MAIWAR (Modelling Australian Industry with Weak-solutions, AMPL and Regions) [1]

By Patrick H. O'Callaghan [2]

Australian Institute for Business and Economics

University of Queensland

Policy makers make decisions regarding medium term policies where capital is evolving. They do so in the face of uncertainty. Existing Computable General Equilibrium models fail to adequately model saving behaviour in the face of uncertainty. Although macroeconomic models typically address this concern, they rely on local first-order approximations around the non-stochastic steady-state: where capital is no longer evolving. Global approximation methods are intractible or slow in multi-sectoral settings.

In this white paper, we extend a simple, yet powerful approximation method of Cai and Judd to the multi-sectoral setting. Unlike other methods, this approach is built for medium-term problems. The solution is *weak*: i.e. characterised by an empirical distribution on the set of paths through time. Similar to structural econometric models, the key measure of accuracy is that the mean (of the simulated sample paths) satisfies the Intertemporal Euler equation.

From the past, the present acts prudently, lest it spoil future action.

Titian: Allegory of Prudence

---

[1] I thank Josh Aberdeen, Patrick Duenow, Cameron Gordon and John Mangan for participating in the creation of MAIWAR. This version has time stamp 00:01 (AEST), 3rd May, 2022. The most recent version (as well as past versions) can be found at `https://arxiv.org/abs/1904.02934` .

[2] Email address `p.ocallaghan@uq.edu.au` and ORCID iD 0000-0003-0606-5465

1. Introduction

1.1. Steady-state models

Beyond the short-term, economic theory provides the lens through which policymakers peer into the future. In medium term, the economy is unlikely to reach its steady state. Moreover, in models with multiple sectors recent research has highlighted the importance of nonlinear effects. For such problems, first-order steady-state approximations are ill-suited. Yet this is the solution method that is most commonly applied today.

Computable General Equilibrium (CGE) models are often used to provide insights into the regional and sectoral impacts of current policies and shocks that propagate through the economy's production networks over time. Such approximations are even less appealing for nonlinear production networks that arise when elasticities of substition are below one. Impressive recent work using Hamiltonian methods still fails to capture the central economic theme of uncertainty.

existing models that assume steady state approximations yet even the impact of long-term policies, such as 2050 net-zero emissions targets should involve a careful analysis of the medium term. For instance, a policy that is unsustainable or sub-optimal in the interim is liable to fail altogether.

Experience is the basis of prediction. Yet outside of stylised settings, even the most experienced forecasters do not claim access to "the full model" that is closed with respect to the relevant states of the world.

Example 1. *The canonical large-world setting is that of the global financial markets. In 2019, all models ommited details of COVID-19. Similarly, in 2007, a clear description of the sub-prime mortgage crisis was beyond reach. The central role of simulation and bootstrap methods in empirical finance points to a prevalence of inductive reasoning.*[3] *That is, to forecasting on the basis of past cases as opposed to a full description of future states. At the same time, market makers need to set prices that are robust to changes that open the door to exploitation via arbitrage.*

[3] Consider **?**, **?**, **?** and **?**.

Incomplete models are a key motivation for recent axiomatic updates to the standard Bayesian framework such as "reverse Bayesianism" of **?**. This and other work (**?**, **?** and **?**) on unawareness and robustness in the state-space sense provide part of our inspiration for the present upgrade to the axiomatic foundations of inductive inference in **?**, henceforth [GS]. By taking rankings as primitive, we extend [GS] to accommodate a forward-looking version of the second-order induction of **?**.

In the present framework, the basic building blocks of the model are observations or (synonymously) past cases. A given past case may be empirical or theoretical, and the predictor's model is naturally bounded in size and scope by the predictor's experience. Our contribution is to extend the framework of [GS] to model less experienced predictors that are prudent. That is to agents that are able to proactively engage in second-order induction: by "pulling themselves up by the bootstraps" and looking at how their model extends to novel cases. Thus allowing them to survive their initial phase of inexperience.

In the remainder of this section, we informally introduce: the model of section 2; the axioms and matrix representation of section 3; and the applications to which we return in section 4. Proofs of main results appear in appendices A to C.

Synopsis of model and results with examples. The predictor is endowed with a qualitative plausibility ranking of eventualities given her current database $D^\star$ of past cases. Moreover, the same is true for every finite resampling of $D^\star$. We identify conditions on the resulting family of (ordinal) rankings for the existence of a suitably unique real-valued matrix $\mathbf{v}$ on eventualities$\times$cases that *represents* the information in these rankings. The form of this representation is linear on databases (i.e. additive over cases) and separable on eventualities, so that for every database D, eventuality y is more likely than x if, and only if,

$$\sum_{c\in D} \mathbf{v}(x,c) \leqslant \sum_{c\in D} \mathbf{v}(y,c). \tag{1}$$

The similarity weight $\mathbf{v}(x,c)$ is the degree of support that case c lends to x.

Example 2. *As a canonical example, consider predicting the slope coefficient β_i from the regression of asset i's returns on the market portfolio (in the two-pass* *method of* **?***). Then values* β_i *are eventualities and* D^* *is the current sample of past returns. In this setting,* $\mathbf{v}$ *is an empirical log-likelihood function that we generate via a generalised notion of bootstrapping of cases in* $D^\star$*.*

Key to the contribution of [GS] is an endogenous notion of case types: a partition of cases according to the marginal information they contribute to a given database. This marginal contribution is measured in terms of the impact on the rankings of eventualities.[4] Case types are analogous to states in the sense that they form the model's dimensions: the lower the dimension the less the experience.

Example 3 (Search Engine Results Page, SERP)**.** *Advertising aside, when users conduct a web search, the search engine compiles a ranking* $\preceq_{D^\star}$ *of (web)pages* $x, y, z, \ldots$ *on the basis of its database* $D^\star$ *of past cases: searches of past users plus feedback from subsequent clicks. Resampling yields other databases* D *and other rankings. At one extreme, past cases may be so similar to one another that the same plausibility ranking arises regardless of how the data is resampled. At the other, past cases may be sufficiently rich that resampling yields every feasible plausibility ranking of eventualities.*

A rich set of past cases is at the heart of the diversity axiom of [GS]–which we refer to as *4*-diversity. This restricts the model to predictors whose current data is sufficiently rich that a resampling exercise generates *all* $4! = 24$ *strict (i.e. total) rankings of every subset of four eventualities.* In this paper, we accommodate less inexperienced predictors by only replacing *4*-diversity with conditional-*2*-diversity. Given the other basic axioms of [GS], conditional-*2*-diversity turns out to be equivalent to requiring that, for every three distinct eventualities x, y and z, resampling generates at least 3 of the $3! = 6$ possible distinct strict rankings of this triple. Conditional-*2*-diversity is minimal in the sense that, in its absence, we lose both existence and uniqueness of the similarity representation (see example 7).

To compensate for a lack of experience, we introduce a subtle and more flexible notion of cases that allows us to capture the predictor's potential awareness of her limited experience. Formally, related notions in the literature go by the name of unforeseen consequences in **?** and shadow propositions in the setting dynamic

[4]This notion of case type is therefore close to Quine's "perceptual similarity" (see section 4).

awareness in **?**. Via a content-free case $\mathfrak{f}$, the predictor can explore the impact on her model of the arrival of a novel case type. In effect, this involves meta-analysis of how her similarity function will evolve over time. Hence the reference to second-order induction.

The prudent predictor ensures her model is robust to the arrival of novel case types. She ensures that, when a novel case arrives, she can accept the ranking it generates without finding herself in the potentially costly position of generating intransitive rankings when she combines past and novel cases.

Example 4 (Second-order inductive inference)**.** *Consider a search-engine startup seeking to establish itself in the face of incumbents with the experience of Google. The start-up engages in second order induction when it is learning the similarity function itself. This may include learning the values* $\mathbf{v}(x, c)$ *of* eq. (1)*, but it may also involve costly updates of the model structure "on the fly", e.g. redefining case types, rankings, etc. The startup is prudent if,* ex ante*, it structures its model to ensure that it is relatively costless to extend to novel case types*: *once they arrive.*

Prudence is only worthwhile when revisions of the predictor's model 'on the fly', once the novel case arrives, is costly. Consider "zero-day attacks" in the setting of cyber security. **?** highlight the essence of time when a novel attack on a computer network arrives; and that such attacks are novel precisely because cyber-security experts have already built in solutions to known vulnerabilities. Also, tradeoffs between time, cost and learning are nowhere more important than in finance. For a bond-market setting, we are able to provide a formal equivalence between prudence and arbitrage pricing in section 4. In section 4, we also discuss: other applications; empirical evidence linking intransitivity, memories and novelty; and connections with the literature on second-order induction in more detail.

2. Model

Following [GS], let the nonempty set X denote the conceivable *eventualities* of the present prediction problem and let $\mathrm{rel}(X)$ denote the set of binary relations or rankings on X. For instance, for a search engine, we identify webpage x with the eventuality "page x is the desired webpage". The predictor is equipped with her current memory $C^\star$: the union of a finite set of past cases $D^\star$ and a *variable* or *free case* $\mathfrak{f}$. The cases in $D^\star$ collectively represent the forecaster's relevant observations or experience. Our first and most fundamental modification of the primitives of [GS] is the inclusion of $\mathfrak{f}$ in the current memory $C^\star$.

Remark. *On a computer, a natural implementation of this setup is the following. Take every case $c \in C^\star$ to consist of a pair $p \times m$: a pointer p that references a memory location and the memory content m. Each $c \in D^\star$ is identical to a case in the setting of* [GS]. *But, for $c = \mathfrak{f}$, there is no meaningful memory content, so m is "empty" or assigned an arbitrary null value. From another perspective, cases in $D^\star$ are constant (of arity zero) whereas $\mathfrak{f}$ is a variable (of positive arity).*

For every nonempty subsample $D \subseteq D^\star$, the predictor has sufficient information to determine a well-defined ranking $\preceq_D$ in $\mathrm{rel}(X)$. *In contrast, since $\mathfrak{f} = \acute{p} \times \acute{m}$ has no meaningful memory content, $\preceq_{\mathfrak{f}}$ is indeterminate and a free variable in* $\mathrm{rel}(X)$. *The prudent predictor gains a better understanding of her current model by assigning a ranking to $\acute{m}$ and exploring the extensions of definition 1.*

Like [GS], we accommodate a forecaster that goes beyond her current memory and includes hypothetical cases $\mathbb{C}$ that she may not have experienced, but which, through reasoning, interpolation or resampling, she can clearly describe. These hypothetical cases are formally constant, like members of $D^\star$. With *case resampling and subsampling* from the literature on bootstrapping in mind, let

$$\mathbb{D} \stackrel{\text{def}}{=} \{D \subseteq \mathbb{C} : \sharp D < \infty\}$$

denote the set of (finite) *determinate or constant databases.* (These are referred to as *memories* in [GS].) Like $D^\star$, each $D \in \mathbb{D}$ contains no copies of $\mathfrak{f}$.

Let $[\mathfrak{f}]$ denote a set of copies of $\mathfrak{f}$. Finally, let $\mathbb{C}^{\mathfrak{f}} \stackrel{\text{def}}{=} \mathbb{C} \cup [\mathfrak{f}]$ and let $\mathfrak{C}$ denote a member of $\{\mathbb{C}, \mathbb{C}^{\mathfrak{f}}\}$. Let $\mathbb{D}^{\mathfrak{f}}$ denote the corresponding set of all finite subsets of $\mathbb{C}^{\mathfrak{f}}$ and take

$$\mathbb{D} = \begin{cases} \mathbb{D} & \text{if, and only if, } \mathfrak{C} = \mathbb{C}, \text{ and} \\ \mathbb{D}^{\mathfrak{f}} & \text{otherwise.} \end{cases}$$

The predictor is endowed with a well-defined plausibility ranking $\preceq_D$ in $\mathrm{rel}(X)$ for each D in $\mathbb{D}$. Denote the symmetric part by $\simeq_D$ and asymmetric part by $\prec_D$. In a minor departure from [GS], the primitive of our model is a point in $\mathrm{rel}(X)^{\mathbb{D}}$

$$\preceq_{\mathbb{D}} \stackrel{\text{def}}{=} \langle \preceq_D : D \in \mathbb{D} \rangle .^{5}$$

For each C in $\mathbb{D}^{\mathfrak{f}} - \mathbb{D}$, the fact that for some $c \in [\mathfrak{f}]$, $c \in C$ means that $\preceq_C$ is indeterminate, free variable in $\mathrm{rel}(X)$. Although, in isolation each such $\preceq_C$ is free, when the axioms we introduce hold, the potential values of the variable $\preceq_{\mathbb{D}^{\mathfrak{f}}} \stackrel{\text{def}}{=} \langle \preceq_C : C \in \mathbb{D}^{\mathfrak{f}} \rangle$ are constrained by the current values of the constant $\preceq_{\mathbb{D}}$.

Case types. As in [GS], two past cases $c, d \in \mathbb{C}$ are of the same *case type* if, and only if, the marginal information of c is everywhere equal to the marginal information of d. Formally, $c \sim^{\star} d$ if, and only if, for every $D \in \mathbb{D}$ such that $c, d \notin D$, $\preceq_{D \cup \{c\}} = \preceq_{D \cup \{d\}}$. By observation 1 of [GS], $\sim^{\star}$ is an equivalence relation on $\mathbb{C}$ and as its collection of equivalence classes generate a partition $\mathbb{T}$ of $\mathbb{C}$.

We extend $\sim^{\star}$ to $\mathbb{C}^{\mathfrak{f}}$ by taking $[\mathfrak{f}]$ to be an equivalence class of its own, so that, for every $c \in \mathbb{C}$, $c \not\sim^{\star} \mathfrak{f}$. We let $\mathbb{T}^{\mathfrak{f}}$ denote the corresponding partition of $\mathbb{C}^{\mathfrak{f}}$.

Like [GS], we also extend $\sim^{\star}$ to $\mathbb{D}^{\mathfrak{f}}$ by treating databases that contain the same number of each case type as equivalent. That is, $C \sim^{\star} D$ if, and only if, for every $t \in \mathbb{T}^{\mathfrak{f}}$, the numbers $\sharp(C \cap t)$ and $\sharp(D \cap t)$ of that case type coincide. To enable a translation of each database to counting vectors $t \mapsto \sharp(D \cap t)$, we impose a

Richness Assumption. *For every $t \in \mathbb{T}^{\mathfrak{f}}$, there are infinitely many cases in t.*

Our key definition is the following.

Definition 1. *$\mathcal{R} \stackrel{\text{def}}{=} \langle \mathcal{R}_D : D \in \mathfrak{D} \rangle$ is an* extension, *and in particular a Y-extension, of $\preceq_{\mathbb{D}}$ if, for some nonempty $Y \subseteq X$, the following all hold:*

1. *for every $D \in \mathfrak{D}$, $\mathcal{R}_D$ belongs to* $\mathrm{rel}(Y)$,
2. *for every $D \in \mathbb{D}$ and every $x, y \in Y$, $x \, \mathcal{R}_D \, y$ if, and only if, $x \preceq_D y$,*
3. *for every $D \in \mathfrak{D}$ and every $c, d \in \mathfrak{C} - D$, if $c \sim^{\star} d$ then $\mathcal{R}_{D \cup \{c\}} = \mathcal{R}_{D \cup \{d\}}$.*

An extension $\mathcal{R}_{\mathfrak{D}}$ is proper *if $\mathfrak{D} = \mathbb{D}^{\mathfrak{f}}$ and improper otherwise.*

[5]The present approach is equivalent to taking $\preceq_{\mathbb{D}} = \{D \times \preceq_D : D \in \mathbb{D}\}$. This way we maintain pairwise distinctness of $\preceq_C = \preceq_D$ such that $C \neq D$. I thank Maxwell B. Stinchcombe for bringing this point to my attention.

By assigning rankings to the free case $\mathfrak{f}$, (potential) extensions of $\preceq_{\mathbb{D}}$ simulate the arrival of novel information. Part 1 of the definition of an extension ensures that, for every proper Y-extension $\mathcal{R}$ and every $D \in \mathbb{D}^{\mathfrak{f}}$, $\mathcal{R}_D$ is a well-defined binary relation on Y. For every extension $\mathcal{R}_{\mathfrak{D}}$ and every $D \in \mathfrak{D}$, let $\mathcal{I}_D$ and $\mathcal{P}_D$ respectively denote the symmetric and asymmetric parts of $\mathcal{R}_D$.

Part 2 of the definition implies that, for every Y-extension $\mathcal{R}$ and every $D \in \mathbb{D}$, $\mathcal{R}$ is simply the restriction $\preceq_D \cap Y^2$ of $\preceq_D$ to Y. We therefore refer to part 2 of the definition of an extension as the preservation or *nonrevision* condition.

Part 3 of the definition ensures that, for proper extensions $\mathcal{R}$, the partition $\mathbb{T}^{\mathfrak{f}}$ of case types generated by $\sim^{\star}$ is at least as fine the partition generated by the equivalence relation generated by $\mathcal{R}$. Two cases $c, d \in \mathbb{C}^{\mathfrak{f}}$ are *equivalent with respect to* $\mathcal{R}$, written $c \sim^{\mathcal{R}} d$, if, for every $D \in \mathbb{D}$ such that $c, d \notin D$, $\mathcal{R}_{D\cup\{c\}} = \mathcal{R}_{D\cup\{d\}}$. This notion allows us to partition the set of proper extensions as follows.

Definition. *A proper extension $\mathcal{R}$ is either* regular *or* novel. *It is novel whenever $[\mathfrak{f}]$ is a distinct equivalence class of $\sim^{\mathcal{R}}$, so that, for every $c \in \mathbb{C}$, $c \not\sim^{\mathcal{R}} \mathfrak{f}$.*

For novel extensions, $\mathfrak{f}$ mimmicks the potential arrival of new information. Yet novel extensions need not feature qualitatively new rankings (i.e. rankings that do not feature in $\preceq_{\mathbb{D}}$). In lemma 2.1, we show that this is because a novel case type is characterised by the quantitative notion of a similarity weight.

For every regular extension $\mathcal{R}$, there exists $c \in \mathbb{C}$ such that $c \sim^{\mathcal{R}} \mathfrak{f}$. Thus, there are as many regular extensions as there are past case types (i.e. $\sharp\mathbb{T}$). Yet every regular Y-extension $\mathcal{R}$ is *equivalent* to the unique improper Y-extension $\langle \preceq_D \cap Y^2 : D \in \mathbb{D}\rangle$ in the sense of

Observation 1. *For every regular Y-extension $\mathcal{R}$ and improper Y-extension $\acute{\mathcal{R}}$, for every $C \in \mathbb{D}^{\mathfrak{f}}$, there exists $D \in \mathbb{D}$ such that $C \sim^{\mathcal{R}} D$ and $\mathcal{R}_C = \acute{\mathcal{R}}_D$.*[6]

Proof of observation 1. Fix $Y \subseteq X$ nonempty and $\acute{\mathcal{R}}$ regular. W.l.o.g., take $C \in \mathbb{D}^{\mathfrak{f}} - \mathbb{D}$, so that C contains at least one copy of $\mathfrak{f}$. For any $c \in C \cap [\mathfrak{f}]$, the fact that $\acute{\mathcal{R}}$ is regular implies that $c \sim^{\acute{\mathcal{R}}} c_1$ for some $c_1 \in \mathbb{C}$. The richness assumption ensures that we may choose c_1 from the complement of C. Then, since neither c

[6]This means that there is a canonical embedding of $\{C \times \mathcal{R}_C : C \in \mathbb{D}^{\mathfrak{f}}\}$ in $\{D \times \acute{\mathcal{R}}_D : D \in \mathbb{D}\}$. The converse embedding follows from the nonrevision condition of definition 1.

nor c_1 belong to $C_1 \overset{\text{def}}{=} C - \{c\}$, $c \sim^{\acute{\mathcal{R}}} c_1$ implies $\acute{\mathcal{R}}_C = \acute{\mathcal{R}}_{C_1 \cup \{c_1\}}$. If c is the unique member of $C \cap [\mathfrak{f}]$, then the proof is complete. Otherwise, using the fact that C is finite, we may proceed by induction until we obtain a set C_n such that $C_n \cap [\mathfrak{f}]$ is empty and $D \overset{\text{def}}{=} C_n \cup \{c_1, \ldots, c_n\}$ belongs to $\mathbb{D}$. Part 2 of definition 1 then implies $\acute{\mathcal{R}}_D = \preceq_D \cap Y^2$, so that, since $\mathcal{R}$ is improper, $\acute{\mathcal{R}}_D = \mathcal{R}_D$. Finally, since $C \sim^{\acute{\mathcal{R}}} D$, $\acute{\mathcal{R}}_C = \acute{\mathcal{R}}_D$, as required. □

3. Axioms and main theorem

The basic axioms of [GS], which we rewrite in terms of extensions, are the following. In each of these axioms, $\mathcal{R}$ is an arbitrary Y-extension of $\preceq_{\mathbb{D}}$.

A0 (Transitivity axiom for $\mathcal{R}$). *For every $D \in \mathfrak{D}$, $\mathcal{R}_D$ is transitive.*

A1 (Completeness axiom for $\mathcal{R}$). *For every $D \in \mathfrak{D}$, $\mathcal{R}_D$ is complete.*

A2 (Combination axiom for $\mathcal{R}$). *For every disjoint $C, D \in \mathfrak{D}$ and every $x, y \in Y$, if $x \; \mathcal{R}_C \; y$ and $x \; \mathcal{R}_D \; y$, then $x \; \mathcal{R}_{C \cup D} \; y$; and if $x \; \mathcal{P}_C \; y$ and $x \; \mathcal{R}_D \; y$, then $x \; \mathcal{P}_{C \cup D} \; y$.*

A3 (Archimedean axiom for $\mathcal{R}$). *For every disjoint $C, D \in \mathfrak{D}$ and every $x, y \in Y$, if $x \; \mathcal{P}_D \; y$, then there exists $k \in \mathbb{Z}_{++}$ such that, for every pairwise disjoint collection $\left\{D_j : \; D_j \sim^{\mathcal{R}} D \text{ and } C \cap D_j = \varnothing\right\}_1^k$ in $\mathfrak{D}$, $x \; \mathcal{P}_{C \cup D_1 \cup \cdots \cup D_k} \; y$.*

The diversity axioms that now follow require that $\mathbb{D}$ is sufficiently rich to support Y-extensions $\mathcal{R}$ with a variety of *total* orderings: i.e. complete, transitive and antisymmetric ($x \; \mathcal{R}_D \; y$ and $y \; \mathcal{R}_D \; x$ implies $x = y$). Let $\text{total}(\mathcal{R})$ denote the set $\{R : \text{for some } C \in \mathfrak{D},\, R = \mathcal{R}_C \text{ is total}\}$ of of total orders that feature in $\mathcal{R}$. For $k = 4$, the following axiom is a restatement of the diversity axiom of [GS].

Diversity (k-Diversity axiom). *For every $Y \subseteq X$ of cardinality $n = 2, \ldots, k$, every regular Y-extension $\mathcal{R}$ of $\preceq_{\mathbb{D}}$ is such that $\sharp\text{total}(\mathcal{R}) = n!$.*

We say *k-diversity holds on Z* if the axiom holds with Z in the place of X. By lowering the bar for the required number of total orders, the following axiom substantively weakens *3*-diversity and *a fortiori* *4*-diversity.

A4′ (Partial *3*-diversity). *For every $Y \subseteq X$ with cardinality $n = 2$ or 3, every regular X-extension $\mathcal{R}$ of $\preceq_{\mathbb{D}}$ is such that $\sharp\text{total}(\mathcal{R}) \geqslant n$.*

Example 6 of the appendix provides an example of $\leq_{\mathbb{D}}$ satisfying A0–A4′, but not *3*-diversity. Partial-*3*-diversity is clearly stronger than *2*-diversity and, moreover, it plays the dual role of guaranteeing uniqueness of the representation and allowing us to avoid restrictions on the cardinality of X. Example 7 shows that partial-*3*-diversity is the weakest axiom with these properties. Moreover, the observation below shows that, in our setting, partial-*3*-diversity is equivalent to

A4 (Conditional-*2*-diversity)**.** *For every three distinct elements* $x, y, z \in X$*, within one of the sets* $\{D' : x \prec_{D'} y\}$ *and* $\{D' : y \prec_{D'} x\}$ *there exists both* C *and* D *such that* $z \prec_C x$ *and* $x \prec_D z$*. If* $\sharp X = 2$*, then* *2*-diversity *holds on* X*.*

Observation 2. *For* $\leq_{\mathbb{D}}$ *satisfying* A0–A3*,* conditional-*2*-diversity *and* partial-*3*-diversity *are equivalent.*

See proof on page 44.

THE PRUDENCE AXIOM that follows is our final requirement. It is distinguished by the fact that it imposes structure on novel extensions. As we will see in the proof of the main theorem (see lemma 2.1), novel extensions are characterised by a cardinal notion. In contrast, the notions of testworthiness and perturbation that we now introduce are ordinal in nature.

Definition. *A proper extension* $\mathcal{R}$ *of* $\leq_{\mathbb{D}}$ *is* testworthy *if it satisfies* A1–A3 *and, for some* $D \in \mathbb{D}$ *such that* $\mathcal{R}_D$ *is total,* $\mathcal{R}_{\mathfrak{f}}$ *is the inverse of* $\mathcal{R}_D$*.*[7]

We now introduce perturbations. The motivation is that, for any given extension $\mathcal{R}$, the predictor knows $\mathcal{R}_{\mathbb{D}}$ and freely chooses $\mathcal{R}_{\mathfrak{f}}$, but the rankings at other databases are somewhat arbitrary. For although the rankings the predictor associates with members of $\mathbb{D}^{\mathfrak{f}} - (\mathbb{D} \cup \{\mathfrak{f}\})$ are constrained, what matters is not the actual rankings, but rather their potential for consistency with the axioms.

Definition. *Let* $\mathcal{R}$ *and* $\acute{\mathcal{R}}$ *be extensions of* $\leq_{\mathbb{D}}$*.* $\acute{\mathcal{R}}$ *is a* perturbation *of* $\mathcal{R}$ *if* $\acute{\mathcal{R}}_{\mathfrak{f}} = \mathcal{R}_{\mathfrak{f}}$*. Moreover,* $\acute{\mathcal{R}}$ *is a* nondogmatic *perturbation if* $\sharp\text{total}(\acute{\mathcal{R}}) \geqslant \sharp\text{total}(\mathcal{R})$*.*

[7]Recall that the inverse $\mathcal{R}_D^{-1}$ of $\mathcal{R}_D$ satisfies $x\,\mathcal{R}_D^{-1}\,y$ if, and only if, $y\,\mathcal{R}_D\,x$.

We are interested in perturbations that are nondogmatic because may reveal intransitivities that the predictor's inexperience conceals. The prudent predictor may exploit them and revise her model *before new cases arrive.*

$\mathit{4}$-Prudence. *For every $Y \subseteq X$ with cardinality 3 or 4, every testworthy Y-extension of $\preceq_{\mathbb{D}}$ that is novel has a nondogmatic perturbation that satisfies* A0–A3*.*

Given that the extensions we consider are nonrevisionistic, it is natural to ask whether A0–A3 are superfluous in the presence of $\mathit{4}$-prudence. Our response is twofold. Firstly, in practice, we would expect A0–A3 to hold more frequently than $\mathit{4}$-prudence which is more cognitively demanding. Secondly, as we will see in the proof of the main theorem, when $\mathbb{T}$ is infinite, for some $Y \subseteq X$, the set of testworthy Y-extensions that are novel may be empty. For every such Y, the following observation confirms that $\mathit{4}$-diversity holds on Y.

Observation 3 (on testworthy extensions)**.** *Let $\preceq_{\mathbb{D}}$ satisfy* A0–A4*. For every $Y \subseteq X$ of cardinality 3 or 4, the set of testworthy Y-extensions is nonempty. If, for some Y, every testworthy Y-extension is regular, then* $\mathit{4}$-diversity *holds on Y.*

See proof on page 44. It is also natural to ask whether $\mathit{4}$-prudence is simply requiring that, on Y such that $\preceq_{\mathbb{J}}$ fails to satisfy $\mathit{4}$-diversity, there exists a testworthy Y-extension that is novel and satisfies $\mathit{4}$-diversity. In the proof of the theorem that now follows, we show that this is not the case.

The main theorem that now follows involves real-valued function $\mathbf{v}$ on the product $X \times \mathbb{C}$. We view $\mathbf{v}$ as a matrix and $\mathbf{v}(x, \cdot)$ as one of its rows. The matrix $\mathbf{v}$ is a *representation of* $\preceq_{\mathbb{D}}$ whenever it satisfies

$$\begin{cases} \text{for every } x, y \in X \text{ and every } D \in \mathbb{D}, \\ x \preceq_D y \quad \text{if, and only if,} \quad \sum_{c \in D} \mathbf{v}(x, c) \leqslant \sum_{c \in D} \mathbf{v}(y, c). \end{cases}$$

The matrix $\mathbf{v}$ *respects case equivalence* (with respect to $\preceq_{\mathbb{D}}$) if, for every $c, d \in \mathbb{C}$, $c \sim^{\star} d$ if, and only if, the columns $\mathbf{v}(\cdot, c)$ and $\mathbf{v}(\cdot, d)$ are equal.

Theorem 1 (Part I, Existence)**.** *Let there be given X, $\mathbb{C}^{\mathfrak{f}}$, $\preceq_{\mathbb{D}}$ and associated extensions, as above, such that the richness condition holds. Then* (1.i) *and* (1.ii) *are equivalent.*

(1.i) A0–A4 *and* 4-prudence *hold for* $\preceq_{\mathbb{D}}$.

(1.ii) *There exists a matrix* $\mathbf{v} : X \times \mathbb{C} \to \mathbb{R}$ *satisfying* (1.a) *and* (1.b):

(1.a) $\mathbf{v}$ *is a representation of* $\preceq_{\mathbb{D}}$ *that respects case equivalence*;

(1.b) *no row of* $\mathbf{v}$ *is dominated by any other row, and for every three distinct elements* $x, y, z \in X$ *and* $\lambda \in \mathbb{R}$, $\mathbf{v}(x, \cdot) \neq \lambda \mathbf{v}(y, \cdot) + (1 - \lambda)\mathbf{v}(z, \cdot)$.[8]

Our uniqueness result is identical to that of [GS].

Theorem 1 (Part II, Uniqueness)**.** *If* (1.i) [*or* (1.ii)] *holds, then the matrix* $\mathbf{v}$ *is unique in the following sense*: *for every other matrix* $\mathbf{u} : X \times \mathbb{C} \to \mathbb{R}$ *that represents* $\preceq_{\mathbb{D}}$, *there is a scalar* $\lambda > 0$ *and a matrix* $\beta : X \times \mathbb{C} \to \mathbb{R}$ *with identical rows (i.e. with constant columns) such that* $\mathbf{u} = \lambda \mathbf{v} + \beta$.

The proof of theorem 1 appears in online appendix A. It relies upon a translation from the abstract database/memory set up of the model to the setting of rational vectors similar to [GS]. We show that the translated theorem theorem 2 (see section A) is equivalent to theorem 1. The proof of theorem 2 is then the subject of online appendix B.

We appeal to the following corollary when we connect the present framework to the concept of arbitrage in finance. Central to this connection is

Definition 2. *For* $Y \in 2^X$, *the matrix* $v^{(\cdot,\cdot)} : Y^2 \times \mathbb{C} \to \mathbb{R}$ *satisfies the* Jacobi identity *whenever, for every* $x, y, z \in Y$, *the rows satisfy* $v^{(x,z)} = v^{(x,y)} + v^{(y,z)}$.

In lemma 2.1, of section B, we show that, for a given extension $\mathcal{R}$, A1–A3 and 2-diversity yield a *pairwise representation* $v^{(\cdot,\cdot)}$ of $\mathcal{R}$.[9]

Corollary 1 (a characterisation of prudence)**.** *Let the number of case types be finite and let* $\preceq_{\mathbb{D}}$ *satisfy* A4. *Then* $\preceq_{\mathbb{D}}$ *satisfies* 4-prudence *if, and only if,* $\preceq_{\mathbb{D}}$ *has a pairwise representation* $v^{(\cdot,\cdot)}$ *that satisfies the Jacobi identity. Moreover, for every other pairwise representation* $u^{(\cdot,\cdot)}$, *there exists* $\lambda > 0$, *such that* $u^{(\cdot,\cdot)} = \lambda v^{(\cdot,\cdot)}$.

See proof on page 44.

[8] Observe that $\mathbf{v}(x, \cdot) - \mathbf{v}(z, \cdot)$ and $\mathbf{v}(y, \cdot) - \mathbf{v}(z, \cdot)$ are noncollinear if, and only if, the affine independence condition of (1.b) holds.

[9] That is, for every $x, y \in X$ and $D \in \mathbb{D}$, $x \, \mathcal{R}_D \, y$ if, and only if $\sum_{c \in D} v^{(x,y)}(c) \geqslant 0$.

4. Discussion

We begin by restating the existence part of the main theorem of [GS].

Theorem (Existence)**.** *Let there be given X, $\mathbb{C}$ and $\preceq_{\mathbb{D}}$, as above, such that the richness condition holds. Then* (i) *and* (ii) *are equivalent.*

(i) A0–A3 *and* 4-diversity *hold for* $\preceq_{\mathbb{D}}$.

(ii) *There exists a matrix* $\mathbf{v} : X \times \mathbb{C} \to \mathbb{R}$ *satisfying* (a) *and* (b) :

  (a) $\mathbf{v}$ *is a representation of* $\preceq_{\mathbb{D}}$ *that respects case equivalence*;

  (b) *if* $\sharp X < 4$, *then no row is dominated by an affine combination of the other rows, and for every four distinct elements* $x, y, z, w \in X$ *and every* $\lambda, \mu, \theta \in \mathbb{R}$ *such that* $\lambda + \mu + \theta = 1$, $\mathbf{v}(x, \cdot) \nleq \lambda\mathbf{v}(y, \cdot) + \mu\mathbf{v}(z, \cdot) + \theta\mathbf{v}(w, \cdot)$.

Although diversity axioms play an important technical role, they are not obviously behavioural. Instead, diversity axioms impose restrictions on what is beyond the predictor's control and on what is central to inductive inference: experience. Our main contention is that $C^{\star}$ may not be so rich as to support $\preceq_{\mathbb{D}}$ satisfying 4-diversity. That is to say, there may exist $Y \subseteq X$ such that $\sharp Y = 4$, and such that the data is insufficiently rich to support all $4! = 24$ strict rankings. A casual comparison of condition (1.b) and (b) confirms that the present framework achieves the main purpose of accommodating the less experienced.

For the remainder of this discussion, we take both X and the set $\mathbb{T}$ of case types to be finite and of cardinality m and n respectively. Via case equivalence, for any $\preceq_{\mathbb{D}}$ (satisfying (1.ii) or (ii)) we may efficiently summarise rankings using a real-valued $m \times n$ matrix $\mathbf{v}$ on the product of X and case types $\mathbb{T}$.

Comparing the complexity of conditional-2-diversity and 4-diversity, in the presence of the other axioms, provides a measure of the value of experience. As an estimate, we compare condition (1.b) with (b) of Gilboa and Schmeidler's theorem. Verifying (b) involves checking n affine dominance constraints: one for each case type. This is well-known to be equivalent to the complexity of linear programming with real variables (**?**). In the absence of knowledge regarding the sparsity of $\mathbf{v}$, the fastest algorithm for achieving this is of order n^3 (**?**). Since this holds for every subset of four distinct eventualities, checking (b) is of order $\binom{m}{4}n^3$. In contrast, verifying $\mathbf{v}(x,\cdot) \nleqslant \mathbf{v}(y,\cdot)$ takes at most n steps for each of the $\binom{m}{2}$ subsets of 2 distinct elements. Likewise, for every three distinct elements x, y, z in X, checking for noncollinearity of two vectors takes at most n steps. Thus, a naïve algorithm for checking (1.b) is of order $\binom{m}{3}n$. Even for $m = 4$ and $n = 4$, the difference is stark: $\binom{m}{3}n = 16$ versus $\binom{m}{4}n^3 = 64$. (The threshold 4 is important as we show in proposition 1 below.)

A robust test of prudence is available precisely when A0–A4 hold and *4*-diversity fails. For, when $n < \infty$, the existence of a similarity representation satisfying item (1.ii) is equivalent to *4*-prudence. The test is robust in that, generically, predictors that fail to check for the arrival of new cases also fail to satisfy item (1.ii). This is because, the row differences $v^{(x,y)} = -\mathbf{v}(x,\cdot) + \mathbf{v}(y,\cdot)$, the Jacobi identity $v^{(x,z)} = v^{(x,y)} + v^{(y,z)}$ holds for every $x, y, z \in X$. This in turn implies that the hyperplanes $H^{\{x,y\}}$, $H^{\{y,z\}}$ and $H^{\{x,z\}}$, to which the row differences are normal, are congruent. Congruence implies the hyperplanes are not in general position. (When $n = 2$, the same argument can be made in terms of extensions: see eq. (16) of the proof of theorem 3.) In other words, there is a zero (conditional) probability of the predictor striking lucky and appearing to be prudent when in fact they are not engaging in this form of second order inductive inference.

Novelty, memory and transitivity. An early test and taxonomy of intransitive behaviour is due to **??**. Weinstein points out that intransivity can sometimes be rational in complex situations. Interestingly, he shows that young are people significantly less transitive in their choices. He also points out that the law itself is designed to accommodate irresponsible under-age decisions. In the psychology literature, **?** show that presenting novel objects is more likely to trigger intransitive choices.

More recently, **?** provide evidence that people with a specific form of memory impairment (lesions in the hippocampus of the brain) are significantly more likely to violate transitivity in pairwise choices of chocolate bars: even though they rank numbers transitively.[10] Whilst this literature does not yet offer a direct test of the present model, it supports the case-based framework of constructing preference as well as our premise that violations of transitivity are often driven by novelty, or equivalently impaired memory.

In line with the case-based approach, the experimental evidence of **?** suggests that agents are constructing their preferences on the basis of past experience. Moreover, it seems natural to interpret agents with hippocampal impairment as inexperienced predictors. The fact that impaired agents then make intransitive decisions is very much in line with what our model predicts as they are, in effect, facing a novel situation and are required to construct their preferences on the fly. It appears that impaired agents are also failing to be prudent, though it is not obvious that chocolates warrant the additional neural computation that accompanies *4*-prudence.

Remark. *A closer look at the relationship between the proportion ρ of hippocampal impairment and the percentage σ of intransitive choices (in figure 2 of* **?***) suggests another interpretation. For ρ above $\frac{1}{4}$, σ is above* 20%: *twice as high as it is for $\rho < \frac{1}{4}$. Memories and the rankings they generate are latent variables to the observable ρ and this threshold is where* conditional-*2*-diversity *fails to hold and our model breaks down. The* 16 *cases in the data are thus partitioned into three groups*: *two that satisfy 3*-diversity *($\rho < 0.05$ and $\sigma < 5\%$)*; 12 *intermediate cases that satisfy* conditional-*2*-diversity *($0.05 \leqslant \rho \leqslant 0.25$ and $5\% \leqslant \sigma \leqslant 10\%$); and* 2 *severe cases that fail to satisfy* conditional-*2*-diversity *($0.25 < \rho$ and $10\% < \sigma$).*

The success or failure of startups. Inexperience raises significant barriers to entry. Overcoming these barriers is either the result of making mistakes and learning by doing "on the fly" or the result of being prudent. Which form of second-order induction bears out in practice will depend on many factors.

The following proposition confirms our thesis that experienced predictors (i.e. those that satisfy *4*-diversity) have indeed encountered a high number of case

[10]The hippocampus is associated with learning and memory. **?** and **?** present evidence showing that the hippocampus plays an important role in imagining future experiences on the basis of past ones. **?** go further by showing that it plays a role in the value-based decision making framework of **?**.

types. As the main theorem of [GS] shows, experienced predictors have no need for the additional structure of extensions. Unless the prediction problem changes (e.g. new eventualities become relevant), they have no need to engage in second order induction. This saving in cognitive effort is the prize that experience confers.

Proposition 1 (experience and case types)**.** *If* $\preceq_{\mathbb{D}}$ *satisfies* A1–A3 *and* *4*-diversity, *then* $n \geqslant \min\{4, m\}$*, and, for every* $Y \subseteq X$ *of cardinality* m' *and regular* Y*-extension* $\mathcal{R}$*, the number* n' *of equivalence classes of* $\sim^{\mathcal{R}}$ *satisfies* $n' \geqslant \min\{4, m'\}$*.*

See proof on page 44. In contrast, conditional-*2*-diversity implies no restrictions on the cardinality of $\mathbb{T}$ beyond $n \geqslant 2$ and this is also a virtue of *2*-diversity.

WHEN IS PRUDENCE WORTH THE TROUBLE? The simple answer to this question is: when revising a model "on the fly", once a novel case arrives, is costly. The following is our main example of such a setting.

Example 5. *Consider a fair market maker of zero-coupon (treasury) bonds.*[11] *The compound-interest formula for the accumulation process of such a bond is*

$$a^{(x,y)} = \left(1 + r^{\{x,y\}}\right)^{-x+y},$$

where $r^{\{x,y\}}$ *is the implied yield on a forward contract that accrues interest between dates* x *and* y*. If* x *is later than* y*, then the contract is to sell, and the market maker pays this yield, so that* $r^{\{y,x\}} = r^{\{x,y\}}$*. Let* $X \subseteq \mathbb{R}_+$ *index a suitable sequence of trading dates with* $0 \in X$ *being the spot date. It is well known that the (normalised) spot bond price* $x \mapsto b(x) = \left(1 + r^{\{0,x\}}\right)^{-x}$ *is arbitrage-free if, and only if, the log-accumulation process satisfies*

$$\text{for every } x, y, z \in X, \quad \log a^{(x,y)} = \log a^{(x,z)} + \log a^{(z,y)}.^{12} \tag{2}$$

This no-arbitrage condition is a special case of the Jacobi identity. We now explain how a market maker might infer the accumulation process from past cases.

---

[11] Similar to a fair insurer, the fair market maker sets the market spread to zero.

[12] To see this, suppose that, for some $x < y$, she sets $a^{(x,y)} < a^{(x,z)}a^{(z,y)}$. Another trader would do well to sell the forward contract (x, y), buy the spot contract (x, z) and sell the spot contract (z, y). A risk-free arbitrage opportunity is also available if the reverse inequality holds.

Let each case $c \in D^\star$ consist of market-relevant data from a given time interval (a block of time periods) in the past. These blocks are chosen so that, for any finite resampling D of cases from $D^\star$, the sequence of cases that makes up D is exchangeable. The free case $\mathfrak{f}$ has no additional structure beyond that of section 2. Next, for every finite resampling D and every date x and y, let $x \preceq_D y$ if, and only if, in answer to the question "At which date will the price be higher?" the market maker finds that y is more plausible than x.

For the following corollary, we introduce the empirical implied yield function that maps triples $x \times y \times c$ to $r_c^{\{x,y\}} \in \mathbb{R}$. This function is characterised by three conditions: for time intervals of length zero, the yield is zero; fair pricing; and case equivalence. These are respectively formalised as follows: for every $x, y \in X$ and every $c, d \in \mathbb{C}$, $r_c^{\{x,x\}} = 0$; $r_c^{\{y,x\}} = r_c^{\{x,y\}}$; and $c \sim^\star d$ if, and only if, $r_c^{\{x,y\}} = r_d^{\{x,y\}}$. The empirical bond price function $B : X \times \mathbb{D} \to \mathbb{R}$ maps every pair $x \times D$ to

$$B(x, D) = \prod_{c \in D} \left(1 + r_c^{\{0,x\}}\right)^{-\frac{x}{|D|}} .$$

When $D^\star$ belongs to $\mathbb{D}$, the number of case types is finite, and we have

Corollary 2. *Let $\preceq_{\mathbb{D}}$ satisfy A4. Then $\preceq_{\mathbb{D}}$ satisfies $\not{4}$-*prudence *if, and only if, there exists empirical implied yield and empirical bond price functions, such that*

$$\begin{cases} \text{for every } x, y \in X \text{ and every } D \in \mathbb{D}, \\ x \preceq_D y \quad \text{if, and only if,} \quad B(x, D) \leqslant B(y, D). \end{cases} \tag{$*$}$$

Moreover, for every $D \in \mathbb{D}$, the bond price $B(\cdot, D)$ is arbitrage-free.

Proof of corollary 2. Via theorem 3, there exists a pairwise Jacobi representation $v^{(\cdot,\cdot)} : X^2 \times \mathbb{C} \to \mathbb{R}$ such that, for every $c, d \in \mathbb{C}$, $v^{(\cdot,\cdot)}(c) = v^{(\cdot,\cdot)}(d)$ if, and only if, $c \sim^\star d$. Moreover, $v^{(\cdot,\cdot)}$ is unique upto multiplication by a positive scalar. Then, for every $x, y, z \in X$, $v^{(x,x)}(\cdot) = 0$, $v^{(y,x)} = -v^{(x,y)}$, and $v^{(x,y)} = v^{(x,z)} + v^{(z,y)}$.

Recalling that $0 \in X$, for every $x \in X$ and $D \in \mathbb{D}$, let

$$B(x, D) = \exp\left(-\tfrac{1}{|D|} \textstyle\sum_{c \in D} v_c^{(x,0)}\right). \tag{3}$$

For the proof of $(*)$, recall that $x \preceq_D y$ if, and only if, $\sum_{c \in D} v^{(x,y)}(c) \geqslant 0$. Since $-v_c^{(y,0)} = v_c^{(0,y)}$ and, via the Jacobi identity, $v^{(x,0)} + v^{(0,y)} = v^{(x,y)}$, we have:

$$-\log B(x, D) + \log B(y, D) = \tfrac{1}{|D|} \textstyle\sum_{c \in D} \left(v_c^{(x,0)} - v_c^{(y,0)}\right) = \tfrac{1}{|D|} \textstyle\sum_{c \in D} v_c^{(x,y)}.$$

It remains for us to confirm that the bond price is a suitable function of the empirical yield function. For every $x, y \in X$ and $c \in \mathbb{C}$, $\log a_c^{(x,y)} = -v^{(x,y)}(c)$: so that, as the solution to $(y-x)\log(1+r_c^{\{x,y\}}) = -v_c^{(x,y)} = v_c^{(y,x)}$, for $x \neq y$,

$$1 + r_c^{\{x,y\}} = \exp\left(\frac{v_c^{(y,x)}}{y-x}\right) = \exp\left(\frac{v_c^{(x,y)}}{x-y}\right) = 1 + r_c^{\{y,x\}}. \tag{4}$$

We therefore observe that $r_c^{\{x,y\}} = r_c^{\{y,x\}}$. For $x = y$, $v_c^{(x,y)} = 0$ ensures that we can take $r_c^{\{x,x\}} = 0$. Finally, note that for $c \sim^\star d$, the property $r_c^{\{x,y\}} = r_c^{\{x,y\}}$ is inherited from $v_c^{(x,y)} = v_d^{(x,y)}$, so that we have an empirical implied yield function.

The fact that, for every D, $B(\cdot, D)$ is arbitrage-free follows by virtue of the fact that $v^{(\cdot,\cdot)}$ satisfies the Jacobi identity. □

In the present setting, because we are modelling a normalised bond price, we obtain a stronger uniqueness result relative to part II of theorem 1. If $\tilde{B}$ is another function that satisfies the present corollary, then, for every $D \in \mathbb{D}$, the spot price $\tilde{B}(0, D) = 1$. Thus, via eq. (3) and corollary 1, for some $\lambda > 0$, $\tilde{B} = \mathrm{e}^{-\lambda} B$.

We now point out an interesting implication of conditional-*2*-diversity *the recent prevalence of negative interest rates. W.l.o.g., fix $x < y$. Then, given* *4*-prudence, *2*-diversity *implies that there exists $c, d \in \mathbb{C}$ such that $v^{(x,y)}(c) < 0 < v^{(x,y)}(d)$. This is equivalent to $v^{(y,x)}(d) < 0 < v^{(y,x)}(c)$, and, via eq. (4), $r_d^{\{x,y\}} < 0 < r_c^{\{x,y\}}$. That is,* *2*-diversity *requires that the market maker's data is rich enough to contain at least one case where the yield between date x and y is negative (as well as one where it is positive).* Conditional-*2*-diversity *extends this notion to require that $r_D^{\{x,y\}} < 0 < r_C^{\{x,y\}}$ for some C and D such that $r_C^{\{x,z\}} \cdot r_D^{\{x,z\}} > 0$.*

DISCUSSION OF SECOND-ORDER INDUCTION. The market maker of example 5 engages in second-order induction when she acts prudently. She reflects on her model by checking that the basic axioms of [GS] will continue to hold when a novel case arrives. By way of contrast, suppose the bond price of the market maker is such that $\preceq_{\mathbb{D}}$ is consistent with the basic axioms, but not *4*-prudence. Then when a novel case arrives, she may be exposed to arbitrage and need to respecify her entire model "on the fly". Such a step corresponds to the intermittent respecification of her similarity weighting function $\mathbf{v}(x, c)$, that **?**, p.10324 describe. In **?**, the "leave-one-out" technique of cross-validating the model by omitting a case of each type is intuitively and operationally close to our inclusion of the free case $\mathfrak{f}$. The difference is that by allowing $\mathfrak{f}$ more degrees of freedom, our market maker can study novel extensions and peer into the future through the lens of her current model. She can exploit the intervals of time inbetween the arrival of novel cases by continuously engaging in second-order induction.

Through an example, we now show that the present framework provides the flexibility to accommodate second-order induction without sacrificing the computational or normative advantages that additive similarity functions provide.

Example (second-order induction, [GS], p.12)**.** *Let c denote a case where Mary chooses restaurant x over restaurant y. In the absence of any further information, it is tempting to assume some similarity between John and Mary. The predictor then finds it plausible that John prefers x to y given $\{c\}$. A separate database D contains no choices between x and y. Thus, in the absence of further information, x and y appear equally likely based on D. Additivity of the similarity function (or* A2*)) implies it is plausible that John prefers x to y given $\{c\} \cup D$. The violation of* A2 *arises when a more careful examination of the contents of D reveals many choices between other pairs of restaurants where John and Mary consistently differ.*

Quine's notion of perceptual similarity (**?**) offers a check on the predictor's inference about John's choice given $\{c\}$. John and Mary may just as well be two drivers passing through an intersection at different times. Although their situations are broadly speaking very similar, if one faces a red light and the other a green light, their responses will differ. In the restaurant setting, some pivotal information is omitted from c. Observing that the evidence in c in favour of x over y is somewhat weak, a prudent predictor instead recasts $\{c\}$ as a database C that combines past observations with copies of the pivotal novel case $\mathfrak{f}$. With a more refined model, that explicitly allows for omitted variables, the predictor can check to see if her model extends to higher dimensions without violating the basic axioms.

Remark 1. *Observe that predictors that only fail to satisfy* A2 *can, with some additional regularity conditions, still be represented by a nonlinear function $u : X \times \mathbb{D} \to \mathbb{R}$ such that for every $D \in \mathbb{D}$ and every $x, y \in X$, $x \preceq_D y$ if, and only if,*

$u(x, D) \leqslant u(y, D)$ *(see* **?***). But predictors that satisfy* A0–A4 *but not 4*-prudence *can also be represented by such a function: because* $\preceq_{\mathbb{D}}$ *is complete and transitive for each* $D \in \mathbb{D}$*. The present framework allows us to disentangle the latter kind of predictor from those who, for good reason, fail to satisfy the combination axiom. (See* [GS] *for examples of such reasons.)*

ON THE VERACITY OF FALSE NEWS. How should a predictor check whether her model consistently extends to higher dimensions (when novel cases arrive)? If our model is a guide then, the most useful rankings that she might wish to examine are those that are far from her own. This is because our definition of testworthy extensions involves assigning to the novel case $\mathfrak{f}$ the inverse of some total ranking $\preceq_D$. This may offer some rationale for why information that differs from our own is intrinsically valuable. Testworthy extensions play a vital role in taming the complexity of our proof. It seems plausible that something similar may be at play when agents encounter radically different information from their own on social media: *even if it is fake.* This may help to explain why false news is significantly more veracious than real news online (**?**). The fact that real news is typically closer to what we have observed in the past means that it is of less value to the prudent predictor that finds it costly to imagine worlds that are far from her own.

APPENDIX A. THE PROOF OF THEOREM 1

As in [GS], we translate the model into one where databases are represented by vectors, the dimensions of which are case types. To allow us to focus on aspects of the present model, proceed directly to rational vectors and present the axioms and a corresponding theorem (theorem 2) which, as we confirm, holds if, and only if, theorem 1 does. The proof of theorem 2 can be found in section B.

CASE TYPES AS DIMENSIONS. From our definition of case types in section 2, $\mathbb{T} = \mathbb{C}_{/\sim^\star}$ and $\mathbb{T}^{\mathfrak{f}} \overset{\text{def}}{=} \mathbb{T} \cup [\mathfrak{f}]$. Let $\mathfrak{T}$ be a free variable in $\{\mathbb{T}, \mathbb{T}^{\mathfrak{f}}\}$. When no possible confusion should arise, we use $\mathfrak{f}$ as shorthand for $[\mathfrak{f}]$. It is straightforward to show that the following construction would work if instead we were to work with any partition T of $\mathbb{C}$ that is at least as fine as $\mathbb{T}$. The present construction is the one with the lowest feasible number $\sharp\mathbb{T}$ of dimensions.

Translation to counting vectors. Let $\mathbb{Z}_+$ denote the set of nonnegative integers and $\mathbb{Z}_{++}$ those that are (strictly) positive. Let $\mathbb{L} \subseteq \mathbb{Z}_+^{\mathbb{T}}$ denote the set of counting vectors $L : \mathbb{T} \to \mathbb{Z}_+$ such that $\{t : L(t) \neq 0\}$ is finite and let $\mathbb{L}^{\mathfrak{f}}$ denote the corresponding subset of $\mathbb{Z}_+^{\mathbb{T}^{\mathfrak{f}}}$. Then let

$$\mathbb{L} = \begin{cases} \mathbb{L} & \text{if, and only if, } \mathfrak{T} = \mathbb{T}, \text{ and} \\ \mathbb{L}^{\mathfrak{f}} & \text{otherwise.} \end{cases}$$

Modulo notation, the following construction is identical to [GS]. For every $D \in \mathbb{D}$, let $L_D : \mathbb{T} \to \mathbb{Z}_+$ denote the function $t \mapsto L_D(t) = \sharp(D \cap t)$. For each $D \in \mathbb{D}$, let $\preceq_{L_D} \stackrel{\text{def}}{=} \preceq_D$. We need to establish that $\preceq_{\mathbb{L}} \stackrel{\text{def}}{=} \langle \preceq_L : L \in \mathbb{L} \rangle$ is well-defined. For every $L \in \mathbb{L}$, the richness assumption (on $\mathbb{T}^{\mathfrak{f}}$) guarantees the existence of $D \in \mathbb{D}$ such that $L_D = L$. By definition, $\sim^\star$ is such that, for every $C, D \in \mathbb{D}$, $C \sim^\star D$ if, and only if, $L_C = L_D$. Straightforward mathematical induction on the cardinality of C shows that $C \sim^\star D$ implies $\preceq_C = \preceq_D$. This construction of $\preceq_{\mathbb{L}}$ ensures that the same notion of equivalence that we introduced in observation 1 also applies here. Thus, $\preceq_{\mathbb{L}} \equiv \preceq_{\mathbb{D}}$.

Translation to rational vectors. Similarly, let $\mathbb{Q}_+$ denote the nonnegative rationals and $\mathbb{Q}_{++}$ those that are (strictly) positive. Take $\mathbb{J} \subseteq \mathbb{Q}_+^{\mathbb{T}}$ to be the set of rational vectors with $\{t \in \mathbb{T} : J(t) \neq 0\}$ finite and take $\mathbb{J}^{\mathfrak{f}}$ to denote the corresponding subset of $\mathbb{Q}_+^{\mathbb{T}^{\mathfrak{f}}}$. For each $J \in \mathbb{J}$, by virtue of the fact that $\mathbb{Z}_{++}$ is well-ordered and J has finite support, there exists (unique) minimal $k_J \in \mathbb{Z}_{++}$ such that $L_J \stackrel{\text{def}}{=} k_J J$ belongs to $\mathbb{L}$. Let $\preceq_J \stackrel{\text{def}}{=} \preceq_{L_J}$. (This definition acquires meaning below once we translate and apply the combination axiom.) In this way, $\preceq_{\mathbb{J}} = \langle \preceq_J : J \in \mathbb{J} \rangle$ is well-defined, and we may introduce axioms for $\preceq_{\mathbb{J}}$ directly: i.e. without first introducing axioms for $\preceq_{\mathbb{L}}$. We first demonstrate that $\preceq_{\mathbb{J}}$ and $\preceq_{\mathbb{D}}$ are equivalent. First note that, for every $I, J \in \mathbb{J}$ such that $L_I = L_J$, $\preceq_I = \preceq_J$. Then, let $L' = L_J$ and take any D such that $L_D = L'$. Then $\preceq_J = \preceq_D$. The reverse embedding follows by virtue of the fact that $\mathbb{L} \subset \mathbb{J}$. Thus, $\preceq_{\mathbb{J}} \equiv \preceq_{\mathbb{D}}$.

Construction of extensions of $\preceq_{\mathbb{J}}$. We follow common practice by letting 2^X denote the collection of nonempty subsets $Y \subseteq X$. For each $Y \in 2^X$, we will denote the set of regular, novel and testworthy Y-extensions (of $\preceq_{\mathbb{D}}$ or $\preceq_{\mathbb{J}}$) by $\mathrm{reg}(Y,\cdot)$, $\mathrm{nov}(Y,\cdot)$ and $\mathrm{test}(Y,\cdot)$ respectively. Recalling that every Y-extension is either regular or novel, let $\mathrm{ext}(Y,\cdot)$ denote the set of all Y-extensions. We now clarify what it means to be an extension of $\preceq_{\mathbb{J}}$.

For each $t \in \mathfrak{T}$, we take $\delta_t : \mathfrak{T} \to \mathbb{R}$ to be the function satisfying $\delta_t(s) = 1$ if $s = t$ and $\delta_t(s) = 0$ otherwise. (When $\mathfrak{T}$ is finite, these are simply the basis vectors for $\mathbb{R}^{\mathfrak{T}}$.) When we wish to emphasise that the vectors belong to in $\mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}$, then, for each $\mathbb{T}^{\mathfrak{f}}$, we will write $\delta^{\mathfrak{f}}_t$. Let

$$\mathbb{I} = \begin{cases} \mathbb{J} & \text{if, and only if, } \mathfrak{T} = \mathbb{T}, \text{ and} \\ \mathbb{J}^{\mathfrak{f}} & \text{otherwise.} \end{cases}$$

For every $I \in \mathbb{J}$ and $J \in \mathfrak{J}$, we write $I \equiv J$ whenever $I = J$ or $J = I \times 0$. (In the latter case, $J(t) = I(t)$ for every $t \in \mathbb{T}$ and $J(\mathfrak{f}) = 0$.) This notion reflects the fact that, for the purposes of the present model, such I and J are equivalent.

Definition 3. *$\mathcal{R} = \langle \mathcal{R}_J : J \in \mathfrak{J} \rangle$ is an* extension*, and in particular a Y-*extension*, of $\preceq_{\mathbb{J}}$ if, and only if, for some nonempty $Y \subseteq X$ both the following hold*

1. *for every $J \in \mathfrak{J}$, $\mathcal{R}_J \in \mathrm{rel}(Y)$, $\mathcal{I}_J \stackrel{\text{def}}{=} \mathcal{R}_J \cap \mathcal{R}_J^{-1}$ and $\mathcal{P}_J \stackrel{\text{def}}{=} \mathcal{R}_J - \mathcal{R}_J^{-1}$;*
2. *for every $J \in \mathbb{J}$ and $L \in \mathfrak{J}$ such that $J \equiv L$, $\mathcal{R}_L = \preceq_J \cap (Y^2)$.*

An extension $\mathcal{R}_{\mathfrak{J}}$ (of $\preceq_{\mathbb{J}}$) is proper if $\mathfrak{J} = \mathbb{J}^{\mathfrak{f}}$ and improper otherwise. A proper extension is either regular or novel. $\mathcal{R}$ is novel if, for every $s \in \mathbb{T}$, there exists I in $\mathbb{J}$ such that, for $J = I \times 0$ (in $\mathbb{J}^{\mathfrak{f}}$), we have $\mathcal{R}_{J+\delta^{\mathfrak{f}}_s} \neq \mathcal{R}_{J+\delta^{\mathfrak{f}}_{\mathfrak{f}}}$.

For every regular Y-extension $\mathcal{R}$ of $\preceq_{\mathbb{D}}$ such that $Y = X$, observation 1 implies $\mathcal{R} \equiv \preceq_{\mathbb{D}}$. And, via $\preceq_{\mathbb{J}} \equiv \preceq_{\mathbb{D}}$ and transitivity of equivalence, we conclude that $\mathcal{R}$ is equivalent to $\preceq_{\mathbb{J}}$. Two sets of extensions are isomorphic if there exists a canonical isomorphism between equivalent extensions.

Lemma 1.1 (proof on page 44)**.** *For every $Y \in 2^X$, $\mathrm{reg}(Y, \preceq_{\mathbb{J}})$ is isomorphic to $\mathrm{reg}(Y, \preceq_{\mathbb{D}})$ and $\mathrm{nov}(Y, \preceq_{\mathbb{J}})$ is isomorphic to $\mathrm{nov}(Y, \preceq_{\mathbb{D}})$.*

Axioms and theorem. We now restate the axioms for Y-extensions $\mathcal{R}$ of $\preceq_{\mathbb{J}}$.

A0$^\flat$ For every $J \in \mathfrak{J}$, $\mathcal{R}_J$ is transitive on Y.

A1$^\flat$ For every $J \in \mathfrak{J}$, $\mathcal{R}_J$ complete on Y.

A2$^\flat$ For every $I, J \in \mathfrak{J}$, every $x, y \in Y$ and every $\lambda, \mu \in \mathbb{Q}_{++}$, if $x\ \mathcal{R}_I\ y$ and $x\ \mathcal{R}_J\ y$, then $x\ \mathcal{R}_{\lambda I+\mu J}\ y$; moreover, if $x\ \mathcal{P}_I\ y$ and $x\ \mathcal{R}_J\ y$, then $x\ \mathcal{P}_{\lambda I+\mu J}$.

A3$^\flat$ For every $I, J \in \mathfrak{J}$ and every $x, y \in Y$ if $x\ \mathcal{P}_J\ y$, then there exists $0 < \lambda < 1$ such that, for every $\mu \in \mathbb{Q} \cap (\lambda, 1)$, $x\ \mathcal{P}_{(1-\mu)I+\mu J}\ y$.

For $k = 2, 3, 4$, k-diversity is defined for extensions of $\preceq_{\mathbb{J}}$ in exactly the same way. We continue to use the term k-diversity in this setting. The following are conditional-*2*-diversity and partial-*3*-diversity respectively.

A4$^\flat$ For every three distinct elements $x, y, z \in Y$, one of the two subsets $\{J' : x \prec_{J'} y\}$ and $\{J' : y \prec_{J'} x\}$ of $\mathbb{J}$ contains both I and J such that $z \prec_I x$ and $x \prec_J z$. If $\sharp Y = 2$, then *2*-diversity holds on Y.

A4$'^\flat$ For every $Y' \subseteq Y$ with cardinality $n = 2$ or 3, every Y'-extension $\mathcal{R}$ of $\preceq_{\mathbb{J}}$ is such that $\sharp\mathrm{total}(\mathcal{R}) \geqslant n$.

A proper extension $\mathcal{R}$ of $\preceq_{\mathbb{J}}$ is *testworthy* if it satisfies A1$^\flat$–A3$^\flat$ and, for some $J \in \mathbb{J}$ such that $\mathcal{R}_{J\times 0}$ is total, $\mathcal{R}_{\mathfrak{f}} = \mathcal{R}_{J\times 0}^{-1}$. Thus, for each $Y \in 2^X$, $\mathrm{test}(Y, \preceq_{\mathbb{D}}) \simeq \mathrm{test}(Y, \preceq_{\mathbb{J}})$. For any pair of extensions $\mathcal{R}$ and $\hat{\mathcal{R}}$, $\hat{\mathcal{R}}$ is a perturbation of $\mathcal{R}$ if $\mathcal{R}_{\mathfrak{f}} = \hat{\mathcal{R}}_{\mathfrak{f}}$. Moreover, $\hat{\mathcal{R}}$ is a nondogmatic perturbation if $\sharp\mathrm{total}(\hat{\mathcal{R}}) \leqslant \sharp\mathrm{total}(\mathcal{R})$.

4-P$^\flat$ For every $Y \subseteq X$ of cardinality 3 or 4, every testworthy Y-extension of $\preceq_{\mathbb{J}}$ that is novel has a nondogmatic perturbation that satisfies A0$^\flat$–A3$^\flat$.

The following result corresponds to claim 2 of [GS]. Its proof is a consequence of mathematical induction and the combination axiom.

Lemma 1.2. *If $\mathcal{R}_{\mathfrak{J}}$ and $\hat{\mathcal{R}}_{\mathfrak{L}}$ are equivalent and the latter satisfies* A2*, then for every $J \in \mathfrak{J}$ and every rational number $q > 0$, we have $\mathcal{R}_{qJ} = \mathcal{R}_J$.*

The fact that $\preceq_{\mathbb{J}} \equiv \preceq_{\mathbb{D}}$ immediately implies that $\preceq_{\mathbb{J}}$ satisfies A0$^\flat$, A1$^\flat$ and A4$^\flat$ if, and only if, the corresponding axiom holds for $\preceq_{\mathbb{D}}$. In general, we have the following result, which then also yields the equivalence for the prudence axiom.

Lemma 1.3 (proof on page 45)**.** *For $\mathcal{R}_{\mathfrak{J}} \equiv \hat{\mathcal{R}}_{\mathfrak{D}}$, $\mathcal{R}_{\mathfrak{J}}$ satisfies* A2$^\flat$–A3$^\flat$ *if, and only if, $\hat{\mathcal{R}}_{\mathfrak{D}}$ satisfies* A2–A3*.*

The matrix $\mathbf{v} : X \times \mathbb{T} \to R$ is a *representation of* $\preceq_{\mathbb{J}}$ whenever it satisfies

$$\begin{cases} \text{for every } x, y \in X \text{ and every } J \in \mathbb{J}, \\ x \preceq_J y \quad \text{if, and only if,} \quad \sum_{t\in\mathbb{T}} \mathbf{v}(x,t)J(t) \leqslant \sum_{t\in\mathbb{T}} \mathbf{v}(y,t)J(t). \end{cases} \tag{$\flat$}$$

We observe that, via the definition of case types, there exists a representation of $\preceq_{\mathbb{D}}$ that respects case equivalence if, and only if there exists a representation of $\preceq_{\mathbb{J}}$. The above translation and results imply that theorem 1 is equivalent to

Theorem 2. *Let there be given X, $\mathbb{T}^{\mathfrak{f}}$, $\preceq_{\mathbb{J}}$ and associated extensions, as above. Then* (2.i) *and* (2.ii) *are equivalent.*

(2.i) A0$^\flat$–A4$^\flat$ *and* *4*-P$^\flat$ *hold for $\preceq_{\mathbb{J}}$ on X.*

(2.ii) *There exists a matrix $\mathbf{v} : X \times \mathbb{T} \to \mathbb{R}$ that satisfies both:*

- (2.a) $\mathbf{v}$ *is a representation of* $\preceq_{\mathbb{J}}$; *and*
- (2.b) *no row of* $\mathbf{v}$ *is dominated by any other row, and, for every three distinct elements $x, y, z \in X$, $\mathbf{v}(x,\cdot) - \mathbf{v}(z,\cdot)$ and $\mathbf{v}(y,\cdot) - \mathbf{v}(z,\cdot)$ are noncollinear (i.e. linearly independent).*

Moreover, $\mathbf{v}$ *is unique in the sense of theorem 1 part II, with* (2.ii) *replacing* (1.ii) *and* $\mathbb{T}$ *replacing* $\mathbb{C}$.

APPENDIX B. THE PROOF OF THEOREM 2

The present proof follows a similar structure to that of [GS]. That is, we begin with the proof for the case of arbitrary (nonempty) X and finite $\mathbb{T}$. We then show that we can "patch" the proof together to account for the case where $0 < \sharp X < \infty = \sharp \mathbb{T}$. We omit the proof for the case where both X and $\mathbb{T}$ are arbitrary since it is identical to that of [GS].

STEP B.1 (for $\sharp\mathbb{T} < \infty$, we characterise A1$^\flat$–A3$^\flat$, *2*-diversity and novel extensions). For any pair of vectors $\acute{v}, J : \mathfrak{T} \to \mathbb{R}$ and consider the inner product

$$\langle \acute{v}, J \rangle \overset{\text{def}}{=} \sum_{t \in \mathfrak{T}} v(t)J(t).$$

Since $\mathfrak{T}$ is finite, the linear operator $J \mapsto \langle \acute{v}, J \rangle$ is real-valued. Let $\acute{H} \overset{\text{def}}{=} \{J : \langle \acute{v}, J \rangle = 0\}$, let $\acute{G} \overset{\text{def}}{=} \{J : \langle \acute{v}, J \rangle > 0\}$, and let $\acute{F} \overset{\text{def}}{=} \{J : \langle \acute{v}, J \rangle \geqslant 0\}$. Similarly, we take $\acute{H}_{++} = \acute{H} \cap \mathbb{R}^{\mathfrak{T}}_{++}$ to be the (strictly) *positive kernel* of $\langle \acute{v}, \cdot \rangle$ and, for $0 \nleqslant \acute{v} \nleqslant 0$, $\acute{G}_{++}$ and $\acute{F}_{++}$ are, respectively, the open and closed half-spaces of $\mathbb{R}^{\mathfrak{T}}_{++}$ associated with $\acute{v}$. (For such $\acute{v}$, we also refer to $\acute{H}_{++}$ as a hyperplane in $\mathbb{R}^{\mathfrak{T}}_{++}$.) For any $Y \in 2^X$ and matrix $\acute{v} : Y^2 \times \mathfrak{T} \to \mathbb{R}$, each row vector $\acute{v}^{(x,y)} : \mathfrak{T} \to \mathbb{R}$ the associated spaces are $\acute{H}^{\{x,y\}}_{++}$, $\acute{G}^{(x,y)}_{++}$ and $\acute{F}^{(x,y)}_{++}$ respectively.

On occasion we refer to the non-negative counterpart of these sets $\acute{H}_{+}, \acute{G}_{+}$ and $\acute{F}_{+}$ in $\mathbb{R}^{\mathfrak{T}}_{+}$. It is natural to ask why we do not work with the latter sets throughout. The answer is that $\mathbb{R}^{\mathfrak{T}}_{++}$ has the same topological structure as $\mathbb{R}^{\mathfrak{T}}$. By working with hyperplanes in $\mathbb{R}^{\mathfrak{T}}_{++}$, the key result, Zaslavski's theorem, from the literature on arrangements (of hyperplanes) applies without modification. We briefly introduce this literature following the next lemma which corresponds to lemma 1 of [GS] and gives meaning to the statement "the arrangement generated by an extension".

Lemma 2.1 (two-diverse pairwise representation). *Let $\acute{\mathcal{R}}$ be a Y-extension of $\preceq_{\mathbb{J}}$. $\acute{\mathcal{R}}$ satisfies* A1$^\flat$–A3$^\flat$ *and 2*-diversity *holds on Y, if, and only if, there exists a matrix $\acute{v}^{(\cdot,\cdot)} : Y^2 \times \mathfrak{T} \to \mathbb{R}$ such that, for every $x, y \in Y$, row $\acute{v}^{(x,y)} : \mathfrak{T} \to \mathbb{R}$ and its associated spaces $\acute{H}^{\{x,y\}}_{++}$ and $\acute{G}^{(x,y)}_{++}$ satisfy*

(i) $\acute{H}^{\{x,y\}}_{++} \cap \mathbb{Q}^{\mathfrak{T}} = \{J : x\, \acute{\mathcal{I}}_J\, y\}$ *and* $\acute{G}^{(x,y)}_{++} \cap \mathbb{Q}^{\mathfrak{T}} = \{J : x\, \acute{\mathcal{P}}_J\, y\}$,

(ii) $\acute{G}^{(x,y)}_{++}$ *and* $\acute{G}^{(y,x)}_{++}$ *are both nonempty if $x \neq y$ and both empty otherwise,*

(iii) $\acute{H}^{\{y,x\}}_{++} = \acute{H}^{\{x,y\}}_{++}$ *(and in particular $\acute{v}^{(y,x)} = -\acute{v}^{(x,y)}$),*

(iv) $\acute{H}^{\{x,y\}}_{++}$ *is the unique hyperplane in $\mathbb{R}^{\mathfrak{T}}_{++}$ that separates $\{J : x\, \acute{\mathcal{P}}_J\, y\}$ and $\{J : y\, \acute{\mathcal{P}}_J\, x\}$ ($\acute{v}^{(x,y)}$ is unique upto multiplication by a positive scalar).*

Moreover, $\acute{\mathcal{R}}$ is novel if, and only if, for every $t \neq \mathfrak{f}$, $\acute{v}^{(\cdot,\cdot)}(t) \neq \acute{v}^{(\cdot,\cdot)}(\mathfrak{f})$.

See proof on page 46. We refer to a matrix $v^{(\cdot,\cdot)}$ that satisfies condition (ii) of lemma 2.1 (for every distinct $x, y \in Y$) as a 2-*diverse matrix.* If $v^{(\cdot,\cdot)}$ satisfies all the conditions of lemma 2.1 with respect to a given extension $\mathcal{R}$ and, in addition, $x = y$ implies $v^{(x,y)} = 0$, then $v^{(\cdot,\cdot)}$ is a 2-diverse pairwise representation of $\mathcal{R}$.[13]

[13] Note that for $\sharp X = 2$, lemma 2.1 constitutes a proof of theorem 2. For we may take $\mathbf{v} : X \times \mathbb{T} \to \mathbb{R}$ such that $\mathbf{v}(x, \cdot) = 0$ and set $\mathbf{v}(y, \cdot) = v^{(x,y)}$, so that $v^{(x,y)} = -\mathbf{v}(x, \cdot) + \mathbf{v}(y, \cdot)$.

AN *arrangement* is a collection of hyperplanes in $\mathbb{R}^{\mathfrak{T}}_{++}$ or $\mathbb{R}^{\mathfrak{T}}$. (A *hyperplane in* $\mathbb{R}^{\mathfrak{T}}_{++}$ is the positive kernel of some nonzero vector.) The chief result, from the mathematics of arrangements, to which we extensively appeal is Zaslavsky's theorem two form of which we state below. For any given extension $\mathcal{R}$, Zaslavsky's theorem allows us to use information about the intersections of hyperplanes in the arrangement to identify $\sharp\text{total}(\mathcal{R})$. It does so by counting the collection $\mathcal{G}_{++}$ of open and connected subsets of $\mathbb{R}^{\mathfrak{T}} - \bigcup\{H_{++} : H_{++} \in \mathcal{A}\}$ are the *chambers* or *regions* of the arrangement. In the present setting, each chamber corresponds to a *CAR ranking* of the elements of Y: a complete, antisymmetric and reflexive (but possibly intransitive) ranking R. Every CAR ranking can be succinctly represented as a CAR list. For instance, take $Y = \{x, y, z\}$ and $xRyRz$, then the corresponding list

$$l = \begin{cases} (x, y, z) & \text{if } R \text{ is transitive} \\ (x, y, z, x) & \text{if } R \text{ is intransitive.}^{14} \end{cases}$$

The notation extends without exception to sets of cardinality 4. E.g., (x, y, z, x, w) represents the CAR ranking that is intransitive over $\{x, y, z\}$ and such that w dominates every other member.

THE INTERSECTION SEMILATTICE of any arrangement $\mathcal{A}$ is the partially ordered (*by reverse inclusion*) set $\mathcal{L}$ of intersections of members of $\mathcal{A}$. The unique minimal element is obtained by taking the intersection $A^{\varnothing}$ over the empty subarrangement $\mathcal{A}^{\varnothing}$ of $\mathcal{A}$ to obtain the ambient space itself. That is $A^{\varnothing} = \mathbb{R}^{\mathfrak{T}}$ or $\mathbb{R}^{\mathfrak{T}}_{++}$, depending on whether we are considering the lattice $\mathcal{L}$ or the lattice $\mathcal{L}_{++}$ respectively. In [GS], as a consequence of *4*-diversity, $\mathcal{H}_{++}$ is always central. In our setting, it is only $\mathcal{H}$ that is guaranteed to be central. In general an arrangement is central, if, and only if, its intersection semilattice has a unique maximal element (**?**, proposition 2.3). Thus, if $\mathcal{H}_{++}$ is *centerless*, then $\mathcal{L}_{++}$ is a meet semilattice with multiple maxima: as in example 6. Extending our notation: if $Y = \{x, y, z, w\}$, then the unique intersection A^{Y} is the (nonempty) *center* of $\mathcal{A}^{Y} = \mathcal{H}$. By $A^{\{x,y,z\}}$, we mean the intersection over $\mathcal{A}^{\{x,y,z\}} \overset{\text{def}}{=} \{H^{\{i,j\}} : i \neq j \text{ in } \{x, y, z\}\}$. Finally, by $A^{\{x,y\}\{z,w\}}$, we mean the intersection over $\mathcal{A}^{\{x,y\}\{z,w\}} \overset{\text{def}}{=} \{H^{\{x,y\}}, H^{\{z,w\}}\}$.

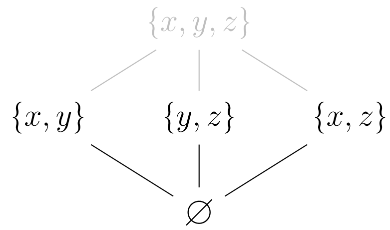

Figure 1: The intersection semilattice $\mathcal{L}_{++} = \mathcal{L} - A^{\{x,y,z\}}$.

Zaslavski's theorem provides two distinct methods for counting the number of regions in an arrangement. The first states that $\sharp\mathcal{G}$ *is equal to the sum of the absolute values of the Möbius function* $\boldsymbol{\mu} : \mathcal{L} \to \mathbb{Z}$ which is defined recursively via

$$\boldsymbol{\mu}(A) = \begin{cases} 1 & \text{if } A = A^{\varnothing} \\ -\sum\{\boldsymbol{\mu}(B) : A \subsetneq B\} & \text{otherwise.}^{15} \end{cases} \tag{5}$$

The above definition of Zaslavski's theorem is explicitly provided by **?**. Specialised to the present setting, the more common (see **???**) "rank" version of Zaslavski's theorem is

$$\sharp\mathcal{G} = \sum_{\substack{\mathcal{A}\subseteq\mathcal{H} \\ \mathcal{A} \text{ central}}} (-1)^{|\mathcal{A}|-\operatorname{rank}(\mathcal{A})},$$

where *central* means that $\bigcap\{H : H \in \mathcal{A}\}$ is nonempty, and $\operatorname{rank}(\mathcal{A})$ is the dimension of the space spanned by the normals to the hyperplanes in $\mathcal{A}$.[16]

Example 6 (a comparison of $\mathcal{L}$ and $\mathcal{L}_{++}$). *Let* $X = \{x,y,z\}$, $\mathbb{T} = \{s,t\}$, *and* $u^{(x,y)} = 1 \times -1$ *and* $u^{(y,z)} = 2 \times -1$ *denote vectors in* $\mathbb{R}^{\mathbb{T}}$. *We now appeal to the Jacobi identity and take* $u^{(x,z)} = u^{(x,y)} + u^{(y,z)} = 3 \times -2$ *and extend to the remaining pairs in* X^2 *using lemma 2.1. Since these vectors are pairwise noncollinear and* $\sharp\mathbb{T} = 2$, *the associated arrangement* $\mathcal{H}_{++} = \left\{H^{\{x,y\}}_{++}, H^{\{y,z\}}_{++}, H^{\{x,z\}}_{++}\right\}$ *consists of three pairwise disjoint lines that partition* $\mathbb{R}^{\mathbb{T}}_{++}$ *and* $\mathcal{G}_{++}$ *has cardinality* 4. *We now confirm this using Zaslavski's theorem.*

In the present setting $A^{\varnothing}_{++} = \mathbb{R}^{\mathbb{T}}_{++}$ *and, via eq. (5),* $\boldsymbol{\mu}(A^{\varnothing}_{++}) = 1$. *Then, since* $\mathbb{R}^{\mathbb{T}}_{++}$ *is the unique element in* $\mathcal{L}_{++}$ *that (strictly) contains each hyperplane in* $\mathcal{H}_{++}$,

[16] Equivalently, $\operatorname{rank}(\mathcal{A})$ is the dimension of the orthogonal complement in $\mathbb{R}^{\mathfrak{T}}$ (or $\mathbb{R}^{\mathfrak{T}}_{++}$) of the intersection over $\mathcal{A}$. Since the intersection over the empty arrangement $\mathcal{A}^{\varnothing}$ is the ambient space $\mathbb{R}^{\mathfrak{T}}$ (or $\mathbb{R}^{\mathfrak{T}}_{++}$), the only vector in $\mathbb{R}^{\mathfrak{T}}$ that is orthogonal to the ambient space is 0, $\operatorname{rank}(\mathcal{A}^{\varnothing}) = 0$.

eq. (5) *yields* $\boldsymbol{\mu}(A) = -\boldsymbol{\mu}(A^{\varnothing}_{++})$ *for each* $A \in \mathcal{H}_{++}$. *Now since the hyperplanes in* $\mathcal{H}_{++}$ *are disjoint, there are no further elements in* $\mathcal{L}_{++}$ *and we observe that*

$$|\mathcal{G}_{++}| = \sum_{A \in \mathcal{L}_{++}} |\boldsymbol{\mu}(A)| = 4.$$

In contrast, although the structure of $\mathcal{L}$ *is otherwise isomorphic to* $\mathcal{L}_{++}$, *since* $\{0\} \subset \mathbb{R}^{\mathbb{T}}$ *is a subset of every hyperplane in* $\mathcal{H}$, $\{0\}$ *is the center* $A^{\{x,y,z\}}$ *of* $\mathcal{H}$ *and the maximal element of* $\mathcal{L}$. *Via eq.* (5) *and the calculations of the previous paragraph, we obtain* $\boldsymbol{\mu}(A^{\{x,y,z\}}) = -\left(\boldsymbol{\mu}(A^{\varnothing}) - 3\boldsymbol{\mu}(A^{\varnothing})\right) = 2$. *Thus,*

$$\sharp\mathcal{G} = \sum_{A \in \mathcal{L}} |\boldsymbol{\mu}(A)| = 6 = 3!.$$

Remark 2 (The relationship between $\mathcal{L}$ and $\mathcal{L}_{++}$). *Let* $\acute{\mathcal{R}}$ *be a* Y*-extension with* 2*-diverse representation* $\acute{u}^{(\cdot,\cdot)}$. *Since, for every distinct* x *and* y *in* Y, $\acute{H}^{\{x,y\}}$ *contains the origin,* $\acute{\mathcal{H}}$ *is centered. As we see in example 6, this is not the case for* $\acute{\mathcal{H}}_{++}$ *where* $\acute{\mathcal{H}}_{++}$ *is centerless and each of its members is maximal in* $\acute{\mathcal{L}}_{++}$.

In [GS], *4*-diversity *guarantees that, for every* $Y \subseteq X$ *of cardinality* $2, 3$ *or* 4, *the improper* Y*-extension generates a centered arrangement in* $\mathbb{R}^{\mathbb{T}}_{++}$. *The fact that* $\mathbb{R}^{\mathfrak{T}}_{++}$ *is open in* $\mathbb{R}^{\mathfrak{T}}$ *ensures that the dimension of any* $L \in \acute{\mathcal{L}}$ *is equal to its counterpart* $L_{++} \in \acute{\mathcal{L}}_{++}$ *provided the latter exists. Thus,* $\acute{\mathcal{L}}_{++}$ *and* $\acute{\mathcal{L}}$ *are isomorphic if, and only if,* $\acute{\mathcal{H}}_{++}$ *is centered. For the same reason,* $\acute{\mathcal{G}}_{++}$ *and* $\acute{\mathcal{G}}$ *are isomorphic if, and only if,* $\acute{\mathcal{H}}_{++}$ *is centered.*

We now abstract a useful property from example 6.

Proposition 2 (polar opposite rankings). *If* $\preceq_{\mathbb{J}}$ *satisfies* A1$^\flat$–A3$^\flat$ *and 2*-diversity, *then, for every* $Y \subseteq X$ *of cardinality* 3 *or* 4, *the improper* Y*-extension* $\mathcal{R}$ *is such that, for some* $J, L \in \mathbb{J}$, $\mathcal{R}_J = \mathcal{R}_L^{-1}$ *belongs to* total$(\mathcal{R})$.

See proof on page 47.

Example 7 (insufficiency of *2*-diversity). *Let* $X = [0,1]^2$ *and let* $\leqslant^{\text{lex}}$ *denote the lexicographic ordering on* X. *Let* $\mathbb{T} = \{s, t\}$, *and, for each* $J \in \mathbb{J}$, *let*

$$\preceq_J = \begin{cases} X^2 & \textit{if } J(s) = J(t); \\ \leqslant^{\text{lex}} & \textit{if } J(s) < J(t); \\ (\leqslant^{\text{lex}})^{-1} & \textit{otherwise.} \end{cases}$$

Recall that if $\leq_J = X^2$*, then* $\leq_J$ *is symmetric and hence equal to* $\simeq_J$*. Thus, for every distinct* $x, y \in X$*,* $H^{\{x,y\}}_{++} = \{J \in \mathbb{R}^{\mathbb{T}}_{++} : J(s) = J(t)\}$*. Via lemma 2.1,* $\leq_{\mathbb{J}}$ *has a two-diverse matrix representation* $v^{(\cdot,\cdot)}$*. But via lemma 2.2, below,* conditional-*2*-diversity *fails to hold. The fact that* $\leq_{\mathbb{J}}$ *fails to satisfy part* (2.a) *of theorem 2 follows from the fact that* $\leq_J$ *is lexicographic for every* J *outside* H*.*

Step B.2 (for $\mathbb{T} < \infty$, characterisations of A4$^\flat$). A matrix $v^{(\cdot,\cdot)}$ that satisfies the conditions of the next lemma is a *conditionally-*2*-diverse (pairwise) representation.*

Lemma 2.2 (conditionally-2-diverse representation). *Let* $\mathcal{R}$ *be a* Y*-extension of* $\leq_{\mathbb{J}}$ *with* 2*-diverse matrix representation* $v^{(\cdot,\cdot)}$*. Then* A4$^\flat$ *holds on* Y *if, and only if, for every three distinct elements* $x, y, z \in Y$*,* $v^{(x,z)}$ *and* $v^{(y,z)}$ *are noncollinear.*

See proof on page 48. The following is a translation of observation 2.

Proposition 3 (on A4$^\flat$ and A4$'^\flat$). *For* $\leq_{\mathbb{J}}$ *satisfying* A0$^\flat$–A3$^\flat$*,* conditional-*2*-diversity *and* partial-*3*-diversity *are equivalent.*

See proof on page 48.

Step B.3 (for $\mathbb{T} < \infty$, a characterisation of *4*-prudence). The following Jacobi identity plays a central role in the proof of [GS].

Definition. *For* $Y \in 2^X$*, the matrix* $v^{(\cdot,\cdot)} : Y^2 \times \mathfrak{T} \to \mathbb{R}$ *satisfies the* Jacobi identity *whenever, for every* $x, y, z \in Y$*,* $v^{(x,z)} = v^{(x,y)} + v^{(y,z)}$*.*

For any given Y-extension $\mathcal{R}$, the *Jacobi identity holds for* $\mathcal{R}$ whenever it holds for some pairwise representation $v^{(\cdot,\cdot)}$ of $\mathcal{R}$. Moreover, in this case, $v^{(\cdot,\cdot)}$ is a *Jacobi representation.* Finally, if the Y-extension $\mathcal{R}$ is improper and the Jacobi identity holds for $\mathcal{R}$, we simply say that the *Jacobi identity holds on* Y. Consider

k-Jac. *For every* $Y \subseteq X$ *with* $3 \leqslant \sharp Y \leqslant k$*, the Jacobi identity holds on* Y*.*

We will work with *3*-Jac and *4*-Jac in particular. The following lemma is the special case of theorem 3 of section C where $\mathbb{T}$ is finite. When $\mathbb{T}$ is finite, for every Y, the set of testworthy Y-extensions that are novel is nonempty. In this case, *4*-prudence implies A0$^\flat$–A3$^\flat$ via nonrevision of rankings (part 2 of definition 3).

Lemma 2.3. *For* $\leq_{\mathbb{J}}$ *satisfying* A4$^\flat$*,* *4*-prudence *holds if, and only if,* *4*-Jac *holds.*

STEP B.4 (for arbitrary X and $\mathbb{T}$, the induction argument). The present step is closely related to lemma 3 and claim 9 of [GS]. There the authors establish that, when *4*-diversity holds, 3-Jac is a necessary and sufficient condition for the (global) Jacobi identity to hold on X. [GS] relies on the fact that *4*-diversity implies linear independence of $\{v^{(x,y)}, v^{(y,z)}, v^{(z,w)}\}$ for every four distinct elements $x, y, z, w \in X$. In the present setting, where conditional-*2*-diversity only implies linear independence of pairs $\{v^{(x,y)}, v^{(y,z)}\}$, the Jacobi identity requires *4*-Jac.

LEMMA 2.4 (Jacobi representation). *Let $\preceq_{\mathbb{J}}$ have a conditionally-*2*-diverse representation $u^{(\cdot,\cdot)}$. Then *4*-Jac holds if, and only if, $\preceq_{\mathbb{J}}$ has a Jacobi representation $v^{(\cdot,\cdot)}$. Moreover, for every Jacobi representation $\mathbf{v}^{(\cdot,\cdot)}$ of $\preceq_{\mathbb{J}}$ there exists $\lambda > 0$ satisfying $\mathbf{v}^{(\cdot,\cdot)} = \lambda v^{(\cdot,\cdot)}$.*

See proof on page 49.

STEP B.5 (for $\sharp\mathbb{T} < \infty$, the concluding arguments in the proof of theorem 2). In step B.2 we showed that A1$^\flat$–A4$^\flat$ hold if, and only if, $\preceq_{\mathbb{J}}$ has a conditionally 2-diverse pairwise representation. In step B.3, we showed that, when $\preceq_{\mathbb{J}}$ satisfies A0$^\flat$–A4$^\flat$, a necessary and sufficient condition for *4*-Jac is *4*-P$^\flat$. In step B.4, via (mathematical) induction, we showed that conditionally 2-diverse Jacobi representations of Y-extensions of $\preceq_{\mathbb{J}}$ such that $\sharp Y = 4$, can be patched together to obtain a conditionally 2-diverse Jacobi representation of $\preceq_{\mathbb{J}}$ (on all of X, regardless of cardinality). The fact that A0$^\flat$ is necessary for a Jacobi representation follows from [GS]. As a consequence, A0$^\flat$–A4$^\flat$ and *4*-P$^\flat$ are necessary and sufficient for a conditionally 2-diverse Jacobi representation of $\preceq_{\mathbb{J}}$. The following argument then completes the proof of theorem 2.

Let $v^{(\cdot,\cdot)}$ be a (conditionally 2-diverse) Jacobi representation of $\preceq_{\mathbb{J}}$ and define $\mathbf{v} : X \times \mathbb{T} \to \mathbb{R}$ as follows. Fix arbitrary $w \in X$, and let $\mathbf{v}(w, \cdot) = 0$. Then, for every $x \in X$, let $\mathbf{v}(x, \cdot) = v^{(w,x)}$. Recalling that $v^{(w,x)} = -v^{(x,w)}$, note that, since the rows of $v^{(\cdot,\cdot)}$ satisfy the Jacobi identity, for every $x, y \in X$, we have

$$v^{(x,y)} = v^{(x,w)} + v^{(w,y)} = -\mathbf{v}(x, \cdot) + \mathbf{v}(y, \cdot).$$

To see that (2.a) holds note that, for every $J \in \mathbb{J}$, we have $x \preceq_J y$, if, and only if, $0 \leqslant \langle v^{(x,y)}, J \rangle$, if, and only if, $\langle v(x, \cdot), J \rangle \leqslant \langle \mathbf{v}(y, \cdot), J \rangle$. For (2.b), note that, since

$v^{(\cdot,\cdot)}$ is a conditionally 2-diverse pairwise representation, for every $x, y \in X$,

$$0 \nleqslant v^{(x,y)} = -\mathbf{v}(x,\cdot) + \mathbf{v}(y,\cdot).$$

Finally, for every $z \in X$, we have, for every $\lambda \in \mathbb{R}$, $v^{(z,x)} \neq \lambda v^{(z,y)}$. Equivalently,

$$v(x,\cdot) \neq (1-\lambda)\mathbf{v}(z,\cdot) + \lambda \mathbf{v}(y,\cdot).$$

Theorem 1 part II, on uniqueness, follows from lemma 2.4 and, without modification, part 3 of the proof of theorem 2 of [GS] (see page 23).

Step B.6 (the case of arbitrary X and $\mathbb{T}$). Note that the inner product we have been working with throughout is well-defined on the infinite-dimensional, linear subspace $\mathbb{R}^{\oplus\mathfrak{T}} \subset \mathbb{R}^{\mathfrak{T}}$ of vectors with finite support. Indeed, for every $v \in \mathbb{R}^{\mathfrak{T}}$ and $J \in \mathbb{R}_{+}^{\oplus\mathfrak{T}}$, the inner product is a finite sum

$$\langle v, J\rangle = \sum_{\{t:J(t)>0\}} v(t)J(t).$$

Next note that $\mathfrak{J}$ is just the set of rational-valued vectors in $\mathbb{R}^{\oplus\mathfrak{T}}$. As such, the notions of orthogonality and hyperplanes carry over to the present, infinite-dimensional, setting. W.l.o.g., let $\acute{\mathcal{R}}$ be an improper Y-extension of $\leq_{\mathbb{J}}$ and note that, for every $Y \subseteq X$ of cardinality 2, 3 or 4, the representation $\acute{u}^{(\cdot,\cdot)}$ of $\acute{\mathcal{R}}$ has finite rank $\acute{\mathbf{r}}$. The *essentialization* $\mathrm{ess}(\acute{\mathcal{H}})$ of the associated arrangement $\acute{\mathcal{H}}$ in $\mathbb{R}^{\oplus\mathbb{T}}$ is the arrangement we obtain via the by orthogonally projecting $\acute{\mathcal{H}}$ onto the ($\acute{\mathbf{r}}$-dimensional) span of $\acute{u}^{(\cdot,\cdot)}$. Let $\mathrm{p} : \mathbb{R}^{\oplus\mathbb{T}} \to S = \mathrm{span}\{\acute{u}^{(\cdot,\cdot)}\}$ denote this projection. Then, for any $J \in \mathfrak{J}$, J is the sum of $\mathrm{p}(J)$ and a term that belongs to the kernel of $u^{(x,y)}$, for every $x, y \in Y$. Thus, for every $J \in \mathbb{R}^{\oplus\mathbb{T}}$, $\langle u^{(x,y)}, J\rangle_{\mathbb{R}^{\oplus\mathbb{T}}} = \langle u^{(x,y)}, \mathrm{p}(J)\rangle_S$. As such, all the structure of $\acute{\mathcal{H}}$ in $\mathbb{R}^{\oplus\mathbb{T}}$ is preserved by $\mathrm{ess}(\acute{\mathcal{H}})$ in the finite-dimensional subspace S.

For $n = \acute{\mathbf{r}}$, take $\mathcal{T} = \{T_i : i = 1, \ldots, n\}$ to be an orthonormal basis for S. Then let $T : R^n \to S$ to be the matrix with columns $\mathcal{T}$. Then, for every $J \in \mathbb{R}^n$, $T(J) = J_1T_1 + \cdots + J_nT_n$, where $J_i \in \mathbb{R}$ for each i. Via orthonormality of $\mathcal{T}$,

$$\langle T_i, T_j\rangle_S = \begin{cases} 1 & \text{if } i = j, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

It then follows that, for every pair of vectors $v, J \in \mathbb{R}^n$, we have

$$\langle T(v), T(J) \rangle_S = \sum_{i,j=1}^{n} v_i J_j \langle T_i, T_j \rangle_S = \langle v, J \rangle_{\mathbb{R}^n}.$$

Now, for each row $u^{(x,y)}$ of $u^{(\cdot,\cdot)}$, let $\acute{\mathbf{u}}^{(x,y)} \stackrel{\text{def}}{=} T^{-1}(\acute{u}^{(x,y)})$, and let $\acute{\mathbf{u}}^{(\cdot,\cdot)}$ denote the matrix with rows $\{\acute{\mathbf{u}}^{(x,y)} : x, y \in Y\}$. Then let $\acute{\mathcal{R}} = \langle \acute{\mathcal{R}}_J : J \in \mathbb{Q}^n_+ \rangle$ denote the extension generated by $\acute{\mathbf{u}}^{(\cdot,\cdot)}$ in $\mathbb{Q}^n_+$, for every $x, y \in Y$ and every $J \in \mathbb{Q}^n_+$, satisfy $x\, \acute{\mathcal{R}}_J\, y$ if, and only if, $\langle \acute{\mathbf{u}}^{(x,y)}, J \rangle_{\mathbb{R}^n} \geqslant 0$. Then, by construction, $\langle \acute{\mathbf{u}}^{(x,y)}, J \rangle_{\mathbb{R}^n} \geqslant 0$ if, and only if $\langle \acute{u}^{(x,y)}, T(J) \rangle_S \geqslant 0$. Thus, $x\, \acute{\mathcal{R}}_J\, y$ if, and only if, $x\, \acute{\mathcal{R}}_{T(J)}\, y$.

Finally, note that, with the exception of lemma 2.4 of step B.4, all results in the proof of theorem 2 involve Y-extensions $\mathcal{R}$ of $\preceq_{\mathbb{J}}$ for $\sharp Y = 2, 3$ or 4. By taking $\mathbb{T}_Y = \{1, \ldots, n\}$ and substituting for $\mathbb{T}$ and then working with $\acute{\mathcal{R}}$ and $\acute{\mathbf{u}}^{(\cdot,\cdot)}$ in $\mathbb{R}^n$, all these results extend to arbitrary X and $\mathbb{T}$. And, since step B.4 holds for arbitrary $\mathbb{T}$ and X, the proof of theorem 2 is indeed complete.

APPENDIX C. STATEMENT AND PROOF OF THEOREM 3

Theorem 3. *For $\preceq_{\mathbb{J}}$ satisfying* A0$^\flat$–A4$^\flat$*,* 4-prudence *holds if, and only if, the Jacobi identity holds (for some pairwise representation of $\preceq_{\mathbb{J}}$).*

Via lemma 2.4, it suffices to show that 4-prudence is equivalent to 4-Jac. Then by construction, 4-prudence and 4-Jac are conditions that apply independently to each subset Y of X that has cardinality 3 or 4. Throughout the present section, $Y \subseteq X$ has cardinality 3 or 4 and $\mathcal{R}$ denotes the improper Y-extension of $\preceq_{\mathbb{J}}$. This will allow us to apply the translation of step B.6 and work in a finite dimensional space. But first we need to rule out the following case.

SUPPOSE THERE IS NO TESTWORTHY Y-EXTENSION THAT IS NOVEL, In this case, 4-prudence holds vacuously on Y. The follow lemma confirms that, in this case, 4-diversity holds and theorem 2 of [GS] applies, so that 4-Jac holds on Y.

Lemma 3.1. *If every testworthy Y-extension of $\preceq_{\mathbb{J}}$ is regular, then $|\mathbb{T}| = \infty$ and* 4-diversity *holds on Y.*

Proof of lemma 3.1. Let $|Y| = 4$. (The proof for case where $|Y| = 3$ is similar and omitted.) Via lemma 3.2, the set of testworthy Y-extensions is nonempty. Let $\hat{\mathcal{R}}$

be a testworthy Y-extension, so that there exists J in $\mathbb{J}$ such that $\hat{\mathcal{R}}_{J\times 0}$ is total and equal to the inverse of $\hat{\mathcal{R}}_{\mathfrak{f}}$. Let $P \stackrel{\text{def}}{=} \hat{\mathcal{P}}_{\mathfrak{f}}$. Thus P is the asymmetric part of a total ordering of Y. Via lemma 2.1, let $\hat{v}^{(\cdot,\cdot)}$ be the matrix representation of $\hat{\mathcal{R}}$ and let $\hat{v}^P$ denote the restriction of $\hat{v}^{(\cdot,\cdot)}$ to $P \times \mathbb{T}^{\mathfrak{f}}$.

Claim 3.1.1. *For every vector $\eta^P = \langle \eta^{(x,y)} \in \mathbb{R}_{++} : (x,y) \in P\rangle$, there exist $s,t \in \mathbb{T}$ such that $\hat{v}^P(s) = \eta^P$ and $\hat{v}^P(t) = -\eta^P$.*

Proof of claim 3.1.1. By way of contradiction, suppose there exists $\acute{\eta}^P \in \mathbb{R}^P_{++}$ such that, for every s in $\mathbb{T}$, $\hat{v}^P(s) \neq \acute{\eta}^P$. That is η^P such that, for every s in $\mathbb{T}$, there exists (x,y) in P such that $\hat{v}^{(x,y)}(s) \neq \acute{\eta}^{(x,y)}$. This property will suffice for the existence of a testworthy Y-extension $\acute{\mathcal{R}}$ that is novel. Define $\acute{v}^{(\cdot,\cdot)} : X^2 \times \mathbb{T}^{\mathfrak{f}} \to \mathbb{R}$ as follows. For each (x,y) in P, let

$$\acute{v}^{(x,y)}(s) \stackrel{\text{def}}{=} \begin{cases} \acute{\eta}^{(x,y)} & \text{if } s = \mathfrak{f}, \\ \hat{v}^{(x,y)}(s) & \text{otherwise.} \end{cases}$$

For every (x,y) in P^{-1}, since (y,x) in P, take $\acute{v}^{(x,y)} = -\acute{v}^{(y,x)}$. Finally, since P is the asymmetric part of a total ordering, for every remaining (x,y) in Y^2, $x = y$, so let $\acute{v}^{(x,y)} = 0$. Observe that, by construction, for every s in $\mathbb{T}$, $\acute{v}^{(\cdot,\cdot)}(s) \neq \acute{v}^{(\cdot,\cdot)}(\mathfrak{f})$. This allows us to appeal to lemma 2.1 and take $\acute{\mathcal{R}}$ to be the novel extension that $\acute{v}^{(\cdot,\cdot)}$ generates. Moreover, $\acute{\mathcal{R}}$ is also testworthy. For together $P = \hat{\mathcal{P}}_{\mathfrak{f}}$ and $\hat{\mathcal{P}}_{\mathfrak{f}} = \hat{\mathcal{P}}^{-1}_{J\times 0}$ imply that $\acute{\mathcal{R}}_{\mathfrak{f}} = \acute{\mathcal{R}}^{-1}_{J\times 0}$ since, for every (x,y) in P, we have

$$\langle \acute{v}^{(x,y)}, J \times 0\rangle = \langle \hat{v}^{(x,y)}, J \times 0\rangle < 0 < \acute{\eta}^{(x,y)} = \acute{v}^{(x,y)}(\mathfrak{f}).$$

This contradiction implies that, for every $\eta^P \in \mathbb{R}^P_{++}$, there exists s in $\mathbb{T}$ such that $\hat{v}^P(s) = \eta^P$. *Mutatis mutandis*, a repetition of the preceding argument by contradiction confirms that there exists t in $\mathbb{T}$ such that $\hat{v}^P(t) = -\eta^P$. □

Claim 3.1.1 implies that, when every testworthy Y-extension is regular, the cardinality of $\mathbb{T}$ is equal to the cardinality of $\mathbb{R}^P$. We now show that 4-diversity holds on Y. Let R denote an arbitrary total ordering of Y. We show that, for some K in $\mathbb{J}$, $\langle \hat{v}^{(x,y)}, K\rangle \geqslant 0$ if, and only if, (x,y) in R. Claim 3.1.1 ensures that we can choose s in $\mathbb{T}$ such that, for some $0 < \epsilon < 1$

$$\hat{v}^{(x,y)}(s) = \begin{cases} 1 + \epsilon & \text{if } (x,y) \text{ in } R \cap P, \\ 1 - \epsilon & \text{if } (x,y) \text{ in } R^{-1} \cap P. \end{cases}$$

Via claim 3.1.1, take t in $\mathbb{T}$ such that, for every (x, y) in P, $\hat{v}^{(x,y)}(t) = -1$. Let $K := \delta_s + \delta_t$ in $\mathbb{J}^{\mathfrak{f}}$, so that $\langle \hat{v}^{(x,y)}, K \rangle = \hat{v}^{(x,y)}(s) + \hat{v}^{(x,y)}(t)$. By evaluating terms and observing that $\epsilon > 0$ we obtain

$$\langle \hat{v}^{(x,y)}, K \rangle = \begin{cases} (1+\epsilon) - 1 > 0 & \text{if } (x,y) \text{ in } R \cap P, \\ (1-\epsilon) - 1 < 0 & \text{if } (x,y) \text{ in } R^{-1} \cap P. \end{cases}$$

Since (x, y) in $R^{-1} \cap P^{-1}$ if, and only if, (y, x) in $R \cap P$ (and, similarly, (x, y) in $R \cap P^{-1}$ if, and only if (y, x) in $R^{-1} \cap P$), we appeal to $\hat{v}^{(x,y)} = -\hat{v}^{(y,x)}$ and obtain

$$\langle \hat{v}^{(x,y)}, K \rangle = \begin{cases} -(1+\epsilon) + 1 < 0 & \text{if } (x,y) \text{ in } R^{-1} \cap P^{-1}, \\ -(1-\epsilon) + 1 > 0 & \text{if } (x,y) \text{ in } R \cap P^{-1}. \end{cases}$$

Since P is the asymmetric part of a total ordering we conclude that, for every $x \neq y$, $\langle \hat{v}^{(x,y)}, K \rangle$ has the right sign. Finally, for $x = y$, $\langle \hat{v}^{(x,y)}, K \rangle = 0$. □

As a consequence of lemma 3.1, for the remainder of the proof of theorem 3 we work under the assumption that the set of testworthy Y-extensions that are novel is nonempty.

For the remainder of this section, redefine $\mathbb{T} := \mathbb{T}_Y$, where recall $\mathbb{T}_Y$ is defined in step B.6. Thus, in the present section, $\mathbb{T}$ is the finite set that indexes the translation to $\mathbb{R}^n$ of step B.6. Then $\mathbb{T}^{\mathfrak{f}} = \mathbb{T} \cup [\mathfrak{f}]$ as before and indeed all other notation remains the same. Whereas $\mathbb{T} = \mathbb{T}_Y$ is now finite, the set $\mathbb{C}_{/\sim^{\mathcal{R}}}$ of equivalence classes of $\sim^{\mathcal{R}}$ in $\mathbb{C}$ may be infinite as we have seen in lemma 3.1.

Since A1$^\flat$–A4$^\flat$ hold, the pairwise representation $u^{(\cdot,\cdot)} : Y^2 \times \mathbb{T} \to \mathbb{R}$ of $\mathcal{R}$ is conditionally 2-diverse. Lemma 2.2 implies that $u^{(\cdot,\cdot)}$ has row rank $\mathbf{r} \geqslant 2$.

Recall the definition of central arrangements. We now show that the set of testworthy extensions with a central arrangement is always nonempty.

Lemma 3.2. *For every $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total, there exists a Y-extension $\acute{\mathcal{R}}$ with $\acute{\mathcal{R}}_{\mathfrak{f}} = \mathcal{R}_J^{-1}$. Moreover, its arrangement $\acute{\mathcal{H}}_{++}$ is central; its representation $\acute{u}^{(\cdot,\cdot)}$ has rank $\acute{\mathbf{r}} = \mathbf{r}$; and 4-*Jac *holds for $\acute{\mathcal{R}}$ if, and only if, it holds for $\mathcal{R}$.*

Proof. Let $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total. If, for some $t \in \mathbb{T}$, $J(t) = 0$, then, via $\binom{\sharp Y}{2}$ applications of A3$^\flat$ (or the continuity properties of the inner product), there exists there $L \in \mathbb{R}^{\mathbb{T}}_{++}$ such that $\mathcal{R}_L = \mathcal{R}_J$. Thus, w.l.o.g., take $J(t) > 0$ for every

$t \in \mathbb{T}$. For some rational $0 < \acute{\iota} < 1$, let $\acute{J} = (1 - \acute{\iota})J \times \acute{\iota}$, so that $\acute{J}$ lies on the (relative) interior of $\mathbb{J}^{\mathfrak{f}}$. Now, for every x, y in Y, let

$$\acute{\eta}^{(x,y)} := -\frac{1-\acute{\iota}}{\acute{\iota}}\langle u^{(x,y)}, J\rangle,$$

and let $\acute{u}^{(x,y)} \stackrel{\text{def}}{=} u^{(x,y)} \times \acute{\eta}^{(x,y)}$, so that $\langle \acute{u}^{(x,y)}, \acute{J}\rangle = 0$. Let $\acute{\mathcal{R}}$ be the associated Y-extension, so that by construction $\acute{\mathcal{R}}_{\mathfrak{f}} = \acute{\mathcal{R}}_J^{-1}$. Since $\acute{J} \in \acute{H}^{\{x,y\}}$ for every distinct $x, y \in Y$, $\acute{\mathcal{H}}$ is central. By construction, note that $\mathbf{r} \leqslant \acute{\mathbf{r}} \leqslant \mathbf{r} + 1$. Moreover, recalling the rank–nullity theorem, we have $\acute{\mathbf{r}} = \sharp\mathbb{T}^{\mathfrak{f}} - \dim(\ker(\acute{u}^{(\cdot,\cdot)}))$. Now, $\acute{J} \in \mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}$ belongs to $\ker(\acute{u}^{(\cdot,\cdot)})$, but not $\ker(u^{(\cdot,\cdot)}) \times 0 \subset \mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}$. Moreover, the latter set belongs to $\ker(\acute{u}^{(\cdot,\cdot)})$. Thus, $\ker(\acute{u}^{(\cdot,\cdot)})$ is of dimension one more than $\ker(u^{(\cdot,\cdot)}) \subset \mathbb{R}^{\mathbb{T}}$. Then since $\sharp\mathbb{T}^{\mathfrak{f}} = \mathbb{T} + 1$, the rank–nullity theorem yields $\acute{\mathbf{r}} = \mathbf{r}$. Finally, via linearity of the inner product, the Jacobi equations (19)–(21) hold for $u^{(\cdot,\cdot)}$ if, and only if, they hold for $\acute{\eta}^{(\cdot,\cdot)}$ and hence $\acute{u}^{(\cdot,\cdot)}$. □

The case where Y has cardinality 3, follows from the next two lemmas.

Lemma 3.3. *If* $Y = \{x, y, z\}$ *and* $\mathbf{r} = 2$, *then* *4*-Jac *and* *4*-prudence *hold on* Y.

Proof of lemma 3.3. Fix arbitrary $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total. We first apply Zaslavski's theorem to prove that the Y-extension $\acute{\mathcal{R}}$ of lemma 3.2 (which is test-worthy relative to J) satisfies $|\acute{\mathcal{G}}_{++}| = 6$. Via lemma 2.2, $|\acute{\mathcal{H}}_{++}| = 3$. And since every subarrangement of $\acute{\mathcal{H}}_{++}$ is central, for every $k = 0, 1, 2, 3$ there are $\binom{3}{k}$ ways to choose $|\mathcal{A}| = k$ hyperplanes from $\acute{\mathcal{H}}_{++}$. For $k < 3$, the rank of every subarrangement is k. For $k = 3$, the rank of the arrangement is $\acute{\mathbf{r}}$. So that, via lemma 3.2, since $\acute{\mathbf{r}} = \mathbf{r} = 2$,

$$|\acute{\mathcal{G}}_{++}| = \binom{3}{3}(-1)^{3-\acute{\mathbf{r}}} + \binom{3}{2}(-1)^{2-2} + \binom{3}{1}(-1)^{1-1} + \binom{3}{0}(-1)^{0-0} = 6. \qquad (6)$$

For both *4*-Jac and *4*-prudence, we require that every member of $\acute{\mathcal{G}}_{++}$ is associated with a total ordering. This ensures that *3*-diversity holds for $\acute{\mathcal{R}}$, so that, via lemma 2 of [GS], *3*-Jac holds for $\acute{\mathcal{R}}$ and lemma 3.2 yields *3*-Jac for $\mathcal{R}$. (When $|Y| = 3$, *3*-Jac and *4*-Jac coincide.) Also, for *4*-prudence, if $\hat{\mathcal{R}}$ is any novel, testable extension that satisfies $\hat{\mathcal{R}}_{\mathfrak{f}} = \mathcal{R}_J^{-1}$, then $\acute{\mathcal{R}}$ is a nondogmatic perturbation of $\hat{\mathcal{R}}$ that satisfies A0$^\flat$-A3$^\flat$.

It therefore remains for us to show that every member of $\acute{\mathcal{G}}_{++}$ is associated with a total ordering of Y. Since $\acute{\mathcal{H}}_{++}$ is central, $\acute{\mathcal{G}}_{++} \equiv \acute{\mathcal{G}}$, we suppress reference to $_{++}$. Let $\{\acute{G}^r : r = 1, \ldots 6\}$ be an enumeration of $\acute{\mathcal{G}}$ such that $\acute{G}^r$ and $\acute{G}^{r+1}$ (mod 6) are separated by a single hyperplane in $\acute{\mathcal{H}}$. Moreover, let the first four members of this enumeration intersect $\mathbb{J} \times \{0\}$. Then, since $\mathcal{R} = \langle \mathcal{R}_I : I \in \mathbb{J} \rangle$ satisfies A0$^\flat$, $\acute{G}^1, \ldots, \acute{G}^4$ are associated with total orders. Take $L \in \acute{G}^5 \cap \mathbb{J}^\mathfrak{f}$ and consider the affine path $\lambda \mapsto \phi(\lambda) = (1-\lambda)\acute{L} + \lambda \acute{J}$, where $\acute{J}$ is defined in lemma 3.2. For some rational λ^* sufficiently close to, but greater than, 1, $L^* = \phi(\lambda^*)$ belongs to $\acute{G}^* \cap \mathbb{J}^\mathfrak{f}$ for some $\acute{G}^* \in \acute{\mathcal{G}}$. Since $\phi(1) = \acute{J}$ belongs to the center of $\mathcal{H}$, L and L^* are separated by all three members of $\mathcal{H}$, so that $\acute{\mathcal{R}}_{L*} = \acute{\mathcal{R}}_L^{-1}$. Thus, $\acute{\mathcal{R}}_L$ is transitive if, and only if, $\acute{\mathcal{R}}_{L*}$ is. Since $\acute{G}^5$ is separated from $\acute{G}^*$ by 3 hyperplanes, $\acute{G}^* = \acute{G}^2$. Thus $\acute{\mathcal{R}}_{L*}$ is transitive. *Mutatis mutandis*, the same argument shows that $\acute{G}^6$ is also associated with a total ordering of Y. □

Lemma 3.4. *If* $Y = \{x, y, z\}$ *and* $\mathbf{r} = 3$*, then neither 4*-prudence *nor 4*-Jac *hold.*

Proof of lemma 3.4. When $\mathbf{r} = 3$, $u^{(x,y)}, u^{(y,z)}$ and $u^{(x,z)}$ are linearly independent, so that *3*-Jac fails to hold.

We now confirm that *3*-prudence also fails to hold. Via lemma 3.2, $\acute{\mathbf{r}} = \mathbf{r}$. The only term in eq. (6) that we adjust for the present lemma is $\mathbf{r} = 3$. Thus,

$$|\acute{\mathcal{G}}| = (-1)^0 + 3(-1)^0 + 3(-1)^0 + (-1)^0 = 8.$$

Then there are $3! = 6$ members of total$(\acute{\mathcal{R}})$, and the two additional regions of $\acute{\mathcal{G}}$ are associated with intransitive CAR rankings. It remains for us to show that every Y-extension $\hat{\mathcal{R}}$ with $\sharp$total$(\hat{\mathcal{R}}) = 6$ is of this form.

In the Hasse diagram of fig. 2, an increase in level corresponds to a decrease in dimension: since $\acute{A}^{\{x,y,z\}}$ is nonempty, it is of dimension at least zero. Since $\acute{J}$ belongs to the interior of $\mathbb{R}^{\mathbb{T}^\mathfrak{f}}_{++}$ and $\acute{A}^{\{x,y\}\{y,z\}}$ is at least one-dimensional, $\acute{A}^{\{x,y\}\{y,z\}}_{++}$ is one-dimensional. Since $\acute{A}^{\{x,y,z\}} \subset \acute{A}^{\{x,y\}\{y,z\}}$, the latter set is nonempty whenever $\acute{\mathcal{H}}_{++}$ is central. The same, of course, applies to other members at the same level. Conversely, if $\acute{A}^{\{x,y\}\{y,z\}}$ is empty, then so is $\acute{A}^{\{x,y,z\}}$. We now use this to show that there is a unique form of Y-extension $\hat{\mathcal{R}}$ such that $\sharp\hat{\mathcal{G}}_{++} = 6$ and, moreover, that any such $\hat{\mathcal{R}}$ fails to satisfy A0$^\flat$.

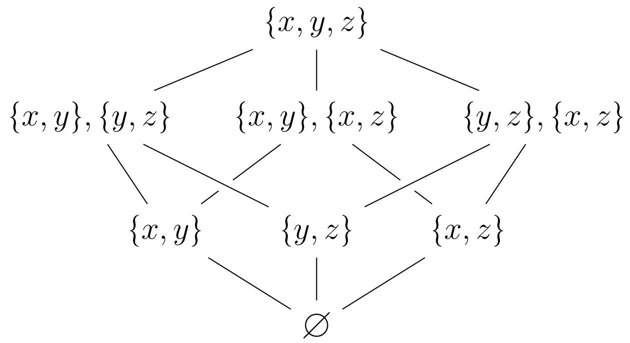

Figure 2: The intersection lattice of a central arrangement for $\sharp Y = 3$ and $\mathbf{r} = 3$.

Recall from example 6 that $\boldsymbol{\mu}(A^{\varnothing}) = 1$ and $\boldsymbol{\mu}(A^{\{i,j\}}) = -\boldsymbol{\mu}(A^{\varnothing})$. Then, since $\hat{A}^{\{x,y\}\{y,z\}} \subset \hat{A}^{\{x,y\}}, \hat{A}^{\{y,z\}}$, we have $\boldsymbol{\mu}(A^{\{x,y\}\{y,z\}}) = -\boldsymbol{\mu}(A^{\varnothing})(1 - 2) = 1$. Finally, by the same argument, since $\hat{A}^{\{x,y,z\}}$ belongs to every member of $\hat{\mathcal{L}}$, we have $\boldsymbol{\mu}(A^{\{x,y,z\}}) = -\boldsymbol{\mu}(A^{\varnothing})(1 - 3 + 3) = -1$. Thus, when $\mathbf{r} = 3$, if $\sharp\hat{\mathcal{G}}_{++} = 6$, then, in addition to $\hat{A}^{\{x,y,z\}}_{++}$, at most one of the intersections $\hat{A}^{\{i,j\}\{k,l\}}_{++}$ is empty. Then, w.l.o.g., suppose $\hat{A}^{\{x,y\}\{y,z\}}_{++}$ is nonempty. We will show that A0$^\flat$ fails to hold. Take $L \in \hat{A}^{\{x,y\}\{y,z\}}_{++}$. If $L \in \mathbb{J}^{\mathfrak{f}}$, then since $x\ \hat{\mathcal{I}}_L\ y\ \hat{\mathcal{I}}_L\ z$ and $\neg(x\ \hat{\mathcal{I}}_L\ z)$, the proof is complete. In any case, L belongs to one $\hat{G}^{(x,z)}_{++}$ or $\hat{G}^{(z,x)}_{++}$. W.l.o.g., suppose $L \in \hat{G}^{(x,z)}_{++}$. Then $\hat{H}^{\{x,y\}}$ and $\hat{H}^{\{y,z\}}$ split $\hat{G}^{(x,z)}_{++}$ into the four (nonempty) open regions formed by expanding

$$\hat{G}^{(x,z)}_{++} - \left(\hat{H}^{\{x,y\}} \cap \hat{H}^{\{y,z\}}\right) = \hat{G}^{(x,z)}_{++} \cap \left(\hat{G}^{(x,y)}_{++} \cup \hat{G}^{(y,x)}_{++}\right) \cap \left(\hat{G}^{(y,z)}_{++} \cup \hat{G}^{(z,y)}_{++}\right).$$

Take any K in the final member $\hat{G}^{(x,z)}_{++} \cap \hat{G}^{(y,x)}_{++} \cap \hat{G}^{(z,y)}_{++}$ of this expansion. Then, since $x\ \hat{\mathcal{P}}_K\ z\ \hat{\mathcal{P}}_K\ y\ \hat{\mathcal{P}}_K\ x$, we observe that $\hat{\mathcal{R}}$ fails to satisfy A0$^\flat$. We conclude that only dogmatic perturbations of $\acute{\mathcal{R}}$ satisfy A0$^\flat$. $\square$

The case where Y has cardinality 4. Note that a failure of *3*-Jac on $Z \subset Y$ such that $|Z| = 3$ implies a failure of *4*-Jac on Y. And since the arguments for the case where $|Y| = 3$ account for the case where *3*-Jac fails, we henceforth assume that *3*-Jac holds on Y. That is, our conditionally *2*-diverse representation $u^{(\cdot,\cdot)}$ will now satisfy equations (19)–(21) with $\hat{\beta} = \beta$ if, and only if, *4*-Jac holds on Y.

First some some useful results that exploit *3*-Jac.

Proposition 4. *If $Y = \{x, y, z, w\}$ and 3-*Jac *holds for $u^{(\cdot,\cdot)}$, then, for every $\acute{v}^{(\cdot,\cdot)} = u^{(\cdot,\cdot)} \times \acute{\eta}^{(\cdot,\cdot)}$ with rank $\acute{\mathbf{r}}$, that satisfies 3-*Jac*, $2 \leqslant \acute{\mathbf{r}} \leqslant 3$.*

Proof of proposition 4. Via proposition 3 and lemma 3.3, $\mathbf{r} \geqslant 2$. Indeed the span of $\{u^{(x,w)}, u^{(y,w)}\}$ is two. Let S denote the span of $\{u^{(x,w)}, u^{(y,w)}, u^{(z,w)}\}$. Since $u^{(y,w)} = -u^{(w,y)}$ and $u^{(\cdot,\cdot)}$ satisfies *3*-Jac, equations (19)–(21) hold for $u^{(\cdot,\cdot)}$. (If *4*-Jac fails to hold, then $\beta \neq \hat{\beta}$, but the equations still hold.) Thus $u^{(x,y)}$, $u^{(y,z)}$ and $u^{(x,z)}$ all belong to S and $\mathbf{r} \leqslant 3$. Now note that above argument does not depend on the cardinality of $\mathbb{T}$, thus take $\acute{\eta}^{(\cdot,\cdot)}$ to satisfy *3*-Jac: indeed with the same parameters that feature in equations (19)–(21) for $u^{(\cdot,\cdot)}$. The preceding argument then extends *mutatis mutandis* to $\acute{v}^{(\cdot,\cdot)}$ and $2 \leqslant \acute{\mathbf{r}} \leqslant 3$. □

Proposition 5. *If $X = \{x, y, z, w\}$, then $4 \leqslant |\mathcal{H}| \leqslant 6$ and these bounds are tight. If, moreover, 3-*Jac *holds for $u^{(\cdot,\cdot)}$ and $|\mathcal{H}| < 6$, then $\mathbf{r} = 2$.*

Proof of proposition 5. The upper bound $|\mathcal{H}| \leqslant 6$ follows from the fact that there are $\binom{4}{2} = 6$ ways to choose distinct pairs of elements from the four-element set X. Via lemma 2.2 only the following equalities are feasible: $H^{\{x,y\}} = H^{\{z,w\}}$, $H^{\{x,z\}} = H^{\{y,w\}}$ and $H^{\{x,w\}} = H^{\{y,z\}}$. Via proposition 2, w.l.o.g., take $G^{(x,y,z,w)}$ and $G^{(w,z,y,x)}$ to belong to $\mathcal{G}$. We consider a convex path ϕ from (some member of) $G^{(x,y,z,w)}$ to $G^{(w,z,y,x)}$ such that $\sharp\{G \in \mathcal{G} : (\operatorname{image}\phi) \cap G \neq \varnothing\}$ is maximal. Via A0$^\flat$, one of $H^{\{x,y\}}, H^{\{y,z\}}$ and $H^{\{z,w\}}$ supports $G^{(x,y,z,w)}$. If it is $H^{\{y,z\}}$, then so does $H^{\{x,w\}}$ and, and since x and w lie at opposite ends of the ordering (x, y, z, w), this would imply that $|\mathcal{H}| = 1$ in violation of lemma 2.2. Since we are seeking a greatest lower bound on $|\mathcal{H}|$, take $H_1 = H^{\{x,y\}} = H^{\{z,w\}}$ to be the first hyperplane to intersect $\operatorname{image}\phi$ and the adjacent chamber is $G^{(y,x,w,z)}$. Via A0$^\flat$, the second hyperplane to intersect $\operatorname{image}\phi$ is $H_2 \stackrel{\text{def}}{=} H^{\{x,w\}}$. If $H^{\{x,w\}} = H^{\{y,z\}}$ then since y and z lie at opposite ends of (y, x, w, z), this would once again lead to a violation of lemma 2.2. Then $G^{(y,w,x,z)}$ is next chamber to intersect $\operatorname{image}\phi$. Since we are seeking a greatest lower bound, via A0$^\flat$, take $H_3 \stackrel{\text{def}}{=} H^{\{y,w\}} = H^{\{x,z\}}$ to be the third hyperplane to intersect $\operatorname{image}\phi$. Then $G^{(w,y,z,x)}$ is the next chamber to intersect $\operatorname{image}\phi$. The final hyperplane to intersect $\operatorname{image}\phi$ is $H_4 = H^{\{y,z\}}$ which brings us to the destination chamber $G^{(w,z,y,x)}$. We conclude that at most two pairs of the six hyperplanes coincide, so that $|\mathcal{H}| \geqslant 4$.

We now prove that *3*-Jac and $H^{\{x,y\}} = H^{\{z,w\}}$ together imply $\mathbf{r} = 2$. Consider equations (19)–(21) (so *3*-Jac holds, but *4*-Jac need not). Via (19), $S = \{u^{(x,w)}, u^{(w,y)}, u^{(x,y)}\}$ is 2-dimensional. Since $u^{(x,y)}$ and $u^{(z,w)}$ are collinear, $u^{(z,w)}$

belongs to S. Finally, equations (20) and (21) yield $u^{yz}, u^{(x,z)} \in S$. □

Lemma 3.5. *If* $Y = \{x, y, z, w\}$, $\mathbf{r} = 3$ *and* *3*-Jac *holds on* Y, *then* *4*-prudence *and* *4*-Jac *hold on* Y

Proof of lemma 3.5. To see that, *4*-Jac holds, we appeal to the proof of lemma 3 of [GS]: if *3*-Jac holds and *4*-Jac does not, then $\{u^{(x,w)}, u^{(y,w)}, u^{(z,w)}\}$ is linearly dependent. This is in contradiction of $\mathbf{r} = 3$.

It remains for us to verify that *4*-prudence also holds. We will show that, for every $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total, the Y-extension $\acute{\mathcal{R}}$ of lemma 3.2 satisfies $|\acute{\mathcal{G}}_{++}| = 4!$. We then confirm that $\acute{\mathcal{R}}$ satisfies A0$^\flat$, so that every member of $\acute{\mathcal{G}}$ is associated with a total ordering of Y. Then total($\acute{\mathcal{R}}$) is maximal and, if $\check{\mathcal{R}}$ is any novel, testworthy extension that satisfies $\check{\mathcal{R}}_{\mathfrak{f}} = \mathcal{R}_J^{-1}$, then $\acute{\mathcal{R}}$ is a nondogmatic perturbation of $\check{\mathcal{R}}$ that satisfies A0$^\flat$-A3$^\flat$.

Since $\mathbf{r} = 3$, the contrapositive of proposition 5 implies that $|\mathcal{H}_{++}| = 6$. The proof that now follows is a simple extension of lemma 3.3 to allow for the fact that $|Y| = 4$. Fix any $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total and w.l.o.g. take $J > 0$. Via lemma 3.2, there exists a testworthy Y-extension $\acute{\mathcal{R}}$ such that $\acute{\mathcal{R}}_{\mathfrak{f}} = \mathcal{R}_J^{-1}$ is total and $L = (1 - \lambda)J \times \lambda$ belongs to both $\mathbb{J}^{\mathfrak{f}}$ and the center of $\acute{\mathcal{H}}_{++}$ and $\acute{v}^{(\cdot,\cdot)}$ satisfies *4*-Jac. Via remark 2, since $\acute{\mathcal{H}}$ is central, $\acute{\mathcal{G}}_{++}$ and $\acute{\mathcal{G}}$ are isomorphic, so we work with $\acute{\mathcal{G}}$. Since $\mathbf{r} = 3$, proposition 5 implies $|\mathcal{H}| = 6$, so that the same is true of $|\acute{\mathcal{H}}|$. The rank of subarrangements with cardinality 4 or more is $\acute{\mathbf{r}}$. Let $\acute{\tau}$ denote the number of subarrangements $\mathcal{A}$ that have cardinality 3 and rank 2. Each of the other $\binom{6}{3} - \acute{\tau}$ subarrangements with cardinality 3 have rank $\acute{\mathbf{r}}$. All other subarrangements have rank equal to their cardinality.

$$
\begin{aligned}
|\acute{\mathcal{G}}| = & \binom{6}{6}(-1)^{6-\acute{\mathbf{r}}} + \binom{6}{5}(-1)^{5-\acute{\mathbf{r}}} + \binom{6}{4}(-1)^{4-\acute{\mathbf{r}}} \\
& + \binom{6}{3}(-1)^{3-\acute{\mathbf{r}}} - \quad \acute{\tau} \quad (-1)^{3-\acute{\mathbf{r}}} + \quad \acute{\tau} \quad (-1)^{3-2} \\
& + \binom{6}{2}(-1)^{2-2} + \binom{6}{1}(-1)^{1-1} + \binom{6}{0}(-1)^{0-0}
\end{aligned} \tag{7}
$$

Via lemma 3.2, $\acute{\mathbf{r}} = \mathbf{r}$, so that $\acute{\mathbf{r}} = 3$. It remains for us to calculate the value of $\acute{\tau}$. Each of the $\binom{4}{3} = 4$ subsets of Y that have cardinality 3 generates a subarrangement of cardinality $\binom{3}{2} = 3$. (For instance, $\mathcal{A}^{\{x,y,z\}} = \{\acute{H}^{\{x,y\}}, \acute{H}^{\{y,z\}}, \acute{H}^{\{x,z\}}\}$.) For such subarrangements, *4*-Jac implies a rank of 2. Arguments from the final step in the proof of proposition 5 confirm that every other subarrangement with cardinality 3 has rank 3. Substituting into eq. (7), we obtain

$$|\acute{\mathcal{G}}| = -1 + 6 - 15 + 20 - 4 - 4 + 15 + 6 + 1 = 24 = 4!.$$

The proof that every member of $\mathcal{G}$ is associated with a transitive ordering is as follows. Via lemma 3.2, *4*-Jac holds for $\acute{\mathcal{R}}$ if, and only if, it holds for $\mathcal{R}$. We have seen that *4*-Jac holds. Then, via a straightforward application of the Jacobi identity and, for arbitary $K \in \mathbb{J}^{\mathfrak{f}}$, the definition of transitivity of $\acute{\mathcal{R}}_K$ yields $\mathrm{A0}^{\flat}$. Thus, every member of $\acute{\mathcal{G}}$ is total. □

In the remaining case, where $Y = \{x, y, z, w\}$ and $\mathbf{r} = 2$, the proof is complicated by the fact that maximally-diverse extensions have centerless arrangements. We begin by choosing $\acute{\eta}^{(\cdot,\cdot)}$ so as to construct $\acute{u}^{(\cdot,\cdot)} = u^{(\cdot,\cdot)} \times \acute{\eta}^{(\cdot,\cdot)}$ with $\acute{\mathbf{r}} = 3$.

Since $\mathbf{r} = 2$, it follows that $u^{(x,z)}, u^{(y,z)}$ and $u^{(w,z)}$ form a linearly dependent set. Thus, for some $\pi, \rho \in \mathbb{R}$,

$$\pi u^{(x,z)} + \rho u^{(y,z)} = u^{(w,z)}. \tag{8}$$

Fix arbitrary $J \in \mathbb{J}$ such that $\mathcal{R}_J$ is total, and, as in lemma 3.2, w.l.o.g., we take $J \in G_{++}^{(x,y,z,w)}$. Then, for $\acute{\iota} = \frac{1}{2}$ and $\acute{J}^{\{x,y,z\}} = (1 - \acute{\iota})J \times \acute{\iota}$, let

$$\acute{\eta}^{(i,j)} = -\tfrac{1-\acute{\iota}}{\acute{\iota}}\langle u^{(i,j)}, J\rangle = -\langle u^{(i,j)}, J\rangle, \quad \text{for every } i, j \in \{x, y, z\}. \tag{9}$$

Note that $\acute{\eta}^{(y,z)} > 0$ is determined by eq. (9). Since $G_{++}^{(x,y,z,w)}$ is open and the inner product is continuous, there exists a compact neighbourhood $N_J \subset G_{++}^{(x,y,z,w)}$ of J such that, for every $L \in N_J$, $\langle u^{(y,z)}, L\rangle < -\langle u^{(w,z)}, L\rangle$ if, and only if, $\langle u^{(y,z)}, J\rangle < -\langle u^{(w,z)}, J\rangle$. For any such $L \neq J$, take $\acute{\lambda}$ and ϵ to be solutions to

$$\acute{\eta}^{(y,z)} = -\tfrac{1-\acute{\lambda}}{\acute{\lambda}}\langle u^{(y,z)}, L\rangle \quad \text{and} \quad \epsilon = \langle u^{(w,z)}, \tfrac{1-\acute{\lambda}}{\acute{\lambda}}L - J\rangle. \tag{10}$$

A rearrangement of the first equality in eq. (10) and, via eq. (9), a substitution for $\acute{\eta}^{(y,z)}$ yields $\acute{\lambda} = \frac{\langle u^{(y,z)},L\rangle}{\langle u^{(y,z)},J\rangle + \langle u^{(y,z)},L\rangle}$ and $\frac{1-\acute{\lambda}}{\acute{\lambda}} = \frac{\langle u^{(y,z)},J\rangle}{\langle u^{(y,z)},L\rangle}$. Since $N_J \subset G^{(y,z)}$, $0 < \acute{\lambda} < 1$ and $\acute{L}^{\{y,z,w\}} \overset{\text{def}}{=} (1 - \acute{\lambda})L \times \acute{\lambda} \in \mathbb{R}_{++}^{\mathbb{T}^{\mathfrak{f}}}$ for every such L. On N_J, the denominator of the map $L \mapsto \acute{L}^{\{x,y,z\}}$ is bounded away from zero, and as the quotient of continuous functions the map continuous, and, furthermore, $\lim_{L \to J} \acute{L}^{\{y,z,w\}} = \acute{J}^{\{y,z,w\}}$.

Similarly, on N_J, the fact that $\langle u^{(y,z)}, L\rangle > 0$ ensures that the map $L \mapsto \epsilon$ is well-defined, continuous and $\lim_{L\to J} \epsilon = 0$.[17]

For every $i, j \in \{y, z, w\}$, let $\acute{\eta}^{(i,j)} = -\frac{1-\acute{\lambda}}{\acute{\lambda}}\langle u^{(i,j)}, L\rangle$. Now note that, via equations (8)–(10), we obtain

$$\pi\acute{\eta}^{(x,z)} + \rho\acute{\eta}^{(y,z)} = \acute{\eta}^{(w,z)} + \epsilon \neq \acute{\eta}^{(w,z)}, \tag{12}$$

so that, for every $\epsilon \neq 0$, $\{\acute{u}^{(i,j)} = u^{(i,j)} \times \acute{\eta}^{(i,j)} : (i,j) = (x,z), (y,z), (w,z)\}$ forms a linearly independent set. Let $\{\acute{u}^{(i,j)} : i, j \in \{x, y, z\} \cup \{y, z, w\}\}$ be defined similarly. To complete the definition of $\acute{u}^{(\cdot,\cdot)}$, we appeal to the fact that, via *3*-Jac, $u^{(\cdot,\cdot)}$ satisfies equations (19)–(21). In particular, from these equations, take parameters α, β, and γ and let $\acute{\eta}^{(x,w)}$ be the (unique) solution to the Jacobi identity

$$\alpha\acute{\eta}^{(x,w)} = \gamma\acute{\eta}^{(x,y)} + \beta\acute{\eta}^{(y,w)} = -\langle\gamma u^{(x,y)}, J\rangle - \tfrac{1-\acute{\lambda}}{\acute{\lambda}}\langle\beta u^{(y,w)}, L\rangle. \tag{13}$$

For the same set of parameter values, $\acute{u}^{(\cdot,\cdot)}$ also satisfies (19)–(21). That is, for $\{x, y, z\}$, via eq. (9) and (21), $\gamma\acute{\eta}^{(x,y)} + \tau\acute{\eta}^{(y,z)} = \phi\acute{\eta}^{(x,z)}$, so that $\acute{u}^{(\cdot,\cdot)}$ satisfies (21). For $\{y, z, w\}$, Via eq. (10) and (20), $\hat{\beta}\acute{\eta}^{(y,w)} + \sigma\acute{\eta}^{(w,z)} = \tau\acute{\eta}^{(y,z)}$, so that $\acute{u}^{(\cdot,\cdot)}$ satisfies (20). For $\{x, y, w\}$, via eq. (13), $\acute{u}^{(\cdot,\cdot)}$ satisfies (19). We now demonstrate that for the final triple $\{x, z, w\}$, the Jacobi identity holds if $\hat{\beta} = \beta$, and $\{\acute{u}^{(x,w)}, \acute{u}^{(w,z)}, \acute{u}^{(x,z)}\}$ has rank 3 otherwise.

First extract the parameters from equations (19)–(21) to obtain the matrix form

$$\begin{array}{c c} \begin{array}{cccccc} (x,w) & (y,w) & (x,y) & (w,z) & (y,z) & (x,z) \end{array} & \\ \left[\begin{array}{ccc|ccc} \alpha & -\beta & -\gamma & 0 & 0 & 0 \\ 0 & \hat{\beta} & 0 & \sigma & -\tau & 0 \\ 0 & 0 & \gamma & 0 & \tau & -\phi \end{array}\right] & \begin{array}{c} (19) \\ (20) \\ (21) \end{array} \end{array} \tag{14}$$

[17]Substituting for $\frac{1-\acute{\lambda}}{\acute{\lambda}}$ in the equation for ϵ in (10) and rearranging yields a linear equation

$$k\langle u^{(y,z)}, L\rangle = \langle u^{(w,z)}, L\rangle \tag{11}$$

in L, where $k = \frac{\epsilon + \langle u^{(w,z)}, J\rangle}{\langle u^{(y,z)}, J\rangle}$. Note that, since $L \in G_{++}^{(z,w)} \cap G_{++}^{(y,z)}$, $\langle u^{(w,z)}, L\rangle < 0$ and $\langle u^{(y,z)}, L\rangle > 0$; moreover, the same is true of J. Thus, $k < 0$ or, equivalently, $\epsilon < -\langle u^{(w,z)}, J\rangle$. If $\langle u^{(y,z)}, L\rangle < -\langle u^{(w,z)}, L\rangle$, then, via eq. (11), $k < -1$. Moreover, our choice of N_J is such that $\frac{\langle u^{(w,z)}, J\rangle}{\langle u^{(y,z)}, J\rangle} < -1$. This ensures that, on N_J, ϵ lies in a neighbourhood of zero, as required.

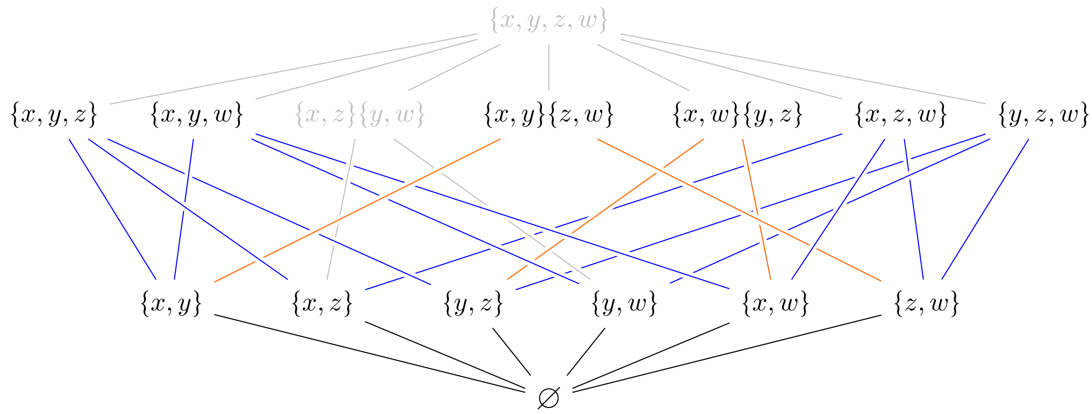

Figure 3: The intersection semilattices $\acute{\mathcal{L}}$ and $\acute{\mathcal{L}}_{++} = \acute{\mathcal{L}} - \{A^Y, \acute{A}^{\{x,z\}\{y,w\}}\}$ when $\sharp\mathbb{T} = 2$, $\hat{\beta} = \beta$ and ϵ is sufficiently small but distinct from zero.

Since the triple $\{\acute{u}^{(i,z)} : i = x, y, w\}$ provides a basis for $\operatorname{span}(\acute{u}^{(\cdot,\cdot)})$, we will write all vectors in terms of this basis. To this end, we derive the reduced row echelon form of eq. (14). In particular, letting r_i denote the rows of the matrix, we perform the operation $r_1 \mapsto r_1 + \frac{\beta}{\hat{\beta}} r_2 + r_3$ to obtain

$$\begin{array}{c}
\begin{array}{cccccc} (x,w) & (y,w) & (x,y) & (w,z) & (y,z) & (x,z) \end{array} \\
\left[\begin{array}{ccc|ccc}
\alpha & 0 & 0 & \frac{\beta}{\hat{\beta}}\sigma & (1-\frac{\beta}{\hat{\beta}})\tau & -\phi \\
0 & \hat{\beta} & 0 & \sigma & -\tau & 0 \\
0 & 0 & \gamma & 0 & \tau & -\phi
\end{array}\right]
\begin{array}{c} (19) \\ (20) \\ (21) \end{array}
\end{array} . \tag{15}$$

In eq. (15), the fact that $\hat{\beta}$ (instead of β) that appears as a pivot in column 2 indicates that, in this derivation, we are, w.l.o.g., choosing $\acute{v}^{(y,w)} = \hat{\beta}\acute{u}^{(y,w)}$. *Mutatis mutandis*, the conclusions we will draw are the same if instead we chose to define $\acute{v}^{(y,w)}$ using β. The other (relevant) rows of $\acute{v}^{(\cdot,\cdot)} : Y^2 \times \mathbb{T}^{\mathfrak{f}} \to \mathbb{R}$ are $\acute{v}^{(x,w)} = \alpha\acute{u}^{(x,w)}$, $\acute{v}^{(x,y)} = \gamma\acute{u}^{(x,y)}$, $\acute{v}^{(z,w)} = \sigma\acute{u}^{(z,w)}$, $\acute{v}^{(y,z)} = \tau\acute{u}^{(y,z)}$ and $\acute{v}^{(x,z)} = \phi\acute{u}^{(x,z)}$. The matrix of the equation that now follows, is invertible if, and only if, $(1-\frac{\beta}{\hat{\beta}}) \neq 0$.

$$\begin{bmatrix} \acute{v}^{(x,w)} \\ \acute{v}^{(x,z)} \\ \acute{v}^{(z,w)} \end{bmatrix} = \begin{bmatrix} -\frac{\beta}{\hat{\beta}} & -(1-\frac{\beta}{\hat{\beta}}) & 1 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \acute{v}^{(w,z)} \\ \acute{v}^{(y,z)} \\ \acute{v}^{(x,z)} \end{bmatrix} \tag{16}$$

Thus, unless $\hat{\beta} = \beta$, we conclude that $\{\acute{v}^{(x,w)}, \acute{v}^{(x,z)}, \acute{v}^{(z,w)}\}$ has the same rank as

$\{\acute{v}^{(w,z)}, \acute{v}^{(y,z)}, \acute{v}^{(x,z)}\}$ which, by construction, has rank 3 for every choice of $\epsilon \neq 0$.

Since $\hat{\beta} = \beta$ if, and only if, *4*-Jac holds for $\mathcal{R}$, we conclude that *4*-Jac holds for $\mathcal{R}$ if, and only if, it holds for $\acute{\mathcal{R}}$. It remains for us to show that, for ϵ sufficiently small $\acute{\mathcal{G}}$ is maximal. For if $\acute{\mathcal{G}}$ is maximal, then via lemma 3.3 and lemma 3.4, A0$^\flat$ holds if, and only if, $\operatorname{rank}\{\acute{v}^{(x,w)}, \acute{v}^{(x,z)}, \acute{v}^{(z,w)}\} = 2$.

By Zaslavski's theorem, it suffices to show that every member of the intersection lattice $\acute{\mathcal{L}}$, other than the center $\acute{A}^{\{x,y,z,w\}}$, has nonempty intersection with $\mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}_{++}$. We now make explicit the dependence of the extension $\acute{\mathcal{R}}$ on our choice of $L \in N_J$, though we do so indirectly via ϵ. Let $d^\epsilon = \max\{\mathrm{d}(\acute{J}, \acute{A}) : \acute{A} \in \acute{\mathcal{L}}^\epsilon$, where $\mathrm{d}(\acute{J}, \acute{A})$ is the minimum (Euclidean) distance between $\acute{J}$ and the (closed) linear subspace $\acute{A}$ of $\mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}$. Note that for $\epsilon = 0$, we obtain a central arrangement of the form of lemma 3.2 with $d^\epsilon = 0$. Moreover, since the Euclidean metric is continuous in its arguments, and, for every ϵ, $\acute{\mathcal{L}}^\epsilon$ is finite, the map $\epsilon \mapsto d^\epsilon$ is continuous and $\lim_{\epsilon \to 0} d^\epsilon = 0$. Thus, for sufficiently small $\epsilon \neq 0$, every $\acute{A} \in \acute{\mathcal{L}}^\epsilon$ intersects $\mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}_{++}$.

Remark. *We note that the above arguments apply without modification to the case where $\sharp\mathcal{H} = 4, 5$. Consider, for example, the Hasse diagram of fig. 3. That case arises when $u^{(x,z)}$ and $u^{(y,w)}$ are collinear, so that $A^{\{xz\}\{y,w\}}$ is a hyperplane of dimension $\sharp\mathbb{T} - 1$. Assuming the same construction, with $\epsilon \neq 0$, so that $\acute{u}^{(\cdot,\cdot)}$ has rank 3 and, via proposition 5, $\acute{u}^{(x,z)}$ and $\acute{u}^{(y,w)}$ are linearly independent. Thus $\acute{A}^{\{xz\}\{y,w\}}$ is of dimension $\sharp\mathbb{T}^{\mathfrak{f}} - 2 = \sharp\mathbb{T} - 1$. Thus, $\acute{A}^{\{x,z\}\{y,w\}} = A^{\{x,z\}\{y,w\}} \times \{0\}$ which belongs to the boundary of $\mathbb{R}^{\mathbb{T}^{\mathfrak{f}}}_{++}$. At $\epsilon = 0$, $\acute{A}^{\{x,z\}\{y,w\}}$ increases by one dimension and the upper two levels of the Hasse diagram collapse to equal $A^{\{x,y,z,w\}}$.*

Online Appendix D. Proofs

Proof of Observation 2. This follows directly from proposition 3 and the construction of $\preceq_{\mathbb{J}}$. $\square$

Proof of Observation 3. Respectively, these two statements follow via lemma 3.2 and lemma 3.1. $\square$

Proof of Corollary 1. This follows from theorem 3, lemma 2.4 and the fact that, via lemma 3.1, 4-prudence implies A0–A3 when the number of case types is finite. $\square$

Proof of Proposition 1. Via lemma 1.3, A1–A3 and 4-diversity hold for $\preceq_{\mathbb{D}}$ if and only if A1$^\flat$–A3$^\flat$ and 4-diversity hold for $\preceq_{\mathbb{J}}$. Let $Y \subseteq X$ be of cardinality $m' = 1, 2, 3$ or 4 and let $\mathcal{R}$ be a regular Y-extension. Via lemma 2.1, there exists a pairwise representation $v^{(\cdot,\cdot)}$. For $m' = 1$, $n' = 1$ because $\mathcal{R}_J$ is constant on $\mathbb{J}^{\mathfrak{f}}$. For $m' = 2$, $n' \geqslant 2$, since via part (ii) of lemma 2.1, $G^{(x,y)}$ and $G^{(y,x)}$ are both nonempty.

By way of contradiction, first suppose $n' = 2$ and $m' \geqslant 3$. Via example 6 of appendix B, $\mathrm{total}(\mathcal{R}) \leqslant 4$. In contrast, 4-diversity requires $\mathrm{total}(\mathcal{R}) = 6$. The remaining case is where $n' = 3$ and $m' \geqslant 4$. If the rank $\mathbf{r}$ of $v^{(\cdot,\cdot)}$ satisfies $\mathbf{r} \geqslant 3$, then the kernel A^Y of $v^{(\cdot,\cdot)}$ is zero-dimensional. Then 0 is the unique element of A^Y. Thus, the positive kernel A^Y_{++} of $v^{(\cdot,\cdot)}$ is empty. Then Zaslavski's theorem implies that $\mathrm{total}(\mathcal{R}) < 4!$, so that 4-diversity fails to hold. If $\mathbf{r} \leqslant 2$, then an application of the rank version Zaslavski's theorem (in particular eq. (7) with $\acute{\mathbf{r}} = \mathbf{r} = 2$) yields

$$|\mathcal{G}_{++}| \leqslant 1 - 6 + 15 - 20 + 15 + 6 + 1 = 12.$$

Thus, once again 4-diversity fails to hold. Thus $n' \geqslant \min\{4, m'\}$, as required. Finally, since $Y \subseteq X$, $m \geqslant m'$, and, via part 3 of definition 1, $n \geqslant n'$. $\square$

Proof of Lemma 1.1. We show that there exists a canonical embedding (a structure preserving injection) of $\mathrm{nov}(Y, \preceq_{\mathbb{J}})$ into $\mathrm{nov}(Y, \preceq_{\mathbb{D}})$. The fact that this map is also surjective follows from the fact that $\mathrm{nov}(Y, \preceq_{\mathbb{D}})$ can be embedded in $\mathrm{nov}(Y, \preceq_{\mathbb{J}})$ in precisely the same way. The proof that the two sets of regular extensions are isomorphic follows via a similar argument plus the observation that every Y-extension is either regular or novel.

Take $\mathcal{R} \in \mathrm{nov}(Y, \preceq_{\mathbb{J}})$ and define $\hat{\mathcal{R}} = \langle \hat{\mathcal{R}}_C : C \in \mathbb{D}^{\mathfrak{f}} \rangle$ via the property: for each $C \in \mathbb{D}^{\mathfrak{f}}$, $\hat{\mathcal{R}}_C \overset{\text{def}}{=} \mathcal{R}_J$ if, and only if, $L_C = L_J$, where, as before, $t \mapsto L_C(t)$ counts the number of cases of type t in C and $L_J = \kappa_J J \in \mathbb{L}^{\mathfrak{f}}$ for some minimal $\kappa_J \in \mathbb{Z}_+$. Now, for any $\mathcal{R}' \neq \mathcal{R}$ in $\mathrm{nov}(Y, \preceq_{\mathbb{J}})$, there exists $J \in \mathbb{J}^{\mathfrak{f}}$ such that $\mathcal{R}'_J \neq \mathcal{R}_J$. If we define $\hat{\mathcal{R}}'$ analogously, so that it is equivalent to $\mathcal{R}'$, then $\hat{\mathcal{R}}' \neq \hat{\mathcal{R}}$. As a consequence, the canonical mapping $\mathcal{R} \mapsto \hat{\mathcal{R}}$ is injective. If we can show that $\hat{\mathcal{R}}$ does in fact belong to $\mathrm{nov}(Y, \preceq_{\mathbb{D}})$, then we have constructed the required embedding. The fact that $\hat{\mathcal{R}}$ satisfies 1 and 2 of definition 1 follows immediately from definition 3. The proof that item 3 of definition 1 holds is as follows. Take any $c, c' \in \mathbb{C}^{\mathfrak{f}}$ and $D \in \mathbb{D}^{\mathfrak{f}}$ such that $c \sim^{\star} c'$ and $c, c' \notin D$. First, observe that $D \cup \{c\} \sim^{\star} D \cup \{c'\}$, and moreover, for some $t \in \mathbb{T}^{\mathfrak{f}}$ we have $c, c' \in t$. Then, for every $t \in \mathbb{T}^{\mathfrak{f}}$, $|D \cup \{c\}| = |D \cup \{c'\}| = L$ for some $L \in \mathbb{L}^{\mathfrak{f}} \cap \mathbb{J}^{\mathfrak{f}}$. Thus $\hat{\mathcal{R}}_{D \cup \{c\}} = \hat{\mathcal{R}}_{D \cup \{c'\}}$, as required for $\hat{\mathcal{R}}$ to be an extension of $\preceq_{\mathbb{D}}$. Finally, via definition 3, the definition of a novel extension ensures that the induced equivalence relation $\sim^{\mathcal{R}}$ on $\mathbb{C}^{\mathfrak{f}}$ satisfies $c \not\sim^{\mathcal{R}} \mathfrak{f}$ for every $c \in \mathbb{C}$. Since $\sim^{\hat{\mathcal{R}}}$ inherits this property, $\hat{\mathcal{R}}$ is novel. □

Proof of Lemma 1.3. Fix $\mathcal{R}_{\mathfrak{J}} \equiv \hat{\mathcal{R}}_{\mathfrak{D}}$ and assume that $\hat{\mathcal{R}}_{\mathfrak{D}}$ satisfies A2. We show that $\mathcal{R}_{\mathfrak{J}}$ satisfies A2$^\flat$. Fix $x, y \in Y$ and $J \in \mathfrak{J}$ such that $x \ \mathcal{R}_J \ y$ and $x \ \mathcal{R}_{J'} \ y$. Fix $\lambda, \mu \in \mathbb{Q}_{++}$ and let κ be the smallest positive integer such that both $L \overset{\text{def}}{=} \kappa\lambda J$ and $L' \overset{\text{def}}{=} \kappa\mu J'$ belong to $\mathfrak{L}$. Then, by lemma 1.2, we have both $x \ \mathcal{R}_L \ y$ and $x \ \mathcal{R}_{L'} \ y$. Moreover, for D, D' such that $L_D = L$ and $L_{D'} = L'$, we have $x \ \hat{\mathcal{R}}_D \ y$ and $x \ \hat{\mathcal{R}}_{D'} \ y$ and, by A2, $x \ \hat{\mathcal{R}}_{D \cup D'} \ y$. Finally, since $L_D + L_{D'} = \kappa(\lambda J + \mu J')$, one further application of lemma 1.2 yields $x \ \mathcal{P}_{\lambda J + \mu J'} \ y$, as required for A2$^\flat$.

The proof that "A2 implies A2$^\flat$" is *mutatis mutandis* a special case of the above argument and ommitted. We now assume $\hat{\mathcal{R}}_{\mathfrak{D}}$ satisfies A2 and A3 and prove that $\mathcal{R}_{\mathfrak{J}}$ satisfies A3$^\flat$. Fix $x, y \in X$ such that $x \ \mathcal{P}_J \ y$ for some $J \in \mathbb{J}$ and take any $J' \in \mathfrak{J}$. Then, by the construction of $\mathcal{R}_{\mathfrak{J}}$, there exists $L, L' \in \mathbb{L}$ such that $jJ = L$ and $j'J' = L'$ for some $j, j' \in \mathbb{Z}_{++}$. By lemma 1.2, $\mathcal{R}_L = \mathcal{R}_J$ and $\mathcal{R}_{L'} = \mathcal{R}_{J'}$. Moreover, by construction, for some D and D' such that $L_D = L$ and $I_{D'} = L'$, $\hat{\mathcal{R}}_D = \mathcal{R}_J$ and $\hat{\mathcal{R}}_{D'} = \mathcal{R}_{J'}$. We therefore conclude that $x \ \hat{\mathcal{P}}_D \ y$, so that A3 implies the existence of $\kappa \in \mathbb{Z}_{++}$ and $\{D_l : D_l \sim^{\hat{\mathcal{R}}} D\}_1^\kappa$ such that $x \ \hat{\mathcal{P}}_{D_1 \cup \cdots \cup D_\kappa \cup D'} \ y$. Then, by the construction of $\mathcal{R}_{\mathfrak{J}}$, $x\ \mathcal{P}_{\kappa L_D + L_{D'}}\ y$. Let $\nu \overset{\text{def}}{=} \frac{1}{\kappa j + j'}$ and take $\lambda = \nu j'$, so that $0 < \lambda < 0$ and $1 - \lambda = \nu\kappa j$. In fact, since $\lambda \in \mathbb{Q}$, we have

$$K \overset{\text{def}}{=} (1-\lambda)J + \lambda J' \in \mathfrak{J}.$$

Simplifying, we obtain $K = \nu(\kappa L + L')$. Since $\nu \in \mathbb{Q}_{++}$ and $\kappa L + L' \in \mathfrak{J}$, lemma 1.2 implies $\mathcal{R}_K = \mathcal{R}_{\kappa L + L'}$. This allows us to conclude that $x\ \mathcal{P}_K\ y$. Finally, take any $\mu \in \mathbb{Q} \cap (0, \lambda)$. From basic properties of the real numbers, there exists $\xi < 1$ such that $\mu = \xi\lambda$ and, moreover, ξ is rational. Next, note that the definition of K implies $\xi(K - J) = \xi\lambda(J' - J)$. Adding J to each side of the latter and applying the definition of μ yields

$$(1-\xi)J + \xi K = (1-\mu)J + \mu J'.$$

Then, since $x\ \mathcal{P}_J\ y$ and $x\ \mathcal{P}_K\ y$, A2$^\flat$ implies $x\ \mathcal{P}_{(1-\mu)J+\mu J'}\ y$, as required for A3$^\flat$.

Conversely, we now assume that $\mathcal{R}_{\mathfrak{J}}$ satisfies A2$^\flat$ and A3$^\flat$ and prove that A3 holds. Take $D, D' \in \mathbb{D}$ such that $x\ \hat{\mathcal{P}}_D\ y$ and any other $D' \in \mathbb{D}$. Let $L = L_D$ and $L' = L_{D'}$. Then, by construction, $x\ \mathcal{P}_L\ y$ and, by A3$^\flat$, there exists $\lambda \in \mathbb{Q} \cap (0,1)$ such that $x\ \mathcal{P}_{(1-\mu)L+\mu L'}\ y$. Then, since μ is rational, $\mu = j/k$ for some $j, k \in \mathbb{Z}_{++}$. Let $q := (1-\mu)/\mu = (k-j)/j$ and let $\kappa = jq$, so that $\kappa = k - j$. The fact that $0 < \mu < 1$ ensures that $\kappa \in \mathbb{Z}_{++}$. To complete the proof, we show that $x\ \mathcal{P}_{\kappa L + L'}\ y$, for then the existence of $D_1, \ldots, D_\kappa$ such that $x\ \hat{\mathcal{P}}_{D_1 \cup \cdots \cup D_\kappa \cup D'}$ immediately follows. Together $x\ \mathcal{P}_{(1-\mu)L+\mu L'}\ y$ and lemma 1.2 imply $x\ \mathcal{P}_{qL+L'}\ y$. Similarly, together $x\ \mathcal{P}_L\ y$ and lemma 1.2 imply $x\ \mathcal{P}_{(j-1)qL}\ y$. Then, since $(j-1)qL + (qL + L') = jqL + L'$ and $\kappa = jq$, an application of A2$^\flat$ yields $x\ \mathcal{P}_{\kappa L + L'}\ y$, as required. □

Proof of Lemma 2.1. In addition to A1$^\flat$–A3$^\flat$, the proof of lemma 1 of [GS] only appeals to *2*-diversity. That lemma, like the present one, does not require A0$^\flat$, or any diversity condition stronger than *2*-diversity since it is only result about distinct pairs of elements x and y in isolation.

W.l.o.g. we suppress reference to the acute accent. Modulo notation, lemma 1 of [GS] and its proof show that A1$^\flat$–A3$^\flat$, in addition to *2*-diversity on Y, imply the existence $v^{(\cdot,\cdot)}$ with rows satisfying properties (i)–(iv) of the present lemma.

We now prove the converse: that properties (i)–(iv) imply the axioms hold. For A1$^\flat$, fix arbitrary $J \in \mathbb{J}$. For every $x, y \in Y$, the fact that $\langle v^{(x,y)}, \cdot\rangle$ is real-valued and, via property (iii), equal to $-\langle v^{(y,x)}, \cdot\rangle$, ensures that J belongs to one of the sets $H^{\{x,y\}}_{++}$, $G^{(x,y)}_{++}$ and $G^{(y,x)}_{++}$. Then property (i) and the fact that J belongs to $\mathbb{Q}^{\mathfrak{T}}$ completes the proof. A2$^\flat$ and A3$^\flat$ hold by virtue of the fact that $\langle v^{(x,y)}, \cdot\rangle$ is linear on $\mathbb{R}^{\mathfrak{T}}_{++}$. Finally, we prove that *2*-diversity holds. Take any distinct $x, y \in Y$. Then, via property (ii), both $G^{(x,y)}_{++}$ and $G^{(y,x)}_{++}$ are nonempty. By continuity of $\langle v^{(x,y)}, \cdot\rangle$, $G^{(x,y)}_{++}$ and $G^{(y,x)}_{++}$ are also open in $\mathbb{R}^{\mathfrak{T}}_{++}$. As such they each contain a rational vector, so that, by property (i), *2*-diversity holds on $\{x, y\}$.

We now prove the characterisation of novel extensions. Let $\mathcal{R}$ be a novel Y-extension with matrix representation $v^{(\cdot,\cdot)}$ satisfying parts (i)–(iv) of the lemma. Fix arbitrary $t \neq \mathfrak{f}$. Then definition 3 implies the existence of $J \in \mathbb{J}$ and $L = J \times 0 \in \mathbb{J}^{\mathfrak{f}}$ such that $\mathcal{R}_{L+\delta_t} \neq \mathcal{R}_{L+\delta_{\mathfrak{f}}}$. W.l.o.g., consider the case where, for some $x, y \in Y$, it holds that both $y \mathcal{R}_{L+\delta_t} x$ and $x \mathcal{P}_{L+\delta_{\mathfrak{f}}} y$. Equivalently,

$$\langle v^{(x,y)}, L + \delta_t\rangle \leqslant 0 < \langle v^{(x,y)}, L + \delta_{\mathfrak{f}}\rangle$$

which, via linearity of $\langle v^{(x,y)}, \cdot\rangle$, we may rearrange to obtain

$$v^{(x,y)}(t) \leqslant -\langle v^{(x,y)}, L\rangle < v^{(x,y)}(\mathfrak{f}). \tag{17}$$

Thus, $v^{(x,y)}(t) \neq v^{(x,y)}(\mathfrak{f})$, as required for the lemma.

For the converse argument, fix arbitrary $t \in \mathbb{T}$. Then $\acute{v}^{(\cdot,\cdot)}(t) \neq \acute{v}^{(\cdot,\cdot)}(\mathfrak{f})$ implies the existence of distinct $x, y \in Y$ such that $\acute{v}^{(x,y)}(t) \neq \acute{v}^{(x,y)}(\mathfrak{f})$. We show that there exists $J \in \mathbb{J}$ and $L = J \times 0 \in \mathbb{J}^{\mathfrak{f}}$ satisfying eq. (17). For then, by retracing (in reverse order) the arguments that lead to eq. (17), we arrive at the conclusion that $\mathcal{R}_{L+\delta_t} \neq \mathcal{R}_{L+\delta_{\mathfrak{f}}}$, as required for $\mathcal{R}$ to be novel.

Take $\mu = v^{(x,y)}(t)$ and $\xi = v^{(x,y)}(\mathfrak{f})$ and consider the case where $\mu < \xi < 0$. Since $\mu < 0$, *2*-diversity implies that there exists $s \in \mathbb{T}$ such that $v^{(x,y)}(s)$ is positive. Then, for some $\lambda \in \mathbb{Q}_{++}$, $-\lambda v^{(x,y)}(s) \in (\mu, \xi)$. Let $L = \lambda\delta^{\mathfrak{f}}_s$ and observe that

$$\mu < -\langle v^{(x,y)}, L\rangle < \xi,$$

as required. *Mutatis mutandis*, the case where both μ and ξ are positive is the same. If $\mu \leqslant 0 \leqslant \xi$, then take $L = 0$, so that $\mu \neq \xi$ yields $\mathcal{R}_{L+\delta^{\mathfrak{f}}_t} \neq \mathcal{R}_{L+\delta^{\mathfrak{f}}_{\mathfrak{f}}}$. □

Proof of Proposition 2. Fix $\sharp Y = 3$ or 4, via lemma 2.1, let $v^{(\cdot,\cdot)}$ denote the 2-diverse matrix representation of the improper Y-extension $\mathcal{R}$. Let $\mathcal{H}_{++}$ denote the associated arrangement of hyperplanes. For every distinct $x, y \in X$, lemma 2.1 implies that $H_{++}^{\{x,y\}}$ intersects $\mathbb{R}_{++}^{\mathbb{T}}$. Then, similar to example 6, the $1 \leqslant n \leqslant \binom{\sharp Y}{2}$ distinct hyperplanes of $\mathcal{H}_{++}$ cut $\mathbb{R}_{++}^{\mathbb{T}}$ into at least $n+1$ regions. At least one pair G and G^* in $\mathcal{G}_{++}$ are therefore separated by all n distinct members of $\mathcal{H}_{++}$. Take $J \in G$, so that, for every distinct $x, y \in Y$, $\langle u^{(x,y)}, J\rangle \neq 0$. Thus $\mathcal{R}_J$ is antisymmetric, complete and, via A0$^\flat$, total. Next, take $L \in G^*$, so that since J and L are separated by every hyperplane in $\mathcal{H}_{++}$, $\mathcal{R}_J = \mathcal{R}_L^{-1}$. □

Proof of Lemma 2.2. Let A4$^\flat$ hold on Y. Since $v^{(\cdot,\cdot)}$ is a 2-diverse pairwise representation, $v^{(x,z)}, v^{(y,z)} \nleqslant 0$. By A4$^\flat$, one of $G_+^{(x,z)}$ and $G_+^{(z,x)}$ contains both J, L such that

$$\langle v^{(y,z)}, L\rangle < 0 < \langle v^{(y,z)}, J\rangle.$$

W.l.o.g., suppose J, L belongs to $G_+^{(x,z)}$. Then $\langle v^{(x,z)}, \cdot\rangle$ is positive on $\{L, J\}$, so that, for every $\lambda \in \mathbb{R}$ $v^{(x,z)} \neq \lambda v^{(y,z)}$, as required.

Conversely, let $x, y, z \in Y$ be such that $v^{(x,z)}$ and $v^{(y,z)}$ are noncollinear. Then $H_{++}^{\{x,z\}} \neq H_{++}^{\{y,z\}}$, and there exists $L \in H_{++}^{\{x,z\}} - H_{++}^{\{y,z\}}$. W.l.o.g., therefore, suppose $L \in H_{++}^{\{x,z\}} \cap G_{++}^{(y,z)}$. Since $v^{(\cdot,\cdot)}$ is 2-diverse, there exists $s, t \in \mathbb{T}$ such that $v^{(x,z)}(s) < 0 < v^{(x,z)}(t)$. Noting that $L \in \mathbb{R}_{++}^{\mathfrak{T}}$, so that $v^{(x,z)} \neq \delta_s, \delta_t$, let ψ_s and ψ_t be the convex paths from L to δ_s and δ_t respectively. For sufficiently small $\lambda > 0$, $\langle v^{(y,z)}, \psi_{s'}(\lambda)\rangle$ remains positive for $s' = s, t$ and, moreover, since $L \in H_{++}^{\{x,z\}}$,

$$\langle v^{(x,z)}, \psi_s(\lambda)\rangle < 0 < \langle v^{(x,z)}, \psi_t(\lambda)\rangle.$$

Finally, since L has finite support, a finite sequence of perturbations of the elements of $\psi_s(\lambda)$ and $\psi_t(\lambda)$ yields (rational-valued) members of $\mathfrak{J}$ with the same properties, as required for A4$^\flat$. □

Proof of Proposition 3. When $X = 2$, A4$^\flat$ and A4$'^\flat$ are identical to *2*-diversity. Let $Y = \{x, y, z\} \subseteq X$ and let $\mathcal{R}$ denote the improper Y-extension of $\preceq_{\mathbb{J}}$. We begin by assuming A4$^\flat$ and showing that $\sharp Y + 1 = 4 \leqslant \sharp\mathrm{total}(\mathcal{R})$. Via lemma 2.2, there are three distinct hyperplanes $H_{++}^{\{x,y\}}, H_{++}^{\{y,z\}}$ and $H_{++}^{\{x,z\}}$ in the associated arrangement $\mathcal{H}_{++}$. Then, as in example 6, $\varnothing = \mathbb{R}_{++}^{\mathbb{T}}$ is the unique element of $\mathcal{L}_{++}$ that lies below each member of $\mathcal{H}_{++}$. Thus, via eq. (5), $\boldsymbol{\mu}(A^{\varnothing}) = 1$ and $\boldsymbol{\mu}(A) = -\boldsymbol{\mu}(A^{\varnothing})$ for all three hyperplanes $A \in \mathcal{H}_{++}$. Thus, Zaslavski's theorem implies that $\sharp\mathcal{G}_{++}$ is bounded below by 4. Thus total$(\mathcal{R}) \geqslant 4$, and since, for every Y-extension $\acute{\mathcal{R}}$, $\sharp$total$(\acute{\mathcal{R}}) \geqslant \sharp$total$(\mathcal{R})$, A4$'^{\flat}$ holds.

Conversely, suppose A4$'^{\flat}$ holds and, once again let $\mathcal{R}$ denote the improper Y-extension of $\preceq_{\mathbb{J}}$, so that total$(\mathcal{R}) \geqslant \sharp Y = 3$. Now A4$'^{\flat}$ implies *2*-diversity, so that, via lemma 2.1, there exists a 2-diverse matrix representation with associated arrangement $\mathcal{H}_{++}$. It is not the case that $|\mathcal{H}_{++}| = 1$, for this would imply that $\sharp$total$(\mathcal{R}) = 2$. W.l.o.g., suppose $H^{\{x,y\}}_{++} \neq H^{\{y,z\}}_{++}$. Observe that A0$^{\flat}$ then implies $H^{\{x,z\}}_{++} \neq H^{\{x,y\}}_{++}$ and $H^{\{x,z\}}_{++} \neq H^{\{y,z\}}_{++}$. This implies that $v^{(x,y)}, v^{(y,z)}$ and $v^{(x,z)}$ are pairwise noncollinear. Finally, an application of lemma 2.2 then yields A4$^{\flat}$. □

Proof of Lemma 2.4. Note that, when $1 \leqslant |X| \leqslant 2$, *4*-Jac holds vacuously and $\preceq_{\mathbb{J}}$ has a Jacobi representation via lemma 2.1. For general X, the fact that *4*-Jac is necessary for $\preceq_{\mathbb{J}}$ to have a Jacobi representation follows simply because if the Jacobi identity holds on X, then it holds on every $Y \subseteq X$. For the sufficiency of *4*-Jac, we proceed by induction. As in lemma 3 and claim 9 of [GS], we assume that X is well-ordered.

In the case that $|X| \leqslant 4$, we only need to show that $v^{(\cdot,\cdot)}$ is unique. W.l.o.g., we take the initial step in our induction argument to satisfy $|X| = 4$. Let $\mathbf{v}^{(\cdot,\cdot)}$ denote any other Jacobi representation of $\preceq_{\mathbb{J}}$. By lemma 2.1, for every distinct $x, y \in Y^2$, there exists $\lambda^{\{x,y\}} > 0$ such that $\mathbf{v}^{(x,y)} = \lambda^{\{x,y\}} v^{(x,y)}$. We need to show that $\lambda^{\{x,y\}} = \lambda$ for every distinct $x, y \in Y$. Let $Y = \{x, y, z, w\}$. By lemma 2.2, the set $\{v^{(x,y)}, v^{(x,z)}, v^{(x,w)}\}$ is pairwise noncollinear. Then, since the Jacobi identity holds for both $v^{(\cdot,\cdot)}$ and $\mathbf{v}^{(\cdot,\cdot)}$, we derive the equation

$$(1 - \lambda^{\{x,y\}})v^{(x,y)} + (1 - \lambda^{\{y,z\}})v^{(y,z)} = (1 - \lambda^{\{x,z\}})v^{(x,z)} \tag{18}$$

Suppose that $1 - \lambda^{\{y,z\}} = 0$. Then, either the other coefficients in eq. (18) are both equal to zero (and our proof is complete), or we obtain a contradiction of lemma 2.2. Thus, $1 - \lambda^{\{y,z\}}$ is nonzero and we may divide through by this term and solve for $v^{(y,z)}$. First note that, since $v^{(\cdot,\cdot)}$ is a Jacobi representation, $v^{(y,x)} + v^{(x,y)} = v^{(y,y)} = 0$. Then, since $v^{(y,x)} = -v^{(x,y)}$,

$$v^{(y,z)} = \frac{1 - \lambda^{xy}}{1 - \lambda^{yz}} v^{(y,x)} + \frac{1 - \lambda^{xy}}{1 - \lambda^{yz}} v^{(x,z)}.$$

We then conclude that both of the coefficients in the latter equation are equal to one. (This follows from linear independence of $v^{(y,x)}$ and $v^{(x,z)}$ together with the

Jacobi identity $v^{(y,z)} = v^{(y,x)} + v^{(x,z)}$.) Thus, $\lambda^{\{x,y\}} = \lambda^{\{y,z\}} = \lambda^{\{x,z\}}$, as required. Repeated application of the same argument to the remaining Jacobi identities yields the desired conclusion, $\mathbf{v}^{(\cdot,\cdot)} = \lambda v^{(\cdot,\cdot)}$.

For the inductive step, take Y to be an initial segment of X. By the induction hypothesis, there exists a Jacobi representation $\mathbf{v}^{(\cdot,\cdot)} : Y^2 \times \mathbb{T} \to \mathbb{R}$ of the improper Y-extension $\mathcal{R} = \preceq_{\mathbb{J}} \cap Y^2$ that is suitably unique.

Claim 3.5.1. *For every $w \in X - Y$ and $W \stackrel{\text{def}}{=} Y \cup \{w\}$, there exists a Jacobi representation $\hat{v}^{(\cdot,\cdot)} : W^2 \times \mathbb{T} \to \mathbb{R}$ of the improper W-extension $\hat{\mathcal{R}}$.*

Proof of claim 3.5.1. Via lemma 2.2, there exists a conditionally 2-diverse pairwise representation $u^{(\cdot,\cdot)}$ of $\preceq_{\mathbb{J}}$. Fix any four distinct elements x, x', y, z in Y. Lemma 2.1 implies the existence of $\phi, \phi' \in \mathbb{R}_{++}$ such that $\phi u^{(x,z)} = \mathbf{v}^{(x,z)}$ and $\phi' u^{(x',z)} = \mathbf{v}^{(x',z)}$. Let $Z = \{x, y, z, w\}$ and $Z' = \{x', y, z, w\}$. Since *3*-Jac holds, there exist positive scalars $\alpha, \beta, \hat{\beta}, \gamma, \sigma$ and τ such that

$$\alpha u^{(x,w)} + \beta u^{(w,y)} = \gamma u^{(x,y)}, \tag{19}$$

$$\hat{\beta} u^{(y,w)} + \sigma u^{(w,z)} = \tau u^{(y,z)}, \text{ and} \tag{20}$$

$$\gamma u^{(x,y)} + \tau u^{(y,z)} = \mathbf{v}^{(x,z)}. \tag{21}$$

Moreover, *4*-Jac ensures that we may take $\beta = \hat{\beta}$. Since $u^{(\cdot,\cdot)}$ is conditionally 2-diverse, $\{u^{(x,y)}, u^{(y,z)}\}$ is linearly independent, and the linear system eq. (21) in the unknowns γ and τ has a unique solution. This, together with the induction hypothesis (which yields $\mathbf{v}^{(x,y)} + \mathbf{v}^{(y,z)} = \mathbf{v}^{(x,z)}$) implies that $\gamma u^{(x,y)} = \mathbf{v}^{(x,y)}$ and $\tau u^{(y,z)} = \mathbf{v}^{(y,z)}$. Similarly, for Z', *4*-Jac yields $\alpha', \beta', \sigma', \gamma', \tau' > 0$ such that

$$\alpha' u^{(x',w)} + \beta' u^{(w,y)} = \gamma' u^{(x',y)}, \tag{22}$$

$$\beta' u^{(y,w)} + \sigma' u^{(w,z)} = \tau' u^{(y,z)}, \text{ and} \tag{23}$$

$$\gamma' u^{(x',y)} + \tau' u^{(y,z)} = \mathbf{v}^{(x',z)}. \tag{24}$$

As in the arguments involving γ and τ, the induction hypothesis yields $\gamma' u^{(x',y)} = \mathbf{v}^{(x',y)}$ and $\tau' u^{(y,z)} = \mathbf{v}^{(y,z)}$. We conclude that $\tau = \tau'$. Substituting for τ' in eq. (23) and appealing to linear independence of $\{u^{(y,w)}, u^{(w,z)}\}$ then yields the desired equalities $\beta = \beta'$ and $\sigma = \sigma'$.

As a consequence of the above argument, for every $y, z \in Y$, take $\hat{v}^{(y,w)}$ and $\hat{v}^{(w,z)}$ to be the unique vectors in $\mathbb{R}^{\mathbb{T}}$ that solve the equation $\hat{v}^{(y,w)} + \hat{v}^{(w,z)} = \mathbf{v}^{(y,z)}$.

For every $y, z \in Y$, let $\hat{v}^{(y,z)} = \mathbf{v}^{(y,z)}$ and $\hat{v}^{(w,w)} = 0$. Then the matrix $\hat{v}^{(\cdot,\cdot)}$ with row vectors $\left\{\hat{v}^{(x,y)} : x, y \in W\right\}$ is a Jacobi representation of $\hat{\mathcal{R}}$. $\square$

Our proof of claim 3.5.1 shows that the extension to W holds for any initial subsegment of Y consisting of four elements. Our proof thereby accounts for the case where X is infinite and w is a limit ordinal. $\square$