Algebraic Properties of Blackwell's Order and A Cardinal Measure of Informativeness

Andrew Kosenko*

April 14, 2022

Abstract

I establish a translation invariance property of the Blackwell order over experiments for dichotomies, show that garbling experiments bring them closer together, and use these facts to define a cardinal measure of informativeness. Experiment A is *inf-norm more informative* (INMI) than experiment B if the infinity norm of the difference between a perfectly informative structure and A is *less* than the corresponding difference for B. The better experiment is "closer" to the fully revealing experiment; distance from the identity matrix is interpreted as a measure of informativeness. This measure coincides with Blackwell's order whenever possible, is complete, order invariant, and prior-independent, making it an attractive and computationally simple extension of the Blackwell order to economic contexts.

JEL Classification: D81, D83, C44, C65.

Keywords: Blackwell order, commutative diagrams, informativeness, garbling, experiments, matrix norms.

*Assistant Professor of Economics, Department of Economics, Accounting, and Finance, School of Management, Marist College. Mailing address: Dyson 359, School of Management, Marist College, 3399 North Road, Poughkeepsie NY, 12601. Email: kosenko.andrew@gmail.com.

1 Introduction

How do we rank two Blackwell experiments? To answer this question I ask, and answer, a related question: "How close is an experiment to the most desirable, the ideal - the fully revealing - experiment?" The closer, in an appropriate sense, an experiment is to an ideal experiment, the better it must be. I use this idea to define a new, complete, order over square experiments that has attractive properties.

In a bedrock contribution (Blackwell (1951, 1953)), David Blackwell established the equivalence of two notions of ranking experiments ordinally - those of informativeness, and payoff-richness (as well as the intimately related notion of statistical sufficiency). An *experiment* is a stochastic mapping from a set of states of the world to a set of signal realizations.[1] Experiment A is Blackwell more informative than experiment B (denoted by $A \succeq_B B$) if every expected utility-maximizing decision maker (DM) prefers A to B, or equivalently, if there exists a "garbling" matrix Γ such that $B = \Gamma A$. This order has become a cornerstone of work in information economics, providing a completely unambiguous ranking of information.

The strength of this result comes at a price: the Blackwell order is not only partial, but, loosely speaking, very partial: "most" experiments are not ranked.[2] This is, perhaps, not surprising - information may be valued differently by DMs with different preferences.

The fundamental nature of Blackwell's order, its ubiquity in economics of information and the study of zero-sum games (e.g. Pęski (2008)), coupled with its partial structure, beg the question: what is the "right" completion of this order? Say that experiment A is *inf-norm more informative* than experiment B (denoted by $A \succeq_{INMI} B$) if the infinity norm of the difference between a perfectly informative experiment, and A is *less* than the norm of the difference between a perfectly informative experiment and B. In other words, the better experiment is closer (in the

[1] "Experiments" are also known as "information structures", and "signals".
[2] In order-theoretic terms, $\succeq_B$ is a chain of the partially ordered set of experiments.

sense of matrix norm distance) to the best possible - the fully revealing one. This paper establishes that $\succeq_B \subsetneq \succeq_{INMI}$: Blackwell dominance implies INMI dominance.

I then define a function (d_{INMI}, based on the $\succeq_{INMI}$ order) over experiments which is computed by taking the norm of the matrix difference between an experiment and the identity matrix, and interpret it as a cardinal measure of informativeness. This measure coincides with Blackwell's order, but ranks all finite square experiments, and is one possible completion of the Blackwell order. I work with dichotomies for initial results, but the main theorem is proved for square matrices of any finite size. There can be many such completions; this paper proposes one that has a clear economic (and geometric, in the case of dichotomies) intuition, is computationally simple, prior-independent, conjecturally order invariant, and as such, useful in economic contexts. In addition, this order has an attractive connection with a translation invariance property of $\succeq_B$, which I also establish here.

A brief review of the literature is in section 2, while section 3 gives the translation invariance result. Section 4 clarifies this by showing that garbling experiments brings them closer together in the sense of (matrix) norm of the difference of the two experiments. Section 5 contains the main result: for a particular matrix norm (namely, the infinity norm), $A \succeq_B B$ implies $\|\mathbb{1} - A\|_\infty \leq \|\mathbb{1} - B\|_\infty$. Finally, for an experiment E I define $d_{INMI}(E)$ to be $||\mathbb{1} - E||_\infty$, discuss its properties, make some observations and a conjecture, and conclude. All proofs appear in the appendix.

2 Related Literature

Other useful completions of $\succeq_B$ have been proposed; Cabrales, Gossner, and Serrano (2013) and Cabrales, Gossner, and Serrano (2017) study completions of $\succeq_B$ related to entropy. They restrict attention to particular classes of utility functions in their 2013 work, and evaluate information-price pairs in the 2017 paper.

Frankel and Kamenica (2019) show that a measure of information (a function over pairs of beliefs) is "valid" (equal to the difference between a DM's expected utility when she is acting optimally under the prior and under the posterior, both evaluated at the posterior) if and only if it satisfies attractive axioms. Importantly, validity is stated for pairs of beliefs; they note that while no metric (over beliefs) is valid in their sense, I conjecture that the INMI measure is a representation of a complete order that does satisfy versions of their axioms, reformulated for experiments. They also characterize measures of uncertainty axiomatically, and link the two notions by giving conditions for compatibility of measures of uncertainty and information.

Mu et al. (2021) study repeated Blackwell experiments; along the way they provide a new characterization of $\succeq_B$ using log-likelihood ratios, and relate it to the Rényi order (also an extension of the Blackwell order, itself linked to Kullback-Leibler divergence). They define a function of an experiment ("perfected log-likelihood ratio") and show that ranking these functions according to first-order stochastic dominance is equivalent to $\succeq_B$.

de Oliveira (2018) is very similar in spirit to the present work; he uses category theoretic tools to give a new proof of Blackwell's seminal result on infomativeness, and applies the techniques to a dynamic information acquisition problem. I study a different problem, but the result on translation invariance of $\succeq_B$ has a strong, and related, category-theoretic flavor.

Moscarini and Smith (2002)

3 Preliminaries

There is a state space $\Omega = \{\omega_1, \ldots, \omega_m\}$ with $m \geq 2$, and a signal realization space $S = \{s_1, \ldots, s_m\}$; note that the two sets are assumed to have the same cardinality, an assumption maintained throughout the paper. A *Blackwell experiment* is a finite square stochastic matrix $P = \{p_{ij}\}$ (i.e. $p_{ij} \geq 0$, and for each $j, \sum_i^m p_{ij} = 1$, so that the matrix is column-stochastic). The columns represent the states, the rows represent the signal realizations, and the matrix entries representing the probabilities of signal realizations in each state. Denote by $\mathbb{1}$ the identity matrix, interpreted as a fully revealing experiment - one in which a signal realization always reveals the true state. Denote by U a rank one matrix, interpreted as the fully uninformative experiment - one in which in each state the probability of each signal realization is equal (i.e. is simply $\frac{1}{m}$).

Experiment A *Blackwell dominates* experiment B if there exists a stochastic matrix Γ, with $\Gamma A = B$. Γ is called the *garbling* (or the stochastic transformation) matrix. The interpretation is that one can mimic the signal distribution from the worse experiment in each state by "garbling" (or adding noise to) signals from the better experiment, without knowing anything about the underlying true state.

Sections 4 and 5 restrict attention to 2×2 matrices, while section 6 states the main result for $n \times n$ matrices.

4 Translation Invariance

I begin by noting a curious feature of the Blackwell order: partial translation invariance. If we garble A (say, using Γ_1 as a garbing matrix) to turn it into B, and then garble *both* A and B by the same garbling M, we obtain not only that MA Blackwell-dominates MB (not an entirely surprising result; denote by Γ_2 the matrix that garbles MA into MB), but there is an additional relationship between the mappings Γ_1 and Γ_2 themselves. Let $A = \begin{pmatrix} a_1 & 1-a_2 \\ 1-a_1 & a_2 \end{pmatrix}$ and call an experiment *straightforward* if $\{a_1, a_2\} \in [\frac{1}{2}, 1]^2$.[3]

Theorem 4.1 (Translation invariance of $\succeq_B$)**.** *Let* Γ_1 *be a straightforward* 2×2 *garbling*

[3] It can be shown that focusing on straightforward experiments involves no loss of generality if the only object of interest is the distribution of posterior beliefs.

$$
\begin{tikzcd}
A \arrow[r, "\Gamma_1"] \arrow[d, "M"] & B \arrow[d, "M"] \\
MA \arrow[r, "\Gamma_2"] & MB
\end{tikzcd}
$$

Figure 1: Translation invariance of $\succeq_B$

matrix, and take a non-singular 2×2 *matrix* A*. Let* $B = \Gamma_1 A$ *(i.e.* $A \succeq_B B$*). For any non-singular* 2×2 *matrix* M*, we have that*

1. *MA Blackwell-dominates MB, and furthermore,*

2. *Since there exists* Γ_1 *with* $\Gamma_1 A = B$*, there exists a matrix* Γ_2*, with* Γ_2 *similar to* Γ_1 *such that* $\Gamma_2 MA = MB$

In other words, the diagram in figure 1 commutes.[4]

The import of the theorem is the garblings Γ_1 and Γ_2 are *similar* matrices - in other words, they represent the same linear transformation, but in different bases.[5] Theorem 4.1 states that the garbling M "shifts" any experiment by an amount "proportional" to the initial distance, because the resulting matrices are still ranked, and the Γ_1 and Γ_2 matrices have a particular relationship. In other words, Blackwell's order is partially *translation invariant*; the moniker "partial" is necessary to reflect the need for the Γ_1 to be straightforward. In more mathematical terms, the garbling matrix is a transformation of the matrix of a linear operator. This observation sheds some light on the idea of Blackwell's order as a linear transformation.

The restriction on Γ_1 is not without loss of generality; the interpretation of assuming Γ_1 to be straightforward is that such a garbling matrix does not "flip" the labels of the signal realizations on average. Finally, restricting Γ_1 to be straightfor-

[4]For a discussion of commutative diagrams Mac Lane (1998) is seminal.

[5]And thus, the features of the linear transformation that have to do with the characteristic polynomial (which does not depend on the choice of basis), such as the determinant, trace and eigenvalues, but also the rank and the normal forms, are preserved. The matrix M^{-1} (notably, *not* M) is the change of basis matrix.

ward is a sufficient, but perhaps not a necessary condition.[6]

Of course, this operation can be repeated - one can continue garbling the matrices B and MA, as illustrated in figure 2:

$$\begin{tikzcd}
A \arrow[r, "\Gamma_1"] \arrow[d, "M_1"] & B \arrow[r, "\Gamma_1^1"] \arrow[d, "M_1"] & C \\
MA \arrow[r, "\Gamma_2"] \arrow[d, "M_1^1"] & MB & \\
M_1^1 MA & &
\end{tikzcd}$$

Figure 2: Repeating the argument

Repeating this procedure, one can consider the "horizontal" and "vertical" limits of this diagram, illustrated in figure 3: $\lim_{k\to\infty} M_1^k M_1^{k-1} \dots M_1^1 A$ and $\lim_{k\to\infty} \Gamma_1^k \Gamma_1^{k-1} \dots \Gamma_1^1 A$. Generally speaking, a stochastic matrix P with the property that $\lim_{n\to\infty} P^n = Q$, with Q having all identical rows is said to be *stochastic, indecomposable, and aperiodic (SIA)* (Wolfowitz (1963)). A sufficient condition for the products $\lim_{k\to\infty} \Gamma_1^k$ and $\lim_{k\to\infty} M_1^k$ to converge is if M_1 and Γ_1 are *Sarymsakov* matrices (introduced in Sarymsakov (1961), redefined in Seneta (1979), and generalized in Xia et al. (2019)).[7] In the present binary setting, this implies that if the garbling matrices are Sarymsakov, the limit exists, and is equal to the completely uninformative matrix U, with all entries equal to $\frac{1}{2}$. In general, the question of characterizing all such SIA matrices is open, and an active area of research.

[6] In fact, the theorem is true for most garblings that are not straightforward, but computations show that for some (rare, but nondegenerate) cases if Γ_1 is not straightforward, Γ_2 fails to be stochastic.

[7] Whether there is any economically meaningful interpretation of the definition of a Sarymsakov matrix is unclear, and for this reason, as well as in the interest of brevity, I refrain from discussing this definition and refer the reader to the original literature. I note here simply that this set of mathematical circumstances - the question of convergence of stochastic matrices to a rank one matrix - has appeared before, and is a complicated (NP-hard, in fact - see Blondel and Olshevsky (2014)) problem in general. It is intriguing that this condition has emerged in the context of Blackwell informativeness as well.

$$
\begin{tikzcd}
A \arrow[r, "\Gamma_1"] \arrow[d, "M_1"] & B \arrow[r, "\Gamma_1^2"] \arrow[d, "M_1"] & C \arrow[r, "\Gamma_1^3"] & \dots \arrow[r, "\Gamma_1^k"] & \underbrace{\lim_{k\to\infty} \Gamma_1^k}_{=U, \text{ if } \Gamma_1 \text{ is SIA}} \quad A \\
M_1 A \arrow[r, "\Gamma_2"] \arrow[d, "M_1^2"] & M_1 B & & & \\
M_1^3 A \arrow[d, "M_1^3"] & & & & \\
\vdots \arrow[d, "M_1^k"] & & & & \\
\underbrace{\lim_{k\to\infty} M_1^k}_{=U, \text{ if } M_1 \text{ is SIA}} \quad A & & & &
\end{tikzcd}
$$

Figure 3: Horizontal and vertical limits of repeated garblings

5 Algebraic Properties of the Blackwell Order

Let us now give a precise meaning to the fact that M "shifts" any experiment by an amount "proportional" to the initial distance. A natural notion of distance is the (matrix) norm; for any subordinate (to the vector norm) matrix norm we have $\|MA - MB\| \leq \|M\|\|A - B\|$. In fact, in our setting, a stronger result is true.

Theorem 5.1. *Suppose A is a straightforward 2×2 experiment, and suppose B is another, arbitrary 2×2 experiment. Then for any subordinate matrix norm (for example, $\|\cdot\|_p$ for $p = 1, 2, \infty$, or $\|\cdot\|_F$) we have*

$$\|MA - MB\| \leq \|A - B\| \tag{1}$$

Thus, garbling experiments brings them closer together in the sense of norm differences, for a large class of standard matrix norms. This sheds some light on the statement "M "shifts" any experiment by an amount "proportional" to the initial distance."

6 A Cardinal Measure of Informativeness

Restricting attention to a particular norm - the infinity norm, computed by taking the maximum absolute row sum of the matrix - we get a further result that relates matrix norms and Blackwell's order.

Theorem 6.1. *Let A and B be two $n \times n$ experiments, and suppose that A is straightforward. Then $A \succeq_B B$ implies $\|\mathbb{1} - A\|_\infty \leq \|\mathbb{1} - B\|_\infty$. In other words, $A \succeq_B B \Rightarrow A \succeq_{INMI} B$.*

Thus, the further a matrix is from full revelation, the "worse" it is. The norm is a continuous function,[8] and thus, if $A \succeq_B B$ are Blackwell ranked experiments, this completion assigns "nearby" unranked experiments values that are "close" to the values for A and B. Its interpretation also has the intuitively attractive features that relate this order to Blackwell and mean preserving spreads; figure 4 illustrates.

Say that f *is one representation* of $\succeq$ if $A \succeq B \Rightarrow f(A) \geq f(B)$. Furthermore, if we have a norm, we can define a metric: $\|\mathbb{1} - A\|_\infty \triangleq d(\mathbb{1}, A)$. Putting these definitions together let $d_{INMI}(A) \triangleq d(\mathbb{1}, A)$; theorem 5.1 implies that d_{INMI} is one representation of the Blackwell order. This representation is an extension (in fact, a completion) of it to elements of the set of straightforward square experiments that are not ranked by $\succeq_B$; in other words, d_{INMI} is a stronger, cardinal version of the Blackwell order. Note also that d_{INMI} is defined without reference to a decision problem, and as such, is prior-independent.

I end with a conjecture: note that $d_{INMI}(A) \geq 0$ with equality if and only if $A = \mathbb{1}$, and furthermore, simulations unmistakeably suggest that $d_{INMI}(A \bigotimes B) =$

[8]Where continuity is understood by "continuous in the topology induced by the norm over the vector space of experiments" (see Barfoot and D'Eleutherio (2002) for details of definition of addition that makes this set into a vector space), and then by focusing on the subspace topology that the space of straightforward experiments inherits.

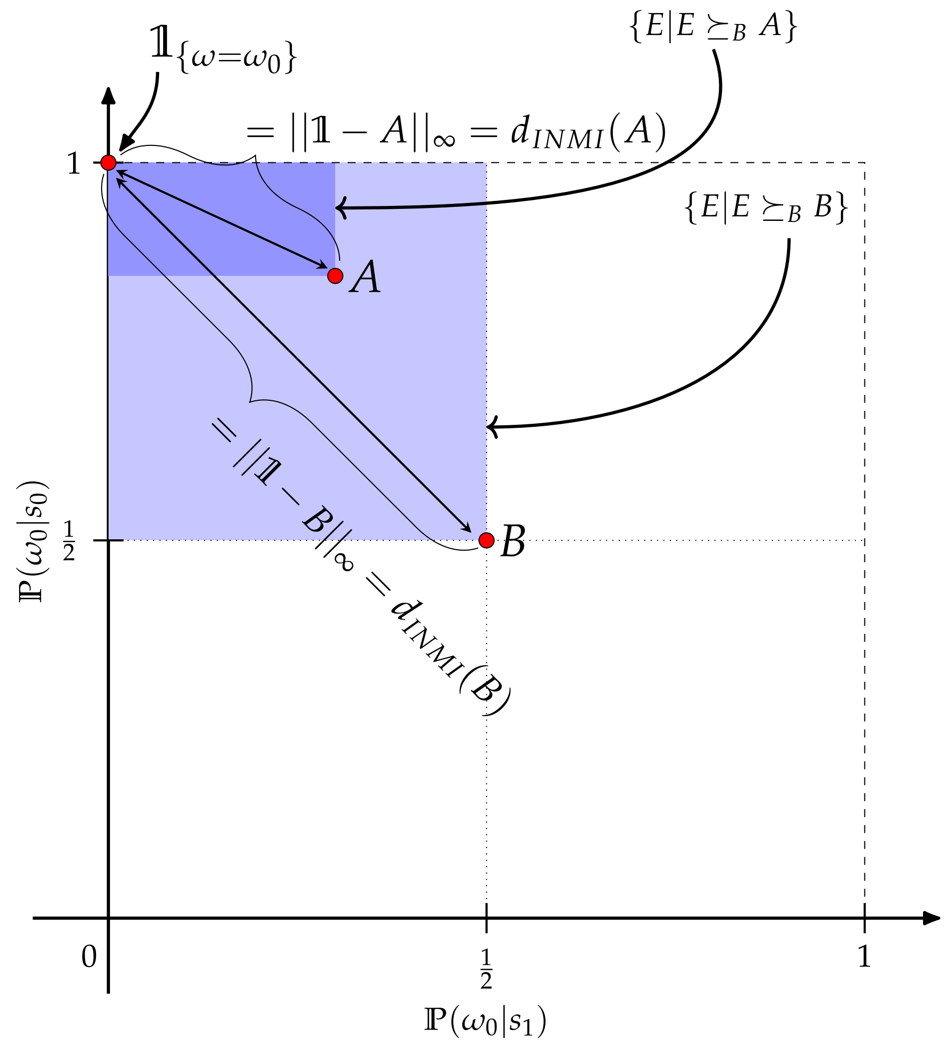

Figure 4: $A \succeq_B B \;\Rightarrow\; A \succeq_{INMI} B$: Blackwell informativeness and norm differences.

In this example there are two possible states, ω_0 and ω_1, and two possible signal realizations, s_0 and s_1. The prior probability of $\omega = \omega_0$ is $\frac{1}{2}$, the true state is ω_0, and A and B are (with abuse of nomenclature) two pairs of posterior beliefs resulting from the eponymous experiments. The possible posterior beliefs after a signal realization are on the axes; in light blue is the set of experiments and posterior belief distributions that are Blackwell better than B (and a mean-preserving spread of posteriors), while in dark blue is the corresponding set for A. E is a generic experiment (and associated posterior belief distribution).

$d_{INMI}(B \otimes A)$,[9] where $\otimes$ is the Kronecker product. In the language of Frankel and Kamenica (2019) this is (an analogue of a) "valid" measure of information. This conjecture provides an intriguing potential link between measures of information and d_{INMI}.

7 Appendix: Proofs

Proof of theorem 3.1. We have that $\Gamma_1 A = B$ by assumption; we need to show the existence of Γ_2 with the stated properties. If it exists, we would have $\Gamma_2 MA = MB$. But then

$$\Gamma_2 MA = MB \iff \Gamma_2 MA = M\Gamma_1 A \tag{2}$$

$$\Rightarrow \Gamma_2 M = M\Gamma_1 \tag{3}$$

$$\Rightarrow \Gamma_2 = M\Gamma_1 M^{-1} \tag{4}$$

Substituting the resulting matrix verifies what was needed to show; the fact that Γ_1 and Γ_2 are similar matrices is immediate from the last equation, which is the definition of similarity. The last equation also gives an explicit formula for Γ_2.

It remains to show that Γ_2 is a garbling - that is, stochastic - matrix. Computing

[9] $A \otimes B$ and $B \otimes A$ are representations of compound experiments where we first observe the realization of the signal from one, and then the other experiment. The interpretation is important - an experiment that represents realizations from multiple information has more rows than columns, while d_{INMI} only ranks square experiments. I exploit the fact that the relevant columns of the Kronecker product of two matrices are numerically equivalent to a matrix representation of a compound experiment; for example, for two binary experiments, the compound experiment is 4×2, while the Kronecker product is 4×4. I construct a square experiment, and ignore the interpretation of the "extra" columns produced by taking the Kronecker product, while retaining them for the purposes of matrix norm difference. While matrix and Kronecker products are not commutative, simulations unequivocally show that d_{INMI} is, althogh the proof is beyond the scope of this note.

the terms explicitly, we obtain

$$M\Gamma_1 M^{-1} = \underbrace{\begin{pmatrix} m_1 & 1-m_2 \\ 1-m_1 & m_2 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} \gamma_1 & 1-\gamma_2 \\ 1-\gamma_1 & \gamma_2 \end{pmatrix}}_{\Gamma_1} \underbrace{\frac{1}{|M|} \begin{pmatrix} m_2 & m_2-1 \\ m_1-1 & m_1 \end{pmatrix}}_{M^{-1}} = \tag{5}$$

$$= \begin{pmatrix} \gamma_1 - \gamma_2 + m_2 - \gamma_1 m_2 + \gamma_2 m_1 & \gamma_1 - \gamma_2 - m_1 - \gamma_1 m_2 + \gamma_2 m_1 + 1 \\ \gamma_2 - \gamma_1 - m_2 + \gamma_1 m_2 - \gamma_2 m_1 + 1 & \gamma_2 - \gamma_1 + m_1 + \gamma_1 m_2 - \gamma_2 m_1 \end{pmatrix} \tag{6}$$

where $|M| = m_1 m_2 - (1-m_2)(1-m_1)$ and $\gamma_1, \gamma_2 \in [\frac{1}{2}, 1]$ by assumption. The columns sum to unity, to confirm that each entry is non-negative one must check cases. For instance, for $\gamma_1 - \gamma_2 + m_2 - \gamma_1 m_2 + \gamma_2 m_1$ to be negative we would need γ_2 and m_2 to be as large as possible (equal to one), which yields a contradiction. This is when the restriction on γ_1 and γ_2 becomes necessary - without it one can obtain cases where the columns of Γ_2 do sum to one, but one of the terms is negative. □

Proof of theorem 4.1. We show this in a sequence of steps; let $\mathbb{1}$ denote an 2×2 identity matrix.

Step 1) $\text{rank}(\mathbb{1} - \Gamma_1) \leq 1$ for any 2×2 column stochastic matrix Γ_1. This is simply because $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \gamma_1 & \gamma_2 \\ 1-\gamma_1 & 1-\gamma_2 \end{pmatrix} = \begin{pmatrix} 1-\gamma_1 & -\gamma_2 \\ \gamma_1 - 1 & \gamma_2 \end{pmatrix}$ for any $\gamma_1, \gamma_2 \in (0,1)$. It is evident that the rank of the resulting matrix is identically 1. If $\gamma_1 = 1$ and $\gamma_2 = 0$ the rank vanishes, since we get the zero matrix. We have assumed that this is not the case (i.e. $A \neq B$) and thus the rank must be equal to unity.

Step 2) $0 < \text{rank}(A - B) = \text{rank}(A - \Gamma_1 A) = \text{rank}((\mathbb{1} - \Gamma_1)A) \leq \min\{\text{rank}(\mathbb{1} - \Gamma_1), \text{rank}(A))\} = 1$.

Step 3) $0 < \text{rank}(MA - MB) = \text{rank}(M(A - \Gamma_1 A)) \leq \min\{\text{rank}(A - \Gamma_1 A), \text{rank}(M)\} = 1$

Step 4) Any rank 1 matrix can be written as an outer product of two vectors (this is a standard result). Thus $A - B = u_1 u_2^T$ and $MA - MB = v_1 v_2^T$ for some 2×1 vectors u_1, v_1, u_2, v_2.

Step 5) We must have $u_1 = v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Let $A = \begin{pmatrix} a_1 & 1 - a_2 \\ 1 - a_1 & a_2 \end{pmatrix}$ and $\Gamma_1 = \begin{pmatrix} \gamma_1 & \gamma_2 \\ 1 - \gamma_1 & 1 - \gamma_2 \end{pmatrix}$ for $\{a_1, a_2\} \in [\frac{1}{2}, 1]^2$ and $\{\gamma_1, \gamma_2\} \in [0, 1]^2$. Then using the previous step, the fact that $\text{rank}(A - B)$, and the fact that these are 2×2 matrices, after some algebra, we obtain the result. Furthermore, in the notation used in this step, we must also have $u_2 = \begin{pmatrix} a_1 - a_1\gamma_1 + \gamma_2(a_1 - 1) \\ \gamma_1(a_2 - 1) - a_2\gamma_2 - a_2 + 1 \end{pmatrix}$. Letting $M = \begin{pmatrix} m_1 & m_2 \\ 1 - m_1 & 1 - m_2 \end{pmatrix}$ for $\{m_1, m_2\} \in [0, 1]^2$, we obtain that

$$v_2 = \begin{pmatrix} a_1 m_1 + [\gamma_2 m_1 - m_2(\gamma_2 - 1)]\,(a_1 - 1) - m_2(a_1 - 1) - a_1\,[\gamma_1 m_1 - m_2(\gamma_1 - 1)] \\ a_2 m_2 + [\gamma_1 m_1 - m_2(\gamma_1 - 1)]\,(a_2 - 1) - m_1(a_2 - 1) - a_2\,[\gamma_2 m_1 - m_2(\gamma_2 - 1)] \end{pmatrix} \tag{7}$$

Step 6) For a matrix A of rank 1 the Frobenius norm and the $p = 2$ norm coincide and are equal to the largest singular value of the matrix, so that $\|A\|_F = \sqrt{\text{tr}(A^T A)}$.

Step 7) Thus $\|A - B\| = \sqrt{\text{tr}(u_2 u_1^T u_1 u_2^T)}$ and $\|MA - MB\| = \sqrt{\text{tr}(v_2 v_1^T v_1 v_2^T)}$. The required difference is equal to

$$\|A-B\| - \|MA - MB\| =$$
$$= \left(2\left[[a_1(1-\gamma_1+\gamma_2)-\gamma_2]^2 + [a_2(1-\gamma_2+\gamma_2)+\gamma_1-1]^2\right]\right)^{\frac{1}{2}} -$$
$$-\left(2\left[[(m_1-m_2)(a_1(1-\gamma_1+\gamma_2)-\gamma_2)]^2 + [(m_2-m_1)(a_2(1-\gamma_2+\gamma_2)+\gamma_1-1)]^2\right]\right)^{\frac{1}{2}}$$
$$\geq 0 \quad (8)$$

$\square$

Proof of theorem 5.1. Let $B = \Gamma A$, and recall that the matrix infinity norm is the maximum absolute row sum of the entries: $||A||_\infty = \max_i \sum_j |a_{ij}| = \sum_{i=1}^n a_{r'i}, \exists r'$. Note that $||\mathbb{1} - A||_\infty = (1 - a_{r_1 r_1}) + \sum_{i \neq r_1}^n a_{r_1 i}$ for some r_1, and analogously, $||\mathbb{1} - B||_\infty = (1 - b_{r_2 r_2}) + \sum_{i \neq r_2}^n b_{r_2 i}$ for some r_2. By definition of matrix multiplication, $b_{ij} = \sum_{k=1}^n \gamma_{ik} a_{kj}$.

We wish to show $\|\mathbb{1} - A\|_\infty \leq \|\mathbb{1} - B\|_\infty$. The contrapositive of this is that for all square A and Γ,

$$\|\mathbb{1} - A\|_\infty > \|\mathbb{1} - \Gamma A\|_\infty = \|\mathbb{1} - B\|_\infty \Rightarrow \tag{9}$$

$$1 - a_{r_1 r_1} + \sum_{i \neq r_1}^n a_{r_1 i} > 1 - b_{r_2 r_2} + \sum_{i \neq r_2}^n b_{r_2 i} \iff \tag{10}$$

$$1 - a_{r_1 r_1} + \sum_{i \neq r_1}^n a_{r_1 i} > 1 - \sum_{k=1}^n \gamma_{r_2 k} a_{k r_2} + \sum_{i \neq r_2}^n \sum_{k=1}^n \gamma_{r_2 k} a_{ki} \iff \tag{11}$$

$$\sum_{i \neq r_1}^n a_{r_1 i} - a_{r_1 r_1} > \sum_{i \neq r_2}^n \sum_{k=1}^n \gamma_{r_2 k} a_{ki} - \sum_{k=1}^n \gamma_{r_2 k} a_{k r_2} \tag{12}$$

Setting $\gamma_{r_2 k}$ to equal the Dirac delta function $\delta_{r_1 k}$ since (eq.(7) has to be true for an arbitrary Γ; note also the change from r_1 to r_2) we obtain the contradiction that

$$\sum_{i\neq r_1}^{n} a_{r_1 i} - a_{r_1 r_1} > \sum_{i\neq r_2}^{n}\sum_{k=1}^{n} \gamma_{r_2 k} a_{ki} - \sum_{k=1}^{n} \gamma_{r_2 k} a_{k r_2} = \sum_{i\neq r_1}^{n} a_{r_1 i} - a_{r_1 r_1} \tag{13}$$

This step shows that there exists a Γ for which eq. (7) is false, and we obtain the contrapositive. The fact that the inequality can be strict can be checked by direct computation. Thus, $\|\mathbb{1} - A\|_\infty \leq \|\mathbb{1} - B\|_\infty$ with a strict inequality in nondegenerate cases. □

Acknowledgments

I am deeply grateful to Navin Kartik for invaluable help and advice, and to an anonymous referee for helpful comments. I thank Yeon-Koo Che and Joseph Stiglitz for immensely beneficial guidance, as well as Joyee Deb, Laura Doval, Guillaume Haeringer, Scott Kominers, Nate Neligh, Anh Nguyen, Pietro Ortoleva, Luca Rigotti, Roberto Serrano, Teck Yong Tan, Roee Teper, and Richard van Weelden for discussions. John Cremin provided expert research assistance. All remaining errors are my own.

References

[1] Barfoot, T.D., and G.M.T. D'Eleuterio. (2002). "An Algebra for the Control of Stochastic Systems: Exercises in Linear Algebra." Fifth International Conference On Dynamics and Control of Systems and Structures in Space. King's College, Cambridge, 14–18 July 2002.

[2] Blackwell, David. (1951). "Comparison of Experiments." Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 93–102, University of California Press, Berkeley, Calif.

[3] Blackwell, David. (1953). "Equivalent Comparisons of Experiments." Ann. Math. Statist. 24, no. 2, 265-272.

[4] de Oliveira, Henrique. (2018). "Blackwell's informativeness theorem using diagrams." Games and Economic Behavior, Volume 109, 126-131. https://doi.org/10.1016/j.geb.2017.12.008

[5] Frankel, Alexander, and Emir Kamenica. (2019). "Quantifying Information and Uncertainty." American Economic Review, 109 (10): 3650-80. https://doi.org/10.1257/aer.20181897

[6] Cabrales, A., Gossner, O., and Serrano, R. (2013). "Entropy and the Value of Information for Investors." The American Economic Review, 103(1), 360-377. http://dx.doi.org/10.1257/aer.103.1.360

[7] Cabrales, Antonio, Olivier Gossner, Roberto Serrano. (2017). "A normalized value for information purchases." Journal of Economic Theory, 170, 266-288. https://doi.org/10.1016/j.jet.2017.05.007

[8] Mac Lane, Saunders. *Categories for the Working Mathematician*. Springer: New York, 1988.

[9] Peski, Marcin. (2008). "Comparison of information structures in zero-sum games." Games and Economic Behavior, 62(2), 732-735. https://doi.org/10.1016/j.geb.2007.06.004

[10] Mu, Xiaosheng, Luciano Pomatto, Philipp Strack, and Omer Tamuz. (2021). "From Blackwell Dominance in Large Samples to Rényi Divergences and Back Again". Econometrica, 89(1), 475-506. https://doi.org/10.3982/ECTA17548

[11] Wolfowitz, J. (1963). Products of Indecomposable, Aperiodic, Stochastic Matrices. Proceedings of the American Mathematical Society, 14(5), 733–737. https://doi.org/10.2307/2034984

[12] Sarymsakov, T.A. (1961). Inhomogeneous Markov Chains, Teor. Veroyatnost. i Primenen., 1961, Volume 6, Issue 2, 194–201

[13] Xia, W., Liu, J., Cao, M., Johansson, K. H., & Basar, T. (2019). Generalized Sarymsakov Matrices. IEEE Transactions on Automatic Control, 64(8), 3085-3100. https://doi.org/10.1109/TAC.2018.2878476

[14] Seneta, E. (1979). "Coefficients of ergodicity: Structure and applications," Advances in Applied Probability, vol. 11, no. 3, pp. 576–590.

[15] Blondel V.D., and A. Olshevsky. (2014). "How to decide consensus: A combinatorial necessary and sufficient condition and a proof that consensus is decidable but NP-hard," SIAM Journal on Control and Optimization, vol. 52, no. 5, pp. 2707–2726.