Cell dynamics inference using Schrödinger Bridges

Application to Tcells data

Happy New Year 🎊🥂

Philippe Martin - 05.01.2026

Topics

Goal: Gain a good understanding of SDEs

1. Motivations
2. Brief History
3. From Langevin Dynamics to Fokker-Planck
    - Master Equation
    - Kramers-Moyal Expansion
    - Moments
4. Application to current data

1. Motivations

Study cell dynamics through Langevin Dyanmics

$$dx(t) = f(t, x)dt + g(t, x)d\mathbf{w}$$

Drift

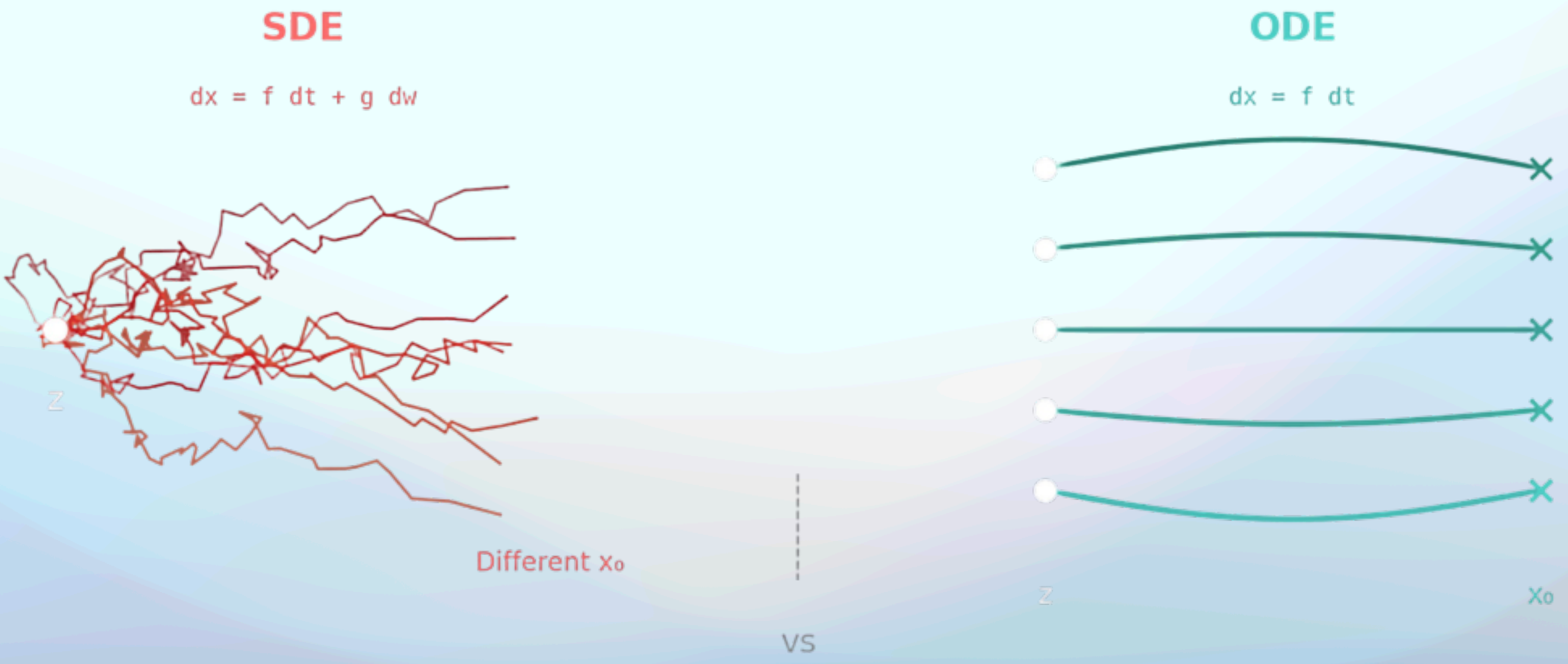
Diffusion

1. Motivations

Study cell dynamics through Langevin Dyanmics

$$dx(t) = f(t, x)dt + g(t, x)d\mathbf{w}$$

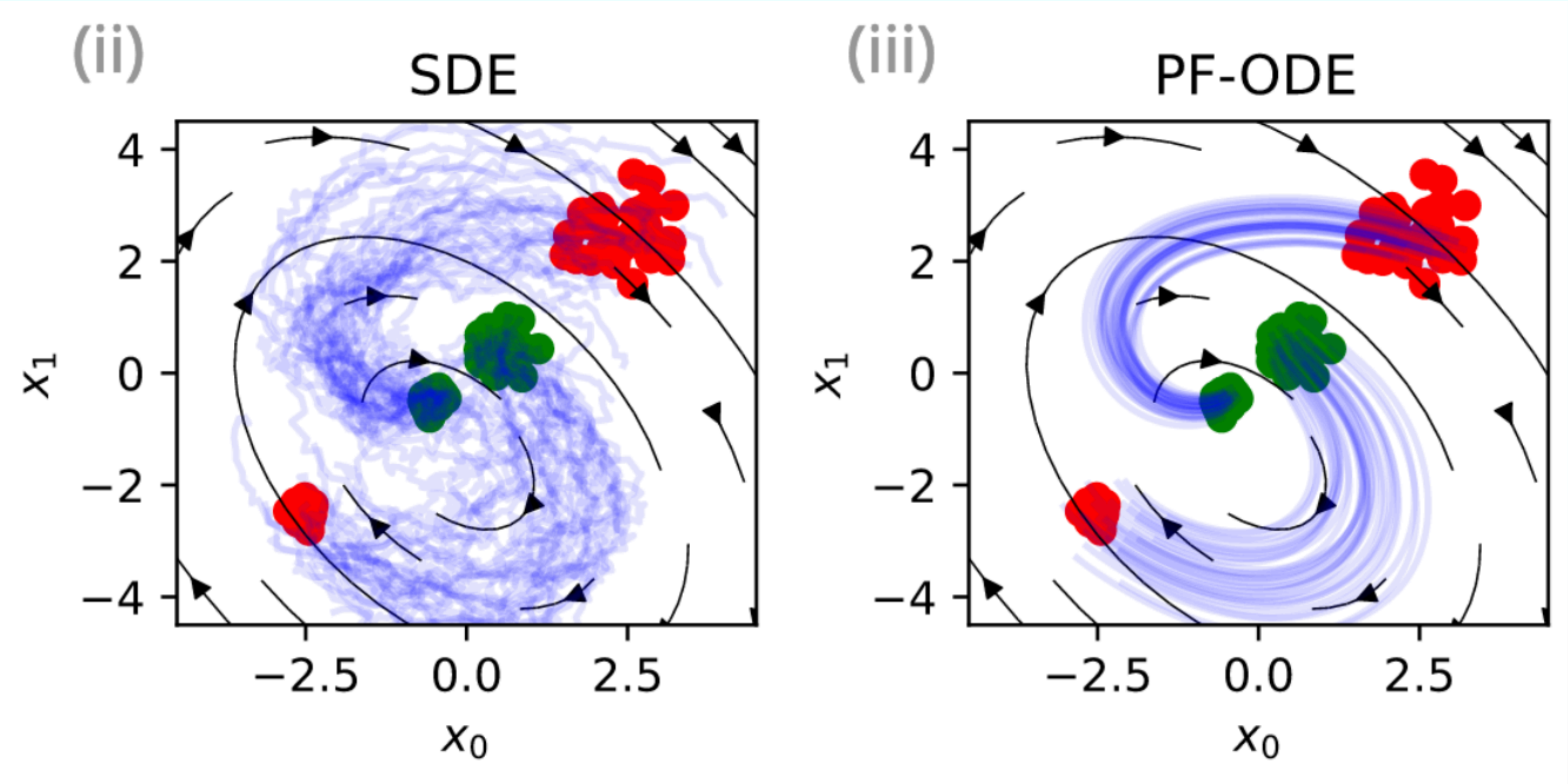
Drift ← $f(t, x)dt$; $g(t, x)d\mathbf{w}$ → Diffusion

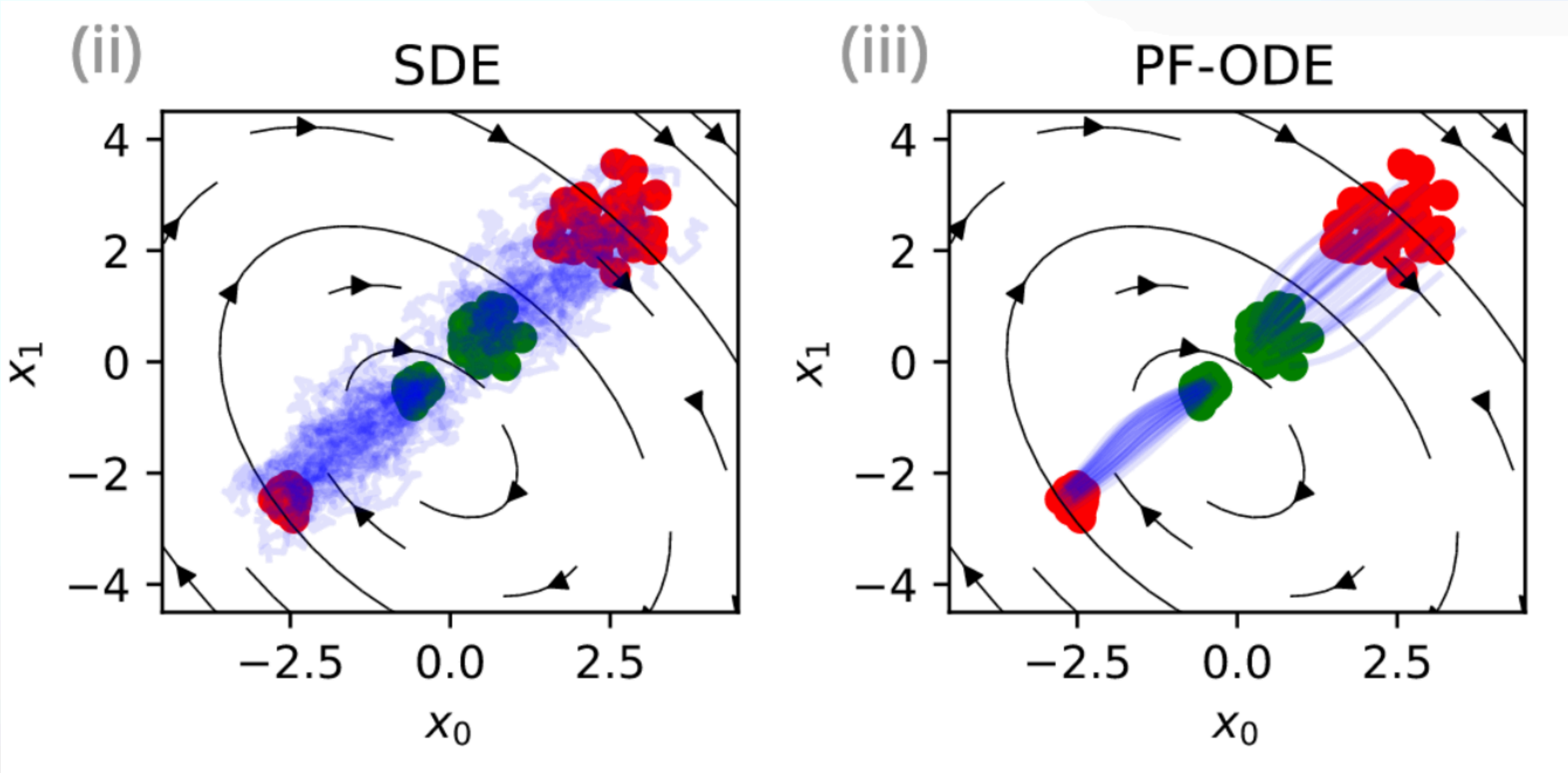


1. Motivations

Always the same: how do we connect two distributions ?

1. Stochastic or Deterministic?
2. Do cell dynamics follow optimal strategy ? Optimal according to what ?
3. If stochastic dynamics, what is the true noise type ?





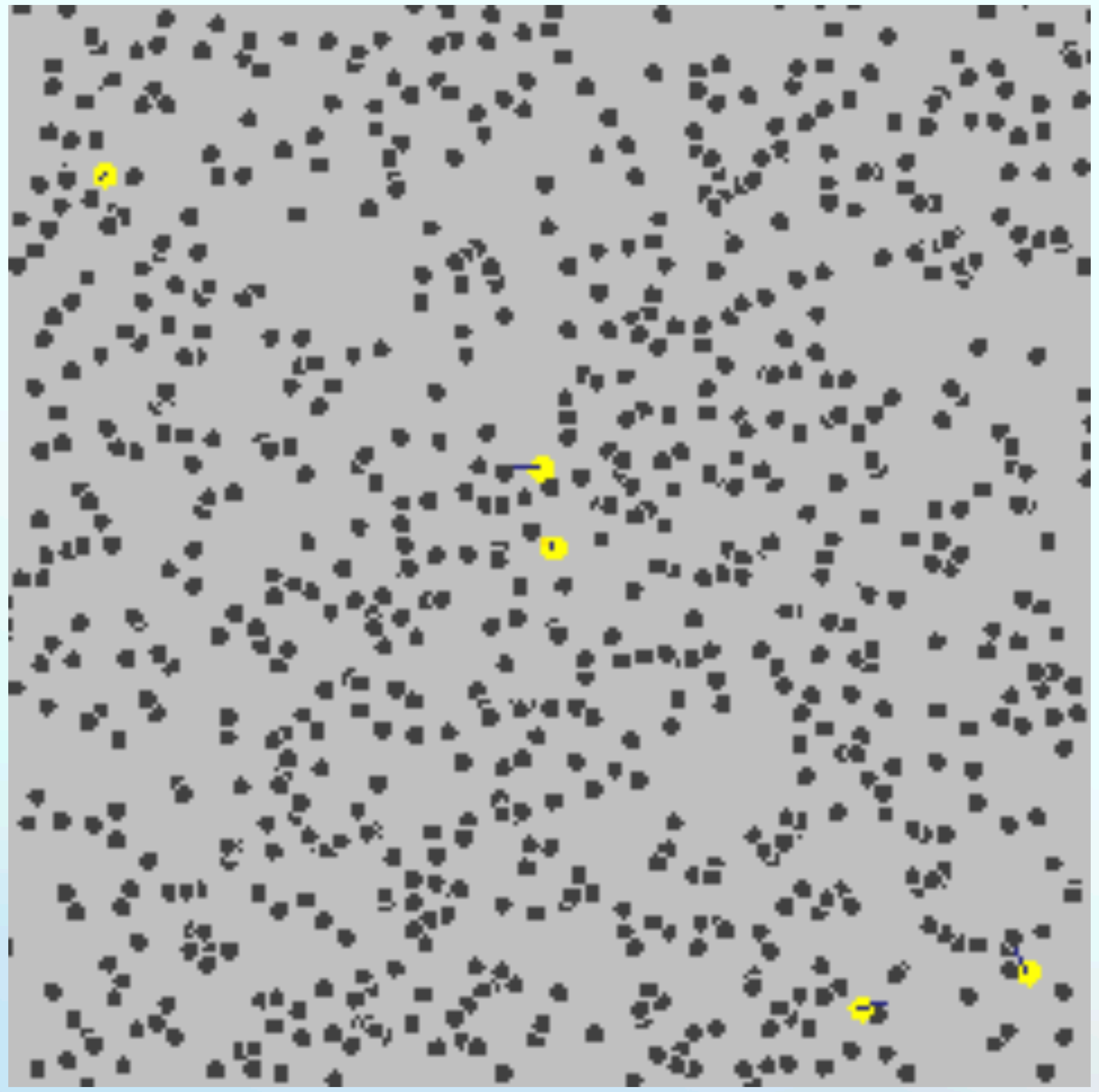
Adapted from S. Zhang, Learning non-equilibrium diffusions with Schrödinger bridges: from exactly solvable to simulation-free

Brief History

2. Brief History

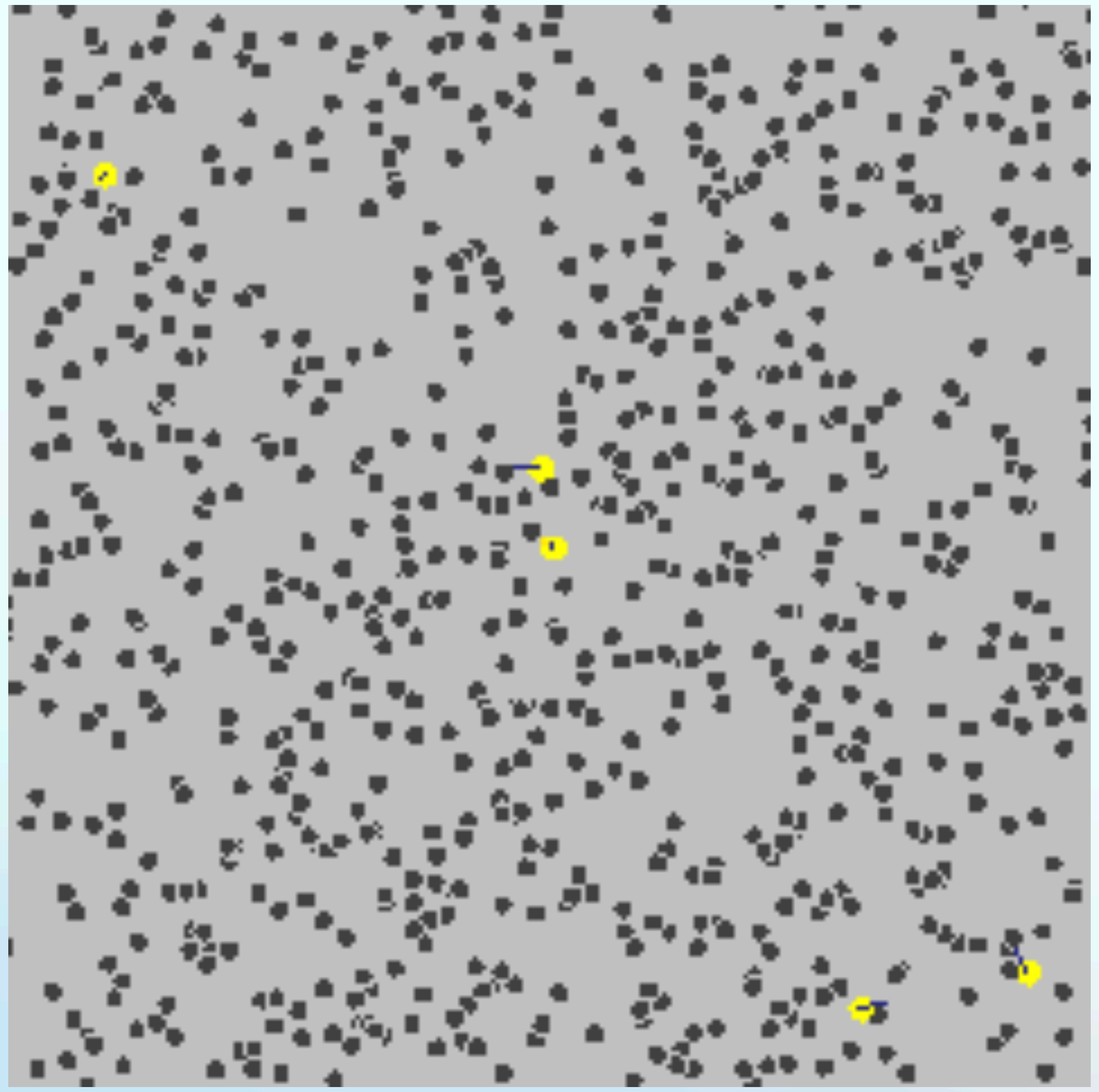
From Brownian Motion to Diffusion Models

(1827) Robert Brown + **Jean Perrin**: small pollen grains have **irregular motion** in water (becomes 'Brownian motion')



2. Brief History

From Brownian Motion to Diffusion Models

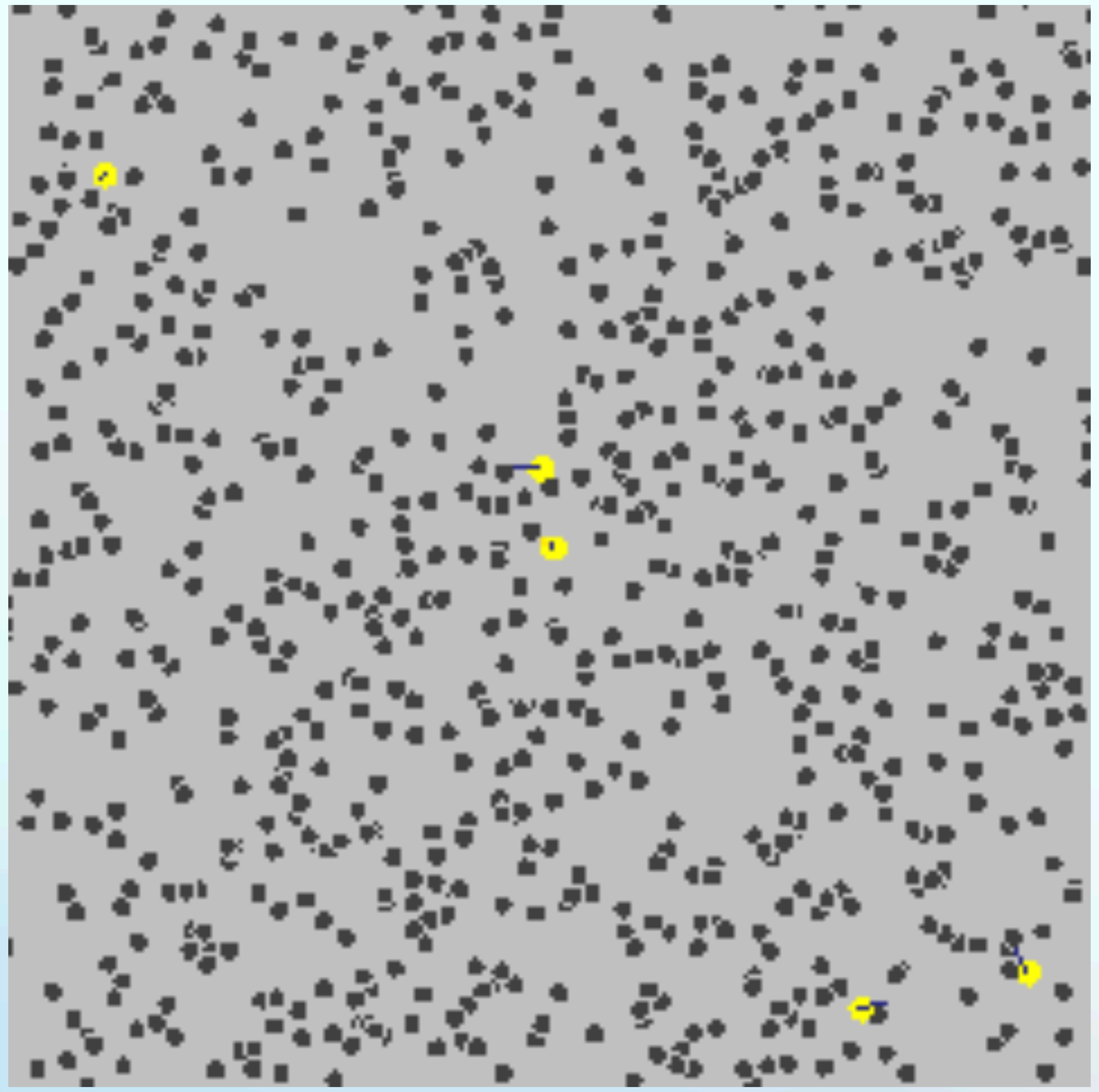


(1827) Robert Brown: small pollen grains have **irregular motion** in water (becomes 'Brownian motion')

(1905) Einstein + Smoluchowski: Motion is caused by the **impact of the water** molecules. Mean squared displacement can be related to a diffusion coefficient using a **probabilistic** approach by considering the **statistical behavior** of the molecules

2. Brief History

From Brownian Motion to Diffusion Models



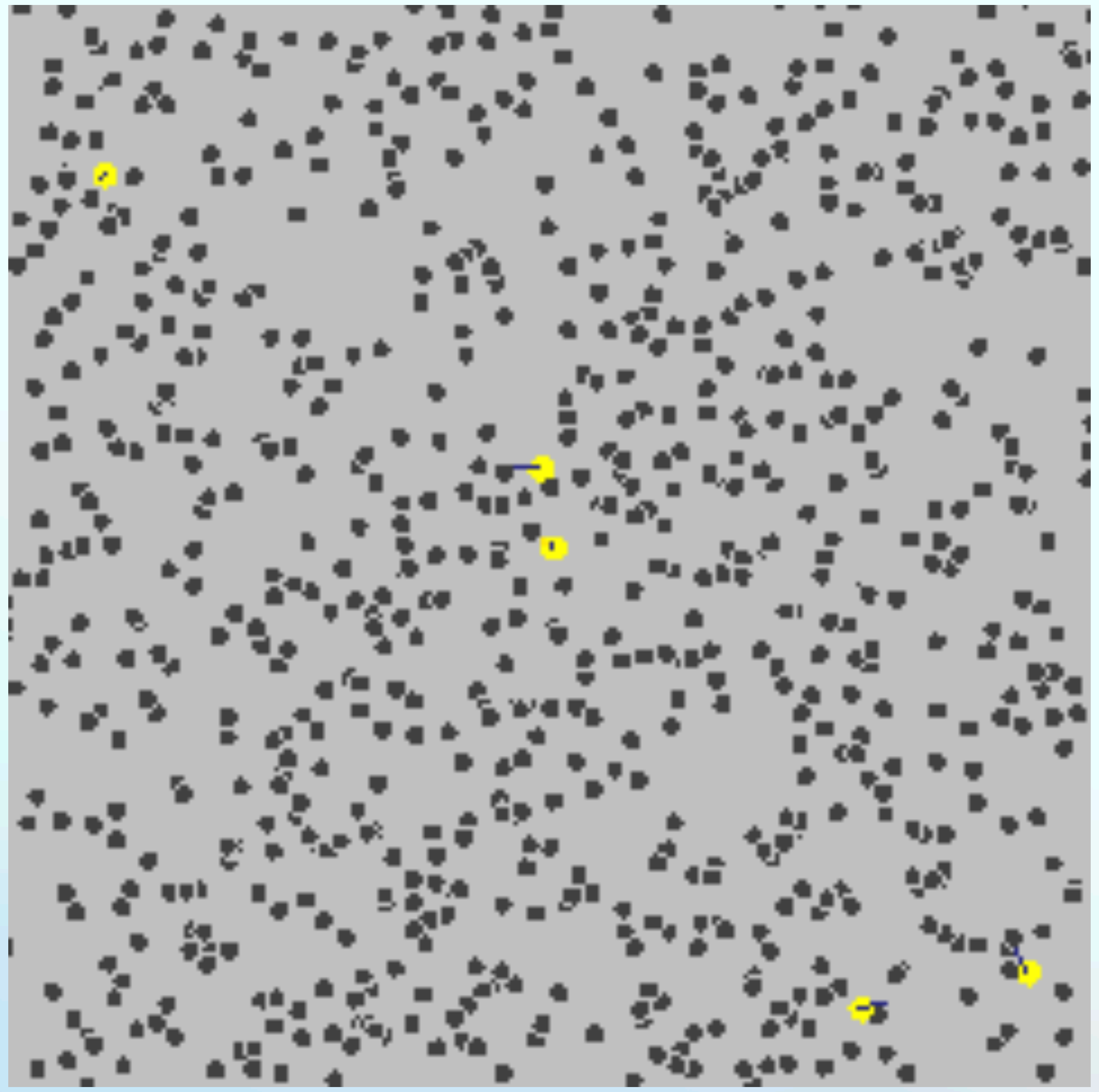
(1827) Robert Brown: small pollen grains have **irregular motion** in water (becomes 'Brownian motion')

(1905) Einstein + Smoluchowski: Motion is caused by the **impact of the water** molecules. Mean squared displacement can be related to a diffusion coefficient using a **probabilistic** approach by considering the **statistical behavior** of the molecules

(1908) Langevin: introduced a stochastic differential equation to describe Brownian motion using **Random Markovian force**

2. Brief History

From Brownian Motion to Diffusion Models



(1827) Robert Brown: small pollen grains have **irregular motion** in water (becomes 'Brownian motion')

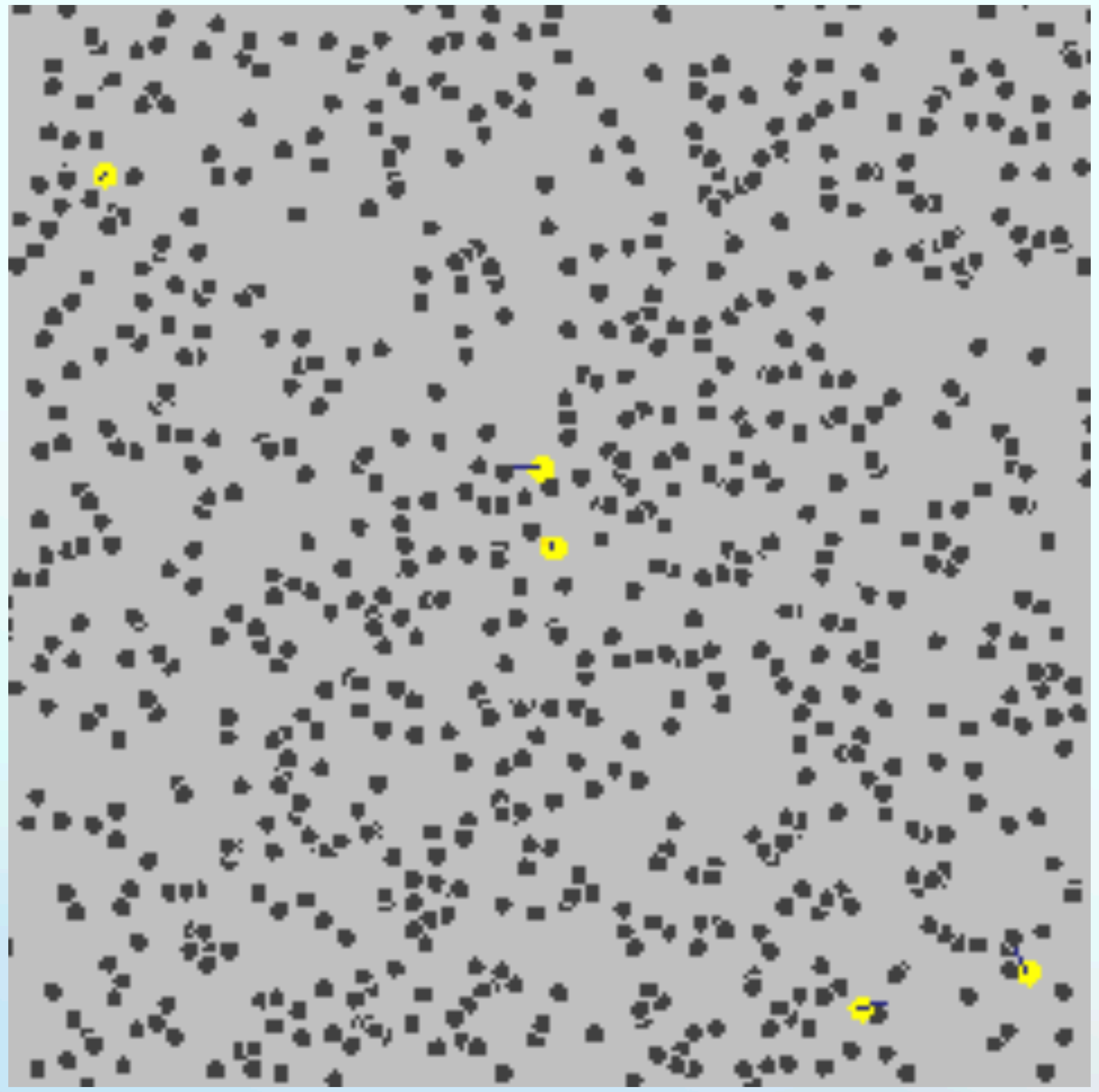
(1905) Einstein + Smoluchowski: Motion is caused by the **impact of the water** molecules. Mean squared displacement can be related to a diffusion coefficient using a **probabilistic** approach by considering the **statistical behavior** of the molecules

(1908) Langevin: introduced a stochastic differential equation to describe Brownian motion using **Random Markovian force**

(1914) Fokker & (1917) Planck: Development of **partial differential equations of probability densities**

2. Brief History

From Brownian Motion to Diffusion Models



(1827) Robert Brown: small pollen grains have **irregular motion** in water (becomes 'Brownian motion')

(1905) Einstein + Smoluchowski: Motion is caused by the **impact of the water** molecules. Mean squared displacement can be related to a diffusion coefficient using a **probabilistic** approach by considering the **statistical behavior** of the molecules

(1908) Langevin: introduced a stochastic differential equation to describe Brownian motion using **Random Markovian force**

(1914) Fokker & (1917) Planck: Development of **partial differential equations of probability densities**

(1940) Kramers & (1949) Moyal: showed a **Taylor expansion** technique to describe the **time evolution of a probability distribution**

From Langevin Dynamics to Fokker Planck

3. From Langevin Dynamics to Fokker Planck

Brownian Motion, a special case of Langevin equation

1. Assume a heavy particule in a fluid

Stokes law $F(t) = -\alpha v(t)$ $\xrightarrow{\text{Newton}}$ $m\dfrac{dv(t)}{dt} + \alpha v(t) = 0$ $\xrightarrow{\gamma = \frac{m}{v}}$ $\dfrac{dv(t)}{dt} + \gamma v(t) = 0$

3. From Langevin Dynamics to Fokker Planck

Brownian Motion, a special case of Langevin equation

1. Assume a heavy particule in a fluid

Stokes law: $F(t) = -\alpha v(t)$ $\xrightarrow{\text{Newton}}$ $m\frac{dv(t)}{dt} + \alpha v(t) = 0$ $\xrightarrow{\gamma = \frac{m}{v}}$ $\frac{dv(t)}{dt} + \gamma v(t) = 0$

2. Assume a light particule in a fluid

Stokes law: $F(t) = -\alpha v(t) + F_f(t)$ $\xrightarrow{\text{Newton}}$ $m\frac{dv(t)}{dt} + \alpha v(t) = F_f(t)$ $\xrightarrow{\begin{cases}\gamma = m/v \\ \Gamma(t) = F_f(t)/m,\end{cases}}$ $\frac{dv(t)}{dt} + \gamma v(t) = \Gamma(t)$

3. From Langevin Dynamics to Fokker Planck

Brownian Motion, a special case of Langevin equation

1. Assume a heavy particule in a fluid

$$F(t) = -\alpha v(t) \xrightarrow[\text{Newton}]{\text{Stokes law}} m\frac{dv(t)}{dt} + \alpha v(t) = 0 \xrightarrow{\gamma = \frac{m}{v}} \frac{dv(t)}{dt} + \gamma v(t) = 0$$

2. Assume a light particule in a fluid

$$F(t) = -\alpha v(t) + F_f(t) \xrightarrow[\text{Newton}]{\text{Stokes law}} m\frac{dv(t)}{dt} + \alpha v(t) = F_f(t) \xrightarrow{\begin{cases} \gamma = m/v \\ \Gamma(t) = F_f(t)/m, \end{cases}} \frac{dv(t)}{dt} + \gamma v(t) = \Gamma(t)$$

3. Brownian Motion: Linear Langevin Equation

$$\boxed{\dot{v} + \gamma v = \Gamma(t)}$$

$\Gamma(t)$:Langevin Force $\begin{cases} \mathbb{E}[\gamma(t)] = 0 \\ \mathbb{E}[\Gamma(t)\Gamma(t')] = q\delta(t - t') \end{cases}$

3. From Langevin Dynamics to Fokker Planck

General SDEs

$$\frac{d\xi(t)}{dt} = f(t,\xi) + g(t,\xi)\Gamma(t) \qquad \text{with } \Gamma(t) \sim \mathcal{N}(0,I)$$

$$\sim$$

$$d\xi(t) = f(t,\xi)dt + g(t,\xi)d\mathbf{w} \qquad \text{with } d\mathbf{w} = \Gamma(t)dt$$

3. From Langevin Dynamics to Fokker Planck

General SDEs

$$\frac{d\xi(t)}{dt} = f(t,\xi) + g(t,\xi)\Gamma(t) \qquad \text{with } \Gamma(t) \sim \mathcal{N}(0,I)$$

$$\sim$$

$$d\xi(t) = f(t,\xi)dt + g(t,\xi)d\mathbf{w} \qquad \text{with } d\mathbf{w} = \Gamma(t)dt$$

Examples:

- (Wiener Process): $\dot{x}(t) = \Gamma(t)$
- (Linear Langevin): $\dot{x}(t) + \gamma x(t) = \Gamma(t)$
- (Non Linear Langevin): $\dot{x}(t) = f(t,x) + g(t,x)d\Gamma(t)$

Solutions:

- $x(t) = \int_0^t e^{-\gamma(t-t')}\Gamma(t')dt'$
- $x(t) = x_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-t')}\Gamma(t')dt'$
- No analytical solution

3. From Langevin Dynamics to Fokker Planck

Example: Diffusion Models

Definition 4.1. Forward Diffusion.

$$d\mathbf{x} = \underbrace{\mathbf{f}(\mathbf{x}, t)}_{\text{drift}}\, dt + \underbrace{g(t)}_{\text{diffusion}}\, d\mathbf{w}. \tag{4.10}$$

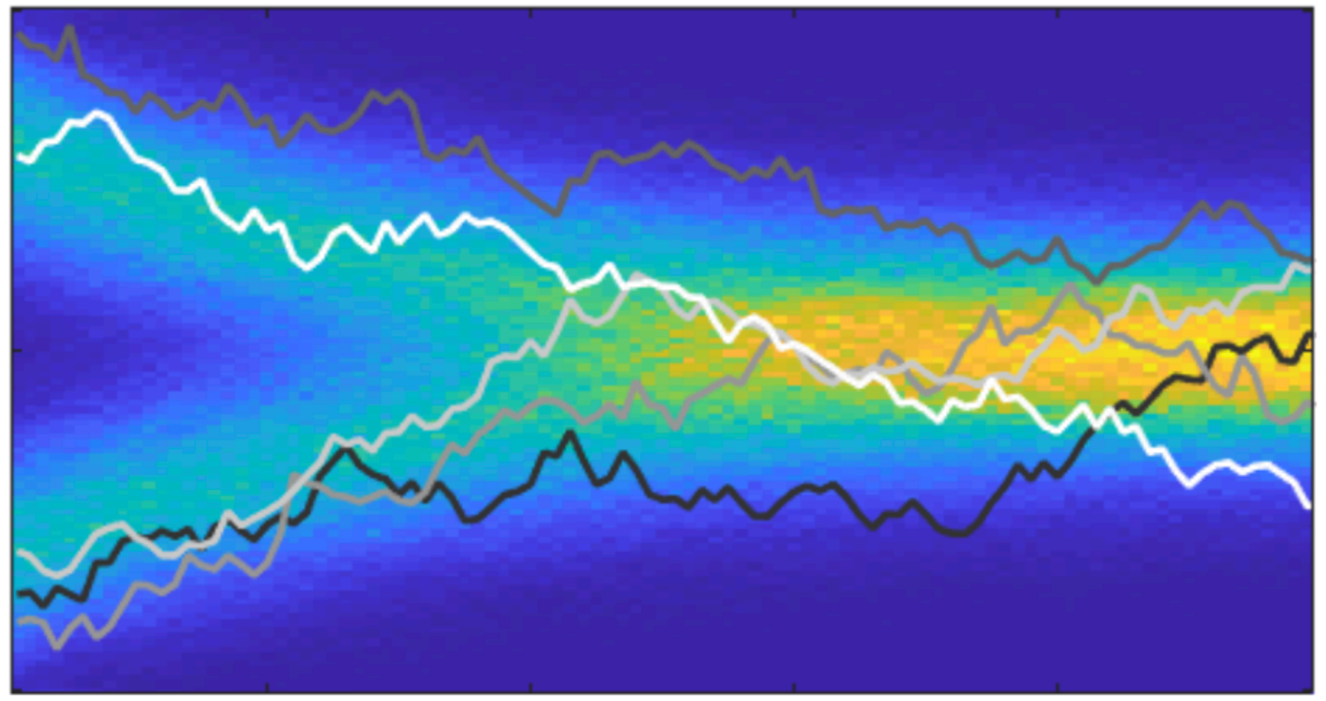
In DDPM: $\begin{cases} f(x,t) = -\dfrac{\beta_t}{2}x \\ g(t) = \sqrt{\beta_t} \end{cases}$

Definition 4.2. Reverse Diffusion.

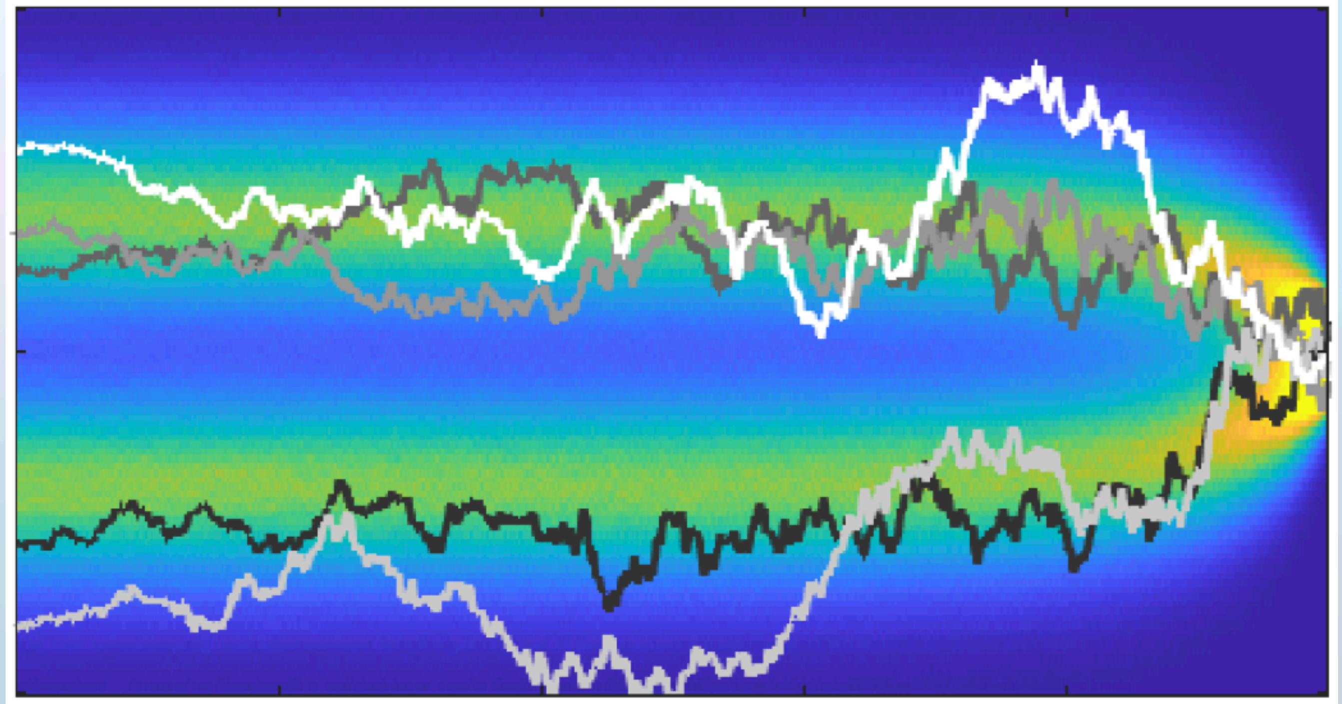
$$d\mathbf{x} = [\underbrace{\mathbf{f}(\mathbf{x}, t)}_{\text{drift}} - g(t)^2 \underbrace{\nabla_{\mathbf{x}} \log p_t(\mathbf{x})}_{\text{score function}}]\, dt \quad + \quad \underbrace{g(t) d\overline{\mathbf{w}}}_{\text{reverse-time diffusion}} \quad , \tag{4.11}$$

where $p_t(\mathbf{x})$ is the probability distribution of $\mathbf{x}$ at time t, and $\overline{\mathbf{w}}$ is the Wiener process when time flows backward.

dt > 0



dt < 0



Can we get a PDE on $p(x,t)$ for a linear and non linear Langevin Equation ?

3. From Langevin Dynamics to Fokker Planck

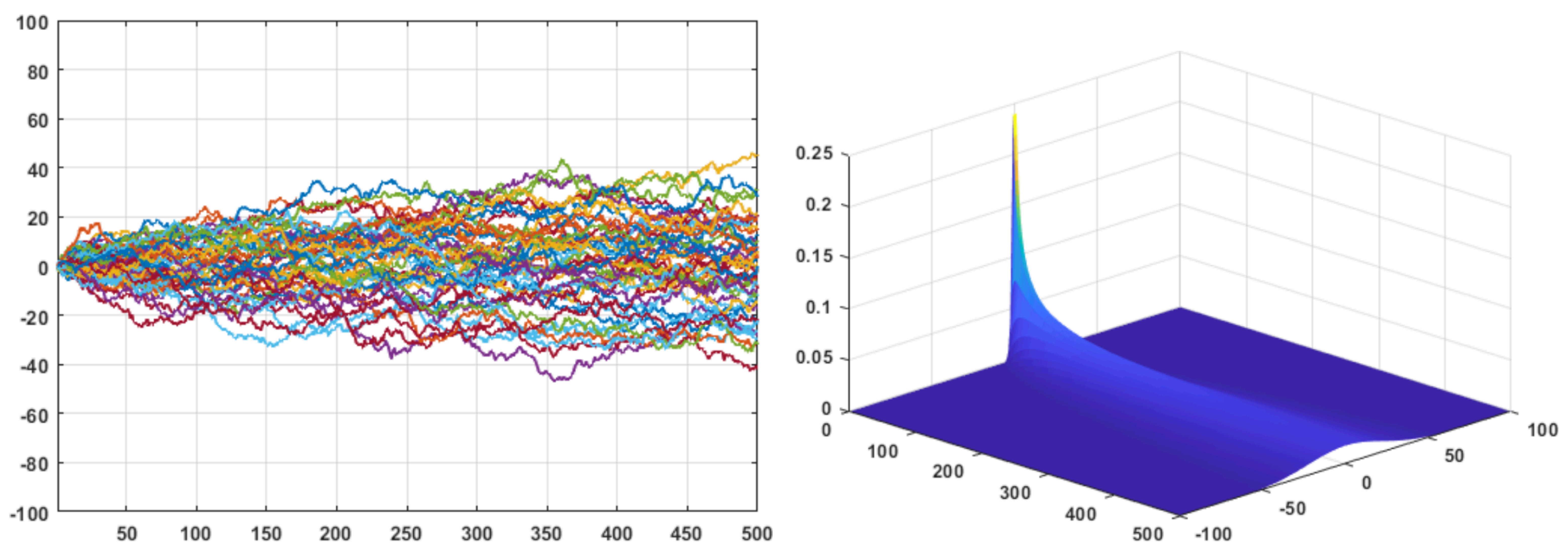
Langevin SDEs have associated ODEs on probability distributions (Example of the Wiener Process)

Theorem 5.4. Wiener Process. Consider the Wiener process

$$\dot{\xi} = \Gamma(t), \tag{5.15}$$

where $\Gamma(t)$ is the Gaussian white noise with $\mathbb{E}[\Gamma(t)] = 0$ and $\mathbb{E}[\Gamma(t)\Gamma(t')] = q\delta(t - t')$. The probability distribution $p(x, t)$ of the solution $\xi(t)$ where $\xi(t) = x$ is

$$p(x,t) = \frac{1}{\sqrt{2\pi q t}} e^{-\frac{(x-\xi_0)^2}{2qt}}. \tag{5.16}$$



(a) Realizations of $\xi(t)$ (b) $p(x, t)$

Figure 5.2: Realization of a Wiener process. (a) The random process follows the stochastic differential equation. We show a few realizations of the random process. (b) The underlying probability distribution $p(x, t)$. As t increases, the variance of the Gaussian also increases.

3. From Langevin Dynamics to Fokker Planck

Master Equation

1. Assume the **generic non linear Langevin Equation** $\dot{x}(t) = f(t, x) + g(t, x)d\Gamma(t)$

2. Assume **Markov Property** $p(x_n, t_n \,|\, \mathbf{x}_{n-1}, \mathbf{t}_{n-1}) = p(x_n, t_n \,|\, x_{n-1}, t_{n-1})$ $\begin{cases} \mathbf{x}_{n-1} = (x_{n-1}, x_{n-2}, \ldots, x_0) \\ \mathbf{t}_{n-1} = (t_{n-1}, t_{n-2}, \ldots, t_0), \end{cases}$

3. The **Masters Equation** states that:

3. From Langevin Dynamics to Fokker Planck

Master Equation

1. Assume the **generic non linear Langevin Equation** $\dot{x}(t) = f(t, x) + g(t, x)d\Gamma(t)$
2. Assume **Markov Property** $p(x_n, t_n \,|\, \mathbf{x}_{n-1}, \mathbf{t}_{n-1}) = p(x_n, t_n \,|\, x_{n-1}, t_{n-1})$ $\begin{cases} \mathbf{x}_{n-1} = (x_{n-1}, x_{n-2}, \ldots, x_0) \\ \mathbf{t}_{n-1} = (t_{n-1}, t_{n-2}, \ldots, t_0), \end{cases}$
3. The **Masters Equation** states that:

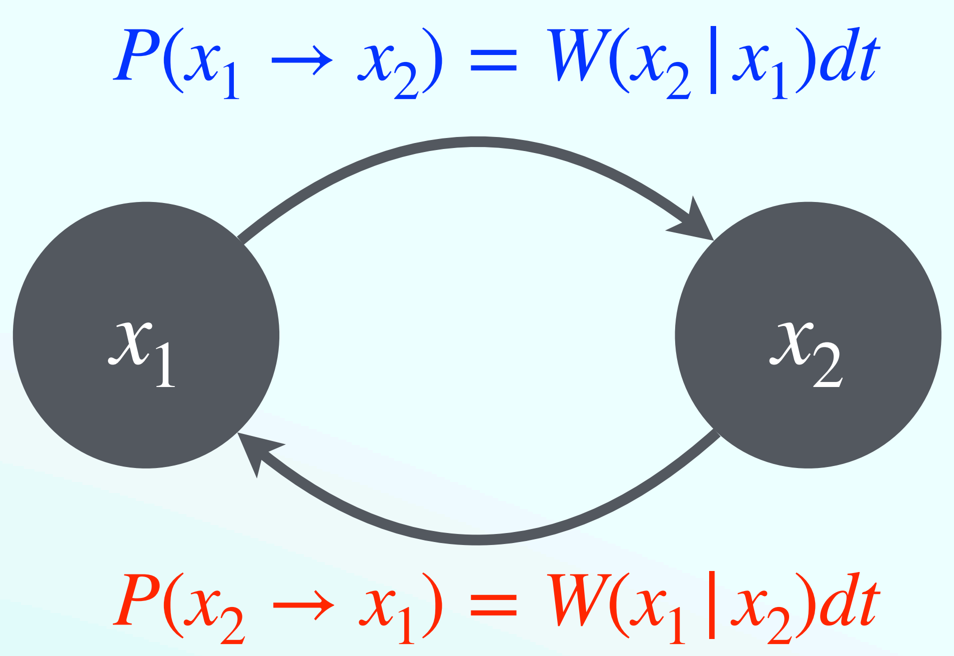
$$\underbrace{\frac{\partial}{\partial t} p(x, t)}_{\text{rate of change}} = \int \underbrace{\Big[W\big(x, t \mid x', t\big)\, p\,(x', t)\Big]}_{\text{in-flow of probability}} - \underbrace{\Big[W\big(x', t \mid x, t\big)\, p(x, t)\Big]}_{\text{out-flow of probability}} \; dx' \,.$$

With $W(x \,|\, x')$ **the probability density function per unit time** $P(x'(t + dt) \,|\, x(t)) = W(x' \,|\, x)dt$

3. From Langevin Dynamics to Fokker Planck

Example: 1D switch

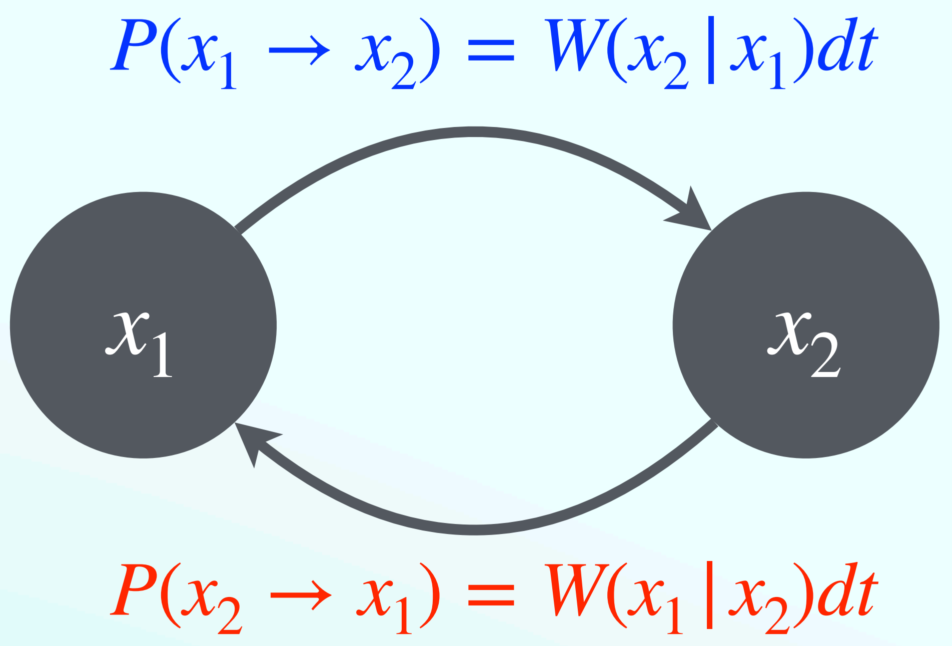
Consider two states 0 and 1 (e.g. a gene which can be either active or inactive). k_0K is the transition rate from 0 to 1 and k_1K transition rate from 1 to 0, with $k_1 + k_0 = 1$.



3. From Langevin Dynamics to Fokker Planck

Example: 1D switch

Consider two states 0 and 1 (e.g. a gene which can be either active or inactive). k_0K is the transition rate from 0 to 1 and k_1K transition rate from 1 to 0, with $k_1 + k_0 = 1$.



$$
\begin{aligned}
p(x_1, t + dt) &= p(x_1, t)\mathbb{P}\left[\text{ stay in } x_1\right] + p(x_2, t)\mathbb{P}\left[\text{ move from } x_2 \text{ to } x_1\right] \\
&= p(x_1, t)\left(1 - \mathbb{P}\left[\text{ move from } x_1 \text{ to } x_2\right]\right) + p(x_2, t)\mathbb{P}\left[\text{ move from } x_2 \text{ to } x_1\right] \\
&= p(x_1, t)\left(1 - W(x_2 | x_1)dt\right) + p(x_2, t)W(x_1 | x_2)dt \; + \; \mathcal{O}(t^2)
\end{aligned}
$$

Jumps during (t, t+dt)

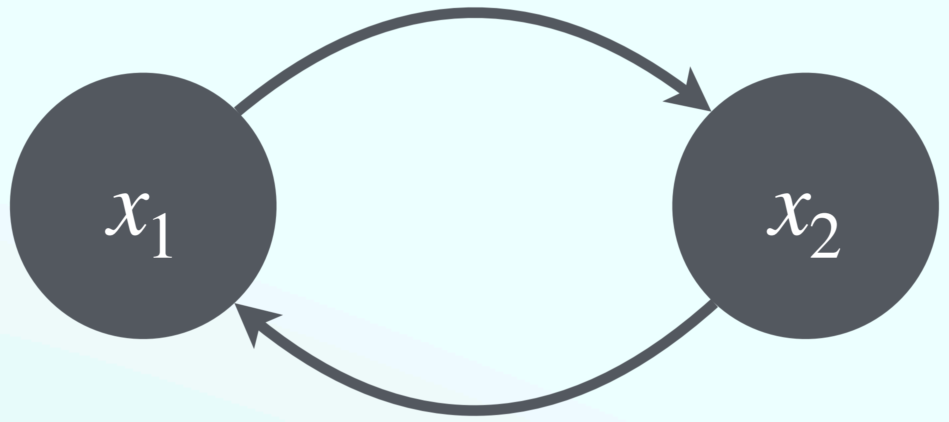
$$
\frac{dp(x_1, t)}{dt} = -W(x_2 | x_1)p(x_1, t) + W(x_1 | x_2)p(x_2, t)
$$

3. From Langevin Dynamics to Fokker Planck

Example: 1D switch

Consider two states 0 and 1 (e.g. a gene which can be either active or inactive). k_0K is the transition rate from 0 to 1 and k_1K transition rate from 1 to 0, with $k_1 + k_0 = 1$.

$P(x_1 \to x_2) = W(x_2|x_1)dt$



$P(x_2 \to x_1) = W(x_1|x_2)dt$

$$
\begin{aligned}
p(x_1, t+dt) &= p(x_1,t)\mathbb{P}\left[\text{ stay in } x_1\right] + p(x_2,t)\mathbb{P}\left[\text{ move from } x_2 \text{ to } x_1\right] \\
&= p(x_1,t)\left(1-\mathbb{P}\left[\text{ move from } x_1 \text{ to } x_2\right]\right) + p(x_2,t)\mathbb{P}\left[\text{ move from } x_2 \text{ to } x_1\right] \\
&= p(x_1,t)\left(1-W(x_2|x_1)dt\right) + p(x_2,t)W(x_1|x_2)dt \ + \ \mathcal{O}(t^2)
\end{aligned}
$$

Jumps during (t, t+dt)

$$
\frac{dp(x_1,t)}{dt} = -W(x_2|x_1)p(x_1,t) + W(x_1|x_2)p(x_2,t)
$$

$$
\begin{aligned}
\frac{dp(s,t)}{dt} &= -W(1-s|s)p(s,t) + W(s|1-s)p(1-s,t) \\
&= K\left(k_{1-s} - p(s,t)\right) \quad \text{if} \begin{cases} W(1-s|s) = k_{1-s} \\ k_s + k_{1-s} = 1 \\ p(1-s,t) + p(s,t) = 1 \end{cases}
\end{aligned}
$$

Solvable !

3. From Langevin Dynamics to Fokker Planck

3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

Langevin

$$dx(t) = f(t, x)dt + g(t, x)d\mathbf{w}$$

↓ *

Fokker Planck

$$\frac{\partial}{\partial t}p(x, t) = -\frac{\partial}{\partial x}f(x, t)p(x, t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}g^2(x, t)p(x, t)$$

* if X(t) is Markov process

3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

Fokker Planck

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial^2}{\partial x^2}g^2p$$

$$= -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial}{\partial x}\bullet\left[g^2\frac{\partial}{\partial x}logp + \frac{\partial}{\partial x}g^2\right]p$$

$$= -\frac{\partial}{\partial x}\left[f - \frac{1}{2}g^2\frac{\partial}{\partial x}logp - \frac{\partial}{\partial x}g^2\right]p$$

$$= -\frac{\partial}{\partial x}Jp$$

3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

Fokker Planck

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial^2}{\partial x^2}g^2 p$$

$$= -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial}{\partial x}\bullet\left[g^2\frac{\partial}{\partial x}logp + \frac{\partial}{\partial x}g^2\right]p$$

$$= -\frac{\partial}{\partial x}\left[f - \frac{1}{2}g^2\frac{\partial}{\partial x}logp - \frac{\partial}{\partial x}g^2\right]p$$

$$= -\frac{\partial}{\partial x}Jp$$

↓

Langevin

$$dx(t) = J(t,x)dt \quad \sim \quad dx(t) = f(t,x)dt + g(x,t)d\mathbf{w}$$

3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

Fokker Planck

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial^2}{\partial x^2}g^2p$$

$$= -\frac{\partial}{\partial x}fp + \frac{1}{2}\frac{\partial}{\partial x}\bullet\left[g^2\frac{\partial}{\partial x}logp + \frac{\partial}{\partial x}g^2\right]p$$

$$= -\frac{\partial}{\partial x}\left[f - \frac{1}{2}g^2\frac{\partial}{\partial x}logp - \cancel{\frac{\partial}{\partial x}g^2}\right]p$$

0 in many ML cases

$$= -\frac{\partial}{\partial x}Jp$$

↓

Langevin $dx(t) = J(t,x)dt \sim dx(t) = f(t,x)dt + g(x,t)d\mathbf{w}$

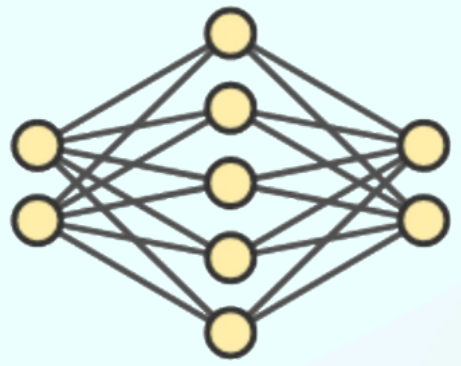
3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}\overbrace{\left[f - \frac{1}{2}g^2\underbrace{\frac{\partial}{\partial x}logp}_{s_t(x)} - \frac{\partial}{\partial x}g^2\right]}^{u_t(x)}p$$

Flow network

$u_t^\theta(x)$

$\boldsymbol{u}_t(\boldsymbol{x}) \approx$ 

Score network

$\boldsymbol{s}_t(\boldsymbol{x}) \approx$ 

$s_t^\theta(x)$

$$f^\theta(x,t) = u_t^\theta(x) + \frac{1}{2}g^2(x,t)s_t^\theta(x,t) + \frac{1}{2}\frac{\partial}{\partial x}g^2(x,t)$$

3. From Langevin Dynamics to Fokker Planck

Kramers Moyal Expansion + Pawula theorem

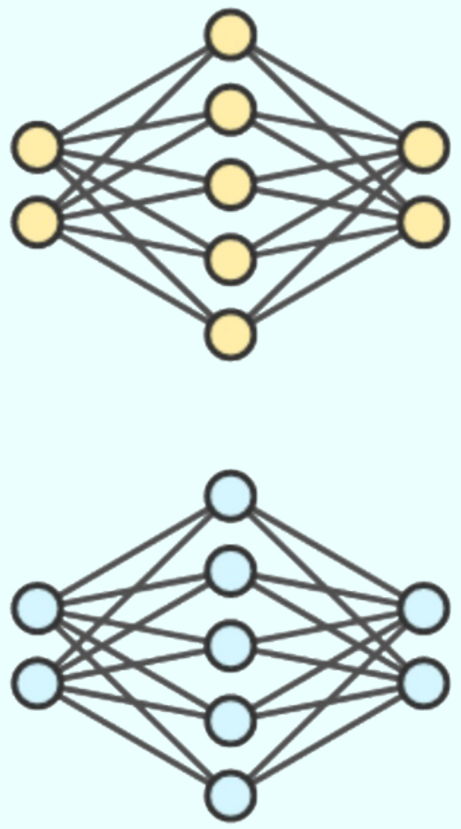
$u_t^\theta(x)$

Flow network

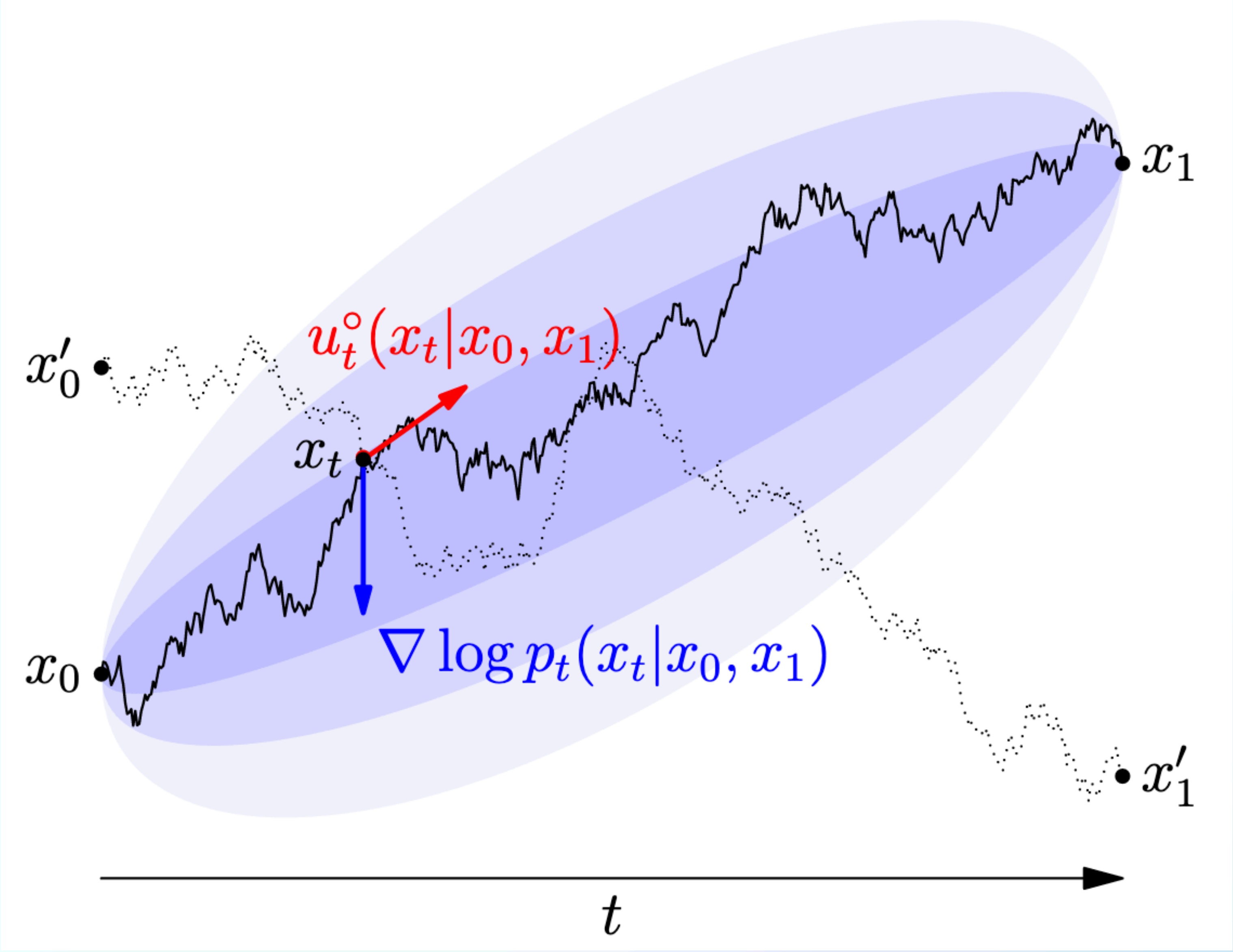
$\boldsymbol{u}_t(\boldsymbol{x}) \approx$

Score network

$\boldsymbol{s}_t(\boldsymbol{x}) \approx$



$s_t^\theta(x)$



Vincent (2011), Lipman (2022): $\begin{cases} s_t^\theta(x) \sim s_t^\theta(x \,|\, x_0, x_1) \\ u_t^\theta(x) \sim u_t^\theta(x \,|\, x_0, x_1) \end{cases}$

3. From Langevin Dynamics to Fokker Planck

Moments

$$M(x, m) = \frac{1}{m!} \lim_{\Delta t \to 0} \left[\frac{1}{\Delta t} \mathbb{E} \left[x(t + \Delta t) - x)^m \right] \Bigg|_{x(t)=x} \right]$$

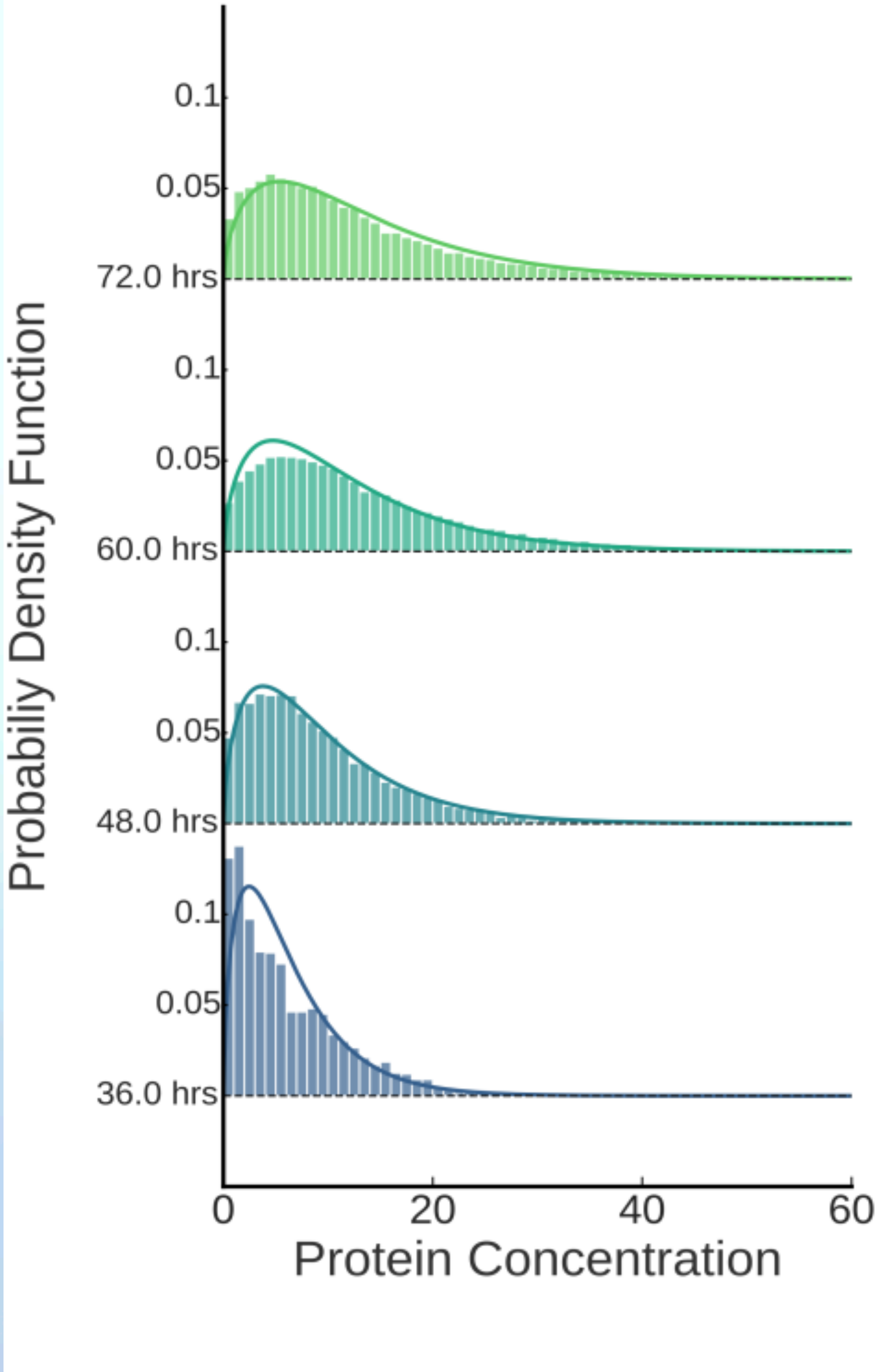
m=1 → Drift

m=2 → Diffusion

Application to current data

4. Application to current data

Kramers Moyal Expansion + Pawula theorem



Cox-Ingersoll-Ross (CIR) model:

$$dx_t = a(b - x_t)\,dt + \sigma\sqrt{x_t}\,dW_t$$

If we assume noise is $\sqrt{x_t}$, can we recover a, b ?

3. From Langevin Dynamics to Fokker Planck

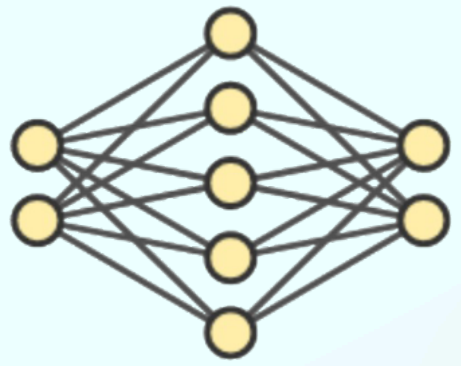
Kramers Moyal Expansion + Pawula theorem

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}\underbrace{\left[f - \frac{1}{2}g^2\underbrace{\frac{\partial}{\partial x}logp}_{s_t(x)} - \frac{\partial}{\partial x}g^2\right]}_{u_t(x)}p$$

Flow network

$u_t^\theta(x)$

$$\boldsymbol{u}_t(\boldsymbol{x}) \approx$$



Score network

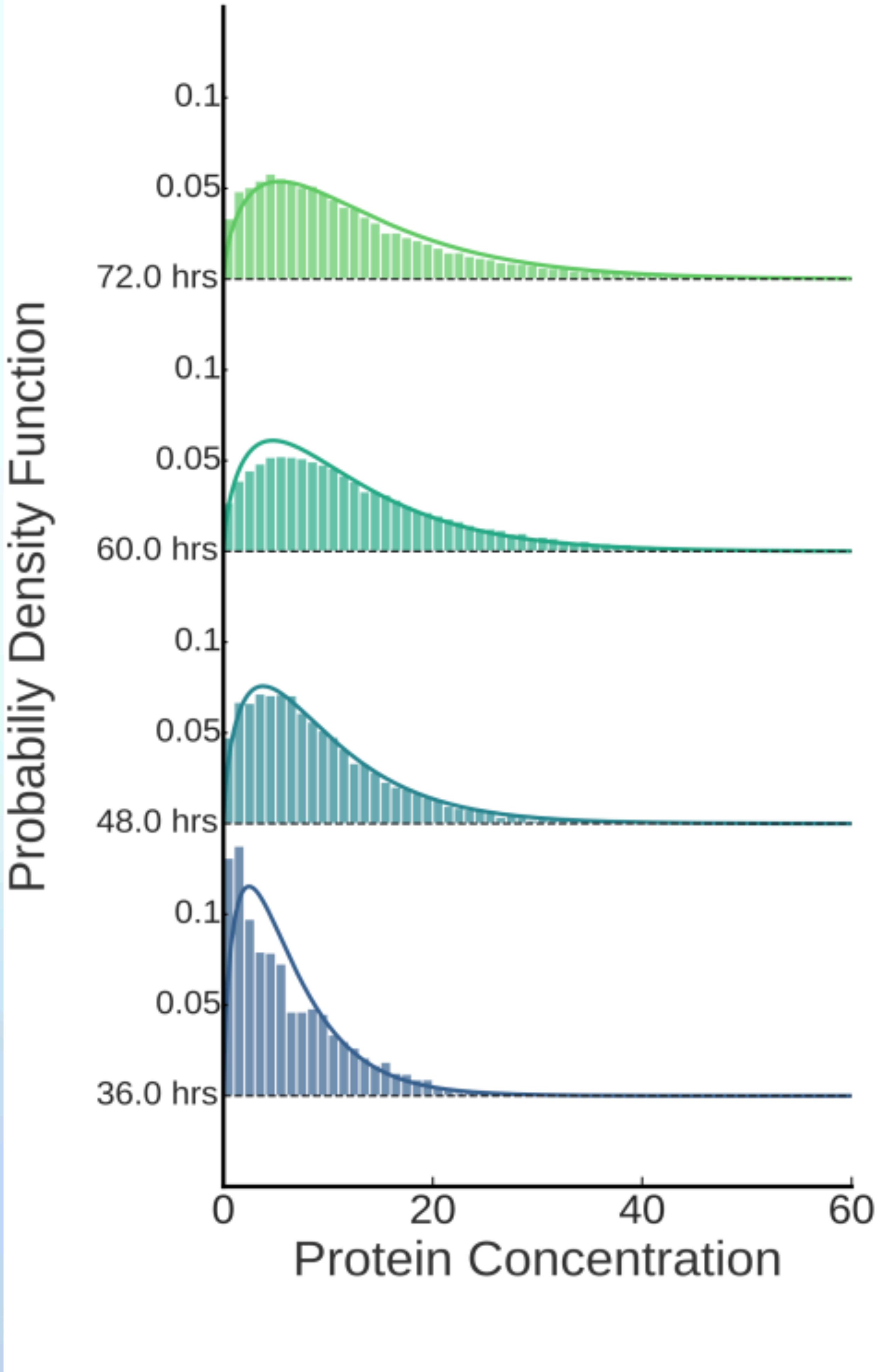
$$\boldsymbol{s}_t(\boldsymbol{x}) \approx$$

$s_t^\theta(x)$

$$f^\theta(x,t) = u_t^\theta(x) + \frac{1}{2}g^2(x,t)s_t^\theta(x,t) + \frac{1}{2}\frac{\partial}{\partial x}g^2(x,t)$$

4. Application to current data

Kramers Moyal Expansion + Pawula theorem



Cox-Ingersoll-Ross (CIR) model:

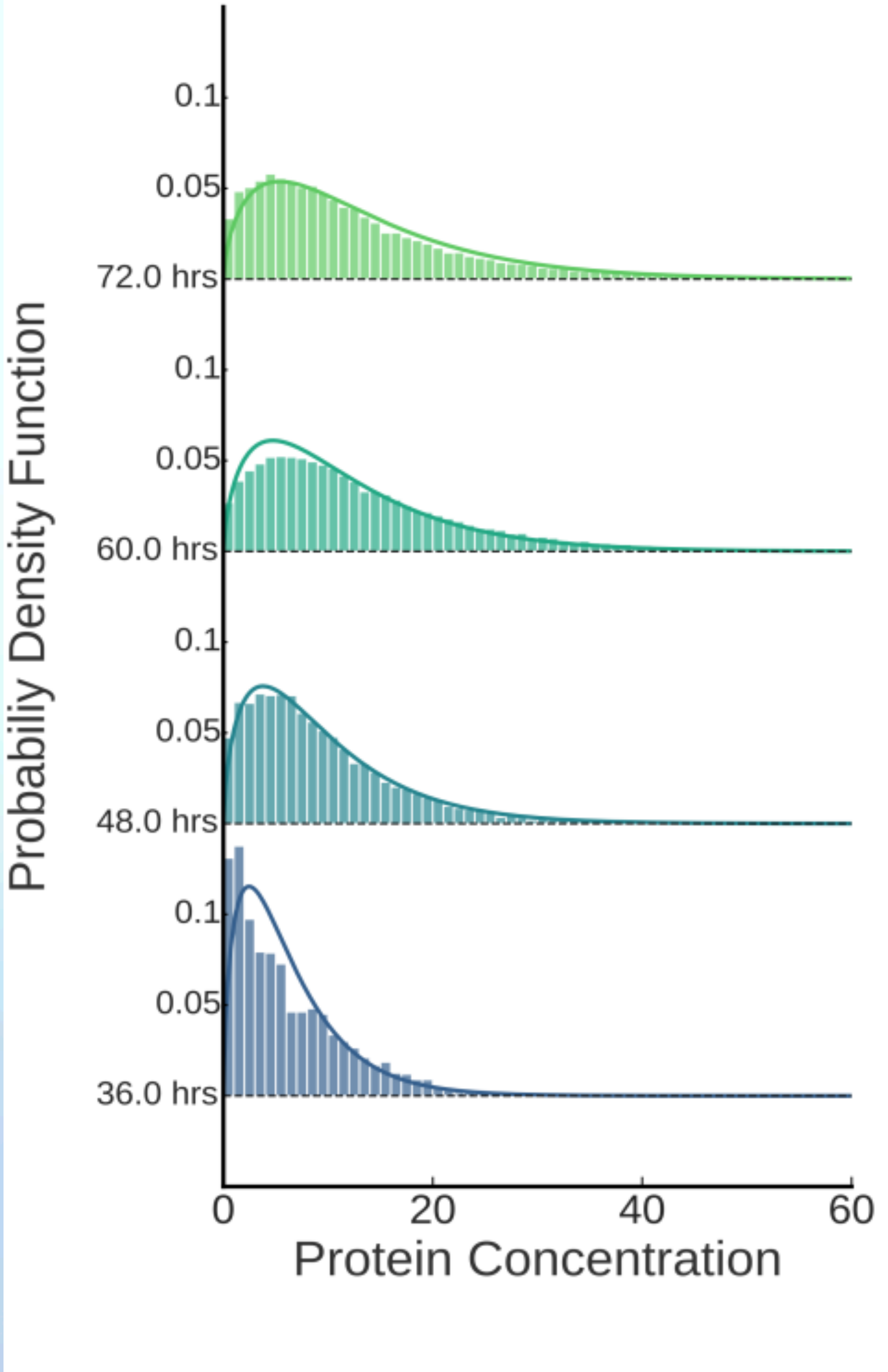
$$dx_t = a(b - x_t)\,dt + \sigma\sqrt{x_t}\,dW_t$$

If we assume noise is $\sqrt{x_t}$, can we recover a, b ?

- change of variable $u_t = \sqrt{x_t}$
- $du_t = \frac{a}{2}(\frac{b}{u_t} - u_t)\,dt + \sigma\,dW_t$
- $f^\theta(x,t) = u_t^\theta(x) + \frac{\sigma^2}{2} s_t^\theta(x,t)$

4. Application to current data

Kramers Moyal Expansion + Pawula theorem



Cox-Ingersoll-Ross (CIR) model:

$$dx_t = a(b - x_t)\,dt + \sigma\sqrt{x_t}\,dW_t$$

If we assume noise is $\sqrt{x_t}$, can we recover a, b ?

- change of variable $u_t = \sqrt{x_t}$
- $du_t = \frac{a}{2}(\frac{b}{u_t} - u_t)\,dt + \sigma\,dW_t$
- $f^\theta(x,t) = u_t^\theta(x) + \frac{\sigma^2}{2}s_t^\theta(x,t)$

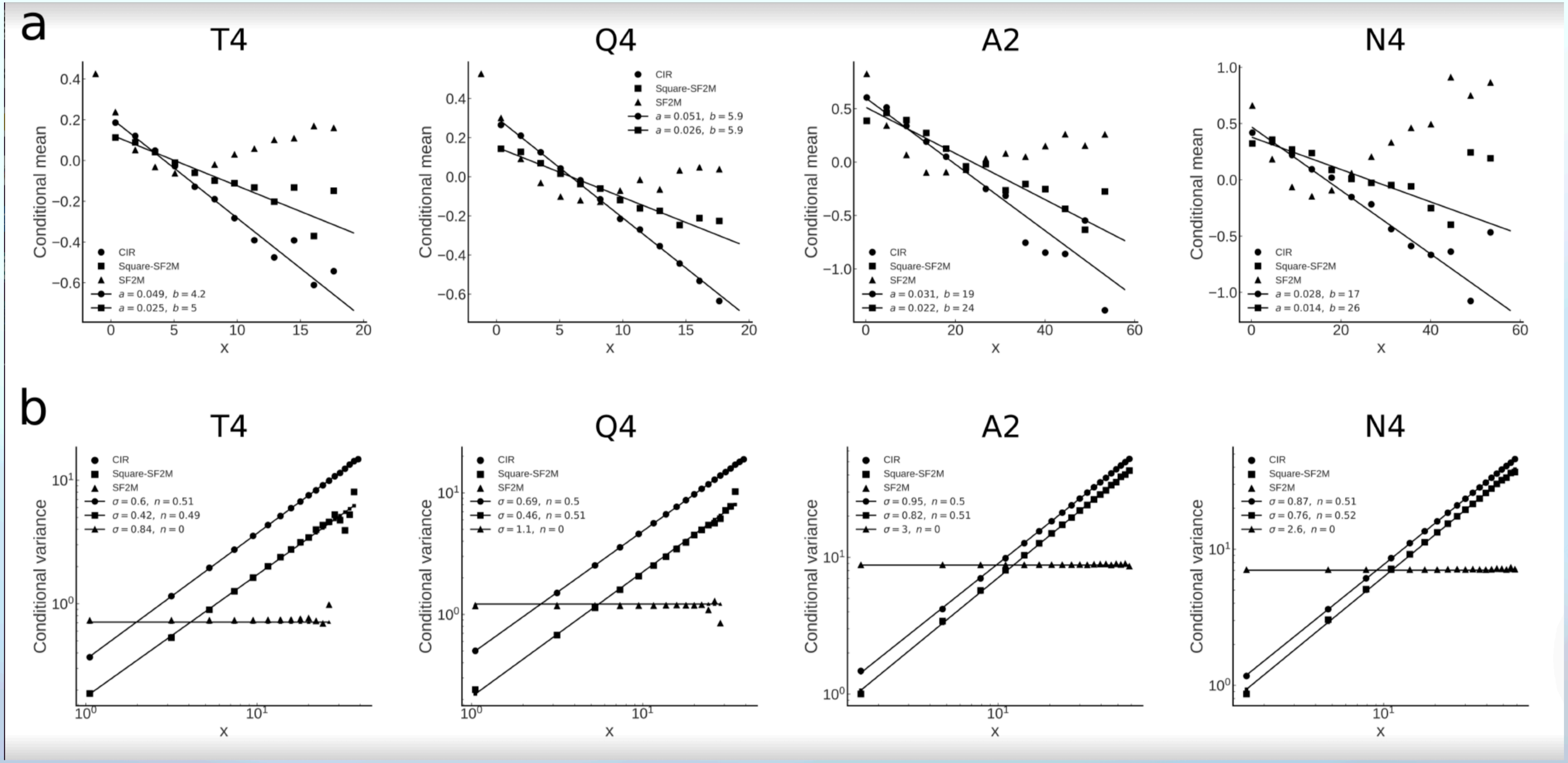
- No change of variable
- $dx_t = a(b - x_t)\,dt + \sigma\sqrt{x_t}\,dW_t$
- Normal inference
- $f^\theta(x,t) = u_t^\theta(x) + \frac{\sigma^2}{2}x_t s_t^\theta(x,t) + \frac{\sigma^4}{2}$

4. Application to current data

$$M(x, m) = \frac{1}{m!} \lim_{\Delta t \to 0} \left[\frac{1}{\Delta t} \mathbb{E} \left[x(t + \Delta t) - x \right)^m \right] \Bigg|_{x(t)=x} \Bigg]$$

m=1

m=2



Thank you