Random Sampling with Removal *

Bernd Gärtner[1], Johannes Lengler[1], and May Szedlák[1]

1 **Department of Computer Science**
Institute of Theoretical Computer Science, ETH Zürich
CH-8092 Zürich, Switzerland
`{gaertner, johannes.lengler, may.szedlak}@inf.ethz.ch`

Abstract

Random sampling is a classical tool in constrained optimization. Under favorable conditions, the optimal solution subject to a small subset of randomly chosen constraints violates only a small subset of the remaining constraints. Here we study the following variant that we call random sampling with removal: suppose that after sampling the subset, we remove a fixed number of constraints from the sample, according to an arbitrary rule. Is it still true that the optimal solution of the reduced sample violates only a small subset of the constraints?

The question naturally comes up in situations where the solution subject to the sampled constraints is used as an approximate solution to the original problem. In this case, it makes sense to improve cost and volatility of the sample solution by removing some of the constraints that appear most restricting. At the same time, the approximation quality (measured in terms of violated constraints) should remain high.

We study random sampling with removal in a generalized, completely abstract setting where we assign to each subset R of the constraints an arbitrary set $V(R)$ of constraints disjoint from R; in applications, $V(R)$ corresponds to the constraints violated by the optimal solution subject to only the constraints in R. Furthermore, our results are parametrized by the dimension δ, i.e., we assume that every set R has a subset B of size at most δ with the same set of violated constraints. This is the first time this generalized setting is studied.

In this setting, we prove matching upper and lower bounds for the expected number of constraints violated by a random sample, after the removal of k elements. For a large range of values of k, the new upper bounds improve the previously best bounds for LP-type problems, which moreover had only been known in special cases. We show that this bound on special LP-type problems, can be derived in the much more general setting of violator spaces, and with very elementary proofs.

1998 ACM Subject Classification F.2.2

Keywords and phrases LP-type problem, violator space, random sampling, sampling with removal

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

On a high level, random sampling can be described as an efficient way of learning something about a problem, by first solving a subproblem of much smaller size. A classical example is the problem of finding the smallest element in a *sorted compact list* [4, Problem 11-3]. Such a list stores its elements in an array, but in arbitrary order. Additional pointers are used to link each element to the next smaller one in the list. Given a sorted compact list of size n,

* Research supported by the Swiss National Science Foundation (SNF Project 200021_150055 / 1)

the smallest element can be found in expected time time $O(\sqrt{n})$ as follows: sample a set of $\lfloor\sqrt{n}\rfloor$ array elements at random. Starting from their minimum, follow the predecessor pointers to the global minimum. The key fact is that the expected number of pointers to be followed is bounded by $\sqrt{n}$, and this yields the expected runtime.

On an abstract level, the situation can be modeled as follows. Let H be a set of size n that we can think of as the set of constraints in an optimization problem, for example the elements in a sorted compact list. Let $V : 2^H \to 2^H$ be a function that assigns to each subset $R \subseteq H$ of constraints a set $V(R) \subseteq H \setminus R$. We can think of $V(R)$ as the set of constraints violated by the optimal solution subject to only the constraints in R. This abstract setting (H, V), denoted by *consistent space*, is introduced here for the first time (for a formal definition see Definition 2.3 below.) In the sorted compact list example, $V(R)$ is the set of elements that are smaller than the minimum of R.

In this setting, the above "key fact" is a concrete answer to the following abstract question: Suppose that we sample a set $R \subseteq H$ of size $r \leq n$ uniformly at random. What can we say about the quantity v_r, the expected size of $V(R)$? What are conditions on V under which v_r is small?

The main workhorse in this context is the *Sampling Lemma* [8]. It states that $v_r = \frac{n-r}{r+1} \cdot x_{r+1}$, where x_r is the expected size of $X(R) = \{h \in R : h \in V(R \setminus \{h\})\}$. In other words, $h \in X(R)$ is a constraint that is not automatically satisfied if the problem is solved without enforcing it. In the sorted compact list example, every nonempty set R has one such "extreme" constraint, namely its minimum. Consequently, we have $x_{r+1} = 1$, and hence $v_r = (n-r)/(r+1)$. With $r = \lfloor\sqrt{n}\rfloor$, $v_r < \sqrt{n}$ follows.

The Sampling Lemma has many other applications in computational geometry when x_{r+1} can be bounded; in a number of relevant cases, we do not only know the expected value v_r but the complete probability distribution $p_\ell = \Pr[|V(R)| = \ell], \ell \leq n$, or tail estimates for $|V(R)|$ [8].

In this paper, we address the following more general question in the abstract setting: Suppose that we sample a set $R \subseteq H$ of size $r \leq n$ uniformly at random, but then we remove a subset $K_R \subseteq R$ of a fixed size k, according to an arbitrary but fixed rule. What can we still say about the expected size of $V(R \setminus K_R)$? If K_R is a random subset of R, the expectation is v_{r-k}, but if K_R is chosen by another (deterministic) rule, then $R \setminus K_R$ is no longer a uniformly random subset, and the Sampling Lemma does not apply.

Our work is originally motivated by *chance-constrained optimization.* In this setting, we have a probability distribution over a (possibly infinite) set of constraints. The goal is to compute a *sufficiently feasible* solution, one that satisfies a randomly chosen constraint with high probability. Such a solution can be obtained by optimizing over a finite sample of constraints drawn from the distribution [2]. Here, a-posteriori removal of constraints is shown to yield a tradeoff between solution quality and violation probability [3]. We are trying to understand the combinatorial essence of this tradeoff.

Intuitively, one would think that if k is constant, the change in the expected number of violated constraints under the removal of k constraints is small. This intuition was proved to be correct if the pair (H, V) is induced by a *nondegenerate LP-type problem*, where the results are parametrized by the dimension δ [5], (for definition of dimension see Definition 4 below). LP-type problems have been introduced and analyzed by Matoušek, Sharir and Welzl as a combinatorial framework that encompasses linear programming and other geometric optimization problems [11, 9]. The quantitative result was that under removal of k elements, the expected number of violated constraints increases by a factor of δ^k at most, which is constant if both δ and k are constant. It was left open whether this factor can be improved for interesting sample sizes (for very specific and rather irrelevant values of δ, r, k, it was shown to be best possible).

In this paper, we improve over the results in [5] in several respects. In Section 3, Theorem 10 we show that the increase factor δ^k can be replaced by $\log n + k$, which is a vast improvement for a large range of values of k. Moreover, the new bound neither requires the machinery of LP-type problems, nor nondegeneracy. It holds in the completely abstract setting of consistent spaces considered above. In this setting, we can also show that the bound is best possible for all sample sizes of the form $r = n^\alpha, 0 < \alpha < 1$ (see Section 5). We also show that this bound is best possible for violator spaces, in the case where $k = \Omega(\delta \log n)$. In general, for violator spaces the gap to the lower bound is $\log n$.

Hence, if anything can be gained over the new bound, additional properties of the violator function V have to be used. Indeed, for small values of k, the increase factor in [5] is better than our new bound for nondegenerate LP-type problems, and most notably, it does not depend on the problem size n. We show in Section 4, Theorem 14 that the same factor can be derived under the much weaker conditions of a *nondegenerate violator space*, and with a much simpler proof, based on a “removal version” of the Sampling Lemma. Furthermore the proof of [5] is given for a specific rule to remove k, whereas our proof works for any rule.

Intuitively, violator spaces are LP-type problems without objective function, and they were introduced to show that many combinatorial properties of LP-type problems and algorithms for LP-type problems do not require the objective function at all [6, 1].

In Section 6, Theorem 18 we show tight upper and lower bounds for the case $\delta = 1$, which shows that the improved bound for nondegenerate violater spaces is best possible for *all* violator spaces. For smaller (and in particular constant) k, the quest for the best bound on the increase factor remains open. In particular, it is not clear whether the exponential growth in k actually happens.

What also remains open is the role of nondegeneracy. In many geometric situations, nondegeneracy can be attained through symbolic perturbation and can therefore be assumed without loss of generality for most purposes. In the abstract setting, this is not necessarily true, as there are examples of LP-type problems for which any “combinatorial perturbation” increases the dimension [10].

2 Basics and Definitions

Throughout the paper we will work with three combinatorial concepts, the LP-type problem, the violator space and the consistent space. The LP-type problem was first introduced by Sharir and Welzl [11], the generalized concept of violator spaces by Gärtner, Matoušek, Rüst, and Škovroň [6]. In this paper we introduce an even more general concept of consistent spaces.

2.1 LP-type Problems

▶ **Definition 1.** An *LP-type* problem is a triple $\mathcal{P} = (H, \Omega, \omega)$ that satisfies the following. H is a finite set (the constraints), Ω a totally ordered set with a smallest element $-\infty$ and $\omega : 2^H \rightarrow \Omega$ a function that assigns an objective function value to $G \subseteq H$, such that $\omega(\emptyset) = -\infty$. For all $F \subseteq G \subseteq H$ and $h \in H$, the following hold.

1. $\omega(F) \leq \omega(G)$, and
2. If $\omega(F) = \omega(G) > -\infty$, then $\omega(G \cup \{h\}) > \omega(G) \Rightarrow \omega(F \cup \{h\}) > \omega(F)$.

The first condition is called *monotonicity*, the second *locality*.

Observe that using locality, by simple induction one can show that if $\omega(F \cup \{h\}) = \omega(F) > -\infty$ for all $h \in G \setminus F$, then $\omega(F) = \omega(G)$.

The classic example of an LP-type problem is the problem of computing the smallest enclosing ball of a finite set of points P in $\mathbb{R}^d$. Let us denote this problem by SEB. We can write this as an LP-type problem by setting $H = P$ and $\Omega = \mathbb{R} \cup \{-\infty\}$. For $G \subseteq H$, $\omega(G)$ is defined as the radius of the smallest enclosing ball of G, with the convention that the radius of the empty set is $-\infty$. Since the smallest enclosing ball of a nonempty set of points exists and is unique [13], ω is well defined.

Monotonicity is clear. To see locality, observe that for $F \subseteq G$, $\omega(F) = \omega(G)$ means that both F and G have the same smallest enclosing ball. If $\omega(G \cup \{h\}) > \omega(G)$, then h is outside this ball, so $\omega(F \cup \{h\}) > \omega(F)$.

▶ **Definition 2.** A constraint $h \in H \setminus G$ is violated by G if $\omega(G \cup \{h\}) > \omega(G)$. We denote the set of violated constraints by $V(G)$.

For SEB, the violated constraints of G are exactly the points lying outside the smallest enclosing ball of G.

2.2 Violator Spaces

Intuitively a violator space is an LP-type problem without an objective function. The advantage is that many things one can prove about LP-type problems do not require the concept of order.

▶ **Definition 3.** A *violator space* is a pair (H, V), $|H| = n$ finite and V a function $2^H \to 2^H$ such that the following is satisfied for all $F \subseteq G \subseteq H$.

1. $G \cap V(G) = \emptyset$.
2. If $G \cap V(F) = \emptyset$, then $V(G) = V(F)$.

The first condition is called *consistency*, the second *locality*.

Observe that the locality condition implies that if $E \subseteq F \subseteq G$ and $V(E) = V(G)$, then $V(E) = V(F) = V(G)$.

The notion of a violator space is more general than the LP-type problem, since every LP-type problem (H, Ω, ω) can naturally be converted into a violator space through $V(R) = \{x \in H \setminus R \mid \omega(R \cup \{x\}) \neq \omega(R)\}$, for $R \subseteq H$. On the other hand, not every violator space can be converted into an LP-type problem. Any unique sink orientation (USO) [12] of a cube or the grid [7] corresponds to a violator space, but not to an LP-type problem in general [6].

▶ **Definition 4.** Let (H, V) be a violator space.

1. $B \subseteq H$ is called a *basis* in (H, V), if for all $F \subsetneq B$, $B \cap V(F) \neq \emptyset$ (or equivalently, $V(F) \neq V(B)$).
2. A basis of $G \subseteq H$ is an inclusion-minimal subset $B \subseteq G$ such that $V(B) = V(G)$. (In particular, a basis of G is a basis in (H, V), and every basis in (H, V) is a basis of itself.)
3. The *combinatorial dimension* of (H, V), denoted $\delta := \delta(H, V)$ is defined by the size of the largest basis in (H, V).

Again, let us illustrate this on SEB. A basis of G is a minimal subset of points with the same enclosing ball of G. In particular all points of the basis are on the ball's boundary. In d-dimensional space, the combinatorial dimension of any SEB-instance is at most $d + 1$, since any enclosing ball can be defined by at most $d + 1$ points on its boundary. However, a basis can be smaller than the combinatorial dimension, and a point set can have more than one basis: in $\mathbb{R}^2$ the set of four corners of a square has two bases, the two pairs of diagonally opposite points.

▶ **Definition 5.** The set of *extreme constraints* $X(G) \subseteq G$ is defined by

$$r \in X(G) \Leftrightarrow r \in V(G \setminus \{r\}).$$

In the SEB case, h is extreme in G if its removal allows for a smaller enclosing ball. Therefore h is necessarily on the boundary of the smallest enclosing ball, but this is not sufficient. For the case $\mathbb{R}^2$, if G consists of the four corners of a square, then G has no extreme point.

It is not hard to see that $X(G)$ is the intersection of all bases of G, hence $|X(G)| \leq \delta$. To bound the expected number of violators, the following result from [8] is known.

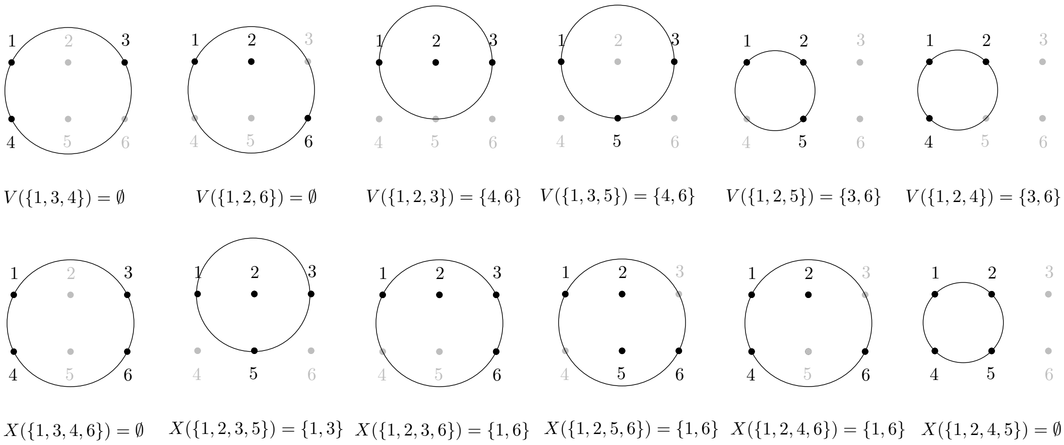
▶ **Lemma 6.** *[Sampling Lemma] Let (H, V) be a violator space with combinatorial dimension δ. Let $R \subseteq H$ a u.a.r. set of size r, $v_r = \mathbb{E}[|V(R)|]$ and $x_r = \mathbb{E}[|X(R)|]$. Then*

$$v_r = \frac{n-r}{r+1} \cdot x_{r+1} \leq \frac{n-r}{r+1} \cdot \delta.$$

The Sampling Lemma can be used to argue that v_r is small if the expected number x_{r+1} of extreme constraints of a random sample of size $r+1$ is small.

In SEB case in $\mathbb{R}^d$, one can use Helly's theorem to show that every set has at most $d+1$ extreme points and therefore $v_r \leq \frac{n-r}{r+1} \cdot (d+1)$. If $d = 2$, then the smallest enclosing ball of a random sample of size $\sqrt{n}$ has in expectation at most $3\sqrt{n}$ points outside.

The figure below shows an example of SEB, on the 2×3 regular grid, in particular $n = 6$. We fix $r = 3$ and therefore look at the violators of sets of size three and extreme points of sets of size four, which are all depicted (up to symmetry). It is not hard to see that eight sets of size three have no violators (corresponding to the first two cases in the figure), while all others have two violators. This implies $v_3 = 1.2$. Looking at the extreme points the sets $\{1,3,4,6\}, \{1,2,4,5\}$ and $\{2,3,5,6\}$ have no extreme elements and all other sets of size four have exactly two extreme elements. Therefore $v_3 = \frac{6-3}{3+1} x_4 = \frac{3}{4} \cdot 1.6 = 1.2$. For six points in general position $v_3 = \frac{3}{4} \cdot 3 = 2.25$.



$V(\{1,3,4\}) = \emptyset$ $V(\{1,2,6\}) = \emptyset$ $V(\{1,2,3\}) = \{4,6\}$ $V(\{1,3,5\}) = \{4,6\}$ $V(\{1,2,5\}) = \{3,6\}$ $V(\{1,2,4\}) = \{3,6\}$

$X(\{1,3,4,6\}) = \emptyset$ $X(\{1,2,3,5\}) = \{1,3\}$ $X(\{1,2,3,6\}) = \{1,6\}$ $X(\{1,2,5,6\}) = \{1,6\}$ $X(\{1,2,4,6\}) = \{1,6\}$ $X(\{1,2,4,5\}) = \emptyset$

▶ **Definition 7.** A violator space (H, V) is called *nondegenerate* if every set $G \subseteq H$ has a unique basis.

Note that SEB it not nondegenerate, since as mentioned in $\mathbb{R}^2$, four points on a square have two bases. For a nondegenerate example see the d-smallest number violator space (Definition 16 below).

It is shown in [10] that in abstract settings one cannot without loss of generality assume nondegeneracy. This is shown by constructing LP-type problems of dimension δ for infinitely many integers δ such that in oder to remove degeneracies, one has to increase the dimension to at least 2δ.

2.3 Consistent Spaces

We now move on to the even more general concept of consistent spaces.

▶ **Definition 8.** A *consistent space* is a pair (H, V), $|H| = n$ finite and V a function $2^H \to 2^H$ such that the following is satisfied for all $G \subseteq H$.

1. $G \cap V(G) = \emptyset$.

Hence a consistent space is a violator space without the locality condition. The basis, combinatorial dimension and extreme constraints of a consistent space can be defined equivalently as in the violator space.

In consistent spaces the first equality $v_r = \frac{n-r}{r+1} \cdot x_{r+1}$ of the Sampling Lemma 6 still holds. However, in general it does not hold that $|X(R)| \leq \delta$ for all $R \subseteq H$. We give an example of a consistent space of dimension 1 such that for some $R \subseteq H$ each element is extreme. Let $R = \{1, 2, \dots, 2m\} \subseteq n$ be of even size. For $i \in [m]$ let $V(i) = \{i + m\}$ and $V(i + m) = \{i\}$. For every $x \in R$ define $V(R \setminus \{x\}) = \{x\}$ and for all other sets define the violators as the empty set. Then this is a consistent space. By definition it also follows that every element in R is extreme. For $x \in R$ it holds that $V(R \setminus \{x\}) = \{x\} = V(x + m \mod 2m)$ and since $x + m \mod 2m \in R$ it follows that the combinatorial dimension is 1.

2.4 Sampling with Removal

As already introduced in [5] for LP-type problems, we are interested in sampling with removal. We define the concept here for the most general case of consistent spaces. All results will then naturally extend for violator spaces and LP-type problems. Suppose we sample uniformly at random $R \subseteq H$ of size r. By some fixed rule P_k, we remove $k < r$ elements of R and obtain a set R_{P_k} of size $r - k$. We define $V_{P_k}(R) := V(R_{P_k})$. Note that in general (H, V_{P_k}) is not a consistent space. We are interested in $\mathbb{E}[|V_{P_k}(R)|]$, for which we will give different bounds in the next two chapters.

Before proceeding to the bounds, we discuss some possible rules for the removal of the k elements. If k elements are removed uniformly at random from R, then $\mathbb{E}[|V_{P_k}(R)|] = v_{r-k}$. Another way to remove k elements is to maximize the number of violators after the removal. In the case of LP-type problems it is intuitive to remove in such a way that the objective function is minimized [5]. For this last rule [5] establishes a bound of $\mathbb{E}[|V_{P_k}(R)|] = O(\frac{n-r}{r+1}\delta^{k+1})$ for fixed k, if the LP-type problem is nondegenerate.

3 An Upper Bound for Consistent Spaces

The main result of this section is Theorem 10, where we show an upper bound on $\mathbb{E}[|V_{P_k}(R)|]$ for consistent spaces. In Lemma 15 and Lemma 17 we show asymptotically matching lower bounds for most relevant values of r.

We start with the following technical lemma.

▶ **Lemma 9.** *Let (H, V), with $|H| = n$, a consistent space of dimension δ and P_k some fixed rule to remove k points. Let R be u.a.r. from all sets of size r, c some suitable constant (e.g. $c = 33$), and $x = c \cdot \max\left\{\frac{n}{r}\delta \log n, \frac{n}{r}k\right\}$. We assume $r = o(n)$, $\delta = o(r)$, $k \leq \frac{r}{c}$, and $r + x \leq n$. Then*

1.

$$Pr[|V_{P_k}(R)| \geq x] \leq \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} \frac{\binom{n}{\alpha}}{\binom{n}{r}} \binom{x}{i} \binom{n-x-\alpha}{r-\alpha-i}.$$

2. *Furthermore for all $0 \leq \alpha \leq \delta$ and $0 \leq i \leq k$,*

$$\frac{\binom{n}{\alpha}}{\binom{n}{r}} \binom{x}{i} \binom{n-x-\alpha}{r-\alpha-i} \leq n^{-3}.$$

Proof of Lemma 9. For $R \subseteq H$ define $\mathcal{B}_{P_k}(R) := \{B \subseteq H \mid B \text{ basis of } R_{P_k}\} \neq \emptyset$, the set of all bases of R_{P_k}. Let $R \subseteq H$ uniformly at random of size r. Then

$$\begin{aligned}\Pr[|V_{P_k}(R)| \geq x] &= \Pr[\exists B \in \mathcal{B}_{P_k}(R) : |V(B)| \geq x] \\ &\leq \Pr[\exists B \subseteq H : |V(B)| \geq x \wedge |B| \leq \delta \wedge B \subseteq R \wedge V(B) = V_{P_k}(R)] \\ &\leq \Pr[\exists B \subseteq H : |V(B)| \geq x \wedge |B| \leq \delta \wedge B \subseteq R \wedge |V(B) \cap R| \leq k].\end{aligned}$$

The last inequality follows since by $V_{P_k}(R) = V(B)$ it follows that $R_{P_k} \cap V(B) = \emptyset$ and because R_{P_k} is obtained from R by removing k elements $|R \cap V(B)| \leq k$. Using union bound we obtain

$$\begin{aligned}\Pr[|V_{P_k}(R)| \geq x] &\leq \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} \Pr[\exists B \subseteq H : |V(B)| \geq x \wedge |B| = \alpha \wedge B \subseteq R \wedge |V(B) \cap R| = i] \\ &\leq \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} \sum_{\substack{B \in \binom{n}{\alpha} \\ |V(B)| \geq x}} \Pr[B \subseteq R \wedge |V(B) \cap R| = i] \\ &= \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} \sum_{\substack{B \in \binom{n}{\alpha} \\ |V(B)| \geq x}} \frac{1}{\binom{n}{r}} \underbrace{\binom{|V(B)|}{i} \binom{n-|V(B)|-\alpha}{r-\alpha-i}}_{(*)}.\end{aligned}$$

We now show that $(*)$ is maximized if $|V(B)| = x$, which concludes the proof. For $i = 0$ this is obvious, hence assume $i \geq 1$.

The claim follows since by basic calculations we can show that for all $y \geq \frac{i(n-\alpha)-r+i+\alpha}{r-\alpha}$,

$$\binom{y}{i} \binom{n-y-\alpha}{r-\alpha-i} \geq \binom{y+1}{i} \binom{n-(y+1)-\alpha}{r-\alpha-i}.$$

Using that $\alpha = o(r)$ and $r, i, \alpha = o(n)$ we get

$$x > (1+o(1))\frac{n}{r}k \geq (1+o(1))\frac{n}{r}i = \frac{i(n-\alpha)-r+i+\alpha}{r-\alpha},$$

and the first part follows.

For the second part again the case of $i = 0$ is easy, hence assume that $i \geq 1$.

$$\frac{\binom{n}{\alpha}}{\binom{n}{r}} \underbrace{\binom{x}{i}}_{\leq \left(\frac{xe}{i}\right)^i} \binom{n-x-\alpha}{r-\alpha-i} \leq \left(\frac{xe}{i}\right)^i \cdot \frac{n!}{\underbrace{\alpha!}_{\geq 1}(n-\alpha)!} \cdot \frac{r!(n-r)!}{n!} \cdot \frac{(n-x-\alpha)!}{(r-\alpha-i)!(n-x-r+i)!}$$

$$\leq \left(\frac{xe}{i}\right)^i \cdot \underbrace{\frac{r\cdots(r-\alpha-i+1)}{(n-\alpha)\cdots(n-\alpha-i+1)}}_{\leq r^\alpha \left(\frac{r}{n}\right)^i} \cdot \frac{(n-r)!}{(n-\alpha-i)!} \cdot \frac{(n-x-\alpha)!}{(n-x-r+i)!}$$

$$\leq \left(\frac{xe}{i}\right)^i \cdot r^\alpha \left(\frac{r}{n}\right)^i \underbrace{\frac{(n-x-\alpha)\cdots(n-x-r+i+1)}{(n-\alpha-i)\cdots(n-r+1)}}_{\leq \left(1-(1+o(1))\frac{x}{n}\right)^{r-\alpha-i} \leq e^{-\frac{x}{2n}r}}$$

$$\leq \left(\frac{xe}{i}\right)^i \cdot r^\alpha \left(\frac{r}{n}\right)^i e^{-\frac{x}{2n}r}.$$

Now in the first case we assume that $k \leq \delta \log n$, hence $x = c\frac{n}{r}\delta \log n$. It follows that

$$\frac{\binom{n}{\alpha}}{\binom{n}{r}} \binom{x}{i} \binom{n-x-\alpha}{r-\alpha-i} \leq \left(\frac{c\frac{n}{r}\delta \log ne}{i}\right)^i \cdot r^\alpha \left(\frac{r}{n}\right)^i e^{-\frac{c}{2}\delta \log n} \leq \left(\frac{c\delta \log ne}{i}\right)^i \cdot r^\alpha e^{-\frac{c}{2}\delta \log n}$$

$$= e^{\overbrace{i \log c}^{\leq \delta \log n \log c} + i \log \delta + i \log \log n + \overbrace{i}^{\leq \delta \log n} - i \log i + \overbrace{\alpha \log r}^{\leq \delta \log n} - \frac{c}{2}\delta \log n}$$

$$\leq e^{i \log \delta + i \log \log n - i \log i - \left(\frac{c}{2} - 2 - \log c\right)\delta \log n}.$$

To show the claim it remains to show that

$$i \log \delta + i \log \log n - i \log i - \left(\frac{c}{2} - 2 - \log c\right) \delta \log n \leq -3 \log n.$$

Since $i \leq k \leq \delta \log n$ we can write $i = \beta\delta \log n$ for some $\beta \in (0, 1]$. Then

$$\begin{aligned} &i \log \delta + i \log \log n - i \log i - \left(\frac{c}{2} - 2 - \log c\right) \delta \log n \\ &\quad = \beta\delta \log n \log \delta + \beta\delta \log n \log \log n - \beta\delta \log n(\log \beta + \log \delta + \log \log n) - \left(\frac{c}{2} - 2 - \log c\right) \delta \log n \\ &\quad = -\beta \log \beta\delta \log n - \left(\frac{c}{2} - 2 - \log c\right) \delta \log n. \end{aligned}$$

It remains to bound $\beta \log \beta$ from below. By taking the derivative we observe that $\beta \log \beta$ attains its minimum when $\beta = \frac{1}{e}$ and hence $\beta \log \beta \geq -\frac{1}{e}$. It therefore follows that

$$\begin{aligned} &i \log \delta + i \log \log n - i \log i - \left(\frac{c}{2} - 2 - \log c\right) \delta \log n \\ &\quad \leq -\left(\frac{c}{2} - 2 - \log c - \frac{1}{e}\right) \delta \log n \leq -3 \log n \text{ for } c \geq 33. \end{aligned}$$

In the second case $k \geq \delta \log n$, hence $x = c \cdot \frac{n}{r} \cdot k$. Again for $i \geq 1$ it follows that

$$\frac{\binom{n}{\alpha}}{\binom{n}{r}} \binom{x}{i} \binom{n-x-\alpha}{r-\alpha-i} \leq \left(\frac{xe}{i}\right)^i \cdot r^\alpha \left(\frac{r}{n}\right)^i e^{-\frac{x}{2n}r} = \left(c\frac{n}{r}k\frac{e}{i}\right)^i r^\alpha \left(\frac{r}{n}\right)^i e^{-\frac{c}{2}k}$$

$$= e^{\overbrace{i \log c}^{\leq k \log c} + i \log k + \overbrace{i}^{\leq k} - i \log i + \overbrace{\alpha \log r}^{\leq \delta \log n \leq k} - \frac{c}{2}k}$$

$$\leq e^{i(\log k - \log i) - (\frac{c}{2} - 2 - \log c)k}$$

Now as before it suffices to show that

$$i(\log k - \log i) - \left(\frac{c}{2} - 2 - \log c\right) k \leq -3 \log n.$$

Since $i \leq k$ we can write $i = \beta k$ for some $\beta \in (0, 1]$. Then

$$\begin{aligned} i(\log k - \log i) &- \left(\frac{c}{2} - 2 - \log c\right) k \\ &= \beta k(\log k - \log k - \log \beta) - \left(\frac{c}{2} - 2 - \log c\right) k \\ &\leq - \left(\frac{c}{2} - 2 - \log c - \frac{1}{e}\right) \underbrace{k}_{\geq \delta \log n} \leq -3 \log n \text{ for } c \geq 33, \end{aligned}$$

where we again used that $\beta \log \beta \geq -\frac{1}{e}$. ◀

▶ **Theorem 10.** *Let (H, V), with $|H| = n$, a* consistent space *of dimension δ and P_k some fixed rule to remove k points. Let $R \subseteq H$ u.a.r. of all sets of size r, for some $r \leq n$. Then*

$$\mathbb{E}[|V_{P_k}(R)|] \leq c \cdot \max\left\{\frac{n}{r} \delta \log n, \frac{n}{r} k\right\} =: x,$$

where c is some suitable constant (e.g. $c = 33$).

Observe that compared to Lemma 6, for most relevant r (e.g. $r = n^{\beta}$, $\beta \in (0, 1)$) and $k = o(\delta \log n)$, there is an additional $\log n$ term.

Proof of Theorem 10. We may assume $r = o(n)$, $\delta = o(r)$, $k \leq \frac{r}{c}$, and $r + x \leq n$, since otherwise the bound is trivial and there is nothing to prove. By definition of expectation

$$\begin{aligned} \mathbb{E}[|V_{P_k}(R)|] &\leq \Pr[|V_{P_k}(R)| < x] \cdot (x - 1) + \Pr[|V_{P_k}(R)| \geq x] \cdot n \\ &\leq x - 1 + \Pr[|V_{P_k}(R)| \geq x] \cdot n. \end{aligned}$$

We will now show that $\Pr[|V_{P_k}(R)| \geq x] \leq n^{-1}$ which will conclude the proof. By Lemma 9,

$$\Pr[|V_{P_k}(R)| \geq x] \leq \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} \frac{\binom{n}{\alpha}}{\binom{n}{r}} \binom{x}{i} \binom{n - x - \alpha}{r - \alpha - i} \leq \sum_{i=0}^{k} \sum_{\alpha=0}^{\delta} n^{-3} \leq n^{-1}.$$

as desired. ◀

4 An Upper Bound for Violator Spaces

In this section we give an upper bound on $\mathbb{E}[|V_{P_k}(R)|]$ for nondegenerate violator spaces, Theorem 14. This is an improvement of the bound given in [5], which stated the same bound for nondegenerate LP-type problems and the specific rule P_k to minimize the objective function after removal. Matching lower bounds for special cases are known [5]. The bound in Theorem 14 is stronger than the (more general) bound in Theorem 10 if δ and k are very small, e.g., if $\delta^k < \log n$; for large δ and k, Theorem 10 is stronger.

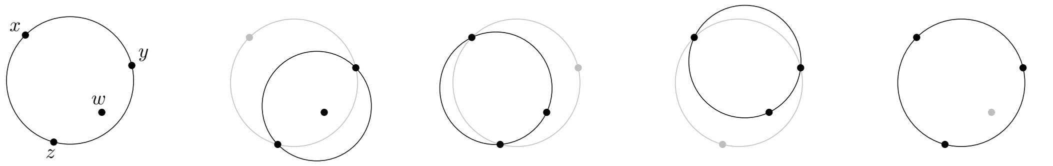
4.1 Extreme Constraints after Removal

Let (H, V) be a violator space of combinatorial dimension δ. In particular, every set has at most δ extreme constraints. For a given natural number k, we want to understand the following quantity:

$$\Delta_k(H, V) := \max_{R \subseteq H} |\{X(R \setminus K) : K \subseteq R, |K| = k\}|.$$

In other words, how many sets of extreme constraints can we get by removing k elements from some set R?

The figure below shows an instance of SEB with $R = \{x, y, z, w\}$, where by removing one point from R we can get four different sets of extreme constraints ($\{x, y, z\}, \{y, z\}, \{x, z, w\}$ and $\{x, y, w\}$. Therefore we conclude that $\Delta_1(H, V) \geq 4$. We will see below that for nondegenerate violator spaces $\Delta_1(H, V) \leq \delta + 1$, so this bound is actually tight.



$X(R) = \{x, y, z\}$ $X(R \setminus \{x\}) = \{y, z\}$ $X(R \setminus \{y\}) = \{x, z, w\}$ $X(R \setminus \{z\}) = \{x, y, w\}$ $X(R \setminus \{w\}) = X(R)$

▶ **Lemma 11.** *Let (H, V) be a nondegenerate violator space. For $R \subseteq H$ let B_R be its unique basis. Then $X(R) = B_R$. Furthermore for $x \in R \setminus X(R)$ we have that $X(R) = X(R \setminus \{x\})$.*

Proof. Let $x \in R \setminus X(R)$. Then by definition $x \notin V(R \setminus \{x\})$ and therefore by locality $V(R) = V(R \setminus \{x\})$. Now $V(R \setminus \{x\}) = V(B_{R \setminus \{x\}})$. By nondegeneracy, $B_R = B_{R \setminus \{x\}}$ and therefore $x \notin B_R$. For the other direction assume $x \in X(R)$. Then $x \in V(R \setminus \{x\})$. If $x \notin B_R$, then by locality $V(B_R) = V(R) = V(R \setminus \{x\})$, which is a contradiction.

The second part follows since by nondegeneracy $X(R) = B_R = B_{R \setminus \{x\}} = X(R \setminus \{x\})$. ◀

Let's bound the easy cases of $\Delta_k(H, V)$ first. We obviously have $\Delta_0(H, V) = 1$ for any violator space (H, V). Moreover for (H, V) nondegenerate we have, $\Delta_1(H, V) \leq \delta + 1$. Indeed, if we remove a non-extreme element x from R, by Lemma 11 we end up with the same set $X(R \setminus \{x\}) = X(R)$ of extreme elements, so only in at most δ cases, we will get a different set. Note that in general this bound does not hold. Consider SEB and assume we have four points on a square and one point in the middle. It is not hard to see that for each point its removal generates a different set of extreme points. Hence $\Delta_1(H, V) \geq 5 > \delta + 1 = 4$.

▶ **Lemma 12.** *Let (H, V) be a nondegenerate violator space. Then $\Delta_k(H, V) \leq \sum_{i=0}^{k} \delta^i$.*

Proof. Let us fix R and a set $K \subseteq R, |K| = k$ to be removed. We claim that we can order the elements of K as $e_1, \ldots, e_k$ such that for some $\ell \in \{0, \ldots, k\}$,

$$e_i \in X(R \setminus \{e_1, \ldots, e_{i-1}\}), \quad i \leq \ell, \text{ and } X(R \setminus K) = X(R \setminus \{e_1, \ldots, e_\ell\}).$$

Indeed, we can do this greedily: as long as we can remove an extreme element from the current set, we do so. At some points, all elements that remain to be removed are non-extreme, and by repeated use of Lemma 11 at this point the removal of all of them does not change the extreme elements anymore.

It follows that all sets $X(R \setminus K)$ can be obtained from R by repeatedly removing an extreme element from the current set, up to k times. In the first round, we therefore have at most δ choices, and for each of them, we have at most δ choices in the second round, and so on. The bound follows. ◀

4.2 Sampling Lemma after Removal

Let (H, V) be a violator space. For $R \subseteq H$ and a natural number k, we define the following two quantities.

$$\begin{aligned} V_k(R) &= \{x \in H \setminus R : x \in V(R \setminus K) \text{ for some } K \subseteq R, |K| = k\}, \\ X_k(R) &= \{x \in R : x \in X(R \setminus K) \text{ for some } K \subseteq R, |K| = k\}. \end{aligned}$$

Clearly, $V(R) = V_0(R)$ and $X(R) = X_0(R)$.

Furthermore, we let $v_{r,k}$ denote the expected size of $V_k(R)$ over a randomly chosen set of size r. Similarly, $x_{r,k}$ is the expected size of $X_k(R)$.

▶ **Lemma 13.** *[Sampling Lemma after Removal]*

$$v_{r,k} = \frac{n-r}{r+1} x_{r+1,k}.$$

Proof. This goes like for the "normal" Sampling Lemma 6 [8]. We define a bipartite graph on the vertex set $\binom{H}{r} \cup \binom{H}{r+1}$, where we connect R and $R \cup \{x\}$ with an edge if and only if $x \in V_k(R)$. Let $x \in H \setminus R$. We have the following equivalences:

$$\begin{aligned}
x \in V_k(R) \quad &\Leftrightarrow \quad x \in V(R \setminus K) \text{ for some } K \subseteq R, |K| = k \\
&\Leftrightarrow \quad x \in X((R \setminus K) \cup \{x\}) \text{ for some } K \subseteq R, |K| = k \\
&\Leftrightarrow \quad x \in X((R \cup \{x\}) \setminus K) \text{ for some } K \subseteq R, |K| = k,\ x \notin K \\
&\Leftrightarrow \quad x \in X((R \cup \{x\}) \setminus K) \text{ for some } K \subseteq R \cup \{x\}, |K| = k,\ x \notin K \\
&\Leftrightarrow \quad x \in X((R \cup \{x\}) \setminus K) \text{ for some } K \subseteq R \cup \{x\}, |K| = k \\
&\Leftrightarrow \quad x \in X_k(R \cup \{x\}),
\end{aligned}$$

where in the fifth step we used that $x \in X((R \cup \{x\}) \setminus K)$ can only occur if $x \in (R \cup \{x\}) \setminus K$.

So we can also define the graph as having an edge between R and $R \cup \{x\}$ if and only if $x \in X_k(R \cup \{x\})$. Since the sum of the degrees of the vertices in $\binom{H}{r}$ is the same as the sum of degrees of the vertices in $\binom{H}{r+1}$, we get

$$\binom{n}{r} v_{r,k} = \sum_{a \in \binom{H}{r}} \deg(a) = \sum_{b \in \binom{H}{r+1}} \deg(b) = \binom{n}{r+1} x_{r,k},$$

which is the claimed result. ◀

Note that as in the "normal" Sampling Lemma, the result holds as well for consistent spaces.

4.3 Violators after Removal

Suppose we sample R at random, and then remove an arbitrary set of k elements K_R according to some fixed rule P_k, and obtain the set $R_{P_k} = R \setminus K_R$. The expected number of violators of R_{P_k} is bounded by $v_{r,k} + k$. This follows since $v_{r,k}$ counts the expected number of violators in $H \setminus R$ that we can possibly get by removing *any* set of k elements and the removed points K_R can also be in $V(R_{P_k})$. Therefore $E[|V_{P_k}(R)|] \leq v_{r,k} + k$.

▶ **Theorem 14.** *Let (H, V) be a nondegenerate violator space of dimension δ, and let R be sampled u.a.r. from all subsets of H of size r. Let P_k be a fixed rule to remove k elements from the random sample. Then*

$$\mathbb{E}[|V_{P_k}(R)|] \leq v_{r,k} + k \leq \sum_{i=1}^{k+1} \delta^i \cdot \frac{n-r}{r+1} + k.$$

Proof. By Lemma 13, we need to bound $x_{r,k}$. To this end, we show that for all R,

$$|X_k(R)| \leq \sum_{i=1}^{k+1} \delta^i.$$

This holds, since by Lemma 12, at most $\sum_{i=0}^{k} \delta^i$ many *sets* of extreme elements can be obtained by removing k elements from R, and each of these sets has at most δ elements. ◀

By [5, Section 7.2], there exists an LP-type problem and a rule P_k, such that $|X_k(R)| = \Theta(\delta^{k+1})$, for $|R| = n-1$. However, the behavior of the bound is unknown for general r. To get better bounds on the expectation $x_{r,k}$, other methods need to be applied.

5 Matching Lower Bounds for Consistent Spaces

In this section we show the matching lower bound of Theorem 10 for consistent spaces for most relevant sizes of r, δ and k.

▶ **Lemma 15.** *Let $r = n^\alpha$, let $\alpha \in (0,1)$, $0 < \epsilon < \alpha$, $\gamma < \alpha - \epsilon$ be constants, and $1 \leq \delta \leq n^\gamma$. Let k, P_k as in Theorem 10. Let $x = \epsilon \frac{n}{r} \delta \log n = o(n)$. Then there exists a consistent space (H, V) of dimension δ, such that*

$$\mathbb{E}[|V_{P_k}(R)|] = (1 + o(1))\epsilon \frac{n}{r} \delta \log n = (1 + o(1))x.$$

Proof. Define (H, V) consistently as follows. The violator set of the empty set is defined as the empty set, $V(\emptyset) = \emptyset$. For all $B \subseteq H$ with $0 < |B| \leq \delta$, its violators are chosen u.a.r. of size $\epsilon \frac{n}{r} \delta \log n$ from $H \setminus B$.

For $R \subseteq H$ of size r we define the violators as follows. If there exists $B \subseteq R$, $0 < |B| \leq \delta$, such that $V(B) \cap R = \emptyset$, then $V(R) = V(B)$. If there exists more than one such B, choose the lexicographically smallest. If no such B exists then set $V(R) = V(\emptyset) = \emptyset$. Therefore for all R we have a basis of size at most δ. Denote the basis for R by B_R.

First we show that it suffices to treat the "worst case" $k = 0$. For $k > 0$ we can reduce the problem to the case $k = 0$ by the following construction, where for all R of size r it holds that $|V(R_{P_k})| \geq |V(R)|$: For R with $V(R) \neq \emptyset$, fix P_k such that none of the k removed elements are in B_R, i.e., $B_R \subseteq R_{P_k}$. Since $R \cap V(B) = \emptyset$ it follows that $R_{P_k} \cap V(B) = \emptyset$ and we can choose $V(R_{P_k}) = V(R)$. If there exists multiple sets of size r with nonempty violator set that are mapped to the same R_{P_k}, choose $V(R_{P_k})$ arbitrary from the set of their violator spaces. For all other sets of size $r - k$, choose their violators as the empty set. It follows that for all R of size r, $|V(R_{P_k})| \geq |V(R)|$. For all other $S \subseteq H$ define $V(S) = \emptyset$.

Hence we may assume that $k = 0$.

$$\mathbb{E}[|V(R)|] = \Pr[|V(R)| = \epsilon \frac{n}{r} \delta \log n] \cdot \epsilon \frac{n}{r} \delta \log n = (1 - \Pr[|V(R)| = 0]) \cdot \epsilon \frac{n}{r} \delta \log n.$$

We now show that $\Pr[|V(R)| = 0] = o(1)$, which concludes the proof. Because we chose the violators of the bases independently

$$\begin{aligned}\Pr[|V(R)| = 0] &= \Pr[\forall B \subseteq R, 0 < |B| \leq \delta \mid V(B) \cap R \neq \emptyset] \\ &= \prod_{\substack{B \subseteq R \\ 0 < |B| \leq \delta}} \Pr[V(B) \cap R \neq \emptyset] = \prod_{\substack{B \subseteq R \\ 0 < |B| \leq \delta}} (1 - \Pr[V(B) \cap R = \emptyset]).\end{aligned}$$

Now we bound $\Pr[V(B) \cap R = \emptyset]$ from below. For B of size β with $0 < \beta \leq \delta$ and $x = \epsilon \frac{x}{r} \delta \log n$, we get

$$\begin{aligned}\Pr[V(B) \cap R = \emptyset] &= \frac{\binom{n-x-\beta}{r-\beta}}{\binom{n-\beta}{r-\beta}} = \frac{(n-x-\beta) \cdots (n-x-r+1)}{(n-\beta) \cdots (n-r+1)} \\ &= e^{\left((1+o(1))\frac{x}{n} + \Theta(\frac{x^2}{n^2})\right) r} = n^{-(1+o(1))\epsilon\delta}.\end{aligned}$$

Using that $\sum_{i=1}^{\delta}\binom{r}{i} \geq \frac{(r-\delta)^\delta}{\delta^\delta}$ and $\delta = o(r)$ we get

$$\begin{aligned}\Pr[|V(R)|=0] &\leq \prod_{\substack{B\subseteq R\\ 0<|B|\leq\delta}} (1-n^{-(1+o(1))\epsilon\delta}) \leq (1-n^{-(1+o(1))\epsilon\delta})^{\frac{(r-\delta)^\delta}{\delta^\delta}} \\ &\leq \exp\left(-(1+o(1))n^{-(1+o(1))\epsilon\delta}\cdot\frac{(1+o(1))^\delta r^\delta}{\delta^\delta}\right).\end{aligned}$$

Plugging in $r = n^\alpha$, we observe that is sufficient to show that $n^{(\alpha-\epsilon+o(1))\delta}(1+o(1))^\delta\cdot\frac{1}{\delta^\delta} = \omega(1)$. By using $\delta \leq n^\gamma$ we get

$$\frac{n^{(\alpha-\epsilon+o(1))\delta}\cdot(1+o(1))^\delta}{\delta^\delta} \geq \left(n^{(\alpha-\epsilon-\gamma+o(1))}\cdot(1+o(1))\right)^\delta = \omega(1),$$

since $\gamma < \alpha-\epsilon$. ◀

We show that using one of the simplest violator spaces, namely the d-smallest number problem, we obtain the bound of $\mathbb{E}[|V_{P_k}(R)|] = \Theta(\frac{n}{r}\cdot(\delta+k))$.

▶ **Definition 16.** We define the *d-smallest number problem* as follows. Let $H = [n] = \{1,2,\ldots,n\}$. For $R \subseteq H$, define $\min_d(R)$ as the d-smallest number in R. Let $V(R) = \{r \in H\setminus R \mid r < \min_d(R)\}$, i.e. all elements smaller than the d-smallest.

We observe that (H,V) is a violator space, with combinatorial dimension d. The basis of R consists of the d smallest elements of R.

▶ **Lemma 17.** *Let $H = [n]$ and $\delta + k \leq r$. For the δ-smallest number problem there exists a rule P_k such that $\mathbb{E}[|V_{P_k}(R)|] = \frac{n-r}{r+1}\cdot(\delta+k)+k$, and therefore for $r = o(n)$, $\mathbb{E}[|V_{P_k}(R)|] = \Theta(\frac{n}{r}\cdot(\delta+k))$.*

Proof of Lemma 17. Let $R \subseteq H$. To maximize the number of violators after the removal of k elements, we remove the k-smallest elements of R. We call this rule P_k and the removed set R_k. Then

$$V_{P_k}(R) = \{r \in H\setminus R \mid r < \min_{\delta+k}(R)\} \cup R_k.$$

We observe that $\{r \in H\setminus R \mid r < \min_{\delta+k}(R)\}$ is exactly the set of violators of R, for the $\delta+k$ smallest problem, whose expected size we know by the Sampling Lemma 6. Hence

$$\mathbb{E}[|V_{P_k}(R)|] = \frac{n-r}{r+1}(\delta+k)+k. \quad ◀$$

Lemma 15 and Lemma 17 show that for consistent spaces the bound of Theorem 10 is tight up to a constant factor for most relevant values of r, δ and k, (i.e., if r,δ and k satisfy the conditions of Lemma 15 or Lemma 17). Furthermore by Lemma 17 if $k \geq \delta\log n$, then the upper bound of Theorem 10 is tight for violator spaces.

6 Tight Bounds for Violator Spaces with Combinatorial Dimension 1

In the case of violator spaces it is open whether (or when) the upper bound of Theorem 10 is tight for $k < \delta\log n$. In this case, there is a gap of up to $\log n$ between upper and lower bounds. Theorem 14 shows that for *non-degenerate* violator spaces, the bound of Theorem 10 is not tight for very small δ and k, but it could still be tight in the general case. For $k = 0$ we know a stronger upper bound of $O(\frac{n-r}{r+1}\delta)$ by the Sampling Lemma 6.

In the following we prove a stronger upper bound for $\delta = 1$ for violator spaces, and we give an example showing that this bound is tight. We prove that there exists only one class of violator spaces of dimension 1, namely the class of the smallest number with repetitions violator space.

▶ **Theorem 18.** *For $\delta = 1$, $\mathbb{E}[|V_{P_k}(R)|] = O(\frac{n}{r}k)$, and this bound is tight.*

It follows that the upper bound given in Theorem 14 is tight for all violator spaces of dimension 1.

▶ **Definition 19.** We define the class of *smallest number with repetitions* violator space as follows. Let $|H| = n$ and H a multiset of $[n]$, i.e., every element of H is in $[n]$ and there might be repetitions. For $R \subseteq H$, let $V(R) = \{x \in H \mid x < \min_{i \in R} i\}$. Finally we require that either $V(\emptyset) = H$ or $V(\emptyset) = V(i)$ for some $i \in H$.

Observe that this a a violator space of dimension 1 and similarly as in the proof of Lemma 17, we can show that $\mathbb{E}[|V_{P_k}(R)|] = O(\frac{n}{r}k)$.

▶ **Lemma 20.** *Let (H, V) be a violator space of dimension 1. If for all $i \in H$, $V(\emptyset) = H$ or $V(i) = \{x \in H \mid x < i\}$ and $V(\emptyset) = V(i)$ for some i, then $V(R) = \{x \in H \mid x < \min_{i \in R} i\}$ for all $R \subseteq H$. This means that the violators of sets of size 0 and 1 uniquely define all other violators.*

Proof. Let $R \subseteq H$, $|R| \geq 2$ with $y := \min_{i \in R} i$. Because the dimension of the violator space is 1 we have $V(R) = V(x)$ for some $x \in R$. Now for all $x > y$, we have $y \in V(x)$, hence $V(R) = V(y)$. ◀

▶ **Lemma 21.** *Every violator space (H, V) of dimension 1 with $|H| = n$, is homeomorphic to an instance of smallest number with repetitions, i.e., those are the only violator spaces of dimension* 1 *that exist.*

Proof. Let (H, V) be a violator space with $|H| = n$. Assume that $V(\emptyset) \neq H$. Then there exists $i \in H \setminus V(\emptyset)$ and therefore $V(i) = V(\emptyset)$. The following holds for all $i \neq j \in H$.

1. If $V(i) \neq V(j)$ then $i \in V(j)$ or $j \in V(i)$.
2. $V(i) \subseteq V(j)$ or $V(j) \subseteq V(i)$.

For the first part assume that $i \notin V(j)$ and $j \notin V(i)$. Then by locality $V(i, j) = V(i) = V(j)$.

For the second part assume that there exists $k \neq l$ such that $k \in V(i) \setminus V(j)$ and $l \in V(j) \setminus V(i)$. We consider $V(i, j, k, l)$. Since $\delta = 1$ we have that $V(i, j, k, l) = V(m)$ for some $m \in \{i, j, k, l\}$. By consistency we know $V(i, j, k, l) \neq V(i)$ since $k \in V(i)$. Similarly $V(i, j, k, l) \neq V(j)$. Therefore w.l.o.g. assume that $V(i, j, k, l) = V(k)$. Then $j \notin V(k)$ and $k \notin V(j)$ and hence by the first part $V(j) = V(k) = V(i, j, k, l)$, which is a contradiction.

We now construct a mapping $f : H \to [n]$, such that $j \in V(i)$ if and only if $f(j) < f(i)$. By Lemma 20 this concludes the proof.

We construct a sequence of pairwise disjoint nonempty sets $V_1, V_2, \ldots V_m \subseteq H$, $m \leq n$, $V_1 \cup \cdots \cup V_m = H$ such that the following holds for all $i \in [m]$: For all $x \in V_i$ we have $V(x) = V_1 \cup \cdots \cup V_{i-1}$.

By setting $f^{-1}(i) = V_i$ for all $i \in [m]$, this is one instance of minimum number with repetitions violator space.

Suppose that for some $i \geq 1$ we have constructed $V_1, \ldots, V_{i-1}$ and $H \setminus (V_1, \ldots, V_{i-1}) \neq \emptyset$. Let V_i be the subset of $H \setminus (V_1, \ldots, V_{i-1})$ with inclusion-minimal violator sets, i.e.,

$x \in H \setminus (V_1, \dots, V_{i-1})$ is in V_i if and only if there exists no $y \in H \setminus (V_1, \dots, V_{i-1})$ such that $V(y) \subsetneq V(x)$. Then obviously V_i is nonempty. We need to show that for such x, $V(x) = V_1 \cup \dots \cup V_{i-1}$. Let $y \in V(x)$. Since $y \notin V(y)$ condition 2. implies that $V(y) \subsetneq V(x)$, hence $y \in V_1 \cup \dots \cup V_{i-1}$. Now let $y \notin V(x)$. If $x \notin V(y)$ then by condition 1. it follows that $V(x) = V(y)$ and hence $y \in V_i$. Otherwise $x \in V(y)$ and therefore by condition 2. $V(x) \subsetneq V(y)$. It follows that $y \notin V_1 \cup \dots \cup V_{i-1}$. ◀

Proof of Theorem 18. The bound follows immediately from Lemma 21 and by the fact that $\mathbb{E}[|V_{P_k}(R)|] = O(\frac{n}{r}k)$, for the smallest number with repetitions LP-type problem. Tightness follows from Lemma 17. ◀

7 Acknowledgments

The authors are grateful to Kenneth Clarkson and Emo Welzl for sharing important insights. Furthermore we thank Luis Barba for useful discussions.

References

1 Y. Brise and B. Gärtner. Clarkson's algorithm for violator spaces. *Computational Geometry*, 44(2):70 – 81, 2011. Special issue of selected papers from the 21st Annual Canadian Conference on Computational Geometry.

2 M. C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM J. Optim.*, 19:1211–1230, 2008.

3 M. C. Campi and S. Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *J. Optim. Theory Appl.*, 148:257–280, 2011.

4 T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms.* The MIT Press, Cambridge, MA., 1990.

5 B. Gärtner. Sampling with removal in LP-type problems. *Journal of Computational Geometry*, 6(2):93–112, 2015.

6 B. Gärtner, J. Matoušek, L. Rüst, and P. Škovroň. Violator spaces: Structure and algorithms. *Discrete Appl. Math.*, 156(11):2124–2141, June 2008.

7 B. Gärtner, W. D. Morris, Jr., and L. Rüst. Unique sink orientations of grids. In *Proc. 11th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, volume 3509 of *Lecture Notes in Computer Science*, pages 210–224. Springer-Verlag, 2005.

8 B. Gärtner and E. Welzl. A simple sampling lemma: Analysis and applications in geometric optimization. *Discrete & Computational Geometry*, 25(4):569–590, 2001.

9 J. Matoušek, M. Sharir, and E. Welzl. A subexponential bound for linear programming. *Algorithmica*, 16:498–516, 1996.

10 J. Matoušek. Removing degeneracy in LP-type problems revisited. *Discrete & Computational Geometry*, 42(4):517–526, 2009.

11 M. Sharir and E. Welzl. A combinatorial bound for linear programming and related problems. In *Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, STACS '92, pages 569–579, London, UK, UK, 1992. Springer-Verlag.

12 T. Szabó and E. Welzl. Unique sink orientations of cubes. In *Proc. 42nd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 547–555, 2000.

13 E. Welzl. Smallest enclosing disks (balls and ellipsoids). In *Results and New Trends in Computer Science*, pages 359–370. Springer-Verlag, 1991.