Parametric Analysis of Polyhedral Iteration Spaces

Philippe Clauss and Vincent Loechner

ICPS, Université Louis Pasteur, Strasbourg,
Pôle API, Bd. Sébastien Brant,
67400 Illkirch, France
{clauss,loechner}@icps.u-strasbg.fr
http://icps.u-strasbg.fr

Abstract

In the area of automatic parallelization of programs, analyzing and transforming loop nests with parametric affine loop bounds requires fundamental mathematical results. The most common geometrical model of iteration spaces, called the polytope model, is based on mathematics dealing with convex and discrete geometry, linear programming, combinatorics and geometry of numbers.

In this paper, we present an automatic method for computing the number of integer points contained in a convex polytope or in a union of convex polytopes. The procedure consists of first, computing the parametric vertices of a polytope defined by a set of parametric linear constraints, and then computing the Ehrhart polynomial, i.e. a parametric expression of the number of integer points.

The paper is illustrated with the computation of the maximum available parallelism of a given loop nest.

1: Introduction

From systolic algorithm synthesis to nested loops parallelization, a common geometric model is used, called the *polytope model*, leading to common fundamental mathematical problems, dealing with discrete and convex geometry, combinatorics, geometry of numbers and linear programming.

One of these fundamental problems is the counting of integer points, i.e. points having integer coordinates, contained in a polytope. Moreover, this counting has to be parameterized, since scientific algorithms often depend on size parameters. Such a counting is necessary when analyzing a parallel algorithm modeled by a polyhedral iteration space. Many kinds of analysis are reduced to this problem such as finding the number of flops executed by a loop, the number of memory locations touched by a loop, the communication volume, the maximum parallelism, and so forth. Such information is essential to optimizing compilers when transforming a parallel program.

In this paper, we present an automatic method for computing the number of integer points contained in a convex polytope or in a union of convex polytopes. This method consists of first computing the parametric vertices of the considered polytope defined by a set of parametric linear constraints, and then computing the *Ehrhart polynomial*, i.e. a parametric expression of the number of integer points.

This second part is based on results due to the mathematician E. Ehrhart, who solved this problem for parametric polytopes that depend on one positive integral parameter. We propose an extension of his results to any number of parameters, showing that this extension is made possible by a new algorithm to find parametric vertices finding algorithm [12].

This algorithm is based on the classical representation of a parametric polytope as a non-parametric polyhedron in the combined parameter and data space. The computation of the faces whose dimension equals the number of parameters, allows the expression of the parametric vertices as affine functions of the parameters. A validity domain partition for the vertices is also computed: each subdomain corresponds to a certain set of vertices of the parametric polytope.

The paper is organized as follows. Section 2 is dedicated to the mathematical background. In section 3, the algorithm to find the parametric vertices is described. It makes possible, in section 4, an extension of Ehrhart's results on the parametric counting of integer points. Finally, we illustrate the method in section 5, by computing the parallelism rate for a given iteration space and a given linear parallelization. This rate is parameterized by the size parameter and the execution instant. In section 6, conclusions and comparisons with related works are given.

2: Definitions and assumptions

We first recall some basic notions dealing with geometry of numbers [9, 17, 1] and enumerative combinatorics [18, 19]. Then some more specific concepts, closely related to the content of the paper, are introduced.

2.1: Polyhedra and polytopes

Let $\mathcal{Q}$ denote the set of rational numbers and $\mathcal{Z}$ the set of integers. A convex polyhedron is defined by a finite set of linear inequalities $P = \{x \in \mathcal{Q}^d \mid A \cdot x \leq b\}$ where A is a rational matrix and b a rational vector. This is called the *implicit representation* of the polyhedron. A bounded convex polyhedron is called a convex polytope, and represents the convex hull of a finite number of points $x_1, \ldots, x_{m+1}$ in $\mathcal{Q}^d$. If $x_1, \ldots, x_{m+1}$ are affinely independent points of $\mathcal{Q}^d$ (which also means that $x_2 - x_1, x_3 - x_1, \ldots, x_{m+1} - x_1$ are linearly independent) then the convex hull of x_1, $\ldots$, x_{m+1} is called an m-dimensional simplex of $\mathcal{Q}^d$ with vertices x_1, $\ldots$, x_{m+1}.

Since our goal is to count the number of integer points contained in a polyhedron, only polytopes are considered in the rest of the paper.

The *affine span* of P is the smallest affine subspace of $\mathcal{Q}^d$ entirely containing P. The *dimension* of P is the dimension of its affine span, and a k-dimensional polytope is called a k-polytope for brevity.

Definition 1 (geometric representation) *A convex polytope P has a dual geometric representation as a combination of vertices:*

$$P = \{x \in \mathcal{Q}^d \mid x = V\nu; \nu_i \geq 0; \sum_{i=1}^{d} \nu_i = 1\}$$

where the columns of matrix V are the coordinates of its vertices and ν is a vector with elements $\nu_i, 1 \leq i \leq d$.

Any polytope P has a unique minimal geometric representation. Procedures exist to compute the dual representations of P, that is, given A and b in the implicit form, compute V and vice versa (see [15] for references).

If $\vec{c}$ is a nonzero vector and γ is defined by $\gamma = MAX\{\vec{c} \cdot x \mid x \in P\}$, the affine hyperplane $\{x \mid \vec{c} \cdot x = \gamma\}$ is called a supporting hyperplane of P. The *faces* of P are $\emptyset$, P, and the intersections of P with its supporting hyperplanes. Each face of P is itself a polyhedron and is called a k-face if it is k-dimensional. If P is a k-polyhedron, faces of P of dimension 0, 1 and $k-1$ are called *vertices, edges* and *facets*, respectively.

In this work, we consider polytopes which may be non-convex, called *normal polyhedra*, according to the following definition. This definition corresponds to the one given by Ehrhart in [5] and [7, p. 47] and includes unions of convex polytopes.

Definition 2 (normal polyhedron) *A normal polyhedron P is a closed polyhedron which admits a finite rectilinear simplicial subdivision. It is then the union of a finite number of simplices $S_1, S_2, ..., S_n$ of same dimension such that,*

- *The intersection between two simplices is either empty or is a common face,*
- *Each facet of one of these simplices belongs to at most one other simplex, faces belonging to only one simplex being the border of the polyhedron,*
- *These simplices can be ordered such that for any $h < n$, if $P_h = S_1 \cup S_2 \cup ... \cup S_h$, the intersection $P_h \cap S_{h+1}$ is either empty, or is the union of common facets of S_{h+1} and the border of P_h.*

2.2: Parameterized lattice polytopes

We consider the parameterized polytopes of the following form:

$$P_N = \{x \in \mathcal{Q}^d \mid A \cdot x \leq B \cdot N + C\}$$

where N is the symbolic p-vector of the parameters, A and B are constant rational matrices of size $(m \times d)$ and $(m \times p)$ respectively, and C is a constant rational m-vector.

Parameterized polytopes can be represented in the *combined* data and parameter space as follows:

$$P' = \left\{ \begin{pmatrix} x \\ N \end{pmatrix} \in \mathcal{Q}^{d+p} \;\middle|\; [A \mid -B] \begin{pmatrix} x \\ N \end{pmatrix} \leq C \right\}$$

Let S_{N_0} be the subspace of $\mathcal{Q}^{d+p}$ associated with a given value of $N = N_0$:

$$S_{N_0} = \left\{ \begin{pmatrix} x \\ N_0 \end{pmatrix} \in \mathcal{Q}^{d+p} \right\}$$

The polytope P_{N_0} can be obtained by intersecting P' with S_{N_0}: $P_{N_0} = P' \cap S_{N_0}$.

A lattice is the set of points $\mathcal{L}^d = \{\alpha_1 \vec{g}_1 + ... + \alpha_d \vec{g}_d \mid \alpha_1, ..., \alpha_d \in \mathcal{Z}\}$ where $\vec{g}_1, ..., \vec{g}_d$ are linearly independent vectors of $\mathcal{Q}^d$, called the basis of the lattice. The set of all the integral points is called *standard lattice* denoted by $\mathcal{Z}^d$. Let G be the matrix whose columns are the vectors $\vec{g}_1, ..., \vec{g}_d$, called the *dilatation matrix*. Any lattice $\mathcal{L}^d$ consists of the points $G \cdot z$, $z \in \mathcal{Z}^d$, i.e. $\mathcal{L}^d = G \cdot \mathcal{Z}^d$. So we get the

most general lattice $\mathcal{L}^d$ by subjecting the standard lattice $\mathcal{Z}^d$ to an arbitrary non-singular transformation G [9]. Hence, without loss of generality, only the standard lattice $\mathcal{Z}^d$ is considered in the rest of the paper. The *enumerator* of any set E is the number of points in $E \cap \mathcal{L}^d$.

The intersection between a lattice $\mathcal{L}^d$ and a polytope is called a lattice polytope P of $\mathcal{L}^d$. A lattice polytope P of $\mathcal{L}^d$ is a $\mathcal{L}^d$-polytope if all of its vertices are in $\mathcal{L}^d$. If $\mathcal{L}^d = \mathcal{Z}^d$, P is an *integral polytope*. If some vertices of P do not belong to $\mathcal{L}^d$ and have rational coordinates, P is a *rational* polytope.

Since by definition, any lattice $\mathcal{L}^d$ can be deduced from $\mathcal{Z}^d$ by a non-singular transformation G, any $\mathcal{L}^d$-polytope is affinely equivalent to an integral polytope. Hence, results given in the rest of the paper and concerning integral polytopes can be directly extended to $\mathcal{L}^d$-polytopes. From now on and without loss of generality, the generic term of integral polytopes will be used.

Our method of counting the number of integer points is based on the decomposition of a parametric polytope into several *homothetic and bordered* systems, associated with *validity domains*, according to the following definition which is an extension of Ehrhart's definition [5] to any number of parameters.

Definition 3 (homothetic system) *A system is homothetic if its constraints are of the form* $\sum a_i x_i = \sum b_j n_j$, $\sum a_i x_i < \sum b_j n_j$, $\sum a_i x_i \leq \sum b_j n_j$, *where the* a_i*'s and the* b_j*'s are given integers, the* x_i*'s are free variables and the* n_j*'s are positive integral parameters.*

Such a system H_N, $N = (n_1, n_2, \ldots, n_q)$, *is significant if and only if* $H_{(1,1,\ldots,1)}$ *defines a bounded polyhedral domain* $P_{(1,1,\ldots,1)}$ *(obviously convex).*

H_N, *which defines the domain* P_N, *is homothetic if all the* p *one-parameter polytopes* $P_{0,0,\ldots,0,n_j,0,\ldots,0}$ *correspond to* $P_{0,0,\ldots,0,1,0,\ldots,0}$ *by the* n_j*-homothetic transformations of center* O. *All the* $P_{0,0,\ldots,0,1,0,\ldots,0}$*'s,* $j = 1..p$, *are called the primitive domains of* H_N.

The properties of such homothetic polyhedra yield the main mathematical result given later in the paper. However this definition has to be extended to *bordered polytopes* in order to handle polytopes modeling loops with *affine* loop bounds. The following definition is also an extension of Ehrhart's [5].

Definition 4 (bordered system) *A system* B_N *is bordered if its constraints are of the form* $\sum a_i x_i = \sum b_j n_j + c$, $\sum a_i x_i < \sum b_j n_j + c$, $\sum a_i x_i \leq \sum b_j n_j + c$, *where the* a_i*'s and the* b_j*'s are given integers, the* c*'s are rationals or not, the* x_i*'s are free variables and the* n_j*'s are positive integral parameters.*

By deleting all the c*'s, an homothetic system* H_N *defining the domain* P_N *is obtained. By adding a parallel hyperplane of the same dimension* $(k-1)$ *to each facet of* P_N, *at a rational distance or not, independently of the parameters* N, *but which may depend on the facet, we obtain the bordered polytope* P'_N *whose kernel is* P_N.

Observe that the borders can cut the kernel while the parameters are less than certain minimal values. Hence, P'_N is said to be *complete* if it exists, i.e. when all the borders are involved as bounds. Observe also that a translation, independent of the parameters, transforms an homothetic polytope P_N into a bordered polytope, complete for all values of the parameters.

In general, multi-parametric polytopes (with several parameters) are described by a collection of homothetic or bordered polytopes, each defined on a disjoint subdomain of parameter values called the *validity domain*. The determination of these subdomains is done by the parametric vertices finding algorithm described in a later section.

2.3: Pseudo-polynomials

In the following, we denote the enumerator of P_N as j_N, i.e. the number of integer points of P_N. It will be formulated using *polynomials* and *pseudo-polynomials* according to the following definitions. We extend Ehrhart's definition of periodic numbers [7] to any dimension.

Definition 5 (periodic number) *A one-dimensional periodic number* $u(n) = [u_1, u_2, \ldots, u_p]_n$ *is equal to the item whose rank is equal to* n *mod* p, p *is called the period of* $u(n)$.

$$u(n) = \begin{cases} u_1 & if\ n = 1\ mod\ p \\ u_2 & if\ n = 2\ mod\ p \\ \ldots & \\ u_p & if\ n = 0\ mod\ p \end{cases}$$

Let $N = (n_j)$ *be a* q*-dimensional vector,* $j \in [1 \ldots q]$. *A* q*-periodic number* $u(N)$ *is defined by a* q*-dimensional array* $(u_I)_N$, $I = (i_1, i_2, \ldots, i_q)$, *of size* $p_1 \times p_2 \times \ldots \times p_q$ *such that*

$$u(N) = u_I \ if\ \forall j \in [1 \ldots q],\ n_j = i_j\ mod\ p_j$$

The vector $p = (p_j)$ *is called the multi-period of* $u(N)$ *and the least common multiple of the* p_i*'s is called the period of* $u(N)$.

For example,

- $(-1)^n = [-1, 1]_n$
- $(-1)^{n-m} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}_{(n,m)}$

Ehrhart in [7] gives the basic properties of one-dimensional periodic numbers which can be easily extended to the multi-dimensional case:

Let $u(N)$ and $u'(N)$ be q-periodic numbers of multi-period p and p' respectively. They can be reduced to the same multi-period $m = (m_i)$, where $m_i = p_i\alpha = p'_i\beta$ is the least common multiple of p_i and p'_i, by repeating α times (respectively β times) the sequence of items in the i^{th} dimension of $u(N)$ (respectively $u'(N)$). Since $u(N)$ and $u'(N)$ are reduced to the same period, $u(N)+u'(N) = (u_I+u'_I)_N$ and $u(N)u'(N) = (u_I u'_I)_N$.

Likewise, we also extend Ehrhart's definition on pseudo-polynomials.

Definition 6 (pseudo-polynomial) *A polynomial $f(N)$ is a pseudo-polynomial if some of its coefficients are periodic numbers instead of constants. The least common multiple of the periods of its periodic coefficients is the pseudo-period of $f(N)$. Thus, a pseudo-polynomial has two characteristics, its degree and its pseudo-period.*

3: A parametric vertices finding algorithm

3.1 Parametric vertices

The method presented below is fully explained in [12]. It computes the vertices $v_i(N)$ of a parametric polyhedron $P_N = \{x \in \mathcal{Q}^d \mid A \cdot x \leq B \cdot N + C\}$, where N is a vector of the form $\begin{pmatrix} n_1 \\ \vdots \\ n_p \end{pmatrix}$. We call P' its associated polyhedron in the combined data and parameter space. We assume in the following that the different n_j are linearly independent, i.e. there's no equality of the form: $DN + d = 0$ where D is a non-null matrix and d a vector. If such an equality exists, it can be easily detected and removed. This has to be verified in order to prove the following theorem:

Theorem 1 *Let $v_i(N)$ be a vertex of the polyhedron P_N. There exists a p-face of P' denoted F_i^p such that $v_i(N)$ corresponds to $F_i^p \cap S_N$.*

Proof The detailed proof can be found in [12]. It is based on the definition of $v_i(N)$ which is a 0-face of P_N: $v_i(N)$ is the intersection of a supporting hyperplane $\mathcal{H}_i$ and P_N. In the combined space, $\mathcal{H}_i$ is the intersection of S_N and a supporting hyperplane of P' denoted $\mathcal{H}'_i$. $\mathcal{H}'_i$ can be chosen such that $\mathcal{H}'_i \cap P'$ is a p-face of P' and the result follows.

Let $Proj$ be the projection of all points in $\mathcal{Q}^{d+p}$ on the parameter space $\mathcal{Q}^p$.

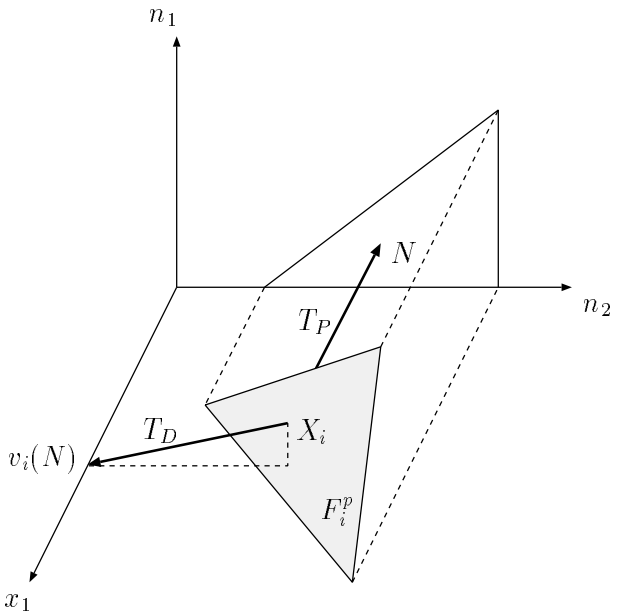

Figure 1. A p-face F_i^p (highlighted) and the computation of $v_i(N)$ as a function of N

Property 1 *The point $v_i(N)$ is defined only for the values of N contained in the polyhedron $Proj(F_i^p)$.*

Proof The range of N over which $F_i^p \cap S_N \neq 0$ is the projection of F_i^p on the parameter space $\mathcal{Q}^p$. Since $\begin{pmatrix} v_i(N) \\ N \end{pmatrix} = F_i^p \cap S_N$, the validity domain for $v_i(N)$ is $N \in Proj(F_i^p)$. $Proj(F_i^p)$ is a convex polyhedron since F_i^p is a convex polyhedron.

Let F_i^p be a p-face of P'. If F_i^p corresponds to a vertex $v_i(N)$ then the intersection between F_i^p and S_N is the single point $\begin{pmatrix} v_i(N) \\ N \end{pmatrix}$. Let $\mathcal{B}$ be a basis of F_i^p and X the coordinates in basis $\mathcal{B}$ of a point in F_i^p.

We can compute the affine transformations T_D and T_P such that:

$$N = T_P(X_i) \tag{1}$$

$$v_i(N) = T_D(X_i) \tag{2}$$

If F_i^p corresponds to a vertex of P_N then the intersection between F_i^p and S_N is a single point (cf. figure 1), and then the transformation T_P is bijective. Equation (1) can be rewritten as:

$$T_P^{-1}(N) = X_i$$

And by replacing X_i by this value in the equation (2) we find:

$$v_i(N) = T_D(T_P^{-1}(N))$$

If T_P is not bijective, then there does not exist one single point in F_i^p corresponding to one value of N, but a whole k-polyhedron, with $k > 0$. The intersection of this polyhedron and S_N does not correspond to a vertex of P_N, but to a k-face of P_N, so this face F_i^p can be ignored by the algorithm.

3.2 Partitioning the parameter space

The output of the above algorithm is a set of vertices and for each vertex, a set of constraints that must be verified by the parameters N for this vertex to be valid.

In order to use these results, we want to find the set of distinct parameter subdomains which separate complete sets of parameterized vertices, i.e. to transform this representation into a collection of sets of conditions on N, such that for each of these sets, a complete list of the vertices of P_N is given.

This is obtained by partitioning the parameter space $\mathbb{Q}^p$ with the intersections of the projections of the p-faces of P'. If for each N inside one of these sets the set of p-faces intersected by S_N is the same, then the set of vertices of P_N is also the same.

The whole parameter validity domain of P_N is $Proj(P')$, the projection of P' on the parameter space $\mathbb{Q}^p$. The projection of the convex p-faces of P' forms a set of polyhedra in $\mathbb{Q}^p$ that covers $Proj(P')$, since P' is entirely defined by its set of p-faces. A set of polyhedra $\{D_j \mid 1 \leq j \leq f\}$ where f is the number of validity domains can be found, such that:

- D_j is a polyhedron of dimension p.
- $D_j \cap D_k$ is a polyhedron of dimension less than p, $\forall j \neq k$.
- $D_1 \cup D_2 \cup \ldots \cup D_f = Proj(P')$
- $\forall N \in D_j, \{v_i(N) = S_N \cap D'\}$ is unique.

These validity domains are essential for the extension of Ehrhart's results presented in the following section. It is shown that their application can be done on each predetermined validity domain.

4: Parametric counting through Ehrhart polynomials extension

In this section, we recall and extend some results due to E. Ehrhart who made a study of the enumeration of lattice points in polyhedra, by means of polynomials along the same lines as Euler. All the initial proofs can be found in [4, 5, 6, 7]. Some aspects of Ehrhart's work were corrected, streamlined, and expanded by I. G. Macdonald [13, 14] and R. P. Stanley [18, 19].

Definition 7 (denominator) *The denominator of a rational point is the lowest common multiple of the denominators of its coordinates. The denominator of a rational polyhedron is the least common multiple of the denominators of its vertices.*

Ehrhart's main result consists in the following theorem dealing with parametric polytopes which depend on only one parameter n.

Theorem 2 (Ehrhart's fundamental theorem) *The enumerator j_n of any homothetic or bordered k-polyhedron P_n is a polynomial in n of degree k if P_n is integral, and it is a pseudo-polynomial in n of degree k whose pseudo-period is the denominator of P_n if P_n is rational.*

We extend this theorem to any number of parameters $N = (n_1, \ldots, n_p)$.

Theorem 3 (fundamental theorem extended) *The enumerator j_N of any k-polytope P_N is expressed at different domains of the parameters values by different multi-variable polynomials in N of degree k if P_N is integral, and by different multi-variable pseudo-polynomials in N of degree k whose pseudo-period is the denominator of P_N if P_N is rational. The parameter validity domains give the parameter ranges where the parametric vertices of P_N are defined. These polynomials or pseudo-polynomials are of the form:*

$$j_N = \sum_{i_1=0}^{k} \sum_{i_2=0}^{k} \ldots \sum_{i_p=0}^{k} c_{i_1,i_2,\ldots,i_p} n_1^{i_1} n_2^{i_2} \ldots n_p^{i_p}$$

where the $c_{i_1,i_2,\ldots,i_p}$'s are periodic numbers whose dimensions are at most p.

The proof uses the following lemma, stating that if a parametric polytope is defined by constraints without any constant c, then the coordinates of its vertices are linear functions of the parameters.

Lemma 1 (parametric vertices) *If P_N is a parametric polytope defined by constraints of the form $a \cdot x \leq b \cdot N$, then the parametric coordinates of its vertices are linear functions of the parameters of the form $l \cdot N$, where l is a constant p-vector.*

Proof (Lemma 1) Without loss of generality, let P_N be a k-polytope defined by a non-redundant system of constraints. Each vertex can be defined as the intersection of at least k constraints, with inequalities getting equalities. The resolution of such a system of equations, in the form of the constraints, obviously leads to solutions expressed as linear functions of the parameters. □

Proof (Theorem 3) Consider a parametric polytope P_n depending on one parameter n, whose vertex coordinates are defined by linear functions of the parameters. P_n has a dual geometric representation $\{x \in \mathcal{Q}^d \mid x = V_n\nu; \nu \geq 0; \sum \nu_i = 1\}$, showing in this case that any point is also a linear function of the parameters. Hence, any instance of P_n can be obtained by the n-homothetic transformation of P_1.

Now, consider a parametric polytope $P_{n,m}$ depending on two parameters n and m, defined by linear constraints of the form $a \cdot x \leq b_1 n + b_2 m$, whose vertex coordinates are defined by linear functions of the parameters. Consider a validity domain where these parametric vertices are defined. For $n = 0$, $P_{0,m}$ is an homothetic polytope, i.e. any instance of $P_{0,m}$ can be obtained by the m-homothetic transformation of $P_{0,1}$.

Moreover, $P_{0,m}$ is the kernel of the one-parameter bordered polytopes $P_{1,m}, P_{2,m}, \ldots$ In the same way, $P_{n,0}$ is homothetic and is the kernel of the one-parameter bordered polytopes $P_{n,1}, P_{n,2}, \ldots$ Hence, $P_{n,m}$ is homothetic and is the kernel of any three-parameters polytope $P_{n,m,p}$ and so on ...

Ehrhart's fundamental theorem can be applied to all these one-parameter bordered polytopes $P_{1,m}, P_{2,m}, \ldots$ or $P_{n,1}, P_{n,2}, \ldots$ and hence, the theorem is proved for any parametric polytope depending on two parameters. In a straightforward manner, this proof can recursively be extended to any number of parameters.

So, the enumerators of $P_{n,1}, P_{n,2}, \ldots$ are pseudo-polynomials in n, whose coefficients depend on the values of m, of the form:

$$j_{n,m} = c_k(n)n^k + c_{k-1}(n)n^{k-1} + \ldots + c_0(n)$$

where $c_0, \ldots, c_k$ are periodic numbers of n, whose values depend on m.

For any value of $m = 1, 2, \ldots$, these coefficients can be computed by substituting n with some initial values 1, 2, ...and by resolving a system of linear equations of the form:

$$\begin{cases} c_0(n) & = & j_{0,m} \\ c_k(n) + c_{k-1}(n) + \ldots + c_0(n) & = & j_{1,m} \\ c_k(n)2^k + c_{k-1}(n)2^{k-1} + \ldots + c_0(n) & = & j_{2,m} \\ \ldots & & \ldots \end{cases}$$

The solutions $c_i(n)$ of such a system are linear functions of some initial pseudo-polynomials in m of degree k, $j_{0,m}, j_{1,m}, j_{2,m}, \ldots$ Hence, $j_{n,m}$ is a two-variables pseudo-polynomial of degree k. It is also straightforward to recursively extend this proof to any number of parameters. □

Due to these theorems, the form of the parametric counting formulae is known: the period is equal to the denominator of the k-polyhedron P_N and the degree is equal to k. Hence, the determination of the Ehrhart pseudo-polynomial can be done by counting some initial values of the enumerator and solving systems of symbolic rational linear equations. Since the number of equations of the system must be equal to the number of unknown values, the complexity depends on the number of parameters, the number of validity domains, the dimension, and the denominator.

The number of unknowns can be reduced using the following theorem, conjectured by Ehrhart and proved by Stanley.

Theorem 4 *Let P_n be a rational k-polytope and let the Ehrhart pseudo-polynomial be*

$$j_n = c_k(n)n^k + c_{k-1}(n)n^{k-1} + \ldots + c_0(n)$$

where $c_0, \ldots, c_k$ are periodic numbers of n. If for some $q \in [0, k]$, the affine span of every q-dimensional face of P_1 contains a lattice point, then for any $p \geq q$, $c_p(n)$ is constant, i.e. of period one.

Example 1 Consider the set of linear constraints $\{0 \leq i \leq \frac{1}{2}n, 0 \leq j \leq \frac{1}{2}m\}$ defining the rational 2-polytope $P_{n,m}$,

$$P_{n,m} = \{\begin{pmatrix} i \\ j \end{pmatrix} \in \mathcal{Z}^2 \mid \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} \leq \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} n \\ m \end{pmatrix}\}$$

$P_{n,m}$ is the closed rectangle whose parametric vertices are $(0,0)$, $(\frac{n}{2}, 0)$, $(0, \frac{m}{2})$, $(\frac{n}{2}, \frac{m}{2})$.

By looking at $P_{1,0}$, we observe that one vertex, $(\frac{1}{2}, 0)$, contains no integer point. Hence, the coefficient of n^0 is periodic. The other coefficients are constant, since all the edges, and obviously the polytope itself, contain at least one integer point.

By looking at $P_{0,1}$, we observe that one vertex, $(0, \frac{1}{2})$, contains no integer point. Hence, the coefficient of m^0 is periodic. The other coefficients are constant, since all the edges, and obviously the polytope itself, contain at least one integer point.

Moreover, the period of the periodic coefficients is equal to the denominator of $P_{n,m}$, i.e. 2. The pseudo-polynomial $j_{n,m} = c_1 n^2 + c_2 n + [c_3, c_4]$ can thus be determined by solving the following system of rational linear equations:

$$\begin{cases} c_4 & = j_{0,m} \\ c_1 + c_2 + c_3 & = j_{1,m} \\ 4c_1 + 2c_3 + c_4 & = j_{2,m} \\ 9c_1 + 3c_2 + c_3 & = j_{3,m} \end{cases}$$

$j_{0,m}, j_{1,m}, j_{2,m}$ and $j_{3,m}$ have to be computed. For example, we compute $j_{0,m} = c_5 m^2 + c_6 m + [c_7, c_8]$:

$$\begin{cases} c_8 & = j_{0,0} \\ c_5 + c_6 + c_7 & = j_{0,1} \\ 4c_5 + 2c_6 + c_8 & = j_{0,2} \\ 9c_5 + 3c_6 + c_7 & = j_{0,3} \end{cases}$$

From $P_{0,0}$ to $P_{0,3}$, we count $j_{0,m} = 1, 1, 2, 2$. It follows that,

$$j_{0,m} = \frac{1}{2}m + [\frac{1}{2}, 1]$$

Finally, $j_{n,m}$ is determined:

$$\begin{aligned} j_{n,m} &= \left(\tfrac{1}{4}m + \left[\tfrac{1}{4}, \tfrac{1}{2}\right]_m\right) n + \left[\tfrac{1}{4}m + \left[\tfrac{1}{4}, \tfrac{1}{2}\right]_m, \tfrac{1}{2}m + \left[\tfrac{1}{2}, 1\right]_m\right]_n \\ &= \tfrac{1}{4}nm + \left[\tfrac{1}{4}, \tfrac{1}{2}\right]_m n + \left[\tfrac{1}{4}, \tfrac{1}{2}\right]_n m + \begin{bmatrix} 1/4 & 1/2 \\ 1/2 & 1 \end{bmatrix}_{n,m} \end{aligned}$$

This method is now applied to the determination of the maximum parallelism of a loop nest.

5: An application of parametric analysis: the maximum parallelism

Consider the following loop nest:

```
for i:= 0 to n do
  for j:= 0 to i do
    for k:= 0 to i-j do
      Statement(s)
    endfor
  endfor
endfor
```

This loop nest is modeled by the one-parameter polyhedral iteration space defined by:

$$P_n = \{\begin{pmatrix} i \\ j \\ k \end{pmatrix} \in \mathcal{Z}^3 | \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} i \\ j \\ k \end{pmatrix} \leq \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} n\}$$

Assume that the dependence analysis of this loop nest yields an optimal linear schedule defined by $t(i,j,k) = i + j + k$: any iteration (i,j,k) is computed at instant $i + j + k$. We use our method to determine the parametric number of simultaneous calculations $par(n, t_0)$ occuring at any instant t_0, in order to determine the parametric maximum parallelism $maxpar(n) = MAX_{t_0}\{par(n, t_0)\}$, i.e. the maximum number of simultaneous calculations.

From a geometrical point of view, all the points (i,j,k) of P_n resulting in the same value of $t_0 = t(i,j,k)$, belong to the two-parameters rational 2-polytope T_{n,t_0} defined by

$$T_{n,t_0} = P_n \cap \{(i,j,k) \in \mathcal{Z}^3 | i + j + k = t_0\}$$

The number of points in T_{n,t_0} is $par(n, t_0)$.

Giving this last parametric definition of T_{n,t_0} as an input to the parametric vertices finding algorithm, it outputs the following results as shown on figures 2 and 3:

- if $0 \leq t_0 \leq n$, the parametric vertices are $v_1 = (t_0, 0, 0)$, $v_2 = (\frac{t_0}{2}, \frac{t_0}{2}, 0)$ and $v_3 = (\frac{t_0}{2}, 0, \frac{t_0}{2})$.
- if $n \leq t_0 \leq 2n$, the parametric vertices are $v_2 = (\frac{t_0}{2}, \frac{t_0}{2}, 0)$, $v_3 = (\frac{t_0}{2}, 0, \frac{t_0}{2})$, $v_4 = (n, t_0 - n, 0)$ and $v_5 = (n, 0, t - n)$.
- otherwise, T_{n,t_0} does not exist and $par(n, t_0) = 0$.

Hence, $par(n, t_0)$ is defined on two validity domains as pseudo-polynomials in n and t_0, of degree 2, and of pseudo-period 2. According to theorem 4 and by looking at $T_{1,0}$, it is deduced that all the coefficients of n^q, $q \leq 2$, are constant:

$$par(n, t_0) = c_1 n^2 + c_2 n + c_3$$

(i) We determine c_1, c_2, c_3 by solving the following system of equations on the first validity domain:

$$\begin{cases} 36c_1 + 6c_2 + c_3 & = par(6, t_0) \\ 49c_1 + 7c_2 + c_3 & = par(7, t_0) \\ 64c_1 + 8c_2 + c_3 & = par(8, t_0) \end{cases}$$

Solving this system yields the equations for $par(6, t_0)$, $par(7, t_0)$ and $par(8, t_0)$. According to theorem 4 and by looking at $T_{1,1}$, it is deduced that only the coefficient of t_0^2 is constant. For example,

$$par(6, t_0) = c_1 t_0^2 + [c_2, c_3] t_0 + [c_4, c_5]$$

such that,

$$\begin{cases} c_5 & = par(6, 0) = 1 \\ c_1 + c_2 + c_4 & = par(6, 1) = 1 \\ 4c_1 + 2c_3 + c_5 & = par(6, 2) = 3 \\ 9c_1 + 3c_2 + c_4 & = par(6, 3) = 3 \\ 16c_1 + 4c_3 + c_5 & = par(6, 4) = 6 \end{cases}$$

where the values of $par(6, 0), \ldots, par(6, 4)$ are given by some initial countings.
$\Rightarrow par(6, t_0) = \frac{1}{8}t_0^2 + [\frac{1}{2}, \frac{3}{4}]t_0 + [\frac{3}{8}, 1]$. In the same way, we determine that the other pseudo-polynomials are the same, i.e $\forall\ n \in [6..8]$, $par(n, t_0) = \frac{1}{8}t_0^2 + [\frac{1}{2}, \frac{3}{4}]t_0 + [\frac{3}{8}, 1]$. Solving the first system gives the parallelism rate of any instant t_0 on the first validity domain:

$$\forall\ 0 \leq t_0 \leq n,\ par(n, t_0) = \frac{1}{8}t_0^2 + \left[\frac{1}{2}, \frac{3}{4}\right] t_0 + \left[\frac{3}{8}, 1\right]$$

Observe that the parallelism rate does not depend on n in this validity domain.

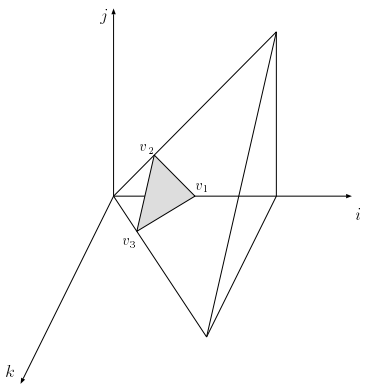

Figure 2. T_{n,t_0} in the first validity domain

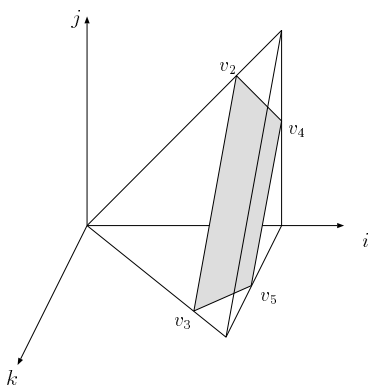

Figure 3. T_{n,t_0} in the second validity domain

(ii) We now consider the second validity domain.

$$\begin{cases} 36c_1 + 6c_2 + c_3 &= par(6,t_0) \\ 49c_1 + 7c_2 + c_3 &= par(7,t_0) \\ 64c_1 + 8c_2 + c_3 &= par(8,t_0) \end{cases}$$

Solving this system yields the equations for $par(6,t_0)$, $par(7,t_0)$ and $par(8,t_0)$. For example,

$$par(6,t_0) = c_1 t_0^2 + [c_2, c_3] t_0 + [c_4, c_5]$$

such that,

$$\begin{cases} 36c_1 + 6c_3 + c_5 &= par(6,6) &= 10 \\ 49c_1 + 7c_2 + c_4 &= par(6,7) &= 9 \\ 64c_1 + 8c_3 + c_5 &= par(6,8) &= 12 \\ 81c_1 + 9c_2 + c_4 &= par(6,9) &= 9 \\ 100c_1 + 10c_3 + c_5 &= par(6,10) &= 11 \end{cases}$$

$\Rightarrow par(6,t_0) = -\frac{3}{8}t_0^2 + [6, \frac{25}{4}]t_0 + [-\frac{117}{8}, -14]$. In the same way, we determine the other pseudo-polynomials:

$$par(7,t_0) = -\tfrac{3}{8}t_0^2 + [7, \tfrac{29}{4}]t_0 + [-\tfrac{165}{8}, -20]$$

$$par(8,t_0) = -\tfrac{3}{8}t_0^2 + [8, \tfrac{33}{4}]t_0 + [-\tfrac{221}{8}, -27]$$

Solving the first system gives the parallelism rate of any instant t_0 on the second validity domain:

$\forall\ n \le t_0 \le 2n,$
$par(n,t_0) = nt_0 - \frac{1}{2}n^2 + \frac{1}{2}n - \frac{3}{8}t_0^2 + \left[0, \frac{1}{4}\right]_{t_0} t_0 - \left[\frac{3}{8}, 1\right]_{t_0}$

(iii) From these results, the maximum parallelism is deduced:

- On the first validity domain ($0 \le t_0 \le n$), $par(n,t_0)$ is maximum when t_0 is maximum, i.e. when $t_0 = n$. Hence, $maxpar(n) = \frac{1}{8}n^2 + \left[\frac{1}{2}, \frac{3}{4}\right] n + \left[\frac{3}{8}, 1\right]$.
- On the second validity domain ($n \le t_0 \le 2n$), we take the derivative of $par(n,t_0)$ with respect to t_0: $par'(n,t_0) = n - \frac{3}{4}t_0 + \left[0, \frac{1}{4}\right]_{t_0}$.

$$par'(n,t_0) \le 0$$

$$\Leftrightarrow \quad n - \tfrac{3}{4}t_0 + \left[0, \tfrac{1}{4}\right]_{t_0} \le 0$$

$$\Leftrightarrow \quad t_0 \ge \tfrac{4n - [0,1]}{3}$$

  Hence, $par(n,t_0)$ increases from $t_0 = n$ to $t_0 = \frac{4n-[0,1]}{3}$, and then decreases until $t_0 = 2n$: it is maximum for $t_0 = t_{max} = \frac{4n-[0,1]}{3}$. Since t_{max} is not always integral, a case study is necessary to get the exact value of $maxpar(n) = par(n, t_{max})$, checking the possible values of $\lfloor t_{max} \rfloor$ and $\lceil t_{max} \rceil$. It yields the following answer: $maxpar(n) = \frac{1}{6}n^2 + \frac{5}{6}n - [1, \frac{23}{24}, 1]$.
- In conclusion, the maximum parallelism of the loop nest is:

$$\forall\ n \ge 0,\ maxpar(n) = \frac{1}{6}n^2 + \frac{5}{6}n - \left[1, \frac{23}{24}, 1\right]$$

6: Conclusion

We have presented a method for computing the parametric number of integer points, i.e. the Ehrhart polynomial, of a convex polytope or a union of convex polytopes. This method includes an original algorithm to find parametric vertices and an extension of Ehrhart's results to any number of parameters.

The procedure is entirely automatic, from a parametric polytope definition as a system of linear constraints to the Ehrhart polynomials. It is partially implemented at this time, and will be soon completed. It uses the *polyhedral library* [22].

This is the first method to be proposed that enumerates both integer and rational polytopes using a single coherent algorithm. Moreover, the output in the form of Ehrhart polynomials is simple and gives exact parametric expressions for the number of integer points. Other work dealing with this same problem is presented below.

M. Haghighat and C. Polychronopoulos present in [10, 11] a method for volume computation. The form of their answers is different because of the *min* and *max* expressions they introduce. Their results tend to be more complicated.

Ferrante, Sarkar and Thrash in [8] present methods for computing the number of distinct memory locations and cache lines accessed by a loop nest. These methods cannot handle unions of polytopes, do not compute symbolic answers, often compute only an approximation and use expensive methods to handle the cache lines touched by a set of references.

The method described by Pugh in [16] consists in a set of techniques, dedicated to simplifying the initial set of constraints. Applying all these tools would be expensive for important sets of constraints with rational coefficients. Moreover, their application is often intuitive: it consists in some cases in adding auxiliary variables, and requires the use of standard formulas for sums of powers (the author expects it will be sufficient to hard code the formulas for powers up to 10). All these facts clearly show some limits in the general use of these techniques.

Tawbi in [21, 20] first uses a polyhedral splitting technique to characterize the validity domains, and an elimination of the variables in a predetermined order. However, her method does not systematically compute exact answers like ours.

We have illustrated this paper with one example from a possible application. There are a lot of other applications in the analysis and transformation of loop nests, some of which can be found in works cited hereabove or in [2, 3]. Other applications may also be found in other fields as sequential program analysis, combinatorics and optimization.

Acknowledgement

We are grateful to Doran K. Wilde for his useful comments and suggestions.

References

[1] M.S. Bazaraa, H.D. Sherali and C.M. Shetty. *Nonlinear Programming, Theory and Algorithms, second edition*. Wiley, New York, 1993.

[2] Ph. Clauss. The volume of a lattice polyhedron to enumerate processors and parallelism. Research Report ICPS 95-11, Submitted to publication, 1995. http://icps.u-strasbg.fr/pub-95/pub-95-11.ps.gz

[3] Ph. Clauss. Counting Solutions to Linear and Nonlinear Constraints through Ehrhart polynomials: Applications to Analyze and Transform Scientific Programs. *10th ACM Int. Conf. on Supercomputing*, Philadelphia, May 1996. Also available as Research Report ICPS 95-20. http://icps.u-strasbg.fr/pub-96/pub-95-20.ps.gz

[4] E. Ehrhart. Sur les polyèdres rationnels homothétiques à n dimensions. *C.R. Acad. Sci. Paris*, 254:616-618, 1962.

[5] E. Ehrhart. Sur un problème de géométrie diophantienne linéaire I. *J. Reine Angew. Math.*, 226:1-29, 1967.

[6] E. Ehrhart. Sur un problème de géométrie diophantienne linéaire II. *J. Reine Angew. Math.*, 227:25-49, 1967.

[7] E. Ehrhart. *Polynômes arithmétiques et Méthode des Polyèdres en Combinatoire*. International Series of Numerical Mathematics, vol.35, Birkhäuser Verlag, Basel/Stuttgart, 1977.

[8] J. Ferrante, V. Sarkar and W. Thrash. On estimating and enhancing cache effectiveness. *Advances in Languages and Compilers for Parallel Processing*, pages 328-343. The MIT Press, 1991.

[9] P.M. Gruber and C.G. Lekkerkerker. *Geometry of Numbers*. North-Holland, Amsterdam, 1987.

[10] M. Haghighat and C. Polychronopoulos. Symbolic analysis: A basis for parallelization, optimization and scheduling of programs. *U. Banerjee et al., editor, Languages and Compilers for Parallel Computing*. Springer-Verlag, August 1993. LNCS 768, *Proc. of the 6th annual worshop on Programming Languages and Compilers for Parallel Computing.*

[11] M. Haghighat and C. Polychronopoulos. Symbolic analysis: A basis for parallelization, optimization and scheduling of programs. Technical Report 1317, CSRD, Univ. of Illinois, August 1993.

[12] V. Loechner and D. K. Wilde. Parameterized polyhedra and their vertices. Research Report ICPS 95-16, Submitted for publication, 1995. http://icps.u-strasbg.fr/pub-95/pub-95-16.ps.gz

[13] I. G. Macdonald. The volume of a lattice polyhedron. *Proc. Camb. Phil. Soc.*, 59:719-726, 1963.

[14] I. G. Macdonald. Polynomials associated with

finite cell-complexes. *J. London Math. Soc.*, 4(2):181-192, 1971.

[15] T. H. Mattheiss and D. Rubin. A survey and comparison of methods for finding all vertices of convex polyhedral sets. *Mathematics of Operations Research*, 5(2):167-185, 1980.

[16] W. Pugh. Counting Solutions to Presburger Formulas: How and Why. *Proc. of the 1994 ACM SIGPLAN Conference on Programming Language Design and Implementation*, 1994.

[17] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley, New York, 1986.

[18] R. P. Stanley. *Combinatorics and Commutative Algebra*. Birkhäuser, Boston, 1983.

[19] R. P. Stanley. *Enumerative Combinatorics Volume I*. Wadsworth & Brooks/Cole, Monterey, California, 1986.

[20] N. Tawbi. Estimation of nested loop execution time by integer arithmetics in convex polyhedra. *Proc. of the 1994 International Parallel Processing Symposium*, April 1994.

[21] N. Tawbi and P. Feautrier. Processor allocation and loop scheduling on multiprocessor computers. *Proc. of the 1992 International Conference on Supercomputing*, pages 63-71, July 1992.

[22] D. K. Wilde. A library for doing polyhedral operations. *Master's thesis*, Oregon State University, Corvallis, Oregon, december 1993. Also published as IRISA technical report PI 785, Rennes, France, december 1993.