Computational Geometry ••• (••••) •••–•••

Contents lists available at ScienceDirect

Computational Geometry: Theory and Applications

www.elsevier.com/locate/comgeo

Multiple covers with balls I: Inclusion–exclusion ☆

Herbert Edelsbrunner, Mabel Iglesias-Ham *

IST Austria (Institute of Science and Technology Austria), Am Campus 1, 3400 Klosterneuburg, Austria

ARTICLE INFO

Article history:
Received 12 March 2015
Accepted 24 October 2016
Available online xxxx

In memory of a good friend and trusted colleague

Keywords:
Multiple cover with balls
Inclusion–exclusion
Voronoi diagrams
Hyperplane arrangements
Exact computation

ABSTRACT

Inclusion–exclusion is an effective method for computing the volume of a union of measurable sets. We extend it to multiple coverings, proving short inclusion–exclusion formulas for the subset of $\mathbb{R}^n$ covered by at least k balls in a finite set. We implement two of the formulas in dimension $n = 3$ and report on results obtained with our software.

1. Introduction

The work reported in this paper is motivated by configurations of balls that model the organization of DNA inside the nuclei of human cells: the *Spherical* 1 *Mega-base-pairs Chromatin Domain*, or *SCD model*, which is supported by high resolution microscopic observations [1,2]. It was recently confirmed that inside the chromosome territories in eukaryotic cells, DNA is compartmentalized in sequences of highly interacting segments of roughly the same length [3]. Each segment consists of about one million base pairs which are rolled up in a shape that resembles a round ball, and the shapes are tightly arranged within a restricted space.

Modeling such a configuration as a *packing* – in which the balls are rigid and allowed to touch but not overlap – is too restrictive because the rolled up base pairs push against each other and deform to cover more empty space than is otherwise possible. Similarly, modeling the configuration as a *covering* – in which the balls overlap and cover space without gaps – is not realistic because some empty space is necessary to facilitate the expression and replication of the DNA. We refer to [4] for a representative text in the rich mathematical literature on packings and coverings with balls. For the reason mentioned before, we are motivated to consider configurations that lie between these two extremes: the balls are allowed to overlap and they do not necessarily cover the entire space; see also [5]. Given such a configuration, we are interested in quantifications. For packings and coverings, it is customary to compute the *density*, which is the expected number of balls that contain a random point. This measure can also be used for more general configurations, but there are other choices. To mention one, we may be interested in the set of points each covered by exactly one ball; its volume is the difference between the volume of the union and of the 2-fold cover of the balls. It requires the ability to measure the set of points covered by at least two balls, which is a special case of the question addressed in this paper.

☆ This work is partially supported by the TopoSys project FP7-ICT-318493-STREP, and by ESF under the ACAT Research Network Programme.
* Corresponding author.
E-mail addresses: edels@ist.ac.at (H. Edelsbrunner), miglesias@ist.ac.at (M. Iglesias-Ham).

http://dx.doi.org/10.1016/j.comgeo.2017.06.014
0925-7721/© 2017 Elsevier B.V. All rights reserved.

1						
1	1					
1	2	1				
1	3	3	1			
1	4	6	4	1		
1	5	10	10	5	1	
1	6	15	20	15	6	1

1						
-1	1					
1	-2	1				
-1	3	-3	1			
1	-4	6	-4	1		
-1	5	-10	10	-5	1	
1	-6	15	-20	15	-6	1

Fig. 1. The first few non-zero rows of the Pascal triangle on the *left*, and of the alternating Pascal triangle on the *right*.

Prior work and results. An effective method for computing the volume of a union of balls, or possibly more general sets, is the principle of inclusion–exclusion. It has a long history in mathematics and is attributed to Abraham de Moivre (1667–1754) but appeared first in writings of Daniel da Silva (1854) and of James Joseph Sylvester (1883). Given a finite collection of measurable sets, $\mathcal{X}$, in $\mathbb{R}^n$, it asserts that the volume of the union is the alternating sum of the volumes of the common intersections of the sets in all subcollections $Q \subseteq \mathcal{X}$. The formula can be generalized to *k-fold covers*, which we define as the set $\mathbb{X}_k$ of points in $\mathbb{R}^n$ that belong to at least k of the sets:

$$\mathrm{Vol}(\mathbb{X}_k) = \sum_{i \geq k} (-1)^{i-k} \binom{i-1}{k-1} \sum_{Q \in \binom{\mathcal{X}}{i}} \mathrm{Vol}(\bigcap Q); \tag{1}$$

in which $\binom{\mathcal{X}}{i}$ denotes the collection of subsets of size i, see for example Chapter IV of Feller's textbook on probability [6, p. 110]. Since we need (1) in the proofs of the short inclusion–exclusion formulas, we give our own proof using the Pascal triangle and its alternating form. If the measurable sets are balls, we write $\mathcal{B}$ for the collection, and $\mathbb{B}_k$ for the k-fold cover. Using the power distance of a point to a ball, the *order-k Voronoi diagram* identifies all collections $Q \subseteq \mathcal{B}$ of size k for which there are points so that the balls in Q are the k closest; see e.g. [7]. Restricting (1) to terms that correspond to cells of the order-k Voronoi diagram, we get a short inclusion–exclusion formula:

$$\mathrm{Vol}(\mathbb{B}_k) = \sum_{\sigma \in \mathcal{V}_k} (-1)^{\operatorname{codim} \gamma(\sigma)} \mathrm{Vol}(\bigcap Q_{\gamma(\sigma)}), \tag{2}$$

see the Order-k Pie Theorem in Section 4 for details. Every γ is a cell of the order-k Voronoi diagram, with at least k and at most $k+n$ balls in the corresponding collection $Q_\gamma \subseteq \mathcal{B}$. Relation (2) generalizes the inclusion–exclusion formula of Naiman and Wynn [8] from the union to more general k-fold covers. We also prove a slightly stronger version of (2) in which the sum ranges over the subcollection of cells that have a non-empty common intersection with the balls that define them. It generalizes the inclusion–exclusion formula based on alpha shapes given in [9]. To reduce the size of the terms, we use levels in hyperplane arrangements in $\mathbb{R}^{n+1}$ and inclusion–exclusion formulas for general polyhedra; see [10,11], and obtain another short inclusion–exclusion formula for the n-dimensional volume of the k-fold cover:

$$\mathrm{Vol}(\mathbb{B}_k) = \sum_{Q \in \mathcal{L}_k} L_Q \cdot \mathrm{Vol}(\bigcap Q); \tag{3}$$

see the Level-k Pie Theorem in Section 5 for details. The collections $Q \subseteq \mathcal{B}$ correspond to affine subspaces of the arrangement, with size between 1 and $n+1$. For $k=1$, the formulas (2) and (3) are the same. Importantly, we have a slightly stronger version of (3) in which all collections of balls are independent. Among other advantages, this additional property eliminates an otherwise necessary case analysis and thus simplifies computer implementations. As mentioned above, the short inclusion–exclusion formulas in (2) and (3) have applications in the study of the spatial organization of chromosomes. We have implemented the formulas in dimension $n=3$, using software supporting exact arithmetic [12,13] and volume formulas for the common intersection of 3-dimensional balls [14].

Outline. Section 2 extends the principle of inclusion–exclusion from unions to k-fold covers. Section 3 provides background on Voronoi diagrams and hyperplane arrangements. Sections 4 and 5 prove short inclusion–exclusion formulas for k-fold covers with balls in $\mathbb{R}^n$. Section 6 presents results of computational experiments. Section 7 concludes the paper.

2. The combinatorial formula

In this section, we explain how the inclusion–exclusion formula for the volume of a union of measurable sets can be extended to k-fold covers. In the context of probability theory, the same extension can be found in [6, p. 110]. We begin with a combinatorial result on Pascal triangles.

Recall that the *Pascal triangle* is a 2-dimensional organization of the binomial coefficients, and the *alternating Pascal triangle* is the same except that the coefficients are listed with alternating sign; see Fig. 1. We think of them as (infinitely large) matrices that can be multiplied. To talk about the product, we introduce notation for the u-th row of the Pascal triangle and the v-th column of the alternating Pascal triangle, $R_u, C_v : \mathbb{Z} \to \mathbb{Z}$ defined by