CAUCHY-STIELTJES FAMILIES WITH POLYNOMIAL VARIANCE FUNCTIONS AND GENERALIZED ORTHOGONALITY

WŁODZIMIERZ BRYC, RAOUF FAKHFAKH, AND WOJCIECH MŁOTKOWSKI

ABSTRACT. This paper studies variance functions of Cauchy-Stieltjes Kernel (CSK) families generated by compactly supported centered probability measures. We describe several operations that allow to construct additional variance functions from known ones. We construct a class of examples which we believe exhausts all cubic variance functions, and provide examples of polynomial variance functions of arbitrary degree. We also relate CSK families with polynomial variance functions to generalized orthogonality.

Our main results are stated solely in terms of classical probability; some proofs rely on analytic machinery of free probability.

1. INTRODUCTION AND MAIN RESULTS

The Cauchy-Stieltjes Kernel (CSK) families of probability measures were introduced in [11] and extended to non-compact setting in [13]. The constructive approach adopted in these papers is based on an idea of kernel family from an unpublished manuscript [41]. The construction emphasizes analogies with exponential families, using the Cauchy-Stieltjes kernel $1/(1-\theta x)$ instead of the exponential kernel $\exp(\theta x)$, and establishing parametrization by the mean. Kernels of the form $h(x\theta)$, including $1/(1-\theta x)^a$, appear also in [26] and the references cited therein.

After re-parametrization by the mean, CSK families are also a special case $q = 0$ of the q-exponential families from [14]. The non-constructive definition from [14, Section 4] is most convenient for our purposes, as it emphasizes the role of the pseudo-variance function, which appears directly in the definition.

Definition 1.1. The *CSK family* with a *pseudo-variance function* $\mathbb{V}$ generated by a compactly supported non-degenerate probability measure ν is a family of probability measures

$$\{Q_m(dx) := f(x, m)\nu(dx) : \ m \in (m_-, m_+)\},$$

where

$$f(x, m) := \begin{cases} \frac{\mathbb{V}(m)}{\mathbb{V}(m) + m(m-x)} & m \neq 0; \\ 1 & m = 0, \mathbb{V}(0) \neq 0; \\ \frac{\mathbb{V}'(0)}{\mathbb{V}'(0) - x} & m = 0, \mathbb{V}(0) = 0. \end{cases} \tag{1.1}$$

The interval (m_-, m_+) is sometimes called the domain of means, but it will not play a major role here. We will only assume that $0 \in (m_-, m_+)$ and $\mathbb{V}(0) \neq 0$. Then (1.1) is the

---

solution of the difference equation

$$\frac{f(x,m) - f(x,0)}{m} = \frac{x-m}{\mathbb{V}(m)} f(x,m), \quad f(x,0) = 1, \tag{1.2}$$

which is a discrete analog of the differential equation for exponential families noted in [40, Theorem 2] (see also [16, Section 5] and [14]).

It is known that measure ν, if it exists, is uniquely determined (up to the mean) by $\mathbb{V}$, see [11]. It is also known that any non-degenerate compactly supported probability measure ν gives rise to a unique (real analytic) function $\mathbb{V}$, which we will sometimes denote by $\mathbb{V}_\nu$. On the other hand, not every function $\mathbb{V}$ can appear as a pseudo-variance function. The question of determining whether a given class of functions $\mathbb{V}$ corresponds to some measures ν generated a sizeable literature both for the exponential and more recently for the CSK families. In the theory of exponential families, all quadratic variance functions were determined in [23] and in [33]. All cubic variance functions up to affine transformations are described in [28]. Ref. [19] characterizes cubic variance functions by generalized orthogonality. Numerous non-polynomial variance functions have also been studied, see [27]; see also [14, Section 2].

The literature about the variance functions of the CSK families is less comprehensive. CSK families with quadratic variance functions were determined in [11, 14], see also [17]. Cubic (pseudo) variance functions with $\mathbb{V}(0) = 0$ have been studied in [13] and they correspond to measures without first moment. In contrast to exponential families, CSK families are not invariant under translation, so cubic variance functions with $\mathbb{V}(0) \neq 0$ cannot be reduced to the case studied in [13] and require separate investigation. This paper is devoted solely to the case $\mathbb{V}(0) \neq 0$.

We now recall some formulas and assumptions that we will rely upon. It is known (see [13, Proposition 3.1] or [14, (3.4)]) that for $m \neq 0$

$$\int x Q_m(dx) = m, \tag{1.3}$$

so family $\{Q_m : m \in (m_-, m_+)\}$ is indeed parameterized by the mean. One can show that if ν has all moments, $0 \in (m_-, m_+)$, and $\mathbb{V}(0) \neq 0$, then (1.3) extends by continuity to

$$\int x \nu(dx) = 0. \tag{1.4}$$

We will simply assume (1.4). It is then known, and easy to check, that the pseudo-variance function that appears in (1.1) is indeed the variance function,

$$\mathbb{V}(m) = \int (x-m)^2 Q_m(dx), \tag{1.5}$$

see [13, Proposition 3.2], and [14, (3.4)] who consider a more general case.

Denote by $\mathcal{V}$ the class of variance functions corresponding to probability measures ν such that ν is compactly supported, centered: $\int x \nu(dx) = 0$, with variance $\int x^2 \nu(dx) = 1$, so that $\mathbb{V}_\nu(0) = 1$. Denote by $\mathcal{V}_\infty$ the class of those $\mathbb{V} \in \mathcal{V}$ that the function $m \mapsto \mathbb{V}(cm)$ is in $\mathcal{V}$ for every real c.

We begin with some algebraic operations that allow to build new variance functions from known ones. (Here we write $\mathbb{V}(m)$ for a function, not its value.)

Theorem 1.2. *Assume that $\mathbb{V}(m) \in \mathcal{V}$, $\mathbb{V}_1(m), \mathbb{V}_2(m) \in \mathcal{V}_\infty$ and $c \geq 1$. Then*

    (i) $\mathbb{V}(m/c) \in \mathcal{V}$;

(ii) $\mathbb{V}(m) + am \in \mathcal{V}$ and $\mathbb{V}_1(m) + am \in \mathcal{V}_\infty$ for any $a \in \mathbb{R}$;

(iii) $\mathbb{V}_1(m) + \mathbb{V}_2(m) - 1 \in \mathcal{V}_\infty$ and $c\mathbb{V}_1(m) - c + 1 \in \mathcal{V}_\infty$;

(iv) $\mathbb{V}_1(m) - m^2 \in \mathcal{V}$.

The proof of this theorem appears in Section 2.3.

Next, we describe of a class of cubic variance functions.

Theorem 1.3. *Fix $a, b, c \in \mathbb{R}$ and consider cubic function $\mathbb{V}(m) = 1 + am + bm^2 + cm^3$. If $(b+1)^3 \geq 27c^2$ then $\mathbb{V}$ is in $\mathcal{V}$. Furthermore, $\mathbb{V}$ is in $\mathcal{V}_\infty$ if and only if $b^3 \geq 27c^2$.*

The proof of this theorem appears in Section 2.4. We conjecture that the condition $(b+1)^3 \geq 27c^2$ is also necessary for $\mathbb{V}$ to belong to $\mathcal{V}$.

Our final result relates polynomial variance functions for a CSK family to generalized orthogonality. Suppose $\{P_n(x) : n = 0, 1, 2 \ldots\}$ is a family of real polynomials, indexed by their degree n with $P_0(x) = 1$; it is sometimes convenient to set $P_k(x) = 0$ for $k < 0$.

We use the following generalized orthogonality condition.

Definition 1.4. *Fix $d \in \mathbb{N}$. We say that polynomials $\{P_n\}$ are d-orthogonal with respect to a probability measure ν if $\int P_n(x)\nu(dx) = 0$ for all $n \geq 1$, and*

$$\int P_n(x)P_k(x)\nu(dx) = 0 \text{ for all } n \geq 2 + (k-1)d, \quad k = 1, 2 \ldots.$$

It is clear that for measures with infinite support, d-orthogonality with $d = 1$ is just the standard orthogonality. For $d = 2$, we recover the 2-orthogonality introduced in Ref. [19]. The concept of d-orthogonality introduced in [38] is different. Condition of pseudo-orthogonality in [24] is also different.

The following result is a generalization of [17, Theorem 3.2] to $d > 1$, and a CSK -version of [19, Theorem 3.1] when $d = 2$.

Theorem 1.5. *Suppose that $\mathbb{V}$ is a variance function of a CSK family generated by a non-degenerate compactly supported probability measure ν with mean 0 and variance 1. Consider the family of polynomials $\{P_n(x)\}$ with generating function*

$$f(x, m) = \sum_{n=0}^{\infty} P_n(x)m^n, \tag{1.6}$$

where $f(x, m)$ is given by (1.1). Then the following statements are equivalent:

(i) *$\mathbb{V}(m) = 1 + \sum_{k=1}^{d+1} a_k m^k$ is a polynomial of degree at most $d + 1$;*

(ii) *There exist constants $\{b_k : k = 1, \ldots, d + 1\}$ such that polynomials $\{P_n\}$ satisfy recursion*

$$xP_n(x) = P_{n+1}(x) + \sum_{k=1}^{(d+1)\wedge n} b_k P_{n+1-k}(x), \quad n \geq 1 \tag{1.7}$$

with initial conditions $P_0(x) = 1$, $P_1(x) = x$.

(iii) *Polynomials $\{P_n(x)\}$ are d-orthogonal with respect to ν.*

Note that the upper limit of the sum on the right hand side of (1.7) is $d + 1$ under the convention that $P_k(x) = 0$ for $k < 0$, and that Proposition 2.7 below provides examples of polynomial variance functions of arbitrarily high degree. The proof of Theorem 1.5 appears in Section 3.1.

The paper is organized as follows. In Section 2 we introduce free probability notation and use it to prove the first two theorems. We also include some additional examples of variance functions. Section 3 is independent of Section 2 and discusses results on polynomials that imply Theorem 1.5. In Section 4 we provide a combinatorial example involving sequences A001764, A098746 and A106228 from OEIS [37]. We also discuss generating functions and sharpness of some results.

2. Variance functions and free probability

Recall that a *dilation* $D_t(\nu)$ of a probability measure ν by a non-zero real number t is a measure $\mu(U) = \nu(U/t)$. $D_{-1}(\nu)$ is called the *reflection of* ν. In the language of probability theory, dilation changes the law of random variable X to the law of tX.

2.1. **Notation from free probability.** For a probability measure μ on $\mathbb{R}$ we put:

$$M_\mu(z) := \int \frac{\mu(dx)}{1 - xz}, \quad G_\mu(z) := \int \frac{\mu(dx)}{z - x}, \quad F_\mu(z) := 1/G_\mu(z),$$

M_μ is called the *moment generating function*, $G_\mu(z)$ is the *Cauchy-Stieltjes transform*. The *free R-transform* can be defined by the equation

$$R_\mu(zM_\mu(z)) + 1 = M_\mu(z). \tag{2.1}$$

The coefficients $\kappa_n(\mu)$ in the Taylor expansion $R_\mu(z) = \sum_{n=1}^\infty \kappa_n(\mu)z^n$ are called *free cumulants*. We will also use

$$r_\mu(z) := R_\mu(z)/z. \tag{2.2}$$

Equation (2.1) can also be written as

$$zM_\mu(z)r_\mu(zM_\mu(z)) = M_\mu(z) - 1. \tag{2.3}$$

Note that for the dilated measure we have

$$M_{D_t(\mu)}(z) = M_\mu(tz), \quad R_{D_t(\mu)}(z) = R_\mu(tz), \quad r_{D_t(\mu)}(z) = tr_\mu(tz). \tag{2.4}$$

If $|z| \neq 0$ is small enough then

$$G_\mu\left(r_\mu(z) + \frac{1}{z}\right) = z, \qquad F_\mu\left(r_\mu(z) + \frac{1}{z}\right) = \frac{1}{z}. \tag{2.5}$$

The sum of two R-transforms is an R-transform and defines the *free additive convolution* of measures $\mu \boxplus \nu$ by $R_{\mu \boxplus \nu}(z) = R_\mu(z) + R_\nu(z)$. For any real $t \geq 1$, it is known that $tR_\mu(z)$ is an R-transform and defines *additive free convolution power* $\mu^{\boxplus t}$ (see [34]).

Probability measure μ is called $\boxplus$-*infinitely divisible* if its free convolution power $\mu^{\boxplus t}$ is well defined for all real $t > 0$. If $\lambda \neq 0$ then μ is $\boxplus$-infinitely divisible if and only if $D_\lambda(\mu)$ is $\boxplus$-infinitely divisible.

It is known, see [6, 20] that a compactly supported μ with the first moment $m_0 = \int x\mu(dx)$ is $\boxplus$-infinitely divisible if and only if there exists a compactly supported finite measure ω on $\mathbb{R}$ such that $\omega(\mathbb{R}) = \int (x - m_0)^2\mu(dx)$ and

$$r_\mu(z) = m_0 + z \int \frac{\omega(dx)}{1 - zx}.$$

In particular, if ν is a generating measure of a CSK family, then under our moment assumptions, ν is free-infinitely divisible if an only if there is a compactly supported probability measure ω such that

$$r_\nu(z) = zM_\omega(z). \tag{2.6}$$

For a probability measure $\mu \neq \delta_0$ with support in $[0, \infty)$, the *S-transform* is a composition inverse of $zR_\mu(z)$, i.e.

$$R_\mu(zS_\mu(z)) = z \quad \text{or} \quad M_\mu\left(\frac{z}{1+z}S_\mu(z)\right) = 1 + z, \tag{2.7}$$

see e.g. [21, (5)]. Note that in particular $\int x\mu(dx) = 1/S_\mu(0)$.

The product of S-transforms is an S-transform and defines the *multiplicative free convolution* $\mu_1 \boxtimes \mu_2$ by $S_{\mu_1 \boxtimes \mu_2}(z) = S_{\mu_1}(z)S_{\mu_2}(z)$. *Multiplicative free convolution powers* $\mu^{\boxtimes p}$ are defined at least for all $p \geq 1$ (see [5, Theorem 2.17]) by $S_{\mu^{\boxtimes p}}(z) = S_\mu(z)^p$.

The *Marchenko-Pastur measure* with parameter $\lambda > 0$:

$$\pi_\lambda(dx) = (1 - \lambda)^+\delta_0 + \frac{\sqrt{4\lambda - (x - 1 - \lambda)^2}}{2\pi x}1_{x \in [(1-\sqrt{\lambda})^2,(1+\sqrt{\lambda})^2]}dx.$$

plays in free probability the role of the Poisson distribution, see [35]. Since $S_{\pi_\lambda}(z) = \frac{1}{\lambda+z}$, we have

$$S_{D_{1/b}(\pi_{1/b})}(z) = \frac{1}{1 + bz}.$$

It is known, see [21, Section 2], [3, Theorem 1.2] and [4, 30], that if $p > 0$, $b > 0$ then $\mu = \left(D_{1/b}(\pi_{1/b})\right)^{\boxtimes p}$ exists if and only if $\max\{p, 1/b\} \geq 1$. This measure μ has compact support in $[0, \infty)$ and its S-transform equals:

$$S_\mu(z) = \frac{1}{(1 + bz)^p}. \tag{2.8}$$

For additional details and background on free probability we refer to [35, 39].

2.2. **Formulas for variance functions.** A variance function $\mathbb{V}$ of a CSK family generated by a compactly supported centered probability measure $\nu \neq \delta_0$ is real-analytic at $m = 0$, so it extends to the analytic mapping $z \mapsto \mathbb{V}(z)$ on an open disk near $z = 0$. Our assumptions on the first two moments of ν imply that $r_\nu(z) = z + \kappa_3(\nu)z^2 + \ldots$ is invertible near $z = 0$ and its composition inverse is $z/\mathbb{V}_\nu(z)$ ([11, Theorem 3.3]), so that

$$r_\nu(z) = z\mathbb{V}_\nu(r_\nu(z)). \tag{2.9}$$

Replacing z by $z/\mathbb{V}(z)$ from equation (2.5) we get

$$F_\nu\left(z + \frac{\mathbb{V}_\nu(z)}{z}\right) = \frac{\mathbb{V}_\nu(z)}{z}. \tag{2.10}$$

(This was first noted in [14, (4.4)] and exploited in [11, 13, 12].)

The following result is known but we prove it for completeness.

Lemma 2.1. *If $z \mapsto \mathbb{V}(z)$ is a variance function then so is $z \mapsto \mathbb{V}(-z)$.*

Proof. Put $\nu_- := D_{-1}(\nu)$. Then, by (2.4), $r_{\nu_-}(z) = -r_\nu(-z)$ and from (2.9) we have $\mathbb{V}_{\nu_-}(z) = \mathbb{V}_\nu(-z)$. $\square$

The following relates class $\mathcal{V}_\infty$ of variance functions to free probability.

Proposition 2.2. *If $\mathbb{V} = \mathbb{V}_\nu$ then $\nu^{\boxplus\lambda^2}$ exists if and only if $\mathbb{V}(z/\lambda) \in \mathcal{V}$. In particular, $\mathcal{V}_\infty$ is the class of those $\mathbb{V}_\nu \in \mathcal{V}$ that ν is $\boxplus$-infinitely divisible.*

Proof. Suppose $\nu^{\boxplus\lambda^2}$ exists and denote $\nu_\lambda := D_{1/\lambda}(\nu^{\boxplus\lambda^2})$. Then, by (2.4), we have $r_{\nu_\lambda}(z) = \lambda r_\nu(z/\lambda)$ and

$$\frac{r_{\nu_\lambda}(z)}{\mathbb{V}(r_{\nu_\lambda}(z)/\lambda)} = \frac{\lambda r_\nu(z/\lambda)}{\mathbb{V}(r_\nu(z/\lambda))} = \lambda\frac{z}{\lambda} = z,$$

which proves that $\mathbb{V}_{\nu_\lambda}(z) = \mathbb{V}(z/\lambda)$. Conversely, from the first equality, if $\mathbb{V}(z/\lambda)$ is a variance function of some ν_λ then $r_{\nu_\lambda}(z)/\lambda = r_\nu(z/\lambda)$ so $\nu^{\boxplus\lambda^2} = D_\lambda(\nu_\lambda)$ exists.

In particular, from Lemma 2.1 we see that $\mathbb{V} \in \mathcal{V}_\infty$ if and only if ν is $\boxplus$-infinitely divisible. $\square$

From (2.9), (2.6) and (2.1) we get the following.

Proposition 2.3. *A function $\mathbb{V}(z)$ belongs to $\mathcal{V}_\infty$ if and only if there is a compactly supported probability measure ω on $\mathbb{R}$ such that $\mathbb{V}(z) = 1 + R_\omega(z)$.*

We remark that $\mathbb{V}(z) = 1 + zr_\omega(z)$ which, when combined with a perturbation theorem in [8] generates a large number of implicit examples of variance functions in $\mathcal{V}_\infty$. In particular, for every $d \geq 3$ there is a $\delta > 0$ such that $\mathbb{V}(z) = 1 + z^2 + \sum_{k=3}^d c_k z^k$ is in $\mathcal{V}_\infty$ when $\max_k |c_k| < \delta$. ([15, Corollary 2.5] yields explicit characterization of such variance functions for $d = 4$.)

2.3. **Proof of Theorem 1.2.** We need the following lemma that we shall use with $\alpha = 1$, $\beta = 0$.

Lemma 2.4. *If ω is a probability distribution on $\mathbb{R}$, $\alpha > 0$, $\beta \in \mathbb{R}$ then there exists a probability distribution μ such that*

$$M_\mu(z) = \frac{1}{1 - \beta z - \alpha z^2 M_\omega(z)}. \tag{2.11}$$

Proof. For the F-transform, $F_\mu(z) = 1/G_\mu(z)$, we have $M_\mu(z) = 1/(zF_\mu(1/z))$, so the relation (2.11) becomes

$$F_\mu(z) = z - \beta - \frac{\alpha}{F_\omega(z)}.$$

Now it suffices to apply Proposition 5.2 from [7] (see also [20, Section 3.3]). $\square$

Proof of Theorem 1.2. Statement (i) follows from Proposition 2.2.

(ii) Let $G(z) = 1/zM_\nu(1/z)$ be the Cauchy-Stieltjes transform of ν and $F(z) = 1/G(z)$. The continued fraction expansion for G gives

$$F(z) = z - b_0 - \cfrac{c_0}{z - b_1 - \cfrac{c_1}{z - b_2 - \cfrac{c_2}{\ddots}}}$$

where b_n, c_n are the Jacobi coefficients in the three-step recursion for the monic orthogonal polynomials with respect to measure ν,

$$xp_n(x) = p_{n+1}(x) + b_n p_n(x) + c_{n-1} p_{n-1}(x), \quad n \geq 0.$$

(This can be read out from [22, Section 2.6]. The recursion and the continued fraction terminates at $c_N = 0$ if ν is purely atomic with $N + 1$ atoms.)

Define $F_a(z) = F(z - a) + a$. Then $F_a(z)$ has the same continued fraction expansion with the same coefficients c_n, the same coefficient b_0, and for $k \geq 1$ coefficient b_k is replaced by $b_k + a$. Therefore, by Favard's theorem (the usual version, or a finite version when $c_N = 0$; the latter can be read out from the first page of [22, Section 2.5]) $F_a(z)$ is the inverse of a Cauchy-Stieltjes transform of a probability measure ν_a. The first two moments of ν_a are not affected by the change of $b_1, b_2, \ldots$, so ν_a has mean 0 and variance 1.

Since F is well defined outside of the support of ν, we have $F(x) > 0$ for $x > K$ and $F(x) < 0$ for $x < -K$. So F_a also extends to the real axis far away from 0, and therefore ν_a has compact support. (This fact is sometimes called Krein's theorem [25], see e.g. [15, Theorem 3.9].)

Since F satisfies (2.10), function $\mathbb{V}_a(z) = \mathbb{V}(z) + az$ satisfies the same identify with F_a in place of F, identifying the variance function.

Suppose now that $\mathbb{V} \in \mathcal{V}_\infty$. Then $\mathbb{V}(cz)$ is a variance function for any real c, so by the previous reasoning with a replaced by ac, we see that $\mathbb{V}(cz) + acz = \mathbb{V}_a(cz)$ is in $\mathcal{V}$, i.e., $\mathbb{V}_a \in \mathcal{V}_\infty$.

(iii) We use Proposition 2.3. If $\mathbb{V}_1(z) = \mathbb{V}_{\nu_1}(z) = 1 + R_{\omega_1}(z)$, $\mathbb{V}_2(z) = \mathbb{V}_{\nu_2}(z) = 1 + R_{\omega_2}(z)$ then

$$\mathbb{V}_1(z) + \mathbb{V}_2(z) - 1 = 1 + R_{\omega_1}(z) + R_{\omega_2}(z) = 1 + R_{\omega_1 \boxplus \omega_2}(z)$$

and similarly

$$c\mathbb{V}_1(z) - c + 1 = 1 + cR_{\omega_1}(z) = 1 + R_{\omega_1 \boxplus c}(z).$$

(iv) Let $\mathbb{V}_1 = \mathbb{V}_{\nu_1}$ and denote $r_1 := r_{\nu_1}$. Then $r_1(z) = zM_\omega(z)$ for some probability measure ω. Using Lemma 2.4, let ν be such a probability measure that $M_\nu(z) = 1/(1 - zr_1(z))$. Denote $M_\nu = M$, $r_\nu := r$ and put $\widetilde{z} := zM(z)$. Then, by (2.3),

$$z = \frac{\widetilde{z}}{M(z)} = \frac{\widetilde{z}}{\widetilde{z}r(\widetilde{z}) + 1}$$

and

$$r_1(z) = \frac{M(z) - 1}{zM(z)} = r(\widetilde{z}).$$

Applying these identities to the equality

$$\frac{r_1(z)}{\mathbb{V}_1(r_1(z))} = z$$

yields

$$\frac{r(\widetilde{z})}{\mathbb{V}_1(r(\widetilde{z}))} = \frac{\widetilde{z}}{\widetilde{z}r(\widetilde{z}) + 1}$$

or equivalently

$$\frac{r(\widetilde{z})}{\mathbb{V}_1(r(\widetilde{z})) - r(\widetilde{z})^2} = \widetilde{z},$$

which proves that $\mathbb{V}_\nu(z) = \mathbb{V}_1(z) - z^2$. $\qquad\qquad\square$

2.4. **Proof of Theorem 1.3.** The proof of Proposition 2.7 uses S-transforms from (2.8).

Lemma 2.5. *Suppose that $\mu \neq \delta_0$ is a probability measure with compact support in $[0, \infty)$ and that $S_\mu(0) = 1$. Then*

$$\mathbb{V}(z) = \frac{1+z}{S_\mu(z)} \tag{2.12}$$

is in $\mathcal{V}_\infty$.

Proof. Define $\omega(dx) = x\mu(dx)$, and note that this is a probability measure since $\omega(\mathbb{R}) = \int x\mu(dx) = 1/S_\mu(0) = 1$. Let ν be the $\boxplus$-infinitely divisible probability measure defined by (2.6). Then $M_\mu(z) = 1 + r_\nu(z)$, so (2.7) gives

$$r_\nu\left(\frac{z}{1+z}S_\mu(z)\right) = z.$$

Recalling that composition inverse of $z \mapsto r(z)$ is $z/\mathbb{V}_\nu(z)$, in a neighborhood of $z = 0$ we get (2.12).

$\square$

Lemma 2.5 yields a class of variance functions in $\mathcal{V}_\infty$ of the following form.

Lemma 2.6. *Let $\beta_1, \ldots, \beta_d > 0$, $p_1, \ldots, p_2 > 0$, with $\max\{p_j, 1/\beta_j\} \geq 1$. Then the function*

$$\mathbb{V}(z) = (1+z)\prod_{j=1}^{d}(1+\beta_j z)^{p_j} \tag{2.13}$$

is in $\mathcal{V}_\infty$.

Proof. For $j \geq 1$, choose μ_j with $S_{\mu_j}(z) = (1+\beta_j z)^{-p_j}$. Define $\mu = \mu_1 \boxtimes \mu_2 \boxtimes \cdots \boxtimes \mu_d$ so that

$$S_\mu(z) = \prod_{j=1}^{d}(1+\beta_j z)^{-p_j}.$$

Lemma 2.5 ends the proof.

$\square$

We will deduce sufficiency in Theorem 1.3 from the following general result.

Proposition 2.7. *Assume that $d \in \mathbb{N}$, $a, b \in \mathbb{R}$, $c > 0$, $b_1, \ldots, b_d > 0$, $p_1, \ldots, p_d > 0$ and that $\max\{p_j, c/b_j\} \geq 1$ for $\leq j \leq d$. Put*

$$\mathbb{V}(z) = az + bz^2 + (1+cz)\prod_{j=1}^{d}(1+b_j z)^{p_j}.$$

If $b \geq -1$ then $\mathbb{V} \in \mathcal{V}$. If $b \geq 0$ then $\mathbb{V} \in \mathcal{V}_\infty$.

In the present paper we are mainly interested in polynomial variance functions, however here we would like to emphasize that the exponents p_j do not have to be integers; for example $(1+z)^{\sqrt{2}}$ or $(1+z)^{3/2}(1+2z)^{3/2}$ are variance functions in $\mathcal{V}_\infty$.

Proof. Put $\mathbb{V}_1(z) := 1 + az + bz^2$, $\mathbb{V}_2(z) := (1+cz)\prod_{j=1}^{d}(1+b_j z)^{p_j}$. Then $\mathbb{V}_2 \in \mathcal{V}_\infty$ in view of Lemma 2.6, combined with Theorem 1.2(i). If $b \geq 0$ then $\mathbb{V}_1 \in \mathcal{V}_\infty$ (see [11, Theorem 3.2] and the comments therein) and consequently $\mathbb{V}(z) = \mathbb{V}_1(z) + \mathbb{V}_2(z) - 1 \in \mathcal{V}_\infty$ by Theorem 1.2(iii), which proves the second part. To conclude, we apply Theorem 1.2(iv). $\square$

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. The case $c = 0$ is well understood: $\mathbb{V} \in \mathcal{V}$ if and only if $b + 1 \geq 0$ and $\mathbb{V} \in \mathcal{V}_\infty$ if and only if $b \geq 0$. In view of Lemma 2.1 we can assume that $c > 0$.

Applying Proposition 2.7 with $d = 1$, $b_1 = c$ and $p_1 = 2$ we get that

$$az + bz^2 + (1 + cz)^3 = 1 + (a + 3c)z + (b + 3c^2)z^2 + c^3 z^3$$

is in $\mathcal{V}$ for any $b \geq -1$, and in $\mathcal{V}_\infty$ for any $b \geq 0$, with any real a, c. Replacing $a + 3c, b + 3c^2, c^3$ by a, b, c respectively we get the sufficient conditions for $\mathbb{V} \in \mathcal{V}$ and for $\mathbb{V} \in \mathcal{V}_\infty$ as stated.

It remains to show that if $b^3 < 27c^2$, then $\mathbb{V}(z) = 1 + az + bz^2 + cz^3$ is not in $\mathcal{V}_\infty$. By Theorem 1.2(ii), without loss of generality we may assume $a = 0$.

We proceed by contradiction. Suppose $\mathbb{V} \in \mathcal{V}_\infty$. If $b > 0$ then by scaling we'd get $1 + z^2 + cz^3 \in \mathcal{V}_\infty$ for some (different) $c^2 > 1/27$. By Proposition 2.3 this would mean that there exists a compactly supported probability measure ω with $r_\omega(z) = z + cz^2$, contradicting [15, Corollary 2.5] which says this to be possible if and only if $c^2 \leq 1/27$.

Suppose now that $b \leq 0$. Then by Theorem 1.2(iii) with $\mathbb{V}_2(z) = 1 + |b|z^2$ we'd get $1 + cz^3 \in \mathcal{V}_\infty$. Since $c > 0$, we'd be able to rescale and get, say, $1 + 2z^3 \in \mathcal{V}_\infty$. Using Theorem 1.2(iii) again, we would get $1 + z^2 + 2z^3 \in \mathcal{V}_\infty$, contradicting [15, Corollary 2.5] again. (In fact, as explained in Remark 4.5 below, $1 + 2z^3$ is not in $\mathcal{V}$.) $\qquad\square$

3. VARIANCE FUNCTIONS AND POLYNOMIALS

In general, if $\mathbb{V}$ is (real) analytic at 0 and $\mathbb{V}(0) \neq 0$, it is easy to see that expansion (1.6) holds, and its coefficients are polynomials $\{P_n(x)\}$ which solve the recursion

$$xP_n(x) = P_{n-1}(x) + \sum_{k=0}^{n} \frac{\mathbb{V}^{(k)}(0)}{k!} P_{n+1-k}(x), \; n \geq 0, \tag{3.1}$$

with initial polynomials $P_{-1}(x) = 0$ and $P_0(x) = 1$. (In particular, polynomials $\{P_n\}$ are monic when $\mathbb{V}(0) = 1$.) To derive (3.1), multiply (1.2) by $m\mathbb{V}(m)$, expand $\mathbb{V}$ into the power series at $m = 0$, expand $f(x, m)$ into power series (recall that $\mathbb{V}(0) \neq 0$), and compare the coefficients at the powers of m.

We therefore consider a slightly more general recursions than (1.7). Suppose that polynomials $\{P_n\}$ satisfy the recursion

$$xP_n(x) = P_{n-1}(x) + \sum_{k=0}^{n} a_k P_{n+1-k}(x), \; n \geq 0 \tag{3.2}$$

with $a_0 \neq 0$ and initial polynomials $P_{-1}(x) = 0$ and $P_0(x) = 1$.

Proposition 3.1. *Let K be a compact set and suppose that there are $A, R > 0$ such that $|a_k| \leq AR^k$ for all $k = 0, 1, \ldots$. Define $\mathbb{V}(z) = \sum_{k=0}^{\infty} a_k z^k$ for $|z| < 1/R$.*

(i) *If polynomials $\{P_n\}$ satisfy recursion (3.2) then*

$$\sum_{n=0}^{\infty} P_n(x)z^n = \frac{\mathbb{V}(z)}{\mathbb{V}(z) + z(z - x)}$$

and the series converges uniformly over $x \in K$. That is, there is $r > 0$ that does not depend on $x \in K$ such that the series converges uniformly over $x \in K$ for all $|z| < r$.

(ii) *If polynomials $\{P_n\}$ satisfy recursion (3.2) and there is a non-degenerate centered probability measure ν such that $\int P_n(x)\nu(dx) = 0$ for all $n \geq 1$, then $a_0 = \mathbb{V}(0) > 0$, measure ν is compactly supported, and $\mathbb{V}(\cdot)$ is the variance function of a CSK family generated by ν.*

(iii) *If $\mathbb{V}(\cdot)$ is a variance function of a CSK family generated by a non-degenerate centered compactly supported probability measure ν and $\{P_n\}$ are polynomials from (1.6) then $\int P_n(x)\nu(dx) = 0$ for $n \geq 1$.*

Proof. (i) Since $P_0(x) = 1$, without loss of generality, we may assume that $A = 1$. Let $M = \sup_{x \in K} |x|$. Choose $C > R$ such that

$$\frac{M}{C} + \frac{1}{C^2} + \frac{R}{C-R} \leq |a_0|. \tag{3.3}$$

We now check by induction that with this choice of C we have

$$\sup_{x \in K} |P_n(x)| \leq C^n \text{ for all } n \geq 0. \tag{3.4}$$

Clearly, $|P_0(x)| \leq 1 \leq C^0$ and $\sup_{x \in K} |P_1(x)| = \sup_{x \in K} |x/a_0| \leq M/|a_0| \leq C$. Suppose that $N \geq 1$ is such that (3.4) holds for all P_n with $n \leq N$. From (3.2) we see that

$$|a_0| \sup_{x \in K} |P_{N+1}(x)| \leq \sup_{x \in K} |xP_N(x)| + \sup_{x \in K} |P_{N-1}(x)| + \sum_{k=1}^{N} R^k \sup_{x \in K} |P_{N+1-k}(x)|$$

$$\leq MC^N + C^{N-1} + C^{N+1} \sum_{k=1}^{N} (\tfrac{R}{C})^k \leq C^{N+1} \left(\tfrac{M}{C} + \tfrac{1}{C^2} + \tfrac{R}{C-R} \right) \leq |a_0|C^{N+1}$$

by (3.3). This proves (3.4) by induction.

From (3.4) it is clear that with $r = 1/C$ the series (1.6) converges uniformly over $x \in K$ for all (complex) $|m| < r$.

To identify the limit, denote the sum of the series by $\varphi(x, z)$. Multiplying (3.2) by $z^n \neq 0$ and summing over n, we get

$$x\varphi(x, z) = z\varphi(x, z) + \frac{1}{z} \sum_{n=0}^{\infty} \sum_{k=0}^{n} z^k a_k z^{n+1-k} P_{n+1-k}(x). \tag{3.5}$$

Changing the order of summation,

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} z^k a_k z^{n+1-k} P_{n+1-k}(x) = \sum_{k=0}^{\infty} z^k a_k \sum_{n=k}^{\infty} z^{n+1-k} P_{n+1-k}(x)$$

$$= \sum_{k=0}^{\infty} z^k a_k (\varphi(x, z) - 1) = \mathbb{V}(z)(\varphi(x, z) - 1).$$

Inserting this into (3.5) we see that

$$x\varphi(x, z) = z\varphi(x, z) + \frac{\mathbb{V}(z)}{z}(\varphi(x, z) - 1).$$

The solution of this equation is $\varphi(x, z) = \frac{\mathbb{V}(z)}{\mathbb{V}(z)+z(z-x)}$ as claimed.

(ii) If $\int P_n(x)\nu(dx) = 0$ for $n \geq 1$, then integrating the uniformly convergent series term-by-term we get $\int \varphi(x,m)\nu(dx) = 1$. Since for real $m \neq 0$ we have $\mathbb{V}(m) \geq 0$, see (1.5), and $a_0 \neq 0$ taking the limit as $m \to 0$ we get $\mathbb{V}(0) = a_0 > 0$.

Measure ν has compact support, as $\mathbb{V}$ is analytic, see [11, Theorem 3.3].

(iii) Suppose that $\{P_n\}$ are polynomials from (1.6) and $\nu(dx)$ has compact support. Then (3.1) implies (3.2) with $a_k = \mathbb{V}^{(k)}(0)/k!$. Since $\mathbb{V}(0) \neq 0$ and $\mathbb{V}$ is real analytic, one can find $R > 1$ such that $|\mathbb{V}^{(k)}(0)| \leq k!\mathbb{V}(0)R^k$, so the assumption on the growth of $|a_k|$ is satisfied. By uniform convergence for all small enough m we can integrate series (1.6) term-by-term. We get

$$1 = \int f(x;m)\nu(dx) = 1 + \sum_{n=1}^{\infty} m^n \int P_n(x)\nu(dx).$$

Thus $\int P_n(x)\nu(dx) = 0$ for all $n \geq 1$.

$\square$

Next, we relate polynomial variance functions to d-orthogonality.

Proposition 3.2. *Suppose that $\mathbb{V}$ is a variance function of a CSK family generated by a non-degenerate compactly supported probability measure ν with mean 0 and variance 1. Consider the family of polynomials $\{P_n(x)\}$ with generating function (1.6), where $f(x,m)$ is given by (1.1). Then*

(i) *$\int P_n(x)\nu(dx) = 0$ for $n \geq 1$.*
(ii) *If in addition polynomial $P_2(x)$ is orthogonal in $L_2(\nu)$ to all polynomials $\{P_n(x) : n \geq 2 + d\}$, then the family $\{P_n(x)\}$ is d-orthogonal, satisfies recursion (1.7) and $\mathbb{V}$ is a polynomial of degree at most $d + 1$.*

Conversely, if the variance function $\mathbb{V}$ of a CSK family generated by measure ν is a polynomial of degree at most $d+1$ then the polynomials from expansion (1.6) are d-orthogonal.

Proof. (i) This is included in Proposition 3.1(iii).

(ii) Since $P_2(x) = x^2 - \mathbb{V}'(0)x - 1$ and $\mathbb{V}(m)$ is given by (1.5), we see that

$$\int P_2(x)f(x,m)\nu(dx) = \mathbb{V}(m) + m^2 - \mathbb{V}'(0)m - 1. \qquad (3.6)$$

On the other hand due to uniform convergence (Lemma 3.1), for all small enough m we can integrate series (1.6) term by term. Since by assumption $\int P_2(x)P_k(x)\nu(dx) = 0$ for $k \geq 2 + d$, we get

$$\int P_2(x)f(x,m)\nu(dx) = \int P_2(x)\sum_{k=0}^{\infty} P_k(x)m^k\nu(dx)$$

$$= \sum_{k=0}^{\infty} m^k \int P_2(x)P_k(x)\nu(dx) = \sum_{k=0}^{d+1} m^k \int P_2(x)P_k(x)\nu(dx). \quad (3.7)$$

Thus, comparing the right hand sides of (3.6) and (3.7) we see that

$$\mathbb{V}(m) = 1 - m^2 + \mathbb{V}'(0)m + \sum_{k=0}^{d+1} c_k m^k$$

is a polynomial of degree at most $d + 1$, where $c_k = \int P_2(x)P_k(x)\nu(dx)$.

We now prove the converse claim. Proposition 3.1(ii) gives $\int P_n(x)\nu(dx) = 0$ for $n \geq 1$. Noting that $\{P_j(x) : j \leq k\}$ span the same subspace as monomials, to prove d-orthogonality it remains to verify that

$$\int x^k P_n(x)\nu(dx) = 0 \text{ for all } n \geq 2 + (k-1)d \tag{3.8}$$

for all $k \in \mathbb{N}$.

The proof proceeds by induction on k. Consider first the case $k = 1$. From (1.7) we see that xP_n is a linear combination of $P_{n+1}, P_n, \ldots, P_{n-d}$. Thus, $\int xP_n(x)\nu(dx) = 0$ if $n \geq d+1$. If $n = 2, \ldots d$, then (1.7) shows that xP_n is a linear combination of $P_{n+1}, \ldots, P_1$, thus $\int xP_n(x)\nu(dx) = 0$, too.

Suppose now that (3.8) holds for some $k \geq 1$. Take $n \geq 2 + kd$. Then $n > d+1$ so from (1.7) we see that polynomial $x^{k+1}P_n(x)$ is a linear combination of polynomials $\{x^k P_j(x) : j = n-d, n-d+1, \ldots, n+1\}$. Since $j \geq n-d \geq 2+kd-d = 2+(k-1)d$, each of the polynomials $x^k P_j(x)$ in the linear combination satisfies the inductive assumption, $\int x^k P_j(x)\nu(dx) = 0$. Thus $\int x^{k+1}P_n(d)\nu(dx) = 0$, proving that (3.8) holds for all $k \in \mathbb{N}$. □

3.1. **Proof of Theorem 1.5.** By Proposition 3.1, for a family of monic polynomials $\{P_n(x)\}$, recursion (1.7) holds if and only if the generating function (1.6) is given by (1.1) with

$$\mathbb{V}(m) = 1 + b_1 m + (b_2 - 1)m^2 + \sum_{k=3}^{d+1} b_k m^k. \tag{3.9}$$

Thus statements (i) and (ii) are equivalent.

Proposition 3.2(ii) gives implication (iii)⇒(i), as it says that already a special case of d-orthogonality implies (i); the implication (i)⇒(iii) is the converse part of Proposition 3.2.

4. ADDITIONAL RESULTS AND COMMENTS

4.1. **A combinatorial example.** Consider the probability distribution on $[0, \infty)$, which in [30] was denoted $\mu(3, 1)$. Its moments are $\frac{1}{3n+1}\binom{3n+1}{n}$ (Fuss numbers of order 3, A001764 in OEIS) and the moment generating function, denoted $\mathcal{B}_3(z)$, is

$$\mathcal{B}_3(z) = \frac{3}{3 - 4\sin^2 \alpha} = \frac{2\sin \alpha}{\sqrt{3z}},$$

where $\alpha = \frac{1}{3}\arcsin\sqrt{27z/4}$. The first expression was obtained in [31], the second can be obtained by elementary manipulations. The density function was described in [36, 32]. We are going to study a probability distribution which is a transformation of $\mu(3, 1)$.

Proposition 4.1. *If μ is a probability measure on $[0, \infty)$, with the moment generating function $M_\mu(z)$, then there exists a probability measure μ_1 on $[0, \infty)$ such that $M_{\mu_1}(z) = \frac{1}{1-zM_\mu(z)}$.*

Proof. This is a consequence of Proposition 6.1 in [7] with $\psi(z) = \frac{zM_\mu(z)}{1-zM_\mu(z)}$. Namely, since $M_\mu(z)$ is $\mathbb{C}^+ \to \mathbb{C}^+$, the function

$$\frac{z}{1 - zM_\mu(z)} = \frac{z - |z|^2\overline{M_\mu(z)}}{|1 - zM_\mu(z)|^2}$$

is also $\mathbb{C}^+ \to \mathbb{C}^+$. □

Let μ denote the probability measure which satisfies

$$M_\mu(z) = \frac{1}{1 - z\mathcal{B}_3(z)} = \frac{3}{3 - 2\sqrt{3z}\sin\alpha},$$

$\alpha = \frac{1}{3}\arcsin\sqrt{27z/4}$. This identity implies that moments $s(n)$ of μ satisfy the following recurrence relation: $s(0) = 1$ and for $n \geq 1$

$$s(n) = \sum_{i=0}^{n-1} \frac{1}{3i+1}\binom{3i+1}{i}s(n-1-i).$$

This sequence appears in OEIS as A098746:

$$1, 1, 2, 6, 23, 102, 495, 2549, 13682, 75714, 428882, \ldots$$

and counts permutations which avoid patterns 4231 and 42513, see [1, 29]. For $n \geq 1$ we have also

$$s(n) = \sum_{i=0}^{n} \frac{n-i}{n+2i}\binom{n+2i}{i}.$$

From the equation $\mathcal{B}_3(z) = 1 + z\mathcal{B}_3(z)^3$ (see [18]) we obtain identity

$$zM_\mu(z)^2\big(M_\mu(z) - 1\big) = z^2 M_\mu(z)^3 + \big(M_\mu(z) - 1\big)^3, \tag{4.1}$$

which yields the free S-transform

$$S_\mu(z) = \frac{1 + z + \sqrt{(1+z)(1-3z)}}{2(1+z)}.$$

Substituting $zM_\mu(z) \mapsto z$ in (4.1) and applying (2.1) we get

$$z\big(R_\mu(z) + 1\big)R_\mu(z) = z^2\big(R_\mu(z) + 1\big) + R_\mu(z)^3.$$

Putting $R_\mu(z) = zr_\mu(z)$ yields

$$r_\mu(z) - 1 = zr_\mu(z)\big(1 - r_\mu(z) + r_\mu(z)^2\big). \tag{4.2}$$

This implies that $r_\mu(z)$ is the generating function for the sequence A106228:

$$1, 1, 2, 6, 21, 80, 322, 1347, 5798, 25512, 114236, 518848, \ldots,$$

which counts Motzkin paths of a special kind. These are free cumulants of μ, namely $\kappa_n(\mu) = A106228(n-1)$ for $n \geq 1$. Note that the shifted sequence

$$1, 2, 6, 21, 80, 322, 1347, 5798, 25512, 114236, 518848, \ldots$$

is not positive definite, for example $\det\big(\kappa_{i+j+2}(\nu)\big)_{i,j=0}^5 = -3374$, so μ is not $\boxplus$-infinitely divisible, see [35].

From (4.2) one can read out that centered measure ν with $r_\nu(z) = r_\mu(z) - 1$ has $\mathbb{V}_\nu(z) = 1 + 2z + 2z^2 + z^3$, and the comment above shows that $\mathbb{V}_\nu \notin \mathcal{V}_\infty$.

4.2. **More on generating functions.** Several authors considered families of polynomials $\{T_n\}$ with the generating function of the form

$$\sum_{n=0}^{\infty} T_n(x)z^n = \frac{M(z)}{N(z) - zx}, \tag{4.3}$$

where $z \mapsto M(z)$ and $z \mapsto N(z)$ are analytic functions in the neighborhood of $0 \in \mathbb{C}$ with $M(0) = N(0) \neq 0$. See [2, Lemma 2] with his $u(z) = z/M(z)$ and $f(z) = N(z)/z$ or the generating function in [17, (3.10)]. (See also [9, 26], and the discussion in [10].)

At first sight (4.3) looks more general than (1.6), but in fact the difference is superficial. The following result was suggested by [17, Section 3.2].

Proposition 4.2. *Let ν be a non-degenerate compactly supported probability measure with mean 0. Suppose that the sequence of polynomials $\{T_n\}$ has generating function (4.3), $\int T_n(x)\nu(dx) = 0$ for $n \geq 1$, and $\int T_n(x)T_1(x)\nu(dx) = 0$ for $n \geq 2$. Let $\mathbb{V}$ be the variance function of the CSK family generated by ν.*

Then, with $t = \mathbb{V}(0)/M(0)$ we have $M(z) = \mathbb{V}(tz)/t$ and $N(z) = \mathbb{V}(zt)/t + tz^2$. In particular, $T_n(x) = t^n P_n(x)$ for all $n = 0, 1, 2 \ldots$, where the sequence $\{P_n\}$ is given by expansion (1.6) for the density of the CSK family generated by ν.

(Polynomials $\{P_n\}$ are monic if the variance of ν is 1.)

We remark that if in addition, $\int T_2(x)T_n(x)\nu(dx) = 0$ for $n \geq d + 2$, then by Proposition 3.2 the variance function of the CSK family generated by ν is a polynomial of degree at most $d + 1$. When $d = 1$, this recovers [17, Section 3.2].

In order to be able to integrate the series term by term, we first confirm that the series converges uniformly over x from any compact set. (Compare Proposition 3.1(i).)

Lemma 4.3. *Fix $M > 0$. Then there is $r > 0$ such that the series (4.3) converges for all $|x| < M$ and all $|m| < r$.*

Proof. The x-dependent radius $r(x)$ of convergence of the series is the minimum modulus root of equation $N(z) - zx = 0$. Since $N(0) \neq 0$ it is clear that for every $M > 0$ there is $r > 0$ such that $|N(z)| > |zx|$ for all $|z| < r$ and all $|x| < M$. So there are no roots in the disk $|z| < r$ and the radius of convergence is at least r. $\qquad\square$

Proof of Proposition 4.2. Choose $r > 0$ such that the series (4.3) converges for all x from the support of ν. Integrating term-by-term with respect to ν we get

$$\int \left(\sum_{n \in \mathbb{N}} T_n(x)z^n \right) \nu(dx) = \sum_{n=0}^{\infty} \int T_n(x)z^n \nu(dx) = \int T_0(x)\nu(dx) = \int \frac{M(0)}{N(0)}\nu(dx) = 1.$$

We therefore get

$$\int \frac{M(z)}{N(z) - zx}\nu(dx) = 1 \tag{4.4}$$

for all real z close enough to 0.

Using this and (4.3) we compute $T_1(x) = (x + M'(0) - N'(0))/M(0)$. Since $\int T_1(x)\nu(dx) = 0$, we see that $M'(0) = H'(0)$ and $T_1(x) = \alpha x$ with $\alpha = 1/M(0) \neq 0$.

Since $T_1(x)$ is bounded on the support of ν and the series converges uniformly, integrating term by term we get

$$\int \left(\sum_{n=0}^{\infty} T_n(x) T_1(x) z^n \right) \nu(dx) = \sum_{n=0}^{\infty} \int T_n(x) T_1(x) z^n \nu(dx)$$

$$= z \int T_1^2(x) \nu(dx) \nu(dx) = \alpha^2 \mathbb{V}(0) z, \quad (4.5)$$

where $\mathbb{V}(0) > 0$ is the variance of ν (recall that ν is non-degenerate). On the other hand, using partial fractions we get

$$\int \frac{M(z)}{N(z) - zx} T_1(x) \nu(dx) = \alpha \int M(z) \left(\frac{N(z)}{z(N(z) - xz)} - \frac{1}{z} \right) \nu(dx)$$

$$= \frac{\alpha N(z)}{z} \int \frac{M(z)}{N(z) - xz} \nu(dx) - \frac{\alpha M(z)}{z} \int 1 \nu(dx) = \alpha \frac{N(z) - M(z)}{z}. \quad (4.6)$$

(Here, we used (4.4) and the fact that ν is a probability measure.) Therefore, with $t = \alpha \mathbb{V}(0) = \mathbb{V}(0)/M(0) \neq 0$, since (4.5) and (4.6) are equal we get $N(z) = M(z) + tz^2$, and (4.4) takes the form

$$\int \frac{M(z)}{M(z) + tz^2 - zx} \nu(dx) = 1.$$

Substituting $z = m/t$ and setting $\mathbb{V}(m) = tM(m/t)$ we see that

$$\int \frac{\mathbb{V}(m)}{\mathbb{V}(m) + m(m - x)} \nu(dx) = 1.$$

This shows that $\mathbb{V}(m) = tM(m/t)$ is the variance function of the CSK family generated by ν, and it defines the corresponding polynomials $\{P_n\}$ via (1.6) .

To relate polynomials T_n and polynomials P_n we use the above identities to re-write (4.3) as follows:

$$\sum_{n=0}^{\infty} \frac{T_n(x)}{t^n} m^n = \frac{M(m/t)}{N(m/t) - mx/t} = \frac{\mathbb{V}(m)}{\mathbb{V}(m) + m(m - x)} = \sum_{n=0}^{\infty} P_n(x) m^n.$$

$\square$

4.3. Sharpness of some results.

Remark 4.4. [15, Corollary 2.5] implies sharp results about general quartic polynomials. For example, one can deduce that $1 + az + bz^2 + dz^4 \in \mathcal{V}_\infty$ if and only if $-b^2/12 \leq d \leq b^2/4$.

Remark 4.5. Theorem 1.2(iii) does not extend to $\mathbb{V}_1, \mathbb{V}_2 \in \mathcal{V}$. To see this, consider $\mathbb{V}_1(z) = \mathbb{V}_2(z) = 1 + z^3/27$, which is in $\mathcal{V}$ by Theorem 1.3. Applying the operation 54 times, we'd get $1 + 2z^3 \in \mathcal{V}$. The latter is not possible. Using recursion (1.7) and Proposition 3.2(i), one can compute low order of the measure corresponding to the variance function $1 + cz^3$. The first six moments are $(m_1, \ldots, m_6) = (0, 1, 0, 2, c, 5)$. The 4×4 Hankel determinant of these moments is $1 - c^2$, so $1 + 2z^3$ is not a variance function.

Remark 4.6. Theorem 1.2(iv) does not extend to $\mathbb{V}_1 \in \mathcal{V}$. To see this, consider $\mathbb{V}_1(z) = 1 + 4z^2 + 2z^3$, which is in $\mathcal{V}$ by Theorem 1.3, and apply the operation 4 times to get $1 + 2z^3$ which is not in $\mathcal{V}$, as was already noted in Remark 4.5.

References

[1] ALBERT, M. H., ALDRED, R. E. L., ATKINSON, M. D., VAN DITMARSCH, H. P., HANDLEY, C. C., AND HOLTON, D. A. Restricted permutations and queue jumping. *Discrete Math. 287*, 1-3 (2004), 129–133.

[2] ANSHELEVICH, M. Free martingale polynomials. *Journal of Functional Analysis 201* (2003), 228–261.

[3] ARIZMENDI, O., AND HASEBE, T. Classical scale mixtures of Boolean stable laws. *Transactions of the American Mathematical Society 368*, 7 (2016), 4873–4905.

[4] BANICA, T., BELINSCHI, S. T., CAPITAINE, M., AND COLLINS, B. Free Bessel laws. *Canad. J. Math 63*, 1 (2011), 3–37.

[5] BELINSCHI, S. T. Complex analysis methods in noncommutative probability. *arXiv preprint math/0602343* (2006).

[6] BERCOVICI, H., AND PATA, V. A free analogue of Hincin's characterization of infinite divisibility. *Proceedings of the American Mathematical Society 128*, 4 (2000), 1011–1015.

[7] BERCOVICI, H., AND VOICULESCU, D. Free convolution of measures with unbounded support. *Indiana University Mathematics Journal 42*, 3 (1993), 733–774.

[8] BERCOVICI, H., AND VOICULESCU, D. Superconvergence to the central limit and failure of the Cramér theorem for free random variables. *Probability Theory and Related Fields 103*, 2 (1995), 215–222.

[9] BOŻEJKO, M., AND DEMNI, N. Generating functions of Cauchy-Stieltjes type for orthogonal polynomials. *Infin. Dimens. Anal. Quantum Probab. Relat. Top. 12*, 1 (2009), 91–98.

[10] BOŻEJKO, M., AND DEMNI, N. Topics on Meixner families. In *Noncommutative harmonic analysis with applications to probability II*, vol. 89 of *Banach Center Publ.* Polish Acad. Sci. Inst. Math., Warsaw, 2010, pp. 61–74.

[11] BRYC, W. Free exponential families as kernel families. *Demonstratio Mathematica XLII*, 3 (2009), 657–672.

[12] BRYC, W., FAKHFAKH, R., AND HASSAIRI, A. On Cauchy-Stieltjes kernel families. *Journal of Mutlivariate Analysis 124* (2014), 296–312.

[13] BRYC, W., AND HASSAIRI, A. One-sided Cauchy-Stieltjes kernel families. *Journal of Theoretical Probabiity 24*, 2 (2011), 577–594.

[14] BRYC, W., AND ISMAIL, M. Approximation operators, exponential, and q-exponential families. arxiv.org/abs/math.ST/0512224, 2005.

[15] CHISTYAKOV, G., AND GÖTZE, F. Characterization problems for linear forms with free summands. *arXiv preprint arXiv:1110.1527* (2011).

[16] DI BUCCHIANICO, A., AND LOEB, D. E. Natural exponential families and umbral calculus. In *Mathematical essays in honor of Gian-Carlo Rota (Cambridge, MA, 1996)*, vol. 161 of *Progr. Math.* Birkhäuser Boston, Boston, MA, 1998, pp. 195–211.

[17] FAKHFAKH, R. Characterization of quadratic Cauchy-Stieltjes families by orthogonality of polynomials. arxiv.org/abs/1706.00704, 2017.

[18] GRAHAM, R. L., KNUTH, D. E., AND PATASHNIK, O. *Concrete mathematics*, second ed. Addison-Wesley Publishing Company, Reading, MA, 1994. A foundation for computer science.

[19] HASSAIRI, A., ZARAI, M., ET AL. Characterization of the cubic exponential families by orthogonality of polynomials. *The Annals of Probability 32*, 3B (2004), 2463–2476.

[20] HIAI, F., AND PETZ, D. *The semicircle law, free random variables and entropy*, vol. 77 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2000.

[21] HINZ, M., AND MŁOTKOWSKI, W. Free powers of the free Poisson measure. *Colloq. Math. 123*, 2 (2011), 285–290.

[22] ISMAIL, M. E. H. *Classical and Quantum Orthogonal Polynomials in one Variable.* Cambridge University Press, Cambridge, 2005.

[23] ISMAIL, M. E. H., AND MAY, C. P. On a family of approximation operators. *J. Math. Anal. Appl. 63*, 2 (1978), 446–462.

[24] KOKONENDJI, C. C. Characterizations of some polynomial variance functions by d-pseudo-orthogonality. *Journal of Applied Mathematics and Computing 19*, 1-2 (2005), 427–438.

[25] KREIN, M. G., AND NUDELMAN, A. *The Markov moment problem and extremal problems.* Translations of Mathematical Monographs Vol. 50. 1977.

[26] KUBO, I., KUO, H.-H., AND NAMLI, S. The characterization of a class of probability measures by multiplicative renormalization. *Commun. Stoch. Anal. 1*, 3 (2007), 455–472.

[27] LETAC, G. *Lectures on natural exponential families and their variance functions*, vol. 50 of *Monografías de Matemática [Mathematical Monographs].* Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1992.

[28] LETAC, G., AND MORA, M. Natural real exponential families with cubic variance functions. *Ann. Statist. 18*, 1 (1990), 1–37.

[29] MARTINEZ, M. A., AND SAVAGE, C. D. Patterns in inversion sequences ii: Inversion sequences avoiding triples of relations. *arXiv preprint arXiv:1609.08106* (2016).

[30] MŁOTKOWSKI, W. Fuss-Catalan numbers in noncommutative probability. *Documenta Mathematica 15* (2010), 939–955.

[31] MŁOTKOWSKI, W., AND PENSON, K. A. Probability distributions with binomial moments. *Infin. Dimens. Anal. Quantum Probab. Relat. Top. 17*, 2 (2014), 1450014, 32.

[32] MŁOTKOWSKI, W., PENSON, K. A., AND ŻYCZKOWSKI, K. Densities of the Raney distributions. *Doc. Math. 18* (2013), 1573–1596.

[33] MORRIS, C. N. Natural exponential families with quadratic variance functions. *Ann. Statist. 10*, 1 (1982), 65–80.

[34] NICA, A., AND SPEICHER, R. On the multiplication of free N-tuples of noncommutative random variables. *Amer. J. Math. 118*, 4 (1996), 799–837.

[35] NICA, A., AND SPEICHER, R. *Lectures on the combinatorics of free probability*, vol. 13. Cambridge University Press, 2006.

[36] PENSON, K. A., AND SOLOMON, A. I. Coherent states from combinatorial sequences. In *Quantum theory and symmetries (Kraków, 2001).* World Sci. Publ., River Edge, NJ, 2002, pp. 527–530.

[37] SLOANE, N. J., ET AL. The on-line encyclopedia of integer sequences, 2003.

[38] VAN ISEGHEM, J. *Approximants de Pade vectoriels.* PhD thesis, Université des sciences et techniques de Lille-Flandres-Artois, 1987.

[39] VOICULESCU, D. V., DYKEMA, K. J., AND NICA, A. *Free random variables.* No. 1. American Mathematical Soc., 1992.

[40] WEDDERBURN, R. W. M. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika 61* (1974), 439–447.

[41] WESOŁOWSKI, J. Kernel families. Unpublished manuscript, 1999.

DEPARTMENT OF MATHEMATICAL SCIENCES,, UNIVERSITY OF CINCINNATI,, PO BOX 210025,, CINCINNATI, OH 45221–0025, USA
E-mail address: Wlodzimierz.Bryc@uc.edu

FACULTY OF SCIENCES, SFAX UNIVERSITY, TUNISIA
E-mail address: fakhfakh.raouf@gmail.com

MATHEMATICAL INSTITUTE, UNIVERSITY OF WROCLAW, PL. GRUNWALDZKI 2/4 50-384 WROCLAW, POLAND
E-mail address: mlotkow@math.uni.wroc.pl