The Relativity of Causal Knowledge

Gabriele D'Acunto[1,2] **Claudio Battiloro**[3]

[1]Information Engineering, Electronics and Telecommunications Dept., Sapienza University, Rome, Italy
[2]National Inter-University Consortium for Telecommunications (CNIT), Parma, Italy
[3]Biostatistics Dept., Harvard University, Cambridge, MA, USA

Abstract

Recent advances in *artificial intelligence* reveal the limits of purely predictive systems and call for a shift toward causal *and* collaborative reasoning. Drawing inspiration from the revolution of Grothendieck in mathematics, we introduce the *relativity of causal knowledge*, which posits structural causal models (SCMs) are inherently imperfect, subjective representations embedded within networks of relationships. By leveraging category theory, we arrange SCMs into a functor category and show that their observational and interventional probability measures naturally form convex structures. This result allows us to encode non-intervened SCMs with convex spaces of probability measures. Next, using sheaf theory, we construct the *network sheaf and cosheaf of causal knowledge*. These structures enable the transfer of causal knowledge across the network while incorporating interventional consistency and the perspective of the subjects, ultimately leading to the formal, mathematical definition of *relative causal knowledge*.

1 INTRODUCTION

"The Book of Why" [23] elects *causality* as the key to overcoming the limit of purely predictive artificial intelligence (AI). This argument recently found mathematical support in the work of Richens and Everitt [27], who gave evidence that robustness to distributional shifts of AI agents is conditioned on the learning of an approximate subjective causal model. The *structural causal model* (SCM) framework [21] is a gold standard for modeling *cause-effect* relationships in complex systems. Informally, a (probabilistic) SCM is made by causal variables and variable-specific sources of noise, together with structural equations – to be read as assignments, like in physics – determining how each variable is causally influenced by others (e.g., the test score of a student depends on hours of study, motivation, and some randomness). An SCM induces the so-called "ladder of causations" [23]: the *associational* layer related to factual information (*seeing*), the *interventional* layer related to the effects of actions (*doing*), and the *counterfactual* layer related to imagine the effect of an action, given that something else occurred (*retrospection*). The ultimate goal of *causal AI* is to empower AI systems with such a ladder of causation for robust and trustworthy decision-making. This work pushes in the same direction.

Motivation. We are driven by a philosophically simple yet technically unexplored concept: *any causal model is an imperfect and subjective representation of the world, and it cannot be severed from the network of relations the subject is immersed in*. As such, our work somehow starts where Richens and Everitt [27] end, as they provided a (specific) formalization of the subjectiveness of causal models without, however, formalizing their dependency on the subject's relationships. In philosophy, this concept has been extensively debated: manipulability theories view causes as "handles" for effecting change, tying them to subjective agency [35]; pluralist approaches suggest multiple, context-dependent concepts of causation [26]; the actor-network theory of Latour and others posits that everything in the social and natural worlds exists in constantly shifting networks of relationships, and nothing exists outside those relationships [17]. A motivating example for our work is *agentic AI*, a frontier paradigm pushing for autonomous AI agents – the subjects – to collaborate in solving complex tasks. However, in our work, the notion of "subject" takes on a broader meaning and could represent, for example, the resolution by which a phenomenon is studied; the sensor that detects pollutants in a specific geographical area, different from that of other sensors; the trading book of a bank seen as a proxy for investment strategy. In all of the above cases, we advocate that each subject likely – and, arguably,

Authors' contributions breakdown in back matter. Correspondence to: <gabriele.dacunto@uniroma1.it> and <cbattiloro@hsph.harvard.edu>.

hopefully – develops a subjective SCM, these SCMs are interconnected, and their interplay can benefit the subject and the entire network. Inherently, this paradigm, which we term *relativity of causal knowledge* and whose core object is *relative causal knowledge* (RCK), is tied to the concept of *perspective*: subjects cannot interact by detaching from their world representation. In our framework, asking for a unique "true" causal description of a system is an ill-posed question, as the very same notion of causality is inherently relative. However, it is important to highlight that the relativity of causal knowledge is different from the *relativism* (in its philosophical meaning) of causal knowledge: *we do not undermine the meaning of things, we question its description as a monolithic object*.

Related Works. RCK has its roots in *category theory* [18] and *sheaf theory* [31], and is inspired by Grothendieck's notion of relativism in the context of mathematics [20]. Grothendieck revolutionized the understanding of mathematical structures by shifting the focus from individual objects to the relationships between them, as expressed through morphisms. This perspective led to a flexible, contextual view of mathematical spaces. We draw a parallel to Grothendieck's paradigm shift by treating SCMs as inherently contextual and interconnected rather than isolated entities. Just as Grothendieck's use of sheaf theory facilitates local-to-global transitions in mathematics, we employ network sheaves–first-order cellular sheaves [5]– to investigate how causal knowledge, i.e., a set of interventional and observational probability measures, is transferred and transformed across different subjects and their perspectives. Previous work focused on a functorial characterization of SCM through mappings between the causal structure and the (discrete) distributions of causal variables [15]. Conversely, we provide a category-theoretic treatment of SCMs – and related interventions – focusing on morphisms between probability spaces, linking the SCM to the category of convex spaces of probability measures [9]. Similar definitions to Definitions 1 and 2 appear in the concurrent work [7].

Relations among SCMs have been investigated over time. The *transportability* problem [22, 2] addresses the transfer of causal knowledge from one environment (the source) to another (the target) under assumptions on *(i)* the knowledge of the underlying causal structure, typically represented by causal graphs, and *(ii)* the types and targets of intervention. In *causal transfer learning* [37, 29, 19], several studies tackle the challenge of transferring causal knowledge from source domains to enhance performance in a target domain (i.e., domain adaptation). Next, *equivalence* of SCMs [3, 33] aims at identifying equivalent (sub)structures to provide insights into how different systems share common causal relationships. Particularly relevant to our work is the theory of *causal abstraction* (CA) [30, 4]. CA formalizes the mappings between SCMs describing the same system at different levels of granularity. In this paper, we will work under the α-abstraction framework proposed by Rischel [28]. The latter is convenient to us since *interventional consistency* (IC) is neatly separated from the definition of the CA.

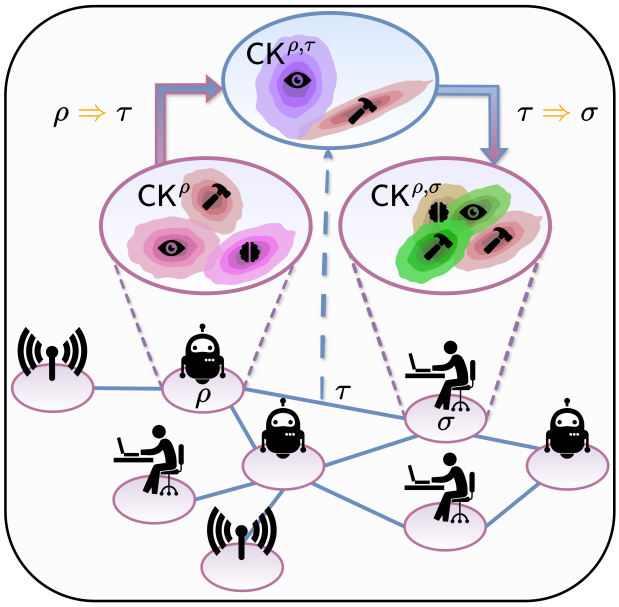


Figure 1: The *relativity of causal knowledge* states that causal knowledge (CK) is subjective and interconnected rather than objective and isolated. Multiple subjects of/in the same system will develop multiple and different instances of CK describing the system. Informally, CK can be seen as a set of probability measures corresponding to 👁 seeing, 🔨 doing, and imagining 🌓. The CK CK^{ρ} of subject ρ is fully accessible exclusively by ρ. As such, another subject σ can only access the *relative causal knowledge* (RCK) $\mathsf{CK}^{\rho,\sigma}$, i.e., the CK of ρ from the perspective of σ. There is a link τ between ρ and σ, i.e., they can communicate, if their CK admits a common interventionally consistent CA acting as *backbone space*. As such, the RCK $\mathsf{CK}^{\rho,\sigma}$ is obtained by first transporting CK^{ρ} on τ, obtaining a more abstract $\mathsf{CK}^{\rho,\tau}$, and then $\mathsf{CK}^{\rho,\tau}$ on σ. Interestingly, subjects that are not directly connected can still access some RCK if there is a path of links among them, but it would be first "filtered" by the perspective of all the other subjects on the path. We elegantly implement RCK using a category-theoretic approach resulting in novel mathematical objects: *The network sheaf and cosheaf of causal knowledge.*

Contributions. A number of results are established to pose the formal definition of RCK. *First*, we introduce the category of SCMs, viz. $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, whose objects are SCMs – expressed as functors – and morphisms are natural transformations. We further characterize hard [21] and soft [8] interventions in the latter category, proving that the set of entailed observational and interventional probability measures is closed under a convex combination operation (cf. Theorem 1). *Second*, aiming at IC, we recast the α-abstraction in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, proving that IC CA morphism corresponding

to endogenous variables is valid in the category of convex spaces of probability measures, viz. CSprob (cf. Theorem 2). Hence, we establish the existence of a functor encoding non-intervened SCMs into objects in CSprob, representing the SCMs' causal knowledge. *Third*, leveraging the encoding functor, we define network sheaf and cosheaf of causal knowledge, finally posing the formal definition of RCK.

Impact. Our proposed network sheaves and cosheaves are particular instances of more general network sheaves and cosheaves in CSprob. Our work ultimately emphasizes their relevance to applications in (but not limited to) AI/ML. In particular, RCK and network sheaf and cosheaf of causal knowledge represent new objects to be investigated, adding to the cellular sheaves valued in Abelian categories, such as those on vector spaces, and Hilbert spaces [5, 12], and the non-Abelian case of sheaves of lattices [10]. Overall, our work opens three major areas of research: *cohomology theory*, *Hodge(-like) theory*, and *learning theory* for network sheaves and cosheaves of causal knowledge.

Notation. Sets and collections are uppercase calligraphic, $\mathcal{A}$. The set of integers from 1 to n is $[n]$. Given a scalar a, we denote by $\bar{a} = 1 - a$. A set $\mathcal{U}$ equipped with a σ-algebra Υ gives a measurable space $(\mathcal{U}, \Upsilon)$. A measurable space $(\mathcal{U}, \Upsilon)$ together with a probability measure μ, i.e., such that $\mu(\mathcal{U}) = 1$, gives a probability space $(\mathcal{U}, \Upsilon, \mu)$. Given two measurable spaces $(\mathcal{U}, \Upsilon)$ and $(\mathcal{V}, \Omega)$, a measurable map $\varphi : \mathcal{U} \to \mathcal{V}$, and a probability measure μ over $(\mathcal{U}, \Upsilon)$, we denote by $\varphi(\mu) \coloneqq \mu \circ \varphi^{-1}$ the pushforward measure over $(\mathcal{V}, \Omega)$. The domain of a function is $\mathcal{D}[]$. The indicator function is $I_{\mathcal{A}}(A)$, 1 if $A \in \mathcal{A}$, 0 otherwise.

Remark 1 *A less technical but comprehensive description together with practical examples can be found in Appendix E. This description is designed to be useful either before or after the reader has parsed the main body of the paper.*

2 CATEGORICAL SCM AND CAUSAL KNOWLEDGE

Category theory [18] is a branch of pure mathematics that studies abstract structures and their relationships through objects and morphisms, focusing on how they compose and interact. A *category* C is composed of objects having a certain structure (e.g., measurable spaces, vector spaces) and arrows (morphisms) between them preserving the structure (e.g., measurable maps, linear maps) and satisfying certain axioms (cf. Definition 12). Objects and morphisms form the collections $\mathcal{C}_0$ and $\mathcal{C}_1$, respectively. Given C, the opposite category C^{op} has the same objects as C and the same, but reversed, arrows. Taking the subcollections $\mathcal{S}_0 \subseteq \mathcal{C}_0$ and $\mathcal{S}_1 \subseteq \mathcal{C}_1$ we can form a *subcategory* of C. Given two categories C and D, we can have an arrow between them, namely a *functor* $F : \mathsf{C} \to \mathsf{D}$. Functors define mappings between categories in a consistent way (cf. Definition 14). Notably, functors cannot destroy relations existing in C when mapping to D. Given two functors F and G from C to D, we can have an arrow $\eta : F \to G$ between them called *natural transformation*, which is a peculiar arrow as it is invariant w.r.t. morphisms between the mapped objects in target category D (cf. Definition 15). Functors can be arranged in categories whose objects are functors and morphisms are natural transformations. More details in Appendix A.

We look at the probabilistic SCM over continuous random variables, denoted by M, as a subjective and imperfect world representation. We consider the *Markovian setting* [21], thus M entails a *directed acyclic graph* (DAG), viz., G^{M}. The nodes of G^{M} correspond to the *endogenous* variables $\mathcal{X}$, that is, the causal variables on which we can intervene on. The edges of G^{M} are determined by structural functions $\mathcal{F} \coloneqq \{f_1, \ldots, f_n\}$ determining the value of each causal variable X_i, $i \in [n]$, in terms of a set of *parents*, viz. $\mathcal{P}_i \subseteq \mathcal{X} \setminus \{X_i\}$, and a node-specific *exogenous* variable, $Z_i \in \mathcal{Z}$. Denote by $\mathcal{Z}^{\mathcal{A}_i} \subseteq \mathcal{Z} \setminus \{Z_i\}$ the set of exogenous variables corresponding to the ancestors of X_i, where $\mathcal{A}_i \subseteq [n] \setminus \{i\}$. $\mathcal{F}$ induces a set of mixing functions $\mathcal{M} \coloneqq \{m_1, \ldots, m_n\}$, such that the values of the endogenous random variables are equivalently expressed as $x_i = m_i\left(\{z_j\}_{j \in \mathcal{A}_i}, z_i\right), \forall\, i \in [n]$.

Accordingly, we can define M as a triple $\langle (\mathcal{U}, \Upsilon, \zeta), (\mathcal{V}, \Omega, \chi), \mathcal{M} \rangle$ made of the probability space of the exogenous $(\mathcal{U}, \Upsilon, \zeta)$, the probability space of the endogenous $(\mathcal{V}, \Omega, \chi)$, and measurable mixing functions $\mathcal{M}$.

Definition 1 (Measure-theoretic SCM) *A Markovian SCM* M *is a triple* $\langle (\mathcal{U}, \Upsilon, \zeta), (\mathcal{V}, \Omega, \chi), \mathcal{M} \rangle$ *where:*

- $(\mathcal{U}, \Upsilon, \zeta)$ *is a probability space associated with exogenous variables* $\mathcal{Z}$*. Specifically, it consists of the product probability measure* $\zeta = \zeta_1 \times \ldots \times \zeta_n$ *on the product measurable space* $(\mathcal{U}, \Upsilon) = (\mathcal{U}_1 \times \ldots \times \mathcal{U}_n, \Upsilon_1 \otimes \ldots \otimes \Upsilon_n)$*. The probability measure is such that, for each* $\mathcal{W}_1 \in \Upsilon_1, \ldots, \mathcal{W}_n \in \Upsilon_n$*, we have*

$$\zeta_1 \times \ldots \times \zeta_n(\mathcal{W}_1 \times \ldots \times \mathcal{W}_n) = \zeta_1(\mathcal{W}_1) \times \ldots \times \zeta_n(\mathcal{W}_n)\,; \tag{1}$$

- $(\mathcal{V}, \Omega, \chi)$ *is a probability space associated with endogenous variables* $\mathcal{X}$ *consisting of a joint probability measure* χ *on the product measurable space* $(\mathcal{V}, \Omega) = (\mathcal{V}_1 \times \ldots \times \mathcal{V}_n, \Omega_1 \otimes \ldots \otimes \Omega_n)$*;*
- $\mathcal{M} \coloneqq \{m_1, \ldots, m_n\}$ *is a collection of measurable mixing functions allowing us to recursively rewrite the causal assignments only in terms of the exogenous variables. Accordingly, the joint probability measure* χ *factorizes as*

$$\chi = \bigtimes_{i=1}^{n} m_i\left(\mu_i\left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i}\right)\right)\,; \tag{2}$$

where (i) $\mathcal{U}^{\mathcal{A}_i} = \bigtimes_{j \in \mathcal{A}_i} \mathcal{U}_j$*; (ii)* μ_i *is a probability measure on the product measurable space* $\left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i}, \Upsilon_i \otimes \Upsilon^{\mathcal{A}_i}\right)$*, with* $\Upsilon^{\mathcal{A}_i} = \bigotimes_{j \in \mathcal{A}_i} \Upsilon_j$*.*

Appendix B provides an example for the case of linear SCM with additive noise. An important property of $\mathcal{M}$ that holds for the main classes of Markovian SCMs, such as additive noise and post-nonlinear models [25, Chapter 4], is invertibility (cf. Appendix C). We will leverage such a property to prove that the proposed $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ preserves expressiveness. At this point, we can leverage Definition 1 to give a functorial definition of M. Consider the following categories, I and Prob, respectively. The former is a small category made only of two objects, a source node and a target node, and a unique arrow from the source to the target node. Specifically, I has shape $\bullet \to \bullet$. The latter category instead has as objects probability spaces and as morphisms measurable maps. The functorial representation follows by viewing M as an arrow between I and Prob, assigning *(i)* to the source node in I the probability space associated with exogenous variables; *(ii)* to the target node the probability space associated with endogenous variables; and *(iii)* to the only arrow in I the collection of measurable maps $\mathcal{M}$.

Definition 2 (Category-theoretic SCM) *An SCM is a functor* $\mathsf{M} : \mathsf{I} \to \mathsf{Prob}$ *defined as follows*

$$\begin{tikzcd} \mathsf{I} \arrow[r, "\mathsf{M}"] & \mathsf{Prob} \\ I \bullet \arrow[d] & (\mathcal{U}, \Upsilon, \zeta) \arrow[d, "\mathcal{M}"] \\ I' \bullet & (\mathcal{V}, \Omega, \chi) \end{tikzcd} \tag{3}$$

Please, refer to Appendix B for an example in the case of linear SCM with additive noise. At this point, we can define a category of SCMs whose objects are functors as in Definition 2, and whose morphisms are natural transformation between the functors.

Definition 3 *The category of SCMs, namely* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$*, consists of (i) functors* $\mathsf{M} : \mathsf{I} \to \mathsf{Prob}$ *as objects; and (ii) natural transformations* $\eta : \mathsf{M} \to \mathsf{M}'$ *as morphisms, such that:*

- *for each* I *in* I*, a measurable map* $\eta_I : \mathsf{M}(I) \to \mathsf{M}'(I)$ *in* Prob*, called component at* I*;*
- *for the unique morphism* $f : I \to I'$ *in* I*, the following commutes:*

$$\begin{tikzcd} \mathsf{M}(I) \arrow[r, "\mathsf{M}^f = \mathcal{M}"] \arrow[d, "\eta_I"'] & \mathsf{M}(I') \arrow[d, "\eta_{I'}"] \\ \mathsf{M}'(I) \arrow[r, "\mathsf{M}'^f = \mathcal{M}'"'] & \mathsf{M}'(I') \end{tikzcd} \tag{4}$$

Exogenous *Endogenous*

In Equation (4), we added dashed rectangles to remark that the functor images *(i)* on the left correspond to the probability spaces of the exogenous variables of M and M'; *(ii)* on the right, to probability spaces of the endogenous variables. The knowledge of an SCM allows us to run *interventions* to act or to simulate "new worlds", that is, obtain new post-interventional distributions. This enables the second and third layers of the ladder of causation [23] described in Section 1. In our work, we consider *hard* and *soft* interventions. A hard intervention $\mathrm{do}\left(\{X_i = x_i\}_{X_i \in \widetilde{\mathcal{X}}}\right)$, where $\widetilde{\mathcal{X}} \subseteq \mathcal{X}$, replaces each assignment function f_i corresponding to $X_i \in \widetilde{\mathcal{X}}$ with the constant x_i, thus generating a new post-interventional SCM, viz. $\mathsf{M}_{\mathcal{H}}$. We associate a hard intervention with a collection $\mathcal{F}_{\mathcal{H}}$ such that $f_i^{\mathcal{H}} = x_i$, for all $X_i \in \widetilde{\mathcal{X}}$; and $f_j^{\mathcal{H}} = f_j$, for all $X_j \in \mathcal{X} \setminus \widetilde{\mathcal{X}}$. Graphically, a hard intervention mutilates G^{M} by removing the incoming edges of the variables in $\widetilde{\mathcal{X}}$. Consequently, according to the truncated product formula [21], an intervention entails a post-interventional distribution $P(\mathcal{X} \mid \mathcal{H}) = \prod_{j \in \mathcal{X} \setminus \widetilde{\mathcal{X}}} P(X_j \mid \mathcal{P}_i, Z_i)$ (evaluated at $\{X_i = x_i\}$).

In $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, the post-interventional SCM is the functor mapping *(i)* the source node in I to $(\mathcal{U}, \Upsilon, \zeta)$, as done by M; *(ii)* the target node to $(\mathcal{V}, \Omega, \chi)_{\mathcal{H}}$, where χ_j degenerates to a point measure $I_{x_j}(X_j)$; *(iii)* and the unique arrow in I to the measurable maps $\mathcal{M}_{\mathcal{H}} := \{m_1^{\mathcal{H}}, \ldots, m_n^{\mathcal{H}}\}$. The components of $\mathcal{M}_{\mathcal{H}}$ are *(i)* constant functions $m_i^{\mathcal{H}} = x_i$ for $X_i \in \mathcal{X}$, and *(ii)* $m_j^{\mathcal{H}} = m_j^{\mathcal{H}}(Z_j \cup \mathcal{Z}^{\widetilde{\mathcal{A}}_j})$ where $\mathcal{Z}^{\widetilde{\mathcal{A}}_j}$ is the set of exogenous corresponding to the ancestors of X_j that are not screened by the intervention (cf. Appendix B). At this point, in the same spirit of the truncation formula above, the post-interventional probability measure reads as

$$\chi_{\mathcal{H}} = \bigtimes_{X_j \in \mathcal{X} \setminus \widetilde{\mathcal{X}}} m_j^{\mathcal{H}} \left(\mu_j \left(\mathcal{U}_j \times \mathcal{U}^{\widetilde{\mathcal{A}}_j}\right)\right) . \tag{5}$$

Running hard interventions on SCMs is not always possible, and in certain cases, this way of intervening is unethical [8]. Therefore, a more general notion of intervention has been considered over the past. Indeed, the mutilation of the DAG is not the only possible informative intervention. Rather, we can be interested in simply modifying the causal mechanisms, without removing any incoming causal relations. Such a family of interventions is called soft.

The soft intervention generates a post-interventional $\mathsf{M}_{\mathcal{S}}$ by substituting $\mathcal{F}$ with $\mathcal{F}_{\mathcal{S}}$, where each f_i associated with the intervened variables $X_i \in \widetilde{X}$ is replaced by another function $\widetilde{f}_i$, and the rest is unchanged. Hence, the soft intervention can change the functional form of an SCM. In principle, soft interventions can potentially add new causal relations. However, in our work, we consider only soft interventions that do not alter the parent set of the endogenous variables. At this point, $\mathcal{F}_{\mathcal{S}}$ induces a new collection of mixing functions $\mathcal{M}_{\mathcal{S}} := \{m_1^{\mathcal{S}}, \ldots, m_n^{\mathcal{S}}\}$ (cf. Appendix B), such that the post-interventional probability measure reads as

$$\chi_{\mathcal{S}} = \bigtimes_{X_i \in \mathcal{X}} m_i^{\mathcal{S}} \left(\mu_i \left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i}\right)\right) . \tag{6}$$

The resulting $\mathsf{M}_{\mathcal{S}}$ is an object in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ mapping *(i)* the source node in I to $(\mathcal{U}, \Upsilon, \zeta)$, as done by M; *(ii)* the

target node to $(\mathcal{V},\, \Omega, \chi)_{\mathcal{S}}$, where $\chi_{\mathcal{S}}$ follows Equation (6); and *(iii)* the unique arrow in I to the measurable map $\mathcal{M}_{\mathcal{S}}$. An important point we must ensure is that the proposed category-theoretic SCM formulation preserves the expressiveness: given either a hard or soft intervention $\mathcal{I}$, we have a suitable morphism in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ from M to $\mathsf{M}_{\mathcal{I}}$. This is ensured by the following.

Lemma 1 *[Interventions in* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ *] Given (i)* M *as in Definition 2; and (ii) the collection of measurable maps* $\mathcal{F}_{\mathcal{I}} \coloneqq \{f_1^i, \ldots, f_n^i\}$*; an intervention in* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ *is a natural transformation* $\eta^{\mathcal{I}} \coloneqq \langle \mathrm{Id}_{\mathsf{M}(I)}, \mathcal{I}\rangle$*, with* $\mathcal{I} = \mathcal{M}_{\mathcal{I}} \circ \mathcal{M}^{-1}$ *such that the following holds*

$$\begin{tikzcd} \mathsf{M}(I) \arrow[r, "\mathcal{M}"] \arrow[d, "\mathrm{Id}_{\mathsf{M}(I)}"'] & \mathsf{M}(I') \arrow[d, "\mathcal{I}"] \\ \mathsf{M}_{\mathcal{I}}(I) \arrow[r, "\mathcal{M}_{\mathcal{I}}"'] & \mathsf{M}_{\mathcal{H}}(I') \end{tikzcd} \implies \mathcal{M}_{\mathcal{I}} = \mathcal{I} \circ \mathcal{M}\,. \quad (7)$$

Proof. See appendix D. □

The proof leverages the fact that the hard intervention acts only on the endogenous, thus the component $\eta_I^{\mathcal{I}}$ mapping $\mathsf{M}(I)$ to $\mathsf{M}_{\mathcal{I}}(I)$ is simply the identity $\mathrm{Id}_{\mathsf{M}(I)}$. Then, for the second component $\eta_{I'}^{\mathcal{I}}$, we leverage invertibility of $\mathcal{M}$ mentioned above. The commutation in Equation (7) follows by construction. Equation (7) highlights that our proposed $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, is as rich as the canonical SCM framework. Specifically, starting from a non-intervened SCM M, we can obtain all possible observational and interventional states of the causal model M through the application of all the possible hard and soft interventions since Lemma 1 guarantees the existence of corresponding measurable maps $\mathcal{I}$.

Remark 2 *By leveraging the invertibility of* $\mathcal{M}$*, Lemma 1 preserves expressiveness not only at the level of probability measures, but also with respect to the values of exogenous and endogenous variables. This is a stronger result, as interventional consistency concerns only the distributional level. In fact,* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ *retains the expressiveness of canonical SCMs also in terms of counterfactual consistency. If counterfactual consistency is relaxed, a more general result on the existence of the morphism* $\mathcal{I}$ *can be established via Brenier's polar factorization theorem [34, Chapter 3], without requiring the invertibility of* $\mathcal{M}$*, as shown in [7].*

Definition 4 (Causal knowledge) *The causal knowledge* $\mathsf{CK}(\mathsf{M})$ *entailed by* M *is the subcategory of* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ *whose objects are* M *together with its intervened states* $\{\mathsf{M}_{\mathcal{I}}\}$ *generated by the application of all the possible interventions* $\{\mathcal{I}\}$*, and whose morphisms are the natural transformations* $\eta^{\mathcal{I}}$ *between these objects in* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$.

Consequently, $\mathsf{CK}(\mathsf{M})$ corresponds to *(i)* a product distribution over the exogenous variables together with the identity morphism, representing the component $\eta_I^{\mathcal{I}}$ of the natural transformations; and *(ii)* the set of probability measures over the endogenous as given in Equations (2), (5) and (6) together with the morphisms $\mathcal{H}$ and $\mathcal{S}$, representing the natural transformation component $\eta_{I'}^{\mathcal{I}}$. Consider now two different non-intervened SCMs, namely M and M', and their corresponding $\mathsf{CK}(\mathsf{M})$ and $\mathsf{CK}(\mathsf{M}')$, respectively. Additionally, suppose that M and M' are not isolated entities; rather, they are immersed in a certain network where other non-intervened SCMs exist. We are interested in relating $\mathsf{CK}(\mathsf{M})$ and $\mathsf{CK}(\mathsf{M}')$ within the network in a way that is both *causally and category-theoretically* consistent. From now on, we will focus on the endogenous layer, since exogenous variables are latent. Hence, when we say $\mathsf{CK}(\mathsf{M})$, with slight abuse of notation we refer to the observational and interventional probability measures over the endogenous.

3 ENCODING CAUSAL KNOWLEDGE WITH CONVEX SPACES

Let us consider the measurable space $(\mathcal{V},\, \Omega)$ associated with the endogenous variables, where Ω is a σ-algebra over $\mathcal{V}$. The set of probability measures on $(\mathcal{V},\, \Omega)$, namely $\Delta_{(\mathcal{V},\, \Omega)}$ is a convex space, subset of the vector space $\mathbb{R}^{\Omega}$. Following [9], we define the latter convex space as follows.

Lemma 2 *[Convex space of probability measures,* $\langle \Delta_{(\mathcal{V},\, \Omega)}, cc_\lambda\rangle$*] The convex space of probability measures, namely* $\langle \Delta_{(\mathcal{V},\, \Omega)}, cc_\lambda\rangle$*, is given by the set* $\Delta_{(\mathcal{V},\, \Omega)}$ *of probability measures* χ *on* $(\mathcal{V},\, \Omega)$ *together with a convex combination operation defined by*

$$cc_\lambda(\chi_1, \chi_2)(\mathcal{O}) \coloneqq \lambda\chi_1(\mathcal{O}) + \bar{\lambda}\chi_2(\mathcal{O})\,, \quad (8)$$

for all $\mathcal{O} \in \Omega$*, with* $\lambda \in [0, 1]$.

Proof. See appendix D. □

Convex spaces $\langle \Delta_{(\mathcal{V},\, \Omega)}, cc_\lambda\rangle$ are the objects of the category of convex spaces of probability measures, namely CSprob, that we introduce below.

Lemma 3 *The category* CSprob *has convex spaces* $\langle \Delta_{(\mathcal{V},\, \Omega)}, cc_\lambda\rangle$ *as objects, and affine measurable maps – i.e., measurable maps commuting with* cc_λ *– as morphisms.*

Proof. See appendix D. □

It turns out that $\mathsf{CK}(\mathsf{M})$ is closed w.r.t. the convex combination operation in Lemma 2.

Theorem 1 *Every convex combination of probability measures corresponding to a causal knowledge* $\mathsf{CK}(\mathsf{M})$ *is a valid soft-interventional probability measure for* $\mathsf{CK}(\mathsf{M})$.

Proof. See appendix D. □

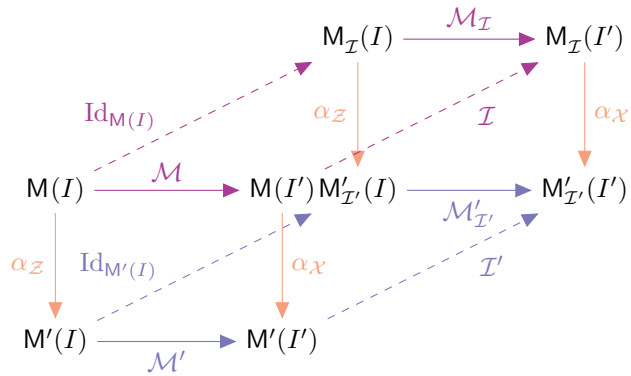
Hence, in light of Theorem 1, it is legitimate to question whether we can establish a functorial encoding of M in a convex space. If on objects, the encoding seems straightforward, on morphisms, we have to be careful. Indeed, given M and M', the natural transformation η in Definition 3 tells us nothing about the existence of natural transformations between the intervened states of M and M'. If the latter natural transformations do not exist in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, an encoding functor will fail in mapping η to an affine morphism between the convex spaces corresponding to M and M', since such a morphism does not exist in CSprob. Conversely, if the intervened states of M and M' are related by a natural transformation in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$, meaning that they are IC, the affine morphism will exist in CSprob.

The latter observation naturally links our work to the theory of *causal abstraction* (CA), electing CA as the necessary formalism for relating causal knowledge. Specifically, in the following, we build upon the $\boldsymbol{\alpha}$-abstraction introduced by [28]. Given a micro-level M and a macro-level M', an $\boldsymbol{\alpha}$-abstraction is a triple $\boldsymbol{\alpha} := \langle \mathcal{R}, a, \alpha \rangle$ where *(i)* $\mathcal{R}$ is a set of endogenous variables in M that are abstracted to the macro-level *(ii)* structurally via the surjective map $a : \mathcal{R} \to \mathcal{X}'$, and *(iii)* functionally by $\alpha : \mathcal{D}[\mathcal{R}] \to \mathcal{D}[\mathcal{X}']$. Here, for each $X'_i \in \mathcal{X}'$, we have a surjective function mapping the values of the micro-level to the macro-level, viz. $\alpha_{X'_i} : \mathcal{D}[a^{-1}(X'_i)] \to \mathcal{D}[X'_i]$. Given an $\boldsymbol{\alpha}$-abstraction between M and M', we say that $\boldsymbol{\alpha}$ is IC if, *for any intervention* $\mathcal{I}$ on the endogenous $\mathcal{X}'_{\mathcal{I}} \subseteq \mathcal{X}'$ and for any set of targets $\mathcal{T}' \subseteq \mathcal{X}' \setminus \mathcal{X}'_{\mathcal{I}}$, we can obtain the values in $\mathcal{D}[\mathcal{T}']$ starting from those in $\mathcal{D}[a^{-1}(\mathcal{X}'_{\mathcal{I}})]$ in two alternative ways: *(i)* by computing the values in $\mathcal{D}[a^{-1}(\mathcal{T}')]$ at the micro-level, then abstracting via $\alpha_{\mathcal{T}'}$; or *(ii)* by abstracting via $\alpha_{\mathcal{X}'_{\mathcal{I}}}$ to $\mathcal{D}[\mathcal{X}'_{\mathcal{I}}]$ and then computing the values in $\mathcal{D}[\mathcal{T}']$ at the macro-level. The $\boldsymbol{\alpha}$-abstraction manifests in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ as a natural transformation between the micro-level M and the macro-level M'. As highlighted in Definitions 2 and 3, our category-theoretic formalism highlights the role of the exogenous. Accordingly, we need to consider the exogenous to properly extend the $\boldsymbol{\alpha}$-abstraction to $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$. Specifically, a and α have two components, the first for the exogenous, and the second for the endogenous, as formalized below.

Definition 5 ($\boldsymbol{\alpha}$-abstraction in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$) *Given micro- and macro-level SCMs,* M *and* M'*, respectively, an* $\boldsymbol{\alpha}$*-abstraction is a tuple* $\boldsymbol{\alpha} = \langle \mathcal{R}, \mathcal{Q}, a, \alpha \rangle$*, where: (i)* $\mathcal{R} \subseteq \mathcal{X}$ *is a set of relevant endogenous variables; (ii)* $\mathcal{Q} \subseteq \mathcal{Z}$ *is a set of relevant exogenous variables given by the union of the set of exogenous corresponding to the endogenous in* $\mathcal{R}$ *and those corresponding to their ancestors; (iii)* $a = \langle a_{\mathcal{Z}}, a_{\mathcal{X}} \rangle$ *is a pair of surjective functions mapping sets,* $a_{\mathcal{Z}} : \mathcal{Q} \to \mathcal{Z}'$ *and* $a_{\mathcal{X}} : \mathcal{R} \to \mathcal{X}'$*, respectively; (iv)* $\alpha = \langle \alpha_{\mathcal{Z}}, \alpha_{\mathcal{X}} \rangle$ *is a natural transformation composed of measurable functions mapping probability spaces,* $\alpha_{\mathcal{Z}}$ *for the exogenous and* $\alpha_{\mathcal{X}}$ *for the endogenous, respectively.*

In addition, interventional consistency translates into commutation of diagrams in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$. Here, we look at an intervention on the endogenous $\mathcal{X}'_{\mathcal{I}} \in \mathcal{X}'$ on the macro-level model M' as a collection of measurable maps, as given in Lemma 1, and denote it by $\mathcal{I}'$. Similarly for the corresponding intervention on the micro-level variables $a_{\mathcal{X}}^{-1}(\mathcal{X}'_{\mathcal{I}}) \in \mathcal{X}$, denoted by $\mathcal{I}$. To aid visualization, we use *(i)* violet for micro-level M layer, *(ii)* gray for macro-level M' layer, and *(iii)* orange for the $\boldsymbol{\alpha}$-abstraction.

Definition 6 (IC $\boldsymbol{\alpha}$-abstraction in $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$) *An* $\boldsymbol{\alpha}$*-abstraction is IC in* $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ *if, for all the interventions* $\eta^{\mathcal{I}'} = \langle \mathrm{Id}_{\mathsf{M}'(I)}, \mathcal{I}' \rangle$ *on the macro-level model* M'*, the faces of the following diagram commute*



Starting from Definition 6, exploiting Theorem 1, we obtain the following.

Theorem 2 *The component* $\alpha_{\mathcal{X}}$ *of* α *within an IC* $\boldsymbol{\alpha}$*-abstraction commutes with* cc_{λ}*, thus is affine.*

Proof. See appendix D. □

At this point, denoted by $\mathsf{NI}(\mathsf{I}, \mathsf{Prob})$ the subcategory of $\mathsf{SCM}(\mathsf{I}, \mathsf{Prob})$ whose objects are non-intervened SCMs, and morphisms are IC $\boldsymbol{\alpha}$-abstraction. Exploiting Theorem 1 and Theorem 2, we have the following.

Theorem 3 *There exists an IC encoding functor* $E : \mathsf{NI}(\mathsf{I}, \mathsf{Prob}) \to \mathsf{CSprob}$ *mapping (i) each non-intervened* $\mathsf{M} := \langle (\mathcal{U}, \Upsilon, \zeta), (\mathcal{V}, \Omega, \chi), \mathcal{M} \rangle$ *to the convex spaces of probability measures* $\Delta_{(\mathcal{V}, \Omega)}$*, and (ii) an IC* $\boldsymbol{\alpha}$*-abstraction between* M *and* M' *with its endogenous component* $\alpha_{\mathcal{X}}$*.*

Proof. See appendix D. □

4 THE NETWORK SHEAF AND COSHEAF OF CAUSAL KNOWLEDGE

The last step toward our proposed *relativity of causal knowledge* is representing causal knowledge over the endogenous within a network. To this aim, each node within the network is attached to the causal knowledge entailed by a certain SCM via Theorem 3. Looking at a single node ρ, we interpret it as the perspective of the causal knowledge on itself. Then, when two nodes ρ and σ are connected by an edge, say $\tau : \rho \sim \sigma$, they are related by a shared IC $\boldsymbol{\alpha}$-abstraction

living on the edge τ. The causal knowledge at ρ can be transported via τ to σ giving life to the relative causal knowledge (RCK): *the causal knowledge of ρ from the perspective of σ*. To make this point clearer, consider two AI agents with their subjective causal models M^ρ and M^σ about a certain phenomenon. By Theorem 3, we can map M^ρ and M^σ to the corresponding causal knowledge, viz. $\mathsf{CK}^\rho \equiv \mathsf{CK}\left(\mathsf{M}^\rho\right)$ and $\mathsf{CK}^\sigma \equiv \mathsf{CK}\left(\mathsf{M}^\sigma\right)$. The causal model M^ρ admits an IC causal abstraction, viz. M^τ. The same abstraction holds for M^σ. Acting on M^τ as well, the encoding functor provides us with a convex space CK^τ which can be interpreted as a *backbone space* on which the CK^ρ and CK^σ can relate each other. In particular, the AI agents use the backbone space to *embed* the causal knowledge of the other into their own. Intuitively, consider that the AI agent ρ *projects* via an abstraction morphism an observational/interventional probability measure onto the backbone space CK^τ, matching an abstracted measure $\chi^{\rho,\tau}$. The latter can be observed by the AI agent σ, who embed $\chi^{\rho,\tau}$ in its own causal knowledge obtaining the relative $\chi^{\rho,\sigma}$, namely the causal probability measure entailed by M^ρ from the perspective of M^σ.

Using the categorical foundations laid in Sections 2 and 3, we formalize RCK through the *network sheaf and cosheaf of causal knowledge*, the latter being the dual construction of the former. Here, the "dual" nomenclature is used with a slight abuse of notation to mimic the jargon used in Abelian categories [5, 12]. These mathematical objects consist of *(i)* a network, *(ii)* convex spaces encoding causal knowledge on nodes and edges, and *(iii)* mappings to move from nodes to edges in the case of the sheaf and vice-versa for the cosheaf. First, we introduce the network as a topological object. Recall that the network is shaped by the existence of an IC CA between SCMs. Such a network is shared by the network sheaf and cosheaf.

Definition 7 (Network) *A (finite) network $G := (\mathcal{N}, \mathcal{E})$ consists of: (i) nodes $\rho \in \mathcal{N}$ homeomorphic to a point (0-dimensional open ball), and (ii) edges $\tau \in \mathcal{E}$ homeomorphic to an open interval (1-dimensional open ball). The closure of each edge is the edge itself plus the two nodes at its boundaries, whereas the node has its closure as there are no lower-dimensional constituents of the network. The face incidence relation induces a partial order $(\Pi, \trianglelefteq)$ on the set of nodes and edges, that is, $\rho \trianglelefteq \tau$ if and only if ρ (node) belongs to the closure of τ (edge).*

Second, the convex spaces of probability measures are objects in CSprob. For the network sheaf, they are dubbed *stalks* and are given by the causal knowledge on the network constituents encoded by the functor E. Consider $\tau : \rho \sim \sigma$. The stalk at ρ is the image $E(\mathsf{M}^\rho)$ in CSprob of the SCM M^ρ; similarly for σ. The stalk at τ instead is the image $E(\mathsf{M}^\tau)$ in CSprob of the causal abstraction shared by ρ and σ, viz. M^τ. The role of the CA is key to relating the causal knowledge at ρ and σ as it provides a *backbone space* where the causal knowledge can be compared. For the network cosheaf instead, the convex spaces are dubbed *costalks*. Consider again $\tau : \rho \sim \sigma$. The costalk at τ, viz. $\widehat{E}(\mathsf{M}^\tau)$, coincides with $E(\mathsf{M}^\tau)$. The costalk at ρ instead is a convex space of probability measures $\widehat{E}(\mathsf{M}^\rho)$, embedded in $E(\mathsf{M}^\rho)$ as specified in the sequel.

Third, the mappings manifest as morphisms in the target category, viz. CSprob. Specifically, in the case of the network sheaf, we *project* $E(\mathsf{M}^\rho)$ onto $E(\mathsf{M}^\tau)$ via the IC CA map, hereinafter *restriction map* and denoted by $\alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau}$. Conversely, in the case of the network cosheaf, we *embed* $E(\mathsf{M}^\tau)$ into $E(\mathsf{M}^\rho)$ obtaining $\widehat{E}(\mathsf{M}^\rho)$ via an affine measurable map $\beta_{\mathcal{X}}^{\rho \trianglelefteq \tau}$, hereinafter *extension map*. At this point, we can formally define the *embedded costalk of causal knowledge* as the convex space $\widehat{E}(\mathsf{M}^\rho) = \langle \widehat{\Delta}_{(\mathcal{V},\Omega)_\rho}, cc_\lambda \rangle$, where the set of probability measures $\widehat{\Delta}_{(\mathcal{V},\Omega)_\rho} \subseteq \Delta_{(\mathcal{V},\Omega)_\rho}$ is

$$\widehat{\Delta}_{(\mathcal{V},\Omega)_\rho} := \{\widehat{\chi}^\rho \text{ on } (\mathcal{V},\Omega)_\rho \,:\, \alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau}(\widehat{\chi}^\rho) = \chi^\tau\}\,, \quad (9)$$

Please notice that Equation (9) and Theorem 2 guarantee that $\widehat{E}(\mathsf{M}^\rho)$ admits $E(\mathsf{M}^\tau)$ as an IC CA. We deliberately use the verbs "to project" and "to embed" to remark that in the node-edge transition formalized by the network sheaf, some of the causal knowledge idiosyncratic to the node is not transported into the backbone space; therefore, in the reverse edge-node transition formalized by the network cosheaf, it is only possible to integrate the transported causal knowledge into the causal knowledge of the node, with no guarantee of perfect reconstruction. This concept is essential for us to properly define the transfer of probability measures on the network. At this point, we are ready to define the network sheaf and cosheaf of causal knowledge.

Definition 8 (Network sheaf of causal knowledge) *Given a network $G := (\mathcal{N}, \mathcal{E})$ with face incidence poset $(\Pi, \trianglelefteq)$, a network sheaf valued in CSprob is a functor $F : (\Pi, \trianglelefteq) \to \mathsf{CSprob}$ assigning (i) to each node ρ and edge τ in $(\Pi, \trianglelefteq)$ stalks consisting of convex spaces of probability measure, $E(\mathsf{M}^\rho)$ and $E(\mathsf{M}^\tau)$, respectively; (ii) to each node-edge incidence relation a restriction map $\alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau} : E(\mathsf{M}^\rho) \to E(\mathsf{M}^\tau)$ being the affine endogenous component within an IC $\boldsymbol{\alpha}$-abstraction between M^ρ and M^τ.*

Definition 9 (Network cosheaf of causal knowledge) *Given a network $G := (\mathcal{N}, \mathcal{E})$ with face incidence poset $(\Pi, \trianglelefteq)$, a network cosheaf valued in CSprob is a functor $\widehat{F} : (\Pi, \trianglelefteq)^{\mathrm{op}} \to \mathsf{CSprob}$ assigning (i) to each node ρ and edge τ in $(\Pi, \trianglelefteq)^{\mathrm{op}}$ costalks consisting of convex spaces of probability measure, $\widehat{E}(\mathsf{M}^\rho)$ and $\widehat{E}(\mathsf{M}^\tau) \equiv E(\mathsf{M}^\tau)$, respectively; (ii) to each node-edge incidence relation an extension map $\beta_{\mathcal{X}}^{\rho \trianglelefteq \tau} : \widehat{E}(\mathsf{M}^\tau) \to \widehat{E}(\mathsf{M}^\rho)$.*

The network sheaf and cosheaf fulfill two complementary functions: the former transports causal knowledge from the nodes to a backbone space where such knowledge can be compared; the latter distributes causal knowledge from the backbone space to the network nodes. Definitions 8 and 9 are particular cases of more general network sheaf and cosheaf in CSprob, in which every causal-related re-

quirements on the (co)stalks and the (extension)restriction maps are dropped. For the network sheaf, we refer to a 0-cochain as a collection of (interventional or observational) probability measures, one for each node. The value of the 0-cochain at the node ρ is χ^ρ. Additionally, the 1-cochain is a collection of probability measures on the edges, representing the IC CA of those on the nodes. Hence, an induced value of the 1-cochain at the edge τ incident to a node ρ is the pushforward probability measure $\chi^\tau = \alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau}(\chi^\rho)$. We denote the space of 0- and 1-cochains by $\mathcal{C}^0(G;F)$ and $\mathcal{C}^1(G;F)$, respectively. An object of interest in sheaf theory [5] is the *global section* of the network sheaf. Loosely speaking, it is a consistent assignment of data to each node of the network that does not break local rules.

Definition 10 (Global section) *Consider F as in Definition 8. A global section χ of F is a choice χ^ρ in $E(\mathsf{M}^\rho)$ for each node ρ of G such that $\chi^\tau = \alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau}(\chi^\rho)$, for all $\rho \trianglelefteq \tau$. The space of global sections of F is denoted by $\Gamma(G;F)$.*

Intuitively, in our RCK, a global section manifests itself as a collection of as many consistent (non-)interventions as the number of connected components of the network G. Specifically, consider a single connected component: according to our modeling, each pair of nodes connected by an edge admits a shared IC CA. Hence, choosing a consistent (non-)intervention for each node would result in a global section vanishing a certain notion of distance – i.e., any suitable information-theoretic metric and ϕ-divergence – between the projected probability measures onto the edges. This is a direct consequence of the interventional consistency inherently ensured by the restriction maps.
Regarding the network cosheaf $\widehat{F}$, the 1-chain agrees with the 1-cochain of F, that is, a collection of probability measures representing the IC CA of those on the nodes. The 0-chain is a collection of probability measures, one for each node, satisfying the projection requirement in Equation (9). We denote by $\mathcal{C}^0(G;\widehat{F})$ and $\mathcal{C}^1(G;\widehat{F})$ the spaces of 0- and 1-chains of $\widehat{F}$. Ultimately, we can pose the formal definition of *relative causal knowledge*:

Definition 11 (Relative Causal Knowledge) *Given a network $G := (\mathcal{N}, \mathcal{E})$, network sheaf and cosheaf of causal knowledge on G, F and $\widehat{F}$, respectively, and two nodes ρ_1 and σ_k connected by a path of k edges $\tau_i : \rho_i \sim \sigma_i$, with $i \in [k]$, the relative causal knowledge is the causal knowledge at ρ_1 from the perspective of σ_k filtered by the shared, abstract, causal knowledge on the edges τ_i, $i \in [k]$ (i.e., path-dependent). Specifically,*

$$\mathsf{CK}^{\rho_1,\sigma_k} := \{ \chi^{\rho_1,\sigma_k} = \beta_{\mathcal{X}}^{\sigma_k \trianglelefteq \tau_k} \circ \alpha_{\mathcal{X}}^{\rho_k \trianglelefteq \tau_k} \circ \ldots \circ \beta_{\mathcal{X}}^{\sigma_1 \trianglelefteq \tau_1} \circ \alpha_{\mathcal{X}}^{\rho_1 \trianglelefteq \tau_1}(\chi^\rho), \; \chi^\rho \in E(\mathsf{M}^\rho) \}. \tag{10}$$

We end the section with a working example on the transfer of causal knowledge. Further discussion and exemplification of RCK is provided in Appendix E.

Setup. Consider a network with three nodes and two edges, forming a chain. We assign to the leftmost node the causal knowledge of subject A, viz. CK^A. Similarly, we attach to the central and rightmost nodes the causal knowledge of subject B and C, viz. CK^B and CK^C. The endogenous variables for the subjects are $\mathcal{A} = \{A_1, A_2, A_3\}$, $\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5\}$, $\mathcal{C} = \{C_1, C_2, C_3\}$, composing non-intervened SMCs M^A, M^B, and M^C. Each subject does not have any information about the SCMs of the others. M^A and M^B admit a shared CA M^X living on the shared edge, consisting of a single causal variable X for simplicity. Specifically, for A the CA structure is $\{A_1, A_2\} \to X$; for B is $\{B_1, B_2, B_3\} \to X$. M^B and M^C admit a shared CA M^Y, given by a single causal variable Y. Here, for B the structure is $\{B_3, B_4, B_5\} \to Y$; for C is $\{C_1, C_2\} \to Y$. For convenience, consider the restriction and extension maps, viz. $\alpha_{\mathcal{X}}^{\rho \trianglelefteq \tau}$ and $\beta_{\mathcal{X}}^{\rho \trianglelefteq \tau}$ to be linear (i.e., matrices). For A, we denote the restriction by $\mathbf{F}^{A \trianglelefteq X} \in \mathbb{R}^{1\times 3}$ and the extension by $\widehat{\mathbf{F}}^{A \trianglelefteq X} \in \mathbb{R}^{3\times 1}$. Similarly, for B we have $\mathbf{F}^{B \trianglelefteq X}$, $\mathbf{F}^{B \trianglelefteq Y}$ in $\mathbb{R}^{1\times 5}$ as restrictions and $\widehat{\mathbf{F}}^{B \trianglelefteq X}$, $\widehat{\mathbf{F}}^{B \trianglelefteq Y}$ in $\mathbb{R}^{5\times 1}$ as extensions. Finally, for C we have $\mathbf{F}^{C \trianglelefteq X} \in \mathbb{R}^{1\times 3}$ and $\widehat{\mathbf{F}}^{C \trianglelefteq X} \in \mathbb{R}^{3\times 1}$. Zero entries in the restriction and extension matrices manage the fact that some variables are not relevant to CA, a.k.a. non-constructive abstraction.

Relative Causal Knowledge. The subjects perform certain tasks through individual soft-intervened SCMs, entailing the soft-intervened probability measures $\chi_{\mathcal{S}}^A = N(0, \mathbf{\Sigma}_{\mathcal{S}}^A)$, $\chi_{\mathcal{S}}^B = N(0, \mathbf{\Sigma}_{\mathcal{S}}^B)$, and $\chi_{\mathcal{S}}^C = N(0, \mathbf{\Sigma}_{\mathcal{S}}^C)$. The collection of these measures constitutes a valuation of the node stalks of the network sheaf, i.e., a 0-cochain. When projected onto the edges via the above CAs, the 0-cochain entails the 1-cochain, that is, a collection of probability measures for the edges X and Y. Then, Definition 11 specifies for instance how B and C see A (i.e., $\chi_{\mathcal{S}}^A$) from their perspectives. Specifically, the relative soft-intervened measures are:
(i) for B, $\chi_{\mathcal{S}}^{A,B} = N(0, \widehat{\mathbf{F}}^{B \trianglelefteq X} \mathbf{F}^{A \trianglelefteq X} \mathbf{\Sigma}_{\mathcal{S}}^A \mathbf{F}^{A \trianglelefteq X^\top} \widehat{\mathbf{F}}^{B \trianglelefteq X^\top}) = N(0, \mathbf{\Sigma}_{\mathcal{S}}^{A,B})$;
(ii) for C, $\chi_{\mathcal{S}}^{A,C} = N(0, \widehat{\mathbf{F}}^{C \trianglelefteq Y} \mathbf{F}^{B \trianglelefteq Y} \mathbf{\Sigma}_{\mathcal{S}}^{A,B} \mathbf{F}^{B \trianglelefteq Y^\top} \widehat{\mathbf{F}}^{C \trianglelefteq Y^\top}) = N(0, \mathbf{\Sigma}_{\mathcal{S}}^{A,C})$.

Global section. A key point of our framework is the possibility of investigating the global agreement among subjects in terms of causal knowledge. Such a global agreement gives rise to the global section in Definition 10. In the working example, we have a global section when the soft-interventions performed by the subjects are such that
$\mathbf{F}^{A \trianglelefteq X} \mathbf{\Sigma}_{\mathcal{S}}^A \mathbf{F}^{A \trianglelefteq X^\top} = \mathbf{F}^{B \trianglelefteq X} \mathbf{\Sigma}_{\mathcal{S}}^B \mathbf{F}^{B \trianglelefteq X^\top}$ (section on X)
and
$\mathbf{F}^{B \trianglelefteq Y} \mathbf{\Sigma}_{\mathcal{S}}^B \mathbf{F}^{B \trianglelefteq Y^\top} = \mathbf{F}^{C \trianglelefteq Y} \mathbf{\Sigma}_{\mathcal{S}}^C \mathbf{F}^{C \trianglelefteq Y^\top}$ (section on Y).

5 DISCUSSION

Conceptually, the relativity of causal knowledge can drive a paradigm shift in how causality is typically understood in AI/ML. By stripping causality of its oracular and absolute meaning, the relativity of causal knowledge situates it within a different ontological setting, where truth is not monolithic but emerges inevitably and relatively from a set of relationships. This is mathematically formalized by the path-dependent RCK in Equation (10). Our framework paves the way to multiple intriguing areas of reasearch, pivotal to fully characterize RCK and make it applicable.

Learning Theory. We developed our theoretical framework with an underlying intuition: RCK could be a better framework to empower AI/ML with the ladder of causation [23]. A learning theory is needed for this. It is surely of interest to develop a statistical theory for models that can ingest network sheaves and cosheaves of causal knowledge to solve downstream learning tasks. In many cases, the restriction and expansion maps – the formal instance of "perspective" – are given, e.g., privacy requirements of the subjects in an agentic network or communication constraints in a sensor network. However, when they are not available, the problem of *(co)sheaf inference*, i.e., learning them in a data-driven fashion given the network connectivity, is fundamental. An even harder problem is *(co)sheaf discovery*, i.e., jointly learning the restriction, extension maps, and network connectivity. Interestingly, both for inference and discovery, the learning losses and models are useful to induce a particular notion of perspective, that could be fairness-, safety-, information-, or trustworthiness-driven, just to name a few.

Cohomology Theory. CSprob is not an Abelian category, as it is not additive, i.e., there is no natural way to "add" two affine maps to get another affine map such that all group axioms are globally satisfied. Consequently, defining a cochain complex – and thus a cohomology theory – on network sheaves of causal knowledge is nontrivial. Developing such a theory would yield rich algebraic invariants that offer global insights into the network structure. These invariants would be instrumental in revealing when and why multiple sheaves of CK share a common structure. Practically, a cohomology theory could help us understand how and why different networks of subjects, each developing causal knowledge about different phenomena, exhibit similarities.

Hodge Theory. The Hodge theorem [11] shows that the kernel of the Laplacian on a Riemannian manifold recovers its de Rham cohomology. Combinatorial versions of the Hodge theory have been developed for cellular sheaves of inner-product spaces (e.g., vector spaces and Hilbert spaces), building on suitable properties of the adjoint maps [12]. Unfortunately, porting these arguments to the CSprob category is nontrivial. In this context, a Hodge(-like) theory could be instrumental in defining diffusion-type operators and developing spectral theories [12] for RCK. Practically, having a Hodge theory might help us understand which are the dynamics that lead to global sections in Definition 10 and why, providing insights on when and why RCK aligns.

6 CONCLUSION

In this paper, we introduced the *relativity of causal knowledge*. We established the formal definition of *relative causal knowledge* (RCK) via a category-theoretic treatment for SCMs, their interventions, and related α-abstractions. Within the relativity of the causal knowledge paradigm, SCMs are inherently imperfect, subjective representations embedded within networks of relationships. By encoding causal knowledge in convex spaces of probability measures and constructing network sheaves and cosheaves, we enabled the transfer of causal knowledge in an interventionally consistent manner while considering diverse perspectives across the network. Our work opens exciting research avenues in cohomology, Hodge(-like), and learning theory for RCK, paving the way for context-aware, robust AI systems.

Author Contributions

The Authors equally contributed to the original idea, the thorough conceptual development of the framework, and the paper writing. GD developed the category-theoretic framework, and established and proved the related theoretical results. GD and CB codeveloped the sheaf-theoretic characterization of RCK.

Acknowledgements

The work of Gabriele D'Acunto was supported by *(i)* the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on " Telecommunications of the Future" (PE00000001 - program " RESTART"), and *(ii)* the SNS JU project 6G-GOALS [32] under the EU's Horizon program Grant Agreement No 101139232.

The work of Claudio Battiloro was supported by the National Institutes of Health (NIH) Grant 1R01ES037156-01.

The Authors thank Dr. Hans Riess from Duke University and Dr. Fabio Massimo Zennaro from University of Bergen for valuable comments on a preliminary version of the work.

References

[1] Elias Bareinboim. *Causal Artificial Intelligence: A Roadmap for Building Causally Intelligent Systems*. Publisher Name, 2025. URL https://causalai-book.net/. (Cited on 16)

[2] Elias Bareinboim and Judea Pearl. Causal inference and the data-fusion problem. *Proceedings of the Na-*

tional Academy of Sciences, 113(27):7345–7352, 2016. (Cited on 2)

[3] Sander Beckers. Equivalent causal models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 6202–6209, 2021. (Cited on 2)

[4] Sander Beckers and Joseph Y Halpern. Abstracting causal models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 2678–2685, 2019. (Cited on 2)

[5] Justin Michael Curry. *Sheaves, cosheaves and applications*. University of Pennsylvania, 2014. (Cited on 2, 3, 7, 8)

[6] Gabriele D'Acunto, Gianmarco De Francisci Morales, Paolo Bajardi, and Francesco Bonchi. Learning multiscale non-stationary causal structures. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL https://openreview.net/forum?id=SQnPE63jtA. (Cited on 13)

[7] Gabriele D'Acunto, Fabio Massimo Zennaro, Yorgos Felekis, and Paolo Di Lorenzo. Causal abstraction learning based on the semantic embedding principle. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=J16AIOkjjY. (Cited on 2, 5)

[8] Frederick Eberhardt and Richard Scheines. Interventions and causal inference. *Philosophy of science*, 74 (5):981–995, 2007. (Cited on 2, 4)

[9] Tobias Fritz. Convex spaces I: Definition and examples. *arXiv preprint arXiv:0903.5522*, 2009. (Cited on 2, 5, 15)

[10] Robert Ghrist and Hans Riess. Cellular sheaves of lattices and the tarski laplacian. *Homology, Homotopy and Applications*, 24(1), 2022. (Cited on 3)

[11] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. John Wiley & Sons, 2014. (Cited on 9)

[12] Jakob Hansen and Robert Ghrist. Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3(4):315–358, 2019. (Cited on 3, 7, 9)

[13] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2):3, 2022. (Cited on 19)

[14] Alexander Immer, Christoph Schultheiss, Julia E Vogt, Bernhard Schölkopf, Peter Bühlmann, and Alexander Marx. On the identifiability and estimation of causal location-scale noise models. In *International Conference on Machine Learning*, pages 14316–14332. PMLR, 2023. (Cited on 14)

[15] Bart Jacobs, Aleks Kissinger, and Fabio Zanasi. Causal inference by string diagram surgery. In *Foundations of Software Science and Computation Structures: 22nd International Conference, FOSSACS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6–11, 2019, Proceedings 22*, pages 313–329. Springer, 2019. (Cited on 2)

[16] David Kaltenpoth and Jilles Vreeken. Nonlinear causal discovery with latent confounders. In *International Conference on Machine Learning*, pages 15639–15654. PMLR, 2023. (Cited on 14)

[17] Bruno Latour. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press, 07 2005. ISBN 9780199256044. doi: 10.1093/oso/9780199256044.001.0001. URL https://doi.org/10.1093/oso/9780199256044.001.0001. (Cited on 1)

[18] Saunders Mac Lane. *Categories for the working mathematician*, volume 5. Springer Science & Business Media, 2013. (Cited on 2, 3, 12)

[19] Sara Magliacane, Thijs Van Ommen, Tom Claassen, Stephan Bongers, Philip Versteeg, and Joris M Mooij. Domain adaptation by using causal inference to predict invariant conditional distributions. *Advances in Neural Information Processing Systems*, 31, 2018. (Cited on 2)

[20] Colin McLarty. The rising sea: Grothendieck on simplicity and generality I. 2003. (Cited on 2)

[21] Judea Pearl. *Causality*. Cambridge University Press, 2009. (Cited on 1, 2, 3, 4)

[22] Judea Pearl and Elias Bareinboim. Transportability of causal and statistical relations: A formal approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 25, pages 247–254, 2011. (Cited on 2)

[23] Judea Pearl and Dana Mackenzie. *The book of why: The new science of cause and effect*. Basic books, 2018. (Cited on 1, 4, 9)

[24] Paolo Perrone. *Starting category theory*. World Scientific, 2024. (Cited on 12)

[25] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: Foundations and learning algorithms*. The MIT Press, 2017. (Cited on 4)

[26] Stathis Psillos. Causal pluralism. In *Worldviews, science and us: Studies of analytical metaphysics: A selection of topics from a methodological perspective*, pages 131–151. World Scientific, 2010. (Cited on 1)

[27] Jonathan Richens and Tom Everitt. Robust agents learn causal world models. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=pOoKI3ouv1. (Cited on 1, 16)

[28] Eigil Fjeldgren Rischel. *The category theory of causal models*. PhD thesis, University of Copenhagen, 2020. (Cited on 2, 6)

[29] Mateo Rojas-Carulla, Bernhard Schölkopf, Richard Turner, and Jonas Peters. Invariant models for causal transfer learning. *Journal of Machine Learning Research*, 19(36):1–34, 2018. (Cited on 2)

[30] P Rubenstein, S Weichwald, S Bongers, J Mooij, D Janzing, M Grosse-Wentrup, and B Schölkopf. Causal consistency of structural equation models. In *33rd Conference on Uncertainty in Artificial Intelligence (UAI 2017)*, pages 808–817. Curran Associates, Inc., 2017. (Cited on 2)

[31] Jean-Pierre Serre. Faisceaux algébriques cohérents. *Annals of Mathematics*, pages 197–278, 1955. (Cited on 2)

[32] Emilio Calvanese Strinati, Paolo Di Lorenzo, Vincenzo Sciancalepore, Adnan Aijaz, Marios Kountouris, Deniz Gündüz, Petar Popovski, Mohamed Sana, Photios A Stavrou, Beatriz Soret, et al. Goal-oriented and semantic communication in 6G AI-native networks: The 6G-GOALS approach. In *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pages 1–6. IEEE, 2024. (Cited on 9)

[33] Thomas S Verma and Judea Pearl. Equivalence and synthesis of causal models. In *Probabilistic and causal inference: The works of Judea Pearl*, pages 221–236. 2022. (Cited on 2)

[34] Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Society, 2021. (Cited on 5)

[35] James Woodward. Causation and manipulability. 2001. (Cited on 1)

[36] Kun Zhang and Aapo Hyvärinen. On the identifiability of the post-nonlinear causal model. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 647–655, 2009. (Cited on 14)

[37] Kun Zhang, Mingming Gong, and Bernhard Schölkopf. Multi-source domain adaptation: A causal view. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 2015. (Cited on 2)

The Relativity of Causal Knowledge
(Supplementary Material)

Gabriele D'Acunto[1,2] **Claudio Battiloro**[3]

[1]Information Engineering, Electronics and Telecommunications Dept., Sapienza University, Rome, Italy
[2]National Inter-University Consortium for Telecommunications (CNIT), Parma, Italy
[3]Biostatistics Dept., Harvard University, Cambridge, MA, USA

A ESSENTIAL MATERIAL ON CATEGORY THEORY

Below are fundamental definitions and examples that are instrumental in providing the necessary background on category theory. For a comprehensive overview of category theory see resources such as [18, 24].

Definition 12 (Category) *A category* C *consists of*

- *A collection* $\mathcal{C}_0$ *whose elements are called objects of* C*;*
- *A collection* $\mathcal{C}_1$ *whose elements are called morphisms of* C*;*

such that:

- *Each morphism* $f : X \to Y$ *has assigned two objects of the category called source and target, respectively;*
- *Each object* X *has a distinguished identity morphism* $\mathrm{Id}_X : X \to X$*;*
- *Given* $f : X \to Y$ *and* $g : Y \to Z$*, than it exists* $g \circ f = h : X \to Z$*.*

These structures satisfy the following axioms:

- *(Unitality)* $\forall f : X \to Y,\ f \circ \mathrm{Id}_X = f$ *and* $\mathrm{Id}_Y \circ f = f$*;*
- *(Associativity) Given* f*,* g*, and* h *such that the compositions are defined, then* $h \circ (g \circ f) = (h \circ g) \circ f$*.*

Example 1 *The following are some notable examples of categories:*

- *Indicate with* $(\mathcal{P}, \leqslant)$ *a partial order set.* $(\mathcal{P}, \leqslant)$ *can be viewed as the category whose objects are the elements* $p \in \mathcal{P}$ *and morphisms are order relations* $p \leq p'$*. Notice that there is at most one morphism between two objects;*
- $\mathsf{Vect}_{\mathbb{R}}$ *is the category whose objects are real vector spaces and morphisms are linear maps;*
- $\mathsf{Hilb}_{\mathbb{R}}$ *is the category whose objects are real complete Hilbert spaces, that is, real vector spaces equipped with an inner product structure and closed under the norm topology. Morphisms are bounded linear maps* $f : X \to Y$*, that is,* $\exists\, b \in \mathbb{R}^+$ *such that* $\|f(x)\| \leq b\|x\|,\ \forall x \in X$*. Every morphism* $f : X \to Y$ *between Hilbert spaces admits an* adjoint $f^\star$ *such that,* $\forall\, x \in X$ *and* $y \in Y$*,* $\langle y,\, f(x)\rangle = \langle f^\star(y),\, x\rangle$*. We also have* $(f^\star)^\star = f$*.*

Starting from Definition 12 we can define the notion of *subcategory*.

Definition 13 *Given a category* C*, a subcategory* S *of* C *consists of the following data:*

- *A subcollection* $\mathcal{S}_0$ *of* $\mathcal{C}_0$ *such that* $\forall\, S$*,* Id_S *is in* S*;*
- *A subcollection* $\mathcal{S}_1$ *of* $\mathcal{C}_1$ *such that* $\forall\, f$ *and* g *for which the composition is defined in* C*, the composite* $h = g \circ f$ *is in* S*.*

Example 2 ($\mathsf{preHilb}_{\mathbb{R}}$) *The category of real vector spaces with an inner product* $\mathsf{preHilb}_{\mathbb{R}}$ *is a subcategory of* $\mathsf{Vect}_{\mathbb{R}}$*. Specifically, objects are real vector spaces equipped with an inner product structure. Morphisms are bounded linear maps* $f : X \to Y$*, that is,* $\exists\, b \in \mathbb{R}^+$ *such that* $\|f(x)\| \leq b\|x\|,\ \forall x \in X$*. Every morphism* $f : X \to Y$ *between inner product vector spaces admits an* adjoint $f^\star$ *such that,* $\forall\, x \in X$ *and* $y \in Y$*,* $\langle y,\, f(x)\rangle = \langle f^\star(y),\, x\rangle$*. We also have* $(f^\star)^\star = f$*.* $\mathsf{preHilb}_{\mathbb{R}}$ *is similar to* $\mathsf{Hilb}_{\mathbb{R}}$*, the difference is that in* $\mathsf{Hilb}_{\mathbb{R}}$ *the vector spaces are also complete.*

Definition 14 (Functor) *Consider* C *and* D *categories. A functor* $F : \mathsf{C} \to \mathsf{D}$ *consists of the following data:*

- *For each object* $X \in \mathcal{C}_0$*, an object* $F(X) \in \mathcal{D}_0$*;*
- *For each object morphism* $\mathcal{C}_1 \ni f : X \to Y$*, a morphism* $\mathcal{D}_1 \ni F(f) : F(X) \to F(Y)$*;*

such that the following axioms hold:

- *(Unitality)* $\forall X \in \mathcal{C}_0,\ F(\mathrm{Id}_X) = \mathrm{Id}_F(X)$*. In other words, the identity in* C *is mapped into the identity in* D*.*
- *(Compositionality)* $\forall f$ *and* $g \in \mathcal{C}_1$ *such that the composition is defined, then* $F(g \circ f) = F(g) \circ F(f)$*. In other words, the composition in* C *is mapped into the composition in* D*.*

To ease the notation, in the sequel, we use F^σ and F^f to denote $F(\sigma)$ and $F(f)$, respectively.

Definition 15 (Natural transformation) *Consider two categories* C *and* D*, and two functors between them, namely* $F : \mathsf{C} \to \mathsf{D}$ *and* $G : \mathsf{C} \to \mathsf{D}$*. A natural transformation* $\alpha : F \to G$ *consists of the following data:*

- *For each object* $X \in \mathcal{C}_0$*, a morphism* $\alpha_X : F^X \to G^X$ *in* $\mathcal{D}$ *called the component of* α *at* X*;*
- *For each morphism* $f : X \to X'$ *in* $\mathcal{C}$*, the following diagram commutes:*

$$\begin{tikzcd} F^X \arrow[r, "F^f"] \arrow[d, "\alpha_X"'] & F^{X'} \arrow[d, "\alpha_{X'}"] \\ G^X \arrow[r, "G^f"'] & G^{X'} \end{tikzcd} \tag{11}$$

A natural transformation can be thought of as a consistent system of arrows between two functors. The *naturality* of α means that it is invariant with respect to maps between the images of two functors.

B EXAMPLES

This appendix is devoted to provide practical examples about the formalism introduced in Section 2. Consider $\mathcal{Z} = \{Z_1, Z_2, Z_3\}$ and $\mathcal{X} = \{X_1, X_2, X_3\}$ as exogenous and endogenous variables of the following linear SCM with additive noise M:

$$\begin{cases} X_1 = Z_1 = m_1(Z_1)\,, \\ X_2 = c_{21}X_1 + Z_2 = c_{21}Z_1 + Z_2 = m_2(Z_1, Z_2)\,, \\ X_3 = c_{32}X_2 + Z_3 = c_{32}c_{21}Z_1 + c_{32}Z_2 + Z_3 = m_3(Z_1, Z_2, Z_3)\,; \end{cases} \tag{12}$$

where $Z_i \sim N(0,1)$ and $Z_i \perp Z_j$, $i, j \in [3]$ and $i \neq j$. Additionally, let the probability space associated with the exogenous to be $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra of $\mathbb{R}^n$. We can equivalently express M as $\mathbf{X} = \mathbf{M}\mathbf{Z}$, where

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 \\ c_{21} & 1 & 0 \\ c_{32}c_{21} & c_{32} & 1 \end{bmatrix}. \tag{13}$$

We call the matrix $\mathbf{M}$ the mixing matrix. See Lemma E.1 in D'Acunto et al. [6] for a proof of the existence of $\mathbf{M}$.

At this point, according to Definition 1, the components of the measure-theoretic M are: *(i)* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))$ for the exogenous; *(ii)* $\mathbf{M}$ as mixing; and *(iii)* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}, \mathbf{M}\mathbf{M}^\top))$ for the endogenous. In detail, the endogenous measure is the pushforward of $N(\mathbf{0}, \mathbf{I}_3)$ given $\mathbf{M}$. Using the notation in the paper for pushforward measures, each X_i follows $m_i(\mu(Z_i \cup \mathcal{Z}^{\mathcal{A}_i})) = m_i(N(\mathbf{0}, \mathbf{I}_{|\mathcal{A}_i|+1})) = N(0, [\mathbf{M}]_{i,:}[\mathbf{M}]_{i,:}^\top)$.

Starting from the measure-theoretic SCM, according to Definition 2, the category-theoretic representation of M is the functor mapping I to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))$, I' to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}, \mathbf{M}\mathbf{M}^\top))$, and the arrow between I and I' to $\mathbf{M}$.

At this point, let us consider hard and soft interventions. Starting with the former, we consider the hard intervention

$\mathrm{do}(X_2 = c)$. Accordingly, Equation (12) becomes

$$\begin{cases} X_1 = Z_1 = m_1^{\mathcal{H}}(Z_1) \sim N(0,1)\,, \\ X_2 = c \quad \text{(constant)}, \\ X_3 = c_{32}c + Z_3 = d + Z_3 = m_3^{\mathcal{H}}(Z_3) \sim N(d,1)\,. \end{cases} \tag{14}$$

Starting from Equation (14), the hard-intervened SCM $\mathsf{M}^{\mathcal{H}}$ in $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$ is the functor mapping I to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))$, I' to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(0,1)\times\delta_c\times N(d,1))$, and the arrow between I and I' to $\mathcal{M}_{\mathcal{H}} = \{m_1^{\mathcal{H}}, c, m_3^{\mathcal{H}}\}$.

Consider now a soft intervention on X_2 which modifies $c_{21} \to c_{21}^{\mathcal{S}}$. Recall that we do not consider soft interventions that modify the parent set of the intervened variable. Accordingly, Equation (12) becomes

$$\begin{cases} X_1 = Z_1 = m_1^{\mathcal{S}}(Z_1)\,, \\ X_2 = c_{21}^{\mathcal{S}}X_1 + Z_2 = c_{21}^{\mathcal{S}}Z_1 + Z_2 = m_2^{\mathcal{S}}(Z_1, Z_2)\,, \\ X_3 = c_{32}X_2 + Z_3 = c_{32}c_{21}^{\mathcal{S}}Z_1 + c_{32}Z_2 + Z_3 = m_3^{\mathcal{S}}(Z_1, Z_2, Z_3)\,; \end{cases} \tag{15}$$

For convenience, represent $\mathcal{M}_{\mathcal{S}}$ as

$$\mathbf{M}_{\mathcal{S}} = \begin{bmatrix} 1 & 0 & 0 \\ c_{21}^{\mathcal{S}} & 1 & 0 \\ c_{32}c_{21}^{\mathcal{S}} & c_{32} & 1 \end{bmatrix}. \tag{16}$$

At this point, the soft-intervened SCM $\mathsf{M}^{\mathcal{S}}$ in $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$ is the functor mapping I to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))$, I' to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}, \mathbf{M}_{\mathcal{S}}\mathbf{M}_{\mathcal{S}}^\top))$, and the arrow between I and I' to $\mathbf{M}_{\mathcal{S}}$.

Finally, we exemplify the category-theoretic representation of the above interventions in $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$. The inverse $\mathcal{M}^{-1}$ exists by definition of the mixing matrix $\mathbf{M}$ itself. Specifically, it is equal to $\mathbf{I}_3 - \mathbf{C}$, where $\mathbf{C}$ is the matrix of causal coefficients corresponding to Equation (12). Using Lemma 1, in $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$ *(i)* the hard intervention corresponds to the natural transformation $\eta^{\mathcal{H}} = \langle \mathrm{Id}_{(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))}, \mathcal{H} = \mathcal{M}_{\mathcal{H}} \circ \mathcal{M}^{-1}\rangle$; and *(ii)* the soft intervention corresponds to the natural transformation $\eta^{\mathcal{S}} = \langle \mathrm{Id}_{(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), N(\mathbf{0}_3, \mathbf{I}_3))}, \mathcal{S} = \mathcal{M}_{\mathcal{S}} \circ \mathcal{M}^{-1}\rangle$.

C ON THE INVERTIBILITY ASSUMPTION

Here we illustrate that invertibility of $\mathcal{M}$ holds for the main classes of Markovian SCMs.

Additive Noise Model. In this case, considering n causal variables, the structural equations read as

$$X_i = f_i(\mathcal{P}_i) + Z_i, \quad \forall\, i \in [n]\,. \tag{17}$$

It is immediate to obtain that the components of $\mathcal{M}^{-1} : \mathcal{V} \to \mathcal{U}$ are $m_i^{-1} = X_i - f_i(\mathcal{P}_i), \forall\, i \in [n]$.

Post-nonlinear Model. In this case, the structural equations read as

$$X_i = g_i(f_i(\mathcal{P}_i) + Z_i), \quad \forall\, i \in [n]\,; \tag{18}$$

where g_i is invertible [36, 16]. It is immediate to obtain that the components of $\mathcal{M}^{-1} : \mathcal{V} \to \mathcal{U}$ are $m_i^{-1} = g_i^{-1}(X_i) - f_i(\mathcal{P}_i)$, $\forall\, i \in [n]$.

Location-scale Noise Models. In this case, the structural equations are [14]

$$X_i = f_i(\mathcal{P}_i) + h_i(\mathcal{P}_i)Z_i, \quad \forall\, i \in [n]\,; \tag{19}$$

where $f_i : \mathcal{V}_{\mathcal{P}_i} \to \mathbb{R}$ and $h_i : \mathcal{V}_{\mathcal{P}_i} \to \mathbb{R}_+$ is strictly positive. It is immediate to obtain that the components of $\mathcal{M}^{-1} : \mathcal{V} \to \mathcal{U}$ are $m_i^{-1} = (X_i - f_i(\mathcal{P}_i))\,/h_i(\mathcal{P}_i)$, $\forall\, i \in [n]$.

D PROOFS

Lemma 1 *[Interventions in* $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$ *] Given (i)* M *as in Definition 2; and (ii) the collection of measurable maps* $\mathcal{F}_{\mathcal{I}} := \{f_1^i, \ldots, f_n^i\}$*; an intervention in* $\mathsf{SCM}(\mathsf{I},\mathsf{Prob})$ *is a natural transformation* $\eta^{\mathcal{I}} := \langle \mathrm{Id}_{\mathsf{M}(I)}, \mathcal{I}\rangle$*, with* $\mathcal{I} = \mathcal{M}_{\mathcal{I}} \circ \mathcal{M}^{-1}$ *such that the following holds*

$$\begin{tikzcd} \mathsf{M}(I) \arrow[r, "\mathcal{M}"] \arrow[d, "\mathrm{Id}_{\mathsf{M}(I)}"'] & \mathsf{M}(I') \arrow[d, "\mathcal{I}"] \\ \mathsf{M}_{\mathcal{I}}(I) \arrow[r, "\mathcal{M}_{\mathcal{I}}"'] & \mathsf{M}_{\mathcal{H}}(I') \end{tikzcd} \implies \mathcal{M}_{\mathcal{I}} = \mathcal{I} \circ \mathcal{M}\,. \tag{(7)}$$

Proof. Let us start from the exogenous component. Since the intervention $\mathcal{I}$, that could be either soft or hard, acts only on the endogenous variables, the probability space corresponding to the exogenous part remains unchanged. That is, for the object $\mathsf{M}(I)$ we simply have

$$\eta_I^{\mathcal{I}} \coloneqq \mathrm{Id}_{\mathsf{M}(I)}\,. \tag{20}$$

This means that for any measure μ on the exogenous, the pushforward along $\eta_I^{\mathcal{I}}$ is exactly μ.

Now, let us consider the endogenous component and recall the collections $\mathcal{F}_{\mathcal{H}}$ and $\mathcal{F}_{\mathcal{S}}$ in Section 2. If the intervention is hard, than we represent it as the collection $\mathcal{F}_{\mathcal{I}} \coloneqq \mathcal{F}_{\mathcal{H}}$. If soft, with $\mathcal{F}_{\mathcal{I}} \coloneqq \mathcal{F}_{\mathcal{S}}$. The collection $\mathcal{F}_{\mathcal{I}}$ entails a new collection of mixing functions; specifically *(i)* $\mathcal{M}_{\mathcal{H}}$ for the hard intervention case, *(ii)* $\mathcal{M}_{\mathcal{S}}$ for the soft one (cf. Section 2). At this point, by invertibility of $\mathcal{M}$, Equation (7) trivially follows by setting the second component $\eta_{I'}^{\mathcal{I}}$ equal to $\mathcal{I} = \mathcal{M}_{\mathcal{I}} \circ \mathcal{M}^{-1}$. □

Lemma 2 *[Convex space of probability measures, $\langle \Delta_{(\mathcal{V},\,\Omega)}, cc_\lambda \rangle$] The convex space of probability measures, namely $\langle \Delta_{(\mathcal{V},\,\Omega)}, cc_\lambda \rangle$, is given by the set $\Delta_{(\mathcal{V},\,\Omega)}$ of probability measures χ on $(\mathcal{V},\,\Omega)$ together with a convex combination operation defined by*

$$cc_\lambda(\chi_1, \chi_2)(\mathcal{O}) \coloneqq \lambda\chi_1(\mathcal{O}) + \bar{\lambda}\chi_2(\mathcal{O})\,, \tag{8}$$

for all $\mathcal{O} \in \Omega$, with $\lambda \in [0, 1]$.

Proof. According to Def. 3.1 in [9], we have to demonstrate that $cc_\lambda(\chi_1, \chi_2)$ satisfies: *(i)* $cc_0(\chi_1, \chi_2) = \chi_2$, *(ii)* $cc_0(\chi_1, \chi_1) = \chi_1$; *(iii)* $cc_\lambda(\chi_1, \chi_2) = cc_{\bar{\lambda}}(\chi_2, \chi_1)$; and $cc_\lambda(cc_\mu(\chi_1, \chi_2), \chi_3) = cc_{\widetilde{\lambda}}(\chi_1, cc_{\widetilde{\mu}}(\chi_2, \chi_3))$ where:

$$\widetilde{\lambda} = \lambda\mu, \qquad \widetilde{\mu} = \begin{cases} \frac{\lambda\bar{\mu}}{\bar{\lambda}\mu}, & \text{if } \lambda\mu \neq 1\,, \\ \text{arbitrary}, & \text{if } \lambda = \mu = 1\,. \end{cases} \tag{21}$$

Since the convex set $\Delta_{(\mathcal{V},\Omega)}$ is a subset of a vector space, the previous properties follow from the axioms of vector space. □

Lemma 3 *The category* CSprob *has convex spaces $\langle \Delta_{(\mathcal{V},\,\Omega)}, cc_\lambda \rangle$ as objects, and affine measurable maps – i.e., measurable maps commuting with cc_λ – as morphisms.*

Proof. From Definition 3.2 in [9] we have that convex spaces as in Lemma 2 together with morphisms commuting with cc_λ form a category. Hence we have to demonstrate that affine measurable maps commute with cc_λ. Recall that a map $f : \Delta_{(\mathcal{V},\,\Omega)} \to \Delta_{(\mathcal{V}',\,\Omega')}$ is affine if for every $\chi_1,\, \chi_2 \in \Delta_{(\mathcal{V},\,\Omega)}$ and every $\lambda \in [0, 1]$ it satisfies $f\left(cc_\lambda(\chi_1, \chi_2)\right) = cc_\lambda\left(f(\chi_1), f(\chi_2)\right)$. Hence, by definition, any affine measurable map necessarily commutes with cc_λ. □

Theorem 1 *Every convex combination of probability measures corresponding to a causal knowledge* $\mathsf{CK}\,(\mathsf{M})$ *is a valid soft-interventional probability measure for* $\mathsf{CK}\,(\mathsf{M})$.

Proof. Consider $\mathsf{M} = \langle (\mathcal{U},\, \Upsilon, \zeta),\, (\mathcal{V},\, \Omega, \chi)\,, \mathcal{M} \rangle$, $\lambda \in [0, 1]$, and w.l.o.g. the probability measures $\chi_{\mathcal{S}_1}$ and $\chi_{\mathcal{S}_2}$ entailed by two (measurable) soft interventions $\eta^{\mathcal{S}_1}$ and $\eta^{\mathcal{S}_2}$ run on M. Hence, the convex combination

$$\chi_{\mathcal{S}_3} = cc_\lambda(\chi_{\mathcal{S}_1}, \chi_{\mathcal{S}_2}) = \lambda\chi_{\mathcal{S}_1} + \bar{\lambda}\chi_{\mathcal{S}_2} \tag{22}$$

is a proper soft-interventional probability measure entailed by $\mathcal{M}_{\mathcal{S}_3} = \{m_1^{\mathcal{S}_3}, \dots, m_n^{\mathcal{S}_3}\}$ according to Equation (6), where $m_i^{\mathcal{S}_3} = \lambda m_i^{\mathcal{S}_1} + \bar{\lambda} m_i^{\mathcal{S}_2}$. Indeed we have

$$\begin{aligned}
\lambda \chi_{\mathcal{S}_1} + \bar{\lambda} \chi_{\mathcal{S}_2} &= \bigtimes_{X_i \in \mathcal{X}} \Big(\lambda m_i^{\mathcal{S}_1} \left(\mu_i \left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i} \right) \right) + \\
&\qquad + \bar{\lambda} m_i^{\mathcal{S}_2} \left(\mu_i \left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i} \right) \right) \Big) \\
&= \bigtimes_{X_i \in \mathcal{X}} \left(\lambda m_i^{\mathcal{S}_1} + \bar{\lambda} m_i^{\mathcal{S}_2} \right) \left(\mu_i \left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i} \right) \right) \\
&= \bigtimes_{X_i \in \mathcal{X}} m_i^{\mathcal{S}_3} \left(\mu_i \left(\mathcal{U}_i \times \mathcal{U}^{\mathcal{A}_i} \right) \right) = \chi_{\mathcal{S}_3} \,.
\end{aligned} \tag{23}$$

Also, recall that a convex combination of measurable functions is still measurable. Starting from Equations (2), (5) and (7), it is straightforward to see that the same holds also when we consider observational and hard-interventional measures within the convex combination. □

Theorem 2 *The component $\alpha_{\mathcal{X}}$ of α within an IC $\boldsymbol{\alpha}$-abstraction commutes with cc_λ, thus is affine.*

Proof. Consider two SCMs M and M' related by an IC $\boldsymbol{\alpha}$-abstraction as in definition 6, and interventions $\mathcal{I}_1'$ and $\mathcal{I}_2\prime$ on M', either hard or soft, corresponding to interventions $\mathcal{I}_1$ and $\mathcal{I}_2$ on M. Then, given $\mathcal{I}_3'$ convex combination of $\mathcal{I}_1'$ and $\mathcal{I}_2\prime$, we have

$$\begin{aligned}
\chi'_{\mathcal{I}_3'} &= \alpha_{\mathcal{X}} \left(\chi_{\mathcal{I}_3} \right) \\
&\overset{(a)}{=} \alpha_{\mathcal{X}} \left(\lambda \chi_{\mathcal{I}_1} + \bar{\lambda} \chi_{\mathcal{I}_2} \right) \\
&\overset{(b)}{=} \lambda \chi'_{\mathcal{I}_1'} + \bar{\lambda} \chi'_{\mathcal{I}_2'} = \chi'_{\mathcal{I}_3'} \,.
\end{aligned} \tag{24}$$

where in *(a)* we apply Theorem 1, and in *(b)* the linearity of the pushforward. □

Theorem 3 *There exists an IC encoding functor $E : \mathsf{NI}(\mathsf{I}, \mathsf{Prob}) \to \mathsf{CSprob}$ mapping (i) each non-intervened $\mathsf{M} := \langle (\mathcal{U}, \Upsilon, \zeta), (\mathcal{V}, \Omega, \chi), \mathcal{M} \rangle$ to the convex spaces of probability measures $\Delta_{(\mathcal{V}, \Omega)}$, and (ii) an IC $\boldsymbol{\alpha}$-abstraction between M and M' with its endogenous component $\alpha_{\mathcal{X}}$.*

Proof. The mapping on objects is well-defined since Theorem 1 guarantees that convex combinations of probability measures associated with observational and interventional states of M are valid soft-interventional probability measures. On morphisms, Theorem 2 implies that $\alpha_{\mathcal{X}}$ is a proper morphism in CSprob. Interventional consistency is inherited because in $\mathsf{NI}(\mathsf{I}, \mathsf{Prob})$ every morphism is an IC $\boldsymbol{\alpha}$-abstraction. □

E FURTHER DISCUSSION AND EXEMPLIFICATION OF RCK

A Paradigm Shift. Our work provides a *novel conceptual foundation* for interpreting causality. Each subject infers causal knowledge from personal experience and can only see the world through its own perspective. Although the subjectiveness of causality has already been partially explored by Richens and Everitt [27] and, very recently, Bareinboim [1], we move further by making causality not only *subjective* but also *relational*. Consequently, asking for a unique "true" causal description of a system is an ill-posed question, as the very same notion of causality is inherently relative in our framework. However, it is important to highlight that the relativity of causality is different from the relativism (in its philosophical meaning) of causality: we do not undermine the meaning of things, we question its description as a monolithic object. In our setting, global causal traits of the system emerge only when there is local agreement throughout the network–the global sections of the sheaf of causal knowledge. This parallels how certain physical quantities in physics remain invariant across all reference frames, while other quantities remain inherently relative. But still, the relative treatment better fits the notion of physical reality we are able to describe. Consider a sudden drop in a company's stock price. Different agents (the subjects) in the financial system (the network sheaf) propose distinct explanations (their CK): equity analysts blame shifting market sentiment; management highlights an unfavorable earnings report; institutional investors worry about looming regulatory risks; retail shareholders point to negative social media coverage fueling panic selling. However, viewed within the broader network, the interplay of these perspectives (the RCK) discloses a nuanced mix of market sentiment, regulatory uncertainty, company-specific fundamentals, and social influence—no single narrative fully explains the drop. A unified description

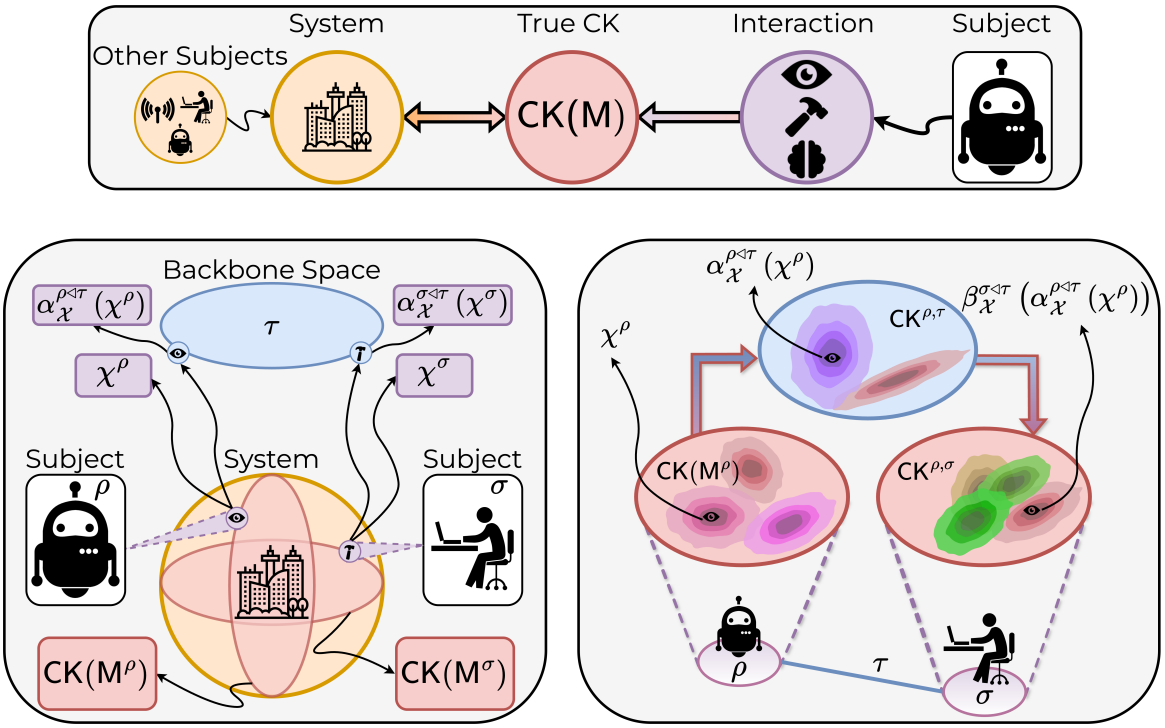


Figure 2: In the top figure, we depict the classical, non-relative approach to causality: a subject interacts with a system that is assumed to be completely describable by an underlying "true" SCM M. We refer to the (convex) space of all the interventional and observational probability measures entailed by M as the "true" Causal Knowledge $\mathsf{CK}(\mathsf{M})$. In this classical setting, the system and M are thus in a bijection, the other subjects are assumed to be part of the system, and the observer subject interacts with the system following (or inferring) $\mathsf{CK}(\mathsf{M})$. The *Relativity of Causal Knowledge* fundamentally challenges this paradigm by making causality subjective and relational, i.e., by breaking the bijection between the system and a "true" SCM, and isolating the subjects from the system while allowing them to interact with each other. The novel, core mathematical object implementing our relative paradigm is the *Network (Co)Sheaf of Causal Knowledge*, which can be informally thought of as a graph on whose nodes and edges are attached certain convex spaces, called *stalks*, that can interact through certain mappings, called *restriction and extension maps*. In our relative setting, as we show in the bottom left figure, the system can be pictured as a spherical planet, and the subjects as satellites orbiting around it. Each subject can only observe the planet from certain angles–its *perspective*–, positioning it on a specific orbit. The orbits represent then the subjective Causal Knowledge of the subjects–$\mathsf{CK}(\mathsf{M}^\rho)$ and $\mathsf{CK}(\mathsf{M}^\sigma)$–, i.e., the (convex) spaces of all the interventional and observational probability measures entailed by the subjective SCMs that the subjects use to describe the underlying system–M^ρ and M^σ. In sheaf jargon, each subject is a node, and each space $\mathsf{CK}(\mathsf{M}^\cdot)$ is a *node stalk*. As such, a point $\chi^\rho \in \mathsf{CK}(\mathsf{M}^\rho)$ on the orbit of a subject ρ is a *specific probability measure*. In sheaf jargon, a collection of points, one per each subject's orbit –$\{\chi^\rho, \chi^\sigma\}$– is a *0-cochain*. Two subjects ρ and σ can communicate if their SCMs M^ρ and M^σ admit *a shared abstraction*, i.e., a backbone space τ. In this case, ρ and σ can map their 0-cochain values χ^ρ and χ^σ to more abstract representations in the backbone space $\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}(\chi^\rho)$ and $\alpha_{\mathcal{X}}^{\sigma\triangleleft\tau}(\chi^\sigma)$ through surjective mappings $\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}$ and $\alpha_{\mathcal{X}}^{\sigma\triangleleft\tau}$, respectively. "More abstract" here means a coarse-grained but interventionally consistent representation, telling us essentially the same story about (subjective) cause-and-effect, but at different levels of detail. In sheaf jargon, ρ and σ are connected by an edge τ, the backbone space is an *edge stalk*, and the mappings are the *restriction maps*. Therefore, a collection of more abstract representations $\{\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}(\chi^\rho), \alpha_{\mathcal{X}}^{\sigma\triangleleft\tau}(\chi^\sigma)\}$ is a *1-cochain*. It is now clear that, in our relative setting, global traits of the underlying system only emerge by studying network-level *invariants*. Among them, "local agreement" of the subjects is particularly important. In sheaf jargon, local agreement refers to *global sections*, i.e., 0-cochains whose values are mapped, per each edge, to the same more abstract value– $\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}(\chi^\rho) = \alpha_{\mathcal{X}}^{\sigma\triangleleft\tau}(\chi^\sigma)$. As we show in the bottom right figure, a probability measure $\chi^\rho \in \mathsf{CK}(\mathsf{M}^\rho)$ in the CK of a subject ρ can be mapped to a (usually less informative) probability measure $\chi^{\rho,\sigma} \in \mathsf{CK}(\mathsf{M}^\rho)$ of another subject σ, connected to ρ through an edge τ, by first applying a restriction map $\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}$ and then an extension map $\beta_{\mathcal{X}}^{\sigma\triangleleft\tau}$, i.e., $\chi^{\rho,\sigma} = \beta_{\mathcal{X}}^{\sigma\triangleleft\tau}(\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}(\chi^\rho))$. The *Relative Causal Knowledge* $\mathsf{CK}^{\rho,\sigma}$ of a subject ρ from the perspective of σ is then the subspace $\mathsf{CK}^{\rho,\sigma} \subseteq \mathsf{CK}(\mathsf{M}^\sigma)$ being the image of $\beta_{\mathcal{X}}^{\sigma\triangleleft\tau}(\alpha_{\mathcal{X}}^{\rho\triangleleft\tau}(\cdot))$. In the general case, RCK is definable for any pair of subjects for which there exists a connecting path in the underlying network, not only for subjects directly connected by an edge (see Definition 11).

emerges only when there is local agreement across the entire network (the global sections of the sheaf of causal knowledge). See Figure 2 for a simple but comprehensive description of our framework.

The Role of Causal Abstractions. A subject cannot simply share all of its CK with some other subject because that would lead to an inherently inconsistent notion of perspective and, thus, of relativity. Think about physics: if we move from one reference frame to another, we don't just use the same measurements, but we transform them to make them consistent in the new frame. The notion of perspective is crucial in our framework because it is useful to model a variety of possible elements in a network of subjects: privacy or fairness constraints, maximizing mutual information, or "simply" modeling the impossibility of a human being to analyze a system if not through their own eyes. In practice, this translates into the uselessness of communicating CK whose distributions have different support on different random variables w.r.t. to the causal knowledge of the receiver subject, which would not know how to use it (see the toy example below). Therefore, the abstractions are a convenient yet rigorous way to model a "shared discourse space" among subjects to enable communication. In this sense, an abstraction is a backbone space: if the subjects operate on different random variables, abstraction and interventional consistency are arguably the best ways to provide them with a rigorous communication medium.

A Clarification on the Meaning of Cochains. A network comprises three subjects σ, ρ, and γ, each of them injecting into the network three soft-intervened probability measures from their own CK. These measures represent their current causal representation of the world. As stated in the main body and in Figure 2, the collection of these three probability measures forms a 0-cochain, i.e., a valuation of the node stalks of the network sheaf. In other words, the value of the 0-cochain at each node is just an object (a probability measure) of the node stalk (the node's CK). Via the restriction maps, the 0-cochain entails a 1-cochain on the edges, i.e., the collection of the abstracted probability measures. Although the subjects share a common CA, the abstracted probability measures might disagree on the edge stalk due to their different understanding of the world at that moment. Then, when a subject, say ρ, embeds onto its own node stalk, the abstract probability measure of subjects σ and γ from the edge stalks, it individuates different probability measures belonging to its CK. The latter represents the understanding of the world of σ and γ expressed in terms of the causal variables of subject ρ. Subject ρ can combine its understanding with those of σ and γ to accomplish its task, leveraging the closure over convex combinations of the CK (Theorem 1). A global section can then be seen as a 0-cochain whose values, for each pair of nodes connected by an edge, are mapped to the same 1-cochain value by the restriction maps.

A Toy Example. Suppose there is some underlying system and consider two subjects ρ and σ. The former observes some random variables X, Z, and Y, and knows that these random variables are related by a structural causal model (SCM) with DAG $X \to Z \to Y$. In particular, ρ knows the joint distribution of these variables and also their joint distribution under every possible intervention–that is, ρ's CK. This SCM is compatible with a more abstract SCM whose random variables are U and M, and that has DAG $U \to M$. "More abstract" here means that there exist surjective mappings porting X, Z and Y into U and M. The more abstract SCM is interventionally consistent, i.e., if we pick a variable in the initial SCM, set it to a specific value (*an intervention*), and do the corresponding intervention in the more abstract model, they both tell us essentially the same story about cause-and-effect, but at different levels of detail. Now, suppose σ observes variables P, W, T, L and knows they are related by an SCM with DAG $P \to W \to T \to L$. This second SCM is also compatible with the same more abstract SCM with DAG $U \to M$. Additionally, σ can access ρ's CK only through this shared abstract representation and vice versa. If σ and ρ are connected by an edge τ, then the Relative CK (RCK) of ρ from the perspective of σ is the result of first applying the restriction map to port the CK of ρ on τ, and then porting that more abstract realization of the CK of ρ to σ by applying the extension map. In the general case, RCK is definable for any pair of subjects for which there exists a connecting path in the underlying network, not only for subjects directly connected by an edge. Moreover, if more than one possible path connects two subjects, the RCK is path-dependent (see Definition 11).

Is RCK Needed? As a little mental exercise to highlight the need for our relative treatment of causality, let us forget for a moment our framework and suppose that there exists a protocol for communicating the entire CK from ρ to σ and vice versa. In this case, σ receives a certain state (*a cochain value*) of the CK over $\{X, Y, Z\}$ from ρ but, in order to utilize this information, σ needs to translate this information onto $\{P, W, T, L\}$. As such, any protocol needs a translation step. The need for this translation step is an informal necessary condition for proving the need of a relative description of causality, beyond our personal philosophical perspective. Consequently, any protocol that uses causal abstraction as the translation step would be a specific instance of our framework. This statement holds true even in more nuanced corner cases: consider again ρ and σ but now suppose that both share the same causal model over the same set of random variables, say $\{X, Y, Z\}$. Consider again the protocol above. This time σ receives from ρ a certain state of its own causal model according to the perspective of ρ in that context. Then, σ can directly combine the ρ 's perspective with its own and accomplish the assigned task. This situation is allowed within our proposed framework. In particular, it is equivalent to set *(i)* the causal abstraction as a causal model over three random variables $\{U, V, W\}$ entailing rotated versions of the probability measures of both ρ

and σ, *(ii)* the restriction maps as rotation matrices, and *(iii)* the extension maps as the transpose of such rotation matrices. The trivial example is when the restrictions and extensions are rotations with a null angle, i.e., the identity.

A Practical and Brief Example in Decentralized Agentic AI. In the context of AI, think about the latent spaces of two autonomous AI agents: they cannot communicate their internal current causal representations (the current 0-cochain) without aligning the latent spaces, i.e., without mapping one latent space into the other. Assume that the latent space of the first agent is a realization of a convex space of probability measures over $\mathbb{R}^d$ and its current causal representation is a point in this space (its current 0-cochain value), and that the latent space of the second agent is a realization of a convex space of probability measures over $\mathbb{R}^n$, with $n \neq d$, and its current causal representation is a point in this space (its current 0-cochain value). The subjects clearly cannot compare or exchange their current causal representations without any additional alignment step. Consider, however, the simplest scenario: linear alignment, where both restriction and extension maps are represented as matrices. In this case, the alignment matrix can be viewed as the product of a restriction matrix and an extension matrix. Notably, this factorization aligns closely with the approach commonly employed in low-rank adaptation methods (e.g., LoRA [13]). Please notice that this is the case even if the agents communicate in natural language, as they would need to be able to explain to each other how to relate their internal causal representations. At this point, it should be clear as well that the internal current causal representation of one agent can be entirely transmitted to the other agent without any transformation only if the two latent spaces coincide. As a consequence, please notice that the restriction and extension maps are not only useful to allow subjects to communicate, but they can also encode how they should communicate (the perspective). This is why we say that they can be designed to enforce privacy, fairness, or whatever other criterion of interest.