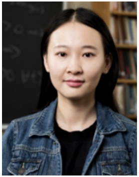
Nonparametric inference of interaction laws in particle/agent systems

Fei Lu

Department of Mathematics, Johns Hopkins University

Joint with:

Mauro Maggioni



Sui Tang

Ming Zhong

July 11, 2019
Applied Math and Comp Sci Colloquium
University of Pennsylvania

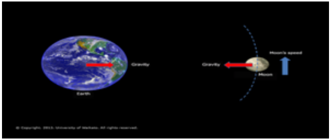
FL acknowledges supports from JHU, NSF

Outline

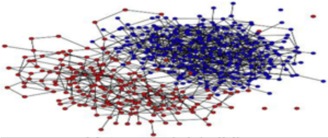
1. Motivation and problem statement
2. Learning via nonparametric regression
3. Numerical examples
4. Ongoing work and open problems

Motivation

Q: What is the law of interaction between particles/agents?



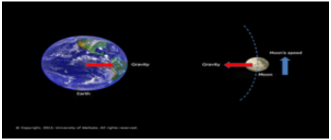
Popkin. Nature(2016)



Voter model (wiki)

Motivation

Q: What is the law of interaction between particles/agents?



$$m\ddot{x}_i(t) = -\nu \dot{x}_i(t) + \frac{1}{N}\sum_{j=1, j\neq i}^{N} K(x_i, x_j),$$

- Newton's law of gravitation:

$$K(x,y) = G\frac{m_1 m_2}{r^2}, r = |x-y|$$

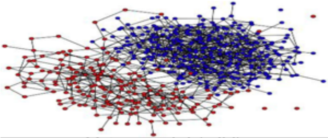
Popkin. Nature(2016)

- Molecular fluid: $K(x,y) = \nabla_x[\Phi(|x-y|)]$
  Lennard-Jones potential: $\Phi(r) = \frac{c_1}{r^{12}} - \frac{c_2}{r^6}$.

- flocking birds/school of fish

$$K(x,y) = \phi(|x-y|)\frac{x-y}{|x-y|}$$

- opinion/voter models, bacteria/cells ...[a]



Voter model (wiki)

[a] (1) Cucker+Smale: On the mathematics of emergence. 2007. (2) Vicsek+Zafeiris: Collective motion. 2012. (3) Mostch+Tadmor: Heterophilious Dynamics Enhances Consensus. 2014 ...

An inference problem:

Infer the rule of interaction in the system

$$m\ddot{x}_i(t) = -\nu\dot{x}_i(t) + \frac{1}{N}\sum_{j=1, j\neq i}^{N} K(x_i - x_j), \quad i = 1, \cdots, N, \; x_i(t) \in \mathbb{R}^d$$

from observations of trajectories.

- x_i is the position of the i-th particle/agent
- **Data**: many independent trajectories $\{\boldsymbol{x}^j(t) : t \in \mathcal{T}\}_{j=1}^{M}$
- **Goal**: infer $\phi : \mathbb{R}+ \to \mathbb{R}$ in

$$K(x) = -\nabla\Phi(|x|) = -\phi(|x|)\frac{x}{|x|}$$

For simplicity, we consider only first-order systems ($m = 0$) $\downarrow$

$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi_{true}(|x_i - x_j|)\frac{x_j - x_i}{|x_j - x_i|} \quad \rightarrow \quad \dot{\boldsymbol{x}} = \mathbf{f}_{\phi_{true}}(\boldsymbol{x}(t))$$

Least squares regression: with $\mathcal{H}_n = \mathrm{span}\{e_i\}_{i=1}^n$,

$$\hat{\phi}_n = \arg\min_{\phi \in \mathcal{H}_n} \mathcal{E}_M(\phi) := \sum_{m=1}^{M} |||\dot{\boldsymbol{x}}^m - \mathbf{f}_{\phi}(\boldsymbol{x}^m)|||^2$$

- Choice of $\mathcal{H}_n$ & function space of learning?
- Inverse problem well-posed/ identifiability?
- Consistency and rate of “convergence”?
  $\rightarrow$ hypothesis testing and model selection

Outline

1. Motivation and problem statement
2. Learning via nonparametric regression:
   - Function space of regression
   - Identifiability: a coercivity condition
   - Consistency and rate of convergence
3. Numerical examples
4. Ongoing work and open problems

Learning via nonparametric regression

The dynamical system:

$$\dot{x} = \mathbf{f}_{\phi_{true}}(\boldsymbol{x}(t))$$

Data: M-trajectories $\{\boldsymbol{x}^m(t) : t \in \mathcal{T}\}_{m=1}^{M}$

- $\boldsymbol{x}^m(0) \overset{i.i.d}{\sim} \mu_0 \in \mathcal{P}(\mathbb{R}^{dN})$
- $\mathcal{T} = [0, T]$ or $\{t_1, \cdots, t_L\}$ with $\dot{\mathbf{x}}(t_i)$

Goal: nonparametric inference[1] of ϕ_{true}

[1] (1) Bongini, Fornasier, Hansen, Maggioni: Inferring Interaction Rules for mean field equations, M3AS, 2017.
(2) Binev, Cohen, Dahmen, Devore and Temlyakov: Universal Algorithms for learning theory, JMLR 2005.
(3) Cucker, Smale: On the mathematical foundation of learning. Bulletin of AMS, 2001.

$$\hat{\phi}_{M,\mathcal{H}} = \underset{\phi \in \mathcal{H}}{\arg\min}\, \mathcal{E}_M(\phi) := \frac{1}{ML} \sum_{l,m=1}^{L,M} \|\mathbf{f}_{\phi}(\boldsymbol{X}^m(t_l)) - \dot{\boldsymbol{X}}^m(t_l)\|^2$$

- $\mathcal{E}_M(\phi)$ is quadratic in ϕ, and $\mathcal{E}_M(\phi) \geq \mathcal{E}_M(\phi_{true}) = 0$
- The minimizer exists for any $\mathcal{H} = \mathcal{H}_n = span\{e_1, \ldots, e_n\}$

Tasks

- Choice of $\mathcal{H}_n$ & function space of learning?
- Inverse problem well-posed/ identifiability?
- Consistency and rate of "convergence"?

$$\begin{tikzcd}
\mathcal{E}_M(\cdot) \arrow[r, "M\to\infty"] \arrow[d] & \mathcal{E}_\infty(\cdot) \arrow[d] \\
\widehat{\phi}_{M,\mathcal{H}} \arrow[r, "?M\to\infty"] \arrow[dr, "??"] & \widehat{\phi}_{\infty,\mathcal{H}} \arrow[d, "?dist(\mathcal{H},\phi_{true})\to 0"] \\
 & \phi_{true}
\end{tikzcd}$$

Review of classical nonparametric regression:

Estimate $y = \phi(z) : \mathbb{R}^D \to \mathbb{R}$ from data $\{z_i, y_i\}_{m=1}^M$.

- $\{z_i, y_j\}$ are iid samples;
- $\hat{\phi}_n := \arg\min_{f \in \mathcal{H}_n} \mathcal{E}_M(f) := \sum_{m=1}^M \|y_i - f(z_i)\|^2 \quad \to \mathbb{E}[Y|Z = z]$
- Optimal rate: if $\mathrm{dist}(\mathcal{H}_n, \phi_{true}) \lesssim n^{-s}$ and $n_* = (M/\log M)^{\frac{1}{2s+1}}$,
$$\|\hat{\phi}_{n_*} - \phi\|_{L^2(\rho_Z)} \lesssim M^{-\frac{s}{2s+D}}$$

[2]

[2] (1) F.Cucker and S.Smale. On the mathematical foundations of learning. Bulletin of the AMS, 2002
(2) L.Györfi, M.Kohler, A.Krzyzak, H.Walk, A Distribution-Free Theoryof Nonparametric Regression (Springer 2002).

Review of classical nonparametric regression:

Estimate $y = \phi(z) : \mathbb{R}^D \to \mathbb{R}$ from data $\{z_i, y_i\}_{m=1}^M$.

- $\{z_i, y_j\}$ are iid samples;
- $\hat{\phi}_n := \underset{f\in\mathcal{H}_n}{\arg\min}\, \mathcal{E}_M(f) := \sum_{m=1}^M \|y_i - f(z_i)\|^2$
- Optimal rate: if $\mathrm{dist}(\mathcal{H}_n, \phi_{true}) \lesssim n^{-s}$ and $n_* = (M/\log M)^{\frac{1}{2s+1}}$,
$$\|\hat{\phi}_{n_*} - \phi\|_{L^2(\rho_Z)} \lesssim M^{-\frac{s}{2s+D}}$$

Our case: learning of kernel $\phi : \mathbb{R}^+ \to \mathbb{R}$ from data $\{\boldsymbol{x}^m(t)\}$

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|) \frac{x_j - x_i}{|x_j - x_i|}$$

- $\{r_{ij}^m(t) := |x_i^m(t) - x_j^m(t)|\}$ not iid
- The values of $\phi(r_{ij}^m(t))$ unknown

Regression measure

Distribution of pairwise-distances $\rho : \mathbb{R}_+ \to \mathbb{R}$

$$\rho_T(r) = \frac{1}{\binom{N}{2} L} \sum_{l,i,i'=1,i<i'}^{L,N} \mathbb{E}_{\mu_0} \delta_{r_{ii'}(t_l)}(r)$$

- unknown, estimated by empirical distribution $\rho_T^M \xrightarrow{M\to\infty} \rho_T$ (LLN)
- intrinsic to the dynamics

Regression function space $L^2(\rho_T)$

- the admissible set $\subset L^2(\rho_T)$
- $\mathcal{H} =$ piecewise polynomials $\subset L^2(\rho_T)$
- singular kernels $\subset L^2(\rho_T)$

Identifiability: a coercivity condition

$$\begin{tikzcd}
\mathcal{E}_M(\cdot) \arrow[r, "M\to\infty"] \arrow[d] & \mathcal{E}_\infty(\cdot) \arrow[d] \\
\widehat{\phi}_{M,\mathcal{H}} \arrow[r, "?M\to\infty"] \arrow[dr, "?"] & \widehat{\phi}_{\infty,\mathcal{H}} \arrow[d, "?"] \\
& \phi_{true}
\end{tikzcd}$$

$$\hat{\phi}_{M,\mathcal{H}} = \underset{\phi\in\mathcal{H}}{\arg\min}\ \mathcal{E}_M(\phi)$$

$$\mathcal{E}_\infty(\hat{\phi}) - \mathcal{E}_\infty(\phi_{true}) = \frac{1}{NT}\int_0^T \mathbb{E}_{\mu_0}\|\mathbf{f}_{\hat{\phi}-\phi_{true}}(\boldsymbol{X}(t))\|^2 dt \geq c\|\hat{\phi} - \phi_{true}\|^2_{L^2(\rho_T)}$$

Coercivity condition. $\exists\ c_{T,\mathcal{H}} > 0$ s.t. for all $\varphi \in \mathcal{H} \subset L^2(\rho_T)$

$$\frac{1}{NT}\int_0^T \mathbb{E}_{\mu_0}\|\mathbf{f}_\varphi(\boldsymbol{x}(t))\|^2 dt = \langle\langle \varphi, \varphi \rangle\rangle \geq c_{T,\mathcal{H}}\|\varphi\|^2_{L^2(\rho_T)}$$

- coercivity: bilinear functional $\langle\langle \varphi, \psi \rangle\rangle := \frac{1}{NT}\int_0^T \mathbb{E}_{\mu_0}\langle \mathbf{f}_\varphi, \mathbf{f}_\psi\rangle(\boldsymbol{x}(t))dt$
- controls condition number of regression matrix

Consistency of estimator

Theorem (L., Maggioni, Tang, Zhong)

Assume the coercivity condition. Let $\{\mathcal{H}_n\}$ be a sequence of compact convex subsets of $L^\infty([0,R])$ such that $\inf_{\varphi \in \mathcal{H}_n} \|\varphi - \phi_{true}\|_\infty \to 0$ as $n \to \infty$. Then

$$\lim_{n\to\infty} \lim_{M\to\infty} \|\widehat{\phi}_{M,\mathcal{H}_n} - \phi_{true}\|_{L^2(\rho_T)} = 0, \text{ almost surely.}$$

- For each n, compactness of $\{\widehat{\phi}_{M,\mathcal{H}_n}\}$ and coercivity implies that $\widehat{\phi}_{M,\mathcal{H}_n} \to \widehat{\phi}_{\infty,\mathcal{H}_n}$ in L^2
- Increasing $\mathcal{H}_n$ and coercivity implies consistency.
- In general, truncation to make $\mathcal{H}_n$ compact

Optimal rate of convergence

Theorem (L. Maggioni, Tang, Zhong)

Let $\{\mathcal{H}_n\}$ be a seq. of compact convex subspaces of $L^\infty[0,R]$ s.t.

$$\dim(\mathcal{H}_n) \le c_0 n, \ \textit{and} \ \inf_{\varphi\in\mathcal{H}_n} \|\varphi - \phi_{true}\|_\infty \le c_1 n^{-s}.$$

Assume the coercivity condition. Choose $n_* = (M/\log M)^{\frac{1}{2s+1}}$: then

$$\mathbb{E}_{\mu_0}[\|\widehat{\phi}_{T,M,\mathcal{H}_{n_*}} - \phi_{true}\|_{L^2(\rho_T)}] \le C\left(\frac{\log M}{M}\right)^{\frac{s}{2s+1}}.$$

- The 2nd condition is about regularity: $\phi \in C^s$
- Choice of $\dim(\mathcal{H}_n)$: adaptive to s and M

Prediction of future evolution

Theorem (L., Maggioni, Tang, Zhong)

Denote by $\widehat{\boldsymbol{X}}(t)$ and $\boldsymbol{X}(t)$ the solutions of the systems with kernels $\widehat{\phi}$ and ϕ respectively, starting from the same initial conditions that are drawn i.i.d from μ_0. Then we have

$$\mathbb{E}_{\mu_0}[\sup_{t\in[0,T]} \|\widehat{\boldsymbol{X}}(t) - \boldsymbol{X}(t)\|^2] \lesssim \sqrt{N}\|\widehat{\phi} - \phi_{true}\|^2_{L^2(\rho_T)},$$

- Follows from Grownwall's inequality

Outline

1. Motivation and problem statement
2. Learning via nonparametric regression:
   - A regression measure and function space
   - Learnability: a coercivity condition
   - Consistency and rate of convergence
3. Numerical examples
   - A general algorithm
   - Lennard-Jones model
   - Opinion dynamics and multiple-agent systems
4. Ongoing work and open problems

The regression algorithm

$$\mathcal{E}_M(\varphi) = \frac{1}{LMN} \sum_{l,m,i=1}^{L,M,N} \left\| \dot{\boldsymbol{x}}_i^{(m)}(t_l) - \sum_{i'=1}^{N} \frac{1}{N} \varphi(r_{i,i'}^m(t_l)) \boldsymbol{r}_{i,i'}^m(t_l) \right\|^2 ,$$

$$\mathcal{H}_n := \{\varphi = \sum_{p=1}^{n} a_p \psi_p(r) : \mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n\},$$

$$\mathcal{E}_{L,M}(\varphi) = \mathcal{E}_{L,M}(\mathbf{a}) = \frac{1}{M} \sum_{m=1}^{M} \|\mathbf{d}^m - \Psi_L^m \mathbf{a}\|_{\mathbb{R}^{LNd}}^2 .$$

$$\frac{1}{M} \sum_{m=1}^{M} A_L^m \mathbf{a} = \frac{1}{M} \sum_{m=1}^{M} b_L^m , \text{ rewrite as } A_M \mathbf{a} = b_M$$

- can be computed parallelly
- Caution: choice of $\{\psi_p\}$ affects condi(A_M)

Assume the coercivity condition: $\langle\!\langle \varphi, \varphi \rangle\!\rangle \geq c_{T,\mathcal{H}} \|\varphi\|^2_{L^2(\rho_T)}$.

Proposition (Lower bound on smallest singular value of A_M)

Let $\{\psi_1, \cdots, \psi_n\}$ be a basis of $\mathcal{H}_n$ s.t.

$$\langle \psi_p, \psi_{p'} \rangle_{L^2(\rho_T^L)} = \delta_{p,p'}, \|\psi_p\|_\infty \leq S_0.$$

Let $A_\infty = \left(\langle\!\langle \psi_p, \psi_{p'} \rangle\!\rangle\right)_{p,p'} \in \mathbb{R}^{n \times n}$. Then $\sigma_{\min}(A_\infty) \geq c_{T,\mathcal{H}}$. Moreover, A_∞ is the a.s. limit of A_M. Therefore, for large M, the smallest singular value of A_M satisfies with a high probability that

$$\sigma_{\min}(A_M) \geq (1 - \epsilon) c_{T,\mathcal{H}}$$

- Choose $\{\psi_p\}$ linearly independent in $L^2(\rho_T)$
- Piecewise polynomials: on a partition of support(ρ_T)
- Finite difference $\approx$ derivatives $\Rightarrow$ an $O(\Delta t)$ error to estimator

Implementation

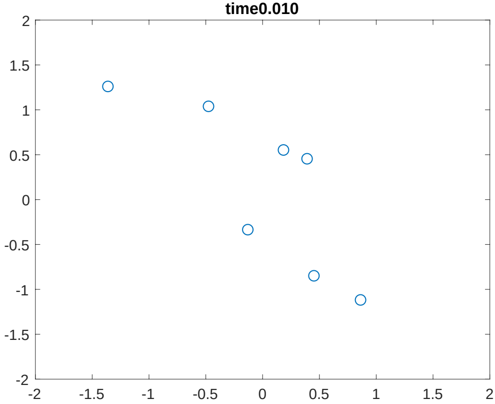
1. Approximate regression measure
   - Estimate the ρ_T with large datasets
   - Partition on support(ρ_T)
2. Construct hypothesis space $\mathcal{H}$:
   - choose the degree of piecewise polynomials
   - set dimension of $\mathcal{H}$ according to sample size
3. Regression:
   - Assemble the arrays (in parallel)
   - Solve the normal equation

Examples: Lennard-Jones Dynamics

The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

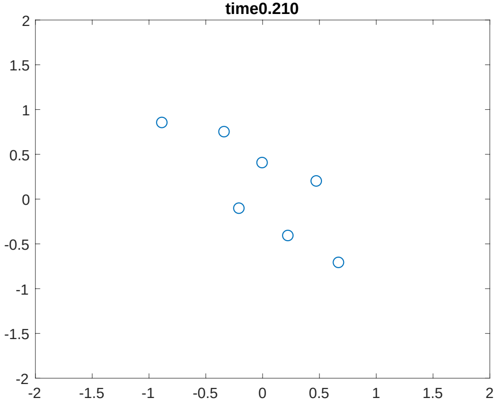
$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|)(x_j - x_i)$$



The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|)(x_j - x_i)$$

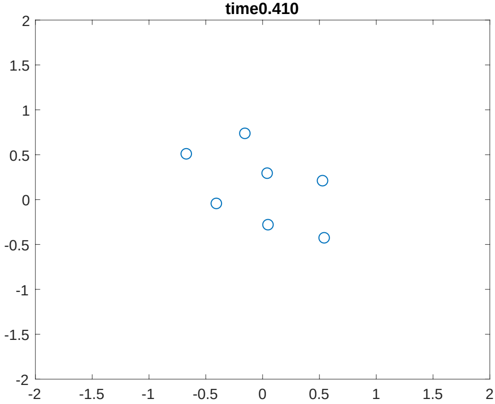


Examples: Lennard-Jones Dynamics

The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

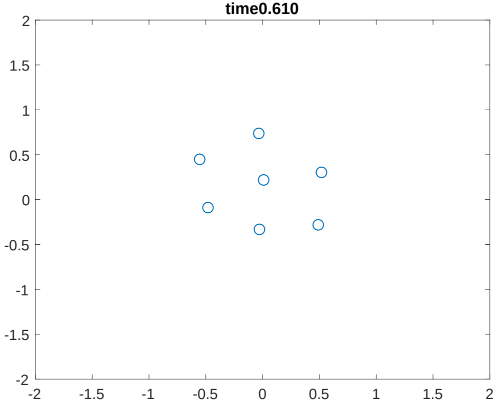
$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|)(x_j - x_i)$$



The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

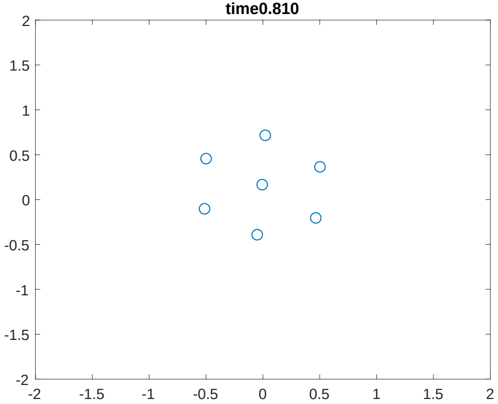
$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|)(x_j - x_i)$$



The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V_{LJ}'(r)$$

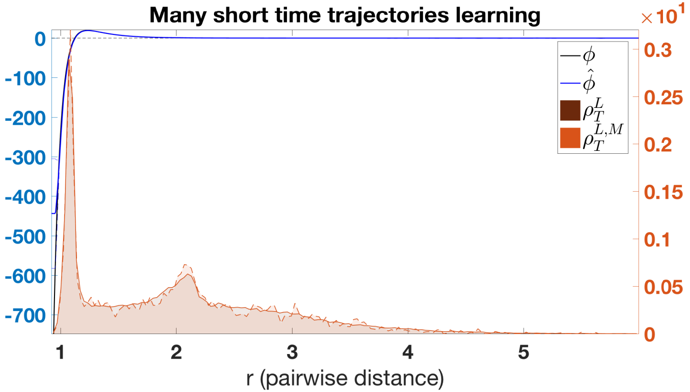
$$\dot{x}_i(t) = \frac{1}{N}\sum_{j=1, j\neq i}^{N} \phi(|x_i - x_j|)(x_j - x_i)$$



The Lennard-Jones potential

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

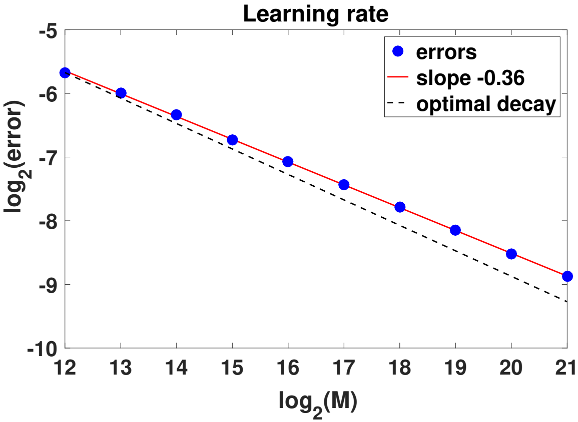
- piecewise linear estimator; Gaussian initial conditions.



Optimal rate

$$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \Rightarrow \quad \phi(r)r = V'_{LJ}(r)$$

- V_{LJ} is highly singular, yet we get close to optimal rate (-0.4).



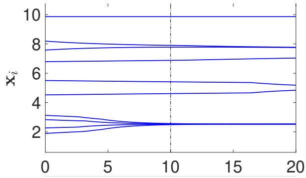
Example: Opinion Dynamics

$N = 10, \boldsymbol{x}_i \in \mathbb{R}$.

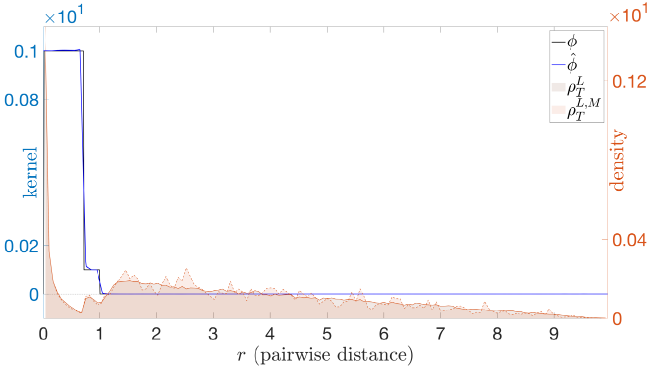
$M = 250, \mu_0 = Unif[0, 10]^{10}$

$\mathcal{T} = [0, 10]$, 200 discrete instances

$\mathcal{H} =$ piecewise constant functions



The estimated kernels:



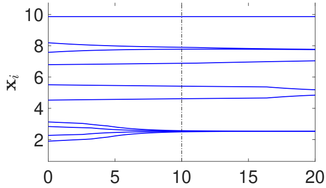
Example: Opinion Dynamics

$N = 10, \boldsymbol{x}_i \in \mathbb{R}$.

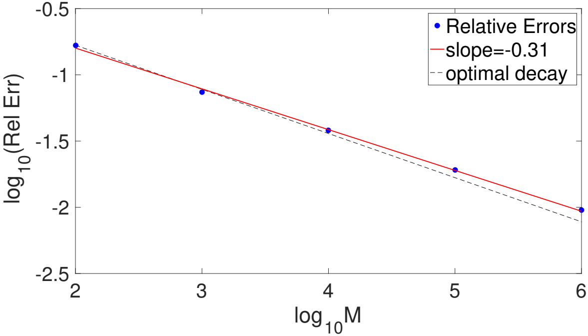
$M = 250, \mu_0 = Unif[0, 10]^{10}$

$\mathcal{T} = [0, 10]$, 200 discrete instances

$\mathcal{H} =$ piecewise constant functions

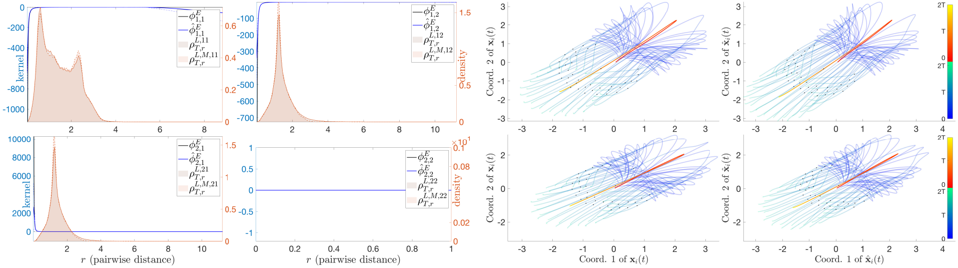


The rate of convergence:



Example: 2nd-order Prey-Predator system

$$\begin{cases} \ddot{\boldsymbol{X}}^m = \mathcal{F}^{\boldsymbol{v}}(\dot{\boldsymbol{X}}^m, \Xi^m) + \mathbf{f}_{\boldsymbol{\phi}^E}(\boldsymbol{X}^m) + \mathbf{f}_{\boldsymbol{\phi}^A}(\boldsymbol{X}^m, \dot{\boldsymbol{X}}^m) \\ \dot{\Xi}^m = \mathcal{F}^{\xi}(\Xi^m) + \mathbf{f}_{\boldsymbol{\phi}^{\xi}}(\boldsymbol{X}^m, \Xi^m), \end{cases}$$

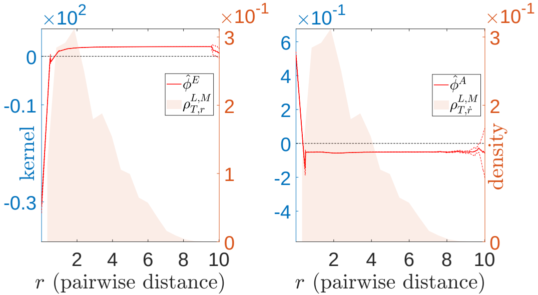


Example: model selection

- **Order selection**

	Learned as 1^{st} order	Learned as 2^{nd} order
1^{st} order system	**0.01** $\pm$ **0.002**	1.6 $\pm$ 1.1
2^{nd} order system	1.7 $\pm$ 0.3	**0.2** $\pm$ **0.06**

- **Interaction type selection**



Summary and open problems

$$\begin{tikzcd}
\mathcal{E}_{T,M}(\cdot) \arrow[r] \arrow[d] & \mathcal{E}_{T,\infty}(\cdot) \arrow[d] \\
\widehat{\phi}_{T,M,\mathcal{H}} \arrow[r, "M\to\infty"] \arrow[dr, "optimal\ \mathcal{H}"] & \widehat{\phi}_{T,\infty,\mathcal{H}} \arrow[d, "dist(\mathcal{H},\phi)\to 0"] \\
 & \phi
\end{tikzcd}$$

Learning theory

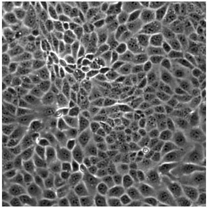
- extended the classical regression theory
- a coercivity condition for identifiability

Theory guided regression algorithms

- Selection of $\mathcal{H}$ (basis functions & dimension)
- Measurement of error of estimators
- Optimal learning rate
- Model selection

Ongoing work

- Different type of systems:
  - 1st- and 2nd-order
  - Multiple type of agents (leader-follower, predator-prey)
  - Stochastic systems
- Coercivity condition
- Adaptive basis functions
- Partial and noisy observations; Mean field equations
- Real data applications:
  learning cell-dynamics



Ongoing work: The coercivity condition

$$\langle\!\langle \varphi, \varphi \rangle\!\rangle \geq c_{\mathcal{H}}^{T} \|\varphi\|_{L^2(\rho_T)}^2, \mathcal{H} \text{ compact}$$

Exchangeability, $g(r) = \phi(r)r \Updownarrow U_t = x_1(t) - x_2(t),\, V_t = x_1(t) - x_3(t)$

$$\int_0^T \underbrace{\mathbb{E}\left[g(|U_t|)g(|V_t|)\frac{\langle U_t, V_t\rangle}{|U_t||V_t|}\right]}_{\int_{\mathbb{R}}^{+}\int_{\mathbb{R}}^{+} g(r)g(s)\mathcal{K}_t(r,s)drds} dt > 0$$

Proposition (Li-Lu19)

Coercivity condition holds for systems with $\Phi(r) = r^{\beta}$, $\beta \in [1,2]$.

- positiveness of integral operator $\leftrightarrow \mathcal{K}(\cdot,\cdot) := \int_0^T \mathcal{K}_t(\cdot,\cdot)dt$
  - Müntz type theorem: $\mathrm{span}\{r^{2n}e^{-r}\}_{n=1}^{\infty}$ dense in $L^2(\mathbb{R}^+)$.
- **Conjecture:** true for general systems

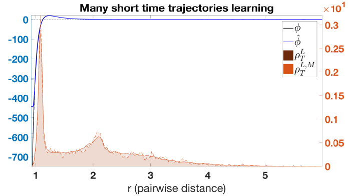
Ongoing work: Adaptive basis functions $\{\psi_p\}$

$$A_M \mathbf{a} = b_M, \text{ with } \mathbb{E}[A_M] = \Big(\underbrace{\langle\langle \psi_p, \psi_{p'} \rangle\rangle}_{\int_{\mathbb{R}}^{+} \int_{\mathbb{R}}^{+} \psi_p(r)\psi_{p'}(s)\mathcal{K}(r,s)drds} \Big)_{p,p' \in 1,\dots,n}$$

Current: piecewise polynomials + uniform partition supp($\bar{\rho}$)

Adaptive strategies:

- Adaptive partition based on $\bar{\rho}$
- Eigenfunctions of integral kernel $\mathcal{K}$
  - $\widehat{\mathcal{K}}$ from data: noisy
  - goal: smooth eigenfunctions



References

- F. Lu, M. Maggioni, and S. Tang. Learning interaction rules in particle/agent systems: the stochastic case. In preparation
- Z. Li, F. Lu, S. Tang, C. Zhang, and M. Maggioni. On the identifiability of interaction functions of particle systems. preprint
- F. Lu, M. Maggioni, and S. Tang. Learning interaction kernels in heterogeneous systems of agents from multiple trajectories. arXiv.
- F. Lu, M. Maggioni, S. Tang and M. Zhong. Nonparametric inference of interaction laws in systems of agents from trajectory data. PNAS, 2019
- M. Bongini, M. Fornasier, M. Maggioni and M. Hansen. Inferring Interaction Rules From Observations of Evolutive Systems I: The Variational Approach. M3AS, 27(05), 909-951, 2017

Thank you!