# (Post)Quantum Brègman divergences, nonlinear resource theories, and renormalisation

arXiv: **1710.01837*** & 1605.02063

*v2 to appear on Monday

Ryszard Paweł Kostecki

*National Quantum Information Centre & Institute of Informatics*
*Faculty of Mathematics, Physics, and Informatics, University of Gdańsk*

June 16, 2019

## Plan

1. Entropic projections as nonlinear state transformations
2. Brègman family of relative entropies
3. Brègman nonexpansive operations
4. Nonlinear resource theories based on Brègman nonexpansive operations
5. Smooth quantum information geometries derived from Brègman divergences
6. Jaynes–Mitchell–Favretti renormalisation

# Quantum information models

- trace class operators: $\mathcal{T}(\mathcal{H}) := \{\rho \in \mathfrak{B}(\mathcal{H}) \mid \rho \geq 0,\ \mathrm{tr}_{\mathcal{H}}|\rho| < \infty\}$
- arbitrary sets of denormalised quantum states: $\mathcal{M}(\mathcal{H}) \subseteq \mathcal{T}(\mathcal{H})^+$
- usually one assumes *a priori* that the morphisms $\mathcal{M}_1(\mathcal{H}_1) \to \mathcal{M}_2(\mathcal{H}_2)$ should be given by some CPTP maps, however there are several assumptions behind it
- our main motivation is to find a reasonable class of (quantum, postquantum) state spaces and morphisms between them which would not be linear and would not be CPTP, yet would provide a consistent description of suitable quantum information processing tasks $\Rightarrow$ characterisation based on $D$ instead of $\otimes$

# Quantum information divergences/relative negentropies

Quantum information divergence $D : \mathcal{M}(\mathcal{H}) \times \mathcal{M}(\mathcal{H}) \to [0, \infty]$
s.t. $D(\rho, \sigma) = 0 \iff \rho = \sigma$.

E.g.

- $D_1(\rho, \sigma) := \mathrm{tr}_{\mathcal{H}}(\rho \log \rho - \rho \log \sigma)$ [Umegaki'62]
- $D_{1/2}(\rho, \sigma) := 2\left\| \sqrt{\rho} - \sqrt{\sigma} \right\|^2_{\mathfrak{G}_2(\mathcal{H})} = 4\mathrm{tr}_{\mathcal{H}}(\frac{1}{2}\rho + \frac{1}{2}\sigma - \sqrt{\rho}\sqrt{\sigma})$
  (Hilbert–Schmidt norm²)
- $D_{L_1(\mathcal{N})}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_{\mathcal{T}(\mathcal{H})} = \frac{1}{2}\mathrm{tr}_{\mathcal{H}}|\rho - \sigma|$ ($L_1$/trace norm)
- $D_\gamma(\rho, \sigma) := \frac{1}{\gamma(1-\gamma)}\mathrm{tr}_{\mathcal{H}}(\gamma\rho + (1-\gamma)\sigma - \rho^\gamma\sigma^{1-\gamma})$; $\gamma \in \mathbb{R} \setminus \{0, 1\}$
  [Hasegawa'93]
- $D_{\alpha,z}(\rho, \sigma) := \frac{1}{1-\alpha} \log \mathrm{tr}_{\mathcal{H}}(\rho^{\alpha/z}\sigma^{(1-\alpha)/z})^z$; $\alpha, z \in \mathbb{R}$
  [Audenauert–Datta'14]
- $D_{\mathfrak{f}}(\rho, \sigma) := \mathrm{tr}_{\mathcal{H}}(\sqrt{\rho}\,\mathfrak{f}(\mathfrak{L}_\rho\mathfrak{R}_\sigma^{-1})\sqrt{\rho})$; $\mathfrak{f}$ operator convex, $\mathfrak{f}(1) = 0$
  [Kosaki'82,Petz'85]

for $\mathrm{ran}(\rho) \subseteq \mathrm{ran}(\sigma)$, and with all $D(\rho, \sigma) := +\infty$ otherwise.

# Entropic paradigm: absolute and relative

- Gibbs'1902, Elsasser'37, Jaynes'57, Ingarden–Urbanik'62,...:

  maximisation of **absolute** entropy
  (e.g., $S(\rho) = -D(\rho, \psi)$ with a fixed prior $\psi = \mathbb{I}/\dim \mathcal{H}$)
  as a method of **model construction**: selecting a specific class of models $\mathcal{M}$ that represent the imposed constraints

- Kullback'59, Good'63, Hobson'69,...:

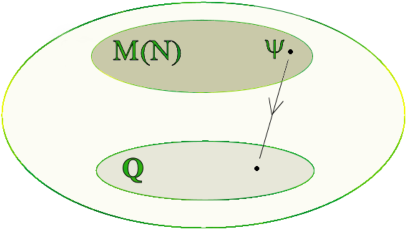
  minimisation of $D(\rho, \psi)$ as a method of state transformation (estimation, learning, updating,...) from $\psi$ onto a set that satisfies given constraints.

# Quantum entropic projections

Let $\mathcal{Q} \subseteq \mathcal{T}(\mathcal{H})^+$ be such that
for each $\psi \in \mathcal{M}(\mathcal{H})$
there exists a unique solution

$$\mathfrak{P}^D_{\mathcal{Q}}(\psi) := \arg\inf_{\rho\in\mathcal{Q}} \{D(\rho,\psi)\}\,.$$

It will be called an entropic projection.




E.g.

- for $D_{1/2}(\rho,\sigma) = 4\mathrm{tr}_{\mathcal{H}}(\frac{1}{2}\rho + \frac{1}{2}\sigma - \sqrt{\rho}\sqrt{\sigma})$,
  and $\mathcal{Q}$ defined as images of closed convex subspaces $\tilde{\mathcal{Q}} \subseteq \mathcal{K}^+ := \mathfrak{G}_2(\mathcal{H})^+$
  under the mapping $\tilde{\mathcal{Q}} \ni \sqrt{\rho} \mapsto (\sqrt{\rho})^2 = \rho \in \mathcal{Q}$
  the entropic projections $\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}$ coincide with the ordinary projection operators on the Hilbert–Schmidt space $\mathfrak{G}_2(\mathcal{H})$.
- for $D_1(\rho,\sigma) = \mathrm{tr}_{\mathcal{H}}(\rho\log\rho - \rho\log\sigma)$
  and $\mathcal{M}(\mathcal{H}) = \mathcal{T}(\mathcal{H})^+_1$, $\psi \in \mathcal{T}(\mathcal{H})^+_1$, $h \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$, then [Araki'77, Donald'90]

$$\exists!\ \psi^h := \arg\inf_{\rho\in\mathcal{T}(\mathcal{H})^+_1} \{D_1(\rho,\psi) + \mathrm{tr}_{\mathcal{H}}(\rho h)\}\,.$$

# Quantum measurement, bayesianity, and maximum relative entropy

- Williams'80, Warmuth'05, Caticha&Giffin'06:

  the Bayes–Laplace rule:

  $$p(x) \mapsto p_{\mathrm{new}}(x) := \frac{p(x)p(b|x)}{p(b)}.$$

  is a special case of

  $$p(\chi) \mapsto p_{\mathrm{new}}(\chi) := \arg\inf_{q \in \mathcal{Q}} \{D_1(q,p)\}\,; \quad D_1(q,p) := \int_{\mathcal{X}} \mu(\chi) q(\chi) \log\left(\frac{q(\chi)}{p(\chi)}\right).$$

- Douven&Romeijn'12: the Bayes–Laplace rule is also a special case of $p \mapsto \arg\inf_{q \in \mathcal{Q}} \{D_1(p,q)\} = \mathfrak{P}^{D_0}_{\mathcal{Q}}(p)$, where $D_0(p,q) = D_1(q,p)$.
- Lüders' rules [Lüders'55]:

  $$\rho \mapsto \rho_{\mathrm{new}} := \sum_i P_i \rho P_i \;\; (\text{'weak'}) \qquad \rho \mapsto \rho_{\mathrm{new}} := \frac{P\rho P}{\mathrm{tr}_{\mathcal{H}}(P\rho)} \;\; (\text{'strong'})$$

- Bub'77'79, Caves–Fuchs–Schack'01, Fuchs'02, Jacobs'02:

  Lüders' rules should be considered as rules of inference (conditioning) that are quantum analogues of the Bayes–Laplace rule

# Quantum bayesian inference from quantum entropic projections

- RPK'13'14, F.Hellmann–W.Kamiński–RPK'14:

1. weak Lüders' rule is a special case of $\rho \mapsto \arg\inf_{\sigma\in\mathcal{Q}}\{D_1(\rho,\sigma)\}$ with
$$\mathcal{Q} = \{\sigma \in \mathcal{T}(\mathcal{H})^+ \mid [P_i,\sigma] = 0\ \forall i\}$$
2. strong Lüders' rule derived from $\rho \mapsto \arg\inf_{\sigma\in\mathcal{Q}}\{D_1(\rho,\sigma)\}$ with
$$\mathcal{Q} = \{\sigma \in \mathcal{T}(\mathcal{H})^+ \mid [P_i,\sigma] = 0,\ \mathrm{tr}_{\mathcal{H}}(\sigma P_i) = p_i\ \forall i\}$$
under the limit $p_2, \ldots, p_n \to 0$.
3. hence, weak and strong Lüders' rules are special cases of quantum entropic projection $\mathfrak{P}^{D_0}_{\mathcal{Q}}$ based on relative entropy $D_0(\sigma,\rho) = D_1(\rho,\sigma)$.

Bayes–Laplace and Lüders' conditionings are special cases of entropic projections
$\Rightarrow$ "quantum bayesianism $\subseteq$ quantum relative entropism".

# Quantum measurements from quantum entropic projections

- Hence: the rule of maximisation of relative entropy (entropic projection on the subspace of constraints) can be considered as a nonlinear generalisation of the dynamics describing *elementary* "quantum measurement".
- F.Hellmann–W.Kamiński–RPK'14: also quantum analogue of Jeffreys' rule follows
- M.Munk–Nielsen'15: partial trace is also entropic projection (for strictly positive states)
- more measurements and more general results: RPK&M.Munk–Nielsen'19 (under construction)
- these results are for $D_0$ and/or $D_1$; however there are many more $D$...
- how general measurements can be derived from entropic projections, allowing both $D$ and $Q$ to vary?

# Generalised pythagorean equation

- The choice of the set $\mathcal{Q}$ for which the entropic projection $\mathfrak{P}^D_{\mathcal{Q}}$ exists and is unique depends very strongly on the structure of $D$: the choice of principle of inference ($D$) determines the accepted data types ($\mathcal{Q}$).
- We need some principle constraining $D$ that would guarantee existence, uniqueness, and good composition properties of $D$-projections.
- We say that $D$ satisfies a generalised pythagorean equation at $\mathcal{Q}$ iff [Chencov'68]

$$D(\phi,\psi) = D(\phi,\mathfrak{P}^D_{\mathcal{Q}}(\psi)) + D(\mathfrak{P}^D_{\mathcal{Q}}(\psi),\psi) \;\; \forall(\phi,\psi)\in\mathcal{Q}\times\mathcal{M}.$$

- Thus, information divergence decomposes additively under a projection onto a suitable subspace, hence we have a nonlinear, yet additive (!), decomposition: **data = signal + noise**
- Example 1: If $\mathcal{Q}$ forms an affine subset of $\mathfrak{G}_2(\mathcal{H})^+$ under $\rho\mapsto\sqrt{\rho}$, then:

$$\left\|x-\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}(z)\right\|^2_{\mathfrak{G}_2(\mathcal{H})} + \left\|\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}(z)-z\right\|^2_{\mathfrak{G}_2(\mathcal{H})} = \|x-z\|^2_{\mathfrak{G}_2(\mathcal{H})}.$$

- Example 2: If $\mathcal{Q} := \{\phi\in\mathfrak{G}_1(\mathcal{H})^+_1 \mid \phi(h) = \mathrm{const}\}$, then [Donald'90]

$$D_1(\phi,\psi^h) + D_1(\psi^h,\psi) = D_1(\phi,\psi) \;\; \forall(\phi,\psi)\in\mathcal{Q}\times\mathfrak{G}_1(\mathcal{H})^+_1.$$

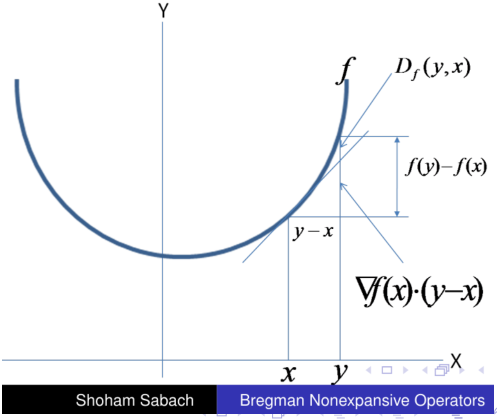
## $D_\Psi$: first idea

Brègman'67:

Let $f : \mathbb{R}^n \to ]-\infty, \infty]$ be convex and proper ($\mathrm{efd}(f) := \{x \in \mathbb{R}^n \mid f(x) \neq \infty\} \neq \varnothing$). Then:

$$D_f(y,x) := f(y) - f(x) - \sum_{i=1}^{n} (y-x)_i [(\mathrm{grad} f)(x)]^i$$

Jones–Byrne'90: $D_f$ is characterised by the generalised pythagorean equation



## $D_\Psi$: reflexive case

Bauschke–Borwein–Combettes'01: Let $X$ be a reflexive Banach space ($X \cong X^{\star\star}$), let $\Psi : X \to ]-\infty, \infty]$ be Legendre ( = convex, proper, lower semi-continuous, Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$, + *some technical conditions*). Then

$$D_\Psi(x, y) := \Psi(x) - \Psi(y) - [\![x - y, \mathfrak{D}^{\mathrm{G}}_y \Psi]\!]_{X \times X^\star}$$

satisfies:

- $D_\Psi(x, y) = 0 \iff x = y$ (information divergence)
- $D_\Psi(x, y) + D_\Psi(y, z) = D_\Psi(x, z) + [\![x - y, \mathfrak{D}^{\mathrm{G}}_z \Psi - \mathfrak{D}^{\mathrm{G}}_y \Psi]\!]_{X \times X^\star}$ (generalised cosine theorem)
- if $C \subseteq X$ is convex and closed then

$$\forall y \in \mathrm{int}(\mathrm{efd}(\Psi) \quad \exists! \mathfrak{P}^{D_\Psi}_C(y) := \arg\inf_{x \in C} \{D_\Psi(x, y)\}$$

- if $C$ is furthermore also an affine subset of $X$ then

$$D_\Psi(x, \mathfrak{P}^{D_\Psi}_C(y)) + D_\Psi(\mathfrak{P}^{D_\Psi}_C(y), y) = D_\Psi(x, y) \quad \forall (x, y) \in C \times X$$

- If $X = \mathcal{H}$ and $\Psi_{1/2} = \frac{1}{2}\|\cdot\|^2_{\mathcal{H}}$ then $\mathfrak{D}^{\mathrm{G}}\Psi_{1/2} = \mathrm{id}_{\mathcal{H}}$ and $D_{\Psi_{1/2}}(x, y) = \frac{1}{2}\|x - y\|^2_{\mathcal{H}}$

## $D_\Psi$: Postquantum Brègman divergences [RPK'17]

- Let $X$ be a reflexive Banach space, $\Psi : X \to ]-\infty, \infty]$ a Legendre function, let $U$ be (a subset of) a positive generating cone of a base norm space $Y$, let $\ell : U \to \ell(U) \subseteq \mathrm{int}(\mathrm{efd}(\Psi)) \subseteq X$ be a bijection.
  We define a postquantum Brègman divergence as:
  $$D_\Psi(\phi, \omega) := \widetilde{D}_\Psi(\ell(\phi), \ell(\omega))$$
  where $\widetilde{D}_\Psi$ is a Brègman divergence on $X$.
- The bijectivity of $\ell$ allows to induce a topology from $X$ onto $U$.
- The existence and uniqueness of the projections onto $\mathcal{Q} \subseteq U$ is guaranteed by requiring $\ell(\mathcal{Q})$ to be convex and closed.
- One can think of $\ell$ as a (nonlinear) coordinate system on $U$, and $X$ as the linear parameter space used for specification of the data required for the entropic projection.
- As a result, all postquantum Brègman divergences satisfy generalised pythagorean theorem for sets that are closed and affine under $\ell$-embeddings.
- If $Y$ is given by a self-adjoint part of a predual of $W^*$-algebra, then $D_\Psi$ is a quantum Brègman divergence.

# $D_\Psi$: Postquantum Brègman divergences

Examples:

(1) [Jenčová'03/'05, Ojima'03/04]:

- $U = \mathcal{N}_\star^+$ for a W*-algebra $\mathcal{N}$
- $X = L_{1/\gamma}(\mathcal{N}, \psi)$: noncommutative $L_p$ space w.r.t. f.n.s. weight $\psi$ on $\mathcal{N}$, $p = \frac{1}{\gamma} \in ]1, \infty[$
- $\ell_\gamma(\phi) = \frac{1}{\gamma}\Delta_{\phi,\psi}$, $\Psi_\gamma(x) = \frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma}$
- $D_\gamma(\omega, \phi) = \frac{1}{1-\gamma}\omega(\mathbb{I}) + \frac{1}{\gamma}\phi(\mathbb{I}) + \frac{1}{\gamma(1-\gamma)}\left[\left[\Delta^\gamma_{\omega,\psi}, \Delta^{1-\gamma}_{\omega,\psi}\right]\right]$
- RPK'11: Generalisation to a construction that is explicitly independent of any $\psi$
- For $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\psi = \mathrm{tr}_\mathcal{H}$, $U = \{\rho \in \mathcal{T}(\mathcal{H}) \mid \mathrm{tr}_\mathcal{H}(\rho\mathbb{I}) = 1\}$, $D_\gamma(\omega, \phi) = \frac{1}{\gamma(1-\gamma)}\mathrm{tr}_\mathcal{H}(\rho_\omega - \rho_\omega^\gamma \rho_\phi^{1-\gamma})$ [Hasegawa'93]

(2) [RPK'17]: a generalisation of (1) to nonassociative $L_p(A, \tau)$ spaces over JBW-algebras $A$ with f.n.s. trace $\tau$

(3) $\lim_{\gamma\to^+1} D_\gamma(\omega, \phi) = D_1(\omega, \phi) =$ denormalised Araki divergence. In finite-dimensional case it is a quantum Brègman divergence. In general, is not a quantum Brègman divergence, yet it is a limit of a family of such divergences. It is not associated naturally with any reflexive Banach space, although it satisfies one-sided generalised cosine and pythagorean theorems.

# Markovian monotonicity

Usual markovian/CPTP setting:

- $T : \mathcal{T}(\mathcal{H})^+ \to \mathcal{T}(\mathcal{H})^+$ is a CPTP map
- information processing inequality
  Petz'85, Tomamichel–Colbeck–Renner'09:

$$D_{\mathfrak{f}}(T(\phi), T(\omega)) \leq D_{\mathfrak{f}}(\phi, \omega) \;\; \forall \phi, \omega \in \mathcal{T}(\mathcal{H})^+ \; \forall T$$

- $D_{\mathfrak{f}}(\rho, \sigma) = \mathrm{tr}_{\mathcal{H}}(\rho^{1/2} \mathfrak{f}(\mathfrak{L}_{\rho} \mathfrak{R}_{\sigma^{-1}}) \rho^{1/2})$ [Kosaki'82, Petz'85'86]
- $\mathfrak{f} : \mathbb{R}^+ \to \mathbb{R}$ is operator convex,

$$\mathfrak{f}(\lambda x + (1 - \lambda) y) \leq \lambda \mathfrak{f}(x) + (1 - \lambda) \mathfrak{f}(y) \;\; \forall x, y \in \mathfrak{B}(\mathcal{H})^+ \;\; \forall \lambda \in [0, 1],$$

  $\mathfrak{f}(0) \leq 0$, $\mathfrak{f}(1) = 0$.

Noticeable facts:

- $D_{\Psi} \cap D_{\mathfrak{f}} = D_{\gamma}$ (characterisation, under some conditions: [Amari'09, RPK'19])
- In quantum resource theories one often considers free operations $\Lambda$ determined as a submonoid of CPTP maps. Hence, there is in principle a larger class $D_{\Lambda}$ that is $\Lambda$-monotone, and a larger space of possible families characterised by $D_{\Psi} \cap D_{\Lambda}$

## Brègman nonexpansive operations

- Let $M \subseteq \mathrm{int}(\mathrm{efd}(\Psi))$ then $T :\to \mathrm{int}(\mathrm{efd}(\Psi)$ is called:
  - (i) Brègman completely nonexpansive $\iff$
  $$\widetilde{D}_\Psi(T(x), T(y)) \leq \widetilde{D}_\Psi(x, y) \;\; \forall x, y \in M$$
  - (ii) Brègman quasi-nonexpansive $\iff$
  $$\widetilde{D}_\Psi(x, T(y)) \leq \widetilde{D}_\Psi(x, y) \;\; \forall (x, y) \in \mathrm{Fix}(T) \times M,$$
  where $\mathrm{Fix}(T) := \{x \in M \mid T(x) = x\} \neq \varnothing$
  - (iii) Brègman firmly quasi-nonexpansive $\iff$
  $$\widetilde{D}_\Psi(p, T(x)) + \widetilde{D}_\Psi(T(x), x) \leq \widetilde{D}_\Psi(p, x) \;\; \forall (x, p) \in K \times \mathrm{Fix}(T)$$
  - (iv) Brègman firmly nonexpansive $\iff$
  $\widetilde{D}_\Psi(T(x), T(y)) + \widetilde{D}_\Psi(T(y), T(x)) + \widetilde{D}_\Psi(T(x), x) + \widetilde{D}_\Psi(T(y), y) \leq$
  $\widetilde{D}_\Psi(T(x), y) + \widetilde{D}_\Psi(T(y), x)$
- (iv) $\Rightarrow$ (iii) $\Rightarrow$ (ii)
- For closed convex C, $\mathfrak{P}_C^{\widetilde{D}_\Psi}$ satisfy (iv)
- Given $\widetilde{\ell} : U \to X$ and $T : \widetilde{\ell}(U) \to \widetilde{\ell}(U)$, an $\widetilde{\ell}$-operation is defined as $\widetilde{\ell}^{-1} \circ T \circ \widetilde{\ell}$
- By composing $\widetilde{\ell}$ with one of $T$ as above, we obtain Brègman nonexpansive $\widetilde{\ell}$-operations
- This is an example of what Mielnik ['69, '74] calls a nonlinear transmitter
- $\mathfrak{P}_C^{D_\Psi}$ is a special case of it

# Brègman nonexpansive nonlinear resource theories

- Let $U$ = a state space
  (e.g., a positive generating cone or a base of a base norm space)
- A resource theory of states: a triple $(P, Q, R)$, where
  - $P :=$ a submonoid of linear endomorphisms = free operations
  - $Q := \{\phi \in U \mid \forall\psi \in U \;\; \exists p \in P \;\; p(\psi) = \phi\}$ = set of free states
  - $R := \{r : U \to \mathbb{R}^+ \mid (r \circ p)(\phi) \leq r(\phi) \;\; \forall\phi \in U\}$ = resource monotones
- In principle, we can drop down the assumption of linearity.
- If $\mathcal{T}$ is a submonoid of Brègman completely nonexpansive $\widetilde{\ell}$-operations
  such that $Q_{\mathcal{T}} := \{\phi \in U \mid \forall\psi \in U \;\; \exists t \in \mathcal{T} \;\; t(\psi) = \phi\} \neq \varnothing$ is $\widetilde{\ell}$-closed $\widetilde{\ell}$-convex
  then $D_{\mathcal{T}} := \inf_{\phi \in Q_{\mathcal{T}}}\{D_\Psi(\phi, \cdot\,)\}$ is a resource monotone
  and $(\mathcal{T}, Q_{\mathcal{T}}, \{D_{\mathcal{T}}\})$ is an example of a nonlinear resource theory
- Other brègmannian resource theories are also possible.

# Smooth quantum information geometries

Under some conditions, $D$ induces a generalisation of smooth riemannian geometry on $\mathcal{M}(\mathcal{N})$.

- Jenčová'05: a general construction of smooth manifold structure on the space of all strictly positive states over arbitrary $W^*$-algebra.
- E.g. $\mathcal{M}(\mathcal{H}) := \{\rho(\theta) \in \mathcal{T}(\mathcal{H}) \mid \rho(\theta) > 0, \theta \in \Theta \subseteq \mathbb{R}^n \text{ open}, \theta \mapsto \rho(\theta) \text{ smooth}\}$
- Eguchi'83/Ingarden et al'82/Lesniewski–Ruskai'99/Jenčová'04:
  Every smooth divergence $D$ with positive definite hessian determines a riemannian metric $\mathbf{g}^D$ and a pair $(\nabla^D, \nabla^{D\dagger})$ of torsion-free affine connections:

$$\mathbf{g}_\phi(u,v) := -\partial_{u|\phi}\partial_{v|\omega} D(\phi,\omega)|_{\omega=\phi},$$
$$\mathbf{g}_\phi((\nabla_u)_\phi v, w) := -\partial_{u|\phi}\partial_{v|\phi}\partial_{w|\omega} D(\phi,\omega)|_{\omega=\phi},$$
$$\mathbf{g}_\phi(v, (\nabla^\dagger_u)_\phi w) := -\partial_{u|\omega}\partial_{w|\omega}\partial_{v|\phi} D(\phi,\omega)|_{\omega=\phi},$$

  which satisfy the characteristic equation of the Norden['37]–Sen['44] geometry,

$$\mathbf{g}^D(u,v) = \mathbf{g}^D(\mathbf{t}_c^{\nabla^D}(u), \mathbf{t}_c^{\nabla^{D\dagger}}(v)) \;\; \forall u,v \in \mathbf{T}\mathcal{M}(\mathcal{N}).$$

- A riemannian geometry $(\mathcal{M}(\mathcal{N}), \mathbf{g}^D)$ has Levi-Civita connection $\bar{\nabla} = (\nabla^D + \nabla^{D\dagger})/2$.
- Example 1: $\mathcal{M}(\mathcal{N}) = \mathcal{T}(\mathcal{H}) \cap \{\rho > 0, \mathrm{tr}_\mathcal{H}(\rho) = 1\}$ and $D_1(\rho,\sigma) = \mathrm{tr}_\mathcal{H}(\rho\log\rho - \rho\log\sigma)$ give Mori['55]–Kubo['56]–Bogolyubov['62] $\mathbf{g}^{D_1}$ and Nagaoka['94]–Hasegawa['95] $(\nabla^{D_1}, \nabla^{D_1\dagger})$:

$$\mathbf{g}^{D_1}_\rho(x,y) = \mathrm{tr}_\mathcal{H}\left(\int_0^\infty \mathrm{d}\lambda\, x\frac{1}{\lambda\mathbb{I}+\rho} y \frac{1}{\lambda\mathbb{I}+\rho}\right), \quad \mathbf{t}^{\nabla^{D_1}}_{\rho,\omega}(x) = x - \mathrm{tr}_\mathcal{H}(\omega x), \quad \mathbf{t}^{\nabla^{D_1\dagger}}_{\rho,\omega}(x) = x.$$

# Smooth quantum information geometries

Taylor expansion of $D$ induces a generalisation of a smooth riemannian geometry on $\mathcal{M}(\mathcal{N})$.

- $\mathcal{M}(\mathcal{H}) := \{\rho(\theta) \in \mathcal{T}(\mathcal{H}) \mid \rho(\theta) > 0, \theta \in \Theta \subseteq \mathbb{R}^n \text{ open}, \theta \mapsto \rho(\theta) \text{ smooth}\}$ is a $C^\infty$-manifold
- Jenčová'05: a general construction of smooth manifold structure on the space of all strictly positive states over arbitrary $W^*$-algebra, with tangent spaces given by noncommutative Orlicz spaces.
- Eguchi'83/Ingarden et al'82/Lesniewski–Ruskai'99/Jenčová'04:
  Every smooth divergence $D$ with positive definite hessian determines
  a riemannian metric $\mathbf{g}^D$ and a pair $(\nabla^D, \nabla^{D^\dagger})$ of torsion-free affine connections:

$$\mathbf{g}_\phi(u,v) := -\partial_{u|\phi}\partial_{v|\omega} D(\phi,\omega)|_{\omega=\phi},$$
$$\mathbf{g}_\phi((\nabla_u)_\phi v, w) := -\partial_{u|\phi}\partial_{v|\phi}\partial_{w|\omega} D(\phi,\omega)|_{\omega=\phi},$$
$$\mathbf{g}_\phi(v, (\nabla^\dagger_u)_\phi w) := -\partial_{u|\omega}\partial_{w|\omega}\partial_{v|\phi} D(\phi,\omega)|_{\omega=\phi},$$

  which satisfy the characteristic equation of the Norden['37]–Sen['44] geometry,

$$\mathbf{g}^D(u,v) = \mathbf{g}^D(\mathbf{t}_c^{\nabla^D}(u), \mathbf{t}_c^{\nabla^{D^\dagger}}(v)) \;\; \forall u,v \in \mathbf{T}\mathcal{M}(\mathcal{N}).$$

- A riemannian geometry $(\mathcal{M}(\mathcal{N}), \mathbf{g}^D)$ has Levi-Civita connection $\bar{\nabla} = (\nabla^D + \nabla^{D^\dagger})/2$.

# Example

- $\mathcal{M}(\mathcal{N}) = \mathcal{T}(\mathcal{H}) \cap \{\rho > 0, \mathrm{tr}_{\mathcal{H}}(\rho) = 1\}$
  $D_1(\rho, \sigma) = \mathrm{tr}_{\mathcal{H}}(\rho \log \rho - \rho \log \sigma)$

  give Mori['55]–Kubo['56]–Bogolyubov['62] riemannian metric:

$$\mathbf{g}^{D_1}_{\rho}(x, y) = \mathrm{tr}_{\mathcal{H}} \left( \int_0^\infty \mathrm{d}\lambda x \frac{1}{\lambda \mathbb{I} + \rho} y \frac{1}{\lambda \mathbb{I} + \rho} \right),$$

  and Nagaoka['94]–Hasegawa['95] affine connections:

$$\mathbf{t}^{\nabla^{D_1}}_{\rho,\omega}(x) = x - \mathrm{tr}_{\mathcal{H}}(\omega x), \;\; \mathbf{t}^{\nabla^{D_1 \dagger}}_{\rho,\omega}(x) = x.$$

# Hessian geometries = dually flat Norden–Sen geometries

If $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ is a Norden–Sen geometry with flat $\nabla$ and $\nabla^\dagger$, then:

1. there exists a unique pair of functions $\Phi : \mathcal{M} \to \mathbb{R}$, $\Phi^{\mathbf{L}} : \mathcal{M} \to \mathbb{R}$ such that $\mathbf{g}$ is their hessian metric,

$$\mathbf{g}(\rho) = \sum_{i,j} \frac{\partial^2 \Phi(\rho(\theta))}{\partial \theta^i \partial \theta^j} \mathrm{d}\theta^i \otimes \mathrm{d}\theta^j,$$

$$\mathbf{g}(\rho) = \sum_{i,j} \frac{\partial^2 \Phi^{\mathbf{L}}(\rho(\eta))}{\partial \eta^i \partial \eta^j} \mathrm{d}\eta^i \otimes \mathrm{d}\eta^j,$$

   where: $\{\theta^i\}$ is a coordinate system s.t. $\Gamma^{\nabla}_{ijk}(\rho(\theta)) = 0 \ \forall \rho \in \mathcal{M}$,
   $\{\eta^i\}$ is a coordinate system s.t. $\Gamma^{\nabla^\dagger}_{ijk}(\rho(\eta)) = 0 \ \forall \rho \in \mathcal{M}$,
   and $\Phi^{\mathbf{L}}$ is a Fenchel conjugate of $\Phi$.

2. the Eguchi equations applied to the Brègman divergence

$$D_\Phi(\rho, \sigma) := \Phi(\rho) + \Phi^{\mathbf{L}}(\sigma) - \sum_i \theta^i(\rho)\eta^i(\sigma)$$

   yield $(\mathbf{g}, \nabla, \nabla^\dagger)$ above.

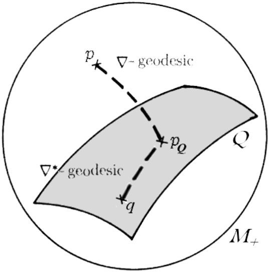
# Smooth generalised pythagorean theorem

Let $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ be a hessian geometry. Then for any $\mathcal{Q} \subseteq \mathcal{M}$ which is:

- $\nabla^\dagger$-autoparallel $:= \nabla^\dagger_u v \in \mathbf{T}\mathcal{Q}\ \forall u, v \in \mathbf{T}\mathcal{Q}$;
- $\nabla^\dagger$-convex $:= \forall \rho_1, \rho_2 \in \mathcal{Q}\ \exists!\ \nabla^\dagger$-geodesics in $\mathcal{Q}$ connecting $\rho_1$ and $\rho_2$;

there exists a unique projection

$$\mathcal{M} \ni \rho \mapsto \mathfrak{P}^{D_\Phi}_{\mathcal{Q}}(\rho) := \arg\inf_{\sigma \in \mathcal{Q}} \{D_\Phi(\sigma, \rho)\} \in \mathcal{Q}.$$




- it is equal to a unique projection of $\rho$ onto $\mathcal{Q}$ along a $\nabla$-geodesic that is $\mathbf{g}$-orthogonal at $\mathcal{Q}$.
- it satisfies a generalised pythagorean equation

$$D_\Phi(\omega, \mathfrak{P}^{D_\Phi}_{\mathcal{Q}}(\rho)) + D_\Phi(\mathfrak{P}^{D_\Phi}_{\mathcal{Q}}(\rho), \rho) = D_\Phi(\omega, \rho) \quad \forall(\omega, \rho) \in \mathcal{Q} \times \mathcal{M}.$$

**Hence, for Brègman divergences $D_\Phi$ the local entropic projections are equivalent with geodesic projections.**

# Mitchell['67]–Jaynes['83] source theory

Consider probabilistic model $\mathcal{M}_3 := \{p(x|\lambda_A, \lambda_B, \lambda_C) = \frac{1}{Z}e^{-\lambda_A A(x) - \lambda_B B(x) - \lambda_C C(x)}\}$, equipped with $((\lambda_i), (\langle A\rangle, \langle B\rangle, \langle C\rangle)) : \mathcal{M}_3 \times \mathcal{M}_3 \to \mathbb{R}^3 \times \mathbb{R}^3$. Consider a source-and-response problem with an additional control variable:

$$\begin{aligned} &\delta\langle A\rangle = 0, \quad \delta\lambda_A \neq 0 \quad \text{'driving variable' (source parameter)} \\ &\delta\langle B\rangle \neq 0, \quad \delta\lambda_B = 0 \quad \text{'heat bath' (response parameter)} \\ &\delta\langle C\rangle = 0, \quad \delta\lambda_C \neq 0 \quad \text{'control variable' (additional source)} \end{aligned}$$

Question: how the presence of the second source affects the relationship between first source and the response parameter? Answer: the correlation matrix

$$\begin{pmatrix} \delta\langle A\rangle \\ \delta\langle B\rangle \\ \delta\langle C\rangle \end{pmatrix} = -\begin{pmatrix} K_{AA} & K_{AB} & K_{AC} \\ K_{BA} & K_{BB} & K_{BC} \\ K_{CA} & K_{CB} & K_{CC} \end{pmatrix} \begin{pmatrix} \delta\lambda_A \\ \delta\lambda_B \\ \delta\lambda_C \end{pmatrix}. \tag{1}$$

can be rewritten as

$$\begin{pmatrix} \delta\langle A\rangle \\ \delta\langle B\rangle \end{pmatrix} = -\begin{pmatrix} K'_{AA} & K'_{AB} \\ K'_{BA} & K'_{BB} \end{pmatrix} \begin{pmatrix} \delta\lambda_A \\ \delta\lambda_B \end{pmatrix}, \tag{2}$$

where $K'_{AB} := K_{AB} - K_{AC}K^{-1}_{CC}K_{CB}$. For $R_{AC} := K_{AC}(K_{AA}K_{CC})^{-1/2}$, one has

$$\delta\langle B\rangle = \frac{K'_{BA}}{K'_{AA}} \frac{1}{1 - R^2_{AC}} \delta\langle A\rangle \tag{3}$$

Thus, one can eliminate control parameter from the dynamics at the price of source term renormalisation (coarse-graining of (1) to (2) and rescaling (3)).

# Favretti's['07] geometrisation

- Let $(\mathcal{M}, \mathbf{g}^{\Psi}, \nabla^{\Psi}, \nabla^{\Psi\dagger})$ be a hessian geometry.
- Let $\dim \mathcal{M} = n + m$ with associated geodesic coordinates $\eta(p) = (\eta_A, \eta_B, \eta_C) \in \mathbb{R}^k \times \mathbb{R}^{n-k} \times \mathbb{R}^m$, $\theta(p) = (\theta^A, \theta^B, \theta^C) \in \mathbb{R}^k \times \mathbb{R}^{n-k} \times \mathbb{R}^m$, related by the Legendre transform: $\theta = (\mathrm{grad}(\Psi))^{-1}(\eta)$, $\eta = (\mathrm{grad}(\Psi))(\theta)$.
- Let the constraints of relative entropic information dynamics $p(t) = \mathfrak{P}^{D_\Psi}_{\mathcal{Q}(t)}(p_0)$ be given by

$$\eta_A(t) = \bar{\eta}_A, \quad \theta^B(t) = \bar{\theta}^B, \quad \eta_C(t) = \bar{\eta}^C.$$

- Using implicit function theorem and the properties of dually flat Norden–Sen geometry, Favretti has shown that

$$\mathrm{d}\eta_B = \bar{\mathbf{g}}^{\Psi}_{BA} \frac{1}{\mathbb{I} - R^2_{AC}} (\bar{\mathbf{g}}^{\Psi}_{AA})^{-1} \mathrm{d}\eta_A, \qquad (4)$$

  where $R^2_{AC} := (\mathbf{g}^{\Psi}_{AA})^{-1} \mathbf{g}^{\Psi}_{AC} (\mathbf{g}^{\Psi}_{CC})^{-1} \mathbf{g}^{\Psi}_{CA}$.
- If $R_{AC}$ has a spectral radius smaller than 1, then one can expand the renormalised expression (4) in terms of corrections that come from the interaction with the additional source:

$$\mathrm{d}\eta_B = (\Gamma_{AB} - \Gamma_{BC}\Gamma_{CA} + \Gamma_{BA}\Gamma_{AC}\Gamma_{CA} - \ldots)\mathrm{d}\eta_A,$$

  where $\Gamma_{ij} := \mathbf{g}^{\Psi}_{ij} (\mathbf{g}^{\Psi}_{jj})^{-1}$.

## Summary

1. Entropic projections as nonlinear state transformations
2. Brègman family of relative entropies
3. Brègman nonexpansive operations
4. Nonlinear resource theories based on Brègman nonexpansive operations
5. Smooth quantum information geometries derived from Brègman divergences
6. Jaynes–Mitchell–Favretti renormalisation

# References


- RPK, 2017, Postquantum Brègman relative entropies and nonlinear resource theories, arXiv:1710.01837
- RPK, 2016, Local quantum information dynamics, arXiv:1605.02063
- M.I.Munk-Nielsen, 2015, Quantum measurements from entropic projections, MSc Thesis, Perimeter Institute, Waterloo, Canada.
- RPK, 2014, Lüders' and quantum Jeffrey's rules as entropic projections, arXiv:1408.3502.
- F. Hellmann, W. Kamiński, RPK, 2014, Quantum collapse rules from maximum relative entropy principle, arXiv:1407.7766.
- RPK, 2013, Information geometric foundations of quantum theory, Ph.D. thesis, Institute for Theoretical Physics, University of Warsaw.
- RPK, 2011, Information dynamics and new geometric foundations for quantum theory, Växjö 2011 Proceedings, AIP Conf. Proc. **1424**, 200. arXiv:1110.4492.
- RPK, 2011, The general form of $\gamma$-family of quantum relative entropies, Open Sys. Inf. Dyn. **18**, 191. arXiv:1106.2225.
- RPK, 2010, Quantum theory as inductive inference, MAXENT 2010 Proceedings, AIP Conf. Proc. **1305**, 28. arXiv:1009.2423.