Bernoulli **12**(2), 2006, 371–379

A characterization of Poisson–Gaussian families by generalized variance

CÉLESTIN C. KOKONENDJI[1] and AFIF MASMOUDI[2]

[1]*Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques Appliquées, UMR 5142 CNRS, IUT STID, avenue de l'Université, 64000 Pau, France.*
E-mail: celestin.kokonendji@univ-pau.fr
[2]*Université de Sfax, Faculté des Sciences de Sfax, BP 802, Sfax, Tunisia.*
E-mail: afif.masmoudi@fss.rnu.tn

We show that if the generalized variance of an infinitely divisible natural exponential family $F = F(\mu)$ in a d-dimensional linear space is of the form $\det K''_\mu(\boldsymbol{\theta}) = \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{b} + c)$, then there exists k in $\{0, 1, \ldots, d\}$ such that F is a product of k univariate Poisson and $(d-k)$-variate Gaussian families. In proving this fact, we use a suitable representation of the generalized variance as a Laplace transform and the result, due to Jörgens, Calabi and Pogorelov, that any strictly convex smooth function f defined on the whole of $\mathbb{R}^d$ such that $\det f''(\boldsymbol{\theta})$ is a positive constant must be a quadratic form.

Keywords: affine variance function; determinant; infinitely divisible measure; Laplace transform; Monge–Ampère equation; r-reducibility

1. Introduction

It is well known that natural exponential families (NEFs) are characterized by their variance functions. In the past ten years, several authors have investigated the so-called *generalized variance*, that is, the determinant of the covariance matrix of an NEF F on $\mathbb{R}^d$ (see, for example, Kokonendji and Seshadri 1996; Hassairi 1999; Kokonendji and Pommeret 2001). While for $d \geqslant 2$ the variance function characterizes F, the generalized variance does not (see Example 1 below), as is the case on the real line, where the generalized variance coincides with the NEF variance. Letac (1989) and Koudou and Pommeret (2002) point out some particularities concerning the basic NEFs, which are the Gaussian and Poisson families. The notion of joint multidimensional Poisson–Gaussian NEFs was introduced by Letac (1989), who characterized these families through their affine variance function. Koudou and Pommeret (2002) presented another characterization of Poisson–Gaussian NEFs in terms of the stability of their finite convolution product.

The aim of this paper is to show that if an NEF F is generated by an infinitely divisible measure μ on $\mathbb{R}^d$ such that its generalized variance is $\det K''_\mu(\boldsymbol{\theta}) = \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{b})$ for some $\mathbf{b} \in \mathbb{R}^d$, then there exists $k \in \{0, 1, \ldots, d\}$ such that $F = F(\mu)$ is a product of k univariate Poisson and $(d-k)$-variate Gaussian NEFs. The present study is motivated by the following question: under what circumstances does generalized variance characterize an

NEF? This is an ambitious project. The tool for studying this problem has hitherto been a suitable representation of the generalized variance as a Laplace transform.

This paper is organized as follows. In Section 2 we recall some of the technical material that we need for this new characterization of Poisson–Gaussian NEFs. In Section 3 we state the main result and make some comments on it. Section 4 is devoted to its proof.

2. Preliminaries

The NEFs represent a very important class of distributions in both probability and statistical theory (Kotz *et al.* 2000, Chapter 54).

2.1. NEFs and generalized variance

Let $\mathcal{M}(\mathbb{R}^d)$ be the set of σ-finite positive measures μ on $\mathbb{R}^d$ not concentrated on an affine subspace of $\mathbb{R}^d$, with the Laplace transform of μ given by

$$L_\mu(\theta) = \int_{\mathbb{R}^d} \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x})\mu(\mathrm{d}\mathbf{x})$$

and such that the interior $\Theta(\mu)$ of the domain $\{\boldsymbol{\theta} \in \mathbb{R}^d;\, L_\mu(\boldsymbol{\theta}) < \infty\}$ is non-empty. Defining the cumulant function as $K_\mu(\boldsymbol{\theta}) = \log L_\mu(\boldsymbol{\theta})$, the NEF generated by $\mu \in \mathcal{M}(\mathbb{R}^d)$, denoted by $F = F(\mu)$, is the family of probability measures $\{P_{\boldsymbol{\theta},\mu}(\mathrm{d}\mathbf{x}) = \exp[\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x} - K_\mu(\boldsymbol{\theta})]\mu(\mathrm{d}\mathbf{x});\, \boldsymbol{\theta} \in \Theta(\mu)\}$. If $\mathbf{X}$ is a random vector distributed according to $P_{\boldsymbol{\theta},\mu}$, then $\mathbb{E}_{\boldsymbol{\theta}}(\mathbf{X}) = \partial K_\mu(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = K'_\mu(\boldsymbol{\theta})$ and $\mathrm{var}_{\boldsymbol{\theta}}(\mathbf{X}) = \partial^2 K_\mu(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^{\mathrm{T}}\partial\boldsymbol{\theta} = K''_\mu(\boldsymbol{\theta})$. The function $\mathbf{m}(\boldsymbol{\theta}) = K'_\mu(\boldsymbol{\theta})$ is a one-to-one transformation from $\Theta(\mu)$ onto $M_F = K'_\mu(\Theta(\mu))$ and thus $\mathbf{m} = \mathbf{m}(\boldsymbol{\theta})$ provides an alternative parametrization of the family $F = \{P(\mathbf{m}, F);\, \mathbf{m} \in M_F\}$, called the mean parametrization. Note that M_F depends only on F, and not on the choice of the generating measure μ of F. The variance matrix of $P(\mathbf{m}, F)$ can be written as a function of the mean parameter $\mathbf{m}$, $\mathbf{V}_F(\mathbf{m}) = K''_\mu(\boldsymbol{\theta})$, called the *variance function* of F. Together with the mean domain M_F, V_F characterizes F within the class of all NEFs. This leads Morris (1982) to establish the first classification of NEFs with quadratic variance function (QVF) on $\mathbb{R}$, containing six basic families, as normal and Poisson, up to affine transformation and convolution power. The multivariate concept of QVF was considered by Letac (1989): $\mathbf{V}_F(\mathbf{m}) = \mathbf{A}(\mathbf{m}, \mathbf{m}) + \mathbf{B}(\mathbf{m}) + \mathbf{C}$, where $\mathbf{A}(\mathbf{m}, \mathbf{m})$, $\mathbf{B}(\mathbf{m})$ and $\mathbf{C}$ are real symmetric $(d \times d)$ matrices of respectively bilinear, linear and constant elements in $\mathbf{m} \in M_F \subseteq \mathbb{R}^d$. Three special cases are the affine variance functions (AVF) with $\mathbf{V}_F(\mathbf{m}) = \mathbf{B}(\mathbf{m}) + \mathbf{C}$ (Letac 1989), the homogeneous QVF with $\mathbf{V}_F(\mathbf{m}) = \mathbf{A}(\mathbf{m}, \mathbf{m})$ (Casalis 1991), and the simple QVF with $\mathbf{V}_F(\mathbf{m}) = \alpha\mathbf{m}\mathbf{m}^{\mathrm{T}} + \mathbf{B}(\mathbf{m}) + \mathbf{C}$, where $\alpha \in \mathbb{R}$ (Casalis 1996); see Consonni *et al.* (2004) and Seshadri (1997) for some properties.

The generalized variance

$$\det \mathbf{V}_F(\mathbf{m}) = \det K''_\mu(\boldsymbol{\theta}) \tag{1}$$

of an NEF F was considered by Kokonendji and Seshadri (1996). Hassairi (1999) showed the following: if μ is an infinitely divisible measure generating F, then there exists a positive measure $\rho(\mu)$ on $\mathbb{R}^d$ such that

$$\det K''_\mu(\boldsymbol{\theta}) = L_{\rho(\mu)}(\boldsymbol{\theta}), \qquad \text{for all } \boldsymbol{\theta} \in \Theta(\mu). \tag{2}$$

In general, the generalized variance does not characterize the NEF.

Example 1. Let F_1 be the NEF on $\mathbb{R}^2$ generated by $\mu_1(\mathrm{d}x, \mathrm{d}y) = \frac{1}{2}\delta_{(0,1)}(\mathrm{d}x, \mathrm{d}y) + \frac{1}{2}\mathrm{e}^{-x}\mathbb{1}_{\mathbb{R}_+}(x)\mathrm{d}x \otimes \delta_0(\mathrm{d}y)$. Its variance function is given by

$$\mathbf{V}_{F_1}(\mathbf{m}) = \begin{bmatrix} m_1^2(1+m_2)(1-m_2)^{-1} & -m_1 m_2 \\ -m_1 m_2 & m_2(1-m_2) \end{bmatrix},$$

and $M_{F_1} = (0, \infty) \times (0, 1)$. Let F_2 be the NEF on $\mathbb{R}^2$ defined as the product of independent gamma and Poisson on the real line, with variance function $\mathbf{V}_{F_2}(\mathbf{m}) = \mathrm{diag}(m_1^2, m_2)$ on $M_{F_2} = (0, \infty)^2$. So we have $\det \mathbf{V}_{F_1}(\mathbf{m}) = \det \mathbf{V}_{F_2}(\mathbf{m}) = m_1^2 m_2$ with $M_{F_1} \neq M_{F_2}$, but also F_1 and F_2 are distinct.

We conclude this subsection by recalling the notion of type of NEF (see Kokonendji and Seshadri 1996, Definition 3.1) and by giving (without proof) the effect of determinant on type of NEF.

Definition 1. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$ and let*

$$\Lambda(\mu) = \{p > 0; \exists \mu_p \in \mathcal{M}(\mathbb{R}^d) : L_{\mu_p}(\boldsymbol{\theta}) = [L_\mu(\boldsymbol{\theta})]^p\}.$$

Two NEFs F_1 and F_2 are said to be of the same type if there exist $\mu \in \mathcal{M}(\mathbb{R}^d)$, $p \in \Lambda(\mu)$ and an affinity φ in $\mathbb{R}^d$ such that $F_1 = F(\mu)$ and $F_2 = F(\tilde{\mu})$, where $\tilde{\mu} = \varphi(\mu_p)$.

Proposition 1. *Let μ, $\tilde{\mu}$ be in $\mathcal{M}(\mathbb{R}^d)$, $F = F(\mu)$, $\tilde{F} = F(\tilde{\mu})$ and $\mathbf{m} \in M_F$.*

(i) *If there exists $(\mathbf{d}, e) \in \mathbb{R}^d \times \mathbb{R}$ such that $\tilde{\mu}(\mathrm{d}\mathbf{x}) = \exp\{\mathbf{d}^{\mathrm{T}}\mathbf{x} + e\}\mu(\mathrm{d}\mathbf{x})$, then $F = \tilde{F}$ with $\Theta(\tilde{\mu}) = \Theta(\mu) + \mathbf{d}$. Moreover, $\det \mathbf{V}_{\tilde{F}}(\tilde{m}) = \det \mathbf{V}_F(\mathbf{m})$ for $\tilde{\mathbf{m}} = \mathbf{m} \in M_F$.*

(ii) *If $\tilde{\mu} = \varphi_* \mu$ is the image measure of μ by $\varphi(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A}$ is a non-degenerate matrix $(d \times d)$ and $\mathbf{b} \in \mathbb{R}^d$, then $\det \mathbf{V}_{\tilde{F}}(\tilde{m}) = (\det \mathbf{A})^2 \det \mathbf{V}_F(\mathbf{m})$ for $\tilde{\mathbf{m}} = \mathbf{A}\mathbf{m} + \mathbf{b} \in \varphi(M_F)$.*

(iii) *If $\tilde{\mu} = \mu_p$ is the pth power measure of μ for $p \in \Lambda(\mu)$, then $\det \mathbf{V}_{\tilde{F}}(\tilde{\mathbf{m}}) = p^d \det \mathbf{V}_F(\mathbf{m})$ for $\tilde{\mathbf{m}} = p\mathbf{m} \in pM_F$.*

2.2. Reducible NEFs

Let $\mathbf{X}$ be a random vector distributed according to an NEF F on $\mathbb{R}^d$. If $\mathbf{X}$ can be partitioned into two independent subvectors $\mathbf{X}_{(1)}$ and $\mathbf{X}_{(2)}$ of dimensions d_1 and $d_2 = d - d_1$, respectively, each following an NEF distribution, then the family F is called reducible (e.g. Bar-Lev *et al.* 1994). The family is said to be *irreducible* if it is not reducible. More generally, we have:

Definition 2. *An NEF F on $\mathbb{R}^d$ is said to be r-reducible if there exists an integer $r = 1, \ldots, d$ such that F is the product of r independent irreducible NEFs $F_1, \ldots, F_r$ defined on $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_r}$ with $d = d_1 + \ldots + d_r$. In this case we write $F = F_1 \circledast \cdots \circledast F_r$, and $\mu = \mu_1 \circledast \cdots \circledast \mu_r$ for their corresponding generator.*

It follows that $K_\mu(\boldsymbol{\theta})$ of an r-reducible $F = F(\mu)$ can be written as $K_\mu(\boldsymbol{\theta}) = K_{\mu_1}(\boldsymbol{\theta}_{(1)}) + \ldots + K_{\mu_r}(\boldsymbol{\theta}_{(r)})$ with $\boldsymbol{\theta} = (\boldsymbol{\theta}^{\mathrm{T}}{}_{(1)}, \cdots, \boldsymbol{\theta}^{\mathrm{T}}{}_{(r)})^{\mathrm{T}} \in \Theta(\mu) = \Theta_1(\mu_1) \times \cdots \times \Theta_r(\mu_r)$, and where $\boldsymbol{\theta}_{(k)}$, K_{μ_k} and $\Theta_k(\mu_k)$ denote respectively the canonical parameter, the cumulant function and the canonical parameter space of the NEF $F_k = F(\mu_k)$, $\mu_k \in \mathcal{M}(\mathbb{R}^{d_k})$ for $k = 1, \ldots, r$. Thus, we have the trivial result:

Proposition 2. *Let $F(\mu)$ be an NEF generated by $\mu \in \mathcal{M}(\mathbb{R}^d)$. If $F(\mu)$ is r-reducible with $\mu = \mu_1 \circledast \cdots \circledast \mu_r$ (up to affinity), where $\mu_k \in \mathcal{M}(\mathbb{R}^{d_k})$ for $k = 1, \ldots, r$, then* $\det K''_\mu(\boldsymbol{\theta}) = \det K''_{\mu_1}(\boldsymbol{\theta}_{(1)}) \times \cdots \times \det K''_{\mu_r}(\boldsymbol{\theta}_{(r)})$, *for all* $\boldsymbol{\theta} = (\boldsymbol{\theta}^{\mathrm{T}}{}_{(1)}, \ldots, \boldsymbol{\theta}^{\mathrm{T}}{}_{(r)})^{\mathrm{T}} \in \Theta(\mu)$.

What about the converse of Proposition 2? Reasonably, we may consider the situation where $r = d$ and each irreducible $\mu_k \in \mathcal{M}(\mathbb{R})$ is infinitely divisible. The assumption of infinite divisibility allows to use (2) as necessary and sufficient condition for each component μ_k.

2.3. Poisson–Gaussian NEFs

Let $(\mathbf{e}_1, \ldots, \mathbf{e}_d)$ be the canonical basis of $\mathbb{R}^d$. A random vector $\mathbf{X}$ on $\mathbb{R}^d$ is said to be a *Poisson–Gaussian* vector of order $k \in \{0, 1, \ldots, d\}$, written $\mathbf{X} \sim PG_k$, if $\mathbf{X} = P_1\mathbf{e}_1 + \ldots + P_k\mathbf{e}_k + Z$ where $P_1, \ldots, P_k, Z$ are independent random variables such that the P_j are Poisson, and Z is a Gaussian variable in $\mathbb{R}^{d-k}$ with given covariance matrix. Any PG_k family is the set of PG_k distributions up to affine transformation. For $k = 0$ we have the Gaussian distribution on $\mathbb{R}^d$, and for $k = d$ we have the (pure) Poisson on $\mathbb{R}^d$. We recall also the following Letac (1989) characterization: an NEF F on $\mathbb{R}^d$ has an AVF if and only if, up to an affine transformation, there exists $k \in \{0, 1, \ldots, d\}$ such that for all $\mathbf{m} = (m_1, \ldots, m_d) \in M_F = (0; \infty)^k \times \mathbb{R}^{d-k}$ the variance function is the $d \times d$ diagonal matrix

$$\mathbf{V}_F(\mathbf{m}) = \mathrm{diag}(m_1, \ldots, m_k, 1, \cdots, 1). \tag{3}$$

This NEF is PG_k with generating measure

$$\mu(\mathrm{d}\mathbf{x}) = \left(\sum_{j \in \mathbb{N}^k} \frac{1}{j!} \delta_j(\mathrm{d}x_1, \ldots, \mathrm{d}x_k) \right) \frac{\exp\left(-\frac{1}{2}\sum_{i=k+1}^d x_i^2\right)}{(2\pi)^{(d-k)/2}} (\mathrm{d}x_{k+1}, \ldots, \mathrm{d}x_d),$$

where δ_j is the Dirac mass at j with $\Theta(\mu) = \mathbb{R}^d$ and $K_\mu(\boldsymbol{\theta}) = \sum_{i=1}^k \mathrm{e}^{\theta_i} + \frac{1}{2}\sum_{i=k+1}^d \theta_i^2$. Then the generalized variance of μ is

$$\det K''_\mu(\boldsymbol{\theta}) = \exp(\theta_1 + \ldots + \theta_k) = \exp(\boldsymbol{\theta}^{\mathrm{T}} \mathbf{1}_k), \tag{4}$$

where $\mathbf{1}_k$ is the vector in $\mathbb{R}^d$ such that the first k components are 1 and the rest 0. From (1) and (3), expression (4) can be written as $\det \mathbf{V}_F(\mathbf{m}) = m_1 \cdots m_k$. We close this subsection by observing that each of the $d+1$ basic families of PG_k in $\mathbb{R}^d$ is infinitely divisible and d-reducible (up to affinity).

3. Result and comments

Let us first state our main result. We present its proof in Section 4.

Theorem 3. *Let $F = F(\mu)$ be an infinitely divisible NEF on $\mathbb{R}^d$. Assume that:*

(i) $\Theta(\mu) = \mathbb{R}^d$, *and*
(ii) *there exist $\mathbf{b} \in \mathbb{R}^d$ and $c \in \mathbb{R}$ such that*

$$\det K''_\mu(\boldsymbol{\theta}) = \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{b} + c), \tag{5}$$

for all $\boldsymbol{\theta}$.

Then there exists $k \in \{0, 1, \ldots, d\}$ such that F is of Poisson–Gaussian PG_k type.

Before commenting on this result, let us present the following consequences. For $n > d$ and $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{R}^d)^n$, let

$$v_{d+1}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{1}{(d+1)!} \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{d+1} \end{bmatrix}.$$

Let $\mu \in \mathcal{M}(\mathbb{R}^d)$ be a probability measure with $\Theta(\mu) = \mathbb{R}^d$. Consider the two probabilities in $(\mathbb{R}^d)^n$: $P_n = \mu \times \cdots \times \mu$ and $\tilde{P}_n = \beta v^2_{d+1} P_n$, where β is a normalization constant. Denote by Q_n and $\tilde{Q}_n$ the respective images of P_n and $\tilde{P}_n$ by the map $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \mapsto \mathbf{x}_1 + \ldots + \mathbf{x}_n$. A reformulation of the Kokonendji and Seshadri (1996, Theorem 2.2) result gives $L_{\tilde{Q}_n}(\boldsymbol{\theta}) = L_{Q_n}(\boldsymbol{\theta}) \det K''_\mu(\boldsymbol{\theta}) / \det K''_\mu(0)$. Then Q_n and $\tilde{Q}_n$ have, up to translation, the same distribution (obviously Poisson–Gaussian) if and only if μ is Poisson–Gaussian. Also, under condition (5) with $\mathbf{b} \neq \mathbf{0}$ ($\Leftrightarrow$ $k \neq 0$), we can show that the maximum likelihood and uniformly minimum variance and unbiased estimators of $\det K''_\mu(\boldsymbol{\theta})$ coincide (Kokonendji and Pommeret 2001).

Remark 1. The hypothesis $\Theta(\mu) = \mathbb{R}^d$ is crucial for many reasons. One of them is that it does not follow from the fact that the function $\boldsymbol{\theta} \mapsto \det K''_\mu(\boldsymbol{\theta})$ is analytic in $\mathbb{R}^d$ that $\boldsymbol{\theta} \mapsto K''_\mu(\boldsymbol{\theta})$ coincides with a positive definite analytic function in $\mathbb{R}^d$; this would imply $\Theta(\mu) = \mathbb{R}^d$ by the principle of maximal analyticity.

Remark 2. Equation (5) in K_μ is of the Monge–Ampère type (see Gutiérrez 2001) which is well known in the area of differential geometry. If $\mathbf{b} = \mathbf{0}$ in the right-hand side of (5) we have the basic Monge–Ampère equation solved using the Jörgens–Calabi–Pogorelov (JCP) result: any strictly convex smooth function f in $\mathbb{R}^d$ such that $\det f''(\boldsymbol{\theta}) = 1$ must be a quadratic form. The latter result was proved by Jörgens (1954) for $d = 2$ (for a comprehensive version, see also Kokonendji 1995), by Calabi (1958) for $d = 3, 4, 5$, and by Pogorelov (1972) for $d \geqslant 6$. A shorter and more analytical proof is given in Cheng and Yau (1986). Recently, Caffarelli and Li (2004) extended this characterization to positive periodic functions in the right-hand side of (5).

Remark 3. The characterization of Gaussian NEFs given by Kokonendji and Seshadri (1996, Theorem 3.5), making direct use of the JCP result without the assumption of infinite divisibility as in Theorem 3, where we consider $\mathbf{b} = \mathbf{0}$ in the right-hand side of (5) and then $k = 0$.

It is a challenging problem to show that there exists a non-quadratic cumulant function K_μ such that $\det K''_\mu(\boldsymbol{\theta}) = 1$ on $\Theta(\mu) \neq \mathbb{R}^d$. The following proof of Theorem 3 relies on the JCP result through equation (2): the positive measure $\rho(\mu)$ on $\mathbb{R}^d$ is associated with $\det K''_\mu$.

4. Proof of Theorem 3

We first recall the following notation and then state the Laplace formula for determinant calculations (see Muir 1960). If $\mathbf{A} = (a_{ij})_{1\leqslant i,j\leqslant d}$ is a $d \times d$ matrix and S is a subset of $\{1, 2, \ldots, d\}$ with $S^c = \{1, 2, \ldots, d\}\backslash S$, we denote by $\mathbf{A}_S$ the matrix $(a_{ij})_{(i,j)\in S\times S}$ with $\det \mathbf{A}_\varnothing = 1$.

Proposition 4. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $d \times d$ *matrices. Then*

$$\det(\mathbf{A} + \mathbf{B}) = \sum_{S\subset\{1,2,\ldots,d\}} \det \mathbf{A}_{S^c} \det \mathbf{B}_S.$$

We also need the following two propositions. The first is an elementary result, so its proof is omitted. The second is due to Bar-Lev *et al.* (1994, Lemma 4.1).

Proposition 5. *Let* $\mathbf{b} \in \mathbb{R}^d$ *and let* $\mu_1, \ldots, \mu_k$ *be* k *independent positive measures in* $\mathbb{R}^d$ *such that* $\mu_1 * \cdots * \mu_k = \delta_{\mathbf{b}}$. *Then there exist* $\mathbf{b}_1, \ldots, \mathbf{b}_k$ *in* $\mathbb{R}^d$ *such that* $\mu_i = \delta_{\mathbf{b}_i}$ *and* $\mathbf{b} = \mathbf{b}_1 + \ldots + \mathbf{b}_k$.

Proposition 6. *Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *map. Then* f *is an affine polynomial if and only if* $\partial^2 f(\boldsymbol{\theta})/\partial\theta_i^2 = 0$, *for* $i = 1, \ldots, d$.

We can now prove Theorem 3. Since μ is infinitely divisible, there exist a symmetric non-negative definite $d \times d$ matrix Σ with rank $r \leqslant d$ and a positive measure ν on $\mathbb{R}^d$ such that

$$K''_\mu(\boldsymbol{\theta}) = \Sigma + \int_{\mathbb{R}^d} \mathbf{x}\mathbf{x}^{\mathrm{T}} \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x})\nu(\mathrm{d}\mathbf{x}) \tag{6}$$

(Gikhman and Skorohod 1973: 342). We let $k = d - r$. For $S = \{i_1, i_2, \ldots, i_l\}$ with

$1 \leqslant i_1 < i_2 < \ldots < i_l \leqslant d$, a non-empty subset of $\{1, 2, \ldots, d\}$, and $\tau_S : \mathbb{R}^d \to \mathbb{R}^l$ the map defined by $\tau_S(\mathbf{x}) = (x_{i_1}, x_{i_2}, \ldots, x_{i_l})^{\mathrm{T}}$, we define by ν_S the image measure of

$$H_l(\mathrm{d}\mathbf{x}_1, \ldots, \mathrm{d}\mathbf{x}_l) = \frac{1}{l!}(\det[\tau_S(\mathbf{x}_1) \ \cdots \ \tau_S(\mathbf{x}_l)])^2 \nu(\mathrm{d}\mathbf{x}_1) \cdots \nu(\mathrm{d}\mathbf{x}_l)$$

by $\varphi_l : (\mathbb{R}^d)^l \to \mathbb{R}^d$, $(\mathbf{x}_1, \ldots, \mathbf{x}_l) \mapsto \mathbf{x}_1 + \ldots + \mathbf{x}_l$. Then the measure $\rho(\mu)$ in (2) can be expressed by using Proposition 4 and (6) as

$$\rho(\mu) = (\det \mathbf{A})\delta_0 + \sum_{\emptyset \neq S \subset \{1,2,\ldots,d\}} (\det \mathbf{A}_{S^c})\ \nu_S, \tag{7}$$

where $\mathbf{A}$ is the diagonal representation matrix of Σ in an orthonormal basis $\mathbf{e} = (\mathbf{e}_1, \cdots, \mathbf{e}_d)$ of $\mathbb{R}^d$ (Hassairi 1999).

Without loss of generality, we assume $c = 0$ in (5); so $\det K''_\mu(\boldsymbol{\theta}) = \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{b})$. From (2) we easily obtain $\rho(\mu) = \delta_{\mathbf{b}}$. If $\mathbf{b} = \mathbf{0}$ then we have $k = 0$, which is the JCP result. Henceforth, we assume $\mathbf{b} \neq \mathbf{0}$ and therefore $k \neq 0$. Thus, the expression (7) implies $\det(\mathbf{A}) = 0$ with $\mathbf{A} = \mathrm{diag}(0, \ldots, 0, \lambda_{k+1}, \ldots, \lambda_d)$, where $\lambda_i > 0$; and for all non-empty subsets S of $\{1, 2, \ldots, d\}$ there exist real numbers $\alpha_S \geqslant 0$ such that

$$(\det \mathbf{A}_{S^c})\ \nu_S = \left(\prod_{i \notin S} \lambda_i\right) \nu_S = \alpha_S \delta_{\mathbf{b}}. \tag{8}$$

The following lemma makes precise the measure ν of (8).

Lemma 7. *Let ν satisfy condition* (8) *and $S_0 = \{1, \ldots, k\}$. Let $J_{\mathbf{e}} = \{i \in \{1, \ldots, d\};$ $x_i^2 \nu(\mathrm{d}\mathbf{x}) \neq 0$, $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_d\mathbf{e}_d\}$. Then $J_{\mathbf{e}} = S_0$ and, for all $i \in S_0$, $x_i^2\nu(\mathrm{d}\mathbf{x}) = \beta_i \delta_{c_i\mathbf{e}_i}(\mathrm{d}\mathbf{x})$, where $\beta_i > 0$ and $c_i \neq 0$.*

Proof. We proceed in three steps.

Step 1. $J_{\mathbf{e}} \supseteq S_0$. Indeed, suppose that there exists $i \in S_0$ such that $i \notin J_{\mathbf{e}}$. According to (6) we obtain

$$\frac{\partial^2 K_\mu(\boldsymbol{\theta})}{\partial \theta_i^2} = K''_\mu(\boldsymbol{\theta})(\mathbf{e}_i, \mathbf{e}_i) = \int_{\mathbb{R}^d} x_i^2 \exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x})\nu(\mathrm{d}\mathbf{x}) = 0,$$

which implies that $K''_\mu(\boldsymbol{\theta})$ is not positive definite. This leads to a contradiction.

Step 2. There exist k linearly independent vectors $(\mathbf{b}_i)_{i \in S_0}$ of $\mathbb{R}^d$ such that $\mathbf{b} = \sum_{i \in S_0} \mathbf{b}_{\mathrm{i}}$ and, for all $i \in S_0$, we have $x_i^2\nu(\mathrm{d}\mathbf{x}) = \beta_i\delta_{\mathbf{b}_i}(\mathrm{d}\mathbf{x})$ with $\beta_i > 0$ and $\tau_{S_0}(\mathbf{b}_i) = b_{ii}\mathbf{e}_i$. Indeed, let D_k be the set of $k \times k$ diagonal matrices. If we denote

$$\begin{aligned} H_k^{(1)}(\mathrm{d}\mathbf{x}_1, \ldots, \mathrm{d}\mathbf{x}_k) &= \mathbb{1}_{\{[\tau_{S_0}(\mathbf{x}_1)\cdots\tau_{S_0}(\mathbf{x}_k)] \in D_k\}}(\mathbf{x}_1, \ldots, \mathbf{x}_k) H_k(\mathrm{d}\mathbf{x}_1, \ldots, \mathrm{d}\mathbf{x}_k) \\ &= \prod_{i \in S_0} x_{ii}^2 \mathbb{1}_{\{\mathbf{x}_i \in \mathbb{R}^d; x_{ij} = 0, \forall j \in S_0 \setminus \{i\}\}}(\mathbf{x}_i)\nu(\mathrm{d}\mathbf{x}_i) \end{aligned}$$

and $H_k^{(2)} = H_k - H_k^{(1)}$, from step 1 we have $x_i^2\nu(\mathrm{d}\mathbf{x}) \neq 0$, for all $i \in S_0$ and hence there exists

$\alpha_1 > 0$ such that $\varphi_k(H_k^{(1)}) = \alpha_1 \delta_{\mathbf{b}}$. Since $H_k^{(1)}$ is the product of k independent positive measures, then $\varphi_k(H_k^{(1)})$ is a convolution product of k independent positive measures. According to Proposition 5 there exist $(\mathbf{b}_i)_{i\in S_0}$ in $\mathbb{R}^d$ such that $x_i^2 \nu(\mathrm{d}\mathbf{x}) = \beta_i \delta_{\mathbf{b}_i}(\mathrm{d}\mathbf{x})$ and $\mathbf{b} = \sum_{i\in S_0} \mathbf{b}_\mathrm{i}$. Note that $(\mathbf{b}_i)_{i\in S_0}$ is linearly independent and is an atom of $H_k^{(1)}$ such that $[\tau_{S_0}(\mathbf{b}_1) \cdots \tau_{S_0}(\mathbf{b}_k)] \in D_k$. Thus we obtain the last part of step 2.

Step 3. $J_\mathbf{e} = S_0$ and $\mathbf{b}_i = b_{ii}\mathbf{e}_i$, for all $i \in S_0$. Otherwise we have $J_\mathbf{e} = \{1, \ldots, k, i_{k+1}, \ldots, i_l\}$, where $l \geqslant k+1$. Similarly to step 2, there also exists $(\mathbf{b}_{i_{k+1}}, \ldots, \mathbf{b}_{i_l})$ such that $x_j^2 \nu(\mathrm{d}\mathbf{x}) = \beta_j \delta_{\mathbf{b}_j}(\mathrm{d}\mathbf{x})$ for all $j \in J_\mathbf{e}$ and $\mathbf{b} = \sum_{j\in J_\mathbf{e}} \mathbf{b}_j$. According to step 2, $\mathbf{b} = \sum_{i\in S_0} \mathbf{b}_\mathrm{i}$ implies $\mathbf{0} = \sum_{i\in J_\mathbf{e}\setminus S_0} \mathbf{b}_\mathrm{i}$ and leads to a contradiction, since $(\mathbf{b}_j)_{j\in J_\mathbf{e}}$ are linearly independent. Therefore $J_\mathbf{e} = S_0$. Since $x_j^2 \nu(\mathrm{d}\mathbf{x}) = 0$ for all $j \in \{k+1, \ldots, d\}$ then for all $i \in S_0$ we have $b_{ij}^2 \nu(\{\mathbf{b}_i\}) = 0$. But $(\mathbf{b}_i)_{i\in S_0}$ are atoms of ν, so for all $j \in \{k+1, \ldots, d\}$ we have $b_{ij}^2 = 0$. Finally, using $\tau_{S_0}(\mathbf{b}_i) = b_{ii}\mathbf{e}_i$ from step 2, we obtain $\mathbf{b}_i = b_{ii}\mathbf{e}_i$ and choose $b_{ii} = c_i$. Hence, the lemma is deduced from these steps. □

Introducing ν from Lemma 7 in (6), we obtain $\partial^2 K_\mu(\boldsymbol{\theta})/\partial\theta_i^2 = \beta_i \exp(c_i\theta_i)\mathbb{1}_{i\in S_0} + \lambda_i \mathbb{1}_{i\notin S_0}$. Let $B(\boldsymbol{\theta}) = \sum_{i\in S_0} c_i^{-2}\beta_i \exp(c_i\theta_i) + \frac{1}{2}\sum_{i\in S_0^c} \lambda_i^2\theta_i^2$. In view of Proposition 6 and the fact that $\partial^2(K_\mu - B)(\boldsymbol{\theta})/\partial\theta_i^2 = 0$, for $i = 1, \ldots, d$, $(K_\mu - B)(\boldsymbol{\theta})$ is an affine function on $\mathbb{R}^d$ and then

$$K_\mu(\boldsymbol{\theta}) = \sum_{i\in S_o} c_i^{-2}\beta_i \exp(c_i\theta_i) + \frac{1}{2}\sum_{i\in S_0^c} \lambda_i^2\theta_i^2 + \mathbf{u}^\mathrm{T}\boldsymbol{\theta} + a,$$

for $(\mathbf{u}, a) \in \mathbb{R}^d \times \mathbb{R}$. Hence $F(\mu)$ is of PG_k type. This concludes the proof of Theorem 3.

Acknowledgements

We sincerely thank the Editor, an Associate Editor and, in particular, the anonymous referee for their valuable comments. We are grateful to Professor Gaston M. N'Guérékata for his attentive reading of an earlier draft of this paper.

References

Bar-Lev, S., Bschouty, D., Enis, P., Letac, G., Lu, I. and Richard, D. (1994) The diagonal multivariate natural exponential families and their classification. *J. Theoret. Probab.*, **7**, 883–929.

Caffarelli, L. and Li, Y.Y. (2004) A Liouville theorem for solutions of the Monge-Ampère equation with periodic data. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **21**, 97–120.

Calabi, E. (1958) Improper affine hyperspheres of convex type and a generalization of a theorem by K. Jörgens. *Michigan Math. J.*, **5**, 105–126.

Casalis, M. (1991) Les familles exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique. *C.R. Acad. Sci. Paris Sér. I*, **312**, 537–540.

Casalis, M. (1996) The $2d+4$ simple quadratic natural exponential families on $\mathbb{R}^d$. *Ann. Statist.*, **24**, 1828–1854.

Cheng, S.Y. and Yau, S.T. (1986) Complete affine hypersurfaces. I. The completeness of affine metrics. *Comm. Pure Appl. Math.*, **39**, 839–866.

Consonni, G., Veronese, P. and Gutiérrez-Peña, E. (2004) Reference priors for exponential families with simple quadratic variance function. *J. Multivariate Anal.*, **88**, 335–364.

Gikhman, I.I. and Skorohod, A.V. (1973) *The Theory of Stochastic Processes II*. New York: Springer-Verlag.

Gutiérrez, C.E. (2001) *The Monge-Ampère Equation*. Boston: Birkhäuser.

Hassairi, A. (1999) Generalized variance and exponential families. *Ann. Statist.*, **27**, 374–385.

Jörgens, K. (1954) Über die Lösungen der Differentialgleichung $rt - s^2 = 1$. *Math. Ann.*, **127**, 130–134.

Kokonendji, C.C. (1995) Une caractérisation des vecteurs gaussiens n-dimensionnels. LAMIFA Technical Report No. 95/03, University of Amiens.

Kokonendji, C.C. and Pommeret, D. (2001) Estimateurs de la variance généralisée pour des familles exponentielles non gaussiennes. *C.R. Acad. Sci. Paris Sér. I*, **332**, 351–356.

Kokonendji, C.C. and Seshadri, V. (1996) On the determinant of the second derivative of a Laplace transform. *Ann. Statist.*, **24**, 1813–1827.

Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000) *Continuous Multivariate Distributions*, *Vol. 1: Models and Applications*, 2nd edn. New York: Wiley.

Koudou, A.E. and Pommeret, D. (2002) A characterization of Poisson-Gaussian families by convolution-stability. *J. Multivariate Anal.*, **81**, 120–127.

Letac, G. (1989) Le problème de la classification des familles exponentielles naturelles sur $\mathbb{R}^d$ ayant une fonction variance quadratique. In H. Heyer (ed.), *Probability Measures on Groups IX*, Lecture Notes in Math. 1306, pp. 194–215. Berlin: Springer-Verlag.

Morris, C.N. (1982) Natural exponential families with quadratic variance functions. *Ann. Statist.*, **10**, 65–80.

Muir, T. (1960) *A Treatise on the Theory of Determinants*. New York: Dover.

Pogorelov, A.V. (1972) On the improper convex affine hyperspheres. *Geom. Dedicata*, **1**, 33–46.

Seshadri, V. (1997) A property of natural exponential families in $\mathbb{R}^n$ with simple quadratic variance function. *J. Statist. Plann. Inference*, **63**, 351–361.

Received November 2004 and revised June 2005