Error Budgeting for Superconducting Modular Quantum Architecture Designs

Evan McKinney[1],
I. Yusuf[2], G. Falstin[1],G. Agarwal[2], M. Hatridge[2], A.K. Jones[3]

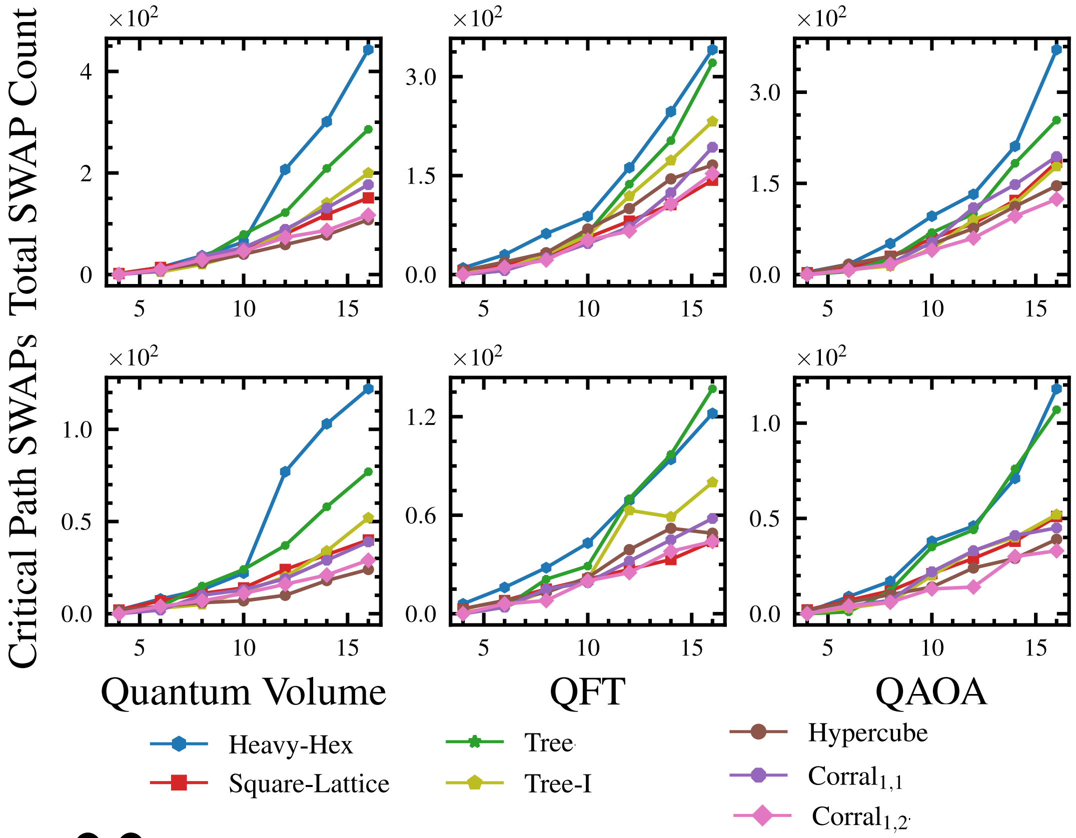
[1]Department of Electrical and Computer Engineering, University of Pittsburgh
[2]Department of Applied Physics, Yale University
[3]Department of Electrical Engineering and Computer Science, Syracuse University

APS March Meeting (2025)

Transpilation as co-design framework

Quantum algorithm
↓
Circuit mapping
↓
Gate decomposition to ISA
↓
Control optimization
↓
Device execution

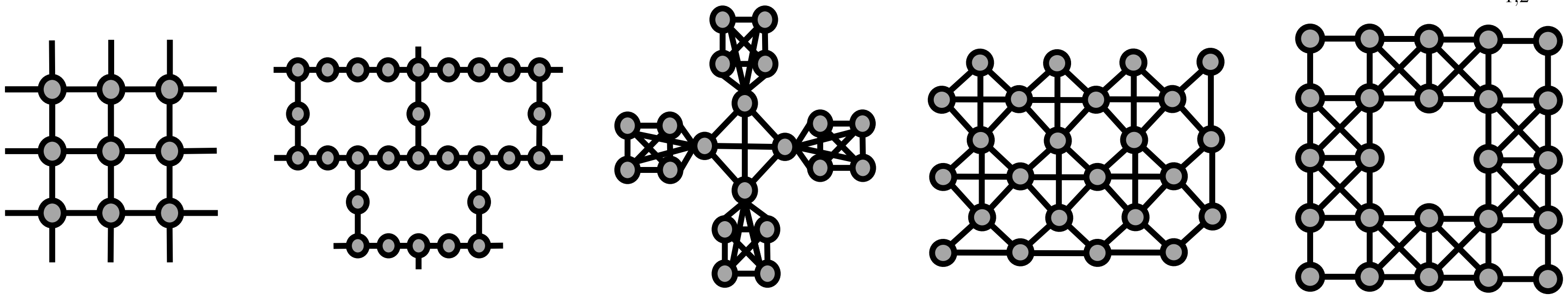
Connectivity reduces circuit costs



Device Constraints
- mode frequency bandwidth
- fabrication precision
- hybridization strengths
- coupler speed-limit

Gate fidelity threshold
- # qubits per module
- qubit-coupler connectivity

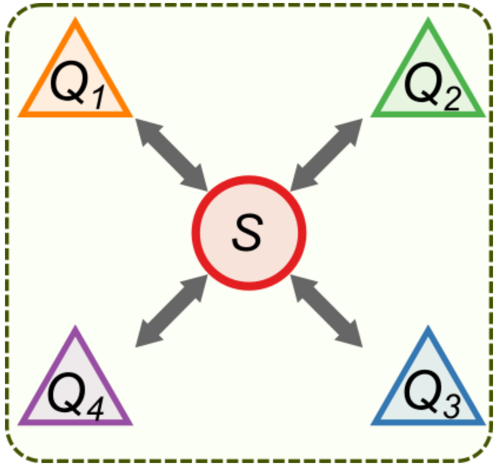
Qubit frequency allocation
- # unique qubit frequencies

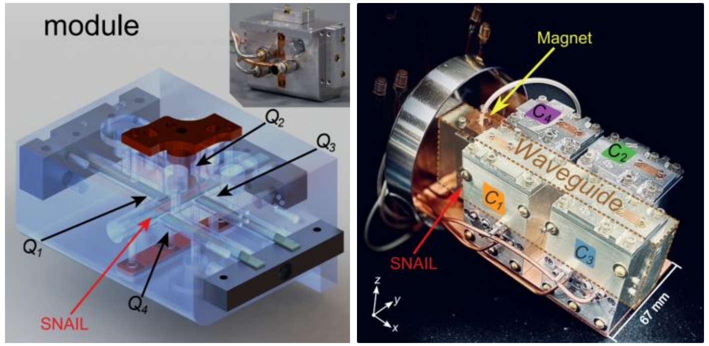


How much qubit connectivity is feasible with high fidelity?

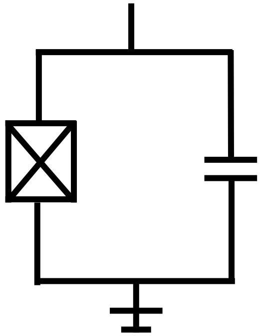
McKinney, et al. **HPCA** (2023)

Four qubit SNAIL-based quantum module





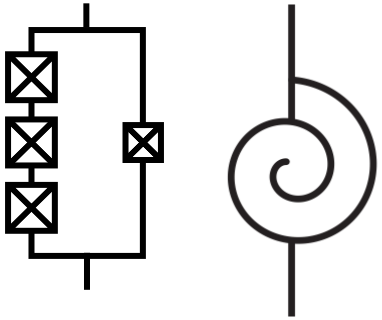
Transmon (qubit)



$$H = \omega_q q^\dagger q + \frac{\alpha}{12}(q^\dagger + q)^4$$

Module:
Yusuf, **MAR-A18.13**
Repicky, **MAR-L18.13**

SNAIL (coupler)



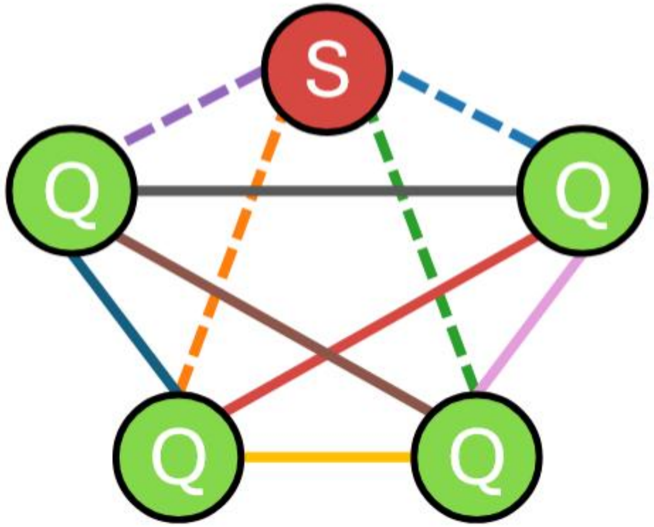
$$H = \omega_s s^\dagger s + g_3(s^\dagger + s)^3$$

SNAIL:
Wang, **MAR-W09.11**
Mesits, **MAR-N17.3**
Nowicki, **MAR-T09.11**

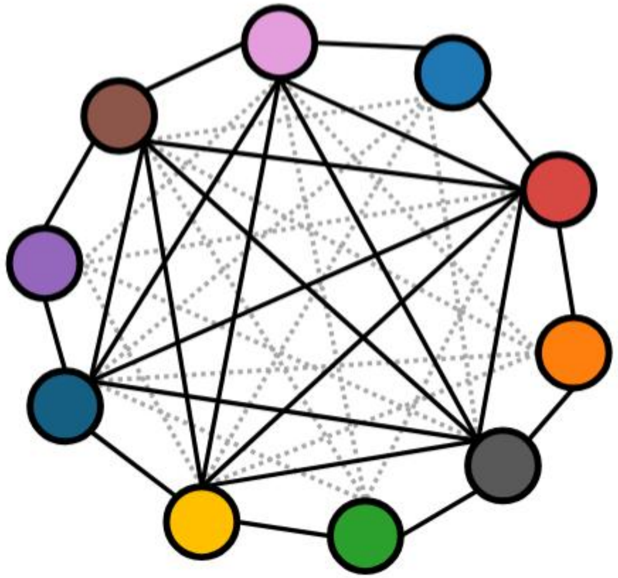
Xia, et al. **arXiv:2306.10162** (2023)
Zhou, et al. **npj Quantum Inf 9**, 54 (2023)
Frattini, et al. **Applied Physics Letters** (2017)

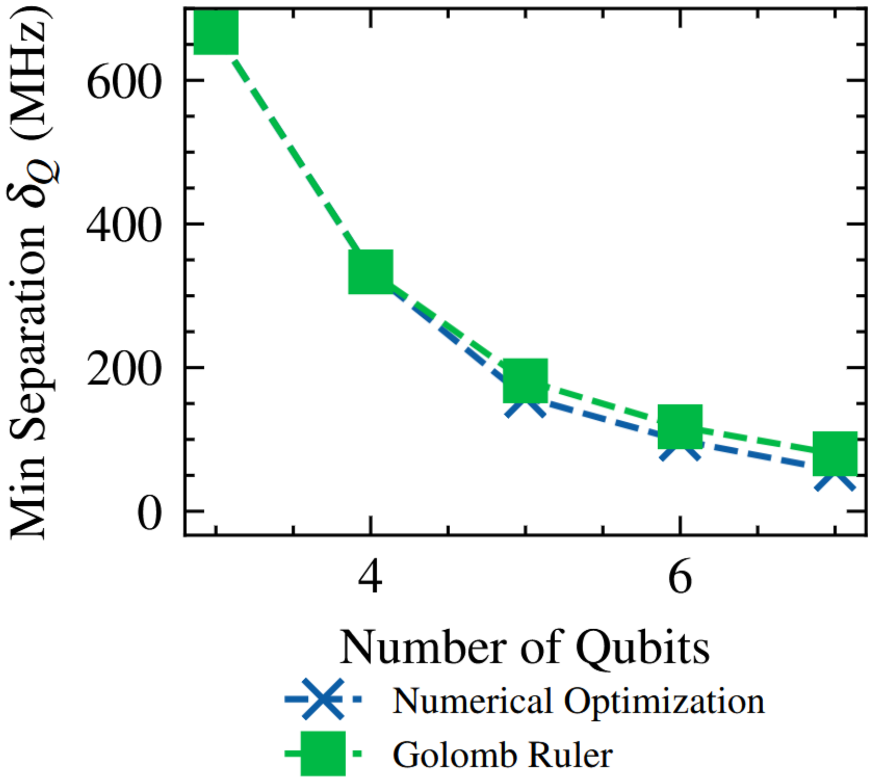
Gate crowding in a module

Pairs of driven interactions
Edge: conversion frequency



"Compatibility graph"
Edge: separation constraint (q)





<u>Frequency Allocation Problem</u>

$$|f(i) - f(j)| \geqslant q_{ij} \quad \forall (i,j),\, i \neq j$$

$$\max \min_{i,j \in X, i \neq j} |f(i) - f(j)|$$

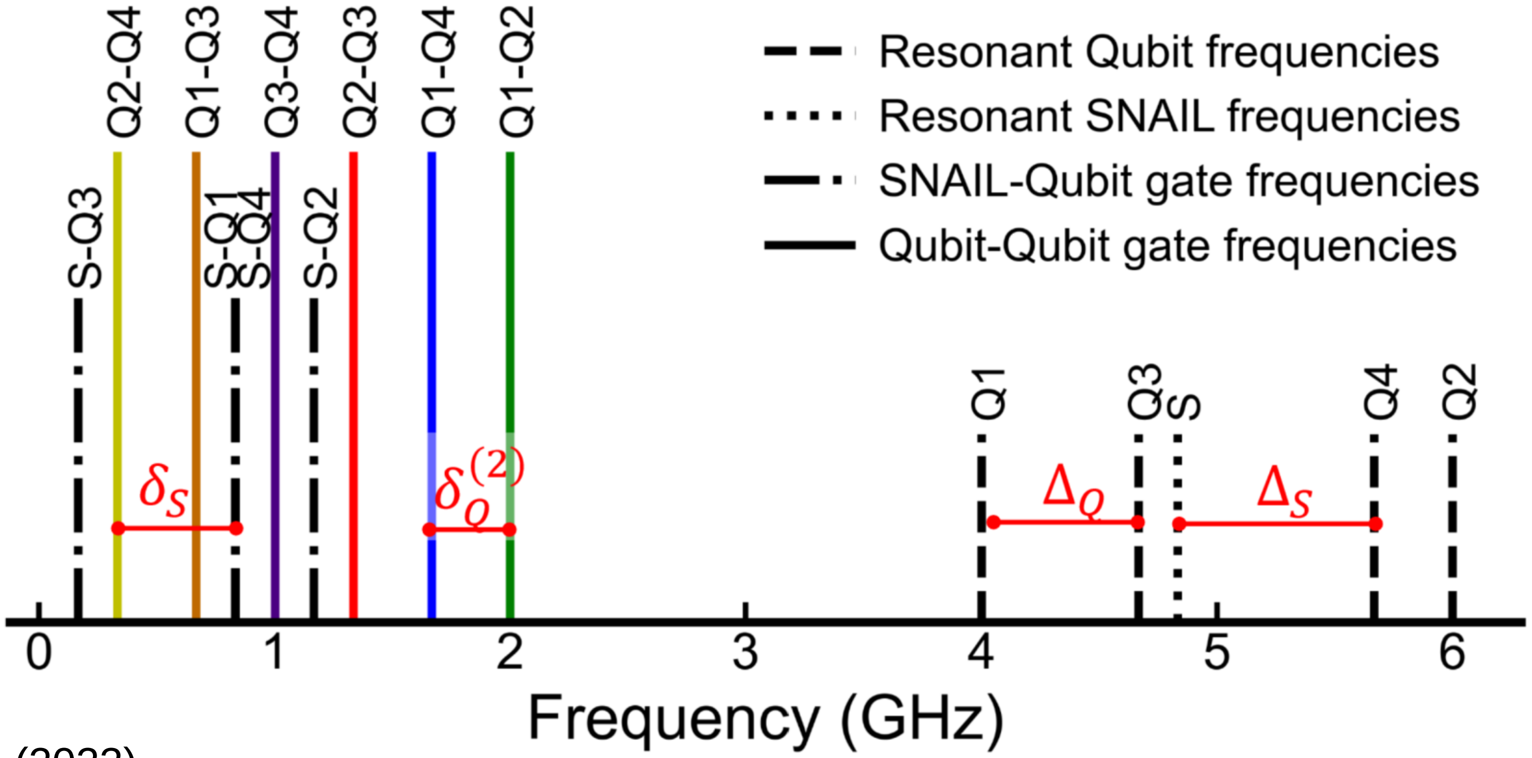
$$2pk + (k^2 \bmod p),\, k \in [0, p-1]$$

de Freitas, et al. **ITOR** (2021)
Park, et al. **J. Oper. Res. Soc. Jpn.** (1996)
Memarsadeghi. **CSE** (2016)

Spectral crowding of terms

$$\tilde{H}_I = g_3\left(se^{-i\tilde{\omega}_s t} + \eta e^{-i\omega_p t} + \sum_i \lambda_{si} q_i e^{-i\tilde{\omega}_{q_i} t} + \text{h.c.}\right)^3$$

$$+\sum_i \frac{\alpha_i}{12}\left(q_i e^{-i\tilde{\omega}_{q_i} t} - \lambda_{si}(se^{-i\tilde{\omega}_s t} + \eta e^{-i\omega_p t}) + \text{h.c.}\right)^4$$

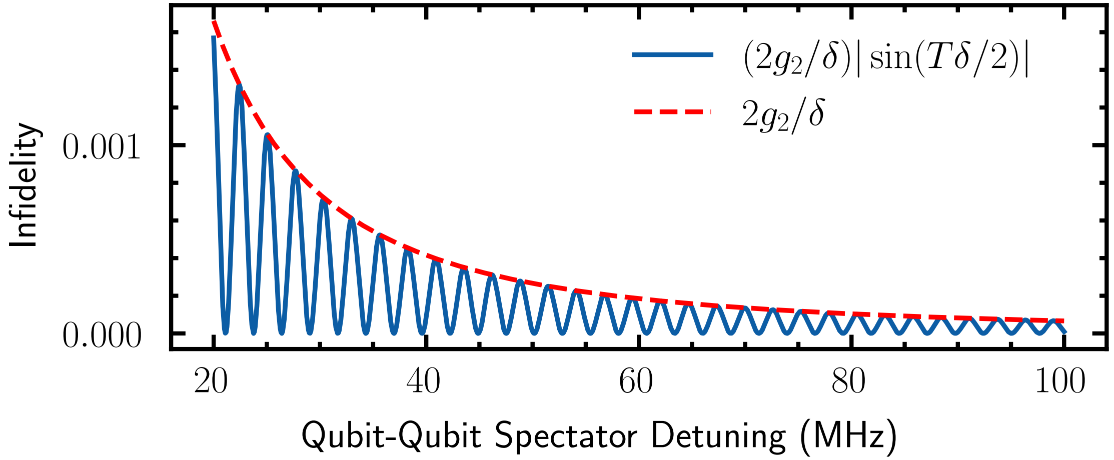


Zhou, Chao. **PhD Thesis** (2023)

Sources of infidelity from module crowding

Coherent error from spectators

Unwanted unitary terms decay with detuning



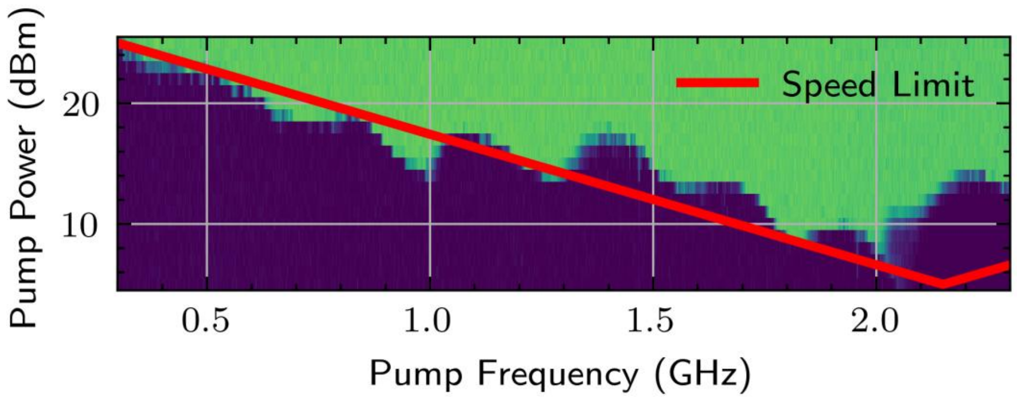
Remove explicit time-dependence with a worst-case approximation.

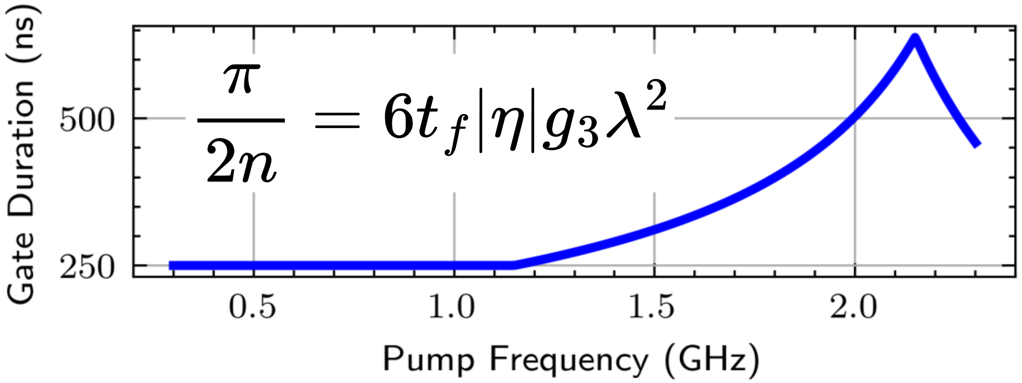
$$\left|\int_0^T e^{-it\delta}dt\right| = 2|\sin(T\delta/2)|/\delta$$

Zhou, Chao. **PhD Thesis** (2023)
McKinney, et al. **ISCA** (2023)

Incoherent loss from speed limits

Experimentally characterized max drive strength



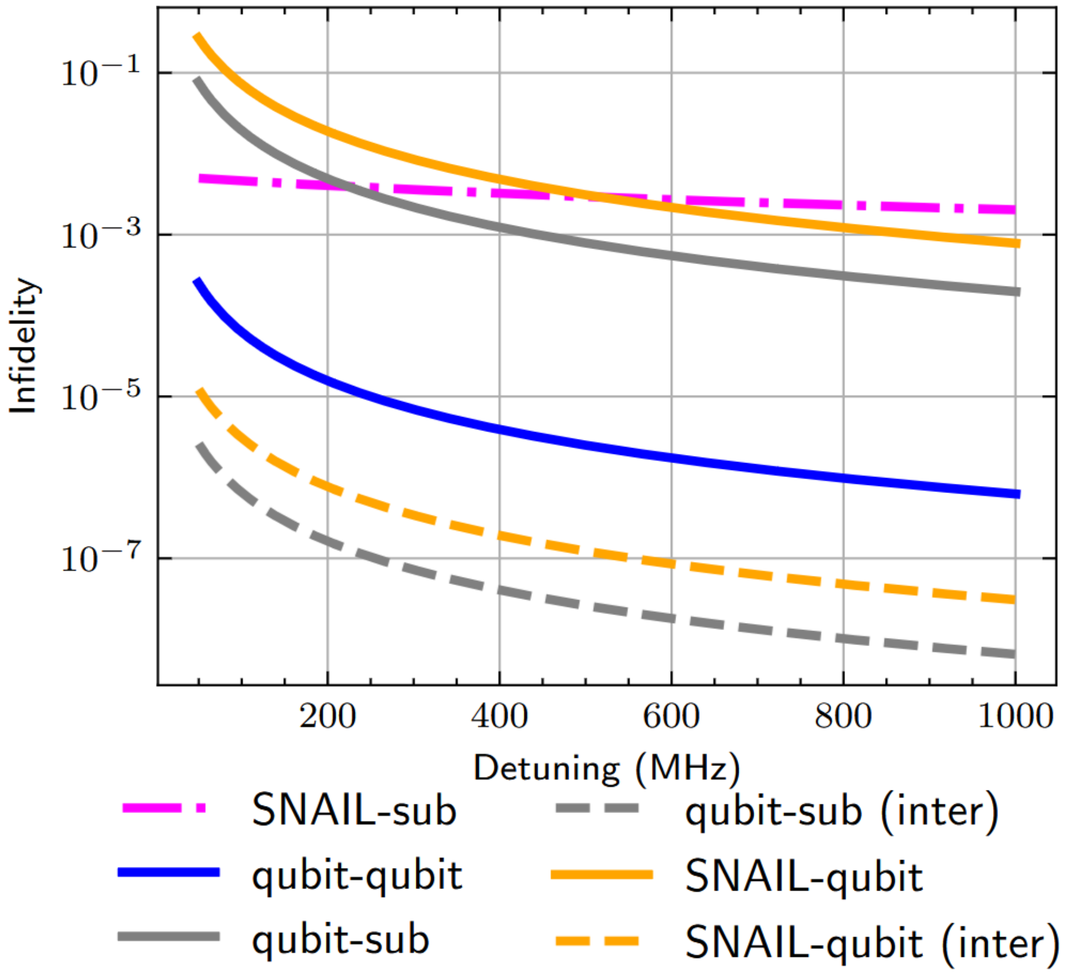


$$|\eta| = \frac{\varepsilon\omega_s}{\omega_p^2 - \omega_s^2}$$

Parametric Speed limits:
Xia, **MAR-L36.2**
Capocci, **MAR-L36.3**

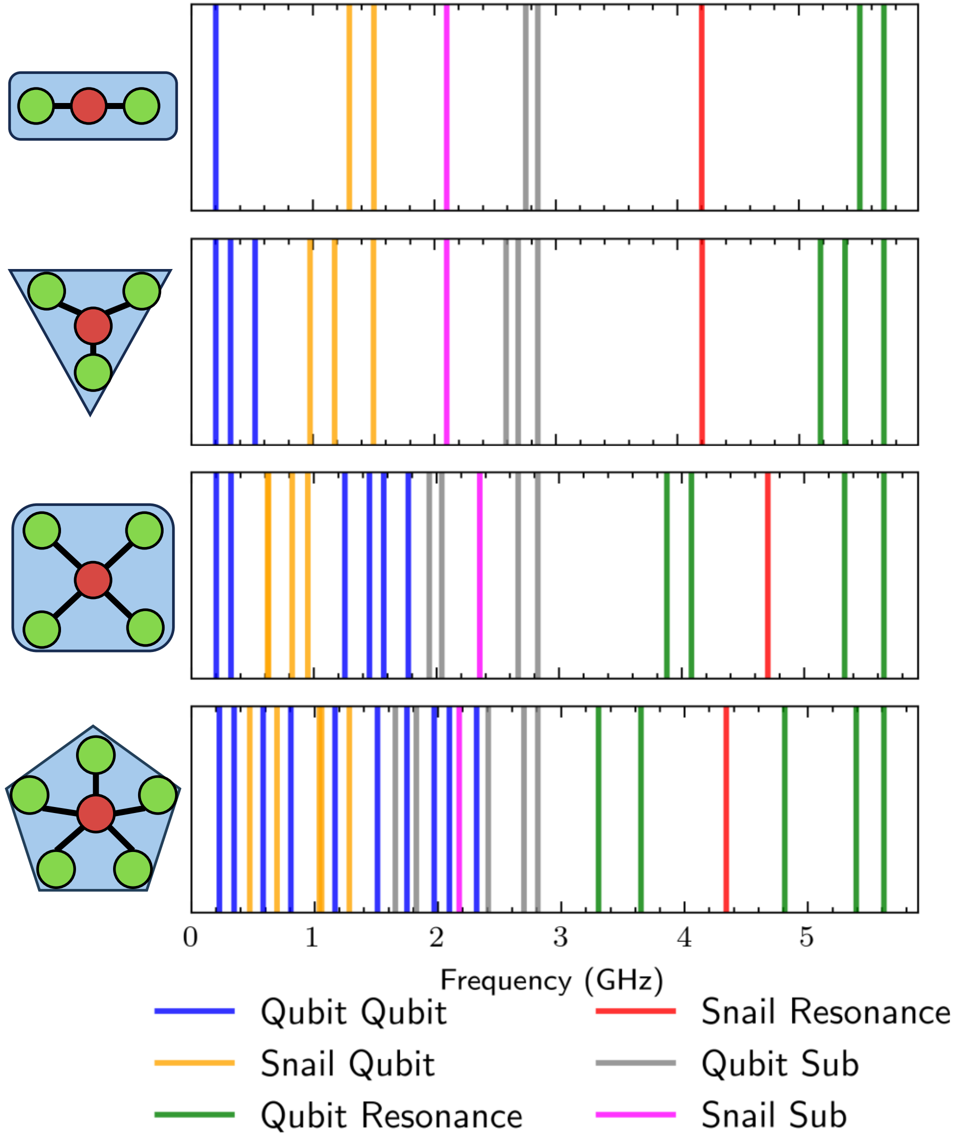
Error budgeting against spectators

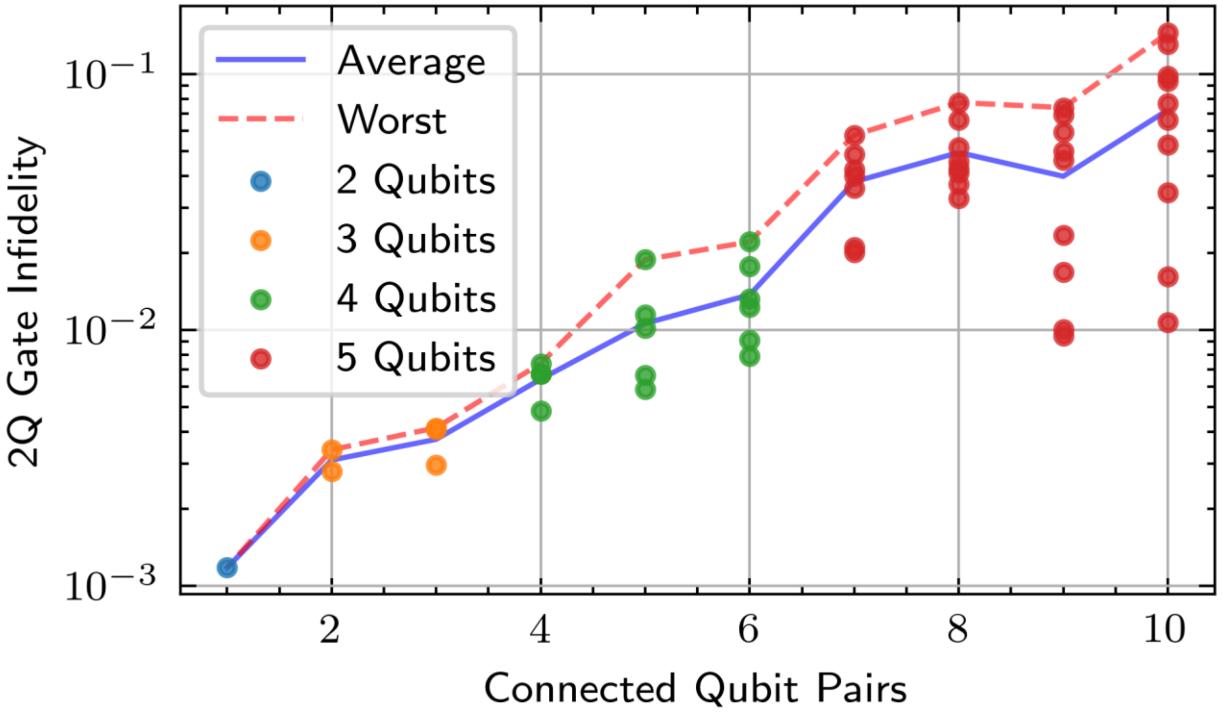
Term	Coefficient	$\omega_p =$	Normalized Prefactor
Driven Term			
$(q_a^\dagger q_b + q_a q_b^\dagger)$	$6\|\eta\|\lambda^2 g_3$	$\|\omega_{q_b} - \omega_{q_a}\|$	1.0
Intra-Module Spectator Terms			
$(s^\dagger + s)$	$3\|\eta\|^2 g_3$	$\omega_s/2$	100.0
$(s^\dagger q_a + s q_a^\dagger)$	$6\|\eta\|\lambda g_3$	$\|\omega_s - \omega_{q_a}\|$	10.0
$(q_a^\dagger + q_a)$	$3\|\eta\|^2 \lambda g_3$	$\omega_{q_a}/2$	10.0
$(s^\dagger q_a + s q_a^\dagger)$	$\alpha\|\eta\|^2 \lambda^3$	$\|\omega_s - \omega_{q_a}\|/2$	0.067
$(q_a^\dagger + q_a)$	$\alpha\|\eta\|^3 \lambda^3/3$	$\omega_{q_a}/3$	0.044
$(s^\dagger + s)$	$N_q \alpha\|\eta\|^3 \lambda^4/3$	$\omega_s/3$	0.018
Inter-Module Spectator Terms			
$(s_n^\dagger + s_n)$	$3\|\eta\|^2 \lambda^2 g_3$	$\omega_{s_n}/2$	1.0
$(s^\dagger q_c + s q_c^\dagger)$	$6\|\eta\|\lambda^3 g_3$	$\|\omega_s - \omega_{q_c}\|$	0.1
$(q_c^\dagger + q_c)$	$3\|\eta\|^2 \lambda^3 g_3$	$\omega_{q_c}/2$	0.1
$(q_a^\dagger q_c + q_a q_c^\dagger)$	$6\|\eta\|\lambda^4 g_3$	$\|\omega_{q_c} - \omega_{q_a}\|$	0.01
$(s_n^\dagger q_a + s_n q_a^\dagger)$	$6\|\eta\|\lambda^5 g_3$	$\|\omega_{s_n} - \omega_{q_a}\|$	0.001
$(q_c^\dagger q_d + q_c q_d^\dagger)$	$6\|\eta\|\lambda^6 g_3$	$\|\omega_{q_d} - \omega_{q_c}\|$	0.0001



For each interaction, use infidelity vs detuning characterization in **Nelder-Mead minimization** to allocate best mode frequencies

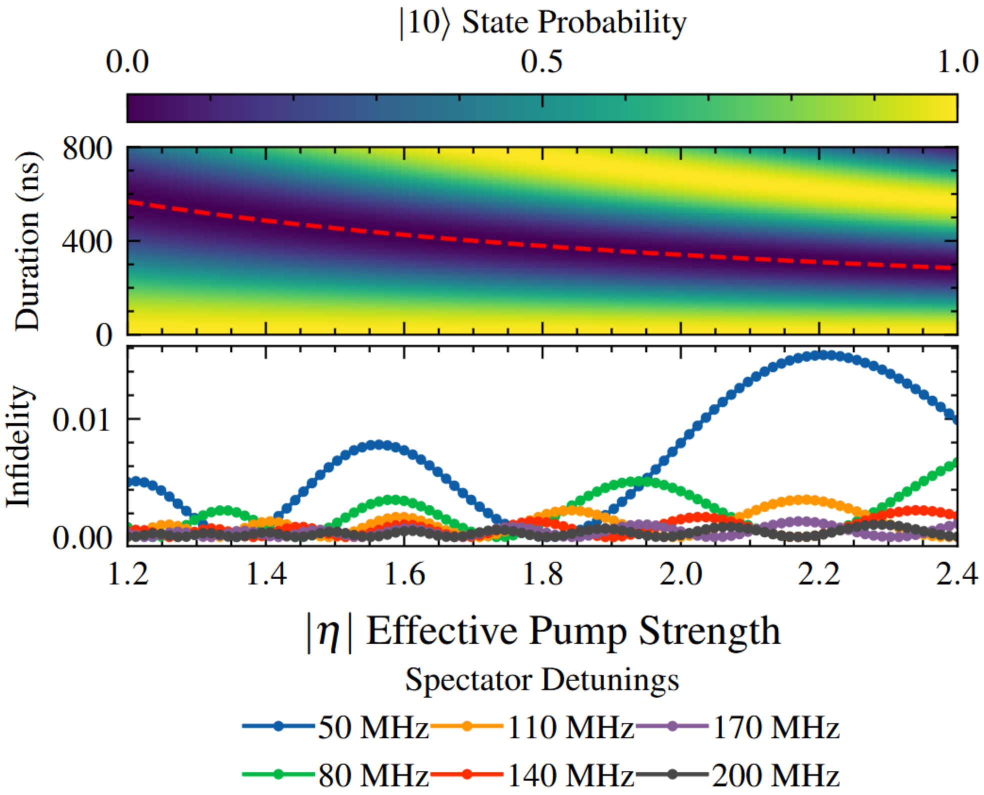
Module fidelities from optimized frequency stacks

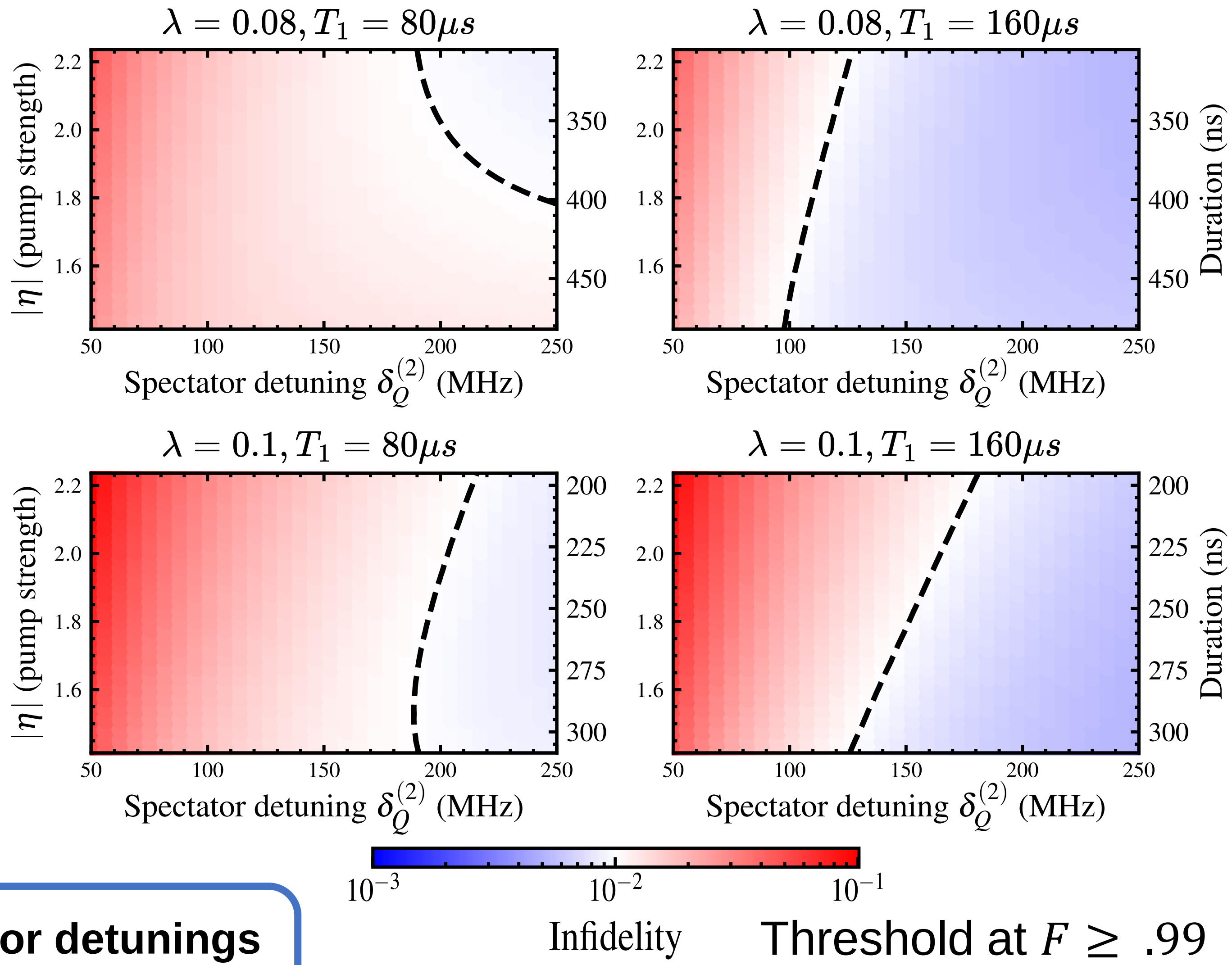




N-qubit module	Avg gate fidelity
2	.996
3	.994
4	**.990**
5	.940

Tradeoffs between gate speed and spectators





Threshold at $F \geq .99$

(a) Larger frequency bandwidth: **increase spectator detunings**

(b) Longer qubit lifetimes: **slower gates, weaker spectators**

Wei, et al. **Phys Rev Applied** (2024)

Conclusions

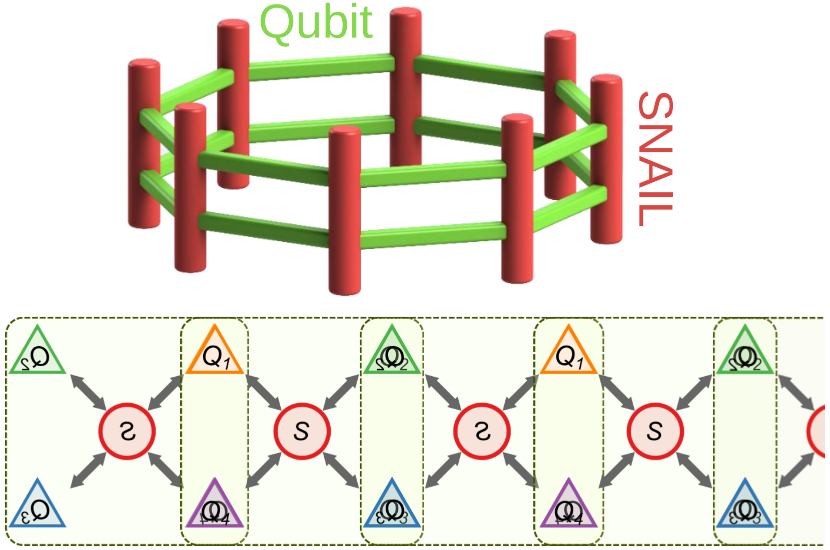
McKinney, et al. ***arXiv:2409.18262*** (2024)

- **Developed a fidelity model from characterized device properties**
  - Quantified tradeoffs between coherent and incoherent noise sources
  - Develop numerical optimization to minimize total average gate infidelity
- **Frequency allocation problem to determine viable module sizes**
  - SNAIL module supports *up to 4 qubits* before fidelity drops below 0.99
  - *Partial frequency allocation* improves base fidelity up to 0.994, yielding a *19.9% relative improvement* in computational accuracy.

Alex K. Jones

Girgis Falstin

Up next: Yusuf, MAR-A18.13

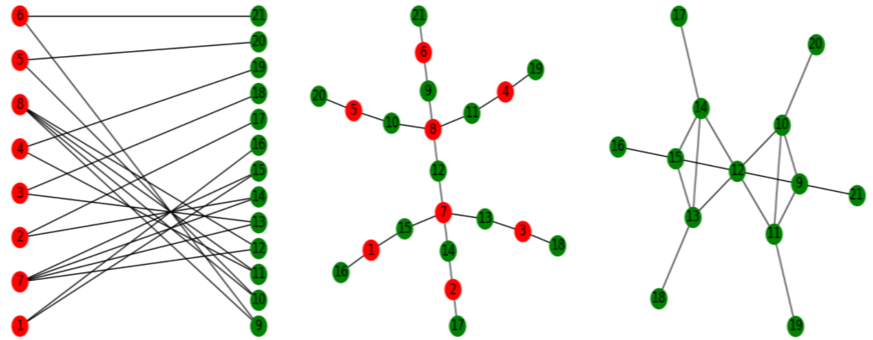


Procedurally generating topologies

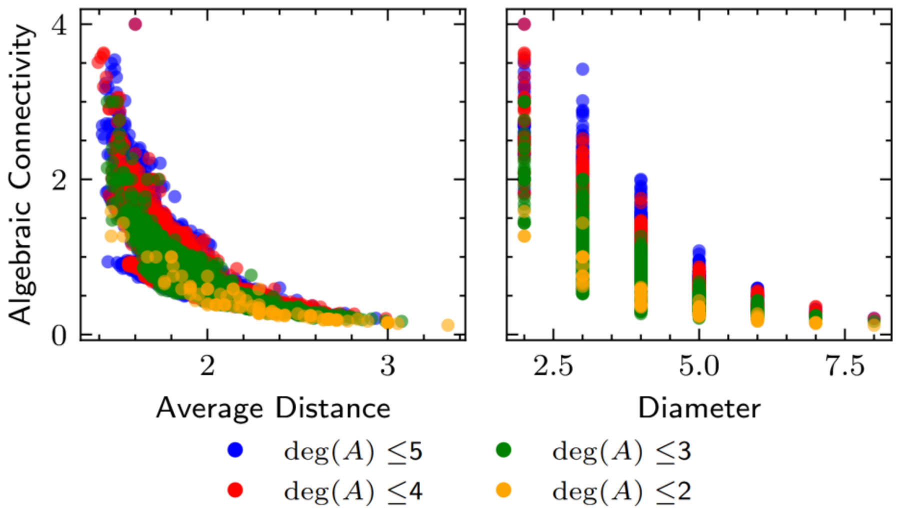
Greedy inverse bipartite projection

1: Satisfy all edges from input G'_A on G_A

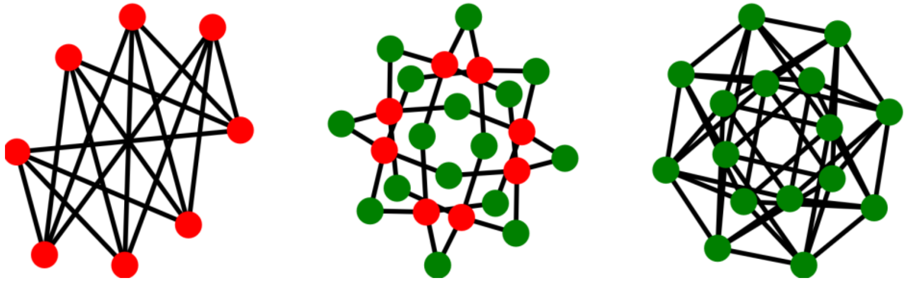
2: Add nodes to G_B until saturate all couplers



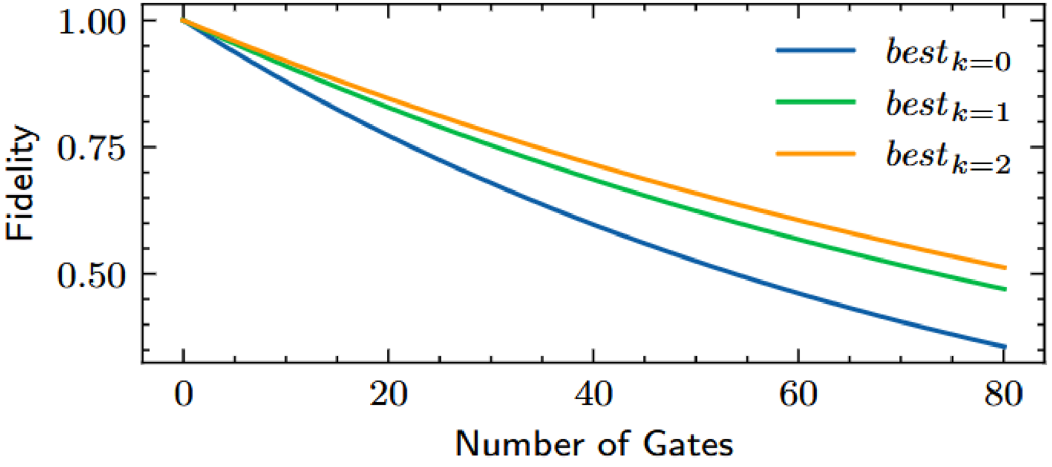
3: Verify isomorphism between input G'_A, output G_A



Selective edge removal



$$F \approx (1 - \epsilon_{\text{gate}})^{(1+3\text{CCR})\times\#\text{gates}}$$



Sacrificing dense connectivity for less spectator crowding is a worthwhile tradeoff!