A Combinatorial Approach to Evaluation of Reliability of the Receiver Output for BPSK Modulation with Spatial Diversity

S. Bliudze D. Krob

{bliudze, dk}@lix.polytechnique.fr

LIX, École Polytechnique, Route de Saclay
91128 Palaiseau Cedex, France

Submitted: Jul 22, 2004; Accepted: Jan 3, 2006; Published: Jan 7, 2006
Mathematics Subject Classifications: 05E05, 05E10

Abstract

In the context of soft demodulation of a digital signal modulated with Binary Phase Shift Keying (BPSK) technique and in presence of spatial diversity, we show how the theory of symmetric functions can be used to compute the probability that the log-likelihood of a recieved bit is less than a given threshold ε. We show how such computation can be reduced to computing the probability that $U - V < \varepsilon$ (denoted $P(U - V < \varepsilon)$) where U and V are two real random variables such that $U = \sum_{i=1}^{N} |u_i|^2$ and $V = \sum_{i=1}^{N} |v_i|^2$ where the u_i's and v_i's are independent centered complex Gaussian variables with variances $\mathbb{E}[\,|u_i|^2\,] = \chi_i$ and $\mathbb{E}[\,|v_i|^2\,] = \delta_i$. We give two expressions in terms of symmetric functions over the alphabets $\Delta = (\delta_1, \ldots, \delta_N)$ and $X = (\chi_1, \ldots, \chi_N)$ for the first $2N-1$ coefficients of the Taylor expansion of $P(U-V < \varepsilon)$ in terms of ε. The first one is a quotient of multi-Schur functions involving two alphabets derived from alphabets Δ and X, which allows us to give an efficient algorithm for the computation of these coefficients. The second expression involves a certain sum of pairs of Schur functions $s_\lambda(\Delta)$ and $s_\mu(X)$ where λ and μ are complementary shapes inside a $N \times N$ rectangle. We show that such a sum has a natural combinatorial interpretation in terms of what we call square tabloids with ribbons and that there is a natural extension of the Knuth correspondence that associates a (0,1)-matrix to each square tabloid with ribbon. We then show that we can completely characterise the (0,1)-matrices that arise from square tabloids with ribbons under this correspondence.

1 Introduction

In this paper we show how combinatorial techniques, such as symmetric functions and the theory of Young tableaux, arise naturally in a rather applied context of digital communications. Let us, therefore, start by introducing the reader to some aspects of the latter.

Modulating numerical signals means transforming them into wave forms. Due to their importance in practice, modulation methods were widely studied in signal processing (see, for instance, chapter 5 of [18]). One of the most important problems in this area is the performance evaluation of the optimum receivers associated with a given modulation method, which leads to the computation of various probabilities of errors (see again [18]).

Among the different modulation protocols used in practical contexts, an important class consists of methods where the modulation reference (i.e. a fixed numerical sequence) is transmitted on the same channel as the usual signal. The demodulation decision is then based on at least two noisy signals, namely, the transmitted signal and the transmitted reference. It happens, however, that one can also take into account in the demodulation process several noisy copies of these two signals: one speaks then of demodulation with diversity. It appears that the probability of errors encountered in such contexts is of the following form:

$$P(U < V) = P\left(U = \sum_{i=1}^{N} |u_i|^2 \ < \ V = \sum_{i=1}^{N} |v_i|^2 \right) , \tag{1}$$

where the u_i and v_i's denote independent centered complex Gaussian random variables with variances equal to $\mathbb{E}[\,|u_i|^2\,] = \chi_i$ and $\mathbb{E}[\,|v_i|^2\,] = \delta_i$ for every $i \in [1, N]$ (cf also Section 3.1).

The problem of computing explicitely probabilities of this last type was studied in signal processing by several researchers (cf [2, 11, 18, 22]). The most interesting result in this direction is due to Barett ([2]) who obtained the following expression

$$P(U < V) = \sum_{k=1}^{N} \left(\prod_{i \neq k} \frac{1}{1 - \delta_k^{-1}\delta_i} \prod_{i=1}^{N} \frac{1}{1 + \delta_k^{-1}\chi_i} \right) ,$$

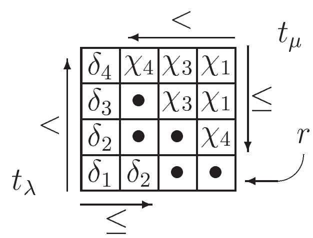
for the probability given by formula (1).

In this paper, we consider the *log-likelihood* of a bit — the value of the so-called *soft bit* obtained at the output of the rake receiver. This value allows one to decide what was the value of the transmitted bit, and is also essential for various decoding algorithms such as MAP and its variants and Soft Output Viterbi Algorithm (SOVA) (see for example Chapter 4 of [10]).

We start by giving some detailed background information on symmetric functions (Section 2), as well as a model describing the Binary Phase Shift Keying (BPSK) modulation (Section 3).

In Section 4, we consider the probability that a bit's log-likelihood is less than a given threshold ε and deduce two expressions in terms of symmetric functions for first coefficients of its Taylor expansion. One of these expressions leads to a stable and efficient algorithm computing these coefficients, whereas the second one allows an interesting combinatorial interpretation that we develop in Section 5.

This combinatorial interpretation involves a class of objects that we call square tabloids with ribbons. These are represented by triples of the form (t_λ, t_μ, r), where t_λ and t_μ are column strict Young tableaux of shapes λ and μ correspondingly, and r is a ribbon ending in the bottom righthand corner of the square N^N. Put together, λ and r (denoted $\lambda \cup r$) also form a Young diagram, of which μ is the complementary one in N^N.



We show that a Robinson-Schensted-Knuth correspondence can be naturally extended to associate a (0,1)-matrix to each square tabloid with ribbon, and we conclude by providing a complete and independent characterisation of the class of (0,1)-matrices that arise in this context.

2 Symmetric functions background

We present here the background on symmetric functions that is used in our paper. More information about symmetric functions can be found in Macdonald's classical textbook ([17]).

Let X be a set of indeterminates. The algebra of symmetric functions over X is then denoted by $Sym(X)$. We define the *complete symmetric functions* $S_k(X)$ by their generating series

$$\sigma_z(X) = \sum_{n=0}^{+\infty} S_n(X)\, z^n = \prod_{x\in X} \frac{1}{1-x\,z} \ . \tag{2}$$

We also define in the same way the *elementary symmetric functions* $\Lambda_k(X)$ by their generating series (which is a polynomial when X is finite)

$$\lambda_z(X) = \sum_{n=0}^{+\infty} \Lambda_n(X)\, z^n = \prod_{x\in X} (1+x\,z) \ . \tag{3}$$

In order to use complete and elementary symmetric functions indexed by any integer $k \in \mathbb{Z}$, we also set $S_k(X) = \Lambda_k(X) = 0$ for every $k < 0$. Every symmetric function can be expressed in a unique way as a product of complete or elementary symmetric functions. For every n-uple $I = (i_1, \ldots, i_n) \in \mathbb{Z}^n$, we now define the *Schur function* $s_I(X)$ as the minor taken over the rows $1, 2, \ldots, n$ and the columns $i_1+1, i_2+2, \ldots, i_n+n$ of the infinite matrix $\mathbb{S} = (S_{j-i}(X))_{i,j\in\mathbb{Z}}$, i.e.

$$s_I(X) = \begin{vmatrix} S_{i_1}(X) & S_{i_2+1}(X) & \ldots & S_{i_n+n-1}(X) \\ S_{i_1-1}(X) & S_{i_2}(X) & \ldots & S_{i_n+n-2}(X) \\ \vdots & \vdots & \vdots & \vdots \\ S_{i_1-n+1}(X) & S_{i_2-n+2}(X) & \ldots & S_{i_n}(X) \end{vmatrix} . \tag{4}$$

We also define more generally for every $I = (i_1, \ldots, i_n) \in \mathbb{Z}^n$ and $J = (j_1, \ldots, j_n) \in \mathbb{Z}^n$, the *skew Schur function* $s_{J/I}(X)$ as the minor of $\mathbb{S}$ taken over the rows $i_1+1, i_2+2, \ldots, i_n+n$ and the columns $j_1+1, j_2+2, \ldots, j_n+n$. The importance of Schur functions comes from the fact that the family of the Schur functions that are indexed by partitions form a classical linear basis of the algebra of symmetric functions (see [17] for the details).

Let us finally introduce the notion of multi-Schur function (see [16]) which is another natural generalization of usual Schur functions that we will use in this paper. Let $(X_i)_{1\le i\le n}$ be a family of n sets of indeterminates. For every n-uple $I = (i_1, \ldots, i_n) \in \mathbb{Z}^n$, one defines then the multi-Schur function $S_I(X_1, \ldots, X_n)$ by the determinantal formula

$$s_I(X_1, \ldots, X_N) = \begin{vmatrix} S_{i_1}(X_1) & S_{i_2+1}(X_2) & \ldots & S_{i_n+n-1}(X_n) \\ S_{i_1-1}(X_1) & S_{i_2}(X_2) & \ldots & S_{i_n+n-2}(X_n) \\ \vdots & \vdots & \vdots & \vdots \\ S_{i_1-n+1}(X_1) & S_{i_2-n+2}(X_2) & \ldots & S_{i_n}(X_n) \end{vmatrix} . \tag{5}$$

Hence the usual Schur function $s_I(X)$ is exactly the multi-Schur function $s_I(X, \ldots, X)$.

2.1 Transformations of alphabets

Let X and Y be two sets of indeterminates. The complete symmetric functions of the formal set $X+Y$ are then defined by their generating series

$$\sigma_z(X+Y) = \sum_{n=0}^{+\infty} S_n(X+Y)\, z^n = \sigma_z(X)\, \sigma_z(Y) . \tag{6}$$

One also defines the complete symmetric functions of the formal set $X-Y$ by setting

$$\sigma_z(X-Y) = \sum_{n=0}^{+\infty} S_n(X-Y)\, z^n = \sigma_z(X)\, \lambda_{-z}(Y) . \tag{7}$$

A symmetric function F of the alphabet $X+Y$ or $X-Y$ is then an element of $Sym(X) \otimes Sym(Y)$ whose expression in this last algebra can be obtained by developing F as a product of complete symmetric functions of $X+Y$ or $X-Y$ that are elements of $Sym(X) \otimes Sym(Y)$ according to the two defining relations (6) and (7). Note also that the complete symmetric functions of the formal set $-X$ are in particular defined by setting

$$\sigma_z(-X) = \sum_{n=0}^{+\infty} S_n(-X)\, z^n = \lambda_{-z}(X) . \tag{8}$$

In other words, if $F(X)$ is a symmetric function of the set X, the symmetric function $F(-X)$ is obtained by applying to F the algebra morphism that replaces $S_n(X)$ by $(-1)^n\,\Lambda_n(X)$ for every $n \geq 0$. Observe that the formal set $X-Y$ can also be defined by setting $X-Y = X+(-Y)$.

The expression of a Schur function of a formal sum of sets of indeterminates is in particular given by the Cauchy formula, which states that one has

$$s_\lambda(X+Y) = \sum_{\mu\subset\lambda} s_\mu(X)\,s_{\lambda/\mu}(Y) \tag{9}$$

for every partition λ (see [17]). One must also point out (see again [17]) that for all partitions μ and λ such that $\mu \subset \lambda$ one has

$$s_{\lambda/\mu}(-X) = s_{\lambda\tilde{}/\mu\tilde{}}(X) \tag{10}$$

where $\lambda\tilde{}$ and $\mu\tilde{}$ are the conjugate partitions of λ and μ correspondingly. Note finally that the resultant of two polynomials can in particular be expressed as a rectangular Schur function of a difference of alphabets. Let X and Y be two sets of respectively N and M indeterminates. The expression

$$R(X,Y) = \prod_{x\in X,\,y\in Y} (x-y)$$

is then the resultant of the polynomials that have X and Y as sets of roots and one can prove that one has $R(X,Y) = S_{N^M}(X-Y)$ (see [16, 17]).

2.2 Vertex operators

In the following, we will also use the vertex operator $\Gamma_z(X)$ that transforms every symmetric function of $Sym(X)$ into a series of $Sym(X)[[z,z^{-1}]]$. As the Schur functions indexed by partitions form a linear basis in $Sym(X)$, it is sufficient to define $\Gamma_z(X)$ only on the elements of the latter. We put

$$\Gamma_z(X)(s_\lambda(X)) = \sum_{m=-\infty}^{\infty} s_{(\lambda,m)}(X)\,z^m$$

for every partition $\lambda = (\lambda_1,\ldots,\lambda_n)$, with $(\lambda,m) = (\lambda_1,\ldots,\lambda_n,m) \in \mathbb{Z}^{n+1}$ for every $m \in \mathbb{Z}$. The following formula due to Thibon (cf [21]) gives then another explicit expression of the action of a vertex operator on a Schur function.

Proposition 2.1 (Thibon; [21]) *Let λ be a partition. Then one has*

$$\Gamma_z(X)(s_\lambda(X)) = \sigma_z(X)\,s_\lambda(X-1/z)\ . \tag{11}$$

2.3 Lagrange's operators

Let $X = \{ x_1, \ldots, x_N \}$ be a finite alphabet of N indeterminates. The *Lagrange interpolating operator* L is the operator that maps every polynomial f of $\mathbb{C}[X]$ symmetric in the last $N-1$ indeterminates, i.e. every element $f(x_1, X\backslash x_1)$ of $Sym(x_1) \otimes Sym(X\backslash x_1)$, onto the symmetric polynomial $L(f)$ of $Sym(X)$ defined by setting

$$L(f) = \sum_{k=1}^{N} \frac{f(x_k, X\backslash x_k)}{R(x_k, X\backslash x_k)}$$

where $R(A, B)$ stands again for the resultant of the two polynomials that have respectively the two sets of indeterminates A and B as sets of roots (cf Section 2.1). The following result, corresponding to the special case of Bott's formula for fibrations in projective lines (see [14, 15] for more details), gives then an interesting property of the Lagrange interpolation operator.

Theorem 2.1 (Lascoux; [14]) *Let $X = \{ x_1, \ldots, x_N \}$ be an alphabet of N indeterminates and let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition that contains $\rho_{N-1} = (N-2, \ldots, 2, 1, 0)$. Then one has*

$$L(x_1^k \, s_\lambda(X\backslash x_1)) = s_{(\lambda, k-N+1)}(X) \tag{12}$$

for every $k \geq 0$, where the Schur function involved in the right hand side of relation (12) is indexed by the sequence $(\lambda, k-N+1) = (\lambda_1, \ldots, \lambda_n, k-N+1)$ of $\mathbb{Z}^{n+1}$.

3 Signal processing background

We consider a model where one transmits a signal $b \in \{-1, +1\}$ on a noisy channel[1]. A reference $r = 1$ is also sent on the noisy channel at the same time as b. We assume that we receive N pairs $(x_i(b), r_i)_{1 \leq i \leq N} \in (\mathbb{C} \times \mathbb{C})^N$ of data (the $x_i(b)$'s) and references (the r_i's)[2] that have the following form

$$\begin{cases} x_i(b) &= \; a_i \, b + \nu_i \quad \text{for every } 1 \leq i \leq N, \\ \quad r_i &= \; a_i \sqrt{\beta_i} + \nu_i' \quad \text{for every } 1 \leq i \leq N, \end{cases}$$

where $a_i \in \mathbb{C}$ is a complex number that models the channel fading associated with $x_i(b)$[3], where $\beta_i \in \mathbb{R}^+$ is a positive real number that represents the signal to noise ratio (SNR) which is available for the reference r_i and where $\nu_i \in \mathbb{C}$ and $\nu_i' \in \mathbb{C}$ denote finally two independent complex white Gaussian noises. We also assume that every a_i is a complex

[1] This is the case, for example, when BPSK modulation is used. For a large number of other modulation methods the information transmitted is more complex, and contains more than one bit. However, performance analysis for these modulations can be reduced to that of BPSK (see [18]).

[2] One speaks in this case of spatial diversity, i.e. when more than one antenna is available, but also of multipath reflexion contexts. These two types of situations typically occur in mobile communications.

[3] Fading is typically the result of the absorption of the signal by buildings. Its complex nature comes from the fact that it models both an attenuation (its modulus) and a dephasing (its argument).

random variable distributed according to a centered Gaussian density of variance α_i for every $i \in [1, N]$.

According to these assumptions, all observables of our model, i.e. the pairs $(x_i(b), r_i)$ for all $1 \leq i \leq N$, are complex Gaussian random variables. We finally also assume that these N observables are N independent random variables which have their image in $\mathbb{C}^2$. Under these hypotheses we have the following formula for the *log-likelihood* that serves as a decision variable in BPSK

$$\Lambda_1 = \log\left(\frac{P(b=+1|X)}{P(b=-1|X)}\right) = \sum_{i=1}^{N} \frac{4\,\alpha_i\,\sqrt{\beta_i}}{1+\alpha_i\,(\beta_i+1)}\,(x_i(b)|r_i) \tag{13}$$

with $X = (x_i(b), r_i)_{1\leq i\leq N}$ and where $(\star|\star)$ denotes the Hermitian scalar product. One indeed decides that b was equal to 1 (resp. to -1) when the right hand side of (13) is positive (resp. negative). Often, when appropriate channel decoding mechanism is used, the actual value of log-likelihood (called in this case a *soft bit*) represents the reliability of the decoder's input. One obtains equation (1) by applying the parallelogram identity to (13).

The situation undesirable for both demodulation (increased chances of taking incorrect decision) and soft decoding algorithms (unreliable input) is when the log-likelihood is close to zero, i.e. $|\Lambda_1| < \varepsilon$. We shall therefore study the probability $P(U-V<\varepsilon)$ [4] generalising (1) where $P(U-V<0)$ is considered instead.

3.1 The analogue of Barret's formula

Let us consider two real random variables U and V defined, as in [6] by setting

$$U = \sum_{i=1}^{N} |u_i|^2 \quad \text{and} \quad V = \sum_{i=1}^{N} |v_i|^2$$

where u_i's and v_i's are independent centered complex Gaussian random variables with variances $\mathbb{E}[|u_i|^2] = \chi_i$ and $\mathbb{E}[|v_i|^2] = \delta_i$ for every $i \in [1, N]$. It is then easy to prove by induction on N that the probability distribution functions of U and V are equal to

$$d_U(x) = \sum_{j=1}^{N} \frac{\chi_j^{N-2}}{\prod\limits_{1\leq i\neq j\leq N}(\chi_j-\chi_i)}\, e^{-\frac{x}{\chi_j}} \quad \text{and} \quad d_V(x) = \sum_{k=1}^{N} \frac{\delta_k^{N-2}}{\prod\limits_{1\leq i\neq k\leq N}(\delta_k-\delta_i)}\, e^{-\frac{x}{\delta_k}} \tag{14}$$

when all variances χ_i and δ_i are distinct. One can then easily obtain

$$P(V>x) = \int_x^{+\infty} d_V(t)\,dt = \sum_{k=1}^{N} \frac{\delta_k^{N-1}}{\prod\limits_{1\leq i\neq k\leq N}(\delta_k-\delta_i)}\, e^{-\frac{x}{\delta_k}}. \tag{15}$$

[4] Probability $P(U-V<\varepsilon)$ can be studied independently as the distribution function of the random variable $U-V$ (cf. [9]).

We then have the following expression for $P(U-V<\varepsilon)$

$$P(U-V<\varepsilon)=\int_0^{+\infty} d_U(x)P(V>x-\varepsilon)\,dx.$$

Substituting relations (14) and (15), this last identity leads to the expression

$$P(U-V<\varepsilon)=\int_0^{+\infty}\sum_{j,k=1}^{N}\frac{\chi_j^{N-2}\,\delta_k^{N-1}}{\prod\limits_{1\le i\ne j\le N}(\chi_j-\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)}\,e^{-\frac{x}{\chi_j}}\,e^{-\frac{x}{\delta_k}}\,e^{\frac{\varepsilon}{\delta_k}}\,dx,$$

from which we immediately get the relation

$$P(U-V<\varepsilon)=\sum_{j,k=1}^{N}\frac{\chi_j^{N-1}\,\delta_k^{N}}{(\delta_k+\chi_j)\prod\limits_{1\le i\ne j\le N}(\chi_j-\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)}\,e^{\frac{\varepsilon}{\delta_k}}.$$

This last formula can now be rewritten as

$$P(U-V<\varepsilon)=\sum_{k=1}^{N}\frac{\delta_k^{N}e^{\frac{\varepsilon}{\delta_k}}}{\prod\limits_{1\le i\le N}(\delta_k+\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)}\left(\sum_{j=1}^{N}\frac{\prod\limits_{1\le i\ne j\le N}(\delta_k+\chi_i)}{\prod\limits_{1\le i\ne j\le N}(\chi_j-\chi_i)}\,\chi_j^{N-1}\right). \quad (16)$$

Finally we can deduce the analogue of Barret's formula (cf. [2, 6]):

$$P(U-V<\varepsilon)=\sum_{k=1}^{N}\frac{\delta_k^{2N-1}e^{\frac{\varepsilon}{\delta_k}}}{\prod\limits_{1\le i\le N}(\delta_k+\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)} \quad (17)$$

due to the fact that the internal sum in relation (16) is just the Lagrange interpolation expression taken at the points $(-\chi_j)_{1\le j\le N}$ for the polynomial δ_k^{N-1} (considered here as a polynomial of $\mathbb{C}[\chi_1,\ldots,\chi_N][\delta_k]$).

4 Symmetric functions expression

We shall try to represent the probability $P(U-V<\varepsilon)$ in terms of Schur functions. In order to do so, we have to get rid of the exponential in the numerator of the right hand side of (17). Replacing it by its Taylor expansion, we obtain

$$P(U-V<\varepsilon)=\sum_{m=0}^{+\infty}\sum_{k=1}^{N}\frac{\delta_k^{2N-m-1}}{\prod\limits_{1\le i\le N}(\delta_k+\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)}\times\frac{\varepsilon^m}{m!}\,.$$

We will now concentrate our efforts on the m-th coefficient of this exponential series, i.e.

$$P_m^{(N)}=P_m^{(N)}(\Delta,X)=\sum_{k=1}^{N}\frac{\delta_k^{2N-m-1}}{\prod\limits_{1\le i\le N}(\delta_k+\chi_i)\prod\limits_{1\le i\ne k\le N}(\delta_k-\delta_i)}\,. \quad (18)$$

This formula can be expressed using the Lagrange operator L. Indeed, let us set $\delta_k = x_k$ and $\chi_k = -y_k$ for every $k \in [1, N]$. Then one can rewrite (18) as

$$P_m^{(N)} = \sum_{k=1}^{N} \frac{x_k^{2N-m-1}}{R(x_k, Y) R(x_k, X \backslash x_k)}$$

where we denoted $X = \{ x_1, \ldots, x_N \}$ and $Y = \{ y_1, \ldots, y_N \}$, and where $R(A, B)$ stands for the resultant of two polynomials having A and B as sets of roots (see Section 2.1). Hence we have

$$P_m^{(N)} = \sum_{k=1}^{N} \frac{g(x_k, X \backslash x_k)}{R(x_k, X \backslash x_k)} = L(g) \tag{19}$$

where g stands for the element of $Sym(x_1) \otimes Sym(X \backslash x_1)$ defined by setting

$$g(x_1, X \backslash x_1) = g(x_1) = \frac{x_1^{2N-m-1}}{R(x_1, Y)} \,.$$

Observe now that one has

$$g(x_1, X \backslash x_1) = \frac{1}{R(X, Y)} \, x_1^{2N-m-1} f(x_1, X \backslash x_1)$$

where f stands for the element of $Sym(x_1) \otimes Sym(X \backslash x_1)$ defined by setting

$$f(x_1, X \backslash x_1) = R(X \backslash x_1, Y) = s_{(N^{N-1})}((X \backslash x_1) - Y) \tag{20}$$

(the last above equality comes from the expression of the resultant in terms of Schur functions given in Section 2.1). Observe now that the resultant $R(X, Y)$, being symmetric in the alphabet X, is a scalar for the operator L. It follows therefore from relation (19) that one has

$$P_m^{(N)} = \frac{L(x_1^{2N-m-1} f(x_1, X \backslash x_1))}{R(X, Y)}. \tag{21}$$

Let us now study the numerator of the right-hand side of relation (21) in order to give another expression for $P_m^{(N)}$. Note first that Cauchy formula leads to the development

$$s_{(N^{N-1})}((X \backslash x_1) - Y) = \sum_{\lambda \subset (N^{N-1})} s_\lambda(X \backslash x_1) s_{(N^{N-1})/\lambda}(-Y) \,. \tag{22}$$

According to the identities (20) and (22), we now obtain for $0 \leq m < 2N$ the relations

$$\begin{aligned} L(x_1^{2N-m-1} f(x_1, X \backslash x_1)) &= \sum_{\lambda \subset (N^{N-1})} L(x_1^{2N-m-1} s_\lambda(X \backslash x_1)) s_{(N^{N-1})/\lambda}(-Y) \\ &= \sum_{\lambda \subset (N^{N-1})} s_{(\lambda, N-m)}(X) s_{(N^{N-1})/\lambda}(-Y), \end{aligned}$$

the latter equality being an immediate consequence of Theorem 2.1. Using the expression of $s_{(N^{N-1})/\lambda}(-Y)$ given by equation (10) and going back to the definition of skew Schur functions, we can rewrite this expression as

$$L(x_1^{2N-m-1} f(x_1, X \backslash x_1)) = \sum_{\lambda \subset (N^{N-1})} (-1)^{|\lambda|} s_{(\lambda, N-m)}(X) s_{\overline{(\lambda, N)}}(Y)$$

where $0 \le m < 2N$, and $\overline{(\lambda, N)}$ denotes the complementary partition of (λ, N) in the square N^N. Going back to the initial variables, the signs disappear in the previous formula by homogeneity of Schur functions. Reporting the identity obtained in such a way into relation (21), we finally get an expression for $P_m^{(N)}$ in terms of Schur functions, i.e.

$$P_m^{(N)} = \frac{\displaystyle\sum_{\lambda \subset (N^{N-1})} s_{(\lambda, N-m)}(\Delta) s_{\overline{(\lambda,N)}}(X)}{\displaystyle\prod_{1 \le i,j \le N} (\chi_i + \delta_j)} \tag{23}$$

where $X = \{\chi_1, \ldots, \chi_N\}$, $\Delta = \{\delta_1, \ldots, \delta_N\}$.

4.1 A determinantal approach

Let us work again with the alphabets X and Y defined in Section 4. We saw there that

$$P_m^{(N)} = \frac{f_m^{(N)}(X, Y)}{R(X, Y)} \tag{24}$$

where $0 \le m < 2N$, and $f_m^{(N)}(X, Y)$ is a symmetric function of $Sym(X) \otimes Sym(Y)$ given by

$$f_m^{(N)}(X, Y) = \sum_{\lambda \subset (N^{N-1})} s_{(\lambda, N-m)}(X) s_{(N^{N-1})/\lambda}(-Y).$$

Let's now compute the action of the vertex operator $\Gamma_z(X)$ on the rectangle Schur function $s_{(N^{N-1})}(X - Y)$. Recall first that Cauchy formula shows that one has

$$s_{(N^{N-1})}(X - Y) = \sum_{\lambda \subset (N^{N-1})} s_\lambda(X)\, s_{(N^{N-1})/\lambda}(-Y).$$

Applying the vertex operator $\Gamma_z(X)$ to this expansion, we obtain

$$\begin{aligned} \Gamma_z(X)(s_{(N^{N-1})}(X - Y)) &= \sum_{\lambda \subset (N^{N-1})} \Gamma_z(X)(s_\lambda(X))\, s_{(N^{N-1})/\lambda}(-Y) \\ &= \sum_{k=-\infty}^{+\infty} \Big(s_{(\lambda,k)}(X)\, s_{(N^{N-1})/\lambda}(-Y) \Big) z^k. \end{aligned}$$

Hence $f_m^{(N)}(X, Y)$ is equal to the coefficient of z^{N-m} in the image of $s_{(N^{N-1})}(X - Y)$ under $\Gamma_z(X)$. On the other hand, using Cauchy formula in connection with relation (11), one can also write

$$\begin{aligned} \Gamma_z(X)(s_{(N^{N-1})}(X - Y)) &= \sigma_z(X) s_{(N^{N-1})}(X - Y - 1/z) \\ &= \sigma_z(X) \Big(\sum_{j=0}^{N-1} s_{(N^{N-1})/(1^j)}(X - Y)\, s_{(1^j)}(-1/z) \Big) \\ &= \Big(\sum_{i=0}^{+\infty} s_i(X) z^i \Big) \Big(\sum_{j=0}^{N-1} s_{(N^{N-1})/(1^j)}(X - Y)(-1/z)^j \Big) \end{aligned}$$

due to the fact that the only non-zero Schur functions of the alphabet $-1/z$ are indexed by column partitions of the form 1^k (and are equal to $(-1/z)^k$). The coefficient of z^{N-m} in the above product gives us then a new expression for $f_m^{(N)}(X,Y)$, i.e.

$$f_m^{(N)}(X,Y) = \sum_{k=0}^{N-1} (-1)^k s_{N-m+k}(X) s_{N^{N-1}/1^k}(X-Y). \tag{25}$$

One can easily verify now that this expression represents the development along the last column of the determinant

$$\begin{vmatrix} s_N(X-Y) & s_{N+1}(X-Y) & \dots & s_{2N-2}(X-Y) & s_{2N-m-1}(X) \\ s_{N-1}(X-Y) & s_N(X-Y) & \dots & s_{2N-3}(X-Y) & s_{2N-m-2}(X) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_2(X-Y) & s_3(X-Y) & \dots & s_N(X-Y) & s_{N-m+1}(X) \\ s_1(X-Y) & s_2(X-Y) & \dots & s_{N-1}(X-Y) & s_{N-m}(X) \end{vmatrix}, \tag{26}$$

itself — an expression of the multi-Schur function $s_{(N^{N-1},N-m)}(X-Y,\dots,X-Y,X)$. Thus relation (26) gives us both a determinantal and a multi-Schur expression for the denominator of the right hand side of formula (24). Using the interpretation of the resultant $R(X,Y)$ as a Schur function (see again Section 2.1), we can conclude that for $0 \le m < 2N$

$$P_m^{(N)} = \frac{s_{(N^{N-1},N-m)}(X-Y,\dots,X-Y,X)}{s_{(N^N)}(X-Y)} \tag{27}$$

where the alphabet $X-Y$ appears $N-1$ times in the numerator of the right hand side of the above formula.

4.2 A Toeplitz system and its solution

Using the determinantal expression of the Schur function $s_{(N^N)}(X-Y)$, we can now observe that for all $0 \le m < 2N$ relation (27) shows that $P_m^{(N)}$ is equal to the quotient of the determinant (26) by the determinant

$$\begin{vmatrix} s_N(X-Y) & s_{N+1}(X-Y) & \dots & s_{2N-1}(X-Y) \\ s_{N-1}(X-Y) & s_N(X-Y) & \dots & s_{2N-2}(X-Y) \\ \vdots & \vdots & \ddots & \vdots \\ s_1(X-Y) & s_2(X-Y) & \dots & s_N(X-Y) \end{vmatrix},$$

which is obtained by replacing the last column of the determinant (26) by the N-dimensional vector $(s_{2N-1}(X-Y), s_{2N-2}(X-Y),\dots,s_N(X-Y))$. Hence the right hand side of relation (27) can be interpreted as the Cramer expression for the last component p_0 of the linear system

$$\begin{pmatrix} s_N(X-Y) & s_{N+1}(X-Y) & \dots & s_{2N-1}(X-Y) \\ s_{N-1}(X-Y) & s_N(X-Y) & \dots & s_{2N-2}(X-Y) \\ \vdots & \vdots & \ddots & \vdots \\ s_1(X-Y) & s_2(X-Y) & \dots & s_N(X-Y) \end{pmatrix} \begin{pmatrix} p_{N-1} \\ p_{N-2} \\ \vdots \\ p_0 \end{pmatrix} = \begin{pmatrix} s_{2N-m-1}(X) \\ s_{2N-m-2}(X) \\ \vdots \\ s_{N-m}(X) \end{pmatrix}.$$

Let us now set $\pi_m(t) = p_0 + p_1\, t + \cdots + p_{N-1}\, t^{N-1}$. The above linear system implies that the coefficients of order N to $2N-1$ in the series $\pi_m(t)\, \sigma_t(X-Y)$ are equal to the coefficients of the same order in $t^m \sigma_t(X)$. Equivalently, this means that there exists a polynomial $\mu_m(t)$ of degree less than or equal to $N-1$ such that one has

$$\pi_m(t)\, \sigma_t(X-Y) - t^m \sigma_t(X) + \mu_m(t) = O(t^{2N}) \ .$$

Going back to the definition of $\sigma_t(X-Y)$, one can notice that this property can be rewritten as

$$(\pi_m(t)\lambda_{-t}(Y) - t^m)\sigma_t(X) + \mu_m(t) = O(t^{2N})$$

that is itself clearly equivalent to

$$\pi_m(t)\lambda_{-t}(Y) + \mu_m(t)\lambda_{-t}(X) = t^m + O(t^{2N}).$$

Since the left hand side of the above identity is a polynomial of degree at most $2N-1$ and we are only considering m such that $0 \le m < 2N$, it follows that its right hand side must be equal to t^m. Hence we have shown that one has

$$\pi_m(t)\lambda_{-t}(Y) + \mu_m(t)\lambda_{-t}(X) = t^m. \tag{28}$$

Thus, for $0 \le m < 2N$, $P_m^{(N)}$ is the constant term $\pi_m(0)$ of the polynomial $\pi_m(t)$, where $\pi_m(t)$ and $\mu_m(t)$ are the polynomials of degree $\le N-1$ defined by (28).

4.3 A Bezoutian algorithm

We now present an algorithm that computes π_m and μ_m iteratively, starting with $m = 0$, and then consecutively deriving π_m and μ_m from π_{m-1} and μ_{m-1} for $m = 1 \ldots 2N-1$.

Algorithm 4.1 (Calculating the polynomials π_m and μ_m)
Input: Alphabets $\Delta = \{\delta_1, \ldots, \delta_N\}$ and $X = \{\chi_1, \ldots, \chi_N\}$.
Output: For all $m = 0 \ldots 2N-1$, a pair of polynomials (π_m, μ_m) satisfying (28).

For $m = 0$, the right hand side of the equality (28) is 1, i.e. the greatest common divisor of $\lambda_{-t}(Y)$ and $\lambda_{-t}(X)$. This implies that we can use the Generalised Euclidean algorithm as first step of our algorithm.

- **Step 0.1.** Consider the two polynomials $X(t)$ and $\Delta(t)$ of $\mathbb{R}[t]$ defined by setting

$$X(t) = \prod_{i=1}^{N} (1 - \chi_i\, t) \quad \text{and} \quad \Delta(t) = \prod_{i=1}^{N} (1 + \delta_i\, t).$$

- **Step 0.2.** Compute the unique polynomial $\pi_0(t)$ of $\mathbb{R}[t]$ of degree $d(\pi_0) \le N-1$ such that

$$\pi_0(t)X(t) + \mu_0(t)\Delta(t) = 1$$

  where $\mu_0(t)$ stands for some polynomial of $\mathbb{R}[t]$ of degree $d(\mu_0) \le N-1$.

Suppose now that, at `Step` $m-1$, we have found the polynomials $\pi_k(t)$ and $\mu_k(t)$ for all $k < m$. Then the following `Step` m provides us the next pair of polynomials $\pi_m(t)$ and $\mu_m(t)$ of degrees $\leq N-1$, satisfying the relation (28).

- **Step *m*.1.** We suppose that $0 < m < 2N$. Let then

$$c = \frac{[t^{N-1}](\mu_{m-1})}{\chi_1 \ldots \chi_N} = (-1)^{N-1} \frac{[t^{N-1}](\pi_{m-1})}{\delta_1 \ldots \delta_N} \ , \tag{29}$$

  where $[t^{N-1}](\pi)$ stands for the coefficient of t^{N-1} in the polynomial $\pi(t)$.

- **Step *m*.2.** We then define

$$\begin{cases} \pi_m(t) \, = t \, \pi_{m-1}(t) + (-1)^N c \, \Delta(t) \\ \mu_m(t) \, = t \, \mu_{m-1}(t) - (-1)^N c \, X(t) \end{cases} \tag{30}$$

  to obtain the required polynomials.

Let us now prove the consistency of this algorithm.

Proposition 4.1 *The polynomials $\pi_m(t)$ and $\mu_m(t)$, produced by Algorithm 4.1, satisfy relation (28).*

Proof – We argue by induction on m. The case $m = 0$ being obvious, we can consider only $m \geq 1$.

Suppose that, at `Step` $m-1$, we have found the two polynomials $\pi_{m-1}(t)$ and $\mu_{m-1}(t)$ of degrees $\leq N-1$ satisfying the relation (28) for $m-1$. First of all, observe that it follows immediately from (28) and the fact that $m < 2N$ that $d(\pi_{m-1}) = N-1$, then also $d(\mu_{m-1}) = N-1$ and vice versa.

If we have $d(\pi_{m-1}), d(\mu_{m-1}) < N-1$, then the two polynomials $\pi_m(t) = t\pi_{m-1}(t)$ and $\mu_m(t) = t\mu_{m-1}(t)$ satisfy (28) for m. Observe that in this case the coefficient c calculated in `Step` $m.1$ is zero, and hence relation (30) holds.

Suppose now that $d(\pi_{m-1}) = d(\mu_{m-1}) = N-1$. Then we can set

$$\begin{cases} \pi_{m-1}(t) = a \, t^{N-1} + \pi'(t) \\ \mu_{m-1}(t) = b \, t^{N-1} + \mu'(t) \end{cases} \quad \text{with } d(\pi'), d(\mu') < N-1 \, . \tag{31}$$

At the same time, one can easily see from the definition of $X(t)$ and $\Delta(t)$ that

$$\begin{cases} X(t) = (-1)^N \chi_1 \ldots \chi_N \cdot t^N + X'(t) \\ \Delta(t) = \delta_1 \ldots \delta_N \cdot t^N + \Delta'(t) \end{cases} \quad \text{with } d(X'), d(\Delta') = N-1 \, . \tag{32}$$

Substituting these four equations into (28) we obtain:

$$\begin{aligned} t^{m-1} \;\; &= \;\; \pi_{m-1}(t) X(t) + \mu_{m-1}(t) \Delta(t) \\ &= \;\; \Big((-1)^N a \, \chi_1 \ldots \chi_N + b \, \delta_1 \ldots \delta_N \Big) t^{2N-1} + O(t^{2N-2}) \end{aligned}$$

As the degree of the left-hand side of this equation is $m-1 < 2N-1$, we conclude that

$$(-1)^N a\,\chi_1\dots\chi_N + b\,\delta_1\dots\delta_N = 0.$$

Therefore we can define the coefficient c as

$$c = (-1)^N \frac{a}{\delta_1\dots\delta_N} = \frac{b}{\chi_1\dots\chi_N}\ . \tag{33}$$

Let us now put $\pi_m(t) = t\,\pi_{m-1}(t) - d\Delta(t)$, where d is a coefficient such that $d(\pi_m) \leq N-1$. Indeed, substituting (31), (32), and (33) into this formula we obtain:

$$\begin{array}{rcl} \pi_m(t) & = & (-1)^{N-1}c\,\delta_1\dots\delta_N\,t^N + t\pi'(t) - d\,\delta_1\dots\delta_N t^N - d\,\Delta'(t) \\ & = & \Big((-1)^{N-1}c - d\Big)\delta_1\dots\delta_N\,t^N + \underbrace{t\,\pi'(t) - d\,\Delta'(t)}_{\deg\ \leq\ N-1}\ , \end{array}$$

thus it is sufficient to take $d = (-1)^{N-1}c$ in order to have $d(\pi_m) \leq N-1$. Applying the same reasoning to $\mu_m(t)$ we obtain the following two expressions:

$$\left\{ \begin{array}{rcl} \pi_m(t) & = & t\,\pi_{m-1}(t) - (-1)^{N-1}c\,\Delta(t) \\ \mu_m(t) & = & t\,\mu_{m-1}(t) - (-1)^N c\,X(t) \end{array} \right. , \tag{34}$$

where $d(\pi_m), d(\mu_m) \leq N-1$. In order to complete our proof we have to check that these two polynomials satisfy (28):

$$\begin{aligned} t^m &= t\,\pi_{m-1}(t)X(t) + t\,\mu_{m-1}(t)\Delta(t) \\ &= \Big(\pi_m(t) + (-1)^{N-1}c\,\Delta(t)\Big)X(t) + \Big(\mu_m(t) + (-1)^N c\,X(t)\Big)\Delta(t) \\ &= \pi_m(t)X(t) + \mu_m(t)\Delta(t) \end{aligned}$$

This ends our proof. ∎

Note 4.1 We recall that $\pi_m(0) = P_m^{(N)}$.

5 Combinatorial interpretation

5.1 A special case

Recall that originally $P_m^{(N)}$ has been defined as the m-th coefficient of the decomposition of $P(U-V<y)$ into an exponential series (cf. Section 4),

$$P(U-V<\varepsilon) = \sum_{m=0}^{\infty} P_m^{(N)} \times \frac{\varepsilon^m}{m!}\,,$$

and therefore, taking $y = 0$, we obtain (see also (23))

$$P(U < V) = P_0^{(N)} = \frac{\displaystyle\sum_{\lambda \subset (N^{N-1})} s_{(\lambda,N)}(\Delta) s_{\overline{(\lambda,N)}}(X)}{\displaystyle\prod_{1 \le i,j \le N} (\chi_i + \delta_j)} \,. \tag{35}$$

This expression for the probability $P(U < V)$ has been obtained in [6], while in [13] it has been given a combinatorial interpretation. Indeed, as $\lambda \subset (N^{N-1})$ we have $\lambda_i \le N$ for all i, and thus (λ, N) is also a partition.

It is well known that a Schur function over an alphabet A indexed by some partition λ can be expressed as a sum

$$s_\lambda(A) = \sum_{t_\lambda} m(t_\lambda)\,,$$

where t_λ runs through all possible Young tableaux over A of shape λ, and $m(t_\lambda)$ is the monomial obtained by taking the product of all elements of A contained in t_λ. For example,

$$s_{(1,2)}(\{\chi_1, \chi_2\}) = \chi_1^2\chi_2 + \chi_1\chi_2^2\,,$$

which corresponds to

$$\begin{array}{|c|c} \cline{1-1} \chi_2 & \\ \hline \chi_1 & \multicolumn{1}{c|}{\chi_1} \\ \hline \end{array} \quad + \quad \begin{array}{|c|c} \cline{1-1} \chi_2 & \\ \hline \chi_1 & \multicolumn{1}{c|}{\chi_2} \\ \hline \end{array}\,.$$

In this manner one can represent the numerator of the fraction in (35) as a sum of the monomials corresponding to (N^N) square tabloids consisting of a Young tableau over the alphabet Δ and a complimentary one over X. For example,

$$P_0^{(2)} = \frac{\chi_1\chi_2(\delta_1^2 + \delta_1\delta_2 + \delta_2^2) + (\chi_1 + \chi_2)(\delta_1^2\delta_2 + \delta_1\delta_2^2) + \delta_1^2\delta_2^2}{(\chi_1 + \delta_1)(\chi_1 + \delta_2)(\chi_2 + \delta_1)(\chi_2 + \delta_2)}$$

corresponds to

$$\begin{array}{|c|c|} \hline \chi_2 & \chi_1 \\ \hline \delta_1 & \delta_1 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \chi_2 & \chi_1 \\ \hline \delta_1 & \delta_2 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \chi_2 & \chi_1 \\ \hline \delta_2 & \delta_2 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \delta_2 & \chi_1 \\ \hline \delta_1 & \delta_1 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \delta_2 & \chi_2 \\ \hline \delta_1 & \delta_1 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \delta_2 & \chi_1 \\ \hline \delta_1 & \delta_2 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \delta_2 & \chi_2 \\ \hline \delta_1 & \delta_2 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \delta_2 & \delta_2 \\ \hline \delta_1 & \delta_1 \\ \hline \end{array}\,.$$

For an arbitrary m such that $0 < m < 2N$, it is possible that $(\lambda, N - m)$ (see again (23)) is not a partition. In order to obtain an analogous representation of $P_m^{(N)}$ we will have to introduce a more complex combinatorial object — *square tabloid with ribbon*.

5.2 Square tabloids with ribbons

Definition 5.1 *A* ribbon *in a Young diagram is a connected chain of boxes not containing a 2×2 square such that any box has at most two neighbours (see Figure 1). The number of boxes in a ribbon is its* length.

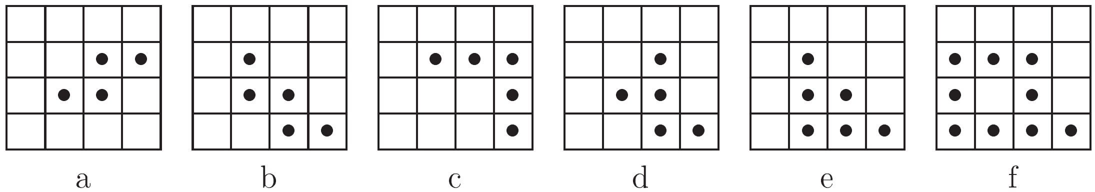


Figure 1: Several examples and counter-examples for the Definition 5.1 of ribbons.

The examples *a*–*c* in the Figure 1 are correct ribbons, while *d*–*f* are not. In this note we will only consider those that start in the lower right-hand corner of the square, and go to the North-West (examples *b*, *c*). For the sake of simplicity we will omit these two conditions when refering to ribbons.

We shall denote by $\mathcal{R}^{(N)}$ the set of all such ribbons in (N^N). Denoting by $\mathcal{R}^{(N)}_m$ the subset of $\mathcal{R}^{(N)}$ consisting of ribbons of a given length m, one obtains the following obvious decomposition:

$$\mathcal{R}^{(N)} = \bigcup_{m=1}^{2N-1} \mathcal{R}^{(N)}_m.$$

In a way analoguous to the one used to represent Young diagrams as a partition of an integer, a ribbon is fully described by a sequence $(r_1, \ldots, r_k)$, where r_i is the number of boxes in its i-th row. For example, the ribbons b and c from Figure 1 are (2, 2, 1) and (1, 1, 3) correspondingly.

For any given $r \in \mathcal{R}^{(N)}$, we shall consider all Young diagrams $\lambda \subset (N^N)$, such that $\lambda \cup r$ is also a Young diagram, together with $\mu = \overline{\lambda \cup r}$ — diagram complementary to the latter. Here the union '$\cup$' is taken in its geometrical sense — the union of two sets of boxes.

We can introduce one of our bijection's domains by considering all possible triplets consisting of a ribbon and two Young tableaux of shapes λ and μ on the alphabets $\Delta = \{\delta_1, \ldots, \delta_N\}$ and $X = \{\chi_1, \ldots, \chi_N\}$ correspondingly, such that put together they form a complete square:

$$\mathcal{T}^{(N)}_m = \{(t_\lambda, t_\mu, r) | r \in \mathcal{R}^{(N)}_m, \lambda \cup r \subset (N^N), \mu = \overline{\lambda \cup r}\},$$

where $1 \le m \le 2N-1$, λ and μ are the shapes of t_λ, and t_μ (see Figure 2).

From the combinatorial point of view this construction means that we take a square N by N, cut out a ribbon of length m, then split the rest into two Young tableaux over alphabets Δ and X.

Proposition 5.1 *We have the following representation for* $P^{(N)}_m$

$$P^{(N)}_m = \frac{\sum\limits_{t \in \mathcal{T}^{(N)}_m} (-1)^{h(t)-1} m(t)}{\prod\limits_{1 \le i,j \le N} (\chi_i + \delta_j)},$$

$$\begin{array}{|c|c|c|c|}\hline \delta_4 & \chi_4 & \chi_3 & \chi_1 \\ \hline \delta_3 & \bullet & \chi_3 & \chi_1 \\ \hline \delta_2 & \bullet & \bullet & \chi_4 \\ \hline \delta_1 & \delta_2 & \bullet & \bullet \\ \hline \end{array}$$

Figure 2: A typical element of $\mathcal{T}_4^{(5)}$.

where $m(t)$ is the monomial corresponding to t and $h(t)$ is the height of the ribbon of t defined as follows. If $t = (t_\lambda, t_\mu, r) \in \mathcal{T}_m^{(N)}$ is a square tabloid with ribbon, and $r = (r_1, \dots, r_N)$ then $h(t) = \max\{i | r_i > 0\}$.

Proof – This proposition becomes an obvious consequence of (23), if we consider the fact that

$$s_{(\dots,i,j,\dots)}(A) = -s_{(\dots,j+1,i-1,\dots)}(A)\,.$$

■

Example 5.1

$$P_2^{(2)} = \frac{\chi_1\chi_2 - \delta_1\delta_2}{(\chi_1+\delta_1)(\chi_1+\delta_2)(\chi_2+\delta_1)(\chi_2+\delta_2)}\,,$$

can be represented as the following difference

$$\begin{array}{|c|c|}\hline \chi_2 & \chi_1 \\ \hline \bullet & \bullet \\ \hline \end{array} \; - \; \begin{array}{|c|c|}\hline \delta_2 & \bullet \\ \hline \delta_1 & \bullet \\ \hline \end{array}\,.$$

■

5.3 Description of the bijection

In the following we shall construct a bijection between the square tabloids with ribbons introduced in the previous section, and a subset $\mathfrak{M}_m^{(N)}$ of $\mathcal{M}_{N\times(N+m)}$ [5] that we will define later on. We shall split a matrix from this set into two parts: the one on the left-hand side containing N columns, and the one on the right-hand side — m columns. The right part is the one that will eventually generate the ribbon in the corresponding tableau, and has only one 1 in each of its columns. Meanwhile the left part will be responsible for the tableau t_λ, and has one 1 for each of its boxes. Going back to the example of Figure 2, the corresponding matrix would be

$$\left(\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right)\,.$$

[5] In the following we will only consider $\{0,1\}$-matrices, therefore we use $\mathcal{M}_{m\times n}$ as a shorthand notation for $\mathcal{M}_{m\times n}(\{0,1\})$.

$$
\begin{array}{|c|c|c|c|}
\hline
\chi_4 & \chi_3 & \chi_2 & \chi_1 \\
\hline
\delta_4 & 6 & \chi_3 & \chi_1 \\
\hline
\delta_2 & 5 & 8 & \chi_4 \\
\hline
\delta_1 & \delta_2 & 7 & 9 \\
\hline
\end{array}
\qquad
\begin{array}{cccc}
 & 4\,3\,2\,1 & & \\
4\,6 & & 3\,1 & \\
2\,5\,8 & & & 4 \\
1\,2\,7\,9 & & & \\
\end{array}
\qquad
P:\ \begin{array}{l}
4\,6 \\
2\,5\,8 \\
1\,2\,7\,9
\end{array}
\qquad
Q:\ \begin{array}{l}
4 \\
3\,3 \\
2\,2\,4 \\
1\,1\,2
\end{array}
$$

a b c

$$
\left(\begin{array}{cccc|ccccc}
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}\right)
$$

d

Figure 3: Applying the algorithm to an element of $\mathcal{T}_4^{(5)}$.

We will now present an algorithm that, given a square tabloid $T \in \mathcal{T}_m^{(N)}$, will construct a matrix from $\mathcal{M}_{N\times(N+m)}$. As this algorithm is based on the Knuth correspondence (cf. [7, 12, 13]), it can be easily seen that it is reversible, thus, denoting by $\mathfrak{M}_m^{(N)}$ the image of the mapping defined by this algorithm, we obtain a bijection between $\mathfrak{M}_m^{(N)}$ and $\mathcal{T}_m^{(N)}$.

Algorithm 5.1 (Construction of a $\{0,1\}$-matrix from a square tabloid)
Input: $T = (t_\lambda, t_\mu, r)$ *— a square tabloid in* $\mathcal{T}_m^{(N)}$.
Output: A $\{0,1\}$*-matrix* $M \in \mathcal{M}_{N\times(N+m)}$.

- **Step 1.** Number each box of r starting with $N+1$ and up to $N+m$ from bottom to top and from left to right (see Figure 3-*a*).
- **Step 2.** Replace all δ_i in t_λ and χ_i in t_μ by i, and join t_λ with r to obtain two Young tableaux P and $\overline{Q}$ of shapes $\lambda \cup r$ and $\mu = \overline{\lambda \cup r}$ correspondingly (Figure 3-*b*).
- **Step 3.** Apply the same procedure as in [13] to obtain a tableau Q of a shape conjugated to $\lambda \cup r$ (Figure 3-*c*; see Example 5.2 below, or [13] for a formal description).
- **Step 4.** Finally apply Knuth's bijection based on column bumping to the pair (P, Q) of Young tableaux of conjugate forms to obtain a matrix from $\mathcal{M}_{N\times(N+m)}$ (Figure 3-*d*).

Example 5.2 Let us elaborate on the example given along the algorithm. We start with the following square tabloid from $\mathcal{T}_4^{(5)}$:

$$
\begin{array}{|c|c|c|c|}
\hline
\chi_4 & \chi_3 & \chi_2 & \chi_1 \\
\hline
\delta_4 & \bullet & \chi_3 & \chi_1 \\
\hline
\delta_2 & \bullet & \bullet & \chi_4 \\
\hline
\delta_1 & \delta_2 & \bullet & \bullet \\
\hline
\end{array}
$$

As the size of the square is 4, and the length of the ribbon — 5, the classical numbering for the ribbon goes from 5 to 9 and is shown in Figure 3-a. Re-labelling the rest of the tabloid and re-arranging it as indicated in the Step 2 of Algorithm 5.1 we obtain the two Young tableaux shown in Figure 3-b.

In Step 3 we take the upper right-hand corner tableau

$$\begin{array}{|c|ccc}\cline{1-1} 4 \\ \cline{1-1} 3 \\ \cline{1-2} 2 & \multicolumn{1}{c|}{3} \\ \cline{1-3} 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{4} \\ \cline{1-3}\end{array}$$

and we transform it into another one of complementary shape by applying the following procedure.

First of all, we consider this tableau as having four columns — the last one being empty. Then we form another tableau by putting in its first column all the numbers between 1 and 4 that are not in the last column of this one. As the last column of the original tableau is empty, we put in the first column of the new one all numbers 1–4. In the second column we put all the numbers that are not in the third one of the original tableau (i.e. 1–3), etc. Thus we obtain the tableau Q from Figure 3-c.

The only thing left to do now is to apply Knuth's bijection to the pair (P, Q) to obtain the matrix in Figure 3-d. ■

Note 5.1 We denote by $\mathfrak{M}^{(N)} = \bigcup_{m=1}^{2N-1} \mathfrak{M}_m^{(N)}$ the set of all matrices that can be obtained by applying this algorithm.

As it has been mentioned in the beginning of the section, it can be easily seen that Algorithm 5.1 is reversible. More precisely, we have the following reciprocal algorithm.

Algorithm 5.2 (Construction of a square tabloid from a $\{0,1\}$-matrix)
Input: A $\{0,1\}$-matrix $M \in \mathfrak{M}_m^{(N)}$.
Output: $T = (t_\lambda, t_\mu, r)$ — a square tabloid in $\mathcal{T}_m^{(N)}$.

- **Step 1.** Applying Knuth's bijection in the opposite direction, we can transform any matrix from $\mathcal{M}_{N\times(N+m)}$ into a pair of Young tableaux P and Q of conjugated shapes: on the alphabets $\{1, \ldots, N+m\}$ and $\{1, \ldots, N\}$ correspondingly.

- **Step 2.** The fact that M belongs to $\mathfrak{M}_m^{(N)}$ implies that both P and Q fit into (N^N), and thus we can again apply to Q the same procedure as in [13] to obtain a new Young tableau t_μ of the shape complementary to that of P.

- **Step 3.** Once again referring to the fact that M belongs to $\mathfrak{M}_m^{(N)}$, we can state that in P there is exactly one occurence of each one of $N+1, \ldots, N+m$, and that the corresponding boxes form a ribbon r numbered from bottom to top and from left to right. Moreover, this ribbon can be cut out of P leaving a Young tableau t_λ on the alphabet $\{1, \ldots, N\}$. (See Section 5.4 for conditions on $\{0,1\}$-matrices defining $\mathfrak{M}_m^{(N)}$ explicitely.)

- **Step 4.** To finalise our algorithm it is sufficient to replace all entries i in t_λ with δ_i, and in t_μ — with χ_i.

Example 5.3 To reverse Example 5.2 we start with the matrix

$$\left(\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

and we transform it into a two-row array representing the positions of 1's: in each column of the array the element in the first row is the row number, and the one in the second row — the column number of a position containing 1. We obtain therefore the following array:

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 4 \\ 7 & 9 & 2 & 5 & 8 & 1 & 6 & 2 & 4 \end{pmatrix}$$

Applying Knuth's bijection consists now in forming one Young tableau by column-bumping in the elements of the second row of the array from left to right, and placing the corresponding elements of the first row into a Young tableau of the conjugated form. This results exactly in the pair of tableaux shown in Figure 3-c.

Applying the procedure described in Example 5.2 we obtain a pair of tableaux of complimentary shapes

$$\begin{array}{|c|c|cc} \cline{1-2} 4 & 6 & & \\ \cline{1-3} 2 & 5 & \multicolumn{1}{c|}{8} & \\ \hline 1 & 2 & \multicolumn{1}{c|}{7} & \multicolumn{1}{c|}{9} \\ \hline \end{array}\ , \quad \begin{array}{|c|cc} \cline{1-1} 4 & & \\ \cline{1-1} 3 & & \\ \cline{1-2} 2 & \multicolumn{1}{c|}{3} & \\ \hline 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{4} \\ \hline \end{array}\ .$$

Finally, re-arranging and re-numbering these two accordingly we end up with the tabloid

$$\begin{array}{|c|c|c|c|} \hline \chi_4 & \chi_3 & \chi_2 & \chi_1 \\ \hline \delta_4 & \bullet & \chi_3 & \chi_1 \\ \hline \delta_2 & \bullet & \bullet & \chi_4 \\ \hline \delta_1 & \delta_2 & \bullet & \bullet \\ \hline \end{array}$$

that was the starting point of Example 5.2. ∎

We will denote the square tabloid obtained by applying Algorithm 5.2 to a matrix M by $\Phi(M)$. Note that $\Phi(M)$ is also defined on some matrices that do not belong to $\mathfrak{M}^{(N)}$, but in that case $\Phi(M) \notin \mathcal{T}^{(N)}$.

5.4 Caracteristics of matrices in $\mathfrak{M}^{(N)}$

In the previous section we presented a bijective mapping from $\mathfrak{M}^{(N)}$ to $\mathcal{T}^{(N)}$. This mapping being defined by an algorithm, we can explicitely calculate its image given an element of $\mathfrak{M}^{(N)}$. However, this set is only defined implicitely as the image of the mapping induced by the Algorithm 5.2. This section is therefore devoted to providing explicit conditions on a matrix from $\mathcal{M}_{N\times(N+m)}$ to be an element of $\mathfrak{M}_m^{(N)}$.

We will use an equivalent of Green's theorem that gives us a way of calculating the shape of the Young tableau obtained by the Robinson–Schensted correspondence from a word on the corresponding alphabet. As Robinson–Schensted correspondence is the base of Knuth's bijection, this theorem can be reformulated in terms of $\{0,1\}$-matrices to be applied to the latter.

First of all, let us introduce some notation. Let M be a $\{0,1\}$-matrix as considered above. We shall denote by $R(M,k)$ the largest possible number of 1's in M that can be arranged in k disjoint (possibly empty) sequences going North-east. Here, as in [7], we will begin each word indicating a direction with a capital letter if the sequence goes strictly in that direction, and with a small one if it does so weakly. Here, for example, "*N*orth-*e*ast" stands for "strictly North and weakly East", i.e. if two 1's in positions (i_1,j_1) and (i_2,j_2) (where $i_2 \le i_1$) belong to the same sequence then we have $i_2 < i_1$ (strictly North), and $j_1 \le j_2$ (weakly East) (see Figure 4). By convention $R(M,0)=0$. Taking $\lambda=(\lambda_1,\dots,\lambda_N)$ to be the shape of the tableau P obtained from M by Knuth

$$\underset{\text{a}}{\begin{pmatrix} 1 & 0 & \boxed{1} \\ 1 & \boxed{1} & 0 \\ 0 & \boxed{1} & 0 \end{pmatrix}} \qquad \underset{\text{b}}{\begin{pmatrix} 1 & 0 & \boxed{1} \\ \boxed{1} & \boxed{1} & 0 \\ 0 & 1 & 0 \end{pmatrix}}$$

Figure 4: The boxed sequence of 1's goes North-east on (a), but not on (b).

correspondence, we can state the following theorem:

Theorem 5.1 (Green) *Given the notation described above, one has*

$$\forall k = 1\dots N, \quad R(M,k) - R(M,k-1) = \lambda_k.$$

Now, if we denote by $(\lambda_1,\dots,\lambda_N)$ and $(r_1,\dots,r_N)$ the shapes of t_λ, and r correspondingly, where $\Phi(M)=(t_\lambda,t_\mu,r)$, and taking M' to be the left-hand $N\times N$ square part of M, we obtain automatically

$$\forall k = 1\dots N, \quad \left\{ \begin{array}{ccccc} R(M',k) & - & R(M',k-1) & = & \lambda_k \\ R(M,k) & - & R(M,k-1) & = & \lambda_k + r_k \end{array} \right. . \tag{36}$$

In other words, (36) provides us a way of calculating the shapes of t_λ and r given a $\{0,1\}$-matrix M. This immediatly delivers the first condition to be satisfied in order for M to be in $\mathfrak{M}^{(N)}$.

Condition 5.1 *Let $\Phi(M)=(t_\lambda,t_\mu,r)$ and $(\lambda_1,\dots,\lambda_N)$ and $(r_1,\dots,r_N)$ be the values provided by (36), then for r to be a correct ribbon as described by the Definition 5.1, it is necessary that*

- *there exists* $h \in [0, N]$ *such that* $r_k > 0$ *for any* $k \in [1, h]$*, and* $r_k = 0$ *when* $k > h$*;*
- $\lambda_k + r_k = \lambda_{k-1} + 1$ *for all* $k \in [2, h]$*;*
- $\lambda_1 + r_1 = N$*.*

The above condition, when fulfilled, guarantees that the shape of the ribbon is correct. It rests therefore to ensure that it's numbering is the required one, i.e. all boxes forming the ribbon must be numbered from bottom to top, and from left to right by the sequence $N + 1, \ldots, N + m$.

First of all, there has to be exactly one box in tableau P for each number between $N + 1$ and $N + m$. This is obviously guaranteed by the following condition.

Condition 5.2 *For any* $k \in [N + 1, N + m]$ *there is exactly one 1 in the* k*-th column of* M*.*

Example 5.4 Let us consider the following matrix

$$M = \left(\begin{array}{cccc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 1 \\ \boxed{1} & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ \boxed{1} & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \boxed{1} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Considering the left-hand side of the matrix we can see that

$$\begin{array}{ll} R(M', 1) = 3 & \text{(boxed sequence),} \\ R(M', 2) = 4 & \text{(boxed and circled sequences),} \\ R(M', 3) = 5 & \text{(boxed, circled, and unmarked sequences).} \end{array}$$

As there are no more 1's left we conclude that $\lambda = (3, 1, 1)$. Now if we consider the whole matrix we obtain

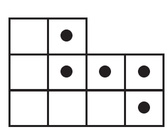
$$\begin{array}{ll} R(M, 1) = 4 & \text{(boxed sequence),} \\ R(M, 2) = 8 & \text{(boxed and circled sequences),} \\ R(M, 3) = 10 & \text{(boxed, circled, and unmarked sequences).} \end{array}$$

and therefore

$$\lambda_1 + r_1 = 4, \quad \lambda_2 + r_2 = 4, \quad \lambda_3 + r_3 = 2,$$

i.e. $r = (1, 3, 1)$. It is easy to see now that both Condition 5.1 and Condition 5.2 are verified.

We can deduce that the lower left-hand side tableau and the ribbon in the image of M will have the following shapes:



$\blacksquare$

In the following we will require some additional notions. As we have seen above, given a matrix $M \in \mathcal{M}_{N\times(N+m)}$ and provided that Condition 5.1 is satisfied, we can calculate the shape of all elements of $\Phi(M)$. Thus, in particular, we know the desired classical numbering of the ribbon. For instance the numbering of the ribbon in the Example 5.4 should be as follows:

$$
\begin{array}{|c|c|c|c|}
\cline{1-2}
 & 6 & \multicolumn{2}{c}{} \\
\hline
 & 5 & 7 & 9 \\
\hline
 & & & 8 \\
\hline
\end{array}\;.
$$

Definition 5.2

- *When we talk about the* columns *of a ribbon, we mean the sequences of numbers in each column of boxes of the ribbon in the classical numbering ((5,6), (7), and (8,9) in the example above).*

- *Two numbers $i, j \in [N+1, \ldots, N+m]$ are said to be in the* same level l*, if each one of them is exactly l boxes down from the top of its column in the classical numbering of the ribbon. We will refer as* levels *to maximal sets of numbers being in the same level. We say that level l_1 is* higher *than level l_2 if $l_1 < l_2$ — in other words if level l_1 is closer to the top.*

Example 5.5 Taking on the previous example, we can say that numbers 6, 7, and 9 form level 0 in this ribbon, and 5 and 8 — level 1. ∎

Example 5.6 More generally, if we have a ribbon numbered as follows,

$$
\begin{array}{|c|c|c|c|}
\cline{1-2}
 & 9 & \multicolumn{2}{c}{} \\
\hline
 & 6 & 7 & 8 \\
\hline
 & & & 5 \\
\hline
\end{array}\;,
$$

we shall say that its *columns in the classical numbering* are (5,6), (7), and (8,9); its *levels [in the classical numbering]* are (6,7,9) and (5,8); and its *columns in the actual numbering* are (6,9), (7), and (5,8). ∎

Our goal now is to ensure that the numbering of the ribbon, obtained by applying Algorithm 5.2, is the same as the classical one.

Condition 5.3

1. *The 1's corresponding to each column of the ribbon form a sequence going south-East.*

2. *The 1's corresponding to each level of the ribbon form a sequence going North-East.*

We can now conclude the section by stating the following theorem:

Theorem 5.2 (Caracterisation of $\mathfrak{M}_m^{(N)}$) *Let $M \in \mathcal{M}_{N\times(N+m)}$ be a $\{0,1\}$-matrix. $M \in \mathfrak{M}_m^{(N)}$ if and only if M satisfies all three conditions 5.1–5.3.*

5.4.1 Proof of the main theorem

It is obvious that $M \in \mathcal{M}_{N\times(N+m)}$ satisfies both Conditions 5.1 and 5.2 if and only if there is exactly one box in $\Phi(M)$ numbered with each one of $N+1,\ldots,N+m$ and these boxes form a correct ribbon in the sense of Definition 5.1. Thus, we only have to show that, when these two conditions are fulfilled, the Condition 5.3 is equivalent to the ribbon in $\Phi(M)$ being numbered correctly.

Recall that Algorithm 5.2 is based on Robinson-Schensted-Knuth correspondence, which has column bumping as its building block. Therefore, when applying this algorithm to M, we perform a certain number Θ of column bumpings. Thus, for each $\theta \in [0,\Theta]$, one can consider a Young tableau T_θ obtained after bumping in θ boxes.

Definition 5.3 *For $a \in [N+1, N+m]$, we shall denote by $d_\theta(a)$ the column of T_θ containing the box numbered a. We take $d_\theta(a) = 0$ if a is not in T_θ.*

Note 5.2 *There is no ambiguity in the definition of $d_\theta(a)$ due to Condition 5.2.*

Example 5.7 Consider the matrix

$$M = \begin{array}{c} \begin{array}{cccccc} \mathit{1} & \mathit{2} & \mathit{3} & \mathit{4} & \mathit{5} & \mathit{6} \end{array} \\ \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{array}\right) \end{array} .$$

To obtain the lambda part and the ribbon of $\Phi(M)$, we have to successively perform column bumping on all letters of the word 562413. This creates the following chain of Young tableaux:

$$\begin{ytableau} 5 \end{ytableau} \to \begin{ytableau} 6 \\ 5 \end{ytableau} \to \begin{ytableau} 6 \\ 2 & 5 \end{ytableau} \to \begin{ytableau} 4 & 6 \\ 2 & 5 \end{ytableau} \to \begin{ytableau} 4 & 6 \\ 1 & 2 & 5 \end{ytableau} \to \begin{ytableau} 3 & 4 & 6 \\ 1 & 2 & 5 \end{ytableau} \rightsquigarrow \begin{ytableau} 3 & \bullet & \bullet \\ 1 & 2 & \bullet \end{ytableau}$$

We have therefore

$$d_1(4) = d_2(4) = d_3(4) = 0 \quad d_4(4) = d_5(4) = 1 \quad d_6(4) = 2\,.$$

$\blacksquare$

We will show now that Condition 5.3.1 is equivalent to the following proposition: at any stage of the column bumping process, the box containing a number from any column of the ribbon in the classical numbering will be no further to the right in the tableau than any box containing another number from the same column but a lower level.

Example 5.8 Taking on the previous example, one can easily verify that, for any $0 \leq \theta \leq 6$, we have $d_\theta(5) \geq d_\theta(6)$.

Lemma 5.1 *Suppose that Conditions 5.1 and 5.2 are satisfied, and let r be the ribbon of $\Phi(M)$. Then for any $a \geq N+1$, such that a and $a+1$ belong to the same column of r in its classical numbering, the relation*

$$d_\theta(a) \geq d_\theta(a+1)$$

is invariant over $\theta \in [0, \Theta]$ such that $d_\theta(a) > 0$.

Proof – Suppose that we have $d_\theta(a) < d_\theta(a+1)$ and $d_{\theta+1}(a) \geq d_{\theta+1}(a+1)$. This means that during the $(\theta+1)^{st}$ column bumping a bumps out some c to take its place in the same column where $a+1$ is (see Figure 5-a). In this case we have $a \leq c$ (by definition of

a b

Figure 5: Snapshots of column-bumping process for Lemma 5.1.

column bumping process), and $c < a+1$ (by definition of Young tableaux), which means that $a = c$. However this is impossible because of the Condition 5.2.

On the other hand, if, at some point, we have $d_\theta(a) \geq d_\theta(a+1)$ and $d_{\theta+1}(a) < d_{\theta+1}(a+1)$, it means that a and $a+1$ are in the same column, and $a+1$ is bumped out by some c (see Figure 5-b). Again, we can notice that, by definition of column bumping, $c \leq a+1$ and $c > a$ implying that $c = a+1$, which contradicts Condition 5.2. ■

Note 5.3 *It can be easily observed that Condition 5.3.1 is equivalent to saying that, for any a as in Lemma 5.1, a is bumped in earlier than $a+1$, i.e. $d_{\theta_a}(a) > d_{\theta_a}(a+1) = 0$, where $\theta_a = \min\{\theta | d_\theta(a) > 0\}$.*

Corollary 5.1 *In the conditions of Lemma 5.1, Condition 5.3.1 implies that $d_\theta(a) \geq d_\theta(a+1)$ for all $\theta \in [0, \Theta]$.*

Proof – This corollary is a trivial consequence of Lemma 5.1 and Note 5.3. ■

Corollary 5.2 *$M \in \mathfrak{M}^{(N)}$ implies the condition 5.3.1.*

Proof – Obviously, if $M \in \mathfrak{M}^{(N)}$, we have $d_\Theta(a) = d_\Theta(a+1)$ for any a as in Lemma 5.1. As Conditions 5.1 and 5.2 also hold in this case, we can deduce from Lemma 5.1 and Note 5.3 that Condition 5.3.1 is verified. ■

We have shown therefore that Condition 5.3.1 is necessary for M to belong to $\mathfrak{M}^{(N)}$. The next step is to prove that Condition 5.3.2 is also necessary. First, we make the following remark that is similar to Note 5.3.

Note 5.4 *Condition 5.3.2 is equivalent to saying that, for any a and b ($a < b$) belonging to the same level in the ribbon, b is bumped in earlier than a, i.e. $0 = d_{\theta_b}(a) < d_{\theta_b}(b)$, where $\theta_b = \min\{\theta | d_\theta(b) > 0\}$.*

Let us now introduce some notation. We consider all columns of the ribbon in the sense of Definition 5.2 numbered from left to right. The conjugated shape of the ribbon is a sequence $(c_1, \ldots, c_N)$, where c_i is the number of boxes in the ribbon's i-th column. In the following discussion we will only consider those columns i that have $c_i > 0$.

For each column i we will denote by t_i its top box in the classical numbering. In the example 5.4 the conjugated shape of the ribbon is (2,1,2), and $t_1 = 6$, $t_2 = 7$, and $t_3 = 9$.

Lemma 5.2 *Suppose that Conditions 5.1, 5.2, and 5.3.1 are satisfied, and let r be the ribbon of $\Phi(M)$. Suppose also that for any $t_{i-1} - k$, and $t_i - k$ — two numbers in the adjacent columns and the same level of r in its classical numbering — we have*

$$d_{\theta+1}(t_{i-1} - k) < d_{\theta+1}(t_i - k)\,,$$

then the same is correct if we replace $\theta + 1$ by θ.

Proof – Suppose this is not the case. Then we have at the same time

$$\begin{array}{c} d_\theta(t_{i-1} - k) \geq d_\theta(t_i - k) \\ d_{\theta+1}(t_{i-1} - k) < d_{\theta+1}(t_i - k) \end{array}\,,$$

which means that $t_i - k$ and $t_{i-1} - k$ are in the same column of T_θ, and during the $(\theta+1)$-st column bumping some t bumps $t_i - k$ out (see Figure 6). From the definition of Young tableaux and column bumping we can conclude then that $t \in [t_{i-1} - k + 1, \ldots, t_i - k - 1]$. However, $t \in [t_{i-1} + 1, \ldots, t_i - k - 1]$ implies that t is in the column i of the final ribbon, and

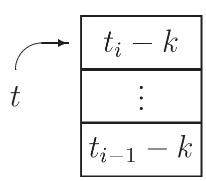


Figure 6: Snapshot of column-bumping process for Lemma 5.2.

thus, by Condition 5.3.1 and Lemma 5.1, we have $d_\theta(t) \geq d_\theta(t_i - k)$, which is impossible, as t bumps out $t_i - k$ at step $\theta + 1$. Thus we can deduce that

$$t \in [t_{i-1} - k + 1, \ldots, t_{i-1}]\ . \tag{37}$$

Notice here that, for $k = 0$, this interval is empty and we obtain a contradiction. We can therefore continue our proof inductively.

Let us suppose that we have proven the assertion of the lemma for all levels higher than k. To prove it for k notice that (37) implies that t is in the column $i - 1$ of the classical numbering of the ribbon, and thus $t = t_{i-1} - k'$, where $0 \leq k' < k$. By the assumption of the induction, we then have

$$d_\theta(t_i - k) - 1 = d_\theta(t) = d_\theta(t_{i-1} - k') < d_\theta(t_i - k') \leq d_\theta(t_i - k) \ , \tag{38}$$

and therefore

$$d_{\theta+1}(t) = d_{\theta+1}(t_{i-1} - k') < d_{\theta+1}(t_i - k') = d_\theta(t_i - k') = d_\theta(t_i - k) = d_{\theta+1}(t) \ . \tag{39}$$

The inequality in (39) is the assumption of the lemma, the first equality is trivial, the second is due to $(t_i - k')$'s not changing its position during $(\theta + 1)$-st column bumping, the third is a consequence of (38), and the last represents the fact that t bumps out $t_i - k$. The relation (39) being contradictory proves the lemma. ■

Let us now generalise Lemma 5.2 to any pair of numbers in the same level of the ribbon.

Lemma 5.3 *Suppose that Conditions 5.1, 5.2, and 5.3.1 are satisfied, and let r be the ribbon of $\Phi(M)$. Suppose also that for any $i < j$ and any $t_i - k$, and $t_j - k$ — two numbers in the same level of r — we have*

$$d_{\theta+1}(t_i - k) < d_{\theta+1}(t_j - k) \, ,$$

then the same is correct if we replace $\theta + 1$ by θ.

Proof – First of all, notice that without loss of generality we can assume that $t_j - k$ and $t_i - k$ are adjacent in the level k of r, i.e. for any h, such that $i < h < j$, there is no box in the level k of the column h of r. Observe also that Lemma 5.2 is a special case of this one with $j = i + 1$, and therefore we only have to discuss the case were $j - i > 1$.

As before we suppose that we have proven the assertion of the lemma for all levels higher than k. We shall suppose that it is not satisfied for level k, and obtain a contradiction. Indeed, using the same reasoning as for (37), we can show that

$$t \in [t_i - k + 1, \ldots, t_{j-1}] \, , \tag{40}$$

which implies that t belongs to some column i' where $i \leq i' < j$, and therefore $t = t_{i'} - k'$ with some $k' \ \in \ [0, \ldots, k - 1]$. Then we have the following chain of relations:

$$d_\theta(t_j - k) - 1 = d_\theta(t) = d_\theta(t_{i'} - k') < d_\theta(t_j - k') \leq d_\theta(t_j - k) \tag{41}$$

that is analogous to (38). We obtain a contradiction in the same manner as in (39), which proves the lemma. ■

Corollary 5.3 *$M \in \mathfrak{M}^{(N)}$ implies the condition 5.3.2.*

Proof – As we have seen before, $M \in \mathfrak{M}^{(N)}$ implies both Conditions 5.1 and 5.2, and by Corollary 5.2 also the Condition 5.3.1. It can also be easily observed that the assumption of Lemma 5.3 holds for $\theta + 1 = \Theta$, and therefore applying this lemma inductively, we prove the validity of Condition 5.3.2 (cf. Note 5.4). ■

Collecting together Corollaries 5.2 and 5.3, we obtain the following proposition.

Proposition 5.2 *$M \in \mathfrak{M}^{(N)}$ implies all the conditions 5.1–5.3.*

5.4.2 Plactic equivalence

To show that Conditions 5.1–5.3 are sufficient for $M \in \mathfrak{M}^{(N)}$, we will have to introduce the notion of *plactic equivalence*. This equivalence has been studied extensively and presented in [12], [13], and [15], therefore we will only briefly present it here.

The column insertion process can be described algebraically by the plactic formalism developed by Lascoux and Schützenberger (cf. [15]) that we will now present. Let A be a totally ordered alphabet. The *plactic monoid* is the monoid constructed over A and subject to the following relations (discovered by Knuth (cf. [12])):

$$\begin{cases} acb \equiv cab\,, & \text{for every } a \leq b < c \text{ in } A \\ bca \equiv bac\,, & \text{for every } a < b \leq c \text{ in } A\,. \end{cases}$$

Two words over A are identified under the plactic relations if and only if the Young tableaux obtained by applying the column insertion process to their mirror images are equal (cf. [7, 15]).

A word $w(T)$ can be associated with a Young tableau T over an alphabet A, by reading the columns of T from top to bottom and from left to right. The words associated with Young tableaux in such a way are called *tableau words*.

Example 5.9 The tableau word associatied with the Young tableau

$$\mathrm{T} = \begin{array}{|c|c|cc} \hline a_3 & a_5 & & \\ \hline a_2 & a_2 & & \\ \hline a_1 & a_1 & \multicolumn{1}{c|}{a_1} & \multicolumn{1}{c|}{a_4} \\ \hline \end{array} \qquad \text{is } w(T) = a_3\, a_2\, a_1\, a_5\, a_2\, a_1\, a_1\, a_4\,.$$

■

Note that applying the column insertion to the mirror image of a tableau word $w(T)$ yields the tableau T. Observe also (see [7, 15]) that a word over A is equivalent with respect to the plactic relations to a unique tableau word (which is therefore associated with the Young tableau given by the column insertion process applied to the mirror image of w). In other words, we can state the following proposition.

Proposition 5.3 *If u is a word over a totally ordered alphabet A, and $\overline{u}$ is its mirror image, then*

$$w(T_{\overline{u}}) \equiv u\,,$$

where $T_{\overline{u}}$ is the Young tableau obtained by applying column insertion to $\overline{u}$.

Note 5.5 *If u and v are two words over A such that $u \equiv v$ then by the previous proposition we have*

$$w(T_{\overline{u}}) \equiv u \equiv v \equiv w(T_{\overline{v}})\,,$$

and therefore $T_{\overline{u}} = T_{\overline{v}}$.

Denoting by $R(w, k)$ the maximum total length of k increasing subsequences of w, we can state the following theorem, which is a consequence of Green's theorem (see [8]).

Theorem 5.3 *If u and v are two words over a totally ordered alphabet A, such that $u \equiv v$, then for any $k \geq 0$ we have*

$$R(u,k) = R(v,k)\,.$$

Finally, we shall mention another important property of plactic equivalence that we will use in the sequel. Let $w = w_1\,w_2\,\ldots w_n$ be a word over a totally ordered alphabet A, and let $a < b$ be two letters of A. We shall denote by $w_{a,b}$ the subword of w consisting of all its letters w_i such that $a \leq w_i \leq b$. In other words, $w_{a,b}$ is the projection of w on the interval $[a,b] \subset A$.

Theorem 5.4 *Let u and v be two words over a totally ordered alphabet A. Let $a < b$ be two letters of A. Then $u \equiv v$ implies $u_{a,b} \equiv v_{a,b}$.*

5.4.3 Proof of the main theorem (continued)

Let us again consider a matrix $M \in \mathcal{M}_{N\times(N+m)}$, and $\Phi(M) = (t_\lambda, t_\mu, r)$ — the square tabloid obtained by applying the algorithm 5.2 to M.

If we take u to be the word obtained by reading the positions (columns) of 1's in M, then keeping the notations of section 5.4.2 we have by definition of Robinson–Schensted–Knuth correspondence

$$T_u = t_\lambda \cup r\,,$$

which implies by Proposition 5.3 that

$$\overline{u} \equiv w(T_u) = w(t_\lambda \cup r)\,.$$

Applying now the theorem 5.4 to this equivalence and the interval $[N+1, N+m]$ we obtain

$$\overline{u''} \equiv w(r)\,,$$

where u'' is the word obtained by reading the positions of 1's in M'' — the right-hand part of M —, and $w(r)$ is the the restriction of the tableau word corresponding to $t_\lambda \cup r$ to the ribbon r. Therefore, by Theorem 5.3, we have for any $k \geq 0$

$$R(\overline{u''},k) = R(w(r),k) \tag{42}$$

Observe that the increasing subsequences in $\overline{u''}$ correspond exactly to the sequences of 1's going North-East in M'', while the decreasing ones correspond to the sequences going south-East. Condition 5.3.1 implies therefore that no two 1's corresponding to boxes of the same column of r in the classical numbering can be part of the same increasing subsequence. Thus the increasing subsequences can only have one 1 per column of r in the classical numbering, which means that they cannot be longer than the subsequences corresponding to levels of r in the classical numbering. More generally, we can conclude that we obtain $R(\overline{u''},k)$ considering the subsequences of $\overline{u''}$ corresponding to the k top levels of r.

On the other hand, Corollary 5.1 implies that an increasing subsequence in $w(r)$ can only have one number per column of r in the classical numbering. At the same time it follows trivially from the definition of a Young tableau that such a sequence can only have one number per column of r in its actual numbering (the one constructed by the algorithm 5.2).

Let us now substitute $k = \min\{c_i | c_i > 0\}$ in (42). In this case k is equal to the number of levels of r of maximum length. Combining the two observations above, we can deduce that in $w(r)$, there are exactly k subsequences each having one letter per column of r in its classical numbering as well as in its actual numbering. Thus, for any i such that $c_i = k$, we can deduce that all numbers $t_i - c_i + 1, \ldots, t_i$ are in the column i of r in its actual numbering, in other words, column i is numbered correctly.

Repeating the same argument for subsequent levels of r we conclude that all columns of r are numbered correctly, which proves that $M \in \mathfrak{M}^{(N)}$, and thus finalises the proof of the theorem 5.2. ■

Acknowledgements

The authors would like to thank the anonymous referees for their comments that have considerably influenced the presentation of our paper bringing forward its combinatorial aspects.

References

[1] G.E. Andrews, *The Theory of Partitions*, Addison Wesley, 1974.

[2] M.J. Barrett, *Error probability for optimal and suboptimal quadratic receivers in rapid Rayleigh fading channels*, IEEE J. Select. Areas in Commun., **5**, Nr.2, 302–304, February 1987.

[3] S. Bliudze, D. Krob, *Performance evaluation of demodulation with diversity — A combinatorial approach III: Threshold analysis*, to appear in the Proceedings of ANALCO, 2005.

[4] L. Comtet, *Analyse combinatoire*, Presses Universitaires de France, 1970.

[5] J.L. Dornstetter, D. Krob, J.Y. Thibon, *Fast and Stable Computation of Error Probability in Rapid Rayleigh Fading Channels*, Proceedings of AlgoTel, 2000.

[6] J.L. Dornstetter, D. Krob, J.Y. Thibon, E.A. Vassilieva, *Performance evaluation of demodulation with diversity — A combinatorial approach I: Symmetric function theoretical methods*, Discrete Mathematics and Theoretical Computer Science, **5**, pp. 191–204, 2002

[7] W. Fulton, *Young Tableaux*, Cambridge University Press, 1997.

[8] C. Greene, *Some partitions associated with a partially ordered set*, J. of Combin. Theory, Ser. A, **20**, 69–79, 1976.

[9] M. Guenach, L. Vandendorpe, *Downlink Performance Analysis of a BPSK-Based WCDMA Using Conventional Rake Receivers With Channel Estimation*, IEEE J. on Select. Areas in Commun., **19**, Nr.11, 2165–2176, November 2001.

[10] L. Hanzo, T.H. Liew, B.L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*, John Wiley & Sons Inc., 2002.

[11] J.P. Imhof, *Computing the distribution of quadratic forms in normal variables*, Biometrika, **48**, 419–426, 1961.

[12] D.E. Knuth, *Permutation, matrices and generalized Young tableaux*, Pacific J. Math., **34**, 709–727, 1970.

[13] D. Krob, E.A. Vassilieva, *Performance analysis of demodulation with diversity – A combinatorial approach II: Bijective methods*, Discrete Applied Mathematics, **145**, (3), 403-421, 2005.

[14] A. Lascoux, *Inversion des matrices de Henkel*, Lin. Alg. and its Appl., **129**, 77–102, 1990

[15] A. Lascoux, M.P. Schützenberger, *Le monoïde plaxique*, Quaderni de la Ricerca Scientifica, A. De Luca Ed., **109**, 129–156, 1981.

[16] A. Lascoux, M.P. Schützenberger, *Formulaire raisonné de fonctions symétriques*, Public. LITP, Paris 7, 1985.

[17] I.G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd Edition, Clarendon Press, 1993.

[18] J. Proakis, *Digital Communications*, 3rd Edition, McGraw-Hill, 1995.

[19] J. Remmel, *The combinatorics of (k, l)-hook Schur functions*, Contemp. Math., **34**, Amer. Math. Soc., Providence, RI, 1984.

[20] J. Remmel, *A bijective proof of a factorization theorem for (k, l)-hook Schur functions*, Linear and Multilinear Algebra, **28**, 119–154, 1990.

[21] J.Y. Thibon, *Hopf algebras of symmetric functions and tensor products of symmetric group representations*, Int. J. of Alg. and Comput., **1**, (2), 207–221, 1991.

[22] G.L. Turin, *The characteristic function of Hermitian quadratic forms in complex normal variables*, Biometrika, **47**, 199–201, 1960.