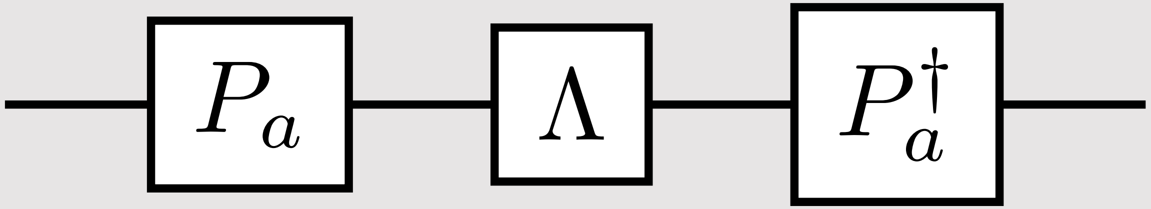
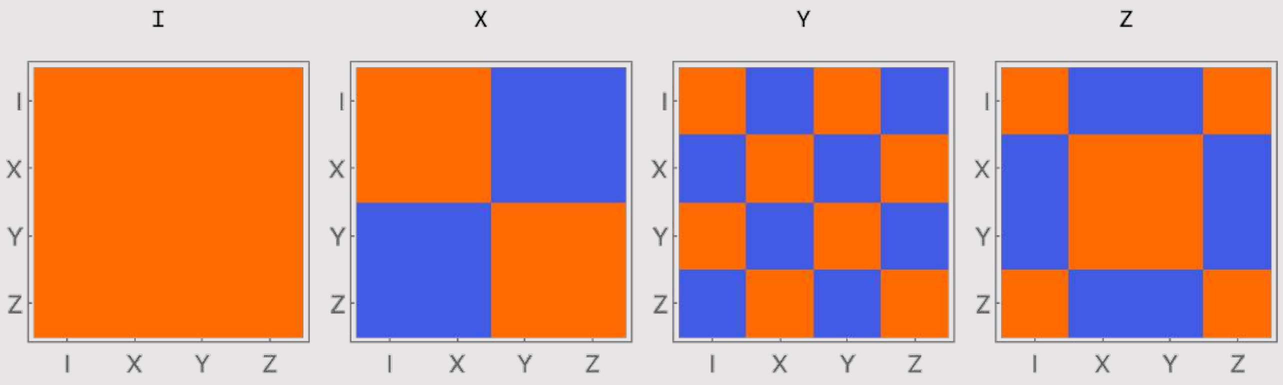
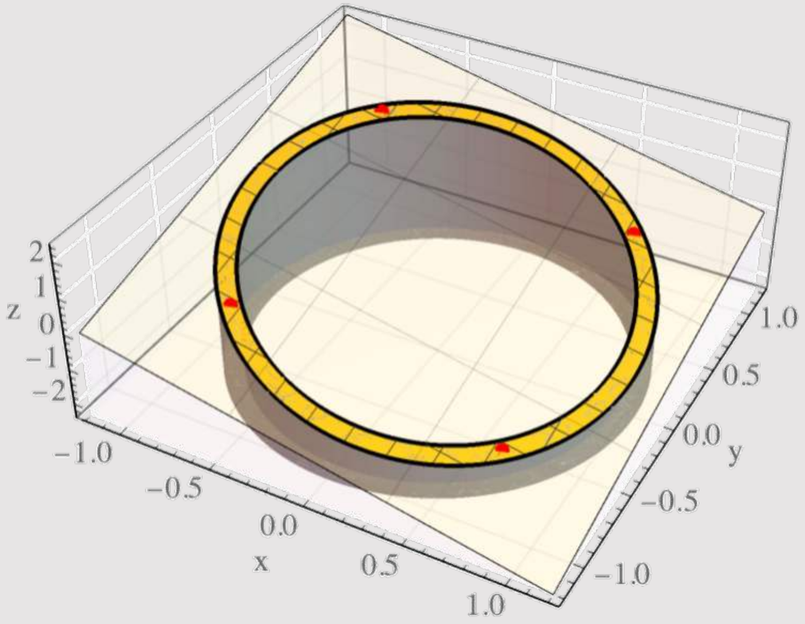
Primer on Pauli Twirling







Zlatko Minev

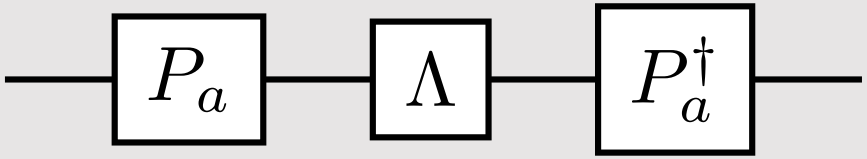
2022-04-20, 07-11

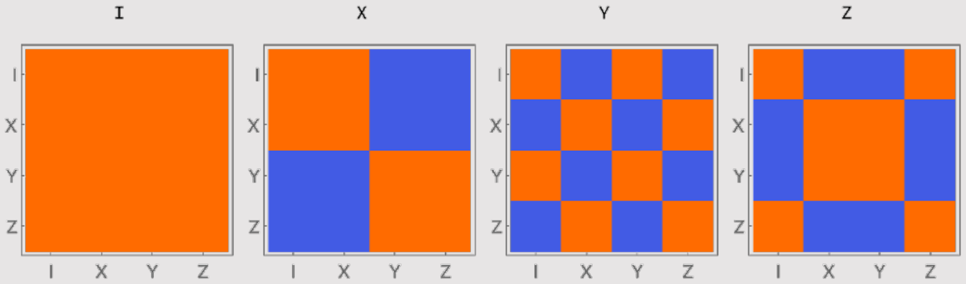
Twirling 101: Overview

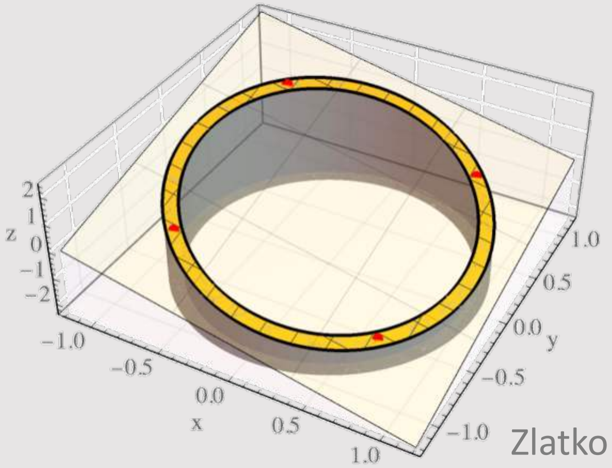
Twirl operationally
- Simple example
- General application
- Summary

Theory of twirling
- Why does twirling work?
- Masking channels

Optional: Advanced
- Why is the Pauli group special for twirling?
- Other twirl groups
- Designs



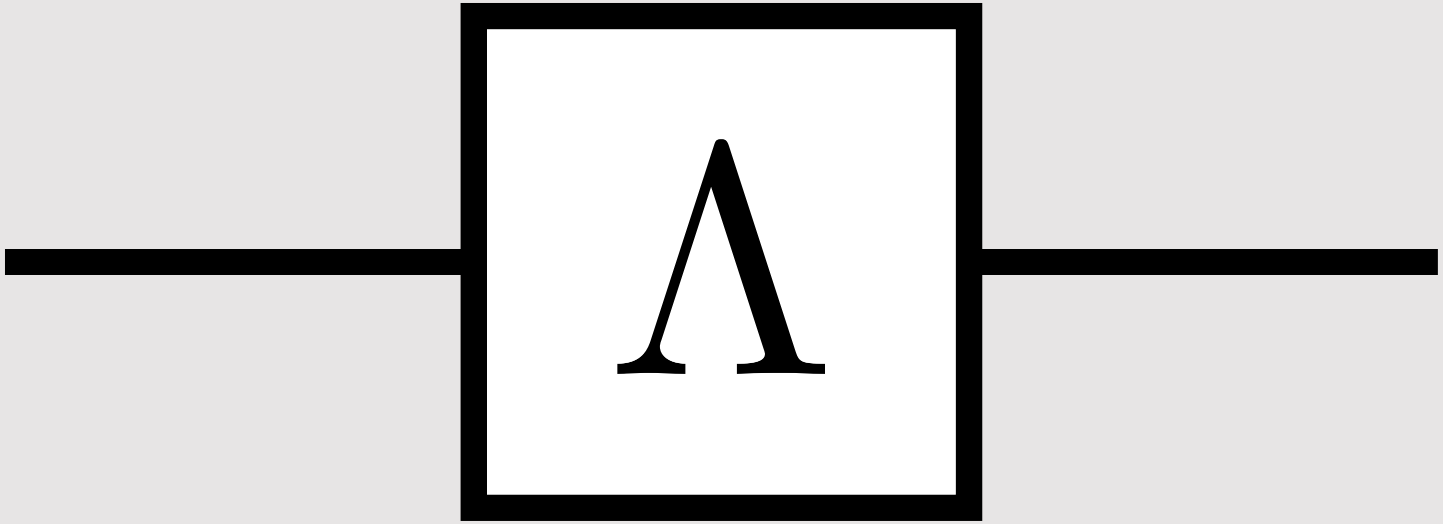




First, let's look at twirling operationally

Quantum channel

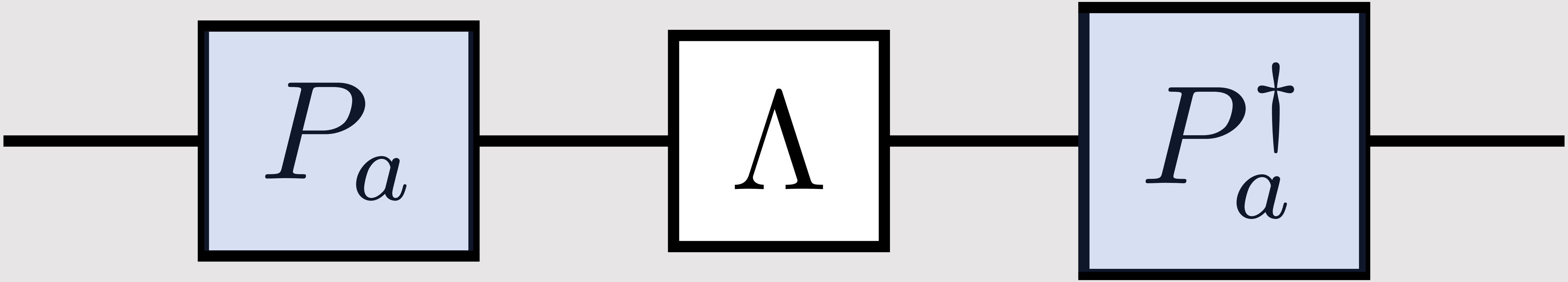
Imagine a quantum channel



Can be any CPTP map
T1, T2, coherent noise,
cross-talk, etc.

Twirl the channel

Imagine a quantum channel sandwiched in the following way

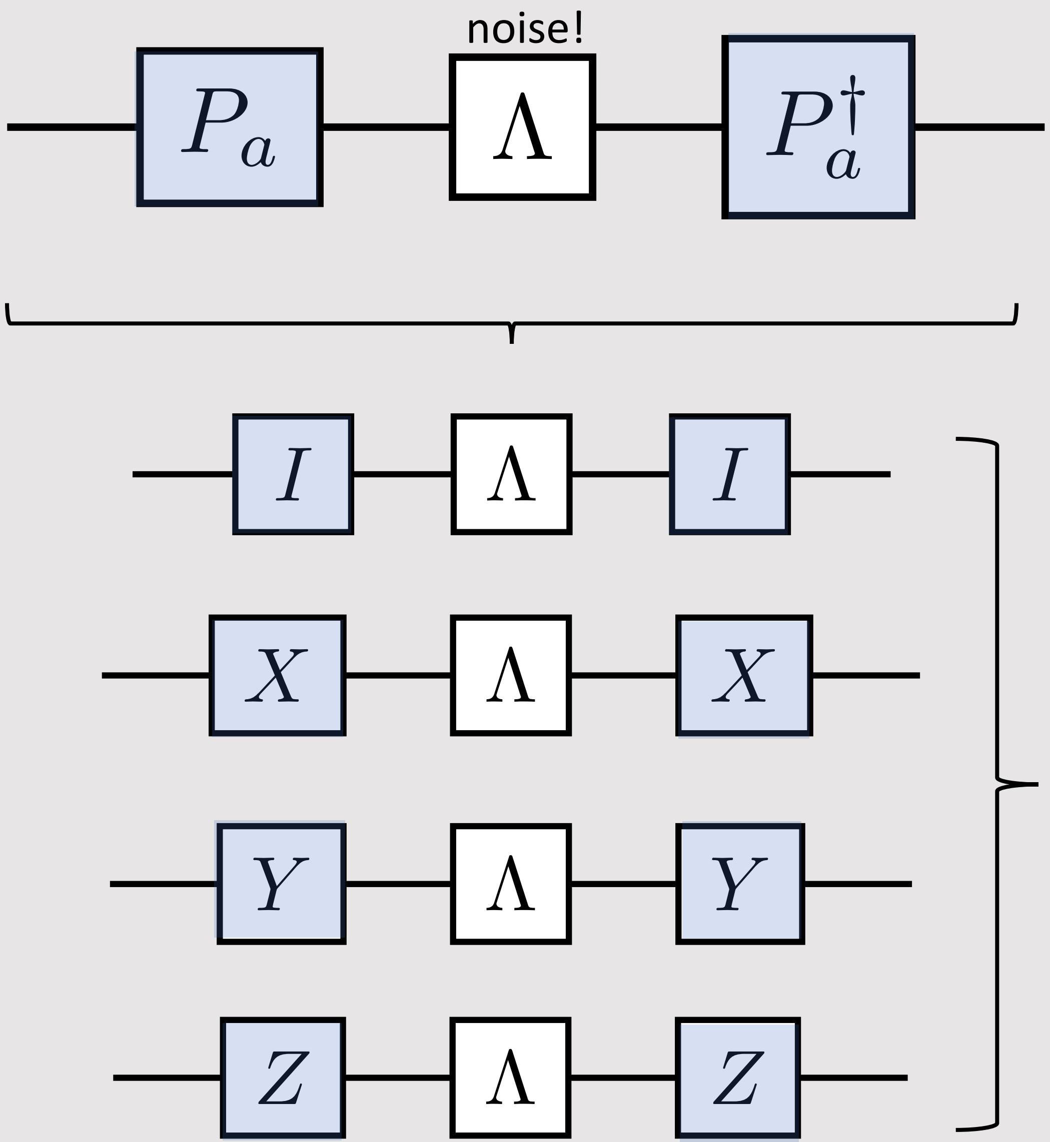


Randomly chosen
unitary gate P_a
from some set

C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Physical Review A 54, 3824 (1996)
C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Physical review letters 76, 722 (1996)
E. Knill, D. Leibfried, R. Reichle, J. Britton, R. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Physical Review A 77, 012307 (2008).
E. Magesan, J. M. Gambetta, and J. Emerson, Physical review letters 106, 180504 (2011)
…

Twirl the noise channel (not gate yet) over the Paulis

What does the twirl construction look like?



Average over some set of gate $\{P_a\}$ that form a group or a design (to be defined later)

Example: qubit channel

twirl over Pauli group

equal probability

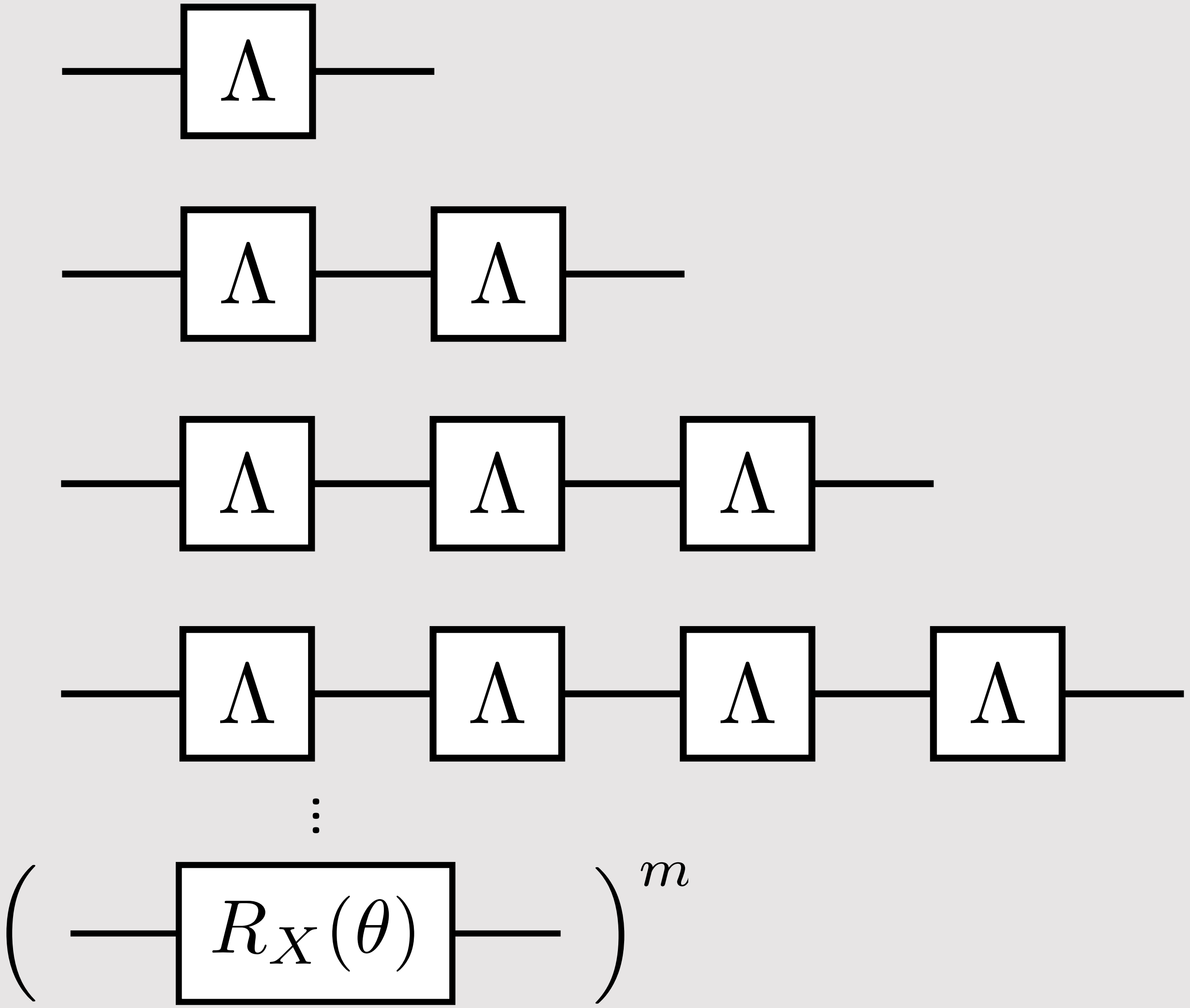
average over instances

Example

Coherent over rotation

$$-\boxed{\Lambda}- \quad = \quad -\boxed{R_X(\theta)}-$$

Amplify error channel to understand situation better



m=1

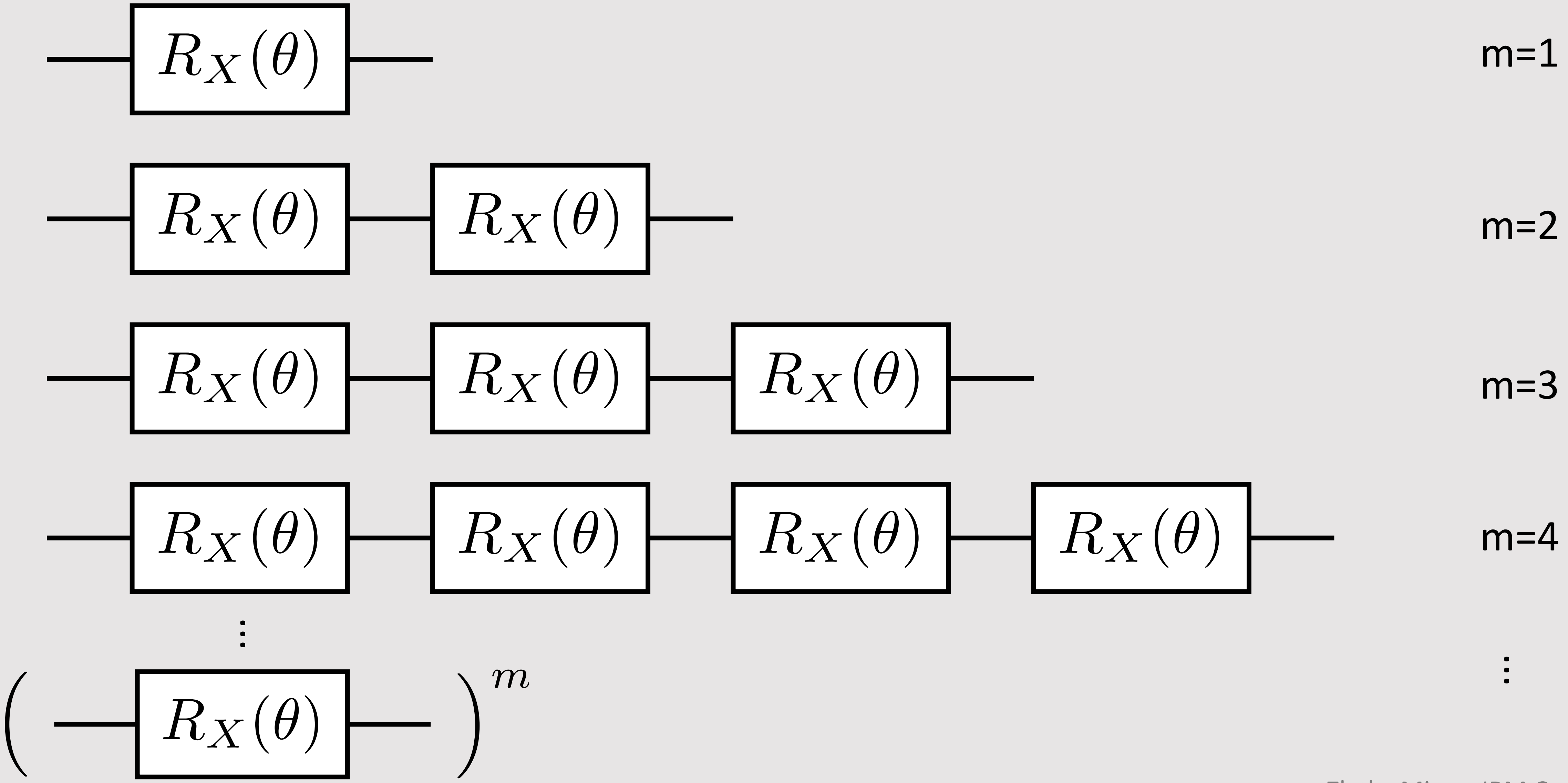
m=2

m=3

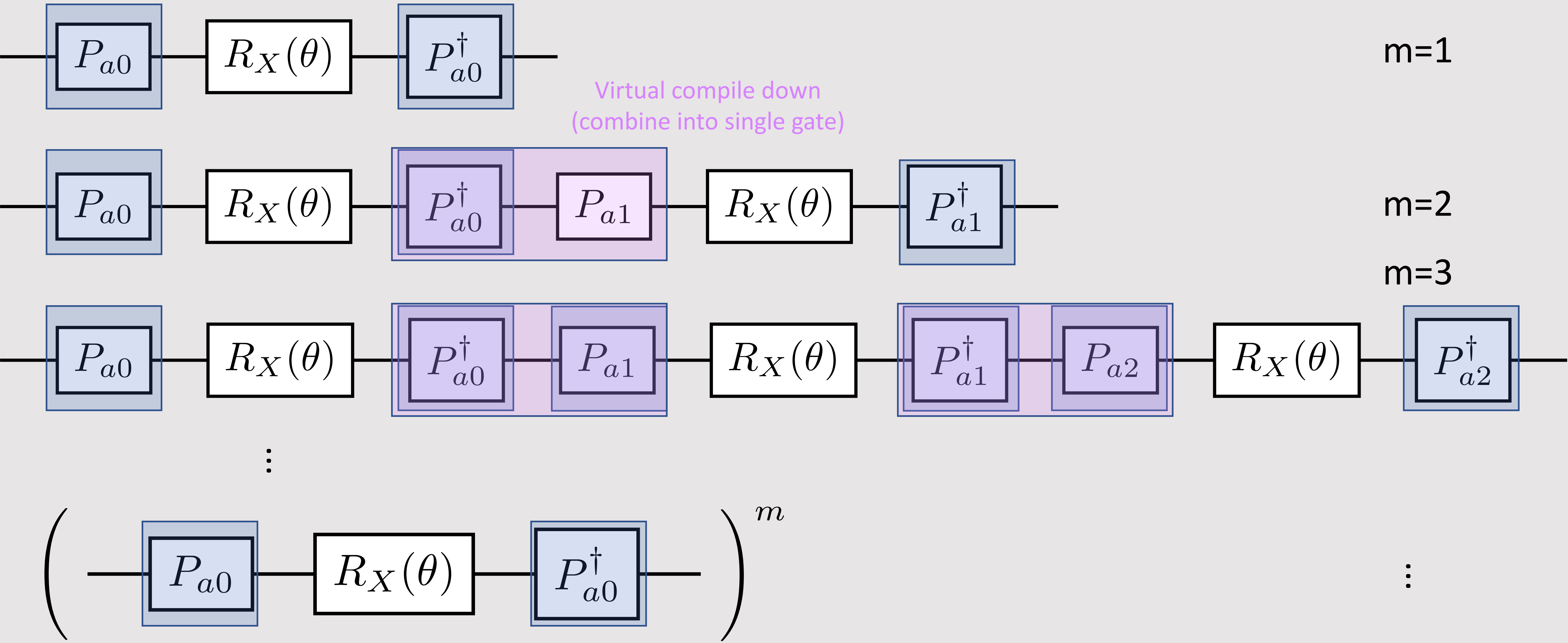
m=4

⋮

Amplify error channel to understand situation better



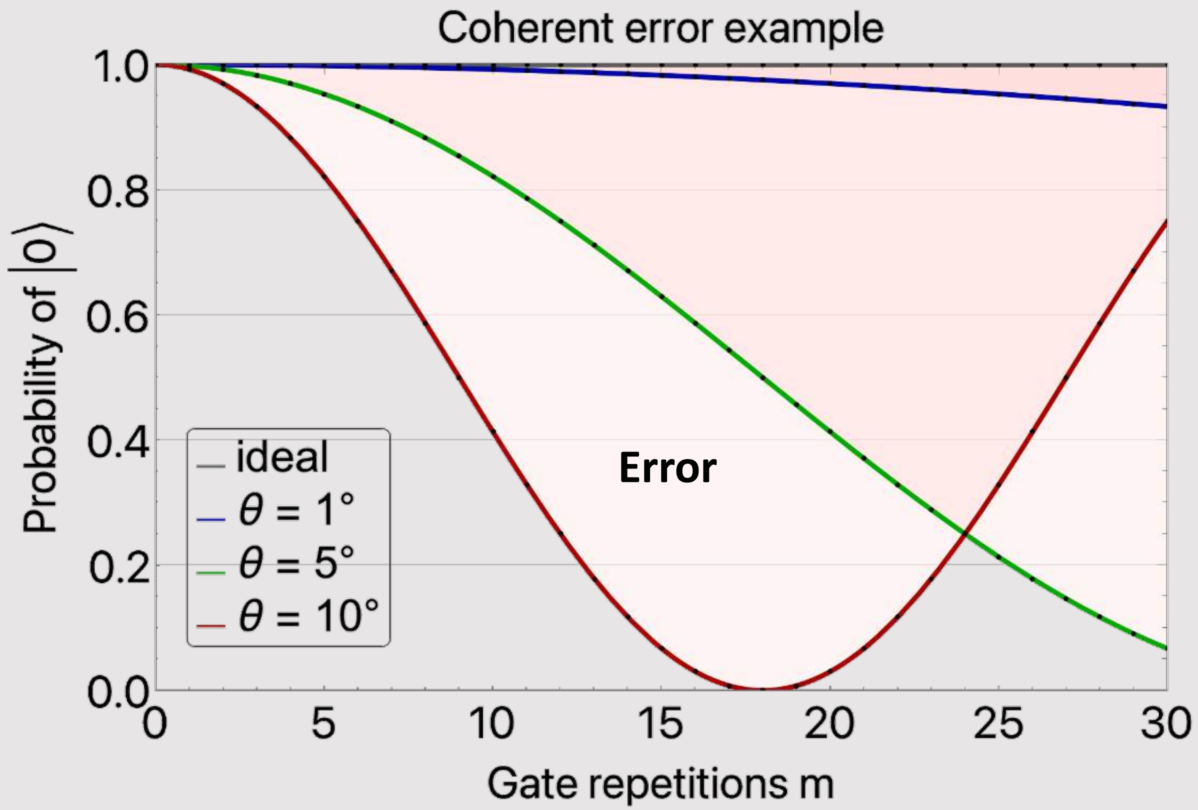
Amplify twirled error to understand situation better

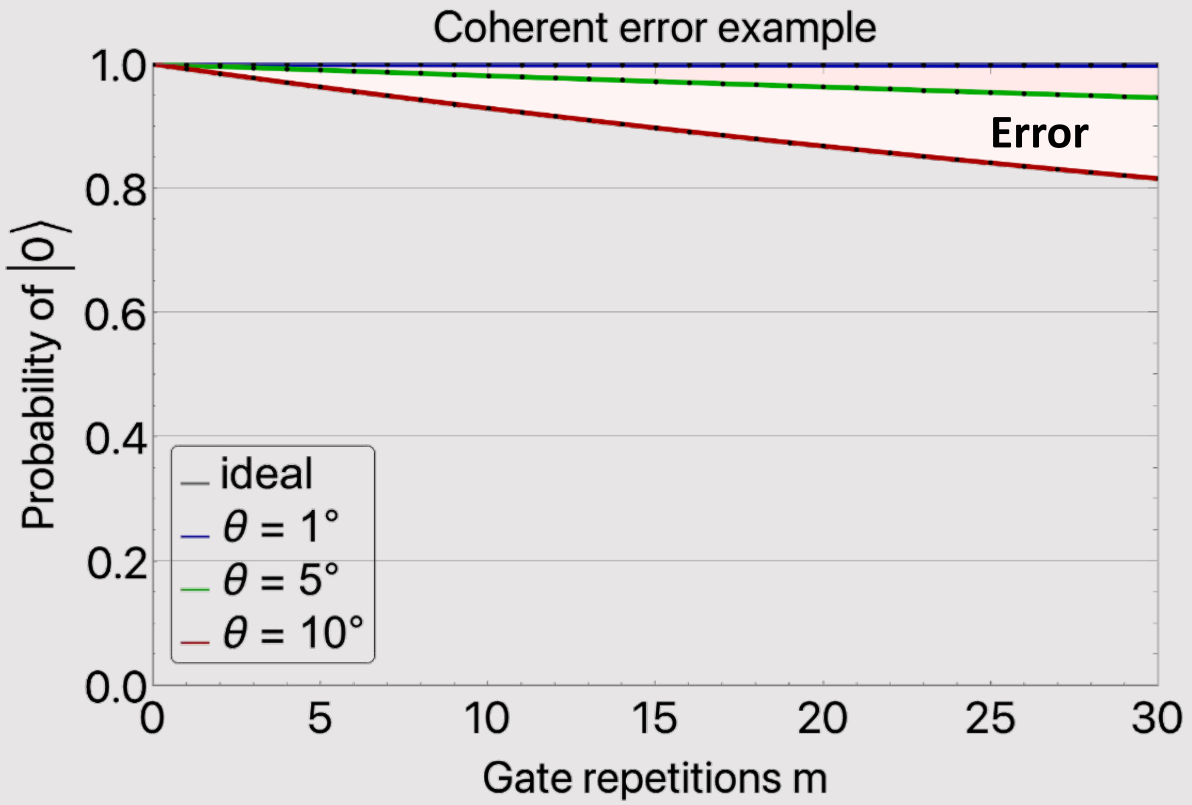


Example: Coherent over rotation

Suppose we meant to do an identity gate, but instead had a small X over rotation of angle theta

$$\left(- \boxed{R_X(\theta)} - \right)^m \qquad\qquad \left(- \boxed{P_{ai}} - \boxed{R_X(\theta)} - \boxed{P^c_{ai}} - \right)^m$$

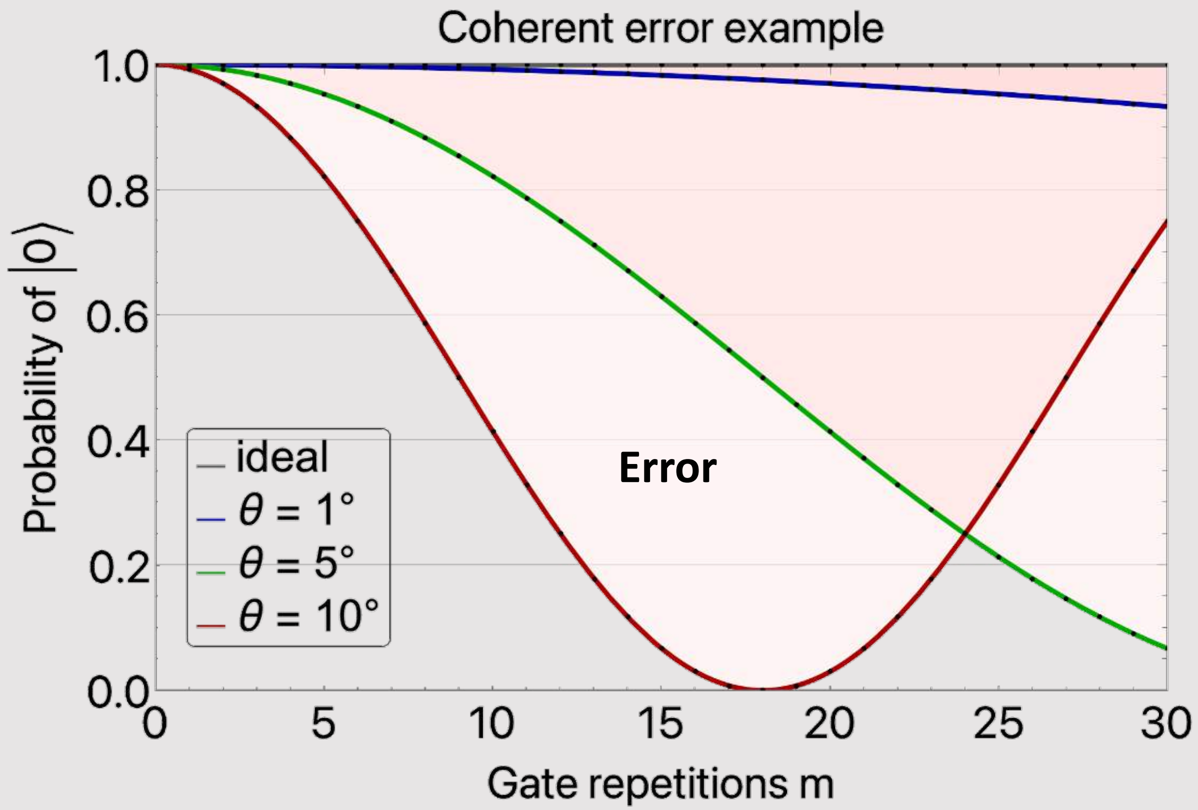


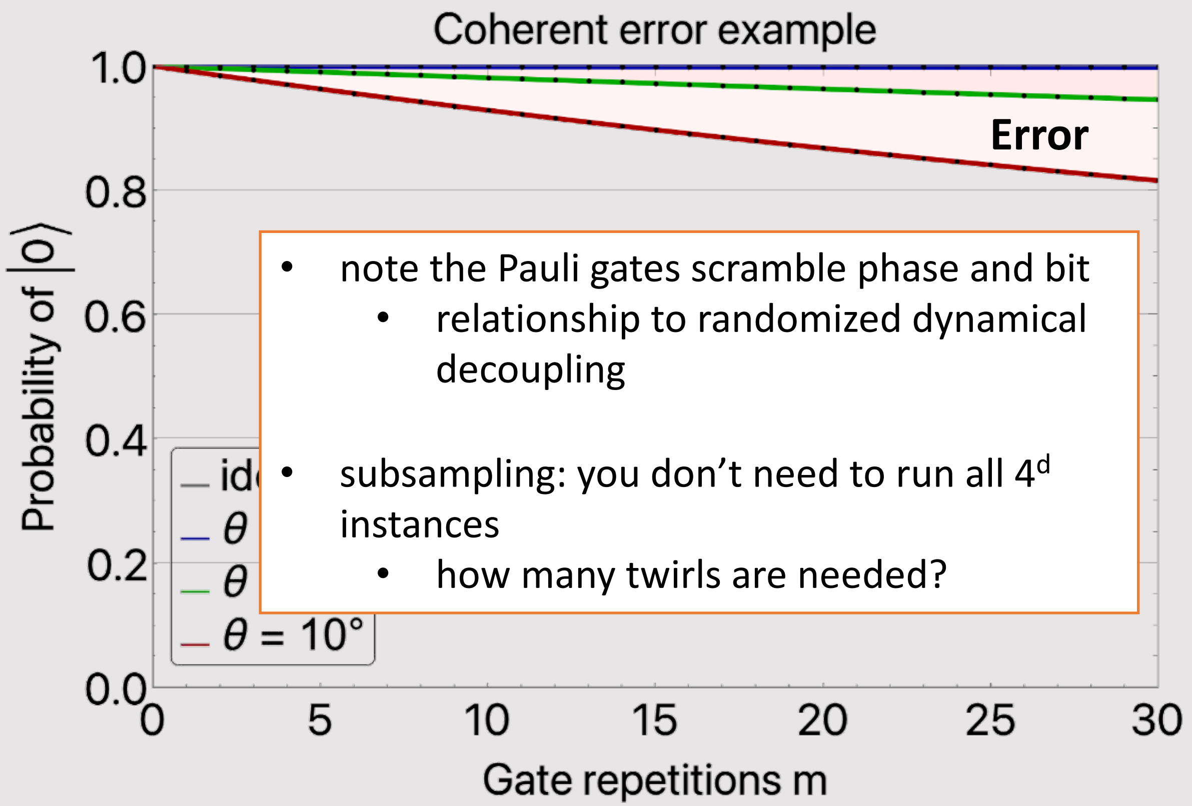


Example: Coherent over rotation

Suppose we meant to do an identity gate, but instead had a small X over rotation of angle theta

$$\left(-\boxed{R_X(\theta)}- \right)^m \qquad \left(-\boxed{P_{ai}}-\boxed{R_X(\theta)}-\boxed{P^c_{ai}}- \right)^m$$





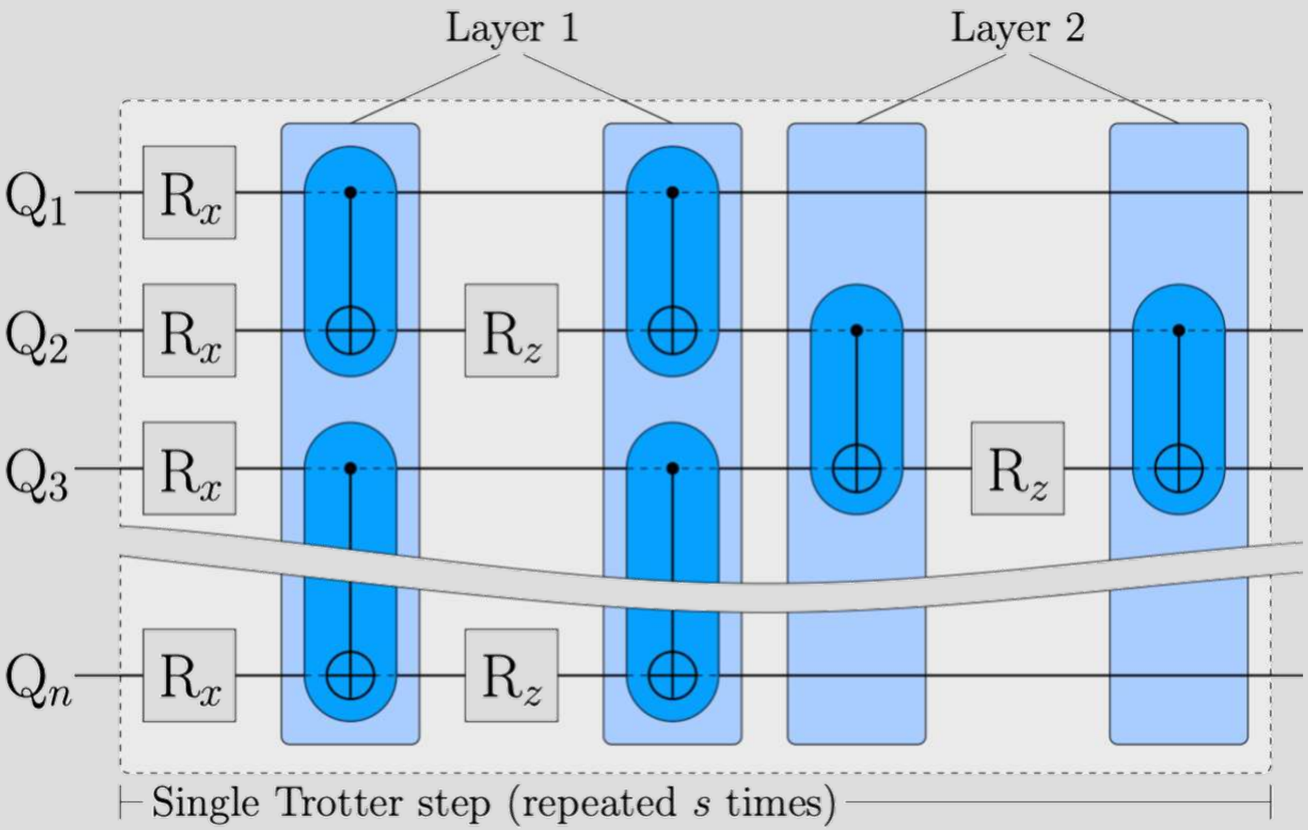
- note the Pauli gates scramble phase and bit
  - relationship to randomized dynamical decoupling
- subsampling: you don't need to run all 4^d instances
  - how many twirls are needed?

Twirling in circuits with gates

Twirling in a general circuit?

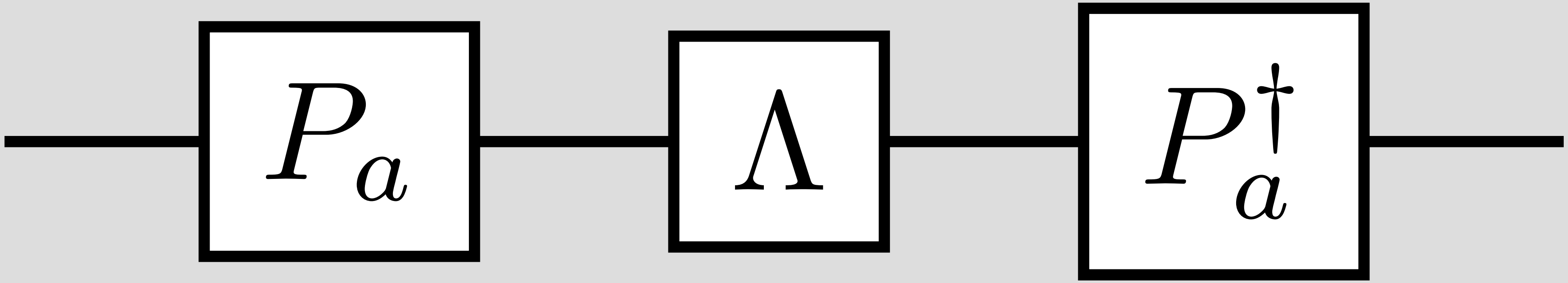
Now, imagine some general structured circuit, e.g.:
(Figure: https://arxiv.org/abs/2201.09866)

How do we describe twirling in this context systematically?



Use in a general algorithm

Use same construction

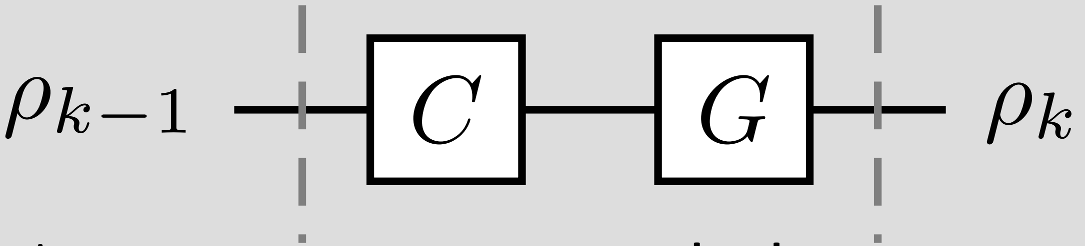


Goal: What is twirling with gates in general?

A method that introduces *independent, random, single-qubit gates* into the logical circuit such that the effective logical circuit remains unchanged but such that the noise is tailored into *stochastic Pauli errors* [Wallman2015].

A cycle of the *ideal* circuit

k-th step



Capital Latin letters (*C, G, P*, ...)

input state	**easy** ideal gate	**hard** ideal gate	output state

Channels

easy gate $\mathcal{C}(\cdot) = C \cdot C^\dagger$

hard gate $\mathcal{G}(\cdot) = G \cdot G^\dagger$

Evolution

$$\rho_k = (\mathcal{G} \circ \mathcal{C})(\rho_{k-1})$$

composition of channels (note ordering is opposite of circuit diagram)

$$\rho_k = \mathcal{G}\mathcal{C}\rho_{k-1}$$

shorthand

Terminology

Gate: an operation on a quantum state, not necessarily a unitary, can be thought of as a quantum channel as well.

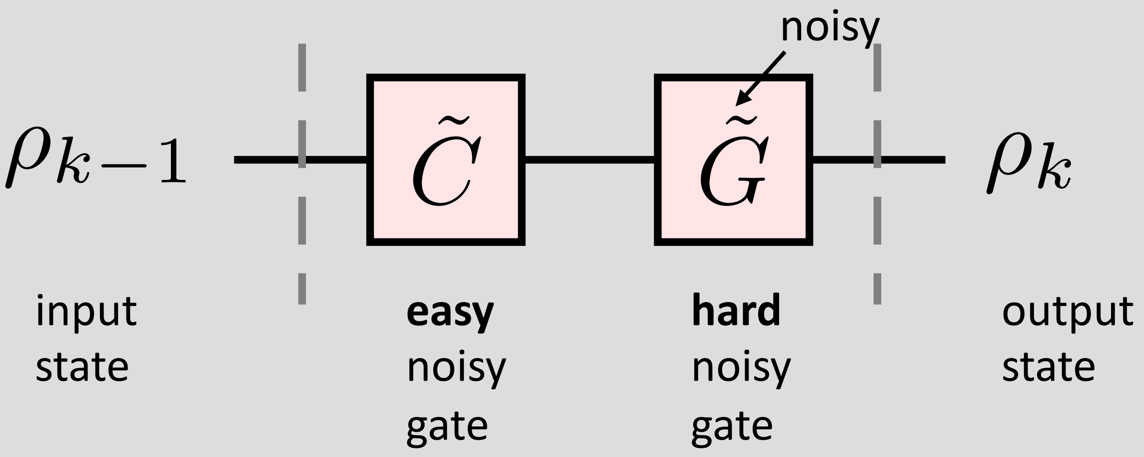
Notation

✍️

Capital Latin letters (*C, G, P*, ...)	Operators (gates)
Capital calligraphic letters ($\mathcal{C}$, $\mathcal{G}$, ...)	Super-operators (gates)
Capital Greek letters (Λ, ...)	Super-operators (noise)

A cycle of the *noisy* circuit

k-th step



Channels

$\tilde{\mathcal{C}}$ total channel of the noisy easy gate *C*

$\tilde{\mathcal{G}}$ total channel of the noisy hard gate *G*

- only thing you have have access to (noisy gates)
- not exactly known!

Evolution

$$\rho_k = (\tilde{\mathcal{G}} \circ \tilde{\mathcal{C}})(\rho_{k-1})$$

composition of channels (note ordering is opposite of circuit diagram)

$$\rho_k = \tilde{\mathcal{G}}\tilde{\mathcal{C}}\rho_{k-1}$$

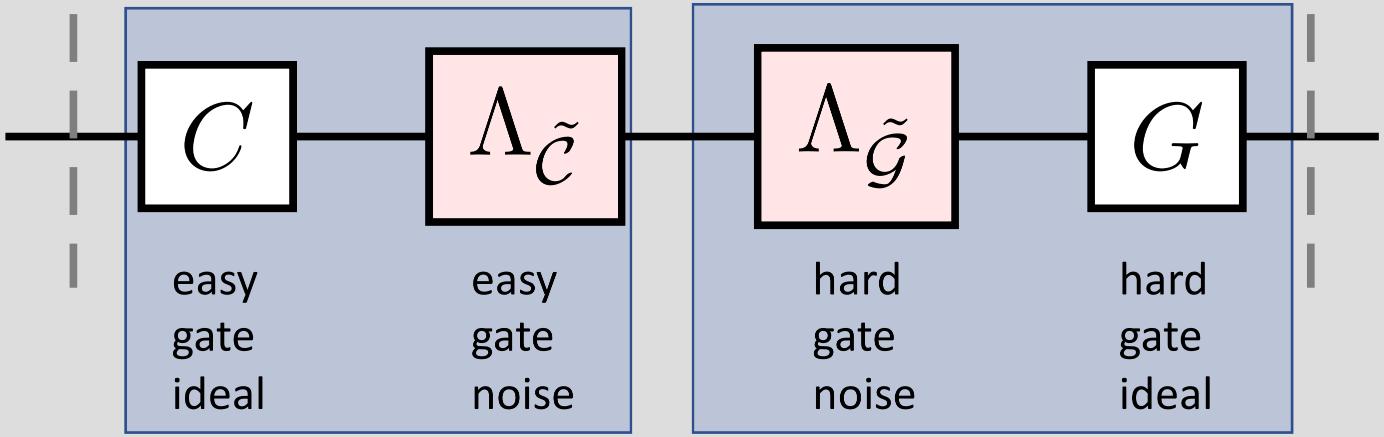
shorthand

Terminology

Noisy gate: the operation that is physically implemented, including all miscalibrations, cross-talk, coherent and incoherent errors.

Model A cycle of the noisy circuit: noise model

k-th step



Notation:
C and G labels in the circuit here indicate the unitary gate. However, when needed we will use the quantum channel version of the gate.

Noise decompositions

$$\tilde{\mathcal{C}} = \Lambda_{\tilde{\mathcal{C}}}\mathcal{C}$$

noisy gate channel — noise channel — ideal gate channel

$$\tilde{\mathcal{G}} = \mathcal{G}\Lambda_{\tilde{\mathcal{G}}}$$

Evolution

$$\tilde{\mathcal{G}}\tilde{\mathcal{C}} = \mathcal{G}\Lambda_{\tilde{\mathcal{G}}}\Lambda_{\tilde{\mathcal{C}}}\mathcal{C}$$

$$\tilde{\mathcal{G}}\tilde{\mathcal{C}} = \mathcal{G}\tilde{\Lambda}\mathcal{C}$$

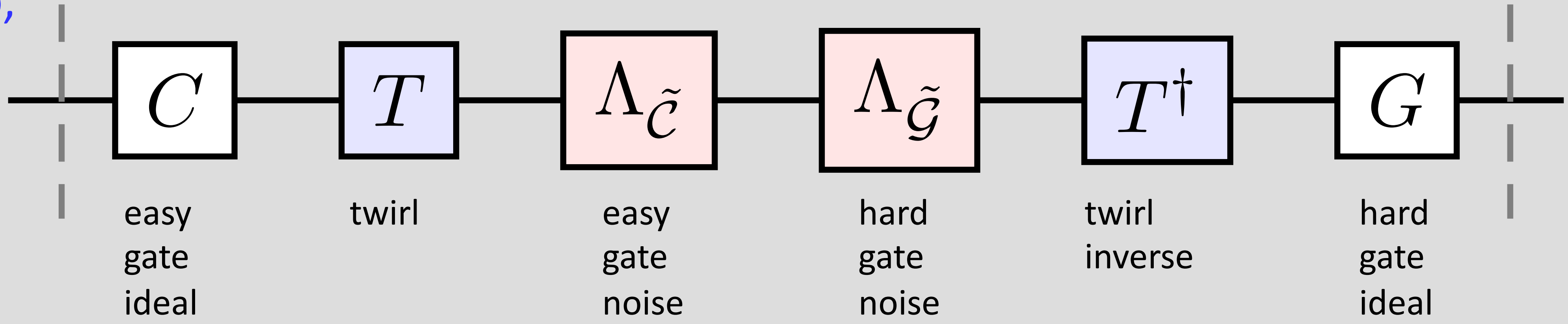
lump easy and hard gate noise into one channel

$$\tilde{\mathcal{G}}\tilde{\mathcal{C}} = \mathcal{G}\Lambda\mathcal{C}$$

after twirling (to come)

Twirling a cycle of the noisy circuit

k-th step, single instance



Virtual twirl

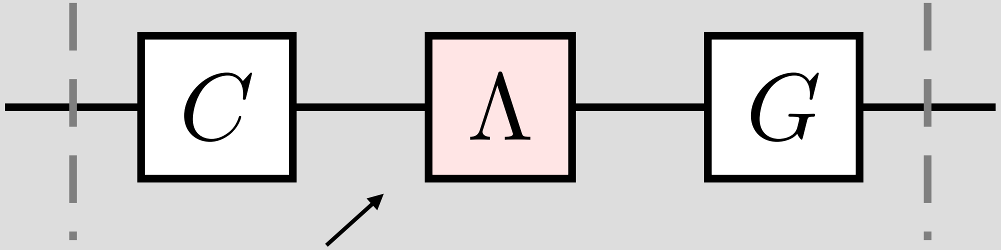
Randomly sample from set $T \in \mathbf{T}$

$$\mathcal{T}(\cdot) := T \cdot T^\dagger$$

Average over all twirl instances yields new effective noise channel, a stochastic Pauli one

$$\mathbb{E}_{T\in\mathbf{T}}\left[\mathcal{T}^\dagger \Lambda_{\tilde{\mathcal{G}}} \Lambda_{\tilde{\mathcal{C}}} \mathcal{T}\right] =: \Lambda$$

Average twirled circuit

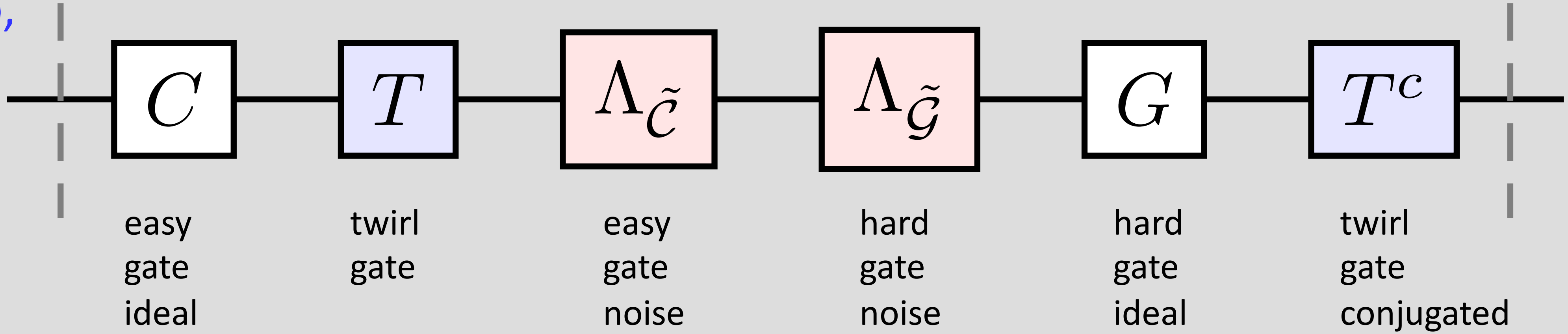


Before: General noise
Now: Stochastic Pauli noise when twirl over Pauli group

Note: the average value is linear functional, but variance is not

Twirling a cycle of the noisy circuit

k-th step, single instance



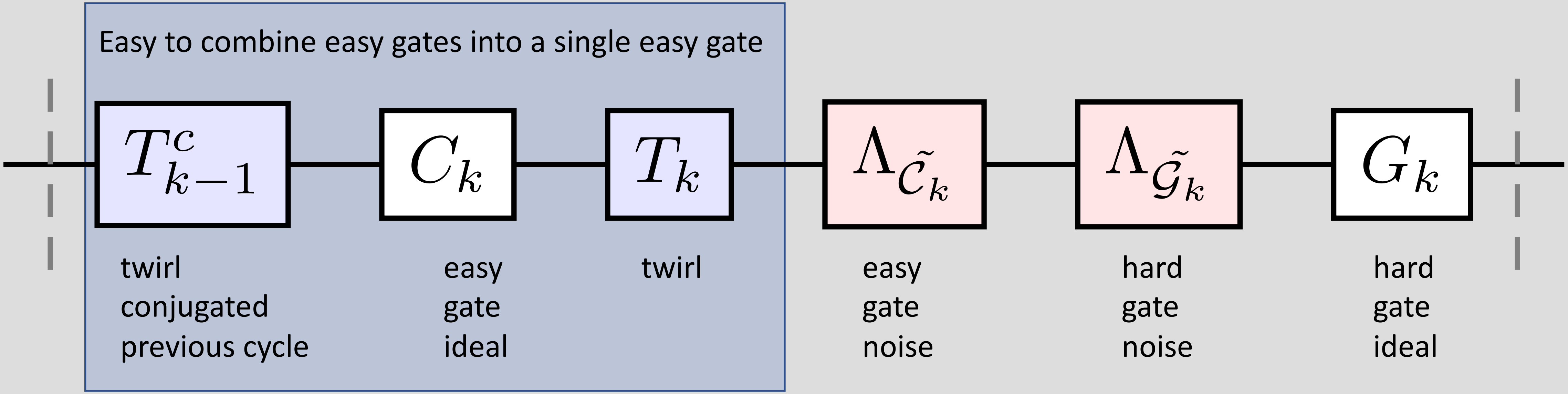
Conjugating the twirl

$$GT^\dagger = T^c G$$

$$T^c = GT^\dagger G^\dagger$$

Putting it together over multiple cycles

k-th step, single instance



Summary of twirling

What does RC achieve? Arbitrary Markovian noise processes can be reduced to effective Pauli processes. Additionally, noise with arbitrary coherence and spatial correlations (such as over rotations and cross-talk) can be tailored into stochastic Pauli noise. RC is fully robust against arbitrary gate-dependent errors on the gates that are most difficult to implement, i.e., the hard gates. Claim: Stochastic Pauli errors with the same average error rate r as a coherent error.

Advantages of RC. Can dramatically *reduce error rates* with little or no experimental overhead. Only requires local gates to tailor general noise on multi-qubit gates into Pauli noise. Reduction in worst-case error. Robust to inevitable variation in errors over the randomizing (easy) gates. Worst-case error rate should be experimentally measurable [18]. Applied directly to gate sets that are universal for quantum computation, including all elements in a large class of fault-tolerant proposals. Also, used in sims of noise since Clifford circuits.

Sampling complexity. *Reputed* to be not only low but even linear or constant with both the number of qubits and the depth of the circuit [6]. Todo: Check.

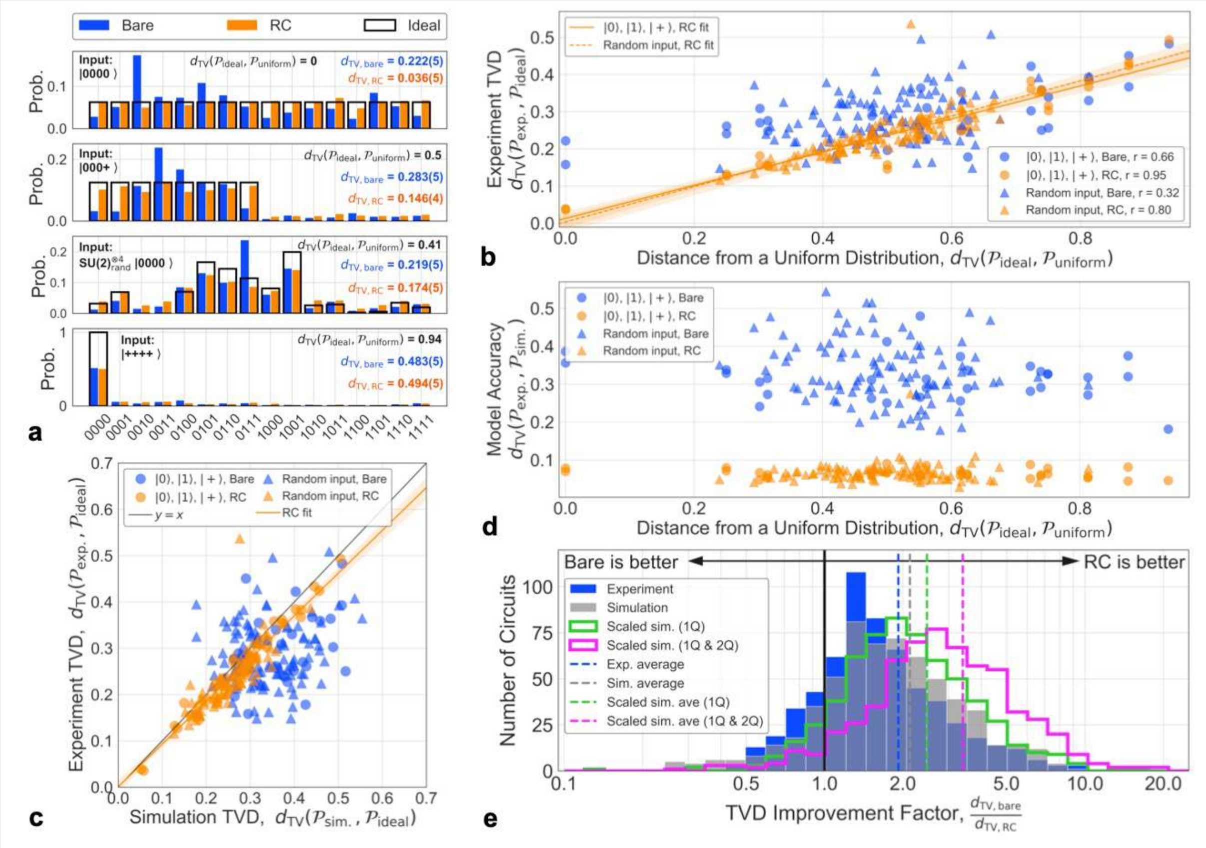
Limitations.

1. *Static noise model for easy gates.* Assumes that the easy gates are all subject to a gate-independent error model $\mathcal{E}_e$. This can however be relaxed, see Eq. (10) of Ref. [18]. Claim is that this introduces very little additional error, especially when $\mathbf{T}$ is a group normalized by $\mathbf{C}$, which is satisfied by the canonical group sets.
2. How large is the amplitude of the tailored noise? If the tailored noise depolarizes the channel in one step, clearly, the protocol is not useful. In a preliminary check of the single qubit case however, we found that the scaling for depolarizing noise is favorable. To be explored more.
3. *Non-Markovian noise.* A particularly significant open problem is the robustness of the RC technique to noise that remains non- Markovian on a time-scale longer than a typical gate time. This could be mitigated with randomized dynamic decoupling, i.e., applying random sequences of Pauli operators to echo out non-Markovian contributions. It is expected that RC thus also benefits from from some dynamical decoupling, though this is not well studied and remains an open problem.

High level: Why twirling?

High-level messages

- Accurately predict algorithm performance using experimentally-measured error rates on universal circuits
- RC: scalable, in situ, low-classical overheard
  - agnostic to noise model
  - closing gap between algorithm performance and RB predictions
  - easy to implement
- Increase accuracy
  - reduction in the total variation distance from dTV_{bare}
    - basis- dependent metric which determines the *probability of obtaining an incorrect solution*
    - R correlation
- PEC / cycle benchmarking & reconstruction
  - residual error syndromes are broadly distributed
  - Only need *R*=20-100 random circuits typically
- For simulations: Pauli channel approximation + Gottesman-Knill theorem for efficient Clifford simulation



Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor

Why does twirling actually work?

Theory and my take on it