Studia Sci. Math. Hungar. **38** (2001) 177–190

Convex cores of measures on $\mathbb{R}^d$

*Imre Csiszár** *and František Matúš*†

Abstract. We define the convex core of a finite Borel measure Q on $\mathbb{R}^d$ as the intersection of all convex Borel sets C with $Q(C) = Q(\mathbb{R}^d)$. It consists exactly of means of probability measures dominated by Q. Geometric and measure-theoretic properties of convex cores are studied, including behaviour under certain operations on measures. Convex cores are characterized as those convex sets that have at most countable number of faces.

1. Introduction

Let Q be a finite Borel measure on the Euclidean space $\mathbb{R}^d$. Borel sets $B \subseteq \mathbb{R}^d$ with the property $Q(B) = Q(\mathbb{R}^d)$ are called full. The intersection of all closed and full sets is the *support* of Q and, similarly, the intersection of all closed, convex and full sets is the *convex support* of Q. The support of Q, $s(Q)$, is closed and full and the convex support of Q, $cs(Q)$, is closed, convex and full. The intersection of all convex and full sets, to be called here *convex core* of Q, $cc(Q)$, seems to have been unmentioned. Yet, the latter is of primary importance for understanding closures of exponential families and problems involving I-divergence projections. These will be investigated in a forthcoming paper, see also [2].

The set $cc(Q)$ is convex but not necessarily full, e.g. for the Lebesque measure on the boundary of a ball, the convex core is the interior of the ball which has measure zero. Obviously, $cc(Q) \subseteq cs(Q)$ and we will show that $cs(Q)$ is the closure of $cc(Q)$. Points $x \in \mathbb{R}^d$ that are of positive measure, $Q(x) > 0$, belong clearly to $cc(Q)$ and thus their convex hull is a Borel set contained in $cc(Q)$. If the support of Q is at most countable then $cc(Q)$ coincides with the convex hull of $s(Q)$. In general it is not obvious that $cc(Q)$ is a Borel set but this follows from our characterization result, Theorem 1.

This work was supported by the HSSS programme of ESF, by the Hungarian National Foundation for Scientific Research, Grants T 26041 and T 032323, and by Grant Agency of Academy of Sciences of the Czech Republic, Grant A 1075801.

2000 Mathematics Subject Classification. Primary 28A99; secondary 52A20, 28A75, 60A10, 62B10.

Key words and phrases. Convex support, convex sets in n dimensions, lattice of faces, means of probabilities, convolution, exponential family.

*I. Csiszár is with the A. Rényi Institute of Mathematics, Hungarian Academy of Sciences, H-1364 Budapest, P.O.Box 127, Hungary; `csiszar@math-inst.hu`

†F. Matúš is with Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic; `matus@utia.cas.cz`

This paper is organized as follows. In the next section we elaborate upon auxiliary lemmas, prove Theorem 1, and give an alternative representation of $\mathit{cc}(Q)$, Theorem 2. Section 3 deals with means of probability measures dominated by a given measure Q. The key result is that the set of these means is equal to $\mathit{cc}(Q)$, Theorem 3. In Section 4 we investigate convex cores of sums, products, images and convolutions of measures. Section 5 discusses relations of our work to some results of Chentsov [1].

2. Faces of convex cores

Throughout this paper we use the terminology and notation of [3], e.g. the *affine hull* of a set $C \subseteq \mathbb{R}^d$ is denoted by $\mathit{aff}(C)$ and the convex hull of C by $\mathit{conv}(C)$. A *face* of a convex set C is a convex subset $F \subseteq C$ such that every closed line segment in C with a relative interior point in F is contained entirely in F. A face $F \subseteq C$ is *proper* if $F \neq C$. The relative interiors $\mathit{ri}(F)$ of nonempty faces F of any nonempty convex set C partition the set C ([3], Theorem 18.2, p. 164).

Lemma 1. $\mathit{cl}(\mathit{cc}(Q)) = \mathit{cs}(Q)$ *and* $\mathit{ri}(\mathit{cc}(Q)) = \mathit{ri}(\mathit{cs}(Q))$.

Proof. First we prove $\mathit{ri}(\mathit{cs}(Q)) \subseteq C$ for each convex Q-full set C. Let $A = C \cap \mathit{aff}(s(Q))$, then $Q(C) = 1$ implies $Q(A) = 1$. Since $\mathit{cl}(A)$ is closed, convex and Q-full, it contains $\mathit{cs}(Q)$. Knowing that the affine dimensions of $s(Q)$, $\mathit{cs}(Q)$ and $\mathit{cl}(A) \subseteq \mathit{aff}(s(Q))$ coincide we conclude, by Corollary 6.5.2 and Corollary 6.3.3 in [3], that $\mathit{ri}(\mathit{cs}(Q))$ is contained in $\mathit{ri}(A) \subseteq A \subseteq C$.

As $\mathit{ri}(\mathit{cs}(Q))$ is a subset of every convex Q-full set, we have $\mathit{ri}(\mathit{cs}(Q)) \subseteq \mathit{cc}(Q)$. This and the inclusion $\mathit{cc}(Q) \subseteq \mathit{cs}(Q)$ imply $\mathit{cl}(\mathit{cc}(Q)) = \mathit{cs}(Q)$ and, again by Corollary 6.5.2 in [3], $\mathit{ri}(\mathit{cc}(Q)) = \mathit{ri}(\mathit{cs}(Q))$. □

Corollary 1. *The assertions* $\mathit{cc}(Q) = \emptyset$, $\mathit{cs}(Q) = \emptyset$ *and* $Q = 0$ *are equivalent.*

For a Borel set A let Q^A denote the *restriction* of Q to A, $Q^A(B) = Q(A \cap B)$ for every $B \subseteq \mathbb{R}^d$ Borel.

Lemma 2. *(i) If C is a convex Borel set then* $\mathit{cc}(Q) \cap C \supseteq \mathit{cc}(Q^C)$.
(ii) Let H be a hyperplane and $H_<, H_>$ the open halfspaces determined by H. If the intersection $\mathit{cc}(Q) \cap H_>$ *is empty then* $\mathit{cc}(Q) \cap H = \mathit{cc}(Q^H)$.

Proof. (i) Since every Q-full set is Q^C-full one has $\mathit{cc}(Q) \supseteq \mathit{cc}(Q^C)$; also as C is Q^C-full $C \supseteq \mathit{cc}(Q^C)$.

(ii) The assumption implies $\mathit{cs}(Q) \subseteq H \cup H_<$ by Lemma 1 and then $Q(H_>) = 0$. If C is convex Q^H-full then $C \cap H$ is convex Q^H-full and $(C \cap H) \cup H_<$ is convex and Q-full. Thus $\mathit{cc}(Q) \subseteq (C \cap H) \cup H_<$ and $\mathit{cc}(Q) \cap H \subseteq C$. This proves $\mathit{cc}(Q) \cap H \subseteq \mathit{cc}(Q^H)$, and the reverse inclusion holds by (i). □

Corollary 2. $\mathit{cc}(Q) = \mathit{ri}(\mathit{cs}(Q))$ *if and only if each nontrivial supporting hyperplane of* $\mathit{cs}(Q)$ *has measure zero.*

Proof. The inclusion $cc(Q) \supseteq ri(cs(Q))$ is by Lemma 1. The difference $cc(Q) \backslash ri(cs(Q))$ is nonempty if and only if there exists a nontrivial supporting hyperplane H of $cs(Q)$ with $cc(Q) \cap H \neq \emptyset$. By Lemma 2 (ii), $cc(Q) \cap H = cc(Q^H)$, which is nonempty if and only if $Q(H) > 0$, due to Corollary 1. $\square$

Lemma 3. *$cc(Q^{cl(F)}) = F$ for every face F of $cc(Q)$.*

Proof. For the empty face it is trivial. If $F = cc(Q)$ then $cl(F) = cs(Q)$ by Lemma 1, and then $Q = Q^{cl(F)}$. We use induction on the affine dimension of $cc(Q)$; the case of zero dimension was covered. Let F be a nonempty proper face of $cc(Q)$. There exists a nontrivial supporting hyperplane H to $cc(Q)$ containing F, cf. [3], Theorem 11.6, p. 100. The intersection $cc(Q) \cap H$ is a proper face of $cc(Q)$, and F is a face of $cc(Q) \cap H$. By Lemma 2 (ii) F is a face of $cc(Q^H)$. By the induction argument on affine dimension of $cc(Q)$, applied to the measure Q^H, we conclude $F = cc((Q^H)^{cl(F)}) = cc(Q^{cl(F)})$. $\square$

Corollary 3. *$Q(cl(F)) > 0$ for every nonempty face F of $cc(Q)$.*

Proof. For a face F of $cc(Q)$, Lemma 3 and Corollary 1 give that $F = \emptyset$ and $Q^{cl(F)} = 0$ are equivalent. Hence $Q(cl(F)) = 0$ if and only if $F = \emptyset$. $\square$

Corollary 4. *$Q(cl(F) \cap cl(G)) = Q(cl(F \cap G))$ for any two faces F and G of $cc(Q)$.*

Proof. The convex core of $Q^{cl(F) \cap cl(G)}$ is contained in $cc(Q^{cl(F)}) \cap cc(Q^{cl(G)}) = F \cap G$. This implies by Lemma 1 that the convex support of $Q^{cl(F) \cap cl(G)}$ is contained in $cl(F \cap G)$ and hence

$$Q(cl(F) \cap cl(G)) = Q^{cl(F) \cap cl(G)}(\mathbb{R}^d) = Q^{cl(F) \cap cl(G)}(cl(F \cap G)) = Q(cl(F \cap G))$$

as desired. $\square$

The family $\mathcal{F}_C$ of all faces of a given convex set C is a complete lattice when the meet is defined by intersection and the join $F_1 \vee F_2$ as the smallest face of C containing both F_1 and F_2. An element $F \in \mathcal{F}_C$ is (join) irreducible if $F = F_1 \vee F_2$ for $F_1, F_2 \in \mathcal{F}_C$ implies $F = F_1$ or $F = F_2$. Every face F can be written, in general nonuniquely, as $F = F_1 \vee \ldots \vee F_k$ where $F_1, \ldots, F_k$ are irreducible and $k \leqslant 2^d$. This follows from the fact that if $F = F_1 \vee F_2$ where F_1 and F_2 are proper faces of F then the affine dimensions of F_1 and F_2 are smaller than the affine dimension of F. The lattice $\mathcal{F}_C$ is at most countable if its subset $\mathcal{F}_C^{\mathrm{irr}}$ of all irreducible elements is at most countable.

Theorem 1. *A convex set $C \subseteq \mathbb{R}^d$ is the convex core of some finite Borel measure if and only if C has at most a countable number of faces.*

Proof. Let $C = cc(Q)$ and F be a nonempty irreducible element of $\mathcal{F}_C$. The join of all proper faces of F is a proper face F_0 of F. Obviously, $Q(cl(F) - cl(F_0)) > 0$ because otherwise $Q^{cl(F)} = Q^{cl(F_0)}$ would imply $F = F_0$ by Lemma 3. We are going to show that

$$Q\big((cl(F) - cl(F_0)) \cap (cl(G) - cl(G_0))\big) = 0$$

for two different nonempty irreducible faces F, G from $\mathcal{F}_C$. Clearly this implies that the family $\mathcal{F}_C^{\mathrm{irr}}$ is at most countable. Since $\mathit{ri}(F)$ and $\mathit{ri}(G)$ are disjoint, there exists a hyperplane H separating F and G properly, see [3], Theorem 11.3, p. 97. By symmetry, let F be not contained in H. The convex core of $(Q^{\mathit{cl}(F)})^H$ is, by Lemma 2 (ii), equal to $E = \mathit{cc}(Q^{\mathit{cl}(F)}) \cap H$. Lemma 3 entails $E = F \cap H$ and as E is a proper face of F we have $E \subseteq F_0$. Using Corollary 4 and $F \cap G \subseteq F \cap H = E$ we can write

$$\begin{aligned} Q(\mathit{cl}(F) \cap \mathit{cl}(G)) = Q(\mathit{cl}(F \cap G)) &= Q^{\mathit{cl}(G)}(\mathit{cl}(F \cap G)) \\ &\leqslant Q^{\mathit{cl}(G)}(\mathit{cl}(E)) \leqslant Q^{\mathit{cl}(G)}(\mathit{cl}(F_0)) = Q(\mathit{cl}(F_0) \cap \mathit{cl}(G)) . \end{aligned}$$

Thus, the Q-measure of the intersection of $\mathit{cl}(F) - \mathit{cl}(F_0)$ and $\mathit{cl}(G)$ must be zero.

Conversely, let $C = \bigcup_{F \in \mathcal{F}_C} \mathit{ri}(F)$ be a convex set with $\mathcal{F}_C$ at most countable. For each nonempty $F \in \mathcal{F}_C$, let $R_{(F)}$ be a probability measure equivalent to the Lebesque measure on $\mathit{aff}(F)$, e.g. a Gaussian measure on $\mathit{aff}(C)$. Then $\mathit{cc}(R_{(F)}^{\mathit{ri}(F)}) = \mathit{ri}(F)$. For Q equal to a convex combination (with all weights positive) of the measures $R_{(F)}^{\mathit{ri}(F)}$, $F \in \mathcal{F}_C$, we have $C = \mathit{cc}(Q)$. Indeed, $\mathit{ri}(F) = \mathit{cc}(R_{(F)}^{\mathit{ri}(F)}) \subseteq \mathit{cc}(Q)$ implies $C \subseteq \mathit{cc}(Q)$, and the opposite inclusion holds since C is convex and Q-full. □

Corollary 5. *Each convex core is a Borel set.*

We will say that a polytope C, i.e. the convex hull of a finite set, is *Q-polytope* if all balls centered at extreme points of C have intersection with $\mathit{aff}(C)$ of positive Q-measure. A Q-polytope with its affine hull equal to $\mathit{aff}(\mathit{s}(Q))$ is referred to have full dimension. In other words, the Q-polytopes of full dimension are the polytopes C having $\mathit{aff}(C) = \mathit{aff}(\mathit{s}(Q))$ and extreme points contained in $\mathit{s}(Q)$.

Theorem 2. *The convex core $\mathit{cc}(Q)$ coincides with the union of the relative interiors of all Q-polytopes. The relative interior of $\mathit{cc}(Q)$ equals the union of relative interiors of all Q-polytopes of full dimension.*

Proof. Let U be the union of relative interiors of all Q-polytopes of full dimension. The convex hull of $\mathit{s}(Q)$ is the set of all convex combinations of form $\sum_{i=1}^{n} \alpha_i x_i$ with $n \geqslant 1$, α_i positive and $x_i \in \mathit{s}(Q)$ while U consists of the convex combinations of this form with $\mathit{aff}(\{x_1, \dots, x_n\}) = \mathit{aff}(\mathit{s}(Q))$. It follows that $\mathit{cl}(U)$ coincides with $\mathit{cl}(\mathit{conv}(\mathit{s}(Q))) = \mathit{cs}(Q)$. Since U is relatively open in $\mathit{aff}(\mathit{s}(Q))$ by definition, and its closure equals $\mathit{cs}(Q)$, we have $U = \mathit{ri}(\mathit{cs}(Q))$; thus by Lemma 1, $U = \mathit{ri}(\mathit{cc}(Q))$.

Let C be any Q-polytope and $A = \mathit{aff}(C)$; the extreme points of C belong to $\mathit{s}(Q^A)$. Then C is a Q^A-polytope of full dimension and we have $\mathit{ri}(C) \subseteq \mathit{cc}(Q^A) \subseteq \mathit{cc}(Q)$ by the argument above. It remains to show that each $a \in \mathit{cc}(Q)$ belongs to the relative interior of some Q-polytope. Let F be the face of $\mathit{cc}(Q)$ for which $a \in \mathit{ri}(F)$. Then by Lemma 3 we have $a \in \mathit{ri}(\mathit{cc}(Q^{\mathit{cl}(F)}))$ and this implies, see above, that a is contained in the relative interior of a $Q^{\mathit{cl}(F)}$-polytope C. Since any $Q^{\mathit{cl}(F)}$-polytope is also a Q-polytope, this completes the proof. □

Remark 1. The convex support of Q is obviously equal to the intersection of all full closed halfspaces. To construct the convex core of Q it is also not necessary to intersect all convex full sets. It suffices to take the full sets of form $H^* = \bigcup_{j=1}^{k} H_j$ where $0 \leqslant k \leqslant d$ and H_1 is an open halfspace of $\mathbb{R}^d$, H_2 is an open halfspace of the boundary hyperplane of H_1, etc. Indeed, let $\mathit{cc}^*(Q)$ be the intersection of all Q-full unions H^*. Clearly, $\mathit{cc}(Q) \subseteq \mathit{cc}^*(Q)$ because H^* is convex. To see the equality here let $x \notin \mathit{cc}(Q)$, i.e. x not in some convex full set C. One can show by induction on the affine dimension of C that $x \notin C$ implies the existence of $H^* \supseteq C$ such that $x \notin H^*$. This means $x \notin \mathit{cc}^*(Q)$.

3. Means, dominance and convex core

In this section P always denotes a (Borel) probability measure (pm) having a mean, i.e. such that $\int_{\mathbb{R}^d} x\, P(dx)$ exists. If Q is a finite measure and $Q(B)$ is positive for a Borel set B, let Q_B be defined by normalizing the restriction Q^B to a pm, $Q_B = Q^B/Q(B)$; this is the *conditioning* by B.

It is well-known ([1] Lemma 21.2) that if a is the mean of a pm P then $a \in \mathit{ri}(\mathit{cs}(P))$. Otherwise, we could have a in a hyperlane H such that $\mathit{cs}(P)$ is contained in $H \cup H_>$, where $H_>$ is an open halfspace determined by H, but not $\mathit{cs}(P) \subseteq H$; then we could argue by $P(H_<) = 0$, $P(H_>) = 0$ and $\mathit{s}(P) \subseteq H$ what contradicts $\mathit{cs}(P) \nsubseteq H$.

Theorem 3. *The convex core of a finite Borel measure Q equals $\left\{ \int_{\mathbb{R}^d} x\, P(dx) :\ P \ll Q \right\}$. Moreover, to each $a \in \mathit{cc}(Q)$ there exists $P \ll Q$ with mean a such that $\frac{dP}{dQ}$ is bounded.*

Proof. If $P \ll Q$ then all Q-full sets are P-full what implies $\mathit{cc}(P) \subseteq \mathit{cc}(Q)$. By the previous observation and Lemma 1, the mean of P is in $\mathit{ri}(\mathit{cs}(P)) = \mathit{ri}(\mathit{cc}(P)) \subseteq \mathit{cc}(P) \subseteq \mathit{cc}(Q)$. This shows one inclusion of Theorem 3.

Let $a \in \mathit{ri}(\mathit{cc}(Q))$. Then a belongs to the relative interior of a Q-polytope C of full dimension due to Theorem 2. For every extreme point y of C, $y \in \mathit{ext}(C)$, which belongs to $\mathit{s}(Q)$ by definition, we choose a closed ball D_y centered at y. The balls are taken sufficiently small to have a in the relative interior of any polytope in $\mathit{aff}(C)$ that intersects each D_y. The conditioning Q_{D_y} makes sense and for the means $z_y = \int_{\mathbb{R}^d} x\, Q_{D_y}(dx)$, obviously in $D_y \cap \mathit{aff}(C)$, one can find positive α_y summing to 1 such that $a = \sum_{y \in \mathit{ext}(C)} \alpha_y z_y$. The mixture $P_a = \sum_{y \in \mathit{ext}(C)} \alpha_y Q_{D_y}$ is dominated by Q, has mean equal to a, and the density w.r.t. Q bounded from above.

Let $a \in \mathit{cc}(Q)$, i.e. $a \in \mathit{ri}(F)$ for a face F of $\mathit{cc}(Q)$. By Lemma 3, $a \in \mathit{ri}(\mathit{cc}(Q^{\mathit{cl}(F)}))$ and, due to the arguments above, a equals the mean of some $P \ll Q^{\mathit{cl}(F)}$ where this P can be chosen so that $\frac{dP}{dQ^{\mathit{cl}(F)}}$ is bounded. Obviously, $P \ll Q$ and $\frac{dP}{dQ}$ is bounded. ☐

Remark 2. Given an increasing sequence of Borel sets B_n, $n \geqslant 1$, with union $\mathbb{R}^d$, we can find to each $a \in \mathit{cc}(Q)$ a sufficiently large n and $P \ll Q^{B_n}$ such that the mean of P equals a. To see this, for $a \in \mathit{ri}(\mathit{cc}(Q))$ we argue as in the previous proof, using that the numbers $Q^{B_n}(D_y)$ are positive from some n. The case $a \in \mathit{ri}(F)$ for a proper face F of $\mathit{cc}(Q)$ reduces to $a \in \mathit{ri}(\mathit{cc}(Q^{\mathit{cl}(F)}))$ as above.

Lemma 4. *If $\mathit{cc}(P)$ is contained in a convex set C and the mean of P belongs to a face F of C then $\mathit{cc}(P) \subseteq F$.*

Proof. Since $\mathit{ri}(\mathit{cc}(P))$ contains the mean of P, it meets F and Theorem 18.1 of [3] implies $\mathit{cc}(P) \subseteq F$. ∎

Corollary 6. *If the mean of $P \ll Q$ is in a face F of $\mathit{cc}(Q)$ then $P \ll Q^{\mathit{cl}(F)}$.*

Proof. The dominance $P \ll Q$ implies $\mathit{cc}(P) \subseteq \mathit{cc}(Q)$ and by Lemma 4 $\mathit{cc}(P) \subseteq F$. Therefore, $\mathit{cs}(P) \subseteq \mathit{cl}(F)$, $P(\mathit{cl}(F)) = 1$, and $P \ll Q^{\mathit{cl}(F)}$. ∎

Lemma 5. $\mathit{ri}(\mathit{cc}(Q)) = \left\{ \int_{\mathbb{R}^d} x\, P(dx) :\ P \equiv Q \right\}$. *Moreover, to each $a \in \mathit{ri}(\mathit{cc}(Q))$ there exists $P \equiv Q$ with mean a and $\frac{dP}{dQ}$ bounded.*

Proof. The inclusion $\supseteq$ is obvious because $P \equiv Q$ implies $\mathit{cc}(P) = \mathit{cc}(Q)$, and then the mean of P is in $\mathit{ri}(\mathit{cs}(P)) = \mathit{ri}(\mathit{cs}(Q)) = \mathit{ri}(\mathit{cc}(Q))$.

Let first Q be a pm with mean c, obviously in $\mathit{cs}(Q)$. For $a \in \mathit{ri}(\mathit{cc}(Q))$ the affine combination $b = \alpha a + (1-\alpha)c$ is in $\mathit{ri}(\mathit{cc}(Q))$ for some $\alpha > 1$. For P_b constructed as in the proof of Theorem 3, having mean b, the pm $\frac{1}{\alpha}P_b + (1-\frac{1}{\alpha})Q$ has mean equal to a. This pm is equivalent to Q and its density w.r.t. Q is bounded. If Q is not a pm but its normalization has a mean then the same argument works using the normalization.

If Q, normalized to a pm if necessary, does not have a mean, all we need is to find a probability measure $R \equiv Q$ that has a mean and such that $\frac{dR}{dQ}$ is bounded. This pm R can be constructed as a convex combination $\sum_{n \geqslant 1} \beta_n Q_{B_n}$ where B_n are bounded Borel sets, $Q(B_n) > 0$, that partition $\mathbb{R}^d$. Let b_n be the mean of Q_{B_n} and β_n be equal to $Q(B_n)\,\gamma 2^{-n}$ when the norm of b_n is less than one and to $Q(B_n)\,\gamma 2^{-n}\|b_n\|^{-1}$ otherwise. All β_n are positive and a unique $\gamma \geqslant 1$ exists so that β_n sum to one. It is easy to see that R is equivalent to Q, the mean of R equals $\sum_{n \geqslant 1} \beta_n b_n$ and $\frac{dR}{dQ} \leqslant \gamma$. ∎

Corollary 7. *If F is a face of $\mathit{cc}(Q)$ then $\mathit{ri}(F) = \left\{ \int_{\mathbb{R}^d} x\, P(dx) :\ P \equiv Q^{\mathit{cl}(F)} \right\}$. Moreover, to each $a \in \mathit{ri}(F)$ there exists $P \equiv Q^{\mathit{cl}(F)}$ with mean a and $\frac{dP}{dQ}$ bounded.*

Proof. By Lemma 3, $\mathit{ri}(F) = \mathit{ri}(\mathit{cc}(Q^{\mathit{cl}(F)}))$. Then Lemma 5 is applied to $Q^{\mathit{cl}(F)}$. ∎

4. Convex core and operations between measures

The convex core behaves neatly with respect to some natural operations between measures and also to limiting considerations. All measures are, as usually, finite and Borel.

We start with a trivial fact that $\mathit{cc}(Q) = \mathit{cc}(\alpha Q)$ for any α positive.

Lemma 6. $\mathit{cc}(Q_1 + Q_2) = \mathit{conv}(\mathit{cc}(Q_1) \cup \mathit{cc}(Q_2))$.

Proof. The inclusion $\supseteq$ is trivial. On the other hand, if $a \in \mathit{cc}(Q_1 + Q_2)$ then a is the mean of a pm P dominated by $Q_1 + Q_2$, cf. Theorem 3. Using the Lebesque decomposition it is easy to see that P is a convex combination of two pm's $P_1 \ll Q_1$ and $P_2 \ll Q_2$. Then a is a convex combination of the means of P_1 and P_2 that belong to $\mathit{cc}(Q_1)$ and $\mathit{cc}(Q_2)$, respectively. ∎

Remark 3. The analogue of Lemma 6 for the support is $\mathsf{s}(Q_1+Q_2)=\mathsf{s}(Q_1)\cup\mathsf{s}(Q_2)$ and for the convex support $\mathsf{cs}(Q_1+Q_2)=\mathsf{cl}(\mathsf{conv}(\mathsf{cs}(Q_1)\cup\mathsf{cs}(Q_2)))$. Here the closure cannot be omitted (in $\mathbb{R}^2$ take $\mathsf{s}(Q_1)$ to be the graph of $x\mapsto x^{-1}$, $x>0$, and $\mathsf{s}(Q_2)$ to be the graph of $x\mapsto -x^{-1}$, $x>0$).

The next assertion concerns the Cartesian product of measures on perhaps different spaces.

Lemma 7. $\mathsf{cc}(Q_1\times Q_2)=\mathsf{cc}(Q_1)\times\mathsf{cc}(Q_2)$.

Proof. The right hand side is intersection of the sets $C_1\times C_2$ over Q_1-full sets C_1 and Q_2-full sets C_2. Since all these sets are $(Q_1\times Q_2)$-full we have the inclusion $\subseteq$.

For $a=(a_1,a_2)$ in $\mathsf{cc}(Q_1)\times\mathsf{cc}(Q_2)$ we know that a_1 is the mean of a measure P_1 dominated by Q_1 and similarly with a_2. But then a is the mean of $P_1\times P_2$ which is dominated by $Q_1\times Q_2$ and must therefore belong to $\mathsf{cc}(Q_1\times Q_2)$ by Theorem 3. $\square$

Remark 4. In Lemma 7 one can replace the convex core by support or convex support.

If T is a Borel measurable transformation between two spaces and Q is a measure on the first one, the image of Q under T will be denoted by TQ, i.e. $TQ(B)=Q(T^{-1}(B))$ for any Borel set B in the second space.

Lemma 8. $\mathsf{cc}(TQ)=T(\mathsf{cc}(Q))$ *for any affine transformation* T.

Proof. We will use Theorem 3. If a pm $P\ll Q$ has mean a then $TP\ll TQ$ has mean $T(a)$; this proves $\mathsf{cc}(TQ)\supseteq T(\mathsf{cc}(Q))$. To prove the opposite inclusion, let $b\in\mathsf{cc}(TQ)$. Then b is the mean of some pm $P'\ll TQ$ with $\frac{dP'}{d(TQ)}$ bounded, by Theorem 3. The pm $P\ll Q$ defined by $\frac{dP}{dQ}(x)=\frac{dP'}{d(TQ)}(T(x))$ satisfies $P'=TP$, and therefore $b=T(a)$ where $a\in\mathsf{cc}(Q)$ is the mean of P, provided that P does have a mean. A sufficient condition for the latter is the existence of integral $\int_{\mathbb{R}^d}x\,Q(dx)$ (since $\frac{dP}{dQ}$ is bounded), thus $\mathsf{cc}(TQ)\subseteq T(\mathsf{cc}(Q))$ follows under this condition. Finally, this condition can be dispensed with, because mutually absolutely continuous measures have the same convex core, and any nonzero measure is equivalent to a pm that has a mean, see the proof of Lemma 5. $\square$

Remark 5. The analogue of Lemma 8 for supports or convex supports requires an additional closure, for example $\mathsf{s}(TQ)=\mathsf{cl}(T(\mathsf{s}(Q)))$ (in $\mathbb{R}^2$ take $(x,y)\mapsto x$ and $\mathsf{s}(Q)$ to be the graph of $x\mapsto x^{-1}$).

Combining Lemma 7 and Lemma 8 we can describe the convex core of convolution by Minkowski sum. The analogues for support and convex support need closures.

Corollary 8. $\mathsf{cc}(Q_1*Q_2)=\mathsf{cc}(Q_1)+\mathsf{cc}(Q_2)$.

As a consequence, the convolution power Q^{*n} has convex core equal to $n\mathsf{cc}(Q)$. This can be easily shown also directly. For convolution powers, $Q^{*n}(n\mathsf{ri}(\mathsf{cc}(Q)))$ always goes to 1 as $n\to\infty$, but it may happen that $Q^{*n}(n\mathsf{cc}(Q))<1$ for all $n\geqslant 1$ and also that $Q^{*n}(n\mathsf{cc}(Q))=0$ for all $n\leqslant\frac{d}{2}$.

Lemma 9. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with the distribution Q. Then*

$$Pr\left\{\tfrac{1}{n}(X_1 + \ldots + X_n) \in \mathit{ri}(\mathit{cc}(Q))\right\} = Q^{*n}(n\,\mathit{ri}(\mathit{cc}(Q))) \to 1\,, \qquad n \to \infty\,.$$

Proof. Where $k \leqslant d$ is the affine dimension of $\mathsf{s}(Q)$, there exist affinely independent points $x_0, \ldots, x_k$ in $\mathsf{s}(Q)$. Obviously, $x_0 + \ldots + x_k$ is in $\mathsf{s}(Q^{*(k+1)})$ and also in $\mathit{ri}((k+1)\mathit{cc}(Q))$ because $\frac{1}{k+1}(x_0 + \ldots + x_k)$ is in $\mathit{ri}(\mathit{cc}(Q))$. We deduce that $Q^{*(k+1)}((k+1)\mathit{ri}(\mathit{cc}(Q)))$ is positive.

Since $X_1 + \ldots + X_n$ is distributed as Q^{*n} we know that

$$Pr\left\{\tfrac{1}{m}(X_1 + \ldots + X_m) \in \mathit{ri}(\mathit{cc}(Q))\right\} > 0$$

for $m = k+1$. This implies that for n sufficiently large, with probability arbitrarily close to 1 at least one of the independent events

$$Y_i = \tfrac{1}{m}(X_{im+1} + \ldots + X_{im+m}) \in \mathit{ri}(\mathit{cc}(Q))\,, \qquad i = 0, 1, \ldots, \lceil \tfrac{n}{m} \rceil - 1,$$

will occur. But, $Y_i \in \mathit{ri}(\mathit{cc}(Q))$ implies that $\frac{1}{n}(X_1 + \ldots + X_n) \in \mathit{ri}(\mathit{cc}(Q))$ because the latter average is a convex combination of Y_i and of $X_1, \ldots, X_{im}, X_{im+m+1}, \ldots, X_n$, each in $\mathit{cs}(Q)$. □

Remark 6. If $a = \int_{\mathbb{R}^d} x\, Q(dx)$ exists then a stronger result holds. Namely, by the law of large numbers, $\frac{1}{n}(X_1 + \ldots + X_n)$ converges to $a \in \mathit{ri}(\mathit{cc}(Q))$ in probability.

Example 1. Let C be the closed cone $z^2 \leqslant x^2 + y^2$, $z \geqslant 0$, in $\mathbb{R}^3$, and Q be the average of the pm δ sitting at $(0,0,0)$ and of the pm R sitting on the boundary of C with the density $(2\pi)^{-1}e^{-r}dr\,d\vartheta$, $r > 0$, $0 \leqslant \vartheta < 2\pi$, in cylindrical coordinates. The convex core of C is union of the point $(0,0,0)$ with $\mathit{int}(C)$ and has only one nonempty proper face. Obviously, $n\mathit{cc}(Q) = \mathit{cc}(Q)$. It is not difficult to show that for the convolution power R^{*n}, $n \geqslant 2$, the interior of C is full. Therefore,

$$Q^{*n} = \frac{1}{2^n}\Big[\delta + nR + \sum_{j=2}^{n} \tbinom{n}{j}\, R^{*j}\Big] \quad \text{and} \quad Q^{*n}(\mathit{cc}(Q)) = \frac{2^n - n}{2^n}\,.$$

Example 2. Let $Q = T\lambda$ be the image of the Lebesque measure λ on [0,1] under the transformation $T(z) = (z, z^2, \ldots, z^d)$. Since every hyperplane of $\mathbb{R}^d$ intersects $\mathsf{s}(Q) = T[0,1]$ in at most d points, all hyperplanes are of zero Q-measure. Hence, by Corollary 2, $\mathit{cc}(Q)$ is equal to the interior of $\mathit{cs}(Q)$. The Q-polytopes of full dimension are of the form $C = \mathit{conv}(T(z_1), \ldots, T(z_k))$ for $k > d$ and $z_1, \ldots, z_k \in [0,1]$ different. These are known as cyclic polytopes, and one of their remarkable properties is that $C' = \mathit{conv}(T(z_{i_1}), \ldots, T(z_{i_n}))$ is always a face of C when $n \leqslant \frac{d}{2}$, see [4], Corollary 0.8, p. 15. Then C' is a face of $\mathit{cs}(Q)$, as well. In particular, the average $\frac{1}{n}(x_1 + \ldots + x_n)$ can belong to $\mathit{cc}(Q)$ for no $n \leqslant \frac{d}{2}$ points $x_i \in \mathsf{s}(Q)$. Therefore $Q^{*n}(n\mathit{cc}(Q)) = 0$ if $n \leqslant \frac{d}{2}$.

Lemma 10. *For Borel sets $B_1 \subseteq B_2 \subseteq \ldots$ with $\bigcup_{n\geqslant 1} B_n = \mathbb{R}^d$*

$$\bigcup\nolimits_{n\geqslant 1} \mathit{cc}(Q^{B_n}) = \mathit{cc}(Q)\,.$$

Proof. The inclusion $\subseteq$ is trivial. For $a \in \mathit{cc}(Q)$ there exists n and $P \ll Q^{B_n}$ with mean a by Remark 2. Theorem 3 implies $a \in \mathit{cc}(Q^{B_n})$. □

Remark 7. The analogues of Lemma 10 for support and convex support require to close the union. For example in $\mathbb{R}$, if $B_n = (-\infty, 0] \cup [\frac{1}{n}, +\infty)$ and $Q = \sum_{n \geqslant 1} 2^{-n} \delta_n$ where δ_n is the pm sitting in $\frac{1}{n}$, then 0 belongs to $\mathsf{s}(Q) \subseteq \mathsf{cs}(Q)$ but not to $(0,1] = \bigcup_{n \geqslant 1} \mathsf{cs}(Q^{B_n}) \supseteq \bigcup_{n \geqslant 1} \mathsf{s}(Q^{B_n})$.

To see $\mathsf{cl}(\bigcup_{n \geqslant 1} \mathsf{s}(Q^{B_n})) = \mathsf{s}(Q)$ and the similar relation for the convex support it suffices to show that the union is dense in $\mathsf{s}(Q)$. In fact, let $x \in \mathsf{s}(Q)$ and D_x be any closed neighbourhood of x. As $Q(D_x) > 0$ the measure $Q^{B_n \cap D_x}$ is nonzero for n sufficiently large. Then $\mathsf{s}(Q^{B_n \cap D_x}) \subseteq D_x$ is nonempty and $\mathsf{s}(Q^{B_n}) \supseteq \mathsf{s}(Q^{B_n \cap D_x})$ intersects D_x.

5. Discussion

Our motivation for studying convex cores of measures has been that this concept emerged in our work [2] about extended exponential families and I-divergence projections. In this context Chentsov's pioneering contributions [1] should be mentioned where convex support rather than convex core was used; however, [1] contains several false statements, see the example below.

A proper nonempty face F of a convex set C is *exposed* (see [3], p. 162) if it is the intersection of C with a nontrivial supporting hyperplane of C; for convenience C and $\emptyset$ will not be considered exposed, in contrast to [3]. Chentsov [1], Definition 21.5, p. 339 (p. 315 of English version[‡]) defined a *ponderable* face of $\mathsf{cs}(Q)$ as an exposed face of positive measure such that $Q(F) > Q(G)$ for each exposed face G of $\mathsf{cs}(Q)$ that is properly contained in F.

An exponential family γ, based on a probability measure Q on $\mathbb{R}^d$, is the set of all pm's $P_\theta \equiv Q$ with Q-densities $e^{\langle \theta, x \rangle - \Lambda_Q(\theta)}$ where $\Lambda_Q(\theta) = \ln \int_{\mathbb{R}^d} e^{\langle \theta, x \rangle} \, Q(dx)$ and θ ranges over the set $\mathsf{dom}(\Lambda_Q) \subseteq \mathbb{R}^d$ where the integral is finite. Chentsov [1] (Definition 21.6) enriched the exponential family γ to $[\gamma]$ by adding his 'boundary at infinity'. This boundary consists of certain pm's equivalent to restrictions of Q to ponderable faces of $\mathsf{cs}(Q)$.

Example 3. On the Euclidean plane, let $Q = \frac{1}{3}(\delta_1 + \delta_2 + \lambda_\square)$ where δ_1 is the pm concentrated in the point $(\frac{1}{3}, 0)$, δ_2 is concentrated in $(\frac{2}{3}, 0)$ and $\lambda_\square$ is the Lebesque measure on the unit square $[0,1]^2$. The convex core of Q is union of the open unit square and the closed segment F with the endpoints $(\frac{1}{3}, 0)$ and $(\frac{2}{3}, 0)$. It has four nonempty faces: the two points, F and $\mathsf{cc}(Q)$. The convex support of Q is the closed unit square. It has one ponderable face: the closed segment G with the endpoints $(0,0)$ and $(1,0)$. It corresponds to the exposed face F of $\mathsf{cc}(Q)$, cf. Lemma 11 below.

Clearly, $\mathsf{dom}(\Lambda_Q) = \mathbb{R}^2$ and

$$e^{\Lambda_Q(\theta_1, \theta_2)} = \tfrac{1}{3} \left(e^{\theta_1/3} + e^{2\theta_2/3} + \theta_1^{-1} \theta_2^{-1} (1 - e^{\theta_1})(1 - e^{\theta_2}) \right)$$

[‡]The English translation (1982) is sometimes confusing.

for nonzero $\theta_1, \theta_2 \in \mathbb{R}$. For $n \geqslant 1$ let P_n be the pm in the exponential family γ based on Q with parameter $(-3\ln n, -3n)$. The point $(\frac{1}{3}, 0)$ has P_n-measure

$$\tfrac{1}{3}\, e^{-\ln n - \Lambda_Q(-3\ln n, -3n)} = \frac{1}{3n}\left[\frac{1}{3}\left(e^{-\ln n} + e^{-2n} + \frac{1-e^{-3\ln n}}{-3\ln n}\cdot\frac{1-e^{-3n}}{-3n}\right)\right]^{-1}$$
$$= \left[1 + ne^{-2n} + \frac{(1-n^{-3})(1-e^{-3n})}{9\ln n}\right]^{-1}$$

converging to 1 as $n \to \infty$. Hence, the Kullback-Leibler divergence $D(\delta_1\|P_n)$ tends to zero.

Since $R = \delta_1 \ll Q$ does not belong to $[\gamma]$ (which is union of the two exponential families based on Q and Q_G), this example contradicts Theorem 23.3 of [1] (p. 371 or p. 346 in English version) that claims existence of a minimizer§ of $D(R\|P)$ over $P \in [\gamma]$ whenever $R \ll Q$. It follows that Lemma 23.7, Theorem 23.4 and following remarks in [1] are false as well.

The above example indicates that the faces of $\mathsf{cc}(Q)$ (rather than the ponderable faces of $\mathsf{cs}(Q)$) might be the right starting points to build extended exponential families, in order to better understand closures ('boundary at infinity') of exponential families and to repair false statements in [1]. Our results in this directions are outlined in [2] and will be published elsewhere.

We conclude this paper by relating the ponderable faces of $\mathsf{cs}(Q)$ to the exposed faces of $\mathsf{cc}(Q)$, and by commenting on the relationship of our Theorem 1 to a result by Chentsov.

Lemma 11. *The mapping $\rho(F) = F \cap \mathsf{cc}(Q)$ is a bijection between the ponderable faces of $\mathsf{cs}(Q)$ and the exposed faces of $\mathsf{cc}(Q)$. In addition, $Q(F)$ is equal to $Q(\mathsf{cl}(\rho(F)))$.*

Proof. We will repeatedly use Lemma 1, 2 and 3 and their Corollaries without explicit reference.

A ponderable face F of $\mathsf{cs}(Q)$ is equal to $H \cap \mathsf{cs}(Q)$ for a nontrivial supporting hyperplane H of $\mathsf{cs}(Q)$. Then $\rho(F) = F \cap \mathsf{cc}(Q) = H \cap \mathsf{cc}(Q) = \mathsf{cc}(Q^H) = \mathsf{cc}(Q^F)$ is nonempty since $Q(F) > 0$, and therefore $\rho(F)$ is an exposed face of $\mathsf{cc}(Q)$. Taking closures we have $\mathsf{cs}(Q^F) = \mathsf{cl}(\rho(F))$. This and $\mathsf{cs}(Q^F) \subseteq F$ imply

$$Q(F) = Q^F(\mathbb{R}^d) = Q^F(\mathsf{cl}(\rho(F))) = Q(\mathsf{cl}(\rho(F)))\ .$$

Given an exposed face E of $\mathsf{cc}(Q)$, $E = H \cap \mathsf{cc}(Q)$, the set $H \cap \mathsf{cs}(Q)$ is an exposed face of $\mathsf{cs}(Q)$ containing E. Let $\pi(E)$ be the intersection of all exposed faces of $\mathsf{cs}(Q)$ containing E. Obviously, $E \subseteq \pi(E) \subseteq H \cap \mathsf{cs}(Q)$. It is easy to see that a nonempty intersection of exposed faces is an exposed face, using that any intersection of faces equals a finite intersection of at most d of these faces. Therefore, $\pi(E)$ is an exposed face of $\mathsf{cs}(Q)$ using also that $\pi(E) \supseteq \mathsf{cl}(E)$ implies $Q(\pi(E)) \geqslant Q(\mathsf{cl}(E)) > 0$. If G is an exposed face of $\mathsf{cs}(Q)$ contained in $\pi(E)$ and $Q(G) = Q(\pi(E))$, then

$$G \supseteq \mathsf{cs}(Q^G) \supseteq \mathsf{cc}(Q^G) = \mathsf{cc}(Q^{\pi(E)}) \supseteq \mathsf{cc}(Q^{\mathsf{cl}(E)}) = E\ .$$

§Chentsov [1] denotes the Kullback-Leibler divergence $D(R\|P)$ by $\boldsymbol{I}[\,P|R\,]$.

This means that $G = \pi(E)$, showing that $\pi(E)$ is a ponderable face of $\mathsf{cs}(Q)$.

It remains to show that π and ρ are mutually inverse mappings. If E is an exposed face of $\mathsf{cc}(Q)$, $E = H \cap \mathsf{cc}(Q)$, then $\pi(E) \subseteq H \cap \mathsf{cs}(Q)$, and therefore $\rho(\pi(E)) = E$.

If F is a ponderable face of $\mathsf{cs}(Q)$, $F = H \cap \mathsf{cs}(Q)$, then

$$G = \pi(\rho(F)) = \pi(F \cap \mathsf{cc}(Q)) = \pi(H \cap \mathsf{cc}(Q)) \subseteq F$$

and $\mathsf{cl}(\rho(F)) \subseteq G$. Using that $Q(F) = Q(\mathsf{cl}(\rho(F)))$, we deduce $Q(F) = Q(G)$. Since G is an exposed face contained in the ponderable face F, this implies $F = G$. $\square$

Remark 8. From our Theorem 1 and Lemma 11 we deduce that the number of ponderable faces is at most countable what is Lemma 21.6 of Chentsov [1]. The proof of this assertion outlined in [1] is based on a notion of *substantially ponderable* face. A ponderable face G of $\mathsf{cs}(Q)$ is substantial if $Q(G - \bigcup F) > 0$ where the union is over all ponderable faces F contained properly in G. No argument is provided however for the measurability of this union, that would be needed to turn the outline in [1] to a complete proof.

Acknowledgement. The authors are indebted to Károly Böröczky, jr., for calling their attention to the property of cyclic polytopes used in Example 2.

References

[1] Chentsov, N.N., *Statistical Decision Rules and Optimal Inference.* Translations of Mathematical Monographs Amer. Math. Soc., Providence – Rhode Island 1982 (Russian original: Nauka, Moskva, 1972).

[2] Csiszár, I. and Matúš, F., Information projections revisited, Book of Abstracts, *International Symposium on Information Theory*, June 25–30, 2000, Sorrento, Italy.

[3] Rockafellar, R.T., *Convex Analysis.* Princeton University Press, 1970.

[4] Ziegler, G.M., *Lectures on Polytopes.* Springer-Verlag, New York 1995.