Séminaire Lotharingien de Combinatoire **67** (2015), Article B67a

SYMBOLIC CALCULUS IN MATHEMATICAL STATISTICS: A REVIEW

ELVIRA DI NARDO

Abstract. In the last ten years, the employment of symbolic methods has substantially extended both the theory and the applications of statistics and probability. This survey reviews the development of a symbolic technique arising from classical umbral calculus, as introduced by Rota and Taylor in 1994. The usefulness of this symbolic technique is twofold. The first is to show how new algebraic identities drive in discovering insights among topics apparently very far from each other and related to probability and statistics. One of the main tools is a formal generalization of the convolution of identical probability distributions, which allows us to employ compound Poisson random variables in various topics that are only somewhat interrelated. Having got a different and deeper viewpoint, the second goal is to show how to set up algorithmic processes performing efficiently algebraic calculations. In particular, the challenge of finding these symbolic procedures should lead to a new method, and it poses new problems involving both computational and conceptual issues. Evidence of efficiency in applying this symbolic method will be shown within statistical inference, parameter estimation, Lévy processes, and, more generally, problems involving multivariate functions. The symbolic representation of Sheffer polynomial sequences allows us to carry out a unifying theory of classical, Boolean and free cumulants. Recent connections within random matrices have extended the applications of the symbolic method.

Contents

1. Introduction 2
2. The symbolic method of moments 6
   2.1. Moments of sampling distributions 9
   2.2. Sheppard's corrections 14
3. Compound Poisson processes 18
   3.1. Time-space harmonic polynomials 23
4. Cumulants 28
   4.1. Randomized compound Poisson r.v.'s 30
   4.2. Lévy processes 33
5. Sheffer polynomial sequences 35
   5.1. Solving linear recurrences 37
   5.2. Lagrange inversion formula 39

5.3. Generalized Bell polynomials 40
6. Fast symbolic computation 41
6.1. k-statistics and polykays 42
6.2. Computing cumulants through generalized Abel polynomials 46
7. Multivariate Faà di Bruno formula 53
7.1. Multivariate cumulants 57
7.2. Multivariate Lévy processes 58
7.3. Multivariate Hermite polynomials 59
7.4. Multivariate Bernoulli polynomials 61
8. Appendix 1: Tables of computation times 62
9. Appendix 2: Families of time-space harmonic polynomials . 64
10. Acknowledgments 67
References 67

1. Introduction

Most practical problems do not have exact solutions, but numerical computing helps in finding solutions to desired precision. Numerical methods give data that can be analyzed to reveal trends or approximations. In contrast to numerical methods, symbolic methods treat objects that are either formal expressions, or are algebraic in nature. In many cases, numerical methods will not give sufficient information about the nature of the problem; this is where symbolic methods fit in. When used properly, they can give us more insight into the problem we are trying to solve. Another advantage is that, while numerical methods may simply fail to compute correct results, symbolic methods yield closed or explicit formulas. Although the term "symbolic method" is also used in different contexts (see for example [42]), here we refer to a set of manipulation techniques aimed at performing algebraic calculations, preferably through an algorithmic approach, in order to find efficient mechanical processes that may be passed to a computer. This is usually called "symbolic computation". Despite the fact that its development took place later than that of numerical or graphical algorithms, symbolic computation has had a comparable impact, especially in theoretical and applied statistical research. Topics range from asymptotic expansions to Edgeworth series, from likelihood functions to saddlepoint approximations.

Since symbolic computation is exact, its accuracy is not a fundamental issue, as it is for numerical methods. Replacing pencil and paper by a computer is not the only aim of symbolic methods. Computational cost is a question naturally arising when problems of complex nature are addressed. The advantage of symbolic methods is that, by approaching a problem from a different mathematical point of view, efficiency is sometimes a by-product.

The symbolic method we are going to review arises from the so-called "classical umbral calculus", a creation of Gian-Carlo Rota, who spent most of his scientific life looking for an elegant and as general as possible formalization of this calculus.

It is certainly due to Rota's efforts that the heuristic method, invented and largely employed by Rev. John Blissard between 1861 and 1868 (see references in [33]), now has a solid foundation and widespread application within the mathematical community. Despite the impressive amount of scientific literature following the introduction of this calculus, Rota, in 1994, had the courage to turn his umbral calculus upside down, outlining what he considered to be a new and correct syntax for this matter. The classical umbral calculus we use consists essentially in a moment calculus, since its basic device is to represent a unital sequence of numbers by symbols α, called umbra, i.e., to associate the sequence $1, a_1, a_2, \ldots$ to the sequence $1, \alpha, \alpha^2, \ldots$ of powers of α via an operator $\mathbb{E}$ that looks like the expectation of random variables (r.v.'s) [35, 75]. The i-th element of the sequence is then called the *i-th moment* of α. This symbolic approach shares with free probability [58, 90] the use of moments as a tool to characterize r.v.'s. An analog of freeness[1] is defined in a manner resembling independence of tensors. In contrast to free probability, where different linear operators may be applied to the same non-commutative r.v., in the symbolic method of moments just one operator is employed, but the same sequence of moments may correspond to more than one symbol. For a pair of commutative r.v.'s, freeness is equivalent to at least one of them having vanishing variance. So freeness is a pure non-commutative device, and the symbolic approach we are going to review can be considered as the natural analog of a non-commutative field.

"As sometimes happens in the practice of the mathematical investigation, the subject we deal with here does not develop the original idea from which our research started in the spring of 1997, but this paper is closely related to it. In that period, Gian-Carlo Rota was visiting professor at the University of Basilicata and, during a conversation just before he left, he shared with us his keen interest in a research project: to develop a combinatorial random variable theory. The delicate question arising on the underlying foundational side and the short time available led us to continue the discussion via email, weaving it with different activities over a period of several months. The following year, Gian-Carlo Rota held his last course in Cortona; we seized the opportunity to spend some time with him. We resumed the thread of our conversations and presented him with the doubts that gradually took hold of us. As usually, his contribution disclosed new horizons that

[1] In free probability, free r.v.'s correspond to classical independent r.v.'s.

have led us to write these pages." This quotation is a part of the introduction of the paper *"Umbral nature of the Poisson random variables"* written by the author together with Domenico Senato and published in a volume dedicated to Gian-Carlo Rota after his death [33]. This paper was the first of a substantial list of articles devoted to the applications of classical umbral calculus in probability and statistics. In these papers, more definitions have been added, and the method has been restyled according to the needs faced when tackling some particular problems. The more these new tools have been developed, the closer the method has come to the method of moments employed in the study of random matrices, which the author believes to be a natural milestone of this theory and at the same time a good starting point for further developments and applications. This is why the version here reviewed is called *symbolic method of moments* in a talk [18] on new challenges and future developments of mathematical statistics given by the author in Dublin in 2011. Two papers have to be mentioned in support of this personal belief: a paper published in the *Annals of Statistics* devoted to the computation of spectral statistics for random matrices [26] and a second one, published in the *Journal of Multivariate Analysis*, where the symbolic method of moments allows us to compute very general joint moments of non-central Wishart distributions [19]. The author chose to not include these topics in this review, since the overall setting is not yet completely settled, and a serious rethinking of the matter in terms of cumulants[2] instead of moments seems necessary.

The first goal of this paper is to give a general overview of how the method has been developed and where it has been used with success. While sticking to the original idea, the current version of the method offers a much richer set of tools. Therefore the introduction of these tools is done gradually along the paper according to the outline of applications. The second goal is to prove its effectiveness in some applications that over the years have contributed to refining the theory. An example particularly meaningful is the computation of k-statistics (Section 6.1), a challenging problem since their first introduction in 1929 as unbiased estimators of cumulants [41]. Many authors have proposed different techniques aiming at performing this computation in a reasonable time. By using the symbolic method here reviewed, an efficient computation of k-statistics can be performed by means of a suitable generalization of randomized compound Poisson r.v.'s.

The paper is organized as follows. Notation is introduced in Section 2 using the semantics of r.v.'s. In this way, the intention is to permit the reader with no prior knowledge of symbolic calculations to become comfortable with them. Section 2.1 shows how to perform these computations in dealing with moments of sampling distributions,

[2]For the definition of cumulants of a number sequence, see Section 4 and in particular Equation (4.1).

reducing the underlying combinatorics of symmetric functions to a few relations covering a great variety of calculations. Section 2.2 is devoted to a special statistical device, called Sheppard's correction, useful in correcting estimations of moments when sample data are grouped into classes. The resulting closed-form formula gives an efficient algorithm to perform computations. In the same section, its multivariate version is sketched, showing what we believe is one of the most interesting feature of this symbolic method: its multivariate generalization. More details are given in Section 7, where a symbolic version of the multivariate Faà di Bruno formula is also introduced. The umbral calculus is particularly suited not only in dealing with number sequences, but also with polynomial sequences. Time-space harmonic polynomials are an example in which the symbolic method allows us to simplify the proofs of many properties, as well as their computation. Introduced in Section 3, these polynomials generate stochastic processes that result in a special family of martingales employed in mathematical finance. Time-space harmonic polynomials rely on the symbolic representation of Lévy processes, as given in Section 4.2, obtained by using a generalization of compound Poisson processes, as given in Section 4.1. A Lévy process is better described by its cumulant sequences than by its moments, as happens for many other r.v.'s, the Gaussian or the Poisson r.v.'s. Section 4 is devoted to introducing umbrae whose moments are cumulants of a given sequence of numbers. Besides classical cumulants, quite recently [90] new families of cumulants have been introduced, the Boolean and the free cumulants by using the algebra of multiplicative functions on suitable lattices. This link is exploited in Section 6.2, where the computational efficiency of the symbolic method of moments is highlighted. In particular, by means of the symbolic representation of Abel polynomials, all families of cumulants can be parametrized by the same formula. This parametrization allows us to write an algorithm giving moments in terms of classical, Boolean and free cumulants, and vice-versa. Other topics on Sheffer sequences are treated in Section 5, including Lagrange inversion and Riordan arrays, and solutions of difference equations. Section 6.1 shows how to compute unbiased estimators of cumulants by using exponential polynomials, which are special Sheffer sequences. The resulting algorithm is faster than others given in the literature, and it has been generalized to the multivariate case. In the last section, the multivariate umbral calculus is introduced, and various applications are given involving special families of multivariate polynomials.

This text is the written version of lectures given at the 67-th Seminaire Lotharingien de Combinatoire in September 2011; joint session with XVII Incontro Italiano di Combinatoria Algebrica. Due to the length and to avoid making dull reading, the formalism is kept to the minimum, and references are given step by step for those wishing to go into details and to understand the technical proofs of formal matters.

2. The symbolic method of moments

The method we are going to introduce in this section consists of shifting powers a^i to indexes a_i by using a linear functional. The idea of dealing with subscripts as they were powers was extensively used by the mathematical community since the nineteenth century, although with no rigorous foundation. The method, sometimes called *Blissard's symbolic method*, is often attributed to J. Sylvester. He was the first who introduced the term *umbral* from latin to denote the shadowy techniques used to prove certain polynomial equations. Roman and Rota have equipped these "tricks" with the right formalism in 1978, by using the theory of linear operators [71]. Despite the impressive amount of applications following this rigorous foundation (see references in [33]), in 1994 Rota proposed a new version of the method in a paper entitled *"The classical umbral calculus"* [75]. The reason of this rethinking is well summarized in the following quotation [75]: "Although the notation of Hopf algebra satisfied the most ardent advocate of spic-and-span rigor, the translation of *classical* umbral calculus into the newly found rigorous language made the method altogether unwieldy and unmanageable. Not only was the eerie feeling of witchcraft lost in the translation, but, after such a translation, the use of calculus to simplify computation and sharpen our intuition was lost by the wayside." The term *classical* is strictly related to which sequence is chosen as the *unity* of the calculus. Rota has always referred to the sequence $\{1\}$, and the same is done in the version we are going to introduce, that from now on will be called *symbolic method of moments.*

Let $\mathbb{R}$ be the real field[3]. The symbolic method of moments consists of a set $\mathscr{A} = \{\alpha, \gamma, \ldots\}$ of elements called *umbrae*, and a linear functional $\mathbb{E} : \mathbb{R}[\mathscr{A}] \to \mathbb{R}$, called *evaluation*, such that $\mathbb{E}[1] = 1$ and

$$\mathbb{E}[\alpha^i \gamma^j \cdots] = \mathbb{E}[\alpha^i] E[\gamma^j] \cdots \quad \text{(uncorrelation property)} \tag{2.1}$$

for non-negative integers $i, j, \ldots$.

Definition 2.1. [33] The sequence $\{1, a_1, a_2, \ldots\} \in \mathbb{R}$ is said to be *umbrally represented* by an umbra α if $\mathbb{E}[\alpha^i] = a_i$ for all positive integers i.

In the following, in place of $\{1, a_1, a_2, \ldots\}$ we will use the notation $\{a_i\}_{i \geq 0}$ assuming $a_0 = 1$. Special umbrae are given in Table 1.

[3]The umbral calculus given in [75] considers a commutative integral domain whose quotient field is of characteristic zero. For the applications given in the following, the real field $\mathbb{R}$ is sufficient.

[4]The i-th Bell number is the number of partitions of a finite non-empty set with i elements or the i-th coefficient times $i!$ in the Taylor series expansion of the function $\exp(e^t - 1)$.

Table 1. Special umbrae

Umbrae	Moments for positive integers i
Augmentation umbra	$\mathbb{E}[\varepsilon^i] = 0$
Unity umbra	$\mathbb{E}[u^i] = 1$
Singleton umbra	$\mathbb{E}[\chi^i] = \delta_{i,1}$, the Kronecker delta
Bell umbra	$\mathbb{E}[\beta^i] = \mathfrak{B}_i$, the i-th *Bell number*[4]
Bernoulli umbra	$\mathbb{E}[\iota^i] = B_i$, the i-th *Bernoulli number*[5]
Euler umbra	$\mathbb{E}[\xi^i] = \mathfrak{E}_i$, the i-th *Euler number*[6]

Associating a measure with a sequence of numbers is familiar in probability theory, when the i-th term of a sequence can be considered as the i-th moment of a r.v., under suitable hypotheses (the so-called Hamburger moment problem [94]). As Rota underlines in *Problem 1: the algebra of probability* [73], all of probability theory could be done in terms of r.v.'s alone by taking an ordered commutative algebra over the reals, and endowing it with a positive linear functional. Following this analogy, a_i is called the *i-th moment* of the umbra α.

Definition 2.2. A r.v. X is said to be *represented* by an umbra α if its sequence of moments $\{a_i\}_{i\geq 0}$ is umbrally represented by α.

In order to avoid misunderstandings, the expectation of a r.v. X will be denoted by E, and its i-th moment by $E[X^i]$.

Example 2.3. If $P(X = 0) = 1$, the r.v. X is represented by the augmentation umbra ε. If $P(X = 1) = 1$, the r.v. X is represented by the unity umbra u. The Poisson r.v. Po(1) of parameter 1 is represented by the Bell umbra β. More examples of umbrae representing classical r.v.'s can be found in [28].

Not all r.v.'s can be represented by umbrae. For example, the Cauchy r.v. does not admit moments. Moreover, not all umbrae represent r.v.'s. For example, the singleton umbra in Table 1 does not represent a r.v. since its variance is negative $\mathbb{E}[\chi^2] - \mathbb{E}[\chi]^2 = -1$, even though this umbra will play a fundamental role in the whole theory. If $\{a_i\}_{i\geq 0}, \{g_i\}_{i\geq 0}$ and $\{b_i\}_{i\geq 0}$ are sequences umbrally represented by the umbrae α, γ and ζ, respectively, then the sequence $\{h_i\}_{i\geq 0}$ with

$$h_i = \sum_{k_1+k_2+\cdots+k_m=i} \binom{i}{k_1, k_2, \ldots, k_m} a_{k_1}\, g_{k_2} \cdots b_{k_m}$$

[5]Many characterizations of Bernoulli numbers can be found in the literature, and each one may be used to define these numbers. Here we refer to the sequence of numbers satisfying $B_0 = 1$ and $\sum_{k=0}^{i} \binom{i+1}{k} B_k = 0$ for $i = 1, 2, \ldots$.

[6]The Euler numbers are the coefficients of the formal power series $2e^z/[e^{2z} + 1]$.

is umbrally represented by $\underbrace{\alpha + \gamma + \cdots + \zeta}_{m}$, that is,

(2.2)
$$\mathbb{E}[(\underbrace{\alpha + \gamma + \cdots + \zeta}_{m})^i] = \sum_{k_1 + k_2 + \cdots + k_m = i} \binom{i}{k_1, k_2, \ldots, k_m} a_{k_1} \, g_{k_2} \cdots b_{k_m}.$$

The right-hand side of Equation (2.2) represents the i-th moment of a summation of independent r.v.'s.

The second fundamental device of the symbolic method of moments is to represent the same sequence of moments by different umbrae.

Definition 2.4. Two umbrae α and γ are said to be *similar* if and only if $\mathbb{E}[\alpha^i] = \mathbb{E}[\gamma^i]$ for all positive integers i. In symbols $\alpha \equiv \gamma$.

If we replace the set $\{\alpha, \gamma, \ldots, \zeta\}$ by a set of m distinct and similar umbrae $\{\alpha, \alpha', \ldots, \alpha''\}$, Equation (2.2) gives

$$\mathbb{E}[(\underbrace{\alpha + \alpha' + \cdots + \alpha''}_{m})^i] = \sum_{\lambda \vdash i} (m)_{\nu_\lambda} d_\lambda a_\lambda, \tag{2.3}$$

where the summation is over all partitions[7] λ of the integer i, $a_\lambda = a_1^{r_1} \, a_2^{r_2} \cdots$ and

$$d_\lambda = \frac{i!}{(1!)^{r_1} (2!)^{r_2} \cdots r_1! \, r_2! \cdots}$$

is the number of i-set partitions with block sizes given by the parts of λ. The new symbol $m{\boldsymbol .}\alpha$ denotes the summation $\alpha + \alpha' + \cdots + \alpha''$ and is called *dot-product* of m and α. This new symbol is referred to as *auxiliary umbra*, in order to underline that it is not in $\mathscr{A}$. So Equation (2.3) may be rewritten as

$$\mathbb{E}[(m{\boldsymbol .}\alpha)^i] = \sum_{\lambda \vdash i} (m)_{\nu_\lambda} d_\lambda a_\lambda. \tag{2.4}$$

The introduction of new symbols is not only a way of lightening notation, but also to represent sequences of moments obtained by suitable generalizations of the right-hand side of Equation (2.4). This will become clearer as we move along further. According to Definition 2.2, the auxiliary umbra $m{\boldsymbol .}\alpha$ represents a summation of m independent and identically distributed (i.i.d.) r.v.'s. Indeed, identically distributed r.v.'s share the same sequence of moments, if they have.

If we denote by $\mathscr{X}$ the alphabet of auxiliary umbrae, the evaluation $\mathbb{E}$ may be defined using the alphabet $\mathscr{A} \cup \mathscr{X}$, and so we can deal with auxiliary umbrae as they were elements of the base alphabet [75]. In

[7]Recall that a partition of an integer i is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_t)$, where λ_j are weakly decreasing integers and $\sum_{j=1}^t \lambda_j = i$. The integers λ_j are called *parts* of λ. The *length* of λ is the number of its parts and will be indicated by ν_λ. A different notation is $\lambda = (1^{r_1}, 2^{r_2}, \ldots)$, where r_j is the number of parts of λ equal to j and $r_1 + r_2 + \cdots = \nu_\lambda$. We use the classical notation $\lambda \vdash i$ to denote that "λ is a partition of i".

the following, we still refer to the alphabet $\mathscr{A}$ as the overall alphabet of umbrae, assuming that auxiliary umbrae are included.

Summations of i.i.d.r.v.'s appear in many sample statistics. So the first application we give in statistics concerns sampling distributions.

2.1. **Moments of sampling distributions.** It is recognized that an appropriate choice of language and notation can simplify and clarify many statistical calculations [54]. In managing algebraic expressions such as the variance of sample means, or, more generally, moments of sampling distributions, the main difficulty is the manual computation. Symbolic computation has removed many of such difficulties. The need of efficient algorithms has increased the attention on how to speed up the computational time. Many of these algorithms rely on the algebra of symmetric functions [1]: U-statistics are an example which arise naturally when looking for minimum-variance unbiased estimators. A U-statistic of a random sample has the form

$$U = \frac{1}{(n)_k} \sum \Phi(X_{j_1}, X_{j_2}, \ldots, X_{j_k}), \tag{2.5}$$

where $X_1, X_2, \ldots, X_n$ are n independent r.v.'s, and the sum ranges over the set of all permutations $(j_1, j_2, \ldots, j_k)$ of k integers chosen in $[n] := \{1, 2, \ldots, n\}$. If $X_1, X_2, \ldots, X_n$ are identically distributed r.v.'s with common cumulative distribution function $F(x)$, U is an unbiased estimator of the population parameter

$$\Theta(F) = \int\int \cdots \int \Phi(x_1, x_2, \ldots, x_k)\, \mathrm{d}F(x_1)\, \mathrm{d}F(x_2) \cdots \mathrm{d}F(x_k).$$

In this case, the function Φ may be assumed to be a symmetric function of its arguments. Often, in applications, Φ is a polynomial in the X_j's, so that U-statistics are symmetric polynomials. By virtue of the fundamental theorem of symmetric polynomials, U-statistics can be expressed in terms of elementary symmetric polynomials. The symbolic method proposed here provides a way to find this expression [22].

The starting point is to replace $\mathbb{R}$ by $\mathbb{R}[x_1, x_2, \ldots, x_n]$, where $x_1, x_2, \ldots, x_n$ are indeterminates. This device allows us to deal with multivariable umbral polynomials $p \in \mathbb{R}[x_1, x_2, \ldots, x_n]$. The uncorrelation property (2.1) is then updated to

$$\mathbb{E}[x_i^{k_1} x_j^{k_2} \cdots \alpha^{k_3} \gamma^{k_4} \cdots] = x_i^{k_1} x_j^{k_2} \cdots \mathbb{E}[\alpha^{k_3}]\mathbb{E}[\gamma^{k_4}] \cdots$$

for any set of distinct umbrae in $\mathscr{A}$, for $i, j \in [n]$, with $i \neq j$, and for all positive integers k_1, k_2, k_3, k_4.

For umbral polynomials $p, q \in \mathbb{R}[x_1, x_2, \ldots, x_n][\mathscr{A}]$, we may relax the similarity equivalence and introduce another umbral equivalence:

$$p \simeq q \quad \text{if and only if} \quad \mathbb{E}[p] = \mathbb{E}[q].$$

A polynomial sequence $\{p_i\}_{i\geq 0} \in \mathbb{R}[x_1, x_2, \ldots, x_n]$, with $p_0 = 1$ and degree$(p_i) = i$, is represented by an umbra ψ, called *polynomial umbra*, if $\mathbb{E}[\psi^i] = p_i$ for all positive integers i. The classical bases of the algebra of symmetric polynomials can be represented by polynomial umbrae [21]. For example, the i-th elementary symmetric polynomial $e_i(x_1, x_2, \ldots, x_n)$ satisfies

$$e_i(x_1, x_2, \ldots, x_n) \simeq \frac{(\chi_1 x_1 + \chi_2 x_2 + \cdots + \chi_n x_n)^i}{i!}, \quad \text{for } i \leq n, \tag{2.6}$$

with $\chi_1, \chi_2, \ldots, \chi_n$ uncorrelated singleton umbrae. If the indeterminates $x_1, x_2, \ldots, x_n$ are replaced by n uncorrelated umbrae $\alpha_1, \alpha_2, \ldots, \alpha_n$ similar to an umbra α, then Equivalence (2.6) becomes

$$e_i(\alpha_1, \alpha_2, \ldots, \alpha_n) \simeq \frac{[n_{\bullet}(\chi\alpha)]^i}{i!}, \quad \text{for } i \leq n. \tag{2.7}$$

The umbral polynomial $e_i(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is called the *i-th umbral elementary symmetric polynomial*. The umbrae $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ in (2.7) may also be replaced by some powers such as $\{\alpha_1^j, \alpha_2^j, \ldots, \alpha_n^j\}$. Let us consider the monomial symmetric polynomial

$$\mathfrak{m}_\lambda(x_1, x_2, \ldots, x_n) = \sum x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}$$

with $\lambda \vdash i \leq n$, where the notation is symbolic and must be interpreted in the way that all different images of the monomial $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}$ under permutations of the variables have to be summed. If the indeterminates $x_1, x_2, \ldots, x_n$ are replaced by n uncorrelated umbrae $\alpha_1, \alpha_2, \ldots, \alpha_n$ similar to an umbra α, then

$$\mathfrak{m}_\lambda(\alpha_1, \alpha_2, \ldots, \alpha_n) \simeq \frac{[n_{\bullet}(\chi\alpha)]^{r_1}}{r_1!} \frac{[n_{\bullet}(\chi\alpha^2)]^{r_2}}{r_2!} \cdots, \tag{2.8}$$

which is a product of umbral elementary polynomials in (2.7). In many statistical calculations, the so-called augmented monomial symmetric polynomials [94] are involved:

$$\tilde{\mathfrak{m}}_\lambda(x_1, x_2, \ldots, x_n) = \sum_{j_1 \neq \ldots \neq j_{r_1} \neq j_{r_1+1} \neq \ldots \neq j_{r_1+r_2} \neq \cdots} x_{j_1} \cdots x_{j_{r_1}} \, x_{j_{r_1+1}}^2 \cdots x_{j_{r_1+r_2}}^2 \cdots.$$

These polynomials are obtained from $\mathfrak{m}_\lambda(x_1, x_2, \ldots, x_n)$ as follows:

$$\mathfrak{m}_\lambda(x_1, x_2, \ldots, x_n) = \frac{\tilde{\mathfrak{m}}_\lambda(x_1, x_2, \ldots, x_n)}{r_1! r_2! \cdots}. \tag{2.9}$$

Then, from Equivalence (2.8) and Equation (2.9), we have

$$\tilde{\mathfrak{m}}_\lambda(\alpha_1, \alpha_2, \ldots, \alpha_n) \simeq [n_{\bullet}(\chi\alpha)]^{r_1} [n_{\bullet}(\chi\alpha^2)]^{r_2} \cdots, \tag{2.10}$$

which is again a product of umbral elementary polynomials (2.7).

Example 2.5. If $\lambda = (1^2, 3)$ and $n \geq 5$ then

$$[n\boldsymbol{.}(\chi\alpha)]^2[n\boldsymbol{.}(\chi\alpha^3)] = (\chi_1\alpha_1 + \cdots + \chi_n\alpha_n)^2(\chi_1\alpha_1^3 + \cdots + \chi_n\alpha_n^3)$$

$$= \left(\sum_{1 \leq j_1 \neq j_2 \leq n} \chi_{j_1}\alpha_{j_1}\chi_{j_2}\alpha_{j_2} \right) \left(\sum_{1 \leq j_3 \leq n} \chi_{j_3}\alpha_{j_3}^3 \right) \simeq \sum_{1 \leq j_1 \neq j_2 \neq j_3 \leq n} \alpha_{j_1}\alpha_{j_2}\alpha_{j_3}^3.$$

This last equivalence follows by observing that $\mathbb{E}[\chi_{j_1}\chi_{j_2}\chi_{j_3}]$ vanishes whenever there is at least one pair of equal indices among $\{j_1, j_2, j_3\}$.

More details on symmetric polynomials and polynomial umbrae are given in [21]. Due to Equivalence (2.10), the fundamental expectation result, at the core of unbiased estimation of moments, may be restated as follows.

Theorem 2.6. [21] *If* $\lambda = (1^{r_1}, 2^{r_2}, \ldots) \vdash i \leq n$, *then*

$$\mathbb{E}\left\{[n\boldsymbol{.}(\chi\alpha)]^{r_1}[n\boldsymbol{.}(\chi\alpha^2)]^{r_2}\cdots\right\} = (n)_{\nu_\lambda} a_\lambda.$$

From Equation (2.4) and Theorem 2.6, the next corollary follows.

Corollary 2.7. *If* $i \leq n$, *then*

$$(m\boldsymbol{.}\alpha)^i \simeq \sum_{\lambda \vdash i} \frac{(m)_{\nu_\lambda}}{(n)_{\nu_\lambda}} d_\lambda \, [n\boldsymbol{.}(\chi\alpha)]^{r_1}[n\boldsymbol{.}(\chi\alpha^2)]^{r_2}\cdots. \tag{2.11}$$

From Equivalence (2.11), if $m = n$, then

$$(n\boldsymbol{.}\alpha)^i \simeq \sum_{\lambda \vdash i} d_\lambda [n\boldsymbol{.}(\chi\alpha)]^{r_1}[n\boldsymbol{.}(\chi\alpha^2)]^{r_2}\cdots. \tag{2.12}$$

Theorem 2.6 discloses a more general result: products of moments are moments of products of umbral elementary symmetric polynomials (2.7).

By analogy with Equation (2.5), the umbral symmetric polynomial on the right-hand side of Equivalence (2.11),

$$\frac{1}{(n)_{\nu_\lambda}}[n\boldsymbol{.}(\chi\alpha)]^{r_1}[n\boldsymbol{.}(\chi\alpha^2)]^{r_2}\cdots,$$

is called the *U-statistic* of uncorrelated and similar umbrae $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then Theorem 2.6 says that U-statistics can be expressed as products of umbral elementary symmetric polynomials, and Corollary 2.7 says that any statistic involving a summation of i.i.d.r.v.'s can be expressed as a linear combination of U-statistics.

A more general problem consists in computing population moments of sample statistics. These are symmetric functions of not independent r.v.'s. Within the symbolic method of moments, the dependence between r.v.'s corresponds to working with umbral monomials with non-disjoint supports. Indeed, the support of an umbral polynomial p is the set of all umbrae occurring in p. If μ and ν are umbral monomials with non-disjoint supports, then $\mathbb{E}[\mu^i \, \nu^j] \neq \mathbb{E}[\mu^i]\mathbb{E}[\nu^j]$ for all positive integers i and j. In the following, unless otherwise specified, we work with umbral monomials with non-disjoint supports and give an example of how to use these monomials to perform computations with not independent r.v.'s.

Example 2.8. [22] Let us consider a bivariate random sample of r.v.'s $(X_1, Y_1), \ldots, (X_n, Y_n)$, that is, $(X_1, \ldots, X_n)$ is a random sample with parent distribution X not necessarily independent from the parent distribution Y of the random sample $(Y_1, \ldots, Y_n)$. In order to compute the expectation of

$$\left(\sum_{i=1}^{n} X_i\right)^2 \left(\sum_{i=1}^{n} Y_i\right) \text{ or } \left(\sum_{i \neq j}^{n} X_i^2 X_j\right) \left(\sum_{i=1}^{n} X_i^2 Y_i\right)^2 \tag{2.13}$$

by the symbolic method of moments, the main tool is to represent the sample statistics in (2.13) as a suitable linear combination of umbral monomials μ_1 and μ_2 representing the populations X and Y, respectively. For the former sample statistic (2.13), we have

$$E\left[\left(\sum_{i=1}^{n} X_i\right)^2 \left(\sum_{i=1}^{n} Y_i\right)\right] = \mathbb{E}[(n_{\bullet}\mu_1)^2 \; (n_{\bullet}\mu_2)].$$

For the latter, let us observe that $E[(\sum_{i \neq j}^{n} X_i^2 X_j)] = \mathbb{E}[(n_{\bullet}\chi\mu_1^2)(n_{\bullet}\chi\mu_1)]$ and $E[(\sum_{i=1}^{n} X_i^2 Y_i)] = \mathbb{E}[n_{\bullet}(\chi\mu_1^2 \; \mu_2)]$, so that the overall expectation of the product is given by

$$\mathbb{E}[n_{\bullet}(\chi_1\mu_1^2) \, n_{\bullet}(\chi_1\mu_1) \, n_{\bullet}(\chi_2\mu_1^2\mu_2) \, n_{\bullet}(\chi_3\mu_1^2\mu_2)]. \tag{2.14}$$

Both evaluations can be performed by using a suitable generalization of Theorem 2.6 involving multisets of umbral monomials.

Let us recall that a multiset of umbral monomials $\{\mu_1, \mu_2, \ldots, \mu_k\}$ is

$$M = \{\underbrace{\mu_1, \ldots, \mu_1}_{f(\mu_1)}, \underbrace{\mu_2, \ldots, \mu_2}_{f(\mu_2)}, \ldots, \underbrace{\mu_k, \ldots, \mu_k}_{f(\mu_k)}\}. \tag{2.15}$$

The length of the multiset M in (2.15) is $|M| = f(\mu_1) + f(\mu_2) + \cdots + f(\mu_k)$. The *support* of the multiset M is $\bar{M} = \{\mu_1, \ldots, \mu_k\}$. A *subdivision* of a multiset M is a multiset S of l non-empty submultisets $M_i = (\bar{M}_i, h_i)$ of M with $l \leq |M|$ that satisfy [21]

a) $\bigcup_{i=1}^{l} \bar{M}_i = \bar{M}$;
b) $\sum_{i=1}^{l} h_i(\mu) = f(\mu)$ for every $\mu \in \bar{M}$.

Roughly speaking, a subdivision is a multiset which plays the same role for M as a partition for a set. A subdivision of the multiset M in (2.15) is denoted by

$$S = \{\underbrace{M_1 \ldots, M_1}_{g(M_1)}, \underbrace{M_2, \ldots, M_2}_{g(M_2)}, \ldots, \underbrace{M_l, \ldots, M_l}_{g(M_l)}\}.$$

Example 2.9. For the multiset $M = \{\mu_1, \mu_1, \mu_2\}$, a subdivision is $S = \{\{\mu_1\}, \{\mu_1, \mu_2\}\}$ or $S = \{\{\mu_1\}, \{\mu_1\}, \{\mu_2\}\}$. In the latter case, we write $S = \{\underbrace{\{\mu_1\}}_{2}, \underbrace{\{\mu_2\}}_{1}\}$.

A `MAPLE` algorithm to find all multiset subdivisions is described in [23]. This procedure is faster than the iterated full partitions of Andrews and Stafford [1], given that it takes into account the multiplicity of all elements of M. The higher this multiplicity is, the more the procedure is efficient. The number of set partitions of $[|M|]$ corresponding to the same subdivision S is

$$d_S = \frac{1}{g(M_1)!\, g(M_2)! \cdots g(M_l)!} \prod_{i=1}^{k} \frac{f(\mu_i)!}{h_1(\mu_i)!\, h_2(\mu_i)! \cdots h_l(\mu_i)!}, \tag{2.16}$$

where $h_j(\mu_i)$ is the multiplicity of μ_i in the submultiset M_j.

Example 2.10. For $M = \{\mu_1, \mu_1, \mu_2\}$ let us consider the subdivision $S = \{\{\mu_1\}, \{\mu_1, \mu_2\}\}$. If we label the elements of $M = \{\mu_1, \mu_1, \mu_2\}$ by the elements of $\{1, 2, 3\}$, the partitions of $\{1, 2, 3\}$ corresponding to $S = \{\{\mu_1\}, \{\mu_1, \mu_2\}\}$ are $1|23$ and $2|13$, so $d_S = 2$. The same result arises from computing (2.16), since for $M_1 = \{\mu_1\}$ and $M_2 = \{\mu_1, \mu_2\}$ we have $g(M_1) = 1$, $g(M_2) = 1$ and $h_1(\mu_1) = 1$, $h_2(\mu_1) = h_2(\mu_2) = 1$. Moreover $f(\mu_1) = 2$ and $f(\mu_2) = 1$.

If we set $(n\boldsymbol{.}\mu)_M = \prod_i (n\boldsymbol{.}\mu_i)^{f(\mu_i)}$, then, from Equivalence (2.12), we obtain

$$(n\boldsymbol{.}\mu)_M \simeq \sum_S d_S\, [n\boldsymbol{.}(\chi\mu)]_S, \tag{2.17}$$

where the summation is over all subdivisions S of the multiset M, the integer d_S is given in (2.16), and

$$[n\boldsymbol{.}(\chi\mu)]_S = \prod_{M_i \in \bar{S}} [n\boldsymbol{.}(\chi\mu_{M_i})]^{\,g(M_i)} \quad \text{with} \quad \mu_M = \prod_{\mu \in \bar{M}} \mu^{f(\mu)}. \tag{2.18}$$

Equivalence (2.17) allows us to generalize Theorem 2.6 to umbral monomials with non-disjoint supports.

Theorem 2.11. [21] *We have*

$$\mathbb{E}\left([n\boldsymbol{.}(\chi\mu)]_S\right) = (n)_{|S|} \prod_{j=1}^{l} m^{g(M_j)}_{h_j(\mu_1), h_j(\mu_2), \ldots, h_j(\mu_k)},$$

where

$$m_{h_j(\mu_1), h_j(\mu_2), \ldots, h_j(\mu_k)} = \mathbb{E}\left[\mu_1^{h_j(\mu_1)} \mu_2^{h_j(\mu_2)} \cdots \mu_k^{h_j(\mu_k)}\right] = \mathbb{E}[\mu_{M_j}].$$

Example 2.8 (continued). In order to compute $\mathbb{E}[(n_{\bullet}\mu_1)^2(n_{\bullet}\mu_2)]$, we use Equivalence (2.17). For $M = \{\mu_1, \mu_1, \mu_2\}$ we have

$$(n_{\bullet}\mu_1)^2(n_{\bullet}\mu_2) \simeq [n.(\chi\mu_1^2\mu_2)] + 2\,[n.(\chi\mu_1)]\,[n.(\chi\mu_1\mu_2)] \\ + [n.(\chi\mu_1^2)][n.(\chi\mu_2)] + [n.(\chi\mu_1)]^2[n.(\chi\mu_2)],$$

and, from Theorem 2.11 with $S = \{\{\mu_1, \mu_1\}, \{\mu_2\}\}$, we get

$$\mathbb{E}[(n_{\bullet}\mu_1)^2(n_{\bullet}\mu_2)] = n\,m_{2,1} + 2(n)_2 m_{1,0}m_{1,1} + (n)_2 m_{2,0}m_{0,1} + (n)_3 m_{1,0}^2 m_{0,1},$$

with $m_{i,j} = E[X^i\,Y^j] = \mathbb{E}[\mu_1^i\,\mu_2^j]$. In order to compute (2.14), Equivalence (2.17) and Theorem 2.11 have to be applied to

$$M = \{\chi_1\mu_1^2, \chi_1\mu_1, \chi_2\mu_1^2\mu_2, \chi_3\mu_1^2\mu_2\}.$$

Since M is a set, the subdivisions S are just set partitions. In difference to the strategy proposed by [105] to solve the same problem, the singleton umbrae in M allow us to speed up the computation. For example, consider the partition $\pi_1 = \{\{\chi_1\mu_1^2, \chi_1\mu_1\}, \{\chi_2\mu_1^2\mu_2, \chi_3\mu_1^2\mu_2\}\}$. We have

$$[n_{\bullet}(\chi\mu)]_{\pi_1} = [n_{\bullet}(\chi\,\chi_1^2\,\mu_1^3)]\,[n_{\bullet}(\chi\,\chi_2\chi_3\,\mu_1^4\mu_2^2)].$$

Since $\chi\,, \chi_2\,, \chi_3$ are uncorrelated umbrae, we have $\mathbb{E}[\chi\,\chi_2\chi_3\,\mu_1^4\mu_2^2] = \mathbb{E}[\mu_1^4\mu_2^2]$, but $\mathbb{E}[\chi\,\chi_1^2\,\mu_1^3] = 0$ since $\mathbb{E}[\chi_1^2] = 0$. Therefore we have

$$\mathbb{E}[n_{\bullet}(\chi\mu)]_{\pi_1}] = 0.$$

This is the trick to speed up the computation: whenever, in one — or more — blocks of the subdivision S, there are at least two umbral monomials involving correlated singleton umbrae, the auxiliary umbra $[n_{\bullet}(\chi\mu)]_S$ has evaluation equal to zero. In the iteration procedure employed to build the subdivision S, this allows us to discard these blocks of the subdivision in the subsequent steps, see [22] for more details.

Further applications of Theorems 2.6 and 2.11 can be found in [22]. See Table 3 of Appendix 1 for a comparison of computation times with procedures existing in the literature.

2.2. **Sheppard's corrections.** In the real world, continuous variables are observed and recorded in finite precision through a rounding or coarsening operation, i.e., a grouping rule. A compromise between the desire to know and the cost of knowing is then a necessary consequence. The literature on grouped data spans different research areas, see for example [46]. Due to the relevance of the method of moments as a tool to learn about a distribution, great attention has been paid in the literature to the computation of moments when data are grouped into classes. The moments computed by means of the resulting grouped frequency distribution are a first approximation to the estimated moments of the parent distribution, but they suffer from the error committed during the process of grouping. The correction for grouping is a sum of two terms, the first depending on the length of the grouping interval, the second being a periodic function of the position. For a continuous r.v., this very old problem was first discussed by Thiele [97], who studied the second term without taking into consideration the first, and then by Sheppard [86] who studied the first term, without taking into consideration the second. Compelling reasons have been given in the literature for neglecting the second term and thus for using Sheppard's corrections, that are nowadays still employed. Quite recently the appropriateness of Sheppard's corrections was re-examined in connection with some applications, see, for example, [15] and [102].

Let us consider a r.v. X with probability density function f, representing the continuous parent distribution underlying data. Denote by a_i its i-th moment (also called *raw moment* to distinguish it from the estimated one),

$$a_i = \int_{-\infty}^{\infty} x^i f(x)\,\mathrm{d}x. \tag{2.19}$$

Let $\tilde{a}_i$ be the i-th moment of the grouped distribution,

$$\tilde{a}_i = \frac{1}{h}\int_{-\infty}^{\infty} x^i \int_{-\frac{1}{2}h}^{\frac{1}{2}h} f(x+z)\,\mathrm{d}z\,\mathrm{d}x, \tag{2.20}$$

with h the length of the grouping intervals. Then the i-th raw moment a_i can be reconstructed via Sheppard's corrections as

$$a_i = \sum_{j=0}^{i} \binom{i}{j} \left(2^{1-j} - 1\right) B_j\, h^j\, \tilde{a}_{i-j}, \tag{2.21}$$

where B_j is the j-th Bernoulli number.

The derivation of Sheppard's corrections was a popular topic in the first half of the last century, see [44] for a historical account. This is because Sheppard deduced Equation (2.21) by using the Euler–Maclaurin summation formula and by assuming that the density function had high order contact with the x-axis at both ends. So there was a considerable controversy on the set of sufficient conditions to be required in order to use formula (2.21). If the rounding lattice is assumed to be random, these sufficient conditions can be removed.

Grouping includes also censoring or splitting data into categories during collection or publication, and so it does not only involve continuous variables. The derivation of corrections for raw moments of a discrete parent distribution followed a different path. They were first given in the Editorial of Vol. 1, no. 1, of *Annals of Mathematical Statistics* (page 111). The method used to develop a general formula was extremely laborious. Some years later, Craig [14] considerably reduced and simplified the derivation of these corrections without requiring any other condition and stating these corrections on the average. Except for the papers of Craig [14] and Baten [7], no attention has been paid to multivariate generalizations of Sheppard's corrections, probably due to the complexity of the resulting formulas.

The symbolic method has allowed us to find such corrections both for continuous and discrete parent distributions in closed-form formulas, which have been naturally extended to multivariate parent distributions [17]. The result is an algorithm to implement all these formulas in any symbolic package [20]. A similar employment of the symbolic method can be performed within wavelet theory. Indeed, the reconstruction of the full moment information a_i in (2.21) through the grouped moments $\tilde{a}_i$ can be seen as some kind of multilevel analysis, see [79, 85].

If the sequence $\{\tilde{a}_i\}_{i\geq 0}$ in (2.20) is umbrally represented by the umbra $\tilde{\alpha}$ and the sequence $\{a_i\}_{i\geq 0}$ in (2.19) is umbrally represented by the umbra α, Sheppard's corrections (2.21) can be obtained by expanding the i-th moment of

$$\alpha \equiv \tilde{\alpha} + h\left(\iota + \frac{1}{2}\right), \tag{2.22}$$

where ι is the Bernoulli umbra. Equivalence (2.22) still holds, if the moments refer to a parent distribution over a finite interval. When a_i is the i-th moment of a discrete population, and $\tilde{a}_i$ is the i-th moment calculated from the grouped frequencies, Equivalence (2.22) needs to be modified to

$$\alpha \equiv \tilde{\alpha} + h\left(\iota + \frac{1}{2}\right) + \frac{h}{m}\left(-1{\boldsymbol .}\iota - \frac{1}{2}\right), \tag{2.23}$$

where we assume that m consecutive values of the discrete population are grouped in a frequency class of width h. In Equivalence (2.23), $-1{\boldsymbol .}\iota$ denotes one more auxiliary umbra, the inverse umbra of the Bernoulli umbra.[8] Sheppard's corrections (2.22) can be recovered from (2.23) letting $m \to \infty$ in a suitable way.

As done in Section 2.1, the generalization of Equivalences (2.22) and (2.23) to the multivariate setting follows by introducing suitable umbral monomials.

Let $(\mu_1, \mu_2, \ldots, \mu_k)$ be a k-tuple of umbral monomials with non-disjoint supports.

Definition 2.12. The sequence $\{m_{t_1,t_2,\ldots,t_k}\}$ is umbrally represented by the k-tuple $(\mu_1, \mu_2, \ldots, \mu_k)$ if $E[\mu_1^{t_1}\mu_2^{t_2}\cdots\mu_k^{t_k}] = m_{t_1,t_2,\ldots,t_k}$ for all positive integers $t_1, t_2, \ldots, t_k$.

[8]Given two umbrae α and γ, they are said to be *inverse* to each other if $\alpha+\gamma \equiv \varepsilon$. We denote the inverse of the umbra α by $-1{\boldsymbol .}\alpha$. An umbra and its inverse are uncorrelated. The inverse of an umbra is not unique, but any two umbrae inverse to any given umbra are similar.

Following the analogy with random vectors, $m_{t_1, t_2, \ldots, t_k}$ is called the multivariate moment of the k-tuple $(\mu_1, \mu_2, \ldots, \mu_k)$ of order $\boldsymbol{t} = (t_1, t_2, \ldots, t_k)$.

According to the latter notation in (2.18), if M is a multiset with support $\bar{M} = \{\mu_1, \mu_2, \ldots, \mu_k\}$ and multiplicities $f(\mu_j) = t_j$ for $j = 1, 2, \ldots, k$, then we may write $E[\mu_M] = m_{t_1, t_2, \ldots, t_k}$. Moreover, assume there exists a multivariate random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_k)$ with joint density $f_{\boldsymbol{X}}(\boldsymbol{x})$ over $\mathbb{R}^k$ such that

$$m_{t_1, t_2, \ldots, t_k} = \int_{\mathbb{R}^k} \boldsymbol{x}^{\boldsymbol{t}} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \tag{2.24}$$

is its multivariate moment of order $\boldsymbol{t}$. Note that, as in the univariate case, without loss of generality we may consider joint densities with range[9] of bounded rectangle type. The moments calculated from the grouped frequencies are

$$\tilde{m}_{t_1, t_2, \ldots, t_k} = \frac{1}{h_1 h_2 \cdots h_k} \int_{A_k} \int_{\mathbb{R}^k} \prod_{j=1}^{k} (x_j + z_j)^{t_j} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{z}, \tag{2.25}$$

where

$$A_k = \Big\{ \boldsymbol{z} = (z_1, z_2, \ldots, z_k) \in \mathbb{R}^k : \\ z_j \in \left(-\frac{1}{2} h_j, \frac{1}{2} h_j \right) \text{ for all } j \in \{1, 2, \ldots, k\} \Big\},$$

and $h_j \in \mathbb{R} - \{0\}$ are the window widths for any component. Then we have the following theorem.

Theorem 2.13. *If the sequence $\{m_{t_1, t_2, \ldots, t_k}\}$ in* (2.24) *is umbrally represented by the k-tuple $(\mu_1, \mu_2, \ldots, \mu_k)$ and the sequence $\{\tilde{m}_{t_1, t_2, \ldots, t_k}\}$ in* (2.25) *is umbrally represented by the k-tuple $(\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_k)$, then*

$$\mu_M \equiv \left[\tilde{\mu} + h \left(\iota + \frac{1}{2} \right) \right]_M, \tag{2.26}$$

where M is the multiset given in (2.15),

$$\left[\tilde{\mu} + h \left(\iota + \frac{1}{2} \right) \right]_M = \prod_{j=1}^{k} \left[\tilde{\mu}_j + h_j \left(\iota_j + \frac{1}{2} \right) \right]^{t_j},$$

and $\{\iota_j\}$ are uncorrelated Bernoulli umbrae.

Equivalence (2.26) is implemented in [20] by means of the following steps. In order to recover expressions of raw multivariate moments in

[9] The range of a joint (probability) density (function) $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the subset of $\mathbb{R}^k$ where $f_{\boldsymbol{X}}(\boldsymbol{x})$ is different from 0.

terms of grouped moments, we need to multiply summations like

$$\sum_{s_j=1}^{i} \binom{i}{s_j} \tilde{\mu}_j^{s_j} h_j^{i-s_j} \left(2^{1-i+s_j} - 1\right) B_{i-s_j},$$

corresponding to the i-th power of $\tilde{\mu}_j + h_j \left(\iota_j + \frac{1}{2}\right)$, and then to replace occurrences of products like $\tilde{\mu}_1^{s_1} \tilde{\mu}_2^{s_2} \cdots \tilde{\mu}_k^{s_k}$ by $\tilde{m}_{s_1, s_2, \ldots, s_k}$. If the multivariate parent distribution is discrete, Equivalence (2.26) has to be updated to

$$\mu_M \equiv \left[\tilde{\mu} + h\left(\iota + \frac{1}{2}\right) + \frac{h}{m}\left(-1.\iota - \frac{1}{2}\right)\right]_M$$

where the symbol on the right-hand side denotes the product

$$\prod_{j=1}^{k} \left[\tilde{\mu}_j + h_j\left(\iota_j + \frac{1}{2}\right) + \frac{h_j}{m_j}\left(-1.\iota_j - \frac{1}{2}\right)\right]^{t_j},$$

and m_j denotes the number of consecutive values grouped in a frequency class of width h_j.

3. Compound Poisson processes

In 1973, Rota [76] claimed that compound Poisson processes are related to polynomial sequences $\{p_i(x)\}_{i \geq 0}$ of binomial type. Such a sequence of polynomials is defined by $p_i(x)$ being of degree i for all positive integers i, $p_0(x) = 1$, and

$$p_i(x+y) = \sum_{j=0}^{i} \binom{i}{j} p_j(x) p_{i-j}(y), \qquad i = 0, 1, \ldots. \tag{3.1}$$

This connection has ramifications into several other areas of analysis and probability: Lagrange expansions, renewal theory, exponential family distributions and infinitely divisible processes. Due to its relevance, different authors have tried to formalize this connection. Stam [91] studied polynomial sequences of binomial type in terms of an integer-valued compound Poisson process. He was especially interested in the asymptotic behavior of the probability generating function $p_i(x)/p_i(1)$ for $i \to \infty$. Partial results are available on the radius of convergence under suitable conditions, involving the power series with coefficients $\{p_i(x)\}_{i \geq 0}$. The resulting theory relies on a complicated system of notations. Constantine and Savit [12] have derived a generalization of Dobińsky's formula by means of compound Poisson processes. Pitman [66] has investigated some probabilistic aspects of Dobińsky's formula.

The theory of Bell umbrae gives a natural way to relate compound Poisson processes to polynomial sequences of binomial type. Historically, this connection was the first application in probability of classical umbral calculus as proposed in 1994. More details are given in [33].

So let us come back to the Bell umbra β, introduced in Section 2. Instead of using Bell numbers, we may characterize this umbra by using its factorial moments. The factorial moments of an umbra are

$$a_{(i)} = \mathbb{E}[(\alpha)_i] = \begin{cases} 1, & \text{if } i = 0, \\ E[\alpha(\alpha-1)\cdots(\alpha-i+1)], & \text{if } i > 0. \end{cases} \tag{3.2}$$

Theorem 3.1. [33] *The Bell umbra β is the unique umbra (up to similarity) with all factorial moments equal to* 1.

In Equation (2.4), set $\mathbb{E}[(m_{\bullet}\alpha)^i] = q_i(m)$ and observe that $q_i(m)$ is a polynomial of degree i in m. Suppose that we replace m by an indeterminate[10] x. The symbol having sequence $\{q_i(x)\}_{i\geq 0}$ as moments is denoted by $x_{\bullet}\alpha$ and called the dot-product of x and α, that is,

$$q_i(x) = \mathbb{E}[(x_{\bullet}\alpha)^i] = \sum_{\lambda \vdash i} (x)_{\nu_\lambda} d_\lambda a_\lambda \tag{3.3}$$

for all positive integers i. A Poisson r.v. Po(x) is represented by $x_{\bullet}\beta$, which is called the *Bell polynomial umbra.*

By using the exponential (partial) Bell polynomials $B_{i,k}$ (cf. [70]), a different way to write $q_i(x)$ in (3.3) is

$$q_i(x) = \mathbb{E}\left[(x_{\bullet}\alpha)^i\right] = \sum_{k=1}^{i} (x)_k B_{i,k}(a_1, a_2, \ldots, a_{i-k+1}). \tag{3.4}$$

Equation (3.4) with α replaced by the Bell umbra β returns the so-called *exponential polynomials* $\Phi_i(x)$,

$$\mathbb{E}\left[(x_{\bullet}\beta)^i\right] = \Phi_i(x) = \sum_{k=1}^{i} (x)_k S(i,k), \tag{3.5}$$

where $S(i,k)$ are Stirling numbers of the second kind. According to [9], these polynomials were introduced by S. Ramanujan in his unpublished notebooks. Later, exponential polynomials were studied by Bell [8] and Touchard [100]. Rota, Kahaner and Odlyzko [76] have stated their basic properties via umbral operators. If, in (3.5), we replace x by a non-negative integer n, the resulting umbra $n_{\bullet}\beta$ is the sum of n similar and uncorrelated Bell umbrae, likewise in probability theory where a Poisson r.v. Po(n) is the sum of n i.i.d. Poisson r.v.'s Po(1). More generally, the closure under convolution of Poisson distributions F_t, that is, $F_s \star F_t = F_{s+t}$, is simply reformulated as[11]

$$(t+s)_{\bullet}\beta \equiv t_{\bullet}\beta + s_{\bullet}\beta$$

[10] We replace $\mathbb{R}$ by $\mathbb{R}[x,y]$. The indeterminate y is added because of (3.1).

[11] According to property *iv)* of Corollary 1 in [33], the umbra $(t+s)_{\bullet}\beta$ is similar to $t_{\bullet}\beta + s_{\bullet}\beta'$, with β and β' uncorrelated Bell umbrae. If $s \neq t$ then $s_{\bullet}\beta$ and $t_{\bullet}\beta$ are two distinct symbols, and so uncorrelated. Then we can use $s_{\bullet}\beta$ instead of $s_{\bullet}\beta'$.

with $t, s \in \mathbb{R}$. If, in (3.5), we replace x by a generic umbra α, the auxiliary umbra $\alpha{\boldsymbol .}\beta$ satisfies

$$(\alpha{\boldsymbol .}\beta)^i \simeq \Phi_i(\alpha) = \sum_{k=1}^{i} (\alpha)_k S(i,k).$$

The umbra $\alpha{\boldsymbol .}\beta$ represents a random sum of independent Poisson r.v.'s Po(1) indexed by an integer r.v. Y represented by α, that is, a randomized Poisson r.v. Po(Y) with parameter Y.

If we swap the umbrae α and β, the resulting umbra is the *α-partition umbra* $\beta{\boldsymbol .}\alpha$ representing a compound Poisson r.v. with parameter 1. Let us recall that a compound Poisson r.v. with parameter 1 is a random sum $S_N = X_1+X_2+\cdots+X_N$, where $N \sim$ Po(1) and $\{X_j\}$ are i.i.d.r.v.'s. In the α-partition umbra $\beta{\boldsymbol .}\alpha$, β represents N and α represents one of $\{X_j\}$. Moreover, since the Poisson r.v. Po(x) is umbrally represented by the Bell polynomial umbra $x{\boldsymbol .}\beta$, a compound Poisson r.v. with parameter x is represented by the *polynomial α-partition* umbra $x{\boldsymbol .}\beta{\boldsymbol .}\alpha$. This encodes the connection between compound Poisson processes and polynomial sequences of binomial type, which is what we have claimed at the beginning of this section. The name "partition umbra" has a probabilistic background. Indeed, the parameter of a Poisson r.v. is usually denoted by $x = \lambda t$, with t representing time. When the (time) interval in which t ranges is partitioned into non-overlapping ones, their contributions are stochastically independent and add up to S_N. This circumstance is umbrally expressed by the relation

$$(x+y){\boldsymbol .}\beta{\boldsymbol .}\alpha \equiv x{\boldsymbol .}\beta{\boldsymbol .}\alpha + y{\boldsymbol .}\beta{\boldsymbol .}\alpha, \tag{3.6}$$

paralleling the property (3.1) for the binomial sequences $p_i(x) = \mathbb{E}[(x{\boldsymbol .}\beta{\boldsymbol .}\alpha)^i]$. If $\{a_i\}_{i\geq 0}$ is umbrally represented by the umbra α, moments of $x{\boldsymbol .}\beta{\boldsymbol .}\alpha$ are

$$\mathbb{E}[(x{\boldsymbol .}\beta{\boldsymbol .}\alpha)^i] = \sum_{k=1}^{i} x^k B_{i,k}(a_1, a_2, \ldots, a_{i-k+1}) = \sum_{\lambda \vdash i} d_\lambda \, x^{\nu_\lambda} a_\lambda, \tag{3.7}$$

where $B_{i,k}$ are the partial Bell exponential polynomials. For $x = 1$, we infer from (3.7) that

$$\mathbb{E}[(\beta{\boldsymbol .}\alpha)^i] = \sum_{k=1}^{i} B_{i,k}(a_1, a_2, \ldots, a_{i-k+1}) = \mathscr{Y}_i(a_1, a_2, \ldots, a_i), \tag{3.8}$$

where $\mathscr{Y}_i = \mathscr{Y}_i(a_1, a_2, \ldots, a_i)$ is the i-th complete Bell exponential polynomial [70]. Then the α-partition umbra represents the binomial sequence $\{\mathscr{Y}_i\}$.

The following example proves that partition umbrae are employed in representing r.v.'s with distribution different from the Poisson distribution.

Example 3.2. A Gaussian r.v. $\mathcal{N}(m, s)$ is represented by the umbra $m + \beta\bullet(s\eta)$, where η is an umbra satisfying $E[\eta^i] = \delta_{i,2}$, for all positive integers i. Indeed, the i-th moment of $m + \beta\bullet(s\eta)$ is

$$\mathbb{E}\big\{[m + \beta\bullet(s\eta)]^i\big\} = \sum_{k=0}^{\lfloor i/2 \rfloor} \left(\frac{s^2}{2}\right)^k \frac{(i)_{2k}}{k!} m^{i-2k},$$

which gives the i-th element of a well-known sequence of orthogonal polynomials, the *Hermite polynomials* $H_i^{(\nu)}(x)$ [70], with $m = x$ and $s^2 = -\nu$.

Theorem 3.3. [33] *The α-partition umbra satisfies the equivalence*

$$(\beta\bullet\alpha)^i \simeq \alpha\,(\beta\bullet\alpha + \alpha)^{i-1}, \quad i = 1, 2, \ldots, \tag{3.9}$$

and conversely every umbra satisfying Equivalence (3.9) *is an α-partition umbra.*

One more property of the α-partition umbra involves *disjoint sums.* The disjoint sum $\alpha \dot{+} \gamma$ of α and γ is an auxiliary umbra representing the sequence $\{a_i + g_i\}_{i\geq 0}$ with $\{a_i\}_{i\geq 0}$ and $\{g_i\}_{i\geq 0}$ umbrally represented by the umbra α and γ, respectively.

Theorem 3.4. [34] $\beta\bullet(\alpha \dot{+} \gamma) \equiv \beta\bullet\alpha + \beta\bullet\gamma$.

If in the α-partition umbra $\beta\bullet\alpha$, we replace the umbra β by a different umbra γ, we get a new auxiliary umbra, the dot-product $\gamma\bullet\alpha$. More formally, let us consider the polynomial $q_i(x)$ in (3.3), and suppose to replace x by an umbra γ with moments $\{g_i\}_{i\geq 0}$. The polynomial

$$q_i(\gamma) = \sum_{\lambda \vdash i} (\gamma)_{\nu_\lambda} d_\lambda a_\lambda \tag{3.10}$$

is an umbral polynomial with support $\{\gamma\}$. The auxiliary umbra $\gamma\bullet\alpha$ is an auxiliary umbra with moments

$$\mathbb{E}[(\gamma\bullet\alpha)^i] = \mathbb{E}[q_i(\gamma)], \tag{3.11}$$

for all positive integers i. A very special dot-product umbra is $\chi\bullet\alpha$, with χ the singleton umbra. This umbra is the keystone in dealing with sequences of cumulants. More details will be given in Section 4. Indeed we will show that any umbra is the partition umbra of $\chi\bullet\alpha$, that is, the symbolic method of moments is like a calculus of measures on Poisson algebras. In the following, we give an example of how to use the umbra $\chi\bullet\alpha$ to disclose a not well-known connection between partition polynomials and Kailath–Segall polynomials.

Example 3.5 (The Kailath–Segall formula). Let $\{X_t\}_{t\geq 0}$ be a centered Lévy process[12] with moments of all orders. The variations of

[12] A centered Lévy process is a zero mean stochastic process with independent and stationary increments [80], see also Section 3.1.

the process $\left\{X_t^{(i)}\right\}_{t\geq 0}$ are defined as $X_t^{(1)} = X_t$ and

$$X_t^{(2)} = [X,X]_t = \sigma^2 t + \sum_{0<s\leq t} (\Delta X_s)^2, \; X_t^{(i)} = \sum_{0<s\leq t} (\Delta X_s)^i, \quad \text{for } i \geq 3,$$

where $\Delta X_s = X_s - X_{s\text{-}}$. The iterated stochastic integrals

$$P_t^{(0)} = 1, \;\; P_t^{(1)} = X_t, \;\; P_t^{(i)} = \int_0^t P_{s\text{-}}^{(i-1)} \mathrm{d}X_s, \;\; \text{for } i \geq 2,$$

are related to the variations $\left\{X_t^{(i)}\right\}_{t\geq 0}$ by the Kailath–Segall formula [47]

$$P_t^{(i)} = \frac{1}{i}\left(P_t^{(i-1)} X_t^{(1)} - P_t^{(i-2)} X_t^{(2)} + \cdots + (-1)^{i+1} P_t^{(0)} X_t^{(i)}\right). \tag{3.12}$$

It follows by induction that $P_t^{(i)}$ is a polynomial in $X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(i)}$, called the i-th *Kailath–Segall polynomial.*

Theorem 3.6. [30] *If* $\{\Upsilon_t\}_{t\geq 0}$ *and* $\{\psi_t\}_{t\geq 0}$ *are two families of umbrae with* $\mathbb{E}[\Upsilon_t^i] = i! E\left[P_t^{(i)}\right]$ *and* $\mathbb{E}[\psi_t^i] = E\left[X_t^{(i)}\right]$, *for all positive integers* i, *then*

$$\Upsilon_t \equiv \beta_{\bullet}[(\chi_{\bullet}\chi)\psi_t] \quad \textit{and} \quad (\chi_{\bullet}\chi)\psi_t \equiv \chi_{\bullet}\Upsilon_t. \tag{3.13}$$

Equivalences (3.13) give the umbral version of the Kailath–Segall formula (the former) and its inversion (the latter). The proof essentially relies on Theorem 3.3, since Equation (3.12) is equivalent to $\mathbb{E}[\Upsilon_t^i] = \mathbb{E}\left[\psi_t(\Upsilon_t + \psi_t)^{i-1}\right]$. By the former Equivalence (3.13), the Kailath–Segall polynomials result in the complete Bell exponential polynomials (3.8) in the variables

$$\left\{(-1)^{i-1}(i-1)! E\left[X_t^{(i)}\right]\right\}.$$

The latter similarity in Equivalence (3.13) is a generalization of

$$\chi_{\bullet}(\chi_1 x_1 + \cdots + \chi_n x_n) \equiv (\chi_{\bullet}\chi)\sigma,$$

which gives the elementary symmetric polynomials in terms of power sum symmetric polynomials, umbrally represented by the umbra σ, see [21]. That is, if we replace the jumps $\{\Delta X_s\}$ in $X_t^{(i)}$ by suitable indeterminates $\{x_s\}$, then the Kailath–Segall polynomials reduce to the polynomials given in [4]. These polynomials deserve further investigations in connection with stochastic integrals [62, 78], for which we believe that an approach via the symbolic method of moments is fruitful.

Theorem 3.6 gives the symbolic representation of Lévy process variations. The next subsection is devoted to the symbolic representation of Lévy processes. This representation extends the claimed connection between binomial sequences and compound Poisson processes to the more general class of infinite divisible distributions.

3.1. **Time-space harmonic polynomials.** Lévy processes are stochastic processes playing a fundamental role in various fields such as physics, engineering and actuarial science [6]. They are particularly popular in mathematical finance as models for jumps or spikes of asset prices in financial markets.

The martingale property is one of the fundamental mathematical properties which underlies many finance models, when no knowledge of the past price process can help to predict the future asset price. In particular, a market is said to be efficient if the price process is modeled by a martingale (for more details see [82]).

Lévy processes are not martingales if they have an increasing or decreasing drift. One tool to compensate this fault is to use Wald's martingale, which has various applications [53]. If $\{X_t\}_{t \geq 0} = \{X_t\}$ is a Lévy process,[13] then the Wald exponential martingale is defined as

$$M_t = \frac{\exp\{z X_t\}}{E[\exp\{z X_t\}]}. \tag{3.14}$$

Definition (3.14) requires the existence of the moment generating function $E[\exp\{z X_t\}]$ in a suitable neighborhood of the origin. A way to overcome this gap is to look for polynomials $P(x,t)$, with the property that, when the indeterminate x is replaced by a Lévy process $\{X_t\}$, the resulting stochastic process is a martingale. These polynomials have been introduced first by Neveu [56] for random walks and then by Sengupta [43] for Lévy processes. They are called *time-space harmonic polynomials.*

Definition 3.7. A family of polynomials $\{P(x,t)\}_{t \geq 0}$ is called *time-space harmonic* with respect to a stochastic process $\{X_t\}_{t \geq 0}$ if

$$E[P(X_t, t) \mid \mathfrak{F}_s] = P(X_s, s),$$

for $0 \leq s \leq t$, where $\mathfrak{F}_s = \sigma\,(X_\tau : 0 \leq \tau \leq s)$ is the natural filtration[14] associated with $\{X_t\}_{t \geq 0}$.

Wald's martingale (3.14) has been recently used in order to characterize time-space harmonic polynomials [84], but without reaching a closed expression.

The aim of this subsection is to show how to recover a very general class of time-space harmonic polynomials by using a symbolic representation of Lévy processes. In particular, our attention focuses on special properties of Lévy processes, namely [80]

a) $X_0 = 0$ a.s. (almost surely);
b) for all $n \geq 1$ and for all $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n < \infty$, the r.v.'s $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent;

[13] When no misunderstanding is possible, we simply use $\{X_t\}$ to denote a stochastic process.

[14] A natural filtration $\mathfrak{F}_t$ is the σ-algebra generated by the pre-images $X_s^{-1}(B)$ for Borel subsets B of $\mathbb{R}$ and times s with $0 \leq s \leq t$.

c) for $0 \leq s \leq t$, $X_{t+s} - X_s$ and X_t are identically distributed (stationarity).

Due to these properties, at time t the information of the process $\{X_t\}$ may be split into i.i.d. subprocesses:

$$X_t \stackrel{d}{=} \underbrace{\Delta X_{t/n} + \Delta X_{t/n} + \cdots + \Delta X_{t/n}}_{n}. \tag{3.15}$$

Property (3.15) characterizes the so-called *infinitely divisible r.v.'s*, which are in one-to-one correspondence with Lévy processes. This is also why Lévy himself refers to these processes as a sub-class of additive processes. Starting from (2.4), a symbolic version of a Lévy process can be obtained by means of the same steps as employed in the construction of $q_i(x)$ at the beginning of Section 3. If the positive integer m is replaced by $t \in \mathbb{R}$ in (2.4), then the auxiliary umbra $t{\boldsymbol .}\alpha$ satisfies

$$\mathbb{E}[(t{\boldsymbol .}\alpha)^i] = q_i(t) = \sum_{\lambda \vdash i} (t)_{\nu_\lambda} \, d_\lambda \, a_\lambda, \tag{3.16}$$

for all positive integers i. The umbra $t{\boldsymbol .}\alpha$ is the *dot-product* of t and α. Since $m{\boldsymbol .}\alpha$ is a summation of m uncorrelated and similar umbrae, $t{\boldsymbol .}\alpha$ parallels the decomposition (3.15). Its properties are the same as the ones of Lévy processes when they admit finite moments. Here there are some examples.

Binomial property. The i-th moment $q_i(t) = E(X_t^i)$ of a Lévy process is a polynomial function of t which satisfies the property (3.1) of binomial sequences. In terms of umbrae, this property is equivalent to

$$(t+s){\boldsymbol .}\alpha \equiv t{\boldsymbol .}\alpha + s{\boldsymbol .}\alpha. \tag{3.17}$$

We cannot miss the analogy between Equivalences (3.6) and (3.17). Indeed, an umbra representing a Lévy process is similar to an umbra representing a compound Poisson process. This is because for any umbra α there exists an umbra γ satisfying $\alpha \equiv \beta{\boldsymbol .}\gamma$, so that $t{\boldsymbol .}\alpha \equiv t{\boldsymbol .}\beta{\boldsymbol .}\gamma$. The proof of this statement is given in Section 4.

Additivity property. If $\{W_t\}$ and $\{Z_t\}$ are two independent Lévy processes, then the process $\{X_t\}$ with $X_t = W_t + Z_t$ is a Lévy process. In terms of umbrae, the additivity property corresponds to

$$t{\boldsymbol .}(\alpha + \gamma) \equiv t{\boldsymbol .}\alpha + t{\boldsymbol .}\gamma, \quad \text{with } \alpha, \gamma \in \mathscr{A}.$$

Homogeneity property. If $\{X_t\}$ is a Lévy process and $c \in \mathbb{R}$, then $\{c\, X_t\}$ is a Lévy process. This homogeneity property is equivalent to

$$t{\boldsymbol .}(c\alpha) \equiv c(t{\boldsymbol .}\alpha), \quad \text{with } \alpha \in \mathscr{A}.$$

Nested processes. If $\{X_t\}$ is a Lévy process, then $\{(X_t)_s\}$ is a Lévy process, that is, a nested Lévy process is again a Lévy process. The analog for $t{\boldsymbol .}\alpha$ is $t{\boldsymbol .}(s{\boldsymbol .}\alpha) \equiv s{\boldsymbol .}(t{\boldsymbol .}\alpha) \equiv (st){\boldsymbol .}\alpha$, with $t, s \in \mathbb{R}$.

According to Definition 3.7, in order to verify that an umbral polynomial is a time-space harmonic polynomial, a suitable notion of conditional evaluation has been introduced in the polynomial ring $\mathbb{R}[x][\mathscr{A}]$, see [30]. Denote by $\mathcal{X}$ the set $\mathcal{X} = \{\alpha\}$.

Definition 3.8. The linear operator $\mathbb{E}(\,\cdot\,|\,\alpha) : \mathbb{R}[x][\mathscr{A}] \longrightarrow \mathbb{R}[\mathcal{X}]$ satisfying

i) $\mathbb{E}(1\,|\,\alpha) = 1$;
ii) $\mathbb{E}(x^m \alpha^n \gamma^i \varsigma^j \cdots \,|\, \alpha) = x^m \alpha^n \mathbb{E}(\gamma^i)\mathbb{E}(\varsigma^j)\cdots$ for uncorrelated umbrae $\alpha, \gamma, \varsigma, \ldots$ and for positive integers $m, n, i, j, \ldots$

is called *conditional evaluation* with respect to α.

In other words, the conditional evaluation with respect to α deals with the umbra α as an indeterminate. As it happens for r.v.'s, the conditional evaluation is an element of $\mathbb{R}[x][\mathscr{A}]$, and the overall evaluation of $\mathbb{E}(p\,|\,\alpha)$ gives $\mathbb{E}(p)$, with $p \in \mathbb{R}[x][\mathscr{A}]$.

The conditional evaluation needs to be carefully handled when dot products are involved. Let us observe that the conditional evaluation with respect to $n.\alpha$ satisfies

$$\mathbb{E}[(n+1).\alpha\,|\,n.\alpha] = \mathbb{E}(n.\alpha + \alpha'\,|\,n.\alpha) = n.\alpha + \mathbb{E}(\alpha'),$$

with α' an umbra similar to α. By similar arguments, for all positive integers n and m, since $\mathbb{E}\{[(n+m).\alpha]^i\,|\,n.\alpha\} = \mathbb{E}\{[n.\alpha + m.\alpha']^i\,|\,n.\alpha\}$, we have

$$\mathbb{E}\{[(n+m).\alpha]^i\,|\,n.\alpha\} = \sum_{j=0}^{i} \binom{i}{j} (n.\alpha)^j \mathbb{E}[(m.\alpha')^{i-j}]. \tag{3.18}$$

Equation (3.18) suggests how to define the conditional evaluation with respect to the auxiliary umbra $s.\alpha$.

Definition 3.9. The conditional evaluation of $t.\alpha$ with respect to the auxiliary umbra $s.\alpha$ is

$$\mathbb{E}[(t.\alpha)^i\,|\,s.\alpha] = \sum_{j=0}^{i} \binom{i}{j} (s.\alpha)^j \mathbb{E}\{[(t-s).\alpha']^{i-j}\}.$$

Theorem 3.10. [17] *For all non-negative integers i, the family of polynomials*

$$Q_i(x,t) = \mathbb{E}[(-t{\boldsymbol .}\alpha + x)^i] \in \mathbb{R}[x] \tag{3.19}$$

is time-space harmonic with respect to $\{t{\boldsymbol .}\alpha\}_{t\geq 0}$, that is,

$$\mathbb{E}\left[Q_i(t{\boldsymbol .}\alpha, t)\,|\,s{\boldsymbol .}\alpha\right] = Q_i(s{\boldsymbol .}\alpha, s) \qquad \textit{for } s,t \geq 0. \tag{3.20}$$

The umbra $-t{\boldsymbol .}\alpha$ in (3.19) is a special auxiliary umbra with the remarkable property

$$-t{\boldsymbol .}\alpha + t{\boldsymbol .}\alpha' \equiv \varepsilon, \tag{3.21}$$

where $\alpha' \equiv \alpha$. Due to Equivalence (3.21), the umbra $-t{\boldsymbol .}\alpha$ is the *inverse*[15] umbra of $t{\boldsymbol .}\alpha$. Its moments can be obtained via Equation (3.16) by replacing t by $-t$. The family $\{Q_i(x,t)\}_{i\geq 0}$ is a basis for the space of time-space harmonic polynomials [29] so that many families of known time-space harmonic polynomials can be recovered as a linear combination of $\{Q_i(x,t)\}_{i\geq 0}$. For example, Hermite polynomials are time-space harmonic with respect to Wiener processes, and Poisson–Charlier polynomials are time-space harmonic with respect to Poisson processes. More examples are given in Appendix 2.

Remark 3.11. Theorem 3.10 may be proved when the Lévy process admits moments up to a finite order m. In this case, the polynomials $Q_i(x,t)$ in Equation (3.19) are defined up to $i \leq m$. Then formal power series are replaced by polynomials of degree m, and operations like summation and multiplication among formal power series can be performed in terms of polynomials [5].

Example 3.12 (Random walks)**.** When the parameter t is replaced by a positive integer n, Theorem 3.10 still holds. The sequence $\{n{\boldsymbol .}\alpha\}_{n\geq 0}$ represents a random walk $\{X_n\}_{n\geq 0}$ with $X_0 = 0, X_n = M_1 + M_2 + \cdots + M_n$, and $\{M_n\}_{n\geq 0}$ a sequence of i.i.d.r.v.'s. Note that the sequence $\{M_n\}_{n\geq 0}$ satisfies $M_0 = X_0 = 0$ and $M_n = X_n - X_{n-1}$, for all positive integers n.

To recover the connection between Wald's martingale (3.14) and the polynomials $Q_i(x,t)$, we have to introduce the *generating function* of an umbra α, that is, the formal power series [34]

$$f(\alpha, z) = 1 + \sum_{i\geq 1} a_i \frac{z^i}{i!} \in \mathbb{R}[x][[z]],$$

whose coefficients are the moments of the umbra α. Formal power series allow us to work with generating functions [106] which do not have a positive radius of convergence or have indeterminate coefficients [92]. Similar umbrae $\alpha \equiv \gamma$ have the same generating function $f(\alpha,z) = f(\gamma,z)$. This property allows us to recover the generating functions of the auxiliary umbrae. Table 2 gives some examples.

In particular, the generating function of the time-space harmonic polynomial umbra $-t{\boldsymbol .}\alpha + x$ is

$$f(-t{\boldsymbol .}\alpha + x, z) = \frac{\exp\{xz\}}{f(\alpha,z)^t} = 1 + \sum_{i\geq 1} Q_i(x,t) \frac{z^i}{i!}. \tag{3.22}$$

[15]Since $-t{\boldsymbol .}\alpha$ and $t{\boldsymbol .}\alpha$ are two distinct symbols, they can be considered uncorrelated. When no confusion arises, we will use the symbol $t{\boldsymbol .}\alpha$ instead of $t{\boldsymbol .}\alpha'$.

TABLE 2. Special generating functions

Umbrae	Generating functions
Unity umbra u	$f(u,z)=1$
Singleton umbra χ	$f(\chi,z)=1+z$
Bell umbra β	$f(\beta,z)=\exp[e^z-1]$
Summation $\alpha+\gamma$	$f(\alpha+\gamma,z)=f(\alpha,z)f(\gamma,z)$
Dot product $n.\alpha$	$f(n.\alpha,z)=f(\alpha,z)^n$
Dot product $t.\alpha$	$f(t.\alpha,z)=f(\alpha,z)^t$
Inverse $-t.\alpha$	$f(-t.\alpha,z)=1/f(\alpha,z)^t$

By replacing x by $t.\alpha$ in (3.22), we can compare Equation (3.22) with the following series expansion for Wald's exponential martingale (3.14)

$$\frac{\exp\{zX_t\}}{E[\exp\{zX_t\}]}=1+\sum_{i\geq 1}R_i(X_t,t)\frac{z^i}{i!}.$$

The algebraic structure of formal power series is isomorphic to sequences endowed with the convolution product, each series corresponding to the sequence of its coefficients [92]. Thus, equality of two formal power series is interpreted as equality of their corresponding coefficients, that is, $E[R_i(X_t,t)]=\mathbb{E}[Q_i(t.\alpha,t)]$.

Since $\mathbb{E}[Q_i(t.\alpha,t)]=0$, we have $E[R_i(X_t,t)]=0$ for $i\geq 1$ and

$$1+\sum_{i\geq 1}E[R_i(X_t,t)]\frac{z^i}{i!}=1$$

in agreement with Wald's identity [56]. Therefore, the sequence $\{E[R_i(X_t,t)]\}_{t\geq 0}$ is umbrally represented by the augmentation umbra ε, with $f(\varepsilon,z)=1$. But this is exactly what happens if in $-t.\alpha+x$ we replace x by $t.\alpha$, due to (3.21).

Time-space harmonic polynomials with respect to $\{t.\alpha\}_{t\geq 0}$ are characterized by the following theorem.

Theorem 3.13. [17] *A polynomial $P(x,t)=\sum_{j=0}^k p_j(t)\,x^j$, of degree k for all $t\geq 0$, is time-space harmonic with respect to a Lévy process $\{X_t\}$, represented by $t.\alpha$, if and only if*

$$p_j(t)=\sum_{i=j}^{k}\binom{i}{j}p_i(0)\,\mathbb{E}[(-t.\alpha)^{i-j}],\qquad \textit{for } j=0,\ldots,k.$$

An open problem is the problem of extending the symbolic method of moments to matrix-valued stochastic processes by using Lévy processes and time-space harmonic polynomials. Matrix-value time-space harmonic polynomials include the matrix-valued counterparts of classical polynomials, as the Hermite, Laguerre or Jacobi polynomials, which have gained increasing interest in multivariate statistics. Different methods have been proposed in the literature, see [51] and references therein, by using eigenvalues of random matrices or hypergeometric functions of matrix argument. It would be interesting to check if dealing with hypergeometric functions by means of the symbolic method leads to more feasible expressions as happens for the multivariate generating functions, see Section 7.

4. Cumulants

A probability measure $\mathbb{P}$ is infinitely divisible if for any positive integer n there is a probability measure $\mathbb{P}_n$ such that $\mathbb{P} = (\mathbb{P}_n)^{n*}$, where $(\mathbb{P}_n)^{n*}$ denotes the n-fold convolution of $\mathbb{P}_n$. This property parallels Equation (3.15) for Lévy processes $\{X_t\}$ and can be generalized. Indeed, if we denote by $\mathbb{P}_t$ the probability measure of X_t, then $\mathbb{P}_t = (\mathbb{P}_1)^{t*}$ with $\mathbb{P}_1$ the probability measure of X_1. Since the convolution of probability measures corresponds to a summation of r.v.'s, the r.v. X_t can be expressed as a summation of t times the r.v. X_1. The symbolic method of moments allows us to generalize this decomposition by replacing $t \in \mathbb{R}$ by the umbra γ. The resulting auxiliary umbra is the dot-product of α and γ in (3.10), that represents a summation of γ times the umbra α. A special dot-product is the umbra $\chi.\alpha$, where χ is the singleton umbra, since its moments are cumulants of the sequence $\{a_i\}_{i\geq 0}$ umbrally represented by α.

Among the number sequences related to r.v.'s, cumulants play a central role characterizing many r.v.'s occurring in classical stochastic processes. For example, a Poisson r.v. is the unique probability distribution for which all cumulants are equal to the parameter. A Gaussian r.v. is the unique probability distribution for which all cumulants vanish beyond the second.

(Formal) cumulants $\{c_j\}$ of a sequence $\{a_i\}$ are usually defined by the identity of formal power series [73]

$$1 + \sum_{i\geq 1} a_i \frac{t^i}{i!} = \exp\left(\sum_{j\geq 1} c_j \frac{t^j}{j!}\right), \tag{4.1}$$

which uniquely determines the non-linear functionals $\{c_j\}$ as polynomials in $\{a_i\}$. In this definition, any questions concerning convergence of involved series may be disregarded [92]. Moreover, many difficulties connected to the so-called problem of cumulants smooth out. Here, with the problem of cumulants, we refer to the characterization of sequences that are cumulants of some probability distribution. The simplest example is that the second cumulant of a probability distribution must always be non-negative, and is zero if and only if all of the higher cumulants are zero. Cumulants are not subject to such constraints when they are analyzed from a symbolic point of view.

Example 4.1 (The variance)**.** The variance $\mathrm{Var}(X) = E(X^2) - E(X)^2$ of a r.v. is the cumulant of second order, an index very useful in statistics. It possesses remarkable properties that may be extended to cumulants of any order:

i) it is invariant under translation: $\mathrm{Var}(X + a) = \mathrm{Var}(X)$ for any constant a;
ii) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$, if X and Y are independent r.v.'s;
iii) $\mathrm{Var}(X)$ is a polynomial in the moments of the r.v. X.

A first approach to the theory of cumulants via the classical umbral calculus was given by Rota and Shen in [74]. By using the symbolic method of moments, the formulas commonly used to express moments of a r.v. in terms of cumulants, and vice-versa, are encoded in a dot-product involving the singleton umbra [34].

Definition 4.2. The α-cumulant umbra κ_α is the auxiliary umbra satisfying $\kappa_\alpha \equiv \chi{\boldsymbol .}\alpha$.

The three main algebraic properties of cumulants can be stated as follows:

i) additivity property:

$$\chi{\boldsymbol .}(\alpha + \gamma) \equiv \chi{\boldsymbol .}\alpha \dot{+} \chi{\boldsymbol .}\gamma,$$

that is, if $\{c_i(X + Y)\}$ is the sequence of cumulants of the summation of two independent r.v.'s X and Y, then $c_i(X+Y) = c_i(X) + c_i(Y)$;

ii) homogeneity property:

$$\chi{\boldsymbol .}(a\alpha) \equiv a(\chi{\boldsymbol .}\alpha), \text{ for all } a \in \mathbb{R};$$

that is, $c_i(aX) = a^i\, c_i(X)$ if $\{c_i(aX)\}$ denotes the sequence of cumulants of aX.

iii) semi-invariance property:

$$\chi{\boldsymbol .}(\alpha + a\, u) \equiv \chi{\boldsymbol .}\alpha \dot{+} a\chi,$$

that is, for all $a \in \mathbb{R}$, we have $c_1(X + a) = c_1(X) + a$ and $c_i(X + a) = c_i(X)$ for all positive integers i greater than 1.

The next proposition follows from Equivalence (3.9) and paves the way to a symbolic handling of Abel sequences, as will become clearer in Section 6.

Proposition 4.3. [32] *If κ_α is the α-cumulant umbra, then, for all positive integers i,*

$$\alpha^i \simeq \kappa_\alpha(\kappa_\alpha + \alpha)^{i-1} \quad \text{and} \quad \kappa_\alpha^i \simeq \alpha(\alpha - 1{\boldsymbol .}\alpha)^{i-1}. \tag{4.2}$$

The former Equivalence (4.2) has been assumed by Rota and Shen as definition of the cumulant umbra [74]. In terms of moments, this equivalence reads

$$a_i = \sum_{j=0}^{i-1} \binom{i}{j} a_j c_{i-j},$$

largely used in statistic framework [65, 87].

Since $\mathbb{E}[(\chi)_i] = (-1)^{i-1}(i-1)!$ for all positive integers i [34], the formula expressing the cumulants $\{c_j\}$ in terms of the moments $\{a_i\}$ is recovered from Equations (3.10) and (3.11), that is,

$$c_j = \sum_{\lambda \vdash j} \mathrm{d}_\lambda (-1)^{\nu_\lambda - 1}(\nu_\lambda - 1)!\, a_\lambda. \tag{4.3}$$

Theorem 4.4 (Inversion theorem)**.** [34] *If κ_α is the α-cumulant umbra, then $\alpha \equiv \beta{\boldsymbol .}\kappa_\alpha$.*

In particular, from (3.8) we have

$$a_i = \mathscr{Y}_i(c_1, c_2, \ldots, c_i) \quad \text{and} \quad a_i = \sum_{\lambda \vdash i} \mathrm{d}_\lambda c_\lambda. \tag{4.4}$$

The inversion theorem in Theorem 4.4 justifies Definition 4.2 since in terms of generating functions we have $f(\alpha, z) = \exp[f(\kappa_\alpha, z) - 1]$, corresponding to Equation (4.1). Moreover, since, from Inversion Theorem 4.4, any umbra α could be seen as the partition umbra of its cumulant umbra κ_α, it is possible to prove a more general result: every polynomial sequence of binomial type is completely determined by the sequence of its (formal) cumulants. On the other hand, we have already seen in the previous section that the polynomial umbra $x{\boldsymbol .}\beta{\boldsymbol .}\kappa_\alpha$ admits a sequence of moments of binomial type, see Equivalence (3.6). Due to Inversion Theorem 4.4, any polynomial sequence of binomial type is umbrally represented by a suitable polynomial umbra $x{\boldsymbol .}\alpha$. This statement gives the connection between Lévy processes and compound Poisson processes. We will get back to this connection in Example 4.5.

If we swap the umbra α and the umbra χ in $\chi{\boldsymbol .}\alpha$, the resulting auxiliary umbra has moments equal to the factorial moments $\{a_{(i)}\}_{i \geq 0}$ of α given in (3.2), that is, $E[(\alpha{\boldsymbol .}\chi)^i] = a_{(i)}$. The umbra $\alpha{\boldsymbol .}\chi$ is called the *α-factorial umbra.* Factorial moments provide very concise expressions for moments of some discrete distributions, such as the binomial distribution [34], for example.

4.1. **Randomized compound Poisson r.v.'s.** The way to characterize the umbra representing a r.v. X is to determine the sequence of its moments $\{a_i\}$. When this sequence exists, this can be done by comparing its moment generating function $E[\exp(zX)]$ with the generating function of the umbra. Recall that, when $E[\exp(zX)] = f(z)$, for z in a suitable neighborhood of the origin, then $f(z)$ admits an exponential expansion in terms of moments, which are completely determined by the corresponding distribution function (and vice-versa). In this case, the moment generating function encodes all the information of X, and the notion of identically distributed r.v.'s corresponds to the similarity among umbrae. In the symbolic method of moments, instead, the convergence of $f(\alpha, z)$ is not relevant [92]. This means that we can define the umbra whose moments are the same as the moments of a log-normal r.v.,[16] even if this r.v. does not admit a moment generating function. In Section 3, generating functions of umbrae have been used to recover Wald's martingale. But generating functions have more advantages. One of which is the encoding of symbolic operations among umbrae in terms of suitable operations among generating functions. For example, Equation (4.1) is encoded in the equivalence $\alpha \equiv \beta.\kappa_\alpha$ of Inversion Theorem 4.4, and, more generally, the exponential of a generating function corresponds to the generating function of the polynomial α-partition umbra

$$f(x{\bf .}\beta{\bf .}\alpha, z) = \exp\left[x\left(f(\alpha, z) - 1\right)\right]. \tag{4.5}$$

Example 4.5 (Lévy processes). If $\{X_t\}$ is a Lévy process, then its moment generating function is $E[e^{zX_t}] = [\phi(z)]^t$, where $\phi(z)$ is the moment generating function of X_1 [80]. This property parallels the infinite divisibility property of $\mathbb{P}_t$ given at the beginning of this section. In particular,

$$E[e^{zX_t}] = \exp[t\,\log\phi(z)] = \exp[t\,k(z)],$$

where $k(z)$ is the cumulant generating function of X_1, satisfying $k(0) = 0$. Comparing $\exp[t\,k(z)]$ with $\exp[t(f(\kappa_\alpha, z) - 1)]$ given in Equation (4.5), with x replaced by t, we infer that

i) a Lévy process is umbrally represented by $t{\bf .}\beta{\bf .}\kappa_\alpha$ representing in turn a compound Poisson process of parameter t;
ii) the umbra κ_α represents the coefficients of $k(z)$, which are the cumulants of X_1;
iii) the compound Poisson process of parameter t involves i.i.d.r.v.'s $\{X_i\}_{i\geq 0}$ whose moments are the cumulants of X_1.

The composition of generating functions corresponds to an iterated dot-product. Indeed, moments of the auxiliary umbra $(\gamma.\beta){\bf .}\alpha$ can be obtained from (3.7), by replacing x by γ:

$$\mathbb{E}[\Phi_i(\gamma{\bf .}\beta)] = \sum_{\lambda\vdash i} d_\lambda\, g_{\nu_\lambda} a_\lambda. \tag{4.6}$$

The auxiliary umbra $(\gamma.\beta){\bf .}\alpha$ is called the *composition umbra* of α and γ, since its i-th moment corresponds to the i-th coefficient of the composition $f(\gamma, f(\alpha, z) - 1)$.

[16] If X is a standard Gaussian r.v. $\mathcal{N}(0,1)$, then the log-normal r.v. is $Y = \exp(X)$.

Theorem 4.6. [33] *If $\alpha, \gamma \in \mathscr{A}$, then $(\gamma \boldsymbol{.} \beta)\boldsymbol{.}\alpha \equiv \gamma\boldsymbol{.}(\beta\boldsymbol{.}\alpha)$.*

Due to Theorem 4.6, we may write the composition umbra as $\gamma\boldsymbol{.}\beta\boldsymbol{.}\alpha$. The composition umbra $(\gamma\boldsymbol{.}\beta)\boldsymbol{.}\alpha$ represents a compound randomized Poisson r.v., i.e., a random sum indexed by a randomized Poisson r.v. represented by the umbra $\gamma\boldsymbol{.}\beta$. The composition umbra $\gamma\boldsymbol{.}(\beta\boldsymbol{.}\alpha)$ generalizes the random sum of i.i.d. compound Poisson r.v.'s with parameter 1, indexed by an integer r.v. X, i.e., a randomized compound Poisson r.v. with random parameter X represented by the umbra γ.

Paralleling the composition of generating functions, the *compositional inverse* of α is the auxiliary umbra $\alpha^{<-1>}$ having generating function $f(\alpha^{<-1>}, z)$ satisfying

$$f[\alpha^{<-1>}, f(\alpha, z) - 1] = f[\alpha, f(\alpha^{<-1>}, z) - 1] = 1 + z$$

or $\alpha\boldsymbol{.}\beta\boldsymbol{.}\alpha^{<-1>} \equiv \alpha^{<-1>}\boldsymbol{.}\beta\boldsymbol{.}\alpha \equiv \chi$, since $f(\chi, z) = 1 + z$.

Theorem 4.7. [27] *An umbra α has a compositional inverse $\alpha^{<-1>}$ if and only if $\mathbb{E}[\alpha] = a_1 \neq 0$.*

A very special umbra is the compositional inverse $u^{<-1>}$ of the unity umbra u which satisfies

$$u\boldsymbol{.}\beta\boldsymbol{.}u^{<-1>} \equiv u^{<-1>}\boldsymbol{.}\beta\boldsymbol{.}u \equiv \chi. \tag{4.7}$$

The following result is very useful in performing computations with dot-products.

Proposition 4.8. [34] $\chi \equiv u^{<-1>}\boldsymbol{.}\beta \equiv \beta\boldsymbol{.}u^{<-1>}$ *and* $\beta\boldsymbol{.}\chi \equiv u \equiv \chi\boldsymbol{.}\beta$.

The compositional inverse of an umbra is closely related to the Lagrange inversion formula, but we will use Sheffer polynomial sequences in Section 5 to compute its moments.

We end this section by introducing the connection between composition umbrae and convolutions of multiplicative functions. Multiplicative functions have been used to define free and Boolean cumulants by means of the lattice of non-crossing and interval partitions, respectively [58, 90]. This is an issue that we will return to again in Section 6.

Example 4.9. [32] Let Π_n be the set of all partitions of $[n]$ equipped with the usual refinement order $\leq$, and denote the minimum and the maximum by $\mathbf{0}_n$ and $\mathbf{1}_n$, respectively. Denote by $|\pi|$ the number of blocks of $\pi \in \Pi_n$. If $\sigma \in \Pi_n$ and $\sigma \leq \pi$, then there exists a unique sequence of non-negative integers $(k_1, k_2, \ldots, k_n)$ with $k_1 + 2\,k_2 + \cdots + n\,k_n = |\sigma|$ and $k_1 + k_2 + \cdots + k_n = |\pi|$ such that

$$[\sigma, \pi] = \Pi_1^{k_1} \times \Pi_1^{k_1} \times \cdots \times \Pi_n^{k_n},$$

with $[\sigma, \pi] = \{\tau \in \Pi_n : \sigma \leq \tau \leq \pi\}$. The sequence $(k_1, k_2, \ldots, k_n)$ is called the *type* of the interval $[\sigma, \pi]$, and k_i is the number of blocks of π that are the union of i blocks of σ. A function $\mathfrak{f} : \Pi_n \times \Pi_n \to \mathbb{C}$ is said to be *multiplicative* if $\mathfrak{f}(\sigma, \pi) = f_1^{k_1} f_2^{k_2} \cdots f_n^{k_n}$ with $\sigma \leq \pi$ and

$f_n := \mathfrak{f}(\mathbf{0}_n, \mathbf{1}_n)$. The Möbius function μ is multiplicative with $\mu_n = (-1)^{n-1}(n-1)!$. Since the sequence $\{(-1)^{n-1}(n-1)!\}$ is represented by the compositional inverse of the unity umbra u, the umbral counterpart of the Möbius function μ is $u^{<-1>}$. The zeta function ζ is multiplicative with $\zeta_n = 1$. The umbral counterpart of the zeta function ζ is u. As the umbra $u^{<-1>}$ is the compositional inverse of u, the Möbius function corresponds to the inverse of the zeta function. Indeed, the convolution $\star$ between two multiplicative functions $\mathfrak{f}$ and $\mathfrak{g}$ is defined by

$$(\mathfrak{f} \star \mathfrak{g})(\sigma, \pi) := \sum_{\sigma \leq \tau \leq \pi} \mathfrak{f}(\sigma, \tau)\, \mathfrak{g}(\tau, \pi).$$

The function $\mathfrak{h} = \mathfrak{f} \star \mathfrak{g}$ is also multiplicative with $h_n = (\mathfrak{f} \star \mathfrak{g})(\mathbf{0}_n, \mathbf{1}_n)$ and

$$h_n = \sum_{\pi \in \Pi_n} f_\pi\, g_{\ell(\pi)}, \tag{4.8}$$

where $f_\pi := f_1^{k_1} f_2^{k_2} \cdots f_n^{k_n}$ and k_i is the number of blocks of π of cardinality i. The convolution of two multiplicative functions (4.8) corresponds to a composition umbra since Equation (4.6) can be indexed by set partitions [21], that is,

$$\mathbb{E}[(\gamma . \beta . \alpha)^n] = \sum_{\pi \in \Pi_n} g_\pi\, a_{|\pi|}.$$

The identity with respect to the convolution $\star$ is the Delta function Δ with $\Delta_n = \delta_{1,n}$. Its umbral counterpart is the singleton umbra χ. The Möbius function μ and the zeta function ζ are inverses of each other with respect to $\star$, that is, $\mu \star \zeta = \zeta \star \mu = \Delta$. So, this last equality parallels Equivalences (4.7).

4.2. **Lévy processes.** In Section 3.1, we have stressed the central role played by Lévy processes in modeling financial markets. They are preferable to Brownian motion since deviation from normality can be explained by the jump or spike component. The symbolic representation of Lévy processes $\{X_t\}$ by means of the family of auxiliary umbrae $\{t . \alpha\}_{t \geq 0}$ takes into account the strong interplay between these processes and infinitely divisible distributions, but does not reveal their jump-diffusion structure. The same considerations hold when α is replaced by $\beta . \kappa_\alpha$, which discloses the role played by cumulants of X_1 in the distribution of X_t, but does not add more information on paths of a Lévy process. Instead, Lévy processes are also called *jump-diffusion processes* since they result from a summation of a Brownian motion and a compensated compound Poisson process,[17] a property that characterizes unexpected changes in the paths, if any. The following theorem gives the jump-diffusion decomposition of a Lévy process.

[17]A compensated Poisson process is $\{N_t - \lambda t\}$ with $\{N_t\}$ a Poisson process $\mathrm{Po}(\lambda t)$.

Theorem 4.10. [80] *If $\{X_t\}$ is a Lévy process, then*

$$X_t = m\,t + s\,\mathcal{B}_t + \sum_{k=1}^{N_t} J_k - t\,a\,\lambda,$$

where $m \in \mathbb{R}, s > 0, \{\mathcal{B}_t\}$ is a standard Brownian motion, $\{N_t\}$ is a Poisson process $\mathrm{Po}(\lambda t)$*, and $\{J_k\}_{k\geq 1}$ are i.i.d.r.v.'s with $E(J) = a < \infty$.*

All sources of randomness are mutually independent. In Theorem 4.10, the diffusion component of a Lévy process is represented by the Wiener process $m\,t + s\,\mathcal{B}_t$, while the jump component is represented by the compensated compound Poisson process $\left(\sum_{k=1}^{N_t} J_k - ta\lambda\right)$. The symbolic method reduces the jump-diffusion decomposition to a compound Poisson process. This symbolic representation is obtained by taking into consideration the so-called Lévy–Khintchine formula, which splits the increment X_1 in a summation according to the jump-diffusion property.

Theorem 4.11 (Lévy–Khintchine formula)**.** [82] *If $\{X_t\}$ is a Lévy process, then $E[e^{zX_t}] = [\phi(z)]^t$, with*

$$\phi(z) = \exp\left\{ z\,m + \frac{1}{2}\,s^2\,z^2 + \int_{\mathbb{R}} \left(e^{zx} - 1 - zx\boldsymbol{1}_{\{|x|\leq 1\}}\right) \mathrm{d}\,\nu(x) \right\}. \tag{4.9}$$

The triple (m, s^2, ν) is called *Lévy triple*, and ν is the *Lévy measure*. Since in Example 4.5 we have seen that $f(t\boldsymbol{.}\beta\boldsymbol{.}\kappa_\alpha, z) = [\phi(z)]^t$, with $f(\kappa_\alpha, z) = \phi(z)$, what we need is to characterize the umbra κ_α according to (4.9). First, let us observe that the Lévy measure ν contains much information on X_t. If $\nu(\mathbb{R}) < \infty$, almost all paths of X_t have a finite number of jumps on every compact interval. Under suitable normalization, this corresponds to the condition $\int_{\mathbb{R}} \mathrm{d}\,\nu(x) = 1$, in agreement with $\mathbb{E}[1] = 1$. If ν admits all moments, then Equation (4.9) may be rewritten as

$$\phi(z) = \exp\left(c_0 z + \frac{1}{2} s^2 z^2\right) \exp\left\{ \int_{\mathbb{R}} (e^{zx} - 1 - zx)\,\mathrm{d}\,\nu(x) \right\} \tag{4.10}$$

where $c_0 = m + \int_{\{|x|\geq 1\}} x\,\mathrm{d}\,\nu(x)$. Equation (4.10) allows us to split the umbra κ_α in a disjoint sum, as the following theorem shows.

Theorem 4.12. *A Lévy process $\{X_t\}$ is umbrally represented by the family of auxiliary umbrae*

$$\{t\boldsymbol{.}\beta\boldsymbol{.}[c_0\chi \dot{+} s\eta \dot{+} \gamma]\}_{t\geq 0},$$

where γ is the umbra associated to the Lévy measure, that is, $f(\gamma, z) = 1 + \int_{\mathbb{R}} (e^{zx} - 1 - zx)\,\mathrm{d}\,\nu(x)$, and η is an umbra with $f(\eta, z) = 1 + z^2/2$.

The jump-diffusion decomposition in Theorem 4.10 parallels the symbolic representation in Theorem 4.12, since from Theorem 3.4 we have

$$t\boldsymbol{.}\beta\boldsymbol{.}[c_0\chi \dot{+} s\eta \dot{+} \gamma] \equiv t\boldsymbol{.}\beta\boldsymbol{.}[c_0\chi \dot{+} s\eta] + t\boldsymbol{.}\beta\boldsymbol{.}\gamma,$$

and the umbra (see Example 3.2)

$$\beta\boldsymbol{.}[c_0\chi \dot{+} s\eta] \equiv c_0 + \beta\boldsymbol{.}(s\eta)$$

represents a Gaussian r.v. with mean c_0 (called *drift*) and variance s^2. The compensated Poisson process is instead umbrally represented by the auxiliary umbra $t\boldsymbol{.}\beta\boldsymbol{.}\gamma$. Indeed, the umbra γ can be further decomposed into a pure point part, corresponding to the moment generating function $\int_{\mathbb{R}} e^{zx}\,\mathrm{d}\,\nu(x)$, and a singular part, corresponding to the moment generating function $\int_{\mathbb{R}}(1+z\,x)\,\mathrm{d}\,\nu(x)$. The singular part of the Lévy measure is encoded in the umbra χ, which indeed does not have a probabilistic counterpart.

The singleton umbra also plays a special role with respect to the martingale property of a Lévy process. A centered Lévy process is a process with $E[X_t]=0$ for all $t \geq 0$, and, in this case, almost all paths of X_t have finite variation.

Theorem 4.13. [3] *A Lévy process is a martingale if and only if its drift is* $c_0 = 0$.

A centered Lévy process is umbrally represented by the auxiliary umbra $\{t\boldsymbol{.}\beta\boldsymbol{.}(s\eta \dot{+} \gamma)\}_{t\geq 0}$, where the explicit contribution of the singleton umbra is not visible anymore.

5. Sheffer polynomial sequences

Many characterizations of Sheffer polynomial sequences are given in the literature. We refer to the one involving generating functions.

Definition 5.1 (Sheffer sequences)**.** A sequence $\{s_i(x)\}$ of polynomials is called a *Sheffer sequence* if and only if its generating function is given by

$$1 + \sum_{i\geq 1} s_i(x)\frac{t^i}{i!} = h(t)\exp\{x\,g(t)\},$$

with $h(t) = h_0 + h_1 t + h_2 t^2 + \cdots$, $g(t) = g_1 t + g_2 t^2 + \cdots$, and $h_0, g_1 \neq 0$.

Applications of Sheffer sequences range from analysis to statistics, from combinatorics to physics. A good account of fields where these polynomials are employed is given in [16]. Their importance stems from the fact that many polynomial sequences, such as Laguerre polynomials, Meixner polynomials of first and second kind, Bernoulli polynomials, Poisson–Charlier polynomials, and Stirling polynomials are Sheffer polynomial sequences. Moreover, the time-space harmonic polynomials introduced in Section 3 are special Sheffer sequences as well.

Sheffer sequences can be considered as the core of the umbral calculus as introduced by Roman and Rota in [71]. There, one finds several chapters devoted to identities, properties and recurrences involving Sheffer sequences. Following these properties, Taylor has characterized their coefficients by using the classical umbral calculus [96]. In order to handle Sheffer polynomials by the symbolic method of moments, a special auxiliary umbra has been introduced in [27]. One of the main advantages of this symbolic representation is that properties of Sheffer sequences are reduced to a few equivalences.

In order to keep the length of the paper within reasonable bounds, here we limit ourselves to presenting some applications of the symbolic method involving Sheffer polynomial sequences. Whenever necessary, we recall the umbral equivalences we need. For a more complete treatment of this subject, the reader should consult [27].

The first application we deal with concerns the finding of solutions to linear recurrences. In many special combinatorial problems, the hardest part in their solution may be the discovery of an effective recursion. Once a recursion has been established, Sheffer polynomials are often a simple and general tool for finding solutions in closed form. The main contributions in this respect are due to Niederhausen [59] (see references therein), by means of finite operator calculus [76]. Using the symbolic method of moments, computations are simplified [36]. We give several examples below.

A second application involves the Lagrange inversion formula. This formula gives the coefficients of the compositional inverse of a formal power series. Indeed, in Definition 5.1 the formal power series $g(t)$ has a compositional inverse, since $g_1 \neq 0$. There are many different proofs of this formula, many of them assuming that the involved power series are convergent and represent analytic functions on a disk. Using the symbolic representation of Sheffer sequences, the proof reduces to a simple computation [27].

A third and last application concerns Riordan arrays. A(n exponential) *Riordan array* is a pair $(g(t), f(t))$ of (exponential) formal power series, where $g(t)$ is an invertible series and $f(0) = 0$. The pair defines an infinite lower triangular array according to the rule

$$d_{i,k} = i\text{-th coefficient of } \; g(t)\frac{[f(t)]^k}{k!}, \quad \text{for } 0 \leq k \leq i < \infty.$$

Riordan arrays are used in a constructive way to find the generating function of many combinatorial sums, to characterize families of orthogonal polynomials, and to compute determinants of Hankel or Toeplitz matrices. Using the symbolic method of moments, we give a characterization of Riordan arrays. This characterization is very general since the same holds for the coefficients of Sheffer polynomials and connection constants [27]. If $\{s_i(x)\}$ and $\{r_i(x)\}$ are Sheffer sequences, the constants $b_{i,k}$ in the expression

$$r_i(x) = \sum_{k=0}^{i} b_{i,k} s_k(x) \tag{5.1}$$

are known as *connection constants*. Several changes of bases among polynomial sequences may be efficiently computed by using connection constants. Connection constants, Riordan arrays, and coefficients of Sheffer polynomials share the same expression in terms of *generalized Bell polynomials* [27]. These polynomials are a generalization of complete Bell exponential polynomials (3.8) and also play a fundamental role in the multivariate version of the symbolic method.

The starting point of the symbolic representation of Sheffer sequences is the following definition.

Definition 5.2. A polynomial umbra ς_x is called a *Sheffer umbra for* (α, γ) if and only if $\varsigma_x \equiv \alpha + x \boldsymbol{.} \gamma^*$, where $\gamma^* \equiv \beta \boldsymbol{.} \gamma^{<-1>}$ is called the *adjoint umbra* of γ.

Of course, to be well posed, in Definition 5.2 the umbra γ has to admit the compositional inverse $\gamma^{<-1>}$, that is, $E[\gamma] = g_1 \neq 0$. From Definition 5.2, the generating function of a Sheffer umbra is

$$f(\varsigma_x, z) = f(\alpha, z) \, \exp\{x \, [f^{<-1>}(\gamma, z) - 1]\},$$

in agreement with Definition 5.1.

5.1. **Solving linear recurrences.** In order to show how to solve linear recurrences by using the symbolic method, we present two examples, the former involving a uniform r.v., the latter Dyck paths [27].

Example 5.3. Suppose we are asked to solve the difference equation

$$q_i(x+1) = q_i(x) + q_{i-1}(x) \tag{5.2}$$

under the condition $\int_0^1 q_i(x) \, \mathrm{d}x = 1$ for all positive integers i. This initial condition is equivalent to $\mathbb{E}[q_i(U)] = 1$, where U is a uniform r.v. on the interval $(0,1)$. Set $q_i(x) = s_i(x)/i!$. Then, from (5.2), we have

$$s_i(x+1) = s_i(x) + i \, s_{i-1}(x). \tag{5.3}$$

A polynomial umbra ς_x is a Sheffer umbra if and only if there exists an umbra γ, equipped with compositional inverse, which satisfies $\varsigma_{\gamma + x \boldsymbol{.} u} \equiv \chi + \varsigma_x$ [27], that is, if and only if $s_i(\gamma + x \boldsymbol{.} u) = s_i(x) + i \, s_{i-1}(x)$ for all positive integers i. By comparing this last equality with (5.3), we have to choose as umbra γ the unity umbra u. Since $u^* \equiv \chi$, we have $q_i(x) \simeq (x \boldsymbol{.} \chi + \alpha)^i / i!$ for α depending on the initial condition. By recalling that $\int_0^1 p(x)\,\mathrm{d}x = \mathbb{E}[p(-1{\boldsymbol .}\iota)]$ for any $p(x) \in \mathbb{R}[x]$ [17], we infer

$$\int_0^1 q_i(x)\,\mathrm{d}x = 1 \quad \text{if and only if} \quad \mathbb{E}[q_i(-1{\boldsymbol .}\iota)] = 1$$
$$\text{if and only if} \quad \mathbb{E}[s_i(-1{\boldsymbol .}\iota)] = i!.$$

Let the sequence $\{i!\}$ be represented by the umbra $\bar{u}$, called the *Boolean unity umbra.*[18] Then, from $\mathbb{E}[q_i(-1{\boldsymbol .}\iota)] = 1$, we have

$$-1{\boldsymbol .}\iota{\boldsymbol .}\chi + \alpha \equiv \bar{u} \quad \text{if and only if} \quad \alpha \equiv \bar{u} + \iota{\boldsymbol .}\chi.$$

Solutions of the recursion (5.2) are moments of the Sheffer umbra $(\iota + \beta + x{\boldsymbol .}u){\boldsymbol .}\chi$ divided by $i!$ □

Example 5.4. A ballot path takes up-steps (u) and right-steps (r), starting at the origin and staying weakly above the diagonal. For example, *ururuur* is a ballot path (starting at the origin and) terminating at the point $(3,4)$. Let $D(i,m)$ be the number of ballot paths terminating at (i,m) which do not contain the pattern (substring) *urru*.

It is not difficult to see that $D(i,m)$ satisfies the recurrence relation

$$D(i,m) - D(i-1,m) = D(i,m-1) - D(i-2,m-1) + D(i-3,m-1)$$

under the initial conditions $D(i,i) = D(i-1,i)$ and $D(0,0) = 1$. Set $D(i,m) = s_i(m)/i!$. The previous recurrence relation may be written as

$$s_i(m) - i\,s_{i-1}(m) = s_i(m-1) - (i)_2\,s_{i-2}(m-1) + (i)_3\,s_{i-3}(m-1), \tag{5.4}$$

with initial condition $s_i(i) = i\,s_{i-1}(i)$. Replace m by x in Equation (5.4). The umbral counterpart of the *Sheffer identity*[19] [27] is $\varsigma_{x+y} \equiv \varsigma_x + y{\boldsymbol .}\gamma^*$, which yields $\varsigma_x \equiv \varsigma_{x-1} + \gamma^*$ for $y = -1$. Since $s_i(x) - i\,s_{i-1}(x) \simeq (\varsigma_x - \chi)^i$, we may replace ς_x by $\varsigma_{x-1} + \gamma^*$ so that

$$(\varsigma_x - \chi)^i \simeq (\varsigma_{x-1} + \gamma^* - \chi)^i. \tag{5.5}$$

Since

$$(\varsigma_{x-1} + \gamma^* - \chi)^i \simeq \varsigma_{x-1}^i + i \sum_{j=0}^{i-1} \binom{i-1}{j} \varsigma_{x-1}^{i-j-1} \left[\frac{(\gamma^*)^{j+1}}{j+1} - (\gamma^*)^j \right],$$

from Equivalence (5.5) and Equation (5.4), we infer

$$f(\gamma^*, t) = 1 + t + \sum_{k \geq 3} t^k = (1 - t^2 + t^3)/(1-t).$$

---

[18] Its moments are called Boolean cumulants, introduced in Section 6. For Boolean cumulants, the umbra $\bar{u}$ plays the same role as the unity umbra u for classical cumulants.

[19] *Sheffer identity:* A sequence $\{s_i(x)\}$ of Sheffer polynomials satisfies $s_i(x+y) = \sum_{k=0}^{i} \binom{i}{k} s_k(x) s_{i-k}(y)$.

The initial condition $s_i(i) = i\, s_{i-1}(i)$ is equivalent to looking for an umbra α satisfying

$$(\alpha + i\,\boldsymbol{.}\gamma^*)^i \simeq i\,(\alpha + i\,\boldsymbol{.}\gamma^*)^{i-1}. \tag{5.6}$$

Assume $\alpha \equiv \bar{u} + \zeta$. In this case the initial condition (5.6) gives $(\zeta + i\,\boldsymbol{.}\gamma^*)^i \simeq \varepsilon^i$, and then $\zeta^i \simeq -i\,[\gamma^{<-1>}]^i$ for all positive integers i. The solution of the recurrence relation (5.4) is

$$\frac{s_i(x)}{i!} \simeq \frac{(\bar{u} + \zeta + x\boldsymbol{.}\gamma^*)^i}{i!}. \qquad \square$$

Two special classes of Sheffer polynomials are often employed in applications: polynomial sequences of binomial type and Appell sequences. In Section 3, we have already seen that polynomial sequences of binomial type are umbrally represented by the polynomial umbra $x\boldsymbol{.}\beta\boldsymbol{.}\gamma$, which is a Sheffer umbra for (ε, γ).

An *Appell sequence* is a sequence of polynomials $\{p_i(x)\}$ satisfying the identity

$$\frac{\mathrm{d}}{\mathrm{d}x}\, p_i(x) = i\, p_{i-1}(x), \quad \text{for } i = 1, 2, \ldots.$$

This sequence of polynomials is represented by a Sheffer umbra for (α, χ). The resulting umbra $\alpha + x\boldsymbol{.}u$ is called *Appell umbra.* Examples of Appell polynomials are the *Bernoulli polynomials* $q_i(x) = E[(\iota + x\boldsymbol{.}u)^i]$ introduced in Section 2, the *Euler polynomials*

$$q_i(x) = E\left[x\boldsymbol{.}u + \left(\frac{1}{2}[\xi - u]\right)^i\right],$$

with ξ the Euler umbra, and the time-space harmonic polynomials $Q_i(x,t) = E[(x + t\boldsymbol{.}\alpha)^i]$ introduced in Section 3.

5.2. **Lagrange inversion formula.** The Lagrange inversion formula gives the coefficients of the compositional inverse of a formal power series and thus of the compositional inverse of an umbra γ. Following Rota, Shen and Taylor [77], the proof we propose relies on the circumstance that any sequence of binomial type can be represented by Abel polynomials $\{x\,(x - i\,a)^{i-1}\}$, $a \in \mathbb{R}$. We use *umbral Abel polynomials* $\{x\,(x - i\boldsymbol{.}\gamma)^{i-1}\}$. An exhaustive treatment of umbral Abel polynomials can be found in [63]. In contrast to [77], the employment of the singleton umbra reduces the proof of the Lagrange inversion formula to a simple computation. We need a preliminary definition.

Definition 5.5. If $\{g_i\}$ is a sequence umbrally represented by an umbra γ, then its *derivative umbra* γ_D has moments $g_{D,i} = i\, g_{i-1}$ for all positive integers i.

The derivative umbra has moments obtained from a symbolic derivation with respect to γ, that is, $i\, g_{i-1} = \mathbb{E}[i\,\gamma^{i-1}] = \mathbb{E}[\partial_\gamma \gamma^i]$.

Theorem 5.6 (Abel representation of binomial sequences). [27] *If γ is an umbra equipped with a compositional inverse, then*

$$(x{\boldsymbol .}\gamma^*_D)^i \simeq x\,(x - i{\boldsymbol .}\gamma)^{i-1}, \quad \textit{for } i = 1, 2, \ldots. \tag{5.7}$$

Theorem 5.6 includes the well-known Transfer Formula [71].

Theorem 5.7 (Lagrange inversion formula). [27]

$$(\gamma_D{}^{<-1>})^i \simeq (-i{\boldsymbol .}\gamma)^{i-1}, \quad \textit{for } i = 1, 2, \ldots.$$

Proof. In (5.7), replace x by the singleton umbra χ and observe that

$$(\chi{\boldsymbol .}\beta{\boldsymbol .}\gamma_D{}^{<-1>})^i \simeq (\gamma_D{}^{<-1>})^i \simeq \chi\,(\chi - i\,{\boldsymbol .}\gamma)^{i-1},$$

due to $\chi.\beta \equiv u$ (see Proposition 4.8). The result follows by applying the binomial expansion to $\chi\,(\chi - i{\boldsymbol .}\gamma)^{i-1}$ and by recalling that $\chi^{k+1} \simeq 0$ for $k = 1, 2, ..., i-1$. $\square$

As $-i{\boldsymbol .}\gamma \equiv i{\boldsymbol .}(-1.\gamma)$, moments of $\gamma_D{}^{<-1>}$ can be computed by using the inverse of γ. The umbra γ_D has first moment equal to 1. It is possible to generalize Theorem 5.7 to umbrae having first moment different from 1, see [27].

5.3. **Generalized Bell polynomials.** By Theorem 5.6, the umbral analog of the well-known *Abel identity*[20] is:

$$(x+y)^i \simeq \sum_{k=0}^{i} \binom{i}{k} (x + k{\boldsymbol .}\gamma)^{i-k} y (y - k{\boldsymbol .}\gamma)^{k-1}. \tag{5.8}$$

The polynomials

$$\mathbb{B}^{(\gamma)}_{i,k}(x) = \binom{i}{k} (x + k{\boldsymbol .}\gamma)^{i-k},$$

given in Equation (5.8), are called *generalized Bell umbral polynomials* since $E[\mathbb{B}^{(\gamma)}_{i,k}(0)] = B_{i,k}\,(g_{D,1}; \ldots\ \ldots; g_{D,i-k+1})$, where $B_{i,k}$ are the partial Bell exponential polynomials given in Equation (3.7), and $\{g_{D,i}\}$ are the moments of the umbra γ_D. Paralleling the complete exponential Bell polynomials, the generalized complete Bell umbral polynomials are

$$\mathbb{Y}^{(\gamma)}_i(x) = (x + \beta{\boldsymbol .}\gamma_D)^i \simeq \sum_{k=0}^{i} \binom{i}{k} (x + k{\boldsymbol .}\gamma)^{i-k} = \sum_{k=0}^{i} \mathbb{B}^{(\gamma)}_{i,k}(x).$$

Generalized Bell polynomials are useful to give a unifying representation of the coefficients of Sheffer polynomials, of Riordan arrays, and of connection constants, as we show in the following. Assume $\mathbb{E}[\gamma] = 1$. Generalizations to the case $\mathbb{E}[\gamma] \neq 1$ are given in [27]. We need a preliminary definition.

Definition 5.8. If $\{g_i\}$ is a sequence umbrally represented by an umbra γ, then its *primitive umbra* γ_P has moments $g_{P,i} = g_{i+1}/(i+1)$ for all positive integers i.

[20] The *Abel identity* is: $(x+y)^i = \sum_{k=0}^{i} \binom{i}{k} (x + k\,a)^{i-k} y (y - k\,a)^{k-1}$, $a \in \mathbb{R}$.

The primitive umbra has moments obtained from a symbolic integration with respect to γ, that is,

$$\frac{g_{i+1}}{i+1} = \mathbb{E}\left(\frac{\gamma^{i+1}}{i+1}\right) = \mathbb{E}\left(\int \gamma^i \mathrm{d}\gamma\right).$$

For $0 \leq k \leq i$, let $c_{i,k}$ denote the k-th coefficient of the i-th Sheffer polynomial $s_i(x)$.

Theorem 5.9 (Coefficients of Sheffer polynomials)**.** [27] *If $\{s_i(x)\}$ is a polynomial sequence umbrally represented by a Sheffer umbra for the pair (α, γ), then $c_{i,k} = \mathbb{E}\left[\mathbb{B}_{i,k}^{(\gamma_P)}(\alpha)\right]$ for all positive integers i and $0 \leq k \leq i$.*

Theorem 5.10 (Riordan array)**.** [27] *The elements of an exponential Riordan array for the pair $(f(\alpha, t), f^{<-1>}(\gamma, t)-1)$ are umbrally represented by $d_{i,k} = \mathbb{E}\left[\mathbb{B}_{i,k}^{(\gamma_P)}(\alpha)\right]$ for all positive integers i and $0 \leq k \leq i$.*

So, the theory of Riordan arrays and that of Sheffer sequences are the two sides of the same coin.

Theorem 5.11 (Connection constants)**.** [27] *If $\{s_i(x)\}$ is a polynomial sequence umbrally represented by a Sheffer umbra for (α, γ) and $\{r_i(x)\}$ is a polynomial sequence umbrally represented by a Sheffer umbra for (ξ, ζ), then $b_{i,k} = \mathbb{E}\left[\mathbb{B}_{i,k}^{(\phi_P)}(\varrho)\right]$ with $\{b_{i,k}\}$ given in (5.1) and $\varrho \equiv (-1.\alpha.\beta.\gamma + \xi.\beta.\zeta).\zeta^*$, $\phi \equiv \zeta.\gamma^*$.*

6. Fast symbolic computation

There are many packages devoted to numerical/graphical statistical toolsets but not doing algebraic/symbolic computations. Quite often, the packages filling this gap are not open source. One of the softwares largely used in the statistical community is `R`.[21] `R` is a much stronger numerical programming environment, and the procedures including symbolic software are not yet specifically oriented for statistical calculations. The availability of a widely spread open source symbolic platform will be of great interest, especially with interface capabilities to external programs. The conceptual aspects related to symbolic methods involve more strictly algebraic and combinatorial issues, and the topics introduced in the previous sections are an example.

What we regard as symbolic computation is now evolving towards a universal algebraic language which aims at combining syntactic elegance and computational efficiency. The classical umbral calculus is no doubt an old topic. However, the version here reviewed has not only the virtue of improving the syntactic elegance of the umbral theory but also to offer a different and deeper viewpoint, for example on the role played by cumulants in working with compound Poisson processes.

[21]Available at (`http://www.r-project.org/`)

In this section, we show how this different and deeper viewpoint allows us to reach computational efficiency as by-product. These examples should push statisticians to climb the steep learning curve of a new method and mathematicians to believe what Terence Tao claimed in his blog: "A different choice of foundations can lead to a different way of thinking about the subject, and thus to ask a different set of questions and to discover a different set of proofs and solutions. Thus it is often of value to understand multiple foundational perspectives at once, to get a truly stereoscopic view of the subject."

Two examples are proposed here. The first concerns efficient algorithms aiming at the computation of experimental measurements of cumulants, known in the literature as *polykays*. The sampling behavior of polykays is much simpler than sample moments [54], but their employment was not so widespread in the statistical community due to the past computational complexity in recovering their expression. Only recently, the symbolic method of moments has contributed in speeding up the procedures for their computations [23]. The second example shows how to recover classical, Boolean, and free cumulants from moments and vice-versa, by using only one family of umbral polynomials. Quite surprisingly this algorithm is an application of the symbolic representation of Sheffer sequences making any of the existing procedures obsolete. This part has a consistent theoretical background devoted to the symbolic representation of classical, Boolean and free cumulants, their generalization to Abel-type cumulants, and some open problems which are still under investigation.

6.1. **k-statistics and polykays.** In dealing with computations, the approach used to compute U-statistics and estimators of multivariate moments is not always efficient. So, further manipulations are required. As an example, we show how to generate k-statistics, which are U-statistics for cumulants, by using a suitable compound Poisson r.v. The i-th k-statistic k_i is the unique symmetric unbiased estimator of the cumulant c_i of a given statistical distribution [94], that is, $E[k_i] = c_i$.

A frequently asked question is: why is it so relevant to get efficiency in these calculations? Usually higher order objects require enormous amounts of data in order to get a good accuracy. Nevertheless, there are different areas, such as astronomy [69], astrophysics [40], biophysics [55], and neuroscience [93], in which there is need to compute high order k-statistics in order to detect a Gaussian population. Undoubtedly, a challenging target is to have efficient procedures to deal with the involved huge amount of algebraic and symbolic computations.

The theory of k-statistics has a long history beginning with Fisher [41], who rediscovered the theory of half-invariants of Thiele [98]. Fisher introduced k-statistics (single and multivariate) as new symmetric functions of a random sample, with the aim of estimating cumulants without using moment estimators. Dressel [38] developed a theory of more general functions, revisited later by Tukey [101], who called them *polykays*. The whole subject is described by Stuart and Ord [94] in great detail. In the 1980s, tensor notation was used by Speed [88] and extended to polykays and single k-statistics. This extension uncovers the coefficients defining polykays to be values of the Möbius function on the lattice of set partitions. As a consequence, Speed used the set-theoretic approach to symmetric functions introduced by Doubilet [37]. In the same period, McCullagh [54] simplified the tensor notation of Kaplan [48] by introducing the notion of generalized cumulants. Symbolic operators for expectation and the derivation of unbiased estimators for multiple sums were introduced by Andrews *et al.* [1].

A classical way to recover k-statistics is to use Theorem 2.6 in (4.3). The resulting U-statistics are expressed in terms of augmented symmetric polynomials. Then, to get k-statistics, we have to express augmented symmetric polynomials in terms of power sum symmetric polynomials.

Theorem 6.1. [21] *For $i \leq n$, we have*[22]

$$(\chi{\boldsymbol .}\alpha)^i \simeq \sum_{\lambda \vdash i} \frac{(\chi{\boldsymbol .}\chi)^{\nu_\lambda}}{(n)_{\nu_\lambda}} d_\lambda \sum_{\pi \in \Pi_{\nu_\lambda}} (\chi{\boldsymbol .}\chi)^{{\boldsymbol .}\pi} (n{\boldsymbol .}\alpha)_{S_\pi}, \tag{6.1}$$

where S_π is the subdivision of the multiset $P_\lambda = \{\alpha^{(r_1)}, \alpha^{2(r_2)}, \ldots\}$ corresponding to the partition $\pi \in \Pi_{\nu_\lambda}$.

A significant computational cost is required to find set partitions or multiset subdivisions, although these procedures are already optimized [22]. By using the symbolic method, an improvement in the performance is achievable. The main idea is to recover k-statistics as cumulants of compound Poisson r.v.'s replacing the augmented symmetric polynomials by exponential polynomials (3.5). The implementation is then sped up since there is no need to compute set partitions.

To this aim, we introduce the *multiplicative inverse* of an umbra.

Definition 6.2. The *multiplicative inverse* of an umbra α is the umbra γ satisfying $\alpha\gamma \equiv u$.

In dealing with a saturated umbral calculus, the multiplicative inverse of an umbra is not unique, but any two multiplicatively inverse umbrae of the umbra α are similar. From Definition 6.2, we have $a_n g_n = 1$ for all non-negative integers n, i.e., $g_n = 1/a_n$, where $\{a_n\}$ and $\{g_n\}$ are umbrally represented by the umbrae α and γ, respectively. In the following, the multiplicative inverse of an umbra α will

[22] For the notation, see Subsection 2.1.

be denoted by the symbol $1/\alpha$. The following proposition shows how to express cumulants of an umbra by suitable polynomials linked to moments of compound Poisson r.v.'s.

Theorem 6.3. [21] *If* $c_i(y) = E[(n_{\bullet}\chi_{\bullet}y_{\bullet}\beta_{\bullet}\alpha)^i]$, *then*

$$(\chi_{\bullet}\alpha)^i \simeq c_i\left(\frac{\chi_{\bullet}\chi}{n_{\bullet}\chi}\right), \quad \text{for } i = 1, 2, \ldots.$$

Thus, moments of the α-cumulant umbra can be computed by evaluating the umbral polynomials $c_i(y)$, after having replaced the indeterminate y by the umbra $\chi_{\bullet}\chi/n_{\bullet}\chi$. The polynomials $c_i(y)$ are moments of a summation of n uncorrelated cumulant umbrae of $y_{\bullet}\beta_{\bullet}\alpha$, the latter in its turn being a polynomial α-partition umbra representing a compound Poisson r.v. of parameter y. The last step is to express the polynomials $c_i(y)$ in terms of power sum symmetric polynomials $n_{\bullet}\alpha^r$ for $r \leq n$.

Theorem 6.4. [21] *If*

$$p_n(y) = \sum_{k=1}^{n} (-1)^{k-1}(k-1)!\, S(n,k)\, y^k,$$

where $S(n,k)$ *are the Stirling numbers of the second kind, then*

$$(\chi_{\bullet}\alpha)^i \simeq \sum_{\lambda \vdash i} d_\lambda\, p_\lambda\left(\frac{\chi_{\bullet}\chi}{n_{\bullet}\chi}\right)(n_{\bullet}\alpha)^{r_1}(n_{\bullet}\alpha^2)^{r_2}\cdots, \tag{6.2}$$

with $p_\lambda(y) = [p_1(y)]^{r_1}[p_2(y)]^{r_2}\cdots$.

Comparing Equivalence (6.2) with Equivalence (6.1), the reduction of complexity is evident since no set partitions are involved.

Theorem 6.4 has been generalized to polykays. Polykays are symmetric statistics $k_{r,\ldots,t}$ satisfying

$$E[k_{r,\ldots,t}] = c_r \cdots c_t,$$

where $c_r, \ldots, c_t$ are cumulants. For simplicity, in the following we just deal with two subindices $k_{r,t}$, the generalization to more than two being straightforward. As a product of uncorrelated cumulants, $(\chi_{\bullet}\alpha)^r(\chi'_{\bullet}\alpha')^t$ is the umbral counterpart of the polykay $k_{r,t}$, with χ, χ' uncorrelated singleton umbrae, and α, α' uncorrelated and similar umbrae. For polykays, the polynomial

$$p_{r,t}(y) = \sum_{(\lambda \vdash r,\, \eta \vdash t)} y^{\nu_\lambda + \nu_\eta}\, (n)_{\nu_\lambda + \nu_\eta}\, d_{\lambda+\eta}\, a_{\lambda+\eta}$$

plays the same role as the polynomial $c_i(y)$ in Theorem 6.3. The following theorem is the analog of Theorem 6.4 for polykays.

Theorem 6.5. [21] *If $q_{r,t}$ is the umbral polynomial obtained via $p_{r,t}(y)$ by replacing $y^{\nu_\lambda+\nu_\eta}$ by*

$$\frac{(\chi.\chi)^{\nu_\lambda}(\chi'.\chi')^{\nu_\eta}}{(n.\chi)^{\nu_\lambda+\nu_\eta}}\frac{d_\lambda d_\eta}{d_{\lambda+\eta}},$$

then $(\chi.\alpha)^r(\chi'.\alpha')^t \simeq q_{r,t}$.

Now, let us consider a polynomial umbra $\rho_{y,k}$ whose first k moments are all equal to y, that is, $\rho_{y,k}^i \simeq (\chi.y.\beta)^i$ for $i = 0, 1, 2, \ldots, k$, and all the remaining moments are zero. If $k = \max\{r, t\}$, then

$$[n.(\rho_{y,k}\,\alpha)]^{r+t} \simeq \sum_{\lambda\vdash(r+t)} d_\xi\,(\chi.\rho_{y,k})_\xi\,(n.\alpha)^{s_1}\,(n.\alpha^2)^{s_2}\cdots,$$

which in turn gives the polynomials $p_{r,t}(y)$ in terms of power sum symmetric polynomials, as $[n.(\rho_{y,k}\,\alpha)]^{r+t} \simeq p_{r,t}(y)$. This device together with Theorem 6.5 speeds up the computation of polykays.

The resulting algorithm can be fruitfully employed also in the estimation of products of moments. Indeed, unbiased estimators of products of moments can be computed by using polykays, as we show in the following. Let us consider a set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of n uncorrelated umbrae similar to α. Set $\alpha^{\cdot\pi} = \alpha_{i_1}^{|B_1|}\alpha_{i_2}^{|B_2|}\cdots\alpha_{i_k}^{|B_k|}$, where $\pi = \{B_1, B_2, \ldots, B_k\}$ is a partition of $[n]$ and $i_1, i_2, \ldots, i_k$ are distinct integers chosen in $[n]$. We have $E[\alpha^{\cdot\pi}] = a_\lambda$, where λ is the type of the partition π. The formula giving products of moments in terms of polykays is [21]

$$\alpha^{\cdot\pi} \simeq \sum_{\substack{\tau\in\Pi_n\\ \tau\leq\pi}} (\chi.\alpha)^{\cdot\tau}. \tag{6.3}$$

By using the Möbius inversion

$$(\chi.\alpha)^{\cdot\pi} \simeq \sum_{\substack{\tau\in\Pi_n\\ \tau\leq\pi}} \mu(\tau,\pi)\,\alpha^{\cdot\pi}, \tag{6.4}$$

we get polykays in terms of products of moments. Equivalences (6.3) and (6.4) have also a different meaning. Taking the evaluation of both sides, formulas connecting classical cumulants and moments are obtained in terms of multiplicative functions. These formulas are

$$a_\pi = \sum_{\tau\leq\pi} c_\tau \text{ for all } \pi, \quad \text{if and only if} \quad c_\pi = \sum_{\tau\leq\pi}\mu(\tau,\pi)a_\tau \text{ for all } \pi,$$

where $a_\pi = a_{|B_1|}a_{|B_2|}\cdots a_{|B_k|}$ (the same for c_π), and

$$\mu(\tau,\pi) = (-1)^{n-r}(2!)^{r_3}(3!)^{r_4}\cdots((n-1)!)^{r_n}$$

is the Möbius function for the classical partition lattice, with $n = |\tau|$, $r = |\pi|$, and $\lambda(\tau,\pi) = (1^{r_1}, 2^{r_2}, \ldots, n^{r_n})$ the class of $[\tau,\pi]$.

Since all these formulas can be generalized to correlated umbral monomials (see Definition 2.12), the next section is devoted to the multivariate umbral calculus and its applications in probability and statistics. Since any umbra is a composition umbra, the key to generalize this symbolic calculus to k-tuples of umbral monomials is the multivariate Faà di Bruno formula, which is related to the composition of multivariate formal power series. For the fast algorithms computing multivariate k-statistics and multivariate polykays by using the symbolic method, the reader is referred to [23]. In Appendix 1, Table 5 and 6, we show the advantage of these fast algorithms compared to existing procedures.

6.2. **Computing cumulants through generalized Abel polynomials.**

6.2.1. *Classical, Boolean, and free cumulants.* In Section 4, cumulants have been introduced as a sequence of numbers providing an alternative to moments of a distribution. Lévy processes are an example because their symbolic representation involves cumulants of X_1. Since cumulants linearize the convolution of probability measures in classical, Boolean, and free probability, they have one more advantage: cumulants are a tool to recognize independent r.v.'s. The linearity of classical convolutions corresponds to tensor independent r.v.'s [39]. The linearity of Boolean convolutions corresponds to Boolean independent r.v.'s for which the factorization rule holds, but without the commutative property of the resulting moments. The Boolean convolution [90] was constructed starting from the notion of partial cumulants, which were extensively used in the context of stochastic differential equations. The linearity of free convolutions corresponds to free r.v.'s [104]. The latter belong to the non-commutative probability theory introduced by Voiculescu [103] with a view to tackle problems in the theory of operator algebras and random matrix theory. The combinatorics underlying classical, Boolean, and free cumulants is closely related to the algebra of multiplicative functions (see Example 4.9). Roughly speaking, one can work with classical, Boolean, and free cumulants using the lattice of set, interval, and non-crossing partitions,[23] respectively.

Since the linearity of cumulants gives a quick access to test whether a given probability measure is a convolution, the availability of a fast procedure for computing cumulants — classical, Boolean, and free — from the sequence of moments turns out to be useful. The symbolic method of moments replaces the combinatorics of partition lattices with the algebra of umbral Abel polynomials given in Theorem 5.6. Indeed, all these families of cumulants share the same parametrization in terms of umbral Abel polynomials. The result is an algorithm providing just one formula for computing all types of cumulants in terms of moments and vice-versa.

[23]The combinatorics underlying non-crossing partitions has been first studied by Kreweras [50] and Poupard [68]. Within free probability, non-crossing partitions are extensively used by Speicher [90].

The aim of this section is to introduce this parametrization, to describe the resulting algorithm and some problems that this parametrization lives open. Moreover we define an Abel-type cumulant and explain why the connection among classical, Boolean, and free cumulants cannot be encoded in a straight way in the formal power series language, as pointed out also in [58].

Let us consider again the umbral Abel polynomials $\{x(x - i{\bf .}\alpha)^{i-1}\}$.

Definition 6.6. The umbral polynomial $\mathfrak{a}_i^{(m)}(x, \alpha) = x\,(x + m{\bf .}\alpha)^{i-1}$, with $m \in \mathbb{Z}$, is called the *generalized umbral Abel polynomial.*

We have already used generalized umbral Abel polynomials in Section 4 for classical cumulants. Indeed, from the latter (4.2), moments of the α-cumulant umbra are umbrally equivalent to $\mathfrak{a}_i^{(-1)}(\alpha, \alpha)$

$$\kappa_{\alpha}^{i} \simeq \alpha(\alpha - 1{\bf .}\alpha)^{i-1} \simeq \mathfrak{a}_i^{(-1)}(\alpha, \alpha). \tag{6.5}$$

The former (4.2) inverts Equivalence (6.5) and can be restated as

$$\alpha^i \simeq \kappa_\alpha(\kappa_\alpha + \alpha)^{i-1} \simeq \mathfrak{a}_i^{(1)}(\kappa_\alpha, \alpha).$$

The connection between these two types of generalized umbral Abel polynomials holds in greater generality as the following proposition shows.

Theorem 6.7 (First Abel inversion theorem)**.** [64]

$$\alpha^i \simeq \mathfrak{a}_i^{(m)}(\gamma, \gamma) \simeq \gamma\,(\gamma + m{\bf .}\gamma)^{i-1} \quad \textit{if and only if} \quad \gamma^i \simeq \mathfrak{a}_i^{(-m)}(\alpha, \gamma) \simeq \alpha\,(\alpha - m{\bf .}\gamma)^{i-1}. \tag{6.6}$$

A natural question arising here is the following: does the sequence $\{\mathfrak{a}_i^{(-m)}(\alpha, \alpha)\}$, resulting from (6.5) for positive integers m, have the three main algebraic properties of cumulants?

Let us consider the umbra $\bar{\alpha} \equiv \alpha\,\bar{u}$ whose moments are $\mathbb{E}[\bar{\alpha}^i] = i!\,a_i$, and the sequence $\{\mathfrak{a}_i^{(-2)}(\bar{\alpha}, \bar{\alpha})\}$. By the First Abel Inversion Theorem, given in Theorem 6.7, we have

$$\bar{\eta}_\alpha^i \simeq \mathfrak{a}_i^{(-2)}(\bar{\alpha}, \bar{\alpha}) \simeq \bar{\alpha}\,(\bar{\alpha} - 2{\bf .}\bar{\alpha})^{i-1} \quad \text{if and only if} \quad \bar{\alpha}^i \simeq \mathfrak{a}_i^{(2)}(\bar{\eta}_\alpha, \bar{\alpha}) \simeq \bar{\eta}_\alpha\,(\bar{\eta}_\alpha + 2{\bf .}\bar{\alpha})^{i-1}. \tag{6.7}$$

The umbra $\bar{\eta}_\alpha$ is called the $\bar{\alpha}$-*Boolean cumulant umbra.* The three main algebraic properties of cumulants can be stated as follows:

i) (*additivity*) $\bar{\eta}_\xi \equiv \bar{\eta}_\alpha \dot{+} \bar{\eta}_\gamma$ if and only if $-1{\bf .}\bar{\xi} \equiv -1{\bf .}\bar{\alpha} \dot{+} -1{\bf .}\bar{\gamma}$;

ii) (*homogeneity*) $\bar{\eta}_{c\alpha} \equiv c\bar{\eta}_\alpha$, $c \in \mathbb{R}$;

iii) (*semi-invariance*) $\bar{\eta}_\xi \equiv \bar{\eta}_\alpha \dot{+} c\,\chi$ if and only if $-1{\bf .}\bar{\xi} \equiv -1{\bf .}\bar{\alpha} \dot{+} -1{\bf .}(c\bar{u})$.

Its moments are $\{\mathbb{E}[\mathfrak{a}_i^{(-2)}(\bar{\alpha},\bar{\alpha})]\}$ whose expansion in terms of moments is

$$\mathbb{E}[\bar{\alpha}(\bar{\alpha}-2.\bar{\alpha})^{i-1}] = \sum_{\lambda \vdash i} d_\lambda (-1)^{\nu_\lambda - 1} \nu_\lambda! \, \bar{a}_\lambda, \quad \text{with } \bar{a}_\lambda = r_1! \, a_1^{r_1} \, r_2! \, a_2^{r_2} \cdots$$

by Definition 6.6.

Theorem 6.8. [32] *If $\bar{\eta}_\alpha$ is the $\bar{\alpha}$-Boolean cumulant umbra, then*

$$\bar{\eta}_\alpha \equiv \bar{u}^{<-1>}.\beta.\bar{\alpha} \quad \text{and} \quad \bar{\alpha} \equiv \bar{u}.\beta.\bar{\eta}_\alpha. \tag{6.8}$$

Comparing Equivalences (6.8) with the equivalences characterizing classical cumulants,

$$\kappa_\alpha \equiv \chi.\alpha \equiv u^{<-1>}.\beta.\alpha \qquad \text{and} \qquad \alpha \equiv \beta.\kappa_\alpha \equiv u.\beta.\kappa_\alpha, \tag{6.9}$$

the analogy is evident. The same analogy has a counterpart within formal power series, as the following remark shows.

Remark 6.9. Let us observe that the ordinary formal power series

$$f(z) = 1 + \sum_{i \geq 1} f_i z^i$$

is the generating function of the umbra $\bar{\alpha}$, that is, $f(\bar{\alpha}, z) = f(z)$. If we consider the umbra γ satisfying $f(\bar{\gamma}, z) = g(z) = 1 + \sum_{i\geq 1} g_i z^i$, then the auxiliary umbra $\bar{\alpha}.\beta.\bar{\gamma}$ has generating function $f[g(t)-1]$. This umbra is the counterpart of the convolution of multiplicative functions on the lattice of *interval partitions*. Indeed, let us consider the lattice of interval partitions $(\mathcal{I}_i, \leq)$. A partition π of Π_i is called an *interval partition* if each block B_j of π is an interval $[a_j, b_j]$ of $[i]$, that is, $B_j = [a_j, b_j] = \{x \in [j] \,|\, a_j \leq x \leq b_j\}$, where $a_j, b_j \in [i]$. The type of each interval $[\sigma, \tau]$ in $\mathcal{I}_i$ is the same as the one given in Example 4.9. A convolution $\diamond$ is defined on the algebra of multiplicative functions by

$$(\mathfrak{g} \diamond \mathfrak{f})(\sigma, \pi) := \sum_{\substack{\sigma \leq \tau \leq \pi \\ \tau \in \mathcal{I}_i}} \mathfrak{g}(\sigma, \tau) \mathfrak{f}(\tau, \pi), \tag{6.10}$$

with $\mathfrak{f}$ and $\mathfrak{g}$ multiplicative functions. Then $\mathfrak{h} = \mathfrak{g} \diamond \mathfrak{f}$ is a multiplicative function and

$$h_i = \sum_{\tau \in \mathcal{I}_i} f_{\ell(\tau)} g_\tau.$$

The convolution (6.10) of two multiplicative functions corresponds to a composition umbra since $E[(\bar{\alpha}.\beta.\bar{\gamma})^i] = \sum_{\pi \in \mathcal{I}_i} a_{|\pi|} \, g_\pi$.

Remark 6.10. There is an interesting parallelism between the umbra having all classical cumulants equal to 1 and the umbra having all Boolean cumulants equal to 1. The Bell umbra β is the unique umbra (up to similarity) having $\{1\}_{i\geq 1}$ as sequence of classical cumulants. Its i-th moment is the i-th *Bell number*, that is, the number of partitions of an i-set. The umbra α satisfying $\bar{\alpha} \equiv (2\bar{u})_{D}$ is the unique umbra (up to similarity) having $\{1\}_{i\geq 1}$ as sequence of Boolean cumulants. The i-th moment of this umbra is 2^{i-1}, that is, the number of interval partitions $|\mathcal{I}_i|$.

Free cumulants are closely related to the Lagrange inversion formula. They were introduced by Voiculescu [104] as semi-invariants via the R-transform. Indeed, let us consider a non-commutative r.v. X, i.e., an element of a unital non-commutative algebra $\mathbb{A}$. Suppose $\phi : \mathbb{A} \to \mathbb{C}$ is a unital linear functional. The i-th moment of X is the complex number $m_i = \phi(X^i)$, while its generating function is the formal power series $M(z) = 1 + \sum_{i \geq 1} m_i z^i$. The *non-crossing* (or *free*) *cumulants of* X are the coefficients r_i of the ordinary power series $R(z) = 1 + \sum_{i \geq 1} r_i z^i$ which satisfies the identity

$$M(z) = R[zM(z)]. \tag{6.11}$$

If we set $M(z) = f(\bar{\alpha}, z)$ and $R(z) = f(\bar{\mathfrak{K}}_\alpha, z)$, then $f(\bar{\alpha}_D, z) = 1 + zM(z)$, and Equation (6.11) gives $\bar{\alpha} \equiv \bar{\mathfrak{K}}_\alpha \boldsymbol{.} \beta \boldsymbol{.} \bar{\alpha}_D$, which justifies the following definition.

Definition 6.11. The umbra $\bar{\mathfrak{K}}_\alpha$ satisfying $(-1 \boldsymbol{.} \bar{\mathfrak{K}}_\alpha)_D \equiv \bar{\alpha}_D^{<-1>}$ is called the *free cumulant umbra of* α.

Denote the auxiliary umbra $\left(\bar{\alpha}_D^{<-1>}\right)_P$ by $\mathfrak{L}_{\bar{\alpha}}$. We call this umbra *Lagrange involution of* α. Then, from Definition 6.11, we have

$$\bar{\mathfrak{K}}_\alpha \equiv -1 \boldsymbol{.} \mathfrak{L}_{\bar{\alpha}},$$

which is an alternative definition of the free cumulant umbra. The three main algebraic properties of cumulants can be stated as follows:

i) (*additivity*) $\bar{\mathfrak{K}}_\xi \equiv \bar{\mathfrak{K}}_\alpha \dot{+} \bar{\mathfrak{K}}_\gamma$ if and only if $-1 \boldsymbol{.} \mathfrak{L}_{\bar{\xi}} \equiv -1 \boldsymbol{.} \mathfrak{L}_{\bar{\alpha}} \dot{+} -1 \boldsymbol{.} \mathfrak{L}_{\bar{\gamma}}$;

ii) (*homogeneity*) $\bar{\mathfrak{K}}_{c\alpha} \equiv c \bar{\mathfrak{K}}_\alpha$, $c \in \mathbb{R}$;

iii) (*semi-invariance*) $\bar{\mathfrak{K}}_\xi \equiv \bar{\mathfrak{K}}_\alpha \dot{+} c\chi$ if and only if $-1 \boldsymbol{.} \mathfrak{L}_{\bar{\xi}} \equiv -1 \boldsymbol{.} \mathfrak{L}_{\bar{\alpha}} \dot{+} -1 \boldsymbol{.} \mathfrak{L}_{c\bar{u}}$.

From Definition 6.11 we have

$$\bar{\mathfrak{K}}_\alpha \equiv \bar{\alpha} \boldsymbol{.} \beta \boldsymbol{.} \bar{\alpha}_D^{<-1>} \quad \text{and} \quad \bar{\alpha} \equiv \bar{\mathfrak{K}}_\alpha \boldsymbol{.} \beta \boldsymbol{.} (-1 \boldsymbol{.} \bar{\mathfrak{K}}_\alpha)_D^{<-1>}.$$

If we compare these two equivalences with those corresponding to the α-cumulant umbra (6.9) and the α-Boolean cumulant umbra (6.8), respectively, the difference is quite obvious. Similarly, while the characterization of classical cumulants and Boolean cumulants is closely related to the composition of generating functions, this is not true for free cumulants. To be more precise, there does not exist a generating function $U(t)$ such that, if $R(t)$ is the free cumulant generating function of $M(t)$, then $M(t) = U[R(t) - 1]$, see also [32]. The same difference arises for convolutions of multiplicative functions on the lattice of non-crossing partitions, as the following remark shows.

Remark 6.12. A *non-crossing partition* $\pi = \{B_1, B_2, \ldots, B_s\}$ of the set $[i]$ is a partition such that, if $1 \leq h < l < k < m \leq i$, with $h, k \in B_j$, and $l, m \in B_{j'}$, then $j = j'$. Let $\mathcal{NC}_i$ denote the set of all non-crossing partitions of $[i]$. Since $\mathcal{NC}_i \subseteq \Pi_i$, we consider the induced subposet $(\mathcal{NC}_i, \leq)$ of $(\Pi_i, \leq)$. If $\pi \in \mathcal{NC}_i$, then the decomposition of the intervals $[\mathbf{0}_i, \pi] = \{\tau \in \mathcal{NC}_i \,|\, \tau \leq \pi\}$ is the same as the one given in Example 4.9, while the intervals $[\pi, \mathbf{1}_i] = \{\tau \in \mathcal{NC}_i \,|\, \pi \leq \tau\}$ have a quite different decomposition, which does not depend on $\ell(\pi)$ (see [89]). The convolution $*$ defined on the multiplicative functions is given by

$$(\mathfrak{g} * \mathfrak{f})(\sigma, \pi) := \sum_{\substack{\sigma \leq \tau \leq \pi \\ \tau \in \mathcal{NC}_i}} \mathfrak{g}(\sigma, \tau)\, \mathfrak{f}(\tau, \pi).$$

Hence, if $\mathfrak{h} = \mathfrak{g} * \mathfrak{f}$, then $h_i = \sum_{\tau \in \mathcal{NC}_i} g_\tau\, \mathfrak{f}(\tau, \mathbf{1}_i)$. Let us observe that the generating function of $f(\mathfrak{L}_{\bar{\alpha}}, z)$ is the *Fourier transform* $(\mathcal{F}\mathfrak{f})(z)$ of multiplicative functions [57]. Indeed, $\mathfrak{L}_{\bar{\omega}} \equiv \mathfrak{L}_{\bar{\gamma}} + \mathfrak{L}_{\bar{\alpha}}$ corresponds to $[\mathcal{F}(\mathfrak{g} * \mathfrak{f})](z) = (\mathcal{F}\mathfrak{g})(z)(\mathcal{F}\mathfrak{f})(z)$ for all unital multiplicative functions $\mathfrak{f}$ and $\mathfrak{g}$ on the lattice of non-crossing partitions.

Remark 6.13. The umbra having all free cumulants equal to 1 has a special meaning. Indeed, the umbra $\bar{\vartheta} \equiv \bar{u}\boldsymbol{.}(-1\boldsymbol{.}\bar{u})_D{}^*$ is the unique umbra satisfying $\bar{\mathfrak{K}}_{\bar{\vartheta}} \equiv \bar{u}$. Since $\vartheta^i \simeq \mathcal{C}_i$, where $\mathcal{C}_i$ denotes the i-th Catalan number, the umbra ϑ is called *Catalan umbra*. The parallelism with the Bell umbra relies on the fact that $\mathcal{C}_i = |\mathcal{NC}_i|$, so the moments of ϑ give the numbers of non-crossing partitions [90].

Example 6.14 (Wigner semicircle distribution)**.** In free probability, the Wigner semicircle r.v.[24] is analogous to the Gaussian r.v. in classical probability theory. Indeed, its free cumulants of degree higher than 2 are zero as for classical cumulants of a Gaussian r.v. By using Equivalence (4.6), it is straightforward to prove that the Wigner semicircle r.v. is represented by the umbra $\bar{\vartheta}\boldsymbol{.}\beta\boldsymbol{.}\bar{\eta}$, where ϑ is the Catalan umbra and η is the umbra introduced in Example 3.2.

Theorem 6.15 (Abel parametrization)**.** [32] *If $\bar{\mathfrak{K}}_{\bar{\alpha}}$ is the free cumulant umbra of α, then*

$$\bar{\alpha}^i \simeq \bar{\mathfrak{K}}_{\bar{\alpha}}(\bar{\mathfrak{K}}_{\bar{\alpha}} + i\boldsymbol{.}\bar{\mathfrak{K}}_{\bar{\alpha}})^{i-1} \quad \textit{and} \quad \bar{\mathfrak{K}}_{\bar{\alpha}}^i \simeq \bar{\alpha}(\bar{\alpha} - i\boldsymbol{.}\bar{\alpha})^{i-1}. \tag{6.12}$$

In terms of generalized Abel polynomials, Equivalences (6.12) satisfy $\bar{\alpha}^i \simeq \mathfrak{a}_i^{(i)}(\bar{\mathfrak{K}}_{\bar{\alpha}}, \bar{\mathfrak{K}}_{\bar{\alpha}})$ and $\bar{\mathfrak{K}}_{\bar{\alpha}}^i \simeq \mathfrak{a}_i^{(-i)}(\bar{\alpha}, \bar{\alpha})$.

Example 6.16. Complete Bell polynomials $a_i = \mathscr{Y}_i(c_1, c_2, \ldots, c_i)$ given in (4.4) return moments in terms of classical cumulants. For free cumulants, the same role is played by the volume polynomials introduced by

[24] The Wigner semicircle r.v. has probability density function given by a semicircle of radius R centered at $(0,0)$ and suitably normalized, that is, $f(x) = 2\sqrt{R^2 - x^2}/(\pi R^2)$ for $x \in [-R, R]$ and $f(x) = 0$ otherwise.

Pitman and Stanley [67]. Let us recall that the i-th volume polynomial $V_i(x)$ is the following homogeneous polynomial of degree i:

$$V_i(x) = \frac{1}{i!} \sum_{\boldsymbol{p} \in \mathrm{park}(i)} x_{\boldsymbol{p}}, \tag{6.13}$$

where $x_{\boldsymbol{p}} = x_{p_1} \cdots x_{p_i}$ and park(i) is the set of all *parking functions* of length i.[25] The volume polynomial (6.13) in the uncorrelated umbrae $\alpha_1, \ldots, \alpha_i$ similar to α satisfy $i!\, V_i(\alpha_1, \ldots, \alpha_i) \simeq \alpha(\alpha + i\boldsymbol{.}\alpha)^{i-1}$, see [63]. In particular, if $\{r_i\}$ denotes the sequence of free cumulants, then $a_i = i!\, V_i(r_1, \ldots, r_i)$, in parallel with (4.4).

6.2.2. *Abel-type cumulants.* A more general class of cumulants can be defined by using the generalized Abel polynomials. To the best of our knowledge, a previous attempt to give a unifying approach to cumulants was given in [2], but the Boolean case does not seem to fit in. Set $\mathfrak{a}_i^{(-m)}(\alpha) = \mathfrak{a}_i^{(-m)}(\alpha, \alpha)$.

Definition 6.17. An *Abel-type cumulant* $c_{i,m}(\alpha)$ of α is

$$c_{i,m}(\alpha) \simeq \mathfrak{a}_i^{(-m)}(\alpha) \simeq \alpha(\alpha - m\boldsymbol{.}\alpha)^{i-1}$$

if i and m are positive integers.

Let us consider the infinite matrix

$$C_\alpha = \begin{pmatrix} \mathfrak{a}_1^{(-1)}(\alpha) & \mathfrak{a}_1^{(-2)}(\alpha) & \ldots \\ \mathfrak{a}_2^{(-1)}(\alpha) & \mathfrak{a}_2^{(-2)}(\alpha) & \ldots \\ \vdots & \vdots & \\ \mathfrak{a}_n^{(-1)}(\alpha) & \mathfrak{a}_n^{(-2)}(\alpha) & \ldots \\ \vdots & \vdots & \end{pmatrix}. \tag{6.14}$$

Classical cumulants can be recovered from the first column, and Boolean cumulants from the second column. The sequence corresponding to the main diagonal of the matrix (6.14) picks out free cumulants.

Theorem 6.18 (Homogeneity property)**.**

$$\mathfrak{a}_i^{(-m)}(c\alpha) \simeq c^i \mathfrak{a}_i^{(-m)}(\alpha), \quad c \in \mathbb{R}.$$

To prove the semi-invariance property, a suitable normalization of $\mathfrak{a}_i^{(-m)}(\alpha)$ is necessary. For example, for Boolean and free cumulants, the normalization coefficient is equal to $\{i!\}$, umbrally represented by the Boolean unity $\bar{u}$. For classical cumulants, no normalization is needed.

Each sequence of cumulants in (6.14) — chosen along columns or along diagonals — linearizes a suitable convolution of umbrae (i.e.,

[25] A *parking function* of length i is a sequence $\boldsymbol{p} = (p_1, \ldots, p_i)$ of i positive integers, whose non-decreasing arrangement $\boldsymbol{p}^{\uparrow} = (p_{j_1}, \ldots, p_{j_i})$ satisfies $p_{j_k} \leq k$.

of moments). More precisely, it is possible to define the convolution $\alpha \overset{(m)}{+} \gamma$ induced by an integer m as

$$\mathfrak{a}_i^{(-m)}(\alpha \overset{(m)}{+} \gamma) \simeq \mathfrak{a}_i^{(-m)}(\alpha) + \mathfrak{a}_i^{(-m)}(\gamma). \tag{6.15}$$

These convolutions are commutative. It is possible to prove the existence of the convolution (6.15) induced by any integer m for any pair α, γ of umbrae [32]. It will be interesting to investigate if further sequences of cumulants known in the literature can be detected in the matrix (6.14), as for example the binary cumulants [95, 99], the tree-cumulants [108] and, more generally, the L-cumulants [109]; or if there are more sequences of cumulants that can be extracted from the matrix (6.14).

6.2.3. *An algorithm to compute cumulants from moments and viceversa.* A procedure to compute free cumulants from moments has been given in [11] by using nested derivatives of (6.11). The symbolic parametrization of cumulants via Abel polynomials suggests a completely different approach to perform such a computation [28]. The main advantage is that the resulting algorithm does not only allow us to compute free cumulants, but also classical and Boolean cumulants.

The algorithm relies on an efficient expansion of the umbral polynomial $\gamma(\gamma + \rho\boldsymbol{.}\gamma)^{i-1}$ for all positive integers i, with $\rho, \gamma \in \mathscr{A}$, that is,

$$\mathbb{E}\left[\gamma(\gamma + \rho\boldsymbol{.}\gamma)^{i-1}\right] = \sum_{\lambda \vdash i} \mathbb{E}[(\rho)_{\nu_\lambda - 1}]\, d_\lambda\, g_\lambda, \tag{6.16}$$

where $\{g_i\}$ is umbrally represented by the umbra γ. To compute the right-hand side of (6.16), we need the factorial moments of ρ. Recall that, if we just know the moments $\{\mathbb{E}[\rho^i]\}$, the factorial moments can be computed by using the well-known change of basis

$$\mathbb{E}[(\rho)_i] = \sum_{k=1}^{i} s(i,k)\mathbb{E}[\rho^k],$$

where $\{s(i,k)\}$ are the *Stirling numbers of the first kind.*

Expansion (6.16) can be fruitfully used to compute cumulants and moments as follows:

i) for classical cumulants in terms of moments: we use $\kappa_\alpha^i \simeq \alpha(\alpha - 1\boldsymbol{.}\alpha)^{i-1}$ in (4.2); then in Equation (6.16) choose $\rho = -1\boldsymbol{.}u$, with $\mathbb{E}[(-1\boldsymbol{.}u)_i] = (-1)_i = (-1)^i i!$, and $\gamma = \alpha$;

ii) for moments in terms of classical cumulants: we use $\alpha^i \simeq \kappa_\alpha(\kappa_\alpha + \alpha)^{i-1}$ in (4.2); then in Equation (6.16) choose $\rho = \beta$, with $\mathbb{E}[(\beta)_i] = 1$, and $\gamma = \kappa_\alpha$;

iii) for Boolean cumulants in terms of moments: we use $\bar{\eta}_\alpha^i \simeq \bar{\alpha}\,(\bar{\alpha} - 2.\bar{\alpha})^{i-1}$ in (6.7); then in Equation (6.16) choose $\rho = -2.u$, with $\mathbb{E}[(-2.u)_i] = (-1)^i (i+1)!$, and $\gamma = \bar{\alpha}$;

iv) for moments in terms of Boolean cumulants: we use $\bar{\alpha}^i \simeq \bar{\eta}_\alpha\,(\bar{\eta}_\alpha + 2.\bar{\alpha})^{i-1}$ in (6.7); then in Equation (6.16) choose $\rho = 2.\bar{u}.\beta$, with $\mathbb{E}[(2.\bar{u}.\beta)_i] = (i+1)!$, and $\gamma = \bar{\eta}_\alpha$;

v) for free cumulants in terms of moments: we use $\bar{\mathfrak{K}}_{\bar{\alpha}}^i \simeq \bar{\alpha}(\bar{\alpha} - i.\bar{\alpha})^{i-1}$ in (6.12); then in Equation (6.16) choose $\rho = -i.u$, with $\mathbb{E}[(-i.u)_j] = (-i)_j$, and $\gamma = \bar{\alpha}$;

vi) for moments in terms of free cumulants: we use $\bar{\alpha}^i \simeq \bar{\mathfrak{K}}_{\bar{\alpha}}(\bar{\mathfrak{K}}_{\bar{\alpha}} + i.\bar{\mathfrak{K}}_{\bar{\alpha}})^{i-1}$ in (6.12); then in Equation (6.16) choose $\rho = i.u$, with $\mathbb{E}[(i.u)_j] = (i)_j$, and $\gamma = \bar{\mathfrak{K}}_{\bar{\alpha}}$.

The symbolic algorithm allows us to compute also Abel-type cumulants by choosing $\rho = -m.u$, with $\mathbb{E}[(-m.u)_i] = (-m)_i$, and $\gamma = \alpha$ in (6.16).

7. Multivariate Faà di Bruno formula

The multivariate Faà di Bruno formula is a chain rule to compute higher derivatives of multivariable functions. This formula has been recently addressed by the following two approaches: combinatorial methods are used by Constantine and Savits [13], and (only in the bivariate case) by Noschese and Ricci [60]; analytical methods involving Taylor series are proposed by Leipnik and Pearce [52]. We refer to this last paper for a detailed list of references on this subject and for a detailed account of its applications. For statistical applications, a good account is the paper of Savits [81]. A comprehensive survey of univariate and multivariate series approximation in statistics is given in [49].

Computing the multivariate Faà di Bruno formula by means of a symbolic software can be done by recursively applying the chain rule. Despite its conceptual simplicity, applications of the chain rule become impractical already for small values, because the number of additive terms becomes enormous and their computation unwieldy. Moreover, the output is often untidy and further manipulations are required to simplify the result. Consequently, a "compressed" version of the multivariate Faà di Bruno formula is more attractive, as the number of variables or the order of derivatives increase. By using the symbolic method, the "compressed" version of the multivariate Faà di Bruno formula is nothing else but a suitable generalization of the well-known multinomial theorem:

$$(x_1 + \cdots + x_k)^i = \sum_{j_1 + \cdots + j_k = i} \binom{i}{j_1, \ldots, j_k} x_1^{j_1} \cdots x_k^{j_k}, \tag{7.1}$$

where the indeterminates are replaced by auxiliary umbrae. Suitable choices of these auxiliary umbrae give rise to an efficient computation of the following compositions: univariate with multivariate, multivariate with univariate, multivariate with the same multivariate, multivariate with different multivariates in an arbitrary number of components [25].

Various attempts to construct a multivariate generalization of the umbral calculus can be found in the literature, see, for example, [61] and [10].

Here, the starting point is the notion of multi-index. If, in the univariate case, the main device is to replace a_n by α^n via the linear functional $\mathbb{E}$, in the multivariate case, the main device is to replace sequences like $\{g_{i_1,\ldots,i_k}\}$, where $\boldsymbol{i} = (i_1, \ldots, i_k) \in \mathbb{N}_0^k$ is a *multi-index*, by a product of powers $\mu_1^{i_1} \cdots \mu_n^{i_k}$, where $\{\mu_1, \ldots, \mu_k\}$ are umbral monomials in $\mathbb{R}[\mathscr{A}]$ with not necessarily disjoint support. The $\boldsymbol{i}$-th power of $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$ is defined by $\boldsymbol{\mu}^{\boldsymbol{i}} = \mu_1^{i_1} \cdots \mu_k^{i_k}$.

Products of powers of umbral monomials have already been employed in Theorem 2.11 and in Definition 2.12, with umbrae indexed by multisets. Indeed there is a connection between the combinatorics of multisets and the Faà di Bruno formula, as highlighted in [45]. But, in facing the problem to set up an algorithm to efficiently compute the multivariate Faà di Bruno formula, multi-index notations have simplified formulae and procedures.

Definition 7.1. A sequence $\{g_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{N}_0^k} \in \mathbb{R}$, with $g_{\boldsymbol{i}} = g_{i_1,\ldots,i_k}$ and $g_{\boldsymbol{0}} = 1$, is umbrally represented by the k-tuple $\boldsymbol{\mu}$ if

$$\mathbb{E}[\boldsymbol{\mu}^{\boldsymbol{i}}] = g_{\boldsymbol{i}}, \qquad \text{for all } \boldsymbol{i} \in \mathbb{N}_0^k.$$

The elements $\{g_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{N}_0^k}$ are called *multivariate moments* of $\boldsymbol{\mu}$.

Definition 7.2. Two k-tuples $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ of umbral monomials are said to be *similar*, if they represent the same sequence of multivariate moments. In symbols: $\boldsymbol{\nu} \equiv \boldsymbol{\mu}$.

Definition 7.3. Two k-tuples $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ of umbral monomials are said to be *uncorrelated* if, for all $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{N}_0^k$, we have $E[\boldsymbol{\nu}^{\boldsymbol{i}} \boldsymbol{\mu}^{\boldsymbol{j}}] = E[\boldsymbol{\nu}^{\boldsymbol{i}}] E[\boldsymbol{\mu}^{\boldsymbol{j}}]$.

If the sequence $\{g_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{N}_0^k}$ is umbrally represented by the k-tuple $\boldsymbol{\mu}$, then the multivariate formal power series

$$f(\boldsymbol{\mu}, \boldsymbol{t}) = 1 + \sum_{\boldsymbol{i} \geq 1} g_{\boldsymbol{i}} \frac{\boldsymbol{t}^{\boldsymbol{i}}}{\boldsymbol{i}!}$$

is the multivariate moment generating function of $\boldsymbol{\mu}$.

Definition 7.4. A random vector $\boldsymbol{X}$ is said to be represented by a k-tuple $\boldsymbol{\mu}$ of umbral monomials, if its sequence of multivariate moments $\{g_{\boldsymbol{i}}\}$ is umbrally represented by $\boldsymbol{\mu}$.

Thus, a summation of m uncorrelated k-tuples similar to $\boldsymbol{\mu}$ represents a summation of m i.i.d. random vectors and can be denoted by $m.\boldsymbol{\mu}$. In order to compute its multivariate moments, we need the notion of *composition of a multi-index* and the notion of *partition of a multi-index*.

Definition 7.5 (Composition of a multi-index)**.** A composition $\boldsymbol{\lambda}$ of a multi-index $\boldsymbol{i}$. In symbols $\boldsymbol{\lambda} \models \boldsymbol{i}$, is a matrix $\boldsymbol{\lambda} = (\lambda_{ij})$ of nonnegative integers, with no zero columns, satisfying $\lambda_{r1} + \lambda_{r2} + \cdots = i_r$ for $r = 1, 2, \ldots, k$.

The number of columns of $\boldsymbol{\lambda}$ is called the *length* of $\boldsymbol{\lambda}$ and is denoted by $l(\boldsymbol{\lambda})$.

Definition 7.6 (Partition of a multi-index)**.** A partition of a multi-index $\boldsymbol{i}$ is a composition $\boldsymbol{\lambda}$, whose columns are in lexicographic order. In symbols $\boldsymbol{\lambda} \vdash \boldsymbol{i}$.

As for integer partitions, the notation $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1^{r_1}, \boldsymbol{\lambda}_2^{r_2}, \ldots)$ means that in the matrix $\boldsymbol{\lambda}$ there are r_1 columns equal to $\boldsymbol{\lambda}_1$, r_2 columns equal to $\boldsymbol{\lambda}_2$, and so on, with $\boldsymbol{\lambda}_1 < \boldsymbol{\lambda}_2 < \cdots$ in lexicographic order. The integer r_i is the multiplicity of $\boldsymbol{\lambda}_i$. We set $\mathfrak{m}(\boldsymbol{\lambda}) = (r_1, r_2, \ldots)$.

Proposition 7.7. [23] *If* $\boldsymbol{i} \in \mathbb{N}_0^k$, *then*

$$\mathbb{E}[(m.\boldsymbol{\mu})^{\boldsymbol{i}}] = \sum_{\boldsymbol{\lambda} \vdash \boldsymbol{i}} \frac{\boldsymbol{i}!}{\mathfrak{m}(\boldsymbol{\lambda})!\, \boldsymbol{\lambda}!} (m)_{l(\boldsymbol{\lambda})} E[\boldsymbol{\mu}^{\boldsymbol{\lambda}_1}]^{r_1} E[\boldsymbol{\mu}^{\boldsymbol{\lambda}_2}]^{r_2} \cdots,$$

where the sum is over all partitions $\boldsymbol{\lambda}$ *of the multi-index* $\boldsymbol{i}$.

Multivariate moments of the auxiliary umbra $(\alpha.\beta.\boldsymbol{\mu})^{\boldsymbol{i}}$ can be computed from Proposition 7.7:

$$\mathbb{E}[(\alpha.\beta.\boldsymbol{\mu})^{\boldsymbol{i}}] = \sum_{\boldsymbol{\lambda} \vdash \boldsymbol{i}} \frac{\boldsymbol{i}!}{m(\boldsymbol{\lambda})!\, \boldsymbol{\lambda}!} a_{l(\boldsymbol{\lambda})} E[\boldsymbol{\mu}^{\boldsymbol{\lambda}_1}]^{r_1} E[\boldsymbol{\mu}^{\boldsymbol{\lambda}_2}]^{r_2} \cdots. \tag{7.2}$$

To obtain Equation (7.2), the arguments are the same as the ones employed to get Equation (4.6). The umbra $\alpha.\beta.\boldsymbol{\mu}$ represents a multivariate randomized compound Poisson r.v., that is, a random sum $S_N = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_N$ of i.i.d. random vectors $\{\boldsymbol{X}_i\}$ indexed by a Poisson r.v. N. The right-hand side of Equivalence (7.2) gives moments of S_N if we replace the $\boldsymbol{i}$-th multivariate moment of $\boldsymbol{\mu}$ by the $\boldsymbol{i}$-th multivariate moment of $\{\boldsymbol{X}_i\}$ and the moments of α by those of N. This result is the same as Theorem 4.1 of [13], obtained with a different proof and different methods.

A way to compute partitions of a multi-index $\boldsymbol{i}$ is by employing subdivisions of the multiset M, when $\boldsymbol{i}$ is its vector of multiplicities.

Example 7.8. The multiset $M = \{\mu_1, \mu_1, \mu_2\}$ corresponds to the multi-index $(2, 1)$, and the subdivisions

$$\{\{\mu_1, \mu_1, \mu_2\}\}, \{\{\mu_1, \mu_1\}, \{\mu_2\}\}, \{\{\mu_1, \mu_2\}, \{\mu_1\}\}, \{\{\mu_1\}, \{\mu_1\}, \{\mu_2\}\}$$

correspond to the multi-index partitions

$$\binom{2}{1}, \binom{2,0}{0,1}, \binom{1,1}{1,0}, \binom{1,1,0}{0,0,1}.$$

By using the auxiliary umbra $\alpha \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\mu}$, we can summarize the different types of composition as follows:

a) (*univariate composed with multivariate*) The composition of a univariate formal power series $f(\alpha, z)$ with a multivariate formal power series $f(\boldsymbol{\mu}, \boldsymbol{z})$ is defined by

$$f(\alpha \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\mu}, \boldsymbol{z}) \;=\; f[\alpha, f(\boldsymbol{\mu}, \boldsymbol{z}) - 1].$$

Equivalence (7.2) gives its $\boldsymbol{i}$-th coefficient.

b) (*multivariate composed with univariate*) The composition of a multivariate formal power series $f(\boldsymbol{\mu}, \boldsymbol{z})$ with a univariate formal power series $f(\gamma, z)$ is defined by

$$f[(\mu_1 + \cdots + \mu_k) \boldsymbol{.} \beta \boldsymbol{.} \gamma, z] \;=\; f[\boldsymbol{\mu}, (f(\gamma, z) - 1, \ldots, f(\gamma, z) - 1)].$$

The i-th coefficient of $f[(\mu_1 + \cdots + \mu_k) \boldsymbol{.} \beta \boldsymbol{.} \gamma, z]$ can be computed by replacing γ by $(\mu_1 + \cdots + \mu_k)$ in (4.6) and α by γ.

c) (*multivariate composed with multivariate*) The composition of a multivariate formal power series $f(\boldsymbol{\mu}, \boldsymbol{z})$ with a multivariate formal power series $f(\boldsymbol{\nu}, \boldsymbol{z})$ is defined by

$$f[(\mu_1 + \cdots + \mu_k) \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}, \boldsymbol{z}] \;=\; f[\boldsymbol{\mu}, (f(\boldsymbol{\nu}, \boldsymbol{z}) - 1, \ldots, f(\boldsymbol{\nu}, \boldsymbol{z}) - 1)].$$

The $\boldsymbol{i}$-th coefficient of $f[(\mu_1 + \cdots + \mu_k) \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}, \boldsymbol{z}]$ can be computed by evaluating Equivalence (7.2) after having replaced $\boldsymbol{\mu}$ by $\boldsymbol{\nu}$ and the umbra α by $(\mu_1 + \cdots + \mu_k)$.

d) (*multivariate composed with multivariate*) The composition of a multivariate formal power series $f(\boldsymbol{\mu}, \boldsymbol{z})$ with a multivariate formal power series $f(\boldsymbol{\nu}, \boldsymbol{z}_{(m)})$, with a different number of indeterminates, is defined by

$$f[(\mu_1 + \cdots + \mu_k) \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}, \boldsymbol{z}_{(m)}] \;=\; f[\boldsymbol{\mu}, (f(\boldsymbol{\nu}, \boldsymbol{z}_{(m)}) - 1, \ldots, f(\boldsymbol{\nu}, \boldsymbol{z}_{(m)}) - 1)].$$

e) (*multivariate composed with different multivariates*) The composition of a multivariate formal power series $f(\boldsymbol{\mu}, \boldsymbol{z})$ with different multivariate formal power series $f(\boldsymbol{\nu}_i, \boldsymbol{z}_{(m)})$, with a different number of indeterminates, is defined by

$$\begin{aligned} (7.3) \quad & f[\mu_1 \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}_1 + \cdots + \mu_k \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}_k, \boldsymbol{z}_{(m)}] \\ & \qquad = f[\boldsymbol{\mu}, (f(\boldsymbol{\nu}_1, \boldsymbol{z}_{(m)}) - 1, \ldots, f(\boldsymbol{\nu}_k, \boldsymbol{z}_{(m)}) - 1)]. \end{aligned}$$

The $\boldsymbol{i}$-th coefficient of the compositions *d)* and *e)* can be obtained by using *generalized Bell polynomials.*

Definition 7.9. The *generalized Bell polynomial* of order $\boldsymbol{i}$ is

$$B_{\boldsymbol{i}}^{(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_k)}(x_1, \ldots, x_k) = (x_1 \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}_1 + \cdots + x_k \boldsymbol{.} \beta \boldsymbol{.} \boldsymbol{\nu}_k)^{\boldsymbol{i}}.$$

The formal power series whose coefficients are generalized Bell polynomials is $\exp\left\{\sum_{i=1}^k x_i[f(\boldsymbol{\nu}_i, \boldsymbol{z}_{(m)}) - 1]\right\}$. These polynomials give the multivariate Faà di Bruno formula for all cases listed before:

a) (*univariate composed with multivariate*) $B_{\boldsymbol{i}}^{\boldsymbol{\nu}}(\alpha) = (\alpha\boldsymbol{.}\beta\boldsymbol{.\mu})^{\boldsymbol{i}}$;

b) (*multivariate composed with univariate*)
$$B_{\boldsymbol{i}}^{\gamma}(\mu_1 + \cdots + \mu_k) = [(\mu_1 + \cdots + \mu_k)\boldsymbol{.}\beta\boldsymbol{.}\gamma]^{\boldsymbol{i}};$$

c);d) (*multivariate composed with multivariate*)
$$B_{\boldsymbol{i}}^{\boldsymbol{\nu}}(\mu_1 + \cdots + \mu_k) = [(\mu_1 + \cdots + \mu_k)\boldsymbol{.}\beta\boldsymbol{.\nu}]^{\boldsymbol{i}};$$

e) (*multivariate composed with different multivariates*)
$$B_{\boldsymbol{i}}^{(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_k)}(\mu_1, \ldots, \mu_k) \simeq (\mu_1\boldsymbol{.}\beta\boldsymbol{.\nu}_1 + \cdots + \mu_k\boldsymbol{.}\beta\boldsymbol{.\nu}_k)^{\boldsymbol{i}}.$$

A `MAPLE` algorithm for computing the evaluation of $B_{\boldsymbol{i}}^{(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_k)}(\mu_1, \ldots, \mu_k)$ is given in [25]. The main idea is to apply the multivariate version of (7.1) to get

$$\begin{aligned}(7.4)\quad & \mathbb{E}\left[B_{\boldsymbol{i}}^{(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_k)}(\mu_1, \ldots, \mu_k)\right] \\ &= \sum_{(\boldsymbol{i}_1,\ldots,\boldsymbol{i}_k):\sum_{j=1}^k \boldsymbol{i}_j = \boldsymbol{i}} \binom{\boldsymbol{k}}{\boldsymbol{i}_1, \ldots, \boldsymbol{i}_k} \mathbb{E}\left[(\mu_1.\beta.\boldsymbol{\nu}_1)^{\boldsymbol{i}_1} \cdots (\mu_k.\beta.\boldsymbol{\nu}_k)^{\boldsymbol{i}_k}\right],\end{aligned}$$

to expand powers $(\mu_i\boldsymbol{.}\beta\boldsymbol{.\nu}_i)^{k_i}$, to multiply the resulting expansions by $(\mu_1.\beta.\boldsymbol{\nu}_1)^{\boldsymbol{i}_1} \cdots (\mu_d.\beta.\boldsymbol{\nu}_d)^{\boldsymbol{i}_d}$, and then to apply evaluation lowering powers to indexes.

7.1. **Multivariate cumulants.** Given the sequence $\{g_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathbb{N}_0^n}$, its *multivariate cumulants* are the coefficients of the formal power series in the left-hand side of the following equation
$$1 + \sum_{\boldsymbol{i}\geq 1} g_{\boldsymbol{i}} \frac{\boldsymbol{t}^{\boldsymbol{i}}}{\boldsymbol{i}!} = \exp\left(\sum_{\boldsymbol{j}\geq 1} c_{\boldsymbol{j}} \frac{\boldsymbol{t}^{\boldsymbol{j}}}{\boldsymbol{i}!}\right).$$

Proposition 7.10. *The auxiliary umbra $\chi\boldsymbol{.\mu}$ represents the multivariate cumulants of $\boldsymbol{\mu}$.*

Proof. The result follows from *a)*, by choosing as umbra α the umbra $\chi\boldsymbol{.}\chi$, and observing that $\chi\boldsymbol{.\mu} \equiv (\chi\boldsymbol{.}\chi)\boldsymbol{.}\beta\boldsymbol{.\mu}$ and
$$f(\chi\boldsymbol{.}\chi, t) = 1 + \log(t + 1). \qquad \square$$

The multivariate analog of the well-known semi-invariance property of cumulants is $\chi\boldsymbol{.}(\boldsymbol{\mu}+\boldsymbol{\nu}) \equiv \chi\boldsymbol{.\mu} \dotplus \chi\boldsymbol{.\nu}$, where $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are uncorrelated n-tuples of umbral monomials, and the *disjoint sum* $\boldsymbol{\mu} \dotplus \boldsymbol{\nu}$ is defined by the multivariate moments $\mathbb{E}[(\boldsymbol{\mu} \dotplus \boldsymbol{\nu})^{\boldsymbol{i}}] = \mathbb{E}[\boldsymbol{\mu}^{\boldsymbol{i}}] + \mathbb{E}[\boldsymbol{\nu}^{\boldsymbol{i}}]$. Thus, multivariate cumulants linearize convolutions of uncorrelated n-tuples of umbral monomials. By using Equation (7.2) with α replaced by $\chi\boldsymbol{.}\chi$, that is, by expanding the $\boldsymbol{i}$ multivariate moments of $(\chi\boldsymbol{.}\chi)\boldsymbol{.}\beta\boldsymbol{.\mu} \equiv \chi\boldsymbol{.\mu}$, we get multivariate cumulants in terms of multivariate moments. Still using Equation (7.2) with α replaced by u and $\boldsymbol{\mu}$ replaced by $\chi.\boldsymbol{\mu}$, that is, expanding the $\boldsymbol{i}$-th multivariate moment of $u{\boldsymbol{.}}\beta{\boldsymbol{.}}(\chi.\boldsymbol{\mu}) \equiv \boldsymbol{\mu}$, we get multivariate moments in terms of multivariate cumulants. The same equations are given in [54] using tensor notations. Umbrae indexed by multisets are employed in [21].

7.2. **Multivariate Lévy processes.** Let $\boldsymbol{X} = \{\boldsymbol{X}_t\}$ be a Lévy process on $\mathbb{R}^k$, with $\boldsymbol{X}_t = \left(X_1^{(t)}, \ldots, X_k^{(t)}\right)$, where $\left\{X_j^{(t)}\right\}_{j=1}^{k}$ are univariate r.v.'s all defined on the same probability space for all $t \geq 0$.

In Section 3, we have shown that every Lévy process $X = \{X_t\}$ on $\mathbb{R}$ is umbrally represented by the family of auxiliary umbrae $\{t{\boldsymbol{.}}\alpha\}_{t\geq 0}$ defined by $\mathbb{E}\left[(t{\boldsymbol{.}}\alpha)^i\right] = E\left[X_t^i\right]$, for all non-negative integers i. This result can be generalized to the multivariate case as follows.

Definition 7.11. A Lévy process $\boldsymbol{X} = \{\boldsymbol{X}_t\}_{t\geq 0}$ on $\mathbb{R}^k$ is umbrally represented by $\{t{\boldsymbol{.}}\boldsymbol{\mu}\}_{t\geq 0}$, if $\boldsymbol{\mu}$ is a k-tuple of umbral monomials representing the k-tuple of random vectors $\boldsymbol{X}_1 = (X_1^{(1)}, \ldots, X_k^{(1)})$ with the property that

$$E[\boldsymbol{\mu}^{\boldsymbol{i}}] = E\left[\left\{X_1^{(1)}\right\}^{i_1} \cdots \left\{X_k^{(1)}\right\}^{i_k}\right],$$

for all $\boldsymbol{i} = (i_1, \ldots, i_k) \in \mathbf{N}_0^k$.

The following theorem gives the multivariate analog of the Lévy–Khintchine formula (4.9).

Theorem 7.12 (Multivariate Lévy–Khintchine formula)**.** [80] *If $\{\boldsymbol{X}_t\}$ is a multivariate Lévy process, then $E\left[e^{\boldsymbol{X}_t \boldsymbol{z}'}\right] = [\phi_{\boldsymbol{X}_1}(\boldsymbol{z})]^t$ with*[26]

$$\begin{aligned}\phi_{\boldsymbol{X}_1}(\boldsymbol{z}) = \exp\Bigg\{\Bigg[&\frac{1}{2}\boldsymbol{z}\Sigma\boldsymbol{z}' + \boldsymbol{m}\boldsymbol{z}' \\ &+ \int_{\mathbb{R}^k}\left(e^{\boldsymbol{z}\boldsymbol{x}'} - 1 - \boldsymbol{z}\,\boldsymbol{x}'\mathbf{1}_{\{|\boldsymbol{x}|\leq 1\}}(\boldsymbol{x})\right)\nu(d\boldsymbol{x})\Bigg]\Bigg\},\end{aligned} \tag{7.5}$$

where $\boldsymbol{m} \in \mathbb{R}^k$ and Σ is a symmetric, positive-definite $k \times k$ matrix.

The triple $(\boldsymbol{m}, \Sigma, \nu)$ is called Lévy triple, Σ is called *covariance matrix*, and ν is the *Lévy measure* on $\mathbb{R}^k$. As in the univariate case, if we set $\boldsymbol{m}_2 = \int_{\mathbb{R}^k} \boldsymbol{x}\,\mathbf{1}_{\{|\boldsymbol{x}|>1\}}(\boldsymbol{x})\,\nu(d\boldsymbol{x})$ and $\boldsymbol{m}_1 = \boldsymbol{m} + \boldsymbol{m}_2$ in Equation (7.5), then

$$\phi_{\boldsymbol{X}_1}(\boldsymbol{z}) = \exp\left\{\left[\frac{1}{2}\boldsymbol{z}\Sigma\boldsymbol{z}' + \boldsymbol{m}_1\boldsymbol{z}' + \int_{\mathbb{R}^k}(e^{\boldsymbol{z}\boldsymbol{x}'} - 1 - \boldsymbol{z}\boldsymbol{x}')\nu(d\boldsymbol{x})\right]\right\}. \tag{7.6}$$

Due to the above equation, a multivariate Lévy process is represented by a family of auxiliary umbrae, as given in the following theorem.

[26]Here, the notation $\boldsymbol{z}'$ denotes the transpose of the row vector $\boldsymbol{z}$.

Theorem 7.13. *A multivariate Lévy process* $\{\boldsymbol{X}_t\}_{t\geq 0}$ *is represented by the family of auxiliary umbrae*

$$\left\{t.\beta.\left(\boldsymbol{\eta}C' \dot{+} \chi.\boldsymbol{m}_1 \dot{+} \boldsymbol{\gamma}\right)\right\}_{t\geq 0}, \tag{7.7}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$ *is a* k*-tuple of uncorrelated umbrae satisfying* $f(\boldsymbol{\eta}, \boldsymbol{z}) = 1 + \boldsymbol{z}\boldsymbol{z}'/2$, *the matrix* C *is the square root of the covariance matrix* Σ *in Equation* (7.6), *and* $\boldsymbol{\gamma}$ *is a* k*-tuple of umbral monomials satisfying*

$$f(\boldsymbol{\gamma}, \boldsymbol{z}) = 1 + \int_{\mathbb{R}^k} \left(e^{\boldsymbol{z}\boldsymbol{x}'} - 1 - \boldsymbol{z}\boldsymbol{x}'\right) \nu(d\boldsymbol{x}).$$

As observed previously, every auxiliary umbra, such as $t.\beta.\boldsymbol{\nu}$, is the symbolic version of a multivariate compound Poisson process of parameter t. Therefore, for t fixed, the symbolic representation (7.7) of a multivariate Lévy process is a multivariate compound Poisson r.v. of parameter t, and the k-tuple $(\boldsymbol{\eta}C' \dot{+} \chi.\boldsymbol{m}_1 \dot{+} \boldsymbol{\gamma})$ umbrally represents the multivariate cumulants of $\boldsymbol{X}_1$.

The auxiliary umbra $\chi.\boldsymbol{m}_1$ does not have a probabilistic counterpart. If $\boldsymbol{m}$ is not equal to the zero vector, this parallels the well-known difficulty to interpret the Lévy measure as a probability measure in Subsection 4.2.

7.3. **Multivariate Hermite polynomials.** In Equation (7.6), assume the Lévy measure ν to be zero. Then the moment generating function of $\{\boldsymbol{X}_t\}_{t\geq 0}$ is $E\left[e^{\boldsymbol{X}_t\boldsymbol{z}'}\right] = [\phi_{\boldsymbol{X}_1}(\boldsymbol{z})]^t$ with

$$\phi_{\boldsymbol{X}_1}(\boldsymbol{z}) = \exp\left\{\left(\frac{1}{2}\boldsymbol{z}\Sigma\boldsymbol{z}' + \boldsymbol{m}_1\boldsymbol{z}'\right)\right\}. \tag{7.8}$$

Theorem 7.14 (Wiener process)**.** [80] *The stochastic process* $\{\boldsymbol{X}_t\}_{t\geq 0}$ *having moment generating function* (7.8) *is a multivariate Wiener process, that is,* $\boldsymbol{X}_t = \boldsymbol{m}_1 t + C\boldsymbol{B}_t$, *where* C *is a* $k \times k$ *matrix whose determinant is non-zero satisfying* $\Sigma = CC'$, *and* $\{\boldsymbol{B}_t\}_{t\geq 0}$ *is the multivariate standard Brownian motion in* $\mathbb{R}^k$.

When $\boldsymbol{m}_1 = \boldsymbol{0}$, the Wiener process reduces to a multivariate non-standard Brownian motion. From Theorem 7.14, we see that the Wiener process $\boldsymbol{X}_t$ is a special multivariate Lévy process, and the following result is a corollary of Theorem 7.13.

Corollary 7.15. *The Wiener process* $\{\boldsymbol{X}_t\}_{t\geq 0}$ *is represented by the family of auxiliary umbrae*

$$\{t.\beta.(\boldsymbol{\eta}C' \dot{+} \chi.\boldsymbol{m}_1)\}_{t\geq 0}.$$

There is a special family of multivariate polynomials which is closely related to the multivariate non-standard Brownian motion: the (multivariate) *Hermite polynomials.* The $\boldsymbol{i}$-th Hermite polynomial $H_{\boldsymbol{i}}(\boldsymbol{x}, \Sigma)$

is defined by

$$H_{\boldsymbol{i}}(\boldsymbol{x},\Sigma) = (-1)^{|\boldsymbol{i}|}\frac{D_{\boldsymbol{x}}^{(\boldsymbol{i})}\phi(\boldsymbol{x};\boldsymbol{0},\Sigma)}{\phi(\boldsymbol{x};\boldsymbol{0},\Sigma)},$$

where $\phi(\boldsymbol{x};\boldsymbol{0},\Sigma)$ denotes the multivariate Gaussian density with mean $\boldsymbol{0}$ and covariance matrix Σ of full rank k. The polynomials $H_{\boldsymbol{i}}(\boldsymbol{x}\Sigma^{-1},\Sigma^{-1})$ are orthogonal with respect to $\phi(\boldsymbol{x};\boldsymbol{0},\Sigma)$, where Σ^{-1} is the inverse matrix of Σ, and is denoted by $\tilde{H}_{\boldsymbol{i}}(\boldsymbol{x},\Sigma)$.

Theorem 7.16. [107] *If $\boldsymbol{Z} \sim N(\boldsymbol{0},\Sigma)$ and $\boldsymbol{Y} \sim N(\boldsymbol{0},\Sigma^{-1})$ are multivariate Gaussian vectors with mean $\boldsymbol{0}$ and covariance matrix Σ and Σ^{-1}, respectively, then, for all $\boldsymbol{i} \in \mathbb{N}_0^k$, we have*

$$H_{\boldsymbol{i}}(\boldsymbol{x},\Sigma) = E[(\boldsymbol{x}\Sigma^{-1} + \mathrm{i}\boldsymbol{Y})^{\boldsymbol{i}}], \text{ and } \tilde{H}_{\boldsymbol{i}}(\boldsymbol{x},\Sigma) = E[(\boldsymbol{x} + \mathrm{i}\boldsymbol{Z})^{\boldsymbol{i}}], \tag{7.9}$$

where i *is the imaginary unit.*

In the literature, Equations (7.9) are called *moment representation* of the Hermite polynomials. Quite recently, many authors have obtained moment representations for various families of polynomials, especially in the univariate case, see references in [31]. The multivariate symbolic method allows us to widen the field of applicability of this moment representation. For the multivariate Hermite polynomials, we have

$$H_{\boldsymbol{i}}(\boldsymbol{x},\Sigma) = \mathbb{E}\left[(-1\boldsymbol{.}\beta\boldsymbol{.}\boldsymbol{\nu} + \boldsymbol{x}\Sigma^{-1})^{\boldsymbol{i}}\right], \quad \text{and } \tilde{H}_{\boldsymbol{i}}(\boldsymbol{x},\Sigma) = \mathbb{E}\left[(-1\boldsymbol{.}\beta\boldsymbol{.}\boldsymbol{\mu} + \boldsymbol{x})^{\boldsymbol{i}}\right], \tag{7.10}$$

with $\boldsymbol{\nu}$ a k-tuple of umbral monomials satisfying $f(\boldsymbol{\nu},\boldsymbol{t}) = 1 + \frac{1}{2}\boldsymbol{t}\Sigma^{-1}\boldsymbol{t}'$ and $\boldsymbol{\mu}$ a k-tuple of umbral monomials satisfying $f(\boldsymbol{\mu},\boldsymbol{t}) = 1 + \frac{1}{2}\boldsymbol{t}\Sigma\boldsymbol{t}'$. In Section 5, we have shown that Appell polynomials are umbrally represented by the polynomial umbra $x\boldsymbol{.}u + \alpha$. It is well known that univariate Hermite polynomials are Appell polynomials. Equations (7.10) then show that also multivariate Hermite polynomials are of Appell type. Let us stress that the umbra $-1\boldsymbol{.}\beta$ enables us to obtain a simple expression for multivariate Hermite polynomials, without the employment of the imaginary unit as done in Equations (7.9). Moreover, since the moments of $\boldsymbol{Z} \sim N(\boldsymbol{0},\Sigma)$ are umbrally represented by the umbra $\beta\boldsymbol{.}\boldsymbol{\mu}$, the k-tuple $\boldsymbol{\mu} \equiv \chi\boldsymbol{.}(\beta\boldsymbol{.}\boldsymbol{\mu})$ in $\tilde{H}_{\boldsymbol{i}}(\boldsymbol{x},\Sigma)$ umbrally represents the multivariate cumulants of $\boldsymbol{Z}$.

Proposition 7.17. $\mathbb{E}\left\{\tilde{H}_{\boldsymbol{i}}[t\boldsymbol{.}\beta(\boldsymbol{\eta}C'),\,\Sigma]\right\} = 0$ *for all* $t \geq 0$.

Proof. The result follows by observing that the k-tuple $\boldsymbol{\mu}$ in (7.10) satisfies $\boldsymbol{\mu} \equiv \boldsymbol{\eta}C'$ and

$$\tilde{H}_{\boldsymbol{i}}[t\boldsymbol{.}\beta(\boldsymbol{\eta}C'),\,\Sigma] = \mathbb{E}\left\{[-t\boldsymbol{.}\beta\boldsymbol{.}\boldsymbol{\mu} + t\boldsymbol{.}\beta(\boldsymbol{\eta}C')]^{\boldsymbol{i}}\right\}. \qquad \square$$

For the family of polynomials $\{\tilde{H}_{\boldsymbol{i}}^{(t)}(\boldsymbol{x},\,\Sigma)\}_{t\geq 0}$, Proposition 7.17 parallels the main property of time-space harmonic polynomials introduced in Section 3.1. Since, in the literature, to the best of our knowledge, a theory of multivariate time-space harmonic polynomials has not yet been given, we believe that the setting introduced here could be a fruitful way to build this theory with respect to multivariate Lévy processes.

More properties on multivariate Hermite polynomials can be found in [25]. We end this subsection by recalling that multivariate Hermite polynomials are special generalized Bell polynomials.

Proposition 7.18. *If $\boldsymbol{\nu}_{\boldsymbol{x}}$ is a k-tuple of umbral polynomials satisfying $f(\boldsymbol{\nu}_{\boldsymbol{x}}, \boldsymbol{t}) = 1 + f(\boldsymbol{\nu}, \boldsymbol{t}+\boldsymbol{x}) - f(\boldsymbol{\nu}, \boldsymbol{t})$, with $f(\boldsymbol{\nu}, \boldsymbol{t}) = 1 + \frac{1}{2}\boldsymbol{t}\Sigma^{-1}\boldsymbol{t}'$, then*

$$H_{\boldsymbol{i}}(\boldsymbol{x}, \Sigma) = (-1)^{|\boldsymbol{i}|}\mathbb{E}\left[(-1.\beta.\boldsymbol{\nu}_{\boldsymbol{x}})^{\boldsymbol{i}}\right].$$

Finally we remark that the efficient computation of the multivariate Hermite polynomials as generalized Bell polynomials can help us in constructing an efficient algorithm for a multivariate Edgeworth approximation of multivariate density functions [49].

7.4. **Multivariate Bernoulli polynomials.** By the symbolic method, it is possible to introduce multivariate Bernoulli polynomials as powers of polynomials whose coefficients involve multivariate Lévy processes. According to [31], this means to give a moment representation for these polynomials.

Definition 7.19. The *multivariate Bernoulli umbra* $\boldsymbol{\iota}$ is the k-tuple $(\iota, \ldots, \iota)$, with ι the Bernoulli umbra.

Definition 7.20. If $\boldsymbol{v} \in \mathbb{N}_0^k$, the *multivariate Bernoulli polynomial* of order $\boldsymbol{v}$ is $B_{\boldsymbol{v}}^{(t)}(\boldsymbol{x}) = \mathbb{E}[(\boldsymbol{x} + t.\boldsymbol{\iota})^{\boldsymbol{v}}]$, where $\boldsymbol{\iota}$ is the multivariate Bernoulli umbra and $t \geq 0$.

As corollary, the multivariate Bernoulli polynomials are

$$B_{\boldsymbol{v}}^{(t)}(\boldsymbol{x}) = \sum_{\boldsymbol{j} \leq \boldsymbol{v}} \binom{\boldsymbol{v}}{\boldsymbol{j}} \boldsymbol{x}^{\boldsymbol{j}}\, \mathbb{E}[(t.\boldsymbol{\iota})^{\boldsymbol{v}-\boldsymbol{j}}].$$

The symbolic representation of the coefficients of $B_{\boldsymbol{v}}^{(t)}(\boldsymbol{x})$ by $\mathbb{E}[(t.\boldsymbol{\iota})^{\boldsymbol{v}}]$ simplifies the calculus and is computationally efficient, see [31].

Proposition 7.21. [31] $\mathcal{B}_{\boldsymbol{v}}^{(t)}(-t.\boldsymbol{\iota}) = \mathcal{B}_{\boldsymbol{v}}^{(t)}[t.(-1.\boldsymbol{\iota})] = 0.$

Let us observe that, when in $B_{\boldsymbol{v}}^{(t)}(\boldsymbol{x})$ we replace $\boldsymbol{x}$ by the auxiliary umbra $-t.\boldsymbol{\mu}$, this means to replace $\boldsymbol{x}^{\boldsymbol{j}}$ by $\mathbb{E}[(-t.\boldsymbol{\mu})^{\boldsymbol{j}}]$. The auxiliary umbra $-t.\boldsymbol{\iota}$ is the symbolic version of a multivariate Lévy process. Indeed, $-1.\boldsymbol{\iota}$ is the umbral counterpart of a k-tuple identically distributed with $(U, \ldots, U)$, where U is a uniform r.v. on the interval $(0, 1)$, and corresponds to the random vector $\mathbf{X}_1 = (X_1^{(1)}, \ldots, X_k^{(1)})$ given in Definition 7.11. By generalizing the definition of conditional evaluation given in Section 3 to the multivariate case, the multivariate Bernoulli polynomials should become time-space harmonic with respect to the multivariate Lévy processes $\{-t.\boldsymbol{\iota}\}_{t\geq 0}$. Indeed, Proposition 7.21 shows that these polynomials share one of the main properties of time-space harmonic polynomials: when the vector of indeterminates is replaced by the corresponding Lévy process, their overall mean is zero. Similar considerations apply for multivariate Euler polynomials, see [31] for details and further connections between these two families of polynomials.

8. Appendix 1: Tables of computation times

In Table 3, computation times are given for computing the mean of statistics such as the second expression in (2.13). The first column refers to the `Mathematica` routine proposed in [105]. The second column refers to a `Maple` routine relying on the symbolic method of moments.

Table 3. Comparison of computation times.

$[1^i 2^j 3^k \cdots]$	SIP	MAPLE
$[5^3\, 9\, 10][1\,2\,3\,4\,5]$	0.7	0.1
$[5^3\, 8\, 9\, 10][1\,2\,3\,4\,5]$	5.6	0.4
$[6\,7\,8\,9\,10][1\,2\,3\,4\,5]$	2.2	0.1
$[6\,7\,8\,9\,10][1\,2][3\,4\,5]$	3.1	0.4
$[6\,7][8\,9\,10][1\,2][3\,4\,5]$	4.7	1.3
$[5\,6\,7\,8\,9\,10][1\,2\,3\,4\,5]$	16.7	0.3
$[5\,6\,7\,8\,9\,10][1\,2\,3\,4\,5\,6]$	348.7	1.5
$[6\,7\,8\,9\,10][6\,7][3\,4\,5][1\,2]$	125.6	16.4

Table 4 lists computation times (in seconds) observed by using the polynomial expansion in (6.16) and Bryc's procedure [11], both implemented in `MAPLE`, release 7, for the conversion from free cumulants to moments.[27]

Table 4. Comparison of computation times needed to compute free cumulants in terms of moments. Tasks performed on Intel (R) Pentium (R), CPU 3.00 GHz, 512 MB RAM.

i	MAPLE (umbral)	MAPLE (Bryc)
15	0.015	0.016
18	0.031	0.062
21	0.078	0.141
24	0.172	0.266
27	0.375	0.703

[27] The output is in the same form as the one given by Bryc's procedure.

Table 5 shows computation times for three different software packages. The first one has been implemented in `Mathematica` and refers to the procedures in [1], here called `AS` algorithms, see also `http://fisher.utstat.toronto.edu/david/SCSI/chap.3.nb`. The second one is the package `MathStatica` from [72]. The third package, `Fast` algorithms, has been implemented in `Maple 10.x` by using the results of Section 6. The source code is given in [24]. Let us remark that, for all procedures, the results are in the same output form and have been performed by the authors on the same platform. To the best of our knowledge, there is no `R` implementation for k-statistics and polykays.

TABLE 5. Computation times in sec. for k-statistics and polykays. Missing computation times mean "greater than 20 hours".

$k_{t,\ldots,l}$	`AS` Algorithms	`MathStatica`	`Fast` algorithms
k_5	0.06	0.01	0.01
k_7	0.31	0.02	0.01
k_9	1.44	0.04	0.01
k_{11}	8.36	0.14	0.01
k_{14}	396.39	0.64	0.02
k_{16}	57982.40	2.03	0.08
k_{18}	-	6.90	0.16
k_{20}	-	25.15	0.33
k_{22}	-	81.70	0.80
k_{24}	-	359.40	1.62
k_{26}	-	1581.05	2.51
k_{28}	-	6505.45	4.83
$k_{3,2}$	0.06	0.02	0.01
$k_{4,4}$	0.67	0.06	0.02
$k_{5,3}$	0.69	0.08	0.02
$k_{7,5}$	34.23	0.79	0.11
$k_{7,7}$	435.67	2.52	0.26
$k_{9,9}$	-	27.41	2.26
$k_{10,8}$	-	30.24	2.98
$k_{4,4,4}$	34.17	0.64	0.08

Table 6 is the same as Table 5, but for multivariate k-statistics and multivariate polykays. Missing computation times in the second column are due to the fact that no procedures are devoted to multivariate polykays in `MathStatica` (release 1).

TABLE 6. Computation times in sec. for multivariate k-statistics and multivariate polykays. In the first column, missing computation times mean "greater than 20 hours".

$k_{t_1 \ldots t_r; l_1 \ldots l_m}$	`AS` Algorithms	`MathStatica`	`Fast`-algorithms
$k_{3\,2}$	0.20	0.02	0.01
$k_{4\,4}$	18.20	0.14	0.02
$k_{5\,5}$	269.19	0.56	0.08
$k_{6\,5}$	1023.33	1.02	0.16
$k_{6\,6}$	-	2.26	0.33
$k_{7\,6}$	-	4.06	0.59
$k_{7\,7}$	-	8.66	1.23
$k_{8\,6}$	-	7.81	1.16
$k_{8\,7}$	-	15.89	2.59
$k_{8\,8}$	-	30.88	5.53
$k_{3\,3\,3}$	1211.05	0.92	0.44
$k_{4\,3\,3}$	-	2.09	0.34
$k_{4\,4\,3}$	-	4.98	1.02
$k_{4\,4\,4}$	-	13.97	2.78
$k_{1\,1;\,1\,1}$	0.05	-	0.01
$k_{2\,1;\,1\,1}$	0.20	-	0.01
$k_{2\,2;\,1\,1}$	1.30	-	0.03
$k_{2\,2;\,2\,1}$	6.50	-	0.06
$k_{2\,2;\,2\,2}$	34.31	-	0.11
$k_{2\,1;\,2\,1;\,2\,1}$	81.13	-	0.17
$k_{2\,2;\,1\,1;\,1\,1}$	30.34	-	0.14
$k_{2\,2;\,2\,1;\,1\,1}$	127.50	-	0.22
$k_{2\,2;\,2\,1;\,2\,1}$	406.78	-	0.47
$k_{2\,2;\,2\,2;\,1\,1}$	467.88	-	0.55
$k_{2\,2;\,2\,2;\,2\,1}$	1402.55	-	1.14
$k_{2\,2;\,2\,2;\,2\,2}$	3787.41	-	2.96

9. Appendix 2: Families of time-space harmonic polynomials

Hermite polynomials

The generalized Hermite polynomials $\left\{H_k^{(s^2 t)}(x)\right\}$ [71] are time-space harmonic polynomials with respect to the Wiener process with zero drift and variance s^2, whose umbral counterpart is the family of auxiliary umbrae $\{t{\boldsymbol .}\beta{\boldsymbol .}(s\eta)\}_{t \geq 0}$. Indeed,

$$H_k^{(s^2 t)}(x) = \mathbb{E}[(x - t{\boldsymbol .}\beta{\boldsymbol .}(s\eta))^k], \qquad k = 1, 2, \ldots,$$

with $s \in \mathbb{R}^+$ and η the umbra introduced in Example 3.2.

Poisson–Charlier polynomials

The Poisson–Charlier polynomials $\left\{\tilde{C}_k(x, \lambda t)\right\}$ [71] are time-space harmonic with respect to the Poisson process $\{N_t\}_{t\geq 0}$ with intensity parameter $\lambda > 0$, whose umbral counterpart is the family of auxiliary umbrae $\{(t\lambda)\boldsymbol{.}\beta\}_{t\geq 0}$. Indeed,

$$\tilde{C}_k(x, \lambda t) = \sum_{j=1}^{k} s(k, j)\, \mathbb{E}[(x - (t\lambda)\boldsymbol{.}\beta)^j], \qquad k = 1, 2, \ldots,$$

where $\{s(k, j)\}$ are the Stirling numbers of the first kind.

Lévy–Sheffer systems

The Lévy–Sheffer system [83] is a sequence of polynomials $\{V_k(x, t)\}_{t\geq 0}$ time-space harmonic with respect to the Lévy process umbrally represented by $\{-t\boldsymbol{.}\alpha\boldsymbol{.}\beta\boldsymbol{.}\gamma^{<-1>}\}_{t\geq 0}$, since

$$V_k(x, t) = \sum_{i=0}^{k} \mathbb{E}[(x - t\boldsymbol{.}\alpha\boldsymbol{.}\beta\boldsymbol{.}\gamma^{<-1>})^i]\, B_{k,i}(g_1, \ldots, g_{k-i+1}), \;\; k = 1, 2, \ldots,$$

where $B_{k,i}$ are partial Bell exponential polynomials and $\{g_i\}$ is the sequence umbrally represented by γ.

Laguerre polynomials

The Laguerre polynomials $\{\mathcal{L}_k^{t-k}(x)\}_{t\geq 0}$ [71] are time-space harmonic with respect to the gamma process $\{G_t(1, 1)\}_{t\geq 0}$, with scale parameter 1 and shape parameter 1, whose umbral counterpart is the family of auxiliary umbrae $\{t\boldsymbol{.}\bar{u}\}_{t\geq 0}$. Indeed,

$$(-1)^k\, k!\, \mathcal{L}_k^{t-k}(x) = \mathbb{E}[(x - t\boldsymbol{.}\bar{u})^k], \qquad k = 1, 2, \ldots,$$

where $\bar{u}$ is the Boolean unity.

Actuarial polynomials

The actuarial polynomials $\{g_k(x, \lambda t)\}_{t\geq 0}$ [71] are time-space harmonic with respect to the gamma process $\{G_t(\lambda, 1)\}_{t\geq 0}$, with scale parameter $\lambda > 0$ and shape parameter 1, whose umbral counterpart is the family of auxiliary umbrae $\{(\lambda t)\boldsymbol{.}\bar{u}\}_{t\geq 0}$. Indeed,

$$g_k(x, \lambda t) = \sum_{i=1}^{k} \mathbb{E}[(x - (\lambda t)\boldsymbol{.}\bar{u})^i]\, B_{k,i}(m_1, \ldots, m_{k-i-1}), \qquad k = 1, 2, \ldots,$$

where $B_{k,i}$ are partial Bell exponential polynomials and $m_i = \mathbb{E}[(\chi\boldsymbol{.}(-\chi))^i]$, for $i = 1, \ldots, k$.

Meixner polynomials

The Meixner polynomials of the first kind $\{M_k(x,t,p)\}_{t\geq 0}$ [71] are time-space harmonic with respect to the Pascal process $\{Pa(p,t)\}_{t\geq 0}$, whose umbral counterpart is the family of auxiliary umbrae $\{t.\bar{u}.d.\beta\}_{t\geq 0}$. Indeed,

$$(-1)^k\,(t)_k\,M_k(x,t,p) = \sum_{i=1}^{k} \mathbb{E}[(x - t.\bar{u}.d.\beta)^i]\,B_{k,i}(m_1,\ldots,m_{k-i-1}),$$

where $B_{k,i}$ are partial Bell exponential polynomials and for $i=1,\ldots,k$

$$m_i = \mathbb{E}\left[\left(\chi.\left(-1.\chi + \frac{\chi}{p}\right)\right)^i\right].$$

Bernoulli polynomials

The Bernoulli polynomials $\{B_k(x,n)\}$ [71] are time-space harmonic with respect to the random walk $\{n.(-1.\iota)\}_{n\geq 0}$. Indeed,

$$B_k(x,n) = \mathbb{E}\left[(x - n.\,(-1.\iota))^k\right], \qquad k=1,2,\ldots.$$

The n-th increment M_n in Example 3.12 is represented by the umbra $-1.\iota$ corresponding to a uniform r.v. on the interval $[0,1]$.

Euler polynomials

The Euler polynomials $\{\mathcal{E}_k(x,n)\}$ [71] are time-space harmonic with respect to the random walk $\{n.\left[\frac{1}{2}\,(-1.\xi + u)\right]\}_{n\geq 0}$. Indeed,

$$\mathcal{E}_k(x,n) = \mathbb{E}\left[\left(x - n.\left[\frac{1}{2}\,(-1.\xi + u)\right]\right)^k\right], \qquad k=1,2,\ldots.$$

The n-th increment M_n in Example 3.12 is represented by the umbra $\frac{1}{2}\,(-1.\xi\;+u)$ corresponding to a Bernoulli r.v. with parameter $1/2$.

Krawtchouk polynomials

The Krawtchouk polynomials $\{\mathcal{K}_k(x,p,n)\}$ [71] are time-space harmonic with respect to the random walk $\{n.(-1.\upsilon)\}_{n\geq 0}$. Indeed,

$$\frac{n!}{(n-k)!}\,\mathcal{K}_k(x,p,n) = \sum_{i=1}^{k} \mathbb{E}[(x + n.\upsilon)^i]\,B_{k,i}(m_1,\ldots,m_{k-i+1}),$$

where $B_{k,i}$ are partial Bell exponential polynomials,

$$m_i = \mathbb{E}\left[\left(\chi.\left(-1.\chi - \frac{\chi}{d}\right)\right)^i\right]$$

for $i=1,\ldots,k$, and $p/q = d$, $p+q=1$. The n-th increment M_n in Example 3.12 is represented by the umbra $-1.\upsilon$ corresponding to a Bernoulli r.v. with parameter p.

Pseudo–Narumi polynomials

The pseudo–Narumi polynomials $\{N_k(x, an)\}$, with $a \in \mathbb{N}$ [71], are time-space harmonic with respect to the random walk $\{(an).(-1{\bf .}\iota)\}_{n \geq 0}$, with $a{\bf .}(-1{\bf .}\iota)$ the umbral counterpart of a sum of $a \in \mathbb{N}$ r.v.'s with uniform distribution on the interval $[0,1]$. Indeed,

$$k!\, N_k(x, an) = \sum_{i=1}^{k} \mathbb{E}[(x - (an){\bf .}(-1{\bf .}\iota))^i] B_{k,i}(m_1, \ldots, m_{k-i+1}),$$

where $B_{k,i}$ are partial Bell exponential polynomials and $m_i = \mathbb{E}[(u^{<-1>})^i]$, for $i = 1, \ldots, k$.

10. Acknowledgments

I would like to thank Domenico Senato and Henry Crapo for helpful discussions and suggestions during the writing of this survey. I would like to also thank Pasquale Petrullo and Rosaria Simone for carefully reading the paper, leading to an improvement of its technical quality. A special thank to Christian Krattenthaler whose remarks and criticisms have stimulated the author in looking for a simplified exposure devoted to readers coming in new to the matter.

References

[1] David F. Andrews and Jamie E. Stafford, *Symbolic Computation for Statistical Inference*, Oxford Statistical Science Series, vol. 21, Oxford University Press, 2000.

[2] Michael Anshelevich, *Partition-dependent stochastic measures and q-deformed cumulants*, Doc. Math. **6** (2001), 343–384.

[3] David Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, 2004.

[4] Florin Avram and Murad S. Taqqu, *Symmetric polynomials of random variables attracted to an infinitely divisible law*, Probab. Theory Rel. Fields **27** (1986), 491–500.

[5] Florin Avram and Murad S. Taqqu, *Noncentral limit theorems and Appell polynomials*, Ann. Probab. **15** (1987), 767–775.

[6] Ole E. Barndorff-Nielsen, Thomas Mikosch and Sidney I. Resnick, *Lévy Processes: Theory and Applications*, Birkhäuser, 2001.

[7] William D. Baten, *Correction for the moments of a frequency distribution in two variables*, Ann. Math. Statist. **2** (1931), 309–319.

[8] Eric T. Bell, *Exponential polynomials*, Ann. of Math. **35** (1934), 258–277.

[9] Bruce C. Berndt, *Ramanujan's Notebooks*, Part 1 , Springer-Verlag, New York, 1985.

[10] James W. Brown, *On multivariable Sheffer sequences*, J. Math. Anal. Appl. **6**9 (1979), 398–410.

[11] Włodzimierz Bryc, *Computing moments of free additive convolution of measures*, Appl. Math. Comput. **194** (2007), 561–567.

[12] Gregory M. Constantine and Thomas H. Savits, *A stochastic representation of partition identities*, SIAM J. Discrete Math. **7** (1994), 194–202.

[13] Gregory M. Constantine and Thomas H. Savits, *A multivariate Faà di Bruno formula with applications*, Trans. Amer. Math. Soc. **348** (1996), 503–520.

[14] Cecil C. Craig, *Sheppard's corrections for a discrete variable*, Ann. Math. Statist. **7** (1936), 55–61.

[15] Arthur P. Dempster and Donald B. Rubin, *Rounding error in regression: the appropriateness of Sheppard's corrections*, J. R. Statist. Soc. Ser. B: Statist. Methodol. **45** (1983), 51–59.

[16] Alessandro Di Bucchianico and Daniel S. Loeb, *A selected Survey of Umbral Calculus*, Electron. J. Combin. **2** Dynamic Survey 3, updated 2000, 28 pp.

[17] Elvira Di Nardo, *A new approach to Sheppard's corrections*, Math. Methods Statist. **12** (2010), 151–162.

[18] Elvira Di Nardo, *Moment symbolic calculus in statistics*, 58-th World Statistics Congress of the International Statistical Institute, Dublin, 2011, `http://www.youtube.com/watch?v=bf7k1NGLQSY&feature=youtu.be`.

[19] Elvira Di Nardo, *On a symbolic representation of non-central Wishart random matrices with applications*, J. Mult. Anal. **1**25 (2014), 121–135.

[20] Elvira Di Nardo and Giuseppe Guarino, *New algorithms for computing Sheppard's corrections*, Worksheet Maple Software, 2010, available at `http://www.maplesoft.com/applications/view.aspx?SID=59515`.

[21] Elvira Di Nardo, Giuseppe Guarino and Domenico Senato, *A unifying framework for k-statistics, polykays and their multivariate generalizations*, Bernoulli **14** (2008), 440–468.

[22] Elvira Di Nardo, Giuseppe Guarino and Domenico Senato, *Symbolic computation of moments of sampling distributions*, Comput. Statist. Data Anal. **52** (2008), 4909–4922.

[23] Elvira Di Nardo, Giuseppe Guarino and Domenico Senato, *A new method for fast computing unbiased estimators of cumulants*, Statist. Comput. **19** (2009), 155–165.

[24] Elvira Di Nardo, Giuseppe Guarino and Domenico Senato, *Fast algorithms for k-statistics, polykays and their multivariate generalizations*, Worksheet Maple Software, 2009, available at `http://www.maplesoft.comapplications/view.aspx?SID=33041`.

[25] Elvira Di Nardo, Giuseppe Guarino and Domenico Senato, *A new algorithm for computing the multivariate Faà di Bruno's formula*, Appl. Math. Comp. **217** (2011), 6286–6295.

[26] Elvira Di Nardo, Peter McCullagh and Domenico Senato, *Natural statistics for spectral samples*, Ann. Stat. **4**1, (2013), 982–1004.

[27] Elvira Di Nardo, Heinrich Niederhausen and Domenico Senato, *A symbolic handling of Sheffer sequences*, Ann. Mat. Pura Appl. **190** (2011), 489–506.

[28] Elvira Di Nardo and Imma Oliva, *On the computation of classical, boolean and free cumulants*, Appl. Math. Comput. **208** (2009), 347–354.

[29] Elvira Di Nardo and Imma Oliva, *On some applications of a symbolic representation of non-centered Lévy processes*, Comm. Statist. Theory Methods **42** (2013), 3974–3988.

[30] Elvira Di Nardo and Imma Oliva, *A new family of time-space harmonic polynomials with respect to Lévy processes*, Ann. Mat. Pura Appl. (4) **192** (2013), 917–929.

[31] Elvira Di Nardo and Imma Oliva, *Multivariate Bernoulli and Euler polynomials via Lévy processes*, Appl. Math. Lett. **25** (2012), 1179–1184.

[32] Elvira Di Nardo, Pasquale Petrullo and Domenico Senato, *Cumulants and convolutions via Abel polynomials*, Europ. J. Combin. **31** (2010), 1792–1804.

[33] Elvira Di Nardo and Domenico Senato, *Umbral nature of the Poisson random variables*, in: Algebraic Combinatorics and Computer Science: a tribute to

Gian-Carlo Rota, Henry Crapo and Domenico Senato (eds.), Springer–Verlag, Milano, 2001, pp. 245–266.

[34] Elvira Di Nardo and Domenico Senato, *An umbral setting for cumulants and factorial moments*, Europ. J. Combin. **27** (2006), 394–413.

[35] Elvira Di Nardo and Domenico Senato, *The eleventh and twelveth problems of Rota's Fubini Lectures: from cumulants to free probability theory*, in: From Combinatorics to Philosophy: the Legacy of G.-C. Rota, E. Damiani, O. D'Antona, V. Marra, F. Palombi (eds.), Springer–Verlag, 2009, pp. 91–130.

[36] Elvira Di Nardo and Domenico Senato, *Symbolic solutions of some linear recurrences*, J. Statist. Plann. Inference **142**(2), 423–429.

[37] Peter Doubilet, *On the foundations of combinatorial theory VII: Symmetric functions through the theory of distribution and occupancy*, Stud. Appl. Math. **51** (1972), 377–396.

[38] Paul L. Dressel, *Statistical seminvariants and their estimates with particular emphasis on their relation to algebraic invariants*, Ann. Math. Statist. **11** (1940), 33–57.

[39] William Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, John Wiley & Sons, New York, 1966.

[40] Pedro G. Ferreira, Joao Magueijo and Joseph Silk, *Cumulants as non-Gaussian qualifiers*, Phys. Rev. D **56** (1997), 4592–4603.

[41] Ronald A. Fisher, *Moments and product moments of sampling distributions*, Proc. Lond. Math. Soc. **30** (1929), 199–238.

[42] Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009.

[43] Anindya Goswami and Arindam Sengupta, *Time-space polynomial martingales generated by a discrete-time martingale*, J. Theoret. Probab. **8** (1995), 417–431.

[44] Anders Hald, *On the history of the correction for grouping* $1873-1922$, Scand. J. Statist. **28** (2001), 417–428.

[45] Michael Hardy, *Combinatorics of partial derivatives*, Electron. J. Combin. **13** (2006), Article #R1, 13 pp.

[46] Daniel F. Heitjan, *Inference from grouped continuous data: a review*, Statist. Sci. **4** (1989), 164–183.

[47] Thomas Kailath and Adrian Segall, *Orthogonal functionals of independent increment processes*, IEEE Trans. Inform. Theory **22** (1976), 287–298.

[48] Edward L. Kaplan, *Tensor notation and the sampling cumulants of k-statistics*, Biometrika **39** (1952), 319–323.

[49] John E. Kolassa, *Series Approximation Methods in Statistics*, Lecture Notes in Statistics, vol. 88, Springer–Verlag, 1994.

[50] Germain Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Math. **1** (1972), 333–350.

[51] Stephan Lawi, *Hermite and Laguerre polynomials and matrix valued stochastic processes*, Elect. Comm. Probab. **13** (2008), 67–84.

[52] Roy B. Leipnik and Charles E. M. Pearce, *The multivariate Faà di Bruno formula and multivariate Taylor expansions with explicit integral remainder term*, Austral. New Zealand Industr. Appl. Math. J. **48** (2007), 327–341.

[53] Yuh-Dauh Lyuu, *Financial Engineering and Computation: Principles, Mathematics, Algorithms*, Cambridge University Press, Cambridge, 1999.

[54] Peter McCullagh, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1987.

[55] Joachim D. Müller, *Cumulant analysis in fluorescence fluctuation spectroscopy*, Biophys J. **86** (2004), 3981–3992.

[56] Jacques Neveu, *Discrete-Parameter Martingales*, North-Holland/American Elsevier, 1975.

[57] Alexandru Nica and Roland Speicher, *A "Fourier transform" for multiplicative functions on non-crossing partitions*, J. Algebraic Combin. **6** (1997), 141–160.

[58] Alexander Nica and Roland Speicher, *Lectures on the Combinatorics of Free Probability*, London Mathematical Society, Lecture Note Series, vol. 335, Cambridge University Press, 2006.

[59] Heinrich Niederhausen, *Finite Operator Calculus*, project book, 2010, available at `http://math.fau.edu/niederhausen/HTML/Research/UmbralCalculus/bookS2010.pdf`.

[60] Silvia Noschese and Paolo E. Ricci, *Differentiation of multivariable composite functions and Bell polynomials*, J. Comput. Anal. Appl. **5** (2003), 333–340.

[61] Christopher Parrish, *Multivariate umbral calculus*, Linear Multilinear Algebra **6** (1978), 93–109.

[62] Giovanni Peccati and Murad S. Taqqu, *Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation*, Bocconi and Springer, Milano, 2011.

[63] Pasquale Petrullo, *A symbolic treatment of Abel polynomials*, in: From Combinatorics to Philosophy: The Legacy of G.-C. Rota, E. Damiani, O. D'Antona, V. Marra, F. Palombi (eds.), Springer–Verlag, Milano, 2009, 210–221.

[64] Pasquale Petrullo, *A Symbolic Approach to Cumulants and Their Connections with The Abel Identity and The Kerov Polynomials.*, Ph.D. thesis, Università degli Studi della Basilicata, Potenza, 2011.

[65] Giovanni Pistone and Henry P. Wynn, *Cumulant varieties*, J. Symbolic Comput. **41** (2006), 210–221.

[66] Jim Pitman, *Some probabilistic aspects of set partitions*, Amer. Math. Monthly **104** (1997), 201–209.

[67] Jim Pitman and Richard P. Stanley, *A polytope related to empirical distributions, plane trees, parking functions and the associahedron*, Discrete Comput. Geom. **27**, (2002), 603–634.

[68] Yves Poupard, *Étude et dénombrement parallèles des partitions non-croisées d'un cycle et des découpages d'un polygone convexe*, Discrete Math. **2** (1972), 279–288.

[69] Sudhakar Prasad and Nicolas Menicucci, *Fisher information with respect to cumulants*, I.E.E.E. Trans. Inform. Theory **50** (2004), 638–642.

[70] John Riordan, *An Introduction to Combinatorial Analysis*, John Wiley and Sons, Inc., New York, 1958.

[71] Steven M. Roman and Gian-Carlo Rota, *The umbral calculus*, Adv. Math. **27** (1978), 95–188.

[72] Colin Rose and Murray D. Smith, *Mathematical Statistics with Mathematica*, Springer Text in Statistics, Springer–Verlag, 2002.

[73] Gian-Carlo Rota, *Twelve problems in probability no one likes to bring up*, in: Algebraic Combinatorics and Computer Science: a tribute to Gian-Carlo Rota, Henry Crapo and Domenico Senato (eds.), Springer–Verlag, Milano, 2001, pp. 57–93.

[74] Gian-Carlo Rota and Jianhong Shen, *On the combinatorics of cumulants*, J. Combin. Theory Ser. A **91** (2000), 283–304.

[75] Gian-Carlo Rota and Brian D. Taylor, *The classical umbral calculus*, SIAM J. Math. Anal. **25** (1994), 694–711.

[76] Gian-Carlo Rota, David Kahaner and Andrew Odlyzko, *On the foundations of combinatorial theory. VIII. Finite operator calculus*, J. Math. Anal. Appl. **42** (1973), 684–760.
[77] Gian-Carlo Rota, Jianhong Shen and Brian D. Taylor, *All polynomials of binomial type are represented by Abel polynomials*, Ann. Sci. Norm. Super. Pisa Cl. Sci. **25** (1998), 731–738.
[78] Gian-Carlo Rota and Timothy C. Wallstrom, *Stochastic integrals: a combinatorial approach*, Ann. Probab. **25** (1997), 1257–1283.
[79] Sandra Saliani and Domenico Senato, *Compactly supported wavelets through the classical umbral calculus*, J. Fourier Anal. Appl. **12** (2006), 27–36.
[80] Ken-Iti Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999.
[81] Thomas H. Savits, *Some statistical applications of Faà di Bruno*, J. Multivariate Anal. **97** (2006), 2131–2140.
[82] Wim Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley Series in Probability and Statistics, 2003.
[83] Wim Schoutens, *Stochastic Processes and Orthogonal Polynomials*, Lecture Notes in Statistics, vol. 146, Springer–Verlag, 2000.
[84] Arindam Sengupta, *Markov processes, time-space harmonic functions and polynomials*, Statist. Probab. Lett. **78** (2008), 3277–3280.
[85] Jianhong Shen, *Combinatorics for wavelets: the umbral refinement equation*, Stud. Appl. Math. **103** (1999), 121–147.
[86] William F. Sheppard, *On the calculation of most probable values of frequency-constants for data arranged according to equidistant division of a scale*, Proc. London Math. Soc. **29** (1898), 353–380.
[87] Peter J. Smith, *A recursive formulation of the old problem of obtaining moments from cumulants and vice versa*, Amer. Statist. **49** (1995), 217–218.
[88] Terry P. Speed, *Cumulants and partition lattices*, Austral. J. Statist. **25** (1983), 378–388.
[89] Roland Speicher, *Multiplicative functions on the lattice of non-crossing partitions and free convolution*, Math. Ann. **298** (1994), 611-628.
[90] Roland Speicher and Reza Woroudi, *Boolean convolution*, in: Free probability theory (Waterloo, ON, 1995), D. Voiculescu (ed.), Fields Inst. Commun., vol. 12, Amer. Math. Soc., Providence, R.I., 1997, pp. 267–279.
[91] Adriaan J. Stam, *Polynomials of binomial type and compound Poisson processes*, J. Math. Anal. Appl. **130** (1988), 493–508.
[92] Richard P. Stanley, *Enumerative Combinatorics*, vol. 1, Cambridge University Press, Cambridge, 1997.
[93] Benjamin Staude, Stefan Rotter and Sonja Grün, *CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains*, J. Comput. Neurosci. **29** (2010), 327–350.
[94] Alan Stuart and Keith Ord, *Kendall's Advanced Theory of Statistics, Vol. 1: Distribution Theory*, sixth ed., Edward Arnold, London, U.K., 1994.
[95] Bernd Sturmfels and Piotr Zwiernik, *Binary cumulant varieties*, Ann. Combin. **17** (2013), 229–250.
[96] Bryan D. Taylor, *Umbral presentations for polynomial sequences*, Comput. Math. Appl. **41** (2001), 1085–1098.
[97] Thorvald N. Thiele, *On an approximation formula*, Tidsskr. Math. **3** (1873), 22–31.

[98] Thorvald N. Thiele, *Elementär Jagttagelseslære*, Gyldendal, Copenhagen, 1897; translated into English: *The theory of observations*, London, 1903; reprinted in Ann. Math. Statist. **2** (1931), 165–308.
[99] David C. Torney, *Binary cumulants*, Adv. Appl. Math. **25** (2000), 34–40.
[100] Jacques Touchard, *Nombres exponentiels et nombres de Bernoulli*, Canad. J. Statist. **8** (1956), 305–320.
[101] John W. Tukey, *Some sampling simplified*, J. Amer. Statist. Assoc. **45** (1950), 501–519.
[102] Stephen B. Vardeman, *Sheppard's correction for variances and "the quantization noise model"*, I.E.E.E. Trans. Instr. Meas. **54** (2005), 2117–2119.
[103] Dan Voiculescu, *Symmetries of some reduced free product C^*-algebras*, in: Lecture Notes in Math., vol. 1132, Springer–Verlag, 1985, pp. 556–588.
[104] Dan Voiculescu, *Lectures on free probability theory*, in: Lectures on probability theory and statistics (Saint-Flour, 1998), P. Bernard (ed.), Lecture Notes in Math., vol. 1738, Springer–Verlag, 2000, pp. 281–349.
[105] Jan Vrbik, *Population moments of sampling distributions*, Comput. Statist. **20** (2005), 611–621.
[106] Herbert S. Wilf, *generatingfunctionology*, Academic Press, 2006.
[107] Christopher S. Withers, *A simple expression for the multivariate Hermite polynomials*, Statist. Probab. Lett. **47** (2000), 165–169.
[108] Piotr Zwiernik and Jim Q. Smith, *Tree cumulants and the geometry of binary tree models*, Bernoulli **18** (2012), 290–321.
[109] Piotr Zwiernik, *L-cumulants, L-cumulant embeddings and algebraic statistics*, J. Algebr. Statist. **3** (2012), 11–43.

Dept. of Mathematics, Computer Science and Economics, University of Basilicata, Campus Macchia Romana, Via dell'Ateneo Lucano 10, 85100 Potenza, Italy

E-mail address: `elvira.dinardo@unibas.it`