Entropy Distance: New Quantum Phenomena

A. Knauf  S. Weis

Department Mathematik
University of Erlangen

Information Geometry and its Applications III
Leipzig, 6th August 2010
<table>
<thead>
<tr>
<th></th>
<th>Motivation</th>
<th>Mean Value Chart</th>
<th>Projection Theorem</th>
<th>Pythagorean Theorem</th>
<th>Applications</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Motivation

Mean Value Chart

Projection Theorem

Pythagorean Theorem

Applications

Summary
1 Motivation

2 Mean Value Chart

3 Projection Theorem

4 Pythagorean Theorem

5 Applications

6 Summary

A. Knauf, S. Weis

Department Mathematik University of Erlangen

Entropy Distance: New Quantum Phenomena
1. Motivation

2. Mean Value Chart

3. Projection Theorem

4. Pythagorean Theorem

5. Applications

6. Summary
Entropy Distance as an Information Functional

Definition

Entropy distance $d_\mathcal{E}$ is the relative entropy distance from an exponential family $\mathcal{E}$ in a finite-dimensional matrix algebra $\mathcal{A}$. Classical algebra $\mathcal{A} \cong \mathbb{C}^N$, quantum otherwise.


Applications of entropy distance include

- **MLE** through the log-likelihood function (classical),
- the stochastic interaction measure of *multi-information*, this is the entropy distance from the independence model (classical & quantum).
Previous Work on $d_\mathcal{E}$ in Classical Probability Theory

Barndorff-Nielsen (’78), Čencov (’82), Ay (’02), Csiszár and Matúš (’03, ’05, ’08)

Rough Idea:

- Pythagorean theorem of relative entropy implies projection theorem along the normal space,
- mean value chart maps $\mathcal{E}$ to the mean value set (convex support),
- extension of $\mathcal{E}$ implies optimal projection theorem and optimal Pythagorean theorem.
Previous Work on $d_\mathcal{E}$ in Classical Probability Theory

Barndorff-Nielsen (’78), Čencov (’82), Ay (’02), Csiszár and Matúš (’03, ’05, ’08)

Rough Idea:

- Pythagorean theorem of relative entropy implies projection theorem along the normal space,
- mean value chart maps $\mathcal{E}$ to the mean value set (convex support),
- extension of $\mathcal{E}$ implies optimal projection theorem and optimal Pythagorean theorem.
Previous Work on $d_E$ in Classical Probability Theory

Barndorff-Nielsen ('78),
Čencov ('82),
Ay ('02),
Csiszár and Matúš ('03, '05, '08)

Rough Idea:

- Pythagorean theorem of relative entropy implies projection theorem along the normal space,
- mean value chart maps $E$ to the mean value set (convex support),
- extension of $E$ implies optimal projection theorem and optimal Pythagorean theorem.
Previous Work on $d_\mathcal{E}$ in the Quantum Case

Classical extension of $\mathcal{E}$ is compact, the mean value set is a polytope and $d_\mathcal{E}$ is continuous. All wrong for quantum case!

Well-known in the quantum case:

- Pythagorean theorem implies (only in a restriction!) a projection theorem along the normal space, see Petz (’08),
- mean value chart and mean value set are known, see Wichmann (’63).

New in our work: Convex structure of mean value set includes non-exposed faces; extensions of $\mathcal{E}$; optimal projection theorem and Pythagorean theorem.
Exponential Families

Definition

State space \( \bar{S} := \{ \rho \in A | \rho \geq 0, \text{tr}(\rho) = 1 \} \). Real vector space \( A_{sa} \subset A \) of self-adjoint matrices, Hilbert-Schmidt inner product. Analytic diffeo. \( \exp_1 : A_{sa}/\mathbb{R} \| \rightarrow S := \{ \rho \in \bar{S} | \rho^{-1} \text{ exists} \} \), \( a \mapsto \frac{e^a}{\text{tr}(e^a)} \), canonical chart \( \ln_0 = \exp_1^{-1} \) to traceless matrices.

Exponential family \( \mathcal{E} := \exp_1( \text{linear subspace of } A_{sa} ) \), \( V := \ln_0(\mathcal{E}) \) tangent space, \( V^\perp \) normal space, orth. projection \( \pi_V : A_{sa} \rightarrow V \), mean value set \( \text{mv}(V) := \pi_V(\bar{S}) \).

Theorem (Wichmann '63)

\( \pi_V \circ \exp_1 \big|_V : V \rightarrow \text{Int}(\text{mv}(V)) \) real analytic bijection.

A. Knauf, S. Weis
Department Mathematik University of Erlangen
Entropy Distance: New Quantum Phenomena
Two Examples

Definition (Mean value chart)

The analytic chart $\pi_V|_E$ onto $\text{Int}(\text{mv}(V))$ is called mean value chart.

$$\exp_1(\text{span}\{\sigma_1 \oplus 1, \sigma_2 \oplus 1\})$$

$$\exp_1(\text{span}\{\sigma_1 \oplus 0, \sigma_2 \oplus 1\})$$

Swallow family

Staffelberg family

$$\mathcal{A} := \text{Mat}(2, \mathbb{C}) \oplus \mathbb{C}$$
2D Mean Value Sets in $\mathcal{A} := \text{Mat}(2, \mathbb{C}) \oplus \mathbb{C}$

- $\overline{S}$ is the 4D cone $\text{conv}(\text{Bloch ball} \oplus 0, 0_2 \oplus 1)$,
- every 2D exponential family $\mathcal{E}$ of $\mathcal{A}$ is included in a 3D cone,
- modulo automorphism take $z := -\frac{1}{2} \mathbb{1}_2 \oplus 1$,
  $\mathcal{W} := \text{span}\{\sigma_1 \oplus 0, \sigma_2 \oplus 0, z\}$, $\mathcal{V} \subset \mathcal{W}$ and the cone
  $\mathcal{C} := (\frac{1}{3} \mathbb{1} + \mathcal{W}) \cap \overline{S} = \exp_1(\mathcal{W}) \supset \mathcal{E}$,
- a complete orbit invariant of 2D planes $\mathcal{V}$ is the angle
  $\angle(\mathcal{V}, z) \in [0, \frac{\pi}{2}]$.

Non-exposed faces are typical for a mean value set!
Exposed Faces, Poonems and Access Sequences

Definition (Rockafellar, Grünbaum, Csiszár and Matúš)

Let $M$ be a convex set. The intersection $F$ of $M$ with a supporting hyperplane of $M$ is called exposed face of $M$. $\emptyset$ and $M$ are exposed faces by definition. In these cases we write $F \overset{\text{Ex}}{<} M$. A sequence $F = F_1 \overset{\text{Ex}}{<} \cdots \overset{\text{Ex}}{<} F_k \overset{\text{Ex}}{<} M$ is called access sequence and $F$ is called poonem of $M$. A poonem, which is not an exposed face is called non-exposed face.

Concept of poonem is equivalent to the more popular concept of face.
The Pythagorean Theorem (Restricted Form)

**Definition**

The relative entropy of two states $\rho, \sigma \in \overline{S}$ is $S(\rho, \sigma) := \infty$ unless $\text{Im}(\sigma) \supset \text{Im}(\rho)$ and then $S(\rho, \sigma) := \text{tr}\rho(\ln \rho - \ln \sigma)$.

**Theorem (Monograph Petz ’08 for an overview)**

If $\rho, \sigma, \tau \in \overline{S}$ are states, $\sigma$ and $\tau$ are invertible and $\rho - \sigma \perp \ln(\tau) - \ln(\sigma)$ holds, then we have

$$S(\rho, \sigma) + S(\sigma, \tau) = S(\rho, \tau).$$

- The Pythagorean theorem induces the projection $\pi_E : \mathcal{E} + V^\perp \to \mathcal{E}, \ a \mapsto (a + V^\perp) \cap \mathcal{E}$ to an exponential family $\mathcal{E}$.
- For a state $\rho \in \mathcal{E} + V^\perp$ we have

$$d_\mathcal{E}(\rho) := \inf_{\sigma \in \mathcal{E}} S(\rho, \sigma) = S(\rho, \pi_\mathcal{E}(\rho)).$$

A. Knauf, S. Weis
Department Mathematik University of Erlangen

Entropy Distance: New Quantum Phenomena
Lattice Isomorphisms and Compressions

Assignment of an orth. projector $F \mapsto p^F$ to each face $F$ of $\text{mv}(V)$ through

- inverse projection $F \mapsto (F + V^\perp) \cap \overline{S}$ lifts faces of $\text{mv}(V)$ to faces of $\overline{S}$,
- (face lattice of $\overline{S}$) $\cong$ (projector lattice of $\mathcal{A}$).

Definition

Orth. projector $p$ defines projection $\mathcal{A} \to p\mathcal{A}p := \{pap | a \in \mathcal{A}\}$ by $a \mapsto pap$. We denote by $\exp_1^p$ and $\ln^p$ the trace normalized exponential and the logarithm in $p\mathcal{A}p$.

A face $F$ of $\text{mv}(V)$ defines the compressed exponential family $\mathcal{E}^{p^F} := \exp_1^{p^F} (p^F Vp^F)$. 
The Geodesic Closure

Definition

The e-geodesic through $\theta \in \mathcal{A}_{sa}$ in the direction of $\nu \in \mathcal{A}_{sa}$ is the curve $\gamma \mapsto g_{\theta, \nu}(\lambda) = \exp_{1}(\theta + \lambda \nu) \subset S$.

If $p$ is the maximal projector of $\nu$ (spectral projector of the largest eigenvalue), then $\mathcal{E}^p = \{ \lim_{\lambda \to \infty} g_{\theta, \nu}(\lambda) \mid \theta \in V \}$.

Definition

The geodesic closure of $\mathcal{E}$ is $\text{cl}_{\text{geo}}(\mathcal{E}) := \bigcup_{F} \mathcal{E}^{\text{F}}$. Here $F$ extends over the exposed faces ($\neq \emptyset$) of $\text{mv}(V)$.

The geodesic closure $\text{cl}_{\text{geo}}(\mathcal{E})$ exceeds $\mathcal{E}$ by the limit points of e-geodesics in $\mathcal{E}$.
The $rI$-Closure

Definition (Csiszár and Matúš)

The $rI$-closure of $\mathcal{E}$ is $\text{cl}_{rI}(\mathcal{E}) := \{ \rho \in \overline{S} \mid d_{\mathcal{E}}(\rho) = 0 \}$.

- E-geodesic asymptotics show $\text{cl}_{\text{geo}}(\mathcal{E}) \subset \text{cl}_{rI}(\mathcal{E})$.
- The proof idea to the following theorem is to concatenate e-geodesic asymptotics. For every poonem $F$ in an access sequence of $\text{mv}(V)$ we form the geodesic closure of the compressed exponential family $\mathcal{E}^{pF}$ and take the union:

$$\text{cl}_{rI}(\mathcal{E}) = \bigcup_F \mathcal{E}^{pF}.$$

- Equality $\text{cl}_{\text{geo}}(\mathcal{E}) = \text{cl}_{rI}(\mathcal{E})$ holds if and only if all faces of $\text{mv}(V)$ are exposed. Examples: independence model, convex family or classical algebra.
### Theorem (Weis ’09)

Let $\mathcal{E}$ be an exponential family with tangent space $V$.

1. If $\rho \in \overline{S}$, then $\rho + V^\perp$ intersects the reverse information closure $\text{cl}_{\text{ri}}(\mathcal{E})$ in a unique point denoted by $\pi_{\mathcal{E}}(\rho)$.

2. If $\rho \in \overline{S}$, then the relative entropy $S(\rho, \cdot)$ has a unique local minimum on $\text{cl}_{\text{ri}}(\mathcal{E})$ and

$$\min_{\sigma \in \text{cl}_{\text{ri}}(\mathcal{E})} S(\rho, \sigma) = S(\rho, \pi_{\mathcal{E}}(\rho)) = S_{\mathcal{E}}(\rho).$$
A Discontinuous Entropy Distance is Possible

Staffelberg family: **Discontinuity** on the base circle of the cone $C$
where $d_E \equiv 0$ except $d_E(\rho) = \ln(2)$ for $\rho := \frac{1}{2}(I_2 \oplus \sigma_2) \oplus 0$.

- The discontinuous $d_E$ at $\varphi = \angle(V, z) = \frac{\pi}{3}$ separates mean value sets with non-exposed faces from those without.
- According to the Pinsker-Csiszár inequality $\text{cl}_I(E) \subset \overline{E}$ holds. Equality $\text{cl}_I(E) = \overline{E} \iff d_E$ is continuous. E.g. indep. model, convex family or (Ay ’02) classical algebra.

A. Knauf, S. Weis
Department Mathematik University of Erlangen

Entropy Distance: New Quantum Phenomena
The Optimal Pythagorean Theorem

Theorem (Weis ’10)

Let $\mathcal{E}$ be an exponential family with tangent space $V$. If $\rho \in \overline{S}$ and $\sigma, \tau \in \text{cl}_{\text{ri}}(\mathcal{E})$ are states such that $\sigma - \rho \perp V$, then $S(\rho, \sigma) + S(\sigma, \tau) = S(\rho, \tau)$ holds.
Maximization of Entropy Distance

Previous work e.g. by Ay, Knauf, Matúš

Definition

If $\rho \in \overline{S}$ then the support projector of $\rho$ is the ortho. projector $s(\rho) \in A$ with the image $\text{Im}(s(\rho)) = \text{Im}(\rho)$. For $a \in A_{sa}$ denote the free energy $F(a) := \ln \text{tr} e^a$ and use $F^p$ for free energy in $pAp$.

Theorem (Knauf and Weis ’10)

Let $\rho \in \overline{S}$, $p := s(\rho)$ and $q := s(\pi_{\mathcal{E}}(\rho))$. For every traceless self-adjoint matrix $u \in pAp$ we have the directional derivative

$$D|_{\rho} d_{\mathcal{E}}(u) = \langle u, \ln^p(\rho) - \ln^q \circ \pi_{\mathcal{E}}(\rho) \rangle.$$  

If $\rho$ is a local maximizer of $d_{\mathcal{E}}$ then for $\theta := \ln^q \circ \pi_{\mathcal{E}}(\rho)$ we have $\rho = \exp^p(p\theta p)$ and $d_{\mathcal{E}}(\rho) = F^q(\theta) - F^p(p\theta p)$.  

A. Knauf, S. Weis  
Department Mathematik University of Erlangen

Entropy Distance: New Quantum Phenomena
Maximum Entropy (Wichmann ’63, Ingarden et al. ’97)

Let $U \subset \mathcal{A}_{sa}$ be a vector space and let $u \in U$.

**Definition**

The **von Neumann entropy** of a state $\rho \in \overline{S}$ is $S(\rho) := -\operatorname{tr} \rho \ln \rho$. The constraint set of $(U, u)$ is $C_{U,u} := \{\rho \in \overline{S} \mid \pi_U(\rho) = u\}$. Let $\mathcal{E} := \exp_1(U)$.

**Theorem (Weis ’09)**

We have $\arg\max_{\rho \in C_{U,u}} S(\rho) = C_{U,u} \cap \text{cl}_{\text{tr}}(\mathcal{E})$.

Idea: replace $\max_\rho S(\rho) - \ln \operatorname{tr} \mathbb{1} = \max_\rho -S(\rho, \frac{\mathbb{1}}{\operatorname{tr} \mathbb{1}})$ by $\min_\rho S(\rho, \frac{\mathbb{1}}{\operatorname{tr} \mathbb{1}})$ and use Pythagorean theorem

\[
\min_\rho [S(\rho, \pi_{\mathcal{E}}(\rho)) + S(\pi_{\mathcal{E}}(\rho), \frac{\mathbb{1}}{\operatorname{tr} \mathbb{1}})] = S(\pi_{\mathcal{E}}(\rho), \frac{\mathbb{1}}{\operatorname{tr} \mathbb{1}}).
\]
Coordinates for Maximum Entropy Ensembles

We fix $a_1, \ldots, a_k \in A_{sa}$ (observables) and put $V := \text{span}\{a_1, \ldots, a_k, \mathbb{I}\} \cap \{\text{tr}(\cdot) = 0\}$. Then we choose a tuple of mean values $(\xi_i)_{i=1}^k \in \{\{\text{tr}(a_i \rho)\}_{i=1}^k \mid \rho \in \mathcal{S}\} \cong \text{mv}(V)$.

**Theorem (Weis ’10)**

*There exists a unique face $G$ of $\text{mv}(V)$ and unique coefficients $\beta_1, \ldots, \beta_k \in \mathbb{R}$ such that for $a(\beta) := -\sum_{i=1}^k \beta_i p^G a_i p^G$ and for $j = 1, \ldots, k$ we have*

$$-\frac{\partial}{\partial \beta_j} F(a(\beta)) = \xi_j \langle p^G, \exp_1(a(\beta)) \rangle.$$

*The maximizer of von Neumann entropy $S$ with mean $(\xi_i)_{i=1}^k$ is $\rho := \exp_1^{p^G}(a(\beta))$ and $S(\rho) = \sum_{i=1}^k \beta_i \xi_i + F^{p^G}(a(\beta))$.***
General Properties of the Entropy Distance $d_\mathcal{E}$

- Entropy distance $d_\mathcal{E}(\rho) = S(\rho, \pi_\mathcal{E}(\rho))$ of a state $\rho$ given by projection $\pi_\mathcal{E} : \overline{\mathcal{S}} \to \text{cl}_{rl}(\mathcal{E})$ along the normal space $V^\perp$.

- Optimal Pythagorean theorem, for states $\rho \in \overline{\mathcal{S}}$ and $\tau \in \text{cl}_{rl}(\mathcal{E})$ we have $S(\rho, \tau) = d_\mathcal{E}(\rho) + S(\pi_\mathcal{E}(\rho), \tau)$.

- Geo. closure $\text{cl}_{geo}(\mathcal{E}) \subset \text{rl-closure} \text{ cl}_{rl}(\mathcal{E}) \subset \text{topo. closure} \overline{\mathcal{E}}$,
  - $\text{cl}_{geo}(\mathcal{E}) = \text{cl}_{rl}(\mathcal{E}) \iff$ no non-exposed faces,
  - $\text{cl}_{rl}(\mathcal{E}) = \overline{\mathcal{E}} \iff$ entropy distance is continuous.
General Properties of the Entropy Distance $d_{\mathcal{E}}$

- Entropy distance $d_{\mathcal{E}}(\rho) = S(\rho, \pi_{\mathcal{E}}(\rho))$ of a state $\rho$ given by projection $\pi_{\mathcal{E}} : \overline{S} \rightarrow \text{cl}_{rI}(\mathcal{E})$ along the normal space $V^\perp$.

- Optimal Pythagorean theorem, for states $\rho \in \overline{S}$ and $\tau \in \text{cl}_{rI}(\mathcal{E})$ we have $S(\rho, \tau) = d_{\mathcal{E}}(\rho) + S(\pi_{\mathcal{E}}(\rho), \tau)$.

- Geo. closure $\text{cl}_{\text{geo}}(\mathcal{E}) \subset rI$-closure $\text{cl}_{rI}(\mathcal{E}) \subset$ topo. closure $\overline{\mathcal{E}}$,
  - $\text{cl}_{\text{geo}}(\mathcal{E}) = \text{cl}_{rI}(\mathcal{E}) \iff$ no non-exposed faces,
  - $\text{cl}_{rI}(\mathcal{E}) = \overline{\mathcal{E}} \iff$ entropy distance is continuous.

A. Knauf, S. Weis

Department Mathematik University of Erlangen

Entropy Distance: New Quantum Phenomena
General Properties of the Entropy Distance $d_{\mathcal{E}}$

- Entropy distance $d_{\mathcal{E}}(\rho) = S(\rho, \pi_{\mathcal{E}}(\rho))$ of a state $\rho$ given by projection $\pi_{\mathcal{E}}: \overline{S} \rightarrow cl_{rl}(\mathcal{E})$ along the normal space $V^\bot$.

- Optimal Pythagorean theorem, for states $\rho \in \overline{S}$ and $\tau \in cl_{rl}(\mathcal{E})$ we have $S(\rho, \tau) = d_{\mathcal{E}}(\rho) + S(\pi_{\mathcal{E}}(\rho), \tau)$.

- Geo. closure $cl_{geo}(\mathcal{E}) \subset rl$-closure $cl_{rl}(\mathcal{E}) \subset$ topo. closure $\overline{\mathcal{E}}$,
  - $cl_{geo}(\mathcal{E}) = cl_{rl}(\mathcal{E}) \iff$ no non-exposed faces,
  - $cl_{rl}(\mathcal{E}) = \overline{\mathcal{E}} \iff$ entropy distance is continuous.
General Properties of the Entropy Distance $d_{\mathcal{E}}$

- Entropy distance $d_{\mathcal{E}}(\rho) = S(\rho, \pi_{\mathcal{E}}(\rho))$ of a state $\rho$ given by projection $\pi_{\mathcal{E}} : \overline{S} \to \text{cl}_r(\mathcal{E})$ along the normal space $V^\perp$.
- Optimal Pythagorean theorem, for states $\rho \in \overline{S}$ and $\tau \in \text{cl}_r(\mathcal{E})$ we have $S(\rho, \tau) = d_{\mathcal{E}}(\rho) + S(\pi_{\mathcal{E}}(\rho), \tau)$.
- Geo. closure $\text{cl}_{\text{geo}}(\mathcal{E}) \subset$ rl-closure $\text{cl}_r(\mathcal{E}) \subset$ topo. closure $\overline{\mathcal{E}}$,
  - $\text{cl}_{\text{geo}}(\mathcal{E}) = \text{cl}_r(\mathcal{E}) \iff$ no non-exposed faces,
  - $\text{cl}_r(\mathcal{E}) = \overline{\mathcal{E}} \iff$ entropy distance is continuous.
General Properties of the Entropy Distance $d_{E}$

- Entropy distance $d_{E}(\rho) = S(\rho, \pi_{E}(\rho))$ of a state $\rho$ given by projection $\pi_{E} : \overline{S} \rightarrow \text{cl}_{rI}(E)$ along the normal space $V_{\perp}$.

- Optimal Pythagorean theorem, for states $\rho \in \overline{S}$ and $\tau \in \text{cl}_{rI}(E)$ we have $S(\rho, \tau) = d_{E}(\rho) + S(\pi_{E}(\rho), \tau)$.

- Geo. closure $\text{cl}_{geo}(E) \subset \text{rl-closure} \text{ cl}_{rI}(E) \subset \text{topo. closure} \overline{E}$,
  - $\text{cl}_{geo}(E) = \text{cl}_{rI}(E) \iff$ no non-exposed faces,
  - $\text{cl}_{rI}(E) = \overline{E} \iff$ entropy distance is continuous.
Open Questions

- What can be said about special families, e.g. independence model, Boltzmann machines?
- Where are the discontinuities of entropy distance?
- What are the one-sided directional derivatives?
- Are there connections to the geometry of entanglement?
Evidence

A. Knauf and S. Weis.
Entropy Distance: New Quantum Phenomena.
*arXiv:1007.5464*

S. Weis.
The Pythagorean Theorem of Relative Entropy.
*arXiv:1003.5671*

S. Weis.
*www.opus.ub.uni-erlangen.de/opus/volltexte/2010/1580*
Thanks!
The Staffelberg Mountain