Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)
Survey Track

Learning Structural Causal Models through Deep Generative Models: Methods, Guarantees, and Challenges

Audrey Poinsot[1,2], **Alessandro Leite**[2], **Nicolas Chesneau**[1], **Michèle Sébag**[2] and **Marc Schoenauer**[2]
[1]Ekimetrics, France
[2]TAU, LISN, INRIA Saclay, France
audrey.poinsot@ekimetrics.com, alessandro.leite@inria.fr

Abstract

This paper provides a comprehensive review of deep structural causal models (DSCMs), particularly focusing on their ability to answer counterfactual queries using observational data within known causal structures. It delves into the characteristics of DSCMs by analyzing the hypotheses, guarantees, and applications inherent to the underlying deep learning components and structural causal models, fostering a finer understanding of their capabilities and limitations in addressing different counterfactual queries. Furthermore, it highlights the challenges and open questions in the field of deep structural causal modeling. It sets the stages for researchers to identify future work directions and for practitioners to get an overview in order to find out the most appropriate methods for their needs.

1 Introduction

The extraction of general and scientific knowledge from extensive and complex datasets has increasingly become a critical expectation in various domains. As highlighted by Pearl [2009b], attaining a deeper understanding that transcends mere associations necessitates an exploration of causation. This is particularly true when seeking to answer "why" questions, which require an insight into the underlying data generating processes. Standard generative approaches typically model the joint density of a set of variables. In contrast, causal generative models emphasize distributions derived from causal interventions and counterfactual operations as classified by the Pearl Causal Hierarchy (PCH) [Pearl and Mackenzie, 2018; Bareinboim *et al.*, 2022]. The PCH comprises three cognitive layers, named association, intervention, and counterfactual, correlating with the cognitive processes of seeing, doing, and imagining [Pearl and Mackenzie, 2018]. A comprehensive understanding of these layers is imperative for addressing diverse causal queries. The first layer, $\mathcal{L}_1$, focuses on associations and factual data. The second one, $\mathcal{L}_2$, delves into intervention data, shedding light on the effects of specific manipulations (a.k.a interventions). The third and final layer, $\mathcal{L}_3$, encompasses counterfactuals, facilitating the exploration of hypothetical scenarios that might have ensued under different interventions, as opposed to actual occurrences. Causal inference tasks aims to derive a quantitative understanding at a higher-level hierarchy when provided with data solely from lower-level ones [Pearl and Mackenzie, 2018]. However, achieving such a goal is inherently challenging without additional assumptions, primarily because the indeterminacy of higher levels by the lower ones poses a fundamental obstacle to such inference efforts [Ibeling and Icard, 2020; Bareinboim *et al.*, 2022]. The so-called indeterminacy means that without making explicit assumptions about the higher level, it is impossible to answer higher-level questions from lower-level data. For instance, answering counterfactual questions requires one to model the influence of the unobserved variables [Bareinboim *et al.*, 2022]. Structural causal models (SCMs) enable one to do so (Section 2). An SCM $\mathcal{M}$ models a set of causal mechanisms $\mathcal{F}$ on the observed $\mathbf{X}$ and unobserved $\mathbf{U}$ variables in addition to distributions over the unobserved variables $P(\mathbf{U})$. Nevertheless, a complete SCM is rarely known in practice, which leads one to question how a counterfactual query ($\mathcal{L}_3$) can be evaluated using a set of data from the underlying levels (i.e., $\mathcal{L}_1$ and $\mathcal{L}_2$). In general, experimental data ($\mathcal{L}_2$) are difficult to access. Thus, in practice, one relies on observational data ($\mathcal{L}_1$) and the knowledge of the causal structure ($\mathcal{L}_2$), formulated as a causal graph [Pearl, 1995; Spirtes *et al.*, 2000].

The recent advancements in deep learning, particularly the development of deep generative models (DGMs), have spawned a multitude of SCMs, the so-called DSCMs [Pawlowski *et al.*, 2020], which rely on deep learning approaches to model causal mechanisms. DGMs enable one to build more expressive SCMs [Xia *et al.*, 2021] thanks to the universal approximation capabilities of neural networks [Hornik, 1991]. As a result, they have been used in various high-stakes domains, including healthcare and fairness, where identifiability guarantees (i.e., existence and uniqueness) of the sought model are fundamental. However, it is essential to acknowledge that different DSCMs may operate under diverse assumptions regarding the functional form of causal mechanisms, leading to distinct guarantees. Understanding these variations is pivotal for informed and effective application of DSCMs in different contexts.

This paper aims to clarify the scope and limitations of the existing DSCMs when provided a known causal graph to an-

swer counterfactual queries. Our primary goal is to make a tour of the state-of-the-art, enabling a practitioner to better identify their needs and the requisites to fulfill them. A secondary one is to list the open questions and challenges. Formally, the overarching question investigated by this paper can be articulated as follows. Given a known causal structure $\mathcal{G}^\star$ and observational data, what are the capabilities of existing DGMs in learning a SCM $\widehat{\mathcal{M}}$ that approximates the true but unknown SCM $\mathcal{M}^\star$? In other words, we are interested in their capability in answering counterfactual questions as if one had access to the true SCM $\mathcal{M}^\star$.

Contributions. The contributions of this paper are threefold. First, it categorizes the existing deep structural causal models (Section 3) following their assumptions and guarantees based solely on observational data and a known causal structure. This classification provides a comprehensive framework for understanding the landscape of deep structural causal modeling when aiming to answer counterfactual questions. Second, analyzing existing DSCM from a theoretical and a empirical viewpoint, this paper assists practitioners when selecting an appropriate DGM to learn an SCM (Section 4) in order to apply these models effectively. Finally, this paper highlights some open questions and challenges (Section 5), serving as suggestions for future research directions in the field. Particularly, this paper stands out as the first one to classify the methods to learn SCMs through DGM based on their hypotheses and guarantees. In contrast, existing surveys [Zhou *et al.*, 2023; Komanduri *et al.*, 2023; Kaddour *et al.*, 2022] primarily classify them according to their deep learning components.

2 Preliminaries of Causal Inference

In this work, we use capital letters for random variables, lowercase letters for their realizations, and bold letters to represent vectors.

2.1 Structural Causal Model

A **Structural Causal Model** [Pearl, 2009b] is a tuple $\mathcal{M} \coloneqq (\mathcal{F}, P(\boldsymbol{U}))$, where $\mathcal{F}$ comprises a set of d structural equations f_i, one for each random variable, $X_i \in \boldsymbol{X}$, $\mathcal{F} = \{X_i \coloneqq f_i(\boldsymbol{PA}(X_i), U_i)\}_{i=1}^d$. Each structural equation f_i computes each variable X_i, also called endogenous variable, from its parents $\boldsymbol{PA}(X_i) \subseteq \boldsymbol{X} \setminus \{X_i\}$, and an exogenous noise variable $U_i \in \boldsymbol{U}$. $P(\mathbf{U})$ is a strictly positive measure over the exogenous noises, $\boldsymbol{U}$. Whenever an unmeasured common cause exists for several endogenous variables, these variables are said to be confounded by a hidden confounder. If one assumes that no hidden confounder exists over the endogenous variables $\mathbf{X}$, the exogenous noise variables $\mathbf{U}$ are considered jointly independent. An SCM characterizes a unique distribution over the variables, $P_{\mathcal{M}}(\boldsymbol{X})$, called the entailed distribution. The conditional distribution of a variable given its parents, $P_{\mathcal{M}}(X_i|\boldsymbol{PA}(X_i))$, is called a causal kernel [Pearl, 2009b]. An SCM describes the so-called **causal graph** $\mathcal{G} \coloneqq (\mathcal{V}, \mathcal{E})$ which encodes the causal dependencies between the variables. Each node corresponds to an endogenous variable, and the set of directed edges represents the parent-child relationships; i.e., $\mathcal{V} = \boldsymbol{X}$ and $\mathcal{E} = \{\boldsymbol{PA}(X_i) \to X_i, \forall i \in [1, d]\}$. A causal graph is usually assumed to be a Directed Acyclic Graph (DAG).

One can use an SCM to answer intervention and counterfactual questions [Pearl, 2009a]. Given an SCM $\mathcal{M}$, an **intervention** consists in changing at least one of the structural equations, for instance, $X_k \coloneqq \widetilde{f}_k(\widetilde{\boldsymbol{PA}}(X_k), \widetilde{U}_k)$. This operation, denoted $\boldsymbol{do}(X_k \coloneqq \widetilde{f}_k(\widetilde{\boldsymbol{PA}}(X_k), \widetilde{U}_k))$, defines a new SCM $\widetilde{\mathcal{M}}$ whose entailed distribution is called the intervention distribution, $P_{\widetilde{\mathcal{M}}}(\boldsymbol{X}) = P_{\mathcal{M}}(\boldsymbol{X}|\boldsymbol{do}(X_k \coloneqq \widetilde{f}_k(\widetilde{\boldsymbol{PA}}(X_k), \widetilde{U}_k)))$. Given an SCM $\mathcal{M}$ and some factual realizations $\boldsymbol{y}$ of $\boldsymbol{Y} \subset \boldsymbol{X}$, a **counterfactual** consists in changing the exogenous noise distribution into the posterior distribution $P_{\boldsymbol{U}|\boldsymbol{Y}=\boldsymbol{y}}$. It defines a new SCM $\mathcal{M}'$, where $P_{\mathcal{M}'}(\boldsymbol{U}) = P_{\mathcal{M}}(\boldsymbol{U}|\boldsymbol{Y} = \boldsymbol{y})$, whose entailed distribution is called the counterfactual distribution. Counterfactuals generally also include an additional intervention. The counterfactual distribution is hence denoted by $P_{\mathcal{M}}^{\boldsymbol{Y}=\boldsymbol{y}}(\boldsymbol{X}|\boldsymbol{do}(X_k \coloneqq \widetilde{f}_k(\widetilde{\boldsymbol{PA}}(X_k), \widetilde{U}_k)))$. Counterfactuals can be derived through the following three-step procedure [Pearl, 2009b]: (a) *Abduction* to change the exogenous noise distribution, (b) *Action* to build the new SCM $\mathcal{M}'$, and (c) *Prediction* to estimate the counterfactual distribution given $\mathcal{M}'$.

2.2 Identifiability

Once a causal task is defined in terms of its causal quantities, it needs to be translated into statistical ones for estimation. The property of identification amounts to showing whether, under the appropriate set of assumptions, such a unique translation exists. There exist several types of identifiability. For instance, structure identifiability states whether the causal graph can be recovered, model identifiability whether the full causal model can be recovered, and, estimand identifiability whether a specific causal estimand, such as a treatment effect, can be estimated. Finally, **intervention and counterfactual identifiability** involves checking whether all intervention and counterfactual queries can be retrieved. Formally, given a class of causal models $\mathbb{M}$ (e.g., SCMs), an $\mathcal{L}_i$-query $Q(\mathcal{M})$ of a model $\mathcal{M} \in \mathbb{M}$ is identifiable from $\mathcal{L}_j$-data, $1 \leq j < i$, if for any pair of models $\mathcal{M}_1$ and $\mathcal{M}_2$ from $\mathbb{M}$, $Q(\mathcal{M}_1) = Q(\mathcal{M}_2)$ whenever $\mathcal{M}_1$ and $\mathcal{M}_2$ match in all $\mathcal{L}_j$ queries [Pearl, 2009b]. Let us note that model identifiability implies intervention and counterfactual identifiability.

3 Classification into SCM and DGM Classes

This section introduces a classification of DGMs designed to learn an approximate structural causal model $\widehat{\mathcal{M}}$ when provided a known causal graph $\mathcal{G}^\star$ and observational data. By abuse of language but without loss of generality, we refer to these methods as DSCMs, since they all model causal mechanisms with deep generative models. The purpose of the approximate structural causal model $\widehat{\mathcal{M}}$ is to answer counterfactual queries. In practical terms, a method aiming to answer counterfactual queries must effectively perform the abduction step by modeling the distribution of the exogenous variables given the realization of the factual data (i.e., $P(\boldsymbol{U} \mid \mathbf{X} = \mathbf{x})$ by using the causal mechanisms (Section 2.1). In addition, assessing whether $\mathcal{L}_3$-identifiability holds is paramount to build

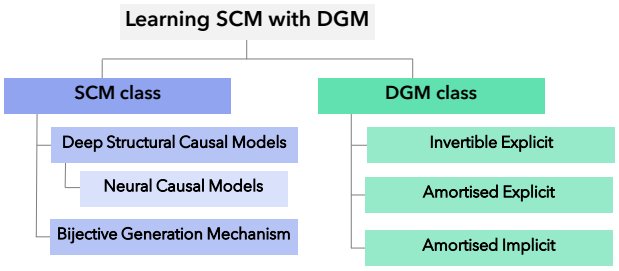

Figure 1: The two-dimension classification of DSCMs

trust in the estimated counterfactuals.

The classification, illustrated in Fig. 1, comprises two dimensions. The first dimension relates to the classes of SCM providing identifiability guarantees (Table 1), while the second one, originally proposed by Pawlowski *et al.* [2020], focuses on the characteristics of the DGM used to learn the SCM, restricting hence the tractability of the abduction step (Table 2). In this classification, DGM classes are exclusive while the SCM ones are not, as one can see in Table 3.

3.1 Classes of Deep Generative Models

DSCMs rely on deep learning components to capture and model the causal mechanisms. The choice of the network architecture can make it hard to compute the abduction step. Hence, a DSCM falls into one of the following classes introduced by Pawlowski *et al.* [2020]: invertible explicit, amortised explicit, and amortised implicit.

Invertible Explicit. Invertible explicit methods use invertible mechanisms with regard to the exogenous noises. The abduction step is hence directly tractable, computing the inverse transformation of the mechanism. Existing invertible explicit methods use conditional normalizing flows to represent the structural equations. In this case, each causal kernel can be computed as $P(X_i|\boldsymbol{PA}(X_i)) = P(U_i) \cdot |det\nabla_{U_i} f_i(\boldsymbol{PA}(X_i), U_i)|^{-1}|_{U_i=f_i^{-1}(U_i,\boldsymbol{PA}(X_i))}$ whenever f_i is assumed to be a diffeomorphic transformation. Examples of methods include **Causal-NF** [Javaloy *et al.*, 2023], **NF-DSCM** [Pawlowski *et al.*, 2020], **NCF** [Parafita and Vitrià, 2020], **CAREFL** [Khemakhem *et al.*, 2021] and, **NF-BGM** [Nasr-Esfahany *et al.*, 2023].

Amortised Explicit. Amortised explicit methods enlarge the considered causal mechanisms to non-invertible functions. By losing the invertibility of the structural equations, the abduction step can no longer be directly computed. Hence, these methods approximate the conditional exogenous noise distribution through auto-encoders. Notably, $f_i(\boldsymbol{PA}(X_i), U_i) = g_i(\boldsymbol{PA}(X_i), U_i)$ and $U_i = e_i(\boldsymbol{PA}(X_i), X_i)$ where g_i and e_i represent the decoder and the encoder[1]. Examples of learning methods belonging to this class include **iVGAE** [Zečević *et al.*, 2021], **VACA** [Sánchez-Martin *et al.*, 2022], and, **DCM** [Chao *et al.*, 2023].

[1]Note that the initial functional form enables "low-level" transformations: $f_i(\boldsymbol{PA}(X_i), U_i) = h_i(\boldsymbol{PA}(X_i), g_i(\boldsymbol{PA}(X_i), \mu_i), \epsilon_i)$ where g_i is a function of any form, h_i an invertible function, and $U_i = (\epsilon_i, \mu_i)$ a noise decomposition such that $P(U_i) = P(\epsilon_i)P(\mu_i)$.

Class	Mechanism	Identifiability Guarantees
DSCM	Neural network	-
NCM	Feedforward neural network	$\mathcal{L}_3$-id. (resp. $\mathcal{L}_2$) iif $\mathcal{L}_3$-id. (resp. $\mathcal{L}_2$) from the true SCM
BGM	Bijective noise	$\mathcal{L}_3$-id. for three settings* cf. Theorems 5.1, 5.2 and 5.3

*Markovian, Instrumental Variable, and Backdoor Criterion

Table 1: Identifiability guarantees of the classes of SCMs

Amortised Implicit. Instead of explicitly learning an encoder, one can also implicitly model it, for instance, using adversarial learning. Such approaches constitute the amortised implicit class. In this case, the abduction step must be conducted using sample rejection [Xia *et al.*, 2023]. In other words, for a given observation $\boldsymbol{Y} = \boldsymbol{y}$ and any SCM, one needs to sample and filter the exogenous noises not generating the desired observation $\boldsymbol{y}$; and then, uses the remained samples to estimate the counterfactual distribution corresponding to $P(U|\boldsymbol{Y} = \boldsymbol{y})$. This class includes the following learning methods: **SCM-VAE** [Komanduri *et al.*, 2022], **Causal-TGAN** [Wen *et al.*, 2022], **CausalGAN** [Kocaoglu *et al.*, 2018], **CFGAN** [Xu *et al.*, 2019], **DECAF** [van Breugel *et al.*, 2021], **WhatIfGAN** [Rahman and Kocaoglu, 2023], **CGN** [Sauer and Geiger, 2021], **DEAR** [Shen *et al.*, 2022], **GAN-NCM** and **MLE-NCM** [Xia *et al.*, 2023].

3.2 Classes of Structural Causal Models

As stated in the previous section, there are various deep generative approaches that can be used to learn an SCM. They all model DSCMs as they rely on deep learning components. However, they make different assumptions about the causal mechanisms leading to differences in the modelled DSCMs. In particular, the Bijective Generation Mechanisms [Nasr-Esfahany *et al.*, 2023] and Neural Causal Models [Xia *et al.*, 2021] are two classes of SCMs to distinguish them.

Bijective Generation Mechanisms

Bijective generation mechanism (BGM) models [Nasr-Esfahany *et al.*, 2023] are structural causal models where the causal mechanisms $f_i : \{\boldsymbol{PA}(X_i), U_i\} \mapsto X_i$ are bijective mappings from the exogenous noises to the endogenous variables. BGM models can also represent hidden confounders: if a hidden confounder has an impact on both X_i and X_j, their exogenous noises U_i and U_j are considered dependent. The bijective mapping enables to uniquely identify the exogenous noise U_i for each realization of the direct causes $\boldsymbol{PA}(X_i)$. As a result, all the invertible explicit methods are notably modeling BGM models. Nasr-Esfahany *et al.* [2023] provides sets of constraints on the causal mechanisms and probability distributions to guarantee $\mathcal{L}_3$-identifiability for three common causal structures: bivariate Markovian, Instrumental Variable, and Backdoor Criterion.

Neural Causal Models

The class of Neural Causal Models (NCMs) is introduced by [Xia *et al.*, 2021]. It defines a set of DSCMs whose causal

Class of DGM	Abduction step		
	Mechanism Inversion	Encoding	Sample Rejection
Invertible Explicit	✓	✓	✓
Amortised Explicit	✗	✓	✓
Amortised Implicit	✗	✗	✓

Table 2: Abduction steps for the classes of DGMs

mechanisms are feedforward neural networks. They involve all generative models except diffusion models (e.g., GANs, VAE, NF). Unlike feedforward neural networks, which process data in a direct, layer-by-layer flow without feedback, diffusion models employ an iterative and inherently stochastic mechanism. It involves gradually adding noise to the data over a series of steps to form a noisy distribution, and then systematically reversing this process through feedback mechanisms to generate new samples from the noisy distribution [Sohl-Dickstein *et al.*, 2015]. As a result, except DCM, all the methods mentioned previously model NCMs. NCMs are defined to consider hidden confounders. Besides the former exogenous noise (i.e., one per observed variable), NCMs consider additional exogenous noises, one for each hidden confounder impacting a subset C of variables: $X_i = f_i(PA(X_i), \{U_C | X_i \in C\})$. The distributions of the exogenous variables are set to $Unif(0,1)^2$. It is proven that an $\mathcal{L}_2$ [Xia *et al.*, 2021] or $\mathcal{L}_3$ [Xia *et al.*, 2023] query is identifiable from an NCM if and only if it is identifiable from the true SCM. Indeed, given that there might exist some hidden confounders, neither $\mathcal{L}_2$ nor $\mathcal{L}_3$ identifiability are guaranteed in general.

4 Analysis of Deep Structural Causal Models

This section analyses the existing DSCMs under the hypotheses, guarantees, evaluation, and applications. Section 4.1 describes the hypotheses and guarantees under five distinct characteristics (Table 3): (a) **causal structure knowledge**: the required type of knowledge on the causal structure, (b) **hidden confounding**: whether the method can handle hidden confounding, (c) **data assumption**: the hypotheses on the causal generating process of the data, (d) **abduction**: the capacity of the method to perform the abduction step, and, (e) the **additional guarantees** highlight supplemental guarantees to the ones provided by the SCM class. Section 4.2 focuses on the practical capabilities of the existing methods based on the experiments reported in the corresponding papers and the existing applications (Table 4).

4.1 Hypotheses and Guarantees

Causal Structure Knowledge. Some approaches only need a causal ordering as input. Indeed, a causal ordering is usually sufficient to build a fully connected DAG compatible, but non-minimal, with the true SCM. However, this is at the expense of additional training, computation time, and complexity. Other methods rely on the known causal graph. Note that CARFEL can also learn the causal structure on top of the causal mechanisms. However, in this case, model identification is only provided in the Markovian bivariate setting.

Hidden confounding. As said, unobserved common causes create spurious correlations that bias $\mathcal{L}_3$ estimations, and real-world applications tend to involve hidden confounding. However, only five methods can deal with them. CGN considers confounded variables in a specific structure configuration, while NCM, BGM, WhatIfGAN, and NCF can deal with any position of unobserved variables in the causal graph as long as this position is known. NeuralID [Xia *et al.*, 2023], an algorithm built on top of the NCM class, can automatically evaluate if a query is identifiable given a DAG with hidden confounders. To do so, two NCMs, constrained to preserve consistency with the observed data, are trained in parallel to maximize and minimize the target query. If the two resulting queries are identical, then the query is considered identifiable. Such a tool is crucial for sensitive applications.

Data Assumptions. Apart from the methods dealing with images (i.e., SCM-VAE, DEAR, and CGN), in general, no assumption is made about the data except those related to the selected DGM. SCM-VAE and DEAR make the hypothesis that one works with high dimensional data, knowing only a causal structure on some attributes characterizing the data, like labels of images. Both methods hence encode and decode the high-dimensional data into their attribute and model the SCM over the attributes. CGN is a computer vision model using a specific DAG over images, liking an image shape, texture, and background. Let us also highlight that WhatIfGAN has been designed to deal with an SCM with variables of different dimensions (e.g., scalars and images). Indeed, it trains the generators independently, when not confounded, to facilitate the convergence and enable one to utilize pre-trained models.

Abduction. On the one hand, as detailed in Section 3, all the explicit (i.e., invertible explicit and amortized explicit) methods are designed to automatically perform the abduction step. On the other hand, the implicit methods must rely on the sample rejection procedure. However, this abduction procedure is only implemented in NCM. As a result, if one wants to estimate counterfactuals with the other implicit methods, this procedure needs to be implemented on top of the existing code. Note that learning an encoder can replace the sample rejection procedure, though it is unclear how expensive and efficient it is compared to this procedure. Indeed, no comparison has been made yet.

Identifiability guarantees. Several results on identifiability exist. The most noticeable one is the NCM identifiability theorem, as the NCM class contains all the methods considered in this work except DCM; see Section 3. This result states that an $\mathcal{L}_2$ [Xia *et al.*, 2021] or $\mathcal{L}_3$ [Xia *et al.*, 2023] query is identifiable from a $\mathcal{G}^\star$-constrained NCM if and only if it is identifiable from the true SCM. This can be interpreted as follows: if the search space of the selected feedforward deep conditional generative models is expressive enough to represent the causal mechanisms, the identifiability property depends only on the causal structure and the available data to train the model. Then, a natural question arises: What if

[2]Note that, as the exogenous noise variable enters as input of the neural model, the assumption of uniform noise is not limiting.

Method	Classification		Additional Hypotheses				Additional Guarantees
	SCM class	DGM class*	Causal Structure	Hidden Confounder	Data Assumptions	Available Abduction	Identifiability, Expressivity, Bounds
NF-BGM	BGM, NCM	IE	DAG	✓	-	Inversion	-
NF-DSCM	BGM, NCM	IE	DAG	✗	f_i diffeomorphic	Inversion	-
GAN-NCM; MLE-NCM	NCM	AI	DAG	✓$^\sharp$	-	Sample Rejection	-
Causal-NF	BGM, NCM	IE	Ordering	✗	f_i diffeomorphic	Inversion	Model id. up to invertible transformation of U
NCF	BGM, NCM	IE	DAG	✓$^\sharp$	f_i diffeomorphic	Inversion	-
CARFEL	BGM, NCM	IE	DAG, $\emptyset$	✗	Affine autoregressive flow	Inversion	Model id. in bivariate case
iVGAE	NCM	AE	DAG	✗	-	✗	-
VACA	NCM	AE	DAG	✗	-	Encoding	$\mathcal{L}_2$-expressivity if the decoder is deep enough cf. Prop.2
DCM	-	AE	Ordering	✗	-	Encoding	$\mathcal{L}_3$-id. with error bounds cf. Corollary 1 & 2
SCM-VAE	NCM	AI	DAG	✗	Additive noise on attributes	✗	-
Causal-TGAN	NCM	AI	DAG	✗	-	✗	-
CausalGAN	NCM	AI	DAG	✗	-	✗	-
CFGAN	NCM	AI	DAG	✗	Categ. outcome & sensitive feature	✗	-
DECAF	NCM	AI	DAG	✗	-	✗	-
WhatIfGAN	NCM	AI	DAG	✓	-	✗	-
CGN	NCM	AI	DAG$^\tau$	✓$^\tau$	Image with attributes	✗	-
DEAR	NCM	AI	Ordering	✗	High-dimensional data with attributes	✗	Data to attribute encoder disentanglement

$^\sharp$A common cause is represented by an additional exogenous noise, $^\tau$Only a confounded trivariate DAG is considered
*Invertible Explicit (IE), Amortised Explicit (AE), and Amortised Implicit (AI)

Table 3: Hypotheses and guarantees of deep structural causal models. The classification (Figure 1) enables one to spot the identifiability results inherited by the SCM class and the compatible abduction step procedures. The methods are described in sections 3.1 and 3.2.

the assumptions made about the causal graph are wrong due to imperfect causal knowledge? We shall return to this point in Section 5.2. DCM also provides the same $\mathcal{L}_3$-identifiability result but in the absence of hidden confounders. In addition, DCM is the only method to furnish $\mathcal{L}_3$-error bounds based on the encode-decoder reconstruction error after training.

Discussion. Table 3 shows that, firstly, half of the method does not implement the abduction step automatically. However, one can rely on the sample rejection procedure [Xia *et al.*, 2023] to answer $\mathcal{L}_3$ queries. Secondly, only a few methods consider hidden confounders. Thirdly, on one hand, in the absence of hidden confounding, knowing a causal ordering and building the associated fully connected DAG is sufficient to compute counterfactuals. On the other hand, if there are hidden confounders, it is preferable to precisely locate them in the causal graph and use the NeuralID algorithm of Xia *et al.* [2023] to evaluate if the desired query is identifiable. Fourthly, for all the methods reviewed in this paper, any $\mathcal{L}_3$-query is identifiable if it is from the true SCM $\mathcal{M}^\star$.

4.2 Evaluation and Applications

Experimental evaluation. There is considerable heterogeneity in the evaluation of the methods: on the datasets, the PCH of the tasks, and the metrics. The PCH heterogeneity comes from the fact that only half of the methods automatically compute the abduction step. Those methods are hence evaluated on $\mathcal{L}_2$-tasks such as intervention estimation, disentanglement, invariant classification, or debiasing by intervening. fMRI [Thompson *et al.*, 2020], Color-MNIST [Kim *et al.*, 2019] are commonly used image datasets to evaluate intervention estimation. Pendulum [Yang *et al.*, 2021] and CelebA [Liu *et al.*, 2015] seem more suited for disentanglement while Morpho-MNIST [Castro *et al.*, 2019] can be used for counterfactual estimation. Regarding tabular data, bnlearn [Scutari, 2010] datasets (ASIA, Child, ALARM, Insurance) are often used to evaluate intervention estimation accuracy as they are extracted from causal bayesian networks. Datasets from the UCI repository [Dua and Graff, 2020] (Adult, Credit Approval, Census, News) are also popular for real data $\mathcal{L}_2$-experiments as causal graphs can be easily approximated from common sense knowledge. However, regarding $\mathcal{L}_3$-estimations, there is a high heterogeneity between the simulated SCMs. As a result, without a unified benchmark, it is a hard task to compare their capabilities to answer counterfactual questions. Moreover, the simulated SCMs are designed without multiple sources of randomness (e.g., DAG

Method	Dataset	PCH	DSCM Comparison	Applications
NF-BGM	Ellips generation simulations	$\mathcal{L}_3$	✗	Video streaming simulations for adaptive bitrate
NF-DSCM	Morpho-MNIST	$\mathcal{L}_3$	✗	Scientific discovery [Yu *et al.*, 2023], cybersecurity data generation [Agrawal *et al.*, 2024]
GAN-NCM; MLE-NCM	Simulated SCMs	$\mathcal{L}_3$	✓GAN, MLE NCM	-
Causal-NF	Simulated SCMs	$\mathcal{L}_3$	✓VACA, CARFEL	Counterfactual fairness & fair regularization of classifier
NCF	Salary simulations using a simulated SCM	$\mathcal{L}_3$	✗	Counterfactual fairness and explainability
CARFEL	4-dimentional polynomial simulated SCM, fMRI	$\mathcal{L}_2$ & $\mathcal{L}_3$	✗	-
iVGAE	ASIA	$\mathcal{L}_2$	✗	-
VACA	Simulated SCMs	$\mathcal{L}_3$	✗	Counterfactual fairness
DCM	Simulated SCMs, fMRI	$\mathcal{L}_3$	✓VACA, CARFEL	-
SCM-VAE	Pendulum, CelebA	$\mathcal{L}_2$	✗	-
Causal-TGAN	ASIA, Child, ALARM, Insurance; Adult, Census, News	$\mathcal{L}_1$	✗	In-domain data augmentation
CausalGAN	CelebA	$\mathcal{L}_2^\star$	✗	Out-of-domain data augmentation
CFGAN	Adult	$\mathcal{L}_2^\sharp$	✗	Fairness debiasing
DECAF	Adult, Credit Approval	$\mathcal{L}_2^\sharp$	✓CFGAN	Fairness debiasing
WhatIfGAN	Color-MNIST	$\mathcal{L}_2$	✓NCM	-
CGN	Color-MNIST; ImageNet [Deng *et al.*, 2009]	$\mathcal{L}_2^\tau$	✗	Out-of-domain data augmentation
DEAR	Pendulum, CelebA	$\mathcal{L}_2^\star$	✗	-

* Disentanglement, $^\sharp$ Fairness debiasing by intervention, $^\tau$ Invariant classification after intervention

Table 4: Existing evaluations and applications of DSCMs. The methods are described in Sections 3.1 and 3.2.

structure). This can introduce a bias in the evaluation, resulting in non-generalizable performance when facing real-world applications.

Applications. Three major goals have been considered in the literature: fairness, explainability, and robustness of machine learning models. Concretely, ImageCFGen[3] [Dash *et al.*, 2022], Causal-NF, NCF, and VACA were used for counterfactual fairness evaluation. Counterfactual fairness consists of assessing the existence of discrimination regarding a sensitive feature by answering the following counterfactual question: "Would the outcome have been the same had the person had a different sensitive feature value?". Once a model is assessed to be counterfactually unfair, counterfactuals can also be used to debiasing a model. For instance, ImageCFGen and Causal-NF have been used to regularize classifiers on unfair classifications by penalizing outcome variation over unfair counterfactuals. Focusing on data fairness instead of model fairness, DECAF and CFGAN have been designed to remove unfair biases from datasets learning the SCM modeling the data generation process and generating a new dataset intervening on the causal graph to break unfair causal paths. Regarding explainability, NCF has also been used for counterfactual explanations of black box image classifiers. Causal data augmentation has been a common approach to increase the robustness of machine learning methods for image classification. CausalTGAN performs in-domain augmentation by sampling data from the learned SCM without intervention. CausalGAN and CGN rely on out-of-domain data augmentation by generating data from an intervened SCM. The interventions are performed on the variables that change the image without changing the label to predict. Cai *et al.* [2024] also use out-of-domain data generation for adversarial attacks of image classifiers. Yu *et al.* [2023] use NF-DSCM for scientific discovery in the case of a Bioinformatics application. They use a counterfactual approach to determine which genetic correlations actually correspond to causal effects on Alzheimer and Glioblastoma illnesses. Finally, Agrawal *et al.* [2024] uses DSCMs to generate synthetic data to train and test cybersecurity models.

[3]ImageCFGEN [Dash *et al.*, 2022] uses an SCM learned with NF-DSCM [Pawlowski *et al.*, 2020].

5 Challenges and Future Directions

5.1 Lack of Evaluation

The lack of a benchmark comprises the major challenge one faces when comparing existing deep structural causal models, as underlined in Table 4. The methods must be evaluated using the same counterfactual distributions and metrics. In addition, they should be assessed regarding their data efficiency, computing time, and robustness to real-world challenges like selection bias, imbalanced data distributions, and imperfect causal knowledge. Furthermore, assessing the methods' robustness against potential misspecification of the causal graph is fundamental. To support a fair comparative evaluation, the community needs to build a referential SCM simulator. Indeed, instead of selecting specific SCMs, the simulator will

enable one to randomly generate DAGs, exogenous noises, and causal mechanisms to fully assess these methods' capabilities, paying attention to the various pitfalls involved in such automatic procedures [Reisach *et al.*, 2021]. Extending the evaluation methodology of Poinsot and Leite [2023] to counterfactuals could be a possible direction.

Finally, a first step towards such a comprehensible benchmarking of DSCMs could focus on the comparison of abduction procedures. Such a study is expected to assess the comparative merit of sample rejection with respect to learning an encoder.

5.2 From Identifiability to Partial Identifiability

Most of the causal generative methods presented in this survey assume knowing the actual causal graph. Yet, assuming complete knowledge of the causal structure is a strong assumption. Indeed, causal graphs are usually built either based on expert knowledge, which is sensitive to human biases, or using causal discovery algorithms, which still have numerous limitations nowadays [Vowels *et al.*, 2022; Faller *et al.*, 2024]. It is thus essential not to dismiss the possibility of relying on an erroneous causal graph, as this can have profound implications for the validity of causal inference results. Some other assumptions, such as additive noise or absence of selection bias, can also be questioned. An envisioned solution relies on moving towards partial identification, which consists of computing bounds for causal queries whenever point identification is impossible due to unsatisfied or un-testable assumptions. Some work has already been done to derive partial identifiability for discrete SCMs [Zaffalon *et al.*, 2020; Zhang *et al.*, 2022]. Hence, more research still needs to be carried out to provide such partial identifiability guarantees to DSCMs in general. However, some partial identifiability guarantees have been proven to be non-informative. Indeed, Melnychuk *et al.* [2023] showed, for instance, that, in general, the counterfactual outcome of (un)treated has non-informative bounds in the class of continuous bivariate SCMs. Even in these cases, it is still possible to carry out a sensitivity analysis to compute bounds on the causal query. Among the studied methods in this paper, only van Breugel *et al.* [2021] carried out a sensitivity analysis regarding the imperfect knowledge of the causal graph. However, the profound impact of any wrong assumption on the validity of causal query estimation should motivate the systematic use of sensitivity analysis for practitioners and theoreticians. Note that some work has already been done on sensitivity analysis under hidden confounding. For example, Frauen *et al.* [2024] developed a neural framework for generalized sensitivity analysis under unobserved confounding for $\mathcal{L}_2$-queries including conditional average treatment effect (CATE). Whereas, Schröder *et al.* [2023] studied the sensitivity of causal fairness ($\mathcal{L}_3$-queries) under unobserved confounding.

5.3 Applications: Opportunities and Challenges

Realistic Causal Simulations. As done by Agrawal *et al.* [2024], we believe DSCMs constitute a great opportunity for practitioners to generate synthetic data similar to real ones following a causal generating process. We particularly think that such approaches could be very useful to benchmark causal inference methods on various types of data coming from numerous applications (e.g., finance, marketing, healthcare), while having a ground truth. Nevertheless, the causal graph, assumed to be known, may, in fact, be misspecified. For this reason, it is crucial to ensure that it was validated by experts beforehand. Otherwise, we recommend limiting the use of such DSCMs to realistic synthetic data generation. Moreover, it is important to note that the existing methods have only been designed to deal with tabular data and images.

Sensitive Applications. As presented in Section 4.1, there exists an algorithm, NeuralID [Xia *et al.*, 2023], compatible with all the methods considered in this survey, automatically testing for query identifiability given the causal graph and the data. This constitutes a great opportunity for sensitive applications requiring identifiability guarantees. However, such an analysis still depends on the hypotheses about the causal graph. This emphasizes again the importance of sensitivity analyses. Moreover, it is also paramount for decision-makers to have access to uncertainty or error measures. DCM is the only method to provide error bounds. Therefore a major direction for forward research is related to uncertainty quantification extending the active research in numerical modeling [Rischel and Weichwald, 2021; Wang *et al.*, 2022] to the field of causal inference. In practice, the reconstruction error used in DCM could be extended to amortised explicit models.

6 Conclusion

This paper reviews and classifies deep structural causal models (DSCMs), examining them from the perspectives of their underlying classes of structural causal model (SCM) and deep generative model (DGM). This dual viewpoint enables one to spotlight the guarantees concerning the identifiability and the ability of these models to address counterfactual queries using only observational data and existing causal knowledge. The analyses reveal that the majority of these guarantees are rooted in the intrinsic properties of the deep learning architectures. Moreover, the studied methods have predominantly been evaluated through heterogeneous and straightforward datasets. This highlights the critical need for developing benchmark datasets that encapsulate the complexities encountered in various high-stakes applications, such as in the medical and financial domains. Lastly, our findings underscore the importance of integrating uncertainty quantification methods and partial counterfactual estimation methods with DSCM. Such a joint approach could significantly improve the resilience and practical utility of these models, especially in complex environments.

Ethical Statement

In this paper, we try to prevent malicious or naive use of deep structural causal models by being transparent about the limitations and guarantees of the various methods. Readers are also warned about the veracity of the made assumptions. To mitigate the risk of misuse of these models, it is emphasized that these methods should not be used to answer causal ques-

tions for decision-making if upstream experts have not validated them.

Acknowledgments

This research received partial funding from the European Commission under the Horizon 2020 program (TRUST-AI Project, Contract No. 952060) and (TAILOR, Contract No. 952215), and from the ANR under the France 2030 program, under the reference 23-PEIA-004.

References

[Agrawal *et al.*, 2024] Garima Agrawal, Amardeep Kaur, and Sowmya Myneni. A review of generative models in generating synthetic attack data for cybersecurity. *Electronics*, 2024.

[Bareinboim *et al.*, 2022] Elias Bareinboim, Juan D. Correa, Duligur Ibeling, and Thomas Icard. *On Pearl's Hierarchy and the Foundations of Causal Inference*. ACM, 2022.

[Cai *et al.*, 2024] Ruichu Cai, Yuxuan Zhu, Jie Qiao, Zefeng Liang, Furui Liu, and Zhifeng Hao. Where and how to attack? a causality-inspired recipe for generating counterfactual adversarial examples. In *AAAI*, 2024.

[Castro *et al.*, 2019] Daniel C. Castro, Jeremy Tan, Bernhard Kainz, Ender Konukoglu, and Ben Glocker. Morpho-MNIST: Quantitative assessment and diagnostics for representation learning. *JMLR*, 2019.

[Chao *et al.*, 2023] Patrick Chao, Patrick Blöbaum, and Shiva Prasad Kasiviswanathan. Interventional and counterfactual inference with diffusion models. *arxiv:2302.00860*, 2023.

[Dash *et al.*, 2022] Saloni Dash, Vineeth N Balasubramanian, and Amit Sharma. Evaluating and mitigating bias in image classifiers: A causal perspective using counterfactuals. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022.

[Deng *et al.*, 2009] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[Dua and Graff, 2020] Dheeru Dua and Casey Graff. UCI machine learning repository. archive.ics.uci.edu/ml/datasets, 2020.

[Faller *et al.*, 2024] Philipp M Faller, Leena C Vankadara, Atalanti A Mastakouri, Francesco Locatello, and Dominik Janzing. Self-compatibility: Evaluating causal discovery without ground truth. In *AIStats*, 2024.

[Frauen *et al.*, 2024] Dennis Frauen, Fergus Imrie, Alicia Curth, Valentyn Melnychuk, Stefan Feuerriegel, and Mihaela van der Schaar. A neural framework for generalized causal sensitivity analysis. In *ICLR*, 2024.

[Hornik, 1991] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 1991.

[Ibeling and Icard, 2020] Duligur Ibeling and Thomas Icard. Probabilistic reasoning across the causal hierarchy. In *AAAI*, 2020.

[Javaloy *et al.*, 2023] Adrián Javaloy, Pablo Sánchez-Martín, and Isabel Valera. Causal normalizing flows: from theory to practice. *NeurIPS*, 2023.

[Kaddour *et al.*, 2022] Jean Kaddour, Aengus Lynch, Qi Liu, Matt J. Kusner, and Ricardo Silva. Causal machine learning: A survey and open problems. *arXiv:2206.15475*, 2022.

[Khemakhem *et al.*, 2021] Ilyes Khemakhem, Ricardo Monti, Robert Leech, and Aapo Hyvarinen. Causal autoregressive flows. In *AIStats*, 2021.

[Kim *et al.*, 2019] Byungju Kim, Hyunwoo Kim, Kyungsu Kim, Sungjin Kim, and Junmo Kim. Learning not to learn: Training deep neural networks with biased data. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.

[Kocaoglu *et al.*, 2018] Murat Kocaoglu, Christopher Snyder, Alexandros G. Dimakis, and Sriram Vishwanath. CausalGAN: Learning causal implicit generative models with adversarial training. In *ICLR*, 2018.

[Komanduri *et al.*, 2022] Aneesh Komanduri, Yongkai Wu, Wen Huang, Feng Chen, and Xintao Wu. SCM-VAE: Learning identifiable causal representations via structural knowledge. In *IEEE International Conference on Big Data*, 2022.

[Komanduri *et al.*, 2023] Aneesh Komanduri, Xintao Wu, Yongkai Wu, and Feng Chen. From identifiable causal representations to controllable counterfactual generation: A survey on causal generative modeling. *arXiv:2310.11011*, 2023.

[Liu *et al.*, 2015] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *IEEE International Conference on Computer Vision*, 2015.

[Melnychuk *et al.*, 2023] Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Partial counterfactual identification of continuous outcomes with a curvature sensitivity model. In *NeurIPS*, 2023.

[Nasr-Esfahany *et al.*, 2023] Arash Nasr-Esfahany, Mohammad Alizadeh, and Devavrat Shah. Counterfactual identifiability of bijective causal models. In *ICML*, 2023.

[Parafita and Vitrià, 2020] Álvaro Parafita and Jordi Vitrià. Causal inference with deep causal graphs. *arXiv:2006.08380*, 2020.

[Pawlowski *et al.*, 2020] Nick Pawlowski, Daniel Coelho de Castro, and Ben Glocker. Deep structural causal models for tractable counterfactual inference. In *NeurIPS*, 2020.

[Pearl and Mackenzie, 2018] Judea Pearl and Dana Mackenzie. *The book of why: the new science of cause and effect*. Basic books, 2018.

[Pearl, 1995] Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 1995.

[Pearl, 2009a] Judea Pearl. Causal inference in statistics: an overview. *Statistics Surveys*, 2009.

[Pearl, 2009b] Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2nd edition, 2009.

[Poinsot and Leite, 2023] Audrey Poinsot and Alessandro Leite. A guide for practical use of ADMG causal data augmentation. In *ICLR 2023 Workshop on Pitfalls of limited data and computation for Trustworthy ML*, 2023.

[Rahman and Kocaoglu, 2023] Md Musfiqur Rahman and Murat Kocaoglu. Towards modular learning of deep causal generative models. In *ICML Workshop on Structured Probabilistic Inference & Generative Modeling*, 2023.

[Reisach *et al.*, 2021] Alexander G. Reisach, Christof Seiler, and Sebastian Weichwald. Beware of the Simulated DAG! Causal Discovery Benchmarks May Be Easy to Game. In *NeurIPS*, 2021.

[Rischel and Weichwald, 2021] Eigil F. Rischel and Sebastian Weichwald. Compositional abstraction error and a category of causal models. In *UAI*, 2021.

[Sánchez-Martin *et al.*, 2022] Pablo Sánchez-Martin, Miriam Rateike, and Isabel Valera. VACA: designing variational graph autoencoders for causal queries. In *AAAI*, 2022.

[Sauer and Geiger, 2021] Axel Sauer and Andreas Geiger. Counterfactual generative networks. In *ICLR*, 2021.

[Schröder *et al.*, 2023] Maresa Schröder, Dennis Frauen, and Stefan Feuerriegel. Causal fairness under unobserved confounding: a neural sensitivity framework. *ICLR*, 2023.

[Scutari, 2010] Marco Scutari. Learning bayesian networks with the bnlearn R package. In *Journal of Statistical Software*, 2010.

[Shen *et al.*, 2022] Xinwei Shen, Furui Liu, Hanze Dong, Qing Lian, Zhitang Chen, and Tong Zhang. Weakly supervised disentangled generative causal representation learning. *JMLR*, 2022.

[Sohl-Dickstein *et al.*, 2015] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *ICML*, 2015.

[Spirtes *et al.*, 2000] Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. The MIT Press, 2000.

[Thompson *et al.*, 2020] WH Thompson, R Nair, H Oya, O Esteban, JM Shine, CI Petkov, RA Poldrack, M Howard, and R Adolphs. Human es-fMRI resource: Concurrent deep-brain stimulation and whole-brain functional MRI. *bioRxiv*, 2020.

[van Breugel *et al.*, 2021] Boris van Breugel, Trent Kyono, Jeroen Berrevoets, and Mihaela van der Schaar. DECAF: Generating fair synthetic data using causally-aware generative networks. In *NeurIPS*, 2021.

[Vowels *et al.*, 2022] Matthew J. Vowels, Necati Cihan Camgoz, and Richard Bowden. D'ya like DAGs? a survey on structure learning and causal discovery. *ACM Computing Surveys*, 2022.

[Wang *et al.*, 2022] Benjie Wang, Matthew R Wicker, and Marta Kwiatkowska. Tractable uncertainty for structure learning. In *ICML*, 2022.

[Wen *et al.*, 2022] Bingyang Wen, Yupeng Cao, Fan Yang, Koduvayur Subbalakshmi, and Rajarathnam Chandramouli. Causal-TGAN: Modeling tabular data using causally-aware GAN. In *ICLR Workshop on Deep Generative Models for Highly Structured Data*, 2022.

[Xia *et al.*, 2021] Kevin Xia, Kai-Zhan Lee, Yoshua Bengio, and Elias Bareinboim. The causal-neural connection: Expressiveness, learnability, and inference. In *NeurIPS*, 2021.

[Xia *et al.*, 2023] Kevin Muyuan Xia, Yushu Pan, and Elias Bareinboim. Neural causal models for counterfactual identification and estimation. In *ICLR*, 2023.

[Xu *et al.*, 2019] Depeng Xu, Yongkai Wu, Shuhan Yuan, Lu Zhang, and Xintao Wu. Achieving causal fairness through generative adversarial networks. In *IJCAI*, 2019.

[Yang *et al.*, 2021] Mengyue Yang, Furui Liu, Zhitang Chen, Xinwei Shen, Jianye Hao, and Jun Wang. CausalVAE: Disentangled representation learning via neural structural causal models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.

[Yu *et al.*, 2023] Fanyang Yu, Rongguang Wang, Pratik Chaudhari, and Christos Davatzikos. Investigating causality between genotype and clinical phenotype in neurological disorders using structural causal model and normalizing flow. In *1st Workshop on Deep Generative Models for Health*, 2023.

[Zaffalon *et al.*, 2020] Marco Zaffalon, Alessandro Antonucci, and Rafael Cabañas. Structural causal models are (solvable by) credal networks. In *10th International Conference on Probabilistic Graphical Models*, 2020.

[Zečević *et al.*, 2021] Matej Zečević, Devendra Singh Dhami, Petar Veličković, and Kristian Kersting. Relating graph neural networks to structural causal models. *arXiv:2109.04173*, 2021.

[Zhang *et al.*, 2022] Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In *ICML*, 2022.

[Zhou *et al.*, 2023] Guanglin Zhou, Shaoan Xie, Guangyuan Hao, Shiming Chen, Biwei Huang, Xiwei Xu, Chen Wang, Liming Zhu, Lina Yao, and Kun Zhang. Emerging synergies in causality and deep generative models: A survey. *arXiv:2301.12351*, 2023.