A Quantum Mechanical Approach for Data Assimilation in Climate Dynamics

Dimitrios Giannakis,[1,*] **Joanna Slawinska,**[2,†] **Abbas Ourmazd**[2,‡]

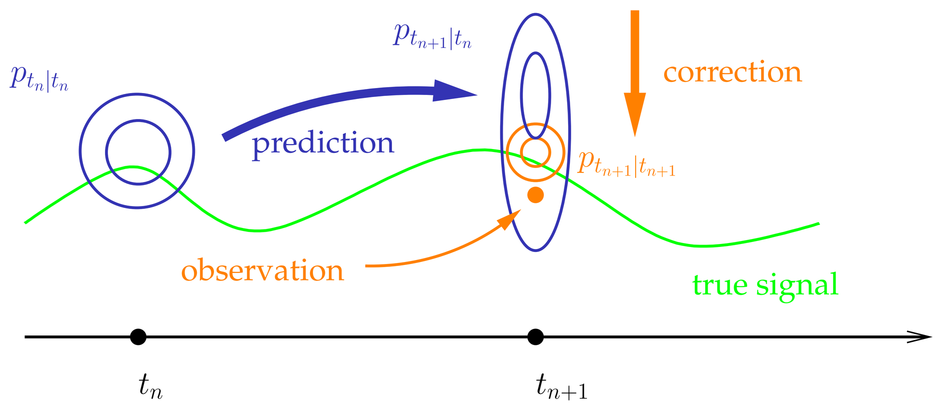
[1]Courant Institute of Mathematical Sciences, NYU, [2]Department of Physics, U. Wisconsin-Milwaukee

*dimitris@cims.nyu.edu, †slawinsk@uwm.edu, ‡ourmazd@uwm.edu

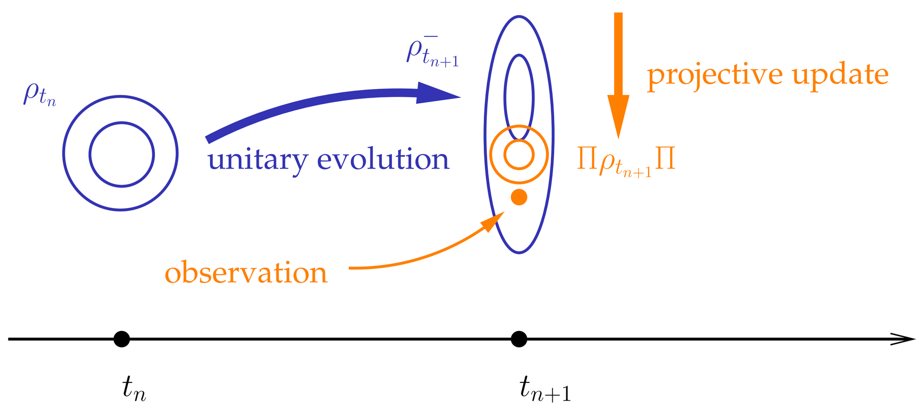
Motivation & Main Achievements

- **Sequential data assimilation** (a.k.a. **filtering**) is a predictor–corrector approach for state estimation and prediction of observables of dynamical systems. Among many applications, it is an integral part of **weather** and **climate forecasting systems**[1].
- Theoretical "gold standard" for filtering is computation of the **Bayesian posterior distribution**, given the full history of past observations of the system. However, this is oftentimes intractable, necessitating the use of *ad hoc* approximations, such as Gaussianity assumptions.
- We propose a new method[2,3] to address these issues, inspired from a conceptual similarity between data assimilation and **quantum mechanics**. Namely, both are inherently statistical theories, alternating between **evolutionary dynamics** between measurements, and **projective dynamics** during measurements.

Data assimilation



Quantum mechanics



- The **quantum mechanical data assimilation** (**QMDA**) framework is realized by mapping the assimilated dynamical system into a quantum system using **Koopman operator techniques**.
- A **data-driven formulation** is also constructed using **kernel methods** for machine learning, enabling data assimilation without prior knowledge of the equations of motion.

QMDA Framework

- We consider a dynamical system $\Phi^t : M \to M$ on a (unknown) state space M, preserving a probability measure μ (climatology). The system is observed at an interval Δt through a function $h : M \to \mathbb{R}$.
- The goal is to infer the probability distribution for future values of $v(t) = h(\Phi^t(x))$, given past measurements $v(t_n)$, $t_n = n\,\Delta t$.
- Associated with the dynamical system is the **Hilbert space of observables** (functions of the state) $L^2(\mu)$ and a group of unitary **Koopman evolution operators**[4]

$$U^t : L^2(\mu) \to L^2(\mu), \quad U^t f(x) = f(\Phi^t(x)).$$

- Following the quantum mechanical formalism, we represent the statistical state of the data assimilation system at time t by a **density operator** ρ_t on $L^2(\mu)$, such that

$$\rho_t \geq 0, \quad \operatorname{tr}\rho_t = 1.$$

  This generalizes the notion of a probability distribution in Bayesian statistics.
- We represent the assimilated observable h by a self-adjoint **multiplication operator** T on $L^2(\mu)$, such that $Tf = hf$. This operator can be decomposed in terms of an **operator-valued measure** E, which generalizes the notion of a spectral measure in time series analysis, viz.

$$T = \int_{\mathbb{R}} \omega\, dE(\omega).$$

- Between measurements, $t_n \leq t < t_{n+1}$, the state ρ_t evolves by **unitary dynamics** under the action of the Koopman operator,

$$\rho_t = U^{\tau *}\rho_{t_n} U^{\tau}, \quad \tau = t - t_n.$$

  The **probability distribution** for $v(t)$ to take values in a set $\Omega \subseteq \mathbb{R}$ is then given by

$$P_t(\Omega) = \operatorname{tr}(\rho_t E(\Omega)).$$

- If the measurement $v(t_{n+1})$ is found to lie in a set $\Xi \subseteq \mathbb{R}$, and the state immediately prior to t_{n+1} is $\rho^-_{t_{n+1}}$, the state $\rho_{t_{n+1}}$ immediately after the measurement is given by

$$\rho_{t_{n+1}} = \frac{E(\Xi)\rho^-_{t_{n+1}}E(\Xi)}{\operatorname{tr}(E(\Xi)\rho^-_{t_{n+1}}E(\Xi))}.$$

  This **projection step** is analogous to the Bayesian update formula in classical statistics.

Data-Driven Approximation

- The scheme is implemented by **finite-rank approximation** (i.e., matrix representation) of all operators involved in a basis of $L^2(\mu)$ learned from training data using **kernel algorithms**[5,6].
- Given time-ordered training data $F(x_n)$ taken through a map $F : M \to \mathbb{R}^d$ on a dynamical trajectory $x_n = \Phi^{t_n}(x_0)$, we compute eigenfunctions $\phi_j(x_n)$ of a self-adjoint kernel integral operator $K : L^2(\mu) \to L^2(\mu)$,

$$Kf(x) = \int_M k(F(x), F(x'))f(x')\,d\mu(x'),$$

  approximating integrals with respect to μ by **ergodic time averages**, i.e., $\int_M g(x)\,d\mu(x) \approx \sum_{n=0}^{N-1} g(x_n)/N$.
- Operators A on $L^2(\mu)$ are then represented by matrices,

$$A_{ij} = \langle \phi_i, A\phi_j\rangle_{L^2(\mu)} \approx \frac{1}{N}\sum_{n=0}^{N-1}\phi_i(x_n)A\phi_j(x_n).$$

  The Koopman operator, in particular, is approximated by the **shift operator** for time series, $U^{q\,\Delta t}\phi_j(x_n) = \phi_j(x_{n+q})$.
- Given the corresponding values $v(t_n) = h(x_n)$ of the assimilated observable, we also approximate the spectral measure E by a **discrete measure** constructed through a histogram of the values of $v(t_n)$.

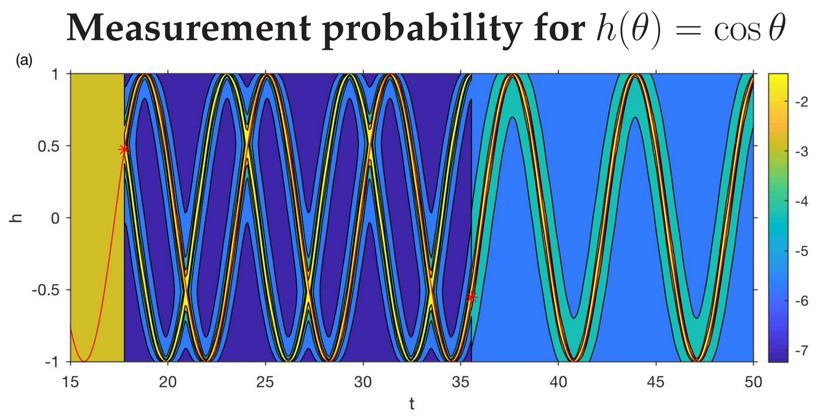
Comparison with Classical Methods

- By expressing data assimilation in terms of **intrinsically linear operators** for the dynamics (U^t), state (ρ_t), and measurement (T), QMDA avoids *ad hoc* approximations such as Gaussianity assumptions and diffusion regularization.
- The method outputs full probability distributions (P_t) in a nonparametric manner, as opposed to low-order moments (e.g., mean, covariance). The availability of P_t is useful for **risk assessment** and **uncertainty quantification**.
- Through basis projection, the cost of operator representation is decoupled from the ambient data space dimension and/or number of training samples.
- Unlike classical spectral approximation techniques, QMDA preserves sign and normalization of predicted probabilities.
- Rigorous **convergence results**[2] are obtained in a limit of infinite training data using techniques from linear operator theory in conjunction with spectral consistency results for kernel algorithms[7].

Periodic Dynamical System

- Dynamical flow is a rotation on the circle $M = S^1$, $\Phi^t(\theta) = \theta + \nu t \mod 2\pi$.
- Assimilated observable is a trigonometric function, $h(\theta) = \cos\theta$.

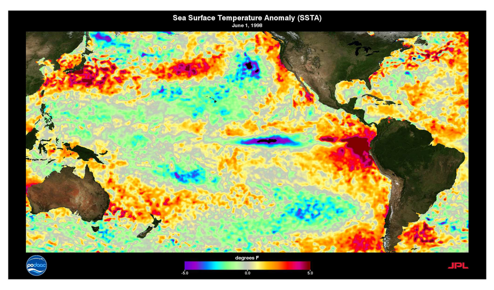
Measurement probability for $h(\theta) = \cos\theta$



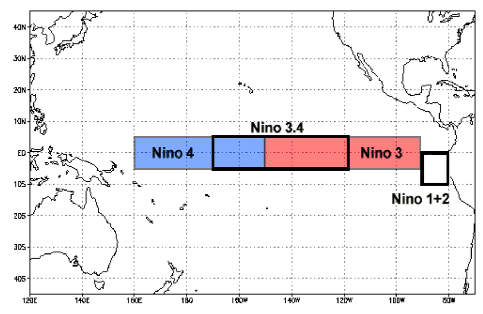
— true signal * observations

- QMDA starts from a stationary state ρ_0, corresponding to an uninformative (uniform) probability distribution P_0.
- When the first measurement is made, P_0 collapses to a **bimodal distribution**, consistent with the fact that $\cos\theta$ is a 2-to-1 function on the circle.
- When the second measurement is made, P_t collapses to a strongly peaked **unimodal distribution** that accurately tracks the true signal. This is consistent with the fact that two successive measurements of $\cos\theta$ are enough to uniquely infer θ.

El Niño Southern Oscillation



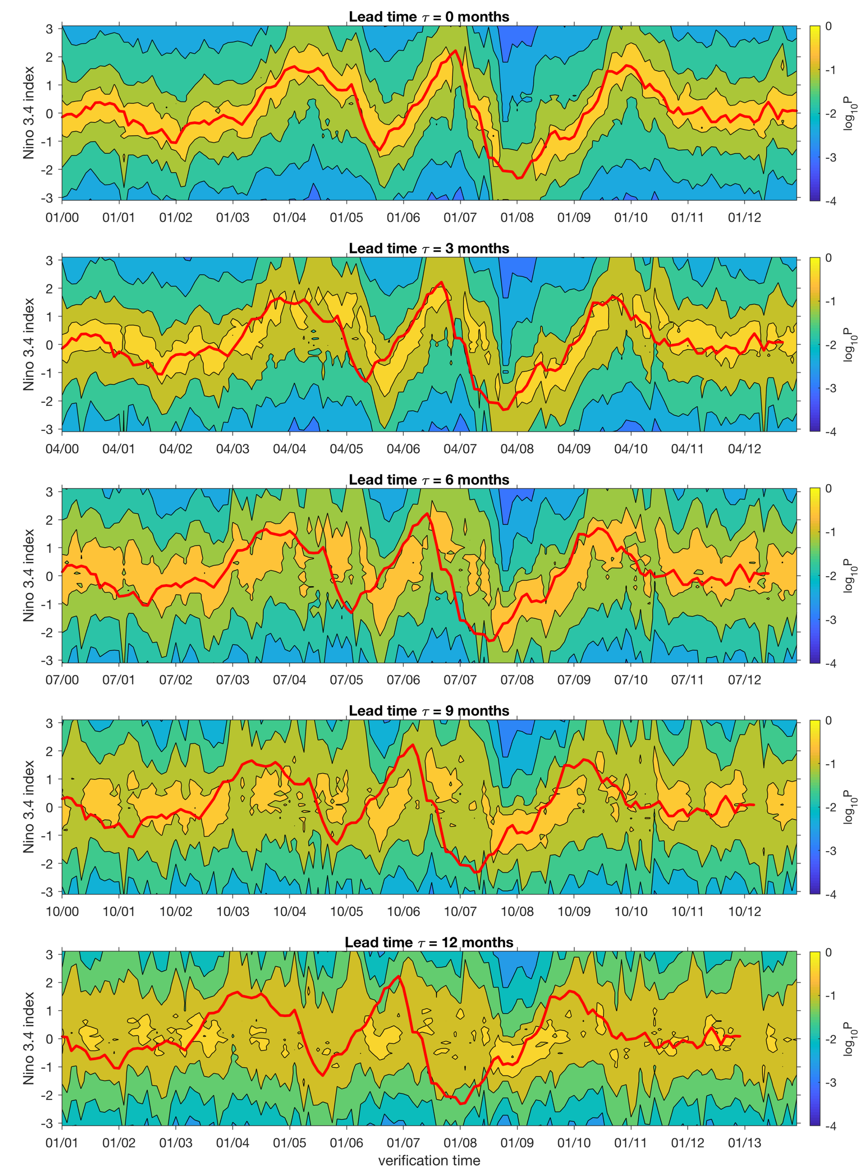
NASA/JPL/PODAAC, NOAA



NOAA

- We apply QMDA to data assimilation of ENSO in the **Community Climate System Model Version 4 (CCSM4)**[8].
- Training data is 1200 years of monthly-averaged **Indo-Pacific SST** fields at 1° resolution ($d \simeq 10^4$ gridpoints).
- Verification data is the **Niño 3.4 index** over the last 100 years of the control integration.
- Assimilated observables (h) are the Niño 1+2, Niño 3, Niño 3.4, and Niño 4 indices, observed monthly.

Probability density P_t for Niño 3.4 index



— true signal

- Starting from an uninformative (climatogological) distribution, the Niño 3.4 distribution P_t output by QMDA is seen to track the true signal.
- In addition to point forecasts (e.g., through the mean), P_t provides meaningful **uncertainty quantification**.
- **El Niño/ La Niña initiation** is oftentimes captured several months in advance. This suggests skillful **seasonal probabilistic ENSO prediction**.

Future Directions

- Extensions to high-dimensional observation functions using **multitask learning** techniques[9].
- Forecasting of ENSO impacts on the climate (e.g., **precipitation**, **sea ice**) and **socio-environmental systems**.
- Applications to closure and **stochastic subgrid-scale modeling** of unresolved dynamics.

References

1. Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge University Press.
2. Slawinska, J., A. Ourmazd, D. Giannakis (2019). A quantum mechanical approach for data assimilation in climate dynamics. *ICML 2019 Workshop on Climate Change: How Can AI Help?*
3. Giannakis, D. (2019). Quantum mechanics and data assimilation. *Phys. Rev. E*, 100:032207.
4. Budisić, M., R. Mohr, I. Mezić (2012). Applied Koopmanism. *Chaos*, 22:047510.
5. Berry, T., D. Giannakis, J. Harlim (2015). Nonparametric forecasting of low-dimensional dynamical systems. *Phys. Rev. E*, 91:032915.
6. Giannakis, D., J. Slawinska, Z. Zhao (2015). Spatiotemporal feature extraction with data-driven Koopman operators. *J. Mach. Learn. Res. Proceedings*, 44:103–115.
7. von Luxburg, U., M. Belkin, O. Bousquet (2008). Consistency of spectral clustering. *Ann. Stat.*, 36(2):555–586.
8. Gent, P. R. et al. (2011). The Community Climate System Model Version 4. *J. Climate*, 24:4973–4991.
9. Giannakis, D., A. Ourmazd, J. Slawinska, and Z. Zhao (2019). Spatiotemporal pattern extraction by spectral analysis of vector-valued observables. *J. Nonlinear Sci.*, 29(5), 2385–2445.

Acknowledgments: This research was supported by ONR YIP grant N00014-16-1-2649 and NSF grant 1842538.