IOP QUANTUM 2020

What Functions can Quantum Models Learn?

JOHANNES JAKOB MEYER, FU BERLIN

arxiv:2008.08605

arxiv:2008.08605

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)

arxiv:2008.08605

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

arxiv:2008.08605

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

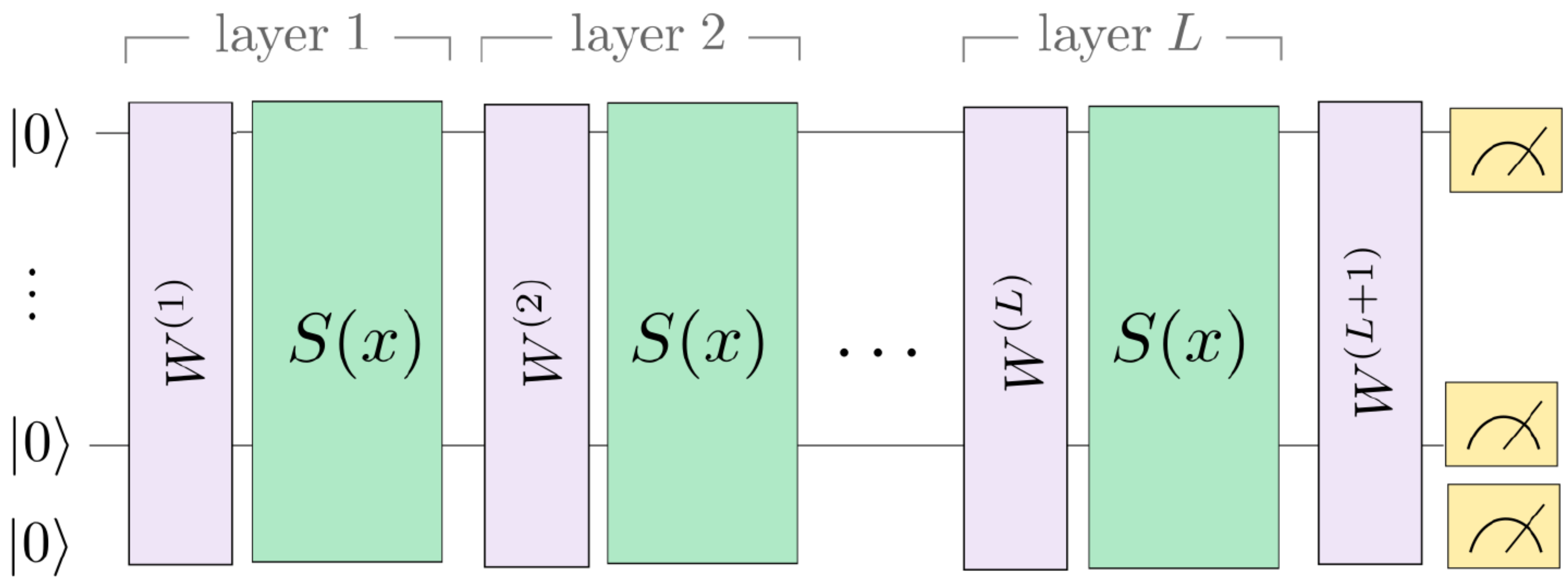
Key Motivation

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

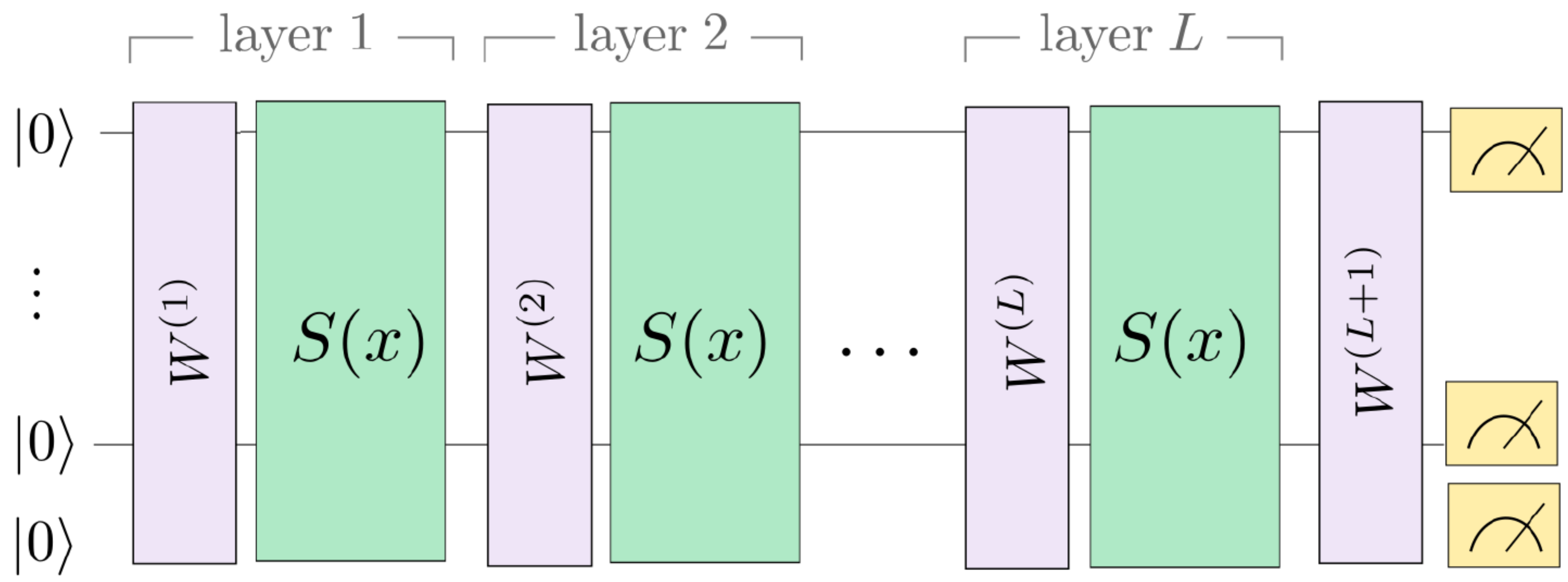
A lot of work has been done to understand the practical side, but very little is known on the theory side

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

Variational Quantum Learning Models



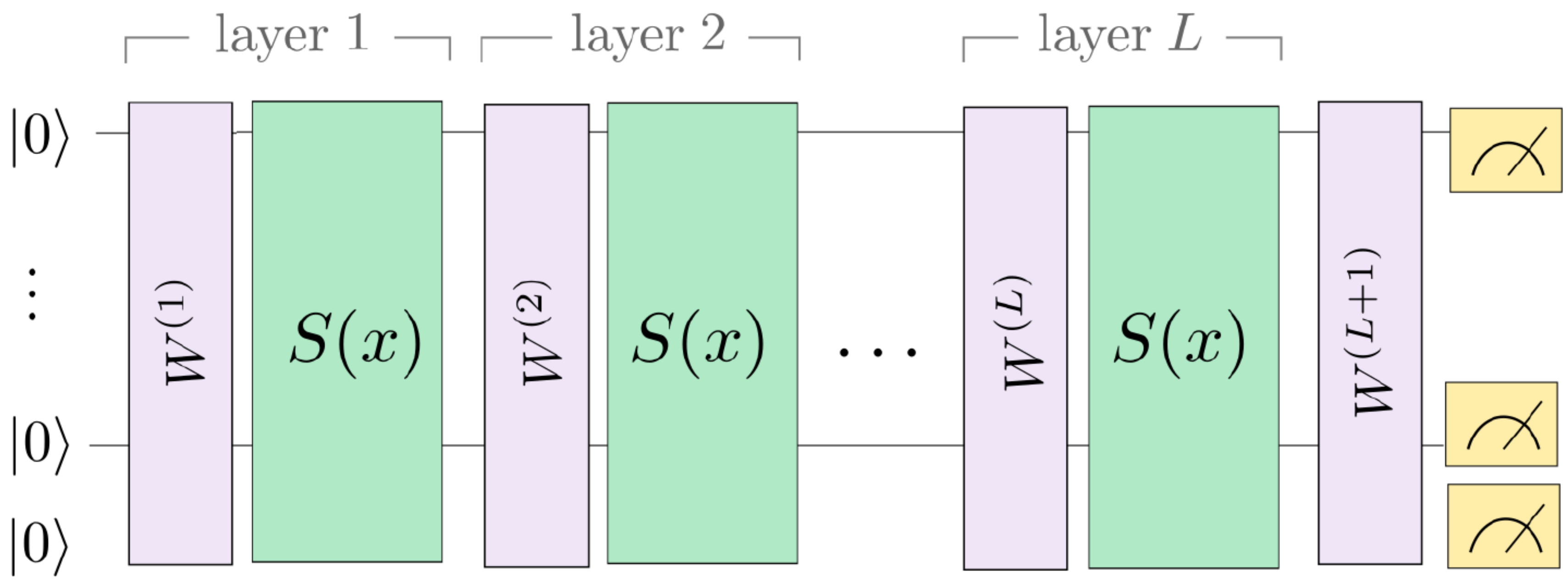
Variational Quantum Learning Models



$$\boxed{f(x) = \langle M \rangle}$$

Model Output

Variational Quantum Learning Models



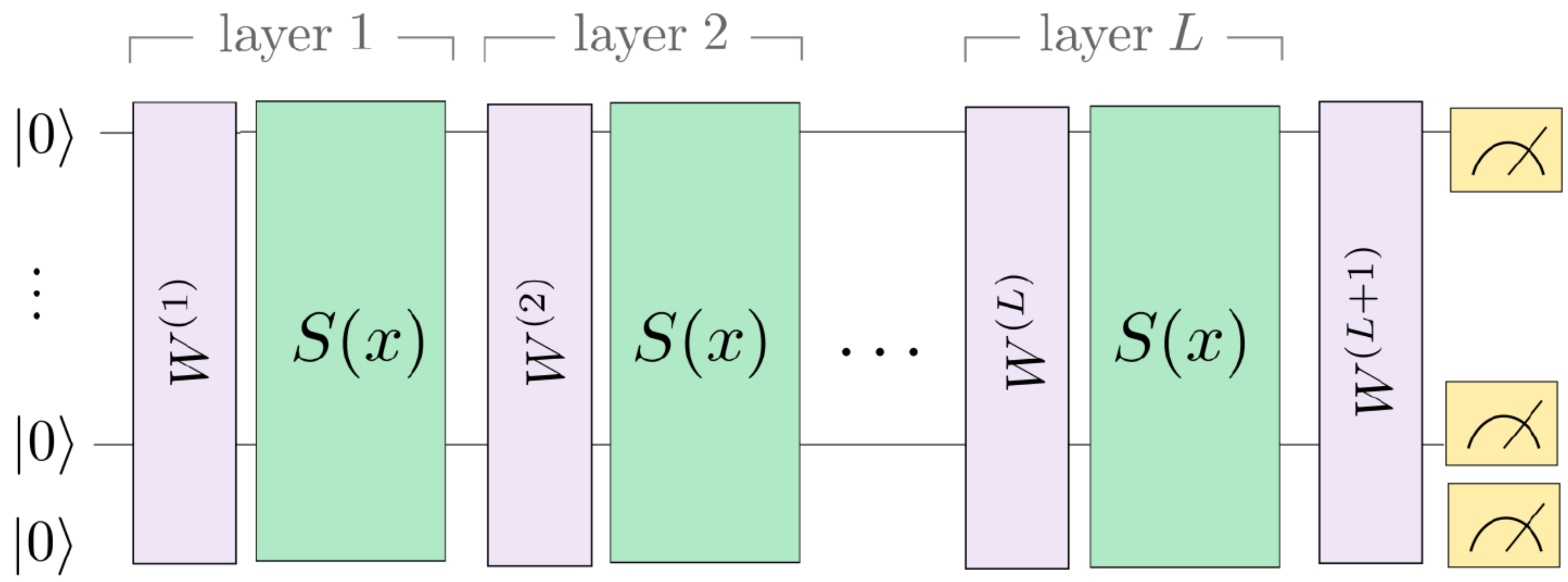
$$f(x) = \langle M \rangle$$

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 #1

> 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

For general encodings the dependence can be exponential:

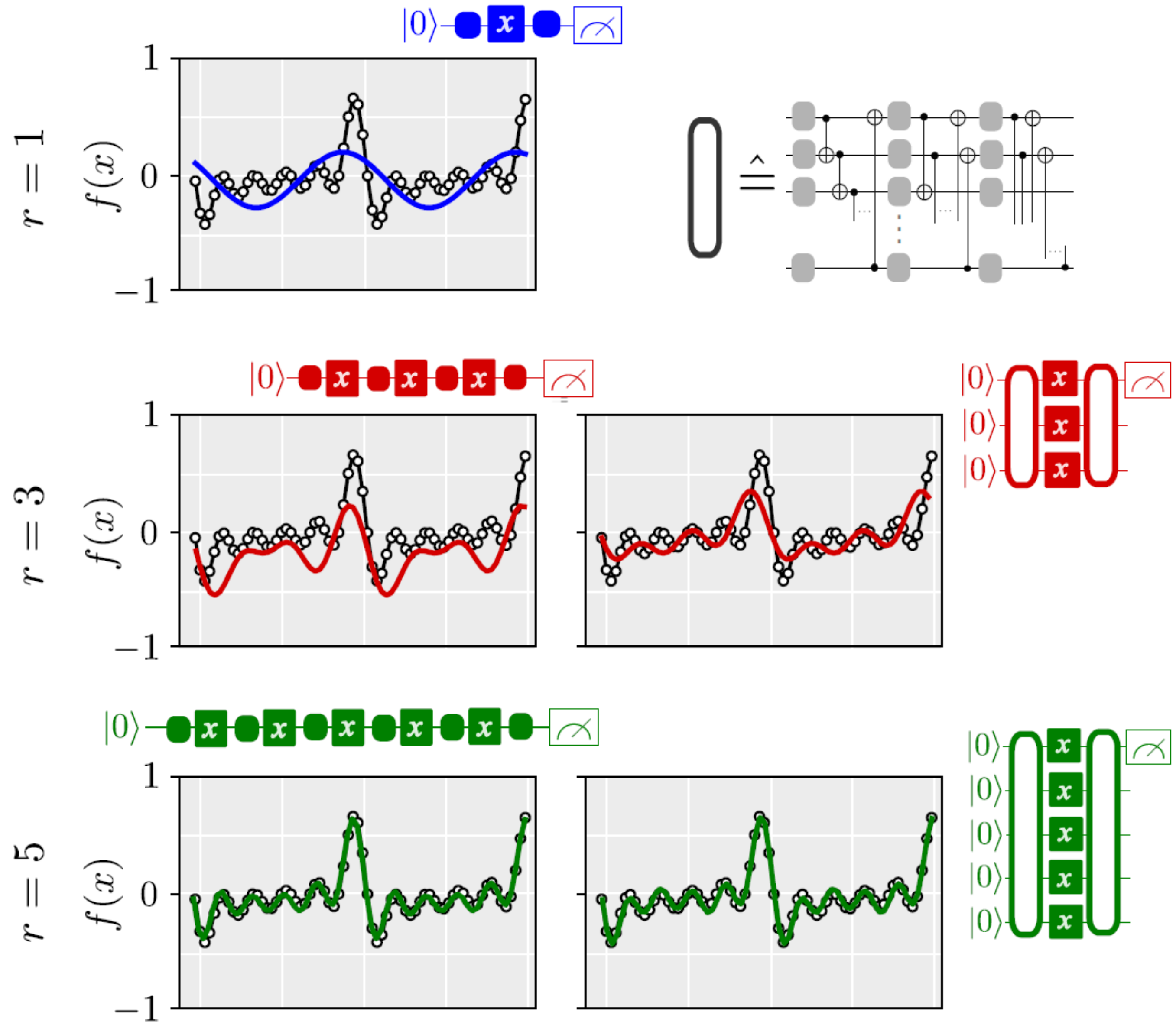
$$\#(\text{frequencies}) \leq \frac{d^{2L}}{2} - 1$$

d System dimension

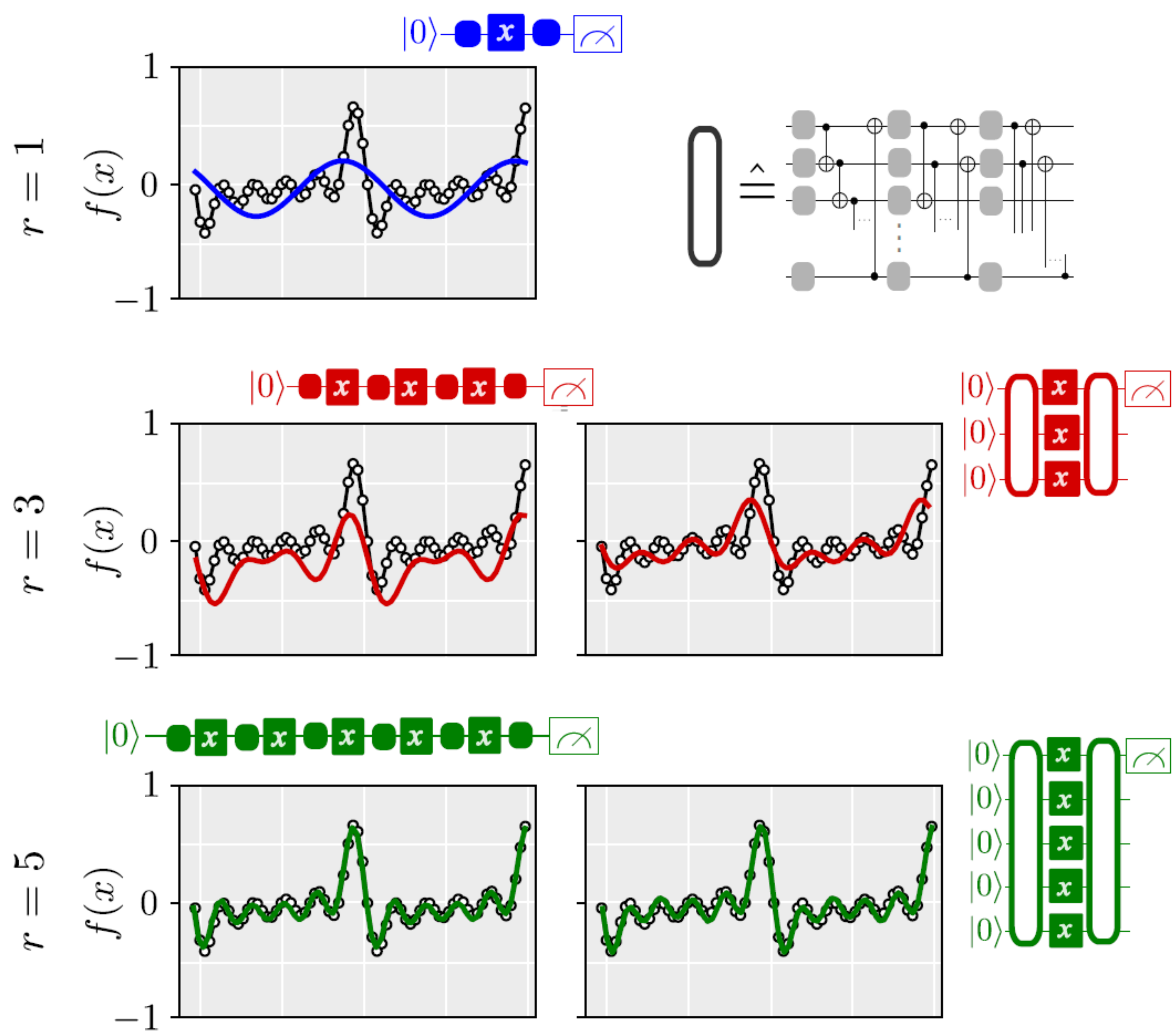
L Number of layers

Consequences for Learning Tasks

Consequences for Learning Tasks



Consequences for Learning Tasks

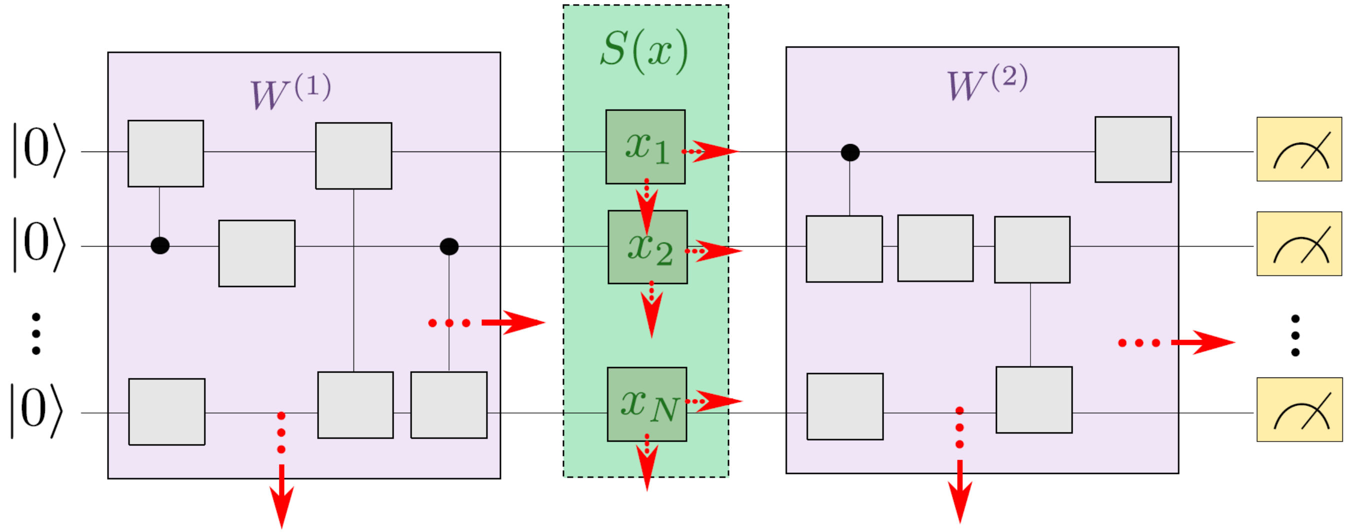


You can reproduce all figures from the paper at home!

Universality of Quantum Models

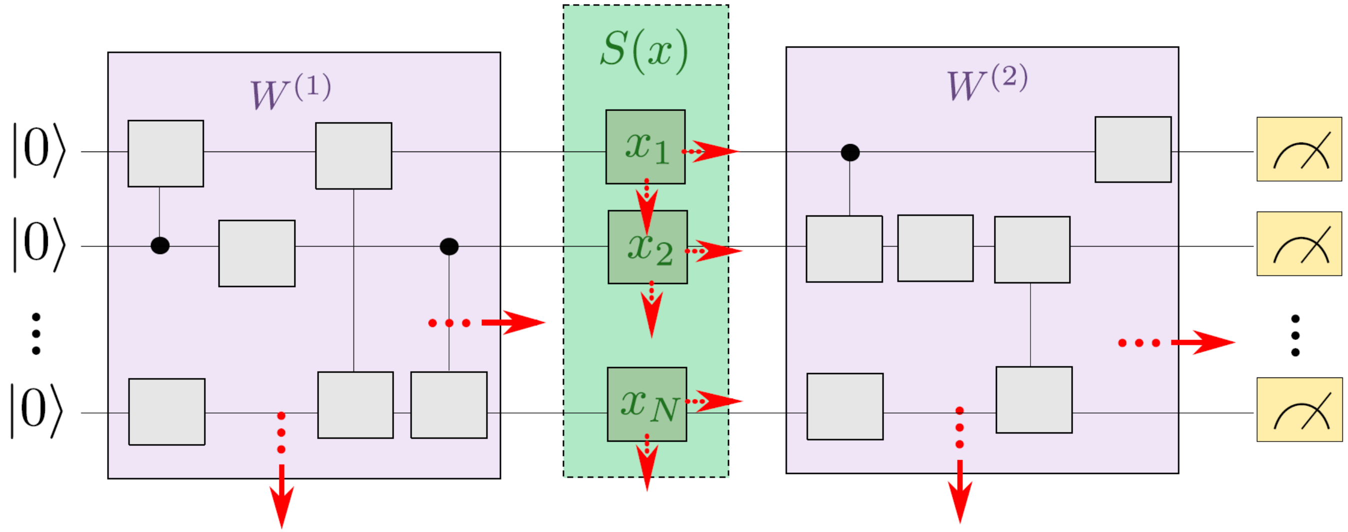
Universality of Quantum Models

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 #2

> 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