Title: Experimental state and measurement tomography for generalised probabilistic theories: bounding deviations from quantum theory via noncontextuality inequality violations

Date: Jul 28, 2017 11:30 AM

URL: http://pirsa.org/17070057

Abstract: In order to perform foundational experiments testing the correctness of quantum mechanics, one requires data analysis tools that do not assume quantum theory. We introduce a quantum-free tomography technique that fits experimental data to a set of states and measurement effects in a generalised probabilistic theory (GPT). (This is in contrast to quantum tomography, which fits data to sets of density operators and POVM elements.) We perform an experiment on the polarization degree of freedom of single photons, and find GPT descriptions of the states and measurements in our experiment. We gather data for a large number of preparation and measurement procedures in order to map out the spaces of allowed GPT states and measurement effects, and we bound their possible deviation from quantum theory. Our GPT tomography method allows us to bound the extent to which nature might be more or less contextual than quantum theory, as measured by the maximum achievable violation of a particular noncontextuality inequality. We find that the maximal violation is confined to lie between 1.2Â±0.1% less than and 1.3Â±0.1% greater than the quantum prediction.

Coauthors: Matthew Pusey, Robert Spekkens, Kevin Resch
Akaike information criterion

• Criterion for model selection
  \[ AIC = -2 \log \mathcal{L} + 2n \]

• Lower AIC value implies higher relative model likelihood
  • Trade-off between not underfitting and overfitting

• For normally-distributed, independent errors
  \[ \mathcal{L} = e^{-\chi^2} \]
  \[ AIC(k) = 2\chi^2(k) + 2n(k) \]
**Akaike information criterion**

<table>
<thead>
<tr>
<th>Model no.</th>
<th>AIC value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$AIC_1$</td>
</tr>
<tr>
<td>2</td>
<td>$AIC_2$</td>
</tr>
<tr>
<td>3</td>
<td>$AIC_3$</td>
</tr>
</tbody>
</table>

Good for **comparing models to each other**
Akaike information criterion

<table>
<thead>
<tr>
<th>Model no.</th>
<th>AIC value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AIC_1</td>
</tr>
<tr>
<td>2</td>
<td>AIC_2</td>
</tr>
<tr>
<td>3</td>
<td>AIC_3</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Good for **comparing models to each other**

\[
\text{likelihood}(k) = \frac{\exp \left\{ -\frac{1}{2} (AIC(k) - AIC_{\min}) \right\}}{\sum_k \exp \left\{ -\frac{1}{2} (AIC(k) - AIC_{\min}) \right\}}
\]
Akaike information criterion

\[ AIC(k) = 2\chi^2(k) + 2n(k) \]
Akaike information criterion

\[ AIC(k) = 2\chi^2(k) + 2n(k) \]

\[ \text{likelihood}(k) = \frac{\exp\left\{-\frac{1}{2}(AIC(k) - AIC_{\text{min}})\right\}}{\sum_k \exp\left\{-\frac{1}{2}(AIC(k) - AIC_{\text{min}})\right\}} \]

Use \( k = 4 \)
Experiment

Raw data

Fit data to rank-k model

Decompose
Measured state and effect spaces
Measured state and effect spaces
Measured state and effect spaces

True state space $S_{\text{true}}$ lies between $S$ and $E_{\text{dual}}$
True effect space lies between $E$ and dual of $S_{\text{true}}$
Measured state and effect spaces

True state space \( S_{\text{true}} \) lies between \( S \) and \( E_{\text{dual}} \)

True effect space lies between \( E \) and dual of \( S_{\text{true}} \)

More preparations and measurements could increase volume ratio

\[ V_S / V_{E_{\text{dual}}} = 0.91267 \pm 0.00001 \]

→ Noise prevents reconstructing QT exactly
1000 preparations and measurements

- 1006 measurements on 6 states
- 6 measurements on 1006 states
- If data is rank 4 $\Rightarrow$ can predict empty 1000x1000 entries

- Chose 6 measurements (vs 4) to allow possibility of revealing additional dimensionality
More preparations and measurements

\[ \frac{V_S}{V_{E_{\text{dual}}}} = 0.968 \pm 0.001 \]
GPT tomography

- Put states and measurements on equal footing (self-consistent)
GPT tomography

• Put states and measurements on equal footing (self-consistent)

• Dimension inferred from data

• Directly extracted qubit-like state and measurement spaces without invoking QM

• Placed a small upper bound on possible deviation from QM
Outline

• GPT framework

• GPT tomography method and application to an experiment

• Noncontextuality inequality
Ontological models framework

- Preparations

- Measurements

- Probabilities

\[ p(X|M, P) = \int d\lambda \, \mu(\lambda|P) \xi(X|\lambda, M) \]
Operational equivalence

\[ P \]

If
\[ p(X|P, M) = p(X|P', M) \quad \forall \ X, \ M \]

then P and P' are operationally equivalent

[Spekkens, PRA 71, 052108 (2005)]
Noncontextuality

Operational equivalence implies equivalence in the ontological model

\[ p(X|P, M) = p(X|P', M) \quad \forall X, M \]
\[ \implies \mu(\lambda|P) = \mu(\lambda|P') \]

Preparation noncontextuality

\[ p(X|P, M) = p(X|P, M') \quad \forall X, P \]
\[ \implies \xi(X|\lambda, M) = \xi(X|\lambda, M') \]

Measurement noncontextuality

[Spekkens, PRA 71, 052108 (2005)]
Parity-oblivious multiplexing “game”

\[ x \in \{00, 01, 10, 11\} \]

\[ y \in \{0, 1\} \]

\[ b \in \{0, 1\} \]

[SBKTP, PRL 102, 010401 (2009)]
Parity-oblivious multiplexing “game”

- Bob guesses $y$-th bit of Alice’s input
  - Bob is not allowed to learn anything about the parity of Alice’s input

\[
p(\text{success}) = \frac{1}{8} \sum_{x,y} p(b = x_y | x, y)
\]

[SBKTP, PRL 102, 010401 (2009)]
Classical strategy

\[ x \in \{00, 01, 10, 11\} \]

- Alice encodes \( x_0 \) and sends it to Bob

\[ p(\text{success}) \leq \frac{3}{4} \]

\[ y \in \{0, 1\} \]

\[ b \in \{0, 1\} \]

[SBKTP, PRL 102, 010401 (2009)]
A noncontextuality inequality

Implement one of 4 preparations: $P_{00}, P_{01}, P_{10}, P_{11}$

\[
\begin{align*}
P_{\text{even}} &= \frac{1}{2} P_{00} + \frac{1}{2} P_{11} \\
P_{\text{odd}} &= \frac{1}{2} P_{01} + \frac{1}{2} P_{10}
\end{align*}
\]

Parity obliviousness $\implies p(b|P_{\text{even}}, M) = p(b|P_{\text{odd}}, M), \quad \forall b, M$

Prep. noncontextuality $\implies p(\lambda|P_{\text{even}}) = p(\lambda|P_{\text{odd}})$

\[
\frac{p(P_{\text{even}}|\lambda)p(\overline{P_{\text{even}}})}{p(\overline{\lambda})} = \frac{p(P_{\text{odd}}|\lambda)p(\overline{P_{\text{odd}}})}{p(\overline{\lambda})}
\]

\[p(P_{\text{even}}|\lambda) = p(P_{\text{odd}}|\lambda)\]

$\implies p(\text{success}) \leq \frac{3}{4}$

[SBKTP, PRL 102, 010401 (2009)]
Quantum strategy

\[ x \in \{00, 01, 10, 11\} \]

\[ y \in \{0, 1\} \]

\[ b \in \{0, 1\} \]

[SBKTP, PRL 102, 010401 (2009)]
Quantum strategy

\[ x \in \{00, 01, 10, 11\} \]

\[ y \in \{0, 1\} \]

\[ p(\text{success}) \leq \cos^2\left(\frac{\pi}{8}\right) \approx 0.85355 \]

[SBKTP, PRL 102, 010401 (2009)]
Experimental violation of inequality

Search over all GPT states and effects to find maximum inequality violation
Experimental violation of inequality

Search over all GPT states and effects to find maximum inequality violation
Parity obliviousness $\implies S_{\text{even}} = S_{\text{odd}}$
Approximate state and effect spaces with spheres

$p(\text{success})_{\text{qubit}} \approx 0.8536$
Experimental violation of inequality

$\textit{p(success)}_{\text{min}} = 0.8303 \pm 0.0002$

Search over all GPT states and effects to find maximum inequality violation

Parity obliviousness $\implies s_{\text{even}} = s_{\text{odd}}$

Approximate state and effect spaces with spheres

Get operational equivalence “for free”

Gives an experimental lower bound on “how contextual” our system is

$p(\text{success})_{\text{qubit}} \approx 0.8536$
Upper bound

Search over all dual GPT states and effects to find maximum inequality violation

Approximate state and effect spaces with spheres
1000x1000 data

\[ p(\text{success})_{\text{min}} = 0.8427 \pm 0.0005 \]

\[ p(\text{success})_{\text{max}} = 0.8647 \pm 0.0005 \]

\[ p(\text{success})_{\text{qubit}} \approx 0.8536 \]
GPT tomography

- Reconstruct full state and effect spaces of a system

- Minimize assumptions about the dimensionality of the system

- GPT tomography allowed us to test a noncontextuality inequality
  - Overcame problem of inexact operation equivalence
  - Found both lower and upper bounds on “amount of contextuality” in our system
Experimental state and measurement tomography for generalised probabilistic theories:
Bounding deviations from quantum theory via noncontextuality inequality violations

Mike Mazurek
Matt Pusey, Rob Spekkens, Kevin Resch

Contextuality: Conceptual Issues, Operational Signatures, and Applications
July 24-28, 2017
Introduction

• Contextuality: Conceptual Issues, Operational Signatures, and Applications

• Data analysis tools should not rely on quantum theory for foundational experiments

• Tomography method within generalised probabilistic theory framework
  • Self-consistent

• Application to experimental test of noncontextuality
Outline

• GPT framework

• GPT tomography method and application to an experiment

• Noncontextuality inequality
GPT framework

\[ p(0 \mid P_i, M_j) \]

[Hardy, quant-ph/0101012 (2001)]
[Barrett, PRA 75, 032304 (2007)]
GPT framework

\[ P_1, \ldots, P_m \]

Operational state: \( s_i = \begin{bmatrix} p(0|P_i, M_1) \\ p(0|P_i, M_2) \\ p(0|P_i, M_3) \\ \vdots \end{bmatrix} \)

Operational effect: \( e_{j,0} = \begin{bmatrix} \vdots \\ 0 \\ 1 \\ \vdots \end{bmatrix} \)
GPT framework

Operational state: \( s_i = \begin{bmatrix} p(0|P_i, M_1) \\ p(0|P_i, M_2) \\ p(0|P_i, M_3) \\ \vdots \end{bmatrix} \)

Operational effect: \( e_{j,0} = \begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} \)

\( p(0|P_i, M_j) = s_i \cdot e_{j,0} \)
Tomographically complete measurement set

- Number of possible measurements might be very large or infinite
- Will exist some **tomographically complete** set of $K$ measurements which fully determine each state

\[
\text{Operational state: } s_i = \begin{bmatrix}
p(0|P_i, M_1) \\
p(0|P_i, M_2) \\
\vdots \\
p(0|P_i, M_K)
\end{bmatrix}
\]

\[
p(0|P_i, M) = f_{M,0}(s_i)
\]
Convex linearity of operational states

\[ P_c \]

\[ \begin{align*}
    w & \rightarrow P_a \\
    1 - w & \rightarrow P_b
\end{align*} \]
Convex linearity of operational states

\[ p(0|P_c, M) = w p(0|P_a, M) + (1 - w) p(0|P_b, M) \]
Convex linearity of operational states

\[ p(0|P_c, M) = wp(0|P_a, M) + (1 - w)p(0|P_b, M) \]

\[ f_{M,0}(\mathbf{s}_c) = wf_{M,0}(\mathbf{s}_a) + (1 - w)f_{M,0}(\mathbf{s}_b) \]

Also true for \( M \)s in tomographically complete set

\[ \mathbf{s}_c = w\mathbf{s}_a + (1 - w)\mathbf{s}_b \]
Convex linearity of operational states

\[ p(0|P_c, M) = wp(0|P_a, M) + (1 - w)p(0|P_b, M) \]

\[ f_{M,0}(s_c) = w f_{M,0}(s_a) + (1 - w) f_{M,0}(s_b) \]

Also true for \( M \)s in tomographically complete set

\[ s_c = w s_a + (1 - w) s_b \]

\[ f_{M,0}(w s_a + (1 - w) s_b) = w f_{M,0}(s_a) + (1 - w) f_{M,0}(s_b) \]

\( f(s) \) is convex linear
Linearity

Convex linearity

\[ s = \sum_i w_i s_i \Rightarrow f_{M,0}(s) = \sum_i w_i f_{M,0}(s_i) \quad w_i \geq 0, \quad \sum_i w_i = 1 \]

implies linearity

\[ s = \sum_i \alpha_i s_i \Rightarrow f_{M,0}(s) = \sum_i \alpha_i f_{M,0}(s_i) \quad \forall \alpha_i \in \mathbb{R} \]

[Hardy, quant-ph/0101012 (2001)]
Linearity

Convex linearity

\[ s = \sum_i w_i s_i \Rightarrow f_{M,0}(s) = \sum_i w_i f_{M,0}(s_i) \quad w_i \geq 0, \quad \sum_i w_i = 1 \]

implies linearity

\[ s = \sum_i \alpha_i s_i \Rightarrow f_{M,0}(s) = \sum_i \alpha_i f_{M,0}(s_i) \quad \forall \alpha_i \in \mathbb{R} \]

\[ \forall M \quad \exists e_{M,0} : f_{M,0}(s) = e_{M,0} \cdot s \]

[Hardy, quant-ph/0101012 (2001)]
2-level classical system (bit)

\[ K = 2 \]

\[ s = (s_0, s_1) = (1, s_1) \quad \text{e} = (e_0, e_1) \]
2-level classical system (bit)

\[ K = 2 \]

\[ \mathbf{s} = (s_0, s_1) = (1, s_1) \quad \quad \quad \mathbf{e} = (e_0, e_1) \]

![Diagram showing 2-level classical system with points (-1,1) and (1,1)]
2-level classical system (bit)

\[ K = 2 \quad p(0|M, P) = e_{M,0} \cdot s_P \]

\[ s = (s_0, s_1) = (1, s_1) \quad e = (e_0, e_1) \]
2-level classical system (bit)

\[ K = 2 \quad p(0|M, P) = e_{M,0} \cdot s_P \]

\[ s = (s_0, s_1) = (1, s_1) \]

\[ e = (e_0, e_1) \]

---

The diagram on the right illustrates the state space of a 2-level classical system, with axes labeled \( s_0 \) and \( s_1 \) for the system state and \( e_0 \) and \( e_1 \) for the effect. Points such as (0,1), (0,0), (1,1), and \((-\frac{1}{2}, \frac{1}{2})\) are marked to show the possible states and effects.
Qubit

\[ \rho = \frac{1}{2}(\mathbb{I} + s_X \sigma_X + s_Y \sigma_Z + s_3 \sigma_Z) \]

\[ E = \frac{1}{2}(e_0 \mathbb{I} + e_X \sigma_X + e_Y \sigma_Z + e_3 \sigma_Z) \]
Qubit

\[
\rho = \frac{1}{2}(\mathbb{1} + s_X\sigma_X + s_Y\sigma_Y + s_Z\sigma_Z)
\]

\[
E = \frac{1}{2}(e_0\mathbb{1} + e_X\sigma_X + e_Y\sigma_Y + e_Z\sigma_Z)
\]

\[
\text{Tr}(\rho E) = \frac{1}{2}(e_0 + e_X e_X + e_Y e_Y + e_Z e_Z)
\]

\[
= (1, s_X, s_Y, s_Z) \cdot \frac{1}{2} (e_0, e_X, e_Y, e_Z)
\]
Qubit

\[ \rho = \frac{1}{2}(\mathbb{I} + s_x \sigma_x + s_y \sigma_y + s_z \sigma_z) \]

\[ E = \frac{1}{2}(e_0 \mathbb{I} + e_x \sigma_x + e_y \sigma_y + e_z \sigma_z) \]

\[ \text{Tr}(\rho E) = \frac{1}{2}(e_0 + s_x e_x + s_y e_y + s_z e_z) \]

\[ = (1, s_x, s_y, s_z) \cdot \frac{1}{2}(e_0, e_x, e_y, e_z) \]
Qubit

\[ \rho = \frac{1}{2}(I + s_X \sigma_X + s_Y \sigma_Z + s_3 \sigma_Z) \]
\[ E = \frac{1}{2}(e_0 I + e_X \sigma_X + e_Y \sigma_Z + e_3 \sigma_Z) \]

\[ \text{Tr}(\rho E) = \frac{1}{2}(e_0 + s_X e_X + s_Y e_Y + s_Z e_Z) \]
\[ = (1, s_X, s_Y, s_Z) \cdot \frac{1}{2}(e_0, e_X, e_Y, e_Z) \]
Generalised non-signalling theory
Generalised non-signalling theory

- Some states are deterministic for multiple effects
- Some effects respond deterministically to multiple states
Some example GPTs

- Qubit
- GNST
- Spekkens’ toy theory
Some example GPTs

Qubit

GNST

Spekkens’ toy theory

Random GPT
Dual state and effect spaces

• There might be logically possible states and effects that aren’t included in the spaces specified by the GPT
Dual state and effect spaces

• There might be **logically possible** states and effects that aren’t included in the spaces specified by the GPT

\[
\mathbf{e}_{\text{possible}} : 0 \leq \mathbf{e}_{\text{possible}} \cdot \mathbf{s} \leq 1 \quad \forall \mathbf{s} \in S
\]
Dual state and effect spaces

- There might be logically possible states and effects that aren’t included in the spaces specified by the GPT

\[ e_{\text{possible}} : 0 \leq e_{\text{possible}} \cdot s \leq 1 \quad \forall \ s \in S \]

\[ s_{\text{possible}} : 0 \leq e \cdot s_{\text{possible}} \leq 1 \quad \forall \ e \in E \]

\[ e_{\text{possible}} \in S_{\text{dual}} \]

\[ s_{\text{possible}} \in E_{\text{dual}} \]
Dual state and effect spaces

- There might be logically possible states and effects that aren’t included in the spaces specified by the GPT

\[
\begin{align*}
e_{\text{possible}} &: 0 \leq e_{\text{possible}} \cdot s \leq 1 \quad \forall s \in S \\
s_{\text{possible}} &: 0 \leq e \cdot s_{\text{possible}} \leq 1 \quad \forall e \in E \\
\end{align*}
\]

- GPTs in which \( S = E_{\text{dual}} \) and \( E = S_{\text{dual}} \) are self dual
  - These satisfy no-restriction hypothesis
Some example GPTs

Qubit

GNST

Spekkens’ toy theory

Random GPT
Outline

- GPT framework

- GPT tomography method and application to an experiment

- Noncontextuality inequality
Infinite-run statistics

\[
\begin{pmatrix}
1 & p(0|P_1, M_2) & p(0|P_1, M_3) & p(0|P_1, M_4) & p(0|P_1, M_5) & \cdots \\
1 & p(0|P_2, M_2) & p(0|P_2, M_3) & p(0|P_2, M_4) & p(0|P_2, M_5) & \cdots \\
1 & p(0|P_3, M_2) & p(0|P_3, M_3) & p(0|P_3, M_4) & p(0|P_3, M_5) & \cdots \\
1 & p(0|P_4, M_2) & p(0|P_4, M_3) & p(0|P_4, M_4) & p(0|P_4, M_5) & \cdots \\
1 & p(0|P_5, M_2) & p(0|P_5, M_3) & p(0|P_5, M_4) & p(0|P_5, M_5) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots 
\end{pmatrix}
\]
Infinite-run statistics

\[
\begin{pmatrix}
1 & s_1^{(1)} & \cdots & s_k^{(1)} \\
1 & s_1^{(2)} & \cdots & s_k^{(2)} \\
1 & s_1^{(3)} & \cdots & s_k^{(3)} \\
1 & s_1^{(4)} & \cdots & s_k^{(4)} \\
1 & s_1^{(5)} & \cdots & s_k^{(5)} \\
\vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
\begin{pmatrix}
1 & e_0^{(2,0)} & e_0^{(3,0)} & e_0^{(4,0)} & e_0^{(5,0)} & \cdots \\
0 & e_1^{(2,0)} & e_1^{(3,0)} & e_1^{(4,0)} & e_1^{(5,0)} & \cdots \\
0 & e_k^{(2,0)} & e_k^{(3,0)} & e_k^{(4,0)} & e_k^{(5,0)} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\end{pmatrix}
\]

"0" "1"

\[P_1, \ldots, P_m \text{ Preparation} \]

\[M_1, \ldots, M_n \text{ Measurement} \]
GPT states and effects

\[
\begin{pmatrix}
1 & s_1^{(1)} & \ldots & s_k^{(1)} \\
1 & s_1^{(2)} & \ldots & s_k^{(2)} \\
1 & s_1^{(3)} & \ldots & s_k^{(3)} \\
1 & s_1^{(4)} & \ldots & s_k^{(4)} \\
1 & s_1^{(5)} & \ldots & s_k^{(5)} \\
\vdots & \vdots & \ddots & \vdots \\
0 & e_1^{(2,0)} & e_1^{(3,0)} & e_1^{(4,0)} & e_1^{(5,0)} & \ldots \\
0 & e_1^{(2,0)} & e_1^{(3,0)} & e_1^{(4,0)} & e_1^{(5,0)} & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
0 & e_k^{(2,0)} & e_k^{(3,0)} & e_k^{(4,0)} & e_k^{(5,0)} & \ldots \\
\end{pmatrix}
\]
GPT states and effects

\[
\begin{pmatrix}
1 & s_1^{(1)} & \cdots & s_k^{(1)} \\
1 & s_1^{(2)} & \cdots & s_k^{(2)} \\
1 & s_1^{(3)} & \cdots & s_k^{(3)} \\
1 & s_1^{(4)} & \cdots & s_k^{(4)} \\
1 & s_1^{(5)} & \cdots & s_k^{(5)} \\
\vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
\begin{pmatrix}
1 & e_0^{(2,0)} & e_0^{(3,0)} & e_0^{(4,0)} & e_0^{(5,0)} & \cdots \\
0 & e_1^{(2,0)} & e_1^{(3,0)} & e_1^{(4,0)} & e_1^{(5,0)} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
0 & e_k^{(2,0)} & e_k^{(3,0)} & e_k^{(4,0)} & e_k^{(5,0)} & \cdots \\
\end{pmatrix}
\]

\[
p(0|P_i, M_j) = s_i \cdot e_{j,0} = \begin{pmatrix}
1 & s_1^{(i)} & \cdots & s_k^{(i)}
\end{pmatrix} \cdot \begin{pmatrix}
e_0^{(j,0)} & \cdots & e_k^{(j,0)}
\end{pmatrix}
\]
Finite-run (noisy) statistics

\[
\begin{pmatrix}
1 & p(0|P_1, M_2) & p(0|P_1, M_3) & p(0|P_1, M_4) & p(0|P_1, M_5) & \cdots \\
1 & p(0|P_2, M_2) & p(0|P_2, M_3) & p(0|P_2, M_4) & p(0|P_2, M_5) & \cdots \\
1 & p(0|P_3, M_2) & p(0|P_3, M_3) & p(0|P_3, M_4) & p(0|P_3, M_5) & \cdots \\
1 & p(0|P_4, M_2) & p(0|P_4, M_3) & p(0|P_4, M_4) & p(0|P_4, M_5) & \cdots \\
1 & p(0|P_5, M_2) & p(0|P_5, M_3) & p(0|P_5, M_4) & p(0|P_5, M_5) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix}
\]

Full rank data matrix -> not factorizable!
Experiment

Raw data
Experiment

Raw data

Fit data to rank-k model

Decompose
Experiment

Raw data

Fit data to rank-k model

Decompose

\[ s_1 \quad s_2 \quad e_1 \quad e_2 \quad e_3 \]

\[ s_1 \quad s_2 \quad e_1 \quad e_2 \quad e_3 \]

\[ s_1 \quad s_2 \quad e_1 \quad e_2 \quad e_3 \]

\[ s_1 \quad s_2 \quad e_1 \quad e_2 \quad e_3 \]

\[ s_1 \quad s_2 \quad e_1 \quad e_2 \quad e_3 \]
Experimental set-up

Polarization preparation

Polarization measurement
\[ p(0 | P_i, M_j) \]

- 100 preparation settings
- 100 measurement settings
- Noise ensures data table is full rank
- Fit data to models of various rank, see which performs best
Fitting to a rank-\(k\) model

Assume noise in \(D\) is independent and Poissonian \(\sim\) Gaussian

\[
D = \begin{array}{|c|c|c|c|c|c|}
\hline
\text{Measurement} & 30 & 40 & 50 & 60 & 70 \\
\text{Parameter} & 0 & 0.2 & 0.4 & 0.6 & 0.8 \\
\hline
\end{array}
\]

\[
\min_{K} \chi^2 = \sum_{i,j} \frac{(D_{i,j} - K_{i,j})^2}{(\Delta D_{i,j})^2}
\]

subj. to \(\text{rank}(K) \leq k,\)

\[
0 \leq K_{i,j} \leq 1 \quad \forall \ i, j
\]
Fitting to a rank-$k$ model

Assume noise in $D$ is independent and Poissonian $\sim$ Gaussian

$$D = \ldots$$

$$\min_{K} \chi^{2} = \sum_{i,j} \frac{(D_{i,j} - K_{i,j})^{2}}{(\Delta D_{i,j})^{2}}$$

subj. to $\text{rank}(K) \leq k,$

$0 \leq K_{i,j} \leq 1 \ \forall \ i, j$
Fitting to a rank-$k$ model

Assume noise in $D$ is independent and Poissonian $\sim$ Gaussian

$$D = \dots$$

$$\min_K \chi^2 = \sum_{i,j} \frac{(D_{i,j} - K_{i,j})^2}{(\Delta D_{i,j})^2}$$

subj. to $\text{rank}(K) \leq k,$

$$0 \leq K_{i,j} \leq 1 \quad \forall i, j$$
Fitting to a rank-$k$ model

Assume noise in $D$ is independent and Poissonian $\sim$ Gaussian

$$D = \text{[Image of a contour plot with a range of values]}$$

$$\min_K \chi^2 = \sum_{i,j} \frac{(D_{i,j} - K_{i,j})^2}{(\Delta D_{i,j})^2}$$

subj. to \text{rank}(K) \leq k,
\quad 0 \leq K_{i,j} \leq 1 \quad \forall \, i, j$$

Rank-$k$ parameterization of $K$:
$$K = SE$$
Determining $k$

Ranks 2 and 3 underfit the data
Akaike information criterion

• Criterion for model selection

\[ AIC = -2 \log \mathcal{L} + 2n \]

• Lower AIC value implies higher relative model likelihood
  • Trade-off between not underfitting and overfitting