Many-body Quantum Score:
A scalable benchmark for digital and analog QPUs and first results on a Pasqal device

Harold Erbin

IPHT, CEA Saclay (France)

TQCI International Conference – 24 June 2025

Collboration CEA-IPHT/Eviden

Outline: 1. Motivations

Why benchmarking?

1. (relatively) objective comparison between different providers
   - capabilities and performance
   - match platforms and tasks

Why benchmarking?

1. (relatively) objective comparison between different providers
   - capabilities and performance
   - match platforms and tasks
2. clear targets for product development and quality control
   - identify bottlenecks or failing components
   - suggest design improvements and validate upgrades

Why benchmarking?

1. (relatively) objective comparison between different providers
   - capabilities and performance
   - match platforms and tasks
2. clear targets for product development and quality control
   - identify bottlenecks or failing components
   - suggest design improvements and validate upgrades
3. improve transparency and trust
   - increase credibility by publishing benchmark results
   - strengthen customer confidence
   - help users make informed decisions

Why benchmarking?

1. (relatively) objective comparison between different providers
   - capabilities and performance
   - match platforms and tasks
2. clear targets for product development and quality control
   - identify bottlenecks or failing components
   - suggest design improvements and validate upgrades
3. improve transparency and trust
   - increase credibility by publishing benchmark results
   - strengthen customer confidence
   - help users make informed decisions
4. standardization
   - better interoperability
   - set and follow industry standards
   - more efficient decision-making

Properties of a good benchmark

1. Relevance: should measure important features
2. Representativeness: metrics should be broadly accepted by industry and academia
3. Equity: all systems should be fairly compared, and metrics specialized to a technology only if not relevant for others
4. Repeatability: results should be unambiguous and verifiable
5. Cost-effectiveness: tests should be resource-efficient
6. Scalability: metrics should be computable for systems of different sizes
7. Transparency: metrics should be readily understandable and easy to compare, and informative for non-experts

Inspired from [1907.03626, Dai-Berleant] (see also [2407.08828, Proctor et al., 2406.03905, Mesman et al.])

Levels of characterization

1. component-level
   - physical qubits
   - single- and multi-qubit gates / operations
2. system-level
   - circuit
   - logical qubits (error correction. . .)
   - architecture design (firmware. . .)
3. application-level
   - algorithmic (subroutines)
   - end-user cases
   - test suites
   - frameworks

Analog- and digital-compatible benchmark

Objective [WIP, CEA-IPHT/Eviden]

Find a scalable benchmark for analog and digital QPUs.

Analog- and digital-compatible benchmark

Objective [WIP, CEA-IPHT/Eviden]

Find a scalable benchmark for analog and digital QPUs.

Challenges for analog QPU

- not universal $\Rightarrow$ less abstraction and more technology-dependent
- continuous process instead of discrete operations
- many-body effective model (spin chains. . .)

Analog- and digital-compatible benchmark

Objective [WIP, CEA-IPHT/Eviden]

Find a scalable benchmark for analog and digital QPUs.

Challenges for analog QPU

- not universal ⇒ less abstraction and more technology-dependent
- continuous process instead of discrete operations
- many-body effective model (spin chains. . .)

⇒ application benchmark (BACQ) based on many-body dynamics

- applications in condensed matter and quantum chemistry
- one of the best place to look for exponential advantage

Here: motivate protocol using Rydberg QPU

Outline: 2. Benchmark protocol

Rydberg spin model

- N spins $1/2$ on a $2d$ lattice
- effective Hamiltonian

$$H_{\mathrm{RI}} \sim \underbrace{\sum_{i<j} \frac{C_6}{\boldsymbol{r}_{ij}^6} n_i n_j}_{\text{spin interaction}} + \underbrace{\frac{\hbar\Omega(t)}{2} \sum_i \sigma_i^x}_{\text{transversal}} - \underbrace{\hbar\delta(t) \sum_i n_i}_{\text{longitudinal}}$$

$n_i = (\sigma_i^z + 1)/2$

Rydberg spin model

- N spins $1/2$ on a $2d$ lattice
- effective Hamiltonian

$$H_{\mathrm{RI}} \sim \underbrace{\sum_{i<j} \frac{C_6}{\boldsymbol{r}_{ij}^6} n_i n_j}_{\text{spin interaction}} + \underbrace{\frac{\hbar\Omega(t)}{2} \sum_i \sigma_i^x}_{\text{transversal}} - \underbrace{\hbar\delta(t) \sum_i n_i}_{\text{longitudinal}}$$

$$\xrightarrow[\substack{\delta\to\delta_* \\ r_{ij}\gg 1}]{} J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z - Jg \sum_i \sigma_i^x$$

$n_i = (\sigma_i^z + 1)/2$

- recover (approximately) nearest-neighbor transverse-field Ising model (TFIM)

Rydberg spin model

- N spins $1/2$ on a $2d$ lattice
- effective Hamiltonian

$$H_{\mathrm{RI}} \sim \underbrace{\sum_{i<j} \frac{C_6}{\boldsymbol{r}_{ij}^6} n_i n_j}_{\text{spin interaction}} + \underbrace{\frac{\hbar\Omega(t)}{2} \sum_i \sigma_i^x}_{\text{transversal}} - \underbrace{\hbar\delta(t) \sum_i n_i}_{\text{longitudinal}}$$

$$\xrightarrow[\substack{\delta \to \delta_* \\ r_{ij} \gg 1}]{} J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z - Jg \sum_i \sigma_i^x$$

$n_i = (\sigma_i^z + 1)/2$

- recover (approximately) nearest-neighbor transverse-field Ising model (TFIM)
- critical point: $g = 1$

Many-body quantum score (1)

MBQS test protocol

1. Setup a spin chain with L spin-$\frac{1}{2}$ equally spaced on a $1d$ ring.
2. Initialize the register with the state $|+\cdots+\rangle$ (σ_i^x eigenstate).
3. Evolve the system (quench) with the Ising Hamiltonian at the critical point for a duration $t_*(L)$ ("peak time").
4. Perform measurements $\{\sigma_i^z\}$ and compute the connected 2-point functions

$$g_\ell^{(2)}(t) := \langle \sigma_1^z \sigma_\ell^z \rangle_c := \langle \sigma_1^z \sigma_\ell^z \rangle_c - \langle \sigma_1^z \rangle \langle \sigma_\ell^z \rangle$$

5. Compute the score function (average relative error with respect to the theoretical values in Ising model)

$$P_2(L) := \frac{1}{\lfloor L/2 \rfloor - 1} \sum_{\ell=2}^{\lfloor L/2 \rfloor} \left| \frac{g_\ell^{(2)\mathrm{exp}}(t_*) - g_\ell^{(2)\mathrm{th}}(t_*)}{g_\ell^{(2)\mathrm{th}}(t_*)} \right|$$

Many-body quantum score (2)

[WIP, CEA-IPHT/Eviden]

MBQS

The score S corresponds to the largest problem size L reached before failing the test with a threshold ϵ, but excluding system sizes below some cut-off:

$$S = L \quad \Longrightarrow \quad \forall L' \in [L_{\min}, L]: \quad P_2(L') \leq \epsilon.$$

Notes

- exclude small L because of geometrical effects
  $L_{\min} = 5$
- may perform readout mitigation

Motivations

- Ising model
  - exact solution with free fermions (Jordan–Wigner transformation) $\rightarrow$ compute properties in polynomial time
  - model close to Rydberg effective Hamiltonian but not identical

Motivations

- Ising model
  - exact solution with free fermions (Jordan–Wigner transformation) $\rightarrow$ compute properties in polynomial time
  - model close to Rydberg effective Hamiltonian but not identical
- peak time $t_*(L)$
  - all spins become highly correlated
  - increases with $L \Rightarrow$ difficulty increases (probe sensitivity to decoherence)
  - increases with interatomic distance
  - highly entangled state

Motivations

- Ising model
  - exact solution with free fermions (Jordan–Wigner transformation) $\rightarrow$ compute properties in polynomial time
  - model close to Rydberg effective Hamiltonian but not identical
- peak time $t_*(L)$
  - all spins become highly correlated
  - increases with $L \Rightarrow$ difficulty increases (probe sensitivity to decoherence)
  - increases with interatomic distance
  - highly entangled state
- critical point: non-trivial dynamics
  - interatomic distance $\approx$ Rydberg blockade radius
  - competition between different terms in Hamiltonian

Motivations

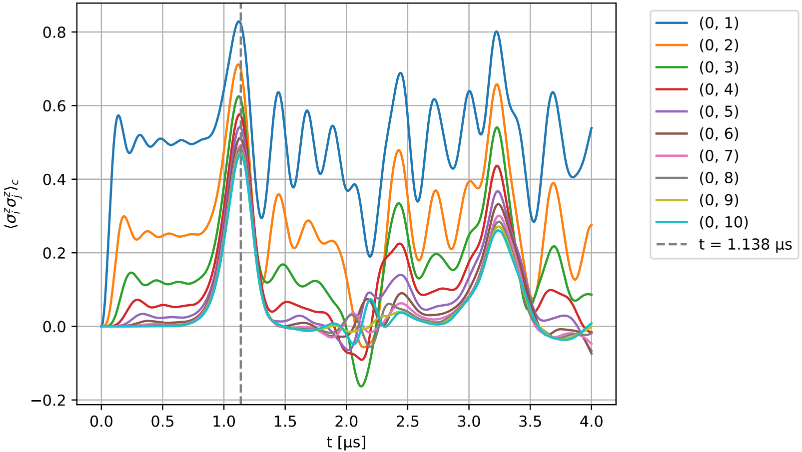
- Ising model
  - exact solution with free fermions (Jordan–Wigner transformation) $\rightarrow$ compute properties in polynomial time
  - model close to Rydberg effective Hamiltonian but not identical
- peak time $t_*(L)$
  - all spins become highly correlated
  - increases with $L \Rightarrow$ difficulty increases (probe sensitivity to decoherence)
  - increases with interatomic distance
  - highly entangled state
- critical point: non-trivial dynamics
  - interatomic distance $\approx$ Rydberg blockade radius
  - competition between different terms in Hamiltonian
- initial state $|+\cdots+\rangle$
  - more challenging for Rydberg QPU (requires pulse sequence to reach the state)
  - simpler theoretically and cleaner dynamics

Classical simulations

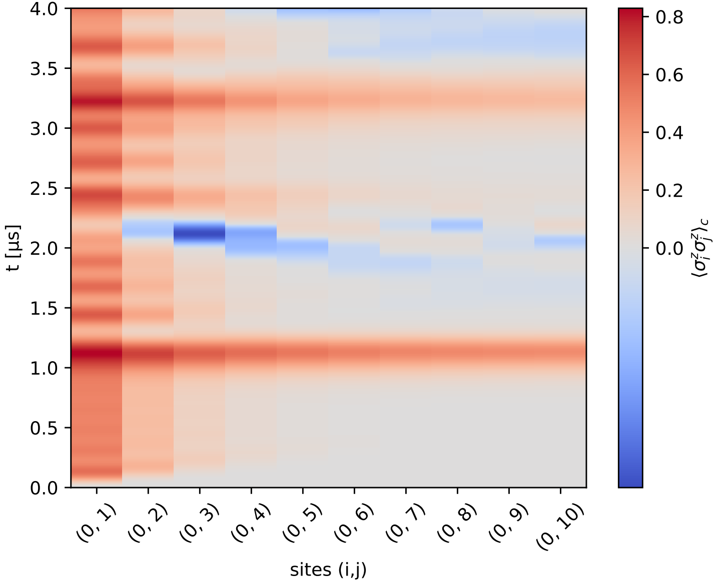
	ground state	dynamics
integrability (free fermions. . .)	easy	easy
MPS	easy	hard
brute-force	very hard	very hard

$\rightarrow$ use MPS to provide score of classical machines for comparison

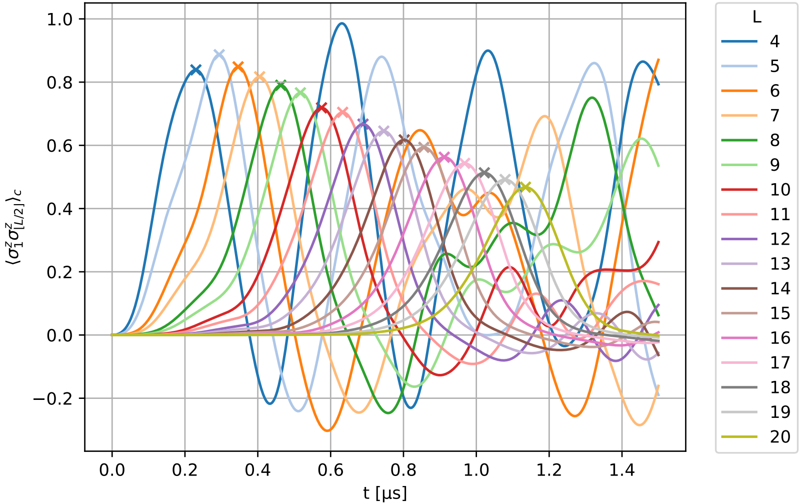
2-point correlation functions



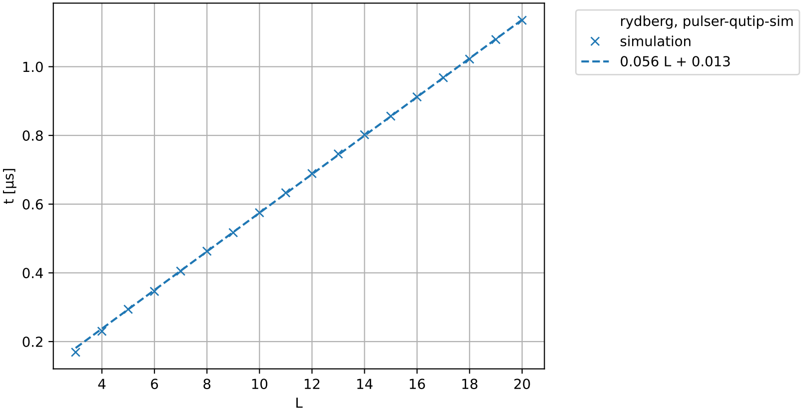
Information propagation



Antipodal correlations



Peak time prediction with linear regression



- $R^2 = 0.995$, RMSE $= 0.016$
- linear behavior $\rightarrow$ ballistic propagation

Outline: 3. Benchmark for Rydberg QPU

Short-term benchmark: `MBQSv0`

`MBQSv0`

Perform the `MBQS` protocol with the following modifications:

- Initial state: $|\downarrow \cdots \downarrow\rangle$ (σ_i^z eigenstate).
- Hamiltonian: Rydberg effective Hamiltonian.

Notes

- `MBQS` is too challenging for current QPU $\rightarrow$ propose a simpler intermediate benchmark
- no analytical solution because of long-range interactions $\rightarrow$ exact simulations: Pasqal's `pulser` + `qutip`

Experiments

- pulse sequence
  - constant Ω to reach critical point $g = 1$
  - constant δ to remove longitudinal field $h_{z,i} \approx 0$

$\rightarrow$ match Ising model up to small corrections (open/periodic bc)
$+$ boundary terms (open bc)

Experiments

- pulse sequence
  - constant Ω to reach critical point $g = 1$
  - constant δ to remove longitudinal field $h_{z,i} \approx 0$

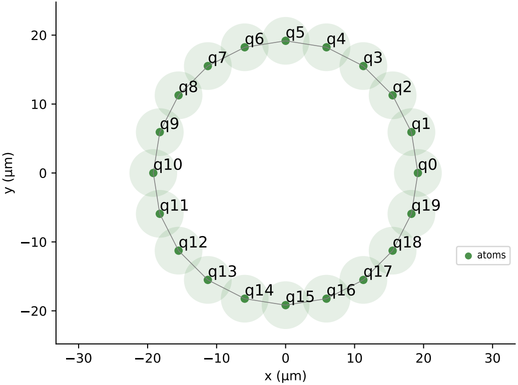
$\rightarrow$ match Ising model up to small corrections (open/periodic bc)
$+$ boundary terms (open bc)

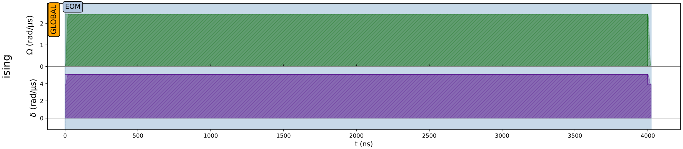
- implementations at TGCC
  - emulation with Eviden's QLM40
  - beta-test Pasqal's Ruby QPU (Orion Beta)

[Special thanks to people from Pasqal and TGCC]

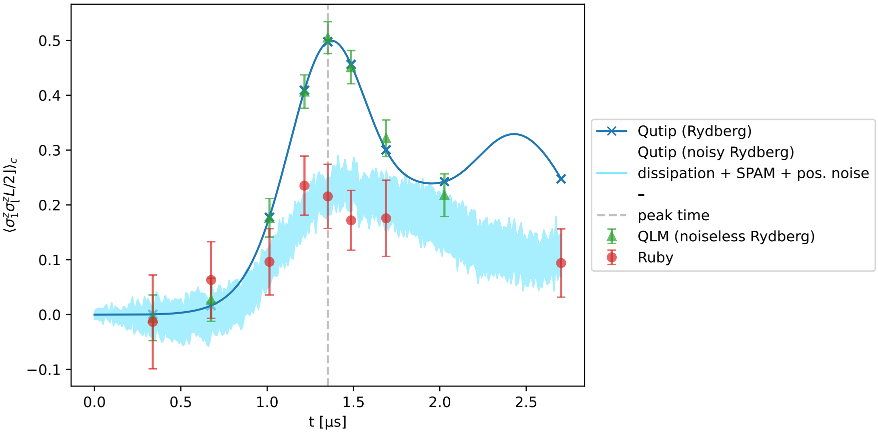
[Pasqal/Genci]

Atom register and pulses



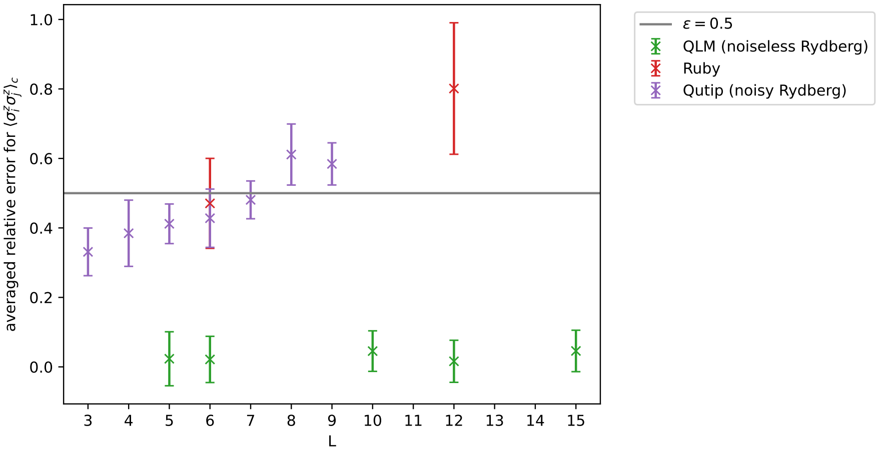


2-point correlation functions



⚠ Preliminary results on Ruby (still being commissioned)

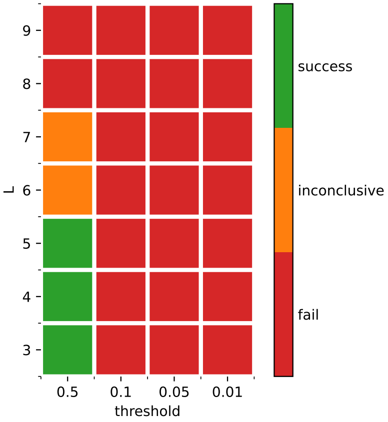
2-point correlation errors



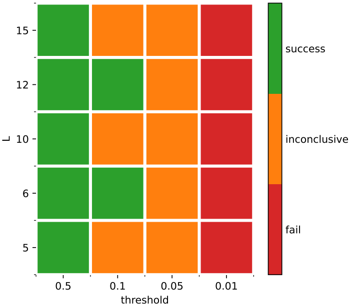
⚠ Preliminary results on Ruby (still being commissioned)

Volumetric success plot

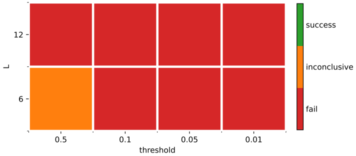
Qutip (noisy Rydberg)



QLM (noiseless Rydberg)



Ruby



⚠ Preliminary results on Ruby (still being commissioned)

Outline: 4. Conclusion

Summary and outlook

Achievements:

- benchmark protocol `MBQS` for analog and digital QPUs
- beta-test of the Ruby QPU at TGCC
- study in details the propagation of correlations in Ising and Rydberg spin chains
- participation to CEN/CENELEC and ISO/IEC standardization committees on quantum simulations and benchmarking

Summary and outlook

Achievements:

- benchmark protocol `MBQS` for analog and digital QPUs
- beta-test of the Ruby QPU at TGCC
- study in details the propagation of correlations in Ising and Rydberg spin chains
- participation to CEN/CENELEC and ISO/IEC standardization committees on quantum simulations and benchmarking

Future directions:

- open-source implementation and compare different platforms, including on gate-based device (trotterization)
- generalize benchmark to $2d$ geometries
- quantum control for Rydberg QPUs [WIP, Carrera-HE-Misguich]
- test many-body physics questions on Rydberg QPU