Exploiting Polyhedral Symmetries in Social Choice[1]

Achill Schürmann

Abstract

One way of computing the probability of a specific voting situation under the Impartial Anonymous Culture assumption is via counting integral points in polyhedra. Here, Ehrhart theory can help, but unfortunately the dimension and complexity of the involved polyhedra grows rapidly with the number of candidates. However, if we exploit available polyhedral symmetries, some computations become possible that previously were infeasible. We show this in three well known examples: Condorcet's paradox, Condorcet efficiency of plurality voting and in Plurality voting vs Plurality Runoff.

1 Introduction

Under the Impartial Anonymous Culture (IAC) assumption, the probability for certain election outcomes can be computed by counting integral solutions to a system of linear inequalities, associated to the specific voting event of interest (see for example [GL11]). There exists a rich mathematical theory going back to works of Ehrhart [Ehr67] in the 1960s that helps to deal with such problems. We refer to [BR07] and [Bar08] for an introduction. The connection to Social Choice Theory was discovered by Lepelley et al. [LLS08] and Wilson and Pritchard [WP07]. A few years earlier a similar theory had been described specifically for the social choice context by Huang and Chua [HC00] (see also [Geh02]). Based on Barvinok's algorithm [Bar94] there now exists specialized mathematical software for performing previously cumbersome or practically impossible computations. The first available program was `LattE`, with its newest version `LattE integrale` (see [LDK$^+$11a]); alternatives are `barvinok` (see [VB08]) and `Normaliz` (see [BIS12]) which are also usable within the `polymake` framework (see [GJ00]).

The purpose of this note is to shed some light on the possibilities for social choice computations that arise through the use of Ehrhart theory and weighted generalizations of it (see [BBL$^+$10]). We in particular show how symmetry of linear systems characterizing certain voting events can be used to reduce computation times, and in some cases, even leads to previously unkown results. As examples, we consider three well studied voting situations with four candidates: *Condorcet's paradox*, the *Condorcet efficiency of plurality voting* and different outcomes in *Plurality vs Plurality Runoff*.

In Section 2 we review some linear models for voting events and introduce some of the used notation. In Section 3 we sketch how counting integral points in polyhedra and Ehrhart's theory can be used to compute probabilities for voting outcomes. In Section 4 we show how the complexity of computations can be reduced by using a symmetry reduced, lower dimensional reformulation. We in particular show how to use integration to obtain exact values for the limiting probability of voting outcomes when the number of voters tends to infinity. As examples, we obtain previously unknown, exact values for two four candidate election events: for the Condorcet efficiency of plurality voting and for Plurality vs Plurality Runoff.

[1] An earlier version of this paper has been published online by *Social Choice and Welfare*, April 13th 2012.

2 Linear systems describing voting situations

Notation

For the start we look at three candidate elections, as everything that will follow can best be motivated and explained in smaller examples. Assume there are n voters, with $n \geq 2$, and each of them has a complete linear (strict) preference order on the three candidates $\mathsf{a}, \mathsf{b}, \mathsf{c}$. We subdivide the voters into six groups

$$(n_{\mathsf{ab}}, n_{\mathsf{ac}}, n_{\mathsf{ba}}, n_{\mathsf{bc}}, n_{\mathsf{ca}}, n_{\mathsf{cb}})\,, \tag{1}$$

according to their six possible preference orders:

$$\mathsf{abc}(n_{\mathsf{ab}}) \qquad \mathsf{acb}(n_{\mathsf{ac}}) \qquad \mathsf{bac}(n_{\mathsf{ba}}) \qquad \mathsf{bca}(n_{\mathsf{bc}}) \qquad \mathsf{cab}(n_{\mathsf{ca}}) \qquad \mathsf{cba}(n_{\mathsf{cb}})$$

For example, there are n_{ab} voters that prefer a over b and b over c. We omit the last preference in the index, as it is determined once we know the others. This type of indexing will show to be useful when we reduce the number of variables in Section 4.

The tuple (1) is referred to as a *voting situation*. In an election with

$$n = n_{\mathsf{ab}} + n_{\mathsf{ac}} + n_{\mathsf{ba}} + n_{\mathsf{bc}} + n_{\mathsf{ca}} + n_{\mathsf{cb}} \tag{2}$$

voters, there are $\binom{n+5}{5}$ possible voting situations. We make the simplifying *Impartial Anonymous Culture (IAC) assumption* that each of these voting situations is equally likely to occur.

Condorcet's Paradox

Maybe the most famous voting paradox goes back to the Marquis de Condorcet (1743–1793). He observed that in an election with three or more candidates, it is possible that pairwise comparison of candidates can lead to an intransitive collective choice. For instance, candidate a could be preferred over candidate b, b could be preferred over candidate c and c could be preferred over candidate a. In this case there is no *Condorcet winner*, that is, someone who beats every other candidate by pairwise comparison.

The condition that candidate a is a Condorcet winner can be described via two linear constraints:

$$n_{\mathsf{ab}} + n_{\mathsf{ac}} + n_{\mathsf{ca}} \;>\; n_{\mathsf{ba}} + n_{\mathsf{bc}} + n_{\mathsf{cb}} \qquad (3) \qquad (\ \mathsf{a} \text{ beats } \mathsf{b}\ )$$

$$n_{\mathsf{ab}} + n_{\mathsf{ac}} + n_{\mathsf{ba}} \;>\; n_{\mathsf{ca}} + n_{\mathsf{bc}} + n_{\mathsf{cb}} \qquad (4) \qquad (\ \mathsf{a} \text{ beats } \mathsf{c}\ )$$

The probability of candidate a being a Condorcet winner in an election with n voters can be expressed as the quotient

$$\mathrm{Prob}(n) \;=\; \frac{\mathrm{card}\left\{(n_{\mathsf{ab}}, \ldots, n_{\mathsf{cb}}) \in \mathbb{Z}^6_{\geq 0} \text{ satisfying (2), (3), (4) }\right\}}{\binom{n+5}{5}}.$$

The denominator is a polynomial of degree 5 in n. It had been observed by Fishburn and Gehrlein [GF76] (cf. [BB83]) that the numerator shows a similar behavior: Restricting to even or odd n it can be expressed as a degree 5 polynomial in n. The leading coefficient of both polynomials is the same and we approach the same probability for large elections (as n tends to infinity). This *limiting probability* is known to be

$$\lim_{n\to\infty} \mathrm{Prob}(n) = \frac{5}{16}.$$

Having the probability for candidate a being a Condorcet winner, we obtain the probability for a Condorcet paradox (no Condorcet winner exists) as $1 - 3 \cdot \mathrm{Prob}(n)$ with an exact limiting probability of $\frac{1}{16}$.

In a similar way we can determine probabilities for other voting events.

Condorcet efficiency of Plurality voting

If there is a Condorcet winner, there is good reason to consider him to be the voter's choice. However, many common voting rules do not always choose the Condorcet winner even if one exists. This is in particular the case for the widely used plurality voting, where the candidate with a majority of first preferences is elected.

The condition that candidate a is a Condorcet winner but candidate b is the plurality winner can be expressed by the two inequalities (3) and (4), together with the two additional inequalities

$$
\begin{array}{rcllr}
n_{\mathsf{ba}} + n_{\mathsf{bc}} & > & n_{\mathsf{ab}} + n_{\mathsf{ac}} & (5) & (\text{ b wins plurality over a }) \\
n_{\mathsf{ba}} + n_{\mathsf{bc}} & > & n_{\mathsf{ca}} + n_{\mathsf{cb}} & (6) & (\text{ b wins plurality over c })
\end{array}
$$

The *Condorcet efficiency* of a voting rule is the conditional probability that a Condorcet winner is elected if one exists. As there are $3 \cdot 2$ possibilities for choosing a Condorcet winner and another plurality winner, we obtain

$$
\mathrm{Prob}(n) \;=\; \frac{6 \cdot \mathrm{card}\left\{(n_{\mathsf{ab}}, \ldots, n_{\mathsf{cb}}) \in \mathbb{Z}^6_{\geq 0} \text{ satisfying (2), (3), (4), (5), (6)}\right\}}{3 \cdot \mathrm{card}\left\{(n_{\mathsf{ab}}, \ldots, n_{\mathsf{cb}}) \in \mathbb{Z}^6_{\geq 0} \text{ satisfying (2), (3), (4)}\right\}}
$$

for the likelihood of a Condorcet winner being a plurality loser. Again, depending on n being odd or even, one obtains polynomials in n in the denominator and the numerator (see [Geh82]). The exact value of the limit $\lim_{n\to\infty} \mathrm{Prob}(n)$ is $16/135$. Therefore, the Condorcet efficiency of plurality voting with three candidates is $119/135 = 88.\overline{148}\%$.

Plurality vs Plurality Runoff

Plurality Runoff voting is a common practice to overcome some of these "problems" of Plurality voting. It is used in many presidential elections, for example in France. After a first round of plurality voting in which none of the candidates has achieved more than 50% of the votes, the first two candidates compete in a second runoff round.

The condition that candidate b is the plurality winner, but candidate a wins the second round of Plurality Runoff can be expressed by the two inequalities (5) and

$$
n_{\mathsf{ab}} + n_{\mathsf{ac}} > n_{\mathsf{ca}} + n_{\mathsf{cb}}, \qquad (7) \qquad (\text{ a wins plurality over c })
$$

together with the linear condition (3) that a beats b in a pairwise comparison. The probability that another candidate is chosen in the second round as the number of voters tends to infinity is known to be $71/576 = 12.3263\overline{8}\%$ (see [LLS08]).

Four and more candidates

Having m candidates we can set up similar linear systems in $m!$ variables. For example, in an election with four candidates a, b, c, d we use the 24-dimensional vector $x^t = (n_{\mathsf{abc}}, \ldots, n_{\mathsf{dcb}})$. Here, indices are taken in lexicographical order. The condition that a is a Condorcet winner is described by the three inequalities that imply a beats b, a beats c and a beats d in a pairwise comparison. As linear systems with 24 variables become hard to grasp, it is convenient to use matrices for their description. We are interested in all non-negative integral (column) vectors x satisfying the matrix inequality $Ax > 0$ for the matrix $A \in \mathbb{Z}^{3\times 24}$ with entries

```
     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
(8)  1  1  1  1  1  1  1  1  1 -1 -1 -1  1  1  1 -1 -1 -1 -1 -1 -1 -1 -1 -1
```

3 Likelihood of voting situations and Ehrhart's theory

Integral points in polyhedral cones

In order to deal with an arbitrary number of candidates, let us put the example above in a slightly more general context. In any of the three voting examples, the voting situations of interest lie in a *polyhedral cone*, that is, in a set $\mathcal{P}$ of points in $\mathbb{R}^d$ (with $d = 6$ or $d = 24$ in case of three or four candidate elections) satisfying a finite number of homogeneous linear inequalities. In addition to the strict inequalities which are different in each of the examples, the condition that the variables n_i are non-negative can be expressed by the homogeneous linear inequalities $n_i \geq 0$.

Let $\mathcal{P}, \mathcal{S} \subset \mathbb{R}^d$ denote two d-dimensional *polyhedral cones*, each defined by some homogeneous linear (possibly strict) inequalities. We may assume that $\mathcal{P}$ is contained in $\mathcal{S}$ and that both polyhedral cones are contained in the orthant $\mathbb{R}^d_{\geq 0}$. If we are interested in elections with n voters, we consider the voting situations (integral vectors) in the intersection of $\mathcal{P}$ and $\mathcal{S}$ with the affine subspace

$$L_n^d = \left\{ (n_1, \ldots, n_d) \in \mathbb{R}^d \mid \sum_{i=1}^{d} n_i = n \right\}.$$

The *expected frequency* of voting situations being in $\mathcal{P}$ among voting situations in $\mathcal{S}$ is then expressed by

$$\mathrm{Prob}(n) = \frac{\mathrm{card}\left(\mathcal{P} \cap L_n^d \cap \mathbb{Z}^d\right)}{\mathrm{card}\left(\mathcal{S} \cap L_n^d \cap \mathbb{Z}^d\right)}. \tag{9}$$

When estimating the probability of candidate a being a Condorcet winner for instance, the homogeneous polyhedral cone $\mathcal{S}$ is simply the non-negative orthant $\mathbb{R}^d_{\geq 0}$ described by the linear inequalities $n_i \geq 0$. In that case the denominator is known to be equal to

$$\binom{n+d-1}{d-1}.$$

This is a polynomial in n of degree $d-1$ (the dimension of $L_n^d \cap \mathcal{S}$).

Ehrhart theory

By Ehrhart's theory, the number of integral solutions in a polyhedral cone intersected with L_n^d can be expressed by a *quasi-polynomial* in n. Roughly speaking, a quasi-polynomial is simply a finite collection $p_1(n), \ldots, p_k(n)$ of polynomials, such that the number of voting situations is given by $p_i(n)$ if $i \equiv n \mod k$.

The degree of the polynomial is equal to the dimension of the polyhedral cone intersected with L_n^d. In the voting events considered here this dimension is always equal to $d-1$. So in the examples with three candidates their degree is always 5. The number k of different polynomials depends on the linear inequalities involved. For the Condorcet paradox we have $k = 2$ polynomials $p_1(n)$ and $p_2(n)$, where $p_1(n)$ gives the answer for odd n ($1 \equiv n \mod 2$) and $p_2(n)$ gives the answer for even n ($0 \equiv 2 \equiv n \mod 2$). For Condorcet efficiency we have $k = 6$ (see [Geh02]) and for Plurality vs Plurality Runoff we have $k = 12$ (see [LLS08]).

Given a polyhedral cone $\mathcal{P}$, the quasi-polynomial $q(n) = \mathrm{card}\left(\mathcal{P} \cap L_n^d \cap \mathbb{Z}^d\right)$ can be explicitly computed using software packages like `LattE integrale` [latte] or `barvinok` [`barvinok`]. The result for the polyhedral cone $\mathcal{P}$ describing candidate a as the Condorcet winner could look like

```
  1/384 * n^5
+ ( 1/64 * { 1/2 * n } + 1/32 ) * n^4
+ ( 17/96 * { 1/2 * n } + 13/96 ) * n^3
+ ( 23/32 * { 1/2 * n } + 1/4 ) * n^2
+ ( 233/192 * { 1/2 * n } + 1/6 ) * n
+ ( 45/64 * { 1/2 * n } + 0 )
```

The curly brackets $\{\cdots\}$ mean the fractional part of the enclosed number, allowing to write the quasi-polynomial in a closed form. In this example we get different polynomials for odd and even n. Note that the leading coefficient (the coefficient of n^5) is in both cases the same. By Ehrhart's theory this is always the case, as it is equal to the *relative volume* of the polyhedron $\mathcal{P} \cap L_1^d$. That is, it is equal to a $\sqrt{d}$-multiple of the standard Lebesgue measure on the affine space L_1^d. The measure is normalized so that the space contains one integral point per unit volume.

One technical obstacle using software like `LattE integrale` or `barvinok` is the use of polyhedral cones described by a mixture of strict and non-strict inequalities. As the software assumes the input to have only non-strict inequalities or equality conditions, one has to avoid the use of strict inequalities. A simple way to achieve this is the replacement of strict inequalities $x > 0$ by non-strict ones $x \geq 1$, in case x is known to be integral. For instance, if x is a linear expression with integer coefficients, and if we are interested in integral solutions as in our examples, this is a possible reformulation.

Altogether, by obtaining quasi-polynomials for numerator and denominator in (9) we get an explicit formula for $\mathrm{Prob}(n)$ via Erhart's theory.

Limiting probabilities via integration

If we want to compute the exact value of $\lim_{n\to\infty} \mathrm{Prob}(n)$ as n tends to infinity, we can use volume computations without using Ehrhart's theory. As mentioned above, the leading coefficients of denominator and numerator correspond to the relative volumes of the sets $\mathcal{P} \cap L_1$ and $\mathcal{S} \cap L_1$:

$$\lim_{n\to\infty} \mathrm{Prob}(n) \;=\; \lim_{n\to\infty} \frac{\mathrm{card}\left(\mathcal{P} \cap L_1^d \cap (\mathbb{Z}/n)^d\right)}{\mathrm{card}\left(\mathcal{S} \cap L_1^d \cap (\mathbb{Z}/n)^d\right)} \;=\; \frac{\mathrm{relvol}\left(\mathcal{P} \cap L_1^d\right)}{\mathrm{relvol}\left(\mathcal{S} \cap L_1^d\right)}$$

In fact, as long as we use the same measure to evaluate the numerator and the denominator, it does not matter what multiple of the standard Lebesgue measure we use to compute volume on the affine space L_1^d. The exact relative volume can be computed using `LattE integrale`. Alternatives are for example `Normaliz` (see [`normaliz`]) or `vinci` (see [BEF00]). Exact computations can be quite involved in higher dimensions (cf. [DF88]). In such cases it is sometimes only possible to compute an approximation, using *Monte Carlo methods* for instance.

4 Reducing the dimension by exploiting polyhedral symmetries

In many models the involved linear systems and polyhedra are quite symmetric. In particular, permutations of variables may lead to equivalent linear systems describing the same polyhedron. Such symmetries are often visible in smaller examples and can automatically be determined for larger problems, for instance by our software `SymPol` (see [RS10]). In the three examples described in Section 2, we can exploit such symmetries to reduce the complexity of computations.

Condorcet's paradox

In case of a being a Condorcet winner in a three candidate election, the variables n_{ab} and n_{ac} occur pairwise (as $n_{\mathsf{ac}} + n_{\mathsf{ab}}$) in inequalities (3), (4) and in equation (2). The same is true for n_{bc} and n_{cb}. By introducing new variables $n_{\mathsf{a}} = n_{\mathsf{ac}} + n_{\mathsf{ab}}$ and $n_{*\mathsf{a}} = n_{\mathsf{bc}} + n_{\mathsf{cb}}$ we can reduce the dimension of the linear system to only four variables:

$$\begin{aligned} n_{\mathsf{a}} + n_{\mathsf{ca}} - n_{*\mathsf{a}} - n_{\mathsf{ba}} &> 0 \\ n_{\mathsf{a}} + n_{\mathsf{ba}} - n_{*\mathsf{a}} - n_{\mathsf{ca}} &> 0 \\ n_{\mathsf{a}} + n_{\mathsf{ca}} + n_{*\mathsf{a}} + n_{\mathsf{ba}} &= n \\ n_{\mathsf{a}}, n_{*\mathsf{a}}, n_{\mathsf{ba}}, n_{\mathsf{ca}} &\geq 0. \end{aligned}$$

The index a indicates that we group all variables which carry candidate a as their first preference and index $*$a stands for grouping of all variables with candidate a ranked last. In the reduced linear system each 4-tuple $(n_{\mathsf{a}}, n_{*\mathsf{a}}, n_{\mathsf{ba}}, n_{\mathsf{ca}})$ represents several voting situations, previously described by 6-tuples. For n_{a} we have $(n_{\mathsf{a}} + 1)$ different possibilities of non-negative integral tuples $(n_{\mathsf{ac}}, n_{\mathsf{ab}})$. Similar is true for $n_{*\mathsf{a}}$. Together we have

$$(n_{\mathsf{a}} + 1)(n_{*\mathsf{a}} + 1)$$

voting situations with three candidates represented by each non-negative integral vector $(n_{\mathsf{a}}, n_{*\mathsf{a}}, n_{\mathsf{ba}}, n_{\mathsf{ca}})$.

In the four candidate case it is possible to obtain a similar reformulation by grouping among 24 variables. We introduce a new variable for sets of variables having same coefficients in the linear system. Having a matrix representation as in (8), this kind of special symmetry in the linear system is easy to find by identifying equal columns. Introducing a new variable for each set of equal columns we get

$$\begin{aligned} n_{\mathsf{a}} - n_{\mathsf{ba}} + n_{\mathsf{ca}} + n_{\mathsf{da}} + n_{*\mathsf{ab}} - n_{*\mathsf{ac}} - n_{*\mathsf{ad}} - n_{*\mathsf{a}} &> 0 \\ n_{\mathsf{a}} + n_{\mathsf{ba}} - n_{\mathsf{ca}} + n_{\mathsf{da}} - n_{*\mathsf{ab}} + n_{*\mathsf{ac}} - n_{*\mathsf{ad}} - n_{*\mathsf{a}} &> 0 \\ n_{\mathsf{a}} + n_{\mathsf{ba}} + n_{\mathsf{ca}} - n_{\mathsf{da}} - n_{*\mathsf{ab}} - n_{*\mathsf{ac}} + n_{*\mathsf{ad}} - n_{*\mathsf{a}} &> 0 \end{aligned} \tag{10}$$

These three inequalities describe voting situations in which candidate a beats candidates b, c and d each in a pairwise comparison. As in all of our examples, we additionally have the condition that the involved variables add up to n and that all of them are non-negative.

As before, the used indices of variables reflect which voter preferences are grouped. As in the three candidate case, n_{a} and $n_{*\mathsf{a}}$ denote the number of voters with candidate a being their first and last preference respectively. Similarly, xy and $*$yx in the index indicate that voters with preference order starting with x, y and ending with y, x have been combined.

Using our software SymPol [sympol] one easily checks that the original system with 24 variables has a symmetry group of order 199065600. The new reduced system with 8 variables, obtained through the above grouping of variables, turns out to have a symmetry group of order 6 only. So most of the symmetry in the original system is of the simple form that is detectable through equal columns in a matrix representation. The remaining 6-fold symmetry comes from the possibility to arbitrarily permute the variables $n_{\mathsf{ba}}, n_{\mathsf{ca}}, n_{\mathsf{da}}$ when at the same time the variables $n_{*\mathsf{ab}}, n_{*\mathsf{ac}}, n_{*\mathsf{ad}}$ are permuted accordingly. This symmetry is due to the fact that candidates b, c and d are equally treated in the linear system (10). The two new variables n_{a} and $n_{*\mathsf{a}}$ each combine six of the former variables. The other six new variables each combine two former ones.

Weighted counting

In general, if we group more than two variables, say if we substitute the sum of k variables $n_1 + \ldots + n_k$ by a new variable N, we have to include a factor of

$$\binom{N+k-1}{k-1}$$

when counting voting situations via N. If we substitute d variables $(n_1, \ldots, n_d)$ by D new variables $(N_1, \ldots, N_D)$, say by setting N_i to be the sum of k_i of the variables n_j, for $i = 1, \ldots, D$, then we count for each D-tuple

$$p(N_1, \ldots, N_D) = \prod_{i=1}^{D} \binom{N_i + k_i - 1}{k_i - 1} \tag{11}$$

many voting situations.

In the example above, with four candidates and candidate a being the Condorcet winner, we have $d = 24$, $D = 8$ and we obtain a degree 16 polynomial

$$\binom{n_{\mathsf{a}}+5}{5}(n_{\mathsf{ba}}+1)(n_{\mathsf{ca}}+1)(n_{\mathsf{da}}+1)(n_{*\mathsf{ab}}+1)(n_{*\mathsf{ac}}+1)(n_{*\mathsf{ad}}+1)\binom{n_{*\mathsf{a}}+5}{5}$$

to count voting situations for each 8-tuple

$$(n_{\mathsf{a}}, n_{\mathsf{ba}}, n_{\mathsf{ca}}, n_{\mathsf{da}}, n_{*\mathsf{ab}}, n_{*\mathsf{ac}}, n_{*\mathsf{ad}}, n_{*\mathsf{a}}) \,.$$

Geometrically, the polyhedral cone $\mathcal{P} \subset \mathbb{R}^d$ is replaced by a new polyhedral cone $\mathcal{P}' \subset \mathbb{R}^D$ in a lower dimension. As the counting is changed we obtain for the probability (9) of voting situations in $\mathcal{P}$ among those in $\mathcal{S}$:

$$\mathrm{Prob}(n) \;=\; \frac{\displaystyle\sum_{x \in \mathcal{P} \cap L_n^d \cap \mathbb{Z}^d} 1}{\displaystyle\sum_{x \in \mathcal{S} \cap L_n^d \cap \mathbb{Z}^d} 1} \;=\; \frac{\displaystyle\sum_{y \in \mathcal{P}' \cap L_n^D \cap \mathbb{Z}^D} p(y)}{\displaystyle\sum_{y \in \mathcal{S}' \cap L_n^D \cap \mathbb{Z}^D} p(y)}. \tag{12}$$

Here, $\mathcal{S}'$ is equal to the corresponding homogeneous polyhedral cone obtained from $\mathcal{S} \subset \mathbb{R}^d$, and $p(y)$ is the polynomial (11) in D variables. In the example of Condorcet's paradox, $\mathcal{S}'$ is simply equal to the full orthant $\mathbb{R}^D_{\geq 0}$.

As seen in Section 3, we can use Ehrhart's theory to determine an explicit formula for $\mathrm{Prob}(n)$. The right hand side of the formula above suggests that we can do this also via *weighted lattice point counting* in dimension D. A corresponding Ehrhart-type theory has recently been considered (see [BBL+10]). A first implementation is available in the package `barvinok` via the command `barvinok_summate`. We successfully tested the software on some reformulations of three candidate elections, but so far `barvinok` seems not capable to do computations for the four candidate case. However, there still seems quite some improvement possible in the current implementation (personal communication with Sven Verdoolaege). It can be expected that future versions of `LattE integrale` will be capable of these computations (personal communication with Matthias Köppe). It appears to be "just" a matter of implementing the ideas in [BBL+10].

We note that, theoretically, it can generally be expected that weighted counting over a smaller dimensional polyhedron is faster than unweighted counting over a corresponding high dimensional polyhedron. However, due to fact that a suitable implementation for weighted counting is not available at the moment, latter approach may practically still be a good choice. For instance, the latest version of `Normaliz` (July 2012) appears to be capable to obtain the Ehrhart quasi-polynomials for the 23-dimensional polyhedra considered in this note (personal communication with Winfried Bruns and Bogdan Ichim).

Limiting probabilities via integration

If we want to compute the exact value of $\lim_{n\to\infty} \operatorname{Prob}(n)$ we may use integration. Using (12) we get through substitution of $y = nz$:

$$\lim_{n\to\infty} \operatorname{Prob}(n) = \lim_{n\to\infty} \frac{\sum\limits_{y\in\mathcal{P}'\cap L_n^D\cap\mathbb{Z}^D} p(y)}{\sum\limits_{y\in\mathcal{S}'\cap L_n^D\cap\mathbb{Z}^D} p(y)} = \lim_{n\to\infty} \frac{\sum\limits_{z\in\mathcal{P}'\cap L_1^D\cap(\mathbb{Z}/n)^D} p(nz)}{\sum\limits_{z\in\mathcal{S}'\cap L_1^D\cap(\mathbb{Z}/n)^D} p(nz)}$$

$$= \lim_{n\to\infty} \frac{\sum\limits_{z\in\mathcal{P}'\cap L_1^D\cap(\mathbb{Z}/n)^D} p(nz)/n^{\deg p}}{\sum\limits_{z\in\mathcal{S}'\cap L_1^D\cap(\mathbb{Z}/n)^D} p(nz)/n^{\deg p}} = \frac{\int_{\mathcal{P}'\cap L_1^D} \operatorname{lt}(z)\, dz}{\int_{\mathcal{S}'\cap L_1^D} \operatorname{lt}(z)\, dz}.$$

Here, the division of numerator and denominator by a degree of p ($\deg p$) power of n shows that the integrals on the right are taken over the leading term $\operatorname{lt}(z)$ of the polynomial $p(z)$ only. Thus determining the exact limiting probability is achieved by integrating a degree $d - D$ monomial over a bounded polyhedron (*polytope*) in the $(D-1)$-dimensional affine space L_1^D. We refer to [LDK+11b] for background on efficient integration methods (cf. [BBL+11] and [Sch98]).

As in the case of relative volume computations in dimension d, the integral is taken with respect to the relative Lebesgue measure – here on the affine space L_1^D. In fact, as we are computing a quotient, any measure being a multiple of the standard Lebesgue measure on L_1^D will give the same value.

For the example with candidate a being a Condorcet winner in a four candidate election, the leading term to be integrated is

$$n_{\mathsf{a}}^5 \cdot n_{\mathsf{ba}} \cdot n_{\mathsf{ca}} \cdot n_{\mathsf{da}} \cdot n_{*\mathsf{ab}} \cdot n_{*\mathsf{ac}} \cdot n_{*\mathsf{ad}} \cdot n_{*\mathsf{a}}^5,$$

which is much simpler than the full polynomial. Integrating this polynomial over the reduced 8-dimensional polyhedron can be done using `LattE integrale` (called with option `valuation=integrate`). In this way one obtains in a few seconds an exact value of 1717/2048 for the probability that a Condorcet winner exists (as n tends to infinity). This value corresponds to the one obtained by Gehrlein in [Geh01] and serves as a test case for our method. The corresponding volume computation with `LattE integrale` (called with option `valuation=volume`) in 24 variables did not finish after several weeks of computation. This is due to the fact that triangulating a 24-dimensional polyhedron is much more involved than integration over a corresponding lower dimensional polyhedron (of dimension 8 in this case). However, Winfried Bruns, Bogdan Ichim and Christof Söger report (May 2012) that the 24-dimensional volume computation is doable with the newest version of their software `Normaliz` (see [`normaliz`]). Nevertheless, their volume computation, using sophisticated heuristics for triangulations (see [BIS12]), is still much slower than the corresponding integration over the 8-dimensional polyhedron.

In a similar way we can deal with other voting situations as well.

Condorcet efficiency of plurality voting

Assuming candidate a is a Condorcet winner, but candidate b wins a plurality voting, we obtain a reduced system in the three candidate case with five variables:

$$\begin{aligned}
n_{\mathsf{a}} - n_{\mathsf{ba}} - n_{\mathsf{bc}} - n_{\mathsf{cb}} + n_{\mathsf{ca}} &> 0 \\
n_{\mathsf{a}} + n_{\mathsf{ba}} - n_{\mathsf{bc}} - n_{\mathsf{cb}} - n_{\mathsf{ca}} &> 0 \\
-n_{\mathsf{a}} + n_{\mathsf{ba}} + n_{\mathsf{bc}} &> 0 \\
n_{\mathsf{ba}} + n_{\mathsf{bc}} - n_{\mathsf{cb}} - n_{\mathsf{ca}} &> 0
\end{aligned}$$

Here the only reduction is the grouping $n_{\mathsf{a}} = n_{\mathsf{ab}} + n_{\mathsf{ac}}$. The corresponding polynomial weight is $n_{\mathsf{a}} + 1$.

The four candidate case is more involved. The linear system with 24 variables has a comparatively small symmetry group of order 92160. We can group six variables into n_{a}. Taking the reduced system (10) of three inequalities with 8 variables (modeling that candidate a is a Condorcet winner) we have to add three inequalities for the condition that candidate b wins plurality. These can be shortly described by $n_{\mathsf{b}} > n_{\mathsf{a}}, n_{\mathsf{c}}, n_{\mathsf{d}}$, but a grouping of variables in n_{b}, n_{c} and n_{d} is incompatible with the other three conditions. Instead we use new variables $n_{\mathsf{b*a}}$, $n_{\mathsf{c*a}}$ and $n_{\mathsf{d*a}}$ (in (10) combined in $n_{\mathsf{*a}}$) for preferences in which a is ranked last. Additionally we have to keep the variables where candidate a is ranked third (in (10) combined in $n_{\mathsf{*ab}}$, $n_{\mathsf{*ac}}$, $n_{\mathsf{*ad}}$).

In the three inequalities (10) we can simply substitute $n_{\mathsf{*a}}$ by $n_{\mathsf{b*a}} + n_{\mathsf{c*a}} + n_{\mathsf{d*a}}$ and $n_{\mathsf{*ad}}$, $n_{\mathsf{*ac}}$ and $n_{\mathsf{*ab}}$ by $n_{\mathsf{bca}} + n_{\mathsf{cba}}$, $n_{\mathsf{bda}} + n_{\mathsf{dba}}$ and $n_{\mathsf{cda}} + n_{\mathsf{dca}}$. The additional three linear inequalities for candidate b being a plurality winner are then:

$$\begin{aligned}
n_{\mathsf{b*a}} + n_{\mathsf{ba}} + n_{\mathsf{bca}} + n_{\mathsf{bda}} - n_{\mathsf{a}} &> 0 \\
n_{\mathsf{b*a}} + n_{\mathsf{ba}} + n_{\mathsf{bca}} + n_{\mathsf{bda}} - n_{\mathsf{c*a}} - n_{\mathsf{ca}} - n_{\mathsf{cba}} - n_{\mathsf{cda}} &> 0 \\
n_{\mathsf{b*a}} + n_{\mathsf{ba}} + n_{\mathsf{bca}} + n_{\mathsf{bda}} - n_{\mathsf{d*a}} - n_{\mathsf{da}} - n_{\mathsf{dba}} - n_{\mathsf{dca}} &> 0
\end{aligned}$$

This reduced linear system has 6 inequalities for 13 variables. It still has a symmetry of order 2 coming from an interchangeable role of candidates c and d. The degree 11 polynomial used for integration is

$$n_{\mathsf{a}}^5 \cdot n_{\mathsf{ba}} \cdot n_{\mathsf{ca}} \cdot n_{\mathsf{da}} \cdot n_{\mathsf{b*a}} \cdot n_{\mathsf{c*a}} \cdot n_{\mathsf{d*a}}.$$

With it, using `LattE integrale`, we obtain an exact limit of

$$\frac{10658098255011916449318509}{14352135440302080000000000} = 74.261410\ldots\%$$

for the Condorcet efficiency of plurality voting with four candidates. To the best of our knowledge this value has not been computed before.

Plurality vs Plurality Runoff

The case of Plurality vs Plurality Runoff has a high degree of symmetry. For three candidates we obtain a reduced four dimensional reformulation:

$$\begin{aligned}
n_{\mathsf{b}} - n_{\mathsf{a}} &> 0 \\
n_{\mathsf{a}} - n_{\mathsf{ca}} - n_{\mathsf{cb}} &> 0 \\
n_{\mathsf{a}} + n_{\mathsf{ca}} - n_{\mathsf{b}} - n_{\mathsf{cb}} &> 0
\end{aligned}$$

Counting is done via the polynomial weight $(n_{\mathsf{a}} + 1)(n_{\mathsf{b}} + 1)$. Integration of $n_{\mathsf{a}} n_{\mathsf{b}}$ over the corresponding 3-dimensional polyhedron yields the known limiting probability.

If we consider elections with m candidates, $m \geq 4$, we can set up a linear system with only $2(m-1)$ variables and m inequalities. We denote the candidates by a, b and $\mathsf{c_i}$ for $i = 1, \ldots, m-2$:

$$\begin{aligned} n_{\mathsf{b}} - n_{\mathsf{a}} &> 0 \\ \text{For } i = 1, \ldots, m-2: \qquad n_{\mathsf{a}} - n_{\mathsf{c_i \cdot a \cdot b}} - n_{\mathsf{c_i \cdot b \cdot a}} &> 0 \\ n_{\mathsf{a}} + \sum_{i=1}^{m-2} n_{\mathsf{c_i \cdot a \cdot b}} - n_{\mathsf{b}} - \sum_{i=1}^{m-2} n_{\mathsf{c_i \cdot b \cdot a}} &> 0 \end{aligned}$$

The first two lines model that candidate b wins plurality over candidate a and that candidate a is second, winning over candidates $\mathsf{c_i}$, for $i = 1, \ldots, m-2$. The last inequality models the condition that candidate a beats b in a pairwise comparison. The variable $n_{\mathsf{c_i \cdot a \cdot b}}$ gives the number of voters with candidate $\mathsf{c_i}$ being their first preference and candidate a being ranked before candidate b. Similarly, $n_{\mathsf{c_i \cdot b \cdot a}}$ is the number of voters with first preference $\mathsf{c_i}$ and candidate b being ranked before candidate a. We use "$\cdot$" to denote any ordering of candidates; in contrast to "$*$" used before we also allow an empty list here. For both variables, $n_{\mathsf{c_i \cdot a \cdot b}}$ and $n_{\mathsf{c_i \cdot b \cdot a}}$, we group $(m-1)!/2$ of the $m!$ former variables. The new variables n_{a} and n_{b} both represent $(m-1)!$ former variables. Therefore, counting is adapted using the polynomial weight

$$(n_{\mathsf{a}} \cdot n_{\mathsf{b}})^{(m-1)!-1} \cdot \prod_{i=1}^{m-2} (n_{\mathsf{c_i \cdot a \cdot b}} \cdot n_{\mathsf{c_i \cdot b \cdot a}})^{(m-1)!/2-1}$$

of degree $m! - 2m + 2$.

The above inequalities assume that candidates b and a are ranked first and second in a plurality voting. So having the probability for the corresponding voting situations, we have to multiply by $m(m-1)$ to get the overall probability of a plurality winner losing in a second Plurality Runoff round.

For four candidates ($m = 4$) we obtain an exact limiting probability of

$$\frac{2988379676768359}{12173449145352192} = 24.548339\ldots\%.$$

This result can be obtained using the weighted, dimension-reduced problem with `LattE integrale`, as well as by a relative volume computation in 24 variables. However, the latter is a few hundred times slower than integration over the dimension reduced polyhedron. A similar result from a volume computation is obtained in [LDK$^+$11b].

To be certain about our new results, we computed the value above, as well as the likelihood for the existence of a Condorcet winner, with a fully independent `Maple` calculation, using the package `Convex` (see [`convex`]). For it, we first obtained a *triangulation* (non-overlapping union of *simplices*) of the dimension-reduced polyhedron and then applied symbolic integration to each simplex.

We also tried to solve the five candidate case, where the polyhedron is only 7-dimensional (in 8 variables). The integration of a polynomial of degree 112, however, seems a bit too difficult for the currently available technology. Nevertheless it seems that we are close to obtain exact five candidate results as well.

5 Conclusions

Using symmetry of linear systems we can obtain symmetry reduced lower dimensional reformulations. These allow to compute exact limiting probabilities for large elections with four candidates. In this work we only gave a few starting examples. Similar calculations are possible for many other voting situations as well. Even during the work on this project, the software packages `LattE integrale` and `Normaliz` for corresponding polyhedral computations have introduced substantial improvements. We can look forward to capabilities of future versions.

At the moment, for elections with five or more candidates further ideas seem necessary. One possibility to reduce the complexity of computations further is the use of additional symmetries which remain in our reduced systems.

References

[Bar94] A. Barvinok, *A polynomial time algorithm for counting integral points in polyhedra when the dimension is fixed*, Math. Oper. Res. **19** (1994), no. 4, 769–779.

[Bar08] ——, *Integer points in polyhedra*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008.

[BBL+10] V. Baldoni, N. Berline, J.A. De Loera, M. Köppe, and M. Vergne, *Computation of the highest coefficients of weighted Ehrhart quasi-polynomials of rational polyhedra*, Found. Comput. Math., to appear, preprint at http://arxiv.org/abs/1011.1602v1.

[BBL+11] ——, *How to integrate a polynomial over a simplex*, Math. Comp. **80** (2011), no. 273, 297–325.

[BB83] S. Berg and B. Bjurulf, *A note on the paradox of voting: anonymous preference profiles and May's formula*, Public Choice **40** (1983), 307–316.

[BIS12] W. Bruns, B. Ichim and C. Sger, *The power of pyramid decomposition in Normaliz*, preprint at `http://arxiv.org/abs/1206.1916`.

[BEF00] B. Büeler, A. Enge, and K. Fukuda, *Exact volume computation for polytopes: a practical study*, Polytopes—combinatorics and computation (Oberwolfach, 1997), DMV Sem., vol. 29, Birkhäuser, Basel, 2000, pp. 131–154.

[BR07] M. Beck and S. Robins, *Computing the continuous discretely*, Undergraduate Texts in Mathematics, Springer, New York, 2007, Integer-point enumeration in polyhedra.

[LDK+11a] J.A. De Loera, B. Dutra, M. Köppe, S. Moreinis, G. Pinto, and J. Wu, *A users guide for* `LattE integrale` *v1.5*, 2011, available at `http://www.math.ucdavis.edu/~latte/`.

[LDK+11b] ——, *Software for Exact Integration of Polynomials over Polyhedra*, 2011, preprint at http://arxiv.org/abs/1108.0117v2.

[DF88] M.E. Dyer and A.M. Frieze, *On the complexity of computing the volume of a polyhedron*, SIAM J. Comput. **17** (1988), no. 5, 967–974.

[Ehr67] E. Ehrhart, *Sur un problème de géométrie diophantienne linéaire. I. Polyèdres et réseaux*, J. Reine Angew. Math. **226** (1967), 1–29.

[Geh82] W.V. Gehrlein, *Condorcet efficiency of constant scoring rules*, Mathematical Social Sciences **2** (1982), no. 2, 123–130.

[Geh01] ———, *Condorcet winners on four candidates with anonymous voters*, Economics Letters **71** (2001), 335–340.

[Geh02] ———, *Obtaining representations for probabilities of voting outcomes with effectively unlimited precision integer arithmetic*, Soc. Choice Welf. **19** (2002), no. 3, 503–512.

[GF76] W.V. Gehrlein and P.C. Fishburn, *The probability of the paradox of voting: a computable solution*, J. Econom. Theory **13** (1976), no. 1, 14–25.

[GJ00] E. Gawrilow and M. Joswig, *polymake: a framework for analyzing convex polytopes*, Polytopes — Combinatorics and Computation (Gil Kalai and Günter M. Ziegler, eds.), Birkhäuser, 2000, pp. 43–74.

[GL11] W.V. Gehrlein and D. Lepelley, *Voting paradoxes and group coherence*, Studies in Choice and Welfare, Springer, Heidelberg, 2011, The Condorcet efficiency of voting rules.

[HC00] H.C. Huang and Vincent C.H. Chua, *Analytical representation of probabilities under the IAC condition*, Soc. Choice Welf. **17** (2000), no. 1, 143–155.

[LLS08] D. Lepelley, A. Louichi, and H. Smaoui, *On Ehrhart polynomials and probability calculations in voting theory*, Soc. Choice Welf. **30** (2008), no. 3, 363–383.

[RS10] T. Rehn and A. Schürmann, *C++ tools for exploiting polyhedral symmetries*, Mathematical Software ICMS 2010, Lecture Notes in Comput. Sci., vol. 6327, Springer, Berlin, 2010, pp. 295–298.

[Sch98] M. Schechter, *Integration over a polyhedron: an application of the Fourier-Motzkin elimination method*, Amer. Math. Monthly **105** (1998), no. 3, 246–251.

[VB08] S. Verdoolaege and M. Bruynooghe, *Algorithms for weighted counting over parametric polytopes: A survey and a practical comparison*, 2008, Eighth ACES Symposium, Edegem, Belgium.

[WP07] M.C. Wilson and G. Pritchard, *Probability calculations under the IAC hypothesis*, Math. Social Sci. **54** (2007), no. 3, 244–256.

Software

[barvinok] `barvinok` *by S. Verdoolaege*, ver. 0.34 (2011),
`http://freshmeat.net/projects/barvinok`

[convex] `Convex` *by M. Franz*, ver. 1.1.3 (2009),
`http://www.math.uwo.ca/~mfranz/convex/`

[latte] `LattE integrale` *by J.A. DeLoera, M. Köppe et al.*, ver. 1.5 (2011),
`http://www.math.ucdavis.edu/~latte/`

[normaliz] `Normaliz` *by W. Bruns, B. Ichim, and C. Söger*, ver. 2.8 (July 2012),
`http://www.mathematik.uni-osnabrueck.de/normaliz/`

[sympol] `SymPol` *by T. Rehn and A. Schürmann*, ver. 0.1.7 (April 2012),
`http://www.geometrie.uni-rostock.de/software/`

Achill Schürmann, Institute of Mathematics, University of Rostock, 18051 Rostock, Germany
Email: `achill.schuermann@uni-rostock.de`