Pfaffian Systems of A-Hypergeometric Systems II  
— Holonomic Gradient Method

K.Ohara, N.Takayama

May 9, 2015

1 Introduction

Let $A$ be a $d \times n$ matrix with entries in $\mathbb{N}_0 = \{0, 1, 2, \ldots \}$. The $i$-th column of the matrix $A$ is denoted by $a_i$. For a given $\beta \in \mathbb{N}_0 A = \{\sum_{i=1}^{n} v_i a_i | v_i \in \mathbb{N}_0\}$, we consider the set of points

$$F(\beta) = \{u \in \mathbb{N}_0^d | Au = \beta\}.$$ 

We are interested in the probability distribution $\frac{1}{u!} p^u / Z(\beta; p)$ on $F$ where $u \in F$, $p \in \mathbb{R}^n$ and $Z$ is the normalizing constant defined by

$$Z(\beta; p) = \sum_{u \in F} \frac{1}{u!} p^u, \quad u! = \prod_{k=1}^{n} u_k!.$$ 

When $a_i$ lies on a hyperplane $h(y) = 1$, the probability distribution can be regarded as a conditional probability distribution induced from the multinomial distribution $\frac{1}{u!} p^u, \sum_{k=1}^{n} p_k = 1$. We will call this probability distribution $A$-distribution following the celebrated work by P.Diaconis and B.Sturmfels on constructions of Markov bases of $F$ in terms of Gröbner bases of the affine toric ideal defined by the matrix $A$ [2], [5]. The class of $A$-distributions includes hypergeometric distributions on contingency tables with a fixed marginal sums.

We will discuss on an exact evaluation of the normalizing constant $Z$ and its derivatives in this paper. The normalizing constant $Z$ satisfies the $A$-hypergeometric system, which is substantially introduced by Gelfand, Kapranov, Zelevinsky in the late 1980’s (see, e.g., [15]), and contiguity relations (recurrence relations) such as $\frac{\partial Z}{\partial p_i}(\beta; p) = Z(\beta - a_i, p)$. By the system and contiguity relations, an discrete analogy of the holonomic gradient method (HGM) introduced in [11] will give an efficient and exact method to evaluate the normalizing constant and its derivatives. For the matrix $A$ standing for the $2 \times n$ contingency tables, the maximal likelihood estimate (MLE) by the HGM is discussed in the recent thesis by M.Ogawa [13]. The MLE needs evaluations of the normalizing constant and its derivatives, which is done by contiguity relations for $2 \times n$ contingency tables given in [3].
In this paper, we will discuss on the discrete HGM for a general matrix $A$. The key step for it is a construction of the Pfaffian system for the $A$-hypergeometric system. We utilize the result of [6], Macaulay type matrices [10], and the Hilbert driven Buchberger algorithm [17] for an efficient construction. Timing data are given to compare several methods.

2 Macaulay Type Matrices

The discussion of this section is well-known, but we do not find a relevant literature for our application. Then, we give a brief explanation on Macaulay type matrices (which has an origin in [10, p.8]). Let $P = K[x_1, \ldots, x_n]$ be the ring of polynomials of $n$ variables where $K$ is a coefficient field. We suppose that a zero-dimensional ideal $I$ of $P$ is generated by polynomials $f_1, \ldots, f_m$. Let $g$ be an element of $I$. Then, $g$ can be written as

$$g = \sum_j \left( \sum_i c_{ij} t_{ij} \right) f_j = \sum_{ij} c_{ij} t_{ij} f_j$$

(1)

where $c_{ij} \in K$ and $t_{ij}$ is a monomial. Let $S$ be a basis, consisting monomials, of $P/I$ as a vector space over $K$. $S$ is a finite set. We assume that a base $S$ of $P/I$ is given. In fact, in our application to $A$-hypergeometric systems, the base $S$ is determined by an algorithmic method in [6] without computing a Gröbner basis for a given $A$-hypergeometric ideal. We define $M$ the union with respect to $i$’s and $j$’s of the set of the monomials which appear in $t_{ij} f_j$. Set $M = M \setminus S$. Let us construct a matrix $F$ of which columns are indexed by $M \cup S$ and whose each row stands for the coefficients of $t_{ij} f_j$. We denote the row consisting of the coefficients of $t_{ij} f_j$ by $C(t_{ij} f_j)$. For $g$ in [11], $C(g)$ is the row vector of the coefficients of the polynomial $g$, which is simplified, indexed by $M \cup S$. It follows from the relation [11] that the vector $C(g)$ can be written as a linear combination of the rows of the matrix $F$.

Let us present an example to illustrate our definitions. Put $f_1 = x_1^2 + x_2^2 - 4$, $f_2 = x_1 x_2 - 1$. Let $<$ be the lexicographic order such that $x_1 > x_2$. Then, the set of the standard monomials $S$ is $\{x_2^2, x_2 x_1, 1\}$. Put $g = x_2 f_1 - x_1 f_2 = x_1 + x_2^2 - 4 x_2$. Then, $M = \{x_1^2 x_2, x_1\}$. We have

$$M \cup S \quad x_1^2 x_2 \quad x_1 \quad x_2^3 \quad x_2^2 \quad x_2 \quad 1$$

$$C(x_2 f_1) \quad \begin{pmatrix} 1 & 0 & 1 & 0 & -4 & 0 \end{pmatrix}$$

$$C(x_1 f_2) \quad \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$C(g) \quad \begin{pmatrix} 0 & 1 & 1 & 0 & -4 & 0 \end{pmatrix}$$

and

$$M \cup S \quad x_1^2 x_2 \quad x_1 \quad x_2^3 \quad x_2^2 \quad x_2 \quad 1$$

$$F = C(x_2 f_1) \quad \begin{pmatrix} 1 & 0 & 1 & 0 & -4 & 0 \end{pmatrix}$$

$$C(x_1 f_2) \quad \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We suppose that the support of $g$, which is an element of $I$ expressed as [11], is $S$ and a monomial $t$ in $M$. We suppose that the coefficient of $t$ in $g$
is normalized to be 1. The monomial $t$ can be uniquely written as a linear combination of the elements of $S$ over $K$ modulo $I$ as $t = \sum_{s \in S} c_s s \mod I$, $c_s \in K$. Then we have $g = t - \sum_{s \in S} c_s s \in I$. Since $C(g)$ is in the row span of $F$, we can obtain $C(g)$ by constructing reduced row echelon form of the matrix $F$ for $C(t_{ij}f_j)$’s with the index order $M > S$. In fact, let $e$ be the number of echelons (the number of the leading terms of the reduced Gröbner basis of the system of linear equations defined by the matrix $F$). Then, the rank of $F$ is equal to $e$. If $C(g)$ is not contained in the reduced row echelon form, then the rank of \( \begin{pmatrix} F \\ C(g) \end{pmatrix} \) is $e + 1$. It contradicts to the assumption that $C(g)$ is contained in the row span of $F$.

For polynomials $f_1, \ldots, f_m$, a positive number $T$, and a basis $S$ of $P/I$ consisting of monomials, we construct a matrix by row vectors of the form $C(t_{fi})$, $i = 1, \ldots, m$ where $t$ runs over the set of the monomials of which degrees are less than or equal to $T$. We will call the matrix the Macaulay type matrix of degree $T$ for $f_1, \ldots, f_m$ and $S$. We denote the matrix by $F_T$.

### 3 Macaulay Type Matrices to Derive Pfaffian Systems

The method finding $C(g)$ by computing the reduced row echelon form from the Macaulay type matrix has no advantage to Gröbner basis methods without new ideas as in the $F_4$ algorithm and its related algorithms, however, as we will see in this paper, this method has an advantage for the numerical analysis of $A$-hypergeometric systems in the ring of differential operators or in the ring of difference-differential operators without them. Consider the ring of differential operators with rational function coefficients $R' = K'(\partial_1, \ldots, \partial_n)$ where $K' = C(x_1, \ldots, x_n)$, which is abbreviated as $C(x)$. Let $I$ be a zero-dimensional left ideal in $R'$, $S$ a basis of $R/I$ over $K'$. Let $s$ be an element of $S$. Then, $\partial_is$ is uniquely written as the linear combination of $S$ over $K'$ modulo $I$ as $\partial_is = \sum_{t \in S} c_i t \mod I$, $c_i \in K$. Let us denote by $g$ the expression $\partial_is - \sum_{t \in S} c_it$. We call $g$ the Pfaffian operator for $\partial_i$, $s$ and $S$. We want to find all Pfaffian operators for all pairs of $\partial_i$, $i = 1, \ldots, n$ and $s \in S$. We utilize Macaulay type matrices explained in the section 2 which can be defined analogously in $R'$.

Let $F = F_T$ be the Macaulay type matrix constructed from $t_{ij}$’s whose total degrees are less than or equal to $T$ and $f_j$’s which are generators of $I$. When the total degree $T$ is sufficiently large, then the ideal element $g$, of which support is in $S \cup \{\partial_i s\}$, can be obtained by computing the reduced row echelon form of $F$ with an index order such that $M > S$ and $\partial_is$ is the smallest index in $M$. We note that the linear algebra elimination over $K'$ to obtain the reduced row echelon form does not perform differentiations of elements of $K' = C(x)$ and then the numerical coefficients of $g$ when the variable $x$ is specialized to a generic constant vector $X$ can be obtained by specializing the matrix $F$ to $x = X$ and by the linear algebra elimination. This will be the key idea to apply Macaulay type matrices to evaluate $A$-hypergeometric polynomials, in which we
avoid a Gröbner basis computation.

Let $A$ be a $d \times n$ matrix. The $i$-th column of $A$ is denoted by $a_i$ and the $(i,j)$ element of $A$ is denoted by $a_{ij}$. We assume that $a_i \in \mathbb{N}^n_0$ and the rank of $A$ is $d$. Let us consider Macaulay type matrices when $I$ is the $A$-hypergeometric ideal $H_A(c)$. We regard the parameters $c_i$’s as indeterminates, and we put

$$K = \mathbb{C}(c_1, \ldots, c_d, x_1, \ldots, x_n) \quad \text{and} \quad R = K\langle \partial_1, \ldots, \partial_n \rangle.$$ 

Generators of $H_A(c)$ consist of the Euler operators $E_j - c_j = \sum_{k=1}^n a_{jk}x_k\partial_k - c_j$, $j = 1, \ldots, d$ and the toric ideal $I_A = \langle \partial^n - \partial^v \mid Au = Av, u, v \in \mathbb{N}^n_0 \rangle$. It is important to distinguish the left ideal $RH_A(c)$ and the left ideal $R'H_A(\beta)$, $\beta \in \mathbb{C}^d$. The former ideal lies in $R$ where $c_i$’s are indeterminates and the latter ideal lies in $R'$. We consider a Macaulay type matrix for $H_A(c)$. We will construct a smaller matrix $F'$ than the matrix $F$ to obtain Pfaffian operators in the following way.

For a given term order $<$, the reduced Gröbner basis of $I_A$ with respect to the order $<$ is denoted by $G$. A basis of $R/RH_A(c)$ can be found by the algorithm in [6] from $G$. We denote it by $S$. Take a monomial $s$ from the set $S$.

**Algorithm 1**

Input: a matrix $A$ which defines an $A$-hypergeometric system and a sufficiently large integer $T$.

Output: A matrix $F'_T$ with entries in $K$, which will be called the **Macaulay type matrix of degree $T$ of the $A$-hypergeometric system**.

1. Compute a Gröbner basis $G$ of $I_A$.
2. Find a basis $S$ by the algorithm in [6]. (A probabilistic version of this method in [7] can be used and is more efficient.)
3. Let $N_T$ be the set of the monomials in $\partial_i$’s whose total order is less than or equal to $T$.
4. Multiply an element $t$ of $N_T$ to an Euler operators $E_j - c_j$ in the ring of differential operators $R$. Reduce the element $t(E_j - c_j)$ by $G$ and obtain the remainder. By the remainders from all the pairs from $N_T$ and the Euler operators, we construct the matrix $F'$ as in the section 2 and set it $F'_T$.

**Theorem 1**

1. When $T$ is sufficiently large, the reduced row echelon form of $F' = F'_T$ contains the Pfaffian operator $\partial_i s - \sum_{t \in S} c_t t$ if $\partial_i s$ is irreducible by $G$.
2. Let $f$ be a solution of $H_A(c)$. The numerical value $(\partial_i s) \cdot f$ at a generic point $x = X \in \mathbb{Q}^n$, $c = \beta \in \mathbb{Q}^d$, can be obtained from the numerical values of $s \cdot f$, $s \in S$ at a point $x = X, c = \beta$ by computing the reduced row echelon form of the numerical matrix $F'|_{x = X, c = \beta}$.
We note that when $\partial_is$ is reducible by $G$, it is equal to an irreducible monomial \( \partial^u \) modulo $G$. When $T$ is sufficiently large $\partial^u - \sum_{t \in S} c_t$ will be in the reduced row echelon form of $F'$. Then, we can obtain the Pfaffian operator $\partial_is - \sum_{t \in S} c_t$ from the reduced row echelon form.

**Proof.** 1. The set $S$ is a basis of $R/RH_A(c)$ over $K = \mathbb{C}(c_1, \ldots, c_d, x_1, \ldots, x_n)$. We multiply $N_T$ from the left to the generators $E_j - c_j$ and $\partial^u - \partial^v$ of $H_A(c)$.

$M$ is the set of the monomials in the obtained elements of $R$ from $N_T \cdot (E_j - c_j)$ and $N_T \cdot G$. Put $M = \tilde{M} \setminus S$ as in the section (2). The set $G$ is a Gröbner basis of the toric ideal $I_A$. We divide $M$ into two disjoint sets $M_r$ and its complement $M_r'$ where the set $M_r$ is the set of reducible monomials by $G$. We define the set $M_i$ the union of $M_r'$ and the set of irreducible monomials obtained by reducing elements of $M_r$ by $G$, which is denoted by $\text{Red}(M_r)$, minus the set $S$. Thus, for $\partial^u \in M_r$, there exists $\partial^v \in M_i \cup S$ such that

$$\partial^u - \partial^v \in I_A.$$  

(2)

The subset of $M_i$, of which element appears in the column index of $F'$, is denoted by $M'$. The set $M \cup (\text{Red}(M_r) \setminus S)$ consists of the disjoint sets $M_r$, $M_i \setminus M'$, $M'$. Construct the Macaulay type matrix $F$ from these binomials (2) and the obtained elements by the multiplication of $N_T$ to the generators of $H_A(c)$. We sort the columns of the matrix by the order $M_r > (M_i \setminus M') > M' > S$. When $T$ is sufficiently large, the reduced row echelon form of the Macaulay type matrix $F$ contains the Pfaffian operator $\partial_is - \sum_{t \in S} c_t$ as we have seen in the previous section. The reduced row echelon form of the smaller matrix $F'$ is contained in the reduced row echelon form of $F$. Since $\partial_is \in M_i \cap M'$ for sufficiently large $T$, the Pfaffian operator is contained in the reduced row echelon form of $F'$.

2. The steps to find the reduced row echelon form do not have steps of multiplication in the ring of differential operators $R$. Then, a specialization of $x$ and $c$ to numbers and constructing the reduced row echelon form commute. Q.E.D.

**Example 1** Consider the $A$-hypergeometric ideal generated by

$$x_1\partial_1 + x_2\partial_2 + x_3\partial_3 + x_4\partial_4 - c_1, x_2\partial_2 + x_4\partial_4 - c_2, x_3\partial_3 + x_4\partial_4 - c_3, \partial_2\partial_3 - \partial_1\partial_4.$$  

For the reverse lexicographic order such that $\partial_1 > \partial_2 > \partial_3 > \partial_4$, the basis by the algorithm (4) is $S = \{1, \partial_4\}$. Put the degree $T = 1$. We multiply the monomials in $N_T = \{1, \partial_1, \partial_2, \partial_3, \partial_4\}$ to the generators. The table below is the result of the multiplication where the index $i_1i_2\cdots i_m$ in the top row stands for
the monomial $\prod_{k=1}^{m} \partial_{h}$ and the index 0 denotes the monomial 1.

We put

$$M_t = \{23, 114, 123, 124, 134, 144, 223, 233, 234\}$$

$$M' = \{1, 2, 3, 11, 12, 13, 14, 22, 24, 33, 34, 44\}.$$  

Note that $M_t = M_r \cup (M_i \setminus M')$ in the proof. The join of the two tables is the matrix $F$ and $M = M_t \cup M'$. Let us use the order of indices such that $M_t > M' > S$. Apply the Gaussian elimination with this order to the first table. We eliminate elements of the column standing for the index 23 by the last row and then remove the last row. Then, we obtain the following matrix which agrees with the matrix $F_T'$, $(T = 1)$ in the algorithm.

We can see, by a calculation, that this matrix can be transformed into the reduced row echelon form whose rank is 12. The reduced row echelon form contains the Pfaffian operators.

## 4 Evaluation of $A$-Hypergeometric Polynomials by Macaulay type matrices

Let $R' = K'(\partial_1, \ldots, \partial_n)$, $K' = Q(x_1, \ldots, x_n)$ be the ring of differential operators with rational function coefficients. Let $I$ be a zero dimensional left ideal of $R'$.
and $S = \{s_1, \ldots, s_r\}$ a basis of $R/I$ as a vector space over $K'$. We assume that the set $S$ consists of monomials of $\partial$. If no confusion arises, the column vector $(s_1, \ldots, s_r)^T$ is also denoted by $S$. There exists a matrix $P_i$ with entries in $K'$ such that $\partial_i S - P_i S = 0$ modulo $I$. The system of differential equations $\partial_i Y - P_i Y = 0$, $i = 1, \ldots, r$ where $Y$ is a column vector of unknown functions is called a Pfaffian system. A basis $S$ and the matrix $P_i$ can be obtained by computing a Gröbner basis of $I$ in the ring of differential operators $R'$ (see, e.g., [5, Chapter 6]). Let $H_A(\beta)$ be the $A$-hypergeometric ideal. Let $Z(\beta; x)$ be an $A$-hypergeometric series for the matrix $A$ and a generic parameter vector $\beta$. It is well known that a contiguity relation

$$\partial_i \bullet Z(\beta; x) = Z(\beta - a_i; x)$$

holds under a suitable normalization of the hypergeometric series, see, e.g., [11, p.x]. In particular, the relation holds when $Z$ is the $A$-hypergeometric polynomial

$$\sum_{\alpha u = \beta, \alpha \in \mathbb{N}_0^n} \frac{1}{u!} x^u, \quad u! = \prod_{i=1}^n u_i!; x^u = \prod_{i=1}^n x_i^{u_i}$$

for $\beta \in \mathbb{N}_0 A = \sum_{i=1}^n \mathbb{N}_0 a_i$. It follows from the contiguity relation that we also have a contiguity relation for derivatives of $Z$

$$\partial_i \bullet (s_j \bullet Z)(\beta; x) = (s_j \bullet Z)(\beta - a_i; x)$$

We define the column vector $Y$ by $(s \bullet Z | s \in S)$. When $S$ is a basis for both $R'/R' H_A(\beta)$ and $R'/R' H_A(\beta - a_i)$, it follows from the contiguity relation and the Pfaffian system, we have the following identity

$$Y(\beta - a_i; x) = \partial_i \bullet Y(\beta; x) = P_i(\beta; x) Y(\beta; x),$$

which is called a recurrence relation or a difference Pfaffian system for the vector valued function $Y$ to the direction $a_i$.

We suppose that the value $Y(c, x)$ at $c = \beta \in \mathbb{N}_0 A$ and $x = X \in \mathbb{Q}^n$ is given. Once we construct the difference Pfaffian system, the value of $Y$ at $c = \beta + a_i$ and $x = X$ is obtained by

$$P_i(\beta + a_i, x)^{-1} Y(\beta, X)$$

when the inverse matrix of $P_i$ exists. When $c = 0$, the hypergeometric polynomial $Z(0; x)$ is equal to 1. Then, the vector valued function $Y(0; x)$ is $(1, 0, \ldots, 0)^T$. We apply the recurrence relation [5] from $Y(0; x)$ iteratively for $i$'s. Then we can obtain the exact value of $Y(\sum_{m=1}^n b_i a_i)$, $h_i \in \mathbb{N}_0$ as long as the inverse matrices of $P_i$’s exist and $S$ is a basis for all $R'/R' H_A(\beta)$. This method to evaluate hypergeometric polynomials is called (difference) holonomic gradient method ((difference) HGM) as an analogy of evaluating normalizing constants for unnormalized probability distributions by utilizing holonomic systems of differential equations [11, 4].
Finding a recurrence relation of $Z$ for indeterminates $x_1, \ldots, x_n$ and $c_1, \ldots, c_d$ is possible by deriving a Pfaffian system by a Gröbner basis, but it requires huge computer resources. When we want to evaluate $Y(c; x)$ at a vector of rational numbers $x = X$ on a line $c = \beta + kH$ parametrized with a parameter $k$ and a direction vector $H$ from the value $Y(\beta; X)$, we only need the matrix $P_i$, restricted on $x = X$ and $c = \beta + kH$. In order find such restricted $P_i$, we can apply the method of Macaulay type matrix and avoid Gröbner basis computation for indeterminates $x$ and $c$. This idea is illustrated in the Figure 1.

### Figure 1: Restriction to $c = \beta + kH$ and $x = X$

| Gröbner basis method | Pfaffian system $P$ of $I$ |
|----------------------|---------------------------|
| Method of Macaulay type matrices | Specialization of $c$ to $\beta + kH$, $x$ to $X$ |
| Restriction of $P$ on a line $c = \beta + kH$ and $x = X$ |

#### Algorithm 2 (Method of Macaulay type matrices for $A$-hypergeometric systems)

**Input:** a matrix $A$, a starting point of the parameter vector $\beta \in \mathbb{N}_0^d$, a vector of numbers $X \in \mathbb{Q}^n$, a basis $S$ of $R_n/R_nH_A(c)$ consisting of monomials of $\partial$, a direction $H \in \mathbb{N}_0^d$ such that $H = \sum_{i=1}^n a_i$.

**Output:** the recurrence relation

$$Y(k-1) = R(k)Y(k)$$

with respect to $k$ for the vector valued hypergeometric functions $Y(k) = (S \cdot Z)(\beta + kH; X)$.

1. Find $h_i \in \mathbb{N}_0$ such that $H = \sum_{i=1}^n h_i a_i$ (express the direction $H$ in terms of $a_i$). This step can be performed by solving the integer program problem of minimizing $w \cdot h$ under the constraint $Ah = H$, $h \in \mathbb{N}_0^d$ for a weight vector $w$.

2. Let $G_{I_A}$ be the Gröbner basis of the toric ideal $I_A$ in the Algorithm. Let $s_j'$ be the normal form $N(\partial^h s_j, G_{I_A})$ of $\partial^h s_j$ with respect to the Gröbner basis $G_{I_A}$. Choose sufficiently large number $T$ such that all $s_j'$, $j = 1, \ldots, r$ are contained in $M' \cup S$ in the Algorithm.

3. Construct the Macaulay type matrix $F_T'$ by the Algorithm.
4. Restrict $F_T'$ to $x = X$ and $c = \beta + kH$ and set the obtained matrix $F''_T$. If the reduced row echelon form of $F''_T$ contains the element standing for

$$s'_i \equiv \sum_{j=1}^{r} t_{ij}s_j, \quad t_{ij} \in \mathbb{Q}(k)$$

for all $j = 1, \ldots, r$, then go to the step 5 else increase $T$ and go to the step 3.

5. Define the matrix $R(k)$ by $(t_{ij})$.

**Example 2** Put $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ and set the direction $H = (1, 1, 1)^T$, which stands for $\partial_4$, put the starting point $\beta = (3, 2, 1)^T$, and put $X = (1, 1, 1/2, 1)$. As we have seen in Example\textsuperscript{1} we can choose $S = (1, \partial_4)^T$. Then, we obtain

$$\begin{pmatrix} Z(\beta - (k-1)H; X) \\ (\partial_4 \bullet Z)(\beta - (k-1)H; X) \end{pmatrix} = \begin{pmatrix} 0 \\ -2(k+1)(k+2) \end{pmatrix} \begin{pmatrix} 1 \\ 3k+5 \end{pmatrix} \begin{pmatrix} Z(\beta - kH; X) \\ (\partial_4 \bullet Z)(\beta - kH; X) \end{pmatrix}$$

by our algorithm. This recurrence relation holds for any hypergeometric polynomial for the parameter $c = \beta + kH$. The hypergeometric polynomial $Z(c;x)$ is $x_1 x_2 x_3^{k+1} \frac{a^{k+1}}{(k+1)!} + x_1^2 x_3^k$ for $c = (3+k, 2+k, 1+k)^T$.

5 **Hilbert Driven Algorithm for $A$-Hypergeometric Ideals**

In the previous section, we discussed a method to evaluate hypergeometric polynomial by contiguity relations and Macaulay type matrices. The evaluation by contiguity relations (the difference HGM) can also be performed if we have a Gröbner basis for the ideal $H_A(c)$. The rank of $RH_A(c)$ where $c_i$’s are indeterminates is equal to the normalized volume of $A$ by Adolphson’s theorem [1], then the Hilbert driven algorithm by Traverso [17] can be applied to the $A$-hypergeometric ideal $H_A(c)$. Let us summarize this method as an algorithm.

**Algorithm 3** (Hilbert driven method for $A$-hypergeometric systems)
Input: a matrix $A$ standing for the $A$-hypergeometric system and $\beta \in \mathbb{N}_0A$.
Output: recurrence relations (difference Pfaffian systems).

1. Evaluate the normalized volume of $A$. Put it $r$. \textsuperscript{1}

\textsuperscript{1}The number $r$ can be obtained by evaluating the multiplicity of $I_A$ or by evaluating the volume by geometric methods.
2. Compute the Gröbner basis $G$ of $H_A(c)$ in the ring of differential operators $R = \mathbb{Q}(c_1, \ldots, c_d, x_1, \ldots, x_n)(\partial_1, \ldots, \partial_n)$. We use the Hilbert driven algorithm to avoid unnecessary $S$-pair checks. In other words, we stop the Buchberger algorithm once the number of standard monomials of an intermediate Gröbner basis equals to $r$.

3. Find a path from $\beta$ to $\beta'$ which is near to 0 by the Algorithm 4.

4. Let $S$ be the vector of the standard monomials for $G$. Compute matrices $P_i$ such that $\partial_i S - P_i S$ modulo $G$ by the normal form algorithm only for the indexes $i$ which appear in the path.

5. We have recurrences $P_i(c + a_i)Y(c + a_i) = Y(c)$ on the path.

At each step of applying recurrences, we specialize parameters $c$ and $x$ in $P_i$ to numbers. It is our heuristic observation that $c_i$ should not be specialized to numbers or parametric polynomials during the Gröbner basis computation and the normal form computation. It seems to make the computation slower.

6. **Comparison of Methods to Evaluate Hypergeometric Polynomials**

In the latter two sections, we have illustrated two methods to evaluate $A$-hypergeometric polynomials numerically. We will compare the two methods and the method by enumerating the fiber $\{u \in \mathbb{N}^d \mid Au = \beta\}$.

The advantage of the method of Macaulay type matrices is that the step 4 of the computation of the reduced row echelon form is performed in the field $\mathbb{Q}(k)$. On the other hand, the Gröbner basis method needs computation in the ring $\mathbb{Q}(c, x)(\partial_1, \ldots, \partial_n)$, which contains $d + 2n$ indeterminates.

All timing data in this section is taken on a machine with Intel Xeon CPU (2.70GHz) with 256 G memory.

**Example 3** Timing data of `test_c11c_cont12` in the package [12] for Risa/Asir.

Put

$$A = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1
\end{pmatrix}, \quad \beta = (3, 2, 1, 1)^T.$$

$X = (1, 1/2, 1/3, 2/3, 1, 1, 1, 1, 1) S = (1, \partial_5, \partial_6, \partial_7, \partial_8, \partial_8^2)^T$.

Our benchmark problem is to get the value of $S \cdot Z$ at $x = X$ and $c = \beta + k(3, 1, 1, 1)^T$. The contiguity for $(1, 1, 1, 1)^T$ is applied to obtain the value at $\beta + k(1, 1, 1, 1)^T$ and next the contiguity for $(1, 0, 0, 0)^T$ is applied to obtain the value at $\beta + k(3, 1, 1, 1)^T$ from the value at $\beta + k(1, 1, 1, 1)^T$. 

10
The following table is timing data of the method of Macaulay type matrices.

| k  | Time (second) |
|----|---------------|
| 0  | 1.45          |
| 10 | 1.48          |
| 20 | 1.70          |
| 30 | 1.79          |
| 40 | 1.89          |
| 50 | 1.99          |
| 60 | 2.16          |
| 70 | 2.35          |
| 80 | 2.58          |
| 90 | 2.82          |
| 100| 3.17          |

The timing for the case $k = 0$ is the time to construct recurrence relations for the direction $(1,1,1,1)^T$ and $(1,0,0,0)^T$. When $k = 10$, our implementation outputs

\[
\text{Val} = [3031806652733247242457/8961925122434933772252249279430656000, 
\ldots].
\]

From the output, we obtain, e.g., values

\[
Z = \frac{3031806652733247242457}{89619251224349337722522492794306560000} = 3.38\ldots \times 10^{-16}
\]

\[
E[U_8] = \partial_8 \cdot \frac{Z}{Z} = \frac{52047189429143224956864/30318066527332447242457}{30318066527332447242457} = 1.71\ldots
\]

where $E[U_8]$ is the expectation of the random variable $U_8 \in \mathbb{N}_0^8$ which satisfies $AU = \beta + k(3,1,1,1)^T$. We compare these data with other two methods for exact evaluation. First one is an exhaustive enumeration of $U \in \mathbb{N}_0^8$ satisfying $AU = \beta + k(3,1,1,1)^T$. It can be easily enumerated for this $A$ by a nested loop of four for statements, because when non-negative integers $U_5, U_6, U_7, U_8$ are given, other $U_i$’s are determined uniquely as integers and the admissible range of $U_5, \ldots, U_8$ can be described in terms of $k$. Moreover coefficients can easily be evaluated by recurrences.
Figure 2: black square=Macaulay type matrix, white square=Hilbert driven GB, white circle=series, horizontal=$k/10$, vertical=seconds

| $k$ | Time (second) | number of fibers |
|-----|---------------|------------------|
| 0   | 0             | 5                |
| 10  | 0.0200        | 1946             |
| 20  | 0.132         | 18436            |
| 30  | 0.476         | 76976            |
| 40  | 1.42          | 220066           |
| 50  | 3.70          | 505206           |
| 60  | 7.68          | 1004896          |
| 70  | 15.9          | 1806636          |
| 80  | 30.9          | 3012926          |
| 90  | 56.0          | 4741266          |
| 100 | 88.7          | 7124156          |

The last table is timing data of evaluating the normalizing constant in the benchmark problem by recurrence relations derived from Pfaffian operators obtained by the algorithm 3. The Hilbert driven Gröbner basis computation in the step 2 takes 5.50 seconds on the yang package for the Risa/Asir. If we do not use the Hilbert driven method in the step 2, it takes 2069.78 seconds on the yang package. In our benchmark problem, we may only obtain $P_1$ and $P_8$. It takes 6.80 seconds to obtain them in the step 3.
| k | Time Recc (second) | Time Recc + GB (second) |
|---|-------------------|------------------------|
| 0 | 0.02              | 12.31                  |
| 10| 0.06              | 12.35                  |
| 20| 0.072             | 12.37                  |
| 30| 0.14              | 12.43                  |
| 40| 0.27              | 12.56                  |
| 50| 0.34              | 12.64                  |
| 60| 0.49              | 12.78                  |
| 70| 0.62              | 12.92                  |
| 80| 0.83              | 13.13                  |
| 90| 1.05              | 13.34                  |
| 100| 1.32             | 13.62                  |

In the table, “Time Recc” is the time to apply the recurrence relations to obtain the value of the normalizing constant. “Time Recc + GB” is the sum of the “Time Recc” and the time of the algorithm.

The Figure 2 tells us a strategy to evaluate $A$-hypergeometric polynomials numerically. When $\beta$ is small, then the enumeration of the fibers will be the best method. When $\beta$ is large, then the method of Macaulay type matrix will be the best choice. When we need to evaluate the hypergeometric polynomial for several $\beta$’s, then the method of Gröbner basis by the Hilbert driven algorithm will be the best choice.

7 Finding a Path to Apply for Recurrence Relations

Suppose that $\beta \in \mathbb{N}_0A$. We want to find a path in $\mathbb{N}_0A$ from $\beta$ to $\beta'$ closer to 0.

Algorithm 4

Input: $A$ and $\beta \in \mathbb{N}_0A$.

Output: a path to $\beta' \in \mathbb{N}_0^d$

1. Express $\beta$ as $\beta = n_1a_{i_1} + \cdots + n_ka_{i_k}$, $n_i \in \mathbb{Z}_{>0}$ by reducing $\partial^\beta$ by a Gröbner basis of $I_A$.

2. Find the maximum in the set

$$\{m \mid \text{path } \beta - \gamma_a \rightarrow \beta - \gamma_a - ma_{i_j} \text{ lies in } \mathbb{N}_0A \text{ for all } \gamma_a = \sum u_ka_k, \partial^\mu \in S \text{ and } m \leq n_j\}$$

When the set is empty or $\beta = 0$, return the path and $\beta$ as the terminating point $\beta'$.

3. Add the pair $(i_j, m)$ to the path. Put $n_j = n_j - m$ and $\beta = \beta - ma_{i_j}$. Go to the step 2.
The definition that “path \( \beta - \gamma_u \rightarrow \beta - \gamma_u - ma_i \) lies in \( N_0A \)” is that \( \beta - \gamma_u - a_i \in N_0A, \beta - \gamma_u - 2a_i \in N_0A, \ldots, \beta - \gamma_u - ma_i \in N_0A. \)

Let \( S \) be a basis of \( R/RHA(c) \) consisting of monomials in \( \partial_i \)'s and \( \partial_i S - P, S, i = 1, \ldots, n \) is the Pfaffian operator for \( S \). The singularity polynomial of the Pfaffian operator is the least common multiple of the denominator polynomials of the elements of the \( P_i \)'s. It is a polynomial in \( C[c, x] \). The basis \( S \) is called a good basis for \( N_0A \) when the following two conditions are satisfied.

1. The singularity polynomial for the Pfaffian operator does not contain the variables \( c_i \)'s.
2. \( S \) is a basis of \( R'/R'H_A(\beta), R' = C(x)\langle \partial_1, \ldots, \partial_n \rangle \) for all \( \beta \in N_0A. \)

Note that when \( A \) admits a rank jumping parameter \( \beta \in N_0A \) \[14\], there exists no good basis.

The following theorem shows that the algorithm \[4\] and HGM reduce the evaluation of \( Z(\beta; x) \) to \( Z(\beta'; x) \) where \( \beta' \) closer to 0 when the condition of the theorem holds.

**Theorem 2** When \( A \) is normal and \( S \) is a good basis, then the matrix \( P_i(\beta + a_i; x) \) has the inverse for \( \beta \) satisfying \( \beta - \sum u_i a_i \in N_0A \) for all \( \partial^u \in S \) and for any \( x \) out of a measure zero set.

**Proof.** We will construct the inverse matrix of \( P_i \). We denote by \( M_A(\beta) \) the left \( D_n \)-module \( D_n/D_nH_A(\beta) \). Since \( A \) is normal, it follows from Saito’s isomorphism \[10\] that \( M_A(\beta) \) and \( M_A(\beta + a_i) \) are isomorphic as the left \( D_n \)-module when \( \beta \in N_0A. \) The isomorphism is given by

\[
M_A(\beta) \ni \ell \mapsto \ell \partial_i \in M_A(\beta + a_i).
\]

In general, they are isomorphic when \( \beta + a_i \notin V(B_i) \), which is the zero set of the \( b \)-ideal for the direction \( i \) \[12\]. When \( A \) is normal, we have \( \beta + a_i \notin V(B_i) \) for all \( i \) and \( \beta \in N_0A. \) Note that we have the induced isomorphism

\[
\text{Hom}_{D_n}(M_A(\beta + a_i), \mathcal{O}) \ni f \mapsto \partial_i \cdot f \in \text{Hom}_{D_n}(M_A(\beta), \mathcal{O}).
\]

Let \( \partial^u \) be an element of \( S \). When \( \beta - \sum u_k a_k \in N_0A, \) there exists an operator \( \ell_{iu} \in D_n \) such that the following diagram of left \( D_n \)-modules commutes and all arrows are isomorphisms.

\[
\begin{array}{ccc}
M_A(\beta - \sum u_k a_k) & \xrightarrow{\ell_{iu}} & M_A(\beta + a_i - \sum u_k a_k) \\
\downarrow \partial^u & & \downarrow \partial^u \\
M_A(\beta) & \xrightarrow{\partial_i} & M_A(\beta + a_i)
\end{array}
\]
Since the diagram commutes, we have
\[ \ell_{iu} \partial^u \partial_i - \partial^u = 0 \text{ in } M_A(\beta + a_i), \]
which implies \( \ell_{iu} \partial^u \partial_i - \partial^u \in D_n H_A(\beta + a_i) \). Let \( f \) be a local solution of \( M_A(\beta) \), which can be regarded as an element of Hom\(_{D_n}(M_A(\beta), O) \). Consider the column vector \( Y_f(\beta) = (s \cdot f \mid s \in S) \). Take a local solution \( g \) in Hom\(_{D_n}(M_A(\beta + a_i), O) \). We have \( \ell_{iu} \partial^u \partial_i \cdot g = \partial^u \cdot g = 0 \). The function \( f = \partial_i \cdot g \) is a solution of \( M_A(\beta) \). Then, we have \( Y_g(\beta + a_i) = (\partial^u \cdot g \mid \partial^u \in S) = (\ell_{iu} \partial^u \cdot f \mid \partial^u \in S) \). Suppose that \( \ell_{iu} \partial^u = \sum c_u(x) \partial^u \). Since \( S \) is a good basis, \( \partial^u \) can be expressed as a linear combination of \( S \) with coefficients in \( C(x) \) modulo \( R'H_A(\beta) \). Therefore, there exists a matrix \( Q_i \) with rational function entries satisfying \( Y_g(\beta + a_i) = Q_i(x) Y_f(\beta) \). We note that \( Q_i \) does not depend on the choice of the function \( g \). It follows from the continuity relation that we have \( Y_f(\beta) = P_1(\beta + a_i, x) Y_g(\beta + a_i) \). Hence, we have \( Y_g(\beta + a_i) = Q_i(x) P_1(\beta + a_i, x) Y_g(\beta + a_i) \). Make \( g \) run over a solution basis. Then, we can see that \( Q_j(x) P_1(\beta + a_i, x) = E \) out of an analytic set in the \( x \)-space. The analytic set defined by a non-zero holomorphic function is a measure zero set. A countable union of measure zero sets is a measure zero set. Q.E.D.

Note that we do not have an example of \( S \) which is not good when \( A \) is normal. It is an interesting question to find such example or to prove the existence of a good basis.

8 Reduction to Non-Negative \( a_i \)'s

We have supposed \( a_i \in \mathbb{N}_0^d \) in the previous sections. We assume that \( a_i \in \mathbb{Z}^d \) and there exists a linear form \( h \in \mathbb{R}^n \) such that \( h(a_i) = 1 \) for all \( i \)'s in this section. We will explain a method to reduce this case to the case \( a_i \in \mathbb{N}_0^d \).

We define the vector \( p \in \mathbb{N}_0^d \) by
\[ p_i = -\min_j (a_{ij}, 0). \]
When \( h(p) = -1 \), we choose a unit vector \( e \) which is not orthogonal to \( h \) and rewrite \( p \) by \( p + e \). Then, we can assume that \( h(p) \neq -1 \). Define the new matrix \( A' \) by \( (a_1 + p, \ldots, a_n + p) \). The toric ideals for \( A \) and for \( A' \) agree, e.g., [5, Lemma 1.5.10].

Let us discuss on a relation between solutions of the \( H_A \) and \( H_{A'} \). When \( f \) is a solution of \( H_A(\beta) \), we have
\[ \sum_{j=1}^n a_{ij} x_j \partial_j f = \beta_i f. \]
Multiplying \( h_i \) and adding with respect to \( i \), we obtain
\[ \sum_{j=1}^n x_j \partial_j f = h(\beta) f \]
by utilizing \( h(a_j) = 1 \). Adding \( p_k(\sum_{j=1}^{n} x_j \partial_j f) = h(\beta)p_k f \) to the \( k \)-th equation \( E_k f = \beta_k f \), we obtain
\[
\sum (a_{kj} + p_k)x_j \partial_j f = (\beta_k + h(\beta)p_k)f
\]
Hence, the solutions of \( H_A(\beta) \) are solutions of \( H_A(\beta') \) where \( \beta' = \beta + h(\beta)c \). In particular, the hypergeometric polynomial for \( H_A(\beta) \) is the hypergeometric polynomial for \( H_A'(\beta') \).

References

[1] A. Adolphson, Hypergeometric Functions and Rings Generated Monomials. Duke Mathematical Journal 73 (1994), 269–290.

[2] P. Diaconis, B. Sturmfels, Algebraic Algorithms for Sampling from Conditional Distributions, Annals of Statistics 26 (1998), 363–397.

[3] Y. Goto, Contiguity Relations of Lauricella’s \( F_D \) revisited. arXiv:1412.3256

[4] H. Hashiguchi, Y. Numata, N. Takayama, A. Takemura, Holonomic gradient method for the distribution function of the largest root of a Wishart matrix, Journal of Multivariate Analysis, 117 (2013) 296-312.

[5] T. Hibi et al, Groebner Bases : Statistics and Software Systems, Springer, 2013.

[6] T. Hibi, K. Nishiyama, N. Takayama, Pfaffian Systems of \( A \)-Hypergeometric Systems I, Bases of Twisted Cohomology Groups. arXiv:1212.6103

[7] T. Hibi, K. Nishiyama, N. Takayama, Computing Bases of Twisted Cohomology Groups for Sparse Polynomials, ACM Communications in Computer Algebra 48 (2014), 118–120.

[8] S. Kuriki, N. Takayama, A. Takemura, \( A \)-Hypergeometric Distributions and Newton Polytopes, preprint.

[9] L. Matusevich, E. Miller, U. Walther, Homological Methods for Hypergeometric Families, Journal of American Mathematical Society 18 (2005), 919–941.

[10] F. S. Macaulay, On Some Formulæ in Elimination, Proceedings of the London Mathematical Society 33 (1902), 3–27.

[11] H. Nakayama, K. Nishiyama, M. Noro, K. Ohara, T. Sei, N. Takayama, A. Takemura, Holonomic Gradient Descent and its Application to the Fisher-Bingham Integral, Advances in Applied Mathematics 47 (2011), 639–658.

[12] ot-hgm-ahg.rr: A Risa/Asir package to evaluate \( A \)-hypergeometric polynomials by HGM. http://www.openxm.org
[13] M. Ogawa, Algebraic Statistical Methods for Conditional Inference of Discrete Statistical Models, PhD. Thesis, University of Tokyo, 2015.

[14] M. Saito, B. Sturmfels, N. Takayama, Hypergeometric polynomials and Integer Programming, Compositio Mathematica, 115, (1999) 185–204.

[15] M. Saito, B. Sturmfels, N. Takayama, Gröbner Deformations of Hypergeometric Differential Equations, Algorithms and Computation in Mathematics 6, Springer, 2000.

[16] M. Saito, Isomorphism Classes of A-Hypergeometric Systems, Compositio Mathematica 128 (2001), 323–338.

[17] C. Traverso, Hilbert Functions and the Buchberger Algorithm, Journal of Symbolic Computation, 22 (1996), 355–376.