Fluctuations in conservative systems and SPDEs

Benjamin Gess
Universität Bielefeld & Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig

Netherlands Mathematical Congress 2023

joint work with Ben Fehrman [https://arxiv.org/abs/1910.11860]
and Daniel Heydecker [https://arxiv.org/abs/2303.11289]

Max-Planck-Institut für Mathematik in den Naturwissenschaften

Universität Bielefeld

Content

General perspective

From interacting particle systems to conservative SPDEs

From large deviations to PDE with irregular drift

Further Applications

Principle viewpoint

- Consider large, interacting, stochastic systems
  (complex dynamics, many model parameters)
- Effective behavior by universal scaling limits.
  Hydrodynamic limits leading to partial differential equations (PDE)
- Stochastic fluctuations often are essential, e.g. meta-stability
- **Aim:** Universal scaling limits, capturing both the average behavior and fluctuations
- **Ansatz**: Conservative SPDEs

$$\partial_t u = \mathrm{div} F(u, Du) + \sqrt{\varepsilon}\mathrm{div}(G(u)\xi),$$

  with ξ space-time white noise, as universal fluctuating continuum models.
- Many open problems, e.g. well-posedness, regularity, stochastic dynamics [Fehrman, G., 2019, 2020, 2022], [Mariani, 2010], [Lions, Souganidis 1998ff], [Gassiat, G., Lions, Souganidis 2021], ...

Nonequilibrium statistical mechanics - fluctuating gradient flows

- Gradient flow on an (infinite dimensional) “Riemannian manifold“ [Jordan, Kinderlehrer, Otto, 1998]

$$\partial_t \rho = -M(\rho)\frac{\partial E}{\partial \rho}(\rho). \quad \text{(e.g. } \dot{X} = -M(X)\nabla E(X)).$$

- Formal non-equilibrium stationary Gibbs state $\mu = \frac{1}{Z}e^{-N^{-1}E(\rho)}" d\rho"$.
- Detailed-balance: Fluctuating gradient flow ([Öttinger 2005], fluctuating hydrodynamics [Spohn 1991])

$$\partial_t \rho = -M(\rho)\frac{\partial E}{\partial \rho}(\rho) + N^{-\frac{1}{2}} M^{\frac{1}{2}}(\rho)\xi.$$

- Decorrelation length $\approx$ typical particle distance / grid-size: $\xi \to \xi^N$
- Nonlinear diffusion

$$\partial_t \rho^N = \partial_{xx}\Phi(\rho^N) + N^{-\frac{1}{2}}\partial_x\left(\Phi^{\frac{1}{2}}(\rho^N)\xi^N\right).$$

Fluctuating hydrodynamics and macroscopic fluctuation theory

[Spohn 1991] [Bertini, De Sole, Gabrielli, Jona-Lasinio, Landim 2015]

Recall: Fluctuating gradient flow

$$\partial_t \rho^N = -M(\rho^N)\frac{\partial E}{\partial \rho}(\rho^N) + N^{-\frac{1}{2}} M^{\frac{1}{2}}(\rho^N)\xi.$$

Rare events: (Im-)probability of a fluctuation ρ in small noise $N^{-\frac{1}{2}} \to 0$ limit

$$\mathbb{P}[\rho^N \approx \rho] = e^{-N\, I(\rho)} \quad N \text{ large.}$$

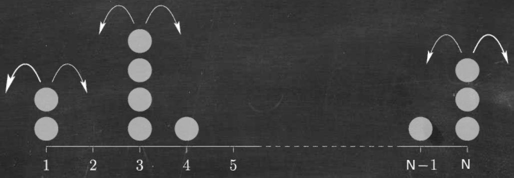
Informally, "by" contraction principle LDP has rate function

$$\begin{aligned} I(\rho) &= \inf\{\int_{t,x} g^2 : \partial_t \rho = -M(\rho)\frac{\partial E}{\partial \rho}(\rho) + M^{\frac{1}{2}}(\rho) g\} \\ &= \|M^{-\frac{1}{2}}(\rho)(\partial_t \rho + M(\rho)\frac{\partial E}{\partial \rho}(\rho))\|_{L^2}^2 \end{aligned}$$

Macroscopic fluctuation theory: Postulate $I(\rho)$ as energy for non-equilibrium systems.

From interacting particle systems to conservative SPDEs

The zero range process (could also consider simple exclusion, independent particles, ..).



- State space $\mathbb{M}_N := \mathbb{N}_0^{\mathbb{T}_N}$, i.e. configurations $\eta : \mathbb{T}_N \to \mathbb{N}_0$: System in state η if container k contains $\eta(k)$ particles.
- Local jump rate function $g : \mathbb{N}_0 \to \mathbb{R}_0^+$.
- Translation invariant, asymmetric, zero mean transition probability $p(k, l)$.
- Markov jump process $\eta(t)$ on $\mathbb{M}_N$.
- $\eta(k, t) = \text{number of particles in box } k \text{ at time } t.$

- **Hydrodynamic limit?** Multi-scale dynamics

- Empirical density field: $\mu^N(x,t) := \frac{1}{N}\sum_k \delta_{\frac{k}{N}}(x)\eta(k, tN^2)$.
- [Hydrodynamic limit - Ferrari, Presutti, Vares; 1987]

$$\mu^N(t) \rightharpoonup^* \bar{\rho}(t)\,dx$$

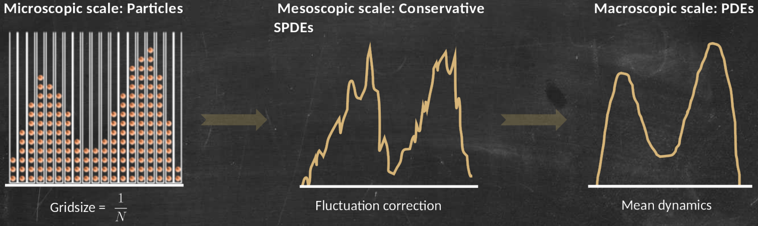
with

$$\partial_t \bar{\rho} = \partial_{xx}\Phi(\bar{\rho})$$

with Φ the mean local jump rate $\Phi(\rho) = \mathbb{E}_{\nu_\rho}[g(\eta(0))]$.

- Loss of information:
  - Fluctuations, rare events?
  - Error: $\mu^N = \bar{\rho} + O(N^{-\frac{1}{2}})$.

Fluctuating Hydrodynamics?



Ansatz: Conservative SPDEs

$$\partial_t \rho^N = \partial_{xx}\Phi(\rho^N) + N^{-\frac{1}{2}}\partial_x\left(\Phi^{\frac{1}{2}}(\rho^N)\xi^N\right).$$

Informal justification:

1. Physics: Fluctuation-dissipation relation, Fluctuating Wasserstein gradient flow, ∞-dim Fokker-Planck equations

From large deviations to parabolic-hyperbolic PDE with irregular drift

Rare events: (Im-)probability to observe a fluctuation ρ: $\mathbb{P}[\mu^N \approx \rho] = e^{-N\,I(\rho)}$.
Zero range process

$$I(\rho) = \inf\{\int_{t,x} |\partial_x H|^2 \Phi(\rho) \ : \quad \underbrace{\partial_t \rho = \partial_{xx}\Phi(\rho) + \partial_x(\Phi(\rho)\overbrace{\partial_x H}^{\in L^2_{\phi(\rho)}})}_{\text{"controlled nonlinear Fokker-Planck equation"}} \quad \}.$$

Theorem ([Large deviation principle, Kipnis, Olla, Varadhan, Benois, Landim, 80's])
For every open set $O \subseteq D([0,T], \mathcal{M}_+)$ we have

$$\mathbb{P}[\mu^N \in \bar{O}] \lesssim e^{-N \inf_{\mu \in \bar{O}} I(\mu)}$$

$$e^{-N \inf_{\mu \in O} J(\mu)} \lesssim \mathbb{P}[\mu^N \in O] \leq \mathbb{P}[\mu^N \in \bar{O}] \lesssim e^{-N \inf_{\mu \in \bar{O}} I(\mu)}$$

where $J = \overline{I_{|A}}$ and A is the set of nice fluctuations $\mu = \rho\, dx$ with ρ a solution to

$$\partial_t \rho = \partial_{xx}\Phi(\rho) + \partial_x(\Phi(\rho)\partial_x H)$$

for some $\underline{H \in C^{1,3}_{t,x}}$.Problem: $\overline{I_{|A}} = I$?

This is a frequently observed problem: E.g. Fluctuations around Boltzmann equation

Theorem (The skeleton equation, [Fehrman, G. 2022])

Let $g \in L^2_{t,x}$, ρ_0 non-negative and $\int \rho_0 \log(\rho_0) dx < \infty$. There is a unique weak solution to

$$\partial_t \rho = \Delta \Phi(\rho) + div(\Phi^{\frac{1}{2}}(\rho) g).$$

The map $g \mapsto \rho$, $L^2_{t,x} \to L^1_{t,x}$ is weak-strong continuous. E.g. $\Phi(\rho) = \rho^m$, $m \in [1, \infty)$.

Comment on the proof:Extending DiPerna-Lions, Ambrosio, Le Bris-Lions to non-linear PDE & combination with kinetic solution theory.

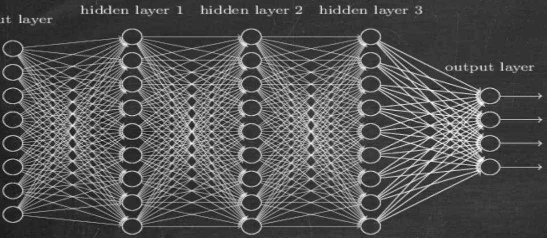
Theorem (LDP for zero range process, [G. Heydecker 2023])

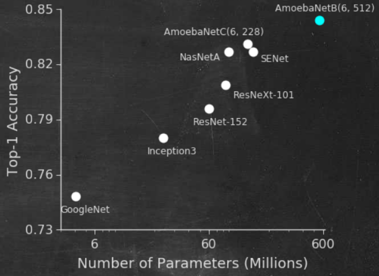
The zero range process satisfies the full large deviations principle with rate function

$$I(\rho) = \|\partial_t \rho - \partial_{xx} \Phi(\rho)\|_{H^{-1}_{\Phi(\rho)}}.$$

Further Examples: Machine learning

Feed-forward neural network





Collecting all parameters $\theta = (\theta_1, \dots, \theta_M) \in \mathbb{R}^M$
Stochastic gradient descent / empirical risk minimization

$$\theta_{n+1} = \theta_n - \eta \nabla_\theta l(f_\theta(X), Y),$$

Scaling limits: Small learning rate η, overparametrization $M \to \infty$.
Shallow network: Empirical distribution $\mu_t^M := \frac{1}{M}\sum_i \delta_{\theta_t^i} \to \mu_t$ solution to

$$\partial_t \mu_t = \mathrm{div}(\nabla V(\mu_t, \cdot)\mu_t) + \sqrt{\sigma}\,\mathrm{div}(T(\mu_t, \cdot)\mu \circ \xi),$$

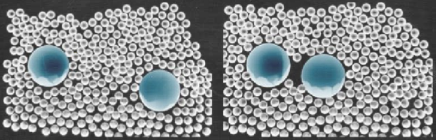
where ξ is space-time white noise and V, A, T_μ are non-local operators, see [Chen, Rotskoff, Bruna, Vanden-Eijnden, 2020], [Gess, Gvalani, Konarovskyi, 2022].

Fluctuating mean field systems

- Interacting agents with common noise

$$dX_t^i = \sigma(X_t^i, \frac{1}{L}\sum_{j\neq i}\delta_{X_t^j}) \circ dB_t + \alpha(X_t^i, \frac{1}{L}\sum_{j\neq i}\delta_{X_t^j})dW_t^i$$

with B_t^i environmental noise, W_t^i individual noise.



- Informally: Conditioned empirical measure

$$\mu^L := \frac{1}{L}\sum_i \delta_{X^i} \to \mu$$

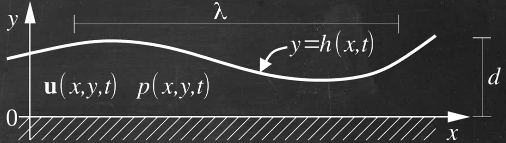
with $\mu = m\,dx$ satisfying the nonlocal conservative SPDE

$$\partial_t m = \frac{1}{2}\Delta\left(\alpha^2(x,m)m\right) + \nabla\cdot(\sigma(x,m)m \circ dB_t).$$

See, e.g. [Kurtz, Xiong 2001], [Coghi, Gess, 2019].

Stochastic thin films

$$\partial_t h = -\mathrm{div}(h^m \nabla \Delta h) + \mathrm{div}(h^{m/2}\xi)$$



(source: Grün, Mecke, Rauscher, 2006)

Fluctuating gradient flow structure

$$\partial_t h = -M(\rho)\frac{\partial E}{\partial h}(h) + M^{\frac{1}{2}}(h)\xi,$$

with $E(h) = \frac{1}{2}\int |\nabla h|^2 dx$ and $M(h) = \mathrm{div}(h^m \nabla \cdot)$,
Relevance to include thermal fluctuations:

- Improved prediction of empirical film rupture time scales
  [Grün, Mecke, Rauscher 2006]
- Corrected spreading rate of droplets
  [Davidovitch, Moro, Stone 2005]

Conclusion:

- Conservative SPDEs as universal fluctuating continuum models.
- Compared to classical hydrodynamic description they include information on fluctuations.
- Central limit fluctuations and large deviations of microscopic systems are correctly captured by the conservative SPDEs.

References

M. Coghi and B. Gess.
Stochastic nonlinear Fokker-Planck equations.
Nonlinear Analysis. Theory, Methods & Applications. An International Multidisciplinary Journal, 187:259–278, 2019.

B. Fehrman and B. Gess.
Non-equilibrium large deviations and parabolic-hyperbolic PDE with irregular drift.
arXiv:1910.11860 [math], Mar. 2022.

B. Gess, R. S. Gvalani, and V. Konarovskyi.
Conservative SPDEs as fluctuating mean field limits of stochastic gradient descent, July 2022.

B. Gess and D. Heydecker.
A Rescaled Zero-Range Process for the Porous Medium Equation: Hydrodynamic Limit, Large Deviations and Gradient Flow, Mar. 2023.