Minimial-Risk Training Samples for QNN Training from Measurements

IAAS Research

University of Stuttgart

Alexander Mandl, Johanna Barzen, Marvin Bechtold, Frank Leymann

[lastname]@iaas.uni-stuttgart.de

Institute of Architecture of Application Systems

EniQmA

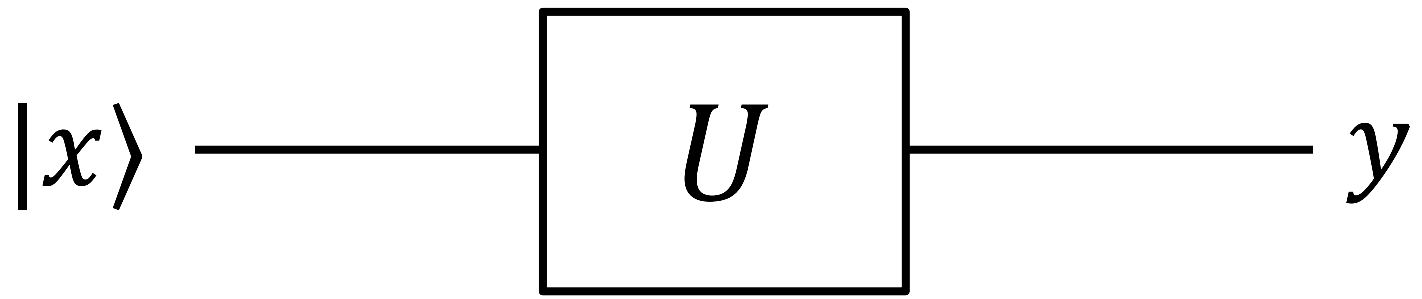
SeQuenC

Motivation and Content

- Preliminaries: Supervised learning on quantum computers
  - Learning from output states vs. learning from measurements
  - Requirements for minimal risk
- Research question
- Analytical results
- Experimental evaluation
- Summary and future work

Supervised Learning using QNNs

- Approximate unknown *target transformation* U
- By using quantum states as training samples

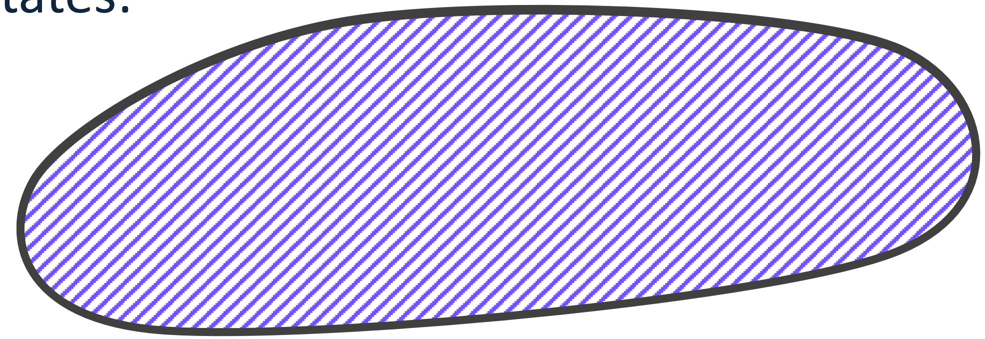


- **Goal:**
  - Obtain a quantum circuit V that behaves the same as U **on the training data**
    - Minimize a **loss function** on the training data
  - **Assumption:** It behaves the same as U **on all possible inputs**

QNN Quality: Risk

Risk: Average QNN-loss on all possible inputs

Set of all quantum states:



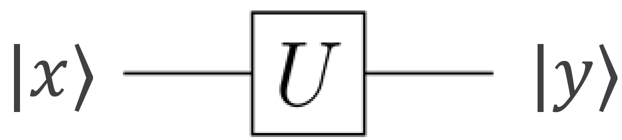
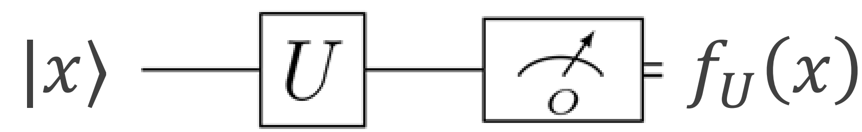
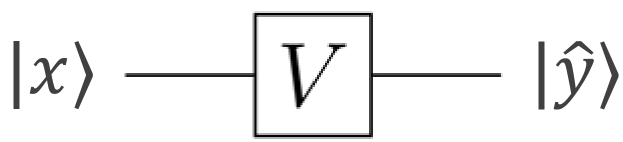
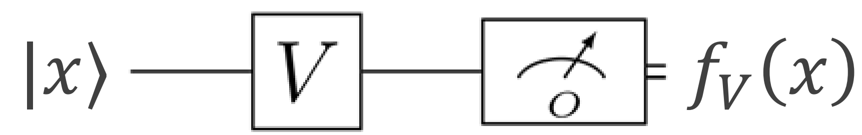
Loss

Average QNN error on finite set of training samples

Risk

Average QNN error on infinite set of possible quantum input states

Different Information/Different Scenarios

	1: Learning output **states**	2: Learning **measurements**
training samples		
QNN output		
Loss function	Mean of $1 - F(\|y\rangle, \|\hat{y}\rangle)$	Mean of $\left(f_U(x) - f_V(x)\right)^2$

U: original transformation, O: observable, V: QNN, F: state fidelity

Supervised Learning using QNNs

- When **learning output states:**
  - Entanglement in training samples reduces expected risk

Expected risk after training $\geq 1 - O\left((rt)^2\right)$

r… degree of entanglement, t… number of samples

- Sharma, Kunal, et al. "Reformulation of the no-free-lunch theorem for entangled datasets." *Physical Review Letters* 128.7 (2022): 070501.
- Mandl, Alexander, et al. "On Reducing the Amount of Samples Required for Training of QNNs: Constraints on the Linear Structure of the Training Data." *arXiv preprint arXiv:2309.13711* (2023).

Requirements for Minimal Risk

- Entanglement proves to be a **valuable resource** when learning output states

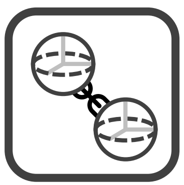
- Learning from output states:
  - General bounds on the risk are proven
  - Mathematical structure of the training data for **minimal expected risk** is described

- Learning from measurements:
  - Some bounds on the expected risk after training are known
  - No complete description of the training data **minimal expected risk** available

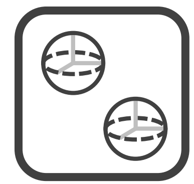
- Sharma, Kunal, et al. "Reformulation of the no-free-lunch theorem for entangled datasets." *Physical Review Letters* 128.7 (2022): 070501.
- Mandl, Alexander, et al. "On Reducing the Amount of Samples Required for Training of QNNs: Constraints on the Linear Structure of the Training Data." *arXiv preprint arXiv:2309.13711* (2023).

Research Questions/Contributions

- Find requirements for minimal risk when learning from measurements



For highly entangled data



Without entanglement

- First step: Limit observables to one-dimensional projectors $O = |o\rangle\langle o|$
  - Provide fundamentals for future generalization

Minimal Risk Training Samples

Methods

Risk: average loss of **infinitely many** possible input states

Training loss impact the risk after training negatively

Reformulate risk in terms of **fidelity** of a pair of states

Use best-case as guideline: QNN V is **perfectly trained**

Infer loss on arbitrary inputs (= risk).

Minimal Risk Training Samples: No Entanglement

- Train QNN V to
  - replicate operator U when measured with observable $O = |o\rangle\langle o|$
  - using a set of training samples S

A **single training** input $|\gamma\rangle = U^{\dagger}|o\rangle$ with its associated output $f_U(\gamma)$ suffices to train V with **zero risk**.

Minimal Risk Training Samples: Entangled Data

- Train QNN V to
  - replicate operator U when measured with observable $O = |o\rangle\langle o|$
  - using a set of training samples S **that are entangled with an auxiliary system**

Schmidt coefficients

Schmidt basis states

$$|x\rangle = c_1|a_1\rangle \otimes |b_1\rangle + c_2|a_2\rangle \otimes |b_2\rangle \ldots$$

If the training input contains $|\gamma\rangle$ as the **basis state with the largest coefficient c_i**, then this entangled input with its associated output $f_U(x)$ suffices to train V with zero risk.

$$|x\rangle = c_1|a_1\rangle \otimes |b_1\rangle + c_2|\gamma\rangle \otimes |b_2\rangle \ldots$$

Research

Minimal Risk Training Samples: Entangled Data

- Train QNN V to
  - replicate operator U when measured with observable $O = |o\rangle\langle o|$
  - using a set of training samples S **that are entangled with an auxiliary system**

Schmidt coefficients

Schmidt basis states

$$|x\rangle = c_1|a_1\rangle \otimes |b_1\rangle + c_2|a_2\rangle \otimes |b_2\rangle \dots$$

If the training input contains $|\gamma\rangle$ as the **basis state with the largest coefficient c_i**, then this entangled input with its associated output $f_U(x)$ suffices to train V with zero risk.

In particular: Always holds if $c_i^2 \geq \frac{1}{2}$

Research

Experimental Evaluation

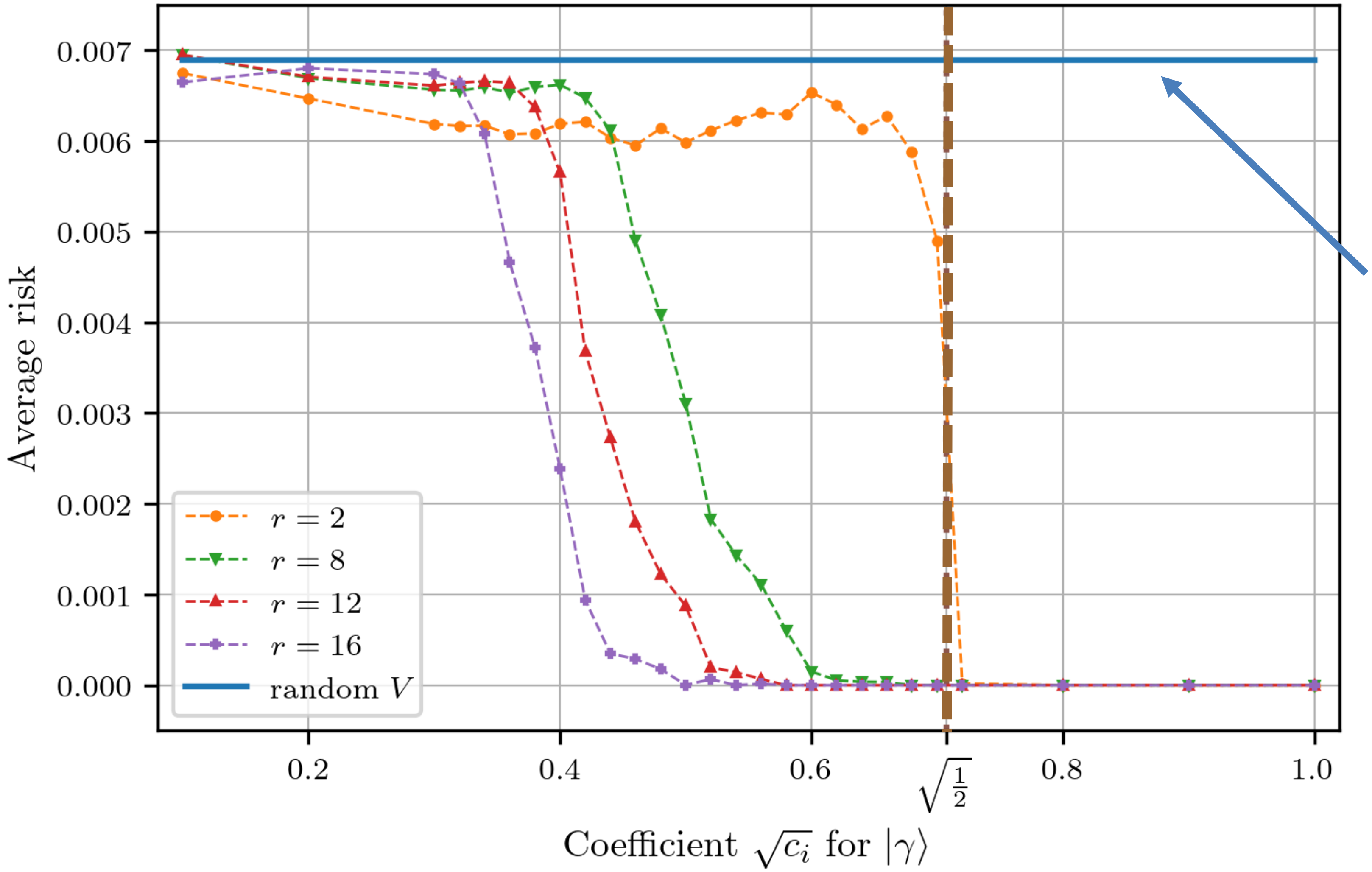
- Analytical investigation found specific training samples that minimize risk
  - Entanglement does not necessarily decrease the risk
- Experiment
  - Evaluate analytical findings
  - Investigate the performance if $|\gamma\rangle$ is **not available** (e.g., random inputs)

Simulated QNN Training

1. Randomly sample 4-qubit state $|o\rangle$ to obtain $O = |o\rangle\langle o|$
2. Randomly sample 4-qubit target operator U
3. Generate training data according to analytical results
4. Optimize parametrized quantum circuit $V(\vec{\theta})$ to minimize training loss
5. Calculate risk after training

Experiment Results

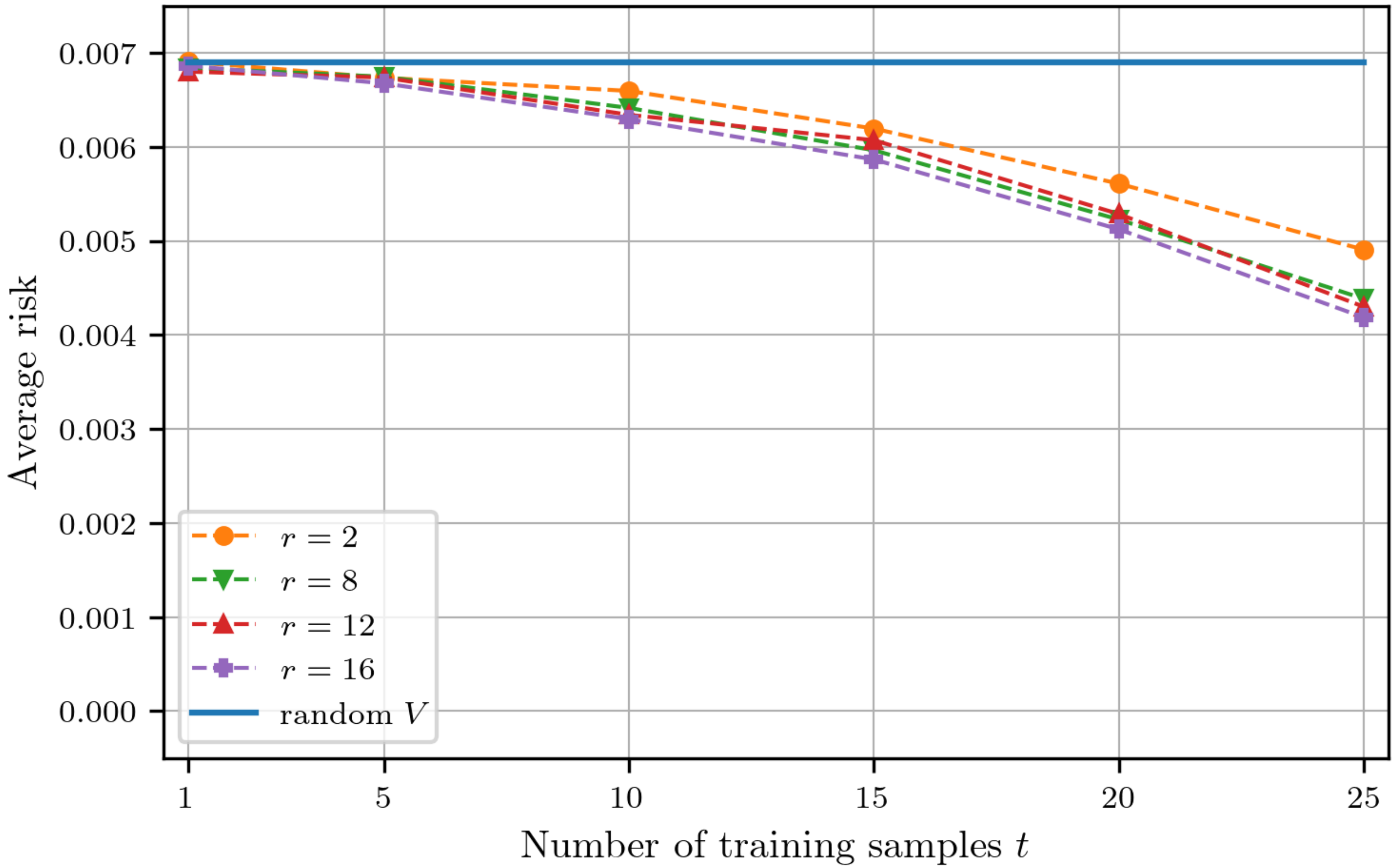
- Effect of the Schmidt coefficient $c_i|\gamma\rangle$ for different Schmidt ranks r for $t = 1$ training sample



Baseline: Expected risk of random quantum circuit

Experiment Results

- Effect of entanglement for randomly sampled training inputs and varying number of training samples



Conclusion and Future Work

- When only the measurement result is known:
  - For one-dimensional projectors: **one training sample is enough**.
  - If sample $|\gamma\rangle$ is known: Entanglement provides no benefit
  - Entanglement produces only **minimal improvement** for random inputs

- Future work:
  - Generalizations for other observables
  - Effect of measurement processes on auxiliary system
  - Perfect training might be hard to achieve: evaluate cost function landscape