Diffusion without fluctuation connecting classical and quantum dynamics

Kenji Fukushima

The University of Tokyo

Based on the work with Yuji Hirono, Akinori Tanaka, on the forthcoming work with Jan Pawlowski, Friederike Ihssen, Renzo Kapust, and on long-term discussions with Francois Gelis

Talk Structure

$$d\Phi = [-K\nabla A - (K - D)\nabla \ln p + \nabla \cdot D]\, dt + \sqrt{2D}\, dW.$$

3-Step Generalizations

D: Stochasticity — does not change p_t for any t

SDE-ODE equivalence / Path-integral Interpretation

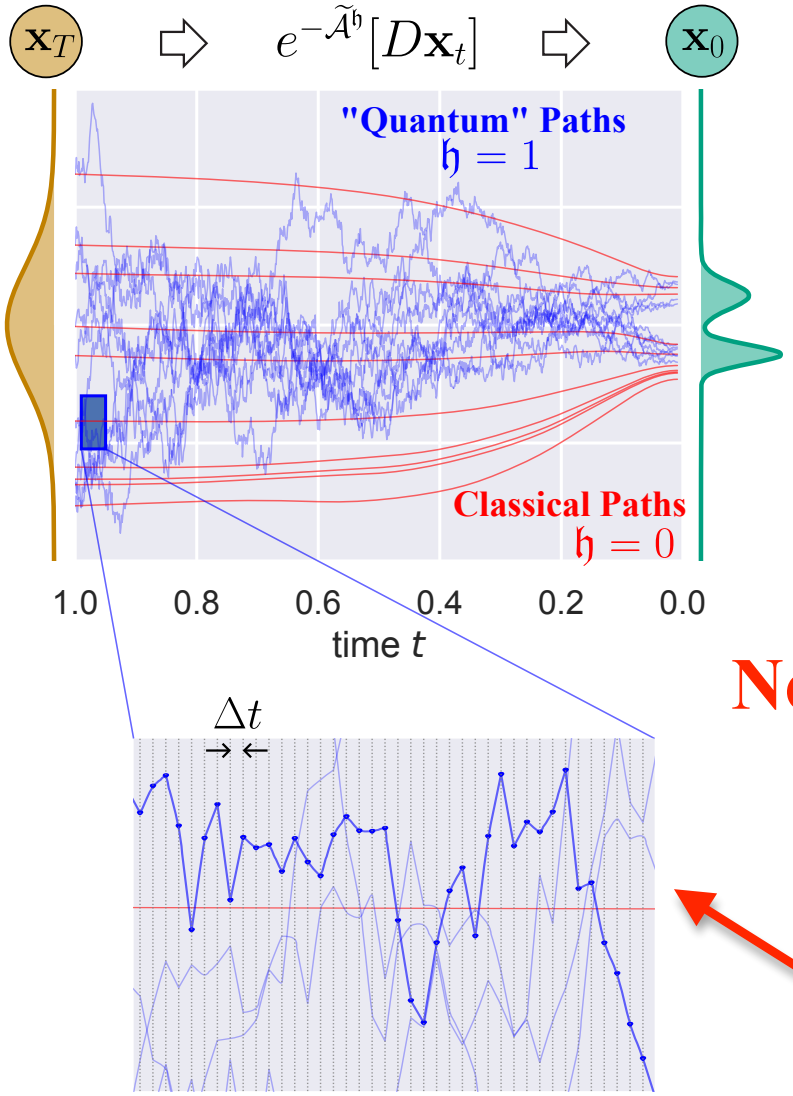
K: Relaxation Kernel — changes p_t with the same p_{eq}

Kernel optimization for the *sign problem*

A: Annealing Potential — $p_t \sim e^{-A_t}$ (quasi-static)

Physics-Informed Kernels (PIKs)

Talk Structure



Forward Process — Noising

Reverse Process — Denoising

Song et al., "*Score-based generative modeling through stochastic differential equations*" (2021)

SDE ~ Quantum-like paths
ODE ~ Classical-like paths

Noise-control Parameter ~ $\hbar$ (analogy)

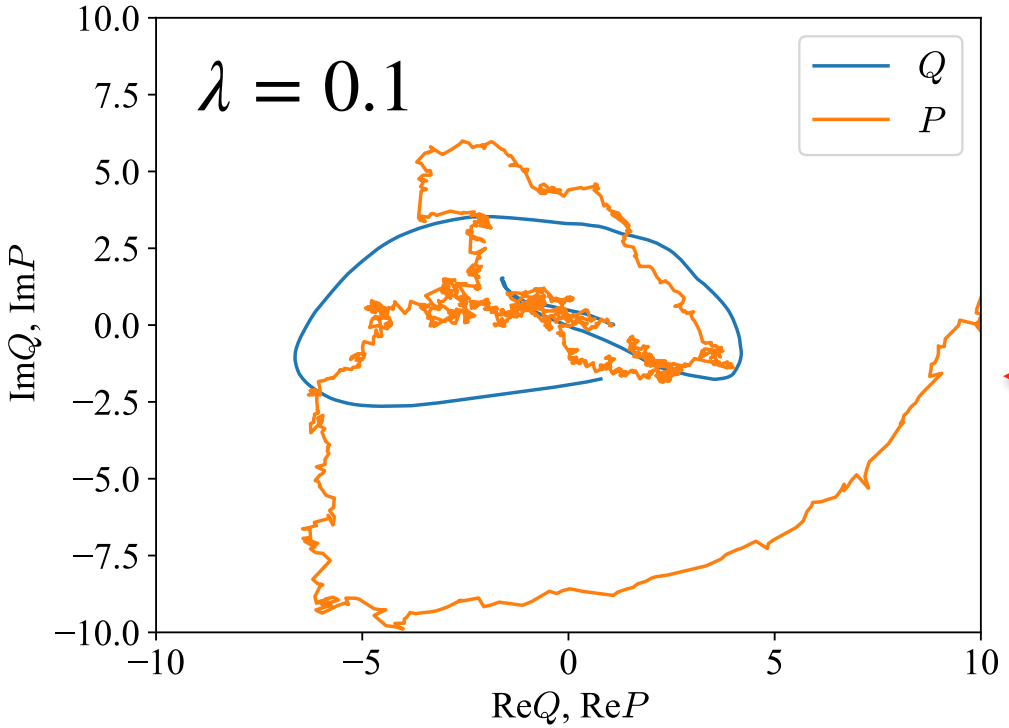
Zhang-Chen, "*Fast Sampling of Diffusion Models with Exponential Integrator*" (2023)

Hirono-Tanaka-Fukushima, "*Understanding Diffusion Models by Feynman's Path Integral*" ICML (2024)

Talk Structure

A Simple but still difficult sign problem

Real-*t* Anharmonic Osc. $H(P,Q) = \frac{P^2}{2} + \frac{Q^2}{2} + \lambda\frac{Q^4}{4}$



Closed-Time-Path Formalism (Schwinger-Keldysh)

SDE with Complex Noise

Deterministic ODE Description

SDE vs ODE

Langevin $dx = f(x,t)dt + g(t)dW_t$ Song et al. (2021)

Fokker-Planck $\partial_t p_t = -\nabla \cdot (p_t f) + \frac{1}{2} g^2 \nabla^2 p_t$

$= -\nabla \cdot (p_t v_t)$

$v_t = f - \frac{1}{2} g^2 s_t \qquad s_t = \nabla \ln p_t$ (score)

SDE vs ODE

Langevin $dx = f(x,t)dt + g(t)dW_t$ Song et al. (2021)

Fokker-Planck $\partial_t p_t = -\nabla \cdot (p_t f) + \frac{1}{2} g^2 \nabla^2 p_t$

$= -\nabla \cdot (p_t v_t)$

Probability-flow ODE

$$dx = \left[f(x,t) - \frac{1}{2} g(t)^2 \nabla \ln p_t(x) \right] dt$$

SDE vs ODE

Reverse Process

Song et al. (2021)

$$dx = \left[f - g^2 s_t\right] dt + g d\bar{W}_t$$

Associated Probability-flow ODE

$$dx = \left[f - \frac{1}{2} g^2 s_t\right] dt$$

$\mathfrak{h} = 1$

$\mathfrak{h} = 0$

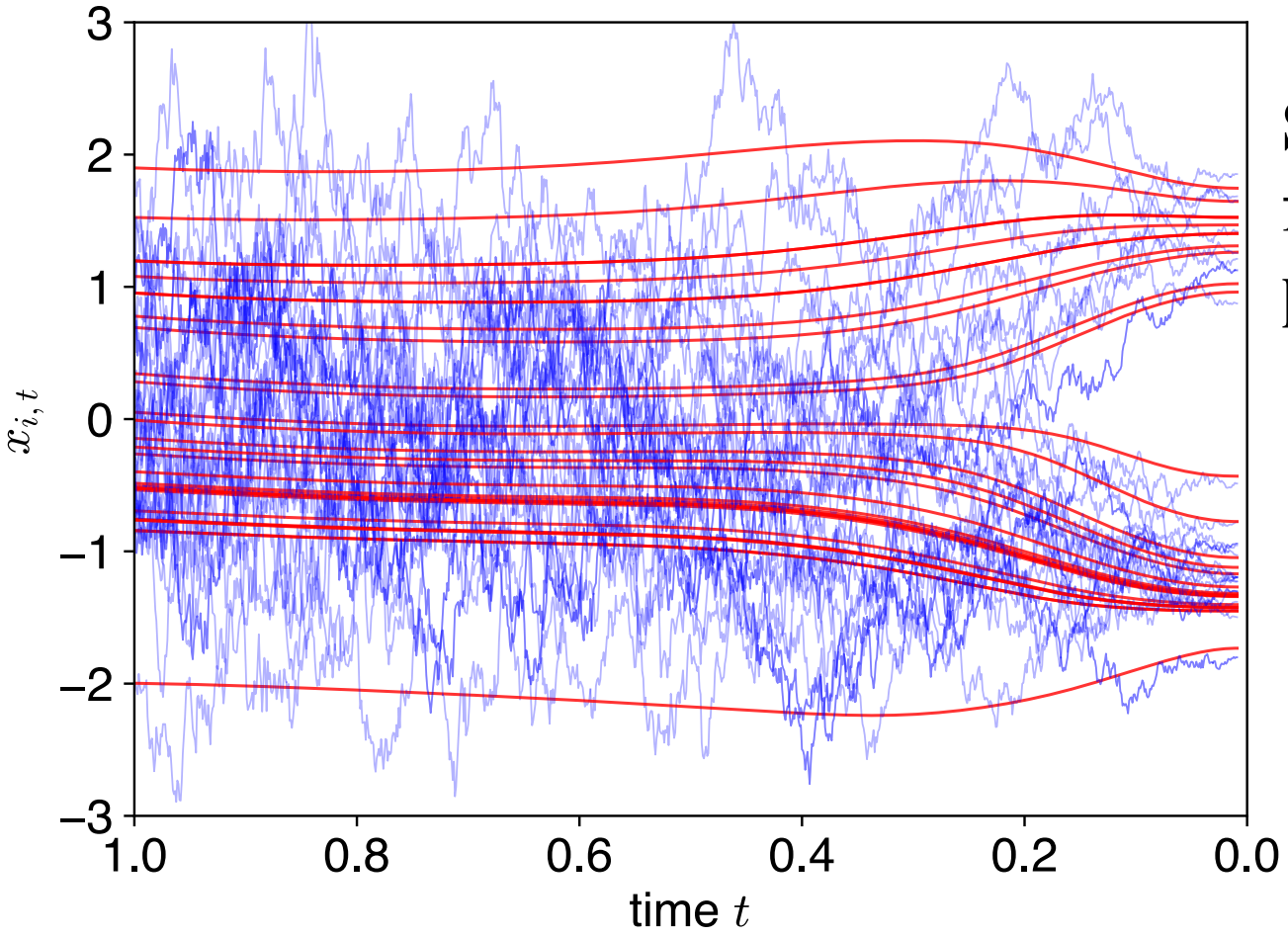
More Generally

Zhang-Chen (2023)

$$dx = \left[f - \frac{1+\mathfrak{h}}{2} g^2 s_t\right] dt + \sqrt{\mathfrak{h}}\, g d\bar{W}_t$$

Toward the Path Integral

Trajectories look very different, yet the same p_t



Similar to quantum fluctuations in the path-integral form?

$$\mathfrak{h} \sim \hbar \, ?$$

Toward the Path Integral

Roughly Speaking…

$$dx = \left[f - \frac{1+\hbar}{2} g^2 s_t \right] dt + \sqrt{\hbar} g d\bar{W}_t$$

Nicolai map

$$\int \mathscr{D}\bar{W}\, e^{-\frac{1}{2}\int dt \bar{W}_t^2}$$

This form allows for the semiclassical (WKB) approximation

$$\int \mathscr{D}x\, J\, \exp\left[- \int dt \frac{\| \dot{x} - f + \frac{1+\hbar}{2} g^2 s_t \|^2}{2\hbar g^2} \right]$$

Why WKB useful?

For numerical experiments, please see the paper…

Hirono-Tanaka-Fukushima (2024)

At the time, comparing the negative log likelihood (NLL) of SDE / ODE samplers was of considerable interest.

IF the learned score is exact, the two are equivalent, but in practical, the learned score is only approximate.

For $s_t \neq \nabla \ln p_t$, which of SDE / ODE generates data closer to p_t ?

But… NLL requires heavy computation… ← WKB !

Why WKB useful?

Application of the WKB approx. Hirono-Tanaka-Fukushima (2024)

Theorem 4.1. *The log-likelihood for the process (9) satisfies*

$$\log q_0^{\hbar}(\mathbf{x}_0) = \log \pi(\mathbf{x}_T) + \int_0^T \mathrm{d}t\, \nabla \cdot \left(\boldsymbol{f}_\theta^{\mathrm{PF}}(\mathbf{x}_t, t) - \frac{\hbar g(t)^2}{2} [\boldsymbol{s}_\theta(\mathbf{x}_t, t) - \nabla \log q_t^{\hbar}(\mathbf{x}_t)] \right), \tag{14}$$

where $\pi(\cdot)$ is a prior and $\mathbf{x}_t$ is the solution for the modified probability flow ODE,

$$\dot{\mathbf{x}}_t = \boldsymbol{f}_\theta^{\mathrm{PF}}(\mathbf{x}_t, t) - \frac{\hbar g(t)^2}{2} [\boldsymbol{s}_\theta(\mathbf{x}_t, t) - \nabla \log q_t^{\hbar}(\mathbf{x}_t)], \tag{15}$$

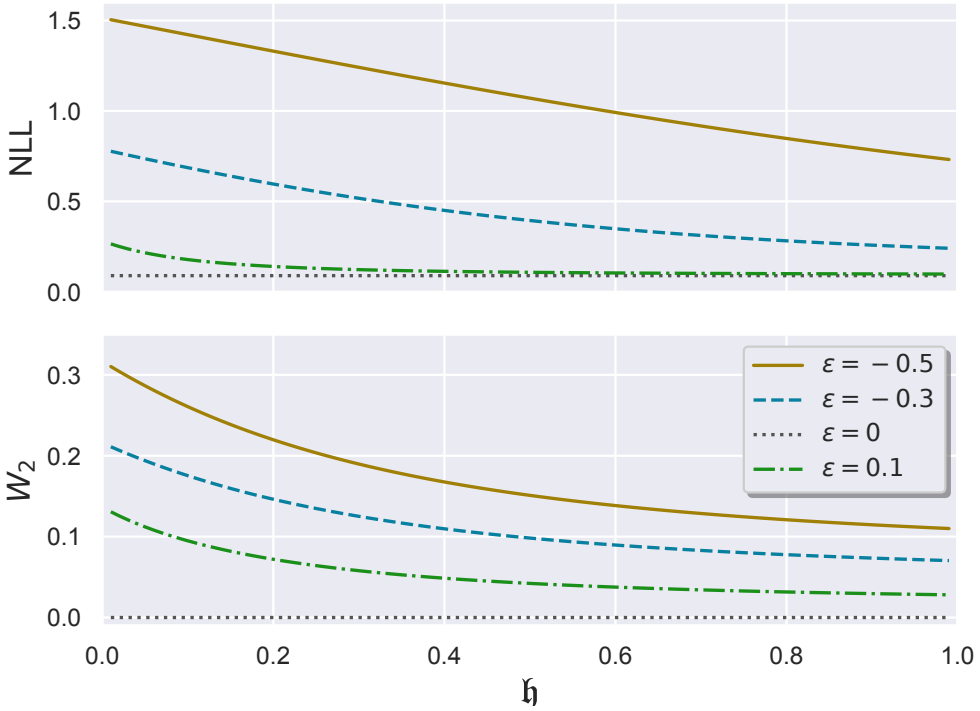
with initial condition $\mathbf{x}_{t=0} = \mathbf{x}_0$.

Why WKB useful?

Application of the WKB approx.

Hirono-Tanaka-Fukushima (2024)

Solvable 1D Gaussian Data

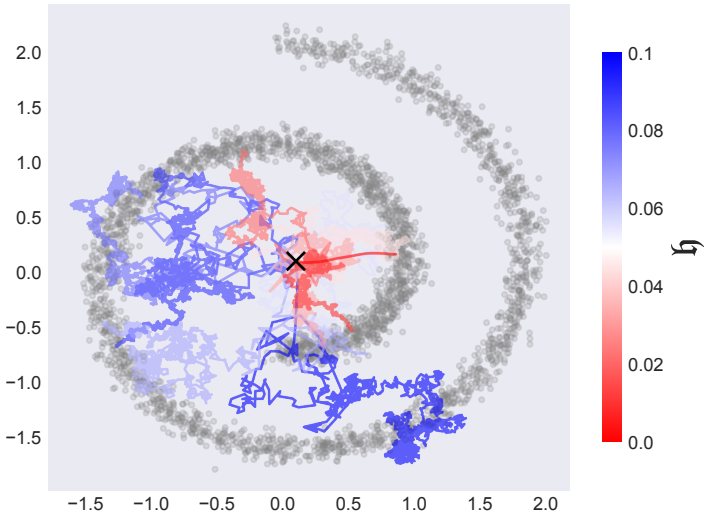


ε: deviation from exact score

2D Synthetic Data

Negative corr.!

		Swiss-roll	
SDE (NLL)	tol	1st-corr	errors
simple (1.39± 0.05)	1e−3	-0.31±0.21	0.13±0.00
	1e−5	-0.44±0.38	0.13±0.00
cosine (1.42± 0.02)	1e−3	-1.59±0.57	0.35±0.00
	1e−5	-3.27±1.11	0.37±0.02



Another Degree of Freedom in SDE

Stochastic Quantization **Damgaard-Hüffel (1987)**

It is possible to consider a more general Langevin equation by introducing a kernel K (see also sections 6 and 10),

$$\dot{\phi}(x,t) = -\int K(x,y)\,\frac{\delta S_{\mathrm{E}}}{\delta\phi(y)}\,\mathrm{d}^n y + \eta(x,t)\,. \tag{3.22}$$

Provided that the noise correlations are changed as well to

$$\langle \eta(x,t)\,\eta(x',t')\rangle_\eta = 2\,K(x,x')\,\delta(t-t')\,, \qquad \text{etc.} \tag{3.23}$$

Kernel choice related to SDE/ODE interpolation?

Answer — No, but yes…

Another Degree of Freedom in SDE

Stochastic Quantization **Damgaard-Hüffel (1987)**

$$d\Phi = -\nabla S\, d\tau + \sqrt{2}\, dW \qquad \langle dW_i dW_j \rangle = \delta_{ij}\, d\tau$$

$$d\Phi_i = \left[-K_{ij}\partial_j S + \partial_j K_{ij}\right] d\tau + \sqrt{2K}_{\,ia}\, dW_a$$

Then, the FP eq. is changed unlike SDE/ODE equivalence:

$$\partial_\tau p = \partial_i \left[K_{ij}\left(\partial_j S\, p + \partial_j p\right)\right]$$

Still, the target distribution is the same:

$$p_{\rm eq}(\Phi) \propto e^{-S(\Phi)}$$

Another Degree of Freedom in SDE

Does it change anything? → **Yes, for the sign problem!**

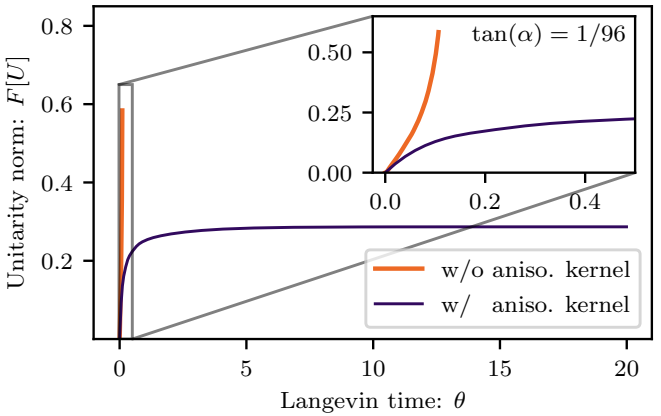
An example from the real-time quantum dynamics

Boguslavski-Hotzy-Müller (2023)

$$U_\mu(x,\theta+\epsilon) = e^{it^a\left[i\epsilon\Gamma_\mu(x)\,\frac{\delta S_{\rm w}}{\delta A^a_\mu(x)}\Big|_\theta + \sqrt{\epsilon\Gamma_\mu(x)}\,\eta^a_\mu(x,\theta)\right]} U_\mu(x,\theta), \tag{5}$$

where ϵ represents the Langevin time step and $\Gamma_\mu(x)$ is the kernel. The Gaussian noise field satisfies

$$\langle \eta^a_\mu(x,\theta)\rangle = 0, \quad \langle \eta^a_\mu(x,\theta)\eta^b_\nu(y,\theta')\rangle = 2\delta_{\theta\theta'}\delta_{xy}\delta^{ab}\delta_{\mu\nu}. \tag{6}$$



An anisotropic kernel can dramatically stabilize the simulation.

Two-Kernel Generalization

$$d\Phi = [-K\nabla S - (K - D)\nabla \ln p + \nabla \cdot D]\, d\tau + \sqrt{2D}\, dW.$$

$$\partial_\tau p = \nabla \cdot [K\,(\nabla S\, p + \nabla p)]$$

$D = K$ **Kernel formulation in Damgaard-Hüffel**

Kernel optimization to suppress the sign problem Alverstad, Larsen, Rothkopf, Lampl, Sexty, …

$D = \mathfrak{h}K$ **SDE/ODE interpolation**

Complex Langevin has runaway problems that may be cured? → **Application**

Three-Component Generalization

$$d\Phi = [-K\nabla A - (K - D)\nabla \ln p + \nabla \cdot D]\, dt + \sqrt{2D}\, dW.$$

A_t : Annealing Schedule

$$\partial_t p_t = \nabla \cdot [K_t (\nabla A_t\, p_t + \nabla p_t)]\,.$$

If we solve this, $p_{\rm eq}$ may not be achieved correctly…

BUT, we already know $p_{\rm eq}$, so we can just set p_t by hand.

After this reinterpretation, the FK eq. is the condition to solve K_t and A_t to satisfy this given p_t .

$$p_t \quad \Rightarrow \quad (A_t, K_t)$$

Three-Component Generalization

$$\boxed{\partial_t p_t = \nabla \cdot \left[K_t \left(\nabla A_t \, p_t + \nabla p_t \right) \right].}$$

$$\Leftrightarrow$$

$$\partial_t p_t = -\nabla \cdot (p_t \dot{\Phi})$$

$$\dot{\Phi} = K_t \nabla (S_t - A_t)$$

$$+$$

$$p_t = e^{-S_t}$$

$$\Rightarrow \partial_t S_t + [K_t \nabla (S_t - \cancel{A_t})] \cdot \nabla S_t = \nabla \cdot [K_t \nabla (S_t - \cancel{A_t})]$$

Is A_t essential? NO, removable

Wegner equation in PIKs (if K_t is eliminated)

Tomorrow's talks by Pawlowski, Kapust, Ihssen

We can generalize PIKs with noise (with the same K_t):

$$d\Phi = \left[(K_t - D_t) \nabla S_t + \nabla \cdot D_t \right] dt + \sqrt{2 D_t} dW_t$$

In this way, PIKs can be related to Complex Langevin.

Application to the Sign Problem

How to solve the real-time quantum dynamics:

$$\langle \mathcal{O}[\phi] \rangle = \int \mathcal{D}\phi_{0\pm}\, \rho[\phi_{0\pm}]$$

$$\times \int_{\phi_\pm(t_0)=\phi_{0\pm}} \mathcal{D}\phi_\pm \int \mathcal{D}\pi_\pm\, \mathrm{e}^{i\int \pi_+\dot{\phi}_+ - \pi_-\dot{\phi}_- - \mathcal{H}(\phi_\pm,\pi_\pm)}\, \mathcal{O}[\phi_+]$$

Classical Fields	**Quantum Fields**
$\varphi \equiv \frac{\phi_+ + \phi_-}{2}$	$\sigma \equiv \phi_+ - \phi_-$
$\Pi_\varphi \equiv \frac{\pi_+ + \pi_-}{2}$	$\Pi_\sigma \equiv \pi_+ - \pi_-$

Application to the Sign Problem

For the anharmonic oscillator:

$$H(P,Q) = \frac{P^2}{2} + \underbrace{\frac{Q^2}{2} + \lambda\frac{Q^4}{4}}_{V(Q)} \qquad q_n^{\pm} \equiv Q_n \pm \frac{\sigma_n}{2}, \quad p_n^{\pm} \equiv P_n \pm \frac{\pi_n}{2}$$

π_n-integration: $Q_n - Q_{n-1} - \Delta P_n = 0$

σ_n-integration: $\int d\sigma_n \, e^{i\sigma_n (P_n - P_{n+1} - \Delta V'(Q_n) - \frac{\lambda}{4}\Delta Q_n \sigma_n^2)}$

This integral is not analytically tractable for a general potential. For the quartic potential, it reduces to an Airy function.

Application to the Sign Problem

So far, no approximation at all:

$$\int d\sigma_n \, e^{i\sigma_n (P_n - P_{n+1} - \Delta V'(Q_n) - \frac{\lambda}{4}\Delta Q_n \sigma_n^2)}$$

This is the pain in the neck…

$$= \int d\zeta_n d\eta_n \, e^{-i\eta_n \zeta_n} \, \delta\left(P_n - P_{n+1} - \Delta V'(Q_n) - \frac{\lambda}{4}\Delta Q_n \zeta_n^2 + \eta_n\right)$$

If this term is neglected somehow, the classical-statistical approx. is reproduced.

$\sim O(\hbar)$

Quantum dynamics is approximated by an ensemble of classical trajectories with fluctuating initial conditions.

Application to the Sign Problem

Let's *not* introduce any approximation:

$$Q_n = Q_{n-1} + \Delta P_n$$

$$P_{n+1} = P_n - \Delta V'(Q_n) - \frac{\lambda}{4}\Delta Q_n \zeta_n^2 + \eta_n$$

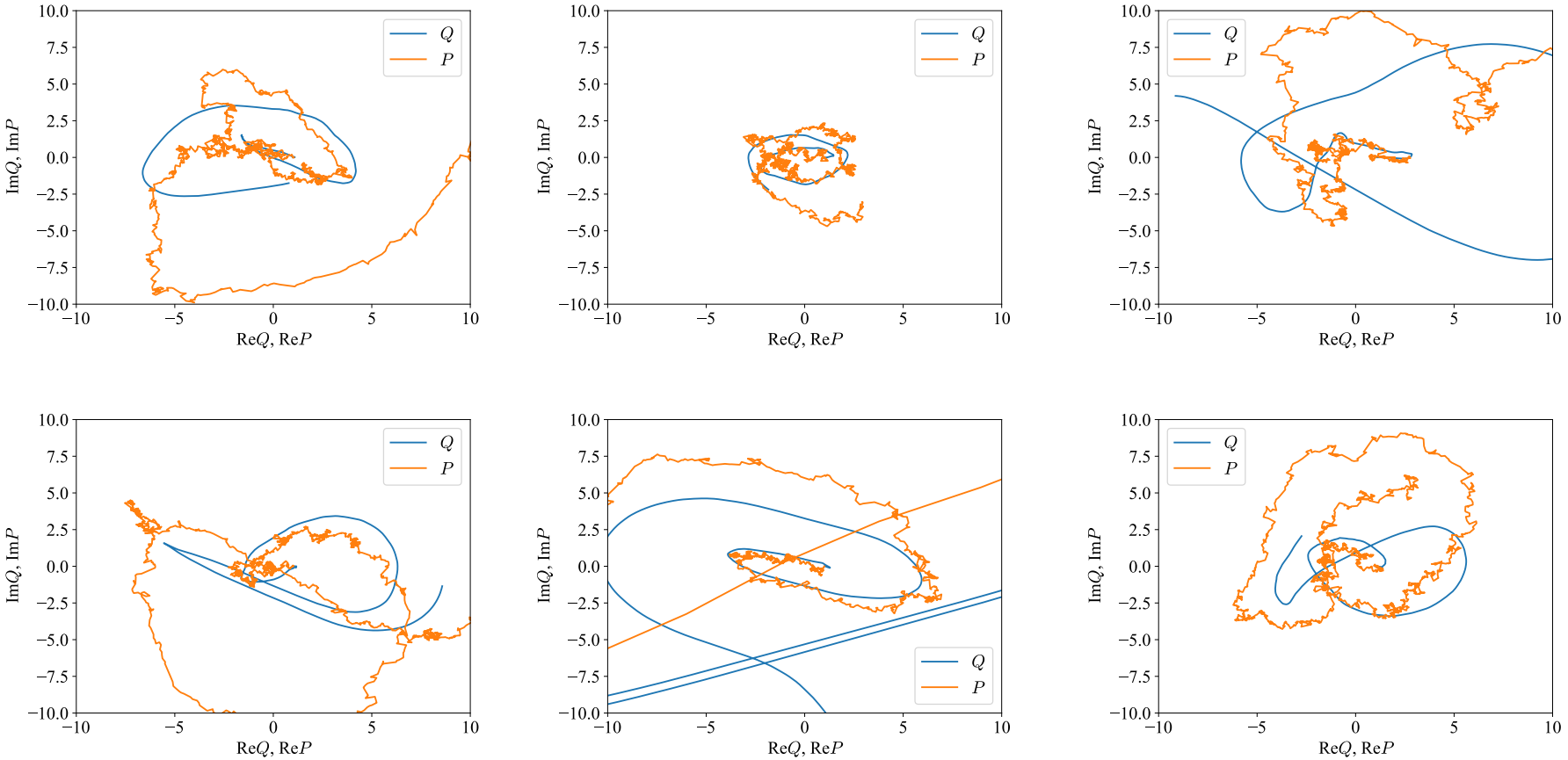
$$\zeta_n = a\,\frac{e^{-i\frac{\pi}{4}}\theta_n + e^{i\frac{\pi}{4}}\rho_n}{\sqrt{2}}, \quad \eta_n = a^{-1}\,\frac{e^{-i\frac{\pi}{4}}\theta_n - e^{i\frac{\pi}{4}}\rho_n}{\sqrt{2}}$$

Complex weighted superposition of Gaussian noises

Can we perform the simulation?

Application to the Sign Problem

No, not at all… then, we have been stuck for many years…



Application to the Sign Problem

Can the Moyal evolution be rewritten as a deterministic flow?

— formally, yes!

Instead of p_t , the Wigner function $W(P, Q, t)$ should be solved.

Fokker-Planck eq. → Moyal eq.

$$\partial_t W = -P\partial_Q W + V'(Q)\partial_P W - \frac{\lambda Q}{4}\partial_P^3 W$$

Deterministic eq. (in the continuum language)

$$\dot{Q} = P$$

$$\dot{P} = -V'(Q) + \frac{\lambda Q}{4}\frac{\partial_P^2 W}{W}$$

Difficulty conservation realized here… *W* can hit zeros as *t* goes.

Application to the Sign Problem

$$\dot{P} = -V'(Q) + \frac{\lambda Q}{4}\frac{\partial_P^2 W}{W}$$

Writing formally $W = e^{-S}$ then

$$\dot{P} = -V'(Q) + \frac{\lambda Q}{4}\left[(\partial_P S)^2 - \partial_P^2 S\right]$$ **cf. Hessian-score flow**

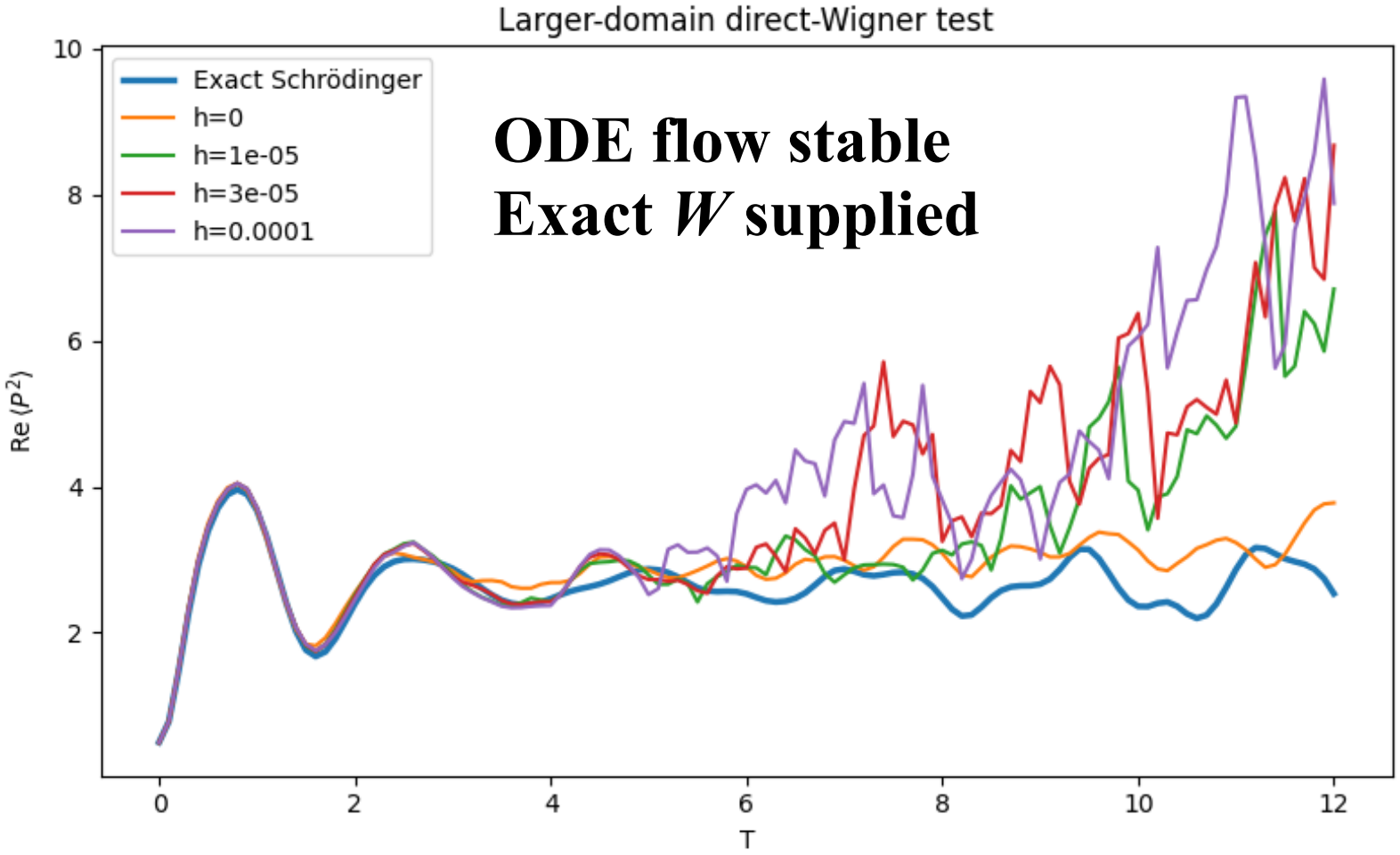
Quantum-fluct. effects re-enter through the score/Hessian terms.

ODE may become singular at $W = 0$?
Weak SDE may avoid such singularity ?
Strong SDE may blow up in complex plane ?

Application to the Sign Problem

Numerical Experiment ($\lambda = 1$):

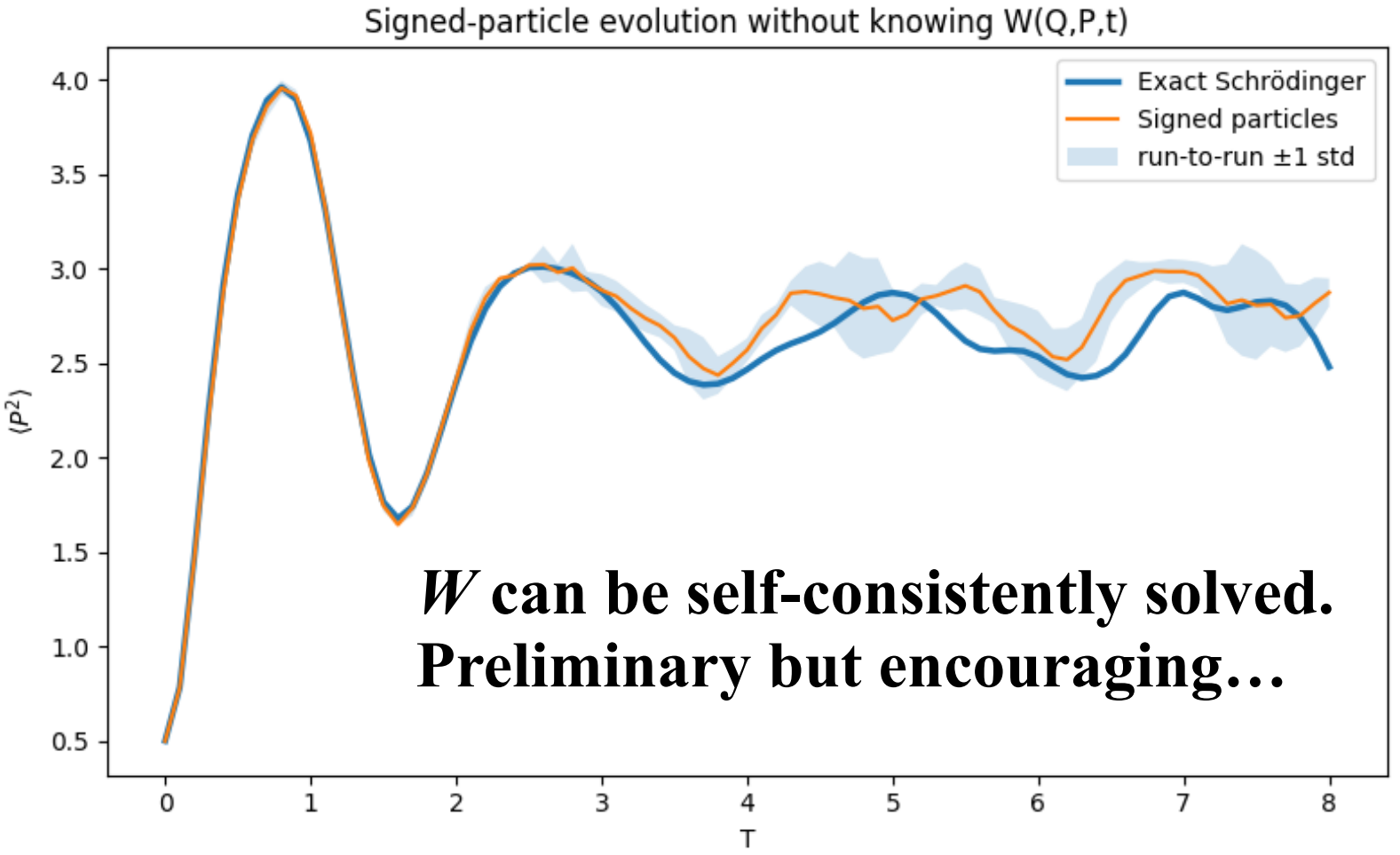
Work in progress



Application to the Sign Problem

Numerical Experiment ($\lambda = 1$):

Work in progress



***W* can be self-consistently solved.**
Preliminary but encouraging…

Summary

- **SDE/ODE equivalence:** **The stochasticity can be varied without changing the Fokker–Planck evolution when the exact score is known. With an approximate score, different noise levels lead to different dynamics.**
- **Kernel optimization:** **The transport kernel changes the transient Fokker–Planck evolution while preserving the stationary distribution. For sign problems, its choice can strongly affect stability and convergence.**
- **Sign problem:** **PIK-type deterministic transport and Complex Langevin can be viewed as different limits of a common (A, K, D) framework. A deterministic or weak-noise reformulation may improve real-time stability, although zeros of the Wigner function remain a central difficulty.**