Mean-field limits for Riesz systems

Matthew Rosenzweig

September 28, 2023

Carnegie Mellon University

“About Entropy in Large classical Particle Systems,” CMI Workshop, Mathematical Institute, Oxford, UK

Outline of the Talk

Introduction

Modulated energy and functional inequalities

Modulated free energy and propagation/generation of chaos

Current and future directions

Introduction

The microscopic model

Microscopic model given by system of N differential equations

$$\begin{cases} dx_i^t = \Big(\frac{1}{N}\sum_{\substack{1\le j\le N \\ j\ne i}} \mathbb{M}\nabla \mathrm{g}(x_i^t, x_j^t) - \nabla V(x_i^t)\Big)dt + \sqrt{\frac{2}{\beta}}dW_i^t \\ x_i^t|_{t=0} = x_i^0 \in \mathbb{R}^{\mathrm{d}}. \end{cases}$$

- g is the (symmetric) interaction potential
- V is the external confining potential (e.g., $V(x) = \frac{1}{2}|x|^2$)
- $W_1, \ldots, W_N$ are independent standard Brownian motions
- $\frac{1}{\beta} \ge 0$ interpreted as a temperature; noise models thermal fluctuations.

Constant real $\mathrm{d} \times \mathrm{d}$ matrix $\mathbb{M}$ is either

- antisymmetric (conservative/Hamiltonian)
- $\mathbb{M} = -\mathbb{I}$ (dissipative/gradient)

N.B.

- No self-interaction in model
- Interaction is *long-range*
- Force/velocity field experienced by a single particle is $\propto$ *average* of fields generated by remaining particles (i.e., *mean-field*)

Relevance of model I

Model case for potential $g(x - y) = g(x, y)$ are *Riesz interactions* indexed by parameter $s < d$:

$$g(x) = \frac{1}{c_{s,d}} \begin{cases} -\log|x|, & s = 0 \\ |x|^{-s}, & s \neq 0 \end{cases}$$

- $s = d - 2$ *Coulomb*
- $s < d - 2$ *sub-Coulomb*
- $s > d - 2$ *super-Coulomb*

Numerous applications & connections to particle systems in physics, particle methods for PDEs, finding equilibrium states for interaction energies, biological and sociological models, large neural networks, approximation theory...

Refer to surveys Jabin 2014, Jabin-Wang 2017, Chaintron-Diez 2022, Golse 2022 and book Borodachov-Hardin-Saff 2019

Questions of interest: mean-field limit I

What are the limiting dynamics of the empirical measure

$$\mu_N^t := \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i^t} \in \mathcal{P}(\mathbb{R}^{\mathrm{d}})$$

as $N \to \infty$?

Formally expect that if $\mu_N^0 \xrightarrow[N\to\infty]{} \mu^0$, then $\mu_N^t \xrightarrow[N\to\infty]{} \mu^t$, where μ^t is a solution to the *mean-field equation*

$$\begin{cases} \partial_t \mu - \operatorname{div}(\mu(\nabla V - \mathbb{M}\nabla \mathsf{g} * \mu)) = \frac{1}{\beta}\Delta\mu \\ \mu|_{t=0} = \mu^0. \end{cases}$$

When $\beta = \infty$, μ_N^t is, in fact, a weak solution to equation (6).

Establishing the mean-field limit refers to proving this convergence.

Questions of interest: propagation of chaos I

Suppose the initial positions $X_N^0 = (x_1^0, \dots, x_N^0)$ are independently and identically distributed with some law μ^0:

$$f_N^0 = (\mu^0)^{\otimes N}.$$

What is the limiting behavior as $N \to \infty$ of the law $f_N^t(x_1, \dots, x_N)$ of the positions $X_N^t = (x_1^t, \dots, x_N^t)$ of the particles at time t?

If μ^t is the solution of the mean-field equation with initial datum μ^0, does it hold that

$$f_N^t \approx (\mu^t)^{\otimes N} \qquad \text{as } N \to \infty \text{ ?}$$

Propagation of chaos refers to the asymptotic factorization of k-point marginals $f_{N;k}^t \rightharpoonup (\mu^t)^{\otimes k}$.[1]

Known that mean-field convergence and propagation of chaos are closely related; qualitatively, they are equivalent Hauray-Mischler 2014.

[1] Recall that the k-point marginal $f_{N;k} := \int_{(\mathbb{R}^d)^{N-k}} f_N(\cdot, x_{k+1}, \dots, x_N) dx_{k+1} \cdots dx_N$.

Questions of interest: generation of chaos I

Related notion of **generation of chaos**.[2] Even when the initial law f_N^0 is not asymptotically chaotic, one still has that $f_{N;k}^t - (\mu^t)^{\otimes k} \to 0$ as $t \to \infty$ and $N \to \infty$.

Interpreted in an *entropic sense*, this means the relative entropy

$$H_N(f_N^t|(\mu^t)^{\otimes N}) := \frac{1}{N} \int_{(\mathbb{R}^d)^N} \log\left(\frac{f_N^t}{(\mu^t)^{\otimes N}}\right) df_N^t$$

tends to zero as $t \to \infty$ and $N \to \infty$.

Question: What is the relation between the two limits $t \to \infty$ and $N \to \infty$? In particular, is $H_N(f_N^t|(\mu^t)^{\otimes N}) = o_N(1)$ uniformly in time (uniform-in-time entropic propagation of chaos)?

Questions of interest: generation of chaos II

For (repulsive) overdamped Langevin dynamics ($\mathbb{M} = -\mathbb{I}$), one expects the law f_N^t weakly converges to the *Gibbs measure*

$$d\mathbb{P}_{N,\beta}(X_N) = \frac{1}{Z_{N,\beta}} e^{-\beta \mathcal{H}_N(X_N)} dX_N.$$

with Hamiltonian

$$\mathcal{H}_N(X_N) = \frac{1}{2N} \sum_{1 \le i \ne j \le N} \mathsf{g}(x_i, x_j) + \sum_{i=1}^{N} V(x_i).$$

If $\mathbb{P}_{N,\beta}$ admits a *logarithmic Sobolev inequality (LSI)*,

$$H(\mathbb{Q}_N | \mathbb{P}_{N,\beta}) \le C_{LS} I(\mathbb{Q}_N | \mathbb{P}_{N,\beta}),$$

then this convergence may be quantified and is exponentially fast. Poincaré inequality or LSI holds as soon as $\mathcal{H}_N$ is *uniformly convex*, i.e. $\mathrm{Hess}\, \mathcal{H}_N \ge c I_{dN \times dN}$, for $c > 0$ Bakry-Émery 1985.

Difficulty is going beyond the uniformly convex case (e.g., Bauerschmidt-Bodineau 2019-21).

Questions of interest: generation of chaos III

The mean-field density μ^t should weakly converge to the *thermal equilibrium measure* μ_β, which is the minimizer among probability measures of the mean-field free energy

$$\mathcal{E}_\beta(\mu) := \frac{1}{2}\int_{(\mathbb{R}^\mathsf{d})^2} \mathsf{g}(x,y)d\mu^{\otimes 2}(x,y) + \int_{\mathbb{R}^\mathsf{d}} V(x)d\mu(x) + \frac{1}{\beta}\int_{\mathbb{R}^\mathsf{d}} \log\mu(x)d\mu(x).$$

If V grows sufficiently fast at infinity, then $\mathcal{E}_\beta$ has a unique minimizer, which is characterized by the existence of a constant $c_\beta \in \mathbb{R}$ such that

$$\mathsf{g} * \mu_\beta + V + \frac{1}{\beta}\log\mu_\beta = c_\beta \quad \text{in } \mathbb{R}^\mathsf{d}.$$

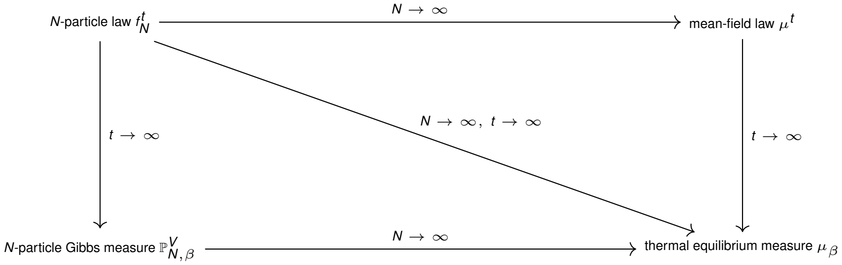
In many cases, it's known that $\mathbb{P}_{N,\beta}^{(k)} \rightharpoonup \mu_\beta^{\otimes k}$. If μ_β admits the nonlinear LSI

$$\mathcal{E}_\beta(\mu) - \mathcal{E}_\beta(\mu_\beta) \le C_{LS}\beta \int_{\mathbb{R}^\mathsf{d}} \left|\frac{1}{\beta}\nabla\log\mu + \nabla V + \nabla\mathsf{g} * \mu\right|^2 d\mu,$$

then μ^t converges exponentially fast to μ_β. Easy to show that a nonlinear LSI for $\mathbb{P}_{N,\beta}$ implies mean-field LSI for μ_β.

Questions of interest: generation of chaos IV

Figure: Large Particle Number and Large Time Limits



[2]This term was recently coined by Jani Lukkarinen.

Previous results: mean-field convergence/propagation of chaos

- Coupling method Sznitman 1991, Hauray-Jabin 2015, Boers-Pickl 2016, Lazarovici-Pickl 2017, Graß 2021, Guillin-Le Bris-Monmarché 2021,...
- Wasserstein stability Braun-Hepp 1977, Dobrushin 1979, Neunzert-Wick 1974, Hauray 2009, Carrillo-Choi-Hauray 2014,...
- Relative entropy Jabin-Wang 2016, 2018, Guillin-Le bris-Monmarché 2021
- Control of microscopic dynamics and compactness for well-chosen point configurations Goodman-Hou-Lowengrub 1990, Schochet 1996,...
- Displacement convexity for Wasserstein gradient flow Carrillo-Ferreira-Precioso 2012, Berman-Onnheim 2015,...
- Compactness via diffusion Osada 1985-1987, Rogers-Shi 1993, Cépa-Lepingle 1997, Fournier-Hauray-Mischler 2014, Wang-Zhao-Zhu 2022,...
- Stability for BBGKY hierarchy Lacker 2021, Han 2022, Jabin-Poyato-Soler 2021, Bresch-Jabin-Soler 2022, Lacker-Le Flem 2023
- **Modulated energy/free energy method** Duerinckx 2016, Serfaty 2020, R. 2020-2021, Nguyen-R.-Serfaty 2021, R.-Serfaty 2021 / Bresch-Jabin-Wang 2019-2020, Chodron de Courcel-R.-Serfaty 2023

Modulated energy and functional inequalities

Modulated-energy method I

$$F_N(X_N, \mu) := \int_{(\mathbb{R}^d)^2 \setminus \triangle} g(x, y) d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\Big)^{\otimes 2}(x, y).$$

Total interaction of system of N discrete charges at x_i against neutralizing background of charge μ, with self-interaction of points (infinite, if $g(x, x) = \infty$) removed.

Quantity first appeared in stat mech of Coulomb/Riesz gases Sandier-Serfaty 2015, Rougerie-Serfaty 2016, Petrache-Serfaty 2017, Leblé-Serfaty 2017-2018 as a *next-order energy*.

Falls out of the *splitting formula*

$$\mathcal{H}_N(X_N) = N\mathcal{E}_\beta(\mu_\beta) + NF_N(X_N, \mu_\beta) - \frac{1}{\beta}\sum_{i=1}^N \log \mu_\beta(x_i).$$

Modulated-energy method II

This modulated energy first used in dynamics context Duerinckx 2016, Serfaty 2020

Idea: establish a Grönwall relation for $F_N(X_N^t, \mu^t)$

- Method goes back to Brenier 2000; similarities with *relative-entropy method* Dafermos 1979, DiPerna 1979, Yau 1991, Saint-Raymond 2009
- Exploits a weak-strong uniqueness principle for limiting equation
- Advantages - quantitative; no need for study of microscopic dynamics
- Disadvantages - typically requires some regularity for or an *a priori* assumption on the limiting solution

Coercivity of modulated energy I

For simplicity, assume $-\Delta \mathsf{g} = \mathsf{c_d}\delta_0$ (Coulomb),[3]

$$\int_{(\mathbb{R}^{\mathsf{d}})^2} \mathsf{g}(x-y)df^{\otimes 2}(x,y) = \mathsf{c_d}\|f\|_{\dot{H}^{-1}}^2.$$

Infinite for $f = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i} - \mu$ if $\mathsf{d} \geq 2$!

We can smear the point masses δ_{x_i} to uniform measure $\delta_{x_i}^{(\eta_i)}$ on sphere $\partial B(0,\eta_i)$ to show that

$$F_N(X_N,\mu) \geq \frac{1}{\mathsf{c_{d,s}}}\left\|\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}^{(\eta_i)} - \mu\right\|_{\dot{H}^{-1}}^2 - \mathrm{Error}_{N,\vec{\eta}}.$$

Both terms blow up as $\eta_i \to 0$, but it turns out that there is a natural scale $N^{-\frac{1}{\mathsf{d}}}$ for the η_i in terms of N, such that the error is negligible as $N \to \infty$.

With more difficulty, this idea of renormalizing the energy through smearing the Dirac masses can be generalized to the full range $0 \leq \mathsf{s} < \mathsf{d}$.

---

[3] Recall that for $a \in \mathbb{R}$, $\|f\|_{\dot{H}^a}^2 := \int_{\mathbb{R}^{\mathsf{d}}}(2\pi|\xi|)^{2a}|\hat{f}(\xi)|^2 d\xi$, where $\hat{f}(\xi)$ denotes the Fourier transform of f.

The modulated energy identity

Suppose that $\beta = \infty$. One computes

$$\frac{d}{dt} F_N(X_N^t, \mu^t)$$

$$\leq - \int_{(\mathbb{R}^\mathsf{d})^2 \setminus \triangle} \left(u^t(x) - u^t(y) \right) \cdot \nabla \mathsf{g}(x-y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i^t} - \mu^t \right)^{\otimes 2}(x,y),$$

where $u^t := \nabla V - \mathbb{M}\nabla \mathsf{g} * \mu^t$ is the vector field associated to the mean-field equation solution μ^t.

If we have a functional inequality of the form

$$|\mathrm{RHS}| \leq C_1(\|u^t\|)\Big(|F_N(X_N^t, \mu^t)| + C_2(\|\mu^t\|)N^{-\alpha}\Big),$$

where C_1, C_2 depends on d, some norm of u^t, μ^t, respectively, and $\alpha > 0$, then Grönwall implies an estimate

$$|F_N(X_N^t, \mu^t)| \leq \left(|F_N(X_N^0, \mu^0)| + N^{-\alpha} \int_0^t C_2(\|\mu^\tau\|) d\tau \right) e^{\int_0^t C_1(\|u^\tau\|) d\tau}.$$

Variation by transport I

In the context of mean-field limits, essential to control quantities that correspond to differentiating F_N along a transport field:

$$\frac{d^n}{dt^n|_{t=0}} F_N\Big((\mathbb{I}+tv)^{\otimes N}(X_N), (\mathbb{I}+tv)\#\mu\Big)$$
$$= \int_{(\mathbb{R}^d)^2\setminus\triangle} \nabla^{\otimes n}\mathrm{g}(x-y) : (v(x)-v(y))^{\otimes n} d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\Big)^{\otimes 2}(x,y),$$

where $(\mathbb{I}+tv)(x) = x + tv(x)$.

Harmonic Analysis Problem: Prove functional inequalities of form

$$|\mathrm{RHS}| \le C(\|v\|)\big(F_N(X_N,\mu) + C(\|\mu\|)N^{-\alpha}\big)$$

for some $\alpha > 0$. Because then this yields an estimate for the time-evolved modulated energy $F_N(X_N^t, \mu^t)$ that implies mean-field convergence.

There are two perspectives on proving these functional inequalities...

The stress-energy tensor perspective I

The first perspective is due Leblé-Serfaty 2018, Serfaty 2020.

Idea: interpret the variation expression in terms of a stress-energy tensor structure. In the Coulomb case, formally letting H_N solve $-\Delta H_N = \mathsf{c}_d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)$, one can use integration by parts to write

$$\int_{(\mathbb{R}^d)^2\setminus\triangle} \Big(v(x) - v(y)\Big)\cdot\nabla \mathsf{g}(x-y)d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\Big)^{\otimes 2}(x,y) = \frac{1}{\mathsf{c}_{d,s}}\int_{\mathbb{R}^d} v\cdot \mathrm{div}\,[H_N, H_N]\,dx,$$

where the *stress-energy tensor* is defined by

$$[h,f]^{ij} := \partial_i h\partial_j f + \partial_i f\partial_j h - \delta^{ij}(\nabla h\cdot\nabla f), \qquad 1\le i,j\le d.$$

Integration by parts and Cauchy-Schwarz allow one to conclude the bound

$$\left|\int_{\mathbb{R}^d} v\cdot\mathrm{div}\,[H_N,H_N]\,dx\right| \le C\|\nabla v\|_{L^\infty}\|\nabla H_N\|_{L^2}^2.$$

The stress-energy tensor perspective II

Formally, $\|\nabla H_N\|_{L^2}^2$ is the Coulomb energy of $\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu$; but $\|\nabla H_N\|_{L^2} = \infty$ due to the singularity of the Dirac masses.

However, the reasoning may be implemented after a <u>renormalization</u>: namely, replace δ_{x_i} with the smeared charge $\delta_{x_i}^{(\eta_i)}$ above, and apply reasoning to

$$H_{N,\vec{\eta}} := \mathsf{g} * \left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - \mu\right), \qquad \vec{\eta} = (\eta_1, \ldots, \eta_N) \in \mathbb{R}_+^N.$$

Estimate the error directly from this replacement after choosing the η_i to be a nearest-neighbor type distance of the order $N^{-\frac{1}{\mathsf{d}}}$.

The stress-tensor approach is elegant, low-tech (i.e., integration by parts), and based on local arguments.

However, it is rigid in the sense that it seems restricted to exact Riesz potentials. Extendable to super-Coulomb case $\mathsf{d}-2 < \mathsf{s} < \mathsf{d}$ by Caffarelli-Silvestre extension; but unclear how to extend it to the sub-Coulomb case $0 \le \mathsf{s} < \mathsf{d}-2$.

The commutator perspective I

The second perspective originates in R. 2020 on developing a new generalization of the modulated-energy method for multiplicative noise.

If $f = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i} - \mu$, then *formally* the first variation may be rewritten

$$\int_{\mathbb{R}^d} \Big(u \cdot (\nabla g * f) - g * (\mathrm{div}(uf))\Big) df(x) = \left\langle f, \left[u^i, \frac{\partial_i}{(-\Delta)^{\frac{d-s}{2}}}\right] f \right\rangle_{L^2}.$$

The first variation is the quadratic form associated to the commutator $\left[u^i, \frac{\partial_i}{(-\Delta)^{\frac{d-s}{2}}}\right]$. The higher-order variations can be similarly formulated in terms of iterated commutators.

By a renormalization argument that involves regularizing the Dirac masses, similar to as in the stress-tensor approach, one can then apply estimates for a class of singular integral operators known as *Calderón* d*-commutators* Calderón 1980, Christ-Journé 1987, Seeger et al. 2019, Lai 2020. The error introduced by renormalization can be estimated directly.

A functional inequality for all Riesz/Riesz-type potentials

Using the commutator perspective, Q.H. Nguyen-R.-Serfaty 2021 proved a first-order functional inequality valid for all Riesz cases $0 \le s < d$, as well as more general “Riesz-type” potentials.

Application: the modulated energy method provides a unified approach to quantitatively proving mean-field convergence of Riesz systems. Previous works only covered varying subcases.

The method of proof we introduced works for higher-order functional inequalities as well.

Sharpness of functional inequality

A natural question is the *size of the exponent* α in the functional inequality.

$$\left|\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}\Big(v(x)-v(y)\Big)\cdot\nabla \mathsf{g}(x-y)d\Big(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}-\mu\Big)^{\otimes 2}(x,y)\right|$$
$$\leq C_1(\|\nabla v\|)\Big(F_N(X_N,\mu)+C_2(\|\mu\|)N^{-\alpha}\Big).$$

More precisely, we want both terms in the RHS to be of the same order as $N\to\infty$; and we say the functional inequality is *sharp* if this is the case.

- By only counting nearest-neighbor (with typical distance of $N^{-\frac{1}{\mathsf{d}}}$) interactions, one expects F_N is at least of order $N^{\frac{\mathsf{s}}{\mathsf{d}}-1}$
- $F_N \geq -CN^{\frac{\mathsf{s}}{\mathsf{d}}-1}$, where $C = C(\|\mu\|_{L^\infty}) > 0$
- Known that $\min F_N$ is of order $N^{\frac{\mathsf{s}}{\mathsf{d}}-1}$ Sandier-Serfaty 2015, Rougerie-Serfaty 2015, Petrache-Serfaty 2017, Hardin et al. 2017
- Optimal exponent $\alpha = 1-\frac{\mathsf{s}}{\mathsf{d}}$ only been shown for Coulomb case $\mathsf{s}=\mathsf{d}-2$ Leblé-Serfaty 2018, Serfaty 2020, R. 2021

Theorem 1 (R.-Serfaty 2022, Chodron de Courcel-R.-Serfaty 2023)

Let $\mathsf{d} \geq 1$, $\mathsf{d}-2 \leq \mathsf{s} < \mathsf{d}$. *There exists a constant* $C = C(\mathsf{d},\mathsf{s}) > 0$ *such that TFH. Let* $\mu \in L^1(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}})$ *with unit mean and* $v : \mathbb{R}^{\mathsf{d}} \to \mathbb{R}^{\mathsf{d}}$ *be Lipschitz. Then for any pairwise distinct* $X_N \in (\mathbb{R}^{\mathsf{d}})^N$*, it holds that*

$$\left|\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle} (v(x)-v(y))\cdot\nabla \mathsf{g}(x-y) d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\Big)^{\otimes 2}(x,y)\right|$$

$$\leq C\|\nabla v\|_{L^\infty}\left(F_N(X_N,\mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2\mathsf{d}N}\mathbf{1}_{\mathsf{s}=0} + C\|\mu\|_{L^\infty}^{\frac{\mathsf{s}}{\mathsf{d}}} N^{\frac{\mathsf{s}}{\mathsf{d}}-1}\right).$$

Comments

- Proof based on the stress-tensor perspective. The improvement to $N^{\frac{s}{d}-1}$ comes from better estimation of the renormalization error.
- We also have sharp estimates for higher-order variations

$$\int_{(\mathbb{R}^d)^2\setminus\triangle} \nabla^{\otimes n}g(x-y) : (v(x)-v(y))^{\otimes n} d\Big(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}-\mu\Big)^{\otimes 2}(x,y).$$

  Proof via an induction argument based on using lower-order variation estimates, which combines both the stress-tensor and commutator perspectives.
- Application: the optimal $N^{\frac{s}{d}-1}$ rate of convergence for the mean-field limit measured in terms of $F_N(X_N^t,\mu^t)$. Previously, the optimal rate was only known for the Coulomb case $s = d-2$ Serfaty 2021, R. 2021. Only the sub-Coulomb case $0 \le s < d-2$ remains.

Further applications of functional inequalities

- Seen that these FIs are crucial for proving mean-field limits of classical particle systems Serfaty 2020, Duerinckx-Serfaty 2020, Bresch-Jabin-Wang 2019-2020, R. 2020-2022, Han Kwan-Iacobelli 2020, Q.H. Nguyen-R.-Serfaty 2021
- Also used to study scaling limits of quantum systems Golse-Paul 2020, R. 2021, Ben Porat 2022
- FIs (first- and higher-order) used to prove central limit theorems for fluctuations of Coulomb gases Leblé-Serfaty 2018, Serfaty 2021
- Second-order FIs also have applications to mean-field limits with special kinds of multiplicative noise R. 2020

Uniform-in-time convergence at positive temperature I

At positive temperature $\beta < \infty$, the noise has a diffusive effect in the limit $N \to \infty$. At the level of the mean-field PDE, this is seen through the decay of solutions as $t \to \infty$: for any $1 \le p \le q \le \infty$

$$\|\nabla^{\otimes n}\mu^t\|_{L^q} \lesssim_{\mathsf{d},p,q} (t/\beta)^{-\frac{n}{2}-\frac{\mathsf{d}}{2}\left(\frac{1}{p}-\frac{1}{q}\right)}\|\mu^0\|_{L^p}.$$

This is exactly the same smoothing effect as satisfied by the linear heat equation! On the torus $\mathbb{T}^{\mathsf{d}}$, one even has exponential decay with μ^t replaced by $\mu^t - 1$.

We saw earlier that the time dependence of our modulated energy estimate depends on quantities involving norms of μ^t:

$$|F_N(X_N^t,\mu^t)| \le \left(|F_N(X_N^0,\mu^0)| + N^{-\alpha}\int_0^t C_2(\|\mu^\tau\|)d\tau\right)e^{\int_0^t C_1(\|u^\tau\|)d\tau}.$$

New idea: Use the decay of μ^t to improve the time dependence, possibly obtaining estimates which are *uniform-in-time*.

Uniform-in-time convergence at positive temperature II

In R.-Serfaty 2021, we developed a stochastic version of the modulated-energy method based on estimating

$$\mathbb{E}\Big(|F_N(X_N^t, \mu^t)|\Big),$$

that works for either conservative or repulsive dissipative dynamics and leads to uniform-in-time estimates (w/ $V = 0$).

Previous works on uniform-in-time convergence (e.g., Malrieu 2003, Cattiaux-Guillin-Malrieu 2008, Salem 2018, Durmus-Eberle-Guillin-Zimmer 2020, Arnaudon-Del Moral 2020, Delarue-Tse 2021, Delgadino-Gvalani-Pavliotis-Smith 2023) impose strong convexity and/or regularity assumptions on the interaction potential g.

Unfortunately, our method breaks down for $s > d - 2$, essentially because g is no longer superharmonic. Also, doesn't work for $s = d - 2$ (cf. Guillin-Le bris-Monmarché 2021).

Modulated free energy and propagation/generation of chaos

Modulated free energy I

A closely related object is the *modulated free energy* of Bresch-Jabin-Wang 2019-2020,

$$E_N(f_N,\mu) := \frac{1}{\beta} H_N(f_N|\mu^{\otimes N}) + \mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right],$$

which is well-suited to studying overdamped Langevin dynamics at positive temperature.

Combines

- relative entropy $H_N(f_N|\mu^{\otimes N})$,
- (average) modulated energy $\mathbb{E}_{f_N}[F_N(X_N,\mu)]$.

Now the points $X_N \in (\mathbb{R}^{\mathrm{d}})^N$ are viewed as randomly distributed according to the law f_N and we take the expectation of the modulated energy $F_N(X_N,\mu)$.

Enlightening to re-express the modulated free energy in a different form...

Modulated Gibbs measure I

Given a density μ, we can define the *modulated Gibbs measure*

$$\mathbb{Q}_{N,\beta}(\mu) := \frac{1}{K_{N,\beta}(\mu)} e^{-\beta N F_N(X_N,\mu)} d\mu^{\otimes N}(X_N),$$

and *modulated partition function*

$$K_{N,\beta}(\mu) := \int_{(\mathbb{R}^{\mathsf{d}})^N} e^{-\beta N F_N(X_N,\mu)} d\mu^{\otimes N}(X_N).$$

Inserting the splitting formula

$$\mathcal{H}_N(X_N) = N\mathcal{E}_\beta(\mu_\beta) + NF_N(X_N,\mu_\beta) - \frac{1}{\beta}\sum_{i=1}^N \log \mu_\beta(x_i).$$

into $\mathbb{P}^V_{N,\beta}$,

$$d\mathbb{P}^V_{N,\beta}(X_N) = \frac{e^{-\beta N\mathcal{E}_\beta(\mu_\beta)}}{Z^V_{N,\beta}} e^{-\beta N F_N(X_N,\mu_\beta)} d(\mu_\beta)^{\otimes N}(X_N).$$

In other words, we have found that

$$\mathbb{P}^V_{N,\beta} = \mathbb{Q}_{N,\beta}(\mu_\beta) \quad \text{and} \quad Z^V_{N,\beta} = K_{N,\beta}(\mu_\beta) e^{-\beta N\mathcal{E}_\beta(\mu_\beta)}.$$

Gibbs measure = modulated Gibbs measure relative to μ_β.

Modulated Gibbs measure II

Conversely, given a probability measure μ, the modulated Gibbs measure $\mathbb{Q}_{N,\beta}(\mu)$ may be seen as a Gibbs measure through a change of the confining potential.

Let

$$V_{\mu,\beta} := -\mathsf{g} * \mu - \frac{1}{\beta} \log \mu.$$

Then retracing the steps of the splitting formula above, one has

$$\mathbb{Q}_{N,\beta}(\mu) = \mathbb{P}_{N,\beta}^{V_{\mu,\beta}}.$$

Modulated Gibbs measure III

Using the explicit form of the modulated Gibbs measure $\mathbb{Q}_{N,\beta}(\mu)$, we may rewrite

$$E_N(f_N, \mu) = \frac{1}{\beta}\left(H_N(f_N|\mathbb{Q}_{N,\beta}(\mu)) + \frac{\log K_{N,\beta}(\mu)}{N}\right).$$

With this rewriting, a crucial condition appearing in all that follows, called *smallness of the free energy*, is

$$|\log K_{N,\beta}(\mu)| = o(N),$$

which corresponds for instance to the “large deviations estimates” in Jabin-Wang 2018. This condition—and even a stronger quantitative one—can be proven in the Riesz cases and for bounded continuous interactions.

Up to a constant related to the smallness of free energy condition, the modulated free energy is another relative entropy!

Dissipation of modulated free energy I

Crucial computation of Bresch et al.,

$$\frac{d}{dt}E_N(f_N^t,\mu^t) \leq$$

$$-\frac{1}{2}\mathbb{E}_{f_N^t}\left[\int_{(\mathbb{R}^d)^2\setminus\triangle}(u^t(x)-u^t(y))\cdot\nabla_1 g(x,y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\right]$$

$$-\frac{1}{\beta^2 N}\mathbb{E}_{f_N^t}\left[\sum_{i=1}^N\left|\nabla_i\log\left(\frac{f_N^t}{(\mu^t)^{\otimes N}}\right)+\frac{\beta}{N}\sum_{j\neq i}\nabla_1 g(x_i,x_j)-\beta\nabla g*\mu^t(x_i)\right|^2\right],$$

where the velocity field associated to the mean-field dynamics is

$$u^t := \frac{1}{\beta}\nabla\log\mu^t+\nabla V+\nabla g*\mu^t.$$

At first pass...

- The second term on the RHS, which is ≤ 0, may be discarded.
- The first term on the RHS can be controlled in the Riesz case by the modulated energy itself through the functional inequality, allowing to close a Grönwall loop. On $\mathbb{T}^d$ (when $V=0$), this is what is done in Bresch-Jabin-Wang 2019, Chodron de Courcel-R.-Serfaty 2023.

Theorem 2 (Chodron de Courcel-R.-Serfaty 2022)

Let $\mathsf{d} \geq 1$*,* $\max(0, \mathsf{d}-2) \leq \mathsf{s} < \mathsf{d}$ *and* $\beta < \infty$*. Let* $\mu^0 \in \mathcal{P}(\mathbb{T}^\mathsf{d}) \cap W^{2,\infty}(\mathbb{T}^\mathsf{d}) \cap \dot{H}^{1+\mathsf{s}-\mathsf{d}}(\mathbb{T}^\mathsf{d})$*, such that* $\inf_{\mathbb{T}^\mathsf{d}} \mu^0 > 0$*. Then the mean-field equation is globally well-posed in* $C\left([0,\infty), \mathcal{P}(\mathbb{T}^\mathsf{d}) \cap W^{2,\infty}(\mathbb{T}^\mathsf{d})\right)$*;* $\inf_{\mathbb{T}^\mathsf{d}} \mu^t \geq \inf_{\mathbb{T}^\mathsf{d}} \mu^0$*; and for any* $n \geq 0$ *and* $1 \leq p \leq \infty$*,*

$$\forall t \geq 1, \qquad \|\nabla^{\otimes n}(\mu^t - 1)\|_{L^p} = O(e^{-Ct}), \qquad \text{as } t \to \infty.$$

For an "entropy solution" f_N *to* N*-particle forward Kolmogorov equation, define the quantity*

$$\mathcal{E}_N^t := E_N(f_N^t, \mu^t) + \frac{\log\left(N\|\mu^t\|_{L^\infty}\right)}{2N\mathsf{d}} \mathbf{1}_{\mathsf{s}=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{\mathsf{s}}{\mathsf{d}}} N^{\frac{\mathsf{s}}{\mathsf{d}}-1},$$

where $\mathsf{C} > 0$ *is a certain constant to ensure that* $\mathcal{E}_N^t \geq 0$*. Then*

$$\mathcal{E}_N(f_N^t, \mu^t) \leq \mathcal{E}_N(f_N^0, \mu^0) \exp\left(C \int_0^t \|\nabla u^\tau\|_{L^\infty} d\tau\right)$$

for some $C > 0$*. In particular, the RHS is bounded uniformly in* t*.*

Modulated Fisher information

In reality, one should not discard the nonpositive term, as it has a dissipative effect crucial to the long-time behavior. We rewrite it as

$$
\begin{aligned}
&-\frac{1}{\beta^2 N}\mathbb{E}_{f_N^t}\left[\sum_{i=1}^N \left|\nabla \log\left(\frac{f_N^t}{(\mu^t)^{\otimes N}}\right) + \frac{\beta}{N}\sum_{j\neq i}\nabla_1 \mathsf{g}(x_i,x_j) - \nabla \mathsf{g}*\mu^t(x_i)\right|^2\right] \\
&\qquad = -\frac{1}{\beta^2 N}\mathbb{E}_{\mathbb{Q}_{N,\beta}(\mu^t)}\left[\left|\nabla\sqrt{\frac{f_N^t}{\mathbb{Q}_{N,\beta}(\mu^t)}}\right|^2\right] = -\frac{1}{\beta^2} I_N(f_N^t|\mathbb{Q}_{N,\beta}(\mu^t)),
\end{aligned}
$$

where I_N is the normalized relative Fisher information. The dissipation identity transforms into

$$
\begin{aligned}
&\frac{d}{dt}E_N(f_N^t,\mu^t) \le -\frac{1}{\beta^2} I_N(f_N^t|\mathbb{Q}_{N,\beta}(\mu^t)) \\
&-\frac{1}{2}\mathbb{E}_{f_N^t}\left[\int_{(\mathbb{R}^\mathsf{d})^2\setminus\triangle}(u^t(x)-u^t(y))\cdot\nabla_1\mathsf{g}(x,y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\right].
\end{aligned}
$$

Goal: exploit a functional inequality relating the relative Fisher information to the modulated free energy to take advantage of the relative FI term.

Modulated LSI

Given data (g, V, β), we say that a *uniform μ-modulated LSI (μ-LSI)* holds if the family of probability measures $\{\mathbb{Q}_{N,\beta}(\mu)\}_{N\geq 1}$ of the form (31) satisfies a uniform LSI.

Main observation: if $\mathbb{Q}_{N,\beta}(\mu)$ satisfies a uniform LSI, then

$$
\begin{aligned}
I_N(f_N|\mathbb{Q}_{N,\beta}(\mu)) &\geq \frac{1}{C_{LS}} H_N(f_N|\mathbb{Q}_{N,\beta}(\mu)) \\
&= \frac{1}{C_{LS}} \left(\beta E_N(f_N, \mu) - \frac{\log K_{N,\beta}(\mu)}{N} \right).
\end{aligned}
$$

In other words, **a uniform LSI for $\mathbb{Q}_{N,\beta}(\mu)$ implies that the relative Fisher information is bounded below by the modulated free energy and an additive error that is $o_N(1)$ assuming smallness of free energy.**

If this estimate holds for $\mu = \mu^t$ along the mean-field flow, then it can be inserted into the dissipation inequality to obtain an exponential decay of the modulated free energy.

An abstract, general result I

Assume the following for the potential $\mathsf{g} : (\mathbb{R}^{\mathsf{d}})^2 \to [-\infty, \infty]$.

(i) $\mathsf{g} \in C^2((\mathbb{R}^{\mathsf{d}})^2 \setminus \triangle)$ is symmetric and for some $\mathsf{s} < \mathsf{d}$, satisfies

$$|\mathsf{g}(x,y)| \leq C \begin{cases} 1 + |\log|x-y||, & \mathsf{s} = 0 \\ 1 + |x-y|^{-\mathsf{s}}, & \mathsf{s} > 0 \end{cases}$$

for some constant $C > 0$.

Ensures all energy expressions are well-defined and that all differential identities can be justified.

(ii) There exists a constant $C_\beta \in [0, \frac{1}{\beta})$ such that for any $f_N \in \mathcal{P}_{ac}((\mathbb{R}^{\mathsf{d}})^N)$ and $\mu \in \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}})$, with $\int_{\mathbb{R}^{\mathsf{d}}} \log(1+|x|) d\mu(x) < \infty$ if $s = 0$,

$$\mathbb{E}_{f_N}\left[F_N(X_N, \mu)\right] \geq -C_\beta H_N(f_N|\mu^{\otimes N}) - o_N(1),$$

where $o_N(1)$ only depends (in an increasing fashion) on μ through $\|\mu\|_{L^\infty}$.

Ensures that the modulated free energy is nonnegative up to $o_N(1)$ error. Note this does not come for free, since we make no sign assumptions on g. In fact, it shows that the modulated free energy controls the relative entropy.

An abstract, general result II

(iii) There exist constants $C_{RE}, C_{ME} \geq 0$, such that

$$\left| \mathbb{E}_{f_N} \left[\int_{(\mathbb{R}^d)^2 \setminus \triangle} (v(x) - v(y)) \cdot \nabla_1 g(x,y) d \left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) \right] \right|$$
$$\leq \|v\|_* \left(C_{RE} H_N(f_N | \mu^{\otimes N}) + C_{ME} \mathbb{E}_{f_N} \left[F_N(X_N, \mu) \right] + o_N(1) \right)$$

for all pairwise distinct configurations $X_N \in (\mathbb{R}^d)^N$, densities $f_N \in \mathcal{P}_{ac}((\mathbb{R}^d)^N)$ and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, and continuous vector fields v with finite homogeneous Sobolev norm $\| \cdot \|_*$ of some order.

An abstract, general result III

Let us introduce the quantity

$$\mathcal{E}_N^t := E_N(f_N^t, \mu^t) + o_N^t(1)$$

as a substitute for the modulated free energy.

- Additive error $o_N^t(1)$ ensures that $\mathcal{E}_N^t \geq 0$, which allows to perform a Grönwall argument on this quantity.
- Depends only on μ^t through the L^∞ norm, hence the t superscript, and is increasing in this dependence.

An abstract, general result IV

Theorem 3 (R.-Serfaty 2023)

Let $\beta > 0$. Assume that the mean-field equation admits a solution $\mu \in C([0,\infty), \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}}))$, such that $\|\mu^t\|_{L^\infty}$ is bounded uniformly in t and $\nabla u^t \in L^\infty$ locally uniformly in t. If $s = 0$, further assume that $\int_{\mathbb{R}^{\mathsf{d}}} \log(1+|x|) d\mu^t < \infty$ for every $t \geq 0$. If $\mathbb{Q}_{N,\beta}(\mu^t)$ satisfies a uniform LSI with constant $C_{LS} > 0$ for every $t \geq 0$, then

$$\forall t \geq 0, \qquad \mathcal{E}_N^t \leq e^{-\frac{4t}{\beta C_{LS}} + \int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2} d\tau} \mathcal{E}_N^0 + e^{-\frac{4t}{\beta C_{LS}} + \int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2} d\tau} \int_0^t e^{\frac{4\tau}{\beta C_{LS}} - \int_0^\tau \frac{\mathsf{C}\|u^{\tau'}\|_*}{2} d\tau'} \times \left[\dot{o}_N^\tau + \frac{4}{\beta C_{LS}} \left(o_N^\tau(1) - \frac{\log K_{N,\beta}(\mu^\tau)}{\beta N} \right) \right] d\tau,$$

where $o_N^\tau(1)$ is as above, and $\dot{o}_N^\tau(1)$ denotes the derivative of $o_N^\tau(1)$ with respect to time.

An abstract, general result V

- Provided $\int_0^\infty \|u^\tau\|_* d\tau < \infty$, the first term on the RHS converges exponentially fast to 0 as $t \to \infty$.
- The second term is $o_N(1)$ uniformly bounded in t, assuming $\log K_{N,\beta}(\mu^\tau) = o(N)$ uniformly in τ and that $\int_0^\infty |\dot{o}_N^\tau(1)| d\tau < \infty$, by the fundamental theorem of calculus and our assumption that $\|\mu^t\|_{L^\infty}$ is uniformly bounded.
- Since $\mathcal{E}_N$ differs from E_N only by additive constants which are $o_N(1)$, and the modulated free energy E_N controls the relative entropy H_N, our estimate implies entropic generation of chaos and also gives a uniform-in-time propagation of chaos if $\mathcal{E}_N^0 = o_N(1)$.

In repulsive Riesz, attractive log, and bounded continuous cases, assumptions (i)-(iii) hold and a more precise form of the main estimate is available.

Cf. previous work on generation of chaos Lacker-Le Flem 2023 (entropic/exp integrable ∇g), Guillin-Le Bris-Monmarché 2021 (entropic/log/conservative), Guillin-Le Bris-Monmarché 2023 (Wasserstein/1D Riesz).

Given a density μ, recall that

$$\mathbb{Q}_{N,\beta}(\mu) = \mathbb{P}_{N,\beta}^{V_{\mu,\beta}}, \qquad \text{where} \quad V_{\mu,\beta} := -\mathsf{g} * \mu - \frac{1}{\beta} \log \mu.$$

Proposition 4

Suppose that $\mu \in \mathcal{P}(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and if $s = 0$, also suppose that $\int \log(1 + |x|) d\mu < \infty$. For $\beta > 0$, suppose that $V_{\mu,\beta}$ is κ-convex, for some $\kappa > 0$. Then the probability measure $\mathbb{Q}_{N,\beta}(\mu)$ has LSI constant $\frac{2}{\beta\kappa}$.

Let $\mu \in \mathcal{P}(\mathbb{R})$ be such that $\log \frac{\mu}{\mu_\beta} \in C^2(\mathbb{R})$. If $V \in C^2$ and $\beta > 0$, then $V_{\mu,\beta}$ is κ-convex with

$$\kappa(\mu) := \inf V'' - \left(\frac{1}{\beta} \| \log \frac{\mu}{\mu_\beta} \|_{\dot{C}^2} + \| \mathsf{g} * (\mu - \mu_\beta) \|_{\dot{C}^2} \right).$$

One can produce $\mu \in \mathcal{P}(\mathbb{R})$ satisfying $\kappa(\mu) > 0$ by choosing $h \in C^2$, setting $\mu := \frac{e^h \mu_\beta}{\int e^h \mu_\beta}$, and taking $\|e^h - 1\|_{C^2}$ arbitrarily small.

Modulated LSI for 1D Riesz case II

$\left\|\log \frac{\mu^t}{\mu_\beta}\right\|_{\dot{C}^2} \to 0$ as $t \to \infty$ J. Huang-R.-Serfaty. So, always exists t_0 such that $\inf_{t \geq t_0} \kappa(\mu^t) > 0$. Then $\mathbb{Q}_{N,\beta}(\mu^t)$ satisfied a uniform LSI with constant independent of t on $[t_0, \infty)$. Combine with propagation of chaos estimate on $[0, t_0]$ to deduce *entropic generation of chaos*.

- Proposition generalizes uniform LSI for 1D log Gibbs measure Chafaï-Lehec 2020
- Proof relies on the convexity of the Riesz potential g on $[0, \infty)$ (only true in 1D). Allows to treat the interaction as a perturbation of the confinement.
- Ordering of particles allows to show the total Hamiltonian is uniformly convex.
- Use contraction theorem Caffarelli 2000 for bounding the Lipschitz seminorm of the Brenier map between two probability measures solely in terms of the convexity constant. Reduces to applying LSI for Gaussian measure Gross 1975.

Current and future directions

Current directions

Modulated energy/free energy combined with these commutator-type functional inequalities are powerful tools for studying the large N behavior of these systems. When combined with analysis of mean-field equation, effective also for large time behavior.

- Uniform-in-time propagation of chaos for attractive case with mildly singular g (e.g., $\mathrm{g} = \log|x|$, Patlak-Keller-Segel). Chodron de Courcel-R.-Serfaty.
- Gaussian CLT for linear statistics/cumulant bounds J. Huang-R.-Serfaty
- Dynamical LDP Hess-Childs

Outlook I

Despite progress, a number of questions remain...

- Sharp functional inequalities in the sub-Coulomb regime $0 \leq s < d - 2$? The singularity of the potential is milder, but the decay at infinity is slower.
- Modulated LSI beyond $d = 1$?
- The case with positive and negative charges (e.g., two-component plasma)? Our methods rely on the density having a definite sign and also when we work at the level of the SDEs, there are no collisions between particles. However, it is known that collisions may occur if the particles are not identically signed.

Outlook II

- What about second-order systems? The prototypical microscopic system is

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = -\frac{1}{N} \sum_{\substack{1 \le j \le N \\ j \neq i}} \nabla \mathsf{g}(x_i - x_j), \end{cases}$$

for which the expected mean-field PDE is the *Vlasov equation*

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = 0 \\ \rho = \int_{\mathbb{R}^\mathsf{d}} df(\cdot, v) \\ E = -\nabla \mathsf{g} * \rho, \end{cases} \qquad (x, v) \in \Omega \times \mathbb{R}^\mathsf{d}.$$

Duerinckx-Serfaty 2020 treated the monokinetic case $f(x, v) = \rho(x)\delta(v - u(x))$, which is amenable to the modulated energy method. But otherwise, this problem is essentially open.

The End

Thank you for your attention!