Quantum state-agnostic work extraction (almost) without dissipation

Josep Lumbreras   Ruo Cheng Huang   Yanglin Hu

Marco Tomamichel   Mile Gu

National University of Singapore

Centre for Quantum Technologies

Nanyang Technological University Singapore

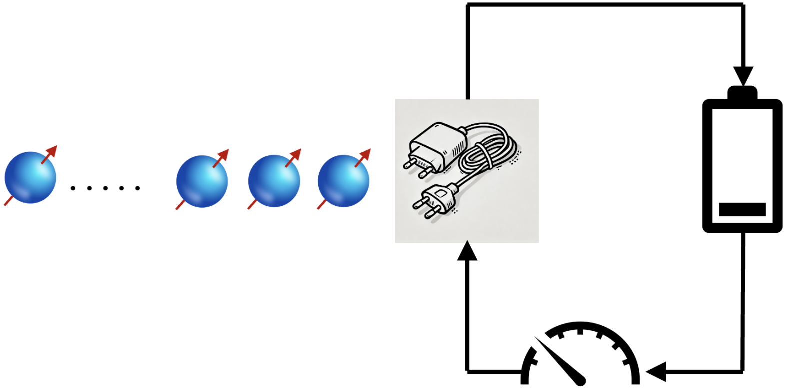
Quantum Information Theory

Nanyang Quantum Hub

Quantum Resources 2025, JeJu

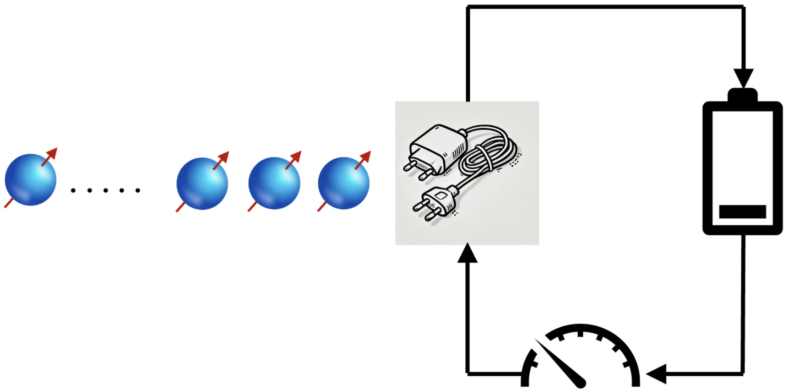
17$^{\text{th}}$ March

The problem



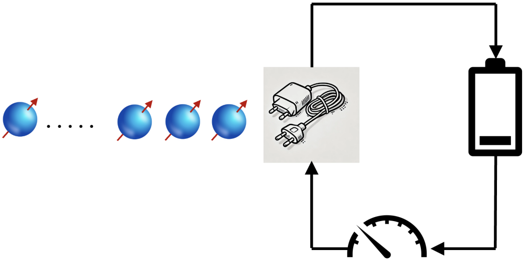
Given sequential access to **finite copies** of an identical, unknown quantum system, what is the optimal approach to extract work from these systems and charge a battery?

General setup



- N copies of **unknown** quantum state $|\psi\rangle$.
- Charging protocol for **single copy**.
- Battery system.
- Measurement feedback.

Main result



- Lack of knowledge of state results in suboptimal extraction of work, leading to dissipation.
- Simple approach: state tomography+extraction, $O(\sqrt{N})$ cumulative dissipation
- Our approach shows how to achieve $O(\log N)$ cumulative dissipation.

Outline

1. Extracting work from knowledge (JC battery)
2. Work dissipation and the exploration-exploitation tradeoff
3. Quantum state tomography under minimal regret
4. Extracting work from thermal bath

Extracting work from knowledge

- consider a source producing qubits in an unknown pure state ψ
- want to learn ψ, but also extract work using partial information
- expecting the state is $\hat{\psi}$, we engineer an interaction that raises a battery system with probability $|\langle\psi|\hat{\psi}\rangle|^2$
- binary reward (charge or not):

$$r_t = \begin{cases} 1 & \text{w.p.} \quad |\langle\psi|\hat{\psi}\rangle|^2 \\ 0 & \text{w.p.} \quad 1 - |\langle\psi|\hat{\psi}\rangle|^2 \end{cases}$$

Jaynes-Cummings battery

Battery system described by $H_B = \omega a^\dagger a$

For $k = 1, 2, ..., N$:

1. Receive unknown $|\psi\rangle$, make a guess $|\psi_k\rangle$, battery state known $|n_k\rangle$.
2. Expose $|\psi\rangle$ to a field that induces a Hamiltonian $H_A = \omega |\psi_k\rangle\langle\psi_k|$.
3. The interaction between the battery and the particle is described by an interaction Hamiltonian

$$H_I = \frac{\Omega}{2}(a \otimes |\psi_k\rangle\langle\psi_k^\perp| + a^\dagger \otimes |\psi_k^\perp\rangle\langle\psi_k|), \tag{1}$$

   that we turn on for a time $t_k = \pi\Omega^{-1}(n_k + 1)^{-\frac{1}{2}}$.

4. Measure the energy of the battery in its energy eigenbasis and update the energy n_{k+1}.

Work dissipation

- The extracted work is defined as $\Delta W_k = \omega(n_{k+1} - n_k)$.
- The expected extracted work is given by

$$\mathbb{E}[\Delta W_k] \leq 2\omega(|\langle\psi_k|\psi\rangle|^2$$

- The dissipation in this round is

$$W_{\text{diss}}^{\text{jc},k} = \max_{|\psi_k\rangle} \mathbb{E}[\Delta W_k] - \mathbb{E}[\Delta W_k] \leq 2\omega(1 - |\langle\psi_k|\psi\rangle|^2).$$

Goal

Minimize the cumulative dissipation over N rounds

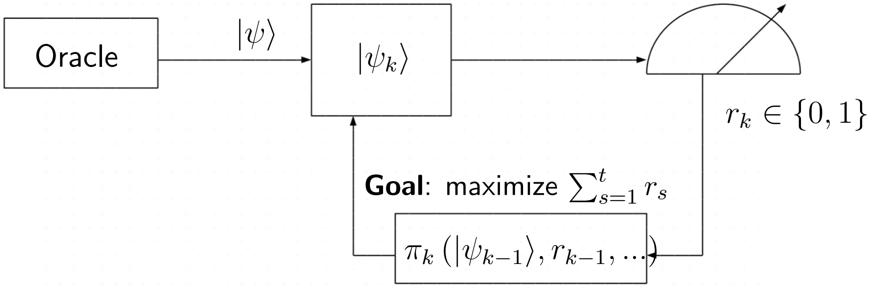
$$W_{\textit{diss}}^{\textit{jc}}(N) = \sum_{k=1}^{N} W_{\textit{diss}}^{\textit{jc},k} \leq 2\omega \sum_{k=1}^{N} (1 - |\langle\psi_k|\psi\rangle|^2).$$

Quantum state tomography under minimal regret

At each round $k \in \{1, 2, ..., N\}$:

- Learner receives unknown $|\psi\rangle$ (fixed, same each round).
- Learner uses policy π_t, chooses probe state $|\psi_k\rangle$ (adaptively) and performs single copy measurement on $|\psi\rangle$ on the direction of $|\psi_k\rangle$.
- Learner receives reward sampled according to Born's rule

$$r_k = \begin{cases} 1 & \text{w.p.} \quad |\langle\psi|\psi_t\rangle|^2 \\ 0 & \text{w.p.} \quad 1 - |\langle\psi|\psi_k\rangle|^2 \end{cases}$$



Quantum state tomography under minimal regret

We want to minimize

$$\mathrm{Regret}(T) = \sum_{t=1}^{T} 1 - \langle \psi | \Pi_t | \psi \rangle$$

- **Tomography**: Obtain accurate estimates of $|\psi\rangle$ such that $1 - \langle \psi | \Pi_t | \psi \rangle$ is small.
- **Reinforcement learning**: Finding a balance between **exploration-exploitation**:
  - Exploration: selecting Π_t that give enough information to estimate $|\psi\rangle$.
  - Exploitation: selecting Π_t close to $|\psi\rangle$ such that minimizes regret.

Exploration-Exploitation

- Many real-life problems can be formulated as an exploration exploitation dilemma.
- Movie recommendation, web advertisement, etc.
- Fundamental problem in reinforcement learning.
- Formalized in the multi-armed bandit framework.

Extracting work after learning

The simplest strategy is

- First αN $(0 \leq \alpha \leq 1)$ copies for learning $|\psi\rangle$ and get estimate $|\hat{\psi}\rangle$
- Last $(1-\alpha)N$ copies fix $|\psi_k\rangle = |\hat{\psi}\rangle$.

This protocol achieves

$$W_{\mathsf{diss}}^{\mathsf{jc}}(N) = O\left(\omega\alpha N + \omega(1-\alpha)(1-\mathbb{E}[\langle\hat{\psi}|\psi\rangle|^2]N\right)$$

Optimal state tomography achieves

$$1-\mathbb{E}[\langle\hat{\psi}|\psi\rangle|^2 \sim \frac{1}{\alpha N}$$

Optimizing over α we get

$$W_{\mathsf{diss}}^{\mathsf{jc}}(N) = O(\omega\sqrt{N})$$

Can we improve $W^{\mathsf{jc}}_{\mathsf{diss}}(N) = O(\omega\sqrt{N})$?

Main result

Theorem 1

There exists a protocol that achieves with probability at least $1-\delta$

$$W^{jc}_{diss}(N) = O(\omega \log(N) \log(N/\delta)).$$

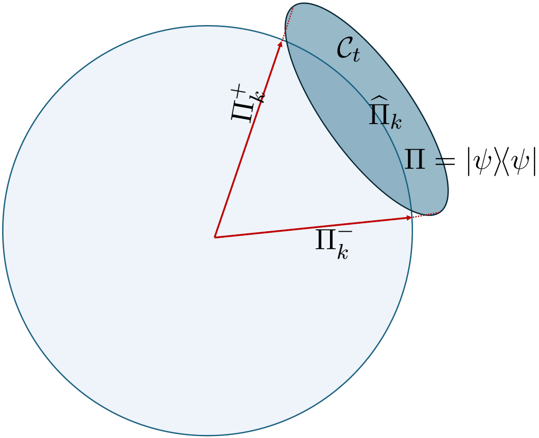
- Proof is constructive: we design and analyse the protocol.
- Almost fully adaptive: uses $T/\log(T)$ rounds of adaptation.

Extracting work while learning

Our protocol is built using:

- **Optimisitc principle:**
  - We construct confidence region $\mathcal{C}_t$ around $|\psi\rangle$ and update $\Pi_t \in \mathcal{C}_t$.
  - Our particular update allows to control the exploration-exploitation at any $t \in [T]$.
- **Median of means weighted online least squares estimator (MoMWLSE):**
  - $\mathcal{C}_t$ is built around an online least-squares estimator.
  - We use weighted version of LSE to put more weight on measurements with low variance outcomes Quantum Part!.
  - The weighted version introduces unbounded random variables $\Rightarrow$ need median of means version (different from classical shadows).

Extracting work while learning



Select measurements $\Pi_t^{\pm}$ along the directions of maximal uncertainty of confidence region.

Median of Means

- For each action $\Pi_s^{\pm}$ we perform $l \sim \log(N)$ independent measurements and construct the following $j \in [l]$ estimators

$$\widetilde{\theta}^{\mathsf{w}}_{k,j} = V_k^{-1} \sum_{s=1}^{k} \frac{1}{\hat{\sigma}_s^2} (a^+_{s,i} r^+_{s,i,j} + a^-_{s,i} r^-_{s,i,j})$$

$$V_k = V_{k-1} + \frac{1}{\hat{\sigma}_k^2} \left(a^+_{k,i} (a^+_{k,i})^{\mathsf{T}} + a^-_{k,i} (a^-_{k,i})^{\mathsf{T}} \right) \quad a^{\pm} \text{ bloch vectors.}$$

- The median of means is defined as (different from classical shadows)

$$\widetilde{\theta}^{\mathsf{wMoM}}_k := \tilde{\theta}^{\mathsf{w}}_{k,l^*} \quad \text{where } k^* = \arg\min_{j \in [l]} y_j,$$

where

$$y_j = \text{median}\{\|\tilde{\theta}^{\mathsf{w}}_{t,j} - \tilde{\theta}^{\mathsf{w}}_{t,i}\|_{V_t} : i \in [l]/j\} \quad \text{for } j \in [l].$$

Confidence region

- The MoM LSE $\widetilde{\theta}_t^{\mathsf{wMoM}}$ defines a confidence region

$$\Pr\left(\theta \in \mathcal{C}_s, \forall s \in [k]\right) \geq 1 - \delta \quad \delta \in (0,1)$$
$$\mathcal{C}_k = \{\theta' \in \mathbb{R}^d : \|\theta' - \tilde{\theta}_k\|_{V_k}^2 \leq \mathsf{poly}(d) \log(1/\delta)\},$$

  if $\hat{\sigma}_k^2$ overestimates the variance of r_s i.e

$$\mathrm{Var}(r_s) = |\langle\psi|\psi_s\rangle|^2 \left(1 - |\langle\psi|\psi_s\rangle|^2\right) \leq \hat{\sigma}_s^2.$$

- It suffices to choose $\hat{\sigma}_s^2 \sim 1/\lambda_{\min}(V_{s-1})$

Extracting work with thermal reservoir

Unknown state $|\psi\rangle$ in a degenerate Hamiltonian $H_A = w\mathbf{1}/2$

Battery system described by $H_B = \int \mu \, |\mu\rangle\langle\mu| \, d\mu$

Tunable thermal reservoir $H_R(\nu) = \nu \, |1\rangle\langle 1|$, defined by a freely-chosen energy gap ν

For $k = 1, 2, ..., N$:

1. Receive unknown $|\psi\rangle$, make a guess ρ_k, battery state known $|\mu_k\rangle$.
2. Apply unitary (depends on ρ_k) to rotate $|\psi\rangle$ to energy eigenstate of H_A.
3. a series of SWAP operation with tailored reservoir state, energetic changes stored to battery system.
4. Measure the energy of the battery in its energy eigenbasis, obtain μ_{k+1}

Dissipation with thermal reservoir

- The extracted work is defined as $\Delta W_k = (\mu_{k+1} - \mu_k)$.
- The expected extracted work is given by

$$\mathbb{E}[\Delta W_k] = \beta^{-1} \left[\mathsf{D}\left(\psi \| \mathbf{1}/2\right) - \mathsf{D}(\psi || \rho_k) \right]$$

- The dissipation in this round is

$$W_{\mathsf{diss}}^{\mathsf{sc},k} = \max_{\rho_k} \mathbb{E}[\Delta W_k] - \mathbb{E}[\Delta W_k] = \beta^{-1} \mathsf{D}(\psi || \rho_k).$$

Theorem 2

There exists protocol that achieves, with probability at least $1 - \delta$

$$W_{diss}^{sc}(N) = O\left(\beta^{-1} \log^2(N) \log\left(\frac{N}{\delta}\right)\right). \quad (2)$$

Conclusions

- We study cumulative dissipation with finite copies.
- We link the problem to the exploration-exploitation dilemma.
- We introduce a protocol that both learns and charges optimally a battery.
- Can we apply similar ideas to the extraction of other resources in the finite copy regime?

The talk is based on:

- with Ruo Cheng Huang, Yanglin Hu, Marco Tomamichel and Mile Gu: *Quantum state-agnostic work extraction (almost) without dissipation* (soon `arXiv`).
- with Mikhail Terekhov and Marco Tomamichel: *Learning pure quantum states (almost) without regret*, `arXiv: 2406.18370` (algorithm)