NISQ applications and the geometry of quantum states

JOHANNES JAKOB MEYER, FU BERLIN

@jj_xyz

What is Geometry?

geometry

/dʒɪˈɒmɪtri/

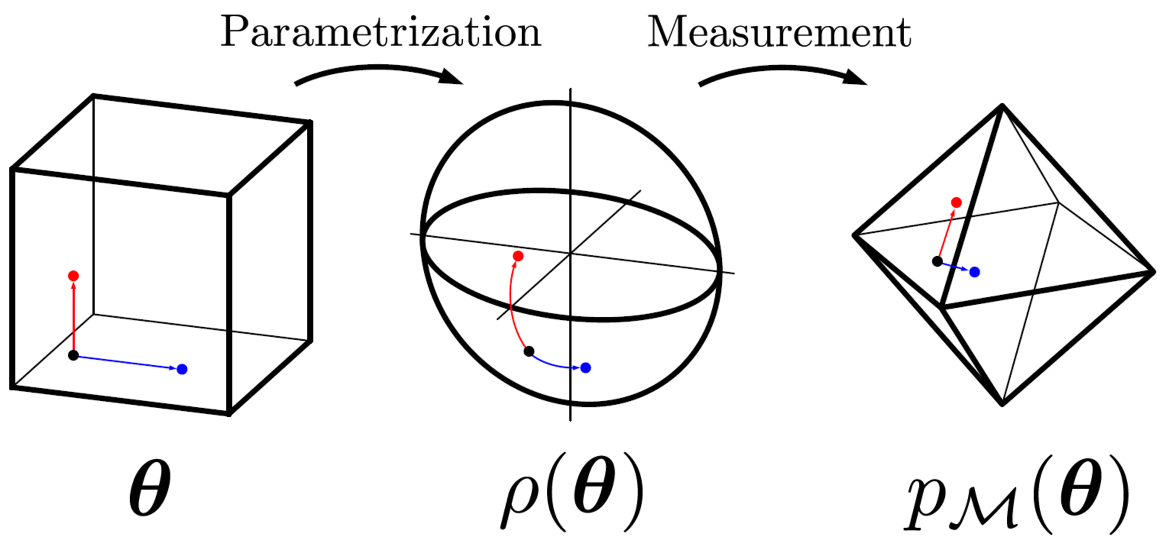
noun

the branch of mathematics concerned with the properties and relations of points, lines, surfaces, solids, and higher dimensional analogues.

Geometry of quantum states

Where are quantum states in relation to each other?

Parametrized Quantum States



Encountered in

- Variational Quantum Algorithms
- Quantum Metrology
- Quantum Optimal Control

Based on Quantum 5, 539

Quantum
the open journal for quantum science

HOME PUBLICATIONS

Fisher Information in Noisy Intermediate-Scale Quantum Applications

Johannes Jakob Meyer

Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
QMATH, Department of Mathematical Sciences, Københavns Universitet, 2100 København Ø, Denmark

Published:	2021-09-09, volume 5, page 539
Eprint:	arXiv:2103.15191v3
Doi:	https://doi.org/10.22331/q-2021-09-09-539
Citation:	Quantum 5, 539 (2021).

GET FULL TEXT PDF READ ON ARXIV VANITY COMMENT ON FERMAT'S LIBRARY

Distances Between Parametrized States

We want to understand parametrized quantum states in *parameter space*. One way to do so is via *distances*

→ Perform a *pullback* of a distance between quantum states

$$d(\boldsymbol{\theta}, \boldsymbol{\theta}') = d(\rho(\boldsymbol{\theta}), \rho(\boldsymbol{\theta}'))$$

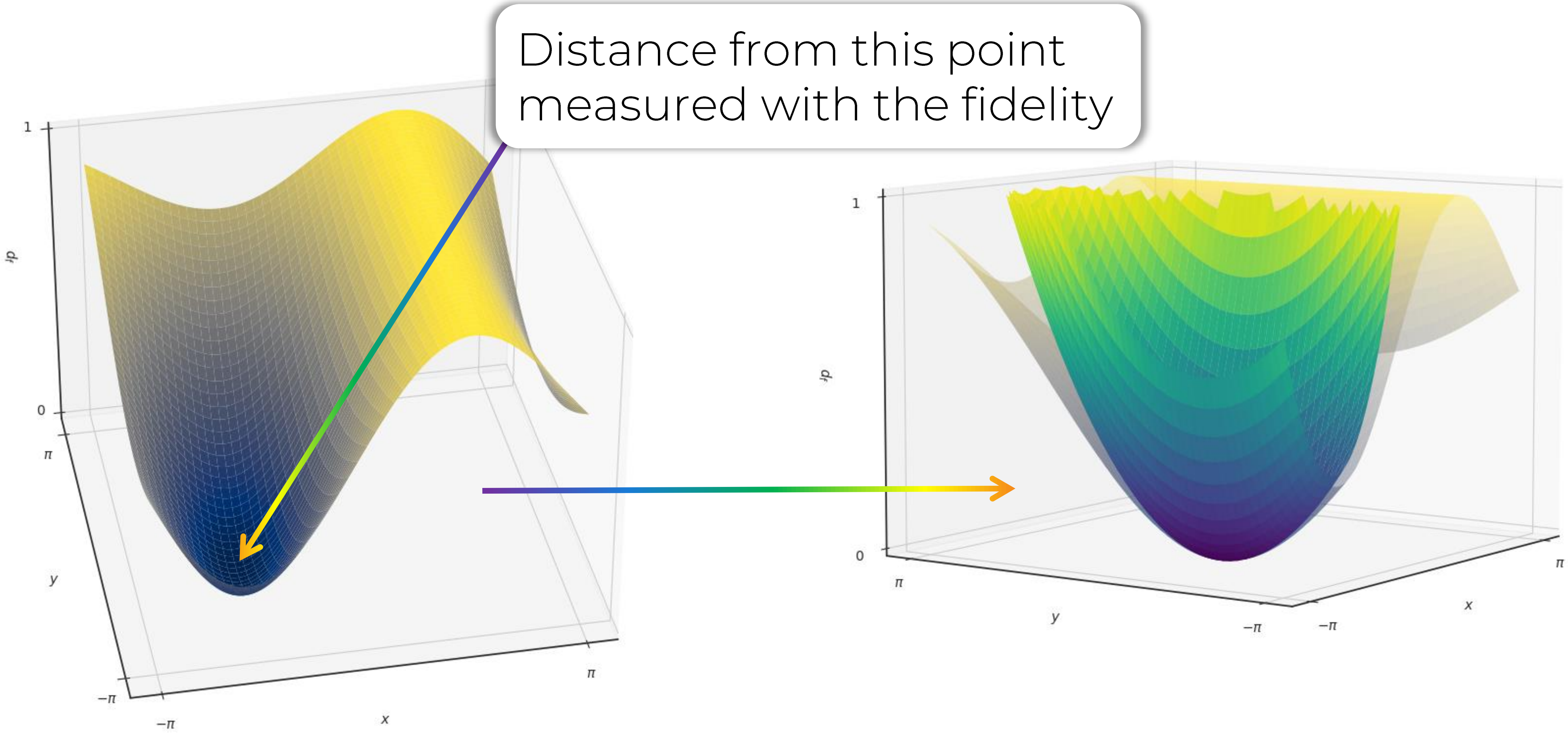
→ or between the output probability distributions

$$d_{\mathcal{M}}(\boldsymbol{\theta}, \boldsymbol{\theta}') = d(p_{\mathcal{M}}(\boldsymbol{\theta}), p_{\mathcal{M}}(\boldsymbol{\theta}'))$$

We require $d(\boldsymbol{\theta}, \boldsymbol{\theta}') \geq 0$, $d(\boldsymbol{\theta}, \boldsymbol{\theta}) = 0$, and *differentiability*.

Small Changes

What happens in the local neighbourhood of the parameter $\boldsymbol{\theta}$?



Small Changes

What happens in the local neighbourhood of the parameter $\boldsymbol{\theta}$?

- Mathematically speaking, the Taylor expansion gives

$$d(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\delta}) = \frac{1}{2}\boldsymbol{\delta}^T M(\boldsymbol{\theta})\boldsymbol{\delta} + O(\|\boldsymbol{\delta}\|^3)$$

with the *Hessian* of the pullback distance relative to a fixed point in parameter space $\boldsymbol{\theta}$,

$$M(\boldsymbol{\theta})_{ij} = \frac{\partial^2}{\partial \delta_i \partial \delta_j} d(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\delta})\Big|_{\boldsymbol{\delta}=0}$$

Information Matrices

The Hessian induces an inner product $\langle \boldsymbol{\delta}, \boldsymbol{\delta}' \rangle_M = \boldsymbol{\delta}^T M \boldsymbol{\delta}'$

With it we can measure

- lengths $\|\boldsymbol{\delta}\| = \sqrt{\langle \boldsymbol{\delta}, \boldsymbol{\delta} \rangle}$
- distances $d(\boldsymbol{\delta}, \boldsymbol{\delta}') = \|\boldsymbol{\delta} - \boldsymbol{\delta}'\| = \sqrt{\langle \boldsymbol{\delta} - \boldsymbol{\delta}', \boldsymbol{\delta} - \boldsymbol{\delta}' \rangle}$
- angles $\sphericalangle(\boldsymbol{\delta}, \boldsymbol{\delta}') = \arccos(\langle \boldsymbol{\delta}, \boldsymbol{\delta}' \rangle / \|\boldsymbol{\delta}\| \|\boldsymbol{\delta}'\|)$

The Hessian contains information about the underlying quantum state, by nature an information-theoretic object. We thus call it *information matrix*.

Classical Fisher Information Matrix

Use the KL divergence (relative entropy) for the pullback

And find the *classical Fisher information matrix (CFIM)* with entries

$$I_{ij}(\boldsymbol{\theta}) = \sum_l \frac{1}{p_l(\boldsymbol{\theta})} \frac{\partial p_l(\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial p_l(\boldsymbol{\theta})}{\partial \theta_j}$$

Intuition: the classical Fisher information quantifies how much a change of parameters changes the underlying probability distribution

How can we calculate it?

To calculate the classical Fisher information matrix, we need

output probabilities

$p_l(\boldsymbol{\theta})$

→ Output of experiment (histogram)

→ Can also use more sophisticated techniques: Bayesian approach, Machine Learning

their derivatives

$\partial_i p_l(\boldsymbol{\theta})$

→ Finite differences

→ Parameter shift rule

Fun Fact: Uniqueness

What happens if we used a different distance?

Theorem (Morozova & Chentsov):

The information matrix associated to any monotonic distance measure between probability distributions will be a positive scalar multiple of the classical Fisher information matrix

Monotonicity: $d(T[p], T[q]) \leq d(p, q)$ for all stochastic maps T

Quantum Fisher Information Matrix (Pure States)

Use the fidelity distance for the pullback

And find the *quantum Fisher information matrix (QFIM)* with entries

$$\mathcal{F}_{ij}(\boldsymbol{\theta}) = 4\,\mathrm{Re}[\langle \partial_i \psi(\boldsymbol{\theta}) | \partial_j \psi(\boldsymbol{\theta}) \rangle - \langle \partial_i \psi(\boldsymbol{\theta}) | \psi(\boldsymbol{\theta}) \rangle \langle \psi(\boldsymbol{\theta}) | \partial_j \psi(\boldsymbol{\theta}) \rangle]$$

Intuition: the quantum Fisher information quantifies how much a change of parameters changes the underlying quantum state

Calculation (Pure States)

For parameters that are encoded via parallel Hamiltonian evolution

$$\mathcal{F}_{ij}(\boldsymbol{\theta}) = 4[\langle\psi(\boldsymbol{\theta})|H_i H_j|\psi(\boldsymbol{\theta})\rangle - \langle\psi(\boldsymbol{\theta})|H_i|\psi(\boldsymbol{\theta})\rangle\langle\psi(\boldsymbol{\theta})|H_j|\psi(\boldsymbol{\theta})\rangle]$$

We can also use perturbations of the fidelity distance itself and tricks using the parameter-shift rule. To get approximation on the matrix level, an analogue of SPSA can be used.

Calculation (Noisy States)

In principle, full tomography is necessary to calculate the quantum Fisher information for noisy states

Variational methods have been proposed, but with high overheads and relying on the success of variational subroutines

An alternative is to approximate the quantum Fisher information, *e.g.* via the *truncated* quantum Fisher information or hierarchical quantities that can be computed from classical shadows

Fun Fact: Non-Uniqueness

Theorem (Petz):

There are infinitely many information matrices arising from monotonic distances between quantum states.

Monotonicity: $d(\Phi[\rho], \Phi[\sigma]) \leq d(\rho, \sigma)$ for all quantum channels Φ

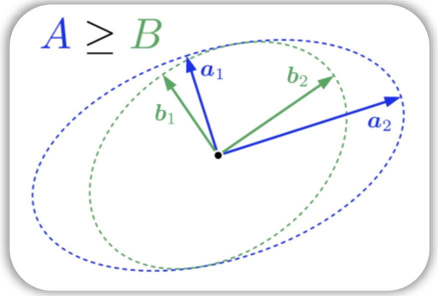
„Our“ quantum Fisher information matrix is also known as *SLD-QFIM*, and it is the smallest of the bunch.

Classical and Quantum Fisher Information

Monotonicity of the underlying distance measure implies that

$$I(T[p(\boldsymbol{\theta})]) \leq I(p(\boldsymbol{\theta})) \quad \text{for all stochastic maps } T$$

$$\mathcal{F}(\Phi[\rho(\boldsymbol{\theta})]) \leq \mathcal{F}(\rho(\boldsymbol{\theta})) \quad \text{for all quantum channels } \Phi$$



But measurements are also channels and the quantum Fisher information for a classical state is equal to the classical Fisher information. Therefore

$$\mathcal{F}(\mathcal{M}[\rho(\boldsymbol{\theta})]) = I(\mathcal{M}[\rho(\boldsymbol{\theta})]) \leq \mathcal{F}(\rho(\boldsymbol{\theta}))$$

The quantum Fisher information is therefore an upper bound for the classical Fisher information matrix arising from *any* measurement.

Roles of Classical and Quantum Fisher Information

$$\mathcal{F}(\boldsymbol{\theta})$$

QUANTUM FISHER INFORMATION

Quantifies ultimate limits for a specific underlying state

Can tell us a lot about the quantum effects influencing our experiments

$$I(\boldsymbol{\theta})$$

CLASSICAL FISHER INFORMATION

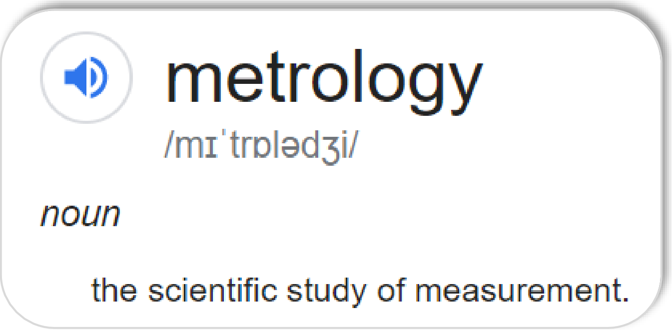
Quantifies behaviour for a fixed measurement

Extremely relevant in practice, as we always have to fix some sort of measurement

Application 1: Quantum Metrology

Physical quantities (magnetic fields, energies, …) need to be **measured** accurately

Study how **quantum effects** can help



Probe State	Encoding Evolution	Measurement	Output Distribution
ρ	$\mathcal{E}(\phi)$	$\mathcal{M}$	$p_l(\phi)$

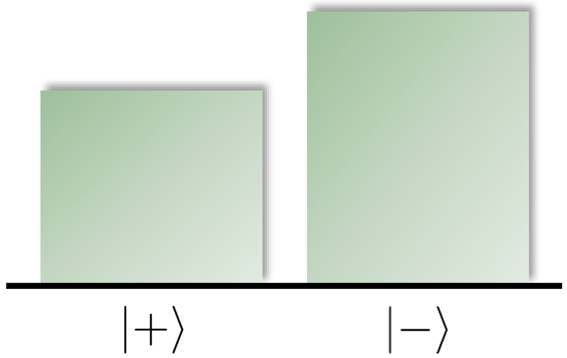
$$\rho \longrightarrow \mathcal{E}(\phi) \longrightarrow \mathcal{M} \longrightarrow p_l(\phi)$$

Gathering Intuition

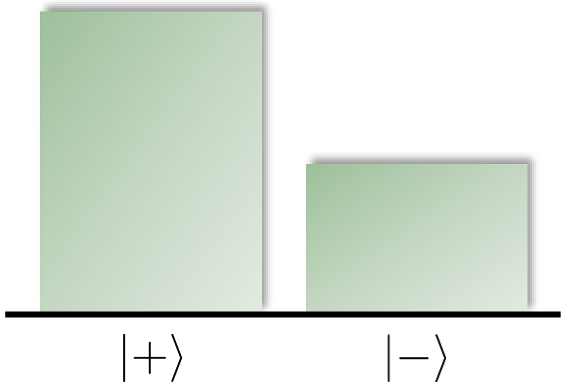
$$|+\rangle \rightarrow e^{-i\phi Z} \rightarrow |+\rangle / |-\rangle$$

THEORY

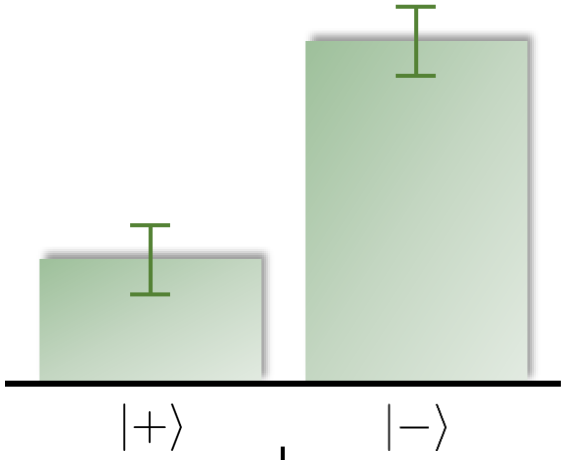
$\phi = 0.9$



$\phi = 0.5$



EXPERIMENT



$$\phi = 1.1 \pm 0.1$$

Cramér-Rao Bound

Formally, we construct an *estimator* for the physical quantity from the output probability distribution

$p(\phi) \mapsto \hat{\varphi}$ unbiased if $\mathbb{E}\{\hat{\varphi}\} = \phi$

The *Cramér-Rao bound* limits the precision of any unbiased estimator

$$\mathrm{Cov}[\hat{\varphi}] \geq \frac{1}{n} I_{\mathcal{M}}(\phi)^{-1} \geq \frac{1}{n} \mathcal{F}(\phi)^{-1}$$

$$\mathrm{Tr}\{\mathrm{Cov}[\hat{\varphi}]\} = \mathrm{MSE}[\hat{\varphi}]$$

Application 2: Quantum Natural Gradient

Gradient descent is a general-purpose method to minimize a cost function

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \eta \nabla C(\boldsymbol{\theta}^{(t)})$$

We can reformulate the gradient update as an optimization problem

$$\boldsymbol{\theta}^{(t+1)} = \underset{\boldsymbol{\vartheta}}{\operatorname{argmin}} \left\{ \boxed{\langle \boldsymbol{\vartheta} - \boldsymbol{\theta}^{(t)}, \nabla C(\boldsymbol{\theta}^{(t)}) \rangle} + \boxed{\frac{1}{2\eta} \| \boldsymbol{\vartheta} - \boldsymbol{\theta}^{(t)} \|_2^2} \right\}$$

Minimal for update opposite to gradient

Regularisation to avoid overstepping

Quantum Natural Gradient Update

But we started this talk realizing that measuring distances between parmeters makes more sense if we take the pullback of a distance between quantum states! Replacing

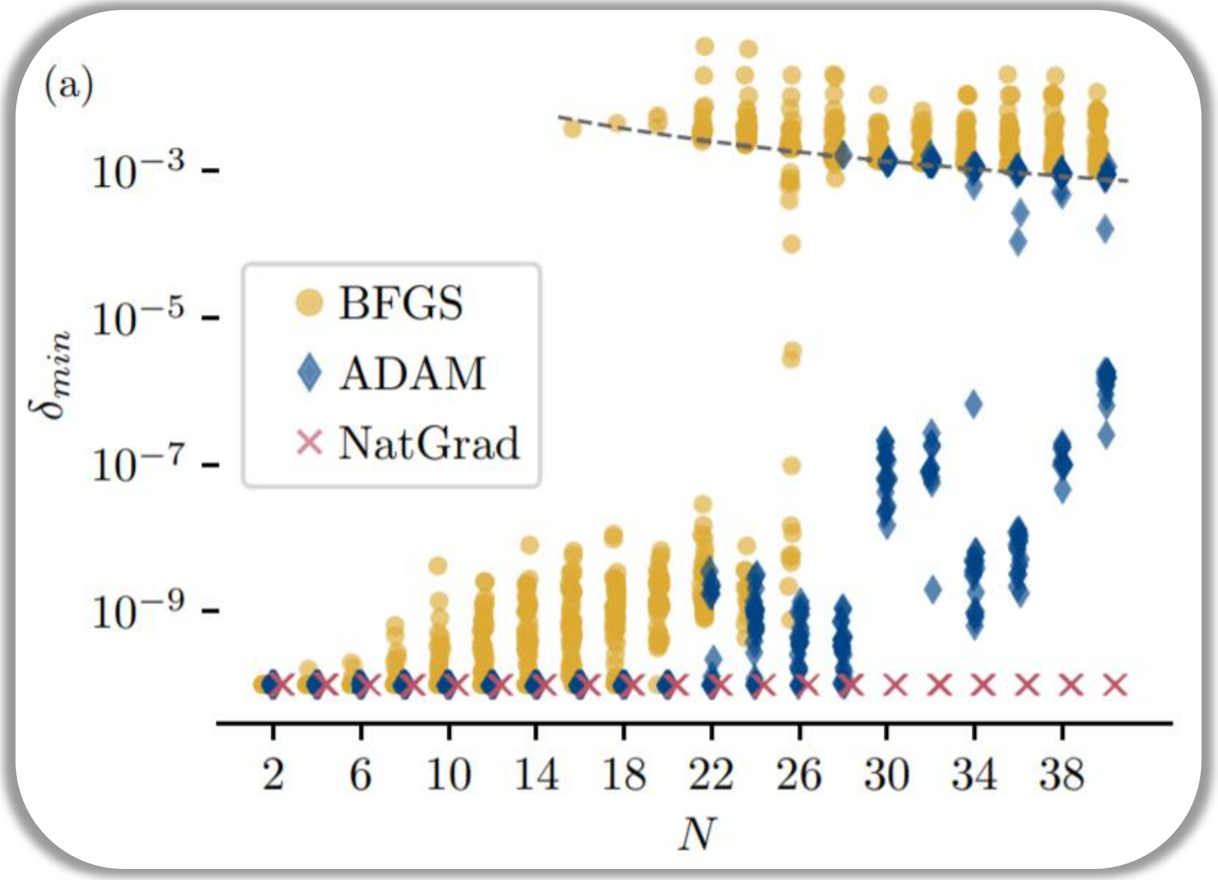
$$\|\boldsymbol{\vartheta} - \boldsymbol{\theta}^{(t)}\|_2^2 \quad \text{with} \quad d_f(\boldsymbol{\vartheta}, \boldsymbol{\theta}^{(t)}) \approx (\boldsymbol{\vartheta} - \boldsymbol{\theta}^{(t)})^T \mathcal{F}(\boldsymbol{\theta}^{(t)})(\boldsymbol{\vartheta} - \boldsymbol{\theta}^{(t)})$$

in the optimization yields the *quantum natural gradient* update step

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \eta \mathcal{F}(\boldsymbol{\theta}^{(t)})^{-1} \nabla C(\boldsymbol{\theta}^{(t)})$$

This update now takes the underlying geometry of the parametrized quantum state into account

Quantum Natural Gradient Helps



David Wierichs, Christian Gogolin, and Michael Kastoryano
Phys. Rev. Research **2**, 043246

Outlook

- Fisher information is a very versatile tool that has found a lot of cool applications in various fields, for example theoretical quantum information, error correction and resource theories

- Parametrized quantum states are fundamental to NISQ applications, so I expect many more interesting applications of Fisher information

- Many open questions remain and much more work is to be done!

Thank you for your attention!

Paper

Slides

www.johannesjakobmeyer.com

@jj_xyz