Approximating the Volume of Unions and Intersections of High-Dimensional Geometric Objects

Karl Bringmann[1] and Tobias Friedrich[2,3]

[1] Universität des Saarlandes, Saarbrücken, Germany
[2] Max-Planck-Institut für Informatik, Saarbrücken, Germany
[3] International Computer Science Institute, Berkeley, CA, USA

Abstract. We consider the computation of the volume of the union of high-dimensional geometric objects. While showing that this problem is **#P**-hard already for very simple bodies (i.e., axis-parallel boxes), we give a fast FPRAS for all objects where one can: (1) test whether a given point lies inside the object, (2) sample a point uniformly, (3) calculate the volume of the object in polynomial time. All three oracles can be weak, that is, just approximate. This implies that Klee's measure problem and the hypervolume indicator can be approximated efficiently even though they are **#P**-hard and hence cannot be solved exactly in time polynomial in the number of dimensions unless $\mathbf{P} = \mathbf{NP}$. Our algorithm also allows to approximate efficiently the volume of the union of convex bodies given by weak membership oracles.

For the analogous problem of the intersection of high-dimensional geometric objects we prove **#P**-hardness for boxes and show that there is no multiplicative polynomial-time $2^{d^{1-\varepsilon}}$-approximation for certain boxes unless $\mathbf{NP} = \mathbf{BPP}$, but give a simple additive polynomial-time ε-approximation.

1 Introduction

Given n bodies in the d-dimensional space, how efficiently can we compute the volume of the union and the intersection? We consider this basic geometric problem for different kinds of bodies. The tractability of this problem highly depends on the representation and the complexity of the given objects. For many classes of objects already computing the volume of one body can be hard. For example, calculating the volume of a polyhedron given either as a list of vertices or as a list of facets is **#P**-hard [6, 14]. For convex bodies given by a membership oracle one can also show that even though there can be no deterministic $(\mathcal{O}(1)d/\log d)^d$-approximation for $d \geq 2$ [2], one can still approximate the volume by an FPRAS (fully polynomial-time randomized approximation scheme). In a seminal paper Dyer, Frieze, and Kannan [7] gave an $\mathcal{O}^*(d^{23})$ algorithm, which was subsequently improved in a series of papers to $\mathcal{O}^*(d^4)$ [16] (where the asterisk hides powers of the approximation ratio and $\log d$).

S.-H. Hong, H. Nagamochi, and T. Fukunaga (Eds.): ISAAC 2008, LNCS 5369, pp. 436–447, 2008.

Volume computation of unions can be hard not only for bodies whose volume is hard to calculate. One famous example for this is *Klee's Measure Problem* (KMP). Given n axis-parallel boxes in the d-dimensional space, it asks for the measure of their union. In 1977, Victor Klee showed that it can be solved in time $\mathcal{O}(n \log n)$ for $d = 1$ [15]. This was generalized to $d > 1$ dimensions by Bentley who presented an algorithm which runs in $\mathcal{O}(n^{d-1} \log n)$, which was later improved by van Leeuwen and Wood [20] to $\mathcal{O}(n^{d-1})$. In FOCS '88, Overmars and Yap [17] obtained an $\mathcal{O}(n^{d/2} \log n)$ algorithm. This was the fastest algorithm for $d \geq 3$ until, on this years SoCG, Chan [5] presented a slightly improved version of Overmars and Yap's algorithm that runs in time $n^{d/2} 2^{\mathcal{O}(\log^* n)}$, where $\log^*$ denotes the iterated logarithm. So far, the only known lower bound is $\Omega(n \log n)$ for any d [8]. Note that the worst-case combinatorial complexity (i.e., the number of faces of all dimensions on the boundary of the union) of $\Theta(n^d)$ does not imply any bounds on the computational complexity. There are various algorithms for special cases, e.g., for hypercubes [1, 11] and unit hypercubes [4]. In this paper we explore the opposite direction and examine the problem of the union of more general geometric objects.

An important special case of KMP is the *hypervolume indicator* (HYP) [21] where all boxes are required to share a common vertex. It is also known as the "Lebesgue measure", the "S-metric" and "hyperarea metric" and is a popular measure of fitness of Pareto sets in multi-objective optimization. There, it measures the number of solutions dominated by a Pareto set. More details can be found in Section 4.

Our Results

It is not hard to see that HYP and KMP are **#P**-hard (see Theorem 1). Hence they cannot be solved in time polynomial in the number of dimensions unless **P** = **NP**. This shows that *exact* volume computation of unions is intractable for all classes of bodies that contain axis-parallel boxes.

This motivates the development of *approximation algorithms* for the volume computation of unions. This question was untackled until now – approaches exist only for discrete sets (see, e.g., Karp, Luby, and Madras [13] for an FPRAS for #DNF which is similar to our algorithm). We give an efficient FPRAS for a huge class of bodies including boxes, spheres, polytopes, schlicht domains, convex bodies determined by an oracle and all affine transformations of those objects mentioned before. The underlying bodies B just have to support the following oracle queries in polynomial time:

- PointQuery(x, B): Is point $x \in \mathbb{R}^d$ an element of body B (approximately)?
- VolumeQuery(B): What is the volume of body B (approximately)?
- SampleQuery(B): Return a random (almost) uniformly distributed point $x \in B$.

PointQuery is a very natural condition which is fulfilled in almost all practical cases. The VolumeQuery condition is important as it could be the case that no efficient approximation of the volume of one of the bodies itself is possible. This, of course, prevents an efficient approximation of the union of such bodies. The SampleQuery is crucial for our FPRAS. In Section 2.3 we will show that it is efficiently computable for a wide range of bodies.

Note that it suffices that all three oracles are weak. More precisely, we allow the following *relaxation* for every body B ($\text{VOL}(B)$ denotes the volume of a body B in the standard Lebesgue measure on $\mathbb{R}^d$, more details are given in Section 2):

- PointQuery(x, B) answers true iff $x \in B'$ for a fixed $B' \subset \mathbb{R}^d$ with $\text{VOL}((B' \setminus B) \cup (B \setminus B')) \leq \varepsilon_{\text{P}} \text{VOL}(B)$.
- VolumeQuery(B) returns a value V' with $(1 - \varepsilon_{\text{V}})\,\text{VOL}(B) \leq V' \leq (1 + \varepsilon_{\text{V}})\,\text{VOL}(B)$.
- SampleQuery(B) returns only an *almost* uniformly distributed random point [10], that is, it suffices to get a random point $x \in B'$ (with B' as above) such that for the probability density f we have for every point x: $|f(x) - 1/\text{VOL}(B')| < \varepsilon_{\text{S}}$.

Let $P(d)$ be the worst PointQuery runtime of any of our bodies, analogously $V(d)$ for VolumeQuery, and $S(d)$ for SampleQuery. Then our FPRAS has a runtime of $\mathcal{O}(nV(d) + \frac{n}{\varepsilon^2}(S(d) + P(d)))$ for producing an ε-approximation[1] with probability $\geq \frac{3}{4}$ if the errors of the underlying oracles are small, i.e., $\varepsilon_{\text{S}}, \varepsilon_{\text{P}}, \varepsilon_{\text{V}} \leq \frac{\varepsilon^2}{47n}$. For example for boxes, that is, for KMP and HYP, this reduces to $\mathcal{O}(\frac{dn}{\varepsilon^2})$ and is the first FPRAS for both problems. In Section 2.3 we also show that our algorithm is an FPRAS for the volume of the union of convex bodies.

The canonic next question is the computation of the volume of the *intersection of bodies* in $\mathbb{R}^d$. It is clear that most of the problems from above apply to this question, too. **#P**-hardness for general, i.e., not necessarily axis-parallel, boxes follows directly from the hardness of computing the volume of a polytope [6, 14]. This leaves open whether there are efficient approximation algorithms for the volume of intersection. In Section 3 we show that there cannot be a (deterministic or randomized) multiplicative $2^{d^{1-\varepsilon}}$-approximation in general, unless $\mathbf{NP} = \mathbf{BPP}$ by identifying a hard subproblem. Instead we give an additive ε-approximation, which is therefore the best we can hope for. It has a runtime of $\mathcal{O}(nV(d) + \varepsilon^{-2}S(d) + \frac{n}{\varepsilon^2}P(d))$, which gives $\mathcal{O}(\frac{dn}{\varepsilon^2})$ for boxes.

2 Volume Computation of Unions

In this section we show that the volume computation of unions is **#P**-hard already for axis-parallel boxes that have one vertex at the origin, i.e., for HYP. After that we give an FPRAS for approximating the volume of the union of bodies which satisfy the three aforementioned oracles and describe several large classes of objects for which the oracles can be answered efficiently.

[1] We will always assume that ε is small, that is, $0 < \varepsilon < 1$.

2.1 Computational Complexity of Union Calculations

Consider the following problem: Let $\mathcal{S}$ be a set of n axis-parallel boxes in $\mathbb{R}^d$ of the form $B = [a_1, b_1] \times \cdots \times [a_d, b_d]$ with $a_i, b_i \in \mathbb{R}, a_i < b_i$. The volume of one such box is $\text{VOL}(B) = \prod_{i=1}^{d}(b_i - a_i)$. To compute the volume of the union of these boxes is known as Klee's Measure Problem (KMP), while we call the problem HYP (for hypervolume) if we have $a_i = 0$ for all $i \in [d]$.

The following Theorem 1 proves **#P**-hardness of HYP and KMP. This is the first hardness result for HYP. To the best of our knowledge there is also no published result that explicitly states that KMP is **#P**-hard. However, without mentioning this implication, Suzuki and Ibaraki [19] sketch a reduction from #SAT to KMP. In the following theorem we reduce #MON-CNF to HYP, which counts the number of satisfying assignments of a Boolean formula in conjunctive normal form in which all variables are unnegated. While the problem of deciding satisfiability of such formula is trivial, counting the number of satisfying assignments is **#P**-hard and even approximating it in polynomial time by a factor of $2^{d^{1-\varepsilon}}$ for any $\varepsilon > 0$ is **NP**-hard, where d is the number of variables (see Roth [18] for a proof).

Theorem 1. HYP *and* KMP *are* **#P**-*hard.*

Proof. To show the theorem, we reduce #MON-CNF to HYP. The hardness of KMP follows immediately. Let $f = \bigwedge_{k=1}^{n} \bigvee_{i \in C_k} x_i$ be a monotone Boolean formula given in CNF with $C_k \subset [d] := \{1, \ldots, d\}$, for $k \in [n]$, d the number of variables, n the number of clauses. Since the number of satisfying assignments of f is equal to 2^d minus the number of satisfying assignments of its negation, we instead count the latter: Consider the negated formula $\bar{f} = \bigvee_{k=1}^{n} \bigwedge_{i \in C_k} \neg x_i$. First, we construct a box $A_k = [0, a_1^k] \times \cdots \times [0, a_d^k]$ in $\mathbb{R}^d$ for each clause C_k with one vertex at the origin and the opposite vertex at $(a_1^k, \ldots, a_d^k)$, where we set

$$a_i^k = \begin{cases} 1, & \text{if} \quad i \in C_k \\ 2, & \text{otherwise} \end{cases}, \qquad i \in [d].$$

Observe that the union of the boxes A_k can be written as a union of boxes of the form $B_{x_1,\ldots,x_d} = [x_1, x_1 + 1] \times \cdots \times [x_d, x_d + 1]$ with $x_i \in \{0, 1\}, i \in [d]$. Moreover, $B_{x_1,\ldots,x_d}$ is a subset of the union $\bigcup_{k=1}^{n} A_k$ iff it is a subset of some A_k iff we have $a_i^k \geq x_i + 1$ for $i \in [d]$ iff $a_i^k = 2$ for all i with $x_i = 1$ iff $i \notin C_k$ for all i with $x_i = 1$ iff $(x_1, \ldots, x_d)$ satisfies $\bigwedge_{i \in C_k} \neg x_i$ for some k iff $(x_1, \ldots, x_d)$ satisfies $\bar{f}$. Hence, since $\text{VOL}(B_{x_1,\ldots,x_d}) = 1$, we have $\text{VOL}(\bigcup_{k=1}^{n} A_k) = |\{(x_1, \ldots, x_d) \in \{0, 1\}^d \mid (x_1, \ldots, x_d) \text{ satisfies } \bar{f}\}|$. Thus a polynomial time algorithm for HYP would result in a polynomial time algorithm for #MON-CNF, which proves the claim. ☐

For integer coordinates it is easy to see that KMP $\in$ **#P** and hence that in this case KMP and HYP are even **#P**-complete.

2.2 Approximation Algorithm for the Volume of Unions

In this section we present an FPRAS for computing the volume of the union of objects for which we can answer PointQuery, VolumeQuery, and SampleQuery in polynomial time. The input of our algorithm ApproxUnion are the approximation ratio ε and the bodies $\{B_1, \ldots, B_n\}$ in $\mathbb{R}^d$ defined by the three oracles. It computes a number $\tilde{U} \in \mathbb{R}$ such that

$$\Pr\left[(1-\varepsilon)\,\mathrm{VOL}\Big(\bigcup_{i=1}^{n} B_i\Big) \le \tilde{U} \le (1+\varepsilon)\,\mathrm{VOL}\Big(\bigcup_{i=1}^{n} B_i\Big)\right] \ge \frac{3}{4}$$

in time polynomial in n, $1/\varepsilon$ and the query runtimes. The number $\frac{3}{4}$ is arbitrary and can be increased to every number $1-\delta$, $\delta > 0$ by a factor of $\log(1/\delta)$ in the runtime by running algorithm ApproxUnion $\log(1/\delta)$ many times and returning the median of the reported values for $\tilde{U}$.

We are following the algorithm of Karp and Luby [12] which the authors used for approximating #DNF and other counting problems on discrete sets. The two main differences are that here we are handling continuous bodies in $\mathbb{R}^d$ and that we allow erroneous oracles.

Algorithm 1. ApproxUnion $(\mathcal{S}, \varepsilon, \varepsilon_{\mathrm{P}}, \varepsilon_{\mathrm{V}}, \varepsilon_{\mathrm{S}})$ calculates an ε-approximation of $\mathrm{VOL}(\bigcup_{i=1}^{n} B_i)$ for a set of bodies $S = \{B_1, \ldots, B_n\}$ in $\mathbb{R}^d$ determined by the oracles PointQuery, VolumeQuery and SampleQuery with error ratios $\varepsilon_{\mathrm{P}}, \varepsilon_{\mathrm{V}}, \varepsilon_{\mathrm{S}}$.

```
M := 0, C := 0, ε̃ := (ε−ε_V)/(1+ε_V), C̃ := (1+ε_S)(1+ε_V)(1+ε_P) / ((1−ε_V)(1−ε_P)), T := 24 ln(2)(1+ε̃)n / (ε̃²−8(C̃−1)n)
for all B_i ∈ S do
   compute V'_i := VolumeQuery(B_i)
od
V' := Σ_{i=1}^n V'_i
while C ≤ T do
   choose i ∈ [n] with probability V'_i/V'
   x := SampleQuery(B_i)
   repeat
      if C > T then return T·V'/(nM)
      choose random j ∈ [n] uniformly
      C := C + 1
   until PointQuery (x, B_j)
   M := M + 1
od
return T·V'/(nM)
```

We assume that we are given upper bounds ε_{P}, ε_{S} and ε_{V} for the error ratios of the oracles. In the algorithm we first compute the runtime T and then via VolumeQuery the volume V'_i for every given body B_i and their sum $V' =$

$\sum_{i=1}^{n} V_i'$. Then, in a loop, we choose a random $i \in [n]$, where we choose an i with probability $\frac{V_i'}{V'}$ and a random (almost) uniformly distributed point $x \in B_i$ via SampleQuery. Then we repeatedly choose a random $j \in [n]$ uniformly and check, whether $x \in B_j$: PointQuery(x, B_j) returns true iff $x \in B_j'$. If this is the case, we leave the inner loop and increase the counter M. This random variable is in the end used to calculate the result $\frac{T \cdot V'}{nM}$.

In the full version of the paper [3] we show correctness of ApproxUnion, that is, we show that it returns an ε-approximation with probability $\geq \frac{3}{4}$ and $T = \mathcal{O}(\frac{n}{\varepsilon^2})$ if $\varepsilon_{\mathrm{S}}, \varepsilon_{\mathrm{P}}, \varepsilon_{\mathrm{V}} \leq \frac{\varepsilon^2}{47n}$. The last inequality reflects the fact that we cannot be arbitrarily accurate if the given oracles are inaccurate. If all oracles can be calculated accurately, i.e., if $\varepsilon_{\mathrm{P}} = \varepsilon_{\mathrm{S}} = \varepsilon_{\mathrm{V}} = 0$, the algorithm runs for just $T = \frac{8 \ln(8)(1+\varepsilon)n}{\varepsilon^2}$ many steps. The runtime of ApproxUnion is clearly

$$\mathcal{O}(n \cdot V(d) + M \cdot S(d) + T \cdot P(d)) = \mathcal{O}(n \cdot V(d) + T \cdot (S(d) + P(d))),$$

where $V(d)$ is the worst VolumeQuery time for any of the bodies, analogously $S(d)$ for SampleQuery and $P(d)$ for PointQuery. This equals $\mathcal{O}(n \cdot V(d) + \frac{n}{\varepsilon^2} \cdot (S(d) + P(d)))$ if $\varepsilon_{\mathrm{S}}, \varepsilon_{\mathrm{P}}, \varepsilon_{\mathrm{V}} \leq \frac{\varepsilon^2}{47n}$.

For boxes all three oracles can be computed exactly in $\mathcal{O}(d)$. This implies that our algorithm ApproxUnion gives an ε-approximation of KMP and HYP with probability $\geq \frac{3}{4}$ in runtime $\mathcal{O}(\frac{nd}{\varepsilon^2})$.

2.3 Classes of Objects Supported by Our FPRAS

To find an FPRAS for the union of a certain class of geometric objects now reduces to calculating the respective PointQuery, VolumeQuery and SampleQuery in polynomial time. We assume that we can get a random real number in constant time. Then all three oracles can be calculated in time $\mathcal{O}(d)$ for d-dimensional boxes. This already yields an FPRAS for the volume of the union of arbitrary boxes, e.g., for KMP and HYP. Note that if we have a body for which we can answer all those queries, all affine transformations of this body fulfill these three oracles, too. We will now present three further classes of geometric objects.

Generalized spheres and boxes: Let $\mathbf{B}_k$ be the class of boxes of dimension k, i.e., $\mathbf{B}_k = \{[a_1, b_1] \times \cdots \times [a_k, b_k] \mid a_i, b_i \in \mathbb{R}, a_i < b_i\}$ and $\mathbf{S}_l$ the class of spheres of dimension l. We can combine any box $B \in \mathbf{B}_k$ and sphere $S \in \mathbf{S}_{d-k}$ to get a d-dimensional object $B \times S$. Furthermore, we can permute the dimensions afterwards to get a generalized "box-sphere". In $\mathbb{R}^3$ this corresponds to boxes, spheres and cylinders. To calculate the respective VolumeQuery, PointQuery and SampleQuery is a standard task of geometry.

Schlicht domains: Let $a_i, b_i \colon \mathbb{R}^{i-1} \to \mathbb{R}$ be continuous functions with $a_i \leq b_i$, where a_1, b_1 are constants. Let $K \subset \mathbb{R}^d$ be defined as the set of all points $(x_1, \ldots, x_d) \in \mathbb{R}^d$ such that $a_1 \leq x_1 \leq b_1, a_2(x_1) \leq x_2 \leq$

$b_2(x_1), \ldots, a_d(x_1, \ldots, x_{d-1}) \leq x_d \leq b_d(x_1, \ldots, x_{d-1})$. K is called a schlicht domain in functional analysis. Fubini's theorem for schlicht domains states that we can integrate a function $f\colon K \to \mathbb{R}$ by iteratively integrating first over x_d, then over x_{d-1}, ..., until we reach x_1. This way, by integrating $f \equiv 1$, we can compute the volume of a schlicht domain, as long as the integrals are computable in polynomial time, and thus answer a VolumeQuery. Similarly, we can choose a random uniformly distributed point inside K: Let $K(y) = \{(x_1, \ldots, x_d) \in K \mid x_1 = y\}$. Then $K(y)$ is another schlicht domain for every $a_1 \leq y \leq b_1$. Assume that we can determine the volume of every such $K(y)$ and the integral $I(y) = \int_{a_1}^{y} K(x)dx$. Then the inverse function $I^{-1}\colon [0, V] \to \mathbb{R}$, where $V = \int_{a_1}^{b_1} K(x)dx$ is the volume of K, allows us to choose a y in $[a_1, b_1]$ with probability proportional to $\mathrm{VOL}(K(y))$. By this we can iteratively choose a value y for x_1 and recurse to find a uniformly random point $(y_2, \ldots, y_d)$ in $K(y)$, plugging both together to get a uniformly distributed point $(y_1, \ldots, y_d)$ in K. Hence, as long as we can compute the involved integrals and inverse functions (or at least approximate them good enough), we can answer SampleQuery. Since PointQuery is trivially computable – as long as we can evaluate a_i and b_i efficiently – this gives an example showing that the classes of objects that fulfill our three conditions include not only convex bodies, but also certain schlicht domains.

Convex bodies: As mentioned in the introduction, exact calculation of VolumeQuery for a polyhedron given as a list of vertices or facets is **#P**-hard [6, 14]. Since there are randomized approximation algorithms (see Dyer et al. [7] for the first one) for the volume of a convex body determined by a membership oracle, we can answer VolumeQuery approximately. The same holds for SampleQuery as these algorithms make use of an almost uniform sampling method on convex bodies. See Lovász and Vempala [16] for a result showing that VolumeQuery can be answered with $\mathcal{O}^*(\frac{d^4}{\varepsilon_V^2})$ questions to the membership oracle and SampleQuery with $\mathcal{O}^*(\frac{d^3}{\varepsilon_S^2})$ queries, for arbitrary errors $\varepsilon_V, \varepsilon_S > 0$ (where the asterisk hides factors of $\log(d)$ and $\log(1/\varepsilon_V)$ or $\log(1/\varepsilon_S)$). PointQuery can naturally be answered with a single question to the membership oracle. By choosing $\varepsilon_V = \varepsilon_S = \frac{\varepsilon^2}{47n}$, we can show [3] that ApproxUnion is an FPRAS for the volume of the union of convex bodies which uses $\mathcal{O}^*(\frac{n^3 d^3}{\varepsilon^4}(d + \frac{1}{\varepsilon^2}))$ membership queries.

Note that all above mentioned classes of geometric objects contain boxes and hence our hardness results still hold and hence an ε-approximation algorithm is the best one can hope for.

3 Volume Computation of Intersections

In this section we are considering the complement to the union problem. We show that surprisingly the volume of a intersection of a set of bodies is often much harder to calculate than its union. For many classes of geometric objects there is even no randomized approximation possible.

As the problem of computing the volume of a polytope is **#P**-hard [6, 14], so is the computation of the volume of the intersection of general (i.e., not necessarily axis-parallel) boxes in $\mathbb{R}^d$. This can be seen by describing a polytope as an intersection of halfplanes and representing these as general boxes.

Now, let us consider the convex bodies again. Trivially, the intersection of convex bodies is convex itself, and from the oracles defining the given bodies $B_1, \dots, B_n$ one can simply construct an oracle, which answers PointQuery for the intersection of those objects: Given a point $x \in \mathbb{R}^d$ it asks all n oracles and returns true iff x lies in all the bodies. One could think now that we can apply the result of Dyer et al. [7] and the subsequent improvements mentioned in the introduction to approximate the volume of the intersection and get an FPRAS for the problem at hand. The problem with that is that the intersection is not "well-guaranteed": There is no point known that lies in the intersection, not to speak of a sphere inside it. However, the algorithm of Dyer et al. [7] relies vitally on the assumption that the given body is well-guaranteed and hence cannot be applied for approximating the volume of the intersection of convex bodies.

We will now present a hard subproblem which shows that the volume of the intersection cannot be approximated (deterministic or randomized) in general.

Definition 1. *Let $p, q \in \mathbb{R}^d_{\geq 0}$. Then $B_p := \{x \mid 0 \leq x_i \leq p_i \ \forall i\}$ is a p-box, $B_{p,q} := B_p \setminus B_q$ is a (p, q)-box, and $\{B_{p,q_1}, B_{p,q_2}, \dots, B_{p,q_n}\}$ is a p-set.*

A (p, q)-box is basically a box where we cut out another box at one corner. The resulting object can itself be a box, too, but in general it is not even convex. It can be seen as the inverse of a box B_p relative to a larger background box B_q. Note that it is easy to calculate the union of a p-set as $\bigcup B_{p,q_i} = B_p \setminus \bigcap B_{q_i}$. On the other hand, the calculation of the intersection of a p-set is **#P**-hard as $\bigcup B_{q_i} = B_p \setminus \bigcap B_{p,q_i}$ by Theorem 1. The following theorem shows that it is not even approximable.

Theorem 2. *Let $p, q_1, \dots, q_n \in \mathbb{R}^d_{\geq 0}$. Then the volume of $\bigcap_{i=1}^n B_{p,q_i}$ cannot be approximated (deterministic or randomized) in polynomial time by a factor of $2^{d^{1-\varepsilon}}$ for any $\varepsilon > 0$ unless* **NP** = **BPP**.

Proof. Consider again the problem #MON-CNF already defined Section 2. We use Roth's result [18] that #MON-CNF cannot be approximated by a factor of $2^{d^{1-\varepsilon}}$ unless **NP** = **BPP** and construct an approximation preserving reduction. Let $f = \bigwedge_{k=1}^n \bigvee_{i \in C_k} x_i$ be a monotone Boolean formula given in CNF with $C_k \subset [d]$, for $k \in [n]$, d the number of variables, n the number of clauses. For every clause C_k we construct a (p, q_k)-box A_k with $p = (2, \dots, 2) \in \mathbb{R}^d$, $q_k = (q_{k,1}, \dots, q_{k,d})$ and $q_{k,i} = 1$ if $i \in C_k$, and $q_{k,i} = 2$ otherwise.

Observe that each A_k can be written as a union of boxes of the form $B_{x_1,\dots,x_d} = [x_1, x_1 + 1] \times \dots \times [x_d, x_d + 1]$ with $x_i \in \{0, 1\}$. Hence, their intersection can also be written as such a union as follows:

$$\bigcap_{k=1}^{n} A_k = \bigcup_{(x_1,\dots,x_d) \in S} B_{x_1,\dots,x_d}$$

for some set $S \subset \{0,1\}^d$. Furthermore, we have $(x_1,\ldots,x_d) \in S$ iff $B_{x_1,\ldots,x_d} \subset A_k$ for all k iff $B_{x_1,\ldots,x_d} \cap \{x \in \mathbb{R}^d_{\geq 0} \mid x \prec q_k\} = \emptyset$ for all k iff $\exists i \in \{1,\ldots,d\} \colon x_i \geq q_{k,i}$ for all k. Since this inequality can only be satisfied if $x_i = 1$ and $q_{k,i} = 1$, which holds iff $i \in C_k$, we have that the former term holds if and only if $\exists i \in C_k \colon x_i = 1$ for all k iff $\bigvee_{j \in C_k} x_j$ is satisfied for all k iff f is satisfied. Hence, we have that the set S equals the set of satisfying assignments of f, so that

$$|\{x \in \{0,1\}^d \mid f(x)=1\}| = |S| \stackrel{(*)}{=} \mathrm{VOL}\Big(\bigcap_{k=1}^{n} A_k\Big) / \mathrm{VOL}\Big(B_{(0,\ldots,0)}\Big) = \mathrm{VOL}\Big(\bigcap_{k=1}^{n} A_k\Big)$$

where $(*)$ comes from the fact that $\bigcap_{k=1}^n A_k$ is composed of $|S|$ many boxes of equal volume and this volume is 1. Hence, we have a polynomial time reduction and inapproximability of the volume of the intersection follows. ☐

This shows that in general there does not exist a polynomial time multiplicative ε-approximation of the volume of the intersection of bodies in $\mathbb{R}^d$. This holds for all classes of objects which include p-sets, e.g. schlicht domains (cf. Section 2.3). Though there is no multiplicative approximation, we can still give an additive approximation algorithm, that is, we can efficiently find a number $\tilde{V}$ such that

$$\Pr[V - \varepsilon \cdot V_{\min} \leq \tilde{V} \leq V + \varepsilon \cdot V_{\min}] \geq \frac{3}{4}$$

where V is the correct volume of the intersection and $V_{\min}$ is the minimal volume of any of the given bodies $B_1,\ldots,B_n$. If we could replace $V_{\min}$ by V in the equation above, we would have an FPRAS. This is not possible in general as the ratio of V and $V_{\min}$ can be arbitrarily small. Hence, such an ε-approximation is not relative to the exact result, but to the volume of some greater body. This is an additive approximation since after rescaling, so that $V_{\min} \leq 1$ we get an additive error of ε. Clearly, we get the result from above quite easily by uniform sampling in the body $B_{\min}$ corresponding to the volume $V_{\min}$. From Bernstein's inequality we know that for N proportional to $1/\varepsilon^2$ and $\tilde{V} = 1/N(Z_1 + \ldots + Z_N)$, where Z_i is a random variable valued 1, if the i-th sample point $x_i = \textsc{SampleQuery}(B_{\min})$ lies in the intersection of $B_1,\ldots,B_n$, and 0 otherwise, $\tilde{V}$ approximates V with absolute error ε. This gives an approximation algorithm with runtime $\mathcal{O}(nV(d) + \frac{1}{\varepsilon^2}S(d) + \frac{n}{\varepsilon^2}P(d))$, yielding $\mathcal{O}(\frac{dn}{\varepsilon^2})$ for boxes.

4 The Hypervolume Indicator

As an application of our results from Section 2 we now analyze the complexity of the hypervolume indicator which is widely used in evolutionary multi-objective optimization. In multi-objective optimization the aim is is to minimize (or maximize) d objective functions $f_i \colon S \to \mathbb{R}$, $1 \leq i \leq d$, over a search space $S \subseteq \mathbb{R}^d$. As these objectives are often conflicting, one does not generally search for a single optimum, but rather for a set of good compromise solutions. In order to compare different solutions, we impose an order $\preceq$ on the d-tuples

Table 1. Results for the computational complexity of the calculation of the volume of union and intersection (asymptotic in the number of dimensions d)

geometric objects	**volume of the union**	**volume of the intersection**
axis-parallel boxes	**#P**-hard + FPRAS	easy
general boxes	**#P**-hard + FPRAS	**#P**-hard
p-sets	easy	**#P**-hard + APX-hard
schlicht domains	**#P**-hard + FPRAS[2]	**#P**-hard + APX-hard
convex bodies	**#P**-hard + FPRAS	**#P**-hard

$f(x) = (f_1(x), \dots, f_d(x))$, $x \in S$, by letting $x_1 \preceq x_2$ whenever $f_i(x_1) \leq f_i(x_2)$ for each i, $1 \leq i \leq d$. If $M \subset S$ is such that $x_1, x_2 \in M$ implies $x_1 \npreceq x_2$, then we call M a "Pareto front" and $f(M)$ a "Pareto set". Pareto fronts correspond to sets of maximal solutions to the optimization problem given by (S, f). The functions f_i are often assumed to be monotone with respect to some ordering on S whence f is bijective and the optimization problem reduces to identifying "good" Pareto sets.

How to compare Pareto sets lies at the heart of research in multi-objective optimization. One measure that has been the subject of much recent study is the so-called "hypervolume indicator" (HYP). It measures the space dominated by the Pareto set relative to a reference point $r \in \mathbb{R}^d$. For a Pareto set M, the hypervolume indicator is

$$\begin{aligned} I(M) &= \text{VOL}\big(\big\{x \in \mathbb{R}^d_{\geq 0} \mid \exists p \in M \text{ so that } r \preceq x \preceq p\big\}\big) \\ &= \text{VOL}\big(\textstyle\bigcup_{p \in M} \big\{x \in \mathbb{R}^d_{\geq 0} \mid r \preceq x \preceq p\big\}\big). \end{aligned}$$

To simplify the presentation we assume $r = (0, \dots, 0)$ which can be achieved by setting $f'(x) := f(x) + r$. Using the notation of Definition 1, we can reduce the hypervolume to the well-understood union problem $I(M) = \text{VOL}(\bigcup_{p \in M} B_p)$.

The hypervolume was first proposed and employed for multi-objective optimization by Zitzler and Thiele [21]. Several algorithms have been developed. It was open so far whether a polynomial algorithm for HYP is possible. Our **#P**-hardness result for HYP (Theorem 1) dashes the hope for a subexponential algorithm (unless **P** = **NP**) and motivates to examine approximation algorithms. Our algorithm ApproxUnion gives an ε-approximation of the hypervolume indicator with probability $1-\delta$ in time $\mathcal{O}(\log(1/\delta)\, nd/\varepsilon^2)$. As its runtime is just linear in n and d, it is not only the first proven FPRAS for HYP, but also a very practical algorithm.

5 Discussion and Open Problems

We have proven **#P**-hardness for the exact computation of the volume of the union of bodies in $\mathbb{R}^d$ as long as the class of bodies includes axis-parallel boxes.

[2] If the integrals are computable in polynomial time (cf. Section 2.3).

The same holds for the intersection if the class of bodies contains general boxes. We have also presented an FPRAS for approximating the volume of the union of bodies that allow three very natural oracles. Very recently, there appeared a few deterministic polynomial-time approximations (FPTAS) for hard counting problems (e.g. [9]). It seems to be a very interesting open question whether there exists a deterministic approximation for the union of some non-trivial class of bodies. Since the volume of convex bodies determined by oracles cannot be approximated to within a factor that is exponential in d [2], the existence of such a deterministic approximation for the union seems implausible. It is also open whether there is a constant C so that HYP or KMP can be efficiently deterministically approximated within a factor of C?

For the intersection we proved that no multiplicative approximation (deterministic or randomized) is possible for p-sets (cf. Definition 1), but we also presented a very simple additive approximation algorithm for the intersection problem. It would be interesting to know if there is a hard class for multiplicative approximation which contains only convex bodies.

Our results are summarized in Table 1. Note the correspondence between axis-parallel boxes and p-sets. The discrete counterpart to their approximability and inapproximability is the approximability of #DNF and the inapproximability of #SAT. One implication of our results is the hardness and a practically efficient approximation algorithm for computing HYP and KMP.

Acknowledgements

We thank Joshua Cooper for suggesting the proof of Theorem 1, Ernst Albrecht for discussions on schlicht domains, and Eckart Zitzler for comments on an earlier version of this paper. This work was partially supported by a postdoctoral fellowship from the German Academic Exchange Service (DAAD).

References

[1] Agarwal, P.K., Kaplan, H., Sharir, M.: Computing the volume of the union of cubes. In: Proc. 23rd annual Symposium on Computational Geometry (SoCG 2007), pp. 294–301 (2007)

[2] Bárány, I., Füredi, Z.: Computing the volume is difficult. Discrete & Computational Geometry 2, 319–326 (1986); Announced at STOC 1986

[3] Bringmann, K., Friedrich, T.: Approximating the volume of unions and intersections of high-dimensional geometric objects (2008), http://arxiv.org/abs/0809.0835

[4] Chan, T.M.: Semi-online maintenance of geometric optima and measures. SIAM J. Comput. 32, 700–716 (2003)

[5] Chan, T.M.: A (slightly) faster algorithm for Klee's measure problem. In: Proc. 24th ACM Symposium on Computational Geometry (SoCG 2008), pp. 94–100 (2008)

[6] Dyer, M.E., Frieze, A.M.: On the complexity of computing the volume of a polyhedron. SIAM J. Comput. 17, 967–974 (1988)
[7] Dyer, M.E., Frieze, A.M., Kannan, R.: A random polynomial time algorithm for approximating the volume of convex bodies. J. ACM 38, 1–17 (1991); Announced at STOC 1989
[8] Fredman, M.L., Weide, B.W.: On the complexity of computing the measure of $\bigcup[a_i, b_i]$. Commun. ACM 21, 540–544 (1978)
[9] Halman, N., Klabjan, D., Li, C.-L., Orlin, J.B., Simchi-Levi, D.: Fully polynomial time approximation schemes for stochastic dynamic programs. In: Proc. 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2008), pp. 700–709 (2008)
[10] Kannan, R., Lovász, L., Simonovits, M.: Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. Random Struct. Algorithms 11, 1–50 (1997)
[11] Kaplan, H., Rubin, N., Sharir, M., Verbin, E.: Counting colors in boxes. In: Proc. 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2007), pp. 785–794 (2007)
[12] Karp, R.M., Luby, M.: Monte-carlo algorithms for the planar multiterminal network reliability problem. J. Complexity 1, 45–64 (1985)
[13] Karp, R.M., Luby, M., Madras, N.: Monte-carlo approximation algorithms for enumeration problems. J. Algorithms 10, 429–448 (1989)
[14] Khachiyan, L.G.: The problem of calculating the volume of a polyhedron is enumerably hard. Russian Mathematical Surveys 44, 199–200 (1989)
[15] Klee, V.: Can the measure of $\bigcup[a_i, b_i]$ be computed in less than $O(n \log n)$ steps? American Mathematical Monthly 84, 284–285 (1977)
[16] Lovász, L., Vempala, S.: Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. J. Comput. Syst. Sci. 72, 392–417 (2006)
[17] Overmars, M.H., Yap, C.-K.: New upper bounds in Klee's measure problem. SIAM J. Comput. 20, 1034–1045 (1991); Announced at FOCS 1988
[18] Roth, D.: On the hardness of approximate reasoning. Artif. Intell. 82, 273–302 (1996)
[19] Suzuki, S., Ibaraki, T.: An average running time analysis of a backtracking algorithm to calculate the measure of the union of hyperrectangles in d dimensions. In: Proc. 16th Canadian Conference on Computational Geometry (CCCG 2004), pp. 196–199 (2004)
[20] van Leeuwen, J., Wood, D.: The measure problem for rectangular ranges in d-space. J. Algorithms 2, 282–300 (1981)
[21] Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans. Evolutionary Computation 3, 257–271 (1999)