



arXiv:2012.07905

Freie Universität Berlin

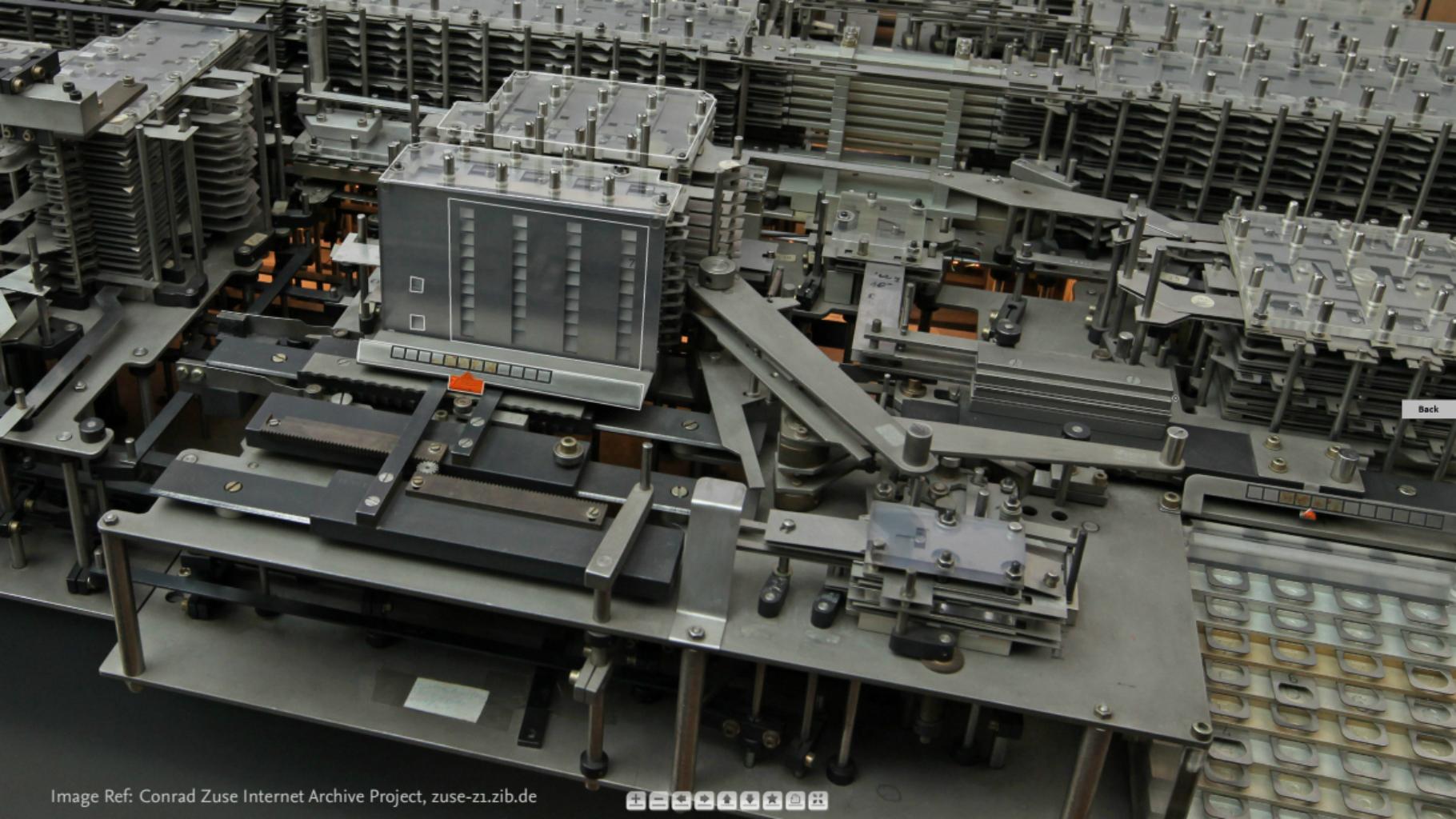


# The quantum sign problem

Dominik Hangleiter

IQC-QuICS Seminar, February 16, 2021

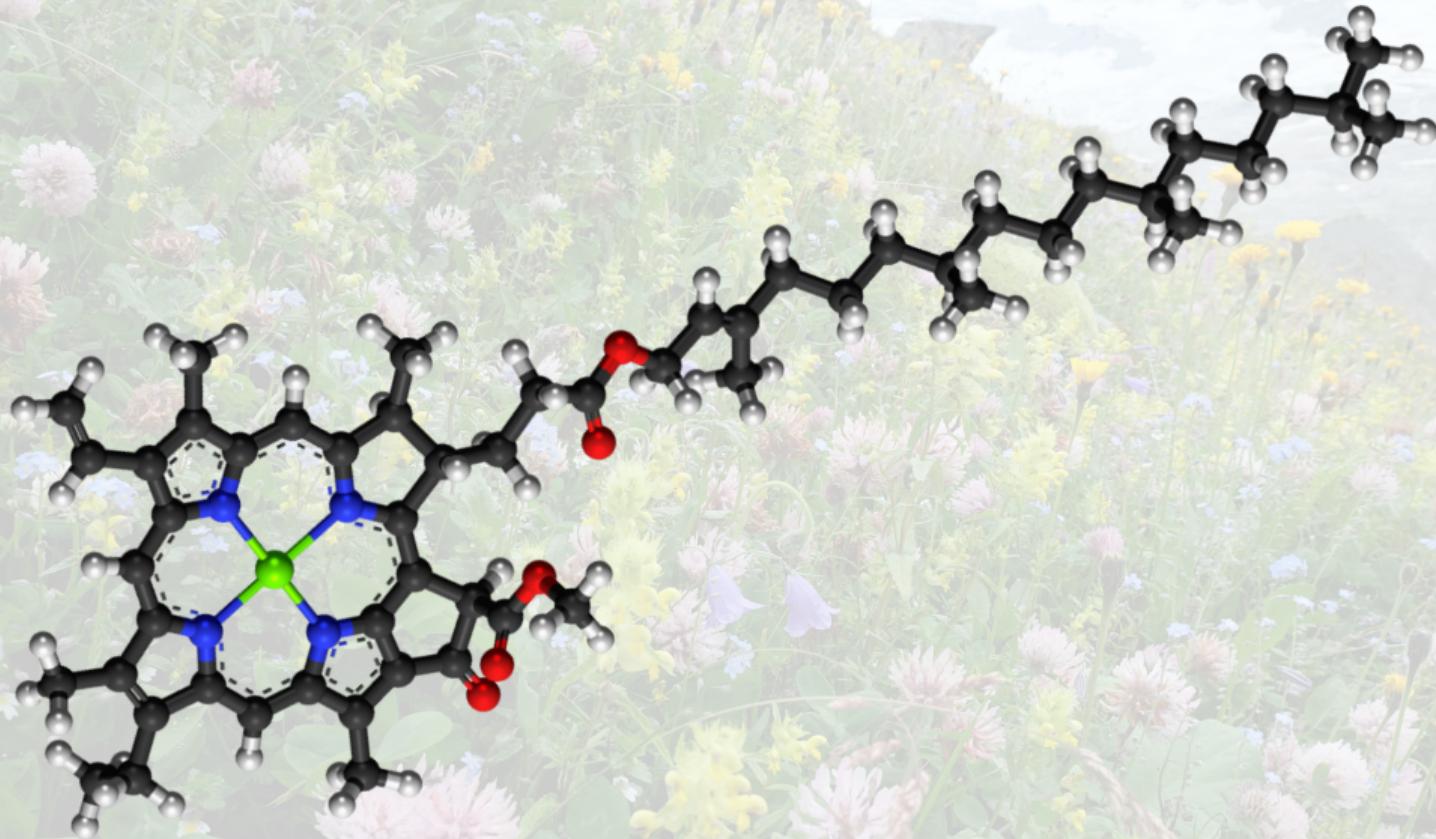
Licensed under CC BY-NC 4.0.





```
circuitOptimizer.py
917         orth = circuit.coreOrths[k]
918         newCircuit = copy.deepcopy(circuit)
919
920         P = polynom_order
921         # get period and steps
922         o = np.max(np.abs(np.linalg.eigvalsh(stepDir)))
923         t = 2*np.pi/fct_order/o
924         mu = np.arange(P+1)*t/P
925         # get translations
926         r_s = sla.expm(-mu[1]*stepDir)
927         R = [np.identity(orth.shape[0], dtype=np.complex128)]
928         for i in range(1,P+1):
929             R.append(np.dot(R[-1],r_s))
930         # get derivatives
931         def _derivative(R):
932             newCircuit.coreOrths[k] = R.dot(orth)
933             return -2*np.real(np.trace(self.calculateEuclideanGradient(newCircuit, k).dot(orth.T.conj()).dot(R.T.conj()).dot(stepDir.T.conj())))
934         dervs = list(map(_derivative,R))
935         mu_mat_inv = np.linalg.inv(mu[1:,np.newaxis]**np.arange(1,P+1))
936         coef = mu_mat_inv.dot(dervs[1:]*dervs[0])
937         coef = list(coef[::-1])
938         coef.append(dervs[0])
939         roots = np.roots(coef)
940         pos_real_roots = [np.real(r) for r in roots if np.abs(np.imag(r))<1E-10 and np.real(r)>=0]
941
942         if showPlot is True:
943             gridSize = 200
944             normPortion = .5
945             orth0 = circuit.coreOrths[k]
946             newCircuit = copy.deepcopy(circuit)
947             gradient = stepDir
948
949             riemGradNormSqr = .5*np.trace(np.dot(gradient.T,gradient))
950             stepSize = normPortion/riemGradNormSqr
951
952             xgrid = np.arange(0,gridSize)*stepSize
953             objval = np.zeros((gridSize))
954
955             for kk in range(0,len(xgrid)):
956                 theOrth = np.dot(sla.expm(-xgrid[kk]*gradient), orth0)
957                 newCircuit.coreOrths[0] = theOrth
958                 objval[kk] = self.objectiveFunction(newCircuit)
```





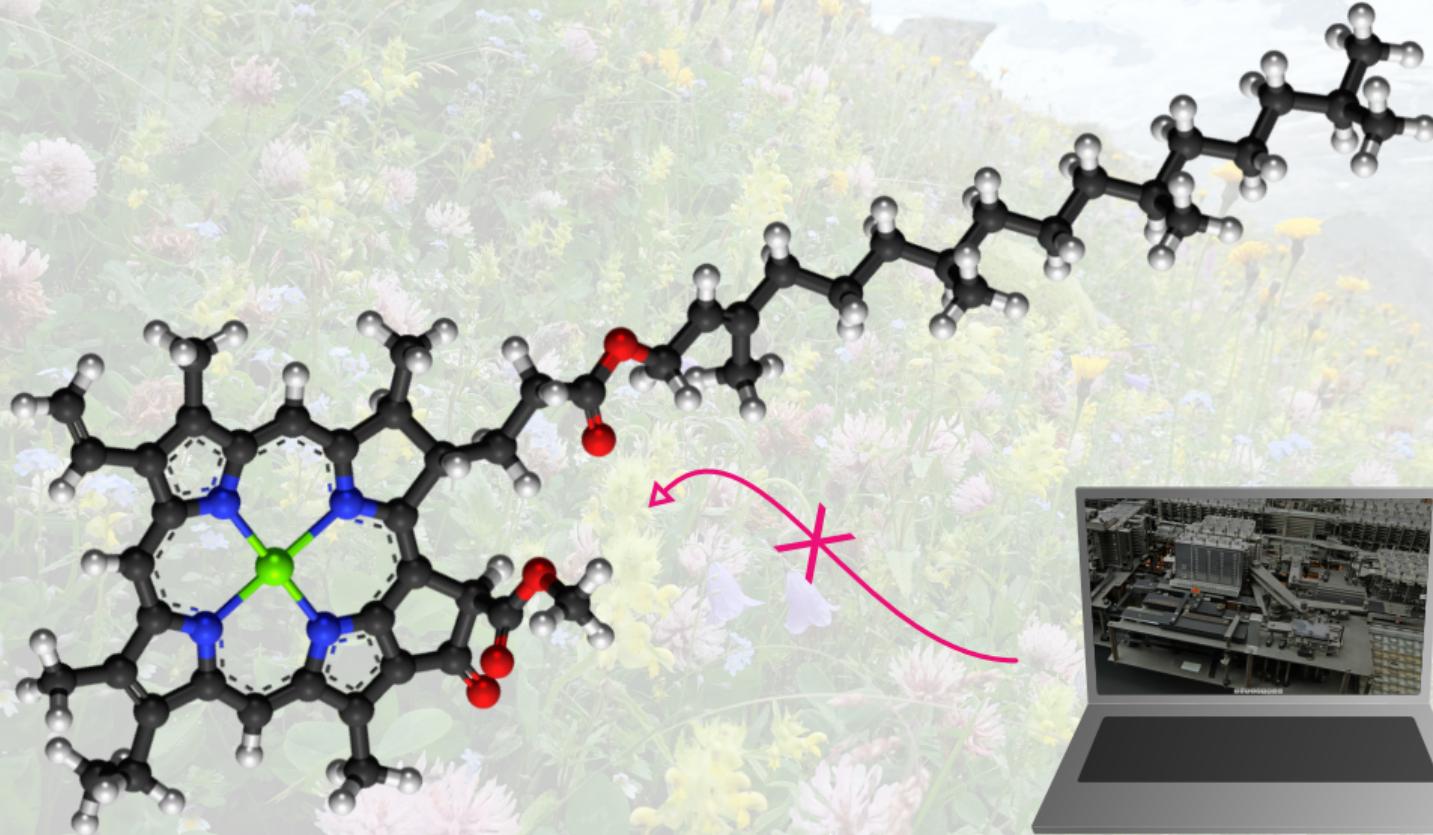
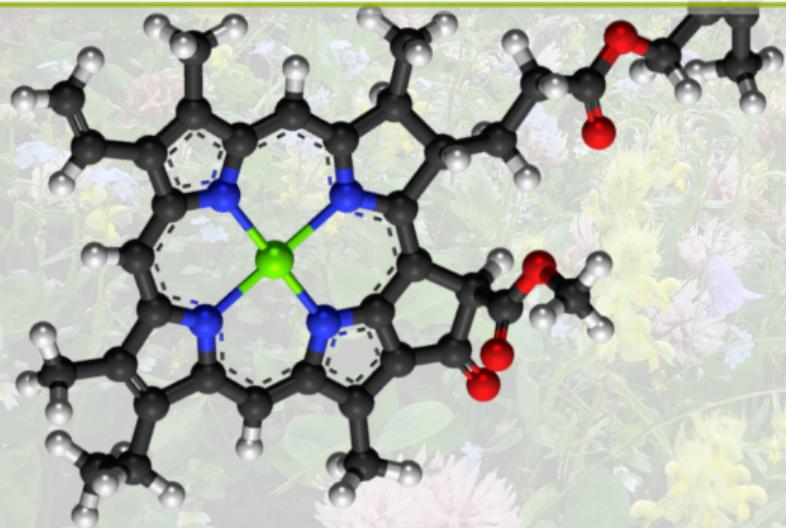
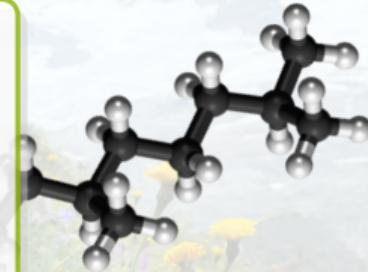


Image Ref.: Jynto, Wikipedia: Chlorophyll & Graham Carlow/IBM

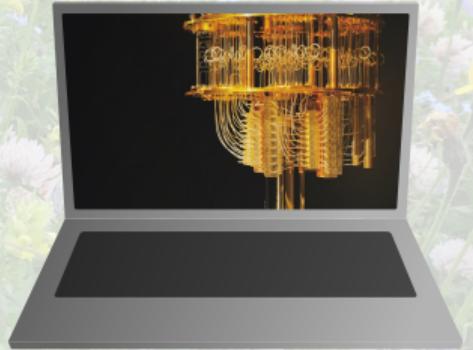
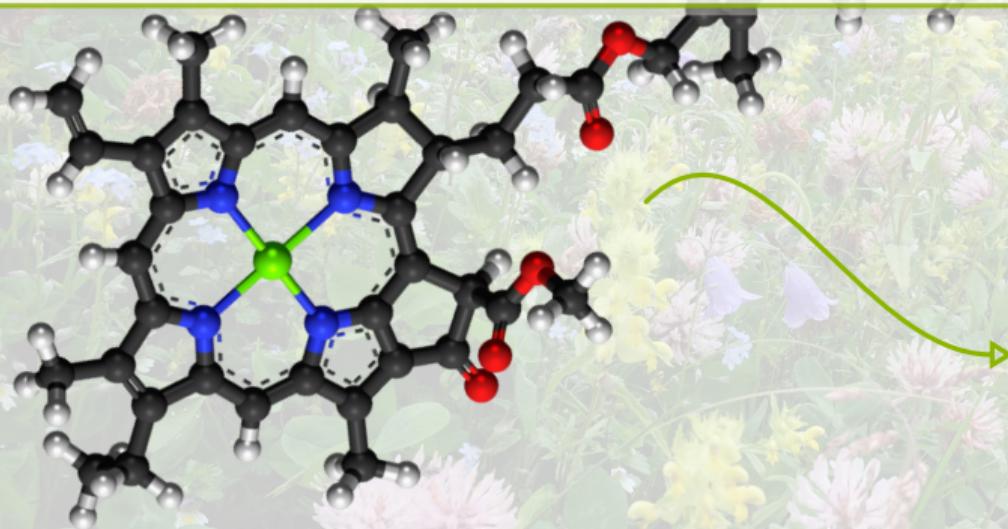
*“We can give up on our rule about what the computer was, we can say: Let the computer itself be built of quantum mechanical elements which obey quantum mechanical laws.”*

— Richard Feynman



*“We can give up on our rule about what the computer was, we can say: Let the computer itself be built of quantum mechanical elements which obey quantum mechanical laws.”*

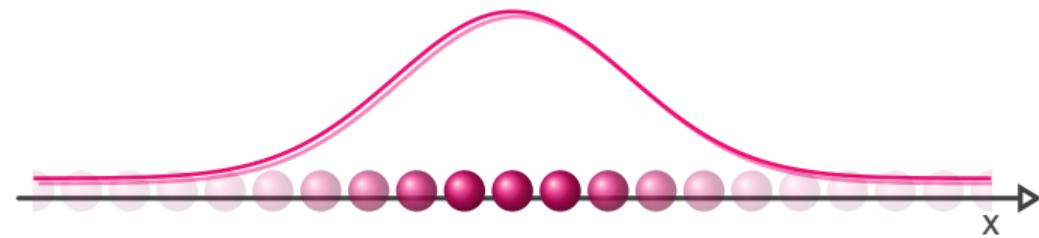
— Richard Feynman



# Quantum measurement

Complex-valued wave function:

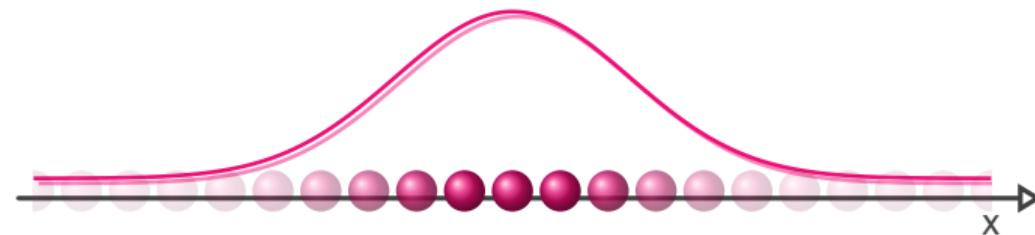
$$\psi : \mathbb{R} \rightarrow \mathbb{C}$$



# Quantum measurement

Complex-valued wave function:

$$\psi : \mathbb{R} \rightarrow \mathbb{C}$$



Wave-function collapse:

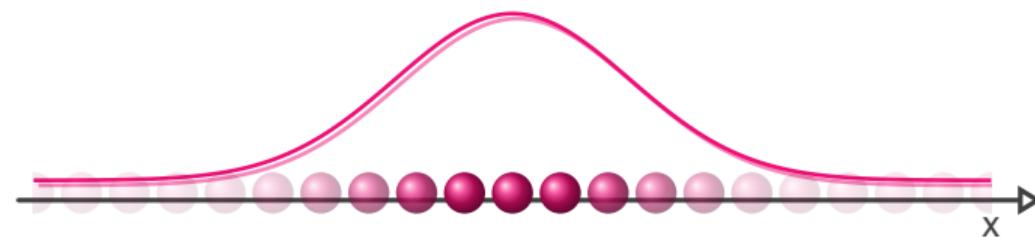
$$\Pr[\text{measured at } x] = |\psi(x)|^2$$



# Quantum measurement

Complex-valued wave function:

$$\psi : \mathbb{R} \rightarrow \mathbb{C}$$



Wave-function collapse:

$$\Pr[\text{measured at } x] = |\psi(x)|^2$$



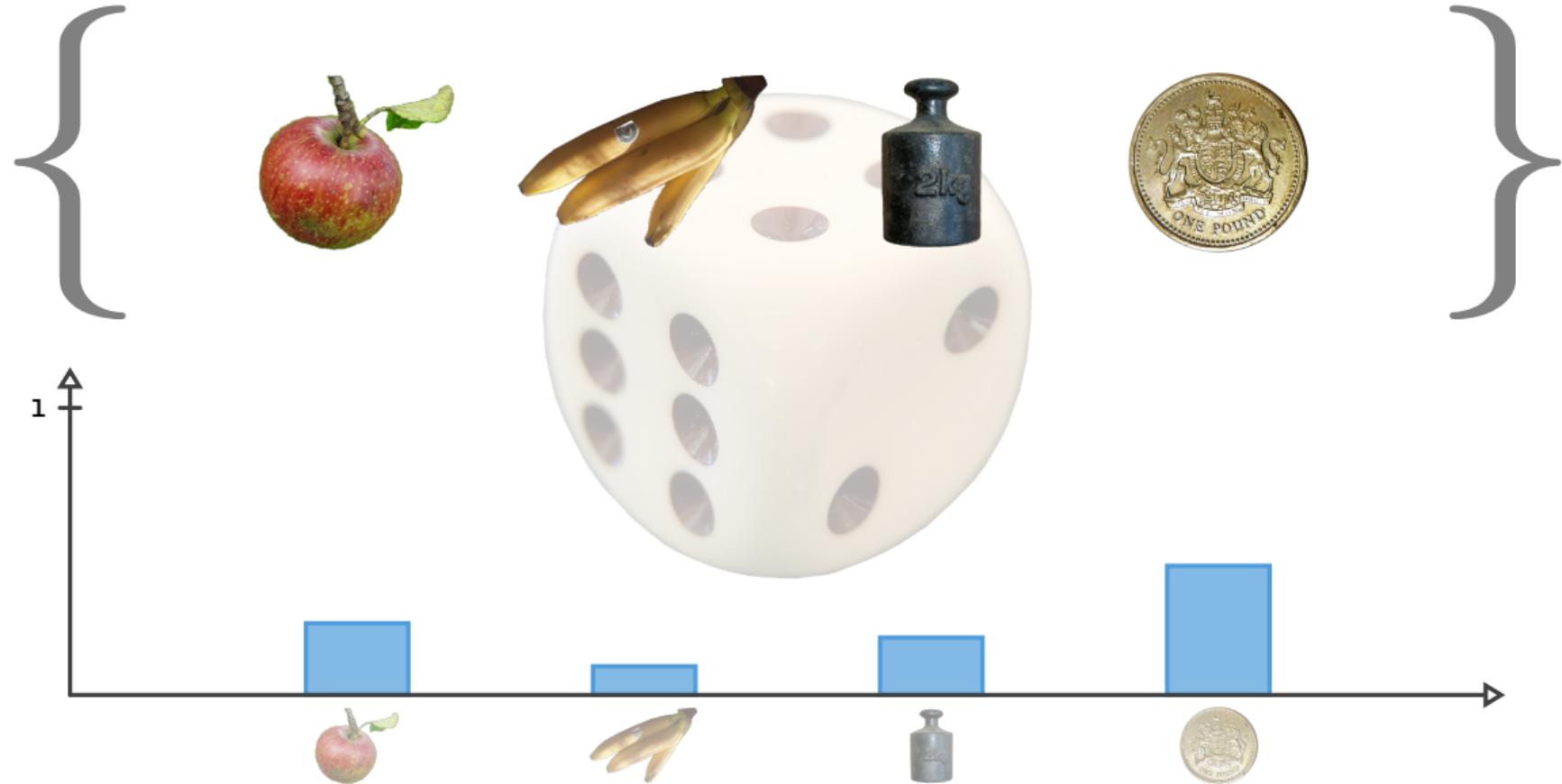
Probability distribution over possible outcomes .....

Sampling

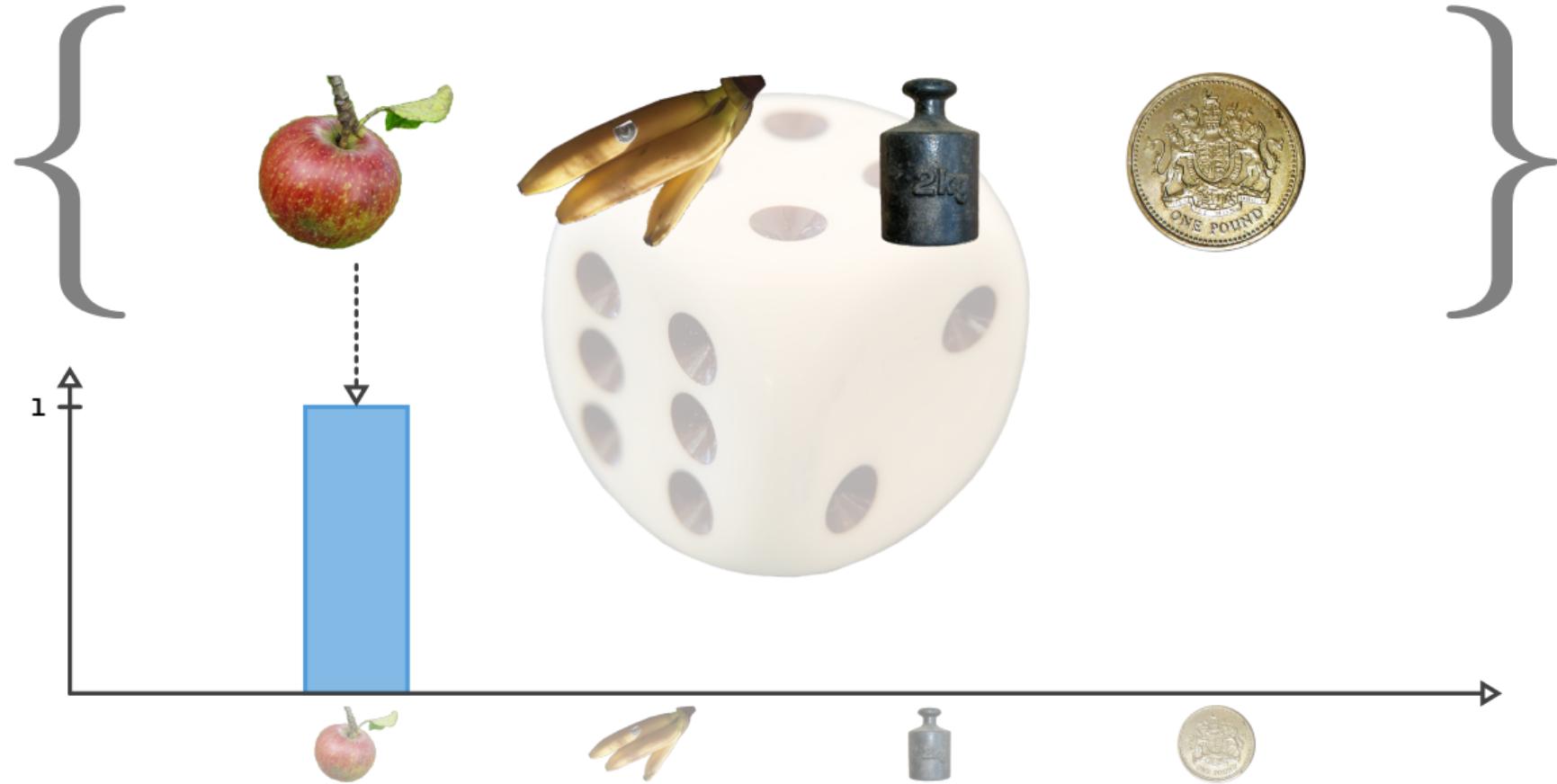
# Sampling



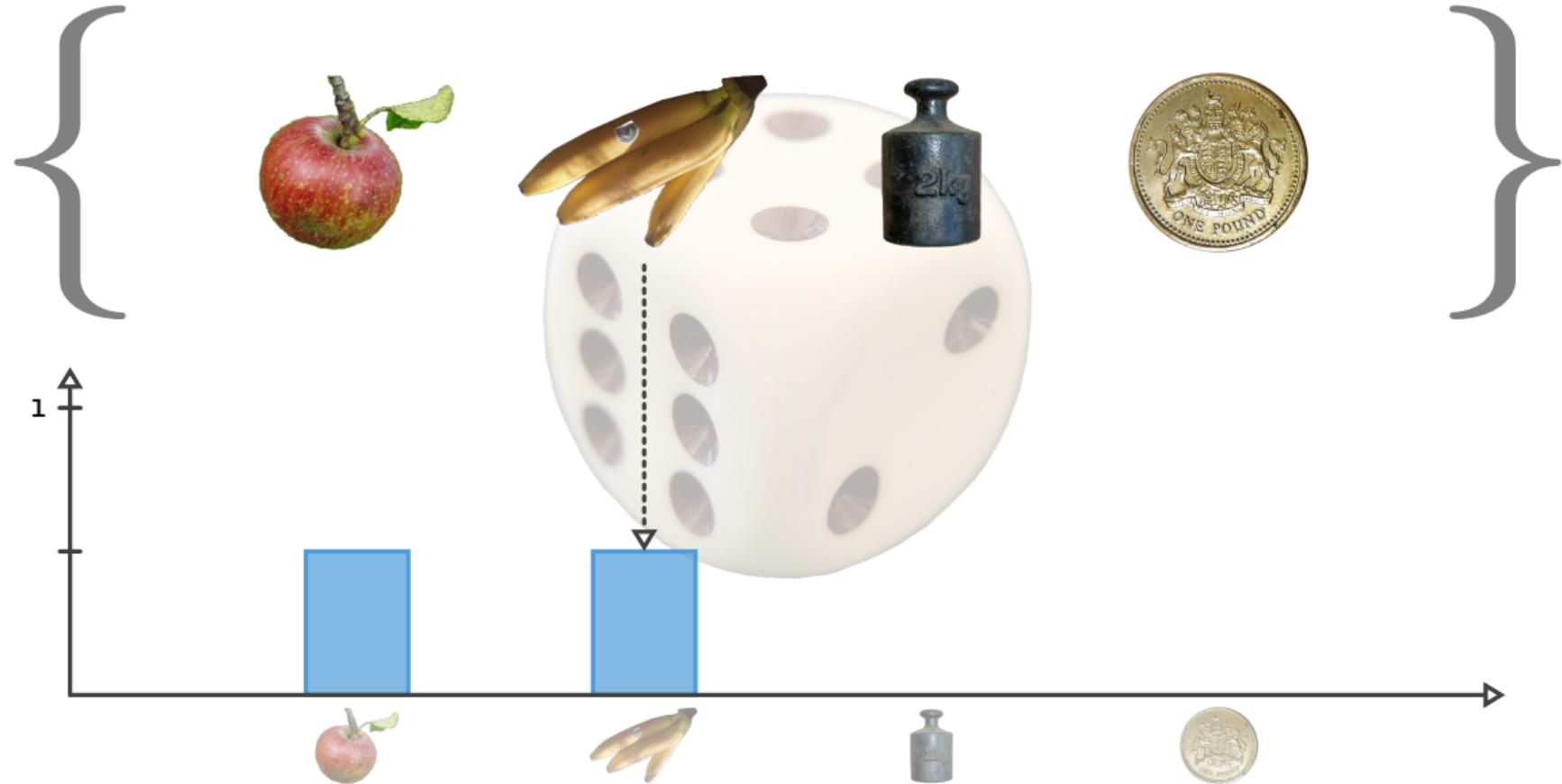
# Sampling



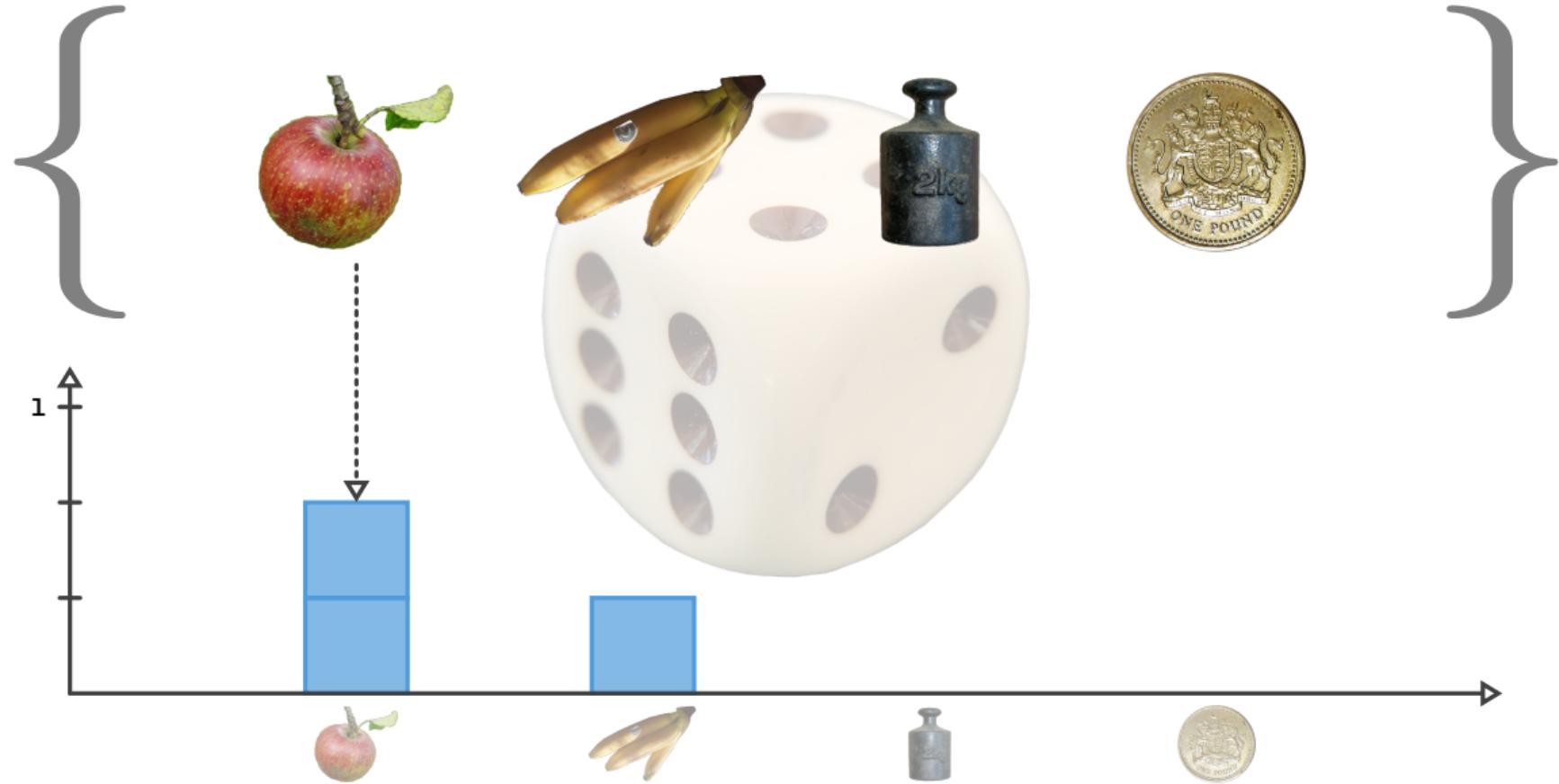
# Sampling



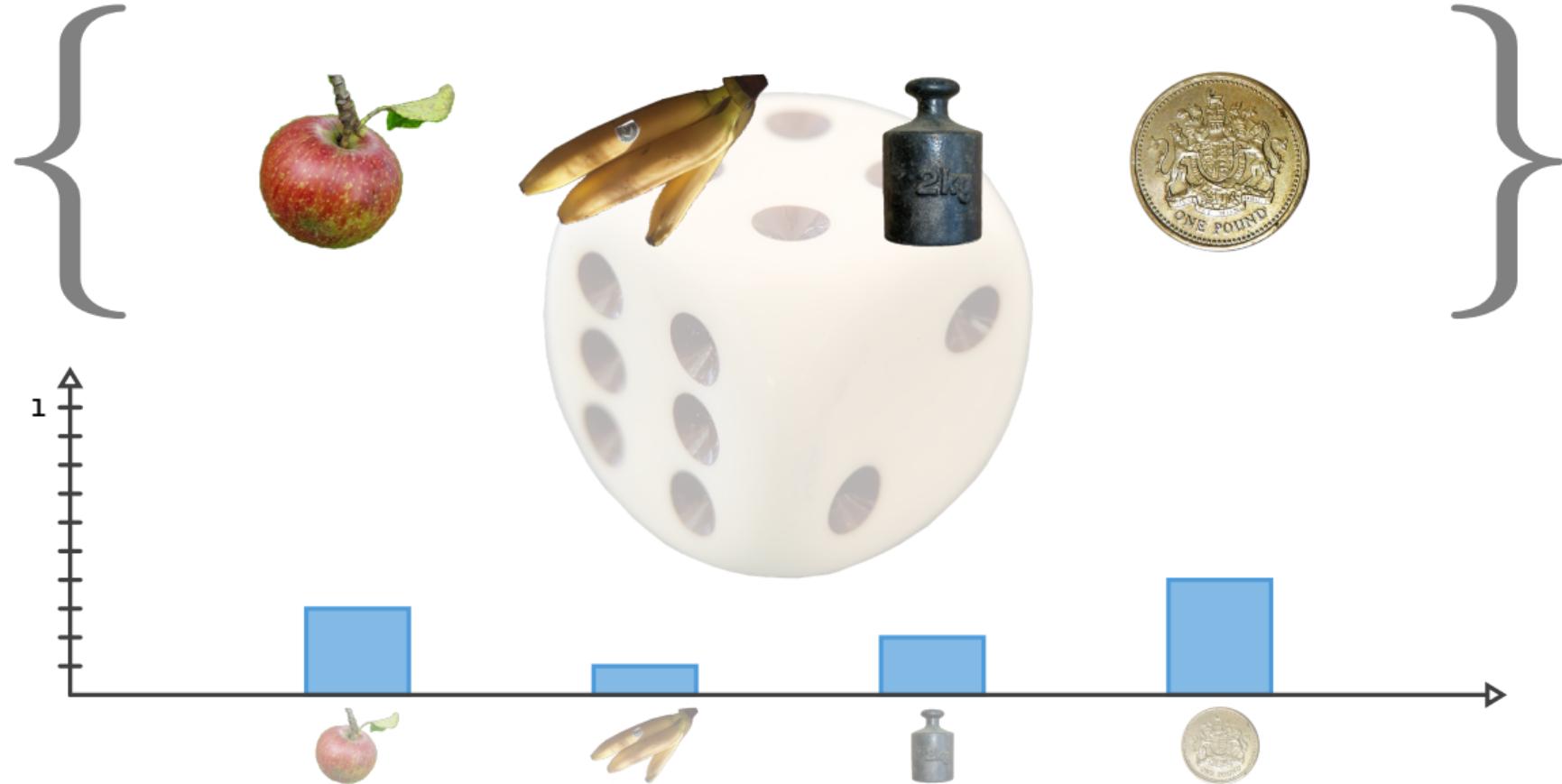
# Sampling



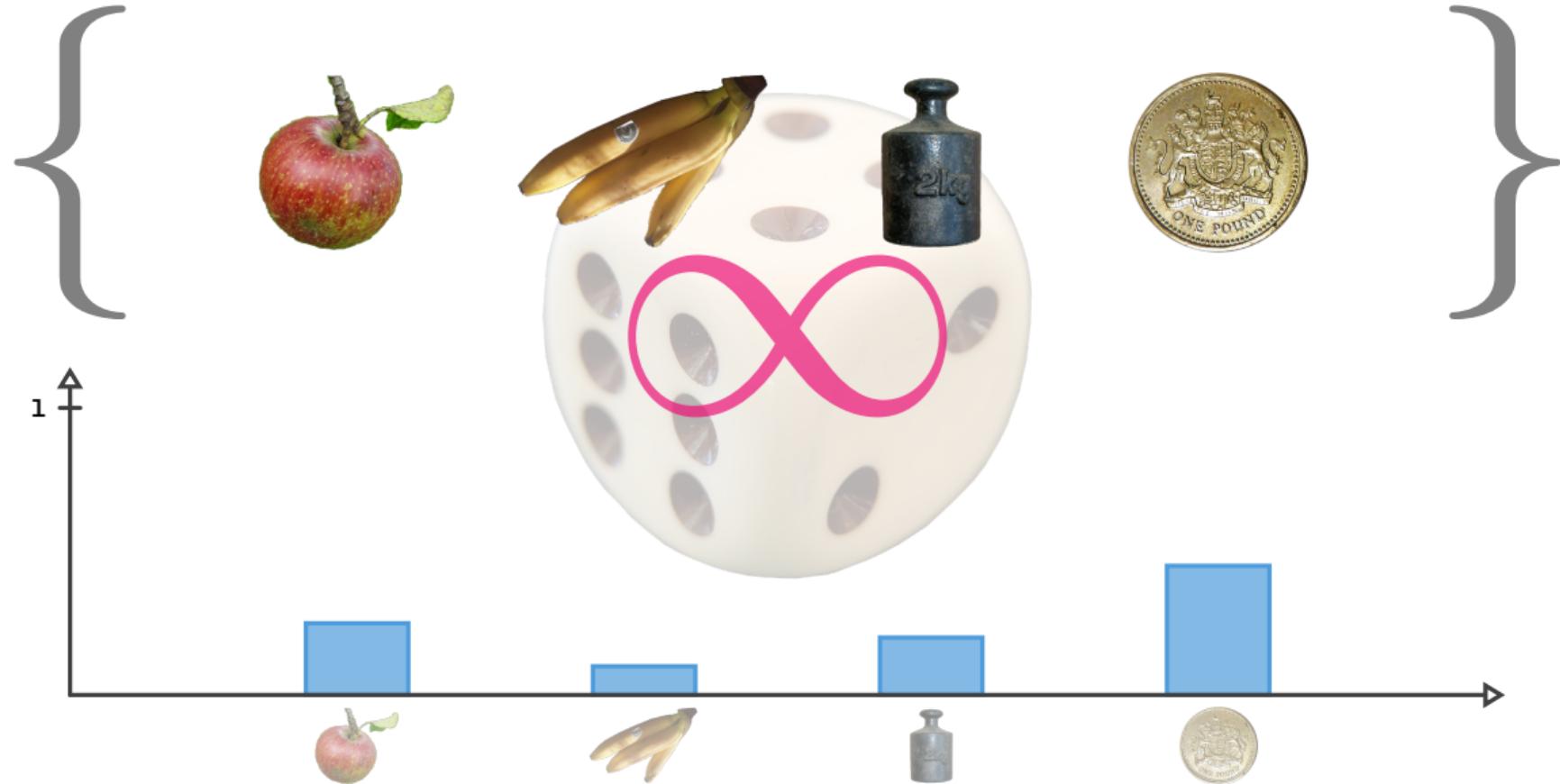
# Sampling



# Sampling



# Sampling



## Computational complexity: scaling of runtime



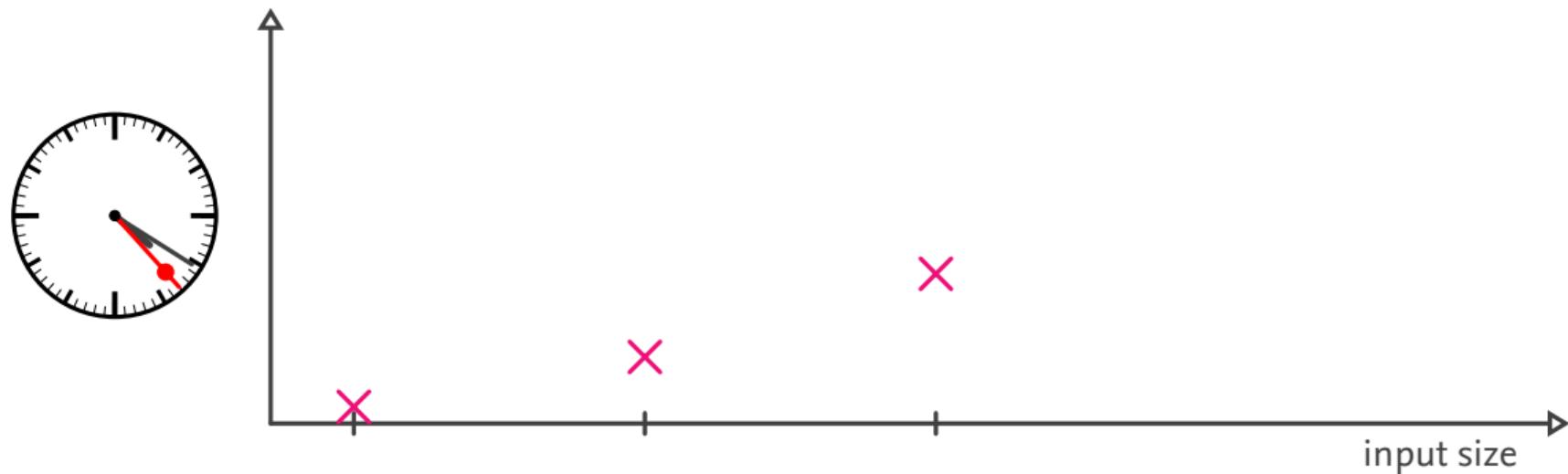
# Computational complexity: scaling of runtime



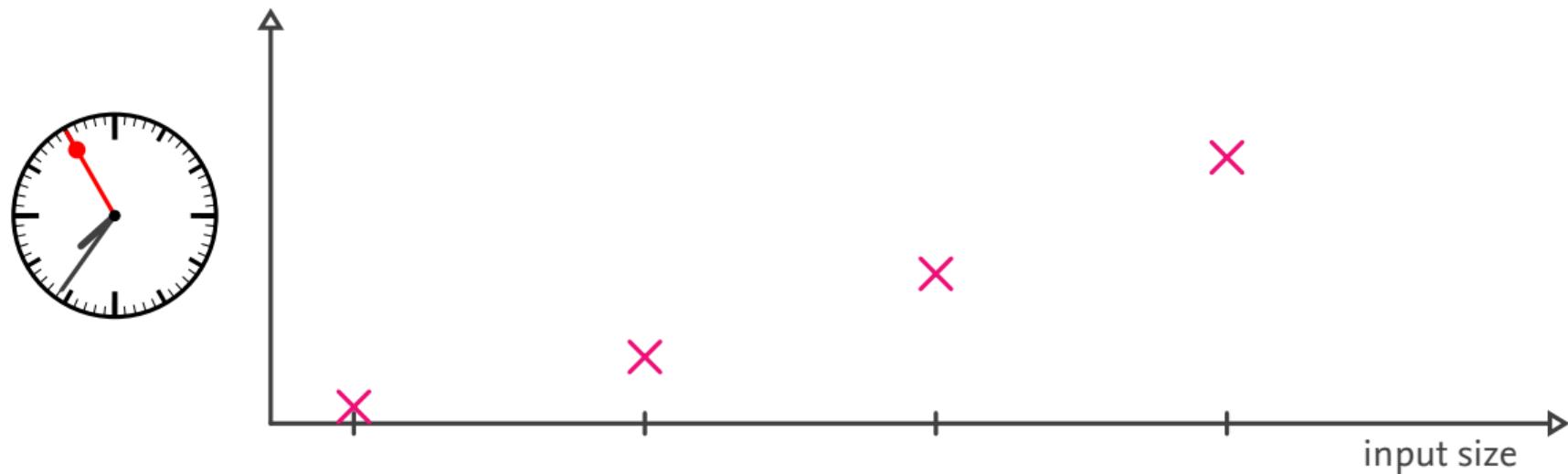
# Computational complexity: scaling of runtime



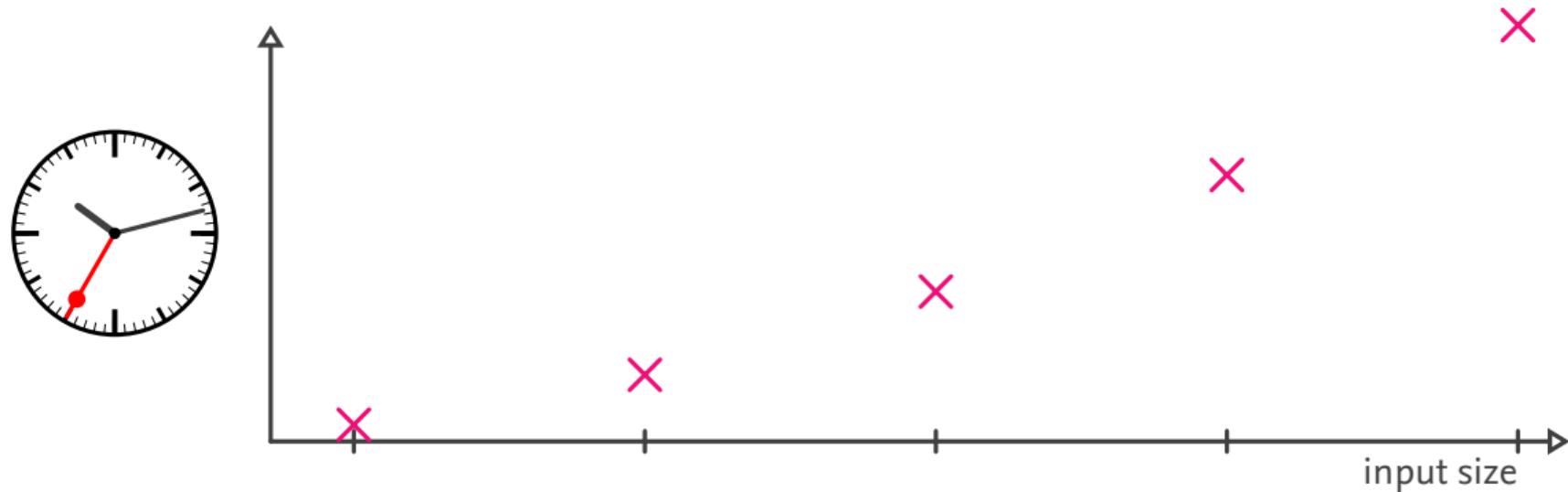
# Computational complexity: scaling of runtime



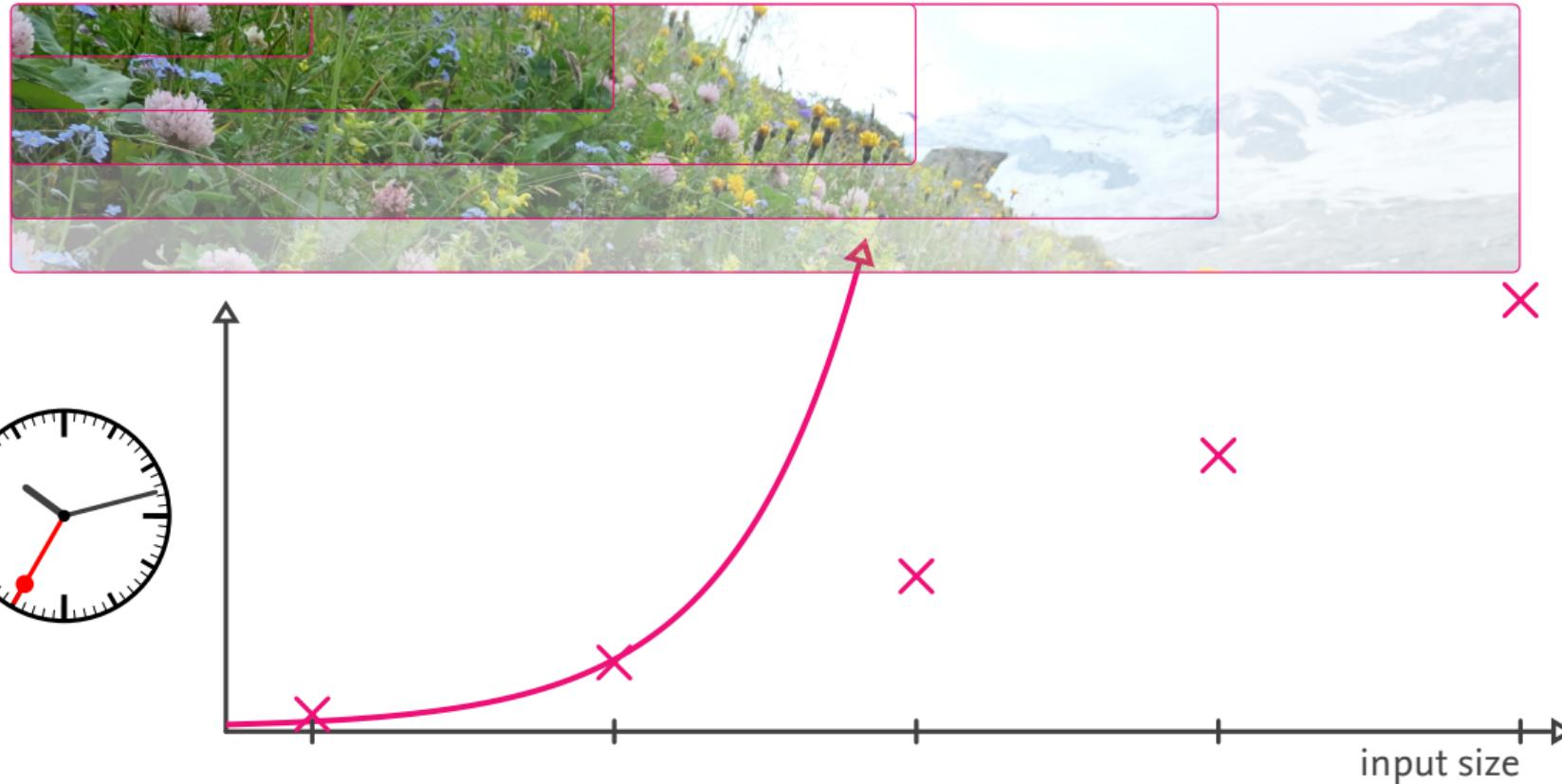
# Computational complexity: scaling of runtime



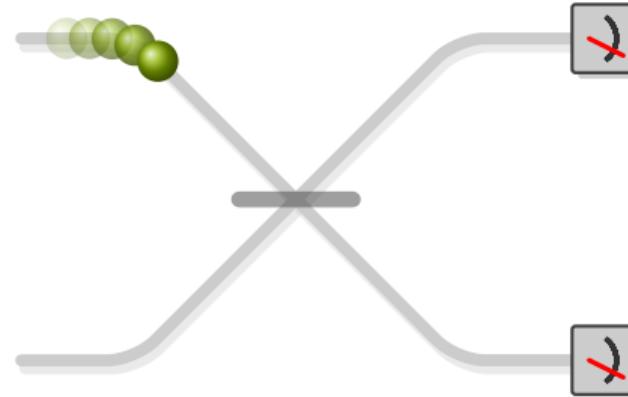
# Computational complexity: scaling of runtime



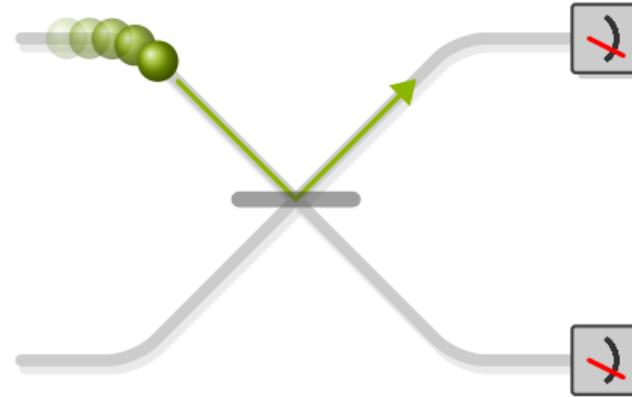
## Computational complexity: scaling of runtime



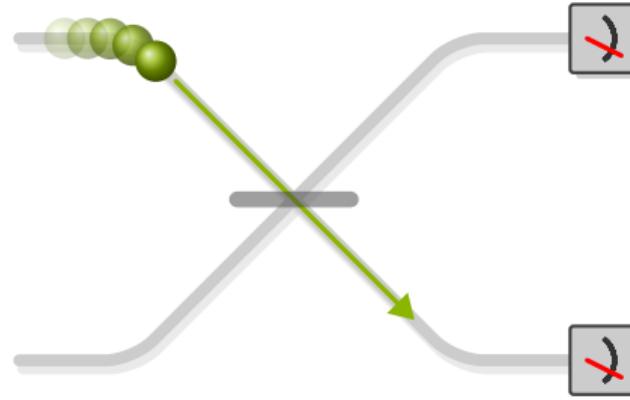
## Quantum sampling: The Hong-Ou-Mandel interferometer



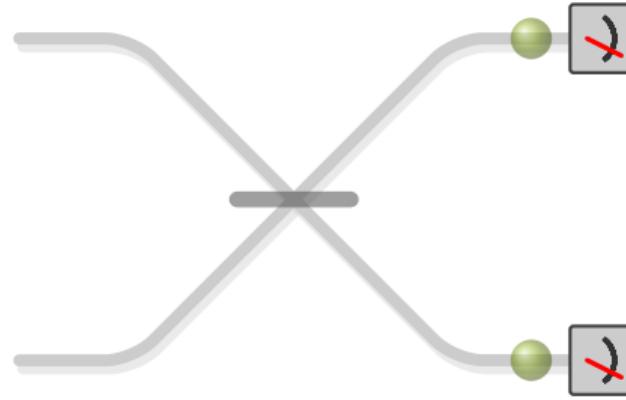
## Quantum sampling: The Hong-Ou-Mandel interferometer



## Quantum sampling: The Hong-Ou-Mandel interferometer

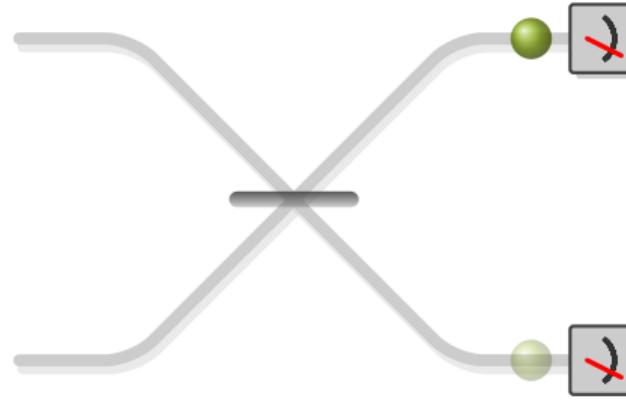


## Quantum sampling: The Hong-Ou-Mandel interferometer



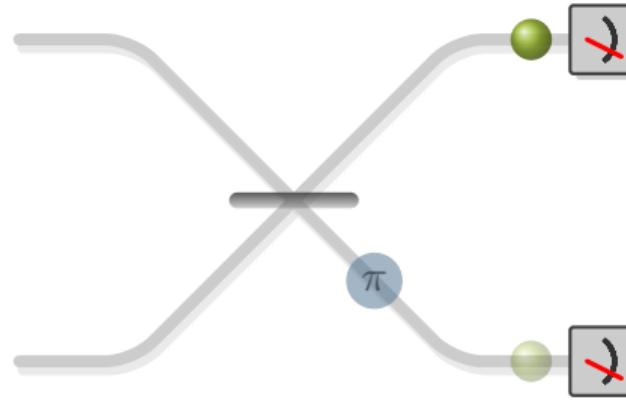
$$\frac{1}{\sqrt{2}} (|0,1\rangle + |1,0\rangle)$$

## Quantum sampling: The Hong-Ou-Mandel interferometer



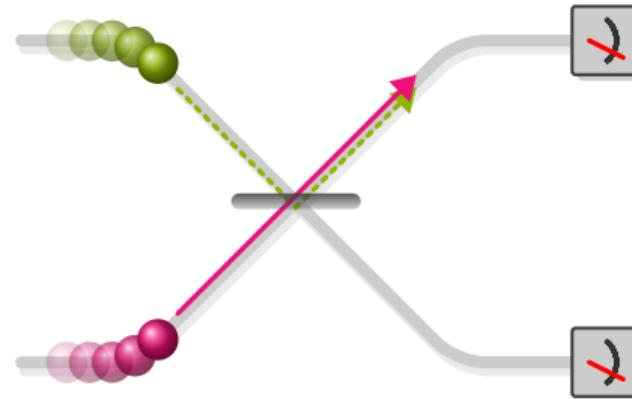
$$\sqrt{\frac{3}{5}}|0,1\rangle + \sqrt{\frac{2}{5}}|1,0\rangle$$

## Quantum sampling: The Hong-Ou-Mandel interferometer

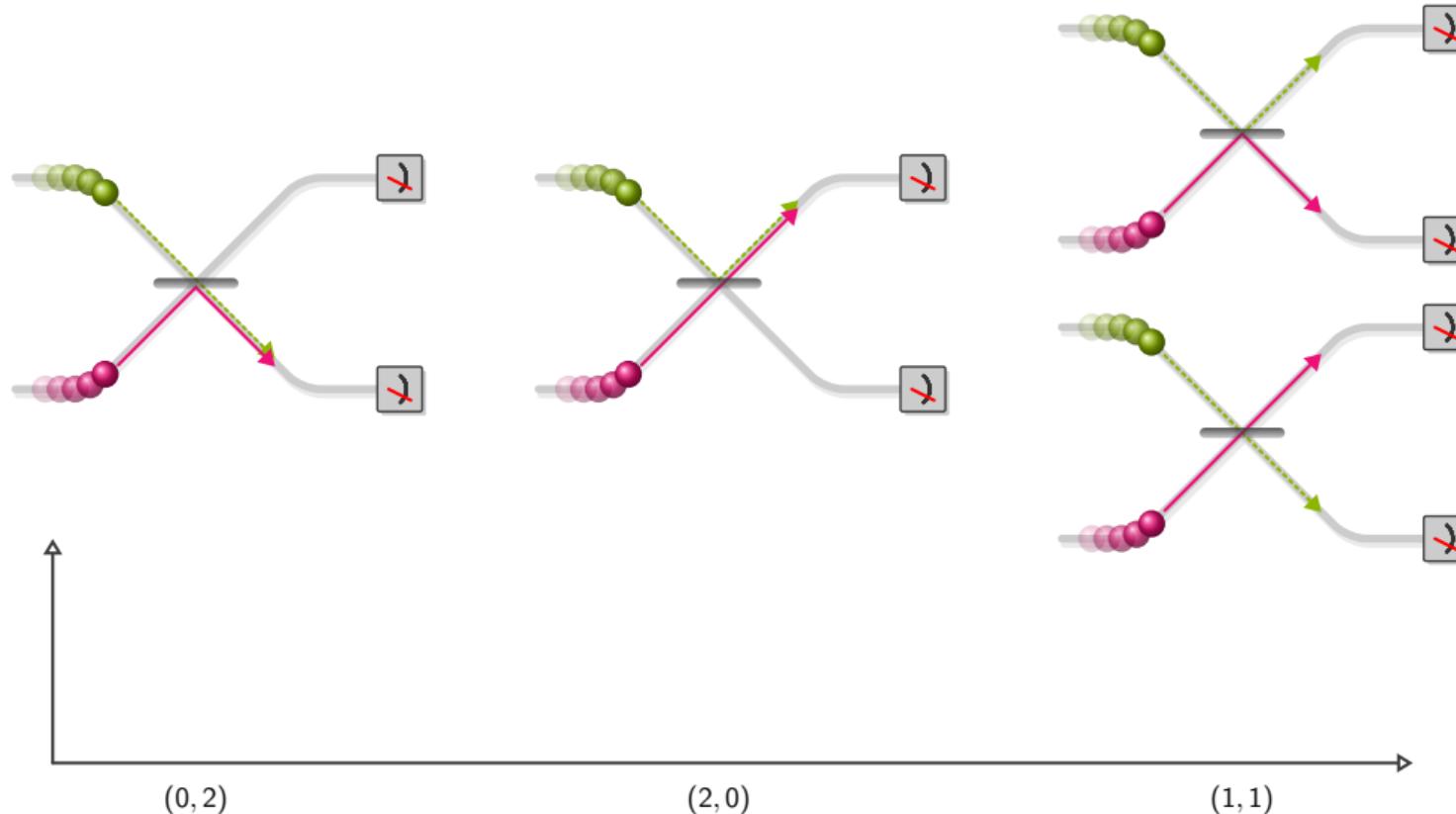


$$\sqrt{\frac{3}{5}}|0,1\rangle - \sqrt{\frac{2}{5}}|1,0\rangle$$

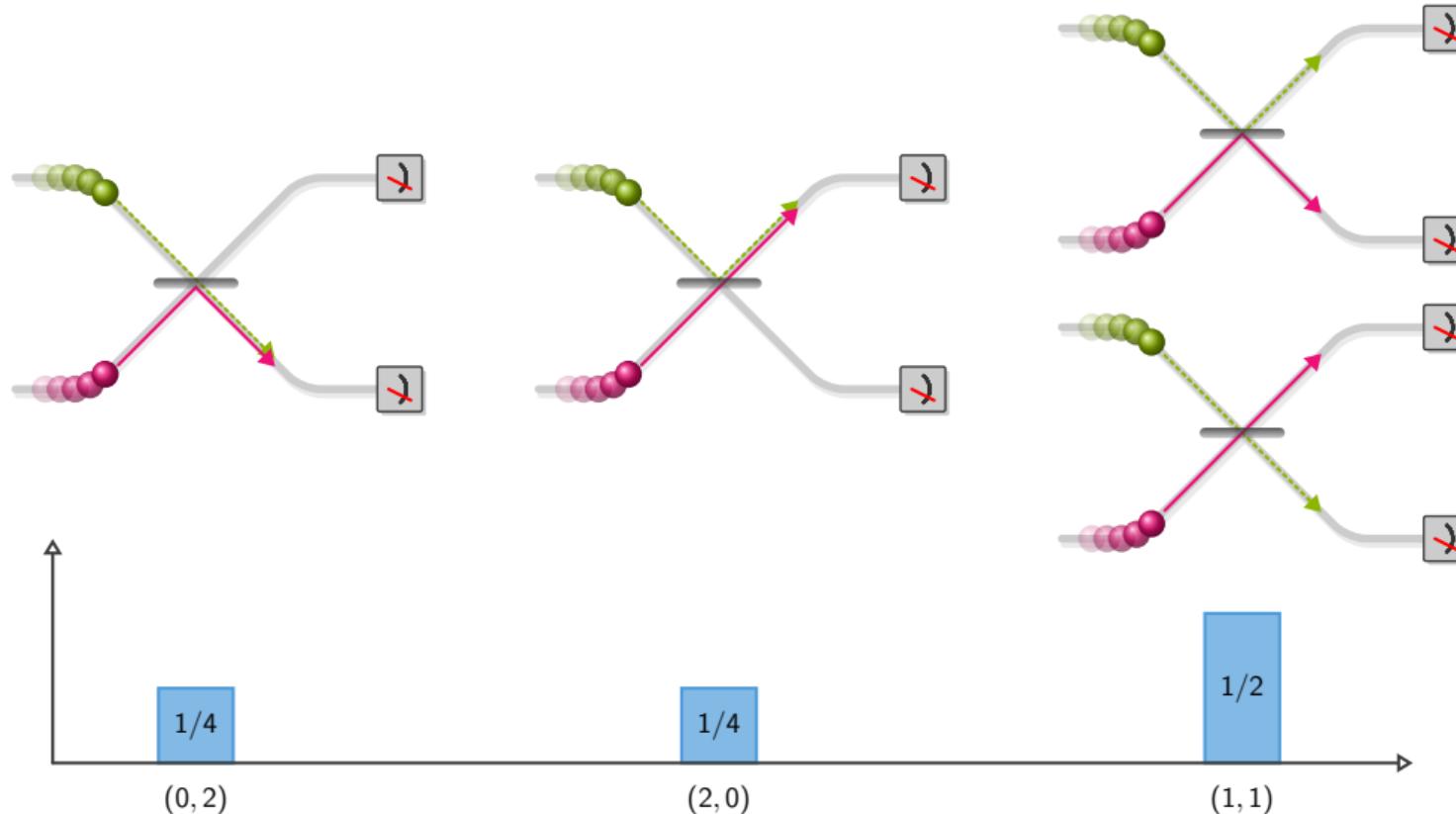
## Quantum sampling: The Hong-Ou-Mandel interferometer



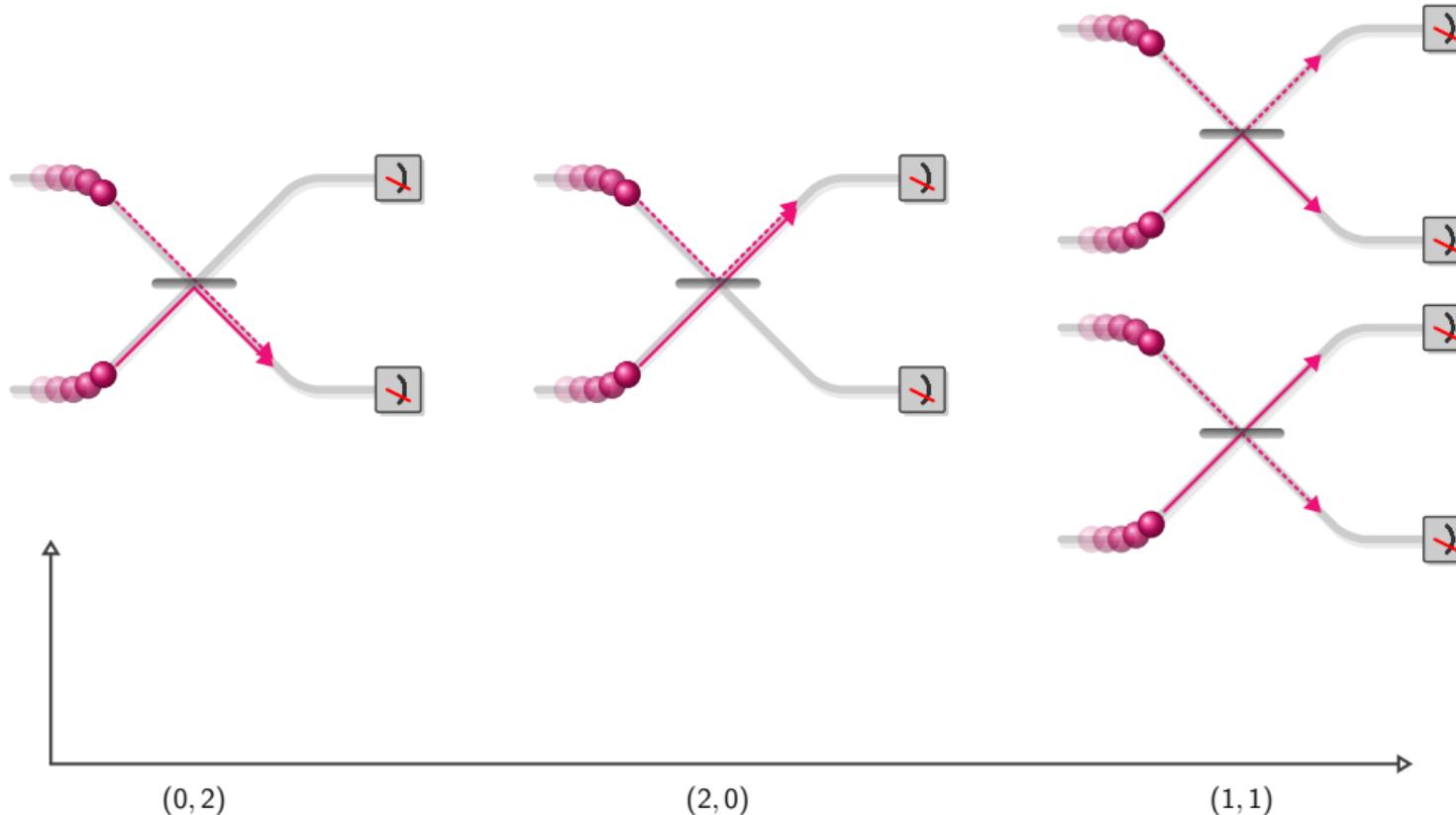
## Quantum sampling: The Hong-Ou-Mandel interferometer – two photons



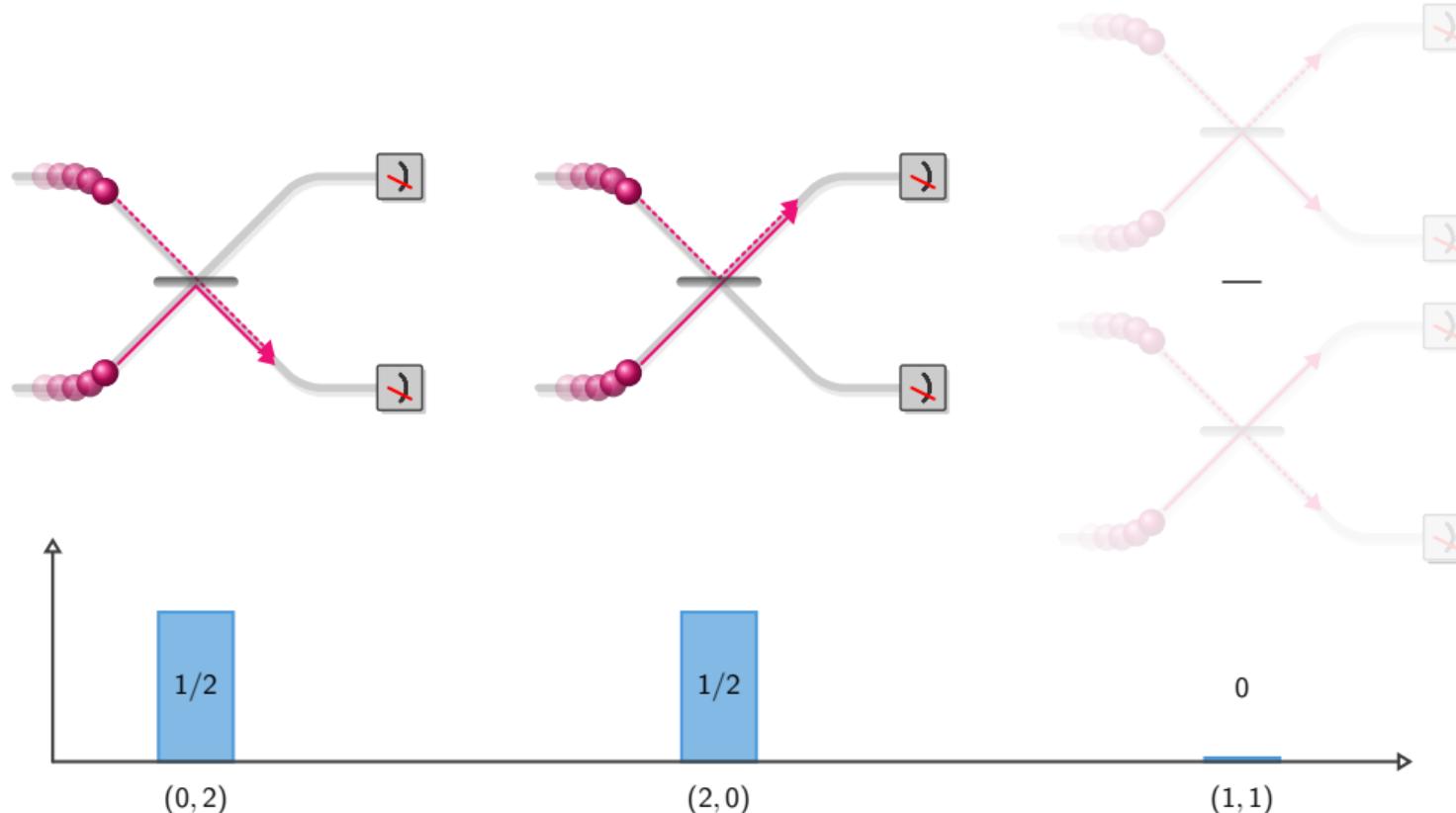
## Quantum sampling: The Hong-Ou-Mandel interferometer – two photons



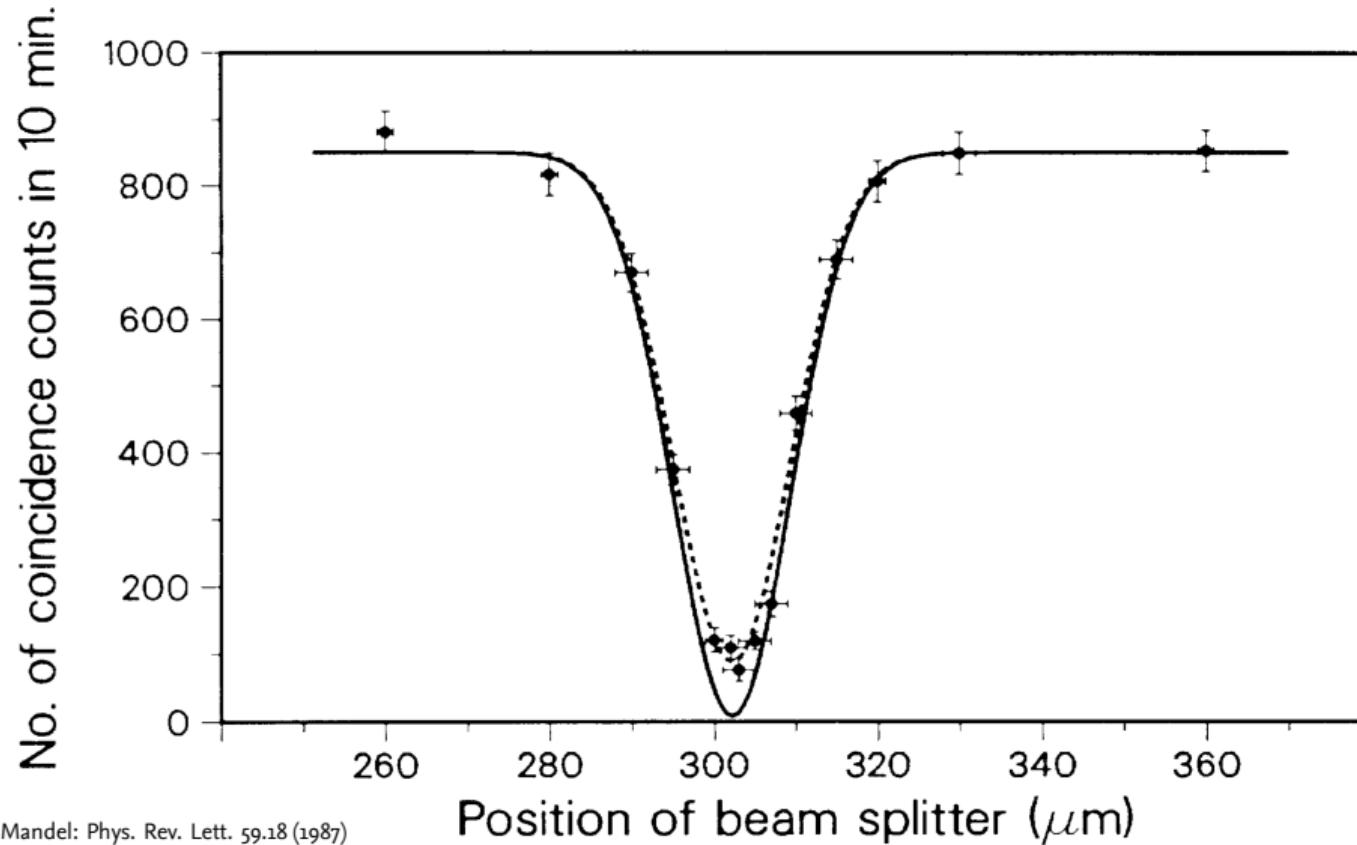
## Quantum sampling: The Hong-Ou-Mandel interferometer – two photons



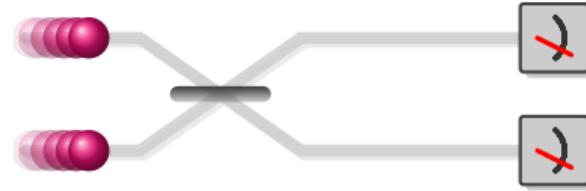
## Quantum sampling: The Hong-Ou-Mandel interferometer – two photons



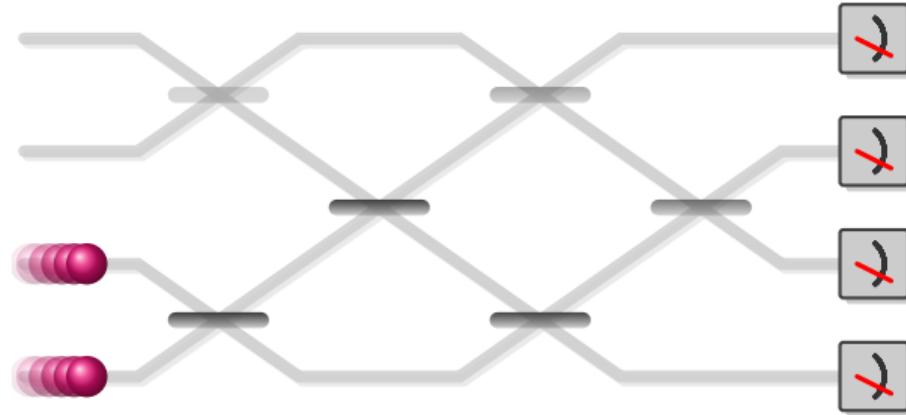
## Quantum sampling: The Hong-Ou-Mandel interferometer – two photons



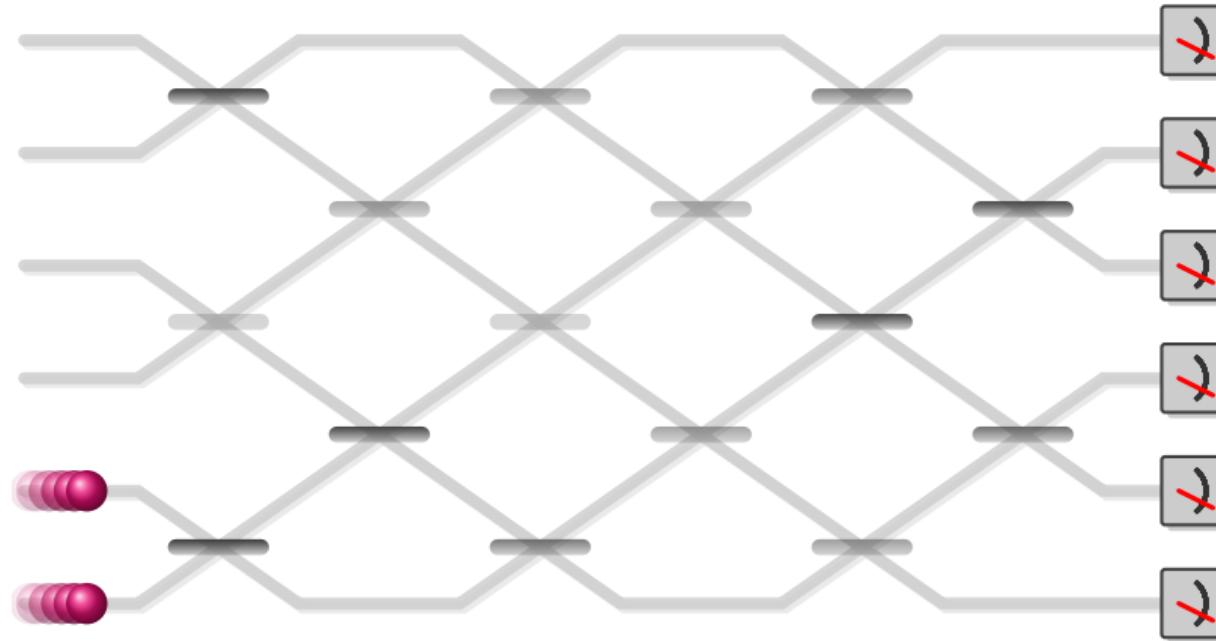
# Complex quantum sampling: A large network of HOM interferometers



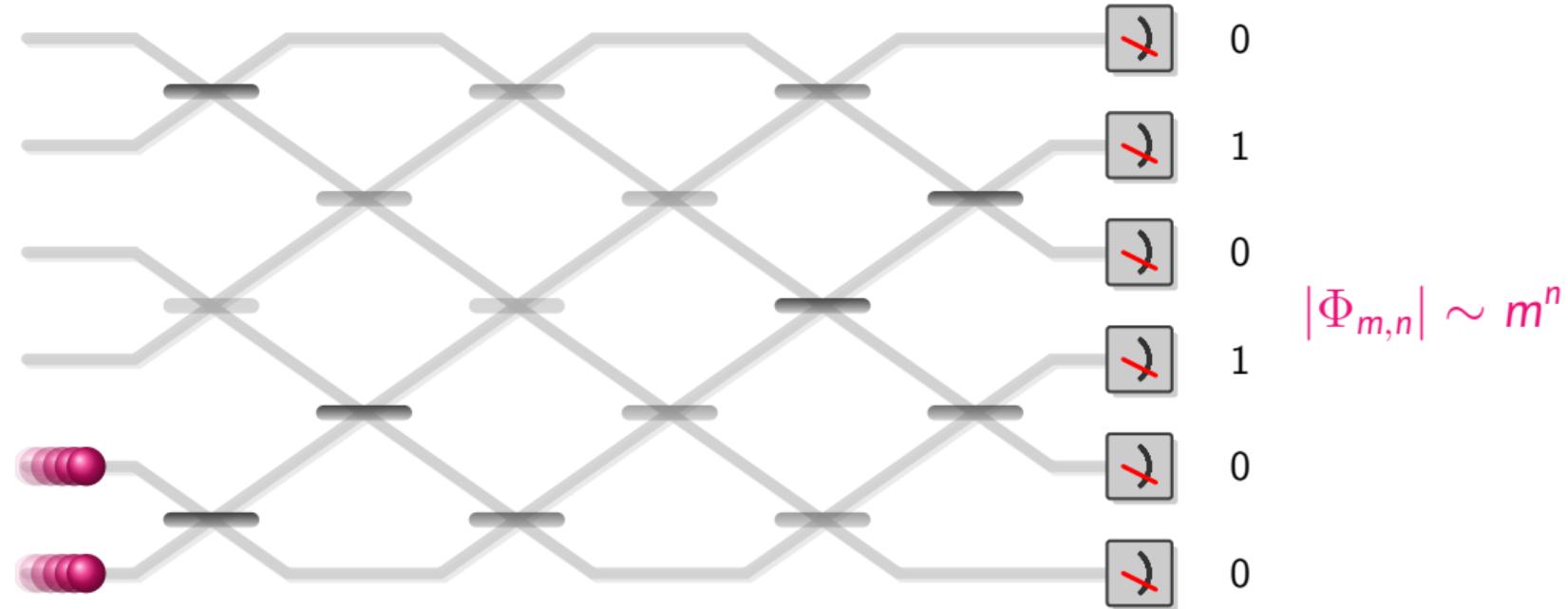
# Complex quantum sampling: A large network of HOM interferometers



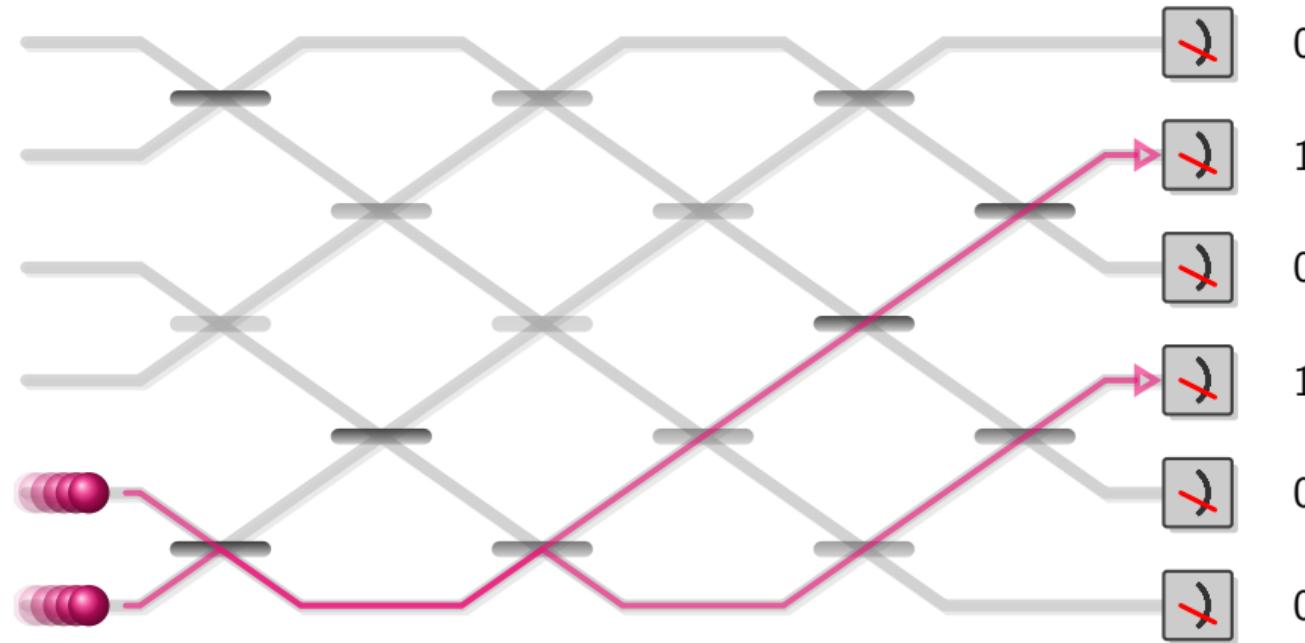
# Complex quantum sampling: A large network of HOM interferometers



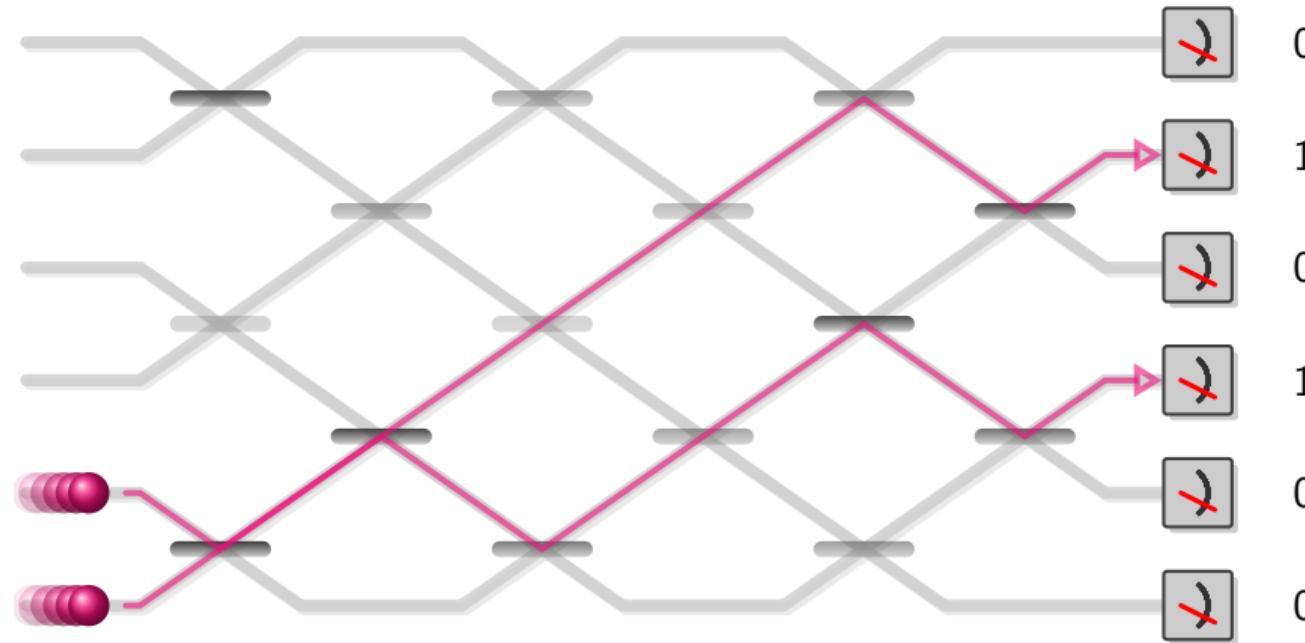
# Complex quantum sampling: A large network of HOM interferometers



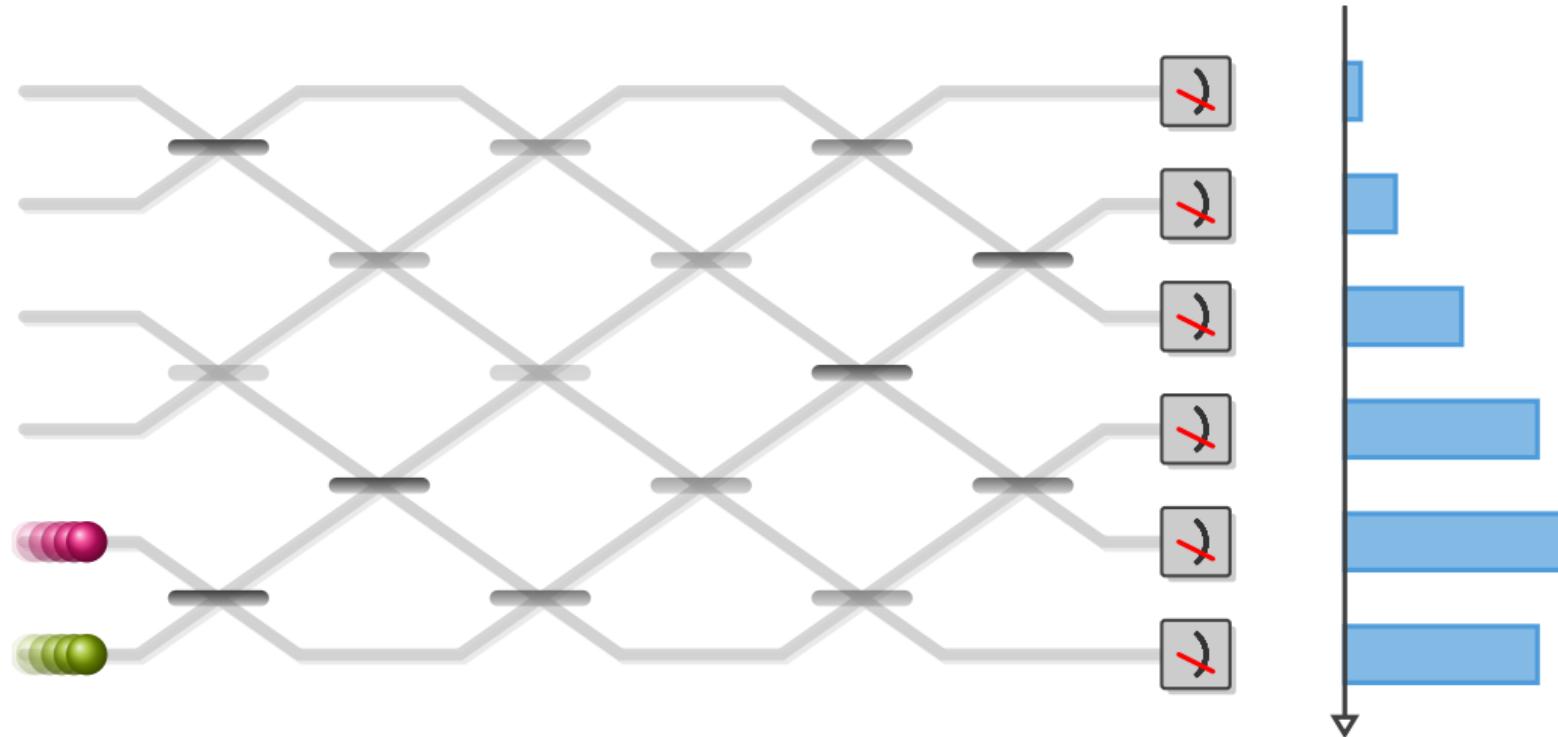
# Complex quantum sampling: A large network of HOM interferometers



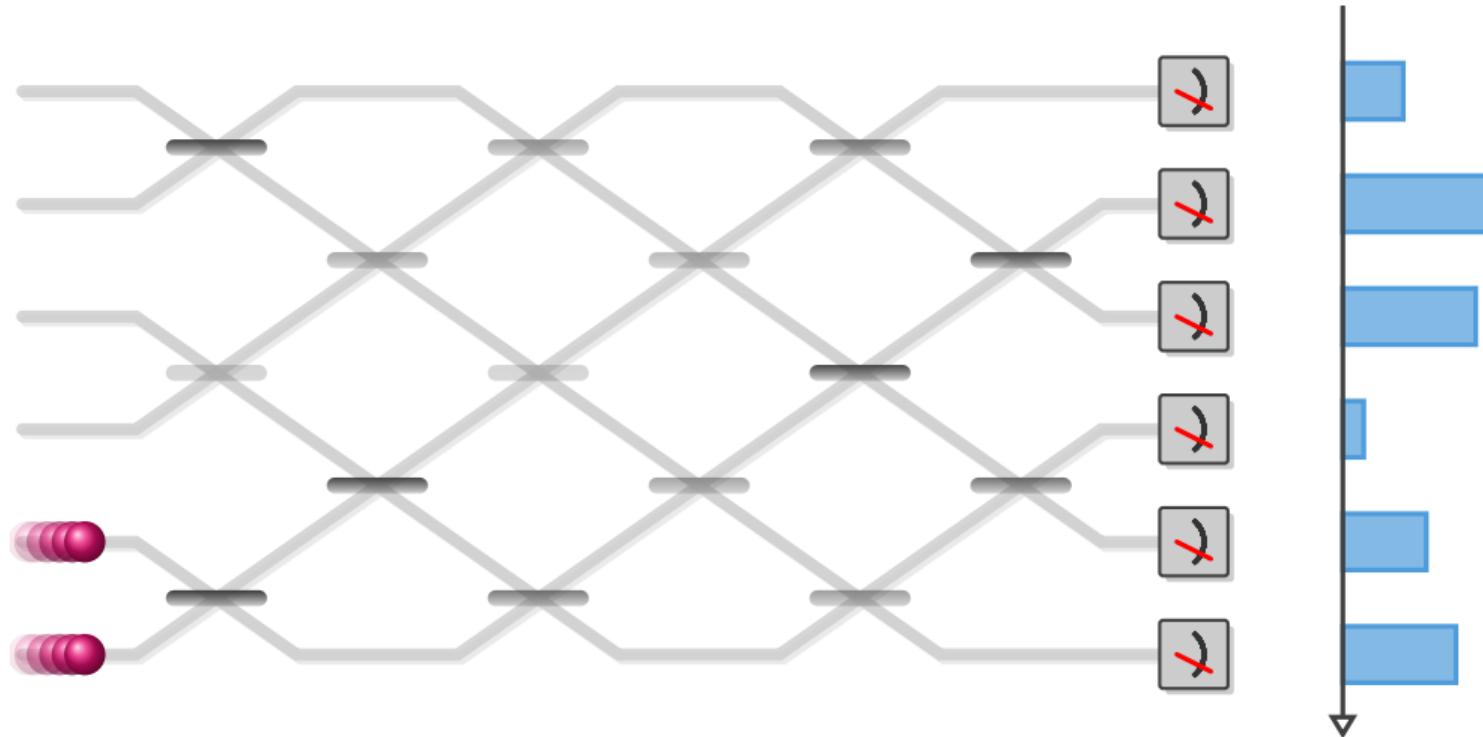
# Complex quantum sampling: A large network of HOM interferometers



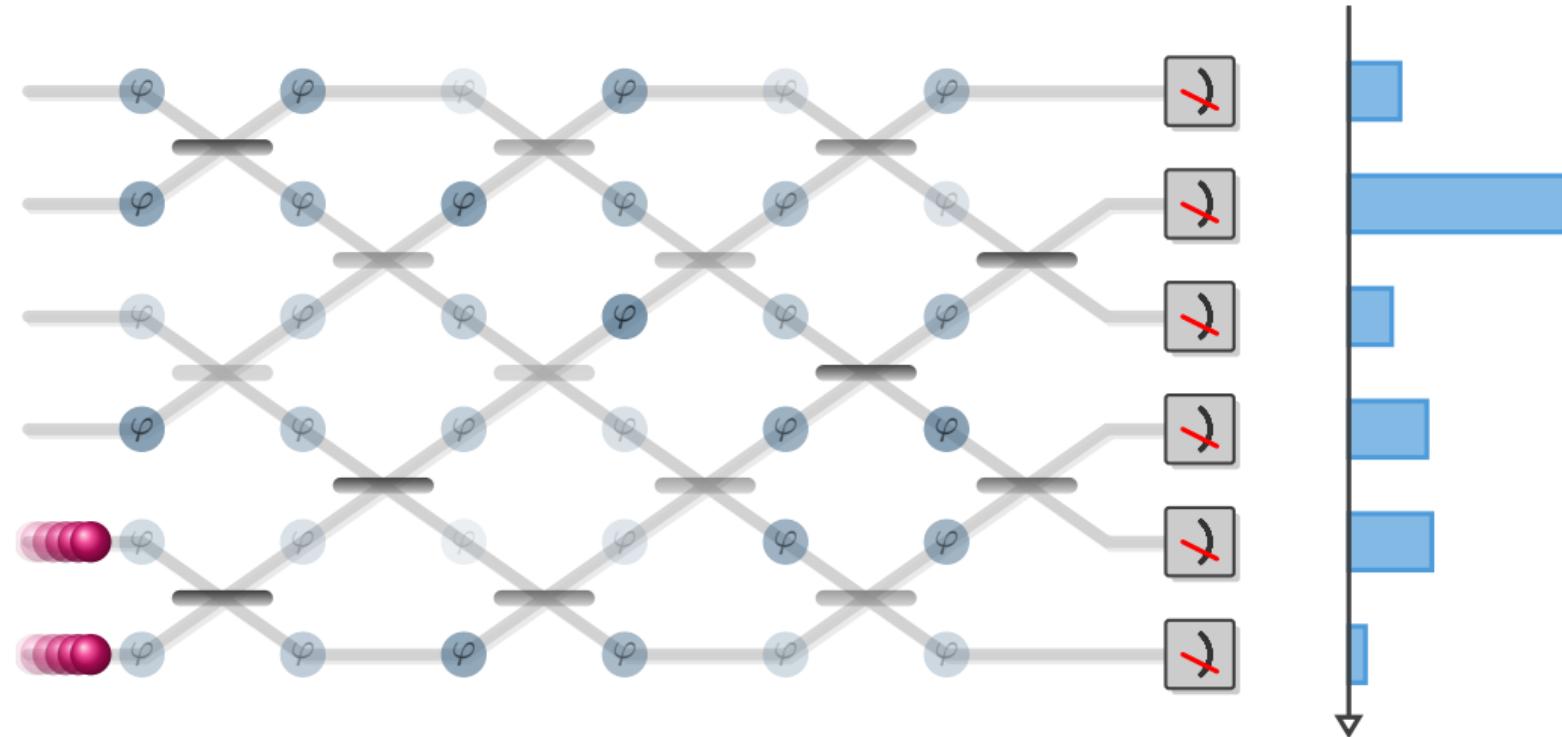
# Complex quantum sampling: A large network of HOM interferometers



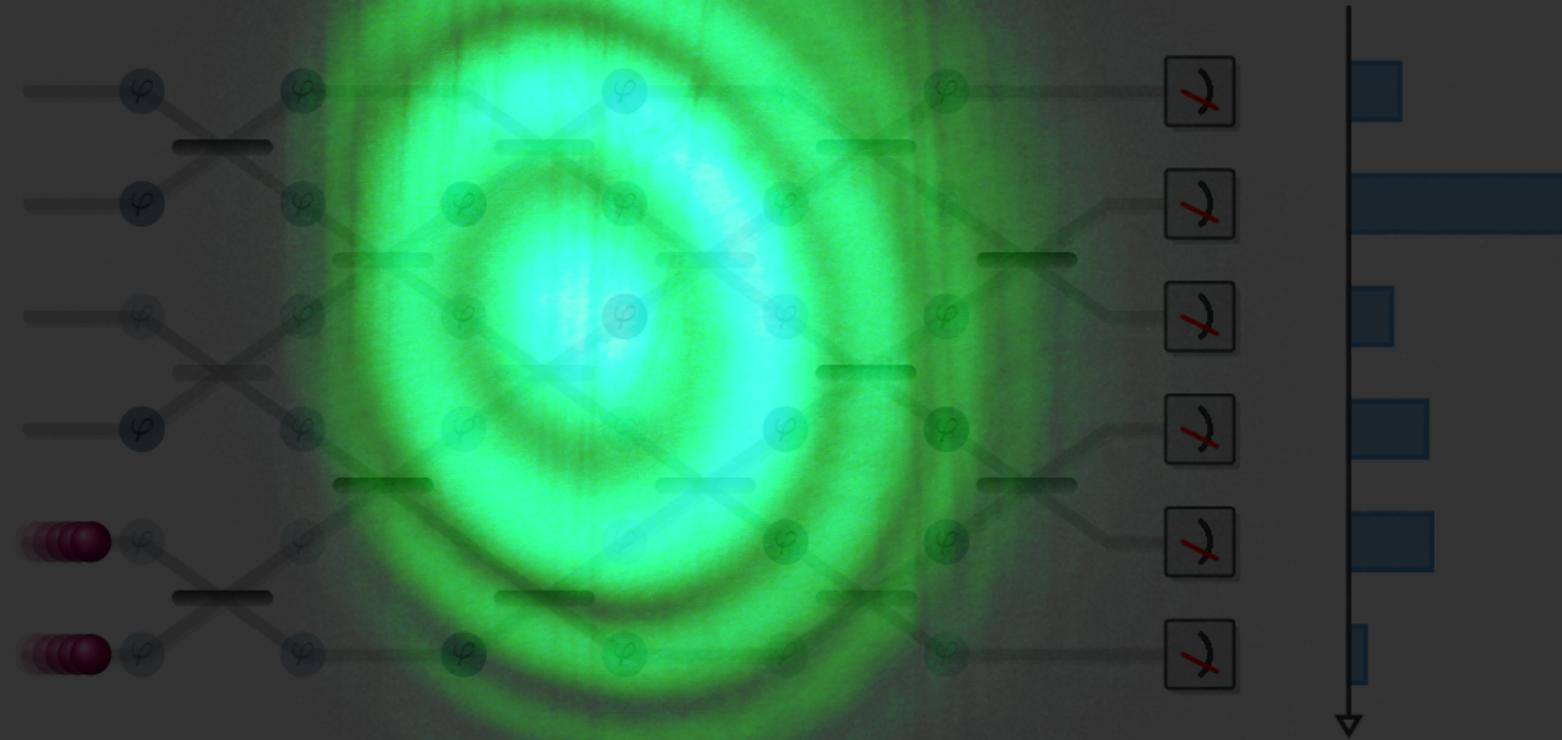
# Complex quantum sampling: A large network of HOM interferometers



# Complex quantum sampling: A large network of HOM interferometers



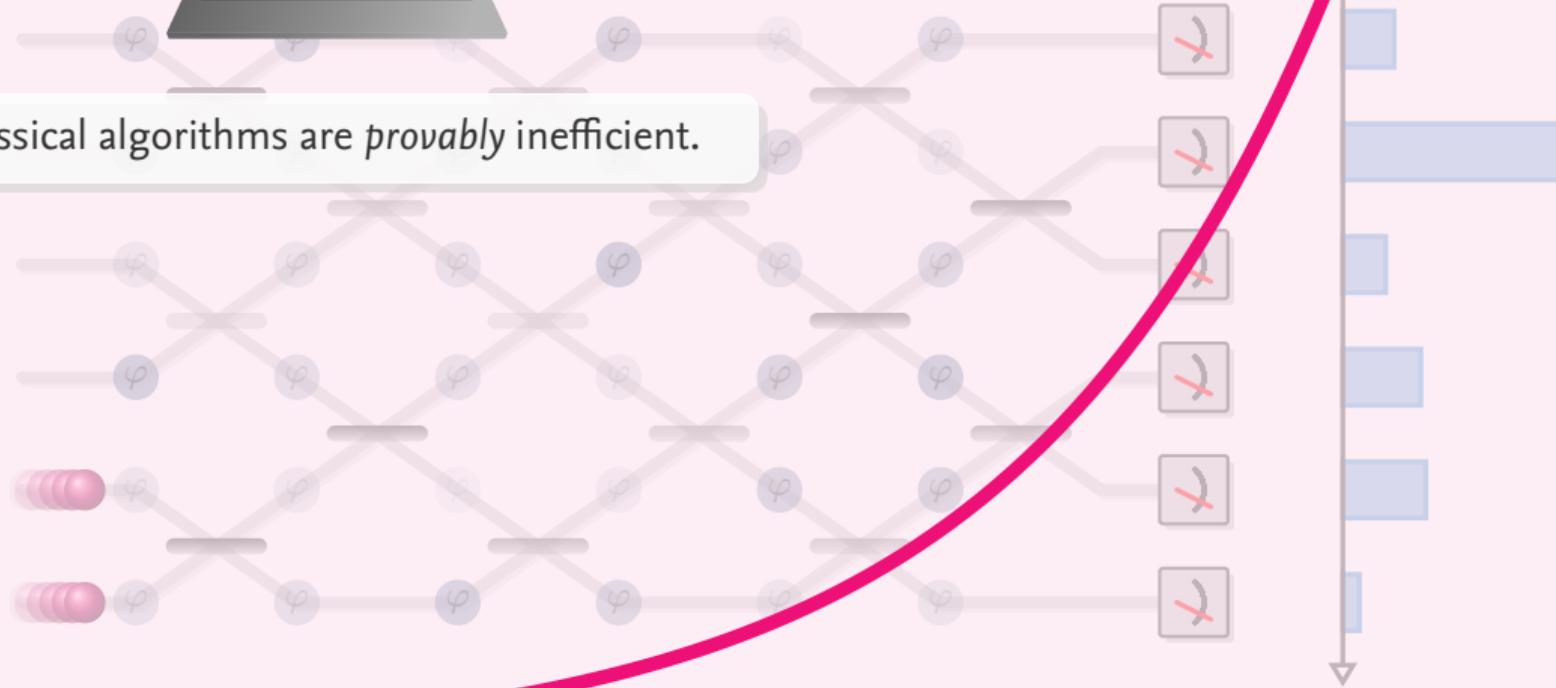
# Complex quantum sampling: A large network of HOM interferometers



# Complex quantum coupling: A large network of HOM interferometers



Classical algorithms are *provably* inefficient.



# Complex quantum coupling: A large network of HOM interferometers



Classical algorithms are *provably* inefficient.

The quantum sign problem

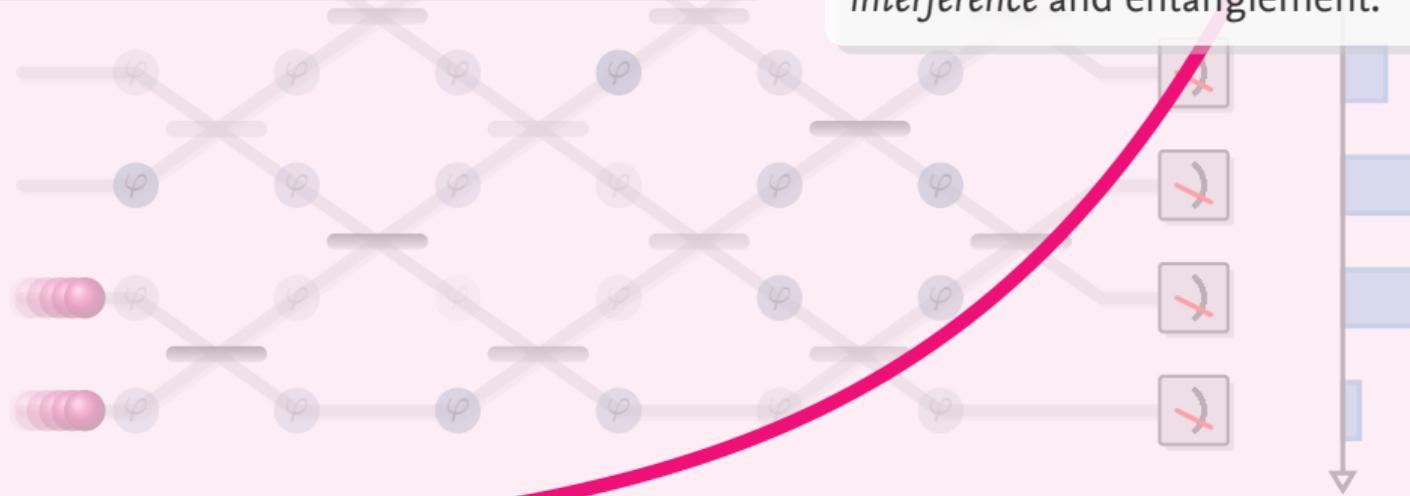


# Complex quantum coupling: A large network of Hadamard gates



Classical algorithms are *provably* inefficient.

Quantum algorithms exploit *high-dimensional interference* and entanglement.



# Complex quantum sampling: A large network of Hadamard meters



Classical algorithms are *provably* inefficient.

Quantum algorithms exploit *high-dimensional interference* and entanglement.



How is the **quantum sign problem** reflected in the computational complexity of different sampling-related tasks?

# Delineating the quantum-classical boundary

Classical

Quantum

# Delineating the quantum-classical boundary

Classical

Quantum

1. Sampling hardness in generic quantum computations

# Delineating the quantum-classical boundary

Classical

2. Estimating outcome probabilities

Quantum

1. Sampling hardness in generic quantum computations

## Perspective 1

### The sign problem in complexity theory

# The sign problem in complexity theory: the basics

## $\exists$ -SAT formula

$$f(x) = (x_1 \vee \bar{x}_7 \vee x_{137}) \wedge (x_5 \vee x_{12} \vee \bar{x}_{17}) \wedge (\bar{x}_{32} \vee \bar{x}_7 \vee x_{17}) \wedge \dots \wedge (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_1)$$

NP: Decide whether  $\exists x : f(x) = 1$ .

#P: Compute  $acc(f) := |\{x : f(x) = 1\}| \in \{0, \dots, 2^n\}$ .

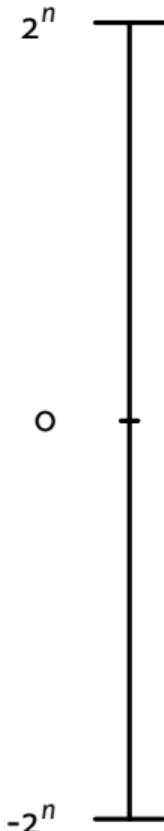
GapP: Compute  $gap(f) := acc(f) - rej(f) \in \{-2^n, \dots, 2^n\}$ .

PP: Decide whether  $gap(f) > 0$  or  $\leq 0$ .



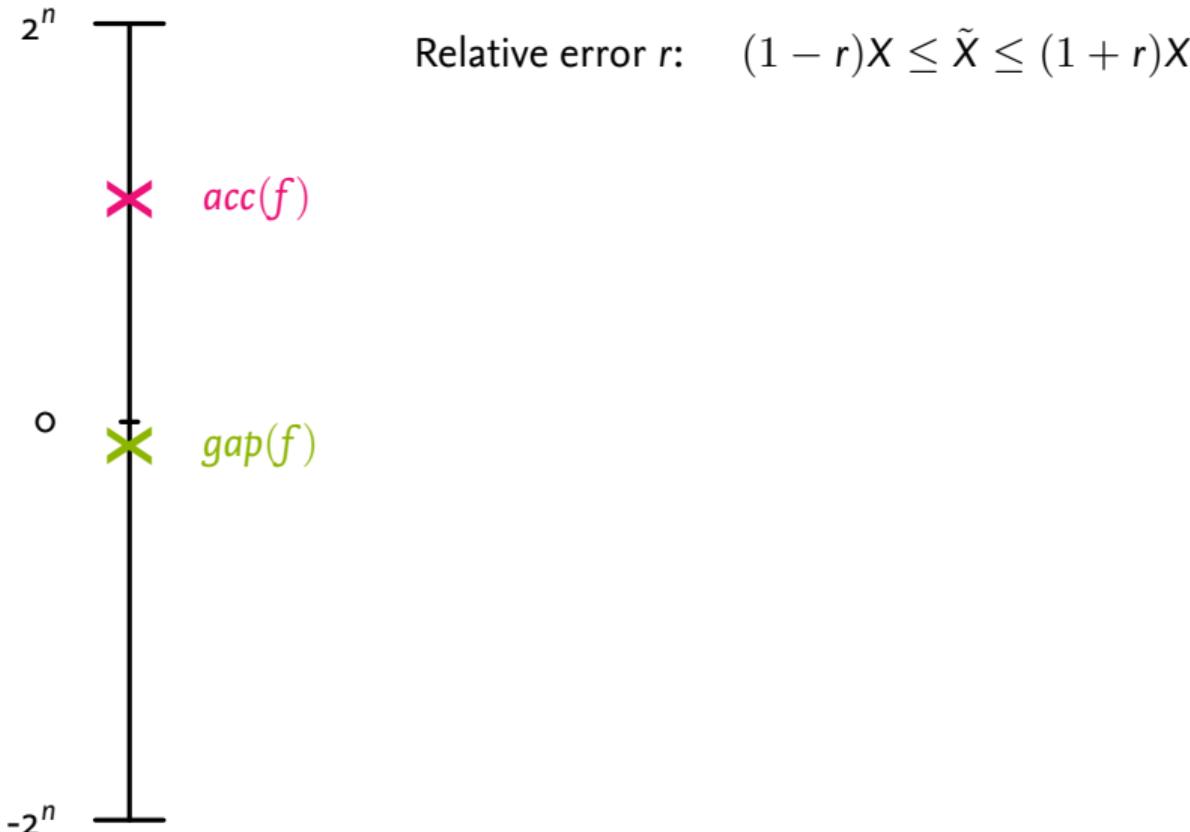
$$P^{\#P} = P^{\text{GapP}} = P^{\text{PP}}$$

# The sign problem in complexity theory: approximations

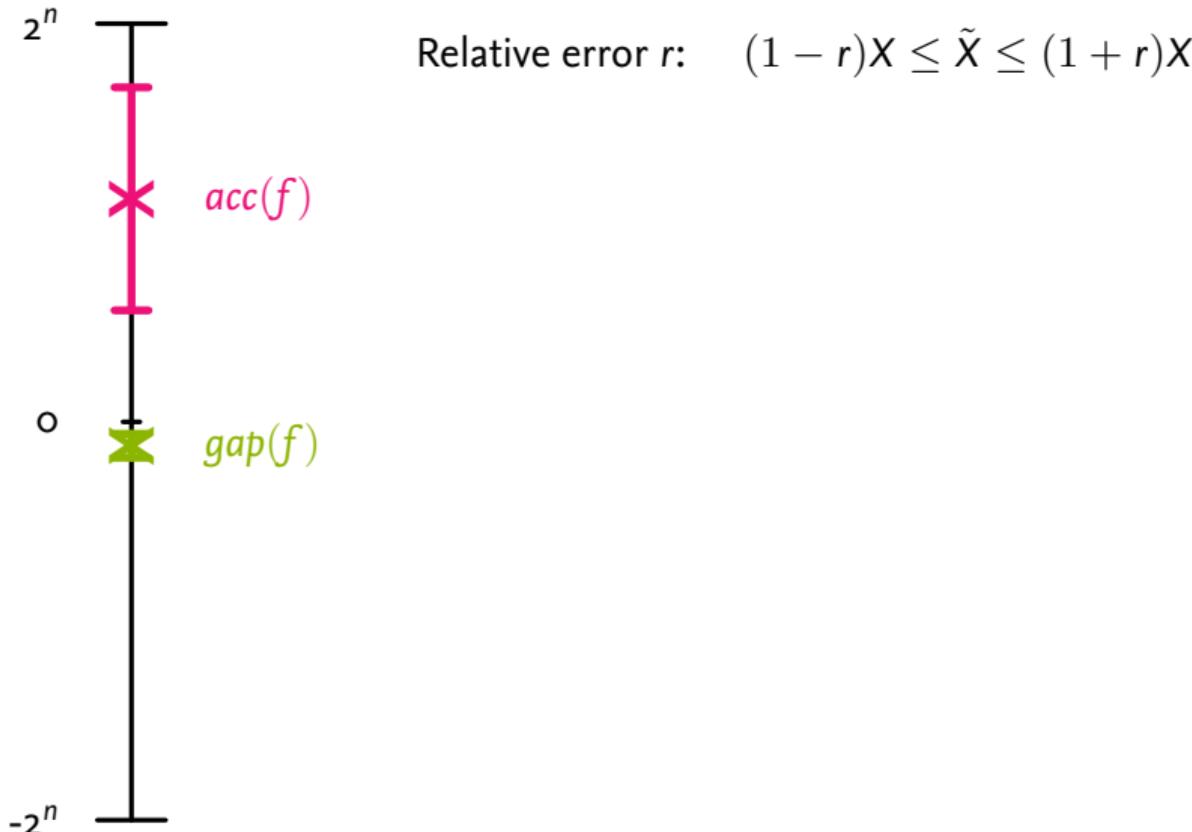


Relative error  $r$ :  $(1 - r)X \leq \tilde{X} \leq (1 + r)X$

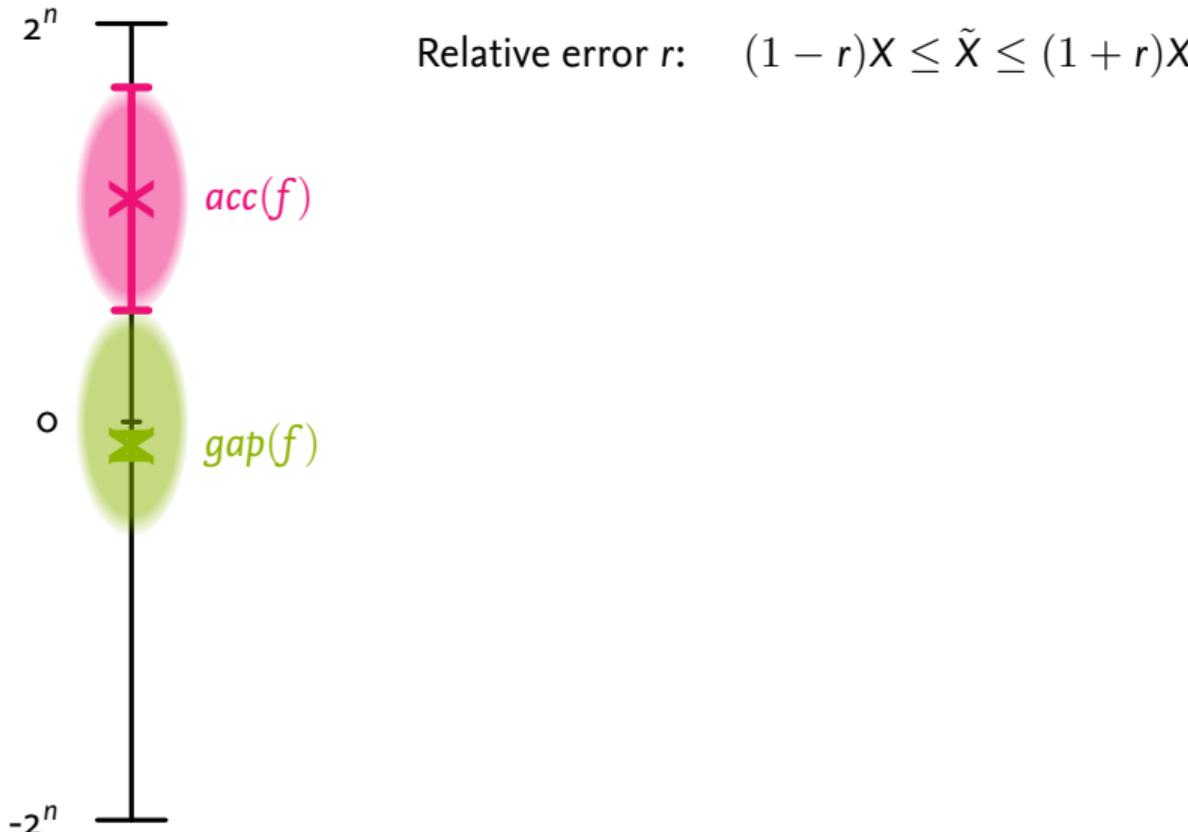
# The sign problem in complexity theory: approximations



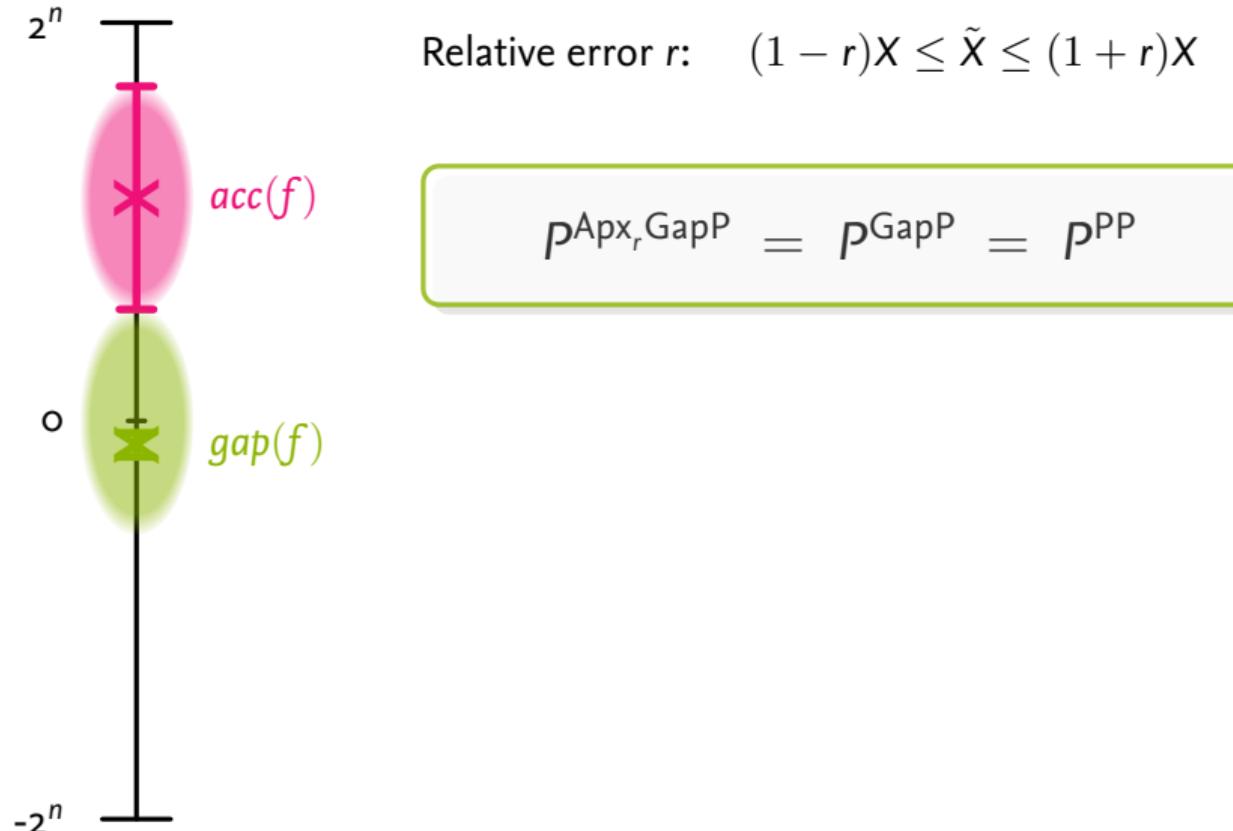
# The sign problem in complexity theory: approximations



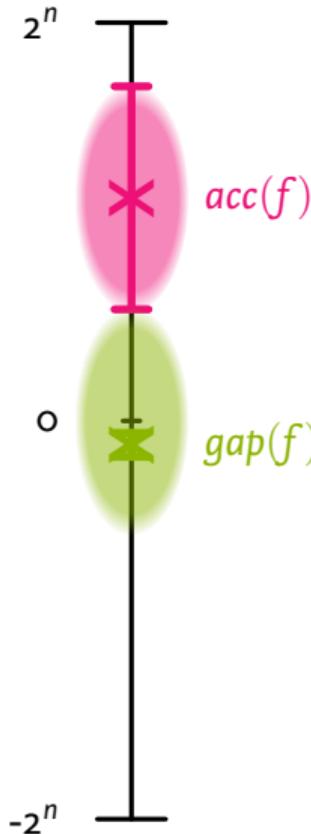
# The sign problem in complexity theory: approximations



# The sign problem in complexity theory: approximations



# The sign problem in complexity theory: approximations

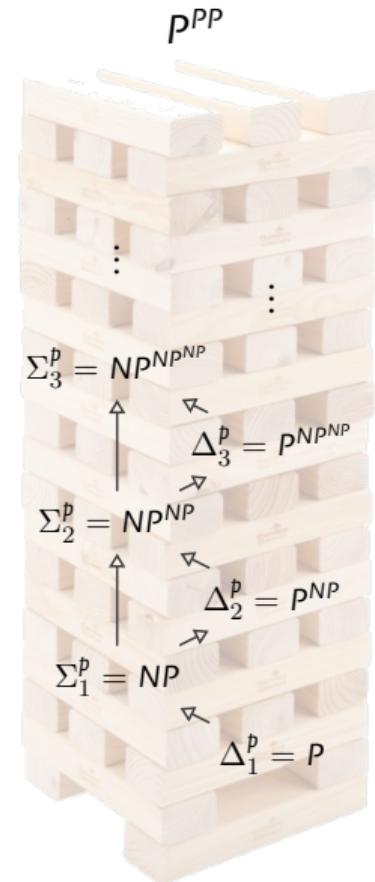


Relative error  $r$ :  $(1 - r)X \leq \tilde{X} \leq (1 + r)X$

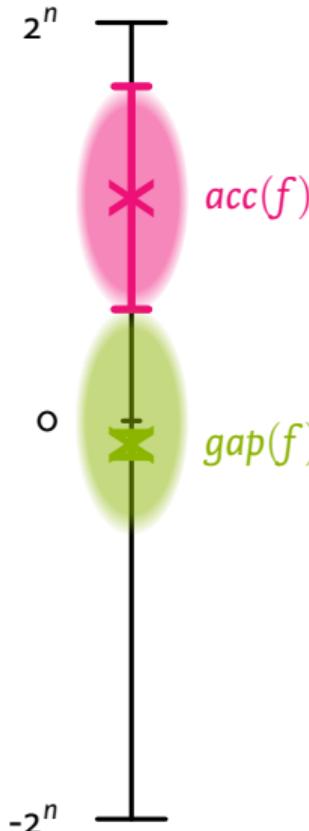
$$P^{\text{Apx}_r \text{GapP}} = P^{\text{GapP}} = P^{\text{PP}}$$

Stockmeyer '83: For any  $r \in 1/\text{poly}(n)$ :

$$\text{Apx}_r \#P \subset \text{FBPP}^{\text{NP}}$$



# The sign problem in complexity theory: approximations

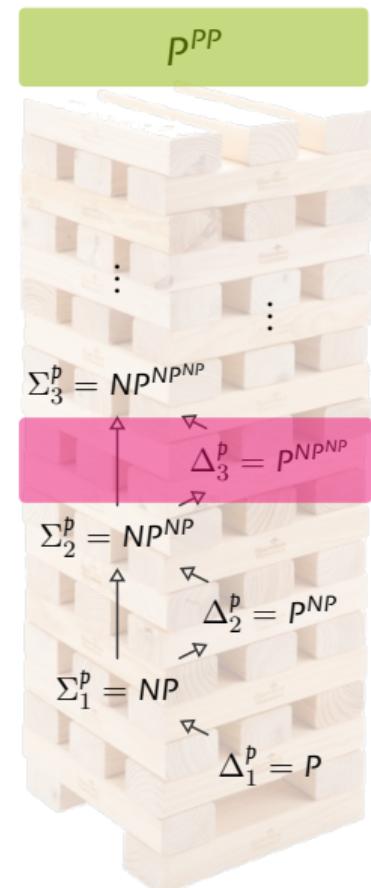


Relative error  $r$ :  $(1 - r)X \leq \tilde{X} \leq (1 + r)X$

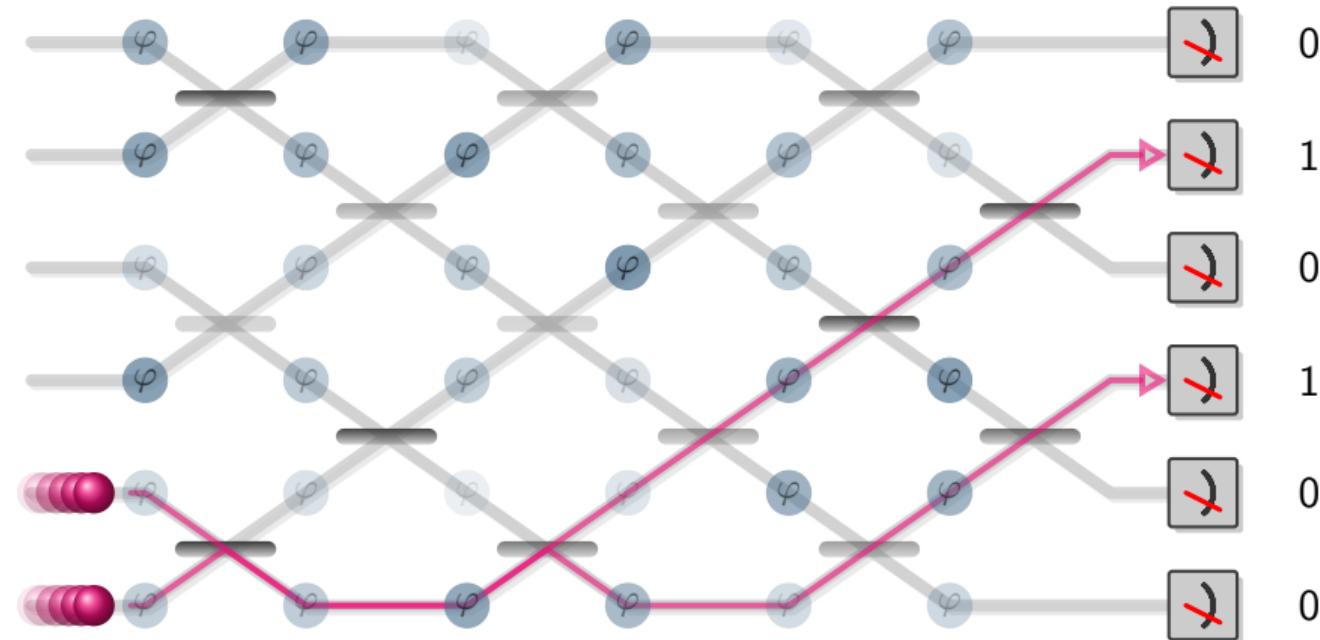
$$P^{\text{Apx}_r \text{GapP}} = P^{\text{GapP}} = P^{\text{PP}}$$

Stockmeyer '83: For any  $r \in 1/\text{poly}(n)$ :

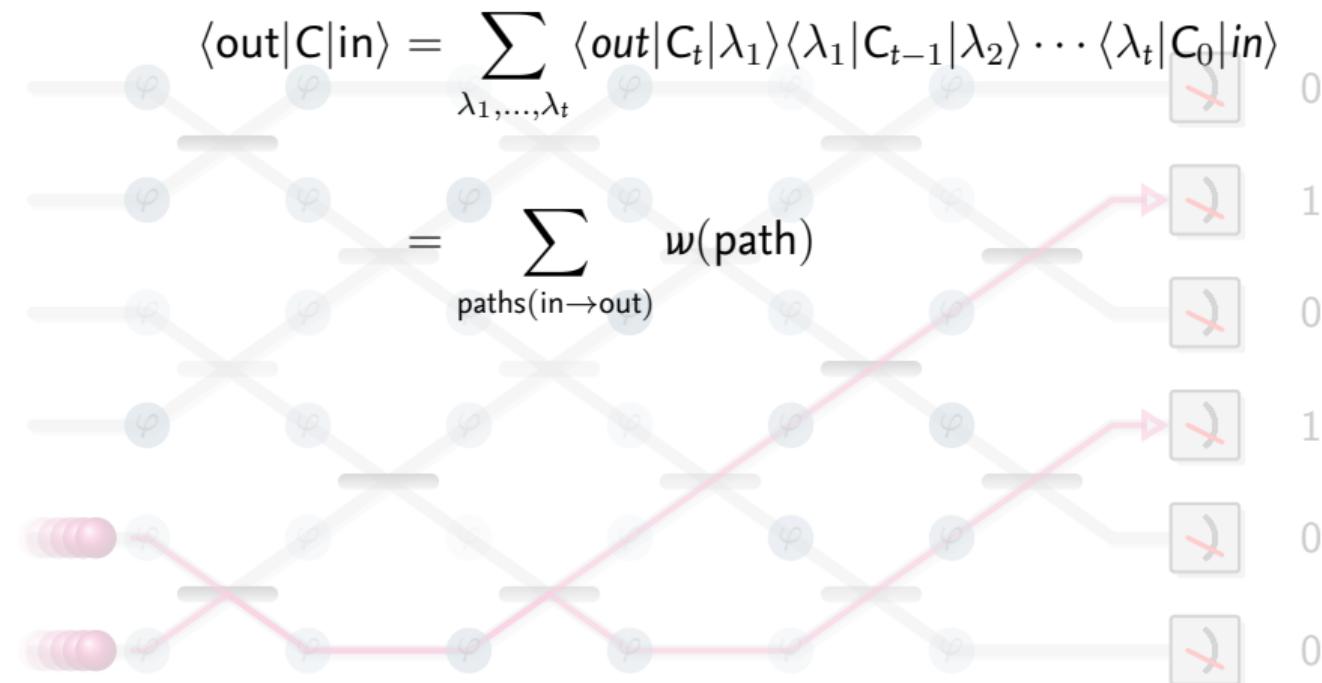
$$\text{Apx}_r \#P \subset \text{FBPP}^{\text{NP}}$$



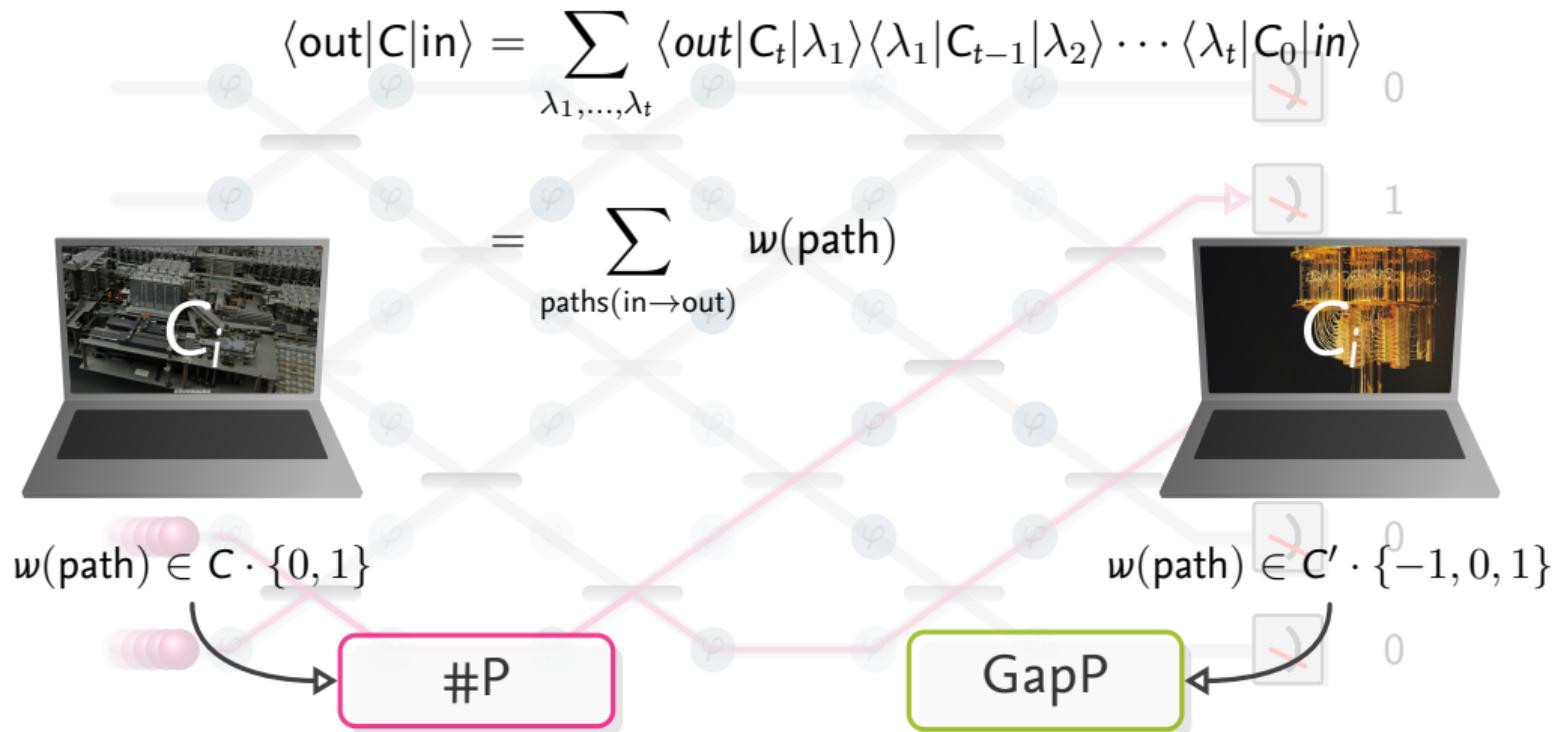
## Success probabilities of quantum and classical circuits



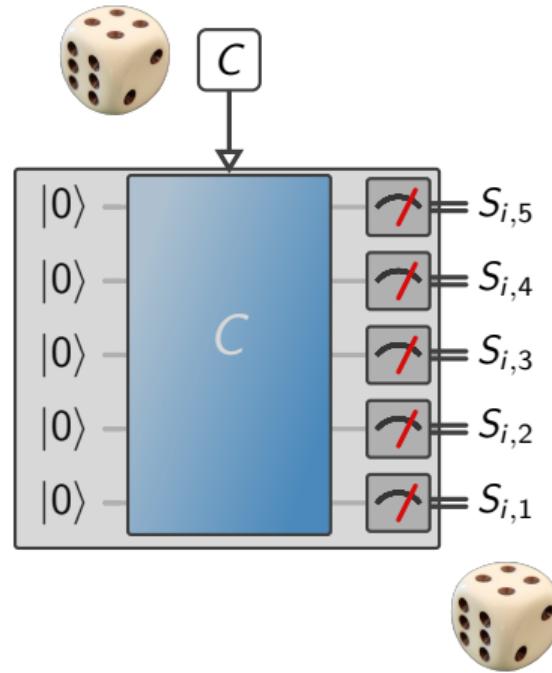
## Success probabilities of quantum and classical circuits



# Success probabilities of quantum and classical circuits



# Leveraging hardness of estimation to hardness of sampling

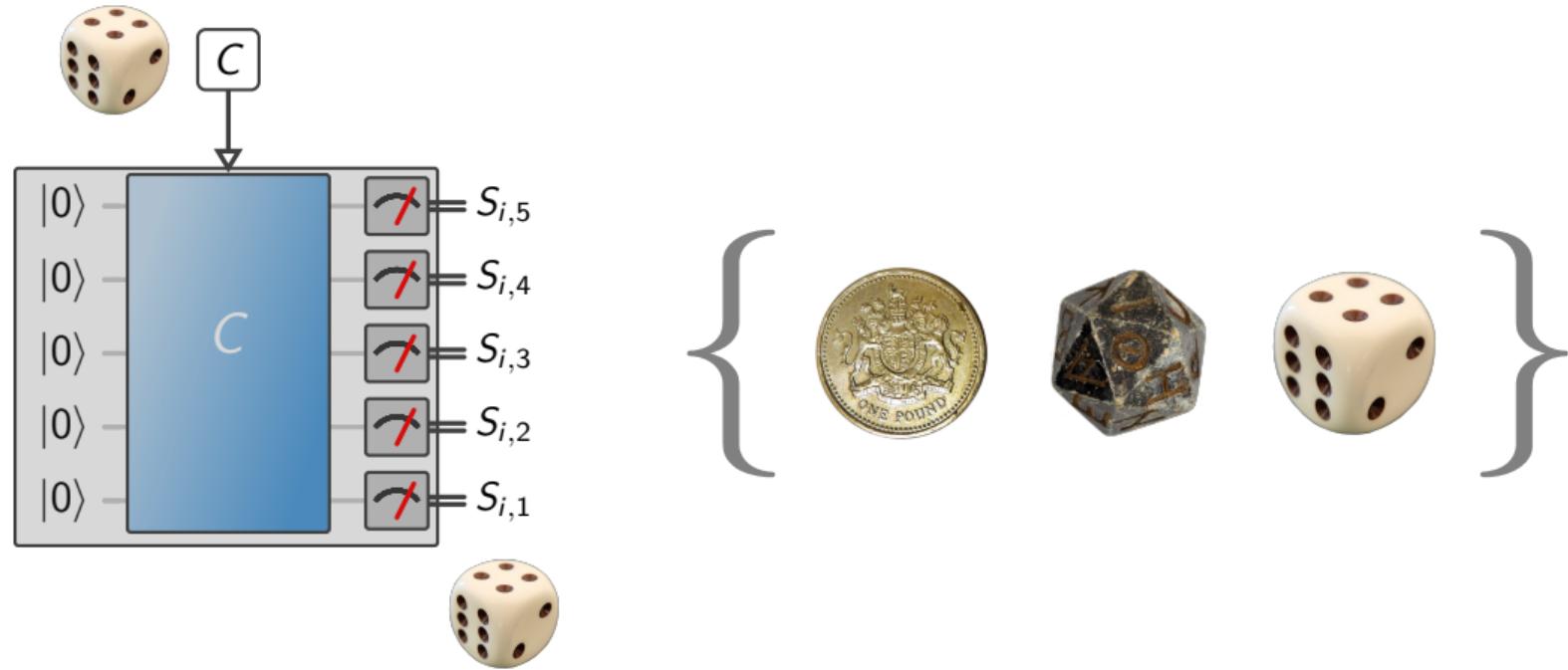


Consider

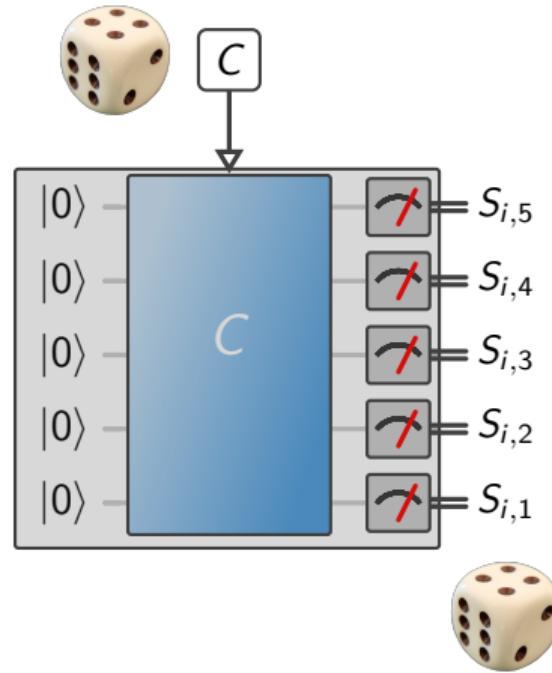
- $n$  qubits
- $\mathcal{U}$  family of quantum circuits, e.g.
  - $\mathcal{U}$  = Universal circuits (e.g. using Clifford + T-gates) [Boi18]
  - $\mathcal{U}$  = Diagonal (IQP) circuits [BMS16]
- $p_{\mathcal{U}}(x) := |\langle x|U|0\rangle^{\otimes n}|^2$

**TASK:** Given  $U \in_R \mathcal{U}$ , sample from  $p_{\mathcal{U}}$ .

## Leveraging hardness of estimation to hardness of sampling



# Leveraging hardness of estimation to hardness of sampling



## CLASSICAL DERANDOMIZABLE SAMPLING

Given  $U \in \mathcal{U}$ , uniformly random  $r$ , output  $y$  s.t.

$$\Pr_x[y] \propto \sum_r f_U(r) = p_U(0^n),$$

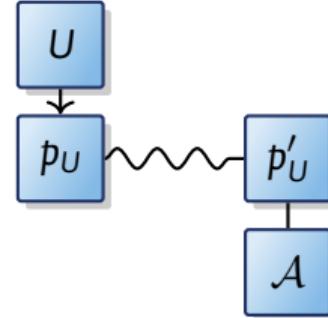
for  $f_U : \{0, 1\}^m \rightarrow \{0, 1\}$  a  $\#P$  function.

## Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .

## Approximate quantum sampling is classically hard

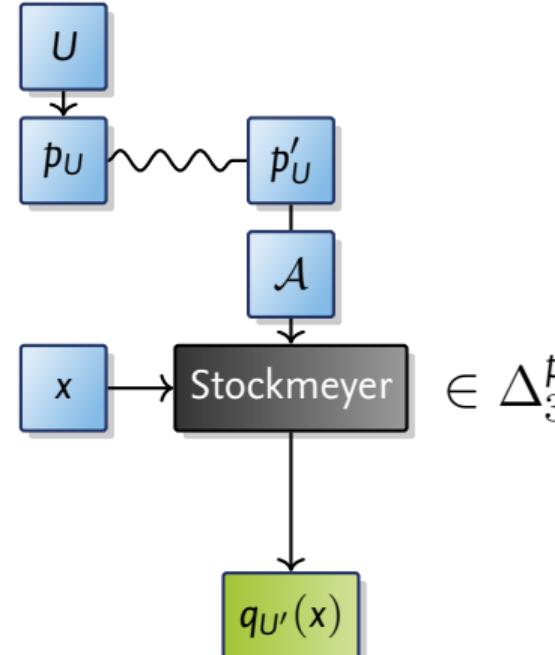
Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .



$$x \in \{0, 1\}^N$$

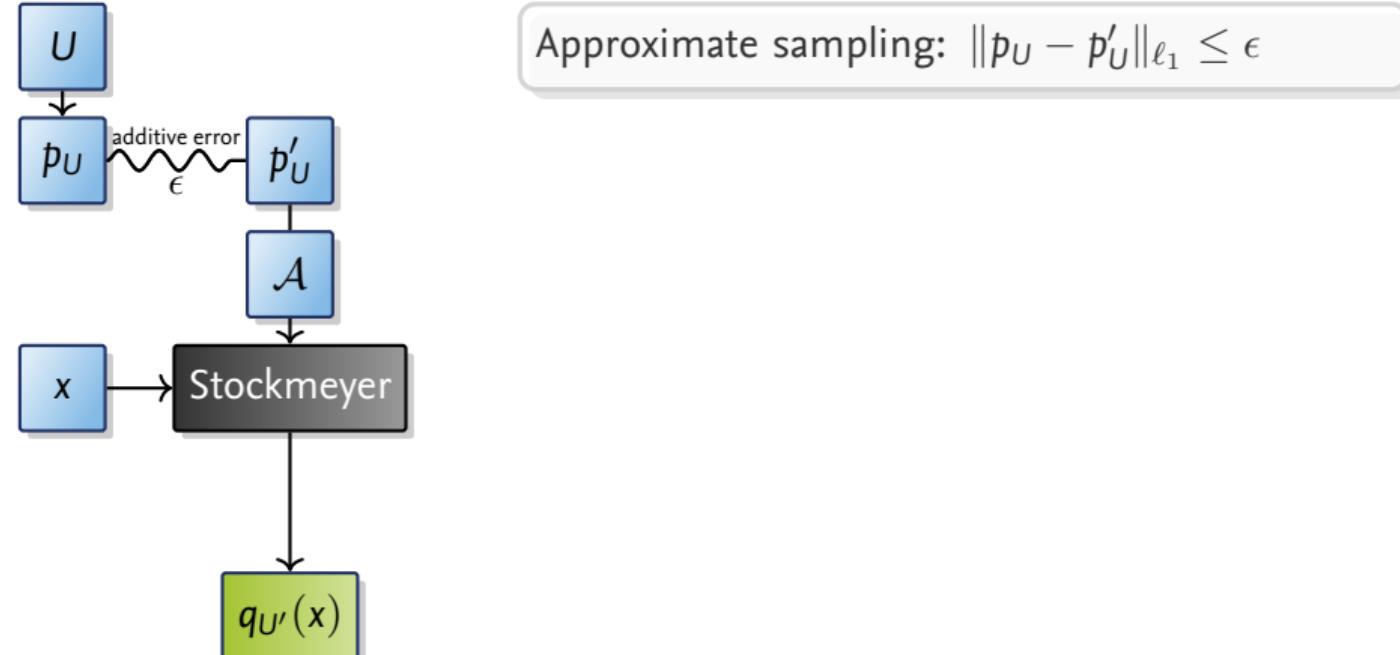
## Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .



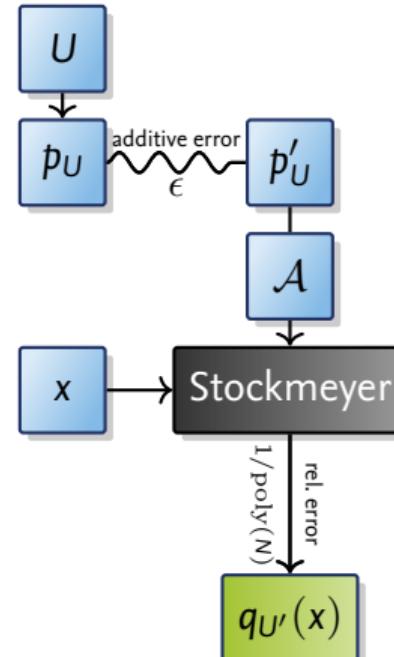
## Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p_U - p'_U\|_{\ell_1} \leq \epsilon$ .



# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p_U - p'_U\|_{\ell_1} \leq \epsilon$ .



Approximate sampling:  $\|p_U - p'_U\|_{\ell_1} \leq \epsilon$

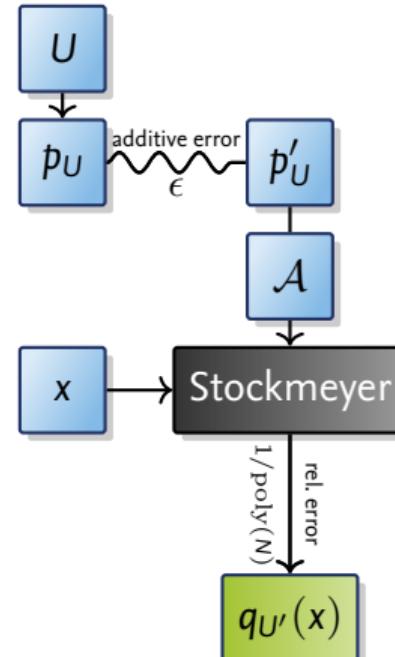
Stockmeyer error:  $|q_U(x) - p'_U(x)| \leq \frac{p'_U(x)}{\text{poly}(n)}$

With probability  $1 - \delta$  over  $U \in_R \mathcal{U}$

$$|q_U(x) - p_U(x)| \leq \frac{p_U(x)}{\text{poly}(n)} + \frac{\epsilon}{2^n \delta} \left(1 + \frac{1}{\text{poly}(n)}\right)$$

# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p_U - p'_U\|_{\ell_1} \leq \epsilon$ .



Approximate sampling:  $\|p_U - p'_U\|_{\ell_1} \leq \epsilon$

Stockmeyer error:  $|q_U(x) - p'_U(x)| \leq \frac{p'_U(x)}{\text{poly}(n)}$

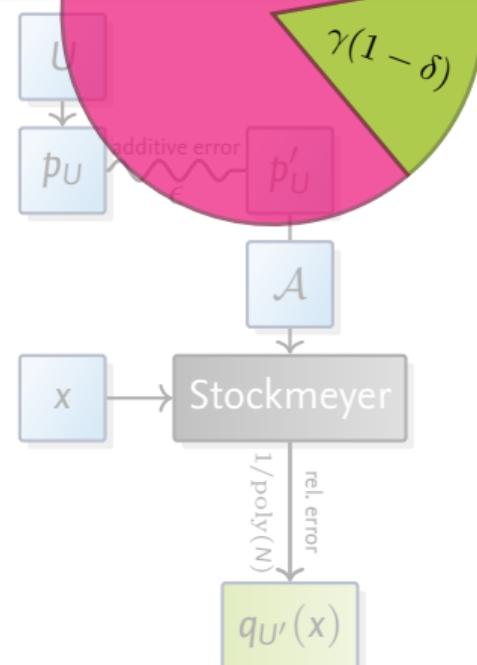
With probability  $1 - \delta$  over  $U \in_R \mathcal{U}$

$$|q_U(x) - p_U(x)| \leq \frac{p_U(x)}{\text{poly}(n)} + \frac{\epsilon}{2^n \delta} \left(1 + \frac{1}{\text{poly}(n)}\right)$$

$$\Pr_{U \in \mathcal{U}} [p_U(x) \geq \frac{\epsilon}{2^n \delta}] \geq \gamma$$

# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .



$$|q_U(x) - p_U(x)| \leq \text{const.} \cdot p_U(x)$$

Approximate sampling:  $|q_U(x) - p_U(x)| \leq \epsilon$

$$\text{Stockmeyer error: } |q_U(x) - p'_U(x)| \leq \frac{p'_U(x)}{\text{poly}(n)}$$

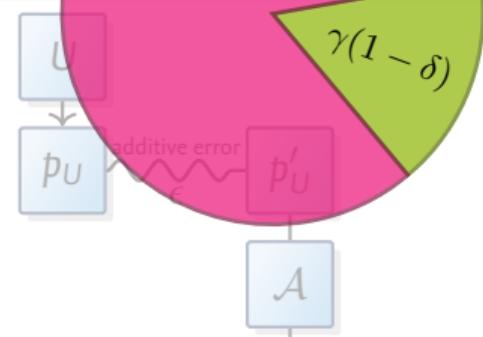
With probability  $1 - \delta$  over  $U \in_R \mathcal{U}$

$$|q_U(x) - p_U(x)| \leq \frac{p_U(x)}{\text{poly}(n)} + \frac{\epsilon}{2^n \delta} \left(1 + \frac{1}{\text{poly}(n)}\right)$$

$$\Pr_{U \in \mathcal{U}} [p_U(x) \geq \frac{\epsilon}{2^n \delta}] \geq \gamma$$

# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .



$$|q_U(x) - p_U(x)| \leq \text{const.} \cdot p_U(x)$$

Approximate sampling:  $|q_U(x) - p_U(x)| \leq \epsilon$

$$\text{Stockmeyer error: } |q_U(x) - p'_U(x)| \leq \frac{p'_U(x)}{\text{poly}(n)}$$

If  $p_U(x)$  is **GapP-hard to estimate** for any  $\gamma(1 - \delta)$  fraction of the instances, then ...

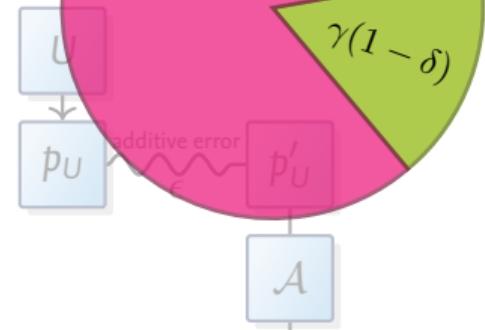
With probability  $1 - \delta$  over  $U \in_R \mathcal{U}$

$$|q_U(x) - p_U(x)| \leq \frac{p_U(x)}{\text{poly}(n)} + \frac{\epsilon}{2^n \delta} \left(1 + \frac{1}{\text{poly}(n)}\right)$$

$$\Pr_{U \in \mathcal{U}} [p_U(x) \geq \frac{\epsilon}{2^n \delta}] \geq \gamma$$

# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that takes a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .



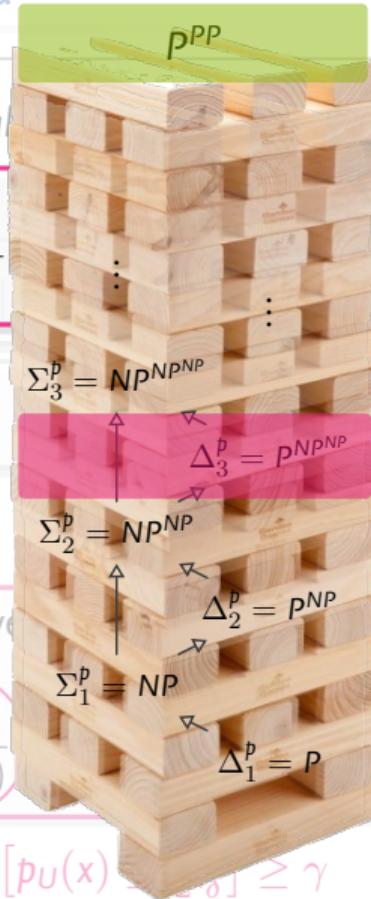
If  $p_U(x)$  is **GapP-hard to estimate** for any  $\gamma(1 - \delta)$  fraction of the instances, then ...

$$q_{U'}(x)$$

Approximate sampling

Stockmeyer error:  $|q_U(x) - p_U(x)| \leq \frac{\epsilon}{\text{poly}(n)}$

$$\Pr_{U \in \mathcal{U}} [p_U(x) > \frac{\epsilon}{2\delta}] \geq \gamma$$

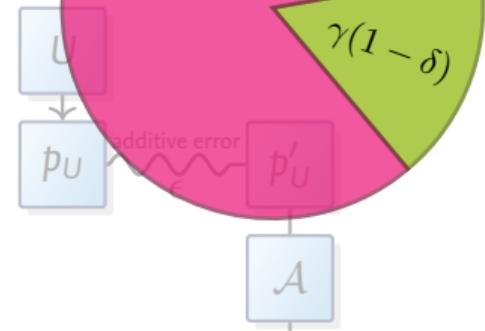


comes from a

$x)$

# Approximate quantum sampling is classically hard

Assume there exists a *classical polynomial-time, derandomizable algorithm*  $\mathcal{A}$  that samples from a distribution  $p'_U$  such that  $\|p'_U - p_U\|_{\ell_1} \leq \epsilon$ .

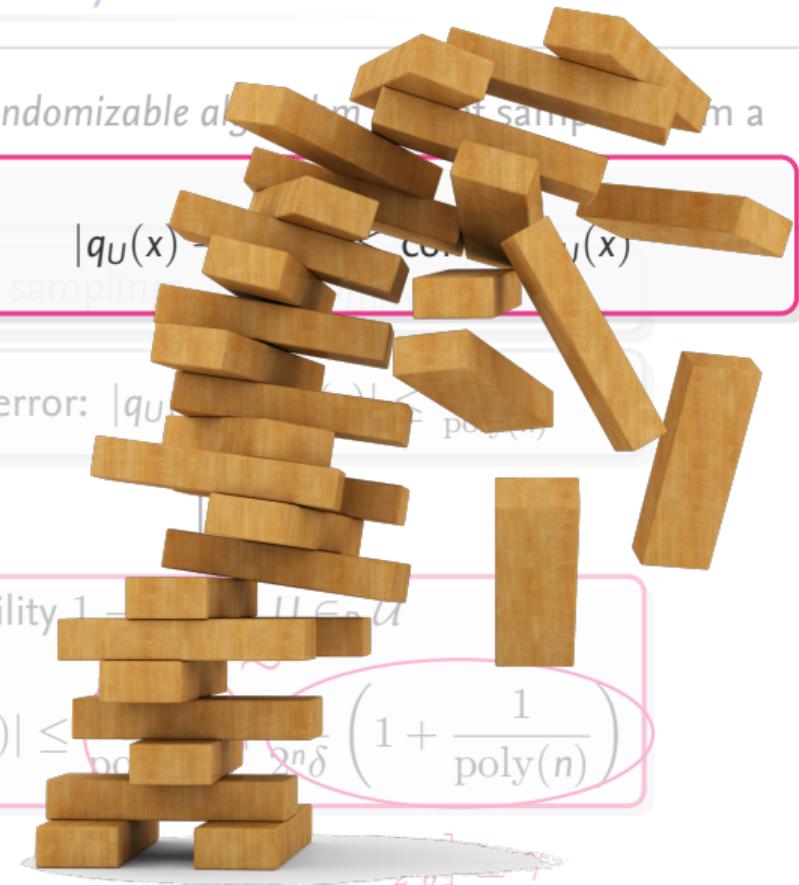


If  $p_U(x)$  is **GapP-hard to estimate** for any  $\gamma(1 - \delta)$  fraction of the instances, then ...

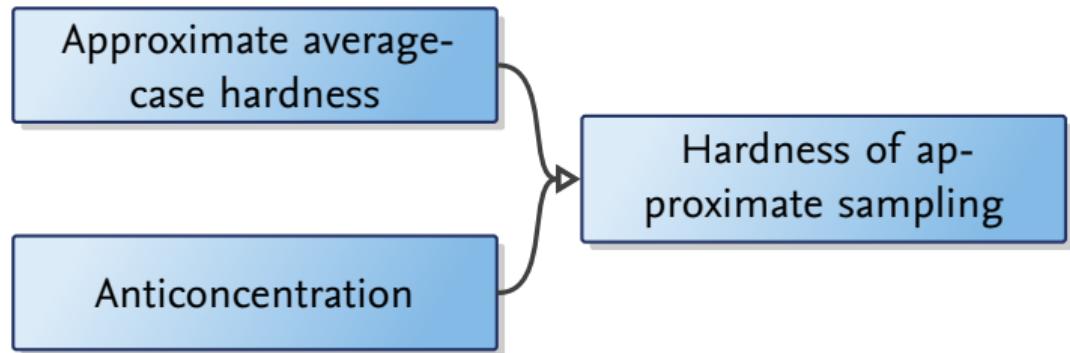
$q_{U'}(x)$

Approximate sampling error:  $|q_{U'}(x) - p_U(x)| \leq \epsilon$

Stockmeyer error:  $|q_{U'}(x) - q_U(x)| \leq \frac{\epsilon}{n\delta} \left(1 + \frac{1}{\text{poly}(n)}\right)$



# Closing loopholes: the complexity-theoretic argument



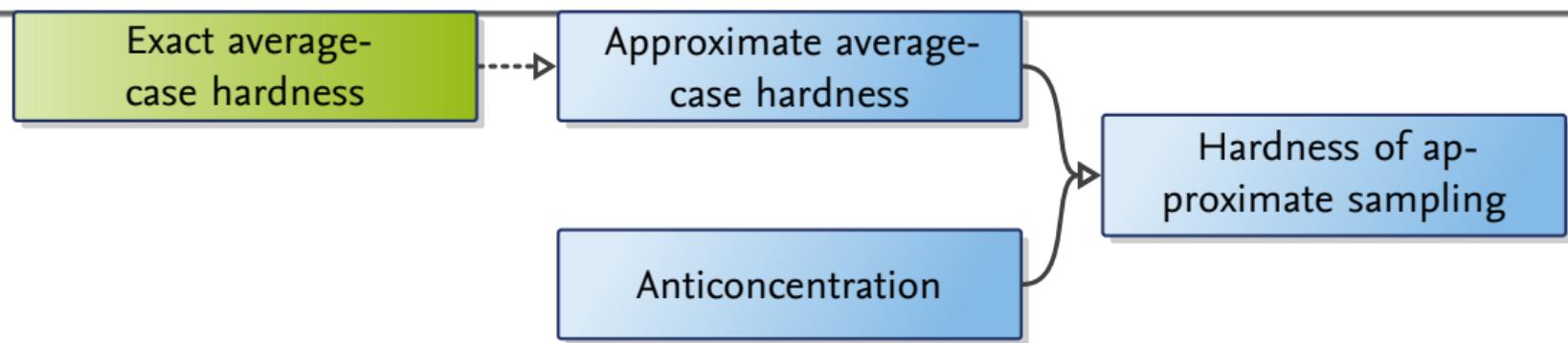
# Closing loopholes: the complexity-theoretic argument

Theorem (Approximate worst-case hardness )

*It is #P-hard to **approximate** the output probabilities of the RQIS scheme to within **relative error  $1/4$** .*

Theorem (Exact average-case hardness )

*It is #P-hard to **exactly compute** any  $3/4 + 1/\text{poly}(n)$  fraction of the output probabilities of the RQIS scheme.*



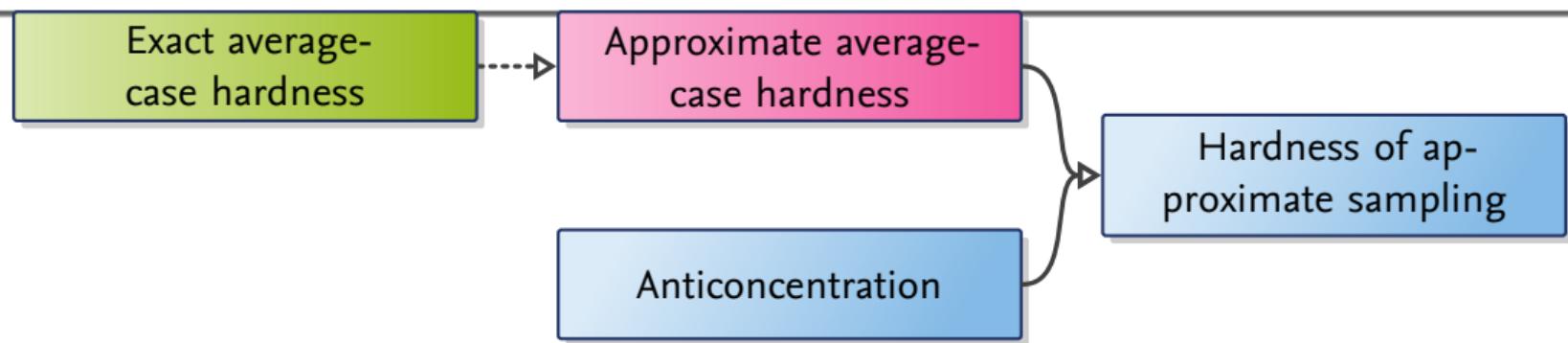
# Closing loopholes: the complexity-theoretic argument

Theorem (Approximate worst-case hardness )

*It is #P-hard to approximate the output probabilities of the RQIS scheme to within relative error  $1/4$ .*

Theorem (Exact average-case hardness )

*It is #P-hard to exactly compute any  $3/4 + 1/\text{poly}(n)$  fraction of the output probabilities of the RQIS scheme.*

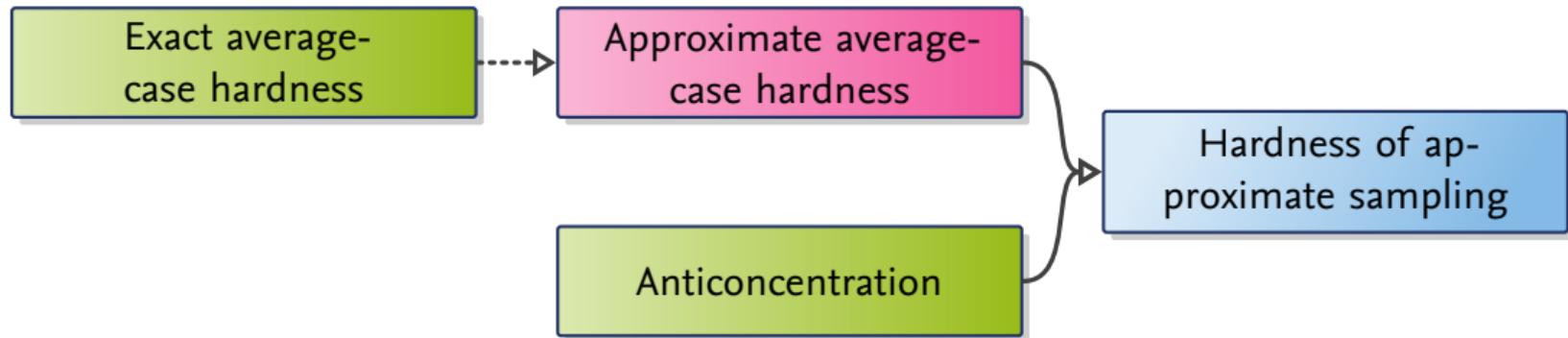


# Closing loopholes: the complexity-theoretic argument

## Theorem (Anticoncentration)

If a circuit family  $\mathcal{U}$  satisfies a second moment bound, then its output probabilities  $|\langle x|U|0\rangle|^2$  anticoncentrate in the following sense: There exist constants  $\alpha, \gamma(\alpha)$  such that

$$\Pr_{U \in_{\text{rand}} \mathcal{U}} \left[ |\langle x|U|0\rangle|^2 > \frac{\alpha}{2^n} \right] \geq \gamma(\alpha)$$

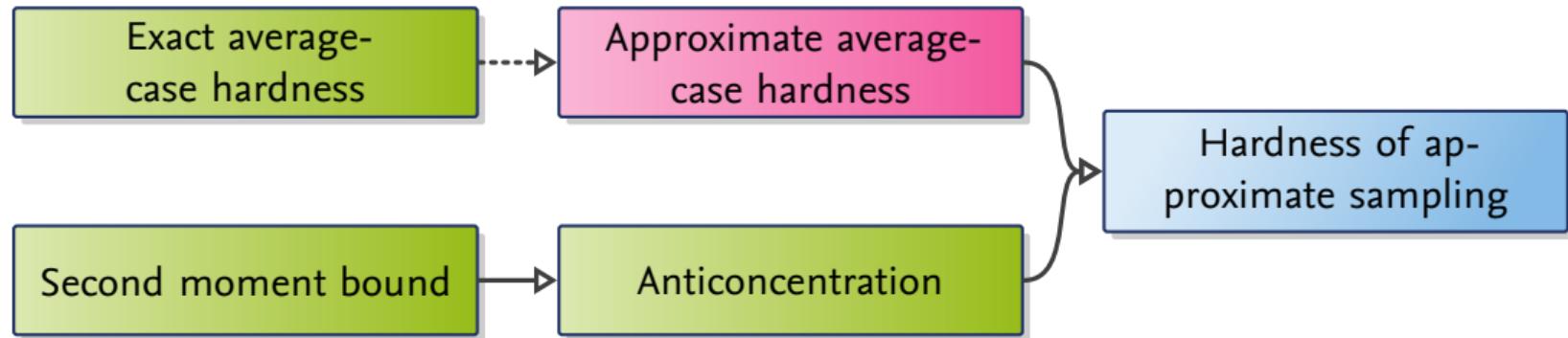


# Closing loopholes: the complexity-theoretic argument

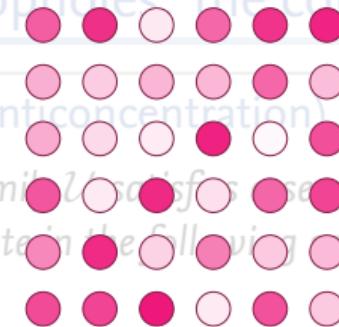
## Theorem (Anticoncentration)

If a circuit family  $\mathcal{U}$  satisfies a second moment bound, then its output probabilities  $|\langle x|U|0\rangle|^2$  anticoncentrate in the following sense: There exist constants  $\alpha, \gamma(\alpha)$  such that

$$\Pr_{U \in_{rand} \mathcal{U}} \left[ |\langle x|U|0\rangle|^2 > \frac{\alpha}{2^n} \right] \geq \gamma(\alpha) \iff \Pr[Z \geq \alpha \mathbb{E}[Z]] \geq (1 - \alpha)^2 \frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]}$$

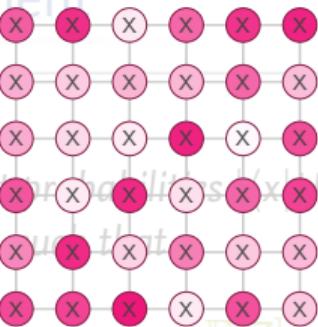
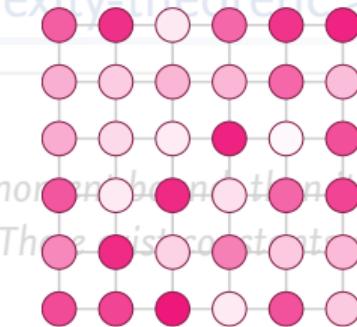


# Closing loopholes: the complexity theoretic argument



Theorem (Anticoncentration)

If a circuit family  $\mathcal{U}$  satisfies second moment bounds on its output probabilities  $|\langle x|U|0\rangle|^2$  anticoncentrate in the following sense: There is a constant  $\alpha, \gamma(\alpha)$  such that

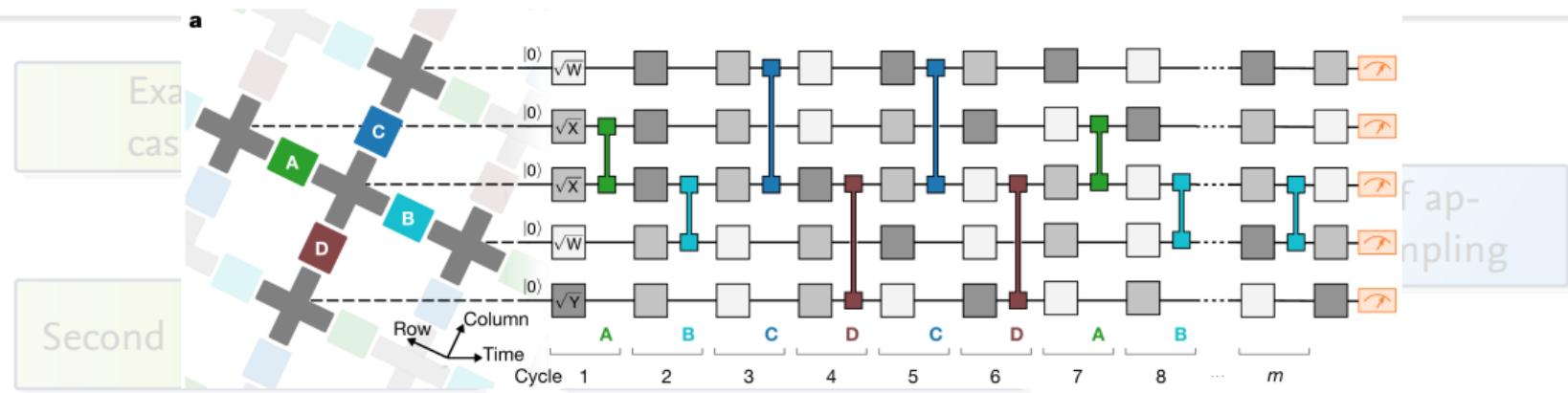


$$\Pr_{U \in \text{rand } \mathcal{U}} \left[ |\langle x|U|0\rangle|^2 \geq \frac{\alpha}{2} \right] \geq \gamma(\alpha)$$

E1: Prepare  $|\psi_\beta\rangle$

E2: Quench with  $H$

E3: Measure in  $X$

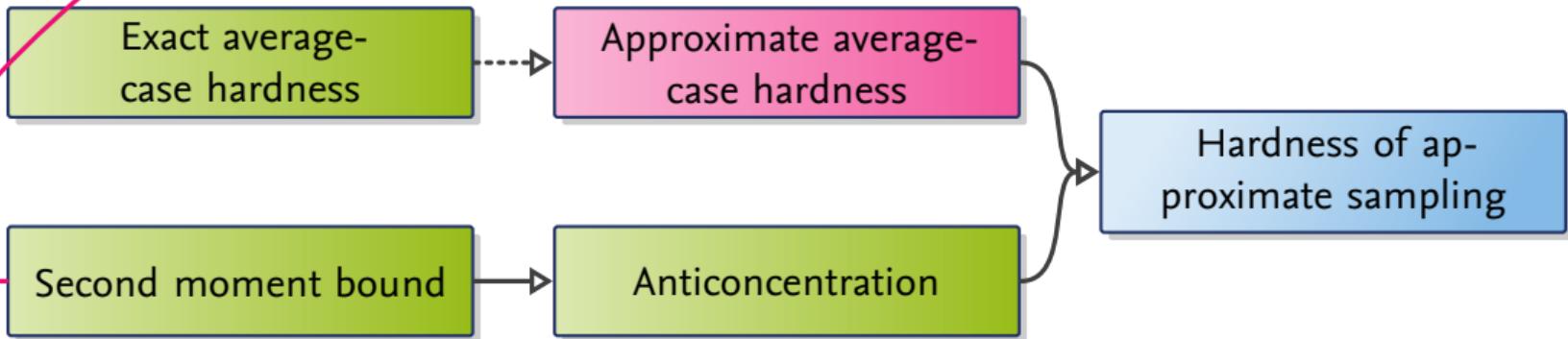
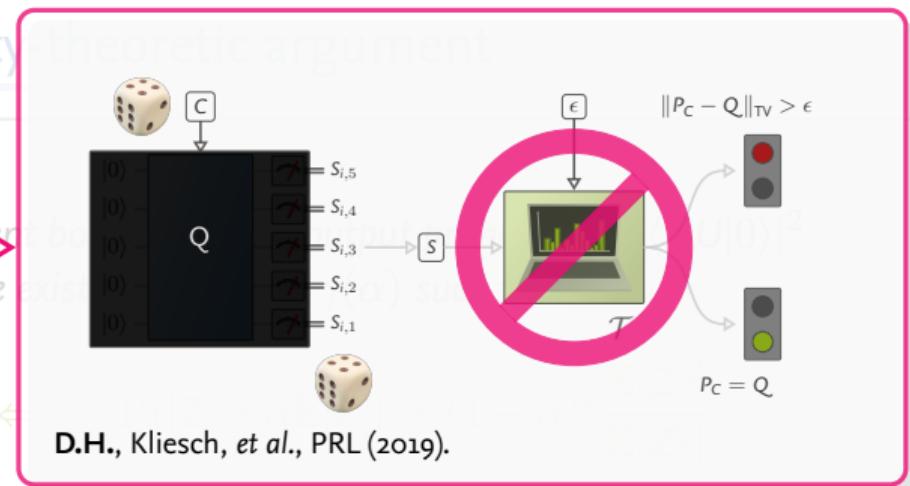


# Closing loopholes: the complexity of exact sampling

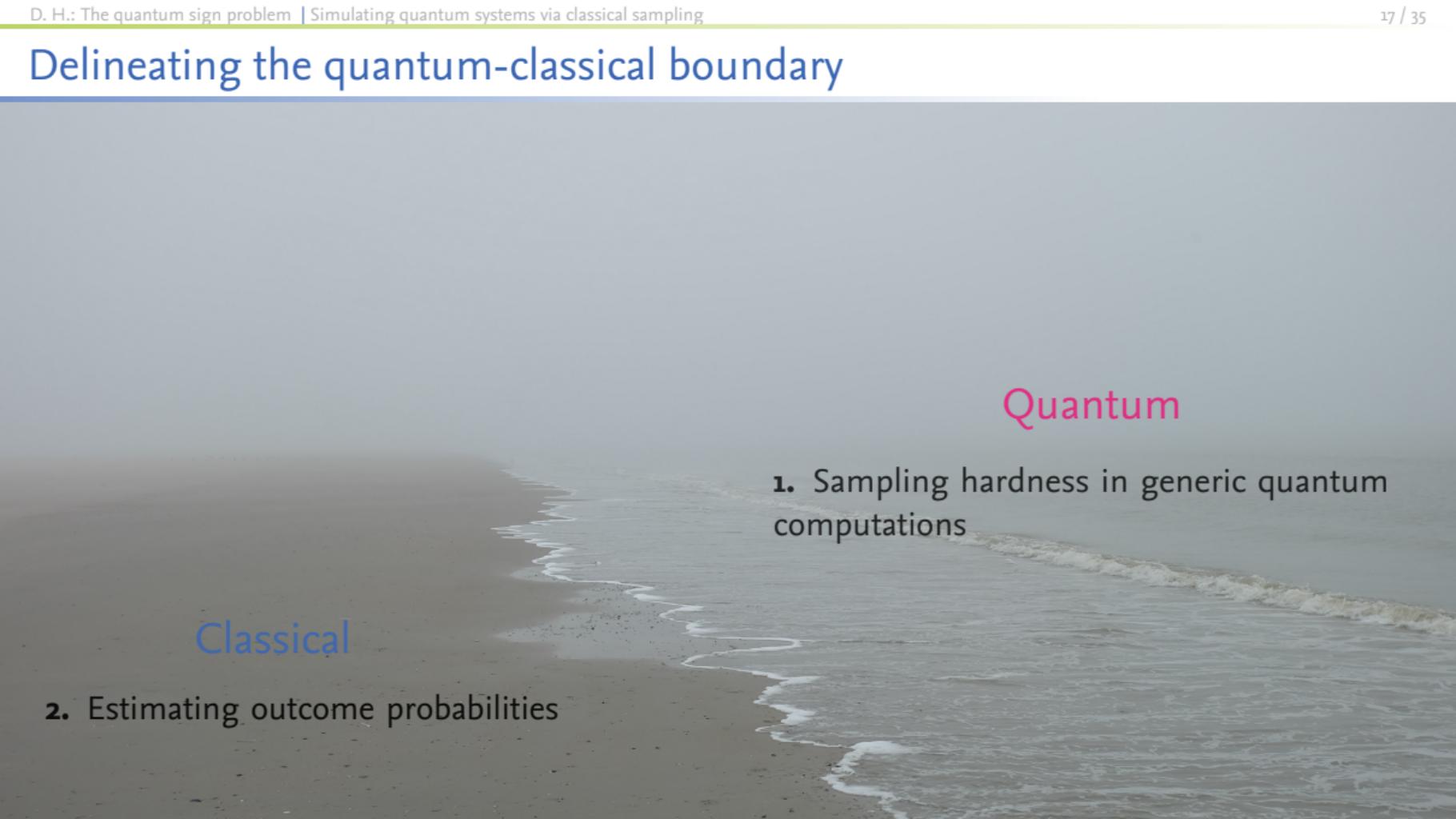
## Theorem (Anticoncentration)

If a circuit family  $\mathcal{U}$  satisfies a second moment bound, then it anticoncentrates in the following sense: There exists a constant  $\gamma(\alpha)$

$$\Pr_{U \in \mathcal{U}_{\text{rand}}} \left[ |\langle x| U |0\rangle|^2 > \frac{\alpha}{2^n} \right] \geq \gamma(\alpha)$$



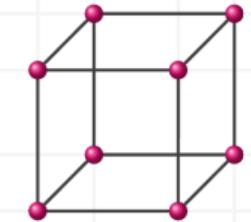
# Delineating the quantum-classical boundary



Classical

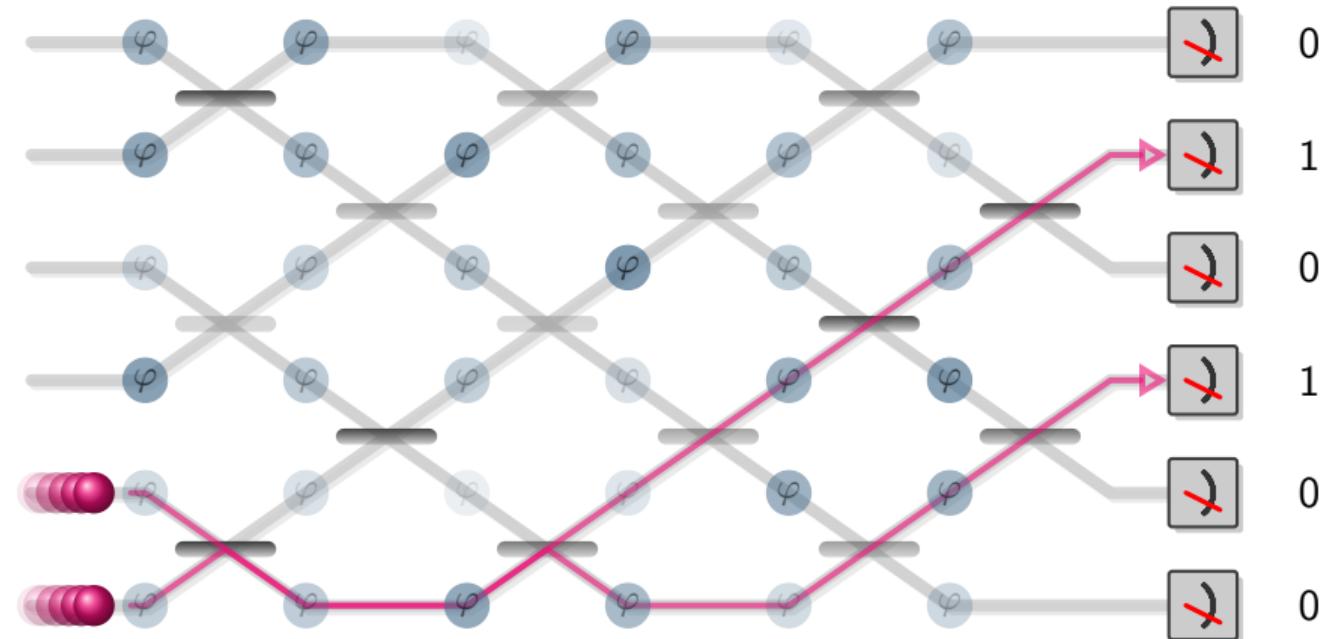
Quantum

1. Sampling hardness in generic quantum computations
2. Estimating outcome probabilities

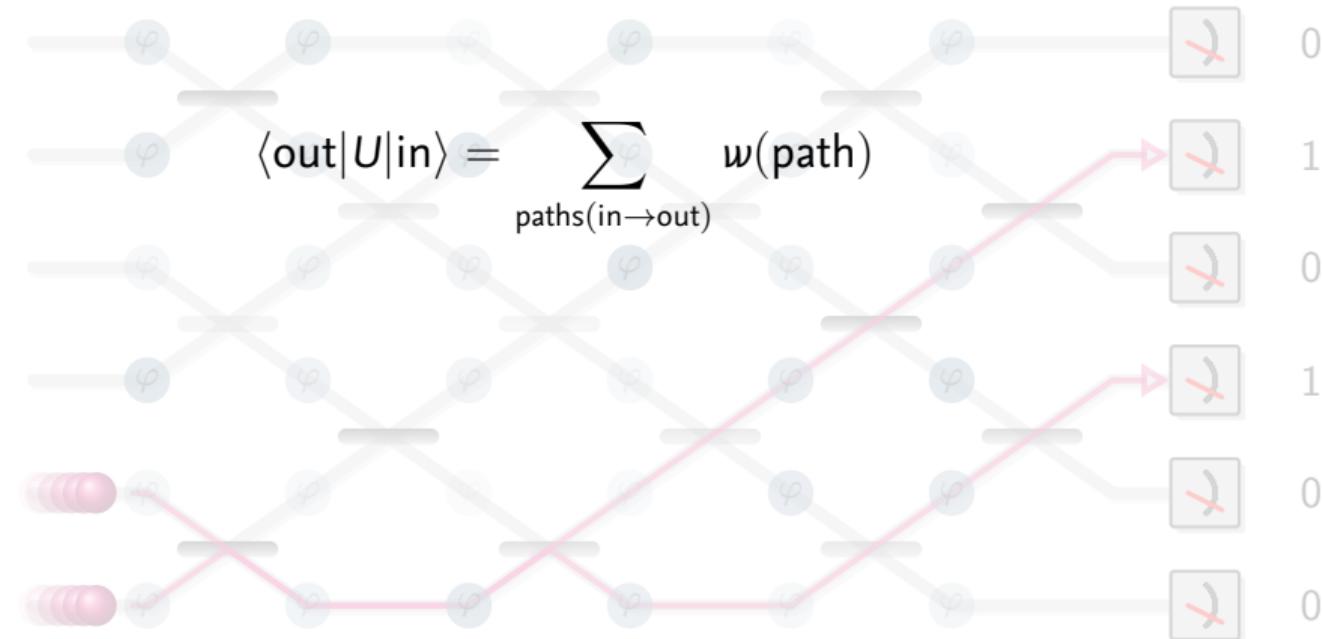


Perspective 2  
Estimating quantum properties via classical sampling

# The quantum sign problem

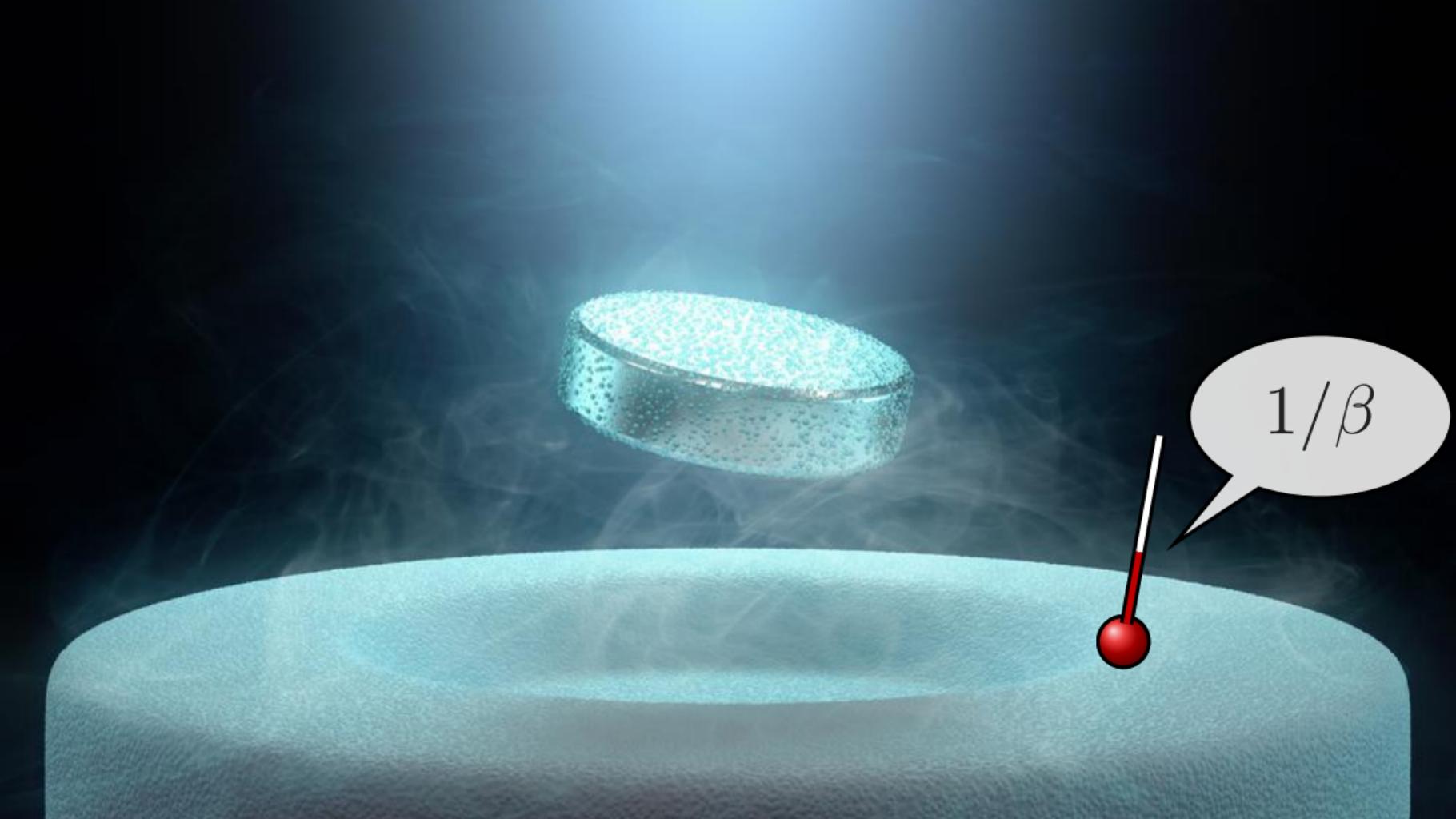


# The quantum sign problem



# Quantum Monte Carlo



 $1/\beta$

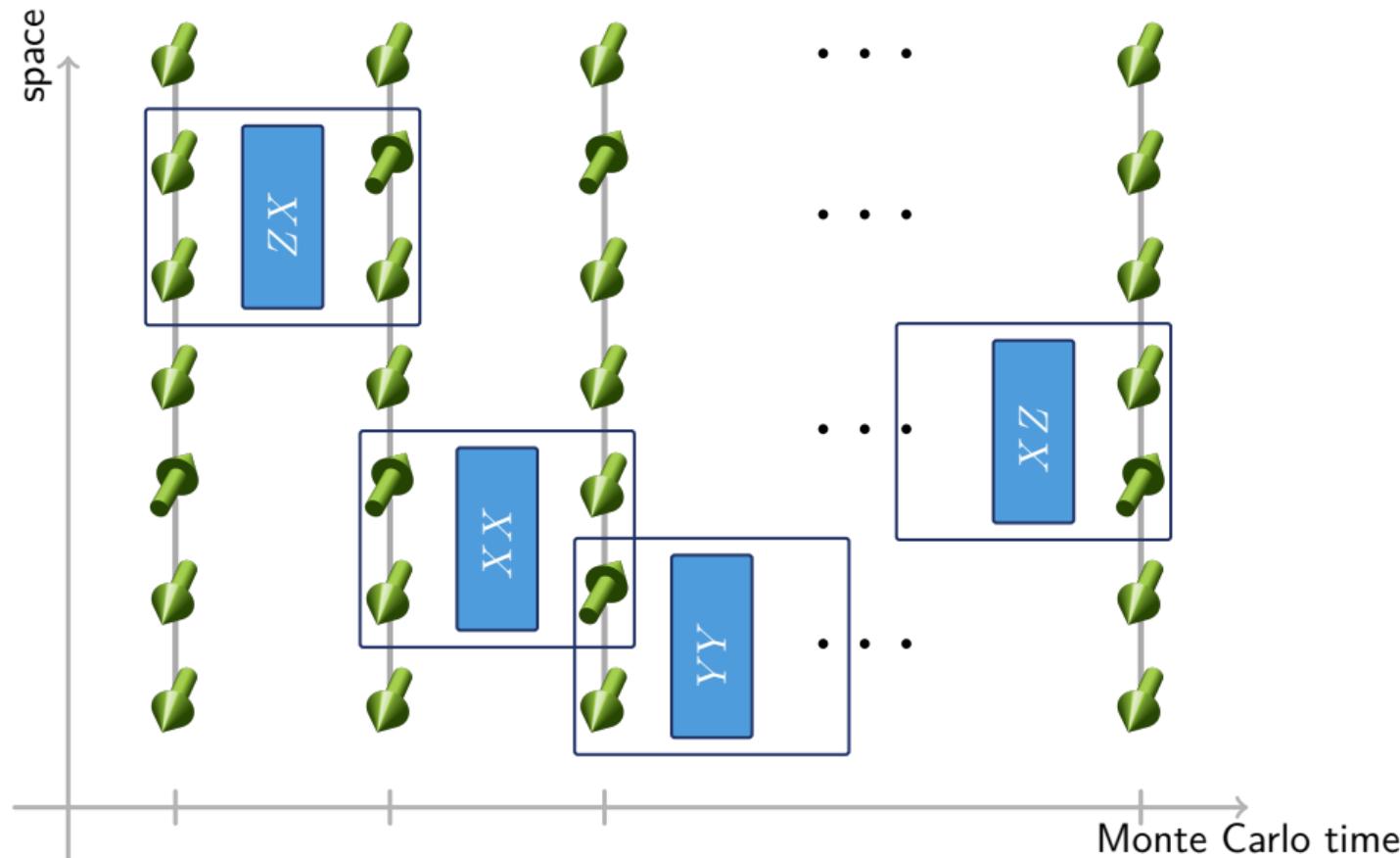




$1/\beta$



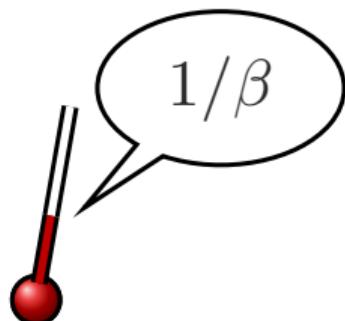
# Quantum Monte Carlo



# Quantum Monte Carlo

Calculating expectation values via expectation values

$$\langle O \rangle_{\beta, H} = \frac{1}{Z} \text{Tr}[e^{-\beta H} O] = \sum_{\lambda} q(\lambda) O(\lambda) = \langle O \rangle_q$$

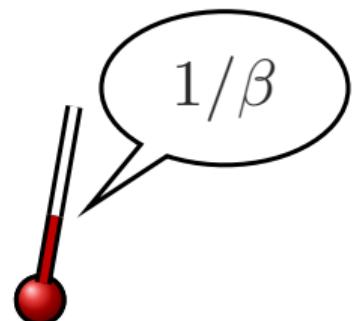


# Quantum Monte Carlo

Calculating expectation values via expectation values

$$\langle O \rangle_{\beta, H} = \frac{1}{Z} \text{Tr}[e^{-\beta H} O] = \sum_{\lambda} q(\lambda) O(\lambda) = \langle O \rangle_q$$

$$e^{-\beta H} = \left( e^{-\beta H/m} \right)^m \approx \left( 1 - \frac{\beta}{m} H \right)^m =: (T_m)^m$$



# Quantum Monte Carlo

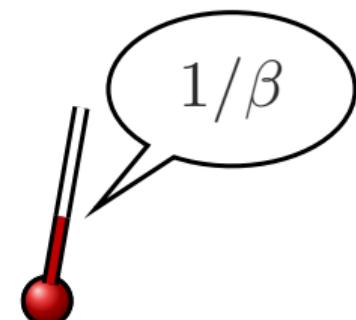
## Calculating expectation values via expectation values

$$\langle O \rangle_{\beta, H} = \frac{1}{Z} \text{Tr}[e^{-\beta H} O] = \sum_{\lambda} q(\lambda) O(\lambda) = \langle O \rangle_q$$

$$e^{-\beta H} = \left( e^{-\beta H/m} \right)^m \approx \left( 1 - \frac{\beta}{m} H \right)^m =: (T_m)^m$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\begin{aligned} \langle O \rangle_{\beta, H} &= \frac{1}{Z} \text{Tr}[e^{-\beta H} O] \approx \frac{1}{Z} \text{Tr}[T_m^m O] \\ &= \frac{1}{Z} \sum_{\lambda} T_m(\lambda_1 | \lambda_2) T_m(\lambda_2 | \lambda_3) \cdots T_m(\lambda_m | \lambda_{m+1}) O(\lambda_{m+1} | \lambda_1) \\ &\equiv \frac{1}{Z} \sum_{\lambda} a(\lambda) O(\lambda) \end{aligned}$$



# Quantum Monte Carlo

Calculating expectation values via expectation values

$$\langle O \rangle_{\beta, H} = \frac{1}{Z} \text{Tr}[e^{-\beta H} O] = \sum_{\lambda} q(\lambda) O(\lambda) = \langle O \rangle_q$$

$$e^{-\beta H} = \left( e^{-\beta H/m} \right)^m \approx \left( 1 - \frac{\beta}{m} H \right)^m =: (T_m)^m$$

$$\text{II} \quad 1/\beta$$

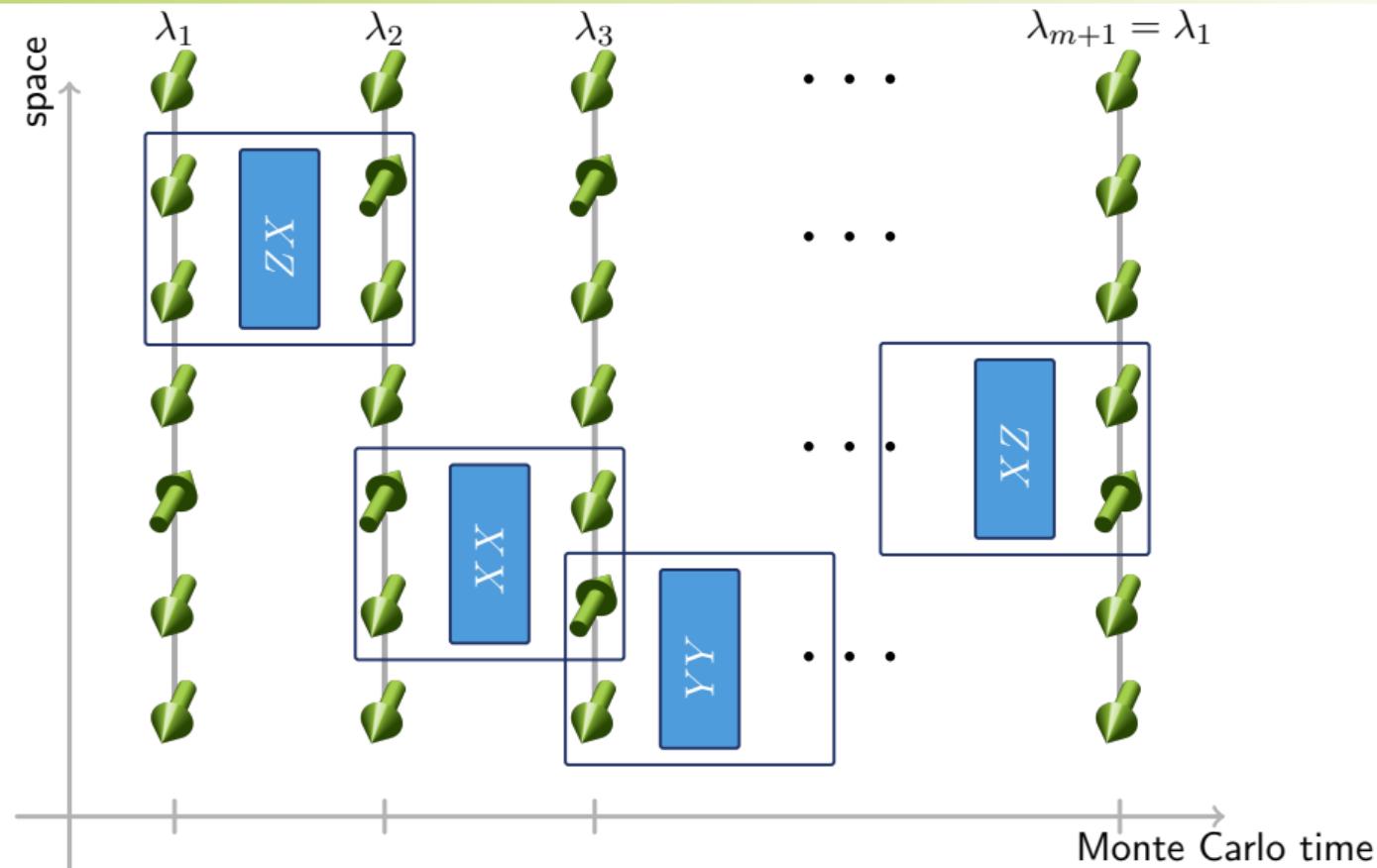
Monte Carlo sampling



$\langle O \rangle_{\beta}$  Draw  $\lambda$  with probability  $q(\lambda) = \frac{a(\lambda)}{\sum_{\lambda} a(\lambda)}$ .

- Metropolis sampling with transition rate  $W_{\lambda \rightarrow \lambda'} = q(\lambda')/q(\lambda)$ :
- Markov chain:  $\lambda^{(1)} \rightarrow \lambda^{(2)} \rightarrow \dots \rightarrow \lambda^{(s)}$

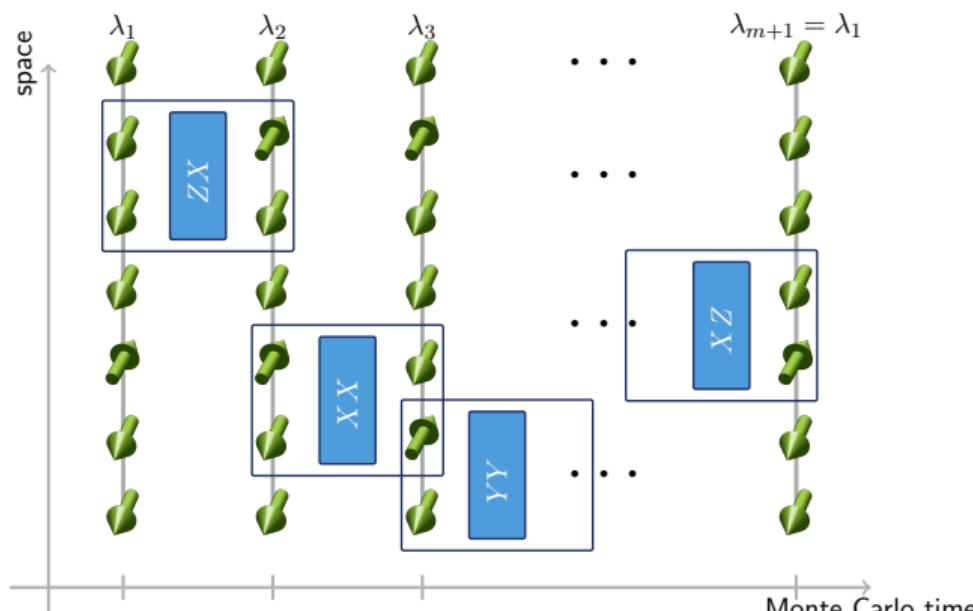
$$\equiv \frac{1}{Z} \sum_{\lambda} a(\lambda) O(\lambda)$$



$$a(\lambda) = T_m(\lambda_1|\lambda_2) \cdot T_m(\lambda_2|\lambda_3) \cdot T_m(\lambda_3|\lambda_4) \cdot \dots \cdot T_m(\lambda_m|\lambda_1)$$

# The Monte Carlo sign problem

$$\langle O \rangle_{\beta, H} = \langle O \rangle_q$$



$$a(\lambda) = T_m(\lambda_1|\lambda_2) \cdot T_m(\lambda_2|\lambda_3) \cdot T_m(\lambda_3|\lambda_4) \cdot \dots \cdot T_m(\lambda_m|\lambda_1)$$

# The Monte Carlo sign problem

$$\langle O \rangle_{\beta, H} = \langle O \rangle_q = \sum_{\lambda} p(\lambda) O'(\lambda)$$

The best Monte Carlo estimator: Absolute value

The variance-optimal estimator

$$q(\lambda) \rightarrow p(\lambda) = |q(\lambda)| / \|q(\lambda)\|_{\ell_1}$$

$$O(\lambda) \rightarrow O'(\lambda) = \text{sign}(q(\lambda)) \|q(\lambda)\|_{\ell_1} O(\lambda)$$

# The Monte Carlo sign problem

$$\langle O \rangle_{\beta, H} = \langle O \rangle_q = \sum_{\lambda} p(\lambda) O'(\lambda) \approx \frac{1}{M} \frac{1}{\langle \text{sign} \rangle} \sum_{\lambda_1, \dots, \lambda_M} \text{sign}(q(\lambda_i))$$

The best Monte Carlo estimator: Absolute value

The variance-optimal estimator

$$q(\lambda) \rightarrow p(\lambda) = |q(\lambda)| / \|q(\lambda)\|_{\ell_1}$$

$$O(\lambda) \rightarrow O'(\lambda) = \text{sign}(q(\lambda)) \|q(\lambda)\|_{\ell_1} O(\lambda)$$



$$\lambda_i \sim_{\text{rand}} p$$

# The Monte Carlo sign problem

$$\langle O \rangle_{\beta, H} = \langle O \rangle_q = \sum_{\lambda} p(\lambda) O'(\lambda) \approx \frac{1}{M} \frac{1}{\langle \text{sign} \rangle} \sum_{\lambda_1, \dots, \lambda_M} \text{sign}(q(\lambda_i)) \pm \frac{1}{\sqrt{M}} \frac{1}{\langle \text{sign} \rangle}$$

The best Monte Carlo estimator: Absolute value  
 The variance-optimal estimator



$$\lambda_i \sim_{\text{rand}} p$$

$$q(\lambda) \rightarrow p(\lambda) = |q(\lambda)| / \|q(\lambda)\|_{\ell_1}$$

$$O(\lambda) \rightarrow O'(\lambda) = \text{sign}(q(\lambda)) \|q(\lambda)\|_{\ell_1} O(\lambda)$$

has sample complexity

$$s \sim \|q\|_{\ell_1}^2 - 1 \equiv \frac{1}{\langle \text{sign} \rangle_q^2} - 1 \quad \text{with} \quad \langle \text{sign} \rangle_q = \frac{\sum_{\lambda} |q(\lambda)| \text{sign}(q(\lambda))}{\sum_{\lambda} |q(\lambda)|}.$$

# The Monte Carlo sign problem

$$\langle O \rangle_{\beta, H} = \langle O \rangle_q = \sum_{\lambda} p(\lambda) O'(\lambda) \approx \frac{1}{M} \frac{1}{\langle \text{sign} \rangle} \sum_{\lambda_1, \dots, \lambda_M} \text{sign}(q(\lambda_i)) \pm \frac{1}{\sqrt{M}} \frac{1}{\langle \text{sign} \rangle}$$

The best Monte Carlo estimator: Absolute value

The variance-optimal estimator

$$q(\lambda) \rightarrow p(\lambda) = |q(\lambda)| / \|q(\lambda)\|_{\ell_1}$$

$$O(\lambda) \rightarrow O'(\lambda) = \text{sign}(q(\lambda)) \|q(\lambda)\|_{\ell_1} O(\lambda)$$



$$\lambda_i \sim_{\text{rand}} p$$

has sample complexity

$$s \sim \|q\|_{\ell_1}^2 - 1 \equiv \frac{1}{\langle \text{sign} \rangle_q^2} - 1 \in \exp(n) \quad \text{with} \quad \langle \text{sign} \rangle_q = \frac{\sum_{\lambda} |q(\lambda)| \text{sign}(q(\lambda))}{\sum_{\lambda} |q(\lambda)|}.$$

# The Monte Carlo sign problem

## The Monte Carlo sign problem

The best Monte Carlo estimator: Absolute value

The variance-optimal estimator

$$q(\lambda) \rightarrow p(\lambda) = |q(\lambda)|/\|q(\lambda)\|_{\ell_1}$$

$$O(\lambda) \rightarrow O'(\lambda) = \text{sign}(q(\lambda))\|q(\lambda)\|_{\ell_1} O(\lambda)$$



$$\lambda_i \sim_{\text{rand}} p$$

has sample complexity

$$s \sim \|q\|_{\ell_1}^2 - 1 \equiv \frac{1}{\langle \text{sign} \rangle_q^2} - 1 \in \exp(n) \quad \text{with} \quad \langle \text{sign} \rangle_q = \frac{\sum_{\lambda} |q(\lambda)| \text{sign}(q(\lambda))}{\sum_{\lambda} |q(\lambda)|}.$$

# The Monte Carlo sign problem

## The Monte Carlo sign problem

- $T_m = (1 - \beta H/m)$
- $a(\lambda) = T_m(\lambda_1|\lambda_2)T_m(\lambda_2|\lambda_3) \cdots T_m(\lambda_m|\lambda_1)$   
 $\Rightarrow q(\lambda) = a(\lambda) / \sum_{\lambda} a(\lambda)$
- Sign problem:  $q(\lambda) < 0$   
 $\Rightarrow p(\lambda) = |q(\lambda)| / \sum_{\lambda} |q(\lambda)|$
- Exponential increase in sample complexity!

has sample complexity

$$s \sim \|q\|_{\ell_1}^2 - 1 \equiv \frac{1}{\langle \text{sign} \rangle_q^2} - 1 \in \exp(n) \quad \text{with} \quad \langle \text{sign} \rangle_q = \frac{\sum_{\lambda} |q(\lambda)| \text{sign}(q(\lambda))}{\sum_{\lambda} |q(\lambda)|}.$$

# Easing the sign problem

## Sign problem 2: Invariance under basis transformations

$$\text{Tr}[e^{-\beta H} O] = \text{Tr}[Z e^{-\beta H} O Z^{-1}] = \text{Tr}[(Z T_m Z^{-1})^m Z O Z^{-1}].$$

→→ The sign problem is a **basis-dependent** property. ←←

Bravyi, DiVincenzo, Oliveira, Terhal:  
arXiv:quant-ph/0606140.  
Bravyi, Terhal: arXiv:0806.1746.  
Cubitt, Montanaro: arXiv:1311.3161.

Troyer and Wiese, PRL, 2008.  
Hastings, J. Math. Phys (2015)

Marvian, Hen, Lidar, arXiv:1802.03408.  
Klassen and Terhal, arXiv:1806.05405.  
Klassen, Marvian *et al.*, arXiv:1906.08800.

# Easing the sign problem

## Sign problem 2: Invariance under basis transformations

$$\text{Tr}[e^{-\beta H} O] = \text{Tr}[Z e^{-\beta H} O Z^{-1}] = \text{Tr}[(Z T_m Z^{-1})^m Z O Z^{-1}].$$

→→ The sign problem is a **basis-dependent** property. ←←



Bravyi, DiVincenzo, Oliveira, Terhal:  
arXiv:quant-ph/0606140.  
Bravyi, Terhal: arXiv:0806.1746.  
Cubitt, Montanaro: arXiv:1311.3161.

Troyer and Wiese, PRL, 2008.  
Hastings, J. Math. Phys (2015)

Marvian, Hen, Lidar, arXiv:1802.03408.  
Klassen and Terhal, arXiv:1806.05405.  
Klassen, Marvian *et al.*, arXiv:1906.08800.

# Easing the sign problem

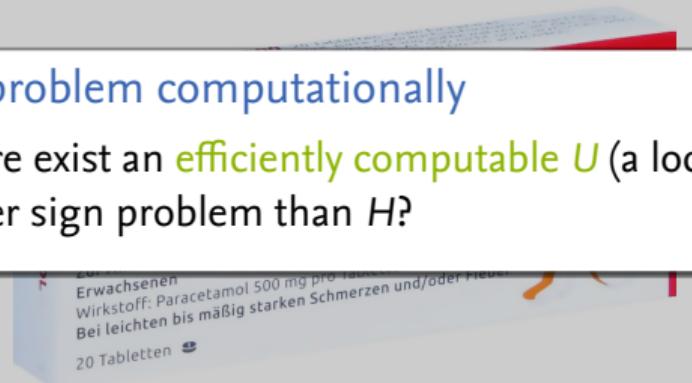
## Sign problem 2: Invariance under basis transformations

$$\mathrm{Tr}[e^{-\beta H} O] = \mathrm{Tr}[Ze^{-\beta H} O Z^{-1}] = \mathrm{Tr}[(ZT_mZ^{-1})^m ZOZ^{-1}].$$

→→ The sign problem is a **basis-dependent** property. ←←

### Easing the sign problem computationally

Given  $H$ , does there exist an **efficiently computable  $U$**  (a local circuit) such that  $UHU^\dagger$  has a smaller sign problem than  $H$ ?



Bravyi, DiVincenzo, Oliveira, Terhal:  
arXiv:quant-ph/0606140.  
Bravyi, Terhal: arXiv:0806.1746.  
Cubitt, Montanaro: arXiv:1311.3161.

Troyer and Wiese, PRL, 2008.  
Hastings, J. Math. Phys (2015)

Marvian, Hen, Lidar, arXiv:1802.03408.  
Klassen and Terhal, arXiv:1806.05405.  
Klassen, Marvian *et al.*, arXiv:1906.08800.

## What is the optimal basis?

Recall the sample complexity of QMC:  $s \sim \langle \text{sign} \rangle_p^{-2} - 1$ .

## What is the optimal basis?

Recall the sample complexity of QMC:  $s \sim \langle \text{sign} \rangle_p^{-2} - 1$ .

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]$$

# What is the optimal basis?

Recall the sample complexity of QMC:  $s \sim \langle \text{sign} \rangle_p^{-2} - 1$ .

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]$$

Computing the average sign via Monte Carlo

$$\frac{\text{Var}[\langle \text{sign} \rangle_q]}{\langle \text{sign} \rangle^2} = \frac{1}{\langle \text{sign} \rangle^2} - 1$$

## What is the optimal basis?

Recall the sample complexity of QMC:  $s \sim \langle \text{sign} \rangle_p^{-2} - 1$ .

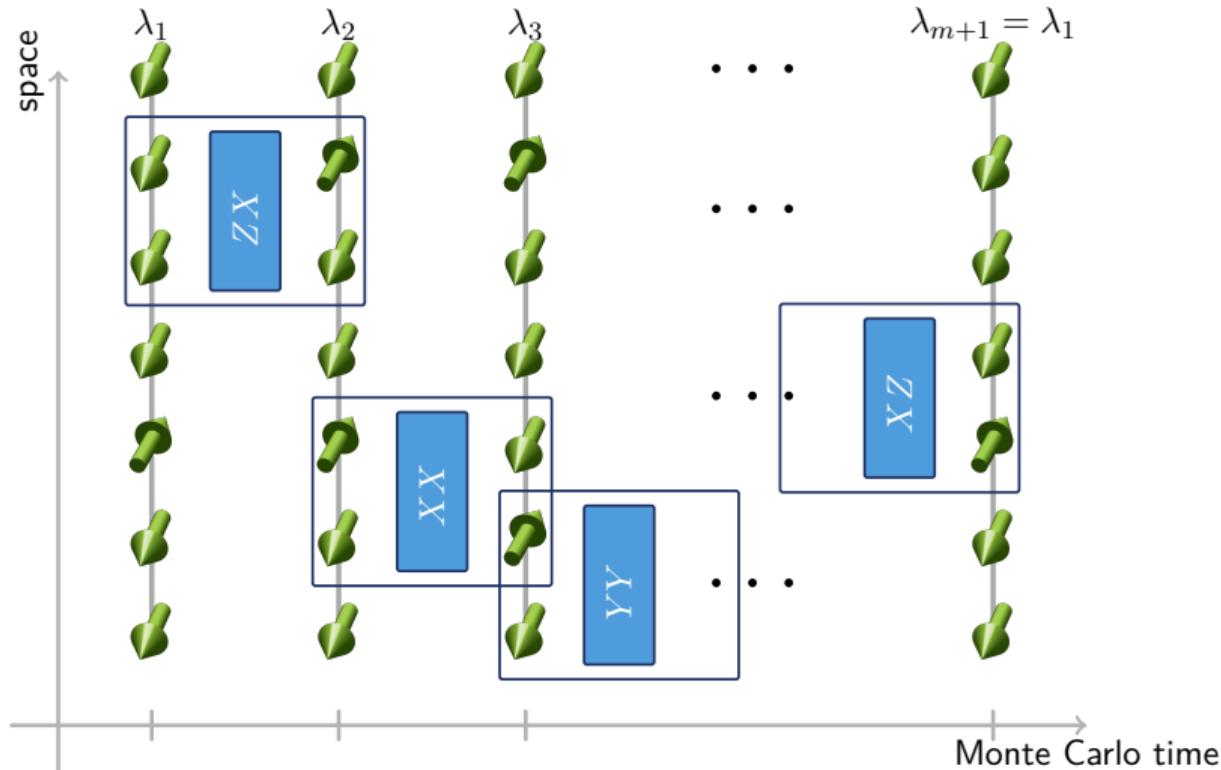
Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]$$

Computing the average sign via Monte Carlo

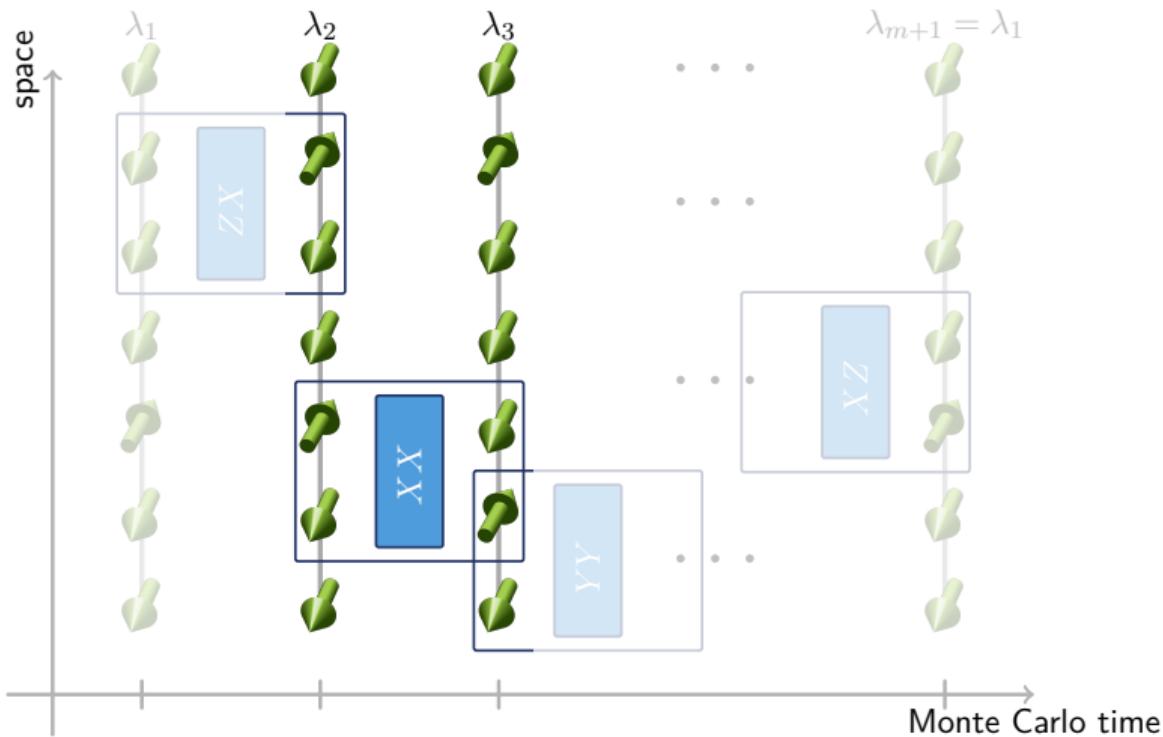
$$\frac{\text{Var}[\langle \text{sign} \rangle_q]}{\langle \text{sign} \rangle^2} = \frac{1}{\langle \text{sign} \rangle^2} - 1$$

# What is the optimal basis?



$$a(\lambda) = |T_m(\lambda_1|\lambda_2)| \cdot |T_m(\lambda_2|\lambda_3)| \cdot |T_m(\lambda_3|\lambda_4)| \cdots \cdot |T_m(\lambda_m|\lambda_1)|$$

# What is the optimal basis?



$$a(\lambda) = T_m(\lambda_1|\lambda_2) \cdot T_m(\lambda_2|\lambda_3) \cdot T_m(\lambda_3|\lambda_4) \cdot \dots \cdot T_m(\lambda_m|\lambda_1)$$

## Non-stoquasticity

$$\begin{pmatrix} - & + & + & 0 \\ + & + & 0 & - \\ + & 0 & - & 0 \\ 0 & - & 0 & 0 \end{pmatrix}$$

## Non-stoquasticity

$$\begin{pmatrix} - & + & + & 0 \\ + & + & 0 & - \\ + & 0 & - & 0 \\ 0 & - & 0 & 0 \end{pmatrix}$$

## Non-stoquasticity

$$\left| \begin{array}{c} | \\ | \\ \left( \begin{array}{cccc} - & + & + & 0 \\ + & + & 0 & - \\ + & 0 & - & 0 \\ 0 & - & 0 & 0 \end{array} \right) \\ | \\ | \end{array} \right| \ell_1$$

## Easing the sign problem: average sign vs. non-stoquasticity

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]$$

## Easing the sign problem: average sign vs. non-stoquasticity

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]$$

The non-stoquasticity

$$\nu_1(UHU^\dagger) := \||UT_m U^\dagger| - UT_m U^\dagger\|_{\ell_1} = \|(UHU^\dagger)_{\text{non-stoq.}}\|_{\ell_1}$$

can be *computed efficiently* for 2-local Hamiltonians

## Easing the sign problem: average sign vs. non-stoquasticity

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \underbrace{\left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]}_{S(U)}$$

The non-stoquasticity

$$\nu_1(UHU^\dagger) := \| |UT_m U^\dagger| - UT_m U^\dagger \|_{\ell_1} = \| (UHU^\dagger)_{\text{non-stoq.}} \|_{\ell_1}$$

can be *computed efficiently* for 2-local Hamiltonians

## Average sign vs. non-stoquasticity: Analytical evidence

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \underbrace{\left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]}_{S(U)}$$

## Average sign vs. non-stoquasticity: Analytical evidence

Minimize  $1/\langle \text{sign} \rangle_p$  over  $U \in \mathcal{U}$

$$\arg \min_{U \in \mathcal{U}} \frac{1}{\langle \text{sign} \rangle_p} = \arg \min_{U \in \mathcal{U}} \text{Tr} \underbrace{\left[ |UT_m U^\dagger|^m - (UT_m U^\dagger)^m \right]}_{S(U)}$$

To first order in the *non-stoquastic matrix entries*:

$$\begin{aligned} S(U) &\approx 2m \||UT_m U^\dagger| - UT_m U^\dagger\|_{\ell_1} \cdot \|(|UT_m U^\dagger| + UT_m U^\dagger)^{m-1}\|_{\ell_\infty} \\ &\propto d \nu_1(H). \end{aligned}$$

## Average sign vs. non-stoquasticity: Analytical evidence

For generic instances, we expect

$$1/\langle \text{sign} \rangle \propto \exp(d\nu_1(H))$$



To first order in the *non-stoquastic matrix entries*:

$$\begin{aligned} S(U) &\approx 2m \||UT_m U^\dagger| - UT_m U^\dagger\|_{\ell_1} \cdot \|(|UT_m U^\dagger| + UT_m U^\dagger)^{m-1}\|_{\ell_\infty} \\ &\propto d \nu_1(H). \end{aligned}$$

## Average sign vs. non-stoquasticity: Analytical evidence

For generic instances, we expect

$$1/\langle \text{sign} \rangle \propto \exp(d\nu_1(H))$$



To first order in the *non-stoquastic matrix entries*:

$$\begin{aligned} S(U) &\approx 2m \|\|UT_mU^\dagger| - UT_mU^\dagger\|_{\ell_1} \cdot \|(|UT_mU^\dagger| + UT_mU^\dagger)^{m-1}\|_{\ell_\infty} \\ &\propto d \nu_1(H). \end{aligned}$$

## Average sign vs. non-stoquasticity: Analytical evidence

For generic instances, we expect

$$1/\langle \text{sign} \rangle \propto \exp(d\nu_1(H))$$

SPARSITY

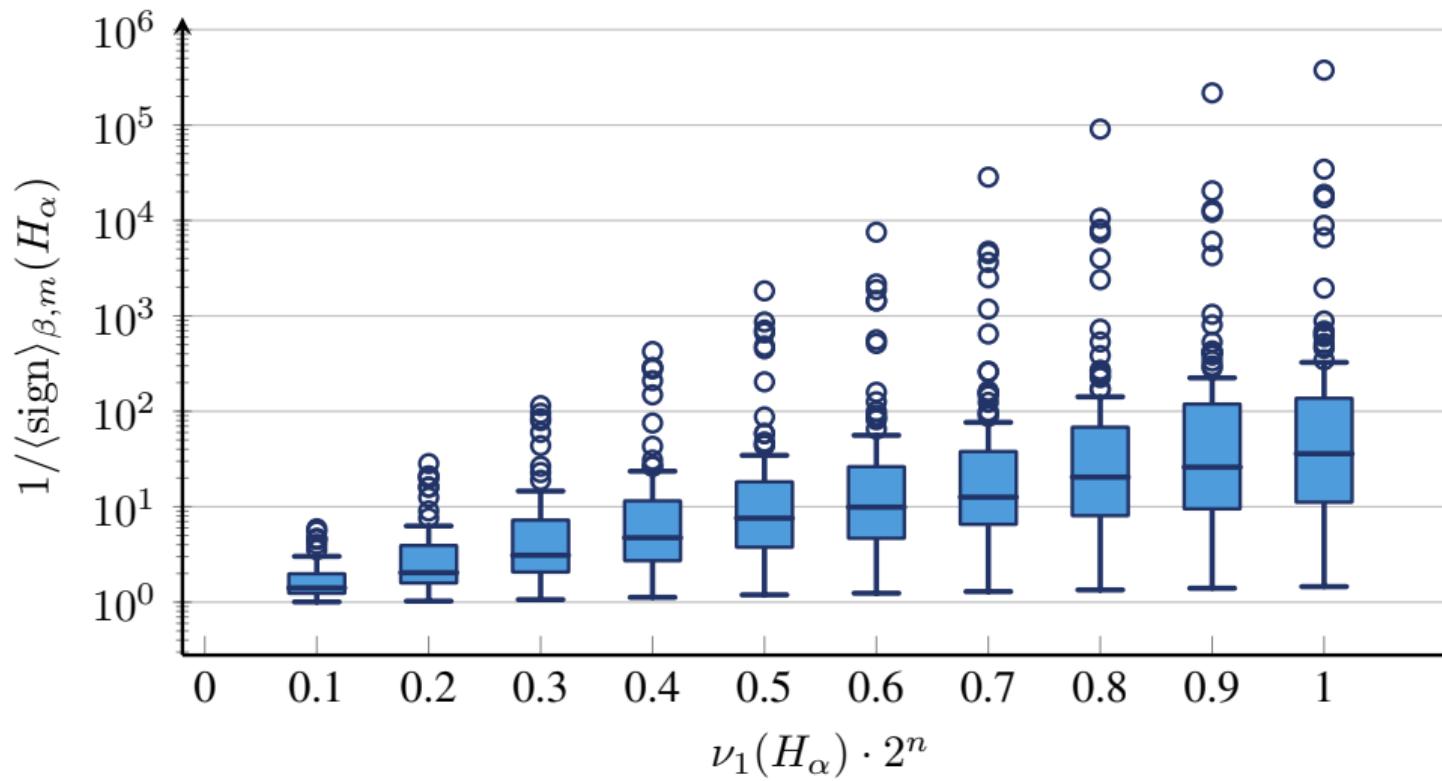
$\ell_1$ -norm

NEGATIVITY

To first order in the *non-stoquastic matrix entries*:

$$\begin{aligned} S(U) &\approx 2m \left\| |UT_mU^\dagger| - UT_mU^\dagger \right\|_{\ell_1} \cdot \|(|UT_mU^\dagger| + UT_mU^\dagger)^{m-1}\|_{\ell_\infty} \\ &\propto d \nu_1(H). \end{aligned}$$

## Average sign vs. non-stoquasticity: Numerical evidence



$$H_\alpha = \frac{H - H_{\text{non-stoq.}} + \alpha H_{\text{non-stoq.}}}{\|H_{\text{non-stoq.}}\|_1}.$$



## Easing in practice: Translation-invariant problems

Translation-invariant non-stoquasticity

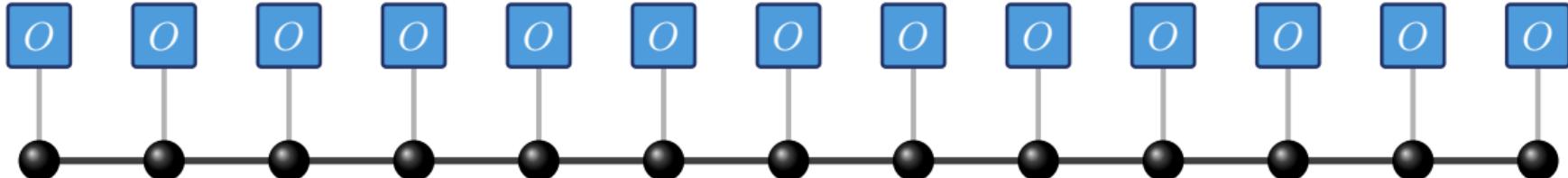
$$H = \sum_{i=1}^n T_i(h)$$
$$\longrightarrow \quad \nu_1(H) \propto \tilde{\nu}_1(h)$$



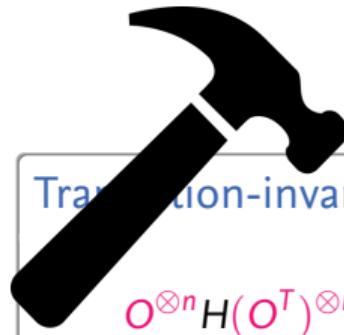
## Easing in practice: Translation-invariant problems

Translation-invariant non-stoquasticity

$$\begin{aligned} O^{\otimes n} H(O^T)^{\otimes n} &= \sum_{i=1}^n T_i((O \otimes O) h(O^T \otimes O^T)) \\ \longrightarrow \quad \nu_1(H) &\propto \tilde{\nu}_1((O \otimes O) h(O^T \otimes O^T)) \end{aligned}$$



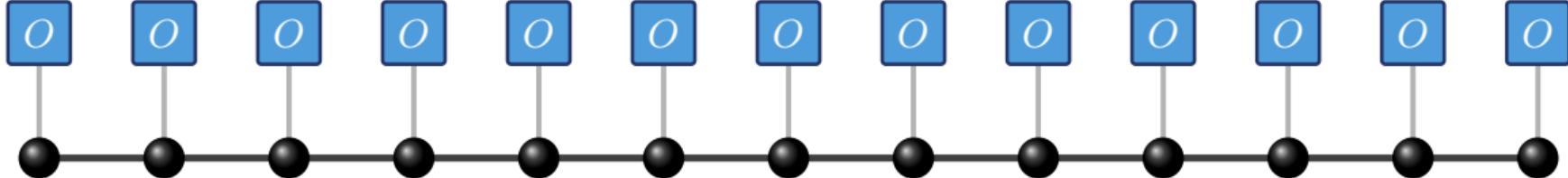
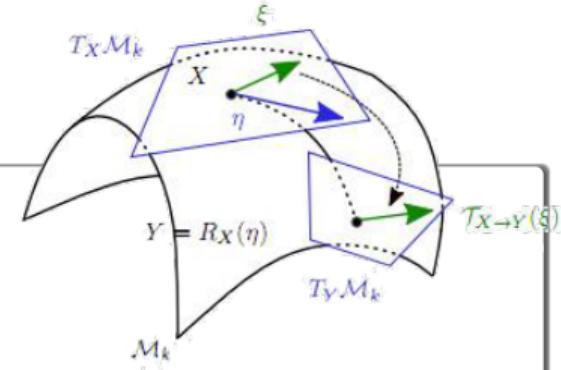
## Easing in practice: Translation-invariant problems



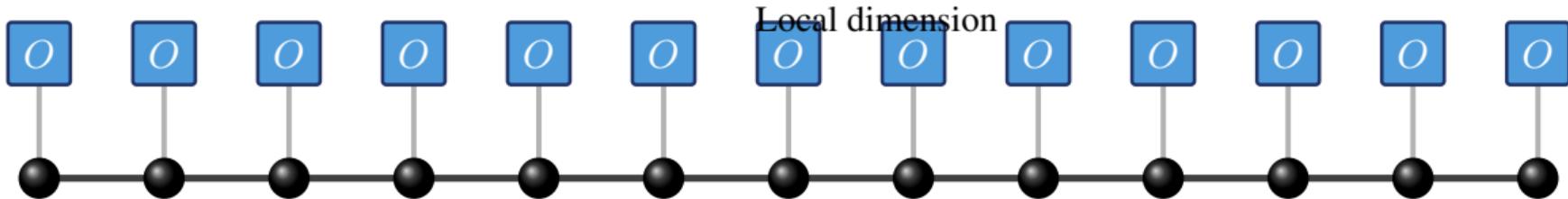
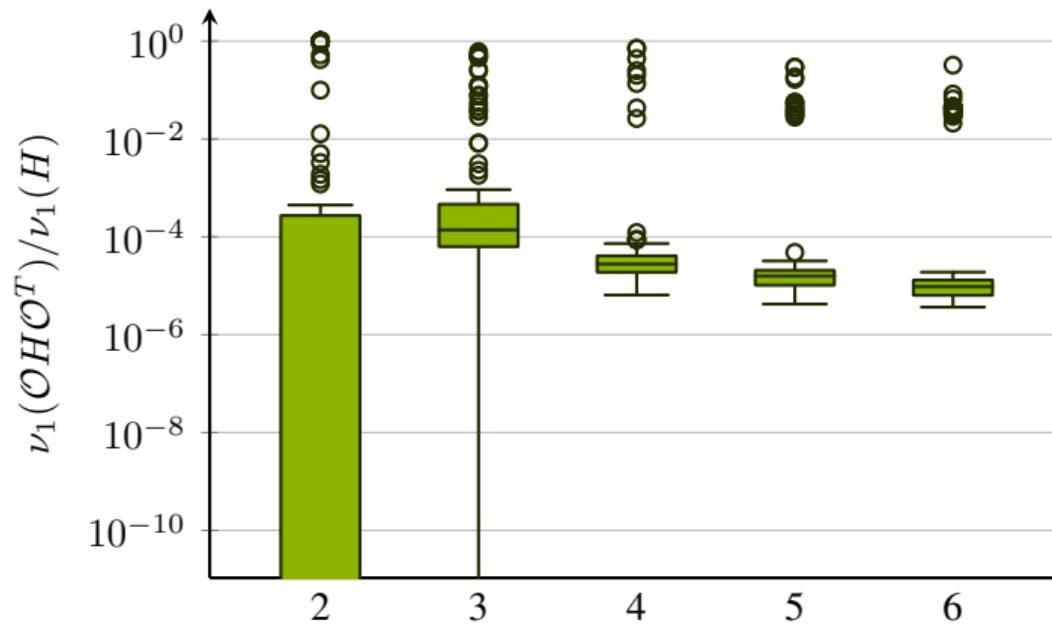
Translation-invariant non-stoquasticity

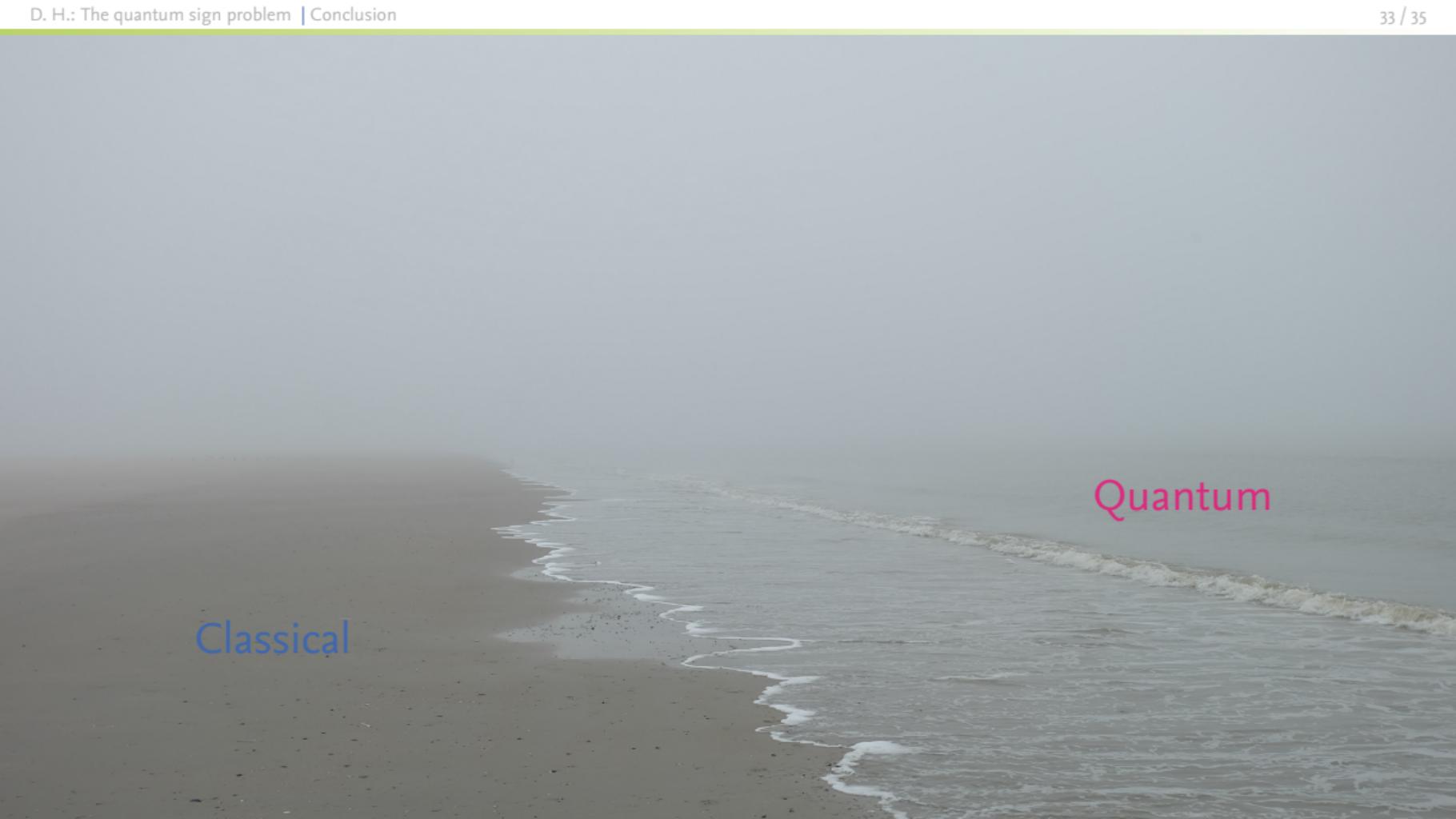
$$O^{\otimes n} H(O^T)^{\otimes n} = \sum_{i=1}^n T_i((O \otimes O) h(O^T \otimes O^T))$$

$$\rightarrow \nu_1(H) \propto \tilde{\nu}_1((O \otimes O) h(O^T \otimes O^T))$$



## Easing in practice: Translation-invariant problems





Classical

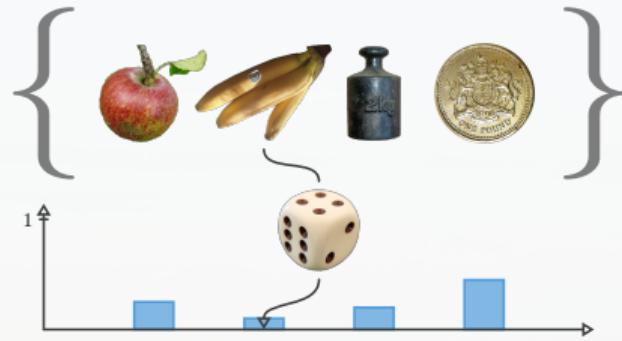
Quantum



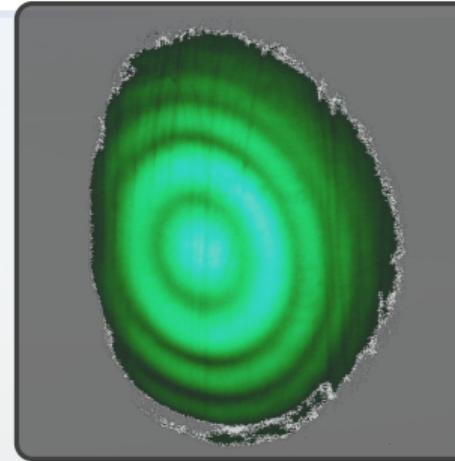
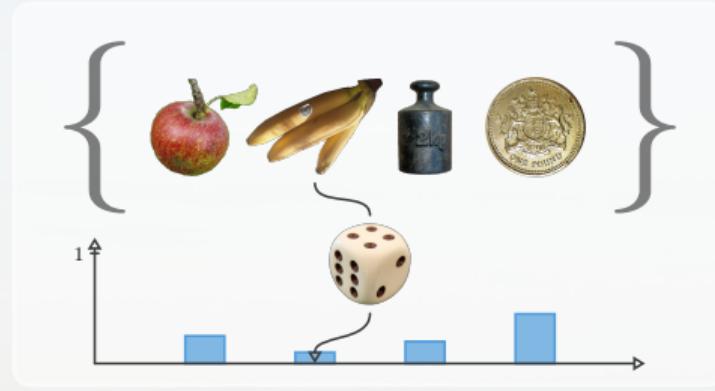
Classical

Quantum

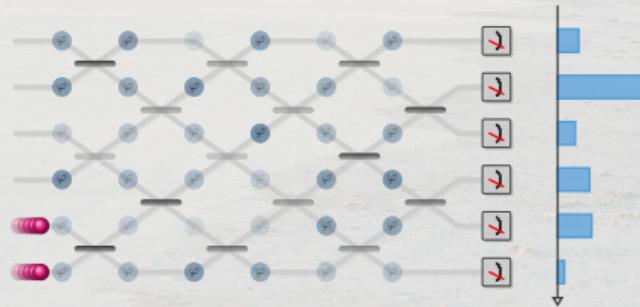
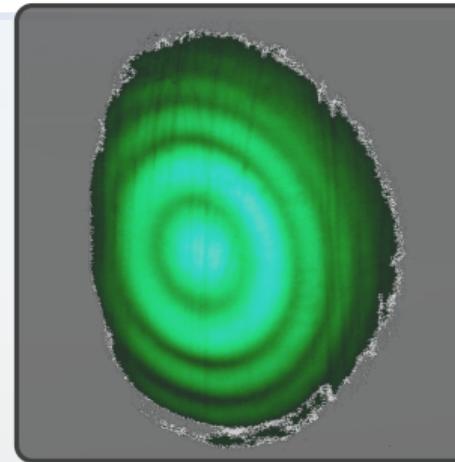
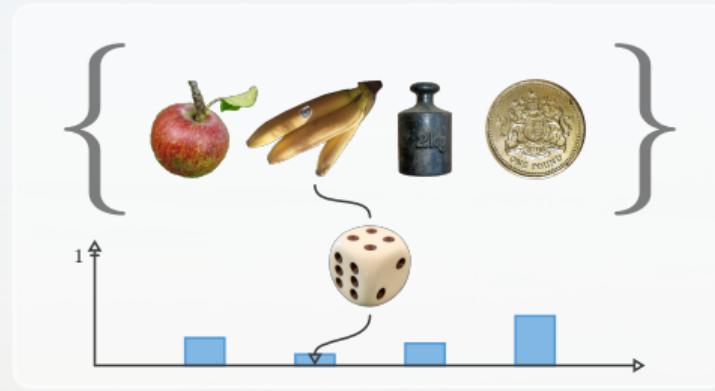
# Outlook



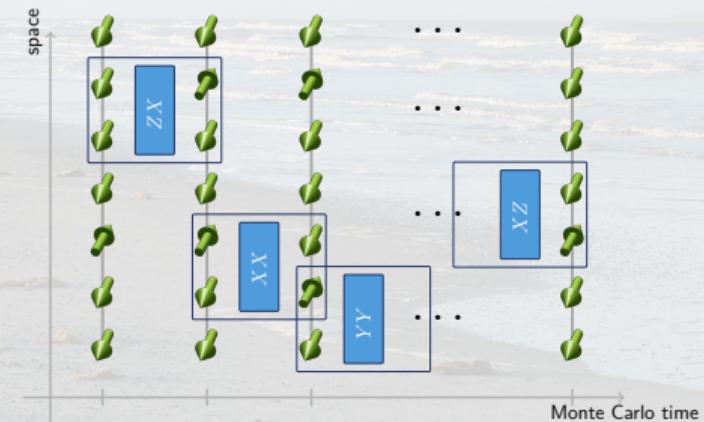
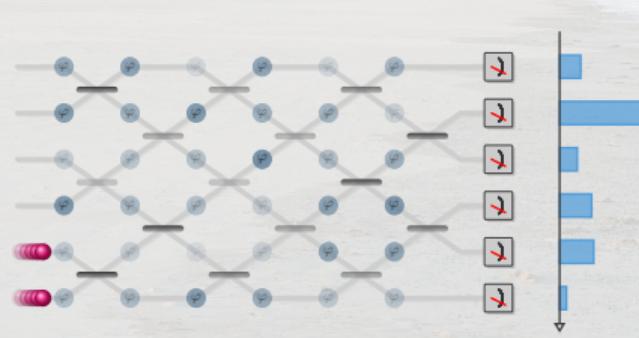
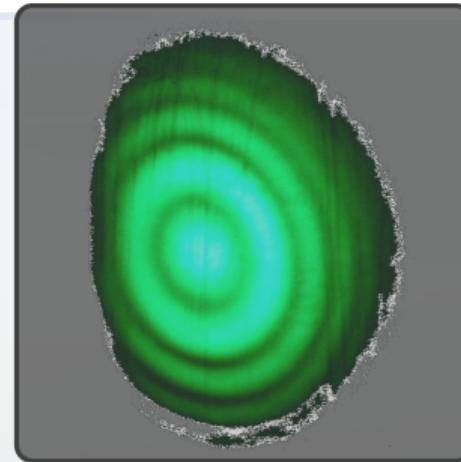
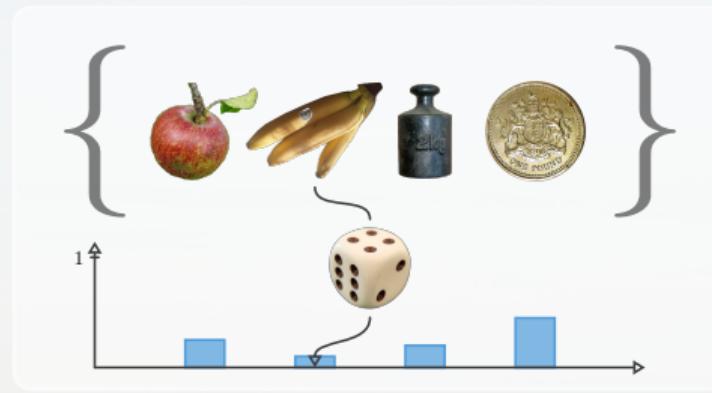
# Outlook



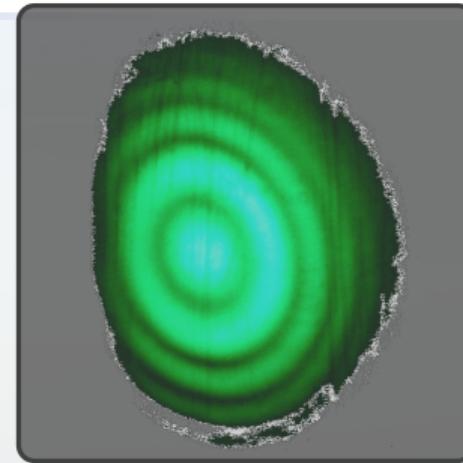
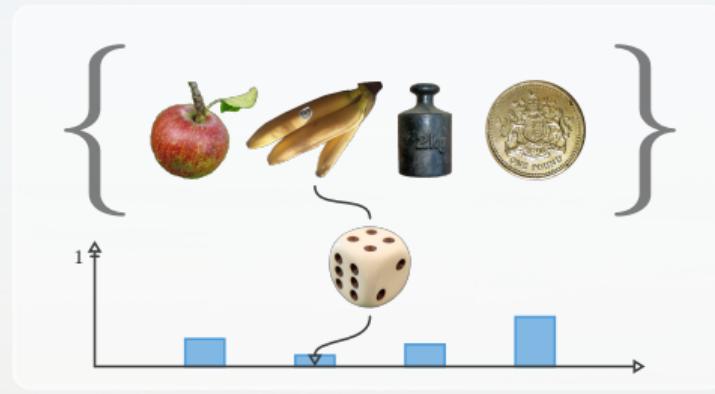
# Outlook



# Outlook



# Outlook



Is the quantum sign problem intrinsic or artificial? [ZOR20]