Adaptively measuring quantum expectation values using the empirical Bernstein stopping rule

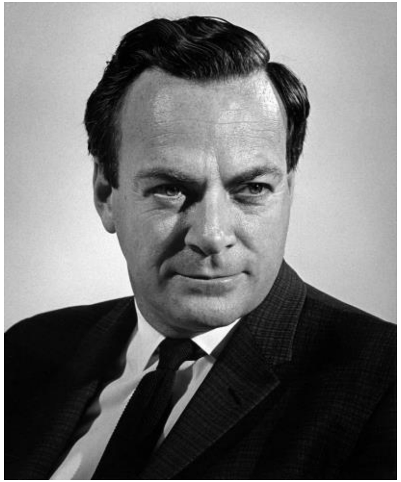
Supervisor: Prof. Martin Kliesch

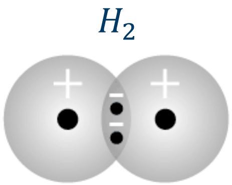
Ugur Tepe

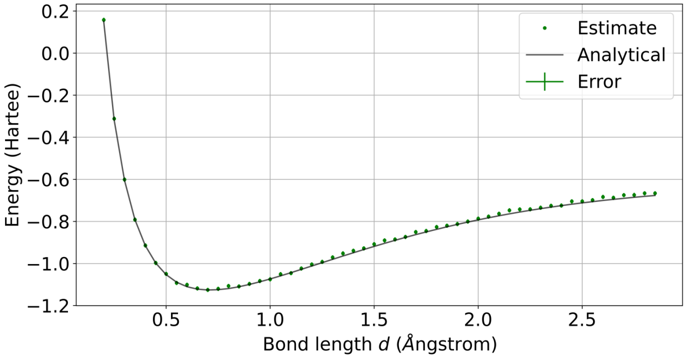
27.09.2023

Why Quantum Computing

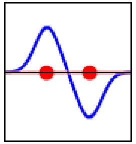
- Physical Problems
  1. Quantum Chemistry
  2. Condensed Matter
  3. …

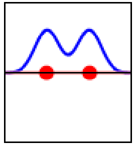






"Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical […]



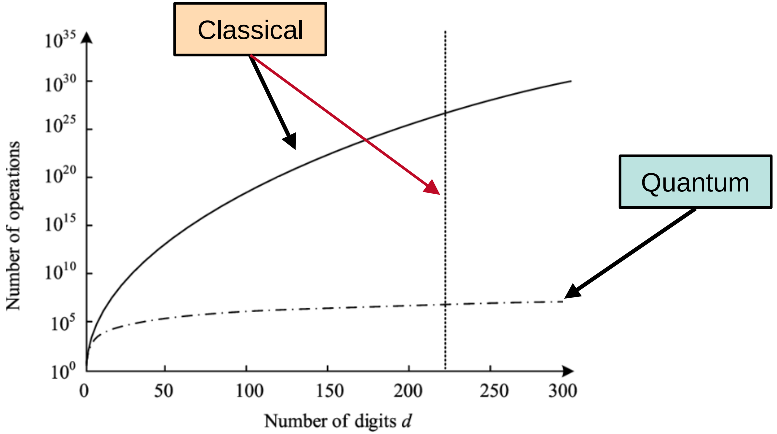


Picture + Quote: see appendix

Why Quantum Computing

- Mathematical Problems
  1. Solve System of Equations
  2. Combinatorial Problems
  3. Finding Prime Factors (Shor‘s algorithm)
  4. …

$$1517 = 41 * 37$$

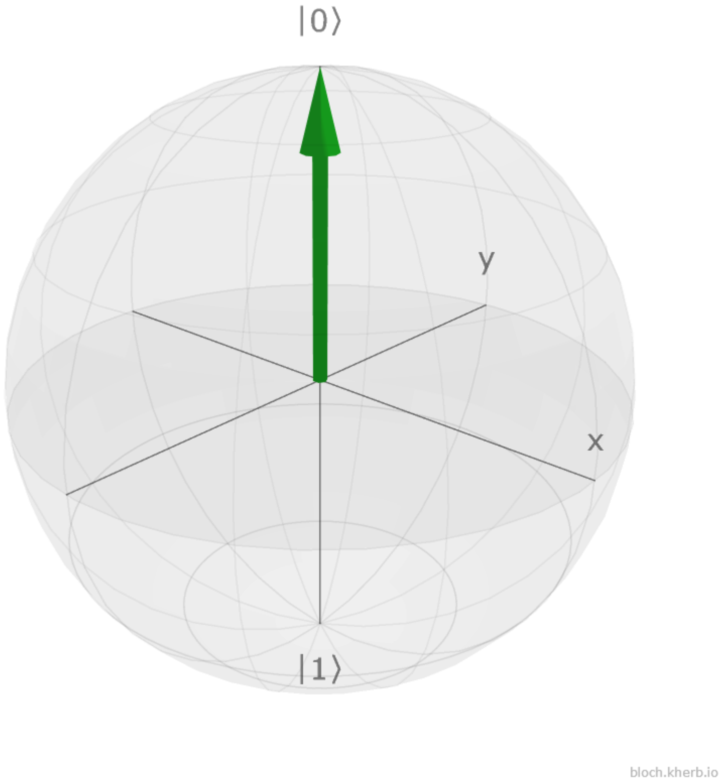


Picture: Dong et. al. Improving the Success Rate of Quantum Algorithm Attacking RSA Encryption System 2022

Quantum Computing

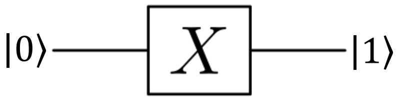
Qubits

- Basis: $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}$, $|1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$
- $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$
- Combine physical systems
  → Tensor product $\otimes$
- $|0\rangle\otimes|0\rangle = |00\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$
- $|01\rangle$, $|10\rangle$, $|11\rangle$ analogous

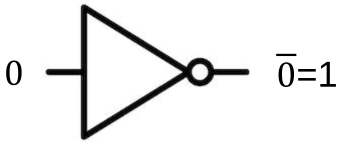


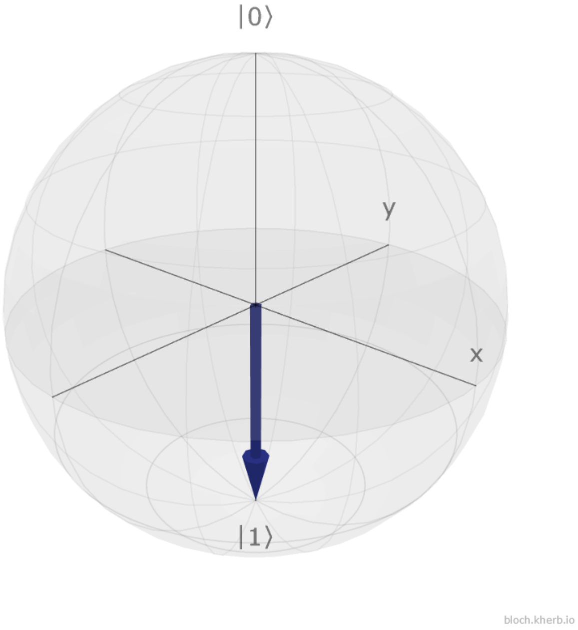
Quantum Gates: X-gate

Quantum



Classical





- $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- X-gate: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

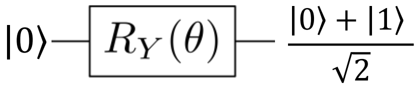
$$X|0\rangle = |1\rangle$$
$$X|1\rangle = |0\rangle$$

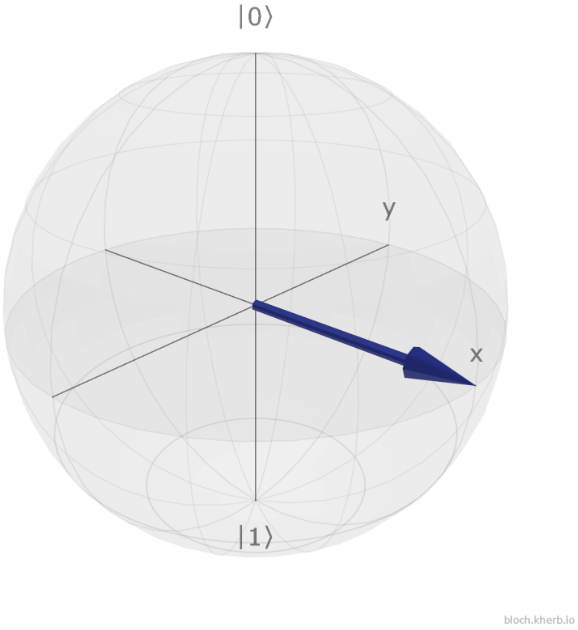
Quantum Computing

Heinrich Heine
Universität Düsseldorf

Quantum Gates: R_y-gate

Quantum





- $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- $R_y(\theta) =$

$$\begin{bmatrix} \cos {}^{\theta}/_{2} & -\sin {}^{\theta}/_{2} \\ \sin \theta & \cos {}^{\theta}/_{2} \end{bmatrix}$$

$$R_y\left(\tfrac{\pi}{2}\right)|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

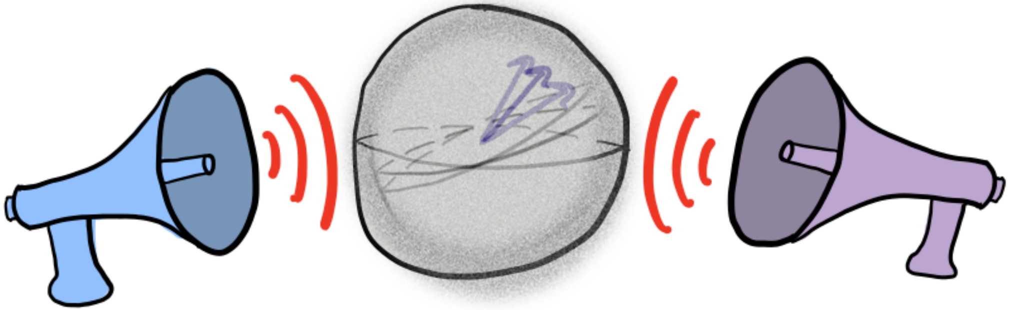
$$R_y\left(\tfrac{\pi}{2}\right)|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

 Picture: https://bloch.kherb.io/

Quantum Computing

Limitation of current quantum computers/ hardware

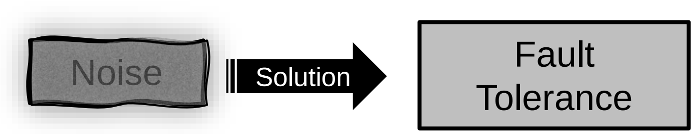
Noise

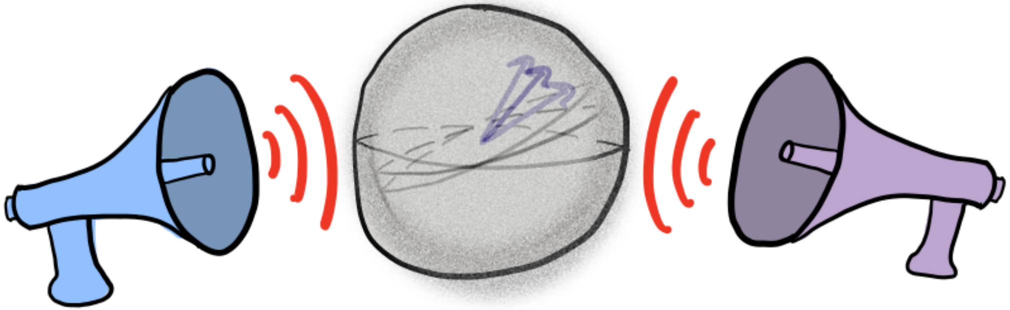


Picture: https://pennylane.ai/qml/demos/tutorial_noisy_circuit_optimization

Quantum Computing

Limitation of current quantum computers/ hardware

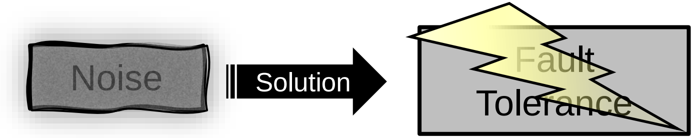


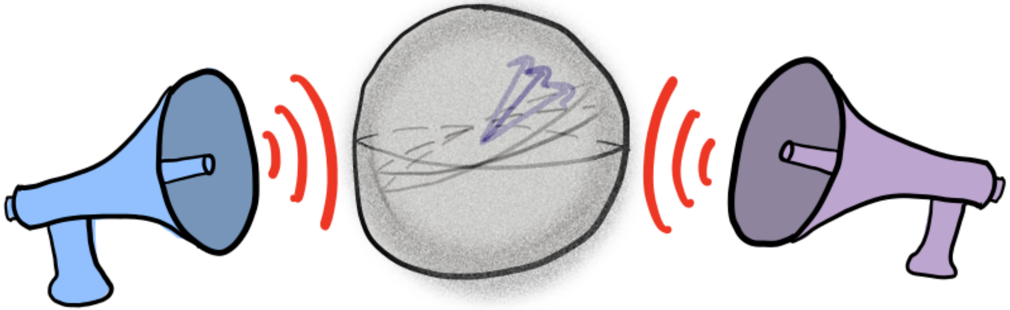


Picture: https://pennylane.ai/qml/demos/tutorial_noisy_circuit_optimization

Quantum Computing

Limitation of current quantum computers/ hardware

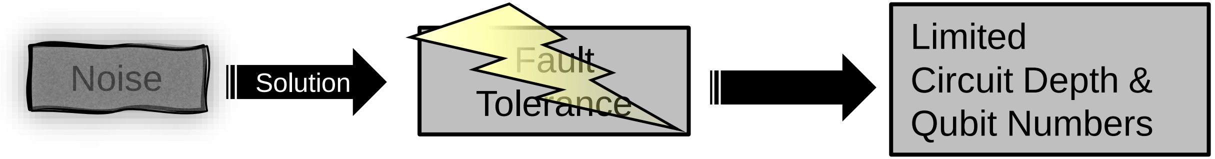


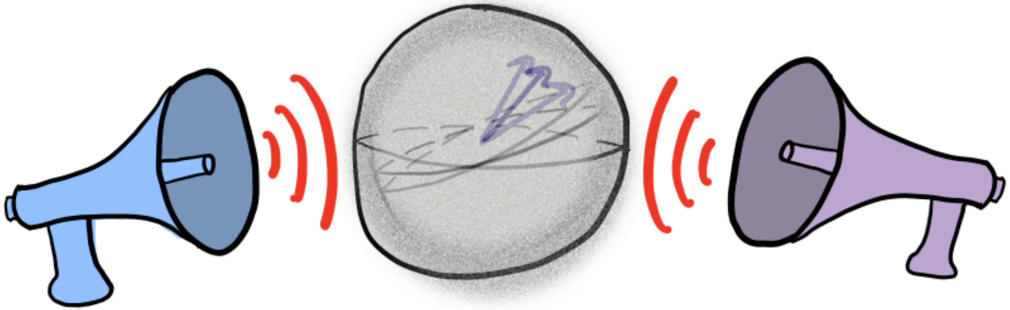


Picture: https://pennylane.ai/qml/demos/tutorial_noisy_circuit_optimization

Quantum Computing

Limitation of current quantum computers/ hardware

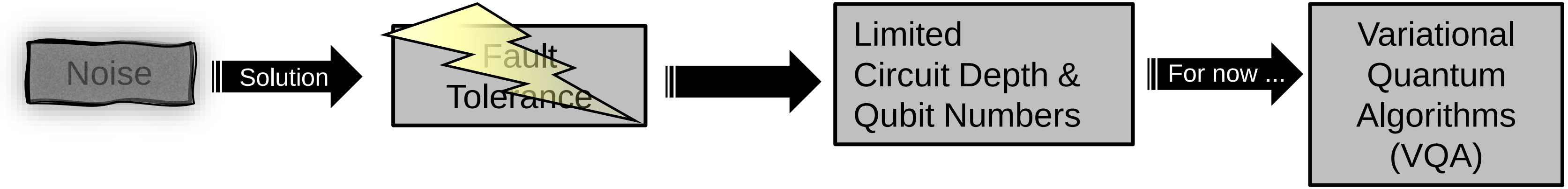


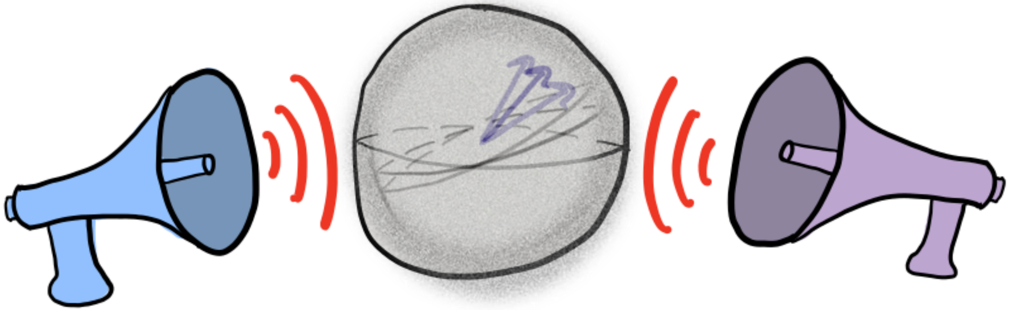


Picture: https://pennylane.ai/qml/demos/tutorial_noisy_circuit_optimization

Quantum Computing

Limitation of current quantum computers/ hardware



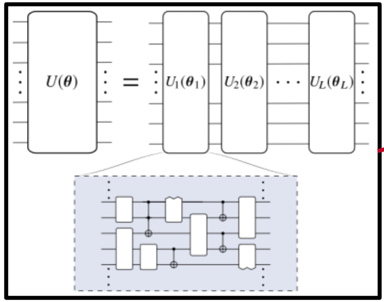


Picture: https://pennylane.ai/qml/demos/tutorial_noisy_circuit_optimization

Schematic Overview of VQA

hhu
Heinrich Heine
Universität Düsseldorf

Quantum | Classical



Find Ansatz

Hybrid loop

Prepare $U(\boldsymbol{\theta})$

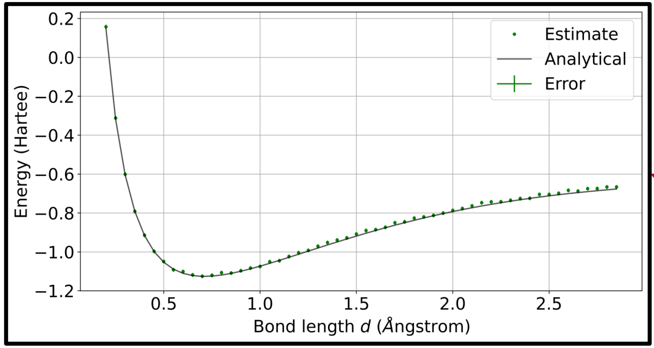
Measure $E(\boldsymbol{\theta})$

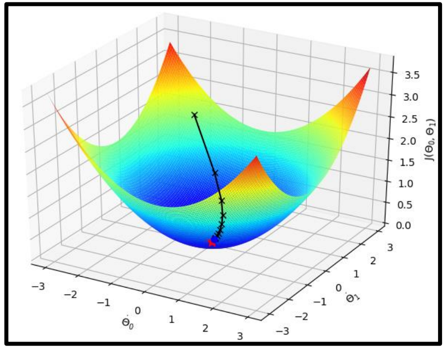
Optimize $\boldsymbol{\theta}$

$|\psi(\boldsymbol{\theta})\rangle \approx |\psi_G\rangle$
$\mathrm{E}(\boldsymbol{\theta}) \approx \mathrm{E}_G$

Rayleigh-Ritz variational principle:
$\mathrm{E}_G \leq \mathrm{E}(\boldsymbol{\theta})$

$E(\theta) = \langle\psi_0|U^{\dagger}HU|\psi_0\rangle$



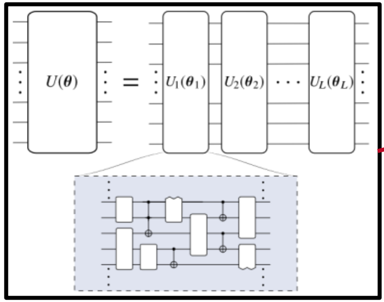


Schematic Overview of VQA

Heinrich Heine
Universität Düsseldorf

Quantum

Classical



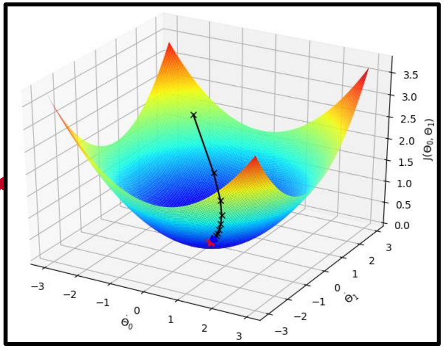
Find Ansatz

Hybrid loop

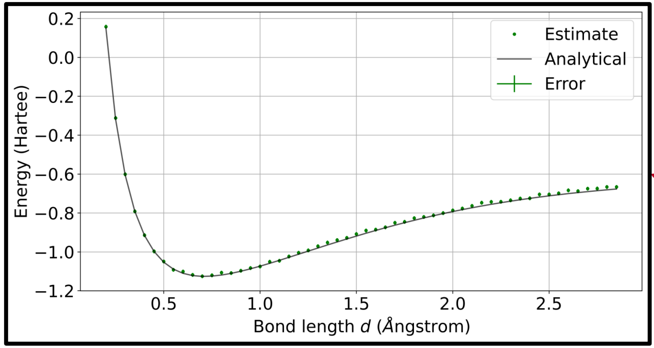
Prepare $U(\boldsymbol{\theta})$

Optimize $\boldsymbol{\theta}$

Measure $E(\boldsymbol{\theta})$



$E(\theta) = \langle\psi_0|U^\dagger H U|\psi_0\rangle$



$|\psi(\boldsymbol{\theta})\rangle \approx |\psi_G\rangle$
$\mathrm{E}(\boldsymbol{\theta}) \approx \mathrm{E}_G$

Require accuracy ϵ
→ Sample complexity bottleneck

Empirical Bernstein Stopping (EBS)

An adaptive sampling algorithm

Tail Bounds

Bernstein‘s Inequality

$$\mathbb{P}[|\bar{X}_t - \mu| \geq \epsilon] \leq exp\left[-\frac{\frac{1}{2}(t\epsilon)^2}{\Sigma^2 + \frac{1}{3}Rt\epsilon}\right] := \delta$$

- So called bound form

$$|\bar{X}_t - \mu| \leq \sqrt{\frac{2\,\Sigma^2 \ln(2/\delta)}{t}} + \frac{R\ln(2/\delta)}{3t}$$

Independent and identically distributed random variable

- $X_1 \ldots X_t \triangleq$ i.i.d random variables
- a $\leq X_i \leq b$
- $\epsilon \in R^+$
- $R \triangleq$ Range

- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$
- $\Sigma^2 = \sum_{i=1}^{t} \sigma_i^2$

Tail Bounds

Heinrich Heine
Universität Düsseldorf

Bernstein's Inequality

Independent and identically distributed random variable

$$\mathbb{P}[|\bar{X}_t - \mu| \geq \epsilon] \leq exp\left[-\frac{\frac{1}{2}(t\epsilon)^2}{\Sigma^2 + \frac{1}{3}Rt\epsilon}\right] := \delta$$

- $X_1 \dots X_t \triangleq$ i.i.d random variables
- $a \leq X_i \leq b$
- $\epsilon \in R^+$
- $R \triangleq$ Range

- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$
- $\Sigma^2 = \sum_{i=1}^{t} \sigma_i^2$

- So called bound form

$$|\bar{X}_t - \mu| \leq \sqrt{\frac{2\,\Sigma^2 \ln(2/\delta)}{t}} + \frac{R\ln(2/\delta)}{3t}$$

- Variance usually unknown !

Heinrich Heine Universität Düsseldorf

Empirical Bernstein Bound

Independent and identically distributed random variable

Variance $\Sigma^2 \xrightarrow{\text{Replace}}$ Empirical variance $\bar{V}_t$

- $X_1 \ldots X_t \triangleq$ i.i.d random variables
- $a \leq X_i \leq b$
- $\epsilon \in R^+$
- $R \triangleq$ Range

- Empirical Bernstein Bound [1]

$$|\bar{X}_t - \mu| \leq \underbrace{\sqrt{\frac{2\bar{\boldsymbol{V}}_t \ln(3/\delta)}{t}}}_{\Theta(1/\sqrt{t})} + \underbrace{\frac{3R\ln(3/\delta)}{t}}_{\Theta(1/t)}$$

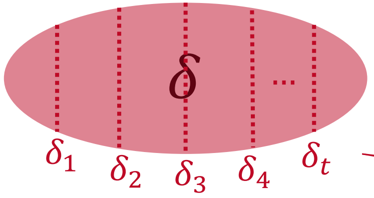
- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$
- $\bar{\boldsymbol{V}}_t = \frac{1}{t}\sum_{i=1}^{t}(X_i - \bar{X}_t)$

- Ideally: $\bar{V}_t \ll R^2$!

1. JY. Audibert et. al., *Tuning bandit algorithms in stochastic environments* (2007)

Empirical Bernstein Stopping (EBS) [2]

Pseudo Code Implementation



- $X_1 \ldots X_t \triangleq$ i.i.d random variables
- $\epsilon \in R^+$
- $1 - \delta \triangleq$ Confidence
- $R \triangleq$ Range

- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$
- $c(t) = \sqrt{\frac{2\bar{V}_t \ln(3/\delta_t)}{t}} + \frac{3R\ln(3/\delta_t)}{t}$

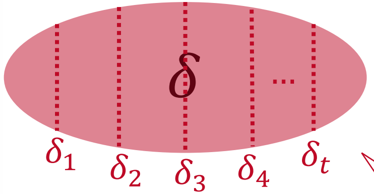
2. V. Mnih, *Efficient stopping rules* (2008)

Empirical Bernstein Stopping (EBS) [2]

Pseudo Code Implementation

```
while ct > eps:
    sample(X)
    update(mean,variance)
    update(ct)

return mean
```



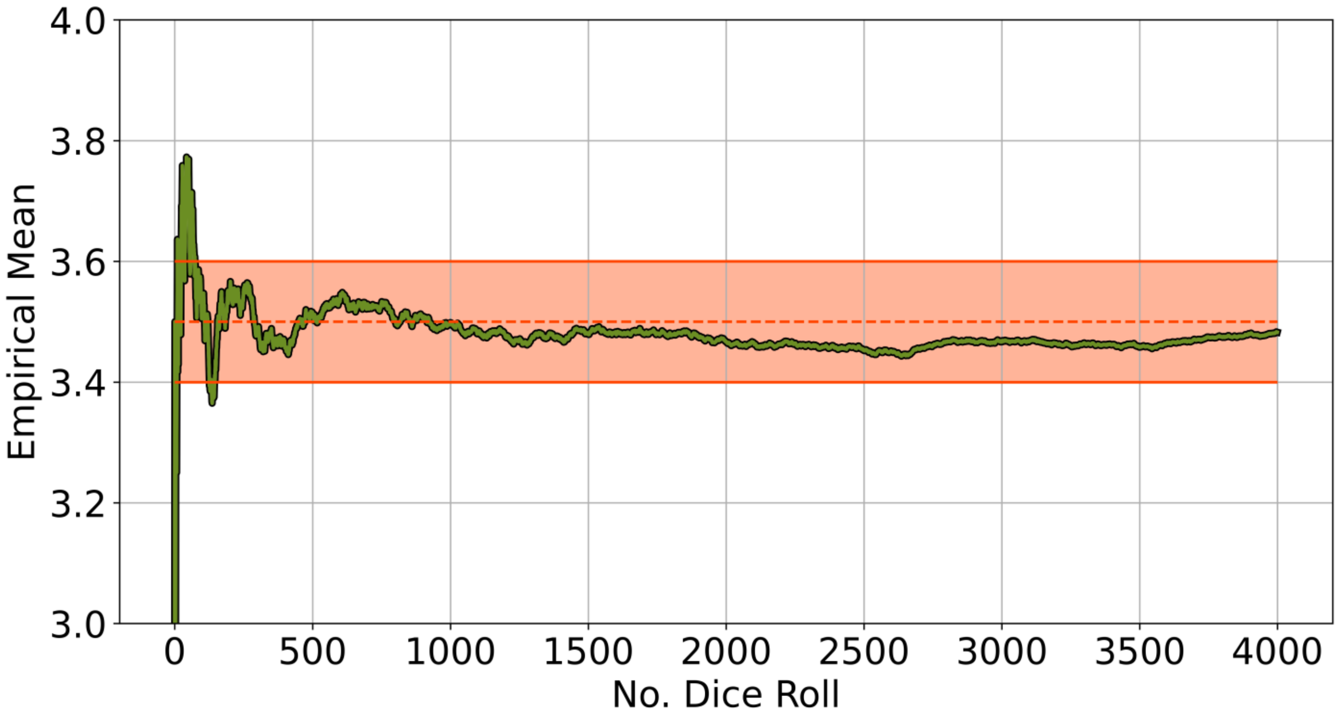
- $X_1 \dots X_t \triangleq$ i.i.d random variables
- $\epsilon \in R^+$
- $1 - \delta \triangleq$ Confidence
- $R \triangleq$ Range

- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$
- $c(t) = \sqrt{\frac{2\bar{V}_t \ln(3/\delta_t)}{t}} + \frac{3R\ln(3/\delta_t)}{t}$

[2]. V. Mnih, *Efficient stopping rules* (2008)

Tail Bounds

Empirical Mean of a Dice: EBS Algorithm

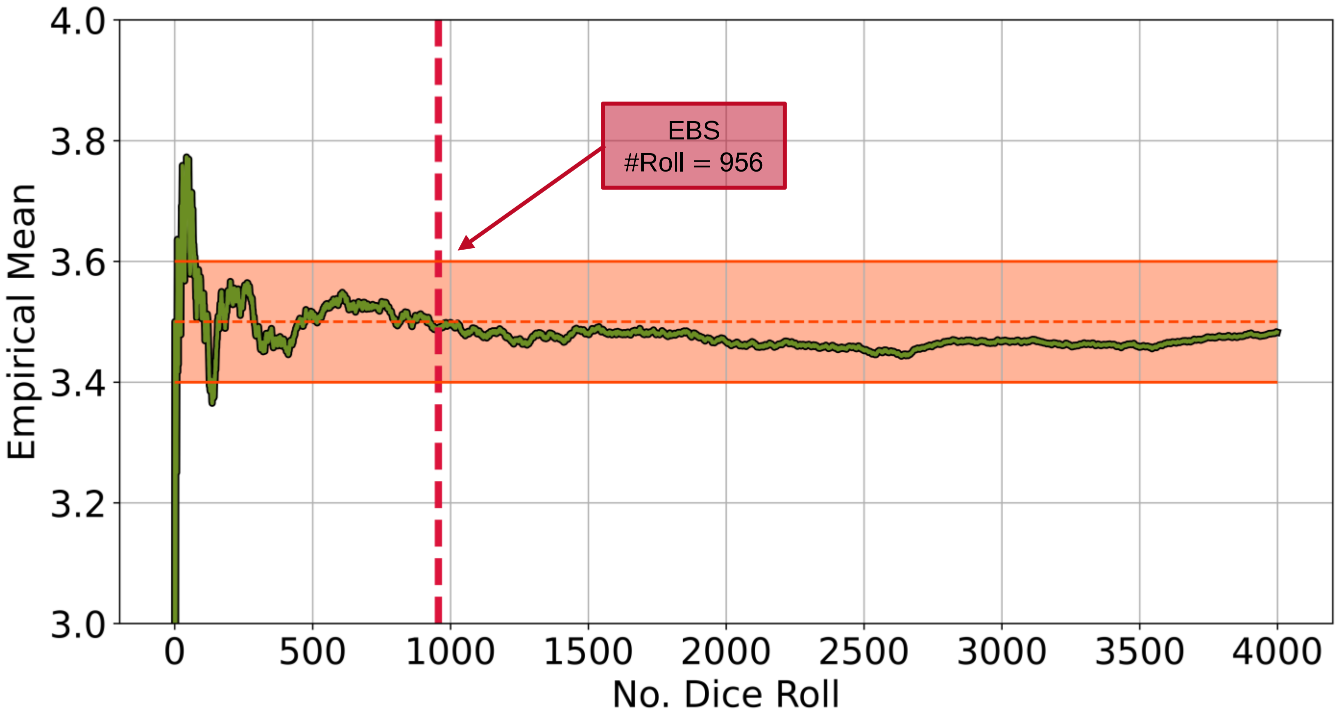


$\epsilon = 0.1$
$\delta = 0.1$
$R = 5$

Tail Bounds

Heinrich Heine Universität Düsseldorf

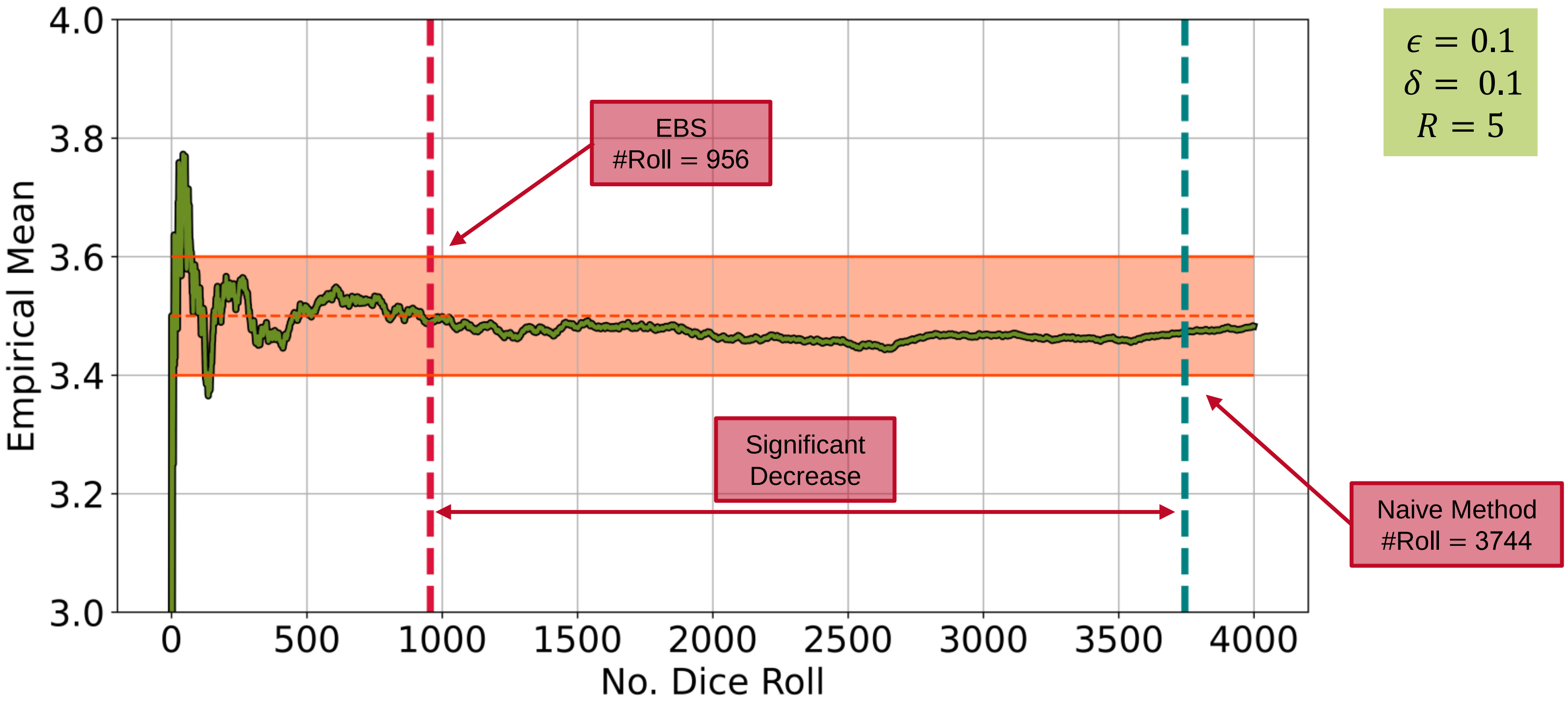
Empirical Mean of a Dice: EBS Algorithm



$\epsilon = 0.1$
$\delta = 0.1$
$R = 5$

Tail Bounds

Empirical Mean of a Dice: EBS Algorithm



$\epsilon = 0.1$
$\delta = 0.1$
$R = 5$

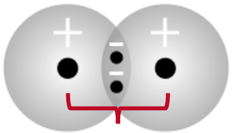
VQAs for solving the Electronic Structure Problem

Electronic Structure Problem

Heinrich Heine Universität Düsseldorf

- $R_i \triangleq$ Nuclei position
- $r_i \triangleq$ Electron position
- $M_i \triangleq$ Nuclei mass
- $m_i \triangleq$ Electron mass
- $Z_i \triangleq$ Nuclei charge
- 1 Hartee $\triangleq$ 27.2 eV

H_2



- Describing whole system

Kinetic Energy | Potenial Energy

$$H = -\sum_i \frac{\nabla^2_{R_i}}{2M_i} - \sum_i \frac{\nabla^2_{r_i}}{2m_i} - \sum_{i,j} \frac{Z_i e^2}{|R_i - r_j|} + \sum_{i,j>i} \frac{Z_i Z_j e^2}{|R_i - R_j|} - \sum_{i,j>i} \frac{e^2}{|r_i - r_j|}$$

Nuclei | Electron | Electron - Nuclei | Nuclei - Nuclei | Electron - Electron

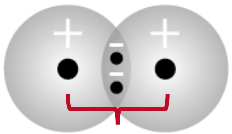
- Born-Oppenheimer approximation + second quantization

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$$

- $a^\dagger \triangleq$ creation operator
- $a \triangleq$ annihilation operator
- p, q, r, s: orbital label

Picture: https://qiskit.org/documentation/stable/0.25/tutorials/chemistry/01_electronic_structure.html

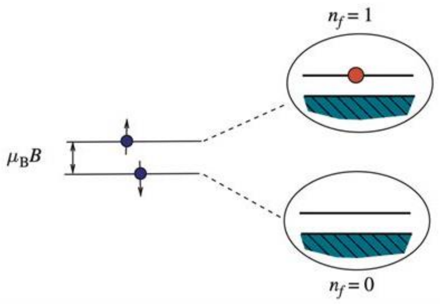
Electronic Structure Problem



$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$$

- $R_i \triangleq$ Nuclei position
- $r_i \triangleq$ Electron position
- $M_i \triangleq$ Nuclei mass
- $m_i \triangleq$ Electron mass
- $Z_i \triangleq$ Nuclei charge
- 1 Hartee $\triangleq$ 27.2 eV

- How to represent this on a quantum computer ?
  - Jordan – Wigner mapping
  - …
- Fermionic operators ⟶ Pauli operators $\sum_i c_i P_i$



- $a^\dagger \triangleq$ creation operator
- $a \triangleq$ annihilation operator
- p, q, r, s: orbital label

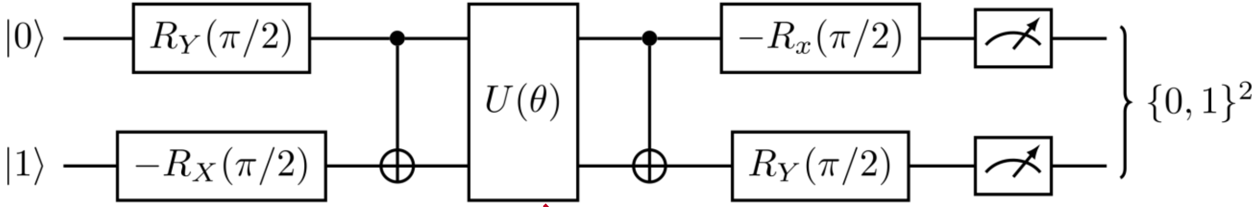
Picture: https://qiskit.org/documentation/stable/0.25/tutorials/chemistry/01_electronic_structure.html

Results:
Total Energy of H_2

Total Energy of H_2

Overview

- $\hat{H}_{H_2}(d) = g_1(d)\mathbb{I} + g_2(d)Z_0 + g_3(d)Z_1 + g_4(d)Z_0 \otimes Z_1 + g_5(d)Y_0 \otimes Y_1 + g_6(d)X_0 \otimes X_1$ [3]
- Guaranteed accuracy ϵ with EBS
- Ansatz for parametrised Circuit: [3]



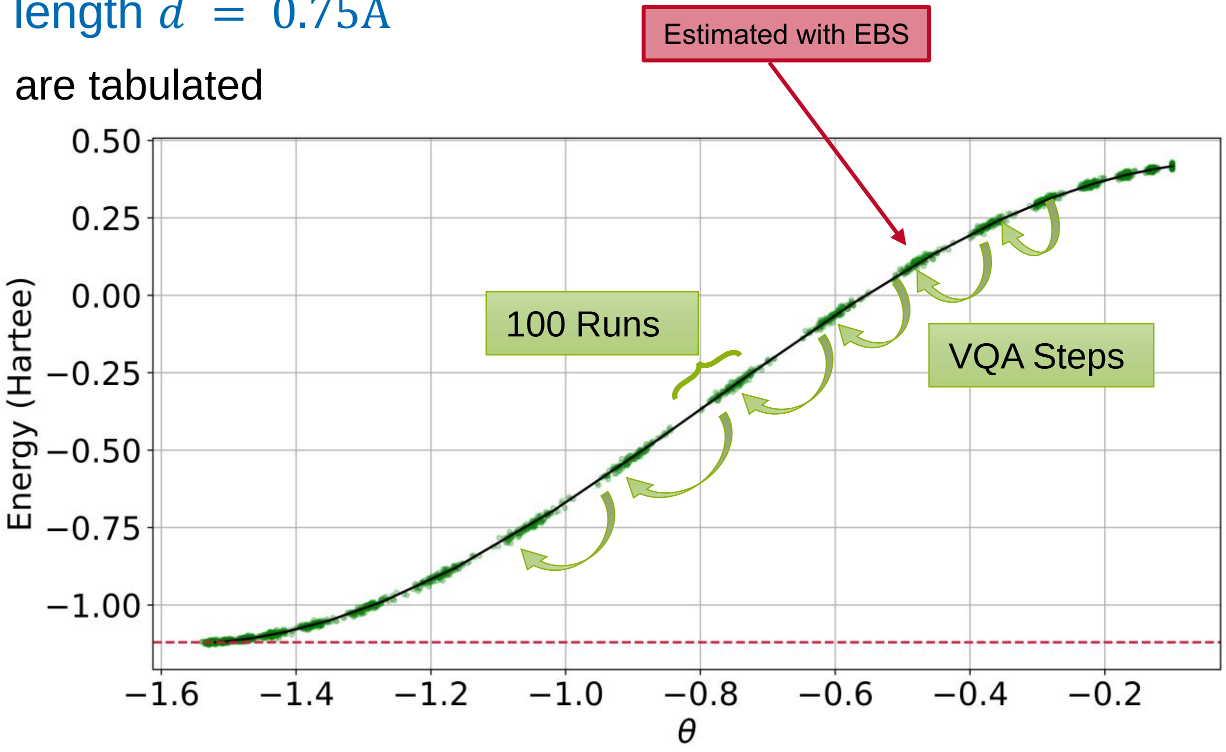
$U(\theta) = e^{-i\theta X_0 \otimes Y_1}$

3. P. J. O'Malley, *Scalable quantum simulation of molecular energies* (2016)

Ground State Energy $H_2(d)$: VQA

Fixed bond length $d = 0.75Å$

- $d \rightarrow \{g_i\}$[3] are tabulated



- $\epsilon = 0.01$
- $\delta = 0.1$
- $R = 3.0636$
- 1 Hartee ≙ $27.2\ eV$

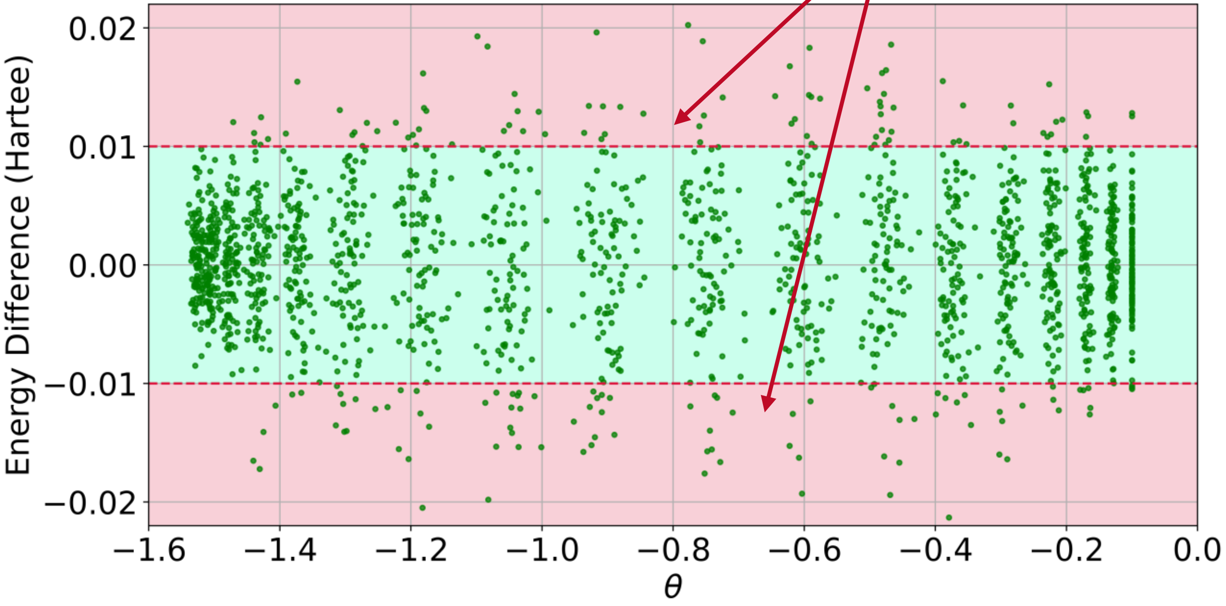
3. P. J. O'Malley, *Scalable quantum simulation of molecular energies* (2016)

Ground State Energy $H_2(d)$: VQA

Fixed bond length $d = 0.75\text{Å}$

- $d \rightarrow \{g_i\}$ [3] are tabulated

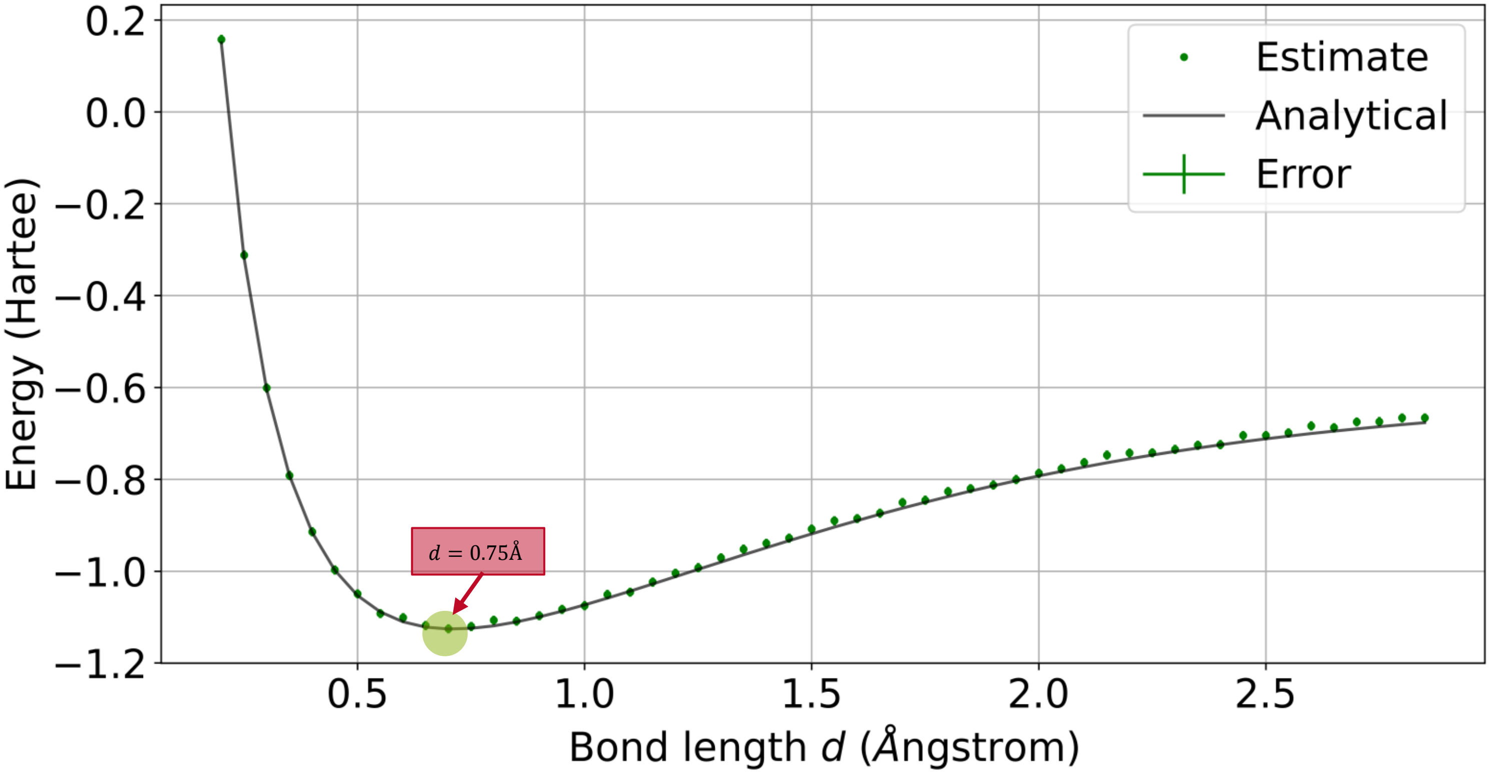
$\delta_{empirical} \approx 0.100$



- $\epsilon = 0.01$
- $\delta = 0.1$
- $R = 3.0636$
- 1 Hartee ≙ $27.2\ eV$

3. P. J. O'Malley, *Scalable quantum simulation of molecular energies* (2016)

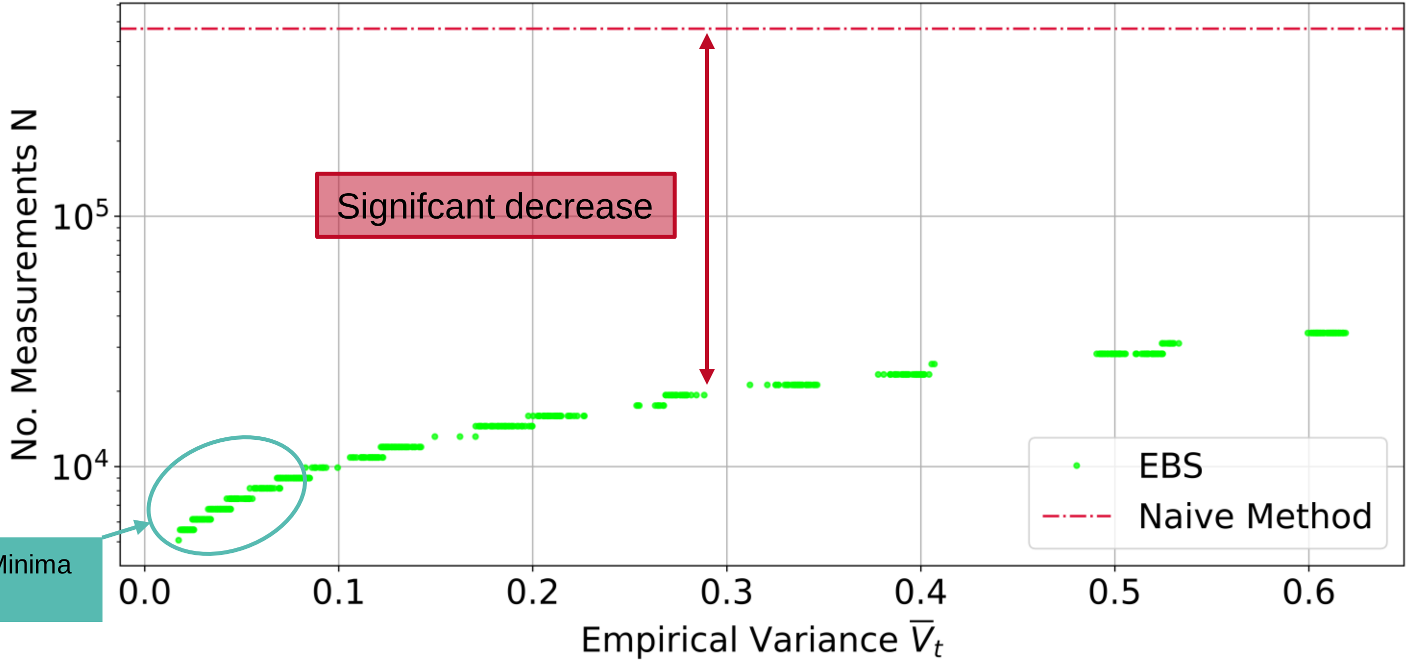
Total Energy of H_2



Reducing the measurement effort

Heinrich Heine Universität Düsseldorf

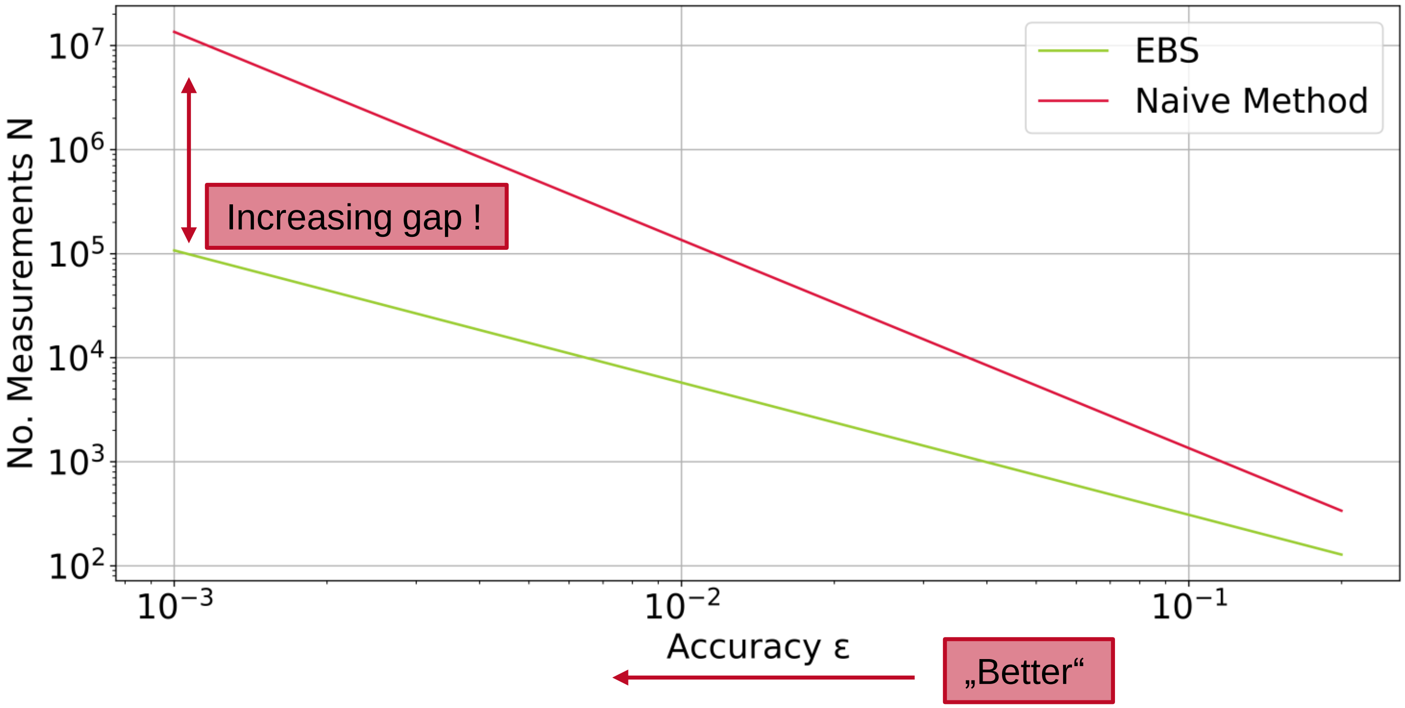
Empirical Variance $\bar{V}_t$ vs Number of Measurements N



Low variance → Minima (local / global)

Reducing the measurement effort

Accuracy ϵ vs Number of Measurements N



Summary & Outlook

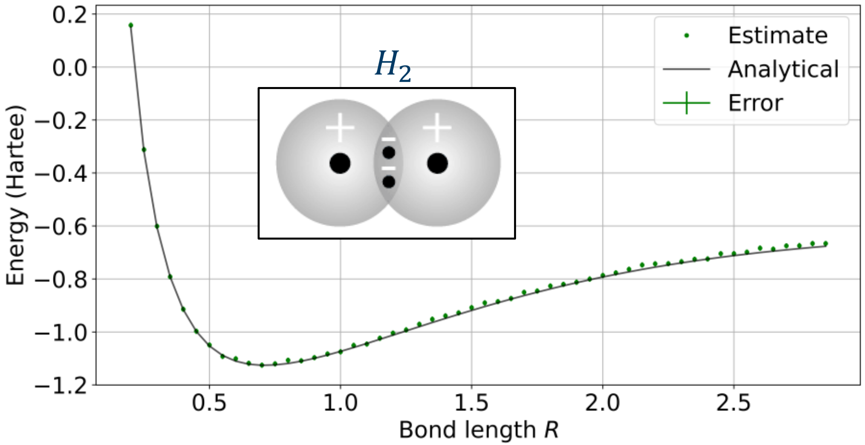
Summary & Outlook

Summary:

- **Significant** reduction of **VQA measurement effort**
- **Guaranteed** confidence $1-\delta$ and accuracy ϵ

However, …

- **Finding** suitable **ansatz** is **hard**→ EBS advantage may not be a pronounced
- $\boldsymbol{H_2}$ is **minimal** example
- Optimal setup → **no noise** simulation
- **Energy** estimator depends on **measurement strategy**



Outlook

- Different applications: e.g., state verifications, etc…
- Test on **real** quantum hardware
- **Improvements** to EBS: Adapt **beyond real-valued random variables**

Sources

1. JY. Audibert et. al., *Tuning bandit algorithms in stochastic environments* (2007)
2. V. Mnih*, Efficient stopping rules* (2008)
3. P. J. O'Malley*, Scalable quantum simulation of molecular energies* (2016)
4. J. Preskill, *Quantum computing in the NISQ era and beyond* (2018)
5. A. Peruzzo, *A variational eigenvalue solver on a photonic quantum processor* (2018)
6. J. K. Blitzstein and J. Hwang*, Introduction to probability* (2015)
7. M. Kliesch, *Characterization, certification, and validation of quantum systems* (2020)
8. M. A. Nielsen, *Quantum computation and quantum information* (2010)
9. Smite-Meister, *Bloch sphere, a geometrical representation of a two-level quantum system* (2009)
10. L. Zhu, *Optimizing shot assignment in variational quantum eigensolver measurement* (2023)
11. Molecular orbit H2 https://commons.wikimedia.org/wiki/File:H2OrbitalsAnimation.gif 25.09.2023 (20:15)
12. Feynman picture https://www.britannica.com/biography/Richard-Feynman 25.09.2023 (20:15)
13. H2 molecule https://byjus.com/chemistry/hydrogen-gas/ 25.09.2023 (20:15)

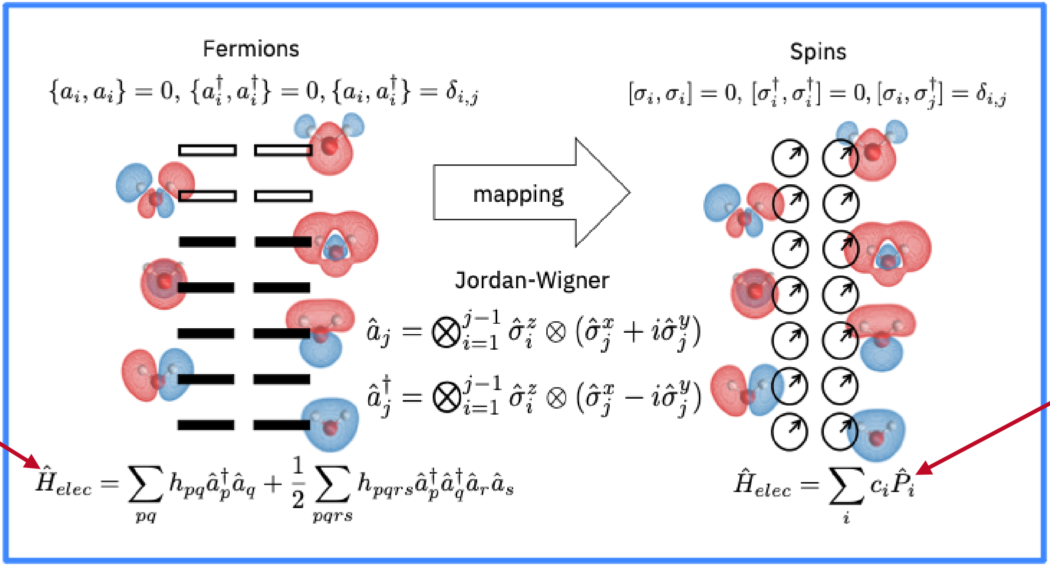
Appendix:
Jordan – Wigner Mapping

Electronic Structure Problem

Jordan-Wigner Mapping

- 1 electron orbital + spin $\longrightarrow$ 2 qubits

Sum of fermionic operators



Sum of Pauli operators

Picture: https://qiskit.org/ecosystem/nature/tutorials/06_qubit_mappers.html

Measurement

Energy curve of H_2

How to measure?

- $\hat{H}_{H_2}(R) = g_1(R)\mathbb{I} + g_2(R)Z_0 + g_3(R)Z_1 + g_4(R)Z_0 \otimes Z_1 + g_5(R)Y_0 \otimes Y_1 + g_6(R)X_0 \otimes X_1$

$\hat{H}_i$	Measurement Basis	$E(R) + g_0 E_0$
Z_0	Z	$g_2 E_2$
Z_1	Z	$g_3 E_3$
$Z_0 \otimes Z_1$	Z	$g_4 E_4$
$Y_0 \otimes Y_1$	Y	$g_5 E_5$
$X_0 \otimes X_1$	X	$g_6 E_6$

Energy curve of H_2

How to measure?

- $\hat{H}_{H_2}(R) = g_1(R)\mathbb{I} + g_2(R)Z_0 + g_3(R)Z_1 + g_4(R)Z_0 \otimes Z_1 + g_5(R)Y_0 \otimes Y_1 + g_6(R)X_0 \otimes X_1$

$\hat{H}_i$	Measurement Basis	$E(R)$ $+g_0E_0$
Z_0	Z	g_2E_2
Z_1		g_3E_3
$Z_0 \otimes Z_1$		g_4E_4
$Y_0 \otimes Y_1$	Y	g_5E_5
$X_0 \otimes X_1$	X	g_6E_6

- 3 x measurement $\longrightarrow$ estimate
- General optimal strategy **unclear**

Quantum Computing

Physical Qubit Implementations

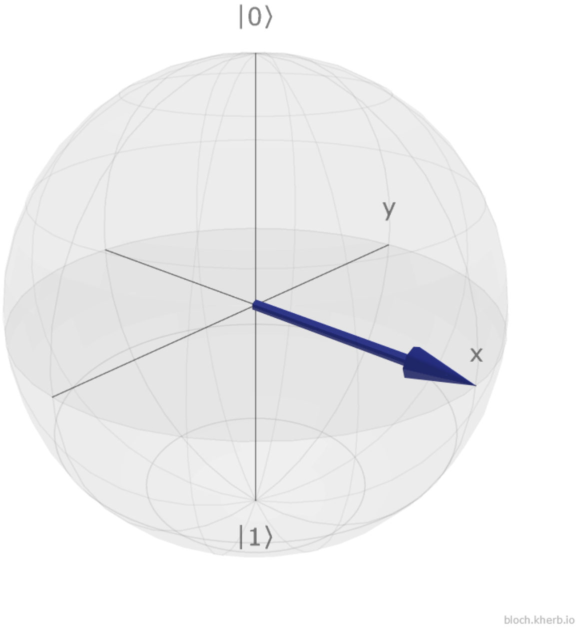
Physical Support	Information	$\|0\rangle$	$\|1\rangle$
Photons	Polarization	$\leftrightarrow$	$\updownarrow$
	Number	Vacuum	Single Photon
Electron	Spin	$\uparrow$	$\downarrow$
	Number	No Electron	One Electron
…	…	…	…

Heinrich Heine Universität Düsseldorf

Quantum Gates

Quantum

$$|0\rangle \text{---} \boxed{H} \text{---} \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

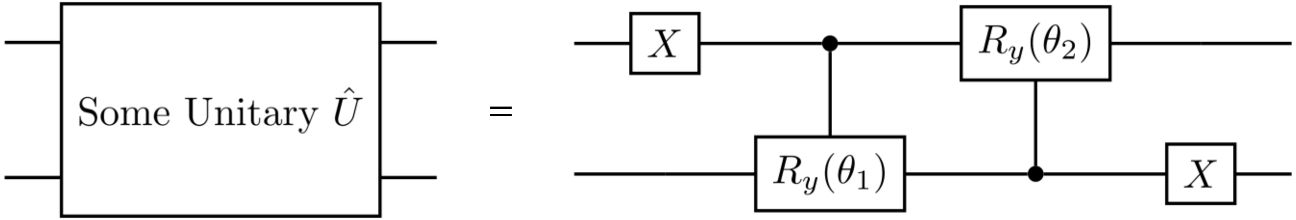


- $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}$
- $|1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$
- Hadamard -gate= $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$

$$H|0\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$
$$H|1\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

Universality of Gates

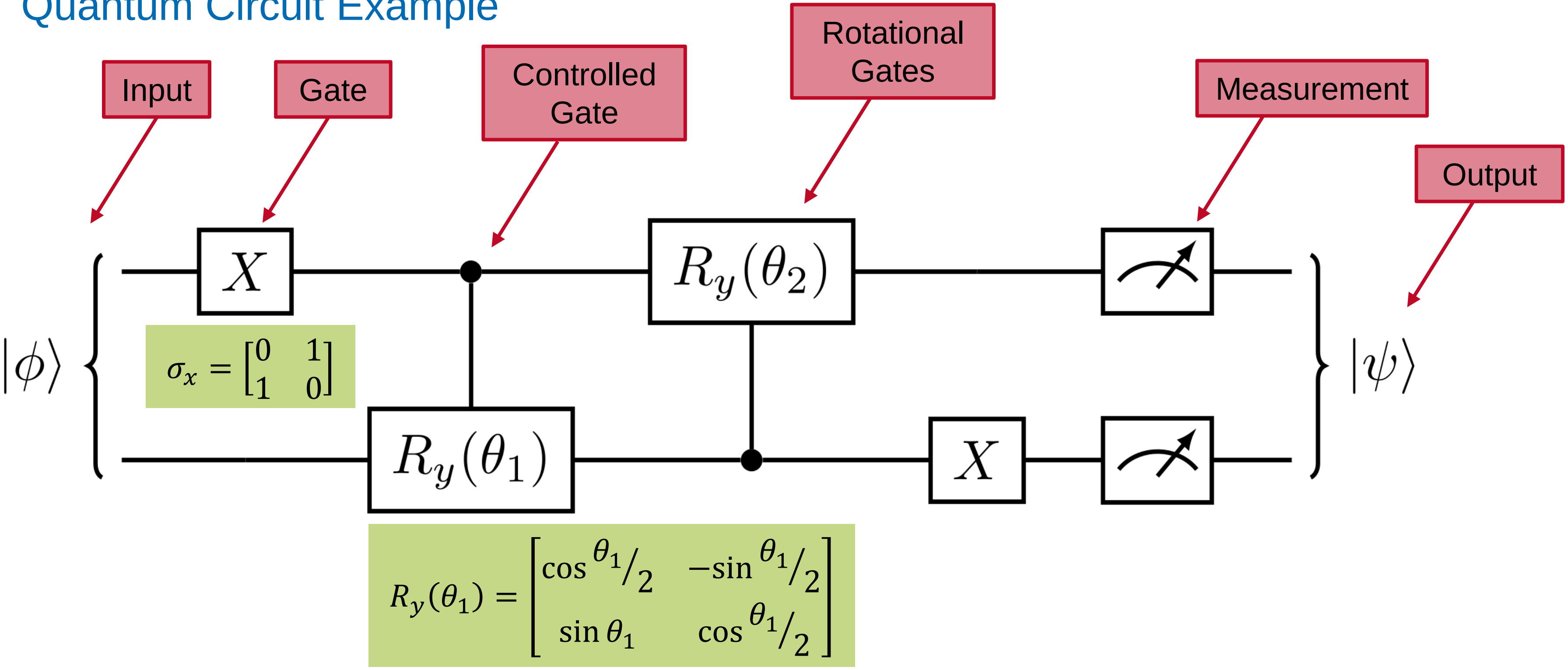
- Certain sets of gates are universal
  - Every unitary operation can be build using these
- Any gate can be decomposed using these
  - $\{\mathrm{CNOT}, \text{all single qubit gates}\}$[1]
  - $\{\mathrm{CNOT}, H, T\}$[2]
  - …



1. A. Barenco, et al., Phys. Rev. A 52, 3457 (1995) 2. P. O. Boykin, et al. Inf. Proc. Lett. 75, 101 (2000)

Quantum Computing

Quantum Circuit Example



$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$R_y(\theta_1) = \begin{bmatrix} \cos {}^{\theta_1}/_2 & -\sin {}^{\theta_1}/_2 \\ \sin \theta_1 & \cos {}^{\theta_1}/_2 \end{bmatrix}$$

Tail Bounds

(ϵ, δ)-Stopping Algorithm

- A stopping algorithm terminates when condition is met, i.e.:

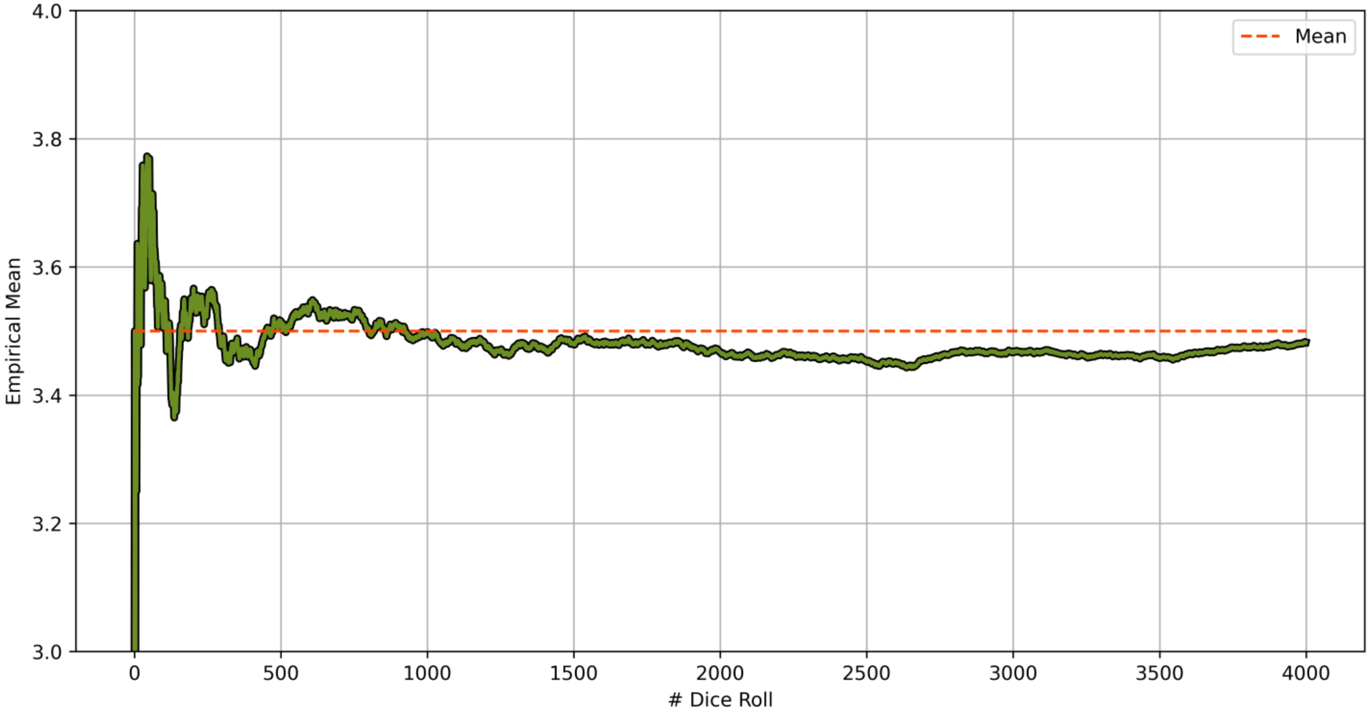
$$\mathbb{P}[|\hat{\mu} - \mu| \leq \epsilon] \geq 1 - \delta$$

- Estimate $\hat{\mu}$ is called an absolute (ϵ, δ)-estimate

• $\hat{\mu} \triangleq$ Estimate
• $\mu \triangleq$ Expected value
• $\epsilon \triangleq$ Error margin
• $1 - \delta \triangleq$ Confidence

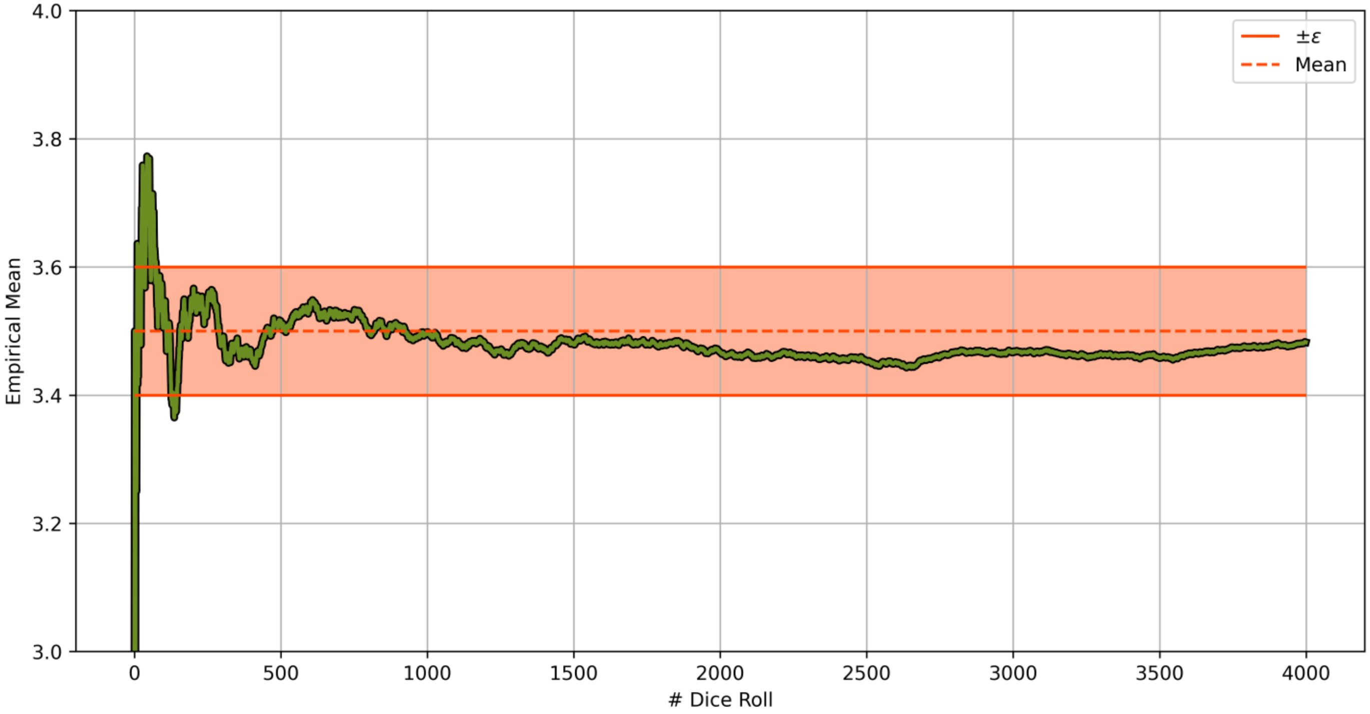
Why Tail Bounds

Empirical Mean of a Dice



Why Tail Bounds

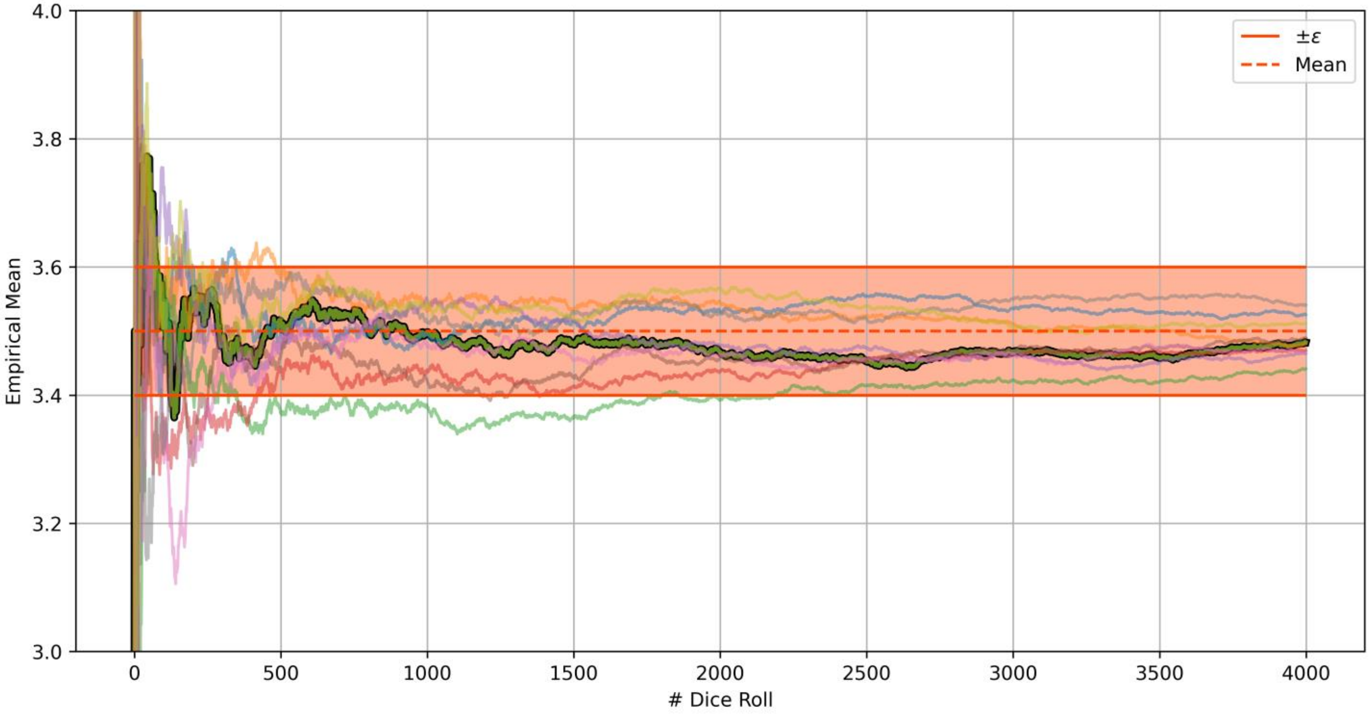
Empirical Mean of a Dice



$\epsilon := 0.1$

Why Tail Bounds

Empirical Mean of a Dice



$\epsilon := 0.1$

Tail Bounds

Heinrich Heine Universität Düsseldorf

Höffding‘s Inequality

$$\mathbb{P}[|\overline{X_t} - \mu| \geq \epsilon] \leq 2exp\left[-\frac{2t\epsilon^2}{R^2}\right] := \delta$$

- Bounded form

$$|\bar{X}_t - \mu| \leq R\sqrt{\frac{\ln(2/\delta)}{2t}} := \epsilon$$

- Solve for t gives minimally required number of samples t_{min}

$$t_{min} = 2\left(\frac{R}{\epsilon}\right)^2 \ln(2/\delta)$$

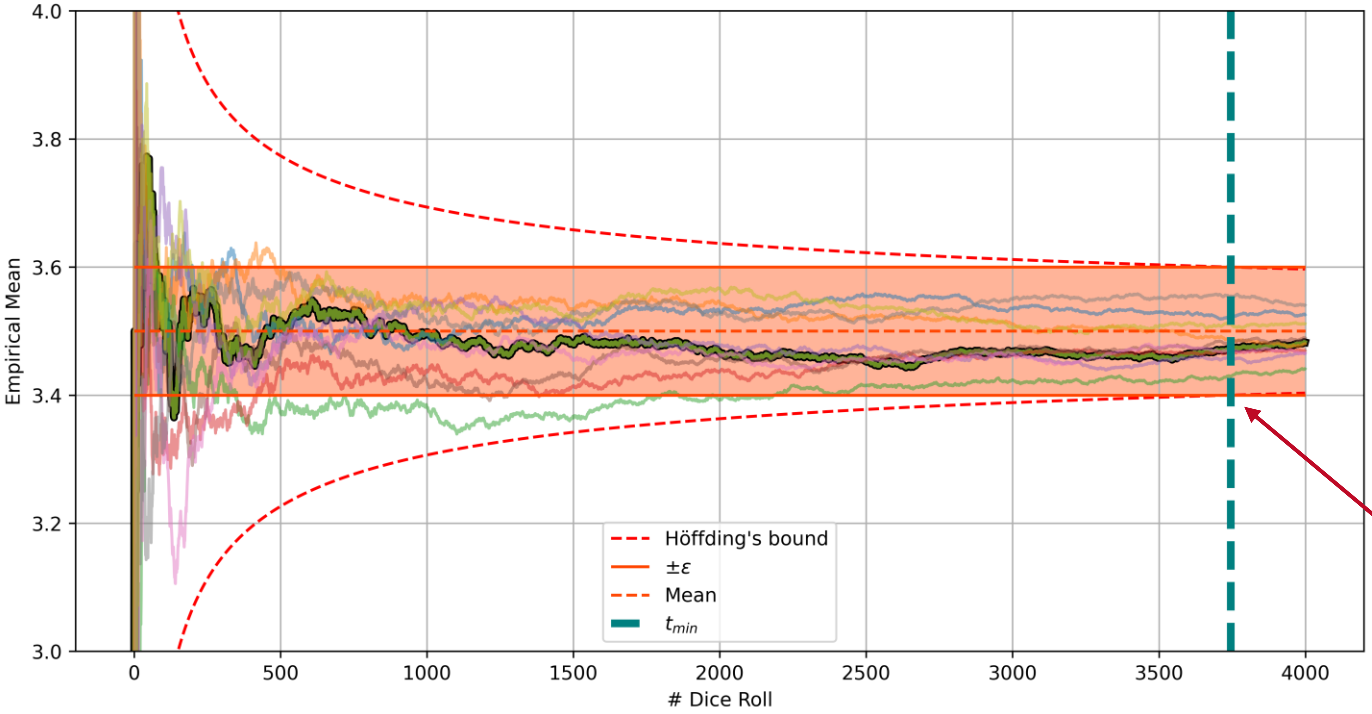
Independent and identically distributed random variable

- $X_1 \ldots X_t \triangleq$ i.i.d random variables
- $a \leq X_i \leq b$
- $\epsilon \in R^+$
- $R \triangleq$ Range

- $\bar{X}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$

Tail Bounds

Empirical Mean of a Dice: Höffing's Bound



$\epsilon := 0.1$
$\delta := 0.1$
$R := 5$

#Roll = 3744

Tail Bounds

Heinrich Heine Universität Düsseldorf

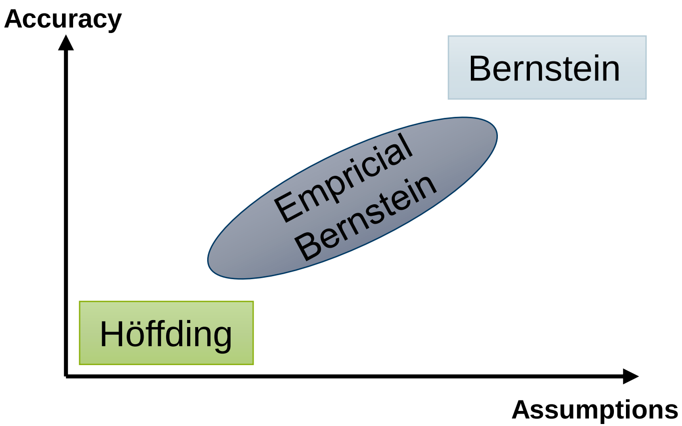
Bernstein‘s bound

- $|\bar{X}_t - \mu| \leq \underbrace{\sqrt{\frac{2\,\Sigma^2 \ln(2/\delta)}{t}}}_{\Theta\left(1/\sqrt{t}\right)} + \underbrace{\frac{R\ln(2/\delta)}{3t}}_{\Theta\left(1/t\right)}$

Höffding‘s bound

- $|\bar{X}_t - \mu| \leq R\underbrace{\sqrt{\frac{\ln(2/\delta)}{2t}}}_{\Theta\left(1/\sqrt{t}\right)}$

For $\Sigma < R \longrightarrow$ Bernstein “better”
For $\Sigma > R \longrightarrow$ Höffding “better”

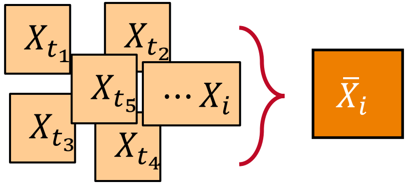


Empirical Bernstein stopping
Improvements

Empirical Bernstein Stopping

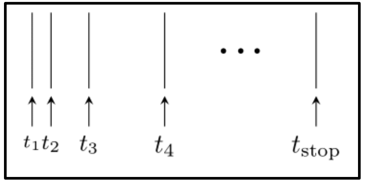
Pseudo Code Implementation: Improvements

- Batch the samples into one sample



- $X_1 \dots X_t \triangleq$ i.i.d random variables
- $\epsilon \in R^+$
- $1 - \delta \triangleq$ Confidence

- Update $c(t)$ with a growing gap



- $R \triangleq$ Range
- $\bar{X}_t = \frac{1}{z} \sum_{i=1}^{t} x_i$
- $c(t) = \sqrt{\frac{2\bar{V}_t \ln(3/\delta_t)}{t}} + \frac{3R \ln(3/\delta_t)}{t}$