Statistical mechanical models for stabiliser codes subject to correlated noise

Joint work with Steve Flammia (USyd/Yale)
Christopher Chubb

Centre for Engineered Quantum Systems
University of Sydney

TQC 2018

christopher.chubb@sydney.edu.au

christopherchubb.com

THE UNIVERSITY OF SYDNEY

EQUS

Statistical mechanical mapping

Stabiliser code & iid noise model	$\longrightarrow$	Disordered statistical mechanical model
Error-correcting threshold	$\longleftrightarrow$	Quenched phase transition
Maximum likelihood decoding	$\longleftrightarrow$	Calculating partition functions/free energy differences

Allows us to reappropriate techniques for studying stat. mech. systems to study quantum codes.

Statistical mechanical mapping

Stabiliser code & iid noise model	$\longrightarrow$	Disordered statistical mechanical model
Error-correcting threshold	$\longleftrightarrow$	Quenched phase transition
Maximum likelihood decoding	$\longleftrightarrow$	Calculating partition functions/free energy differences

Allows us to reappropriate techniques for studying stat. mech. systems to study quantum codes.

Statistical mechanical mapping

Stabiliser code & iid noise model	$\longrightarrow$	Disordered statistical mechanical model
Error-correcting threshold	$\longleftrightarrow$	Quenched phase transition
Maximum likelihood decoding	$\longleftrightarrow$	Calculating partition functions/free energy differences

Allows us to reappropriate techniques for studying stat. mech. systems to study quantum codes.

Our results

- Generalise mapping to arbitrary codes and correlated noise
- Show how to apply our mapping to circuit noise via the history code
- Numerically demonstrate that mild correlations can lower the threshold of the toric code considerably
- Show that the stat. mech. mapping gives tensor network ML decoders which generalise the MPS decoder of Bravyi, Suchara and Vargo

Our results

- Generalise mapping to arbitrary codes and correlated noise
- Show how to apply our mapping to circuit noise via the history code
- Numerically demonstrate that mild correlations can lower the threshold of the toric code considerably
- Show that the stat. mech. mapping gives tensor network ML decoders which generalise the MPS decoder of Bravyi, Suchara and Vargo

Our results

- Generalise mapping to arbitrary codes and correlated noise
- Show how to apply our mapping to circuit noise via the history code
- Numerically demonstrate that mild correlations can lower the threshold of the toric code considerably
- Show that the stat. mech. mapping gives tensor network ML decoders which generalise the MPS decoder of Bravyi, Suchara and Vargo

Our results

- Generalise mapping to arbitrary codes and correlated noise
- Show how to apply our mapping to circuit noise via the history code
- Numerically demonstrate that mild correlations can lower the threshold of the toric code considerably
- Show that the stat. mech. mapping gives tensor network ML decoders which generalise the MPS decoder of Bravyi, Suchara and Vargo

Independent case

Let $\llbracket A, B \rrbracket$ be the scalar commutator of two Paulis, $AB =: \llbracket A, B \rrbracket BA$.

For a stabiliser code generated by $\{S_k\}_k$, and an error Pauli E,

$$H_E(\vec{s}) := -\sum_i \sum_{\sigma \in \mathcal{P}_i} J_i(\sigma) \llbracket \sigma, E \rrbracket \prod_{k : \llbracket \sigma, S_k \rrbracket = -1} s_k$$

for some couplings $J_i(\sigma) \to \mathbb{R}$.

This has the symmetry that flipping a spin $s_j \to -s_j$ and applying a stabiliser $E \to ES_j$ cancel out.

Independent case

Let $\llbracket A, B \rrbracket$ be the scalar commutator of two Paulis, $AB =: \llbracket A, B \rrbracket BA$.

For a stabiliser code generated by $\{S_k\}_k$, and an error Pauli E,

$$H_E(\vec{s}) := -\sum_i \sum_{\sigma \in \mathcal{P}_i} J_i(\sigma) \llbracket \sigma, E \rrbracket \prod_{k : \llbracket \sigma, S_k \rrbracket = -1} s_k$$

for some couplings $J_i(\sigma) \to \mathbb{R}$.

This has the symmetry that flipping a spin $s_j \to -s_j$ and applying a stabiliser $E \to ES_j$ cancel out.

Independent case

Let $[\![A,B]\!]$ be the scalar commutator of two Paulis, $AB =: [\![A,B]\!]\, BA$.

For a stabiliser code generated by $\{S_k\}_k$, and an error Pauli E,

$$H_E(\vec{s}) := -\sum_i \sum_{\sigma\in\mathcal{P}_i} J_i(\sigma)\,[\![\sigma,E]\!] \prod_{k:[\![\sigma,S_k]\!]=-1} s_k$$

for some couplings $J_i(\sigma) \to \mathbb{R}$.

This has the symmetry that flipping a spin $s_j \to -s_j$ and applying a stabiliser $E \to ES_j$ cancel out.

Independent case

Suppose we have an independent error model

$$\Pr(E) = \prod_i p_i(E_i).$$

Nishimori condition: $$\beta J_i(\sigma) = \frac{1}{4}\sum_{\tau\in\mathcal{P}} \log p_i(\tau)\,[\![\sigma,\tau]\!]\,,$$

Using the Fourier-like orthogonality relation $\frac{1}{4}\sum_\sigma [\![\sigma,\tau]\!] = \delta_{\tau,I}$ we get that

$$-\beta H_E(\vec{1}) = \sum_i \sum_\tau \log p_i(\tau) \frac{1}{4} \sum_{\sigma,\tau} [\![\sigma, E\tau]\!] = \sum_i \log p_i(E_i) = \log \Pr(E).$$

The previous symmetry gives us that

$$Z_E = \sum_{\vec{s}} e^{-\beta H_E(\vec{s})} = \sum_S e^{-\beta H_{ES}(\vec{1})} = \sum_S \Pr(ES) = \Pr(\overline{E}).$$

Independent case

Suppose we have an independent error model

$$\Pr(E) = \prod_i p_i(E_i).$$

Nishimori condition: $$\beta J_i(\sigma) = \frac{1}{4} \sum_{\tau \in \mathcal{P}} \log p_i(\tau) \llbracket \sigma, \tau \rrbracket ,$$

Using the Fourier-like orthogonality relation $\frac{1}{4} \sum_\sigma \llbracket \sigma, \tau \rrbracket = \delta_{\tau, I}$ we get that

$$-\beta H_E(\vec{1}) = \sum_i \sum_\tau \log p_i(\tau) \frac{1}{4} \sum_{\sigma,\tau} \llbracket \sigma, E\tau \rrbracket = \sum_i \log p_i(E_i) = \log \Pr(E).$$

The previous symmetry gives us that

$$Z_E = \sum_{\vec{s}} e^{-\beta H_E(\vec{s})} = \sum_S e^{-\beta H_{ES}(\vec{1})} = \sum_S \Pr(ES) = \Pr(\overline{E}).$$

Error correction threshold as a quenched phase transition

Consider the free energy cost of a logical operator L,

$$\Delta_E(L) = -\frac{1}{\beta}\log Z_{EL} + \frac{1}{\beta} Z_E = \frac{1}{\beta}\log\frac{\Pr(E)}{\Pr(EL)}.$$

For a fixed error E

	Quantum code	Stat. mech. system
Below threshold	$\{\Pr(EL)\}_L$ peaked	$\Delta(L)\to\infty$ (in mean)
Above threshold	$\{\Pr(EL)\}_L$ flat	$\Delta(L)\to 0$ (in prob.)

Error correction threshold as a quenched phase transition

Consider the free energy cost of a logical operator L,

$$\Delta_E(L) = -\frac{1}{\beta}\log Z_{EL} + \frac{1}{\beta}Z_E = \frac{1}{\beta}\log\frac{\Pr(E)}{\Pr(EL)}.$$

For a fixed error E

	Quantum code	Stat. mech. system
Below threshold	$\{\Pr(EL)\}_L$ peaked	$\Delta(L)\to\infty$ (in mean)
Above threshold	$\{\Pr(EL)\}_L$ flat	$\Delta(L)\to 0$ (in prob.)

Error correction threshold as a quenched phase transition

Consider the free energy cost of a logical operator L,

$$\Delta_E(L) = -\frac{1}{\beta}\log Z_{EL} + \frac{1}{\beta} Z_E = \frac{1}{\beta}\log\frac{\Pr(E)}{\Pr(EL)}.$$

For a fixed error E

	Quantum code	Stat. mech. system
Below threshold	$\{\Pr(EL)\}_L$ peaked	$\Delta(L) \to \infty$ (in mean)
Above threshold	$\{\Pr(EL)\}_L$ flat	$\Delta(L) \to 0$ (in prob.)

Error correction threshold as a quenched phase transition

Consider the free energy cost of a logical operator L,

$$\Delta_E(L) = -\frac{1}{\beta}\log Z_{EL} + \frac{1}{\beta} Z_E = \frac{1}{\beta}\log\frac{\Pr(E)}{\Pr(EL)}.$$

For a fixed error E

	Quantum code	Stat. mech. system
Below threshold	$\{\Pr(EL)\}_L$ peaked	$\Delta(L) \to \infty$ (in mean)
Above threshold	$\{\Pr(EL)\}_L$ flat	$\Delta(L) \to 0$ (in prob.)

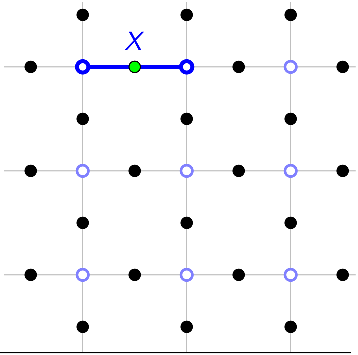
Independent examples

Toric code

Bit-flip → Random-bond Ising[1]

Indep. $X\&Z$ → 2×RBIM (uncoupled)

Depolarising → 2×RBIM (coupled)[2]



Color code

Depolarising → Eight-vertex model [2]

[1] Dennis et.al., JMP 2002, doi:10/cs2mtf, arXiv:quant-ph/0110143

[2] Bombin et.al., PRX 2012, doi:10/crz5, arXiv:1202.1852

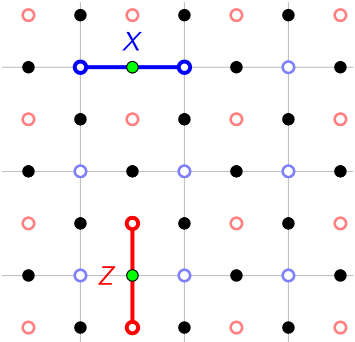
Independent examples

Toric code

Bit-flip → Random-bond Ising[1]

Indep. $X\&Z$ → 2×RBIM (uncoupled)

Depolarising → 2×RBIM (coupled)[2]



Color code

Depolarising → Eight-vertex model [2]

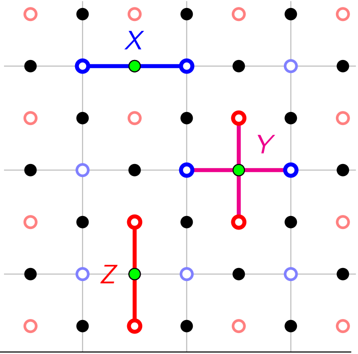
[1] Dennis et.al., JMP 2002, `doi:10/cs2mtf`, `arXiv:quant-ph/0110143`

[2] Bombin et.al., PRX 2012, `doi:10/crz5`, `arXiv:1202.1852`

Independent examples

Toric code

Bit-flip → Random-bond Ising[1]
Indep. $X\&Z$ → 2×RBIM (uncoupled)
Depolarising → 2×RBIM (coupled)[2]



Color code

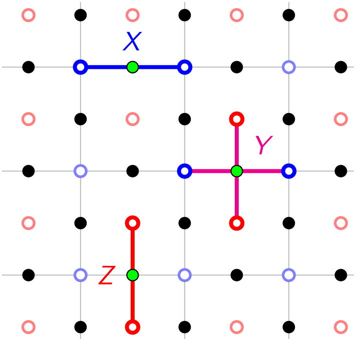
Depolarising → Eight-vertex model [2]

[1]Dennis et.al., JMP 2002, `doi:10/cs2mtf`, `arXiv:quant-ph/0110143`
[2]Bombin et.al., PRX 2012, `doi:10/crz5`, `arXiv:1202.1852`

Independent examples

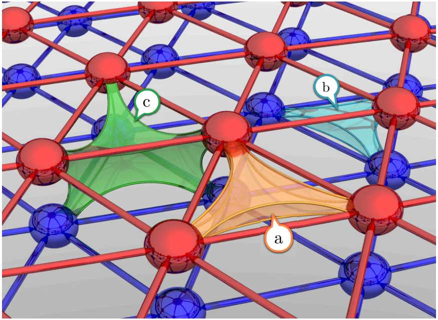
Toric code

Bit-flip → Random-bond Ising[1]
Indep. $X\&Z$ → 2×RBIM (uncoupled)
Depolarising → 2×RBIM (coupled)[2]



Color code

Depolarising → Eight-vertex model [2]



[1] Dennis et.al., JMP 2002, `doi:10/cs2mtf`, `arXiv:quant-ph/0110143`
[2] Bombin et.al., PRX 2012, `doi:10/crz5`, `arXiv:1202.1852`

Correlated case

The key point independence gave us was the ability to factor our noise model

$$\Pr(E) = \prod_i p_i(E_i).$$

We can generalise this to correlated models:

Factored distribution

An error model factors over regions $\{R_j\}_j$ if there exist $\phi_j : \mathcal{P}_{R_j} \to \mathbb{R}$ such that

$$\Pr(E) = \prod_j \phi_j\left(E_{R_j}\right)$$

This model includes many probabilistic graphical models, such as Bayesian Networks and Markov Random Fields.

Correlated case

The key point independence gave us was the ability to factor our noise model

$$\Pr(E) = \prod_i p_i(E_i).$$

We can generalise this to correlated models:

Factored distribution

An error model factors over regions $\{R_j\}_j$ if there exist $\phi_j : \mathcal{P}_{R_j} \to \mathbb{R}$ such that

$$\Pr(E) = \prod_j \phi_j \left(E_{R_j}\right)$$

This model includes many probabilistic graphical models, such as Bayesian Networks and Markov Random Fields.

Correlated case

The construction works pretty much the same as before, changing $\sigma \in \mathcal{P}_i$ to $\sigma \in \mathcal{P}_{R_j}$:

$$H_E(\vec{s}) := -\sum_j \sum_{\sigma \in \mathcal{P}_j} J_j(\sigma) \left[\!\left[\sigma, E\right]\!\right] \prod_{k : [\![\sigma, S_k]\!] = -1} s_k$$

$$\text{Nishimori condition:} \qquad \beta J_j(\sigma) = \frac{1}{|\mathcal{P}_{R_j}|} \sum_{\tau \in \mathcal{P}_{R_j}} \log \phi_j(\tau) \left[\!\left[\sigma, \tau\right]\!\right],$$

As before we get that $Z_E = \Pr(\overline{E})$, and so the threshold manifests as a phase transition.

Correlated case

The construction works pretty much the same as before, changing $\sigma \in \mathcal{P}_i$ to $\sigma \in \mathcal{P}_{R_j}$:

$$H_E(\vec{s}) := -\sum_j \sum_{\sigma \in \mathcal{P}_j} J_j(\sigma) \llbracket \sigma, E \rrbracket \prod_{k : \llbracket \sigma, S_k \rrbracket = -1} s_k$$

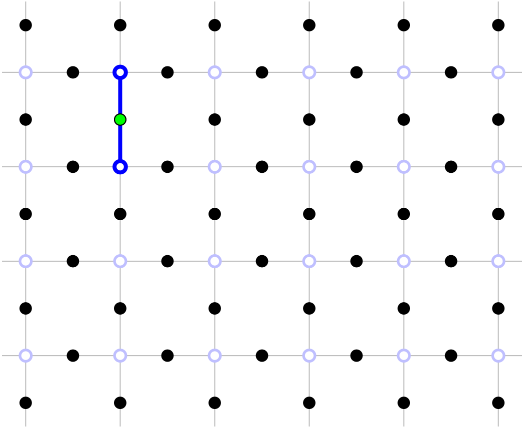
$$\text{Nishimori condition:} \qquad \beta J_j(\sigma) = \frac{1}{|\mathcal{P}_{R_j}|} \sum_{\tau \in \mathcal{P}_{R_j}} \log \phi_j(\tau) \llbracket \sigma, \tau \rrbracket ,$$

As before we get that $Z_E = \Pr(\overline{E})$, and so the threshold manifests as a phase transition.

Correlated example

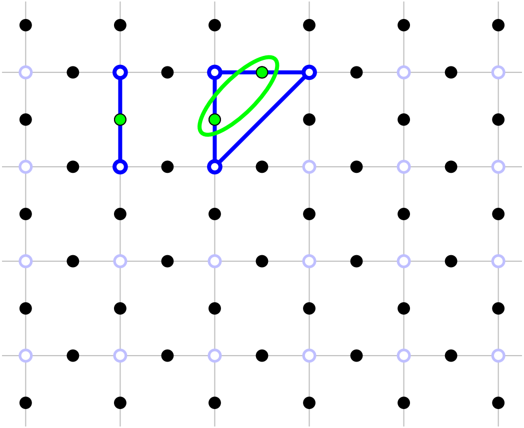
Toric code with correlated bit-flips

Correlations induce longer-range interactions



Toric code with correlated bit-flips

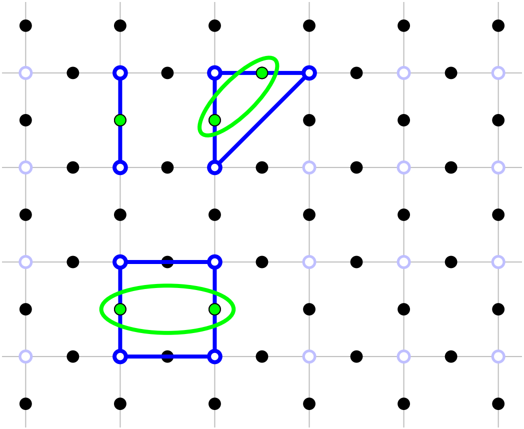
Correlations induce longer-range interactions



Correlated example

Toric code with correlated bit-flips

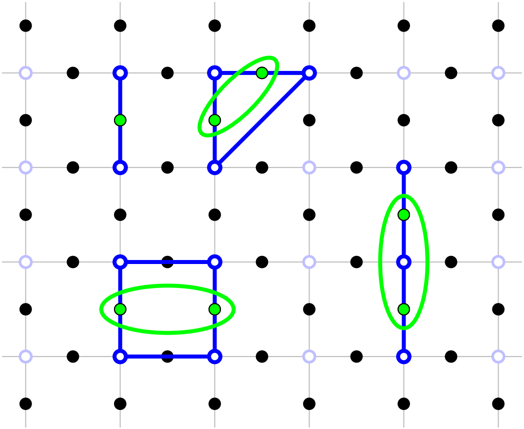
Correlations induce longer-range interactions



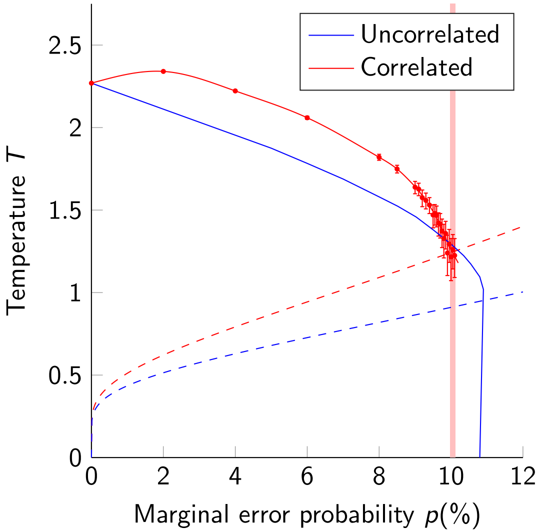
Correlated example

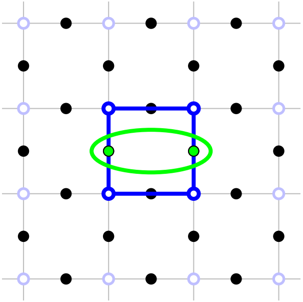
Toric code with correlated bit-flips

Correlations induce longer-range interactions



Monte Carlo simulations





Thresholds

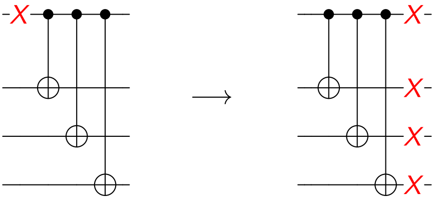
Indep.[1]: $p_{\mathrm{th}} = 10.94(2)\%$
Corr.: $p_{\mathrm{th}} = 10.04(6)\%$

[1] Dennis et.al., JMP 2002, `doi:10/cs2mtf`, `arXiv:quant-ph/0110143`

Circuit noise

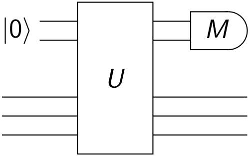
Noise followed by ideal measurements is unrealistic. In reality, circuits will be faulty.

Applying measurement circuits will tend to spread around and correlate noise:



Circuit noise

We will consider measurement circuits of the form

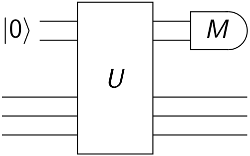


where U is a Clifford and M is a Pauli.

For convenience we only consider independent noise on each circuit. We also will push noise through until after the unitary:

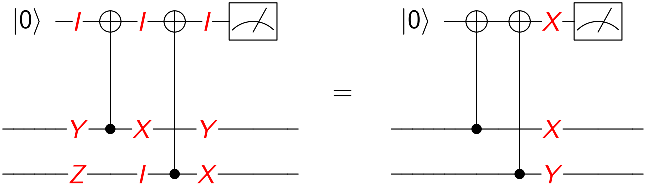
Circuit noise

We will consider measurement circuits of the form

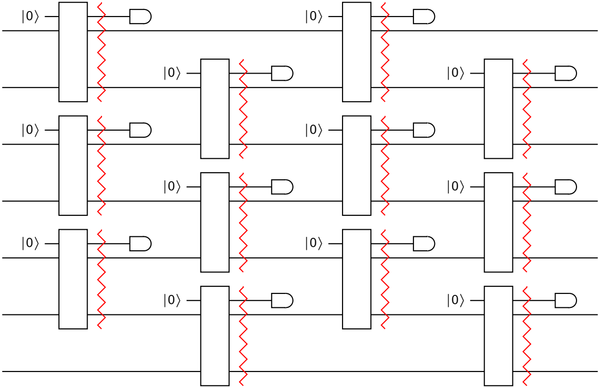


where U is a Clifford and M is a Pauli.

For convenience we only consider independent noise on each circuit. We also will push noise through until after the unitary:

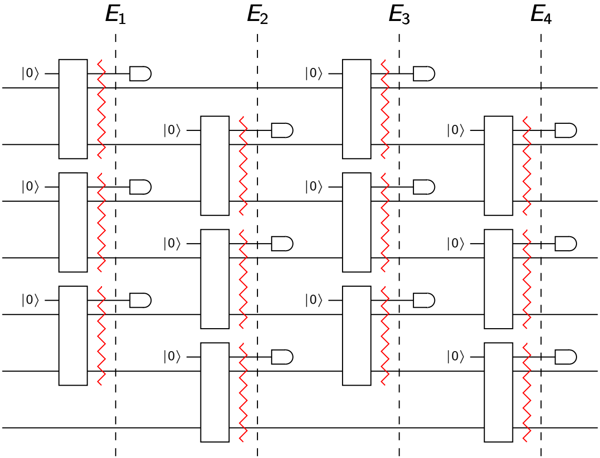


Factorising circuit noise



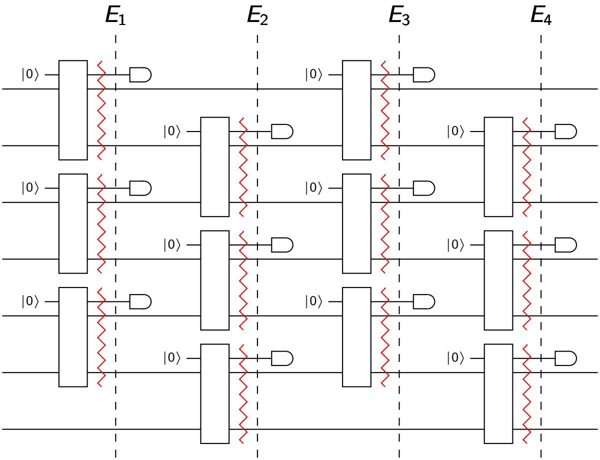
The distribution of the error history $(E_1, \ldots, E_T)$ factors.

Factorising circuit noise



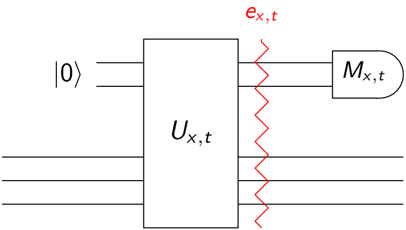
The distribution of the error history $(E_1, \ldots, E_T)$ factors.

Factorising circuit noise



The distribution of the error history $(E_1, \ldots, E_T)$ factors.

Factorising circuit noise



$(E_t)_t$ and $(e_{x,t})_{x,t}$ are 1-to-1, so

$$\Pr\Big((E_t)_t\Big) = \Pr\Big((e_{x,t})_{x,t}\Big).$$

If $e_{x,t} \sim p_{x,t}$ independently, then

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}(e_{x,t}).$$

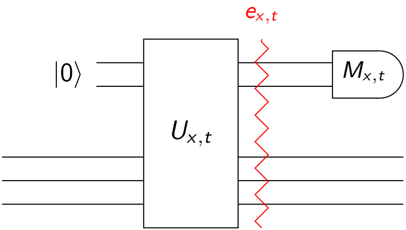
In terms of the accumulated errors:

$$E_t|_{R^+_{x,t}} = e_{x,t}\left(U_{x,t}\, E_{t-1}|_{R^-_{x,t}}\, U^\dagger_{x,t}\right)$$

This implies the distribution of error histories factor over $\{R^+_{x,t} \cup R^-_{x,t}\}_{x,t}$:

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}\left(E_t|_{R^+_{x,t}}\left(U_{x,t}\, E^\dagger_{t-1}\Big|_{R^-_{x,t}}\, U^\dagger_{x,t}\right)\right)$$

Factorising circuit noise



$(E_t)_t$ and $(e_{x,t})_{x,t}$ are 1-to-1, so

$$\Pr\Big((E_t)_t\Big) = \Pr\Big((e_{x,t})_{x,t}\Big).$$

If $e_{x,t} \sim p_{x,t}$ independently, then

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}(e_{x,t}).$$

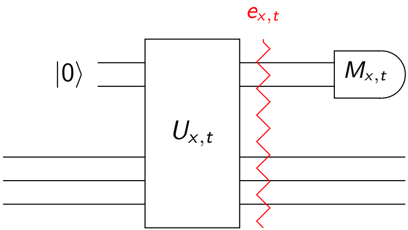
In terms of the accumulated errors:

$$E_t|_{R^+_{x,t}} = e_{x,t}\left(U_{x,t}\, E_{t-1}|_{R^-_{x,t}} U^\dagger_{x,t}\right)$$

This implies the distribution of error histories factor over $\{R^+_{x,t} \cup R^-_{x,t}\}_{x,t}$:

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}\left(E_t|_{R^+_{x,t}}\left(U_{x,t}\, E^\dagger_{t-1}\Big|_{R^-_{x,t}} U^\dagger_{x,t}\right)\right)$$

Factorising circuit noise



$(E_t)_t$ and $(e_{x,t})_{x,t}$ are 1-to-1, so

$$\Pr\Big((E_t)_t\Big) = \Pr\Big((e_{x,t})_{x,t}\Big).$$

If $e_{x,t} \sim p_{x,t}$ independently, then

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}(e_{x,t}).$$

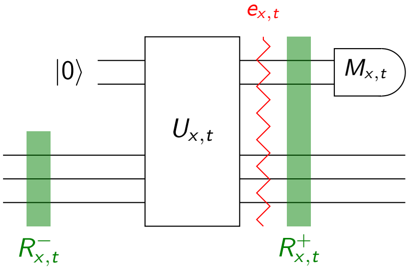
In terms of the accumulated errors:

$$E_t|_{R^+_{x,t}} = e_{x,t}\left(U_{x,t}\, E_{t-1}|_{R^-_{x,t}}\, U^\dagger_{x,t}\right)$$

This implies the distribution of error histories factor over $\{R^+_{x,t} \cup R^-_{x,t}\}_{x,t}$:

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}\left(E_t|_{R^+_{x,t}}\left(U_{x,t}\, E^\dagger_{t-1}\Big|_{R^-_{x,t}} U^\dagger_{x,t}\right)\right)$$

Factorising circuit noise



$(E_t)_t$ and $(e_{x,t})_{x,t}$ are 1-to-1, so

$$\Pr\Big((E_t)_t\Big) = \Pr\Big((e_{x,t})_{x,t}\Big).$$

If $e_{x,t} \sim p_{x,t}$ independently, then

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}(e_{x,t}).$$

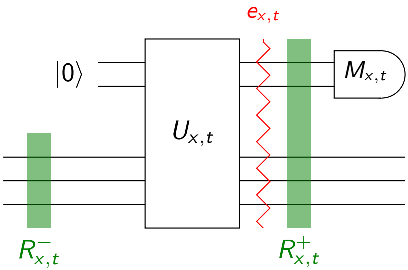
In terms of the accumulated errors:

$$E_t|_{R^+_{x,t}} = e_{x,t}\left(U_{x,t}\, E_{t-1}|_{R^-_{x,t}} U^\dagger_{x,t}\right)$$

This implies the distribution of error histories factor over $\{R^+_{x,t} \cup R^-_{x,t}\}_{x,t}$:

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}\left(E_t|_{R^+_{x,t}}\left(U_{x,t}\, E^\dagger_{t-1}\Big|_{R^-_{x,t}} U^\dagger_{x,t}\right)\right)$$

Factorising circuit noise



$(E_t)_t$ and $(e_{x,t})_{x,t}$ are 1-to-1, so

$$\Pr\Big((E_t)_t\Big) = \Pr\Big((e_{x,t})_{x,t}\Big).$$

If $e_{x,t} \sim p_{x,t}$ independently, then

$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}(e_{x,t}).$$

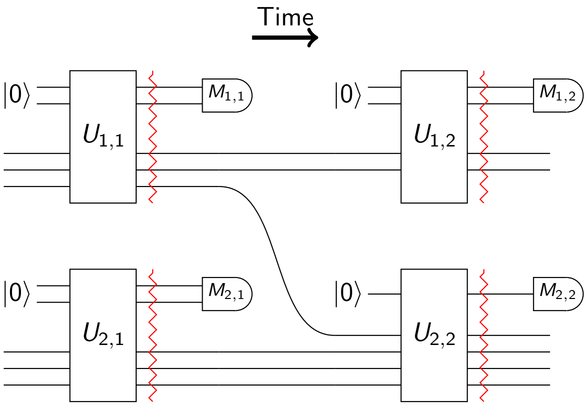
In terms of the accumulated errors:

$$E_t|_{R^+_{x,t}} = e_{x,t}\left(U_{x,t}\, E_{t-1}|_{R^-_{x,t}} U^\dagger_{x,t}\right)$$

This implies the distribution of error histories factor over $\{R^+_{x,t} \cup R^-_{x,t}\}_{x,t}$:

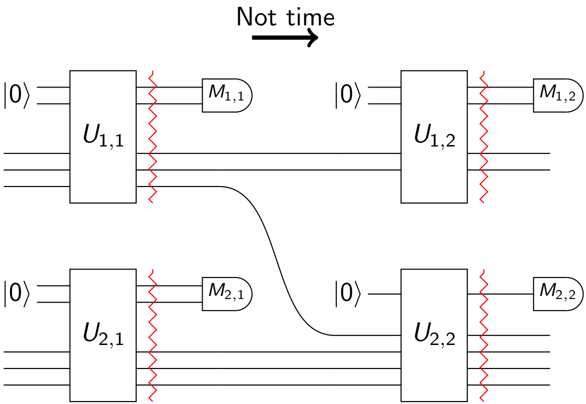
$$\Pr\Big((E_t)_t\Big) = \prod_{x,t} p_{x,t}\left(E_t|_{R^+_{x,t}}\left(U_{x,t}\, E^\dagger_{t-1}\Big|_{R^-_{x,t}} U^\dagger_{x,t}\right)\right)$$

History code



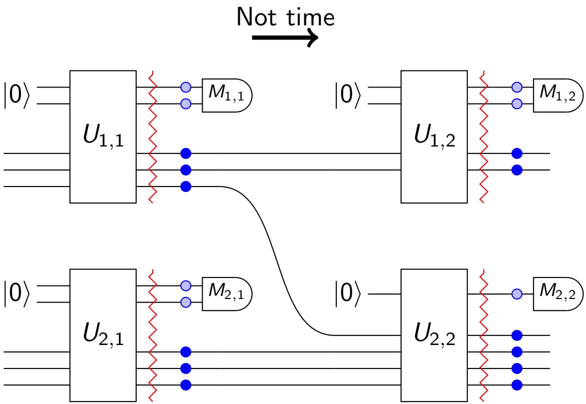
- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

History code



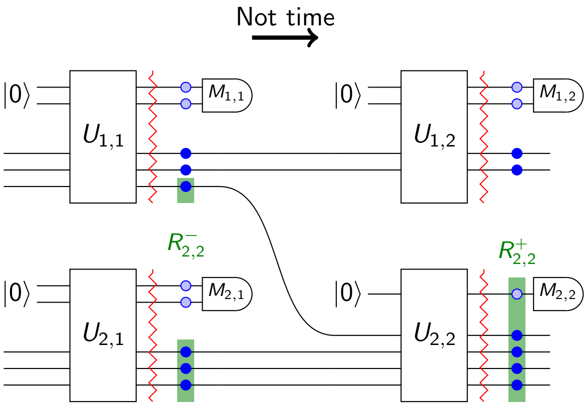
- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

History code



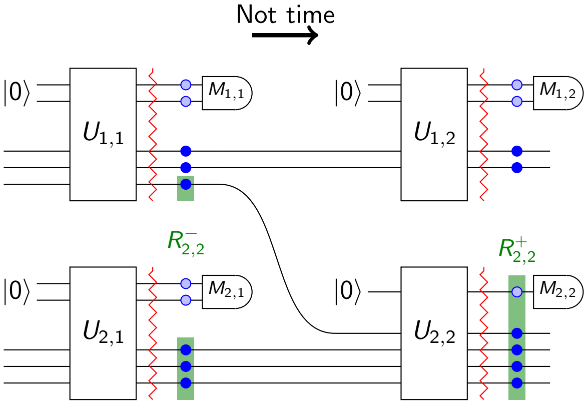
- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

History code



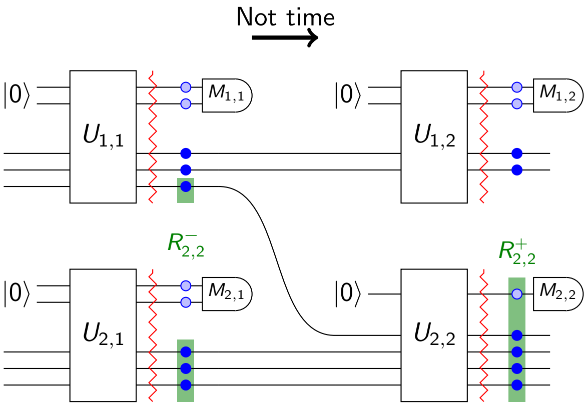
- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

History code



- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

History code



- The error model is given by error history, which factor.
- Stabilisers are the measurements $(M_{x,t})_{x,t}$, and stabilisers are final time-slice.
- Logicals are logicals at final time-slice.

Conclusions and further work

- Extended the stat. mech. mapping to correlated models
- Can apply stat. mech. mapping to circuit noise via the history code
- Stat. mech. mapping gives tensor network ML decoders

- Can we apply this to experimentally relevant correlated models?
- Can we use the decoders to understand to better understand the connection between correlation and the threshold (ongoing work).