Sarang Gopalakrishnan Boulder, 2023

Learning from measurements (and related topics)

Lec 1: Error correcting codes (where is the info?)

Lec 2: Learning classical labels (how do you get at it?)

Lec 3: Using measurement data for state preparation (what do you do with it?)

KEY REFS:

- Preskill, Notes on QI, Ch. 3, Ch. 10.
- Li and Fisher, PRB 103, 104306 (2021).
- Choi, Bao, Qi, Altman, PRL 125, 030505 (2020).
- Gullans, Huse, PRL 125, 070606 (2020).
- Hayden, Preskill, JHEP 09 (2007), 120.
- Schumacher, Nielsen, PRA 54, 2629 (1996).

- Barratt et al., PRL 129, 120604; 200602 (2022).

- Napp et al., PRX 12, 021021 (2022).
- Bao, Block, Altman, 2110.06963.
- Lin, Ye, Zou, Sang, Hsieh, 2205.05692.

1/ Information and errors

1.1/ Error correction from 35,000 feet.

× Classical case (repetition)

Encoding is a map from little to big strings: e.g.

$$\underset{\text{message}}{\{0,1\}} \longrightarrow \underset{\text{codewords}}{\{000, 111\}} \subseteq (\mathbb{Z}_2)^{\otimes 3}$$

The point is to protect info through redundancy:

e.g. $0 \xrightarrow[\text{encode}]{} 000 \xrightarrow[\text{error}]{} 001 \xrightarrow[\text{decode}]{} 000 \ (\longrightarrow 0)$

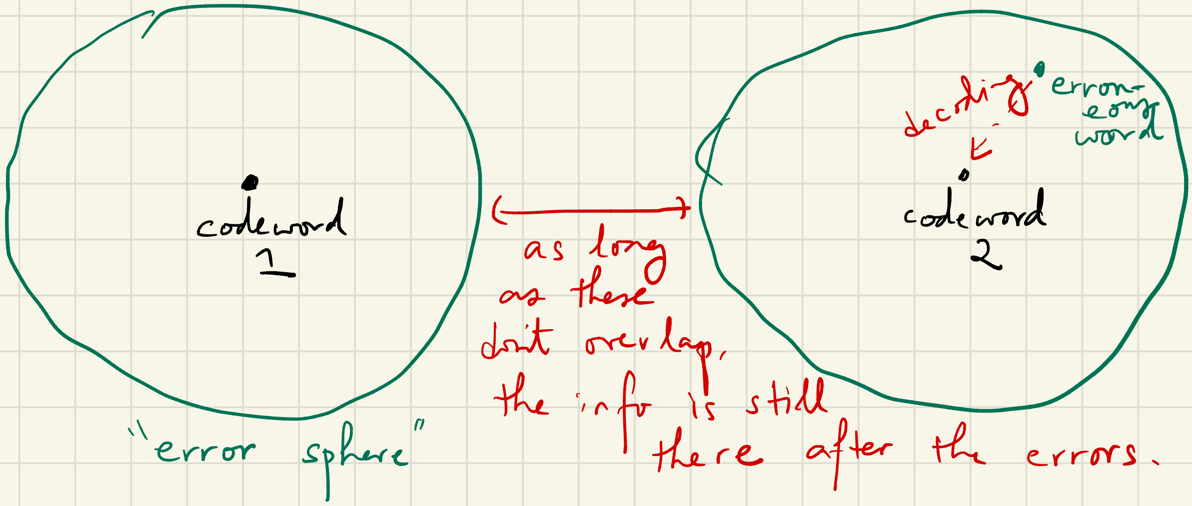
This procedure can correct errors provided that errors map distinct codewords to disjoint sets of strings.

If distinct codewords have "Hamming distance" d, they can tolerate $< d/2$ errors.

i.e. differ in d locations

Code distance = minimum Hamming distance between codewords.

Note: classically you can copy data, so it is irrelevant whether the environment learned the message.



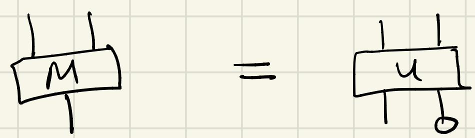
Note: we can talk about error correction w/o thinking about explicit decoders.

× Quantum case.

Some reminders:

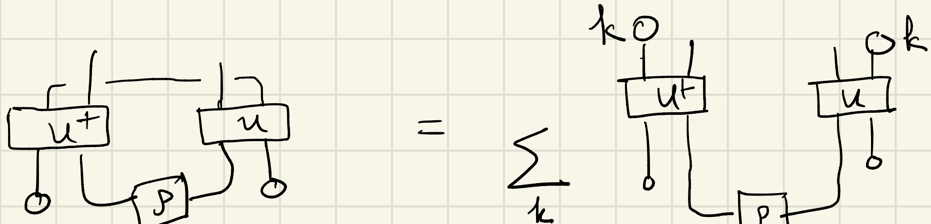
× An <u>isometry</u> $M : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ $(\dim \mathcal{H}_2 \geq \dim \mathcal{H}_1)$ preserves inner products. $M^\dagger M = \mathbb{1}_{\mathcal{H}_1}$.

Diagrammatically:



(i.e, it's a unitary w/ fixed input legs.)

× A quantum channel is a map from legal density matrices to legal density matrices — involving a system interacting w/ an environment that is then traced out.



Stinespring form Kraus form ("measure then throw out")

Quantum error correction: embed states in R in a larger space Q

$$\rho \in R \xrightarrow[\substack{\text{isometry} \\ \mathcal{I}}]{\text{encoding}} \tilde{\rho} \in Q \xrightarrow[\substack{\text{channel} \\ \mathcal{E}}]{\text{error}} \sigma \in Q \; \boxed{\xrightarrow[\text{channel}]{\text{recovery}} \tilde{\rho} \in Q}$$

want to avoid thinking about this

Under what conditions is info retrievable?

× Consider a Kraus decomposition of the error channel

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger$$

× $\mathcal{I}$ should not be able to collapse superpositions of code words in R by measuring the environment.

For each k, we require

$$A_k \mathcal{I} : R \longrightarrow Q$$

to act as an isometry up to an overall multiplicative factor.

× Any two vectors $|\psi\rangle, |\varphi\rangle \in R$ must have

× $\| A_k \mathcal{I} |\psi\rangle \| = \| A_k \mathcal{I} |\varphi\rangle \| = f(k)$

× $\langle \varphi | \mathcal{I}^\dagger A_k^\dagger A_k \mathcal{I} | \psi \rangle = f(k) \langle \varphi | \psi \rangle$.

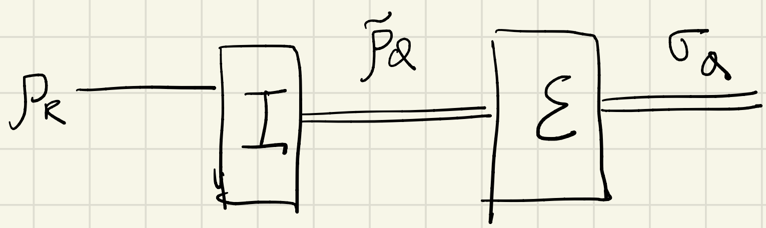
Think of $M_k^\dagger M_k$ as a POVM: the environment learns nothing about the code space.

Cf - purification in Romain's lectures: two initial states in the codespace are not distinguished by measurements

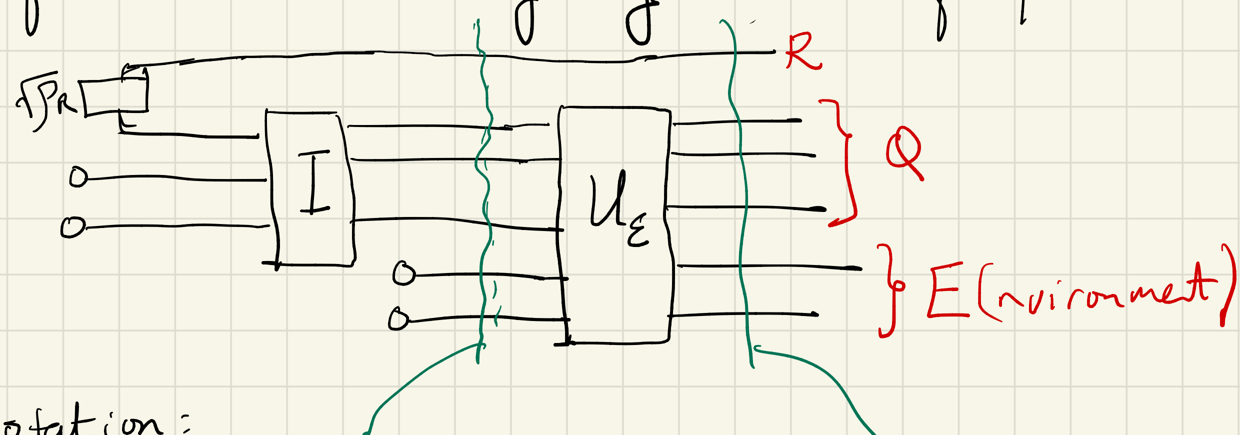
$\Updownarrow$

an initially mixed state from the codespace remains mixed after conditioning on measurements.

1.2/ Decoupling Principle.



Useful to rewrite everything in terms of pure states:



Notation:

state on this time slice is

$|\Psi\rangle_{RQ_0} \otimes |0\rangle_E$

state on this time slice

$|\Psi\rangle_{RQE}$

Claim (decoupling):

Exact error correction is possible iff $I(R:E)=0$.
(i.e., there's a recovery channel on Q that restores the original state on RQ.).

(Here, $I(A:B) = S(A)+S(B)-S(AB)$ is the "mutual information.")

Proof sketch:

1/ Suppose $I(R:E)>0$.

Note that
$$I(R:Q)+I(R:E)$$
$$= S(R)+S(Q)-S(E)+S(R)+S(E)-S(Q)$$
$$= 2S(R).$$

So $I(R:Q)$ has decreased from its pre-error value.

We can use:

Quantum data processing inequality:
$$I(A:B) \geq I(A:\mathcal{E}(B))$$
for any quantum channel acting on B.

to see that this loss of MI is unrecoverable.

2/ The more interesting direction.

Suppose $I(R:E)=0$.

Then $\rho_{RE} = \rho_R \otimes \rho_E$.

Write in product eigenbasis:

$$\rho_{RE} = \sum_{ij} r_i e_j \, |r_i e_j\rangle\langle r_i e_j| .$$

Can write the full state on RQE (undoing Schmidt decomp):

$$|\Psi_{RQE}\rangle = \sum_{ij} \sqrt{r_i e_j} \, |r_i e_j\rangle \otimes |q_{ij}\rangle .$$

Measure Q in the basis $\Pi_j = \sum_i |q_{ij}\rangle\langle q_{ij}|$.

(i.e., disentangle E by "measuring it out").

For outcome j you get

$$\Pi_j |\Psi_{RQE}\rangle = |e_j\rangle \otimes \sum_i \sqrt{r_i} \, |r_i\rangle \otimes |q_{ij}\rangle .$$

This is a purification of ρ_R, so it can be unitarily rotated back to the initial state. $\square$.

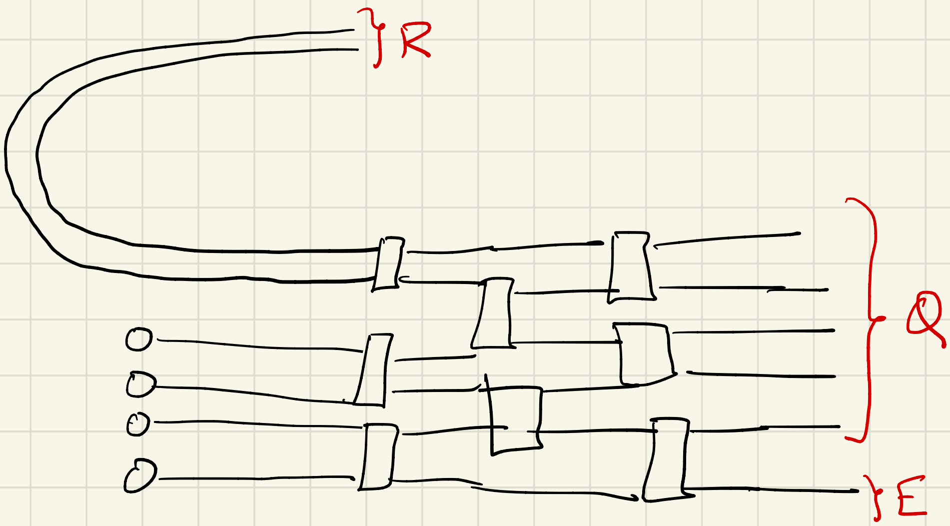
Approximate decoupling: (Schumacher - Westmoreland)

Suppose $I(R:E) \leq \varepsilon$. Then recovery is possible w/ fidelity $F \geqslant 1 - \sqrt{\varepsilon}$.

1.3/ Code distances and stat mech. (Li + Fisher)

Simple error model: some qubits from the system are swapped into the environment / "erased".

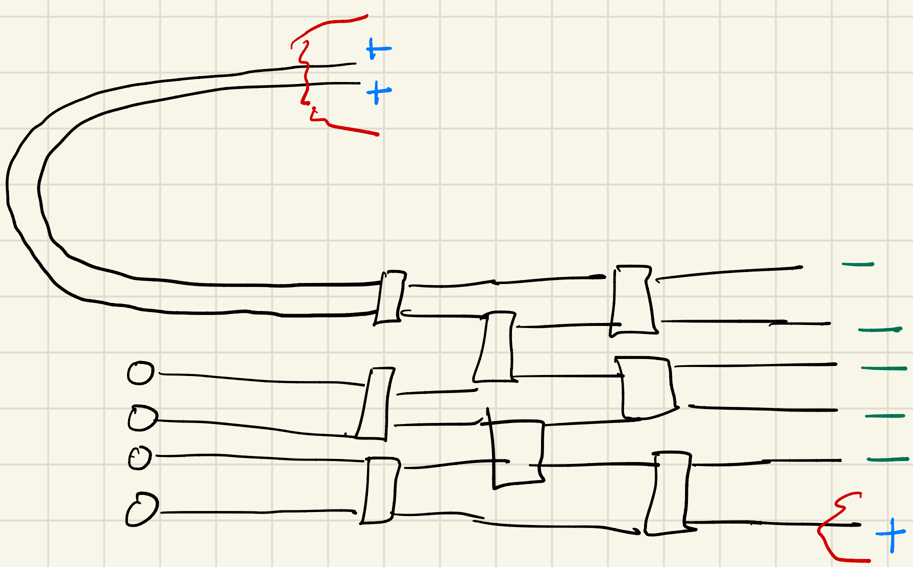
Can this be corrected?



Can compute $I(R:E) = S(R) + S(E) - S(RE)$

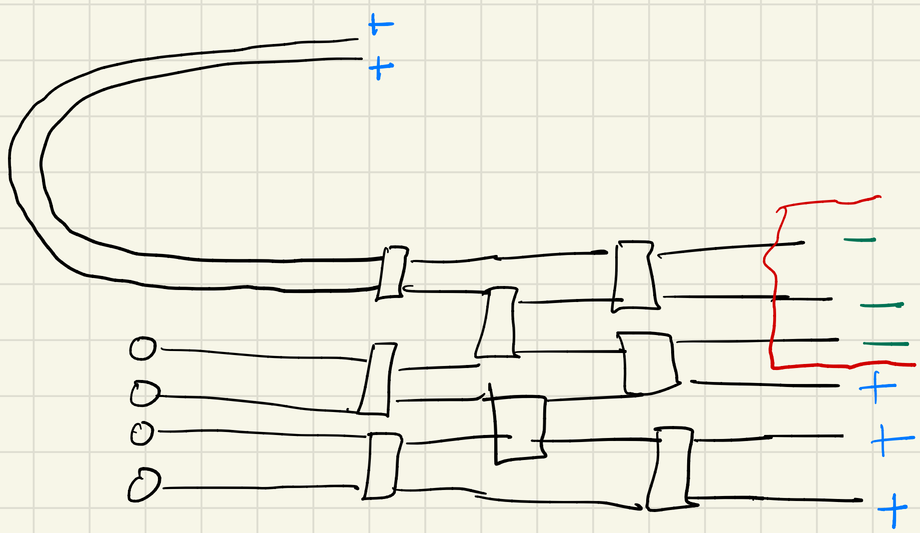
as domain wall energy costs.

e.g. $S(RE)$:

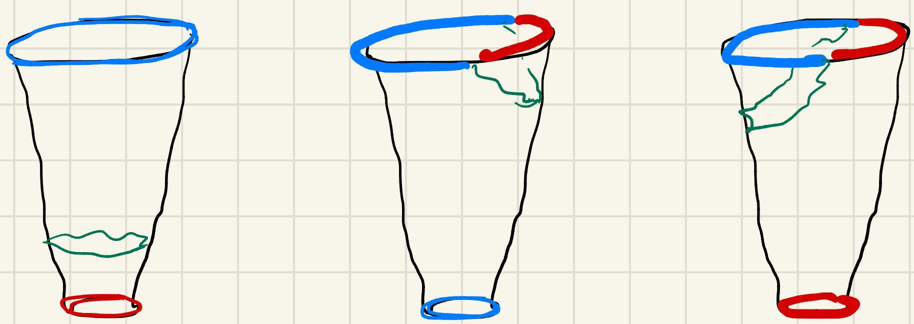


The DW for $S(RE)$ is just two disconnected pieces so $I(R:E) = 0$.

Now suppose you erase more qubits:



The code distance corresponds to a change in the topology of the domain walls. It is easier to unfold & visualize as a cylinder:



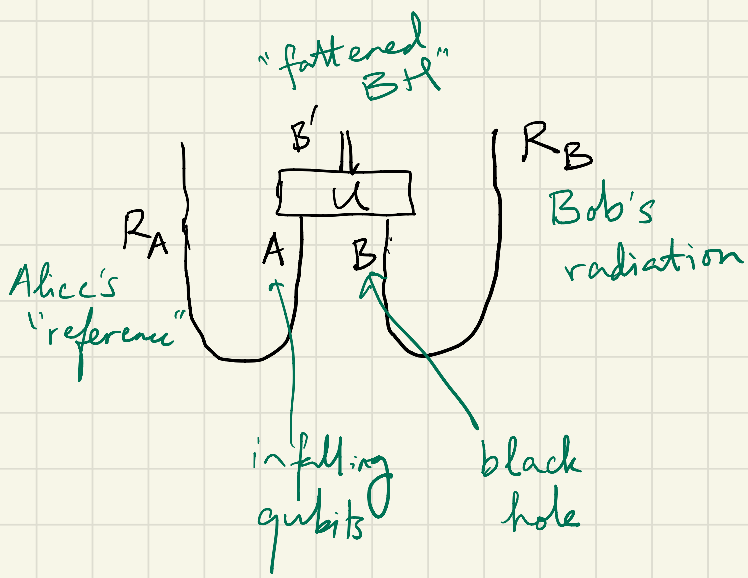
S(R) S(E) S(RE) = S(Q)

In the first two cases the domain wall is obvious. In the third case there are two possibilities, depending on the size of the erased region.

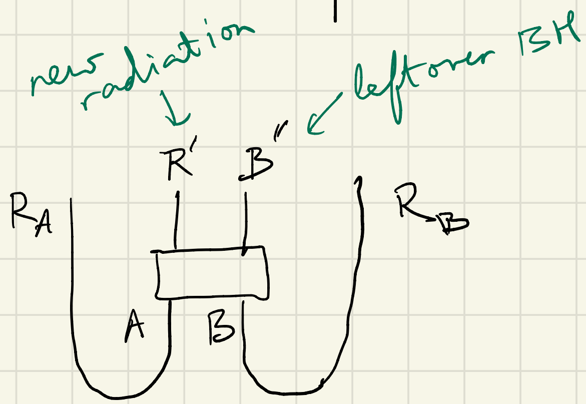
The (contiguous) code distance is the smallest size for which R and E "connect".

1.4/ Hayden-Preskill

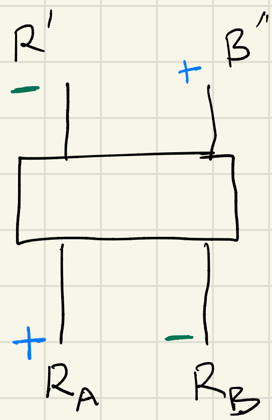
- Alice drops some quantum info into an old* black hole.
  ↳ more-than-half-evaporated
- Bob has been collecting all the radiation ever emitted. So Bob holds a purification of the black hole.
- Story up to this point:



- Fattened BH keeps radiating, Bob keeps collecting.



Q/ When does the info leak to Bob?
(or, when do R_A and B'' decouple?)



Decoupling happens when

$$|R'| \sim |A|.$$

Black hole almost immediately spits the info back out!

But how would you recover it?? .... Long story

Classical Hayden-Preskill

- Black hole is a long string of N−k bits, which Bob knows.
- Alice appends k bits (initially unknown).
- Dynamics is a random permutation on N bits (known to Bob) followed by s bits leaking out to Bob.
- How many bits does Bob need to decipher the message?
- Prob. of random strings matching in s places, $\sim 2^{-s}$.
- "Phase space" for a spurious match, 2^k.

$\longrightarrow$ error probability, 2^{k-s}, $\ll 1$ when $k \ll s$.

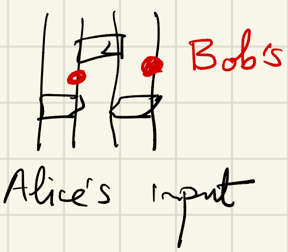
2/ Learning labels from measurements

From earlier: mixed phase $\longrightarrow$ measurement record is blind.

pure phase $\xrightarrow{??}$ meas. record is useful?

Idea: Alice prepares one of two states $\{|\psi_1\rangle, |\psi_2\rangle\}$.

The state evolves and is frequently measured by Bob.



Bob uses his measurement data to <u>infer</u> which state Alice sent in.

Bob checks his guess by calling Alice.

This gets around post-selection: it's using correlations in the measurement record!

A problem:

Consider an initial state drawn from $\{|+x\rangle, |-x\rangle\}$.
Bob measures the z component.
The wavefn collapses but nothing was learned!

What happened?

$$2^{-1/2}\left(|0\rangle \pm |1\rangle\right) \xrightarrow[\text{isometry}]{\text{measure-ment}} 2^{-1/2}\left(|00\rangle \pm |11\rangle\right)$$

measuring apparatus interacts w/ environment

$$2^{-1/2}\left(|000\rangle \pm |111\rangle\right)$$

trace environ-ment

$$\frac{1}{2}\left[|00\rangle\langle 00| + |11\rangle\langle 11|\right]$$

which-state info has been lost in high-order correlations...

× For a fixed trajectory / Kraus map $M_{\vec{k}}$, two ways for a mixed state $\rho = a|\psi\rangle\langle\psi| + b|\varphi\rangle\langle\varphi|$ to purify:

"coalescence"	"learning"
$M_{\vec{k}}|\psi\rangle \,/\!/\, M_{\vec{k}}|\varphi\rangle$	$\lVert M_{\vec{k}}|\psi\rangle\rVert / \lVert M_{\vec{k}}|\varphi\rangle\rVert \to 0$

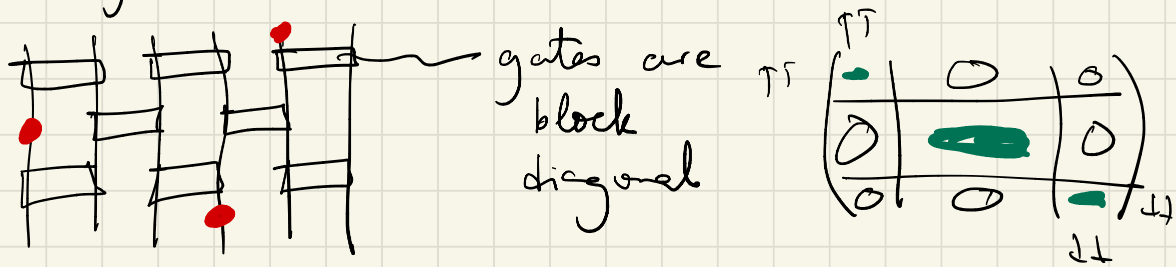
we would like a model w/ the latter effect dominant.

2.1 / Classical labels, the U(1) circuit

One way to achieve learning w/o coalescence: have a good quantum number distinguishing Alice's input states.

Easiest example: charge.

U(1) symmetric circuit:



measure the conserved U(1) charge, Q

initial state $\rho_0 = a|\psi_1\rangle\langle\psi_1| + b|\psi_2\rangle\langle\psi_2|$

$$\hat{Q}|\psi_1\rangle = q_1|\psi_1\rangle, \quad \hat{Q}|\psi_2\rangle = q_2|\psi_2\rangle.$$

Q/ How well do measurements "learn" charge?

2.2/ Three classifiers

a/ Obvious classifier

* Bob knows the circuit & outcomes, & can simulate it w/ the outcomes $\vec{k}$ forced.

Bob computes $\|M_{\vec{k}}|\psi_1\rangle\|$, $\|M_{\vec{k}}|\psi_2\rangle\|$ and guesses the one w/ larger weight.

This is information-theoretically optimal but computationally hard (even w/ QC).

We'll return to it in a bit...

b/ Very simple classifier.

Treat measurement outcomes as uncorrelated samples.

* Informal argument:

for charge Q_1 on L sites: $p(\text{empty}) = 1 - Q_1/L$
$p(\text{full}) = Q_1/L$.

likewise for $Q_2 \equiv Q_1 + 1$

to distinguish these possibilities we require a resolution $\sim |Q_1 - Q_2|/L$

$\Longrightarrow \left(\frac{L}{|Q_1 - Q_2|}\right)^2$ shots by CLT. (assuming $O(1)$ variance)

Need $\sim L^2$ shots to resolve charge, have $\sim pLt$ shots at time t

$\Longrightarrow$ charge is well-resolved at $t_\# \sim L/p$.

In fact it is resolved much sooner for large p...

* Less informal argument

$p(N \text{ empty out of } M \text{ shots})$ conditional on Q_1 $\xrightarrow{CLT} \frac{1}{\sqrt{2\pi c M}} \exp\left(-\left(N - \frac{Q_1}{L}M\right)^2 \Big/ (cM)\right)$.

$\parallel$

$p(N, M | Q_1)$

By Bayes' rule, $p(Q_1 | N, M) = \frac{p(N,M|Q_1)\, p(Q_1)}{p(N,M|Q_1) + p(N,M|Q_2)}$

Ratio $\frac{p(Q_1|N,M)}{p(Q_2|N,M)} = \frac{p(N,M|Q_1)}{p(N,M|Q_2)} \simeq \exp\left[-\frac{1}{cL^2}(MQ - NL)\right]$

This becomes small for typical samples (w/ $N \simeq MQ/L + O(\sqrt{M})$) when $L^2 \sim M$.

c/ Slightly better-informed classifier

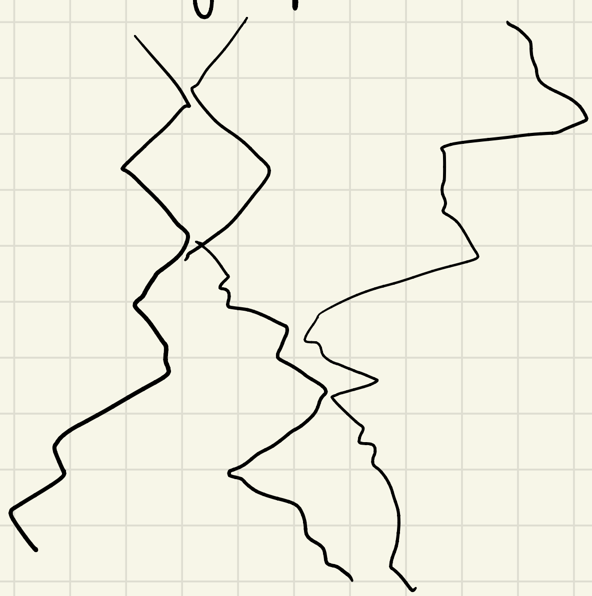
The classifier in (b) threw out
- × measurement location info
- × gate info.

Let's throw out the latter but keep the former.

This treats each gate as a maximally random process that conserves charge.

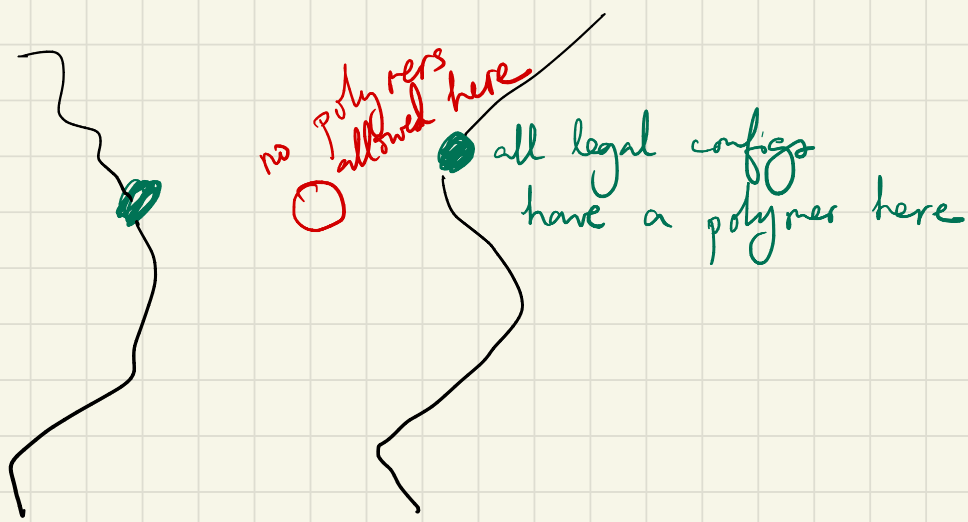
$\longrightarrow$ symmetric exclusion process.

Each history of the SEP looks like directed polymers:



Effects of measurement:

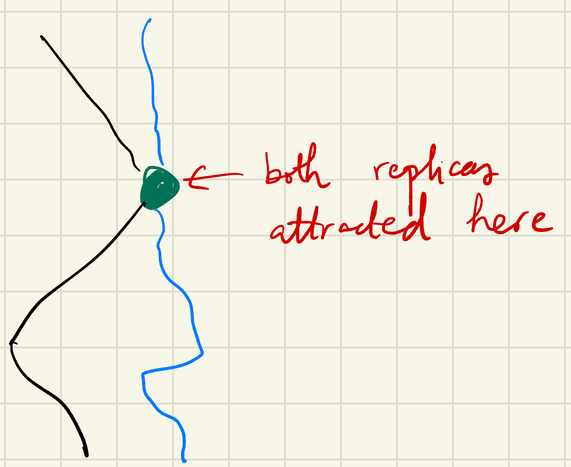
× fix $\vec{m}$. Then the prob. $p(\vec{m}|Q_1)$ is a partition fn. of SEP w/ quenched constraints at all measured sites:



× can again find $p(Q_1|\vec{m})$ by Bayes' rule.

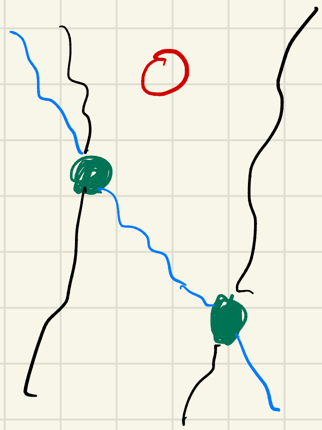
Replica theory (very schematic):

constraints induce interactions among replicas



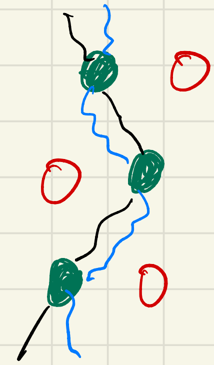
Q/ Do the interactions successfully bind polymers in different replicas?

A/ Two phases:



"unbound"

this looks like an "inter-replica phase slip".



"bound"

Bottom line: apparent phase-slip unbinding transition in BKT class.

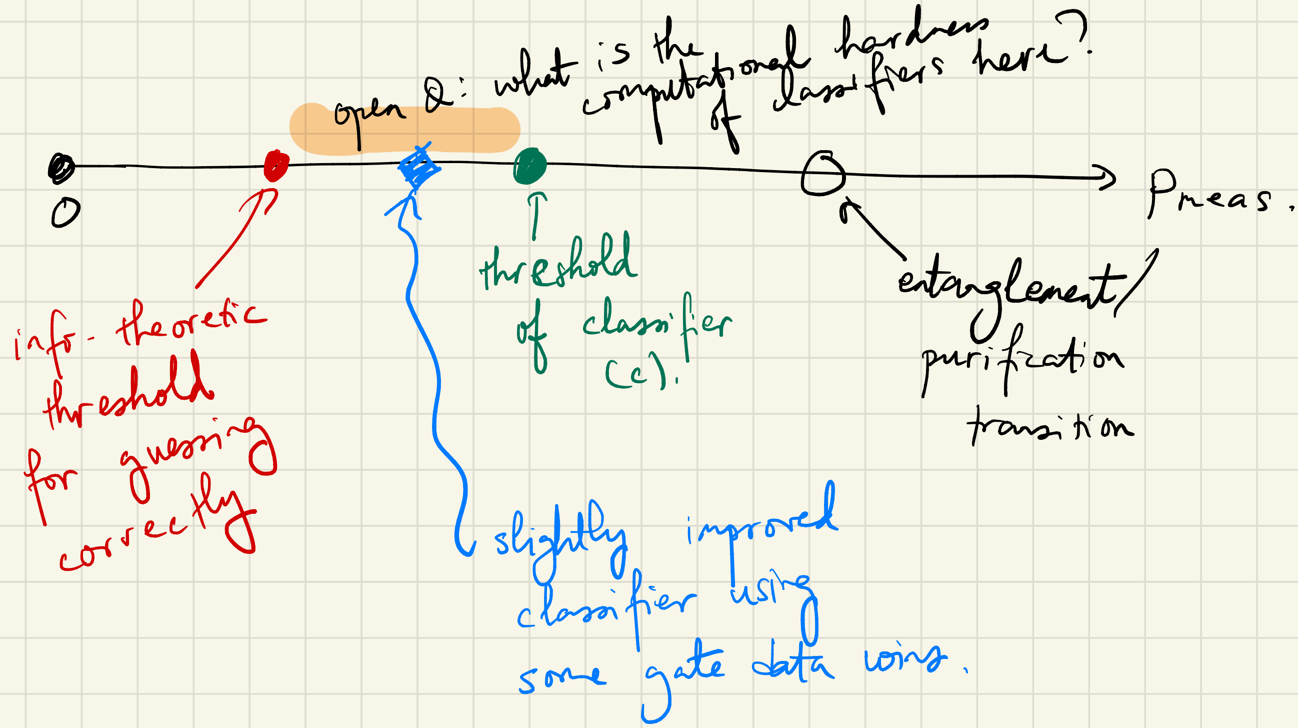
Meaning of the phases:

* unbound phase: $p(\vec{m}|Q_i)$ reasonably broad across Q_i, observations don't fix charge
* bound phase: observations fix charge w/o knowledge of gates.
* in bound phase, fluctuations are confined to finite scale ξ. For evolution times $\gg \xi \log L$, measurements fix charge.
* in unbound phase, for size L, correlation time $t \sim L$ (Luttinger liquid physics).

2.3/ Big picture

fix a time $\log L \ll t \ll L$.

take $L \to \infty$



Other extensions: learning in higher dimensions, w/ non-abelian symmetries etc.