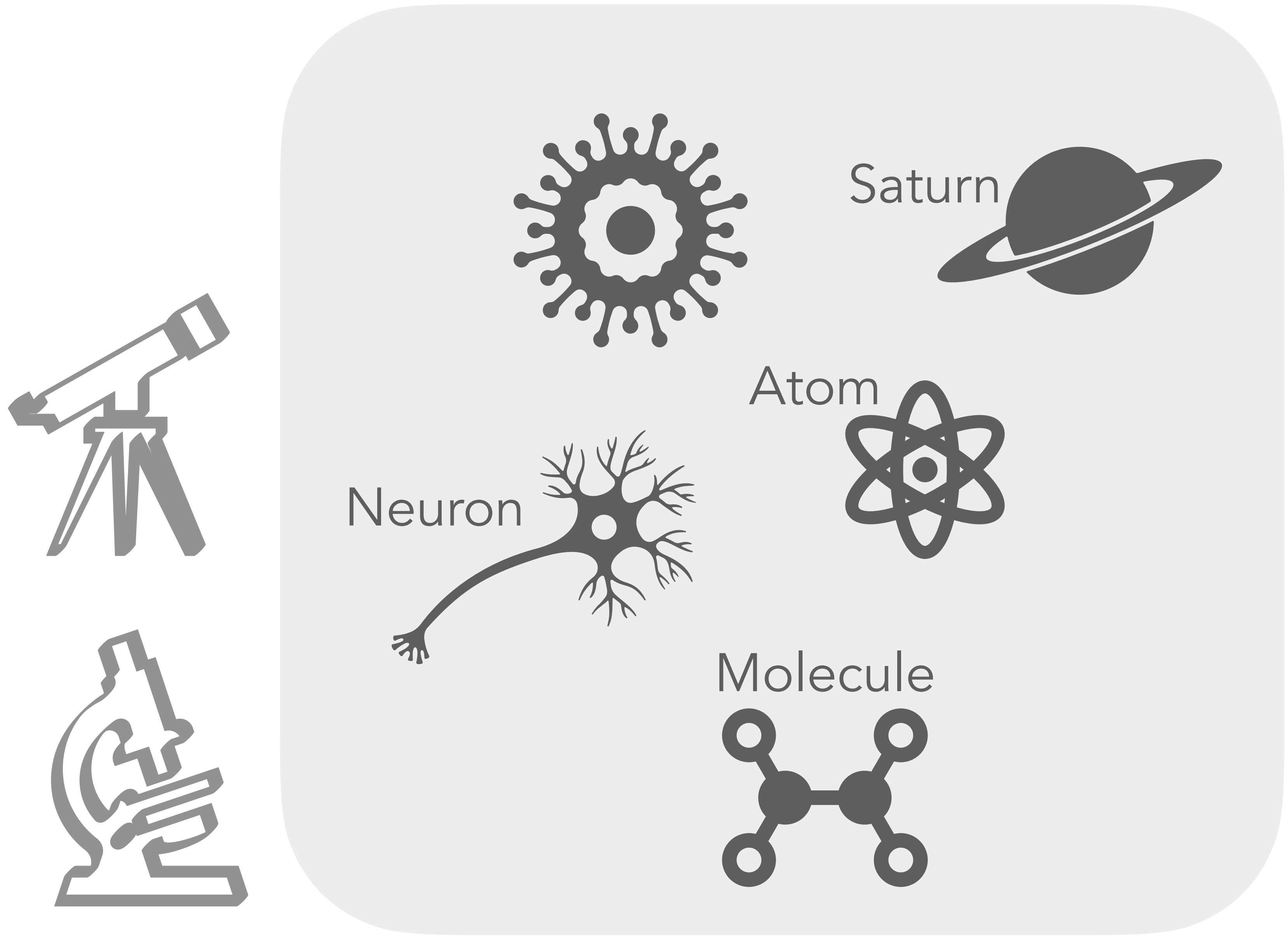
Learning from quantum experiments

Presenter: Hsin-Yuan (Robert) Huang
Caltech

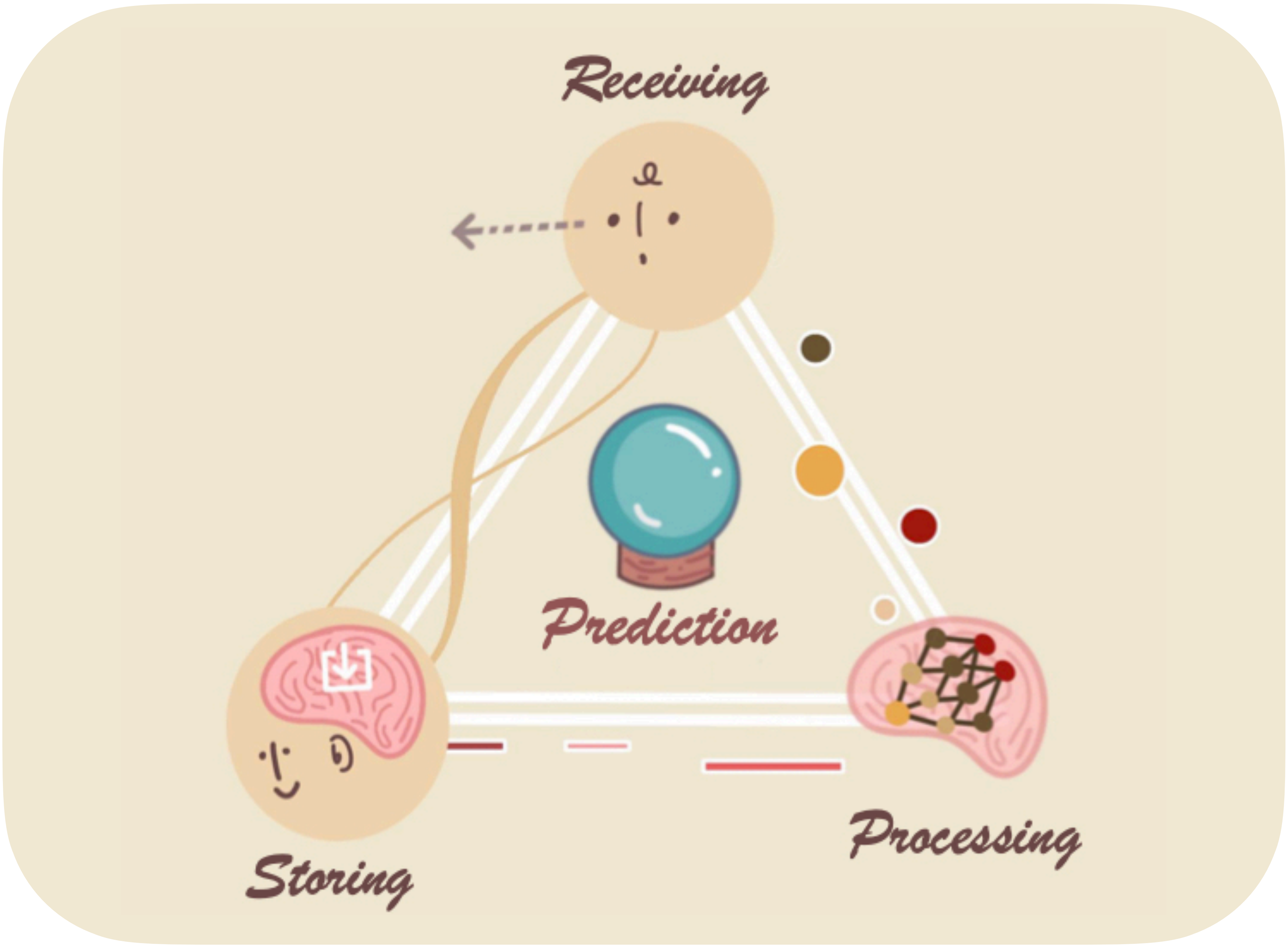
INSTITUTE FOR QUANTUM INFORMATION AND MATTER
Caltech

Motivation

- One of the central goals of science is to learn how the physical world operates.
- By performing experiments, humans can **receive information** about the physical world, and **process information** to form predictive models.



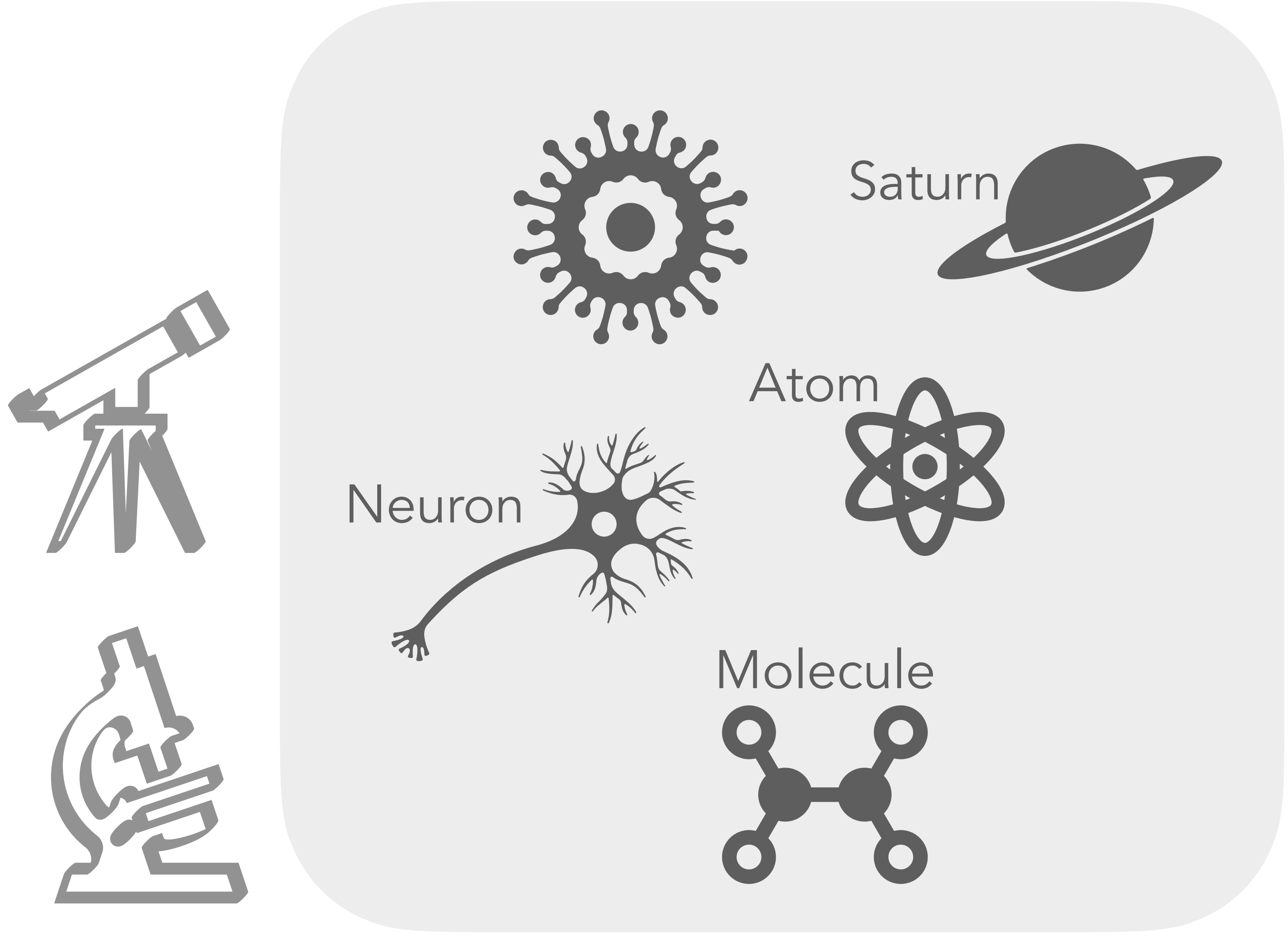
Examples of scientific disciplines



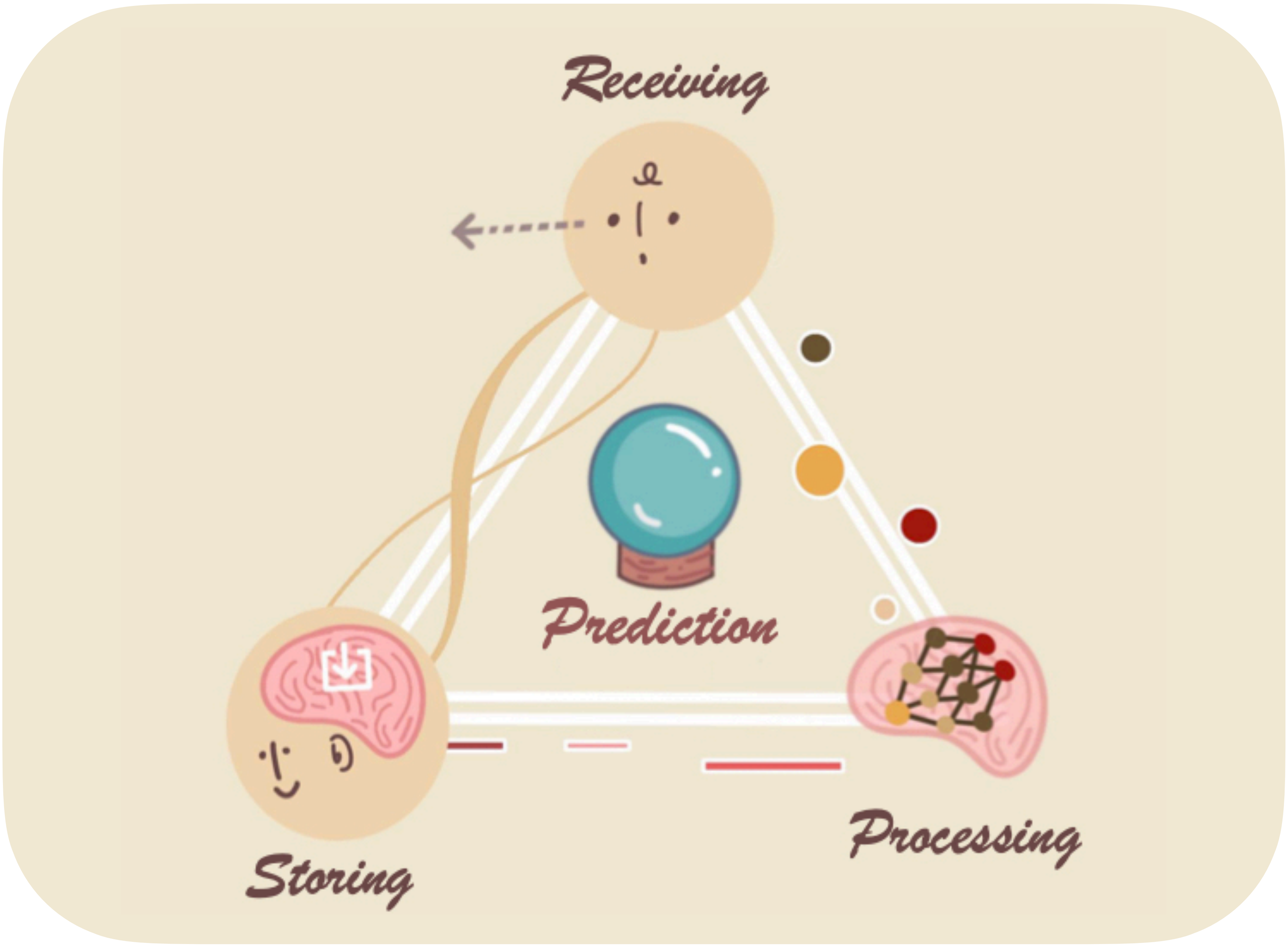
A cartoon depiction of learning

Motivation

- To accelerate and automate scientific development, it is important to understand how to **design better algorithms** to learn from experiments.
- A burgeoning field in QI considers the task of learning from **quantum experiments**.



Examples of scientific disciplines



A cartoon depiction of learning

Outline

- Basic setting and examples

- Key ideas:

  Part I — Designing good learning algorithms

  Part II — Proving no good learning algorithms exist

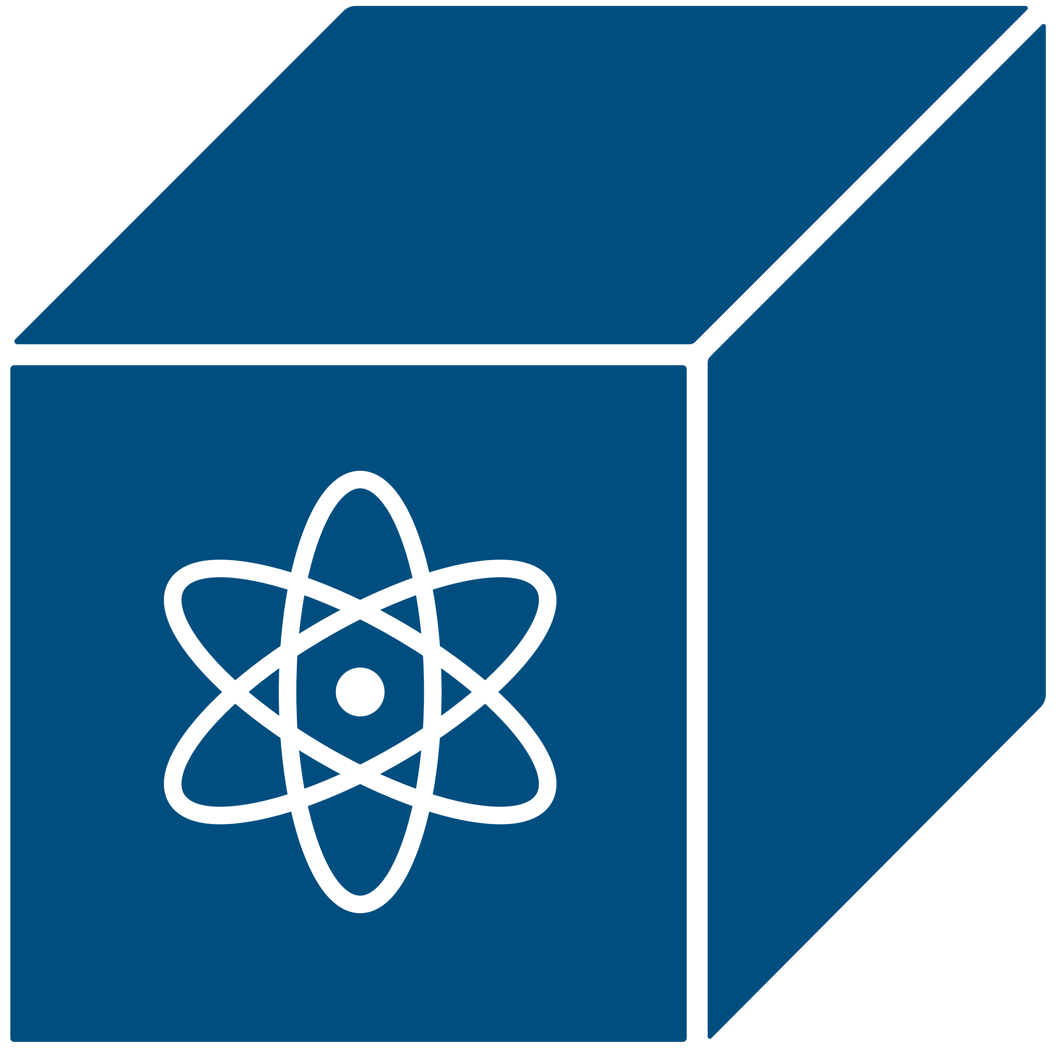
- Outlook and open questions

Outline

- Basic setting and examples
- Key ideas:
  Part I — Designing good learning algorithms
  Part II — Proving no good learning algorithms exist
- Outlook and open questions

Basic setting

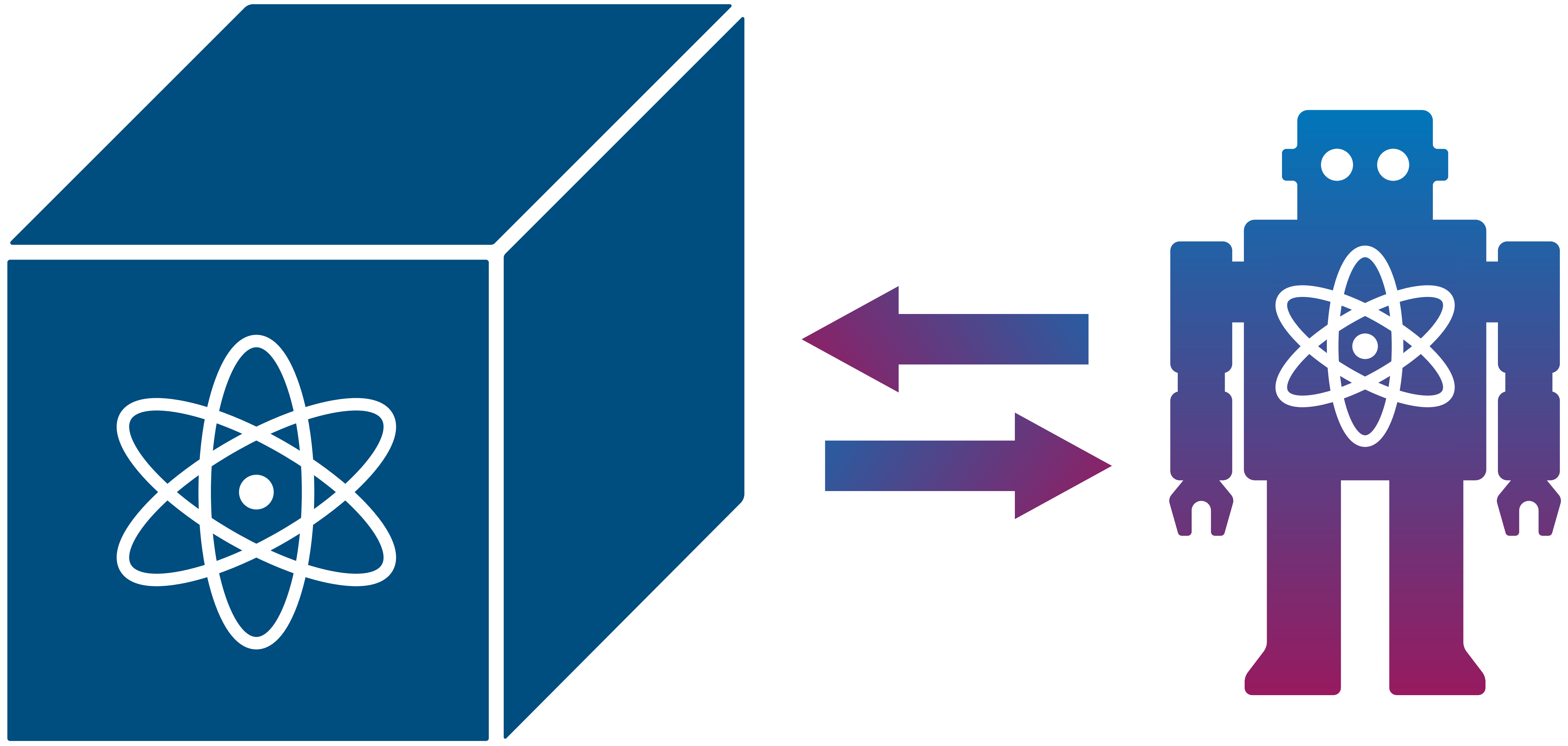
- There is an unknown quantum object (states, processes, entire phase diagram, …).
- Learn that object from experiments. So it becomes (approximately) known.



Unknown quantum object

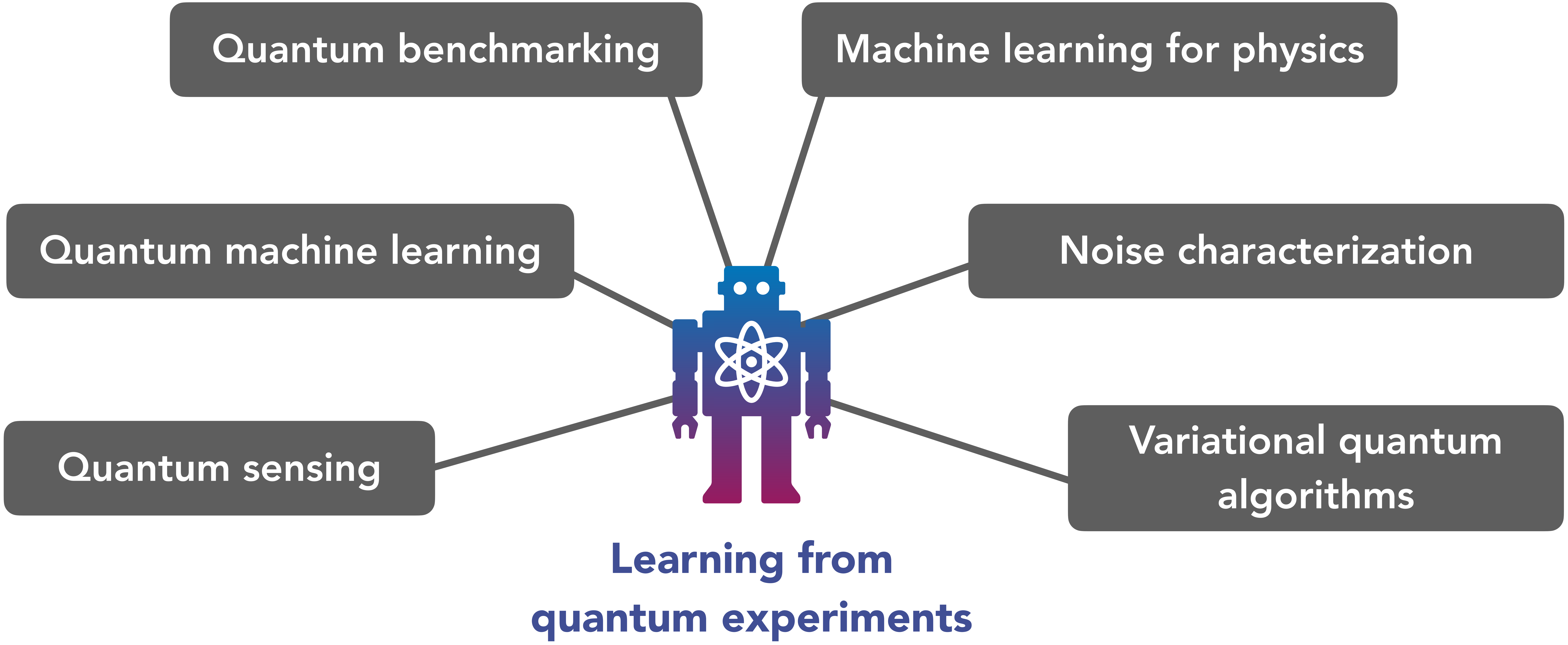
Basic setting

- There is an unknown quantum object (states, processes, entire phase diagram, …).
- Learn that object from experiments. So it becomes (approximately) known.
- How many experiments are needed? (Sample and query complexity)



Unknown quantum object

Overview



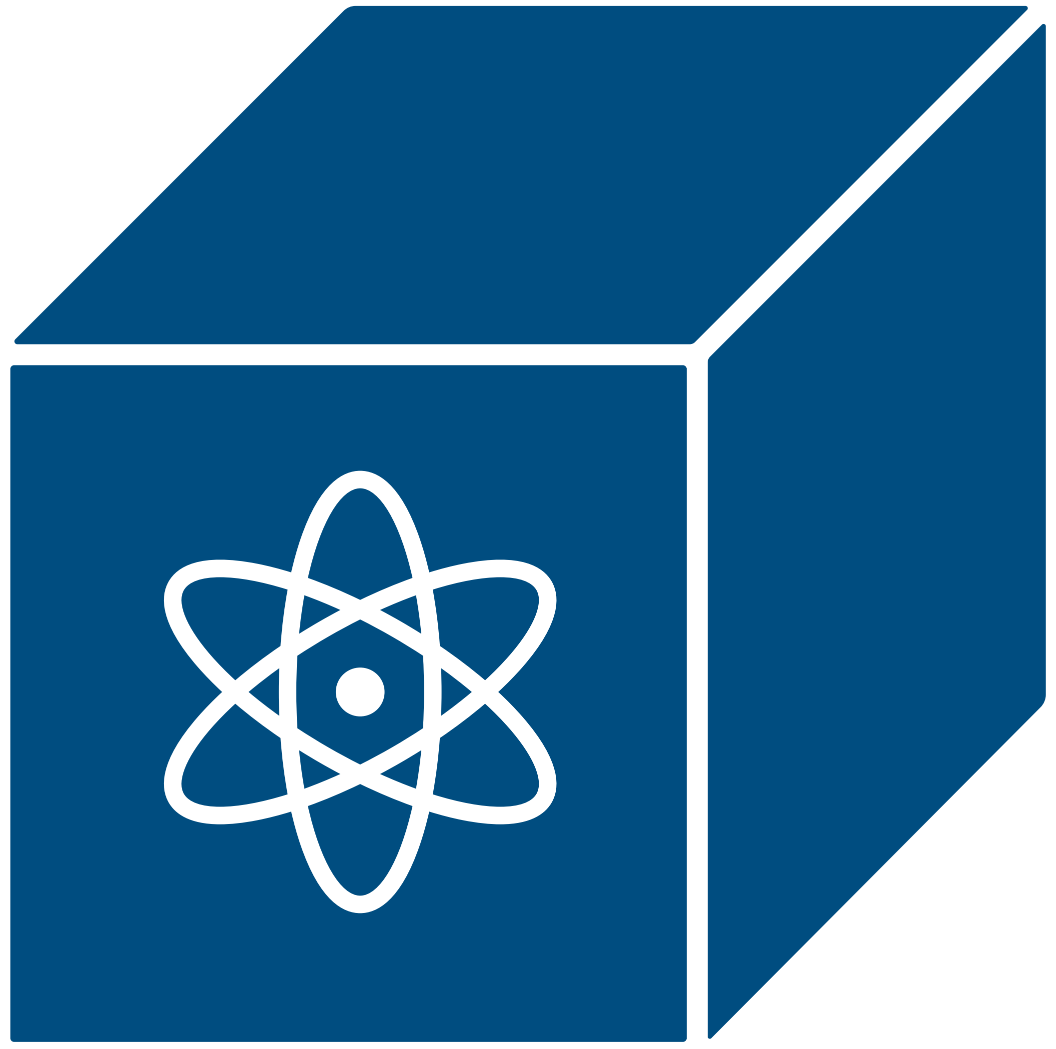
Goal: Provide a learning-theoretic foundation for various applications

Example 1: Quantum state tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\hat{\rho} \approx \rho$ under trace norm $\|\cdot\|_1$.

Motivations:

- The most basic quantum learning problem



Unknown quantum state

References:

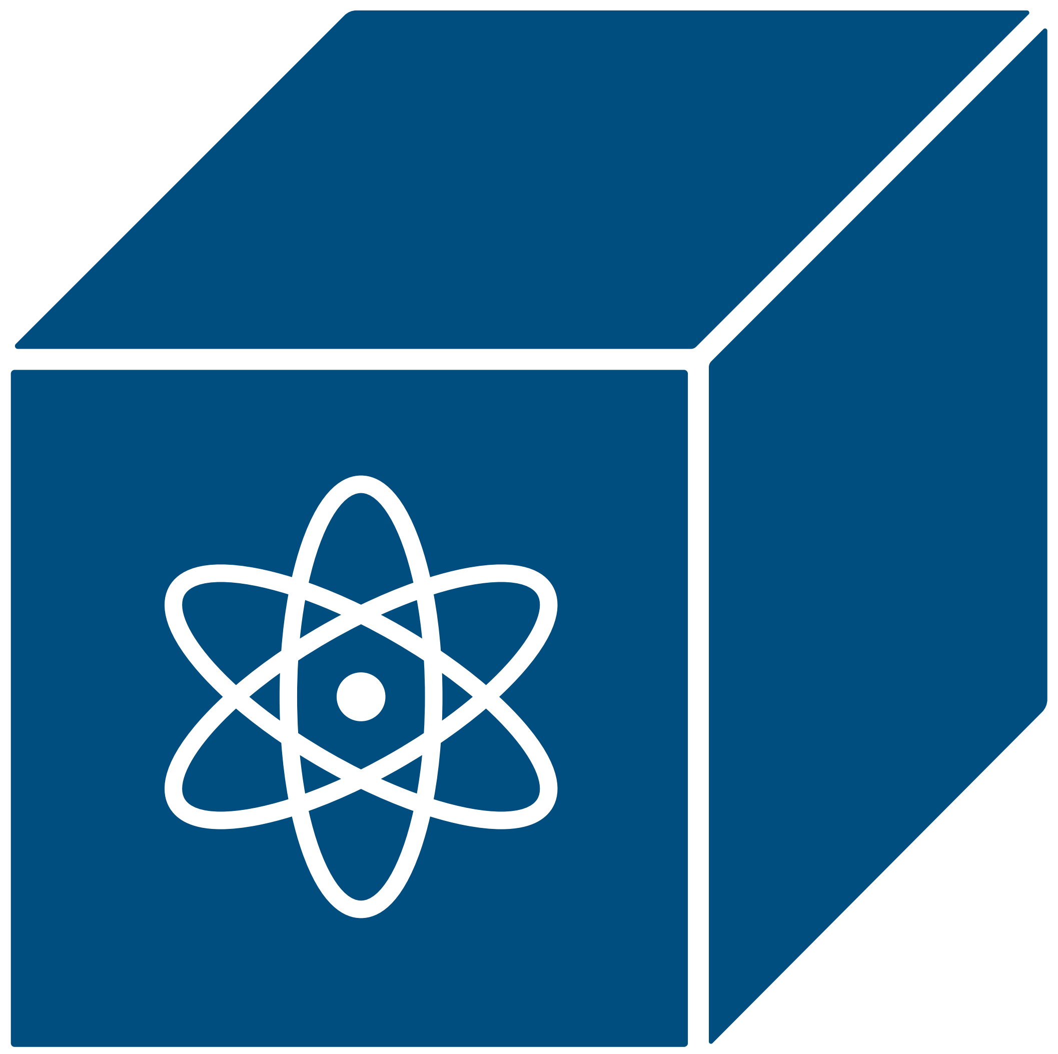
[1] Leonhardt, Ulf. "Quantum-state tomography and discrete Wigner function." *Physical review letters* 74.21 (1995): 4101.
[2] Gross, David, et al. "Quantum state tomography via compressed sensing." *Physical review letters* 105.15 (2010): 150401.
[3] O'Donnell, Ryan, and John Wright. "Efficient quantum tomography." *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016.
[4] Haah, Jeongwan, et al. "Sample-optimal tomography of quantum states." *IEEE Transactions on Information Theory* 63.9 (2017): 5628-5641.

Example 1: Quantum state tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\hat{\rho} \approx \rho$ under trace norm $\|\cdot\|_1$.

Motivations:

- The most basic quantum learning problem
- Benchmark quantum systems



Unknown quantum state

References:

[1] Leonhardt, Ulf. "Quantum-state tomography and discrete Wigner function." *Physical review letters* 74.21 (1995): 4101.

[2] Gross, David, et al. "Quantum state tomography via compressed sensing." *Physical review letters* 105.15 (2010): 150401.

[3] O'Donnell, Ryan, and John Wright. "Efficient quantum tomography." *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016.

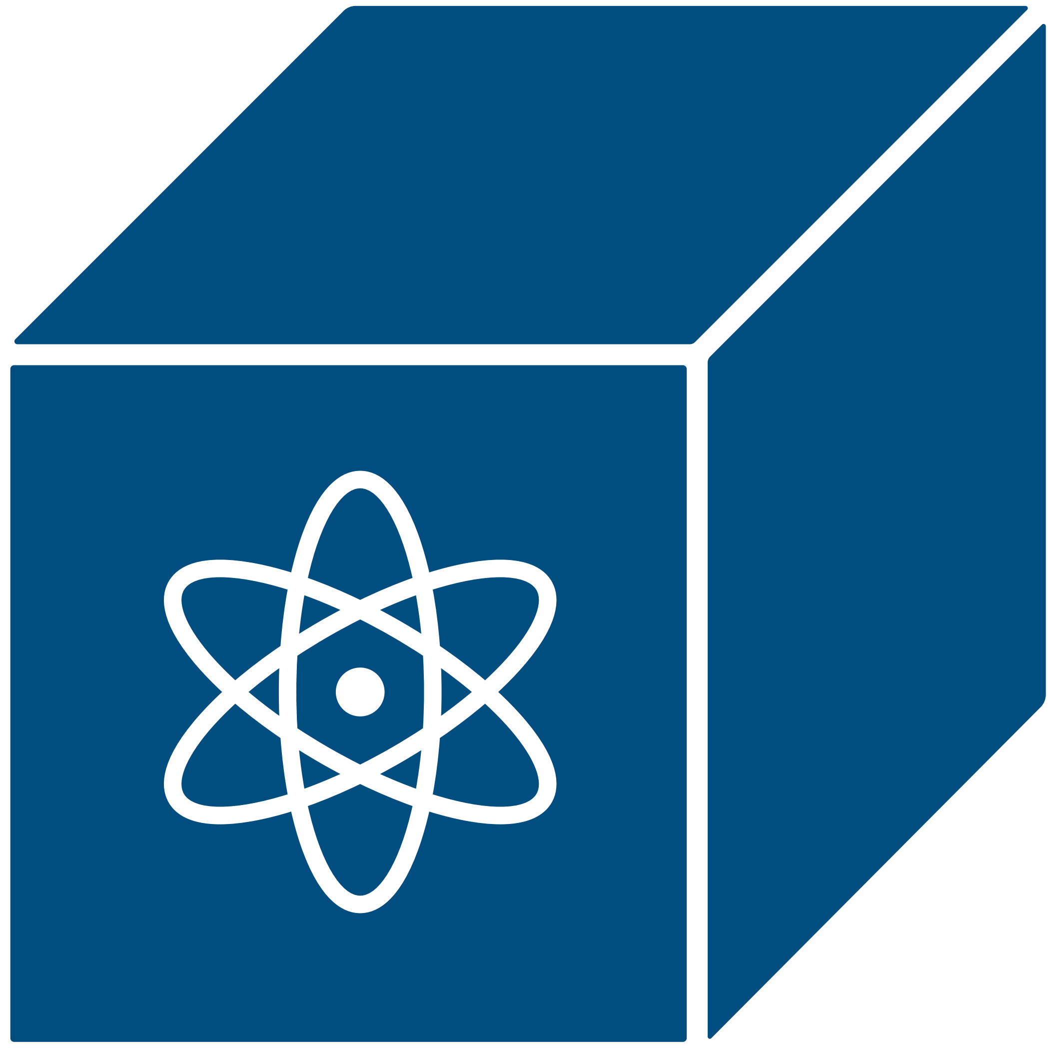
[4] Haah, Jeongwan, et al. "Sample-optimal tomography of quantum states." *IEEE Transactions on Information Theory* 63.9 (2017): 5628-5641.

Example 1: Quantum state tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\hat{\rho} \approx \rho$ under trace norm $\|\cdot\|_1$.

Motivations:

- The most basic quantum learning problem
- Benchmark quantum systems



Unknown quantum state

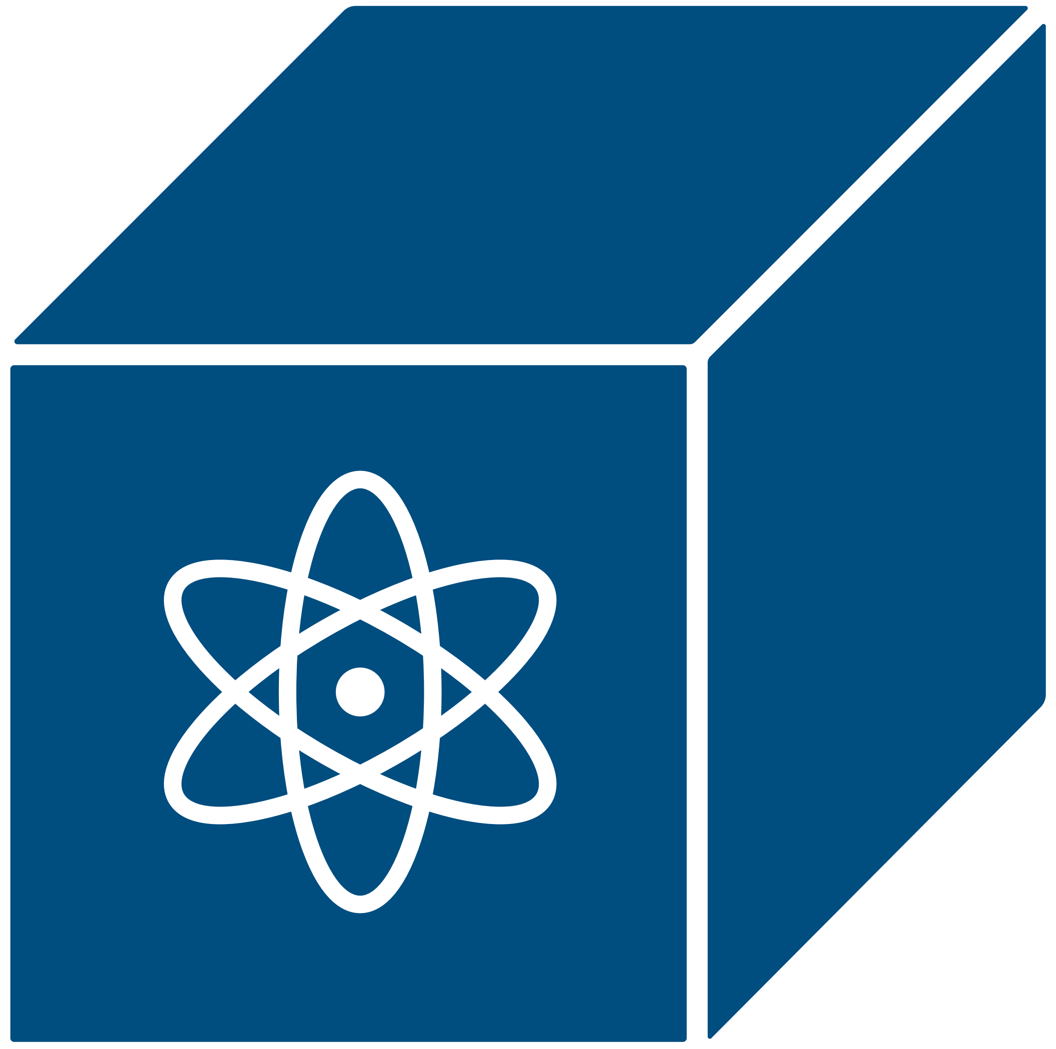
Complexity is exponential in n

References:

[1] Leonhardt, Ulf. "Quantum-state tomography and discrete Wigner function." *Physical review letters* 74.21 (1995): 4101.
[2] Gross, David, et al. "Quantum state tomography via compressed sensing." *Physical review letters* 105.15 (2010): 150401.
[3] O'Donnell, Ryan, and John Wright. "Efficient quantum tomography." *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016.
[4] Haah, Jeongwan, et al. "Sample-optimal tomography of quantum states." *IEEE Transactions on Information Theory* 63.9 (2017): 5628-5641.

Example 2: Shadow tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O_i \hat{\rho}) \approx \mathrm{Tr}(O_i \rho)$ for observables $O_1, \ldots, O_M$.



Unknown quantum state

References:

[1] Aaronson, Scott. "Shadow tomography of quantum states." *SIAM Journal on Computing* 49.5 (2019): STOC18-368.

[2] Bădescu, Costin, and Ryan O'Donnell. "Improved quantum data analysis." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

[3] Huang, Hsin-Yuan, Richard Kueng, and John Preskill. *"Predicting many properties of a quantum system from very few measurements." Nature Physics 16.10 (2020): 1050-1057.*

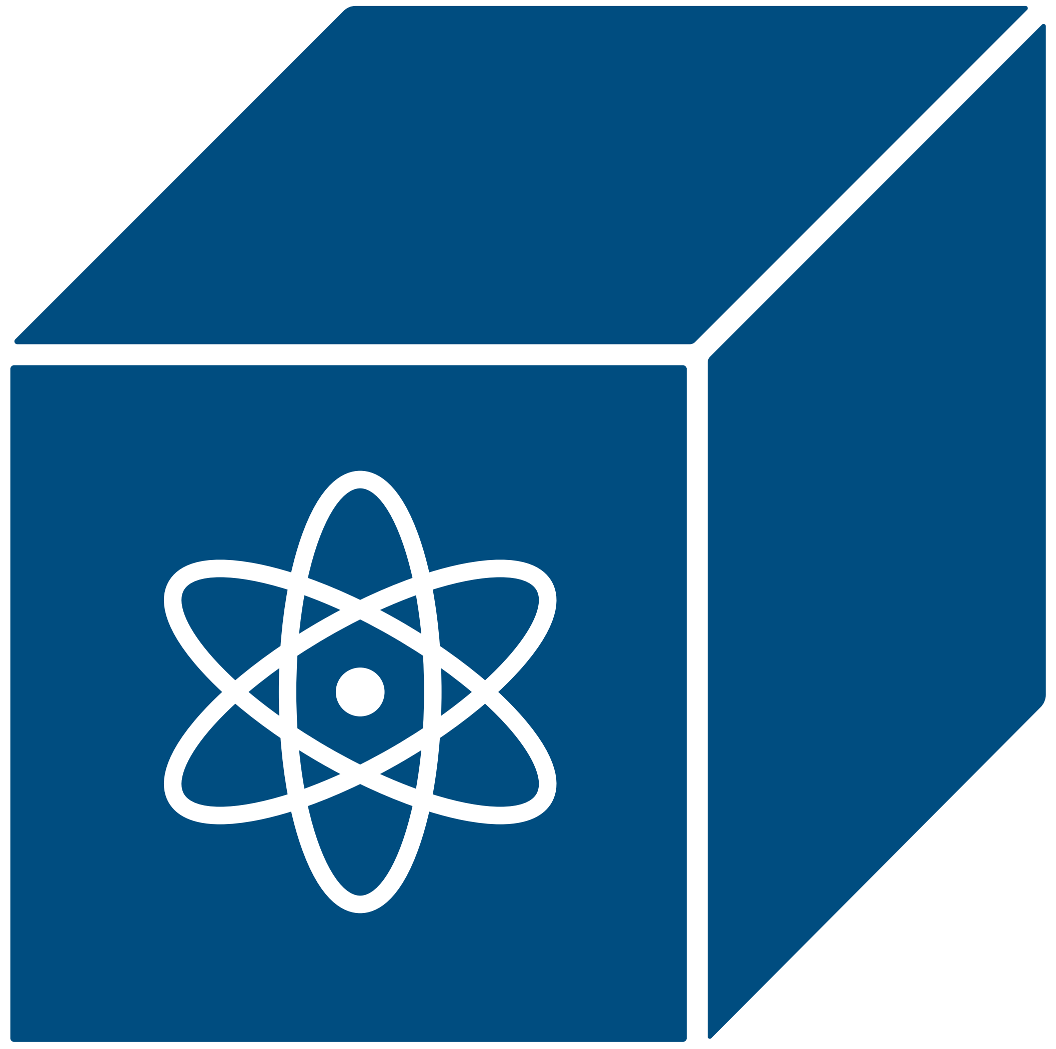
[4] Cotler, Jordan, and Frank Wilczek. "Quantum overlapping tomography." *Physical review letters* 124.10 (2020): 100401.

Example 2: Shadow tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O_i \hat{\rho}) \approx \mathrm{Tr}(O_i \rho)$ for observables $O_1, \ldots, O_M$.

Motivations:

- 2nd most basic quantum learning problem



Unknown quantum state

References:

[1] Aaronson, Scott. "Shadow tomography of quantum states." *SIAM Journal on Computing* 49.5 (2019): STOC18-368.

[2] Bădescu, Costin, and Ryan O'Donnell. "Improved quantum data analysis." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

[3] Huang, Hsin-Yuan, Richard Kueng, and John Preskill. *"Predicting many properties of a quantum system from very few measurements." Nature Physics 16.10 (2020): 1050-1057.*

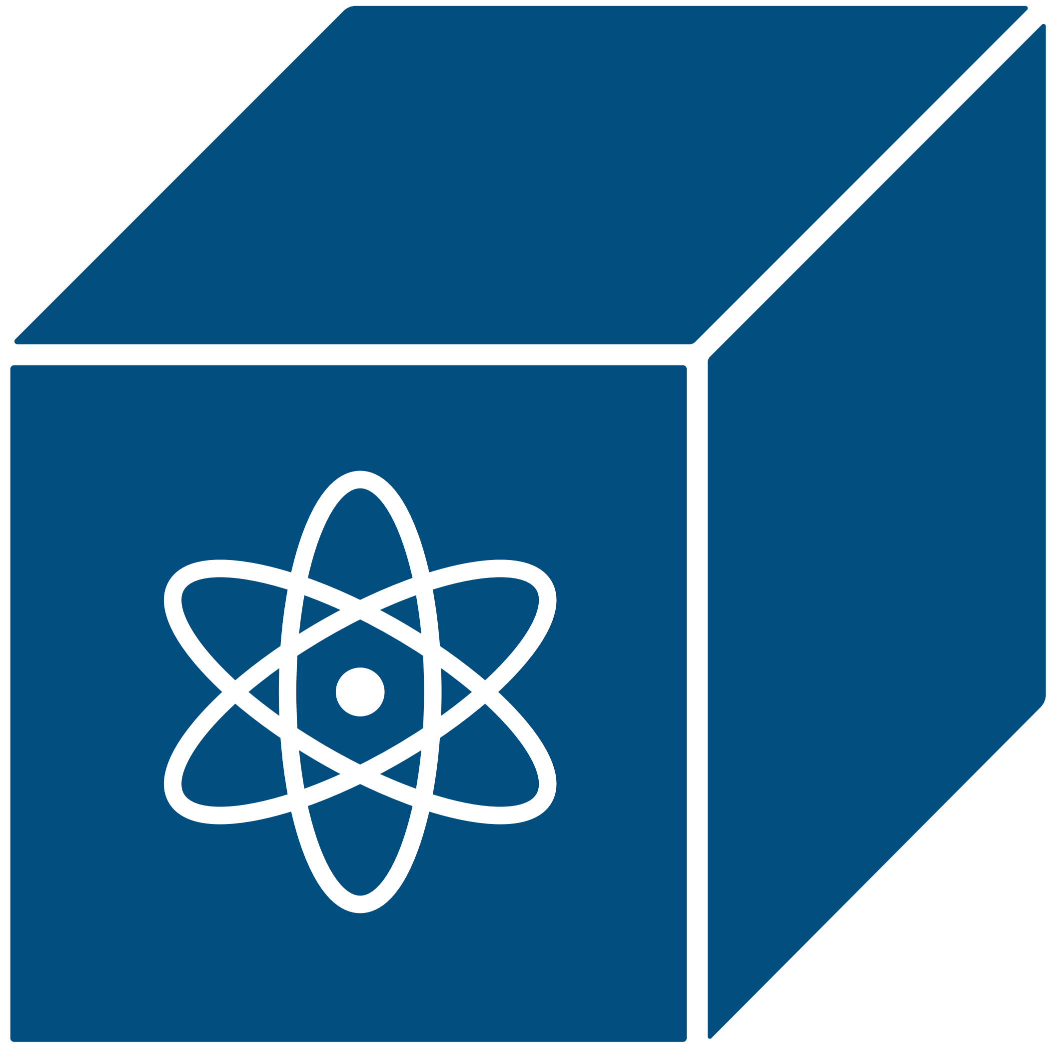
[4] Cotler, Jordan, and Frank Wilczek. "Quantum overlapping tomography." *Physical review letters* 124.10 (2020): 100401.

Example 2: Shadow tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O_i \hat{\rho}) \approx \mathrm{Tr}(O_i \rho)$ for observables $O_1, \ldots, O_M$.

Motivations:

- 2nd most basic quantum learning problem
- Benchmark quantum systems w/ good scaling in n



Unknown quantum state

References:

[1] Aaronson, Scott. "Shadow tomography of quantum states." *SIAM Journal on Computing* 49.5 (2019): STOC18-368.

[2] Bădescu, Costin, and Ryan O'Donnell. "Improved quantum data analysis." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

[3] Huang, Hsin-Yuan, Richard Kueng, and John Preskill. *"Predicting many properties of a quantum system from very few measurements." Nature Physics 16.10 (2020): 1050-1057.*

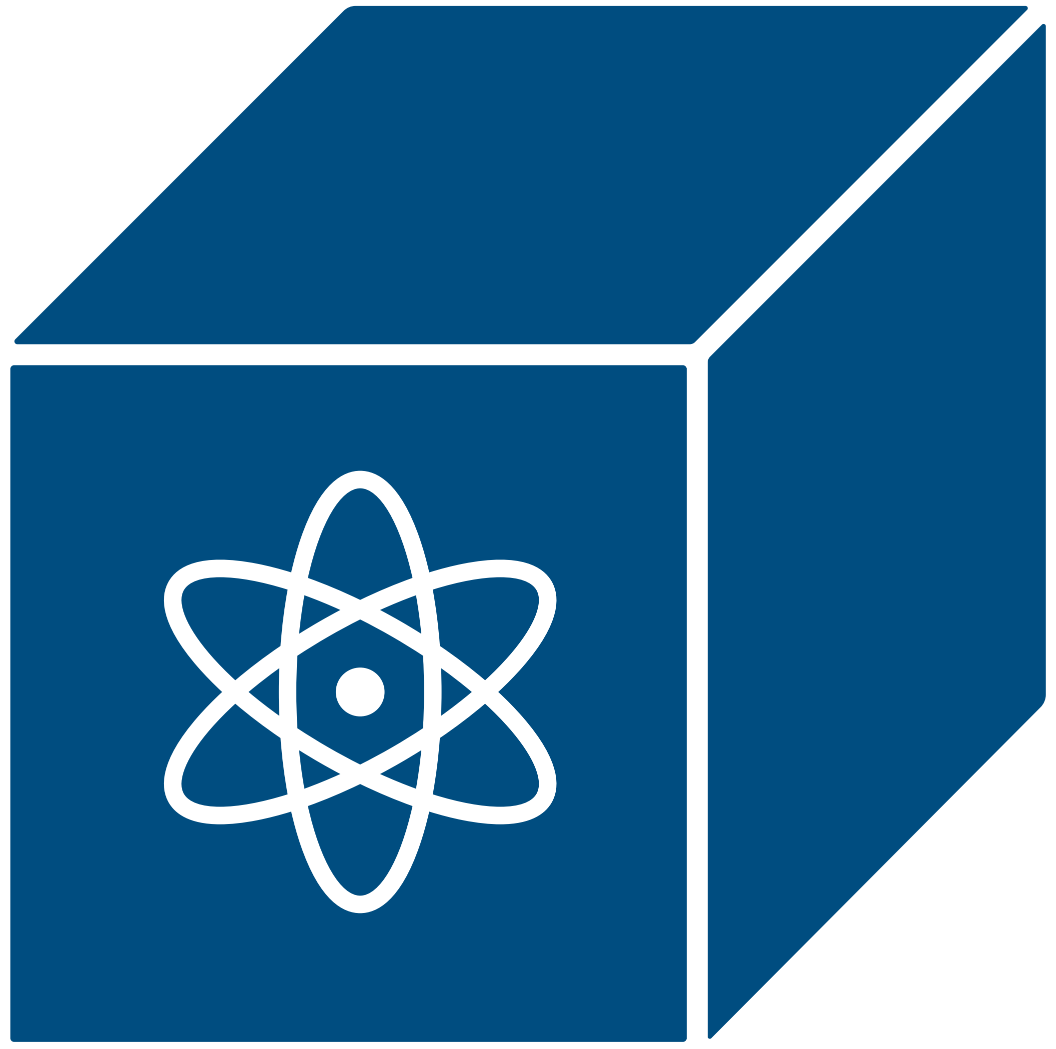
[4] Cotler, Jordan, and Frank Wilczek. "Quantum overlapping tomography." *Physical review letters* 124.10 (2020): 100401.

Example 2: Shadow tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O_i \hat{\rho}) \approx \mathrm{Tr}(O_i \rho)$ for observables $O_1, \ldots, O_M$.

Motivations:

- 2nd most basic quantum learning problem
- Benchmark quantum systems w/ good scaling in n
- A basic primitive in hybrid quantum/classical algorithms



Unknown quantum state

References:

[1] Aaronson, Scott. "Shadow tomography of quantum states." *SIAM Journal on Computing* 49.5 (2019): STOC18-368.

[2] Bădescu, Costin, and Ryan O'Donnell. "Improved quantum data analysis." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

[3] Huang, Hsin-Yuan, Richard Kueng, and John Preskill. *"Predicting many properties of a quantum system from very few measurements." Nature Physics 16.10 (2020): 1050-1057.*

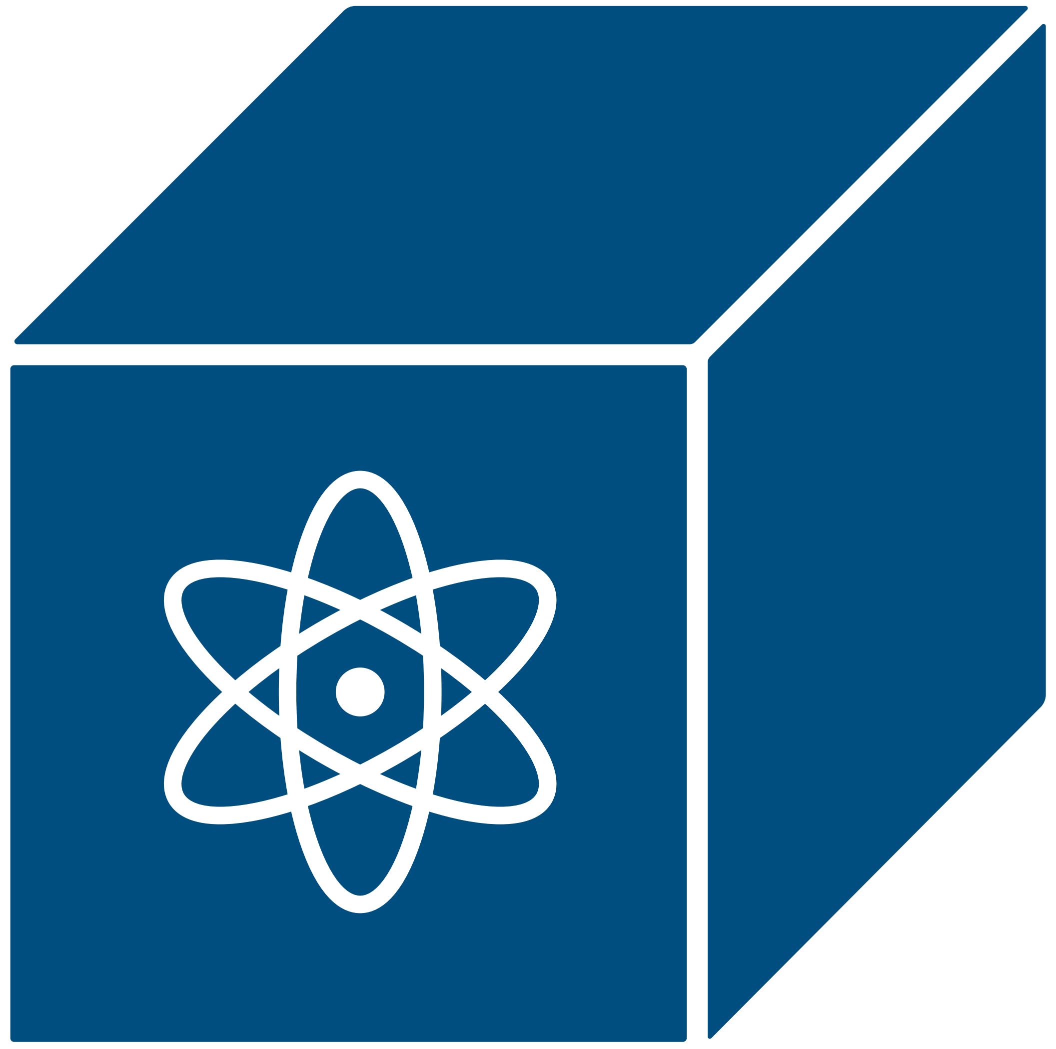
[4] Cotler, Jordan, and Frank Wilczek. "Quantum overlapping tomography." *Physical review letters* 124.10 (2020): 100401.

Example 2: Shadow tomography

- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn a classical description $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O_i \hat{\rho}) \approx \mathrm{Tr}(O_i \rho)$ for observables $O_1, \ldots, O_M$.

Motivations:

- 2nd most basic quantum learning problem
- Benchmark quantum systems w/ good scaling in n
- A basic primitive in hybrid quantum/classical algorithms



Unknown quantum state

Complexity is linear or independent in n

References:

[1] Aaronson, Scott. "Shadow tomography of quantum states." *SIAM Journal on Computing* 49.5 (2019): STOC18-368.

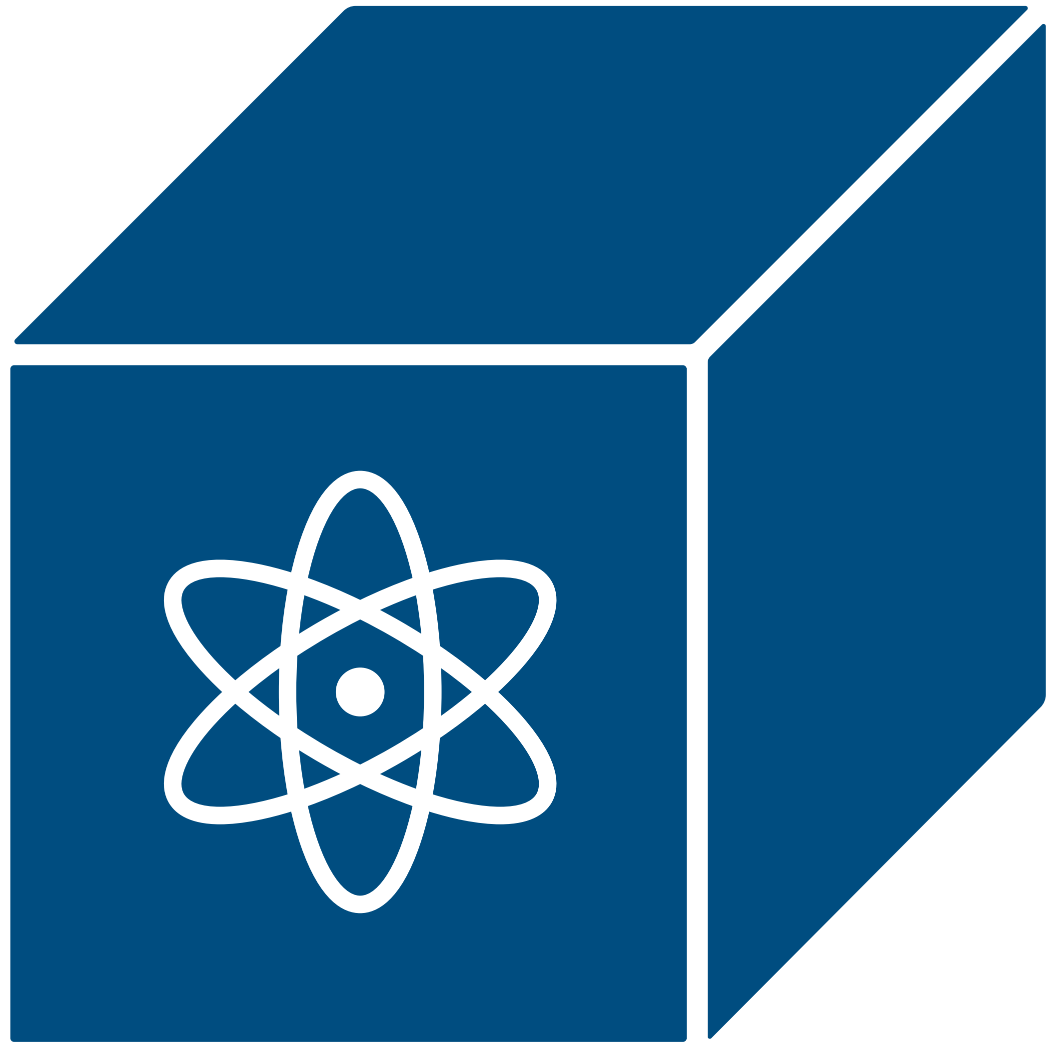
[2] Bădescu, Costin, and Ryan O'Donnell. "Improved quantum data analysis." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

[3] Huang, Hsin-Yuan, Richard Kueng, and John Preskill. *"Predicting many properties of a quantum system from very few measurements." Nature Physics 16.10 (2020): 1050-1057.*

[4] Cotler, Jordan, and Frank Wilczek. "Quantum overlapping tomography." *Physical review letters* 124.10 (2020): 100401.

Example 3: Pauli channel tomography

- There is an unknown n-qubit Pauli channel $\mathscr{P}$.
- Learn $\hat{\mathscr{P}}$ by preparing input states, evolving under $\mathscr{P}$, and measuring output states.
- After learning, we want $\hat{\mathscr{P}} \approx \mathscr{P}$ under diamond norm.



Unknown Pauli channel

References:

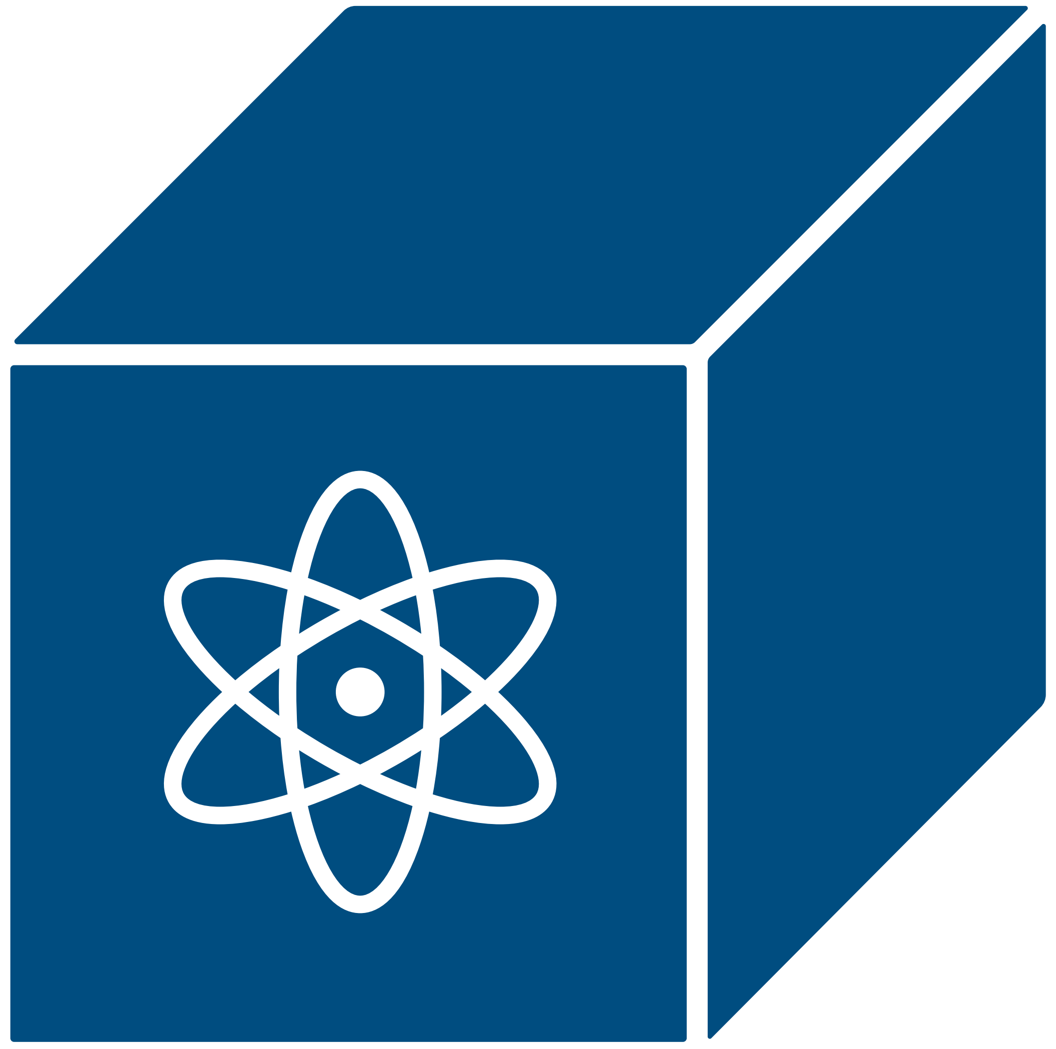
[1] Flammia, Steven T., and Joel J. Wallman. "Efficient estimation of Pauli channels." *ACM Transactions on Quantum Computing* 1.1 (2020): 1-32.
[2] *Harper, Robin, Steven T. Flammia, and Joel J. Wallman. "Efficient learning of quantum noise." Nature Physics 16.12 (2020): 1184-1188.*
[3] Flammia, Steven T., and Ryan O'Donnell. "Pauli error estimation via Population Recovery." *Quantum* 5 (2021): 549.
[4] Chen, Senrui, et al. "Quantum advantages for pauli channel estimation." *Physical Review A* 105.3 (2022): 032435.

Example 3: Pauli channel tomography

- There is an unknown n-qubit Pauli channel $\mathscr{P}$.
- Learn $\hat{\mathscr{P}}$ by preparing input states, evolving under $\mathscr{P}$, and measuring output states.
- After learning, we want $\hat{\mathscr{P}} \approx \mathscr{P}$ under diamond norm.

Motivations:

- Characterize quantum noise
- Useful for quantum error correction, error mitigation



Unknown Pauli channel

References:

[1] Flammia, Steven T., and Joel J. Wallman. "Efficient estimation of Pauli channels." *ACM Transactions on Quantum Computing* 1.1 (2020): 1-32.

[2] *Harper, Robin, Steven T. Flammia, and Joel J. Wallman. "Efficient learning of quantum noise." Nature Physics 16.12 (2020): 1184-1188.*

[3] Flammia, Steven T., and Ryan O'Donnell. "Pauli error estimation via Population Recovery." *Quantum* 5 (2021): 549.

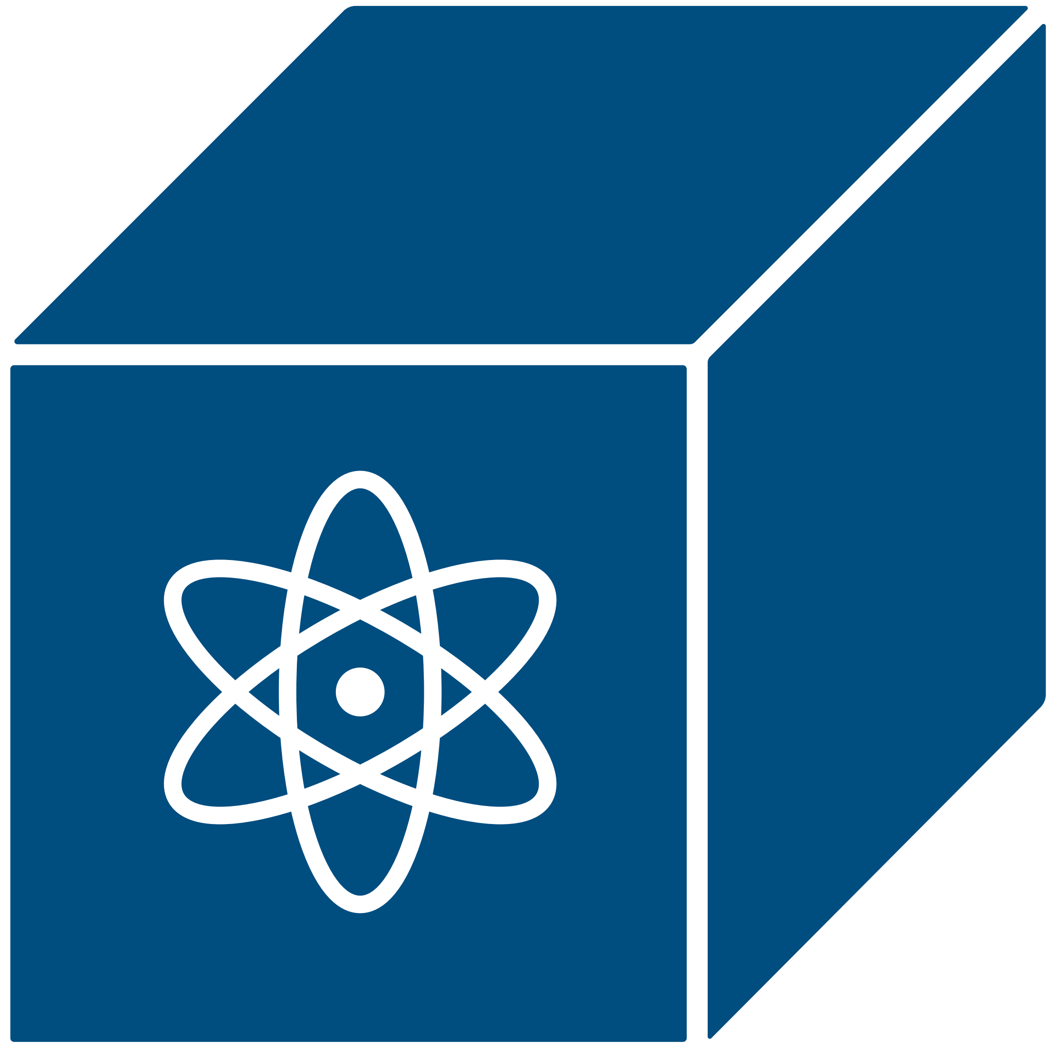
[4] Chen, Senrui, et al. "Quantum advantages for pauli channel estimation." *Physical Review A* 105.3 (2022): 032435.

Example 3: Pauli channel tomography

- There is an unknown n-qubit Pauli channel $\mathscr{P}$.
- Learn $\hat{\mathscr{P}}$ by preparing input states, evolving under $\mathscr{P}$, and measuring output states.
- After learning, we want $\hat{\mathscr{P}} \approx \mathscr{P}$ under diamond norm.

Motivations:

- Characterize quantum noise
- Useful for quantum error correction, error mitigation



Complexity varies under additional assumptions

References:

[1] Flammia, Steven T., and Joel J. Wallman. "Efficient estimation of Pauli channels." *ACM Transactions on Quantum Computing* 1.1 (2020): 1-32.

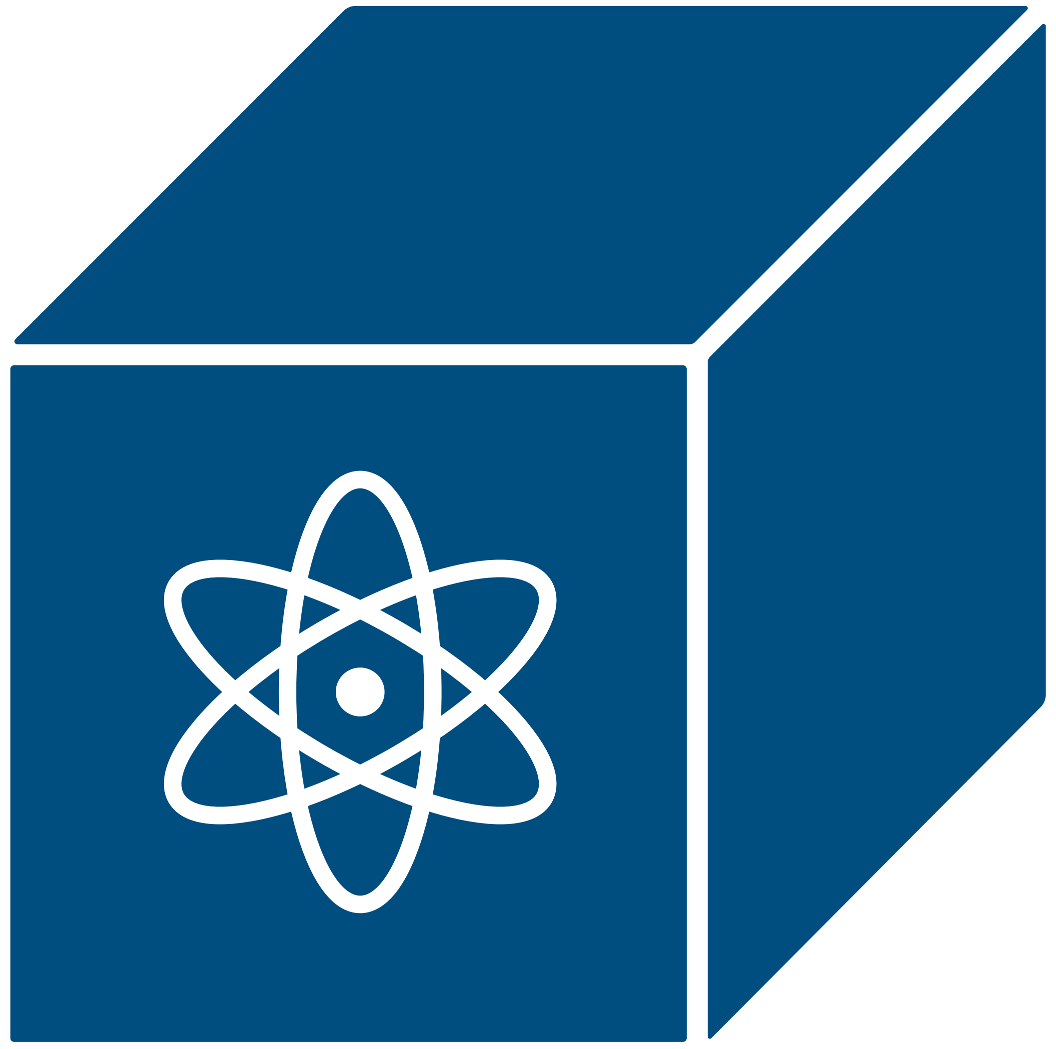
[2] *Harper, Robin, Steven T. Flammia, and Joel J. Wallman. "Efficient learning of quantum noise." Nature Physics 16.12 (2020): 1184-1188.*

[3] Flammia, Steven T., and Ryan O'Donnell. "Pauli error estimation via Population Recovery." *Quantum* 5 (2021): 549.

[4] Chen, Senrui, et al. "Quantum advantages for pauli channel estimation." *Physical Review A* 105.3 (2022): 032435.

Example 4: Predicting ground states

- There is an unknown $f(x)$ mapping parameter x to the ground state of $H(x)$.
- Learn $\hat{f}$ by preparing ground states under different x's, and measuring the states.
- After learning, we want $\hat{f}(x) \approx f(x)$ for most of x.



Unknown phase diagram

References:

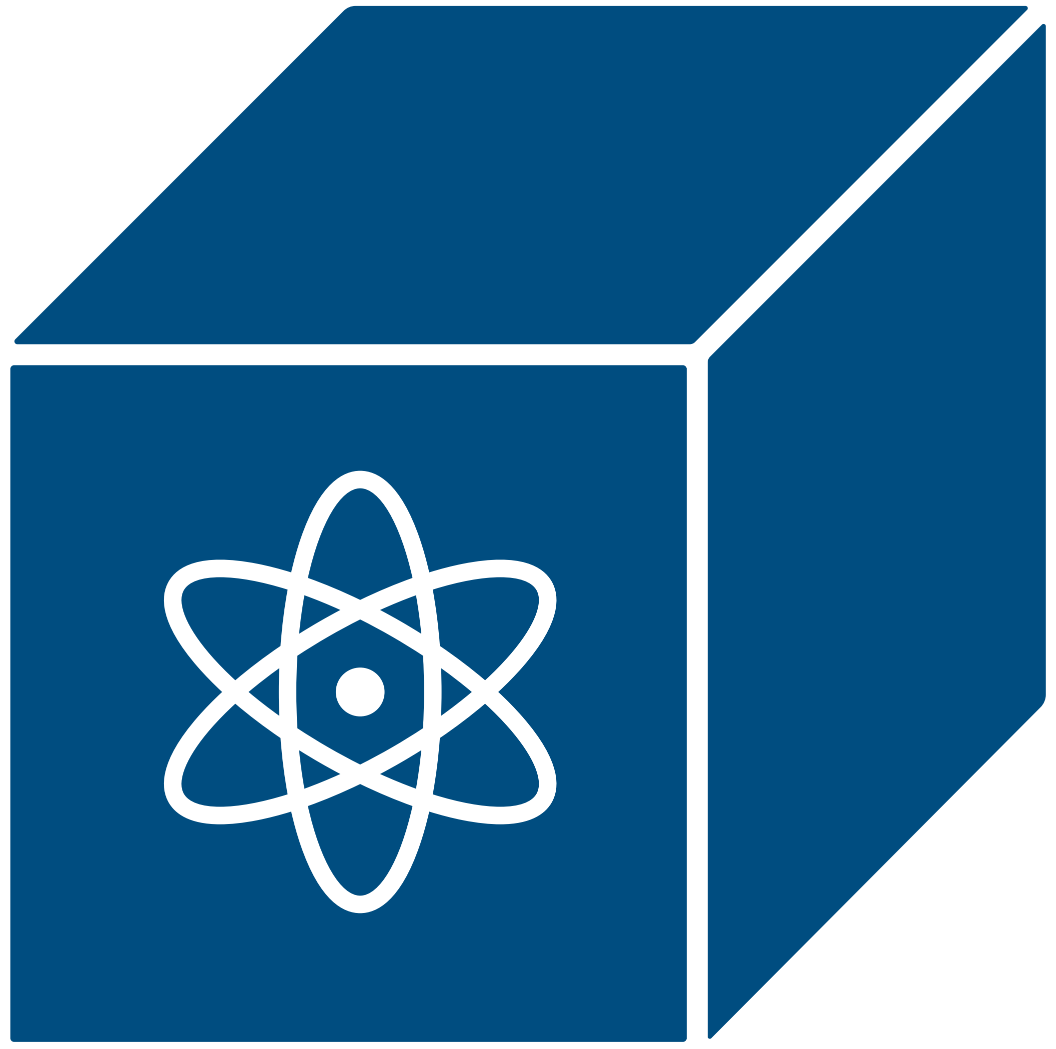
[1] Carleo, Giuseppe, and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks." *Science* 355.6325 (2017): 602-606.
[2] Gilmer, Justin, et al. "Neural message passing for quantum chemistry." *International conference on machine learning*. PMLR, 2017.
[3] Qiao, Zhuoran, et al. "OrbNet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features." *The Journal of chemical physics* 153.12 (2020): 124111.
[4] Huang, Hsin-Yuan, et al. "Provably efficient machine learning for quantum many-body problems." *arXiv preprint arXiv:2106.12627 (2021).*

Example 4: Predicting ground states

- There is an unknown $f(x)$ mapping parameter x to the ground state of $H(x)$.
- Learn $\hat{f}$ by preparing ground states under different x's, and measuring the states.
- After learning, we want $\hat{f}(x) \approx f(x)$ for most of x.

Motivations:

- Machine learning for quantum chemistry/physics
- Speed up computation with ML



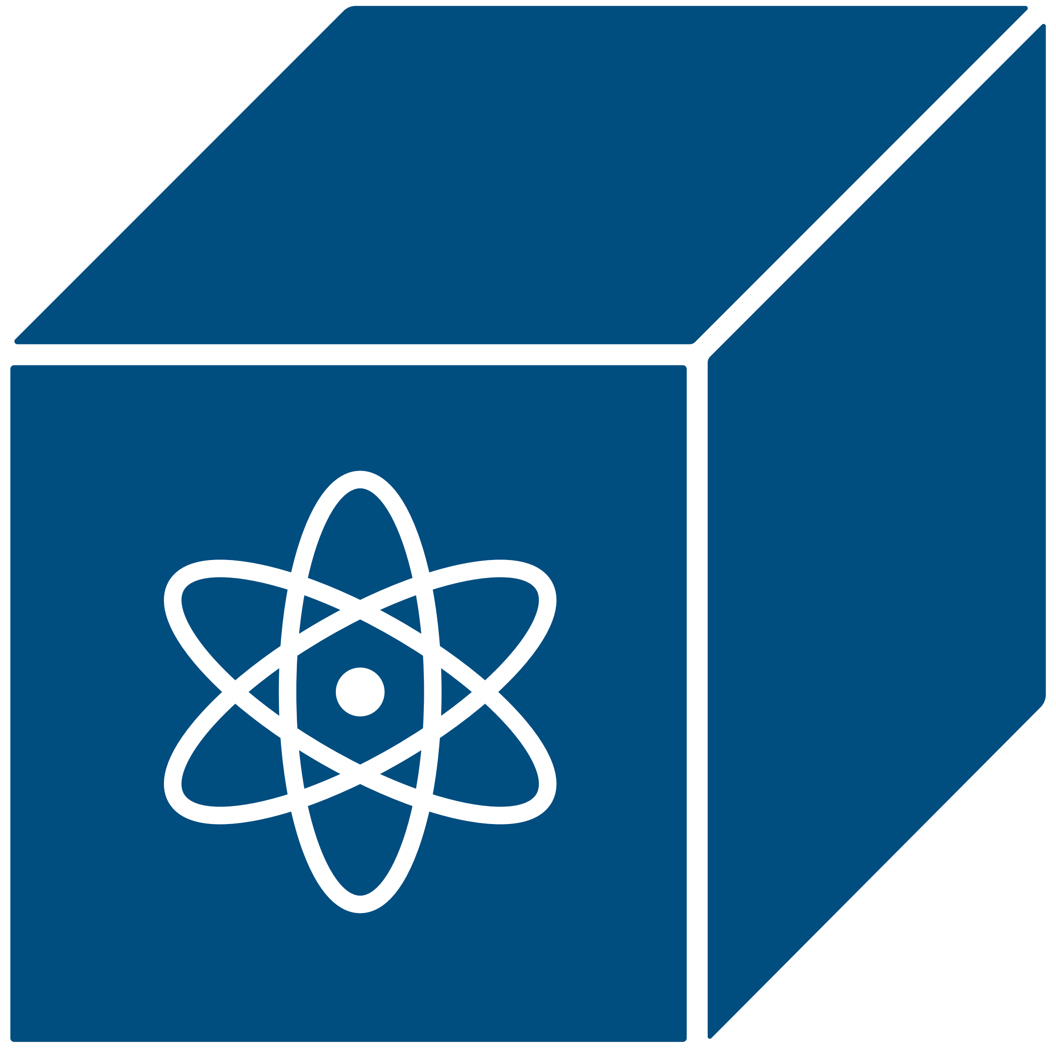
Unknown phase diagram

References:

[1] Carleo, Giuseppe, and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks." *Science* 355.6325 (2017): 602-606.
[2] Gilmer, Justin, et al. "Neural message passing for quantum chemistry." *International conference on machine learning*. PMLR, 2017.
[3] Qiao, Zhuoran, et al. "OrbNet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features." *The Journal of chemical physics* 153.12 (2020): 124111.
[4] Huang, Hsin-Yuan, et al. "Provably efficient machine learning for quantum many-body problems." *arXiv preprint arXiv:2106.12627 (2021).*

Type your problem here

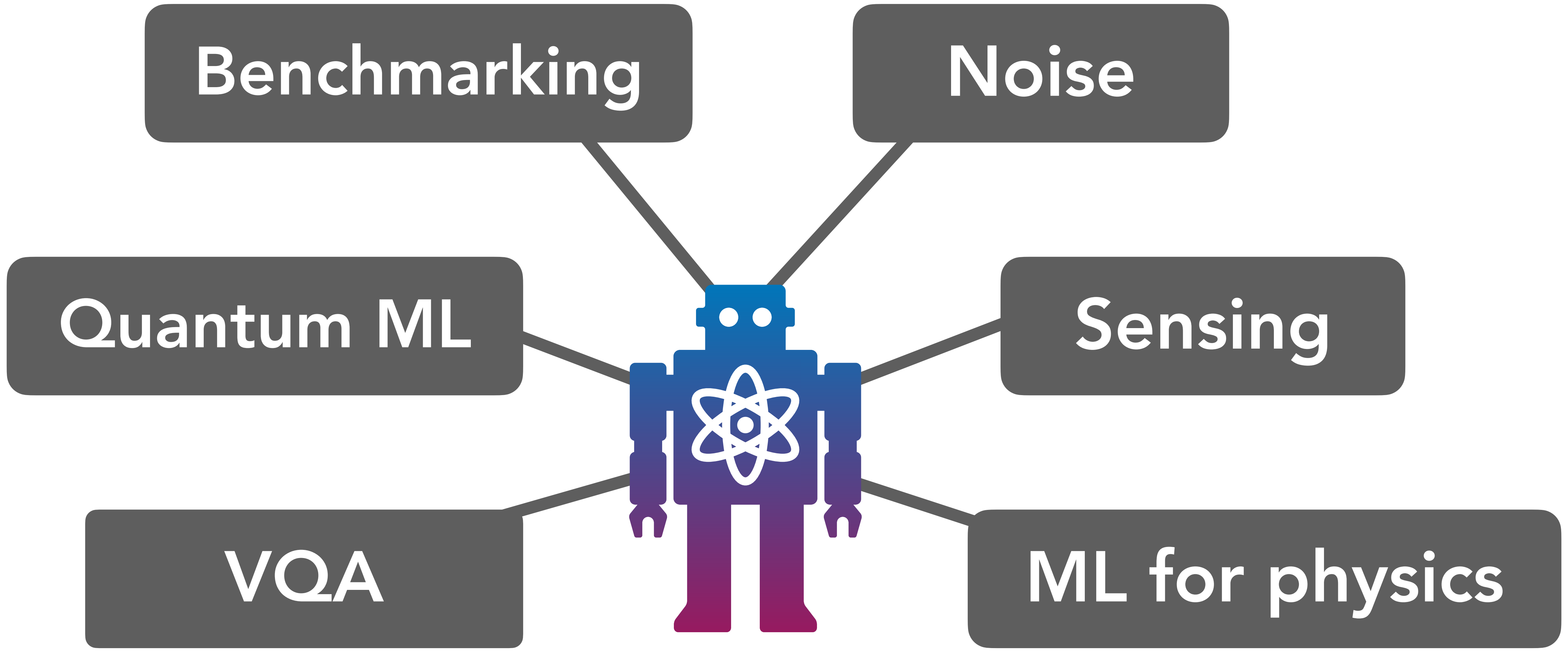
- Almost all problems contain some aspects about learning an unknown object.



Unknown [ ]

More problems

- What quantum circuits/algorithms can we learn? (QML & VQA)
- What aspects of an unknown quantum machine is learnable? (Benchmarking & Noise)
- How to learn a good quantum sensor given an unknown quantum machine? (Sensing)
- Can a learning algorithm discover "new physics"? (ML for physics)
- The list goes on …



Outline

- Basic setting and examples

- Key ideas:

  Part I — Designing good learning algorithms

  Part II — Proving no good learning algorithms exist

- Outlook and open questions

Outline

- Basic setting and examples

- Key ideas:

  Part I — Designing good learning algorithms

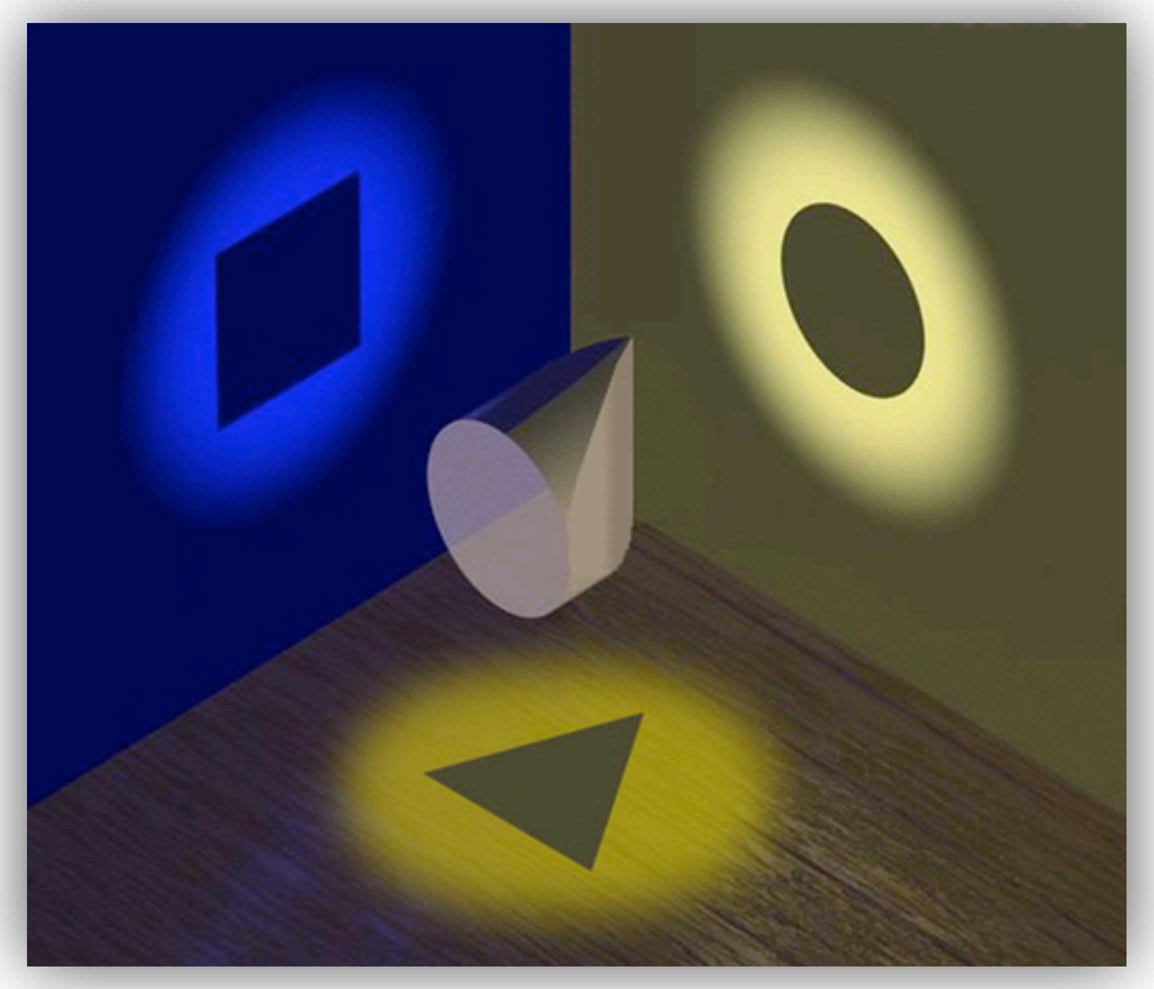
  Part II — Proving no good learning algorithms exist

- Outlook and open questions

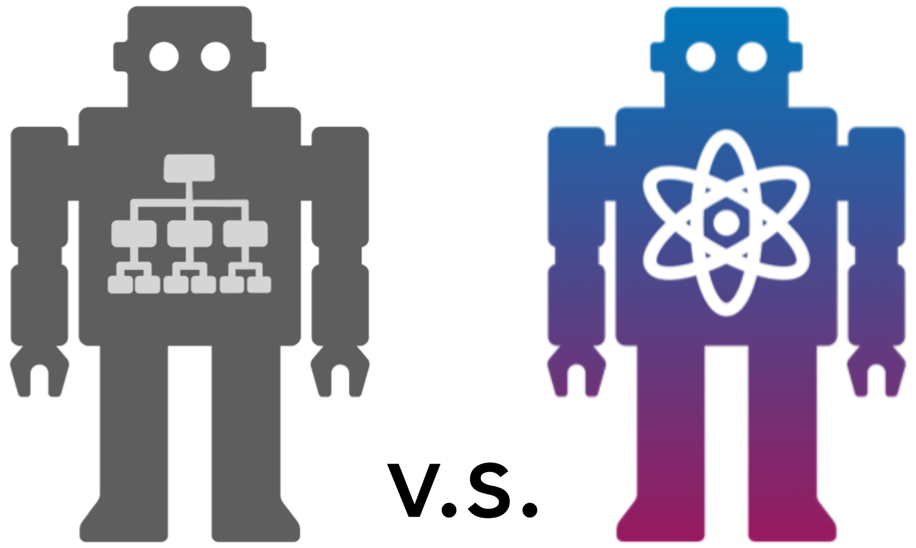
Key Ideas

- Part I focuses on upper bounds (how to design good learning algorithms).
- Part II focuses on lower bounds (how to show that no good algorithms exist).

Part I



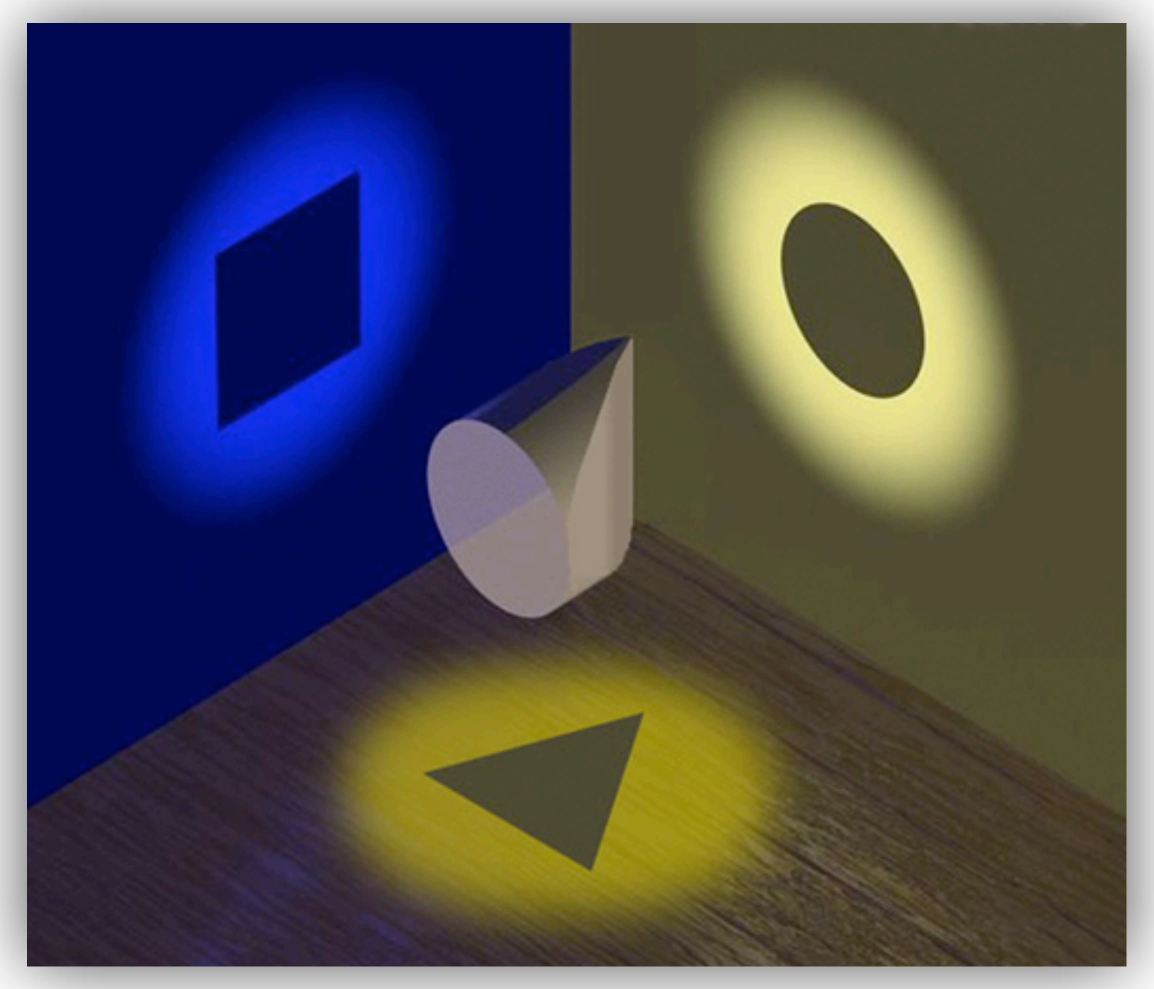
Part II



Key Ideas

Algorithmic side: randomized experiments + data processing
Analysis side: geometric analysis + concentration inequality

Part I



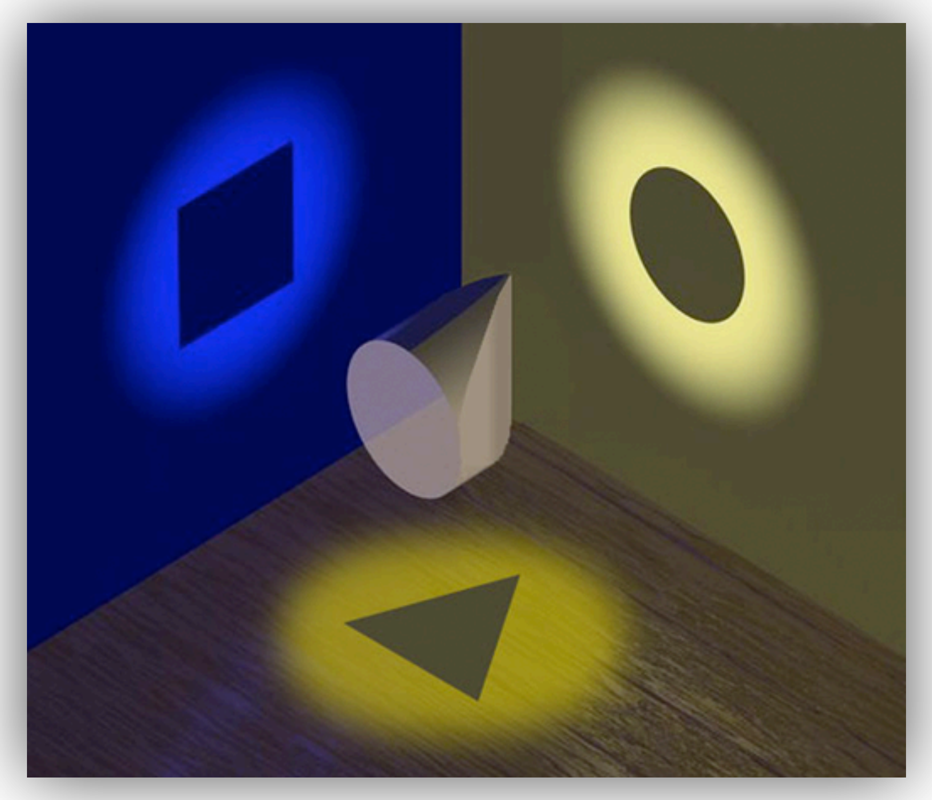
Part II

A dive into shadow tomography

Algorithmic side: randomized experiments + data processing
Analysis side: geometric analysis + concentration inequality

Recall the task of shadow tomography:

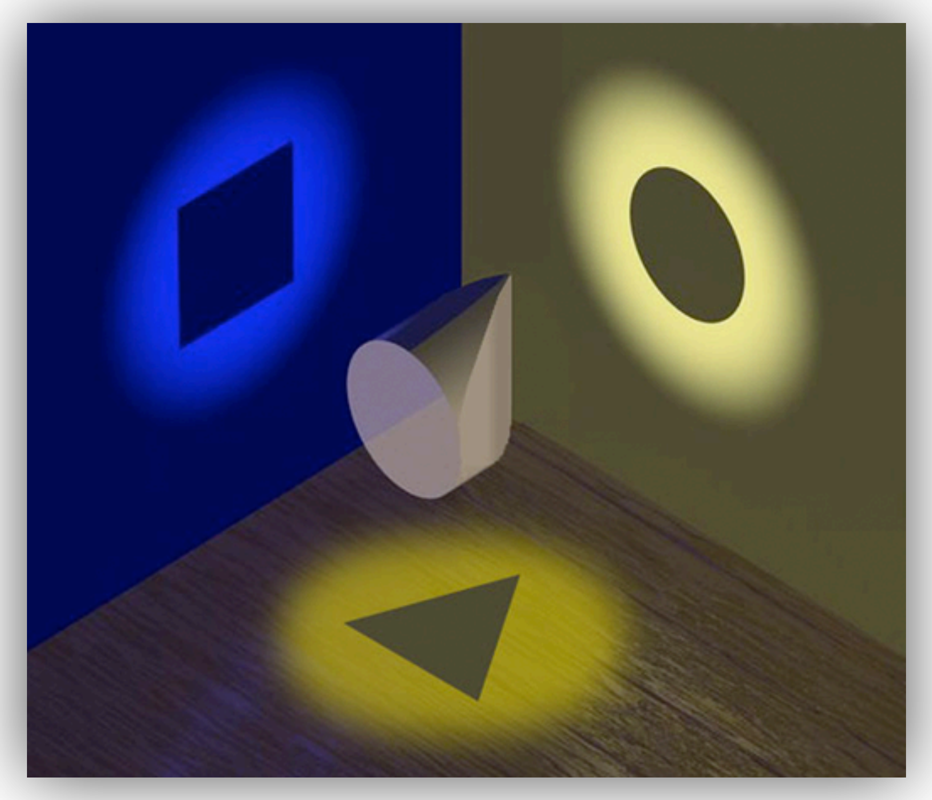
- There is an unknown n-qubit quantum state described by $\rho \in \mathbb{C}^{2^n \times 2^n}$.
- Learn $\hat{\rho}$ by performing measurements on copies of ρ.
- After learning, we want $\mathrm{Tr}(O\hat{\rho}) \approx \mathrm{Tr}(O\rho)$ for observables $O_1, \ldots, O_M$.



A dive into shadow tomography

Randomized experiment:

- Sample a random Clifford U_i to rotate the quantum state ρ.
- Measure the state in the computational basis $|b_i\rangle \in \{0,1\}^n$.



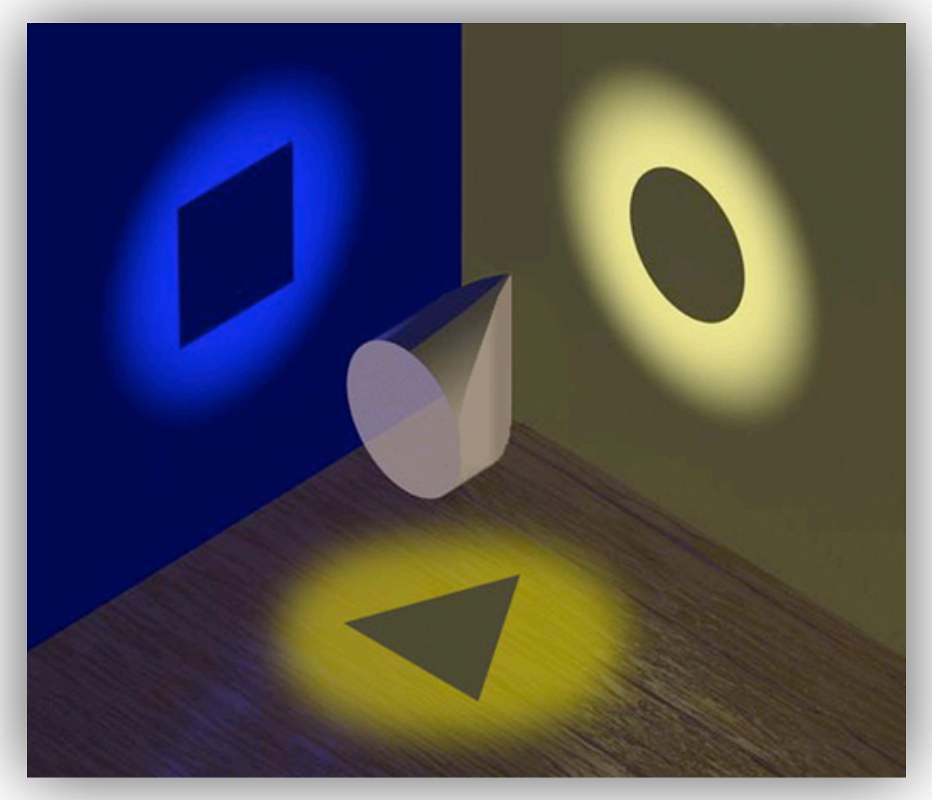
A dive into shadow tomography

Randomized experiment:

- Sample a random Clifford U_i to rotate the quantum state ρ.
- Measure the state in the computational basis $|b_i\rangle \in \{0,1\}^n$.

Data processing:

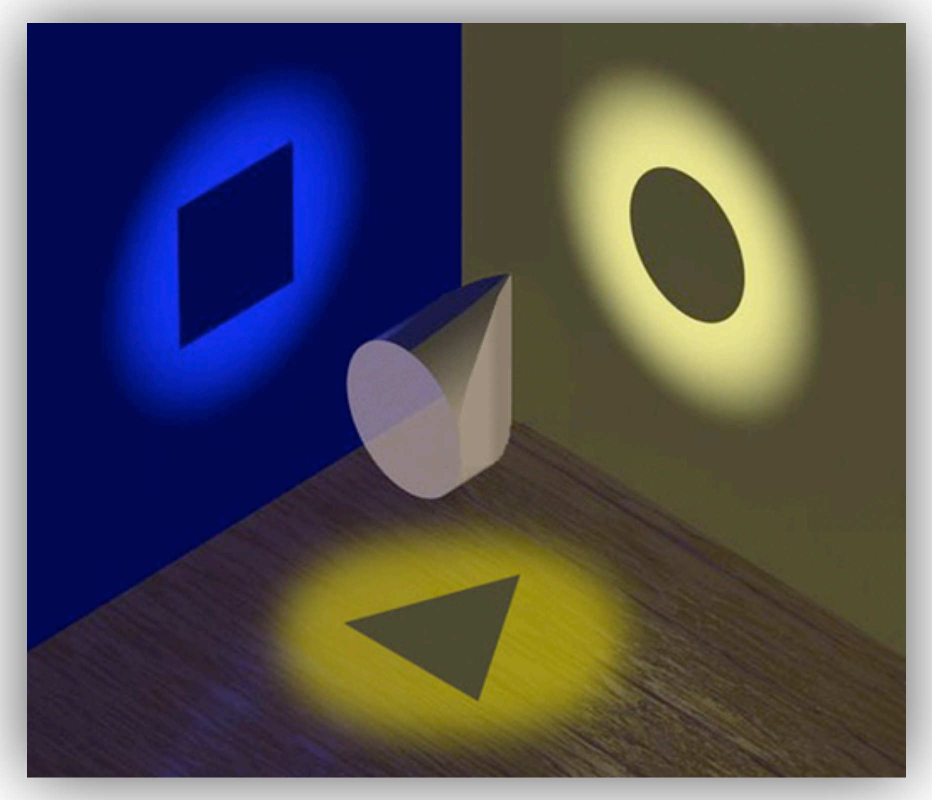
- Construct $\hat{\rho}_i = \left[(2^n + 1)U_i^\dagger |b_i\rangle\langle b_i| U_i - I\right]$ for each experiment.



A dive into shadow tomography

Randomized experiment:

- Sample a random Clifford U_i to rotate the quantum state ρ.
- Measure the state in the computational basis $|b_i\rangle \in \{0,1\}^n$.



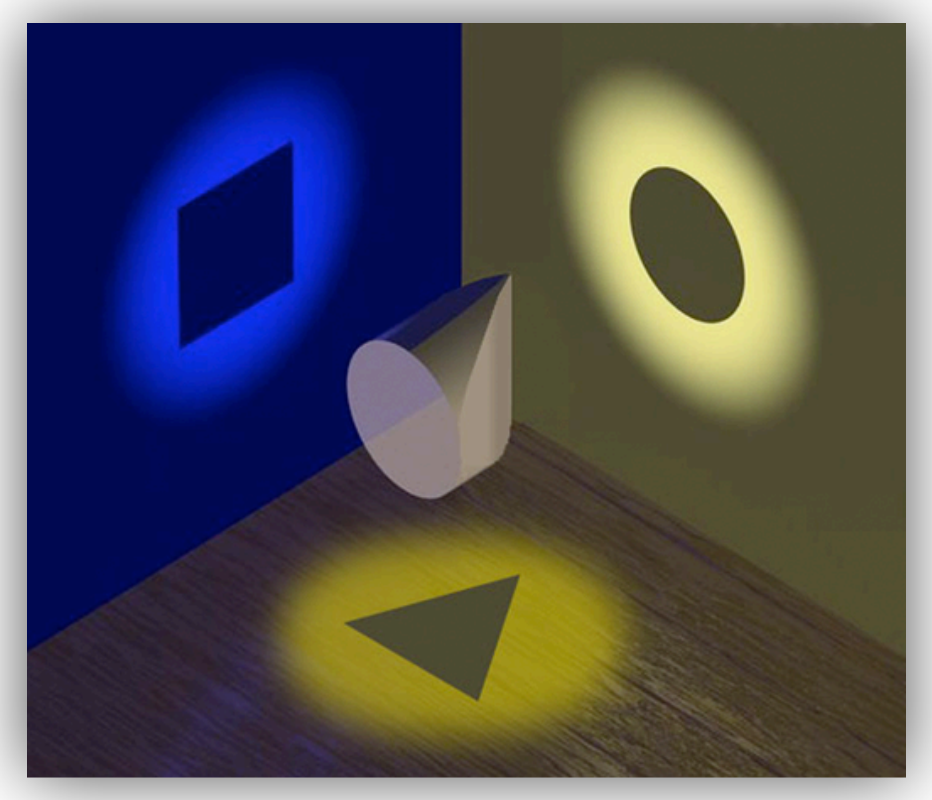
Data processing:

- Construct $\hat{\rho}_i = \left[(2^n+1)U_i^\dagger |b_i\rangle\langle b_i| U_i - I\right]$ for each experiment.

Geometric analysis:

- The geometry of Clifford unitary says $\mathbb{E}[\mathrm{Tr}(O\hat{\rho})_i] = \mathrm{Tr}(O\rho)$ and $\mathrm{Var}[\mathrm{Tr}(O\hat{\rho}_i)] \leq 3\mathrm{Tr}(O^2)$.

A dive into shadow tomography



Randomized experiment:

- Sample a random Clifford U_i to rotate the quantum state ρ.
- Measure the state in the computational basis $|b_i\rangle \in \{0,1\}^n$.

Data processing:

- Construct $\hat{\rho}_i = \left[(2^n + 1)U_i^\dagger |b_i\rangle\langle b_i| U_i - I\right]$ for each experiment.

Geometric analysis:

- The geometry of Clifford unitary says $\mathbb{E}[\mathrm{Tr}(O\hat{\rho})_i] = \mathrm{Tr}(O\rho)$ and $\mathrm{Var}[\mathrm{Tr}(O\hat{\rho}_i)] \leq 3\mathrm{Tr}(O^2)$.

Concentration bound:

- For $O_1, \ldots, O_M$ with $\mathrm{Tr}(O^2) = \mathcal{O}(1)$, we can predict $\mathrm{Tr}(O_i\rho)$ after $\mathcal{O}(\log(M))$ measurements.

A dive into shadow tomography

Theorem (Huang et al.; 2020)

1. Given $B, \epsilon > 0$, the procedure learns a classical representation of an unknown quantum state ρ from
$$N = \mathcal{O}(B \log(M)/\epsilon^2) \text{ measurements.}$$
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \|O_i\|_2^2$, the procedure can use the classical representation to predict $\hat{o}_1, \ldots, \hat{o}_M$, where $|\hat{o}_i - \mathrm{tr}(O_i \rho)| < \epsilon$, for all i.

For example:

- $M = 10^6$, $B = 1$, then naively we need $10^6/\epsilon^2$ measurements.
- This theorem shows that we only need $6 \log(10)/\epsilon^2$ measurements.

Furthermore, we don't need to know $O_1, \ldots, O_M$ in advance.

A dive into shadow tomography

Theorem (Huang et al.; 2020)

1. Given $B, \epsilon > 0$, the procedure learns a classical representation of an unknown quantum state ρ from
$$N = \mathcal{O}(B \log(M)/\epsilon^2) \text{ measurements.}$$
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \|O_i\|_2^2$, the procedure can use the classical representation to predict $\hat{o}_1, \ldots, \hat{o}_M$,
where $|\hat{o}_i - \mathrm{tr}(O_i \rho)| < \epsilon$, for all i.

Small for low-rank observables

For example:

- $M = 10^6$, $B = 1$, then naively we need $10^6/\epsilon^2$ measurements.
- This theorem shows that we only need $6 \log(10)/\epsilon^2$ measurements.

Furthermore, we don't need to know $O_1, \ldots, O_M$ in advance.

Application:
Quantum fidelity $|\psi\rangle\langle\psi|$

A dive into shadow tomography

Theorem (Huang et al.; 2020)

1. Given $B, \epsilon > 0$, the procedure learns a classical representation of an unknown quantum state ρ from
$$N = \mathcal{O}(B \log(M)/\epsilon^2) \text{ measurements.}$$
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \|O_i\|_2^2$, the procedure can use the classical representation to predict $\hat{o}_1, \ldots, \hat{o}_M$, where $|\hat{o}_i - \mathrm{tr}(O_i \rho)| < \epsilon$, for all i.

Application: fidelity $|\psi\rangle\langle\psi|$

For example:

- $M = 10^6$, $B = 1$, then naively we need $10^6/\epsilon^2$ measurements.
- This theorem shows that we only need $6 \log(10)/\epsilon^2$ measurements.

Furthermore, we don't need to know $O_1, \ldots, O_M$ in advance.

A dive into shadow tomography

Theorem (Huang et al.; 2020)

1. Given $B, \epsilon > 0$, the procedure learns a classical representation of an unknown quantum state ρ from
$$N = \mathcal{O}(B \log(M)/\epsilon^2) \text{ measurements.}$$
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \|O_i\|^2_{\text{shadow}}$, the procedure can use the classical representation to predict $\hat{o}_1, \ldots, \hat{o}_M$, where $|\hat{o}_i - \text{tr}(O_i \rho)| < \epsilon$, for all i.

Depends on randomized experiments

For example:

- $M = 10^6$, $B = 1$, then naively we need $10^6/\epsilon^2$ measurements.
- This theorem shows that we only need $6 \log(10)/\epsilon^2$ measurements.

Furthermore, we don't need to know $O_1, \ldots, O_M$ in advance.

Other applications

Algorithmic side: randomized experiments + data processing
Analysis side: geometric analysis + concentration inequality

- Cross platform verification [1, 2]
- Characterizing topological order [3, 4]
- Probing entanglement entropy [5, 6]
- Diagnosing quantum chaos [7]
- Learning quantum noise [8, 9]

See more examples in the review [10].

References:

[1] Elben, Andreas, et al. "Cross-platform verification of intermediate scale quantum devices." Physical review letters 124.1 (2020): 010504.
[2] Anshu, Anurag, Zeph Landau, and Yunchao Liu. "Distributed quantum inner product estimation." arXiv preprint arXiv:2111.03273 (2021).
[3] Elben, Andreas, et al. "Many-body topological invariants from randomized measurements in synthetic quantum matter." Science advances 6.15 (2020).
[4] Huang, Hsin-Yuan, et al. "Provably efficient machine learning for quantum many-body problems." arXiv preprint arXiv:2106.12627 (2021).
[5] Brydges, Tiff, et al. "Probing Rényi entanglement entropy via randomized measurements." Science 364.6437 (2019): 260-263.
[6] Elben, Andreas, et al. "Mixed-state entanglement from local randomized measurements." Physical Review Letters 125.20 (2020): 200501.
[7] Vermersch, Benoît, et al. "Probing scrambling using statistical correlations between randomized measurements." *Physical Review X* 9.2 (2019): 021061.
[8] Flammia, Steven T., and Joel J. Wallman. "Efficient estimation of Pauli channels." *ACM Transactions on Quantum Computing* 1.1 (2020): 1-32.
[9] *Helsen, Jonas, et al. "Estimating gate-set properties from random sequences." arXiv preprint arXiv:2110.13178 (2021).*
[10] Elben, Andreas, et al. "The randomized measurement toolbox." arXiv preprint arXiv:2203.11374 (2022)

Why randomized experiments?

A. In many cases, they are **asymptotically optimal!** Adaptively choosing experiments based on new information seems better, but often don't [1, 2, 3].

B. Randomization turns **bad scenarios** into low-probability events.

C. If information is **distributed evenly**, random sampling quickly converges.

References:
[1] Chen, Sitan, et al. "Exponential separations between learning with and without quantum memory." *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2022.
[2] Anshu, Anurag, Zeph Landau, and Yunchao Liu. "Distributed quantum inner product estimation." *arXiv preprint arXiv:2111.03273* (2021).
[3] Chen, Sitan, et al. "Tight Bounds for State Tomography with Incoherent Measurements." arXiv preprint *arXiv:2206.05265* (2022)

Outline

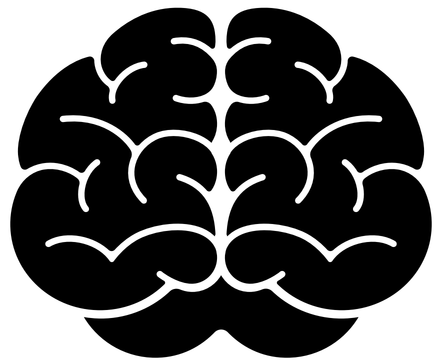
- Basic setting and examples

- Key ideas:

  Part I — Designing good learning algorithms

  Part II — Proving no good learning algorithms exist

- Outlook and open questions



Outline

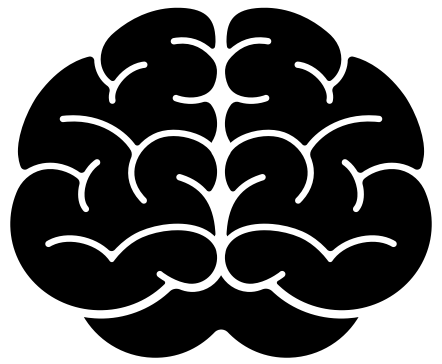
- Basic setting and examples

- Key ideas:

  Part I — Designing good learning algorithms

  Part II — Proving no good learning algorithms exist

- Outlook and open questions



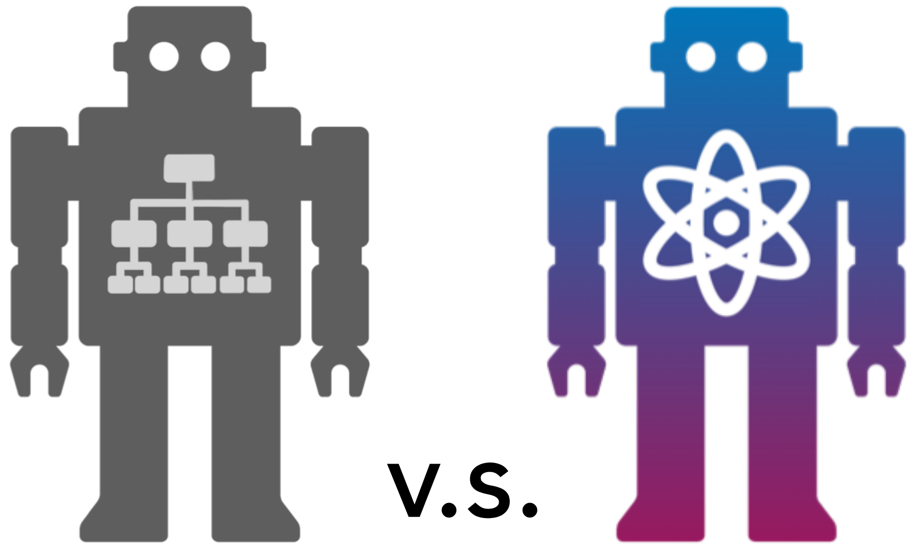
Key Ideas

Mostly focus on
Conventional Experiments

- Part II focuses on lower bounds (showing no good algorithms exist).

Part I

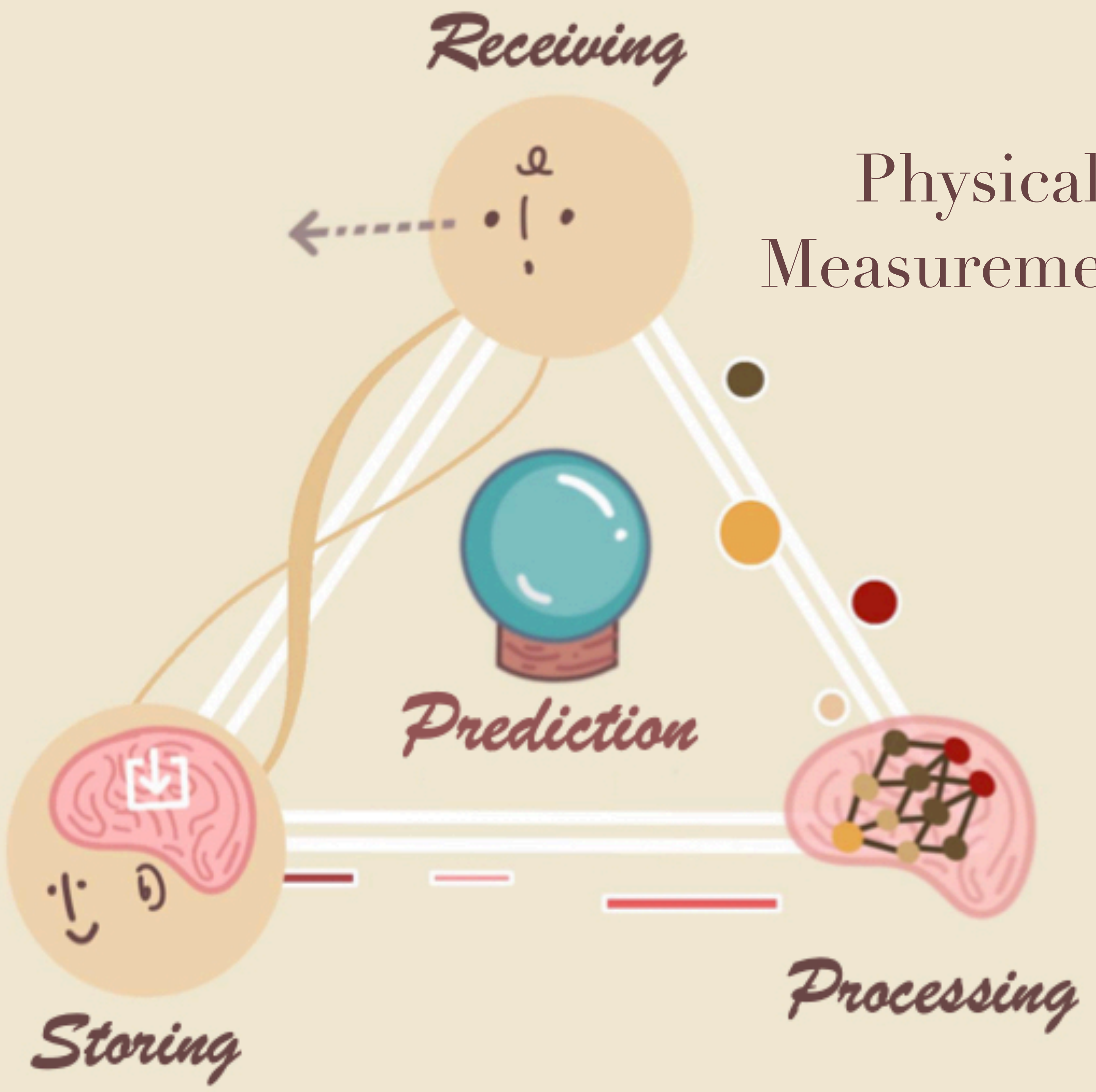
Part II



Conventional Experiments

What scientists currently do in the lab

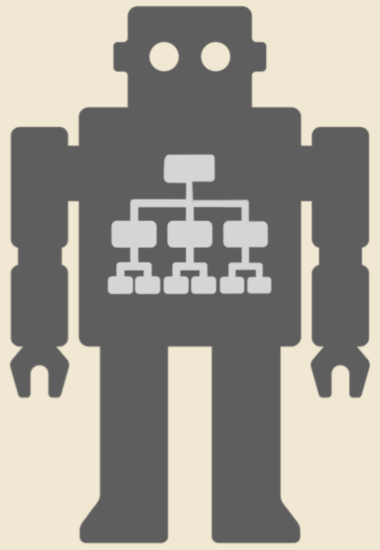
Receive, process, and store classical information



Physical Measurements

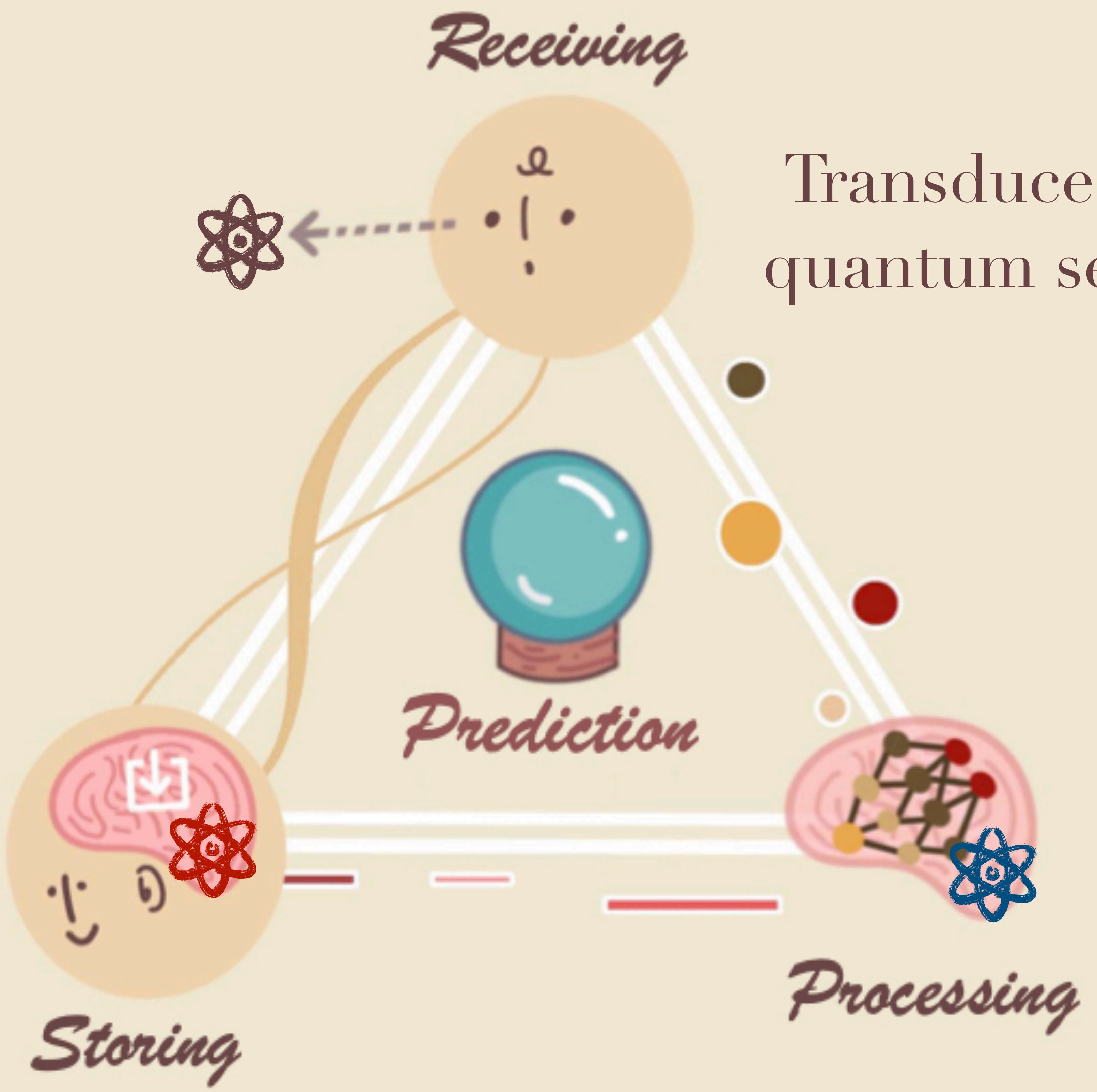
Classical Memory

Classical Computation



Quantum-enhanced Experiments

What future experiments could be like

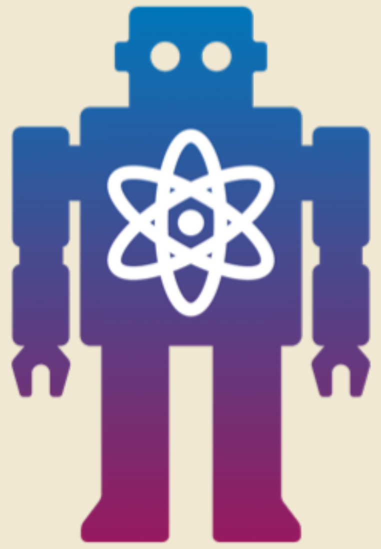


Transduce from quantum sensors

Quantum Memory

Quantum Computation

Receive, process, and store quantum information

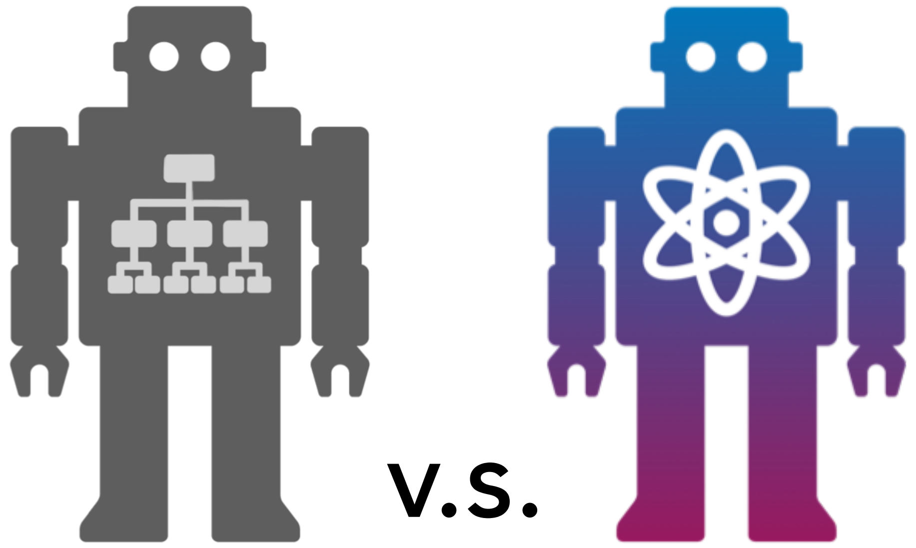


Key Ideas

- Proving lower bounds for conventional experiments (classical agents) helps us understand the potential quantum advantage in learning from experiments.

Part I

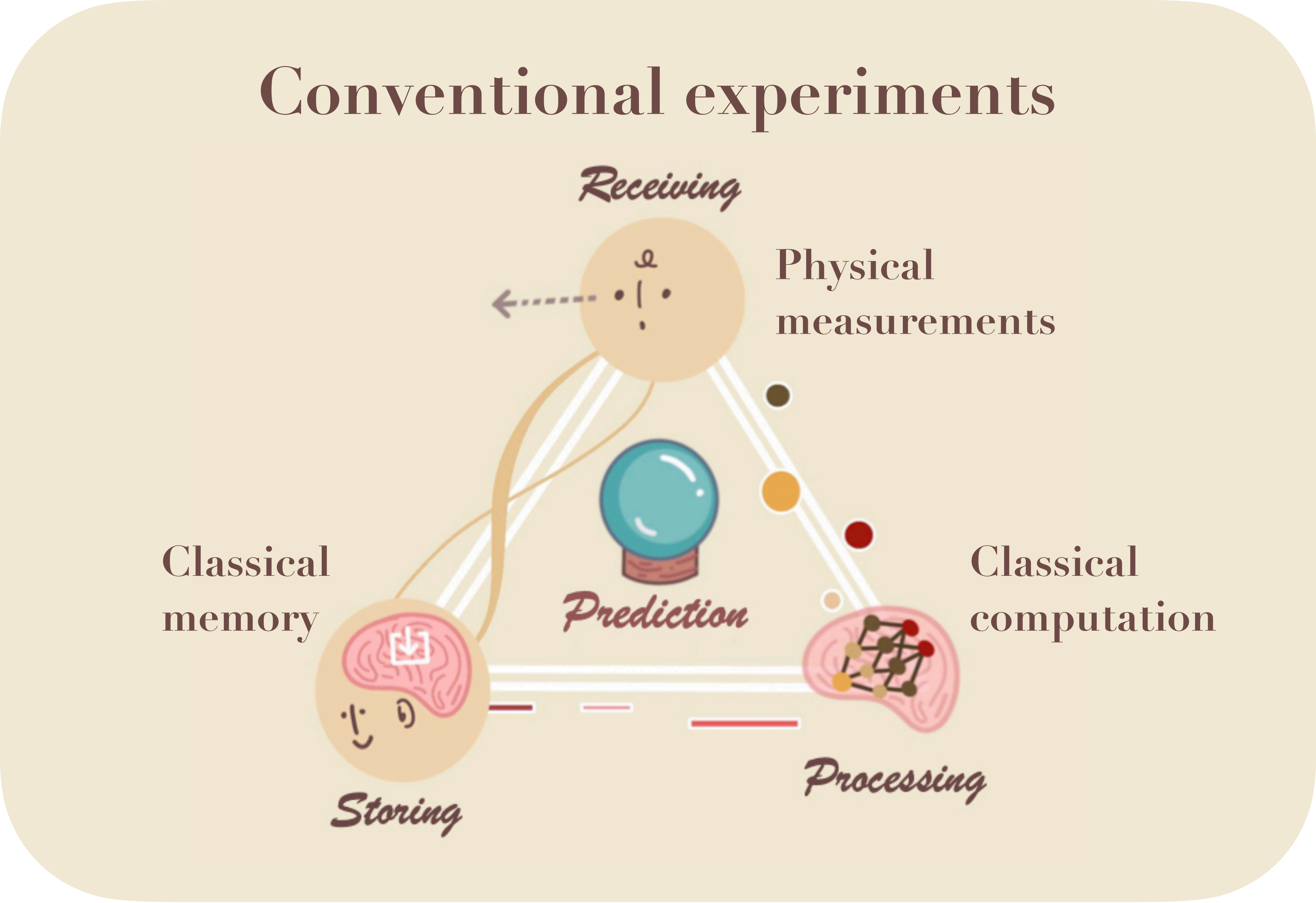
Part II

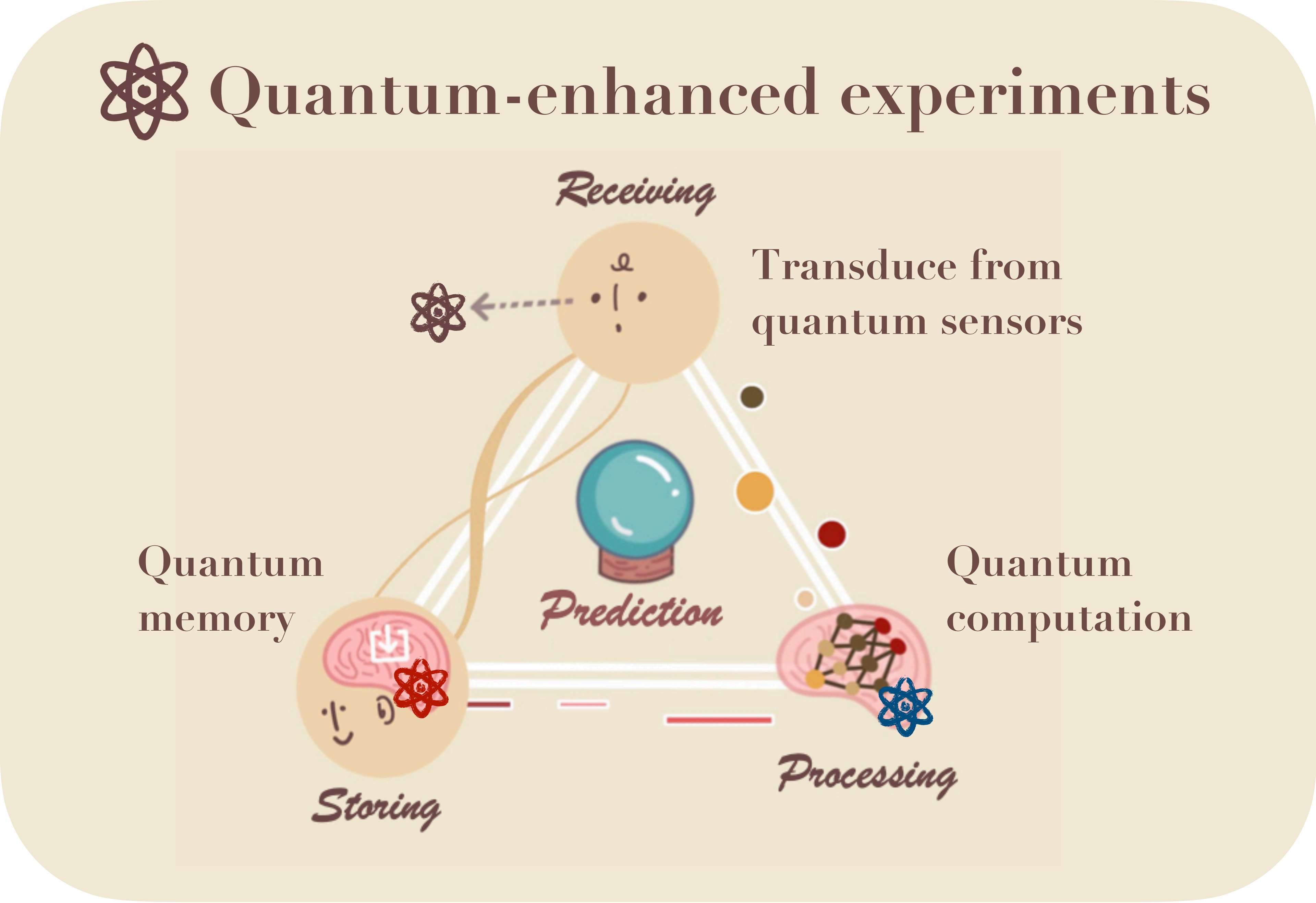


Mathematical Framework

- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

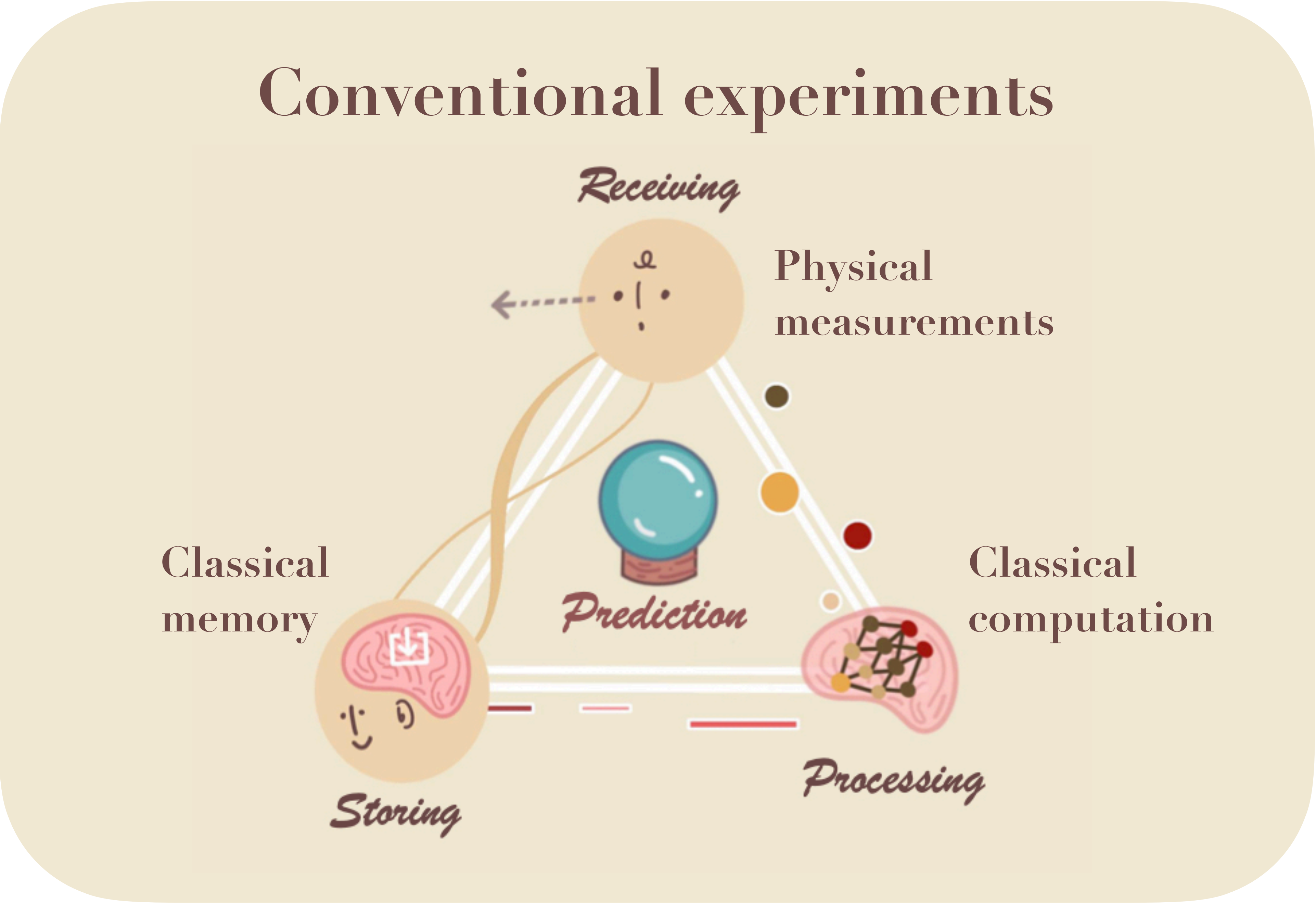
Related framework has been considered in [Bubeck, Chen, Li, FOCS'20], [Huang, Kueng, Preskill, PRL'21], [Aharonov, Cotler, Qi, '21]





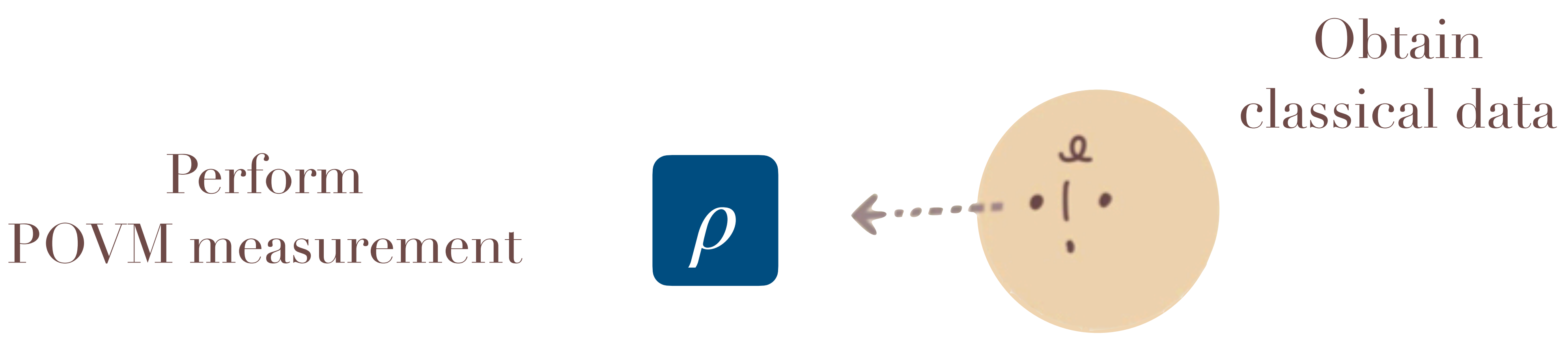
Mathematical Framework

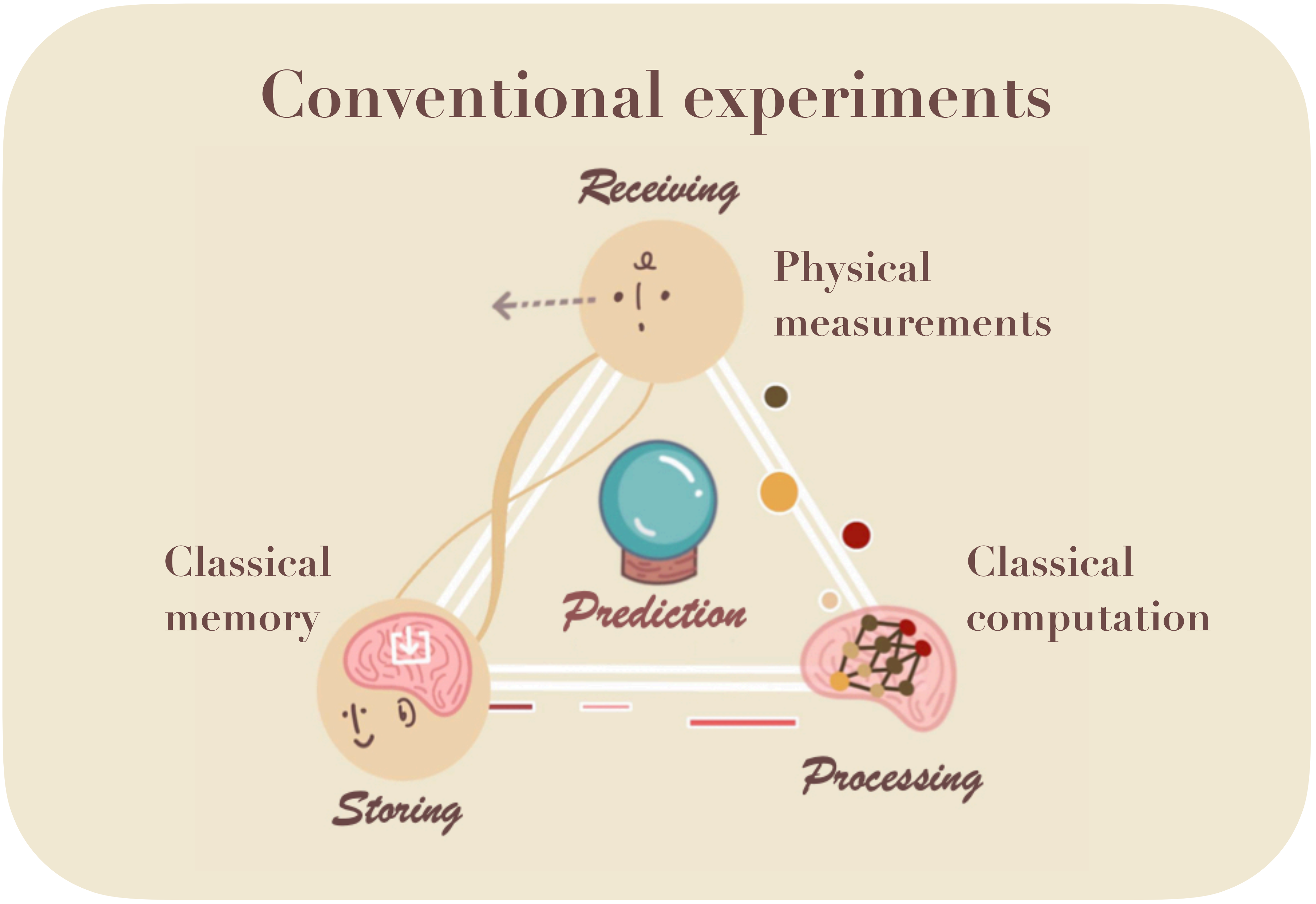
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.



Mathematical Framework

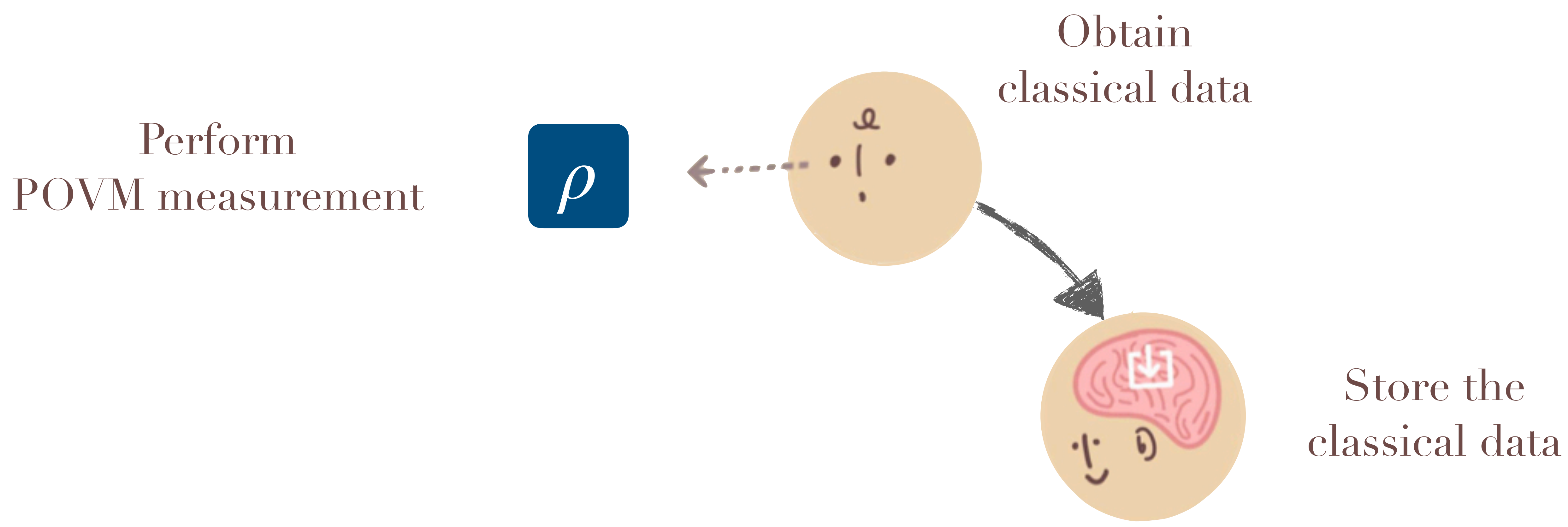
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

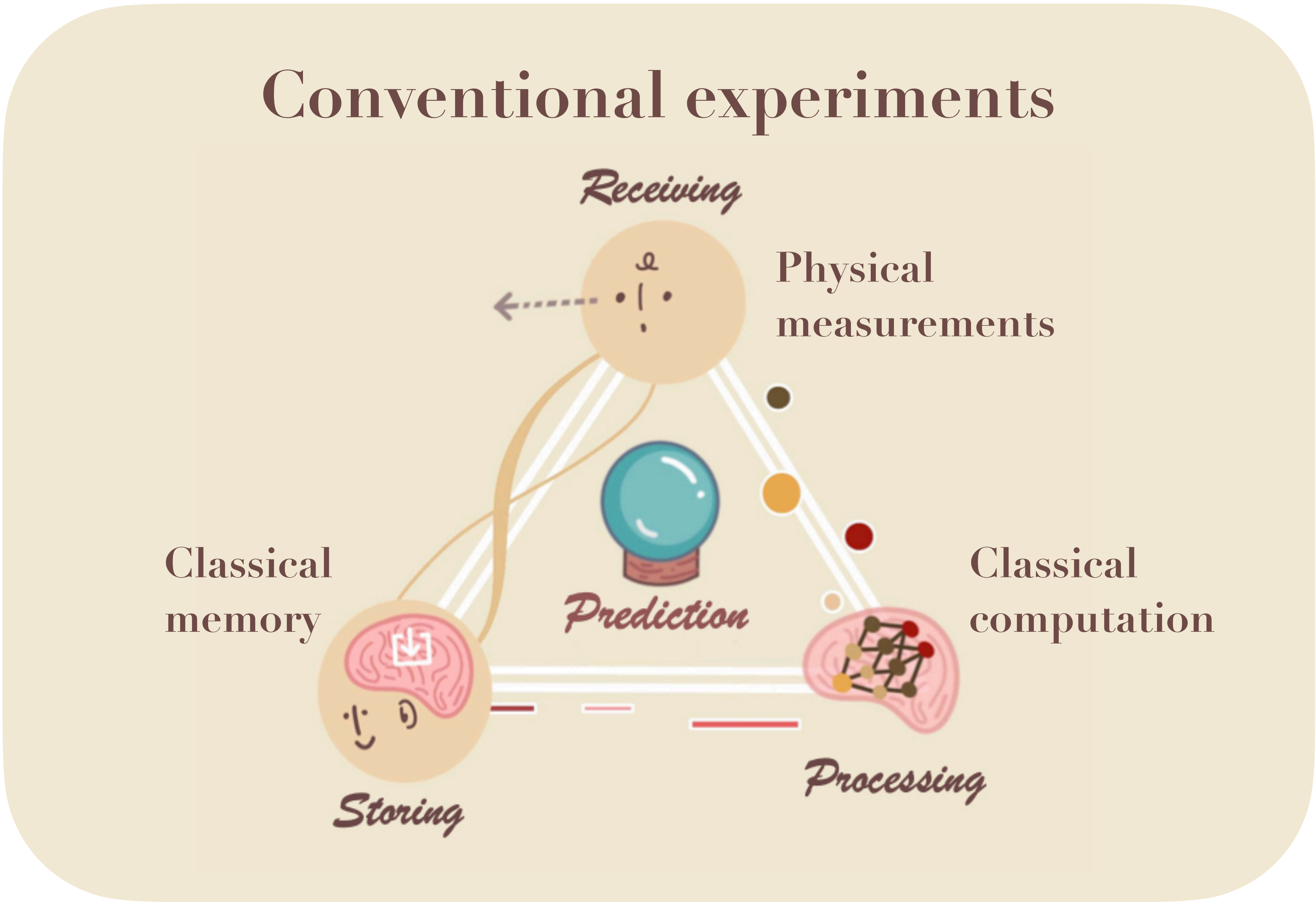




Mathematical Framework

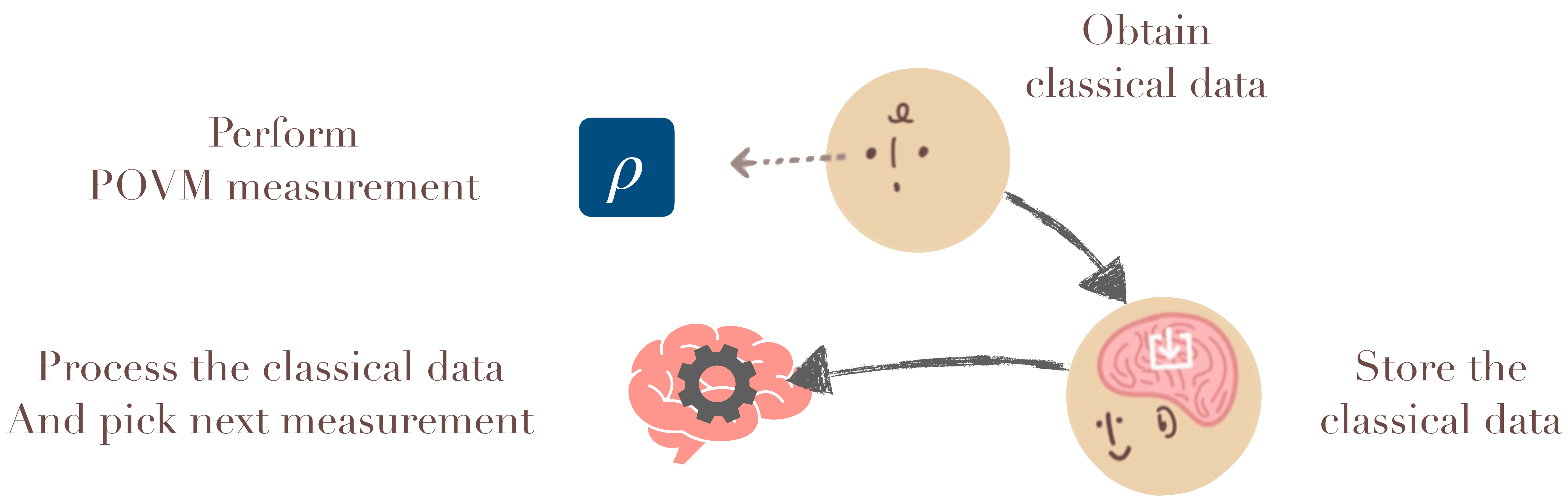
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

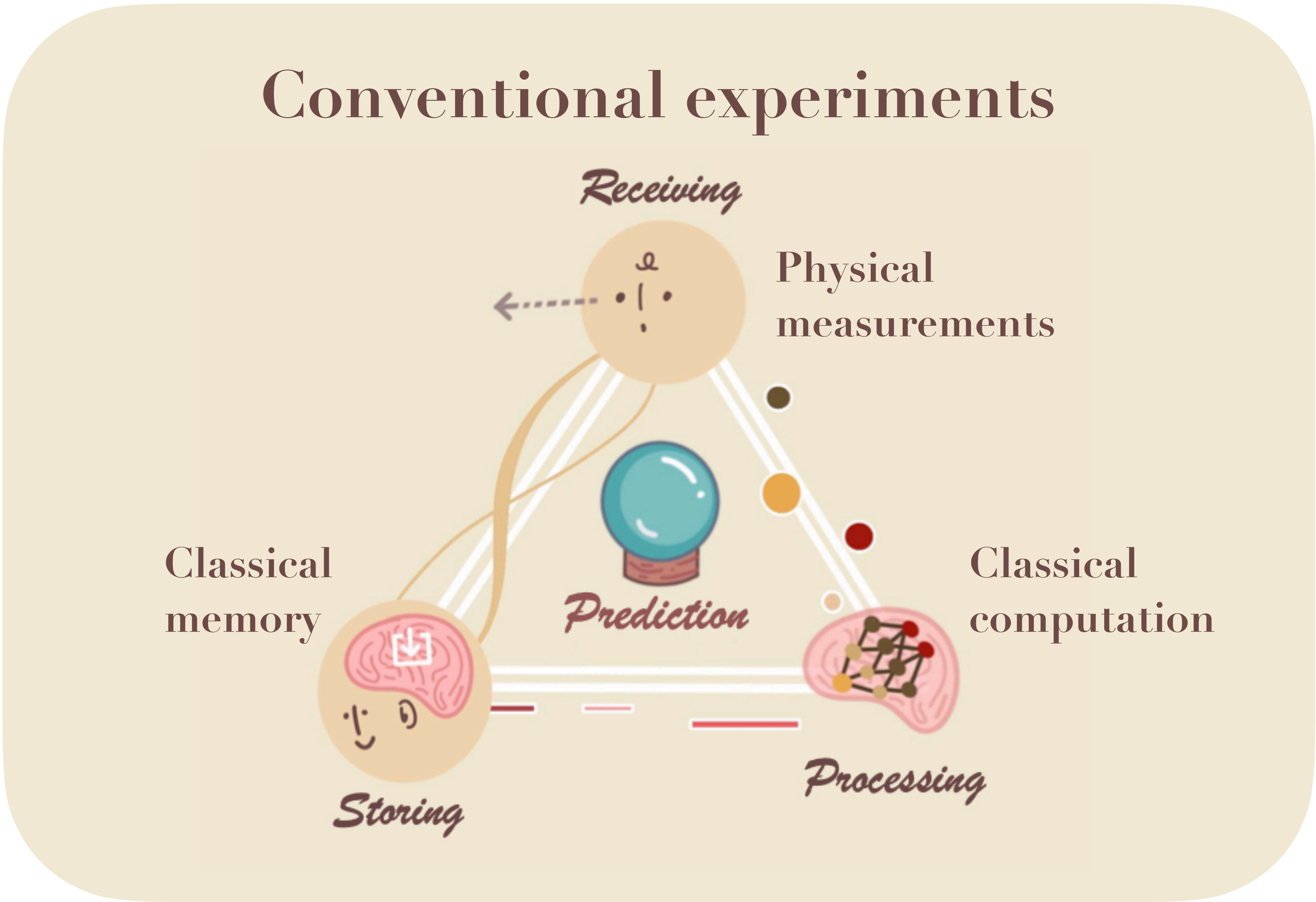




Mathematical Framework

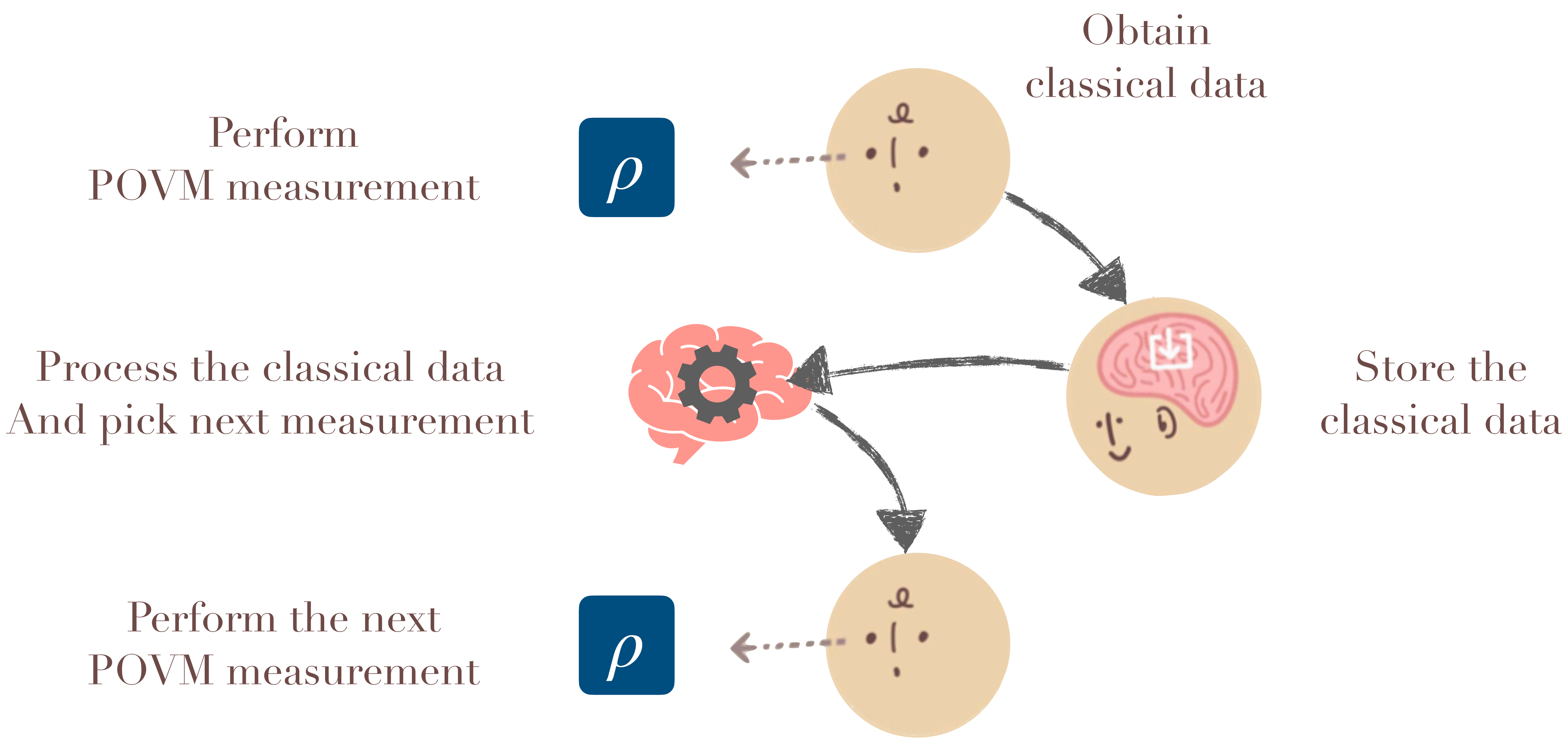
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

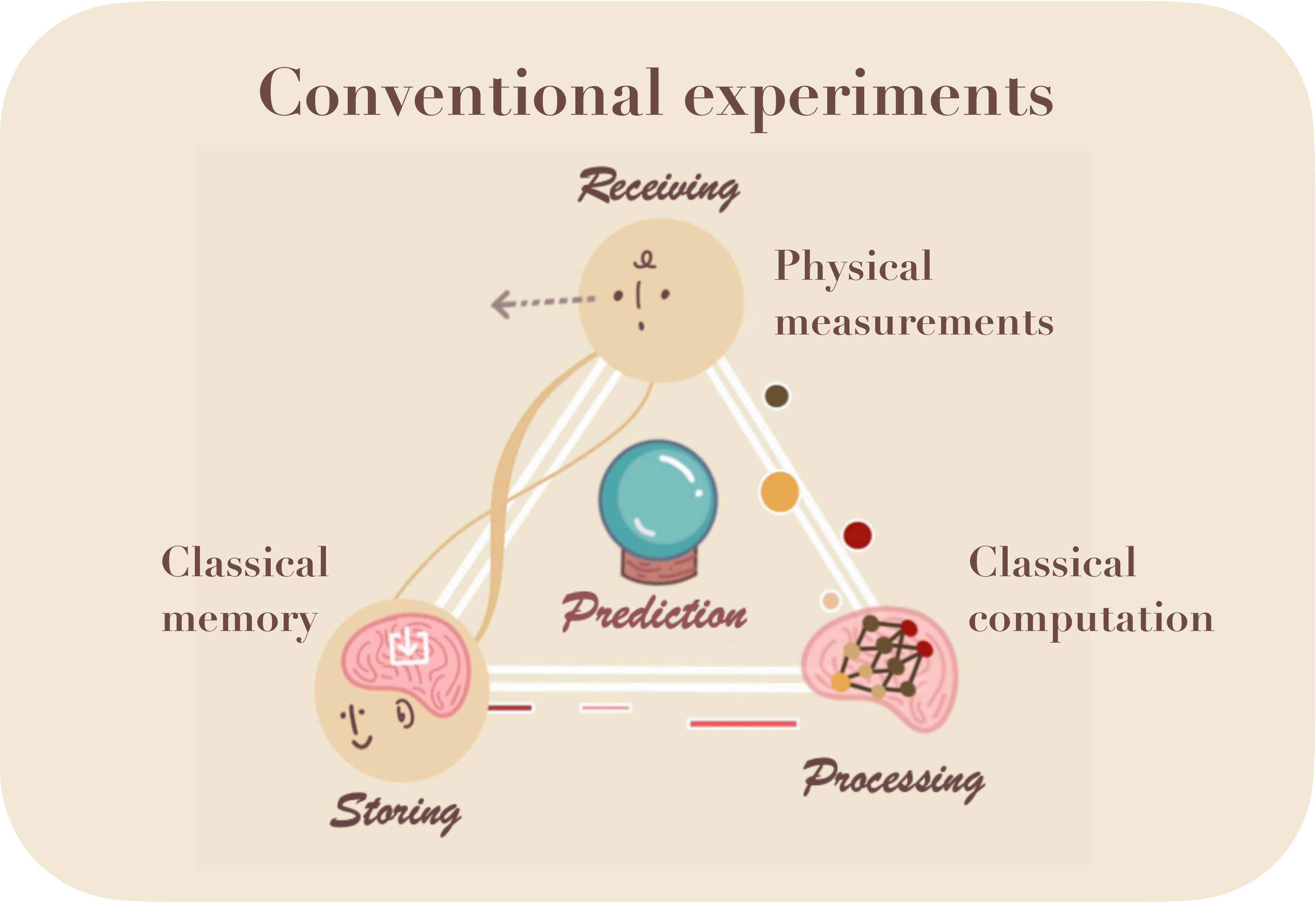




Mathematical Framework

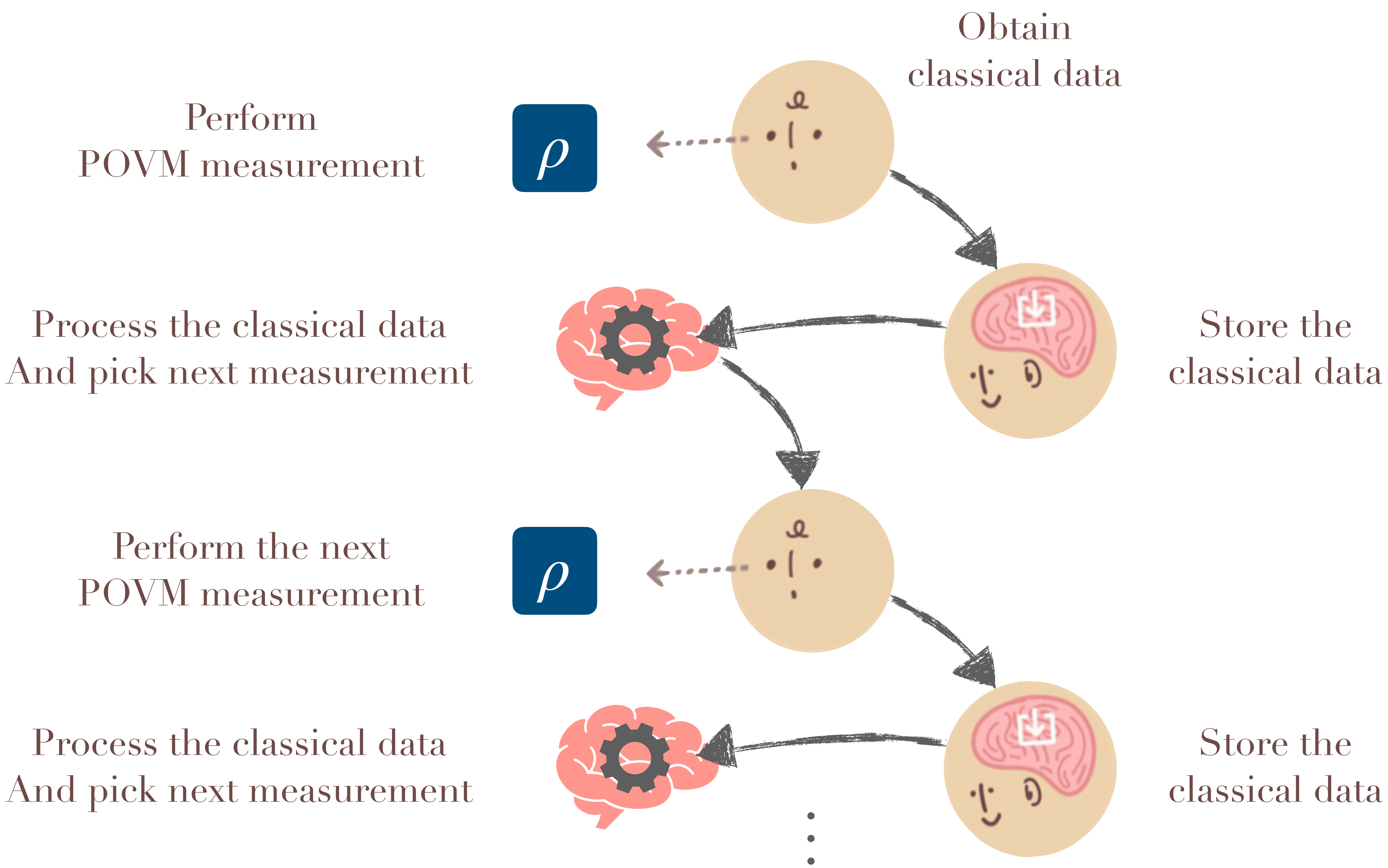
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

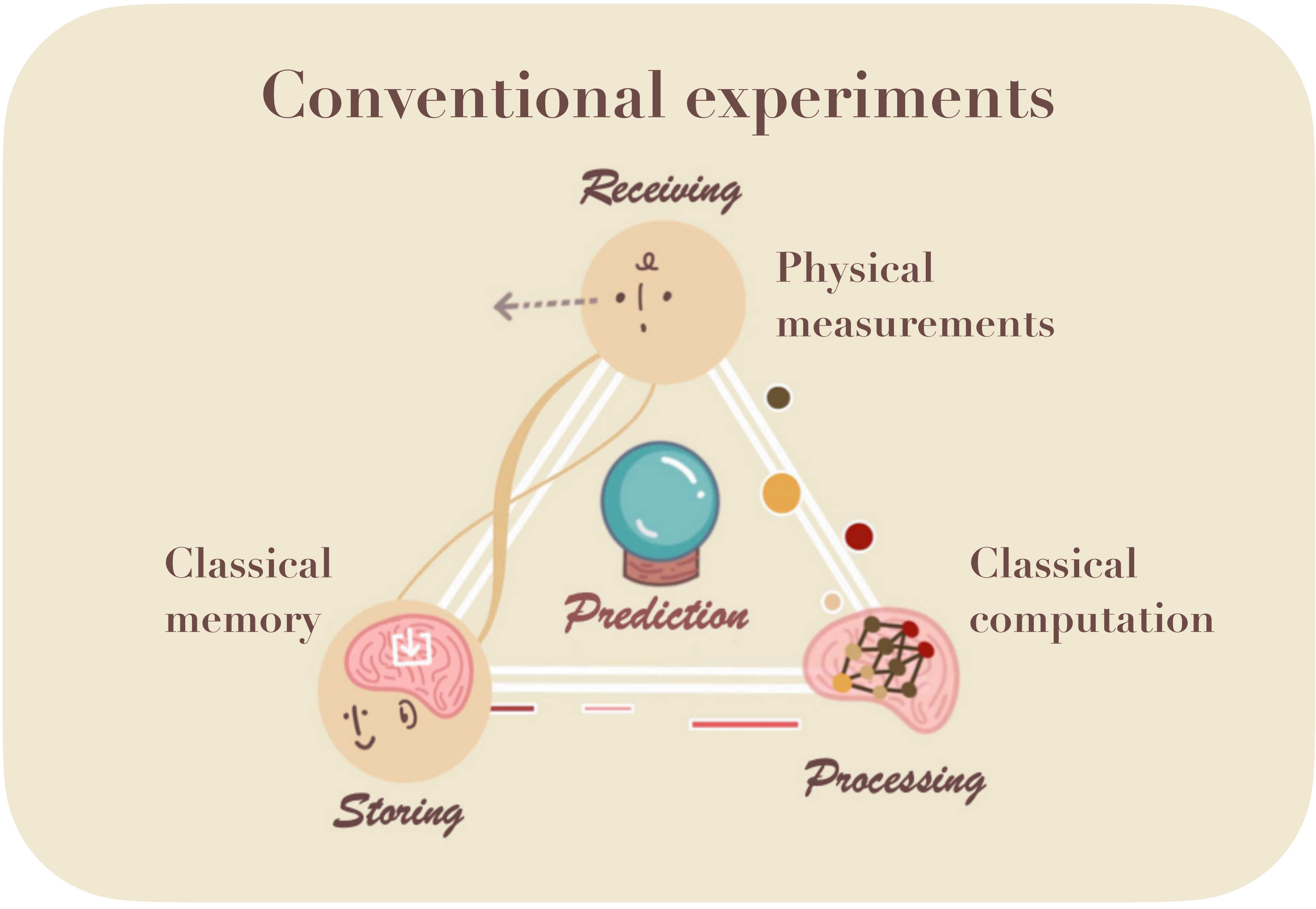




Mathematical Framework

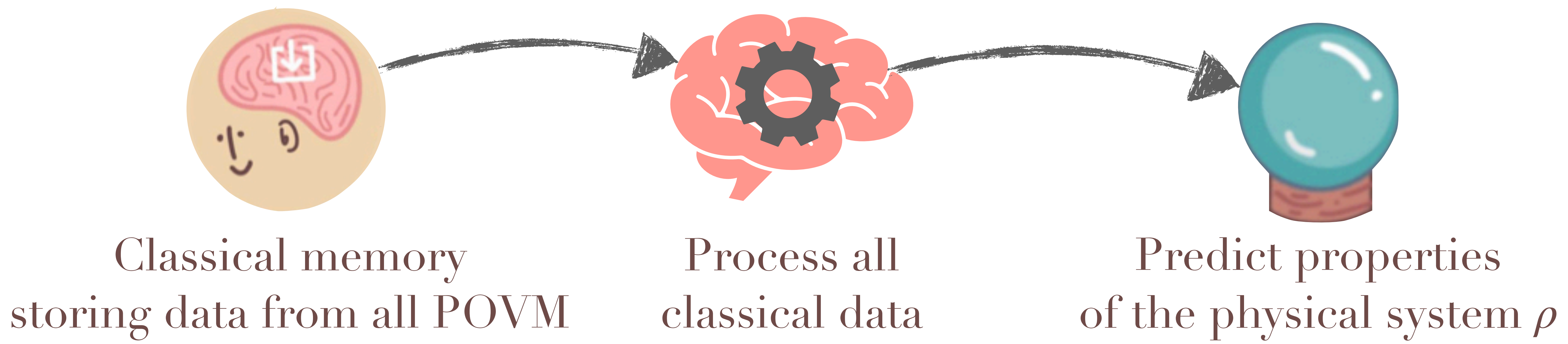
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

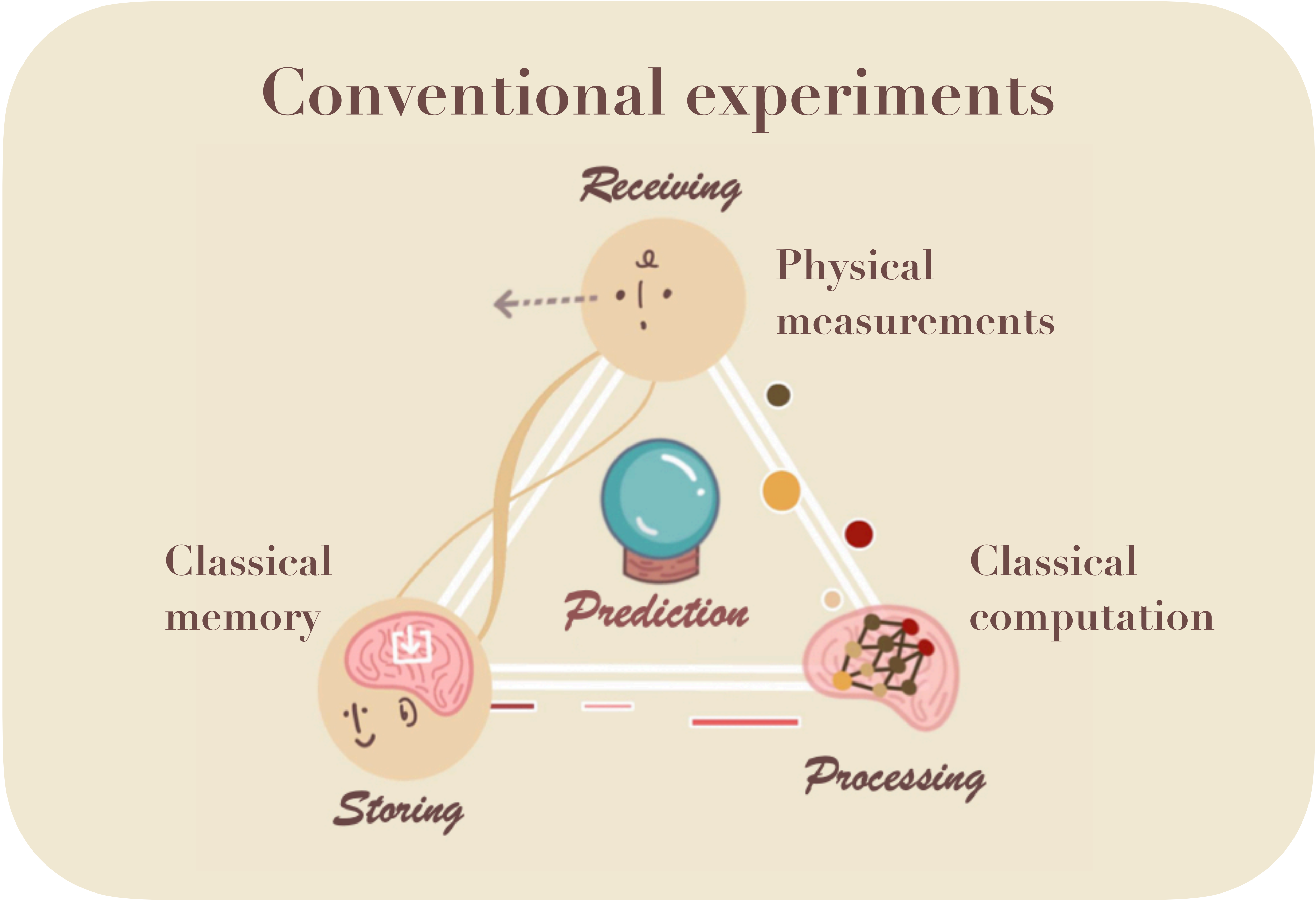




Mathematical Framework

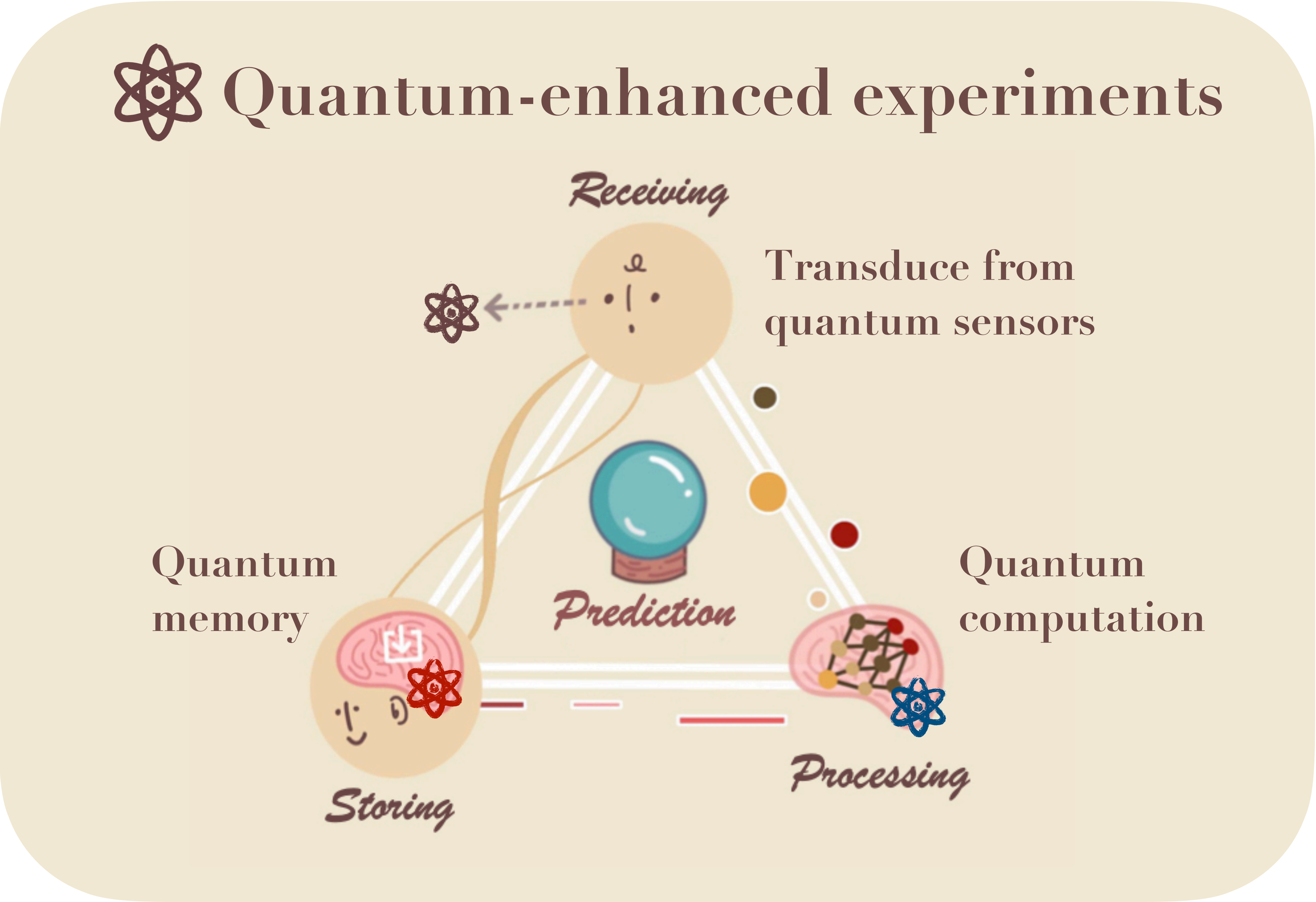
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.





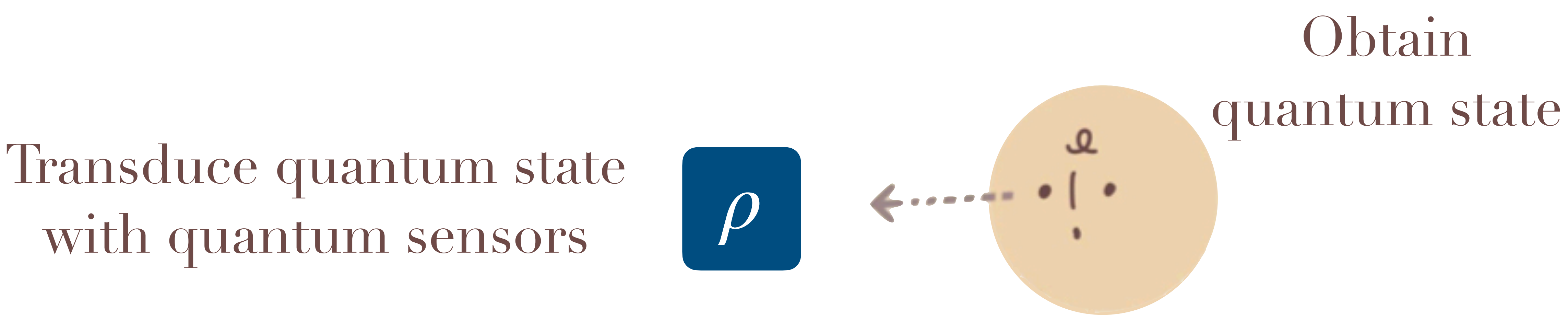
Mathematical Framework

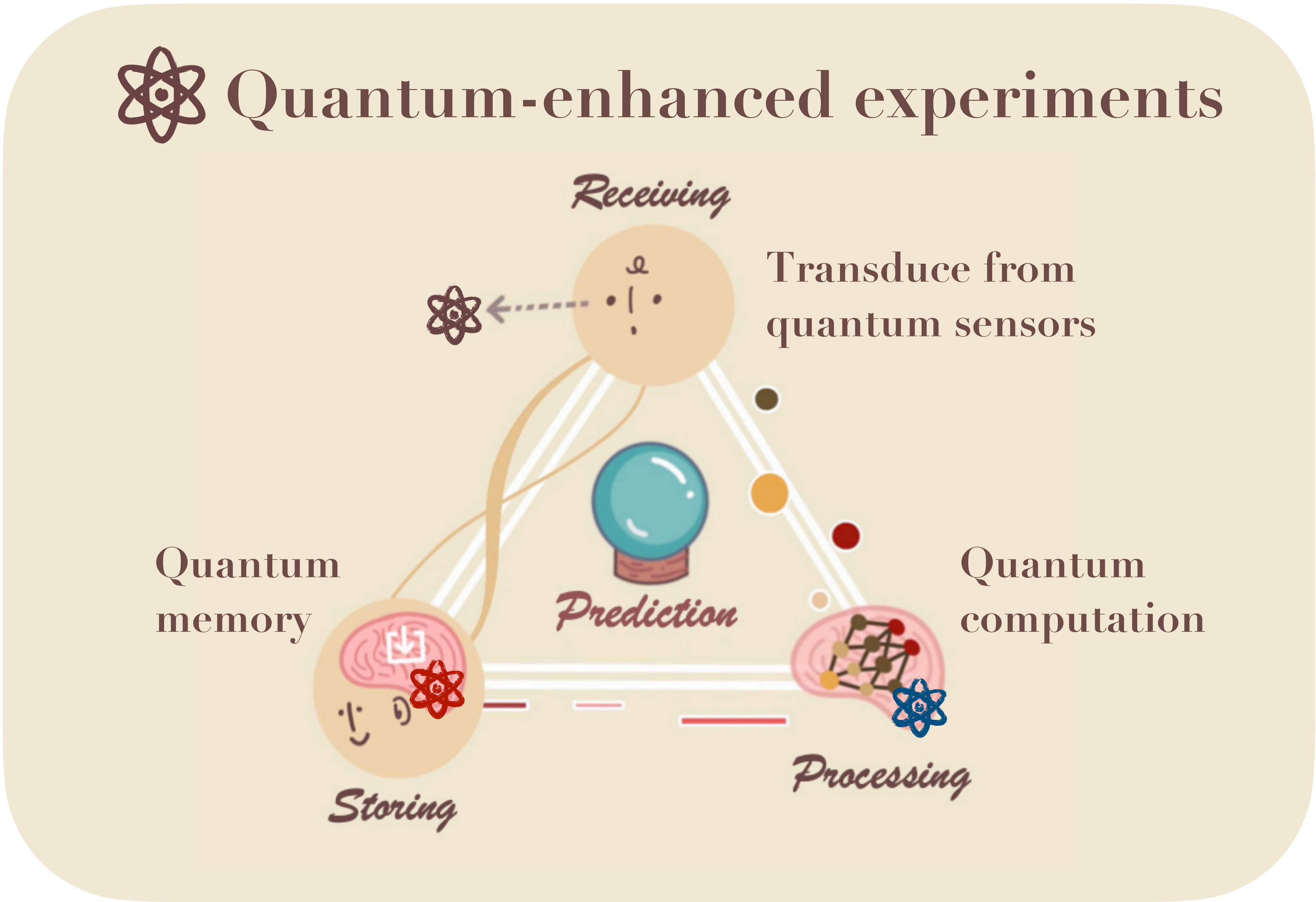
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.



Mathematical Framework

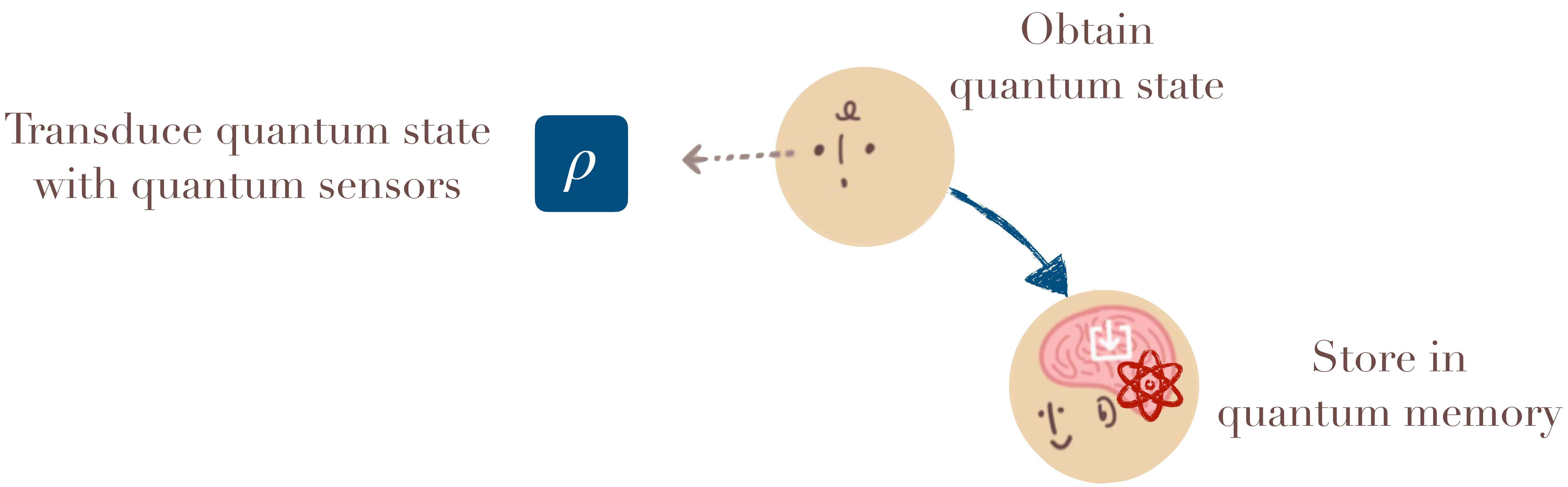
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

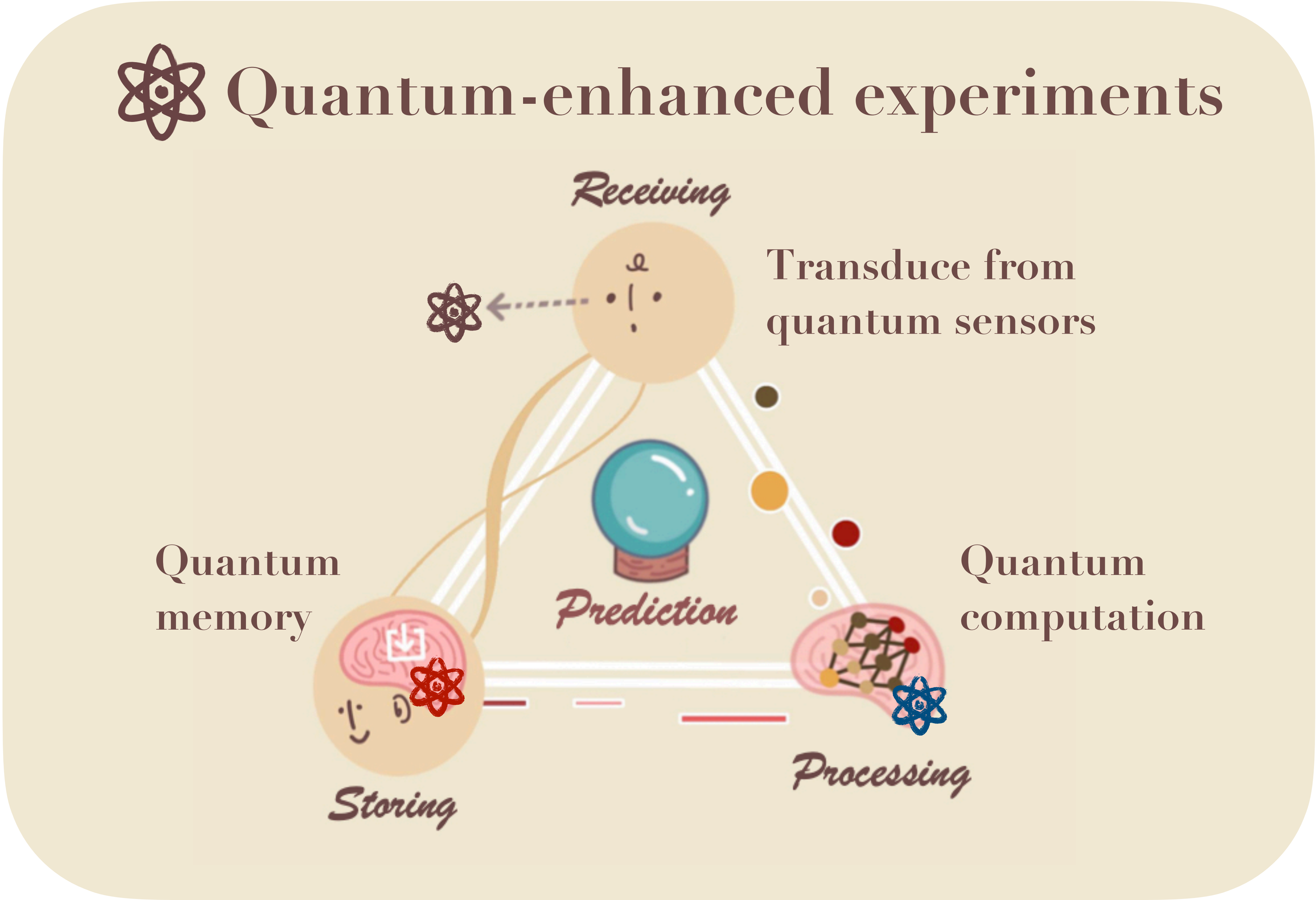




Mathematical Framework

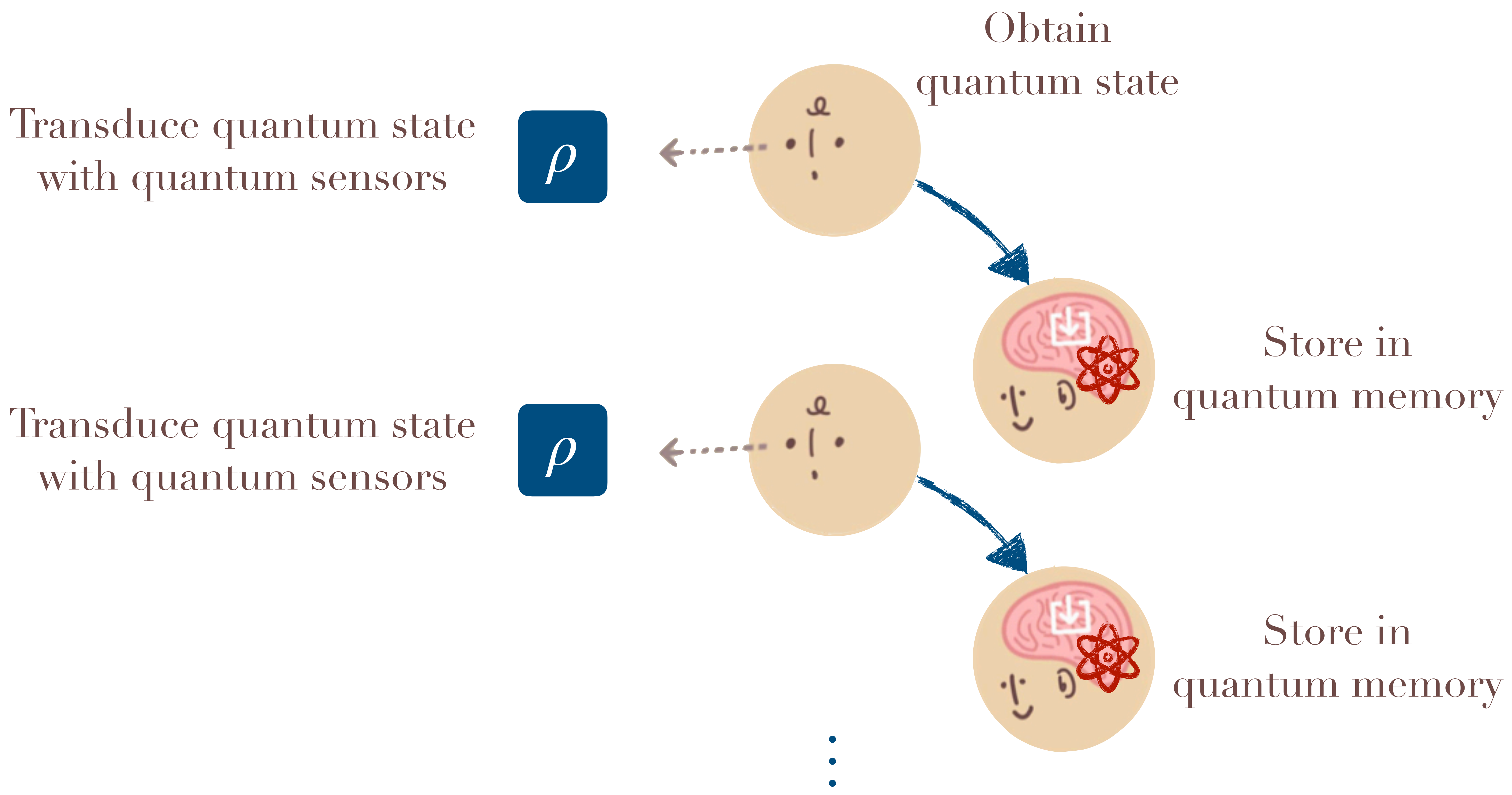
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

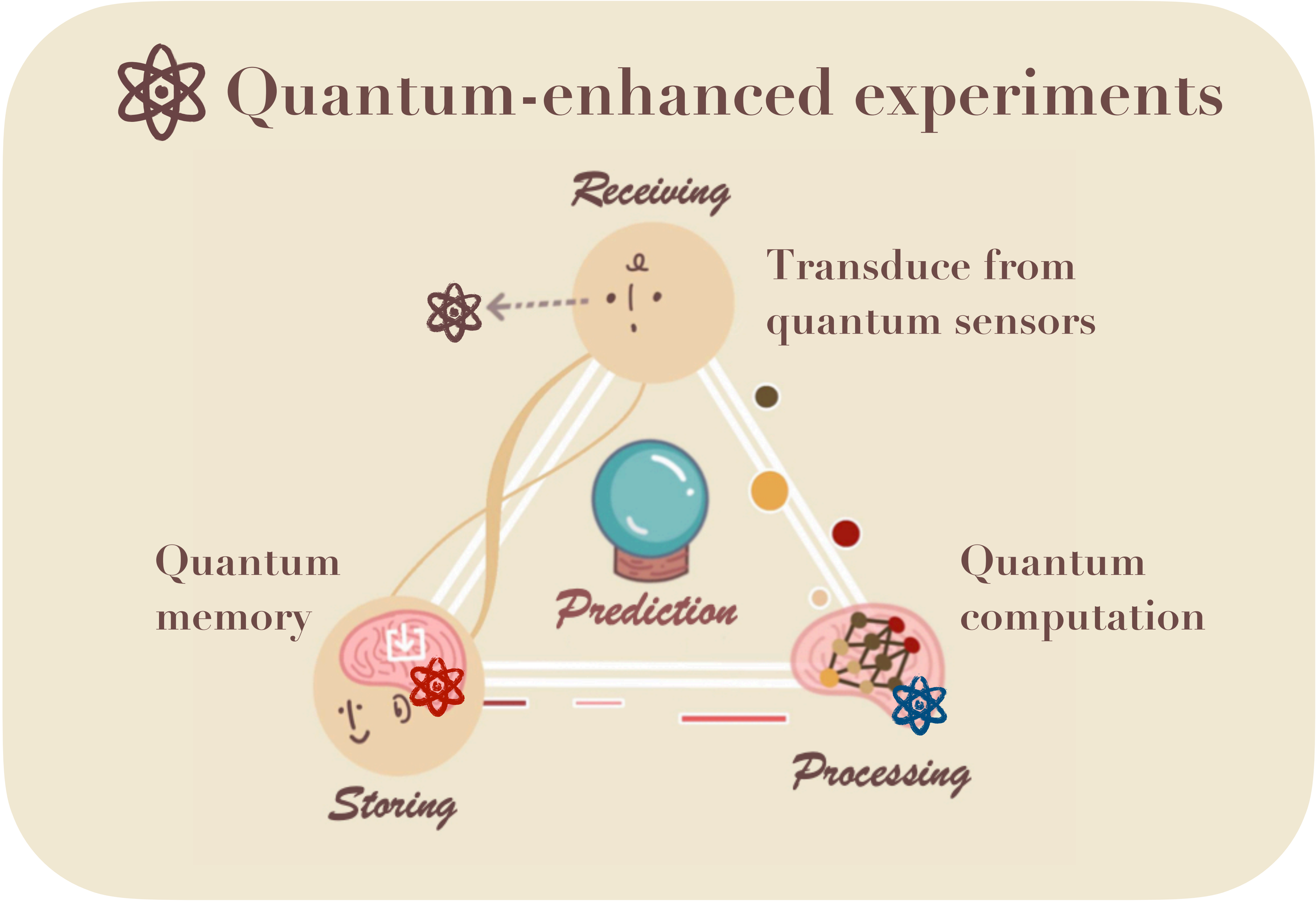




Mathematical Framework

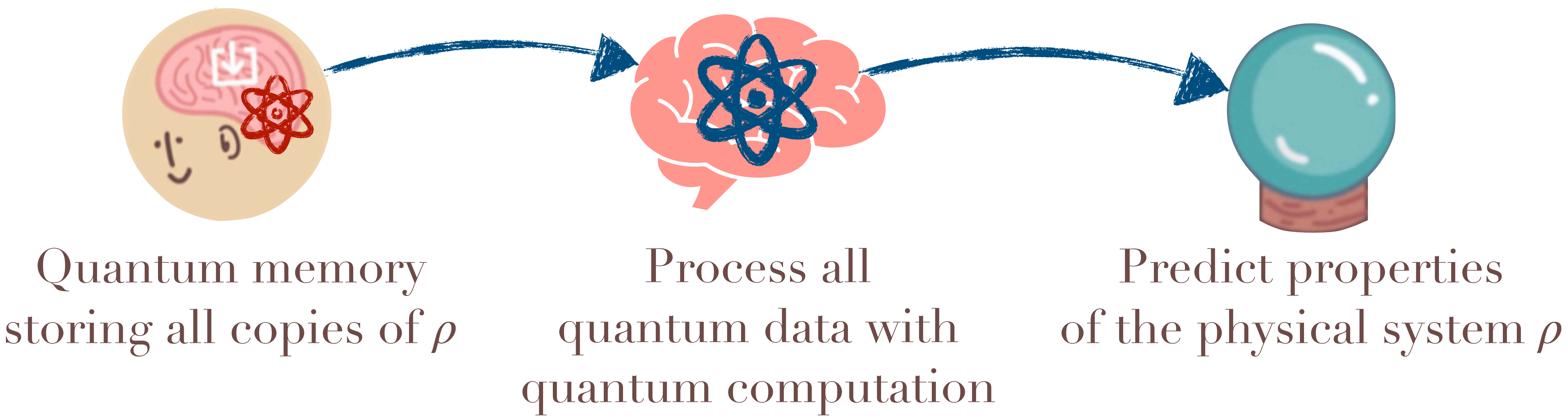
- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.

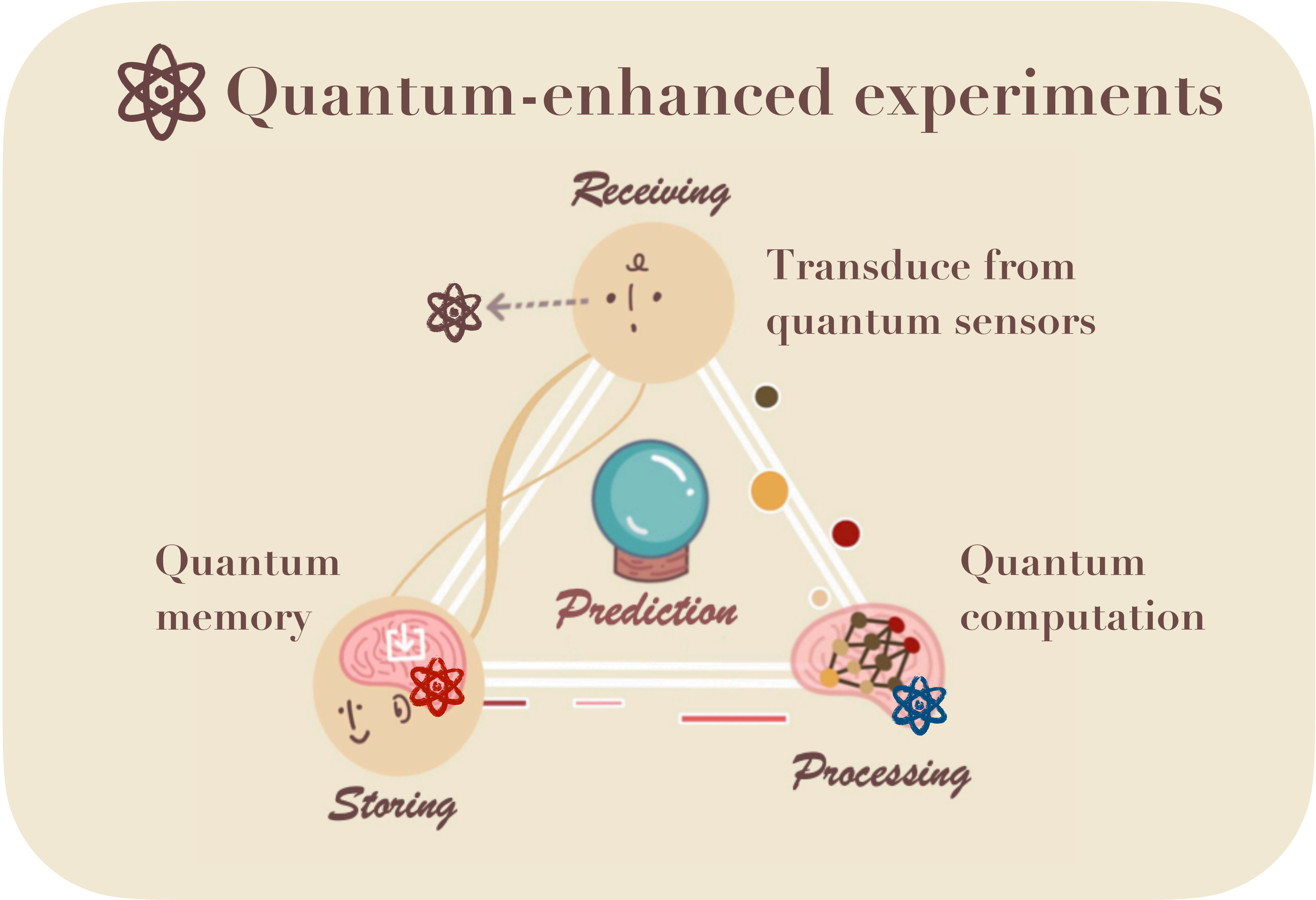




Mathematical Framework

- We consider a simple task of learning about an unknown physical system ρ (density matrix).
- Assume that a physical source that could generate a single copy of ρ at a time.



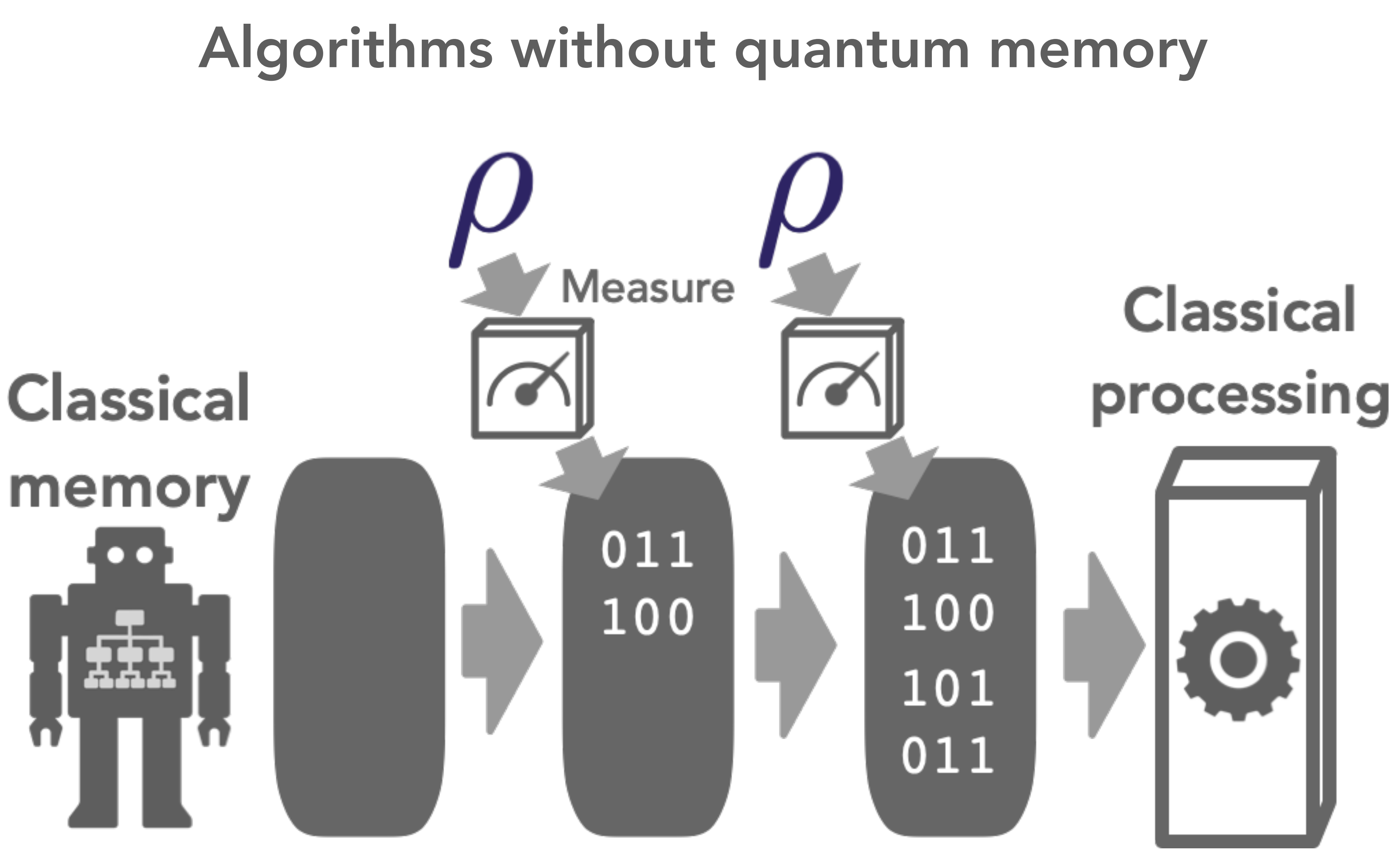


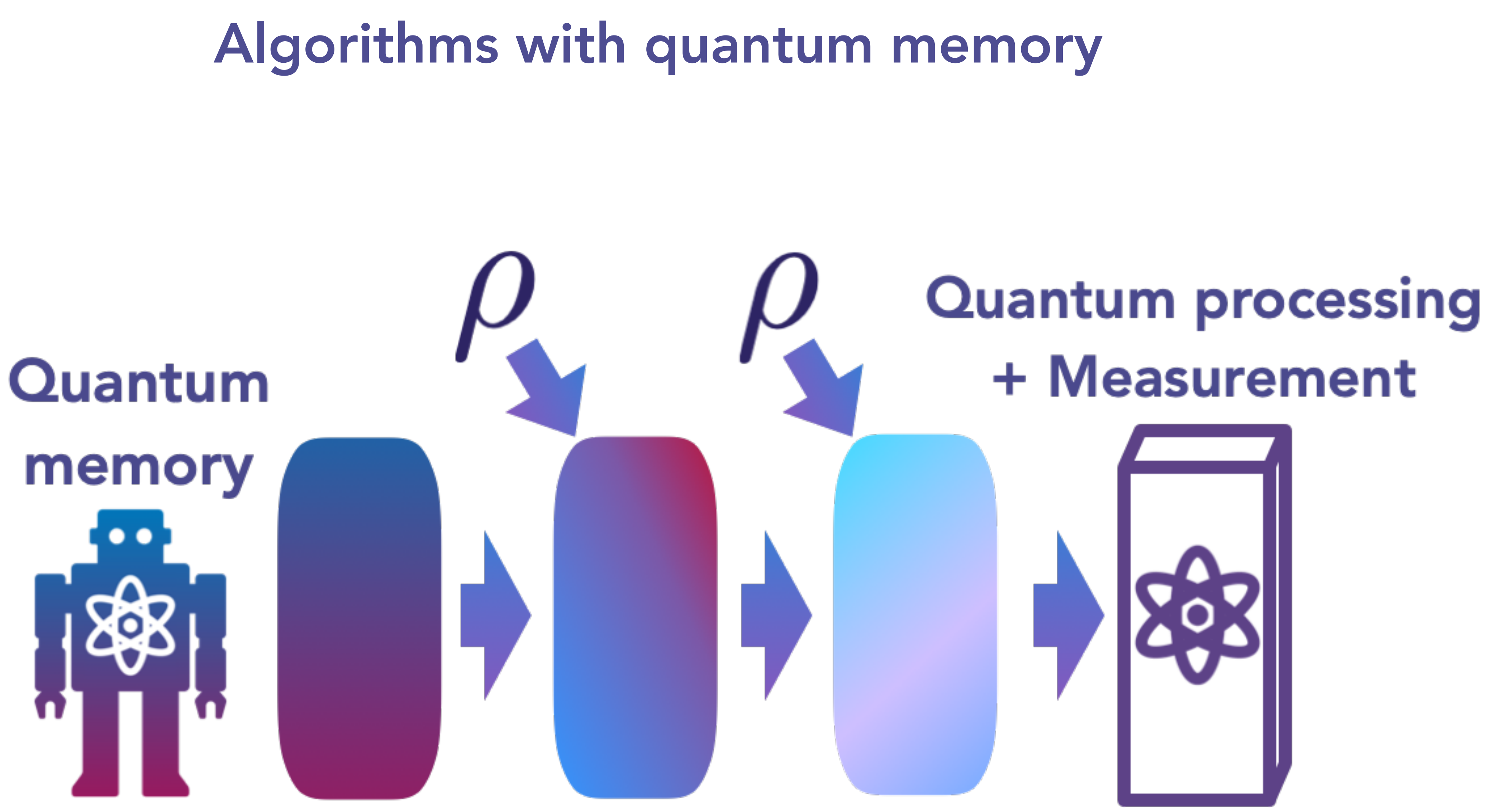
Mathematical Framework

- Main difference:

  Having quantum memory for entangling quantum information from past and future experiments.

- We can then analyze the possible protocols/algorithms to study their learning ability.





Quantum advantage in predicting properties

- The classical/quantum agent learns a classical model of the n-qubit state ρ.
- Subsequently, one can use the classical model to predict $|\mathrm{Tr}(P\rho)|$ for an observable P chosen from $\{I, X, Y, Z\}^{\otimes n}$.

Theorem

Classical agent needs $\Omega(2^n)$ experiments to predict observable from the set, but quantum agent only need $\mathcal{O}(n)$ experiments to predict all 4^n observables.

Quantum advantage in predicting properties

- The classical/quantum agent learns a classical model of the n-qubit state ρ.
- Subsequently, one can use the classical model to predict $|\mathrm{Tr}(P\rho)|$ for an observable P chosen from $\{I, X, Y, Z\}^{\otimes n}$.

Theorem

Classical agent needs $\Omega(2^n)$ experiments to predict observable from the set, but quantum agent only need $\mathcal{O}(n)$ experiments to predict all 4^n observables.

Uncertainty principle significantly hinders the learning ability of classical agent, but surprisingly not the ability of a quantum agent.

Quantum advantage in predicting properties

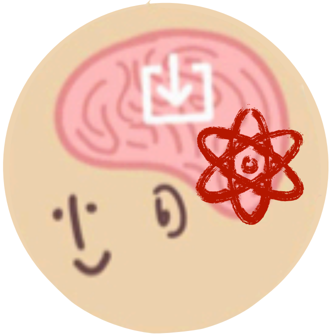
- The classical/quantum agent learns a classical model of the n-qubit state ρ.
- Subsequently, one can use the classical model to predict $|\mathrm{Tr}(P\rho)|$ for an observable P chosen from $\{I, X, Y, Z\}^{\otimes n}$.

Theorem

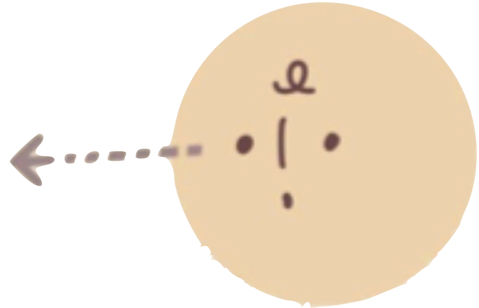
Classical agent needs $\Omega(2^n)$ experiments to predict observable from the set, but quantum agent only need $\mathcal{O}(n)$ experiments to predict all 4^n observables.

Exponential quantum advantage is present even when the state ρ is a classical distribution over product states (no entanglement!).

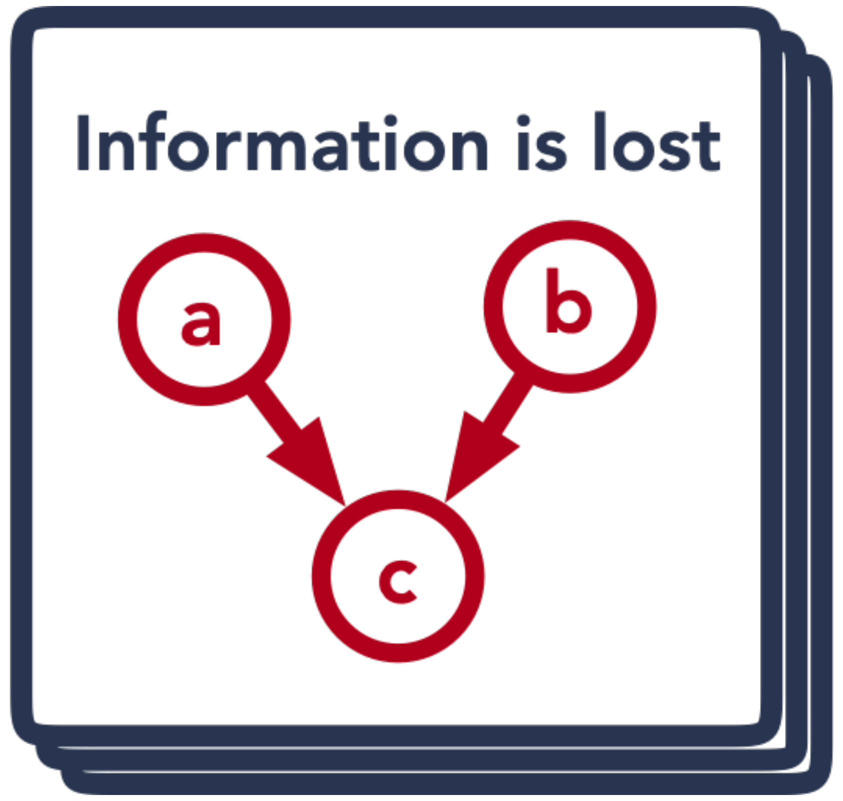
[CCHL21] Chen, Cotler, Huang, Li. Exponential separations between learning with and without quantum memory, FOCS 2021.
[HKP21] Huang, Kueng, Preskill. Information-theoretic bounds on quantum advantage in machine learning, PRL 2021.

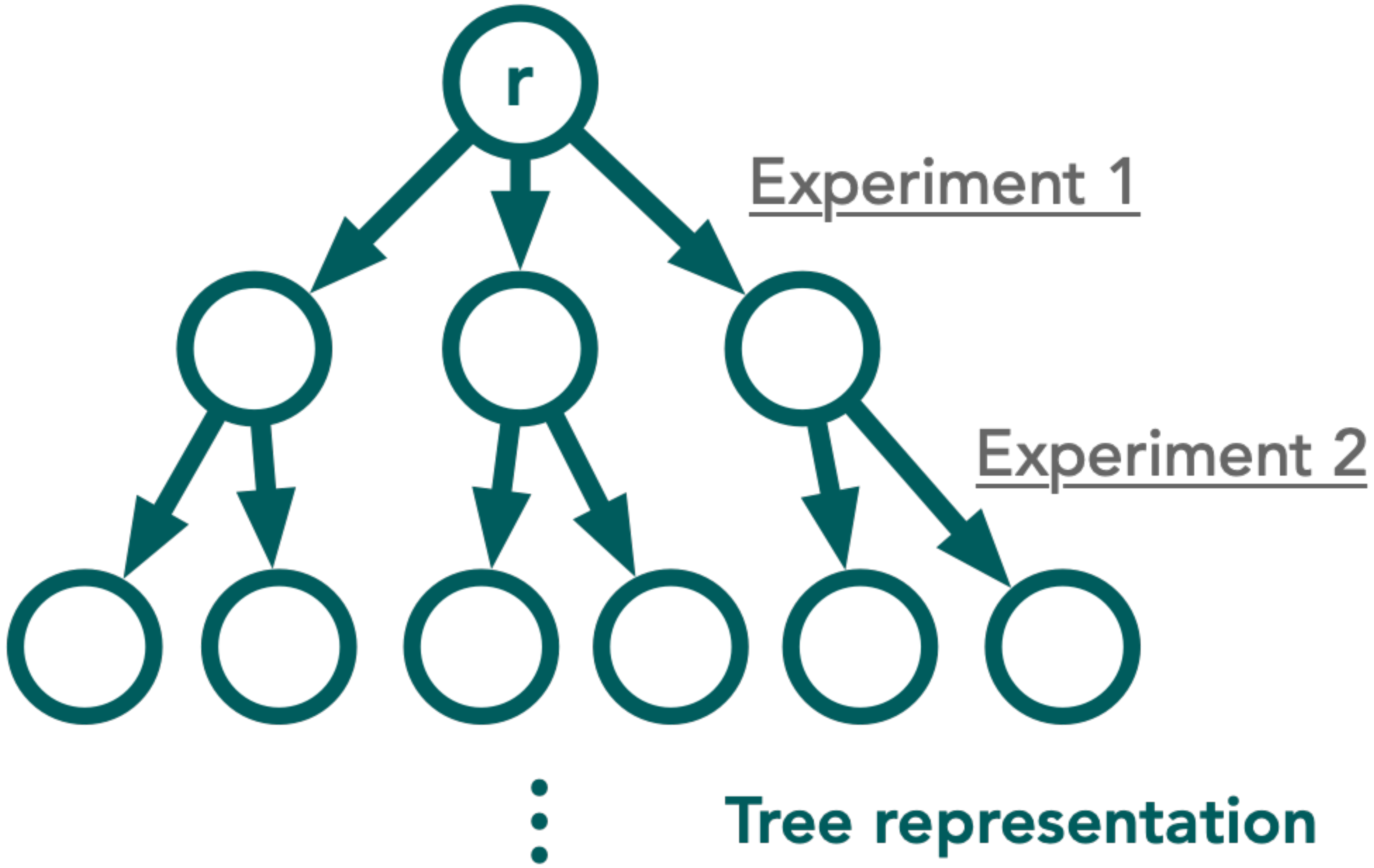


Proof Sketch: Tree representation

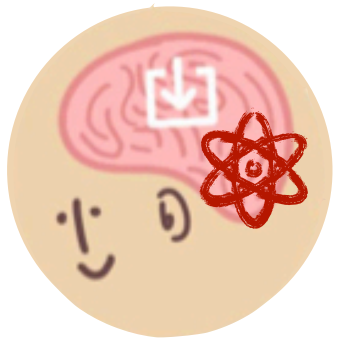


- Consider the lower bound $\Omega(2^n)$ for classical agents; See [HKP21] for upper bound $\mathcal{O}(n)$.
- We consider a graphical representation for the memory state of the classical agent.

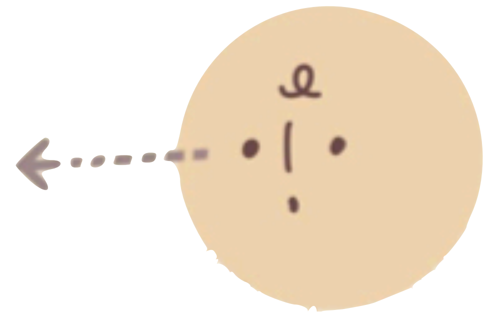




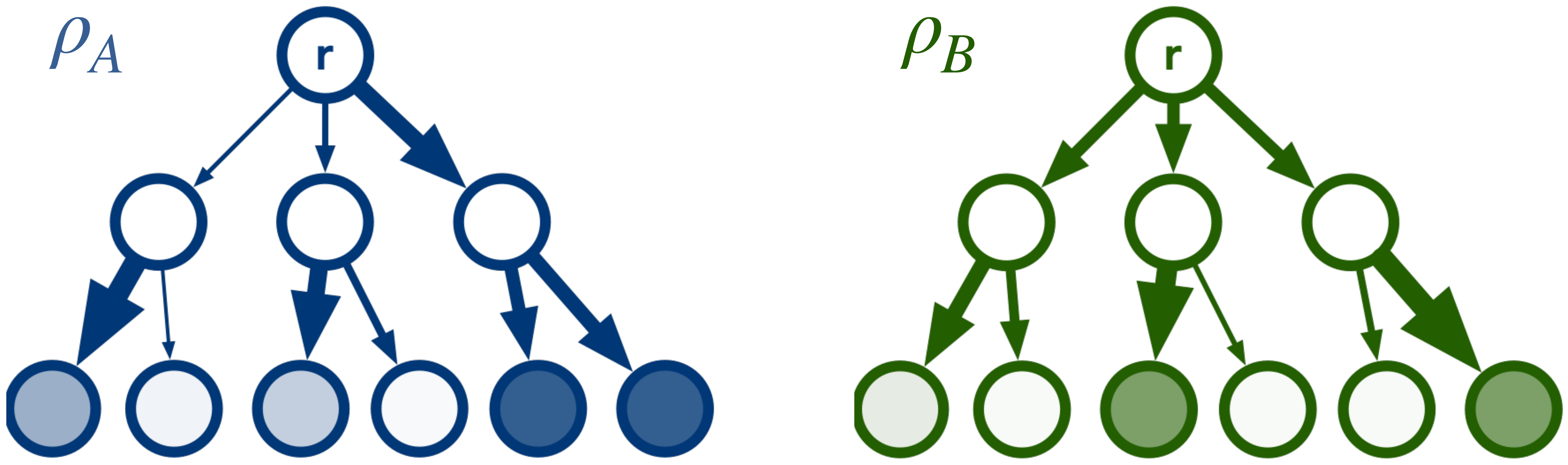
[CCHL21] Chen, Cotler, Huang, Li. Exponential separations between learning with and without quantum memory, FOCS 2021.
[HKP21] Huang, Kueng, Preskill. Information-theoretic bounds on quantum advantage in machine learning, PRL 2021.



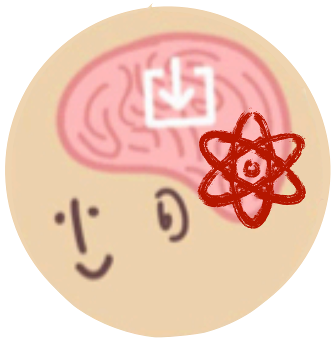
Proof Sketch: Tree representation



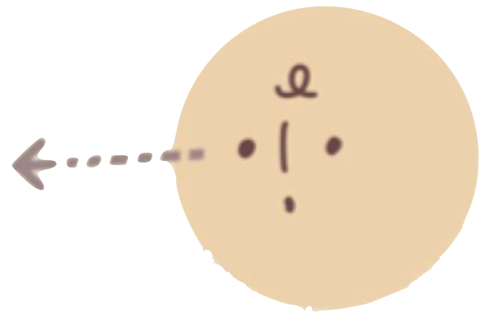
- Consider the lower bound $\Omega(2^n)$ for classical agents; See [HKP21] for upper bound $\mathcal{O}(n)$.
- We consider a graphical representation for the memory state of the classical agent.



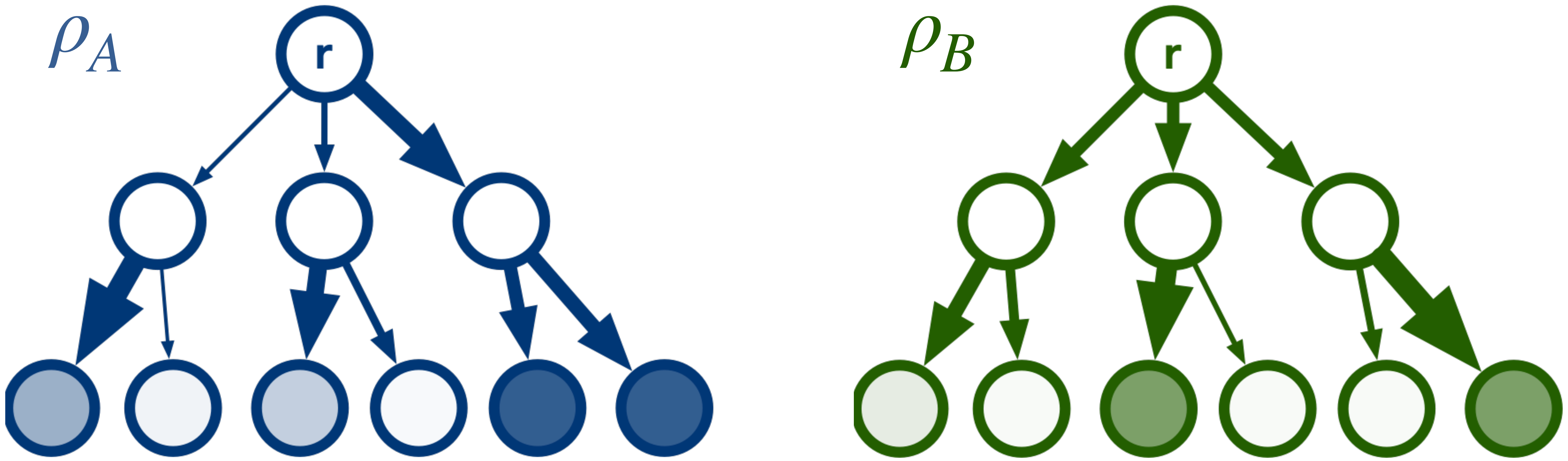
[CCHL21] Chen, Cotler, Huang, Li. Exponential separations between learning with and without quantum memory, FOCS 2021.
[HKP21] Huang, Kueng, Preskill. Information-theoretic bounds on quantum advantage in machine learning, PRL 2021.



Proof Sketch: Tree representation



- Consider the lower bound $\Omega(2^n)$ for classical agents; See [HKP21] for upper bound $\mathcal{O}(n)$.
- We consider a graphical representation for the memory state of the classical agent.

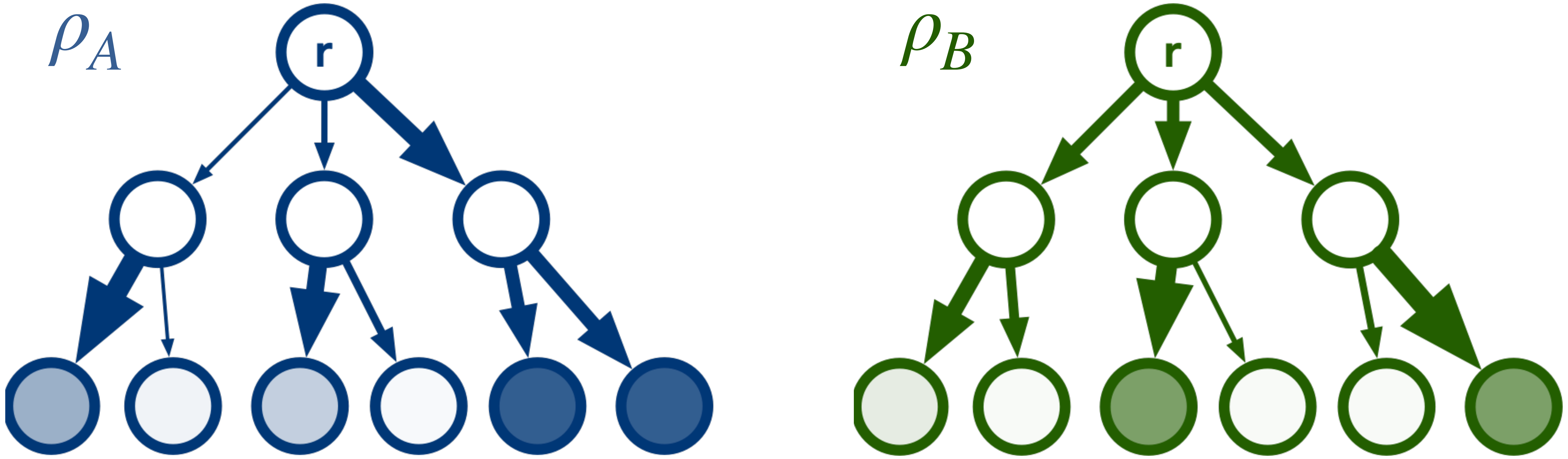


Probability distribution (bottom layer) sufficiently different $\equiv$ Classical agent can distinguish ρ_A and ρ_B

[CCHL21] Chen, Cotler, Huang, Li. Exponential separations between learning with and without quantum memory, FOCS 2021.
[HKP21] Huang, Kueng, Preskill. Information-theoretic bounds on quantum advantage in machine learning, PRL 2021.

Proof Sketch: Tree representation

- Consider the lower bound $\Omega(2^n)$ for classical agents; See [HKP21] for upper bound $\mathcal{O}(n)$.
- We consider a graphical representation for the memory state of the classical agent.

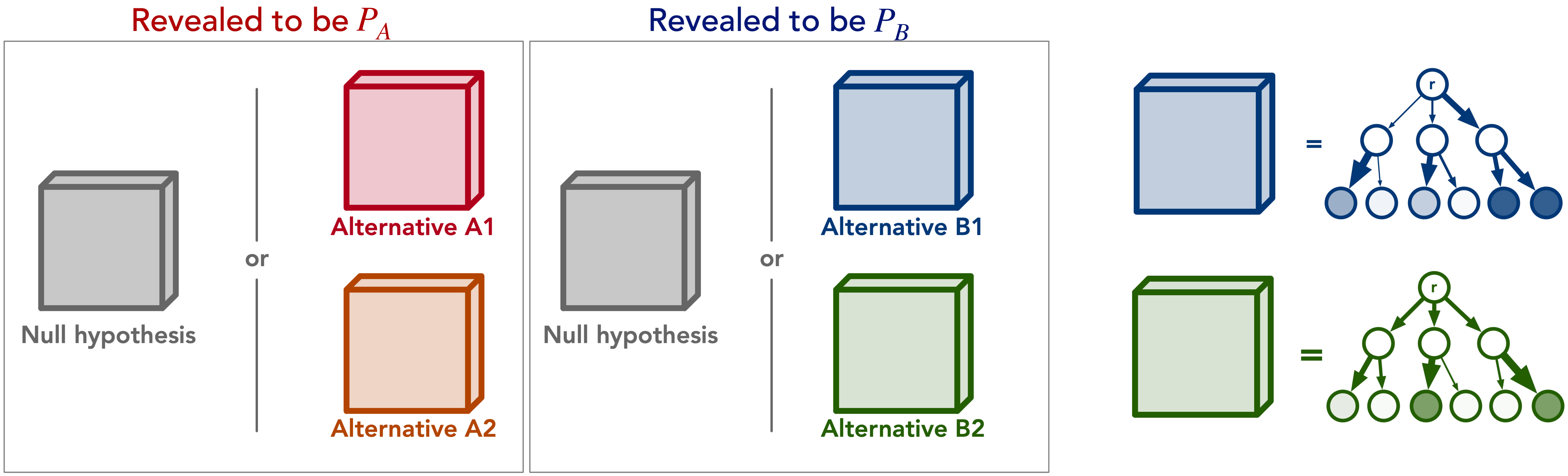


Probability distribution (bottom layer) sufficiently different $\equiv$ Classical agent can distinguish ρ_A and ρ_B

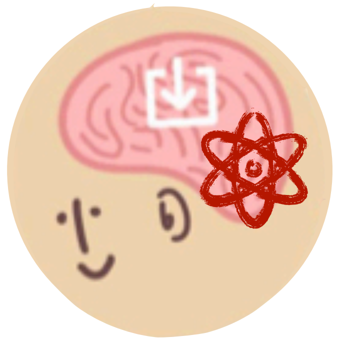
More experiments done $\equiv$ Deeper the tree $\equiv$ More distinct the distribution

Many-vs-one distinguishing task

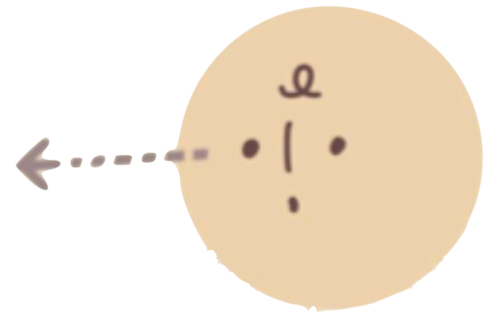
- If the classical agent succeeds in the prediction task, then he/she must succeeds in a corresponding distinguishing task.



Partially-revealed many-versus-one distinguishing task



Quantum advantage in predicting properties



- The classical/quantum agent learns a classical model of the n-qubit state ρ.
- Subsequently, one can use the classical model to predict $|\mathrm{Tr}(P\rho)|$ for an observable P chosen from $\{I, X, Y, Z\}^{\otimes n}$.

Theorem

Classical agent needs $\Omega(2^n)$ experiments to predict observable from the set, but quantum agent only need $\mathcal{O}(n)$ experiments to predict all 4^n observables.

Other examples

Current proof techniques vary rather substantially for different tasks.

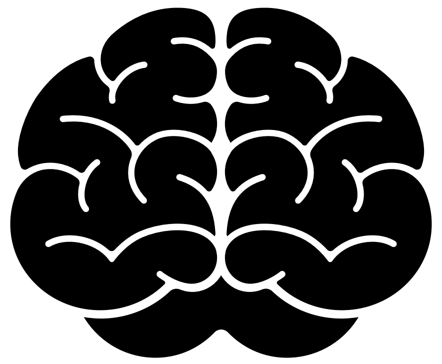
- Estimating $\mathrm{Tr}(\rho\sigma)$ [1]
- Quantum state tomography [2]
- Classifying symmetry [3]
- Quantum state certification [4]
- Quantum PCA [5]
- Learning Pauli channel [6]

References:

[1] Anshu, Anurag, Zeph Landau, and Yunchao Liu. "Distributed quantum inner product estimation." arXiv preprint arXiv:2111.03273 (2021).
[2] Chen, Sitan, et al. "Tight Bounds for State Tomography with Incoherent Measurements." arXiv preprint *arXiv:2206.05265* (2022)
[3] Chen, Sitan, et al. "Exponential separations between learning with and without quantum memory." *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2022.
[4] Chen, Sitan, et al. "Tight Bounds for Quantum State Certification with Incoherent Measurements." *arXiv preprint arXiv:2204.07155* (2022).
[5] Huang, Hsin-Yuan et al. "Quantum advantage in learning from experiments." *Science (New York, N.Y.)* vol. 376,6598 (2022): 1182-1186.
[6] Chen, Senrui, et al. "Quantum advantages for pauli channel estimation." *Physical Review A* 105.3 (2022): 032435.

Outline

- Basic setting and examples
- Key ideas:
  Part I — Designing good learning algorithms
  Part II — Proving no good learning algorithms exist
- Outlook and open questions



Outline

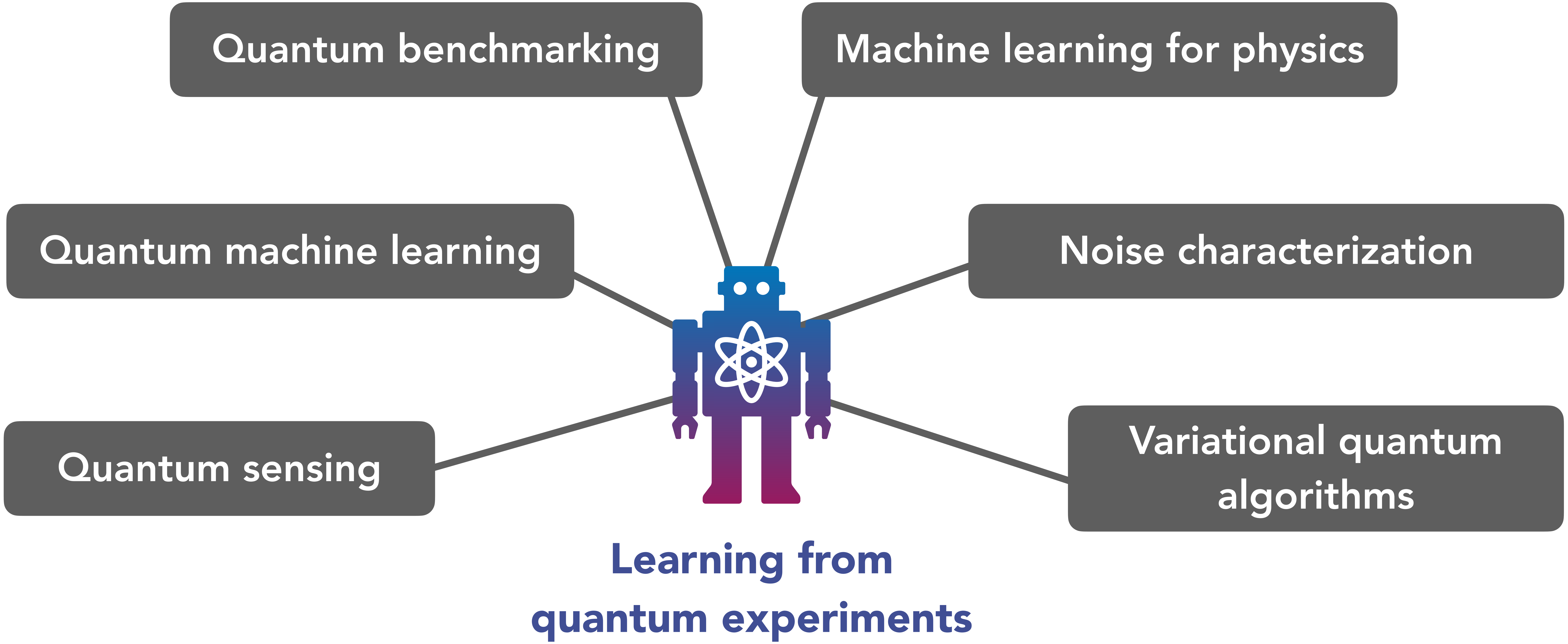
- Basic setting and examples

- Key ideas:
  Part I — Designing good learning algorithms
  Part II — Proving no good learning algorithms exist

- Outlook and open questions

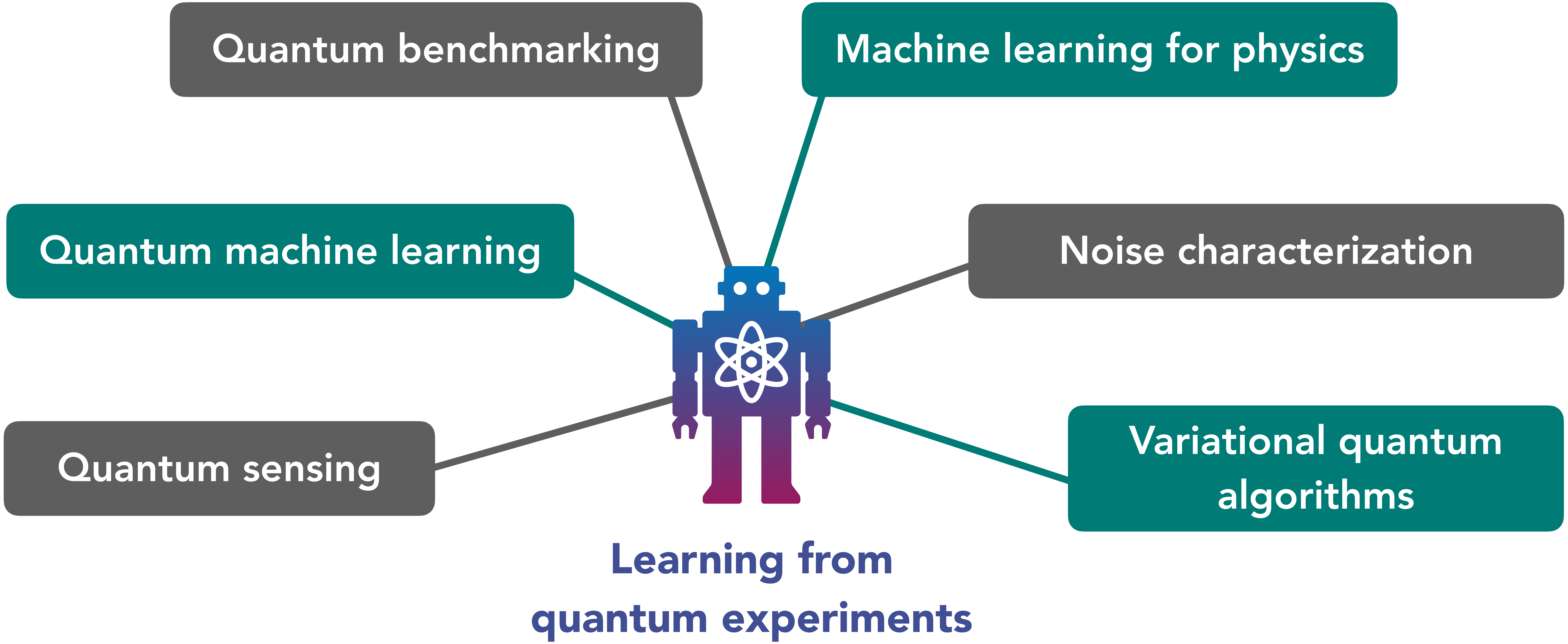
Outlook

We know a little bit about what we can learn efficiently.



Outlook

A lot of problems in these fields are yet to be studied rigorously.



Open questions

- Can we perform shadow tomography on broader classes of observables computationally efficiently?
  We only know how to do it efficiently for low-weight and Pauli observables.

- What classes of quantum dynamics/circuits are efficiently learnable?
  Many natural classes are either not efficiently learnable quantumly or efficiently learnable classically (e.g., local Hamiltonian evolution). Is there something in between?

- Could we efficiently learn if there are non-local quantum noise?
  Fault-tolerant quantum computers require local noise. Can we experimentally test this?