

OPTIMAL QUANTUM CONTROL AND AN APPLICATION

Frank K. Wilhelm, Nicolas Wittler, Kevin Pack, **Forschungszentrum Jülich**

Shai G. Machnes, Anurag Saha Roy, **Qruise GmbH**

Stefan Fillip, Federico Roy, **Walther-Meissner-Institute Garching**

Daniel Egger, **IBM Research Rüschlikon**

BENASQUE SCHOOL ON SUPERCONDUCTING QUBITS | APRIL 19, 2023

arXiv:2003.10132; arXiv:2009.09866 arXiv:1508.00442

Mitglied der Helmholtz



QUANTUM
FLAGSHIP



FIDELITY IS THE LIMITING FACTOR

Google supremacy experiment

[Arute *et al.* Nature 574, 505 (2019)]

99.8% noise

30,000,000 shots

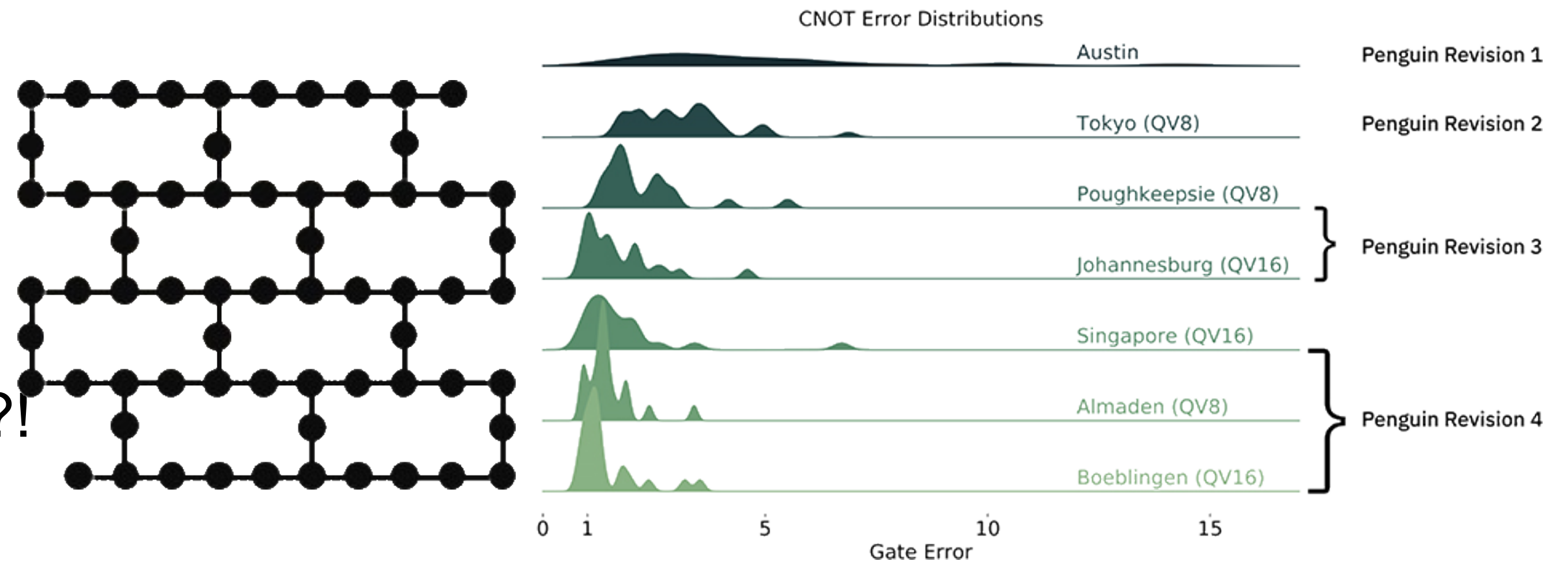
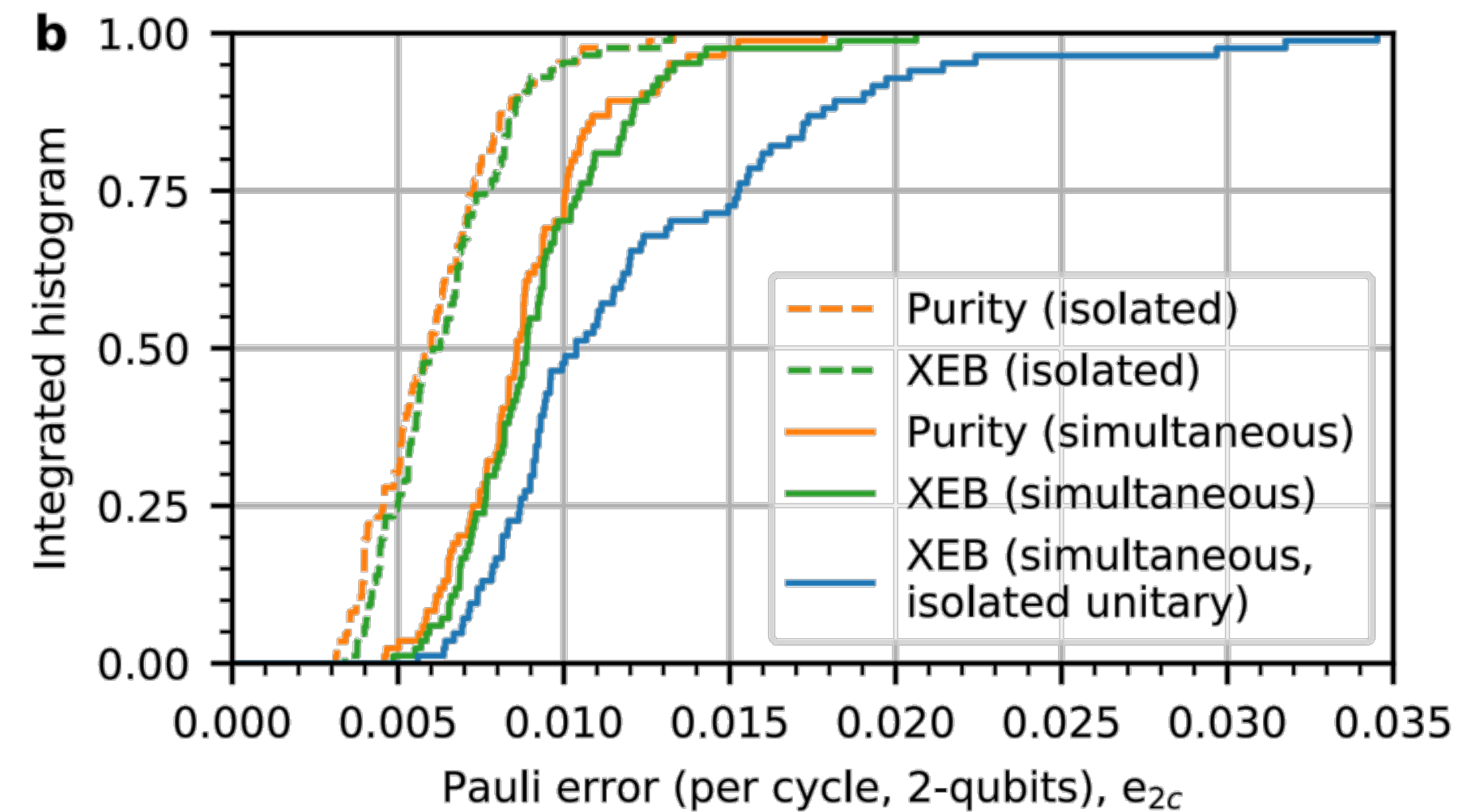
IBM's best

[Jurcevic *et al.* arXiv:2008.08571 (2020)]

Quantum Volume of 64

A 6x6 circuit has 1/3 noise rate

What's the point of a 65 qubit device?!



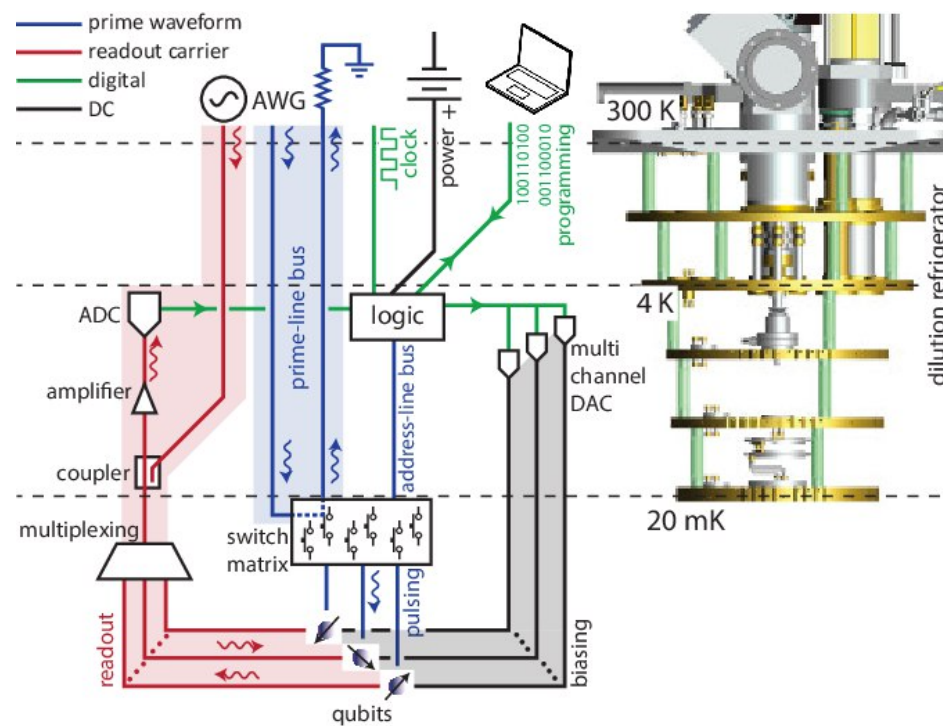
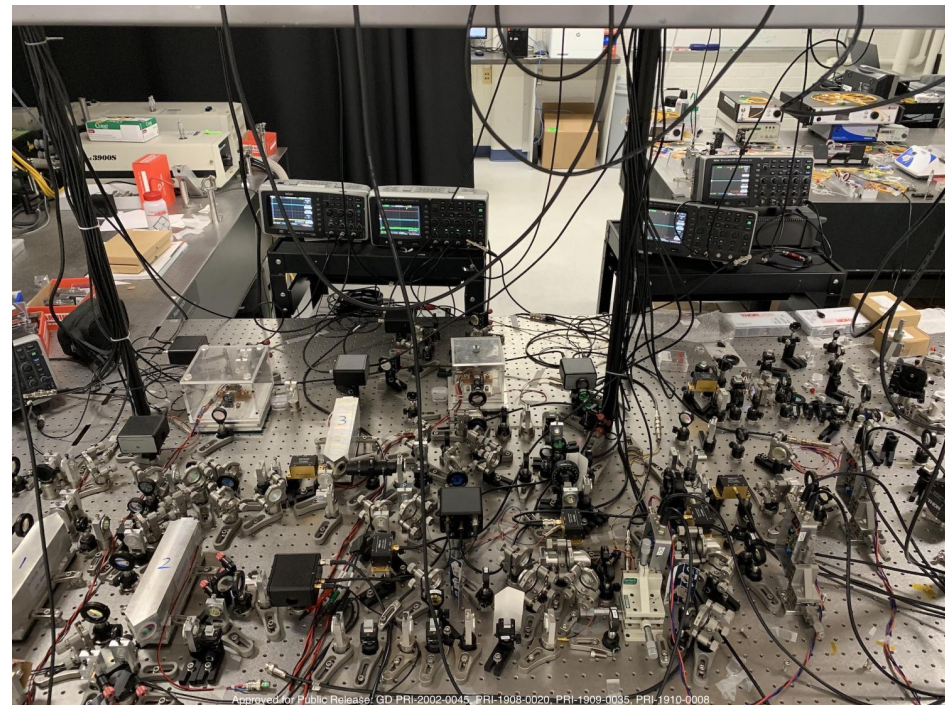
OPERATING QUANTUM COMPUTERS IS HARD

Extreme complexity



Error rates not improving

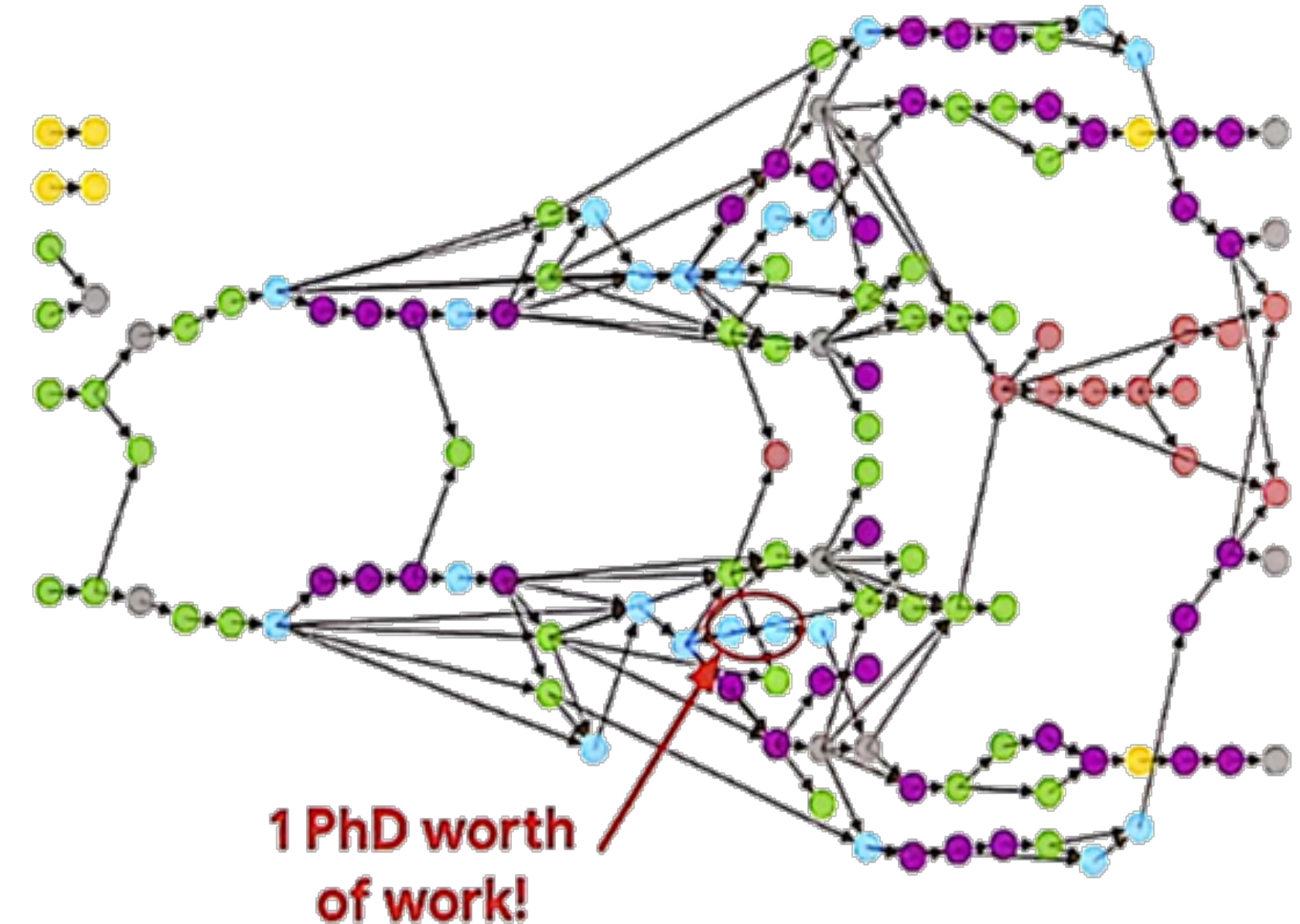
Effort unscalable



2014: 0.25%

2019: 0.3%

2022: 0.5%



AVENUES TOWARDS HIGHER FIDELITY

What can the community do?

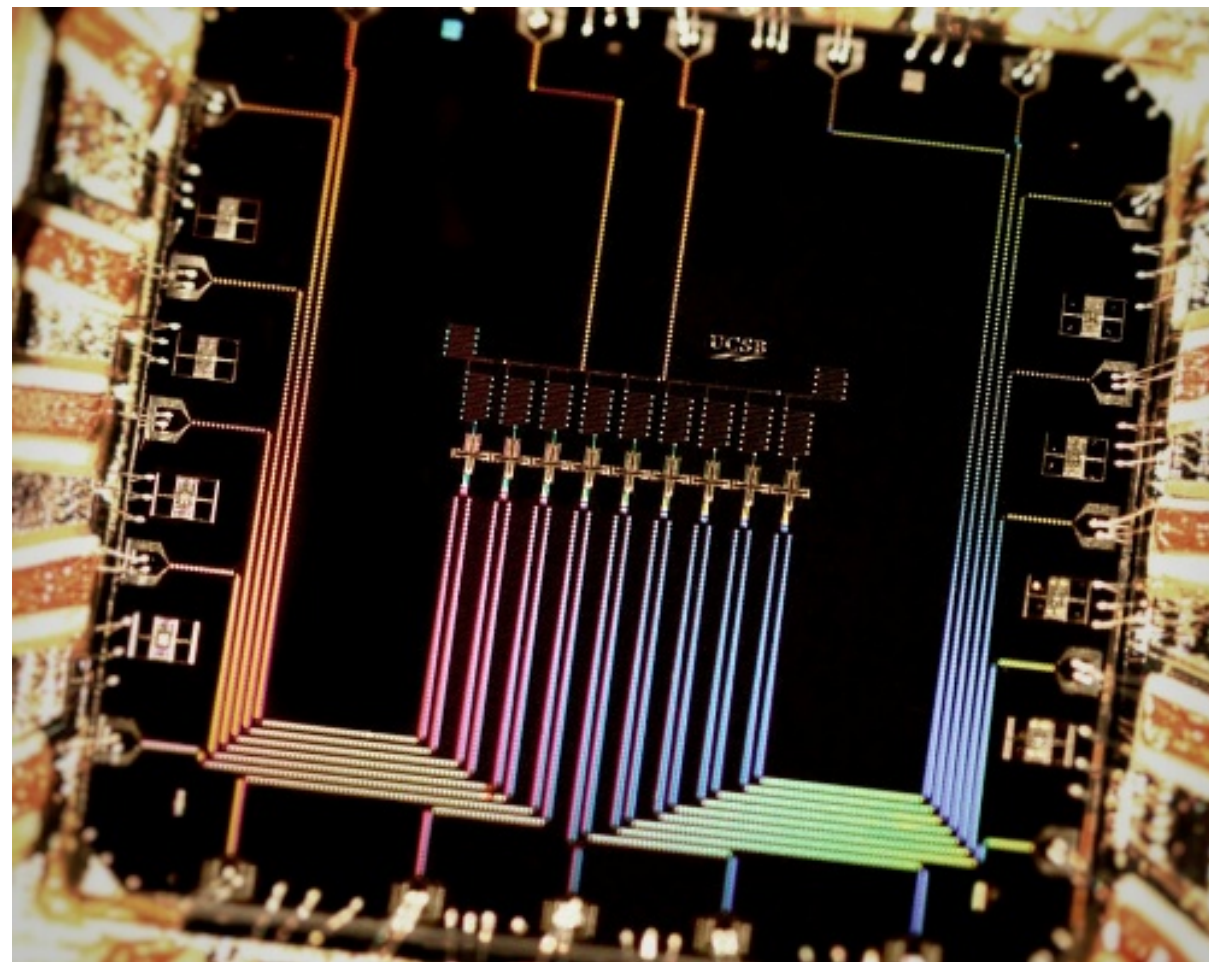
- Remove sources of decoherence - can lead to less flexibility / more complex designs
- Outrun decoherence - operate faster if coherence times remain fixed
- Avoid the impact of decoherence by smart strategies - can be found by optimization
- Remove coherent errors by better understanding and then compensating them
- Address any unknown unknowns

Applications of optimal quantum control

GOALS OF GATE DESIGN

Find pulses making gates that are good enough in the shortest time possible to beat decoherence

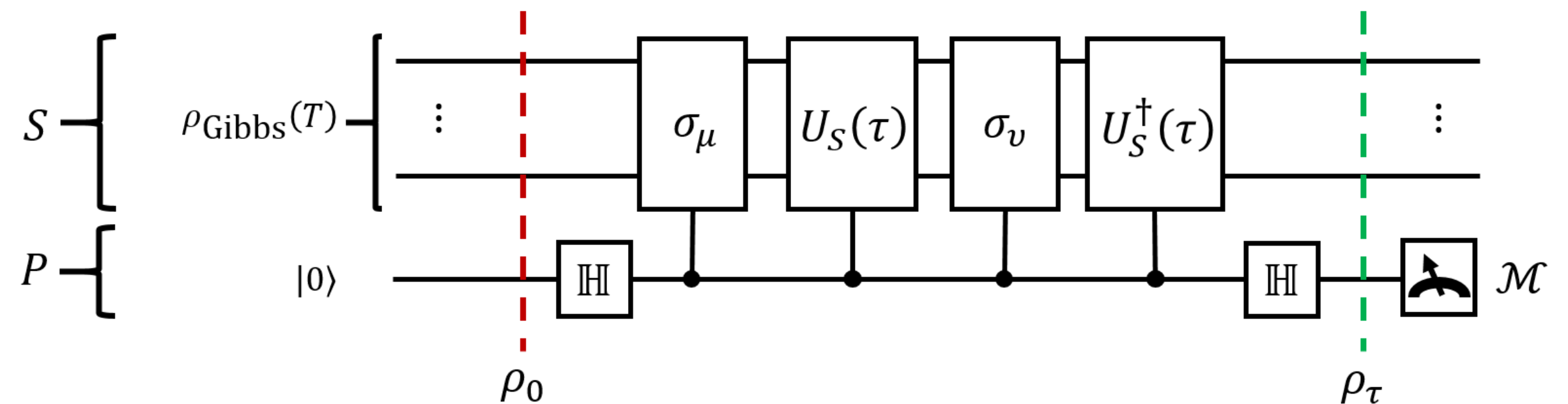
Take this



Use this



Make this



Mimimize gate error, 1-F

Make it work under realistic conditions

Use systematic mathematical methods towards this goal

CONTENT

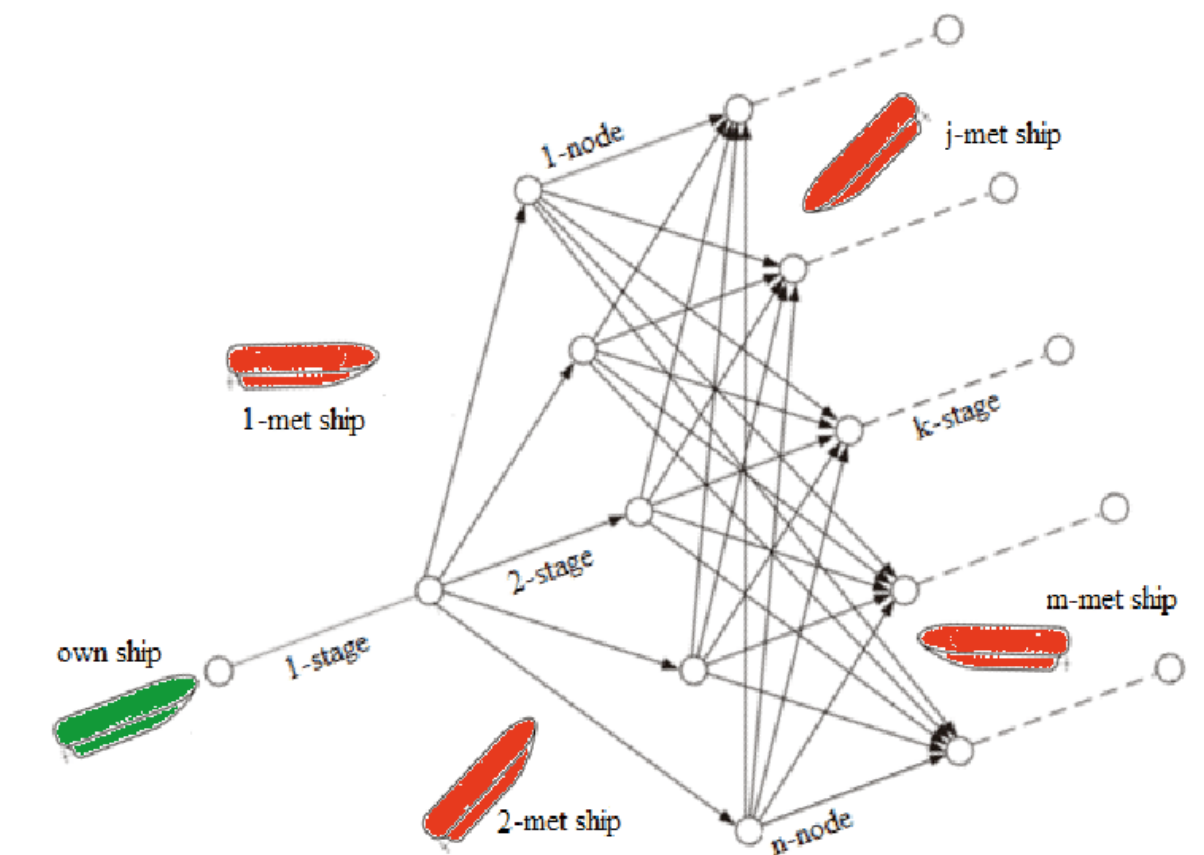
- Optimal control for classical systems
- An example for quantum control : GRAPE
- Arbitrary parameterizations for sparsity: RedCRAB and GOAT
- The need to close the loop
- Randomized Benchmarking+ Combined characterization and control
- An application: Ultrafast single-qubit gates in a transmon
- Digital twins and AI applications

OPTIMAL CONTROL FOR CLASSICAL SYSTEMS

EXAMPLES

Rockets and ships

- Optimal control historically invented for steering classical objects: Ships, rockets, planes
- e.g. Apollo mission design problem
 $\min \text{Fuel}[\vec{r}(t), \dot{\vec{r}}(t)]$ constrained to Newton's equations as running cost + trajectory from earth to the moon as boundary conditions
- Note: Most modern control works with feedback which assumes non-invasive measurement
- This lecture: Control without real-time feedback



OPTIMAL CONTROL OF A CLASSICAL SHO

Does not get simpler than that

Setting: Find (dimensionless) force $f(t)$ in the EOM $\ddot{x}(t) + \Omega^2 x(t) = f(t)$ that steers the system from $x(0), \dot{x}(0)$ to $x(T), \dot{x}(T)$

Naive solution: Use the HO Green's function $G(\tau) = \frac{\theta(\tau)}{\Omega} \sin \Omega \tau$ and use it to solve for the coordinate and velocity. This gets us

$$x(T) - x(0)\cos \Omega T - \frac{\dot{x}(0)}{\Omega} \sin \Omega T = \int_0^T dt' \frac{\sin \Omega(T - t')}{\Omega} f(t') \text{ and}$$
$$\dot{x}(T) - \dot{x}(0)\cos \Omega T + \Omega x(0)\sin \Omega T = \int_0^T dt' \cos \Omega(T - t') f(t')$$

This allows us to find $f(t)$ e.g. by Fourier analysis

OBSERVATIONS FROM THE SIMPLEST PROBLEM

- LHS of the equations describes the *drift* of the system, i.e., the dynamics without external force
- Optimal control is used to direct / correct the drift
- There are usually multiple solutions
- In practice, one would impose an energy constraint or penalty $\int_0^T dt f^2(t) \leq A$
- It turns out that if a time-optimality is reached, there are fewer solutions

VARIATIONAL CALCULUS FOR OPTIMAL CONTROL

A generalizable technique

- Suppose we have a set of dynamical equations $\dot{x} = f[x(t), u(t), t]$ for $0 \leq t \leq T$ where x are the state variables and u are the controls.
- We want to optimize a cost function at the end $J[x(T), T]$ for a trajectory following the dynamical equation
- We introduce Lagrange multipliers to enforce that constraint and thus optimize

$$\bar{J} = J[x(T), T] + \int_0^T dt \lambda^T(t) (f[x(t), u(t), t] - \dot{x})$$

- We introduce the associated Hamilton's function (aka the adjoint function)

$H[x(t), u(t), t] = \lambda(t)f[x(t), u(t), t]$ and rewrite by integration by parts

$$\bar{J} = J[x(T), T] - \lambda^T(T)x(T) + \lambda^T(0)x(0) + \int_0^T dt \left\{ H[x(t), u(t), t] + \dot{\lambda}x(t) \right\}$$

EULER-LAGRANGE EQUATIONS

$$\bar{J} = J[x(T), T] - \lambda^T(T)x(T) + \lambda^T(0)x(0) + \int_0^T dt \left\{ H[x(t), u(t), t] + \dot{\lambda}^T x(t) \right\}$$

We vary u and by this we vary x and find
$$\delta \bar{J} = \left(\frac{\partial J}{\partial x} - \lambda^T \right) \delta x \Big|_{t=T} + \lambda^T \delta x \Big|_{t=0} + \int_0^T dt \left[\left(\frac{\partial H}{\partial x} + \dot{\lambda}^T \right) \delta x + \frac{\partial H}{\partial u} \delta u \right]$$

Note that the $\delta x(0) = 0$ and that δx and δu are not independent given the EOM - which is fixed by the Lagrange multiplier

The Lagrange multiplier thus follows the end-value problem (i.e., a time-inverse initial value problem) $\dot{\lambda}^T = - \frac{\partial H}{\partial x} = - \lambda^T \frac{\partial f}{\partial x}$ and

$$\lambda^T(T) = \frac{\partial J}{\partial x(T)}$$

This is a necessary condition. If it is satisfied, we have $\delta \bar{J} = \int_0^T dt \frac{\partial H}{\partial u} \delta u$. For an extremum we need $\frac{\partial H}{\partial u} = \lambda^T \frac{\partial f}{\partial u} = 0 \quad \forall t$

These equations constitute the **Pontryagin maximum principle (PMP)**

PRACTICAL STRATEGY

Start from an initial guess for the controls

Solve $\dot{x} = f[x(t), u(t), t]$

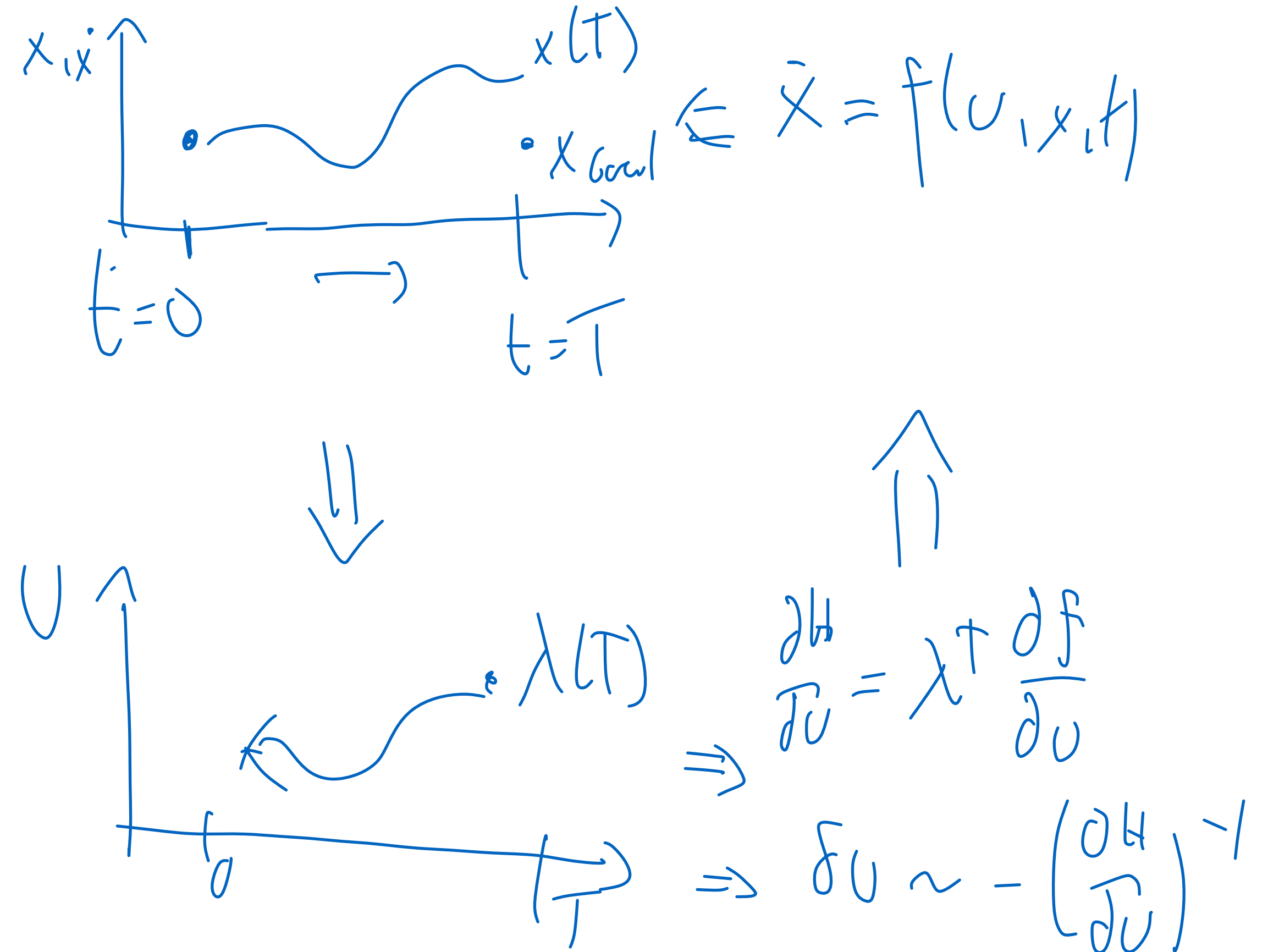
$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} = -\lambda^T \frac{\partial f}{\partial x} \text{ and } \lambda^T(T) = \frac{\partial J}{\partial x(T)}$$

Compute the gradient $\frac{\partial H}{\partial u} = \lambda^T \frac{\partial f}{\partial u} \quad \forall t$ and upgrade u

according to the gradient

Repeat until converged

Homework: Apply this to the driven harmonic oscillator going from rest at $x=0$ to being at rest at $x(T)$.



OPTIMAL CONTROL FOR THE SCHRÖDINGER EQUATION

PROBLEM SETTING

State-to-state transfer - Schrödinger equation as a dynamical system

Want to transfer $|\psi(0)\rangle \mapsto |\psi_f\rangle$ up to global phase, so $J = \left| \langle \psi_f | \psi(T) \rangle \right|^2$

Dynamical variables x : A representation of $|\psi(t)\rangle$

Dynamical equation: Schrödinger, so $\dot{f} = -i \left(H_0 + \sum_i u_i(t) H_i \right) |\psi(t)\rangle$ with WLOG a bilinear Hamiltonian

This allows us to directly apply the PMP

Challenge: Analytically calculate the gradient $\frac{\partial H}{\partial u} = \lambda^T \frac{\partial f}{\partial u} \quad \forall t$

ANALYTICAL GRADIENTS

For time-sliced pulses

Assume a piecewise constant pulse (real or good approximation)

$$U(T) = U_N U_{N-1} \cdots U_2 U_1 \text{ with } U_k = \exp \left(-i\delta t (H_0 + \sum_i u_i(j) H_i) \right)$$

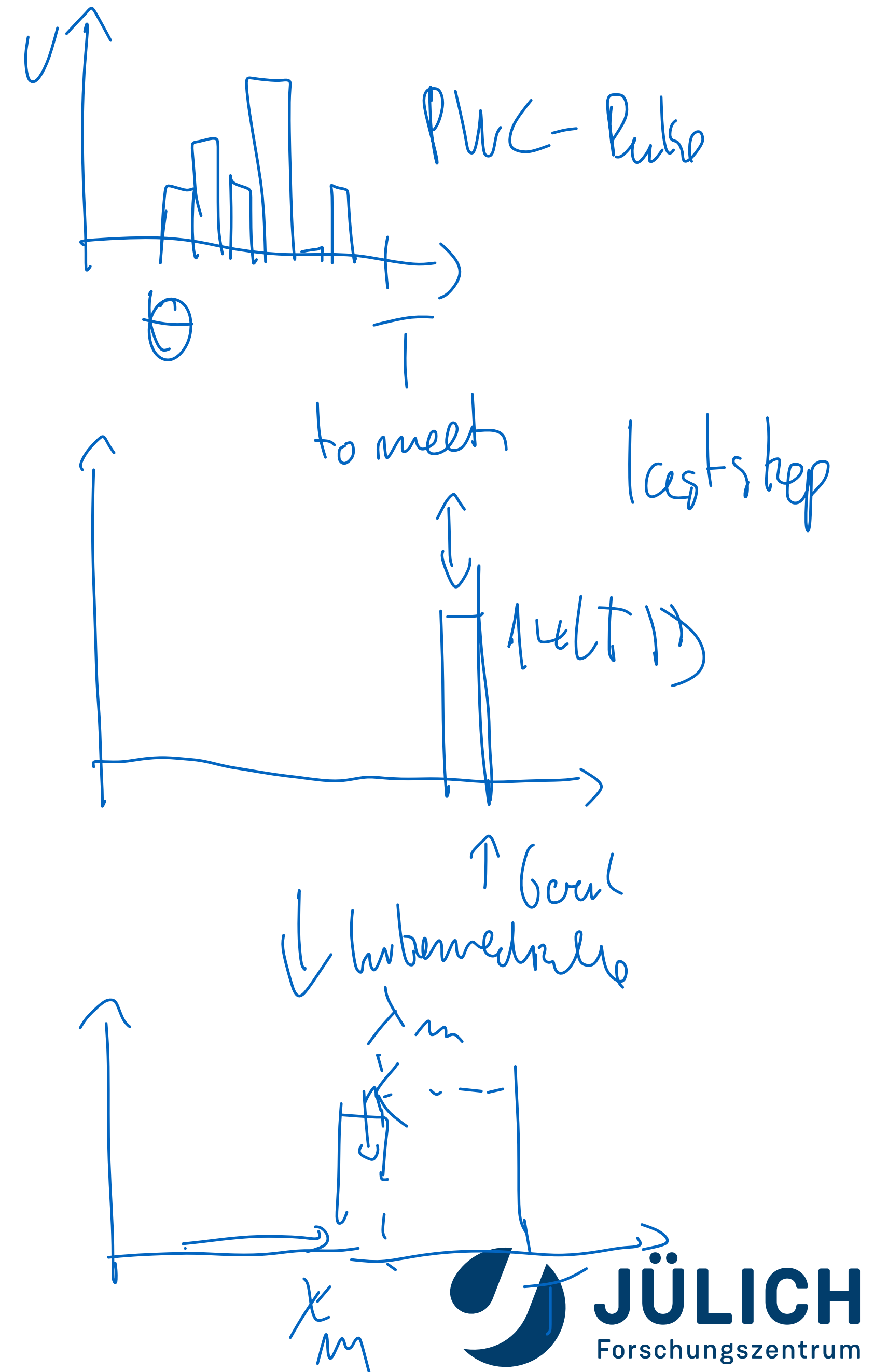
Rewrite the performance index as

$$J = \left| \langle \psi_f | U_N U_{N-1} \cdots U_2 U_1 \psi_0 \rangle \right|^2 = \left| \langle U_{m+1}^\dagger \cdots U_N^\dagger \psi_f | U_m \cdots U_1 \psi_0 \rangle \right|^2 \equiv \left| \langle \lambda_m | \chi_m \rangle \right|^2$$

with the propagated initial state $|\chi_m\rangle$ and the back-propagated target $|\lambda_m\rangle$

Now for small enough time step we can show

$$\frac{\partial J}{\partial u_i(j)} = -i\delta t \langle \lambda_j | H_i | \rho_j \rangle - \text{voilà, analytical gradient}$$



RECIPE

Start from an initial guess of the controls

Compute the propagated initial states $|\rho_m\rangle$ and the back-propagated target states $|\rho_m\rangle$ by iterative matrix multiplication

Compute the gradient of the the performance index und update the performance index

Repeat until converged

Note: Can also be done for gates, then the performance index is

$$\| U_f - U(T) \|_2^2 = \text{Tr} \left[\left(U_f^\dagger - U^\dagger(T) \right) \left(U_f - U(T) \right) \right] = 2d - 2\text{Tr} U_f^\dagger U(T) \dots \text{made phase-}$$

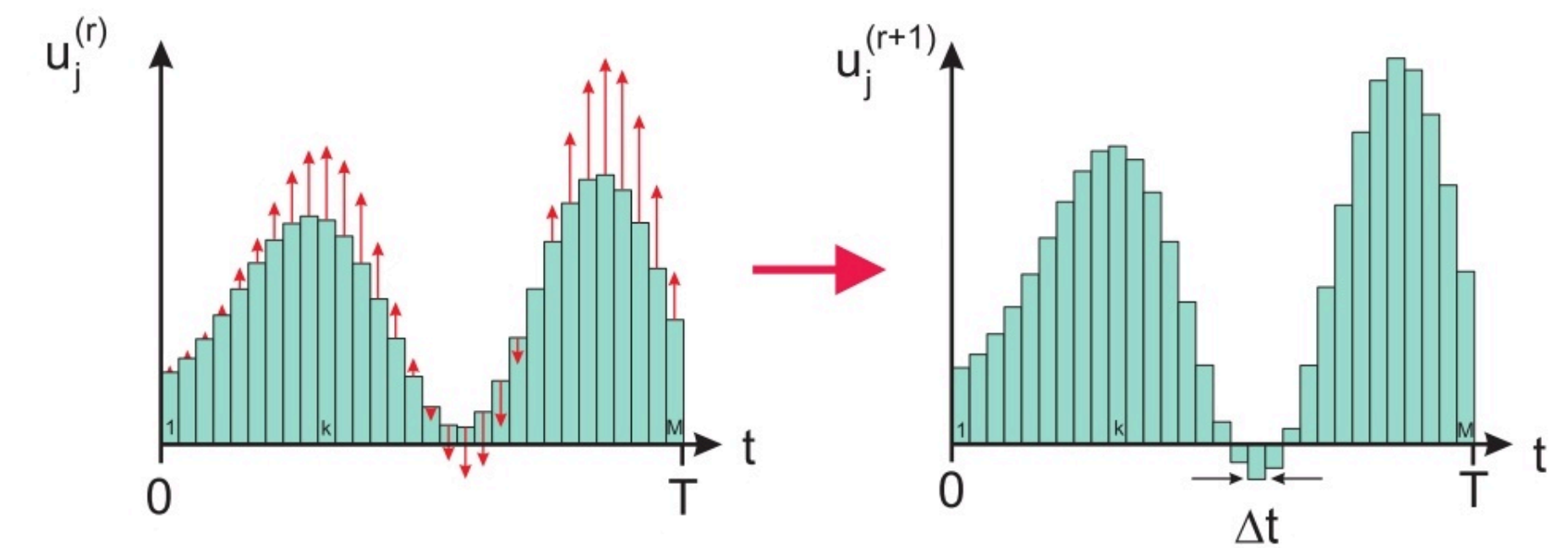
insensitive my maximising $\left| \text{Tr} U_f^\dagger U(T) \right|$

UPDATE STRATEGIES

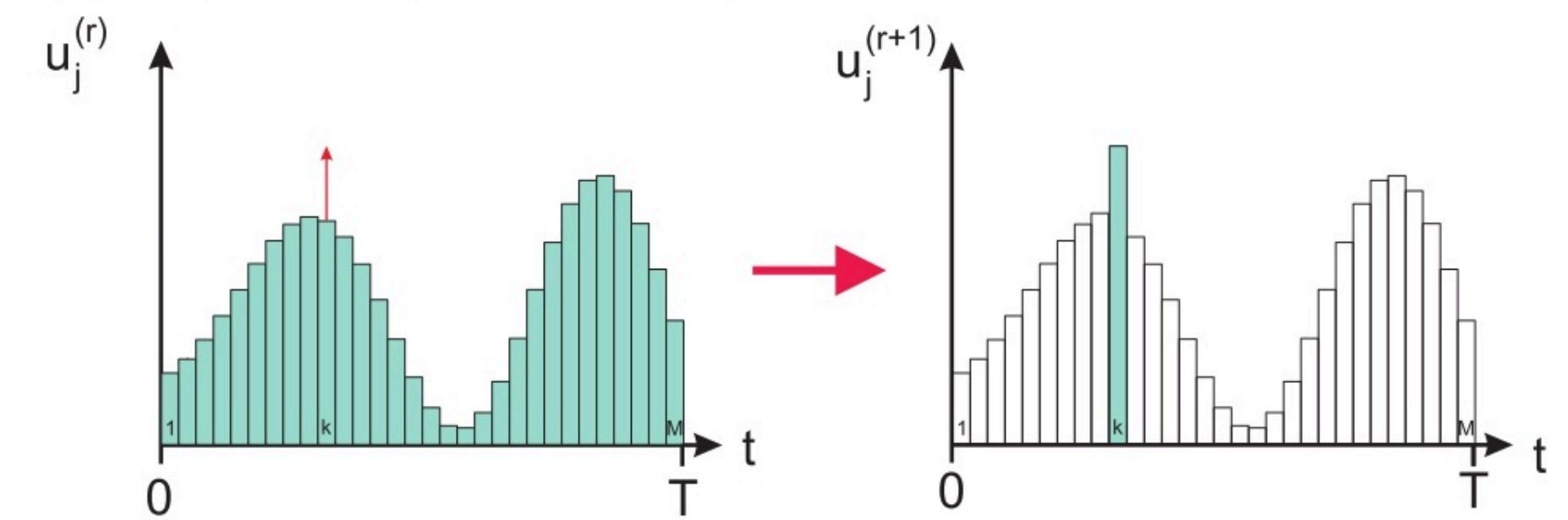
Krotov is monotonically convergent, but the steps take longer

All techniques can be boosted by L-BFGS, exact gradients and many other tricks ...

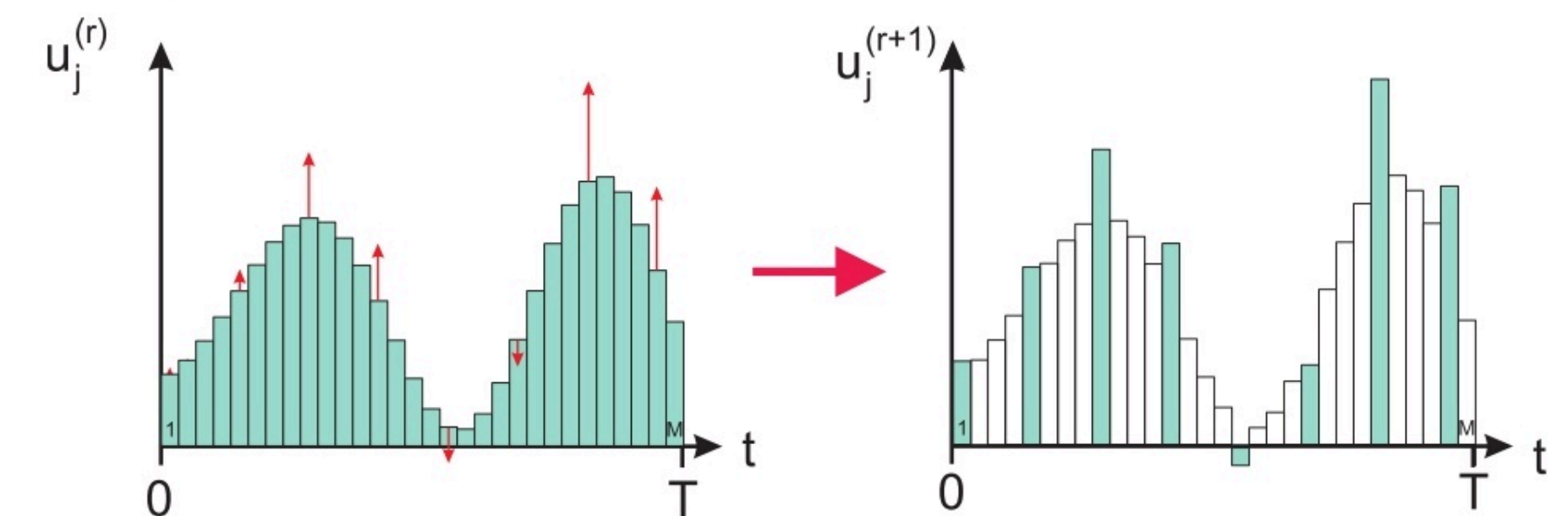
(a) concurrent (GRAPE-type)



(b) sequential (Krotov-type)



(c) hybrid



APPLIED PERSPECTIVE

CONTROL FOR NONLINEARITY AND SPECTRAL CROWDING

EPL, **123** (2018) 60001
doi: 10.1209/0295-5075/123/60001

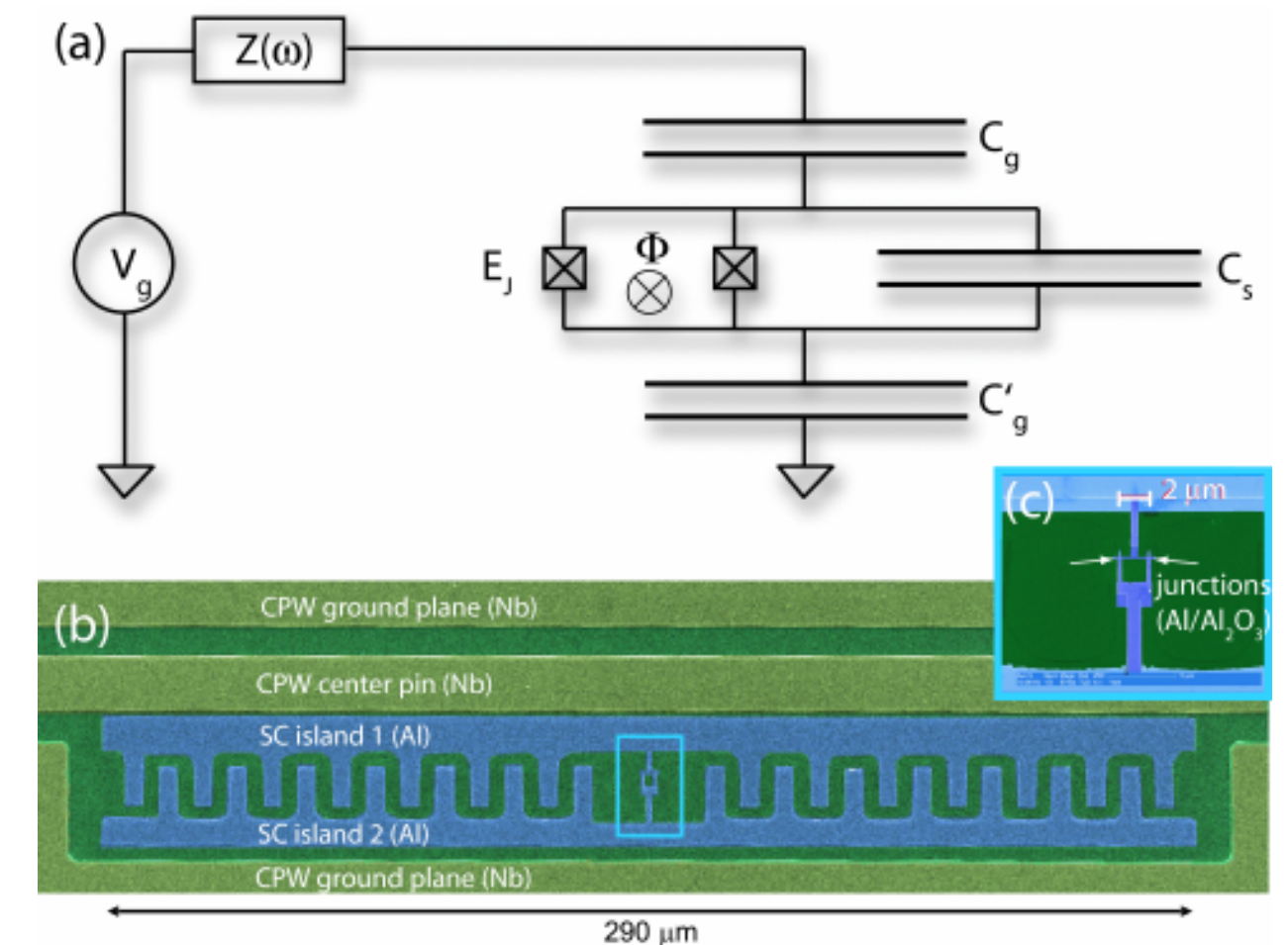
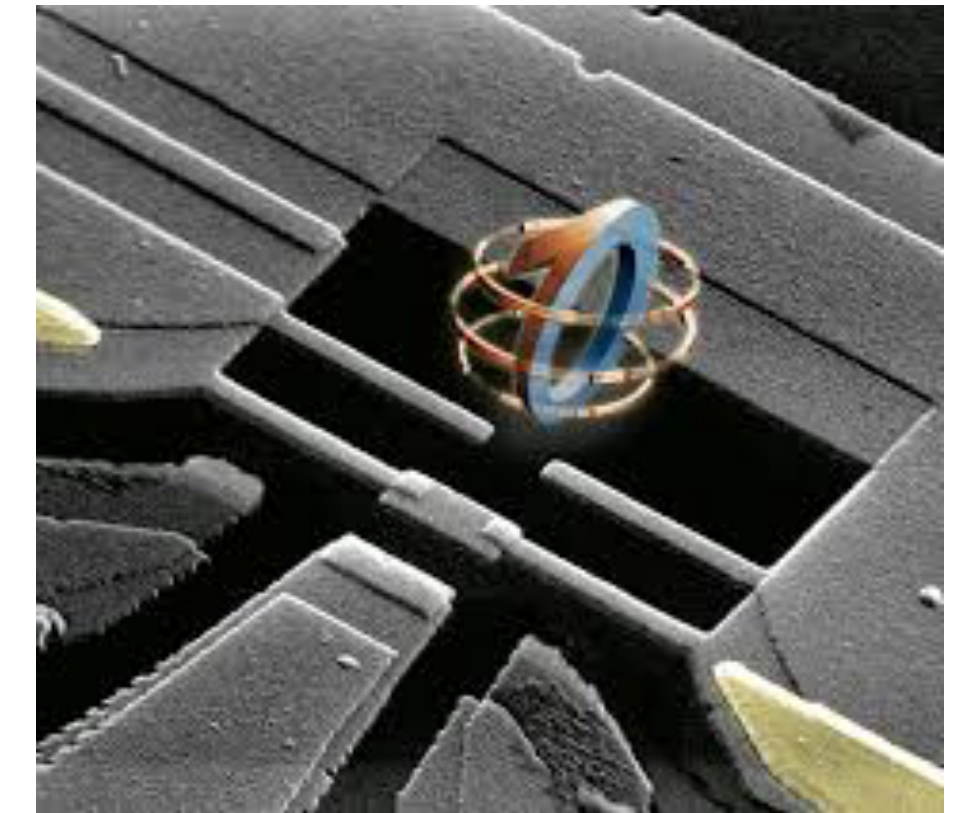
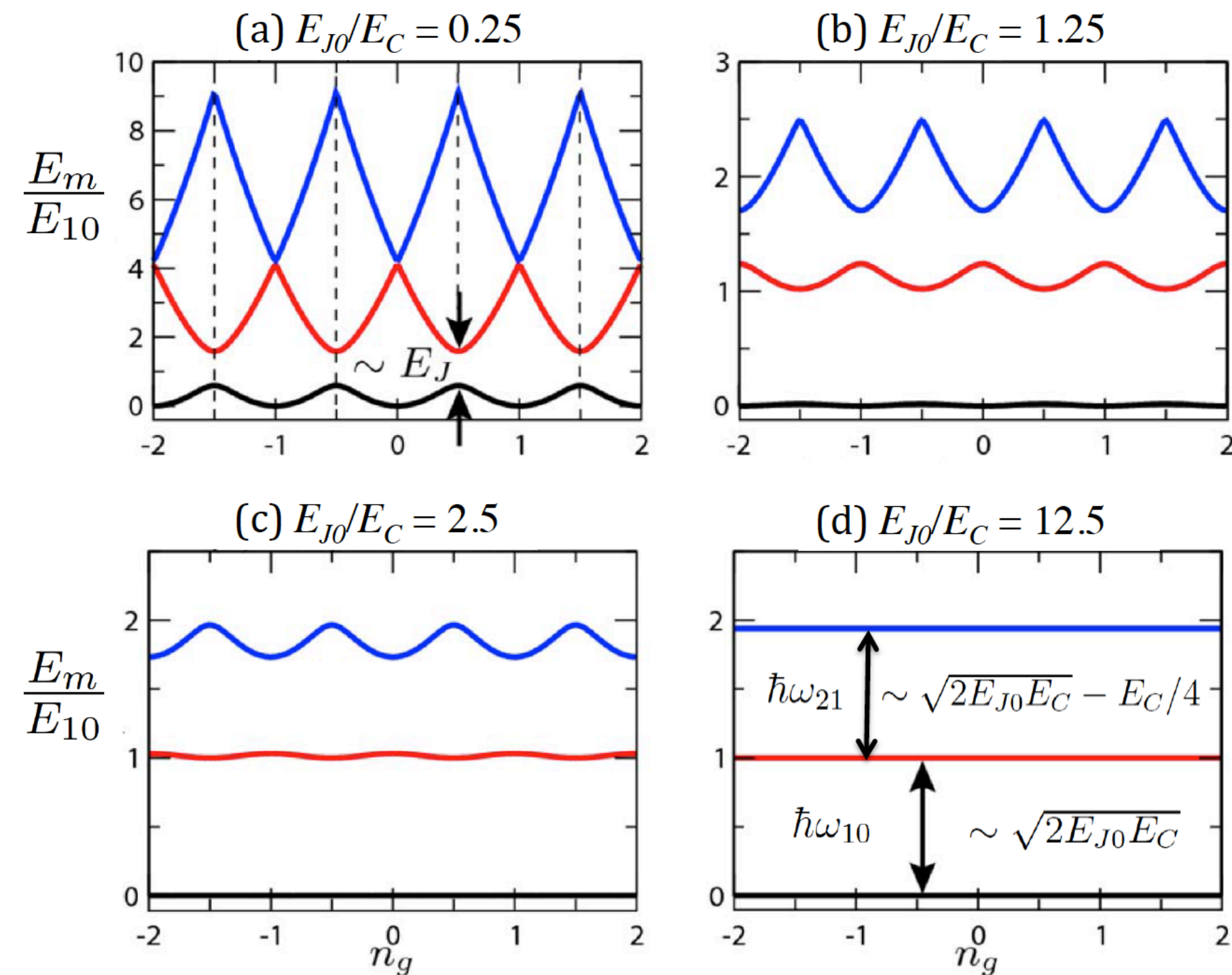
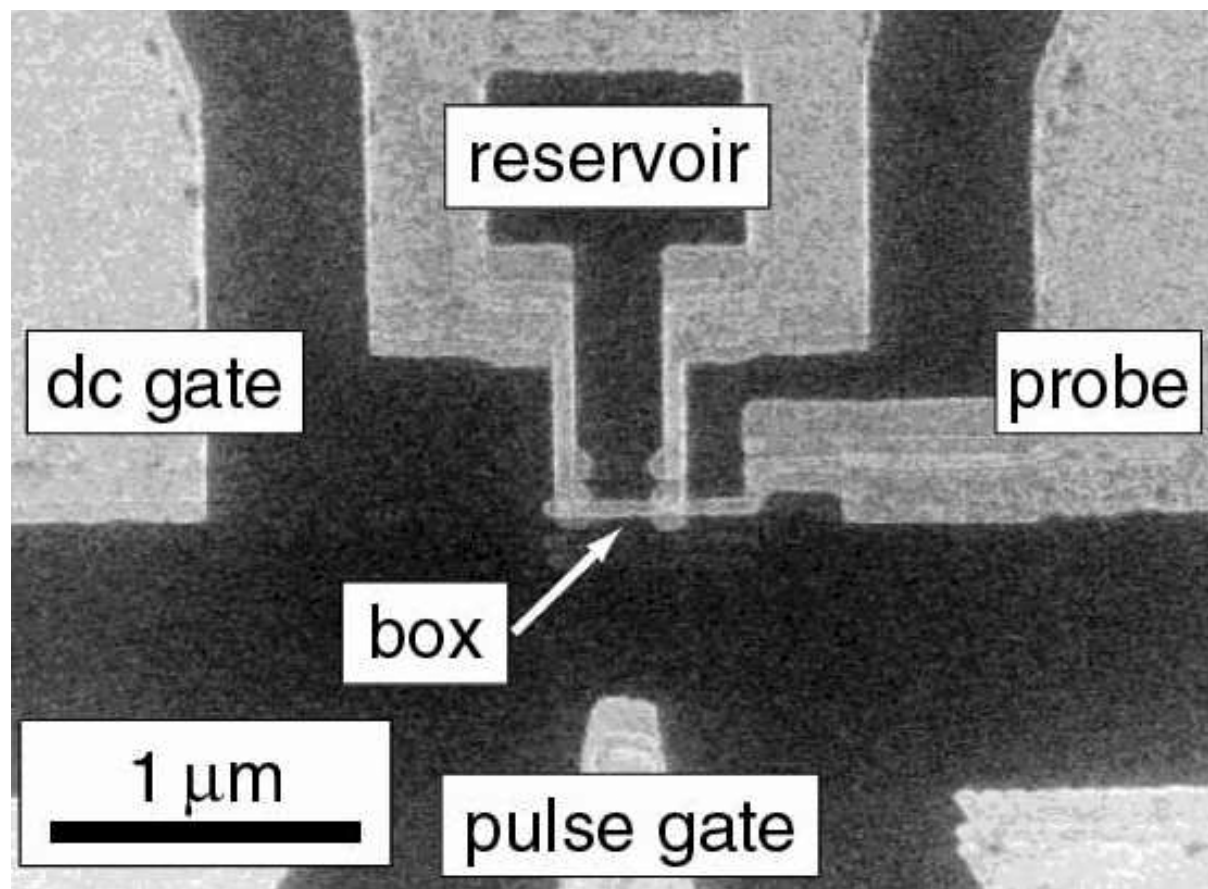
www.epljournal.org

Focus Article

Counteracting systems of diabaticities using DRAG controls: The status after 10 years^(a)

L. S. THEIS¹, F. MOTZOI^{2(b)}, S. MACHNES¹ and F. K. WILHELM¹

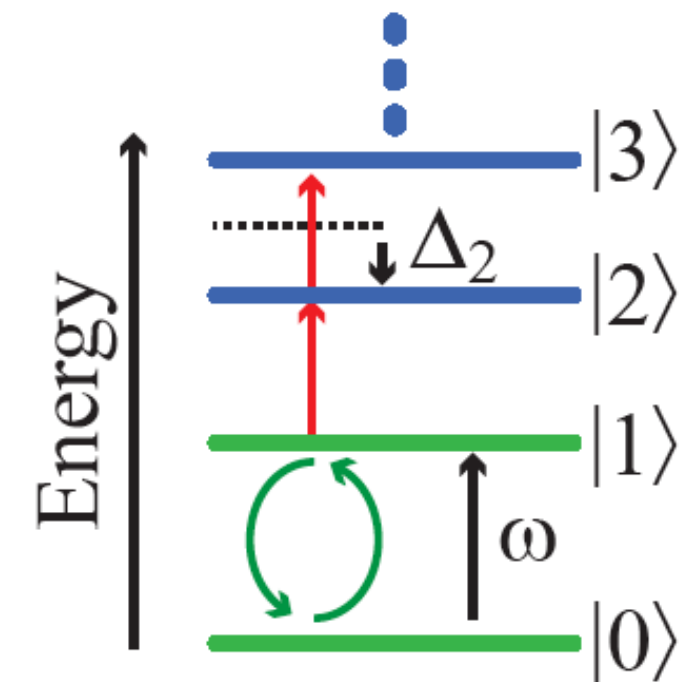
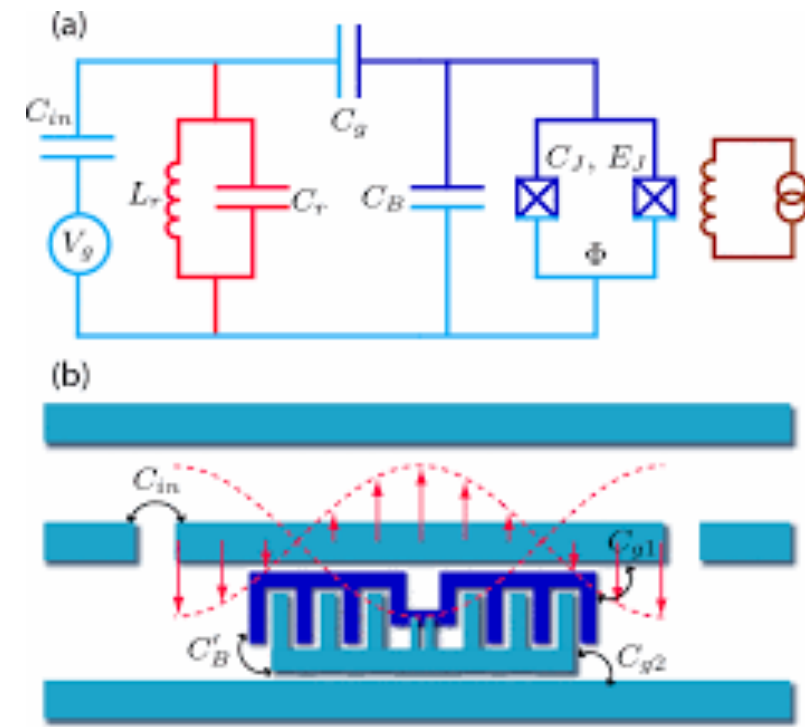
EVOLUTION OF NONLINEARITIES



Bigger qubits have better temporal coherence,
but get closer to the (semiclassical) HO

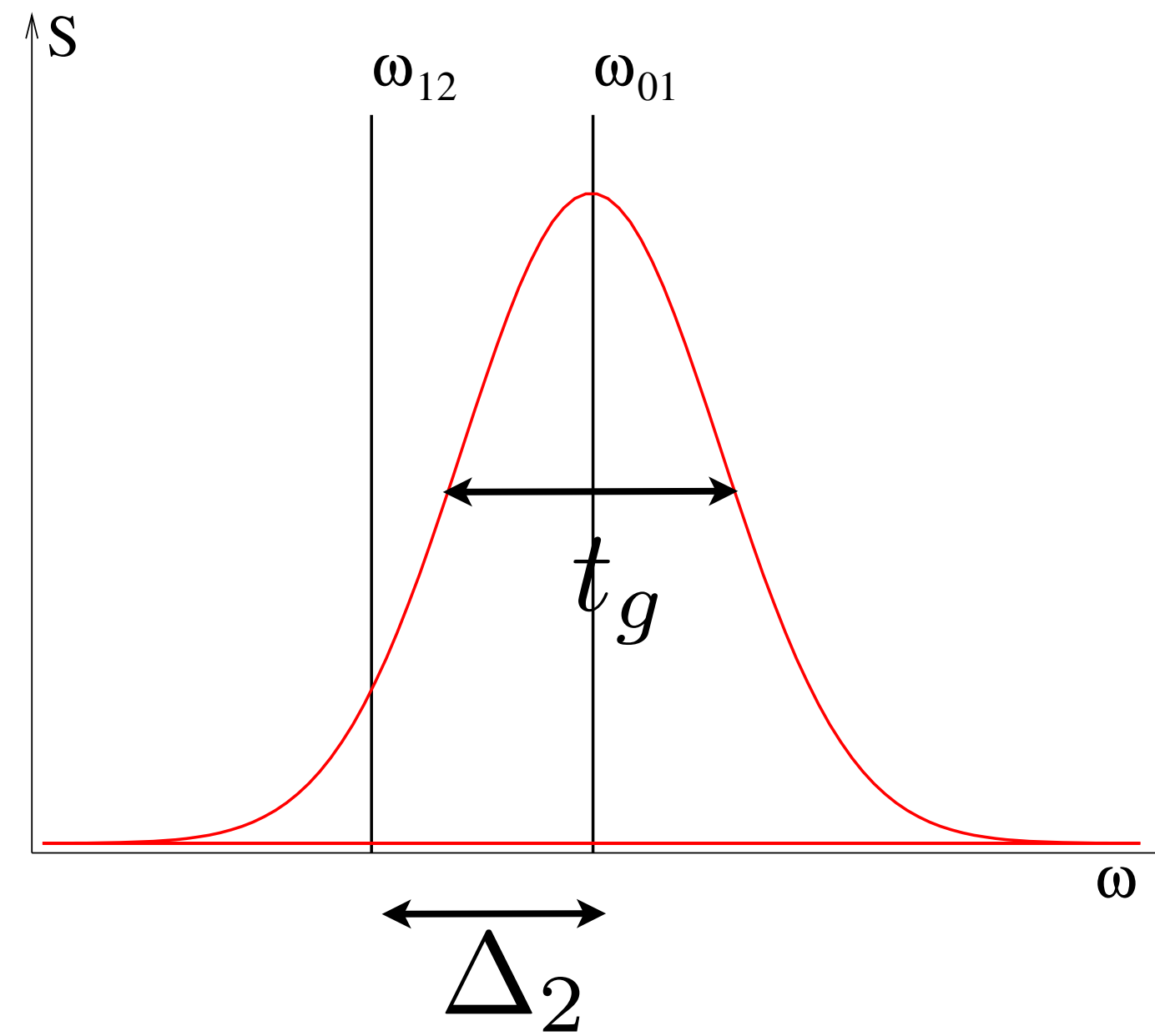
Mitglied der Helmholtz-Gemeinschaft

DRAW: GETTING MOST OUT OF NONLINEARITY

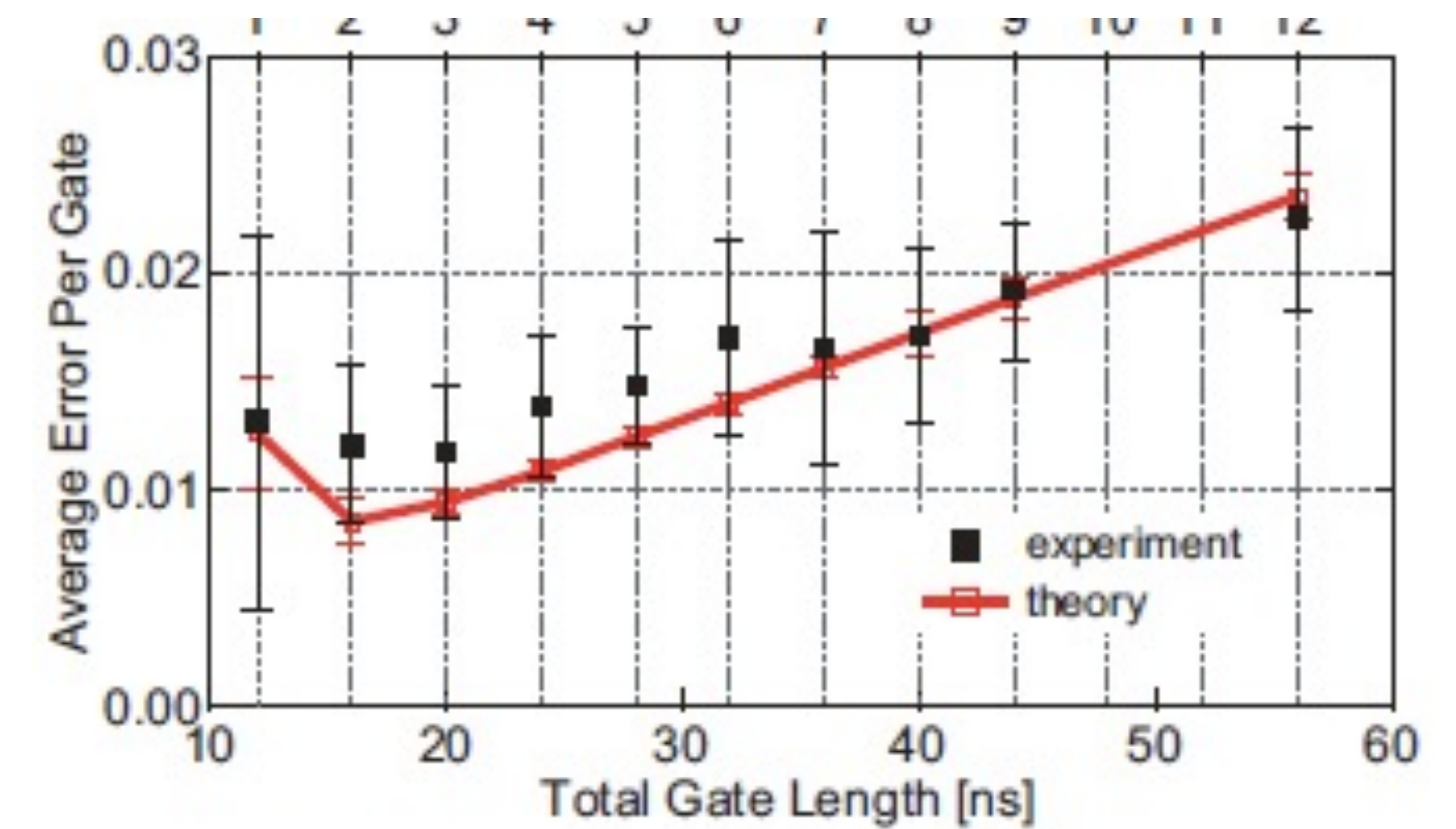
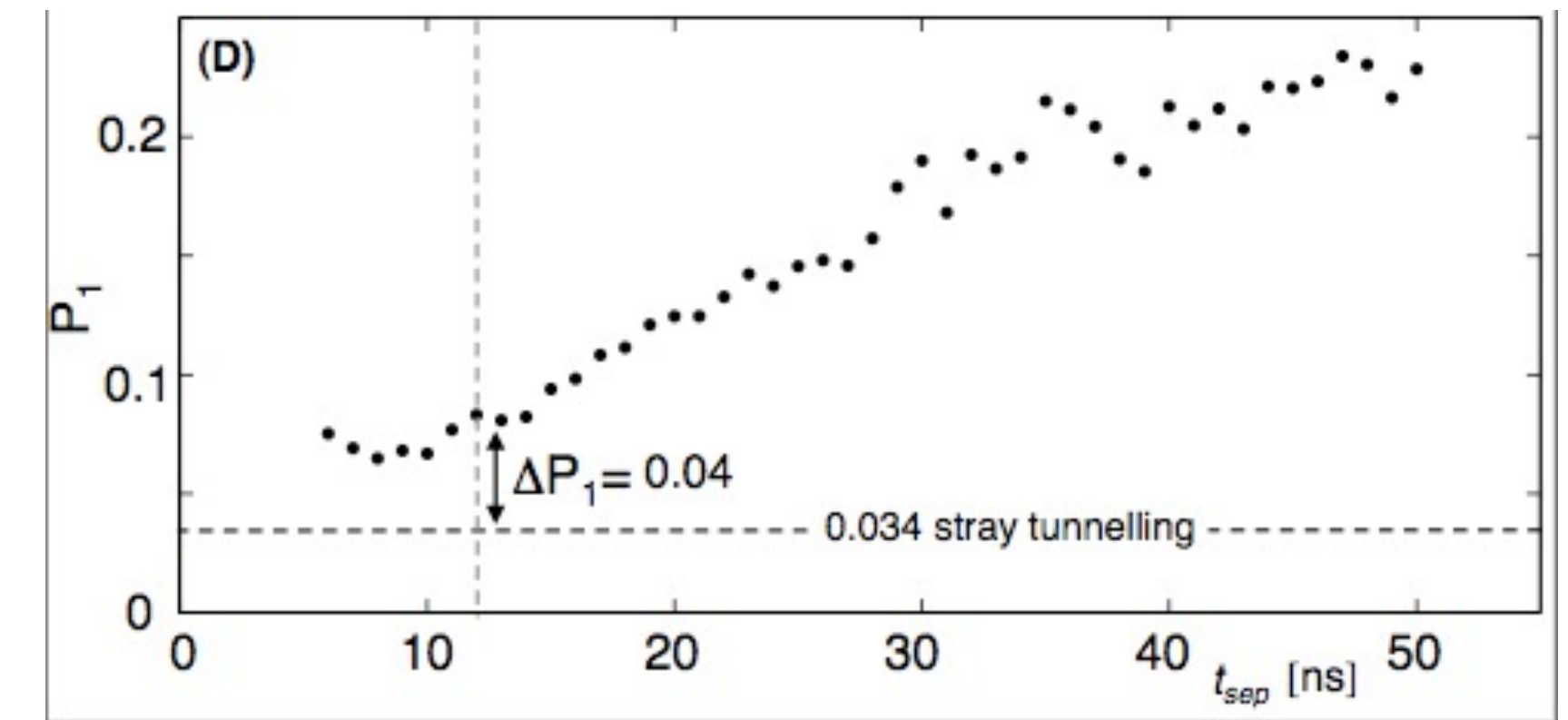


Transmon
= weakly nonlinear

Spectral limitation:
Duration/bandwidth uncertainty



Heisenberg relation of control



Crossover: Coherence vs bandwidth limit

WHY DOES IT WORK?

Iterative frame cancellation:

Pi-pulse on the $|0\rangle$ to $|1\rangle$ transition + adiabatically returning in the $|1\rangle$ to $|2\rangle$ transition

However: Going through the pulse (as a frame change) leads to a counterdiabatic force proportional to the pulse derivative

Apply another force to correct it: Needs to go from 0 to 0 so the frames match in the beginning and the end

Wait, we introduce a new force: repeat
(Drawing on the board)

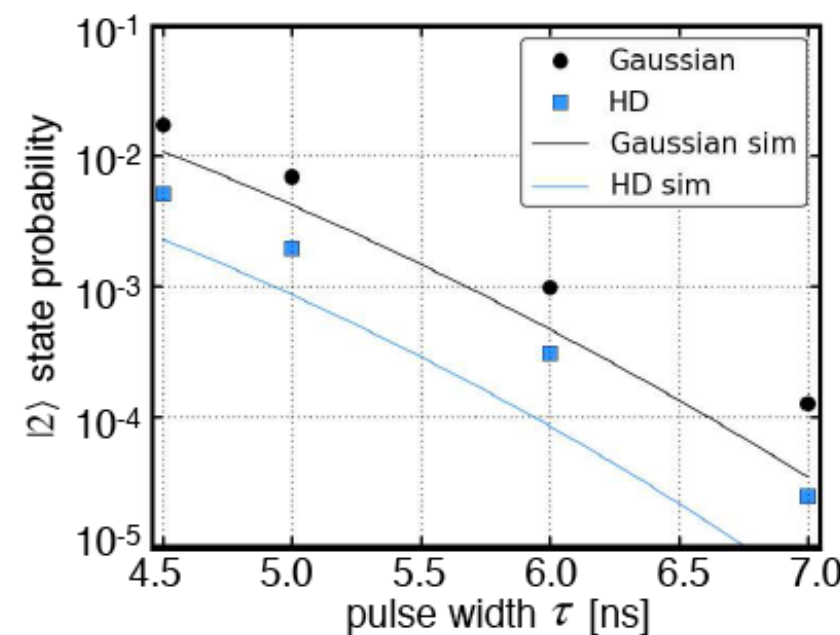
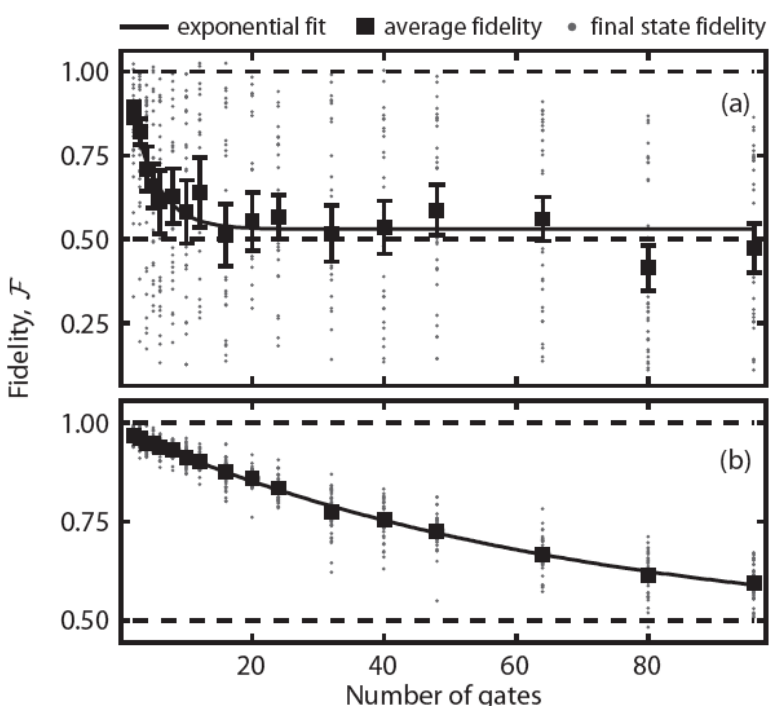
DRAG, WAHWAH AND FRIENDS

DRAG

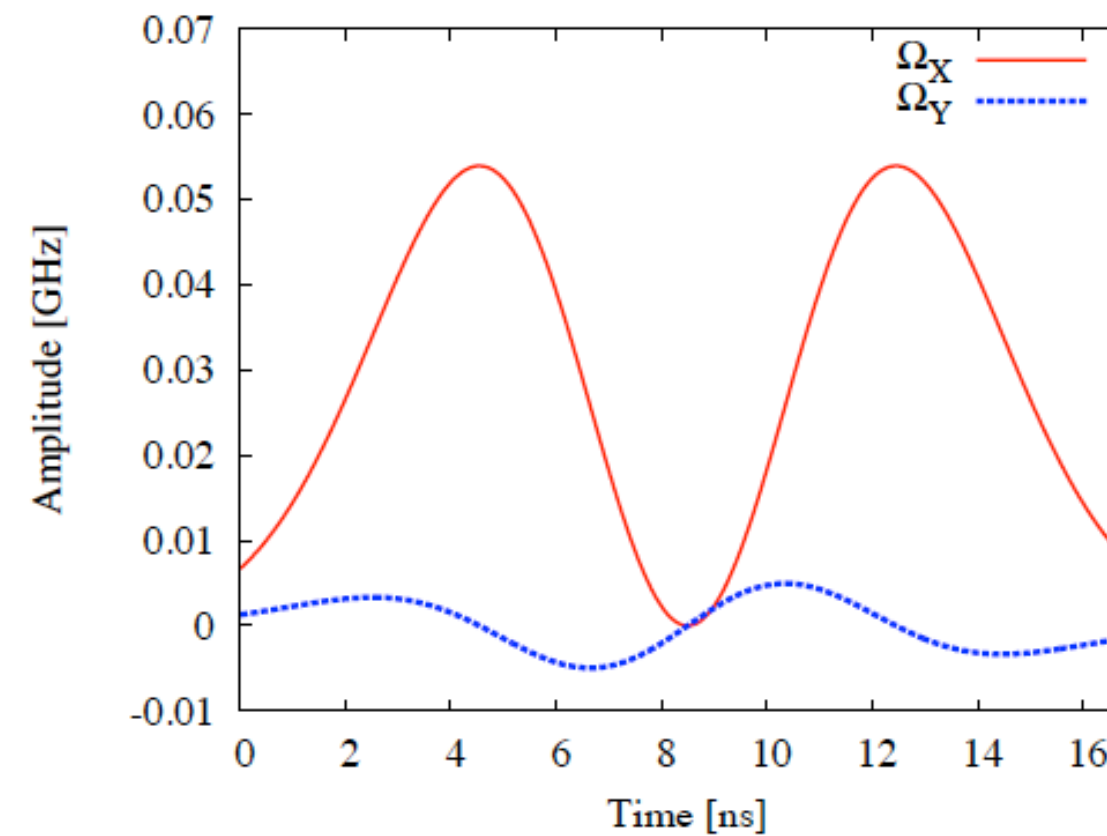
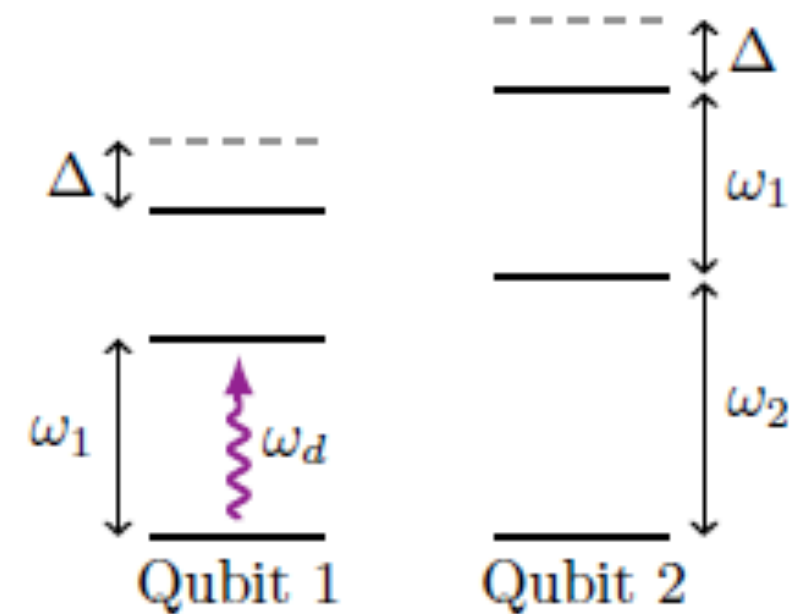
$$u_1(t)\cos\omega t + u_2(t)\sin\omega t$$

$$u_2 = \frac{\dot{u}_1}{\Delta_2}$$

Simple parameterization of numerical result:
Implementable pulse

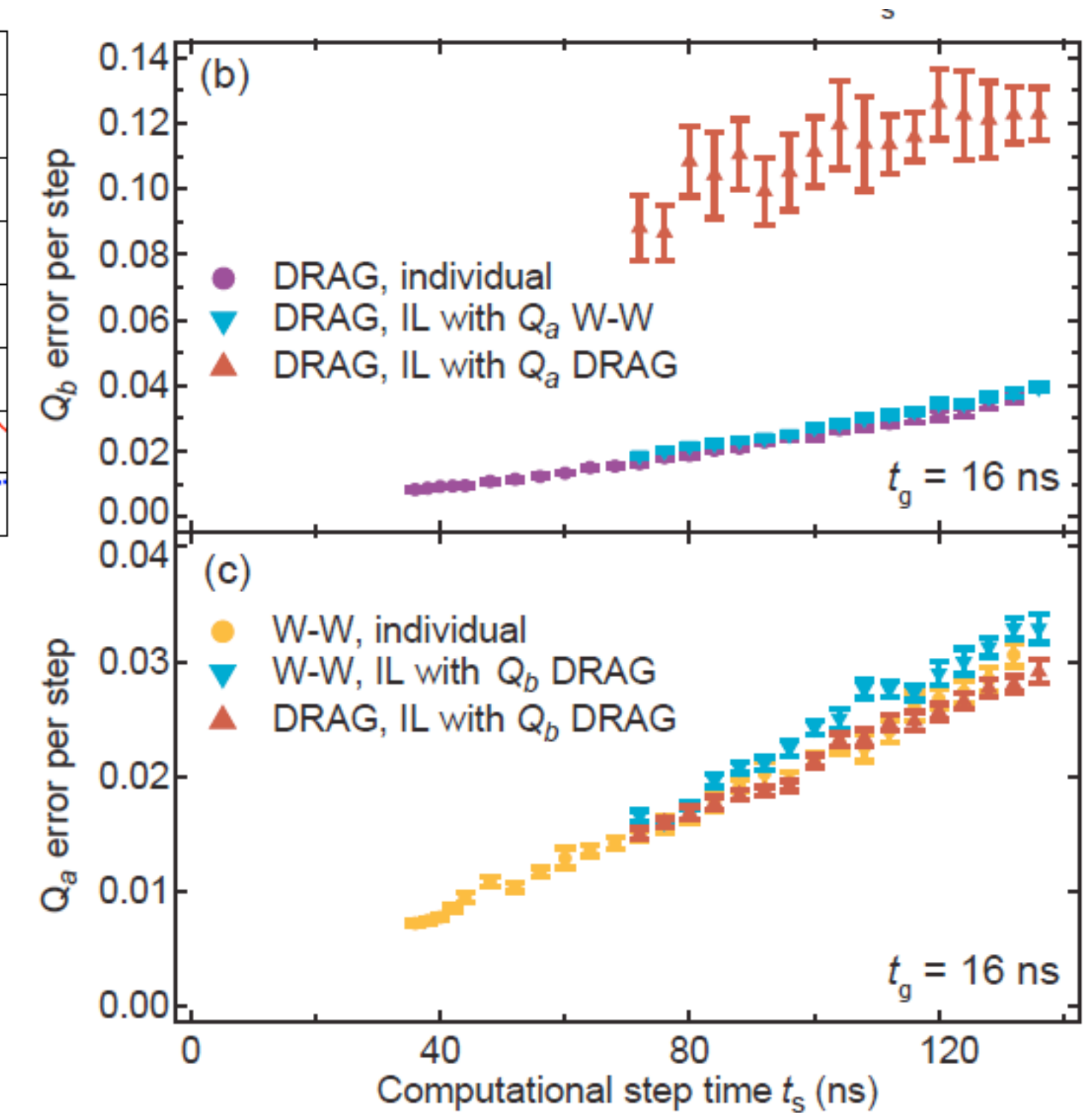


WAHWAH



$$\Omega_x(t) = A_\pi e^{-\frac{18}{t_g^2}(t-\frac{t_g}{2})^2} \cdot \overbrace{\left\{1 - A \cos\left[\frac{\delta}{2}\left(t - \frac{t_g}{2}\right)\right]\right\}}^{\text{Removes } |1\rangle_2 \leftrightarrow |2\rangle_2}$$

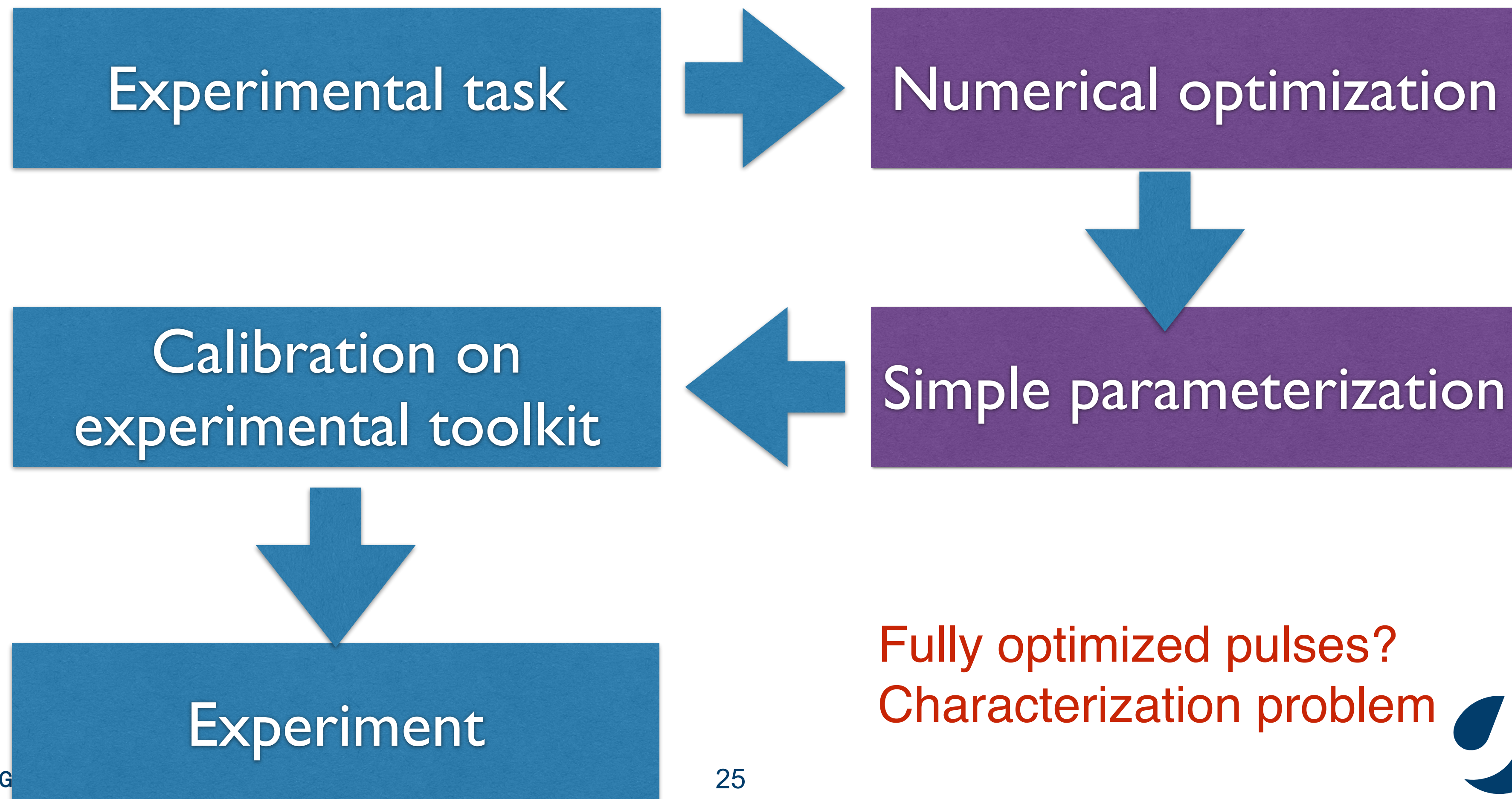
$$\Omega_y(t) = -\frac{1}{2\Delta}\dot{\Omega}_x(t)$$



Weak AnHarmonicity With Average Hamiltonian = WAHWAH

Last year's Nature is this year's subroutine (J. Martinis)

FEW-PARAMETER WORKFLOW



Fully optimized pulses?
Characterization problem

BASIC OPTIMAL CONTROL

Goal: For N qubits, generate $SU(2^N)$

Standard choice: Single qubit rotations+ perfect entangler - go figure how to make them

„Go figure“ is a bad idea, think about the right technique

More systematic: $\hat{H} = \hat{H}_0 + \sum_i u_i(t) \hat{H}_i$ H_0 : Drift, u_i : Control fields,
 H_i : Control Hamiltonians

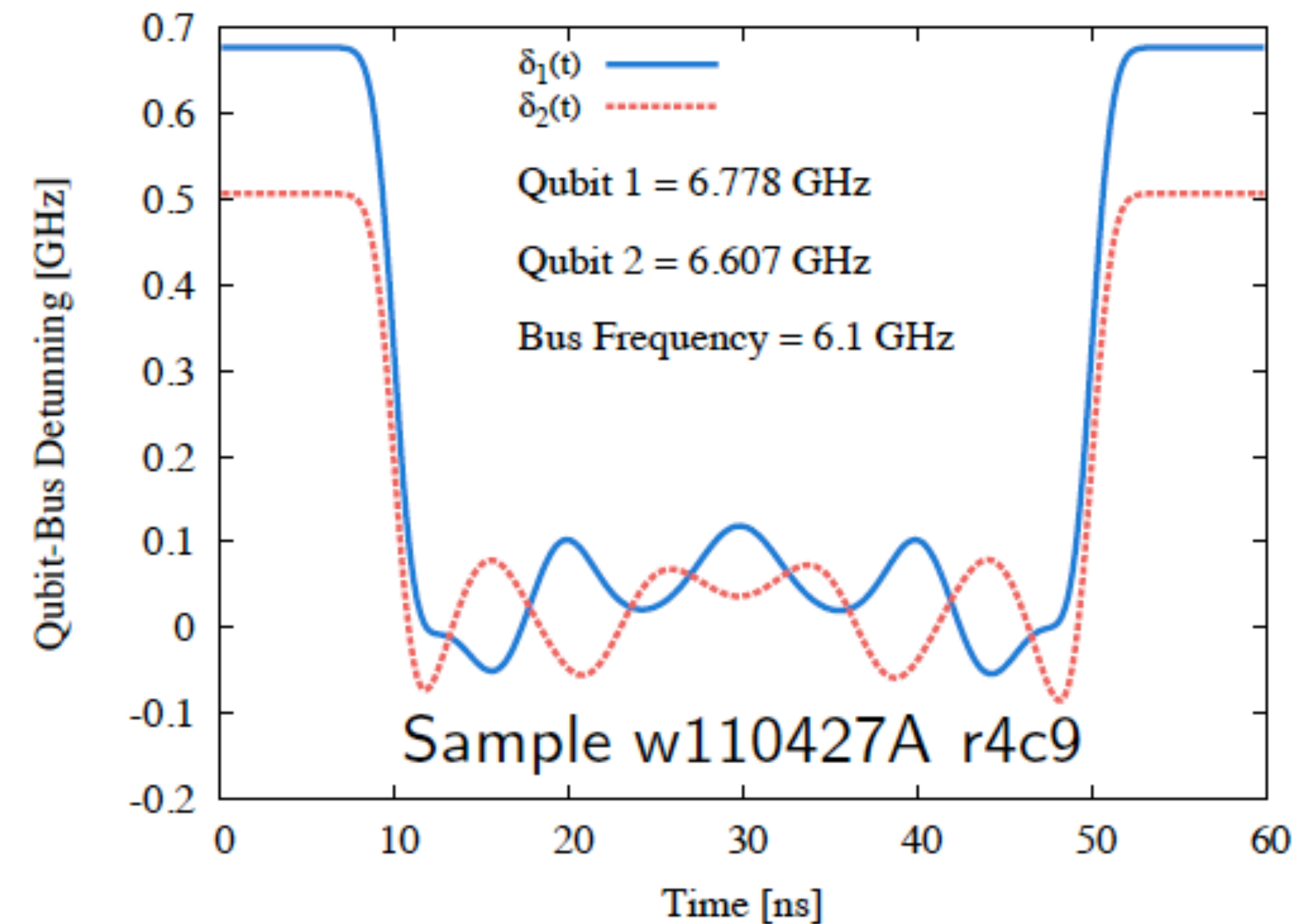
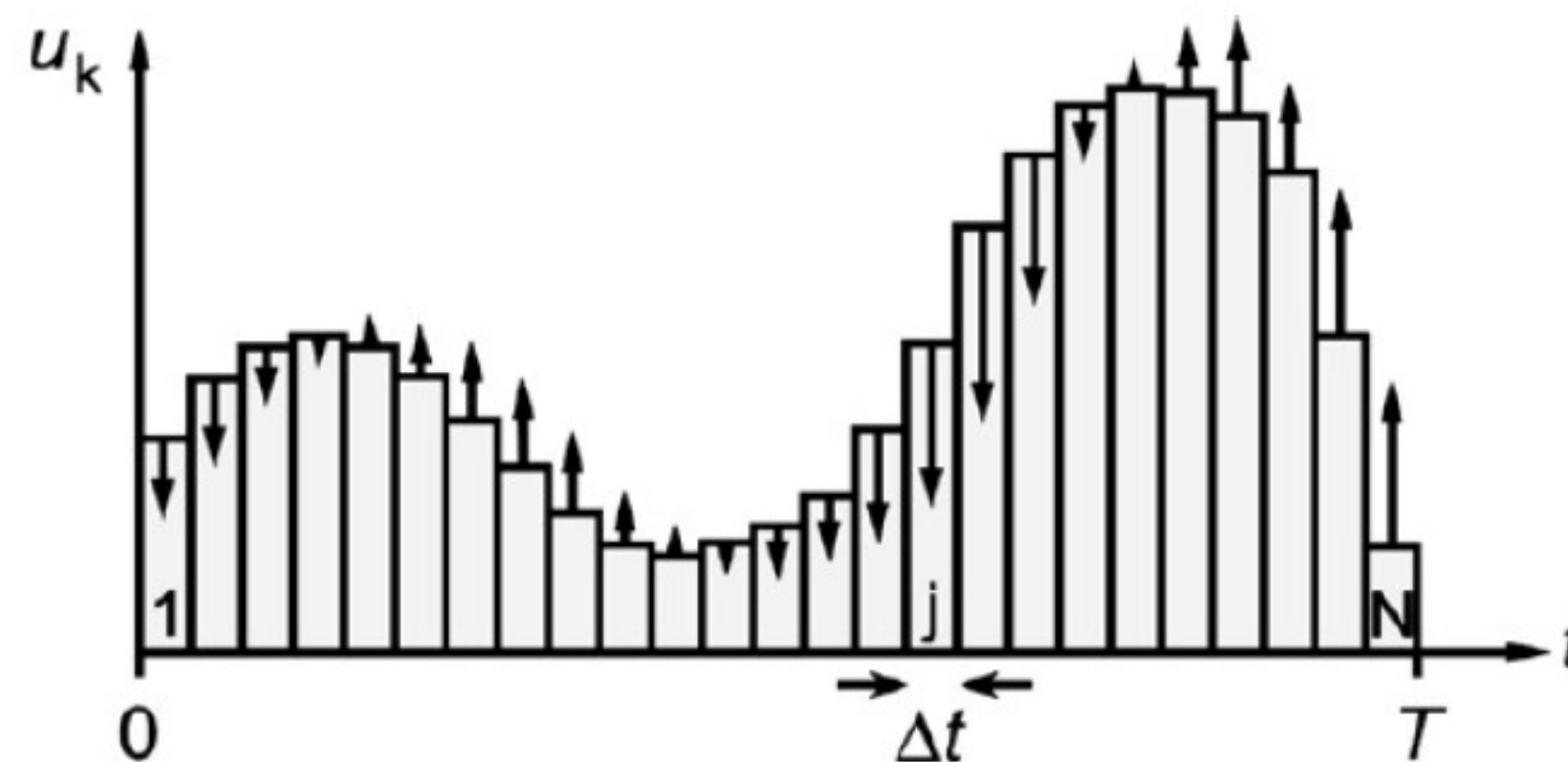
System fully controllable if Lie closure: $\langle H_i | i = 0 \dots N \rangle = \mathfrak{su}(2^N)$

Find $u_i(t)$ to reach

with search based on analytical gradients

$$U_{\text{gate}} = U(t,0) = \mathbb{T} \exp \left(-i \int_0^t d\tau H(\tau) \right)$$

Find controls that maximize fidelity



How to debug something complex, non-intuitive?

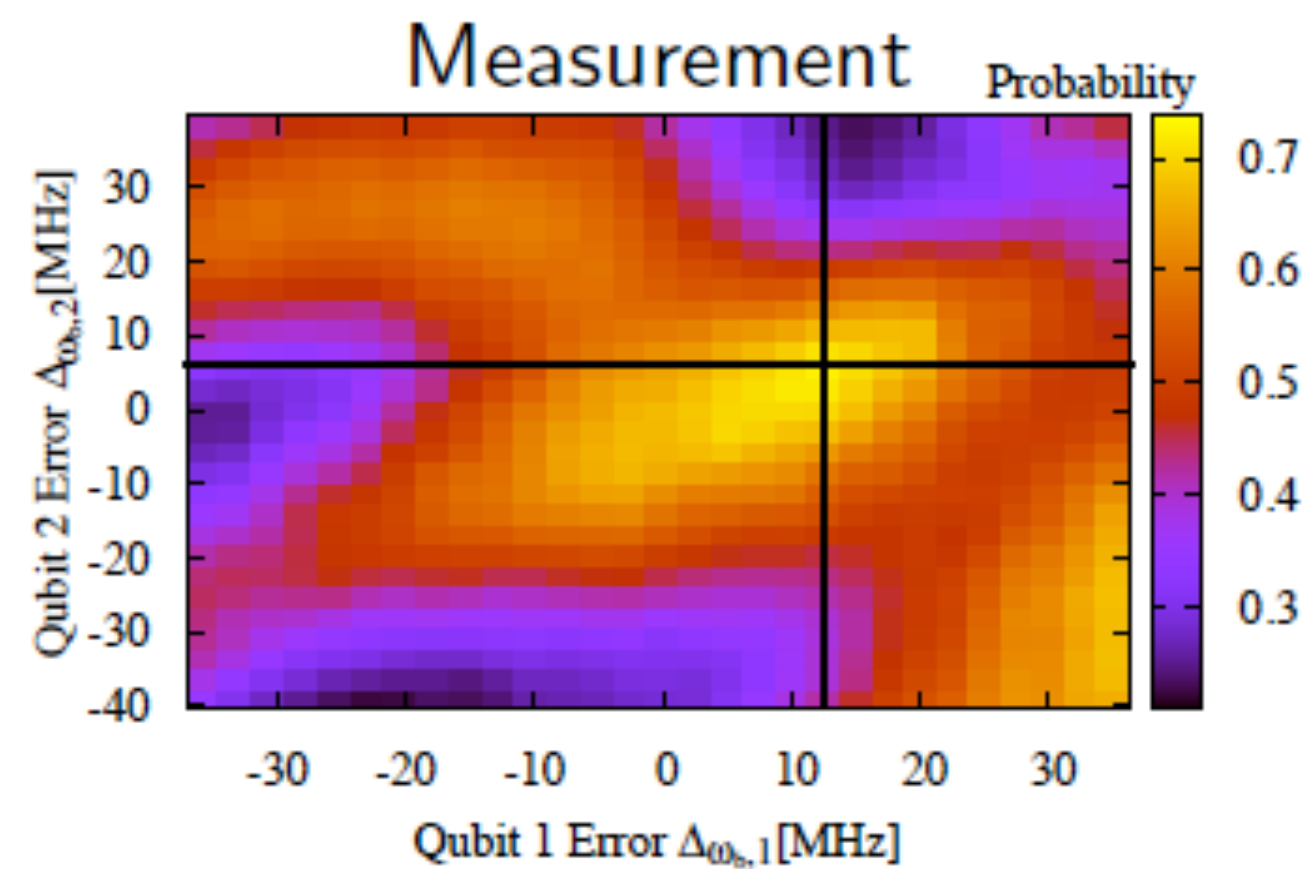
ERROR LANDSCAPE

Extremum is a flat point

Strong curvature in the landscape leads to strong sensitivity to error

hard to debug due to multitude of parameters

.... but at least theory matches experiment



TUNEUP CHALLENGE

Fabrication uncertainty

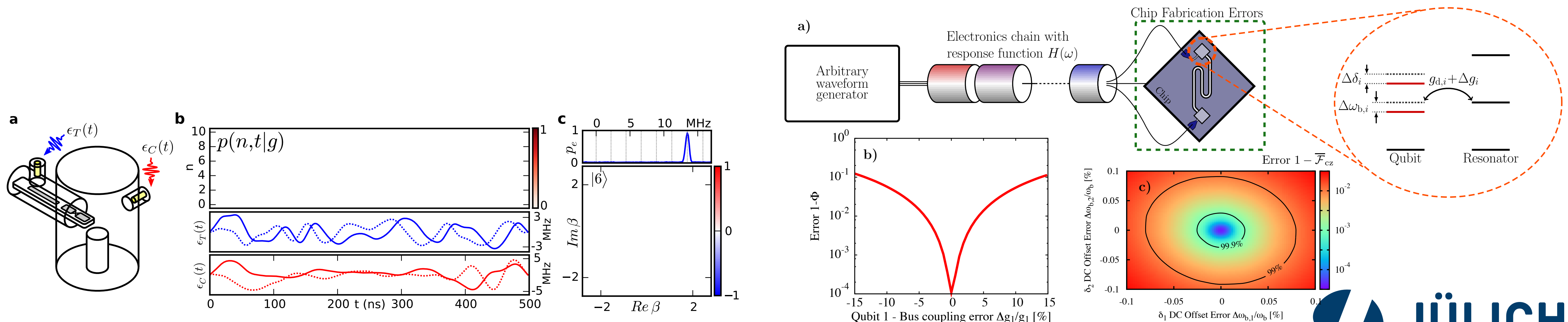
Transfer function uncertainty

Best detector: The qubit itself

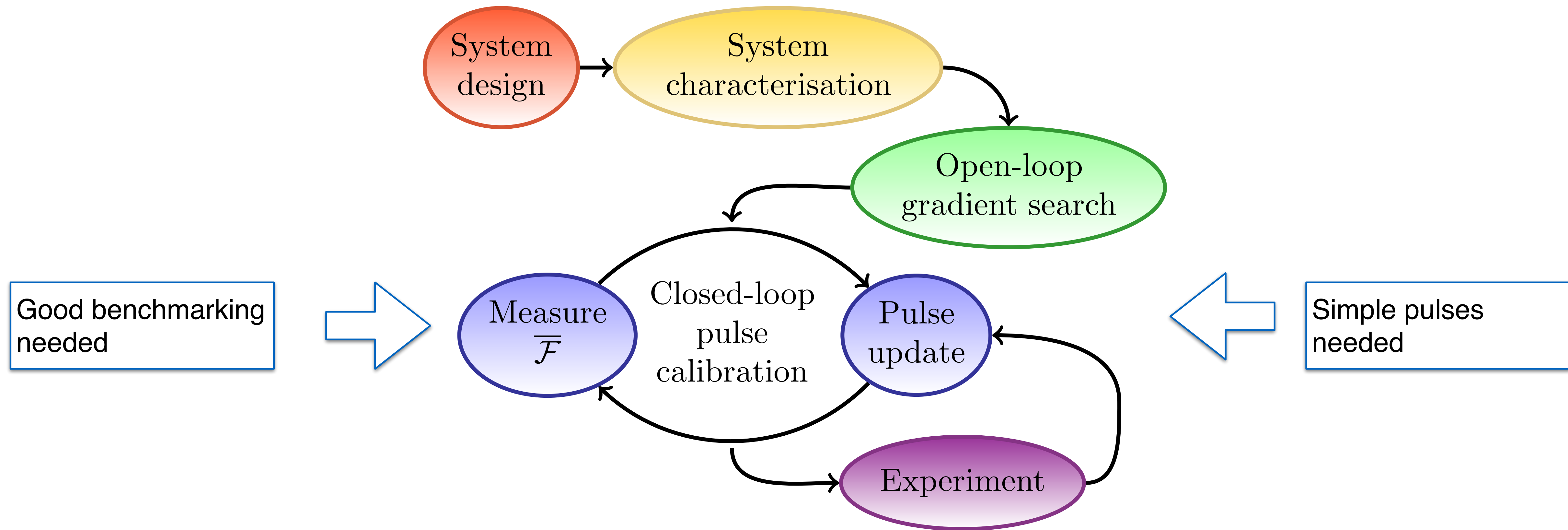
One solution: Be like the other fields ([Heeres et al., 2016](#)): Extreme precision at limited bandwidth (not exploring all of OC potential)

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{junk}} + \sum_i u_i(t) \left(\hat{H}_i + \hat{H}_{i,\text{junk}} \right)$$

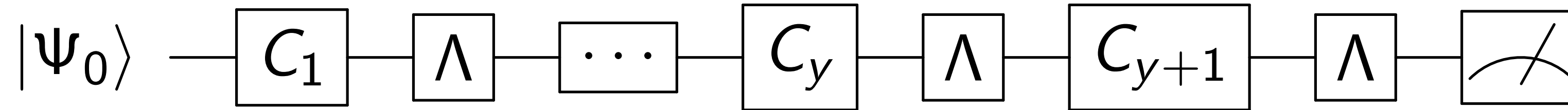
Unwanted degrees of freedom: i) non-computational energy levels ii) spurious DOFs
[Markovian decoherence usually beaten by speed]



ADAPTIVE HYBRID OPTIMAL CONTROL

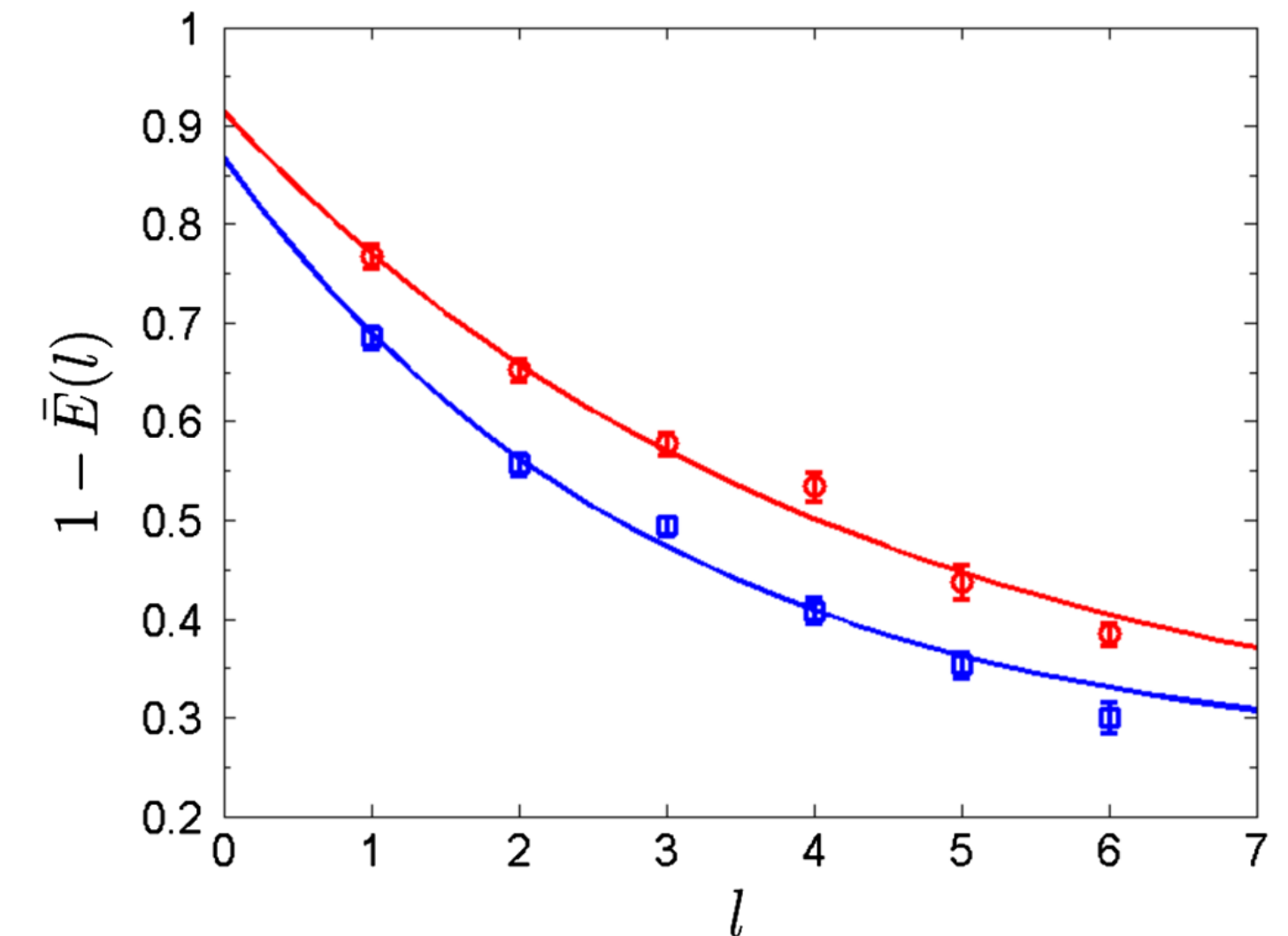


RANDOMIZED BENCHMARKING



- Clifford gates: Normalizer of Pauli group
- Gottesman-Knill theorem: Can be efficiently simulated
- Can be inverted (Clifford **group**)
- Random sequence leads to depolarizing channel (twirling)
- Fidelity related to survival probability

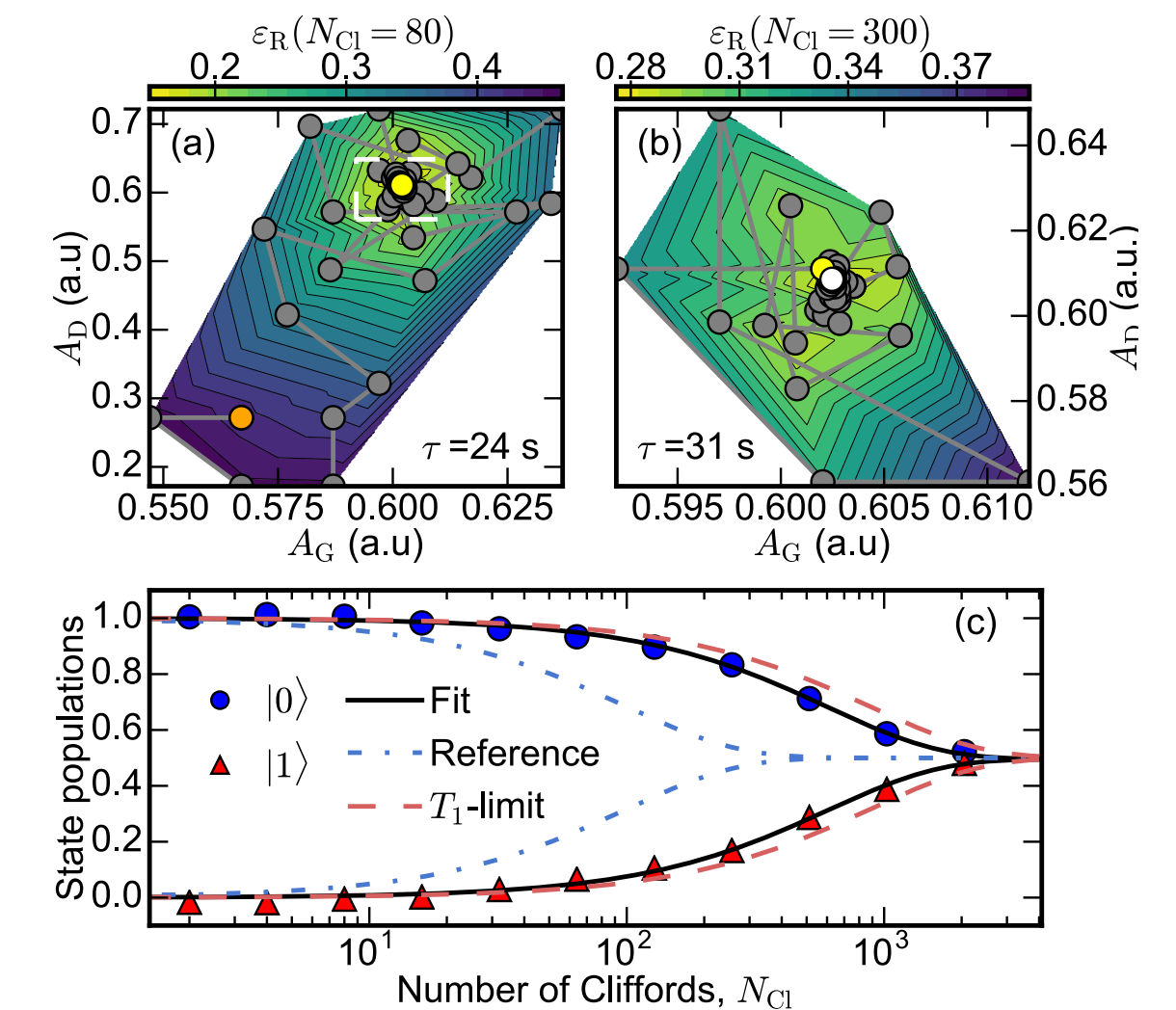
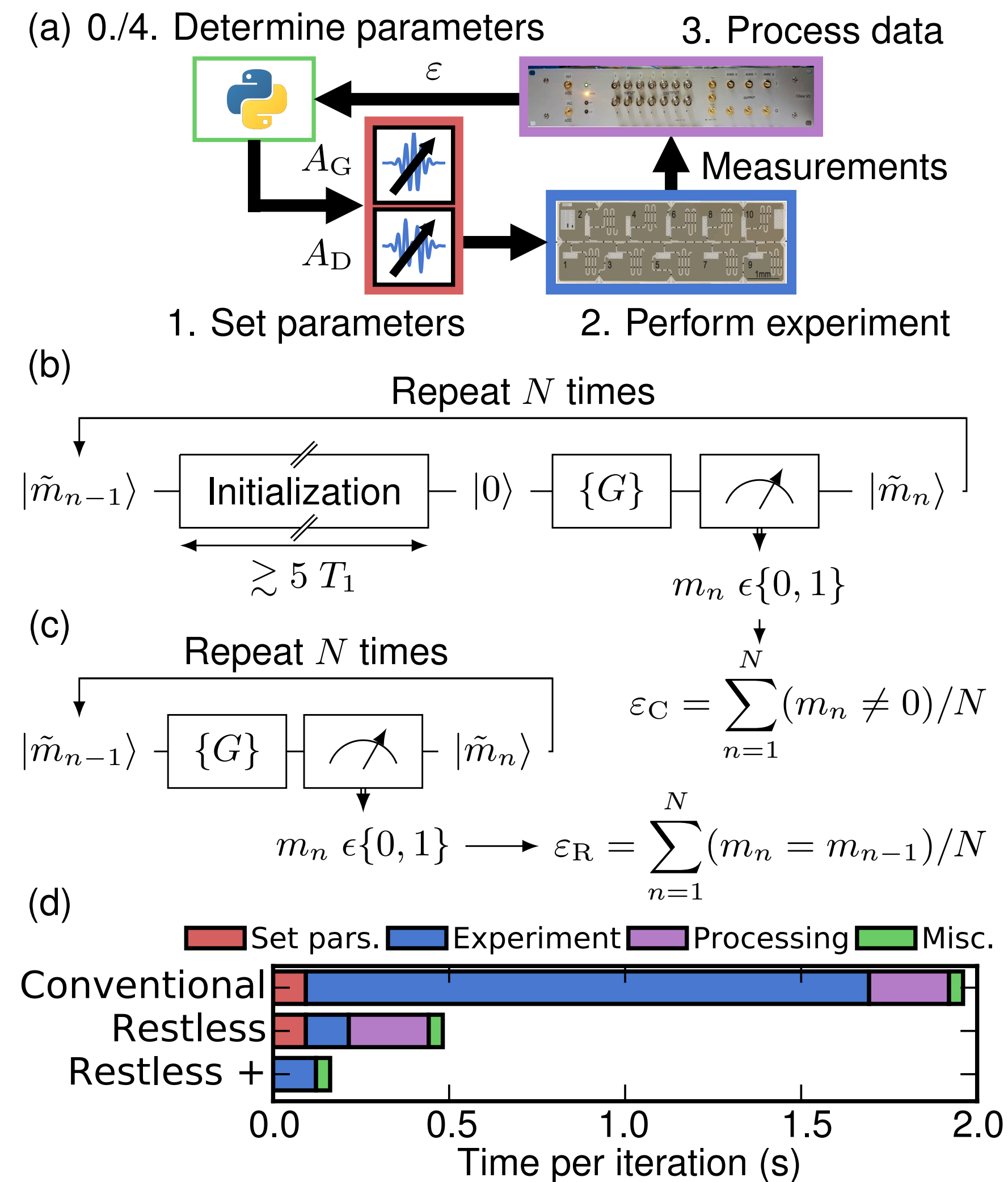
$$\Delta_\lambda(\rho) = \lambda\rho + \frac{1-\lambda}{d}I$$



Offset: SPAM
Decay constant:
Error per gate

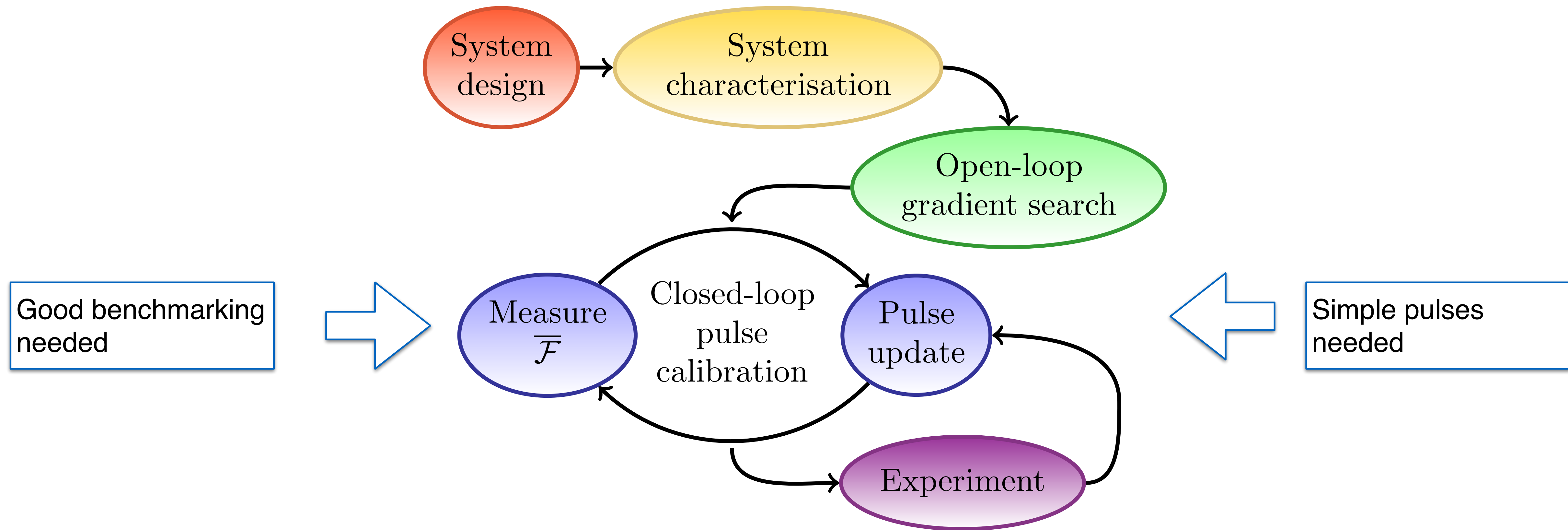
Fidelity measurement

- Randomized benchmarking
- Fast and reliable convergence
- Can be extended to measure more details
- Time consuming re-start
- Better: **Restless tuneup**



- Avoid waiting for initialization
- measure correlations instead of absolute results

ADAPTIVE HYBRID OPTIMAL CONTROL



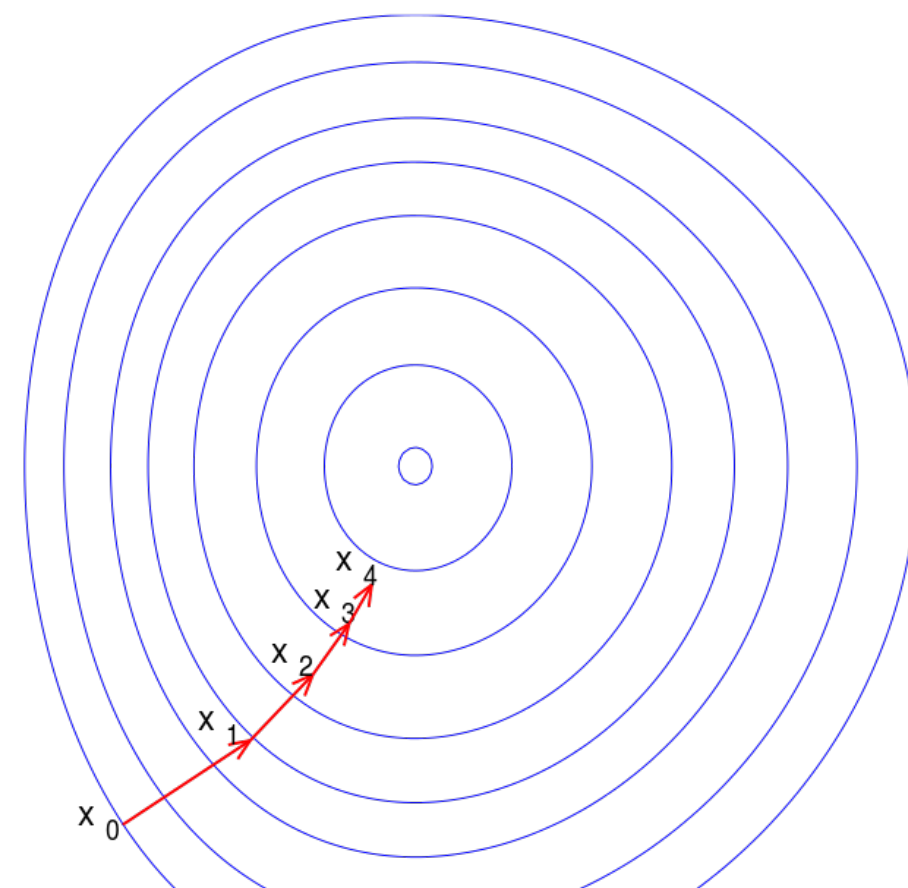
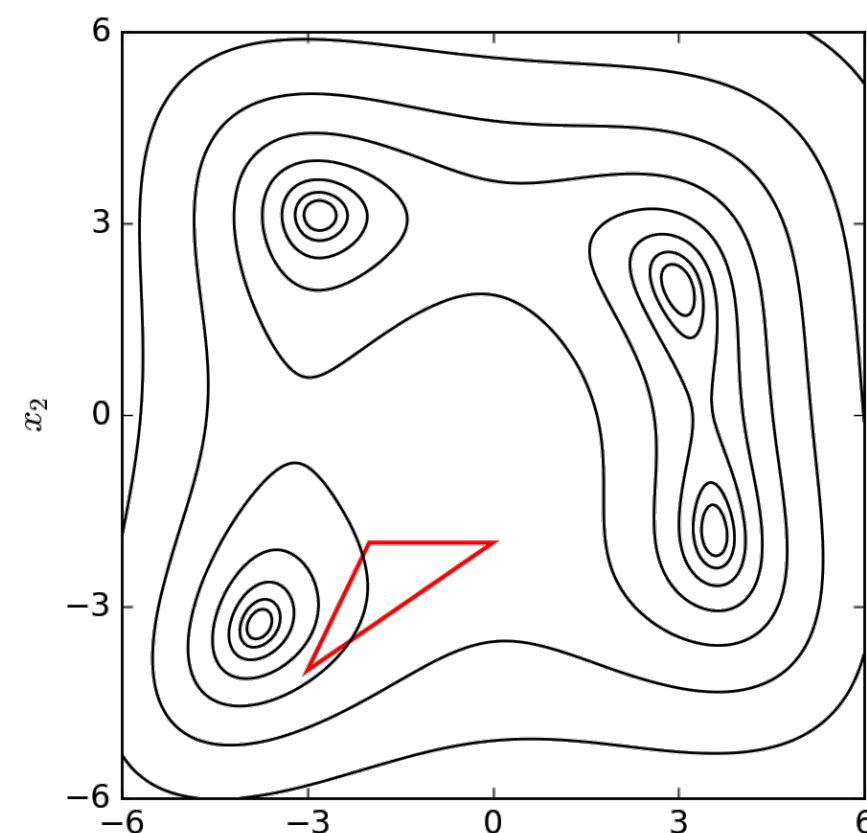
GOAT: Making pulses as
simple as possible

(but not simpler)

= get away from piecewise constant

Direct vs. Gradient methods

- Direct method:
 - Parametrize the control functions $c_k(\bar{\alpha}, t)$
 - Evaluate at several points
 - Figure out the next set of points to sample
 - Repeat
- Pro: Great for closed loop calibration
- Con: Slow (very slow for many parameters)



- Requires computing the gradient
- Start somewhere, follow the gradient
- Pro:
 - Fast
 - Can handle large parameter spaces
- Con:
 - +computing the gradient
 - “Krotov” - Based on the Pontryagin Max. Principle (PMP)
 - Non-trivial mathematically – calculus of variations
 - Requires backwards-in-time propagation of an adjoint state

L. S. Pontryagin, V. G. Bol'tanskii, R. S. Gamkre-lidze, and E. F. Mischenko.
The Mathematical Theory of Optimal Processes. Pergamon Press, New York (1964)



OPEN-LOOP OPTIMAL CONTROL WITH GOAT

From $H(\bar{\alpha}, t) = H_0 + \sum_{k=1}^K c_k(\bar{\alpha}, t) H_k$ we compute $\partial_{\bar{\alpha}} H(\bar{\alpha}, t)$

We propagate

$$\partial_t \begin{pmatrix} U \\ \partial_{\bar{\alpha}} U \end{pmatrix} = -\frac{i}{\hbar} \begin{pmatrix} H & 0 \\ \partial_{\bar{\alpha}} H & H \end{pmatrix} \begin{pmatrix} U \\ \partial_{\bar{\alpha}} U \end{pmatrix}$$

to get $\partial_{\bar{\alpha}} U(\bar{\alpha}, T)$

Finally, $\partial_{\bar{\alpha}} g(\bar{\alpha}) = -\text{real} \left(\frac{g^*}{|g|} \frac{1}{\dim(U)} \text{Tr} \left(U_{goal}^\dagger \partial_{\bar{\alpha}} U(\bar{\alpha}, T) \right) \right)$

We follow the gradient using standard algorithms,
e.g. LBFGS.

GOAT RESULTS: CROSS-RESONANCE GATES

Current best: 0.992 at 160ns

Single carrier : 0.999 (coherent)

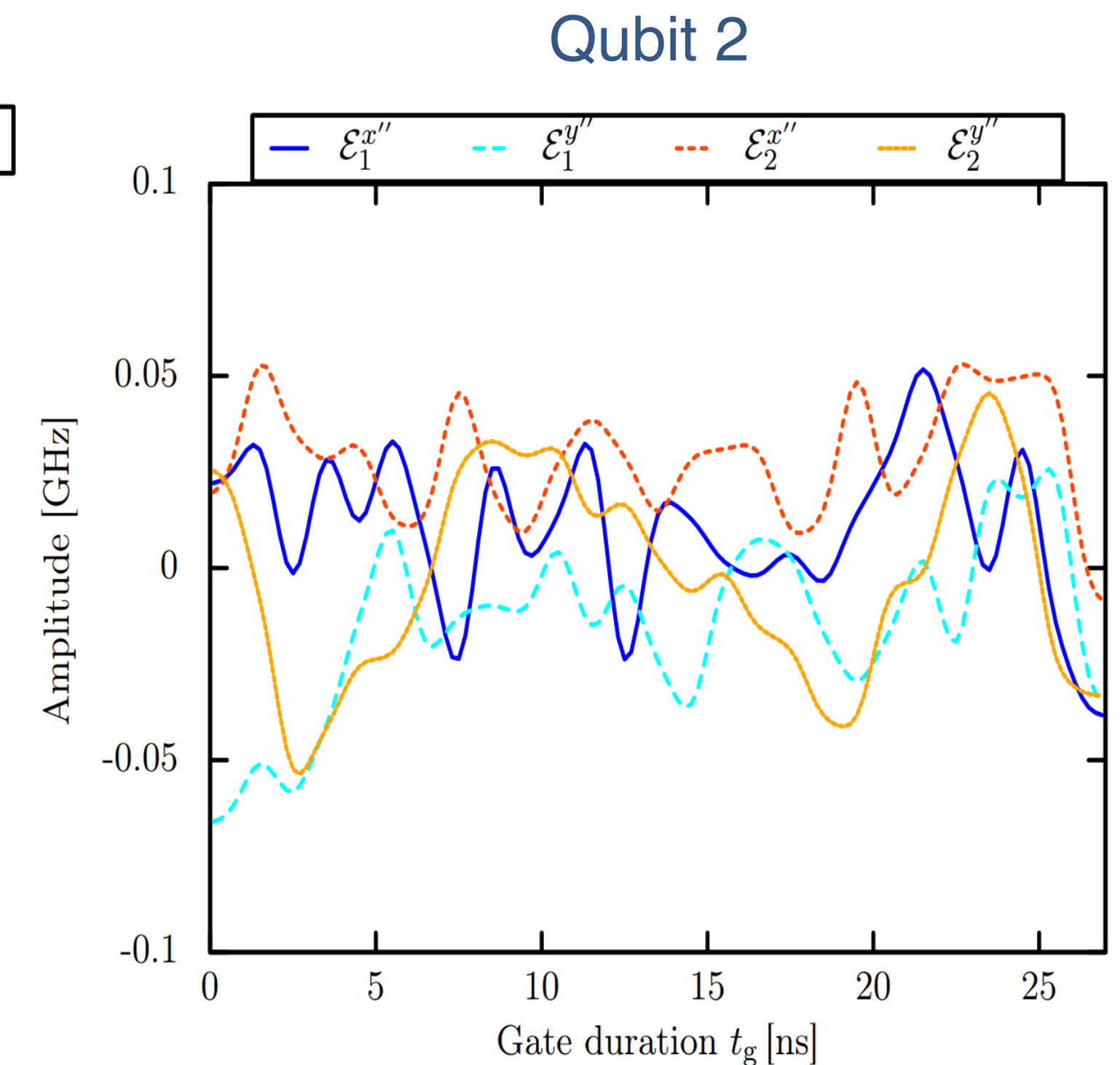
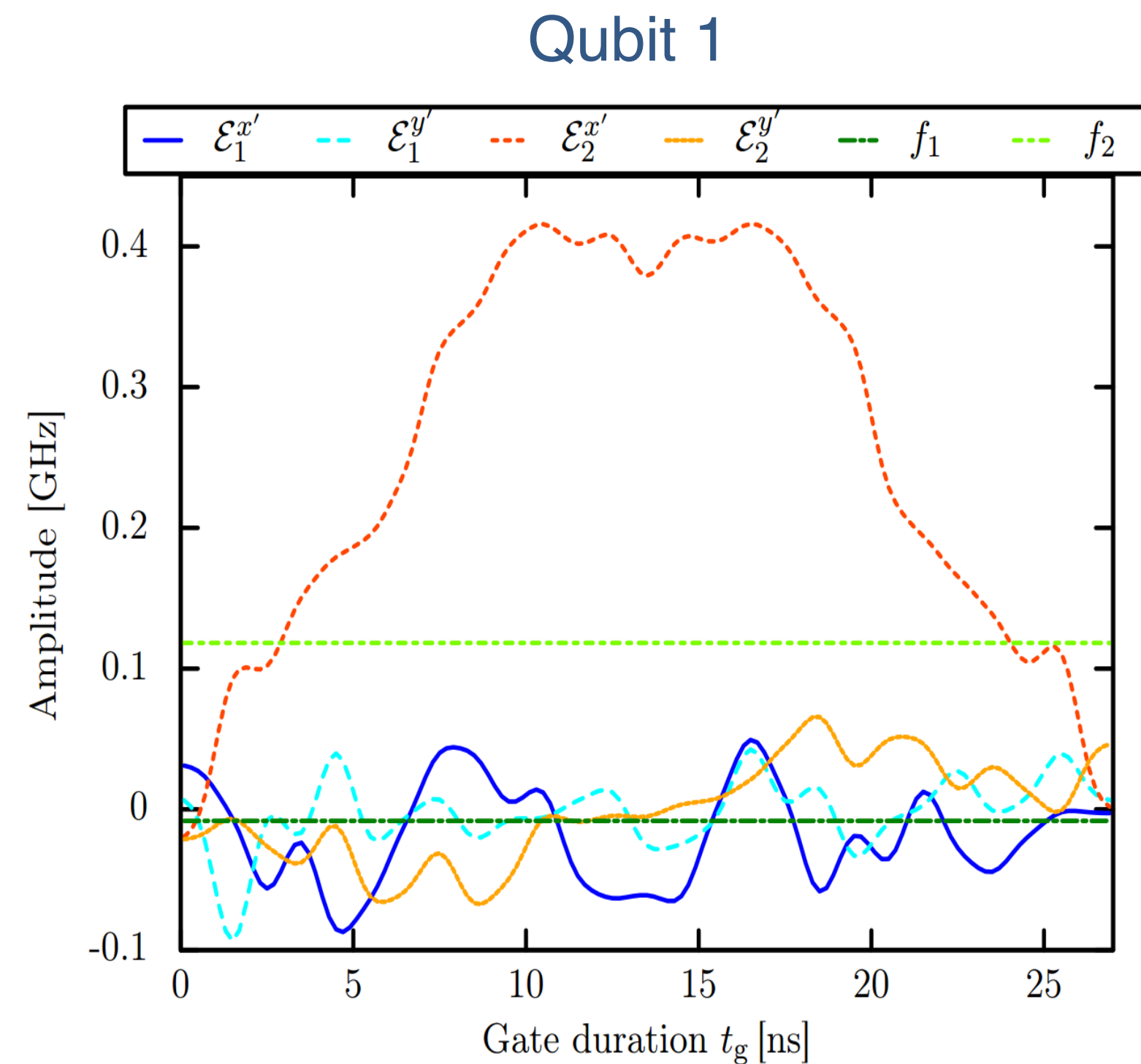
70ns for 9 Fourier components

15ns for 167 components

Two carriers, i.e. drive both qubits

Result: 27ns 0.998 (coherent)

PWC with 330MHz filter



GOAT CAN RESOLVE IBM'S FREQUENCY COLLISIONS

IBM utilizes fixed-frequency transmon qubits with static resonator coupling.

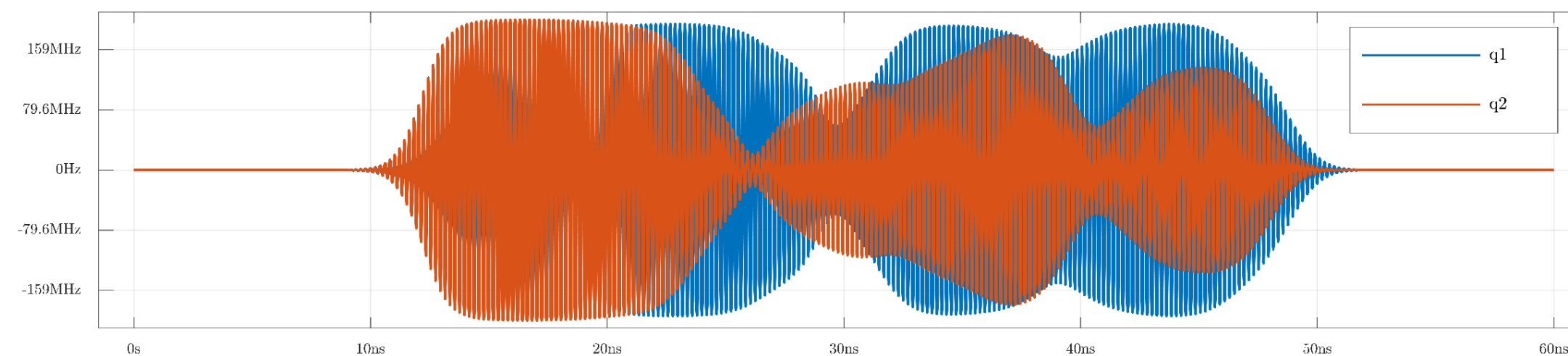
Current limitations in transmon manufacturing limit control of qubit frequency.

WAHWAH+

Inevitable result is high probability of freq. collisions in 50+ qubit chips.

Example: 01 transition of qubit A close to 12 transition of neighboring qubit B.

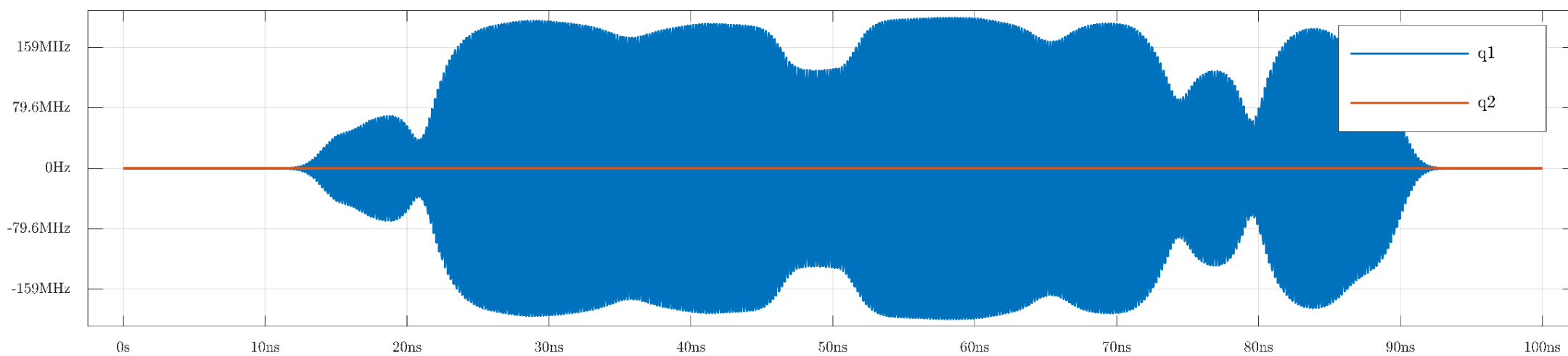
Issue resolvable with GOAT optimal control:



Identity gate when $01_A = 01_B$

Both qubits driven

Fidelity (coherent): 0.997



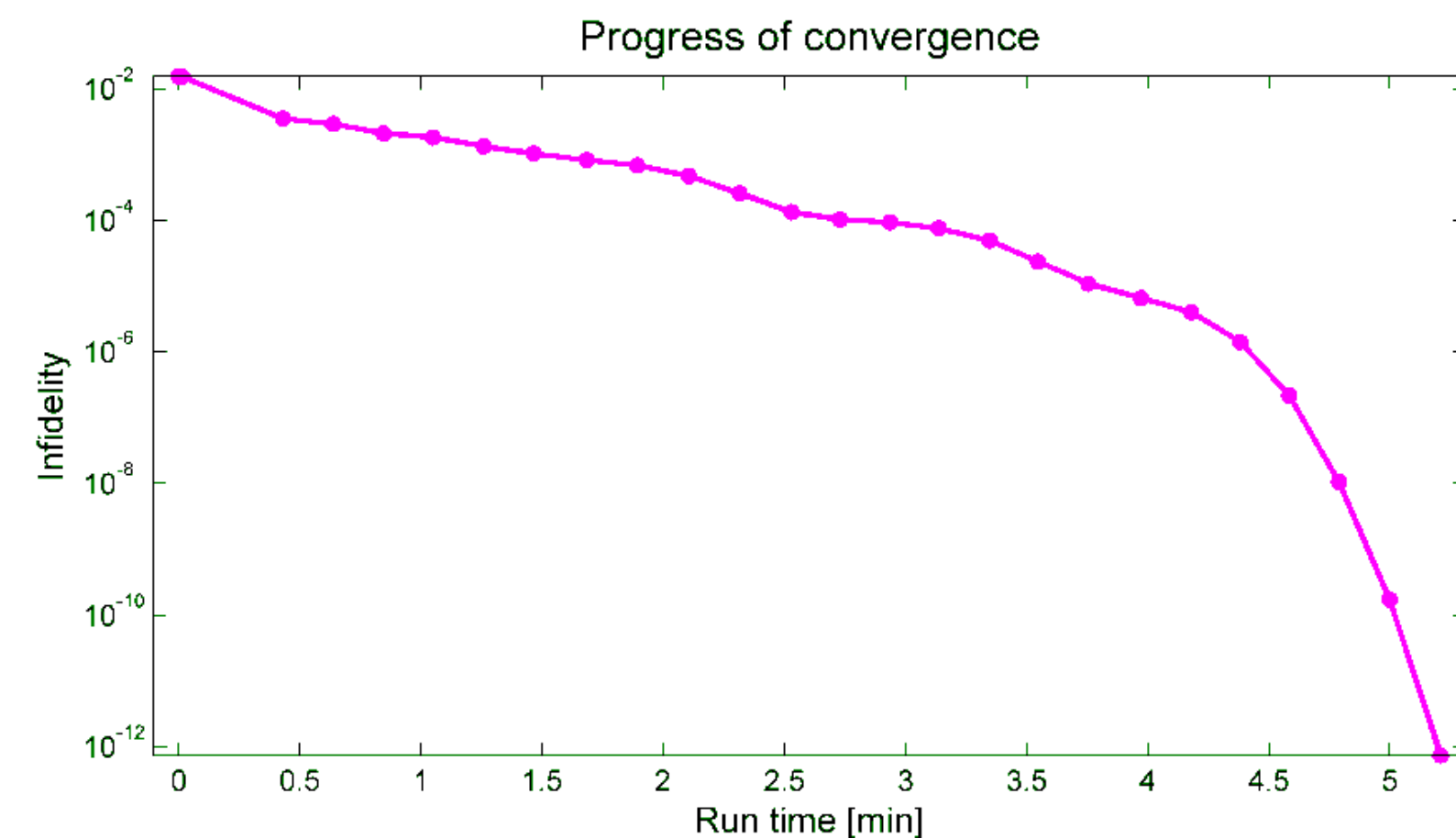
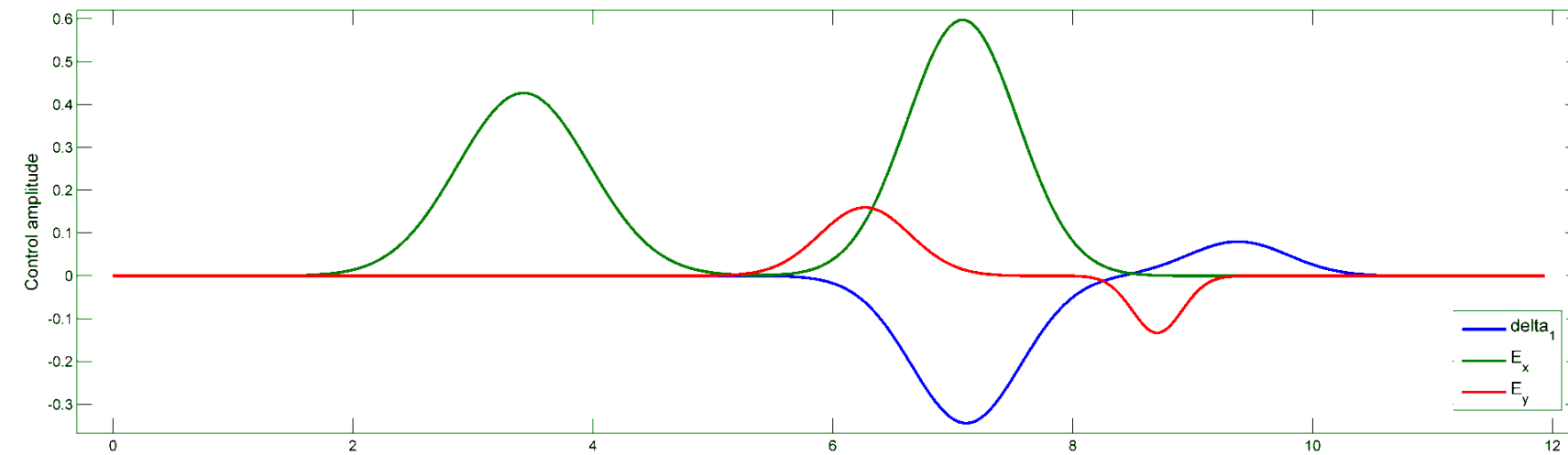
CNOT gate for *identical* qubits
(coupling to resonator different)

Fidelity (coherent): 0.992

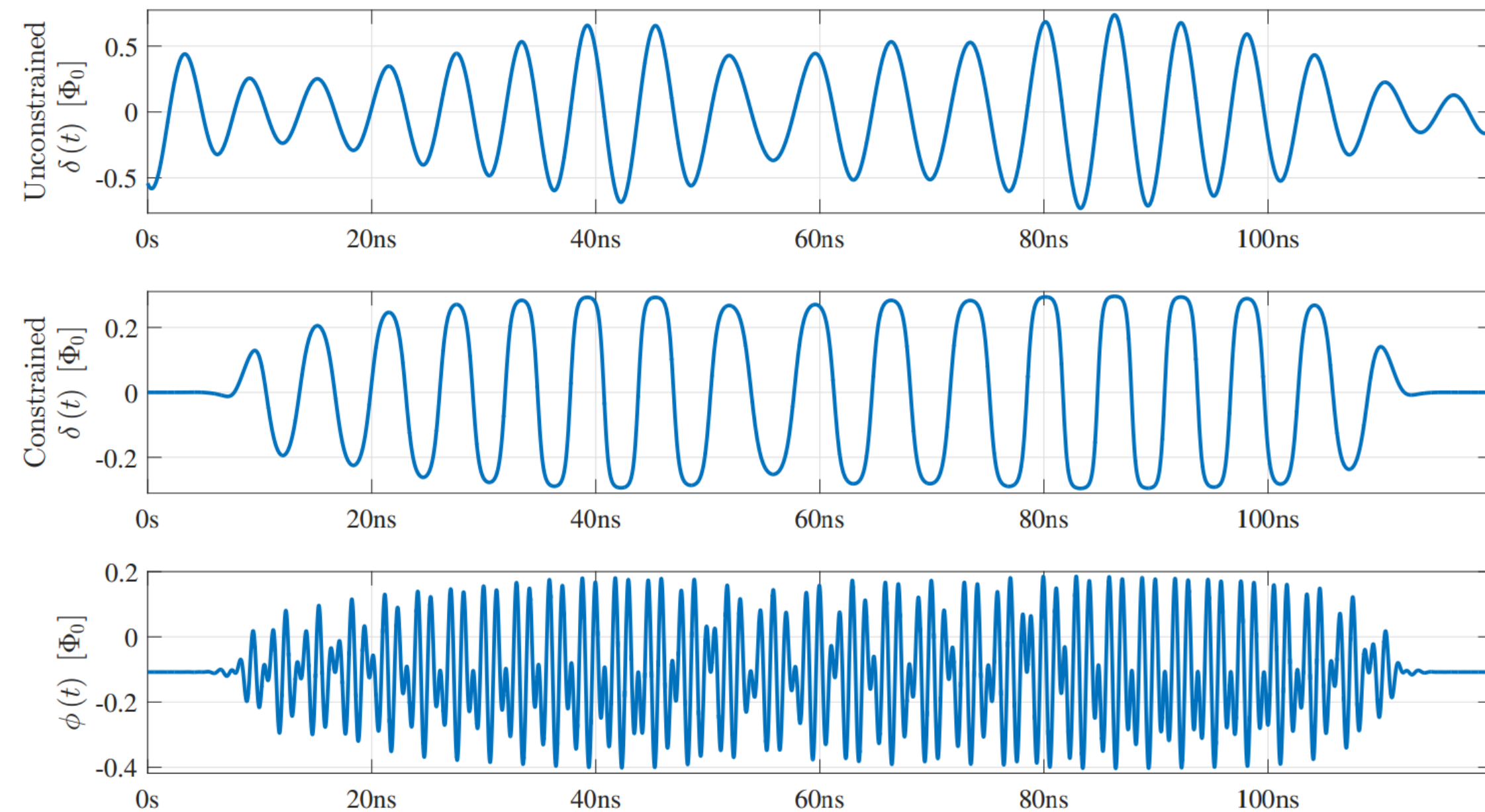
GOAT was used to do this

Xmon-like system, 10ns single qubit gate

2 Gaussians per control channel



Parametric gate with tunable coupler



6 Fourier components, full system model
 10^{-12} infidelity (ignoring dephasing)

C³: Combined characterization and control

BACK TO THE DRAWING BOARD

If we need pulse calibration,
it means our model is wrong

We learn nothing about the model
from the calibration

We need an error budget to improve design
of next-gen hardware

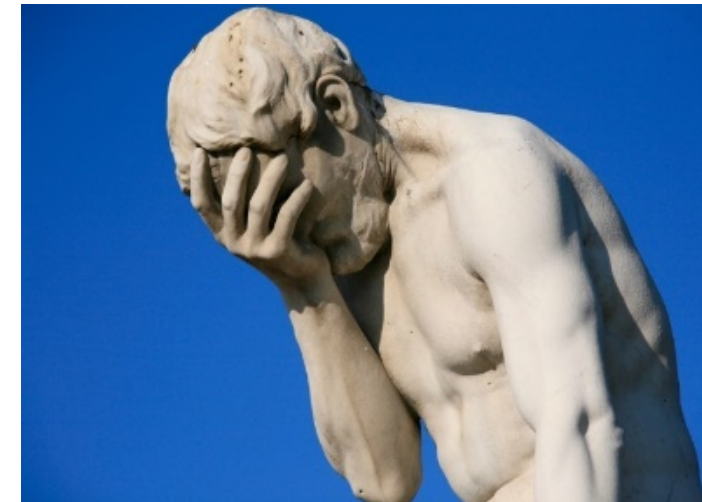
We need a good model for a detailed error budget

A Good Model

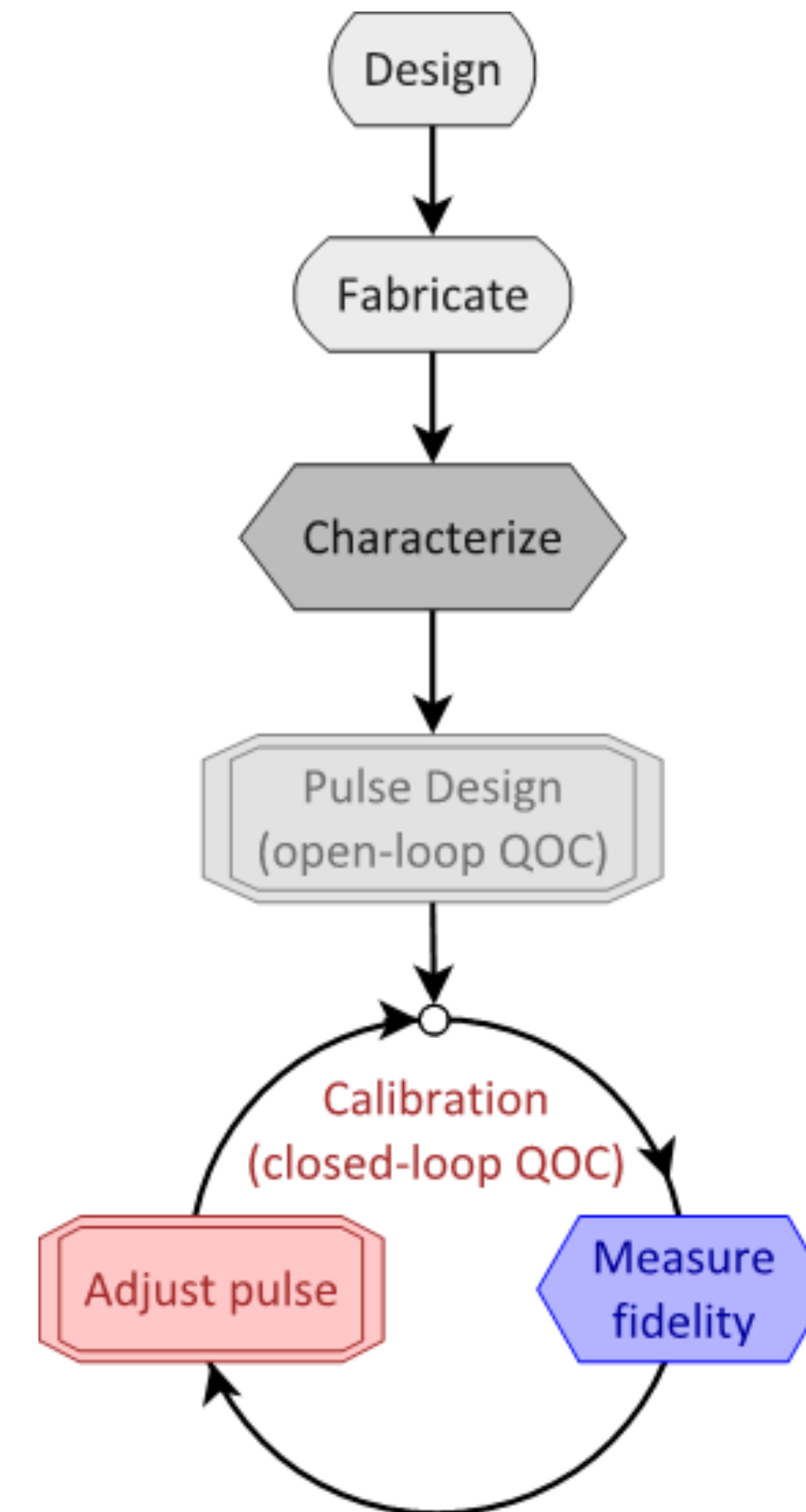
Predicts behavior for the actions we care about
(gates in a multi-qubit system)

To the accuracy we care about (0.9999)

From a Good Model, best achievable gates
and error budget are derivable



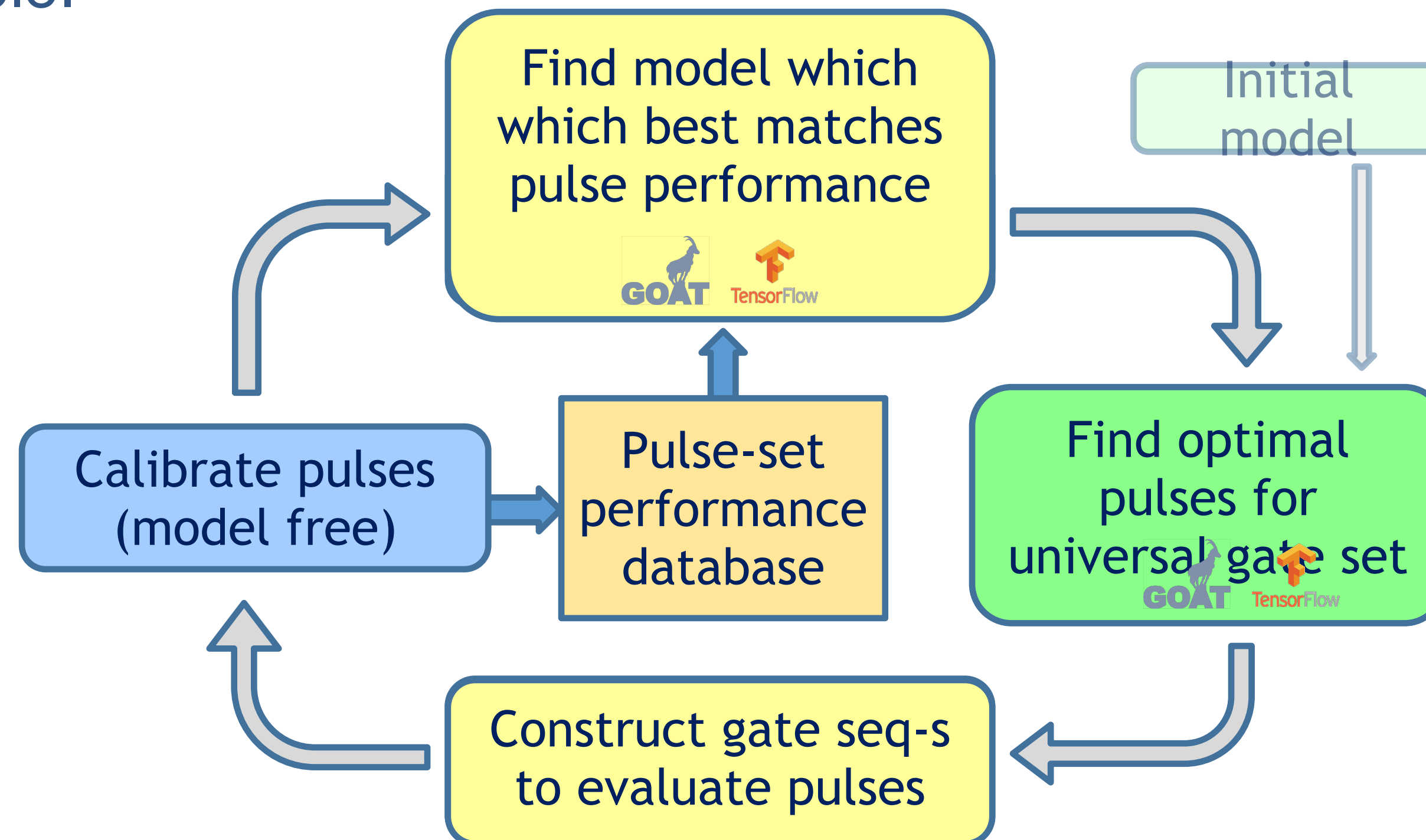
Cain / Henri Vidal



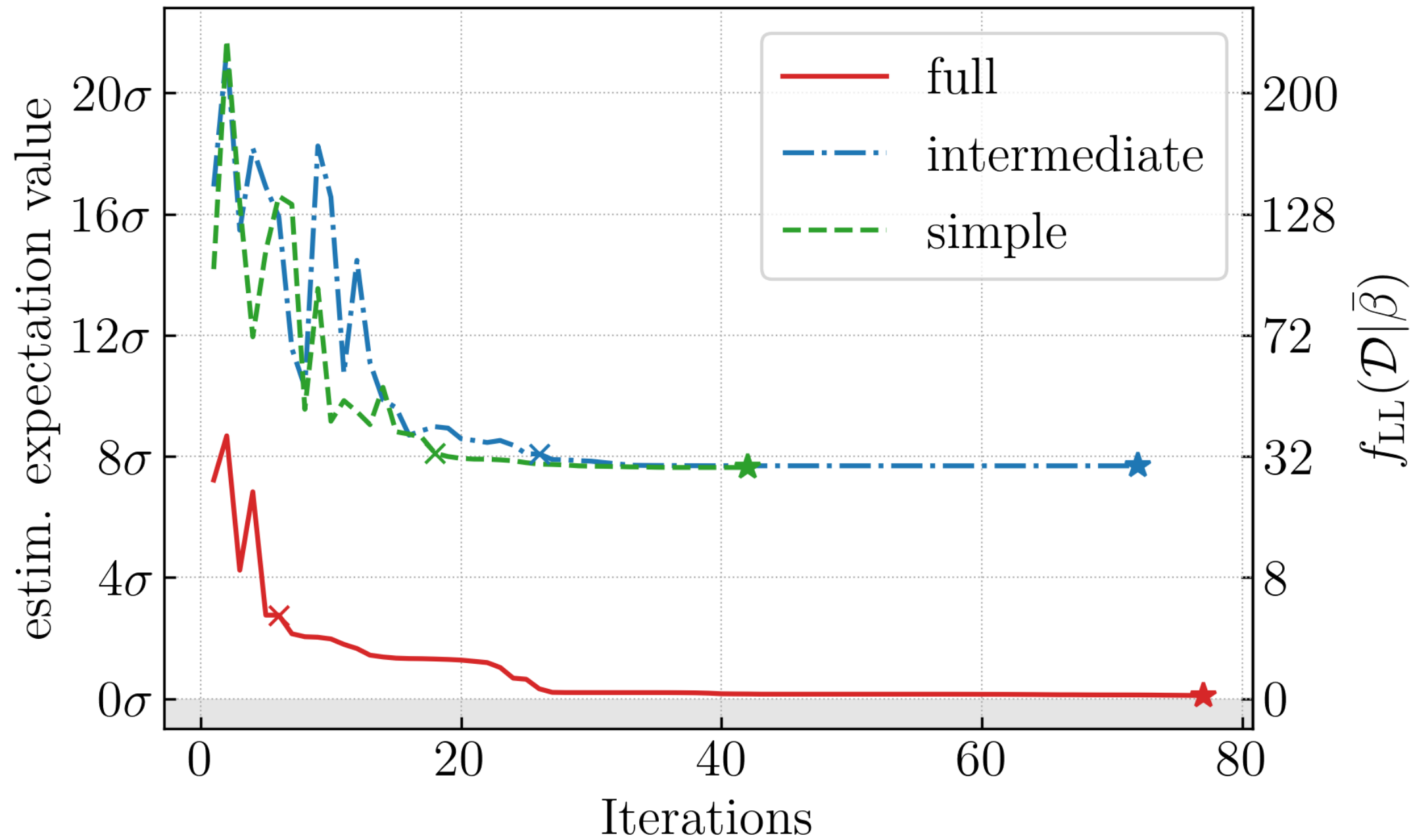
C³ – Combined Calibration and Characterization

- Best gates given current model-based via GOAT optimal control
- Model-free calibration with advanced algorithms
- Improve model based on observed pulse fidelities, using GOAT
- C³ converges when model is *Good*, i.e. accurately predicts fidelities. From a *Good Model*, best achievable gates and error budget are derivable.

Nonintuitive arrows:
Machine learning



FULL MODEL MATCHING



Models

simple uncoupled Duffing oscillators

intermediate added coupling

full same as black-box

The dataset

We store sequences and results as

$$\mathcal{D} = \{S_k(\alpha_j) \rightarrow m_{j,k}\}$$

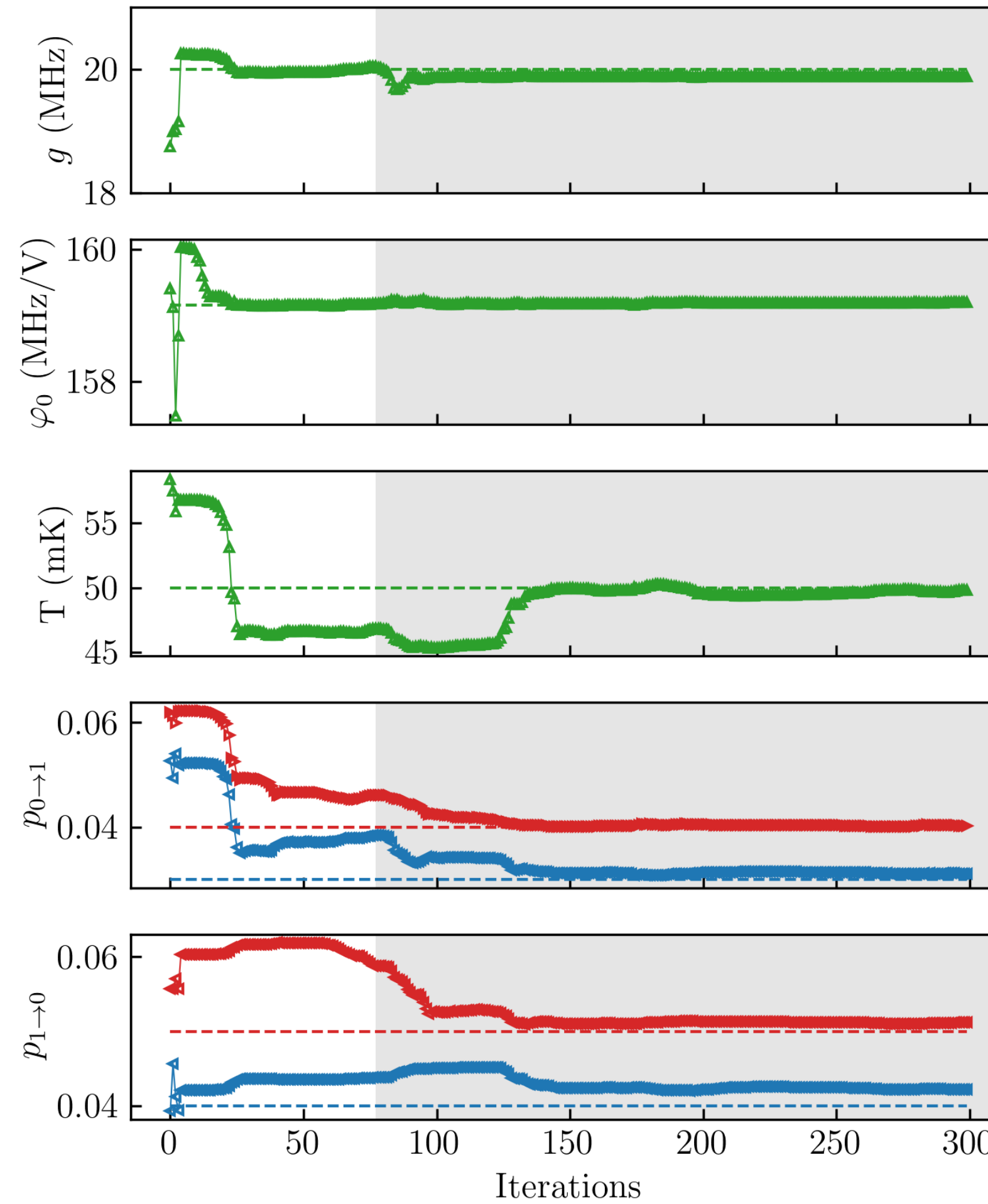
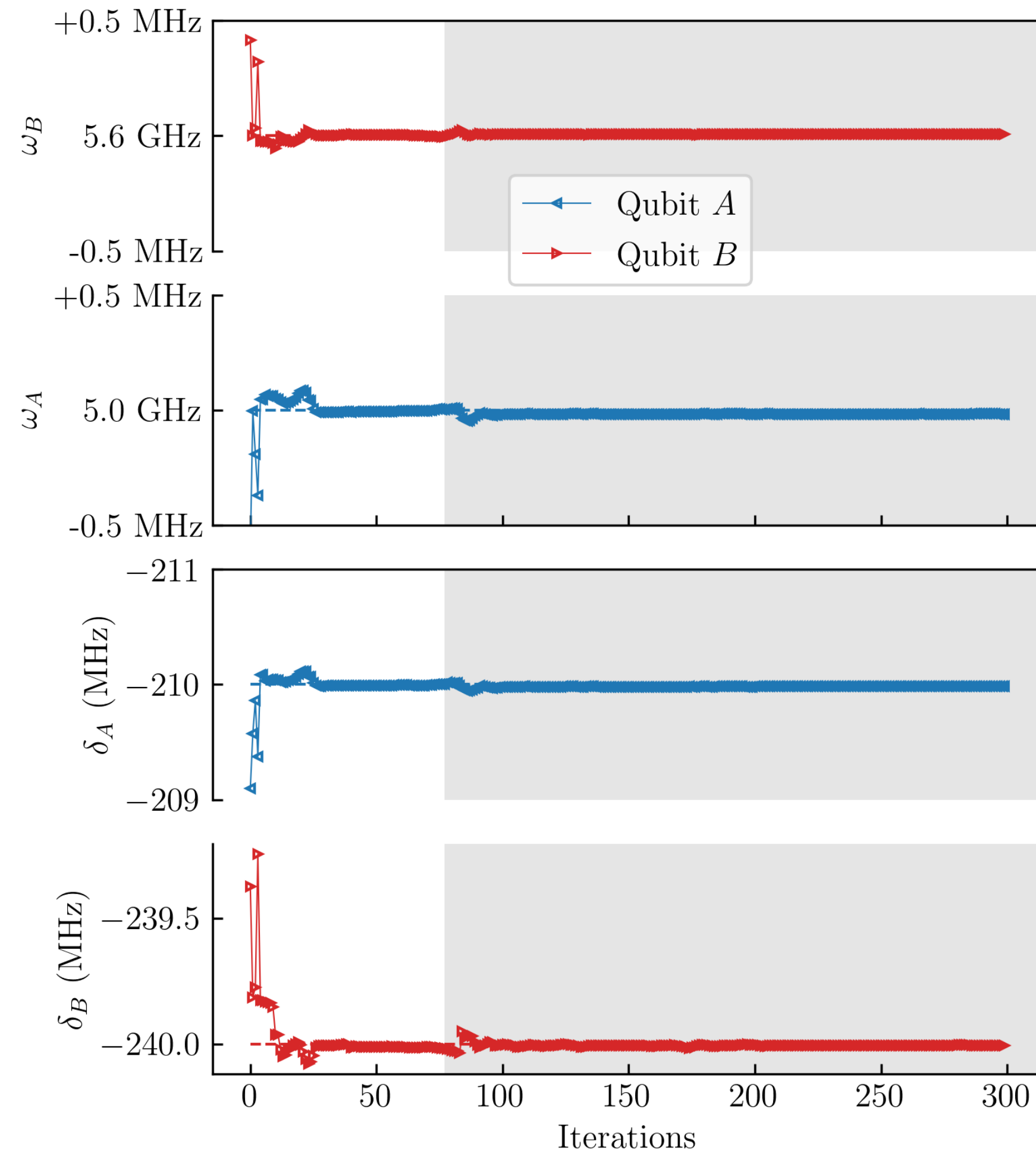
Model match score

Simulate $\tilde{m}_n(\beta, \alpha)$ and compare with (noisy!) experiment

$$f_{\text{LL}}(\mathcal{D}|\beta) = \frac{1}{2N} \sum_{n=1}^N \left[\left(\frac{m_n - \tilde{m}_n}{\tilde{\sigma}_n} \right)^2 - 1 \right]$$

where $\tilde{\sigma}_n$ is the std of \tilde{m}_n
and β are the model parameters.

FINDING MANY PARAMETERS AT THE SAME TIME



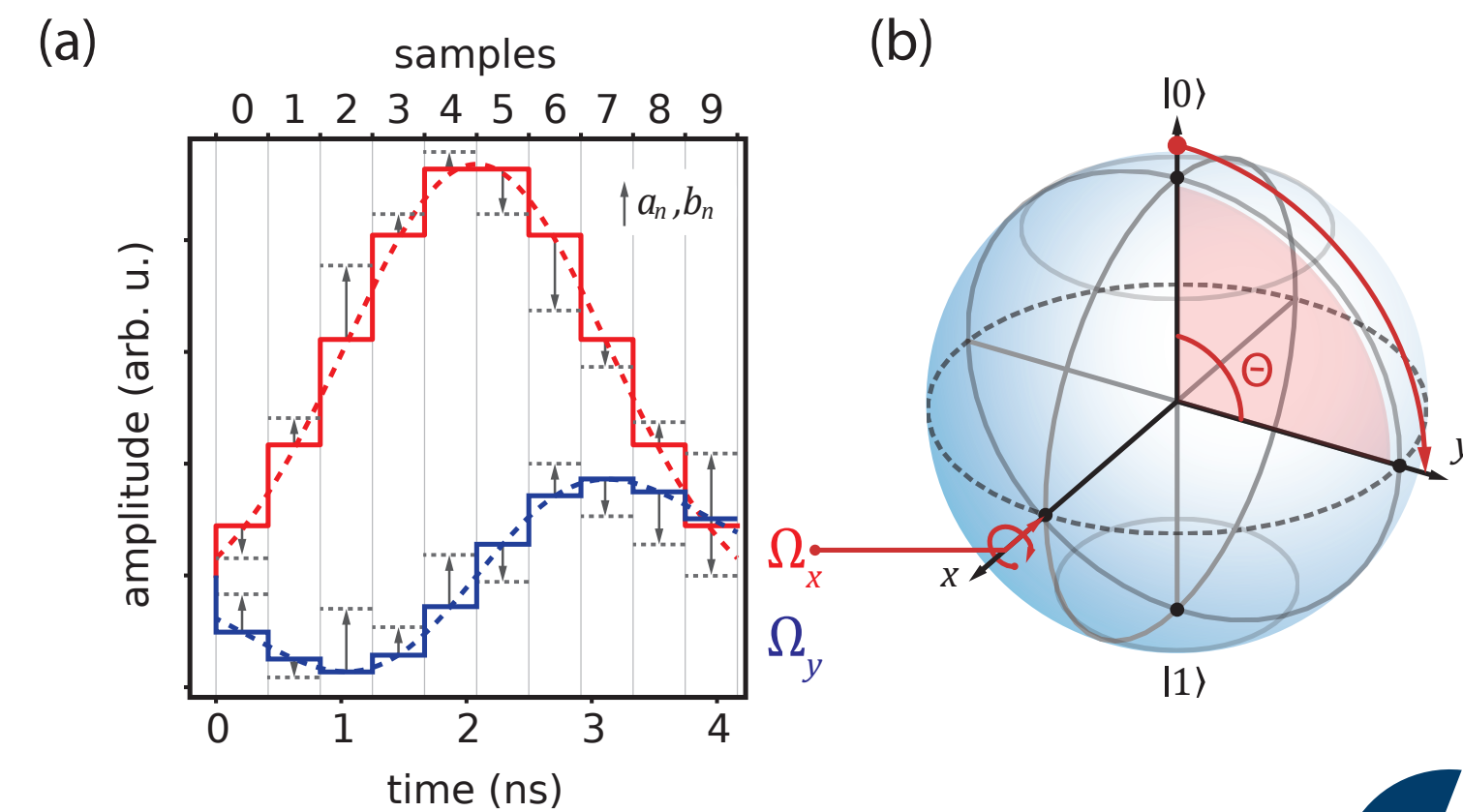
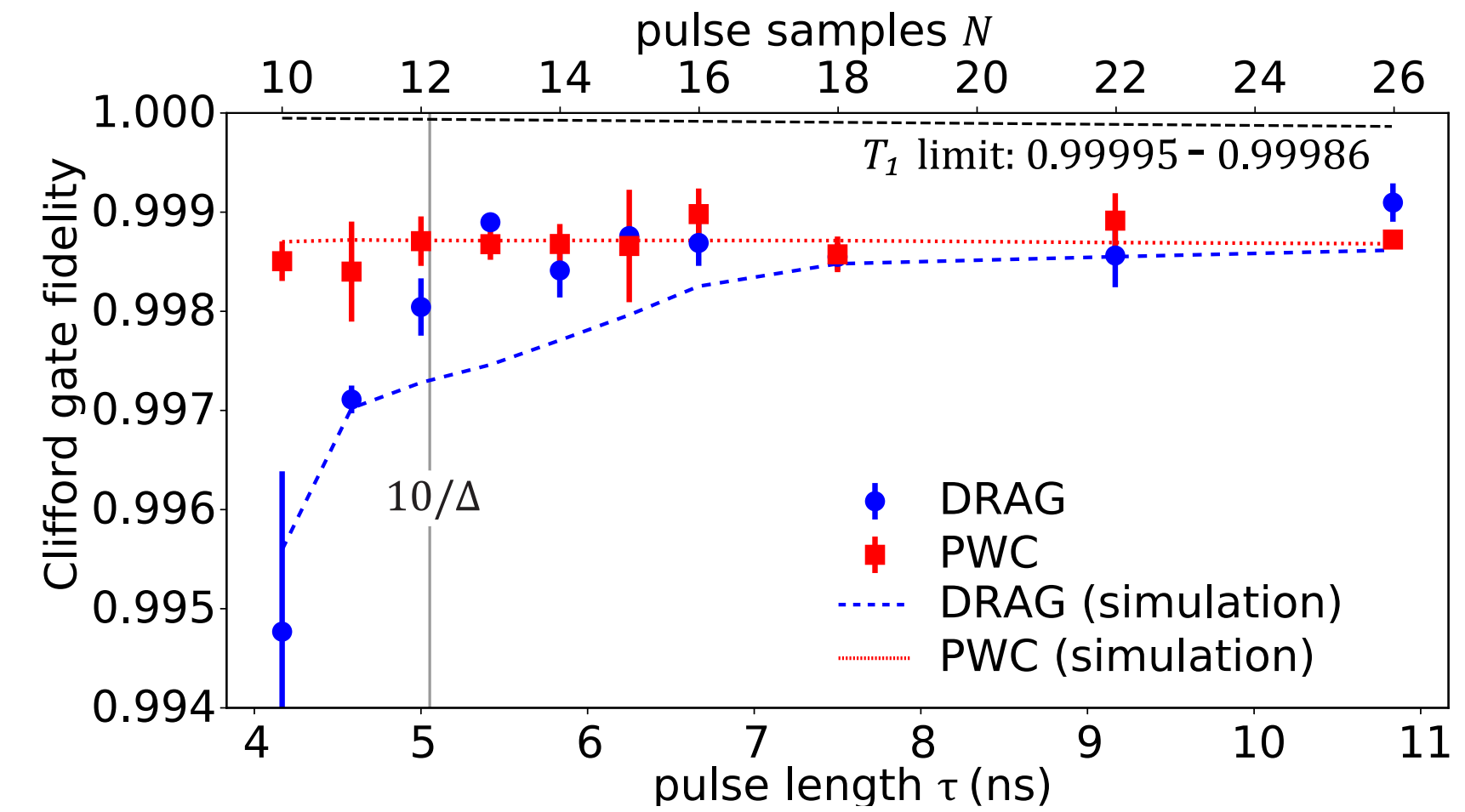
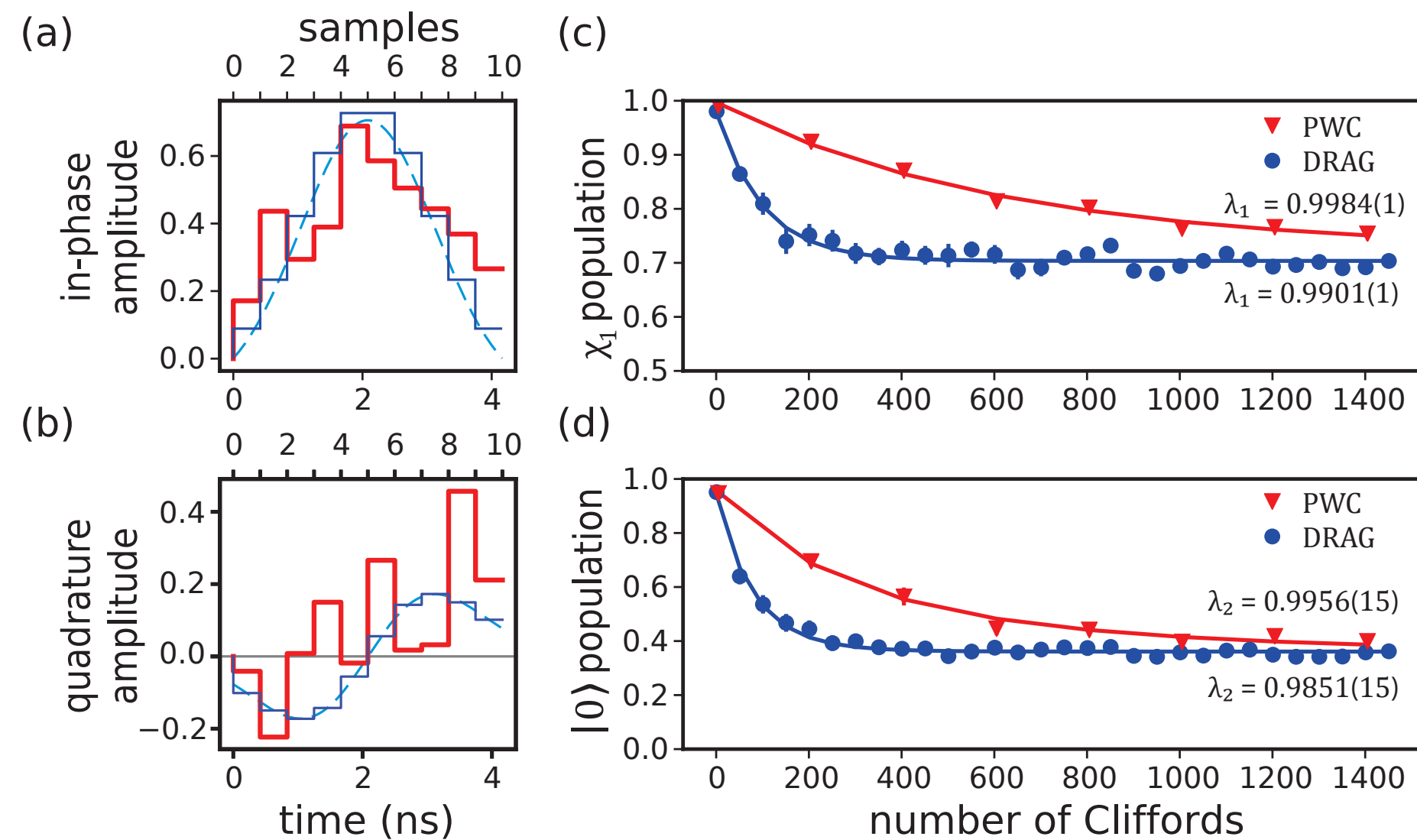
- ▶ Qubit frequencies ω_i
- ▶ Anharmonicities δ_i
- ▶ Coupling g
- ▶ System temperature T
- ▶ Field conversion φ_0
- ▶ Confusion $p_{i \rightarrow j}$

GOING BEYOND DRAG

trace out quantum speed limit

7-fold reduction of error

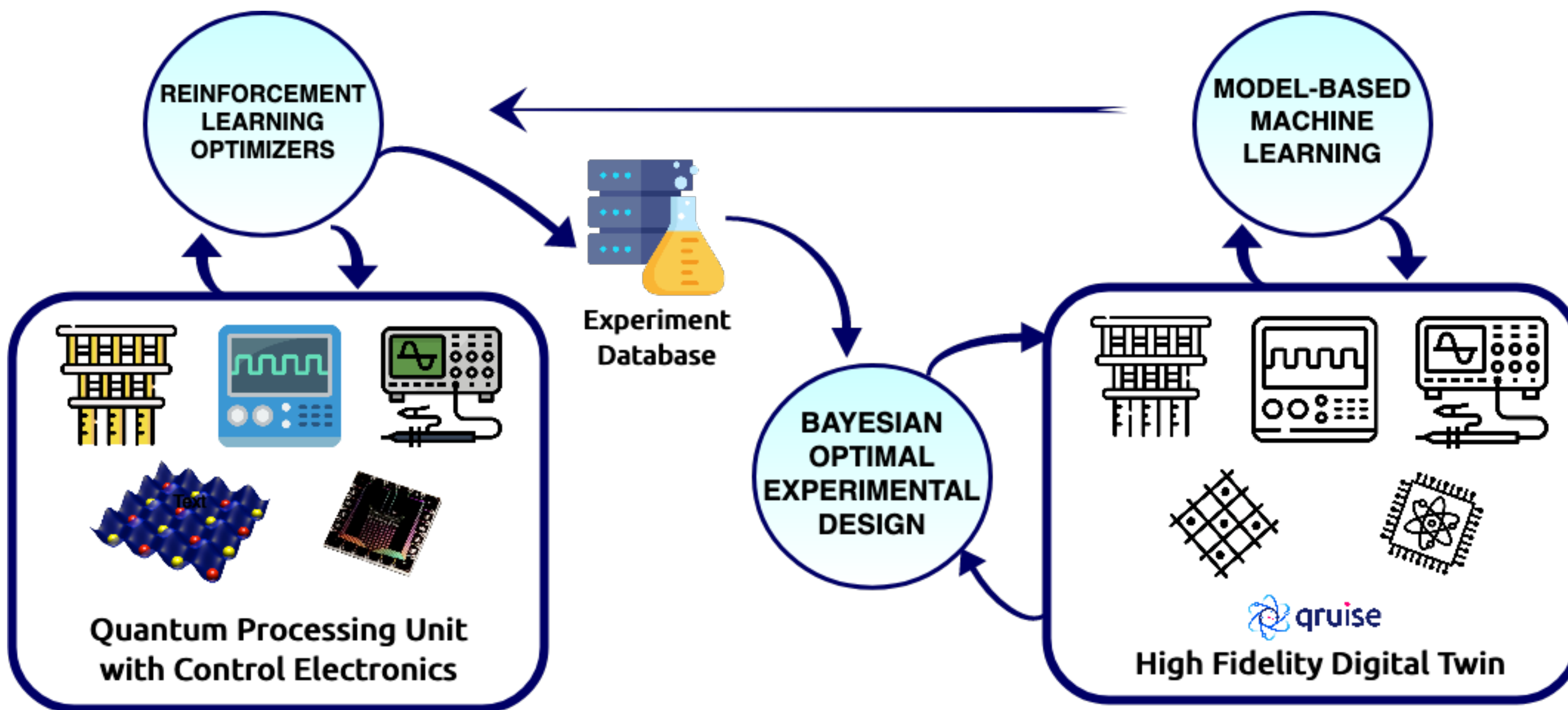
strong deviation from DRAG



Putting it all together



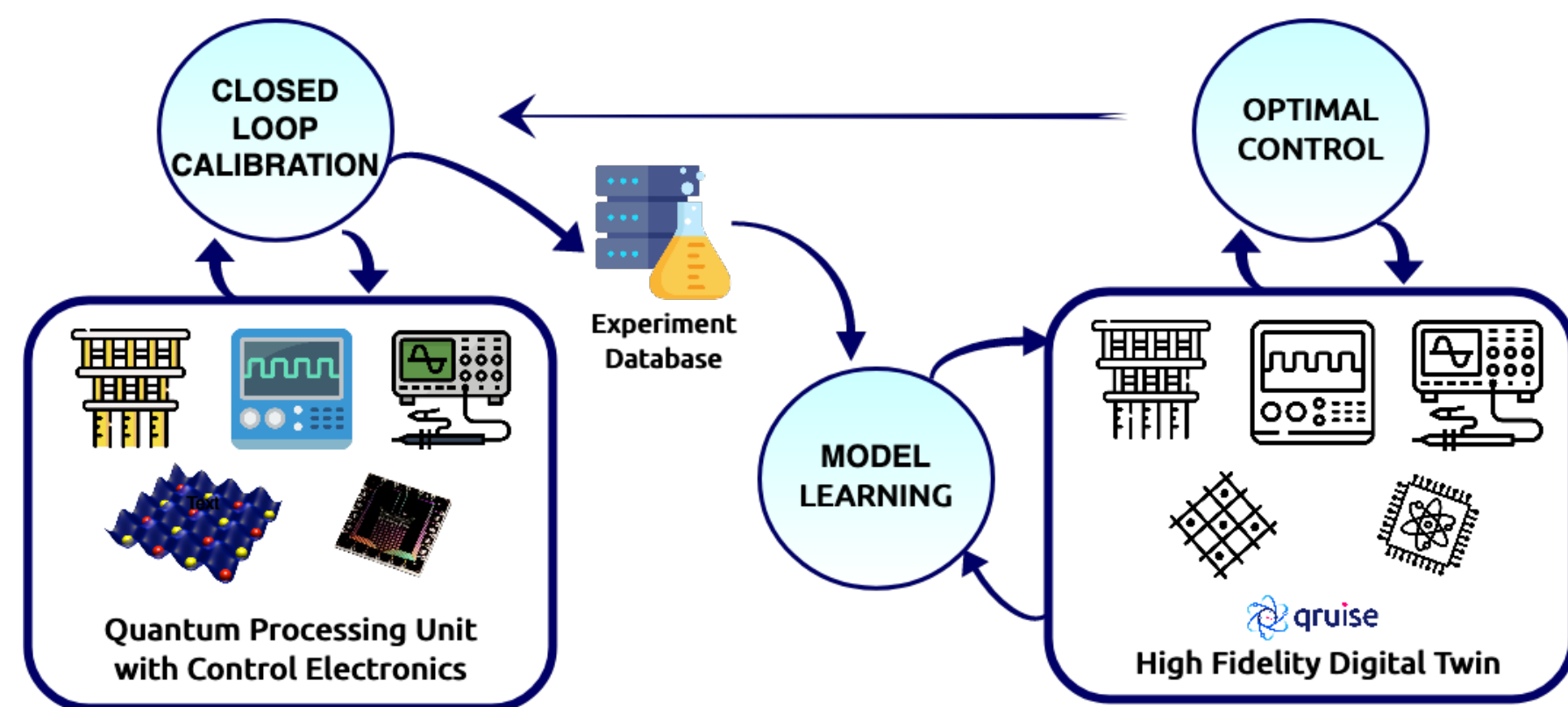
Qruise Stack – it's all ML



White Box Physics based Explainable AI

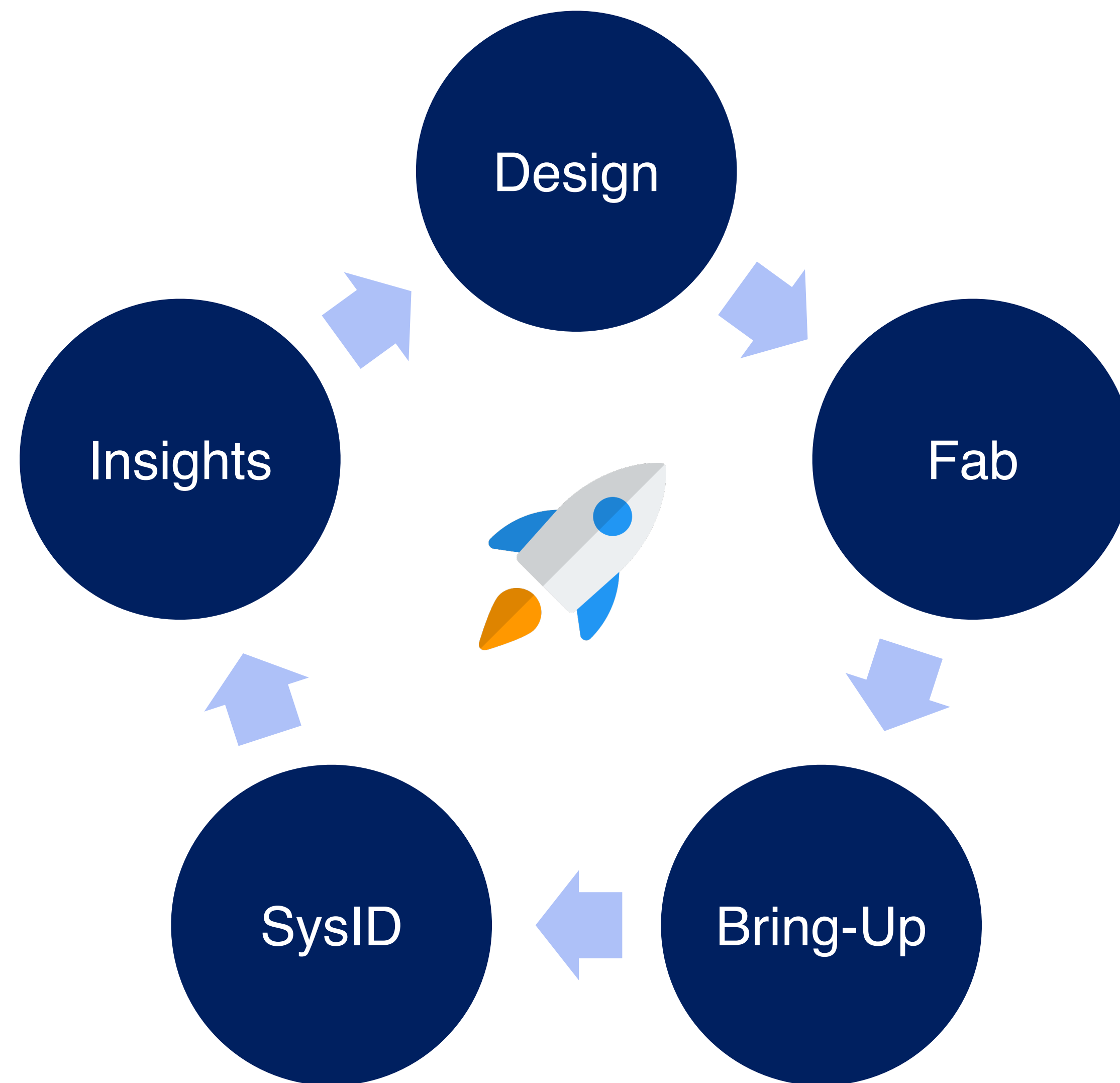


Qruise Stack



Reverse Engineering your Quantum Hardware

- Identify and Isolate error sources
- Understand impact of individual error sources
- Identify where to focus efforts
- Model what-if situations





QruiseOS powering FZ Jülich QC

Zero to Hero Workflow

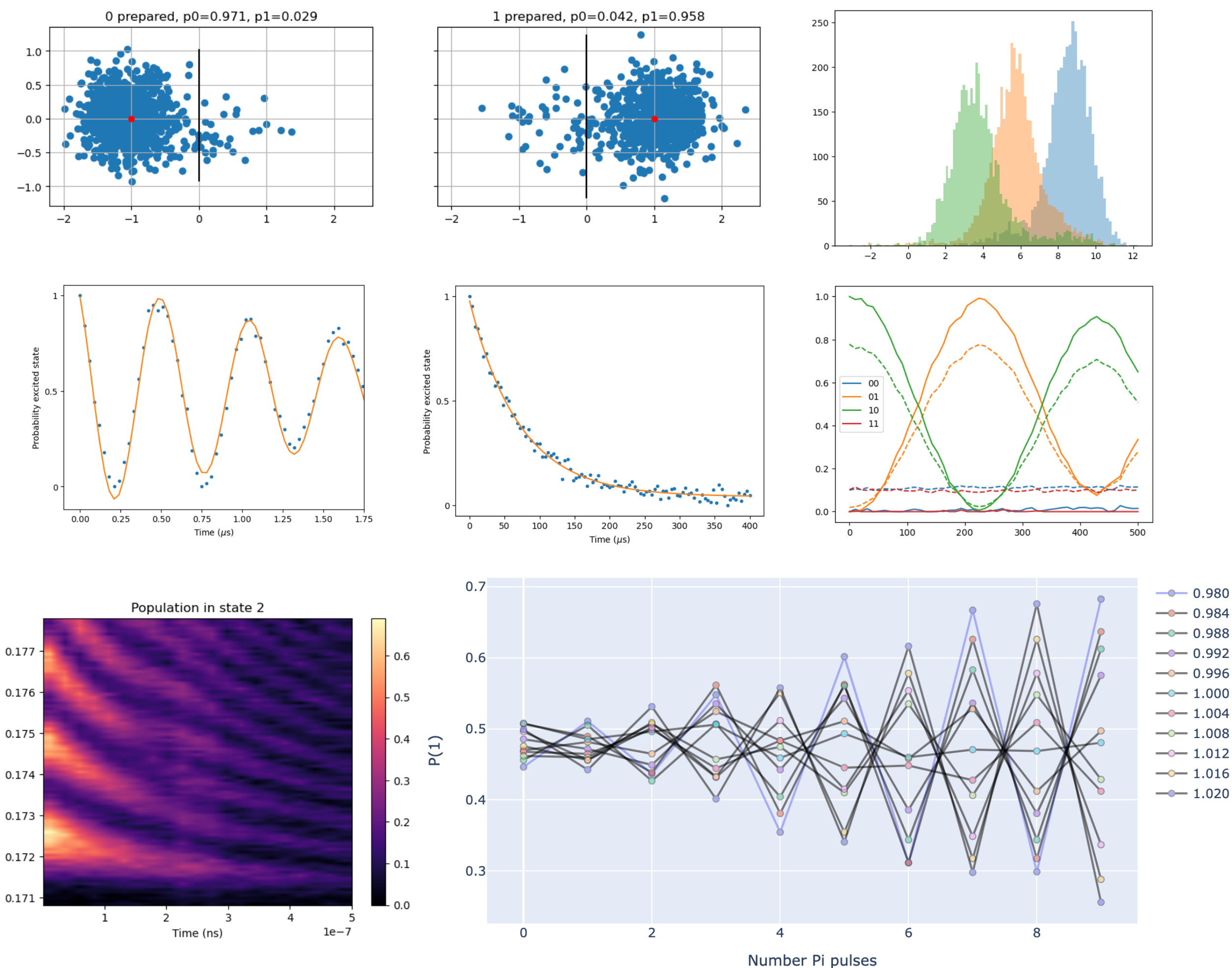
From cool-down to fully functional chip with best-achievable gates and extremely detailed characterization. Automated and at scale.

Ongoing recalibration, removing the need for downtime and keeping fidelities always at the optimum.

Using this data, we perform ongoing recharacterization.

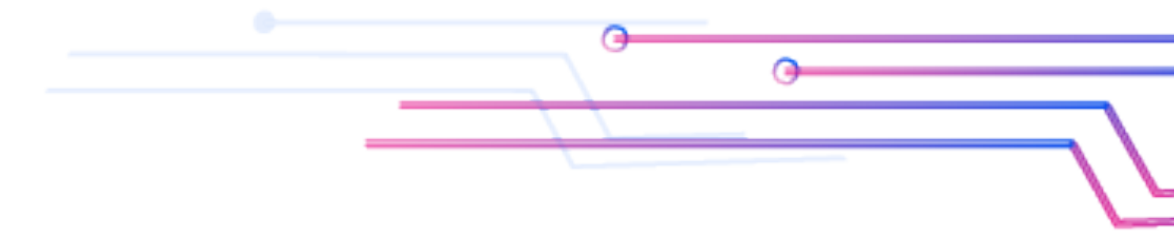
v1 – in operation

v2 – currently in development



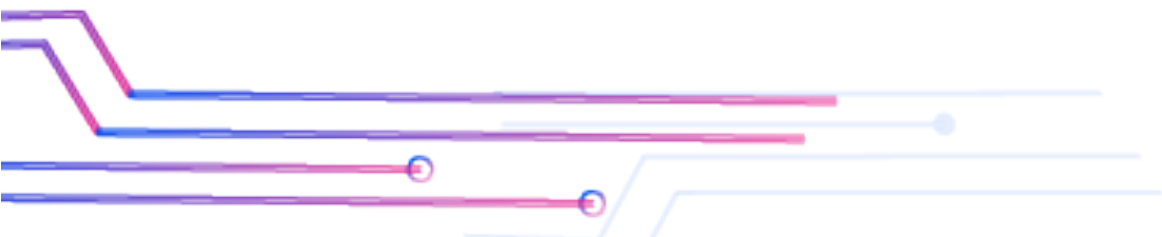


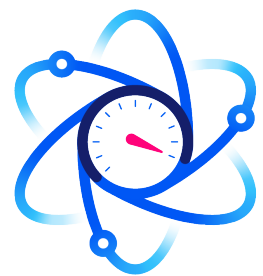
QruiseOS powering FZ Jülich QC



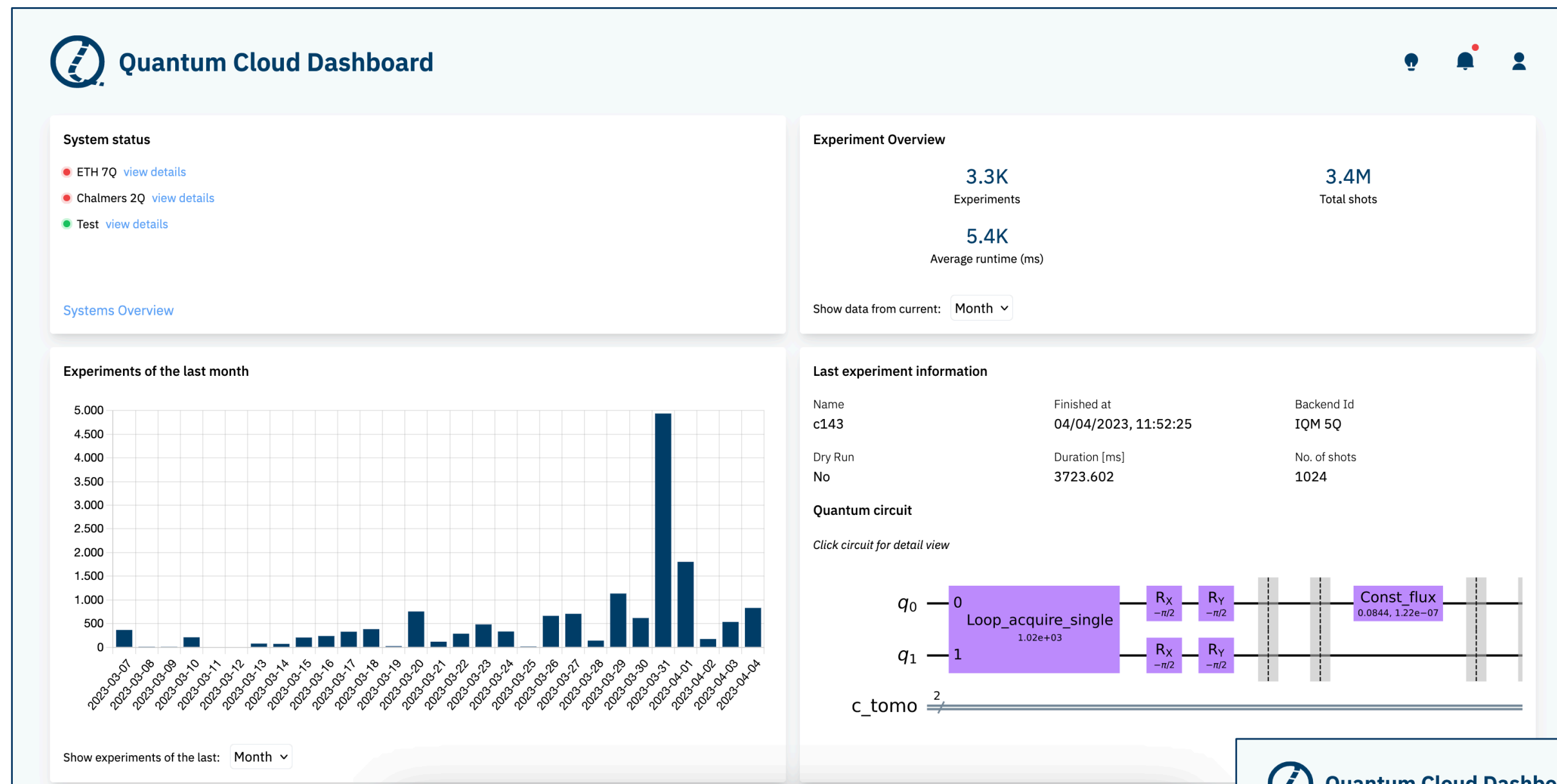
Digital Twin Characterization of Tunable Qubit – Tunable Coupler Architecture

- Frequencies and couplings for all components (including flux tunability)
- Control line transfer functions, amplitude non-linearities and cross-talks
- Readout confusion matrix
- Initialization errors / temperatures
- Lindblad channels and non-Markovian noise spectra, allowing for robust and DD gates
- Spurious component couplings



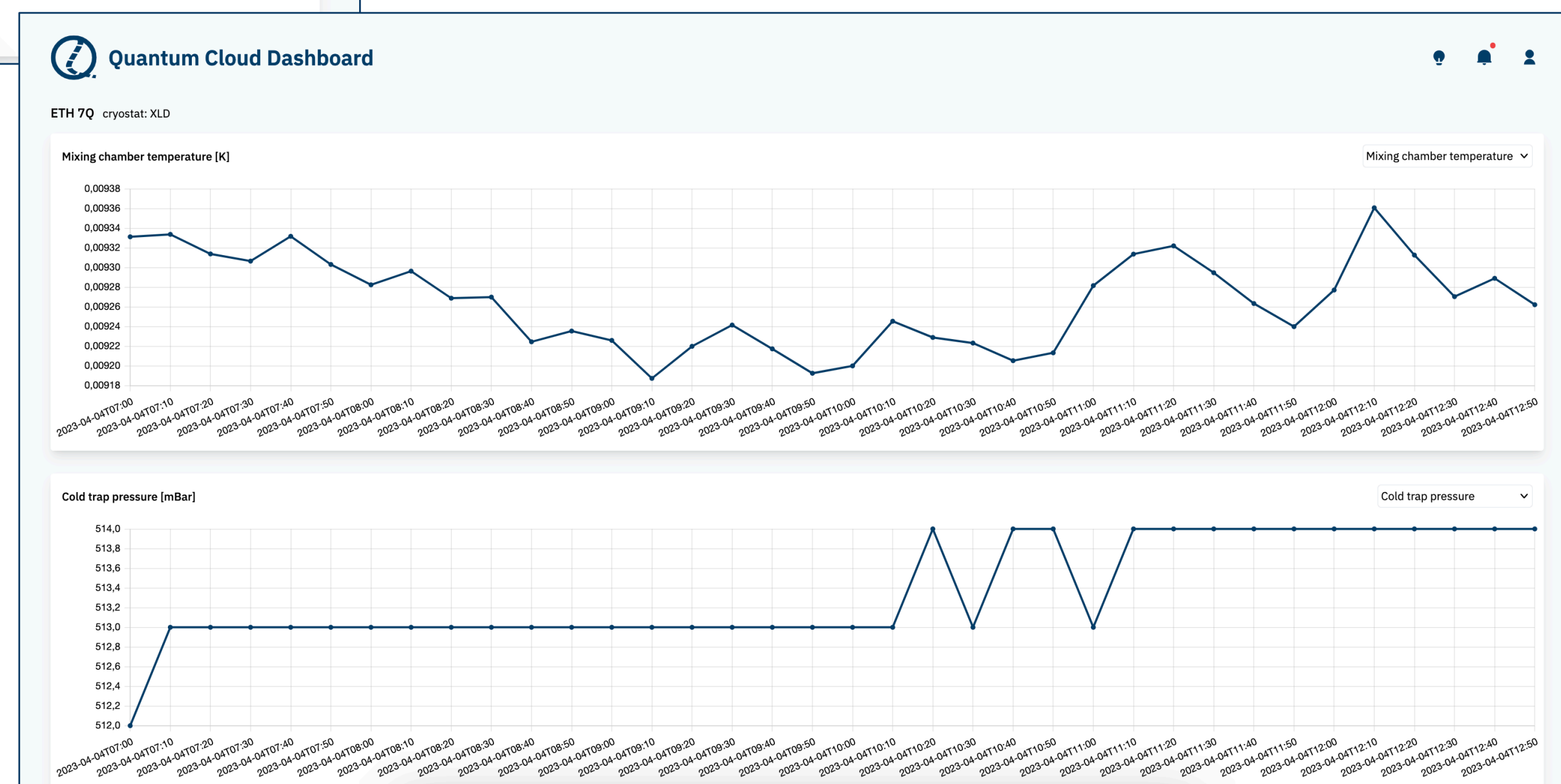


QruiseOS powering FZ Jülich QC



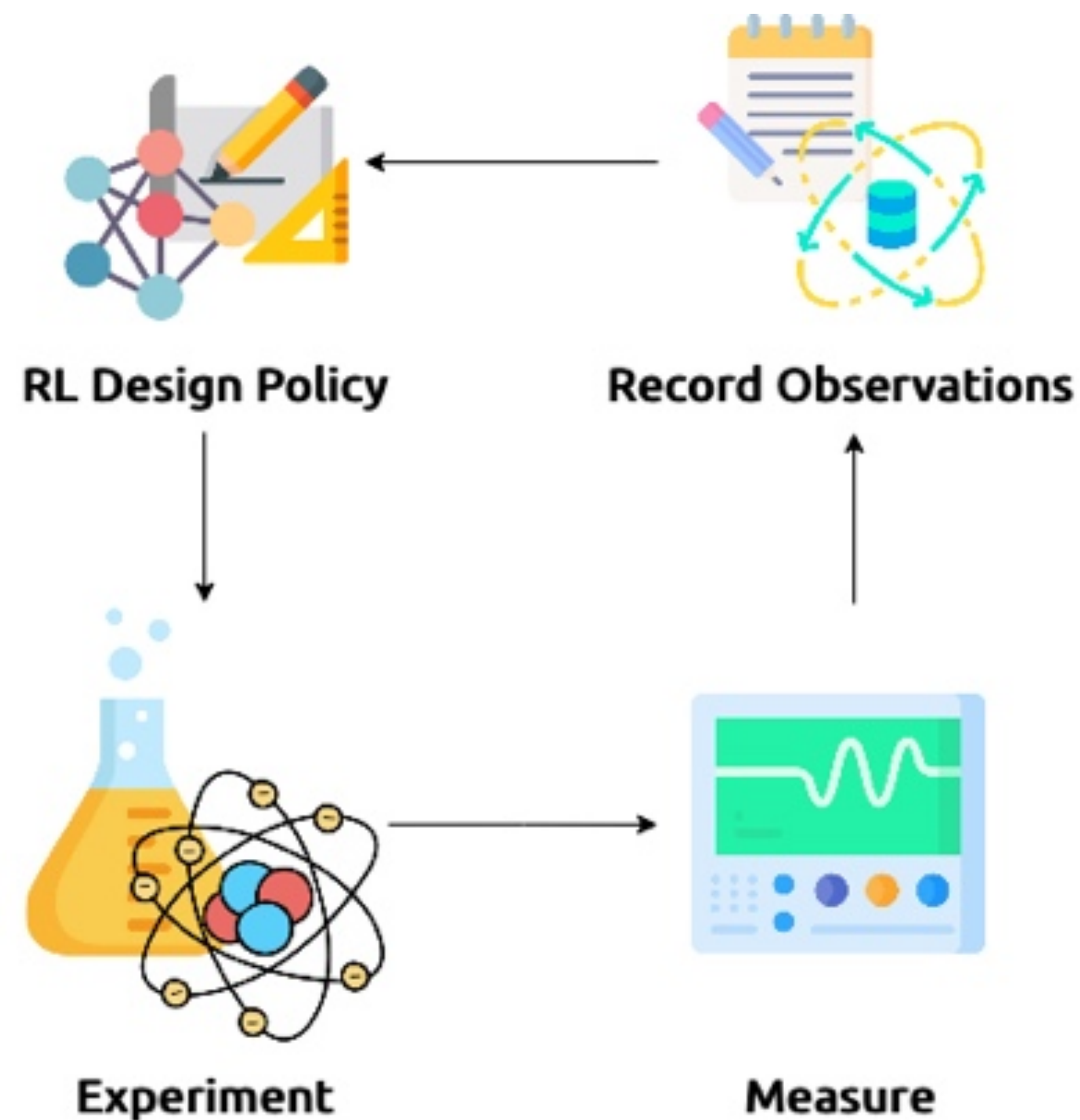
Live dashboard for tracking quantum device usage

Monitor and Analyse
Live & Historical data
from Essential Infra





Bayesian Adaptive Experimental Design



Adaptive Characterization and
Calibration of quantum hardware
based on Statistical Learning

Given a system and its model, devise the optimal set of experiments for
constraining arbitrary model parameter(s) within a given tolerance bound.

Closing the Loop – for Large Scale Noisy devices

qruise - One slide company overview

- Spinoff from FZJ & Padua U. via Helmholtz Validation Project
- Founded and funded in late 2021.
- 15 people and growing.
- Multiple research grants, including EIC Transition (2.5M EUR).
- 3 year runway.
- Initial focus was quantum control for q. computation.
Superconducting, Rydberg, Ions, NV centers, ...
- Expanded to q. sensing: NMR, atom interf., NMR, optical clocks
- Expanded further to silicon photonics
- Expanding ambition to building a general-purpose ML Physicist
- Looking for physicists with ML experience of visa-versa

