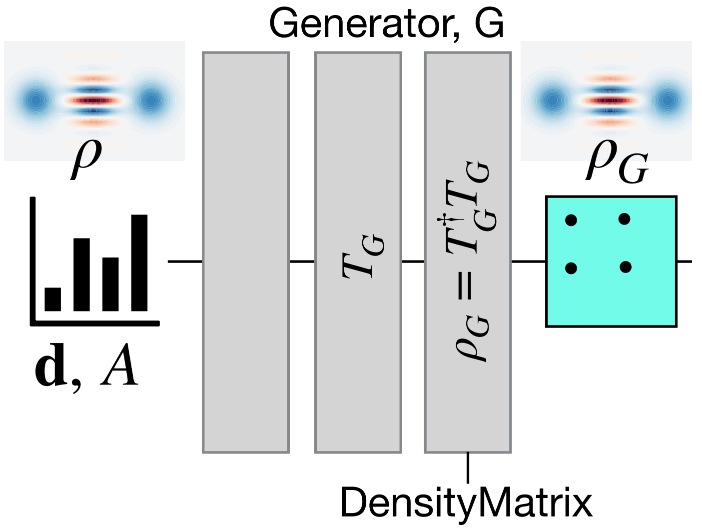
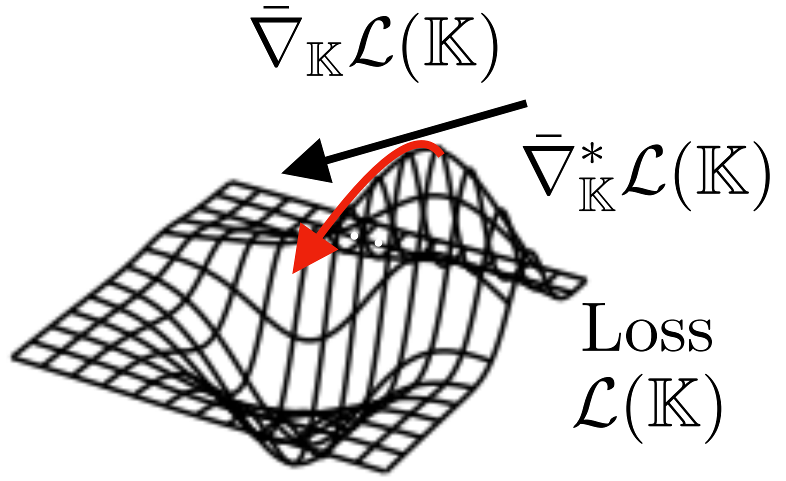
WACQT | Wallenberg Centre for Quantum Technology

Quantum state and process tomography with machine learning and gradient descent



2nd Workshop of Machine Learning for Quantum Technology,
Erlangen, Germany, November 5-8, 2024



Anton Frisk Kockum

Associate Professor, WACQT/Chalmers

CHALMERS

Knut och Alice Wallenbergs Stiftelse

Outline

- Part 1: Quantum state tomography with conditional generative adversarial networks
- Part 2: Quantum process tomography with gradient-descent learning of Kraus operators
- Bonus slides: Photoelectrons and GD-QST
- Summary

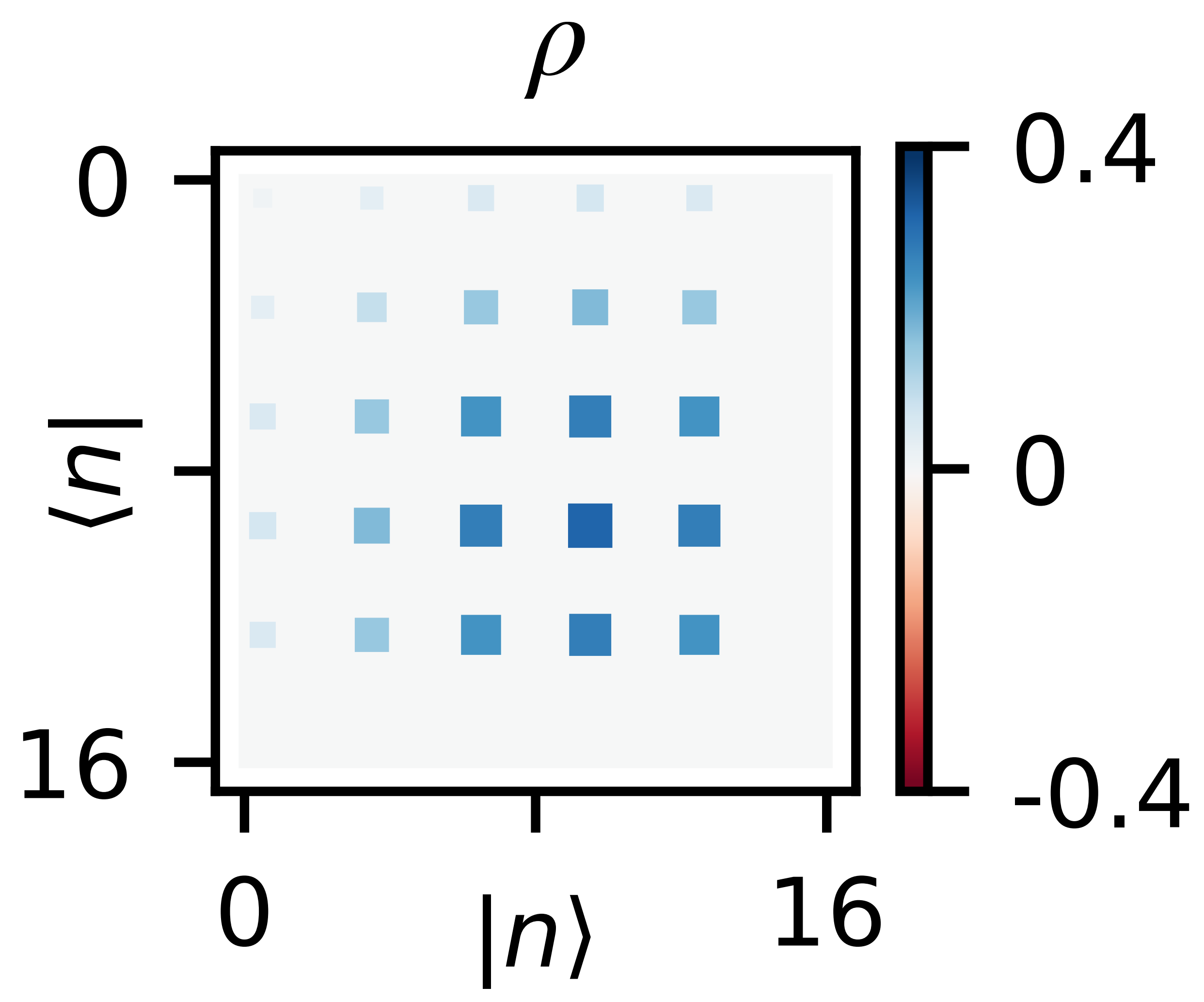
S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Lett. **127**, 140502 (2021)
S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Res. **3**, 033278 (2021)
S. Ahmed, F. Quijandría, and A. F. Kockum, Phys. Rev. Lett. **130**, 150402 (2023)
Code available on GitHub: quantshah/qst-cgan and quantshah/gd-qpt

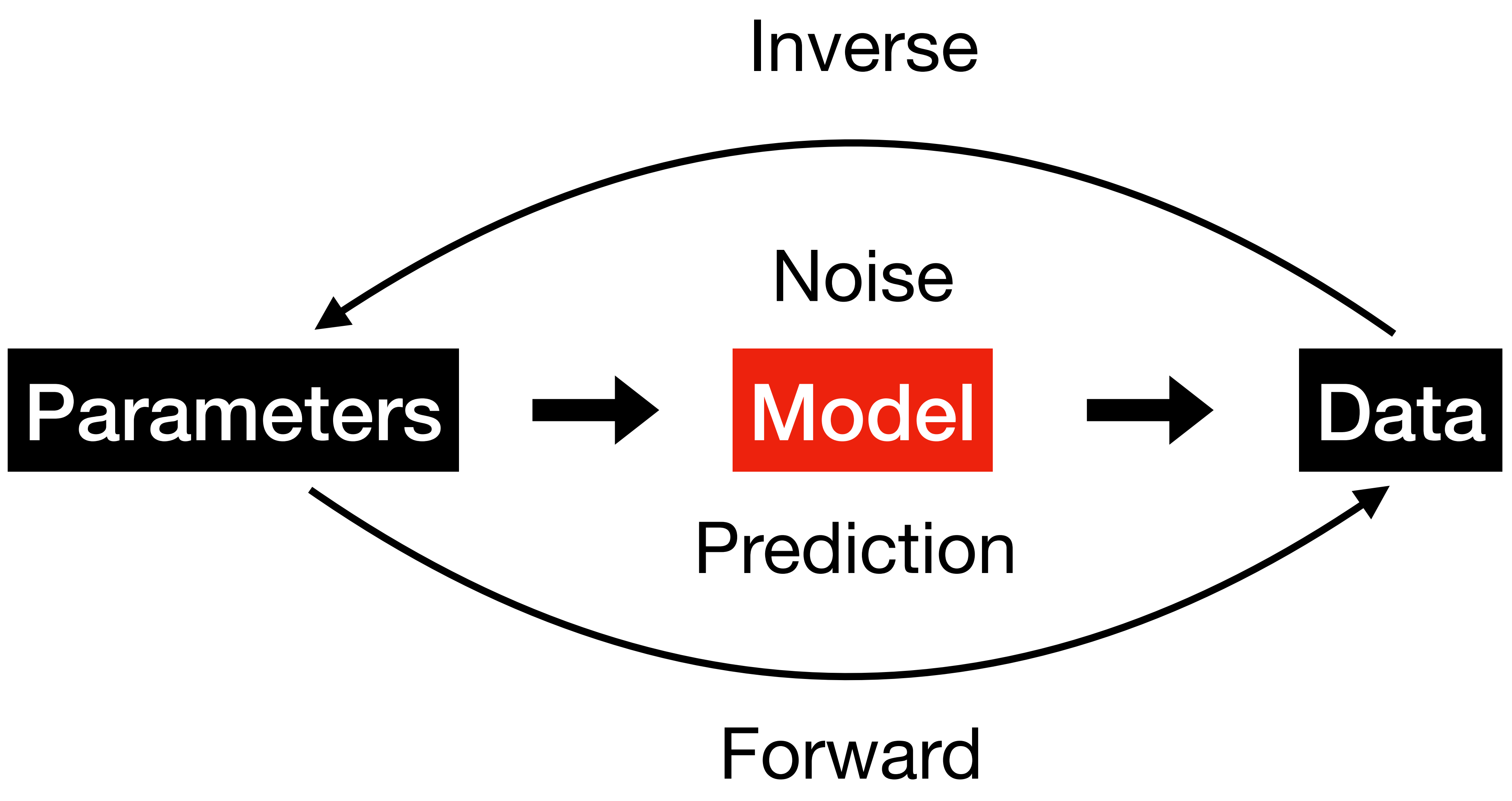
Outline

- Part 1: Quantum state tomography with conditional generative adversarial networks
- Part 2: Quantum process tomography with gradient-descent learning of Kraus operators
- Bonus slides: Photoelectrons and GD-QST
- Summary

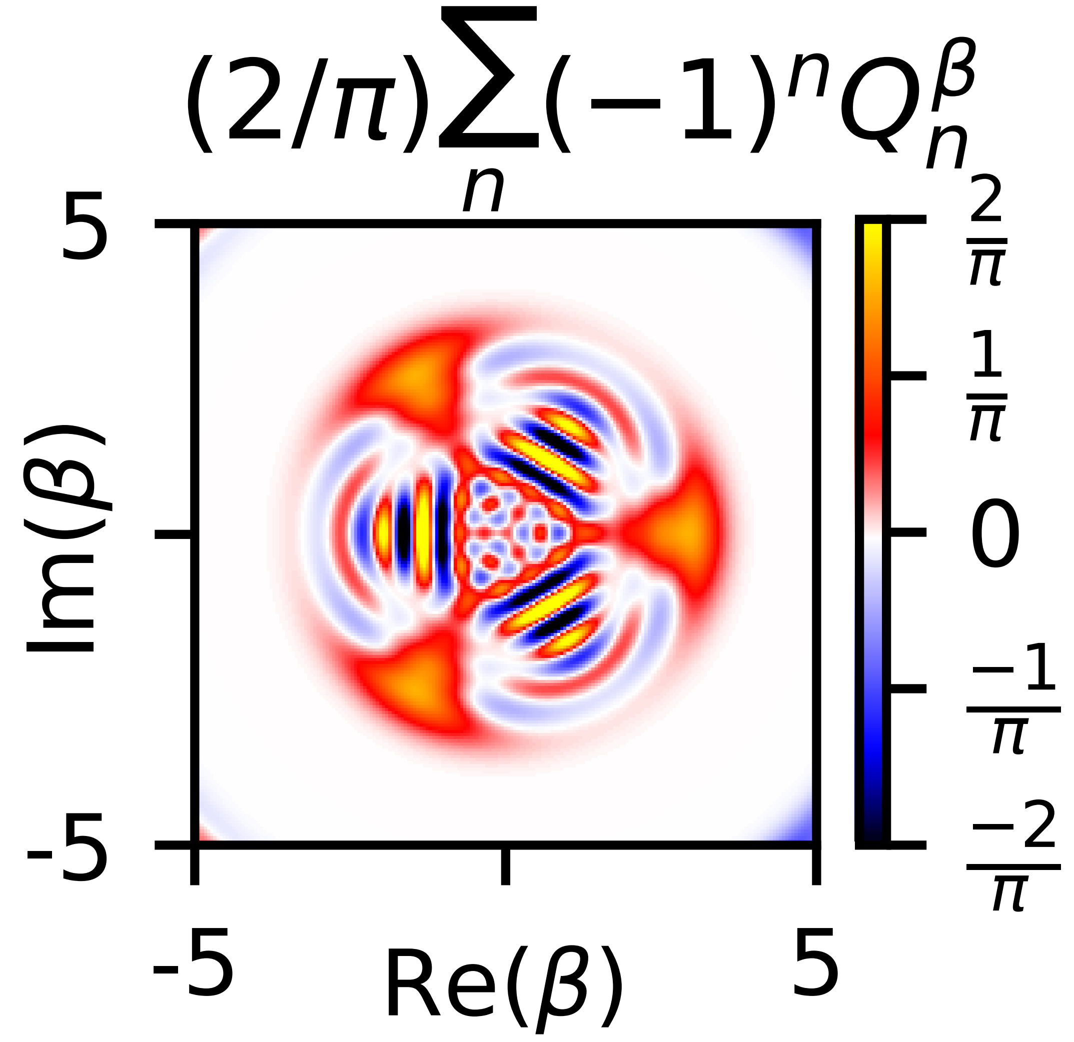
Quantum state tomography

Density matrix





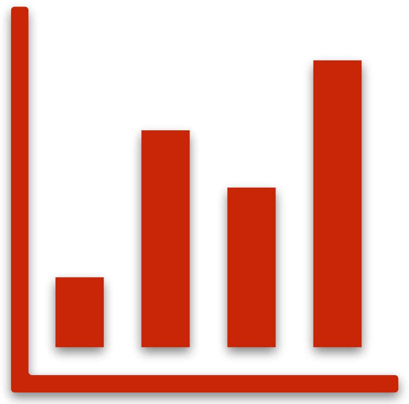
Wigner function



Quantum state tomography is a tricky inverse problem, important for many experiments

- Many parameters $(N^2 - 1)$ in the density matrix to determine
- A lot of measurement data required
- Calculations can be time-consuming

Generative models



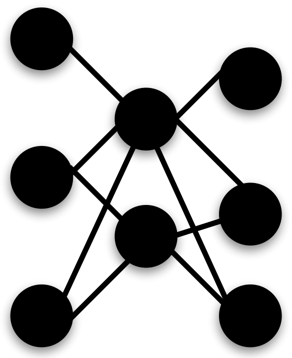
d

Data samples

unknown
data distribution

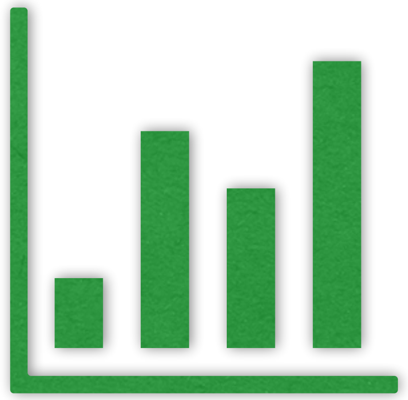
Probability density

p_{data}



Generative model

sample new data



d'

d'~ p_{data}

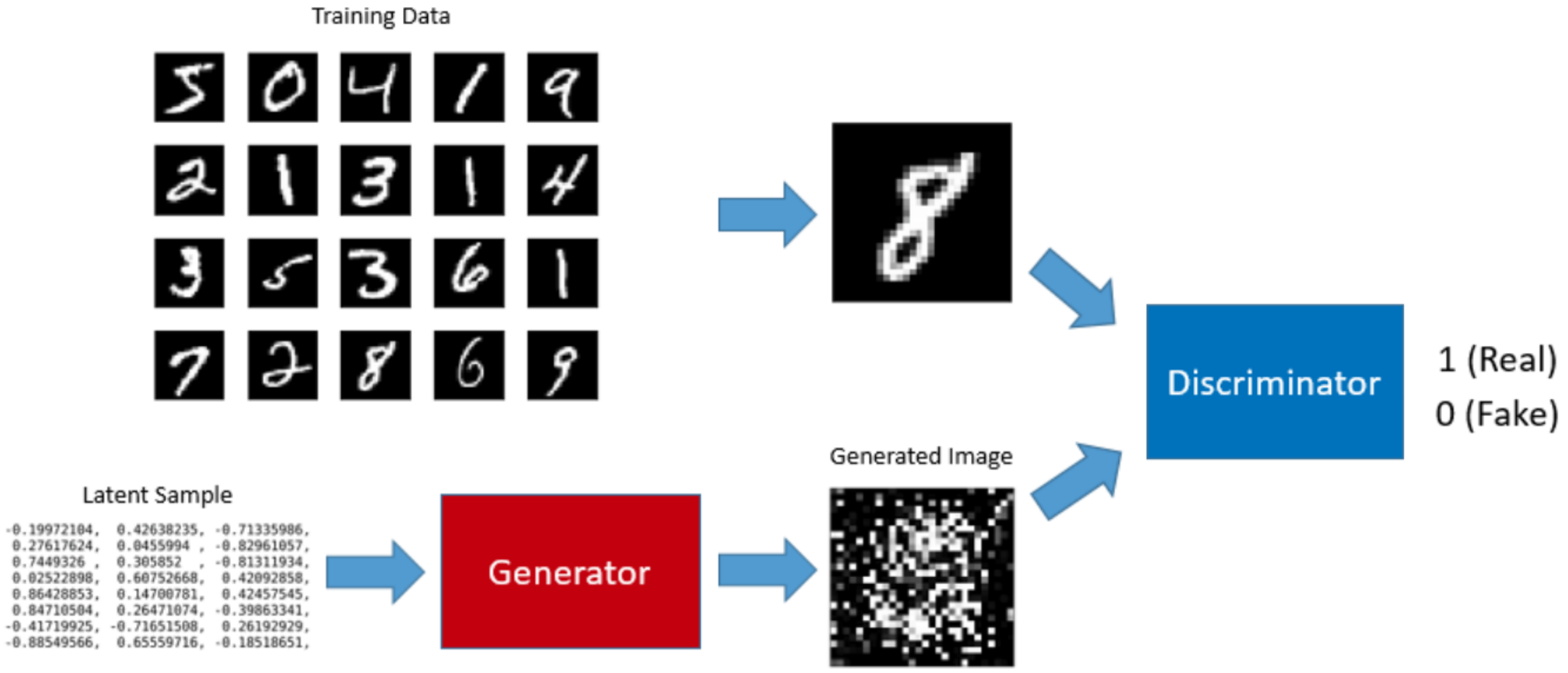
Flickr-Faces-HQ Dataset (FFHQ)
https://github.com/NVlabs/ffhq-dataset

ThisPersonDoesNotExist.com

Learn p_{data}, approximate it, or sample from it

Generative adversarial networks

Duelling networks train to generate plausible fake data and to distinguish it from real data



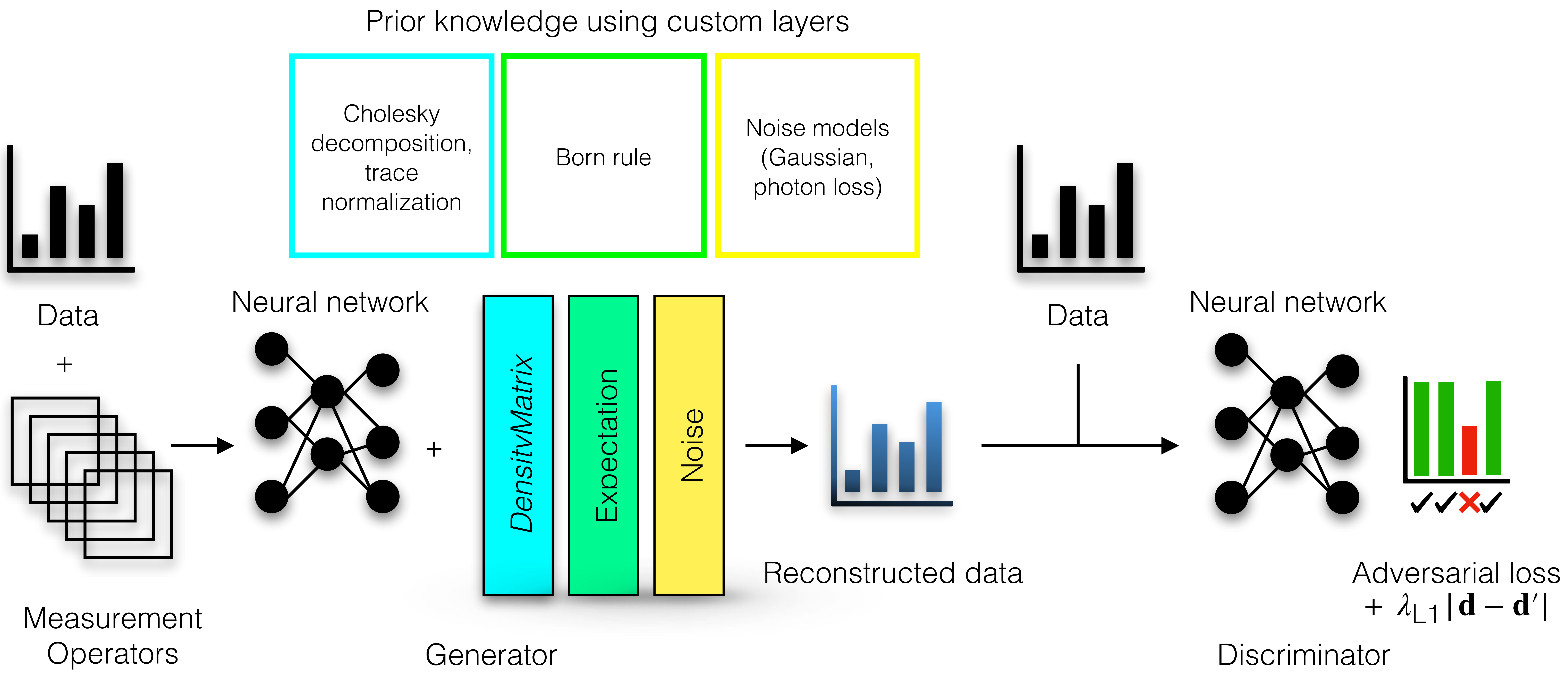
https://naokishibuya.medium.com/understanding-generative-adversarial-networks-4dafc963f2ef

Conditional generative adversarial networks

Increased control over the output from a GAN by conditioning on some variable

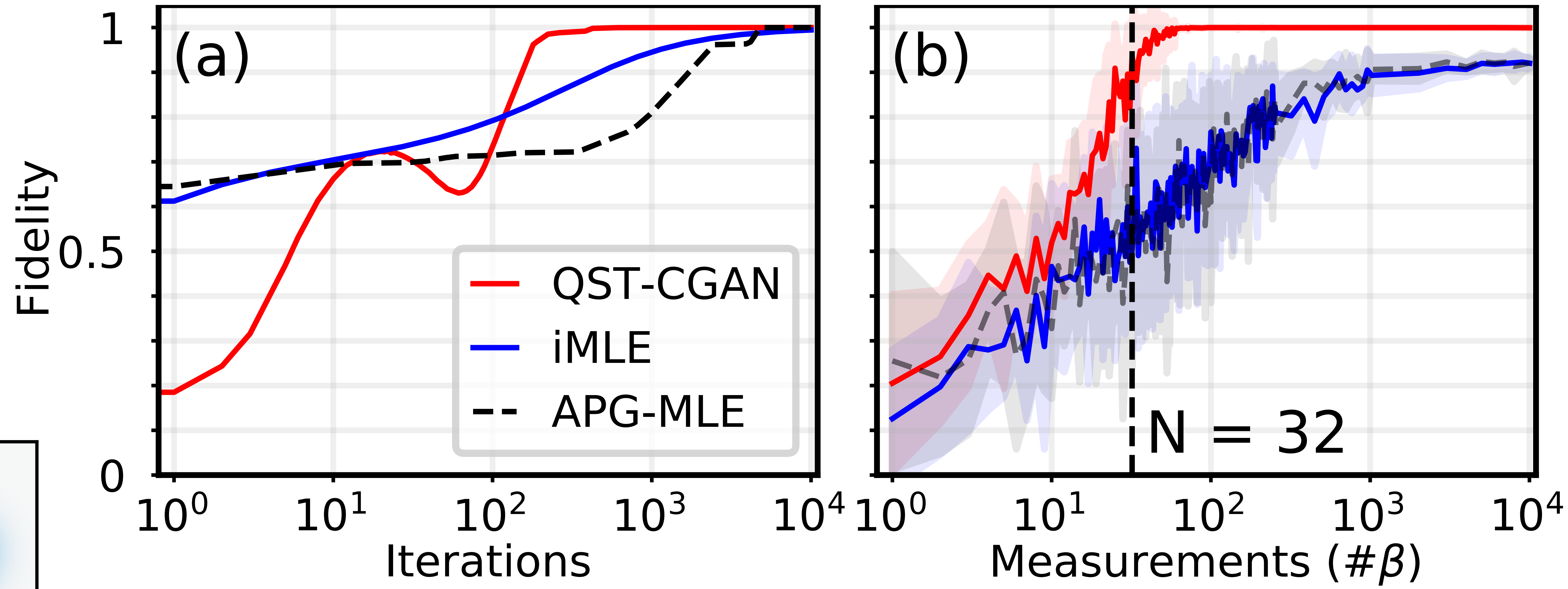
AttGAN (arXiv:1711.10678)

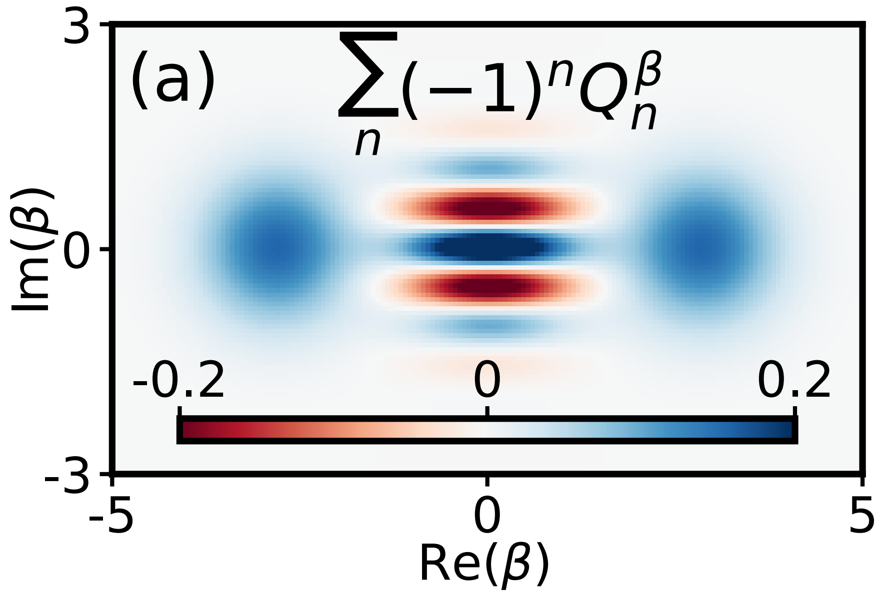
Quantum state tomography with a CGAN



Calculation time and measurement data

Comparing to standard methods based on maximum-likelihood estimation (MLE): iterative MLE (iMLE) and accelerated projected-gradient-based MLE (APG-MLE)





Cat state with 2 photons, 32x32 grid of Husimi Q data

The QST-CGAN converges to high fidelity using much fewer iterations and much less measurement data

Applications

Made the code freely available on GitHub: quantshah/qst-cgan

33 stars

2 watching

8 forks

”I used QST-CGAN extensively for this work. It was REALLY useful. Without it, all quantitative analysis in my manuscript would be impossible.”
— Sangil Kwon, RIKEN, Japan
Nature Communications **15**, 86 (2024)

Summary for QST-CGAN

We have applied conditional generative adversarial networks to the problem of reconstructing quantum states from measurement data

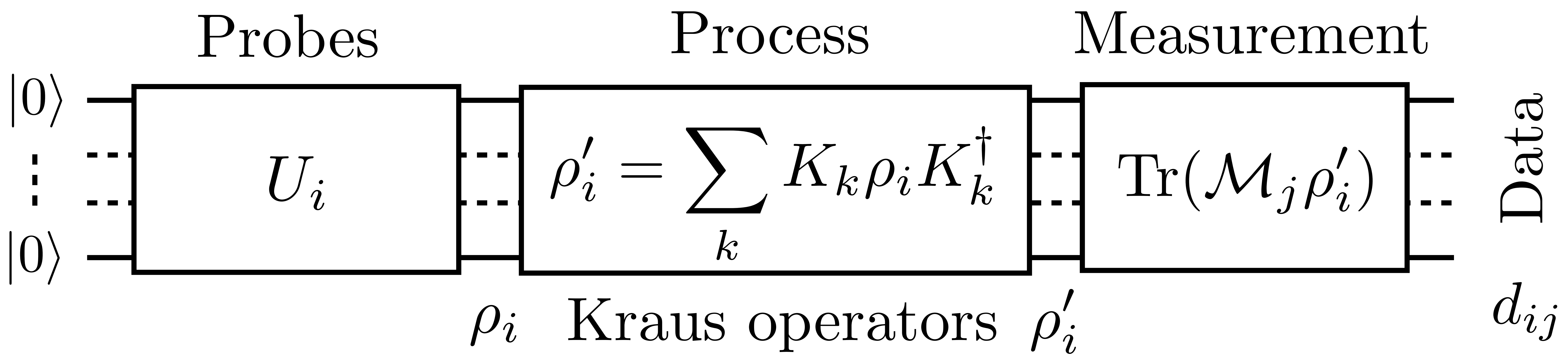
Faster reconstruction (fewer iterations), less data needed, single-shot reconstruction with pre-training, handling various types of noise

S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Lett. **127**, 140502 (2021)
S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Res. **3**, 033278 (2021)
Code available on GitHub: quantshah/qst-cgan

Outline

- Part 1: Quantum state tomography with conditional generative adversarial networks
- Part 2: Quantum process tomography with gradient-descent learning of Kraus operators
- Bonus slides: Photoelectrons and GD-QST
- Summary

Quantum process tomography



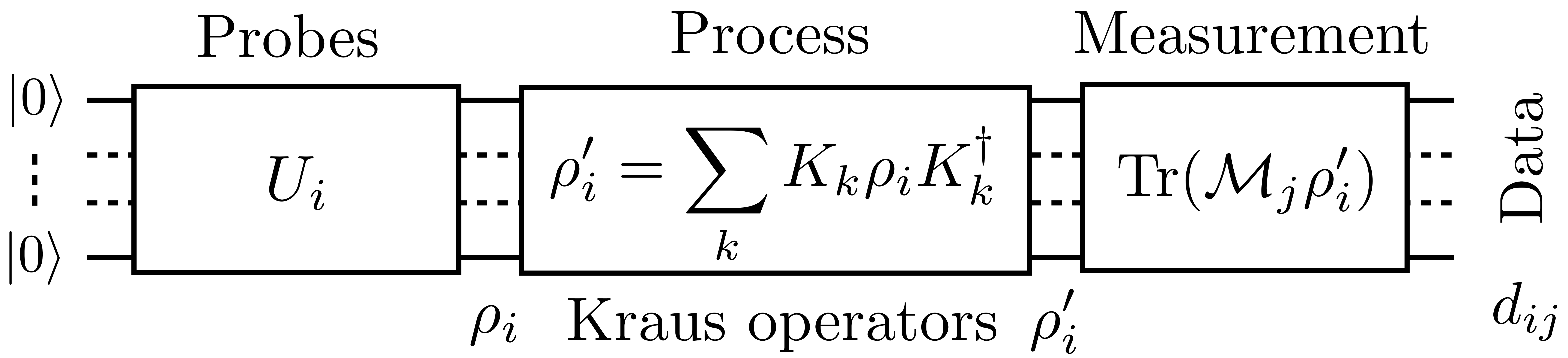
A quantum operation is a completely positive and trace-preserving (TP) linear map $\mathcal{E}(\rho) = \rho'$

Characterizing a quantum operation is called quantum process tomography (QPT)

QPT requires preparing many initial states, applying the operation, and then measuring many different observables

We use the Kraus representation $\mathcal{E}(\rho) = \sum_{l=1}^{k} K_l \rho K_l{}^\dagger$ TP if $\sum_{l=1}^{k} K_l^\dagger K_l = \mathbb{I}$

QPT with gradient descent



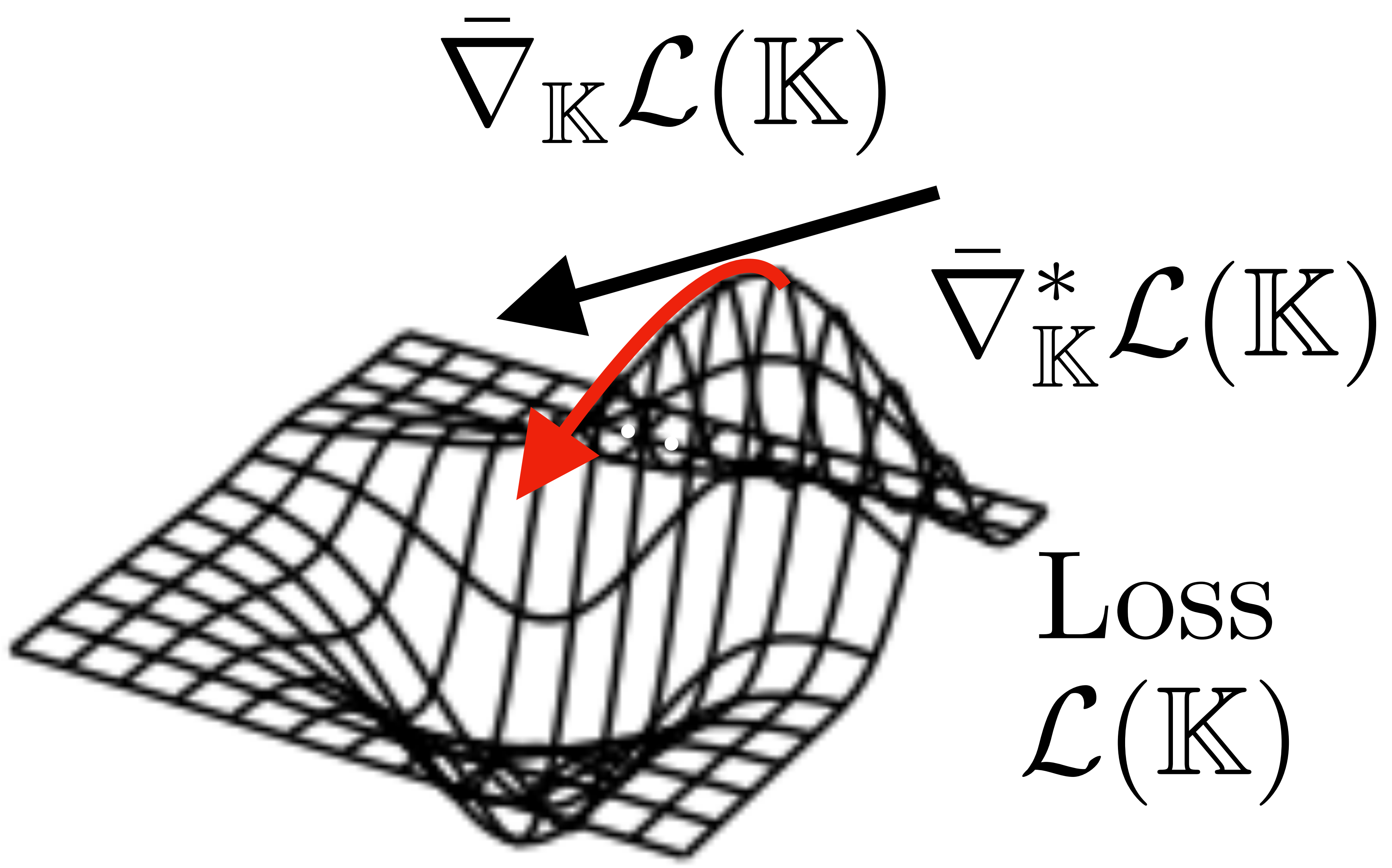
Finding the best Kraus operators for representing the process can be cast as the task of minimising a loss function describing the discrepancy between the data and our estimate

$$\mathcal{L}(\mathbb{K}) = \sum_{ij} \left[d_{ij} - \mathrm{Tr}\left[\mathcal{M}_j \left(\sum_k K_k \rho_i K_k{}^\dagger \right) \right] \right]^2 + \lambda ||\mathbb{K}||_1$$

We use gradient descent for the minimisation

QPT with gradient descent

Keeping the Kraus operators trace-preserving requires restricting them to the so-called Stiefel manifold



A modified update rule for the gradient descent takes care of this

$$\bar{\nabla}^*_{\mathbb{K}}\mathcal{L}(\mathbb{K}) = A(\mathbb{I} + \frac{\eta}{2}B^\dagger A)^{-1}B^\dagger\mathbb{K}$$

$$A = [G \quad \mathbb{K}] \quad B = [\mathbb{K} \quad -G]$$

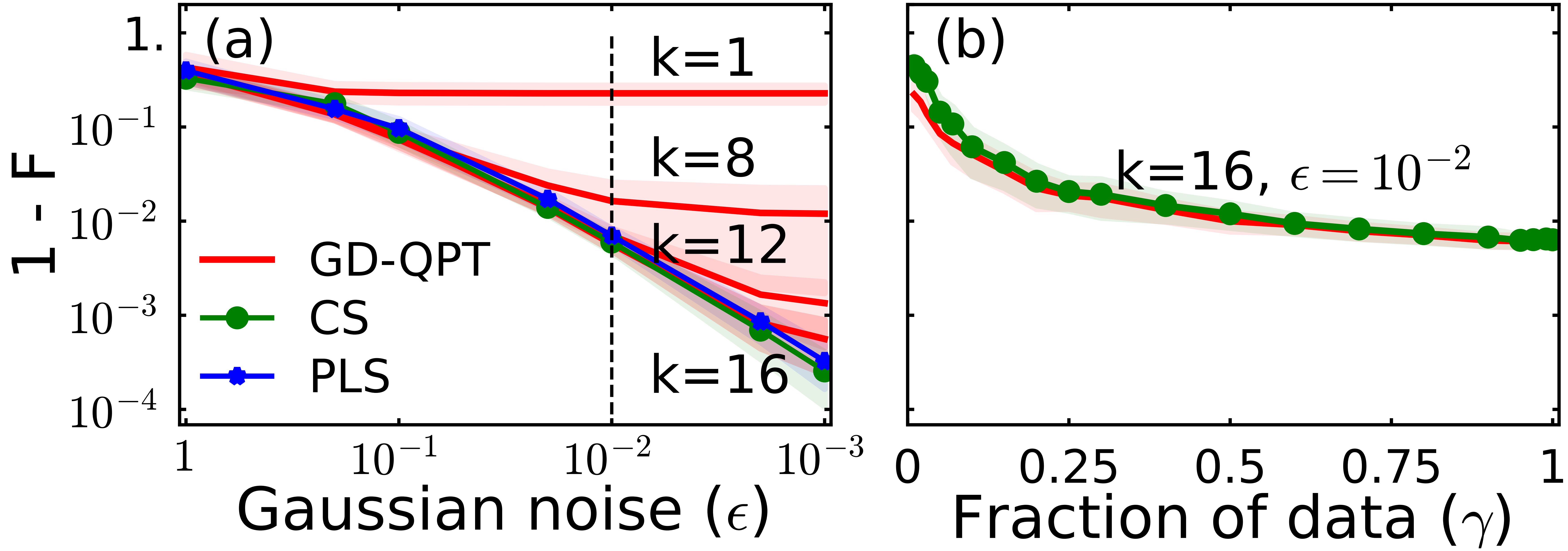
$$G' = \bar{\nabla}_{\mathbb{K}}\mathcal{L}(\mathbb{K}) \quad G = G'/||G'||_2$$

η is a learning rate

$$\mathbb{K}' = \mathbb{K} - \eta\bar{\nabla}^*_{\mathbb{K}}\mathcal{L}(\mathbb{K})$$

Benchmarking GD-QPT

Comparing against compressed sensing (CS) and projected least squares (PLS)

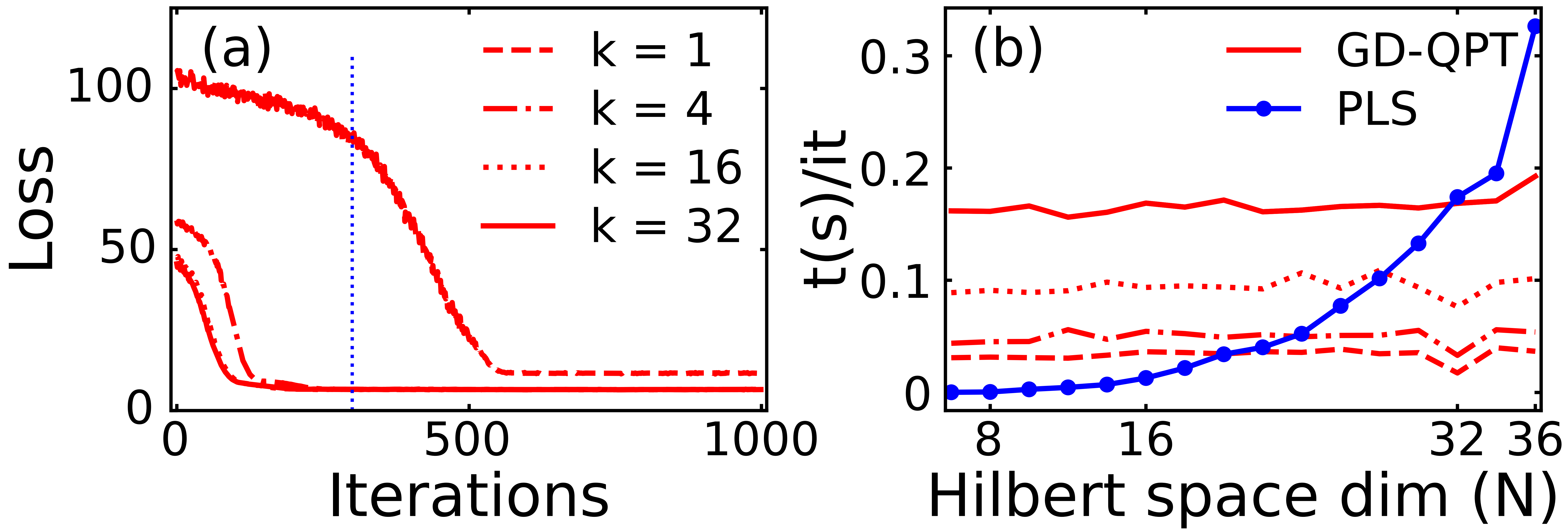


Random two-qubit processes of full rank $k = 16$

GD-QPT matches CS in its specialty — dealing with little data

Benchmarking GD-QPT

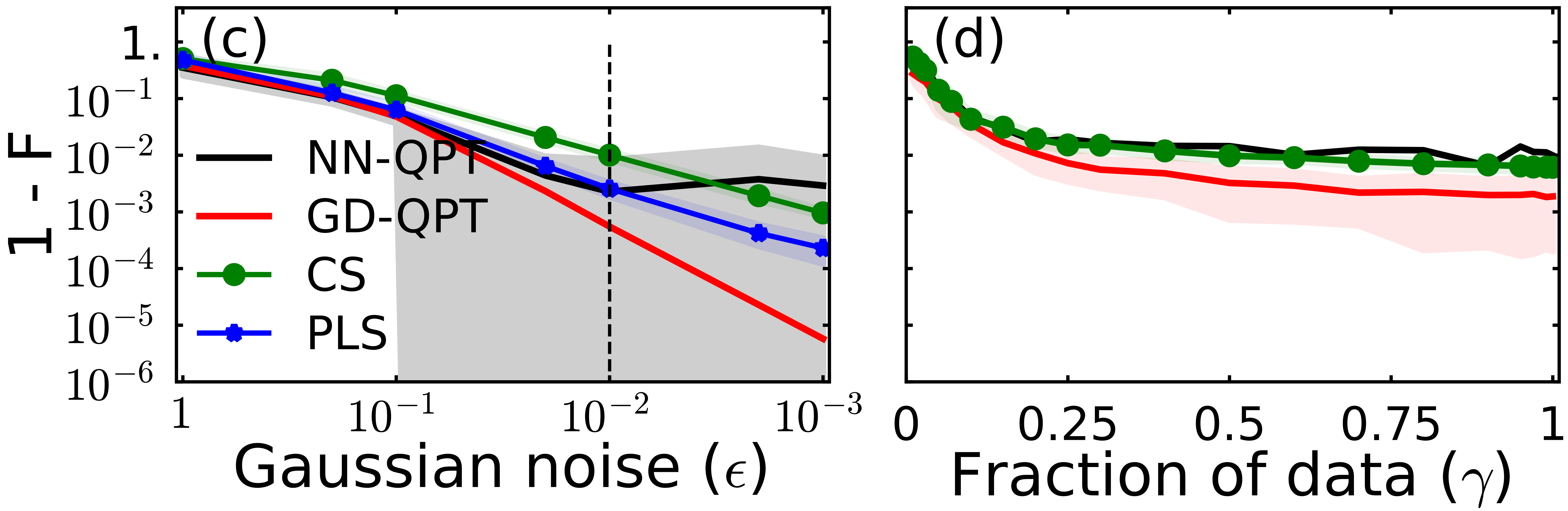
CS cannot handle larger systems, but PLS can — GD-QPT does at least as well as PLS here



Random five-qubit processes of rank $k = 3$

What about neural networks?

We trained a feed-forward neural network to output Kraus operators optimising agreement with data

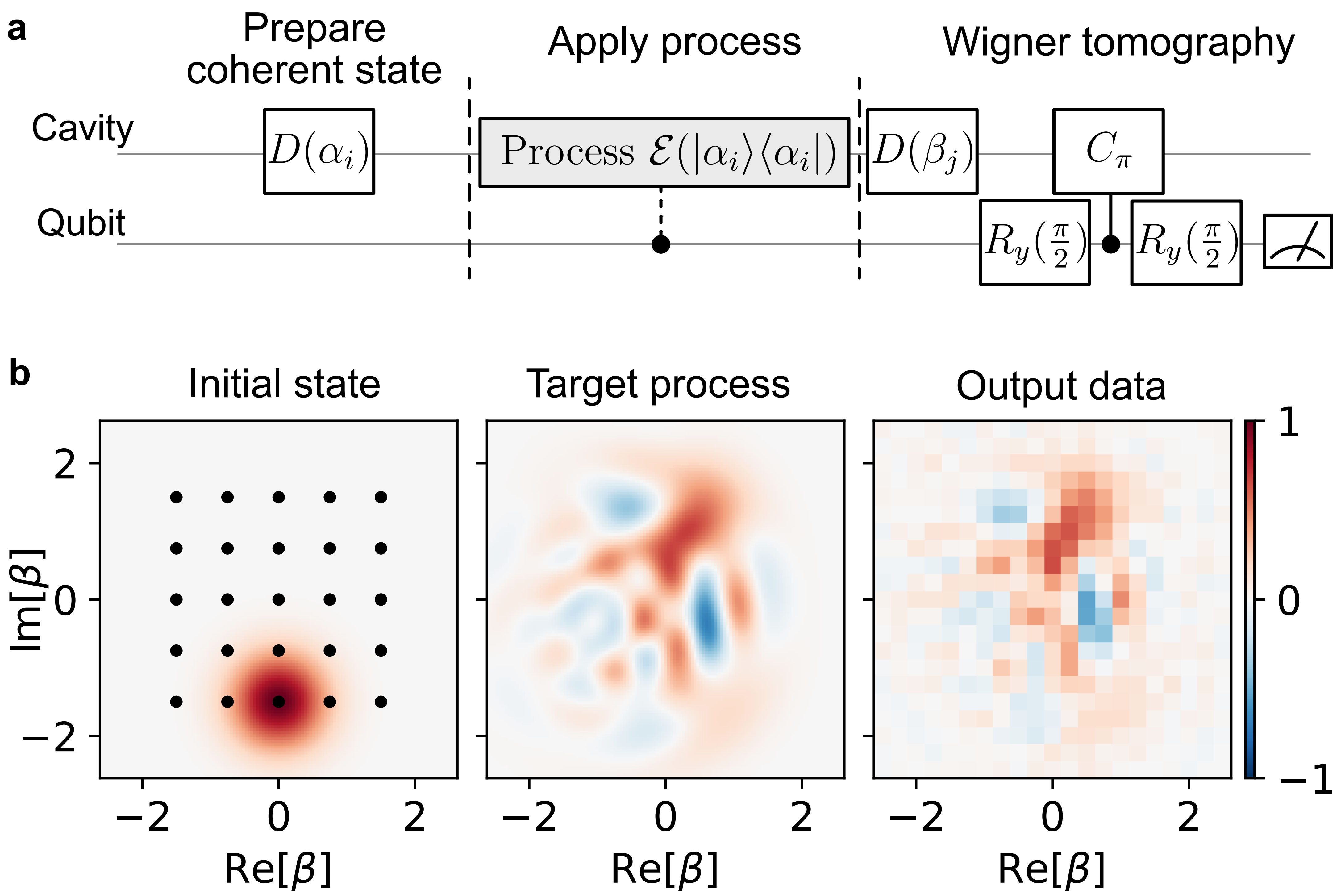


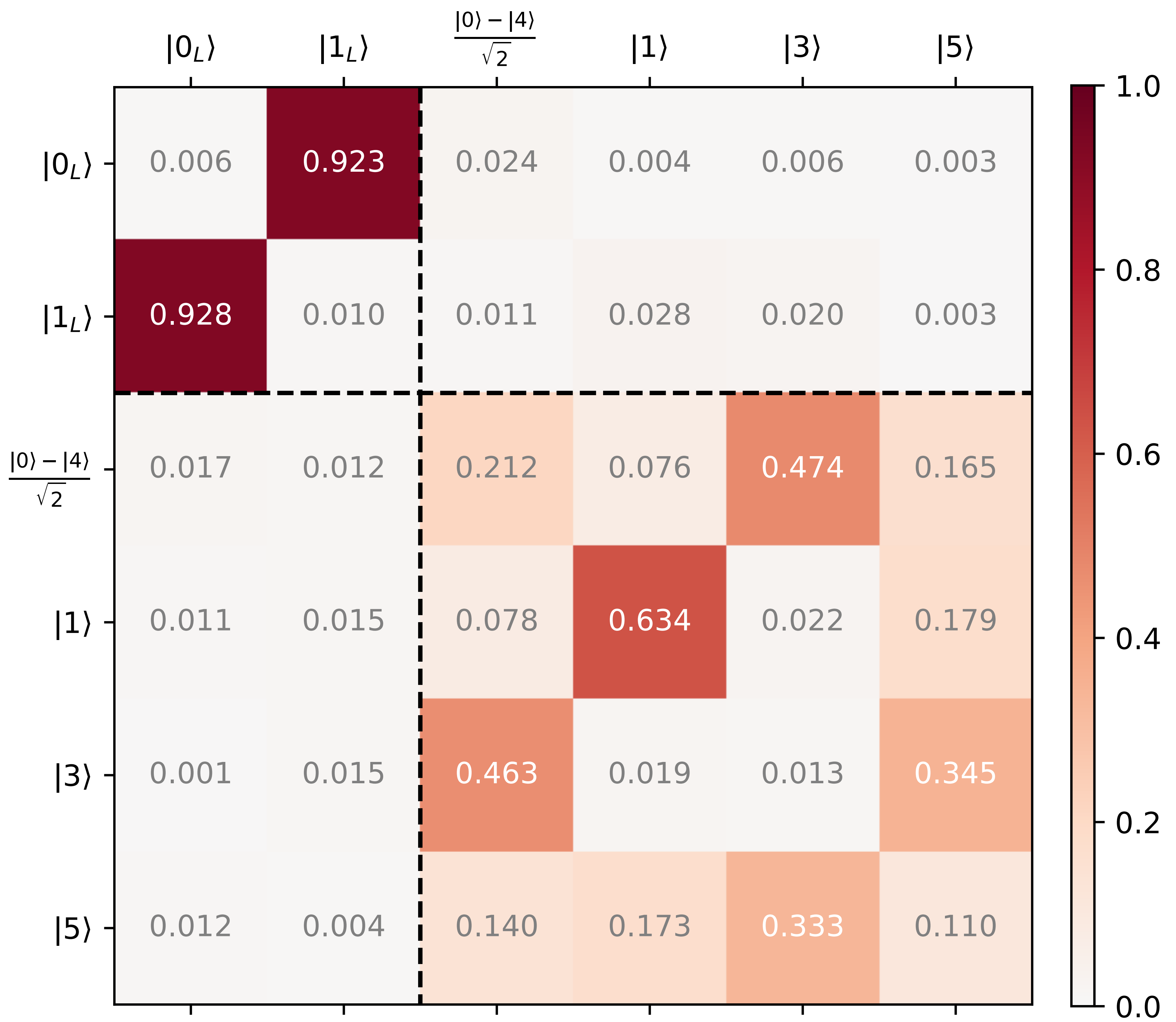
Random two-qubit processes of rank $k = 4$

No advantage from the NN

Applications

Enabled extensive characterization at Chalmers of a logical gate on a qubit encoded in states of a harmonic oscillator





M. Kervinen et al., Physical Review A **110**, L020401 (2024)

Summary for GD-QPT

We have applied gradient descent on the Stiefel manifold to the problem of reconstructing quantum processes from measurement data

GD-QPT combines the best of two worlds — it can both reconstruct processes from little data and in relatively large Hilbert spaces

Neural networks don't seem to improve on GD-QPT

S. Ahmed, F. Quijandría, and A. F. Kockum, Phys. Rev. Lett. **130**, 150402 (2023)
Code available on GitHub: quantshah/gd-qpt

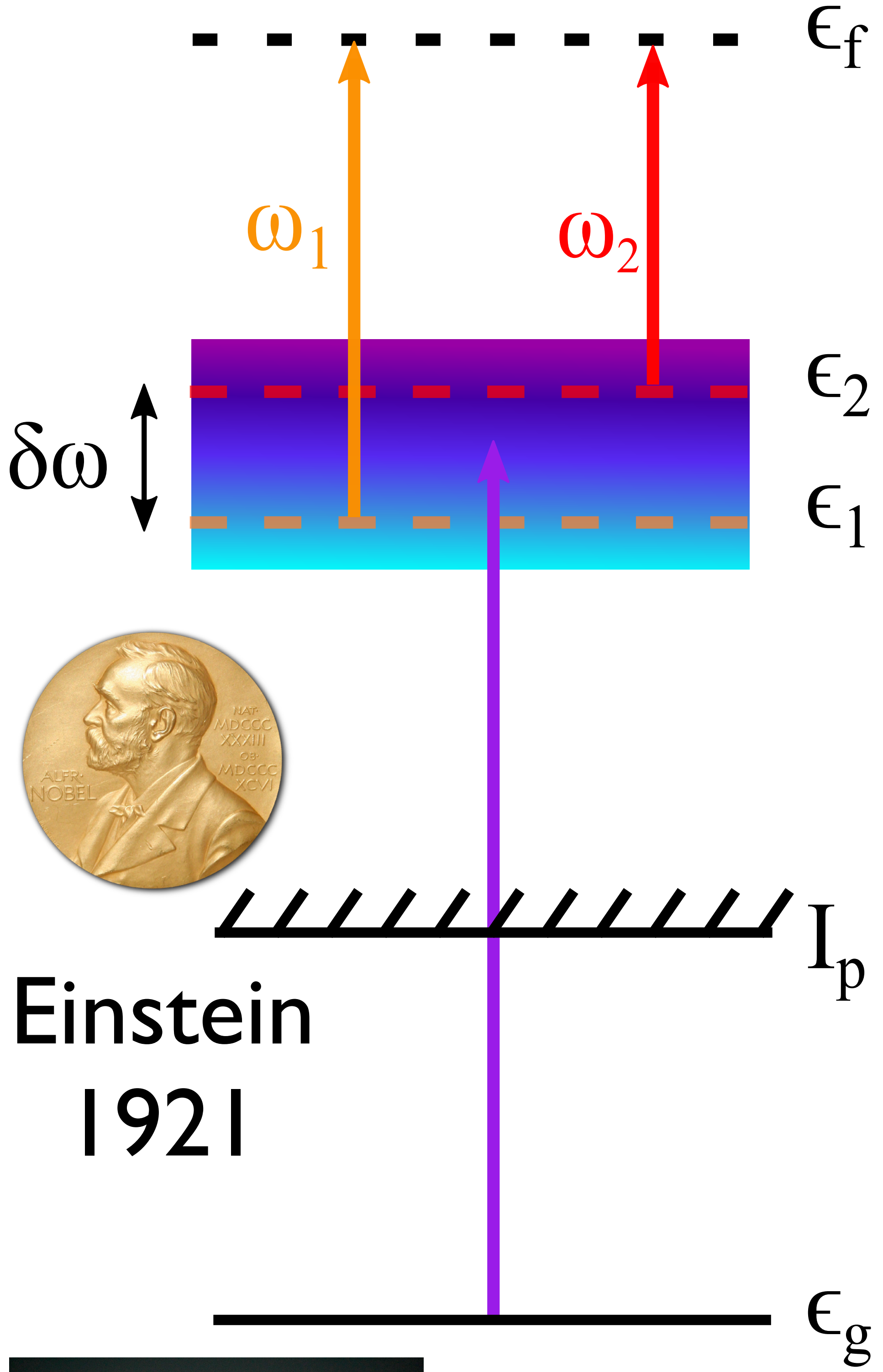
19 stars

2 watching

Outline

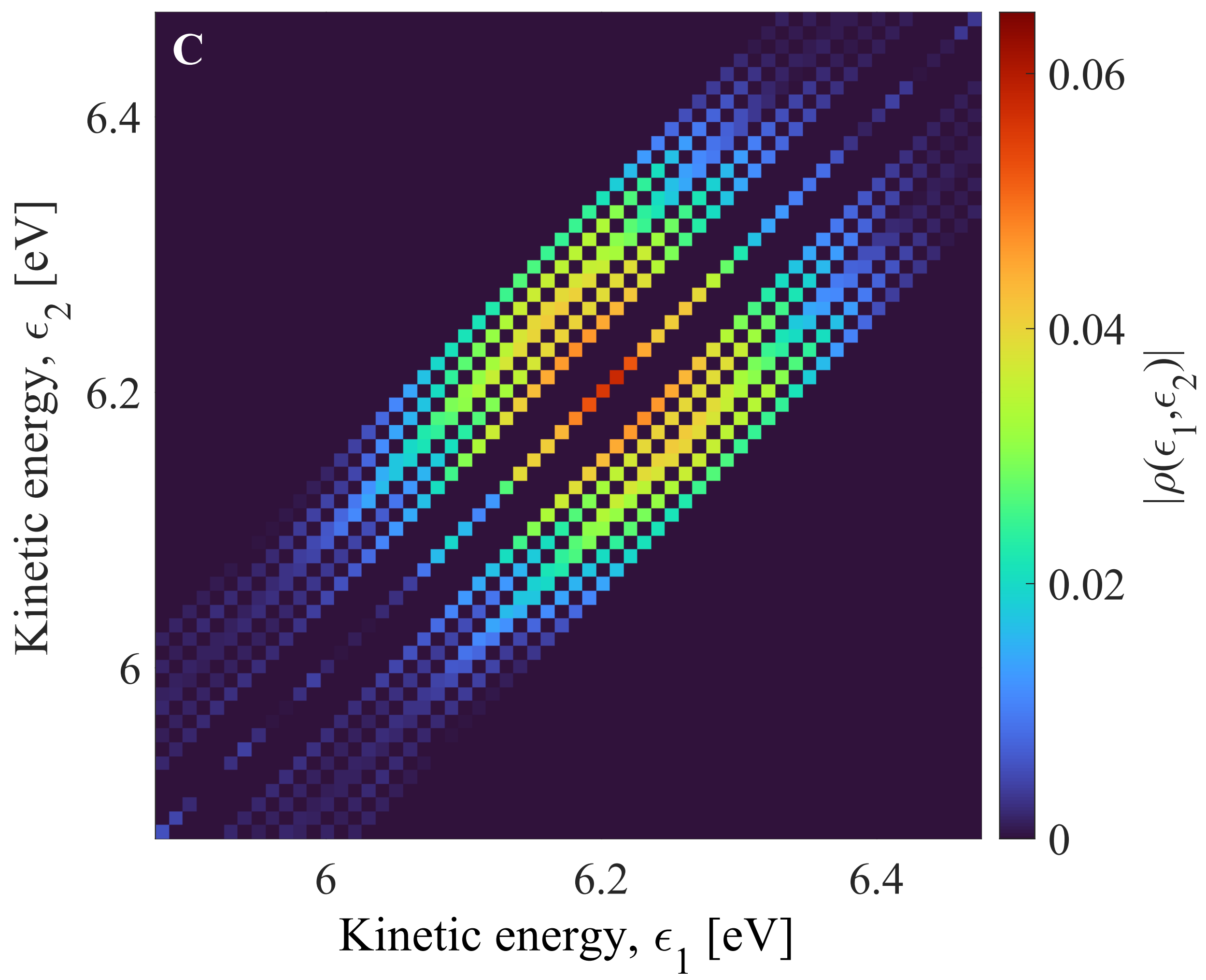
- Part 1: Quantum state tomography with conditional generative adversarial networks
- Part 2: Quantum process tomography with gradient-descent learning of Kraus operators
- Bonus slides: Photoelectrons and GD-QST
- Summary

Nobel bonus: Measuring the state of photoelectrons

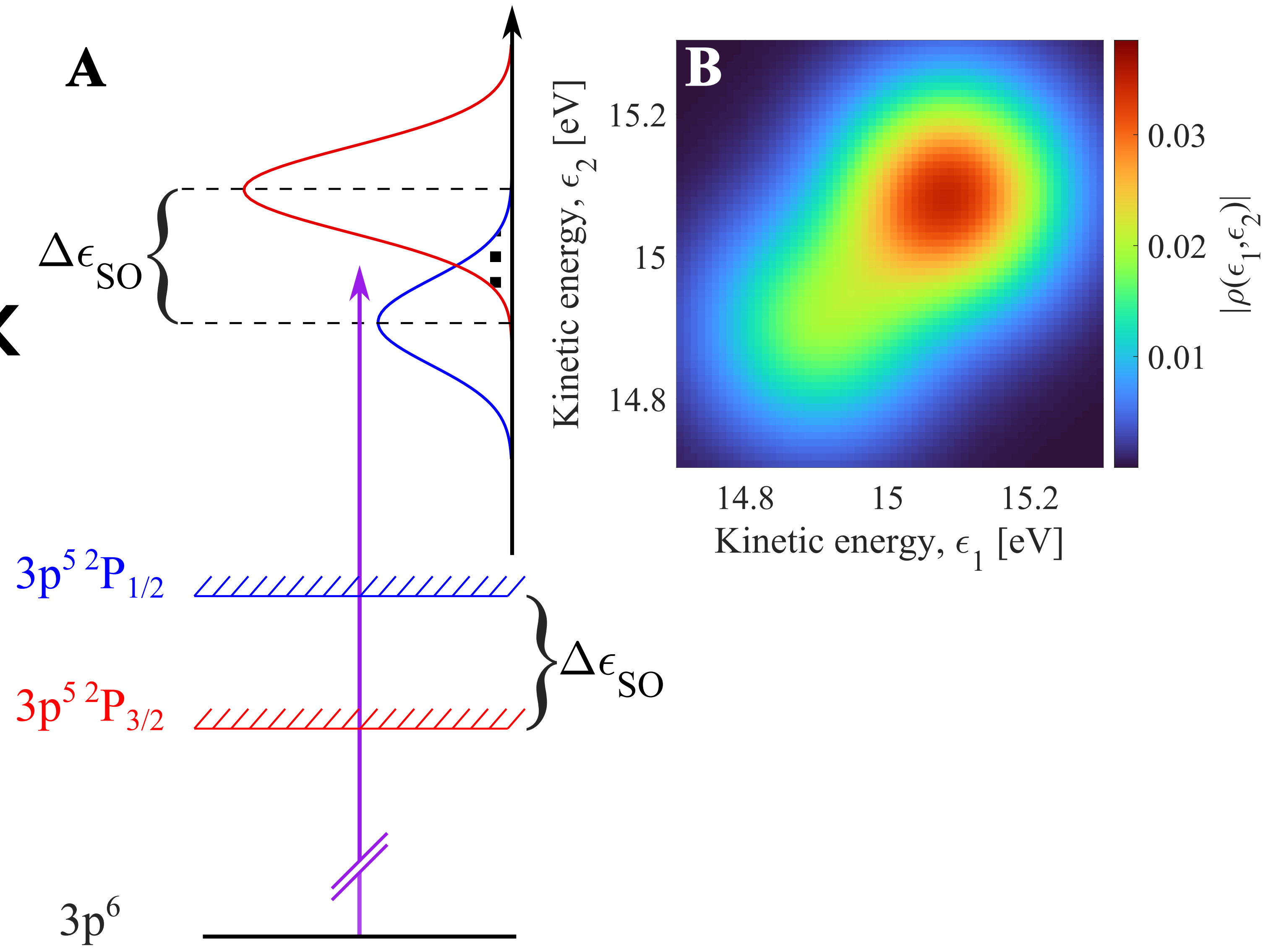


Einstein
1921

Ionizing atoms, adding a bichromatic IR probe, vary detuning to extract some diagonals in the photoelectron density matrix



Need priors to fill out the full density matrix correctly from data — Bayesian estimation



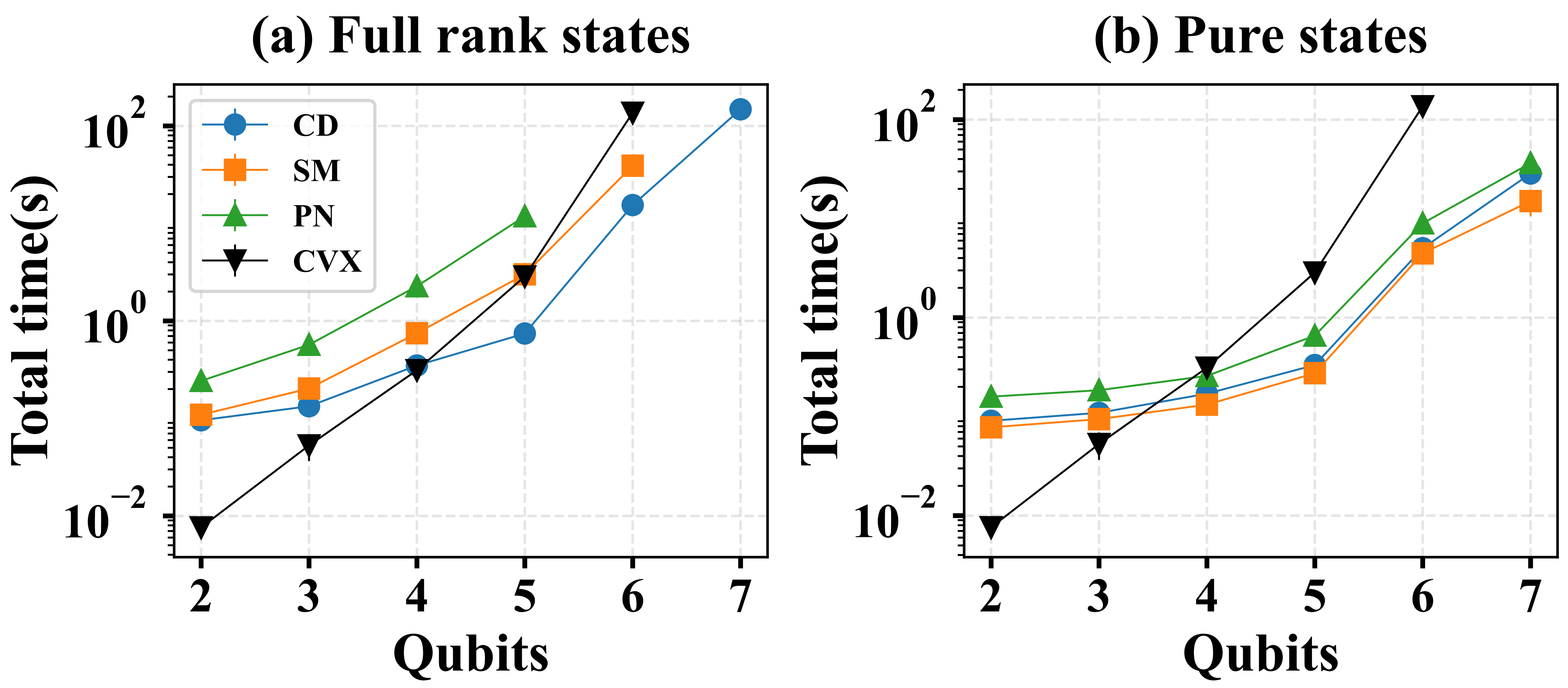
Spin-orbit interaction in argon reflected in the photoelectron state

H. Laurell, …, S. Ahmed, …, A. F. Kockum, Anne L'Huillier, and D. Busto, arXiv:2309.13945 (2023)
to appear in Nature Photonics; code to be made available

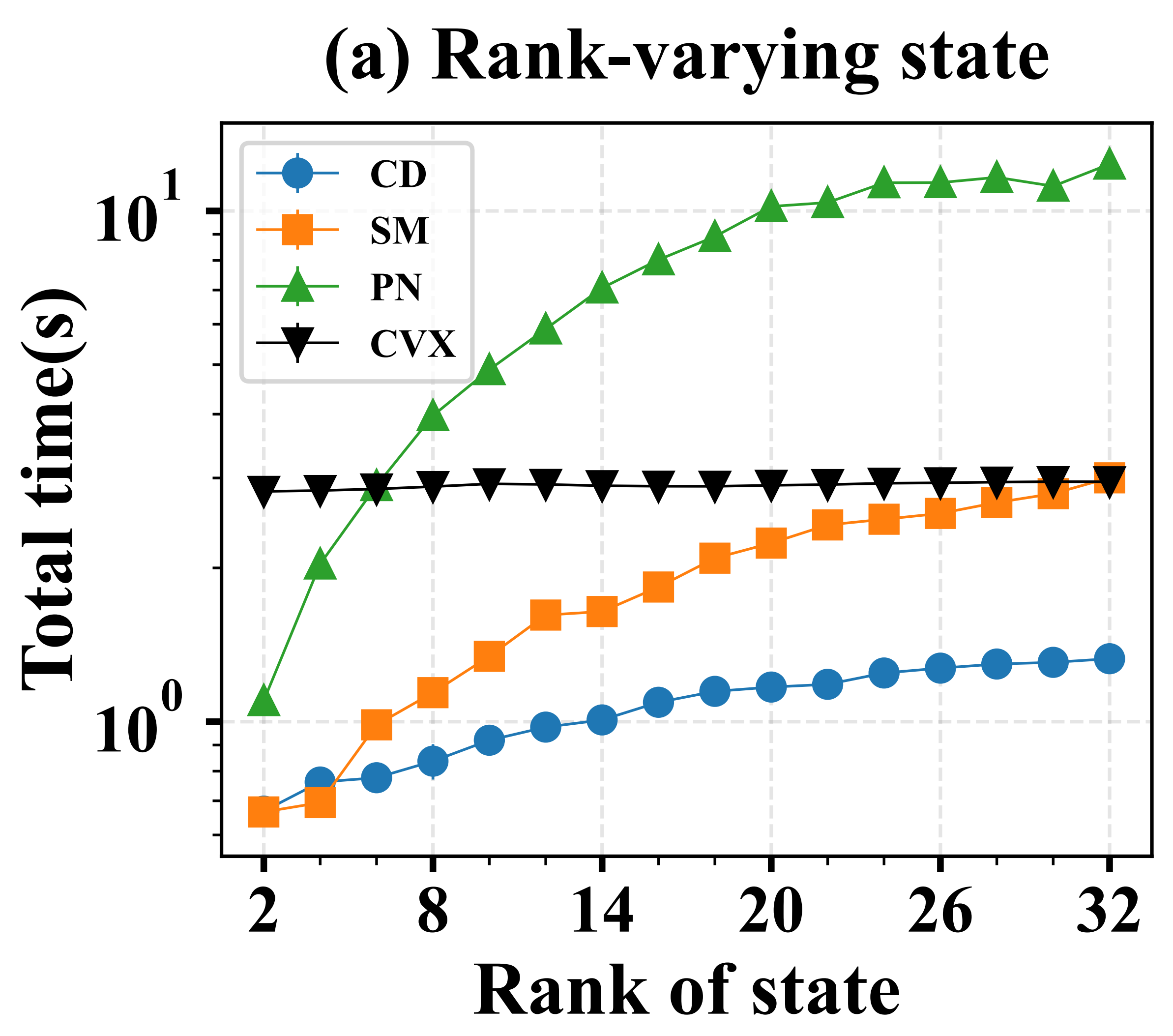
Quantum state tomography with gradient descent

3 parameterizations: Cholesky decomposition, Stiefel manifold, projective normalization

Able to control the rank of the ansatz in all three



5 qubits:



Here comparing to convex optimization;
other benchmarks in progress

A. Gaikwad, M. S. Torres Hernandez, S. Ahmed, and A. F. Kockum, in preparation (2024)

Outline

- Part 1: Quantum state tomography with conditional generative adversarial networks
- Part 2: Quantum process tomography with gradient-descent learning of Kraus operators
- Bonus slides: Photoelectrons and GD-QST
- Summary

Summary and outlook

@quantshah
@AntonFKockum

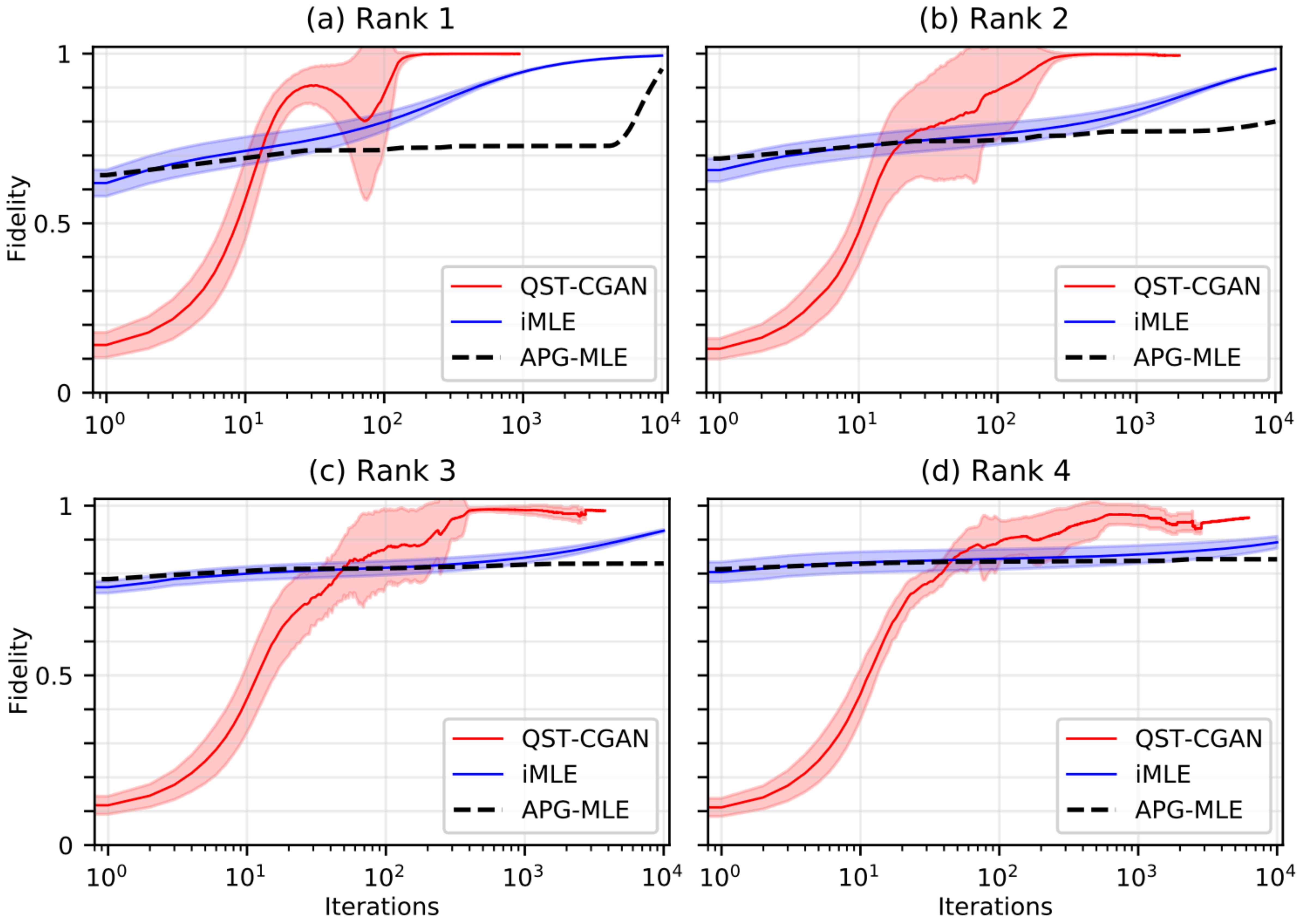
- We have tackled quantum state tomography and quantum process tomography with machine learning and gradient descent
- Improvements over state-of-the-art methods when it comes to computation time, amount of data needed, handling of noise, …
- Important design choice: build in physics knowledge in the state/process representation
- Ongoing/next: revisit QST without NNs, approximate state/process representations to handle larger systems, elucidate error bars, …
- Look for applications beyond quantum computing!

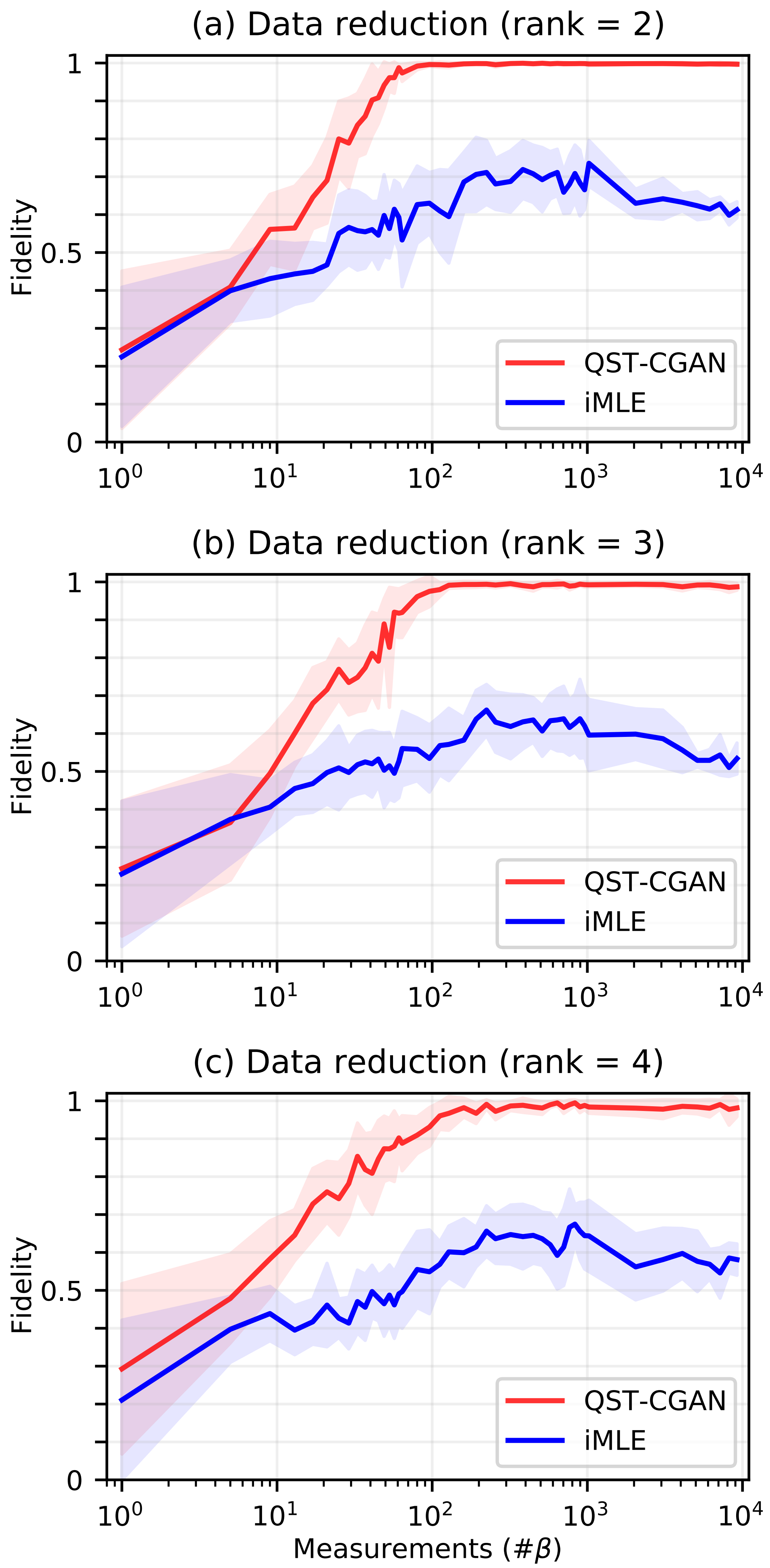
S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Lett. **127**, 140502 (2021)
S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Phys. Rev. Res. **3**, 033278 (2021)
S. Ahmed, F. Quijandría, and A. F. Kockum, Phys. Rev. Lett. **130**, 150402 (2023)
H. Laurell, …, S. Ahmed, …, A. F. Kockum, A. L'Huillier, and D. Busto, arXiv:2309.13945 (2023)
A. Gaikwad, M. S. Torres Hernandez, S. Ahmed, and A. F. Kockum, in preparation (2024)
Code available on GitHub: quantshah/qst-cgan and quantshah/gd-qpt

Reconstructing mixed states

Mixed states (higher rank) are harder to parameterize

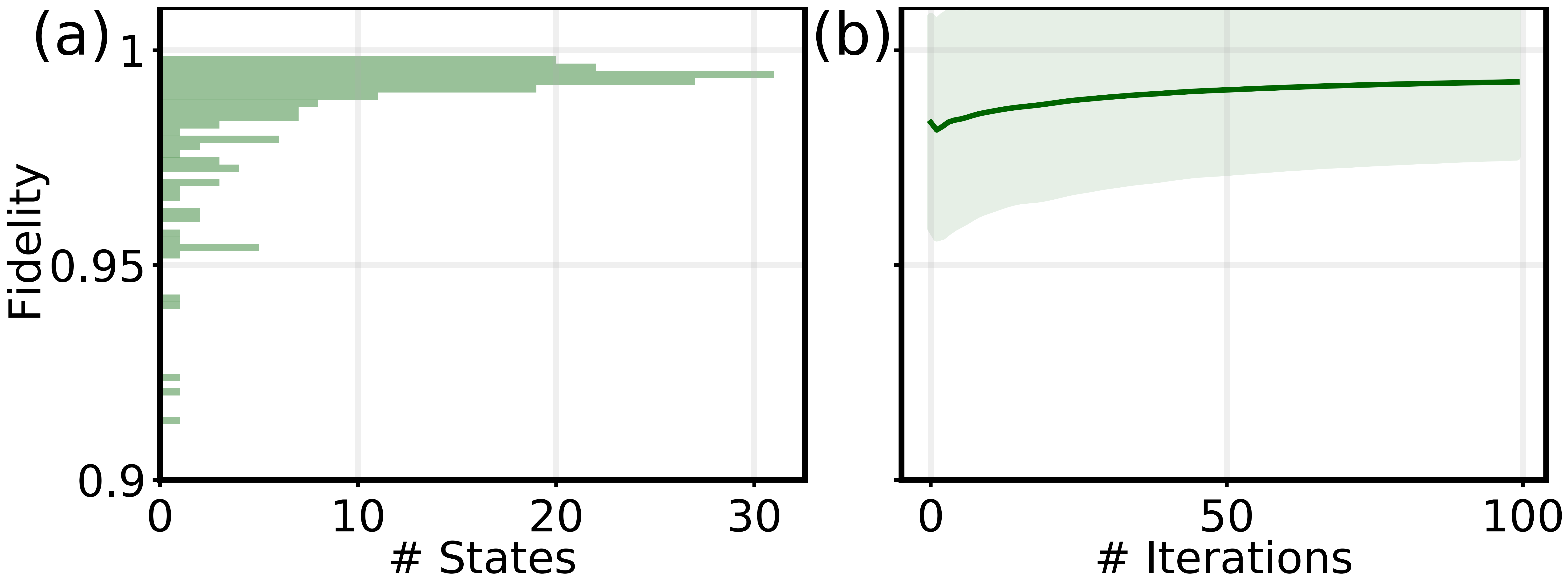
$$\rho' = 0.8 * \text{cat} + (0.2/(r-1)) \sum_{n=0}^{r-2} \text{fock}(n); r = 2,3,\ldots$$





Single-shot reconstruction

Training the QST-CGAN on simulated data from a certain type of states allows it to directly output a density matrix for new data without iterating

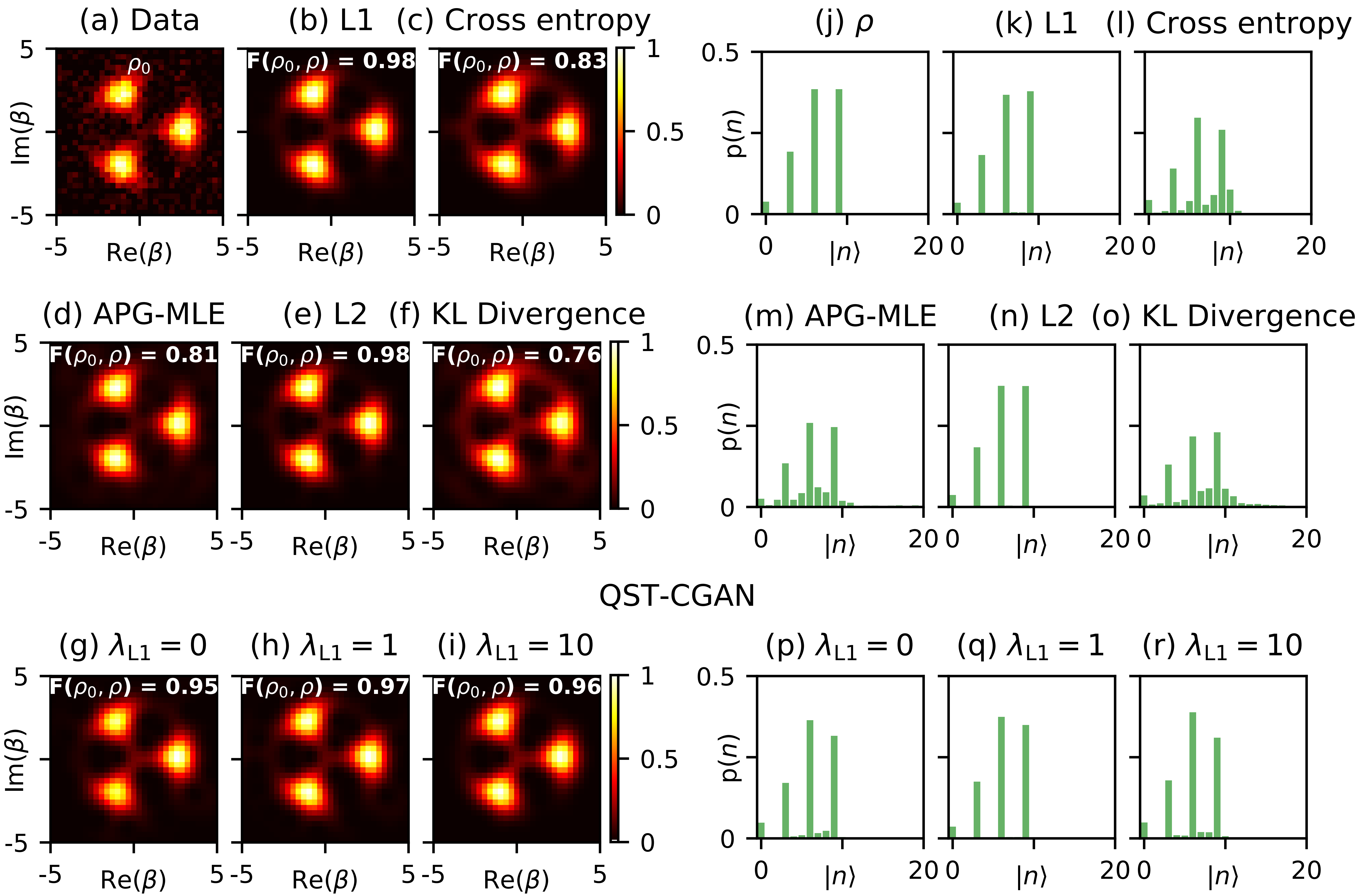


Dealing with noisy data

Data with additive Gaussian noise, e.g., from insufficient averaging

Same noise level included in the noise layer of the QST-CGAN

The discriminator adapts well



WACQT facts

Knut och Alice Wallenbergs Stiftelse

wacqt.se
@wacqt_sweden

Main goals:	i) To build a broad competence base in Sweden for Quantum Technology ii) To build a quantum computer based on superconducting circuits
Two parts:	Core project on quantum computing Excellence program including all of Quantum Technology
Universities:	Chalmers (Director: Per Delsing) KTH, Lund Univ, Stockholm Univ, Linköping Univ
Duration:	12 years (3+4+3+2 years), started 1/1 2018
Involving industry:	SME for enabling technology Big industry for applications
Funding:	~150 M$ KAW Universities Industry partners Quantum technology flagship: OpenSuperQ, EuroQCI