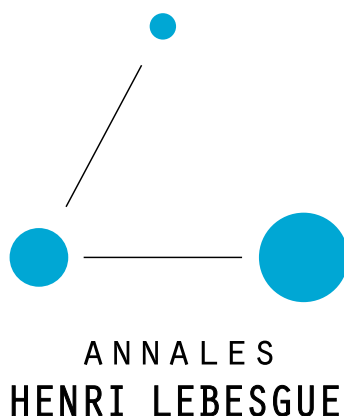




SÉBASTIEN MARTINEAU

RÉMY POUDEVIGNE-AUBOIRON

PAUL RAX

STOCHASTIC DOMINATION AND LIFTS OF RANDOM VARIABLES IN PERCOLATION THEORY

DOMINATION STOCHASTIQUE ET
RELÈVEMENTS DE VARIABLES
ALÉATOIRES EN THÉORIE DE LA
PERCOLATION

ABSTRACT. — Consider some matrix waiting for its coefficients to be written. For each column, sample independently a Bernoulli random variable of some parameter p . Seeing all this and possibly using extra randomness, Alice then chooses one spot in each column, in any way she wants. When the Bernoulli random variable of some column is equal to 1, the number 1 is written in the chosen spot. When the Bernoulli random variable of a column is 0, nothing is done on this column. We prove that, using extra randomness, it is possible for Bob to fill the empty entries with well chosen 0's and 1's so that the entries of the matrix are independent Bernoulli random variables of parameter p . We investigate various generalisations and variations of this problem, and use this result to revisit and generalise (nonstrict) monotonicity of the percolation

Keywords: percolation, stochastic domination, lift.

2020 *Mathematics Subject Classification:* 60E15, 82B43.

DOI: <https://doi.org/10.5802/ahl.284>

(*) Rémy Poudevigne-Auboiron has been supported by the project ANR LOCAL (ANR-22-CE40-0012-02) operated by the Agence Nationale de la Recherche (ANR).

threshold p_c with respect to a form of graph-quotienting, namely fibrations. We also use this result to revisit the BK inequality.

In a second part, which is independent of the first one, we revisit strict monotonicity of p_c with respect to fibrations, a result that naturally requires more assumptions than its nonstrict counterpart. We reprove the bond-percolation case of the result of Martineau–Severo without resorting to essential enhancements, using couplings instead.

RÉSUMÉ. — Considérons une matrice dont les coefficients sont encore à écrire par deux personnes, Alice et Bob. Pour chaque colonne, tirons indépendamment une variable aléatoire suivant une loi de Bernoulli d'un certain paramètre p . En ayant toutes ces informations, et usant potentiellement d'une source d'aléa externe, Alice sélectionne une case dans chaque colonne, selon une règle de son choix. Quand la variable aléatoire de Bernoulli de la colonne vaut 1, Alice écrit 1 dans la case sélectionnée, et passe sinon à la colonne suivante sans rien faire. Nous établissons que, en utilisant de l'aléa supplémentaire, il est toujours possible pour Bob de remplir les cases restantes avec des 0 et des 1 de sorte que les cases de la matrice soient des variables aléatoires *i.i.d.* de loi de Bernoulli de paramètre p . Nous explorons diverses généralisations de ce problème. À la lueur de ces nouvelles techniques, nous retrouvons deux résultats classiques de la théorie de la percolation : l'inégalité BK, et le fait que quotienter un graphe ne peut pas augmenter son seuil critique de percolation. La notion de quotient pertinente ici est celle de « fibration ».

Dans une seconde partie, indépendante de la première, nous revisitons le caractère strictement monotone de p_c vis-à-vis des fibrations, un résultat nécessitant plus d'hypothèses que son analogue large. Nous retrouvons le cas de percolation par arêtes du résultat de Martineau–Severo sans recourir aux *essential enhancements*, en n'utilisant que des arguments de couplage.

1. Introduction

1.1. Stochastic domination of lifts

One purpose of this article is to prove the following theorem and study several generalisations of it. Throughout the text, the notation $\pi^{-1}(b)$ stands for the set $\pi^{-1}(\{b\})$. Theorem 1.1 is best read with an eye on Figure 1.1.

THEOREM 1.1. — *Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. Let $p \in [0, 1]$. Let $(X_b)_{b \in B}$ be a collection of independent random variables with Bernoulli distribution of parameter p . On the same probability space, let $(S(b))_{b \in B}$ be a collection of random variables such that, for every $b \in B$, the random variable $S(b)$ takes values in $\pi^{-1}(b)$. At last, for every $a \in A$, set Y_a to be $X_{\pi(a)}$ if $S \circ \pi(a) = a$ and 0 otherwise.*

Then, the distribution of $(Y_a)_{a \in A}$ is stochastically dominated by $\text{Bernoulli}(p)^{\otimes A}$.

The conclusion of Theorem 1.1 means that it is possible to find a new probability space and random variables $(Y'_a)_{a \in A}$ and $(Z_a)_{a \in A}$ defined on this new probability space such that the following conditions hold:

- (1) the joint distribution of $(Y'_a)_{a \in A}$ is equal to that of $(Y_a)_{a \in A}$,
- (2) the random variables Z_a are i.i.d. with Bernoulli distribution of parameter p ,
- (3) and for every $a \in A$, we almost surely have $Y'_a \leq Z_a$.

Remark 1.2. — In Theorem 1.1, the random variables $S(b)$ are allowed to be highly correlated, both relative to each other and relative to (X_b) . It is good to have in mind the case when every $S(b)$ is of the form $f_b((X_{b'})_{b' \in B})$ for some measurable function f_b .

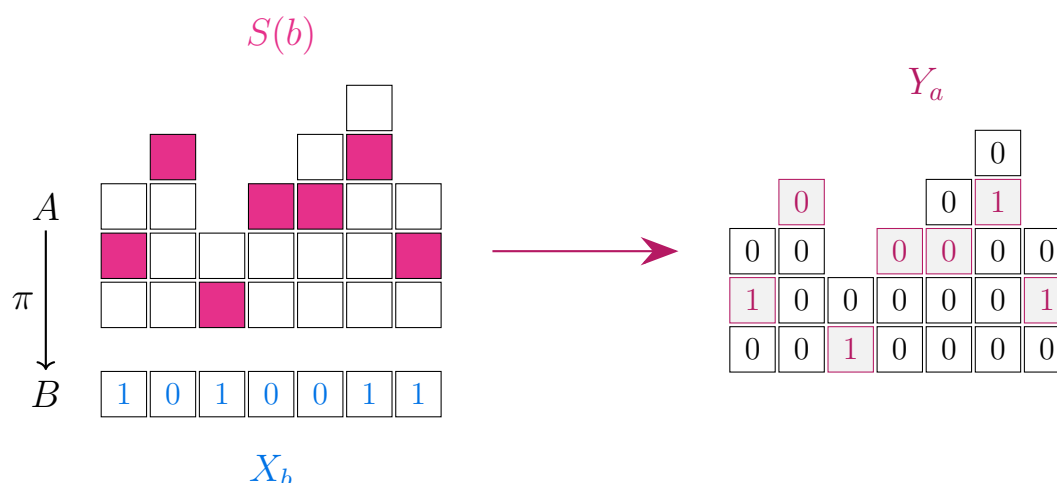


Figure 1.1. Illustration of the notation of Theorem 1.1.

What drew us to Theorem 1.1 lies in the realm of percolation theory. In Section 1.2, we provide general context, definitions and setup regarding percolation theory. Connections between percolation theory and Theorem 1.1 are then explained in Section 1.3.

1.2. General background regarding percolation theory

Let $\mathcal{G}$ be some countable locally finite graph.⁽¹⁾ *Site percolation* (resp. *bond percolation*) of parameter p consists in declaring each vertex (resp. edge) to be open with probability p , independently. It can be thought of as a random subgraph as follows. In the site case, the new set of vertices is that of open vertices, and two vertices are said to be adjacent if they were adjacent in the original graph. In the bond case, all vertices are kept but only open edges are retained. In both cases, connected components of this random subgraph are called *clusters*. Percolation theory is devoted to understanding these clusters. In particular, one is interested in the *infinite* clusters: is there any? if so, how many are there? what is their geometry?

The probability that there is an infinite cluster is a function of p which is weakly increasing and can only take values in $\{0, 1\}$, due to Kolmogorov's zero-one law.

⁽¹⁾By “graph”, we mean a simple graph, i.e. the edges are undirected and we do not allow self-loops or multiple edges. “Locally finite” means that each vertex has a finite number of neighbours. A graph is “countable” if its vertex-set is countable (which entails countability of its edge-set). We take “countable” to mean “finite or countable”. Being interested in infinite connected components, the finite case, although not excluded, will have no interest: finite graphs have critical parameters equal to 1.

There is therefore a unique parameter $p_c(\mathcal{G}) \in [0, 1]$ such that, for every $p \in [0, 1]$, this probability is zero if $p < p_c(\mathcal{G})$ and one if $p > p_c(\mathcal{G})$. The number $p_c(\mathcal{G})$ is called the *critical parameter* of $\mathcal{G}$. More precisely, depending on whether one works with site or bond percolation, each graph gives rise to two critical parameters, namely $p_c^{\text{site}}(\mathcal{G})$ and $p_c^{\text{bond}}(\mathcal{G})$.

For most graphs, the critical parameters $p_c^{\text{site}}(\mathcal{G})$ and $p_c^{\text{bond}}(\mathcal{G})$ are nontrivial, meaning that they are neither equal to 0 nor to 1. Whenever these values are indeed nondegenerate, percolation gives rise to two regimes (one below p_c and one above), and we say that this process undergoes a phase transition.

Percolation has been extensively studied since its introduction by Broadbent and Hammersley almost seventy years ago [BH57], particularly on the d -dimensional cubic lattice $\mathbb{Z}^d$. It was introduced as a model for propagation of a fluid in a porous medium, random removal accounting for porosity and clusters for propagation. Since then, percolation has continued to attract attention for several reasons: it is interesting as a model of statistical mechanics; it is connected to diverse mathematical domains; and it is charming and challenging in itself. See [DC18] and references therein for a survey.

Since the seminal papers [BS96, BLPS99], percolation on rather general graphs has been investigated in a systematic way, revealing a deep interplay between percolation, geometry, and group theory. A fair amount of what we revisit in this article lies in this lineage.

1.3. Revisiting a percolation result via Theorem 1.1

It was proved in [BS96, Theorem 1] (see also [CMT23, Proposition 4.1]) that if a graph $\mathcal{S}$ is a “quotient” of a graph $\mathcal{L}$, then, for both site and bond percolation, we have $p_c(\mathcal{S}) \geq p_c(\mathcal{L})$. Here is a more precise statement, which encompasses quotients but also other situations, via the general concept of fibration.

Notation 1.3. — We denote the large graph by the symbol $\mathcal{L}$, which one may read “large” rather than “L”. Likewise, we denote the small graph by the symbol $\mathcal{S}$, which one may read “small” rather than “S”. Given any graph $\mathcal{G}$, we denote by $V_{\mathcal{G}}$ its set of vertices and by $E_{\mathcal{G}}$ its set of edges.

In this paper, given two countable locally finite graphs $\mathcal{L}$ and $\mathcal{S}$, a *fibration* from $\mathcal{L}$ to $\mathcal{S}$ is a surjective map $\pi : V_{\mathcal{L}} \rightarrow V_{\mathcal{S}}$ such that for every vertex x of $\mathcal{L}$ and every neighbour v of $\pi(x)$ in $\mathcal{S}$, there is a neighbour y of x in $\mathcal{L}$ such that $\pi(y) = v$.

Remark 1.4. — Apart from surjectivity, the definition of fibration states that for every vertex u in $\mathcal{S}$ and every neighbour v of u , we are able to “lift” the edge $\{u, v\}$ in $\mathcal{S}$ in order to put the endpoint corresponding to u *anywhere* in $\pi^{-1}(u)$.

THEOREM 1.5 (Benjamini–Schramm 96). — *Let $\mathcal{L}$ and $\mathcal{S}$ be two countable locally finite graphs. Assume that there is a fibration from $\mathcal{L}$ to $\mathcal{S}$.*

Then, we have $p_c^{\text{site}}(\mathcal{L}) \leq p_c^{\text{site}}(\mathcal{S})$ and $p_c^{\text{bond}}(\mathcal{L}) \leq p_c^{\text{bond}}(\mathcal{S})$.

The proof of Theorem 1.5 proceeds by exploring (a spanning tree of) the cluster of the origin step by step. Doing so properly ensures that vertices (or edges) yet to

be revealed are still, conditionally on previous steps, open independently and with probability p . As π can lift edges, one is able to lift trees as well. By performing p -percolation on $\mathcal{S}$, lifting what one sees in the exploration (both closed and open vertices or edges) to $\mathcal{L}$, and then putting i.i.d. Bernoulli(p) random variables on what remains of $\mathcal{L}$, one gets a coupling such that whenever the origin is connected to infinity in $\mathcal{S}$, this is the case as well in $\mathcal{L}$.

This raises the question of whether it is necessary to proceed via such a cautious exploration algorithm, and to lift both positive and negative information. If $p > p_c(\mathcal{S})$, would it not be easier to simply pick an infinite open path in $\mathcal{S}$ and lift it to $\mathcal{L}$, thus proving that $p \geq p_c(\mathcal{L})$? If we lift an infinite path and put i.i.d. Bernoulli(p) random variables on its complement, it will not work: some vertices will be open with probability larger than p , and independence would be problematic as well. Nevertheless, Theorem 1.1 ensures that it is possible to fill the complement of the lifted path with well chosen random variables so that the final result is p -percolation on $\mathcal{L}$.

Sketch of proof of Theorem 1.5 using Theorem 1.1. — We treat the case of site percolation, as the case of bond percolation can either be tackled in the same way or reduced to that of site percolation — see Section 4.1.

Let p be such that there is an infinite open path in $\mathcal{S}$ almost surely. Let us show that this is also the case for $\mathcal{L}$. We set $A = V_{\mathcal{L}}$ and $B = V_{\mathcal{S}}$. Let $(X_b)_{b \in B}$ be a p -percolation on $\mathcal{S}$, where $X_b = 1$ means that the vertex b is open and $X_b = 0$ means that it is closed. Almost surely, there is an infinite open path: pick one and call it κ . One can lift it via π to some path in $\mathcal{L}$, which we call $\tilde{\kappa}$. For each $b \in B$, if b belongs to κ , set $S(b)$ to be its corresponding vertex in $\tilde{\kappa}$, otherwise define $S(b)$ in an arbitrary way. Define (Y_a) as in Theorem 1.1. Notice that the configuration (Y_a) contains an infinite open path, namely $\tilde{\kappa}$. Theorem 1.1 now states that (Y_a) is stochastically dominated by p -percolation on $\mathcal{L}$. As there is almost surely an infinite open path in (Y_a) , so is the case for p -percolation on $\mathcal{L}$. This completes the proof, except that we did not take care of measurability issues regarding κ and $\tilde{\kappa}$. We defer these details to Section 4.2. $\square$

1.4. Other results we revisit

Still in the seminal paper [BS96], Benjamini and Schramm asked a precise question which can be very roughly summarised as: is p_c strictly increasing under strict quotient? The statement they asked for was proved in [MS19] by using essential enhancements.⁽²⁾ We provide a new proof of the bond-percolation case of this result. Our new proof does not use essential enhancements or differential inequalities but couplings instead. This new proof does not rely on Theorem 1.1 (it is more traditional and relies on appropriate exploration algorithms).

⁽²⁾For a reference on essential enhancements, see the original paper [AG91]. As some lemma in this paper is not properly established, see [BBR14] for a discussion of what exactly is rigourously known.

At last, let us mention that even though we did not investigate stochastic domination of lifts with the BK inequality in mind, our Theorem 1.1 is tailored to revisit it. This is performed in the short Section 5. The BK inequality itself is presented there.

Structure of the paper

In Section 2, we state and prove our main theorem, of which Theorem 1.1 is a particular case. Section 3 is devoted to investigating variations of the main theorem: a rather delicate statement appears when we try to lift several random variables per column (see Theorem 3.1). In Section 4, we apply our main theorem to not only give a new proof of Theorem 1.5 but also generalise it. The BK inequality is revisited in Section 5. At last, we revisit strict monotonicity of p_c under quotient. This is performed in Section 6, which is completely independent of the other sections. All sections can be read mostly independently of each other, except that we advise to read Section 2.1 before Section 3.

Acknowledgments

This project originated as a side project from an internship of the third author under the supervision of the first one and David García-Zelada, whom we thank for the very nice working experience as well as for his clear view of geometry. We are grateful to François Ledrappier and Florent Martineau for stimulating conversations. We are also indebted to David García-Zelada, Thierry Lévy and Romain Tessera for leading us to understand that the terminology “weak covering map” from Martineau–Severo was poorly chosen: these maps are much closer to fibrations in homotopy theory than to covering maps. We are grateful to Franco Severo for asking whether our techniques, then used to prove non-strict inequalities for p_c , could be used to recover the strict monotonicity theorem of [MS19]: we did not manage to do so but our attempts have led to the results of Section 3.1 and Section 6. We thank Vibhu Dalal for proofreading this article, and the reviewer for their feedback. At last, S. Martineau has a warm thought for Vincent Tassion, as Section 6 has a flavour very close to what we were trying to do at the very beginning of our doctorate.

2. Main theorem

In this section, we establish a vastly generalised version of Theorem 1.1, namely Theorem 2.3. The main point of Theorem 2.3 is to yield Corollary 2.4, which is similar to Theorem 1.1 but with $[0, \infty]$ -valued random variables, still taken to be independent⁽³⁾ but not necessarily identically distributed. This corollary will be stated using a different framework. Stated in the same way as Theorem 1.1, it goes as follows.

⁽³⁾ Only between π -fibres, not within them.

For completeness, we first recall what we mean in general by stochastic domination. Let $(E, \mathcal{E})$ be a measurable space. We further assume that E is endowed with a partial order $\leq$ such that $\{(x, y) \in E^2 : x \leq y\}$ belongs to the σ -field $\mathcal{E} \otimes \mathcal{E}$. Let μ and ν be two probability measures on $(E, \mathcal{E})$. We say that μ is *stochastically dominated* by ν if there are random variables X and Y defined on the same probability space such that the following three conditions hold:

- (1) the distribution of X is μ ,
- (2) the distribution of Y is ν ,
- (3) the inequality $X \leq Y$ holds almost surely.

When working on $[0, \infty]^A$, the implicit order will always be the product order, given by

$$x \leq y \iff \forall a \in A, \quad x_a \leq y_a.$$

The set $[0, \infty]$ is endowed with its usual topology, making it a compact space. The space $[0, \infty]^A$, endowed with the product topology, is then compact as well.

COROLLARY 2.1 (reformulation of Corollary 2.4). — *Recall the notations of Theorem 1.5. Then, for every $a \in A$, set $Y_a := X_{\pi(a)} \mathbf{1}_{S \circ \pi(a)=a}$.*

For every $b \in B$, let ρ_b be a probability measure on $[0, \infty]^{\pi^{-1}(b)}$. Assume that for every $b \in B$ and every $a \in \pi^{-1}(b)$, the distribution of X_b is stochastically dominated by the a -marginal of ρ_b . Then, the distribution of $(Y_a)_{a \in A}$ is stochastically dominated by $\bigotimes_{b \in B} \rho_b$.

Remark 2.2. — Even though this statement seems much more general than Theorem 1.1, it can actually be easily deduced from the $p = 1/2$ case of Theorem 1.1. Indeed, by the trick of the pseudoinverse⁽⁴⁾ cumulative distribution function, it suffices to deal with the case where each X_b is uniformly distributed on $[0, 1]$. Instead of assuming marginals of ρ_b to stochastically dominate $\text{Unif}([0, 1])$, we can assume that we are in the least favourable case, namely when these marginals are equal to the uniform distribution on $[0, 1]$. By considering binary digits, which are then independent Bernoulli random variables of parameter $1/2$, up to multiplying A and B by $\mathbb{N}$, the general case is reduced to the balanced case of Theorem 1.1.

2.1. Statement

Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. We will be interested in a specific kind of probability measures on $[0, \infty]^A$. Say that a probability measure μ on $[0, \infty]^A$ is a π -*lift* if, for μ -almost surely x , for every $b \in B$, at most one $a \in \pi^{-1}(b)$ satisfies $x_a \neq 0$.

Notice that a probability measure is a π -lift if and only if it can be obtained as the distribution of a random variable Y built as follows:

- (1) sample some random variable X taking values in $[0, \infty]^B$,
- (2) on the same probability space as X , for every b , sample some random variable $S(b)$ taking values in $\pi^{-1}(b)$,

⁽⁴⁾More formally, this is a left adjoint in the sense of Galois connections.

(3) for every $a \in A$, set $Y_a = X_{\pi(a)} \mathbb{1}_{S \circ \pi(a)=a}$.

One purpose of the present paper is to study results of the following type: *Let μ and ρ be two probability measures on $[0, \infty]^A$. Assume that μ is a π -lift and “some suitable assumptions”. Then, the probability measure μ is stochastically dominated by ρ .*

Our main result is the following one. Recall that a *section* s of π is a map $s : B \rightarrow A$ such that $\pi \circ s = \text{id}_B$. Informally, it is a way to select one element in each $\pi^{-1}(b)$.

THEOREM 2.3. — *Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. Let μ and ρ be two probability measures on $[0, \infty]^A$. Assume that μ is a π -lift. Further, assume it is possible to fix a section s of π such that the following two conditions simultaneously hold:*

A $_{\rho}$ *Let Z be a ρ -distributed random variable. For every $b \in B$, let W_b denote the random variable $(Z_c : c \in A \text{ such that } \pi(c) \neq b)$. For every b , every $a \in \pi^{-1}(b)$ and every event $H \in \sigma(W_b)$ of positive probability, the conditional probability distribution of $Z_{s(b)}$ given H is stochastically dominated by that of Z_a given H .*

B $_{\mu}^{\rho}$ *Let Y be a μ -distributed random variable. For every $b \in B$, let $X_b := \max_{a \in \pi^{-1}(b)} Y_a$. Notice that this maximum is almost surely well defined since μ is a π -lift. For every $a \in A$, let $Y'_a = X_{\pi(a)} \mathbb{1}_{s \circ \pi(a)=a}$. The condition we ask is that the distribution of Y' is stochastically dominated by ρ .*

Then, the probability measure μ is stochastically dominated by ρ .

The assumptions of Theorem 2.3 are not meant to be natural a priori but to capture the essence of the proof (and we will see in Section 3.2 that a natural generalisation of it fails). From Theorem 2.3, we deduce the two following more concrete results, which are derived in Section 2.5.

If μ is a π -lift, the pushforward of μ by $x \mapsto (b \mapsto \sup_{a \in \pi^{-1}(b)} x_a)$ is called the *pushdown* of μ . This pushdown can alternatively be described as follows. Let Y be a μ -distributed random variable. For every $b \in B$, let $Z_b := \max_{a \in \pi^{-1}(b)} Y_a$. Notice that this maximum is almost surely well defined since μ is a π -lift. Then, the distribution of (Z_b) is the pushdown of μ .

COROLLARY 2.4. — *Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. For every $b \in B$, let ν_b be a probability measure on $[0, \infty]$ and let ρ_b be a probability measure on $[0, \infty]^{\pi^{-1}(b)}$. Assume that, for every $b \in B$ and every $a \in \pi^{-1}(b)$, the probability measure ν_b is stochastically dominated by the a -marginal of ρ_b . Let μ be a π -lift such that its pushdown is stochastically dominated by $\bigotimes_{b \in B} \nu_b$.*

Then, the probability measure μ is stochastically dominated by $\bigotimes_{b \in B} \rho_b$.

Remark 2.5. — This version of Corollary 2.4 is slightly stronger than Corollary 2.1, as we only assume the pushforward of μ to be stochastically dominated by $\bigotimes_b \nu_b$ instead of equal to it. It is actually not hard to derive the strong version from the weak one. Indeed, any π -lift with such a dominated pushforward measure is stochastically dominated by a π -lift such that its pushforward measure is equal to $\bigotimes_{b \in B} \nu_b$ (if Y is constructed using (X, S) , Y' using (X', S) , and if $X \leq X'$, then we have $Y \leq Y'$).

Let $\mathfrak{S}_\pi$ denote the group of all permutations σ of A that satisfy $\forall a \in A, \pi \circ \sigma(a) = \pi(a)$. This group acts on $[0, \infty]^A$ via $\sigma \cdot x : a \mapsto x_{\sigma^{-1}(a)}$. We say that a probability measure ρ on $[0, \infty]^A$ is π -exchangeable if, for every $\sigma \in \mathfrak{S}_\pi$, the pushforward of ρ by $x \mapsto \sigma \cdot x$ is equal to ρ .

Example 2.6. — Sample a random variable Θ . Depending on Θ , for every b , choose some probability measure $\nu_{b,\Theta}$ on $[0, \infty]$. Then, conditionally on Θ , let $(Z_a)_{a \in A}$ be a collection of independent random variables such that, for every a , the random variable Z_a has distribution $\nu_{\pi(a),\Theta}$. Then, the (unconditional) distribution of $(Z_a)_{a \in A}$ is π -exchangeable.

Example 2.7. — Let $\pi : A \rightarrow B$ be a surjective map between finite sets. Let $k \in \{0, 1, \dots, |A|\}$. Pick a random subset of A uniformly at random among all sets of cardinality k . Then, the distribution of the indicator function of this random set is π -exchangeable.

More generally, say that ρ is *sufficiently symmetric* if it is possible to choose, for every b , a subgroup G_b of the group of permutations of $\pi^{-1}(b)$ that acts transitively on $\pi^{-1}(b)$ in such a way that for every $\sigma \in \prod_b G_b$, the pushforward of ρ by $x \mapsto \sigma \cdot x$ is equal to ρ .

Example 2.8. — Every π -exchangeable probability measure ρ is in particular sufficiently symmetric.

Example 2.9. — Let $A = B \times \mathbb{Z}/4\mathbb{Z}$ and let $\pi : (b, x) \mapsto b$. For every b , independently and equiprobably, select either $\{(b, 0), (b, 2)\}$ or $\{(b, 1), (b, 3)\}$. Then, the distribution of the indicator function of all selected elements of A is sufficiently symmetric but not π -exchangeable.

Let ρ be a probability measure on $[0, \infty]^A$. For every section s , let the s -marginal of ρ denote the pushforward of ρ by $x \mapsto (x_{s(b)})_{b \in B}$. Observe that if ρ is sufficiently symmetric, then all its s -marginals are equal. We call this univoquely defined probability measure the *horizontal marginal* of ρ .

COROLLARY 2.10. — Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. Let μ and ρ be two probability measures on $[0, \infty]^A$. Assume that:

- (1) μ is a π -lift,
- (2) ρ is sufficiently symmetric,
- (3) the pushdown of μ is stochastically dominated by the horizontal marginal of ρ .

Then, the probability measure μ is stochastically dominated by ρ .

In order to prove Theorem 2.3, the key step will be to establish it in the finite setup, namely when A and B are finite and $[0, \infty]$ is replaced by $[N] := \{0, 1, \dots, N\}$. This is performed in Section 2.3. The general case then follows by general arguments of discrete approximation and compactness (see Section 2.4).

The finite case is tackled by first taking care of the one-column case, which means that we further assume B to be a singleton. For this, a mere greedy algorithm works (see Section 2.2). To handle the general (still finite) case, we proceed by induction,

putting carefully each element of B into play one after the other. Adding one such element is made possible because of Assumption $\mathbf{A}_\rho$ of Theorem 4.2 and the one-column case. Performing this strategy requires some care: one important subtlety in the proof is that when a column is yet to be considered, it should be treated as a singleton, not as an empty column.

Before moving on to the proof, we would like to explain why this latter, “natural” approach does not suffice. Let us take the one-column case for granted and try to deduce from it the two-columns case, i.e. the case where B has cardinality 2. For concreteness, let us try to work out the following example. We take $A = \{\text{left}, \text{right}\} \times \{\text{top}, \text{bottom}\}$, $B = \{\text{left}, \text{right}\}$ and $\pi : A \rightarrow B$ the obvious projection. Let X be a random variable having the uniform distribution on $\{0, 1\}^B$, so that its components are independent Bernoulli random variables of parameter $1/2$. Let ρ be the uniform probability measure on $\{0, 1\}^A$. The random section S we consider is the following: if $X = (1, 1)$ then $S : i \mapsto (i, \text{top})$, otherwise we take $S : i \mapsto (i, \text{bottom})$. Looking only at the first component and using the one-column case, we can stochastically dominate what happens on the left column as desired. To conclude, it would suffice to argue that, conditionally on that, X_{right} still has a Bernoulli distribution of parameter $1/2$.

At first sight, this may seem to work, as the components of X have been taken to be independent with this distribution. However, in order to take care of the left column, we needed to know not only X_{left} but also $S(\text{left})$, which revealed information on X_{right} .

In particular, if we lifted a 1 in the top-left position, then we know for sure that X_{right} is equal to 1. For this reason, adding columns in the naive way is not enough to solve the problem at hand. We have to take into account the fact that the conditional distribution of X_{right} need not be equal to (or stochastically dominated by) its unconditional distribution. This is why columns that are not handled yet will not be fully forgotten: it will be useful to remember the *values* that are lifted but not *where* they are lifted. Assumptions $\mathbf{A}_\rho$ and $\mathbf{B}_\mu^\rho$ then emerge as the natural setup to suitably perform our “surgery of couplings”.

At last, let us mention that it is possible to give an explicit example of a suitable coupling on the first column that cannot, in any way, be extended into a suitable coupling on two columns. For instance in our setting, we can consider the following coupling:

$$\begin{array}{cccc} \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \end{bmatrix}. \end{array}$$

This suitable coupling cannot be extended into two columns for the following reason. Conditioned on the left column being $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we know that we need a 1 on the top right corner for the coupling to work. However, this can only happen with probability at most $1/2$, therefore the coupling cannot be extended.

2.2. The one-column case

Let us start by proving the following easy lemma. For $n \geq 0$, we write $[n] := \{0, 1, \dots, n\}$.

LEMMA 2.11. — *Let N be a positive integer and let X be an $[N]$ -valued random variable. Let C be a finite set and let ρ be a probability measure on $[N]^C$. Assume that for every $c \in C$, the distribution of X is stochastically dominated by the c -marginal of ρ .*

Let H be a C -valued random variable. Let Y denote the random element of $[N]^C$ defined by $Y : c \mapsto X \mathbf{1}_{H=c}$. Then, the distribution of Y is stochastically dominated by ρ .

Proof. — Without loss of generality, we will assume that $C = [M]$, for some M . We consider the set $\mathcal{J} = (\{1, \dots, N\} \times [M]) \sqcup \{0\}$. This set encodes all possible values for Y . More precisely, the state 0 encodes $\varphi_0 := 0 \in [N]^C$ and (i, j) encodes $\varphi_{i,j} := i \mathbf{1}_{\{j\}}$. We endow $\mathcal{J}$ with the following well-ordering of lexicographic nature. We declare 0 to be larger than any element of $\mathcal{J}$, and $(i, j) \preceq (i', j')$ holds if and only if $i > i'$ or both $i = i'$ and $j \leq j'$. Please note that the inequalities for i and j go in reverse direction. In words, introducing $\preceq$ means that we want to take care of the possible states in the following order:

$$(N, 1), (N, 2), \dots, (N, M), (N-1, 1), \dots, (N-1, M), \dots, (1, 1), \dots, (1, M), 0.$$

The rest of the proof can be summarised as follows: proceeding greedily in this order works.

Let us introduce some formalism. Denote by μ the distribution of Y . A *subcoupling* is a measure ν on $[N]^C \times [N]^C$ such that the following two conditions hold:

- for every $\alpha \in [N]^C$, we have $\sum_{\beta} \nu(\alpha, \beta) \leq \mu(\alpha)$,
- for every $\beta \in [N]^C$, we have $\sum_{\alpha} \nu(\alpha, \beta) \leq \rho(\beta)$.

Given some $I \in \mathcal{J}$, say that a subcoupling is *suitable for step I* if it satisfies the following two conditions:

- for every $J \preceq I$, we have $\sum_{\beta} \nu(\varphi_J, \beta) = \mu(\varphi_J)$
- for every α which is not of the form φ_J for $J \preceq I$, we have $\sum_{\beta} \nu(\alpha, \beta) = 0$.

Our purpose is to prove that there is a subcoupling that is suitable for step 0. If we reach that goal, then we are done. Indeed, as the support of μ is a subset of $\{\varphi_I : I \in \mathcal{J}\}$, the first marginal will be equal to μ , thus yielding that the total mass of the subcoupling is 1; and as ρ has mass 1 as well, the second marginal will also be equal to ρ instead of satisfying only the inequality of subcouplings.

Let us prove by induction on $(\mathcal{J}, \preceq)$ that, for every $I \in \mathcal{J}$, there is a subcoupling suitable for step I . We start with $I = (N, 1)$. Using the assumption of the lemma for $c = 1$, we have $\sum_{\beta: \beta(1)=N} \rho(\beta) \geq \mathbf{P}(X = N) \geq \mathbf{P}(Y = \varphi_{N,1})$, so indeed ρ has enough room in $\{\beta : \beta(1) = N\}$ to allow such a subcoupling.

Now let I be such that there is a subcoupling suitable for step I , and let us pick ν such a subcoupling. If $I = 0$, then we are done, so we assume that $I = (i, j)$. There are three cases to handle:

- (1) when $j < M$,
- (2) when $j = M$ and $i > 1$,
- (3) when $(i, j) = (1, M)$.

Let us consider the first case. We want to prove that there is a subcoupling suitable for step $(i, j + 1)$. Using our assumption for $c = j + 1$, we have

$$\text{Space} := \sum_{\beta: \beta(j+1) \geq i} \rho(\beta) \geq \mathbf{P}(X \geq i) = \mathbf{P}(X = i) + \sum_{i' > i} \mathbf{P}(X = i'),$$

where the last sum is the null empty sum if $i = N$. Let us compute how much of this space has been used already, and how much we need to use now. We need to store a mass of $\text{Need} := \mathbf{P}(X = i, H = j + 1)$. On the other hand, the total mass of ν is $\text{Used} := \sum_{j' \leq j} \mathbf{P}(X = i, H = j') + \sum_{i' > i} \mathbf{P}(X = i')$. Therefore, the inequality

$$\mathbf{P}(X = i) \geq \mathbf{P}(X = i, H = j + 1) + \sum_{j' \leq j} \mathbf{P}(X = i, H = j')$$

guarantees that $\text{Space} - \text{Used} \geq \text{Need}$, thus yielding the existence of a subcoupling that is suitable for step $(i, j + 1)$.

Let us move on to the second case. We want to prove that there is a subcoupling suitable for step $(i - 1, 1)$. Let us reset the notation ($\text{Space}, \text{Need}, \text{Used}$). By assumption, we have $\text{Space} := \sum_{\beta: \beta(1) \geq i-1} \rho(\beta) \geq \mathbf{P}(X \geq i - 1) = \mathbf{P}(X = i - 1) + \sum_{i' \geq i} \mathbf{P}(X = i')$. We need to store a mass of $\text{Need} := \mathbf{P}(X = i - 1, H = 1)$, and the total mass of ν is $\text{Used} := \sum_{i' \geq i} \mathbf{P}(X = i')$. Therefore, the trivial inequality

$$\mathbf{P}(X = i - 1) \geq \mathbf{P}(X = i - 1, H = 1)$$

guarantees that $\text{Space} - \text{Used} \geq \text{Need}$, thus yielding the existence of a subcoupling that is suitable for step $(i - 1, 1)$.

At last, we need to take care of the case $(i, j) = (1, M)$. Then, the only remaining configuration to handle is the null configuration. Since it is the minimal configuration, it imposes no constraint at all on the mass repartition, so this step holds trivially. We could also say that $\text{Space} := 1$, $\text{Need} := \mathbf{P}(X = 0)$, $\text{Used} := \sum_{i \geq 1} \mathbf{P}(X = i)$, and that the inequality $\text{Space} - \text{Used} \geq \text{Need}$ clearly holds (and is an equality). This concludes the proof of Lemma 2.11. $\square$

Remark 2.12. — Lemma 2.11 can also be proved using Strassen's Theorem for monotone couplings [Str65, Lin99, Arm16], which itself admits simple proofs for finite spaces — for instance using Farkas' lemma or Hall's *Marriage Theorem*. We deemed more instructive to provide the above proof, which requires no tool at all and provides a straightforward construction of a suitable coupling.

2.3. From one column to the finite case

The goal of this section is to prove Proposition 2.13, which is a finite version of Theorem 2.3. The reason why we restrict ourselves to a finite setup is to avoid technical questions of measurability.

PROPOSITION 2.13. — *Let N be a positive integer. Let $\pi : A \rightarrow B$ be a surjective map between nonempty finite sets. Let $\mathfrak{s}$ be a section of π . Let μ and ρ be two probability measures on $[N]^A$. Assume that μ is a π -lift and that the conditions $\mathbf{A}_\rho$ and $\mathbf{B}_\mu^\rho$ are met. Then, the probability measure μ is stochastically dominated by ρ .*

We interpret A as an array with columns indexed by B , and $\mathfrak{s}$ as selecting one favourite position in each column. The condition $\mathbf{B}_\mu^\rho$ ensures that the “ $\mathfrak{s}$ -flattened” version of μ is dominated by ρ . This fact will serve as the base case of an induction proof to show that we can “unflatten” the columns one by one while keeping stochastic domination by ρ . More precisely, we will use the stochastic domination of the “flattened” measure to dominate the one before flattening. To do so, we need two tools. The first one is a way to pull back our monotone coupling, which is the object of Lemma 2.14. The second one is a way to use Lemma 2.11 in order to resample a column in a monotonic way, without breaking the rest of our coupling. This is enabled by Lemma 2.16.

It will be convenient to study couplings for probability measures. Given ξ_1 and ξ_2 two probability measures on spaces E_1 and E_2 , a *coupling* of ξ_1 and ξ_2 is a probability measure η on $E_1 \times E_2$ such that, for every i , the pushforward of η by $(x_1, x_2) \mapsto x_i$ is ξ_i .

LEMMA 2.14. — *Let $h_1 : E_1 \rightarrow E'_1$ and $h_2 : E_2 \rightarrow E'_2$ be two functions between countable sets. Let ν_1 be a probability measure on E_1 and ν_2 be a probability measure on E_2 . Let η be a coupling between the pushforward measures $\nu_1 \circ h_1^{-1}$ and $\nu_2 \circ h_2^{-1}$. Then, there is a coupling between ν_1 and ν_2 such that its pushforward by (h_1, h_2) is η .*

Proof. — Let X_1 be a ν_1 -distributed random variable and X_2 be a ν_2 -distributed random variable. For every $(x'_1, x'_2) \in E'_1 \times E'_2$ such that $\eta(x'_1, x'_2) > 0$, we consider some probability measure on $E_1 \times E_2$ with first marginal the conditional distribution of X_1 given $h_1(X_1) = x'_1$ and with second marginal the conditional distribution of X_2 given $h_2(X_2) = x'_2$ (such a probability measure exists, one can for example take the product of these measures). Lemma 2.14 follows by considering the convex combination of these measures, with convexity coefficients given by $\eta(x'_1, x'_2)$. $\square$

Lemma 2.14 can be used to reprove transitivity of stochastic domination. In this section, we use the notation $\preceq$ for stochastic domination: we write $\xi \preceq \zeta$ to mean that ξ is stochastically dominated by ζ .

COROLLARY 2.15. — *Let ξ_1, ξ_2 and ξ_3 be three probability measures on a countable ordered space E such that $\xi_1 \preceq \xi_2$ and $\xi_2 \preceq \xi_3$. Then, we have $\xi_1 \preceq \xi_3$.*

Proof. — Take ν_1 to be the distribution of a monotone coupling between ξ_1 and ξ_2 , and set $\nu_2 := \xi_3$. Let $h_1 : E \times E \rightarrow E$ be the projection on the second coordinate. As we assumed $\xi_2 \preceq \xi_3$, we can pick η the distribution of some monotone coupling between $\nu_1 \circ h_1^{-1} = \xi_2$ and ν_2 . Applying Lemma 2.14 with $h_2 = \text{id}_E$, we get a triplet of random variables (X_1, X_2, X_3) such that (X_1, X_2) is ν_1 -distributed and (X_2, X_3) is η -distributed. As both ν_1 and η are monotone couplings, the chain of inequalities $X_1 \leq X_2$ and $X_2 \leq X_3$ holds almost surely. Therefore, (X_1, X_3) is a monotone coupling between ξ_1 and ξ_3 . $\square$

The next lemma enables us to integrate monotone couplings.

LEMMA 2.16. — *Let C be a countable ordered set and D a countable set. Let f be a map from $C \times C$ to D . Let ν_1 and ν_2 be probability measures on C . Let (X_1, X_2) be a coupling of ν_1 and ν_2 . This coupling is not assumed to be monotone.*

Assume that for all $d \in D$ such that $E_d := \{f(X_1, X_2) = d\}$ has positive probability, there is a monotone (increasing) coupling (X_1^d, X_2^d) between the conditional distribution of X_1 given E_d and that of X_2 given E_d .

Then, the probability measure ν_1 is stochastically dominated by ν_2 .

Proof. — For every d such that $\mathbb{P}(E_d) > 0$, let ρ_d be the distribution of a suitable (X_1^d, X_2^d) . The distribution $\sum_d \mathbb{P}(E_d) \rho_d$ is then that of a monotone coupling between ν_1 and ν_2 . $\square$

We are now ready to establish Proposition 2.13.

Proof of Proposition 2.13. — Take $(A, B, \pi, \mathfrak{s}, \mu, \rho)$ as in Proposition 2.13. Without loss of generality, we assume that $B = \llbracket 1, K \rrbracket := \{1, \dots, K\}$. Our goal is to prove that $\rho \succcurlyeq \mu$. To do so, we will work by considering “flattened” versions of μ , and recursively “unflatten” them in a proper coupling.

We define a sequence of measures $(\mu_n)_{n \in \llbracket 0, K \rrbracket}$ by induction from K until 0. We set $\mu_K := \mu$. For $n \in \llbracket 1, K \rrbracket$, let $f_n : \llbracket 0, N \rrbracket^A \rightarrow \llbracket 0, N \rrbracket^A$ be defined as follows

- if $\pi(a) \neq n$, then $f_n(y)_a = y_a$,
- if $a = \mathfrak{s}(n)$, then $f_n(y)_a = \max_{a' \in \pi^{-1}(n)} y_{a'}$,
- otherwise, $f_n(y)_a = 0$.

For $n \in \llbracket 0, K-1 \rrbracket$, the measure μ_n is then defined as the pushforward of μ_{n+1} by f_{n+1} . By definition and because of Assumption $\mathbf{B}_\mu^\rho$, we have $\mu_0 \preccurlyeq \rho$. Let us now establish the induction step $(\mu_n \preccurlyeq \rho) \Rightarrow (\mu_{n+1} \preccurlyeq \rho)$.

Let $n \in \llbracket 0, K-1 \rrbracket$ be such that $\mu_n \preccurlyeq \rho$. Our goal is to show that $\mu_{n+1} \preccurlyeq \rho$. Let us apply Lemma 2.14 with $E_1 = E'_1 = E_2 = E'_2 = \llbracket 0, N \rrbracket^A$, $h_1 = f_{n+1}$, h_2 the identity map, $\nu_1 = \mu_{n+1}$, $\nu_2 = \rho$, and η a monotone coupling between μ_n and ρ (such an η exists as we assumed $\mu_n \preccurlyeq \rho$). This enables us to take a coupling (Y, Z) of μ_{n+1} and ρ such that the following is true. If ℓ_{n+1} is set to be $\arg\max_{a \in \pi^{-1}(n+1)} (Y_a)$ when it is well defined and $\mathfrak{s}(n+1)$ otherwise then, we almost surely have:

- $Y_a = 0$ for every $a \in \pi^{-1}(n+1) \setminus \{\ell_{n+1}\}$,
- $Z_a \geq Y_a$ for every $a \neq \ell_{n+1}$,
- $Z_{\mathfrak{s}(n+1)} \geq Y_{\ell_{n+1}} = \max_{a \in \pi^{-1}(n+1)} Y_a$.

Note that this coupling has no reason to be monotone, so more work is needed to deduce stochastic domination.

Let us consider the forgetful function $g_{n+1} : \llbracket 0, N \rrbracket^A \times \llbracket 0, N \rrbracket^A \rightarrow \llbracket 0, N \rrbracket^{A \setminus \pi^{-1}(n+1)}$ defined by $((y_a)_{a \in A}, (z_a)_{a \in A}) \mapsto (z_a)_{a \notin \pi^{-1}(n+1)}$. We want to apply Lemma 2.16 with g_{n+1} , Z and Y . Before doing so, let us gather a few observations:

- (1) Recall that we almost surely have domination outside $\pi^{-1}(n+1)$. Therefore, for any given $d \in \llbracket 0, N \rrbracket^{A \setminus \pi^{-1}(n+1)}$, conditioned on the event $E_d := \{g_{n+1}(Y, Z) = d\}$, the random variable Y is almost surely dominated by d outside $\pi^{-1}(n+1)$.

- (2) Thanks to this and Lemma 2.14, for any given d , we can find a monotone coupling (Y^d, Z^d) as soon as there is stochastic domination of the conditional distribution of $(Y_a)_{a \in \pi^{-1}(n+1)}$ given E_d by that of $(Z_a)_{a \in \pi^{-1}(n+1)}$ given E_d .
- (3) By construction, the distribution of $(Y_a)_{a \in \pi^{-1}(n+1)}$ conditioned on E_d is a π -lift.
- (4) As the inequality $\max_{a \in \pi^{-1}(n+1)} Y_a \leq Z_{\mathfrak{s}(n+1)}$ holds almost surely, it also holds almost surely conditioned on E_d . In particular, the distribution of $\max_{a \in \pi^{-1}(n+1)} Y_a$ conditioned on E_d is stochastically dominated by that of $Z_{\mathfrak{s}(n+1)}$ conditioned on E_d .
- (5) Finally, Assumption $\mathbf{A}_\rho$ assures us that for any $a \in \pi^{-1}(n+1)$, conditioned on E_d , the distribution of Z_a stochastically dominates that of $Z_{\mathfrak{s}(n+1)}$.

Due to Observations (4) and (5), transitivity of stochastic domination ensures that for any given d and every $a \in \pi^{-1}(n+1)$, conditioned on E_d , the distribution of Z_a stochastically dominates that of $\max_{\pi^{-1}(n+1)} Y$. This allows us to use Lemma 2.11, which yields that the distribution of $(Y_a)_{a \in \pi^{-1}(n+1)}$ conditioned on E_d is stochastically dominated by that of $(Z_a)_{a \in \pi^{-1}(n+1)}$ conditioned on E_d . Coupled with Observation (2), this result implies that conditioned on any event of the form E_d , the distribution of Z stochastically dominates that of Y . We can now use Lemma 2.16 to conclude that the distribution of Z (in other words ρ) stochastically dominates that of Y (that is, μ_{n+1}). The induction step is thus established. By induction, we get $\rho \succcurlyeq \mu_K = \mu$. $\square$

2.4. From the finite setup to the general case

Let us now use Proposition 2.13 to establish Theorem 2.3. Naturally, the strategy is to discretise the problem, apply Proposition 2.13, and then take limits.

For $n \geq 0$, let f_n be the function defined on $[0, \infty]$ by

$$f_n : x \mapsto \max \left\{ \frac{k}{n} : \frac{k}{n} \leq x \text{ and } 0 \leq k \leq n^2 \right\}.$$

Each f_n is weakly increasing and the sequence (f_n) converges pointwise to the identity of $[0, \infty]$.

Now, let us enumerate $A = \{a_1, a_2, \dots\}$ in such a way that for all i and j , we have $a_i = \mathfrak{s} \circ \pi(a_j) \Rightarrow i \leq j$. Such an enumeration can for instance be obtained by starting with an arbitrary enumeration and then, for each $b \in B$, swapping $\mathfrak{s}(b)$ with the first element of $\pi^{-1}(b)$ to appear in this enumeration. For $n \geq 0$, define the map $g_n : [0, \infty]^A \rightarrow [0, \infty]^A$ by $g_n(x) : a_i \mapsto f_n(x_{a_i}) \mathbb{1}_{i \leq n}$. Each map g_n is weakly increasing and the sequence (g_n) converges pointwise to the identity of $[0, \infty]^A$. For $n \geq 0$, define the probability measure μ_n (resp. ρ_n) to be the pushforward of μ (resp. ρ) by g_n . In other words, if a random variable Y has distribution μ (resp. ρ), then the distribution of $g_n(Y)$ is denoted by μ_n (resp. ρ_n).

Observe that for every n , the measures μ_n and ρ_n satisfy the assumptions of Theorem 2.3, as we assumed that it was the case for μ and ρ . On top of that, (μ_n, ρ_n) fits the finite setup: the entries outside of $\{a_1, \dots, a_n\}$ are all null and irrelevant; also notice that if Proposition 2.13 holds for all $\{0, 1, \dots, N\}$, then it holds for all $\{0, 1, \dots, N^2\}$, hence also for labels in $\{\frac{k}{N}, : 0 \leq k \leq N^2\}$. Therefore, we can apply Proposition 2.13 and get some monotone coupling ν_n between μ_n and ρ_n .

The space $[0, \infty]^A$, endowed with the product topology, is compact Hausdorff, thus so is its Cartesian square. The set $\{(x, y) \in [0, \infty]^A \times [0, \infty]^A : \forall a \in A, x_a \leq y_a\}$ is closed inside this square, hence compact itself. Therefore, the space of monotone couplings, endowed with the smallest topology making continuous the integration against all continuous bounded functions, is compact. We can thus pick ν some accumulation point of the sequence (ν_n) : we may write $\nu = \lim_{k \rightarrow \infty} \nu_{n_k}$. By construction, ν -almost every (x, y) satisfies $\forall a \in A, x_a \leq y_a$. What remains to check is that ν has the correct marginals.

By symmetry, it suffices to take care of the first marginal. It suffices to check that for every continuous (automatically bounded) $h : [0, \infty]^A \rightarrow \mathbb{R}$, we have $\int h(x) d\nu(x, y) = \int h(x) d\mu(x)$. As h is continuous bounded, we know that

$$\int h(x) d\nu(x, y) = \lim_{k \rightarrow \infty} \int h(x) d\nu_{n_k}(x, y).$$

As the first marginal of ν_{n_k} is μ_{n_k} , we have $\int h(x) d\nu(x, y) = \lim_{k \rightarrow \infty} \int h(x) d\mu_{n_k}(x)$. It thus suffices to establish that (μ_n) converges to μ . To see this, let Y be a μ -distributed random variable: as $g_n(Y)$ converges pointwise to Y , the distribution of $g_n(Y)$ converges to that of Y , which is exactly what we had left to prove. This concludes the proof of Theorem 2.3. $\square$

2.5. Proofs of the corollaries

This section is dedicated to deriving Corollaries 2.4 and 2.10 from Theorem 2.3.

Proof of Corollary 2.4. — Without loss of generality, we assume that A and B are disjoint. Let $\tilde{A} := A \sqcup B$. Let $\tilde{\pi} : \tilde{A} \rightarrow B$ be defined as π on A and as the identity on B . For every $b \in B$, pick some probability measure $\tilde{\rho}_b$ on $\tilde{\pi}^{-1}(b) = \pi^{-1}(b) \sqcup \{b\}$ in such a way that the $\pi^{-1}(b)$ -marginal of $\tilde{\rho}_b$ is ρ_b and its b -marginal is ν_b . This is indeed possible, for instance by setting $\tilde{\rho}_b := \rho_b \otimes \nu_b$. We consider the section $\mathfrak{s}$ of $\tilde{\pi}$ given by the inclusion of B as a subset of $\tilde{A} = A \sqcup B$. Let $\tilde{\rho} := \bigotimes_b \tilde{\rho}_b$ and let $\tilde{\mu}$ be the pushforward of μ by the map $g : [0, \infty]^A \rightarrow [0, \infty]^{A \sqcup B}$ given by

$$g(x)_c = \begin{cases} x_c & \text{if } c \in A, \\ 0 & \text{if } c \in B. \end{cases}$$

We want to apply Theorem 2.3 to $(\tilde{A}, B, \tilde{\pi}, \mathfrak{s}, \tilde{\rho}, \tilde{\mu})$.

It is the case that $\tilde{\pi}$ is a surjective map between nonempty countable sets. The probability measure $\tilde{\mu}$ is indeed a $\tilde{\pi}$ -lift, as the map g attributes automatically the value zero to all entries that were not originally present in A . Assumption $\mathbf{A}_{\tilde{\rho}}$ holds: indeed, by independence, there is no conditioning to make, and we assumed that for every b and every $a \in \pi^{-1}(b)$, the probability measure ν_b is stochastically dominated by the a -marginal of ρ_b . Besides, the last assumption of Corollary 2.4 guarantees that the pushdown of $\tilde{\mu}$ is stochastically dominated by $\bigotimes_b \nu_b$, which gives Assumption $\mathbf{B}_{\tilde{\mu}}^{\tilde{\rho}}$.

We can thus apply Theorem 2.3, yielding that $\tilde{\mu}$ is stochastically dominated by $\tilde{\rho}$. Restricting a corresponding coupling to entries in A instead of the whole $A \sqcup B$ establishes the stochastic domination of μ by ρ , hereby completing the proof of Corollary 2.4. $\square$

Proof of Corollary 2.10. — Let $\mathfrak{s}$ be any section of π . It suffices to check Assumptions $\mathbf{A}_\rho$ and $\mathbf{B}_\mu^\rho$. Observe that for every b , the fixator⁽⁵⁾ of $A \setminus \pi^{-1}(b)$ can be identified with G_b and acts transitively on $\pi^{-1}(b)$. Therefore, the conditional distribution appearing in $\mathbf{A}_\rho$ depends only on b , not on $a \in \pi^{-1}(b)$. Assumption $\mathbf{A}_\rho$ follows, as a probability measure is always stochastically dominated by itself. As for $\mathbf{B}_\mu^\rho$, it follows from our third assumption. $\square$

Remark 2.17. — In the previous argument, the condition that really matters as far as $\mathbf{A}_\rho$ is concerned is the following transitive-fixator condition. Consider the group Sym_ρ of all $\sigma \in \mathfrak{S}_\pi$ such that the pushforward of ρ by $x \mapsto \sigma \cdot x$ is equal to ρ . The condition then reads as follows: Sym_ρ satisfies that, for every b , the fixator of $A \setminus \pi^{-1}(b)$ acts transitively on $\pi^{-1}(b)$. This condition is indeed satisfied as soon as ρ is sufficiently symmetric. Let us explain why, conversely, the transitive-fixator condition implies that ρ is sufficiently symmetric. For each b , using the transitive-fixator condition yields a transitive subgroup G_b of the permutations of $\pi^{-1}(b)$ such that letting it act naturally on $\pi^{-1}(b)$ and as the identity elsewhere preserves the probability measure ρ . We deduce that ρ is left invariant under the action of the group $\bigoplus_b G_b$ — which is defined to be the group of the elements (g_b) of $\prod_b G_b$ such that all but finitely many g_b 's are the identity. Finally, as probability measures on $[0, \infty]^A$ are characterised by their finite-dimensional marginals, invariance under $\bigoplus_b G_b$ implies invariance under the larger group $\prod_b G_b$. In the end, the following three conditions are equivalent: the transitive-fixator condition, the existence of transitive G_b 's such that Sym_ρ contains $\bigoplus_b G_b$, and the same with $\prod_b G_b$ instead of $\bigoplus_b G_b$. We have chosen to define “sufficiently symmetric” by using $\prod_b G_b$ because this definition is simple and because it makes clear that the horizontal marginal is well defined.

3. Variations on Theorem 2.3

In this section, we investigate in which directions Theorem 2.3 may or may not be improved. This section is not used in subsequent sections.

3.1. Lifting several variables per column

In this subsection, we do not try to give a maximally general statement fitting the purpose announced in the title (this would probably be hard to read). We will see that a new phenomenon occurs, and we have chosen to focus on the simplest statement exhibiting this phenomenon.

The purpose will no longer be to prove that two fully specified probability measures satisfy stochastic domination. Instead, the lifting process will leave some room for indeterminacy, and the theorem will state that there is a way to proceed such that the lifting constraints are satisfied and stochastic domination holds.

⁽⁵⁾When a group G acts on a set E , the *fixator* of a subset P of E is the group of all g 's satisfying $\forall x \in P, g \cdot x = x$. Here, we consider $\prod_b G_b$ acting on A .

A convenient way to state our result uses the notion of extension of probability spaces. An *extension* of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the data of a probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ together with a measure-preserving map $\varphi : \Omega^* \rightarrow \Omega$, i.e. a measurable map such that the pushforward of $\mathbb{P}^*$ by φ is $\mathbb{P}$. Any random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ can be interpreted in the extended space by considering $X^* := X \circ \varphi$. By abuse of notation, if the context is unambiguous, we may continue to use the notation X instead of X^* for extended random variables.

THEOREM 3.1. — *Let $\pi : A \rightarrow B$ be a surjective map between nonempty countable sets. Let $p \in [0, 1]$. For every $b \in B$, let X_b and $X_b^\dagger$ be Bernoulli random variables of parameter p . Assume that the collection of random variables formed by all the X_b 's and all the $X_b^\dagger$'s is independent. Let S and $S^\dagger$ be random sections of π such that, for every $b \in B$, we almost surely have $S(b) \neq S^\dagger(b)$.*

Then, up to extending the probability space, there are $\{0, 1\}$ -valued random variables Y_a and Z_a such that the following conditions hold:

- (1) (a) for every a , we have $a \notin \{S \circ \pi(a), S^\dagger \circ \pi(a)\} \Rightarrow Y_a = 0$,
 (b) for every b satisfying $X_b = 0$ and $X_b^\dagger = 0$, we have $Y_{S(b)} = 0$ and $Y_{S^\dagger(b)} = 0$,
 (c) for every b satisfying $X_b = 1$ and $X_b^\dagger = 0$, we have $Y_{S(b)} = 1$ and $Y_{S^\dagger(b)} = 0$,
 (d) for every b satisfying $X_b = 1$ and $X_b^\dagger = 1$, we have $Y_{S(b)} = 1$ and $Y_{S^\dagger(b)} = 1$,
 (e) for every b satisfying $X_b = 0$ and $X_b^\dagger = 1$, exactly one a in $\{S(b), S^\dagger(b)\}$ satisfies $Y_a = 1$,
 (2) the distribution of (Z_a) is $\text{Bernoulli}(p)^{\otimes A}$,
 (3) for all a , we have $Y_a \leq Z_a$.

Remark 3.2. — The indeterminacy is captured by (1e), where we have the possibility to lift the “weak bit” $X_b^\dagger$ in $\{S(b), S^\dagger(b)\}$ but not to decide where. On the contrary, (1c) states that the “strong bit” X_b can be lifted exactly where one wants. (1a) and (1b) are useless, in the sense that if some (Y_a) satisfies the conclusion apart from these items, setting $Y'_a = Y_a \mathbb{1}_{a \in \{S \circ \pi(a), S^\dagger \circ \pi(a)\}} \mathbb{1}_{(X_{\pi(a)}, X_{\pi(a)}^\dagger) \neq (0,0)}$ will satisfy the full conclusion for free. Likewise, in (1e), it is not “exactly one” that truly matters but “at least one”. We decided to include these conditions in order to try and construct Y_a as explicitly as before, and see exactly where it fails. If we remove all useless items from the conclusion, then we can further take $Y_a = Z_a$.

Our motivation for investigating this was to get a new proof of Theorem 6.1 using techniques in the spirit of Theorem 1.1. Theorem 6.1, first obtained in [MS19], is a version of Theorem 1.5 where, under additional assumptions, one gets the strict inequality. The general idea there is to use the fact that π is (at least 2)-to-1 to try and “get a second chance”. We have not managed to use Theorem 3.1 or variations of it to obtain such a proof. Still, Section 6 provides a new proof of this result, based on explorations and couplings.

The strategy to prove Theorem 3.1 follows an easy scheme: from some $\pi : A \rightarrow B$ and probability measures μ and ρ , construct new $\tilde{\pi} : \tilde{A} \rightarrow B$ and probability measures

$\tilde{\mu}$ and $\tilde{\rho}$; then apply Theorem 2.3 to these new objects. It will be easy for the reader to craft their own variations.

Proof of Theorem 3.1. — Let $\tilde{A} := \{(a_1, a_2) \in A^2 : a_1 \neq a_2 \text{ and } \pi(a_1) = \pi(a_2)\}$. Let $\tilde{\pi} : \tilde{A} \rightarrow B$ be defined by mapping any (a_1, a_2) to the element $\pi(a_1)$, which is also $\pi(a_2)$. The set $\tilde{A}$ is indeed countable. As $\tilde{S} : b \mapsto (S(b), S^\dagger(b))$ is a random section of $\tilde{\pi}$, such sections exist, entailing surjectivity of $\tilde{\pi}$ (and nonemptiness of $\tilde{A}$, as B is nonempty).

For every $b \in B$, let $\tilde{X}_b := 2X_b + X_b^\dagger$. For every $b \in B$, let $\rho_b := \text{Bernoulli}(p)^{\otimes \pi^{-1}(b)}$, and let $\tilde{\rho}_b$ be the pushforward of the measure ρ_b by the map

$$(x_a)_{a \in \pi^{-1}(b)} \longmapsto (2x_{a_1} + x_{a_2})_{(a_1, a_2) \in \pi^{-1}(b)}.$$

As the $\tilde{X}_b$'s are independent, we can apply Corollary 2.1 to $\tilde{X}_b$, $\tilde{S}(b)$, and $\tilde{\rho}_b$. Therefore, letting $\tilde{Y}_{\mathbf{a}} := \tilde{X}_{\tilde{\pi}(\mathbf{a})} \mathbb{1}_{\tilde{S} \circ \tilde{\pi}(\mathbf{a}) = \mathbf{a}}$, we get that the distribution of $(\tilde{Y}_{\mathbf{a}})$ is stochastically dominated by $\bigotimes_{b \in B} \tilde{\rho}_b$.

Up to extending our probability space, we can thus define random variables Z_a , for $a \in A$, in such a way that:

- the random variables Z_a are i.i.d. with Bernoulli distribution of parameter p ,
- for every $\mathbf{a} = (a_1, a_2) \in \tilde{A}$, we have $\tilde{S} \circ \tilde{\pi}(\mathbf{a}) = \mathbf{a} \Rightarrow \tilde{Y}_{\mathbf{a}} \leq 2Z_{a_1} + Z_{a_2}$.

We defer justification of this point to the end of the proof. For now, observe that the second condition can be rewritten as

$$\tilde{S} \circ \tilde{\pi}(\mathbf{a}) = \mathbf{a} \implies (2X_{\tilde{\pi}(\mathbf{a})} + X_{\tilde{\pi}(\mathbf{a})}^\dagger) \mathbb{1}_{\tilde{S} \circ \tilde{\pi}(\mathbf{a}) = \mathbf{a}} \leq 2Z_{a_1} + Z_{a_2}$$

holding for every $\mathbf{a}$, which in turn can be written as

$$2Z_{S(b)} + Z_{S^\dagger(b)} \geq 2X_b + X_b^\dagger$$

holding for every b .

Observe that, to establish the result, there is no need to speak of the Y_a 's. It suffices to prove (1a)–(1e) with lower bounds on Z_a instead of equalities for Y_a . More precisely, it suffices to prove the following three conditions:

- for every b satisfying $X_b = 1$ and $X_b^\dagger = 0$, we have $Z_{S(b)} = 1$,
- for every b satisfying $X_b = 1$ and $X_b^\dagger = 1$, we have $Z_{S(b)} = 1$ and $Z_{S^\dagger(b)} = 1$,
- for every b satisfying $X_b = 0$ and $X_b^\dagger = 1$, at least one a in $\{S(b), S^\dagger(b)\}$ satisfies $Z_a = 1$,

Let us establish that these three conditions hold. When $(X_b, X_b^\dagger) = (1, 1)$, we have $2Z_{S(b)} + Z_{S^\dagger(b)} \geq 3$, which gives $Z_{S(b)} = Z_{S^\dagger(b)} = 1$, as desired. When $(X_b, X_b^\dagger) = (1, 0)$, we have $2Z_{S(b)} + Z_{S^\dagger(b)} \geq 2$, which indeed guarantees that $Z_{S(b)} = 1$. At last, when $(X_b, X_b^\dagger) = (0, 1)$, we have $2Z_{S(b)} + Z_{S^\dagger(b)} \geq 1$, which indeed guarantees that there is at least one 1 among $Z_{S(b)}$ and $Z_{S^\dagger(b)}$.

It remains to justify how to implement stochastic domination via extensions, which is in the spirit of Lemma 2.14. Let $\mathbf{Z} = (Z_a)_{a \in A}$ be a family of independent Bernoulli random variables of parameter p , defined on some probability space. On this same probability space, we can also define $\mathbf{W} = (Z_{a_1}, Z_{a_2})$, which has distribution $\bigotimes_{b \in B} \tilde{\rho}_b$. As the random variables $\mathbf{Z}$ and $\mathbf{W}$ take values in Polish spaces, we can define a probability kernel κ such that the conditional distribution of $\mathbf{Z}$ given $\mathbf{W}$ is $\kappa_{\mathbf{W}}$.

On our original probability space, we have the random variable $\mathbf{Y} = (\tilde{Y}_{\mathbf{a}})$. By Corollary 2.1, we know that there are yet another probability space and random variables $\mathbf{Y}'$ and $\mathbf{W}'$ defined on it such that $\mathbf{Y}'$ has the same distribution as $\mathbf{Y}$, $\mathbf{W}'$ the same distribution as $\mathbf{W}$, and the inequality $\mathbf{Y}' \leq \mathbf{W}'$ holds almost surely. Let κ' be a probability kernel such that the conditional distribution of $\mathbf{W}'$ given $\mathbf{Y}'$ is $\kappa'_{\mathbf{Y}'}$. If $\mathbf{Z}$ takes values in E , then a suitable extension is given by $\Omega^* := \Omega \times E$, the probability measure $\mathbb{P}^* := \iint \delta_\omega \otimes \kappa_{\mathbf{w}} d\kappa'_{\mathbf{Y}(\omega)}(\mathbf{w}) d\mathbb{P}(\omega)$, and the map $\varphi : (\omega, \mathbf{z}) \mapsto \omega$. The random variable $\mathbf{Z}^* : (\omega, \mathbf{z}) \mapsto \mathbf{z}$ behaves as desired. $\square$

Counterexample 3.3. — Please note that Theorem 3.1 does not hold if the condition of the subtle (1e) is replaced by “ $Y_{S(b)} = 0$ and $Y_{S^\dagger(b)} = 1$ ”. Indeed, a counterexample is possible with $B = \{o\}$ and $A = \{1, 2\}$. Simply take $(S(o), S^\dagger(o)) = (1, 2)$ if $X_o = 1$ and $(S(o), S^\dagger(o)) = (2, 1)$ otherwise. Then, the parameter of the Bernoulli random variable lifted at 1 via this too strong procedure would have parameter $1 - (1 - p)^2$, which is strictly larger than p as soon as $p \in (0, 1)$.

3.2. Beyond Conditions $\mathbf{A}_\rho$ $\mathbf{B}_\mu^\rho$

In this section, we investigate a generalisation of Theorem 2.3 that may seem reasonable but turns out to be wrong. We show that if we drop Assumptions $\mathbf{A}_\rho$, $\mathbf{B}_\mu^\rho$ and assume something weaker instead, then Theorem 2.3 becomes false. The weak assumption we consider is:

$\mathbf{C}_\mu^\rho$ for every section s , the pushdown of μ is stochastically dominated by the s -marginal of ρ , i.e. by the pushforward of ρ by $x \mapsto (x_{s(b)})_{b \in B}$.

Let us build an example where this condition is satisfied but where μ is not stochastically dominated by ρ .

Our sets A and B are defined by: $A := \{1, 2, 3\} \times \{1, 2, 3\}$ and $B := \{1, 2, 3\}$, and the function π is the projection on the second coordinate. We will represent elements of $\{0, 1\}^A$ and $\{0, 1\}^B$ respectively by 3×3 matrices with coefficients in $\{0, 1\}$ and horizontal vectors of size 3 with coefficients in $\{0, 1\}$. Let μ be the uniform measure on the set given by these four matrices:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

In particular, the pushdown of μ is the uniform measure the following vectors:

$$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}.$$

Similarly let ρ be the uniform probability measure on these four matrices:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Among these matrices, only the matrices

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

are larger than either of the three matrices

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

As a result, μ cannot be stochastically dominated by ρ . We still need to show that $\mathbf{C}_\mu^\rho$ holds. By symmetry, we may assume that $s(1) = 1$. We still have nine possible sections to check, which is done in Table 3.1. Every row corresponds to a given section and the coupling between μ and ρ is given by which matrix is in which column.

Remark 3.4. — Likewise, we can see that if we try to keep Assumptions $\mathbf{A}_\rho$ and $\mathbf{B}_\mu^\rho$ and go beyond $[0, \infty]$ -labels, we get stopped right away. Take labels to live in $\{0, 1\}^2$, endowed with its product ordering. Let B be a singleton and $A := \{1, 2\}$. Let μ be the uniform probability measure on the set given by these two configurations:

$$\begin{bmatrix} 10 \\ 00 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 00 \\ 01 \end{bmatrix}$$

and let ρ be the uniform probability measure on the two configurations

$$\begin{bmatrix} 10 \\ 01 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 01 \\ 10 \end{bmatrix}.$$

Then all assumptions hold but the conclusion does not. This minimal counterexample can be implemented in any label space that is not totally ordered and contains a global minimum (which plays the role of the element 0 of $[0, \infty]$).

It is legitimate to ask whether such a counterexample can be made to satisfy the following more stringent version of $\mathbf{C}_\mu^\rho$:

$\tilde{\mathbf{C}}_\mu^\rho$ for every section s , the pushdown of μ is equal to the s -marginal of ρ , i.e. equal to the pushforward of ρ by $x \mapsto (x_{s(b)})_{b \in B}$.

Such a counterexample was found by Vibhu Dalal and is given in Appendix A.2.

Table 3.1. This table justifies that our construction satisfies Assumption $\mathbf{C}_\mu^\rho$.

[illegible]

4. Application of Theorem 1.1 to percolation

The purpose of this section is to use Theorem 1.1 to reprove and generalise Theorem 1.5. Site percolation, bond percolation, and fibrations have been defined in the introduction. Recall that, given a graph $\mathcal{G}$, we write $V_{\mathcal{G}}$ for its set of vertices and $E_{\mathcal{G}}$ for its set of edges.

4.1. Site percolation encompasses bond percolation

The bond-version of Theorem 1.5 follows from its site-version. Indeed, recall that given a countable locally finite graph $\mathcal{G}$, one can build another countable locally finite graph, $\mathcal{G}^*$, by setting $V_{\mathcal{G}^*} = E_{\mathcal{G}}$ and declaring two elements of $V_{\mathcal{G}^*}$ to be adjacent if and only if the corresponding edges in $\mathcal{G}$ have exactly one vertex in common. Then, we have $p_c^{\text{site}}(\mathcal{G}^*) = p_c^{\text{bond}}(\mathcal{G})$.

It remains to take care of π . To do so, introduce $\mathcal{L}_{\pi}^*$ defined as follows. Its vertex-set is given by the set of all edges of $\mathcal{L}$ that are mapped by π to edges of $\mathcal{S}$. As for the edge-structure on $\mathcal{L}_{\pi}^*$, it is simply that induced by $\mathcal{L}^*$: in other words, two vertices of $\mathcal{L}_{\pi}^*$ are adjacent if and only if, seen as edges of $\mathcal{L}$, they have exactly one vertex in common. Every fibration π from $\mathcal{L}$ to $\mathcal{S}$ induces a fibration from $\mathcal{L}_{\pi}^*$ to $\mathcal{S}^*$. As $\mathcal{L}_{\pi}^*$ is a subgraph of $\mathcal{L}^*$, we have $p_c^{\text{site}}(\mathcal{L}^*) \leq p_c^{\text{site}}(\mathcal{L}_{\pi}^*)$. Combining these observations with the equalities $p_c^{\text{site}}(\mathcal{L}^*) = p_c^{\text{bond}}(\mathcal{L})$ and $p_c^{\text{site}}(\mathcal{S}^*) = p_c^{\text{bond}}(\mathcal{S})$ reduces the case of bond percolation to that of site percolation.

4.2. Measurability issues

To complete the proof of Theorem 1.5 via Theorem 1.1 sketched in the introduction, it remains to take care of the measurability of κ and $\tilde{\kappa}$. There are two rather easy ways to take care of this: we may either reduce the problem to a finite setup where there is no measurability to check or indeed check the measurability we need.

4.2.1. Reduction to a finite setup

Let $p \in [0, 1]$ and let us consider p -site percolation on both $\mathcal{L}$ and $\mathcal{S}$. It suffices to prove that for every $x \in V_{\mathcal{L}}$, the probability that x is connected to infinity in $\mathcal{L}$ is larger than or equal to the probability that $\pi(x)$ is connected to infinity in $\mathcal{S}$. Therefore, it suffices to prove that, for every n , the probability that there is an open self-avoiding path⁽⁶⁾ starting at x of length n is larger than or equal to the probability that there is an open self-avoiding path starting at $\pi(x)$ of length n . We are then able to conclude, as this inequality results from Proposition 2.13, with A the n -ball centred at x and B the n -ball centred at $\pi(x)$.

We do not claim that π induces a fibration from A to B , simply that every self-avoiding path of length n started at $\pi(x)$ (which necessarily lies in B) can be lifted to a self-avoiding path of length n in x (which necessarily lies in A).

⁽⁶⁾We say that a path is *self-avoiding* if no vertex is visited more than once. We take the *length* of a self-avoiding path to be the number of edges it uses, which is the number of vertices it visits minus one.

Remark 4.1. — If we only care about Theorem 1.5, we do not need to go through Section 2.4, as Proposition 2.13 suffices.

4.2.2. Checking measurability

We can enumerate all vertices of $\mathcal{S}$ and do the same for $\mathcal{L}$. This provides a well-ordering of these vertex-sets. Then, by lexicographic ordering, we get total orders on the set of $\mathbb{N}$ -indexed self-avoiding paths in $\mathcal{S}$ and $\mathcal{L}$, respectively. These orders are such that any nonempty closed⁽⁷⁾ set admits a smallest element. A percolation configuration being fixed, the set of *open* $\mathbb{N}$ -indexed self-avoiding paths is topologically closed. Likewise, some $\mathbb{N}$ -indexed self-avoiding paths in $\mathcal{S}$ being fixed, the set of all its lifts is closed. Therefore, picking κ to be the smallest open such path (when some infinite open path exists) and $\tilde{\kappa}$ to be the smallest lift of κ provides a well-defined construction. The construction of κ is indeed measurable because, by local finiteness, an open finite self-avoiding path admits an infinite open extension if and only if it admits finite open extensions of all sizes. Likewise, the construction of $\tilde{\kappa}$ is measurable: define $\tilde{\kappa}_0$ to be the smallest vertex in $\pi^{-1}(\kappa_0)$, then κ_1 to be the smallest neighbour of $\tilde{\kappa}_0$ in $\pi^{-1}(\kappa_1)$, etc.

4.3. Generalisation of Theorem 1.5

Actually, we can prove a more general result.

THEOREM 4.2. — *Let $\mathcal{L}$ and $\mathcal{S}$ be countable locally finite graphs. Let π be a surjective map from $V_{\mathcal{L}}$ to $V_{\mathcal{S}}$. Assume that there is a measurable way to assign to every infinite self-avoiding path $\gamma : \mathbb{N} \rightarrow V_{\mathcal{S}}$ some infinite self-avoiding path $\tilde{\gamma} : \{n_{\gamma}, n_{\gamma} + 1, \dots\} \rightarrow V_{\mathcal{L}}$ such that for every $m \geq n_{\gamma}$, we have $\pi(\tilde{\gamma}_m) = \gamma_m$.*

For every vertex v in $V_{\mathcal{S}}$, let $p_v \in [0, 1]$. On $\mathcal{S}$, consider the random configuration where each vertex is kept independently with probability p_v .

Let $\mathfrak{X}$ be some random subset of $V_{\mathcal{L}}$. We assume that the family of random variables $(\mathfrak{X} \cap \pi^{-1}(v))_{v \in V_{\mathcal{S}}}$ is independent. Besides, for every $x \in V_{\mathcal{L}}$, assume that $\mathbb{P}(x \in \mathfrak{X}) \geq p_v$.

Then, the probability that there is an infinite path of retained vertices in $\mathcal{L}$ is larger than or equal to the probability that there is an infinite path of retained vertices in $\mathcal{S}$.

Remark 4.3. — This theorem has two kinds of assumptions: the first paragraph is of geometric nature while the second and third paragraphs are probabilistic. The techniques of [BS96] can be used to handle this level of probabilistic generality: we can get a classical proof by exploration of the above theorem if the geometric assumption is replaced by the assumption that there is a fibration from $\mathcal{L}$ to $\mathcal{S}$. However, the techniques of [BS96] really need the stronger geometric condition of having a fibration: we need to be able to continue lifting our path wherever the past exploration led us, which imposes us to ask for the possibility to lift edges $\{u, v\}$

⁽⁷⁾For the topology of pointwise convergence, where vertex-sets are seen as discrete.

to all $x \in \pi^{-1}(u)$. Theorem 4.2 relaxes this condition, only asking for every infinite self-avoiding path to admit at least one lift, which may start where it wants to (and can be measurably picked). The condition is actually relaxed further, by allowing to forget the beginning of the path.

Proof. — Apply our new proof of Theorem 1.5 but use Corollary 2.4 instead of Theorem 1.1. Regarding the measurability of κ , the argument of Section 4.2.1 does not apply anymore but that of Section 4.2.2 still applies. As for $\tilde{\kappa}$, we assumed the measurability we need. $\square$

As explained in Remark 4.3, the main gain in Theorem 4.2 is about geometry. Convincing examples appear to be lacking so far: in the situations we came up with, we could use a few tricks to get back to the case of fibrations. Still, we find conceptual value in not having to *assume* that we are in the fibration setup or to use ad hoc tricks. This conceptual value says more about the nature of our arguments and our understanding than about the list of concrete examples we can handle.

Regarding the probabilistic gain, we indeed extend the scope of Theorem 1.5, but in a way that Benjamini and Schramm could have covered with their techniques if they wanted to. It is easy to come up with diverse examples. Let us provide one that we find natural and interesting.

Example 4.4. — Assume that every $\pi^{-1}(x)$ has cardinality 2 and that all p_v 's are equal to $1/2$. Independently, for each $x \in V_{\mathcal{S}}$, prescribe $\mathfrak{X} \cap \pi^{-1}(x)$ to be a uniformly chosen singleton included in $\pi^{-1}(x)$. In other words, the percolation model on $\mathcal{L}$ is defined by picking independently and uniformly in each π -fibre one vertex that we keep, letting the other aside.

Using Theorem 2.3 instead of Corollary 2.4 would lead to an even more general version of Theorem 4.2. Using Corollary 2.10 would lead to a variation of it, and other examples. A priori, the probabilistic gain in this generalisation and this variation cannot be obtained by direct adaptation of [BS96], as their exploration argument relies heavily on independence.

Let us conclude Section 4.3 by an amusing observation. Let $\mathcal{S}$ be a countable locally finite graph. We define the graph $\mathcal{L}$ to be a collection of disjoint, noninteracting infinite rays, one per $\mathbb{N}$ -indexed self-avoiding path in $\mathcal{S}$. There is a surjective map from $V_{\mathcal{L}}$ to $V_{\mathcal{S}}$ satisfying the condition of Theorem 4.2. Therefore, by Theorem 4.2, it seems that $p_c^{\text{site}}(\mathcal{S}) \geq p_c^{\text{site}}(\mathcal{L})$. But we have $p_c^{\text{site}}(\mathcal{L}) = 1$, as it consists of a bunch of rays and each ray has $p_c = 1$. We seem to get into a contradiction, namely that every $\mathcal{S}$ satisfies $p_c^{\text{site}}(\mathcal{S}) = 1$. It turns out that we glossed over a single but important detail: for this argument to be sound, we need the vertex-set of $\mathcal{L}$ to be countable, which seldom happens. For sure, whenever $p_c^{\text{site}}(\mathcal{S}) < 1$, the set of its $\mathbb{N}$ -indexed self-avoiding paths is uncountable: this fact is very easy but a convoluted proof is given by the proof by contradiction of the present paragraph. In some loose sense, this remark recalls that percolation does not behave well if the ways to reach infinity are decomposed into uncountably many types, and conversely Theorem 4.2 says that it does behave well for countably many types.

4.4. A question

For every $n \geq 0$, let $\mathcal{I}_n$ denote the graph with vertex-set $\{0, 1, \dots, n\}$ and with edges corresponding to $\{i, j\}$ such that $|i - j| = 1$. Likewise, for $n \geq 3$, let $\mathcal{C}_n$ denote the n -cycle, i.e. the graph with vertex-set $\mathbb{Z}/n\mathbb{Z}$ and with edges corresponding to $\{i, j\}$ such that $j = i \pm 1$. Also, given two graphs $\mathcal{G}$ and $\mathcal{H}$, define the product $\mathcal{G} \times \mathcal{H}$ by letting $V_{\mathcal{G} \times \mathcal{H}} := V_{\mathcal{G}} \times V_{\mathcal{H}}$ and declaring (u, x) to be adjacent to (v, y) if either “ $u = v$ and y is a neighbour of x ” or “ $x = y$ and v is a neighbour of u ”. In this section, p_c can be taken to mean either always p_c^{site} or always p_c^{bond} , both leading to a correct reading.

Let $\mathcal{G}$ be a countable locally finite graph. By graph-inclusion, it is clear that

$$m \geq n \implies p_c(\mathcal{G} \times \mathcal{I}_m) \leq p_c(\mathcal{G} \times \mathcal{I}_n).$$

Because of Theorem 1.5, it is also the case that for every $m, n \geq 3$, whenever m is a multiple of n , we have $p_c(\mathcal{G} \times \mathcal{C}_m) \leq p_c(\mathcal{G} \times \mathcal{C}_n)$.

QUESTION 4.5. — *Is it the case that for every countable locally finite graph $\mathcal{G}$ and every $m \geq n \geq 3$, we have $p_c(\mathcal{G} \times \mathcal{C}_m) \leq p_c(\mathcal{G} \times \mathcal{C}_n)$?*

This question asks whether the monotonicity that holds for the “slab” $\mathcal{G} \times \mathcal{I}_n$ also holds for its “periodised” version $\mathcal{G} \times \mathcal{C}_n$.

5. Revisiting the BK inequality

Let C be a finite set. A subset $\mathcal{E} \subset \{0, 1\}^C$ is said to be *increasing* if its indicator function is weakly increasing for the product order. In other words, $\mathcal{E}$ is increasing if and only if for every $\omega \in \mathcal{E}$ and every $\omega' \in \{0, 1\}^C$ satisfying $\forall i, \omega_i \leq \omega'_i$, we have $\omega' \in \mathcal{E}$.

Given two subsets $\mathcal{E}_1$ and $\mathcal{E}_2$ of $\{0, 1\}^C$, the *disjoint occurrence* of $\mathcal{E}_1$ and $\mathcal{E}_2$ is denoted by $\mathcal{E}_1 \circ \mathcal{E}_2$ and defined as follows. It is the set of all $\omega \in \{0, 1\}^C$ such that there are disjoint subsets P_1 and P_2 of C such that the following two conditions hold:

- for every $\omega' \in \{0, 1\}^C$ such that ω' and ω agree on their restriction to P_1 , we have $\omega' \in \mathcal{E}_1$,
- for every $\omega' \in \{0, 1\}^C$ such that ω' and ω agree on their restriction to P_2 , we have $\omega' \in \mathcal{E}_2$.

In other words, we can find two disjoint “witnesses”, one guaranteeing that $\mathcal{E}_1$ holds and the other one doing the same for $\mathcal{E}_2$.

The BK inequality, due to van den Berg and Kesten [BK85], is very useful in percolation theory and goes as follows.

THEOREM 5.1 (BK inequality). — *Let C be a finite set. Let $\mathbb{P}$ be a probability measure on $\{0, 1\}^C$ of the form $\bigotimes_i \text{Bernoulli}(p_i)$. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two increasing subsets of $\{0, 1\}^C$.*

Then, we have $\mathbb{P}(\mathcal{E}_1 \circ \mathcal{E}_2) \leq \mathbb{P}(\mathcal{E}_1)\mathbb{P}(\mathcal{E}_2)$.

A typical application of the BK inequality in percolation theory goes as follows. Let o be a vertex in some graph, and let $n \geq 1$. Perform bond percolation of some fixed parameter p . Denote by $a_1(n, p)$ the probability that there is some open path starting at o that touches the sphere of radius n centred at o . Denote by $a_2(n, p)$ the probability that there are two *edge-wise disjoint* open paths starting at o that touch the sphere of radius n centred at o . As the existence of an open path from the origin to the n -sphere is an increasing event, the BK inequality yields $a_2(n, p) \geq a_1(n, p)^2$. Therefore, any lower bound on $a_1(n, p)$ automatically yields some lower bound for $a_2(n, p)$. Getting sharper inequalities than $a_2(n, p) \geq a_1(n, p)^2$ requires hard work; see [BN11, GM25, RT26].

Proof of the BK inequality given our results. — For now, let us assume that all p_i are equal to a fixed value $p \in [0, 1]$, which is the case we are usually interested in percolation theory. Let $A := C \times \{1, 2\}$, $B := C$ and $\pi : A \rightarrow B$ be the map $(i, j) \mapsto i$. Endow $\{0, 1\}^B$ with the probability measure $\mathbb{P} = \text{Bernoulli}(p)^{\otimes B}$. For every $i \in B$, let us define $X_i : \{0, 1\}^B \rightarrow \{0, 1\}$ by $X_i(\omega) = \omega_i$. At last, let $\mathbf{P} := \text{Bernoulli}(p)^{\otimes A}$.

Let us pick a random section S as follows. Whenever $X \in \mathcal{E}_1 \circ \mathcal{E}_2$, pick disjoint witnesses P_1 and P_2 and make sure that S maps every element of P_1 to 1 and every element of P_2 to 2 — there is no constraint outside $P_1 \cup P_2$. Such a choice is possible as P_1 and P_2 are disjoint. When $X \notin \mathcal{E}_1 \circ \mathcal{E}_2$, the random section S can be arbitrarily chosen. Let $Y_i^{(j)} := X_i \mathbb{1}_{S(i)=(i,j)}$.

Let $\mathcal{E}_{\text{lift}}$ be the set of all $(\omega^{(1)}, \omega^{(2)}) \in \{0, 1\}^A$ such that $\omega^{(1)} \in \mathcal{E}_1$ and $\omega^{(2)} \in \mathcal{E}_2$. This event is increasing, as $\mathcal{E}_1$ and $\mathcal{E}_2$ are themselves increasing. By construction, we have $\mathcal{E}_1 \circ \mathcal{E}_2 \subset \{(Y^{(1)}, Y^{(2)}) \in \mathcal{E}_{\text{lift}}\}$, so that $\mathbb{P}(\mathcal{E}_1 \circ \mathcal{E}_2) \leq \mathbb{P}((Y^{(1)}, Y^{(2)}) \in \mathcal{E}_{\text{lift}})$. Because of Theorem 1.1 and as $\mathcal{E}_{\text{lift}}$ is increasing, we have $\mathbb{P}((Y^{(1)}, Y^{(2)}) \in \mathcal{E}_{\text{lift}}) \leq \mathbf{P}(\mathcal{E}_{\text{lift}})$. By independence, we have $\mathbf{P}(\mathcal{E}_{\text{lift}}) = \mathbb{P}(\mathcal{E}_1)\mathbb{P}(\mathcal{E}_2)$. We conclude by composing the inequalities:

$$\mathbb{P}(\mathcal{E}_1 \circ \mathcal{E}_2) \leq \mathbb{P}((Y^{(1)}, Y^{(2)}) \in \mathcal{E}_{\text{lift}}) \leq \mathbf{P}(\mathcal{E}_{\text{lift}}) = \mathbb{P}(\mathcal{E}_1)\mathbb{P}(\mathcal{E}_2).$$

To get the version with different p_i , it suffices to have a version of Theorem 1.1 with different p_i . Corollary 2.1 is more than sufficient to this end. $\square$

As Corollary 2.1 and Theorem 2.3 are much more general than Theorem 1.1, we can obtain variations of the BK inequality. For instance, with the same proof as before but using Corollary 2.1 instead of Theorem 1.1, we get the following result. It may well not be new but we currently do not know of a place where such a statement is written.

PROPOSITION 5.2. — *Let C be a finite set. Let $\mathbb{P}$ be a probability measure on $\{0, 1\}^C$ of the form $\bigotimes_i \text{Bernoulli}(p_i)$. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two increasing subsets of $\{0, 1\}^C$. Let $\mathbf{P}$ be a probability measure on $\{0, 1\}^{C \times \{1, 2\}}$ of the form $\bigotimes_{i \in C} \rho_i$, where ρ_i is a probability measure on $\{0, 1\}^2$ such that both its marginals are Bernoulli of parameter at least p_i .*

Then, we have $\mathbb{P}(\mathcal{E}_1 \circ \mathcal{E}_2) \leq \mathbf{P}(\mathcal{E}_1 \times \mathcal{E}_2)$.

For instance, applying Proposition 5.2 in the situation of Example 4.4 with $p_i = 1/2$, we recover [BN11, Theorem 3], under the additional assumptions that we consider

increasing events. Beffara and Nolin attribute their Theorem 3 to Reimer [Rei00, Theorem 1.2].

Remark 5.3. — Reimer proved that the BK inequality holds without assuming that the events under consideration are increasing [Rei00]. We cannot use our techniques to prove this stronger inequality. Indeed, it is crucial that $\mathcal{E}_{\text{lift}}$ is increasing in order to obtain the inequality $\mathbb{P}((Y^{(1)}, Y^{(2)}) \in \mathcal{E}_{\text{lift}}) \leq \mathbf{P}(\mathcal{E}_{\text{lift}})$. To get Reimer’s inequality, it would suffice to have a lifting theorem which, in the phrasing of the abstract, also enables Alice to lift zeroes in the same way she lifts ones, without Bob being able to overwrite. However, this “theorem” fails badly, even for $|A| = 2$ and $|B| = 1$.

6. Revisiting strict monotonicity of p_c with respect to quotient

In [MS19], Martineau and Severo proved a result regarding *strict* monotonicity of p_c for bond percolation and site percolation. This answered in particular [BS96, Question 1]. In this section, we provide a new proof of this result for the case of bond percolation.

This proof does not rely on the new coupling results of other sections but it still has a very strong scent of coupling. Contrary to [MS19], our proof does not use the theory of essential enhancements. This section shares with [Van25] the philosophy of studying percolation with couplings instead of differential inequalities. Before stating the main result of this section, namely Theorem 6.1, let us introduce some relevant definitions.

Given a graph $\mathcal{G}$, a *subtree* is a graph $\mathcal{T}$ that is a tree and satisfies $V_{\mathcal{T}} \subset V_{\mathcal{G}}$ and $E_{\mathcal{T}} \subset E_{\mathcal{G}}$. Let $\mathcal{L}$ and $\mathcal{S}$ denote two countable locally finite graphs. A map $\pi : V_{\mathcal{L}} \rightarrow V_{\mathcal{S}}$ satisfies the *disjoint tree-lifting property* if for any two distinct vertices x_1 and x_2 in $\mathcal{L}$ satisfying $\pi(x_1) = \pi(x_2)$, for every subtree $\mathcal{T}$ of $\mathcal{S}$ containing the vertex $\pi(x_1)$, there are subtrees $\mathcal{T}_1$ and $\mathcal{T}_2$ of $\mathcal{L}$ such that:

- the vertex-sets of $\mathcal{T}_1$ and $\mathcal{T}_2$ are disjoint,
- for every $i \in \{1, 2\}$, the vertex x_i belongs to $\mathcal{T}_i$,
- for every $i \in \{1, 2\}$, the map π induces a well-defined graph isomorphism from $\mathcal{T}_i$ to $\mathcal{T}$.

In this paper, a *cycle* in a graph $\mathcal{G}$ is a finite sequence of vertices $(v_0, \dots, v_\ell)$ such that each vertex is adjacent *or equal* to the previous one, and such that v_ℓ is equal to v_0 . The number ℓ is called the *length* of the cycle. Given $\pi : V_{\mathcal{L}} \rightarrow V_{\mathcal{S}}$, say that *we can switch floors by lifting short cycles* if there is a constant c such that the following holds: for every vertex x in $\mathcal{L}$, there is a distinct vertex y in $\mathcal{L}$ satisfying $\pi(x) = \pi(y)$ and such that there is a cycle of length at most c in $\mathcal{S}$ that admits a lift starting at x and ending at y . In the previous sentence, a lift of a cycle $(v_0, \dots, v_\ell)$ is a path $(x_0, \dots, x_\ell)$ in $\mathcal{L}$ such that, for every $i \leq \ell$, we have $\pi(x_i) = v_i$. We allow consecutive vertices of a cycle to be equal because we want Theorem 6.1 to cover situations such as $\mathcal{L}$ being the cubic lattice, $\mathcal{S}$ the square lattice and $\pi : (x, y, z) \mapsto (x, y)$.

At last, recall that a graph has *bounded degree* if there is some constant D such that each vertex has at most D neighbours.

THEOREM 6.1 ([MS19]). — *Let $\mathcal{L}$ and $\mathcal{S}$ be two countable locally finite graphs. Assume that $\mathcal{S}$ has bounded degree. Further, assume that there is a surjective map $\pi : V_{\mathcal{L}} \rightarrow V_{\mathcal{S}}$ satisfying the disjoint tree-lifting property and such that we can switch floors by lifting short cycles.*

If $p_c^{\text{bond}}(\mathcal{L}) < 1$, then we have $p_c^{\text{bond}}(\mathcal{L}) < p_c^{\text{bond}}(\mathcal{S})$.

Remark 6.2. — It would be natural to further assume that π is a fibration. One can put this condition in the theorem or not without changing the scope of the result, as any π satisfying the conditions of Theorem 6.1 is necessarily a fibration: any time we want to lift an edge to some $x_1 \in V_{\mathcal{L}}$, we can use floor-switching to find a distinct $x_2 \in V_{\mathcal{L}}$ such that $\pi(x_1) = \pi(x_2)$, and then apply the weak disjoint-tree lifting property to (x_1, x_2) and a tree consisting of a single edge. The fact that π is a fibration will indeed be used in Section 6.

Remark 6.3. — The statement we give here does not correspond exactly to [MS19, Theorem 1.1]. Our Theorem 6.1 directly implies theirs (for bond percolation only). The techniques of [MS19] actually prove the above statement for both site and bond percolation, even though they did not state it in this way. Theorem 1.1 of [MS19] corresponds to further assuming that $\mathcal{L}$ has bounded degree, that π is 1-Lipschitz, and that $\mathcal{L}$ and $\mathcal{S}$ are nonempty and connected. Apart from the connectedness of $\mathcal{S}$, these additional assumptions are actually not used in the proof of [MS19]. Connectedness of $\mathcal{S}$ is really used. However, it could be removed, as in our proof, by noticing that the bounds obtained in [MS19] are not mere strict inequalities but are in some sense “uniform”.

Remark 6.4. — We decided not to assume connectedness of the graphs for two reasons. First, obtaining such a result is strictly more informative and captures some “uniformity” of the strict inequality: if we have sequences of connected graphs $(\mathcal{L}_n)$ and $(\mathcal{S}_n)$ such that $(\mathcal{L}_n, \mathcal{S}_n)$ satisfies the assumptions of Theorem 6.1 with bounds on the degree and on the c of “short cycles” that are uniform in n , then our result applied to $\bigsqcup_n \mathcal{L}_n$ and $\bigsqcup_n \mathcal{S}_n$ guarantees that $\inf_n p_c^{\text{bond}}(\mathcal{L}_n) < \inf_n p_c^{\text{bond}}(\mathcal{S}_n)$. If we assumed connectedness in Theorem 6.1, we would know that, for every n , the inequality $p_c^{\text{bond}}(\mathcal{L}_n) < p_c^{\text{bond}}(\mathcal{S}_n)$ holds but we would not be able to derive strict inequality for the infima.

A second reason not to assume connectedness is that the setup we consider is well behaved if we want to transfer knowledge from site-percolation to bond-percolation: see Section 4.1. Even when $\mathcal{S}$ and $\mathcal{L}$ are connected, the graph $\mathcal{L}_{\pi}^*$ may not be connected, and our point is that it is not necessary to find a connected substitute for $\mathcal{L}_{\pi}^*$. In the end, this second reason is not concretely used here, as we do not prove the site-version but only the bond-version of [MS19, Theorem 1.1]. In other words, we have decided to work in a good framework, even though we do not use its goodness.

Informal overview of the proof strategy

As in [MS19], we want to introduce some sort of new percolation model on $\mathcal{S}$ (or a closely related graph) such that we have $p_c(\mathcal{L}) \leq p_c^{\text{new}}(\mathcal{S}) < p_c(\mathcal{S})$. To do so, we will first deterministically partition $\mathcal{S}$ (or something closely related to it) into blocks, which will be called “cells”. Then, we will take the usual percolation and allow some “bonus” inside each cell, and the whole game is to do so in such a way that at the same time:

- (1) this bonus can be “implemented, via π , in the usual percolation on $\mathcal{L}$ ”,
- (2) this bonus is sufficiently strong for the inequality $p_c^{\text{new}}(\mathcal{S}) < p_c(\mathcal{S})$ to be true (and provable).

A noteworthy feature of this proof is that each of the two inequalities $p_c(\mathcal{L}) \leq p_c^{\text{new}}(\mathcal{S})$ and $p_c^{\text{new}}(\mathcal{S}) < p_c(\mathcal{S})$ is proved by using some exploration algorithm but that the algorithms are different (one specific for each inequality). There is no need for any form of “compatibility” between these algorithms, apart from correctly describing our percolation process.

Step 1: reduction of the problem

To begin with, let us explain why it suffices to prove $p_c(\mathcal{L}) < p_c(\mathcal{S})$ in the following setup.

Notation 6.5. — Given a graph $\mathcal{G}$, for $A \subset V_{\mathcal{G}}$, we denote by E_A the set of all edges of $\mathcal{G}$ the two endpoints of which belong to A . We define the vertex-boundary $\partial_V A$ to be the set of all vertices in A having a neighbour outside A . The interior of A is defined as $\mathring{A} := A \setminus \partial_V A$. The edge-boundary of A is $\partial_E A = \{e \in E_{\mathcal{G}} : |e \cap A| = 1\}$. The notation $B_u(\alpha)$ stands for the closed ball of radius α centred at u , i.e. $B_u(\alpha) = \{v \in V_{\mathcal{G}} : d(u, v) \leq \alpha\}$.

The setup. — The graphs $\mathcal{L}$ and $\mathcal{S}$ have nonempty countable vertex-sets. The graph $\mathcal{L}$ is locally finite. The graph $\mathcal{S}$ has bounded degree and satisfies $0 < p_c^{\text{bond}}(\mathcal{S}) < 1$. We have a surjective map $\pi : V_{\mathcal{L}} \rightarrow V_{\mathcal{S}}$ with the disjoint tree-lifting property. We also have positive integers r and R such that the following holds. For every vertex x in $\mathcal{L}$, there is a distinct vertex y in $\mathcal{L}$ satisfying $\pi(x) = \pi(y)$ and such that there is a cycle of length at most $2r$ in $\mathcal{S}$ that admits a lift starting at x and ending at y . Furthermore, we have a set of vertices $\mathfrak{C} \subset V_{\mathcal{S}}$ such that, for every $u \in \mathfrak{C}$, there exists a set $\mathcal{C}_u \subset V_{\mathcal{S}}$ such that the following conditions hold:

- (1) when u ranges over $\mathfrak{C}$, the $E_{\mathcal{C}_u}$ ’s form a partition of $E_{\mathcal{S}}$,
- (2) for every $u \in \mathfrak{C}$, the interior of $\mathcal{C}_u$ is connected,
- (3) for every $u \in \mathfrak{C}$, we have the inclusions $B_u(r) \subset \mathring{\mathcal{C}}_u$ and $\mathcal{C}_u \subset B_u(R)$.

We start with $(\mathcal{L}_0, \mathcal{S}_0, \pi_0)$ satisfying the assumptions of Theorem 6.1. As $\mathcal{S}_0$ has bounded degree, $p_c^{\text{bond}}(\mathcal{S}_0) > 0$. Without loss of generality, assume that $p_c^{\text{bond}}(\mathcal{S}_0) < 1$, as the statement is trivial when $p_c^{\text{bond}}(\mathcal{S}_0) = 1$. In particular, $V_{\mathcal{S}_0}$ is nonempty and so is $V_{\mathcal{L}_0}$, by surjectivity of π . We can pick $r_0 \geq 1$ such that for every vertex x in $\mathcal{L}_0$, there is a distinct vertex y in $\mathcal{L}$ satisfying $\pi(x) = \pi(y)$ and such that there

is a cycle of length at most $2r_0 + 1$ in $\mathcal{S}$ that admits a lift starting at x and ending at y . Removing edges from $\mathcal{L}_0$ cannot decrease its p_c^{bond} , so we may remove from it all edges that are not mapped, via π_0 , to edges of $\mathcal{S}_0$. Notice that doing so preserves the assumptions made on π_0 , which are only concerned by how π_0 lifts edges of $\mathcal{S}_0$. In other words, without loss of generality, we may and will assume that π_0 induces a well defined map from $E_{\mathcal{L}_0}$ to $E_{\mathcal{S}_0}$. We want to prove that $p_c^{\text{bond}}(\mathcal{L}_0) < p_c^{\text{bond}}(\mathcal{S}_0)$.

Let $\mathfrak{C}$ denote a maximal $(2r_0 + 1)$ -separated subset of $V_{\mathcal{S}_0}$, i.e. a set any two distinct elements of which are at distance at least $2r_0 + 1$ and that is maximal with respect to inclusion among all such sets. Fix, for every $u \in \mathfrak{C}$, a set $\tilde{C}_u \subset V_{\mathcal{S}_0}$ in such a way that the following conditions hold:

- (i) when u ranges over $\mathfrak{C}$, the $\tilde{C}_u$'s form a partition of $V_{\mathcal{S}_0}$,
- (ii) every $\tilde{C}_u$ is connected,
- (iii) for every $u \in \mathfrak{C}$, we have the inclusions $B_u(r_0) \subset \tilde{C}_u \subset B_u(2r_0)$.

Such a choice is indeed possible: for instance, one can take the Voronoi cells associated with $\mathfrak{C}$ and tie break arbitrarily all cases of equality (see Figure 6.1).

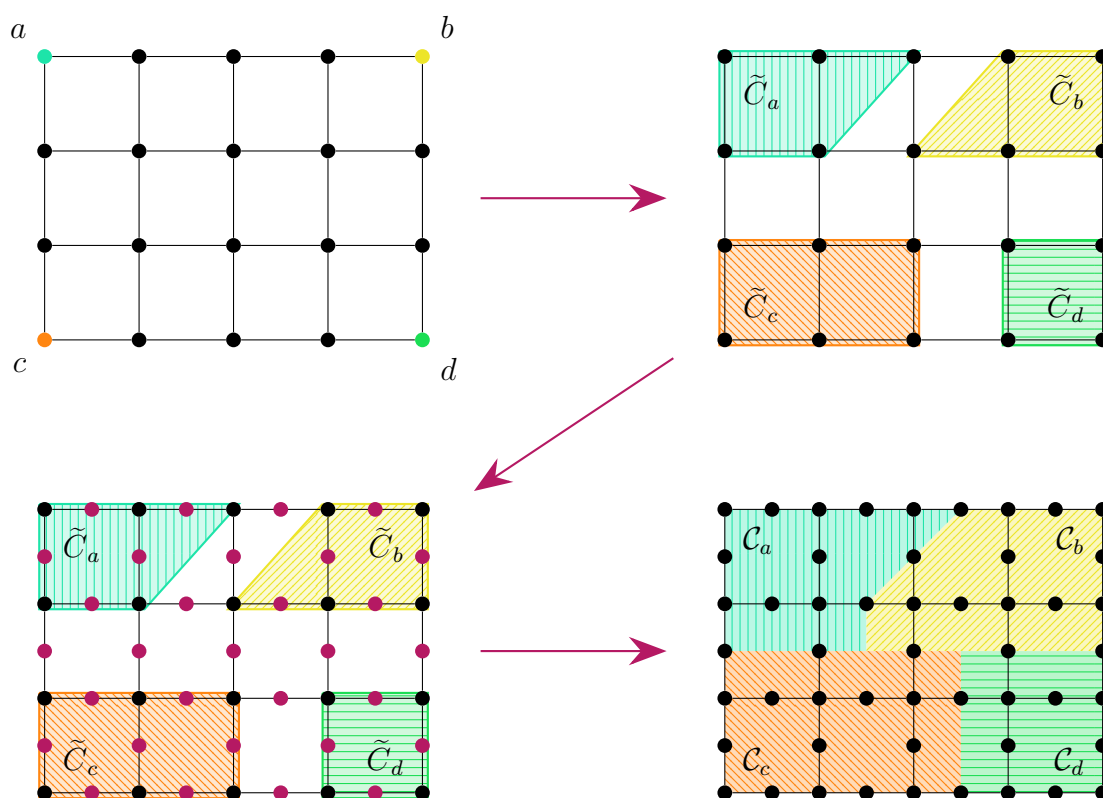


Figure 6.1. Going from $\mathcal{S}_0$ to $\mathcal{S}$ and defining the cells.

Given a graph $\mathcal{G}$, define a new graph $\mathcal{G}'$ by performing a so-called *edge subdivision* at each edge, i.e. by adding a new vertex at the middle of each edge. Formally, one has $V_{\mathcal{G}'} = V_{\mathcal{G}} \sqcup E_{\mathcal{G}}$. As for the edges, for every $u \in V_{\mathcal{G}}$ and every edge $e \in E_{\mathcal{G}}$ containing u , we declare u and e to be adjacent in $\mathcal{G}'$, and no other edge is introduced. Set $\mathcal{L} = \mathcal{L}'_0$ and $\mathcal{S} = \mathcal{S}'_0$, and π the map $V_{\mathcal{L}_0} \sqcup E_{\mathcal{L}_0} \rightarrow V_{\mathcal{S}_0} \sqcup E_{\mathcal{S}_0}$ induced by π_0 .

Notice that, as $\mathfrak{C}$ is a subset of $V_{\mathcal{S}_0}$, it is also a subset of $V_{\mathcal{S}}$. For $u \in \mathfrak{C}$, we define $\mathcal{C}_u = \tilde{\mathcal{C}}_u \sqcup \{e \in E_{\mathcal{S}_0} : e \cap \tilde{\mathcal{C}}_u \neq \emptyset\}$. Observe the edge-boundary of $\tilde{\mathcal{C}}_u$ in $\mathcal{S}_0$ is equal to the vertex-boundary of $\mathcal{C}_u$ in $\mathcal{S}$. At last, we set $r = 2r_0$ (to take into account that distances are twice larger in $\mathcal{S}$ than in $\mathcal{S}_0$) and $R = 4r_0 + 1$. Making all these choices indeed produces data fitting the setup we were targeting.

To conclude the reduction, it remains to explain why the inequality $p_c(\mathcal{L}) < p_c(\mathcal{S})$ suffices to yield $p_c(\mathcal{L}_0) < p_c(\mathcal{S}_0)$, which is easy as, for any graph, we have $p_c(\mathcal{G}) = p_c(\mathcal{G}')^2$.

Remark 6.6. — The only reason why we restrict ourselves to bond percolation goes as follows. In the remaining of the proof, we work in the aforementioned setup to prove the desired conclusion. The arguments will adapt readily to site percolation, provided we replace (1) of the setup by the stronger condition “the $\mathcal{C}_u$ ’s form a partition of $V_{\mathcal{S}}$ ” — and they would still work if we further weaken (ii) to “for every $u \in \mathfrak{C}$ and every $v \in \partial_V \mathcal{C}_u$, there is a path from u to v in $\mathcal{C}_u$ that, apart from its final step, avoids $\partial_V \mathcal{C}_u$ ”. It is however not clear how to reduce the general situation to this modified setup.

Step 2: Introducing the augmented percolation model on $\mathcal{S}$

Let $p \in [0, 1]$ and let $(X_e)_{e \in E_{\mathcal{S}}}$ be a collection of independent Bernoulli random variables of parameter p . An edge e satisfying $X_e = 1$ is said to be *p-open*. Let $s \in [0, 1]$ and let $(Y_{\mathcal{C}})$ be a collection of independent Bernoulli random variables of parameter s , indexed by the cells. The families (X_e) and $(Y_{\mathcal{C}})$ are also taken to be independent of one another.

For $A \subset V_{\mathcal{S}}$, we define $K_{p,s}^+(A)$, the augmented cluster of A , to be the smallest set $K \subset V_{\mathcal{S}}$ containing A and satisfying the following “closure” properties:

- (1) for every $v \in K$ and every neighbour w of v , if $X_{\{v,w\}} = 1$, then w belongs to K ,
- (2) for every cell $\mathcal{C}$, every $v \in \partial_V \mathcal{C}$ and every $w \in \overset{\circ}{\mathcal{C}}$ adjacent to v , if $v \in K$ and $X_{\{v,w\}} = 1$ and all edges $e \subset \overset{\circ}{\mathcal{C}}$ satisfy $X_e = 1$ and $Y_{\mathcal{C}} = 1$, then K contains the whole cell $\mathcal{C}$.

The augmented cluster of a vertex v is defined to be $K_{p,s}^+(v) := K_{p,s}^+(\{v\})$. Notice that if u belongs to $K_{p,s}^+(v)$, it does not imply that v belongs to $K_{p,s}^+(u)$; see Figure 6.2.

Given $s \in [0, 1]$, there is a unique $p_c^{\text{aug}}(\mathcal{S}, s, v)$ such that:

- for every $p < p_c^{\text{aug}}(\mathcal{S}, s, v)$, we have $\mathbb{P}(K_{p,s}^+(v) \text{ is infinite}) = 0$,
- for every $p > p_c^{\text{aug}}(\mathcal{S}, s, v)$, we have $\mathbb{P}(K_{p,s}^+(v) \text{ is infinite}) > 0$.

If u and v are in the same connected component of $\mathcal{S}$, then we have $p_c^{\text{aug}}(\mathcal{S}, s, u) = p_c^{\text{aug}}(\mathcal{S}, s, v)$. Therefore, if $\mathcal{C}$ denotes a connected component of $\mathcal{S}$, the quantity $p_c^{\text{aug}}(\mathcal{C}, s)$ is well defined. Besides, by Kolmogorov’s zero-one law, the following conditions hold:

- for every $p < p_c^{\text{aug}}(\mathcal{C}, s)$, almost surely, for every $v \in \mathcal{C}$, the augmented cluster of v is finite,

- for every $p > p_c^{\text{aug}}(\mathcal{C}, s)$, there is almost surely a vertex in $\mathcal{C}$ such that its augmented cluster is infinite.

It makes sense to set $p_c^{\text{aug}}(\mathcal{S}, s) := \inf_{\mathcal{C}} p_c^{\text{aug}}(\mathcal{C}, s)$, as we then have this dichotomy:

- for every $p < p_c^{\text{aug}}(\mathcal{S}, s)$, almost surely, for every $v \in V_{\mathcal{S}}$, the augmented cluster of v is finite,
- for every $p > p_c^{\text{aug}}(\mathcal{S}, s)$, there is almost surely a vertex in $\mathcal{S}$ such that its augmented cluster is infinite.

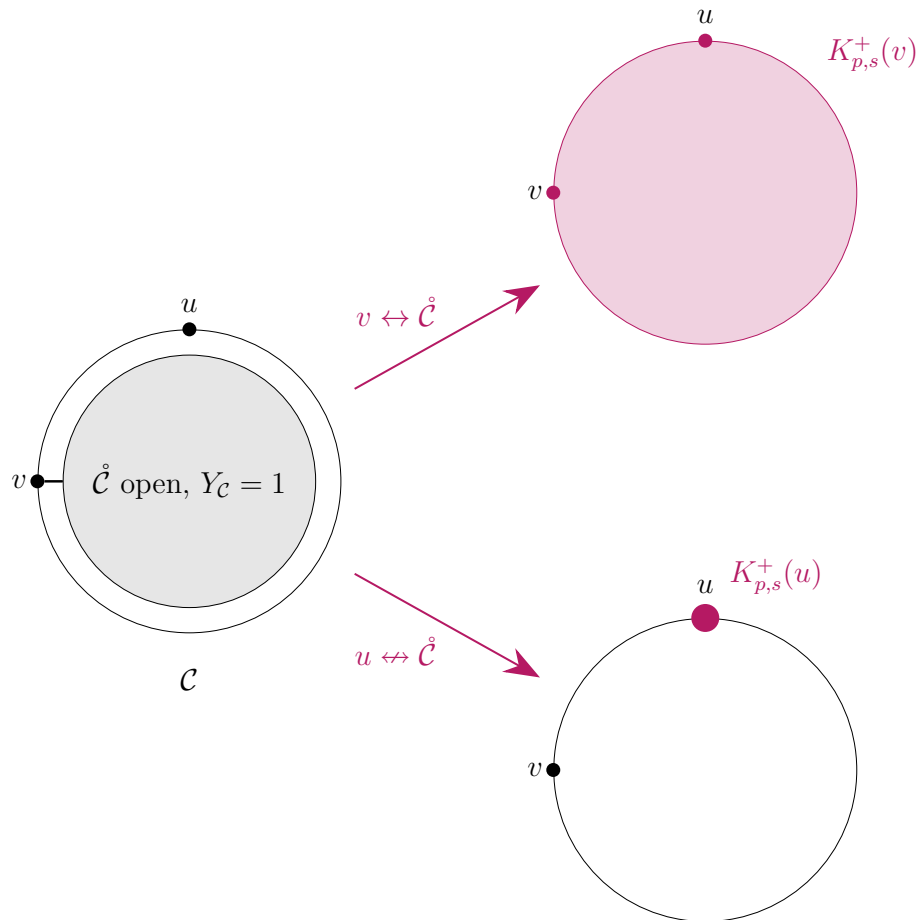


Figure 6.2. Assuming that the X 's and Y 's are 0 in the neighbouring cells, this is a situation where u belongs to $K_{p,s}^+(v)$ but v does not belong to $K_{p,s}^+(u)$.

What will play the role of $p_c^{\text{new}}(\mathcal{S})$ from the heuristics of page 1256 will be $p_c^{\text{aug}}(\mathcal{S}, s)$, for small enough values of s . More precisely, in Step 3, we shall prove that for every $s \in (0, 1]$, we have $p_c^{\text{aug}}(\mathcal{S}, s) < p_c^{\text{bond}}(\mathcal{S})$. Then, we shall see in Step 4 that for s small enough, we have $p_c^{\text{bond}}(\mathcal{L}) \leq p_c^{\text{aug}}(\mathcal{S}, s)$. Step 4 is established by using a standard exploration algorithm, revealing edges at the boundary of the cluster one at a time. The novelty lies in Step 3, which is proved by using *another* algorithm: there, we proceed cell by cell rather than edge by edge.

Step 3: For every $s \in (0, 1]$, we have $p_c^{\text{aug}}(\mathcal{S}, s) < p_c^{\text{bond}}(\mathcal{S})$

Let $s \in (0, 1]$. Our purpose is to find some $v_0 \in V_{\mathcal{S}}$ such that $p_c^{\text{aug}}(\mathcal{S}, s, v_0) < p_c^{\text{bond}}(\mathcal{S})$.

Let v_0 be a vertex lying at the boundary of its cell. Then, one way to explore its p -cluster⁽⁸⁾ is by using the following algorithm. It does not really explore the full cluster, only what happens at the boundaries of the cells (which is enough, as the cluster is infinite if and only if its intersection with $\bigcup_{\mathcal{C}} \partial_V \mathcal{C}$ is infinite). We will call an element of $\bigcup_{\mathcal{C}} \partial_V \mathcal{C}$ a *boundary-vertex*.

INSTRUCTIONS. —

- (1) *Before applying the algorithm, enumerate the cells: when we will say “pick a cell such that something”, it will mean that we pick the cell with smallest label among all cells satisfying something. This will guarantee univoque definition and measurability of the process, conditional independence (we do not pick a cell depending on what is inside), and that any cell that could be picked will be picked at some point even if the algorithm goes on forever.*⁽⁹⁾
- (2) *Initialise the process by declaring all cells to be unexplored, and set the current cluster to be $\{v_0\}$.*
- (3) *Pick some unexplored cell $\mathcal{C}$ such that its boundary contains a vertex of the current cluster. Reveal all X_e ’s for e in the cell $\mathcal{C}$. We do not need to remember all these values, only the answer to each of the following questions: for any v and w in $\partial_V \mathcal{C}$, the question is “is there a p -open path connecting v to w inside $\mathcal{C}$?”. When the answer is yes, we say that v is p -connected to w in $\mathcal{C}$. Then, the cell $\mathcal{C}$ becomes explored.*
- (4) *Add to the current cluster all boundary-vertices v such that there are a boundary-vertex w in the current cluster and an explored cell $\mathcal{C}$ containing both v and w such that v is p -connected to w in $\mathcal{C}$.*
- (5) *As long as it adds new vertices, iterate Instruction (4).*
- (6) *As long as there are unexplored cells, iterate Instructions (3)–(5).*

The algorithm stops after finitely many steps if and only if the p -cluster of v_0 is finite. If it stops after finitely many steps, define the final cluster to be the current cluster at the end of the algorithm. If the algorithm goes on forever, the current cluster is an increasing sequence of sets: in this case, we can define the final cluster to be the union of these sets. Notice that the final cluster produced by the algorithm consists of the intersection of the true percolation cluster with the set of boundary-vertices.

How can this be helpful to us? To explain how, let us introduce, for each cell $\mathcal{C}$, the space $\mathfrak{B}_{\mathcal{C}}$ of all equivalence relations on $\partial_V \mathcal{C}$. Elements of $\mathfrak{B}_{\mathcal{C}}$ will typically serve to capture equivalence relations such as “being p -connected in $\mathcal{C}$ ”. For each $\mathcal{C}$, the finite set $\mathfrak{B}_{\mathcal{C}}$ is endowed with a partial order as follows. We say that η is smaller

⁽⁸⁾ Namely its connected component for the graph defined by the p -open edges: this p -cluster is simply a cluster of bond percolation with parameter p .

⁽⁹⁾ We could make our argument to work without taking care of this last point.

than η' if, for any v and w in $\partial_V \mathcal{C}$, whenever v and w are in the same η -class, then they are also in the same η' -class.

The strategy is to define an algorithm similar to the previous one exploring the boundary-vertices of the (p, s) -augmented percolation cluster of v_0 . Due to the augmentation, we will be able to find some $q > p$ such that for every $\mathcal{C}$, the random equivalence relation “being q -connected in $\mathcal{C}$ ” is stochastically dominated by “being (p, s) -connected in $\mathcal{C}$ ”. Therefore, the algorithm will provide a coupling such that if the q -cluster of v_0 is infinite, then so is the (p, s) -augmented cluster. With some extra care regarding how to choose p and q , we will be able to conclude Step 3 from there.

To define the exploration algorithm for (p, s) -augmented percolation, we only modify Instructions (3) and (4). Besides, we consider that we picked the same enumeration in Instruction 1 for both algorithms. Instructions (3) and (4) are as follows.

- (3) Pick some unexplored cell $\mathcal{C}$ such that its boundary contains a vertex of the current cluster. Reveal all X_e 's for e in the cell $\mathcal{C}$, as well as $Y_{\mathcal{C}}$. We do not need to remember all these values, only the answer to each of the following questions: for any v and w in $\partial_V \mathcal{C}$, the question is “is it the case that either there is a p -open path from v to w in $\mathcal{C}$ or something else happens?”. In the previous sentence, “something else” refers to all the following conditions holding simultaneously: there is a p -open edge in $\mathcal{C}$ such that one of its endpoints belongs to $\mathring{\mathcal{C}}$ and the other one belongs to the current cluster; all edges lying in $\mathring{\mathcal{C}}$ are p -open; $Y_{\mathcal{C}}$ is equal to 1. When the answer to the question is yes, we say that v is (p, s) -connected to w in $\mathcal{C}$. Then, the cell $\mathcal{C}$ becomes explored. *We insist that v being (p, s) -connected to w in $\mathcal{C}$ is a Boolean that is defined now and will not be updated later. In other words, at any subsequent step, when we ask if vertices are (p, s) -connected in $\mathcal{C}$, the “current cluster” appearing in the definition of “ (p, s) -connected” is the cluster at the beginning of the instruction revealing $\mathcal{C}$.*
- (4) Add to the current cluster all boundary-vertices v such that there are a boundary-vertex w in the current cluster and an explored cell $\mathcal{C}$ containing both v and w such that v is (p, s) -connected to w in $\mathcal{C}$.

The final cluster produced by the algorithm is a subset⁽¹⁰⁾ of the intersection of the true augmented cluster with the set of boundary-vertices. In particular, if the cluster produced by this algorithm is infinite, then $K_{p,s}^+(v_0)$ is infinite.

Denote by $Z_p(\mathcal{C})$ the random equivalence relation defined by “being p -connected in $\mathcal{C}$ ”. For A a nonempty subset of $\partial_V \mathcal{C}$, denote by $Z_{p,s}^A(\mathcal{C})$ the random equivalence relation defined as in Instruction (3), but with A instead of “the current cluster”. Because of the above algorithms, if the first of the following conditions holds, then so does so second:

- for every $\mathcal{C}$ and A , the distribution of $Z_q(\mathcal{C})$ is stochastically dominated by that of $Z_{p,s}^A(\mathcal{C})$,

⁽¹⁰⁾ This is generally not an equality because, in the closure property (2) of page 1258, it may well be that suitable v 's exist but that none of them is in the *current* version of the cluster.

- the intersection of the q -cluster of v_0 with the boundary-vertices has a distribution stochastically dominated by the same for the (p, s) -augmented cluster of v_0 .

Combining this observation with the following lemma, we will be able to conclude.

LEMMA 6.7. — *Let $\varepsilon > 0$. There is some $\delta \in (0, \varepsilon]$ depending only on R, s, ε and our upper bound on the maximal degree of $\mathcal{S}$ such that for every $p \in [\varepsilon, 1 - \varepsilon]$, for every $\mathcal{C}$ and A , the distribution of $Z_{p+\delta}(\mathcal{C})$ is stochastically dominated by that of $Z_{p,s}^A(\mathcal{C})$.*

Remark 6.8. — As for Lemma 2.11, this lemma follows rather easily from Strassen's Theorem for monotone couplings. Here again, we opted for the Strassen-free proof because, despite being longer, it seemed to us it was more instructive.

Proof. — Once we know R and an upper bound on the maximal degree of $\mathcal{S}$, there are, up to isomorphism, only finitely many possibilities for $(\mathcal{C}, A)$. Therefore, it suffices, given ε , to find some δ depending on $(s, \varepsilon, \mathcal{C}, A)$ such that for every $p \in [\varepsilon, 1 - \varepsilon]$, the distribution of $Z_{p+\delta}(\mathcal{C})$ is stochastically dominated by that of $Z_{p,s}^A(\mathcal{C})$. Let us prove this.

Let $\varepsilon > 0$. Let $\mathcal{C}$ be a cell and A a nonempty finite subset of $\partial_V \mathcal{C}$. Let $p \in [\varepsilon, 1 - \varepsilon]$. Let μ_p be the distribution of $(Z_p(\mathcal{C}), Z_{p,s}^A(\mathcal{C}))$. Notice that μ_p is a monotone coupling between its first and its second marginal. We use the notation **empty** for the minimum of $\mathfrak{B}_{\mathcal{C}}$, namely the equivalence relation with singleton classes. We use the notation **full** for the maximum of $\mathfrak{B}_{\mathcal{C}}$, namely the equivalence relation with a single class. Because $\mathring{\mathcal{C}}$ is connected, if we have the following conditions:

- edges lying in $\mathring{\mathcal{C}}$ are all open,
- other edges in $\mathcal{C}$ are closed except for a single one, which besides has an endpoint in A ,
- $Y_{\mathcal{C}} = 1$,

then $Z_p(\mathcal{C})$ is **empty** and $Z_{p,s}^+(\mathcal{C})$ is **full**. The event “ $Z_p(\mathcal{C}) = \text{empty}$ and $Z_{p,s}^+(\mathcal{C}) = \text{full}$ ” has therefore probability at least $\alpha := s \cdot \varepsilon^{|E_{\mathcal{C}}|} > 0$.

Recall that μ_p is a probability measure on $(\mathfrak{B}_{\mathcal{C}})^2$. Notice that, for every $\eta \in \mathfrak{B}_{\mathcal{C}}$, the quantity $\nu_p(\eta) := \sum_{\eta'} \mu_p(\eta, \eta')$ depends polynomially, hence uniformly continuously, on $p \in [0, 1]$. Therefore, we can pick $\delta \in (0, \varepsilon]$ such that for every $(p, p') \in [0, 1]^2$, if $|p - p'| \leq \delta$, then we have $\beta := \sum_{\eta} |\nu_{p'}(\eta) - \nu_p(\eta)| \leq \alpha$. By taking $p \in [\varepsilon, 1 - \varepsilon]$ and $p' = p + \delta$, this means that, considering probability measures as vectors, the L^1 -distance between the first marginal of μ_p and $\nu_{p+\delta}$ is $\beta \leq \alpha$. In order to get a monotone coupling between $\nu_{p+\delta}$ and the distribution of $Z_{p,s}^A(\mathcal{C})$, we start from μ_p (which has the desired second marginal) and “move mass around” to adjust the first marginal.

If $\nu_{p+\delta}(\eta) > \nu_p(\eta)$, we want to take mass $\nu_{p+\delta}(\eta) - \nu_p(\eta)$ from (**empty**, **full**) and reallocate it to (η, full) : this transforms the weight of η into the desired one without changing the second marginal and without breaking monotonicity of the coupling, since $\eta \leq \text{full}$. If $\nu_{p+\delta}(\eta) \leq \nu_p(\eta)$, we want to take mass $\nu_p(\eta) - \nu_{p+\delta}(\eta)$ from one or several (η, η') and reallocate it to (**empty**, η'): this transforms the weight of η into the desired one without changing the second marginal and without breaking

monotonicity of the coupling, since $\eta' \geq \text{empty}$. Doing so for all $\eta \neq \text{empty}$, we get a monotone coupling with correct marginals except possibly for empty (we do not run into negative weights precisely because⁽¹¹⁾ $\beta \leq \alpha$). Then, because probability measures have total mass 1, the marginals must also be correct for empty . $\square$

We are now ready to conclude Step 3. Pick ε such that $p_c(\mathcal{S}) \in (\varepsilon, 1 - \varepsilon)$. Pick some δ satisfying the conclusion of Lemma 6.7. We can pick v_0 such that $p_c^{v_0}(\mathcal{S}) < p_c(\mathcal{S}) + \delta$, where $p_c^{v_0}(\mathcal{S})$ denotes p_c^{bond} of the connected component of v_0 in $\mathcal{S}$. We can thus pick q such that $\varepsilon < p_c(\mathcal{S}) \leq p_c^{v_0}(\mathcal{S}) < q < p_c(\mathcal{S}) + \delta$. Therefore, we can pick p such that $p + \delta \geq q$ and $\varepsilon < p < p_c(\mathcal{S})$ — either $p = q - \delta$ or $p = \frac{p_c(\mathcal{S}) + \varepsilon}{2}$ does the job. As $p + \delta \geq q > p_c^{v_0}(\mathcal{S})$, the $(p + \delta)$ -cluster of v_0 is infinite with positive probability. Because of Lemma 6.7 and the observation made just before it, this means that the (p, s) -augmented cluster of v_0 is infinite with positive probability. This gives $p_c^{\text{aug}}(\mathcal{S}, s) \leq p < p_c(\mathcal{S})$, thus ending Step 3.

Step 4: For some $s \in (0, 1]$, we have $p_c^{\text{bond}}(\mathcal{L}) \leq p_c^{\text{aug}}(\mathcal{S}, s)$.

This part is performed exactly as in the proof of [MS19, Proposition 2.1], to which the reader is referred for details. It is even simpler in the present case as the sets E_C are disjoint, so that there is no need for the “multiple edges” trick. Let us still provide a sketch of the argument.

Let $p \in (0, 1]$. We explore a spanning tree of the p -cluster of v_0 and lift it to $\mathcal{L}$ in the usual way, revealing one by one the edges adjacent to the explored part of the cluster. But if our p -open explored edges fill $B_u(r)$ in $\mathcal{S}$, then, since we can switch floors by lifting cycles of length at most some constant c , we will be able to connect to some unexplored copy⁽¹²⁾ of $\mathcal{C}_u$ with probability at least $p^c > 0$. Performing p -percolation on this copy, there is some probability at least $p^M > 0$ that this copy turns out to be fully open (where $M := \max \deg(\mathcal{S})^R$ is an upper bound on the cardinality of the edge-set of a spanning tree of $\mathcal{C}_u$). Setting $s_p := p^{M+c} > 0$, these observations yield a coupling such that, almost surely, the image by π of the p -cluster of v'_0 contains the (p, s_p) -augmented cluster of v_0 .

How to conclude from there? If $p_c^{\text{bond}}(\mathcal{L}) = 0$, then there is nothing to prove. Otherwise, setting $s = s_{p_c^{\text{bond}}(\mathcal{L})}$ yields the statement of Step 4, thus concluding the proof of Theorem 6.1. $\square$

In light of Remarks 6.3 and 6.6, we conclude by asking the following question.

QUESTION 6.9. — *Is it possible to adapt our proof of Theorem 6.1 to cover the case of site percolation?*

Addendum

Posterior to the paper getting finalised and accepted for publication, Ivailo Hartarsky brought to our attention the reference [CTSS24]. Their Proposition 2.5 and

⁽¹¹⁾ Actually, as the total increase of the weights matches its total decrease, the inequality $\beta \leq 2\alpha$ would suffice.

⁽¹²⁾ Formally an unexplored lift of a spanning tree of $\mathcal{C}_u$.

our Theorem 1.1 can be deduced from each other in a rather straightforward way, by using Strassen's theorem. Their approach and ours differ, and theirs does not yield Corollary 2.1 (even for $\{0, 1\}$ -valued random variables) or Theorem 2.3.

Appendix A. Counterexamples for C_μ^ρ and $\tilde{C}_\mu^\rho$

The appendix was written by Vibhu Dalal.

In this appendix, we present two counterexamples obtained through an algorithmic search. The first one is a smaller version of the counterexample constructed in Section 3.2: it shows that the condition C_μ^ρ does not imply stochastic domination. The second one answers negatively the analogous question for the stronger condition $\tilde{C}_\mu^\rho$ defined at the end of Section 3.2.

Throughout the appendix, when we refer to the size of a measure, we mean the cardinality of its support.

A.1. A smaller counterexample for the condition C_μ^ρ

One can systematically search for counterexamples of this type by exploring finite configurations which satisfy the condition C_μ^ρ and then testing whether stochastic domination fails. Using this approach, we found the following smaller counterexample, involving only three matrices of size 2×3 . Let μ be the uniform measure on the three matrices

$$M_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let ρ be the uniform measure on the three matrices

$$N_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad N_2 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad N_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

A direct verification shows that μ and ρ satisfy the condition C_μ^ρ .

However, the measure ρ does not stochastically dominate μ . Indeed, the matrices M_1 and M_2 have total μ -mass $2/3$. But among the matrices in the support of ρ , only the all-ones matrix N_3 dominates at least one of M_1 and M_2 . This gives only ρ -mass $1/3$. Therefore, no monotone coupling from μ to ρ can exist.

A.2. A counterexample for the condition $\tilde{C}_\mu^\rho$

We now give a counterexample showing that the condition $\tilde{C}_\mu^\rho$ does not imply stochastic domination. Recall that the condition $\tilde{C}_\mu^\rho$ requires that, for every section s , the pushdown measure associated with μ coincides with the s -marginal of ρ , that is, with the image of ρ under the map

$$x \longmapsto (x_{s(b)})_{b \in B}.$$

Let μ be the uniform measure on the following eight matrices:

$$\begin{aligned} M_1 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & M_2 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, & M_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & M_4 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \\ M_5 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & M_6 &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, & M_7 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & M_8 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \end{aligned}$$

For every section s , the s -marginal of μ is the uniform distribution on $\{0, 1\}^3$.

Let ρ be the uniform measure on the following eight matrices:

$$\begin{aligned} N_1 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, & N_2 &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, & N_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, & N_4 &= \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \\ N_5 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, & N_6 &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, & N_7 &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & N_8 &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \end{aligned}$$

One can check that every s -marginal of ρ is uniform on $\{0, 1\}^3$. Therefore the condition $\tilde{\mathbf{C}}_\mu^\rho$ is satisfied. Nevertheless, ρ does not stochastically dominate μ . Indeed, consider the four matrices

$$M_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad M_6 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_7 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad M_8 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

These four matrices have total μ -mass $1/2$. However, among the matrices $N_1, \dots, N_8$, only N_3, N_6, N_8 dominate at least one of them. This gives only ρ -mass $3/8$. Therefore, no monotone coupling between μ and ρ can exist.

BIBLIOGRAPHY

- [AG91] Michael Aizenman and Geoffrey Grimmett, *Strict monotonicity for critical points in percolation and ferromagnetic models*, J. Stat. Phys. **63** (1991), no. 5-6, 817–835. $\uparrow$ 1231
- [Arm16] Benjamin Armbruster, *A Short Proof of Strassen's Theorem Using Convex Analysis*, 2016, <https://arxiv.org/abs/1603.00137>. $\uparrow$ 1238
- [BBR14] Paul Balister, Béla Bollobás, and Oliver Riordan, *Essential enhancements revisited*, 2014, <https://arxiv.org/abs/1402.0834>. $\uparrow$ 1231
- [BH57] Simon Broadbent and John Hammersley, *Percolation processes: I. Crystals and mazes*, Proc. Camb. Philos. Soc. **53** (1957), no. 3, 629–641. $\uparrow$ 1230
- [BK85] Jacob van Berg and Harry Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. **22** (1985), no. 3, 556–569. $\uparrow$ 1252
- [BLPS99] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm, *Group-invariant percolation on graphs*, Geom. Funct. Anal. **9** (1999), no. 1, 29–66. $\uparrow$ 1230
- [BN11] Vincent Beffara and Pierre Nolin, *On monochromatic arm exponents for 2D critical percolation*, Ann. Probab. **39** (2011), no. 4, 1286–1304. $\uparrow$ 1253
- [BS96] Itai Benjamini and Oded Schramm, *Percolation Beyond $\mathbb{Z}^d$, Many Questions And a Few Answers*, Electron. Commun. Probab. **1** (1996), 71–82. $\uparrow$ 1230, 1231, 1250, 1251, 1254
- [CMT23] Daniel Contreras, Sébastien Martineau, and Vincent Tassion, *Locality of percolation for graphs with polynomial growth*, Electron. Commun. Probab. **28** (2023), article no. 1 (9 pages). $\uparrow$ 1230

- [CTSS24] Natalia Cardona-Tobón, Anja Sturm, and Jan M. Swart, *A min-max random game on a graph that is not a tree*, 2024, <https://arxiv.org/abs/2409.02660>. ↑1263
- [DC18] Hugo Duminil-Copin, *Sixty years of percolation*, Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018, World Scientific; Sociedade Brasileira de Matemática, 2018, pp. 2829–2856. ↑1230
- [GM25] Loïc Gassmann and Ioan Manolescu, *Comparison of arm exponents in planar FK-percolation*, ALEA, Lat. Am. J. Probab. Math. Stat. **22** (2025), no. 2, 1053–1065. ↑1253
- [Lin99] Torgny Lindvall, *On Strassen’s Theorem on Stochastic Domination*, Electron. Commun. Probab. **4** (1999), 51–59. ↑1238
- [MS19] Sébastien Martineau and Franco Severo, *Strict monotonicity of percolation thresholds under covering maps*, Ann. Probab. **47** (2019), no. 6, 4116–4136. ↑1231, 1232, 1244, 1254, 1255, 1256, 1263
- [Rei00] David Reimer, *Proof of the van den Berg–Kesten conjecture*, Comb. Probab. Comput. **9** (2000), no. 1, 27–32. ↑1254
- [RT26] Ritvik R. Radhakrishnan and Vincent Tassion, *Strict inequalities for arm exponents in planar percolation*, Probab. Theory Relat. Fields **195** (2026), no. 1-2, 659–688. ↑1253
- [Str65] Volker Strassen, *The existence of probability measures with given marginals*, Ann. Math. Stat. **36** (1965), no. 2, 423–439. ↑1238
- [Van25] Hugo Vanneuville, *Exponential decay of the volume for Bernoulli percolation: a proof via stochastic comparison*, Ann. Henri Lebesgue **8** (2025), 101–112. ↑1254

Manuscript received on 3rd July 2025,
revised on 12th March 2026,
accepted on 19th May 2026.

Recommended by Editors S. Gouëzel and I. Benjamini.
Published under license CC BY 4.0.



eISSN: 2644-9463

This journal is a member of Centre Mersenne.



Sébastien MARTINEAU
LPSM, Sorbonne Université
4 Place Jussieu, Case 158
75252 Paris Cedex 05 (France)
smartineau@lpsm.paris

Rémy POUDEVIGNE-AUBOIRON
LPSM, Sorbonne Université
4 Place Jussieu, Case 158
75252 Paris Cedex 05 (France)
rpoudevigne@lpsm.paris

Paul RAX
MathNet Team
INRIA Paris
48 Rue Barrault
75013 Paris 13 (France)
paul-pierre.rax@inria.fr
Vibhu DALAL
Appendix Writer
Department of Statistics
Columbia University
1255 Amsterdam Avenue
New York, NY 10027 (USA)
vd2512@columbia.edu