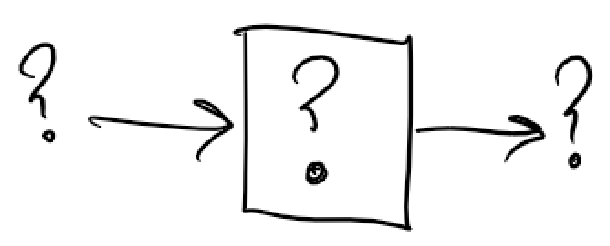
Data-Driven Inference and Observationally Complete Devices

joint work with: M. Dall'Arno, A. Bisio, A. Tosini

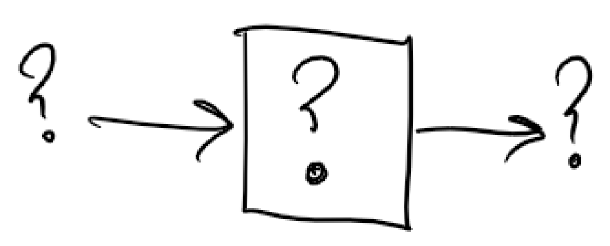
Francesco Buscemi (Nagoya University)

51st Symposium on Mathematical Physics
Toruń, Poland, 16 June 2019

An unknown device:



An unknown device:



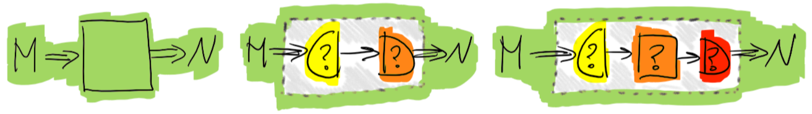
how can we infer anything about it?

The Starting Point

- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment

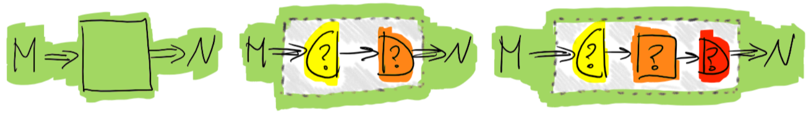
The Starting Point

- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment
- given is also a **hypothesis** (prior information) about the structure of the circuit that generated the data:



The Starting Point

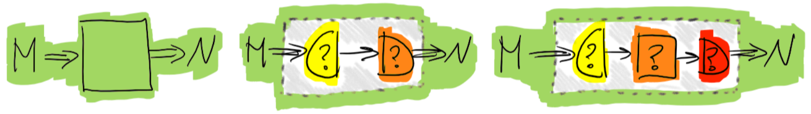
- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment
- given is also a **hypothesis** (prior information) about the structure of the circuit that generated the data:



- **Aim:** to construct an inference, consistent with the hypothesis, about the pieces composing the circuit that generated the dataset

The Starting Point

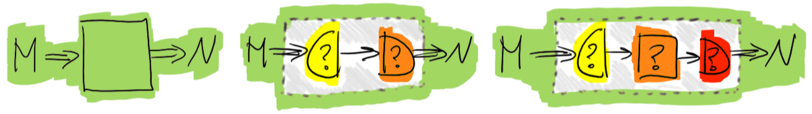
- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment
- given is also a **hypothesis** (prior information) about the structure of the circuit that generated the data:



- **Aim:** to construct an inference, consistent with the hypothesis, about the pieces composing the circuit that generated the dataset
- in the negative: if the dataset is incompatible with the hypothesis, the hypothesis is falsified (like in a Bell test)

The Starting Point

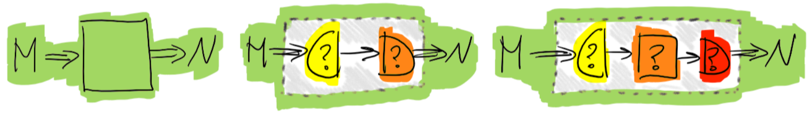
- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment
- given is also a **hypothesis** (prior information) about the structure of the circuit that generated the data:



- **Aim:** to construct an inference, consistent with the hypothesis, about the pieces composing the circuit that generated the dataset
- in the negative: if the dataset is incompatible with the hypothesis, the hypothesis is falsified (like in a Bell test)
- in the positive: the hypothesis is "corroborated," but also some information about the device can be inferred (given an inference rule)

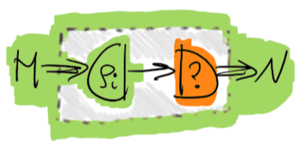
The Starting Point

- given is a **set of data** in the form $p(j|i)$, where $i \in [1, M]$ labels the setups (input) and $j \in [1, N]$ the outcomes (output) of an experiment
- given is also a **hypothesis** (prior information) about the structure of the circuit that generated the data:



- **Aim:** to construct an inference, consistent with the hypothesis, about the pieces composing the circuit that generated the dataset
- in the negative: if the dataset is incompatible with the hypothesis, the hypothesis is falsified (like in a Bell test)
- in the positive: the hypothesis is "corroborated," but also some information about the device can be inferred (given an inference rule)
- case-study in this talk: measurement inference

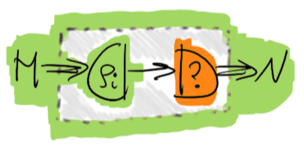
Conventional tomography



- probe: input states
- inference target: measurement
- probe states known

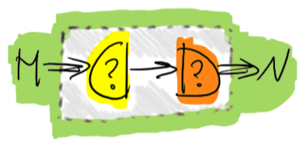
Tomography VS Data-Driven Inference

Conventional tomography



- probe: input states
- inference target: measurement
- probe states known

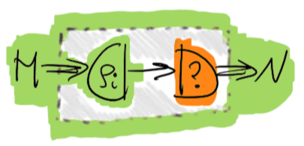
Data-driven inference (this talk)



- probe: input states
- inference target: measurement
- probe states unknown

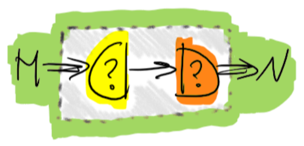
Tomography VS Data-Driven Inference

Conventional tomography



- probe: input states
- inference target: measurement
- probe states known

Data-driven inference (this talk)



- probe: input states
- inference target: measurement
- probe states unknown

Motivation: to break (or at least to loosen) the circular argument on which conventional tomography relies

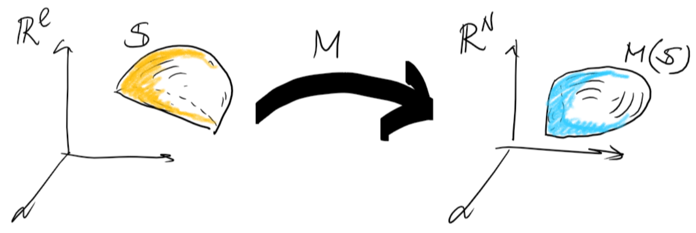
Wigner's Other Chain

As Wigner put it:

> *[...] the experimentalist uses certain apparatus to measure the position, let us say, or the momentum, or the angular momentum. Now, how does the experimentalist know that this apparatus will measure for him the position? "Oh," you say, "he observed the apparatus. He looked at it." Well that means that he carried out a measurement on it. How did he know that the apparatus with which he carried out that measurement will tell him the properties of the apparatus? Fundamentally, this is again a chain which has no beginning.* **And at the end we have to say, "We learned that as children how to judge what is around us.**"

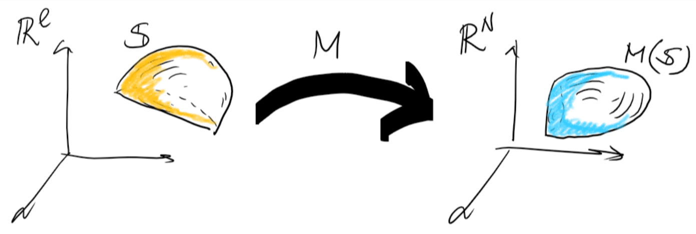
[E.P. Wigner, Lecture at the Conference on the Foundations of Quantum Mechanics, Xavier University, Cincinnati, 1962.]

Measurement Representation



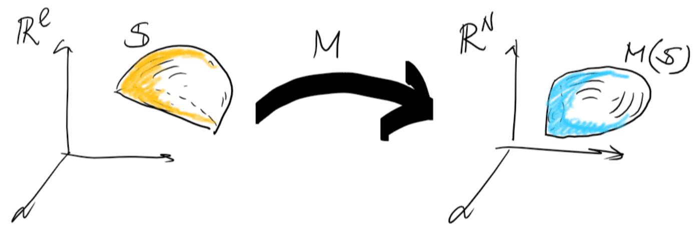
- measurement: linear mapping M from state set $\mathbb{S} \subset \mathbb{R}^\ell$ to probability distributions in $\mathbb{R}^N$

Measurement Representation



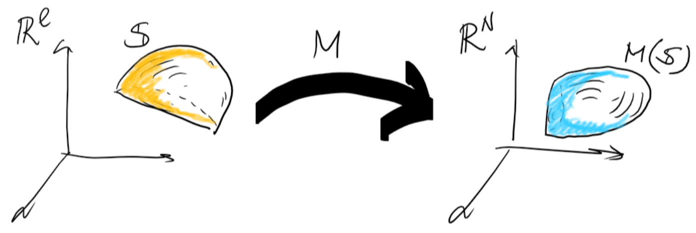
- measurement: linear mapping M from state set $\mathbb{S} \subset \mathbb{R}^\ell$ to probability distributions in $\mathbb{R}^N$
- assumption in this talk: measurements are informationally complete (otherwise conditions become more technical)

Measurement Representation



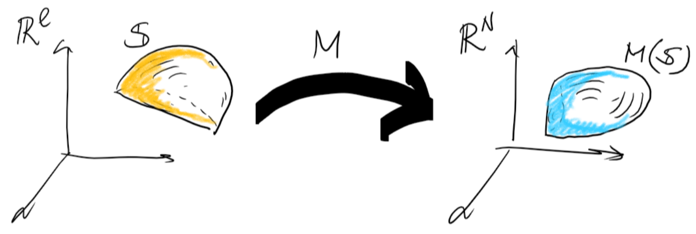
- measurement: linear mapping M from state set $\mathbb{S} \subset \mathbb{R}^\ell$ to probability distributions in $\mathbb{R}^N$
- assumption in this talk: measurements are informationally complete (otherwise conditions become more technical)
- measurement range: $M(\mathbb{S}) \triangleq \{\boldsymbol{p} \in \mathbb{R}^N : \boldsymbol{p} = M(\rho), \rho \in \mathbb{S}\}$

Measurement Representation



- measurement: linear mapping M from state set $\mathbb{S} \subset \mathbb{R}^\ell$ to probability distributions in $\mathbb{R}^N$
- assumption in this talk: measurements are informationally complete (otherwise conditions become more technical)
- measurement range: $M(\mathbb{S}) \triangleq \{\boldsymbol{p} \in \mathbb{R}^N : \boldsymbol{p} = M(\rho), \rho \in \mathbb{S}\}$
- gauge symmetry: any transformation U such that $U(\mathbb{S}) = \mathbb{S}$

Measurement Representation



- measurement: linear mapping M from state set $\mathbb{S} \subset \mathbb{R}^\ell$ to probability distributions in $\mathbb{R}^N$
- assumption in this talk: measurements are informationally complete (otherwise conditions become more technical)
- measurement range: $M(\mathbb{S}) \triangleq \{\boldsymbol{p} \in \mathbb{R}^N : \boldsymbol{p} = M(\rho), \rho \in \mathbb{S}\}$
- gauge symmetry: any transformation U such that $U(\mathbb{S}) = \mathbb{S}$
- **Theorem:** the range $M(\mathbb{S})$ identifies M up to gauge symmetries

Quiz

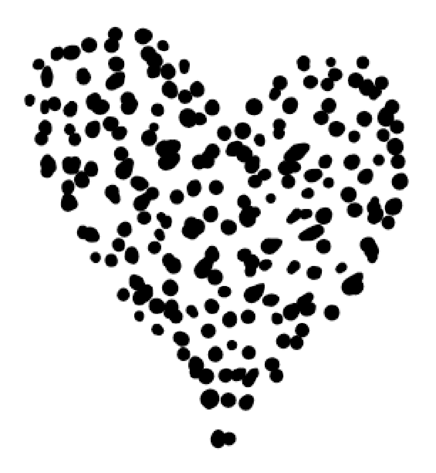
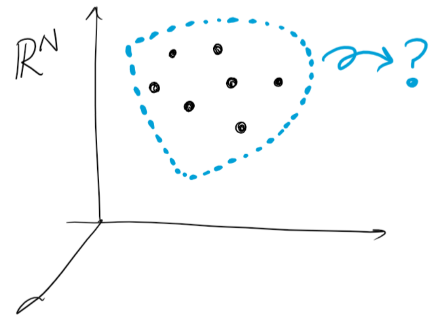


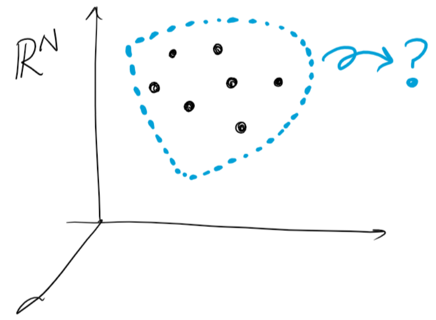
Figure 1: What do you see?

Inferring a Range from the Dataset



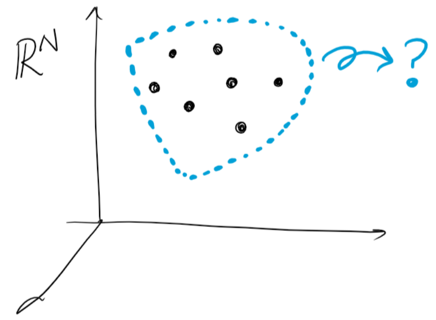
- hypothesis: let us assume a theory $(\mathbb{S}, \mathbb{E})$

Inferring a Range from the Dataset



- hypothesis: let us assume a theory $(\mathbb{S}, \mathbb{E})$
- this tells us how measurement ranges look like

Inferring a Range from the Dataset



- hypothesis: let us assume a theory $(\mathbb{S}, \mathbb{E})$
- this tells us how measurement ranges look like
- **Data-Driven Inference (DDI) Rule:** in the face of data $D = \{\boldsymbol{p}_x \in \mathbb{R}^N\}$, infer the range which:
  1. contains the convex hull of D and
  2. is of minimum euclidean volume

Some Comments

- "minimum volume" in the affine variety spanned by D

Some Comments

- "minimum volume" in the affine variety spanned by D
  - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)

Some Comments

- "minimum volume" in the affine variety spanned by D
    - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)
    - why minimum? because we want to infer "as little as possible" in the face of the data, that is, the least committal inference consistent with the data

Some Comments

- "minimum volume" in the affine variety spanned by D
  - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)
  - why minimum? because we want to infer "as little as possible" in the face of the data, that is, the least committal inference consistent with the data
- the output of **DDI** may be not unique: the inference rule may return *a set* of compatible minimum-volume ranges

Some Comments

- "minimum volume" in the affine variety spanned by D
  - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)
  - why minimum? because we want to infer "as little as possible" in the face of the data, that is, the least committal inference consistent with the data
- the output of **DDI** may be not unique: the inference rule may return *a set* of compatible minimum-volume ranges
- **DDI** may fail: for example, if the data are incompatible with the hypothesis $(\mathbb{S}, \mathbb{E})$

Some Comments

- "minimum volume" in the affine variety spanned by D
  - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)
  - why minimum? because we want to infer "as little as possible" in the face of the data, that is, the least committal inference consistent with the data
- the output of **DDI** may be not unique: the inference rule may return *a set* of compatible minimum-volume ranges
- **DDI** may fail: for example, if the data are incompatible with the hypothesis $(\mathbb{S}, \mathbb{E})$
- **Problem 1:** in order to apply **DDI**, one first needs to know the shape of $\mathbb{S}$

Some Comments

- "minimum volume" in the affine variety spanned by D
  - why volume? because in this way the inference does not change under linear transformations (and these are all that matter for a linear theory)
  - why minimum? because we want to infer "as little as possible" in the face of the data, that is, the least committal inference consistent with the data
- the output of **DDI** may be not unique: the inference rule may return *a set* of compatible minimum-volume ranges
- **DDI** may fail: for example, if the data are incompatible with the hypothesis $(\mathbb{S}, \mathbb{E})$
- **Problem 1:** in order to apply **DDI**, one first needs to know the shape of $\mathbb{S}$
- **Problem 2:** empirical data are not probability distributions but finite-statistics frequencies

When Is the Inference Correct?

- assume that there is a "true but unknown" measurement to be inferred, and that the hypothesis about the underlying theory is "correct"

When Is the Inference Correct?

- assume that there is a “true but unknown” measurement to be inferred, and that the hypothesis about the underlying theory is “correct”
- what data are needed so that **DDI** returns the correct range?

When Is the Inference Correct?

- assume that there is a "true but unknown" measurement to be inferred, and that the hypothesis about the underlying theory is "correct"
- what data are needed so that **DDI** returns the correct range?
- denote by $\mathcal{S} \subseteq \mathbb{S}$ the set of probe states $\{\rho_i : i \in [1, M]\}$ that are used to generate the statistics (i.i.d. assumptions everywhere)

When Is the Inference Correct?

- assume that there is a “true but unknown” measurement to be inferred, and that the hypothesis about the underlying theory is “correct”
- what data are needed so that **DDI** returns the correct range?
- denote by $\mathcal{S} \subseteq \mathbb{S}$ the set of probe states $\{\rho_i : i \in [1, M]\}$ that are used to generate the statistics (i.i.d. assumptions everywhere)
- **Observational Completeness:** $\mathcal{S} \subseteq \mathbb{S}$ is observationally complete for measurement M whenever **DDI**$[M(\mathcal{S})] = M(\mathbb{S})$

When Is the Inference Correct?

- assume that there is a “true but unknown” measurement to be inferred, and that the hypothesis about the underlying theory is “correct”
- what data are needed so that **DDI** returns the correct range?
- denote by $\mathcal{S} \subseteq \mathbb{S}$ the set of probe states $\{\rho_i : i \in [1, M]\}$ that are used to generate the statistics (i.i.d. assumptions everywhere)
- **Observational Completeness:** $\mathcal{S} \subseteq \mathbb{S}$ is observationally complete for measurement M whenever **DDI**$[M(\mathcal{S})] = M(\mathbb{S})$
- the entire $\mathbb{S}$ is obviously observationally complete for any measurement

When Is the Inference Correct?

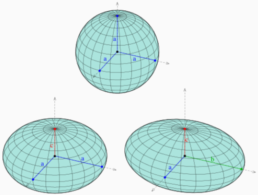
- assume that there is a "true but unknown" measurement to be inferred, and that the hypothesis about the underlying theory is "correct"
- what data are needed so that **DDI** returns the correct range?
- denote by $\mathcal{S} \subseteq \mathbb{S}$ the set of probe states $\{\rho_i : i \in [1, M]\}$ that are used to generate the statistics (i.i.d. assumptions everywhere)
- **Observational Completeness:** $\mathcal{S} \subseteq \mathbb{S}$ is observationally complete for measurement M whenever **DDI**$[M(\mathcal{S})] = M(\mathbb{S})$
- the entire $\mathbb{S}$ is obviously observationally complete for any measurement
- Question: are there less demanding OC sets?

When Is the Inference Correct?

- assume that there is a "true but unknown" measurement to be inferred, and that the hypothesis about the underlying theory is "correct"
- what data are needed so that **DDI** returns the correct range?
- denote by $\mathcal{S} \subseteq \mathbb{S}$ the set of probe states $\{\rho_i : i \in [1, M]\}$ that are used to generate the statistics (i.i.d. assumptions everywhere)
- **Observational Completeness:** $\mathcal{S} \subseteq \mathbb{S}$ is observationally complete for measurement M whenever **DDI**$[M(\mathcal{S})] = M(\mathbb{S})$
- the entire $\mathbb{S}$ is obviously observationally complete for any measurement
- Question: are there less demanding OC sets?
- **Theorem:** $\mathcal{S} \subseteq \mathbb{S}$ is observationally complete *for any* measurement whenever **DDI**$[\mathcal{S}] = \mathbb{S}$

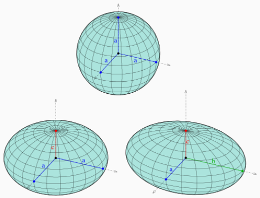
The special case of spherical theories

When $\mathbb{S}$ Is a Hypersphere...



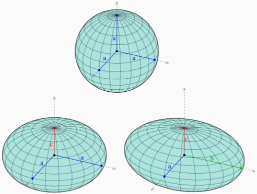
- ...any measurement range is an ellipsoid

When $\mathbb{S}$ Is a Hypersphere...



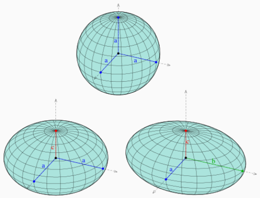
- ...any measurement range is an ellipsoid
- hence, **DDI** returns the minimum-volume enclosing ellipsoid which is **efficiently computed and always unique** *for any* dataset D (John, 1948)

When $\mathbb{S}$ Is a Hypersphere...



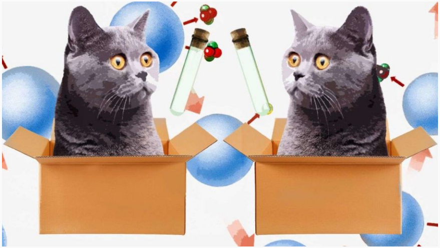
- ...any measurement range is an ellipsoid
- hence, **DDI** returns the minimum-volume enclosing ellipsoid which is **efficiently computed and always unique** *for any* dataset D (John, 1948)
- in fact, hyperspherical theories are *exactly those* that allow a unique inference *for any* dataset

When $\mathbb{S}$ Is a Hypersphere...



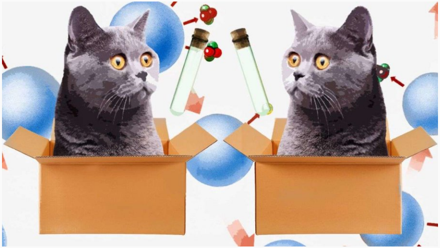
- ...any measurement range is an ellipsoid
- hence, **DDI** returns the minimum-volume enclosing ellipsoid which is **efficiently computed and always unique** *for any* dataset D (John, 1948)
- in fact, hyperspherical theories are *exactly those* that allow a unique inference *for any* dataset
- **DDI** may still return an ellipsoid which is not the range of a valid measurement: in this case a failure is announced

The Case of Qubits



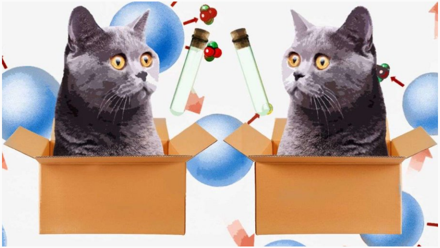
- gauge symmetries are unitary and antiunitary transformations

The Case of Qubits



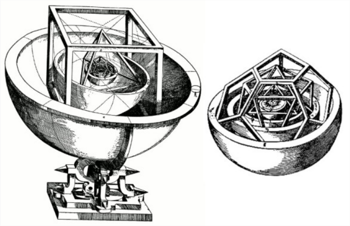
- gauge symmetries are unitary and antiunitary transformations
- hence, **DDI** is able to return a qubit measurement up to unitaries or antiunitaries

The Case of Qubits



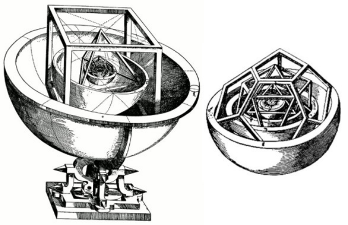
- gauge symmetries are unitary and antiunitary transformations
- hence, **DDI** is able to return a qubit measurement up to unitaries or antiunitaries
- moreover, a representative measurement can be explicitly constructed for any range (closed formula)

Observationally Complete Sets for Qubits



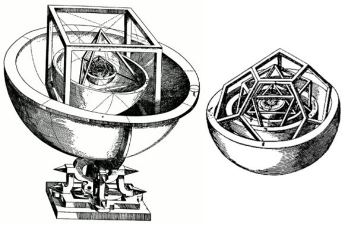
- a set $\mathcal{S}$ is OC iff $\textbf{MVVE}(\mathcal{S}) = \mathbb{S}$

Observationally Complete Sets for Qubits



- a set $\mathcal{S}$ is OC iff **MVVE**$(\mathcal{S}) = \mathbb{S}$
- **Fact:** in any real dimension ℓ, the minimum-volume ellipsoid enclosing $\ell + 1$ points is a hypersphere iff the points form a regular simplex

Observationally Complete Sets for Qubits



- a set $\mathcal{S}$ is OC iff $\textbf{MVVE}(\mathcal{S}) = \mathbb{S}$
- **Fact:** in any real dimension ℓ, the minimum-volume ellipsoid enclosing $\ell + 1$ points is a hypersphere iff the points form a regular simplex
- hence, SIC ensembles are OC

Example: Observational VS Informational Completeness

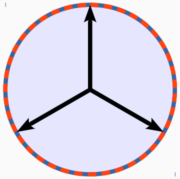


Figure 2: A regular simplex is OC.

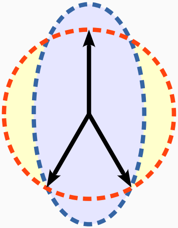


Figure 3: An irregular simplex is not OC.

Example: Observational VS Informational Completeness

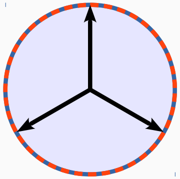


Figure 2: A regular simplex is OC.

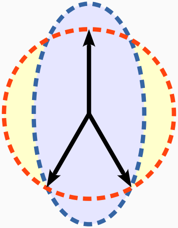


Figure 3: An irregular simplex is not OC.

In particular:

- a pure SIC ensemble is also OC

Example: Observational VS Informational Completeness

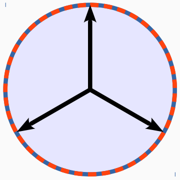


Figure 2: A regular simplex is OC.

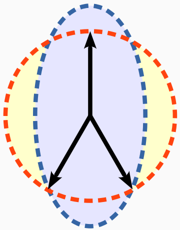


Figure 3: An irregular simplex is not OC.

In particular:

- a pure SIC ensemble is also OC
- a depolarized SIC ensemble still is IC and "symmetric" but is not OC anymore

Example: DDI in Action

- suppose the dataset comprises three probability distributions in $\mathbb{R}^4$, that is $D = \{\boldsymbol{p}_1 = (\frac{1}{2}, 0, \frac{1}{4}, \frac{1}{4}), \boldsymbol{p}_2 = (\frac{1}{8}, \frac{3}{8}, \frac{2+\sqrt{3}}{8}, \frac{2-\sqrt{3}}{8}), \boldsymbol{p}_3 = (\frac{1}{8}, \frac{3}{8}, \frac{2-\sqrt{3}}{8}, \frac{2+\sqrt{3}}{8})\}$

Example: DDI in Action

- suppose the dataset comprises three probability distributions in $\mathbb{R}^4$, that is $D = \{\boldsymbol{p}_1 = (\frac{1}{2}, 0, \frac{1}{4}, \frac{1}{4}), \boldsymbol{p}_2 = (\frac{1}{8}, \frac{3}{8}, \frac{2+\sqrt{3}}{8}, \frac{2-\sqrt{3}}{8}), \boldsymbol{p}_3 = (\frac{1}{8}, \frac{3}{8}, \frac{2-\sqrt{3}}{8}, \frac{2+\sqrt{3}}{8})\}$
- suppose that, for the inference, we assume a theory with a spherical state set: for example, a qubit

Example: DDI in Action

- suppose the dataset comprises three probability distributions in $\mathbb{R}^4$, that is $D = \{\boldsymbol{p}_1 = (\frac{1}{2}, 0, \frac{1}{4}, \frac{1}{4}), \boldsymbol{p}_2 = (\frac{1}{8}, \frac{3}{8}, \frac{2+\sqrt{3}}{8}, \frac{2-\sqrt{3}}{8}), \boldsymbol{p}_3 = (\frac{1}{8}, \frac{3}{8}, \frac{2-\sqrt{3}}{8}, \frac{2+\sqrt{3}}{8})\}$
- suppose that, for the inference, we assume a theory with a spherical state set: for example, a qubit
- **DDI:** the four effects are coplanar and arranged in a square

Conclusions

- *inference of quantum devices* from classical data

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference
- *emergence of SIC* qubit measurements as the minimal OC qubit measurements

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference
- *emergence of SIC* qubit measurements as the minimal OC qubit measurements
- are there finite minimal OC sets for all dimensions? would these always be SIC?

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference
- *emergence of SIC* qubit measurements as the minimal OC qubit measurements
- are there finite minimal OC sets for all dimensions? would these always be SIC?
- OC-ness does not need the Hilbert space structure

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference
- *emergence of SIC* qubit measurements as the minimal OC qubit measurements
- are there finite minimal OC sets for all dimensions? would these always be SIC?
- OC-ness does not need the Hilbert space structure

References:

- initial idea (case of qubit channels): F.B. and M. Dall'Arno. `arXiv:1805.01159`
- experiment: I. Agresti, D. Poderini, G. Carvacho, L. Serra, R. Chaves, F.B., M. Dall'Arno, F. Sciarrino. `arXiv:1806.00380`
- this talk: M. Dall'Arno, F.B., A. Bisio, A. Tosini. `arXiv:1812.08470`
- M. Dall'Arno, A. Ho, F.B., V. Scarani. `arXiv:1905.04895`

Conclusions

- *inference of quantum devices* from classical data
- inference based on idea of *self-consistent minimality*
- *observationally complete sets* allow correct inference
- *emergence of SIC* qubit measurements as the minimal OC qubit measurements
- are there finite minimal OC sets for all dimensions? would these always be SIC?
- OC-ness does not need the Hilbert space structure

References:

- initial idea (case of qubit channels): F.B. and M. Dall'Arno. arXiv:1805.01159
- experiment: I. Agresti, D. Poderini, G. Carvacho, L. Serra, R. Chaves, F.B., M. Dall'Arno, F. Sciarrino. arXiv:1806.00380
- this talk: M. Dall'Arno, F.B., A. Bisio, A. Tosini. arXiv:1812.08470
- M. Dall'Arno, A. Ho, F.B., V. Scarani. arXiv:1905.04895

thank you