# Information Hierarchies

Benjamin Brooks, Alexander Frankel, and Emir Kamenica[*]

August 2021

### Abstract

If experiment $A$ is Blackwell more informative than experiment $B$, it is always possible that $A$ and $B$ are induced by signals $A'$ and $B'$ such that $A'$ is a refinement of $B'$, that is, $A'$ entails observing $B'$ plus some additional information. We first show that this result does not extend beyond pairs of experiments: There exist collections of experiments that cannot be induced by a collection of signals so that whenever two experiments are Blackwell ordered, the associated signals are refinement ordered. In other words, sometimes it is impossible for more informed agents to know everything that less informed agents know. More broadly, define an *information hierarchy* to be a partially ordered set that ranks experiments in terms of informativeness. Is it the case that for any choice of experiments indexed on the hierarchy such that higher experiments are Blackwell more informative, there are signals that induce these experiments with higher signals being refinements of lower signals? We show that the answers is affirmative if and only if the undirected graph of the information hierarchy is a forest.

---

# 1  Introduction

There are two distinct things we might mean when we say that $A$ is more informed than $B$. One is that $A$'s information about some state of the world is more accurate than $B$'s. The other is that $A$ knows everything that $B$ knows; that is, $A$ observes what $B$ observes, plus something else. The first notion is typically formalized by representing a source of information as an experiment, i.e., a distribution of posterior beliefs (Blackwell, 1953).[1] We then say that $A$ is Blackwell more informed than $B$ if $A$'s beliefs are a mean-preserving spread of $B$'s beliefs. The second notion is formalized by representing a source of information as a partition of some expanded state space $\Omega \times X$ (Green and Stokey 1978). This formulation distinguishes payoff-relevant states ($\Omega$) from the realizations of signals conditional on those states ($X$). We then say that $A$ knows everything that $B$ knows if $A$'s partition is a refinement of $B$'s partition.

These two formalisms are closely related. Every signal, i.e., partition of the expanded state space, induces an experiment. If one signal is a refinement of another, then the experiment it induces is Blackwell more informative. Moreover, Green and Stokey (1978) establish a partial converse: given two Blackwell-ordered experiments, there exist refinement-ordered signals that induce those experiments. In other words, if $A$'s information is more accurate than $B$'s, it is always possible that $A$ knows everything that $B$ knows.

A natural question is whether Green and Stokey's result extends beyond pairs of experiments. Suppose we have some collection of experiments. Is it always the case that we can construct a collection of signals that induce those experiments, and whenever two experiments are Blackwell ordered, the two corresponding signals are refinement ordered?

We show that the answer is negative. We construct an example where $A$ is Blackwell more informed than $B$ and $C$, who are in turn more informed than $D$, and yet it cannot be that $A$ knows everything that $B$ and $C$ know and that all three know everything that $D$ knows.

To examine this issue more generally, we introduce the notion of an *information hierarchy*, which is simply a partially ordered set. We consider experiment allocations that assign an experiment to each element of the hierarchy, with the property that a higher element's experiment is

---

[1]This definition is equivalent to the prior-free formulation of an experiment as a map from the state to distributions over signal realizations.

Blackwell more informative than a lower element's experiment. We say an experiment allocation is constructible if it is possible to assign a signal to every element of the hierarchy so that each signal induces the corresponding experiment, and a higher element's signal is a refinement of a lower element's signal. If every monotone experiment allocation is constructible, we say that the information hierarchy is universally constructible.

Our main theorem characterizes the set of information hierarchies that are universally constructible. A partially ordered set (and thus an information hierarchy) is associated with an undirected graph, whose nodes are the elements of the set and whose edges are determined by the partial order.[2] We establish that an information hierarchy is universally constructible if and only if this graph is a forest.

We also establish a generalization of this result. We consider a class of other economically meaningful orders on signals, and we show that, if we replace refinement with any one of these orders in our notion of constructibility, our theorem continues to hold.

We discuss two applications of our theorem. First, suppose there is an agent who observes information from several sources and an econometrician who does not know the data generating process behind these sources. We analyze the conditions under which the econometrician can rationalize the agent's reactions to this information. In particular, we show that two necessary conditions, Bayes plausibility and Blackwell monotonicity,[3] do not suffice for rationalizability when there are more than two sources of information. Second, we study information design in organizations under the constraint that managers must have access to the information of their subordinates.

The rest of this paper proceeds as follows. Section 2 establishes basic definitions. Section 3 presents motivating examples. Section 4 contains the remaining definitions and our main result. Section 5 discusses extensions. Section 6 provides two applications. We discuss related literature in probability theory and information economics in Section 7. Section 8 is a conclusion.

---

[2]The standard representation of a partially ordered set as a directed graph encodes the partial order by placing an edge from $n$ to $n'$ if $n$ covers $n'$, i.e., if $n > n'$ and there is no $n''$ such that $n > n'' > n'$. We associate with each information hierarchy the undirected version of this directed graph.

[3]Bayes plausibility requires that the agent's average belief does not vary across the sets of observed information sources. Blackwell monotonicity requires that when an agent observes a superset of sources, her beliefs are more dispersed than when she observes a subset.

## 2 States, signals, and experiments

There is a finite state space $\Omega$ and a prior $\mu_0 \in \Delta(\Omega)$.

A *signal* $\pi$ is a finite partition of $\Omega \times [0,1]$ s.t. $\pi \subset S$, where $S$ is the set of non-empty Lebesgue-measurable subsets of $\Omega \times [0,1]$ (Green and Stokey, 1978; Gentzkow and Kamenica, 2017). An element $s \in S$ is a *signal realization*. The interpretation of this formalism is that a random variable $x$, drawn uniformly from $[0,1]$, determines the signal realization conditional on the state. Thus, the conditional probability of $s$ given $\omega$ is $p(s|\omega) = \lambda(\{x | (\omega, x) \in s\})$ where $\lambda(\cdot)$ denotes the Lebesgue measure. Importantly, if we know that one agent observed some signal, say $\pi_A$, and another agent observed some signal $\pi_B$, we can deduce not only how informed each agent is about $\omega$, but also the joint distribution of their signal realizations (and thus their higher-order beliefs).

In contrast, a (Blackwell) experiment only specifies how informed an agent is about $\omega$, without specifying her information about the beliefs of other agents. The standard definition of an experiment is a map from $\Omega$ to distributions over signal realizations, but for ease of exposition we identify each experiment with the distribution of beliefs it induces. Thus, we define an *experiment*, denoted by $\tau$, as an element of $\Delta(\Delta(\Omega))$ that has finite support and satisfies $\mathbb{E}_\tau[\mu] = \mu_0$.

We write $\pi \trianglerighteq \pi'$ if $\pi$ is a *refinement* of $\pi'$, i.e., every element of $\pi$ is a subset of some element of $\pi'$.[4] If $\pi \trianglerighteq \pi'$, an agent who observes $\pi$ has access to all the information available to the agent who observes $\pi'$ (and thus knows those agent's beliefs). We write $\tau \succsim \tau'$ if $\tau$ is *Blackwell more informative* than $\tau'$, i.e., $\tau$ is a mean-preserving spread of $\tau'$. If $\tau \succsim \tau'$, an agent who observes $\tau$ obtains a higher payoff than the agent who observes $\tau'$ in any decision problem.

We denote the set of all signals by $\Pi$. The pair $(\Pi, \trianglerighteq)$ is a lattice and we let $\vee$ denote the join, i.e., $\pi \vee \pi'$ is the coarsest refinement of both $\pi$ and $\pi'$. Note that $\pi \vee \pi'$ is the signal that is equivalent to observing both $\pi$ and $\pi'$.

Each signal $\pi$ induces an experiment, denoted $\langle \pi \rangle$, according to

$$\langle \pi \rangle (\mu) = \sum_{\{s \in \pi | \mu_s = \mu\}} p(s)$$

where $p(s)$ is the unconditional probability of signal realization $s$ and $\mu_s$ is the posterior belief about

---

[4] We then also say $\pi'$ is *coarser* than $\pi$.

$\omega$ conditional on $s$.[5] Moreover, given signal $\pi$, we let $\tilde{\mu}_\pi$ denote the belief-valued random variable on $\Omega \times [0,1]$ that is the posterior induced by the observation of the signal realization from $\pi$. Note that the refinement order implies a belief-martingale property: if $\pi \trianglerighteq \pi'$, then $\mathbb{E}\left[\tilde{\mu}_\pi | \tilde{\mu}_{\pi'}\right] = \tilde{\mu}_{\pi'}$.[6] In other words, observing additional information cannot change one's belief on average.

## 3   Illustrative examples

There are four experiments, $\tau_A$, $\tau_B$, $\tau_C$, and $\tau_D$, some of which might be Blackwell more informative than others. We wish to know whether there exist signals $\pi_i$ for $i \in \{A, B, C, D\}$ such that

(C1)  $\langle \pi_i \rangle = \tau_i$ and

(C2)  $\tau_i \succsim \tau_j$ implies $\pi_i \trianglerighteq \pi_j$.

As we discussed in the introduction, the answer to this question will shed light on (i) whether Green and Stokey's (1978) result on the equivalence of representing information as experiments vs. signals generalizes beyond pairwise comparisons, and (ii) whether it is always possible that a more informed agent knows what a less informed agent knows.

As we will see, for some Blackwell rankings of the experiments, the answer is yes, no matter what the the exact experiments are. For other rankings, this is not the case; there exist experiments for which there do not exist signals satisfying (C1) and (C2). We illustrate these results through three examples, depicted in Figure 1.

**Chain**

Suppose $\tau_A \succsim \tau_B \succsim \tau_C \succsim \tau_D$. When experiments are ordered in this way, it is always possible to construct signals that satisfy (C1) and (C2), regardless of the particular experiments. This follows from Green and Stokey (1978): For any $\pi'$ and $\tau$ with $\tau \succsim \langle \pi' \rangle$, there exists a $\pi$ such that (i) $\langle \pi \rangle = \tau$ and (ii) $\pi \trianglerighteq \pi'$.[7] In other words, take any signal that induces some experiment; there is a refinement of this signal that induces any particular more-informative experiment.

---

[5]Note that $p(s) = \sum_{\omega \in \Omega} p(s|\omega) \mu_0(\omega)$. For any $s$ with $p(s) > 0$, we have $\mu_s(\omega) = \frac{p(s|\omega)\mu_0(\omega)}{p(s)}$. For those $s$ with $p(s) = 0$, we can set $\mu_s$ to be any arbitrary belief.

[6]Throughout the paper, when we say two random variables are equal, we mean almost surely.

[7]This result first appears as Theorem 1 in Green and Stokey (1978). This result is stated and proved using the notation in our paper by Gentzkow and Kamenica (2017).

Figure 1: Example Hierarchies



(a) Chain     (b) Tree     (c) Diamond
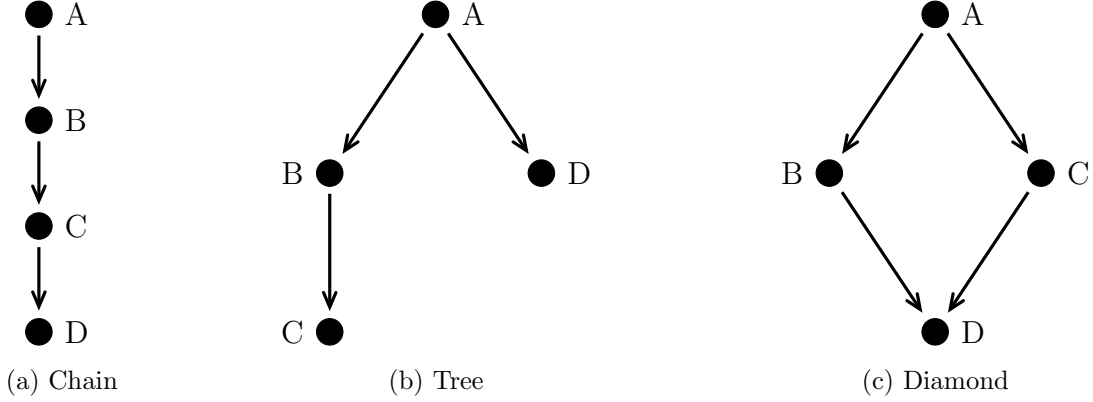
Now, consider any $\pi_D$ that induces $\tau_D$. Since $\tau_C \succsim \tau_D$, there exists some signal $\pi_C \trianglerighteq \pi_D$ that induces $\tau_C$. Similarly, there is a $\pi_B \trianglerighteq \pi_C$ that induces $\tau_B$. Finally, there is a $\pi_A \trianglerighteq \pi_B$ that induces $\tau_A$. Hence, we constructed signals that satisfy (C1) and (C2).

**Tree**

Next, suppose $\tau_A \succsim \tau_B \succsim \tau_C$, $\tau_A \succsim \tau_D$, and $\tau_D$ is not comparable with $\tau_B$ or $\tau_C$. In this case, it is again always possible to construct signals so that $\langle \pi_i \rangle = \tau_i$ and $\tau_i \succsim \tau_j$ implies $\pi_i \trianglerighteq \pi_j$. The argument, however, is more subtle than for the chain. To see the issue, suppose we first go "up the tree" and proceed as in the previous example: we construct $\pi_D$ and then $\pi_A \trianglerighteq \pi_D$ that induce $\tau_D$ and $\tau_A$. Now, to assign a signal $\pi_B$, we have to go "down the tree." But there is no guarantee that, given the $\pi_A$ we constructed, there exists a $\pi_B$ that is coarser than $\pi_A$ and induces $\tau_B$.

We establish a result (Lemma 1), however, that tells us we can construct $\pi_B$ that induces $\tau_B$ and has the property that $\langle \pi_A \vee \pi_B \rangle = \langle \pi_A \rangle = \tau_A$. (To do so, we build a $\pi_B$ whose realizations are independent of the state given the realization of $\pi_A$.) Then, we replace the "provisional" $\pi_A$ with $\pi_A \vee \pi_B$ (which is of course also finer than $\pi_D$). We then apply a similar procedure to assign $\pi_C$ and update the provisional $\pi_A$ and $\pi_B$. This algorithm is discussed in greater detail in Section 4. Hence, we constructed signals that satisfy (C1) and (C2).

Figure 2: The diamond hierarchy is not universally constructible
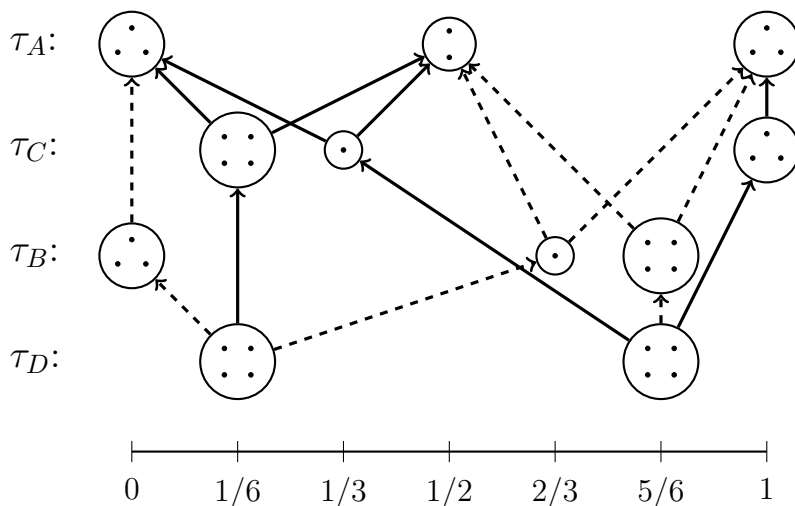
## Diamond

Finally, suppose $\tau_A \succsim \tau_B \succsim \tau_D$, $\tau_A \succsim \tau_C \succsim \tau_D$, and $\tau_B$ and $\tau_C$ are not comparable. We now give an example of four experiments ordered in this manner such that there do not exist signals that satisfy (C1) and (C2).[8]

Let $\Omega = \{0, 1\}$ with a uniform prior. Since the state space is binary, we associate each belief with $\Pr(\omega = 1) \in [0, 1]$. The rows in Figure 2 depict the four experiments, with the number of dots in a circle proportional to the probability mass on that belief. It is easy to see that these four experiments satisfy the given Blackwell ordering, with arrows (solid or dashed) showing how to obtain a higher-ranked distribution as a mean-preserving spread of a lower-ranked distribution.

Now, to obtain a contradiction, suppose that the signals $\pi_A$, $\pi_B$, $\pi_C$, and $\pi_D$ satisfy (C1) and (C2). First, note that these signals do not merely induce the respective marginal distributions of beliefs (i.e., the experiments), but they also determine the joint distribution of these beliefs. In particular, as noted earlier, whenever $\pi \trianglerighteq \pi'$ we must have $\mathbb{E}[\tilde{\mu}_\pi | \tilde{\mu}_{\pi'}] = \tilde{\mu}_{\pi'}$.

We use this fact to show that when $\tilde{\mu}_{\pi_D} = \frac{1}{6}$, (i) $\pi_A \trianglerighteq \pi_B \trianglerighteq \pi_D$ implies that $\tilde{\mu}_{\pi_A} \neq 1$, while (ii) $\pi_A \trianglerighteq \pi_C \trianglerighteq \pi_D$ implies that $\tilde{\mu}_{\pi_A} = 1$ with a strictly positive probability. To see (i), note that when $\tilde{\mu}_{\pi_D} = \frac{1}{6}$, we must have $\tilde{\mu}_{\pi_B} = \frac{1}{6}$ and thus $\tilde{\mu}_A \neq 1$ (since $\tilde{\mu}_{\pi_A} = 1$ if and only if $\tilde{\mu}_{\pi_B} = 1$). To see (ii), note that when $\tilde{\mu}_{\pi_D} = \frac{1}{6}$, there is a strictly positive probability that $\tilde{\mu}_{\pi_C} = \frac{2}{3}$ (since $\tilde{\mu}_{\pi_C} = \frac{5}{6}$ if

---

[8]The mathematical details of this example are presented in the proof of Lemma 6 in the Appendix.

and only if $\tilde{\mu}_{\pi_D} = \frac{5}{6}$); and when $\tilde{\mu}_{\pi_C} = \frac{2}{3}$, there is a strictly positive probability that $\tilde{\mu}_{\pi_A} = 1$ (since all of the other beliefs in the support of $\tilde{\mu}_{\pi_A}$ are strictly lower than $\frac{2}{3}$). Thus, we have reached a contradiction.

## 4  Universal constructibility

### 4.1  Information hierarchies and constructibility

These examples motivate the following issue. Partially order a collection of experiments so that higher experiments are Blackwell more informative than lower experiments. Is it possible to find signals that induce these experiments, and that satisfy the property that signals inducing higher experiments are refinements of signals inducing lower ones? We argued that this was in fact always possible when the ranking on experiments corresponded to one of the chain or tree examples above but was not necessarily possible when experiments were ranked as a diamond. The main result of this paper is an if-and-only-if characterization of the rankings of experiments such that suitable refinement-ordered signals exist for any experiments whose Blackwell order is consistent with the ranking.

We now provide the remaining definitions needed to state our result. An *information hierarchy* $H$ is a finite partially ordered set $(N, \geq)$. We refer to elements of $N$ as *nodes*. Given $n, n' \in N$, we say that $n$ *covers* $n'$ if $n > n'$ and there does not exist $n'' \in N$ with $n > n'' > n'$. The *graph of $H$*, denoted $G(H)$, is an undirected graph whose nodes are the elements of $N$ with an edge between two nodes if one covers the other. $G(H)$ is a *forest* if there is at most one path between any two nodes.[9]

Fix an information hierarchy $H$ and $(\Omega, \mu_0)$. An *experiment allocation* on $H$ is a map that

---

[9]It is standard to represent a partially ordered set as a directed graph, e.g., as we do in Figure 1. Our main theorem relates constructibility of a hierarchy to a property of its undirected analogue; this is why we define $G(H)$ as the undirected version. Formally, the directed graph $\overrightarrow{G}(H)$, is the pair $\left(N, \overrightarrow{E}\right)$, where $N$ is the set of nodes, $\overrightarrow{E} \subseteq N \times N$ is the set of directed edges, and $(n, n') \in \overrightarrow{E}$ if $n$ covers $n'$. A directed path from $n$ to $n'$ is an alternating sequence of nodes and directed edges $(n_0, \overrightarrow{e}_0, ..., n_{L-1}, \overrightarrow{e}_{L-1}, n_L)$, where $L > 0$, $n_0 = n$, $n_L = n'$, $n_l \in N$ for all $l \in \{0, ..., L\}$, $\overrightarrow{e}_l = (n_l, n_{l+1}) \in \overrightarrow{E}$ for all $l \in \{0, ..., L-1\}$, and $l \neq l' \Rightarrow \overrightarrow{e}_l \neq \overrightarrow{e}_{l'}$. $G(H)$ is the pair $(N, E)$, where $N$ is the set of nodes, $E \subseteq \{e \subseteq N \mid |e| = 2\}$ is the set of undirected edges, and $\{n, n'\} \in E$ if $n$ covers $n'$ or $n'$ covers $n$. An undirected path from $n$ to $n'$ is an alternating sequence of nodes and undirected edges $(n_0, e_0, ..., n_{L-1}, e_{L-1}, n_L)$, where $L > 0$, $n_0 = n$, $n_L = n'$, $n_l \in N$ for all $l \in \{0, ..., L\}$, $e_l = \{n_l, n_{l+1}\} \in E$ for all $l \in \{0, ..., L-1\}$, and $l \neq l' \Rightarrow e_l \neq e_{l'}$.

assigns an experiment to every node; experiment allocation $\beta$ is *monotone* if $n \geq n'$ implies that $\beta(n) \succsim \beta(n')$, i.e., higher nodes are Blackwell more informed than lower nodes. A *signal allocation* on $H$ is a map that assigns a signal to every node; signal allocation $\sigma$ is *monotone* if $n \geq n'$ implies that $\sigma(n) \trianglerighteq \sigma(n')$, i.e., signals associated with higher nodes refine signals associated with lower nodes. A signal allocation $\sigma$ *induces* an experiment allocation $\beta$ if for all $n$, $\beta(n) = \langle \sigma(n) \rangle$. A (monotone) experiment allocation on $H$ is *constructible* if there exists a monotone signal allocation that induces it.[10]

We say that $H$ is *universally constructible* if for every $\Omega$ and $\mu_0$, every monotone experiment allocation is constructible.

## 4.2  Main result

We now present the main result of the paper:

**Theorem 1.** *An information hierarchy is universally constructible if and only if its graph is a forest.*

To relate this result to our motivating examples, both the chain and the tree are examples of information hierarchies whose graphs are forests; hence, every monotone experiment allocation is constructible. The graph of the diamond, however, is not a forest, because there are two paths from $A$ to $D$, namely the path going through $B$ and the path going through $C$; hence, we were able to present a monotone experiment allocation that is not constructible.

A rigorous proof of Theorem 1 is in the Appendix. The remainder of this section sketches the argument.

## 4.3  Proof sketch: If

We begin by explaining why, if $G(H)$ is a forest, then any monotone experiment allocation can be constructed. We illustrate this for the special case when $G(H)$ is a tree. One can then obtain the general case by applying the same argument on each of the disjoint trees that make up the forest.

The proof relies on the following result, which is of independent information-theoretic interest. We say that $\pi'$ is *statistically redundant* given $\hat{\pi}$ if $\langle \hat{\pi} \vee \pi' \rangle = \langle \hat{\pi} \rangle$, i.e., observing $\hat{\pi}$ and $\pi'$ yields the

---

[10]It is immediate that any constructible experiment allocation is monotone because $\pi \trianglerighteq \pi' \Rightarrow \langle \pi \rangle \succsim \langle \pi' \rangle$.

same beliefs as observing $\hat{\pi}$ only.[11]

**Lemma 1.** *Fix signals $\overline{\pi}$ and $\pi$, with $\overline{\pi} \trianglerighteq \pi$, and an experiment $\tau'$ that is Blackwell less informative than $\langle \pi \rangle$. There exists a signal $\pi'$ that induces $\tau'$ and is statistically redundant given any signal that is between $\pi$ and $\overline{\pi}$ in the refinement order.*

To better understand the content of Lemma 1, it is helpful to note that the following, stronger, conjecture does not hold. One might think that, analogously to the aforementioned result of Green and Stokey (1978), for any $\pi$ and $\tau'$ with $\tau'$ Blackwell less informative than $\langle \pi \rangle$, there exists a $\pi'$ that induces $\tau'$ and is coarser than $\pi$. This is not the case.[12] Lemma 1 implies, however, that we can nonetheless find a $\pi'$ that induces $\tau'$ and is statistically redundant given $\pi$, even though we cannot guarantee that $\pi \trianglerighteq \pi'$. Moreover, the Lemma further implies we can find a $\pi'$ so that $\pi'$ is statistically redundant given any $\hat{\pi}$ such that $\pi \trianglelefteq \hat{\pi} \trianglelefteq \overline{\pi}$.

Why do we need the upper bound $\overline{\pi}$? Once we fix any $\overline{\pi}$ and $\pi$, we can construct $\pi'$ that is statistically redundant given $\overline{\pi}$, $\pi$, and anything in between. But, suppose we only fix $\pi$ and consider an arbitrary $\pi'$. As long as $\pi'$ is not coarser than $\pi$, there is always some refinement of $\pi$, say $\pi^*$, such that $\pi'$ is not statistically redundant given $\pi^*$.

The proof of Lemma 1 is constructive. To see how the construction works, consider the example in Figure 3. The state is either $L$ or $R$. Each of the four rows represents a signal, i.e., a partition of $\{L, R\} \times [0, 1]$. The two rectangles in each row represent the unit interval crossed with the two states. Each signal realization is indicated by its pattern. Note that $\overline{\pi} \trianglerighteq \pi$, and while $\langle \pi \rangle \succsim \langle \tilde{\pi} \rangle$, it is not the case that $\pi \trianglerighteq \tilde{\pi}$ nor that $\langle \pi \vee \tilde{\pi} \rangle = \langle \pi \rangle$.

To establish the claim in Lemma 1, we need to construct a signal $\pi'$ that induces the same beliefs as $\tilde{\pi}$, but is statistically redundant given $\pi$, and given any refinement of $\pi$ up to $\overline{\pi}$. The bottom row illustrates such a construction. Each signal realization of $\pi'$ corresponds to a signal realization of $\tilde{\pi}$, with the same likelihood in each state. However, the "locations" of the signal realizations in $\pi'$ are re-arranged so that the conditional probability of each signal realization of $\pi'$ in state $\omega$ given

---

[11]It may be helpful to contrast statistical redundancy to the stronger notion of refinement/coarsening. If $\hat{\pi} \trianglerighteq \pi'$, we do not merely have that $\langle \hat{\pi} \vee \pi' \rangle = \langle \hat{\pi} \rangle$, we also have that $\langle \hat{\pi} \vee \pi' \vee \pi'' \rangle = \langle \hat{\pi} \vee \pi'' \rangle$ for any signal $\pi''$. For example, in Figure 3, $\pi'$ is not a coarsening of $\pi$, but (as explained in the discussion below) $\pi'$ is statistically redundant given $\pi$.

[12]For example, suppose that $\langle \pi \rangle \succsim \tau'$ but the support of $\tau'$ has more elements than the number of signal realizations in $\pi$. Then, no $\pi'$ that induces $\tau'$ could be a coarsening of $\pi$.

Figure 3: Illustration of Lemma 1

$$\omega = L \qquad\qquad\qquad \omega = R$$
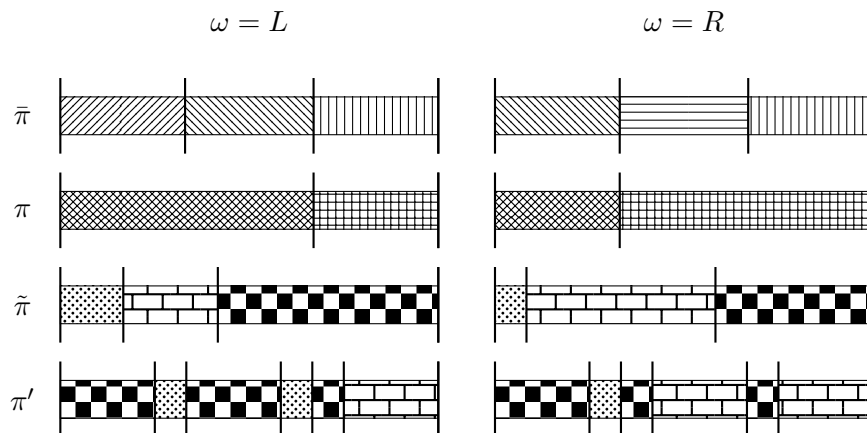


$\bar{\bar{\pi}}$ is (i) the same for $\omega = L$ and $\omega = R$, and (ii) the same for any elements of $\bar{\bar{\pi}}$ that refine the same element of $\pi$. Property (i) ensures that $\pi'$ is statistically redundant given $\bar{\bar{\pi}}$, while (i) and (ii) together ensure that it is also redundant given any $\hat{\pi}$ s.t. $\bar{\bar{\pi}} \trianglerighteq \hat{\pi} \trianglerighteq \pi$. The proof in the Appendix generalizes this construction.

With Lemma 1 in hand, we can now establish the universal constructibility of trees. Consider the tree hierarchy described in Section 3. Take some monotone experiment allocation $\beta$ on this hierarchy. We construct a sequence of monotone signal allocations on progressively larger subsets of $N$, adding one node at a time. The signal assigned to a given node may be revised as the sequence progresses.

In Step 1 we pick an arbitrary node, say node $D$. Given $\beta(D)$, let $\sigma^1(D)$ be any signal that induces $\beta(D)$. In Step 2 we pick a node connected to $D$, node $A$. Node $A$ is above node $D$ and so we follow the same procedure as in the chain example to tentatively assign a suitable signal $\sigma^2(A)$ to $A$, with $\langle \sigma^2(A) \rangle = \beta(A)$ and $\sigma^2(A) \trianglerighteq \sigma^1(D)$; the signal to $D$ is unchanged in this step, i.e., we let $\sigma^2(D) = \sigma^1(D)$. In Step 3 we pick an unassigned node connected to $A$ or $D$, which here must be node $B$. Since $B$ is below $A$, a complication in assigning a signal to $B$ is that there may not exist a signal $\tilde{\pi}$ such that $\langle \tilde{\pi} \rangle = \beta(B)$ and yet $\sigma^2(A) \trianglerighteq \tilde{\pi}$. This is where Lemma 1 comes into play. Applying Lemma 1 with $\pi = \bar{\pi} = \sigma^2(A)$ and $\tau = \beta(B)$, we conclude that there exists a signal $\pi'$ that induces $\beta(B)$ and is statistically redundant with $\sigma^2(A)$, so that $\langle \sigma^2(A) \vee \pi' \rangle = \langle \sigma^2(A) \rangle$. We then set $\sigma^3(B) = \pi'$, we replace the initial assignment to $A$ with $\sigma^3(A) = \sigma^2(A) \vee \pi'$, and we leave the signal at $D$ unchanged at $\sigma^3(D) = \sigma^2(D)$. Finally, in Step 4, we move to the final node

11

$C$, below the previously assigned node $B$. We now apply Lemma 1 with $\bar{\pi} = \sigma^3(A)$,[13] $\pi = \sigma^3(B)$, and $\tau = \beta(C)$ to obtain a new signal $\pi'$ that induces $\beta(C)$ and is statistically redundant with the previously assigned signals at nodes above $C$. We then set $\sigma^4(C) = \pi'$ and replace the signals of the previously assigned nodes above $C$ as in Step 3.

This completes the construction of a monotone signal allocation inducing experiment allocation $\beta$. The full details of this procedure, applied to any hierarchy whose graph is a forest, are the heart of the formal proof of the if direction.

Which properties of a tree are used in this argument? When we apply this procedure to trees (or forests more generally), we can add the nodes in a way so that the new node $n$ always covers or is covered by exactly one node $n'$ to which we have previously assigned a signal. When $n$ covers $n'$, we can apply Green and Stokey's (1978) result to obtain the new signal at $n$. When $n$ is covered by $n'$, we can apply Lemma 1 instead. If the hierarchy is not a forest, however, no matter how we add the nodes, we will inevitably find ourselves in a situation where the new node covers or is covered by more than one node, in which case neither Green and Stokey's (1978) nor Lemma 1 is of use.

### 4.4 Proof sketch: Only if

We now sketch the proof of the other direction of Theorem 1, that is, a hierarchy is universally constructible only if its graph is a forest. The proof consists of three main steps. Step 1 introduces a notion of a closed subhierarchy and establishes that a hierarchy is universally constructible only if its closed subhierarchies are as well. Step 2 shows that every hierarchy that is not a forest contains a closed subhierarchy that is one of two types. Step 3 shows that both of these types are not universally constructible, thus completing the proof.

Step 1: Given a hierarchy $H = (N, \geq)$, a subset of nodes $N' \subseteq N$ induces the hierarchy $H' = (N', \geq)$, which we refer to as a *subhierarchy* of $H$, with the partial order being the restriction of $\geq$ to $N'$. Say that $H'$ is *closed* if for every $n', n'' \in N'$ and $n \in N$, $n' \geq n \geq n''$ implies $n \in N'$. In other words, $N'$ contains all the nodes from $N$ that are "between" the nodes of $N'$. Lemma 2 in the Appendix shows that if $H$ is universally constructible, then every closed subhierarchy of $H$ is also universally constructible. The argument is as follows. Fix a closed subhierarchy $H'$. We show that

---

[13] In general, $\bar{\pi}$ is taken to be the join of all signals previously assigned to nodes above the new node, which in this case is simply $\sigma^3(A)$.

any monotone experiment allocation $\beta$ on $H'$ can be extended to a monotone experiment allocation on all of $H$. This step uses the hypothesis that $H'$ is closed, since it means that in extending $\beta$, we never have to fill in beliefs "in between" nodes in $H'$. As $H$ is universally constructible, there is a monotone signal allocation on $H$ that induces the extension of $\beta$, and the restriction of this signal allocation to $N'$ is a monotone signal allocation on $H'$ that induces $\beta$. Since $\beta$ is arbitrary, $H'$ is universal constructible.

Step 2: Suppose $G(H)$ is not a forest. A hierarchy is *cyclic* if $G(H)$ is not a forest. Given that $N$ is finite and $H$ is cyclic, it follows that $H$ contains a *minimal cyclic closed subhierarchy (MCC)*, i.e., a subhierarchy $H' = (N', \geq)$ such that: $H'$ is cyclic, closed, and there does not exist $N'' \subsetneq N'$ such that $H'' = (N'', \geq)$ is cyclic and closed. Any MCC $H'$ must fall into one of the following categories:

- $H'$ is a *union of non-comparable paths* (UNP): It contains a maximal node $\overline{n}$ and a minimal node $\underline{n}$, and its graph consists of at least two paths between $\overline{n}$ and $\underline{n}$. Moreover, nodes in $N \setminus \{\overline{n}, \underline{n}\}$ are comparable only if they are in the same path. The diamond is an example of a UNP; another example is depicted in Figure 4a.

- $H'$ is a *crown*: It contains exactly four nodes, which we label $F$(ather), $M$(other), $S$(on), $D$(aughter), and the partial order consists of $F \geq S$, $F \geq D$, $M \geq S$, and $M \geq D$. A crown is depicted in Figure 6a.

This taxonomy takes considerable effort to prove formally, but the high level argument is as follows. An MCC $H'$ contains maximal nodes, which are not covered by any other node, and minimal nodes, which do not cover any other nodes. We distinguish two cases on these maximal and minimal nodes.

In the first case, every maximal node in $H'$ covers every minimal node. Then, the fact that $H'$ is minimal, together with the existence of a cycle, implies that $H'$ contains exactly four nodes and is in fact a crown.

Alternatively, there is a maximal node $\overline{n}$ that does not cover a minimal node $\underline{n}$. Then, we show that $H'$ is a UNP. To see this, note the following two subcases. First, it may be that $H'$ is simply the set of nodes that are between $\overline{n}$ and $\underline{n}$, i.e., a *between set*. Then, $H'$ consists of a series of directed paths between $\overline{n}$ and $\underline{n}$. Now, if nodes in distinct paths were comparable, then it would

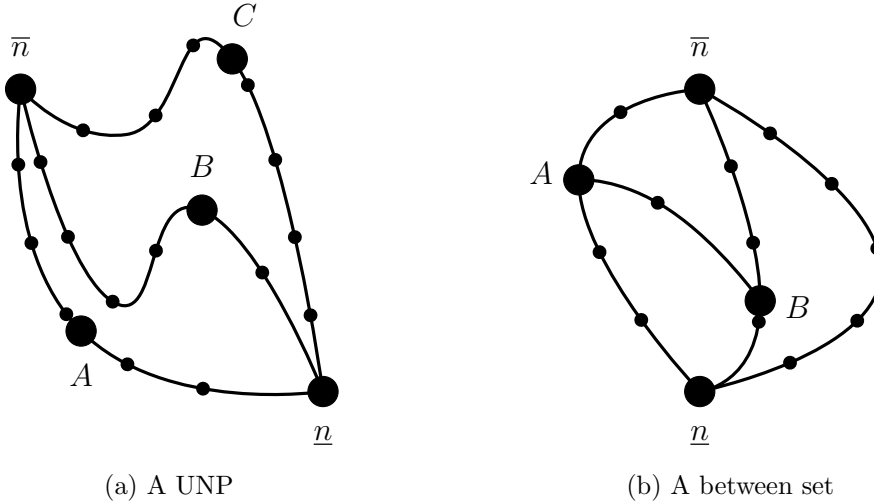(a) A UNP                                      (b) A between set

Figure 4: For adjacent nodes, the node that is higher in the figure covers the node that is lower. (a) A UNP. Nodes $A$, $B$, and $C$ are in different paths, and therefore are not comparable. (b) A between set which is not an MCC. There are two overlapping (directed) paths that go from $\overline{n}$ to $A$ to $\underline{n}$, one that goes through $B$ and one that does not. There is a smaller cyclic closed subhierarchy, namely the nodes between $A$ and $\underline{n}$.

be possible to find a smaller cyclic closed subhierarchy, as illustrated in Figure 4b, violating the fact that $H'$ is minimal. Thus, nodes must not be comparable across paths and $H'$ is a UNP. The second subcase is that $H'$ is not a between set. Lemma 4 in the Appendix shows that then every cycle in $H'$ must contain every node in $N'$. Any such *spanning cycle* can be decomposed into two undirected paths between a maximal node and a minimal node. If any nodes in these two paths were comparable, we could find a smaller cycle that does not contain every node in $N'$, which would contradict Lemma 4. As a result, $H'$ is a UNP.

Step 3: Finally, we show that UNPs and crowns are not universally constructible. The former is demonstrated in Lemma 6 in the Appendix using a similar construction as we used for the diamond in Section 3. Indeed, the diamond is an example of a UNP in which there are exactly two undirected paths from $\overline{n}$ to $\underline{n}$.

For the crown, we construct a particular monotone experiment allocation that is not constructible: Let $\Omega = \{0, 1, 2\}$ and $\mu_0 (\omega = 0) = \frac{1}{2}$, $\mu_0 (\omega = 1) = \frac{1}{4}$, and $\mu_0 (\omega = 2) = \frac{1}{4}$. We represent each belief as a pair $(x, y)$ in the triangle $\{(x, y) \mid x \geq 0, y \geq 0, x + y \leq 1\}$, where $Pr (\omega = 1) = x$ and $Pr (\omega = 2) = y$. Consider the experiment allocation $\beta$ that assigns to $F$, $M$, $S$, and $D$ the experiments indicated in Figure 5a. Each belief realization is a circle, the letters inside the circle

14

Figure 5: The crown is not universally constructible



(a) The experiments.

(b) $D$ to $M$.

(c) $S$ to $M$.

(d) $D$ to $F$.

(e) $S$ to $F$.
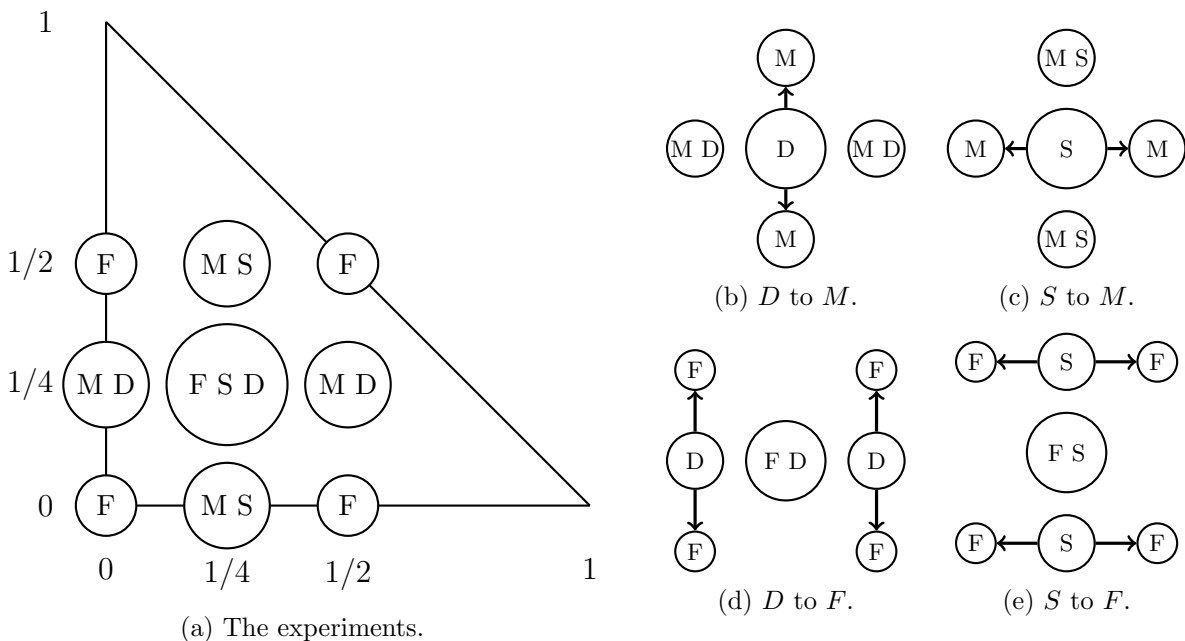
indicate the nodes for which this belief is in the support, and the area of the circle is the likelihood of the belief (which is the same for all nodes attached to the belief). For example, the belief $(1/4, 1/4)$ can be realized for $F$, $S$, and $D$, with likelihood of $1/2$ in each case. It is easy to see from Figure 5 that $\beta$ assigns higher nodes more informative signals. For example, going from $D$ to $M$ (Figure 5b), the mass on the east and west beliefs stays the same, but the mass on the central belief is spread out in the north-south directions.

Lemma 7 in the Appendix formally defines this experiment allocation and shows that it is not constructible. To see what goes wrong, suppose there is a monotone signal allocation that induces $\beta$. Such a signal allocation induces a joint distribution over beliefs at all four nodes. Consider first the joint distribution of beliefs between $D$, $M$, and $S$. The center $D$ belief occurs if and only if the $M$ belief is north or south (Figure 5b). The center belief of $S$ occurs if and only if the $M$ belief is east or west (Figure 5c). Hence, the probability of both $D$ and $S$ having the center belief is zero. Now consider the joint distribution of beliefs between $D$, $F$, and $S$. The belief at $D$ is in the center if and only if the belief at $F$ is in the center (Figure 5d), and the belief at $S$ is in the center if and only if the belief at $F$ is in the center (Figure 5e). So, the probability that both $D$ and $S$ have the center belief is positive, yielding a contradiction.
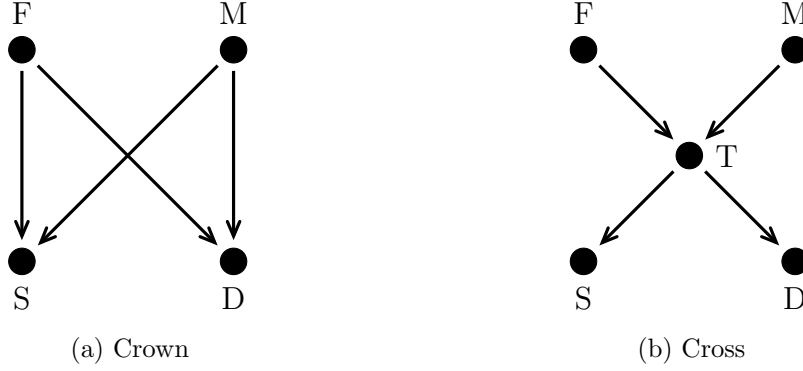
(a) Crown            (b) Cross

Figure 6: The crown and cross hierarchies.

# 5  Discussion

## 5.1  Binary states

A key step in our proof of the only-if direction of Theorem 1 shows that the crown is not universally constructible. Specifically, Figure 5 presents a particular experiment allocation involving three states that is not constructible.

It is natural to ask whether there is an example of a non-constructible allocation on the crown using two states. No such example exists: when the state space is binary, every monotone experiment allocation on the crown is constructible. The reason is as follows. Take an arbitrary monotone experiment allocation $\beta$ on $\{F, M, S, D\}$. It is known that the set of experiments under the Blackwell order is a lattice when the state space is binary (Kertz and Rösler, 2000; Müller and Scarsini, 2006). Thus, there exists an experiment $\beta(S) \vee \beta(D)$ such that $\beta(S) \vee \beta(D) \succsim \beta(S)$, $\beta(S) \vee \beta(D) \succsim \beta(D)$, and $\tau \succsim \beta(S) \vee \beta(D)$ for any $\tau \succsim \beta(S), \beta(D)$. We can therefore expand the crown hierarchy by adding a $T$(utor) whom $F$(ather) and $M$(other) hired to oversee the children: $F \geq T$, $M \geq T$, $T \geq S$, and $T \geq D$. We refer to this hierarchy as the *cross*, which is depicted in Figure 6b. Let $\hat{\beta}$ be the experiment allocation on the cross that sets $\hat{\beta}(n) = \beta(n)$ for $n \in \{F, M, S, D\}$ and $\hat{\beta}(T) = \beta(S) \vee \beta(D)$. The allocation $\hat{\beta}$ is monotone by construction. The cross is a forest, and hence by Theorem 1, it is universally constructible. Thus, there is a monotone signal allocation $\hat{\sigma}$ that induces $\hat{\beta}$. Restricting $\hat{\sigma}$ to $\{F, M, S, D\}$ yields a monotone signal allocation on the crown that induces $\beta$.

This example illustrates that the binary state space is special. With three or more states, every

monotone experiment allocation is constructible if and only if the hierarchy is a forest. With binary states, being a forest is sufficient for constructibility, but it is not necessary. A natural direction for future research is to characterize necessary and sufficient conditions for every monotone experiment allocation to be constructible for the case of binary states.

## 5.2 Beyond the refinement order

As we mentioned at the outset, there are various things one might mean by "$A$ is more informed than $B$." One is that $A$'s signal is Blackwell more informative than $B$'s, i.e., $\langle \pi_A \rangle \succsim \langle \pi_B \rangle$. Another is that $A$ has observed all of $B$'s information, i.e., $\pi_A \unrhd \pi_B$. These, however, are not the only economically relevant comparisons of signals. For instance, it might be that $A$ knows $B$'s belief about the state. Or, it might be that if $A$ observed $B$'s information, $A$'s belief about the state of the world would not change, i.e., that $\pi_B$ is statistically redundant given $\pi_A$. In a companion paper (Brooks *et al.*, 2020), we explore these and other relations on signals in more detail. In this section, we explain how our results extend to a large class of notions of "more informed," including those just mentioned.

First, consider the belief-martingale relation on $\Pi$, denoted $\mathcal{B}$, defined by $\pi \, \mathcal{B} \, \pi'$ if $\mathbb{E}\left[\tilde{\mu}_\pi | \tilde{\mu}_{\pi'}\right] = \tilde{\mu}_{\pi'}$. (We use the word relation rather than order because $\mathcal{B}$ is not transitive.)[14] Note that the refinement order implies the belief-martingale relation, which in turn implies the Blackwell order:

$$\pi \unrhd \pi' \Rightarrow \pi \, \mathcal{B} \, \pi' \Rightarrow \langle \pi \rangle \succsim \langle \pi' \rangle.$$

Many economically relevant comparisons of signals are "in between" refinement and the belief-martingale relation, i.e., they are weaker than the former and stronger than the latter. For example, the two aforementioned comparisons—$A$ knowing $B$'s belief or $B$'s signal being statistically redundant given $A$'s—fall in this in-between category. Formally, we say a binary relation $\mathcal{R}$ on $\Pi$ is *proper* if $\pi \unrhd \pi' \Rightarrow \pi \, \mathcal{R} \, \pi' \Rightarrow \pi \, \mathcal{B} \, \pi'$.

Our main result can be extended by replacing the refinement order with any proper relation. In particular, say a signal allocation $\sigma$ on $H$ is $\mathcal{R}$-*monotone* if $n \geq n'$ implies that $\sigma(n) \, \mathcal{R} \, \sigma(n')$. The experiment allocation $\beta$ on $H$ is $\mathcal{R}$-*constructible* if $\beta$ is induced by some $\mathcal{R}$-monotone signal

---

[14]See Brooks *et al.* (2020).

allocation on $H$. A hierarchy is said to be $\mathcal{R}$-*universally constructible* if every monotone experiment allocation is $\mathcal{R}$-constructible for any $\Omega$ and $\mu_0$. With this terminology, we have the following:

**Theorem 2.** *Fix any proper relation $\mathcal{R}$. A hierarchy $H$ is $\mathcal{R}$-universally constructible if and only if its graph is a forest.*

The if direction follows immediately from Theorem 1, since $\unrhd$-universal constructibility implies $\mathcal{R}$-universal constructibility for any proper $\mathcal{R}$. For the only if direction, the key is to recognize that the proofs of Lemmas 6 and 7 in the Appendix establish not only that UNPs and crowns are not $\unrhd$-universally constructible, but also that they are not $\mathcal{B}$-universally constructible. With that change, the remainder of the proof of the only if direction of Theorem 1 (replacing every instance of "universally constructible" with "$\mathcal{B}$-universally constructible") establishes that if $G(H)$ is not a forest, it is not $\mathcal{B}$-universally constructible. A fortiori, it is not $\mathcal{R}$-universally constructible for any proper $\mathcal{R}$.

# 6 Applications

## 6.1 Rationalizing reaction to unknown sources of information

Consider an agent who obtains information from multiple sources. If we do not know the information-generating process, what restrictions does the agent's rationality impose on her potential reactions to this information? Concretely, suppose a decision maker has access to a set of information sources $\{x_1, x_2, ..., x_M\}$. Suppose further that our dataset $\mathcal{D} = \{\tau_S\}_{S \subseteq \{x_1, ..., x_M\}}$ tells us the distribution of beliefs conditional on observing any non-empty subset of information sources. When can we rationalize a given dataset $\mathcal{D}$ in the sense that we can associate each information source $x_i$ with some signal (i.e., an element of $\Pi$) and conclude that belief formation is consistent with Bayes' rule?

To be rationalized, belief distributions (i.e., experiments) in $\mathcal{D}$ have to satisfy two obvious properties. First, there is *Bayes plausibility*: the average belief cannot differ across sets of information sources, i.e., $\mathbb{E}_{\tau_S}[\mu] = \mathbb{E}_{\tau_{S'}}[\mu]$ for any two subsets $S$ and $S'$. Second, there is *Blackwell monotonicity*: observing a superset of sources necessarily induces a more dispersed distribution of beliefs, i.e.,
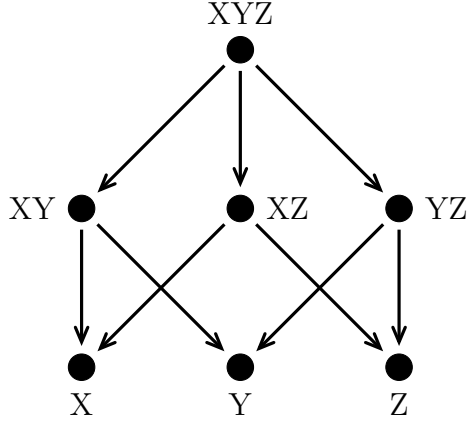
Figure 7: The set-inclusion hierarchy.

$\tau_S$ is a mean-preserving spread of $\tau_{S'}$ if $S' \subseteq S$. A natural question is whether these are the only properties imposed by Bayesian updating.

Theorem 1 tells us that, when there are three or more sources, the answer is no.[15] In this case, Bayesian updating requires more than just Bayes plausibility and Blackwell monotonicity. To see why, consider the set-inclusion information hierarchy $H$ where each non-empty collection of sources $S \subseteq \{x_1, ..., x_M\}$ is associated with a node $n_S$ and the partial order is the superset order: $n_S \geq n_{S'}$ if $S' \subseteq S$. As illustrated in Figure 7, the graph of this information hierarchy is not a forest. By Theorem 1, this means that there is some monotone experiment allocation on $H$, call it $\beta$, that cannot be induced by any monotone signal allocation on $H$. Now, we can associate with this $\beta$ a dataset $\mathcal{D} = \{\tau_S\}_{S \subseteq \{x_1, ..., x_M\}}$ by setting $\tau_S = \beta(n_S)$. Note that $\mathcal{D}$ necessarily satisfies Bayes plausibility and Blackwell monotonicity (since $\beta$ is monotone). If we could rationalize $\mathcal{D}$ by associating each $x_i$ with some signal $\pi(x_i) \in \Pi$, then the signal allocation $\sigma(n_s) = \bigvee_{x_i \in S} \pi(x_i)$ would induce $\beta$ and be monotone (since $S' \subseteq S$ implies $\bigvee_{x_i \in S} \pi(x_i) \trianglerighteq \bigvee_{x_i \in S'} \pi(x_i)$), contradicting Theorem 1. Thus, we know that there are datasets that satisfy Bayes plausibility and Blackwell monotonicity, yet cannot be rationalized.

A fruitful direction for future research would be to fully characterize which reactions to unknown sources of information are rationalizable.

---

[15]When there are only two sources, $M = 2$, it is easy to show that the answer is indeed affirmative. Any reaction to two unknown sources of information that satisfies Bayes plausibility and Blackwell monotonicity is consistent with Bayesian updating.

## 6.2 Information design

Suppose Sender provides information to a set of agents. Moreover, she face certain types of mono-tonicity constraints such as some agents must know the beliefs of some other agents or some agents must have access to others' information. Then, we can think of agents as elements of an information hierarchy, and the information design problem consists of selecting a (suitably monotone) signal allocation on this hierarchy. Our results shed light on how such monotonicity constraints affect the information design problem.

For example, consider organizations. A long literature in organization economics emphasizes the importance of the hierarchical structure of managerial relationships (Williamson, 1967). One important aspect of organizational design is deciding how much information to provide—about individuals' prospects for promotion, about the overall performance of the organization, etc.—to each member of the organization. It is often suboptimal to provide full transparency and share full information with everyone (Fuchs, 2007; Jehiel, 2015; Smolin, 2017).

A natural constraint that an information designer might face is that anyone in the organization ought to have access to the information that is available to her subordinates, i.e., that the allocation of signals within the organization should be monotone with respect to the management structure, so that the signal of a superior refines the signals of their subordinates. This constraint interacts with the organization structure. Theorem 1 implies that, if an organization has the feature that every subordinate has at most one superior, the aforementioned constraint can always be satisfied as long as individuals who are higher up in the organization are more informed in the Blackwell sense. Furthermore, if the information designer's objectives only depend on each agent's experiment, then one could reformulate the information design problem in terms of the choice of monotone experiment allocation, rather than choosing monotone signal allocations directly. With richer managerial relationships,[16] however, Theorem 1 also tells us there could be desirable allocations of information which are incompatible with the monotonicity constraint, even though they provide (Blackwell) more information to those higher up in the organization.

---

[16]For instance, suppose that the CEO oversees two middle managers who share the oversight of an employee.

# 7    Related literature

We now discuss the connection between our results and related literature in probability theory and in information economics.[17]

## 7.1    Connections to probability theory

Mathematically, our results are close to a literature in probability theory on the relationship between orders on probability distributions and stochastic processes. Because the connection to this literature is technical, our discussion is somewhat formal and detailed.

Fix two distributions $\tau, \tau' \in \Delta\left(\mathbb{R}^k\right)$. We say that $\tau$ *is greater than* $\tau'$ *in the convex order* if for every convex function $\phi : \mathbb{R}^k \to \mathbb{R}$, $\int \phi d\tau \geq \int \phi d\tau'$. The Sherman-Stein theorem says that the following are equivalent: (i) $\tau$ is greater than $\tau'$ in the convex order; (ii) $\tau$ can be obtained from $\tau'$ via mean-preserving spreads. Blackwell (1953, Section 3) observes that his result is implied by the Sherman-Stein theorem, since we can interpret $\tau$ as an experiment and the convex function $\phi$ as the maximum utility obtained in a decision problem given a belief. Importantly, the mean-preserving spread from $\tau$ to $\tau'$ is a probability transition kernel which defines a martingale on $\mathbb{R}^k$. Thus, an equivalent way of stating (ii) is that there exists a two-period martingale on $\mathbb{R}^k$ such that the marginals are $\tau$ and $\tau'$.

The equivalence was later generalized by Strassen (1965, Theorem 8): $\tau_1, \ldots, \tau_M \in \Delta\left(\mathbb{R}^k\right)$ are increasing in the convex order if and only if there exists an $M$-period martingale with marginals $\tau_1, \ldots, \tau_M$. This result is closely related to Green and Stokey (1978) and, more broadly, whether experiments can be represented as partitions: When the set of realizations of the martingale is finite, the martingale generates a filtration $(\Sigma_m)_{m=1}^M$, where each $\Sigma_m$ is generated by a finite partition $\pi_m$ of the underlying probability space, and the partitions are increasing in the refinement order. We can therefore view the probability space itself as an expanded state space, so that the $\pi_m$ are refinement-ordered signals that induce the given experiments. In our terminology, Strassen's theorem implies that the chain is universally constructible.

There is a substantial literature that further generalizes and strengthens Strassen's theorem.

---

[17]Our inquiry also leads us to a pure graph-theoretic question of whether a partially ordered set contains subsets of a particular form. This subject has been studied in combinatorics and graph theory (e.g., Lu, 2014); within economics, it is used by Curello and Sinander (2019) to study rankings on preference relations.

See Hirsch *et al.* (2011) for a survey.[18] To our knowledge, Juillet (2016) is the only paper in this literature to study collections of distributions that are indexed on partially ordered sets. He presents examples that illustrate various ways in which Strassen's theorem does not generalize. Most closely related to the present study, Juillet (2016, Section 4.1) constructs a collection $(\tau_x)_{x \in X}$, where $X$ is the diamond and $\tau_x \in \Delta(\mathbb{R})$, such that the collection is increasing in the convex order but there does not exist a real-valued martingale on $X$ with the given marginals. The distributions in Juillet's example are similar in spirit to (but distinct from) our example of a non-constructible belief allocation given in Section 3.[19] Our Theorem 1 implies that the equivalence between $(\tau_x)_{x \in X}$ increasing in the convex order and induced by a martingale holds if and only if $X$ is a forest.

Also related is Fill and Machida (2001), who study an analogue of the problem of Juillet (2016), but where the convex order is replaced by a generalization of first-order stochastic dominance. Specifically, Fill and Machida (2001) consider collections of distributions indexed on a partially ordered set $(\tau_x)_{x \in X}$, where each $\tau_x$ is a distribution on a different partially ordered set $S$. They define a notion of stochastic monotonicity for distributions on $S$ that reduces to first-order stochastic dominance when $S$ is totally ordered. A pair $(X, S)$ is *monotonicity equivalent* if the following condition holds: a collection is $(\tau_x)_{x \in X}$ is stochastically monotone if and only if there is a distribution over non-decreasing functions $f : X \to S$ such that $\tau_x$ is the marginal distribution of $f(x)$. Monotonicity equivalence is analogous to universal constructibility, but with stochastic monotonicity in place of the Blackwell order. In the special case that $S$ contains a sub-poset that is either a cycle or a crown, Fill and Machida (2001) show that $(X, S)$ is monotonicity equivalent if and only if $X$ is a forest. The cases where $X$ is a diamond or a crown also feature prominently in their proof, which follows a similar strategy of showing that if $X$ is not a forest, then it contains a sub-poset on which monotonicity equivalence fails, and then extending the counterexample to all of $X$.

## 7.2 Connections to information economics

Our paper connects to various threads in the economics literature studying information and beliefs. In addition to the aforementioned classical literature on information orders, Bergemann and Morris

---

[18]We thank Ludvig Sinander and Aniko Oery for drawing our attention to this literature.

[19]It is straightforward to adapt Juillet's example to show that the diamond is not universally constructible. But unlike our example from Section 3, his example is not easily adapted for general UNPs.

(2016) study an extension of the Blackwell order to type spaces, and Mu *et al.* (2019) consider comparisons of repeated experiments.

We contribute to the literature on higher order beliefs (Harsanyi, 1967; Mertens and Zamir, 1985; Brandenburger and Dekel, 1993). Specifically, we characterize when restrictions of the form "player $i$ knows player $j$'s type" place constraints on $i$'s and $j$'s first-order beliefs, beyond the obvious constraint that $i$ must be Blackwell more informed than $j$.[20]

We also study how signals can be combined to produce more informative signals. Gentzkow and Kamenica (2017) study this issue in the context of a communication game with a receiver who combines information provided by multiple senders. Börgers, Hernando-Veciana and Krähmer (2013) study the interaction between signals from the perspective of whether signals are substitutes or complements.

We also contribute to the growing literature on information design (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2016). Arieli *et al.* (2021) characterize feasible joint belief distributions of a group of agents in a binary state case. Mathevet and Taneva (2020) analyze the implications of information design for organizational structure.

## 8    Conclusion

We conclude with a brief discussion of an important direction for future work. Our analysis has focused on the question of when an information hierarchy is universally constructible. That condition requires that *every* monotone experiment allocation is constructible. Of course, even when not every monotone experiment allocation is constructible, some are. A natural goal would therefore be to characterize, for a given information hierarchy, the set of constructible monotone experiment allocations. Such a characterization would be of particular interest in the case of the set-inclusion hierarchy (Figure 7), which would identify the exact sufficient and necessary conditions for rationalizability of reactions to unknown sources of information, as discussed in Section 6.1.

---

[20]In the language of type spaces, "player $i$ knows player $j$'s type" is analogous to the refinement order. As discussed in Section 5.2, we also study relations that are weaker than refinement.

# References

ARIELI, I., BABICHENKO, Y., SANDOMIRSKIY, F. and TAMUZ, O. (2021). Feasible joint posterior beliefs. *Journal of Political Economy*, **129** (9).

BERGEMANN, D. and MORRIS, S. (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, **11** (2), 487–522.

BLACKWELL, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, pp. 265–272.

BÖRGERS, T., HERNANDO-VECIANA, A. and KRÄHMER, D. (2013). When are signals complements or substitutes? *Journal of Economic Theory*, **148** (1), 165–195.

BRANDENBURGER, A. and DEKEL, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, **59** (1), 189–198.

BROOKS, B., FRANKEL, A. and KAMENICA, E. (2020). Orders and relations on signals. In progress.

CURELLO, G. and SINANDER, L. (2019). *The Lattice Structure of Preference Comparison*. Tech. rep., University of Oxford and Northwestern University.

FILL, J. A. and MACHIDA, M. (2001). Stochastic monotonicity and realizable monotonicity. *Annals of probability*, pp. 938–978.

FUCHS, W. (2007). Contracting with repeated moral hazard and private evaluations. *American Economic Review*, **97** (4), 1432–1448.

GENTZKOW, M. and KAMENICA, E. (2017). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior*, **104**, 411–429.

GREEN, J. R. and STOKEY, N. L. (1978). Two representations of information structures and their comparisons.

HARSANYI, J. C. (1967). Games with incomplete information played by "bayesian" players, i–iii part i. the basic model. *Management science*, **14** (3), 159–182.

HIRSCH, F., PROFETA, C., ROYNETTE, B. and YOR, M. (2011). *Peacocks and associated martingales, with explicit constructions.* Springer Science & Business Media.

JEHIEL, P. (2015). Investment strategy and selection bias: An equilibrium perspective on overconfidence.

JUILLET, N. (2016). Peacocks parametrised by a partially ordered set. In *Séminaire de probabilités XLVIII*, Springer, pp. 13–32.

KAMENICA, E. and GENTZKOW, M. (2011). Bayesian persuasion. *American Economic Review*, **101** (6), 2590–2615.

KERTZ, R. P. and RÖSLER, U. (2000). Complete lattices of probability measures with applications to martingale theory. *Lecture Notes-Monograph Series*, pp. 153–177.

LU, L. (2014). On crown-free families of subsets. *Journal of Combinatorial Theory, Series A*, **126**, 216–231.

MATHEVET, L. and TANEVA, I. (2020). Optimal information hierarchies.

MERTENS, J.-F. and ZAMIR, S. (1985). Formulation of bayesian analysis for games with incomplete information. *International Journal of Game Theory*, **14** (1), 1–29.

MU, X., POMATTO, L., STRACK, P., TAMUZ, O. *et al.* (2019). Blackwell dominance in large samples. *arXiv preprint arXiv:1906.02838.*

MÜLLER, A. and SCARSINI, M. (2006). Stochastic order relations and lattices of probability measures. *SIAM Journal on Optimization*, **16** (4), 1024–1043.

SMOLIN, A. (2017). *Dynamic Evaluation Design.* Tech. rep., University Library of Munich, Germany.

STRASSEN, V. (1965). The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, **36** (2), 423–439.

WILLIAMSON, O. E. (1967). Hierarchical control and optimum firm size. *Journal of political economy*, **75** (2), 123–138.

# A Proofs

## A.1 Proof of Lemma 1

*Proof of Lemma 1.* Let $\tilde{\pi}$ be a signal s.t. $\langle\tilde{\pi}\rangle = \tau'$. Since $\langle\pi\rangle \succsim \tau'$, there exists a garbling $g : \pi\times\tilde{\pi} \to [0,1]$ such that $\sum_{\tilde{s}\in\tilde{\pi}} g(s,\tilde{s}) = 1 \ \forall s \in \pi$, and $p(\tilde{s}|\omega) = \sum_{s\in\pi} g(s,\tilde{s}) p(s|\omega)$. For every $\overline{s} \in \overline{\pi}$, let $\underline{s}(\overline{s})$ denote the element of $\pi$ s.t. $\overline{s} \subseteq \underline{s}(\overline{s})$. (This element exists since $\overline{\pi} \trianglerighteq \pi$.) Now, $\forall \overline{s} \in \overline{\pi}$, let $\{X_{\tilde{s}}^{\overline{s}}\}_{\tilde{s}\in M^{\overline{s}}}$ be a partition of $\overline{s}$ s.t. $\forall\omega$, $\lambda(\{x|(\omega,x) \in X_{\tilde{s}}^{\overline{s}}\}) = \lambda(\{x|(\omega,x) \in \overline{s}\}) g(\underline{s}(\overline{s}),\tilde{s})$, where $M^{\overline{s}} = \{\tilde{s} \in \tilde{\pi} | g(\underline{s}(\overline{s}),\tilde{s}) > 0\}$. Such a partition exists because $\sum_{\tilde{s}\in\tilde{\pi}} g(\underline{s}(\overline{s}),\tilde{s}) = 1$ for all $\underline{s}(\overline{s}) \in \pi$. Let $\pi' = \{Z^{\tilde{s}}\}_{\tilde{s}\in\tilde{\pi}}$ with $Z^{\tilde{s}} = \bigcup_{\overline{s}\in\overline{\pi} \text{ s.t.} \tilde{s}\in M^{\overline{s}}} X_{\tilde{s}}^{\overline{s}}$. We now show that $\pi'$ satisfies (i) and (ii). To show (i), it suffices to show that $p(Z^{\tilde{s}}|\omega) = p(\tilde{s}|\omega)$ for every $\tilde{s}$ and $\omega$. We have

$$
\begin{aligned}
p(Z^{\tilde{s}}|\omega) &= \lambda\left(\left\{x|(\omega,x) \in \bigcup_{\overline{s}\in\overline{\pi} \text{ s.t.}\tilde{s}\in M^{\overline{s}}} X_{\tilde{s}}^{\overline{s}}\right\}\right) \\
&= \sum_{\overline{s}\in\overline{\pi} \text{ s.t.}\tilde{s}\in M^{\overline{s}}} \lambda(\{x|(\omega,x) \in X_{\tilde{s}}^{\overline{s}}\}) \\
&= \sum_{\overline{s}\in\overline{\pi} \text{ s.t.}\tilde{s}\in M^{\overline{s}}} \lambda(\{x|(\omega,x) \in \overline{s}\}) g(\underline{s}(\overline{s}),\tilde{s}) \\
&= \sum_{\overline{s}\in\overline{\pi}} \lambda(\{x|(\omega,x) \in \overline{s}\}) g(\underline{s}(\overline{s}),\tilde{s}) \\
&= \sum_{s\in\pi} \sum_{\overline{s} \text{ s.t. } \underline{s}(\overline{s})=s} \lambda(\{x|(\omega,x) \in \overline{s}\}) g(\underline{s}(\overline{s}),\hat{s}) \\
&= \sum_{s\in\pi} g(s,\hat{s}) \sum_{\overline{s} \text{ s.t. } \underline{s}(\overline{s})=s} \lambda(\{x|(\omega,x) \in \overline{s}\}) \\
&= \sum_{s\in\pi} g(s,\tilde{s}) \lambda(\{x|(\omega,x) \in s\}) \\
&= \sum_{s\in\pi} g(s,\tilde{s}) p(s|\omega) \\
&= p(\tilde{s}|\omega).
\end{aligned}
$$

To show (ii), consider some $\hat{\pi}$ s.t. $\overline{\pi} \trianglerighteq \hat{\pi} \trianglerighteq \pi$ and some $\hat{s} \in \hat{\pi}$. Since $\hat{\pi} \vee \pi' \trianglerighteq \hat{\pi}$, there is a partition of $\hat{s}$, say $\{s_i^{\vee}\}_{i\in I}$ s.t. $s_i^{\vee} \in \hat{\pi} \vee \pi'$ for all $i$. It will suffice to show that for every $\omega$, $\omega'$, and $s_i^{\vee}$, we have

$$
\frac{p(s_i^{\vee}|\omega)}{p(s_i^{\vee}|\omega')} = \frac{p(\hat{s}|\omega)}{p(\hat{s}|\omega')}.
$$

Consider some $s_i^{\vee}$. Note that there exists $\underline{s} \in \pi$ with $\hat{s} \subseteq \underline{s}$ since $\hat{\pi} \trianglerighteq \pi$. Let $Q = \{\overline{s} \in \overline{\pi} | \overline{s} \subseteq \hat{s}\}$.

Since $\overline{\pi} \trianglerighteq \hat{\pi}$, for every $\omega$, $\lambda\left(x | \left(\omega, x\right) \in \hat{s}\right) = \sum_{\overline{s} \in Q} \lambda\left(x | \left(\omega, x\right) \in \overline{s}\right)$. Note that $\overline{s} \subseteq \underline{s}$ for all $\overline{s} \in Q$.

Now, we know that $s_i^{\vee} = s' \cap \hat{s}$ for some $s' \in \pi'$. By definition of $\pi'$, we know that $s' = \underset{\overline{s} \in \overline{\pi} \text{ s.t.} \tilde{s} \in M^{\overline{s}}}{\cup} X_{\tilde{s}}^{\overline{s}}$

for some $\tilde{s} \in \tilde{\pi}$. Hence,

$$
s_i^{\vee} = \left(\underset{\overline{s} \in \overline{\pi} \text{ s.t.} \tilde{s} \in M^{\overline{s}}}{\cup} X_{\tilde{s}}^{\overline{s}}\right) \cap \hat{s}
$$

$$
= \underset{\overline{s} \in \overline{\pi} \text{ s.t.} \tilde{s} \in M^{\overline{s}}}{\cup} \left(X_{\tilde{s}}^{\overline{s}} \cap \hat{s}\right)
$$

$$
= \underset{\overline{s} \in Q \text{ s.t.} \tilde{s} \in M^{\overline{s}}}{\cup} X_{\tilde{s}}^{\overline{s}},
$$

where the last equality follows from the fact that $X_{\tilde{s}}^{\overline{s}} \subseteq \overline{s}$, and hence $X_{\tilde{s}}^{\overline{s}} \cap \hat{s} = X_{\tilde{s}}^{\overline{s}}$ if $\overline{s} \in Q$ and $X_{\tilde{s}}^{\overline{s}} \cap \hat{s}$ is empty if $\overline{s} \notin Q$. Hence,

$$
p\left(s_i^{\vee} | \omega\right) = \lambda\left(\left\{x | \left(\omega, x\right) \in s_i^{\vee}\right\}\right)
$$

$$
= \lambda\left(\left\{x | \left(\omega, x\right) \in \underset{\overline{s} \in Q \text{ s.t.} \tilde{s} \in M^{\overline{s}}}{\cup} X_{\tilde{s}}^{\overline{s}}\right\}\right)
$$

$$
= \sum_{\overline{s} \in Q \text{ s.t.} \tilde{s} \in M^{\overline{s}}} \lambda\left(\left\{x | \left(\omega, x\right) \in X_{\tilde{s}}^{\overline{s}}\right\}\right)
$$

$$
= \sum_{\overline{s} \in Q \text{ s.t.} \tilde{s} \in M^{\overline{s}}} \lambda\left(\left\{x | \left(\omega, x\right) \in \overline{s}\right\}\right) g\left(\underline{s}, \tilde{s}\right)
$$

$$
= \sum_{\overline{s} \in Q} \lambda\left(\left\{x | \left(\omega, x\right) \in \overline{s}\right\}\right) g\left(\underline{s}, \tilde{s}\right)
$$

$$
= g\left(\underline{s}, \tilde{s}\right) \sum_{\overline{s} \in Q} \lambda\left(\left\{x | \left(\omega, x\right) \in \overline{s}\right\}\right)
$$

$$
= g\left(\underline{s}, \tilde{s}\right) \lambda\left(\left\{x | \left(\omega, x\right) \in \hat{s}\right\}\right)
$$

$$
= g\left(\underline{s}, \tilde{s}\right) p\left(\hat{s} | \omega\right).
$$

Hence,

$$
\frac{p\left(s_i^{\vee} | \omega\right)}{p\left(s_i^{\vee} | \omega'\right)} = \frac{g\left(\underline{s}, \tilde{s}\right) p\left(\hat{s} | \omega\right)}{g\left(\underline{s}, \tilde{s}\right) p\left(\hat{s} | \omega'\right)} = \frac{p\left(\hat{s} | \omega\right)}{p\left(\hat{s} | \omega'\right)},
$$

which completes the proof of Lemma 1. $\qquad\square$

## A.2 Proof of Theorem 1: If

Let $H$ be an information hierarchy and suppose $G(H)$ is a forest. Let $\beta$ be a monotone experiment allocation on $H$. We will construct a monotone signal allocation that induces $\beta$. To do so, we construct a sequence of subhierarchies of $H$, adding nodes of $H$ one by one, until we reach the full hierarchy $H$. At each step, we assign a signal to the newly added node and potentially reassign the signals allocated to the previously added nodes.

We begin with some notation and terminology. A *construction procedure* $f$ is a bijection from $\{1, ..., |N|\}$ to $N$ that specifies the order in which the nodes are added. Let $N_l^f = \{f(1), ..., f(l)\}$. If $f(l) = n$, we say that $n$ was added at time $l$, and we refer to $N_{l-1}^f$ as the previously added nodes. For any subset $N' \subseteq N$, let $CoveredBy(N') = \{n \in N \setminus N' | \exists n' \in N' \text{ that covers } n\}$, $Covering(N') = \{n \in N \setminus N' | \exists n' \in N' \text{ that is covered by } n\}$, and

$$Disconnected(N') = \{n \in N \setminus N' | \nexists n' \in N' \text{ s.t. there is a path from } n \text{ to } n' \text{ in } G(H)\}.$$

Now, consider a construction procedure $f$ of the following form. Let $f(1)$ be any node in $N$. For $l \in \{2, 3, ..., |N|\}$, let $f(l)$ be an arbitrary element of $CoveredBy\left(N_{l-1}^f\right) \cup Covering\left(N_{l-1}^f\right) \cup Disconnected\left(N_{l-1}^f\right)$. Note that for any $N' \subsetneq N$, $CoveredBy(N') \cup Covering(N') \cup Disconnected(N')$ is not empty.

*Claim 1.* For each $l \geq 2$, there is at most one edge in $G(H)$ between $f(l)$ and nodes in $N_{l-1}^f$.

*Proof of Claim 1.* Suppose toward contradiction that $f(l)$ has an edge in $G(H)$ with distinct $n, n' \in N_{l-1}^f$. Since $n$ and $n'$ both have an edge with $f(l)$, they must belong to the same tree in $G(H)$. Moreover, there must be a path between $n$ and $n'$ in $G\left(\left(N_{l-1}^f, \geq\right)\right)$. To see this, let $\underline{n}$ be the node that was added earliest to $N_{l-1}^f$ among the nodes in the tree to which $n$ and $n'$ belong. For every other node $f(l') \in N_{l-1}^f$ from this tree, we must have $f(l') \in CoveredBy\left(N_{l'-1}^f\right) \cup Covering\left(N_{l'-1}^f\right)$, which in turn means that there is a path from $f(l')$ to $\underline{n}$ in $G\left(\left(N_{l'-1}^f, \geq\right)\right)$ and thus in $G\left(\left(N_{l-1}^f, \geq\right)\right)$. Hence, there is a path from both $n$ and $n'$ to $\underline{n}$ and thus a path between $n$ and $n'$ in $G\left(\left(N_{l-1}^f, \geq\right)\right)$. So, there must be a path between $n$ and $n'$ in $G(H)$ that does not go through $f(l)$. But, because $f(l)$ has an edge with both $n$ and $n'$, there is another path from $n$ to $n'$ that goes through $f(l)$. However, $G(H)$ is a forest, so there cannot be multiple paths between

two nodes; we have reached a contradiction. □

Now, given this construction procedure $f$, we assign signals to nodes as follows. At step $l$, we expand $N_{l-1}^f$ to $N_l^f = N_{l-1}^f \cup f(l)$ and assign signals according to $\sigma^l : N_l^f \to \Pi$. We proceed by induction and show that, as long as the signals previously allocated to nodes in $N_{l-1}^f$ are monotone on $\left(N_{l-1}^f, \geq\right)$ and induce appropriate experiments (i.e., for all $m \in N_{l-1}^f$, $\langle \sigma^{l-1}(m) \rangle = \beta(m)$), then the $\sigma^l$ we specify is monotone on $\left(N_l^f, \geq\right)$ induces appropriate experiments.

First, to node $f(1)$, we assign an arbitrary signal $\sigma^1(f(1))$ such that $\langle \sigma^1(f(1)) \rangle = \beta(f(1))$. Note we are vacuously satisfying the base case of the induction argument: the signal allocation to the single node in $N_1^f$ induces appropriate experiments. For $l \geq 2$, there are three cases: $f(l) \in CoveredBy\left(N_{l-1}^f\right)$, $f(l) \in Covering\left(N_{l-1}^f\right)$, and $f(l) \in Disconnected\left(N_{l-1}^f\right)$.

We first consider the case $f(l) \in Covering\left(N_{l-1}^f\right)$. Note that, by Claim 1, $f(l)$ covers exactly one node in $N_{l-1}^f$ (call this node $\bar{m}$) and is not covered by any nodes in $N_{l-1}^f$. Since $\beta(f(l)) \succsim \beta(\bar{m})$, there exists some $\pi \trianglerighteq \sigma^{l-1}(\bar{m})$ such that $\langle \pi \rangle = \beta(f(l))$ (which follows from Theorem 1 of Green and Stokey (1978)). We set $\sigma^l(f(l)) = \pi$ and we keep the signal allocation for nodes in $N_{l-1}^f$ unchanged, i.e., $\sigma^l(m) = \sigma^{l-1}(m)$ for all $m \in N_{l-1}^f$. It is clear that $\sigma^l$ induces appropriate experiments (by the inductive hypothesis for $m \in N_{l-1}^f$ and by construction for $f(l)$). We also need to show that this signal allocation is monotone on $N_l^f$. Consider any $m, m' \in N_l^f$ such that $m > m'$. Since $f(l) \in Covering\left(N_{l-1}^f\right)$, either $m, m' \in N_{l-1}^f$ or $f(l) = m$. In the former case, we know $\sigma^l(m) = \sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(m') = \sigma^l(m')$ by the inductive hypothesis. If $f(l) = m$, we know $f(l) > \bar{m} \geq m'$. By the inductive hypothesis, $\sigma^l(\bar{m}) = \sigma^{l-1} \trianglerighteq \sigma^{l-1}(m') = \sigma^l(m')$ and thus $\sigma^l(f(l)) \trianglerighteq \sigma^l(\bar{m}) \trianglerighteq \sigma^l(m')$. That completes the proof for this case.

Now consider the case where $f(l) \in CoveredBy\left(N_{l-1}^f\right)$. Let $\underline{m}$ be the node in $N_{l-1}^f$ that covers $f(l)$. Denote $\tau' = \beta(f(l))$, $\pi = \sigma^{l-1}(\underline{m})$, and $\bar{\pi} = \bigvee_{m \in N_{l-1}^f} \sigma^{l-1}(m)$. By Lemma 1, we know $\exists \pi'$ such that (i) $\langle \pi' \rangle = \tau'$, and (ii) $\forall \hat{\pi}$ s.t. $\bar{\pi} \trianglerighteq \hat{\pi} \trianglerighteq \pi$, $\langle \hat{\pi} \vee \pi' \rangle = \langle \hat{\pi} \rangle$. We set $\sigma^l(f(l)) = \pi'$. For $m \in N_{l-1}^f$, if $m \geq f(l)$, we set $\sigma^l(m) = \sigma^{l-1}(m) \vee \pi'$; otherwise, we set $\sigma^l(m) = \sigma^{l-1}(m)$. We need to show that $\sigma^l$ is monotone on $\left(N_l^f, \geq\right)$ and induces appropriate experiments. We have that $\langle \sigma^l(f(l)) \rangle = \langle \pi' \rangle = \tau' = \beta(f(l))$. For $m \in N_{l-1}^f$, first consider cases where $m \geq f(l)$, so $\langle \sigma^l(m) \rangle = \langle \sigma^{l-1}(m) \vee \pi' \rangle$. Since $m \geq \underline{m}$ (recall that $\underline{m}$ covers $f(l)$), by the inductive hypothesis, $\sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(\underline{m}) = \pi$; moreover, $\bar{\pi} = \bigvee_{m' \in N_{l-1}^f} \sigma^{l-1}(m') \trianglerighteq \sigma^{l-1}(m)$;

29

hence, $\langle \sigma^l(m) \rangle = \langle \sigma^{l-1}(m) \vee \pi' \rangle = \langle \sigma^{l-1}(m) \rangle$. For $m \in N_{l-1}^f$ s.t. $m \not\geq f(l)$, $\langle \sigma^l(m) \rangle = \langle \sigma^{l-1}(m) \rangle$.

Since by the inductive hypothesis, $\langle \sigma^{l-1}(m) \rangle = \beta(m)$, we have established that $\langle \sigma^l(m) \rangle = \beta(m)$

for all $m \in N_{l-1}^f$. We now need to show that $\sigma^l$ is monotone. Consider any $m, m' \in N_{l-1}^f$ s.t.

$m \geq m'$. There are three cases. First, suppose $m \geq m' \geq f(l)$. In that case, we know that

$\sigma^l(m) = \sigma^{l-1}(m) \vee \pi'$ and $\sigma^l(m') = \sigma^{l-1}(m') \vee \pi'$. Since (by the inductive hypothesis) $\sigma^{l-1}(m) \unrhd$

$\sigma^{l-1}(m')$, we know that $\sigma^{l-1}(m) \vee \pi' \unrhd \sigma^{l-1}(m') \vee \pi'$, and hence $\sigma^l(m) \unrhd \sigma^l(m')$. The second case

is where $m \geq f(l)$ and $m' \not\geq f(l)$. Then, $\sigma^l(m) = \sigma^{l-1}(m) \vee \pi'$ and $\sigma^l(m') = \sigma^{l-1}(m')$. Since (by

the inductive hypothesis) $\sigma^{l-1}(m) \unrhd \sigma^{l-1}(m')$, we have that $\sigma^l(m) = \sigma^{l-1}(m) \vee \pi' \unrhd \sigma^{l-1}(m) \unrhd$

$\sigma^{l-1}(m') = \sigma^l(m')$. Finally, suppose that $m \not\geq f(l)$ and $m' \not\geq f(l)$. Then, $\sigma^l(m) = \sigma^{l-1}(m)$

and $\sigma^l(m') = \sigma^{l-1}(m')$. Since (by the inductive hypothesis) $\sigma^{l-1}(m) \unrhd \sigma^{l-1}(m')$, we have that

$\sigma^l(m) \unrhd \sigma^l(m')$.

Finally, suppose $f(l) \in Disconnected\left(N_{l-1}^f\right)$. We assign an arbitrary signal $\sigma^l(f(l))$ to $f(l)$

such that $\langle \sigma^l(f(l)) \rangle = \beta(f(l))$, and we keep the signal allocation to nodes in $N_{l-1}^f$ unchanged, i.e.,

$\sigma^l(m) = \sigma^{l-1}(m)$ for all $m \in N_{l-1}^f$. It is clear that $\sigma^l$ induces appropriate experiments (by the

inductive hypothesis for $m \in N_{l-1}^f$ and by construction for $f(l)$). Since $f(l)$ is not comparable to

any node in $N_{l-1}^f$, the fact that $\sigma^{l-1}$ is a signal allocation on $\left(N_{l-1}^f, \geq\right)$ implies that $\sigma^l$ is monotone

on $\left(N_l^f, \geq\right)$. This completes the proof. $\square$

## A.3    Proof of Theorem 1: Only if

**Lemma 2.** *If an information hierarchy $H$ is universally constructible and $H'$ is a closed subhierarchy*

*of $H$, then $H'$ is universally constructible.*

*Proof of Lemma 2.* Suppose $H$ is an information hierarchy that is universally constructible under

$\mathcal{B}$. Fix a state space $\Omega$ and a prior $\mu_0$. Suppose $H' = (N', \geq)$ is a closed subhierarchy of $H$. Let

$\beta'$ be a monotone experiment allocation on $H'$. We need to construct a monotone signal allocation

on $H'$ that induces $\beta'$. Let $\beta : N \to \Delta(\Delta(\Omega))$ defined as follows: (i) if $n \in N'$, let $\beta(n) = \beta'(n)$;

(ii) if $n \notin N'$ and $\exists n' \in N'$ such that $n > n'$, let $\beta(n) = \bar{\tau}$; and (iii) if $n \notin N'$ and $\nexists n' \in N'$ such

that $n > n'$, let $\beta(n) = \underline{\tau}$.

*Claim 2.* $\beta$ is monotone on $H$.

*Proof of Claim 2.* Consider $n, n' \in N$ with $n \geq n'$. We show that $\beta(n) \succsim \beta(n')$ by considering four exhaustive cases:

If $n$ and $n'$ are both in $N'$, this follows from the fact that $\beta'$ is monotone on $H'$.

If $n$ and $n'$ are both not $N'$, consider two subcases. If $\exists n'' \in N'$ such that $n > n''$, then $\beta(n) = \bar{\tau} \succsim \beta(n')$. Otherwise, since $n \geq n'$ and $\nexists n'' \in N'$ such that $n > n''$, it must be that $\nexists n'' \in N'$ such that $n' > n''$, so $\beta(n) \succsim \beta(n') = \underline{\tau}$.

If $n \notin N'$ and $n' \in N'$, $\beta(n) = \bar{\tau} \succsim \beta(n')$.

Finally, if $n \in N'$ and $n' \notin N'$, then there cannot exist an $n'' \in N'$ with $n' > n''$. If such an $n''$ did exist, then since $H'$ is closed and $n, n'' \in N'$, we would have that $n' \in N'$, a contradiction. Thus, $\beta(n) \succsim \beta(n') = \underline{\tau}$. $\diamondsuit$

Since $\beta$ is a monotone experiment allocation on $H$ and $H$ is universally constructible, there exists a monotone signal allocation $\sigma$ on $H$ that induces $\beta$. Clearly, the restriction of $\sigma$ to $N'$ induces $\beta'$ and is monotone on $H'$. $\square$

Recall that $\overrightarrow{G}(H)$ is the directed graph associated with the hierarchy $H$ (see footnote 9). The next result shows that if $H'$ is a closed subhierarchy of $H$, then $\overrightarrow{G}(H')$ is the subgraph of $\overrightarrow{G}(H)$ obtained by dropping edges with nodes that are not in $H'$.

**Lemma 3.** *Fix a hierarchy $H = (N, \geq)$ and a closed subhierarchy $H' = (N', \geq)$. Let $\overrightarrow{E}$ be the set of edges in $\overrightarrow{G}(H)$. Then $\overrightarrow{G}(H') = \left(N', \overrightarrow{E'}\right)$, where $\overrightarrow{E'} = \left\{(n, n') \in \overrightarrow{E} \mid n, n' \in N'\right\}$.*

*Proof of Lemma 3.* Fix $n, n' \in N'$. We need to show that $(n, n') \in \overrightarrow{E'}$ if and only if $(n, n') \in \overrightarrow{E}$. If $(n, n') \in \overrightarrow{E}$, then $n$ covers $n'$ in $H$, i.e., $n > n'$ and there is no $n'' \in N$ with $n > n'' > n'$. A fortiori, there is no $n'' \in N' \subset N$ with $n > n'' > n'$; hence, $n$ covers $n'$ in $H'$, so $(n, n') \in \overrightarrow{E'}$. If $(n, n') \in \overrightarrow{E'}$, then $n$ covers $n'$ in $H'$. As a result, $n > n'$. If there exists $n'' \in N$ such that $n > n'' > n'$, then, since $H'$ is closed, $n'' \in N'$, so $n$ must not cover $n'$ in $H'$, a contradiction. As a result, $n$ covers $n'$ in $H$ as well, so $(n, n') \in \overrightarrow{E}$. $\square$

The *between set of $(n, n')$ in $H = (N, \geq)$* is defined as

$$Btw(n, n', H) = \left\{\hat{n} \in N \mid n \geq \hat{n} \geq n'\right\}.$$

Clearly, the subhierarchy induced by the between set of any pair of nodes is closed. Moreover, a subhierarchy $H' = (N', \geq)$ is closed if and only if $N'$ contains $Btw(n, n', H)$ for all $n, n' \in N'$. We say that $Btw(n, n', H)$ is *simple* if every node in $Btw(n, n', H) \setminus \{n, n'\}$ belongs to exactly one directed path in $\overrightarrow{G}(H)$ from $n$ to $n'$. $H'$ is a *minimal cyclic closed subhierarchy (MCC)* of $H$ if it is cyclic, closed, and there is no cyclic and closed subhierarchy $H'' = (N'', \geq)$ of $H$ with $N'' \subsetneq N'$. We say that a cycle in $G(H')$ is a *spanning cycle* if every node in $N'$ is in the cycle.

**Lemma 4.** *Fix a hierarchy $H = (N, \geq)$ and a subhierarchy $H' = (N', \geq)$. Suppose $H'$ is an MCC of $H$. Then, either (i) $N'$ is a simple between set in $H$, or (ii) every cycle in $G(H')$ is a spanning cycle.*

*Proof of Lemma 4.* For this proof, all between sets are defined relative to $H$, and we simply write $Btw(n, n')$ for $Btw(n, n', H)$. Similarly, by closed we mean closed in $H$. Note that by Lemma 3, $\overrightarrow{G}(H')$ is the subgraph of $\overrightarrow{G}(H)$ obtained by dropping edges with nodes that are not in $H'$. In particular, $H'$ is cyclic if and only if $H$ contains a cycle whose nodes are in $H'$. This fact is used freely below.

We consider two cases. First, suppose there are two nodes $n, n' \in N'$ such that $n \geq n'$ and there are two distinct paths from $n$ to $n'$ in the directed graph $\overrightarrow{G}(H)$. Note that $Btw(n, n')$ is closed, so that these paths are in $\overrightarrow{G}((Btw(n, n'), \geq))$ as well, so that $(Btw(n, n'), \geq)$ is cyclic. Moreover, since $H'$ is closed, we have that $Btw(n, n') \subseteq N'$. Hence, since $H'$ is an MCC, we must have that $N' = Btw(n, n')$. It remains to show that the between set $N'$ is simple. Suppose to the contrary there is some node $\hat{n} \in N' \setminus \{n, n'\}$ such that $\hat{n}$ belongs to two distinct paths from $n$ to $n'$ in $\overrightarrow{G}(H)$. Then, there are either two distinct directed paths from $n$ to $\hat{n}$ or two distinct directed paths from $\hat{n}$ to $n'$; thus, either $Btw(n, \hat{n})$ or $Btw(\hat{n}, n')$ must be cyclic. Since both $Btw(n, \hat{n})$ and $Btw(\hat{n}, n')$ are closed and strict subsets of $N'$, $H'$ must not be an MCC, so we have reached a contradiction. Thus, we have established that $N'$ must be a simple between set.

Now consider the second case where for every $n, n' \in N'$, there is at most one path from $n$ to $n'$ in the directed graph $\overrightarrow{G}(H)$. Given a path $P$ (either directed or undirected), let $N_P$ denote the set of nodes that appear in $P$. Since $H'$ is an MCC, $G(H') = (N', E)$ contains a cycle $C = (n_0, e_0, ... n_{L-1}, e_{L-1}, n_L)$ where $L > 1$, $n_0 = n_L$. We will argue that $(N_C, \geq)$ is closed. The fact that $N_C = N'$ will then follow directly from the hypothesis that $H'$ is an MCC.

Let us then suppose that $(N_C, \geq)$ is not closed, in order to reach a contradiction. Given a directed path $(n_0, \overrightarrow{e}_0, ..., n_{L-1}, \overrightarrow{e}_{L-1}, n_L)$, its *undirected analogue* is the undirected path $(n_0, e_0, ...n_{L-1}, e_{L-1}, n_L)$ where $e_i = \{n_i, n_{i+1}\}$. Say that a directed path $P$ in $\overrightarrow{G}(H)$ *only contains edges in $C$* if every edge in the undirected analogue of $P$ is in $C$. A directed path $P$ in $\overrightarrow{G}(H)$ is an *external directed connection* (EDC) from $n$ to $n'$ if (i) $P$ is a directed path from $n$ to $n'$; (ii) $n, n' \in N_C$; and (iii) $P$ does not only contain edges in $C$. Say that $(n, n') \in N_C$ are an *externally connected pair* (ECP) if there is an external directed connection from $n$ to $n'$ or from $n'$ to $n$. An ECP $(n_i, n_j)$ is said to be *minimally close* if for every $i \leq \underline{l} < \bar{l} \leq j$, $(n_{\underline{l}}, n_{\bar{l}})$ is an ECP only if $\underline{l} = i$ and $\bar{l} = j$.

*Claim 3.* Given any two nodes $n, n' \in N'$, if $P$ is the unique directed path in $\overrightarrow{G}(H)$ from $n$ to $n'$, then $N_P = Btw(n, n')$.

*Proof of Claim 3.* If there are two non-comparable nodes in $Btw(n, n')$, there would be two distinct directed paths from $n$ to $n'$. Hence, all nodes in $Btw(n, n')$ are comparable. Therefore, there is a directed path from $n$ to $n'$ whose nodes are $Btw(n, n')$. Since there is a unique directed path from $n$ to $n'$, the set of nodes in $P$ is $Btw(n, n')$. $\diamondsuit$

*Claim 4.* There exist $i, j$ such that $(n_i, n_j)$ is a minimally close ECP.

*Proof of Claim 4.* We know there is a pair of nodes in $N_C$ that are an ECP. Otherwise, $(N_C, \geq)$ would be closed. Moreover, since $L$ is finite, there is a pair of nodes in $N_C$ that are a minimally close ECP. $\diamondsuit$

Let $(n_i, n_j)$ be a minimally close ECP s.t. $\{n_i, n_{i+1}, ..., n_j\} \subsetneq N_C$. Let $\bar{n} = \max\{n_i, n_j\}$ and $\underline{n} = \min\{n_i, n_j\}$. Let $P^e$ denote the external directed connection from $\bar{n}$ to $\underline{n}$. Let $\tilde{P}$ be the undirected path $(n_i, \tilde{e}_i, ..., \tilde{e}_{j-1}, n_j)$ in $G(H)$ from $n_i$ to $n_j$ in $C$. Let $\tilde{Q}$ denote the undirected path from $n_i$ to $n_j$ that "goes in the other direction" from $\tilde{P}$ in $C$, i.e., $\tilde{Q} = (n_i, e_{i-1}, n_{i-1}, ..., e_0, n_0, e_{L-1}, n_{L-1}, ..., e_j, n_j)$. Let $S = N_{\tilde{P}} \cup Btw(\bar{n}, \underline{n})$.

*Claim 5.* $(S, \geq)$ is cyclic.

*Proof of Claim 5.* It suffices to show there are two distinct undirected paths from $n_i$ to $n_j$ in $G((S, \geq))$. One path is $\tilde{P}$. The other path is the undirected analogue of the external directed connection $P^e$. Since $P^e$ is external, these two undirected paths must be distinct. $\diamondsuit$

*Claim* 6. $S$ is closed.

*Proof of Claim 6.* Let $Y = \cup_{n,n' \in N_{\tilde{P}}} Btw(n, n')$. We will show that $Y$ is closed and that $Y = S$.

First we show that $Y$ is closed. Consider any $n', n'' \in Y$ and $n \in Btw(n', n'')$. By definition of $Y$, $n' \in Btw(n_1, n_2)$ and $n'' \in Btw(n_3, n_4)$, where $n_l \in N_{\tilde{P}}$ for $l = 1, 2, 3, 4$. Hence, $n_1 \geq n' \geq n \geq n'' \geq n_4$ and thus $n \in Btw(n_1, n_4) \subseteq Y$.

It remains to show that $S = Y$. Given $n \in N_{\tilde{P}}$, $n \in Btw(n, n) \subseteq Y$. Moreover, $Btw(\overline{n}, \underline{n}) \subseteq Y$. Hence, $S \subseteq Y$.

Now, consider some $n \in Y$. We need to show that $n \in S$. If $n \in N_{\tilde{P}}$, then we are done. Otherwise, $n \notin N_{\tilde{P}}$. We know $n \in Btw(n', n'')$ for some $n', n'' \in N_{\tilde{P}}$. If $(n', n'') = (\overline{n}, \underline{n})$, $n \in Btw(\overline{n}, \underline{n}) \subseteq S$. Suppose instead that $(n', n'') \neq (\overline{n}, \underline{n})$. We will reach a contradiction. Let $P$ denote the directed path from $n'$ to $n''$ whose nodes include $n$. Because $(\overline{n}, \underline{n})$ is a minimally close ECP, path $P$ must only include edges in $C$. Since $n \notin N_{\tilde{P}}$, the nodes in path $P$ cannot be a subset of $N_{\tilde{P}}$. Thus, the nodes in $P$ contain the nodes in $\tilde{Q}$, including $\overline{n}$ and $\underline{n}$. The sequence of nodes and edges in $P$ between $\overline{n}$ and $\underline{n}$ is a directed path between those nodes, and thus is equal to $P^e$ (by uniqueness of the directed path). Since $P^e$ contains an edge which is not in $C$, we have contradicted the hypothesis that $P$ only contains edges in $C$. Thus, we have established that $Y \subseteq S$. $\diamond$

We have established that $(S, \geq)$ is cyclic and closed and that $S \subseteq N'$. Since $H'$ is an MCC, it must be that $S = N'$. But since $S = N_{\tilde{P}} \cup Btw(\overline{n}, \underline{n})$, it must be that $N_{\tilde{Q}} \subseteq Btw(\overline{n}, \underline{n})$. All nodes in a between set are comparable, by the hypothesis that directed paths are unique, and so all nodes in $N_{\tilde{Q}}$ are comparable. Hence, $\tilde{Q}$ must be the undirected analogue of $P^e$. This contradicts the hypothesis that $(\overline{n}, \underline{n})$ is an ECP. $\square$

**Lemma 5.** *If the hierarchy $H = (N, \geq)$ is such that every cycle in $G(H)$ is a spanning cycle, then for any pair of nodes $n, n' \in N$, there exist two undirected paths from $n$ to $n'$ such that the union of the nodes in the two paths is $N$ and the intersection of the nodes in the two paths is $\{n, n'\}$.*

*Proof of Lemma 5.* Since there exists a spanning cycle $G(H)$, for any pair of nodes $n, n' \in N$, there exist two paths $P = \left(n, e_0^P, n_1^P ..., n_{L^P-1}^P, e_{L^P-1}^P, n'\right)$ and $Q = \left(n, e_0^Q, n_1^Q ..., n_{L^Q-1}^Q, e_{L^Q-1}^Q, n'\right)$ in $G(H)$ such that $N = N_P \cup N_Q$ and $E_P \cap E_Q = \emptyset$, where $E_P = \left\{e_0^P, ..., e_{L^P-1}^P\right\}$ and $E_Q = \left\{e_0^Q, ..., e_{L^Q-1}^Q\right\}$. We need to show that $N_P \cap N_Q = \{n, n'\}$. Suppose to the contrary that there

exists $\hat{n} \in N_P \cap N_Q$ with $\hat{n} \notin \{n, n'\}$. We know there exist $l^P \in \{1, ..., L^P - 1\}$ and $l^Q \in \{1, ..., L^Q - 1\}$ such that $\hat{n} = n_{l^P}^P = n_{l^Q}^Q$. Now, consider the path $\left(n, e_0^P, n_1^P, ..., e_{l^P-1}^P, n_{l^P}^P, e_{l^Q-1}^Q, n_{l^Q-1}^Q, e_{l^Q-2}^Q, ..., n\right)$ in $G(H)$. Since $E_P \cap E_Q = \emptyset$, this is a well-defined path. But it is a cycle that is not spanning and thus we have reached a contradiction. $\qquad \square$

Suppose a hierarchy $H$ has nodes

$$N = \{\overline{n}\} \cup \{\underline{n}\} \cup_{l=1}^{L} \left\{n_k^l\right\}_{k=1}^{K_l},$$

with $L \geq 2$, such that for all $l = 1, \ldots, L$, (i) $K_l \geq 1$, (ii) $\overline{n} > n_1^l$, (iii) $n_{K_l}^l > \underline{n}$, (iv) $n_k^l$ and $n_{k+1}^l$ are comparable for every $k = 1, \ldots, K_l - 1$, and (v) $n_k^l$ is not comparable to $n_{k'}^{l'}$ if $l \neq l'$. In this case, we say that $H$ is a *union of non-comparable paths (UNP)*.

**Lemma 6.** *If the hierarchy $H$ is a UNP, then $H$ is not universally constructible.*

*Proof of Lemma 6.* Let $\Omega = \{0, 1\}$. We will present a monotone experiment allocation based on the allocation on the diamond in Section 3 and show that it is not constructible. Consider the experiment allocation given by[21] $\beta(\underline{n}) = \left(\frac{1}{2}, \left(\frac{1}{6}\right); \frac{1}{2}\left(\frac{5}{6}\right)\right)$; $\beta(\overline{n}) = \left(\frac{3}{8}, (0); \frac{1}{4}, \left(\frac{1}{2}\right); \frac{3}{8}, (1)\right)$, $\beta(n_k^1) = \left(\frac{3}{8}, (0); \frac{1}{8}, \left(\frac{2}{3}\right); \frac{1}{2}, \left(\frac{5}{6}\right)\right)$ for $k = 1, \ldots, K_1$; and $\beta(n_k^l) = \left(\frac{1}{2}, \left(\frac{1}{6}\right); \frac{1}{8}, \left(\frac{1}{3}\right); \frac{3}{8}, (1)\right)$ for all $l = 2, \ldots, L$ and $k = 1, \ldots, K_l$. These are the experiments depicted in Figure 2, where $\beta(\overline{n}) = \tau_A$; $\beta(\underline{n}) = \tau_D$; for $k = 1, \ldots, K_1$, let $\beta(n_k^1) = \tau_B$; and for all $l = 2, \ldots, L$ and $k = 1, \ldots, K_l$, let $\beta(n_k^l) = \tau_C$. Monotonicity of $\beta$ is immediate from the mean-preserving spreads depicted in Figure 2.[22]

Toward a contradiction, suppose $\sigma$ is monotone signal allocation that induces $\beta$. We begin by establishing that, it must be that case that, along any undirected path from $\overline{n}$ to $\underline{n}$, the realized beliefs must be equal across the interior nodes on that path.

*Claim 7.* For all $l = 1, \ldots, L$, we have $\tilde{\mu}_{\sigma(n_k^l)} = \tilde{\mu}_{\sigma(n_1^l)}$ for all $k = 1, \ldots, K_l$.

*Proof of Claim 7.* It suffices to establish that for any $k = 1, ..., K_l - 1$, we have $\tilde{\mu}_{\sigma(n_{k+1}^l)} = \tilde{\mu}_{\sigma(n_k^l)}$. We know $n_k^l$ and $n_{k+1}^l$ are comparable. Assume that $n_{k+1}^l > n_k^l$; the case where $n_k^l > n_{k+1}^l$ is

---

[21]The experiment is given as a sequence of ordered pairs of a probability and a belief.

[22]To see that $\tau_B$ is a mean-preserving spread of $\tau_D$, note that we can obtain $\tau_B$ from $\tau_D$ by spreading the realization $\mu = \frac{5}{6}$ in $\tau_D$ to $\{\frac{1}{3}, 1\}$ in $\tau_B$ and leaving the realization $\mu = \frac{1}{6}$ in $\tau_D$ unchanged. To see that $\tau_A$ is a mean-preserving spread of $\tau_B$, note that we can obtain $\tau_A$ from $\tau_B$ by spreading the realizations $\mu = \frac{1}{6}$ and $\mu = \frac{1}{3}$ in $\tau_B$ to $\{0, \frac{1}{2}\}$ in $\tau_A$ and leaving the realization $\mu = 1$ in $\tau_B$ unchanged. The argument for why $\tau_A \succsim \tau_C \succsim \tau_D$ is symmetric.

analogous and omitted. Since $\sigma$ is monotone, we have that $\sigma\left(n_{k+1}^l\right) \trianglerighteq \sigma\left(n_k^l\right)$. As the refinement order implies the belief-martingale property we have that $\mathbb{E}[\tilde{\mu}_{\sigma(n_{k+1}^l)}|\tilde{\mu}_{\sigma(n_k^l)}] = \tilde{\mu}_{\sigma(n_k^l)}$. Now suppose toward contradiction that $\tilde{\mu}_{\sigma(n_{k+1}^l)} \neq \tilde{\mu}_{\sigma(n_k^l)}$. We would conclude that $\tilde{\mu}_{\sigma(n_{k+1}^l)}$ is a strict mean-preserving spread of $\tilde{\mu}_{\sigma(n_k^l)}$.[23] But that would mean that $\tilde{\mu}_{\sigma(n_{k+1}^l)}$ is not equal to $\tilde{\mu}_{\sigma(n_k^l)}$ in distribution and thus that $\langle \sigma\left(n_{k+1}^l\right)\rangle \neq \langle \sigma\left(n_k^l\right)\rangle$. $\diamond$

We will now reach a contradiction by establishing that $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0 \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right)$ is equal to zero and is strictly bigger than zero.

Step 1: We show that $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0 \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) = 0$. First, note that $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0|\tilde{\mu}_{\sigma(n_1^1)} = 0\right) = 1$ since $\mathbb{E}\left[\tilde{\mu}_{\sigma(\overline{n})}|\tilde{\mu}_{\sigma(n_1^1)} = 0\right] = 0$ and the support of $\tilde{\mu}_{\sigma(\overline{n})}$ lies above 0. Moreover, since $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0\right) = \frac{3}{8} = Pr\left(\tilde{\mu}_{\sigma(n_1^1)} = 0\right)$, we must also have $Pr\left(\tilde{\mu}_{\sigma(n_1^1)} = 0|\tilde{\mu}_{\sigma(\overline{n})} = 0\right) = 1$. A similar argument establishes that $Pr\left(\tilde{\mu}_{\sigma\left(n_{K_1}^1\right)} = \frac{5}{6}|\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) = 1$ and $Pr\left(\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}|\tilde{\mu}_{\sigma\left(n_{K_1}^1\right)} = \frac{5}{6}\right) = 1$. Claim 7 tells us that $Pr\left(\tilde{\mu}_{\sigma(n_1^1)} = 0 \ \& \ \tilde{\mu}_{\sigma(n_{K_1}^1)} = \frac{5}{6}\right) = 0$. Hence, $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0 \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) = 0$.

Step 2: We show that $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0 \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) > 0$. It suffices to show that (a) $Pr\left(\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}|\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) > 0$, and (b) $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3} \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) > 0$.

Arguments analogous to the ones in Step 1 yield $Pr\left(\tilde{\mu}_{\sigma\left(n_{K_2}^2\right)} = \frac{1}{6}|\tilde{\mu}_{\sigma(\underline{n})} = \frac{1}{6}\right) = 1$. Because $Supp\left(\tilde{\mu}_{\sigma(\underline{n})}\right) = \left\{\frac{1}{6}, \frac{5}{6}\right\}$, this in turn implies that $Pr\left(\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}|\tilde{\mu}_{\sigma\left(n_{K_2}^2\right)} = \frac{1}{3}\right) = 1$. Thus, Claim 7 tells us that $Pr\left(\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right) = 1$. Therefore $Pr\left(\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}|\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) = \frac{Pr\left(\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right)}{Pr\left(\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right)} > 0$, establishing part (a).

Now, note that $Pr\left(\tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right) = 1$ implies $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3} \ \& \ \tilde{\mu}_{\sigma(\underline{n})} = \frac{5}{6}\right) = Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right)$. Moreover, since $Supp\left(\tilde{\mu}_{\sigma(\overline{n})}\right) = \left\{0, \frac{1}{2}, 1\right\}$, we have $\mathbb{E}\left[\tilde{\mu}_{\sigma(\overline{n})}|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right] = \frac{1}{3} \Rightarrow Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0|\tilde{\mu}_{\sigma\left(n_1^2\right)} = \frac{1}{3}\right) > 0$, establishing (b). $\square$

**Lemma 7.** *The crown is not universally constructible.*

*Proof.* Consider the experiment allocation given by $\beta\left(F\right) = \left(\frac{1}{8}, \left(0, \frac{1}{2}\right); \frac{1}{8}, \left(\frac{1}{2}, \frac{1}{2}\right); \frac{1}{8}, \left(\frac{1}{2}, 0\right); \frac{1}{8}, \left(0, 0\right); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right)\right)$, $\beta\left(M\right) = \left(\frac{1}{4}, \left(0, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{2}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{4}, 0\right); \frac{1}{4}, \left(\frac{1}{4}, \frac{1}{4}\right)\right)$, $\beta\left(S\right) = \left(\frac{1}{4}, \left(\frac{1}{4}, 0\right); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{4}, \frac{1}{2}\right)\right)$, and $\beta\left(D\right) = \left(\frac{1}{4}, \left(0, \frac{1}{4}\right); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{2}, \frac{1}{4}\right)\right)$. These experiments are depicted in Figure 5. To see that $\beta\left(F\right)$ is a mean-preserving spread of $\beta\left(S\right)$, note that we can obtain $\beta\left(F\right)$ from $\beta\left(S\right)$ by (i) spreading the

---

[23]For any two random variables $X$ and $Y$, if $\mathbb{E}\left[X|Y\right] = Y$ and $X \neq Y$, $X$ must be a strict mean-preserving spread of $Y$.

realization $\left(\frac{1}{4}, 0\right)$ in $\beta\left(S\right)$ to $\left\{\left(0, 0\right), \left(\frac{1}{2}, 0\right)\right\}$ in $\beta(F)$, (ii) leaving the realization $\left(\frac{1}{4}, \frac{1}{4}\right)$ in $\beta(S)$ unchanged, and (iii) spreading the realization $\left(\frac{1}{4}, \frac{1}{2}\right)$ in $\beta\left(S\right)$ to $\left\{\left(0, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right)\right\}$. To see that $\beta\left(M\right)$ is a mean-preserving spread of $\beta\left(S\right)$, note that we can obtain $\beta\left(M\right)$ from $\beta\left(S\right)$ by (i) leaving the realization $\left(\frac{1}{4}, 0\right)$ in $\beta\left(S\right)$ unchanged, (ii) spreading the realization $\left(\frac{1}{4}, \frac{1}{4}\right)$ in $\beta\left(S\right)$ to $\left\{\left(0, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{4}\right)\right\}$ in $\beta(M)$, and (iii) leaving the realization $\beta\left(S\right) = \left(\frac{1}{4}, \frac{1}{2}\right)$ unchanged. The argument for why $\beta\left(F\right)$ and $\beta\left(M\right)$ are mean-preserving spreads of $\beta\left(D\right)$ is symmetric. Hence, $\beta$ is monotone.

We claim that $\beta$ is not constructible. As before, consider the joint distribution of beliefs on $\{F, M, S, D\}$ induced by any monotone signal allocation $\sigma$, and recall that $\sigma(n) \trianglerighteq \sigma(n')$ implies $\mathbb{E}\left[\tilde{\mu}_{\sigma(n)} | \tilde{\mu}_{\sigma(n')}\right] = \tilde{\mu}_{\sigma(n')}$. Specifically, consider the conditional probability of $\tilde{\mu}_{\sigma(S)} = \mu_0$ given $\tilde{\mu}_{\sigma(D)} = \mu_0$. Since $\sigma\left(M\right) \trianglerighteq \sigma\left(S\right)$, we have that $\tilde{\mu}_{\sigma(S)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(M)} \in \left\{\left(0, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{4}\right)\right\}$. Since $\sigma\left(M\right) \trianglerighteq \sigma\left(D\right)$, we have that $\tilde{\mu}_{\sigma(D)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(M)} \in \left\{\left(\frac{1}{4}, 0\right), \left(\frac{1}{4}, \frac{1}{2}\right)\right\}$. Hence, the joint probability of $\tilde{\mu}_{\sigma(S)} = \mu_0$ and $\tilde{\mu}_{\sigma(D)} = \mu_0$ must be zero, i.e., $Pr\left(\tilde{\mu}_{\sigma(S)} = \mu_0 | \tilde{\mu}_{\sigma(D)} = \mu_0\right) = 0$. But, $\sigma\left(F\right) \trianglerighteq \sigma\left(S\right)$ implies $\tilde{\mu}_{\sigma(S)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(F)} = \mu_0$, and $\sigma\left(F\right) \trianglerighteq \sigma\left(D\right)$ implies $\tilde{\mu}_{\sigma(D)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(F)} = \mu_0$. Hence, $Pr\left(\tilde{\mu}_{\sigma(S)} = \mu_0 | \tilde{\mu}_{\sigma(D)} = \mu_0\right) = 1$, a contradiction. $\qquad\square$

**Lemma 8.** *If the hierarchy $H = (N, \geq)$ is cyclic and $N$ is a simple between set in $H$, then $H$ is not universally constructible.*

*Proof of Lemma 8.* Since $H = (N, \geq)$ is cyclic and $N$ is a simple between set, $H$ must be a UNP. Therefore Lemma 6 establishes that $H$ is not universally constructible. $\qquad\square$

**Lemma 9.** *If the hierarchy $H$ is cyclic and not a crown, and every cycle in $G\left(H\right)$ is a spanning cycle, then $H$ is not universally constructible.*

*Proof of Lemma 9.* Suppose that $H = (N, \geq)$ is cyclic and not a crown, and that every cycle in $G\left(H\right)$ is a spanning cycle. We say an element $\overline{n}$ of $N$ is maximal if there is no $n \in N$ such that $n > \overline{n}$. We say an element $\underline{n}$ of $N$ is minimal if there is no $n \in N$ such that $\underline{n} > n$.

*Claim 8.* There exist $\overline{n}$ and $\underline{n}$ in $N$ that are maximal and minimal, respectively, such that $\overline{n}$ does not cover $\underline{n}$.

*Proof of Claim 8.* Suppose toward contradiction that every maximal element covers every minimal element. Hence, $G\left(H\right)$ is a complete bipartite graph of maximal and minimal elements. If there

were only one maximal element or only one minimal element, there could not be a cycle. So, there must be at least two of each. Take any $N' \subseteq N$ consisting of exactly two maximal and two minimal elements, and let $H' = (N', \geq)$. $H'$ is a crown and therefore is cyclic. Moreover, $H'$ is clearly closed, so that by Lemma 3, the nodes in $H'$ are part of a cycle in $G(H)$ as well. Since every cycle in $G(H)$ is a spanning cycle, we must have $N = N'$, and thus $H$ is a crown subhierarchy, and we have reached a contradiction. $\diamondsuit$

By Claim 8, we can find $\bar{n}$ and $\underline{n}$ in $N$ that are maximal and minimal, respectively, such that $\bar{n}$ does not cover $\underline{n}$. By Lemma 5, there are two distinct undirected paths $P$ and $Q$ in $G(H)$ from $\bar{n}$ to $\underline{n}$ such that the union of the nodes in the two paths is $N$ and the intersection of the nodes in the two paths is $\{\bar{n}, \underline{n}\}$. As a result, $H$ is a UNP. Lemma 6 therefore implies that $H$ is not universally constructible. $\square$

*Proof of Theorem 1: Only if.* Suppose $G(H)$ is not a forest, i.e., it contains a cycle. Since $N$ is finite, $H$ contains a subhierarchy $H' = (N', \geq)$ that is an MCC of $H$. By Lemma 4, either (i) $N'$ is a simple between set in $H$, or (ii) every cycle in $G(H')$ is a spanning cycle. Consider case (i). Because $H'$ is closed and $N'$ is a simple between set in $H$, $N'$ is also a simple between set in $H'$. Thus, Lemma 8 implies that $H'$ is not universally constructible. Now consider case (ii). If $H'$ is not a crown, then Lemma 9 implies that is not universally constructible. If $H'$ is a crown, Lemma 7 implies it is not universally constructible. Therefore, by Lemma 2, $H$ is not universally constructible. $\square$