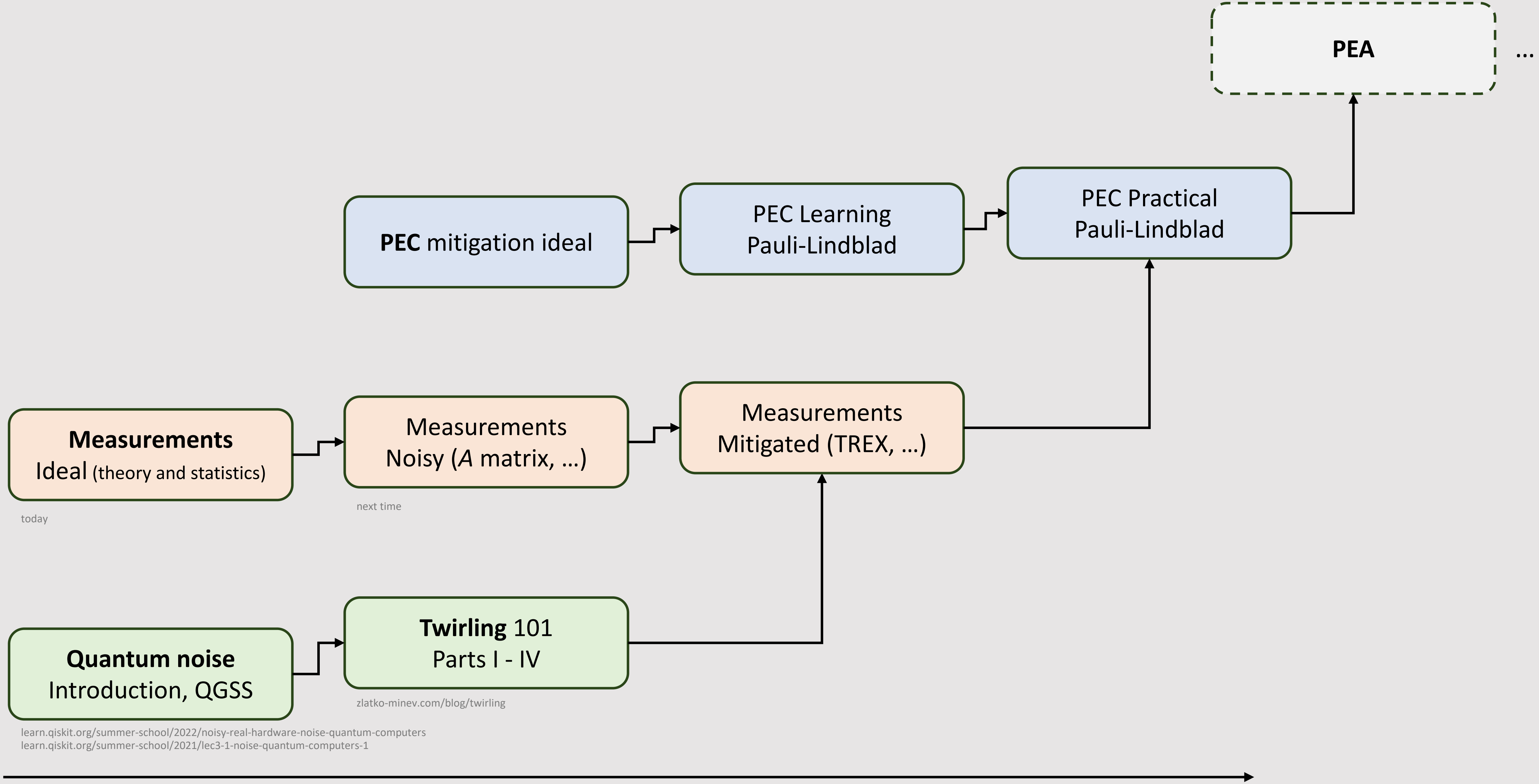
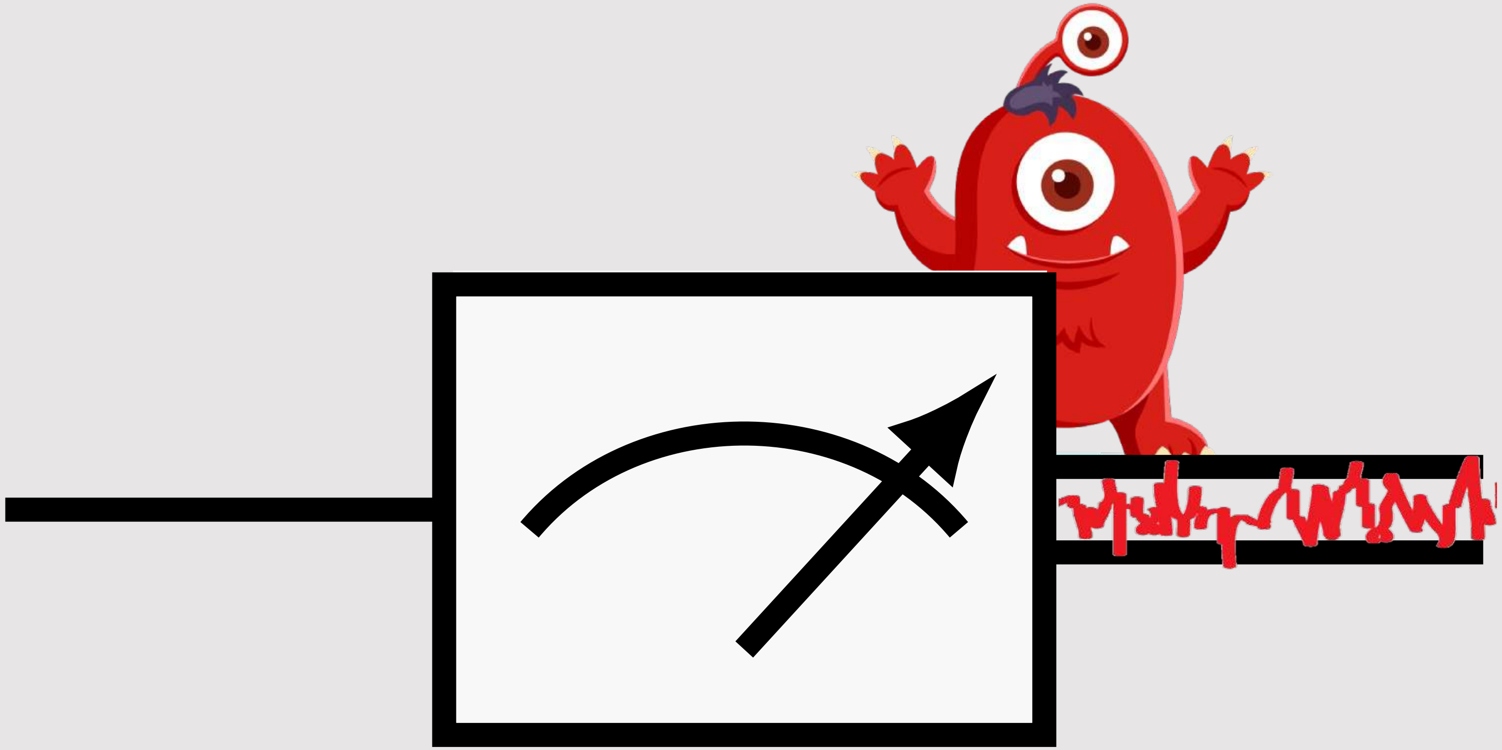
Tutorial map



Quantum Measurement Tutorial

Ideal, Noisy, Mitigated



Zlatko K. Minev

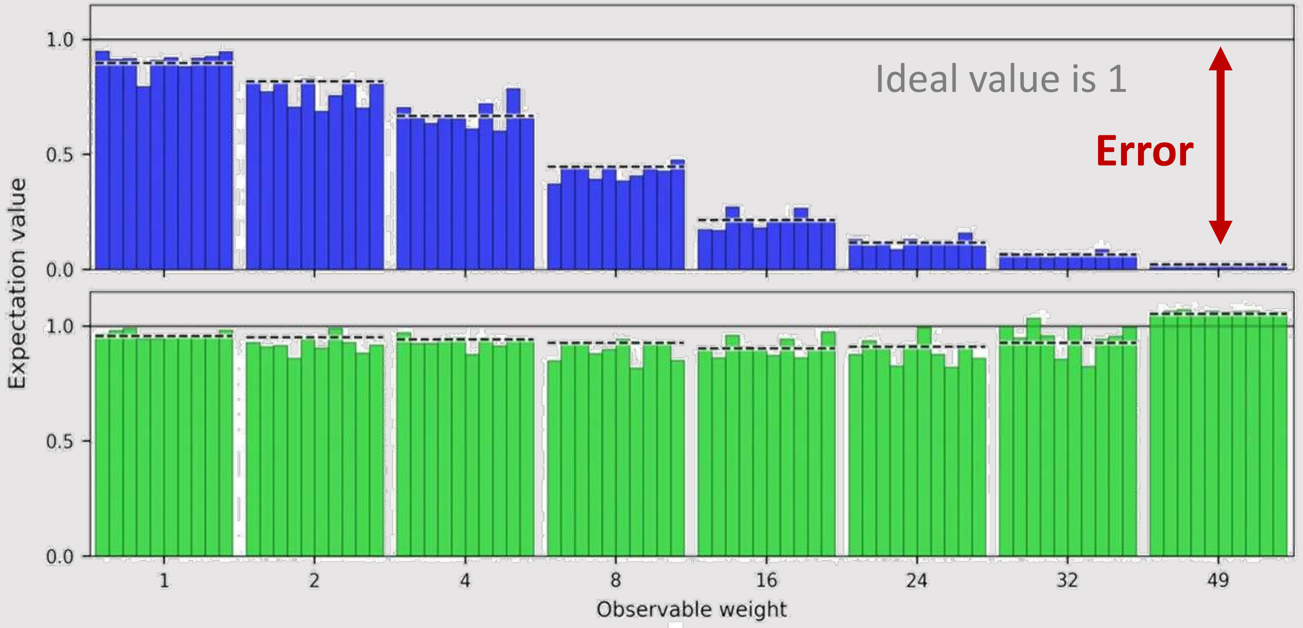
Monster image by jcomp on Freepik

@zlatko_minev zlatko-minev.com/blog

Motivation: Measurement matters

Example data before and after mitigation

(PEC + TREX on a 50-qubit circuit)



Z ZZ ZZZZ ZZZZZZZZ ...

Experiment:

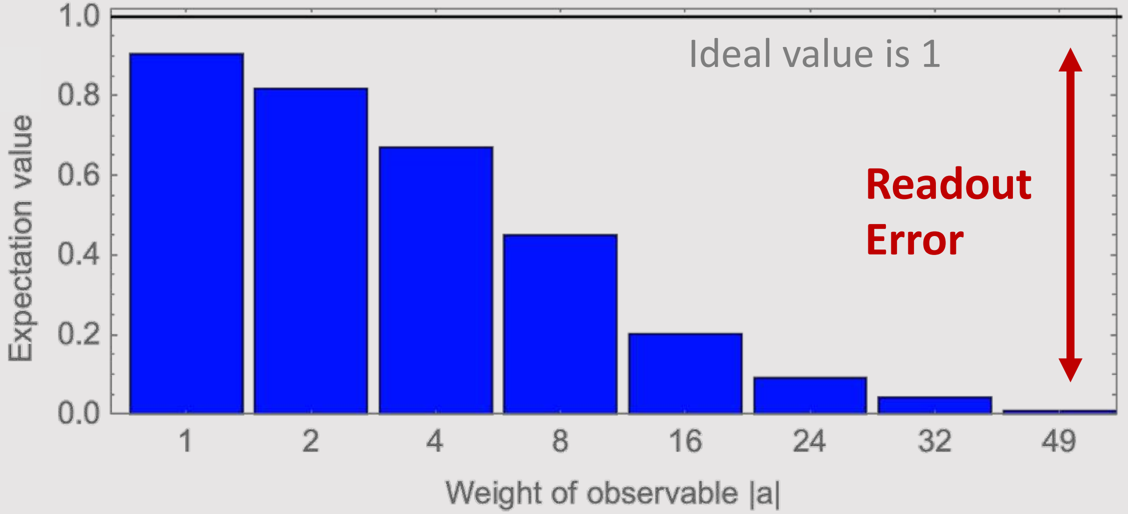
50 qubits

2 depth in cX layers

Plot from research.ibm.com/blog/gammabar-for-quantum-advantage

Readout-error-only model

Let's see what happen even if no gate errors



Z ZZ ZZZZ ZZZZZZZZ ...

A model:

0.05 readout error per-qubit (bit-flip model)

What do we theoretically expect? How do we model it? ...

What are the error bars? Confidence? Which method to use? ...

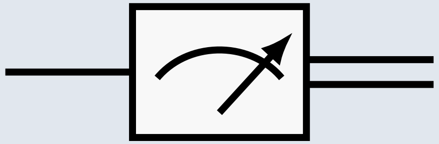
Can we extract more Paulis at once? How do we measure entanglement? ...

Measurement in Quantum Computers

Ideal

Measurement theory 101

- Why care?
- Formulation



Single qubit example

- Single shot
- Many shots
- Statistics, unbiased estimators
- Bounds, Chernoff-Hoeffding inequality
- Error bars on experimental data: likelihood
- Bootstrapping
- Measurement bases

Scaling to *n* qubits

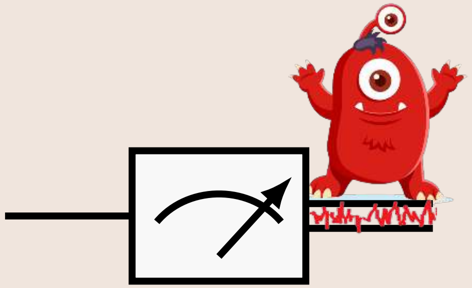
- Single shot preliminaries
- Binary and Pauli strings
- Walsh–Hadamard
- Many shots
- Entangled measurements
- …

Noisy

Qubit example

A matrix, cross-talk

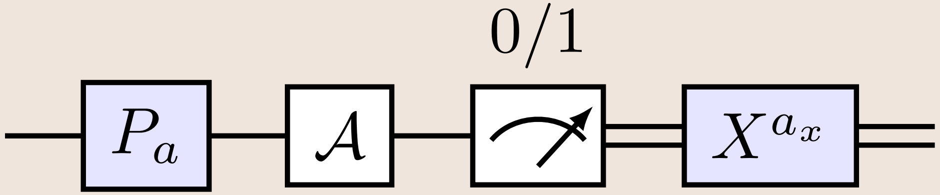
Scaling



Mitigated

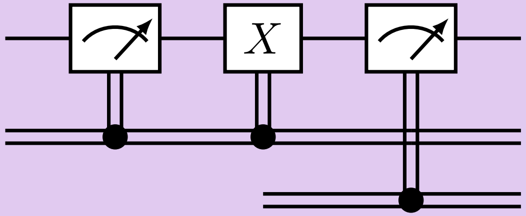
A matrix mitigation

Twirling and T-REX



Possible other topics of interest

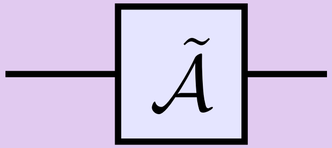
Measurement theory for dynamic circuits



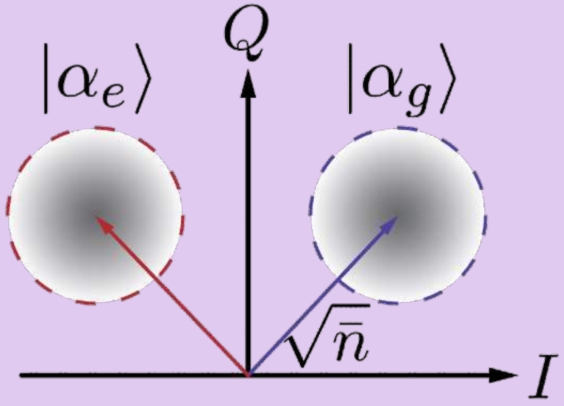
Mitigation for dynamic circuits

Complete and SIC-POVMs

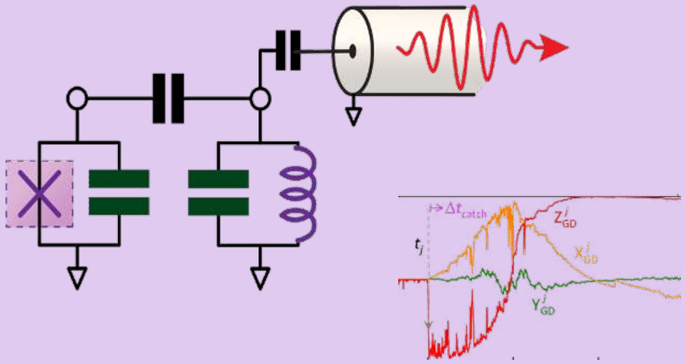
Pauli-Linbdlad formulation for measurement noise



Heterodyne readout I/Q planes



cQED readout and quantum stochastic trajectories



...

Guideposts

Evil monster

Eve the Evil monster is here to terrorize the quantum world

Random Sorcery

Randomness and unpredictability abounds

Key result

Used to highlight a key result

Key challenge

Used to highlight a key challenge

Try it yourself

Pause the video and try to work it out, before I share the answer

Dangerous bend sign

Denotes extra information,
with appendix-like depth or digression.
You can skip this material on a first reading

Caution

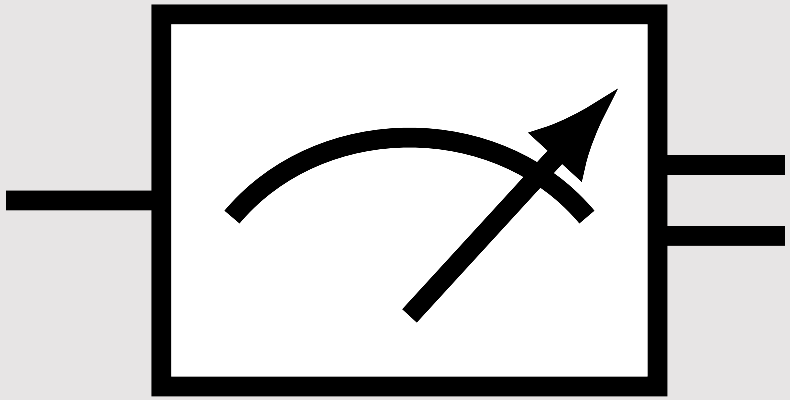
Caution or common pitfall

Monster image by jcomp on Freepik; coin-toss icon: Good Ware; pen: bclogo latex

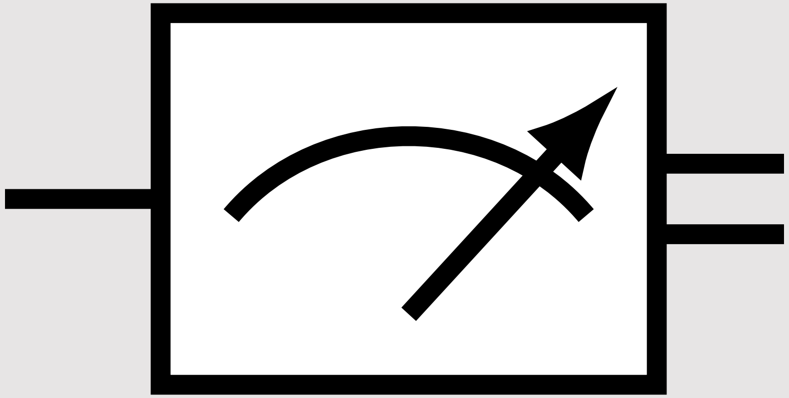
Measurement in Quantum Computers

indimart

meter



Refresher



Why measurement?

Quantum-to-classical interface

- how we get information out of a quantum computer
- fundamentally different from any classical measurement
- where all the randomness and most of the weirdness in quantum happens -> ‘Measurement problem’
- any results we want *must* pass through a measurement
- any noise in measurement will unavoidably degrade our results
- we must get to know them well to master them

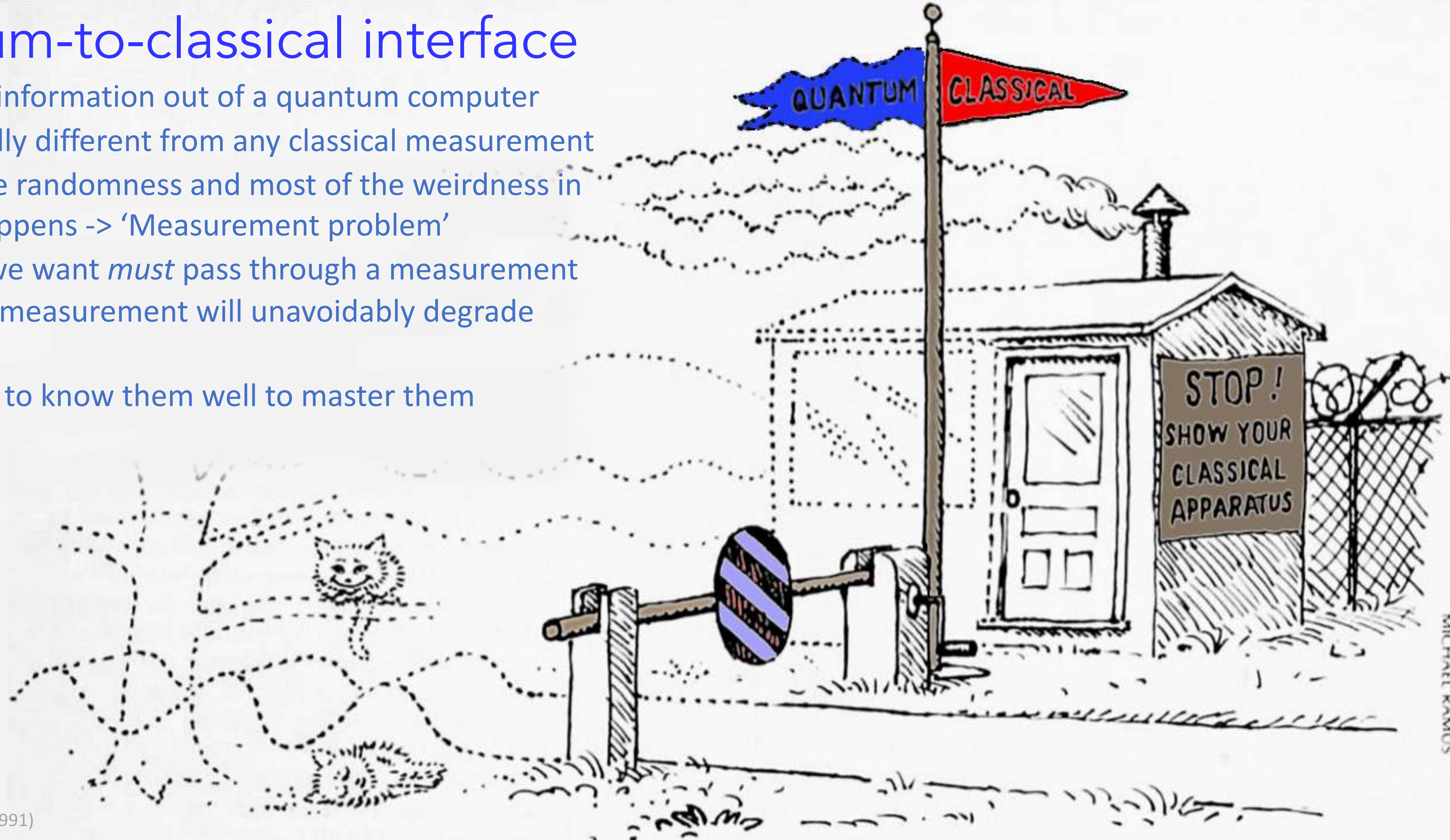
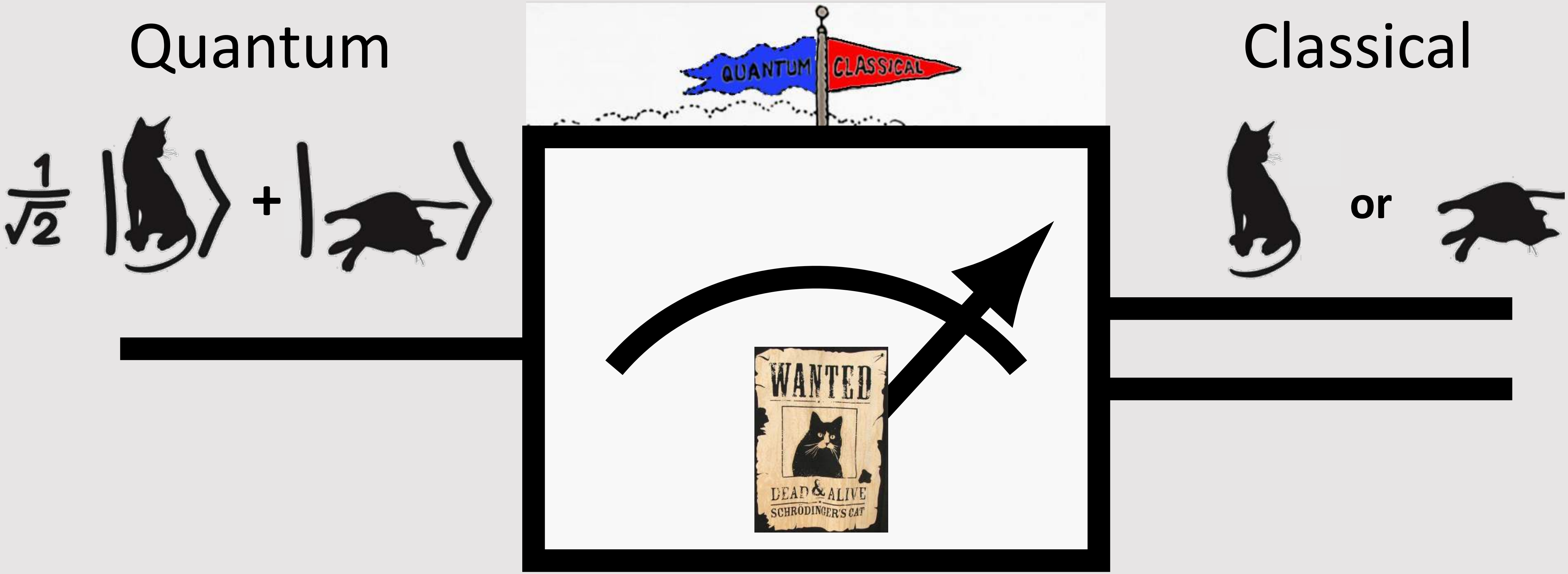


Image: Zurek, Phys. Today (1991)

Quantum-to-classical interface



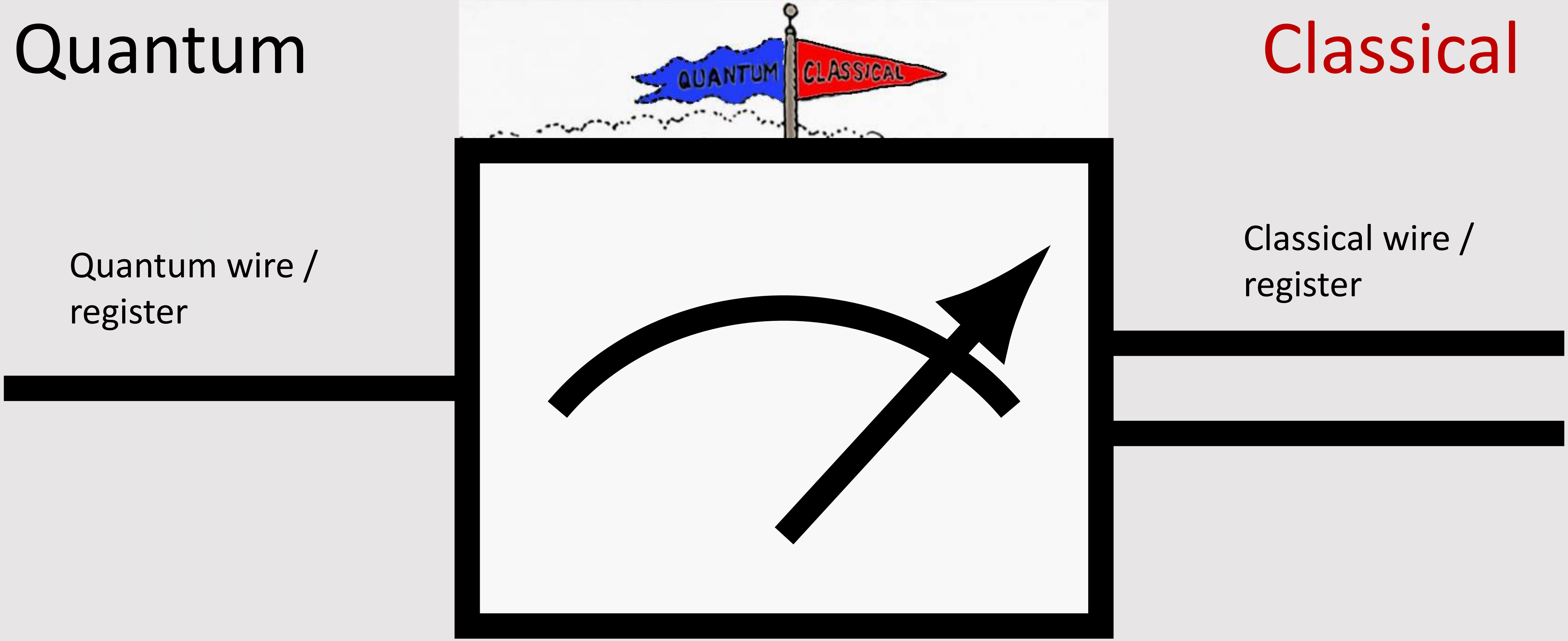
Randomness
Projection / collapse
Connection to classical world

cat image: docencia

The meter

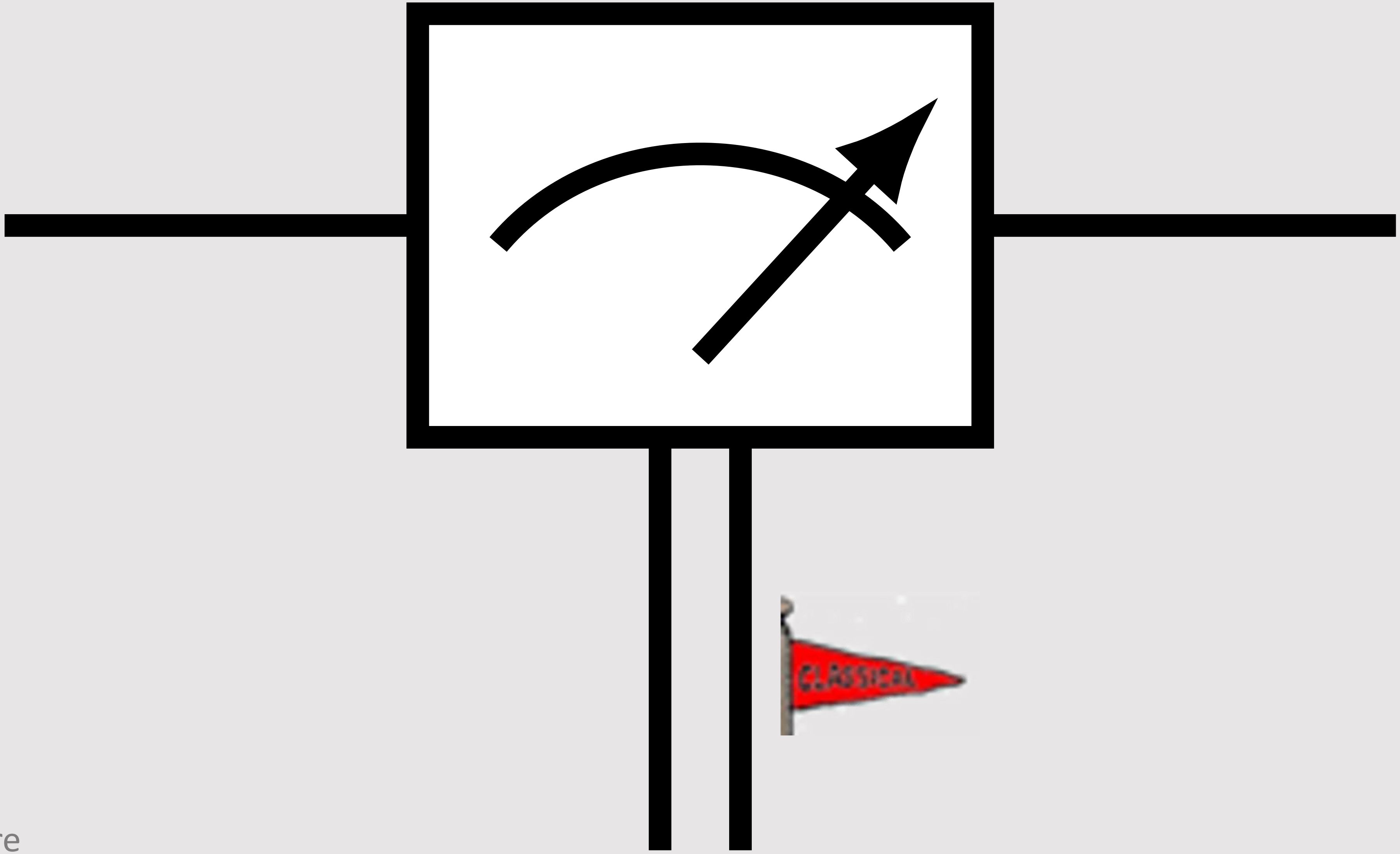
Quantum

Classical



cat image: docencia

The meter with the posterior state



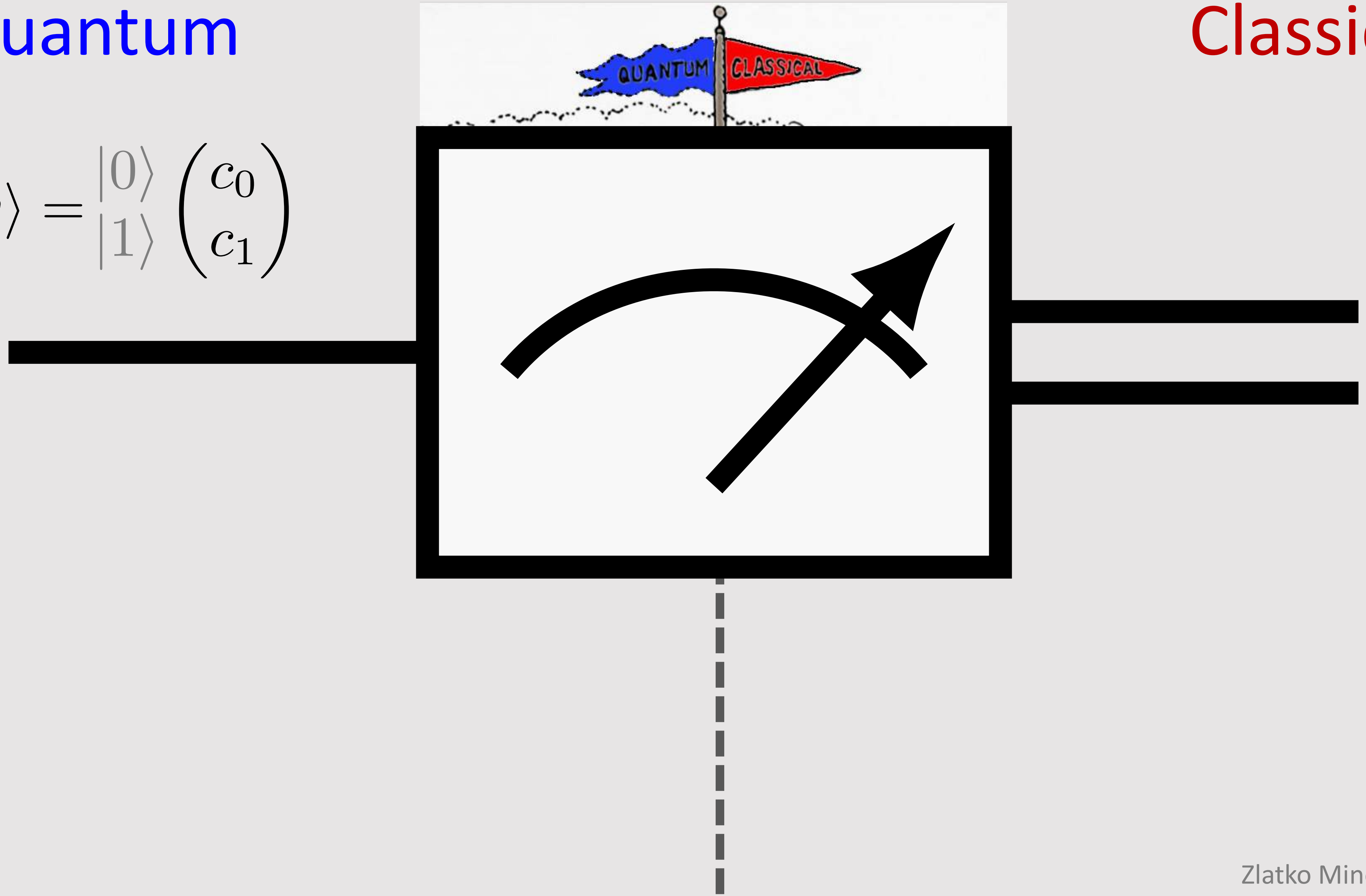
For a later lecture

Qubit example

Quantum

Classical

$$|\psi\rangle = \begin{array}{c} |0\rangle \\ |1\rangle \end{array} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$$



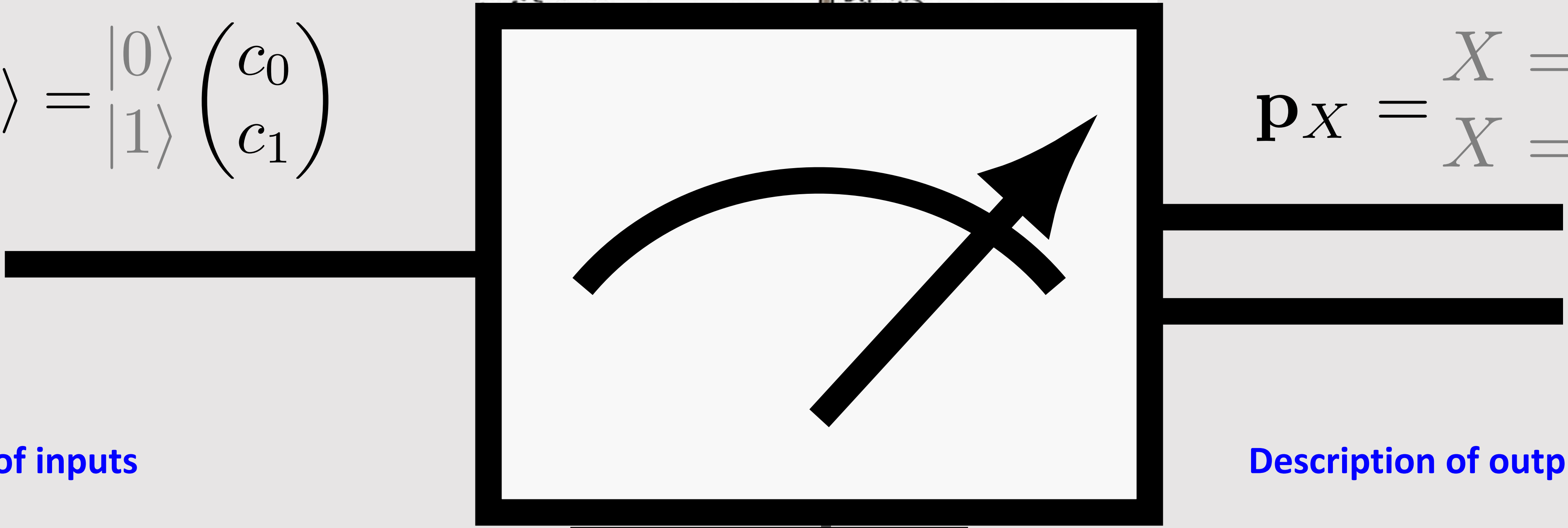
Qubit example

Quantum

Classical

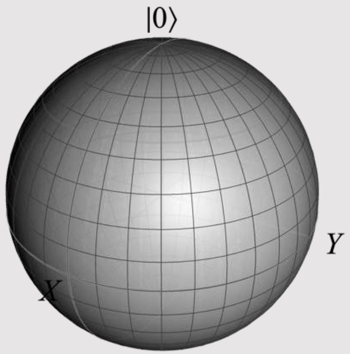
$$|\psi\rangle = \begin{array}{c} |0\rangle \\ |1\rangle \end{array} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$$

$$\mathbf{p}_X = \begin{array}{c} X=0 \\ X=1 \end{array} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$$



Description of inputs

Description of outputs



$$\Sigma = \{0, 1\}$$

$$\mathcal{H} = \mathbb{C}^{\Sigma}$$

$$|\psi\rangle \in \mathcal{H}$$

Measure in computational basis
$|0\rangle, |1\rangle$

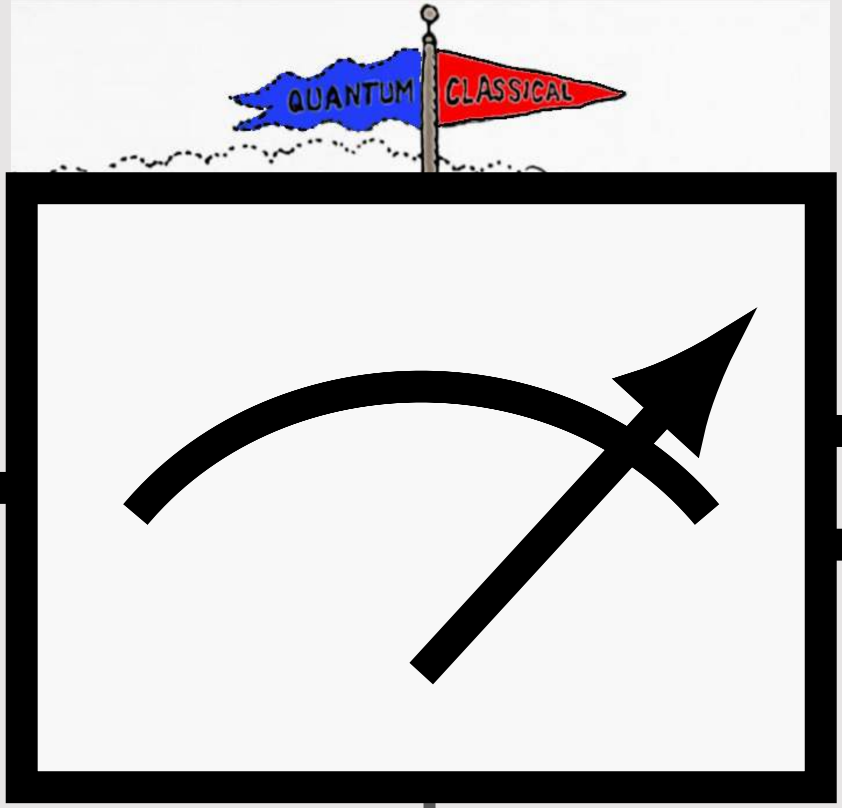
outcome   outcome probability

$$\begin{cases} X=0 & : \quad p\,(X=0) = |\langle 0|\psi\rangle|^2 \\ X=1 & : \quad p\,(X=1) = |\langle 1|\psi\rangle|^2 \end{cases}$$

Qubit example

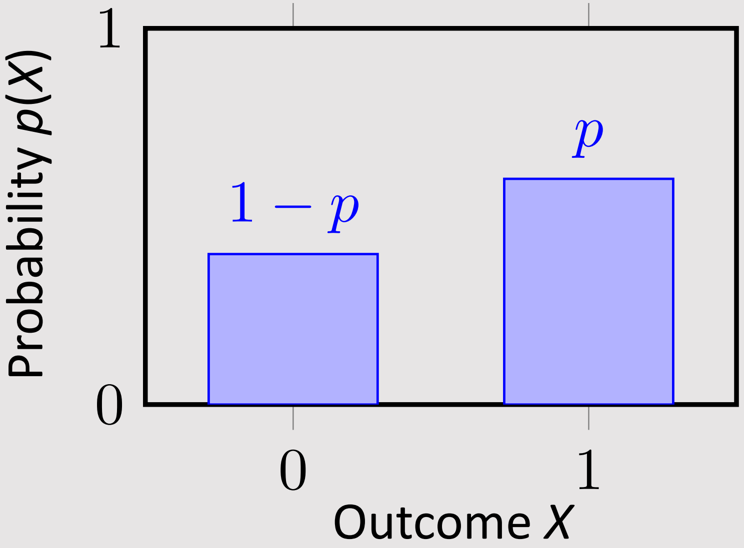
Quantum

$$|\psi\rangle = \begin{array}{c} |0\rangle \\ |1\rangle \end{array} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$$

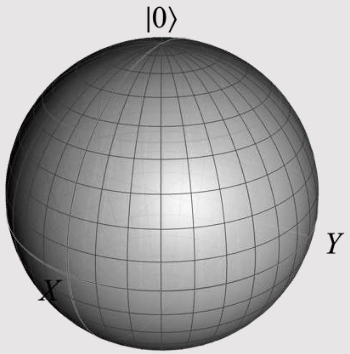


Classical

$$\mathbf{p}_X = \begin{array}{c} X = 0 \\ X = 1 \end{array} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$$



Description of inputs



$$\Sigma = \{0, 1\}$$

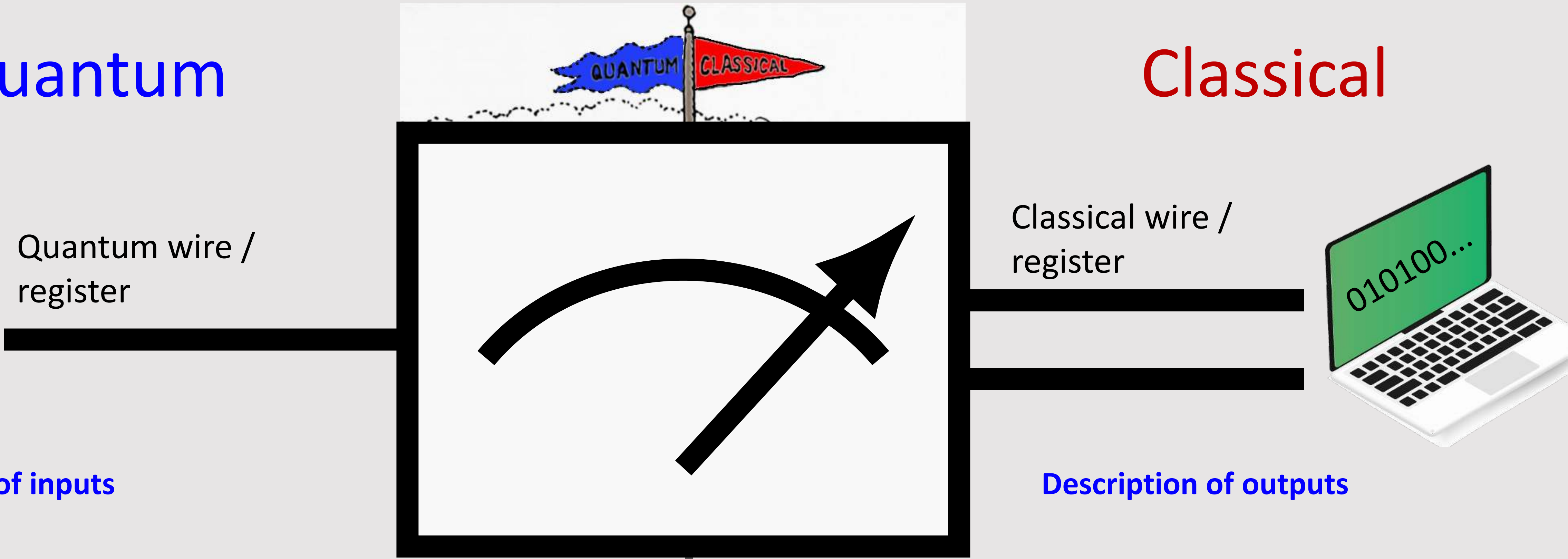
$$\mathcal{H} = \mathbb{C}^{\Sigma}$$

$$|\psi\rangle \in \mathcal{H}$$

The input and outputs of the meter interface: Formal

Quantum

Classical



Description of inputs

Description of outputs

$$|\psi\rangle \in \mathcal{H}$$

A pure state in the Hilbert space.

$$\rho \in D\left(\mathcal{H}\right) \subset L\left(\mathcal{H}\right)$$

A state in the space of density operators, a subspace of the operator space on the Hilbert pure-state space. $D\left(\mathcal{H}\right) := \{\rho \in \mathrm{Pos}\left(\mathcal{H}\right) : \mathrm{Tr}\left(\rho\right) = 1\}$

$$X \in \Sigma$$

Classical random variable X over an alphabet Σ (assume computational basis)

$$\mathbf{p}_X \in \mathcal{P}\left(\Sigma\right)$$

Probability vector for X, in the space of probability vectors over the alphabet Σ

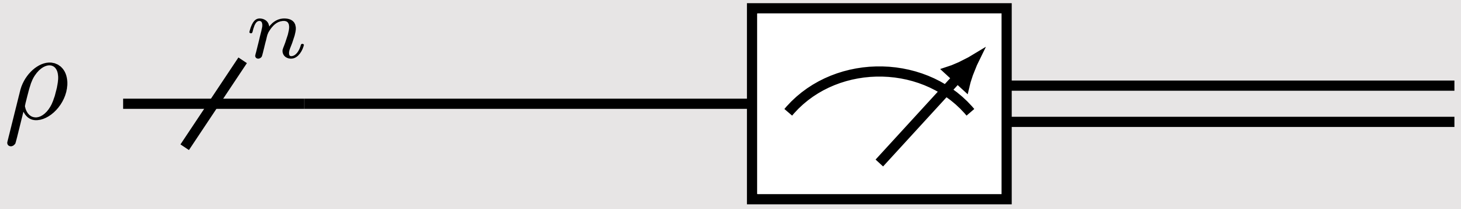
Quantum Meter: Elements and notation

measurement operator associated with outcome x

possible classical outcome

set of possible classical measurement outcomes

$$\{\hat{\mu}(x) : x \in \Sigma\}$$



$$X, \mathbf{p}_X$$

input state

on n qubits

meter

classical random variable representing the measurement outcome in an instance

probability distribution as (as a vector) over the possible outcomes

Coin-toss icon created by Good Ware - Flaticon

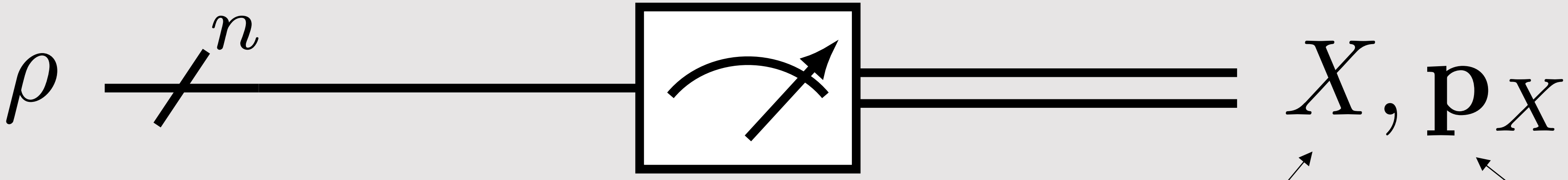
Meter: Measurement operators

measurement operator associated with outcome x

possible classical outcome

set of possible classical measurement outcomes

$$\{\hat{\mu}(x) : x \in \Sigma\}$$



input state

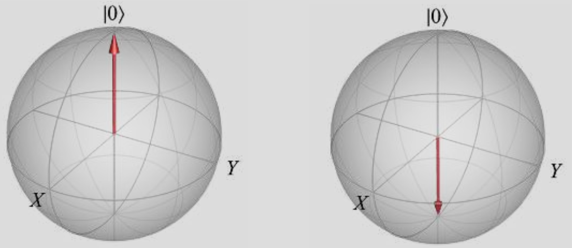
on n qubits

meter

classical random variable representing the measurement outcome in an instance

probability distribution as (as a vector) over the possible outcomes

Qubit example



$$\hat{\mu}\,(0) = |0\rangle\,\langle 0|$$

$$\hat{\mu}\,(1) = |1\rangle\,\langle 1|$$

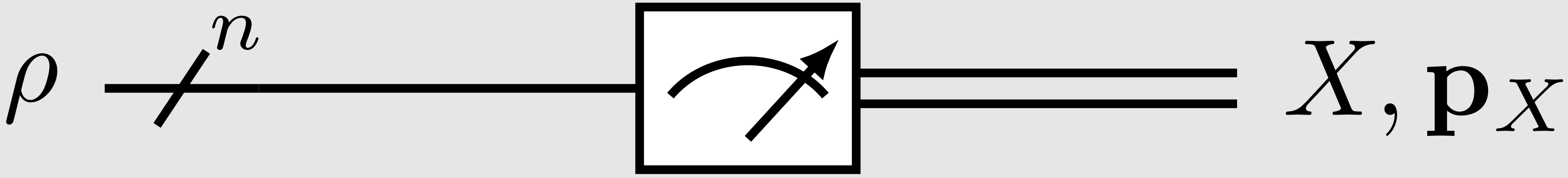
Meter: Measurement operators

measurement operator associated with outcome x

possible classical outcome

set of possible classical measurement outcomes

$$\{\hat{\mu}(x) : x \in \Sigma\}$$



input state

on n qubits

meter

General description of the meter

Measurement function and operators:

$$\mu : \Sigma \to \mathrm{Pos}(\mathcal{H})$$

Positive semidefinite operators

$$\{\hat{\mu}(x) : x \in \Sigma\} \subset \mathrm{Pos}(\mathcal{H})$$

Subject to identity resolution:

$$\sum_{x \in \Sigma} \hat{\mu}(x) = \hat{I}_{\mathcal{H}}$$

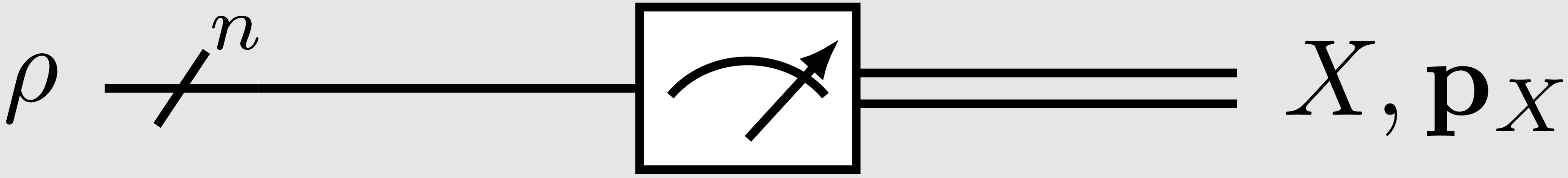
Meter: Quantum-to-classical rule

measurement operator associated with outcome x

possible classical outcome

set of possible classical measurement outcomes

$$\{\hat{\mu}(x) : x \in \Sigma\}$$



input state

on n qubits

meter

Description of the meter

The **quantum-to-classical rule** (postulate of quantum)

$$p\left(x\right) := \Pr\left(X = x\right) = \langle \hat{\mu}\left(x\right), \rho \rangle = \mathrm{Tr}\left(\hat{\mu}^{\dagger}\left(x\right)\rho\right)$$
$$= \langle\!\langle \hat{\mu}\left(x\right) | \rho \rangle\!\rangle$$

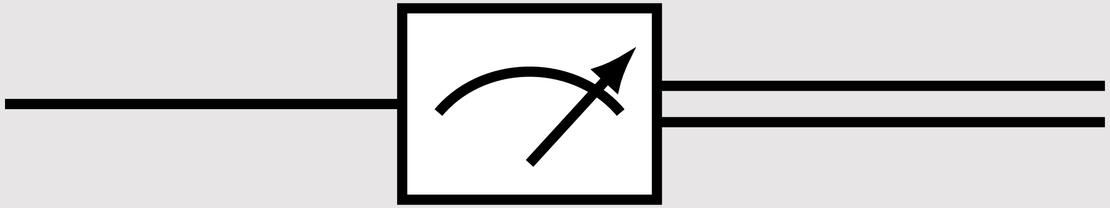
John von Neumann
Image source: LANL

Max Born
Image source: Public domain

Meter: Quantum-to-classical rule

measurement operator
associated with outcome x

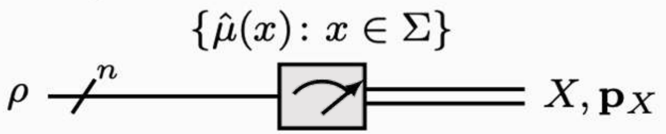
$$\{\hat{\mu}(x) : x \in \Sigma\}$$



Positive operator-valued measure (POVM): Acronyms

- The measurement operator is also known under a few equivalent names:
  positive operator-valued measure (**POVM)**, and
  probability-operator-valued measure (**POM**)
- Provide a mathematical formalism for general measurements, including loss of information.
  - Generalizes the traditional projective or Von Neumann measurements, sometimes called projection-valued measures (**PVM**)
- POVMs do not account for the effect on the quantum state after the measurement.
  For that, we need to introduce the *effect-valued measures* we would touch on in dynamic circuits, which can be used to construct POVMs.

Defined as measurement operators



A *measurement* is a function of the form, or a collection of positive semidefinite operators of the form, [3, Defn. 2.34]

$$\mu : \Sigma \to \mathrm{Pos}(\mathcal{H}) \,, \quad \{\hat{\mu}\,(x) : x \in \Sigma\} \subset \mathrm{Pos}(\mathcal{H}) \,, \quad \text{subject to} \quad \sum_{x\in\Sigma} \hat{\mu}\,(x) = \hat{I}_{\mathcal{H}} \,,$$

for some alphabet Σ and Hilbert space $\mathcal{H}$, satisfying the resolution constraint of the identity $\hat{I}_{\mathcal{H}}$ on the Hilbert space for non-sub-ensemble measurement (i.e., non-normalized trajectories).

Σ	set of possible *measurement outcomes* of the measurement apparatus.
x	a measurement outcome value, $x \in \Sigma$.
$\hat{\mu}\,(x)$	*measurement operator* (*positive semidefinite*) associated with outcome x (POVM).
X	random sample variable, $X \in \Sigma$, sampled according to the probability distribution $p\,(x)$
$p\,(x)$	probability distribution over outcomes $p\,(x) := \Pr\,(X = x) = \langle \hat{\mu}\,(x)\,, \rho\rangle = \langle\!\langle \hat{\mu}\,(x)\,\vert\rho\rangle\!\rangle = \mathrm{Tr}\left(\hat{\mu}^{\dagger}\,(x)\,\rho\right)$.
$\mathbf{p}_X$	probability distribution as a column vector with elements $\{p\,(x) = \langle\!\langle \hat{\mu}\,(x)\,\vert\rho\rangle\!\rangle : x \in \Sigma\}$ over the outcome space. Lives in in the set of probability vectors $\mathbf{p}_X \in \mathcal{P}\,(\Sigma) \subset \mathcal{X}$. Sometimes the subscript is omitted.
$\mathcal{X}$	classical outcome register space $\mathcal{X} := \mathbb{C}^{\Sigma}$; note, $\mathbf{p}_X \in \mathcal{X}$.

Measurement notation cheat sheet

Properties

- ➡ Probability vector $\mathbf{p}$ depends *linearly* on ρ.
- ➡ Positive semidefinite: both $\hat{\mu}\,(x)$ and ρ.
  - ➡ $\hat{\mu}^{\dagger} = \hat{\mu}$, every positive semidefinite operator is Hermitian.
  - ➡ Bounding norms of positive Hermitian operators, see (C62).

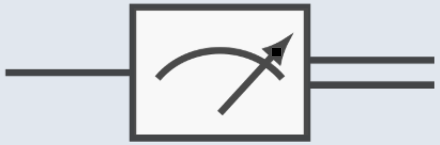
…

Measurement in Quantum Computers

Ideal

Measurement theory 101

- Why care?
- Formulation



Single qubit example

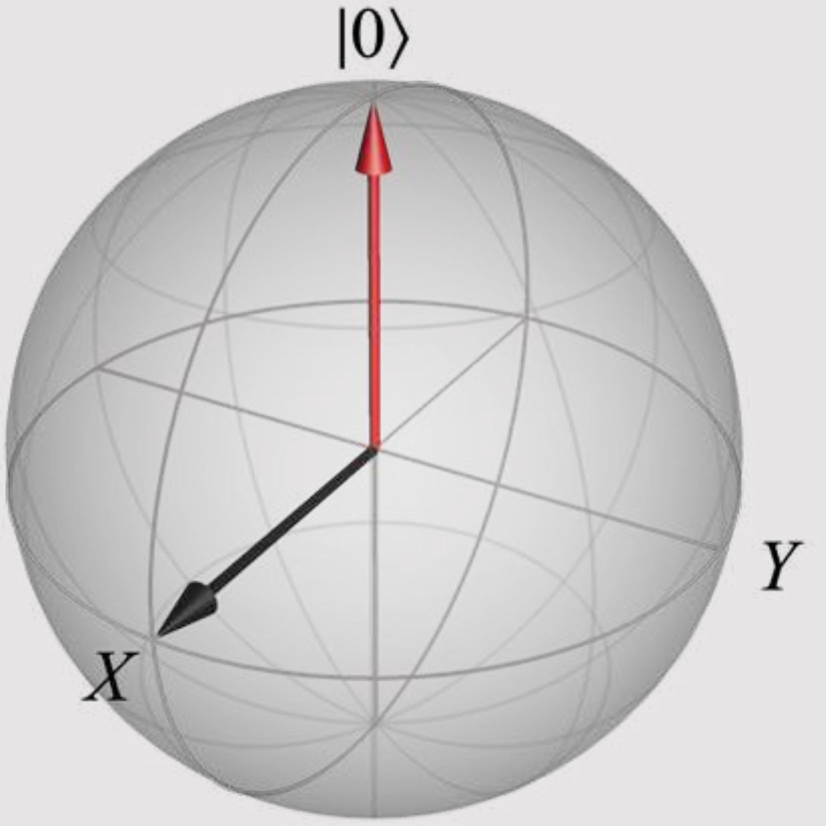
- Single shot
- Many shots
- Statistics, unbiased estimators
- Bounds, Chernoff-Hoeffding inequality
- Bootstrapping
- Measurement bases

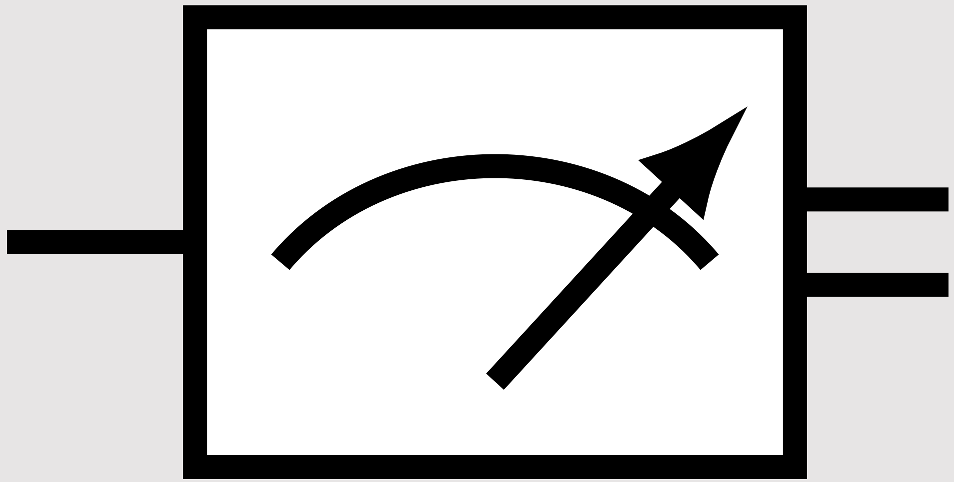
Scaling to *n* qubits

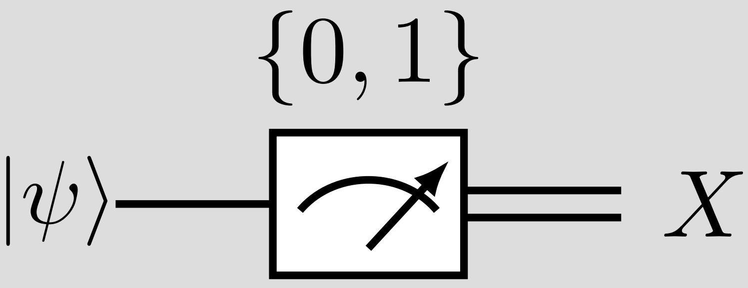
- Single shot preliminaries
- Binary and Pauli strings
- Walsh–Hadamard
- Many shots
- Entangled measurements
- …

* Coin-toss icon created by Good Ware - Flaticon

Single qubit example







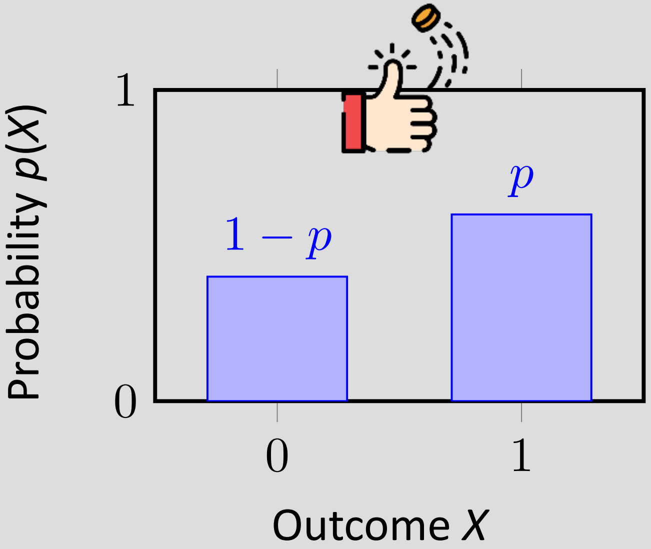
Set of possible outcomes

$$\Sigma = \{0, 1\}$$

$$X \in \Sigma$$

Measurement operators

$$\begin{array}{ll} x & \hat{\mu}(x) \\ 0: & \hat{\mu}(0) = |0\rangle\langle 0| \\ 1: & \hat{\mu}(0) = |1\rangle\langle 1| \end{array}$$



Probability to measure outcome

$\rho = |\psi\rangle\langle\psi|$ to be introduced later in detail

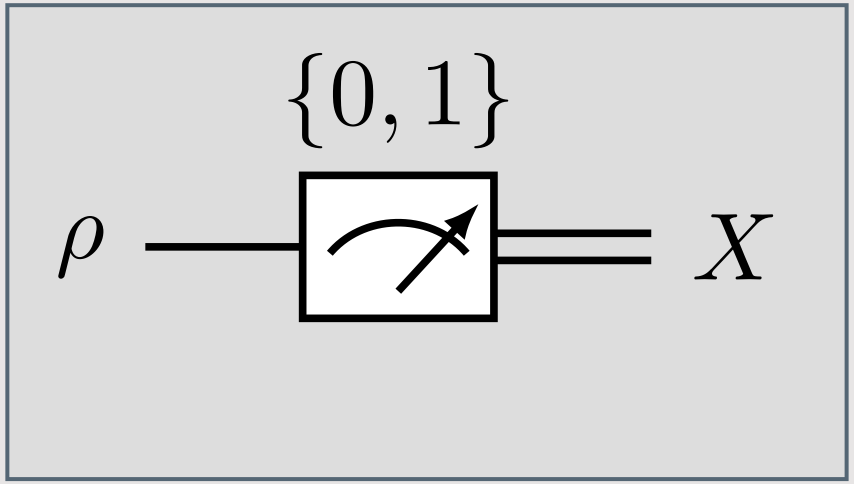
$$\begin{cases} X = 0: & p(X=0) = \mathrm{Tr}\left(\hat{\mu}(0)^\dagger \rho\right) = \mathrm{Tr}\left(|0\rangle\langle 0|\rho\right) = \frac{1}{2}\left(1 + \langle Z\rangle\right) \\ X = 1: & p(X=1) = \mathrm{Tr}\left(\hat{\mu}(1)^\dagger \rho\right) = \mathrm{Tr}\left(|1\rangle\langle 1|\rho\right) = \frac{1}{2}\left(1 - \langle Z\rangle\right) \end{cases}$$

Bernoulli distribution. Single shot outcome follows a Bernoulli distribution:

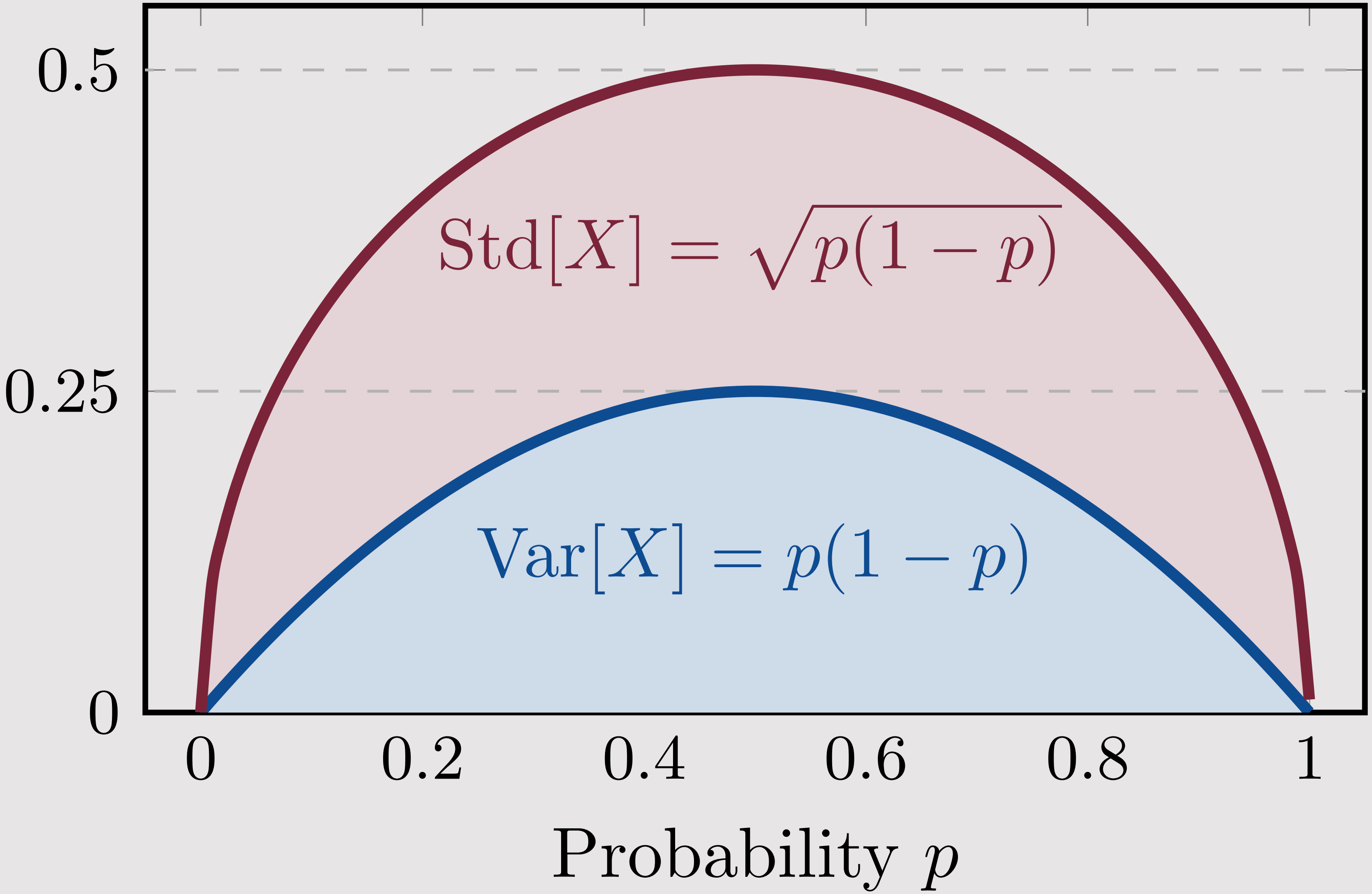
$$X \sim \mathrm{Bernoulli}(p) \qquad p := \mathrm{Tr}\left(|1\rangle\langle 1|\rho\right) \in [0, 1]$$

$$\mathrm{E}[X] = p \qquad \mathrm{Var}[X] = p(1-p)$$

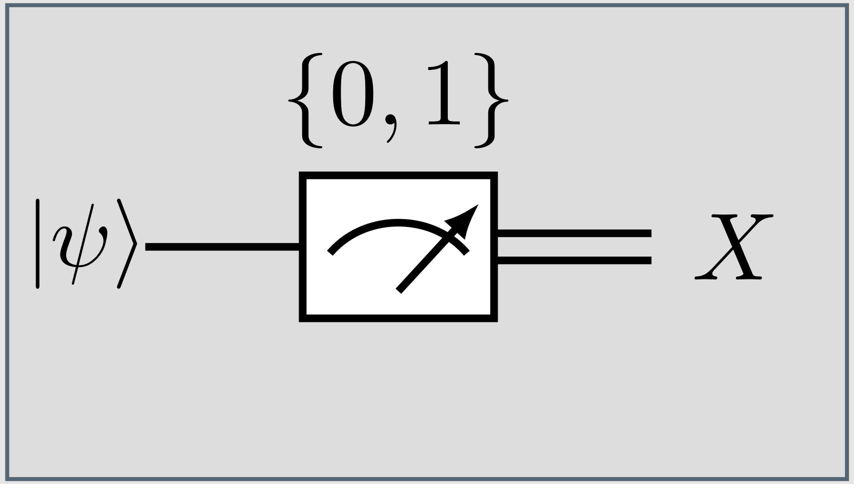
Quantum projection noise for a single shot

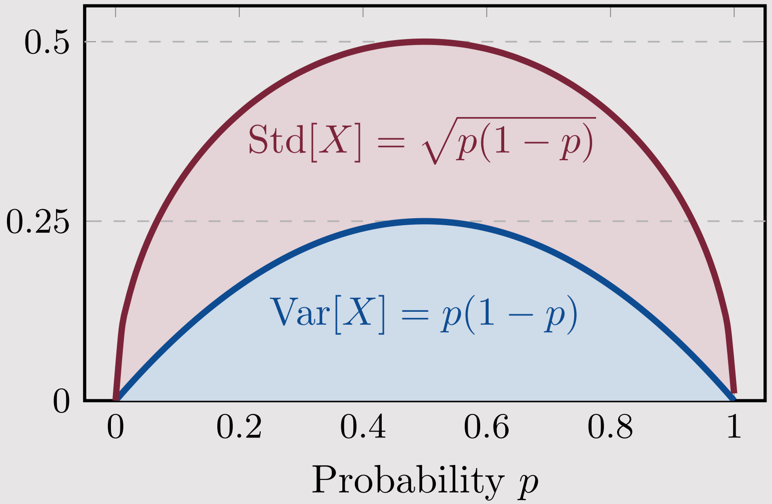


Variance of the random classical variable vs. probability to obtain 1



Uncertainty in observable (expected empirical variance)





The *uncertainty of the observable* $\hat{O}$ is a measure of the spread of results around the mean $\langle \hat{O} \rangle$. Empirically, defined as the difference between each measured result and the mean is calculated, squared, and then averaged.

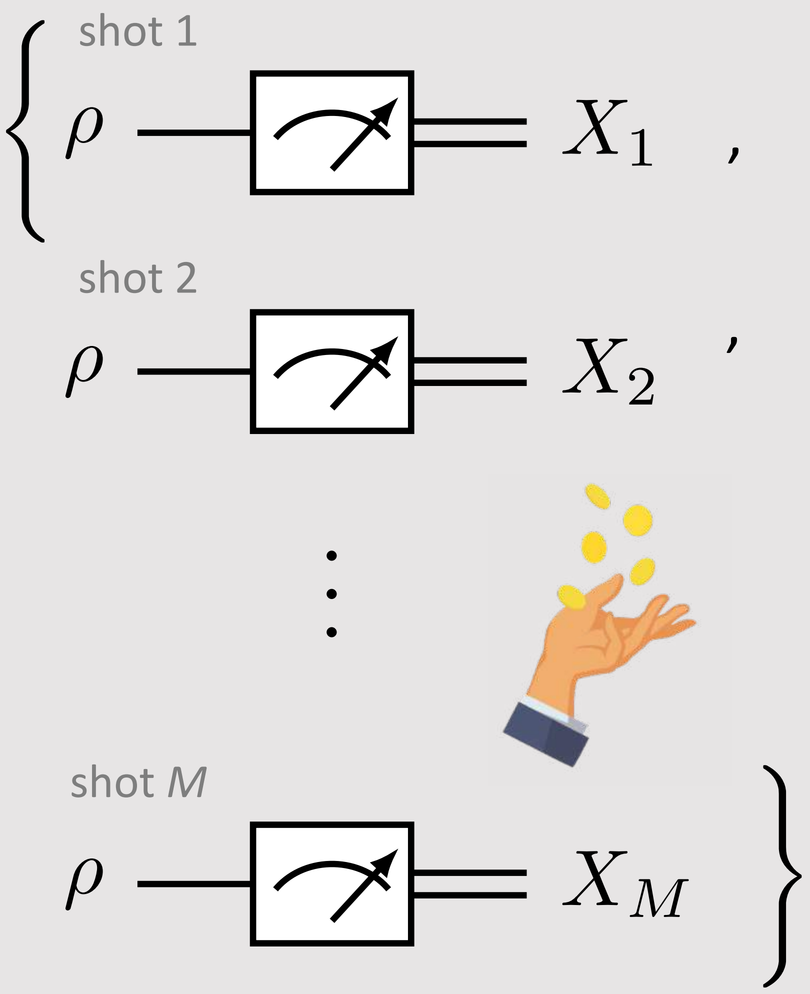
$$\mathrm{Var}\left[X\right] = \mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}\left[X\right]^2 \ ,$$

$$\therefore \sigma_O := \left\langle \left(\hat{O} - \left\langle \hat{O} \right\rangle\right)^2 \right\rangle = \left\langle \hat{O}^2 \right\rangle - \left\langle \hat{O} \right\rangle^2$$

This quantity is *not an observable*, since it is *not a linear function* of ρ; i.e., it can depend on products such as $\rho_x \times \rho_z$.

Hand Image by iconicbestiary on Freepik

Ideal single qubit measurement with M shots



M shots with IID distribution

M outcomes: independent and identically distributed (iid) random variables

$$X_1, X_2, \ldots, X_M \in \Sigma$$

$$X_1, X_2, \ldots, X_M \sim \Pr\left[X = x\right] = \langle \hat{\mu}\left(x\right) \rangle$$

Empirical mean random variable (sample mean statistic)

$$S = \frac{1}{M} \sum_{m=1}^{M} X_m$$

Find the expectation value and variance of the empirical mean

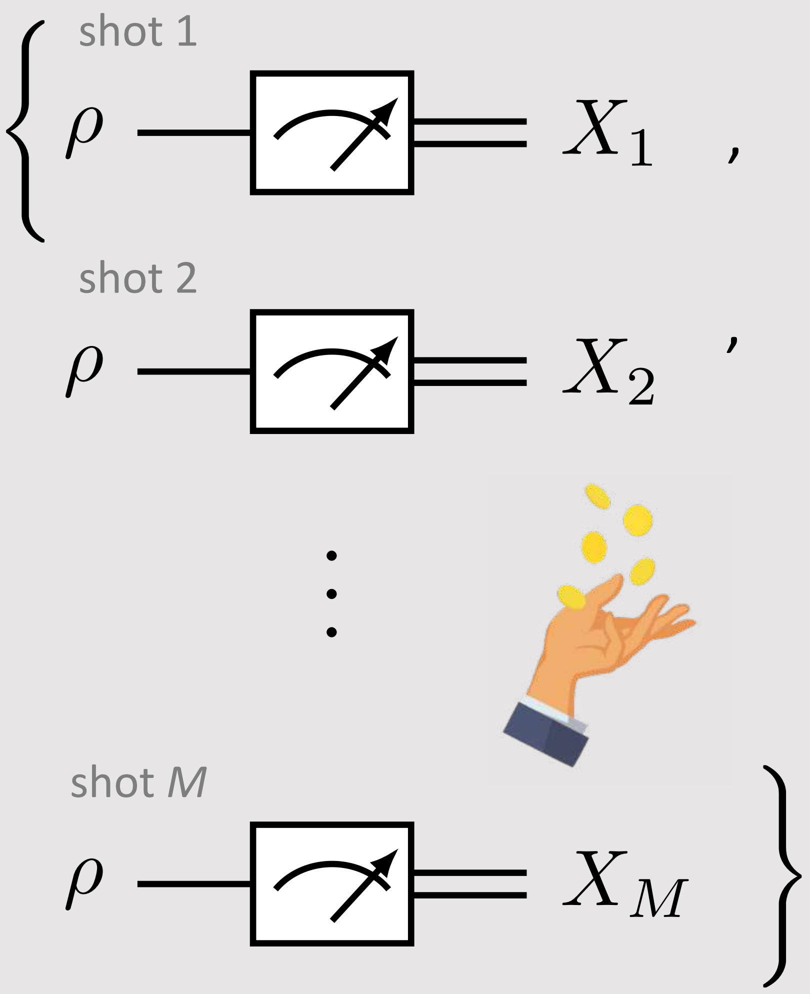
$$\mathrm{E}\left[X_m\right] = \mathrm{E}\left[X\right] = p \quad \forall m \in \{1, \ldots, M\}$$

$$\mathbb{V}\left[X_m\right] = \mathbb{V}\left[X\right] = p\left(1 - p\right) \quad \forall\, m \in \{1, \ldots, M\}$$

$$\mathrm{E}\left[aX_m + bX_n\right] = a\mathrm{E}\left[X_m\right] + b\mathrm{E}\left[X_n\right]$$

$$\forall m, n, \quad a, b \in \mathbb{C}$$

Ideal single qubit measurement with M shots



M shots with IID distribution

M outcomes: independent and identically distributed (iid) random variables

$$X_1, X_2, \ldots, X_M \in \Sigma$$

$$X_1, X_2, \ldots, X_M \sim \Pr\left[X = x\right] = \langle \hat{\mu}\left(x\right) \rangle$$

Empirical mean random variable

$$S = \frac{1}{M} \sum_{m=1}^{M} X_m$$

$$\begin{aligned} \mathrm{E}\left[S\right] &= \mathrm{E}\left[O(X)\right] \\ &= \sum_{x \in \Sigma} O(x) p(x) \quad = \quad \mathbf{o}^{\intercal} \mathbf{p} \\ &= \left\langle \sum_{x \in \Sigma} O(x) \mu\left(x\right), \rho \right\rangle \\ &= \langle\langle \hat{O} | \rho \rangle\rangle \quad = \quad \mathrm{Tr}\left(\hat{O}^{\dagger} \rho\right) \\ &= \left\langle \hat{O} \right\rangle \quad \text{(unbiased est)} \end{aligned}$$

$$\begin{aligned} \mathbb{V}\left[S\right] &= \mathbb{V}\left[\frac{1}{M} \sum_{m=1}^{M} X_m\right] \\ &= \frac{1}{M^2} \sum_{m=1}^{M} \mathbb{V}\left[X\right] \\ &= \frac{1}{M} \mathbb{V}\left[X\right] \\ &= \frac{p\left(1-p\right)}{M} \end{aligned}$$

Observable

Map outcomes:

$$x \mapsto O(x)$$

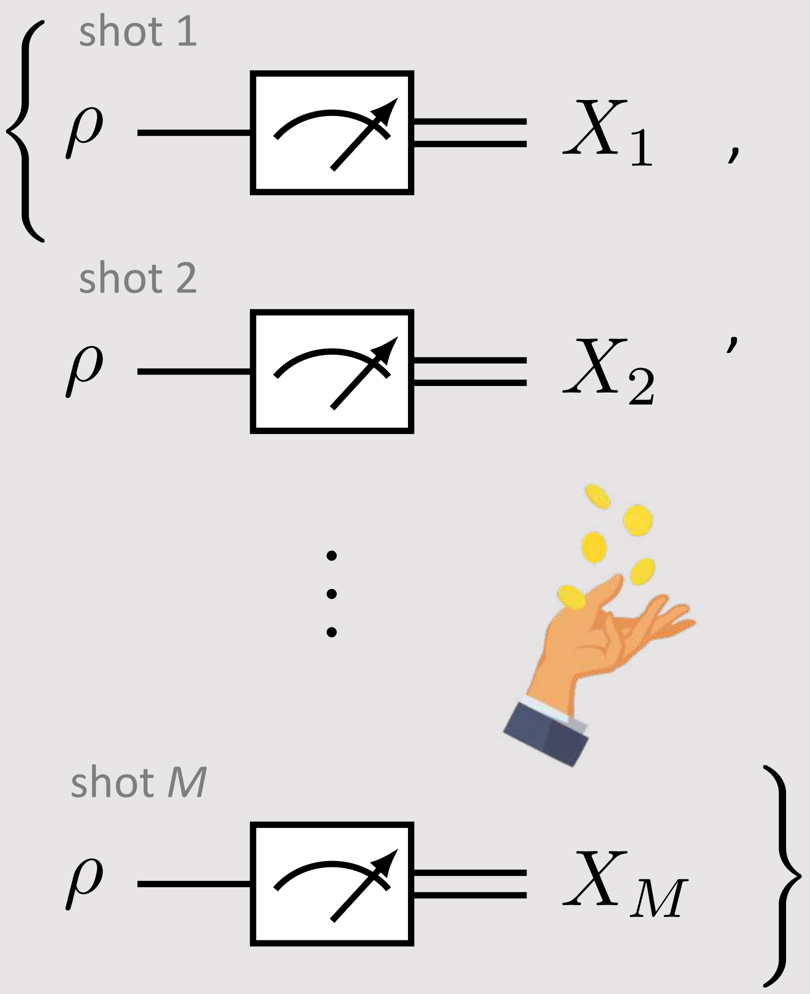
Example Pauli Z

$$\hat{O} = \sum_{x \in \Sigma} O\left(x\right) \hat{\mu}\left(x\right)$$

$$O(0) = +1$$

$$O(1) = -1$$

Connection to distributions



$$S = \frac{1}{M} \sum_{m=1}^{M} X_m$$

The sum of the zero-one outcomes *X*s is

$$M \times S \sim \mathrm{Binomial}\,(M, p)$$

The distribution of the average is a scaled binomial distribution

$$\Pr\left[S = \frac{k}{M}\right] = \binom{M}{k} p^k (1-p)^{M-k}$$

for $k = 0,\ldots,M$.

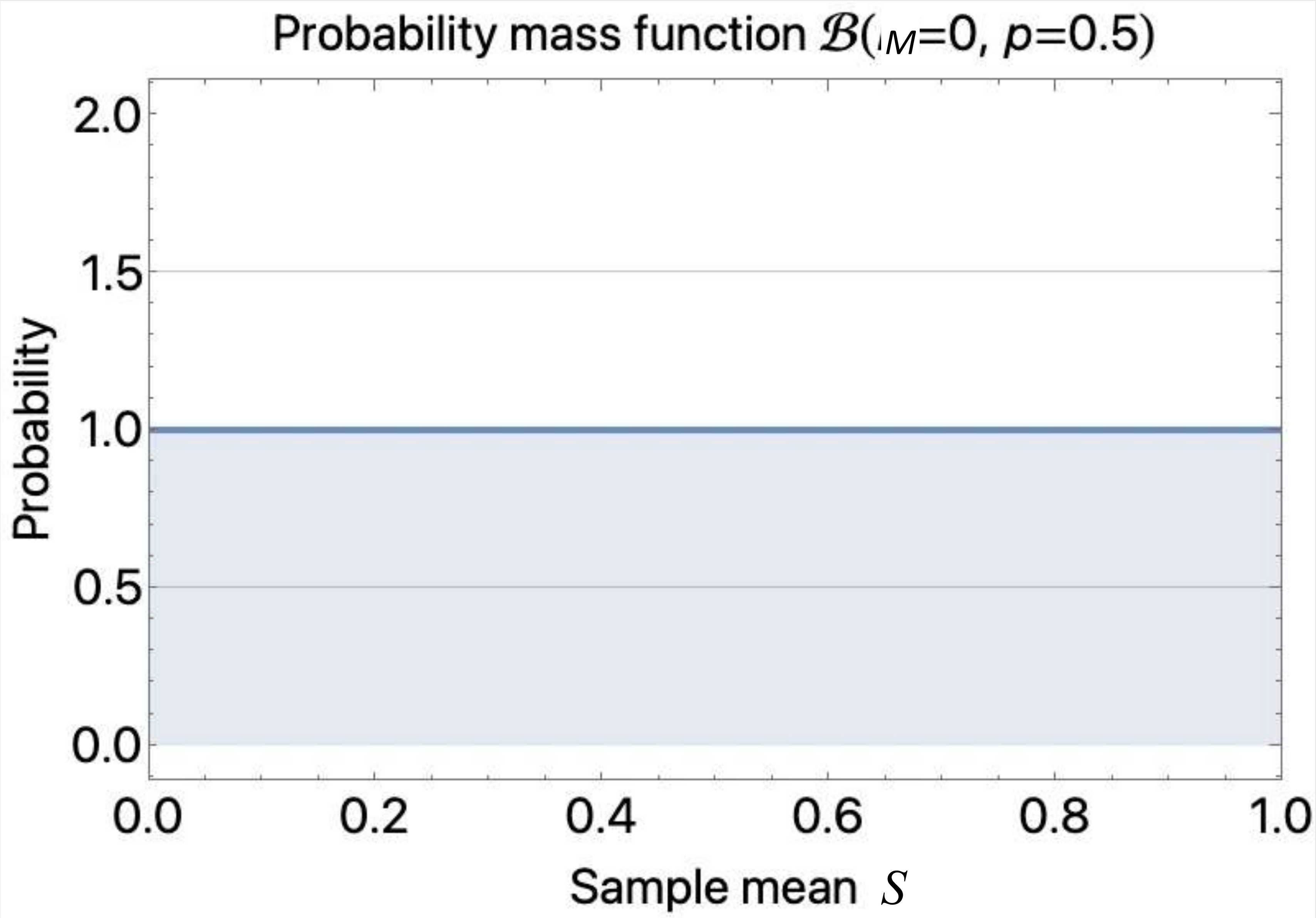
For large number of shots M, it is well approximated (central limit theorem; see next slide) by the normal distribution parameterized by its mean and standard deviation:

$$S \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{M}}\right)$$

The approximation becomes more accurate as M increases.

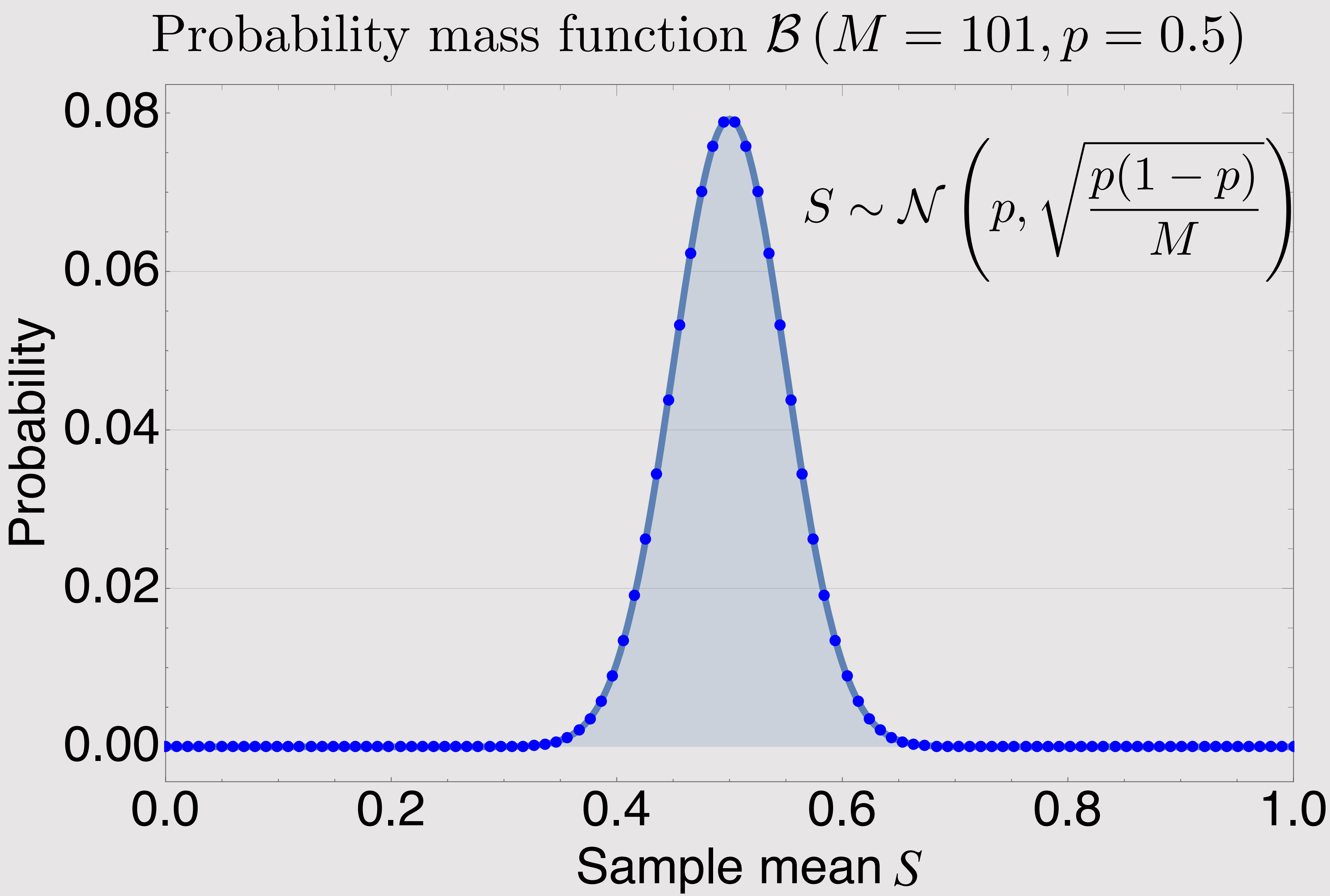
Hand Image by iconicbestiary on Freepik

Animation of convergence of shots expectation value and mean

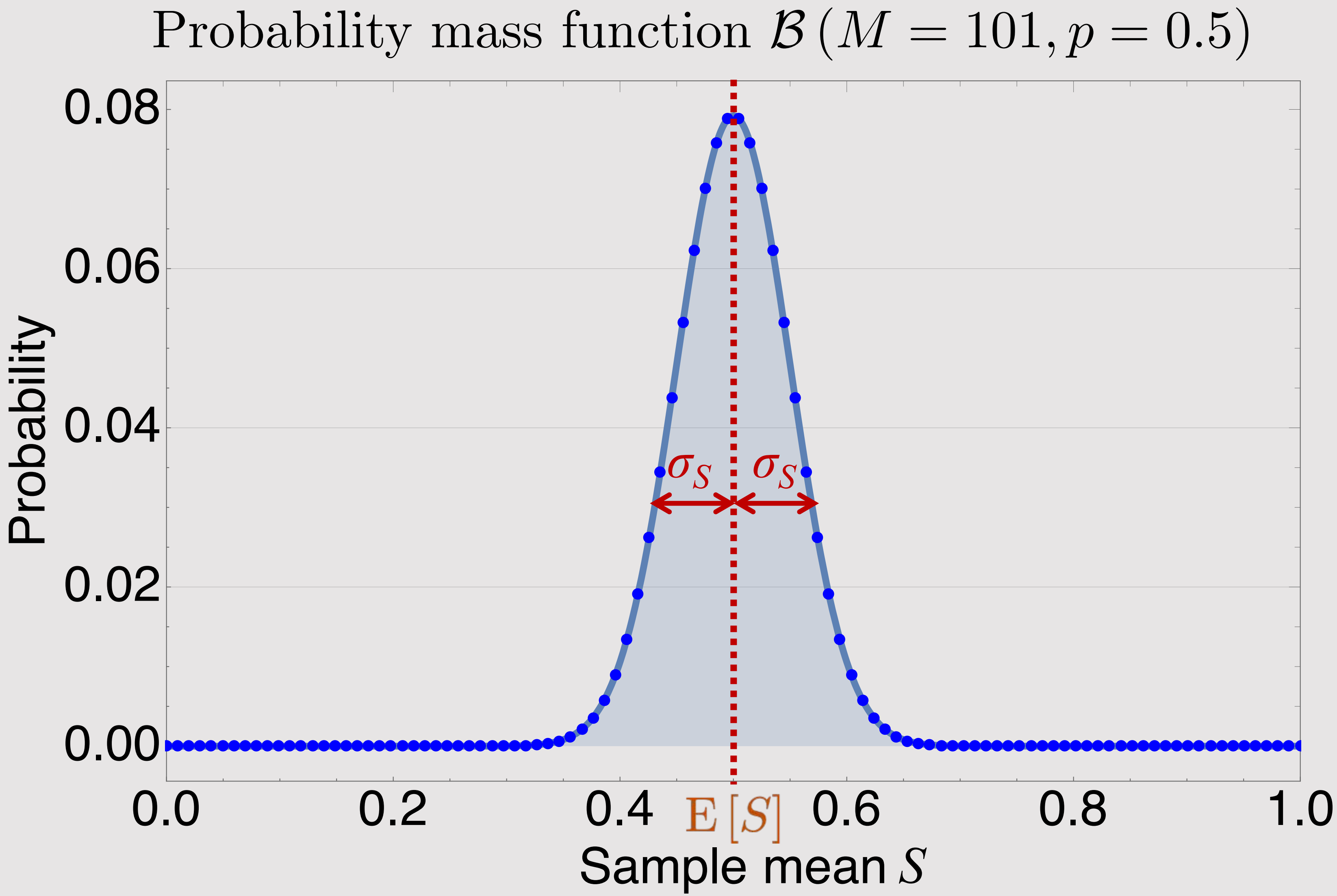


$$\Pr\left[S = \frac{k}{M}\right] = \binom{M}{k} p^k (1-p)^{M-k}$$

Sampled output distribution



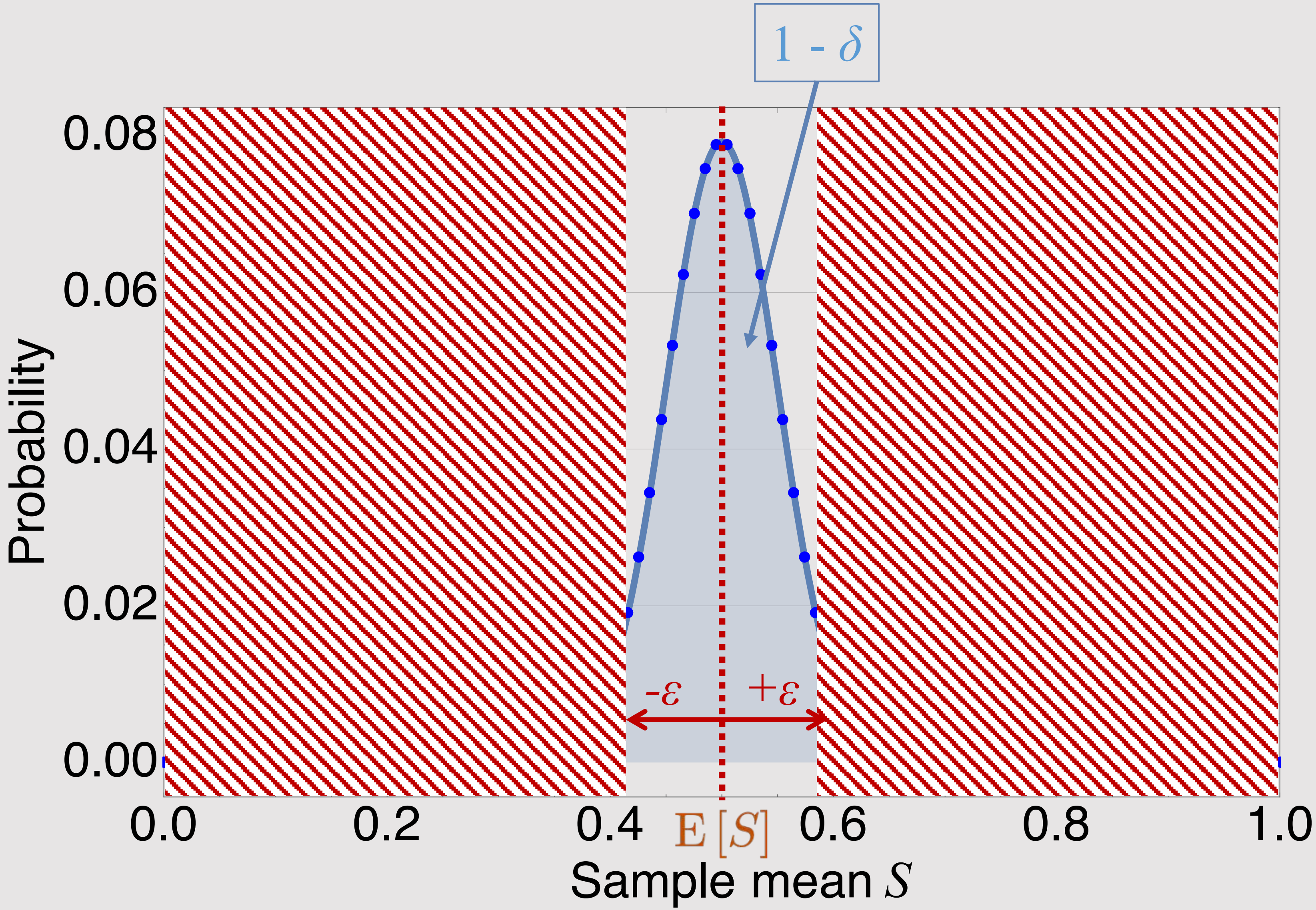
Properties of the output distribution



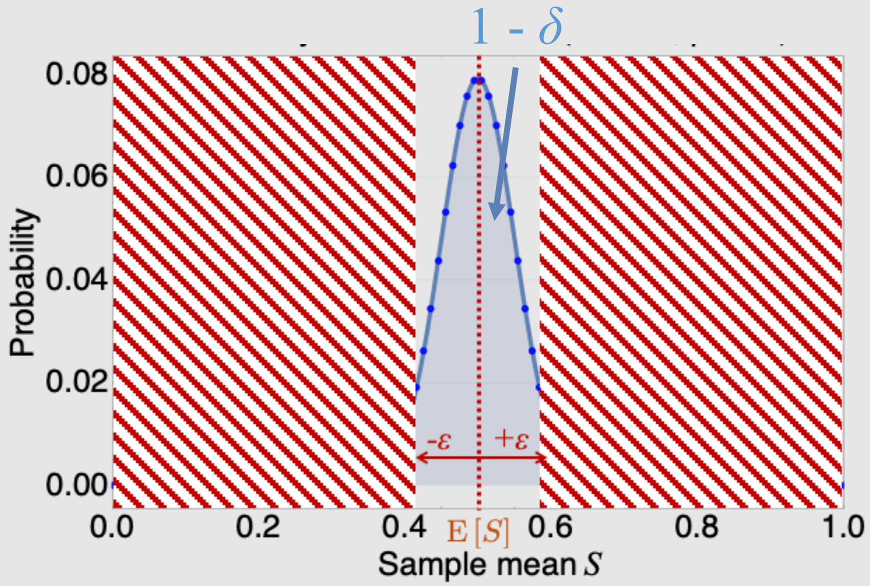
$$\mathrm{E}\left[S\right] = p$$

$$\sigma_S = \sqrt{\mathrm{Var}\left[S\right]} = \sqrt{\frac{p(1-p)}{M}}$$

Concentration Measure for Sampling Expectation Values



Error Bound on Quantum Expectation Values
Concentration Measure for Sampling Expectation Values



Chernoff-Hoeffding two-sided tail bound, given $|O(x)| \leq 1$ for

ε specified (additive) precision (worst case additive error) for S with success probability at least $1 - \delta$.

$$\Pr\left[\left|S - \left\langle \hat{O} \right\rangle\right| > \epsilon\right] \leq \delta := 2\exp\left(-\frac{1}{2}M\epsilon^2\right)$$

δ specific failure probability for meeting precision ε empirically.

Empirical mean & sample properties

$$S = \frac{1}{M}\sum_{m=1}^{M} O(X_m)$$

Tf. required number of shots is at least [or with high probability (greater than 2/3)]

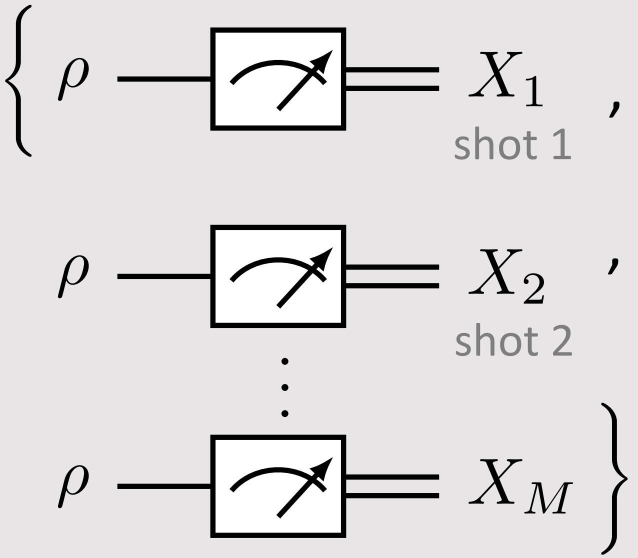
$$M \geq 2\epsilon^{-2}\log\left(2\delta^{-1}\right) \qquad \left[M \gtrapprox 4\epsilon^{-2}\right] \; .$$

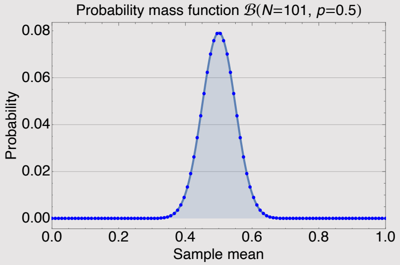
* Can find even tighter bound here owing to smaller [0,1] range

Note that this scales same way (mod δ) as the variance bound with ε = σ :

$$M \geq \frac{1}{4}\epsilon^{-2}$$

Ideal single qubit measurement with *M* shots



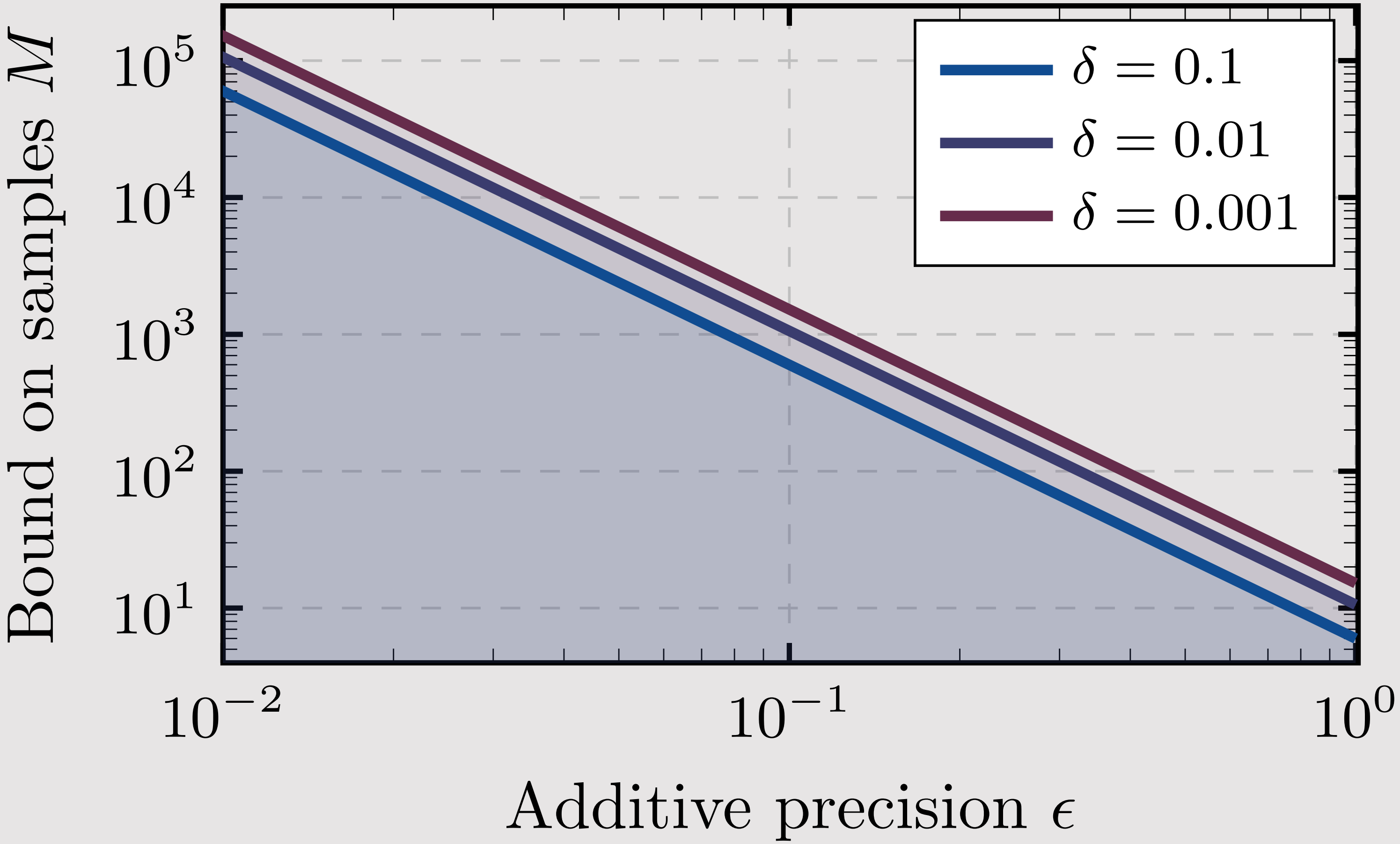


Observe that the **probability** δ is much cheaper than the **precision** ε.
Observe, n is not part of the equation.

"Knowing you are *not* wrong is cheaper than knowing you are right." - Derek

Tail bound on sample complexity

$$M \geq 2\epsilon^{-2} \log\left(2\delta^{-1}\right) \qquad \left[M \gtrsim 4\epsilon^{-2}\right]$$



Concentration inequalities and tail bounds

Making a list,
checking it twice,
going to see
which inequality
is nice!

Markov? Hoeffding?
Jensen? Chebyshev?
Chernoff?

https://www.zlatko-minev.com/blog/inequalities

1. Probability (Technical note 11.9 v0.6)

1A. Concentration inequalities and tail bounds

Unless otherwise specified, all variables are real $\mathbb{R}$. Inequalities come as one-sided $\Pr(\cdots \leq \cdots)$ and two-sided $\Pr(\lvert\cdots\rvert \leq \cdots)$. Notation: X is a random variable, $\mu := \mathbb{E}[X]$, $\sigma^2 := \mathrm{Var}[X]$, $S_n := X_1 + \cdots + X_n$.

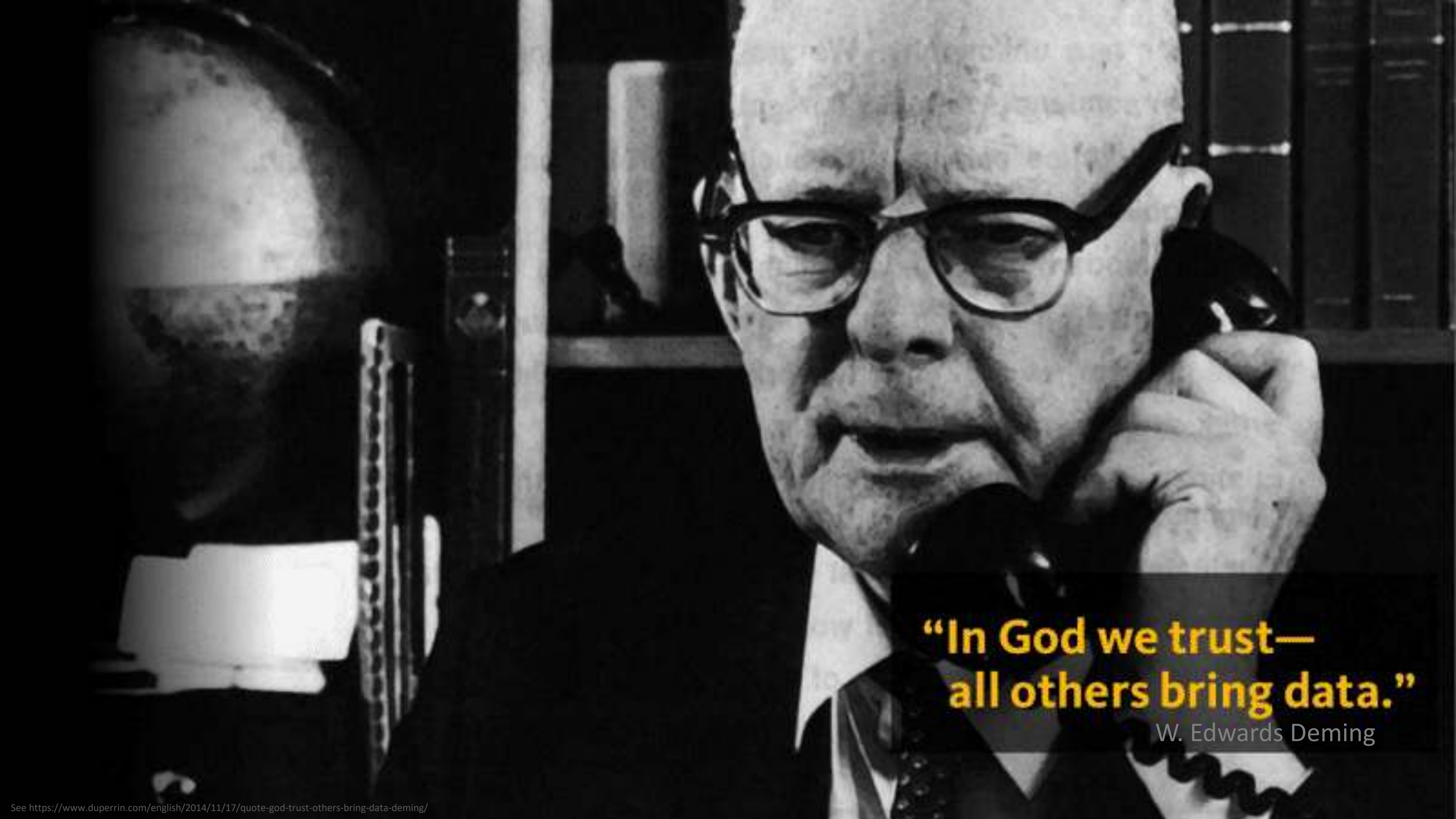
Inequality	Conditions		Common form	Notes / Alternate form	
			Single random variable		
Markov[1]	Non-negative	$X \geq 0$	$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$ $\forall a > 0$	$\Pr[X \geq kE[X]] \leq \frac{1}{k}$ $k > 1$	[3, Sec. 5.1][6, Thm 1.13]
extension	+ non-negative, strictly increasing func Φ	$X \geq 0$ $\Phi(X) \geq 0$ increasing	$\Pr[X \geq a] = \Pr[\Phi(X) \geq \Phi(a)] \leq \frac{\mathrm{E}(\Phi(X))}{\Phi(a)}$	$\forall a > 0$	Wiki
Reverse Markov	upper-bounded by U (can be positive)	$\max X = U$	$\Pr[X \leq a] \leq \frac{U - \mathrm{E}[X]}{U - a}$	$\forall a > 0$	[1, Sec. 3.1]
Chebyshev[2]	Finite mean and variance	$\mathbb{E}[X], \mathrm{Var}[X]$ finite	$\Pr[\lvert X - \mathbb{E}[X]\rvert \geq a] \leq \frac{\sigma^2}{a^2}$	$\Pr[\lvert X - \mathbb{E}[X]\rvert \geq a \cdot \sigma] \leq \frac{1}{a^2}$ $\forall a > 0,\ \sigma^2 = \mathrm{Var}[X]$	[1, Sec. 3.2] [3, Sec. 5.1][2, Thm 18.11]
Cantelli	Improved Chebyshev	(same; but one-sided)	$\Pr[X - \mathbb{E}[X] \geq a]) \leq \frac{\sigma^2}{\sigma^2 + a^2}$	$\forall a > 0,\ \sigma^2 = \mathrm{Var}[X]$	Wiki
Chernoff[3]	Generic		$\Pr[X \geq a] = \Pr[e^{tX} \geq e^{ta}]$	$\forall t > 0, \quad a \in \mathbb{R}$	[1, Sec. 3.3]
Jensen		$f : \mathbb{R} \to \mathbb{R}$; f convex	$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$		[3, Prob. 5.3][6, Thm 1.14]
Hoeffding's lemma		$\mathbb{E}[X] = \mu$ $a \leq X \leq b$	$\mathbb{E}[e^{\lambda X}] \leq e^{\lambda\mu} e^{\frac{\lambda^2 (b-a)^2}{8}}$	$\lambda \in \mathbb{R}$	[1, Sec. 3.4]
			Sum of random variables		
Chernoff-Hoeffding (one-sided)	n independent random vars	$X_1, \ldots, X_n$ indep $S_n = X_1 + \cdots + X_n$ $X_i \in [a_i, b_i]$ $\forall i$	$\Pr[S_n - \mathbb{E}[S_n] \geq t] \leq \exp\left(\frac{-2t^2 n^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$		[1, Sec. 3.5]
(two-sided)[4]	(same as above)		$\Pr[\lvert S_n - \mathbb{E}[S_n]\rvert > t] \leq 2\exp\left(\frac{-2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$	$\forall t \in (0, \frac{1}{2})$	[5, Thm.1.1]
(two-sided iid)	same plus iid, range, mean μ for each	$X_1, \ldots, X_n \in [0,1]$ $\mathbb{E}[X_i] = \mu$ iid	$\Pr\left[\left\lvert \frac{S_n}{n} - \mu \right\rvert \geq \epsilon\right] \leq 2\exp(-2n\epsilon^2)$	$\forall \epsilon > 0$	[6, Thm 1.16]
Thm 1.3	n independent random vars	$X_1, \ldots, X_n$ indep $S_n = X_1 + \cdots + X_n$	$\Pr[S_n - \mathbb{E}[S_n] > \epsilon] \leq 2\exp\left(\frac{-\epsilon^2}{4\sum_{i=1}^n \mathrm{Var}[X_i]}\right)$	$\epsilon \in (0, 2\,\mathrm{Var}[S_n] / (\max_i \lvert X_i - \mathbb{E}[X_i]\rvert))$	[5, Thm. 1.3]
Azuma					
Weak law of large numbers	n independent iid random vars	$X_1, \ldots, X_n$ indep $\mathbb{E}[X_i] = \mu$ iid	$\lim_{n\to\infty} \Pr[\lvert \frac{1}{n} S_n - \mu\rvert \geq \epsilon] = 0$	$\forall \epsilon > 0$	[3, Sec. 5.2][6, Thm 1.15]
Strong law of large numbers	(same)	(same)	$\Pr[\lim_{n\to\infty} \frac{1}{n} S_n = \mu] = 1$		[3, Sec. 5.5]
	Advanced				
Bennett	n independent zero-mean	$X_1, \ldots, X_n$ indep $\mathbb{E}[X_i] = 0$ iid	$\Pr[S_n > \epsilon] \leq \exp\left(-n\sigma^2 h\left(\frac{\epsilon}{n\sigma^2}\right)\right)$	$\sigma^2 := \frac{1}{n}\sum_{i=1}^n \mathrm{Var}[X_i]$, $\forall \epsilon > 0$, $h(a) := (1+a)\log(1+a) - a$ for $a \geq 0$	[1, 4.1]
Bernstein	(same)	(same)	$\Pr[S_n > \epsilon] \leq \exp\left(\frac{-n\epsilon^2}{2(\sigma^2 + \epsilon/3)}\right)$	(same)	[1, 4.2]
Efron-Stein	scalar func of vars $f : \chi^n \to \mathbb{R}$	$X_1, \ldots, X_n$ indep w/ values in set χ	$\mathrm{Var}[Z] \leq \sum_{i=1}^n \mathbb{E}\left[(Z - \mathbb{E}_i[Z])^2\right]$	$Z := g(X_1, \ldots, X_n)$ $\mathbb{E}_i[Z] := \mathbb{E}[Z \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_n]$	[1, 4.3]
McDiarmid's	scalar func of vars $f : \chi^n \to \mathbb{R}$	$X_1, \ldots, X_n$ indep w/ values in set χ	$\Pr[f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)] \geq \epsilon] \leq \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^n c_i^2}\right)$	condition: c-bounded difference property $\forall \epsilon > 0$ $\lvert f(X_1, \ldots, X_i, \ldots, X_n) - f(X_1, \ldots, X_i', \ldots, X_n)\rvert \leq c_i$	[1, 4.4]

[1] Markov's inequality bounds the first moment of random variable. Use it when a constant probability bound is sufficient [1, Sec. 3.3].

[2] Chebyshev is derived from Markov. It bounds the second moment. It is the appropriate one when the variance σ is known. If σ is unknown, we can use the bounds of $X \in [a, b]$.

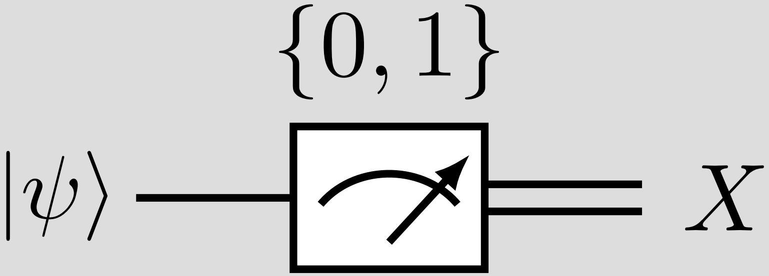
[3] Chernoff bound is used to bound the tails of the distribution for a sum of independent random variables. By far the most useful tool in randomized algorithms [1, Sec. 3.3].

[4] This probability can be interpreted as the level of significance ϵ (probability of making an error) for a confidence interval around the mean of size 2ϵ. Therefore, we require at least $\log(2\alpha)/2t^2$ samples to acquire $1 - \alpha$ confidence interval $\mathbb{E}[X] \pm t$.



See https://www.duperrin.com/english/2014/11/17/quote-god-trust-others-bring-data-deming/

From data to error estimations



Suppose we took M=10 measurements with all zero outcomes:
0000000000

What is our estimate of p?
What is the error bar on our estimate of p?

p = 0?

The sunrise problem (see also German tank problem, Rule of succession)

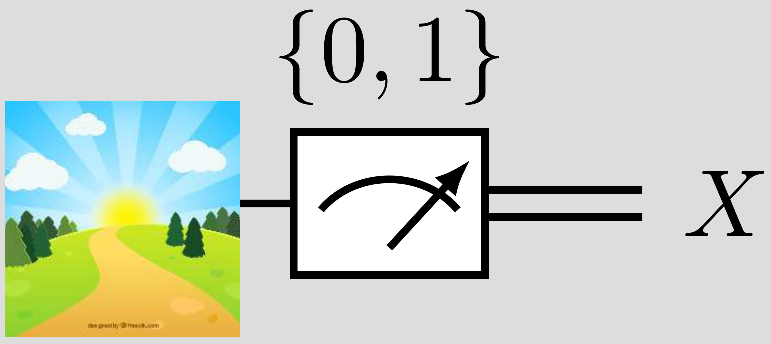
$$\Pr\left[\text{sun rise tomorrow}|\text{It rose } M \text{ times before}\right] = \frac{M+1}{M+2}$$

For small M

Bayesian and likelihood estimation

Connections to
- Maximum likely estimation (MLE)
- Bayesian estimation
- …

From data to error estimations

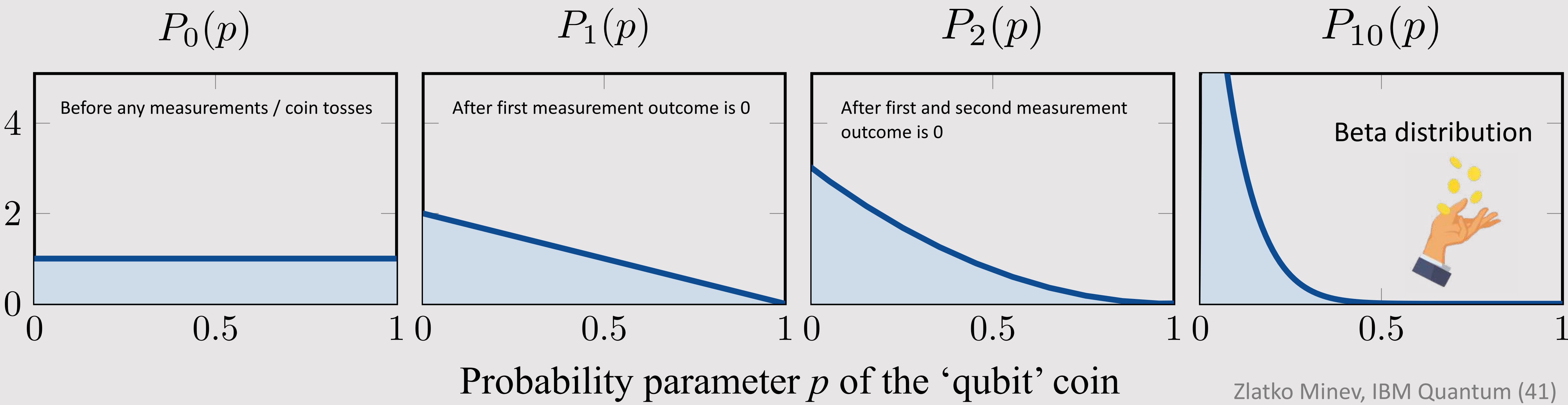


Suppose we took M=10 measurements with all zero outcomes: 0000000000

What is our estimate of p? What is the error bar on our estimate of p?

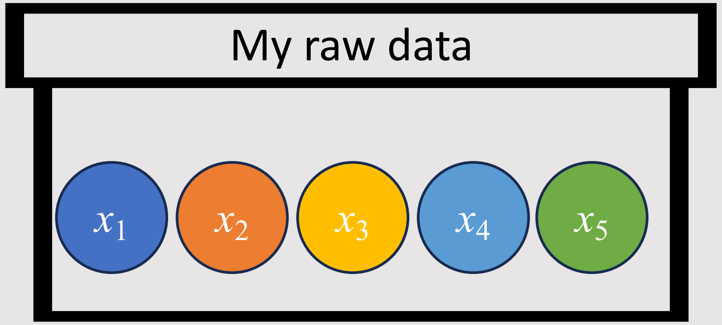
Bayesian update

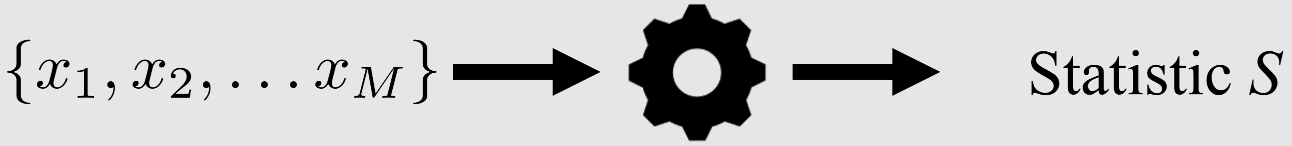
$$P'(p|x) = \frac{P(x|p)\,P(p)}{P(x)} = \frac{P(x|p)}{P(x)}$$



Bootstrapping: A very brief overview

Bootstrapping is a powerful resampling technique used for estimating properties or even the distribution of a statistic.



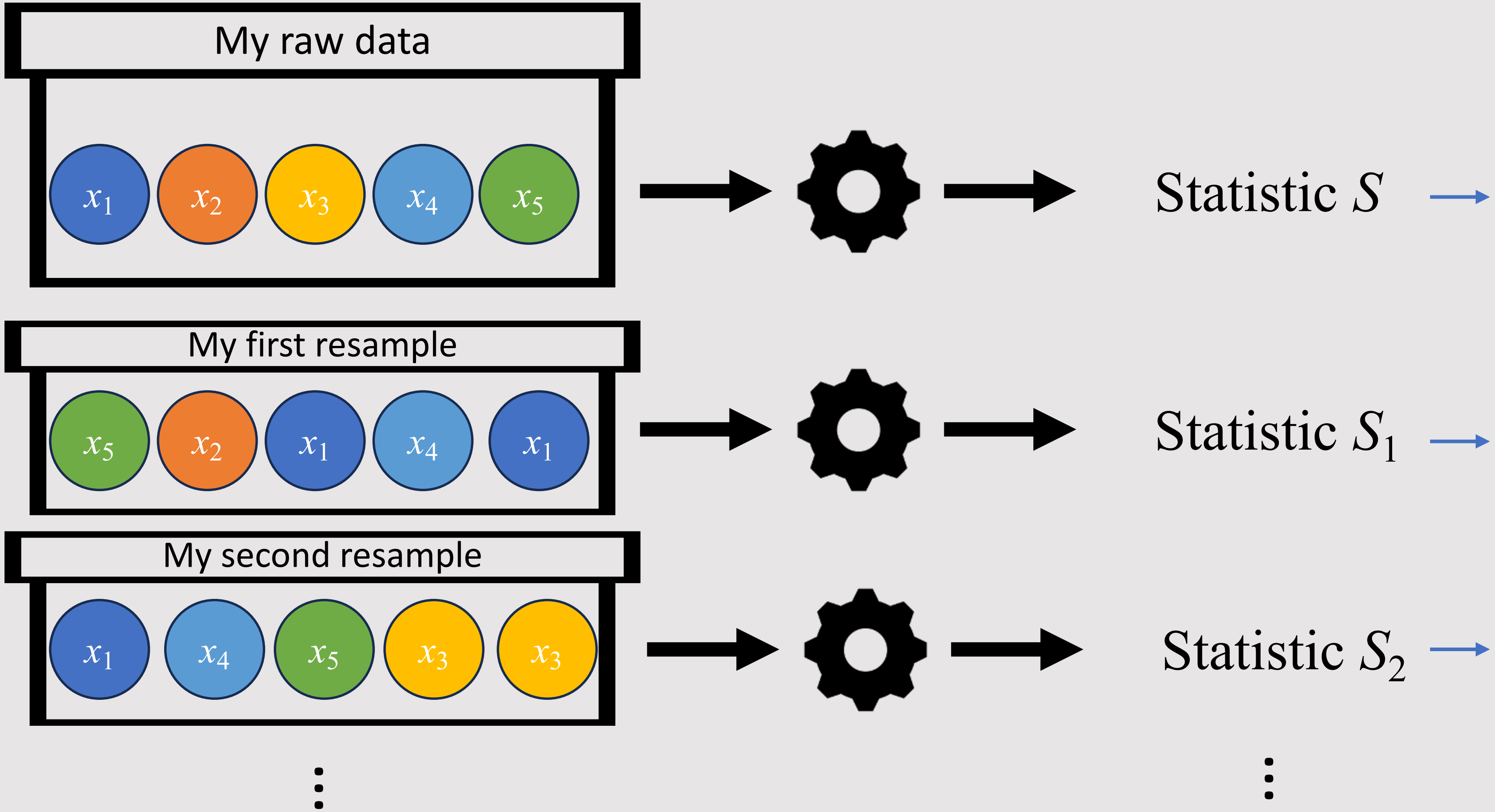


Statistic	Description / Estimator
$P_x := \frac{1}{M}\sum_{m=1}^{M}[X_m = x]$	Empirical probability to measure $X = x$
$S := \frac{1}{M}\sum_{m=1}^{M}[X_m = x]$	Sample mean
$V := \frac{1}{M-1}\sum_{m=1}^{M}\left(X_m - \frac{1}{M}\sum_{m=1}^{M} X_m\right)^2$	Sample variance (unbiased; Bessel's correction)
$f(X_1, \ldots, X_M) = \frac{1}{M}\sum_{m=1}^{M} f(X_m)$	Average of generic function of the outcome
$H := -\sum_{x\in\Sigma} P_x \log(P_x)$	Classical entropy of the outcomes
$H_\alpha := \frac{1}{1-\alpha}\log\left(\sum_{x\in\Sigma} P_x^\alpha\right)$	Classical Renyi entropy of the outcomes

⋮

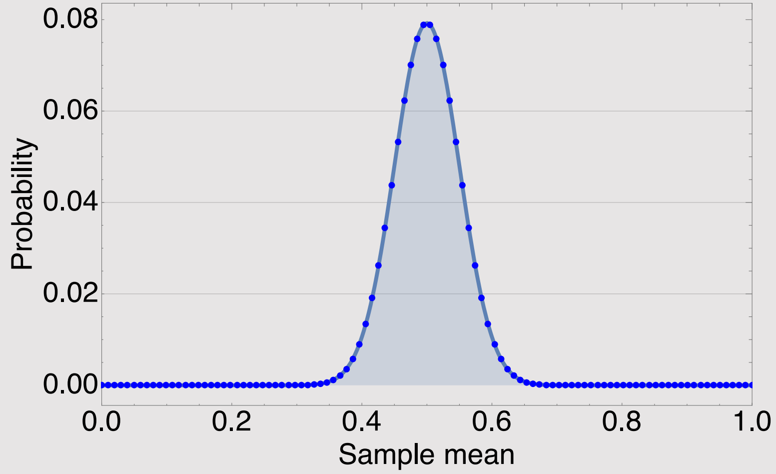
Bootstrapping: A very brief overview

Bootstrapping is a powerful resampling technique used for estimating properties or even the distribution of a statistic. By repeatedly resampling with replacement from the observed raw data, we calculate the statistic on each resample.



Bootstrap Empirical Mass Function

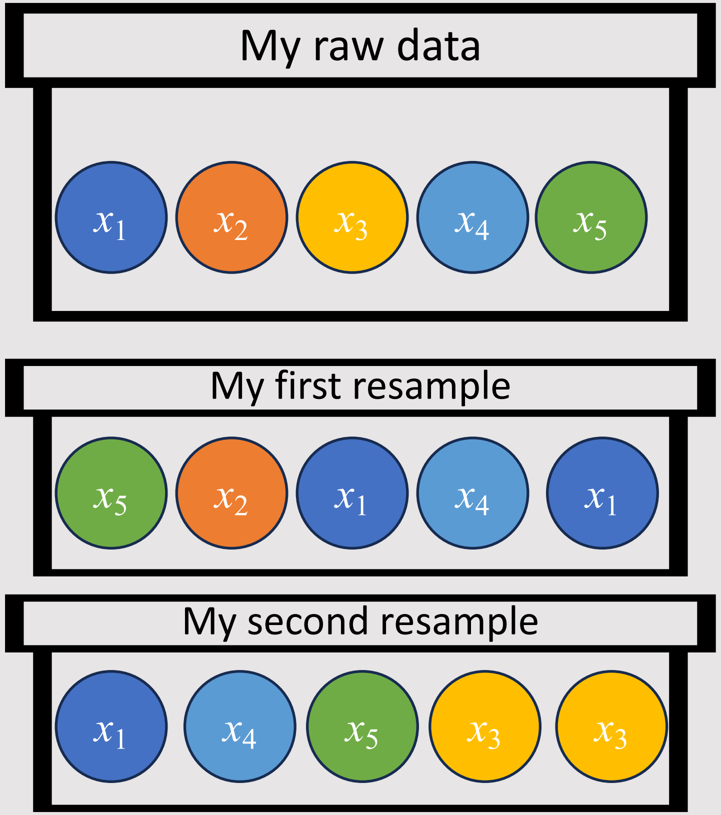
$$\hat{p}(x) = \frac{1}{M} \sum_{m=1}^{M} \mathbb{I}(x_m = x)$$



This yields an empirical estimate of the **standard error** – through the bootstrapped static ensamble used to calculate the standard deviations σ_S.

Bootstrapping: A very brief overview

Bootstrapping is a powerful resampling technique used for estimating properties or even the distribution of a statistic. By repeatedly resampling with replacement from the observed raw data, we calculate the statistic on each resample.



Advantages

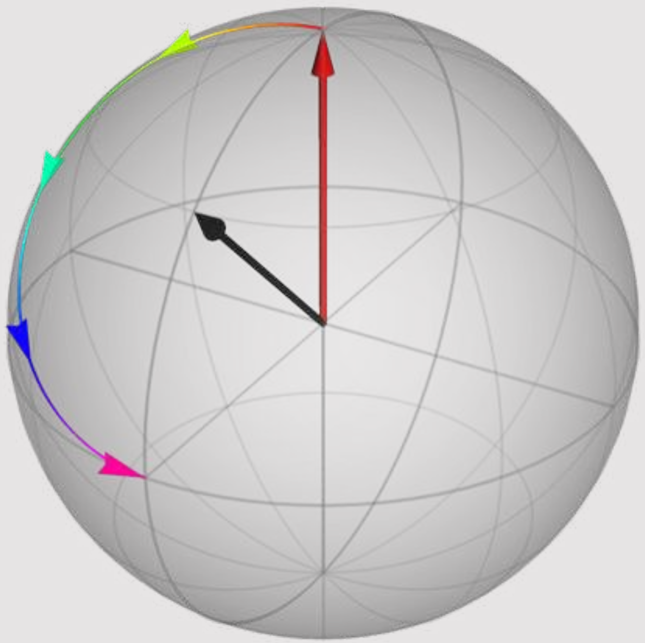
- applicable to any non-linear data processing, complex statistical estimates, etc. where theory is unclear.
  - versatility: can be applied to almost any data and any statistical estimator.
- minimal assumptions
  - non-parametric method
  - requires fewer assumptions about data
- simplicity, ease of implementation, broadly used

Key things to remember

- Mean of the bootstrap ensemble: same as the raw data, no new information
- Convergence: the bootstrap distribution converges to the true distribution of the statistic as the sample size increases. It is asymptotically correct

Statisticians have a different version of
The Boy Who Cried Wolf.

smbc-comics.com/comic/false-positives



Rotating measurement bases

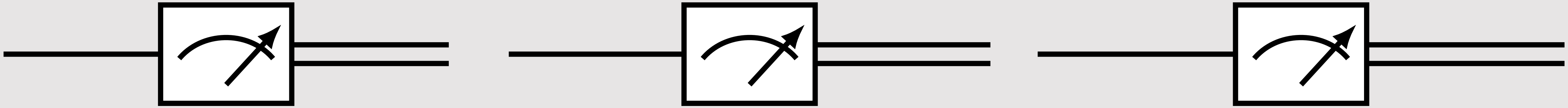
Measuring in bases other than the computational basis

Z $\{|0\rangle, |1\rangle\}$

X $\{|+\rangle, |-\rangle\}$

Y $\{|+i\rangle, |-i\rangle\}$

ρ



$$\hat{\mu}(+) = |+\rangle\langle+|$$
$$\hat{\mu}(-) = |-\rangle\langle-|$$

Description of the meter

The **quantum-to-classical rule** (postulate of quantum)

$$p(x) \coloneqq \Pr(X = x) = \langle \hat{\mu}(x), \rho \rangle = \mathrm{Tr}\left(\hat{\mu}^{\dagger}(x)\,\rho\right)$$
$$= \langle\langle \hat{\mu}(x) | \rho \rangle\rangle$$

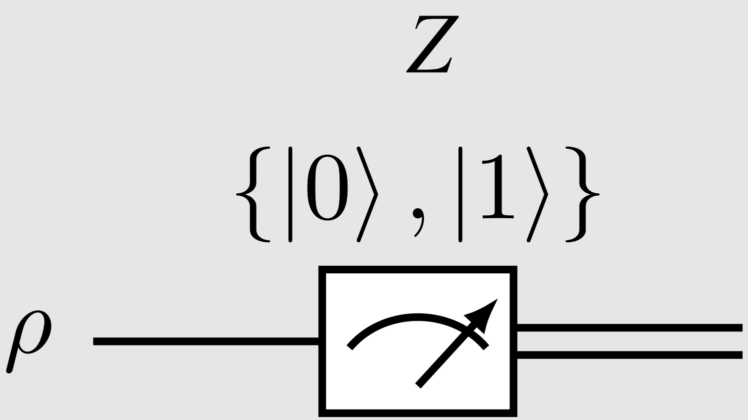
John von Neumann

Image source: LANL

Max Born

Image source: Public domain

Relating to the measurement we have at our disposal



$$
\begin{aligned}
\mathrm{Tr}\left[\mu^{\dagger}(x)\,\rho\right] &= \mathrm{Tr}\left[\mu^{\dagger}(x)\,I\rho I\right] \\
&= \mathrm{Tr}\left[\mu^{\dagger}(x)\,U^{\dagger}U\rho' U^{\dagger}U\right] \\
&= \mathrm{Tr}\left[\left(U\mu^{\dagger}(x)\,U^{\dagger}\right)\left(U\rho' U^{\dagger}\right)\right] \\
&= \mathrm{Tr}\left[\left(U^{\dagger}\mu(x)\,U\right)^{\dagger}\left(U\rho' U^{\dagger}\right)\right] \\
&= \mathrm{Tr}\left[\mu'(x)\,\rho'\right]
\end{aligned}
$$

$$\mu'(x) \coloneqq U^{\dagger}\mu(x)\,U$$

Description of the meter

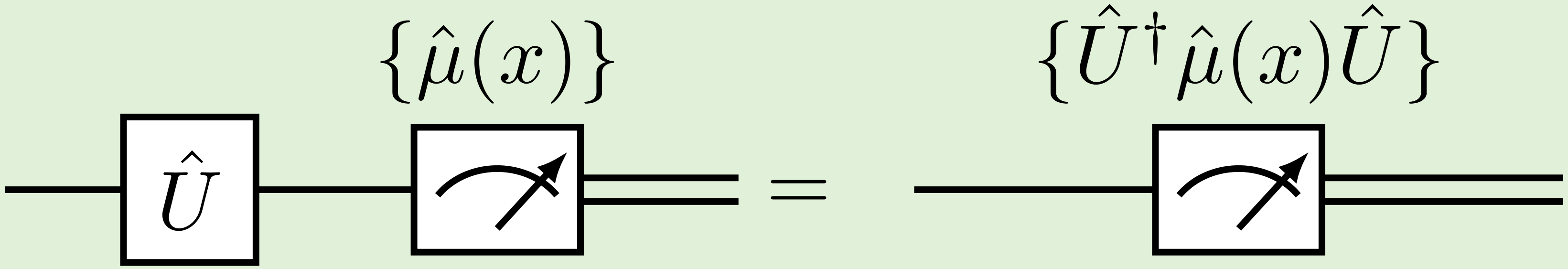
The **quantum-to-classical rule** (postulate of quantum)

$$
\begin{aligned}
p(x) \coloneqq \Pr(X = x) &= \langle \hat{\mu}(x), \rho \rangle = \mathrm{Tr}\left(\hat{\mu}^{\dagger}(x)\,\rho\right) \\
&= \langle\!\langle \hat{\mu}(x) | \rho \rangle\!\rangle
\end{aligned}
$$

John von Neumann
Image source: LANL

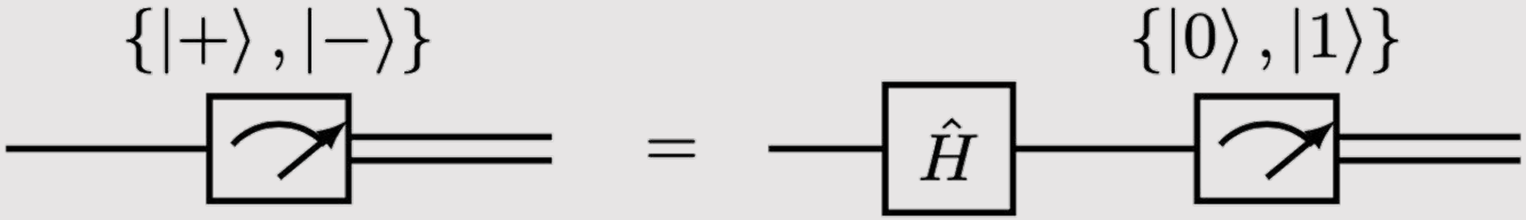
Max Born
Image source: Public domain

Example: Change of measurement basis

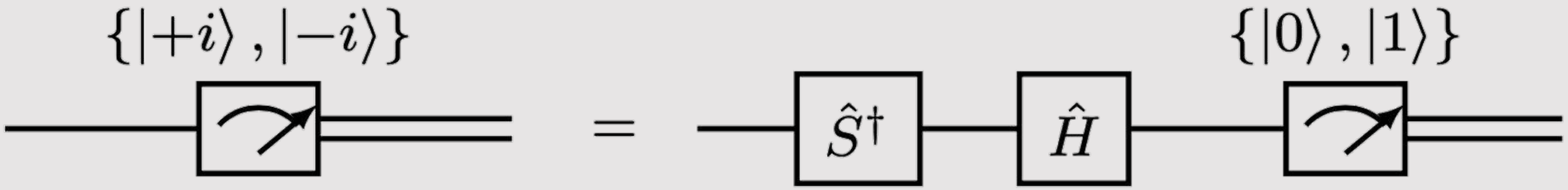


Examples:

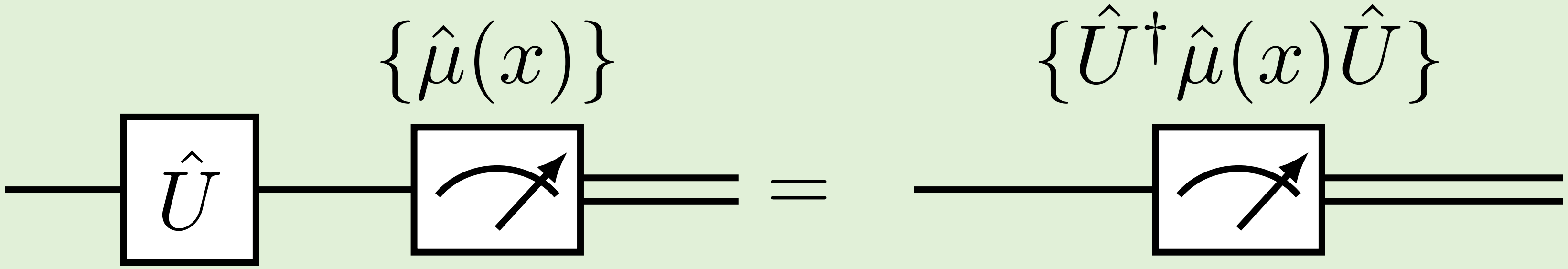
X:



Y:



Example: Change of measurement basis



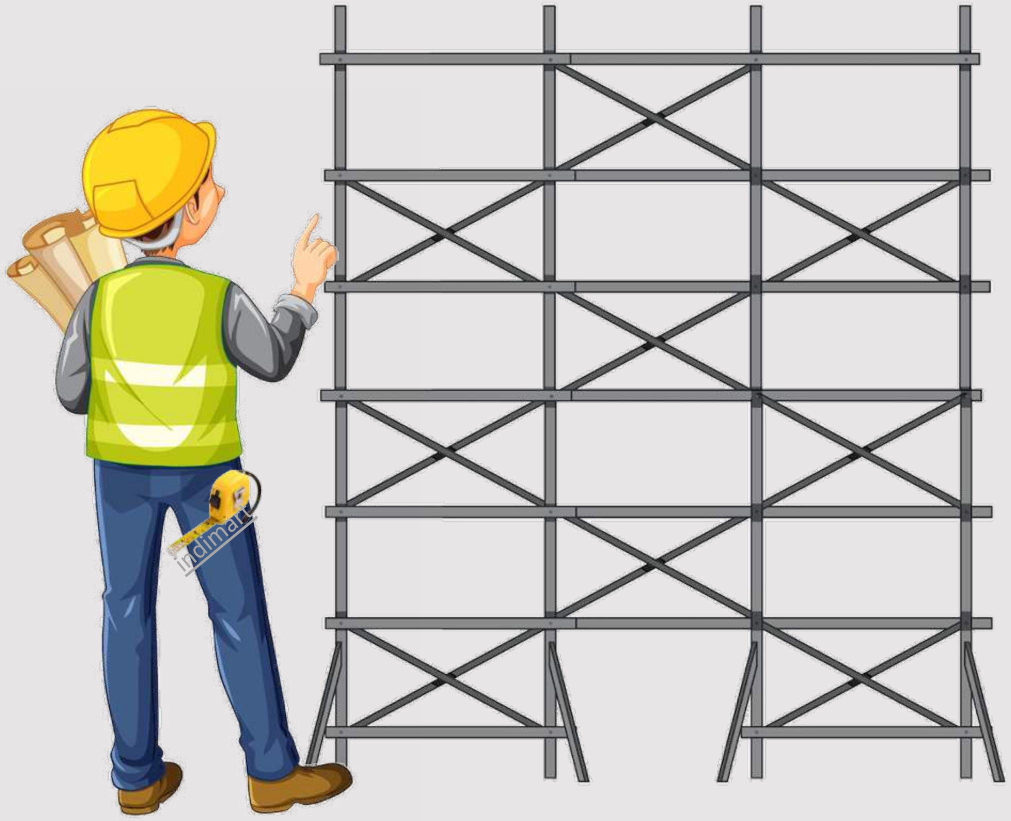
Examples:

	Z basis (standard)	X basis (from Z)	Y basis (from Z)
$\hat{U} =$	$\hat{I}$	$\hat{H}^\dagger = \hat{H}$	$\left(\hat{S}\hat{H}\right)^\dagger = \hat{H}\hat{S}^\dagger$
$\hat{\mu}\,(0) =$	$\lvert 0\rangle\langle 0\rvert = \frac{1}{2}\left(\hat{I}+\hat{Z}\right)$	$\lvert +\rangle\langle +\rvert = \frac{1}{2}\left(\hat{I}+\hat{X}\right)$	$\lvert +i\rangle\langle +i\rvert = \frac{1}{2}\left(\hat{I}+\hat{Y}\right)$
$\hat{\mu}\,(1) =$	$\lvert 1\rangle\langle 1\rvert = \frac{1}{2}\left(\hat{I}-\hat{Z}\right)$	$\lvert -\rangle\langle -\rvert = \frac{1}{2}\left(\hat{I}-\hat{X}\right)$	$\lvert -i\rangle\langle -i\rvert = \frac{1}{2}\left(\hat{I}-\hat{Y}\right)$

That’s our scaffolding

All else follows from here

Building models and assumption



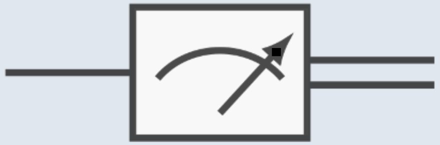
cartoon: brgfx from freepik

Measurement in Quantum Computers

Ideal

Measurement theory 101

- Why care?
- Formulation



Single qubit example

- Single shot
- Many shots
- Statistics, unbiased estimators
- Bounds, Chernoff-Hoeffding inequality
- Error bars on experimental data: likelihood
- Bootstrapping
- Measurement bases

Scaling to *n* qubits

- Single shot preliminaries
- Binary and Pauli strings
- Walsh–Hadamard
- Many shots
- Entangled measurements
- …

* Coin-toss icon created by Good Ware - Flaticon