

CAN COHERENT PREDICTIONS BE CONTRADICTIONARY?

KRZYSZTOF BURDZY* ** AND
SOUMIK PAL,* *** *University of Washington*

Abstract

We prove the sharp bound for the probability that two experts who have access to different information, represented by different σ -fields, will give radically different estimates of the probability of an event. This is relevant when one combines predictions from various experts in a common probability space to obtain an aggregated forecast. The optimizer for the bound is explicitly described. This paper was originally titled ‘Contradictory predictions’.

Keywords: Conditional probability; opinion; joint distribution of conditional expectations

2010 Mathematics Subject Classification: Primary 60E15

1. Introduction

Imagine two experts, with access to different information, but sharing the same worldview. We model this by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with two distinct sub- σ -fields \mathcal{G} and \mathcal{H} of \mathcal{F} . The sub- σ -fields represent the information accessible to the two experts, while the common probability space represents their worldview, in the sense that, if one of the experts knew exactly what the other knows, he/she would arrive at exactly the same conclusions. This set-up is sometimes called the problem of combining experts’ opinions under partial information, or more colloquially ‘the wisdom of crowds’, and is a popular topic in statistics and decision theory. In general there are N experts with their individual sub- σ -fields who are all trying to predict the probability of a common event A (such as the event that a particular candidate will win an election). The usual question is whether there is a *coherent* way (see Definition 1 below) to combine or aggregate their predictions to come up with a better forecast; see [6, 8, 9, 13]. This field has attracted renewed interest in the current age of social networks (see [12, 18]). In practice most aggregation techniques are simply weighted linear combination of the individual forecasts. However, [21 and 14] show that linear aggregations are not coherent; they recommend nonlinear combinations. Building on these earlier works, [22] develops a mathematical framework to combine predictions when experts use ‘partially overlapping information sources’, and [7] uses this framework for the case of $N = 2$ experts in prediction markets who take turns in updating their beliefs. Also see [4, 15, 20] for applications to economics, [17] for applications to banking and finance, [19] for applications to meteorology, [23] for applications

Received 23 December 2019; revision received 29 June 2020.

* Postal address: Department of Mathematics, Box 354350, University of Washington, Seattle, WA 98195.

** Email address: burdzy@uw.edu

*** Email address: soumik@uw.edu

© The Author(s), 2021. Published by Cambridge University Press on behalf of Applied Probability Trust.

to maintenance of wind turbines, and [1, 11] for philosophical implications. The problem is also related to modeling insider trading in finance [16], where the insider has more information than the rest of the traders, i.e., $\mathcal{G} \subseteq \mathcal{H}$, although the general non-containment scenario makes sense for two different insiders. For recent developments on the pure-mathematics side, related to [3] and the present article, see [5].

We ask a different, complementary question, not about the aggregate of the various predictions, but about their spread. More specifically, we consider $N = 2$ experts predicting the possibility of a common event A and derive sharp probabilistic bounds on the range of their predictions. In particular, we are interested in when this spread is large, a phenomenon that we call ‘contradictory predictions’. Two people make contradictory predictions if one of them asserts that the probability of A is very small and the other one says that it is very large. We will present a theorem formalizing the idea that the two experts are unlikely to make contradictory predictions even if they have different information sources.

Let us formulate the problem in a quantitative and rigorous manner.

Definition 1. Let A be an event in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, so $A \in \mathcal{F}$, and let $X = \mathbb{P}(A | \mathcal{G})$ and $Y = \mathbb{P}(A | \mathcal{H})$ for two sub- σ -fields \mathcal{G} and \mathcal{H} of \mathcal{F} . Random variables X and Y satisfying these conditions are called *coherent*, and so is their joint distribution. Random variables X and Y are coherent if and only if, for some $A \in \mathcal{F}$,

$$0 \leq X, Y \leq 1 \quad \text{and} \quad X = \mathbb{P}(A | X) \quad \text{and} \quad Y = \mathbb{P}(A | Y)$$

(see [3, (1.2)]).

Let

$$\lambda(\delta) = \sup \mathbb{P}(|X - Y| \geq 1 - \delta), \tag{1.1}$$

where the supremum is taken over all probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$, all events $A \in \mathcal{F}$, and all sub- σ -fields \mathcal{G} and \mathcal{H} of \mathcal{F} . It was proved in Theorem 14.1 of [1] that $\lambda(\delta) \leq 5\delta$ for $\delta \leq 1/10$. The following stronger result was proved by Jim Pitman and was published in [2, Thm. 18.1] with his permission. For all $\delta \in (0, 1)$,

$$\frac{2\delta}{1 + \delta} \leq \lambda(\delta) \leq 2\delta. \tag{1.2}$$

The purpose of this article is to remove the gap between the lower and upper bounds, even though the gap is very small for small δ , i.e., in the most interesting case.

Theorem 1. For all $\delta \in (0, 1/2)$,

$$\lambda(\delta) = \frac{2\delta}{1 + \delta}. \tag{1.3}$$

It has been shown in [3] that $\lambda(\delta)$ is, curiously, discontinuous at $\delta = 1/2$ (see Figure 1). More precisely, $\lambda(\delta) = 1$ for $\delta \geq 1/2$. To see this, construct X and Y so that $\mathbb{P}(X = 1/2) = 1$ and $\mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = 1/2$ (see the discussion around (1.10) in [3]).

As mentioned above, one can interpret X and Y as the opinions of two experts about the probability of A given different sources of information \mathcal{G} and \mathcal{H} , assuming the experts agree on some initial assignment of probability \mathbb{P} to events in \mathcal{F} . The authors of [6, p. 284] pointed out the following:

If X and Y are both produced by ‘experts’, then one should not expect them to be wildly different. For example, it would seem paradoxical if, with X say uniform on $[0, 1]$, one always had $Y = 1 - X$.

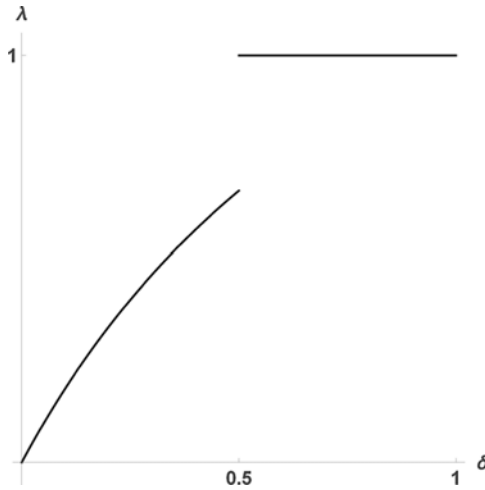


FIGURE 1: The function $\lambda(\delta)$ is discontinuous at $\delta = 1/2$.

This suggests that X and Y cannot be too negatively dependent. However, elementary examples in [6, Section 4.1] show that for any prescribed value of $\mathbb{E}X = \mathbb{E}Y = \mathbb{P}(A) \in (0, 1)$, the correlation between X and Y can take any value in $(-1, 1]$. Consider for instance, for $\delta \in (0, 1)$, the distribution of (X, Y) concentrated on the three points $(1 - \delta, 1 - \delta)$, $(0, 1 - \delta)$, and $(1 - \delta, 0)$, with

$$\mathbb{P}(X = Y) = \mathbb{P}(1 - \delta, 1 - \delta) = \frac{1 - \delta}{1 + \delta} \quad \text{and} \quad \mathbb{P}(0, 1 - \delta) = \mathbb{P}(1 - \delta, 0) = \frac{\delta}{1 + \delta}. \quad (1.4)$$

This example from [10] gives coherent X and Y with correlation $\rho(X, Y) = -\delta$ which can be any value in $(-1, 0)$.

We end this section with a proof of (1.2) borrowed from [2, Theorem 18.1], because it is simple and it underscores the huge gap in complexity between the proof of the upper bound in (1.2) and our proof of the upper bound in (1.3).

Proposition 1. (J. Pitman.) *For all $\delta \in (0, 1)$,*

$$\frac{2\delta}{1 + \delta} \leq \lambda(\delta) \leq 2\delta. \quad (1.5)$$

Proof. We will use notation matching that in the proof of our main theorem.

The lower bound for $\lambda(\delta)$ is attained by the example in (1.4) with $A = \{X = Y\}$.

For the upper bound, it will suffice to consider the case $0 < \delta < 1/2$. Note that for any $0 < \delta < 1/2$ and any random variables X and Y with $0 \leq X \leq 1$ and $0 \leq Y \leq 1$,

$$\{|X - Y| \geq 1 - \delta\} \subset \{X \leq \delta, Y \geq 1 - \delta\} \cup \{Y \leq \delta, X \geq 1 - \delta\}. \quad (1.6)$$

Since $X = \mathbb{E}(\mathbb{1}_A | X)$,

$$\mathbb{P}(X \leq \delta, Y \geq 1 - \delta, A) \leq \mathbb{P}(X \leq \delta, A) = \mathbb{E}(\mathbb{1}_{\{X \leq \delta\}} X) \leq \delta \mathbb{P}(X \leq \delta).$$

We have $1 - Y = \mathbb{E}(\mathbb{1}_{A^c} | Y)$, so

$$\begin{aligned} \mathbb{P}(X \leq \delta, Y \geq 1 - \delta, A^c) &\leq \mathbb{P}(Y \geq 1 - \delta, A^c) = \mathbb{E}(\mathbb{1}_{\{1 - Y \leq \delta\}}(1 - Y)) \\ &\leq \delta \mathbb{P}(Y \geq 1 - \delta). \end{aligned}$$

It follows that

$$\mathbb{P}(X \leq \delta, Y \geq 1 - \delta) \leq \delta(\mathbb{P}(X \leq \delta) + \mathbb{P}(Y \geq 1 - \delta)) \tag{1.7}$$

and similarly

$$\mathbb{P}(Y \leq \delta, X \geq 1 - \delta) \leq \delta(\mathbb{P}(Y \leq \delta) + \mathbb{P}(X \geq 1 - \delta)). \tag{1.8}$$

For $0 < \delta < 1/2$ the events $X \leq \delta$ and $X \geq 1 - \delta$ are disjoint, so $\mathbb{P}(X \leq \delta) + \mathbb{P}(X \geq 1 - \delta) \leq 1$, and the same holds for Y . Add (1.7) and (1.8) and use (1.6) to obtain the upper bound in (1.5). □

2. Proof of Theorem 1

The proof of Theorem 1, our main result, will consist of a sequence of lemmas.

The first of these, Lemma 1, is essentially the proof of Theorem 1 in the case when the σ -fields \mathcal{G} and \mathcal{H} are very simple. The lemma is applied only once, at the very end of the proof of Theorem 1 (at the end of the paper).

Lemma 2 contains an elementary argument showing that one can discretize the problem, i.e., that it is enough to prove Theorem 1 in the case when \mathcal{G} and \mathcal{H} are finite.

The remaining lemmas form a reduction argument. Each lemma presents and analyzes a transformation that either decreases the size of \mathcal{G} or \mathcal{H} (and does not increase the size of the other σ -field), or moves the probability mass around the probability space to make it more ‘structured’. The latter aspect of the argument resembles a sliding puzzle with 15 squares, or a Rubik’s Cube. A more detailed and rigorous description of this part of the proof is given below, between the proof of Lemma 2 and the statement of Lemma 3.

We end our introductory remarks about the proof with the observation that our proof is constructive in the following sense. If a probability space is given explicitly, i.e., all relevant probabilities have known numerical values, then all steps of our argument can be implemented in an algorithm that will generate a finite sequence of probability spaces such that, for a fixed δ , the value of $\mathbb{P}(|X - Y| \geq 1 - \delta)$ will increase or stay the same with every transformation. As a result, for any probability space one can effectively construct a much simpler probability space with $\mathbb{P}(|X - Y| \geq 1 - \delta)$ greater than or equal to that for the original probability space.

Fix any $\delta \in (0, 1/2)$. Most elements of the model (sets, probabilities) will change within this section, but the value of δ will remain fixed.

Lemma 1. *Consider any two events G and H such that $\mathbb{P}(G) > 0$ and $\mathbb{P}(H) > 0$. If*

$$|\mathbb{P}(A | G) - \mathbb{P}(A | H)| \geq 1 - \delta$$

then

$$\mathbb{P}(G \cap H) \leq \frac{\delta}{1 + \delta}(\mathbb{P}(G) + \mathbb{P}(H)). \tag{2.1}$$

Proof. Assume without loss of generality that

$$\mathbb{P}(A | G) - \mathbb{P}(A | H) \geq 1 - \delta.$$

Suppose that we can prove that

$$\mathbb{P}((G \cap H^c) \cup (H \cap G^c) | G \cup H) \geq \mathbb{P}(A | G) - \mathbb{P}(A | H). \tag{2.2}$$

Then

$$\begin{aligned} \frac{\mathbb{P}(G \cap H)}{\mathbb{P}(G \cup H)} &= \mathbb{P}(G \cap H \mid G \cup H) = 1 - \mathbb{P}((G \cap H^c) \cup (H \cap G^c) \mid G \cup H) \\ &\leq 1 - (\mathbb{P}(A \mid G) - \mathbb{P}(A \mid H)) \leq \delta, \end{aligned}$$

and, therefore,

$$\begin{aligned} \mathbb{P}(G \cap H) &\leq \delta \mathbb{P}(G \cup H) = \delta(\mathbb{P}(G) + \mathbb{P}(H)) - \delta \mathbb{P}(G \cap H), \\ (1 + \delta)\mathbb{P}(G \cap H) &\leq \delta(\mathbb{P}(G) + \mathbb{P}(H)). \end{aligned}$$

This implies (2.1), so it will suffice to prove (2.2).

Let

$$\begin{aligned} p_1 &= \mathbb{P}(A \mid G \cap H^c), & p_2 &= \mathbb{P}(A \mid G \cap H), & p_3 &= \mathbb{P}(A \mid G^c \cap H), \\ q_1 &= \mathbb{P}(G \cap H^c), & q_2 &= \mathbb{P}(G \cap H), & q_3 &= \mathbb{P}(G^c \cap H). \end{aligned}$$

Then

$$\mathbb{P}(A \mid G) - \mathbb{P}(A \mid H) = \frac{p_1 q_1 + p_2 q_2}{q_1 + q_2} - \frac{p_2 q_2 + p_3 q_3}{q_2 + q_3} \leq \frac{q_1 + p_2 q_2}{q_1 + q_2} - \frac{p_2 q_2}{q_2 + q_3}. \tag{2.3}$$

For $p_2 = 0$ we obtain

$$\frac{q_1 + p_2 q_2}{q_1 + q_2} - \frac{p_2 q_2}{q_2 + q_3} = \frac{q_1}{q_1 + q_2} \leq \frac{q_1 + q_3}{q_1 + q_2 + q_3} = \mathbb{P}((G \cap H^c) \cup (H \cap G^c) \mid G \cup H). \tag{2.4}$$

For $p_2 = 1$,

$$\begin{aligned} \frac{q_1 + p_2 q_2}{q_1 + q_2} - \frac{p_2 q_2}{q_2 + q_3} &= \frac{q_1 + q_2}{q_1 + q_2} - \frac{q_2}{q_2 + q_3} \leq \frac{q_1 + q_3}{q_1 + q_2 + q_3} \\ &= \mathbb{P}((G \cap H^c) \cup (H \cap G^c) \mid G \cup H). \end{aligned} \tag{2.5}$$

The right-hand side of (2.3) is a linear function of p_2 , so (2.3)–(2.5) imply that for all $p_2 \in [0, 1]$,

$$\mathbb{P}(A \mid G) - \mathbb{P}(A \mid H) \leq \frac{q_1 + p_2 q_2}{q_1 + q_2} - \frac{p_2 q_2}{q_2 + q_3} \leq \mathbb{P}((G \cap H^c) \cup (H \cap G^c) \mid G \cup H).$$

This completes the proof of (2.2). □

Lemma 2. Recall $\lambda(\delta)$ defined in (1.1), and let

$$\lambda'(\delta) = \sup \mathbb{P}(|X - Y| \geq 1 - \delta),$$

where the supremum is taken over all probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$, all events $A \in \mathcal{F}$, and all finitely generated sub- σ -fields \mathcal{G} and \mathcal{H} of \mathcal{F} . If $\lambda'(\delta) = 2\delta/(1 + \delta)$ for all $\delta \in (0, 1/2)$, then $\lambda(\delta) = 2\delta/(1 + \delta)$ for all $\delta \in (0, 1/2)$.

Proof. It is obvious that $\lambda'(\delta) \leq \lambda(\delta)$. It remains to prove the opposite inequality.

For $n > 1$, let \mathcal{G}_n be the σ -field generated by the events $\{k/n \leq X < (k + 1)/n\}$, $0 \leq k \leq n$ (note that $k = n$ is included). Let $X_n = \mathbb{P}(A \mid \mathcal{G}_n)$. Since

$$\mathbb{P}(\{X_n \notin [k/n, (k + 1)/n)\} \cap \{k/n \leq X < (k + 1)/n\}) = 0,$$

we have

$$\mathbb{P}(|X_n - X| \leq 1/n) = 1.$$

Similarly, let \mathcal{H}_n be the σ -field generated by the events $\{k/n \leq Y < (k + 1)/n\}$, $0 \leq k \leq n$, and $Y_n = \mathbb{P}(A | \mathcal{H}_n)$. Then

$$\mathbb{P}(|Y_n - Y| \leq 1/n) = 1.$$

It follows that

$$\mathbb{P}(|X - Y| \geq 1 - \delta) \leq \mathbb{P}(|X_n - Y_n| \geq 1 - \delta - 2/n). \tag{2.6}$$

Since n can be arbitrarily large, (2.6) implies that if one can prove that $\lambda'(\delta) = 2\delta/(1 + \delta)$ for all $\delta \in (0, 1/2)$, then $\lambda(\delta) = 2\delta/(1 + \delta)$ for all $\delta \in (0, 1/2)$. \square

Notation 1. In view of Lemma 2 we can and will assume from now on that \mathcal{G} and \mathcal{H} are finitely generated. Let $\{G_1, G_2, \dots, G_{m(\mathcal{G})}\}$ be the partition of Ω generating \mathcal{G} , and let the family $\{H_1, H_2, \dots, H_{m(\mathcal{H})}\}$ be defined in the analogous way relative to \mathcal{H} .

We will assume without loss of generality that $\mathbb{P}(G_k) > 0$ and $\mathbb{P}(H_k) > 0$ for all k .

Let

$$p_k = \mathbb{P}(G_k), \quad 1 \leq k \leq m(\mathcal{G}), \tag{2.7}$$

$$q_k = \mathbb{P}(H_k), \quad 1 \leq k \leq m(\mathcal{H}), \tag{2.8}$$

$$C_{j,k} = G_j \cap H_k, \tag{2.9}$$

$$x_k = X(\omega) \quad \text{for } \omega \in G_k, \quad 1 \leq k \leq m(\mathcal{G}), \tag{2.10}$$

$$y_k = Y(\omega) \quad \text{for } \omega \in H_k, \quad 1 \leq k \leq m(\mathcal{H}). \tag{2.11}$$

Events $C_{k,j}$ will be called cells. Assume without loss of generality that

$$x_1 \leq x_2 \leq \dots \leq x_{m(\mathcal{G})}, \quad y_1 \leq y_2 \leq \dots \leq y_{m(\mathcal{H})}. \tag{2.12}$$

In Figure 2, the values of x_k grow from the left to the right, and the values of y_k grow from the bottom to the top.

Let

$$m_-(\mathcal{G}) = \max\{k : \exists j \text{ such that } y_j - x_k \geq 1 - \delta\}, \tag{2.13}$$

$$m_+(\mathcal{G}) = \min\{k : \exists j \text{ such that } x_k - y_j \geq 1 - \delta\}, \tag{2.14}$$

$$m_-(\mathcal{H}) = \max\{k : \exists j \text{ such that } x_j - y_k \geq 1 - \delta\}, \tag{2.15}$$

$$m_+(\mathcal{H}) = \min\{k : \exists j \text{ such that } y_k - x_j \geq 1 - \delta\}. \tag{2.16}$$

By convention, $\max(\emptyset) = 0$ and $\min(\emptyset) = \infty$. Let

$$B = \{|X - Y| \geq 1 - \delta\}. \tag{2.17}$$

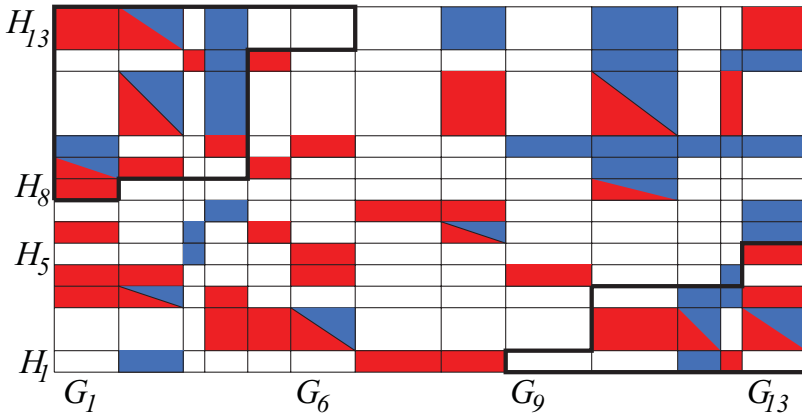


FIGURE 2: White represents ‘empty’ cells, i.e., cells $C_{k,j}$ such that $\mathbb{P}(C_{k,j}) = 0$. Red designates ‘cells in A^c ’, i.e., cells $C_{k,j}$ such that $\mathbb{P}(C_{k,j}) > \mathbb{P}(C_{k,j} \cap A) = 0$. Blue designates ‘cells in A ’, i.e., cells $C_{k,j}$ such that $\mathbb{P}(C_{k,j}) > \mathbb{P}(C_{k,j} \cap A^c) = 0$. Rectangles colored both red and blue represent cells $C_{k,j}$ that contain both A and A^c , i.e., $\mathbb{P}(C_{k,j} \cap A) > 0$ and $\mathbb{P}(C_{k,j} \cap A^c) > 0$. The event B is contained within two thick closed polygonal lines. In this illustration, $m(\mathcal{G}) = m(\mathcal{H}) = 13$, $m_-(\mathcal{G}) = 6$, $m_+(\mathcal{G}) = 9$, $m_-(\mathcal{H}) = 5$, and $m_+(\mathcal{H}) = 8$. Cells have different areas to illustrate the point that their probabilities may be unequal, but no conditions on probability values should be inferred from cell areas.

The definitions (2.13)–(2.16) are illustrated in Figure 2 as follows. The projection of the set B (the cells within two closed thick polygonal lines) onto the horizontal axis consists of the two disjoint intervals $[1, m_-(\mathcal{G})]$ and $[m_+(\mathcal{G}), m(\mathcal{G})]$. The projection of the set B onto the vertical axis consists of the two disjoint intervals $[1, m_-(\mathcal{H})]$ and $[m_+(\mathcal{H}), m(\mathcal{H})]$.

We will define a number of transformations of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, event A , and σ -fields \mathcal{G} and \mathcal{H} . We will denote the transformed objects by $(\Omega', \mathcal{F}', \mathbb{P}, A', \mathcal{G}', \text{ and } \mathcal{H}')$. We will also write in a similar manner X', p_k', x_k' , etc. Note that the only exception to this rule is \mathbb{P} . We will write \mathbb{P} instead of \mathbb{P}' (except in Lemma 3) even though the transformed probability measure is not necessarily equal to the original one. This should not cause any confusion. We will denote our transformations by $\mathcal{T}_1, \mathcal{T}_2$, etc., and we will write

$$\mathbf{S} = (\Omega, \mathcal{F}, \mathbb{P}, A, \mathcal{G}, \mathcal{H}), \quad \mathbf{S}' = (\Omega', \mathcal{F}', \mathbb{P}, A', \mathcal{G}', \mathcal{H}'), \quad \mathcal{T}(\mathbf{S}) = \mathbf{S}'.$$

All objects that are not listed in the definition of \mathbf{S} (for example X, Y , and those in (2.7)–(2.11)) are uniquely defined given \mathbf{S} .

Our definitions of transformations will contain some assumptions about \mathbf{S} . If the assumptions are not satisfied, the transformation should be interpreted as the identity, i.e., $\mathcal{T}(\mathbf{S}) = \mathbf{S}$.

Note that $\mathbb{P}(B) = \sum_{1 \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B)$. Most of our transformations will satisfy the following three conditions:

$$\begin{aligned} \mathbb{P}(B') &= \sum_{1 \leq j \leq m'(\mathcal{G}'), 1 \leq k \leq m'(\mathcal{H}')} \mathbb{P}(C'_{j,k} \cap B') \geq \sum_{1 \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B) = \mathbb{P}(B), & (2.18) \\ m'(\mathcal{G}') &\leq m(\mathcal{G}), & m'(\mathcal{H}') &\leq m(\mathcal{H}). & (2.19) \end{aligned}$$

If

$$\sum_{1 \leq j \leq m_-(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B) > 0 \text{ and } \sum_{m_+(\mathcal{G}) \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B) > 0, \quad (2.20)$$

then

$$\sum_{1 \leq j \leq m'_-(\mathcal{G}'), 1 \leq k \leq m'(\mathcal{H}')} \mathbb{P}(C'_{j,k} \cap B') > 0 \text{ and} \tag{2.21}$$

$$\sum_{m'_+(\mathcal{G}') \leq j \leq m'(\mathcal{G}'), 1 \leq k \leq m'(\mathcal{H}')} \mathbb{P}(C'_{j,k} \cap B') > 0.$$

The meaning of the condition in (2.20)–(2.21) is that if the part of B in the upper left corner in Figure 2 is non-empty and has a strictly positive \mathbb{P} -measure, and the same is true for the part of B in the lower right corner, then the same holds for S' .

Lemma 3. *Suppose that S is given and $\mathbb{P}(B) > 0$. Then for every $\varepsilon > 0$ there exists S' such that $\mathbb{P}(B') > \mathbb{P}(B) - \varepsilon$, $m'(\mathcal{G}') \leq m(\mathcal{G}) + 1$, $m'(\mathcal{H}') \leq m(\mathcal{H}) + 1$, and (2.21) holds.*

Proof. Since $\mathbb{P}(B) > 0$, we must have either $m_-(\mathcal{G}) \geq 1$, $m_+(\mathcal{H}) \leq m(\mathcal{H})$, and

$$\sum_{1 \leq j \leq m_-(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B) > 0, \tag{2.22}$$

or $m_-(\mathcal{H}) \geq 1$, $m_+(\mathcal{G}) \leq m(\mathcal{G})$, and

$$\sum_{m_+(\mathcal{G}) \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \mathbb{P}(C_{j,k} \cap B) > 0. \tag{2.23}$$

By symmetry, it will suffice to discuss only one of these cases.

Suppose that $m_-(\mathcal{G}) \geq 1$, $m_+(\mathcal{H}) \leq m(\mathcal{H})$, and (2.22) holds. If $m_-(\mathcal{H}) \geq 1$, $m_+(\mathcal{G}) \leq m(\mathcal{G})$, and (2.23) is true, then we set $S' = S$ and we are done. Suppose otherwise. We will assume that $m_-(\mathcal{H}) = 0$ and $m_+(\mathcal{G}) = \infty$. It is easy to see that the argument given below applies also in the case when $m_-(\mathcal{H}) \geq 1$, $m_+(\mathcal{G}) \leq m(\mathcal{G})$, and (2.23) does not hold.

Let

$$\Omega' = \Omega \cup \{\omega'_{k,0} : 1 \leq k \leq m(\mathcal{G}) + 1\} \cup \{\omega'_{m(\mathcal{G})+1,j} : 1 \leq j \leq m(\mathcal{H})\},$$

where the $\omega'_{k,j}$ are all distinct and do not belong to Ω . Let

$$\begin{aligned} C'_{k,j} &= C_{k,j}, & 1 \leq k \leq m(\mathcal{G}), 1 \leq j \leq m(\mathcal{H}), \\ C'_{k,0} &= \{\omega'_{k,0}\}, & 1 \leq k \leq m(\mathcal{G}) + 1, \\ C'_{m(\mathcal{G})+1,j} &= \{\omega'_{m(\mathcal{G})+1,j}\}, & 1 \leq j \leq m(\mathcal{H}), \\ G'_k &= \bigcup_{0 \leq j \leq m(\mathcal{H})} C'_{k,j}, & 1 \leq k \leq m(\mathcal{G}) + 1, \\ H'_j &= \bigcup_{1 \leq k \leq m(\mathcal{G})+1} C'_{k,j}, & 0 \leq j \leq m(\mathcal{H}), \\ \mathcal{G}' &= \sigma\{G'_k, 1 \leq k \leq m(\mathcal{G}) + 1\}, \\ \mathcal{H}' &= \sigma\{H'_j, 0 \leq j \leq m(\mathcal{H})\}, \\ A' &= A \cup G'_{m(\mathcal{G})+1}. \end{aligned}$$

Fix any $\varepsilon > 0$. Let $\varepsilon_1 > 0$ be so small that

$$\mathbb{P}(B)(1 - \varepsilon_1/2 - \varepsilon_1/4) > \mathbb{P}(B) - \varepsilon. \tag{2.24}$$

It is elementary to see that we can define a probability measure \mathbb{P}' on Ω' with the following properties:

$$\mathbb{P}'(H'_0 \setminus C'_{m(\mathcal{G})+1,0}) = \varepsilon_1/2,$$

$$\mathbb{P}'(C'_{m(\mathcal{G})+1,0}) > 0,$$

$$\mathbb{P}'(G'_{m(\mathcal{G})+1}) = \varepsilon_1\delta/4,$$

$$\mathbb{P}'(F) = (1 - \varepsilon_1/2 - \varepsilon_1\delta/4)\mathbb{P}(F), \quad F \subset \Omega.$$

Note that $x_{m(\mathcal{G})+1} = 1$ and

$$y_0 = \frac{\mathbb{P}(A' \cap H'_0)}{\mathbb{P}(H'_0)} \leq \frac{\mathbb{P}(C'_{m(\mathcal{G})+1,0})}{\mathbb{P}(H'_0 \setminus C'_{m(\mathcal{G})+1,0})} \leq \frac{\mathbb{P}(G'_{m(\mathcal{G})+1})}{\mathbb{P}(H'_0 \setminus C'_{m(\mathcal{G})+1,0})} = \frac{\varepsilon_1\delta/4}{\varepsilon_1/2} = \delta/2.$$

Hence, $C'_{m(\mathcal{G})+1,0} \subset B'$. Let us shift the index for the generators H'_j of \mathcal{H}' so that the generators are $H'_1, H'_2, \dots, H'_{m(\mathcal{H})+1}$. We see that $m'(\mathcal{G}') = m(\mathcal{G}) + 1$, $m'(\mathcal{H}') = m(\mathcal{H}) + 1$, $m'_-(\mathcal{H}') \geq 1$, $m'_+(\mathcal{G}') \leq m'(\mathcal{G}')$, and (2.21) holds. It remains to note that, in view of (2.24),

$$\mathbb{P}'(B') \geq \mathbb{P}'(B) = \mathbb{P}(B)(1 - \varepsilon_1/2 - \varepsilon_1\delta/4) > \mathbb{P}(B) - \varepsilon. \quad \square$$

Lemma 4. Suppose that for some \mathbf{S} , j_1 , and k_1 , we have $\mathbb{P}(C_{j_1,k_1} \cap A) = p_1 > 0$, and consider any $p_2 \in (0, p_1)$. There exists \mathbf{S}' such that $m'(\mathcal{G}') = m(\mathcal{G})$, $m'(\mathcal{H}') = m(\mathcal{H})$,

$$\mathbb{P}(C'_{j,k} \cap A') = \mathbb{P}(C_{j,k} \cap A), \quad \mathbb{P}(C'_{j,k} \cap (A')^c) = \mathbb{P}(C_{j,k} \cap A^c),$$

for all $1 \leq j \leq m(\mathcal{G})$ and $1 \leq k \leq m(\mathcal{H})$, and there exists an event $F \subset C'_{j_1,k_1} \cap A'$ with $\mathbb{P}(F) = p_2$. Moreover, the conditions (2.18)–(2.21) are satisfied.

Proof. Let Ω' be defined as follows:

$$\Omega' = \{\omega'_{j,k,1}\}_{1 \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \cup \{\omega'_{j,k,2}\}_{1 \leq j \leq m(\mathcal{G}), 1 \leq k \leq m(\mathcal{H})} \cup \{\omega'_{j_1,k_1,3}\},$$

where all listed elements are distinct. Let

$$C'_{j,k} \cap A' = \{\omega'_{j,k,1}\}, \quad C'_{j,k} \cap (A')^c = \{\omega'_{j,k,2}\},$$

for all $1 \leq j \leq m(\mathcal{G})$ and $1 \leq k \leq m(\mathcal{H})$, except that $C'_{j_1,k_2} \cap A' = \{\omega'_{j_1,k_1,1}, \omega'_{j_1,k_1,3}\}$. Let $F = \{\omega'_{j_1,k_1,1}\}$.

Define the probability measure on Ω' by

$$\mathbb{P}(C'_{j,k} \cap A') = \mathbb{P}(C_{j,k} \cap A), \quad \mathbb{P}(C'_{j,k} \cap (A')^c) = \mathbb{P}(C_{j,k} \cap A^c),$$

for all $1 \leq j \leq m(\mathcal{G})$ and $1 \leq k \leq m(\mathcal{H})$, and

$$\mathbb{P}(F) = \mathbb{P}(\{\omega'_{j_1,k_1,1}\}) = p_2, \quad \mathbb{P}(\{\omega'_{j_1,k_1,3}\}) = p_1 - p_2.$$

It is elementary to check that \mathbf{S}' satisfies the lemma. □

Lemma 5. *Suppose that there exists $k \in \{1, \dots, m(\mathcal{G}) - 1\}$ such that for every $j \in \{1, \dots, m(\mathcal{H})\}$ we have*

$$C_{k,j} \cup C_{k+1,j} \subset B \quad \text{or} \quad \mathbb{P}((C_{k,j} \cup C_{k+1,j}) \cap B) = 0. \tag{2.25}$$

Then there exists \mathbf{S}' such that (2.18) and (2.20)–(2.21) are satisfied, and

$$m'(\mathcal{G}') = m(\mathcal{G}) - 1, \quad m'(\mathcal{H}') \leq m(\mathcal{H}). \tag{2.26}$$

In graphical terms, the condition (2.25) means that the thick polygonal line extends in a straight fashion along the boundaries of at least two consecutive cells in the interior of the rectangle in Figure 2. There are several such ‘long’ thick line segments in Figure 2.

Proof of Lemma 5. Let

$$G'_j = G_j, \quad 1 \leq j < k, \tag{2.27}$$

$$G'_k = G_k \cup G_{k+1}, \tag{2.28}$$

$$G'_j = G_{j+1}, \quad k + 1 \leq j \leq m(\mathcal{G}) - 1. \tag{2.29}$$

Let \mathcal{G}' be the σ -field generated by $\{G'_j\}_{1 \leq j \leq m(\mathcal{G})-1}$. All other objects in \mathbf{S} remain unchanged; for example, $A' = A$, $\mathcal{H}' = \mathcal{H}$, etc. It follows from (2.27)–(2.29) that for $1 \leq j \leq m(\mathcal{H})$,

$$\begin{aligned} C'_{r,j} &= C_{r,j}, \quad 1 \leq r < k, \\ C'_{k,j} &= C_{k,j} \cup C_{k+1,j}, \\ C'_{r,j} &= C_{r+1,j}, \quad k + 1 \leq r \leq m(\mathcal{G}) - 1. \end{aligned}$$

Hence, to prove (2.18), it will suffice to show that for all j ,

$$\mathbb{P}(C'_{k,j} \cap B) \geq \mathbb{P}(C_{k,j} \cap B) + \mathbb{P}(C_{k+1,j} \cap B). \tag{2.30}$$

It is elementary to check that (2.28) implies that $x_k \leq x'_k \leq x_{k+1}$. In view of (2.25), for every j , we have either

$$|x_k - y_j| \geq 1 - \delta \quad \text{and} \quad |x_{k+1} - y_j| \geq 1 - \delta \tag{2.31}$$

or

$$|x_k - y_j| < 1 - \delta \quad \text{and} \quad |x_{k+1} - y_j| < 1 - \delta. \tag{2.32}$$

If (2.31) is true then $|x'_k - y_j| \geq 1 - \delta$, so $C'_{k,j} \subset B$, and therefore (2.30) is satisfied. In the case (2.32), the condition (2.30) holds because the right-hand side is 0. We have finished the proof of (2.18).

It follows from (2.27)–(2.29) that (2.26) is satisfied.

Some cells that belong to B can coalesce in the upper left corner in Figure 2, if there are any to start with, but they cannot disappear. The same remark applies to the lower right corner, so the condition (2.20)–(2.21) holds. □

Lemma 6. For any \mathbf{S} there exists \mathbf{S}' such that (2.18)–(2.21) are satisfied and

$$m'_-(\mathcal{G}') = m'(\mathcal{H}') - m'_+(\mathcal{H}') + 1, \tag{2.33}$$

$$m'_-(\mathcal{H}') = m'(\mathcal{G}') - m'_+(\mathcal{G}') + 1, \tag{2.34}$$

$$y'_{m'(\mathcal{H}')-k+1} - x'_k \geq 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{G}'), \tag{2.35}$$

$$y'_{m'(\mathcal{H}')-k+1} - x'_{k+1} < 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{G}'), \tag{2.36}$$

$$y'_{m'(\mathcal{H}')-k} - x'_k < 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{G}'), \tag{2.37}$$

$$x'_{m'(\mathcal{G}')-k+1} - y'_k \geq 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{H}'), \tag{2.38}$$

$$x'_{m'(\mathcal{G}')-k+1} - y'_{k+1} < 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{H}'), \tag{2.39}$$

$$x'_{m'(\mathcal{G}')-k} - y'_k < 1 - \delta, \quad 1 \leq k \leq m'_-(\mathcal{H}'), \tag{2.40}$$

$$x'_1 < x'_2 < \dots < x'_{m(\mathcal{G}')} \quad \text{and} \quad y'_1 < y'_2 < \dots < y'_{m(\mathcal{H}')}. \tag{2.41}$$

Moreover,

$$\mathbf{S}' = \mathbf{S} \quad \text{or} \quad m'(\mathcal{G}') < m(\mathcal{G}) \quad \text{or} \quad m'(\mathcal{H}') < m(\mathcal{H}). \tag{2.42}$$

Proof. We apply the transformation defined in Lemma 5 repeatedly, as long as there is some k satisfying (2.25). In view of (2.26), these transformations strictly decrease $m(\mathcal{G})$, so they have to stop at some point, either when $m(\mathcal{G}) = 1$ or when (2.25) is not satisfied any more. Then we exchange the roles of \mathcal{G} and \mathcal{H} and we collapse pairs of ‘rows’ H_k and H_{k+1} into H'_k in a similar manner until the condition analogous to (2.25) fails.

If \mathbf{S}' is the final result of the transformations described above then \mathbf{S}' may be represented graphically as follows. The thick polygonal line must turn at every cell corner in the interior of the rectangle (see Figure 4), so that the condition (2.25) is not satisfied, and neither is its analogue with the roles of \mathcal{G} and \mathcal{H} exchanged. In other words, the boundary of B in the interior of the rectangle is a zigzag line that turns at every opportunity. It is elementary to check that the conditions (2.33)–(2.40) are a rigorous version of this graphical description.

To see that (2.41) holds, recall that we have weak inequalities in view of (2.12). If any weak inequality in the first set is an equality, say $x'_k = x'_{k+1}$, then (2.25) is satisfied and we can reduce the number of generators $m'(\mathcal{G}')$ using the transformation defined in Lemma 5. A similar argument applies to the second set of inequalities in (2.41). However, by the argument given in the first part of this proof, $m'(\mathcal{G}')$ and $m'(\mathcal{H}')$ cannot be decreased any more by the transformation defined in Lemma 5, so (2.41) must be true.

The conditions (2.18)–(2.21) are satisfied because this is the case for the transformation defined in Lemma 5. The alternative stated in (2.42) follows from (2.26). \square

Notation 2. If $(k, j) = (k, m(\mathcal{H}) - k + 1)$ and the conditions (2.35)–(2.37), without primes, are satisfied, then we will write $(k, j) \in \mathcal{D}_-$. If $(j, k) = (m(\mathcal{G}) - k + 1, k)$ and the conditions (2.38)–(2.40), without primes, are satisfied, then we will write $(j, k) \in \mathcal{D}_+$. In Figure 4, the cells belonging to \mathcal{D}_- and \mathcal{D}_+ are crossed. The set $\mathcal{D}_- \cup \mathcal{D}_+$ is the ‘internal border’ of B . Note that if $(k, j) \in \mathcal{D}_-$ then $x_k \leq \delta$ and $y_j \geq 1 - \delta$. If $(k, j) \in \mathcal{D}_+$ then $x_k \geq 1 - \delta$ and $y_j \leq \delta$.

Lemma 7. Suppose that \mathbf{S} satisfies (2.33)–(2.41) and there exist $k \in \{1, \dots, m(\mathcal{G}) - 1\}$ and $i \in \{1, \dots, m(\mathcal{H})\}$ such that

$$C_{k,i} \subset B, \quad C_{k+1,i} \subset B^c, \quad \mathbb{P}(C_{k,i}) = 0. \tag{2.43}$$

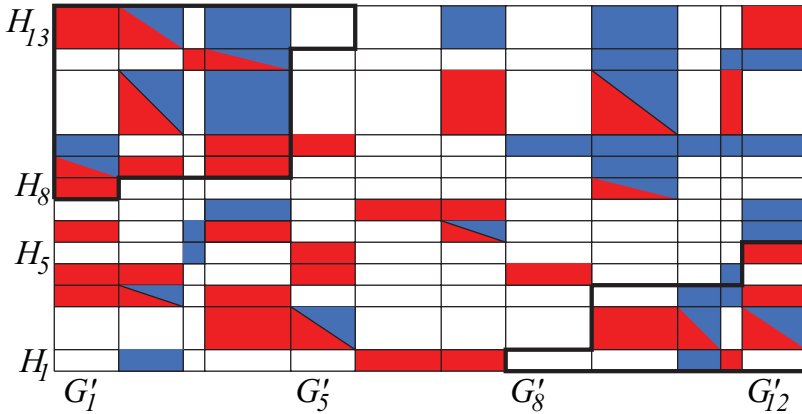


FIGURE 3: This configuration S' has been obtained by transforming the S of Figure 2 according to (2.46). For the meanings of the colors and thick lines, see Figure 2. Columns G_4 and G_5 in Figure 2 have been combined to form column G'_4 in the present figure. This is because in row H_9 in Figure 2, the rightmost cell in B (within the thick polygonal line) is white, i.e., its probability is 0.

Then there exists S' such that (2.18) and (2.20)–(2.21) are satisfied, and

$$m'(\mathcal{G}') = m(\mathcal{G}) - 1, \quad m'(\mathcal{H}') \leq m(\mathcal{H}). \tag{2.44}$$

The condition (2.43) is illustrated in Figure 2 as follows: in row H_9 , the rightmost cell in B (within the thick polygonal line) is white; i.e., its probability is 0. The transformation defined in Lemma 7 is illustrated in Figures 2 and 3 (see the caption of Figure 3).

Proof of Lemma 7. Let

$$G'_j = G_j, \quad 1 \leq j < k, \tag{2.45}$$

$$G'_k = G_k \cup G_{k+1}, \tag{2.46}$$

$$G'_j = G_{j+1}, \quad k + 1 \leq j \leq m(\mathcal{G}) - 1. \tag{2.47}$$

Let \mathcal{G}' be the σ -field generated by $\{G'_j\}_{1 \leq j \leq m(\mathcal{G})-1}$. All other objects in S remain unchanged; for example, $A' = A$, $\mathcal{H}' = \mathcal{H}$, etc. It follows from (2.45)–(2.47) that for $1 \leq j \leq m(\mathcal{H})$,

$$C'_{r,j} = C_{r,j}, \quad 1 \leq r < k, \tag{2.48}$$

$$C'_{k,j} = C_{k,j} \cup C_{k+1,j}, \tag{2.49}$$

$$C'_{r,j} = C_{r+1,j}, \quad k + 1 \leq r \leq m(\mathcal{G}) - 1. \tag{2.50}$$

Hence, to prove (2.18), it will suffice to show that for all j ,

$$\mathbb{P}(C'_{k,j} \cap B') \geq \mathbb{P}(C_{k,j} \cap B) + \mathbb{P}(C_{k+1,j} \cap B). \tag{2.51}$$

It is elementary to check that (2.46) implies that $x_k \leq x'_k \leq x_{k+1}$.

Consider $j \neq i$. It follows from the conditions (2.33)–(2.40) (see the remark about the zigzag line in the proof of Lemma 6) that

$$C_{k,j} \cup C_{k+1,j} \subset B \quad \text{or} \quad \mathbb{P}((C_{k,j} \cup C_{k+1,j}) \cap B) = 0.$$

In the first case,

$$|x_k - y_j| \geq 1 - \delta \quad \text{and} \quad |x_{k+1} - y_j| \geq 1 - \delta.$$

This implies that $|x'_k - y'_j| = |x'_k - y_j| \geq 1 - \delta$, so $C'_{k,j} \subset B'$, and therefore

$$\mathbb{P}(C'_{k,j} \cap B') = \mathbb{P}(C'_{k,j}) = \mathbb{P}(C_{k,j}) + \mathbb{P}(C_{k+1,j}) \geq \mathbb{P}(C_{k,j} \cap B) + \mathbb{P}(C_{k+1,j} \cap B).$$

Hence (2.51) is satisfied.

If $\mathbb{P}((C_{k,j} \cup C_{k+1,j}) \cap B) = 0$, then (2.51) holds because its right-hand side is 0.

In the case when $j = i$, our assumptions (2.43) imply that $\mathbb{P}(C_{k,i} \cap B) = \mathbb{P}(C_{k+1,i} \cap B) = 0$, so we conclude that (2.51) is true. This completes the proof of (2.18).

It follows from (2.45)–(2.47) that (2.44) is satisfied.

The condition (2.20)–(2.21) follows from (2.48)–(2.51). □

Lemma 8. *For any \mathbf{S} there exists \mathbf{S}' such that (2.18)–(2.21) and (2.33)–(2.41) are satisfied, and $\mathbb{P}(C'_{i,j}) > 0$ for all $(i, j) \in \mathcal{D}'_- \cup \mathcal{D}'_+$.*

Proof. Let \mathcal{T}_1 denote the transformation defined in Lemma 6.

Lemma 7 can be obviously generalized to any situation when $C_{k,i} \subset B$, $\mathbb{P}(C_{k,i}) = 0$, and $C_{k,i}$ is adjacent to a cell in B^c . Fix any order, denoted \prec , for the set of pairs of cells. Let \mathcal{T}_2 denote the transformation defined in Lemma 7, applied to the first pair of cells (according to the order \prec) satisfying the generalized conditions described at the beginning of this paragraph. Recall that, by convention, if there is no such pair of cells then $\mathcal{T}_2(\mathbf{S}) = \mathbf{S}$.

Let $\mathbf{S}_0 = \mathbf{S}$ and $\mathbf{S}_k = \mathcal{T}_1(\mathcal{T}_2(\mathbf{S}_{k-1}))$ for $k \geq 1$. The transformations \mathcal{T}_1 and \mathcal{T}_2 decrease $m(\mathcal{G})$ or $m(\mathcal{H})$, unless they act as the identity transformation, by (2.42) and (2.44). Hence, for some k_1 and all $k \geq k_1$, $\mathbf{S}_k = \mathbf{S}_{k_1}$. We let $\mathbf{S}' = \mathbf{S}_{k_1}$ and note that $\mathcal{T}_1(\mathbf{S}') = \mathcal{T}_2(\mathbf{S}') = \mathbf{S}'$. This implies that \mathbf{S}' must satisfy the properties listed in Lemmas 6 and 7, i.e., (2.33)–(2.41), and the property that there is no cell $C_{k,i} \subset B$ that is adjacent to a cell in B^c and such that $\mathbb{P}(C_{k,i}) = 0$. A different way of saying this is that $\mathbb{P}(C'_{i,j}) > 0$ for all $(i, j) \in \mathcal{D}'_- \cup \mathcal{D}'_+$.

The conditions (2.18)–(2.21) hold for the transformation defined in this proof because they hold for \mathcal{T}_1 and \mathcal{T}_2 . □

Lemma 9. *Suppose that \mathbf{S} satisfies the conditions (2.33)–(2.41). There exists \mathbf{S}'' such that (2.18)–(2.21) and (2.33)–(2.41) hold, and*

$$\mathbb{P}(C''_{k,j}) > \mathbb{P}(C''_{k,j} \cap (A'')^c) = 0 \quad \text{or} \quad \mathbb{P}(C''_{k,j}) > \mathbb{P}(C''_{k,j} \cap A'') = 0, \tag{2.52}$$

for all $(k, j) \in \mathcal{D}''_- \cup \mathcal{D}''_+$.

We have used double primes in the statement of the lemma so that \mathbf{S}'' will not be confused with the \mathbf{S}' constructed in the first step of the proof.

Proof of Lemma 9. The main idea of the proof is to fix $(k, j) \in \mathcal{D}_- \cup \mathcal{D}_+$ and move as much as possible of A^c inside $C_{k,j}$ to A without destroying those properties of \mathbf{S} that need to be preserved. If this is not possible, a similar transformation is applied to A in place of A^c . Then the transformation is repeatedly applied to all $(k, j) \in \mathcal{D}_- \cup \mathcal{D}_+$.

First we apply the transformation defined in Lemma 8, if necessary, so that we can assume that (2.20) and (2.33)–(2.41) are satisfied for \mathbf{S} , and $\mathbb{P}(C_{i,j}) > 0$ for all $(i, j) \in \mathcal{D}_- \cup \mathcal{D}_+$.

Step 1. Consider any $(k, j) \in \mathcal{D}_- \cup \mathcal{D}_+$. If

$$\mathbb{P}(C_{k,j} \cap (A)^c) = 0 \quad \text{or} \quad \mathbb{P}(C_{k,j} \cap A) = 0, \tag{2.53}$$

then we let $S' = S$ and we let $\mathcal{T}_3(k, j)$ act as the identity transformation, i.e., $\mathcal{T}_3(k, j)(S) = S$. In this case at least one of the conditions in (2.65) (see below) holds.

Otherwise we proceed as follows. Assume that $(k, j) \in \mathcal{D}_-$ and

$$p_k \geq q_j. \tag{2.54}$$

Let $\alpha \geq 0$ be the largest real number such that the following three conditions are satisfied:

$$\frac{\alpha}{p_k} + x_k \leq x_{k+1} \wedge \delta, \tag{2.55}$$

$$\frac{\alpha}{q_j} + y_j \leq y_{j+1} \wedge 1, \tag{2.56}$$

$$\alpha \leq \mathbb{P}(C_{k,j} \cap A^c). \tag{2.57}$$

Here, if $k = m(\mathcal{G})$ then, by convention, $x_{k+1} = \infty$, and similarly, if $j = m(\mathcal{H})$ then $y_{j+1} = \infty$. At least one of the inequalities (2.55)–(2.57) is an equality.

We make \mathcal{F} finer, if necessary (see Lemma 4), and find an event $A_* \subset C_{k,j} \cap A^c$ such that $\mathbb{P}(A_*) = \alpha$. Then we let $A' = A \cup A_*$. The other elements of S remain unchanged. We have assumed that (2.53) does not hold, so

$$\mathbb{P}(C'_{k,j}) \geq \mathbb{P}(C'_{k,j} \cap A') \geq \mathbb{P}(C_{k,j} \cap A) > 0. \tag{2.58}$$

We have

$$x'_k = \frac{\alpha}{p_k} + x_k, \tag{2.59}$$

$$x'_i = x_i, \quad \text{for } i \neq k, \tag{2.60}$$

$$y'_j = \frac{\alpha}{q_j} + y_j, \tag{2.61}$$

$$y'_i = y_i, \quad \text{for } i \neq j. \tag{2.62}$$

Clearly, the condition (2.19) is satisfied. To prove (2.18) and (2.20)–(2.21), it will suffice to show that for all i and n ,

$$\text{if } |x_n - y_i| \geq 1 - \delta \quad \text{then} \quad |x'_n - y'_i| \geq 1 - \delta. \tag{2.63}$$

In view of (2.59)–(2.62), (2.63) holds if $n \neq k$ and $i \neq j$.

The conditions (2.41) and (2.35)–(2.37) and the assumption that $(k, j) \in \mathcal{D}_-$ imply that $|x_k - y_i| \geq 1 - \delta$ if and only if $i \geq j$. It follows from (2.54), (2.59), and (2.61)–(2.62) that if $i \geq j$, then

$$y'_i - x'_k \geq y'_j - x'_k = \frac{\alpha}{q_j} + y_j - \frac{\alpha}{p_k} - x_k \geq y_j - x_k \geq 1 - \delta.$$

This proves (2.63) for $n = k$ and all i .

The conditions (2.41) and (2.35)–(2.37) and the assumption that $(k, j) \in \mathcal{D}_-$ imply that $|x_n - y_j| \geq 1 - \delta$ if and only if $n \leq k$. It follows from (2.54), (2.55), (2.59)–(2.60), and (2.61) that if $n \leq k$, then

$$y'_j - x'_n \geq y'_j - x'_k = \frac{\alpha}{q_j} + y_j - \frac{\alpha}{p_k} - x_k \geq y_j - x_k \geq 1 - \delta.$$

This completes the proof of (2.63) for $i = j$ and all n . Hence, (2.18) and (2.20)–(2.21) hold true.

Recall that at least one of the inequalities (2.55)–(2.57) is an equality. We list all possible cases below:

- If $x'_k = \alpha/p_k + x_k = \delta$ then we must have $y'_j = 1$ in view of (2.63). This implies that $\mathbb{P}(C'_{k,j} \cap (A')^c) = 0$.
- If $x'_k = \alpha/p_k + x_k = x_{k+1} < \delta$ then $x'_k = x'_{k+1}$.
- If $y'_j = \alpha/q_j + y_j = 1$ then $\mathbb{P}(C'_{k,j} \cap (A')^c) = 0$.
- If $y'_j = \alpha/q_j + y_j = y_{j+1} < 1$ then $y'_j = y'_{j+1}$.
- If $\alpha = \mathbb{P}(C_{k,j} \cap A^c)$ then $\mathbb{P}(C'_{k,j} \cap (A')^c) = 0$.

These observations and (2.58) allow us to conclude that $x'_k = x'_{k+1}$ or $y'_j = y'_{j+1}$ or $\mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap (A')^c) = 0$.

A completely analogous argument shows that if (2.54) does not hold then we can transform \mathbf{S} into \mathbf{S}' such that (2.18)–(2.21) are true, and we have the alternatives $x'_k = x'_{k-1}$ or $y'_j = y'_{j-1}$ or $\mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap A') = 0$.

We next replace the assumption that $(k, j) \in \mathcal{D}_-$ with $(k, j) \in \mathcal{D}_+$. Once again, we can apply the analogous argument to conclude that we can construct \mathbf{S}' satisfying (2.18)–(2.21) and such that $x'_k = x'_{k+1}$ or $y'_j = y'_{j+1}$ or $\mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap (A')^c) = 0$ or $x'_k = x'_{k-1}$ or $y'_j = y'_{j-1}$ or $\mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap A') = 0$.

We see that if $(k, j) \in \mathcal{D}_- \cup \mathcal{D}_+$, then

$$x'_k = x'_{k+1} \quad \text{or} \quad y'_j = y'_{j+1} \quad \text{or} \quad x'_k = x'_{k-1} \quad \text{or} \quad y'_j = y'_{j-1}, \tag{2.64}$$

or

$$\mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap (A')^c) = 0 \quad \text{or} \quad \mathbb{P}(C'_{k,j}) > \mathbb{P}(C'_{k,j} \cap A') = 0. \tag{2.65}$$

Let $\mathcal{T}_3(k, j)$ be the transformation defined in this step.

Step 2. Let \prec denote an arbitrary order for the set of pairs (k, j) . Let \mathcal{T}_3 be defined as $\mathcal{T}_3(k, j)$ applied to the first pair (k, j) (in the sense of the order \prec) such that (2.53) does not hold.

Let \mathcal{T}_4 denote the transformation defined in Lemma 6.

Let $\mathbf{S}_0 = \mathbf{S}$ and $\mathbf{S}_k = \mathcal{T}_3(\mathcal{T}_4(\mathbf{S}_{k-1}))$ for $k \geq 1$. The transformation \mathcal{T}_4 strictly decreases $m(\mathcal{G})$ or $m(\mathcal{H})$, unless it acts as the identity transformation. Hence, for some n_1 and all $k \geq n_1$, $\mathbf{S}_k = \mathcal{T}_3(\mathbf{S}_{k-1})$. This implies that for $k \geq n_1$, \mathbf{S}_k satisfies (2.41), and this in turn implies that (2.64) cannot be true for \mathbf{S}_k .

Since none of the conditions in (2.64) can hold, one of the conditions in (2.65) must be true for \mathbf{S}_k , $k \geq n_1$. Note that $\mathcal{T}_3(k, j)$ does not change $C_{n,i} \cap A$ or $C_{n,i} \cap A^c$ unless $(n, i) = (k, j)$. Hence, for some $n_2 \geq n_1$ and all $k \geq n_2$, $\mathbf{S}_k = \mathcal{T}_3(\mathbf{S}_{n_2})$. We let $\mathbf{S}'' = \mathbf{S}_{n_2}$.

The conditions (2.18)–(2.21) hold for the transformation defined in this proof because they hold for \mathcal{T}_3 and \mathcal{T}_4 . □

Consider four cells that lie at the corners of a rectangle, i.e., the cells C_{k_1,j_1} , C_{k_2,j_2} , C_{k_1,j_2} , and C_{k_2,j_1} for some j_1, j_2, k_1 , and k_2 . The transformation defined in the next lemma moves as much as possible of A from the upper left corner to the lower left corner, and compensates by moving the same amount of A from the lower right corner to the upper right corner. The result of the transformation is that one of the four cells will not hold any A .

Lemma 10. *Suppose that for some \mathbf{S} , $1 \leq k_1 < k_2 \leq m(\mathcal{G})$, and $1 \leq j_2 < j_1 \leq m(\mathcal{H})$, we have*

$$p := \min(\mathbb{P}(A \cap C_{k_1,j_1}), \mathbb{P}(A \cap C_{k_2,j_2})) > 0. \tag{2.66}$$

Assume that

$$C_{k_1,j_1} \cup C_{k_2,j_2} \cup C_{k_1,j_2} \cup C_{k_2,j_1} \subset B, \quad \text{or} \tag{2.67}$$

$$C_{k_1,j_1} \cup C_{k_1,j_2} \subset B \quad \text{and} \quad C_{k_2,j_1} \cup C_{k_2,j_2} \subset B^c, \quad \text{or} \tag{2.68}$$

$$C_{k_1,j_1} \cup C_{k_1,j_2} \subset B^c \quad \text{and} \quad C_{k_2,j_1} \cup C_{k_2,j_2} \subset B, \quad \text{or} \tag{2.69}$$

$$C_{k_1,j_1} \cup C_{k_2,j_1} \subset B \quad \text{and} \quad C_{k_1,j_2} \cup C_{k_2,j_2} \subset B^c, \quad \text{or} \tag{2.70}$$

$$C_{k_1,j_1} \cup C_{k_2,j_1} \subset B^c \quad \text{and} \quad C_{k_1,j_2} \cup C_{k_2,j_2} \subset B. \tag{2.71}$$

Then there exists \mathbf{S}' such that $m'(\mathcal{G}') = m(\mathcal{G})$, $m'(\mathcal{H}') = m(\mathcal{H})$,

$$\mathbb{P}(C'_{k_1,j_1}) = \mathbb{P}(C_{k_1,j_1}) - p, \tag{2.72}$$

$$\mathbb{P}(C'_{k_2,j_2}) = \mathbb{P}(C_{k_2,j_2}) - p, \tag{2.73}$$

$$\mathbb{P}(C'_{k_1,j_2}) = \mathbb{P}(C_{k_1,j_2}) + p, \tag{2.74}$$

$$\mathbb{P}(C'_{k_2,j_1}) = \mathbb{P}(C_{k_2,j_1}) + p, \tag{2.75}$$

and $\mathbb{P}(C'_{k,j}) = \mathbb{P}(C_{k,j})$ for all other values of (k, j) .

Moreover, (2.20)–(2.21) hold. The condition (2.18) holds with equality.

Proof. We can make \mathcal{F} finer, if necessary (see Lemma 4), so that there exist sets $A_1 \subset A \cap C_{k_1,j_1}$ and $A_2 \subset A \cap C_{k_2,j_2}$ such that $\mathbb{P}(A_1) = \mathbb{P}(A_2) = p$. Let

$$\begin{aligned} C'_{k_1,j_1} &= C_{k_1,j_1} \setminus A_1, & C'_{k_2,j_2} &= C_{k_2,j_2} \setminus A_2, \\ C'_{k_1,j_2} &= C_{k_1,j_2} \cup A_1, & C'_{k_2,j_1} &= C_{k_2,j_1} \cup A_2. \end{aligned}$$

For all other values of (k, j) , we let $C'_{k,j} = C_{k,j}$. These transformations redefine the G_k , the H_j , \mathcal{G} , and \mathcal{H} . The other elements of \mathbf{S} are unaffected.

It is clear that (2.72)–(2.75) are satisfied and $\mathbb{P}(C'_{k,j}) = \mathbb{P}(C_{k,j})$ for all other values of (k, j) .

It is easy to check that $p'_k = p_k$, $q'_k = q_k$, $x'_k = x_k$, and $y'_k = y_k$ for all k . Therefore, $C'_{k,j} \subset B'$ if and only if $C_{k,j} \subset B$, for every (k, j) .

Suppose that (2.68) is true. Then

$$\begin{aligned} & \mathbb{P}(C'_{k_1,j_1} \cap B') + \mathbb{P}(C'_{k_1,j_2} \cap B') \\ &= \mathbb{P}(C'_{k_1,j_1}) + \mathbb{P}(C'_{k_1,j_2}) = \mathbb{P}(C_{k_1,j_1}) - p + \mathbb{P}(C_{k_1,j_2}) + p \\ &= \mathbb{P}(C_{k_1,j_1}) + \mathbb{P}(C_{k_1,j_2}) = \mathbb{P}(C_{k_1,j_1} \cap B) + \mathbb{P}(C_{k_1,j_2} \cap B), \end{aligned} \tag{2.76}$$

and

$$\mathbb{P}(C'_{k_2,j_1} \cap B') + \mathbb{P}(C'_{k_2,j_2} \cap B') = 0 = \mathbb{P}(C_{k_2,j_1} \cap B) + \mathbb{P}(C_{k_2,j_2} \cap B).$$

These formulas and the fact that $\mathbb{P}(C'_{k,j}) = \mathbb{P}(C_{k,j})$ for all other values of (k, j) imply that (2.18) holds with equality.

It follows from (2.68) that either $j_1, j_2 \leq m_-(\mathcal{H})$ or $j_1, j_2 \geq m_+(\mathcal{H})$. This and (2.76) imply that (2.20)–(2.21) hold.

One can prove that (2.20)–(2.21) hold and (2.18) holds with equality under any of the assumptions (2.67)–(2.71) in a similar manner.

The condition (2.19) obviously holds. □

Notation 3. We will use $\mathcal{T}_{\not\rightarrow}((k_1, j_1), (k_2, j_2))$ to denote the transformation defined in Lemma 10. We define a transformation $\mathcal{T}_{\not\rightarrow}^c((k_1, j_1), (k_2, j_2))$ by replacing A with A^c in the assumption (2.66) and the construction of $\mathcal{T}_{\not\rightarrow}((k_1, j_1), (k_2, j_2))$.

If the assumptions of Lemma 10 do not hold then $\mathcal{T}_{\not\rightarrow}((k_1, j_1), (k_2, j_2))$ will denote the identity transformation, and similarly for $\mathcal{T}_{\not\rightarrow}^c((k_1, j_1), (k_2, j_2))$.

The transformation defined in the next lemma converts the top right family of cells in Figure 4 into subsets of A . It also converts the bottom left family of cells in Figure 4 into subsets of A^c .

Lemma 11. *Given \mathbf{S} , let*

$$A' = \left(A \cup \bigcup_{m_-(\mathcal{G}) < k \leq m(\mathcal{G}), m_-(\mathcal{H}) < j \leq m(\mathcal{H})} C_{k,j} \right) \setminus \left(\bigcup_{1 \leq k < m_+(\mathcal{G}), 1 \leq j < m_+(\mathcal{H})} C_{k,j} \right). \tag{2.77}$$

Let all other elements of \mathbf{S}' be the same as those of \mathbf{S} . Then the conditions (2.18)–(2.21) hold.

Proof. The condition (2.19) clearly holds.

Note that

$$\begin{aligned} x'_k &\leq x_k, & \text{for } 1 \leq k \leq m_-(\mathcal{G}), \\ y'_k &\leq y_k, & \text{for } 1 \leq k \leq m_-(\mathcal{H}), \\ x'_k &\geq x_k, & \text{for } m_+(\mathcal{G}) \leq k \leq m(\mathcal{G}), \\ y'_k &\geq y_k, & \text{for } m_+(\mathcal{H}) \leq k \leq m(\mathcal{H}). \end{aligned}$$

These observations and (2.13)–(2.16) imply that $B \subset B'$, and therefore (2.18) and (2.20)–(2.21) hold. □

The transformation defined in the next lemma empties all cells in the ‘bottom right corner of the upper left corner’ (marked in green in Figure 4), and similarly on the other side of the

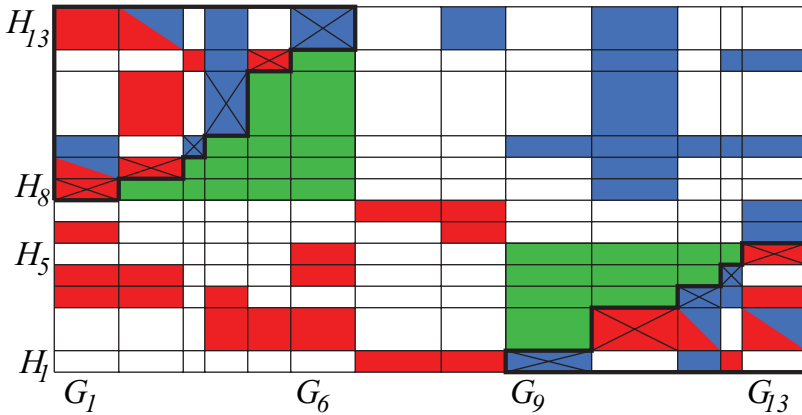


FIGURE 4: (Color coded) For the meanings of the white, red, and blue colors and the thick lines, see the caption of Figure 2. Green has the same meaning as white: it represents ‘empty’ cells, i.e., cells $C_{k,j}$ such that $\mathbb{P}(C_{k,j}) = 0$. The green cells are the cells that were ‘emptied’ in Step 1 of the proof of Lemma 12. Every crossed cell belongs either to \mathcal{D}_- or to \mathcal{D}_+ and must satisfy $\mathbb{P}(C_{k,j}) > 0$ and either $\mathbb{P}(C_{k,j} \cap A) = 0$ or $\mathbb{P}(C_{k,j} \cap A^c) = 0$. All cells below \mathcal{D}_- are in A or empty. All cells above \mathcal{D}_+ are in A or empty. All cells to the right of \mathcal{D}_- are in A or empty. All cells to the left of \mathcal{D}_+ are in A^c or empty. If a cell is below a cell in \mathcal{D}_- and to the right of a (different) cell in \mathcal{D}_- , then it is empty. If a cell is above a cell in \mathcal{D}_+ and to the left of a (different) cell in \mathcal{D}_+ , then it is empty.

diagonal. The result of the transformations described in Lemmas 11–12 is that large regions in Figure 4 do not have cells that would hold both A and A^c .

Lemma 12. Consider an \mathbf{S} such that (2.20) holds. Then there exists \mathbf{S}' such that the conditions (2.18)–(2.21), (2.33)–(2.41), and (2.52) hold, and

$$\mathbb{P}((G'_k \setminus B') \cap A') = 0 \quad \text{for} \quad 1 \leq k \leq m_-(\mathcal{G}'), \tag{2.78}$$

$$\mathbb{P}((H'_k \setminus B') \cap A') = 0 \quad \text{for} \quad 1 \leq k \leq m_-(\mathcal{H}'), \tag{2.79}$$

$$\mathbb{P}((G'_k \setminus B') \cap (A')^c) = 0 \quad \text{for} \quad m_+(\mathcal{G}) \leq k \leq m(\mathcal{G}), \tag{2.80}$$

$$\mathbb{P}((H'_k \setminus B') \cap (A')^c) = 0 \quad \text{for} \quad m_+(\mathcal{H}) \leq k \leq m(\mathcal{H}). \tag{2.81}$$

Proof. Step 1. It follows from (2.20) that $m_-(\mathcal{G}) \geq 1$, $m_+(\mathcal{H}) \leq m(\mathcal{H})$, $m_-(\mathcal{H}) \geq 1$, and $m_+(\mathcal{G}) \leq m(\mathcal{G})$. Consider k and j such that $k \leq m_-(\mathcal{G})$, $j \geq m_+(\mathcal{H})$, and $C_{k,j} \subset B^c$. Let $C'_{k,j} = \emptyset$.

If $\mathbb{P}(A \mid C_{k,j}) < 1 - \delta$ then let $A' = A \setminus C_{k,j}$ and $C'_{k,1} = C_{k,1} \cup C_{k,j}$.

If $\mathbb{P}(A \mid C_{k,j}) \geq 1 - \delta$ then let $A' = A \cup C_{k,j}$ and $C'_{m(\mathcal{G}),j} = C_{m(\mathcal{G}),j} \cup C_{k,j}$.

These changes will affect \mathcal{G} and \mathcal{H} . Other elements of \mathbf{S} will be unchanged.

It is easy to check that $x'_k \leq x_k$, $x'_{m(\mathcal{G})} \geq x_{m(\mathcal{G})}$, $y'_1 \leq y_1$, and $y'_j \geq y_j$. All other x_i and y_i will be unaffected. For all i and n such that $C_{i,n} \subset B$, $\mathbb{P}(C'_{i,n}) = \mathbb{P}(C_{i,n})$. Hence, in view of (2.13)–(2.16), we see that (2.18)–(2.21) hold. We also have

$$A \cap \bigcup_{m_-(\mathcal{G}) < k \leq m(\mathcal{G}), m_-(\mathcal{H}) < j \leq m(\mathcal{H})} C_{k,j} \subset A' \cap \bigcup_{m'_-(\mathcal{G}') < k \leq m'(\mathcal{G}'), m'_-(\mathcal{H}') < j \leq m'(\mathcal{H}')} C'_{k,j}, \tag{2.82}$$

$$A' \cap \bigcup_{1 \leq k < m'_+(\mathcal{G}'), 1 \leq j < m'_+(\mathcal{H}')} C'_{k,j} \subset A \cap \bigcup_{1 \leq k < m_+(\mathcal{G}), 1 \leq j < m_+(\mathcal{H})} C_{k,j}. \tag{2.83}$$

We repeat the transformation for all k and j such that $k \leq m_-(\mathcal{G}), j \geq m_+(\mathcal{H})$, and $C_{k,j} \subset B^c$. The result is that $C'_{k,j} = \emptyset$ for all such pairs (k, j) . Moreover, (2.18)–(2.21) and (2.82)–(2.83) hold.

We apply an analogous sequence of transformations corresponding to all pairs (k, j) such that $k \geq m_+(\mathcal{G}), j \leq m_-(\mathcal{H})$, and $C_{k,j} \subset B^c$. The result is, once again, that $C'_{k,j} = \emptyset$ for all such pairs (k, j) . The conditions (2.18)–(2.21) and (2.82)–(2.83) still hold. Note that A and A' do not exchange roles in (2.82)–(2.83); i.e., these conditions hold as stated.

We will denote the composition of all transformations defined in this step by \mathcal{T}_5 .

For later reference, we record the following properties of $\mathbf{S}' = \mathcal{T}_5(\mathbf{S})$:

$$C'_{k,j} = \emptyset \quad \text{if } k \leq m_-(\mathcal{G}), j \geq m_+(\mathcal{H}), \text{ and } C'_{k,j} \subset B^c; \tag{2.84}$$

$$C'_{k,j} = \emptyset \quad \text{if } k \geq m_+(\mathcal{G}), j \leq m_-(\mathcal{H}), \text{ and } C'_{k,j} \subset B^c. \tag{2.85}$$

Step 2. Let $\mathcal{T}_6, \mathcal{T}_7$, and \mathcal{T}_8 denote the transformations defined in Lemmas 11, 9, and 6. Let $\mathbf{S}_0 = \mathbf{S}$ and $\mathbf{S}_k = \mathcal{T}_5(\mathcal{T}_6(\mathcal{T}_7(\mathcal{T}_8(\mathbf{S}_{k-1}))))$ for $k \geq 1$. In view of (2.42) and the fact that $\mathcal{T}_5, \mathcal{T}_6$, and \mathcal{T}_7 have the property (2.19), \mathcal{T}_8 must act as the identity eventually; i.e., there must exist n_1 such that $\mathbf{S}_k = \mathcal{T}_5(\mathcal{T}_6(\mathcal{T}_7(\mathbf{S}_{k-1})))$ for $k \geq n_1$.

The transformations \mathcal{T}_5 and \mathcal{T}_6 do not change $m(\mathcal{G})$ and $m(\mathcal{H})$. They also do not change the values of $\mathbb{P}(C_{k,j})$ and $\mathbb{P}(C_{j,k} \cap A)$ for $(k, j) \in \mathcal{D}_- \cap \mathcal{D}_+$. It follows that \mathcal{T}_7 must act as the identity eventually; i.e., there must exist $n_2 \geq n_1$ such that $\mathbf{S}_k = \mathcal{T}_5(\mathcal{T}_6(\mathbf{S}_{k-1}))$ for $k \geq n_2$.

It follows from (2.77) and (2.82)–(2.83) that there must exist $n_3 \geq n_2$ such that $\mathbf{S}_k = \mathbf{S}_{k-1}$ for $k \geq n_3$. Let $\mathbf{S}' = \mathbf{S}_{n_3}$. Since $\mathcal{T}_i(\mathbf{S}') = \mathbf{S}'$ for $i = 5, 6, 7, 8$, the conditions (2.18)–(2.21), (2.33)–(2.41), (2.52), and (2.78)–(2.81) are satisfied. \square

Lemma 13. *If $\mathbb{P}(B) > 0$ and $\varepsilon > 0$, then there exists \mathbf{S}' satisfying*

- (i) $\mathbb{P}(B') > \mathbb{P}(B) - \varepsilon$,
- (ii) $m_-(\mathcal{G}) = 1$, or $m_-(\mathcal{G}) = 2$ and $\mathbb{P}(C_{1,m(\mathcal{H})}) = 0$, and
- (iii) $m_-(\mathcal{H}) = 1$, or $m_-(\mathcal{H}) = 2$ and $\mathbb{P}(C_{m(\mathcal{G}),1}) = 0$.

Proof. Suppose that $\mathbb{P}(B) > 0$ and $\varepsilon > 0$. We apply the transformation defined in Lemma 3 and obtain \mathbf{S}' satisfying $\mathbb{P}(B') > \mathbb{P}(B) - \varepsilon, m'(\mathcal{G}') \leq m(\mathcal{G}) + 1, m'(\mathcal{H}') \leq m(\mathcal{H}) + 1$, and (2.21).

Next we apply the transformation defined in Lemma 12 and obtain \mathbf{S}'' satisfying (2.18), (2.21), (2.33)–(2.41), (2.52), (2.78)–(2.81), $m''(\mathcal{G}'') \leq m(\mathcal{G}) + 1$, and $m''(\mathcal{H}'') \leq m(\mathcal{H}) + 1$. These properties of \mathbf{S}'' are illustrated in Figure 4.

The logical scheme of the remaining part of the proof is represented by the flowchart in Figure 5.

The leaves of our branching argument will be denoted by \textcircled{X} . These are places where logical contradictions are reached so another branch of the proof must be considered.

We will now give names to the transformations that we will apply in this proof.

Let \mathcal{T}_9 be the transformation defined in Lemma 12.

Recall that $\mathcal{T}_{\nearrow}((k_1, j_1), (k_2, j_2))$ denotes the transformation defined in Lemma 10. The transformation $\mathcal{T}_{\nearrow}^c((k_1, j_1), (k_2, j_2))$ was defined in an analogous way, by replacing A with A^c in the assumption (2.66) and the construction of $\mathcal{T}_{\nearrow}((k_1, j_1), (k_2, j_2))$.

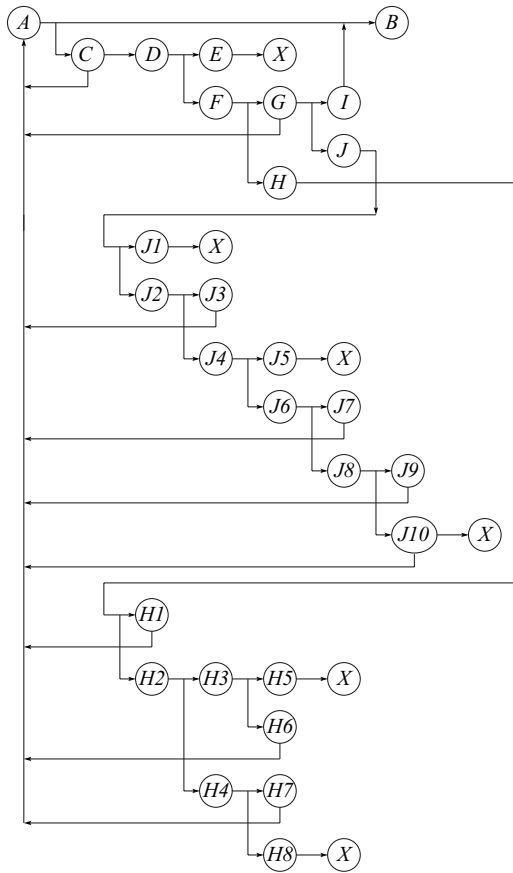


FIGURE 5: The logical structure of the proof of Lemma 13.

Recall that $\mathcal{T}_3(k, j)$ is the transformation defined in Step 1 of the proof of Lemma 9. If $\mathbf{S}' = \mathcal{T}_3(k, j)(\mathbf{S})$ then at least one of the conditions (2.64)–(2.65) holds.

The root of the flowchart in Figure 5 is (A). Our argument will require that we jump to (A) repeatedly. We will argue in the main body of the proof that every jump to (A) is associated with the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$. Therefore, there can be only a finite number of jumps to (A) from later parts of the proof.

We will repeatedly apply the transformations named above, but instead of using the notation with primes, e.g. $C'_{k,j}$, for the resulting objects, we will always write $C_{k,j}$, etc., without primes, to simplify the notation. We hope that this convention will not be confusing in this proof.

Suppose that \mathbf{S} is given. In view of the initial part of the proof, we can assume that it has all the properties of \mathbf{S}'' , i.e., it satisfies (2.18), (2.21), (2.33)–(2.41), (2.52), and (2.78)–(2.81).

(A) Apply \mathcal{T}_9 to \mathbf{S} .

If $m_-(\mathcal{G}) \leq 1$ then we jump to (B).

(C) Suppose that $m_-(\mathcal{G}) > 1$. We apply the transformations $\mathcal{T}_{\nearrow}^c((k, m(\mathcal{H})), (m_-(\mathcal{G}), j))$ repeatedly for all $k < m_-(\mathcal{G})$ and $j < m_+(\mathcal{H})$. The resulting \mathbf{S} satisfies (2.18) because the following case of (2.68)–(2.71) is satisfied for $k < m_-(\mathcal{G})$ and $j < m_+(\mathcal{H})$:

$$C_{k,m(\mathcal{H})} \cup C_{m_-(\mathcal{G}),m(\mathcal{H})} \subset B \quad \text{and} \quad C_{k,j} \cup C_{m_-(\mathcal{G}),j} \subset B^c.$$

In later applications of \mathcal{T}_{\nearrow} or \mathcal{T}_{\nearrow}^c , we will leave it to the reader to check that one of the conditions (2.67)–(2.71) is satisfied.

In view of (2.84)–(2.85), the result of these transformations has the property that

$$\mathbb{P}\left(A^c \cap \bigcup_{k < m_-(\mathcal{G})} C_{k,m(\mathcal{H})}\right) = 0 \quad \text{or} \tag{2.86}$$

$$\mathbb{P}\left(A^c \cap \bigcup_{j < m(\mathcal{H})} C_{m_-(\mathcal{G}),j}\right) = 0. \tag{2.87}$$

We apply the transformation $\mathcal{T}_3(m_-(\mathcal{G}), m(\mathcal{H}))$. Note that it changes the ‘A-contents’ of only $C_{m_-(\mathcal{G}),m(\mathcal{H})}$. The resulting \mathbf{S} satisfies (2.64) or (2.65) with $(k, j) = (m_-(\mathcal{G}), m(\mathcal{H}))$.

If \mathbf{S} satisfies (2.64) then we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(D) Assume that (2.65) holds, i.e., that we have either $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A) = 0$ or $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A^c) = 0$.

(E) Assume that (2.87) holds and recall (2.78)–(2.81). These formulas imply that $\mathbb{P}(\bigcup_{j < m(\mathcal{H})} C_{m_-(\mathcal{G}),j}) = 0$. Since $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})}) > 0$ and either $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A) = 0$ or $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A^c) = 0$, we must have $x_{m_-(\mathcal{G})} = 0$ or $x_{m_-(\mathcal{G})} = 1$. If $x_{m_-(\mathcal{G})} = 0$ then this contradicts the facts that $m_-(\mathcal{G}) > 1$ and $x_{m_-(\mathcal{G})} > x_1 \geq 0$. We cannot have $x_{m_-(\mathcal{G})} = 1$ because that would contradict (2.13). We conclude that (2.87) cannot hold. (X)

(F) Assume that (2.86) is true. The transformations which generated (2.86)–(2.87) did not change the x_i and y_i , and they also did not change the fact that $\mathbb{P}(C_{m_-(\mathcal{G}),i} \cap A) = 0$ for all $i < m(\mathcal{H})$. We are in the current branch of the proof because we have not jumped to (A) after applying $\mathcal{T}_3(m_-(\mathcal{G}), m(\mathcal{H}))$. Hence, $x_{m_-(\mathcal{G})} > x_{m_-(\mathcal{G})-1}$. If $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A) = 0$ then $x_{m_-(\mathcal{G})} = 0$, but this contradicts the facts that $m_-(\mathcal{G}) > 1$ and $x_{m_-(\mathcal{G})} > x_1 \geq 0$. Hence, we must have $\mathbb{P}(C_{m_-(\mathcal{G}),m(\mathcal{H})} \cap A^c) = 0$. This, (2.78)–(2.81), and (2.86) imply that

$$\mathbb{P}(A^c \cap H_{m(\mathcal{H})}) = \mathbb{P}\left(A^c \cap \bigcup_{1 \leq k \leq m(\mathcal{G})} C_{k,m(\mathcal{H})}\right) = 0. \tag{2.88}$$

At this point our argument branches as follows. We will consider the case $m_-(\mathcal{G}) = 2$ in (G) and (I), the case $m_-(\mathcal{G}) = 3$ in (G) and (J), and the case $m_-(\mathcal{G}) \geq 4$ in (H).

(G) If $m_-(\mathcal{G}) = 2$ or $m_-(\mathcal{G}) = 3$, then we interchange the roles of A and A^c and those of \mathcal{G} and \mathcal{H} , and argue as follows. We apply the transformations $\mathcal{T}_{\nearrow}((k, m_+(\mathcal{H})), (1, j))$ repeatedly for all $k > m_-(\mathcal{G})$ and $j > m_+(\mathcal{H})$. Then we apply the transformation $\mathcal{T}_3(1, m_+(\mathcal{H}))$.

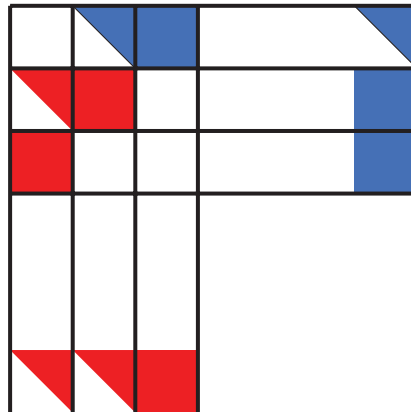


FIGURE 6: See the body of the article for the description.

If \mathbf{S} satisfies (2.64) then we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

Otherwise we will reach the conclusion analogous to (2.88), namely that

$$\mathbb{P}(A \cap G_1) = 0. \tag{2.89}$$

We will now argue that the transformations described between (2.88) and (2.89) do not affect the validity of (2.88), assuming that there was no jump to (A). The transformations $\mathcal{T}_{\nearrow}((k, m_+(\mathcal{H})), (1, j))$ do not change any of the $x_i, y_i, p_i,$ or q_i . The transformation $\mathcal{T}_3(1, m_+(\mathcal{H}))$ does not affect any cells in $H_{m(\mathcal{H})}$ because $m_+(\mathcal{H}) < m(\mathcal{H})$. Hence, (2.88) remains true.

(I) Suppose that $m_-(\mathcal{G}) = 2$ and use (2.88)–(2.89) to conclude that $\mathbb{P}(C_{1,m(\mathcal{H})}) = 0$. We now jump to (B).

(J) Next assume that $m_-(\mathcal{G}) = 3$. This case is rather complicated, so it is the only part of the proof whose steps are illustrated one by one in Figures 6–12. We will now explain how different events are represented in these figures. The three columns represent $G_1, G_2,$ and G_3 . The three rows represent $H_{m_+(\mathcal{H})} = H_{m(\mathcal{H})-2}, H_{m(\mathcal{H})-1},$ and $H_{m(\mathcal{H})}$. The colors have the same meaning as in Figure 4. Cells containing both white and some other color may be empty or contain either A or A^c , depending on the color. The colored areas at the bottom represent the family of cells $C_{k,j}$ (not individual cells) with $1 \leq k \leq 3$ and $j < m_+(\mathcal{H})$, and the colored areas on the right represent the family of cells $C_{k,j}$ with $k > m_-(\mathcal{G})$ and $H_{m_+(\mathcal{H})} \leq j \leq H_{m(\mathcal{H})}$.

The starting point is illustrated in Figure 6. This \mathbf{S} satisfies (2.88)–(2.89). We have either

$$\mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A) = 0 \tag{2.90}$$

or $\mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A^c) = 0$. We will discuss only the case in (2.90), depicted in Figure 6. The other case can be dealt with in an analogous way.

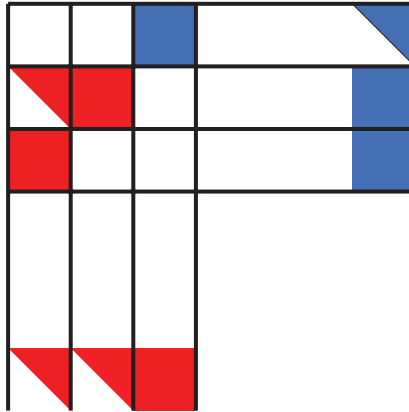


FIGURE 7: See the body of the article for the description.

We apply the transformations $\mathcal{T}_{\neq}((2, m(\mathcal{H})), (k, m(\mathcal{H}) - 1))$ repeatedly for all $k > m_-(\mathcal{G})$. The result of these transformations has the property that

$$\mathbb{P} \left(A \cap \bigcup_{k > m_-(\mathcal{G})} C_{k, m(\mathcal{H})-1} \right) = 0 \quad \text{or} \quad (2.91)$$

$$\mathbb{P} (A \cap C_{2, m(\mathcal{H})}) = 0. \quad (2.92)$$

(J1) If (2.91) holds then $\mathbb{P}(H_{m(\mathcal{H})-1} \cap A) = 0$ because of (2.78)–(2.81), (2.89), and (2.90). But then $y_{m(\mathcal{H})-1} = 0$. This is impossible because $m(\mathcal{H}) - 1 > m_+(\mathcal{H})$. (X)

(J2) Assume that (2.92) is true. We apply $\mathcal{T}_3(2, m(\mathcal{H}) - 1)$. The resulting \mathbf{S} satisfies (2.64) or (2.65) with $(k, j) = (2, m(\mathcal{H}) - 1)$.

(J3) If \mathbf{S} satisfies (2.64) then we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(J4) Assume that (2.65) holds, i.e., either $\mathbb{P}(C_{2, m(\mathcal{H})-1} \cap A) = 0$ or $\mathbb{P}(C_{2, m(\mathcal{H})-1} \cap A^c) = 0$ (see Figures 7–8).

(J5) Assume that $\mathbb{P}(C_{2, m(\mathcal{H})-1} \cap A) = 0$ (see Figure 7). This assumption, combined with (2.78)–(2.81) and (2.92), implies that $x_2 = \mathbb{P}(G_2 \cap A) = 0$. This is impossible because $x_2 > x_1 \geq 0$. (X)

(J6) Hence, we must have $\mathbb{P}(C_{2, m(\mathcal{H})-1} \cap A^c) = 0$ (see Figure 8). We apply the transformations $\mathcal{T}_{\neq}^c((1, m_+(\mathcal{H}) + 1), (2, j))$ repeatedly for all $j < m_+(\mathcal{H})$. The result of these transformations has the property that

$$\mathbb{P} (A^c \cap C_{1, m_+(\mathcal{H})+1}) = 0 \quad \text{or} \quad (2.93)$$

$$\mathbb{P} \left(A^c \cap \bigcup_{j < m_+(\mathcal{H})} C_{2, j} \right) = 0. \quad (2.94)$$

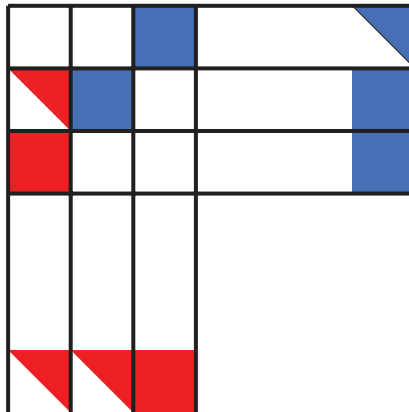


FIGURE 8: See the body of the article for the description.

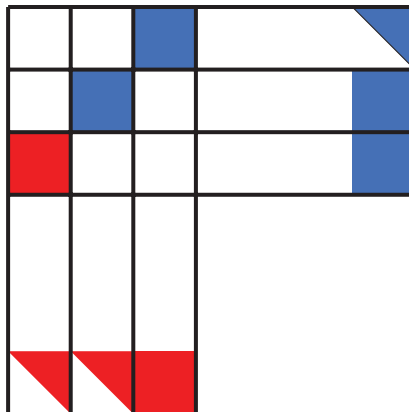


FIGURE 9: See the body of the article for the description.

We apply the transformation $\mathcal{T}_3(2, m(\mathcal{H}) - 1)$. The resulting \mathbf{S} satisfies (2.64) or (2.65) with $(k, j) = (2, m(\mathcal{H}) - 1)$.

(J7) If \mathbf{S} satisfies (2.64) then we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(J8) Assume that (2.65) holds (see Figures 9–10); that is,

$$\mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A) = 0 \quad \text{or} \quad \mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A^c) = 0. \tag{2.95}$$

(J9) Suppose (2.93) holds.

If $\mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A^c) = 0$, then $\mathbb{P}(H_{m(\mathcal{H})-1} \cap A^c) = 0$, because of (2.78)–(2.81) (see Figure 9). Then $y_{m(\mathcal{H})-1} = 1 = y_{m(\mathcal{H})}$ and we jump to (A). An application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

If (2.93) holds and $\mathbb{P}(C_{2,m(\mathcal{H})-1} \cap A) = 0$, then $\mathbb{P}(G_2 \cap A) = 0$, because of (2.78)–(2.81) and (2.92) (see Figure 10). Then $x_2 = 0 = x_1$ and we jump to (A). An application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

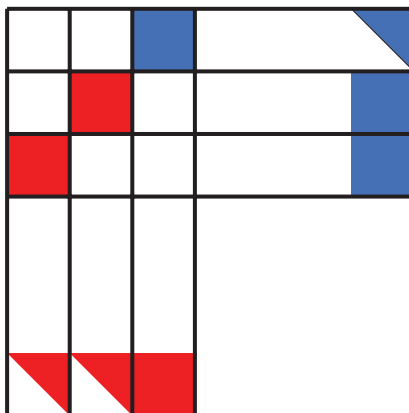


FIGURE 10: See the body of the article for the description.

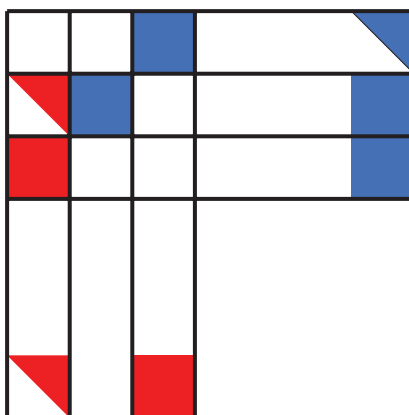


FIGURE 11: See the body of the article for the description.

(J10) Next suppose that (2.94) holds; see Figures 11–12. In view of (2.84)–(2.85), (2.92), and (2.95), either $x_2 = 1$ or $x_2 = 0$. The first case is impossible because $2 < m_-(\mathcal{G})$. (X)

In the second case we have $x_2 = 0 = x_1$ and we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(H) We now assume that $m_-(\mathcal{G}) \geq 4$. Recall (2.88). We will analyze $H_{m(\mathcal{H})-1} \cap B$. We apply the transformations $\mathcal{T}_{\downarrow}^c((k, m(\mathcal{H}) - 1), (m_-(\mathcal{G}) - 1, j))$ repeatedly for all $k < m_-(\mathcal{G}) - 1$ and $j < m_+(\mathcal{H})$. The resulting \mathbf{S} has the property that

$$\mathbb{P}\left(A^c \cap \bigcup_{j < m_+(\mathcal{H})} C_{m_-(\mathcal{G})-1, j}\right) = 0 \quad \text{or} \tag{2.96}$$

$$\mathbb{P}\left(A^c \cap \bigcup_{k < m_-(\mathcal{G})-1} C_{k, m(\mathcal{H})-1}\right) = 0. \tag{2.97}$$

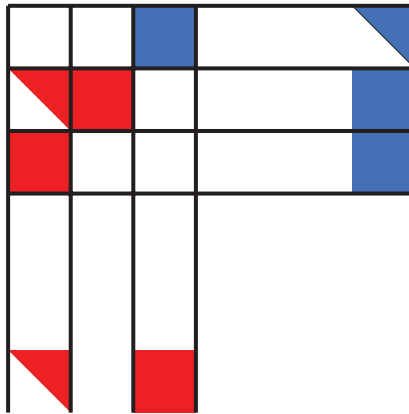


FIGURE 12: See the body of the article for the description.

We apply the transformation $\mathcal{T}_3(m_-(\mathcal{G}) - 1, m(\mathcal{H}) - 1)$. The resulting \mathbf{S} satisfies (2.64) or (2.65) with $(k, j) = (m_-(\mathcal{G}) - 1, m(\mathcal{H}) - 1)$.

(H1) If \mathbf{S} satisfies (2.64) then we jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(H2) Assume that (2.65) holds, i.e., either $\mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})-1} \cap A) = 0$ or $\mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})-1} \cap A^c) = 0$.

(H3) Suppose that

$$\mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})-1} \cap A^c) = 0. \tag{2.98}$$

(H5) Assume that (2.96) holds. The transformations which generated (2.97)–(2.96) did not affect the conditions (2.78)–(2.81). This, (2.84)–(2.85), and (2.96) imply that $\mathbb{P}(\bigcup_{j < m(\mathcal{H})-1} C_{m_-(\mathcal{G})-1, j}) = 0$. We have $\mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})} \cap A^c) = 0$ because of (2.88). These observations and (2.98) imply that $x_{m_-(\mathcal{G})-1} = 1$. But this contradicts the definition of $m_-(\mathcal{G})$.

Hence, (2.96) cannot be true. (X)

(H6) Next suppose that (2.97) is true and recall (2.84)–(2.81). We conclude that $\mathbb{P}(A^c \cap H_{m(\mathcal{H})-1}) = 0$. Hence, $y_{m(\mathcal{H})-1} = 1$. We have shown earlier that $y_{m(\mathcal{H})} = 1$, so $y_{m(\mathcal{H})-1} = y_{m(\mathcal{H})}$. We jump to (A). Then an application of \mathcal{T}_9 will result in the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in view of (2.41) and (2.42).

(H4) Suppose that $\mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})-1} \cap A) = 0$. Then, in view of (2.88),

$$\mathbb{P}(C_{m_-(\mathcal{G}), m(\mathcal{H})} \cap A^c) = 0 \quad \text{and} \quad \mathbb{P}(C_{m_-(\mathcal{G})-1, m(\mathcal{H})-1} \cap A) = 0. \tag{2.99}$$

Recall that we are assuming here that $m_-(\mathcal{G}) \geq 4$. The argument given between (H) and (H6) can be applied with the roles of \mathcal{G} and \mathcal{H} , and those of A and A^c , interchanged. Just as in the case of the original argument, some branches will end with (X) and some will lead to (A). The only branch that will not end this way will generate the following analogue of (2.99):

$$\mathbb{P}(C_{1, m_+(\mathcal{H})} \cap A) = 0 \quad \text{and} \quad \mathbb{P}(C_{2, m(\mathcal{H})+1} \cap A^c) = 0. \tag{2.100}$$

The argument which led to (2.99) did not affect the cells in (2.100), so, by analogy, the argument used to prove (2.100) does not affect the cells in (2.99). Hence, both (2.99) and (2.100) hold.

For all $(k, j) \in \mathcal{D}_-$ we have $\mathbb{P}(C_{k,j}) > 0$ and either $\mathbb{P}(C_{k,j} \cap A^c) = 0$ or $\mathbb{P}(C_{k,j} \cap A) = 0$. It follows from (2.99) and (2.100) that there exists $(k, j) \in \mathcal{D}_-$ with $\mathbb{P}(C_{k,j} \cap A^c) = 0$ and $\mathbb{P}(C_{k+1,j+1} \cap A) = 0$. Fix (k, j) with these properties.

We apply the transformations $\mathcal{T}_{\nearrow}^c((k + 1, j + 1), (k, j_1))$ repeatedly, for all $((k + 1, j + 1), (k, j_1))$ such that $j_1 \neq j, j + 1$. The resulting \mathbf{S} has the property that

$$\mathbb{P}(A^c \cap C_{k+1,j+1}) = 0 \quad \text{or} \tag{2.101}$$

$$\mathbb{P}\left(A^c \cap \bigcup_{j_1 \neq j, j+1} C_{k,j_1}\right) = 0. \tag{2.102}$$

(H7) If (2.101) holds, then $\mathbb{P}(C_{k+1,j+1}) = 0$, because we have assumed that $\mathbb{P}(C_{k+1,j+1} \cap A) = 0$. We now apply the transformation defined in Lemma 7 (or one of its variants described at the beginning of the proof of Lemma 8). This transformation decreases $m(\mathcal{G})$ or $m(\mathcal{H})$. Then we jump to (A).

(H8) If (2.102) holds then we combine it with the assumption that $\mathbb{P}(C_{k,j} \cap A^c) = 0$ to obtain

$$\mathbb{P}(G_k \cap A^c) = \mathbb{P}(C_{k,j+1} \cap A^c). \tag{2.103}$$

Reversing the roles of \mathcal{G} and \mathcal{H} , (k, j) and $(k + 1, j + 1)$, and A and A^c , we obtain the following formula analogous to (2.103):

$$\mathbb{P}(H_{j+1} \cap A) = \mathbb{P}(C_{k,j+1} \cap A). \tag{2.104}$$

It follows from (2.103), (2.104), $k \leq m_-(\mathcal{G})$, and $j + 1 \geq m_+(\mathcal{H})$ that

$$\begin{aligned} \mathbb{P}(A^c \mid C_{k,j+1}) &= \frac{\mathbb{P}(A^c \cap C_{k,j+1})}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1})} \\ &\geq \frac{\mathbb{P}(A^c \cap C_{k,j+1})}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1}) + \mathbb{P}(G_k \setminus C_{k,j+1})} \\ &= \frac{\mathbb{P}(A^c \cap G_k)}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1}) + \mathbb{P}(G_k \setminus C_{k,j+1})} \\ &= \frac{\mathbb{P}(A^c \cap G_k)}{\mathbb{P}(G_k)} = 1 - x_k \geq 1 - \delta, \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}(A \mid C_{k,j+1}) &= \frac{\mathbb{P}(A \cap C_{k,j+1})}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1})} \\ &\geq \frac{\mathbb{P}(A \cap C_{k,j+1})}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1}) + \mathbb{P}(H_{j+1} \setminus C_{k,j+1})} \\ &= \frac{\mathbb{P}(A \cap H_{j+1})}{\mathbb{P}(A^c \cap C_{k,j+1}) + \mathbb{P}(A \cap C_{k,j+1}) + \mathbb{P}(H_{j+1} \setminus C_{k,j+1})} \\ &= \frac{\mathbb{P}(A \cap H_{j+1})}{\mathbb{P}(H_{j+1})} = y_{j+1} \geq 1 - \delta. \end{aligned}$$

The two inequalities contradict each other since $\delta < 1/2$. (X)

(B) We jump here from two points in the proof. We can jump here from (A) in the case when $m_-(\mathcal{G}) \leq 1$. We can also jump here from (I); in this case we have $m_-(\mathcal{G}) = 2$ and $\mathbb{P}(C_{1,m(\mathcal{H})}) = 0$. All branches of the proof corresponding to $m_-(\mathcal{G}) > 2$ (sub-branches of (J) and (H)) ended with (X) or returned to (A).

We have pointed out within the proof that every jump to (A) was associated with the decrease of $m(\mathcal{G})$ or $m(\mathcal{H})$, in most cases because of the application of \mathcal{T}_9 . Hence, after a finite number of jumps to (A), \mathbf{S} must have been transformed so that $m_-(\mathcal{G}) \leq 2$.

We can now exchange the roles of \mathcal{G} and \mathcal{H} and apply the same argument to cells $C_{k,j}$ with $k \geq m_+(\mathcal{G})$ and $j \leq m_-(\mathcal{H})$. In this way, we will eliminate the case $m_-(\mathcal{H}) > 2$ and will be left with the case $m_-(\mathcal{H}) \leq 2$. Moreover, if $m_-(\mathcal{H}) = 2$, we will have $\mathbb{P}(C_{m(\mathcal{G}),1}) = 0$. Some transformations in the new part of the proof are of the type \mathcal{T}_{\nearrow} or \mathcal{T}_{\nearrow}^c ; these transformations do not change any of the x_k, y_j, p_k , or q_j . Transformations of the type $\mathcal{T}_3(\cdot, \cdot)$ in the new part of the proof do not affect cells $C_{k,j}$ with $k \leq m_-(\mathcal{G})$ and $j \geq m_+(\mathcal{H})$. Hence, we have constructed \mathbf{S} satisfying $m_-(\mathcal{G}) \leq 1$, or $m_-(\mathcal{G}) = 2$ and $\mathbb{P}(C_{1,m(\mathcal{H})}) = 0$, and also satisfying $m_-(\mathcal{H}) \leq 1$, or $m_-(\mathcal{H}) = 2$ and $\mathbb{P}(C_{m(\mathcal{G}),1}) = 0$.

Part (i) of the lemma is satisfied because after the initial application of Lemma 3, at the very beginning of the proof, all other transformations satisfied (2.18). □

Proof of Theorem 1. Recall that $\delta \in (0, 1/2)$ is fixed.

For the proof of the lower bound in (1.3), see Proposition 1.

We turn our attention to the upper bound. In view of Lemma 2, we can assume that \mathcal{G} and \mathcal{H} are finitely generated.

Consider any \mathbf{S} . If $\mathbb{P}(B) = 0$ then we are done. Assume that $\mathbb{P}(B) > 0$ and consider any $\varepsilon > 0$. Let \mathbf{S}' be the result of the transformation defined in Lemma 13. Then $\mathbb{P}(B') \geq \mathbb{P}(B) - \varepsilon$.

Parts (ii) and (iii) of Lemma 13 imply that for each $1 \leq k \leq m'(\mathcal{G}')$ there is at most one j such that $C'_{k,j} \subset B'$ and $\mathbb{P}(C'_{k,j}) > 0$, and, similarly, for each $1 \leq j \leq m'(\mathcal{H}')$ there is at most one k such that $C'_{k,j} \subset B'$ and $\mathbb{P}(C'_{k,j}) > 0$. Let \mathcal{B}' be the family of all $1 \leq k \leq m'(\mathcal{G}')$ such that there is a (unique) $j(k)$ such that $C'_{k,j(k)} \subset B'$. Note that $j(k_1) \neq j(k_2)$ if $k_1 \neq k_2$. We apply Lemma 1 as follows:

$$\begin{aligned} \mathbb{P}(B) - \varepsilon &\leq \mathbb{P}(B') = \sum_{k \in \mathcal{B}'} \mathbb{P}(C'_{k,j(k)}) = \sum_{k \in \mathcal{B}'} \mathbb{P}(G'_k \cap H'_{j(k)}) \\ &\leq \sum_{k \in \mathcal{B}'} \frac{\delta}{1 + \delta} (\mathbb{P}(G'_k) + \mathbb{P}(H'_{j(k)})) = \frac{\delta}{1 + \delta} \left(\sum_{k \in \mathcal{B}'} \mathbb{P}(G'_k) + \sum_{k \in \mathcal{B}'} \mathbb{P}(H'_{j(k)}) \right) \\ &\leq \frac{2\delta}{1 + \delta}. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrarily small, $\mathbb{P}(B) \leq 2\delta/(1 + \delta)$. □

Acknowledgements

We are grateful to David Aldous and Jim Pitman for very useful advice. We thank the anonymous referees for many suggestions for improvement. The first author is supported in part by Simons Foundation Grant 506732. The second author is supported in part by NSF Grant DMS-1612483.

References

- [1] BURDZY, K. (2009). *The Search for Certainty. On the Clash of Science and Philosophy of Probability*. World Scientific, Hackensack, NJ.
- [2] BURDZY, K. (2016). *Resonance—from Probability to Epistemology and Back*. Imperial College Press, London.
- [3] BURDZY, K. AND PITMAN, J. (2020). Bounds on the probability of radically different opinions. *Electron. Commun. Probab.* **25**, 14.
- [4] CASARIN, R., MANTOAN, G. AND RAVAZZOLO, F. (2016). Bayesian calibration of generalized pools of predictive distributions. *Econometrics* **4**, 17.
- [5] CICHOMSKI, S. (2020). Maximal spread of coherent distributions: a geometric and combinatorial perspective. Preprint. Available at <https://arxiv.org/abs/2007.08022>.
- [6] DAWID, A. P., DEGROOT, M. H. AND MORTERA, J. (1995). Coherent combination of experts' opinions. *Test* **4**, 263–313.
- [7] DAWID, A. P. AND MORTERA, J. (2018). A note on prediction markets. Preprint. Available at <https://arxiv.org/abs/1702.02502>.
- [8] DEGROOT, M. H. (1988). A Bayesian view of assessing uncertainty and comparing expert opinion. *J. Statist. Planning Infer.* **20**, 295–306.
- [9] DEGROOT, M. H. AND MORTERA, J. (1991). Optimal linear opinion pools. *Manag. Sci.* **37**, 546–558.
- [10] DUBINS, L. E. AND PITMAN, J. (1980). A divergent, two-parameter, bounded martingale. *Proc. Amer. Math. Soc.* **78**, 414–416.
- [11] EASWARAN, K., FENTON-GLYNN, L., HITCHCOCK, C. AND VELASCO, J. D. (2016). Updating on the credences of others: disagreement, agreement, and synergy. *Philosophers' Imprint* **16**, 1–39.
- [12] FRENCH, S. (2011). Aggregating expert judgement. *RACSAM Rev. R. Acad. A* **105**, 181–206.
- [13] GENEST, C. (1984). A characterization theorem for externally Bayesian groups. *Ann. Statist.* **12**, 1100–1105.
- [14] GNEITING, T. AND RANJAN, R. (2013). Combining predictive distributions. *Electron. J. Statist.* **7**, 1747–1782.
- [15] KAPETANIOS, G., MITCHELL, J., PRICE, S. AND FAWCETT, N. (2015). Generalised density forecast combinations. *J. Econometrics* **188**, 150–165.
- [16] KCHIA, Y. AND PROTTER, P. (2015). Progressive filtration expansions via a process, with applications to insider trading. *Internat. J. Theoret. Appl. Finance* **18**, 1550027.
- [17] KRÜGER, F. AND NOLTE, I. (2016). Disagreement versus uncertainty: evidence from distribution forecasts. *J. Banking Finance* **72**, S172–S186.
- [18] LORENZ, J., RAUHUT, H., SCHWEITZER, F. AND HELBING, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proc. Nat. Acad. Sci. USA* **108**, 9020–9025.
- [19] MÖLLER, A. AND GROSS, J. (2016). Probabilistic temperature forecasting based on an ensemble autoregressive modification. *Quart. J. R. Meteor. Soc.* **142**, 1385–1394.
- [20] MORAL-BENITO, E. (2015). Model averaging in economics: an overview. *J. Econom. Surveys* **29**, 46–75.
- [21] RANJAN, R. AND GNEITING, T. (2010). Combining probability forecasts. *J. R. Statist. Soc. B [Statist. Methodology]* **72**, 71–91.
- [22] SATOPÄÄ, V. A., PEMANTLE, R. AND UNGAR, L. H. (2016). Modeling probability forecasts via information diversity. *J. Amer. Statist. Assoc.* **111**, 1623–1633.
- [23] TAYLOR, J. W. AND JEON, J. (2018). Probabilistic forecasting of wave height for offshore wind turbine maintenance. *Europ. J. Operat. Res.* **267**, 877–890.