# Information Hierarchies

Benjamin Brooks, Alexander Frankel, and Emir Kamenica*

June 2020

**Abstract**

If Anne knows more than Bob about the state of the world, she may or may not know what Bob thinks, but it is always possible that she does. In other words, if the distribution of Anne's belief about the state is a mean-preserving spread of the distribution of Bob's belief, we can construct signals for Anne and Bob that induce these distributions of beliefs and provide Anne with full information about Bob's belief. We establish that with more agents, the analogous result does not hold. It might be that Anne knows more than Bob and Charles, who in turn both know more than David, yet what they know about the state precludes the possibility that Anne knows what Bob and Charles think and that everyone knows what David thinks. More generally, we define an *information hierarchy* as a partially ordered set and ask whether higher elements being Blackwell more informed always makes the hierarchy compatible with higher elements having more information (under various notions of that term) than lower elements. We show that the answer is affirmative if and only if the graph of the hierarchy is a forest. We discuss applications of this result to rationalizing a decision maker's reaction to unknown sources of information and to information design in organizations.

*JEL classification*: C70; D82; D83; D85
*Keywords*: Information, beliefs, signals, Blackwell order, networks

*Brooks: Department of Economics, University of Chicago, babrooks@uchicago.edu. Frankel: Booth School of Business, University of Chicago, afrankel@uchicago.edu. Kamenica: Booth School of Business, University of Chicago, emir@uchicago.edu. We thank Ozan Candogan, Eddie Dekel, Simone Galperti, Stephen Morris, Xiaosheng Mu, Ludvig Sinander, Alex Smolin, and Bruno Strulovici for helpful comments. We thank Oguz Bayraktar for research assistance. Kamenica thanks the Neubauer Family Faculty Fellowship for research support.

# 1 Introduction

There are various things we might mean when we say, "Anne is more informed than Bob." We might mean that Anne's information about some state of the world is more accurate than Bob's (in the sense of Blackwell 1951). Or, we might mean that Anne knows everything that Bob knows. Of course, the latter implies the former: if Anne knows everything that Bob knows, her information is necessarily more accurate. Moreover, if Anne's information is more accurate, then, whatever the extent of Anne's and Bob's knowledge about the state, it is always possible that Anne knows everything that Bob knows. In formal terms, if the distribution of Anne's belief about the state[1] is a mean-preserving spread of the distribution of Bob's belief, we can always construct signals for Anne and Bob that induce these belief distributions such that Anne's signal is a refinement of Bob's signal. This fundamental result means that informational comparisons of belief distributions can often be interpreted as comparisons of signals.

In this paper, we explore the relationship between various notions of "more informed" in the presence of many agents. For each notion, we ask whether being more informed imposes the same restrictions on the distribution of beliefs about the state as the Blackwell order does. Perhaps surprisingly, we find that the observations from the previous paragraph do not extend to more than three agents. We construct an example where Anne's information is more accurate than Bob's and Charles's, whose information is in turn more accurate than David's, and yet it cannot be that Anne knows everything that Bob and Charles know and that all three know everything that David knows. Formally, we construct four distributions of posteriors, $\tau_A$, $\tau_B$, $\tau_C$, and $\tau_D$ such that $\tau_A$ is a mean-preserving spread of $\tau_B$ and $\tau_C$, which are in turn mean-preserving spreads of $\tau_D$, and show there do not exist four signals $\pi_A$, $\pi_B$, $\pi_C$, $\pi_D$ that induce these four distributions of posteriors and have the property that $\pi_A$ refines $\pi_B$, $\pi_C$, and $\pi_D$, and $\pi_B$ and $\pi_C$ refines $\pi_D$. It is always the case that a less-informed person cannot know what the more-informed people think; in our example, a more-informed person cannot know what less-informed people think.

To examine this issue in full generality, we introduce the notion of an *information hierarchy.* An information hierarchy is simply a partially ordered set. We consider allocations of distributions

[1]Throughout the paper, we use the terms "belief about the state" and "first-order belief" interchangeably, which – when there is no risk of confusion with higher-order beliefs – we abbreviate to simply "belief."

of beliefs to the elements of the hierarchy that are *monotone*, meaning that higher elements have more accurate information in the sense of Blackwell. Given such a belief allocation, we ask whether it is *constructible under refinement*, i.e., whether it can be induced by a refinement-monotone signal allocation in which higher elements know the signal realizations of lower elements. If every monotone belief allocation is constructible in this way, we say that the information hierarchy is *universally constructible under refinement*.

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] Our main theorem shows that an information hierarchy is universally constructible under refinement if and only if its undirected graph is a forest, i.e., there is at most one path between any two elements. Thus, if the undirected graph is a forest, requiring that higher elements' signals refine lower elements' signals implies no additional restrictions on beliefs than that higher nodes' beliefs are Blackwell more informative than lower nodes' beliefs. If the undirected graph is not a forest, then refinement does imply additional restrictions on beliefs relative to the Blackwell order.

The aforementioned four-person example entails an information hierarchy whose graph is not a forest: there are two paths from Anne to David, one through Bob and one through Charles. In contrast, the undirected graph in the two-person example, from the first paragraph, is a forest.[3] This is precisely why in the four-person but not in the two-person example, it was possible for the extent of knowledge about the state to preclude an individual with more accurate information from knowing everything known by the less informed.

We also consider a number of other notions of "more informed," which are weaker than the requirement that Anne knows everything that Bob knows (refinement), but stronger than the mere fact that Anne's signal is more accurate than Bob's (the Blackwell order). For instance, it may be that Anne knows Bob's belief about the state but does not know all of Bob's information. Or, it may be that Bob's information does not add to Anne's information—her belief about the state would not change if she were to observe his information. These are examples of a class that we call *proper relations*, which are binary relations on signals that are weaker than refinement and satisfy

[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] Indeed, the graph of any information hierarchy with three or fewer elements is necessarily a forest.

a natural belief-martingale property.[4] Our theorem establishes that, given any proper relation, an information hierarchy is universally constructible under that relation if and only if the undirected graph of the hierarchy is a forest.

The proof of the "if" direction of the theorem is relatively straightforward, though it does require establishing a novel information-theoretic result that might be of independent interest (see Lemma 2). Under the hypothesis that the undirected graph is a forest, we use this result to iteratively construct a refinement-monotone signal allocation inducing any given monotone belief allocation. The "only if" direction is considerably more involved. The proof relies on three key ideas: First, we show that a hierarchy is universally constructible only if its *closed subhierarchies* are also universally constructible. (By closed, we mean that all of the elements in the hierarchy that are between two elements of the subhierarchy are also in the subhierarchy.) Second, we show that any hierarchy that is not a forest must contain a closed subhierarchy taking one of two forms: either its undirected graph is a *crown* or it is a *union of non-comparable paths.* The latter can be seen as a generalization of the four-person example described above. Finally, for both of these subhierarchy forms, we present monotone belief allocations that are not constructible.

We discuss two applications of our Theorem. First, suppose there is an agent who observes information from several sources. Consider an econometrician who does not know the data generating process behind the information sources, but observes the agent's reaction to the information. In particular, the econometrician sees the distribution of the agent's beliefs given any subset of information sources. We know these reactions must satisfy two simple properties. Bayes plausibility requires that the agent's average belief is equal across the subsets. Blackwell monotonicity requires that when an agent observes a superset of sources, her beliefs are more dispersed than when she observes a subset. A natural conjecture might be that these two are the only constraints imposed by Bayesian rationality. Our Theorem implies that this conjecture is false.

Second, we apply the Theorem to constrained information design. For instance, take a Sender who selects what information to provide to members of an organization but must ensure that managers have access to all information provided to their subordinates. Our result identifies the types of organizational structures under which this requirement is strictly more binding than the

[4] The property says that the expectation of Anne's belief given Bob's belief is equal to Bob's belief.

weaker requirement that higher-ups should have more accurate information.

In a broad sense, this paper explores the relationship between different orders on experiments, building on Blackwell (1951). Relatedly, Bergemann and Morris (2016) study an extension of the Blackwell order to type spaces. Mu *et al.* (2019) consider comparisons of repeated experiments.

One way to frame our contribution is as a study of the implications that the common prior assumption imposes on the universal type space (Mertens and Zamir, 1985; Brandenburger and Dekel, 1993). In particular, given a set of players, we examine whether a collection of restrictions of the form "player $i$ knows player $j$'s type" places constraints on $i$'s and $j$'s first-order beliefs (beyond the obvious constraint that $i$ must be Blackwell more informed than $j$).[5]

We 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.* (2020) 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.

Finally, our inquiry 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.

The rest of the paper proceeds as follows. Section 2 describes our model of information hierarchies. Section 3 presents several examples of hierarchies and discusses which ones are universally constructible; it also presents examples of proper relations. Section 4 presents our main result and a sketch of the proof. Section 5 discusses the applications of our results. Section 6 briefly concludes. All omitted proofs are in the Appendix.

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

# 2 Set-up

## 2.1 States, signals, and distributions of posteriors

Given 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).[6] 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. Let $p(s|\omega) = \lambda(\{x|(\omega,x) \in s\})$ and $p(s) = \sum_{\omega\in\Omega} p(s|\omega)\mu_0(\omega)$, where $\lambda(\cdot)$ denotes the Lebesgue measure. That is, $p(s|\omega)$ is the conditional probability of $s$ given $\omega$ and $p(s)$ is the unconditional probability of $s$. We denote the set of all signals by $\Pi$.

A *distribution of beliefs*, denoted by $\tau$, is an element of $\Delta\Delta\Omega$ that has finite support and satisfies $\mathbb{E}_\tau[\mu] = \mu_0$. We partially order distributions of beliefs by informativeness in the sense of Blackwell (1951) and write $\tau \succsim \tau'$ if $\tau$ is a mean-preserving spread of $\tau'$. We let $\overline{\tau}$ denote the maximally informative distribution of beliefs (whose support contains only degenerate beliefs) and $\underline{\tau}$ denote the minimally informative distribution of beliefs (that puts probability one on the prior).

Observing a signal realization $s$ s.t. $p(s) > 0$ generates a unique posterior belief $\mu_s$, where the probability of $\omega$ given $s$ is[7]

$$\mu_s(\omega) = \frac{p(s|\omega)\mu_0(\omega)}{p(s)}.$$

Each signal $\pi$ induces a distribution of posteriors denoted by $\langle\pi\rangle$, according to

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

We denote the refinement order on $\Pi$ by $\trianglerighteq$, i.e., given $\pi, \pi' \in \Pi$, we write $\pi \trianglerighteq \pi'$ if every element of $\pi$ is a subset of some element of $\pi'$. 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 $\pi'$. For any set $P$, we denote the join of all its

[6] Throughout we assume $|\Omega| \geq 2$.

[7] For those $s$ with $p(s) = 0$, set $\mu_s$ to be an arbitrary belief.

elements by $\bigvee P$.

Given signal $\pi$, let $\tilde{\mu}_\pi$ denote the belief-valued random variable that reflects the posterior induced by the observation of the signal realization from $\pi$. We define a belief-martingale relation on $\Pi$, denoted $\pi \mathcal{B} \pi'$ if $\mathbb{E}\left[\tilde{\mu}_\pi | \tilde{\mu}_{\pi'}\right] = \tilde{\mu}_{\pi'}$.[8] (We use the word relation rather than order because this relation is not transitive.)[9]

Note that the refinement order implies the belief-martingale relation, which in turn implies the Blackwell order on the induced distributions of posteriors; i.e.,

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

We say a binary relation $\mathcal{R}$ on $\Pi$ is *proper* if $\pi \trianglerighteq \pi' \Rightarrow \pi \, \mathcal{R} \, \pi' \Rightarrow \pi \, \mathcal{B} \, \pi'$. We provide some examples of proper relations in Section 3.3.

## 2.2 Information hierarchies

An *information hierarchy* $H$ is a finite partially ordered set $(N, \geq)$ with the corresponding strict order $>$. Since we will heavily rely on graph-theoretic representations of $(N, \geq)$, we refer to elements of $N$ as *nodes*. Nodes $n$ and $n'$ are *comparable* if $n \geq n'$ or $n' \geq n$. Given $n, n' \in N$, $n$ *covers* $n'$ if $n > n'$ and there does not exist $n'' \in N$ with $n > n'' > n'$.

The *directed graph* of $H$, denoted $G(H)$, is the pair $(N, E)$, where $N$ is the set of nodes, $E \subseteq N \times N$ is the set of directed edges, and $(n, n') \in E$ if $n$ covers $n'$. A *directed path* from $n$ to $n'$ in $G(H)$ is an alternating sequence of vertices and directed 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 $l \neq l' \Rightarrow e_l \neq e_{l'}$. The *undirected graph* of $H$, denoted $\tilde{G}(H)$, is the pair $\left(N, \tilde{E}\right)$, where $N$ is the set of nodes, $\tilde{E} \subseteq \{\tilde{e} \subseteq N' | \, |\tilde{e}| = 2\}$ is the set of undirected edges, and $\{n, n'\} \in \tilde{E}$ if $n$ covers $n'$ or $n'$ covers $n$. An *undirected path* from $n$ to $n'$ in $\left(N, \tilde{E}\right)$ is an alternating sequence of vertices and undirected edges $(n_0, \tilde{e}_0, ..., n_{L-1}, \tilde{e}_{L-1}, n_L)$, where $L > 0$, $n_0 = n$, $n_L = n'$, $n_l \in N$ for all $l \in \{0, ..., L\}$, $\tilde{e}_l = \{n_l, n_{l+1}\} \in \tilde{E}$ for all $l \in \{0, ..., L-1\}$, and $l \neq l' \Rightarrow \tilde{e}_l \neq \tilde{e}_{l'}$. A *cycle* in $\tilde{G}(H)$ is an undirected path from $n$ to $n$. We say $H$ is *cyclic* if $\tilde{G}(H)$ contains a cycle. The graph depicts

[8] Throughout the paper, when we say two random variables are equal, we mean almost surely.
[9] See Brooks *et al.* (2020).

what is often termed the transitive reduction of $H$: if $n > n'$, there is a path from $n$ to $n'$, but there is an edge from $n$ to $n'$ only if there is no node between them.

A subset of nodes $N' \subseteq N$ induces the information hierarchy $H' = (N', \geq)$, which we refer to as a *subhierarchy* of $H$, with the partial order being the restriction of $\geq$ on $N'$. Note that if $H'$ is a subhierarchy of $H$, $G(H') = (N', E')$ need not be a subgraph of $G(H) = (N, E)$; specifically, $E'$ may contain edges that are not in $E$.

An undirected graph is a *tree* if there is exactly one undirected path between any two nodes. An undirected graph is a *forest* if there is at most one undirected path between any two nodes, i.e., there are no cycles. Thus, a forest is a union of disjoint trees.

### 2.3 Beliefs and signals in hierarchies

A *belief allocation* on $H$ given $(\Omega, \mu_0)$ is a map that assigns a distribution of beliefs to every node in $N$. A belief allocation $\beta$ is *monotone* (with respect to $H$) if $n \geq n' \Rightarrow \beta(n) \succsim \beta(n')$, i.e., if higher nodes are Blackwell more informed than lower nodes.

A *signal allocation* on $H$ given $\Omega$ is a map that assigns a signal to every node in $N$. Given a binary relation $\mathcal{R}$ on the set of signals $\Pi$, we say that a signal allocation $\sigma$ is $\mathcal{R}$-*monotone* (with respect to $H$) if $n \geq n' \Rightarrow \sigma(n)\, \mathcal{R}\, \sigma(n')$. For instance, a signal allocation is $\trianglerighteq$-monotone if higher nodes have signals that refine lower nodes' signals. Given $\Omega$ and $\mu_0$, a signal allocation $\sigma$ *induces* a belief allocation $\beta$ if for all $n$, $\beta(n) = \langle \sigma(n) \rangle$.

### 2.4 Universal constructibility

Given $(\Omega, \mu_0)$ and a proper relation $\mathcal{R}$, we say that a monotone belief allocation $\beta$ on $H$ is *constructible under* $\mathcal{R}$ if $\beta$ is induced by some $\mathcal{R}$-monotone signal allocation on $H$.[10] The main question we explore in this paper is whether, given a proper $\mathcal{R}$, every monotone $\beta$ on $H$ is constructible under $\mathcal{R}$. As we will see, the answer will not depend on the choice of $\mathcal{R}$. However, the answer can depend on the cardinality of $\Omega$. Accordingly, given $\Omega$, say that the information hierarchy $H$ is $\Omega$-*universally constructible under* $\mathcal{R}$ if for any $\mu_0 \in \Delta\Omega$, every monotone belief allocation on $H$ is constructible under $\mathcal{R}$. A hierarchy is said to be *universally constructible under* $\mathcal{R}$ if it is $\Omega$-

[10] It is immediate that only a monotone $\beta$ is $\mathcal{R}$-constructible for any proper $\mathcal{R}$ (because $\pi\, \mathcal{R}\, \pi' \Rightarrow \pi\, \mathcal{B}\, \pi' \Rightarrow \langle \pi \rangle \succsim \langle \pi' \rangle$).

universally constructible under $\mathcal{R}$ for any $\Omega$. Note that under any $\mathcal{R}$, if a hierarchy is $\Omega$-universally constructible, then it is $\Omega'$-universally constructible if $|\Omega'| \leq |\Omega|$.

While the definition of universal constructibility requires a condition to hold across all priors, our characterization of universally constructible hierarchies would remain unchanged if we fixed any particular interior prior.[11]

# 3 Examples

In this section, we present several examples of information hierarchies and discuss which of them are universally constructible. Along the way, we establish some Lemmas and intuitions that will play a central role in the proof of our main result. We also present some examples of proper relations.

Whenever we show that a hierarchy is universally constructible under the refinement order, this will immediately imply that it is universally constructible under any proper relation. Conversely, when we show a hierarchy is not universally constructible under the belief-martingale relation, this will immediately imply that it is not universally constructible under any proper relation. Consequently, we focus on establishing that certain hierarchies are universally constructible under the refinement order and some other hierarchies are not universally constructible under the belief-martingale relation.

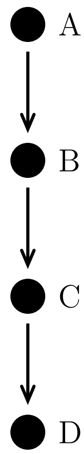
## 3.1 Examples of information hierarchies

**Example 1** (Chain)**.** There are four individuals—$A$, $B$, $C$, and $D$—ranked in alphabetical order. The set of nodes is $N = \{A, B, C, D\}$, and $\geq$ reflects the ranking relation, $A \geq B \geq C \geq D$. Figure 1a depicts the directed graph of the information hierarchy $(N, \geq)$.

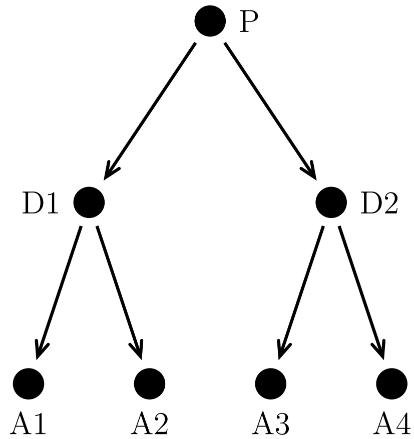
**Example 2** (Tree)**.** There is a small organization that consists of a president ($P$) who has two deputies ($D1$ and $D2$), each of whom has two assistants ($A1$ and $A2$; $A3$ and $A4$). The partial order reflects the organizational hierarchy, with $P \geq D1$, $D2$; $D1 \geq A1, A2$; and $D2 \geq A3, A4$. Figure 1b depicts the directed graph.

[11]Formally, fix any $H$, $\Omega$, $\mathcal{R}$, and an interior $\mu_0^* \in \Delta\Omega$. If every monotone belief allocation on $H$ given $(\Omega, \mu_0^*)$ is constructible under $\mathcal{R}$, then for all $\mu_0 \in \Delta\Omega$, every monotone belief allocation on $H$ given $(\Omega, \mu_0)$ is constructible under $\mathcal{R}$.

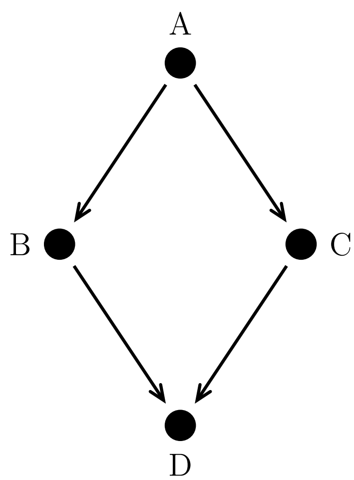
Figure 1: Example Hierarchies



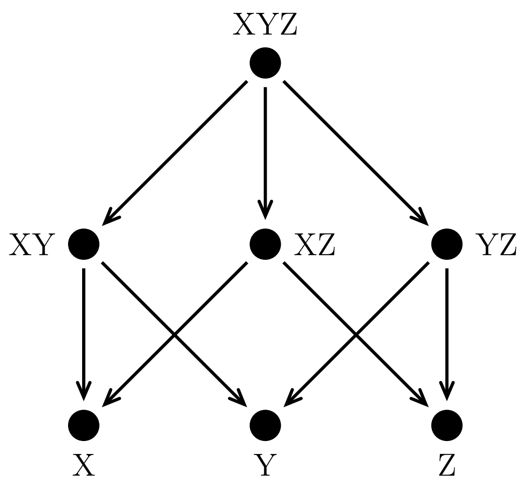
(a) Chain



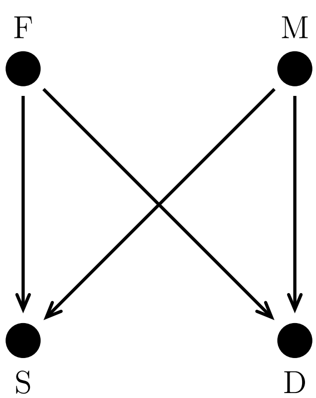
(b) Tree



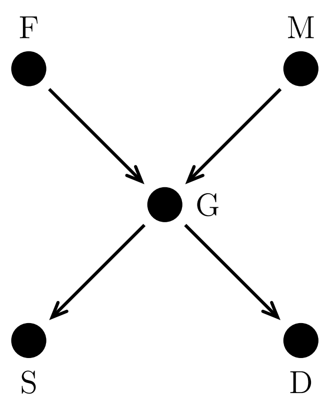
(c) Diamond



(d) Set-inclusion



(e) Crown



(f) Cross

**Example 3** (Diamond)**.** Returning to the example from the introduction, there is an organization whose president ($A$) has two deputies ($B$ and $C$) that share an assistant ($D$). We refer to this hierarchy—with $A \geq B, C$ and $B, C \geq D$—as the *diamond*, depicted in Figure 1c.

**Example 4** (Set-inclusion)**.** Elements of an information hierarchy need not represent individuals. Suppose there are three sources of information: $X$, $Y$, and $Z$. A decision maker has access to some of these sources. The nodes in this *set-inclusion* hierarchy are the possible collections of the information sources, i.e., $N = \{X, Y, Z, XY, XZ, YZ, XYZ\}$, and $\geq$ denotes the inclusion order: $XYZ \geq XY, XZ, YZ$; $XY \geq X, Y$; $XZ \geq X, Z$; $YZ \geq Y, Z$. Figure 1d depicts the directed graph of the hierarchy.

**Example 5** (Crown)**.** There are two parents, $F$ and $M$, who have two children, $S$ and $D$. We refer to this information hierarchy, with $\geq$ reflecting the parenting relation ($F \geq S, D$; $M \geq S, D$), as the *crown*, depicted in Figure 1e.[12]

**Example 6** (Cross)**.** The parents from the crown in Example 5 have hired a governess $G$, who manages the children and reports to the parents. The new partial order is $F, M \geq G$; $G \geq S, D$. The resulting *cross* hierarchy is depicted in Figure 1f. Note that the subhierarchy of the cross induced by the subset $\{F, M, S, D\}$ is a crown. However, the subgraph obtained by dropping the node $G$ is *not* the graph of the crown. As we will see, this observation will play an important role when we consider universal constructibility of the crown versus the cross.

### 3.2 Universal constructibility of the example hierarchies

#### A chain is universally constructible under the refinement order

Every information hierarchy that is a chain (i.e., that is totally ordered) is universally constructible under the refinement order. To see why, it is helpful to note the following result.[13]

**Lemma 1** (Lemma 4 from Gentzkow and Kamenica, 2017)**.** *For any $\pi'$ and $\tau$ with $\tau \succsim \langle \pi' \rangle$, $\exists \pi$ s.t. (i) $\langle \pi \rangle = \tau$ and (ii) $\pi \trianglerighteq \pi'$.*

[12] More specifically, this hierarchy is a 4-crown, where an $n$-crown is defined as a partially ordered set in which half of the nodes are maximal and half are minimal, and there is a single cycle that contains all of the nodes. Since this is the only crown we consider in this paper, we refer to it as *the* crown. Note that "crown" also has distinct meanings in graph theory.

[13] This result first appears as Theorem 1 in Green and Stokey (1978). The proof in Gentzkow and Kamenica (2017) uses the same notation as this paper.

In other words, take any signal, which induces some distribution of beliefs. There is a refinement of this signal that induces any more-informative distribution of beliefs.

Now, consider the chain hierarchy $H$ from Example 1. Consider some monotone belief allocation $\beta$ on $H$. Given $\beta(D)$, let $\sigma(D)$ be any signal that induces $\beta(D)$. Since $\beta(C) \succsim \beta(D)$, by Lemma 1, there exists some signal $\pi_C \trianglerighteq \sigma(D)$ that induces $\beta(C)$; let $\sigma(C) = \pi_C$. Similarly, since $\beta(B) \succsim \beta(C)$, there is a signal $\pi_B \trianglerighteq \sigma(C)$ that induces $\beta(B)$; let $\sigma(B) = \pi_B$. Finally, there is a $\pi_A \trianglerighteq \sigma(B)$ that induces $\beta(A)$; let $\sigma(A) = \pi_A$.

**A tree is universally constructible under the refinement order**

Establishing that a tree is universally constructible under refinement is somewhat more complicated than for a chain. The proof relies on the following result, which might be 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 same beliefs as observing $\hat{\pi}$ only.

**Lemma 2.** *For any $\overline{\pi}$, $\pi$, and $\tau'$ with $\overline{\pi} \trianglerighteq \pi$ and $\langle \pi \rangle \succsim \tau'$, $\exists \pi'$ s.t. (i) $\langle \pi' \rangle = \tau'$, and (ii) $\forall \hat{\pi}$ s.t. $\overline{\pi} \trianglerighteq \hat{\pi} \trianglerighteq \pi$, $\pi'$ is statistically redundant given $\hat{\pi}$.*

In words, given a signal $\pi$ and a distribution of beliefs $\tau'$ that is less informative than $\langle \pi \rangle$, there is a signal $\pi'$ that induces $\tau'$ but is statistically redundant given $\pi$ (or given any signal that refines $\pi$ up to an upper bound $\overline{\pi}$).[14] We sketch the proof of Lemma 2 in Section 4.2.

To better understand the content of Lemma 2, it is helpful to note that the following, stronger, conjecture does not hold. One might think that, analogously to Lemma 1, for any $\pi$ and $\tau'$ with $\langle \pi \rangle \succsim \tau'$, $\exists \pi'$ s.t. (i) $\langle \pi' \rangle = \tau'$ and (ii) $\pi \trianglerighteq \pi'$. This is not the case.[15] Lemma 2 implies, however, that we can nonetheless find a $\pi'$ s.t. $\langle \pi' \rangle = \tau'$ and $\pi'$ 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 $\pi \trianglelefteq \hat{\pi} \trianglelefteq \overline{\pi}$.

With Lemma 2 in hand, we can now establish the universal constructibility of trees. Consider the tree hierarchy $H$ from Example 2. Take some monotone $\beta$ on $H$. We construct a $\trianglerighteq$-monotone

[14] The upper bound is a technical condition related to the fact that signals are finite partitions. Given that $\pi'$ is finite, it cannot be statistically redundant given *all* refinements of $\pi$.

[15] 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 $\pi$.

$\sigma$ that induces $\beta$ as follows. Given $\beta(A1)$, let $\sigma(A1)$ be any signal that induces $\beta(A1)$. We follow the same procedure as in the case of the chain to (tentatively) assign suitable $\sigma(D1)$ and $\sigma(P)$. Now, consider assigning a signal to $A2$. The complication is that there may not exist a signal $\tilde{\pi}$ such that $\langle\tilde{\pi}\rangle = \beta(A2)$ and yet $\sigma(D1) \trianglerighteq \tilde{\pi}$. This is where Lemma 2 comes into play. By Lemma 2, we know there is a signal $\pi'$ such that $\langle\pi'\rangle = \beta(A2)$, $\langle\sigma(D1) \vee \pi'\rangle = \langle\sigma(D1)\rangle = \beta(D1)$, and $\langle\sigma(P) \vee \pi'\rangle = \langle\sigma(P)\rangle = \beta(P)$. Thus, we can replace the initial assignment of signals to $D1$ and $P$ with $\sigma(D1) \vee \pi'$ and $\sigma(P) \vee \pi'$, respectively. A similar procedure (with repeated reassignment of the previously assigned signals) can then be used to sequentially assign signals to $D2$, $A3$, and $A4$. The details of this procedure, applied to any hierarchy whose graph is a forest, are the heart of the proof of Proposition 1 below.

**The diamond is not universally constructible under the belief-martingale relation**

Consider the diamond hierarchy from Example 3. We will show this hierarchy is not $\Omega$-universally constructible for any $\Omega$.

Suppose $\Omega = \{0, 1\}$ with an equiprobable prior. Since the state space is binary, we associate each belief with $\Pr(\omega = 1)$; a belief is depicted as a number in the unit interval.

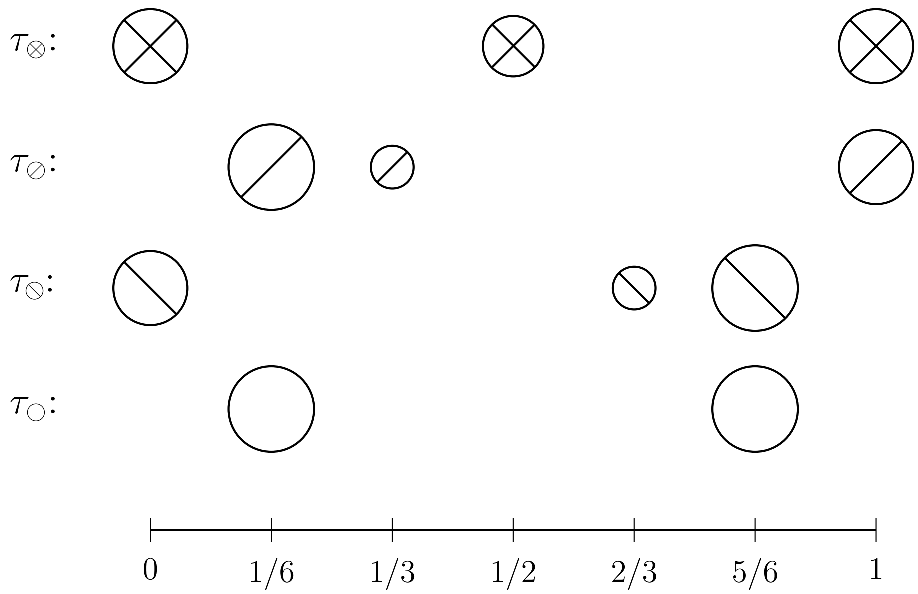
Consider the belief allocation $\beta$ that respectively assigns to nodes $A$, $B$, $C$, and $D$ the distributions of beliefs $\tau_{\otimes}$, $\tau_{\oslash}$, $\tau_{\circledbslash}$, and $\tau_{\bigcirc}$, as indicated in Figure 2. (We will refer these four distributions of beliefs again below, so it is helpful to give them names.) In the figure, we depict each distribution of beliefs as a collection of circles with matching markings. Each circle represents a belief in the support of the distribution, where the area of a circle is proportional to the probability mass on that belief. Denote a distribution that puts probability $p_i$ on belief $\mu_i$ by $(p_1, \mu_1; p_2, \mu_2; ..., ; p_n, \mu_n)$; we have $\tau_{\otimes} = \left(\frac{3}{8}, 0; \frac{1}{4}, \frac{1}{2}; \frac{3}{8}, 1\right)$, $\tau_{\oslash} = \left(\frac{1}{2}, \frac{1}{6}; \frac{1}{8}, \frac{1}{3}; \frac{3}{8}, 1\right)$, $\tau_{\circledbslash} = \left(\frac{3}{8}, 0; \frac{1}{8}, \frac{2}{3}; \frac{1}{2}, \frac{5}{6}\right)$, and $\tau_{\bigcirc} = \left(\frac{1}{2}, \frac{1}{6}; \frac{1}{2}, \frac{5}{6}\right)$.

It is easy to see that $\beta$ is monotone.[16] We will now argue that $\beta$ is not constructible under the belief-martingale relation $\mathcal{B}$.

Suppose $\sigma$ is a $\mathcal{B}$-monotone signal allocation that induces $\beta$. Consider the joint distribution of

[16] To see that $\beta(B) = \tau_{\oslash}$ is a mean-preserving spread of $\beta(D) = \tau_{\bigcirc}$, note that we can obtain $\tau_{\oslash}$ from $\tau_{\bigcirc}$ by spreading the realization $\mu = \frac{5}{6}$ in $\tau_{\bigcirc}$ to $\left\{\frac{1}{3}, 1\right\}$ in $\tau_{\oslash}$ and leaving the realization $\mu = \frac{1}{6}$ in $\tau_{\bigcirc}$ unchanged. To see that $\beta(A) = \tau_{\otimes}$ is a mean-preserving spread of $\beta(B) = \tau_{\oslash}$, note that we can obtain $\tau_{\otimes}$ from $\tau_{\oslash}$ by spreading the realizations $\mu = \frac{1}{6}$ and $\mu = \frac{1}{3}$ in $\tau_{\oslash}$ to $\left\{0, \frac{1}{2}\right\}$ in $\tau_{\otimes}$ and leaving the realization $\mu = 1$ in $\tau_{\oslash}$ unchanged. The argument for why $\beta(A) \succsim \beta(C) \succsim \beta(D)$ is symmetric.

Figure 2: Distributions over $\Omega = \{0, 1\}$ showing the diamond hierarchy is not universally constructible under $\mathcal{B}$




beliefs on $\{A, B, C, D\}$ induced by $\sigma$. Specifically, consider the conditional probability of $\tilde{\mu}_{\sigma(A)} = 1$ given $\tilde{\mu}_{\sigma(D)} = \frac{1}{6}$. Since $\sigma(B)\ \mathcal{B}\ \sigma(D)$, we must have that $\tilde{\mu}_{\sigma(B)} = \frac{1}{6}$ whenever $\tilde{\mu}_{\sigma(D)} = \frac{1}{6}$.[17] Similarly, since $\sigma(A)\ \mathcal{B}\ \sigma(B)$, we must have $\tilde{\mu}_{\sigma(A)} = 1$ whenever $\tilde{\mu}_{\sigma(B)} = 1$. Finally, since $Pr\left(\tilde{\mu}_{\sigma(A)} = 1\right) = Pr\left(\tilde{\mu}_{\sigma(B)} = 1\right)$, we must have that $Pr\left(\tilde{\mu}_{\sigma(A)} = 1 | \tilde{\mu}_{\sigma(B)} \neq 1\right) = 0$. Combining these observations, we obtain $Pr\left(\tilde{\mu}_{\sigma(A)} = 1 | \tilde{\mu}_{\sigma(D)} = \frac{1}{6}\right) = 0$.

However, applying a similar logic to the assumption that $\sigma(A)\ \mathcal{B}\ \sigma(C)$ and $\sigma(C)\ \mathcal{B}\ \sigma(D)$ tells us that $Pr\left(\tilde{\mu}_{\sigma(C)} = \frac{2}{3} | \tilde{\mu}_{\sigma(D)} = \frac{1}{6}\right) > 0$ and $Pr\left(\tilde{\mu}_{\sigma(A)} = 1 | \tilde{\mu}_{\sigma(C)} = \frac{2}{3}\ \&\ \tilde{\mu}_{\sigma(D)} = \frac{1}{6}\right) > 0$. Combining these two inequalities yields $Pr\left(\tilde{\mu}_{\sigma(A)} = 1 | \tilde{\mu}_{\sigma(D)} = \frac{1}{6}\right) > 0$, which contradicts the conclusion we derived from the fact that $\sigma(A)\ \mathcal{B}\ \sigma(B)$ and $\sigma(B)\ \mathcal{B}\ \sigma(D)$. Therefore, no $\mathcal{B}$-monotone signal allocation can induce $\beta$.

In addition to the formal argument above, here is a simple intuition for why the diamond is not universally constructible. The belief allocation we constructed is such that along each edge of the diamond, there is a unique way to spread the less-informative belief distribution to produce the more-informative belief distribution. In particular, there is a unique conditional distribution over beliefs at $B$ given the realized belief at $D$, a unique conditional distribution over beliefs at $A$ given

[17] Because the support of $\beta(B)$ is $\left\{\frac{1}{6}, \frac{1}{3}, 1\right\}$, the only way to have $\mathbb{E}\left[\tilde{\mu}_{\sigma(B)} | \tilde{\mu}_{\sigma(D)} = \frac{1}{6}\right] = \frac{1}{6}$ is to have $\tilde{\mu}_{\sigma(B)} = \frac{1}{6}$ whenever $\tilde{\mu}_{\sigma(D)} = \frac{1}{6}$.

a realized belief at $B$, etc. Thus, given a belief realization at $D$, we can derive a distribution of belief realizations at $A$ by integrating over beliefs at $B$ or by integrating over beliefs at $C$. If there were a $\mathcal{B}$-monotone signal allocation that induced these beliefs, then the conditional distribution of the belief at $A$ given the realized belief at $D$ must be "independent of path" up the diamond. As we showed above, this cannot be the case.

Since we established that the diamond is not constructible for $\Omega = \{0, 1\}$, it is not constructible for any $\Omega$.[18]

#### A set-inclusion hierarchy is not universally constructible under the belief-martingale relation

Consider the environment, described in Example 4, where a decision maker has access to three unknown sources of information. This environment induced the set-inclusion hierarchy depicted in Figure 1d. Note that the graph of the diamond hierarchy can be seen as a subgraph of the graph in Figure 1d, if we associate $A$ with $XYZ$, $B$ with $XY$, $C$ with $YZ$, and $D$ with $Y$. Consequently, under the belief-martingale relation, the set-inclusion hierarchy is not universally constructible for the same reason that the diamond hierarchy is not universally constructible. In particular, any belief allocation $\beta$ on the set-inclusion hierarchy that assigns $\beta(XYZ) = \tau_{\otimes}$, $\beta(XY) = \tau_{\oslash}$, $\beta(YZ) = \tau_{\obslash}$, and $\beta(Y) = \tau_{\bigcirc}$ is not constructible under $\mathcal{B}$.[19]

It is worthwhile to note that the mere fact that $A$, $B$, $C$, and $D$ in the diamond are ordered the same way as $XYZ$, $XY$, $YZ$, and $Z$ in the set-inclusion hierarchy does not by itself mean that non-constructibility of the diamond implies non-constructibility of the set-inclusion hierarchy. More broadly, it may be that a hierarchy $H$ is universally constructible under some $\mathcal{R}$, but $H'$ is a subhierarchy of $H$ and $H'$ is universally constructible under $\mathcal{R}$. The next two examples illustrate this possibility. Section 4.3.1 discusses the issue in detail.

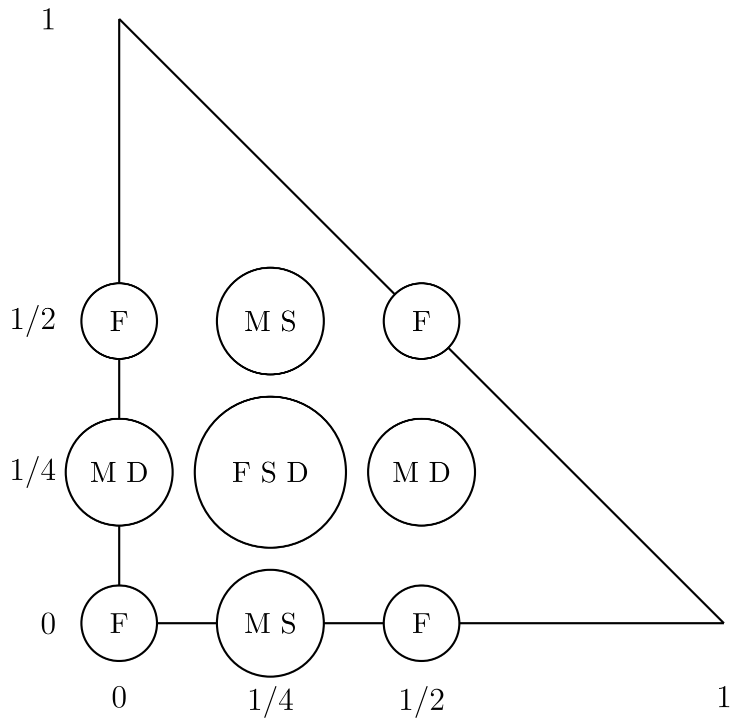
#### The crown is not universally constructible under the belief-martingale relation

Consider the crown hierarchy from Example 5. We will show that this hierarchy is not $\Omega$-universally constructible under $\mathcal{B}$ if $|\Omega| \geq 3$. (As we discuss below, it is $\Omega$-universally constructible if $|\Omega| = 2$.)

[18] Recall that if a hierarchy is universally constructible for some state space, it is universally constructible for any state space of lower cardinality.

[19] Moreover, it is easy to see that there is a monotone belief allocation on the set-inclusion hierarchy that assigns these distributions of beliefs to these nodes. For example, we could set $\beta(X) = \beta(Z) = \beta(XZ) = \underline{\tau}$.

Figure 3: Distributions over $\Omega = \{0, 1, 2\}$ showing the crown is not universally constructible under $\mathcal{B}$




Suppose $\Omega = \{0, 1, 2\}$ with a prior $\mu_0(\omega = 0) = \frac{1}{2}$, $\mu_0(\omega = 1) = \frac{1}{4}$, and $\mu_0(\omega = 2) = \frac{1}{4}$. We represent each belief as a pair $(x, y)$ in the unit square with $x + y \leq 1$, where $Pr(\omega = 1) = x$ and $Pr(\omega = 2) = y$. Consider the belief allocation $\beta$ that assigns to $F$, $M$, $S$, and $D$ the distributions of belief indicated in Figure 3. As before, we depict each distribution of beliefs as a collection of circles. If a letter $n \in \{F, M, S, D\}$ appears inside a circle, then the belief indicated by this circle is in the support of $\beta(n)$. If a belief is in the support of both $\beta(n)$ and $\beta(n')$, both $n$ and $n'$ appear inside that circle; moreover, these two distributions put the same probability mass on that belief (with the mass indicated by the area of the circle).[20]

It is easy to see from Figure 3 that $\beta$ is monotone.[21] We will now argue that $\beta$ is not constructible

[20]Formally, the belief allocation is given by $\beta(F) = \left(\frac{1}{8}, \left(0, \frac{1}{2}\right); \frac{1}{8}, \left(\frac{1}{2}, \frac{1}{2}\right); \frac{1}{8}, \left(\frac{1}{2}, 0\right); \frac{1}{8}, (0, 0); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right)\right)$, $\beta(M) = \left(\frac{1}{4}, \left(0, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{2}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{4}, 0\right); \frac{1}{4}, \left(\frac{1}{4}, \frac{1}{2}\right)\right)$, $\beta(S) = \left(\frac{1}{4}, \left(\frac{1}{4}, 0\right); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{4}, \frac{1}{2}\right)\right)$, and $\beta(D) = \left(\frac{1}{4}, \left(0, \frac{1}{4}\right); \frac{1}{2}, \left(\frac{1}{4}, \frac{1}{4}\right); \frac{1}{4}, \left(\frac{1}{2}, \frac{1}{4}\right)\right)$.

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

under $\mathcal{B}$. As before, consider the joint distribution of beliefs on $\{F, M, S, D\}$ induced by any $\mathcal{B}$-monotone signal allocation $\sigma$. Specifically, consider the conditional probability of $\tilde{\mu}_{\sigma(S)} = \mu_0$ given $\tilde{\mu}_{\sigma(D)} = \mu_0$. Since $\sigma(M)\ \mathcal{B}\ \sigma(S)$, 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 $\sigma(M)\ \mathcal{B}\ \sigma(D)$, 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, $\sigma(F)\ \mathcal{B}\ \sigma(S)$ implies $\tilde{\mu}_{\sigma(S)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(F)} = \mu_0$, and $\sigma(F)\ \mathcal{B}\ \sigma(D)$ implies $\tilde{\mu}_{\sigma(D)} = \mu_0 \Leftrightarrow \tilde{\mu}_{\sigma(F)} = \mu_0$. Hence, $Pr\left(\tilde{\mu}_{\sigma(S)} = \mu_0 | \tilde{\mu}_{\sigma(D)} = \mu_0\right) = 1$. We have reached a contradiction: no $\mathcal{B}$-monotone $\sigma$ can induce the belief allocation $\beta$.

For future reference, we summarize this discussion with the following formal result:

**Lemma 3.** *A crown is not $\Omega$-universally constructible under $\mathcal{B}$ if $|\Omega| \geq 3$.*

#### The cross is universally constructible under the refinement order

The cross hierarchy from Example 6 is universally constructible under $\trianglerighteq$. The argument is analogous to the argument for why the tree hierarchy from Example 2 is universally constructible. The procedure we discussed in that example for how to construct a $\trianglerighteq$-monotone $\sigma$ to induce any monotone $\beta$ applies to the cross as well.

Note that the subhierarchy $(\{F, M, S, D\}, \geq)$ is a crown, which is not universally constructible under $\trianglerighteq$, even though the cross is. This might seem puzzling at first. One might think that the impossibility of constructing a $\trianglerighteq$-monotone $\sigma$ that induces the belief allocation $\beta$ on $\{F, M, S, D\}$ in the previous example means that the cross is also not universally constructible under $\trianglerighteq$. But, even though that $\beta$ is monotone with respect to the crown, there is no way of extending that belief allocation to the cross (by assigning some $\beta(G)$ to node $G$) in a way that would preserve monotonicity.

That the cross is universally constructible under $\trianglerighteq$ does, however, imply that the crown is $\Omega$-universally constructible under $\trianglerighteq$ if $|\Omega| = 2$. This follows from the fact that the set of distributions of beliefs under the Blackwell order is a lattice when the state space is binary (Kertz and Rösler, 2000; Müller and Scarsini, 2006). Take an arbitrary monotone belief allocation $\beta$ on $\{F, M, S, D\}$. If the state space is binary, the lattice property implies there exists a unique distribution of beliefs $\beta(S) \vee \beta(D)$ such that $\beta(S) \vee \beta(D) \succsim \beta(S), \beta(D)$ and $\tau \succsim \beta(S) \vee \beta(D)$ for any $\tau \succsim \beta(S), \beta(D)$.

Now, let $\hat{\beta}$ be the belief allocation on the cross that sets $\hat{\beta}(n) = \beta(n)$ for $n \in \{F, M, S, D\}$ and $\hat{\beta}(G) = \beta(S) \vee \beta(D)$. It is immediate that $\hat{\beta}$ is monotone. By the universal constructibility of the cross under refinement, there is a $\trianglerighteq$-monotone $\hat{\sigma}$ that induces $\hat{\beta}$. Restricting $\hat{\sigma}$ to $\{F, M, S, D\}$ yields a $\trianglerighteq$-monotone signal allocation on the crown that induces $\beta$.

### 3.3 Examples of proper relations

As we mentioned at the outset, there are various things one might mean by "Anne is more informed than Bob." One is that Anne's signal is Blackwell more informative than Bob's, i.e., $\langle \pi_A \rangle \succsim \langle \pi_B \rangle$. Another is that Anne has observed all of Bob's information, i.e., $\pi_A \trianglerighteq \pi_B$. These, however, are not the only economically relevant comparisons of signals. For instance, it might be that Anne knows Bob's belief about the state. Or, it might be that if Anne were to observe Bob's information, she would not change her belief about the state of the world, 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.

It is easy to see that $\pi_A \trianglerighteq \pi_B$ implies that Anne knows Bob's belief about the state, which in turn implies the belief-martingale relation, $\pi_A \,\mathcal{B}\, \pi_B$. Similarly, it is easy to see that $\pi_A \trianglerighteq \pi_B$ implies that $\pi_B$ is statistically redundant given $\pi_A$, which in turn implies $\pi_A \,\mathcal{B}\, \pi_B$. Hence, Anne knowing Bob's belief about the state is a proper relation. Similarly, statistical redundancy is a proper relation.

Thus, the fact we have established that trees are universally constructible under refinement, a fortiori establishes they are universally constructible under these other relations: given any monotone belief allocation $\beta$ on a tree, we can construct a signal allocation $\sigma$ that induces $\beta$ and has the property that for every $n > n'$, observing $\sigma(n)$ suffices to know the beliefs at $\sigma(n')$. Likewise, we can construct a signal allocation $\sigma$ that induces $\beta$ such that for every $n > n'$, $\sigma(n')$ is statistically redundant given $\sigma(n)$.

In the other direction, since a diamond and the crown are not universally constructible under $\mathcal{B}$, we know there are monotone belief allocations $\beta$ on those hierarchies such that it is impossible to construct a signal allocation $\sigma$ that induces $\beta$ and has the property that for all $n > n'$, observing $\sigma(n)$ suffices to know the beliefs at $\sigma(n')$. Moreover, there are monotone belief allocations $\beta$ such that it is impossible to construct a signal allocation $\sigma$ that induces $\beta$ and has the property that for all $n > n'$, $\sigma(n')$ is statistically redundant given $\sigma(n)$.

# 4 Universal constructibility

## 4.1 Main result

We now present our main result, a characterization of universally constructible hierarchies under any proper relation.

**Theorem 1.** *Fix any proper relation $\mathcal{R}$ on signals. An information hierarchy is universally constructible under $\mathcal{R}$ if and only if its undirected graph is a forest.*

The proof of Theorem 1 is broken up into two propositions, which separately establish sufficiency and necessity of the forest condition.

**Proposition 1.** *An information hierarchy is universally constructible under the refinement order if its undirected graph is a forest.*

**Proposition 2.** *Suppose $|\Omega| \geq 3$. An information hierarchy is $\Omega$-universally constructible under the belief-martingale relation only if its undirected graph is a forest.*

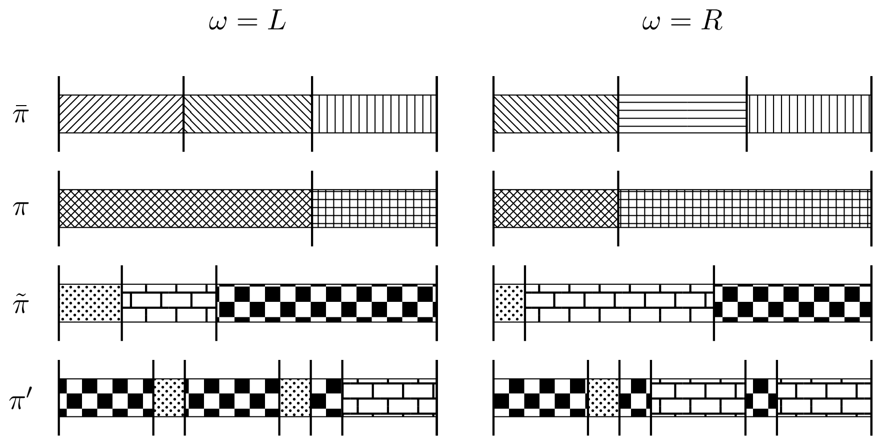
Rigorous proofs are in the Appendix. The next two subsections provide outlines of these proofs.

## 4.2 Outline of the proof of Proposition 1

Since a forest is a union of disjoint trees, Proposition 1 is a straightforward consequence of the fact that trees are universally constructible under the refinement order. As we discuss in Section 3, universal constructibility of trees follows from Lemma 2. Here we provide the intuition behind Lemma 2.

Consider the example in Figure 4. 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. For every signal, each signal realization, i.e., each element of the partition, is indicated by its shading. 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 $\langle\pi \vee \tilde{\pi}\rangle = \langle\pi\rangle$.

Figure 4: Illustration of Lemma 2




To establish the claim in Lemma 2, 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 "location" of the signal realizations in $\pi'$ are re-arranged so that the conditional probability of each signal realization of $\pi'$ in state $\omega$ given $\overline{\pi}$ is (i) the same for $\omega = L$ and $\omega = R$, and (ii) the same for any elements of $\overline{\pi}$ that refine the same element of $\pi$. The first property ensures that $\pi'$ is statistically redundant given $\overline{\pi}$, while the addition of the second property ensures that it is also redundant given any $\hat{\pi}$ s.t. $\overline{\pi} \trianglerighteq \hat{\pi} \trianglerighteq \pi$. The proof in the Appendix provides an algorithm for constructions that satisfy these two properties in general.

### 4.3 Outline of the proof of Proposition 2

Recall that in Section 3, we gave two examples of hierarchies which are not universally constructible under $\mathcal{B}$, the diamond and the crown. For each of these hierarchies, we presented monotone belief allocations that are not constructible. These two belief allocations turn out to be "canonical" in the sense that for any hierarchy that is not universally constructible under $\mathcal{B}$, we can generalize one of those belief allocations to establish non-universal constructibility.

#### 4.3.1 Constructibility and closed subhierarchies

Sometimes, we can establish that a hierarchy is not universally constructible by noting that it has a subhierarchy that is not universally constructible. As a simple example, suppose we take the diamond in Figure 1c, and form a new hierarchy $H'$ by adding nodes that are above $A$ or below $D$. Then, we can extend any monotone belief allocation $\beta$ on the diamond to $H'$ by assigning full information to the nodes above $A$ and no information to the nodes below $B$. The resulting belief allocation $\beta'$ will be monotone on $H'$, and it is constructible under $\mathcal{B}$ if and only if $\beta$ was constructible under $\mathcal{B}$ on the diamond. Thus, any hierarchy which "embeds" the diamond in this sense is not $\Omega$-constructible under $\mathcal{B}$ for any $\Omega$.

A natural conjecture might therefore be that a hierarchy is not universally constructible under $\mathcal{B}$ if it contains a subhierarchy that is not universally constructible under $\mathcal{B}$. Without further conditions, this conjecture is false. This was previously demonstrated by the crown and the cross: the cross is universally constructible under $\mathcal{B}$, but if we drop its center node, the resulting subhierarchy is a crown, which is not universally constructible under $\mathcal{B}$.

The conjecture is true, however, if we add an additional hypothesis on the subhierarchy. Given $H = (N, \geq)$, a subhierarchy $H' = (N', \geq)$ is *closed (in $H$)* 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 4 in the Appendix establishes that a hierarchy is universally constructible under $\mathcal{B}$ only if its closed subhierarchies are universally constructible under $\mathcal{B}$. (This observation holds for any proper relation, but for the purposes of proving Proposition 2, it suffices to establish it for $\mathcal{B}$.)

The basic idea behind the proof is as follows. Consider some hierarchy $H$ that is universally constructible under $\mathcal{B}$. Let $H'$ be a closed subhierarchy and let $\beta'$ be a monotone belief allocation on $H'$. We extend $\beta'$ to a monotone belief allocation $\beta$ on $H$ as follows. Since $H'$ is closed, any node in $H$ that is not in $H'$ falls into one of three mutually exclusive categories: (i) it is above some node in $H'$, (ii) it is below some node in $H'$, or (iii) it is not comparable with any node in $H'$. In case (i), we allocate full information to the node, and in cases (ii) and (iii), we allocate no information to the node. The resulting $\beta$ is monotone. Moreover, since $H$ is universally constructible under $\mathcal{B}$, there is a $\mathcal{B}$-monotone signal allocation $\sigma$ that induces $\beta$. The restriction of $\sigma$ to $H'$ is also $\mathcal{B}$-monotone and induces $\beta'$. Since $\beta'$ was arbitrary, we conclude that $H'$ is universally constructible under $\mathcal{B}$.

Thus, to prove Proposition 2, it suffices to show that if $H$ is not a forest, it contains a closed subhierarchy that is not universally constructible under $\mathcal{B}$.

#### 4.3.2 Unions of non-comparable paths

The next step in the argument utilizes an important class of hierarchies, termed *unions of non-comparable paths* (UNPs). We adapt the belief construction for the diamond to establish that UNPs are not universally constructible under $\mathcal{B}$. UNPs do not necessarily contain a diamond as a closed subhierarchy, so establishing non-constructibility of UNPs does not directly follow from the observations in Section 3.2.

An information hierarchy is a UNP if its graph is a union of at least two undirected paths between the pair of nodes $A'$ and $D'$ satisfying three properties: (i) node $A'$ is maximal; (ii) node $D'$ is minimal; and (iii) a pair of nodes are comparable only if they are in the same undirected path. Such a graph is depicted in Figure 5. Note that we have not assumed that all the nodes in the same path are comparable.

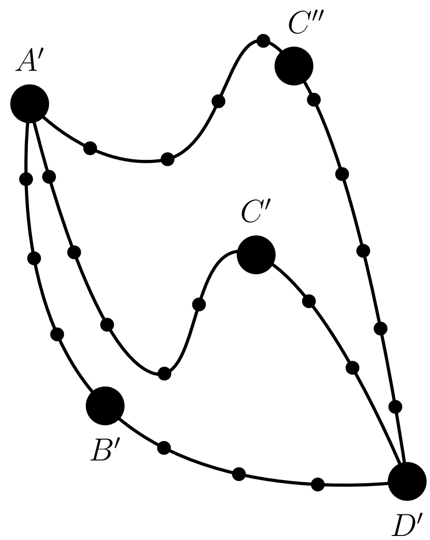
Lemma 8 in the Appendix formally establishes that UNPs are not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$. Here, we provide a sketch of that proof.

We first establish that if a belief allocation on a UNP allocates the same belief distribution on any pair of nodes $n$ and $n'$ that are in the same path (neither node being $A'$ or $D'$), then any $\mathcal{B}$-monotone signal allocation that induces it yields beliefs at $n$ and $n'$ that are perfectly correlated.

Now, consider some belief allocation $\beta$ on a UNP $H$ based on the belief allocation on the diamond from Section 3. In particular, set $\beta(A') = \tau_{\otimes}$; $\beta(D') = \tau_{\bigcirc}$; for nodes $n$ in one of the paths of $H$, say the path including $B'$ in Figure 5 (excluding $A'$ and $D'$), set $\beta(n) = \tau_{\oslash}$; and for all remaining nodes $n$ (namely the nodes in the paths including $C'$ and $C''$ in Figure 5) set $\beta(n) = \tau_{\bigcirc\!\!\!\!\diagdown}$. We show that $\beta$ is not constructible under $\mathcal{B}$. Roughly speaking, the nodes in the path with $B'$ (whose belief realizations are perfectly correlated) collectively serve the role of node $B$ in the diamond, and the other paths serve the role of node $C$ from the diamond.[22]

[22]Note, however, that this argument relies on details of this particular belief allocation. The argument does not go through if we replace $\tau$ with an arbitrary non-constructible belief allocation on the diamond. There exist belief allocations on the diamond that are non-constructible but whose "extension" to a UNP is constructible. The reason why this is possible is that in the diamond $A \geq D$, but the UNP does not require that $A' \geq D'$; thus signal allocations

Figure 5: An example of a UNP. Here, the curves depict undirected paths between $A'$ and $D'$, and slopes of the curve denote "local" comparisons between nodes. For example, the path through $B'$ is decreasing, indicating that it is a directed path with $B'$ below $A'$ and above $D'$. In contrast, the paths through $C'$ and $C''$ are not directed; while $C'$ is above $D'$, it is non-comparable with $A'$. This a UNP because it satisfies three properties: (i) node $A'$ is maximal; (ii) node $D'$ is minimal; and (iii) a pair of nodes are comparable only if they are in the same undirected path.




#### 4.3.3 Minimal cyclic closed (MCC) subhierarchies

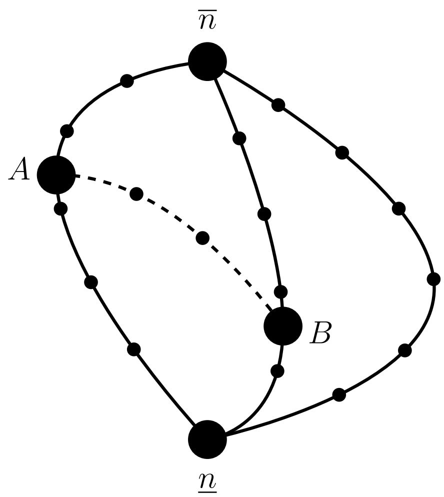
Given the observations from the previous two subsections, the proof of Proposition 2 is completed by showing that if a hierarchy is not a forest, then it contains a closed subhierarchy that is either a UNP or a crown. Below are the main steps in the argument.

Suppose that $H$ is not a forest, so that it contains a cycle. Since any hierarchy is a closed subhierarchy of itself, $H$ also contains a subhierarchy that is cyclic and closed. Given that $N$ is finite, it follows that $H$ contains a *minimal cyclic closed subhierarchy (MCC)*, i.e., a subhierarchy $H' = (N', \geq)$ such that: (i) $H'$ is cyclic and closed, and (ii) there does not exist a subhierarchy $H'' = (N'', \geq)$ with $N'' \subsetneq N'$ such that $H''$ is cyclic and closed.

We show that any MCC must be either a UNP or a crown. This result takes considerable effort to prove formally, but the basic idea 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.

$\hat{\sigma}$ on the diamond and $\sigma$ on $H$ with $\hat{\sigma}(A) = \sigma(A')$ and $\hat{\sigma}(D) = \sigma(D')$ might be $\mathcal{B}$-monotone on $H$ but not on the diamond.

Figure 6: A between set which is not an MCC. Since the nodes $A$ and $B$ are comparable, there is a smaller closed cyclic subhierarchy, namely the nodes that are between $A$ and $\underline{n}$.




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 be possible to find a smaller cyclic closed subhierarchy, as illustrated in Figure 6, 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 6 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 6. As a result, $H'$ is a UNP. This completes the outline of the proof of Proposition 2.

# 5 Applications

## 5.1 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 Anne is a decision maker with access to a set of Blackwell experiments $\{x_1, x_2, ..., x_M\}$. Suppose further that we see the distribution of Anne's beliefs that arises after she observes any subset of these experiments; our dataset $\mathcal{D} = \{\tau_S\}_{S \subseteq \{x_1,...,x_M\}}$ tells us the distribution of Anne's beliefs for every non-empty subset of experiments that she observes. When can we rationalize a given dataset $\mathcal{D}$ in the sense that we can associate each experiment $x_i$ with some signal (i.e., an element of $\Pi$) and conclude that Anne's belief formation is consistent with Bayes' rule?

To be rationalized, belief distributions in $\mathcal{D}$ have to satisfy two obvious properties. First, there is *Bayes plausibility*: the average belief cannot differ across sets of experiments, 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 larger set of experiments necessarily induces a more dispersed distribution of beliefs, i.e., $\tau_S$ 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 the answer is No. When there are three or more experiments,[23] 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 experiments $S \subseteq \{x_1, ..., x_M\}$ 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 1d, the undirected graph of this information hierarchy $H$ contains a cycle. By Theorem 1, this means that there is some monotone belief allocation on $H$, call it $\beta$, that cannot be induced by any refinement-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

[23] When there are only two experiments, $M = 2$, 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.

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 yet be refinement-monotone (since $S' \subseteq S$ implies $\bigvee_{x_i \in S} \pi(x_i) \trianglerighteq \bigvee_{x_i \in S'} \pi(x_i)$). This would contradict Theorem 1. Thus, we know that there are datasets that satisfy Bayes plausibility and Blackwell monotonicity, yet cannot be rationalized.

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

### 5.2 Information design

Consider a Sender who provides information to a set of agents. Moreover, suppose the designer faces certain types of monotonicity 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 a superior's signal refines every subordinate's signal. This constraint interacts with the organization structure. Proposition 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. With richer managerial relationships,[24] however, Proposition 2 tells us there could be desirable allocations of information which are incompatible with the constraint, even though they provide (Blackwell) more information to those higher up in the

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

organization.

## 6 Conclusion

We study the relationship between various notions of informativeness in a general model of distributed information. Take some information hierarchy, i.e., a specification of which elements should be more informed than others. We analyze whether every belief allocation on this hierarchy that is monotone in the Blackwell order (higher elements know more about the state of the world) is compatible with a signal allocation that is monotone in a stronger sense, e.g., higher elements know everything lower elements know. Our main result is that the answer is affirmative if and only if the undirected graph of the information hierarchy is a forest.

Importantly, our analysis has focused on whether *every* monotone belief allocation is constructible. Another natural goal would be to characterize, given an arbitrary information hierarchy and a proper relation, the set of monotone belief allocations that are constructible.[25]

[25] An additional, narrower, question is: which information hierarchies are universally constructible when the state space is binary? The answer must be non-trivial, since the crown is universally constructible in the binary state case, while the diamond is not.

## References

Arieli, I., Babichenko, Y., Sandomirskiy, F. and Tamuz, O. (2020). Feasible joint posterior beliefs. *arXiv preprint arXiv:2002.11362*.

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. (1951). Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, The Regents of the University of California.

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.

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.

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

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.

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 2

*Proof of Lemma 2.* 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 $\left\{X^{\overline{s}}_{\tilde{s}}\right\}_{\tilde{s}\in M^{\overline{s}}}$ be a partition of $\overline{s}$ s.t. $\forall\omega$, $\lambda\left(\left\{x|\,(x, \omega) \in X^{\overline{s}}_{\tilde{s}}\right\}\right) = \lambda\left(\{x|\,(x, \omega) \in \overline{s}\}\right) g\left(\underline{s}(\overline{s}), \tilde{s}\right)$, where $M^{\overline{s}} = \{\tilde{s} \in \tilde{\pi}|g\left(\underline{s}(\overline{s}), \tilde{s}\right) > 0\}$. Such a partition exists because $\sum_{\tilde{s}\in\tilde{\pi}} g\left(\underline{s}(\overline{s}), \tilde{s}\right) = 1$ for all $\underline{s}(\overline{s}) \in \pi$. Let $\pi' = \left\{Z^{\tilde{s}}\right\}_{\tilde{s}\in\tilde{\pi}}$ with $Z^{\tilde{s}} = \bigcup_{\overline{s}\in\overline{\pi} \text{ s.t.} \tilde{s}\in M^{\overline{s}}} X^{\overline{s}}_{\tilde{s}}$. We now show that $\pi'$ satisfies (i) and (ii). To show (i), it suffices to show that $p\left(Z^{\tilde{s}}|\omega\right) = p(\tilde{s}|\omega)$ for every $\tilde{s}$ and $\omega$. We have

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

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 $\{s^{\vee}_i\}_{i\in I}$ s.t. $s^{\vee}_i \in \hat{\pi} \vee \pi'$ for all $i$. It will suffice to show that for every $\omega$, $\omega'$, and $s^{\vee}_i$, we have

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

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

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

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,

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

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 2. □

### A.2 Proof of Proposition 1

*Proof of Proposition 1.* Let $H$ be an information hierarchy and suppose $\tilde{G}(H)$ is a forest. Let $\beta$ be a monotone belief allocation on $H$. We will construct a $\trianglerighteq$-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

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

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 $\tilde{G}(H)$ between $f(l)$ and nodes in $N_{l-1}^f$.

*Proof of Claim 1.* Suppose toward contradiction that $f(l)$ has an edge in $\tilde{G}(H)$ 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 $\tilde{G}(H)$. Moreover, there must be a path between $n$ and $n'$ in $\tilde{G}\left(\left(N_{l-1}^f, \geq\right)\right)$. 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 $\tilde{G}\left(\left(N_{l'-1}^f, \geq\right)\right)$ and thus in $\tilde{G}\left(\left(N_{l-1}^f, \geq\right)\right)$. Hence, there is a path from both $n$ and $n'$ to $\underline{n}$ and thus a path between $n$ and $n'$ in $\tilde{G}\left(\left(N_{l-1}^f, \geq\right)\right)$. So, there must be a path between $n$ and $n'$ in $\tilde{G}(H)$ 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, $\tilde{G}(H)$ is a forest, so there cannot be multiple paths between two nodes; we have reached a contradiction. $\diamondsuit$

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

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 $N^f_1$ induces appropriate beliefs and is $\trianglerighteq$-monotone. For $l \geq 2$, there are three cases: $f(l) \in \mathit{CoveredBy}\left(N^f_{l-1}\right)$, $f(l) \in \mathit{Covering}\left(N^f_{l-1}\right)$, and $f(l) \in \mathit{Disconnected}\left(N^f_{l-1}\right)$.

We first consider the case $f(l) \in \mathit{Covering}\left(N^f_{l-1}\right)$. Note that, by Claim 1, $f(l)$ covers exactly one node in $N^f_{l-1}$ (call this node $\bar{m}$) and is not covered by any nodes in $N^f_{l-1}$. Since $\beta(f(l)) \succsim \beta(\bar{m})$, there exists some $\pi \trianglerighteq \sigma^{l-1}(\bar{m})$ such that $\langle \pi \rangle = \beta(f(l))$ (cf: Lemma 1). We set $\sigma^l(f(l)) = \pi$ and we keep the signal allocation to nodes in $N^f_{l-1}$ unchanged, i.e., $\sigma^l(m) = \sigma^{l-1}(m)$ for all $m \in N^f_{l-1}$. It is clear that $\sigma^l$ induces appropriate beliefs (by the inductive hypothesis for $m \in N^f_{l-1}$ and by construction for $f(l)$). We also need to show that this signal allocation on $N^f_l$ is $\trianglerighteq$-monotone. Consider any $m, m' \in N^f_l$ such that $m > m'$. Since $f(l) \in \mathit{Covering}\left(N^f_{l-1}\right)$, either $m, m' \in N^f_{l-1}$ 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 $f(l) \in \mathit{CoveredBy}\left(N^f_{l-1}\right)$. Let $\underline{m}$ be the node in $N^f_{l-1}$ that covers $f(l)$. Denote $\tau' = \beta(f(l))$, $\pi = \sigma^{l-1}(\underline{m})$, and $\overline{\pi} = \bigvee_{m \in N^f_{l-1}} \sigma^{l-1}(m)$. By Lemma 2, we know $\exists \pi'$ such that (i) $\langle \pi' \rangle = \tau'$, and (ii) $\forall \hat{\pi}$ s.t. $\overline{\pi} \trianglerighteq \hat{\pi} \trianglerighteq \pi$, $\langle \hat{\pi} \vee \pi' \rangle = \langle \hat{\pi} \rangle$. We set $\sigma^l(f(l)) = \pi'$. For $m \in N^f_{l-1}$, 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 $\sigma^l$ induces appropriate beliefs and is $\trianglerighteq$-monotone. We have that $\langle \sigma^l(f(l)) \rangle = \langle \pi' \rangle = \tau' = \beta(f(l))$. For $m \in N^f_{l-1}$, 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, $\overline{\pi} = \bigvee_{m' \in N^f_{l-1}} \sigma^{l-1}(m') \trianglerighteq \sigma^{l-1}(m)$; hence, $\langle \sigma^l(m) \rangle = \langle \sigma^{l-1}(m) \vee \pi' \rangle = \langle \sigma^{l-1}(m) \rangle$. For $m \in N^f_{l-1}$ s.t. $m \ngeq f(l)$, $\langle \sigma^l(m) \rangle = \langle \sigma^{l-1}(m) \rangle$. 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 $m \in N^f_{l-1}$. We now need to show that $\sigma^l$ is $\trianglerighteq$-monotone. Consider any $m, m' \in N^f_{l-1}$ 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) $\sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(m')$, we know that $\sigma^{l-1}(m) \vee \pi' \trianglerighteq \sigma^{l-1}(m') \vee \pi'$, and hence $\sigma^l(m) \trianglerighteq \sigma^l(m')$. The second case is where $m \geq f(l)$ and $m' \ngeq f(l)$. Then, $\sigma^l(m) = \sigma^{l-1}(m) \vee \pi'$ and $\sigma^l(m') = \sigma^{l-1}(m')$. Since (by the inductive hypothesis) $\sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(m')$, we have that $\sigma^l(m) = \sigma^{l-1}(m) \vee \pi' \trianglerighteq \sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(m') = \sigma^l(m')$. Finally, suppose that $m \ngeq f(l)$ and $m' \ngeq f(l)$. Then, $\sigma^l(m) = \sigma^{l-1}(m)$ and $\sigma^l(m') = \sigma^{l-1}(m')$. Since (by the inductive hypothesis) $\sigma^{l-1}(m) \trianglerighteq \sigma^{l-1}(m')$, we have that $\sigma^l(m) \trianglerighteq \sigma^l(m')$.

Finally, suppose $f(l) \in \textit{Disconnected}\left(N^f_{l-1}\right)$. 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 $N^f_{l-1}$ unchanged, i.e., $\sigma^l(m) = \sigma^{l-1}(m)$ for all $m \in N^f_{l-1}$. It is clear that $\sigma^l$ induces appropriate beliefs (by the inductive hypothesis for $m \in N^f_{l-1}$ and by construction for $f(l)$). Since $f(l)$ is not comparable to any node in $N^f_{l-1}$, the fact that the signal allocation on $N^f_{l-1}$ is $\trianglerighteq$-monotone implies that the signal allocation on $N^f_l$ is also $\trianglerighteq$-monotone. This completes the proof. □

### A.3 Proof of Proposition 2

**Lemma 4.** *If an information hierarchy $H$ is $\Omega$-universally constructible under $\mathcal{B}$ and $H'$ is a closed subhierarchy of $H$, then $H'$ is $\Omega$-universally constructible under $\mathcal{B}$.*

*Proof of Lemma 4.* Fix $\Omega$. Suppose $H$ is an information hierarchy that is $\Omega$-universally constructible under $\mathcal{B}$. Consider some $\mu_0$. Suppose $H' = (N', \geq)$ is a closed subhierarchy of $H$. Let $\beta'$ be a monotone belief allocation on $H'$. We need to construct a $\mathcal{B}$-monotone signal allocation on $H'$ that induces $\beta'$. Let $\beta$ be a belief allocation on $H$ 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 $\beta(n) = \overline{\tau}$; 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 $\beta(n) = \overline{\tau} \succsim \beta(n')$. 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 $n' \in N'$, $\beta(n) = \overline{\tau} \succsim \beta(n')$.

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, $\beta(n) \succsim \beta(n') = \underline{\tau}$. ♢

Since $\beta$ is monotone on $H$, and $H$ is $\Omega$-universally constructible under $\mathcal{B}$, there exists a $\mathcal{B}$-monotone signal allocation $\sigma$ on $H$ that induces $\beta$. Clearly, the restriction of $\sigma$ to $N'$ induces $\beta'$ and is $\mathcal{B}$-monotone on $H'$. □

The next result shows that if $H'$ is a closed subhierarchy of $H$, then $G(H')$ is the subgraph of $G(H)$ obtained by dropping edges with nodes that are not in $H'$.

**Lemma 5.** *Fix a hierarchy $H = (N, \geq)$ and a closed subhierarchy $H' = (N', \geq)$. Let $E$ be the set of edges in $G(H)$. Then $G(H') = (N', E')$, where $E' = \{(n, n') \in E | n, n' \in N'\}$.*

*Proof of Lemma 5.* Fix $n, n' \in N'$. We need to show that $(n, n') \in E'$ if and only if $(n, n') \in E$. If $(n, n') \in E$, 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 $(n, n') \in E'$. If $(n, n') \in E'$ , 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 $(n, n') \in E$. □

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

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

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 $G(H)$ 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 $\tilde{G}(H')$ is a *spanning cycle* if every node in $N'$ is in the cycle.

**Lemma 6.** *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 $\tilde{G}\,(H')$ is a spanning cycle.*

*Proof of Lemma 6.* 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 5, $G\,(H')$ is the subgraph of $G\,(H)$ 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 $G(H)$. Note that $Btw\,(n, n')$ is closed, so that these paths are in $G\,((Btw\,(n, n')\,, \geq))$ 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 $G(H)$. 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 $G(H)$. 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, $\tilde{G}\,(H') = \left(N', \tilde{E}\right)$ contains a cycle $C = (n_0, \tilde{e}_0, ...n_{L-1}, \tilde{e}_{L-1}, n_L)$ 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 $(n_0, e_0, ..., n_{L-1}, e_{L-1}, n_L)$, its *undirected analog* is the undirected path $(n_0, \tilde{e}_0, ...n_{L-1}, \tilde{e}_{L-1}, n_L)$ where $\tilde{e}_i = \{n_i, n_{i+1}\}$. Say that a directed path $P$ in $G(H)$ *only contains edges in* $C$ if every edge in the undirected analog of $P$ is in $C$. A directed path $P$ in $G(H)$ 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 $i \leq \underline{l} < \bar{l} \leq j$, $\left(n_{\underline{l}}, n_{\bar{l}}\right)$ 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 $G(H)$ 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 $\overline{n} = \max\{n_i, n_j\}$ and $\underline{n} = \min\{n_i, n_j\}$. Let $P^e$ denote the external directed connection from $\overline{n}$ to $\underline{n}$. Let $\tilde{P}$ be the undirected path $(n_i, \tilde{e}_i, ..., \tilde{e}_{j-1}, n_j)$ in $\tilde{G}\,(H)$ 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., $\tilde{Q} = (n_i, \tilde{e}_{i-1}, n_{i-1}, ..., \tilde{e}_0, n_0, \tilde{e}_{L-1}, n_{L-1}, ..., \tilde{e}_j, n_j)$. Let $S = N_{\tilde{P}} \cup Btw\,(\overline{n}, \underline{n})$.

*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 $\tilde{G}\,((S, \geq))$. One path is $\tilde{P}$. The other path is the undirected analog 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 $Y \subseteq S$. $\diamondsuit$

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 7.** *If the hierarchy $H = (N, \geq)$ is such that every cycle in $\tilde{G}\,(H)$ 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 7.* Since there exists a spanning cycle $\tilde{G}\,(H)$, for any pair of nodes $n, n' \in N$, there exist two undirected paths $P = \left(n, \tilde{e}_0^P, n_1^P ..., n_{L^P-1}^P, \tilde{e}_{L^P-1}^P, n'\right)$ and $Q = \left(n, \tilde{e}_0^Q, n_1^Q ..., n_{L^Q-1}^Q, \tilde{e}_{L^Q-1}^Q, n'\right)$ such that $N = N_P \cup N_Q$ and $\tilde{E}_P \cap \tilde{E}_Q = \emptyset$, where $\tilde{E}_P = \left\{\tilde{e}_0^P, ..., \tilde{e}_{L^P-1}^P\right\}$ and $\tilde{E}_Q = \left\{\tilde{e}_0^Q, ..., \tilde{e}_{L^Q-1}^Q\right\}$. 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 $\hat{n} = n^P_{l^P} = n^Q_{l^Q}$. Now, consider the undirected path $\left(n, \tilde{e}^P_0, n^P_1, ..., \tilde{e}^P_{l^P-1}, n^P_{l^P}, \tilde{e}^Q_{l^Q-1}, n^Q_{l^Q-1}, \tilde{e}^Q_{l^Q-2}, ..., n\right)$. Since $\tilde{E}_p \cap \tilde{E}_q = \emptyset$, this is a well-defined path. But it is a cycle that is not spanning and thus we have reached a contradiction. $\square$

Suppose a hierarchy $H$ has nodes

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

with $L \geq 2$, such that for all $l = 1, \ldots, L$, (i) $K_l \geq 1$, (ii) $\overline{n} > n^l_1$, (iii) $n^l_{K_l} > \underline{n}$, (iv) $n^l_k$ and $n^l_{k+1}$ are comparable for every $k = 1, \ldots, K_l - 1$, and (v) $n^l_k$ is not comparable to $n^{l'}_{k'}$ if $l \neq l'$. In this case, we say that $H$ is a *union of non-comparable paths (UNP).*

**Lemma 8.** *If the hierarchy $H$ is a UNP, then $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$.*

*Proof of Lemma 8.* To establish the Lemma, it suffices to show that if the hierarchy is a UNP, it is not $\Omega$-universally constructible under $\mathcal{B}$ for $\Omega = \{0, 1\}$. We will present a monotone belief allocation $\beta$, based on the belief allocation on the diamond in Section 3, and show that $\beta$ is not constructible under $\mathcal{B}$. Let $\beta(\overline{n}) = \tau_{\otimes}$; $\beta(\underline{n}) = \tau_{\bigcirc}$; for $k = 1, \ldots, K_1$, let $\beta(n^1_k) = \tau_{\diagdown\!\!\!\!\bigcirc}$; and for all $l = 2, \ldots, L$ and $k = 1, \ldots, K_l$, let $\beta(n^l_k) = \tau_{\oslash}$. Recall that $\tau_{\otimes}$ has support $\left\{0, \frac{1}{2}, 1\right\}$, $\tau_{\bigcirc}$ has support $\left\{\frac{1}{6}, \frac{5}{6}\right\}$, $\tau_{\diagdown\!\!\!\!\bigcirc}$ has support $\left\{0, \frac{2}{3}, \frac{5}{6}\right\}$, and $\tau_{\oslash}$ has support $\left\{\frac{1}{6}, \frac{1}{3}, 1\right\}$. By the argument in footnote 16, $\beta$ is monotone.

Toward a contradiction, suppose $\sigma$ is $\mathcal{B}$-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 $\tilde{\mu}_{\sigma\left(n^l_k\right)} = \tilde{\mu}_{\sigma\left(n^l_1\right)}$ 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 $\tilde{\mu}_{\sigma\left(n^l_{k+1}\right)} = \tilde{\mu}_{\sigma\left(n^l_k\right)}$. We know $n^l_k$ and $n^l_{k+1}$ are comparable. Assume that $n^l_{k+1} > n^l_k$; the case where $n^l_k > n^l_{k+1}$ is analogous and omitted. Since $\sigma$ is $\mathcal{B}$-monotone, we have $\mathbb{E}[\tilde{\mu}_{\sigma(n^l_{k+1})} | \tilde{\mu}_{\sigma(n^l_k)}] = \tilde{\mu}_{\sigma\left(n^l_k\right)}$. Now suppose toward contradiction that $\tilde{\mu}_{\sigma\left(n^l_{k+1}\right)} \neq \tilde{\mu}_{\sigma\left(n^l_k\right)}$. We would conclude that $\tilde{\mu}_{\sigma\left(n^l_{k+1}\right)}$ is a strict mean-preserving spread of $\tilde{\mu}_{\sigma\left(n_k^l\right)}$.[26] But that would mean that $\tilde{\mu}_{\sigma\left(n_{k+1}^l\right)}$ is not equal to $\tilde{\mu}_{\sigma\left(n_k^l\right)}$ in distribution and thus that $\langle\sigma\left(n_{k+1}^l\right)\rangle \neq \langle\sigma\left(n_k^l\right)\rangle$. $\diamondsuit$

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 $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0 | \tilde{\mu}_{\sigma\left(n_1^1\right)} = 0\right) = 1$ since $\mathbb{E}\left[\tilde{\mu}_{\sigma(\overline{n})} | \tilde{\mu}_{\sigma\left(n_1^1\right)} = 0\right] = 0$ and the support of $\tilde{\mu}_{\sigma(\overline{n})}$ lies above 0. Moreover, since $Pr\left(\tilde{\mu}_{\sigma(\overline{n})} = 0\right) = \frac{3}{8} = Pr\left(\tilde{\mu}_{\sigma\left(n_1^1\right)} = 0\right)$, we must also have $Pr\left(\tilde{\mu}_{\sigma\left(n_1^1\right)} = 0 | \tilde{\mu}_{\sigma(\overline{n})} = 0\right) = 1$. 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 9.** *If the hierarchy $H = (N, \geq)$ is cyclic and $N$ is a simple between set in $H$, then $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$.*

*Proof of Lemma 9.* Since $H = (N, \geq)$ is cyclic and $N$ is a simple between set, $H$ must be a UNP. Therefore Lemma 8 establishes that $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$. $\square$

**Lemma 10.** *If the hierarchy $H$ is cyclic and not a crown, and every cycle in $\tilde{G}(H)$ is a spanning cycle, then $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$.*

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

*Proof of Lemma 10.* Suppose that $H = (N, \geq)$ is cyclic and not a crown, and that every cycle in $\tilde{G}(H)$ 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, $\tilde{G}(H)$ 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 5, the nodes in $H'$ are part of a cycle in $\tilde{G}(H)$ as well. Since every cycle in $\tilde{G}(H)$ is a spanning cycle, we must have $N = N'$, and thus $H$ is a crown subhierarchy, and we have reached a contradiction. ◊

By Claim 8, we can find $\overline{n}$ and $\underline{n}$ in $N$ that are maximal and minimal, respectively, such that $\overline{n}$ does not cover $\underline{n}$. By Lemma 7, there are two distinct undirected paths $P$ and $Q$ in $\tilde{G}(H)$ from $\overline{n}$ 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 $\{\overline{n}, \underline{n}\}$. As a result, $H$ is a UNP. Lemma 8 therefore implies that $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$. □

*Proof of Proposition 2.* Suppose $\tilde{G}(H)$ 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 6, either (i) $N'$ is a simple between set in $H$, or (ii) every cycle in $\tilde{G}(H')$ 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 9 implies that $H'$ is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$. Now consider case (ii). If $H'$ is not a crown, then Lemma 10 implies that is not $\Omega$-universally constructible under $\mathcal{B}$ for any $\Omega$. If $H'$ is a crown, Lemma 3 implies it is not $\Omega$-universally constructible under $\mathcal{B}$ if $|\Omega| \geq 3$. Hence, if $|\Omega| \geq 3$, $H'$ is not $\Omega$-universally constructible under $\mathcal{B}$. Therefore, by Lemma 4, $H$ is not $\Omega$-universally constructible under $\mathcal{B}$ if $|\Omega| \geq 3$. □