

Quantum Voronoi Diagrams and Holevo Channel Capacity for 1-Qubit Quantum States

Frank Nielsen
 Sony Computer Science Laboratories, Inc.
 École Polytechnique, LIX
 Palaiseau, France
 Email: nielsen@lix.polytechnique.fr

Richard Nock
 CEREGMIA
 Antilles-Guyane University
 Martinique, France
 Email: rnock@martinique.univ-ag.fr

Abstract—In this paper, we first introduce a smooth parametric family of Bregman-Csiszár quantum entropies including the von Neumann and Burg quantum entropies. We then describe the dualistic nature of Voronoi diagrams for 1-qubit quantum states inside the 3D Bloch ball representation. We show that these diagrams can be computed as Bregman Voronoi diagrams for the corresponding Bregman generator acting on Hermitian density matrices. This implies that these dual diagrams can be derived from power diagrams of balls in the Laguerre geometry, and allows one to prove by equivalence that the von Neumann quantum Voronoi diagram on the degenerated Bloch sphere of pure quantum states coincides with the ordinary Euclidean Voronoi diagram, bypassing the fact that the quantum divergence is not defined there. We then show how to compute the Holevo channel capacity of 1-qubit quantum states, and provide a practical approximation algorithm based on Bregman core-sets. Finally, we define the quantum sided centroids that yield practical upper bounds on the Holevo capacity in linear time.

I. INTRODUCTION AND PRELIMINARIES

The 21st century attests the accelerated rise of the deployment of quantum mechanics into various industrial prototypes like the prominent quantum cryptographic systems. Recent breakthroughs in experimental physics bridged the gap between mathematical theory and practice, and the analysis of quantum channel characteristics such as its *capacity* become a fundamental problem associated with related open problems.¹

In quantum information theory [1], particle state distributions are analyzed probabilistically by means of *density* matrices \mathbf{X} . A d -level system is characterized by a $d \times d$ matrix $\mathbf{X} \in \mathbb{C}_{d \times d}$ with complex coefficients that satisfies the following three properties:

- 1) \mathbf{X} is Hermitian. That is, \mathbf{X} is equal to its *conjugate transpose*: $\mathbf{X} = \mathbf{X}^{*T}$,
- 2) \mathbf{X} has *unit trace*. That is, the sum of diagonal elements sums up to one and has no imaginary part:
 $\text{Tr}(\mathbf{X}) = \sum_{i=1}^d X_{i,i} = 1$,
- 3) \mathbf{X} is *semi-positive definite*. That is, \mathbf{X} belongs to the positive cone: $\mathbf{x}^T \mathbf{X} \mathbf{x} \geq 0 \forall \mathbf{x} \neq \mathbf{0}$. This condition implies positive determinant: $\det \mathbf{X} \geq 0$.

Let $\mathcal{S}(\mathbb{C}^d)$ denote the space of such semi-positive definite density square matrices of size $d \times d$. One qubit (quantum bit

for short) systems² are the simplest fundamental case, obtained for $d = 2$. The above three conditions imply the following characterization of the family of 2×2 complex matrices [2]:

$$\mathbf{X} = \left\{ \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \mid x^2 + y^2 + z^2 \leq 1 \right\}, \quad (1)$$

where i denotes the imaginary number $i^2 = -1$. The condition $x^2 + y^2 + z^2 \leq 1$ is derived from the semi-positive definiteness assumption ($\det \mathbf{X} \geq 0$). Thus 1-qubit states \mathbf{X} can be represented equivalently by a $\mathbf{x} = (x, y, z)$ triple of reals, a 3D point $\mathbf{x} \in \mathbb{R}^3$, and the set $\mathcal{S}(\mathbb{C}^2)$ of 1-qubits is referred to as the *Bloch ball*³. We distinguish between *pure* states which have degenerated density matrices of rank 1 (noninvertible matrix), and *mixed* states of full rank 2. The pure state condition is geometrically visualized by density matrices lying on the boundary of the Bloch ball: The *Bloch sphere*. The state \mathbf{X} of a 1-qubit is expressed using three reals $\mathbf{x} = (x, y, z)$ that can be reinterpreted in spherical coordinates as $\mathbf{x} = (r, \theta, \phi)$ where r denotes the *radius* of the state to the origin, and θ and ϕ encode the *latitude* and *longitude* rotation angles: $(r, \theta, \phi) \leftrightarrow (x = r \sin \theta \cos \phi, y = r \sin \theta \sin \phi, z = r \cos \theta)$. In order to define the *quantum divergence* that is the generalization of the Kullback-Leibler divergence (better known as relative entropy [1]) to density matrices, we first define the *logarithm* of a density matrix using its *spectral decomposition*. Consider the singular value decomposition (SVD) of a Hermitian matrix \mathbf{X} : $\mathbf{X} = \mathbf{V} \text{Diag}(\boldsymbol{\lambda}) \mathbf{V}^*$, with both \mathbf{V} and $\mathbf{V}^* = \mathbf{V}^T$ unitary orthonormal matrices, and all eigenvalues $\lambda_i \geq 0$ real and positive. The diagonal matrix represents the *eigenspectra* and the complex orthonormal rotation matrix \mathbf{V} the associated *eigenspace*. Using the spectrum decomposition of a matrix \mathbf{X} , we define the logarithm of a density matrix as: $\log \mathbf{X} = \mathbf{V} \text{Diag}(\log \lambda_1, \dots, \log \lambda_d) \mathbf{V}^*$. The *quantum von Neuman entropy* $H(\mathbf{X})$ (matrix entropy) is a generalization of the classical *Shannon entropy*⁴ to density matrices:

²In general, n -qubit systems require dimension $d = 2^n$.

³Named after physicist Felix Bloch, first director of the CERN institute.

⁴The Shannon entropy of a discrete d -dimensional distribution \mathbf{p} is defined as $H(\mathbf{p}) = -\sum_{i=1}^d p_i \log \frac{1}{p_i} = -\sum_{i=1}^d p_i \log p_i$. The distance between “classical bits” is the Hamming distance.

¹See the quantum information problems list at:
<http://www.imaph.tu-bs.de/qi/problems/>

$$H(\mathbf{X}) = -\text{Tr}(\mathbf{X} \log \mathbf{X}). \quad (2)$$

It can be shown that the quantum entropy is equal to the Shannon entropy for the eigenvalue distribution:

$$H(\mathbf{X}) = H(\boldsymbol{\lambda}) = -\sum_{i=1}^d \lambda_i \log \lambda_i. \quad (3)$$

The quantum information divergence I (matrix relative entropy) generalizes the Kullback-Leibler divergence [18] (KL) by considering the following distortion measure:

$$I(\mathbf{P}||\mathbf{Q}) = \text{Tr}(\mathbf{P}(\log \mathbf{P} - \log \mathbf{Q})) \geq 0. \quad (4)$$

This divergence is *not* symmetric nor does it satisfy the triangle inequality. It is therefore not a metric. Note that the quantum divergence is defined for $\mathbf{P} \rightarrow \mathbf{0}$ by taking the limit: $\lim_{\mathbf{X} \rightarrow \mathbf{0}} \text{Tr}(\mathbf{X} \log \mathbf{X}) = 0$ (since $\lim_{x \rightarrow 0} x \log x = 0$). However, the divergence is *not* properly defined when \mathbf{Q} is *not* full rank (i.e., \mathbf{Q} encodes a pure state) because of the undefined logarithm. The quantum information divergence is *reflexive*: $I(\mathbf{P}||\mathbf{Q}) = 0 \Leftrightarrow \mathbf{P} = \mathbf{Q}$. We further have the following quantum/classical information inequality:

$$I(\mathbf{P}||\mathbf{Q}) \geq \text{KL}(\boldsymbol{\lambda}_{\mathbf{P}}||\boldsymbol{\lambda}_{\mathbf{Q}}) \geq 0, \quad (5)$$

where $\boldsymbol{\lambda}_{\mathbf{P}}$ and $\boldsymbol{\lambda}_{\mathbf{Q}}$ are the eigenvalue distributions of the spectral decomposition of the density matrices \mathbf{P} and \mathbf{Q} , respectively. The inequality is *strict* if and only if the eigenspaces of \mathbf{P} and \mathbf{Q} differ. Interestingly, this von Neumann quantum information divergence belongs to a broader class of parametric divergences called Bregman divergences [3]. Bregman divergences are parameterized families of distortion measures induced by a strictly convex and differentiable convex function $F : \mathcal{S}(\mathbb{C}^d) \rightarrow \mathbb{R}$ such that:

$$D_F(\mathbf{P}||\mathbf{Q}) = F(\mathbf{P}) - F(\mathbf{Q}) - \langle \mathbf{P} - \mathbf{Q}, \nabla F(\mathbf{Q}) \rangle \quad (6)$$

where the real inner product is defined as: $\langle \mathbf{P}, \mathbf{Q} \rangle = \text{Tr}(\mathbf{P}\mathbf{Q}^*)$, and $\nabla F(\cdot)$ is the Gâteaux derivative: Namely, the gradient. We have $D_F(\mathbf{P}||\mathbf{Q}) = 0$ if and only if $\mathbf{P} = \mathbf{Q}$ (positive-definiteness generalizing Gibb's inequality [1], [8]). Bregman divergences can also be interpreted *locally* as quadratic distance measures by considering the Taylor expansion of F with an *exact remainder* term: $D_F(\mathbf{P}||\mathbf{Q}) = (\mathbf{P} - \mathbf{Q})^* \frac{\nabla F^2(\boldsymbol{\varepsilon})}{2} (\mathbf{P} - \mathbf{Q})$, where $\boldsymbol{\varepsilon}$ depends on both \mathbf{P} and \mathbf{Q} . The quantum information divergence is a Bregman divergence obtained for the Bregman generator $F(\mathbf{X}) = \text{Tr}(\mathbf{X} \log \mathbf{X})$. Dhillon and Tropp [4] thoroughly investigated Bregman matrix distortion measures for "matrix nearness" decompositions with a special care given to the squared Fröbenius, von Neumann information and the log det divergences obtained respectively for the generators $F(\mathbf{X}) = \frac{1}{2} \|\mathbf{X}\|^2$, $F(\mathbf{X}) = \text{Tr}(\mathbf{X} \log \mathbf{X})$ and $F(\mathbf{X}) = -\log \det \mathbf{X}$. It follows that the quantum relative entropy $I = D_F$ (von Neumann Bregman divergence) has thus a neat axiomatic characterization [5], and can further

be *extended* following Csiszár least square projection characterization [5], by using for the Bregman generator the *extended negative entropy* $F(\mathbf{X}) = \text{Tr}(\mathbf{X} \log \mathbf{X} - \mathbf{X})$. The gradient of $F(\mathbf{X})$ is $\nabla F(\mathbf{X}) = \log \mathbf{X}$, the quantum Burg entropy. In summary, the *extended* von Neumann quantum divergence is a Bregman divergence in disguise for generator $F(\mathbf{X}) = \text{Tr}(\mathbf{X} \log \mathbf{X} - \mathbf{X})$:

$$I(\mathbf{P}||\mathbf{Q}) = D_F(\mathbf{P}||\mathbf{Q}) = \text{Tr}(\mathbf{P}(\log \mathbf{P} - \log \mathbf{Q}) - \mathbf{P} + \mathbf{Q}). \quad (7)$$

Bregman divergences are invariant by affine terms and enjoy a remarkable bijection with probability distributions of the statistical exponential families [6]. Carrying out the calculations for the 3D Bloch ball of 1-qubits \mathbf{P} and \mathbf{Q} (with r_P and r_Q denoting the respective radii in the 3D Bloch ball representation), we obtain [8]:

$$I(\mathbf{P}||\mathbf{Q}) = \frac{1+r_P}{2} \log \frac{1+r_P}{2} + \frac{1-r_P}{2} \log \frac{1-r_P}{2} - \frac{1}{2} \log \frac{1-r_Q^2}{4} - \frac{\langle \mathbf{p}, \mathbf{q} \rangle}{2r_Q} \log \frac{1+r_Q}{1-r_Q} \quad (8)$$

$$\stackrel{\text{def}}{=} a(r_P) + b(r_Q) - c(r_Q) \langle \mathbf{p}, \mathbf{q} \rangle. \quad (9)$$

II. QUANTUM BREGMAN-CSISZÁR DIVERGENCES

The choice of the proper quantum divergence may depend upon the situation implied by the underlying study or application needs [1]. It is therefore interesting to design a *flexible generic* quantum divergence by generalizing parametric divergences proposed in classical information theory. We propose to extend the von Neumann quantum/Log det divergences [4] to the class of *quantum Bregman-Csiszár divergences* for density matrices. First, define the smooth family of strictly convex and differentiable Bregman generators F_α on density matrices for a single parameter $\alpha \in [0, 1]$ as:

$$F_\alpha(\mathbf{X}) = \frac{1}{\alpha(1-\alpha)} \text{Tr}(-\mathbf{X}^\alpha + \alpha \mathbf{X} - \alpha \mathbf{I} + \mathbf{I}), \quad (10)$$

where \mathbf{I} denotes the identity matrix, and \mathbf{X}^α is defined from the spectral decomposition of $\mathbf{X} = \mathbf{V} \text{Diag}(\boldsymbol{\lambda}) \mathbf{V}^*$ as $\mathbf{X}^\alpha = \mathbf{V} \text{Diag}(\lambda_1^\alpha, \dots, \lambda_d^\alpha) \mathbf{V}^*$. From the Bregman divergence of Eq. 6, it follows the α -quantum Bregman divergence:

$$D_\alpha(\mathbf{P}||\mathbf{Q}) = \frac{1}{\alpha(1-\alpha)} \text{Tr}(\mathbf{Q}^\alpha - \mathbf{P}^\alpha + \alpha \mathbf{Q}^{\alpha-1}(\mathbf{P} - \mathbf{Q})). \quad (11)$$

Note that since Bregman generators are equivalent up to affine terms [10], we find that F_0 is the quantum Burg entropy and F_1 is the usual von Neumann entropy. Observe that we have in the limit case $\lim_{\alpha \rightarrow 1} F_\alpha(\mathbf{X}) = F(\mathbf{X}) = \text{Tr}(\mathbf{X} \log \mathbf{X} - \mathbf{X})$ and $\lim_{\alpha \rightarrow 1} \nabla F_\alpha(\mathbf{X}) = \lim_{\alpha \rightarrow 1} \frac{1}{1-\alpha} (\mathbf{I} - \mathbf{X}^{\alpha-1}) = \log \mathbf{X}$.



Fig. 1. Quantum Voronoi diagram of 1-qubit pure states: Voronoi cells annotated with density matrices on the latitude-longitude map (left) and Voronoi cells visualized on the 3D Bloch sphere (right). Zoom in pdf please.

III. QUANTUM VORONOI DIAGRAMS

Kato et al. [7], [8] studied the Voronoi diagram for 1-qubit systems with respect to the Fubini-Study D_{FS} and Bures D_B metric distances [9] for pure states. Let \mathbf{P} and \mathbf{Q} denote two pure state density matrices represented on the Bloch sphere by 3D points \mathbf{p} and \mathbf{q} , respectively. We have $D_{FS}(\mathbf{P}, \mathbf{Q}) = \arccos \sqrt{\text{Tr}(\mathbf{P}\mathbf{Q})} = \arccos \sqrt{\frac{1+\langle \mathbf{p}, \mathbf{q} \rangle}{2}}$ and $D_B(\mathbf{P}, \mathbf{Q}) = \sqrt{1 - \text{Tr}(\mathbf{P}\mathbf{Q})} = \frac{1}{\sqrt{2}} \|\mathbf{p} - \mathbf{q}\|$. Observe that we just need the diagonal elements⁵ selected by the trace operator for computing these quantum distances. They thus show that for the case of pure state 1-qubits, these Voronoi diagrams are *equivalent* to the ordinary Voronoi diagram on the sphere. This spherical Voronoi diagram can in turn be simply obtained as the ordinary 3D Euclidean Voronoi diagram restricted to the unit (Bloch) sphere. Figure 1 depicts such a quantum Voronoi diagram under these metrics with density matrix annotations in each Voronoi cell. Moreover, Kato et al. [7], [8] investigated the Voronoi diagram with respect to the quantum information divergence and carried out calculations in the limit case of \mathbf{Q} being a pure state. They deduce that the quantum Voronoi diagram of pure states is identical to the conventional spherical Voronoi diagram although they differ for mixed states. We revisit concisely these results under the framework of Bregman Voronoi diagrams [10] and show how to naturally extend these diagrams to pure states from corresponding affine power diagrams fully defined over \mathbb{R}^3 . Since Bregman divergences are usually asymmetric [10], we consider the *left-sided* and *right-sided* Bregman Voronoi diagrams of density matrix set $\mathcal{P} = \{\mathbf{P}_1, \dots, \mathbf{P}_n\}$ defined as the cell complex induced by the left- and right-sided *Bregman bisectors*. The Voronoi cells $\text{Vor}_F(\mathbf{P}_i) = \cap_{j \neq i} H_F(\mathbf{P}, \mathbf{P}_j)$ and $\text{Vor}'_F(\mathbf{P}_i) = \cap_{j \neq i} H'_F(\mathbf{P}, \mathbf{P}_j)$ are defined respectively using the sided bisectors as follows: $H_F(\mathbf{P}, \mathbf{Q}) = \{\mathbf{X} \mid D_F(\mathbf{X} \parallel \mathbf{P}) = D_F(\mathbf{X} \parallel \mathbf{Q})\}$, and $H'_F(\mathbf{P}, \mathbf{Q}) = \{\mathbf{X} \mid D_F(\mathbf{P} \parallel \mathbf{X}) = D_F(\mathbf{Q} \parallel \mathbf{X})\}$.

These bisectors match only for symmetric Bregman divergences that are generalized quadratic distances [10]. Let $\mathbf{X}' = \nabla F(\mathbf{X})$ denote the gradient of Hermitian matrix \mathbf{X} . The left-sided bisector is always a *hyperplane* [10] whatever the considered generator F :

$$H_F(\mathbf{P}, \mathbf{Q}) : \langle \mathbf{X}, \mathbf{P}' - \mathbf{Q}' \rangle + F(\mathbf{P}) - \langle \mathbf{P}, \mathbf{P}' \rangle - F(\mathbf{Q}) + \langle \mathbf{Q}, \mathbf{Q}' \rangle = 0 \quad (12)$$

⁵In general, for computing the quantum divergence between any two $d \times d$ states \mathbf{P} and \mathbf{Q} , we just need $O(d^2)$ operations for computing the matrix product of density matrices along the diagonal only.

This equation becomes for the case of the *extended* negative von Neuman entropy (with $\mathbf{P}' = \log \mathbf{P}$), the following hyperplane equation in dimension d :

$$H_F(\mathbf{P}, \mathbf{Q}) : \{\mathbf{X} \mid \text{Tr}(\mathbf{X}(\log \mathbf{P} - \log \mathbf{Q}) - \mathbf{P} + \mathbf{Q}) = 0\}. \quad (13)$$

Using the a, b, c notations of the spherical coordinates of Eq. 9, it follows that for 1-qubit states on the 3D Bloch ball we have the bisector *plane* equation:

$$H_F(\mathbf{p}, \mathbf{q}) : \{\mathbf{x} \mid \langle \mathbf{x}, c(r_Q)\mathbf{q} - c(r_P)\mathbf{p} \rangle + b(r_P) - b(r_Q) = 0\} \quad (14)$$

Next, we show that the right-type bisector is not linear but dually linear in the gradient space $\nabla F(\mathbf{X})$.

A. Legendre duality, dual divergence and bisector

Since generator F of Eq. 6 is strictly convex and differentiable, we associate to F a unique dual conjugate function F^* via the Legendre-Fenchel slope transformation such that:

$$F^*(\mathbf{Y}) = \sup_{\mathbf{X} \in \mathcal{S}(\mathbb{C}^d)} \{\langle \mathbf{Y}, \mathbf{X} \rangle - F(\mathbf{X})\}. \quad (15)$$

The unique supremum is reached at point $\mathbf{Y} = \nabla F(\mathbf{X}) = \mathbf{X}'$. The Legendre transformation defines a dual quantum Bregman divergence for the dual generator $F^*(\mathbf{X}) = \text{Tr}(\exp \mathbf{X})$ ($\exp \mathbf{X}$ is again defined using the spectral decomposition $\exp \mathbf{X} = \mathbf{V} \text{Diag}(\exp \lambda_1, \dots, \exp \lambda_d) \mathbf{V}^*$):

$$D_F(\mathbf{P} \parallel \mathbf{Q}) = F(\mathbf{P}) + F^*(\mathbf{Q}') - \langle \mathbf{P}, \mathbf{Q}' \rangle = D_{F^*}(\mathbf{Q}' \parallel \mathbf{P}') \quad (16)$$

Thus although the right-type bisector is not linear, it is dually linear by considering the associated dual Bregman generator: $H'_F(\mathbf{P}, \mathbf{Q}) \equiv H_{F^*}(\mathbf{Q}', \mathbf{P}')$.

For 1-qubit states represented by a 3D point inside the Bloch ball, we have the Legendre conjugate explicit using 3D coordinates as [7]:

$$(x^*, y^*, z^*) = \nabla F_B(x, y, z) = \frac{1}{2r} \log \frac{1+r}{1-r}(x, y, z), \quad (17)$$

where ∇F_B is the gradient for the 3D Bloch generator function, and r is the radius $r = \sqrt{x^2 + y^2 + z^2}$. Observe that the dual Legendre function F_B^* does not admit any closed-form formula although we can easily tabulate it in practice for fine approximations by using a 1D look-up-table array.

B. Affine and dually affine quantum Voronoi diagrams

Since the left-type Voronoi diagram is affine, we use the handy *universal construction* of affine diagram from power diagram [11] to define quantum Voronoi diagrams as power diagrams of balls in the Laguerre geometry. We associate to density matrix \mathbf{X}_i the ball with Hermitian matrix center:

$$\nabla F(\mathbf{X}_i) = \log \mathbf{X}_i = \mathbf{V}_i \begin{bmatrix} \log \lambda_{i,1} & 0 \\ 0 & \log \lambda_{i,2} \end{bmatrix} \mathbf{V}_i^* \quad (18)$$

with $\mathbf{V}_i = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{x_i - iy_i}{\sqrt{x_i^2 + y_i^2}} \sqrt{\frac{r_i + z_i}{r_i}} & \frac{x_i - iy_i}{\sqrt{x_i^2 + y_i^2}} \sqrt{\frac{r_i - z_i}{r_i}} \\ \sqrt{\frac{r_i - z_i}{r_i}} & -\sqrt{\frac{r_i + z_i}{r_i}} \end{bmatrix}$ and squared radius [10]

$$r_i^2 = \langle \nabla F(\mathbf{X}_i), \nabla F(\mathbf{X}_i) \rangle + 2(F(\mathbf{X}_i) - \langle \mathbf{X}_i, \nabla F(\mathbf{X}_i) \rangle), \quad (19)$$

that is potentially imaginary and infinite for pure states (since one eigenvalue $\frac{1-r}{2}$ is zero [7]). The power bisector of two 3D Euclidean balls $B(\mathbf{p}, r_P)$ and $B(\mathbf{q}, r_Q)$ centered at 3D points \mathbf{p} and \mathbf{q} is the radical hyperplane of equation [10]:

$$2 \langle \mathbf{x}, \mathbf{q} - \mathbf{p} \rangle + \langle \mathbf{p}, \mathbf{p} \rangle - \langle \mathbf{q}, \mathbf{q} \rangle + r_Q^2 - r_P^2 = 0. \quad (20)$$

Since pure states have the *same* equivalent ball radius (being infinite in this limit case, see Eq. 19) it follows that the dual quantum Voronoi diagrams on the Bloch sphere has *in the limit case* (proof omitted) matching affine bisectors which coincides exactly with the bisector equation⁶ for the 3D ordinary Euclidean Voronoi diagram for 3D points \mathbf{p} and \mathbf{q} on the Bloch sphere.

Theorem 3.1: The von Neumann quantum Voronoi diagrams are Bregman Voronoi diagrams that can be computed from equivalent power diagrams. In particular, the quantum Voronoi diagrams on the Bloch sphere of pure states coincides with an ordinary Euclidean Voronoi diagram restricted to the sphere.

This reduction to power diagrams is attractive since power diagrams are defined on the *full Euclidean space* \mathbb{E}^3 (i.e., inside mixed states, on pure states and outside the Bloch sphere). The 1-qubit quantum Voronoi diagram can be computed easily using 3D power diagrams [11] and $\mathbf{X} \leftrightarrow \nabla F(\mathbf{X})$ conversions [10]. Although the left-side and right-side quantum Voronoi diagrams on pure states match and coincide with the Euclidean Voronoi diagram, it is not anymore the case *inside* the Bloch ball of mixed states where they provably differ [10].

IV. HOLEVO CHANNEL CAPACITY

A quantum channel is a *linear transform*. That is, an affine map that maps quantum states to other quantum states: $T : \mathcal{S}(\mathbb{C}^2) \rightarrow \mathcal{S}(\mathbb{C}^2)$. Geometrically, the effect of a channel is to map the 3D Bloch ball of pure/mixed states to a deformed 3D ellipsoid contained inside the Bloch ball (thus T is a particular affine map). The Holevo capacity [12] of this channel T is defined as the *radius* of the smallest enclosing ball of the ellipsoid:

$$C(T) = \inf_{\mathbf{P} \in \mathcal{S}(\mathbb{C}^2)} \sup_{\mathbf{Q} \in \mathcal{S}(\mathbb{C}^2)} I(T(\mathbf{Q}) \| T(\mathbf{P})). \quad (21)$$

Kato et al. [7] proposed an approximation algorithm for computing the Holevo capacity using the *farthest quantum Voronoi diagram* as follows:

⁶Indeed, we have in Euclidean geometry:

$$\|\mathbf{x}\mathbf{p}\| = \|\mathbf{x}\mathbf{q}\| \Leftrightarrow \langle \mathbf{x} - \mathbf{p}, \mathbf{x} - \mathbf{p} \rangle = \langle \mathbf{x} - \mathbf{q}, \mathbf{x} - \mathbf{q} \rangle$$

(by squaring norms) from which we get the ordinary bisector equation: $2 \langle \mathbf{x}, \mathbf{q} - \mathbf{p} \rangle + \langle \mathbf{p}, \mathbf{p} \rangle - \langle \mathbf{q}, \mathbf{q} \rangle = 0$.

- Sample uniformly n “points” \mathcal{P} (i.e., density matrices) on the 3D Bloch sphere,
- Map these pure states using the quantum channel (i.e., apply the linear transform $\mathcal{P} \mapsto T(\mathcal{P})$),
- Compute the smallest enclosing ball of $T(\mathcal{P})$ and retrieve the center (i.e., deduce the capacity as the radius).

The farthest quantum Voronoi diagram as well as k -order quantum Voronoi diagrams [10] are similarly affine diagrams that can be derived from power diagrams in the Laguerre geometry [10]. Further, these diagrams can be computed from the lower envelopes of hyperplanes in dimension $d+1$ tangent to the hypersurface $(\mathbf{X}, F(\mathbf{X}))$ derived from the lifting to the potential function F , as described in [7], [10]. Instead of computing the full quantum Voronoi diagram, we better use the linear-time generalization [13] of Welzl’s smallest enclosing ball algorithm to Bregman divergences to compute exactly the smallest enclosing ball of the discrete sample point set \mathcal{P} . This fully explains⁷ why at most $d+1 = 4$ points of $T(\mathbf{P})$ (the deformed ellipsoid) are lying on the smallest enclosing sphere. The error analysis due to the input sampling has been reported in Theorem 1 of [8]. Since the *quantum circumcenter* of $T(\mathcal{P})$ yields anyway an approximation of the channel capacity, we rather use a practical approximation algorithm based on Bregman core-sets [19] for matrix divergences.

The *simple iterative* algorithm works as follows [19]: Choose an initial circumcenter seed \mathbf{c}_1 (i.e., random point of \mathcal{P} or the centroid $\bar{\mathcal{P}}$), find the farthest point \mathbf{f} of \mathcal{P} wrt. to the quantum (Bregman) divergence and update the circumcenter to $\mathbf{c}_{i+1} \leftarrow \nabla F^{-1}(\nabla F(\mathbf{c}_i) + \frac{1}{i+1} \nabla F(\mathbf{f}))$ by walking on the geodesic $\mathbf{c}_i \mathbf{f}$. At the i -th iteration, we get a $O(1 + \sqrt{i})$ -approximation [19] of the circumcenter. Therefore it takes $O(\frac{dn}{\epsilon^2}) = O(\frac{d}{\epsilon^2})$ time to get a $(1 + \epsilon)$ -approximation, by first sampling $n = \frac{1}{\epsilon}$ points [7]. We can further generalize this method to the improved $O(\frac{dn}{\epsilon}) = O(\frac{d}{\epsilon})$ algorithm by generalizing the core-set approach of Panigrahy [14].

Theorem 4.1: A fine approximation of the Holevo capacity channel of 1-qubit states can be computed in $O(\frac{1}{\epsilon^2})$ -time.

The radius (Holevo capacity of the quantum communication channel) is given by the Jensen-Shannon divergence [15], [19]:

$$r^* = \min_{\mathbf{w}} \left(H \left(\sum_i w_i \mathbf{X}_i \right) - \sum_i w_i H(\mathbf{X}_i) \right), \quad (22)$$

where the $w_i > 0$ indicates the states participating to the basis of the smallest enclosing ball of the ellipsoid. This Holevo capacity is thus based on Jensen’s remainder, a measure of the convexity degree [15] of the von Neumann entropy. The quantum capacity can be extended to the α -entropies defined in Section II and yields a parametric generalization of Jensen-Shannon divergences: The Burbea-Rao divergences [16].

Computing experimentally the channel capacity is important for tackling the conjecture problems of quantum information theory [7], [8], first raised in 1996: Show that additivity

⁷In [8], Kato et al. described this as “a mysterious geometric structure of the space of quantum states.” This property holds for any Bregman divergence, including the α -Bregman Csizsár quantum divergences.

property of quantum channels or provide a counterexample. The additivity conjecture of the Holevo capacity expresses the power of entangled states in quantum communication.

V. QUANTUM CENTROIDS AND INFORMATION RADIUS

To get quick upper bounds on the Holevo capacity, we may reconsider Eq. 22 and fix the weight vector \mathbf{w} . We then need to find the center \mathbf{X}^+ by minimizing the *average quantum information divergence*. This MINAVG optimization problem [17] yields to the notion of sided left- and right-type quantum centroids that have been investigated recently under the framework of Bregman divergences [17]. It is shown in [17] that these Bregman centroids are *generalized means* [17] with identical information radius r_F^+ defined as the Burbea-Rao divergence [16]: $r_F^+(\mathbf{w}) = F(\sum_i w_i \mathbf{X}_i) - \sum_i w_i F(\mathbf{X}_i)$. For uniform weight vector \mathbf{w} , we get the sided centroids; For non-uniform weight vectors \mathbf{w} , this yields the notion of barycenter. Thus the circumcenter is itself the barycenter of the basis points of the smallest enclosing ball. For sake of simplicity, we have considered the spaces of density matrices \mathcal{X} and its “dual” gradient space $\nabla F(\mathcal{X})$ sitting in Euclidean geometry \mathbb{E}^3 . However, the quantum states form a *manifold* \mathcal{M} in information geometry [18]. Namely, the 1-qubit state space is a flat manifold \mathcal{M} that admits a dual pair of *biorthogonal* coordinate systems [18]: The θ - and $\eta = \nabla F(\theta)$ -affine coordinate systems related to the $\nabla^{(m)}$ and $\nabla^{(e)}$ mixture and exponential connections [18]. The Bregman generator functions are then called *potential* or *contrast* functions [18] and play the role of *canonical* divergences in these dually flat manifolds. In that framework, Spellman and Vemuri [20] called these sided information-theoretic (Kullback-Leibler) circumcenters the *e*-center and *m*-center. They compute these centers using convex programming but did not notice that the optimization problem was ill-defined because it is not differentiable on the furthest Voronoi diagram. Pelletier [21] studied the sided barycenters in information geometry using projection and considered interpolation applications of these barycenters for statistical problems.

VI. CONCLUSION

We have presented a generalization of the quantum Voronoi diagrams to the family of quantum divergences based on α -entropic Bregman-Csiszár functions [5]. We showed that these dual diagrams can be derived equivalently from power (affine) diagrams. Further, we described a *simple* and *practical* $(1 + \epsilon)$ -approximation $O(\frac{1}{\epsilon^2})$ time algorithm for computing the Holevo capacity. Although we considered in this paper quantum Voronoi diagrams and the Holevo channel capacity, we would like to point out that the scope of computational geometry in dually flat spaces for (Hermitian) matrices extends beyond quantum information theory. For example, the von Neuman divergence has been successfully used for machine learning algorithms [22]. In future work, we would like to study geometrically the 1-qubit states for another prominent class of distortion measures: Csiszár *f*-divergences based on the “likelihood” ratio $\frac{p}{q}$: $I_f(p||q) = \int qf(\frac{p}{q})dx$. In that case,

the underlying geometry is not anymore necessarily flat as shown by Yoshizawa and Tanabe [23] for the special case of multivariate normal distributions. The intersection of the class of Bregman divergences with the class of Csiszár *f*-divergences is the only Kullback-Leibler divergence.

ACKNOWLEDGMENT

This work was partially sponsored by French ANR GAIA 07-BLAN-0328-04.

REFERENCES

- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, ISBN 0521635039, 2000.
- [2] D. Petz and C. Sudár, “Geometries of quantum states,” *Journal of Mathematical Physics*, vol. 37, no. 6, pp. 2662–2673, 1996.
- [3] L. M. Bregman, “The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming,” *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 200–217, 1967.
- [4] I. S. Dhillon and J. A. Tropp, “Matrix nearness problems with Bregman divergences,” *SIAM Journal on Matrix Analysis and Applications*, vol. 29, no. 4, pp. 1120–1146, 2007.
- [5] I. Csiszar, “Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems,” *The Annals of Statistics*, vol. 19, no. 4, pp. 2032–2066, 1991.
- [6] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, “Clustering with Bregman divergences,” *Journal of Machine Learning Research (JMLR)*, vol. 6, pp. 1705–1749, 2005.
- [7] K. Kato, M. Oto, H. Imai, and K. Imai, “Voronoi diagrams for pure 1-qubit quantum states,” quant-ph/0604101, 2006.
- [8] K. Kato, H. Imai, and K. Imai, *Error Analysis of a Numerical Calculation about One-qubit Quantum Channel Capacity*, 2007. See PhD thesis: arXiv:0803.3109
- [9] D. Bures, “An extension of kakutani’s theorem on infinite product measures to the tensor product of semifinite w^* -algebras,” *Transactions on American Mathematical Society*, vol. 135, pp. 199–212, 1969.
- [10] F. Nielsen, J.-D. Boissonnat, and R. Nock, “Bregman Voronoi diagrams: Properties, algorithms and applications,” 2007, arXiv.org:0709.2196 (ACM-SIAM SODA’07).
- [11] F. Aurenhammer, “Power diagrams: properties, algorithms and applications,” *SIAM J. Comput.*, vol. 16, no. 1, pp. 78–96, 1987.
- [12] A. S. Holevo, “The capacity of quantum channel with general signal states,” *IEEE Transactions on Information Theory*, vol. 44, quant-ph/9611023, 1998.
- [13] F. Nielsen and R. Nock, “On the smallest enclosing information disk,” *Information Processing Letters*, vol. 105, no. 3, pp. 93–97, 2008.
- [14] R. Panigrahy, “Minimum enclosing polytope in high dimensions,” arXiv cs.CG/0407020, 2004.
- [15] J. Lin, “Divergence measures based on the shannon entropy,” *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991.
- [16] J. Burbea and C. Rao, “On the convexity of some divergence measures based on entropy functions,” *IEEE Transactions on Information Theory*, vol. 28, no. 3, pp. 489–495, 1982.
- [17] F. Nielsen and R. Nock, “On the centroids of symmetrized Bregman divergences,” 2007, arXiv.org:0711.3242.
- [18] S. Amari and H. Nagaoka, *Methods of Information Geometry*, A. M. Society, Ed. Oxford University Press, 2000.
- [19] F. Nielsen and R. Nock, “On Approximating the Smallest Enclosing Bregman Balls,” 2006, *ACM Symposium on Computational Geometry*, pp. 485–486. <http://www.sonycsl.co.jp/person/nielsen/BregmanBall/>
- [20] E. Spellman and B. Vemuri, “Efficient shape indexing using an information theoretic representation,” 2005, pp. 145–153.
- [21] B. Pelletier, “Informative barycentres in statistics,” *Annals of the Institute of Statistical Mathematics*, vol. 57, no. 4, pp. 767–780, 2005.
- [22] K. Tsuda, G. Rätsch, and M. K. Warmuth, “Matrix exponentiated gradient updates for on-line learning and Bregman projection,” *J. Mach. Learn. Res.*, vol. 6, pp. 995–1018, 2005.
- [23] S. Yoshizawa and K. Tanabe, “Dual differential geometry associated with Kullback-Leibler information on the Gaussian distributions and its 2-parameter deformations,” *SUT Journal of Mathematics*, vol. 35, no. 1, pp. 113–137, 1999.