Learning dynamical systems: an example from open quantum system dynamics.

Pietro Novelli
Istituto Italiano di Tecnologia
pietro.novelli@iit.it

Abstract

Machine learning algorithms designed to learn dynamical systems from data can be used to forecast, control and interpret the observed dynamics. In this work we exemplify the use of one of such algorithms, namely Koopman operator learning, in the context of open quantum system dynamics. We will study the dynamics of a small spin chain coupled with dephasing gates and show how Koopman operator learning is an approach to efficiently learn not only the evolution of the density matrix, but also of *every physical observable* associated to the system. Finally, leveraging the spectral decomposition of the learned Koopman operator, we show how symmetries obeyed by the underlying dynamics can be inferred directly from data.

1 A primer on Koopman operator learning

The Koopman operator is a mathematical object concerned with the evolution of a dynamical system. Formally it is defined as the linear operator evolving any *observable*[1] in a pre-specified set. It owes its name to the works of Bernard Koopman on the evolution of Hamiltonian systems [Koo31], but current applications of the Koopman operator embrace the whole field of dynamical systems [BMM12, MMS20, BBKK22]. We now provide a minimalistic introduction to the Koopman operator, referring to [MMS20, KNM$^+$22] for a mathematically rigorous presentation.

Given a discrete dynamical system over a state space $\mathcal{X}$, the Koopman operator U is defined as

$$(Uf)(x_t) := f(x_{t+1}) \quad \text{for all } t \in \mathbb{N},$$

for every observable $f : \mathcal{X} \to \mathbb{R}$ in a suitable set. The Koopman operator, therefore, acts by evolving the observables of the system by one step of the dynamics. The above definition is appropriate for deterministic dynamical systems, but can be consistently extended to stochastic systems via conditional expectation operators [MT93]. In this work we hinge on the formulation of Koopman operator learning in [KNM$^+$22], in which the space of observables is restricted to a reproducing kernel Hilbert space $\mathcal{H}$ [Aro50, SC08]. Within this setting, an estimator $\hat{U} : \mathcal{H} \to \mathcal{H}$ of the Koopman operator can be *learned* from a single dataset of observations $(x_i, y_i)_{i=1}^n$ consisting of sampled states $x_i \in \mathcal{X}$ along with their one-step-forward evolutions $y_i \in \mathcal{X}$.

An obvious use of the learned operator $\hat{U}$ is forecasting both future states and *every observable* in $\mathcal{H}$ (see Remark 1). The linearity of $\hat{U}$, however, also allows to compute its spectral decomposition, and if $(\lambda_i, \xi_i, \psi_i)_{i=1}^r$ are eigen-triplets[2] of $\hat{U}$, the evolution of an observable $f : \mathcal{X} \to \mathbb{R}$ can be dissected as a sum of *modes*

$$(\hat{U}f)(x_t) = \sum_{i=1}^{r} \lambda_i^t \gamma_i^f \psi_i(x_0), \tag{1}$$

[1] That is, a scalar function of the state of the dynamical system.

[2] Namely: eigenvalues, left and right eigenfunctions.

Machine Learning and the Physical Sciences workshop, NeurIPS 2022.

where γ_i^f is a linear functional of the observable f linked to the i-th left eigenfunction ξ_i. The Koopman mode decomposition [BMM12, AM17] can be used to infer properies of the dynamical system such as constant of motions, correlation and dephasing between different observables, mixing times, normal modes, metastable states and many others. These quantities are of paramount importance in the physical sciences and machine-learning-enabled tools to compute the Koopman mode decomposition unlock the possibility to study complex physical systems too hard to be approached with analytical theories alone.

Remark 1 (Forecasting the observables)**.** *Suppose we have observed T states $(x_t)_{t\leq T}$ of a (possibly noisy) dynamical system and we are interested in learning the evolution of its observables $f : \mathcal{X} \to \mathbb{R}$. A straightforward approach is to first learn an map $\hat{S}$ between the states such that $x_{t+1} \approx \hat{S}(x_t)$ and then predict any observable by simply composing it with the learned map $\hat{S}$, i.e. $f(x_{t+1}) \approx f(\hat{S}(x_t))$. Here, $\hat{S} : \mathcal{X} \to \mathcal{X}$ can be anything: a neural network, a vector-valued kernel etc. The framework developed in [KNM+22] for the Koopman operator learning, on the contrary, actively capitalize on the structure of the space of observables $\mathcal{H}$ to* directly *provide the best forecast as $f(x_{t+1}) \approx (\hat{U}f)(x_t)$ for* any *observable $f \in \mathcal{H}$, without the middle step of learning the dynamical map $\hat{S}$. The two approaches coincide if the underlying dynamics is deterministic, whereas for stochastic dynamical systems the first approach is sub-optimal (i.e. leading to larger errors).*

2 Quantum dynamics and the Lindblad equation

In the context of open quantum system dynamics, the state of a quantum mechanical system is described by a *density matrix* (or operator) $\rho : \mathbb{H} \to \mathbb{H}$, where $\mathbb{H}$ is a Hilbert space characterizing the quantum system [NC12, Chapter 2]. Physical requirements prescribe that the density matrix is a positive definite, hermitian operator satisfying $\mathrm{tr}(\rho) = 1$. Any physical observable A is an hermitian operator acting on $\mathbb{H}$, and its quantum mechanical expectation value for a given state ρ is $\langle A\rangle_\rho := \mathrm{tr}(A\rho)$. Noticing that $\mathrm{tr}(A\rho)$ is the scalar product between A and ρ in the Hilbert-Schmidt sense, we conclude that the observables of physical interest are *linear* functionals of $\rho \in \mathrm{HS}(\mathbb{H})$, where $\mathrm{HS}(\mathbb{H})$ is the space of Hilbert-Schmidt operators on $\mathbb{H}$.

For quantum systems with discrete energy levels, the Hilbert space $\mathbb{H}$ is isomorphic to $\mathbb{C}^d$ for some (possibly large) d. In the case of N two-level systems such as spins, one has $\mathbb{H} \sim \mathbb{C}^{2^N}$.

If the physical system is closed, i.e. isolated from any interaction with the environment, its evolution is unitary and governed by the Schrödinger equation. Conversely, when the system is open its equation of motion is also affected by the environment. Under proper assumptions and approximations[3], the dynamics of an open quantum system is described by the Lindblad equation [Lin76, WM09, Chapter 3]

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}\left[H, \rho(t)\right] + \sum_k \left[L_k\rho(t)L_k^\dagger - \frac{\rho(t)L_k^\dagger L_k + L_k^\dagger L_k\rho(t)}{2}\right],$$

where H is the Hamiltonian of the system and $L_k : \mathbb{H} \to \mathbb{H}$ are *collapse operators* describing disspiative processes. The Lindblad equation is a linear ordinary differential equation for the density matrix ρ. Its form guarantees that for all t, $\rho(t)$ is a density matrix (i.e. satisfy the aforementioned requirements of positivity, hermiticity and unit trace).

Remark 2 (Relation with Heisenberg and Schrödinger pictures)**.** *It is well known that the evolution of quantum systems can be described by two equivalent* pictures*. In the Schrödinger picture, which we employed above, the states evolve in time, while the operators are fixed. In the Heisenberg picture the opposite is true: states are fixed and operators evolve. Clearly, because of theirs equivalence, the choice of the picture do not affect the evolution of the expected value $\langle A\rangle_\rho(t) = \mathrm{tr}(A^{\mathrm{H}}(t)\rho) = \mathrm{tr}(A\rho^{\mathrm{S}}(t))$. The Koopman operator acts directly on the expected value $\langle A\rangle_\rho(t)$ and is therefore well defined in both pictures. Koopman operator learning, however, is based on a dataset of* evolved states *and is arguably more natural to adopt the Schrödinger picture to study its properties.*

[3]Namely: separability, Markovianity, Born and secular approximations.

3 Koopman operator learning on a spin chain

We now illustrate Koopman operator learning on a simple spin chain model by generating the training dataset from a simulated Lindblad evolution. This synthetic setting is mostly useful as a benchmark, as we can compare the *learned* dynamical system with the ground truth. A more sensible use of Koopman operator learning that we envision, however, is to make use of experimental measurements on complex quantum systems (e.g. quantum computers) to distill an accurate and interpretable model of theirs dynamical evolution.

We consider a spin chain with N spins described by the Hamiltonian

$$H := -\frac{1}{2}\sum_{i=1}^{N-1} J_{\parallel}\left(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y\right) + J_{\perp}\sigma_i^z\sigma_{i+1}^z, \tag{2}$$

where $\sigma_i^{x,y,z}$ are Pauli matrices corresponding to the i-th spin. The model (2) is also known as quantum Heisenberg model, and under proper assumptions on the coupling constants $J_{\parallel}, J_{\perp}$ can be exactly solved using the Bethe ansatz [KMGT97, KHM98].

We let the spin chain (2) interact with the environment via dephasing channels [MP08]. To this end, we choose the collapse operators of the Lindblad equation as N pure dephasings $L_i = \sigma_i^z\sqrt{\Gamma/2}$ with rate $\sqrt{\Gamma/2}$. As explained in [MP08], a pure dephasing $\sigma^z\sqrt{\Gamma/2}$ for a two level system is responsible for the decay of the off-diagonal elements of the density matrix at a rate $\propto e^{-\Gamma t}$.

Experimental details. The data are generated by simulating a chain of $N = 5$ spins with the QuTiP framework [JNN12]. We construct a dataset of quantum states by saving the density matrix of the simulated system $\rho(t)$ at times $t = k\Delta t$ with $k = 1, \ldots, 200$ and $\Delta t := 0.5$. The first half of the data, corresponding to the initial 100 timesteps, is used to train the Koopman operator estimator $\hat{U}$, while the second half is used for testing. The initial state $\rho(0)$ for the integration of the Lindblad equation is chosen as the pure state $|\downarrow\uparrow\uparrow\uparrow\uparrow\rangle$, corresponding to the first spin down and the rest up. The coupling constants and dephasing rate are set to $J_{\parallel} := 0.1\pi$, $J_{\perp} := 0.2\pi$, and $\Gamma = 0.2$, respectively.

The estimator $\hat{U}$ of the Koopman operator is computed using the *reduced rank regression* algorithm presented in [KNM+22] with rank $r = 19$, Tikhonov regularization $\lambda = 10^{-6}$ and a linear kernel. The *reduced rank regression* algorithm with linear kernel on the dataset $(x_i, y_i)_{i=1}^n$ returns the minimizer

$$\hat{U} := \operatorname*{arg\,min}_{U\in \mathrm{HS}_r(\mathcal{X})} \frac{1}{n}\sum_{i=1}^{n}\|y_i - U^* x_i\|^2 + \lambda\|U\|_{\mathrm{HS}}^2,$$

where $\mathrm{HS}_r(\mathcal{X})$ is the set of Hilbert-Schmidt operators on $\mathcal{X}$ with rank smaller or equal than r and U^* is the hermitian conjugate of U.

The code to reproduce the experiments is available at the link https://github.com/CSML-IIT-UCL/kooplearn. The experiments have been conducted on a workstation equipped with an Intel(R) Core^{TM} i9-9900X CPU @ 3.50GHz, 48GB of RAM and a NVIDIA GeForce RTX 2080 Ti GPU.

3.1 Forecasting the physical observables

We now compare the dynamics of selected physical observables as forecasted by the Koopman operator estimator $\hat{U}$ against the ground truth obtained by direct integration of the Lindblad equation. We focus on the spin polarization and on the spin current. The spin polarization (along z) of the i-th spin is the observable associated with the spin operator $S_i^z := \hbar\sigma_i^z/2$. A non-vanishing spin polarization $\mathrm{tr}(S_i^z\rho)$ across many sites signals the onset of a magnetic order for the state ρ. The Heisenberg model (2), indeed, is a rudimentary model of a magnetic material and depending on the coupling constants $J_{\parallel}$ and $J_{\perp}$ might exhibit either ferromagnetic or antiferromagnetic order [KMGT97, KHM98].

The spin current at site i can be derived from the continuity equation as detailed in [SKK04] and reduces to the operator

$$j_i^z := \frac{iJ_{\parallel}}{4}\left[(\sigma_i^+\sigma_{i+1}^- - \sigma_i^-\sigma_{i+1}^+) - (\sigma_{i-1}^+\sigma_i^- - \sigma_{i-1}^-\sigma_i^+)\right],$$

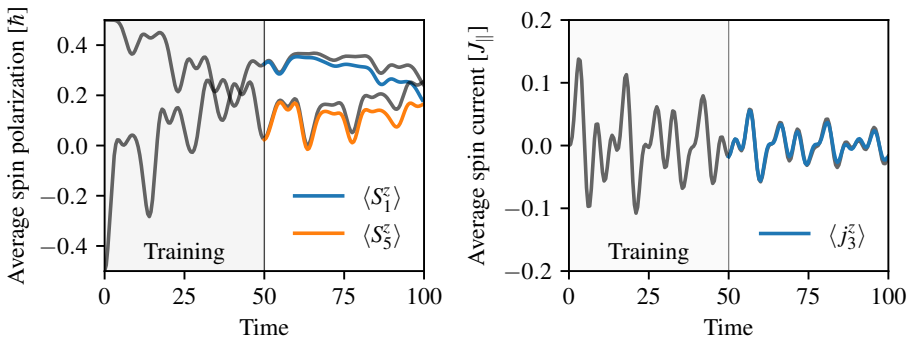

Figure 1: Left panel: forecast of the average spin polarization of the first (1) and last (5) spin of the chain in units of $\hbar$. Right panel: forecast of the average spin current of the central (3) spin of the chain in units of $J_{\parallel}$. In both panels the gray traces correspond to the ground truth and the first half of the data were used to train the Koopman operator estimator.

where $\sigma_i^{\pm} := \sigma_i^x \pm i\sigma_i^y$ are the usual ladder operators for the i-th spin. The first term in round parenthesis in the definition of j_i^z accounts for *outgoing* currents while the second term for *incoming* currents at site i. Manipulation and control of spin currents is of great interest in the field of *spintronics*, dealing with the use electrons' spin degrees of freedom for information processing. In spintronic devices, data storage and transmission can be achieved at higher efficiency compared to standard electronic devices [HYN$^+$20].

In Figure 1 we show the forecast of, respectively, the average spin polarization and the average spin current of selected sites along the chain. In both panels, the left portion of the plot corresponds to the data used to learn the Koopman operator estimator $\hat{U}$, while in the right portion of the plot are compared the forecasted dynamics on unseen data against the true dynamics. The forecast is accurate both for spin polarization and current at short times (i.e. for $50 \leq t \lessapprox 70$). At later times, however, the forecasted spin polarization diverge from the ground truth, while the forecasted spin current stays remarkably close to the true spin current.

3.2 Decay rates, normal modes, constants of motion and symmetries

We now highlight how the Koopman mode decomposition (1) might be used to infer quantities of physical interest such as decay rates, frequency of the normal modes and constants of motion of the dynamics.

Decay rates and frequency of the normal modes are easily obtained noticing that the time-dependence in (1) is expressed only by the eigenvalues λ_i of $\hat{U}$. The theory of dynamical systems dictates that $|\lambda_i| \leq 1$ for stable systems. The powered eigenvalues λ_i^t appearing in (1) can be conveniently expressed in polar form, yielding $\lambda_i^t := e^{-t/\tau_i} e^{i2\pi\omega_i t}$, with $\tau_i = -t/\log(|\lambda_i|)$ and $\omega_i = \arg(\lambda_i)/(2\pi t)$.

Computing the decay rates of the learned $\hat{U}$ we get that 16 out of 19 eigenvalues have decay rates in the interval $[0.018, 0.021]$, which is quite close to the true rate $\Gamma = 0.02$ selected for the pure dephasing gates. The remaining three decay rates are, respectively $0.014, 0.026$ and 0.003, the last one of which corresponds to the steady-state solution. Similarly, the computed frequencies ω_i lie in the range $[0.05, 0.2]$, of the same order of the (properly normalized) coupling constants $J_{\parallel}/2\pi = 0.05$ and $J^z/2\pi = 0.1$.

The last point we would like to comment concerns the mode corresponding to the steady-state of the system, i.e. the one with eigenvalue one. This mode, in practice, is characterized by an *estimated* Koopman eigenvalue very close to 1 and in our calculation is given by $\lambda_1 = 0.9985$. By construction, the right eigenfunction associated to the steady-state mode is a constant of motion. Indeed, $\psi_1(\rho_t) = Ut\psi_1(\rho_0) = \lambda_1^t \psi_1(\rho_0) = \psi_1(\rho_0)$ for all t.

In the present case we are dealing with *linear* observables[4], and every eigenfunction $\psi_i(\rho)$ can be expressed, by construction, as $\psi_i(\rho) = \mathrm{tr}(\Psi_i \rho)$ for some operator Ψ_i. For the steady-state mode, the operator Ψ_1 is a symmetry transformation for which the dynamics is invariant. We verified that the estimated Ψ_1 commutes with the total spin operator along z, as expected for the pure dephasing channel [MP08].

Broader impact

The abstract focus on two areas of active research: machine learning and quantum technologies. The techniques discussed in this abstract do not pose any direct ethical concern. Machine learning, here, is used only as a tool to aid and foster scientific research in the field of quantum technologies. The results presented might impact (hopefully in a positive way) the cohort of highly specialized researchers working on quantum technologies, and are very unlikely to impact the society at large.

Checklist

1. (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [No]
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [No]
3. (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

References

[AM17] Hassan Arbabi and Igor Mezić. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, 16(4):2096–2126, 2017.

[4] That is, linear functionals of the density matrix.

[Aro50] Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American mathematical society*, 68(3):337–404, 1950.

[BBKK22] Steven L. Brunton, Marko Budišić, Eurika Kaiser, and J. Nathan Kutz. Modern Koopman Theory for Dynamical Systems. *SIAM Review*, 64(2):229–340, May 2022.

[BMM12] Marko Budišić, Ryan Mohr, and Igor Mezić. Applied Koopmanism. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4):047510, December 2012.

[HYN+20] Atsufumi Hirohata, Keisuke Yamada, Yoshinobu Nakatani, Ioan-Lucian Prejbeanu, Bernard Diény, Philipp Pirro, and Burkard Hillebrands. Review on spintronics: Principles and device applications. *Journal of Magnetism and Magnetic Materials*, 509:166711, September 2020.

[JNN12] J.R. Johansson, P.D. Nation, and Franco Nori. QuTiP: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, August 2012.

[KHM98] Michael Karbach, Kun Hu, and Gerhard Muüller. Introduction to the bethe ansatz II. *Computers in Physics*, 12(6):565, 1998.

[KMGT97] Michael Karbach, Gerhard Muüller, Harvey Gould, and Jan Tobochnik. Introduction to the bethe ansatz i. *Computers in Physics*, 11(1):36, 1997.

[KNM+22] Vladimir Kostic, Pietro Novelli, Andreas Maurer, Carlo Ciliberto, Lorenzo Rosasco, and Massimiliano Pontil. Learning dynamical systems via koopman operator regression in reproducing kernel hilbert spaces, 2022.

[Koo31] B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, May 1931.

[Lin76] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, June 1976.

[MMS20] Alexandre Mauroy, Igor Mezić, and Yoshihiko Susuki, editors. *The Koopman Operator in Systems and Control*. Springer International Publishing, 2020.

[MP08] Florian Marquardt and Annett Püttmann. Introduction to dissipation and decoherence in quantum systems, 2008.

[MT93] Sean P. Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability*. Communications and Control Engineering. Springer London, 1993.

[NC12] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, June 2012.

[SC08] Ingo Steinwart and Andreas Christmann. *Support Vector Machines*. Springer New York, 2008.

[SKK04] F. Schütz, P. Kopietz, and M. Kollar. What are spin currents in heisenberg magnets? *The European Physical Journal B*, 41(4):557–560, October 2004.

[WM09] Howard M. Wiseman and Gerard J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, November 2009.