Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Journal of Multivariate Analysis 92 (2005) 239–256
Journal of Multivariate Analysis
http://www.elsevier.com/locate/jmva

# Extension of the variance function of a steep exponential family

A. Hassairi* and A. Masmoudi

*Faculté des Sciences, Université de Sfax, B.P 802, Sfax, Tunisia*

Received 29 January 2002

**Abstract**

Let $F=\{P(m,F);\ m\in M_F\}$ be a multidimensional steep natural exponential family parameterized by its domain of the means $M_F$ and let $V_F(m)$ be its variance function. This paper studies the boundary behaviour of $V_F$. Necessary and sufficient conditions on a point $\overline{m}$ of $\partial M_F$ are given so that $V_F$ admits a continuous extension $V_F(\overline{m})$ to the point $\overline{m}$ . It is also shown that the existence of $V_F(\overline{m})$ implies the existence of a limit distribution $P(\overline{m},F)$ concentrated on an exposed face of $\overline{M}_F$ containing $\overline{m}$. The relation between $V_F(\overline{m})$ and $P(\overline{m},F)$ is established and some illustrating examples are given.
© 2003 Elsevier Inc. All rights reserved.

*Keywords:* Exponential family; Variance function; Steepness; Convex; Exposed face

## 1. Introduction

Exponential families have been a distinguished topic of theoretical statistics and probability theory for several decades. As they provide a general framework for many practical optimization problems in statistics, their behaviour on the boundary of natural parameter or mean parameter spaces is of theoretical interest. Our approach in the present work is based on the natural exponential families and their description through means (see [7]). In this context, the variance function of a natural exponential family (NEF) appears as the most appropriate tool, and so it has received a great deal of attention in the statistical literature. Its importance stems from the fact that it characterizes the family within the class of all natural

*Corresponding author.
*E-mail address:* abdelhamid.hassairi@fss.rnu.tn (A. Hassairi).

0047-259X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2003.09.010

exponential families [9,11]. Several classifications of NEFs by means of variance functions have been defined (see, [2,3,8,9]). Also, many characteristic properties of classes of distributions have been established using variance functions (see, [4,6]). Beside its role for the study of NEFs, the variance function itself has many nice intrinsic algebraic properties. Jørgensen [5] asked what properties are shared by all members of a natural exponential family $F$ in terms of its variance function $V_F$. He considered the one-dimensional version of the problem. In this case, $V_F$ is a real valued function and the domain of the means $M_F$ is an interval of $\mathbb{R}$ which, naturally, has at most two extremities. The problem involves studying the behaviour of $V_F$ at one of the extremities of $M_F$. We are concerned with multidimensional steep NEFs, that is, NEFs with the domain of their means equal to the interior of the convex hull of the support. This global assumption is satisfied in all reasonable multidimensional cases and it is justified from a technical point of view. The steepness enables to apply the convex analysis methods to the convex supports of NEFs. A natural problem within this approach is to identify the points of the boundary of the domain of the means where the variance function $V_F$ admits an extension. In Section 2, we show that $V_F$ extends continuously to a point on the boundary if and only if $V_F$ is bounded in its neighbourhood. We also show that this is equivalent to the existence of a special bounded neighbourhood of the point. Section 3 is devoted to continuous extensions of the mapping $m \mapsto P(m, F)$ in the weak topology. We prove that boundedness of $V_F$ is sufficient for these extensions and that limit distributions $P(\overline{m}, F)$ are concentrated on faces of the closure $\overline{M}_F$. In this section we also give the link between $V_F(\overline{m})$ and the variance of the limit distribution $P(\overline{m}, F)$. Proofs are postponed to Section 4.

## 2. Extension of a variance function

We introduce first some notation and review some basic concepts concerning exponential families and their variance functions. For more details, we refer the reader to [7].

For a positive Radon measure $\mu$ on $\mathbb{R}^d$, we denote

$$L_\mu : \mathbb{R}^d \to ]0, +\infty[ : \theta \mapsto \int_{\mathbb{R}^d} \exp\langle \boldsymbol{\theta}, \mathbf{x} \rangle \mu(dx)$$

the Laplace transform, where $\langle \theta, x \rangle$ is the ordinary scalar product of $\theta$ and $x$ in $\mathbb{R}^d$. Also we denote

$$\Theta(\mu) = interior\{\theta \in \mathbb{R}^d;\ L_\mu(\theta) < +\infty\},$$

$$k_\mu = \log L_\mu;$$

$k_\mu$ is the cumulant generating function of $\mu$.

The set $\mathcal{M}(\mathbb{R}^d)$ is defined as the set of positive measures $\mu$ that are not concentrated on an affine hyperplane and $\Theta(\mu)$ is not empty.

For $\mu$ in $\mathscr{M}(\mathbb{R}^d)$, the set of probabilities

$$F = F(\mu) = \{P(\theta,\mu) = \exp(\langle \theta, x\rangle - k_\mu(\theta))\mu(dx);\ \theta \in \Theta(\mu)\}$$

is called the natural exponential family (NEF) generated by $\mu$. Of course, $\mu$ and $\mu'$ in $\mathscr{M}(\mathbb{R}^d)$ are such that $F(\mu) = F(\mu')$ if and only if there exists $(a,b)$ in $\mathbb{R}^d \times \mathbb{R}$ such that

$$\mu'(dx) = \exp(\langle a, x\rangle + b)\mu(dx).$$

For $\mu$ in $\mathscr{M}(\mathbb{R}^d)$, $k_\mu$ is strictly convex and real analytic on $\Theta(\mu)$, so that $k'_\mu$ defines a diffeomorphism from $\Theta(\mu)$ to its image $M_F$ called the mean domain of $F$. Let $\psi_\mu : M_F \to \Theta(\mu)$ be the inverse function of $k'_\mu$ and, for $m$ in $M_F$, $P(m,F) = P(\psi_\mu(m),\mu)$. We denote by $L_s(\mathbb{R}^d)$ the set of symmetric linear maps of $\mathbb{R}^d$. For $m$ in $M_F$, $V_F(m) = k''_\mu(\psi_\mu(m))$ is the covariance operator of $P(m,F)$. The map from $M_F$ to $L_s(\mathbb{R}^d)$, defined by $V_F : M_F \to L_s(\mathbb{R}^d);\ m \mapsto V_F(m)$, is called the variance function of the natural exponential family $F$. It is easily proved that, for all $m$ in $M_F$, $\psi'_\mu(m)$ is the reciprocal of $V_F(m)$.

The importance of the variance function stems from the fact that it characterizes the family $F$ in the following sense: If $F$ and $F_1$ are two NEFs on $\mathbb{R}^d$ such that $V_F$ and $V_{F_1}$ coincide on an open subset of $M_F \cap M_{F_1}$, then $F = F_1$.

We examine the influence of an affine transformation on the elements of a NEF $F$. Let $\varphi$ be in the affine group of $\mathbb{R}^d$, i.e., $x \to \varphi(x) = a(x) + b$, where $b$ is in $\mathbb{R}^d$ and $a$ is in the linear group $GL(\mathbb{R}^d)$. The following facts are easily checked.

$$\varphi(F) = F(\varphi(\mu))$$

$$M_{\varphi(F)} = \varphi(M_F)$$

$$V_{\varphi(F)}(m) = aV_F(\varphi^{-1}(m))^t a \quad \text{for all } m \in M_{\varphi(F)}.$$

We evoke now the notion of steepness for a natural exponential family $F = F(\mu)$ generated by a measure $\mu$ belonging to $\mathscr{M}(\mathbb{R}^d)$. It was introduced by Barndorff-Nielsen [1]. We say that the family $F$ is steep if, for all $\overline{\theta} \in \partial\Theta(\mu)$,

$$\lim_{\theta \to \overline{\theta}} ||k'_\mu(\theta)|| = +\infty,$$

where $||.||$ denotes the Euclidean norm in $\mathbb{R}^d$.

Denote by $conv(supp(\mu))$ the closed convex hull of the support of $\mu$. We always have $M_F \subset int(conv(supp(\mu)))$. Barndorff-Nielsen [1] has shown that $F$ is steep if and only if $M_F = int(conv(supp(\mu)))$.

We now come to our results concerning the extension of the variance function of a steep NEF $F = F(\mu)$. We first prove a lemma concerning the behaviour of $\psi_\mu$. It is useful because of the link between the variance function $V_F$ and $\psi_\mu$.

**Lemma 2.1.** *Let F be a steep NEF on* $\mathbb{R}^d$. *Then for all* $\overline{m} \in \partial M_F$,

$$\lim_{m \to \overline{m}} ||\psi_\mu(m)|| = +\infty.$$

**Proof.** Suppose the contrary, so that there exists a sequence $(m_n)$ in $M_F$ such that $m_n \xrightarrow[n\to+\infty]{} \overline{m}$ and $\psi_\mu(m_n) = \theta_n \xrightarrow[n\to+\infty]{} \overline{\theta}$.

For $\theta$ in $\Theta(\mu)$, we define the following maps:

$$\varphi_n : [0,1] \to \mathbb{R}; \quad \lambda \mapsto \varphi_n(\lambda) = k_\mu((1-\lambda)\theta_n + \lambda\theta)$$

and

$$\varphi : ]0,1] \to \mathbb{R}; \quad \lambda \mapsto \varphi(\lambda) = k_\mu((1-\lambda)\overline{\theta} + \lambda\theta).$$

We consider two cases.

*Case* 1: $\overline{\theta} \in D(\mu) \backslash \Theta(\mu)$. As $\varphi_n$ is strictly convex, $\varphi'_n$ is strictly increasing on $[0,1]$. Hence we have, for all $\lambda \in ]0,1[$,

$$\varphi'_n(0) < \varphi'_n(\lambda) < \varphi'_n(1).$$

This implies that

$$\langle k'_\mu(\theta_n), \theta - \theta_n \rangle \leqslant \varphi'_n(\lambda) \leqslant \langle k'_\mu(\theta), \theta - \theta_n \rangle. \tag{2.1}$$

Since the sequence $s_n = k'_\mu(\theta_n)$ is bounded, there exists a subsequence $s_{n_k} = k'_\mu(\theta_{n_k})$ converging to a point $\overline{s}$.

Taking in (2.1) the limits when $k \to +\infty$, we obtain, for all $\lambda \in ]0,1[$,

$$\langle \overline{s}, \theta - \overline{\theta} \rangle \leqslant \varphi'(\lambda) \leqslant \langle k'_\mu(\theta), \theta - \overline{\theta} \rangle.$$

But, the steepness of the family $F$ implies

$$\lim_{\lambda \searrow 0} \varphi'(\lambda) = \lim_{\lambda \searrow 0} \langle k'_\mu((1-\lambda)\overline{\theta} + \lambda\theta), \theta - \overline{\theta} \rangle = -\infty,$$

which is a contradiction.

*Case* 2: $\overline{\theta} \in \overline{\Theta(\mu)} \backslash D(\mu)$. By the Rolle's formula there exists $\lambda_n \in ]0,1[$ such that

$$\varphi_n(1) - \varphi_n(0) = \varphi'_n(\lambda_n).$$

Since $\varphi'_n$ is strictly increasing on $[0,1]$, we have

$$\varphi_n(0) = \varphi_n(1) - \varphi'_n(\lambda_n) < \varphi_n(1) - \varphi'_n(0).$$

This implies that

$$k_\mu(\theta_n) < k_\mu(\theta) - \langle k'_\mu(\theta_n), \theta - \theta_n \rangle.$$

As $\overline{\theta} \notin D(\mu)$, Fatou lemma implies that

$$\lim_{n\to+\infty} k_\mu(\theta_n) = +\infty.$$

This contradicts the fact that the sequence $(k_\mu(\theta) - \langle k'_\mu(\theta_n), \theta - \theta_n \rangle)_n$ is bounded. Thus we have the desired conclusion. □

Recall that the norm of a linear endomorphism $A$ of $\mathbb{R}^d$ is defined by

$$||A|| = \sup\{||A.h||;\ ||h|| = 1\},$$

and let $B_f(x,r)$ and $B(x,r)$ denote the respective closed and open balls in $\mathbb{R}^d$, both with centre $x$ and radius $r$.

Next we give a technical result concerning the behaviour of the function $\psi_\mu$ in a neighbourhood of a boundary point in which the variance function is bounded.

**Proposition 2.2.** *Let $F$ be a steep NEF and let $\overline{m}$ be in $\partial M_F$. Suppose that there exists $\varepsilon > 0$ such that*

$$\sup_{m \in B_f(\overline{m},\varepsilon) \cap M_F} (||V_F(m)||) = \lambda_\varepsilon < +\infty.$$

*Then, for all $m_1$ and $m_2$ in $B_f(\overline{m},\varepsilon) \cap M_F$, we have*

$$||m_1 - m_2|| \leqslant \lambda_\varepsilon ||\psi_\mu(m_1) - \psi_\mu(m_2)||.$$

**Proof.** As

$$1 = ||V_F(m)(V_F(m))^{-1}|| \leqslant ||V_F(m)|| ||(V_F(m))^{-1}||,$$

for all $m \in B_f(\overline{m},\varepsilon) \cap M_F$,

$$||(V_F(m))^{-1}|| \geqslant \frac{1}{||V_F(m)||} \geqslant \frac{1}{\lambda_\varepsilon}.$$

Writing the Taylor expansion with integral remainder, we obtain

$$\begin{aligned}&\langle \psi_\mu(m_1) - \psi_\mu(m_2), m_1 - m_2 \rangle \\ &\quad = \int_0^1 (V_F((1-t)m_1 + tm_2))^{-1}(m_1 - m_2, m_1 - m_2)\,dt.\end{aligned}$$

For an arbitrary positive definite element of $L_s(\mathbb{R}^d)$ $A$ and all $x \in \mathbb{R}^d$,

$$\langle A^{-1}x, x \rangle \geqslant \frac{1}{||A||} ||x||^2.$$

This fact applied to $A = V_F((1-t)m_1 + tm_2)$ gives

$$\int_0^1 (V_F((1-t)m_1 + tm_2))^{-1}(m_1 - m_2, m_1 - m_2)\,dt \geqslant \frac{1}{\lambda_\varepsilon} ||m_1 - m_2||^2.$$

Hence

$$\langle \psi_\mu(m_1) - \psi_\mu(m_2), m_1 - m_2 \rangle \geqslant \frac{1}{\lambda_\varepsilon} ||m_1 - m_2||^2.$$

On the other hand, we have that

$$\langle \psi_\mu(m_1) - \psi_\mu(m_2), m_1 - m_2 \rangle \leqslant ||\psi_\mu(m_1) - \psi_\mu(m_2)|| ||m_1 - m_2||.$$

Consequently,

$$||\psi_\mu(m_1) - \psi_\mu(m_2)|| \geqslant \frac{1}{\lambda_\varepsilon} ||m_1 - m_2||. \qquad \square$$

The following result gives necessary and sufficient conditions for the extension of $V_F$ to a point $\overline{m}$ on the boundary $\partial M_F$ of $M_F$. We will say that the variance function $V_F$ can be extended to $\overline{m}$ if the limit of $V_F(m)$ exists for $m$ in $M_F$ tending to $\overline{m}$. In this case we write

$$V_F(\overline{m}) = \lim_{m \to \overline{m}} V_F(m).$$

**Theorem 2.3.** *Let $F$ be a steep NEF on $\mathbb{R}^d$ and let $\overline{m}$ be in $\partial M_F$. The following properties are equivalent*

(i) *$V_F$ extends to $\overline{m}$*
(ii) *$V_F$ is bounded in a neighbourhood of $\overline{m}$ in $M_F$*
(iii) *There exists a neighbourhood $B_f(\overline{m}, \varepsilon_o) \cap M_F$ such that the two conditions*

$$I(\overline{m}) = \begin{cases} \text{(a)} \quad \psi_\mu(B_f(\overline{m}, \varepsilon_o) \cap M_F) + B_f(o, r) \subset \Theta(\mu) \\ \text{(b)} \quad k'_\mu(\psi_\mu(B_f(\overline{m}, \varepsilon_o)) \cap M_F + B_f(o, r)) \text{ is bounded.} \end{cases}$$

*are satisfied.*

The following is an appealing corollary, because it concerns an important class of natural exponential families.

**Corollary 2.4.** *If the support of $F$ is bounded, then the variance function $V_F$ extends to $\overline{M}_F$.*

**Proof.** Since the support of $\mu$ is bounded, $\Theta(\mu) = \mathbb{R}^d$ and $M_F$ is bounded. Hence, from Theorem 2.3, $V_F$ extends to $\overline{M}_F$. $\square$

The following example illustrates the result in Theorem 2.3.

**Example 2.1.** Let $F$ be the NEF on $\mathbb{R}^2$ generated by the measure

$$\mu(dx, dy) = N(0,1)(dx) \otimes \delta_o(dy) + \delta_{(0,1)}(dx, dy).$$

Then

$$\Theta(\mu) = \mathbb{R}^2, \; L_\mu(\theta_1, \theta_2) = e^{\theta_2} + e^{\frac{1}{2}\theta_1^2}, \quad M_F = \mathbb{R} \times ]0, 1[ \quad \text{and}$$

$$V_F(m_1, m_2) = \begin{pmatrix} \dfrac{(1 - m_2)^2 + m_1^2 m_2}{1 - m_2} & -m_1 m_2 \\ -m_1 m_2 & m_2(1 - m_2) \end{pmatrix}.$$

For the point $\overline{m} = (m_1, 1) \in \partial M_F$, condition (a) in $I(\overline{m})$ is satisfied because $\Theta(\mu) = \mathbb{R}^2$, however, $\lim_{m \to \overline{m}} ||V_F(m)|| = +\infty$. This shows that condition (a) in $I(\overline{m})$ is not sufficient for the extension of the variance function $V_F$ to $\overline{m} \in \partial M_F$.

**Proposition 2.5.** *Let F be a steep NEF on* $\mathbb{R}^d$. *Then*

$$M_F^* = \{\overline{m} \in \partial M_F;\ \ V_F(\overline{m}) \ \textit{exists}\}$$

*is an open subset of* $\partial M_F$.

**Proof.** Suppose that $M_F^* \neq \emptyset$ and let $\overline{m}$ be an element of $M_F^*$. According to Theorem 2.3, the variance function is bounded in a neighbourhood $B(\overline{m}, \varepsilon_o) \cap M_F$ of $\overline{m}$. Let $x$ be an element of $B(\overline{m}, \varepsilon_o) \cap \partial M_F$ and take $\rho = \inf(\varepsilon_o, ||\overline{m} - x||)$. It is clear that the variance function is also bounded in $B(x, \rho) \cap M_F$. Again from Theorem 2.3, we deduce that $x \in M_F^*$. Hence $B(\overline{m}, \rho) \cap \partial M_F \subset M_F^*$. □

Note that the extended variance function is continuous on $M_F^* \cup M_F$.

Next we give another property of the points on the boundary of the domain of the means in which the variance function can be extended.

Recall that an exposed face $H$ of a convex set $C$ in $\mathbb{R}^d$ is defined as the intersection of $C$ with a non-trivial supporting hyperplane of $C$ (see [10, p. 162]).

With the same notation and hypothesis, in particular that $F = F(\mu)$ is a steep NEF, we have.

**Theorem 2.6.** *Let H be an exposed face of* $\overline{M}_F$ *and let* $\overline{m} \in H$. *If* $V_F(\overline{m})$ *exists, then* $\overline{m} \in conv(supp(\mu) \cap H)$.

Note that the converse of this theorem is not true. For instance, in the setting of Example 2.1, $H = \mathbb{R} \times \{1\}$ is an exposed face of $\overline{M}_F$, $\overline{m} = (0, 1) \in conv(supp(\mu) \cap H)$ and $\lim_{m \to \overline{m}} ||V_F(m)|| = +\infty$.

## 3. Limit distributions

Let $F$ be a steep NEF on $\mathbb{R}^d$ and let $\overline{m}$ be on the boundary $\partial M_F$ of $M_F$. We show that if $V_F$ extends to $\overline{m}$, then there exists a limit distribution $P(\overline{m}, F)$ concentrated on an exposed face $H$ containing $\overline{m}$. In this case, we determine the link between $V_F(\overline{m})$ and $P(\overline{m}, F)$.

**Definition 3.1.** Let $\overline{m}$ be a point of $\partial M_F$. If $P(m, F)$ has a tight limit $P(\overline{m}, F)$ when $m \in M_F$ tends to $\overline{m}$, then $P(\overline{m}, F)$ is called a limit distribution.

With this notation, we may consider the statistical model

$$\{P(m, F);\ \ m \in M_F \cup M_F^*\}$$

as a full NEF parametrized by the mean. It is the closure of the NEF $F$ in the weak topology.

**Theorem 3.1.** *Let $F$ be a steep NEF and let $\overline{m}\in\partial M_F$. Suppose $V_F$ extends to $\overline{m}$. Then*

(i) *there exists a limit distribution $P(\overline{m},F)$ in $\overline{m}$;*
(ii) *$V_F(\overline{m})$ is the variance of $P(\overline{m},F)$.*

For the proof of Theorem 3.1 we need the following lemma.

**Lemma 3.2.** *Let $\overline{m}\in\partial M_F$. Suppose that the two conditions* (a) *and* (b) *in $I(\overline{m})$ are satisfied; then for all $\theta$ in $B_f(o,r)$,*

(i) *for $m\in M_F$, the map defined on the interval $]0,1]$ by*
$$t\mapsto \langle k'_\mu(\psi_\mu(\overline{m}+t(m-\overline{m}))+\theta), m-\overline{m}\rangle$$
*extends by continuity to 0*
(ii) *the map $\varphi: m\mapsto k'_\mu(\psi_\mu(m)+\theta)$ extends by continuity to $\overline{m}$.*

**Proof.** (i) Consider the function
$$f(t)=\langle k'_\mu(\psi_\mu(\overline{m}+t(m-\overline{m}))+\theta), m-\overline{m}\rangle.$$
It is differentiable on $]0,1]$ with
$$f'(t)=\langle k''_\mu(\psi_\mu(\overline{m}+t(m-\overline{m}))+\theta)(V_F(\overline{m}+t(m-\overline{m})))^{-1}(m-\overline{m}), m-\overline{m}\rangle.$$
As $k_\mu$ is strictly convex, $k''_\mu(\psi_\mu(\overline{m}+t(m-\overline{m}))+\theta)$ and $(V_F(\overline{m}+t(m-\overline{m})))^{-1}$ are positive definite and symmetric. Hence so is
$$k''_\mu(\psi_\mu(\overline{m}+t(m-\overline{m}))+\theta)(V_F(\overline{m}+t(m-\overline{m})))^{-1}.$$
This implies that $f'(t)>0$ for all $t\in]0,1]$ and so $f$ is strictly increasing in $]0,1]$. As $f$ is bounded, it admits a finite limit at zero.

(ii) Follows from (i). □

**Proof of Theorem 3.1.** (i) Let $(m_n)_{n\in\mathbb{N}}$ be a sequence in $M_F$ converging to $\overline{m}$. We know that for all $\theta\in B_f(o,r)$,
$$k'_{P(m_n,F)}(\theta)=k'_\mu(\psi_\mu(m_n)+\theta).$$
According to Lemma 3.2, $\lim_{n\to+\infty} k'_{P(m_n,F)}(\theta)$ exists for all $\theta\in B_f(o,r)$; we denote it by $g(\theta)$. If we set
$$\lambda_o=\sup\{||k'_{P(m,F)}(\theta)||;\ \theta\in B_f(o,r) \text{ and } m\in B(\overline{m},\varepsilon_o)\},$$
then, for all $\theta$ and $\theta'\in B_f(o,r)$,
$$||k_{P(m_n,F)}(\theta)-k_{P(m_n,F)}(\theta')||\leqslant\lambda_o||\theta-\theta'||. \tag{3.1}$$

Hence, according to Ascoli theorem, $k_{P(m_n,F)}$ has a sub-sequence that converges uniformly on the closed ball $B_f(o,r)$ to a function $L$. The inequality (3.1) implies then that

$$||L(\theta) - L(\theta')|| \leqslant \lambda_o ||\theta - \theta'||,$$

and so $L$ is continuous on $B_f(o,r)$. We verify now that the limit $L$ is independent of the sequence $(m_n)_{n\in\mathbb{N}}$ and depends only on $\overline{m}$. In fact, let $(m_n)_{n\in\mathbb{N}}$ and $(m'_n)_{n\in\mathbb{N}}$ be two sequences in $M_F$ converging to $\overline{m}$ and let $L_1$ and $L_2$ be defined on $B_f(o,r)$ by

$$L_1(\theta) = \lim_{n\to+\infty} k_{P(m_n,F)}(\theta) \quad \text{and} \quad L_2(\theta) = \lim_{n\to+\infty} k_{P(m'_n,F)}(\theta).$$

Then $L_1$ and $L_2$ are differentiable on $B_f(o,r)$ and we have

$$L'_1(\theta) = g(\theta) \quad \text{and} \quad L'_2(\theta) = g(\theta).$$

This implies that

$$L_1(\theta) = L_2(\theta) + c.$$

Since $L_1(0) = L_2(0) = 0$, we obtain $c = 0$. Finally, Levy theorem guarantees the existence of a distribution $P(\overline{m},F)$ such that $k_{P(\overline{m},F)}(\theta) = L(\theta)$ for all $\theta \in B_f(o,r)$.

(ii) Since $\lim_{n\to+\infty} k''_{P(m_n,F)}(0) = V_F(\overline{m})$, we conclude that $V_F(\overline{m})$ is the variance of $P(\overline{m},F)$. □

Theorem 3.1 provides a sufficient condition for the existence of a limit distribution $P(\overline{m},F)$. We show now that, under this condition, $P(\overline{m},F)$ is concentrated on an exposed face of $\overline{M_F}$ containing $\overline{m}$.

**Proposition 3.3.** *Let $F$ be a steep NEF and let $\overline{m}$ be an element of an exposed face $H$ of $\overline{M}_F$. If $V_F(\overline{m})$ exists, then the limit distribution $P(\overline{m},F)$ is concentrated on $H$.*

**Proof.** Suppose that the dimension of $H$ is equal to $k$. Then, without loss of generality, we may assume that $H \subset \mathbb{R}^k \times \{0\}^{d-k}$.

Let $u$ be an exterior normal vector on $H$ such that,

$$H = \overline{M}_F \cap \{x \in \mathbb{R}^d;\ \langle x,u \rangle = 0\} \quad \text{and} \quad M_F \subset \{x \in \mathbb{R}^d; \langle x,u \rangle \leqslant 0\}.$$

For $\varepsilon > 0$, we set

$$A_\varepsilon(u) = \{x \in \mathbb{R}^d;\ \langle x,u \rangle \leqslant -\varepsilon\}.$$

As

$$P(m,F)(A_\varepsilon(u)) \leqslant -\frac{1}{\varepsilon} \int_{A_\varepsilon(u)} \langle x,u \rangle P(m,F)(dx) \leqslant -\frac{1}{\varepsilon} \langle m,u \rangle$$

and

$$\lim_{m\to\overline{m}} \langle m,u \rangle = \langle \overline{m},u \rangle = 0,$$

we obtain that

$$\lim_{m \to \overline{m}} P(m, F)(A_\varepsilon(u)) = 0.$$

Hence, for any $\varepsilon > 0$, we have

$$P(\overline{m}, F)(A_\varepsilon(u)) = 0.$$

Thus

$$P(\overline{m}, F) \text{ is concentrated on } H. \qquad \square$$

The following corollary gives a relation between the dimension of a face $H$ containing the point $\overline{m}$ and the rank of $V_F(\overline{m})$. In this case, $V_F(\overline{m})$ is a degenerate matrix.

**Corollary 3.4.** *Let $F$ be a steep NEF and let $H$ be an exposed face of $\overline{M_F}$ with dimension $k$. Suppose that, for $\overline{m} \in H$, $V_F(\overline{m})$ exists. Then*

(i) *rank* $V_F(\overline{m}) \leqslant k$;
(ii) $V_F(\overline{m}) = 0$ *if and only if* $P(\overline{m}, F) = \delta_{\overline{m}}$.

**Proof.** (i) Since $k''_{P(m,F)}(0) = V_F(m)$, we have $k''_{P(\overline{m},F)}(0) = V_F(\overline{m})$.

From Proposition 3.3, we know that the limit distribution $P(\overline{m}, F)$ is concentrated on $H$. This implies that

$$\mathit{rank}\ V_F(\overline{m}) \leqslant k.$$

(ii) $V_F(\overline{m}) = k''_{P(\overline{m},F)}(0) = 0$ if and only if the limit distribution $P(\overline{m}, F) = \delta_{\overline{m}}$. $\square$

Note that for $k = 0$, this means that $H = \{\overline{m}\}$ and so $V_F(\overline{m}) = 0$. This generalizes the result of Jørgensen [5] for a real NEF.

**Example 3.1.** The bivariate inverse Gaussian distribution (see [3]).

Let $F$ be the NEF generated by the measure $\mu$ concentrated on $]0, +\infty[\times \mathbb{R}$ defined by

$$\mu(dx, dy) = \frac{1}{2\pi x^2} exp\left[-\frac{1}{2x}(1 + y^2)\right] 1_{]0,+\infty[}(x)\, dx\, dy.$$

Its variance function is defined on $M_F = ]0, +\infty[\times \mathbb{R}$ by

$$V_F(m_1, m_2) = \begin{pmatrix} m_1^3 & m_1^2 m_2 \\ m_1^2 m_2 & (1 + m_2^2) m_1 \end{pmatrix}.$$

For an element $\overline{m} = (0, m_2)$ of the exposed face $H = \{0\} \times \mathbb{R}$, we have $\lim_{m \to \overline{m}} V_F(m) = 0$. Hence in this case, for all $\overline{m}$ in $H$, the limit distribution $P(\overline{m}, F)$ is equal to $\delta_{\overline{m}}$. This example shows in particular that the inequality *rank* $V_F(\overline{m}) \leqslant k$ in Corollary 3.4 may be strict.

We conclude this paragraph by another example.

**Example 3.2.** Let $F$ be the NEF on $\mathbb{R}^2$ generated by the measure

$$\mu(dx,dy) = \frac{1}{2}(\delta_{(0,1)}(dx,dy) + e^{-x}1_{\mathbb{R}_+}(x)dx \otimes \delta_o(dy)).$$

The Laplace transform of $\mu$ is defined on $\Theta(\mu) = ]-\infty, 1[\times\mathbb{R}$ by

$$L_\mu(\theta_1,\theta_2) = \frac{1}{2}\left(e^{\theta_2} + \frac{1}{1-\theta_1}\right).$$

The calculation of the variance function leads to

$$M_F = ]0,+\infty[\times]0,1[ \quad \text{and}$$

$$V_F(m_1,m_2) = \begin{pmatrix} \frac{m_1^2(1+m_2)}{1-m_2} & -m_1m_2 \\ -m_1m_2 & m_2(1-m_2) \end{pmatrix}.$$

We observe that the variance function $V_F$ does not extend to the face $H_1 = [0,+\infty[\times\{1\}$. However $V_F$ extends to the face $H_2 = [0,+\infty[\times\{0\}$ and for all $\overline{m} = (m_1, 0) \in H_2\backslash\{(0,0)\}$, we have

$$P(\overline{m},F)(dx) = \frac{1}{m_1}e^{-\frac{1}{m_1}x}1_{\mathbb{R}+}(x)dx$$

and

$$P(0,F) = \delta_0.$$

## 4. Proof of Theorems 2.3 and 2.6

We first prove in the following proposition a technical result which gives an equivalent version to the condition $I(\overline{m})$ in Theorem 2.3.

**Proposition 4.1.** *Let $\overline{m}$ be in $\partial M_F$. Then the following statements are equivalent*

(i) (iii) *in Theorem* 2.3 *is satisfied.*
(ii) *there exist $\varepsilon_o > 0$ and $\rho > 0$ such that*

$$\begin{cases} \text{(a)} & \psi_\mu(B_f(\overline{m},\varepsilon_o)\cap M_F) + B_f(o,\rho) \subset \Theta(\mu), \\ \text{(b)} & L_\mu(\psi_\mu(B_f(\overline{m},\varepsilon_o)\cap M_F) + B_f(o,\rho)) \text{ is bounded}. \end{cases}$$

**Proof.** (i) $\Rightarrow$ (ii) Assumption (i) implies that

$$\sup\{||k'_{P(m,F)}(\theta)||; \quad m \in B_f(\overline{m},\varepsilon_o)\cap M_F, \quad \theta \in B_f(o,r)\} = \lambda_o < +\infty.$$

This, with the finite increments theorem, implies that

$$|k_{P(m,F)}(\theta) - k_{P(m,F)}(0)| \leqslant \lambda_o||\theta||.$$

Since $k_{P(m,F)}(0)=0$ and $k_{P(m,F)}=\log(L_{P(m,F)})$, we obtain that, for all $m\in B_f(\overline{m},\varepsilon_o)\cap M_F$ and $\theta\in B_f(o,r)$, $L_{P(m,F)}(\theta)\leqslant e^{\lambda_o r}$, which proves (ii).

(ii) $\Rightarrow$ (i) Let

$$K_o=\sup\{L_{P(m,F)}(\theta);\ m\in B_f(\overline{m},\varepsilon_o)\cap M_F,\ \theta\in B_f(o,\rho)\}$$

and let $h$ be in the sphere $S(o,\frac{\rho}{3})$ with centre $o$ and radius $\frac{\rho}{3}$. If $\theta\in B_f(o,\frac{\rho}{3})$, then

$$\begin{aligned}|\langle L'_{P(m,F)}(\theta),h\rangle|\leqslant&\left|\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle\geqslant 0\}}\langle x,h\rangle e^{\langle\theta,x\rangle}P(m,F)(dx)\right|\\&+\left|\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle<0\}}\langle x,h\rangle e^{\langle\theta,x\rangle}P(m,F)(dx)\right|.\end{aligned}$$

Using the inequality $u\leqslant e^u$, we obtain that

$$\begin{aligned}&\left|\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle\geqslant 0\}}\langle x,h\rangle\ e^{\langle\theta,x\rangle}P(m,F)(dx)\right|\\&\quad\leqslant\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle\geqslant 0\}}e^{\langle\theta+h,x\rangle}P(m,F)(dx)\leqslant K_o\end{aligned}$$

$$\begin{aligned}&\left|\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle<0\}}\langle x,h\rangle e^{\langle\theta,x\rangle}P(m,F)(dx)\right|\\&\quad\leqslant\int_{\{x\in\mathbb{R}^d;\ \langle x,h\rangle<0\}}e^{\langle\theta-h,x\rangle}P(m,F)(dx)\leqslant K_o.\end{aligned}$$

This implies

$$|\langle L'_{P(m,F)}(\theta),h\rangle|\leqslant 2K_o.$$

As the function $\theta\mapsto k_{P(m,F)}(\theta)$ is convex and $k'_{P(m,F)}(0)=m$, we have

$$k_{P(m,F)}(\theta)-k_{P(m,F)}(0)\geqslant\langle m,\theta\rangle.$$

Therefore

$$-k_{P(m,F)}(\theta)\leqslant-\langle m,\theta\rangle\leqslant||m||||\theta||\leqslant\frac{1}{3}\rho(\varepsilon_o+||\overline{m}||)=c.$$

As we have

$$\langle k'_{P(m,F)}(\theta),h\rangle=\int_{\mathbb{R}^d}\langle x,h\rangle e^{\langle\theta,x\rangle-k_{P(m,F)}(\theta)}P(m,F)(dx),$$

we deduce that

$$|\langle k'_{P(m,F)}(\theta),h\rangle|\leqslant 2K_oe^c.$$

Let now $u$ be in the unit sphere $S(o,1)$ and take $h=\frac{\rho}{3}u$. Then

$$|\langle k'_{P(m,F)}(\theta),u\rangle|\leqslant\frac{6K_oe^c}{\rho}.$$

Thus

$$||k'_{P(m,F)}(\theta)||\leqslant\frac{6K_oe^c}{\rho}$$

which proves (i). □

The proof of Theorem 2.3 relies on the following lemma. For a proof we can consult [7, p. 40].

**Lemma 4.2.** *Let $\mu_n$ be a sequence of probability measures on $\mathbb{R}^d$ which converges tightly to $\mu$. Consider $g:\mathbb{R}^d\to\mathbb{R}$ a continuous function such that, for all $\varepsilon>0$, there exists $A_\varepsilon>0$ such that, for all $A\geqslant A_\varepsilon$ and $n\in\mathbb{N}$,*

$$\int_{\{x\in\mathbb{R}^d;\ ||x||\geqslant A\}}|g(x)|\mu_n(dx)<\varepsilon.$$

*Then*

$$\lim_{n\to+\infty}\int_{\mathbb{R}^d}g(x)\mu_n(dx)=\int_{\mathbb{R}^d}g(x)\mu(dx).$$

**Proof of Theorem 2.3.** (iii) ⇒ (i) Let $\overline{m}$ be in $\partial M_F$ and let $(m_n)_{n\in\mathbb{N}}$ be a sequence in $M_F$ converging to $\overline{m}$. We will verify that, for $\varepsilon_o>0$, there exists, $n_o\in\mathbb{N}$ such that, for all $n\geqslant n_o$,

$$m_n\in B_f(\overline{m},\varepsilon_o)\cap M_F\quad\text{and}\quad\sup_{n\geqslant n_o}\int_{\mathbb{R}^d}||x||^2P(m_n,F)(dx)\ \text{is finite.}$$

In fact, let $A\in\mathbb{R}_+^*h\in\mathbb{R}^d$ and denote

$$E(A)=\{x\in\mathbb{R}^d;\ ||x||\geqslant A\},$$

$$E_+(A)=\{x\in\mathbb{R}^d;\ \langle x,h\rangle\geqslant 0\ \text{and}\ ||x||\geqslant A\},$$

$$E_-(A)=\{x\in\mathbb{R}^d;\ \langle x,h\rangle\leqslant 0\ \text{and}\ ||x||\geqslant A\}.$$

According to Proposition 4.1, we set

$$K_o=\sup\{L_{P(m,F)}(\theta);\ m\in B_f(\overline{m},\varepsilon_o)\cap M_F\ \text{and}\ \theta\in B_f(o,r)\}.$$

As $u^2 \leqslant e^u$ for $u \geqslant 0$, we obtain, for all $h$ in $B_f(o, \frac{r}{2})$,

$$\begin{aligned}\int_{E(A)} \langle x,h\rangle^2 P(m_n,F)(dx) &\leqslant \int_{E_+(A)} \langle x,h\rangle^2 P(m_n,F)(dx) \\ &\quad + \int_{E_-(A)} \langle x,h\rangle^2 P(m_n,F)(dx) \\ &\leqslant \int_{E_+(A)} e^{\langle x,h\rangle} P(m_n,F)(dx) \\ &\quad + \int_{E_-(A)} e^{-\langle x,h\rangle} P(m_n,F)(dx).\end{aligned}$$

So

$$\int_{E_+(A)} \langle x,h\rangle^2 P(m_n,F)(dx) \leqslant 2K_o. \tag{4.1}$$

Using Hölder inequality, we obtain

$$\begin{aligned}\int_{E_+(A)} e^{\langle x,h\rangle} P(m_n,F)(dx) &\leqslant (L_{P_{(m_n,F)}}(2h))^{\frac{1}{2}}(P(m_n,F)(E_+(A)))^{\frac{1}{2}} \\ &\leqslant \sqrt{K_o}(P(m_n,F)(E_+(A))^{\frac{1}{2}} \\ &\leqslant \sqrt{K_o}(P(m_n,F)(E(A))^{\frac{1}{2}}.\end{aligned}$$

Similarly, we have

$$\int_{E_-(A)} e^{-\langle x,h\rangle} P(m_n,F)(dx) \leqslant \sqrt{K_o}(P(m_n,F)(E(A))^{\frac{1}{2}}.$$

Hence

$$\int_{E(A)} \langle x,h\rangle^2 P(m_n,F)(dx) \leqslant 2\sqrt{K_o}(P(m_n,F)(E(A))^{\frac{1}{2}}.$$

As the Chebyshev inequality implies

$$P(m_n,F)(E(A)) \leqslant \frac{1}{A^2}\int_{\mathbb{R}^d} ||x||^2 P(m_n,F)(dx),$$

we obtain that

$$\int_{E(A)} \langle x,h\rangle^2 P(m_n,F)(dx) \leqslant \frac{2\sqrt{K_o}}{A}\left(\int_{\mathbb{R}^d} ||x||^2 P(m_n,F)(dx)\right)^{\frac{1}{2}}.$$

Let now $(e_i)_{1\leqslant i\leqslant d}$ be any orthonormal basis in $\mathbb{R}^d$ and make $h = \frac{r}{2}e_i$ in (4.1), for $i = 1, \ldots, d$. Then we obtain, for $x = \sum_{1\leqslant i\leqslant d} x_i e_i$,

$$\int_{\mathbb{R}^d} \langle x,h\rangle^2 P(m_n,F)(dx) = \frac{r^2}{4}\int_{\mathbb{R}^d} x_i^2 P(m_n,F)(dx) \leqslant 2K_o.$$

Therefore

$$\int_{\mathbb{R}^d} ||x||^2 P(m_n, F)(dx) = \sum_{1 \leqslant i \leqslant d} \int_{\mathbb{R}^d} x_i^2 P(m_n, F)(dx) \leqslant \frac{8dK_o}{r^2}. \tag{4.2}$$

This done, we have that

$$\int_{E(A)} \langle x, h \rangle^2 P(m_n, F)(dx) \leqslant \frac{2\sqrt{K_o}}{A} \left( \int_{\mathbb{R}^d} ||x||^2 P(m_n, F)(dx) \right)^{\frac{1}{2}} \leqslant \frac{C}{A}, \tag{4.3}$$

where $C = \frac{4K_o\sqrt{2d}}{r}$.

According to (4.3) and using the Hölder inequality, we have, for $h, k$ in $B_f(o, \frac{r}{4})$,

$$\int_{E(A)} |\langle x, h \rangle \langle x, k \rangle| P(m_n, F)(dx)$$

$$\leqslant \left( \int_{E(A)} \langle x, h \rangle^2 P(m_n, F)(dx) \right)^{\frac{1}{2}} \left( \int_{E(A)} \langle x, k \rangle^2 P(m_n, F)(dx) \right)^{\frac{1}{2}} \leqslant \frac{C}{A}.$$

It follows that for $\varepsilon > 0$, there exists $A_\varepsilon > 0$ such that, for all $A > A_\varepsilon$,

$$\int_{E(A)} |\langle x, h \rangle \langle x, k \rangle| P(m_n, F)(dx) < \varepsilon.$$

As, from Theorem 3.1, $P(m_n, F)$ converges tightly to $P(\overline{m}, F)$ when $n \to +\infty$, Lemma 4.2 implies that

$$\lim_{n \to +\infty} \int_{\mathbb{R}^d} \langle x, h \rangle \langle x, k \rangle P(m_n, F)(dx) = \int_{\mathbb{R}^d} \langle x, h \rangle \langle x, k \rangle P(\overline{m}, F)(dx).$$

Since

$$V_F(m_n)(h, k) = \int_{\mathbb{R}^d} \langle x, h \rangle \langle x, k \rangle P(m_n, F)(dx) - \langle m_n, h \rangle \langle m_n, k \rangle,$$

we deduce that $\lim_{n \to +\infty} V_F(m_n)$ exists and does not depend on the choice of the sequence $(m_n)_{n \in \mathbb{N}}$. More precisely, we have

$$V_F(\overline{m}) = \lim_{n \to +\infty} V_F(m_n) = k''_{P(\overline{m}, F)}(0).$$

(i) $\Rightarrow$ (ii) Since $V_F$ extends to $\overline{m}$, then $V_F$ is bounded in a neighbourhood of $\overline{m}$ in $M_F$.

(ii) $\Rightarrow$ (iii) Let $\overline{m} \in \partial M_F$. Suppose that there exists $\varepsilon > 0$ such that the variance function $V_F$ is bounded on $I_\varepsilon = B_f(\overline{m}, \varepsilon) \cap M_F$. This means that

$$\sup_{m \in I_\varepsilon} ||V_F(m)|| = \lambda_\varepsilon < +\infty.$$

On the other hand, the steepness of the family $F$ implies that

$$\partial(\psi_\mu(I_\varepsilon)) \subset \Theta(\mu).$$

For $m\in B_f(\overline{m},\frac{\varepsilon}{2})\cap M_F$, denote

$$r_m=\inf\{||\psi_\mu(m)-\theta||;\ \theta\in\partial(\psi_\mu(I_\varepsilon))\}.$$

Then

$$r_m=\inf\{||\psi_\mu(m)-\theta||;\ \theta\in\partial(\psi_\mu(I_\varepsilon))\}=\inf\{||\psi_\mu(m)-\theta||;\ \theta\in\partial(\psi_\mu(I_\varepsilon))\cap\Theta(\mu)\}.$$

As $\psi_\mu$ is a diffeomorphism from $M_F$ into $\Theta(\mu)$ we have that

$$\psi_\mu(S(\overline{m},\varepsilon)\cap M_F)=\partial(\psi_\mu(I_\varepsilon))\cap\Theta(\mu).$$

In fact, writing $\partial(I_\varepsilon)=\overline{I}_\varepsilon\cap\overline{I^c_\varepsilon}$, where $I^c_\varepsilon$ is the complement of $I_\varepsilon$ in $\mathbb{R}^d$, we have

$$\begin{aligned}\psi_\mu(\partial(I_\varepsilon)\cap M_F)&=\psi_\mu(\overline{I^c_\varepsilon},\cap\overline{I}_\varepsilon\cap M_F)\\&=\psi_\mu(\overline{I}_\varepsilon\cap M_F)\cap\psi_\mu(\overline{I^c_\varepsilon}\cap M_F)\\&=\overline{\psi_\mu(I_\varepsilon)}\cap\overline{(\psi_\mu(I_\varepsilon))^c}\cap\Theta(\mu)\\&=\partial(\psi_\mu(I_\varepsilon))\cap\Theta(\mu).\end{aligned}$$

On the other hand,

$$\partial(I_\varepsilon)\cap M_F=S(\overline{m},\varepsilon)\cap M_F.$$

Hence

$$r_m=\inf\{||\psi_\mu(m)-\psi_\mu(x)||;\ x\in S(\overline{m},\varepsilon)\cap M_F\}.$$

Finally we show that $\liminf_{m\to\overline{m}} r_m\neq 0$.

Suppose the contrary, that is, there exists a sequence $(m_n)_{n\in\mathbb{N}}$ in $M_F$ converging to $\overline{m}$ so that $\lim_{n\to+\infty} r_{m_n}=0$. Then, there exists $(x^k_n)_k\subset S(\overline{m},\varepsilon)\cap M_F$ such that

$$x^k_n\xrightarrow[k\to+\infty]{}x_n\quad\text{and}\quad||\psi_\mu(m_n)-\psi_\mu(x^k_n)||\xrightarrow[k\to+\infty]{}r_n.$$

For $x_n\notin M_F$, Lemma 2.1 implies that $\lim_{k\to+\infty}||\psi_\mu(x^k_n)||=+\infty$. This contradicts the fact that $||\psi_\mu(m_n)-\psi_\mu(x^k_n)||\xrightarrow[k\to+\infty]{}r_n$. Hence, for all $n\in\mathbb{N}, x_n\in M_F\cap S(\overline{m},\varepsilon)$ and so $(x_n)_{n\in\mathbb{N}}$ is bounded. We may then assume that $(x_n)_{n\in\mathbb{N}}$ converges to a point $x$ of $S(\overline{m},\varepsilon)\cap\overline{M}_F$.

Using the inequality in Proposition 2.2, we obtain

$$||\psi_\mu(x_n)-\psi_\mu(m_n)||\geqslant\frac{1}{\lambda_\varepsilon}||x_n-m_n||.$$

Letting $n\to+\infty$, we get

$$0\geqslant\frac{1}{\lambda_\varepsilon}||x-\overline{m}||=\frac{\varepsilon}{\lambda_\varepsilon},$$

which is a contradiction. Therefore

$$\liminf_{m\to\overline{m}} r_m=r\neq 0.$$

We can now verify that, for all $m$ in $B_f(\overline{m},\frac{\varepsilon}{2})\cap M_F$,

$$B(\psi_\mu(m),r)\subset\psi_\mu(I_\varepsilon)\subset\Theta(\mu).\tag{4.4}$$

Suppose that there exists $\theta$ in $B(\psi_\mu(m), r)\backslash\psi_\mu(I_\varepsilon)$. Then the segment $[\theta, \psi_\mu(m)]$ cuts $\psi_\mu(I_\varepsilon)$ and cuts also its complement. As $[\theta, \psi_\mu(m)]$ is connected,

$$[\theta, \psi_\mu(m)] \cap \partial(\psi_\mu(I_\varepsilon)) \neq \emptyset.$$

Let $\theta_o$ be an element of $[\theta, \psi_\mu(m)] \cap \partial(\psi_\mu(I_\varepsilon))$. We have that

$$r \leqslant r_m \leqslant ||\psi_\mu(m) - \theta_o|| \leqslant ||\theta - \psi_\mu(m)||$$

which contradicts the fact that $\theta \in B(\psi_\mu(m), r)$.

Inclusion (4.4) is then established. Thus, for all $m \in B_f(\overline{m}, \frac{\varepsilon}{2}) \cap M_F$, we have $\psi_\mu(m) + B(o, r) \subset \psi_\mu(I_\varepsilon)$, and so $\psi_\mu(B_f(\overline{m}, \frac{\varepsilon}{2}) \cap M_F) + B(o, r) \subset \psi_\mu(I_\varepsilon)$.

Taking the images by $k'_\mu$, we get

$$k'_\mu\left(\psi_\mu\left(B_f\left(\overline{m}, \frac{\varepsilon}{2}\right) \cap M_F\right) + B(o, r)\right) \subset I_\varepsilon = B_f(\overline{m}, \varepsilon) \cap M_F.$$

This ends the proof of Theorem 2.3 and we give now the proof of Theorem 2.6. □

**Proof of Theorem 2.6.** Let $\overline{m}$ be in $H\backslash conv(supp(\mu) \cap H)$ and suppose that $V_F(\overline{m})$ exists. According to Theorem 3.1, the existence of $V_F(\overline{m})$ guarantees the existence of a limit distribution $P(\overline{m}, F)$. From Proposition 3.3, we have

$$supp(P(\overline{m}, F)) \subset H.$$

Let us suppose that

$$supp(P(\overline{m}, F)) \subset supp(\mu) \cap H. \tag{4.5}$$

Then

$$conv(supp(P(\overline{m}, F))) \subset conv(supp(\mu) \cap H).$$

As $\overline{m} \notin conv(supp(\mu) \cap H)$, then $\overline{m} \notin conv(supp(P(\overline{m}, F)))$. This contradicts the fact that $k'_{P(\overline{m},F)}(0) = \overline{m} \in conv(supp(P(\overline{m}, F)))$.

It remains to prove (4.5). Suppose that $x$ is a point of $H$ which does not belong to $supp(\mu)$. Then there exists $\varepsilon_o > 0$ such that

$$\mu(B_f(x, \varepsilon_o)) = 0.$$

Choose $\varepsilon$ in $]0, \varepsilon_o]$ such that $P(\overline{m}, F)(S(x, \varepsilon)) = 0$ and let $(m_n)_{n\in\mathbb{N}}$ be a sequence in $M_F$ converging to $\overline{m}$. Since $P(m_n, F)$ converges tightly to $P(\overline{m}, F)$, then

$$P(m_n, F)(B_f(x, \varepsilon)) \underset{n\to+\infty}{\longrightarrow} P(\overline{m}, F)(B_f(x, \varepsilon)).$$

The fact that $P(m_n, F)(dx) = e^{\langle \psi_\mu(m_n), x\rangle - k_\mu(\psi_\mu(m_n))}\mu(dx)$ and $\mu(B_f(x, \varepsilon)) = 0$ implies that

$$P(m_n, F)(B_f(x, \varepsilon)) = 0, \quad \text{for all } n \in \mathbb{N}.$$

Therefore

$$P(\overline{m}, F)(B_f(x, \varepsilon)) = 0$$

and in particular $x \notin supp(P(\overline{m}, F))$. □

## Acknowledgments

We sincerely thank the Editor and a referee for various suggestions and comments.

## References

[1] O. Barndorff-Nielsen, Information and Exponential Families in Statistical Theory, Wiley, New York, 1978.
[2] M. Casalis, The 2d + 4 simple quadratic natural exponential families on $\mathbb{R}^d$, Ann. Statist. 24 (4) (1994) 1828–1854.
[3] A. Hassairi, La classification des familles exponentielles naturelles sur $\mathbb{R}^n$ par l'action du groupe linaire de $\mathbb{R}^{n+1}$, C. R. Acad. Sci. Paris sér. I, t. 315 (1992) 207–210.
[4] A. Hassairi, S. Lajmi, Riesz Exponential families on symmetric cones, J. Theoret. Probab. 14 (4) (2001) 927–948.
[5] B. Jørgensen, R. Martinez, M. Tsao, Asymptotic behaviour of the variance function, Scand. J. Statist. 21 (1994) 223–243.
[6] G. Letac, A characterization of the Wishart exponential families by an invariance property, J. Theoret. Probab. 2 (1) (1989) 71–86.
[7] G. Letac, Lectures on Natural exponential families and their variance functions, Monograf. Mat. 50, Instituto de Matematica Pura e Applicada, Rio de Janeiro, 1992.
[8] G. Letac, M. Mora, Natural real exponential families with cubic variance function, Ann. Statist. 18 (1990) 1–37.
[9] C.N. Morris, Natural exponential families with quadratic variance functions, Ann. Statist. 10 (1) (1982) 65–80.
[10] T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
[11] M.C.K. Tweedie, Functions of a statistical variate with given means, with special reference to Laplacian distribution, Proc. Cambridge Philos. Soc. 43 (1947) 41–49.

## Further reading

L.D. Brown, Fundamentals of Statistical Exponential Families, Inst. of Mathematics and Statistics, Lecture Notes—Monograph Series, Vol. 9, 1986.
C. Kokonendji, V. Seshadri, The Lindsay transform of natural exponential families, Canad. J. Statist. 22 (2) (1994) 259–272.
A. Masmoudi, Structures asymptotiques d'une fonction variance, C. R. Acad. Sci. Paris Sr. I, t. 239 (1999) 177–182.
G.K. Pedersen, Analysis Now, Springer, New York, 1989.
R. Webster, Convexity, Oxford University Press, Oxford, New York, Tokyo, 1994.