Online PCA in Converging Self-consistent Field Equations

Xihan Li[1], Xiang Chen[2], Rasul Tutunov[3], Haitham Bou Ammar[3], Lei Wang[4], Jun Wang[1]
[1]University College London [2]Huawei Noah's Ark Lab [3]Huawei R&D U.K.
[4]Institute of Physics, Chinese Academy of Sciences

NEURAL INFORMATION PROCESSING SYSTEMS

Scan for the full paper

Problem Setting

This is called eigen decomposition:

An $n \times n$ matrix — A scalar (eigenvalue)

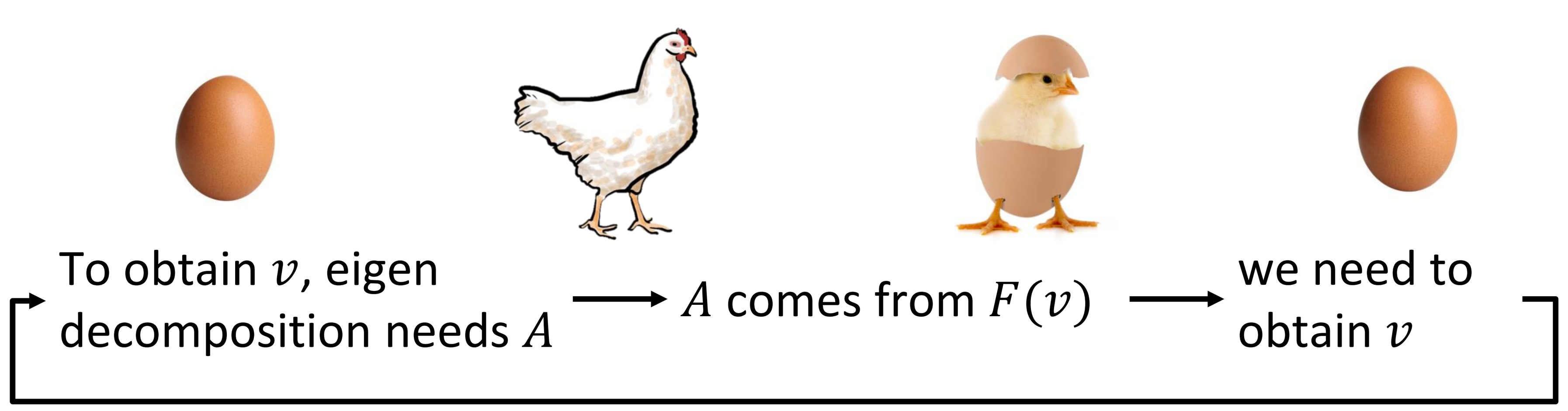
$$Av = \lambda v$$

What will happen if matrix A is not directly given, but a given function of v?

F is given — An $n \times n$ matrix — A scalar (eigenvalue)

$$A \doteq F(v) \quad F(v)v = \lambda v$$

Eigen decomposition cannot be directly applied anymore!



Equations in such form are called Self-consistent Field Equations

They are important in Quantum Physics!

$$H|\Psi\rangle = E|\Psi\rangle$$

Traditional Methods

Assign an initial v_0 $\xrightarrow{F_0 = F(v_0)}$ F_0 $\xrightarrow[\text{Eigen decomposition}]{F_0 v_1 = \lambda v_1}$ v_1 $\longrightarrow F_1 \longrightarrow v_2 \longrightarrow$ ... (until convergence)

Easily fails to converge (**infinite oscillation** between two or more states)

Two current main research directions:

1. Generate a better initial solution v_0
2. Mix F_t with those in previous iterations F_{t-1}, F_{t-2}, ... to stabilize the iteration

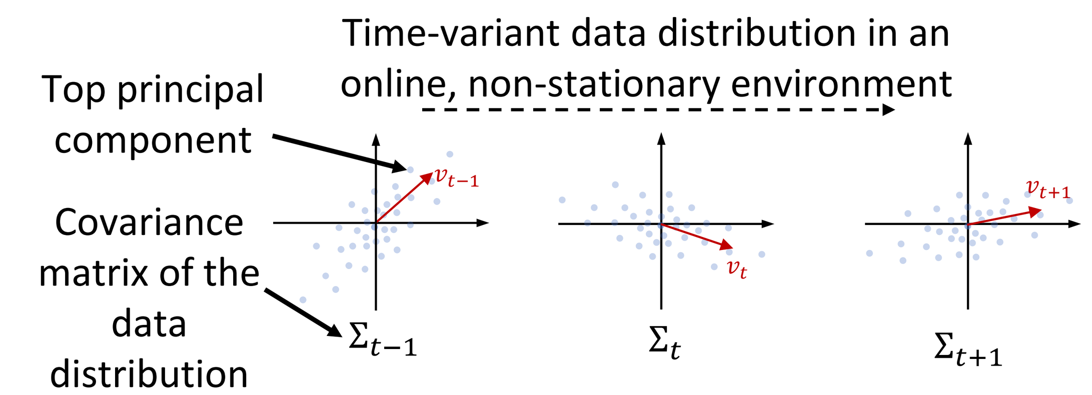
Motivation

We find a connection between two very different problems in different fields

An $n \times n$ matrix — A scalar (eigenvalue)

$$F(v)v = \lambda v$$

SCF Equation



Online PCA

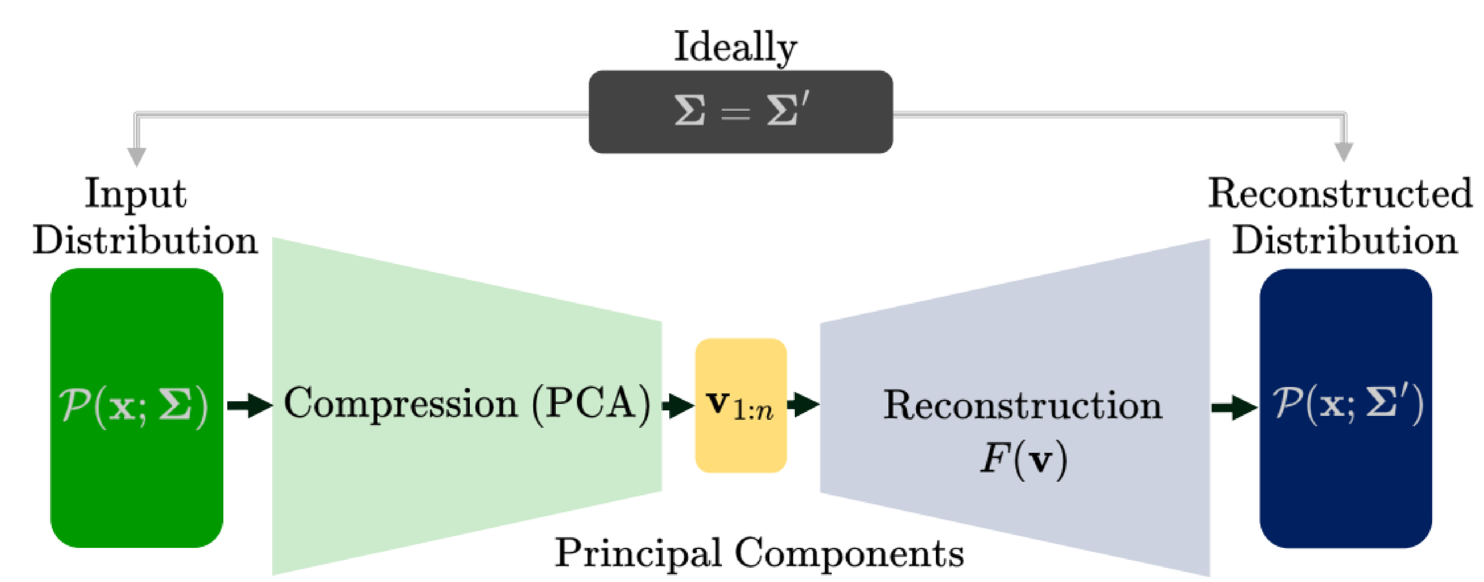
(1) They involve eigen decomposition

(2) Their matrices to be decomposed are not determined

However, Online PCA is **incrementally updated**

Can this help avoid infinite oscillation between states?

Our Method



$F(v)v = \lambda v$ is to say, if we have a matrix Σ, then

Step 1: Decompose Σ to get its top eigenvector v

New interpretation: "compress" Σ with PCA to have v

Step 2: Compute a new matrix $\Sigma' = F(v)$

New interpretation: "reconstruct" Σ from v with $F(\cdot)$

Then we will have $\Sigma' = \Sigma$

Equivalent Target: finding a distribution that is invariant before and after being processed by an auto-encoder structure involving PCA.

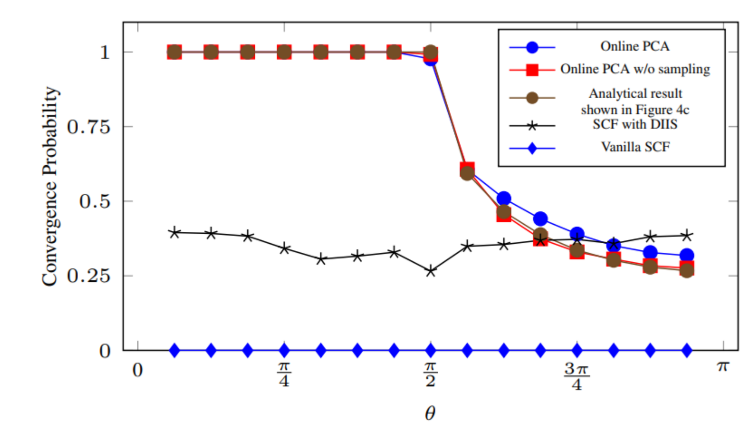
Then, the fixed-point iteration $v_0 \to F_0 \to v_1 \to F_1 \to \cdots$ can be regarded as

Compress (PCA) → reconstruct → compress (PCA) → reconstruct → ...

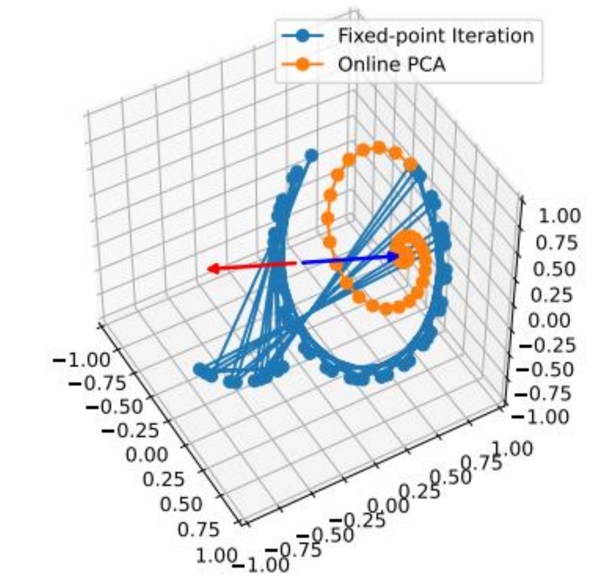
We are continuously running PCA in a non-stationary environment!

Now we can apply Online PCA to **update v incrementally** to avoid **infinite oscillation**

Experimental Results



(a) Convergence ratio

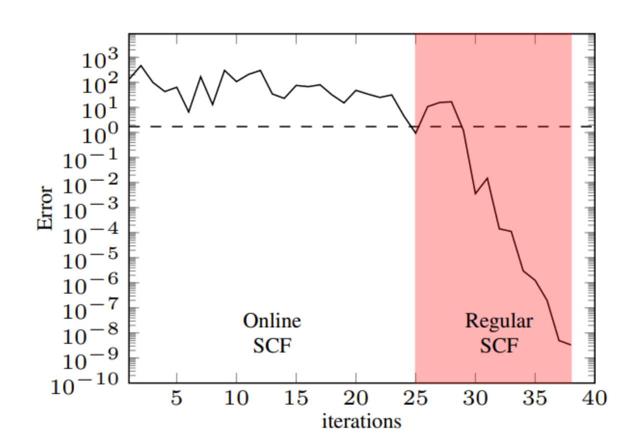


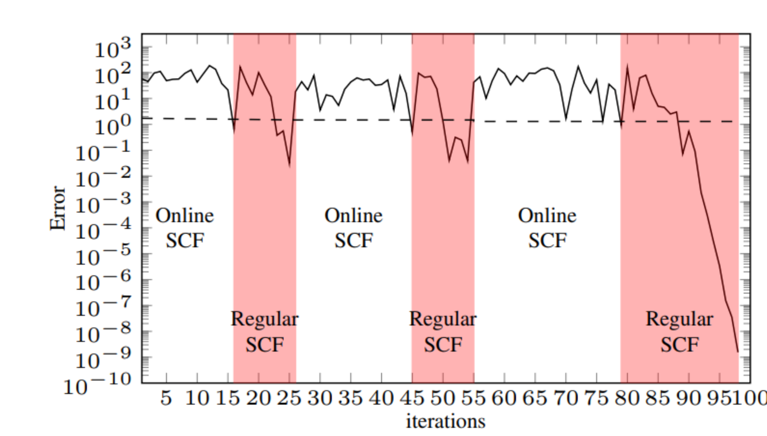
Case study: solve $(Avv^\top A^\top)v = \lambda v$

- Vanilla fixed-point method: does not work at all (0% convergence)
- DIIS: ~40% convergence ratio
- Online PCA: the top curves, half has 100% convergence ratio

Methods	Hartree-Fock			DFT with B3LYP		
	#(Nonconverged molecules)		Average #(iterations)	#(Nonconverged molecules)		Average #(iterations)
Regular SCF	124	(9.27%)	25.49	407	(30.42%)	21.09
Full Online SCF	13	(0.97%)	584.68	217	(16.22%)	1835.24
Adaptive Online SCF	0	(0%)	42.97	0	(0%)	60.58

Results on 1,338 randomly sampled molecules in QM9 dataset. All methods are initialized with core Hamiltonian and accelerated by DIIS





For real-world SCF equations such as Hartree-Fock and DFT, our proposed method with adaptations (Online SCF) can also achieve high convergence ratio with a moderate increase of iterations. We also proposed an adaptive switching mechanism between online and regular mode, to balance efficiency and convergency.