An OSSS-type inequality for uniformly drawn subsets of fixed size

Jacob van den Berg[†]
[†]CWI and VU University Amsterdam ; email: J.van.den.Berg@cwi.nl

Abstract

The OSSS inequality (O'Donnell, Saks, Schramm and Servedio [16]) gives an upper bound for the variance of a function f of independent 0-1 valued random variables in terms of the influences of these random variables and the computational complexity of a (randomised) algorithm for determining the value of f.

Duminil-Copin, Raoufi and Tassion [6] obtained a generalization to monotonic measures and used it to prove new results for Potts models and random-cluster models. Their generalization of the OSSS inequality raises the question if there are still other measures for which a version of that inequality holds.

We derive a version of the OSSS inequality for a family of measures that are far from monotonic, namely the k-out-of-n measures (these measures correspond with drawing k elements from a set of size n uniformly). We illustrate the inequality by studying the event that there is an occupied horizontal crossing of an $R \times R$ box on the triangular lattice in the site percolation model where exactly half of the vertices in the box are occupied.

Key words and phrases: OSSS inequality, percolation, randomized algorithm.

1 Introduction

1.1 Background and main results of the paper

The OSSS inequality [16] gives an upper bound for the variance of a function f of independent 0-1 valued random variables in terms of their influences on f and their revealment probabilities with respect to an algorithm to determine the value of f. Since the break-through work by Duminil-Copin, Raoufi and Tassion around 2017 it has become one of the main tools to prove sharp phase transition in a number of important models from statistical mechanics. In particular, in [5] they applied the OSSS inequality to Voronoi percolation on $\mathbb{R}^d$, and in [6] they extended the inequality to monotonic measures (measures satisfying the FKG lattice condition) and used that extension to prove sharp phase transition for random-cluster and Potts models. That version of OSSS was further extended, but still under the condition that the measure is monotonic, by Hutchcroft [11] (who developed a two-function form and used it to obtain new critical exponent inequalities) and by Dereudre and Houdebert [4] (who gave a generalization in a continuum setting and used it to prove sharp phase transition for the Widom-Rowlinson model).

The above raises the natural question whether the OSSS inequality can be extended beyond the class of monotonic measures. Our main result, Theorem 1.1 below, is a version of the OSSS inequality for a family of measures which are clearly non-monotonic, namely k-out-of-n measures.

Before stating the theorem, we will introduce the necessary definitions and notation. Let E be a finite set and let $k \leq |E|$, where $|E|$ denotes the size (i.e. cardinality) of $|E|$. If $\omega \in \{0,1\}^E$, we will use the notation $|\omega|$ for $\sum_{e \in E} \omega_e$.
Informally, the k-out-of-E distribution (notation $P_{k,E}$) is the uniform distribution on the set of all subsets of E of size k. Formally, we define $P_{k,E}$ as the following distribution on $\{0,1\}^E$: For each $\omega \in \{0,1\}^E$,

$$P_{k,E}(\omega) = \begin{cases} 0, & \text{if } |\omega| \neq k \\ \frac{1}{\binom{|E|}{k}}, & \text{if } |\omega| = k, \end{cases} \tag{1}$$

Often, when the set E is clear from the context, or only its size $|E|$ matters, we will simply write $P_{k,n}$ instead of $P_{k,E}$, with $n = |E|$ (and call it a k-out-of-n distribution).

An *increasing event* is a set $A \subset \{0,1\}^E$ with the property that, for all pairs ω, σ with $\omega \in A$ and $\sigma \in \{0,1\}^E$, $\omega \leq \sigma$ implies that $\sigma \in A$. Here, as usual, $\omega \leq \sigma$ means that $\omega_e \leq \sigma_e$ for all $e \in E$.

Let $\omega \in \{0,1\}^E$ and $A \subset \{0,1\}^E$. An element $e \in E$ is said to be *pivotal* (w.r.t. ω and A) if exactly one of ω and $\omega^{(e)}$ is in A. Here $\omega^{(e)}$ denotes the element of $\{0,1\}^E$ obtained from ω by replacing ω_e by $1 - \omega_e$. We say that e is 1-pivotal (respectively 0-pivotal) if e is pivotal and $\omega_e = 1$ (respectively 0). The probability (w.r.t. the measure $P_{k,E}$) that e is pivotal, also called the (absolute) influence of e (on the event A), will be denoted by $I^A_{k,E}(e)$. In cases where the event A is clear from the context, we simply write $I_{k,E}(e)$.

In addition, we will use the notation $I_{k,E}^{A,1}(e)$, and $I_{k,E}^{A,0}(e)$, for the probability that e is 1- pivotal, and the probability that e is 0-pivotal respectively.

Somewhat informally, a *decision tree* is (in the present context) an algorithm to check whether an 'input string' ω is in A, by 'examining' one by one the elements of E (i.e. querying their ω-value ω_e), where at each step the next element of E to be examined depends on the already examined elements of E and their ω-values. Note that if, after some step, the up to then revealed ω-values already determine whether the input string is in A or not, the algorithm may stop. That step wil be denoted by τ.

We will use the *formal* definition of a decision tree used in [6]: A decision tree is a pair $T = (e_1, \phi)$, with $e_1 \in E$ and ϕ of the form $((\phi_2, \cdots, \phi_n)$, where each ϕ_t is a function which assigns, to each pair $\big((e_1, \cdots, e_{t-1}),\ \omega_{(e_1,\cdots,e_{t-1})}\big)$ an element of $E \setminus \{e_1, \cdots, e_{t-1}\}$.

The corresponding algorithm is then as follows: First e_1 is examined, i.e. its value ω_{e_1} is queried (and revealed). Depending on that value the next element of E, namely the element $e_2 = \phi_2(e_1, \omega_{e_1})$ is selected and examined. After the value ω_{e_2} of e_2 has been revealed, the next element, namely $e_3 = \phi_3((e_1, e_2), (\omega_{e_1}, \omega_{e_2})$ is examined etcetera.

The time τ mentioned before is then formally defined as

$$\tau(\omega) := \min\{t \geq 1 \ : \ \text{for all } \omega' \in \{0,1\}^E \text{ with } \omega'_{e_{[t]}} = \omega_{e_{[t]}},\ I_A(\omega') = I_A(\omega)\}. \tag{2}$$

Note that τ depends on A and T although, this is not visible in the notation.

If the string ω is generated according to a probability distribution $\mathbb{P}$ under which some elements of $\{0,1\}^E$ have probability 0 (which is the case with the k-out-of-n measures studied in the current paper), we slightly change the above definition of τ by replacing 'all' by '$\mathbb{P}$-almost all'.

One could say that a given decision tree is 'efficient' (w.r.t. the probability distribution on the set of input strings) if 'typically', or 'on average', the number τ is 'small'.

In this paper we focus on the case where $\mathbb{P} = P_{k,E}$, and we will denote the probability that the value at a given 'site' $e \in E$ is revealed (for the decision tree T) by $\delta_e = \delta_e(A, T)$. So, formally,

$$\delta_e(A, T) = P_{k,E}(e_t = e \text{ for some } t \leq \tau(\omega)).$$

Our main result, Theorem 1.1 below, is a version of the OSSS inequality for k-out-of-n measures.

Theorem 1.1. *Let E be a finite set with $|E| \geq 2$, and let $A \subset \{0,1\}^E$ be an increasing event. For every $k \leq |E|$ and every decision tree T,*

$$\begin{aligned} & P_{k,E}(A)(1 - P_{k,E}(A)) \\ \leq & \sum_{e \in E} I_{k,E}^{A,0}(e) \; \delta_e(A,T) \;+\; \log(|E|) \sum_{e \in E} I_{k,E}^{A,0}(e) \; \frac{1}{|E|-1} \sum_{e' \in E} \delta_{e'}(A,T). \end{aligned} \tag{3}$$

Observing that $1/(|E|-1) \leq 2/|E|$ if $|E| \geq 2$, it is clear that the following, more compact form of the upper bound for $P_{k,E}(A)(1 - P_{k,E}(A))$ also holds (and is at most a constant factor 2 larger than the upper bound in (3)).

Corollary 1.2. *With E, A, k and T as in Theorem 1.1,*

$$\begin{aligned} & P_{k,E}(A)(1 - P_{k,E}(A)) \\ \leq & \sum_{e \in E} I_{k,E}^{A,0}(e) \; \delta_e(A,T) \;+\; 2\log(|E|) \sum_{e \in E} I_{k,E}^{A,0}(e) \; \bar{\delta}(A,T), \end{aligned} \tag{4}$$

where $\bar{\delta}(A,T) = \frac{1}{|E|} \sum_{e \in E} \delta_e(A,T)$, the 'average revealment per vertex'.

We also obtained the following version of Theorem 1.1, where, roughly speaking, the factor $\log(|E|)$ has been replaced by $\frac{1}{\varepsilon}$, and where an extra term of order $P_{k,E}(\tau \geq (1-\varepsilon)n)$ has been added.

Theorem 1.3. *Let E be a finite set and let $A \subset \{0,1\}^E$ be an increasing event. Let $n = |E|$ and let $\varepsilon \in (0,1)$. For every $k \leq |E|$ and every decision tree T, we have*

$$\begin{aligned} & P_{k,E}(A)(1 - P_{k,E}(A)) \\ \leq & \sum_{e \in E} I_{k,E}^{A,0}(e) \; \delta_e(A,T) \;+\; \frac{1}{\varepsilon} \sum_{e \in E} I_{k,E}^{A,0}(e) \; \bar{\delta}(A,T) \\ & + \frac{1}{2} P_{k,E}(\tau \geq (1-\varepsilon)n), \end{aligned} \tag{5}$$

with $\bar{\delta}(A,T)$ as in (4).

Remark 1.4. *(a) As mentioned before, the OSSS inequality in [6] gives an upper bound for the variance of a function. In many applications in that and other papers, the corresponding function is the indicator function of an event. In our theorem we restrict to such functions.*

(b) The term with the logarithmic factor in (3) *is not present in the OSSS inequality for monotonic measures in [6]. This logarithmic factor seems to be quite harmless for many applications, for instance those in Section 5 of our paper. However, since it may disturb other potential applications, we also stated Theorem 1.3 which does not have the logarithmic factor. It does have an extra term* $P_{k,E}(\tau \geq (1-\varepsilon)n)$. *In most percolation applications concerning connection events in a large box (with, say,* n *vertices), that term is (if* ε *is sufficiently small) typically very small for large* n. *In such applications the decision tree usually corresponds with a so-called exploration procedure (or path), where roughly speaking an occupied cluster or its boundary is explored step-by-step. Sites in the interior of the box that are vacant and whose neigbours are also vacant, are typically not inspected by such a procedure. By standard arguments one can easily see that then (if* k/n *is bounded away from* 1, *which is a natural condition in percolation studies, and* ϵ *is sufficiently small, depending on* k/n*) the probability (under the measure* $P_{k,E}$*) of the event that there are less than* ϵn *such vertices (and hence the extra term in Theorem 1.3) tends to* 0 *exponentially fast as* n *tends to* ∞.

In Section 5 we illustrate Theorem 1.1 by studying the event that there is a horizontal crossing of an $R \times R$ box on the triangular lattice in the site percolation model where exactly k of the $n := R^2$ vertices are occupied. We show there that the expected number of pivotal sites (and, consequently, the 'discrete derivative' of the crossing probability with respect to the fraction of occupied sites) at the value $k = R^2/2$ is larger than some positive power of R, see Theorem 5.2. The proof uses, besides Theorem 1.1, only a minimum of preliminaries from Bernoulli percolation (i.e. the usual percolation model where the states of the vertices are independent of each other). It can be proved without using Theorem 1.1, but I don't know a proof which neither uses Theorem 1.1, nor quite heavy results from Bernoulli percolation; see the comments and discussion in Section 5.2.

1.2 Other related work

The OSSS inequality, seen from right to left, gives a *lower* bound for the expectation of the number of pivotals, and that is how it has been successfully used in percolation theory and related fields, for instance in [6] and other papers mentioned above. A different well-known inequality for product measures, which does not involve decision trees but also provides a lower bound for that expectation, is the KKL inequality [12] (and a related inequality by Talagrand)[20]).

The KKL inequality (or, rather, a consequence of it) and Talagrand's inequality say, roughly speaking, that, for product measures, the expected number of pivotals for the event A is at least some constant times the probability of A, times (1 minus the probability of A) times $\log(1/\mathcal{M})$, where $\mathcal{M}$ is the maximum over all e of the probability that e is pivotal. In situations where no suitable decision tree exists (or is known) but where some upper bound on the quantity $\mathcal{M}$ is known, the OSSS inequality is often useless while KKL (and Talagrand's) inequalities still give a useful result. On the other hand, in specific situations, where some suitable decision tree does exist, OSSS can be substantially stronger than KKL and Talagrand's inequality.

While the proof of KKL has an analytic/algebraic flavour (Fourier expansion, hypercontractivity), the OSSS inequality is, essentially, proved from suitable coupling arguments which are more 'probabilistic' in nature.

The KKL inequality was generalised to k-out-of-n measures (and similar measures on state spaces with a larger 'alpahbet', e.g. $\{0, 1, 2\}^n$) by O'Donnell and Wimmer [17] (see also [7]).

Finally, we remark here that the paper [1] extends yet another inequality from product measures to k-out-of-n measures. However, that inequality (and its proof) are very different in nature from OSSS.

1.3 Organization of the paper

In Section 2 we give ingredients for the proof of our main result, Theorem 1.1: general ingredients (mainly from the first part of Section 2 of [6]) in Section 2.1, and specific ingredients for k-out-of-n measures in Section 2.2. In Section 3 we will then, usung these ingredients, present the proof of Theorem 1.1. Most of the proof of Theorem 1.3 is the same as that of Theorem 1.1. The differences will be stated and explained in Section 4.

In Section 5 we illustrate Theorem 1.1 by applying it to box-crossing probabilities for a percolation model on a box in the triangular lattice where a fixed number of vertices is occupied: see Theorem 5.2 in that section. Section 5.2 gives several remarks concerning, among other things, other (potential) ways to prove Theorem 5.2.

2 Ingredients for the proof of the main theorem

2.1 General ingredients

In this subsection we present some results which hold for all distributions, not only for k-out-of-n. These results come mainly from [6]. For convenience we use the same notation as in that paper.
The following lemma in [6] about 'encoding' a random string with an arbitrary distribution in terms of independent random variables uniformly distributed on $[0,1]$, is intuitively appealing and has been used implicitly and in different contexts in the literature before. We follow quite closely its formulation in Lemma 2.1 of [6]. (The lemma is stated there only for monotonic measures, although it is true for general probability measures).

Let, as before, E be a finite set and let $n = |E|$. Let μ be a probability measure on $\{0,1\}^E$. For $u \in [0,1]^n$ and $e = (e_1, \cdots, e_n)$ a permutation of E, define the element $F_e(u) = x \in \{0,1\}^n$ as follows:

$$x_{e_1} := \begin{cases} 1, & \text{if } u_1 \geq \mu(\omega_{e_1} = 0), \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

and, inductively, for $2 \leq t \leq n$,

$$x_{e_t} := \begin{cases} 1, & \text{if } u_t \geq \mu[\omega_{e_t} = 0 \,|\, \omega_{e_{[t-1]}} = x_{e_{[t-1]}}], \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We will often write $F_e^\mu(u)$, instead of just $F_e(u)$, to emphasize the dependence on μ.

Lemma 2.1. [[6], Lemma 2.1] *Let $\mathbf{U} = (\mathbf{U}_1, \cdots, \mathbf{U}_n)$ be a sequence of independent uniformly on $[0,1]$ distributed random variables, and let $\mathbf{e} = (\mathbf{e}_1, \cdots, \mathbf{e}_n)$ be a random (not necessarily uniform) permutation of the elements of E. If, for every $1 \leq t \leq n$, $\mathbf{U}_t$ is independent of $(\mathbf{e}_{[t]}, \mathbf{U}_{[t-1]})$, then $F_{\mathbf{e}}^\mu(\mathbf{U})$ has distribution μ.*

Let, as in Theorem 1.1, $A \subset \{0,1\}^n$ be an increasing event and let $\mathbb{I}_A$ denote its indicator function.
We will use the construction in the first part of Section 2 of [6] of a sequence of suitably coupled random strings $\mathbf{Y}^0, \cdots, \mathbf{Y}^n$, each with distribution μ, with the property that each pair of consecutive strings is 'highly correlated' and where $\mathbb{I}_A(\mathbf{Y}^0)$ and $\mathbb{I}_A(\mathbf{Y}^n)$ are independent. To make our paper self-contained, we describe their construction here in detail. Again we remark

that although [6] presents the construction for monotonic measures μ, the same construction works for general μ.

Let $\mathbf{U} = (\mathbf{U}_1, \cdots, \mathbf{U}_n)$ and $\mathbf{V} = (\mathbf{V}_1, \cdots, \mathbf{V}_n)$ be independent strings of iid random variables, all uniformly distributed on the interval $[0, 1]$. The idea is, loosely speaking, to define first, using the $\mathbf{U}_i$'s only, a triple $\mathbf{e}$, $\mathbf{X}$, τ, corresponding to the decision tree in the statement of Theorem 1.1. Then, with the same $\mathbf{e}$, but with some $\mathbf{U}_i$'s suitably (depending on the value of τ) replaced by $\mathbf{V}_i$'s, the strings $\mathbf{Y}^k$, $k = 0, \cdots, n$ are defined in a way similar to (6) and (7). Here are the details, with the same notation as in [6].

Define inductively, for $t \geq 1$, (with ϕ and e_1 as in the definition of a decision tree T in the paragraphs preceding Theorem 1.1),

$$\mathbf{e}_t = \begin{cases} e_1 & \text{if } t = 1, \\ \phi_t(\mathbf{e}_{[t-1]}, \mathbf{X}_{\mathbf{e}_{[t-1]}}) & \text{if } t > 1, \end{cases} \tag{8}$$

and

$$\mathbf{X}_{\mathbf{e}_t} = \begin{cases} 1 & \text{if } \mathbf{U}_t \geq \mu(\omega_{\mathbf{e}_t} = 0 \,|\, \omega_{\mathbf{e}_{[t-1]}} = \mathbf{X}_{\mathbf{e}_{[t-1]}}, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Further define (equivalently to the stopping time in (2)),

$$\tau := \min\{t \geq 1 \,:\, \forall x \in \{0,1\}^E \text{ with } \mu(x) > 0 \text{ and } x_{\mathbf{e}_{[t]}} = \mathbf{X}_{\mathbf{e}_{[t]}},\, I_A(x) = I_A(\mathbf{X})\}. \tag{10}$$

The strings $\mathbf{Y}^t$, $0 \leq t \leq n$, are then definined by

$$\mathbf{Y}^t = F^{\mu}_{\mathbf{e}}(\mathbf{W}^t), \text{ where} \tag{11}$$

$$\mathbf{W}^t := (\mathbf{V}_1, \cdots, \mathbf{V}_t, \mathbf{U}_{t+1}, \cdots, \mathbf{U}_\tau, \mathbf{V}_{\tau+1}, \cdots, \mathbf{V}_n). \tag{12}$$

Remark 2.2. *For all clarity we remark, as done in [6], that in particular* $\mathbf{W}^t = \mathbf{V}$ *if* $t \geq \tau$.

We still follow (a straightforward suitable adaptation of) the first part of Section 2 of [6]: From its definition it is clear that $\mathbf{X}$ is $\mathbf{U}$-measurable. By Lemma 2.1, it is easy to see that $\mathbf{X}$, and each $\mathbf{Y}^t$, has distribution μ. Note that $\mathbf{Y}^n$ is not in general independent of $\mathbf{Y}^0$. Also note that, although $\mathbf{W}^n = \mathbf{V}$, $\mathbf{Y}^n$ is not necessarily $\mathbf{V}$-measurable (because via $\mathbf{e}$ it involves $\mathbf{U}$). However, Lemma 2.1 says that, no matter what the outcome of $\mathbf{e}$ is, the conditional distribution of $\mathbf{Y}^n$ is always μ. Hence $\mathbf{X}$ and $\mathbf{Y}^n$ are independent, so $2\mu(A)(1 - \mu(A)) = \mathbb{E}[|\mathbb{I}_A(\mathbf{X}) - \mathbb{I}_A(\mathbf{Y}^n)|]$. Hence, since (by the definition of τ) $\mathbb{I}_A(\mathbf{X}) = \mathbb{I}_A(\mathbf{Y}^0)$, it holds that (in spite of the fact that, as mentioned above, $\mathbf{Y}^0$ and $\mathbf{Y}^n$ are not necessarily independent)

$$\begin{aligned} 2\mu(A)(1-\mu(A)) &= \mathbb{E}[|\mathbb{I}_A(\mathbf{Y}^0) - \mathbb{I}_A(\mathbf{Y}^n)|] \\ &\leq \sum_{t=1}^{n} \mathbb{E}[|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})|] \\ &= \sum_{t=1}^{n} \mathbb{E}[|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})| \, \mathbb{I}_{t \leq \tau}]. \end{aligned} \tag{13}$$

The above equation will be the starting point of the proof of Theorem 1.1 in Section 3. We emphasize again that the equation is very general: it followed immediately from the generally valid coupling construction from Section 2 of [6] and is, essentially, the variance upper bound on the bottom of page 84 in that paper (recall that the function f there is, in our case, an indicator function). To really make use of that equation for our situation we will first, in the next subsection, develop specific ingredients for k-out-of-n measures. The key result presented in that subsection is Proposition 2.6, the proof of which uses Lemma 2.5.

2.2 Specific ingredients for k-out-of-n measures

In this subsection we introduce some results which are quite specific for k-out-of-n measures, and which will be used in Section 3 to prove Theorem 1.1. We use the notation and terminology (adopted from [6]) of the previous subsection. In particular, F_e^{μ} was defined in the paragraph containing (6) and (7).

Lemma 2.3. *Let $1 \leq k \leq n-1$. Let E be a finite set with $|E| = n$, and let $e \in E$. Let $\mathbf{U} = (\mathbf{U_1}, \cdots, \mathbf{U_n})$ be i.i.d. random variables, uniformly distributed on $[0,1]$. Let $\mathbf{e} = (\mathbf{e}_1, \cdots, \mathbf{e}_n)$ be a random (not necessarily uniformly drawn) permutation of the elements of E, with the following properties: $\mathbf{e}_1 = e$ and, for each $t \geq 2$, $\mathbf{e}_t$ is $\mathbf{U}_{[t-1]}$-measurable.*

Let $\mathbf{Z} = F_{\mathbf{e}}^{P_{k,E}}(\mathbf{U})$ and $\mathbf{Z}' = F_{\mathbf{e}}^{P_{k+1,E}}(\mathbf{U})$.
Then the pair $(\mathbf{Z}, \mathbf{Z}')$ has the following distribution:

$$\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \begin{cases} \frac{1}{n-k} P_{k,E}(\alpha), & \text{if } |\alpha| = k, |\beta| = k+1 \text{ and } \beta \geq \alpha \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Remark: *Informally, the lemma says that, given that* $\mathbf{Z} = \alpha$, $\mathbf{Z}'$ *is 'obtained' by randomly and uniformly choosing one of the* 0*'s in the string* α *and replacing it by* 1*. So the pair* $(\mathbf{Z}, \mathbf{Z}')$ *represents the intuitively simplest monotone coupling of* $P_{k,E}$ *and* $P_{k+1,E}$*.*

Proof. We give a proof by induction in n. The case $n = 1$ is trivial. Now suppose $n \geq 2$. It is easy to see from the definitions and Lemma 2.1 that $\mathbf{Z}' \geq \mathbf{Z}$ and that $\mathbf{Z}$ has distribution $P_{k,E}$ and $\mathbf{Z}'$ has distribution $P_{k+1,E}$. So we assume that α and β satisfy the conditions in the first line of (14). We distinguish three cases: (i): $\alpha_e = \beta_e = 1$. (ii): $\alpha_e = 0, \beta_e = 1$. (iii): $\alpha_e = \beta_e = 0$.
We will treat case (i) here; the other two cases can be treated analogously. Note that, by the definitions, the event that $\mathbf{Z}_e = \mathbf{Z'}_e = 1$ is $\mathbf{U}_1$-measurable and has probability k/n. Also note that (in case (i)) $\mathbf{Z}_{E \setminus \{e\}} = F^{P_{k-1,E\setminus\{e\}}}_{(\mathbf{e}_2,\cdots,\mathbf{e}_n)}(\mathbf{U}_2, \cdots, \mathbf{U}_n)$ and $\mathbf{Z'}_{E\setminus\{e\}} = F^{P_{k,E\setminus\{e\}}}_{(\mathbf{e}_2,\cdots,\mathbf{e}_n)}(\mathbf{U}_2, \cdots, \mathbf{U}_n)$. Hence the induction hypothesis gives that

$$\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \frac{k}{n} \frac{1}{(n-1)-(k-1)} P_{k-1,E\setminus\{e\}}(\alpha_{E\setminus\{e\}}).$$

Further, evidently, (recall that we are in case(i)),

$$P_{k,E}(\alpha) = \frac{k}{n} P_{k-1,E\setminus\{e\}}(\alpha_{E\setminus\{e\}}).$$

These two equations together immediately give $\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \frac{1}{n-k} P_{k,E}(\alpha)$, as desired. ∎

Lemma 2.4. *Let* E, $\mathbf{U} = (\mathbf{U}_1, \cdots, \mathbf{U}_n)$*, and* $\mathbf{e}$ *be as in Lemma 2.3. Further, let* $\mathbf{V}_1$ *be uniformly distributed on* $[0, 1]$*, independently of* $\mathbf{U}$*. Let* $\mathbf{Z} := F^{P_{k,E}}_{\mathbf{e}}(\mathbf{U})$ *and* $\mathbf{Z}' := F^{P_{k,E}}_{\mathbf{e}}(\mathbf{V}_1, \mathbf{U}_2, \cdots, \mathbf{U}_n)$*.*
Let α *and* β *be two* different *elements of* $\{0, 1\}^E$*. Then*

$$\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \begin{cases} \frac{1}{n} P_{k,E}(\alpha), & \text{if } |\alpha| = |\beta| = k \text{ and } \exists f \neq e \text{ s.t. } \beta = \alpha^{(e,f)} \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

where $\alpha^{(e,f)}$ *is the string obtained from* α *by swopping* α_e *and* α_f*.*

Proof. If $n = 1$, then $\mathbb{P}(\mathbf{Z} \neq \mathbf{Z}') = 0$, so that statement of the lemma obviously holds. So we may assume that $n \geq 2$. It is easy to see from the definitions that if the pair α, β does not satisfy the conditions in the first line of (15), then $\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = 0$. So we assume that $|\alpha| = k$ and that $\beta = \alpha^{(e,f)}$ for some $f \neq e$. Note that then (since $\beta \neq \alpha$) $\alpha_f \neq \alpha_e$. We distinguish two cases: (i): $\alpha_e = 1$ (and hence $\beta_e = 0$, $\alpha_f = 0$ and $\beta_f = 1$). (ii): $\alpha_e = 0$ (and hence $\beta_e = 1$, $\alpha_f = 1$ and $\beta_f = 0$).

Here we treat case (i) (the proof of the other case is similar). In this case we have that $\mathbf{Z}_e = 1$ and $\mathbf{Z'}_e = 0$. The former event is $\mathbf{U}_1$-measurable and has probability k/n, and the latter is $\mathbf{V}_1$-measurable and has probability $(n-k)/n$. Also note that, in case (i), $\mathbf{Z}_{E\setminus\{e\}} = F^{P_{k-1,E\setminus\{e\}}}_{(\mathbf{e}_2,\cdots,\mathbf{e}_n)}(\mathbf{U}_2, \cdots, \mathbf{U}_n)$ and $\mathbf{Z'}_{E\setminus\{e\}} = F^{P_{k,E\setminus\{e\}}}_{(\mathbf{e}_2,\cdots,\mathbf{e}_n)}(\mathbf{U}_2, \cdots, \mathbf{U}_n)$. Finally, note that $\beta_{E\setminus\{e\}} = (\alpha_{E\setminus\{e\}})^{(f)}$ So, by independence and application of Lemma 2.3, we get

$$\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \frac{k}{n}\,\frac{n-k}{n}\,\frac{1}{(n-1)-(k-1)} P_{k-1,E\setminus\{e\}}(\alpha_{E\setminus\{e\}}) = \frac{k}{n^2} P_{k-1,E\setminus\{e\}}(\alpha_{E\setminus\{e\}}).$$

Combined with the obvious equation $P_{k,E}(\alpha) = \frac{k}{n} P_{k-1,E\setminus\{e\}}(\alpha_{E\setminus\{e\}})$, this gives that $\mathbb{P}(\mathbf{Z} = \alpha, \mathbf{Z}' = \beta) = \frac{1}{n} P_{k,E}(\alpha)$, as desired. □

In fact, we need the following, more general form, with a stopping time, of the previous lemma.

Lemma 2.5. *Let* $\mathbf{e}$*,* e *and* $\mathbf{U}$ *be as in Lemma 2.3, and let* σ *be a stopping time w.r.t.* $\mathbf{U}$*. (So, for each* $1 \leq m \leq n$*, the event* $\{\sigma = m\}$ *is* $\mathbf{U}_{[m]}$*-measurable). Further, let* $\mathbf{V} = (\mathbf{V}_1, \cdots, \mathbf{V}_n)$ *be iid random variables uniformly distributed on* $[0, 1]$ *independent of* $\mathbf{U}$*.*
Let $\mathbf{Z} := F^{P_{k,E}}_{\mathbf{e}}(\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_\sigma, \mathbf{V}_{\sigma+1}, \cdots, \mathbf{V}_n)$
and $\mathbf{Z}' := F^{P_{k,E}}_{\mathbf{e}}(\mathbf{V}_1, \mathbf{U}_2, \cdots, \mathbf{U}_\sigma, \mathbf{V}_{\sigma+1}, \cdots, \mathbf{V}_n)$*.*
(In particular, if $\sigma = 1$*, then* $\mathbf{Z} = F^{P_{k,E}}_{\mathbf{e}}(\mathbf{U}_1, \mathbf{V}_2, \cdots, \mathbf{V}_n)$ *and*
$\mathbf{Z}' = F^{P_{k,E}}_{\mathbf{e}}(\mathbf{V}_1, \mathbf{V}_2, \cdots, \mathbf{V}_n)$*).*
Then the pair $(\mathbf{Z}, \mathbf{Z}')$ *satisfies* (15) *for all* $\alpha, \beta \in \{0, 1\}^E$ *with* $\alpha \neq \beta$*.*

Proof. of Lemma 2.5 from Lemma 2.4:
First some terminology: We say that a pair of random strings of length n of 0's and 1's "satisfies (15)" if, for all $\alpha, \beta \in \{0, 1\}^E$ with $\alpha \neq \beta$, the probability that the first string has outcome α and the second string has outcome β, is equal to the r.h.s. of (15).

Let $\tilde{\mathbf{e}}$ be a 'modification' of $\mathbf{e}$ with the following properties: $\tilde{\mathbf{e}}_t = \mathbf{e}_t$ for all $t \leq \sigma + 1$, and $\tilde{\mathbf{e}}_{[\sigma+2,n]}$ is a $\mathbf{U}_{[\sigma]}$-measurable permutation of the elements of $E \setminus \{\mathbf{e}_1, \cdots, \mathbf{e}_{\sigma+1}\}$.
In particular, $\tilde{\mathbf{e}}_1 = e$ and $\tilde{\mathbf{e}}_t$ is $\mathbf{U}_{[t-1]}$-measurable for each $t \geq 2$. Hence, by Lemma 2.4, the pair

$$(F^{P_{k,E}}_{\tilde{\mathbf{e}}}(\mathbf{U}_1, \cdots, \mathbf{U}_n), (F^{P_{k,E}}_{\tilde{\mathbf{e}}}(\mathbf{V}_1, \mathbf{U}_2, \cdots, \mathbf{U}_n)) \tag{16}$$

satisfies (15). Hence, since now $\mathbf{U}_{\sigma+1}, \cdots, \mathbf{U}_n$ are no longer used for determining the order in which edges are chosen, we can replace them in the expression (16) by $\mathbf{V}_{\sigma+1}, \cdots, \mathbf{V}_n$, respectively, without changing the joint distribution of the pair of random strings. In other words, the pair

$$(F_{\tilde{\mathbf{e}}}^{P_{k,E}}(\mathbf{U}_1, \cdots, \mathbf{U}_\sigma, \mathbf{V}_{\sigma+1}, \cdots, \mathbf{V}_n), (F_{\tilde{\mathbf{e}}}^{P_{k,E}}(\mathbf{V}_1, \mathbf{U}_2, \cdots, \mathbf{U}_\sigma, \mathbf{V}_{\sigma+1}, \cdots, \mathbf{V}_n)) \tag{17}$$

has the same distribution as the pair in (16) and hence satisfies (15). However, since this is true for any such modification $\tilde{\mathbf{e}}$ of $\mathbf{e}$, it remains true if, after time σ we use additional randomness, independent of $\mathbf{U}_1, \cdots, \mathbf{U}_\sigma$ and $\mathbf{V}$, to determine the order in which edges are chosen. In particular the pair in (17) still satisfies (15) if we replace $\tilde{\mathbf{e}}$ back by $\mathbf{e}$. Noting that that replacement gives exactly the pair of strings $(\mathbf{Z}, \mathbf{Z}')$ in the statement of Lemma 2.5, completes the proof of this lemma.

□

Recall the notation $I_{k,E}^{A,0}(\cdot)$ from Section 1 for the probability of being pivotal *and* having value 0.

Proposition 2.6. *Let $\mathbf{Z}$ and $\mathbf{Z}'$ be as in Lemma 2.5 (and recall that $e = \mathbf{e}_1$ is a deterministic element of E). Let $A \subset \{0,1\}^E$ be an increasing event. Then*

$$\mathbb{P}(\textit{ exactly one of } \mathbf{Z} \textit{ and } \mathbf{Z}' \textit{ is in } A) \leq \frac{2}{n} \sum_{f \in E \setminus \{e\}} \left(I_{k,E}^{A,0}(e) + I_{k,E}^{A,0}(f) \right). \tag{18}$$

Further, (18) *also holds with the superscript* 0 *replaced by* 1.
Moreover, (18) *also holds without the superscript* 0 *and the prefactor* 2*, i.e.,:*

$$\mathbb{P}(\textit{ exactly one of } \mathbf{Z} \textit{ and } \mathbf{Z}' \textit{ is in } A) \leq \frac{1}{n} \sum_{f \in E \setminus \{e\}} \left(I_{k,E}^{A}(e) + I_{k,E}^{A}(f) \right). \tag{19}$$

Proof.

$$\mathbb{P}(\textit{exactly one of } \mathbf{Z} \textit{and } \mathbf{Z}' \textit{is in } A) = \sum_{\alpha} \sum_{f}^{(*)} \frac{1}{n} P_{k,n}(\alpha) \tag{20}$$

$$= \frac{1}{n} \sum_{f \in E \setminus \{e\}} \sum_{\alpha}^{(**)} P_{k,n}(\alpha)$$

$$= \frac{1}{n} \sum_{f \in E \setminus \{e\}} (\text{ SUM(I) } + \text{ SUM(II) } + \text{ SUM(III) } + \text{ SUM(IV) }),$$

where the first equality follows from Lemma 2.5, the superscript $(*)$ means that the summation is over all f such that exactly one of α and $\alpha^{(e,f)}$ is in A, the superscript $(**)$ means that the summation is over all α such that exactly one of α and $\alpha^{(e,f)}$ is in A, and where:

$$\mathrm{SUM(I)} = \sum_{\alpha:\, \alpha\in A,\, \alpha^{(e,f)}\notin A,\, \alpha_e=0} P_{k,n}(\alpha); \quad \mathrm{SUM(II)} = \sum_{\alpha:\, \alpha\notin A,\, \alpha^{(e,f)}\in A,\, \alpha_e=0} P_{k,n}(\alpha);$$

$$\mathrm{SUM(III)} = \sum_{\alpha:\, \alpha\in A,\, \alpha^{(e,f)}\notin A,\, \alpha_f=0} P_{k,n}(\alpha); \quad \mathrm{SUM(IV)} = \sum_{\alpha:\, \alpha\notin A,\, \alpha^{(e,f)}\in A,\, \alpha_f=0} P_{k,n}(\alpha).$$

By symmetry (using, for fixed e and f, the 1-1 map $\alpha \leftrightarrow \alpha^{(e,f)}$), SUM(I) $=$ SUM(IV) and SUM(II) $=$ SUM(III), and hence the r.h.s. of (20) is equal to $\frac{2}{n}\sum_{f\neq e}(\mathrm{SUM(II)} + \mathrm{SUM(IV)})$. Finally, by the observation that the restriction on α in the definition of SUM(II) implies (since A is increasing) that e is 0-pivotal, and the restriction on α in the definition of SUM(III) implies that f is 0-pivotal, we get (18). The version of (18) with superscript 1, and the result (19), are obtained in (practically) the same way. $\square$

3 Proof of Theorem 1.1

See the notation and terminology given in Section 2.1. The strategy of the proof is roughly as follows. Our starting point is the general equation (13), which came from [6] and holds for all measures (not only monotone ones). We will see below that, after some general manipulations (suitable summing and conditioning) the proof deviates crucially from that in [6], since we (have to) invoke special properties of k-out-of-n measures. Applying such a property, namely Proposition 2.6, to the situation where we condition on the first $t-1$ selected 'posisitions' and their 0-1 values, and then summing over all possibilities (and using a simple standard inequality) will give the required result.

Now we start with the proof in detail: We have (recall that we take $n = |E|$),

$$
\begin{aligned}
&2P_{k,E}(A)(1 - P_{k,E}(A)) \qquad (21)\\
&\le \sum_{t=1}^{n} \mathbb{E}[\,|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})|\; \mathbb{I}_{t\le\tau}]\\
&= \sum_{t=1}^{n} \sum_{e\in E} \sum_{\tilde{E}\subset E\setminus\{e\}} \sum_{\beta\in\Omega_{\tilde{E}}} \mathbb{E}\left[\,|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})|\,\mathbb{I}_{\mathbf{e}_t=e,\,\mathbf{e}_{[t-1]}=\tilde{E},\,\mathbf{Y}^{t-1}_{\tilde{E}}=\beta,\,t\le\tau}\right]\\
&= \sum_{t=1}^{n} \sum_{e\in E} \sum_{\tilde{E}\subset E\setminus\{e\}} \sum_{\beta\in\Omega_{\tilde{E}}} \mathbb{E}\left[\mathbb{P}\left[!(\mathbf{Y}^{t-1}, \mathbf{Y}^t, A) \;|\; \mathbf{U}_{[t-1]}, \mathbf{V}_{[t-1]}\right] \mathbb{I}_{\mathbf{e}_t=e,\,\mathbf{e}_{[t-1]}=\tilde{E},\,\mathbf{Y}^{t-1}_{\tilde{E}}=\beta,\,t\le\tau}\right],
\end{aligned}
$$

where the inequality comes from (13), the notation $!(\mathbf{Y}^{t-1}, \mathbf{Y}^t, A)$ is used for the event that exactly one of $\mathbf{Y}^{t-1}$ and $\mathbf{Y}^t$ is in A, and where, with some abuse of notation, we wrote $\mathbf{e}_{[t-1]} = \tilde{E}$ for $\{\mathbf{e}_1, \cdots, \mathbf{e}_{t-1}\} = \tilde{E}$. Note that the last equality in (21) holds because, for each t, the event $\{\mathbf{e}_t = e, \mathbf{e}_{[t-1]} = \tilde{E}, \mathbf{Y}^{t-1}_{\tilde{E}} = \beta, t \le \tau\}$ is measurable w.r.t. the collection of random variables $(\mathbf{U}_{[t-1]}, \mathbf{V}_{[t-1]})$. Also note that, in the last two lines of equation (21) we may, in fact, replace $\mathbf{Y}^{t-1}_{\tilde{E}} = \beta$ by $\mathbf{Y}^{t-1}_{\tilde{E}} = \mathbf{Y}^t_{\tilde{E}} = \beta$.

Now observe that, while we are conditioning on $\mathbf{U}_{[t-1]}, \mathbf{V}_{[t-1]}$ with $\mathbf{e}_t = e, \mathbf{e}_{[t-1]} = \tilde{E}, \mathbf{Y}^{t-1}_{\tilde{E}} (= \mathbf{Y}^t_{\tilde{E}}) = \beta$ and $t \le \tau$, we have

$$
\mathbf{Y}^{t-1}_{E\setminus\tilde{E}} = F^{P_{k-|\beta|,E\setminus\tilde{E}}}_{(e,\mathbf{e}_{t+1},\cdots,\mathbf{e}_n)}(\mathbf{U}_t, \mathbf{U}_{t+1}, \cdots, \mathbf{U}_\tau, \mathbf{V}_{\tau+1}, \cdots, \mathbf{V}_n),
$$

$$
\mathbf{Y}^{t}_{E\setminus\tilde{E}} = F^{P_{k-|\beta|,E\setminus\tilde{E}}}_{(e,\mathbf{e}_{t+1},\cdots,\mathbf{e}_n)}(\mathbf{V}_t, \mathbf{U}_{t+1}, \cdots, \mathbf{U}_\tau, \mathbf{V}_{\tau+1}, \cdots, \mathbf{V}_n),
$$

and the event $!(\mathbf{Y}^{t-1}, \mathbf{Y}^t, A)$ in (21) is the event that exactly one of $\mathbf{Y}^{t-1}_{E\setminus\tilde{E}}$ and $\mathbf{Y}^{t}_{E\setminus\tilde{E}}$ is in $A(\beta)$, where $A(\beta)$ is the set of all $\omega \in \{0,1\}^{E\setminus\tilde{E}}$ for which $\omega \times \beta$ (the element of $\{0,1\}^E$ which coincides with β on $\tilde{E}$ and with ω on $E \setminus \tilde{E}$) is in A.

So, for fixed t, and with the above mentioned conditioning on $\mathbf{U}_{[t-1]}, \mathbf{V}_{[t-1]}$, we are in the context of Proposition 2.6, with proper adaptations (e.g. the $\mathbf{U}_1$ and $\mathbf{V}_1$ in that proposition correspond, in the current situation, with $\mathbf{U}_t$ and $\mathbf{V}_t$ respectively; and the n and k in that proposition are, in the current situation, $n - t + 1$ and $k - |\beta|$ respectively). Hence, applying Proposition 2.6, we obtain from (21) that

$$
\begin{aligned}
&2P_{k,E}(A)(1-P_{k,E}(A)) && (22)\\
&\leq \sum_{t=1}^{n}\sum_{e\in E}\sum_{\tilde{E}\subset E\setminus\{e\}}\sum_{\beta\in\Omega_{\tilde{E}}} \mathbb{P}(\mathbf{e}_t=e,\mathbf{e}_{[t-1]}=\tilde{E},\mathbf{Y}^{t-1}_{\tilde{E}}=\beta,t\leq\tau)\\
&\times\frac{2}{n-t+1}\sum_{f\in E\setminus(\tilde{E}\cup\{e\})}\left(I^{A(\beta),0}_{k-|\beta|,\,E\setminus\tilde{E}}(e)+I^{A(\beta),0}_{k-|\beta|,\,E\setminus\tilde{E}}(f)\right)\\
&=\ \mathrm{TERM(I)}+\mathrm{TERM(II)},
\end{aligned}
$$

where

$$
\begin{aligned}
&\mathrm{TERM(I)} && (23)\\
&=2\sum_{t=1}^{n}\sum_{e\in E}\sum_{\tilde{E}\subset E\setminus\{e\}}\sum_{\beta\in\Omega_{\tilde{E}}}\mathbb{P}\left(\mathbf{e}_t=e,\mathbf{e}_{[t-1]}=\tilde{E},\mathbf{Y}^{t-1}_{\tilde{E}}=\beta,t\leq\tau\right)\ I^{A(\beta),0}_{k-|\beta|,\,E\setminus\tilde{E}}(e)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{TERM(II)} && (24)\\
&=\sum_{t=1}^{n}\sum_{e\in E}\sum_{\tilde{E}\subset E\setminus\{e\}}\sum_{\beta\in\Omega_{\tilde{E}}}\mathbb{P}(\mathbf{e}_t=e,\mathbf{e}_{[t-1]}=\tilde{E},\mathbf{Y}^{t-1}_{\tilde{E}}=\beta,t\leq\tau)\\
&\times\frac{2}{n-t+1}\sum_{f\in E\setminus(\tilde{E}\cup\{e\})}I^{A(\beta),0}_{k-|\beta|,\,E\setminus\tilde{E}}(f).
\end{aligned}
$$

We first rewrite TERM(I): By the definition of $\mathbf{Y}^{t-1}$ (and using a version of Lemma 2.1), we have that, given the event that $\mathbf{e}_t=e,\mathbf{e}_{[t-1]}=\tilde{E}$ and $t\leq\tau$ (which is a $\mathbf{U}_{[t-1]}$ measurable event), the conditional probability that $\mathbf{Y}^{t-1}_{\tilde{E}}=\beta$ is equal to $P_{k,E}(\omega_{\tilde{E}}=\beta)$. Also note that $I^{A(\beta),0}_{k-|\beta|,\,E\setminus\tilde{E}}(e)$ is equal to $P_{k,E}(e \text{ is } 0-\text{pivotal w.r.t.} A\,|\,\omega_{\tilde{E}}=\beta)$. Hence we get

$$
\begin{aligned}
&\text{TERM(I)} \qquad (25)\\
&= 2\sum_{t=1}^{n}\sum_{e\in E}\sum_{\tilde{E}\subset E\setminus\{e\}} \mathbb{P}(\mathbf{e}_t = e, \mathbf{e}_{[t-1]} = \tilde{E}, t \le \tau)\\
&\times \sum_{\beta\in\{0,1\}^{\tilde{E}}} P_{k,E}(\omega_{\tilde{E}} = \beta)\, I^{A(\beta),0}_{k-|\beta|,\, E\setminus\tilde{E}}(e)\\
&= 2\sum_{t=1}^{n}\sum_{e\in E}\mathbb{P}(\mathbf{e}_t = e, t \le \tau)\, I^{A,0}_{k,E}(e)\\
&= 2\sum_{e\in E}\delta_e(A,T)\ I^{A,0}_{k,E}(e).
\end{aligned}
$$

Note that this is similar to the upper bound in the OSSS inequality for independent random variables. The treatment of TERM(II) takes a little more work and will give a somewhat different contribution which does not appear in the 'classical' OSSS. First, with practically the same arguments as used in (25), we get

$$
\begin{aligned}
&\text{TERM(II)} \qquad (26)\\
&= \sum_{t=1}^{n}\frac{2}{n-t+1}\sum_{e\in E}\sum_{\tilde{E}\subset E\setminus\{e\}} \mathbb{P}(\mathbf{e}_t = e, \mathbf{e}_{[t-1]} = \tilde{E}, t \le \tau) \sum_{f\in E\setminus(\tilde{E}\cup\{e\})} I^{A,0}_{k,E}(f).
\end{aligned}
$$

Next, summing over $\tilde{E}$ (and using the trivial observation that the sum over f in the expression (26) does not decrease if the restriction on f in that sum is omitted), gives

$$
\begin{aligned}
&\text{TERM(II)} \qquad (27)\\
&\le \sum_{t=1}^{n}\frac{2}{n-t+1}\sum_{e\in E}\mathbb{P}(\mathbf{e}_t = e, t \le \tau)\sum_{f\in E} I^{A,0}_{k,E}(f)\\
&= \sum_{t=1}^{n}\frac{2}{n-t+1}\mathbb{P}(t \le \tau)\ \sum_{f\in E} I^{A,0}_{k,E}(f).
\end{aligned}
$$

Note that, in the first sum in the r.h.s. of (27), the factor $1/(n-t+1)$ is increasing in t while the other factor, $\mathbb{P}(t \le \tau)$, is decreasing. Because of this (and also using that, since under $P_{k,E}$ the number of 1's and 0's is fixed, τ is a.s. at most $n-1$) that first sum in the r.h.s. of (27) is less than or equal to

$$\sum_{t=1}^{n-1} \frac{2}{n-t+1} \times \frac{1}{n-1} \sum_{t=1}^{n-1} \mathbb{P}(\tau \geq t) \leq 2 \log n \, \frac{1}{n-1} \, \mathbb{E}[\tau].$$

Hence, since $\mathbb{E}[\tau] = \sum_{e \in E} \delta_e(A,T)$, we have that

$$\text{TERM(II)} \; \leq 2 \log n \sum_{e \in E} I_{k,E}^{A,0}(e) \, \frac{1}{n-1} \sum_{e \in E} \delta_e(A,T). \tag{28}$$

Theorem 1.1 now follows immediately from (22), (25) and (28). $\square$

4 Proof of Theorem 1.3

The proof of Theorem 1.3 is essentially the same as that of Theorem 1.1. Here we present and explain the differences. First, by the same arguments used for the general result (13) (which came essentially from [6]), it is clear that the l.h.s. of that equation array is also less than or equal to

$$\sum_{t=1}^{(1-\varepsilon)n} \mathbb{E}[|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})| \, \mathbb{I}_{t \leq \tau}] + \mathbb{P}(\tau \geq ((1-\varepsilon)n),$$

where, for brevity, we write $(1-\varepsilon)n$ for $\lfloor (1-\varepsilon)n \rfloor$. Hence, in Section 3, the first inequality in (21) becomes

$$2P_{k,E}(A)(1-P_{k,E}(A)) \leq \sum_{t=1}^{(1-\varepsilon)n} \mathbb{E}[|\mathbb{I}_A(\mathbf{Y}^t) - \mathbb{I}_A(\mathbf{Y}^{t-1})| \, \mathbb{I}_{t \leq \tau}] + \mathbb{P}(\tau \geq ((1-\varepsilon)n). \tag{29}$$

To the summation in the r.h.s. of (29) we can then apply the same specific (for k-out-of-n measures) arguments and computations which led from (21) to (22) in Section 3. Instead of (22) we thus get

$$2P_{k,E}(A)(1-P_{k,E}(A)) \leq \text{TERM(I')} + \text{TERM(II')} + \mathbb{P}(\tau \geq (1-\varepsilon)n), \tag{30}$$

where TERM(I') and TERM(II') are as TERM(I) and TERM(II) in (25) and (26), respectively, except that the summation over t is now from 1 to

$\lfloor n(1-\varepsilon)\rfloor$. For TERM(I') we can thus trivially use the same upperbound (25) we had for TERM(I), i.e.,

$$\text{TERM(I')} \leq 2\sum_{e\in E} \delta_e(A,T)\ I_{k,E}^{A,0}(e), \tag{31}$$

Getting a suitable upper bound for TERM(II') requires a bit more attention (recall that we wanted to get rid of the logarithmic factor in (28)) and is done as follows. Instead of (27) we get

$$\begin{aligned}
&\text{TERM(II')} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (32)\\
&\leq \sum_{t=1}^{(1-\varepsilon)n} \frac{2}{n-t+1} \sum_{e\in E} \mathbb{P}(\mathbf{e}_t = e, t \leq \tau) \sum_{f\in E} I_{k,E}^{A,0}(f)\\
&\leq \sum_{t=1}^{n} \frac{2}{n-(1-\varepsilon)n} \mathbb{P}(t \leq \tau) \sum_{f\in E} I_{k,E}^{A,0}(f)\\
&\leq \frac{2}{\varepsilon n} \sum_{t=1}^{n} \mathbb{P}(t \leq \tau) \sum_{f\in E} I_{k,E}^{A,0}(f)\\
&= \frac{2}{\varepsilon n} \mathbb{E}(\tau) \sum_{f\in E} I_{k,E}^{A,0}(f).
\end{aligned}$$

Finally, combining (30), (31) and (32) (and using that $\mathbb{E}(\tau) = \sum_{e\in E} \delta_e$) immediately gives Theorem 1.3.

5 Box crossing probabilities in a percolation model with a fixed number of occupied vertices

Now we illustrate our main result, Theorem 1.1, by using it in the study of a percolation model on a box where the number of occupied vertices is fixed. We will also use a few basic results from the literature on the 'standard' percolation model (Bernoulli percolation), where the states of the vertices (or edges) are *independent* of each other. See e.g. [8] and [9] for a general introduction to Bernoulli percolation, and many results and references.

Consider an $R \times R$ box in the triangular lattice. More precisely, in terms of the standard embedding of this lattice in the plane (which we identify with the set of complex numbers $\mathbb{C}$), the box we consider is the graph with vertex set $V_R := \{x + y\exp(i\pi/3) \,:\, x, y \in \mathbb{Z},\, 0 \leq x, y \leq R-1\}$, and where two vertices v and w share an edge iff $|v-w|=1$. Each vertex can be *vacant* (which corresponds with value 0) or *occupied* (value 1).

Let A_R denote the event that there is an occupied path which crosses the box horizontally. In ordinary (Bernoulli) percolation, where the vertices are independently occupied with probability p and vacant with probability $1-p$ (the corresponding distribution will be denoted by P_p), it is well-known that $P_{\frac{1}{2}}(A_R) = 1/2$ (which follows from a simple symmetry argument), and that the expected number of pivotal vertices (and hence, by the well-known Margulis-Russo formula, also $\frac{d}{dp}P_p(A_R)\mid_{p=1/2}$) grows at least as a power of R. In fact, very sharp versions of this result are known, see [19].

We will study the same event A_R but now for the model where a fixed number, denoted by k, of the vertices is occupied. So the probability measure is now $P_{k,R^2} = P_{k,V_R}$. In particular we will study the case where R is even and $k = R^2/2$ (i.e. exactly half of the vertices is occupied).

Again, from symmetry, $P_{\frac{R^2}{2},R^2}(A_R) = 1/2$. We will show an analog of the Bernoulli percolation result mentioned above, namely that (roughly speaking) the 'discrete derivative' (with respect to the fraction of $1's$) of the above probability is again larger than a constant times a power of R. More precisely we will, using Theorem 1.1, prove Theorem 5.2 below, from which Corollary 5.4 below follows easily. (Analogs of this theorem and its corollary for *bond* percolation on the *square* lattice can be proved in practically the same way).

Recall the definitions of "pivotal" and "0-pivotal" from Section 1 (a few paragraphs below (1)).

Remark 5.1. *Although we don't use this in the proof of Theorem 5.2, we note that (as is well-known), for this specific event A_R, a vertex $v \in V_R$ is pivotal iff there are four disjoint paths, each starting from a neighbour of v to the boundary of the box: two occupied paths to the left and the right side, respectively, and two vacant paths to the top and and the bottom side, respectively.*

Theorem 5.2. *Let N_R^0 denote the number of vertices that are* 0*-pivotal for the event A_R, and let $E_{\frac{R^2}{2},R^2}(N_R^0)$ denote its expectation w.r.t. the distribution $P_{\frac{R^2}{2},R^2}$. There are $\alpha > 0$ and $C > 0$ such that, for all even $R \geq 2$,*

$$E_{\frac{R^2}{2},R^2}(N_R^0) \geq CR^\alpha. \tag{33}$$

Remark 5.3. (a) *The proof (in Section 5.1) of Theorem 5.2 uses our OSSS-type result Theorem 1.1 and a minimal amount of knowledge of Bernoulli*

percolation: essentially only RSW (and FKG), which are pre-1979 results.
(b) *Theorem 5.2 (and a stronger version) can also be proved without using Theorem 1.1, but such a proof would (as far as I know) require much heavier results from Bernoulli percolation, see Section 5.2.1.*
(c) *I am not aware of earlier mathematically rigorous work on percolation with a fixed number of occupied vertices or edges. In the physics literature, such models have been studied and compared with Bernoulli percolation: see the paper [10] where (heuristic) predictions are given for finite-size corrections of various quantities.*

Corollary 5.4. *For all even R,*

$$\frac{(P_{\frac{R^2}{2}+1,R^2}(A_R) - P_{\frac{R^2}{2},R^2}(A_R))}{1/R^2} \geq 2CR^\alpha, \tag{34}$$

with C and α as in Theorem 5.2.

Note that the denominator in the l.h.s. of (34) is the increase of the fraction of occupied vertices when the parameter changes from $R^2/2$ to $R^2/2+1$; so the l.h.s. of (34) can indeed be interpreted as a 'discrete derivative'.

The proof of Theorem 5.2 will be given in the next subsection. We will now first show how Corollary 5.4 follows from that theorem. For that we use and state Observation 5.5 below, which holds for all increasing events. This observation is an analog of the well-known Margulis-Russo formula for product measures. Its proof (which we omit), uses a straightforward coupling of $P_{k+1,n}$ and $P_{k,n}$, is simpler than that of the Margulis-Russo formula, and is probably, implicitly or explicitly, already in the litereature.

Observation 5.5. *Let $n \geq 1$ and let $A \subset \{0,1\}^n$ be an increasing event. Let N_A^0 denote the number of vertices that are 0-pivotal for A, and $E_{k,n}(N_A^0)$ its expectation w.r.t. the distribution $P_{k,n}$. For all $k \leq n-1$,*

$$P_{k+1,n}(A) - P_{k,n}(A) = \frac{1}{n-k} E_{k,n}(N_A^0). \tag{35}$$

Proof. (of Corollary 5.4 from Theorem 5.2). The corollary follows immediately from Theorem 5.2 by applying Observation 5.5, with $n = R^2$ and $k = R^2/2$, to the event A_R. $\square$

For the proof of Theorem 5.2 we will use the following almost trivial inequality relating $P_{\frac{n}{2},n}$ and $P_{1/2}$:

Observation 5.6. *For all even $n \geq 2$ and all increasing events $A \subset \{0,1\}^n$,*

$$P_{\frac{n}{2},n}(A) \leq 2P_{\frac{1}{2}}(A).$$

(This Observation follows immediately from the obvious facts that the event B that there are at least $\frac{n}{2}$ $1's$ has, under $P_{1/2}$, probability larger than 1/2, and that $P_{1/2}(A\,|B) \geq P_{\frac{n}{2},n}(A)$.)

As said in Remark 5.3, the proof of Theorem 5.2 uses only a minimal amount of knowledge from Bernoulli percolation, namely, FKG-Harris (the result that, for product measures, increasing events are positively correlated) and RSW. The form of RSW we will use is that there is a $c_1 > 0$ such that for all $R \geq 1$,

$$P_{\frac{1}{2}}(\exists \text{ an occupied vertical crossing of a given } 3R \times R \text{ box }) < c_1. \tag{36}$$

As is well-known, this (combined with FKG) easily implies that there is a $c_2 < 1$ such that, for all $R \geq 1$, the probablity (under $P_{\frac{1}{2}}$) that there is an occupied path crossing the annulus between two concentric boxes, one of size $R \times R$ and the other of size $3R \times 3R$, is smaller than c_2. Since a path from 0 to a point at distance m from 0 has to cross of order $\log m$ specific annuli of the above shape, it follows immediately (and is a classical result) that there are c_3, $c_4 > 0$ such that, for all $m \geq 1$,

$$\pi(m) < c_3 m^{-c_4}, \tag{37}$$

where $\pi(m) = P_{\frac{1}{2}}(\exists$ a path from 0 to a point at distance m from 0).

Remark 5.7. *Since the 1980's the result* (37) *has been dramatically sharpened ([14]), but it is, together with Theorem 1.1, sufficient for our purpose: to prove Theorem 5.2.*

5.1 Proof of Theorem 5.2

To obtain a lower bound for the expected number of pivotal vertices we will use our main result, Theorem 1.1, and the existence of a suitable algorithm (decision tree) for checking the existence of an occupied horizontal crossing in the box, where 'suitable' means that the revealment probabilities, under the measure $P_{\frac{R^2}{2},R^2}$ are in some sense small. (As in many applications of OSSS in the literature, we will average over a number of decision trees to obtain a desired 'average' smallness of revealments).

For discovering a crossing from the left to the right side of a box, there is a well-known algorithm involving a so-called *exploration path*. The construction of such a path is best explained in terms of the graph obtained by representing each vertex of the triangular lattice as (the midpoint of) a small hexagon in the hexagonal lattice (which is the dual of the triangular lattice). We colour a hexagon 'white' if its corresponding vertex in the triangular lattice is occupied, and 'black' otherwise. The most common exploration path is a path (in the box) on the mentioned hexagonal lattice, which starts in a corner of the box, say the lower-right corner, and, informally speaking, develops step-by-step in such a way that at each step there is white on the left and black on the right. There is an occupied horizontal crossing of the box if and only if the exploration path reaches the left side of the box before the top side. So, to discover whether the event A_R holds, only the colours of those vertices that are reached by the exploration path before it hits the left or top side of the box, have to be revealed.

This algorithm in itself is not 'suitable' yet, beacuse vertices close to the mentioned corner have a 'high' probability to be examined. To take care of this, Schramm and Steif [18] (see Section 4 of their paper) presented, in the context of Bernoulli percolation, an adaptation of the algorithm, which is informally as follows. For each point v_0 (denoted by p_0 in their paper) on the right side of the box, consider a straightforward modification $\beta = \beta_{v_0}$ of the above exploration path to discover if there is a horizontal occupied path from the left side of the box to the part of the right side above v_0. And, similarly, a modification β'_{v_0} to determine the existence of a horizontal occupied path from the left side of the box to the part of the right side below v_0. Together, these two exploration paths determine whether the event A_R occurs. So, for each v_0 one now has a decision tree for the event A_R. Schramm and Steif showed that, for each vertex w in the box, the average over all the above mentioned v_0's of the corresponding revealment probabilities for w, is 'small'.

As said, the above mentioned work by Schramm and Steif concerns Bernoulli percolation. For the percolation model with fixed number of occupied vertices we can, and will, consider the same decision trees (with exactly the same definition of β_{v_0} and β'_{v_0}). The revealment probabilities will of course differ from those in the Bernoulli case. Below is a somewhat informal outline of the computation leading to an upper bound on these revealment probabilities, where we focus on the differences and adaptations compared to that for Bernoulli percolation in [18]. See Section 4 of [18] for more precise (deterministic) properties (and pictures) of the exploration paths.

Let T_{v_0} denote the 'decision tree' corresponding to v_0 (i.e., with exploration paths β_{v_0} and β'_{v_0}). For brevity we will write $\mathbb{P}$ for $P_{\frac{R^2}{2},R^2}$. We have, for each hexagon H,

$$\mathbb{P}(H \text{ examined under } T_{v_0}) \leq \mathbb{P}(H \text{ on } \beta_{v_0}) + \mathbb{P}(H \text{ on } \beta'_{v_0}), \tag{38}$$

where "H on β" means that at least one of the sides of H is on β.

It is easy to see that if H touches β_{v_0} then there is a black or a white path from (a neighbour of) H to distance at least $|z_0(H) - v_0| - K$ from H, where $z_0(H)$ is the closest point to H on the right side of the box, and K is a (universal) constant. (Schramm and Steif obtained a considerably stronger statement, but for our purpose the one above is sufficient).
Hence

$$\begin{aligned} &\mathbb{P}(H \text{ on } \beta_{v_0}) \\ &\leq \mathbb{P}(\exists \text{ a white or a black path from } H \text{ to distance } |z_0(H) - v_0| - K \text{ from } H) \\ &= 2\mathbb{P}(\exists \text{ a white path from } H \text{ to distance } |z_0(H) - v_0| - K \text{ from } H). \end{aligned} \tag{39}$$

The same inequality also holds for β'_{v_0}, so together with (38) this gives

$$\begin{aligned} &\mathbb{P}(H \text{ examined under } T_{v_0}) \\ &\leq 4\mathbb{P}(\exists \text{ white path from } H \text{ to distance } |z_0(H) - v_0| - K \text{ from } H)). \end{aligned}$$

Summing this over all v_0 on the r.h.s. of the box (there are roughly R of these) and using Observation 5.6 and (37), gives, for every H,

$$\begin{aligned} &\sum_{v_0} \mathbb{P}(H \text{ examined under } T_{v_0}) \\ &\leq c_5 \sum_{m=1}^{R} \mathbb{P}(\exists \text{ white path from } H \text{ to distance } m \text{ from } H) \\ &\leq 2c_5 \sum_{m=1}^{R} c_3 m^{-c_4} \leq c_6 R^{1-c_4}, \end{aligned} \tag{40}$$

where c_5 and c_6 are constants > 0. Hence, for some constant $c_7 > 0$,

$$[\text{The average over all } v_0 \text{ of } \mathbb{P}(H \text{ examined under } T_{v_0})] \leq c_7 R^{-c_4}. \tag{41}$$

Now for each v_0 we apply our main result, Theorem 1.1 (with E the set of vertices in the $R \times R$ box, A the above mentioned crossing event A_R, and $T = T_{v_0}$). So for each v_0 this gives an inequality of the form (3). Note that the l.h.s. of each of these inequalities is 1/4. 'Averaging' these inequalities and using (41) yields

$$1/4 \leq (1 + \log R) \sum_{w} \mathbb{P}(w \text{ is } 0 - \text{pivotal for } A_R)\; c_7\, R^{-c_4},$$

(where the sum is over all vertices w in the $R \times R$ box) and hence that the l.h.s. of (33) is $\geq \frac{c_8}{\log R} R^{c_4}$, where c_8 is a universal constant. This completes the proof of Theorem 5.2. □

Remark 5.8. *We emphasize again that Schramm and Steif [18] worked on Bernoulli percolation (not on the percolation model with fixed number of occupied vertices). We also remark that it was not the goal of their paper to obtain lower bounds for the expected number of pivotals but to prove a form of quantitative noise sensitivity (and use that to prove the existence of exceptional times in a dynamical percolation model). For their purpose they used, apart from the above mentioned exploration paths, not OSSS but a different inequality (Theorem 1.8 in their paper), which involves discrete Fourier analysis. In fact, roughly speaking, the special case where a parameter in their Theorem 1.8 is equal to* 1*, produces an* upper *bound for the expected number of pivotals.*

5.2 Further comments on Theorem 5.2 and its proof

5.2.1 Sketch of an alternative proof

As said, the proof of Theorem 5.2 in the previous subsection uses, besides our OSSS-like result (Theorem 1.1), only a minimal amount of (Bernoulli) percolation theory: essentially only RSW. I don't know a proof of Theorem 5.2 which neither uses the OSSS-like result, nor more than only mild results from Bernoulli percolation.

One *can* prove Theorem 5.2 (and even a stronger version) without using OSSS, but instead using quite heavy results from Bernoulli percolation. I am not aware of such a proof in the literature, and will give an informal sketch how to do it:

First, recall that one step (namely, the last inequality but one in (40)) in the proof of Theorem 5.2 uses an obvious comparison (Observation 5.6) between $P_{\frac{1}{2}}$ and $\mathbb{P}$ $(= P_{\frac{R^2}{2}, R^2})$. That comparison was used to bound the probability (under $\mathbb{P}$) that there is an occupied path from a given vertex to a region at a certain distance from that vertex. Also recall (see Remark 5.1) that the event that a given vertex v is pivotal for A_R is the event that there are four distinct 'arms' with certain properties from neighbours of v to the boundary of the $R \times R$ box. It is natural to ask if the probability of that event under the probability measure $\mathbb{P}$ can also be suitably 'compared' with that under the measure $P_{\frac{1}{2}}$ (which would then yield another way to obtain Theorem 5.2).

Such comparison can indeed be made, but it involves, roughly speaking, a change of the parameter (which was 1/2) of the Bernoulli process. Such change is needed to 'control' the effect of the typical fluctuation of the number of occupied vertices in the $R \times R$ box in the Bernoulli model. This fluctuation is of course of order of the square root of the number of vertices in the box, i.e. of order $\sqrt{R}^2 = R$. This can be controlled ('compensated') by a change of the Bernoulli parameter by roughly $\frac{1}{R^2} \times$ the mentioned typical fluctiation, i.e., roughly $R/R^2 = 1/R$. The characteristic length scale of Bernoulli percolation at parameter $1/2 + \delta$, denoted by $L(1/2 + \delta)$, is known to be of order at least $1/\delta$. (Another way to say this, is that the correlation length critical exponent is at least 1. Much more is known but not needed here).

At length scales below that characteristic length, the Bernoulli model with the new parameter $(1/2 + \delta)$ behaves typically like the critical model, i.e. the Bernoulli model with parameter 1/2). That is essentially the main, and technically complicated, result in the celebrated paper [13] by Kesten, see also [15]. (Also see, for a slightly more subtle version of Kesten's result, which is actually needed in the argument above, with an extra state of the vertices, the paper [3] by Damron et al, in particular Lemma 6.3).

Since in our case $\delta \asymp 1/R$, the characteristic length is at least of order R (the length of the box). Hence, the mentioned four-arm event has, under the measure $\mathbb{P}$ roughly the same probability as under $P_{\frac{1}{2}}$. Summing that probability over all vertices in the 'bulk' of the box then yields Theorem 5.2 from its (well-known) analog for Bernoulli percolation.

5.2.2 Ideas concerning yet another proof

Another proof of Theorem 5.2 may be obtained by mimicking a classical proof for the Bernoulli percolation analog of that theorem: By first showing that with 'reasonable' probability there is a vertex in the bulk of the box with 'five armes', namely, besides the mentioned four arms an additional occupied arm to the boundary of the box (see e.g. Section 5.2 in [15] how to do this for Bernoulli percolation). And, next, showing that this extra path 'costs' at least a probability $R^{-\varepsilon}$ for some $\varepsilon > 0$. Finally, by summing over all vertices in the bulk of the box, this would then give (33) with $\alpha = \varepsilon$. However, the analog for the k-out-of-n model of the first part of this proof (i.e. showing that there is with reasonable probability a vertex 'with five arms') would again require a comparison with near-critical Bernoulli percolation and use of the above mentioned heavy work by Kesten [13].

5.2.3 Comparison with KKL in this box-crossing situation

We also briefly return to the k-out-of-n version of KKL from [17], mentioned in Section 1.2. Since each vertex v in the box has distance at least $R/2 - 1$ to the left or to the right side of the box, the probability that v is pivotal is at most

$$\mathbb{P}(\exists \text{ occupied path from a neighbour of } v \text{ to distance } \frac{R}{2} - 1 \text{ from } v),$$

which by Observation 5.6 and (37) is at most $2c_3(R/2-1)^{-c_4}$. Since this holds for each v in the box, application of the above mentioned version of KKL then gives that the expected number of pivotal vertices is at least of order $\log R$, which is much weaker than (33).

5.2.4 Final Remark

The main purpose of this paper is to present Theorem 1.1 (and Theorem 1.3), a version of OSSS for k-out-of-n measures. Theorem 5.2 and its practically self-contained proof (without using Bernoulli percolation results beyond RSW) are meant as an illustration of how that version of OSSS can be used, not as a major application.

Acknowledgments

I thank Jeff Kahn and Gabor Pete for some valuable comments and questions on the first version of this paper. In particular, JK brought the paper [17] by O'Donnell and Wimmer to my attention, and GP asked stimulating questions about the logarithmic factor in Theorem 1.1.

Further, I thank Pierre Nolin for helpful information concerning near-critical Bernoulli percolation.

References

[1] J. van den Berg and J. Jonasson, A BK inequality for randomly drawn subsets of fixed size, *Probab. Theory Relat. Fields* **154**, 835-844 (2012).

[2] S. Boucheron, G. Lugosi and P. Massart, *Concentration Inequalities*, Oxford University Press (2013).

[3] M. Damron, A. Sapozhnikov and B. Vágvölgyi, Relations between invasion percolation and critical percolation in two dimensions, *Ann. Probab.* **37**, 2297-2331 (2009).

[4] D. Dereudre and P. Houdebert, Sharp phase transition for the continuum Widom-Rowlinson model, *Ann. Inst. H. Poincaré Probab. Statist.* **57**, 387-407 (2021).

[5] H. Duminil-Copin, A. Raoufi and V. Tassion, Exponential decay of connection probabilities for subcritical Voronoi percolation in $\mathbb{R}^d$, *Probab. Theory Relat. Fields* **173**, 479-490 (2019).

[6] H. Duminil-Copin, A. Raoufi, and V. Tassion, Sharp phase transition for the random-cluster and Potts models via decision trees, *Annals of Mathematics* **189**, 75-99 (2019).

[7] Y. Filmus, R. O'Donnell and X. Wu, Log-Sobolev inequality for the multislice, with applications, *Electron. J. Probab.* **27**, article no. 33, 1-30 (2022).

[8] G.R. Grimmett. *Percolation*, second edition, Springer (1999).

[9] G.R. Grimmett. *Probability on graphs*, Second Edition, IMS Textbooks 8, Cambridge University Press (2018).

[10] H. Hu, H.W.J. Blöte and Y. Deng, Percolation in the canonical ensemble, *Journal of Physics A: Mathematical and Theoretical* **45** 494006 (2012).

[11] T. Hutchcroft, New critical exponent inequalities for percolation and the random cluster-cluster model, *Probab. and Math. Phys.* **1**, No. 1 (2020).

[12] J. Kahn, G. Kalai and N. Linial, The influence of variables on boolean functions, *Proc. 29th IEEE Symp. on Foundations of Comp. Sci.*, 68-80 (1988).

[13] H. Kesten, Scaling relations for 2-D percolation, *Comm. Math. Phys.* **109**, 109-156 (1987).

[14] G.F. Lawler, O. Schramm and W. Werner, One-arm exponent for critical 2D percolation, *Electr. J. Probab.* **7**, Paper 2 (2002).

[15] P. Nolin, Near-critical percolation in two dimensions. *Electron. J. Probab.* **13**, no.55, 1562-1323 (2008).

[16] R. O'Donnell, M. Saks, O. Schramm and R. Servedio, Every decision tree has an influential variable, 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05), Pittsburgh (2005).

[17] R. O'Donnell and K. Wimmer, KKL, Kruskal-Katona, and monotone sets, *SIAM Journal on Computing* **42**, 2375-2399 (2013).

[18] O. Schramm and J. Steif, Quantitative noise-sensitivity and exceptional times for percolation, *Ann. Math.* **171** (2), 619-672 (2010).

[19] S. Smirnov and W. Werner, Critical exponents for two-dimensional percolation, *Math. Res. Letters* **8**, 729-744 (2001).

[20] M. Talagrand, On Russo's approximate zero-one law, *Ann. Probab.* **22**, 1576–1587 (1994).