(\mathbb{R})$$

is the probability distribution of the Kac process $K(t)$ starting in 0.

- Let $X_0 \sim f_0$ have probability density. Then, the solution $u(t,x)$ is the probability density of the Kac process $(X_t)_{t\geq 0}$ starting in X_0.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

Theorem

Let $f_0 = \delta_0$. Then μ_t associated to $u(t,\cdot)$ solves (CE)

$$\partial_t \mu_t = -\partial_x(\mu_t v_t)$$

in a weak sense, where the velocity field $v_t(x) = v(t,x)$ is given by

$$v(t,x) := \begin{cases} \frac{x}{t+\frac{r_t(x)}{c}\frac{I_0(\beta r_t(x))}{I_1(\beta r_t(x))}} & \text{if} \quad x \in (-ct, ct), \\ c & \text{if} \quad x = ct, \\ -c & \text{if} \quad x = -ct, \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

Compare to diffusion starting in δ_0: $v_t(x) = \frac{x}{2t}$

When starting in δ_{x_0}, we know $v(t,x)$ and we can sample from $u(t,x)$!

Outline

Proposition (Kac' insertion) (Griego/Hersh 1971)

For any initial function $f_0 \in H^2(\mathbb{R}^d)$, $d \geq 1$, let $w_c(t,x)$ be the solution of the undamped wave equation with velocity $c > 0$

$$\begin{aligned} \partial_{tt} w(t,x) &= c^2 \Delta w(t,x), \\ w(0,x) &= f_0(x), \\ \partial_t w(0,x) &= 0. \end{aligned}$$

Then, the function $u : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$u(t,x) := \mathbb{E}\big[w_c(\tau_t, x)\big]$$

solves the multidimensional damped wave equation.

Drawback: For $d > 1$, the $u(t,x)$ obtained by Kac' insertion method might not stay a probability density if we start with a probability density f_0,
see example in Tautz/Lerche 2016.

Theorem (Griego/Hersh 1971)

For $f_0 \in H^2(\mathbb{R}^d)$. Let W be a one-dimensional Wiener process. Then

$$h(t,x) := \mathbb{E}\left[w_1\left(\sigma W(t), x\right)\right], \quad \sigma > 0$$

solves the heat equation

$$\partial_t h(t,x) = \frac{\sigma^2}{2}\Delta h(t,x),$$
$$h(0,x) = f_0(x).$$

- If $\mathbf{X}_0 \sim f_0$ with probability density f_0, then h is the probability density of a d-dimensional Wiener process $\sigma\mathbf{W}(t)$ starting in $\mathbf{X}_0$.

Theorem

For $f_0 \in H^2(\mathbb{R}^d)$, let u be given by the Kac' insertion with $a > 0$, $c \geq c_0 > 0$, and h the solution of the heat equation with $\sigma^2 := \frac{c^2}{a}$. Then

$$\|h(t,\cdot) - u(t,\cdot)\|_{L_2(\mathbb{R}^d)} \leq C(c_0, d)\left(\frac{1}{at} + \frac{1}{c(ta)^{\frac{1}{2}}}\right) \quad \text{for all } t > 0.$$

- h and u converge to each other for $t \to \infty$
- for any $t \geq 0$, the solution $u^{a,c}(t,\cdot)$ converges to solution $h(t,\cdot)$ for $a, c \to \infty$ with fixed $\sigma^2 = \frac{c^2}{a}$
- in particular, this holds true in one dimension, where we will need it

3.2 Multidimensional Kac Process

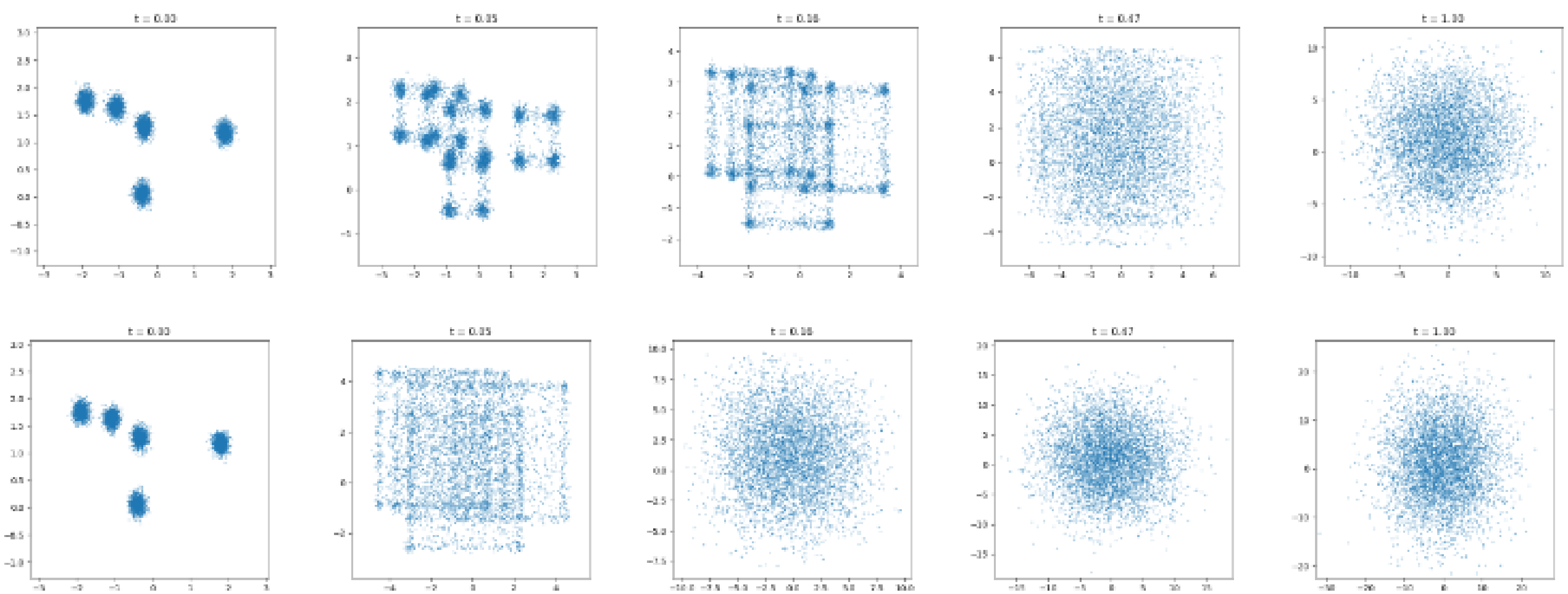
For a random variable $\mathbf{X}_0 \in \mathbb{R}^d$, $d \geq 1$, we consider i.i.d. copies $K^i(t)$, $i = 1, ..., d$, of a Kac process starting in $\mathbf{0}$.

- d-dimensional Kac process starting in $\mathbf{0} \in \mathbb{R}^d$:

$$\mathbf{K}(t) := (K^1(t), ..., K^d(t))$$

- d-dimensional Kac process starting in $\mathbf{X}_0$:

$$\mathbf{X}_t := \mathbf{X}_0 + \mathbf{K}(t)$$



Kac process $\mathbf{X}_t$ with $(a, c) = (1, 1)$ and $(a, c) = (25, 5)$

Proposition (CE and Lipschitz Continuity)

Let $P_{\mathbf{X}_0} \in \mathcal{P}_2(\mathbb{R}^d)$. Then, for the probability distribution flow $(P_{\mathbf{X}_t})_{t\geq 0}$ of the d-dimensional Kac process $\mathbf{X}_t$ starting in $\mathbf{X}_0$, the following holds true:

i) $(P_{\mathbf{X}_t})_{t\geq 0}$ is Lipschitz continuous in the Wasserstein space with constant $d^{\frac{1}{2}}c$, and $\mathbf{X}_t$ is almost surely Lipschitz continuous in $\mathbb{R}^d$ with the same constant.

ii) $(P_{\mathbf{X}_t})_{t\geq 0}$ is an absolutely continuous curve in $AC^2([0,T];\mathcal{P}_2(\mathbb{R}^d))$ for any $T>0$, and admits a velocity field fulfilling

$$\|v_t\|^2_{L_2(\mu_t)} \leq dc^2 \quad \text{for a.e. } t \in [0,\infty),$$

iii) The support of $(P_{\mathbf{X}_t})_{t\geq 0}$ grows at most linearly in time.

Theorem (Convergence)

Let f_0 with $f_0^i \in H^2(\mathbb{R})$ be a probability density of a random variable $\mathbf{X}_0 = (X_0^1, \ldots, X_0^d)$ with independent components such that

$$f_0(x) = f_0(x^1, \ldots, x^d) = \prod_{i=1}^{d} f_0^i(x^i),$$

$\mathbf{X}_0 + \mathbf{K}(t)$ with probability density u_t, and
$\mathbf{X}_0 + \sigma \mathbf{W}(t)$, $\sigma^2 := \frac{c^2}{a}$ with probability density h_t.
Assume that $\mathbf{X}_0$ is independent of $\mathbf{K}(t)$ and $\sigma \mathbf{W}(t)$, respectively.
Then it holds

$$\|h(t,\cdot) - u(t,\cdot)\|_{L_2(\mathbb{R}^d)} \le C_d(c_0, \sigma) \left(\frac{1}{a t^{1+\frac{d-1}{4}}} + \frac{1}{c a^{\frac{1}{2}} t^{\frac{1}{2}+\frac{d-1}{4}}} \right) + R_d(a,c,t),$$

where the function $R_d(a,c,t)$ satisfies $\lim_{a,c\uparrow\infty} R_d(a,c,t) = 0$ for any t, and $R_d(a,c,\cdot) \in \mathcal{O}(t^{-(\frac{1}{2}+\frac{d}{4})})$ for any a, c.

Hence, the density of our multidimensional Kac process L_2-converges to the density of the Wiener process with rate $\mathcal{O}(t^{-(\frac{1}{2}+\frac{d-1}{4})})$.

Outline

1. Basic Idea of Generative Flows
2. Kac Model and Telegrapher's Equation in 1D
3. Multidimensional Generalization of the Telegrapher's equation
4. Velocity Field of the Multidimensional Kac Flow
5. Numerical Examples

Lemma

Let $\mu_t \in AC^2(I;\mathcal{P}_2(\mathbb{R}^d))$ such that

$$\mu_t(x) = \prod_{i=1}^{d} \mu_t^i(x^i) \quad \text{for all } x = (x^1, \ldots, x^d).$$

Furthermore, assume that the one-dimensional measure flows μ_t^i satisfy $\mu_t^i \in AC^2(I;\mathcal{P}_2(\mathbb{R}))$ and a CE with some vector field v_t^i, i.e.

$$\partial_t \mu_t^i(x^i) = -\partial_{x^i}(\mu_t^i(x^i) v_t^i(x^i)) \quad \text{for all } x^i \in \mathbb{R}.$$

Then μ_t satisfies the CE

$$\partial_t \mu_t(x) = -\nabla \cdot (\mu_t(x) v_t(x))$$

with the velocity field

$$v_t(x) = (v_t^1(x^1), \ldots, v_t^d(x^d)) \quad \text{for all } x = (x^1, \ldots, x^d).$$

Theorem

For a fixed $x_0 \in \mathbb{R}^d$, let $\mathbf{X}_t := x_0 + \mathbf{K}(t)$ be the Kac process starting in x_0. Then, $\mu_t(\cdot|x_0) := P_{\mathbf{X}_t}$ satisfies a continuity equation (**??**) with the velocity field

$$v_t(x|x_0) := (v_t^1(x^1|x_0^1), \dots, v_t^d(x^d|x_0^d)) \quad \text{for all } x = (x^1, \dots, x^d),$$

where the univariate components $v_t^i(x^i|x_0^i)$ are given by

$$v_t^i(x^i|x_0^i) := \begin{cases} (x^i - x_0^i)\left(t + \dfrac{r_t(x^i - x_0^i)}{c} \dfrac{I_0(\beta r_t(x^i - x_0^i))}{I_1(\beta r_t(x^i - x_0^i))}\right)^{-1} & \text{if } x^i \in (x_0^i - ct, x_0^i + ct) \\ c & \text{if } x^i = x_0^i + ct \\ -c & \text{if } x^i = x_0^i - ct \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

Furthermore, it holds

$$v_t^i(x^i|x_0^i) = \mathbb{E}[\dot{X}^i(t) \mid X_t^i = x^i, X_0^i = x_0^i].$$

Corollary

Let $\mathbf{X}_t$ be the multidimensional Kac process starting in $\mathbf{X}_0 \sim \mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Assume that $\mathbf{X}_0$ is independent of $\mathbf{K}(t)$, and assume that $\mathbf{X}_0$ has a density f_0. Then, the marginal distribution of $\mathbf{X}_t$ admits the density u_t given by

$$u_t(x) = \int u_t(x|x_0) f_0(x_0) \, \mathrm{d}x_0,$$

where $u_t(\cdot|x_0)$ is the conditional distribution of $\mathbf{X}_t$ conditioned to $\mathbf{X}_0 = x_0$.

Denote the conditional velocity field corresponding to $u_t(\cdot|x_0)$ by $v_t(\cdot|x_0)$. Then, the vector field

$$v_t(x) := \frac{1}{u_t(x)} \int u_t(x|x_0)\, v_t(x|x_0) f_0(x_0) \, \mathrm{d}x_0$$

satisfies the CE

$$\partial_t u_t(x) = -\nabla \cdot \big(u_t(x) v_t(x)\big)$$

in the distributional sense.

Loss Function for Learning the Velocity Field

Loss function:

$$\mathcal{L}(\theta) := \mathbb{E}_{(t,x)\sim\mathcal{U}(0,T)\times u_t}\left[\left\|v^\theta(t,x) - v(t,x)\right\|^2\right]$$

Up to a constant

$$\mathcal{L}(\theta) = \mathbb{E}_{t\sim\mathcal{U}(0,T),\, x_0\sim\mu_0,\, x\sim u_t(\cdot|x_0)}\left[\left\|v^\theta(t,x) - v(t,x|x_0)\right\|^2\right]$$

Remark: Suppose that the times at which the Poisson process jumps are given by $0 < j_1 < j_2 < \cdots < j_{N(t)} < t$. Then τ_t can be expressed exactly as

$$\tau_t = j_1 + \sum_{k=1}^{N(t)-1} (-1)^k (j_{k+1} - j_k) + (-1)^{N(t)}(t - j_{N(t)}).$$

where we define $j_0 := 0$, see Zhang et al., Mathematics and Computers in Simulation 2018

Outline

1. Basic Idea of Generative Flows
2. Kac Model and Telegrapher's Equation in 1D
3. Multi-dimensional Generalizations of the Model
4. Velocity Field of the Multidimensional Kac Flow
5. Numerical Examples (in Progress)

5. Numerical Examples

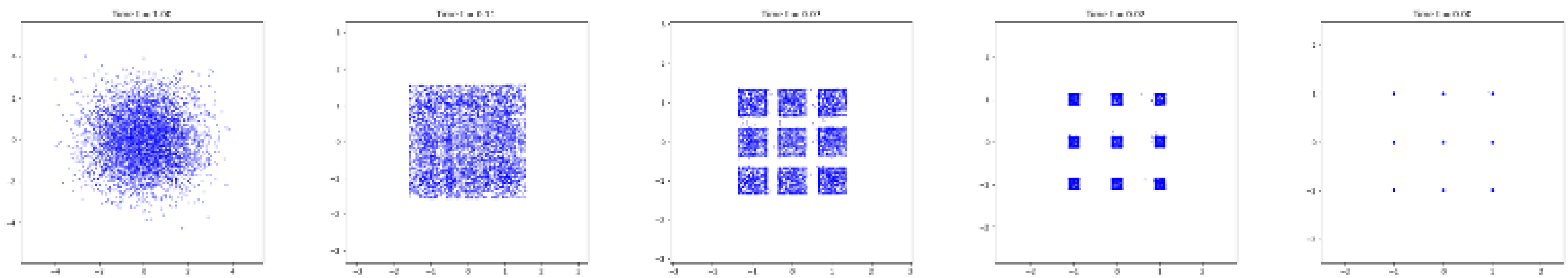
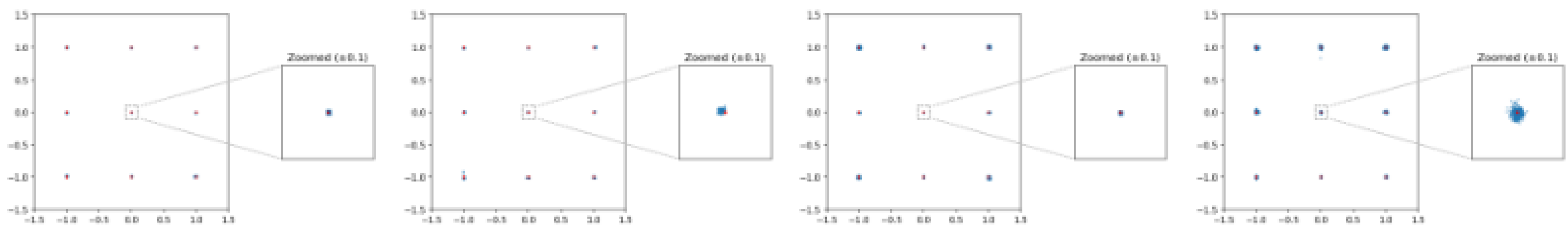


Figure 4: Backward evolution of the *learned* Kac flow for $(a, c) = (25, 5)$, see also Figure 5.



$(a, c) = (25, 5)$ 80k iterations | $(a, c) = (64, 8)$ 280k iterations | $(a, c) = (1600, 40)$ 460k iterations | diffusion 420k iterations

Figure 5: Generated samples (blue) vs. ground truth (red) at the indicated iteration for each model. The Kac model can precisely recover the small modes, while the diffusion model creates "blobs".

5. Numerical Examples

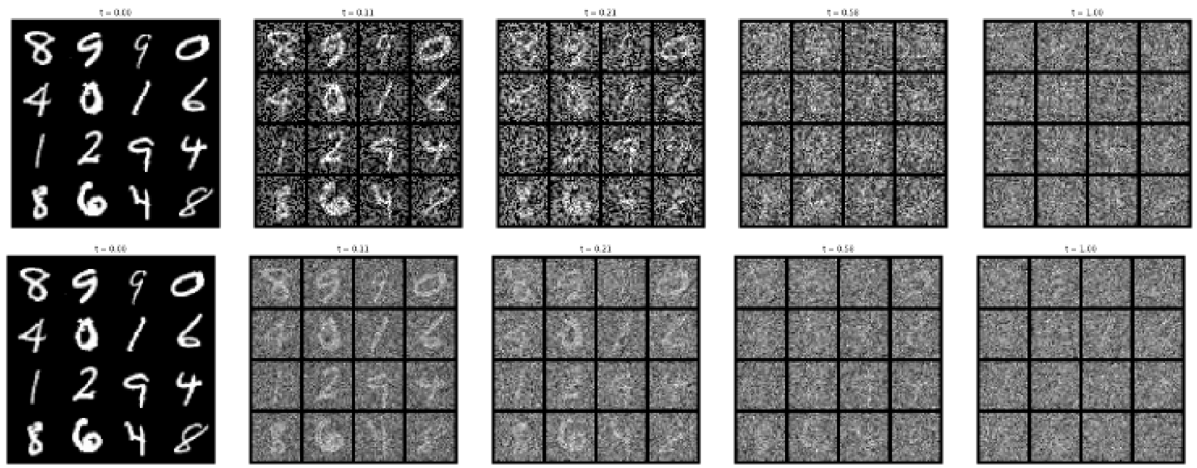
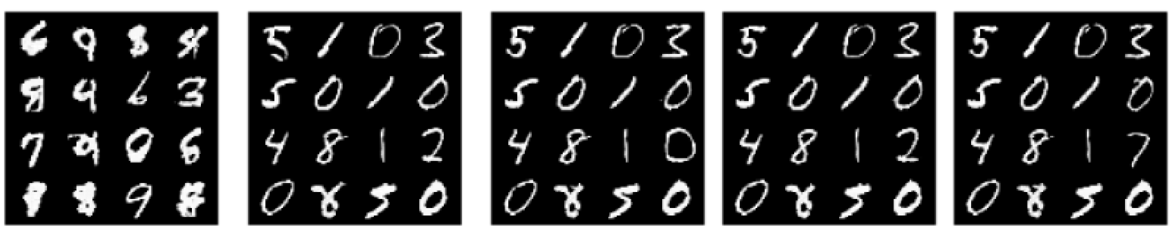
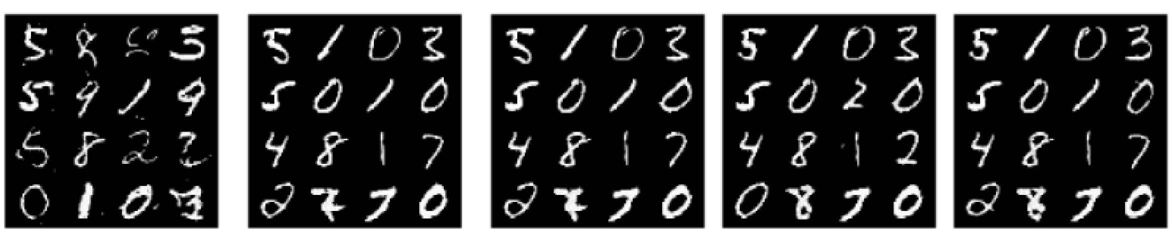


Figure 6: The standard forward Kac process starting in MNIST data, simulated until time $T = 10$ with $(a, c) = (1, 1)$ (upper row) and $(a, c) = (25, 5)$ (lower row).



$(a, c) = (1, 1)$ $(a, c) = (25, 5)$ $(a, c) = (100, 10)$ $(a, c) = (225, 15)$ $(a, c) = (900, 30)$

Figure 7: Generation results of the mean-reverting Kac model for $T = 1$.

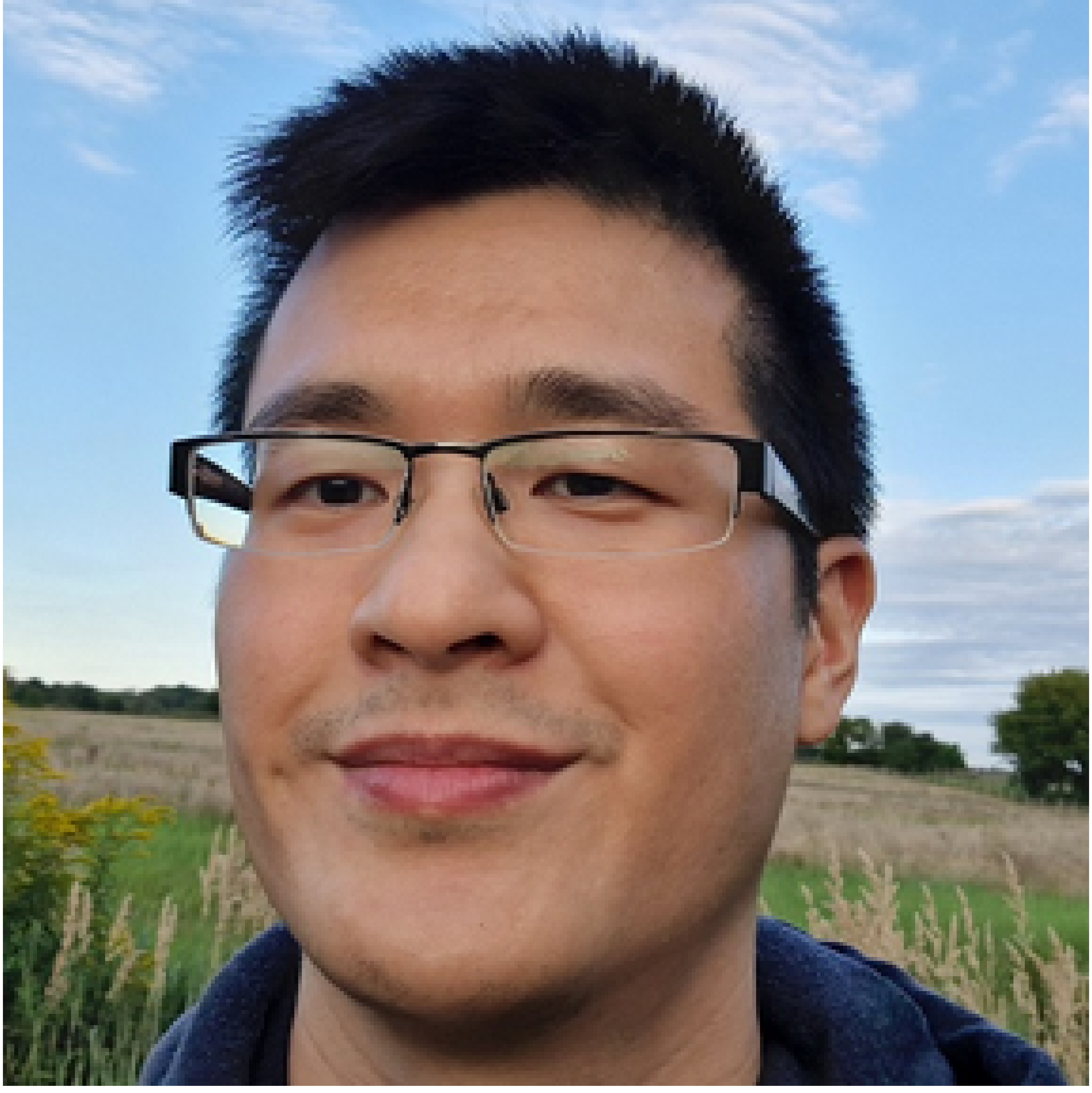


$(a, c) = (1, 1)$ $(a, c) = (25, 5)$ $(a, c) = (100, 10)$ $(a, c) = (225, 15)$ $(a, c) = (900, 30)$

Figure 8: Generation results of the mean-reverting Kac model for $T = 5$.

MANY THANKS FOR YOUR ATTENTION!

HAVE a NICE SOMMER!

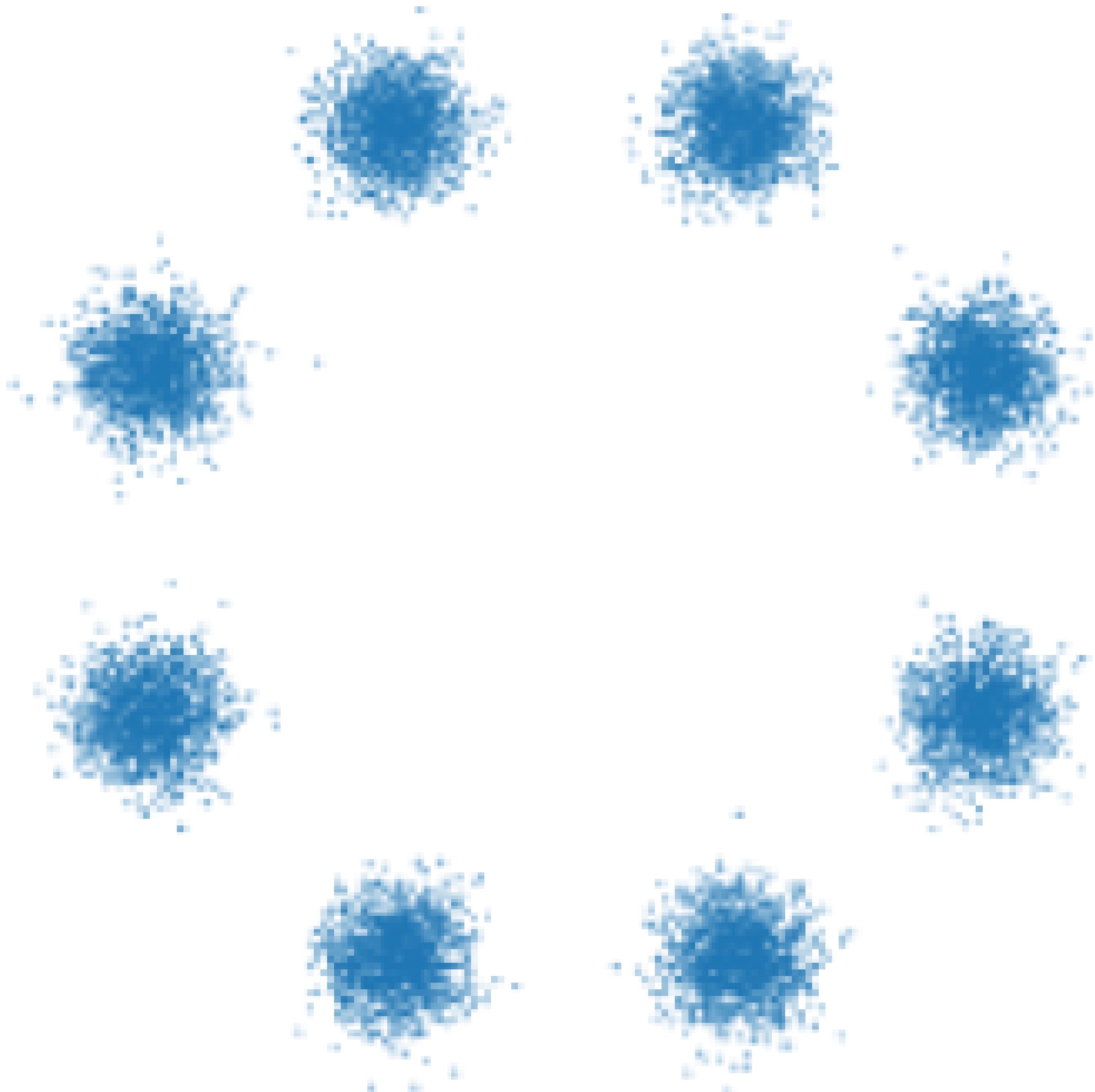


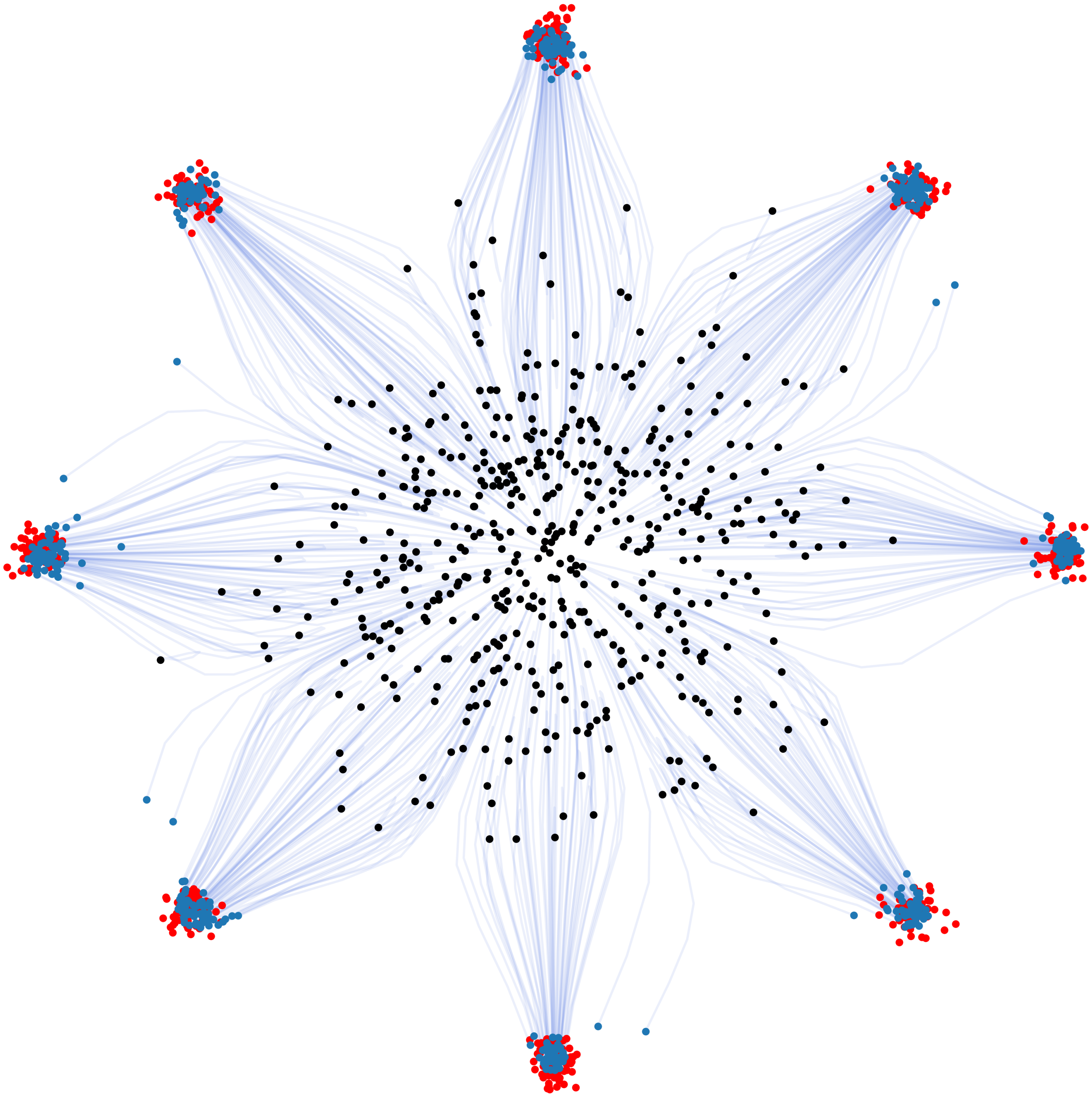
Berlin Mathematics Research Center

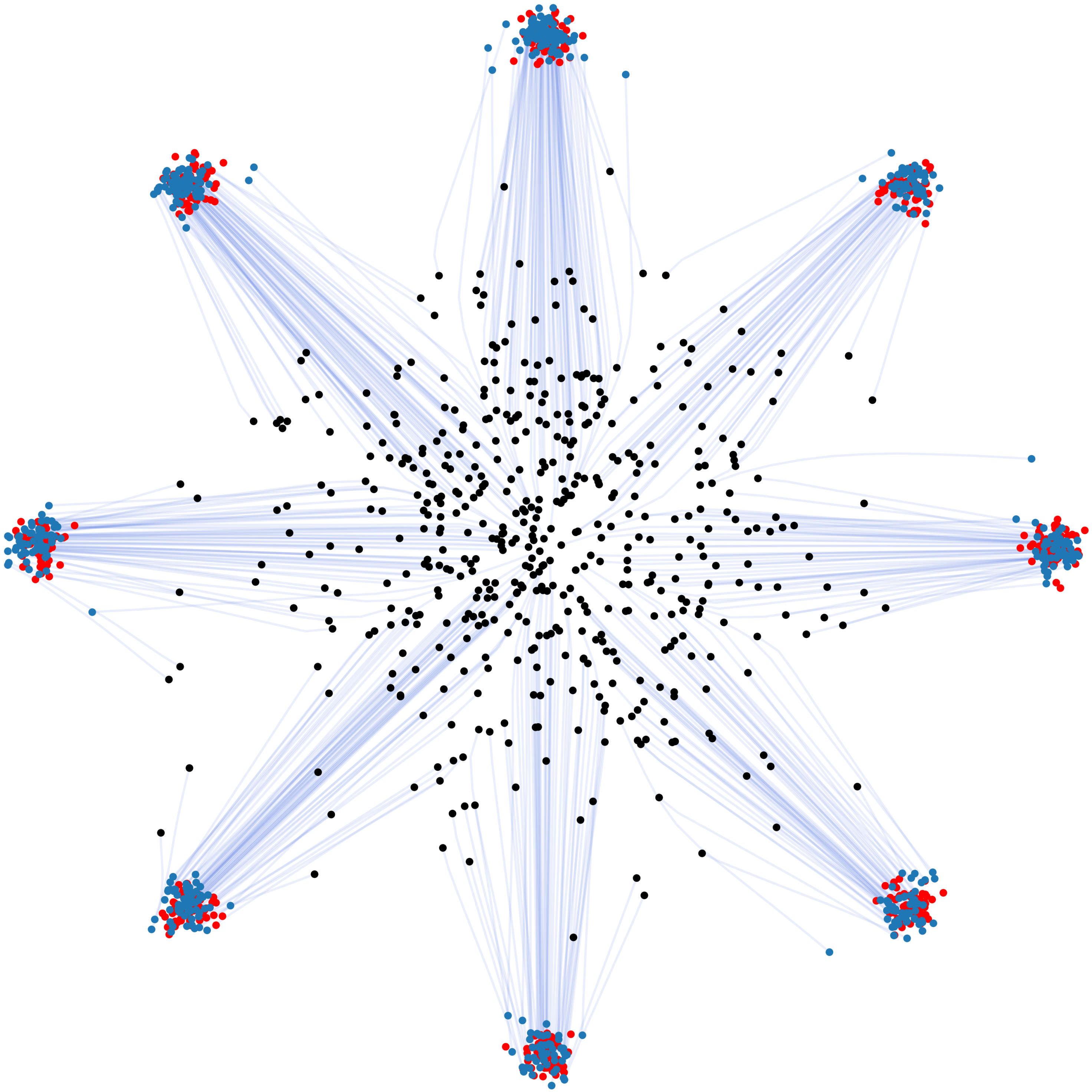
MATH+

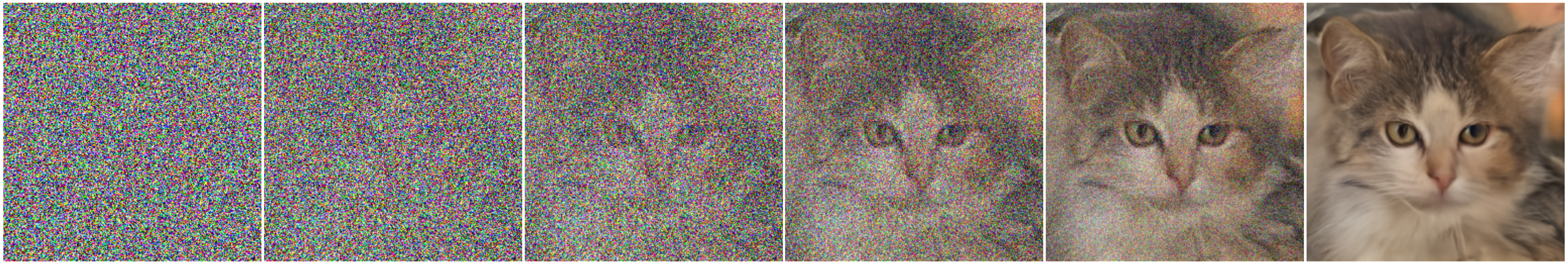
Funded under Germany's Excellence Strategy by

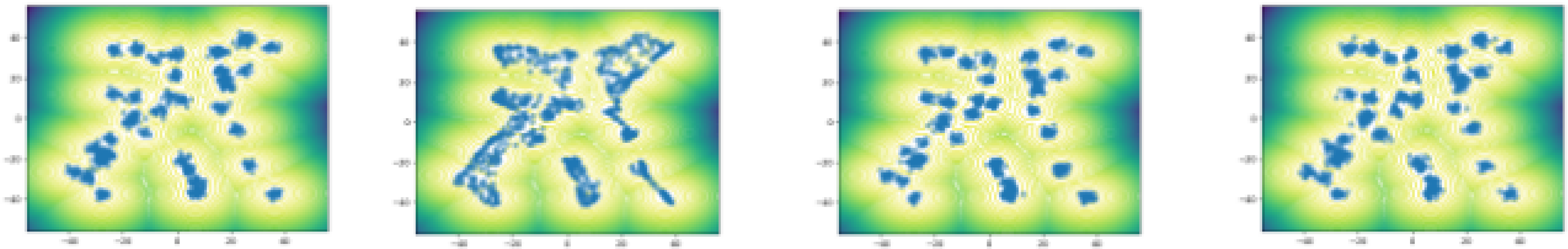
DFG Deutsche Forschungsgemeinschaft











ground truth linear learned gradient flow

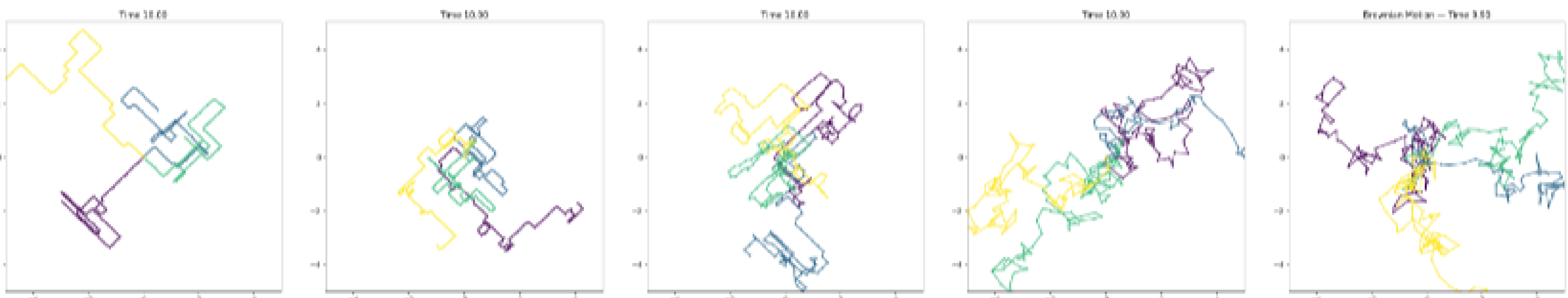
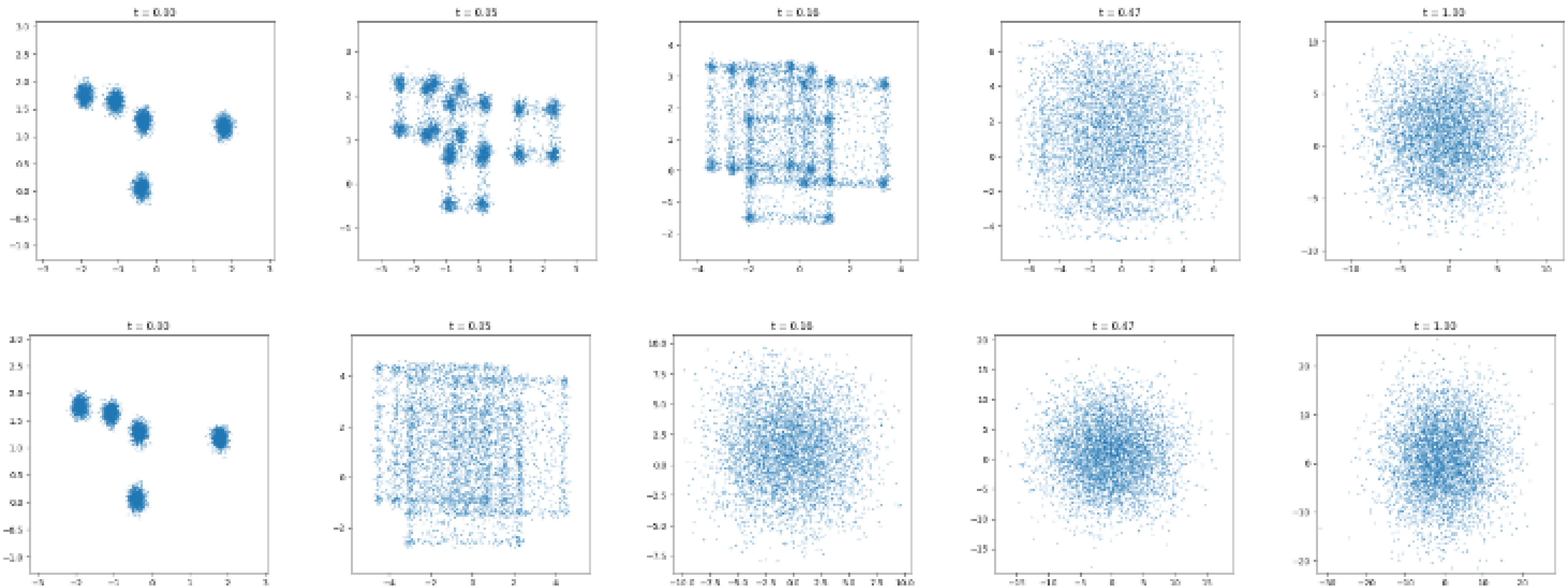


Figure 1: Paths of the componentwise Kac walk in 2D, simulated until time $T = 10$ with damping/velocity parameters $(a, c) = (1, 1), (2, 1), (4, 2), (25, 5)$, and a standard Brownian motion (right).


t = 0.00
t = 0.05
t = 0.16
t = 0.47
t = 1.00
t = 0.00
t = 0.05
t = 0.16
t = 0.47
t = 1.00

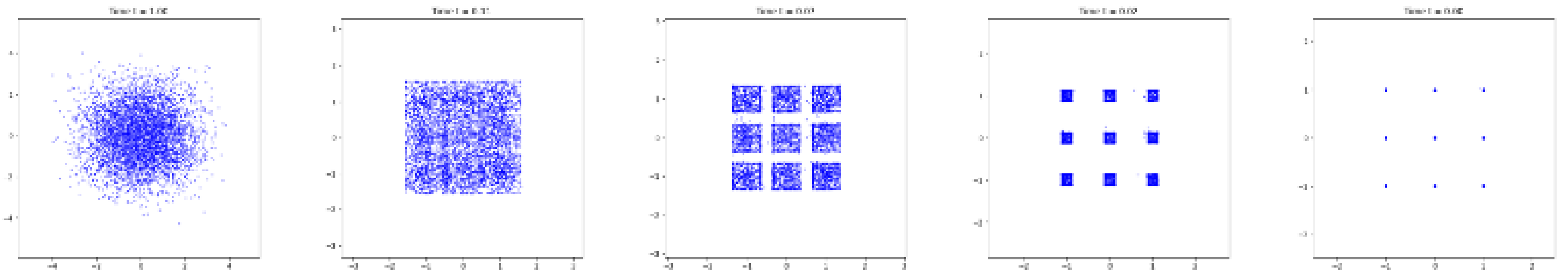
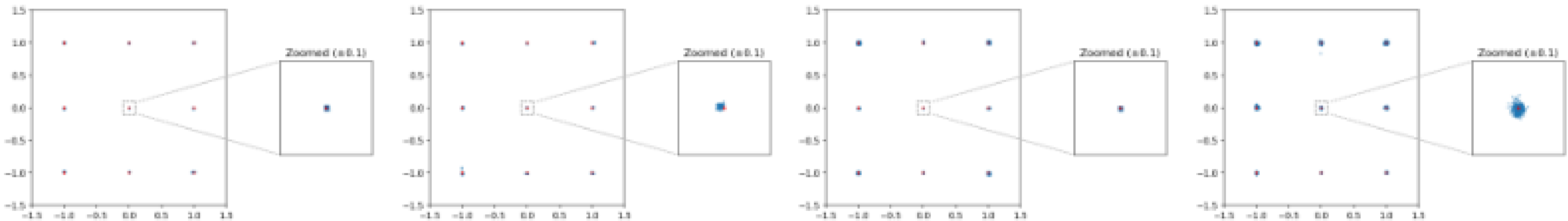


Figure 4: Backward evolution of the *learned* Kac flow for $(a, c) = (25, 5)$, see also Figure 5.



$(a, c) = (25, 5)$
80k iterations

$(a, c) = (64, 8)$
280k iterations

$(a, c) = (1600, 40)$
460k iterations

diffusion
420k iterations

Figure 5: Generated samples (blue) vs. ground truth (red) at the indicated iteration for each model. The Kac model can precisely recover the small modes, while the diffusion model creates "blobs".

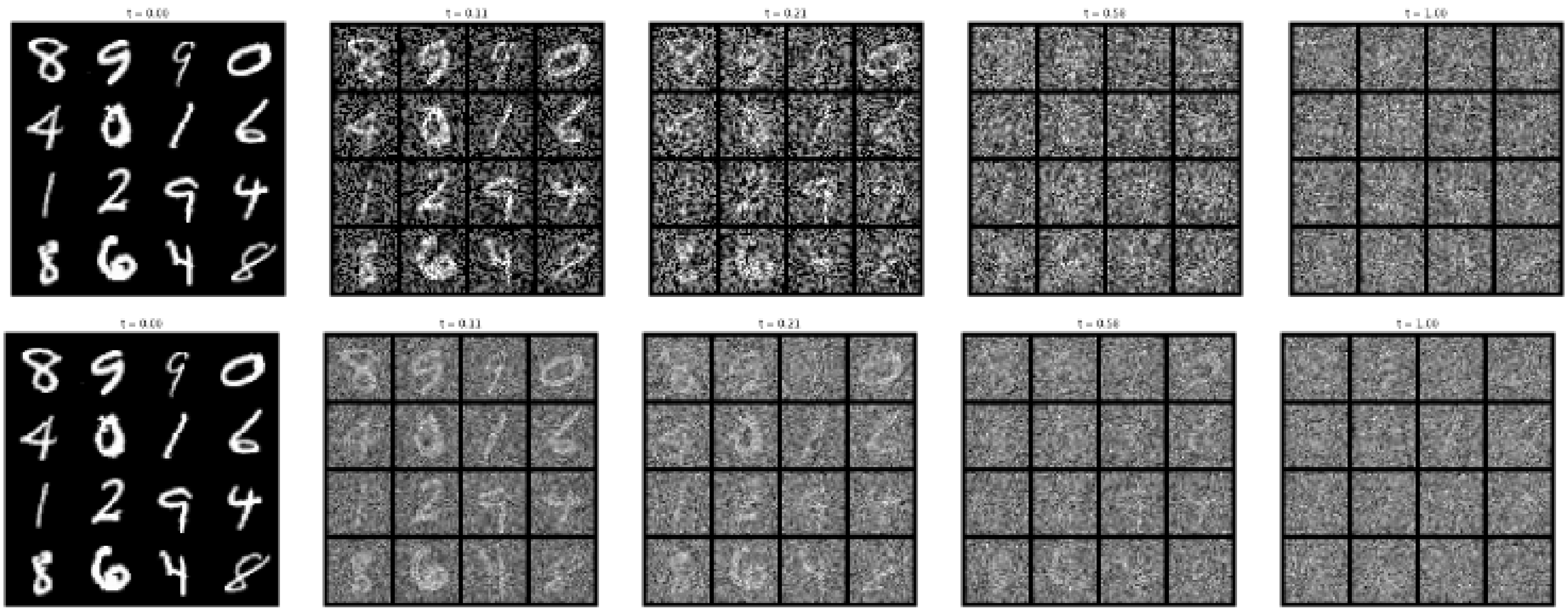
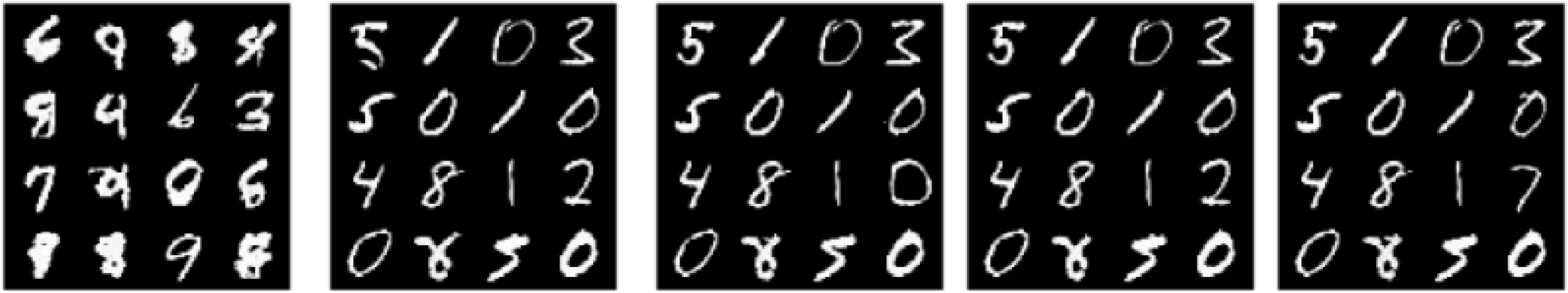
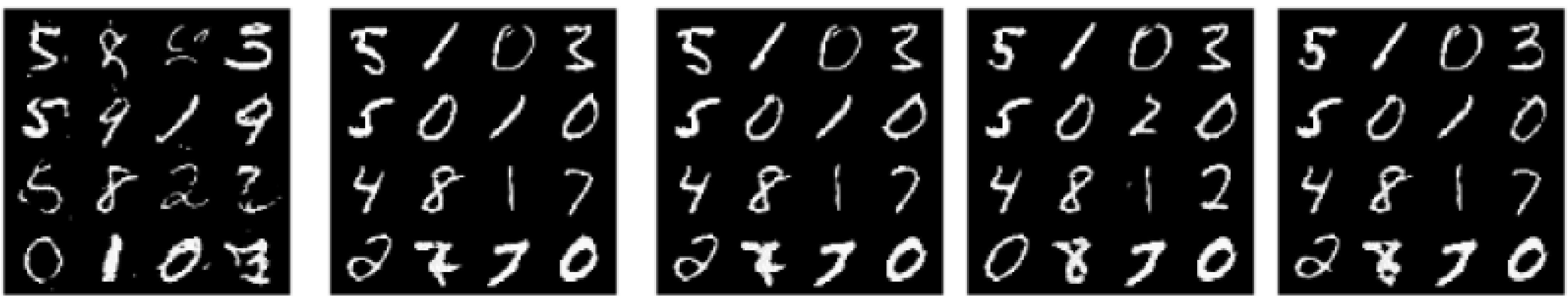


Figure 6: The standard forward Kac process starting in MNIST data, simulated until time $T = 10$ with $(a, c) = (1, 1)$ (upper row) and $(a, c) = (25, 5)$ (lower row).



$(a, c) = (1, 1)$ $(a, c) = (25, 5)$ $(a, c) = (100, 10)$ $(a, c) = (225, 15)$ $(a, c) = (900, 30)$

Figure 7: Generation results of the mean-reverting Kac model for $T = 1$.



$(a, c) = (1, 1)$ $(a, c) = (25, 5)$ $(a, c) = (100, 10)$ $(a, c) = (225, 15)$ $(a, c) = (900, 30)$

Figure 8: Generation results of the mean-reverting Kac model for $T = 5$.