CHALLENGING COMPUTATIONS OF HILBERT BASES OF CONES ASSOCIATED WITH ALGEBRAIC STATISTICS

WINFRIED BRUNS, RAYMOND HEMMECKE, BOGDAN ICHIM, MATTHIAS KÖPPE, AND CHRISTOF SÖGER

Abstract. In this paper we present two independent computational proofs that the monoid derived from $5 \times 5 \times 3$ contingency tables is normal, completing the classification by Hibi and Ohsugi. We show that Vlach's vector disproving normality for the monoid derived from $6 \times 4 \times 3$ contingency tables is the unique minimal such vector up to symmetry. Finally, we compute the full Hilbert basis of the cone associated with the non-normal monoid of the semi-graphoid for $|N| = 5$. The computations are based on extensions of the packages LattE-4ti2 and Normaliz.

1. Introduction

Let $S = \mathrm{monoid}(G)$ be an affine monoid generated by a finite set $G \subseteq \mathbb{Z}^n$ of integer vectors. We call S *normal* if $S = \mathrm{cone}(G) \cap \mathrm{lattice}(G)$, where $\mathrm{cone}(G) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum \lambda_i \mathbf{g}_i, \lambda_i \in \mathbb{R}_+, \mathbf{g}_i \in G\}$ denotes the rational polyhedral cone generated by G and where $\mathrm{lattice}(G) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum \lambda_i \mathbf{g}_i, \lambda_i \in \mathbb{Z}, \mathbf{g}_i \in G\}$ denotes the sublattice of $\mathbb{Z}^n$ generated by G. In this paper, we will stick to the case that $\mathrm{lattice}(G) = \mathbb{Z}^n$. Then, normality of S is equivalent to saying that G contains the Hilbert basis of $\mathrm{cone}(G)$, i.e., every lattice point in $\mathrm{cone}(G)$ can be written as a nonnegative integer linear combination of elements in G. Lattice points in $\mathrm{cone}(G) \setminus \mathrm{monoid}(G)$ are called *holes* (or *gaps*). Clearly, $\mathrm{monoid}(G)$ is nonnormal if and only there exists at least one hole.

By the *Hilbert basis* $\mathcal{H}(C)$ of a pointed rational cone C we mean the unique minimal system of generators of the monoid M of lattice points in C. The Hilbert basis of C consists of the *irreducible elements* of M, i.e., those elements of M that do not have a nontrivial representation as a sum of two elements of M (see [2, Ch. 2] for a comprehensive discussion). Note that deciding normality of an affine monoid is NP-hard [6].

An $r_1 \times r_2 \times \cdots \times r_N$ *contingency table* is a function $T : \{1, \ldots, r_1\} \times \cdots \times \{1, \ldots, r_N\} \to \mathbb{Z}_+$ where $\mathbb{Z}_+$ denotes the nonnegative integers. It can be imagined as an N-dimensional

---

Date: April 20, 2010.

Acknowledgement: B. Ichim was partially supported by CNCSIS grant RP-1 no. 7/01.07.2009 during the preparation of this work. M. Köppe was partially supported by grant DMS-0914873 of the National Science Foundation.

array of size $r_1 \times r_2 \times \cdots \times r_N$ with nonnegative integer entries. Such a contingency table arises when one classifies a sample of individuals according to the values of N random variables X_j, $j = 1, \dots, N$, where X_j takes values in $\{1, \dots, r_j\}$. The j-th $(N-1)$-marginal T_j of T is the $r_1 \times \cdots \times r_{j-1} \times r_{j+1} \times \cdots \times r_N$ contingency table defined by

$$T_j(i_1, \dots, i_{j-1}, i_{j+1}, \dots, i_N) = \sum_{k=1}^{r_j} T(i_1, \dots, i_{j-1}, k, i_{j+1}, \dots, i_N).$$

The marginals are the basic tool for testing the independence of X_j from the compound random variable $(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_N)$.

The $r_1 \times r_2 \times \cdots \times r_N$ contingency tables form the monoid $\mathcal{O}$ of integral points in the nonnegative orthant of $\mathbb{R}^D$ where $D = r_1 \cdots r_N$. The assignment $T \mapsto (T_1, \dots, T_N)$ is a monoid homomorphism $\mathcal{M}$ from $\mathcal{O}$ into the monoid of nonnegative integer points in $\mathbb{R}^{d_1+\cdots+d_N}$ where $d_j = r_1 \cdots r_{j-1} r_{j+1} \cdots r_N$. In the following, the image $\mathcal{M}(\mathcal{O})$ is called the *monoid derived from* $r_1 \times r_2 \times \cdots \times r_N$ *contingency tables* (by taking line sums). For the role of these monoids and their normality in algebraic statistics we refer the reader to Ohsugi and Hibi [12], Drton, Sturmfels and Sullivant [5] and Sullivant [15].

Normality of monoids derived from $r_1 \times r_2 \times \cdots \times r_N$ contingency tables was settled almost completely in [12]. In this paper we close the last open cases by showing computationally, via two different approaches and independent implementations, that $5 \times 5 \times 3$ has a normal monoid. The normality for $5 \times 5 \times 3$ implies normality for the other two open cases $5 \times 4 \times 3$ and $4 \times 4 \times 3$ by [12, 3.2] (or can also be verified computationally).

In the last section we report on a partial verification of a conjecture of Sturmfels and Sullivant [14] on the normality of cut monoids of graphs.

2. Results

The defining matrix $A_{5\times5\times3}$ whose columns generate the monoid associated to $5 \times 5 \times 3$ contingency tables is given in the following table in which every $\cdot$ corresponds to an entry 0.

```
1..............1..............1..............1..............1..............
.1..............1..............1..............1..............1.............
..1..............1..............1..............1..............1............
...1..............1..............1..............1..............1...........
....1..............1..............1..............1..............1..........
.....1..............1..............1..............1..............1.........
......1..............1..............1..............1..............1........
.......1..............1..............1..............1..............1.......
........1..............1..............1..............1..............1......
.........1..............1..............1..............1..............1.....
..........1..............1..............1..............1..............1....
...........1..............1..............1..............1..............1...
............1..............1..............1..............1..............1..
```

```
.............1..............1..............1..............1..............1.
..............1..............1..............1..............1..............1
1..1..1..1..1..............................................................
.1..1..1..1..1.............................................................
..1..1..1..1..1............................................................
...............1..1..1..1..1...............................................
................1..1..1..1..1..............................................
.................1..1..1..1..1.............................................
..............................1..1..1..1..1................................
...............................1..1..1..1..1...............................
................................1..1..1..1..1..............................
.............................................1..1..1..1..1.................
..............................................1..1..1..1..1................
...............................................1..1..1..1..1...............
............................................................1..1..1..1..1..
.............................................................1..1..1..1..1.
..............................................................1..1..1..1..1
111........................................................................
...111.....................................................................
......111..................................................................
.........111...............................................................
............111............................................................
...............111.........................................................
..................111......................................................
.....................111...................................................
........................111................................................
...........................111.............................................
..............................111..........................................
.................................111.......................................
....................................111....................................
.......................................111.................................
..........................................111..............................
.............................................111...........................
................................................111........................
...................................................111.....................
......................................................111..................
.........................................................111...............
............................................................111............
...............................................................111.........
..................................................................111......
.....................................................................111...
........................................................................111
```

Note that this normality problem cannot be settled directly (except for $4 \times 4 \times 3$) by computing the Hilbert basis of the associated cone using state-of-the-art software such as `Normaliz` v2.2 [3, 4] or `4ti2` v1.3.2 [1, 7]. Both codes fail to return an answer due to time and to memory requirements of intermediate computations. Using the computational approaches presented below, we can now show the following.

Lemma 1. *The monoid derived from $5 \times 5 \times 3$ contingency tables by taking line sums (= two-marginals) is normal.*

This completes the normality classification of the monoids derived from $r_1 \times r_2 \times \cdots \times r_N$ contingency tables by taking line sums as given in [12]:

Theorem 2. *Let $r_1 \geq r_2 \geq \ldots \geq r_N \geq 2$ be integer numbers. Then the monoid derived from $r_1 \times r_2 \times \cdots \times r_N$ contingency tables by taking line sums is normal if and only if the contingency table is of size*

- $r_1 \times r_2$, $r_1 \times r_2 \times 2 \times \ldots \times 2$, *or*
- $r_1 \times 3 \times 3$, *or*
- $4 \times 4 \times 3$, $5 \times 4 \times 3$, *or* $5 \times 5 \times 3$.

For the monoid of $6 \times 4 \times 3$ contingency tables, a vector disproving normality was presented by Vlach [16]. Let M be the monoid derived from $6 \times 4 \times 3$ contingency tables and $\mathbf{f}$ be the vector in $\mathbb{R}^{4\times 3} \oplus \mathbb{R}^{6\times 3} \oplus \mathbb{R}^{6\times 4}$ given by the following three matrices:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

The unique point in the $6\times4\times3$ (transportation) polytope $\{\mathbf{z} \in \mathbb{R}^{6\times 4\times 3} : A\mathbf{z} = \mathbf{f}, \mathbf{z} \geq \mathbf{0}\}$ is

$$\mathbf{z}^* = \frac{1}{2} \left(\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{array} \right).$$

(We have written the $6 \times 4 \times 3$ contingency table $\mathbf{z}^*$ as a sequence of 6 matrices of size 4×3.) This equation shows on the one hand that $\mathbf{f}$ indeed belongs to the cone C generated by M (since $2\mathbf{f} \in M$), and on the other hand, by the uniqueness of the solution, that $\mathbf{f} \notin M$. Since $\mathbb{Z}^{6\times 4\times 3}/\operatorname{lattice}(M)$ is torsionfree (as one can verify computationally), $\mathbf{f}$ lies in $\operatorname{lattice}(M)$, and it follows that M is not normal. We can show the following more precise result.

Lemma 3. *The vector $\mathbf{f}$ presented by Vlach [16] is the unique vector (up to the underlying $S_6 \times S_4 \times S_3$ symmetry) in the Hilbert basis of the cone of $6 \times 4 \times 3$ contingency tables that is not an extreme ray.*

The treatment in [8] now completely describes *all* holes of the cone, that is, all lattice points in $\operatorname{cone}(A_{6\times4\times3})$ that cannot be written as a nonnegative linear integer combination of the (integer) generators of the cone:

Corollary 4. *Let* $\mathbf{f}$ *be the hole in* $\mathrm{cone}(A_{6\times 4\times 3})$ *and let* $\mathbf{z}^* \in \mathbb{R}_+^{6\times 4\times 3}$ *be the unique solution to* $A_{6\times 4\times 3}\mathbf{z} = \mathbf{f}, \mathbf{z} \in \mathbb{R}_+^{6\times 4\times 3}$*, as stated above. Moreover, let* G *denote the set of those* 24 *columns of* $A_{6\times 4\times 3}$ *for which* $\mathbf{z}_i^* > 0$.

Then the set of holes in $\mathrm{cone}(A_{6\times 4\times 3})$ *is the set of all points that can be written uniquely as* $\sigma(\mathbf{f} + \mathbf{s})$ *with* $\sigma \in S_6 \times S_4 \times S_3$ *and with* $\mathbf{s} \in \mathrm{monoid}(G)$.

Finally, we have computed the Hilbert basis of the cone associated to the semi-graphoid for $|N| = 5$ [13]. It was already shown in [9] that the corresponding monoid is not normal by constructing a hole via a different method. The computation of the full Hilbert basis was not possible at that time, neither with `Normaliz`, nor with `4ti2`. Here is the defining matrix whose columns generate the monoid associated to the semi-graphoid for $|N| = 5$. Every dot corresponds to an entry 0. The symbols $+$ and $-$ represent entries 1 and -1.

```
++++++++++......................................................................
----......++++++................................................................
-...---.........++++++..........................................................
.-..-..--................++++++.................................................
..-..-.-.-............................++++++....................................
...-..-.--..............................................++++++..................
+.........---...---...+++.......................................................
.+........-..--..........---...+++..............................................
..+........-.-.-......................---...+++.................................
...+........-.--........................................---...+++...............
....+...........-..--....-..--....+++...........................................
.....+...........-.-.-................-..--....+++..............................
......+...........-.--..................................-..--....+++............
.......+..................-.-.-........-.-.-.......+++..........................
........+..................-.--..........................-.-.-.......+++........
.........+..............................-.--..............-.--............+++...
..........+.....+.....--.+.....--.--.+..........................................
...........+.....+....-.-.............+.....--.--.+.............................
............+.....+....--...............................+.....--.--.+...........
.............+............+....-.-.....+....-.-....--.+.........................
..............+............+....--.......................+....-.-....--.+.......
...............+........................+....--...........+....--.........--.+..
...................+........+.....-.-....+.....-.-.-.-.+........................
....................+........+.....--......................+.....-.-.-.-.+......
.....................+....................+.....--..........+.....--......-.-.+.
..............................+............+........--.......+........--...--..+
......................+........+..+..-......+..+..-+..--........................
.......................+........+..+.-........................+..+..-+..--......
........................+....................+..+.-............+..+.-.....+..--.
.................................+............+.....+.-.........+.....+.-..+.-.-
....................................+............+...+.-...........+...+.-..+.--
.....................................+............+...++............+...++...+++
```

Lemma 5. *The Hilbert basis of the cone associated to the semi-graphoid for* $|N| = 5$ *has* 1300 *elements that come in* 21 *orbits under the underlying symmetry group* $S_5 \times S_2$. *These are represented by the* 21 *rows of the following matrix:*

$$
\begin{array}{*{32}{r}}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
2 & 0 & 0 & 0 & 0 & -2 & -1 & -1 & 1 & 1 & 0 & -1 & 1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -2 & 1 & -1 & -1 & 0 & -1 & 0 & 1 & 1 & 0 & 1 \\
2 & 0 & 0 & 0 & -2 & 0 & -1 & -1 & 1 & 0 & 0 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 0 & 1 & -1 & -1 & 1 & 0 & -1 & 0 & -2 & 0 & 0 & 0 & 2 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -1 & 2 & -1 & 2 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 2 & -1 & -1 & -1 & 2 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & -1 & -1 & 1 & -1 & 1 & -2 & 0 & 0 & -1 & 0 & 1 & 0 & 1 & -1 & 0 & -2 & 0 & 1 & 1 & -1 & -1 & -1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & -1 & 0 & -2 & 1 & 0 & 0 & -2 & 0 & 0 & -1 & -1 & 1 & 0 & 1 & -2 & 0 & -1 & 1 & 1 & -1 & 0 & 1 & -1 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & -1 & 0 & -2 & 0 & 0 & 0 & -2 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 0 & 0 & -2 & 0 & 0 & 2 \\
1 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & 2 & -1 & 2 & 0 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & -2 & -1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -2 & 1 & 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -2 & 2 \\
1 & 1 & 0 & 1 & 0 & -1 & -1 & -1 & -1 & 0 & -1 & 1 & 1 & -1 & 0 & 1 & 1 & 0 & -1 & 1 & 1 & -1 & 0 & -1 & -1 & -1 & -1 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 & -1 & 1 & -2 & 0 & 0 & 0 & -2 & 1 & -1 & -1 & -2 & 1 & 0 & 0 & 0 & 1 & -2 & -1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
2 & -1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & 0 & 1 & -2 & -1 & -1 & -2 & 0 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & -2 & -1 & -1 & 0 & -1 & 0 & -1 & -1 & -1 & -2 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & -1 & -1 & -2 & -1 & -1 & 3 \\
0 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 2 & 2 & 0 & 0 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -2 & 3 \\
1 & 0 & 0 & 1 & 2 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -2 & 0 & 0 & 2 & 1 & 0 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -2 & 0 & 0 & 0 & 0 & 2 \\
2 & 0 & 0 & 0 & 0 & -2 & -1 & -1 & 1 & 2 & 1 & -1 & 1 & -1 & 1 & 1 & 0 & 0 & -1 & 0 & -1 & -2 & 0 & -2 & -1 & -1 & 0 & 1 & 1 & 1 & 2 & 0 \\
3 & -1 & -1 & 1 & -1 & -2 & 0 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 0 & -2 & -1 & 1 & -1 & -1 & 3 \\
3 & 1 & -1 & -1 & -1 & -2 & -1 & -1 & -1 & 2 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & -2 & 0 & 0 & 0 & 1 & 2 \\
2 & 1 & 0 & 0 & 0 & -2 & -1 & -1 & -1 & 2 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 2 & -1 & -1 & -1 & -2 & 0 & 0 & 0 & 1 & 2
\end{array}
$$

3. Computational Approaches

In this section we present the two computational approaches that allowed us to solve the three challenging Hilbert basis computations of the cones associated to $5 \times 5 \times 3$-tables, to $6 \times 4 \times 3$-tables, and to semi-graphoids for $|N| = 5$. In the first approach, we iteratively decompose the cone into smaller cones and exploit the underlying symmetry and set inclusion to avoid a lot of unnecessary computations. An implementation of this approach is freely available in the new release `latte-for-tea-too-1.4` of "LattE for tea, too" (`http://www.latte-4ti2.de`), a joint source code distribution of the two software packages `LattE macchiato` and `4ti2`. In the second approach, we exploit the fact that the cones are nearly compressed; hence many cones in any pulling triangulation are unimodular, and the same holds in placing triangulations. Using our second approach, none of these unimodular cones is constructed, saving a lot of computation time. An implementation of this approach will be freely available in the next release of Normaliz (`http://www.math.uos.de/normaliz`), together with the input files of the examples of this paper.

3.1. **First approach: exploiting symmetry.** Let us assume that we wish to compute the Hilbert basis of a rational polyhedral cone $C = \operatorname{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_s) \subseteq \mathbb{R}^n$. Moreover, assume that C has a coordinate-permuting symmetry group S, that is, if $\mathbf{v} \in V$ and $\sigma \in S$ then also $\sigma(\mathbf{v}) \in C$. Herein, the vector $\sigma(\mathbf{v})$ is obtained by permuting the components of $\mathbf{v}$ according to the permutation σ.

One approach to find the Hilbert basis of C is to find a regular triangulation of C into simplicial cones $C_1, \ldots, C_k$ and to compute the Hilbert bases of the simplicial cones $C_1, \ldots, C_k$. Clearly, the union of these Hilbert bases is a (typically non-minimal) system of generators of the monoid of lattice points in C. The drawback of this approach is that a complete triangulation of C is often too hard to accomplish.

Instead of computing a full triangulation, we compute only a (regular) subdivision of C into few cones. To this end we remove one of the generators of the cones, say $\mathbf{r}_s$, compute the convex hull of the cone $C' = \text{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_{s-1})$, and find all facets $\mathcal{F}$ of C' that are visible from $\mathbf{r}_s$. By $\mathcal{F}'$ we denote the set of all cones that we get as the convex hull of a facet in $\mathcal{F}$ with the ray generated by $\mathbf{r}_s$. Then $\mathcal{F}' \cup \{C'\}$ gives a regular subdivision of C, called the *subdivision with distinguished generator* $\mathbf{r}_s$. Before we now subdivide those cones in $\mathcal{F}'$ further into smaller cones, we use the following simple observation to remove cones that can be avoided due to the underlying symmetry given by S.

Lemma 6. *Let $C, C_1, \ldots, C_k \subseteq \mathbb{R}^n$ be rational polyhedral cones such that $C = \bigcup_{i=1}^k C_i$ (not necessarily a disjoint union). Suppose that there is a permutation σ and indices i and j such that $C_i \subseteq \sigma(C_j) \subseteq C$. Then the Hilbert basis of C is contained in the union of the Hilbert bases of the cones $C_1, \ldots, C_{i-1}, \sigma(C_j), C_{i+1}, \ldots, C_k$.*

Proof. The result follows by observing that all lattice points in C_i also belong to $\sigma(C_j)$ and thus can be written as a nonnegative integer linear combination of the Hilbert basis of $\sigma(C_j)$. □

If successful, this test whether C_i can be dropped is a very efficient way of removing unnecessary cones. However, the fewer generators are present in the cones $C_1, \ldots, C_k$, the higher the chance that this test fails. So one has to make a trade-off between a simple test (that may fail more and more often) and a direct treatment of each cone C_i. As we compute only regular subdivisions whose cones are spanned by some of the vectors $\mathbf{r}_1, \ldots, \mathbf{r}_s$, each of the cones $C_1, \ldots, C_k$ can be represented by a characteristic 0-1-vector $\chi(C_1), \ldots, \chi(C_k)$ of length s that encodes which of the generators of C are present in this cone. This makes the test $C_i \subseteq \sigma(C_j)$ comparatively cheap, as we only need to check whether $\chi(C_i) \leq \sigma(\chi(C_j))$.

Summarizing these ideas, the symmetry-exploiting approach can be stated as follows:

(1) Let $C = \text{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_s) \subseteq \mathbb{R}^n$ and $\mathcal{C} = \{C\}$.
(2) $i := 0$
(3) While $\mathcal{C} \neq \emptyset$ do
   (a) $i := i + 1$
   (b) For all $K \in \mathcal{C}$ that contain the i-th generator compute a subdivision with distinguished ith generator.
   (c) Let $\mathcal{T}$ be the set of all cones in these subdivisions.
   (d) Let $\mathcal{M}$ be the set of those cones with a maximum number of rays.
   (e) Let $\mathcal{C} \neq \emptyset$ be the set $\mathcal{M}$ together with all cones $T \in \mathcal{T}$ that are not covered by a cone $\sigma(M)$ with $M \in \mathcal{M}$ and $\sigma \in S$, see Lemma 6.

   (f) Remove from $\mathcal{C}$ all simplicial cones and compute their Hilbert bases.
(4) For each computed Hilbert basis element $\mathbf{h}$ compute its full orbit $\{\sigma(\mathbf{h}) : \sigma \in S\}$ and collect them in a set $\mathcal{H}$.
(5) Remove the reducible elements from $\mathcal{H}$.
(6) Return the set of irreducible elements as the minimal Hilbert basis of C.

This quite simple approach via triangulations and elimination of cones by symmetric covering already solves all three presented examples. In particular, it gives a computational proof to Lemma 1. The candidates for the representatives of Hilbert basis elements can be computed using "LattE for tea, too" by calling

```
dest/bin/hilbert-from-rays-symm --hilbert-from-rays="dest/bin/hilbert-from-rays"
                                    --dimension=26 S5.rays
dest/bin/hilbert-from-rays-symm --hilbert-from-rays="dest/bin/hilbert-from-rays"
                                    --dimension=43 355.short.rays
dest/bin/hilbert-from-rays-symm --hilbert-from-rays="dest/bin/hilbert-from-rays"
                                    --dimension=42 346.short.rays
```

The data files can be found on `http://www.latte-4ti2.de`. (For typographical reasons each command has been printed in two lines.)

3.2. **Second approach: partial triangulation.** In the second approach, we build up a triangulation of the given cone $C = \operatorname{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_s) \subseteq \mathbb{R}^n$. However, by using the following Lemma 7 and its Corollary 8, we can avoid regions of the triangulation that consist only of unimodular cones (for which the extreme ray generators already constitute a Hilbert basis). More precisely, we try to omit simplicial cones whose non-extreme Hilbert basis elements are contained in previously computed simplicial cones.

In the following we describe the facets of a full-dimensional rational cone by (uniquely determined) primitive integral exterior normal vectors. In other words, $F = \{\mathbf{x} \in C : \mathbf{c}^\intercal \mathbf{x} = 0\}$ where c has coprime integer entries and $\mathbf{c}^\intercal \mathbf{y} \leq 0$ for all $\mathbf{y} \in C$.

Lemma 7. *Let $C = \operatorname{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_k) \subseteq \mathbb{R}^n$ be a rational polyhedral cone such that*

- $\mathbf{r}_1, \ldots, \mathbf{r}_k \in \mathbb{Z}^n$,
- $\mathbf{r}_1, \ldots, \mathbf{r}_{k-1}$ *lie in a facet of C defined by the hyperplane* $\mathbf{c}^\intercal \mathbf{x} = 0$,
- $\mathbf{c}^\intercal \mathbf{r}_k = 1$.

Then the Hilbert basis of C is the union of $\{\mathbf{r}_k\}$ and the Hilbert basis of $\operatorname{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_{k-1})$.

Proof. Let $\mathbf{z} \in C \cap \mathbb{Z}^n$. Then $\mathbf{z} = \sum_{i=1}^k \lambda_i \mathbf{r}_i$ for some nonnegative real numbers $\lambda_1, \ldots, \lambda_k$. Multiplying by $\mathbf{c}^\intercal$, we obtain

$$\mathbf{c}^\intercal \mathbf{z} = \sum_{i=1}^k \lambda_i \mathbf{c}^\intercal \mathbf{r}_i = \lambda_k \mathbf{c}^\intercal \mathbf{r}_k = \lambda_k.$$

As $\mathbf{c}, \mathbf{z} \in \mathbb{Z}^n$, we obtain $\lambda_k \in \mathbb{Z}$. Hence, $\mathbf{z}$ is the sum of a nonnegative integer multiple of $\mathbf{r}_k$ and a lattice point $\mathbf{z} - \lambda_k \mathbf{r}_k \in \operatorname{cone}(\mathbf{r}_1, \dots, \mathbf{r}_{k-1})$, which can be written as a nonnegative integer linear combination of elements from the Hilbert basis of this cone. The result now follows. □

This lemma implies the following fact, which excludes many regions when searching for missing Hilbert basis elements.

Corollary 8. *Let $\mathbf{r}_1, \dots, \mathbf{r}_k \in \mathbb{Z}^n$ such that $C' = \operatorname{cone}(\mathbf{r}_1, \dots, \mathbf{r}_{k-1})$ has dimension n, and $C = C' + \operatorname{cone}(\mathbf{r}_k)$. Suppose that $\mathbf{r}_k \notin C'$. Moreover, let $F_1, \dots, F_q$ be the facets of C' visible from $\mathbf{r}_k$ and let $\mathbf{c}_1, \dots, \mathbf{c}_q$ the normal vectors of these facets as introduced above. Then*

$$\mathcal{H}(C') \cup \{\mathbf{r}_k\} \cup \bigcup \{\mathcal{H}(F_i + \operatorname{cone}(\mathbf{r}_k)) : |\mathbf{c}_i^\intercal \mathbf{r}_k| \geq 2,\ i = 1, \dots, q\}$$

generates $C \cap \mathbb{Z}^n$.

Proof. Evidently we obtain a system of generators of $C \cap \mathbb{Z}^n$ if we extend the union in the Corollary over all facets F_i, $i = 1, \dots, q$. It remains to observe that

$$\mathcal{H}(F_i + \operatorname{cone}(\mathbf{r}_k)) = \{\mathbf{r}_k\} \cup \mathcal{H}(C' \cap F_i)$$

if $|\mathbf{c}_i^\intercal \mathbf{r}_k| = 1$. But this is the statement of Lemma 7. □

Corollary 8 yields an extremely efficient computation of Hilbert bases—provided the case $|\mathbf{c}_i^\intercal \mathbf{r}_k| \geq 2$ occurs only rarely, or, in other words, the system $\mathbf{r}_1, \dots, \mathbf{r}_k$ of generators is not too far from a Hilbert basis.

A thoroughly consequent application of Corollary 8 could be realized as follows, collecting the list $\mathcal{A}(C)$ of critical simplicial cones in a recursive algorithm.

(1) Initially $\mathcal{A}(C)$ is empty.
(2) One searches lexicographically for the first linearly independent subset $\{\mathbf{r}_{i_1}, \dots, \mathbf{r}_{i_d}\}$. If the cone generated by these elements is not unimodular, it is added to $\mathcal{A}(C)$.
(3) Now the remaining elements among $\mathbf{r}_1, \dots, \mathbf{r}_s$ (if any) are inserted into the algorithm in ascending order. Suppose that C' is the cone generated by the elements processed already, and let $\mathbf{r}_j$ be the next element to be inserted. Then for all facets F_i of C' such that $\mathbf{c}_i^\intercal \mathbf{r}_k \geq 2$ the list $\mathcal{A}(C)$ is augmented by $\mathcal{A}(F_i + \operatorname{cone}(\mathbf{r}_j))$.

After all the critical simplicial cones have been collected, it remains to compute their Hilbert bases and to reduce their union globally, together with $\{\mathbf{r}_1, \dots, \mathbf{r}_s\}$.

Let us add some remarks on this approach.

(a) It is not hard to see that the list $\mathcal{A}(C)$ constitutes a subcomplex of the lexicographical triangulation obtained by inserting $\mathbf{r}_1, \dots, \mathbf{r}_s$. However, this fact is irrelevant for the computation of Hilbert bases.

(b) In an optimal list of simplicial cones each candidate for the Hilbert basis of C would appear exactly once. (The candidates are the elements of the Hilbert bases of the simplicial cones.) The algorithm above cannot achieve this goal since the cones $F+\mathrm{cone}(\mathbf{r}_j)$ for fixed j are treated independently of each other.

(c) The drawback of the algorithm above is that it uses the Fourier-Motzkin elimination recursively for subcones. Therefore Normaliz applies the algorithm above only on the top level and produces a full triangulation of the cones $F_i+\mathrm{cone}(\mathbf{r}_k)$ for which $\mathbf{c}_i^\intercal \mathbf{r}_k \geq 2$ (instead of the list $\mathcal{A}(F_i+\mathrm{cone}(\mathbf{r}_j))$).

(d) It is a crucial feature of the partial triangulation that it reduces memory usage drastically.

	Contingency tables				Semigraphoid
	$4\times 4\times 3$	$5\times 4\times 3$	$5\times 5\times 3$	$6\times 4\times 3$	$N=5$
emb-dim	40	47	55	54	32
dim	30	36	43	42	26
# rays	48	60	75	72	80
# HB	48	60	75	4,392	1,300
# supp hyp	4,948	29,387	306,955	153,858	117,978
# full tri	2,654,000	102,538,980	9,248,466,183	3,100,617,276	1,045,346,320
# partial tri	48	4,320	775,800	206,064	3,109,495
# cand	96	1,260	41,593	10,872	168,014
real time par.	<0.1	0.2	38	9	14
real time ser.	<0.1	1.5	813	201	225

TABLE 1. Data of challenging Hilbert basis computations

We illustrate the size of the computation and the gain of the improved algorithm by the data in Table 1. In the table we use the following abbreviations: emb-dim is the dimension of the space in which the cone (or monoid) is embedded, dim denotes its dimension, # rays is the number of extreme rays, # HB is the number of elements in the Hilbert basis, # full tri is the number of simplicial cones in a full triangulation computed by `Normaliz`, # partial tri is the number of cones in the partial triangulation, # cand is the number of candidates for the Hilbert basis, and # supp hyp is the number of support hyperplanes.

In addition to the improved algorithm just presented, parallelization has contributed substantially to the rather short computation times (given in minutes) that (the experimental version of) `Normaliz` needs for the cones considered. The computation times were measured on a SUN Fire X4450 with 24 Xeon cores where we had limited the number of threads to 1 for the strictly serial computation. Even on a single processor machine computation times are moderate, as the last line of Table 1 shows.

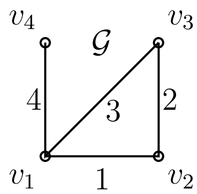
We should add that `Normaliz` cannot compute the full triangulations for $5 \times 5 \times 3$, $6 \times 4 \times 3$ and the semi-graphoid. The numbers were determined by a special program that just produced and counted the simplicial cones.

4. On a conjecture of Sturmfels and Sullivant

In this short section we report on a partial verification of a conjecture of Sturmfels and Sullivant on the normality of cut monoids of graphs.

Let $\mathcal{G}$ be a simple, undirected graph without loops on the vertex set V with edge set E. We label the edges $1, \dots, e$. A *cut* of $\mathcal{G}$ is a decomposition $V = A \cup B$ into disjoint subsets. Each cut defines a 0-1-vector $\mathbf{c}_{\{A,B\}}$ in $\mathbb{Z}^{2e}$ as follows: (i) for $j = 1, \dots, e$ the j-th entry of $\mathbf{c}_{\{A,B\}}$ is 1 if and only if the vertices x, y of edge j satisfy $\{x, y\} \subset A$ or $\{x, y\} \subset B$; (ii) for $j = e + 1, \dots, 2e$ the j-th entry of $\mathbf{c}_{\{A,B\}}$ is 1 if and only if the vertices of edge $j - e$ belong to different sets in the decomposition.

The *cut monoid* of $\mathcal{G}$ is the submonoid of $\mathbb{Z}^{2e}$ generated by the vectors $\mathbf{c}_{\{A,B\}}$ where $\{A, B\}$ runs through the cuts of $\mathcal{G}$.The eight 0-1-vectors below the figure generate the cut monoid of the graph $\mathcal{G}$:



$$\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{array} \qquad \begin{array}{cccccccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}$$

Cut monoids have been introduced to the algebraic statistics literature by Sturmfels and Sullivant [14]. They stated the very interesting conjecture that cut monoids of graphs without K_5-minors are normal. (A minor of a graph $\mathcal{G}$ is a graph $\mathcal{H}$ that can be produced from $\mathcal{G}$ by a composition of (i) deletion of a vertex and (ii) contraction of an edge.) In fact, the cut monoid of K_5 is nonnormal, this implies nonnormality for every graph with a K_5-minor. Sturmfels and Sullivant verified their conjecture for graphs with at most 6 vertices.

For graphs with 7 and 8 vertices we used the approach via partial triangulations (and parallelization) in order to verify the conjecture. We generated all these graphs (up to symmetry) with the help of nauty [10] and then excluded the graphs which have a K_5-minor. For the remaining graphs no counterexample could be found. The computations took 1 minute for 689 graphs with 7 vertices and 20 hours for 6708 graphs with 8 vertices.

For recent progress on this problem we refer the reader to Ohsugi [11].

References

[1] 4ti2 team. `4ti2`–*A software package for algebraic, geometric and combinatorial problems on linear spaces.* Available at www.4ti2.de.

[2] W. Bruns and J. Gubeladze. *Polytopes, rings, and K-theory.* Springer Monographs in Mathematics (2009).

[3] W. Bruns and B. Ichim. *Normaliz. Computing normalizations of affine semigroups. With contributions by C. Söger.* Available at http://www.math.uos.de/normaliz.

[4] W. Bruns and B. Ichim. *Normaliz: algorithms for affine monoids and rational cones.* J. Algebra, in press.

[5] M. Drton, B. Sturmfels and S. Sullivant, *Lectures on algebraic statistics.* Oberwolfach Seminars **39**, Birkhäuser (2009).

[6] A. Durand, M. Hermann and L. Juban. *On the complexity of recognizing the Hilbert Basis of a linear diophantine system.* In: Proceedings of the 24th International Symposium on Mathematical Foundations of Computer Science, LNCS **1672** (1999), 92–102.

[7] R. Hemmecke. *On the computation of Hilbert bases of cones.* In: Mathematical Software, ICMS 2002, A. M. Cohen, X.-S. Gao, N. Takayama (eds.), World Scientific (2002), 307–317.

[8] R. Hemmecke, A. Takemura, and R. Yoshida. *Computing holes in semi-groups and its application to transportation problems.* Contrib. Discrete Math. **4** (2009), 81–91.

[9] R. Hemmecke, J. Morton, A. Shiu, B. Sturmfels, and O. Wienand. *Three counterexamples on semigraphoids.* Comb. Probab. Comput. **17** (2008), 239–257.

[10] B. McKay. `nauty`–*A program for computing automorphism groups of graphs and digraphs.* Available at http://cs.anu.edu.au/∼bdm/nauty/.

[11] H. Ohsugi. *Normality of cut polytopes of graphs is a minor closed property.* Discrete Math. **310** (2010), 1160–1166.

[12] H. Ohsugi and T. Hibi. *Toric ideals arising from contingency tables.* In: Commutative Algebra and Combinatorics. Ramanujan Mathematical Society Lecture Note Series **4** (2006), 87–111.

[13] M. Studený. *Probabilistic conditional independence structures.* Springer Series in Information Science and Statistics, Springer, London, 2005.

[14] B. Sturmfels and S. Sullivant. *Toric geometry of cuts and splits.* Mich. Math. J. **57** (2008), 689–709.

[15] S. Sullivant, *Normal binary graph models.* Preprint arXiv:0906.178. To appear in Ann. Inst. Stat. Math.

[16] M. Vlach. *Conditions for the existence of solutions of the three-dimensional planar transportation problem.* Discrete Appl. Math. **13** (1986), 61–78.

Winfried Bruns, Universität Osnabrück, FB Mathematik/Informatik, 49069 Osnabrück, Germany

E-mail address: wbruns@uos.de

Raymond Hemmecke, Zentrum Mathematik, M9, Technische Universität München, Boltzmannstr. 3, 85747 Garching, Germany

E-mail address: hemmecke@ma.tum.de

Bogdan Ichim, Institute of Mathematics, C.P. 1-764, 70702 Bucharest, Romania

E-mail address: bogdan_ichim@yahoo.com

Matthias Köppe, University of California, Davis, Department of Mathematics, One Shields Avenue, Davis, CA 95616, USA

E-mail address: mkoeppe@math.ucdavis.edu

Christof Söger, Universität Osnabrück, FB Mathematik/Informatik, 49069 Osnabrück, Germany

E-mail address: csoeger@uos.de