

# EIGENSTATE EXTRACTION WITH NEURAL-NETWORK TOMOGRAPHY

Abhijeet Melkani, Clemens Gneiting, and Franco Nori

Presenter: Parsa Rangriz  
Sharif University of Technology

Phys. Rev. A 102, 022412 (2020)

# AUTHORS



Abhijeet Melkani  
University of Oregon



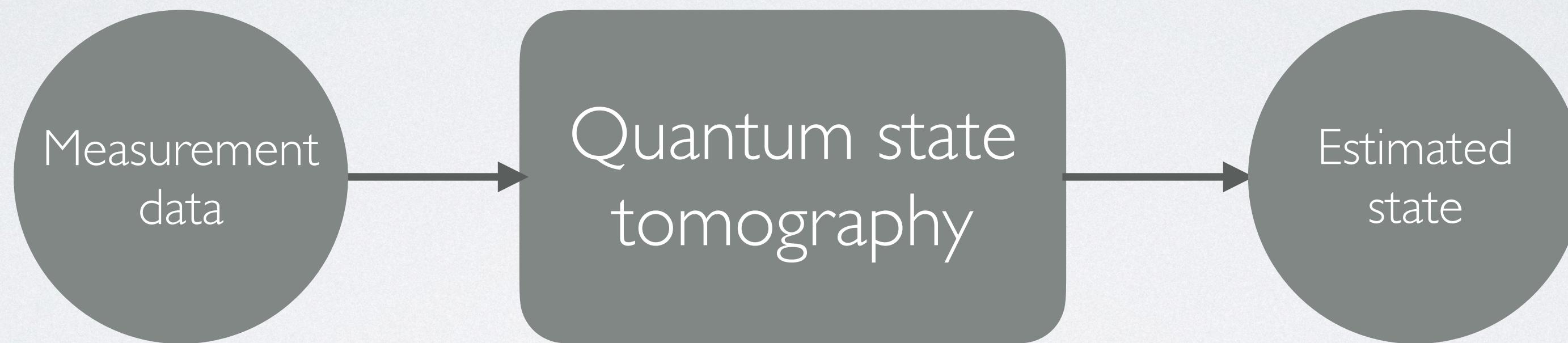
Clemens Gneiting  
RIKEN Cluster for Pioneering Research



Franco Nori  
RIKEN Cluster for Pioneering Research  
University of Michigan

# INTRODUCTION

- *Quantum state tomography* is the process of reconstructing the states produced in quantum experiments or devices from their measurement data.



- Our scheme is based on the insight that highly efficient methods for *pure-state tomography* can also be used to robustly recover the (dominant) eigenstates of mixed states.

# INTRODUCTION

- Our scheme promotes any pure-state tomography method, i.e., any measurement data-based state estimation that is constrained within the set of pure states, to mixed-state tomography.
- To demonstrate our reconstruction scheme, we here use and adjust a recently developed method for pure-state tomography based on neural-network representations of quantum states.
- Indeed, Neural Quantum States (NQS) have been shown to be viable for tomography of complex, high-dimensional pure states, leveraging both the efficient and scalable representation of neural networks and their great expressional power.

# OPTIMAL LOW-RANK APPROXIMATIONS OF DENSITY MATRICES

Let us assume  $\rho$  to be the density matrix to be reconstructed,  $\rho = \sum_{i=1}^n p_i |\psi\rangle\langle\psi|$  where,  $p_1 \geq p_2 \geq \dots \geq p_n$ .

**Proposition 1:** In terms of fidelity, the unique closest pure state to  $\rho$  is its dominant eigenstate  $|\psi_1\rangle$ , with fidelity  $F(\rho, |\psi\rangle\langle\psi|) = p_1$ .

**Proposition 2:** In terms of trace distance, the unique closest pure state to  $\rho$  is its dominant eigenstate  $|\psi_1\rangle$ , with trace distance  $T(\rho, |\psi_1\rangle\langle\psi_1|) = 1 - p_1$ .

**Proposition 3:** In terms of fidelity, the unique closest rank- $r$  approximation to  $\rho$  is

$$\sigma = \frac{1}{\kappa(r)} \sum_{i=1}^r p_i |\psi_i\rangle\langle\psi_i|.$$

With fidelity  $F(\rho, \sigma) = \kappa(r) = p_1 + p_2 + \dots + p_r$

**Proposition 4:** There are infinitely many rank- $r$  approximations to  $\rho$  which achieve the same trace distance as  $\sigma$ ,  $T(\rho, \sigma) = 1 - \kappa(r)$ .

# ITERATIVE EIGENSTATE RECONSTRUCTION

In the following, let  $P_m$  denote a family of projectors, corresponding to measurements in the experiment.

**Step 1:** Based on the measurement statistics  $\text{Tr}(P_m \rho)$ , employ a chosen method for pure-state tomography to determine the pure state  $|\hat{\psi}_1\rangle$  which is closest to  $\rho$ .

**Step 2:** Numerically calculate the measurement statistics for the eigenstate approximation  $|\hat{\psi}_1\rangle$ ,  $\text{Tr}(P_m |\hat{\psi}_1\rangle\langle\hat{\psi}_1|) = \langle\hat{\psi}_1 | P_m |\hat{\psi}_1\rangle$ .

**Step 3:** Determine the dominant eigenvalue  $\hat{p}_1$  corresponding to  $|\hat{\psi}_1\rangle$ .

Since  $1 - p_1 = T(\rho, |\psi_1\rangle\langle\psi_1|)$ , the trace distance can be estimated from the measurement statistics according to

$$T(\rho, |\psi_1\rangle\langle\psi_1|) = \max_P |\text{Tr}(P\rho) - \langle\psi_1 | P |\psi_1\rangle|$$

We then estimate,

$$1 - \hat{p}_1 \approx \max_{P_m} |\text{Tr}(P_m \rho) - \langle\hat{\psi}_1 | P_m |\hat{\psi}_1\rangle|$$

# ITERATIVE EIGENSTATE RECONSTRUCTION

**Step 4:** Then calculate the measurement statistics for the hypothetical state  $\rho'$  according to

$$Tr(P_m\rho') = \frac{1}{1 - \hat{p}_1} (Tr(P_m\rho) - \hat{p}_1 \langle \hat{\psi}_1 | P_m | \hat{\psi}_1 \rangle)$$

where,  $\rho' = \frac{1}{1 - \hat{p}_1} (\rho - \hat{p}_1 |\hat{\psi}_1\rangle\langle\hat{\psi}_1|) \approx \frac{1}{1 - \hat{p}_1} \sum_{i=2}^n p_i |\psi_i\rangle\langle\psi_i|$

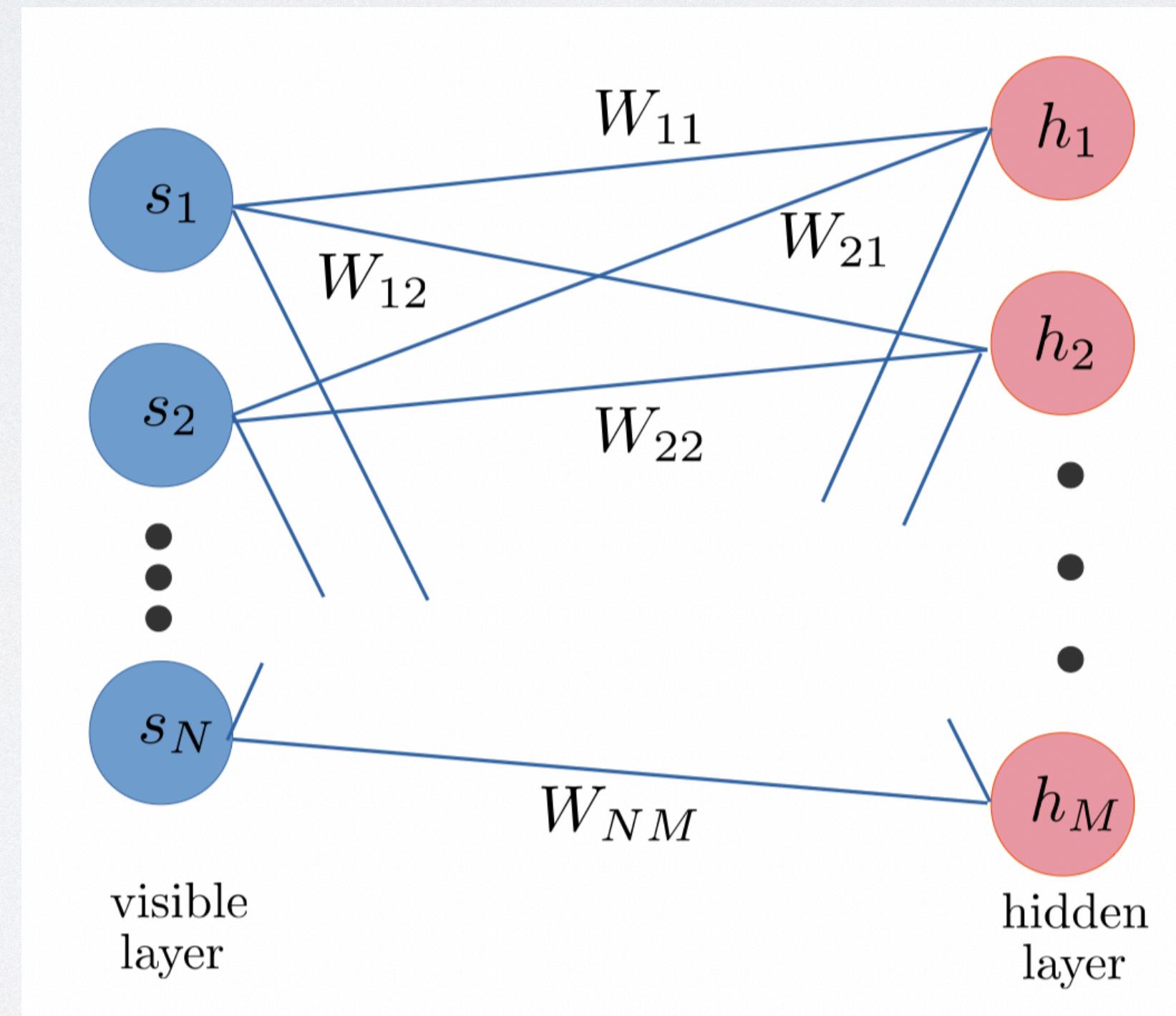
**Step 5:** Return to Step 1 with the new measurement statistics  $Tr(P_m\rho')$  and do the other steps, again.

**Termination:** Terminate after reaching the desired rank  $k$ . The constructed density matrix then becomes,

$$\hat{\rho} = \frac{1}{\sum_{i=1}^k \hat{p}_i} \sum_{i=1}^k \hat{p}_i |\hat{\psi}_i\rangle\langle\hat{\psi}_i|$$

# RESTRICTED BOLTZMANN MACHINE (RBM)

An RBM consists of two layers: the visible layer with  $N$  nodes  $s_i$  (visible neurons), and the hidden layer with  $M$  auxiliary nodes  $h_i$  (hidden neurons). The hidden neurons are coupled to the visible ones, but there is no coupling among neurons in the same layer.



Schematic depiction of a restricted Boltzmann machine

## RESTRICTED BOLTZMANN MACHINE (RBM)

Consequently, an RBM can be expressed in the succinct form:

$$p_\kappa(\vec{s}, \vec{h}) = \exp \left( \sum_{ij} W_{ij}^\kappa s_i h_j + \sum_i a_i^\kappa s_i + \sum_j b_j^\kappa h_j \right)$$

The edge weight  $W_{ij}$  and the bias weight  $a_i$  and  $b_i$  form the parameters to be optimized, subject to some data-based training.

# TOMOGRAPHY WITH NEURAL-NETWORK QUANTUM STATES

Consider a quantum systems composed of  $n$  qubits, with its Hilbert space spanned by some reference basis  $\vec{s} = (s_1, s_2, \dots, s_n)$ , with  $s_i = \pm 1$ . A pure quantum state is then completely characterized by the  $2^n$  coefficients  $\langle \vec{s} | \psi \rangle = \psi(\vec{s})$ .

In the definition of the neural-network quantum state ansatz which we implement here, these coefficients are approximated by two real-valued neural networks,  $p_\lambda$  and  $p_\mu$ , based on the RBM architecture, such that

$$\psi_{\lambda,\mu}(\vec{s}) = \sqrt{\frac{p_\lambda}{Z_\lambda}} \exp[i\phi_\mu(\vec{s})/2]$$

where  $\phi_\mu(\vec{s}) = \log p_\mu(\vec{s})$ , and  $Z_\lambda$  denotes a normalization constant.

# TOMOGRAPHY WITH NEURAL-NETWORK QUANTUM STATES

The starting point for the reconstruction is a series of independent projection measurements on a pure state,  $P_b(\vec{s}^{[b]})$ . Here the basis rotations  $b$  are applied to  $\vec{s}$  to obtain a collection of projection bases  $\vec{s}^{[b]}$ .

The RBMs are then trained on this data set such that the network parameters,  $\lambda$  and  $\mu$ , maximize the data-set likelihood, i.e.,  $|\psi_{\lambda,\mu}(\vec{s}^{[b]})|^2 \approx P_b(\vec{s}^{[b]})$

The Kullback-Leibler divergence, which quantifies the statistical distance between two probability distributions, can be used as cost function,

$$C = \sum_b \sum_{\vec{s}^{[b]}} P_b(\vec{s}^{[b]}) \log \frac{P_b(\vec{s}^{[b]})}{|\psi_{\lambda,\mu}(\vec{s}^{[b]})|^2}$$

Typically, the cost function is minimized iteratively by gradient descent.

Thanks for your attention!