SANZ GROUP SEMINAR

Variational Methods for Quantum Sensing and Learning

JOHANNES JAKOB MEYER, FU BERLIN & QMATH

@jj_xyz

arxiv:2006.06303

A variational toolbox for quantum multi-parameter estimation

Johannes Jakob Meyer,[1] Johannes Borregaard,[2,3] and Jens Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Qutech and Kavli Institute of Nanoscience, Delft University of Technology, 2628 CJ Delft, The Netherlands*
[3]*Mathematical Sciences, Universitetsparken 5, 2100 København Ø, Matematik E, Denmark*

(Dated: June 11, 2020)

Johannes Borregaard
TU Delft

Jens Eisert
FU Berlin

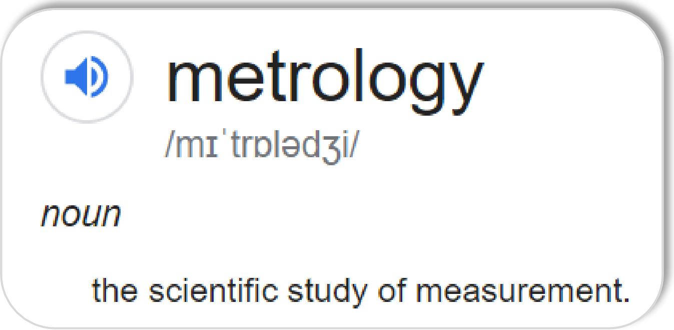
Quantum Metrology

Quantum Metrology

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

Quantum Metrology

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



Quantum Metrology

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

Study how **quantum effects** can help

metrology
/mɪˈtrɒlədʒi/

noun

the scientific study of measurement.

Quantum Metrology

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

Study how **quantum effects** can help

metrology
/mɪˈtrɒlədʒi/
noun
the scientific study of measurement.

Probe
State

ρ

Quantum Metrology

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

Study how **quantum effects** can help

metrology
/mɪˈtrɒlədʒi/
noun
the scientific study of measurement.

Probe State

Encoding Evolution

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

Quantum Metrology

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

Study how **quantum effects** can help

metrology
/mɪˈtrɒlədʒi/
noun
the scientific study of measurement.

Probe State | Encoding Evolution | Measurement

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

Quantum Metrology

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

Study how **quantum effects** can help

metrology
/mɪˈtrɒlədʒi/

noun

the scientific study of measurement.

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

Gathering Intuition

Gathering Intuition

$$e^{-i\phi Z}$$

Gathering Intuition

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

Gathering Intuition

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

Gathering Intuition

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

THEORY

Gathering Intuition

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

THEORY

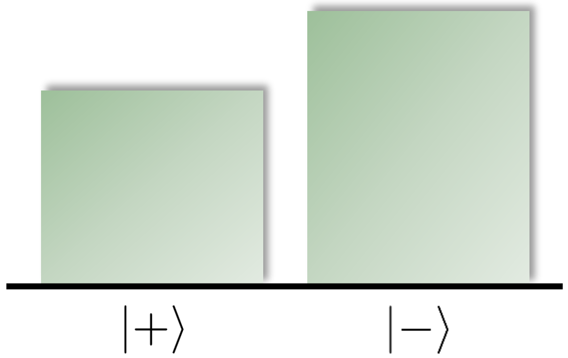
$$\phi = 0.9$$

Gathering Intuition

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

THEORY

$$\phi = 0.9$$

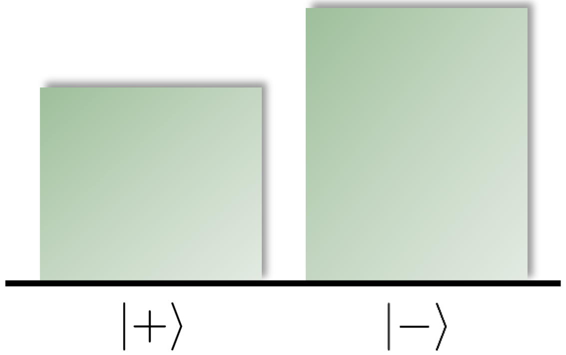


Gathering Intuition

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

THEORY

$\phi = 0.9$



$\phi = 0.5$

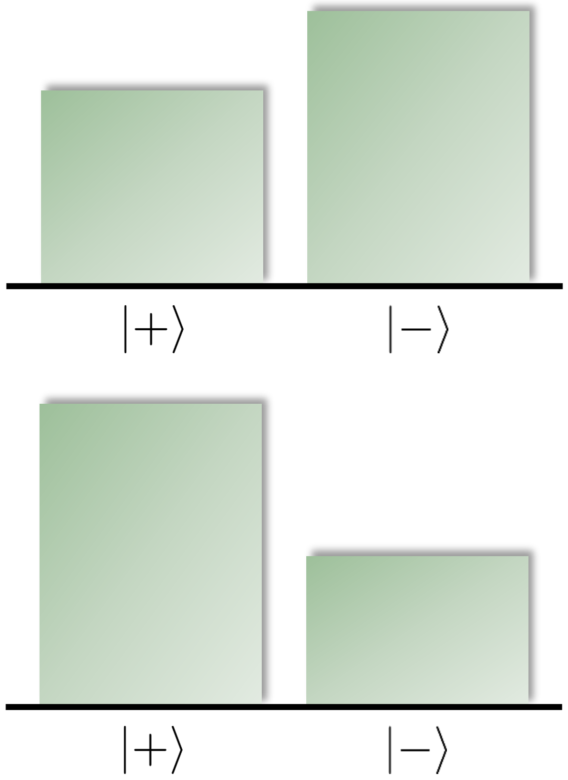
Gathering Intuition

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

THEORY

$\phi = 0.9$

$\phi = 0.5$

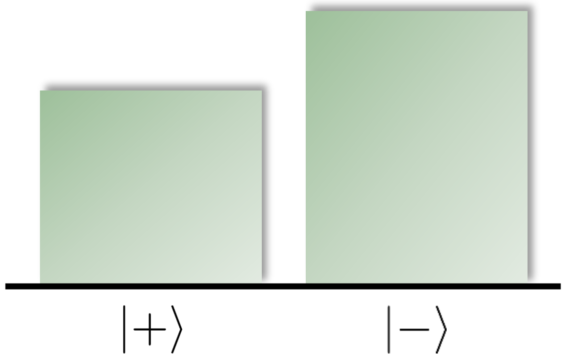


Gathering Intuition

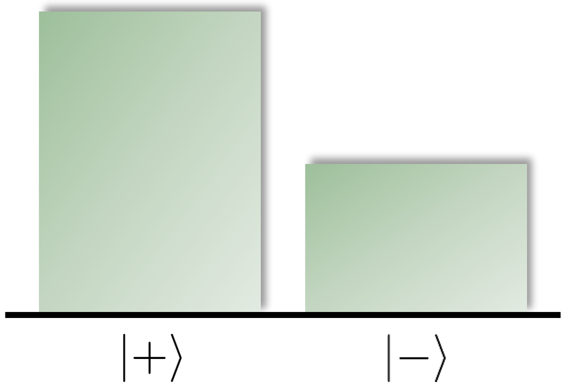
$$|+\rangle \rightarrow e^{-i\phi Z} \rightarrow |+\rangle/|-\rangle$$

THEORY

$$\phi = 0.9$$



$$\phi = 0.5$$



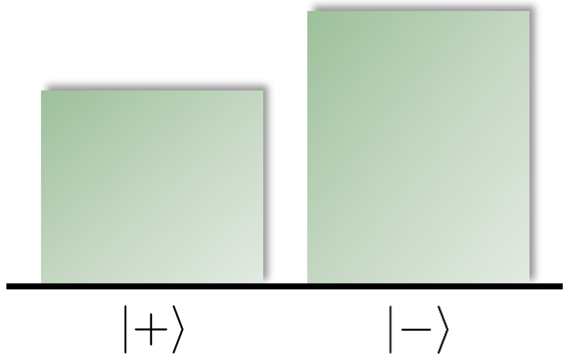
EXPERIMENT

Gathering Intuition

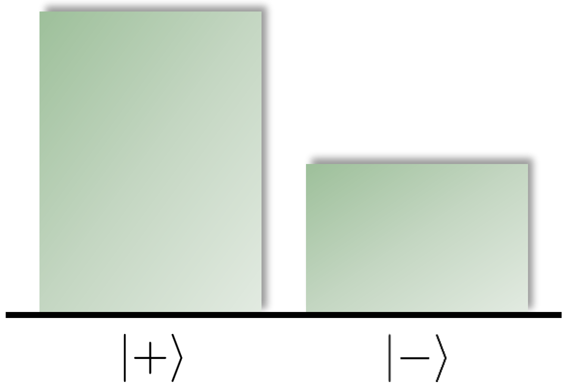
$$|+\rangle \rightarrow e^{-i\phi Z} \rightarrow |+\rangle/|-\rangle$$

THEORY

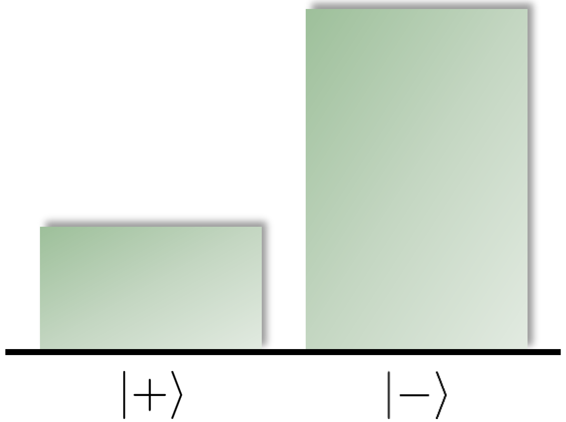
$\phi = 0.9$



$\phi = 0.5$



EXPERIMENT

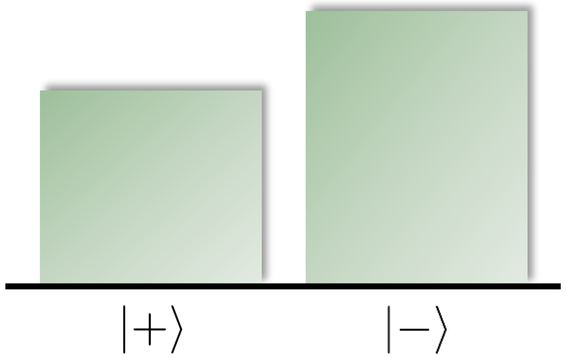


Gathering Intuition

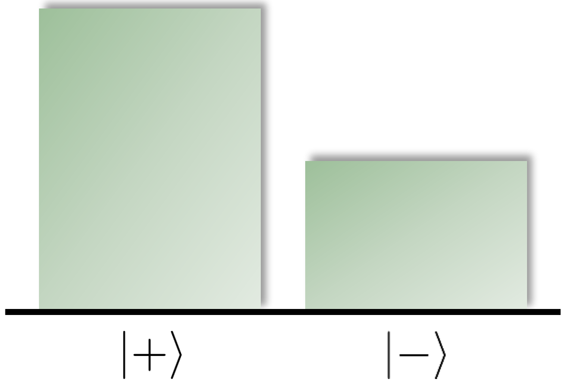
$$|+\rangle \rightarrow e^{-i\phi Z} \rightarrow |+\rangle/|-\rangle$$

THEORY

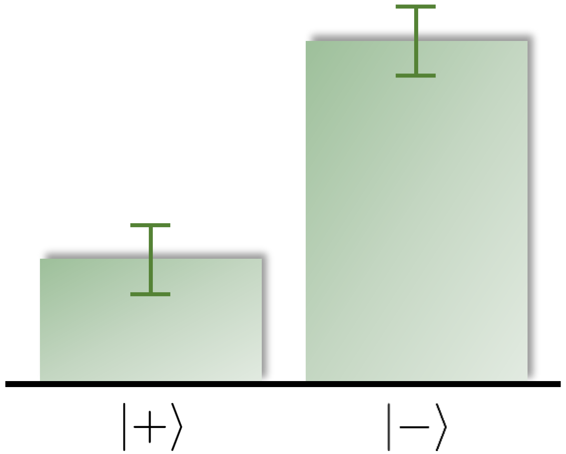
$\phi = 0.9$



$\phi = 0.5$

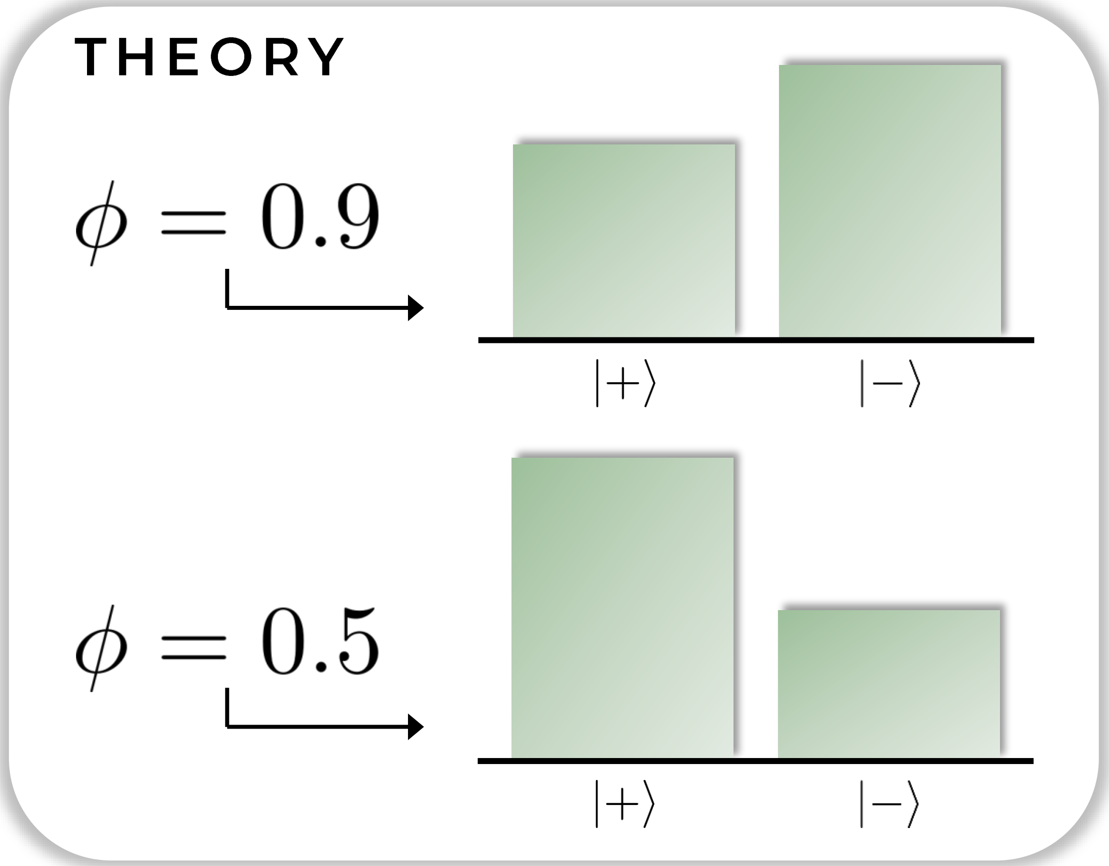


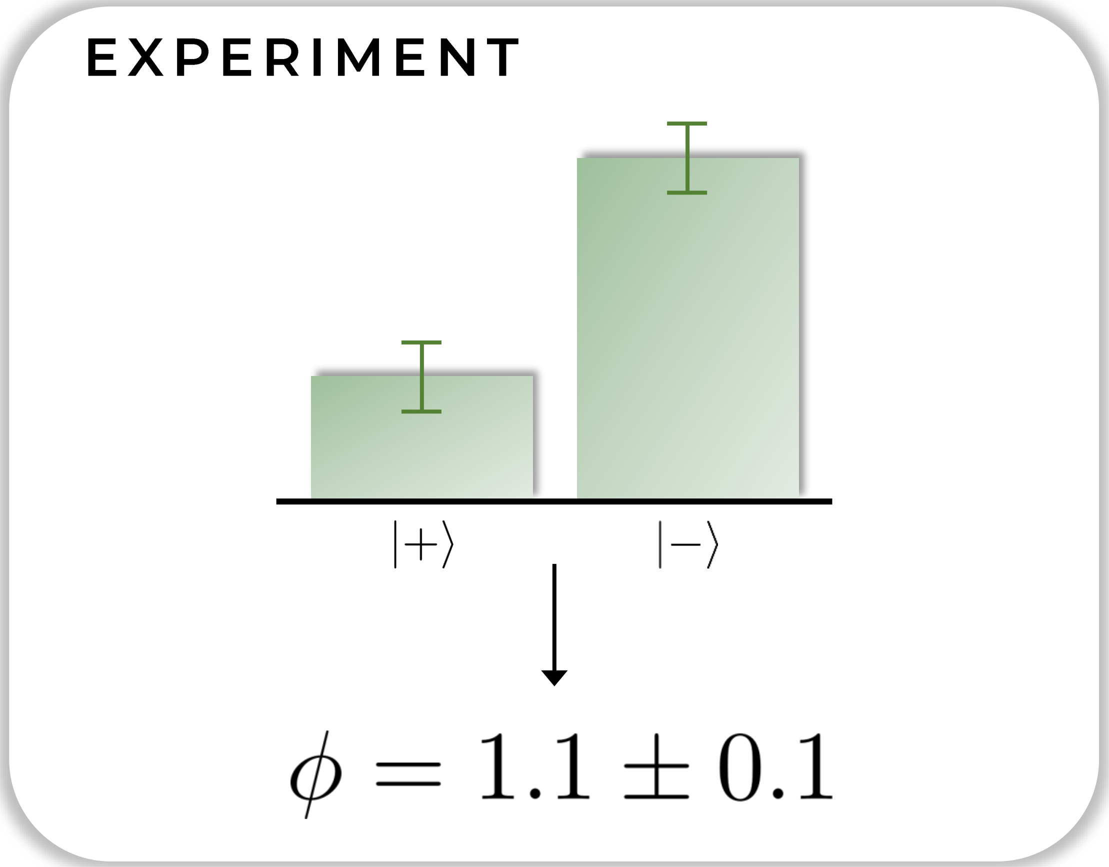
EXPERIMENT



Gathering Intuition

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





Measuring Performance

Measuring Performance

Task: Compute an **estimator** from the output distribution

$$p_l(\boldsymbol{\phi}) \;\rightarrow\; \hat{\boldsymbol{\varphi}} \colon \mathbb{E}\{\hat{\boldsymbol{\varphi}}\} = \boldsymbol{\phi}$$

Measuring Performance

Task: Compute an **estimator** from the output distribution

$$p_l(\phi) \;\rightarrow\; \hat{\boldsymbol{\varphi}}\colon \mathbb{E}\{\hat{\boldsymbol{\varphi}}\} = \boldsymbol{\phi}$$

Precision from n samples is limited by **Cramér-Rao bound**

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

Measuring Performance

Task: Compute an **estimator** from the output distribution

$$p_l(\phi) \;\rightarrow\; \hat{\varphi} \colon \mathbb{E}\{\hat{\varphi}\} = \phi$$

Precision from n samples is limited by **Cramér-Rao bound**

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

$$\boxed{\mathrm{Tr}\{\mathrm{Cov}(\hat{\varphi})\} = \mathrm{MSE}(\hat{\varphi})}$$

Measuring Performance

Task: Compute an **estimator** from the output distribution

$$p_l(\phi) \;\to\; \hat{\boldsymbol{\varphi}}\colon \mathbb{E}\{\hat{\boldsymbol{\varphi}}\} = \phi$$

Precision from n samples is limited by **Cramér-Rao bound**

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

Measuring Performance

Fisher Information in Noisy Intermediate-Scale Quantum Applications

Johannes Jakob Meyer[1,2]

[1]Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
[2]QMATH, Department of Mathematical Sciences, Københavns Universitet, 2100 København Ø, Denmark
28-03-2021

arXiv:2103.15191

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

Recap #1

Recap #1

Sensing amounts to estimating the underlying physical parameters from a classical probability distribution

Recap #1

Sensing amounts to estimating the underlying physical parameters from a classical probability distribution

Estimation precision is limited by the Cramér-Rao bound

Recap #1

Sensing amounts to estimating the underlying physical parameters from a classical probability distribution

Estimation precision is limited by the Cramér-Rao bound

Attainble precision is quantified by the classical Fisher information

Recap #1

Sensing amounts to estimating the underlying physical parameters from a classical probability distribution

Estimation precision is limited by the Cramér-Rao bound

Attainble precision is quantified by the classical Fisher information

The quantum Fisher information bounds the attainable Fisher information

Recap #1

Sensing amounts to estimating the underlying physical parameters from a classical probability distribution

Estimation precision is limited by the Cramér-Rao bound

Attainble precision is quantified by the classical Fisher information

The quantum Fisher information bounds the attainable Fisher information

→ Classical Fisher information should be used to judge sensing quality!

Optimal Metrology

Optimal Metrology

We need to find optimal probes and measurements

Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

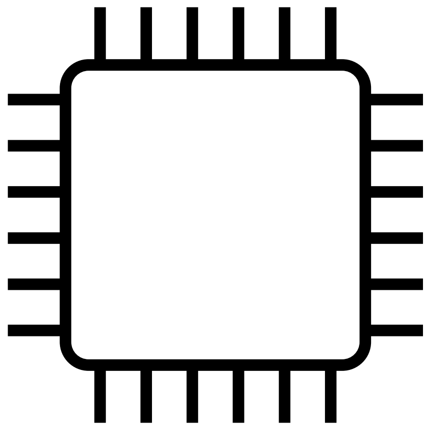
NISQ techniques come to the rescue:
use variational approaches

Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

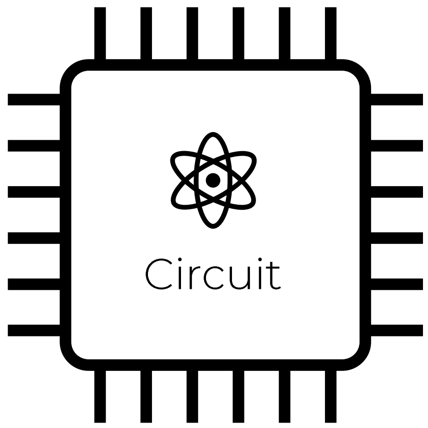


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

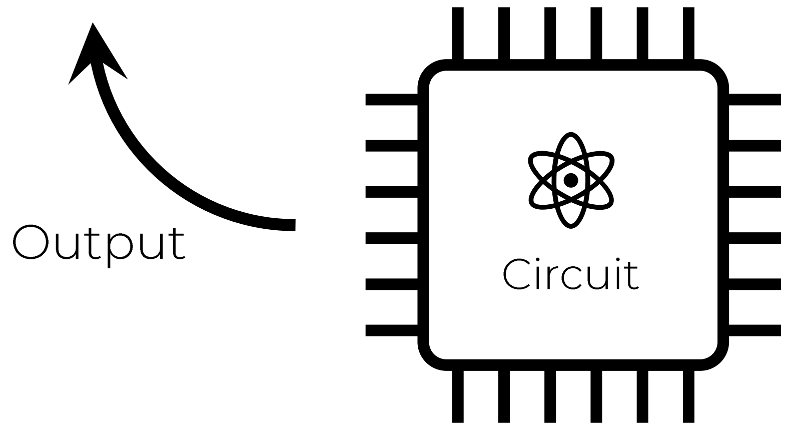


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

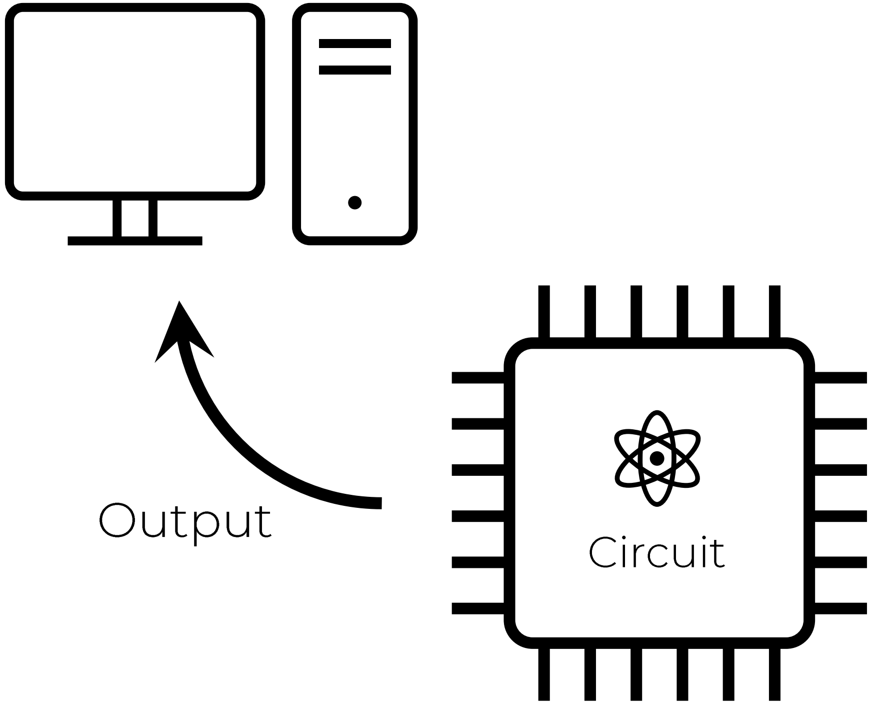


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

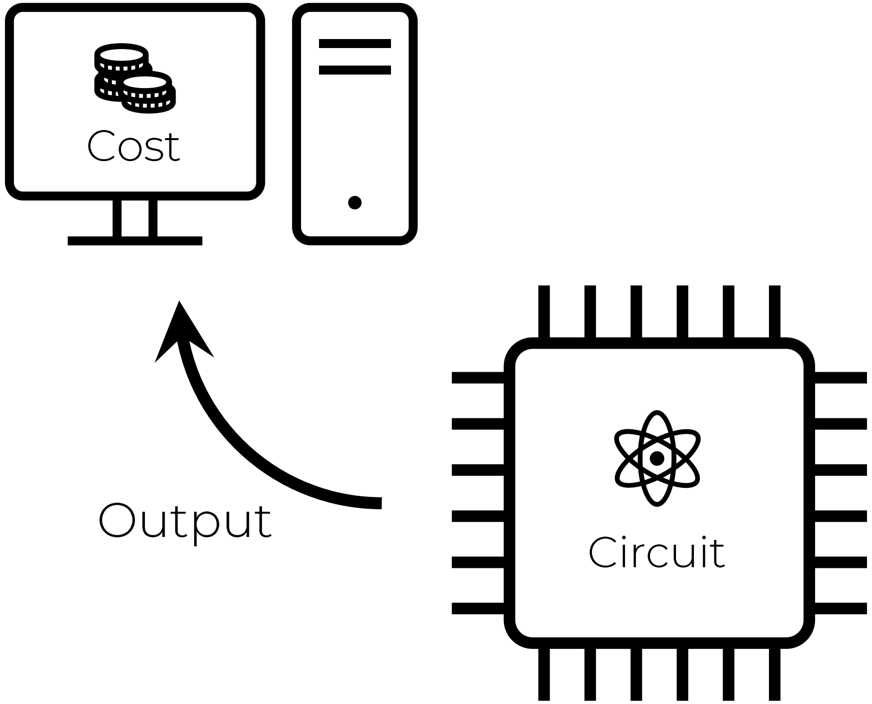


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

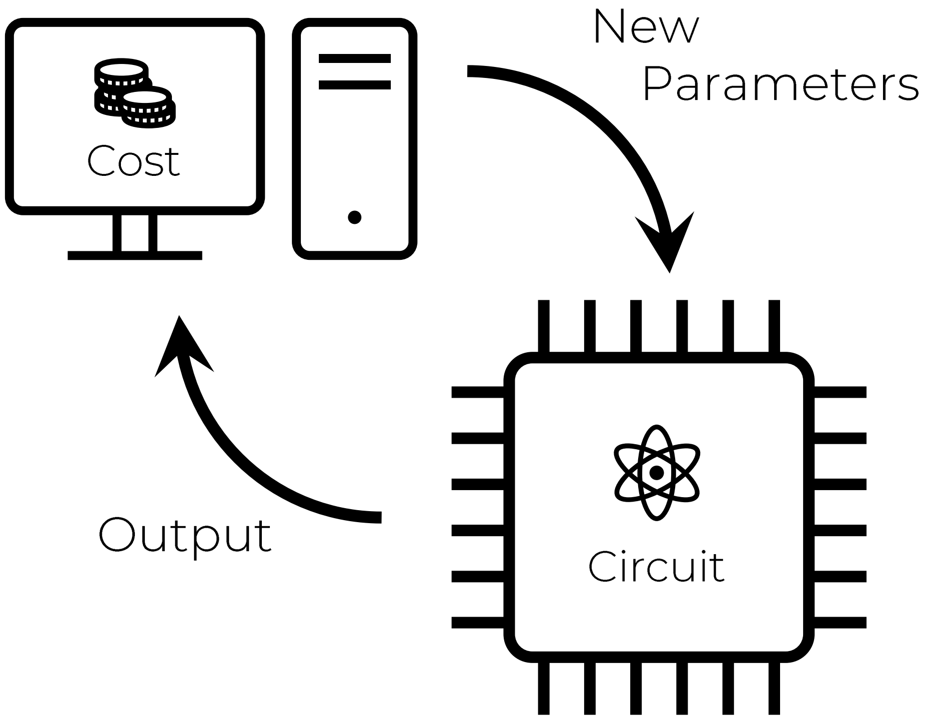


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches

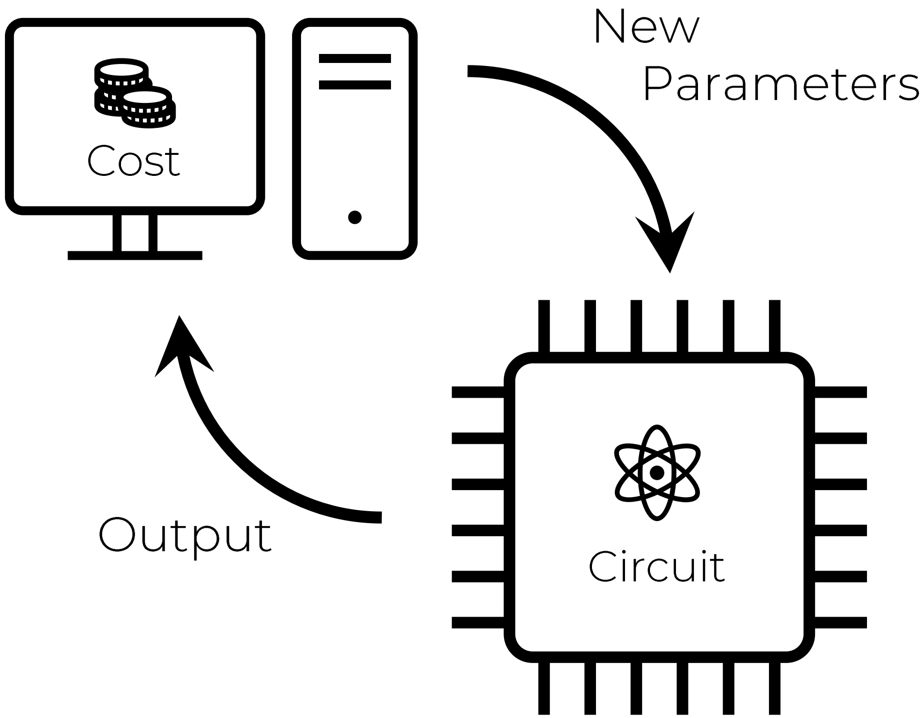


Optimal Metrology

We need to find optimal probes and measurements

Complicated under noise and device limitations

NISQ techniques come to the rescue:
use variational approaches



Prior work[1,2] focused on single-parameter metrology and surrogates for the Quantum Fisher Information

[1]Kaubrügger, Raphael, et al. *Physical Review Letters* 123.26 (2019): 260505.
[2]Koczor, Bálint, et al. "Variational-State Quantum Metrology." *New Journal of Physics* (2020).

Cost Function

Cost Function

Include post-processing by considering a function of the parameters:
Exploit transformation rule of Fisher Information Matrix

Cost Function

Include post-processing by considering a function of the parameters:
Exploit transformation rule of Fisher Information Matrix

$$\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\phi}) \;\rightarrow\; I_{\boldsymbol{f}} = J^T I_{\boldsymbol{\phi}} J$$

Cost Function

Include post-processing by considering a function of the parameters:
Exploit transformation rule of Fisher Information Matrix

$$\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\phi}) \;\rightarrow\; I_{\boldsymbol{f}} = J^T I_{\boldsymbol{\phi}} J$$

Need a scalar cost function:
Apply weighted trace to both sides of the CRB!

Cost Function

Include post-processing by considering a function of the parameters:
Exploit transformation rule of Fisher Information Matrix

$$\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\phi}) \;\rightarrow\; I_{\boldsymbol{f}} = J^T I_{\boldsymbol{\phi}} J$$

Need a scalar cost function:
Apply weighted trace to both sides of the CRB!

$$\mathrm{Tr}\{W\,\mathrm{Cov}(\hat{\boldsymbol{f}})\} \geq \frac{1}{n}\,\mathrm{Tr}\{W I_{\boldsymbol{f}}^{-1}\} = \frac{1}{n} C_W$$

Calculation of Fisher Information

Calculation of Fisher Information

Fisher Information Matrix w.r.t. the physical parameters

$$[I_{\boldsymbol{\phi}}]_{jk} = \sum_l \frac{1}{p_l} \frac{\partial p_l}{\partial \phi_j} \frac{\partial p_l}{\partial \phi_k}$$

Calculation of Fisher Information

Fisher Information Matrix w.r.t. the physical parameters

$$[I_{\boldsymbol{\phi}}]_{jk} = \sum_l \frac{1}{p_l} \frac{\partial p_l}{\partial \phi_j} \frac{\partial p_l}{\partial \phi_k}$$

Exploit parameter-shift rule[1,2] to calculate derivatives

$$\partial_j p_l(\boldsymbol{\phi}) = \frac{1}{2} \left[p_l \left(\boldsymbol{\phi} + \frac{\pi}{2} \boldsymbol{e}_j \right) - p_l \left(\boldsymbol{\phi} - \frac{\pi}{2} \boldsymbol{e}_j \right) \right]$$

[1] Schuld, Maria, et al. *Physical Review A* 99.3 (2019): 032331.
[2] Banchi, Leonardo, and Gavin E. Crooks. arXiv preprint arXiv:2005.10299 (2020).

Recap #2

Recap #2

The classical Fisher information matrix is calculated from the output probabilities and their derivatives w.r.t. the physical parameters

Recap #2

The classical Fisher information matrix is calculated from the output probabilities and their derivatives w.r.t. the physical parameters

The derivatives can be calculated on the device via the parameter-shift rule

Recap #2

The classical Fisher information matrix is calculated from the output probabilities and their derivatives w.r.t. the physical parameters

The derivatives can be calculated on the device via the parameter-shift rule

The classical Fisher information matrix w.r.t. post-processed parameters can be computed using the post-processing‘s Jacobian

Recap #2

The classical Fisher information matrix is calculated from the output probabilities and their derivatives w.r.t. the physical parameters

The derivatives can be calculated on the device via the parameter-shift rule

The classical Fisher information matrix w.r.t. post-processed parameters can be computed using the post-processing‘s Jacobian

The cost function is obtained from a weighted trace of the Cramér-Rao bound

Have a Look at the Paper for

Have a Look at the Paper for

Multiple extensions of the algorithm, for example including prior knowlege (Bayesian approach)

Have a Look at the Paper for

Multiple extensions of the algorithm, for example including prior knowlege (Bayesian approach)

A parameter-shift rule for noise channels

Have a Look at the Paper for

Multiple extensions of the algorithm, for example including prior knowlege (Bayesian approach)

A parameter-shift rule for noise channels

Details on the implementation of parameter-shift rules in experiments

Have a Look at the Paper for

Multiple extensions of the algorithm, for example including prior knowlege (Bayesian approach)

A parameter-shift rule for noise channels

Details on the implementation of parameter-shift rules in experiments

Numerical experiments that showcase the performance of the approach

The Algorithm Landscape

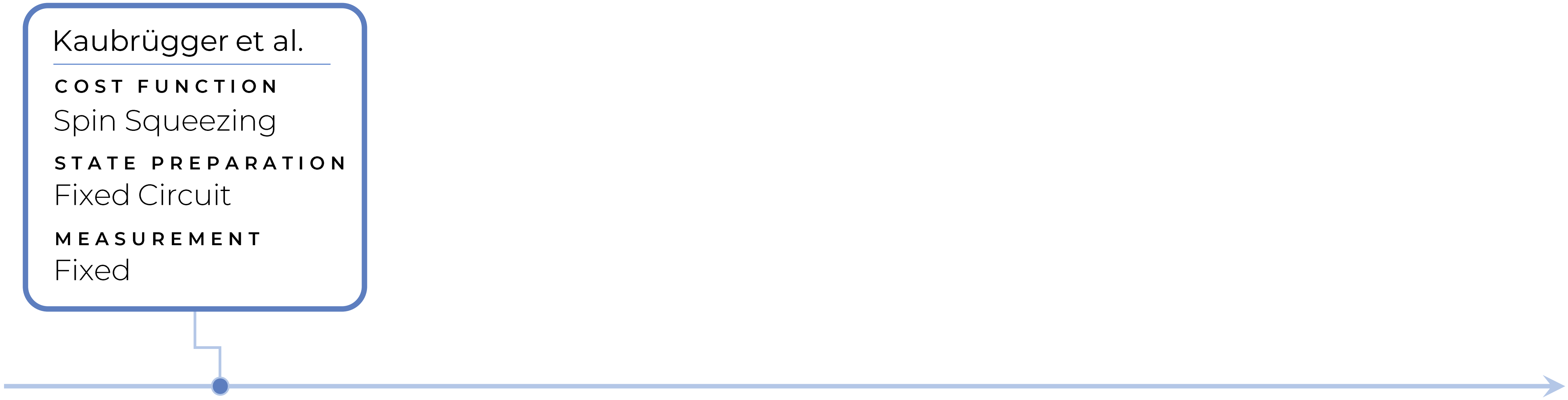
The Algorithm Landscape

Single Parameter

Multiparameter

The Algorithm Landscape

Single Parameter
Multiparameter



The Algorithm Landscape

Single Parameter   Multiparameter

Kaubrügger et al.

COST FUNCTION
Spin Squeezing

STATE PREPARATION
Fixed Circuit

MEASUREMENT
Fixed

Koczor et al.

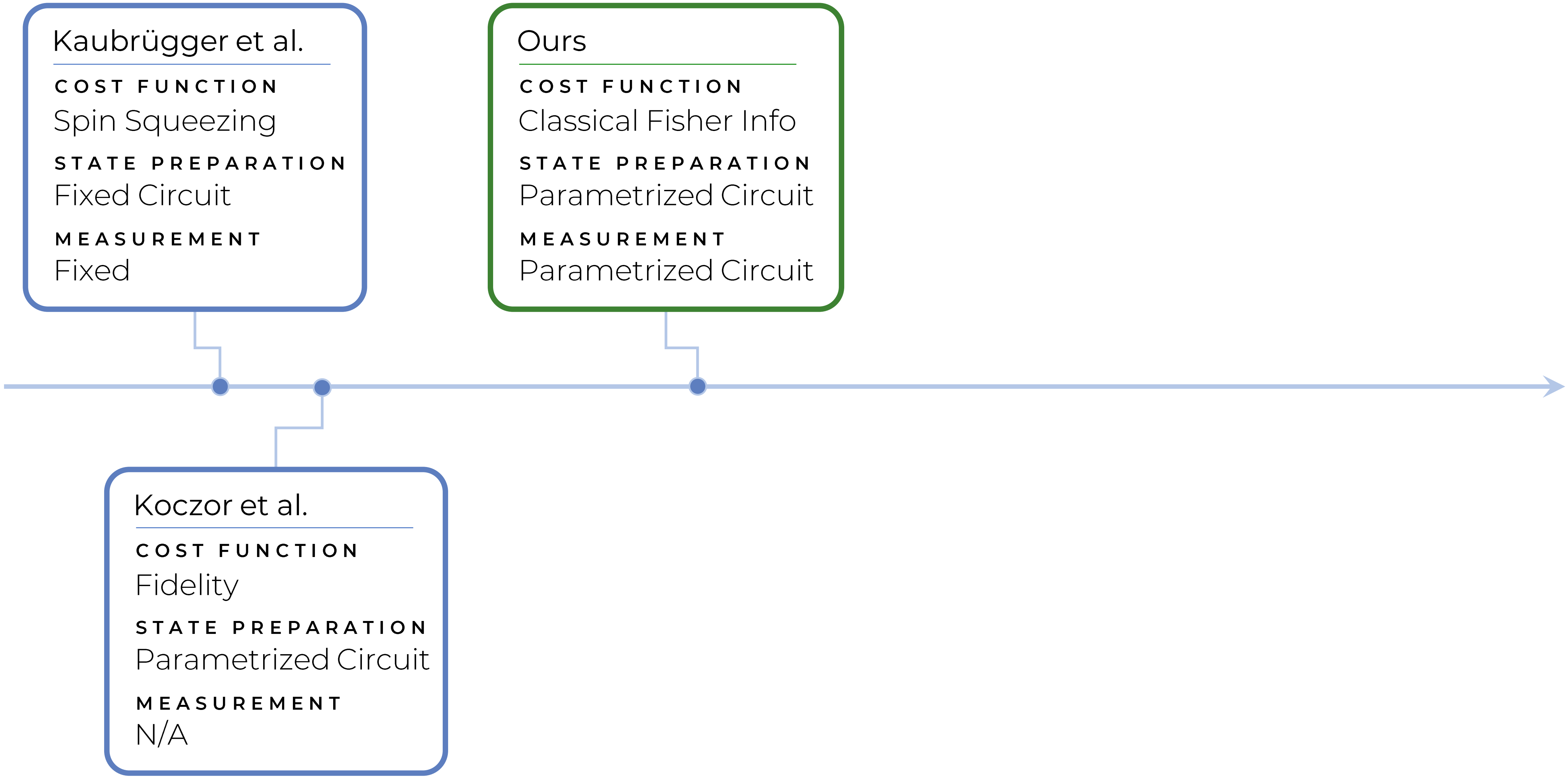
COST FUNCTION
Fidelity

STATE PREPARATION
Parametrized Circuit

MEASUREMENT
N/A

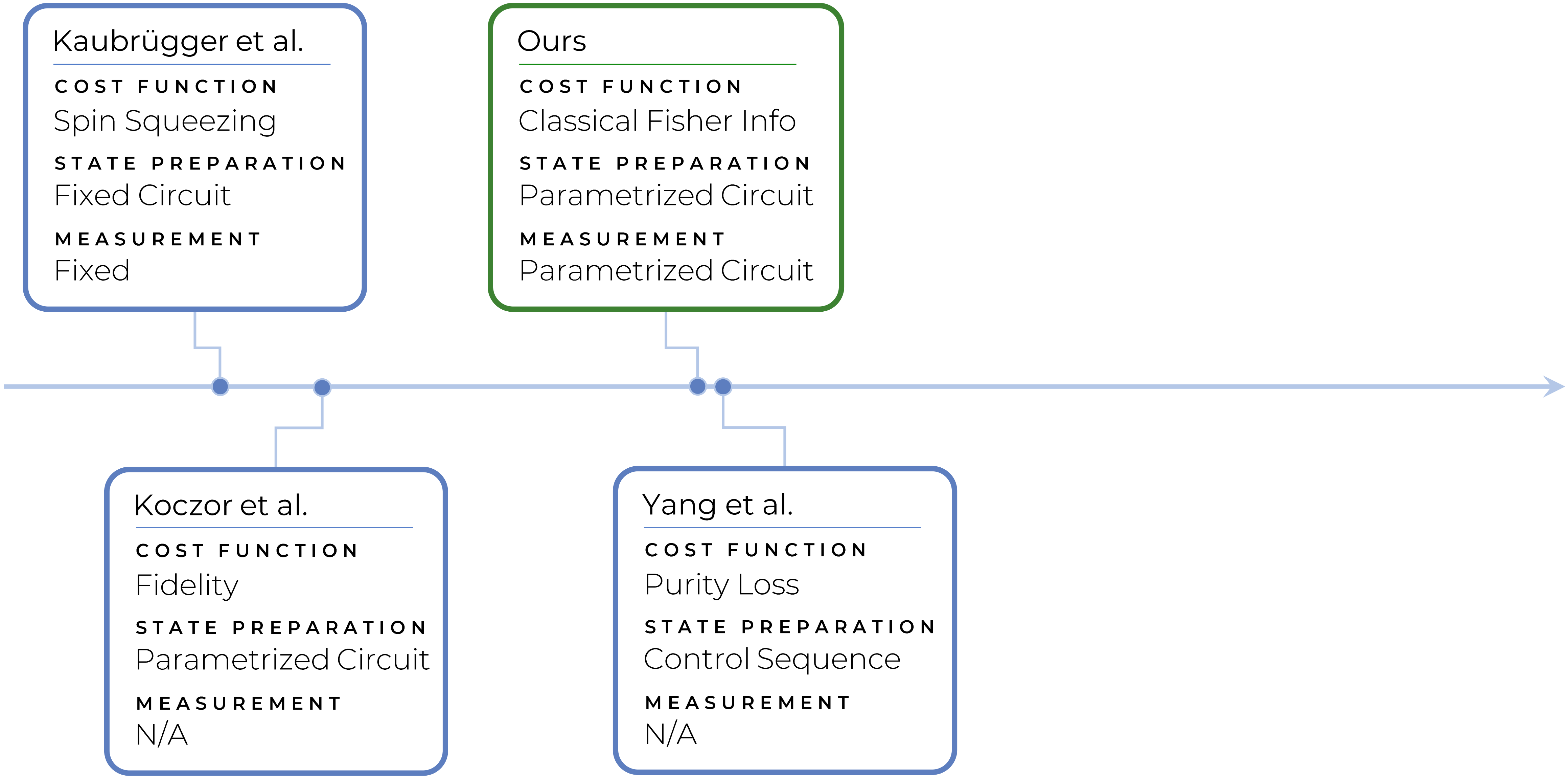
The Algorithm Landscape

Single Parameter Multiparameter



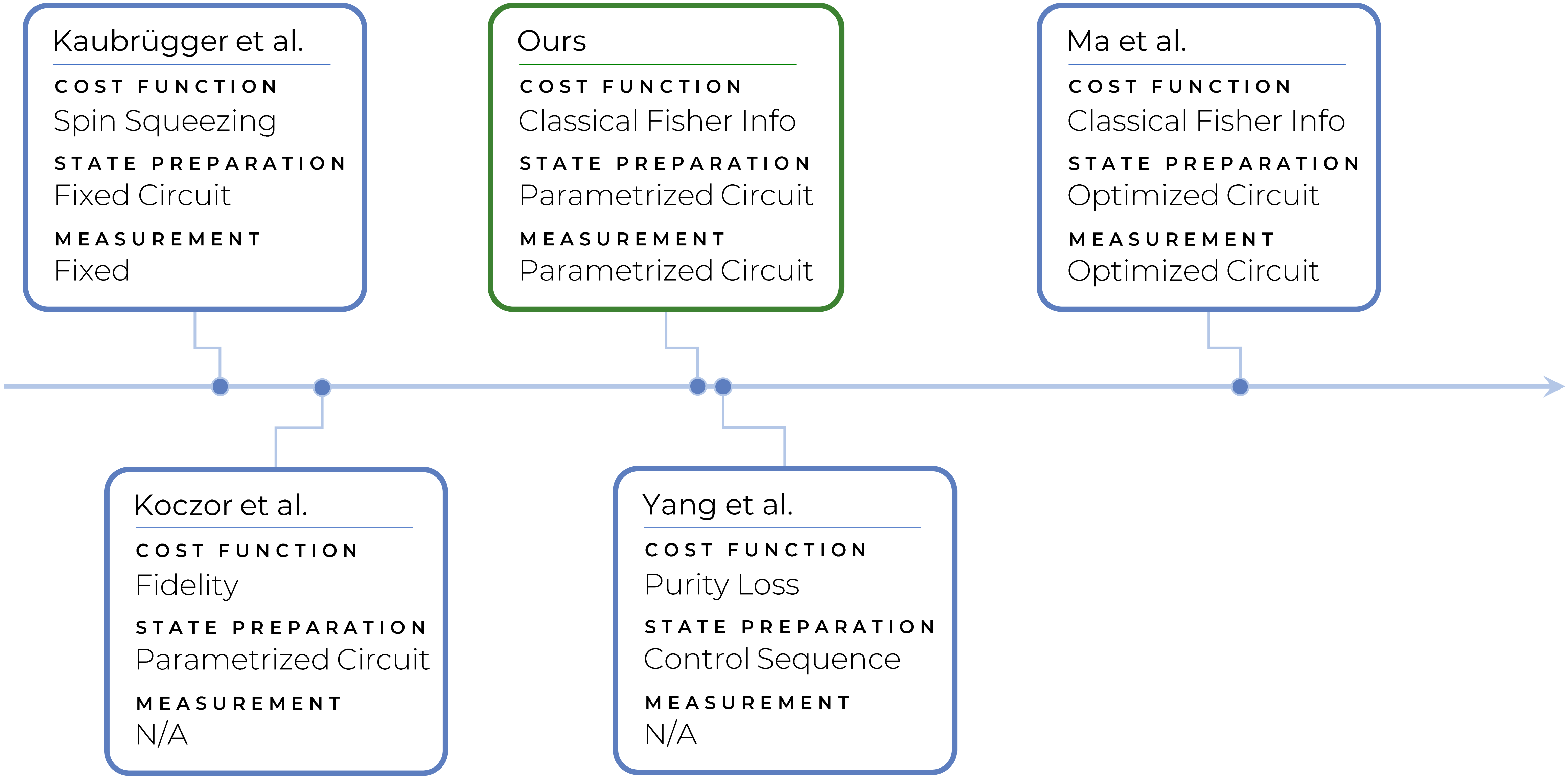
The Algorithm Landscape

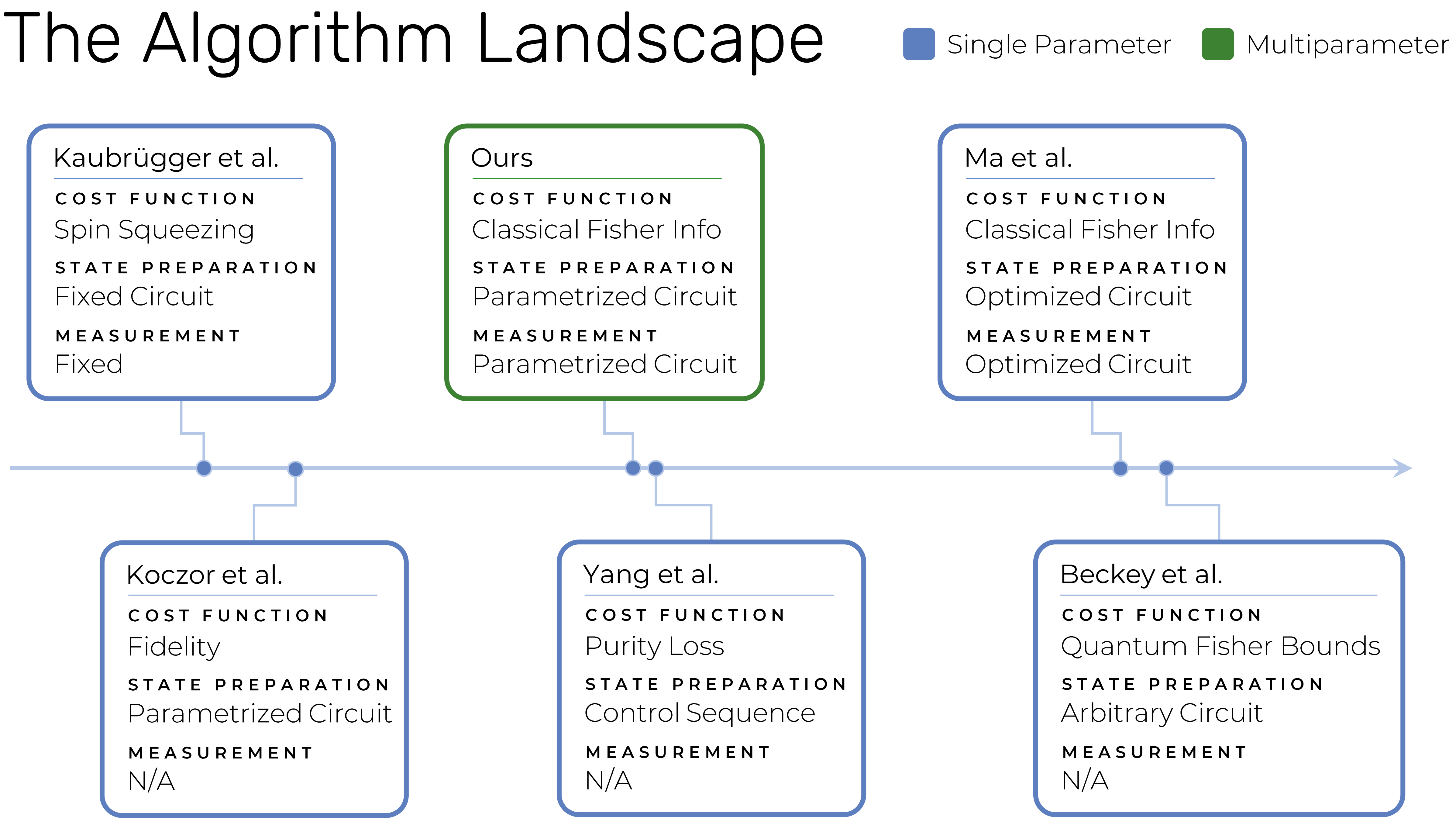
Single Parameter Multiparameter



The Algorithm Landscape

Single Parameter Multiparameter




The Algorithm Landscape
Single Parameter
Multiparameter
Kaubrügger et al.
COST FUNCTION
Spin Squeezing
STATE PREPARATION
Fixed Circuit
MEASUREMENT
Fixed
Ours
COST FUNCTION
Classical Fisher Info
STATE PREPARATION
Parametrized Circuit
MEASUREMENT
Parametrized Circuit
Ma et al.
COST FUNCTION
Classical Fisher Info
STATE PREPARATION
Optimized Circuit
MEASUREMENT
Optimized Circuit
Koczor et al.
COST FUNCTION
Fidelity
STATE PREPARATION
Parametrized Circuit
MEASUREMENT
N/A
Yang et al.
COST FUNCTION
Purity Loss
STATE PREPARATION
Control Sequence
MEASUREMENT
N/A
Beckey et al.
COST FUNCTION
Quantum Fisher Bounds
STATE PREPARATION
Arbitrary Circuit
MEASUREMENT
N/A

Take-Home Message #1

> Variational methods on near-term quantum computers can be used to improve quantum sensors

Paper

Demo

Slides

Fisher Note

10.1103/PhysRevA.103.032430

10.1103/PhysRevA.103.032430

The effect of data encoding on the expressive power of variational quantum machine learning models

Maria Schuld,[1] Ryan Sweke,[2] and Johannes Jakob Meyer[2]

[1] *Xanadu, Toronto, ON, M5G 2C8, Canada*
[2] *Dahlem Center for Complex Quantum Systems, Freie Universitt Berlin, 14195 Berlin, Germany*

(Dated: August 21, 2020)

10.1103/PhysRevA.103.032430

The effect of data encoding on the expressive power of variational quantum machine learning models

Maria Schuld,[1] Ryan Sweke,[2] and Johannes Jakob Meyer[2]

[1] *Xanadu, Toronto, ON, M5G 2C8, Canada*

[2] *Dahlem Center for Complex Quantum Systems, Freie Universitt Berlin, 14195 Berlin, Germany*

(Dated: August 21, 2020)

Maria Schuld
Xanadu

Ryan Sweke
FU Berlin

10.1103/PhysRevA.103.032430

Disclaimer: After uploading to the arxiv we were notified of work by Vidal and Theis[1] that has significant overlap with ours.

[1]Gil Vidal, Francisco Javier, and Dirk Oliver Theis. "Input Redundancy for Parameterized Quantum Circuits." *Frontiers in Physics* 8 (2020): 297.

Maria Schuld
Xanadu

Ryan Sweke
FU Berlin

Key Motivation

Key Motivation

Using quantum models for learning tasks is one of the key fields where NISQ devices are hoped to bring forth a quantum advantage

Key Motivation

Using quantum models for learning tasks is one of the key fields where NISQ devices are hoped to bring forth a quantum advantage

A lot of work has been done to understand the practical side, but very little is known on the theory side

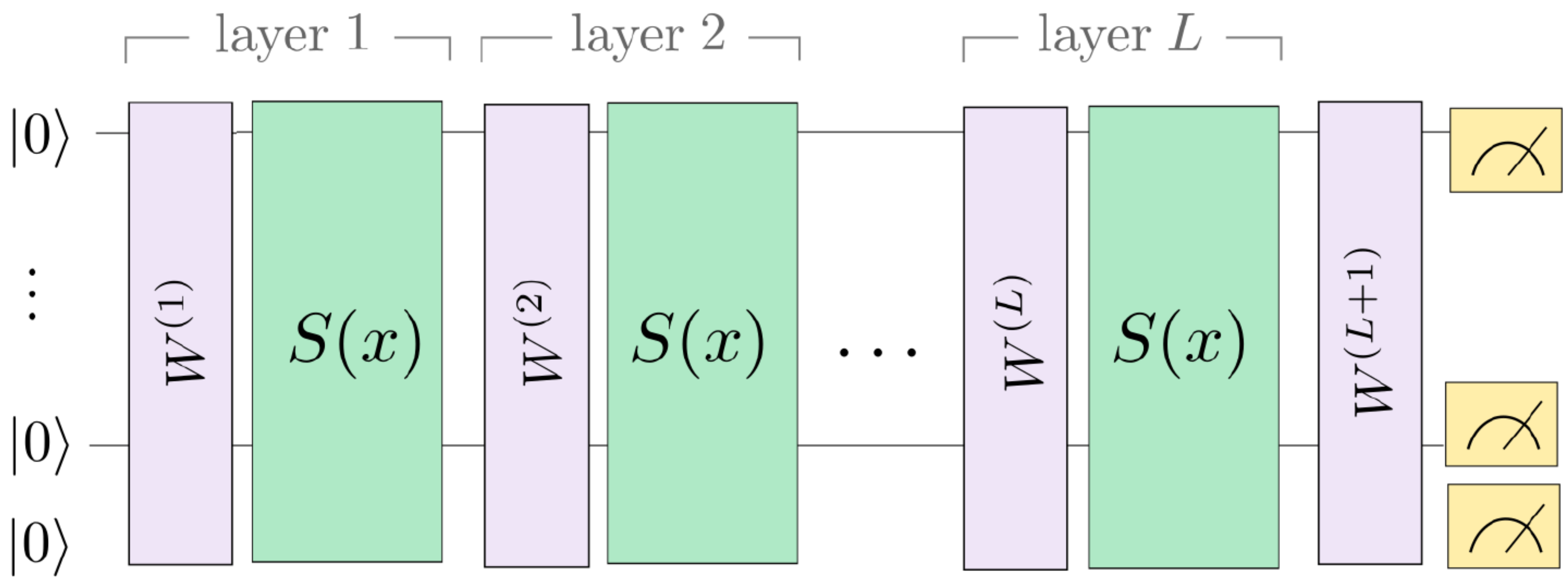
Key Motivation

Using quantum models for learning tasks is one of the key fields where NISQ devices are hoped to bring forth a quantum advantage

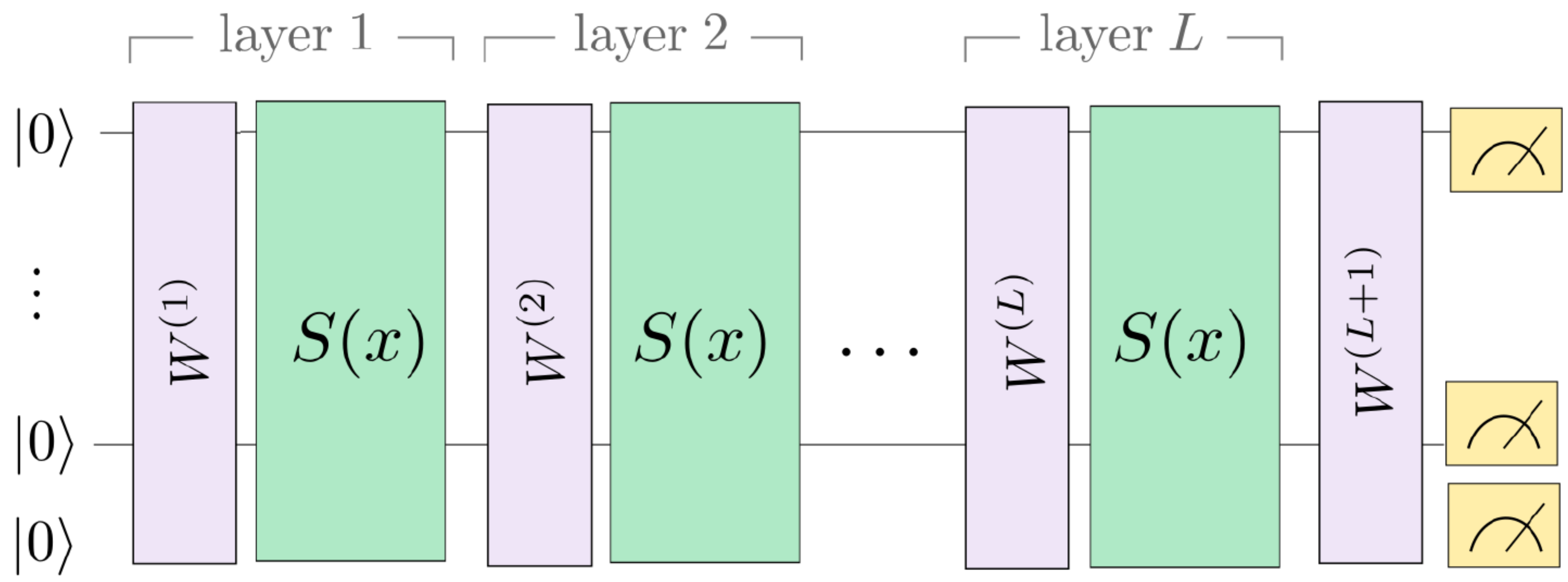
A lot of work has been done to understand the practical side, but very little is known on the theory side

So we asked ourselves: **What functions can such models learn?**

Variational Quantum Learning Models



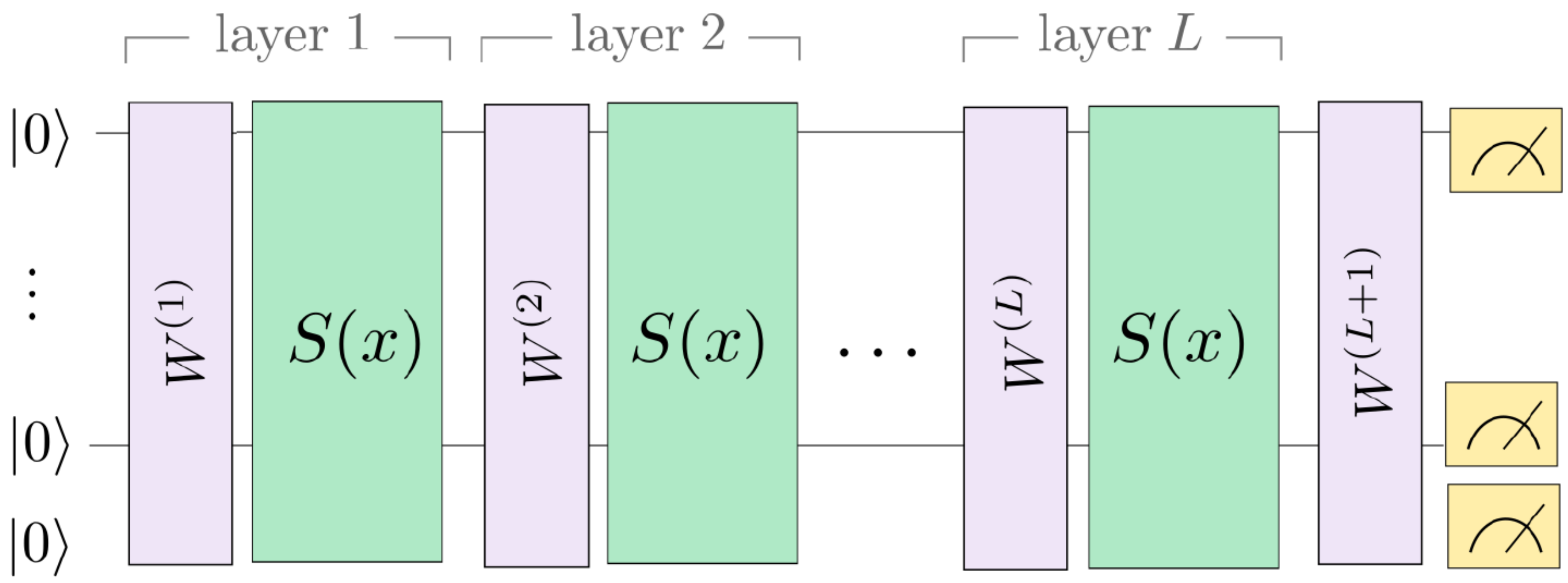
Variational Quantum Learning Models



$$f(x) = \langle M \rangle$$

Model Output

Variational Quantum Learning Models



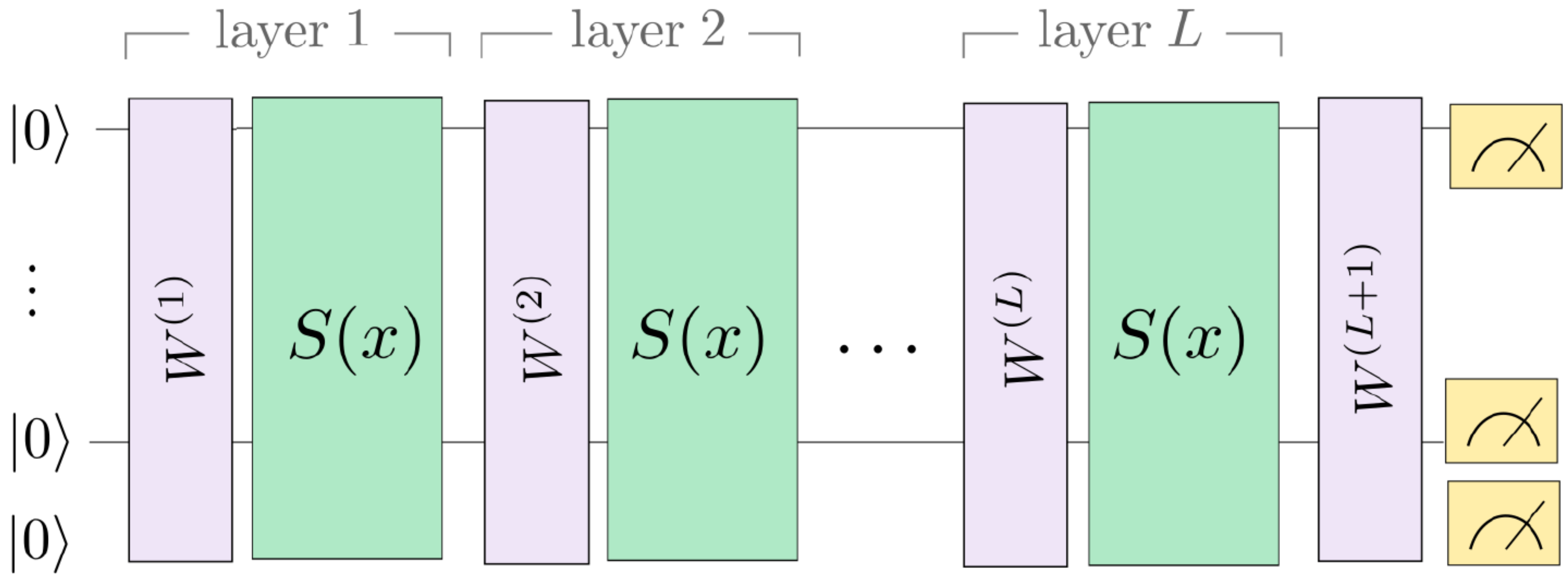
$$f(x) = \langle M \rangle$$

Model Output

$$S(x) = e^{-ixH}$$

Hamiltonian Evolution

Variational Quantum Learning Models



$$f(x) = \langle M \rangle$$

Model Output

$$S(x) = e^{-ixH}$$

Hamiltonian Evolution

$$W^{(l)}(\boldsymbol{\theta})$$

Trainable Blocks

Output State of the Quantum Model

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$

$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$

$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$

$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

$$|\psi\rangle = \sum_{\lambda} \psi_{\lambda}|\lambda\rangle$$

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$
$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

$$|\psi\rangle = \sum_{\lambda} \psi_{\lambda}|\lambda\rangle$$
$$S(x)|\psi\rangle = \sum_{\lambda} \psi_{\lambda} e^{-ix\lambda}|\lambda\rangle$$

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$
$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

$$|\psi\rangle = \sum_{\lambda} \psi_\lambda|\lambda\rangle$$
$$S(x)|\psi\rangle = \sum_{\lambda} \psi_\lambda e^{-ix\lambda}|\lambda\rangle$$
$$WS(x)|\psi\rangle = \sum_{\lambda'\lambda} W_{\lambda'\lambda}\psi_\lambda e^{-ix\lambda}|\lambda'\rangle$$

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$

$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

$$|\psi\rangle = \sum_{\lambda} \psi_{\lambda}|\lambda\rangle$$

$$S(x)|\psi\rangle = \sum_{\lambda} \psi_{\lambda} e^{-ix\lambda}|\lambda\rangle$$

$$WS(x)|\psi\rangle = \sum_{\lambda'\lambda} W_{\lambda'\lambda}\psi_{\lambda} e^{-ix\lambda}|\lambda'\rangle$$

$$S(x)WS(x)|\psi\rangle = \sum_{\lambda'\lambda} W_{\lambda'\lambda}\psi_{\lambda} e^{-ix(\lambda+\lambda')}|\lambda'\rangle$$

Output State of the Quantum Model

The eigenvalues of the generator determine the frequencies

$$H|\lambda\rangle = \lambda|\lambda\rangle$$

$$S(x)|\lambda\rangle = e^{-ix\lambda}|\lambda\rangle$$

and the frequencies accumulate between layers:

$$|\psi\rangle = \sum_{\lambda} \psi_{\lambda}|\lambda\rangle$$

$$S(x)|\psi\rangle = \sum_{\lambda} \psi_{\lambda} e^{-ix\lambda}|\lambda\rangle$$

$$WS(x)|\psi\rangle = \sum_{\lambda'\lambda} W_{\lambda'\lambda}\psi_{\lambda} e^{-ix\lambda}|\lambda'\rangle$$

$$S(x)WS(x)|\psi\rangle = \sum_{\lambda'\lambda} W_{\lambda'\lambda}\psi_{\lambda} e^{-ix(\lambda+\lambda')}|\lambda'\rangle$$

The output state contains all possible sums of frequencies

Output of the Quantum Model

Output of the Quantum Model

Output is expectation value and therefore contains a complex conjugation

$$f_{M,\boldsymbol{\theta}}(x) = \langle \psi_{\boldsymbol{\theta}}(x) | M | \psi_{\boldsymbol{\theta}}(x) \rangle = \sum_{\omega \in \Omega} c_\omega(M, \boldsymbol{\theta}) e^{i\omega x}$$

Output of the Quantum Model

Output is expectation value and therefore contains a complex conjugation

$$f_{M,\boldsymbol{\theta}}(x) = \langle \psi_{\boldsymbol{\theta}}(x)|M|\psi_{\boldsymbol{\theta}}(x)\rangle = \sum_{\omega \in \Omega} c_\omega(M, \boldsymbol{\theta}) e^{i\omega x}$$

For one layer of encoding

$$\Omega = \{\lambda_j - \lambda_k \,|\, \lambda_j, \lambda_k \in \mathrm{spec}(H)\}$$

Output of the Quantum Model

Output is expectation value and therefore contains a complex conjugation

$$f_{M,\boldsymbol{\theta}}(x) = \langle \psi_{\boldsymbol{\theta}}(x) | M | \psi_{\boldsymbol{\theta}}(x) \rangle = \sum_{\omega \in \Omega} c_\omega(M, \boldsymbol{\theta}) e^{i\omega x}$$

For L layers of encoding

$$\Omega = \{\lambda_{j_1} + \cdots + \lambda_{j_L} - \lambda_{k_1} - \cdots - \lambda_{k_L} \,|\, \lambda_{j_l}, \lambda_{k_l} \in \mathrm{spec}(H)\}$$

Output of the Quantum Model

Output is expectation value and therefore contains a complex conjugation

$$f_{M,\boldsymbol{\theta}}(x) = \langle \psi_{\boldsymbol{\theta}}(x) | M | \psi_{\boldsymbol{\theta}}(x) \rangle = \sum_{\omega \in \Omega} c_\omega(M, \boldsymbol{\theta}) e^{i\omega x}$$

For L layers of encoding

$$\Omega = \{\lambda_{j_1} + \cdots + \lambda_{j_L} - \lambda_{k_1} - \cdots - \lambda_{k_L} \,|\, \lambda_{j_l}, \lambda_{k_l} \in \mathrm{spec}(H)\}$$

The accessible spectrum consists of all sums of differences of eigenvalues of the generator of the data encoding

Take-Home Message #2

Quantum learning models output Fourier series, repeating data encoding gives access to higher frequencies

Pauli-Encodings

Pauli-Encodings

Pauli rotations are the most popular encoding strategy, e.g.

$$S(x) = R_Z(x) = e^{-ixZ/2}$$

Pauli-Encodings

Pauli rotations are the most popular encoding strategy, e.g.

$$S(x) = R_Z(x) = e^{-ixZ/2}$$

Pauli rotations give an integer spectrum

Pauli-Encodings

Pauli rotations are the most popular encoding strategy, e.g.

$$S(x) = R_Z(x) = e^{-ixZ/2}$$

Pauli rotations give an integer spectrum

The number of available frequencies grows linearly in depth and width

Pauli-Encodings

Pauli rotations are the most popular encoding strategy, e.g.

$$S(x) = R_Z(x) = e^{-ixZ/2}$$

Pauli rotations give an integer spectrum

The number of available frequencies grows linearly in depth and width

But for general encodings the dependence can be exponential:

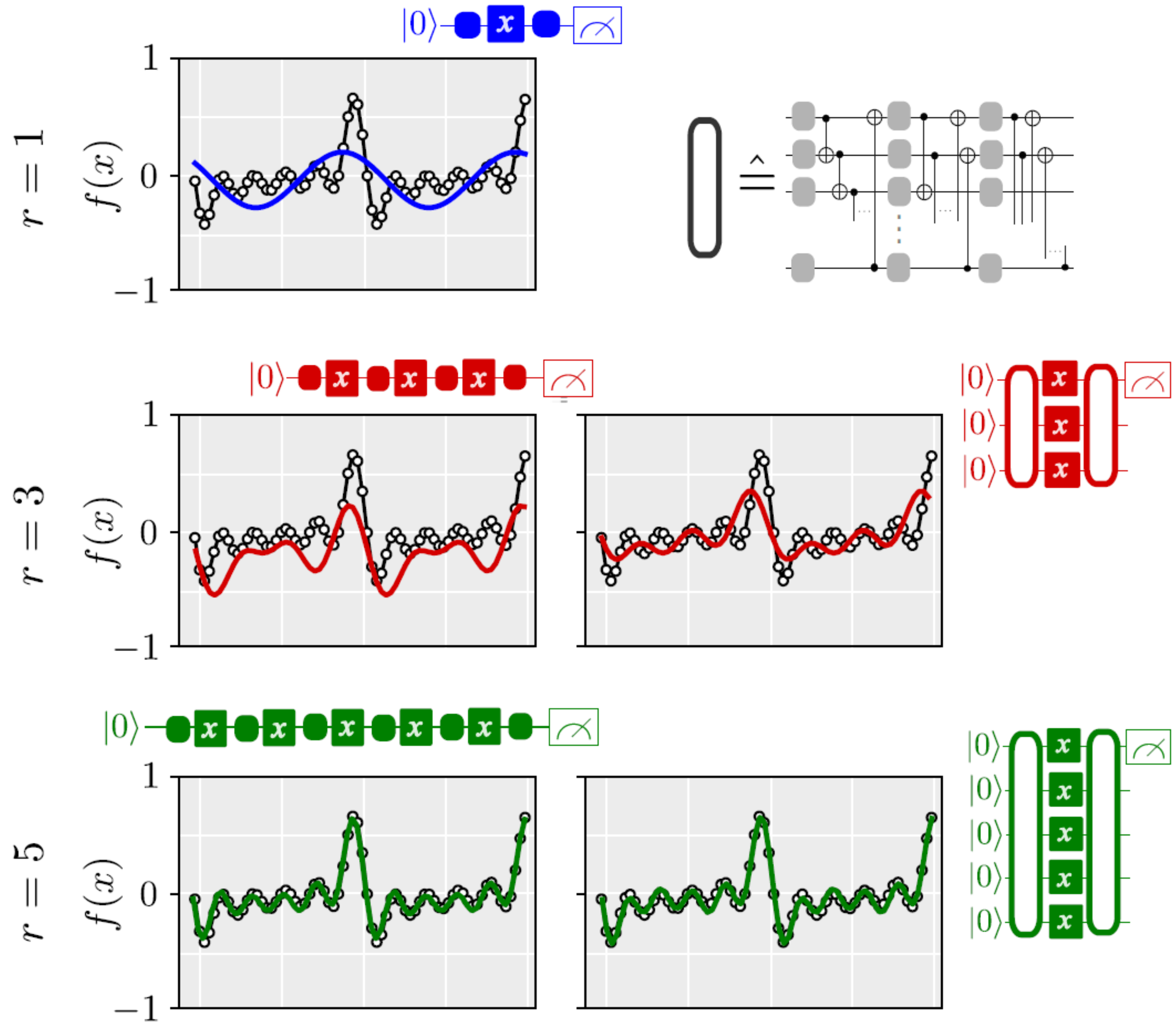
$$\#(\text{frequencies}) \leq \frac{d^{2L}}{2} - 1$$

d System dimension

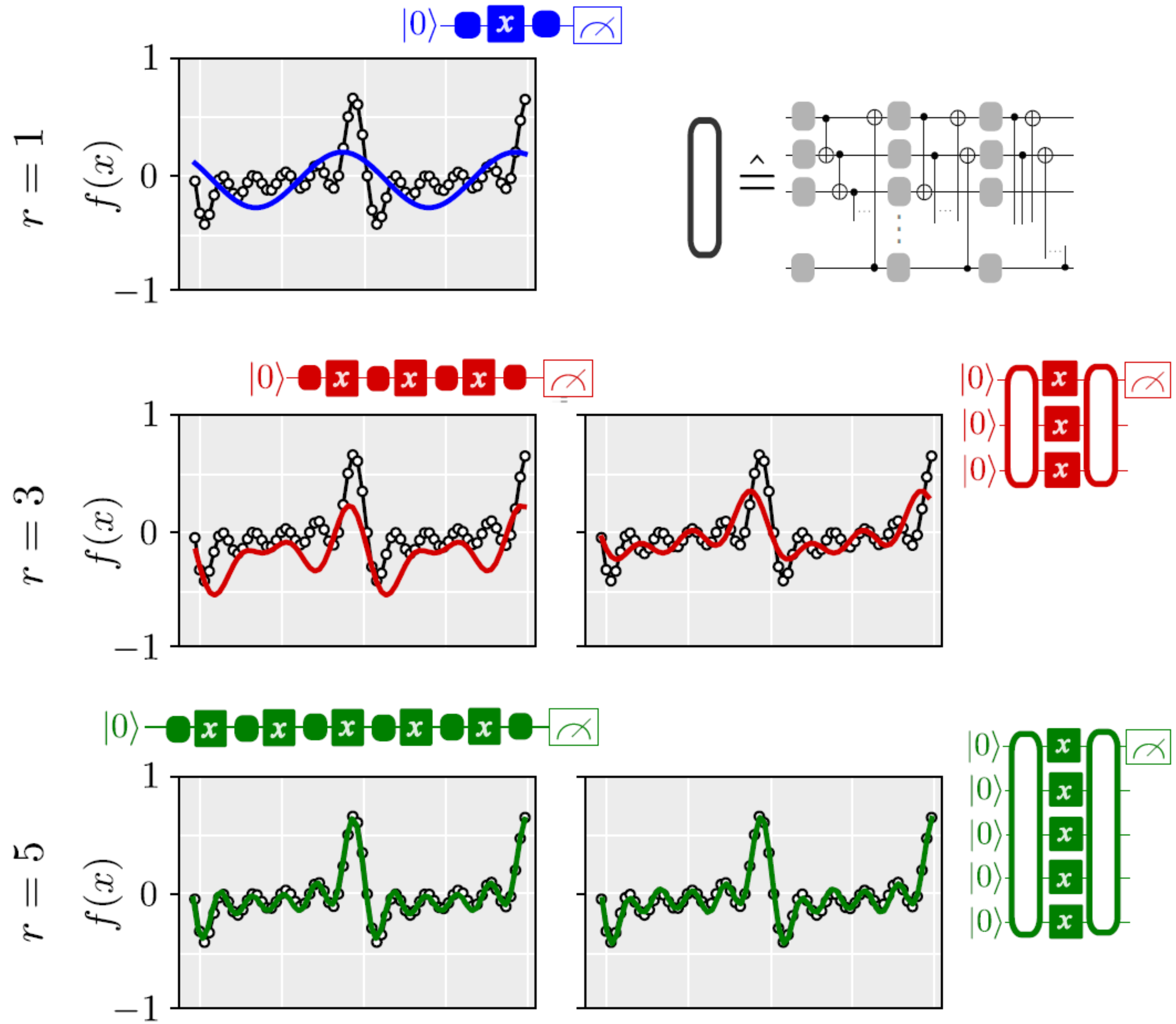
L Number of layers

Consequences for Learning Tasks

Consequences for Learning Tasks



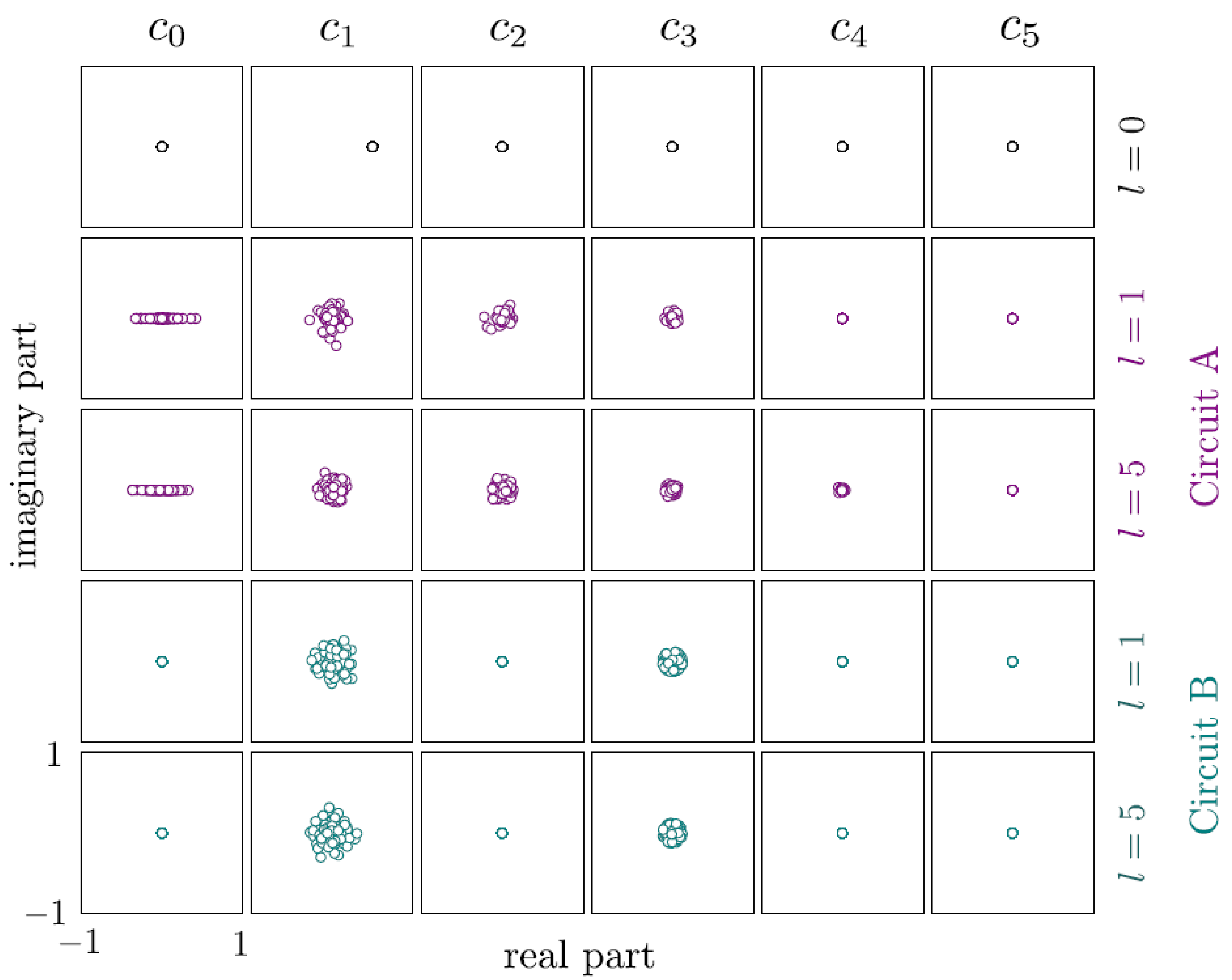
Consequences for Learning Tasks

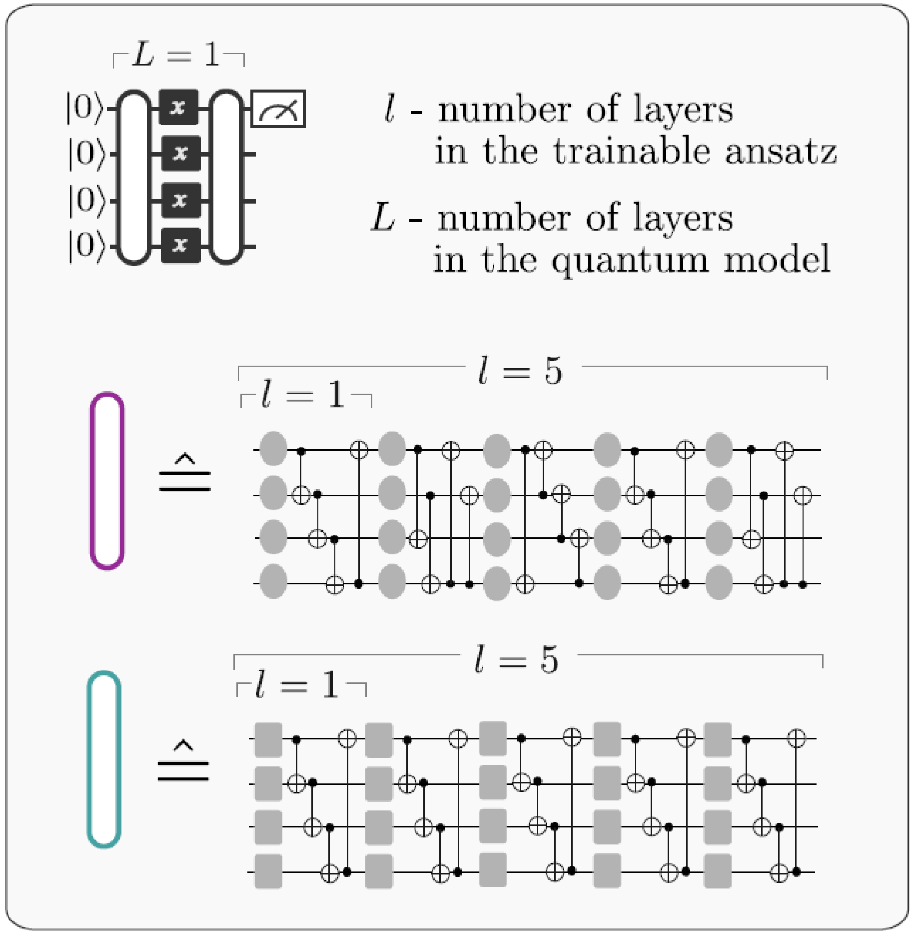


You can reproduce all figures from the paper at home!

Distribution of Fourier Coefficients

Distribution of Fourier Coefficients

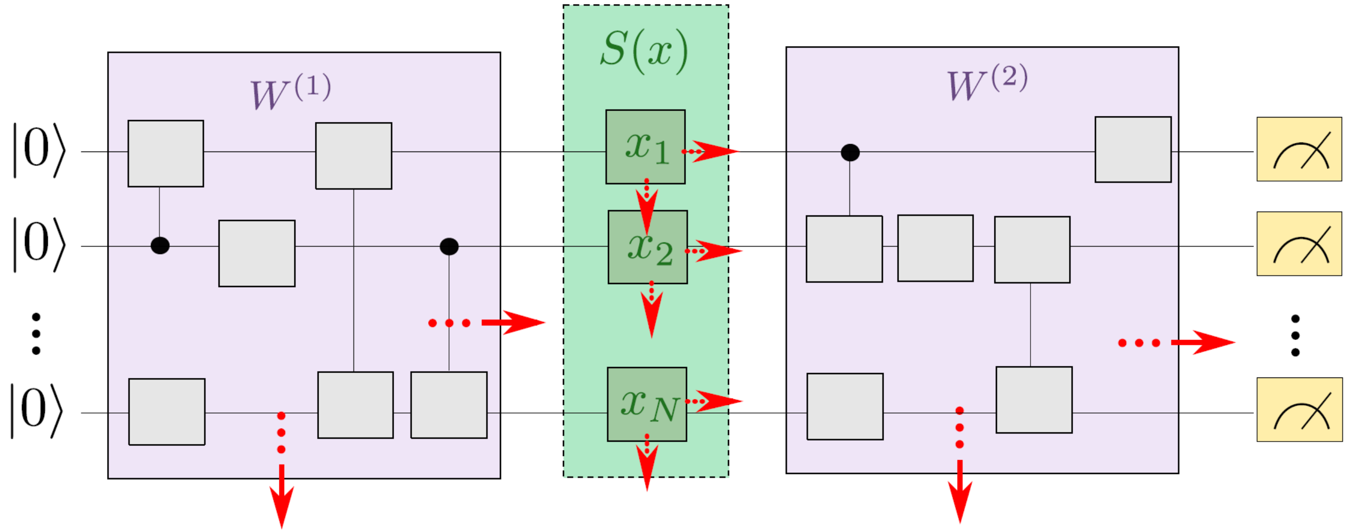




Universality of Quantum Models

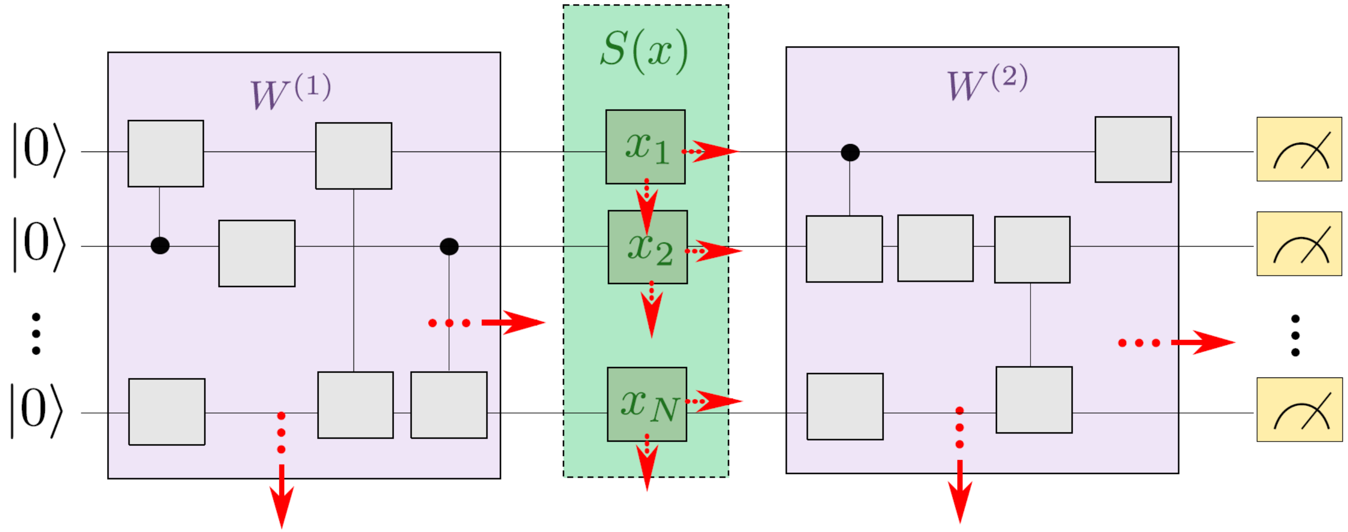
Universality of Quantum Models

A model with one layer of data encodings generated by a *universal Hamiltonian family* and arbitrary unitaries is a universal function approximator



Universality of Quantum Models

A model with one layer of data encodings generated by a *universal Hamiltonian family* and arbitrary unitaries is a universal function approximator



A universal Hamiltonian family asymptotically has access to all integer frequencies. Repeated single-qubit Pauli rotation encodings are a universal Hamiltonian family!

Take-Home Message #3

> Quantum learning models are universal function approximators

Practical Implications for Model Building

Practical Implications for Model Building

1. Know your data encoding, it fundamentally limits what you can learn!

Practical Implications for Model Building

1. Know your data encoding, it fundamentally limits what you can learn!
2. Powerful quantum computers can make stupid models

Practical Implications for Model Building

1. Know your data encoding, it fundamentally limits what you can learn!
2. Powerful quantum computers can make stupid models
3. Rescale your data wisely

Practical Implications for Model Building

1. Know your data encoding, it fundamentally limits what you can learn!
2. Powerful quantum computers can make stupid models
3. Rescale your data wisely
4. Classical pre-processing can alter the model‘s output dramatically

Practical Implications for Model Building

1. Know your data encoding, it fundamentally limits what you can learn!
2. Powerful quantum computers can make stupid models
3. Rescale your data wisely
4. Classical pre-processing can alter the model‘s output dramatically
5. Make your observables trainable

Open Questions

Open Questions

1. Can we use this approach to calculate meaningful generalisation measures?

Open Questions

1. Can we use this approach to calculate meaningful generalisation measures?
2. Can we link specific ansatz classes for the trainable blocks to the output Fourier coefficients?

Open Questions

1. Can we use this approach to calculate meaningful generalisation measures?
2. Can we link specific ansatz classes for the trainable blocks to the output Fourier coefficients?
3. Is universal approximation possible with fixed qubit numbers?

Open Questions

1. Can we use this approach to calculate meaningful generalisation measures?
2. Can we link specific ansatz classes for the trainable blocks to the output Fourier coefficients?
3. Is universal approximation possible with fixed qubit numbers?
4. Are quantum models good for signal processing?

Thank you for your attention!

Paper

Demo

Slides