

A general circuit framework for consistent logical reasoning in Wigner's friend scenarios

V. Vilasini

ETH zürich

Joint work with Mischa Prebin Woods (Inria, Grenoble)



Based on: Vilasini and Woods 2022, [arXiv:2209.09281](https://arxiv.org/abs/2209.09281)

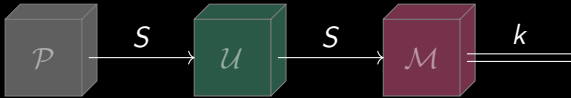
Motivation

Standard use of quantum theory



$$P(k|\mathcal{P}, \mathcal{U}, \mathcal{M})$$

Standard use of quantum theory



$$P(k|\mathcal{P}, \mathcal{U}, \mathcal{M})$$

- All agents outside the purview of theory

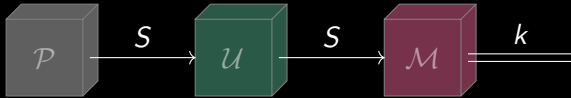
Standard use of quantum theory



$$P(k|\mathcal{P}, \mathcal{U}, \mathcal{M})$$

- All agents outside the purview of theory
- Agents not part of the boxes/wires in the circuit

Standard use of quantum theory



$$P(k|\mathcal{P}, \mathcal{U}, \mathcal{M})$$

- All agents outside the purview of theory
- Agents not part of the boxes/wires in the circuit
- Objective measurement probabilities (Born rule)

Standard use of quantum theory (agents outside Heisenberg cut)

effectively classical

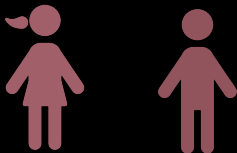


fully quantum



Standard use of quantum theory (agents outside Heisenberg cut)

effectively classical



fully quantum



Can QT be consistently extended to include agents as systems of the theory?

Universal use of quantum theory (agents within Heisenberg cut)

effectively classical



fully quantum



Wigner's Friend Scenarios (WFS): agents as fully quantum systems

Universal use of quantum theory (motile Heisenberg cut)

effectively classical



fully quantum



Wigner's Friend Scenarios (WFS): Protocols where agents can

Universal use of quantum theory (motile Heisenberg cut)

effectively classical



fully quantum



Wigner's Friend Scenarios (WFS): Protocols where agents can

- model other agents' lab as unitarily evolving closed q. systems

Universal use of quantum theory (motile Heisenberg cut)

effectively classical



fully quantum



Wigner's Friend Scenarios (WFS): Protocols where agents can

- model other agents' lab as **unitarily evolving closed q. systems**
- have **full quantum control** over lab of another agent



Frauchiger and Renner (FR): Agents modelling each other as **quantum systems** and reasoning using **classical logic** in WFS will run into logical paradoxes!



Frauchiger and Renner (FR): Agents modelling each other as **quantum systems** and reasoning using **classical logic** in WFS will run into logical paradoxes!



A consistent formalism using which quantum agents can reason logically, make and test physical predictions?



Frauchiger and Renner (FR): Agents modelling each other as quantum systems and reasoning using classical logic in WFS will run into logical paradoxes!



A consistent formalism using which quantum agents can reason logically, make and test physical predictions?



Vilasini and Woods (VW): Yes! Generalised q. circuits



Frauchiger and Renner (FR): Agents modelling each other as **quantum systems** and reasoning using **classical logic** in WFS will run into logical paradoxes!



A consistent formalism using which quantum agents can reason logically, make and test physical predictions?



Vilasini and Woods (VW): Yes! Generalised q. circuits

- Consistency w/o giving up **quantum theory** or **classical logic**



Frauchiger and Renner (FR): Agents modelling each other as **quantum systems** and reasoning using **classical logic** in WFS will run into logical paradoxes!



A consistent formalism using which quantum agents can reason logically, make and test physical predictions?



Vilasini and Woods (VW): Yes! Generalised q. circuits

- Consistency w/o giving up **quantum theory** or **classical logic**
- Operational formalisation of H-cuts



Frauchiger and Renner (FR): Agents modelling each other as **quantum systems** and reasoning using **classical logic** in WFS will run into logical paradoxes!



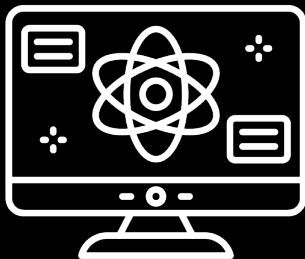
A consistent formalism using which quantum agents can reason logically, make and test physical predictions?



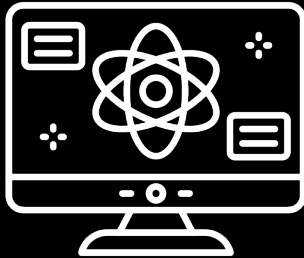
Vilasini and Woods (VW): Yes! Generalised q. circuits

- Consistency w/o giving up **quantum theory** or **classical logic**
- Operational formalisation of H-cuts
- Precise necessary condition for FR-type paradoxes

Agents \neq conscious human beings. Can be quantum computers!



Agents \neq conscious human beings. Can be quantum computers!



Necessary conditions for agency:

- Can measure another quantum system and store outcome
- Can compute probabilities, perform basic logical deductions

The FR no-go theorem and apparent paradox

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

($K_A(a = i)$: Agent A knows with certainty that $a = i$)

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

($K_A(a = i)$: Agent A knows with certainty that $a = i$)

C $K_B K_A(a = i) \Rightarrow K_B(a = i)$

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

($K_A(a = i)$: Agent A knows with certainty that $a = i$)

C $K_B K_A(a = i) \Rightarrow K_B(a = i)$

S $K_A(a = i) \Rightarrow \neg K_A(a = i')$, $\forall i \neq i'$

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

($K_A(a = i)$: Agent A knows with certainty that $a = i$)

C $K_B K_A(a = i) \Rightarrow K_B(a = i)$

S $K_A(a = i) \Rightarrow \neg K_A(a = i')$, $\forall i \neq i'$

U Agents can model others' labs **unitarily**, have **full q. control**

Q If Born rule assigns probability 1 to a **measurement outcome**, an agent can be certain of that **outcome**.

($K_A(a = i)$: Agent A knows with certainty that $a = i$)

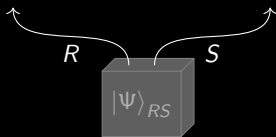
C $K_B K_A(a = i) \Rightarrow K_B(a = i)$

S $K_A(a = i) \Rightarrow \neg K_A(a = i')$, $\forall i \neq i'$

U Agents can model others' labs **unitarily**, have **full q. control**

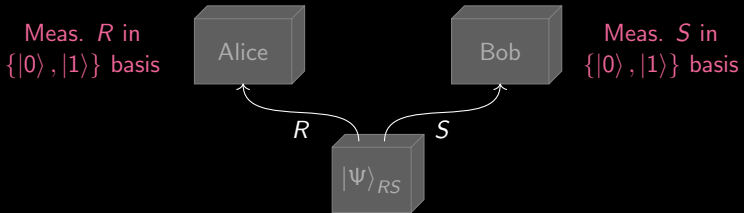
Theorem (FR): \exists a protocol where agents reasoning using Q , U , C and S will arrive at logically contradictory predictions.

The FR protocol (entanglement version)



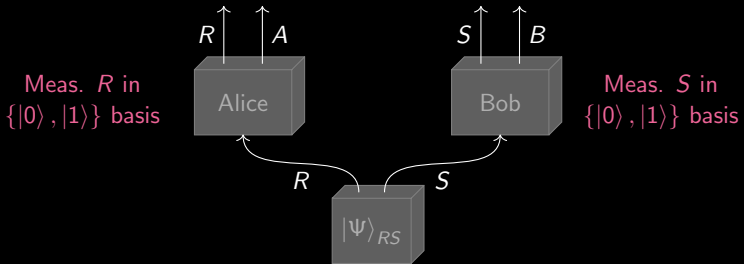
$$|\Psi\rangle_{RS} = \frac{1}{\sqrt{3}} (|00\rangle + |10\rangle + |11\rangle)_{RS}$$

The FR protocol (entanglement version)



$$|\Psi\rangle_{RS} = \frac{1}{\sqrt{3}}(|00\rangle + |10\rangle + |11\rangle)_{RS}$$

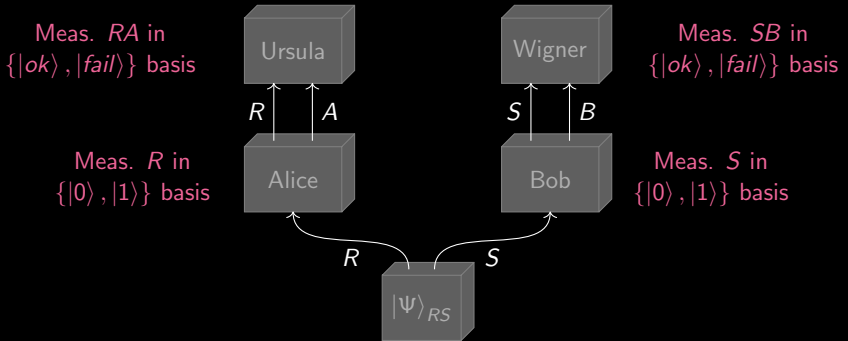
The FR protocol (entanglement version)



$$|\Psi\rangle_{RS} = \frac{1}{\sqrt{3}}(|00\rangle + |10\rangle + |11\rangle)_{RS}$$

RA : Alice's lab, SB : Bob's lab

The FR protocol (entanglement version)

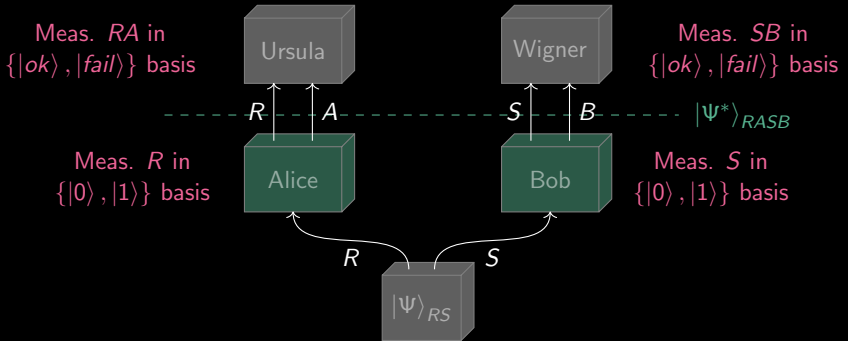


$$|\Psi\rangle_{RS} = \frac{1}{\sqrt{3}}(|00\rangle + |10\rangle + |11\rangle)_{RS}$$

RA : Alice's lab, SB : Bob's lab

$$|ok\rangle := \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad |fail\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Applying U to Alice and Bob's labs

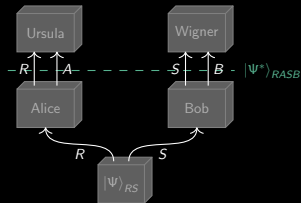


$$|\Psi\rangle_{RS} = \frac{1}{\sqrt{3}} (|00\rangle + |10\rangle + |11\rangle)_{RS}$$

RA : Alice's lab, SB : Bob's lab

$$|\Psi^*\rangle_{RASB} = \frac{1}{\sqrt{3}} (|0000\rangle + |1100\rangle + |1111\rangle)_{RASB}$$

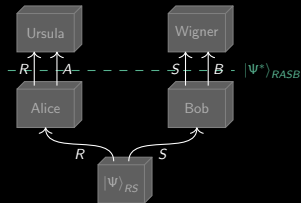
FR's steps in proof of theorem (reasoning using Q , U and C)



(1) Ursula reasons about Bob:
 $\langle ok|_{RA} \otimes \langle 00|_{SB} \cdot |\Psi^*\rangle_{RASB} = 0$
 she concludes $P(b = 1|u = ok) = 1$,
 or $K_U(u = ok \Rightarrow b = 1)$.

$$|\Psi^*\rangle_{RASB} = \frac{1}{\sqrt{3}} (|0000\rangle + |1100\rangle + |1111\rangle)_{RASB}$$

FR's steps in proof of theorem (reasoning using Q , U and C)

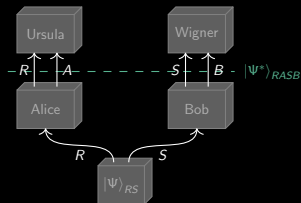


(1) Ursula reasons about Bob:
 $\langle ok|_{RA} \otimes \langle 00|_{SB} \cdot |\Psi^*\rangle_{RASB} = 0$
 she concludes $P(b = 1|u = ok) = 1$,
 or $K_U(u = ok \Rightarrow b = 1)$.

$$|\Psi^*\rangle_{RASB} = \frac{1}{\sqrt{3}}(|0000\rangle + |1100\rangle + |1111\rangle)_{RASB}$$

(2) Ursula reasons about Bob's reasoning about Alice
 $K_U K_B(b = 1 \Rightarrow a = 1)$

FR's steps in proof of theorem (reasoning using Q , U and C)



$$|\Psi^*\rangle_{RASB} = \frac{1}{\sqrt{3}}(|0000\rangle + |1100\rangle + |1111\rangle)_{RASB}$$

(1) Ursula reasons about Bob:
 $\langle ok|_{RA} \otimes \langle 00|_{SB} \cdot |\Psi^*\rangle_{RASB} = 0$
 she concludes $P(b = 1|u = ok) = 1$,
 or $K_U(u = ok \Rightarrow b = 1)$.

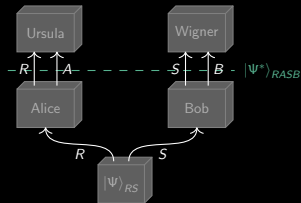
(2) Ursula reasons about Bob's reasoning about Alice

$$K_U K_B(b = 1 \Rightarrow a = 1)$$

(3) Ursula reasons about Bob's reasoning about Alice's reasoning
 about Wigner

$$K_U K_B K_A(a = 1 \Rightarrow w = fail)$$

FR's steps in proof of theorem (reasoning using Q , U and C)



$$|\Psi^*\rangle_{RASB} = \frac{1}{\sqrt{3}} (|0000\rangle + |1100\rangle + |1111\rangle)_{RASB}$$

(1) Ursula reasons about Bob:
 $\langle ok|_{RA} \otimes \langle 00|_{SB} \cdot |\Psi^*\rangle_{RASB} = 0$
 she concludes $P(b = 1 | u = ok) = 1$,
 or $K_U(u = ok \Rightarrow b = 1)$.

(2) Ursula reasons about Bob's reasoning about Alice

$$K_U K_B (b = 1 \Rightarrow a = 1)$$

(3) Ursula reasons about Bob's reasoning about Alice's reasoning about Wigner

$$K_U K_B K_A (a = 1 \Rightarrow w = fail)$$

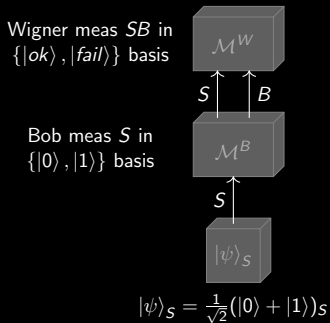
Violation of S in any round where $u = w = ok$ (non-zero prob)

$$K_U (w = fail) \text{ and } K_U (w = ok)$$

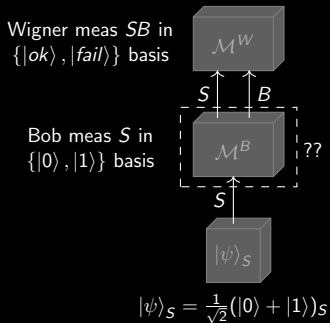
PARADOX!

A simple resolution to FR paradoxes

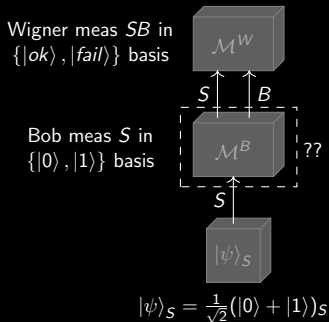
Wigner's original expt: unitarity vs projection postulate



Wigner's original expt: unitarity vs projection postulate



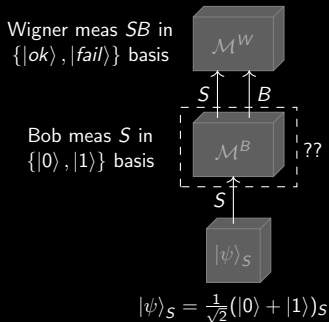
Wigner's original expt: unitarity vs projection postulate



Unitary $\mathcal{M}_{unitary}^B$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

Wigner's original expt: unitarity vs projection postulate



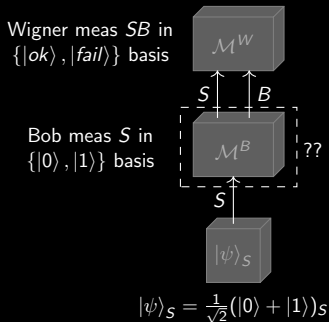
Unitary $\mathcal{M}_{unitary}^B$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

Projection $\mathcal{M}_{projection}^B$

$$|\psi\rangle_S \mapsto |00\rangle \text{ or } |11\rangle \text{ or mixture}$$

Wigner's original expt: unitarity vs projection postulate



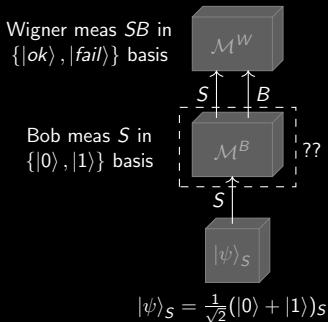
Unitary $\mathcal{M}_{unitary}^B : x_B = 0$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

Projection $\mathcal{M}_{projection}^B : x_B = 1$

$$|\psi\rangle_S \mapsto |00\rangle \text{ or } |11\rangle \text{ or mixture}$$

Wigner's original expt: unitarity vs projection postulate



Unitary $\mathcal{M}_{unitary}^B : x_B = 0$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

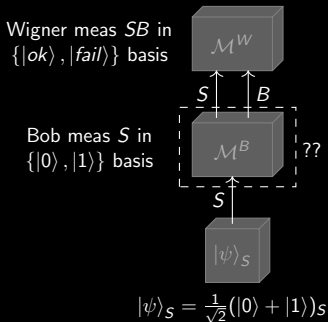
$$\Rightarrow P(w = ok | x_B = 0) = 0$$

Projection $\mathcal{M}_{projection}^B : x_B = 1$

$$|\psi\rangle_S \mapsto |00\rangle \text{ or } |11\rangle \text{ or mixture}$$

$$\Rightarrow P(w = ok | x_B = 1) > 0$$

Wigner's original expt: unitarity vs projection postulate



Unitary $\mathcal{M}_{unitary}^B : x_B = 0$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

$$\Rightarrow P(w = ok | x_B = 0) = 0$$

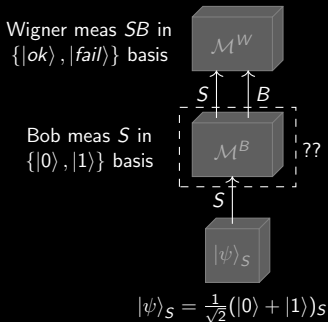
Projection $\mathcal{M}_{projection}^B : x_B = 1$

$$|\psi\rangle_S \mapsto |00\rangle \text{ or } |11\rangle \text{ or mixture}$$

$$\Rightarrow P(w = ok | x_B = 1) > 0$$

- Predictions do depend on how mmt is modelled ($x_B \in \{0, 1\}$)!

Wigner's original expt: unitarity vs projection postulate



Unitary $\mathcal{M}_{unitary}^B : x_B = 0$

$$|\psi\rangle_S \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB} = |fail\rangle_{SB}$$

$$\Rightarrow P(w = ok | x_B = 0) = 0$$

Projection $\mathcal{M}_{projection}^B : x_B = 1$

$$|\psi\rangle_S \mapsto |00\rangle \text{ or } |11\rangle \text{ or mixture}$$

$$\Rightarrow P(w = ok | x_B = 1) > 0$$

- Predictions do depend on how mmt is modelled ($x_B \in \{0, 1\}$)!
- Formalises the H-cut: agents' "memory" as q. system vs storing effectively classical values

FR paradox disappears once “settings” /H-cuts are accounted for

FR's statements

- $u = ok \Rightarrow b = 1$ “ $P(b = 1|u = ok) = 1$ ”
- $b = 1 \Rightarrow a = 1$ “ $P(a = 1|b = 1) = 1$ ”
- $a = 1 \Rightarrow w = fail$ “ $P(w = fail|a = 1) = 1$ ”
- $P(u = w = ok) = \frac{1}{12} > 0$

Explicit statements in our framework

- $u = ok \wedge (x_A = 0, x_B = 1) \Rightarrow b = 1$
- $b = 1 \wedge (x_A = 1, x_B = 1) \Rightarrow a = 1$
- $a = 1 \wedge (x_A = 1, x_B = 0) \Rightarrow w = fail$
- $P(u = w = ok|(x_A = 0, x_B = 0)) = \frac{1}{12} > 0$

Cannot be chained together by any axiom of classical logic

Necessary condition for FR-type apparent paradoxes

- | outcome probabilities of one mmt are independent of another mmt's setting $x \in \{0, 1\}$ (unitary vs projection)

Necessary condition for FR-type apparent paradoxes

I outcome probabilities of one mmt are independent of another mmt's setting $x \in \{0, 1\}$ (unitary vs projection)

Q, U, C, S formalise FR's assumptions Q, U, C, S , but taking into account settings $x \in \{0, 1\}$.

Necessary condition for FR-type apparent paradoxes

I outcome probabilities of one mmt are independent of another mmt's setting $x \in \{0, 1\}$ (unitary vs projection)

Q, U, C, S formalise FR's assumptions Q, U, C, S , but taking into account settings $x \in \{0, 1\}$.

Theorem:

1. Q, U, C and S are always consistent in any WFS.

Necessary condition for FR-type apparent paradoxes

I outcome probabilities of one mmt are independent of another mmt's setting $x \in \{0, 1\}$ (unitary vs projection)

Q, U, C, S formalise FR's assumptions Q, U, C, S, but taking into account settings $x \in \{0, 1\}$.

Theorem:

1. Q, U, C and S are always consistent in any WFS.
2. Any logical paradox can only arise if agents reason using Q, U, C, S and I in a WFS where I actually fails.

Necessary condition for FR-type apparent paradoxes

I outcome probabilities of one mmt are independent of another mmt's setting $x \in \{0, 1\}$ (unitary vs projection)

Q, U, C, S formalise FR's assumptions Q, U, C, S, but taking into account settings $x \in \{0, 1\}$.

Theorem:

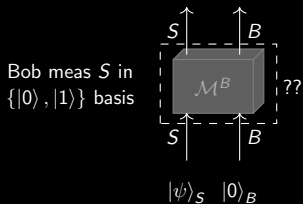
1. Q, U, C and S are always consistent in any WFS.
2. Any logical paradox can only arise if agents reason using Q, U, C, S and I in a WFS where I actually fails.

Universality of QT does not threaten logic!



Quantum circuit framework for WFS

What is the channel \mathcal{M}_B ?

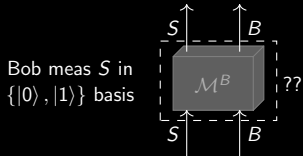


$$|\psi\rangle_S = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_S$$

Unitary: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB}$

Projection: $|00\rangle, |11\rangle$, mixture

What is the channel \mathcal{M}_B ?



$$|\psi\rangle_S |0\rangle_B$$

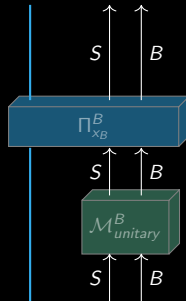
$$|\psi\rangle_S = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_S$$

Unitary: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB}$

Projection: $|00\rangle, |11\rangle$, mixture

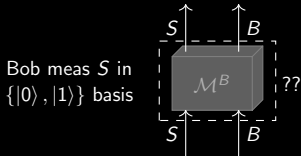
Explicit description of \mathcal{M}_B

$$b \in \{\perp, 0, 1\}$$



$$x_B \in \{0, 1\} \quad |\psi\rangle_S |0\rangle_B$$

What is the channel \mathcal{M}_B ?



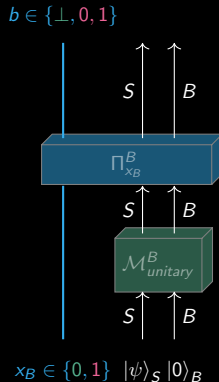
$$|\psi\rangle_S |0\rangle_B$$

$$|\psi\rangle_S = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_S$$

Unitary: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{SB}$

Projection: $|00\rangle, |11\rangle$, mixture

Explicit description of \mathcal{M}_B

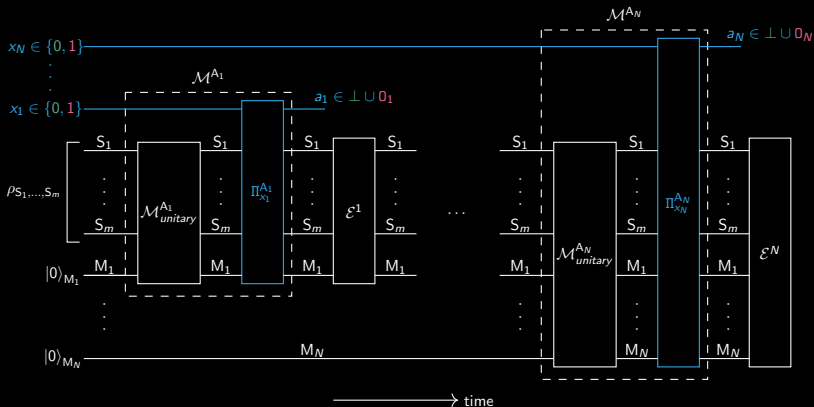


Setting-dependent projectors

$$\Pi_{x_B=0}^B = \mathcal{I}_{SB} \text{ (trivial outcome } b = \perp)$$

$$\Pi_{x_B=1}^B = \{|00\rangle\langle 00|_{SB}, |11\rangle\langle 11|_{SB}\} \text{ (non-trivial } b \in \{0, 1\})$$

Generalises to arbitrary WFS over N agents A_1, \dots, A_N , performing arbitrary quantum operations on each other's labs/memories



An augmented circuit for a general WFS

Completeness, logical and causal consistency

Completeness, logical and causal consistency

Theorem (informal): An augmented circuit for a WFS

- (1) Encodes all predictions that can be made in that WFS
- (2) Never leads to contradictory predictions
- (3) Predictions only depend on settings in past light cone

Completeness, logical and causal consistency

Theorem (informal): An augmented circuit for a WFS

- (1) Encodes all predictions that can be made in that WFS
- (2) Never leads to contradictory predictions
- (3) Predictions only depend on settings in past light cone

Further...

- Unifying framework for previous responses
- Reasoning rules for quantum agents
- Allows subjective H-cuts, non-absolute measurement events

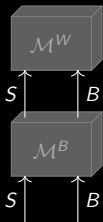
Emergence of objective measurement events

❓ How do objective predictions emerge in standard quantum expts when there can be subjectivity in WFS?

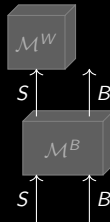
? How do objective predictions emerge in standard quantum expts when there can be subjectivity in WFS?

(Causal) structural distinction

WFS (quantum control of full lab)



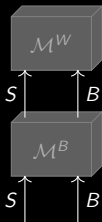
Standard quantum expt



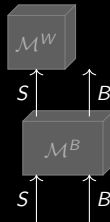
? How do objective predictions emerge in standard quantum expts when there can be subjectivity in WFS?

(Causal) structural distinction

WFS (quantum control of full lab)



Standard quantum expt

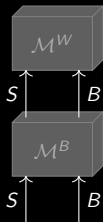


- Theorem: Predictions in standard q. expts are setting-indep

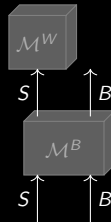
? How do objective predictions emerge in standard quantum expts when there can be subjectivity in WFS?

(Causal) structural distinction

WFS (quantum control of full lab)



Standard quantum expt



- Theorem: Predictions in standard q. expts are setting-indep
- Objective probabilities, independent of H-cuts emerge

Summary and conclusions

FR: Logical paradox in a particular WF scenario with q . agents

FR: Logical paradox in a particular WF scenario with q . agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q . agents

FR: Logical paradox in a particular WF scenario with q . agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q . agents

- Resolution: account for H-cut (unitary vs projection channel)

FR: Logical paradox in a particular WF scenario with q . agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q . agents

- Resolution: account for H-cut (unitary vs projection channel)
- Absolute notion of mmt events not necessary for sound logic

FR: Logical paradox in a particular WF scenario with q. agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q. agents

- Resolution: account for H-cut (unitary vs projection channel)
- Absolute notion of mmt events not necessary for sound logic
- Objectivity/absoluteness emerges in standard q. expts

FR: Logical paradox in a particular WF scenario with q. agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q. agents

- Resolution: account for H-cut (unitary vs projection channel)
- Absolute notion of mmt events not necessary for sound logic
- Objectivity/absoluteness emerges in standard q. expts
- Framework interpretation independent (just like q. circuits)

FR: Logical paradox in a particular WF scenario with q. agents

Our work: Quantum circuit framework that ensures logical and causal consistency in any WF scenario with q. agents

- Resolution: account for H-cut (unitary vs projection channel)
- Absolute notion of mmt events not necessary for sound logic
- Objectivity/absoluteness emerges in standard q. expts
- Framework interpretation independent (just like q. circuits)

Take home: Quantum agents in WFS can always reason consistently without giving up quantum theory or classical logic, as long as they don't ignore non-trivial dependences on H-cuts.

Outlook

- WFS and measurement problem beyond quantum theory
(Ormrod, Vilasini, Barrett 2023: QPL Talk on Wednesday)

- WFS and measurement problem beyond quantum theory
(Ormrod, Vilasini, Barrett 2023: QPL Talk on Wednesday)
- Relation to logical contextuality, semantic paradoxes
(Nurgalieva, Vilasini 2023: QPL Talk on Thursday)

- WFS and measurement problem beyond quantum theory
(Ormrod, Vilasini, Barrett 2023: QPL Talk on Wednesday)
- Relation to logical contextuality, semantic paradoxes
(Nurgalieva, Vilasini 2023: QPL Talk on Thursday)
- Quantum causal models and relativistic causality in WFS
(Ongoing work)

- WFS and measurement problem beyond quantum theory
(Ormrod, Vilasini, Barrett 2023: QPL Talk on Wednesday)
- Relation to logical contextuality, semantic paradoxes
(Nurgalieva, Vilasini 2023: QPL Talk on Thursday)
- Quantum causal models and relativistic causality in WFS
(Ongoing work)
- Broader research program on WFS theory and experiments,
other no-go theorems

- WFS and measurement problem beyond quantum theory
(Ormrod, Vilasini, Barrett 2023: QPL Talk on Wednesday)
- Relation to logical contextuality, semantic paradoxes
(Nurgalieva, Vilasini 2023: QPL Talk on Thursday)
- Quantum causal models and relativistic causality in WFS
(Ongoing work)
- Broader research program on WFS theory and experiments,
other no-go theorems

THANK YOU!
