Bernoulli **17**(1), 2011, 441–455
DOI: 10.3150/10-BEJ277

From Schoenberg to Pick–Nevanlinna: Toward a complete picture of the variogram class

EMILIO PORCU[1] and RENÉ L. SCHILLING[2]

[1]*Department of Mathematics, University Jaume I of Castellón, Campus Riu Sec, E-12071 Castellón, Spain. E-mail: eporcu@uni-goettingen.de*
[2]*Institute of Mathematical Stochastics, Technical University Dresden, D-01062 Dresden, Germany. E-mail: rene.schilling@tu-dresden.de*

We show that a large subclass of variograms is closed under products and that some desirable stability properties, such as the product of special compositions, can be obtained within the proposed setting. We introduce new classes of kernels of Schoenberg–Lévy type and demonstrate some important properties of rotationally invariant variograms.

Keywords: complete Bernstein functions; isotropy; Schoenberg–Lévy kernels; variograms

1. Introduction

Positive and conditionally positive definite functions on groups or semigroups have a long history and appear in many applications in probability theory, operator theory, potential theory, moment problems and various other areas. They constitute an important chapter in all treatments of harmonic analysis and their origins can be traced back to papers by Carathéodory, Herglotz, Bernstein and Matthias (see [3] and references therein), culminating in Bochner's theorem from 1932; see the surveys by Berg [3] and Sasvári [28]. Schoenberg's theorem explains the possibility of constructing rotationally invariant positive definite and (the negatives of) conditionally positive definite functions on Euclidean spaces via completely monotone functions and Bernstein functions. Positive and conditionally positive definite functions are a cornerstone of spatial statistics where they are known, respectively, as *covariances* (or kernels) and *variograms*. The theory of random fields, which began in the 1940s with the early works of Kolmogorov (see [10] and references quoted therein) and was further developed by Gandin [13] and Matheron [24], among others, is based on the specification of these classes. In particular, the kriging predictor, that is to say, the best linear unbiased predictor, depends exclusively on the underlying covariance or variogram and we refer to the tour de force in Stein [33] for a rigorous assessment of this framework.

Let $\{Z(\xi), \xi \in \mathbb{R}^d\}$ be a stationary Gaussian random field. The associated covariance function $C: \mathbb{R}^d \to \mathbb{R}$ is positive definite, that is, for any finite collection of points $\{\xi_i\}_{i=1}^n \in \mathbb{R}^d$, the matrix $(C(\xi_i - \xi_j))_{i,j=1}^n$ is positive definite:

$$\text{for all } a_1, a_2, \ldots, a_n \in \mathbb{C} \qquad \sum_{i,j=1}^n a_i C(\xi_i - \xi_j)\overline{a}_j \geq 0.$$

Thus, a function $C:\mathbb{R}^d \to \mathbb{R}$ is positive definite if and only if there exists a stationary Gaussian random field having $C(\cdot)$ as covariance function. If $C(\cdot)$ is *rotationally invariant*, then the associated Gaussian random field is called *isotropic*.

It is well known that the family of covariance functions is a convex cone which is closed under products, pointwise convergence and scale mixtures; for these basic facts, the reader is referred to standard textbooks on geostatistics such as Chilès and Delfiner [10].

A *variogram* $\gamma:\mathbb{R}^d \to \mathbb{R}$ is the variance of the increments of an intrinsically stationary random field, that is, for any two points $\xi_1, \xi_2 \in \mathbb{R}^d$, $\mathbb{V}\mathrm{ar}(Z(\xi_1) - Z(\xi_2)) := \gamma(\xi_1 - \xi_2)$. Note that $\gamma(0) = 0$, $\gamma(\xi) = \gamma(-\xi)$ and that $-\gamma$ is *conditionally positive definite*, that is, for any finite collection of points $\{\xi_i\}_{i=1}^n \in \mathbb{R}^d$, we have

$$\text{for all } a_1, \dots, a_n \in \mathbb{C} \text{ such that} \qquad \sum_{i=1}^n a_i = 0, \qquad -\sum_{i,j=1}^n a_i \gamma(\xi_i - \xi_j)\overline{a}_j \geq 0. \tag{1}$$

With a slight abuse of notation, we will also use the name *variogram* for a function $\gamma:\mathbb{R}^d \to \mathbb{R}$ with $\gamma(0) \geq 0$ and such that $\gamma(\xi) - \gamma(0)$ is the variance of the increments of an intrinsically stationary random field.

There is a close relationship between variograms γ and stationary covariance functions C. The elementary estimate $|C(\xi)| \leq C(0) =: \mathbb{V}\mathrm{ar}\, Z$ shows that stationary covariance functions are necessarily bounded; in particular, $\gamma(\xi) := C(0) - C(\xi)$ is a variogram. Indeed, variograms may be unbounded, as in the case of fractional Brownian motion. If, however, the variogram is bounded, then it is necessarily of the form $C(0) - C(\xi)$, $\xi \in \mathbb{R}^d$, for some stationary covariance function $C(\cdot)$; see, for instance, [10] or [4], Proposition 7.13, and for a more general result due to Harzallah, see [18].

The terminology concerning positive and conditional positive definiteness is not uniform throughout the literature; it depends very much on the mathematical context or the scientific application. Christakos [11] and many other applied scientists use the notion of *permissibility* for both concepts. We will use both conventions alongside each other whenever no confusion can arise.

In this paper, we are mainly interested in rotationally invariant covariances and variograms. This means that the associated Gaussian random field is weakly or intrinsically stationary and isotropic. Isotropy and stationarity are independent assumptions, but we will assume both to keep things simple. An isotropic covariance function, rescaled by its value at the origin, is the characteristic function of a rotationally symmetric random vector on the sphere of $\mathbb{R}^d$. This class of covariances is well understood and we refer to Gneiting [14,15] and the references therein for an extensive survey of this topic. Much less is known about variograms. For instance, it is common knowledge that the class of variograms is a convex cone which is closed in the weak topology of pointwise convergence, but the product of two variograms is not necessarily a variogram. This is a point that deserves a thorough discussion, in the light of a recent beautiful result in [23], Theorem 3(i), where a simple permissibility condition is given for the product of two exponential variograms composed with a homogeneous function.

We shall give a general answer to this question, as well as a complete characterization of those variograms whose product is again permissible. We shall then focus on other challenging problems related to special compositions of variograms, as well as to quasi-arithmetic compositions of them.

The use of kernels of Schoenberg–Lévy type has been persistently emphasized in both old and recent literature. In this paper, we give new forms of kernels of this type that may be appealing for modeling in spatial statistics.

Another crucial problem faced in this paper regards the potential trade-off between, on the one hand, the computational advantages induced by the use of compactly supported kernels and, on the other hand, the fact that compactly supported kernels can be positive definite only on finite-dimensional spaces, by a striking and beautiful result due to Wendland [35]. We consider this problem from the point of view of variograms; this makes sense since variograms, which are possibly unbounded, represent a larger class than covariance functions.

The paper is organized as follows. Section 2 contains the basic material required for a self-contained exposition and for understanding the technical proofs of our statements. Section 3 assesses new stability properties of the variogram class, while Section 4 is dedicated to kernels of Schoenberg–Lévy type.

2. Complete Bernstein functions and complete monotonicity

This section is mainly expository and we collect here some basic material needed later. We will frequently use the following characterization of variograms, for which a proof can be found in [4], Proposition 7.5.

Theorem 1. *A function $\gamma : \mathbb{R}^d \to \mathbb{R}$ is a variogram if and only if the following three conditions are satisfied*:

(i) $\gamma(0) \geq 0$;
(ii) $\gamma(\xi) = \gamma(-\xi)$;
(iii) $-\gamma$ *is conditionally positive definite*, *that is*, *equation* (1) *holds for all* $\xi_1, \dots, \xi_n \in \mathbb{R}^d$.

Let us remark that in harmonic analysis, functions satisfying conditions (i)–(iii) of Theorem 1 are often called *negative definite functions*. We will not use this notion in this paper.

Often, Pólya's theorem (see [4], Theorem 5.4) is useful if one wants to construct concrete examples of variograms.

Theorem 2. *A continuous function $\phi : \mathbb{R} \to [0, \infty)$ which is even* (*i.e.*, $\phi(x) = \phi(-x)$), *decreasing and convex on the interval* $(0, \infty)$ *is positive definite*.

Clearly, $\phi(0) - \phi(x)$ is increasing, concave and a variogram; see, for example, [4], Corollary 7.7.

Recall that a function $f : (0, \infty) \to \mathbb{R}$ is called *completely monotone* if it is arbitrarily often differentiable and

$$(-1)^n f^{(n)}(x) \geq 0 \qquad \text{for } x > 0, n = 0, 1, \dots.$$

By Bernstein's theorem, the set $\mathcal{CM}$ of completely monotone functions coincides with the set of Laplace transforms of positive measures μ on $[0, \infty)$, that is,

$$f(x) = \mathcal{L}\mu(x) = \int_{[0,\infty)} \mathrm{e}^{-xt}\, \mathrm{d}\mu(t),$$

where we only require that e^{-xt} is μ-integrable for any $x > 0$. $\mathcal{CM}$ is a convex cone which is closed under multiplication and pointwise convergence.

Definition 3. *A function $f : (0, \infty) \to \mathbb{R}$ is called a* Stieltjes function *if it is of the form*

$$f(x) = a + \int_{[0,\infty)} \frac{\mathrm{d}\mu(t)}{x+t}, \tag{2}$$

where $a \geq 0$ and μ is a positive measure on $[0, \infty)$ such that $\int_{[0,\infty)} (1+t)^{-1}\, \mathrm{d}\mu(t) < \infty$.

The following properties of the family $\mathcal{S}$ of Stieltjes functions can be found in [4], Section 14, and [3]. $\mathcal{S}$ is a convex cone such that $\mathcal{S} \subset \mathcal{CM}$. For every $f \in \mathcal{S}$, the fractional power $f^\alpha \in \mathcal{S} \subset \mathcal{CM}$, $0 < \alpha \leq 1$, is again a Stieltjes function. Thus, for $f \in \mathcal{S}$, we see that f^α is completely monotone for any $\alpha > 0$, so f belongs to the set $\mathcal{L}$ of logarithmically completely monotone functions discussed in, for example, [3], Section 2.6. The formula

$$\frac{1}{x(1+x^2)} = \int_{[0,\infty)} \mathrm{e}^{-xt}(1 - \cos t)\, \mathrm{d}t$$

shows that $x^{-1}(1 + x^2)^{-1}$ is completely monotone; however, it cannot be a Stieltjes function since it has poles at $\pm i$ and (2) indicates that a Stieltjes function has a holomorphic extension to the cut plane $\mathbb{C} \setminus (-\infty, 0]$. From the integral representation of f, it is immediate that this extension satisfies $\operatorname{Im} z \operatorname{Im} f(z) \leq 0$, that is, f maps the upper complex half-plane to the lower and vice versa.

Definition 4. *A function $f : (0, \infty) \to [0, \infty)$ is called a Bernstein function if it is infinitely often differentiable and $f' \in \mathcal{CM}$.*

The set of Bernstein functions is denoted $\mathcal{BF}$; it is a convex cone which is closed under pointwise convergence. Since a Bernstein function is non-negative and increasing, it has a non-negative limit $f(0+)$. Integrating the Bernstein representation of the completely monotone function f' gives the following integral representation of $f \in \mathcal{BF}$:

$$f(x) = \alpha x + \beta + \int_{(0,\infty)} (1 - \mathrm{e}^{-xt})\nu(\mathrm{d}t), \tag{3}$$

where $\alpha, \beta \geq 0$ are constants and ν is the *Lévy measure*, that is, a positive measure on $(0, \infty)$ satisfying

$$\int_{(0,\infty)} \frac{t}{1+t}\nu(\mathrm{d}t) < \infty.$$

The following composition result will be useful throughout the paper; see [3].

Theorem 5. *Let $\mathcal{X}$ be either of the sets $\mathcal{BF}, \mathcal{CM}$. Then*

$$f \in \mathcal{X}, g \in \mathcal{BF} \quad \Longrightarrow \quad f \circ g \in \mathcal{X}.$$

If we assume that the representing measure $\nu(\mathrm{d}t)$ in (3) is of the form $\nu(\mathrm{d}t) = m(t)\,\mathrm{d}t$, where $m(t)$ is completely monotone, then we get the family of *complete Bernstein functions*. We denote the collection of all complete Bernstein functions by $\mathcal{CBF}$. It is not hard to see that $\mathcal{CBF}$ is, like $\mathcal{BF}$, a convex cone which is closed under pointwise limits. Complete Bernstein functions are widely used in various fields and they are closely related to the following concepts: Bondesson T_2-class (see [9] for the original definition and [5] for further information), operator-monotone functions (the classical source is [12]) and Pick functions (which are also known as Nevanlinna functions, i.e., holomorphic functions in the upper half-plane with non-negative imaginary part there). A detailed survey can be found in [29], and short introductions in [3,20,30]. Among the most prominent examples of complete Bernstein functions are

$$x \mapsto \frac{\lambda x}{\lambda + x} \qquad (\lambda > 0), \qquad x \mapsto x^\alpha \qquad (0 < \alpha < 1),$$
$$x \mapsto \log(1+x), \qquad x \mapsto \sqrt{x}\arctan\frac{1}{\sqrt{x}}.$$

Further examples are given below in Table 1. Many Bernstein functions given in closed form are already in $\mathcal{CBF}$. There are not many known examples of functions in $\mathcal{BF} \setminus \mathcal{CBF}$ and they are all finite or infinite sums of the form $\sum_i p_i(1-\mathrm{e}^{-\lambda_i x})$; see [3]. Some interesting examples are given in terms of the q-versions of the digamma function $\psi_q(x)$ and Euler's constant γ_q: the function $x \mapsto \psi_q(x+1) + \gamma_q$ is in $\mathcal{BF} \setminus \mathcal{CBF}$; see [22].[1]

The following statements are taken from [29].

Table 1. Examples of complete Bernstein functions ($\Gamma(a;x) := \int_x^\infty t^{a-1}\mathrm{e}^{-t}\,\mathrm{d}t$ is the incomplete Gamma function)

Function	Parameter restrictions	Function	Parameters restriction
$1 - \frac{1}{(1+x^\alpha)^\beta}$	$0 < \alpha, \beta \le 1$	$\mathrm{e}x - x(1+\frac{1}{x})^x - \frac{x}{x+1}$	
$(\frac{x^\rho}{1+x^\rho})^\gamma$	$0 < \gamma, \rho < 1$	$\frac{1}{a} - \frac{1}{x}\log(1+\frac{x}{a})$	$a > 0$
$\frac{x^\alpha - x(1+x)^{\alpha-1}}{(1+x)^\alpha - x^\alpha}$	$0 < \alpha < 1$	$\sqrt{\frac{x}{2}}\frac{\sinh^2\sqrt{2x}}{\sinh(2\sqrt{2x})}$	
$\sqrt{x}(1-\mathrm{e}^{-2a\sqrt{x}})$	$a > 0$	$x^{1-\nu}\mathrm{e}^{ax}\Gamma(\nu; ax)$	$a > 0, 0 < \nu < 1$
$\frac{x(1-\mathrm{e}^{-2\sqrt{x+a}})}{\sqrt{x+a}}$	$a > 0$	$x^\nu \mathrm{e}^{a/x}\Gamma(\nu; \frac{a}{x})$	$a > 0, 0 < \nu < 1$

[1] We are grateful to a referee supplying this reference.

Theorem 6. *A function $f:(0,\infty)\to[0,\infty)$ such that $f(0+)$ exists is a complete Bernstein function if and only if it has an analytic extension to the cut complex plane $\mathbb{C}\setminus(-\infty,0]$ such that* $\operatorname{Im} z\cdot\operatorname{Im} f(z)\geq 0$, *that is, f preserves upper and lower half-planes. In particular, all complete Bernstein functions are of the form*

$$f(z)=bz+a+\int_{(0,\infty)}\frac{z}{z+t}\sigma(\mathrm{d}t) \tag{4}$$

with $a,b\geq 0$ and a measure σ satisfying $\int_{(0,\infty)}(1+t)^{-1}\,\mathrm{d}t<\infty$.

Proofs of this classic result can also be found in [3,20,30]. Theorem 6 can be used to show that, for any $f\not\equiv 0$,

$$f\in\mathcal{CBF}\quad\iff\quad\left[x\mapsto\frac{f(x)}{x}\right]\in\mathcal{S}\quad\iff\quad\left[x\mapsto\frac{x}{f(x)}\right]\in\mathcal{CBF}\quad\iff\quad\frac{1}{f}\in\mathcal{S}. \tag{5}$$

Let us briefly indicate the argument: if $f\in\mathcal{CBF}$, then we can use (4) and divide by z. Comparing the resulting formula with (2) reveals that $f(z)/z$ is (the extension to $\mathbb{C}\setminus(-\infty,0]$ of) a Stieltjes function. Therefore (see the comment following Definition 3), we know that $f(z)/z$ maps the upper to the lower complex half-plane. Consequently, the inverse $g(z):=z/f(z)$ preserves upper and lower half-planes and is, by Theorem 6, in $\mathcal{CBF}$. Using the integral representation (4) for g and dividing by z, we get that $g(z)/z=1/f(z)$ is (the extension of) a Stieltjes function. As before, we see that $f=1/(1/f)$ preserves upper and lower half-planes and is, therefore, a complete Bernstein function. This proves all equivalences in (5).

Using the fact that (the extensions of) functions in $\mathcal{CBF}$ preserve, and those in $\mathcal{S}$ swap, complex half-planes, we immediately get the following result. If we let $\mathcal{X}$ be either $\mathcal{CBF}$ or $\mathcal{S}$, then

$$f,g\in\mathcal{X}\quad\Longrightarrow\quad f\circ g\in\mathcal{CBF}.$$

The following stability properties are less obvious.

Theorem 7. *Let $f,g,h\in\mathcal{CBF}$ and $f\not\equiv 0$. Then:*

(i) $(f^\alpha(x)+g^\alpha(x))^{1/\alpha}\in\mathcal{CBF}$ *for all* $\alpha\in[-1,1]\setminus\{0\}$;
(ii) $(f(x^\alpha)+g(x^\alpha))^{1/\alpha}\in\mathcal{CBF}$ *for all* $\alpha\in[-1,1]\setminus\{0\}$;
(iii) $f(x^\alpha)\cdot g(x^{1-\alpha})\in\mathcal{CBF}$ *for all* $\alpha\in[0,1]$;
(iv) $h(f(x))\cdot g(\frac{x}{f(x)})\in\mathcal{CBF}$.

Assertion (iv) was discovered by Uchiyama [34], Lemma 2.1, and since fractional powers $f(x)=x^\alpha$, $0\leq\alpha\leq 1$, are in $\mathcal{CBF}$, (iv) implies (iii). For positive $\alpha>0$, assertions (i), (ii) are in [26] – his proofs are easily adapted to $\alpha<0$ since $f\in\mathcal{CBF}$ if and only if $1/f\in\mathcal{S}$; see (5). A unified approach will appear in [29].

Letting $\alpha\to 0$ in Theorem 7 proves $\lim_{\alpha\downarrow 0}(\frac{1}{2}f^\alpha+\frac{1}{2}g^\alpha)^{1/\alpha}=\sqrt{fg}$ and since pointwise limits of complete Bernstein functions are complete Bernstein, we see that $\sqrt{fg}\in\mathcal{CBF}$ whenever

$f, g \in \mathcal{CBF}$. From this, we can easily deduce a new proof of the so-called log-convexity of the convex cone $\mathcal{CBF}$:

$$f, g \in \mathcal{CBF}, \alpha \in [0,1] \quad \Longrightarrow \quad f^{\alpha} \cdot g^{1-\alpha} \in \mathcal{CBF}. \tag{6}$$

Alternative proofs can be found in [2] and [29].

Indeed, if α is a dyadic number of the form $\alpha = \sum_{i=1}^{n} \alpha_i 2^{-i}$ with $\alpha_i \in \{0,1\}$ and $\alpha_n = 1$, then $\alpha' = 1 - \alpha$ is of the same type with $\alpha'_n = 1$. This is because

$$\alpha' = \sum_{i=1}^{\infty} 2^{-i} - \sum_{i=1}^{n} \alpha_i 2^{-i} = \sum_{i=1}^{n-1} (1-\alpha_i) 2^{-i} + \sum_{i=n+1}^{\infty} 2^{-i} = \sum_{i=1}^{n-1} \alpha'_i 2^{-i} + 2^{-n}$$

with $\alpha'_i = 1 - \alpha_i$, $i = 1, \dots, n-1$. This means that

$$f^{\alpha} g^{1-\alpha} = \prod_{i=1}^{n} \sqrt[2^i]{f^{\alpha_i} g^{\alpha'_i}} = \sqrt{h_1 \sqrt{h_2 \cdots \sqrt{h_{n-2} \sqrt{h_{n-1} \sqrt{f_n g_n}}}}},$$

where h_i stands for either f_i or g_i if $\alpha_i = 1$ or $\alpha_i = 0$, respectively. Thus, repeated applications of (6) with $\alpha = \alpha' = \frac{1}{2}$ lead to (6) for all dyadic $\alpha \in (0,1)$. Since $(0,1) \ni \alpha \mapsto f^{\alpha}$ is continuous, we get (6) for all $\alpha \in (0,1)$.

3. Variograms and their stability properties

As already emphasized in Section 1, the starting point for this work is a result in [23], Theorem 3(i), which is reported below with a short alternative proof.

Theorem 8 ([23]). *Let $\gamma : \mathbb{R}^d \to \mathbb{R}$ be a homogeneous function. Then*

$$\bigl(1 - \mathrm{e}^{-a_1 \gamma(\xi)}\bigr)\bigl(1 - \mathrm{e}^{-a_2 \gamma(\xi)}\bigr), \tag{7}$$

$a_i > 0$, $i = 1, 2$, is a variogram if and only if $\gamma(\xi) = |A\xi|$ for the Euclidean norm $|\cdot|$ and a $d \times d$ matrix A.

It is natural to ask whether Ma's theorem works only for the exponential class of variograms or whether it can be generalized. The subsequent result gives an answer to this problem, supplying a wide class of variograms closed under products.

Here and hereafter, we will use a famous result of Schoenberg and Bochner; see [31] (in the context of covariance functions and complete monotonicity) and [8], page 99 (in the context of variograms and Bernstein functions). We restate Bochner's version in the setting of the current paper. Alternative proofs can be found in the Appendix of Jacob and Schilling [21] and Steerneman and van-Perlo-ten Kleij [32].

Lemma 9. *All variograms γ which are rotationally invariant and permissible in all (or at least infinitely many) dimensions $d = 1, 2, \ldots$ are of the form $\gamma(\xi) = f(|\xi|^2)$ with a Bernstein function $f \in \mathcal{BF}$.*

The next result is not only a generalization of Ma's result, but also the key to a simple proof of Theorem 8.

Theorem 10. *Let g_1, g_2 be Bernstein functions and $0 \le \alpha_1, \alpha_2$ such that $\alpha_1 + \alpha_2 \le 1$. Then $g_1(x^{\alpha_1})g_2(x^{\alpha_2})$ is a Bernstein function.*

Proof. Set $h_{\alpha,\beta}(x) := g_1(x^\alpha) \cdot g_2(x^\beta)$, $x > 0$. It is enough to show that $h'_{\alpha,\beta} \in \mathcal{CM}$. Clearly,

$$h'_{\alpha,\beta}(x) = x^{\alpha+\beta-1}\left(\alpha g_1'(x^\alpha)\frac{g_2(x^\beta)}{x^\beta} + \beta g_2'(x^\beta)\frac{g_1(x^\alpha)}{x^\alpha}\right).$$

Since $g_i \in \mathcal{BF}$, we have that $g_i' \in \mathcal{CM}$ and $x^{-1}g_i(x) \in \mathcal{CM}$. This will also be the case for the compositions $g_1'(x^\alpha)$ and $g_2'(x^\beta)$, $g_1(x^\alpha)/x^\alpha$ and $g_2(x^\beta)/x^\beta$, by a straightforward application of Theorem 5. Moreover, for $\alpha + \beta \le 1$, $x \mapsto x^{\alpha+\beta-1}$ is completely monotone. The proof is completed since completely monotone functions form a convex cone which is closed under products. □

Corollary 11. *Let $\mathbb{R}^d \ni \xi \mapsto \gamma_i(\xi) = g_i(|\xi|^2)$ be rotationally invariant variograms for all $d \in \mathbb{N}$, $i = 1, 2$. Let $\alpha, \beta \in [0, 1]$ be such that $\alpha + \beta \le 1$ and let A be a $d \times d$ matrix. Then*

$$f_{\alpha,\beta}(\xi) := g_1(|A\xi|^{2\alpha})g_2(|A\xi|^{2\beta})$$

is still a variogram on $\mathbb{R}^d$ for all $d \in \mathbb{N}$.

Remark 12. The result of Theorem 10 extends immediately to the product of n Bernstein functions: for $\sum_{i=1}^n \alpha_i \le 1$, $\alpha_i \ge 0$ and $g_i \in \mathcal{BF}$, the function $h(x) := \prod_{i=1}^n g_i(x^{\alpha_i})$ is again in $\mathcal{BF}$. This generalizes the case where $\alpha_i = \frac{1}{n}$, $g_i = g$, $i = 1, \ldots, n$, leading to $h(x) = (g(x^{1/n}))^n$, which is due to [7].

The proof of the result above offers a considerably easier way to show Ma's result.

Proof of Theorem 8. If $\gamma(\xi) = |A\xi|$, Corollary 11 with $g_i(x) = 1 - \exp(-a_i x)$, $i = 1, 2$ and $\alpha = \beta = \frac{1}{2}$ shows that (7) is a variogram.

Now, assume that (7) is a variogram. Then

$$\xi \mapsto \frac{(1 - \mathrm{e}^{-a_1\gamma(\xi)})}{a_1} \cdot \frac{(1 - \mathrm{e}^{-a_2\gamma(\xi)})}{a_2}$$

is a variogram for all $a_1, a_2 > 0$ and so is its pointwise limit $\gamma^2(\xi)$ as $a_1, a_2 \to 0$; thus, $\gamma^2(\xi)$ is a real-valued variogram. As such, it has a Lévy–Khinchine representation

$$\gamma^2(\xi) = Q\xi \cdot \xi + \int_{x\neq 0} \bigl(1 - \cos(x \cdot \xi)\bigr)\nu(\mathrm{d}x),$$

where $Q \in \mathbb{R}^{d\times d}$ is positive semi-definite and ν is a measure with $\int_{x\neq 0} |x|^2 \wedge 1\nu(\mathrm{d}x) < \infty$. Since $\gamma(\xi)$ is homogeneous, we get

$$\gamma^2(\xi) = \frac{\gamma^2(n\xi)}{n^2} \overset{n\to\infty}{\longrightarrow} Q\xi \cdot \xi = \bigl|\sqrt{Q}\xi\bigr|^2$$

for the uniquely determined, positive semidefinite square root $A = \sqrt{Q}$ of Q. □

Several examples of Bernstein functions may be found in [3,4] or in [21]; an extensive list will be included in the monograph [29]. Three celebrated classes of Bernstein functions are well known in the spatial statistics literature:

(1) the *Matérn* class [25] $f_{\alpha,\nu} = 1 - 2^{1-\nu}/\Gamma(\nu)(\alpha\sqrt{x})^\nu K_\nu(\alpha\sqrt{x})$, $x > 0$, for $\alpha, \nu > 0$ and where K_ν is the modified Bessel function of the second kind of order ν;
(2) the *Cauchy* class [16] $f_{\alpha,\beta}(x) := 1 - (1 + x^\alpha)^{-\beta}$, $x > 0$, where $0 < \alpha \leq 1$ and $0 < \beta$;
(3) the *Dagum* class [6] $f_{\rho,\gamma}(x) := (\frac{x^\rho}{1+x^\rho})^\gamma$, $x > 0$, where $\rho, \gamma \in (0, 1)$.

Let us mention a few more stability properties that make some classes of functions appealing for their use in spatial statistics. We again work within the framework of rotationally invariant functions.

Proposition 13. *Let $\gamma : \mathbb{R}^d \to \mathbb{R}$ be rotationally invariant for all dimensions $d = 1, 2, \ldots$ such that $\gamma(\xi) = g(|\xi|^2)$ for some $g \in \mathcal{CBF}$. Then*:

(i) $\mathbb{R}^d \ni \xi \mapsto \frac{|\xi|^2}{g(|\xi|^2)}$ *is a rotationally invariant variogram which is permissible for every* $d \in \mathbb{N}$;
(ii) $\mathbb{R}^d \ni \xi \mapsto \frac{1}{g(1/|\xi|^2)}$ *and* $\xi \mapsto |\xi|^2 g(\frac{1}{|\xi|^2})$ *are rotationally invariant variograms which are permissible for every* $d \in \mathbb{N}$.

Proof. Part (i) is a simple application of the first equivalence in (5) which states that $g \in \mathcal{CBF}$ if and only if $g(x)/x$ is a Stieltjes function.

Part (ii) follows immediately by noting that, for $g \in \mathcal{CBF}$, the function $x \mapsto 1/g(1/x)$ is a composition of the type $\sigma \circ g \circ \sigma(x)$, where σ is the Stieltjes function $x \mapsto \frac{1}{x}$. Since the composition $\sigma \circ g$ is a Stieltjes function and since the composition of two Stieltjes functions is in $\mathcal{CBF}$, we have the first assertion of part (ii). If we apply part (i) to this variogram, the second assertion follows. □

For further (stability) properties of the class $\mathcal{CBF}$, the reader is referred to [29]; some examples of complete Bernstein functions are given below.

Another interesting problem arises when quasi-arithmetic operators, in the sense of Hardy, Littlewood and Pólya [17], are applied to variograms. This means that we seek conditions which preserve the permissibility of the underlying structure. This has been considered in [27] for quasi-arithmetic composition of covariance functions. We believe that the same question in connection with variograms is even more challenging from the mathematical point of view and is equally important as far as statistics are concerned.

Recall that a *power mean* is a mapping of the form $(u,v) \mapsto \psi_\alpha(u,v) := (u^\alpha + v^\alpha)^{1/\alpha}$ for $(u,v) \in \mathbb{R}^2$ and $\alpha \in \mathbb{R} \setminus \{0\}$.

Proposition 14. *Let $\gamma_i : \mathbb{R}^d \to \mathbb{R}$, $i = 1, 2$, be rotationally invariant variograms for all dimensions $d \in \mathbb{N}$. We write g_i for the radial function such that $\gamma_i(\xi) = g_i(|\xi|^2)$:*

(i) *If $g_1, g_2 \in \mathcal{CBF}$, then $\xi \mapsto (\gamma_1^\alpha(\xi) + \gamma_2^\alpha(\xi))^{1/\alpha}$ is a variogram for all $\alpha \in [-1, 1] \setminus \{0\}$.*
(ii) *If $g_1, g_2 \in \mathcal{CBF}$, then $\xi \mapsto (g_1(|\xi|^{2\alpha}) + g_2(|\xi|^{2\alpha}))^{1/\alpha}$ is a variogram for all $\alpha \in [-1, 1] \setminus \{0\}$.*
(iii) *$\xi \mapsto g_1(|\xi|^{2\alpha}) g_2(|\xi|^{2-2\alpha})$ is a variogram for all $0 < \alpha < 1$.*

Proof. Since, by Lemma 9, $g_i \in \mathcal{BF}$, assertion (iii) is a simple consequence of Corollary 11. We should mention at this point that for $g_1, g_2 \in \mathcal{CBF}$, the resulting rotationally invariant variogram would again be of the form $h(|\xi|^2)$ with $h \in \mathcal{CBF}$; see Theorem 7(iv). Both (i) and (ii) follow immediately from 7(i) and (ii), respectively. □

Finally, we combine two aspects treated separately until now. Given two or three isotropic variograms, we seek permissibility conditions for the products of special compositions. The proposition below results from a simple application of Theorem 7(iv) with $h(s) = s$, $f = g_1$, $g = g_2$, respectively, $h = g_3$, $f = g_1$, $g = g_2$.

Proposition 15. *Let $\mathbb{R}^d \ni \xi \mapsto \gamma_i(\xi)$, $i = 1, 2, 3$, be rotationally invariant and isotropic variograms for all $d \in \mathbb{N}$ and assume that $\gamma_i(\xi) = g_i(|\xi|^2)$, where $g_i \in \mathcal{CBF}$. Then*

$$\xi \mapsto \gamma_1(\xi)\gamma_2\left(\frac{\xi}{\sqrt{\gamma_1(\xi)}}\right) \quad \text{and} \quad \gamma_3\big(\sqrt{\gamma_1(\xi)}\big)\gamma_2\left(\frac{\xi}{\sqrt{\gamma_1(\xi)}}\right)$$

are still permissible for all $d \in \mathbb{N}$ and of the form $h(|\xi|^2)$ with some $h \in \mathcal{CBF}$.

We conclude this section by presenting another curious way to construct continuous variograms and, more generally, complex-valued conditionally positive definite functions, with the help of Bernstein functions. The interesting fact in the example below is the product structure, which is quite unusual for conditionally positive definite functions.

Proposition 16. *Let f be a Bernstein function such that the representing measure ν in the Lévy–Khinchine formula (3) has a monotone decreasing density m, that is, $f(x) = \alpha + \beta x + \int_{(0,\infty)} (1 - e^{-xt}) m(t)\,dt$.*

Then $\xi \mapsto \mathrm{i}\xi f(\mathrm{i}\xi)$ is conditionally positive definite and $\xi \mapsto -\mathrm{Re}(\mathrm{i}\xi f(\mathrm{i}\xi))$ is a continuous variogram.

Proof. By the monotonicity of m, we see that $m(t) = \nu[t, \infty)$ for a (Lévy) measure ν, that is, a measure ν on $(0, \infty)$ satisfying $\int_{(0,\infty)} t(1+t)^{-1}\nu(\mathrm{d}t)$. The integration properties of ν become clear from the calculation below since we have only used Fubini's theorem for positive integrands to swap integrals. For $x \geq 0$, we get

$$\begin{aligned} xf(x) &= \alpha x + \beta x^2 + \int_0^\infty x(1 - \mathrm{e}^{-xt}) \int_t^\infty \nu(\mathrm{d}s)\,\mathrm{d}t \\ &= \alpha x + \beta x^2 + \int_0^\infty \int_t^\infty x(1 - \mathrm{e}^{-xt})\nu(\mathrm{d}s)\,\mathrm{d}t \\ &= \alpha x + \beta x^2 + \int_0^\infty \int_0^s x(1 - \mathrm{e}^{-xt})\,\mathrm{d}t\,\nu(\mathrm{d}s) \\ &= \alpha x + \beta x^2 + \int_0^\infty [\mathrm{e}^{-xs} - 1 + sx]\nu(\mathrm{d}s), \end{aligned}$$

which, as by-product, shows that $\int_0^\infty s^2 \wedge s\nu(\mathrm{d}s) < \infty$. We may, therefore, plug in $z = \mathrm{i}\xi$ and get

$$\mathrm{i}\xi f(\mathrm{i}\xi) = -\left(-\mathrm{i}\alpha\xi + \beta\xi^2 + \int_0^\infty [1 - \mathrm{e}^{-\mathrm{i}s\xi} - \mathrm{i}s\xi]\nu(\mathrm{d}s)\right).$$

Thus, $-\gamma(\xi) := \mathrm{i}\xi f(\mathrm{i}\xi)$ is conditionally positive definite and $\operatorname{Re}\gamma(\xi)$ is a variogram. □

4. Kernels and variograms of the Schoenberg–Lévy type

This section explores some results that may be obtained when working with kernels of the Schoenberg–Lévy type. These kernels are extensively used in the literature and we refer to Ma [23] and the references therein. For $\xi_1, \xi_2 \in \mathbb{R}^d$, these are non-stationary covariance functions obtained from a non-negative function $g : [0, \infty) \to [0, \infty)$ such that $g(0) = 0$ through the linear combination

$$g(|\xi_1|) + g(|\xi_2|) - g(|\xi_1 - \xi_2|).$$

A celebrated example is the fractional Brownian sheet [1] with $g(\xi) = |\xi|^\alpha$, $\alpha \in (0, 2]$. Ma [23] shows that for a fixed $\xi_0 \in \mathbb{R}^d$, the function

$$C_{\xi_0}(\xi) := g(|\xi + \xi_0|) + g(|\xi - \xi_0|) - 2g(|\xi|)$$

is a covariance function, provided that $g(|\xi|)$ is a variogram. Indeed, we are going to show that this is a simple consequence of the following, more general, result.

Lemma 17. *Let $\gamma : \mathbb{R}^d \to \mathbb{R}$ be a continuous variogram and let $\xi, \eta \in \mathbb{R}^d$, $d \in \mathbb{N}$. Then*

$$\phi_\eta(\xi) := \gamma(\xi + \eta) + \gamma(\xi - \eta) - 2\gamma(\xi)$$

is a continuous covariance function as a function of ξ. *Moreover, if*

$$\gamma_\eta(\xi) := 2\gamma(\eta) + 2\gamma(\xi) - \gamma(\xi+\eta) - \gamma(\xi-\eta),$$

then $\xi \mapsto \gamma_\eta(\xi)$ *is a continuous variogram.*

Note that in Lemma 17, we have $\gamma_\eta(\xi) = \gamma_\xi(\eta)$, that is, $\eta \mapsto \gamma_\eta(\xi)$ is also a continuous variogram.

Proof of Lemma 17. Recall the following elementary formula for the cosine: $\cos(a+b) + \cos(a-b) = 2\cos a \cos b$. Since $\gamma(\xi)$ has the Lévy–Khinchine representation

$$\gamma(\xi) = Q\xi \cdot \xi + \int_{x \neq 0} (1 - \cos x \cdot \xi)\nu(\mathrm{d}x),$$

we find that

$$\begin{aligned}
\phi_\eta(\xi) &= Q(\xi+\eta)\cdot(\xi+\eta) + Q(\xi-\eta)\cdot(\xi-\eta) - 2Q\xi\cdot\xi \\
&\qquad + \int_{x\neq 0} \big(2\cos x\cdot\xi - \cos x\cdot(\xi+\eta) - \cos x\cdot(\xi-\eta)\big)\nu(\mathrm{d}x) \\
&= 2Q\eta\cdot\eta + \int_{x\neq 0} (2\cos x\cdot\xi - 2\cos x\cdot\xi \cos x\cdot\eta)\nu(\mathrm{d}x) \\
&= 2Q\eta\cdot\eta + 2\int_{x\neq 0} (1-\cos x\cdot\eta)\cos x\cdot\xi\,\nu(\mathrm{d}x).
\end{aligned}$$

This shows that $\xi \mapsto \phi_\eta(\xi)$ is symmetric and positive definite, hence a covariance function. Now, consider

$$\begin{aligned}
\gamma_\eta(\xi) &= 2\gamma(\eta) - \phi_\eta(\xi) \\
&= 2a + 2Q\eta\cdot\eta + 2\int_{x\neq 0}(1-\cos x\cdot\eta)\nu(\mathrm{d}x) \\
&\qquad - 2Q\eta\cdot\eta - 2\int_{x\neq 0}(1-\cos x\cdot\eta)\cos x\cdot\xi\,\nu(\mathrm{d}x) \\
&= 2a + 2\int_{x\neq 0}(1-\cos x\cdot\eta)(1-\cos x\cdot\xi)\nu(\mathrm{d}x).
\end{aligned}$$

Thus, $\gamma_\eta(\xi)$ is a variogram in ξ. The proof is thus complete. □

Lemma 17 has an obvious extension to continuous *complex*-valued functions $\gamma : \mathbb{R}^d \to \mathbb{C}$ satisfying $\gamma(0) \geq 0$, $\gamma(\xi) = \overline{\gamma(-\xi)}$ and the permissibility condition (1) for all $\xi_1, \ldots, \xi_n \in \mathbb{R}^d$. Since such functions also enjoy a (complex) Lévy–Khinchine representation (see [4]), exactly the same argument as in the proof of Lemma 17 shows that for every fixed $\xi_0 \in \mathbb{R}^d$,

$$\gamma_{\xi_0}(\xi) := 2\gamma(\xi) + 2\operatorname{Re}\gamma(\xi_0) - \gamma(\xi - \xi_0) - \gamma(\xi + \xi_0)$$

is permissible and has the Lévy–Khinchine representation

$$\gamma_{\xi_0}(\xi) = 2\int_{y\neq 0}(1 - \mathrm{e}^{\mathrm{i}y\cdot\xi})\big(1 - \cos(y\cdot\xi_0)\big)\nu(\mathrm{d}y),$$

where ν is the Lévy measure of γ. Lemma 17 is a very special case of [4], Proposition 18.2, which goes back to Harzallah [19].

Acknowledgements

We are grateful to three anonymous referees for their valuable comments which helped to improve the presentation of this paper. This work was initiated when the first-named author was visiting the Technical University of Dresden. He is grateful to Professor Zoltán Sasvári for the invitation and the hospitality he was shown. Emilio Porcu acknowledges the DFG-SNF Research Group FOR916, subproject A2.

References

[1] Ayache, A., Leger, S. and Pontier, M. (2002). Drap Brownien fractionnaire. *Potential Anal.* **17** 31–43. MR1906407

[2] Berg, C. (1979). The Stieltjes cone is logarithmically convex. In *Complex Analysis, Joensuu 1978* (I. Laine, O. Lehto and T. Sorvali, eds.). *Lecture Notes in Math.* **747** 46–54. Berlin: Springer. MR0553029

[3] Berg, C. (2008). Stieltjes–Pick–Bernstein–Schoenberg and their connection to complete monotonicity. In *Positive Definite Functions: From Schoenberg to Space–Time Challenges* (E. Porcu and J. Mateu, eds.). Castellón de la Plana. ISBN: 978-84-612-8282-1.

[4] Berg, C. and Forst, G. (1975). *Potential Theory on Locally Compact Abelian Groups*. *Ergeb. Math. Grenzgeb. Bd.* **87**. Berlin: Springer. MR0481057

[5] Berg, C. and Forst, G. (1982). A convolution equation relating the generalized Γ convolutions and the Bondesson class. *Scand. Actuar. J.* 171–175. MR0685883

[6] Berg, C., Mateu, J. and Porcu, E. (2008). The Dagum family of isotropic correlation functions. *Bernoulli* **14** 1134–1149. MR2543589

[7] Berg, C., Boyadzhiev, K. and deLaubenfels, R. (1993). Generation of generators of holomorphic semigroups. *J. Austral. Math. Soc. Ser. A* **55** 246–269. MR1232759

[8] Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. Berkeley, CA: California Univ. Press. MR0072370

[9] Bondesson, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Gebiete* **57** 39–71. MR0623454

[10] Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. Chichester: Wiley. MR1679557

[11] Christakos, G. (1984). On the problem of permissible covariance and variogram models. *Water Resources Res.* **20** 251–265.

[12] Donoghue Jr., W.F. (1974). *Monotone Matrix Functions and Analytic Continuation*. *Grundlehren Math. Wiss.* **207**. Berlin: Springer. MR0486556

[13] Gandin, L.S. (1965). General problem of optimal interpolation and extrapolation of meteorological fields. *Trudy MGO* **168** 75–83.

[14] Gneiting, T. (1999). Isotropic correlation functions on d-dimensional balls. *Adv. in Appl. Probab.* **31** 625–631. MR1742685

[15] Gneiting, T. (2001). Criteria of Pólya type for radial positive definite functions. *Proc. Amer. Math. Soc.* **129** 2309–2318. MR1823914

[16] Gneiting, T. and Schlather, M. (2004). Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Rev.* **46** 269–282. MR2114455

[17] Hardy, G.H., Littlewood, J.E. and Pólya, G. (1952). *Inequalities*, 2nd ed. Cambridge: Cambridge Univ. Press. MR0046395

[18] Harzallah, K. (1967). Fonctions opérant sur les fonctions définies-négatives. *Ann. Inst. Fourier* **17** 443–468. MR0223830

[19] Harzallah, K. (1969). Sur une démonstration de la formule de Lévy–Khintchine. *Ann. Inst. Fourier* **19** 527–532. MR0268598

[20] Jacob, N. (2001). *Pseudo Differential Operators and Markov Processes. Fourier Analysis and Semigroups* **1**. London: Imperial College Press. MR1873235

[21] Jacob, N. and Schilling, R.L. (2005). Function spaces as Dirichlet spaces (about a paper by Maz'ya and Nagel). *Z. Anal. Anwendungen* **24** 3–28. MR2146549

[22] Krattenthaler, C. and Srivastava, H.M. (1996). Summations for basic hypergeometric series involving a q-analogue of the Digamma function. *Comput. Math. Appl.* **32** 73–91. MR1398550

[23] Ma, C. (2007). Why is isotropy so prevalent in spatial statistics? *Trans. Amer. Math. Soc.* **135** 865–871. MR2262884

[24] Matheron, G. (1973). The intrinsic random functions and their applications. *Adv. in Appl. Probab.* **5** 439–468. MR0356209

[25] Matérn, B. (1986). *Spatial Variation*, 2nd ed. *Lecture Notes in Statistics* **36**. Berlin: Springer. MR0867886

[26] Nakamura, Y. (1989). Classes of operator monotone functions and Stieltjes functions. In *The Gohberg Anniversary Collection, Vol. II: Topics in Analysis and Operator Theory* (H. Dym *et al*., eds.). *Operator Theory: Advances and Applications* **41** 395–404. Basel: Birkhäuser. MR1038348

[27] Porcu, E., Mateu, J. and Christakos, G. (2007). Quasi-arithmetic means of covariance functions with potential applications to space–time data. Submitted.

[28] Sasvári, Z. (2008). On the extension problem for covariance functions. In *Positive Definite Functions: From Schoenberg to Space–Time Challenges* (E. Porcu and J. Mateu, eds.). Castellón de la Plana. ISBN: 978-84-612-8282-1.

[29] Schilling, R.L., Song, R. and Vondraček, Z. (2010). *Bernstein Functions: Theory and Applications*. *de Gruyter, Studies in Mathematics* **37**. Berlin: Springer. MR2598208

[30] Schilling, R.L. (1998). Subordination in the sense of Bochner and a related functional calculus. *J. Austral. Math. Soc. Ser. A* **64** 368–396. MR1623282

[31] Schoenberg, I.J. (1938). Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* **44** 522–536. MR1501980

[32] Steerneman, A.G.M. and van Perlo ten Kleij, F. (2005). Spherical distributions: Schoenberg (1938) revisited. *Expo. Math.* **23** 281–287. MR2164779

[33] Stein, M. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer. MR1697409

[34] Uchiyama, M. (2006). A new majorization between functions, polynomials and operator inequalities. *J. Funct. Anal.* **231** 221–244. MR2190170
[35] Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree. *Adv. Comput. Math.* **4** 389–396. MR1366510

Received December 2008 and revised March 2010