ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Functional Analysis 235 (2006) 69–89

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Characterization of the simple cubic multivariate exponential families

A. Hassairi *, M. Zarai

Faculté des Sciences, Université de Sfax, B.P. 802, Sfax, Tunisie

Received 23 May 2005; accepted 1 September 2005

Available online 23 November 2005

Communicated by Paul Malliavin

Dedicated to Professor G. Letac on his 65th birthday

Abstract

The Letac–Mora class of real cubic natural exponential families has been characterized by a property of 2-orthogonality of an associated sequence of polynomials (see [G. Letac, M. Mora, Natural real exponential families with cubic variance functions, Ann. Statist. 18 (1990) 1–37; A. Hassairi, M. Zarai, Characterization of the cubic exponential families by orthogonality of polynomials, Ann. Probab. 32 (2004) 2463–2476]). The present paper introduces a notion of transorthogonality for a sequence of polynomial on $\mathbb{R}^d$ to extend the characterization to the multivariate version of the Letac–Mora class of real natural exponential families.

Keywords: Natural exponential family; Variance function; Simple cubic exponential family; Transorthogonality

1. Introduction

Let $F = \{P(m, F);\ m \in M_F\}$ be a natural exponential family (NEF) on $\mathbb{R}^d$ parameterized by its domain of the means M_F and let $V_F(m)$ denote the covariance operator of the probability distribution $P(m, F)$. Then the function $m \mapsto V_F(m)$, called the variance

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

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2005.09.005

function of the family F plays an important role in the study of F. Indeed, V_F characterizes the family F within the class of all natural exponential families. Furthermore several classifications of NEFs by the form of the variance function have been realized. Morris [9] describes the class of real NEFs such that $V_F(m)$ is a polynomial of degree at most 2 in the mean. Letac and Mora [8] have extended the work of Morris by classifying all real cubic NEFs with polynomial variance function of degree less than or equal to three. These classes cover the most common real distributions and have many interesting characteristic properties. A remarkable result concerning the Morris class is due to Feinsilver [2] who shows that certain sequences of polynomials naturally associated to a NEF $F(\mu)$ are μ-orthogonal if and only if the family is in the Morris class. This characterization has been extended to the Letac–Mora class of cubic NEFs by Hassairi and Zarai [4]. These authors have introduced a notion of 2-orthogonality for a sequence of polynomials and they have shown that the cubicity of the variance function of a NEF $F(\mu)$ is equivalent to the $(\mu - 2)$-orthogonality of the associated polynomials. In higher dimensions, the whole class of quadratic NEFs on a linear space E is not completely determined. However Casalis [1] has described the so called class of simple quadratic NEFs on $\mathbb{R}^d$ that is NEFs such that the variance function is of the form $V_F(m) = am \otimes m + B(m) + C$, with $m \otimes m(\theta) = \langle \theta, m \rangle m$ and B linear in m. This class represents the multivariate version of the Morris class and is also characterized by the orthogonality of the associated polynomials (see [2,5]). Let us mention that the simple quadratic NEFs are not the only NEFs which have a quadratic variance function. A Wishart NEF has an homogeneous polynomial variance function of degree two in the mean. The present paper is a continuation of the characteristic results concerning the classes of NEFs with polynomial variance functions and based on the notion of orthogonal polynomials. In fact, after the classification of simple quadratic NEFs, Hassairi [3] has introduced a class of natural exponential families on $\mathbb{R}^d$ with polynomial variance function of degree less than or equal to three in the mean. This class represents the multivariate version of the Letac–Mora class of real cubic NEFs and contains the simple quadratic ones, we will call it the class of simple cubic NEFs. Our aim is then to extend the notion of 2-orthogonality for a sequence of polynomials on $\mathbb{R}$ in a notion of transorthogonality for a sequence of polynomials on $\mathbb{R}^d$ and to show that a simple cubic NEF on $\mathbb{R}^d$ is characterized by the transorthogonality of the associated family of polynomials. Having done this, one may ask if we can go further, that is if we can find a property of orthogonality which characterizes the multivariate exponential families whose the variance function is a polynomial of degree $k \geqslant 4$. In fact, the characterizations of the simple quadratic and the simple cubic NEFs are motivated by the fact that these families are entirely described and that they cover the most common multivariate distributions. In these classes, the variance function is not any quadratic or any cubic polynomial, it has a specific form, and the most important in the characterizations is that the property of transorthogonality implies that the family is not only cubic but simple cubic and that the property of ordinary orthogonality implies that the family is not only quadratic but simple quadratic. Out of the simple cubic and the Wishart NEFs, we have yet no other examples of multivariate NEFs with polynomial variance functions and so we have not any idea about the form of a multivariate polynomial variance function of degree $k \geqslant 4$. Accordingly, we cannot define a property of orthogonality which correspond to a real situation and we think that assuming that the variance function has the general form of a polynomial of degree $k \geqslant 4$ is artificial and has no mathematical interest. Also talking about the characterization of a class of NEFs without any example of distribution belonging to this class, has no probabilistic interest. The paper is organized as follows. In Section 2, we present the class of simple cubic NEFs on $\mathbb{R}^d$. It is obtained from the simple quadratic NEFs by some specific action of the linear group $GL(\mathbb{R}^{d+1})$ on the space $\mathcal{L}_s(\mathbb{R}^d)$ of symmetric linear operators on $\mathbb{R}^d$. In Section 3, we state and prove our main result concerning the characterization of the simple cubic NEFs of $\mathbb{R}^d$.

2. Simple cubic exponential families

2.1. Natural exponential families

For an accurate presentation of the simple cubic natural exponential families, we first introduce some general facts concerning exponential families and their variance functions. Our notations are the ones used by Letac [6]. Let E be a linear vector space with finite dimension d, let E^* be its dual and let $E^* \times E \to \mathbb{R} : (\theta, x) \mapsto \langle \theta, x \rangle$ be the duality bracket. We denote by $\mathcal{L}_s(E^*, E)$ (respectively $\mathcal{L}_s(E, E^*)$) the space of the symmetric linear maps from E^* to E (respectively from E to E^*). If μ is a positive Radon measure on E, we denote

$$L_\mu(\theta) = \int_E \exp\langle \theta, x \rangle \mu(dx) \leqslant +\infty$$

its Laplace transform and we denote by $\Theta(\mu)$ the interior of the convex set $D(\mu) = \{\theta \in E^*;\ L_\mu(\theta) < \infty\}$. $\mathcal{M}(E)$ will denote the set of measures μ such that $\Theta(\mu)$ is not empty and μ is not concentrated on an affine hyperplane of E. If μ is in $\mathcal{M}(E)$, we also denote

$$k_\mu(\theta) = \log L_\mu(\theta), \quad \theta \in \Theta(\mu),$$

the cumulant function of μ. To each μ in $\mathcal{M}(E)$ and θ in $\Theta(\mu)$, we associate the probability distribution on E

$$P(\theta, \mu)(dx) = \exp\bigl\{\langle \theta, x \rangle - k_\mu(\theta)\bigr\}\mu(dx). \tag{2.1}$$

The set

$$F = F(\mu) = \bigl\{P(\theta, \mu);\ \theta \in \Theta(\mu)\bigr\}$$

is called the natural exponential family generated by μ. If μ and μ' are in $\mathcal{M}(E)$, then $F(\mu) = F(\mu')$ if and only if there exists (a, b) in $E \times \mathbb{R}$ such that $\mu'(dx) = \exp\{\langle a, x \rangle + b\}\mu(dx)$. Therefore, if μ is in $\mathcal{M}(E)$ and $F = F(\mu)$,

$$\mathcal{B}_F = \bigl\{\mu' \in \mathcal{M}(E);\ F(\mu') = F\bigr\}$$

is the set of basis of F.

The function k_μ is strictly convex and real analytic. Its first derivative k'_μ defines a diffeomorphism between $\Theta(\mu)$ and its image M_F. Since $k'_\mu(\theta) = \int x P(\theta, \mu)(dx)$, M_F is called the domain of the means of F. The inverse function of k'_μ is denoted by ψ_μ and setting $P(m, F) = P(\psi(m), \mu)$ the probability of F with mean m, we have

$$F = \{P(m, F);\ m \in M_F\},$$

which is the parameterization of F by the mean. The density of $P(m, F)$ with respect to μ,

$$f_\mu(x, m) = \exp\{\langle \psi_\mu(m), x\rangle - k_\mu(\psi_\mu(m))\}, \tag{2.2}$$

will play an important role in the definition of a sequence of polynomials associated to F. Now the covariance operator of $P(m, F)$ is denoted by $V_F(m)$. Clearly

$$V_F(m) = k''_\mu(\psi_\mu(m)) = (\psi'_\mu(m))^{-1} \in \mathcal{L}_s(E^*, E). \tag{2.3}$$

We now examine the influence of an affine transformation on the elements of a NEF. Let φ be in the affine group of E, i.e., $\varphi(x) = A(x) + b$ where A is in the linear group $GL(E)$ and b is in E, and let $F = F(\mu)$ be a NEF on E. The following facts are easily checked:

$$\begin{aligned}
&\varphi(F) = F(\varphi(\mu)),\\
&M_{\varphi(F)} = \varphi(M_F),\\
&V_{\varphi(F)}(m) = A V_F(\varphi^{-1}(m))\, {}^t A.
\end{aligned}$$

Eventually, we consider the influence of taking powers p of convolution in NEFs, where p is a positive number, but not necessarily an integer. If μ is in $\mathcal{M}(E)$, let us introduce the Jørgensen set

$$\Lambda = \{\alpha > 0;\ \text{there exists } \mu_\alpha \text{ in } \mathcal{M}(E) \mid \Theta(\mu_\alpha) = \Theta(\mu) \text{ and } k_{\mu_\alpha}(\theta) = \alpha k_\mu(\theta)\}.$$

Denoting $F_\alpha = F(\mu_\alpha)$ where α is in Λ, one has

$$M_{F_\alpha} = \alpha M_F \quad \text{and} \quad V_{F_\alpha}(m) = \alpha V_F\left(\frac{m}{\alpha}\right).$$

The transformation $F \mapsto F_\alpha$ is the Jørgensen transformation with parameter α.

The importance of the variance function stems from the fact that it characterizes the family F in the following sense: If F and F' are two NEFs such that $V_F(m)$ and $V_{F'}(m)$ coincide on a nonempty open subset of $M_F \cap M_{F'}$, then $F = F'$.

A NEF F on E is said to be reducible if there exist two subspaces E_1 and E_2 and two NEFs F_1 on E_1 and F_2 on E_2 such that $E = E_1 \oplus E_2$ and $F = F_1 \times F_2$. In this case, $M_F = M_{F_1} \times M_{F_2}$ and $V_F(m_1, m_2) = V_{F_1}(m_1) \otimes V_{F_2}(m_2)$.

2.2. Polynomials associated to a NEF

In the course of his study of the different characterization of the Morris class of quadratic real natural exponential families, Letac and Mora [8] has introduced the notion of polynomials associated to a real natural exponential family. He has used these polynomials to give an interpretation of the Meixner and Feinsilver results as characterizations of the Morris class. This notion has then been extended to the multivariate exponential families by Labeye-Voisin and Pommeret [5] who has also extended the Meixner and Feinsilver characterizations to the Casalis class of multivariate simple quadratic exponential families. Next, we show how these polynomials are defined, our notations are the ones used in [5–7].

Let $(e_1,\dots,e_d)$ be the canonical basis of $\mathbb{R}^d$. Then, for $n=(n_1,\dots,n_d)=\sum_{i=1}^d n_i e_i$ in $\mathbb{N}^d$, we set $|n|=n_1+\cdots+n_d$, $n!=n_1!\dots n_d!$, and for $x\in\mathbb{R}^d$, we write $x^n=x_1^{n_1}\dots x_d^{n_d}$.

A polynomial of degree k in x can be written,

$$Q(x)=\sum_{q\in\mathbb{N}^d;|q|\leqslant k} c_q x^q,$$

where at least one of the coefficient c_q such that $|q|=k$ is nonnull.

Consider now a natural exponential family $F=F(\mu)$ and take $\mu=P(m_0,F)$ with m_0 fixed in M_F. For $m\in M_F$, the density $f_\mu(\cdot,m)$ of $P(m,F)$ with respect to μ is given by (2.2) with $f_\mu(\cdot,m_0)$ equal to 1. The Taylor expansion in m of the analytic function $m\mapsto f_\mu(x,m)$ in the neighborhood of m_0 is

$$f_\mu(x,m)=\exp\bigl\{\bigl\langle\psi_\mu(m),x\bigr\rangle-k_\mu\bigl(\psi_\mu(m)\bigr)\bigr\}=\sum_{n\in\mathbb{N}^d}\frac{(m-m_0)^n}{n!}P_n(x), \tag{2.4}$$

where for all n in $\mathbb{N}^d$,

$$P_n(x)=\frac{\partial^{|n|}}{\partial m^n}f_\mu(x,m_0), \tag{2.5}$$

is a polynomial in x of degree $|n|$.

Hence, for each m_0 in M_F, we have a sequence of polynomials $P_n(x)$ associated to the natural exponential family $F=F(\mu)$. These polynomials have been used by Letac [7] and by Labeye-Voisin and Pommeret [5] for the characterization of the real and multivariate simple quadratic NEFs. We will use them to characterize the multivariate simple cubic NEFs.

2.3. Simple cubic exponential families

The class of the simple cubic exponential families in a real linear space E with dimension d is the natural extension of the Letac–Mora [5] class of real cubic exponential families. It has been defined by Hassairi and Zarai [4] as the class of exponential families obtained from the Casalis class of simple quadratic exponential families of E by the action of the linear group $G = GL(\mathbb{R}, E)$ on the natural exponential families of E. For instance, an element $g = \left[\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right]$ of the linear group G is defined by its blocks $(\alpha, \beta, \gamma, \delta)$ in $\mathbb{R} \times E^* \times E \times \mathcal{L}(E)$ ($\mathcal{L}(E)$ is the space of endomorphisms on E). Its respective actions on $\mathbb{R} \times E$ and $\mathbb{R} \times E^*$ are defined by

$$\begin{pmatrix} x_0 \\ x \end{pmatrix} \mapsto g \begin{pmatrix} x_0 \\ x \end{pmatrix} = \begin{pmatrix} \alpha x_0 + \langle \beta, x \rangle \\ \gamma x_0 + \delta(x) \end{pmatrix}$$

and

$$(k, \theta) \mapsto (k, \theta) g = \big(k\alpha + \langle \theta, \gamma \rangle, \beta k + \delta^*(\theta)\big),$$

where δ^* is the adjoint of δ. For g in G and m in E, we denote

$$d_g(m) = \alpha + \langle \beta, m \rangle \quad \text{and} \quad h_g(m) = \big(d_g(m)\big)^{-1}\big(\gamma + \delta(m)\big).$$

If U is an open set of E and g is in G, we denote

$$U_g = \big\{m \in E;\ d_g(m) > 0 \text{ and } h_g(m) \in U\big\}.$$

It is easily seen that if m is in E such that $d_g(m) = \alpha + \langle \beta, m \rangle \neq 0$, then the differential $h'_g(m)$ of h_g at m is an isomorphism of E (see Hassairi [3]).

Let g be an element of G and let U be a nonempty open subset of E such that $U_g \neq 0$. For $V : U \to \mathcal{L}_s(E^*, E)$, we define $T_g V : U_g \to \mathcal{L}_s(E^*, E)$ by

$$(T_g V)(m) = \big(d_g(m)\big)^{-1}\big(h'_g(m)\big)^{-1} V\big(h_g(m)\big)\big(h'_g(m)\big)^{*-1}.$$

In particular, if $E = \mathbb{R}$ the action of an element g of $GL(\mathbb{R}^2)$ on a real NEF F is given by

$$(T_g V_F)(m) = \frac{(\alpha + \beta m)^3}{(\alpha\delta - \beta\gamma)^2} V_F\left(\frac{\gamma + \delta m}{\alpha + \beta m}\right).$$

It is easy to see that if $\alpha = 1$ and $\beta = 0$, the image F_1 of F by the affinity $x \mapsto \delta(x) + \gamma$ satisfies

$$V_{F_1} = T_g V_F.$$

Also if α is in the Jørgensen set of F and if we denote by I the identity operator of the space E, then, for $g = \left[\begin{smallmatrix}\alpha & 0\\ 0 & I\end{smallmatrix}\right]$, T_g corresponds to the Jørgensen transformation with parameter α.

Let G_0 be the subgroup of G whose the elements are such that $\beta = 0$ and $\alpha > 0$. An element g_0 of G_0 may be written as a product of a Jørgensen transformation and an affine transformation. More precisely, we have

$$g_0 = \begin{bmatrix} \alpha & 0 \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} 1 & 0 \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \gamma/\alpha & \delta \end{bmatrix}\begin{bmatrix} \alpha & 0 \\ 0 & I \end{bmatrix}.$$

All the classifications of NEFs are done up to affine transformations and Jørgensen transformations, that is up to G_0-orbits. In fact, the Morris class of real quadratic NEFs has six G_0-orbits and the Letac–Mora class of real cubic NEFs has twelve G_0-orbits distributed in four G-orbits in the following way:

1st G-orbit	Gaussian 1		Inverse Gaussian m^3	
2nd G-orbit	Poisson m	Gamma m^2	Abel $m(1+m)^2$	Ressel $m^2(1+m)$
3rd G-orbit	Binomial $m(1-m)$	Negative-binomial $m(1+m)$	Takács $m(1+m)\big(1+\big(\frac{1+a}{a}\big)m\big)$ $a>0$	
4th G-orbit	Hyperbolic $1+m^2$		Large arcsine $m\big(1+2m+\frac{1+a^2}{a^2}m^2\big)$ $a>0$	Strict arcsine $m(1+m^2)$

Observe that each variance function of degree three is in the G-orbit of a quadratic one. More generally, if E is a linear space with dimension d and if we denote by $C(E)$ the class of simple quadratic NEFs on E, then $C(E)$ contains $(2d+4)$ G_0-orbits; $(d+1)$ Poisson–Gaussian G_0-orbits, $(d+1)$ negative multinomial-gamma G_0-orbits, a multinomial G_0-orbit, and an hyperbolic G_0-orbit built from particular mixtures of families of normal, Poisson, gamma, hyperbolic on $\mathbb{R}$ and (negative) multinomial distributions (see [1] for more details). We come now to the definition of the simple cubic natural exponential families. It is based on the fact that, for all g in G and F a simple quadratic exponential family, $T_g V_F$ is a polynomial of degree less than or equal to three. In fact, according to this result and to the observation concerning the real case, the class of simple cubic natural exponential families on E is defined as the class $G(C(E))$ of natural exponential families on E obtained by the action of the linear group $G = GL(\mathbb{R}, E)$ on the class $C(E)$ of simple quadratic NEFs on E. The classification of simple cubic NEFs is entirely determined by Hassairi [3], in $(d+3)$ G-orbits.

3. Characterization of the simple cubic NEFs on $\mathbb{R}^d$

In this section, we extend the characterization of the real cubic natural exponential families in the Feinsilver way (see [4]) to the class of simple cubic natural exponential families of $\mathbb{R}^d$.

3.1. Main results

We first introduce a specific norm on $\mathbb{R}^d$ which will serve in the definition of a notion of transorthogonality of a sequence of polynomials.

Let $(e_1,\dots,e_d)$ be the canonical basis of $\mathbb{R}^d$. Then it is easy to verify that the map $x = (x_1,\dots,x_d) \mapsto \|x\|_+ = \max(-\sum x_i^-, \sum x_i^+)$ defines a norm on $\mathbb{R}^d$ such that for $n \in \mathbb{N}^d$, $\|n\|_+ = |n| = n_1 + \cdots + n_d$.

Definition 3.1. A family $(Q_n)_{n\in\mathbb{N}^d}$ of polynomials on $\mathbb{R}^d$ is said to be μ-transorthogonal if for all n and $q\in\mathbb{N}^d$, $\int Q_n(x)Q_q(x)\mu(dx)=0$ when $\|n-q\|_+\geqslant \inf(|n|;|q|)$.

Note that the notion of transorthogonality generalizes the 2-orthogonality defined in Hassairi and Zarai [4], since when $d=1$, $\|n-q\|_+\geqslant \inf(|n|;|q|)$ is nothing but $n\geqslant 2q$ or $q\geqslant 2n$.

Next, we state our main result which characterizes the class of multivariate simple cubic natural exponential families by the transorthogonality of the associated polynomials. For the sake of simplification, we will restrict our statement to the simple cubic natural exponential families F such that for an element m_0 of M_F, $V_F(m_0)$ is diagonal. This is justified by the fact that, up to a linear transformation, any exponential family has this property and the class of simple cubic natural exponential families is invariant by linear transformations.

Theorem 3.1. *Let F be a irreducible natural exponential family on $\mathbb{R}^d$ and let m_0 be in M_F. Then F is simple cubic with $V_F(m_0)$ diagonal if and only if the sequence of polynomials $(P_n)_{n\in\mathbb{N}^d}$ defined by* (2.5) *is $P(m_0,F)$-transorthogonal.*

3.2. Proof of Theorem 3.1

($\Rightarrow$) If we write the variance function $V_F(m)$ as $V_F(m)=(V_{ij}(m))_{1\leqslant i,j\leqslant d}$, then from the simple cubicity of the NEF F, there exist real numbers α, $(v_i)_{1\leqslant i\leqslant d}$, $(a_i)_{1\leqslant i\leqslant d}$, and $(b_{ij}^k)_{0\leqslant i,j,k\leqslant d}$ such that

$$\begin{aligned}V_{ij}(m)={}&\sum_{s=1}^{d}a_s(m-m_0)_i(m-m_0)_j(m-m_0)_s+\alpha(m-m_0)_i(m-m_0)_j\\&+\sum_{l,k\in T}a_lb_{lj}^k(m-m_0)_i(m-m_0)_l+\sum_{l,k\in T}a_lb_{li}^k(m-m_0)_j(m-m_0)_l\\&+\sum_{k=1}^{d}b_{ij}^k(m-m_0)_k+v_i\delta_{ij}.\end{aligned}$$

On the other hand, we know that there exists $r>0$ such that, for all $m\in B(m_0,r)$ and for all $x\in\mathbb{R}^d$,

$$\sum_{n\in\mathbb{N}^d}\frac{(m-m_0)^n}{n!}P_n(x)=\exp\bigl\{\bigl\langle\psi_\mu(m),x\bigr\rangle-k_\mu\bigl(\psi_\mu(m)\bigr)\bigr\}.$$

Denoting $\theta=\psi_\mu(m)$, this may be written as

$$\exp(\theta x)=\Biggl(\sum_{n\in\mathbb{N}^d}\bigl(k_\mu'(\theta)-m_0\bigr)^n\frac{P_n(x)}{n!}\Biggr)\exp\bigl\{k_\mu(\theta)\bigr\}.$$

Taking the derivative with respect to θ_i gives

$$x_i \exp\langle\theta, x\rangle = \sum_{n\in\mathbb{N}^d} \frac{P_n(x)}{n!}\left(\sum_{s=1}^{d} n_s(m-m_0)^{n-e_s}\frac{\partial^2 k_\mu}{\partial\theta_i\partial\theta_s}(\theta) + (m-m_0)^n\frac{\partial k_\mu}{\partial\theta_i}(\theta)\right)$$
$$\times \exp\{k_\mu(\theta)\},$$

which is equivalent to

$$\sum_{n\in\mathbb{N}^d}\frac{(m-m_0)^n}{n!}x_i P_n(x) = \sum_{n\in\mathbb{N}^d}\frac{P_n(x)}{n!}\left(\sum_{s=1}^{d} n_s(m-m_0)^{n-e_s}V_{is}(m) + (m-m_0)^n m_i\right).$$

By identification, we get

$$\begin{aligned}
x_i P_n(x) = \mathbf{1}_{A_{e_i}}(n)\Bigg[&\sum_{s,k=1}^{d} a_k n_i(n_s-\delta_{is})(n_k-\delta_{ik}-\delta_{ks})P_{n-e_i-e_k}(x)\\
&+\sum_{s=1}^{d}\alpha n_i(n_s-\delta_{is})P_{n-e_i}(x) + \sum_{s,k,l=1}^{d} a_l b_{il}^k n_k(n_s-\delta_{ks})P_{n-e_k}(x)\\
&+\sum_{s,k,l=1}^{d} a_l b_{ls}^k n_i(n_k-\delta_{ik})P_{n+e_s-e_i-e_k}(x)\\
&+\sum_{s,k,l=1}^{d} a_l b_{is}^k n_k(n_s-\delta_{ks})P_{n+e_s-e_k}(x)\Bigg]\\
&+n_i P_{n-e_i}(x) + m_{0i}P_n(x) + v_i P_{n+e_i}(x),
\end{aligned}$$
$$\text{with } A_q = \{n\in\mathbb{N}^d \mid |n| \geqslant 2|q|\}. \tag{3.1}$$

Now, to show the transorthogonality of the polynomials $P_n(x)$, we need only to verify the two following facts.

(a) For all $n\in\mathbb{N}^d\setminus\{0\}$, $\int P_n(x)\mu(dx) = 0$.

(b) There exist real numbers $\alpha_{n,q}^s$, $\beta_{n,q}^s$ and $\lambda_{n,q}^s$ such that, for all $n, q\in\mathbb{N}^d\setminus\{0\}$ such that $|n|\geqslant|q|$ and $\|n-q\|_+\geqslant|q|$,

$$\begin{aligned}
x^q P_n(x) = \mathbf{1}_{A_q}(n)\Bigg[&\sum_{p\in\mathbb{N}^d;|p|=|q|}\alpha_{n,q}^p P_{n-q-p}(x) + \sum_{|n|-2|q|+1\leqslant|s|\leqslant|n|-|q|}\beta_{n,q}^s P_s(x)\Bigg]\\
&+\lambda_{n,q}^0 P_{n-q}(x) + \sum_{|n|-|q|+1\leqslant|s|\leqslant|n|+|q|}\lambda_{n,q}^s P_s(x),
\end{aligned}$$

where

$$\begin{cases}\lambda_{n,q}^0 = 0, & \text{if } n-q\notin\mathbb{N}^d,\\ \alpha_{n,q}^p = 0, & \text{if } n-q-p\notin\mathbb{N}^d.\end{cases} \tag{3.2}$$

Proof of (a). We first observe that

$$\int \frac{\partial}{\partial m} f_\mu(x,m)\mu(dx) = \psi'_\mu(m)\int (x-m) f_\mu(x,m)\mu(dx)$$
$$= \psi'_\mu(m)\int (x-m)P(m,F)(dx)$$
$$= 0.$$

Since, for all n, we have

$$\int \left| \frac{\partial^n}{\partial m^n} f_\mu(x,m)\right| \mu(dx) = \int \left|P_n(x-m)\right| f_\mu(x,m)\mu(dx)$$
$$= \int \left|P_n(x-m)\right| P(m,F)(dx) < +\infty$$

then

$$\int \frac{\partial}{\partial m}\left\{\frac{\partial^n}{\partial m^n} f_\mu(x,m)\right\}\mu(dx) = \frac{\partial}{\partial m}\int \frac{\partial^n}{\partial m^n} f_\mu(x,m)\mu(dx).$$

Hence we obtain that, for all $n \in \mathbb{N}^d \setminus \{0\}$,

$$\int \frac{\partial^n}{\partial m^n} f_\mu(x,m)\mu(dx) = 0.$$

This, for $m = m_0$, gives $\int P_n(x)\mu(dx) = 0$.

Proof of (b). We can write (3.1) as

$$x_i P_n(x) = \mathbf{1}_{A_{e_i}}(n)\left[\sum_{k=1}^{d} \alpha_{n,e_i}^{e_k} P_{n-e_i-e_k}(x) + \sum_{|s|=|n|-1} \beta_{n,e_i}^{s} P_s(x)\right]$$
$$+ \lambda_{n,e_i}^{0} P_{n-e_i}(x) + \sum_{|n|\leqslant|s|\leqslant|n|+1} \lambda_{n,e_i}^{s} P_s(x),$$

where

$$\begin{cases} \lambda_{n,e_i}^{0} = 0, & \text{if } n-e_i \notin \mathbb{N}^d,\\ \alpha_{n,e_i}^{e_k} = 0, & \text{if } n-e_i-e_k \notin \mathbb{N}^d. \end{cases} \tag{3.3}$$

For a fixed n in $\mathbb{N}^d \setminus \{0\}$, let us show by induction on $k \leqslant |n|$ that for all $q \in \mathbb{N}^d \setminus \{0\}$ such that $|q| = k$ and $\|n-q\|_+ \geqslant k$,

$$x^q P_n(x) = \mathbf{1}_{A_q}(n)\Bigg[\sum_{p\in\mathbb{N}^d;\,|p|=|q|} \alpha_{n,q}^p P_{n-q-p}(x) + \sum_{|n|-2|q|+1\leqslant|s|\leqslant|n|-|q|} \beta_{n,q}^s P_s(x)\Bigg] + \lambda_{n,q}^0 P_{n-q}(x) + \sum_{|n|-|q|+1\leqslant|s|\leqslant|n|+|q|} \lambda_{n,q}^s P_s(x),$$

where

$$\begin{cases} \lambda_{n,q}^0 = 0, & \text{if } n-q\notin\mathbb{N}^d, \\ \alpha_{n,q}^p = 0, & \text{if } n-q-p\notin\mathbb{N}^d. \end{cases}$$

For $k=1$, this is nothing but the equality (3.3).

Suppose now that (3.2) is true for $1\leqslant k<n$, and let (q',q) be in $(\mathbb{N}^d)^2$ such that $|q'|=k+1$, $q'=q+e_i$ and $\|n-q'\|_+\geqslant k+1$. Then, from the assumptions, we get

$$\begin{aligned}
x^{q'}P_n(x) &= x_i\big(x^q P_n(x)\big) \\
&= \mathbf{1}_{A_q}(n)\Bigg[\sum_{p\in\mathbb{N}^d;\,|p|=|q|} \alpha_{n,q}^p x_i P_{n-q-p}(x) + \sum_{|n|-2|q|+1\leqslant|s|\leqslant|n|-|q|} \beta_{n,q}^s x_i P_s(x)\Bigg] \\
&\quad + \lambda_{n,q}^0 x_i P_{n-q} + \sum_{|n|-|q|+1\leqslant|s|\leqslant|n|+|q|} \lambda_{n,q}^s x_i P_s(x) \\
&= \mathbf{1}_{A_q}(n)\Bigg[\sum_{p\in\mathbb{N}^d;\,|p|=|q|} \alpha_{n,q}^p \Bigg\{\mathbf{1}_{A_{e_i}}(n-q-p)\Bigg(\sum_{k=1}^d \alpha_{n-q-p,e_i}^{e_k} P_{n-q-p-e_i-e_k}(x) \\
&\quad + \sum_{|s|=|n-q-p|-1} \beta_{n-q-p,e_i}^s P_s(x)\Bigg) + \lambda_{n-q-p,e_i}^0 P_{n-q-p-e_i}(x) \\
&\quad + \sum_{|n-q-p|\leqslant|s|\leqslant|n-q-p|+1} \lambda_{n-q-p,e_i}^s P_s(x)\Bigg\} \\
&\quad + \sum_{|n|-2|q|+1\leqslant|s|\leqslant|n|-|q|} \beta_{n,q}^s \Bigg\{\mathbf{1}_{A_{e_i}}(s)\Bigg(\sum_{k=1}^d \alpha_{s,e_i}^{e_k} P_{s-e_i-e_k}(x) \\
&\quad + \sum_{|s'|=|s|-1} \beta_{s,e_i}^{s'} P_{s'}(x)\Bigg) + \lambda_{s,e_i}^0 P_{s-e_i}(x) + \sum_{|s|\leqslant|s'|\leqslant|s|+1} \lambda_{s,e_i}^{s'} P_{s'}(x)\Bigg\}\Bigg] \\
&\quad + \lambda_{n,q}^0 \Bigg\{\mathbf{1}_{A_{e_i}}(n-q)\Bigg(\sum_{k=1}^d \alpha_{n-q,e_i}^{e_k} P_{n-q-e_i-e_k}(x) + \sum_{|s''|=|n-q|-1} \beta_{n-q,e_i}^{s''} P_{s''}(x)\Bigg) \\
&\quad + \lambda_{n-q,e_i}^0 P_{n-q-e_i}(x) + \sum_{|n-q|\leqslant|s''|\leqslant|n-q|+1} \lambda_{n-q,e_i}^{s''} P_{s''}(x)\Bigg\}
\end{aligned}$$

$$+\sum_{|n|-|q|+1\leqslant|s|\leqslant|n|+|q|}\lambda_{n,q}^{s}\Bigg\{\mathbf{1}_{A_{e_i}}(s)\Bigg(\sum_{k=1}^{d}\alpha_{s,e_i}^{e_k}P_{s-e_i-e_k}(x)$$

$$+\sum_{|s'|=|s|-1}\beta_{s,e_i}^{s'}P_{s'}(x)\Bigg)+\lambda_{s,e_i}^{0}P_{s-e_i}(x)+\sum_{|s|\leqslant|s''|\leqslant|s|+1}\lambda_{s,e_i}^{s''}P_{s''}(x)\Bigg\}.$$

Hence there exist $(\alpha_{n,q'}^{s})_{s\in\mathbb{N}^d}$, $(\beta_{n,q'}^{s})_{s\in\mathbb{N}^d}$ and $(\lambda_{n,q'}^{s})_{s\in\mathbb{N}^d}$ such that

$$\begin{aligned}x^{q'}P_n(x)=\mathbf{1}_{A_{q'}}(n)\Bigg[&\sum_{p'\in\mathbb{N}^d||p'|=|q'|}\alpha_{n,q'}^{p'}x_iP_{n-q'-p'}(x)\\&+\sum_{|n|-2|q'|+1\leqslant|s|\leqslant|n|-|q'|}\beta_{n,q'}^{s}x_iP_s(x)\Bigg]\\&+\lambda_{n,q'}^{0}x_iP_{n-q'}(x)+\sum_{|n|-|q'|+1\leqslant|s|\leqslant|n|+|q'|}\lambda_{n,q'}^{s}x_iP_s(x),\end{aligned}$$

where

$$\begin{cases}\lambda_{n,q'}^{0}=\lambda_{n,q}^{0}\lambda_{n-q-p,e_i}^{0}=0, & \text{if } n-q'\notin\mathbb{N}^d,\\ \alpha_{n,q'}^{p'}=\alpha_{n,q}^{p}\alpha_{n-q-p,e_i}^{e_k}=0, & \text{if } n-q'-p'\notin\mathbb{N}^d.\end{cases}$$

($\Leftarrow$) We will do this in three steps. In the first step, we show that $V_F(m_0)$ is diagonal, in the second step, we show that F is cubic and in the third step we show that it is precisely simple cubic. The following property of symmetry of a variance function due to Letac [6] will play a crucial role in the proof.

$$\big(V_F'(m)\cdot V_F(m)(u)\big)(v)=\big(V_F'(m)\cdot V_F(m)(v)\big)(u),\quad \forall u,v\in\mathbb{R}^d.$$

In particular, when F is a cubic NEF on $\mathbb{R}^d$ with variance function

$$V_F(m)=X(m,m,m)+A(m,m)+B(m)+C,$$

the condition of symmetry is equivalent to the fact that, for all (u, v, m) in $(\mathbb{R}^d)^2 \times M_F$, the following terms are symmetric in (u, v):

- $X(X(m,m,m)u,m,m)v + X(m,X(m,m,m)u,m)v + X(m,m,X(m,m,m)u)v + A(X(m,m,m)u,m)v + A(m,X(m,m,m)u)v$;
- $X(B(m)u,m,m)v + X(m,B(m)u,m)v + X(m,m,B(m)u)v + A(A(m,m)u,m)v + A(m,A(m,m)u)v + B(X(m,m,m)u)v$;
- $X(B(m)u,m,m)v + X(m,B(m)u,m)v + X(m,m,B(m)u)v + A(A(m,m)u,m)v + A(m,A(m,m)u)v + B(X(m,m,m)u)v$;
- $X(Cu,m,m)v + X(m,Cu,m)v + X(m,m,Cu)v + A(B(m)u,m)v + A(m,B(m)u)v + B(A(m,m)u)v$;
- $A(m,Cu)v + A(Cu,m)v + B(B(m)u)v$;
- $B(Cu)v$.

(3.4)

Step 1. (2.4) allows us to claim the existence of $r > 0$ such that, for all m in $B(m_0, r)$ and for all x in $\mathbb{R}^d$,

$$f_\mu(x,m) = \exp\{\langle \psi_\mu(m), x\rangle - k_\mu(\psi_\mu(m))\} = \sum_{n\in\mathbb{N}^d} \frac{(m-m_0)^n}{n!} P_n(x).$$

The μ-transorthogonality of the polynomials (P_n) implies that, for all $(m, m') \in (B(m_0, r))^2$,

$$\begin{aligned} g(m,m') &= \exp\{k_\mu(\psi_\mu(m)+\psi_\mu(m')) - k_\mu(\psi_\mu(m)) - k_\mu(\psi_\mu(m'))\} \\ &= \int f_\mu(x,m) f_\mu(x,m')\mu(dx) \\ &= \int \sum_{n,q\in\mathbb{N}^d} \frac{(m-m_0)^q(m'-m_0)^n}{n!q!} P_n(x)P_q(x)\mu(dx) \\ &= 1 + \sum_{\|n-q\|_+<\inf(|n|;|q|)} \frac{(m-m_0)^q(m'-m_0)^n}{n!q!} \int P_n(x)P_q(x)\mu(dx) \\ &= 1 + \sum_{\|n-q\|_+<\inf(|n|;|q|)} a_{nq}(m-m_0)^q(m'-m_0)^n \end{aligned}$$

with $a_{nq} = \frac{1}{n!q!}\int P_n(x)P_q(x)\mu(dx)$.

Taking the derivative with respect to m_i, we get, for all $(m, m') \in (B(m_0, r))^2$,

$$\begin{aligned} &\langle k'_\mu(\psi_\mu(m)+\psi_\mu(m')) - k'_\mu(\psi_\mu(m)), \psi'_\mu(m)e_i\rangle g(m,m') \\ &\quad = \sum_{\|n-q\|_+<\inf(|n|;|q|)} q_i a_{nq}(m-m_0)^{q-e_i}(m'-m_0)^n. \end{aligned} \tag{3.5}$$

Making $m=m_0$ in (3.5), then since $\psi_\mu(m_0)=0$, we get

$$\big\langle (m'-m_0), \psi'_\mu(m_0)e_i\big\rangle = a_{e_i e_i}(m'-m_0)_i,$$

which is true for all $m' \in B(m_0,r)$. Denoting $\psi_\mu = (\psi_1, \psi_2, \dots, \psi_d)$, we get for $1 \leqslant k \leqslant d$,

$$\frac{\partial \psi_k}{\partial m_i}(m_0) = a_{e_i e_i}\delta_{ik}.$$

Hence $V_F(m_0) = (\psi'_\mu(m_0))^{-1}$ is a diagonal matrix with diagonal elements $(a^{-1}_{e_1 e_1}, \dots, a^{-1}_{e_d e_d}) = (v_1, \dots, v_d)$.

Step 2. Again we take the derivative of (3.1) with respect to m_j. For all (m,m') in $(B(m_0,r))^2$,

$$\begin{aligned}
&\Big[\big\langle k''_\mu\big(\psi_\mu(m)+\psi_\mu(m')\big)\big(\psi'_\mu(m)e_j\big) - k''_\mu\big(\psi_\mu(m)\big)\big(\psi'_\mu(m)\big)e_j, \psi'_\mu(m)e_i\big\rangle \\
&\qquad + \big\langle k'_\mu\big(\psi_\mu(m)+\psi_\mu(m')\big) - k'_\mu\big(\psi_\mu(m)\big), \big(\psi''_\mu(m)e_j\big)e_i\big\rangle + \big\langle k'_\mu\big(\psi_\mu(m)+\psi_\mu(m')\big) \\
&\qquad - k'_\mu\big(\psi_\mu(m)\big), \psi'_\mu(m)e_i\big\rangle\big\langle k'_\mu\big(\psi_\mu(m)+\psi_\mu(m')\big) - k'_\mu\big(\psi_\mu(m)\big), \psi'_\mu(m)e_j\big\rangle\Big] g(m,m') \\
&\quad = \sum_{\|n-q\|_+ < \inf(|n|;|q|)} q_i(q_j+\delta_{ij}) a_{nq}(m-m_0)^{q-e_i-e_j}(m'-m_0)^n.
\end{aligned} \tag{3.6}$$

If we set $m=m_0$ in (3.6), then the transorthogonality of the polynomials $P_n(x)$ gives, for all $m' \in B(m_0,r)$,

$$\begin{aligned}
&\big\langle k''_\mu\big(\psi_\mu(m')\big) - k''_\mu(0)\big(\psi'_\mu(m_0)\big)e_j, \psi'_\mu(m_0)e_i\big\rangle + \big\langle m'-m_0, \big(\psi''_\mu(m_0)e_j\big)e_i\big\rangle \\
&\qquad + \big\langle m'-m_0, \psi'_\mu(m)e_i\big\rangle\big\langle m'-m_0, \psi'_\mu(m)e_j\big\rangle \\
&\quad = \sum_{\|n-e_i-e_j\|_+ < \inf(|n|;2)} (1+\delta_{ij}) a_{n,e_i+e_j}(m'-m_0)^n \\
&\quad = \sum_{n\in\{e_i+e_l; e_j+e_l; e_i+e_j+e_s\}} (1+\delta_{ij}) a_{n,e_i+e_j}(m'-m_0)^n \\
&\quad = (1+\delta_{ij})\Bigg[\sum_{l=1}^{d} a_{e_i+e_l,e_i+e_j}(m'-m_0)^{e_i+e_l} + \sum_{l=1}^{d} a_{e_j+e_l,e_i+e_j}(m'-m_0)^{e_j+e_l} \\
&\qquad + \sum_{s=1}^{d} a_{e_i+e_j+e_s,e_i+e_j}(m'-m_0)^{e_i+e_j+e_s}\Bigg].
\end{aligned}$$

Therefore, for all $m' \in B(m_0,r)$,

$$V_{ij}(m') = \sum_{s=1}^{d} a_{ij}^{s}(m'-m_0)_i(m'-m_0)_j(m'-m_0)_s + \alpha_{ij}(m'-m_0)_i(m'-m_0)_j$$
$$+ \sum_{l=1}^{d} \beta_{ij}^{l}(m'-m_0)_i(m'-m_0)_l + \sum_{k=1}^{d} \beta_{ji}^{l}(m'-m_0)_j(m'-m_0)_l$$
$$+ \sum_{k=1}^{d} b_{ij}^{k}(m'-m_0)_k + V_{ij}(m_0),$$

where

$$\begin{cases} b_{ij}^{k} = -v_i v_j \frac{\partial^2 \psi_k}{\partial m_i \partial m_j}(m_0), \\ \beta_{ij}^{l} = (1+\delta_{ij}) v_i v_j a_{(e_i+e_l)(e_i+e_j)}, \\ \alpha_{ij} = \beta_{ij}^{j} + \beta_{ji}^{i} - 1, \\ a_{ij}^{s} = (1+\delta_{ij}) v_i v_j a_{(e_i+e_j+e_s)(e_i+e_j)}. \end{cases} \tag{3.7}$$

This implies that the degree of V_F is less than or equal to three and so we can write,

$$V_F(m) = X(m-m_0, m-m_0, m-m_0) + A(m-m_0, m-m_0) + B(m-m_0) + V_F(m_0),$$

where

$$X(m,m,m) = \left(m_i m_j \sum_{s=1}^{d} a_{ij}^{s} m_s\right)_{1 \leqslant i,j \leqslant d},$$
$$A(m,m) = \left(\alpha_{ij} m_i m_j + \sum_{l \neq j} \beta_{ij}^{l} m_i m_l + \sum_{l \neq i} \beta_{ji}^{l} m_j m_l\right)_{1 \leqslant i,j \leqslant d} \quad \text{and}$$
$$B(m) = \left(\sum_{k=1}^{d} b_{ij}^{k} m_k\right)_{1 \leqslant i,j \leqslant d}.$$

Step 3. In order to show that F is simple cubic, we will prove that if for all i, j in $T = \{1, \ldots, d\}$, there exists l_0 such that $a_{ij}^{l_0} \neq 0$, then F is simple cubic, and if for all i, j, s in T, $a_{ij}^{s} = 0$, then F is simple quadratic. Any other situation corresponds to a reducible variance function such that each block satisfies one of the two last conditions.

Writing the condition of symmetry of V_F for $u = e_i$ and $v = e_j$ (see (3.4)), we get the symmetry in (i, j) of the following expressions:

$$a_{kj}^{s}\left(a_{ij}^{l} + a_{si}^{l}\right) + a_{kj}^{l}\left(a_{ij}^{s} + a_{li}^{s}\right), \tag{3.8}$$

$$\sum_{s=1}^{d} a_{kj}^{s}\Biggl(\alpha_{ki}m_i m_k + \sum_{l\in T,l\neq k}\beta_{ik}^{l}m_i m_l + \sum_{l\neq i}\beta_{ki}^{l}m_k m_l\Biggr)m_s m_j$$
$$+\sum_{s=1}^{d} a_{kj}^{s}\Biggl(\alpha_{ij}m_i m_j + \sum_{l\in T,l\neq j}\beta_{ij}^{l}m_i m_l + \sum_{l\neq s}\beta_{ji}^{l}m_j m_l\Biggr)m_k m_s$$
$$+\sum_{s=1}^{d} a_{kj}^{s}\Biggl(\alpha_{si}m_i m_s + \sum_{l\in T,l\neq i}\beta_{si}^{l}m_s m_l + \sum_{l\neq s}\beta_{is}^{l}m_i m_l\Biggr)m_k m_j$$
$$+\alpha_{kj}\sum_{l=1}^{d} a_{ik}^{l}m_i m_j m_k m_l + \alpha_{kj}\sum_{l=1}^{d} a_{ij}^{l}m_i m_j m_k m_l$$
$$+\sum_{s,l\in T,l\neq j}\beta_{kj}^{l}\bigl(a_{ik}^{s}+a_{il}^{s}\bigr)m_i m_s m_k m_l + \sum_{s,l\in T,l\neq k}\beta_{jk}^{l}\bigl(a_{ij}^{s}+a_{il}^{s}\bigr)m_i m_j m_s m_l, \quad (3.9)$$

$$\sum_{l,s=1}^{d} a_{kj}^{l}b_{li}^{s}m_k m_j m_s + \sum_{l,s=1}^{d} a_{kj}^{l}b_{ij}^{s}m_k m_l m_s + \sum_{l,s=1}^{d} a_{kj}^{l}b_{ki}^{s}m_l m_j m_s + \sum_{l,s=1}^{d} b_{kj}^{l}a_{li}^{s}m_l m_i m_s$$
$$+\alpha_{kj}\alpha_{ij}m_i m_j m_k + \sum_{l\neq k}\alpha_{kj}\beta_{ik}^{l}m_i m_j m_l + \sum_{l\neq i}\alpha_{kj}\beta_{ji}^{l}m_l m_k m_j + \sum_{l\neq j}\alpha_{kj}\beta_{ij}^{l}m_i m_k m_l$$
$$+\sum_{l\neq j}\beta_{kj}^{l}\alpha_{li}m_i m_k m_l + \sum_{l\neq k}\beta_{jk}^{l}\alpha_{li}m_i m_j m_l + \sum_{l\neq k}\beta_{jk}^{l}\alpha_{ij}m_i m_j m_l$$
$$+\sum_{l\neq j,s\neq i}\beta_{kj}^{l}\beta_{li}^{s}m_s m_k m_l + \sum_{l\neq j,s\neq l}\beta_{il}^{s}\beta_{kj}^{l}m_s m_k m_i + \sum_{l\neq k,s\neq i}\beta_{li}^{s}\beta_{jk}^{l}m_s m_l m_j$$
$$+\sum_{l\neq k,s\neq l}\beta_{il}^{s}\beta_{jk}^{l}m_s m_i m_j + \sum_{l\neq j,s\neq k}\beta_{ik}^{s}\beta_{kj}^{l}m_s m_l m_i + \sum_{l\neq k,s\neq i}\beta_{ji}^{s}\beta_{jk}^{l}m_s m_l m_j$$
$$+\sum_{l\neq k,s\neq j}\beta_{ij}^{s}\beta_{jk}^{l}m_s m_l m_i, \quad (3.10)$$

$$v_i\Biggl(\delta_{ik}m_j\sum_{l=1}^{d} a_{kj}^{l}m_l + m_j m_k a_{kj}^{i}\Biggr) + \sum_{l,s=1}^{d} b_{kj}^{l}\Biggl(\alpha_{li}m_l m_i + \sum_{s\neq i}\beta_{li}^{s}m_l m_s + \sum_{s\neq l}\beta_{il}^{s}m_s m_i\Biggr), \quad (3.11)$$

$$v_i\Biggl(\delta_{ik}\alpha_{kj}m_j + \beta_{kj}^{i}m_k + \sum_{l=1}^{d}\beta_{kj}^{l}\delta_{ik}m_l + \sum_{s,s'=1}^{d} b_{kj}^{s}b_{si}^{s'}m_{s'}\Biggr), \quad (3.12)$$

and

$$v_i b_{kj}^{i} = v_j b_{ki}^{j}. \quad (3.13)$$

We will first show that if all the entries of V_F are cubic, then a_{ij}^s does not depend on i and j.

Lemma 3.1. *Suppose that for all $i, j \in T = \{1, \ldots, d\}$, the degree of V_{ij} is 3, then for $l_0 \in T$, we have the following statements*:

(a) *If $\forall s \in T \setminus \{l_0\}$ $a_{l_0 s}^{l_0} \neq 0$ then $\forall i, j \in T$ $a_{ij}^{l_0}$ does not depend on i and j.*
(b) *If $\forall s \in T$ $a_{l_0 s}^{l_0} = a_{ss}^{l_0} = 0$ then $\forall i, j \in T$ $a_{ij}^{l_0} = 0$.*
(c) *If there exists $i_0 \in T$ such that $a_{l_0 i_0}^{l_0} = 0$ then $\forall i, j \in T$ $a_{ij}^{l_0} = 0$.*

Proof. (a) Without loss of generality, we suppose that $l_0 = 1$. From (3.8), we have

$$a_{kj}^1\left(a_{ij}^1 + a_{1i}^1\right) = a_{ki}^1\left(a_{ij}^1 + a_{1j}^1\right).$$

Setting $k = i = 1$ and $j \neq 1$, we get

$$a_{1j}^1\left(a_{1j}^1 + a_{11}^1\right) = a_{11}^1\left(a_{1j}^1 + a_{1j}^1\right),$$

and it follows that $a_{1j}^1 = a_{11}^1$, $\forall j \neq 1$.

Hence, for $j = 1$ and $k, i \in T$, we have

$$a_{11}^1\left(a_{11}^1 + a_{11}^1\right) = a_{ki}^1\left(a_{11}^1 + a_{11}^1\right),$$

that is $a_{ki}^1 = a_{11}^1$.

(b) We use again the fact that $\forall i$, j and $k \in T$

$$a_{kj}^1\left(a_{ij}^1 + a_{1i}^1\right) = a_{ki}^1\left(a_{ij}^1 + a_{1j}^1\right).$$

As $\forall s \in T$ $a_{1s}^1 = a_{ss}^1 = 0$, then $a_{kj}^1 a_{ij}^1 = a_{ki}^1 a_{ij}^1$. Let $k = i$, since $a_{ii}^1 = 0$, then $a_{ij}^1 = 0$.

(c) From (b) it is sufficient to show that $\forall s \in T$ $a_{1s}^1 = a_{ss}^1 = 0$.

Suppose first that there exists $s_0 \in T$ such that $a_{1s_0}^1 \neq 0$. Then taking $s = l = j = k = 1$ and $i = s_0$ in (3.8), we get

$$a_{11}^1\left(a_{1s_0}^1 + a_{1s_0}^1\right) = a_{1s_0}^1\left(a_{1s_0}^1 + a_{11}^1\right).$$

Hence

$$a_{s_0 1}^1 = a_{11}^1 \neq 0. \tag{3.14}$$

On the other hand, for $s = l = j = 1$ and $i = i_0$ in (3.8) we get

$$a_{ki_0}^1\left(a_{1s_0}^1 + a_{11}^1\right) = a_{1k}^1\left(a_{1i_0}^1 + a_{1i_0}^1\right).$$

Thus, $a^1_{ki_0}=0$, for all $k\in T$. For $k=l=j=1$ and $i=i_0$ in (3.8) we get

$$a^1_{11}\big(a^s_{1i_0}+a^s_{1i_0}\big)+a^s_{11}\big(a^1_{1i_0}+a^1_{si_0}\big)=a^1_{1i_0}\big(a^s_{1i_0}+a^s_{11}\big)+a^s_{1i_0}\big(a^1_{1i_0}+a^1_{s1}\big).$$

Therefore

$$2a^1_{11}a^s_{1i_0}=a^s_{1i_0}a^1_{1s}.$$

From (3.14), $\forall s\in T\ a^s_{1i_0}=0$, i.e., degree of V_{1i_0} is less or equal two. This is in contradiction with the hypothesis and so, $\forall s\in T\ a^1_{1s}=0$ and we deduce that

$$a^1_{ij}a^1_{sj}=a^1_{ij}a^1_{si}. \tag{3.15}$$

Now suppose that there exists $s_0\in T$ such that $a^1_{s_0s_0}\neq 0$. Then setting $l=i=1$ and $k=j=s_0$ in (3.8), we get

$$a^s_{1s_0}\big(a^1_{1s_0}+a^1_{ss_0}\big)+a^1_{1s_0}\big(a^s_{1s_0}+a^s_{1s_0}\big)=a^1_{s_0s_0}\big(a^s_{1s_0}+a^s_{11}\big)+a^s_{s_0s_0}\big(a^1_{1s_0}+a^1_{11}\big)$$

and so

$$a^1_{s_0s_0}\big(a^s_{1s_0}+a^s_{11}\big)=a^s_{1s_0}a^1_{ss_0}. \tag{3.16}$$

If $a^1_{ss_0}\neq 0$ then, from (3.15) and (3.16), we get $a^1_{ss_0}=a^1_{s_0s_0}=a^1_{ss}\neq 0$. It follows that $a^s_{11}=0$.

If $a^1_{ss_0}=0$ then, from (3.15), we get $a^s_{1s_0}+a^s_{11}=0$ and from (3.16), $a^s_{ss_0}+a^s_{1s}=0$.

Taking in (3.16), $l=i=1$ and $j=s_0$, we obtain

$$a^s_{11}\big(a^s_{1s_0}+a^s_{ss_0}\big)=a^s_{1s_0}\big(a^s_{1s_0}+a^s_{1s}\big).$$

Thus $\forall s\in T\ a^s_{11}=0$, which contradicts the hypothesis. Hence from (b),

$$a^1_{ij}=0,\quad \forall i,j\in T. \qquad \square$$

The following corollary is an immediate consequence of Lemma 3.1.

Corollary 3.1. *With the same notations we have the following statements*:

(1) $\forall i,j,s\in T\ a^s_{ij}$ *does not depend on* (i,j).
(2) *For all* $i,j\in T$ *and* $l\in T\setminus\{i,j\}$, *the* β^l_{ij} *does not depend on* i.
(3) F *is simple cubic and*

$$
\begin{aligned}
V_{ij}(m) = {} & \sum_{s=1}^{d} a_s (m-m_0)_i (m-m_0)_j (m-m_0)_s + \alpha_{ij}(m-m_0)_i (m-m_0)_j \\
& + \sum_{l=1}^{d} h_j^l (m-m_0)_i (m-m_0)_l + \sum_{k=1}^{d} h_i^l (m-m_0)_j (m-m_0)_l \\
& + \sum_{k=1}^{d} b_{ij}^k (m-m_0)_k + v_i \delta_{ij}.
\end{aligned}
$$

Proof. (1) follows from (a) and (b) of Lemma 3.1.

(2) We know that there exists $s_0 \in T$ such that $a_{s_0} \neq 0$.

Take first $l = s_0$. Then identifying the coefficients of $m_{s_0}^2 m_i^2$ in (3.9), we get

$$
a_{s_0} \beta_{ij}^{s_0} = a_{s_0} \beta_{s_0 j}^{s_0} \quad \text{and it follows that} \quad \beta_{ij}^{s_0} = \beta_{s_0 j}^{s_0}. \tag{3.17}
$$

Now assuming that $l \in T \setminus \{i, j, s_0\}$ and identifying the coefficients of $m_{s_0} m_i^2 m_{s_0}$ in (3.9), we get

$$
a_l \beta_{ij}^{s_0} + a_{s_0} \beta_{ij}^l = a_l \beta_{s_0 j}^{s_0} + a_{s_0} \beta_{s_0 j}^l
$$

and so

$$
\beta_{ij}^l = \beta_{s_0 j}^l. \tag{3.18}
$$

Also identifying the coefficients of $m_{s_0}^2 m_j m_l$ in (3.9), we obtain

$$
2\left(a_{s_0} \beta_{jj}^l + a_l \beta_{jj}^{s_0}\right) = 2 a_{s_0} \beta_{s_0 j}^l + a_l \beta_{lj}^{s_0} + a_l \beta_{s_0 j}^{s_0}.
$$

Hence

$$
\beta_{s_0 j}^l = \beta_{jj}^l. \tag{3.19}
$$

Finally from (3.17)–(3.19), we conclude that β_{ij}^l does not depend on i.

(3) From (1) and (2) we can write

$$
\begin{aligned}
V_{ij}(m) = {} & \sum_{s=1}^{d} a_s (m-m_0)_i (m-m_0)_j (m-m_0)_s + \alpha_{ij}(m-m_0)_i (m-m_0)_j \\
& + \sum_{l=1}^{d} h_j^l (m-m_0)_i (m-m_0)_l + \sum_{k=1}^{d} h_i^l (m-m_0)_j (m-m_0)_l \\
& + \sum_{k=1}^{d} b_{ij}^k (m-m_0)_k + v_i \delta_{ij}
\end{aligned}
$$

with $a_s = a_{ij}^s$ and $h_j^l = \beta_{ij}^l$. $\square$

Lemma 3.2. *If there exists $i_0 \in T$ such that the degree of $V_{i_0 i_0}$ is less than or equal* 2*, then $\forall j \in T$ degree of $V_{i_0 j}$ is less than or equal* 2.

Proof. If we set $s = l$ in (3.8), we get

$$a_{kj}^s\big(a_{ij}^s + a_{si}^s\big) = a_{ki}^s\big(a_{ij}^s + a_{sj}^s\big). \tag{3.20}$$

Let $k = i_0$ and $i = s$ in (3.20). Then since $a_{i_0 i_0}^s = 0$, we obtain that $a_{si_0}^s = 0$.

Now setting $k = i = i_0$ in (3.20), then since $a_{i_0 i_0}^s = 0$ and $a_{si_0}^s = 0$, we get $a_{i_0 j}^s = 0$ $\forall s, j \in T$. Hence the degree of $V_{i_0 j}$ is less than or equal 2. □

Proposition 3.1. *If F is quadratic, then it is simple quadratic.*

Proof. If we suppose that F is quadratic, then for all $i, j, s \in T$, $a_{ij}^s = 0$ and it follows that

$$V_{ij}(m) = \alpha_{ij}(m - m_0)_i(m - m_0)_j + \sum_{l \in T\setminus\{j\}} \beta_{ij}^l (m - m_0)_i (m - m_0)_l + \sum_{l \in T\setminus\{i\}} \beta_{ji}^l (m - m_0)_j (m - m_0)_l + \sum_{k=1}^{d} b_{ij}^k (m - m_0)_k + V_{ij}(m_0). \tag{3.21}$$

In order to show that F is simple quadratic, we have to show that $\forall i, j$ and $\forall l \in T \setminus \{j\}$, $\beta_{ij}^l = 0$.

We first make $k = j$ in (3.10) and identify the coefficients of m_i^3 to obtain

$$\beta_{ij}^i\big(\beta_{ij}^i - \beta_{jj}^i\big) = 0 \quad \text{for all } i \neq j. \tag{3.22}$$

Identifying in (3.10) the coefficients of $m_l^2 m_i$, we get for all $l \in T \setminus \{j\}$,

$$\beta_{ij}^l\big(\beta_{ij}^l + \beta_{lj}^l - 2\beta_{jj}^l\big) = 0. \tag{3.23}$$

Suppose that $\beta_{ij}^l \neq 0$. Then from (3.23) we get

$$\beta_{ij}^l + \beta_{lj}^l = 2\beta_{jj}^l,$$

and from (3.22) we can deduce that $\beta_{ij}^l = (1 + \epsilon_{lj})\beta_{jj}^l$ with

$$\epsilon_{lj} = \begin{cases} 0 & \text{if } \beta_{lj}^l \neq 0, \\ 1 & \text{if } \beta_{lj}^l = 0. \end{cases}$$

Hence β_{ij}^l does not depend on i, we write it β_{lj}. Now identifying the coefficients of m_i^3 in (3.22) gives $\beta_{ij}^2 = 0$, which is a contradiction.

It remains to prove that if F is such that

$$V_{ij}(m) = \alpha_{ij}(m - m_0)_i(m - m_0)_j + \sum_{k=1}^{d} b_{ij}^k (m - m_0)_k + V_{ij}(m_0),$$

then necessarily F is simple quadratic, that is the α_{ij} are independent of i and j. This has been done in [5]. □

References

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

[2] P. Feinsilver, Some classes of orthogonal polynomials associated with martingales, Proc. Amer. Math. Soc. 98 (1986) 298–302.

[3] A. Hassairi, La classification des familles exponentielles naturelles sur $\mathbb{R}^n$ par l'action du groupe linéaire de $\mathbb{R}^{n+1}$, C. R. Acad. Sci. Paris Sér. I 315 (1992) 207–210.

[4] A. Hassairi, M. Zarai, Characterization of the cubic exponential families by orthogonality of polynomials, Ann. Probab. 32 (2004) 2463–2476.

[5] E. Labeye-Voisin, D. Pommeret, Polynômes orthogonaux associés aux familles exponentielles de $\mathbb{R}^d$, C. R. Acad. Sci. Paris Sér. I 320 (1995) 79–84.

[6] G. Letac, Lectures on Natural Exponential Families and Their Variance Functions, Monografias de Matemática, vol. 50, Instituto de Matemática Pure e Aplicada, Rio de Janeiro, 1992.

[7] G. Letac, Introduction to Morris (1982) Natural Exponential Families with Quadratic Variance Functions, Breakthroughs in Statistics 3 (1997) 369–373.

[8] G. Letac, M. Mora, Natural real exponential families with cubic variance functions, Ann. Statist. 18 (1990) 1–37.

[9] C.N. Morris, Natural exponential families with quadratic variance functions, Ann. Statist. 10 (1982) 65–82.