Annales Henri Lebesgue
Article à paraître.

ANNALES
HENRI LEBESGUE

IMRE BÁRÁNY
MATTHIEU FRADELIZI
XAVIER GOAOC
ALFREDO HUBARD
GÜNTER ROTE

RANDOM POLYTOPES AND THE WET PART FOR ARBITRARY PROBABILITY DISTRIBUTIONS

POLYTOPES ALÉATOIRES ET PARTIES IMMERGÉES DE MESURES DE PROBABILITÉS ARBITRAIRES

Abstract. — We examine how the measure and the number of vertices of the convex hull of a random sample of n points from an arbitrary probability measure in $\mathbb{R}^d$ relate to the

Keywords: Random polytope, floating body, ε-nets.
2010 *Mathematics Subject Classification:* 52B05, 52A22.
DOI unassigned yet
(*) I. B. was supported by the Hungarian National Research, Development and Innovation Office NKFIH Grants K 111827 and K 116769, and by the Bézout Labex (ANR-10-LABX-58). X. G. was supported by Institut Universitaire de France. The authors are grateful for the hospitality during the ASPAG (ANR-17-CE40-0017) workshop on geometry, probability, and algorithms in Arcachon in April 2018.

wet part of that measure. This extends classical results for the uniform distribution from a convex set proved by Bárány and Larman in 1988. The lower bound of Bárány and Larman continues to hold in the general setting, but the upper bound must be relaxed by a factor of $\log n$. We show by an example that this is tight.

RÉSUMÉ. — Nous examinons comment la mesure et le nombre de sommets de l'enveloppe convexe d'un échantillon aléatoire de n points tirés selon une mesure de probabilité arbitraire sur $\mathbb{R}^d$ sont reliés à la partie immergée de la mesure. Cela étend des résultats classiques pour la mesure uniforme sur un convexe montrés par Bárány et Larman en 1988. La minoration de Bárány et Larman est toujours vraie dans ce cadre général mais la majoration doit être affaiblie d'un facteur $\log n$. Nous montrons par un exemple que cette borne est optimale.

1. Introduction and Main Results

Let K be a convex body (convex compact set with non-empty interior) in $\mathbb{R}^d$, and let $X_n = \{x_1, \dots, x_n\}$ be a random sample of n uniform independent points from K. The set $P_n = \operatorname{conv} X_n$ is a *random polytope* in K. For $t \in [0, 1)$ we define the *wet part* K_t of K:

$$K_t = \{\, x \in K : \text{there is a halfspace } h \text{ with } x \in h \\ \text{and } \operatorname{Vol}(K \cap h) \leqslant t \operatorname{Vol} K \,\}$$

The name "wet part" comes from the mental picture when K is in $\mathbb{R}^3$ and contains water of volume $t \operatorname{Vol} K$. Bárány and Larman [2] proved that the measure of the wet part captures how well P_n approximates K in the following sense:

THEOREM 1.1 ([2, Theorem 1]). — *There are constants c and N_0 depending only on d such that for every convex body K in $\mathbb{R}^d$ and for every $n > N_0$*

$$\frac{1}{4} \operatorname{Vol} K_{1/n} \leqslant \mathbb{E}[\operatorname{Vol}(K \setminus P_n)] \leqslant \operatorname{Vol} K_{c/n}.$$

By Efron's formula (see (1.2) below), this directly translates into bounds for the expected number of vertices of P_n, see Section 1.2.

1.1. Results for general measures.

The notions of random polytope and wet part extend to a general probability measure μ defined on the Borel sets of $\mathbb{R}^d$. The definition of a μ-random polytope P_n^μ is clear: X_n is a sample of n random independent points chosen according to μ, and $P_n^\mu = \operatorname{conv} X_n$. The wet part W_t^μ is defined as

$$W_t^\mu = \{\, x \in \mathbb{R}^d : \text{ there is a halfspace } h \text{ with } x \in h \text{ and } \mu(h) \leqslant t \,\}.$$

The μ-measure of the wet part is denoted by $w^\mu(t) := \mu(W_t^\mu)$. Here is an extension of Theorem 1.1 to general measures:

THEOREM 1.2. — *For any probability measure μ in $\mathbb{R}^d$ and $n \geqslant 2$,*

$$\frac{1}{4} w^\mu(\tfrac{1}{n}) \leqslant \mathbb{E}[1 - \mu(P_n^\mu)] \leqslant w^\mu((d+2)\tfrac{\ln n}{n}) + \tfrac{\varepsilon_d(n)}{n},$$

where $\varepsilon_d(n) \to 0$ as $n \to +\infty$ and is independent of μ.

A similar upper bound, albeit with worse constants, follows from a result of Vu [15, Lemma 4.2], which states that P_n^μ contains $\mathbb{R}^d \setminus W^\mu_{c \ln n/n}$ with high-probability. Since a containment with high probability is usually stronger than an upper bound in expectation, one may have hoped that the $\log n/n$ in the upper bound of Theorem 1.2 can be reduced. Our main result shows that this is not possible, not even in the plane:

Theorem 1.3. — *There exists a probability measure ν on $\mathbb{R}^2$ such that*

$$\mathbb{E}[1-\mu(P_n^\nu)] > \frac{1}{2} \cdot w^\nu(\log_2 n/n)$$

for infinitely many n.

The measure that we construct actually has compact support and can be embedded into $\mathbb{R}^d$ for any $d \geqslant 2$. It will be apparent from the proof that the same construction has the stronger property that for *every* constant $C > 0$, the inequality $\mathbb{E}[1-\mu(P_n^\nu)] > \frac{1}{2} \cdot w^\nu(C \log_2 n/n)$ holds for infinitely many values n.

1.2. Consequences for f-vectors

Let $f_0(P_n^\mu)$ denote the number of vertices of P_n^μ. For *non-atomic* measures (measures where no single point has positive probability), Efron's formula [7] relates $E\left[f_0(P_n^\mu)\right]$ and $\mathbb{E}\left[\mu(P_n^\mu)\right]$:

$$\mathbb{E}[f_0(P_n^\mu)] = \sum_{i=1}^{n} \Pr\left[x_i \notin \operatorname{conv}(X_n \setminus \{x_i\})\right] \tag{1.1}$$

$$= n \cdot \int_x \Pr\left[x \notin P_{n-1}^\mu\right] d\mu(x) = n(1-\mathbb{E}[\mu(P_{n-1}^\mu)]) \tag{1.2}$$

For *any* measure, this still holds as an inequality:

$$\mathbb{E}[f_0(P_n^\mu)] \geqslant \sum_{i=1}^{n} \Pr\left[x_i \notin \operatorname{conv}(X_n \setminus \{x_i\})\right] = n(1-\mathbb{E}[\mu(P_{n-1}^\mu)]) \tag{1.3}$$

The measure that is constructed in Theorem 1.3 is non-atomic. As a consequence, Theorems 1.2 and 1.3 give the following bounds for the number of vertices:

Theorem 1.4. —

(i) *For any non-atomic probability measure μ in $\mathbb{R}^d$,*

$$\frac{1}{e} n w^\mu(\tfrac{1}{n}) \leqslant \mathbb{E}\left[f_0(P_n^\mu)\right] \leqslant n w^\mu\Big((d+2)\tfrac{\ln n}{n}\Big) + \varepsilon_d(n),$$

where $\varepsilon_d(n) \to 0$ as $n \to +\infty$ and is independent of μ.

(ii) *There exists a non-atomic probability measure ν on $\mathbb{R}^2$ such that*

$$\mathbb{E}\left[f_0(P_n^\nu)\right] > \frac{1}{2} n \cdot w^\nu(\log_2 n/n)$$

for infinitely many n.

Theorem 1.4 follows from Theorems 1.2 and 1.3 except that Efron's Formula (1.2) induces a shift in indices, as it relates $f_0(P_n^\mu)$ to $\mu(P_{n-1}^\mu)$. This shift affects only the constant in the lower bound of Theorem 1.4(i), which goes from $\frac{1}{4}$ to $\frac{1}{e}$, see Section 3.1.

The upper bound of Theorem 1.4(i) fails for general distributions. For instance, if μ is a discrete distribution on a finite set, then $w^\mu(t) = 0$ for any t smaller than the mass of any single point and the upper bound cannot hold uniformly as $n \to \infty$. Of course, in that case Inequality (1.3) is strict.

For convex bodies, the number $f_i(P_n)$ of i-dimensional faces of P_n can also be controlled via the measure of the wet part since Bárány [1] proved that $\mathbb{E}[f_i(P_n)] = \Theta(n \operatorname{Vol} K_{1/n})$ for every $0 \leqslant i \leqslant d-1$. No similar generalization is possible for Theorem 1.2. Indeed, consider a measure μ in $\mathbb{R}^4$ supported on two circles, one on the (x_1, x_2)-plane, the other in the (x_3, x_4)-plane, and uniform on each circle; P_n^μ has $\Omega(n^2)$ edges almost surely.

Before we get to the proofs of Theorems 1.2 (Section 3.2) and 1.3 (Section 4), we discuss in Section 2 a key difference between the wet parts of convex bodies and of general measures.

2. Wet part: convex sets versus general measures

A key ingredient in the proof of the upper bound of Theorem 1.1 in [2] is that for a convex body K in $\mathbb{R}^d$, the measure of the wet part K_t cannot change too abruptly as a function of t: If $c \geqslant 1$, then

$$\operatorname{Vol} K_t \leqslant \operatorname{Vol} K_{ct} \leqslant c' \operatorname{Vol} K_t \tag{2.1}$$

where c' is a constant that depends only on c and d [2, Theorem 7]. In particular, a multiplicative factor can be taken out of the volume parameter of the wet part and the upper bound in Theorem 1.1 can be equivalently expressed as

$$\mathbb{E}[\operatorname{Vol}(K \setminus P_n)] \leqslant c' \operatorname{Vol} K_{1/n}. \tag{2.2}$$

(This is actually the way how the upper bound of Theorem 1.1 is formulated in [2, Theorem 1].) This alternative formulation shows immediately that the lower bound of Theorem 1.1 (and hence also of Theorem 1.2) cannot be improved by more than a constant.

2.1. Two circles and a sharp drop

The right inequality in (2.1) does not extend to general measures. An easy counterexample to the corresponding bound

$$w^\mu(ct) \leqslant c' w^\mu(t) \tag{2.3}$$

is the following "drop construction". It is a probability measure μ in the plane supported on two concentric circles, uniform on each of them, and with measure p

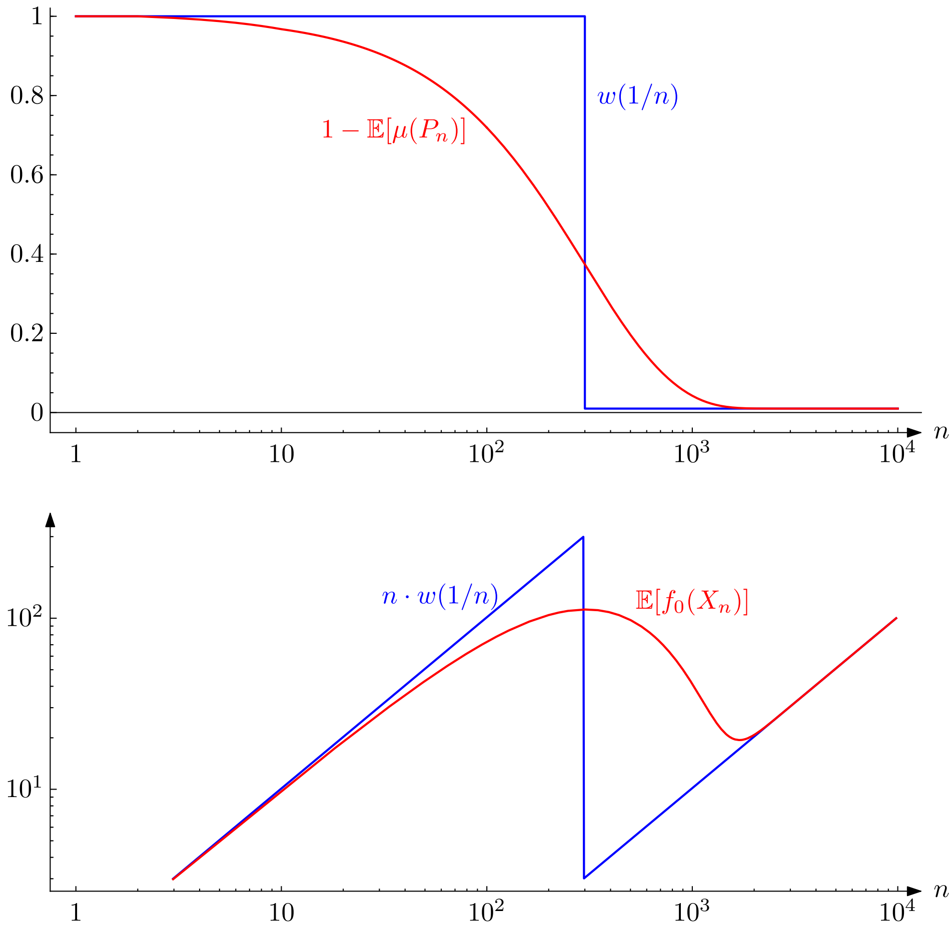

Figure 2.1. The quantities involved in Theorems 1.1–1.4 for the drop construction with $p = 1/100$, when the outer circle has twice the radius of the inner circle. Top: $\mathbb{E}[1-\mu(P_n^\mu)]$ and $w(1/n)$, the x-axis being a logarithmic scale. Bottom: $\mathbb{E}[f_0(P_n^\mu)]$ and $n \cdot w(1/n)$ on a doubly-logarithmic scale.

on the outer circle. Let τ denote the measure of a halfplane externally tangent to the inner circle; remark that $\tau < p/2$. As t is decreased, the measure $w^\mu(t)$ of the wet part drops at $t = \tau$:

$$w^\mu(t) = \begin{cases} p, & \text{if } t < \tau \\ 1, & \text{if } t \geqslant \tau \end{cases} \tag{2.4}$$

We can make this drop arbitrarily sharp by choosing a small p. In particular, for any given c', setting $p < \frac{1}{c'}$ makes it impossible to fulfill (2.3) when t is chosen in the range $t < \tau < ct$.

This example also challenges Inequality (2.2). As shown in Figure 2.1 (top), the function $w^\mu(1/n)$ has a sharp drop, while $\mathbb{E}\,[1-\mu(P_n^\mu)]$ shifts from the higher to the lower branch of the step in a gradual way. For this construction, the straightforward extension of Theorem 1.1 would imply that $\mathbb{E}[1-\mu(P_n^\mu)]$ remains within a constant multiplicative factor of $w^\mu(1/n)$. Thus, $\mathbb{E}[1-\mu(P_n^\mu)]$ would have to follow the steep drop.

2.2. A drop for the number of vertices.

The fact that $\mathbb{E}[1-\mu(P_n^\mu)]$ cannot drop too sharply is more easily seen by examining $\mathbb{E}\left[f_0(P_n^\mu)\right]$. Since the measure defined in Equation (2.4) is non-atomic, Efron's Formula (1.2) applies, so let us compare $\mathbb{E}\left[f_0(P_n^\mu)\right]$ and $n \cdot w^\mu(1/n)$. As illustrated in Figure 2.1 (bottom), $n \cdot w^\mu(1/n)$ has a sawtooth shape with a sharp drop from 300 to 3 at $n = 300$, and $\mathbb{E}[f_0(P_n)]$ does actually shift from the higher to the lower branch of the sawtooth, in a gradual way.

The fact that $\mathbb{E}[f_0(P_n^\mu)]$ can *decrease* is perhaps surprising at first sight, but this phenomenom is easy to explain: We pick random points one by one. As long as all points lie on the inner circle, $f_0(P_n^\mu) = n$. The first point to fall on the outer circle swallows a constant fraction of the points into the interior of P_n^μ, while adding only a single new point on the convex hull, causing a big drop. This happens around $n \approx 1/p$.

Again, the straightforward extension of Theorem 1.1 would imply that $\mathbb{E}[f_0(P_n)]$ follows the steep drop. Yet, on average, a single additional point can reduce $f_0(P_n)$ by a factor of at most 1/2. Hence, the drop of $\mathbb{E}[f_0(P_n)]$ cannot be so abrupt as the drop of $n \cdot w^\mu(1/n)$, for p small enough.

2.3. A sequence of drops

We prove Theorem 1.3 in Section 4 by an explicit construction that sets up a sequence of such drops. The function $n \cdot w^\mu(1/n)$ reaches larger and larger peaks as n increases, while dropping down more and more steeply between those peaks. Our proof of Theorem 1.3 will not actually refer to any drop or oscillating behavior. We will simply identify a sequence of values $n = n_1, n_2, \ldots$ for which $\mathbb{E}[1-\mu(P_n^\mu)]$ is larger than $\frac{1}{2}w^\mu(\log_2 n/n)$.

2.4. Open questions

It is an outstanding open problem whether a drop as exhibited by our two-circle construction can occur for the uniform selection from a *convex* body: Can the expectation of the number of vertices of a random polytope decrease in such a setting? This is impossible in the plane [6] or for the three dimensional ball [4], but open in general. See [5] and the discussion therein; see also [10] for the same problem for some particular rotationally invariant measures.

Perhaps Theorem 1.1 remains valid for some restricted class of measures μ, for instance, logconcave measures. One approach to circumvent the "impossibility result" of Theorem 1.3 would be to first extend (2.1) and establish that for $c > 1$ there is c' such that for all $t > 0$

$$w^\mu(t) \leqslant w^\mu(ct) \leqslant c' \cdot w^\mu(t).$$

The second step would derive from this property the extension of Theorem 1.1. We don't know if any of these two steps is valid.

We can weaken the claim of Theorem 1.1 in a different way, while maintaining it for all measures. For example, it is plausible that the upper bound in the theorem holds for a subset of numbers $n \in \mathbb{N}$ of positive density. On the other hand we do not know if there is a measure for which the bound of Theorem 1.1 is valid only for a finite number of natural numbers.

3. Proof of Theorem 1.2

Let μ be a probability measure in $\mathbb{R}^d$. For better readability we drop all superscripts μ.

3.1. Lower bound

The proof of the lower bound is similar to the one in the convex-body case. For every fixed point $x \in W_t$, by definition, there exists a half-space h with $x \in h$ and $\mu(h) \leqslant t$. If $h \cap P_n$ is empty, then x is not in P_n, and therefore, for $x \in W_t$,

$$\Pr[x \notin P_n] \geqslant \Pr[h \cap P_n = \emptyset] = (1 - \mu(h))^n \geqslant (1-t)^n. \tag{3.1}$$

Then, for any t,

$$\begin{aligned} 1 - \mathbb{E}[\mu(P_n)] &= \int_{x \in \mathbb{R}^d} \Pr[x \notin P_n] d\mu(x) \\ &\geqslant \int_{x \in W_t} \Pr[x \notin P_n] d\mu(x) \\ &\geqslant \int_{x \in W_t} (1-t)^n d\mu(x) = (1-t)^n w(t). \end{aligned}$$

We choose $t = 1/n$. Since the sequence $(1 - \frac{1}{n})^n$ is increasing, for $n \geqslant 2$ we have $1 - \mathbb{E}[\mu(P_n)] \geqslant \frac{1}{4} w(\frac{1}{n})$. □

To obtain the analogous lower bound from Theorem 1.4(i), we write

$$\mathbb{E}\left[f_0(P_n)\right] = n \mathbb{E}\left[1 - \mu(P_{n-1})\right] \geqslant n(1-t)^{n-1} w(t).$$

Again, choosing $t = 1/n$ yields the claimed lower bound

$$\mathbb{E}\left[f_0(P_n)\right] \geqslant n \left(1 - \frac{1}{n}\right)^{n-1} w(\tfrac{1}{n}) \geqslant \frac{1}{e} n w(\tfrac{1}{n}),$$

since the sequence $(1 - \frac{1}{n})^{n-1}$ is now decreasing to $\frac{1}{e}$.

3.2. Floating bodies and ε-nets

Before we turn our attention to the upper bound, we will point out a connection to ε-nets. Consider a probability space (U, μ) and a family $\mathcal{H}$ of measurable subsets of U. An *ε-net* for $(U, \mu, \mathcal{H})$ is a set $S \subseteq U$ that intersects every $h \in \mathcal{H}$ with $\mu(h) \geqslant \varepsilon$ [12, §10.2]. In the special case where $U = (\mathbb{R}^d, \mu)$ and $\mathcal{H}$ consists of all half-spaces, if a set S is an ε-net, then the convex hull P of S contains $\mathbb{R}^d \setminus W_\varepsilon$. Indeed, assume that there exists a point x in $\mathbb{R}^d \setminus W_\varepsilon$ and not in P. Consider a closed halfspace h that contains x and is disjoint from P. Since $x \notin W_\varepsilon$ we must have $\mu(h) > \varepsilon$ and S cannot be an ε-net.

We call the region $\mathbb{R}^d \setminus W_\varepsilon$ the *floating body* of the measure μ with parameter ε, by analogy to the case of convex bodies. The relation between floating bodies and ε-nets was first observed by Van Vu, who used the ε-net Theorem to prove that P_n^μ contains $\mathbb{R}^d \setminus W_{c \log n/n}$ with high probability [15, Lemma 4.2] (a fact previously established by Bárány and Dalla [3] when μ is the normalized Lebesgue measure on a convex body). This implies that, with high probability, $1 - \mu(P_n) \leqslant w(c \log n/n)$. The analysis we give in Section 3.3 refines Vu's analysis to sharpen the constant. Note that Theorem 1.3 shows that Vu's result is already asymptotically best possible.

3.3. Upper bound

For $d = 1$, the proof of the upper bound is straightforward and may actually be improved. Indeed, we have $w(t) = \min\{2t, 1\}$, and Efron's Formula (1.3) yields

$$\mathbb{E}\left[1 - \mu(P_n)\right] \leqslant \frac{1}{n+1}\mathbb{E}\left[f_0(P_{n+1})\right] \leqslant \frac{2}{n+1} \leqslant w\left(\frac{1}{n+1}\right) \leqslant w\left(3\frac{\ln n}{n}\right).$$

We will therefore assume $d \geqslant 2$.

We use a lower bound on the probability of a random sample of U to be an ε-net for $(U, \mu, \mathcal{H})$. We define the *shatter function* (or growth function) of the family $\mathcal{H}$ as

$$\pi_{\mathcal{H}}(N) = \max_{X \subseteq U, |X| \leqslant N} |\{\, X \cap h : h \in \mathcal{H} \,\}|,$$

A proof is in Appendix A.

Lemma 3.1 ([11, Theorem 3.2]). — *Let (U, μ) be a probability space and $\mathcal{H}$ a family of measurable subsets of U. Let X_s be a sample of s random independent elements chosen according to μ. For any integer $N > s$, the probability that X_s is not a ε-net for $(U, \mu, \mathcal{H})$ is at most*

$$2\pi_{\mathcal{H}}(N) \cdot (1 - \tfrac{s}{N})^{(N-s)\varepsilon - 1}.$$

Lemma 3.1 is a quantitative refinement of a foundational result in learning theory [14, Theorem 2]. It is commonly used to prove that small ε-nets exist for range spaces of bounded Vapnik-Chervonenkis dimension [9], see also [11, Theorem 3.1] or [13, Theorem 15.5]. For that application, it is sufficient to show that the probability of failure is less than 1; This works for $\varepsilon \approx d \ln n/n$ (with appropriate lower-order terms), where d is the Vapnik-Chervonenkis dimension. In our proof, we will need a smaller failure probability of order $o(1/n)$, and we will achieve this by setting $\varepsilon \approx (d+2)\ln n/n$. We will apply the lemma in the case where $U = \mathbb{R}^d$ and $\mathcal{H}$ is the set of halfspaces in $\mathbb{R}^d$. We mention that by increasing ε more agressively, the probability of failure can be made exponentially small.

For the family $\mathcal{H}$ of halfspaces in $\mathbb{R}^d$, we have the following sharp bound on the shatter function [8]:

$$\pi_{\mathcal{H}}(N) \leqslant 2\sum_{i=0}^{d} \binom{N-1}{i}.$$

The proof of the upper bound of Theorem 1.2 starts by remarking that for any $\varepsilon \in [0,1]$ we have:

$$\begin{aligned}\mathbb{E}\left[1-\mu(P_n)\right] &= \int_{\mathbb{R}^d} \Pr[x \notin P_n] d\mu(x)\\ &= \int_{\mathbb{R}^d\setminus W_\varepsilon} \Pr[x \notin P_n] d\mu(x) + \int_{W_\varepsilon} \Pr[x \notin P_n] d\mu(x)\\ &\leqslant \int_{\mathbb{R}^d\setminus W_\varepsilon} \Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] d\mu(x) + \int_{W_\varepsilon} d\mu(x)\\ &\leqslant \Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] + w(\varepsilon).\end{aligned}$$

Here, the first inequality between the probabilities holds since the event $x \notin P_n$ trivially implies that $\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n$ when $x \in \mathbb{R}^d \setminus W_\varepsilon$. We thus have

$$\mathbb{E}\left[1-\mu(P_n)\right] \leqslant w(\varepsilon) + \Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n].$$

We now want to set ε so that $\Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n]$ is $\frac{\varepsilon_d(n)}{n}$ with $\varepsilon_d(n) \to 0$ as $n \to \infty$. As shown in Section 3.2, the event $\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n$ implies that P_n fails to be an ε-net. The probability can thus be bounded from above using Lemma 3.1 with $s = n$. Taking logarithms, for any $N > n$,

$$\ln \Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] \leqslant \ln \pi_{\mathcal{H}}(N) + ((N-n)\varepsilon - 1)\ln(1-\tfrac{n}{N}) + \ln 2.$$

Since we assume that $d \geqslant 2$, we have

$$\pi_{\mathcal{H}}(N) \leqslant 2\sum_{i=0}^{d}\binom{N-1}{i} \leqslant N^d \quad \text{and} \quad \ln \pi_{\mathcal{H}}(N) \leqslant d\ln N.$$

We set $N = n\lceil \ln n\rceil$, so that:

$$\begin{aligned}\ln \Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] \leqslant\ & d\ln n + d\ln\lceil \ln n\rceil\\ &+ ((N-n)\varepsilon - 1)\ln(1-\tfrac{n}{N}) + \ln 2.\end{aligned}$$

We then set $\varepsilon = \delta\frac{\ln n}{n}$, with $\delta \approx d$ to be fine-tuned later. If n is large enough, the factor $((N-n)\varepsilon - 1) \approx \delta \ln^2 n$ is nonnegative, and we can use the inequality $\ln(1-x) \leqslant -x$ for $x \in [0,1)$ in order to bound the second term:

$$\begin{aligned}((N-n))\varepsilon - 1)\ln(1-\tfrac{n}{N}) &\leqslant -\Big((N-n)\varepsilon - 1\Big)\frac{n}{N}\\ &= -n\varepsilon + \frac{n}{\lceil \ln n\rceil}\varepsilon + \frac{1}{\lceil \ln n\rceil} \leqslant -\delta\ln n + \delta + 1.\end{aligned}$$

Altogether, we get

$$\Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] \leqslant 2^{\delta+1}e\cdot n^{d-\delta}\lceil \ln n\rceil^d$$

so for every $\delta > d+1$ we have $\Pr[\mathbb{R}^d \setminus W_\varepsilon \not\subseteq P_n] = \frac{\varepsilon_d(n)}{n}$ with $\varepsilon_d(n) \to 0$ as $n \to \infty$. Setting $\delta = d+2$ yields the claimed bound. □

4. Proof of Theorem 1.3

In this section, logarithms are base 2. For better readability we drop the superscripts ν.

4.1. The construction

The measure ν is supported on a sequence of concentric circles $C_1, C_2, \ldots$, where C_i has radius

$$r_i = 1 - \frac{1}{i+1}.$$

On each C_i, ν is uniform, implying that ν is rotationally invariant. We let $D_i = \bigcup_{j \geqslant i} C_j$. For $i \geqslant 1$ we put

$$\nu(D_i) = s_i := 4 \cdot 2^{-2^i}$$

and remark that $\nu(\mathbb{R}^2) = s_1 = 1$, so ν is a probability measure. The sequence $\{s_i\}_{i \in \mathbb{N}}$ decreases very rapidly. The probabilities of the individual circles are

$$p_i := \nu(C_i) = s_i - s_{i+1} = 4 \cdot \left(2^{-2^i} - 2^{-2^{i+1}}\right) = s_i \left(1 - \frac{s_i}{4}\right) \approx s_i,$$

for $i \geqslant 1$.

The infinite sequence of values n for which we claim the inequality of Theorem 1.3 is

$$n_i := 2^{2^i + 2i} \approx \tfrac{1}{s_i} \log^2 \tfrac{1}{s_i}.$$

In Section 4.2, we examine the wet part and prove that $w(\frac{\log n_i}{n_i}) \leqslant s_i$. We then want to establish the complementary bound $\mathbb{E}\left[1 - \nu(P_{n_i})\right] > s_i/2$. Since ν is non-atomic, Efron's formula yields

$$\mathbb{E}\left[1 - \nu(P_{n_i})\right] = \frac{1}{n_i + 1} \mathbb{E}\left[f_0(P_{n_i+1})\right]$$

and it suffices to establish that $\mathbb{E}\left[f_0(P_{n_i+1})\right] > (n_i + 1) s_i/2$. This is what we do in Section 4.3.

4.2. The wet part

Let us again drop the superscript ν. Let h_i be a closed halfplane that has a single point in common with C_i, so its bounding line is tangent to C_i. We have

$$w(t) = s_i, \ \text{ for } \nu(h_i) \leqslant t < \nu(h_{i-1}).$$

So, as t decreases, $w(t)$ drops step by step, each step being from s_i to s_{i+1}. In particular,

$$w(t) \leqslant s_i \iff t < \nu(h_{i-1}). \tag{4.1}$$

For $j > i$, the portion of C_j contained in h_i is equal to $2 \arccos(r_i/r_j)$. Hence,

$$\nu(h_i \cap C_j) = \frac{\arccos(r_i/r_j)}{\pi} \cdot p_j.$$

We will bound the term $\arccos(r_i/r_j)$ by a more explicit expression in terms of i. To get rid of the arccos function, we use the fact that $\cos x \geqslant 1 - x^2/2$ for all $x \in \mathbb{R}$. We obtain, for $0 \leqslant y \leqslant 1$,

$$\arccos(1 - y) \geqslant \sqrt{2y}.$$

Moreover, the ratio r_i/r_j can be bounded as follows:

$$\frac{r_i}{r_j} \leqslant \frac{r_i}{r_{i+1}} = \frac{i}{i+1} \Big/ \frac{i+1}{i+2} = 1 - \frac{1}{(i+1)^2}.$$

Thus we deduce that

$$\frac{\arccos(r_i/r_j)}{\pi} \geqslant \frac{\arccos(1-1/(i+1)^2)}{\pi} \geqslant \frac{\sqrt{2}}{\pi(i+1)}.$$

We have established a bound on $\arccos(r_i/r_j)/\pi$, which is the fraction of a single circle C_j that is contained in h_i. Hence, considering all circles C_j with $j > i$ together, we get

$$\nu(h_i) \geqslant \frac{\sqrt{2}}{\pi(i+1)} s_{i+1}.$$

We check that for $i \geqslant 4$,

$$\frac{\log n_i}{n_i} = \frac{2^i + 2i}{2^{2^i+2i}} = 2^{-2^i}2^{-2i}(2^i+2i) = \frac{s_i}{4}2^{-i}(1+2^{1-i}i) < s_i\frac{\sqrt{2}}{\pi i} \leqslant \nu(h_{i-1}),$$

because $2^{-i}(1+2^{1-i}i) < \frac{\sqrt{2}}{\pi i}$ for all $i \geqslant 4$. Using (4.1), this gives our desired bound:

$$w(\tfrac{\log n_i}{n_i}) \leqslant s_i,$$

for all $i \geqslant 4$. With little effort, one can show that actually $w(\frac{\log n_i}{n_i}) = s_i$. One can also see that, for any $C > 0$, the condition $w(C\frac{\log n_i}{n_i}) \leqslant s_i$ holds if i is large enough, because the exponential factor 2^{-i} dominates any constant factor C in the last chain of inequalities. This justifies the remark that we made after the statement of Theorem 1.3.

4.3. The random polytope

Assume now that X_n is a set of n points sampled independently from ν. We intend to bound from below the expectation $\mathbb{E}\left[f_0(\operatorname{conv} X_{n_i+1})\right]$. Observe that for any $n \in \mathbb{N}$ one has

$$\mathbb{E}|X_n \cap C_i| = np_i \quad \text{and} \quad \Pr(X_n \cap D_{i+1} = \emptyset) = (1-s_{i+1})^n.$$

Intuitively, as n varies in the range near n_i, many points of X_n lie on C_i and yet no point of X_n lies in D_{i+1}. So P_n has, in expectation, at least $np_i \approx ns_i$ vertices. At the same time, the term $w(\log n/n)$ in the claimed lower bound drops to s_i. So the expected number of vertices is about ns_i which is larger than $\frac{1}{2}ns_i = \frac{n}{2}w(\log n/n)$.

Formally, we estimate the expected number of vertices when $n = n_i + 1$:

$$\begin{aligned}
&\mathbb{E}\left[f_0(\operatorname{conv} X_{n_i+1})\right] \\
&\quad\geqslant \mathbb{E}\left[f_0(\operatorname{conv} X_{n_i+1}) \mid X_{n_i+1} \cap D_{i+1} = \emptyset\right] \cdot \Pr(X_{n_i+1} \cap D_{i+1} = \emptyset) \\
&\quad\geqslant \mathbb{E}[|X_{n_i+1} \cap C_i|] \cdot (1-s_{i+1})^{n_i+1} \\
&\quad= (n_i+1)p_i(1-s_{i+1})^{n_i+1} \\
&\quad= (n_i+1)s_i\left[\frac{p_i}{s_i}(1-s_{i+1})^{n_i+1}\right]
\end{aligned}$$

The last square bracket tends to 1 as $i \to \infty$. In particular, it is larger than $\frac{1}{2}$ for $i \geqslant 4$. This shows that for all $i \geqslant 4$

$$\mathbb{E}\left[f_0(\operatorname{conv} X_{n_i+1})\right] > \frac{1}{2}(n_i+1)s_i \geqslant \frac{1}{2}(n_i+1)w(\tfrac{\log n_i}{n_i}). \qquad \square$$

4.4. Higher dimension

We can embed the plane containing ν in $\mathbb{R}^d$ for $d \geqslant 3$. The analysis remains true but the random polytope is of course flat with probability 1. To get a full-dimensional example, we can replace each circle by a $(d-1)$-dimensional sphere, all other parameters being kept identical: all spheres are centered in the same point, C_i has radius $1 - \frac{1}{i+1}$, the measure is uniform on each C_i and the measure of $\cup_{j \geqslant i} C_i$ is $4 \cdot 2^{-2^i}$. The analysis holds *mutatis mutandis*.

As another example, which does not require new calculations, we can combine ν with the uniform distribution on the edges of a regular $(d-2)$-dimensional simplex in the $(d-2)$-dimensional subspace orthogonal to the plane that contains the circles, mixing the two distributions in the ratio 50 : 50.

In all our constructions, the measure is concentrated on lower-dimensional manifolds of $\mathbb{R}^d$, circles, spheres, or line segments. If a continuous distribution is desired, one can replace each circle in the plane by a narrow annulus and each sphere by a thin spherical shell, without changing the characteristic behaviour.

5. An alternative treatment of atomic measures

Even for measures with atoms, one can give a precise meaning to Efron's formula: The expression in (1.1) counts the expected number of convex hull vertices of P_n that are unique in the sample X_n. From this, it is obvious that Efron's formula (1.2) is a lower bound on $\mathbb{E}[f_0(P_n)]$ (1.3).

For dealing with atomic measures, there is an alternative possibility. The resulting statements involve different quantities than our original results, but they have the advantage of holding for every measure. We denote by $\bar{f}_0(X_n)$ the number of points of the sample X_n that lie *on the boundary* of their convex hull P_n, counted with multiplicity in case of coincident points. We denote by $\breve{P}_n$ the interior of P_n. Then a derivation analogous to (1.1–1.2) leads to the following variation of Efron's formula:

$$\mathbb{E}[\bar{f}_0(X_n)] = n(1 - \mathbb{E}[\mu(\breve{P}_{n-1})]) \tag{5.1}$$

We emphasize that we mean the boundary and interior with respect to the ambient space $\mathbb{R}^d$, not the *relative* boundary or interior.

Even for some non-atomic measures, this gives different results. Consider the uniform distribution on the boundary of an equilateral triangle. Then $\mathbb{E}[\bar{f}_0(X_n)] = n$, while $\mathbb{E}[f_0(P_n)] \leqslant 6$. Accordingly, $\mathbb{E}[\mu(\breve{P}_n)] = 0$, while $\mathbb{E}[\mu(P_n)]$ converges to 1.

We denote the closure of the wet part W_t^μ by $\bar{W}_t^\mu$ and its measure by $\bar{w}^\mu(t) := \mu(\bar{W}_t^\mu)$.

With these concepts, we can prove the following analogs of Theorems 1.2–1.4. Observe that for a measure μ for which for every hyperplane H, $\mu(H) = 0$ the content of this theorem is the same as the previous ones.

THEOREM 5.1. — (i) *For any probability measure μ in $\mathbb{R}^d$ and $n \geqslant 2$,*

$$\frac{1}{4}\bar{w}^{\mu}(\tfrac{1}{n}) \leqslant \mathbb{E}[1-\mu(\breve{P}_n^{\mu})] \leqslant \bar{w}^{\mu}\Big((d+2)\tfrac{\ln n}{n}\Big) + \tfrac{\varepsilon_d(n)}{n}, \tag{5.2}$$

and

$$\frac{1}{e}n\bar{w}^{\mu}(\tfrac{1}{n}) \leqslant \mathbb{E}\left[\bar{f}_0(X_n^{\mu})\right] \leqslant n\bar{w}^{\mu}\Big((d+2)\tfrac{\ln n}{n}\Big) + \varepsilon_d(n), \tag{5.3}$$

where $\varepsilon_d(n) \to 0$ as $n \to +\infty$ and is independent of μ.

(ii) *There is a non-atomic probability measure ν on $\mathbb{R}^2$ such that*

$$\mathbb{E}[1-\mu(\breve{P}_n^{\nu})] > \tfrac{1}{2}\cdot\bar{w}^{\nu}(\log_2 n/n)$$

and

$$\mathbb{E}\left[\bar{f}_0(X_n^{\nu})\right] > \tfrac{1}{2}n\cdot\bar{w}^{\nu}(\log_2 n/n)$$

for infinitely many n.

Proof sketch. — Since the derivation is parallel to the proofs in Sections 3–4, we only sketch a few crucial points.

(i) For proving the lower bound in (5.2), we modify the initial argument leading to (3.1): For every fixed $x \in W_t$, there is a *closed* half-space h with $x \in h$ whose corresponding open halfspace $\breve{h}$ has measure $\mu(\breve{h}) \leqslant t$. Therefore,

$$\Pr[x \notin \breve{P}_n] \geqslant \Pr[h \cap \breve{P}_n = \emptyset] = \Pr[\breve{h} \cap P_n = \emptyset] = (1-\mu(\breve{h}))^n \geqslant (1-t)^n.$$

The remainder of the proof can be adapted in a straightforward way.

In Section 3.2, we have established that for an ε-net S, its convex hull P contains $\mathbb{R}^d \setminus W_\varepsilon$. Since the interior operator is monotone, this implies that $\mathbb{R}^d \setminus \bar{W}_\varepsilon \subseteq \breve{P}$. Therefore, the ε-net argument of Section 3.3 applies to the modified setting and establishes the upper bound in (5.2).

Finally, by Efron's modified formula (5.1), the result (5.2) carries over to (5.3) as in our original derivation.

(ii) The lower-bound construction of Theorem 1.3 gives zero measure to every hyperplane, and therefore all quantities in part (ii) are equal to the corresponding quantites in Theorem 1.3 and Theorem 1.4(ii). □

Acknowledgements.

I. B. was supported by the Hungarian National Research, Development and Innovation Office NKFIH Grants K 111827 and K 116769, and by the Bézout Labex (ANR-10-LABX-58). X. G. was supported by Institut Universitaire de France. The authors are grateful for the hospitality during the ASPAG (ANR-17-CE40-0017) workshop on geometry, probability, and algorithms in Arcachon in April 2018.

BIBLIOGRAPHY

[1] I. Bárány, Intrinsic volumes and f-vectors of random polytopes, *Mathematische Annalen* **285** (1989), 671–699. ↑4

[2] I. Bárány, D. G. Larman, Convex bodies, economic cap coverings, random polytopes, *Mathematika* **35** (1988), 274–291. ↑2, 4

[3] I. Bárány, L. Dalla, Few points to generate a random polytope, *Mathematika* **44** (1997), 325–331. ↑8

[4] M. Beermann, *Random polytopes.* Ph. D. thesis, University of Osnabrück, 2015. Available at `https://repositorium.ub.uni-osnabrueck.de/bitstream/urn:nbn:de:gbv:700-2015062313276/1/thesis_beermann.pdf`. ↑6

[5] M. Beermann, M. Reitzner, *Monotonicity of functionals of random polytopes.* Preprint arXiv:1706.08342, (2017). ↑6

[6] O. Devillers, M. Glisse, X. Goaoc, G. Moroz, M. Reitzner, The monotonicity of f-vectors of random polytopes. *Electron. Commun. Probab.* **18** (2013), 1–8. ↑6

[7] B. Efron, The convex hull of a random set of points. *Biometrika* **52** (1965), 331–343. ↑3

[8] E. F. Harding, The number of partitions of a set of n points in k dimensions induced by hyperplanes. *Proceedings of the Edinburgh Mathematical Society* **15(4)** (1967), 285–289. ↑8

[9] D. Haussler E. Welzl, ε-Nets and simplex range queries, *Discrete & Computational Geometry* **2** (1987), 127–151. ↑8

[10] Z. Kabluchko, C. Thäle, D. Zaporozhets, Beta polytopes and Poisson polyhedra: f-vectors and angles, arXiv:1805.01338v2, (2019). ↑6

[11] J. Komlós, J. Pach G. Woeginger, Almost tight bounds for ε-nets, *Discrete Comput. Geom.* **7** (1992), 163–173. ↑8, 14

[12] J. Matoušek, *Lectures on Discrete Geometry.* Springer (2002). ↑7

[13] J. Pach, P. K. Agarwal, *Combinatorial Geometry.* John Wiley and Sons (1995). ↑8, 14

[14] V. N. Vapnik, A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory of Probability and its Applications* **2** (1971), 264–280. ↑8

[15] V. H. Vu, Sharp concentration of random polytopes, *Geom. Funct. Anal.* **15** (2005), 1284–1318. ↑3, 8

Appendix A. Proof of Lemma 3.1

For completeness, this appendix, which does not appear in the published version in the Annales Henri Lebesgue, includes a slightly expanded proof of Lemma 3.1 from [11] and [13, pp. 249–251].

We denote X_s by X, and we extend it by adding a random sample Y of size $N-n$. All elements are drawn independently (with replacement) and the result is a sample $Z = X \cup Y$ of size N. We consider all samples as *multisets*, i.e., we abstract from the order but we keep multiplicities.

For a sample S we denote by $\#_h(S)$ the number of elements in h, counted with multiplicity. For a range h, the random variable $\#_h(S)$ is binomially distributed with success probability $p = \mu(h)$.

We denote by m_h the median of this distribution for $S = Y$:

$$\Pr[\#_h(Y) > m_h] \leqslant \tfrac{1}{2} \leqslant \Pr[\#_h(Y) \geqslant m_h]$$

The median of a binomial distribution differs from the mean by less than 1 [A2, A1]. Thus:

$$|m_h - (N-n)\mu(h)| < 1 \tag{A.1}$$

We denote by $\mathcal{H}_\varepsilon := \{\, h \in \mathcal{H} \mid \mu(h) \geqslant \varepsilon \,\}$ the set of "large enough" ranges.

We introduce the abbreviation

$$\mathrm{skew}_h(X, Y) \iff \#_h(X) = 0 \wedge \#_h(Y) \geqslant m_h,$$

characterizing a very unbalanced distribution of the elements of $Z \cap h$ over X and Y. We are interested in the event

$$\exists h \in \mathcal{H}_\varepsilon \colon \mathrm{skew}_h(X, Y).$$

The independence of X and Y implies following inequality about the probability of this event:

$$\begin{aligned} &\Pr[\exists h \in \mathcal{H}_\varepsilon \colon \mathrm{skew}_h(X, Y)] \\ = &\Pr[\exists h \in \mathcal{H}_\varepsilon \colon \#_h(X) = 0 \wedge \#_h(Y) \geqslant m_h] \\ \geqslant &\Pr[\exists h \in \mathcal{H}_\varepsilon \colon \#_h(X) = 0] \times \min_{h \in \mathcal{H}_\varepsilon} \Pr[\#_h(Y) \geqslant m_h] \end{aligned}$$

This is justified as follows: Suppose that the first event on the last line happens, and some range $h \in \mathcal{H}_\varepsilon$ with $\#_h(X) = 0$ exists. If *that particular* h also fulfills the second condition $\#_h(Y) \geqslant m_h$, then the event on the first line happens; and the probability for the second condition is a least the min-expression on the right side.

Since this minimum is at least $\frac{1}{2}$, we obtain

$$\Pr[\exists h \in \mathcal{H}_\varepsilon \colon \#_h(X) = 0] \leqslant 2 \cdot \Pr[\exists h \in \mathcal{H}_\varepsilon \colon \mathrm{skew}_h(X, Y)]. \tag{A.2}$$

The left side describes the event that X is not an ε-net, which is what we are interested in.

Now we will fix the multiset Z and consider all possibilities how it can arise as the union of two samples X and Y. For every h, we have

$$\Pr[\mathrm{skew}_h(X, Y) \mid Z\,] = \begin{cases} 0, & \text{if } \#_h(Z) < m_h \\ \dfrac{\binom{N-n}{m}}{\binom{N}{m}}, & \text{if } \#_h(Z) = m \geqslant m_h \end{cases}$$

If $h \in \mathcal{H}_\varepsilon$, the expression for the second case is bounded by

$$(\tfrac{N-n}{N})^m \leqslant (1 - \tfrac{n}{N})^{m_h} \leqslant (1 - \tfrac{n}{N})^{(N-n)\mu(h)-1} \leqslant (1 - \tfrac{n}{N})^{(N-n)\varepsilon-1},$$

where we have used (A.1) to bound m_h. This bound is trivially also valid for the first case.

We now want to take the union over all ranges $h \in \mathcal{H}_\varepsilon$. If two ranges $h, h' \in \mathcal{H}$ have equal intersections with Z: $h \cap Z = h' \cap Z$, then the events $\mathrm{skew}_h(X, Y)$ and $\mathrm{skew}_{h'}(X, Y)$ are identical, and we can ignore such duplications. The multiset Z is intersected by the ranges $h \in \mathcal{H}$ in at most $\pi_{\mathcal{H}}(N)$ different ways. Thus, the union bound leads to the following estimate:

$$\Pr[\exists h \in \mathcal{H}_\varepsilon \colon \mathrm{skew}_h(X, Y) \mid Z\,] \leqslant \pi_{\mathcal{H}}(N)(1 - \tfrac{n}{N})^{(N-n)\varepsilon-1}$$

Since this bound holds for every multiset Z, it also holds for a random Z, and with (A.2), we obtain the claimed bound of the lemma:

$$\Pr[X \text{ is not an } \varepsilon\text{-net}] \leqslant 2\pi_{\mathcal{H}}(N) \cdot (1 - \tfrac{n}{N})^{(N-n)\varepsilon - 1} \qquad \square$$

BIBLIOGRAPHY

[A1] K. Hamza, The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions. *Statistics & Probability Letters*, **23** (1995), 21–25. doi:10.1016/0167-7152(94)00090-U. ↑15

[A2] R. Kaas and J. M. Buhrman, Mean, median and mode in binomial distributions, *Statistica Neerlandica*, **34** (1980), 13–18. ↑15

Manuscript received on 23rd February 2019,
revised on 1st January 2000,
accepted on 1st January 2000.

Recommended by Editor H. Lebesgue.

Imre BÁRÁNY
Rényi Institute of Mathematics, Hungarian Academy of Sciences, PO Box 127, 1364 Budapest, and Department of Mathematics University College London, Gower Street, London WC1E 6BT, England
barany@renyi.hu

Matthieu FRADELIZI
Université Paris-Est, LAMA (UMR 8050), UPEM, UPEC, CNRS, F-77454, Marne-la-Vallée, France
matthieu.fradelizi@u-pem.fr

Xavier GOAOC
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
xavier.goaoc@loria.fr

Alfredo HUBARD
Université Paris-Est, Marne-la-Vallée, Laboratoire d'Informatique Gaspard Monge, 5 Boulevard Descartes, 77420 Champs-sur-Marne, France
alfredo.hubard@u-pem.fr

Günter ROTE
Freie Universität Berlin, Institut für Informatik, Takustraße 9, 14195 Berlin
rote@inf.fu-berlin.de