Journal of Big Data

RESEARCH Open Access

Modeling psychological profiles in volleyball via mixed-type Bayesian networks

Maria Iannario[1†], Dae-Jin Lee[2†] and Manuele Leonelli[2*†]

[†]The authors Maria Iannario, Dae-Jin Leue, and Manuele Leonelli contributed equally to this work.

*Correspondence:
Manuele Leonelli
manuele.leonelli@ie.edu
[1]Department of Political Sciences, University of Naples Federico II, Via Rodinó 22, 80133 Napoli, Italy
[2]School of Science and Technology, IE University, Paseo de la Castellana 259, 28046 Madrid, Spain

Abstract

Psychological attributes rarely operate in isolation: coaches and practitioners reason about networks of related traits rather than single indicators. We analyze a new dataset of 164 female volleyball players from Italy's C and D leagues that combines standardized psychological profiling with background information. To learn directed relationships among mixed-type variables (ordinal questionnaire scores, categorical demographics, and continuous indicators), we introduce a hybrid structure learning approach that combines a latent Gaussian copula representation with a constraint-based skeleton and a score-based refinement to produce a single directed acyclic graph. We also study a bootstrap-aggregated variant to improve stability. In simulation studies spanning sample size, sparsity, and dimensionality, the proposed method achieves lower structural error and higher edge recovery than recent copula-based alternatives while maintaining high specificity. Applied to volleyball, the learned network organizes mental skills around goal setting and self-confidence, with emotional arousal linking motivation and anxiety, and places key personality traits, most notably neuroticism and extraversion, upstream of skill clusters. Scenario analyses quantify how improvements in specific skills propagate through the network to shift preparation, confidence, and self-esteem. Overall, the approach provides an interpretable, data-driven framework for profiling psychological traits in sport and for supporting decisions in athlete development.

Keywords Bayesian networks, Sports analytics, Sports psychology, Structural learning

Introduction

Elite and amateur sport increasingly recognize that psychological factors are central to performance. Meta-analytic reviews show that key psychological constructs (such as motivation, self-efficacy, affective states, and trained mental skills) are systematically related to competitive outcomes [1, 2]. Personality traits and self-esteem further shape how athletes respond to pressure and to coaching interventions [3, 4]. In parallel, the routine collection of standardized psychological questionnaires alongside demographic, training, and performance records has expanded, creating opportunities for data-driven modeling of mental skills in sport [5].

Beyond single-variable analyses, coaches and practitioners typically reason about networks of interrelated traits and contexts [4]. Recent work across both team and individual sports adopts an explicitly multivariate perspective that links personality, coping and mental skills, and performance, showing that effects frequently operate indirectly through intermediate psychological processes [6, 7]. Understanding how psychological attributes influence one another, and how they collectively relate to performance, requires models that represent multivariate dependence and conditional structure.

Bayesian networks (BNs) provide such a framework: they encode conditional dependencies in a directed acyclic graph (DAG) and support probabilistic reasoning under uncertainty [8]. In psychology and psychopathology, BNs are now widely used to assess interrelations among traits and symptoms and to probe admissible causal structures from observational data [9]. In sport, however, particularly for modeling athletes' psychological profiles, BN applications remain rather limited [10, 11]. More broadly across sports analytics, BN methods are still underrepresented relative to regression, SEM, and black-box machine-learning approaches, despite successful demonstrations in outcome prediction, performance assessment, injury prognosis, and tactical decision modeling [12–14].

Sport psychology datasets deriving from evaluation surveys integrate multiple data types, including Likert-scale questionnaire responses (ordinal variables), demographic or role-related information (categorical variables), and indicators of training load or performance, which are often measured on a continuous scale (i.e. [6, 7]). In much applied BN work, this heterogeneity is handled by discretizing continuous variables or collapsing categories before analysis. However, ad hoc binning can distort dependence structure, reduce statistical power, and yield graphs that change with the chosen cut points [15, 16]. These limitations have been noted in recent performance-analytics applications as well [13]. Accordingly, approaches that avoid ad hoc discretization are desirable.

There is growing interest in BNs for mixed-type variables, a setting historically less developed than purely discrete or Gaussian formulations. Such data are precisely the focus of the present work. Classical conditional linear Gaussian models allow discrete and continuous variables but impose parent–child restrictions that limit expressiveness [8]. More recent work leverages latent Gaussian copulas with rank-based estimation to model mixed marginals while retaining tractable structure learning [17–20]. Most proposals to date are *constraint-based*, recovering the graph through sequences of conditional-independence tests that add, remove, and orient edges.

Building on these developments, we adopt the "hybrid" strategy familiar from discrete and Gaussian BNs, where a constraint-based phase is followed by a score-based refinement whose search space is restricted by the first phase, e.g., Max–Min Hill-Climbing (MMHC) [21]. To our knowledge, we are the first to bring this approach to mixed data under a latent Gaussian copula: we extend latent–copula learners [18, 19] with a constrained score-based refinement. We term the resulting algorithm *latent MMHC*. Unlike purely constraint-based copula learners, which may leave some edge directions unresolved, latent MMHC returns a single oriented DAG. We also study a bootstrap-aggregated "stable" variant that repeats the procedure on resamples and aggregates edge-selection frequencies to produce a consensus DAG [22].

In this paper, we study the problem of learning interpretable dependency structures from mixed-type data in settings where ad hoc discretization is undesirable. We evaluate

the proposed approach through a comprehensive simulation study spanning sample size, sparsity, and dimensionality, focusing on structural accuracy, edge recovery, and computational cost. To support reproducibility and applied use, we provide an open-source `R` implementation of the proposed approach.

The methodological development is motivated by a common problem in applied sport psychology: routinely collected profiling data combine ordinal questionnaire responses, aggregated continuous scores, and categorical background variables, yet practitioners are interested in understanding how psychological traits conditionally relate to one another rather than in isolated associations. In the next section, we introduce a dataset of female volleyball players and employ it to illustrate the substantive goals and modeling challenges that motivate the proposed approach.

The paper is structured as follows. Section "Motivating application and data" introduces the volleyball dataset and discusses the substantive goals and modeling challenges that motivate the proposed framework. Section "Methodology" presents the mixed-type BN framework and introduces the latent MMHC algorithm and its stable variant. Section "Simulation study" benchmarks the proposed methods across sample sizes, sparsity levels, and dimensions. Section "Modeling psychological traits of volleyball athletes" applies the approach to the volleyball dataset to derive interpretable, coaching-relevant insights. Section "Conclusions" summarizes the findings and outlines avenues for future work.

Motivating application and data

Psychological profiling data in volleyball

To illustrate the applied setting that motivates the methodological framework developed in this paper, we consider a dataset of competitive volleyball athletes that combines psychological profiling with background information. The data consist of 164 female players competing in the C and D leagues of the Campania region in Italy, collected in 2020 within the Laboratory for Statistical Data Analysis (LASt) at the University of Naples Federico II. Alongside demographic and sport-related variables (e.g., age group, playing role, weekly training volume, competition level), the study collected standardized psychological measures commonly used in applied sport psychology. Due to the limited sample size, playing positions were aggregated into two functional groups: hitters (middle blocker, outside hitter, opposite hitter) and non-hitters (setter, libero).

Personality traits were assessed through the NEO Five-Factor Inventory [23], a 15-item short form that evaluates the Big Five dimensions: extraversion, agreeableness, conscientiousness, neuroticism (emotional stability), and openness. Mental skills were measured using the sport performance psychological inventory (IPPS-48) [24], which covers eight psychological skill areas: self-talk, goal setting, self-confidence, emotional arousal control, cognitive anxiety, concentration disruption, mental practice, and race preparation. Finally, self-esteem was evaluated with the Rosenberg Self-Esteem Scale [25], a 10-item uni-dimensional scale capturing global self-worth.

For each construct, we computed the average of the underlying item responses, yielding continuous scores that reflect the latent trait rather than individual ordinal responses. This aggregation is standard practice in psychometrics and reflects the assumption that the latent constructs are inherently continuous. The resulting variables, however, still retain features of their ordinal measurement origin, such as bounded

support and asymmetric distributions. The full list of variables is given in Table 1 and the empirical distributions of the continuous variables are reported in Appendix A.

Together with demographic and experience variables, these measures form a heterogeneous dataset comprising continuous, ordinal, and binary variables, which is representative of routinely collected psychological profiling data in sport.

Modeling challenges and methodological implications

The volleyball dataset described above highlights a set of modeling challenges that are typical in applied sport psychology and, more broadly, in behavioral and social sciences.

Table 1 List of variables used in the BN analysis of volleyball athletes, with descriptions and states

Group	Variable	Description	States/Range
Sports	league_division	Competition level of player's team	C, D
	num_competitions	Number of official competitions played	None, 1–3, 4–10, 11–20, >20
	player_position	Volleyball playing position	Non_Hitter, Hitter
	plays_other_sports	Whether the player practices other sports	No, Yes
	referee_experience	Whether the player has experience as a referee	No, Yes
	weekly_training_hours	Weekly training volume	<3h, 3–6h, 6–9h, >9h
Lifestyle	age_group	Age group of respondent	18–19, 20–24, 25–29, 30+
	drinks_alcohol	Alcohol consumption status	No, Yes
	in_relationship	Whether the player is in a relationship	No, Yes
	working_status	Main life activity	Not_Working, Working
	smokes	Smoking status	No, Yes
Big five	agreeableness	Compassion, helpfulness, and cooperation	Continuous average, original range 1–5
	conscientiousness	Self-control, diligence, attention to detail	Continuous average, original range 1–5
	extraversion	Boldness, energy, and social interactivity	Continuous average, original range 1–5
	neuroticism	Moodiness, irritability, and emotional instability	Continuous average, original range 1–5
	openness	Curiosity, creativity, and openness to new ideas	Continuous average, original range 1–5
Self-esteem	self_esteem	Global evaluation of self-worth	Continuous average, original range 1–4
IPPS-48	concentration_issues	Difficulties in maintaining attentional focus	Continuous average, original range 1–6
	emotional_arousal	Regulation of emotional activation during play	Continuous average, original range 1–6
	goal_setting	Tendency to define and pursue performance goals	Continuous average, original range 1–6
	match_preparation	Strategies used before competition for mental readiness	Continuous average, original range 1–6
	mental_practice	Use of imagery or mental rehearsal strategies	Continuous average, original range 1–6
	self_confidence	Belief in own performance capacity	Continuous average, original range 1–6
	self_talk	Use of internal or external speech for focus	Continuous average, original range 1–6
	worry	Cognitive and emotional concern during performance	Continuous average, original range 1–6

Continuous variables correspond to averaged psychometric scores and retain their original ordinal range

First, the data comprise variables of heterogeneous measurement types, including categorical background characteristics, ordinal questionnaire responses, and continuous scores obtained by aggregating multiple ordinal items. Second, the available sample size is moderate relative to the number of variables, a common situation in applied settings where data collection is resource-intensive and target populations are specialized.

A common strategy in applied BN modeling is to discretize continuous variables and proceed with a fully discrete analysis. While this approach simplifies estimation, it can substantially increase model complexity. In discrete BNs, the number of free parameters grows rapidly with both the number of parents and the number of categories per variable, leading to highly parameterized models even for moderately sized graphs. In small to medium samples, this can result in unstable structure learning, sensitivity to arbitrary choices of cut points, and the inclusion of spurious dependencies [15, 16]. Moreover, when data are sparse relative to the size of the conditional probability tables, parameter estimates become strongly influenced by the choice of prior or smoothing scheme (e.g., Laplace or Dirichlet priors), so that inferred probabilities may reflect prior assumptions as much as, or more than, empirical evidence.

To illustrate this issue in a setting closely aligned with our application, we conducted a simple experiment based on the variables in the volleyball dataset. We considered the full set of sport, lifestyle, and psychological variables reported in Table 1, imposing basic structural constraints consistent with domain knowledge. Specifically, we prevented psychological traits (Big Five, self-esteem, and IPPS-48 scores) from acting as parents of sport and lifestyle variables. Within this constrained space, we generated random DAGs with varying edge probabilities and, for each graph, counted the number of free parameters required under two modeling strategies: a fully discrete specification using the original categorical ranges of all variables, and a mixed-data formulation based on latent Gaussian representations of the psychological scores, as commonly adopted in recent BN literature.

Table 2 reports the median number of free parameters as a function of graph density under these two formulations. Even for relatively sparse structures, the fully discrete specification leads to a dramatic increase in the number of parameters, whereas models that retain continuous representations for psychological scores remain substantially more parsimonious. This gap widens rapidly as graph density increases, exacerbating the risk of overfitting and unstable estimation in finite samples.

Table 2 Median number of free parameters as a function of edge probability for latent Gaussian and fully discrete Bayesian networks, computed over randomly generated DAGs on the volleyball dataset variables

Edge probability	Latent Gaussian BN	Discrete BN
0.05	40	496
0.10	55	1,964
0.15	71	14,408
0.20	84	55,954
0.25	102	185,110
0.30	116	969,489
0.35	131	6,183,270
0.40	146	21,509,442
0.45	158	78,031,660
0.50	177	277,928,138

These considerations motivate the need for structure learning approaches that can accommodate mixed-type data without relying on ad hoc discretization, while retaining interpretability and scalability. In the next section, we introduce a hybrid structure learning framework that is designed to address these challenges.

Methodology

This section introduces the modeling framework and learning procedure used throughout the paper. We first review DAG-based BNs and standard structure-learning strategies, then describe the latent Gaussian copula representation that enables mixed-type learning, and finally present the proposed latent MMHC algorithm and its stable bootstrap variant.

Graphical models and directed acyclic graphs

Let $\mathrm{X} = (X_1, \ldots, X_p)$ be a vector of random variables. A *directed acyclic graph* (DAG) $\mathcal{G} = (V, E)$ consists of a set of nodes $V = \{1, \ldots, p\}$ and a set of directed edges $E \subseteq V \times V$ with no directed cycles. Each node $j \in V$ corresponds to a random variable X_j, and an edge $(i, j) \in E$ indicates a directed link from *i* to *j*, so that X_i is a direct parent of X_j.

The DAG encodes a factorization of the joint distribution *P* of X according to the local Markov property [8]:

$$P(\mathrm{X}) = \prod_{j=1}^{p} P(X_j \mid \mathrm{X}_{\mathrm{pa}(j)}), \tag{1}$$

where $\mathrm{pa}(j)$ denotes the set of indices corresponding to the parents of node *j* in $\mathcal{G}$. The presence of a directed edge (*i*, *j*) therefore indicates that X_i enters directly into the conditional distribution of X_j through conditioning on its parents. The overall collection of conditional independence relations encoded by the graph is not specified edge by edge, but emerges from the global structure of the DAG. By breaking down the joint distribution into local components, the DAG provides a compact and interpretable representation of potentially high-dimensional dependencies. In our application, the learned parent sets $\mathrm{pa}(j)$ identify which background and psychological variables directly predict each trait in the conditional sense of Eq. (1), providing an interpretable map of conditional relationships among athlete-profile measures.

The set of conditional independencies encoded by the DAG can be formally characterized using the concept of *d-separation* [8]. Let π be a path between nodes *i* and *j* in the graph. The path π is said to be *blocked* by a set of nodes $S \subset V$ if there exists a node *k* on the path such that one of the following conditions holds:

1. The path passes through a chain or fork structure $X_u \to X_k \to X_v$ or $X_u \leftarrow X_k \to X_v$, and $X_k \in S$.
2. The path passes through a collider structure $X_u \to X_k \leftarrow X_v$, and neither X_k nor any of its descendants belong to *S*.

If all paths between X_i and X_j are blocked by *S*, we say that X_i and X_j are *d-separated* given *S* in $\mathcal{G}$, and denote this by $X_i \perp\!\!\!\perp X_j \mid \mathrm{X}_S$. The global Markov property, stating that all d-separation statements in the graph correspond to conditional independencies in

the distribution, is equivalent to the local Markov property expressed in Eq. (1) under standard assumptions [8]. These two characterizations offer distinct but equivalent perspectives on the dependency structure encoded by the DAG.

In practice, several different DAGs can encode the same set of conditional independencies. This motivates the notion of *Markov equivalence*: two DAGs are said to be Markov equivalent if they entail the same set of d-separation statements [8]. A fundamental result establishes that two DAGs are Markov equivalent if and only if they have the same skeleton (i.e., the undirected graph obtained by removing edge directions) and the same set of colliders. The collection of all DAGs that are Markov equivalent forms a *Markov equivalence class*, which can be uniquely represented by a *completed partially directed acyclic graph* (CPDAG) [26]. A CPDAG is a mixed graph in which an edge is directed if its orientation is shared across all DAGs in the equivalence class, and undirected if at least one DAG assigns each possible orientation. Fig. 1 illustrates these concepts.

Learning graphical models from data

Learning a DAG is a central task in probabilistic modeling when the dependence structure among a collection of variables is unknown and must be inferred from data. Given a dataset $\mathcal{D} = \{\mathrm{x}^{(1)}, \ldots, \mathrm{x}^{(n)}\}$ consisting of n independent observations of a random vector X, the goal is to recover a DAG $\mathcal{G}$ whose structure encodes a set of conditional independence relations that are compatible with the joint distribution of X. In practice, this means finding a graph such that the conditional independencies implied by the DAG closely match those supported by the observed data.

Three broad classes of algorithms have been developed to tackle this problem: constraint-based, score-based, and hybrid methods.

Constraint-based methods rely on performing a series of conditional independence tests among the observed variables. These methods use the principle that conditional independencies in the distribution can be mapped to d-separation statements in a DAG under appropriate modeling assumptions. A prototypical example is the PC algorithm [27], which starts from a fully connected undirected graph and iteratively removes edges between variables found to be conditionally independent given some subset of the remaining variables. Once the skeleton is estimated, a set of orientation rules is applied to identify colliders and propagate edge directions without introducing cycles or new colliders. The output is typically a CPDAG representing the Markov equivalence class of DAGs consistent with the observed conditional independencies.

Score-based methods, in contrast, treat the problem of structure learning as an optimization task. A scoring function is defined over the space of DAGs, quantifying how well each graph explains the data according to a chosen statistical criterion, such as the Bayesian Information Criterion (BIC). The graph with the highest score is typically sought using a greedy or heuristic search strategy (e.g. tabu search), where each step considers local modifications to the current graph by adding, deleting, or reversing a

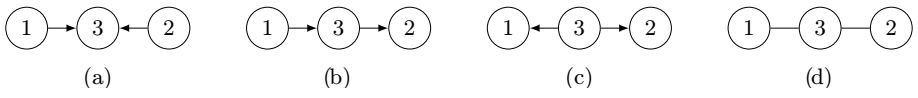

Fig. 1 Illustration of skeletons, colliders, and Markov equivalence. DAGs (**a**), (**b**), and (**c**) share the same skeleton $1 - 3 - 2$, but only DAG (**a**) contains the collider $1 \to 3 \leftarrow 2$. DAGs (**b**) and (**c**) are Markov equivalent and belong to the same equivalence class, represented by the CPDAG in panel (**d**), where undirected edges reflect ambiguous orientations

single edge while maintaining acyclicity [28]. These methods always return a single DAG that maximizes the chosen score over the explored search space.

Hybrid approaches combine features from both constraint-based and score-based methods. One widely used example is the Max-Min Hill-Climbing (MMHC) algorithm [21]. In its first phase, a constraint-based procedure is used to estimate the skeleton of the graph. In the second phase, a score-based search is performed, where edge additions, deletions, or reversals are only allowed if they are consistent with the previously identified skeleton. This is typically enforced by defining a *blacklist* that prohibits edges not present in the skeleton from being considered during the search.

No single approach consistently outperforms the others; their relative performance depends on factors such as sample size, dimensionality, and distributional properties of the data [29].

Structure learning in the Gaussian case

Gaussian BNs [8] provide a natural and analytically tractable setting for structure learning, forming the basis for methods applicable to more general or mixed data types. Under this model, each variable is expressed as a linear function of its parents in the DAG, plus an independent Gaussian noise term. That is, for each node *j* in the graph,

$$X_j = \sum_{k \in \mathrm{pa}(j)} \beta_{jk} X_k + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \sigma_j^2), \tag{2}$$

where the residuals $(\varepsilon_1, \ldots, \varepsilon_p)$ are mutually independent. This formulation defines a structural equation model (SEM), in which the regression coefficients β_{jk} encode the edge structure of the DAG.

Let L be a $p \times p$ lower triangular matrix with unit diagonal, such that $L_{uv} \neq 0$ if and only if $u \to v$ in $\mathcal{G}$, and let D be a diagonal matrix with entries σ_j^2. The SEM system can then be compactly written as $L^\top \mathrm{X} = \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}_p(0, D)$. It follows that the joint distribution of X is multivariate normal, with covariance matrix $\Sigma = L^{-\top} D L^{-1}$ and precision matrix $\Omega = \Sigma^{-1} = L D^{-1} L^\top$ [8]. This decomposition makes explicit how the DAG structure encoded in L determines the joint distribution, and is particularly useful for efficient score evaluation in structure learning.

In score-based learning of Gaussian BNs, the structure is inferred by optimizing a likelihood-based criterion. To simplify notation, we assume without loss of generality that the data are standardized, with $\mathbb{E}[X_j] = 0$ and $\mathrm{Var}(X_j) = 1$ for all *j*. Under this setup, and adopting a SEM perspective, the log-likelihood can be decomposed into a sum of nodewise contributions:

$$\ell(\mathcal{G}) = -\frac{n}{2} \sum_{j=1}^{p} \log(\widehat{\sigma}_j^2), \tag{3}$$

where $\widehat{\sigma}_j^2$ is the residual variance from the regression of X_j on its parents.

In constraint-based approaches, structure learning proceeds by testing conditional independencies implied by the Gaussian distribution. Under multivariate normality, the conditional independence $X_i \perp\!\!\!\perp X_j \mid \mathrm{X}_S$ is equivalent to $\rho_{ij \cdot S} = 0$, where $S \subset V \setminus \{i, j\}$ denotes a conditioning set of variables and $\rho_{ij \cdot S}$ is the partial correlation between X_i and X_j given X_S. This allows conditional independence testing to be

performed using standard tools from Gaussian theory, based on estimates of partial correlations and their sampling distribution [30]. This testing procedure forms the basis of the Gaussian version of the PC algorithm and related constraint-based methods.

Mixed data and latent Gaussian copula models

In many real-world applications, datasets consist of a vector of observed variables $\mathrm{X} = (X_1, \ldots, X_p)$, which may include a combination of continuous, ordinal, binary, and count components. While BNs are widely used to model such dependencies, structure learning from mixed data remains comparatively underexplored. The most common approach is the *conditional linear Gaussian* model [8], which assumes Gaussian linear regressions for continuous variables and multinomial distributions for discrete ones. However, it imposes restrictive constraints on the structure of the graph, notably disallowing continuous parents of discrete nodes, and can struggle to represent complex dependencies.

To address these limitations, recent approaches model mixed data using *latent Gaussian copulas* [18–20, 31]. In this framework, each observed variable is linked to an underlying Gaussian latent variable, and the DAG is learned in the latent space. This choice enables the use of Gaussian structure-learning methods, while avoiding ad hoc discretization of the continuous psychological measures. Specifically, for each observed variable X_j, we assume the existence of a latent variable $Z_j \sim \mathcal{N}(0, 1)$ such that

$$Z_j = \Phi^{-1}(F_j(X_j)),$$

where F_j is the marginal cumulative distribution function of X_j, and Φ^{-1} is the quantile function of the standard normal. The vector $\mathrm{Z} = (Z_1, \ldots, Z_p)$ is then modeled as a multivariate Gaussian:

$$\mathrm{Z} \sim \mathcal{N}_p(0, \Sigma_Z),$$

where Σ_Z is a correlation matrix that encodes the dependence structure among the latent variables. This construction induces a Gaussian copula over the observed variables X, while allowing arbitrary marginal distributions for each component. Alternative approaches not relying on a latent Gaussian layer, such as ordinal regression, conditional likelihoods, or structured local models, have also been explored in the literature [32, 33].

The key idea in latent Gaussian copula models is that the dependence structure among the observed variables X is governed by the latent correlation matrix Σ_Z defining the joint distribution of the latent variables Z. When marginals are continuous, conditional independencies in Z carry over to X [34]; with discrete margins, additional dependencies may arise but are typically viewed as secondary [35].

One strategy to estimate Σ_Z is based on Hoff's extended rank likelihood [17], which avoids explicit modeling of the marginal distributions by treating the observed ranks as constraints on the latent variables. This approach underlies the *copula PC* method [18], where the correlation matrix is estimated via Gibbs sampling. A fully Bayesian extension has also been proposed [20], combining the rank likelihood with a DAG-Wishart prior. A known challenge in this context is the effective loss of sample size when estimating partial correlations involving sparse levels.

An alternative strategy, used in the *latent PC* method [19], estimates the latent correlation matrix directly from the observed data using pairwise Kendall's τ coefficients. For

each pair of variables (X_j, X_k), the latent correlation σ_{jk} is estimated via the identity $\sigma_{jk} = \sin(\pi\tau_{jk}/2)$, where τ_{jk} denotes the empirical Kendall correlation between X_j and X_k. This provides a closed-form, rank-based estimator of Σ_Z that is simple to compute and robust to marginal distributional assumptions.

Irrespective of the method used to estimate the latent correlation matrix Σ_Z, structure learning proceeds by testing conditional independencies via partial correlations computed from Σ_Z. These tests form the basis of constraint-based algorithms such as the PC algorithm, applied directly to the latent space. By leveraging the Gaussian properties of the latent variables, standard methods for continuous data can be extended to mixed settings without requiring parametric assumptions on the observed marginals.

Hybrid structure learning from latent Gaussian copulas

Most existing structure learning approaches based on latent Gaussian copulas rely on constraint-based algorithms such as the PC algorithm [18, 19]. While these methods can consistently recover the Markov equivalence class under standard assumptions, they do not return a unique DAG. The only exception is the fully Bayesian approach of [20], which samples from the posterior distribution over DAGs. However, this method is computationally demanding and may become impractical as the number of variables increases.

We introduce a new class of hybrid algorithms for structure learning from mixed data, which we call *latent MMHC*. An intuitive and graphical overview of the procedure is shown in Fig. 2. Our approach extends existing PC-based copula approaches by incorporating a score-based refinement phase, following the MMHC framework [21]. The graph skeleton is first estimated via the *latent PC* algorithm, using Kendall's τ to estimate the latent correlation matrix. A constrained score-based search is then used to orient edges

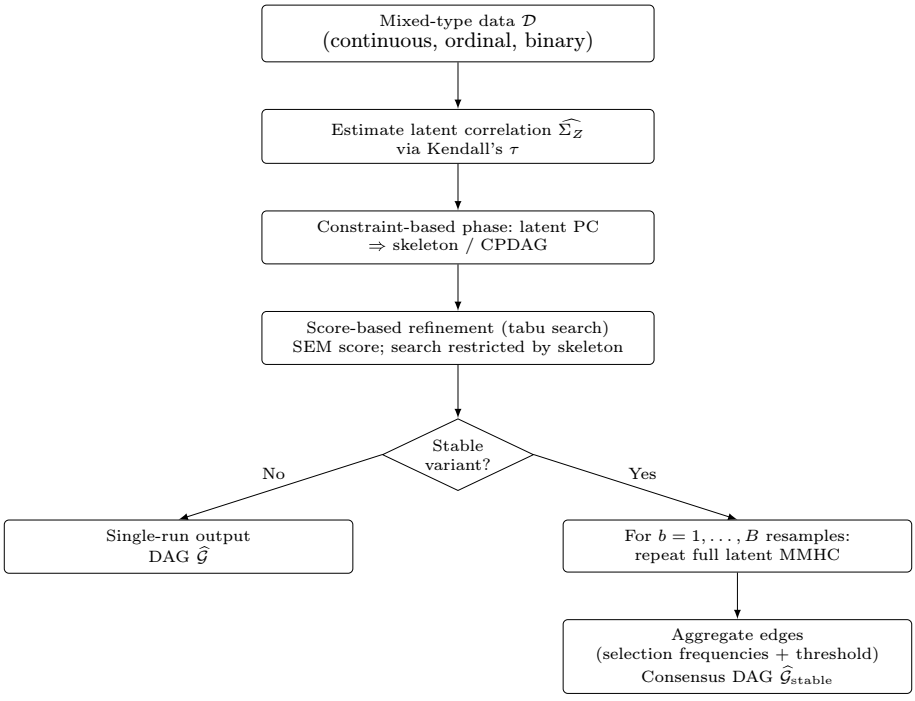

Fig. 2 Flowchart of latent MMHC using the SEM score. The stable variant repeats the full learning procedure on bootstrap resamples and aggregates edge frequencies to obtain a consensus structure

and return a DAG. To our knowledge, this is the first hybrid structure learning approach for mixed data based on latent Gaussian copulas, addressing a key gap in the literature: existing latent copula methods rely almost exclusively on constraint-based techniques, which return only equivalence classes and may struggle to orient edges reliably.

In the refinement step, we define a *blacklist* that prohibits any edge not present in the skeleton obtained during the constraint-based phase. We then apply a tabu search over the space of DAGs, scoring candidate structures using the Gaussian likelihood implied by the underlying structural equation model in Eq. (3), computed from the rank-based latent variables. The resulting procedure refines the initial equivalence class returned by the latent PC algorithm into a single oriented DAG.

We focus on latent MMHC because prior work has shown the latent PC method to outperform copula PC in both accuracy and scalability [19], particularly when dealing with high-dimensional or sparse data. However, the same hybrid strategy could be applied using copula PC in the first phase, resulting in a copula MMHC variant.

Although both copula PC and latent PC are consistent under appropriate conditions, our hybrid procedure does not inherit these guarantees. As discussed in [21], the score-based refinement phase in MMHC lacks theoretical consistency guarantees, since it relies on greedy local search which may converge to suboptimal DAGs even in the infinite-sample limit. Our method shares this limitation. Nevertheless, hybrid algorithms like MMHC have been shown to perform competitively in practice, often improving edge orientation and reducing structural errors compared to purely constraint-based methods. In this spirit, our approach seeks to combine the robustness of latent copula estimators with the practical strengths of score-based refinement.

To further improve the stability of the learned structures, we implement a bootstrap-based aggregation scheme that we refer to as *stable latent MMHC*. This procedure is based on the widely used nonparametric bootstrap for BNs [22], in which the structure learning algorithm is applied to multiple bootstrap replicates of the data. The resulting DAGs are aggregated into a consensus structure using the `bnlearn` functions `averaged.network` and `inclusion.threshold`, which compute edge selection frequencies and produce a final graph containing only those edges that appear with sufficient support across replicates. This extension enhances the robustness of our method to small-sample variability and complements the underlying statistical strength of the latent copula framework.

All proposed methods are implemented in `R`, using the `pcalg` package [36] for skeleton estimation via the PC algorithm and the `bnlearn` package [28] for the score-based refinement. The implementation is freely available at https://github.com/manueleleonelli/latent_mmhc. Both libraries support flexible, high-dimensional structure learning, and `bnlearn` allows custom scoring functions, enabling direct implementation of the Gaussian likelihood induced by the structural equation model. Since the final output is a `bnlearn` DAG object, it integrates seamlessly with the package's tools for probabilistic inference, model comparison, visualization, and bootstrapping, and is fully compatible with the broader ecosystem of `bnlearn`-based packages for extended analysis [37].

Simulation study

The aim of the simulation study is to benchmark structure learning methods across regimes of increasing statistical and combinatorial difficulty, rather than to reproduce a specific applied dataset. We therefore vary the number of variables p, the expected neighborhood size s, and the sample size n to span settings ranging from relatively sparse graphs with extensive data to denser graphs under more limited statistical power. These regimes are known to strongly affect the size of the space of admissible DAGs and, consequently, the difficulty of recovering the underlying structure.

Data-generating mechanism

We simulate data from a Gaussian copula model with an underlying DAG and mixed-type marginals. The procedure follows [18, 20], and is designed to reflect realistic features observed in applied problems, such as our volleyball dataset.

Given integers n (sample size), p (number of variables), and s (edge sparsity parameter), data are generated as follows:

1. **DAG structure:** We construct a weighted adjacency matrix $A \in \mathbb{R}^{p\times p}$ corresponding to a DAG where edges are only allowed from lower-indexed to higher-indexed nodes (i.e., $A_{ij} \neq 0$ only if $j < i$). This guarantees acyclicity by design, as the resulting graph respects a topological ordering. For each pair (i, j) with $i > j$, an edge from j to i is included independently with probability $s/(p-1)$, and the corresponding entry A_{ij} is assigned a random weight in [0.1, 1] with a random sign. This procedure yields a sparse DAG in which the expected number of neighbors per node is s. Although the DAGs are generated randomly, the imposed ordering and sparsity constraints yield layered sparse dependence patterns (many predictors feeding into downstream variables), which is a common abstraction in psychological and behavioral modeling where covariates and constructs act jointly but not all-to-all.
2. **Latent Gaussian data:** Latent variables $\mathrm{Z} \in \mathbb{R}^{n\times p}$ are generated by sampling $\boldsymbol{\varepsilon} \sim \mathcal{N}_p(0, I)$ and solving the structural equation model $\mathrm{Z} = \boldsymbol{\varepsilon}(\mathrm{I} - A)^{-T}$. This yields samples from a multivariate Gaussian distribution faithful to the generated DAG.
3. **Observed mixed data:** The first $p/2$ variables are transformed into discrete observations using monotonic transformations of the corresponding latent variables:
   - With probability 0.5, a variable is made *binary* by thresholding the standard normal CDF $\Phi(Z_j)$ at $1-\theta$ with $\theta \sim \mathrm{Uniform}(0.2, 0.8)$.
   - Otherwise, it is made *ordinal* by applying the quantile function of a Binomial(4, θ) distribution to $\Phi(Z_j)$, again with $\theta \sim \mathrm{Uniform}(0.2, 0.8)$.

This construction places discrete variables *upstream* of continuous ones, mirroring a common setting in psychological research, where background variables and item-level questionnaire responses are categorical or ordinal, while psychological constructs are represented by continuous scores obtained by aggregating multiple items. These aggregated scores are treated as latent traits and often modeled downstream of background factors. The resulting discrete margins are *unbalanced*, as frequently observed in practice. Our volleyball dataset is an instance of this broader setting. The remaining $p/2$ variables are kept continuous and correspond directly to their respective columns in Z.

Experimental setup and evaluation metrics

For each combination of number of variables $p \in \{10, 30\}$, sample size $n \in \{500, 1000, 2500, 5000\}$, and sparsity level $s \in \{2, 5\}$, we generated 50 random DAGs. For each DAG, a dataset of size n was simulated independently. The chosen parameter combinations are intended to explore how algorithmic performance varies with graph density and effective sample size. For fixed p, increasing s increases graph density and typically increases structural ambiguity (i.e., more competing graphs can yield similar conditional independence patterns), making recovery more challenging. Conversely, for fixed s, increasing p leads to relatively sparser graphs in proportional terms, which can improve edge recovery despite higher dimensionality. The effect of sample size n should be interpreted relative to graph complexity: smaller graphs with low s are comparatively data-rich, whereas denser graphs require substantially larger samples for reliable structure learning.

We considered a suite of structure learning methods, starting with the established *copula PC* and *latent PC* algorithms, with implementations given in [18] and [19], respectively. Following suggestions from both papers, we fixed the critical level for conditional independence testing to $\alpha = 0.01$. We then evaluated our proposed *latent MMHC* algorithm, which combines a latent PC skeleton with a score-based refinement step using the Gaussian likelihood implied by the structural equation model in Eq. (3). To assess the impact of a more permissive constraint phase, we also considered a variant with a higher significance level $\alpha = 0.05$ in the initial PC step.

Lastly, to isolate the contribution of combining a constraint-based skeleton with score-based refinement, we included a single purely score-based baseline. This *SEM score-based* method optimizes the Gaussian likelihood implied by the structural equation model in Eq. (3) using tabu search, without imposing any skeleton constraints. Comparing this approach with latent MMHC allows us to directly assess the gain achieved by integrating a constraint-based phase prior to score-based optimization. Additional score-based and copula-based variants were also explored during preliminary experimentation, but did not yield competitive performance relative to the methods reported here and are therefore omitted for clarity.

We evaluated each method using four performance metrics. The *Structural Hamming Distance* (SHD) counts the number of edge insertions, deletions, and reversals required to transform the estimated DAG into the true DAG; lower SHD values indicate better structural accuracy [21]. Following its original definition, SHD is computed on CPDAGs: DAG-returning methods are first converted to their corresponding CPDAGs, ensuring that methods which natively return equivalence classes (e.g., copula PC and latent PC) are compared on the same footing. *Sensitivity* (SEN) measures the proportion of true edges correctly identified (true positives), while *Specificity* (SPE) measures the proportion of absent edges correctly excluded (true negatives). Finally, we report the *runtime* of each method. All metrics are computed over 50 independent replications (each replication corresponds to one randomly generated DAG and one simulated dataset) for each (p, n, s) configuration, and summarized by their median values.

Results

The Structural Hamming Distance (SHD) results in Table 3 show that the proposed latent MMHC approach yields the lowest or near-lowest SHD across most configurations,

Table 3 Median Structural Hamming Distance (SHD) across simulation settings varying sample size, sparsity, and number of variables

Method	s = 2				s = 5			
	Sample size *n*				Sample size *n*			
	500	1000	2500	5000	500	1000	2500	5000
p = 10								
copula PC	7	7	6	5	22	22	**20**	21
latent PC	7	6	5	5.5	**21**	22	**20**	**19**
SEM score-based	7	7	5.5	5	23	22	23	23
latent MMHC ($\alpha = 0.01$)	**5**	4.5	**3**	**2**	**21**	**21**	**20**	19.5
latent MMHC ($\alpha = 0.05$)	**5**	**4**	**3**	2.5	**21**	**21**	**20**	**19**
p = 30								
copula PC	24	22	21	18.5	64.5	61	58	56.5
latent PC	30	30	27	25	60	53	48	44.5
SEM score-based	30	25	23	23	72	74.5	87	100
latent MMHC ($\alpha = 0.01$)	**17**	**12**	**8**	**8**	**55**	49	**44**	**41**
latent MMHC ($\alpha = 0.05$)	18	13	9	9	56	**48**	46	42.5

Bold values indicate the best-performing method for each configuration

Table 4 Median sensitivity across simulation settings varying sample size, sparsity, and number of variables

Method	s = 2				s = 5			
	Sample size *n*				Sample size *n*			
	500	1000	2500	5000	500	1000	2500	5000
p = 10								
copula PC	0.00	0.13	0.29	0.38	0.09	0.14	0.23	0.28
latent PC	0.39	0.41	0.50	0.56	0.17	0.21	0.26	0.32
SEM score-based	**0.50**	**0.58**	**0.67**	**0.75**	**0.35**	**0.41**	**0.48**	**0.50**
latent MMHC ($\alpha = 0.01$)	**0.50**	0.57	**0.67**	**0.75**	0.25	0.30	0.30	0.36
latent MMHC ($\alpha = 0.05$)	**0.50**	**0.58**	**0.67**	**0.75**	0.25	0.30	0.32	0.36
p = 30								
copula PC	0.11	0.23	0.33	0.46	0.12	0.22	0.34	0.43
latent PC	0.41	0.48	0.58	0.62	0.27	0.39	0.44	0.50
SEM score-based	**0.54**	**0.67**	**0.74**	**0.78**	**0.46**	**0.53**	**0.61**	**0.65**
latent MMHC ($\alpha = 0.01$)	0.53	0.65	**0.74**	0.77	0.34	0.43	0.50	0.53
latent MMHC ($\alpha = 0.05$)	**0.54**	0.65	**0.74**	0.77	0.36	0.44	0.49	0.54

Bold values indicate the best-performing method for each configuration

particularly as graph complexity increases. For smaller graphs ($p = 10$), latent MMHC consistently improves upon both copula PC and latent PC, with the $\alpha = 0.01$ variant performing best at smaller sample sizes and the $\alpha = 0.05$ variant providing slight gains at larger n. For larger graphs ($p = 30$), the advantage of latent MMHC becomes more pronounced: the $\alpha = 0.01$ variant attains the lowest SHD in all settings with $s = 2$ and in three out of four settings with $s = 5$, while the $\alpha = 0.05$ variant performs comparably in denser regimes. Overall, these results indicate that combining a constraint-based skeleton with score-based refinement substantially improves structural recovery relative to purely constraint-based or purely score-based alternatives.

The sensitivity results in Table 4 reveal a clear trade-off between aggressive edge recovery and structural conservatism. As expected, sensitivity increases monotonically with the strength of the underlying edge in the data-generating model, with strong dependencies being recovered with high probability even at moderate sample sizes (see

Appendix ""B.B.1 Edge-strength–stratified sensitivity Edge-strength–stratified sensitivity""). The purely score-based SEM approach consistently achieves the highest sensitivity across most configurations, particularly in denser graphs ($s = 5$) and for larger networks ($p = 30$). This behavior is expected, as the absence of a preliminary constraint-based phase means that true edges are not discarded early and can be recovered during the score-based search. The hybrid latent MMHC approach shows slightly lower sensitivity, especially in dense regimes, reflecting the fact that the initial constraint-based skeleton may remove some true edges before refinement. Nevertheless, latent MMHC consistently outperforms both copula PC and latent PC across all settings, indicating that combining constraint-based screening with score-based refinement substantially improves edge recovery relative to purely constraint-based methods. Increasing the PC threshold to $\alpha = 0.05$ yields modest gains in sensitivity, particularly for larger p, by allowing a less restrictive skeleton in the first phase.

To further examine whether sensitivity patterns could be influenced by discretization rather than by the learning algorithm itself, we conducted an additional simulation study comparing the proposed mixed-type method with discretization-based baselines using different numbers of quantile bins for continuous variables. As shown in Appendix ""B.B.2 Effect of discretization on edge recovery Effect of discretization on edge recovery"", discretization generally leads to a reduction in sensitivity for continuous–continuous edges and, to a lesser extent, for discrete–continuous edges, while having little effect on discrete–discrete relationships. These results suggest that discretization primarily impacts the recovery of edges involving continuous variables, whereas relationships among discrete variables are largely unaffected.

Specificity results in Table 5 highlight the complementary strengths of the hybrid approach. While the purely score-based SEM method attains high sensitivity, it exhibits noticeably lower specificity, particularly in dense settings, reflecting a tendency to introduce spurious edges when no structural constraints are imposed. In contrast, latent MMHC maintains very high specificity across all configurations, often matching or exceeding the performance of purely constraint-based methods. This effect is especially pronounced for larger graphs ($p = 30$), where both latent MMHC variants achieve

Table 5 Median specificity across simulation settings varying sample size, sparsity, and number of variables

Method	**s = 2**				**s = 5**			
	Sample size *n*				**Sample size *n***			
	500	**1000**	**2500**	**5000**	**500**	**1000**	**2500**	**5000**
p = 10								
copula PC	0.93	0.91	0.91	0.92	0.85	0.83	0.82	0.82
latent PC	0.91	0.92	0.92	0.92	0.87	0.86	0.85	0.87
SEM score-based	0.95	0.94	0.94	0.94	0.83	0.80	0.76	0.74
latent MMHC ($\alpha = 0.01$)	0.94	0.95	0.94	0.94	0.90	0.89	**0.89**	0.87
latent MMHC ($\alpha = 0.05$)	**0.96**	**0.96**	**0.96**	**0.97**	**0.91**	**0.90**	0.88	**0.88**
p = 30								
copula PC	0.97	0.97	0.97	0.97	0.96	0.96	0.95	0.95
latent PC	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.97
SEM score-based	0.97	0.98	0.98	0.97	0.94	0.93	0.91	0.89
latent MMHC ($\alpha = 0.01$)	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**	**0.98**	**0.98**
latent MMHC ($\alpha = 0.05$)	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**	**0.98**	**0.98**

Bold values indicate the best-performing method for each configuration

near-perfect specificity across all sample sizes and sparsity levels. The constraint-based skeleton plays a central role in controlling false positives, while the subsequent score-based refinement improves orientation without substantially inflating the false discovery rate. Although increasing the PC threshold to $\alpha = 0.05$ slightly reduces specificity in a few small-sample cases, the overall level remains high, illustrating that latent MMHC achieves a favorable balance between edge recovery and false positive control.

Finally, Table 6 highlights the computational trade-offs among the considered methods. As expected, purely constraint-based approaches are the fastest, with copula PC and latent PC exhibiting very low runtimes across all configurations. In contrast, the purely score-based SEM method is the most computationally demanding, particularly as the number of variables and the sample size increase. The proposed latent MMHC variants incur a moderate additional cost relative to constraint-based methods due to the score-based refinement step, but remain substantially faster than the purely score-based alternative. In particular, the $\alpha = 0.01$ variant shows runtimes only marginally larger than latent PC in most settings, while delivering markedly improved structural accuracy. Overall, latent MMHC achieves a favorable balance between computational efficiency and statistical performance, effectively combining the strengths of constraint-based and score-based learning.

Across Tables 3, 4, 5 and 6, the expected effect of the design parameters is visible. Increasing s yields denser graphs and typically increases the difficulty of distinguishing candidate structures from finite samples, which leads to larger SHD and lower sensitivity unless n is sufficiently large; at the same time, specificity remains comparatively high for most methods, indicating that false positive control is generally easier than complete edge recovery in the dense regime. Increasing n improves recovery (lower SHD and higher sensitivity) as conditional independence testing and score-based refinement become more reliable when the effective sample size grows relative to structural complexity. Differences between $p = 10$ and $p = 30$ reflect the trade-off between relative density (for fixed s, the edge probability $s/(p-1)$ decreases with p) and the increased dimensionality of conditioning, so performance need not vary monotonically with p.

Table 6 Median runtime (in seconds) across simulation settings varying sample size, sparsity, and number of variables

Method	s = 2				s = 5			
	Sample size *n*				Sample size *n*			
	500	1000	2500	5000	500	1000	2500	5000
p = 10								
copula PC	**0.01**	**0.01**	**0.01**	**0.02**	**0.03**	**0.05**	0.09	0.12
latent PC	0.02	0.02	0.02	0.03	**0.03**	**0.05**	**0.07**	**0.08**
SEM score-based	0.11	0.13	0.22	0.38	0.14	0.19	0.33	0.62
latent MMHC ($\alpha = 0.01$)	0.05	0.06	0.10	0.17	0.08	0.11	0.17	0.28
latent MMHC ($\alpha = 0.05$)	0.06	0.07	0.11	0.18	0.09	0.12	0.20	0.32
p = 30								
copula PC	**0.05**	**0.07**	**0.12**	**0.20**	**0.26**	**0.61**	1.73	3.71
latent PC	0.42	0.75	0.58	0.59	0.37	**0.61**	**0.98**	**1.20**
SEM score-based	0.81	1.01	1.64	2.69	1.22	1.70	3.31	6.72
latent MMHC ($\alpha = 0.01$)	0.55	0.88	0.88	1.09	0.50	0.79	1.32	1.87
latent MMHC ($\alpha = 0.05$)	1.04	2.27	2.37	2.53	0.74	1.13	1.59	2.29

Bold values indicate the fastest method for each configuration. Runtimes refer only to the DAG learning step, excluding the construction of the latent covariance matrix

Overall, the simulation results indicate that latent MMHC provides a favorable balance between structural accuracy, edge recovery, and false positive control. Compared to purely constraint-based methods, it achieves substantially lower structural error, while avoiding the loss of specificity observed for purely score-based learning. This balance becomes increasingly important as graph complexity grows. At the same time, the computational overhead of latent MMHC remains moderate and close to that of constraint-based approaches, particularly for the $\alpha = 0.01$ variant. Together, these findings support latent MMHC as an effective and practical approach for structure learning in mixed-type data.

Assessing the stable latent MMHC variant

We conducted a focused simulation to evaluate the potential benefits of bootstrap aggregation in our setting. As the bootstrap procedure is computationally intensive, we restricted the analysis to the case $p = 10$, with sample sizes $n \in \{500, 1000\}$. For the stable variant, we applied $B = 200$ bootstrap resamples within each replicate. We focused on the latent MMHC model with $\alpha = 0.01$, as it showed the best overall performance in the full simulation study. We compared this baseline against its bootstrap-aggregated (stable) counterpart.

Figure 3 summarizes the SHD results. The bootstrap-aggregated stable latent MMHC shows only modest gains over the baseline: improvements are visible either in slightly lower medians or in reduced variability across replications, depending on the configuration. The effect is not dramatic, but the stable variant generally delivers more consistent performance, in line with the idea that aggregation can dampen spurious edges and reinforce those that recur across resamples.

Modeling psychological traits of volleyball athletes

Building on the methodological framework and simulation results presented above, we now apply the proposed approach to the volleyball dataset introduced in Sect. "Motivating application and data". The goal is to infer an interpretable dependency structure among psychological traits and to evaluate how different methodological choices affect substantive conclusions in an applied context.

Model learning

Based on the simulation study in Sect. "Simulation study", we adopted the stable latent MMHC algorithm with the SEM score and significance level $\alpha = 0.01$, which provided

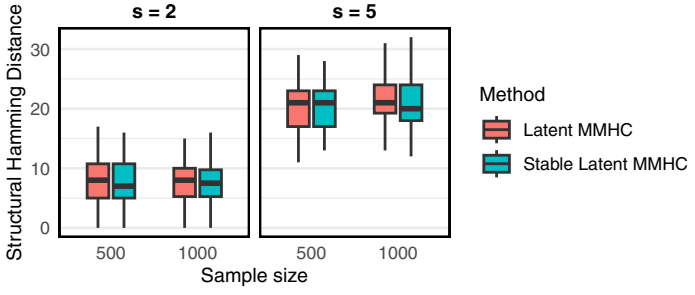

Fig. 3 SHD for latent MMHC and its stable (bootstrap-aggregated) variant across sample sizes and sparsity levels, for $p = 10$. Each panel corresponds to a different sparsity setting ($s = 2$ on the left and $s = 5$ on the right), and boxplots compare the standard and stable versions at $n = 500$ and $n = 1000$

the most favorable balance between structural accuracy, edge recovery, and false positive control across a wide range of settings. Given the relatively small sample size, we increased the number of bootstrap replicates to $B = 500$ to enhance the stability of the learned structure. In addition, edges from psychological traits to lifestyle and sport-related background variables were forbidden, as the latter are more naturally interpreted as exogenous covariates in this context.

The consensus structure obtained from the bootstrap aggregation step was used as the backbone of the final network. At this stage, the continuous psychological trait scores were discretized back to their original ordinal scales (e.g., 1–5 or 1–6) by rounding to the nearest category. A fully discrete Bayesian network was then fitted using Bayesian parameter estimation with a Dirichlet prior equal to one for all conditional probability tables. This prior acts as a regularization device, which is particularly important given the limited sample size: it avoids zero probabilities, improves the robustness of inference, and supports more reliable out-of-sample predictions. This two-step strategy leverages the information contained in the continuous scores during structure learning, while yielding a final model that remains interpretable in terms of the original psychological categories and compatible with standard discrete BN inference tools.

For comparison, we also learned a BN directly from the discretized data by applying standard discrete structure learning without the latent Gaussian step. This additional model serves as a baseline to assess how discretization at the structure-learning stage affects the resulting network in terms of sparsity and stability.

Results

The dependency structure learned using the stable latent MMHC algorithm is shown in Fig. 4 and is freely available within the `bnRep` R package [38]. None of the sport- or lifestyle-related background variables exhibit directed connections to the psychological traits in the learned network; for clarity, these variables are therefore omitted from the visualization. The resulting structure focuses exclusively on the conditional relationships among psychological constructs.

Several patterns stand out. Personality traits appear as upstream factors: *neuroticism* influences both *conscientiousness* and *worry*, while *conscientiousness* further drives *openness* and *agreeableness*. *Extraversion* is linked to *agreeableness* and to higher *self-esteem*. Together these pathways suggest that the Big Five, and in particular neuroticism, play a central role in shaping both other traits and anxiety-related outcomes.

Within the mental skills cluster, *goal setting* and *self-confidence* occupy a central position. *Goal setting* predicts *self-confidence*, and propagates to *mental practice* and *match preparation*, the latter of which leads to greater use of *self-talk*. *Self-confidence* also feeds directly into *self-esteem*, highlighting its role as a pivotal node bridging performance-related skills and global self-evaluations.

Finally, *emotional arousal* emerges as an important connector, connected to *extraversion*, *goal setting*, *self-confidence*, and *worry*. Together with the influence of *concentration disruption* on *worry*, this underscores the role of emotion regulation and attentional control in shaping anxiety responses.

Overall, the learned network points to a layered structure in which basic personality traits shape coping-related constructs such as *worry*, while motivational and

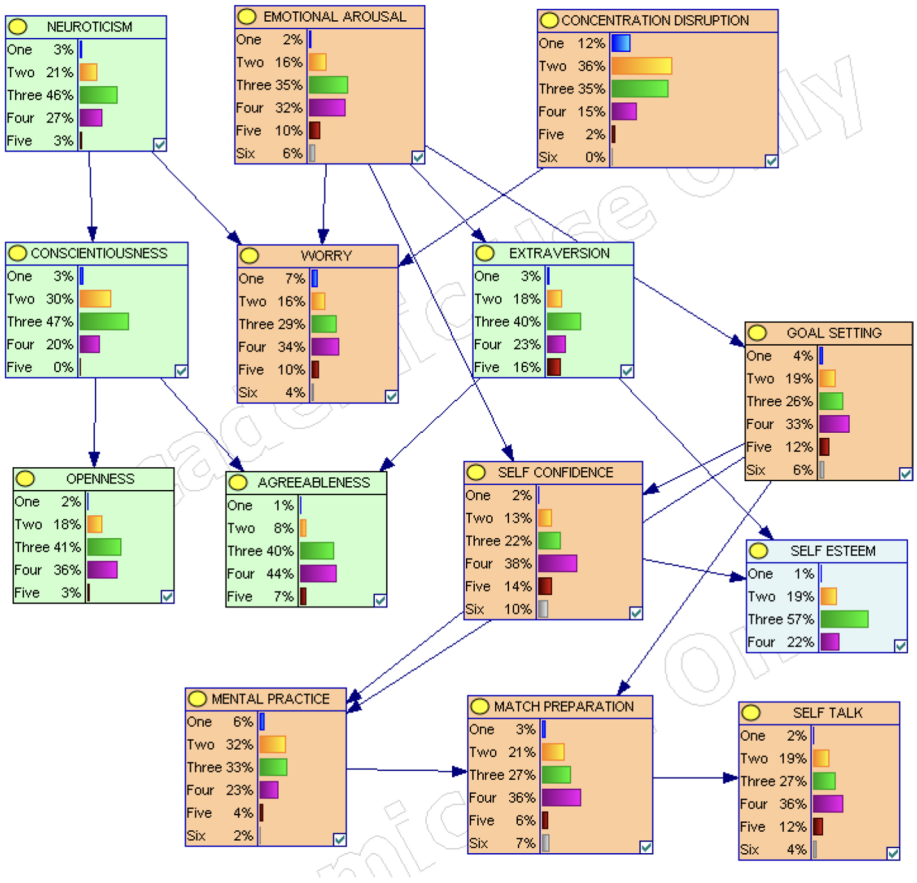

Fig. 4 Consensus BN of psychological traits for volleyball athletes, visualized using GeNIe. Nodes are color-coded by construct: green for Big Five traits, orange for IPPS-48 skills, and light blue for self-esteem. Each node includes a bar plot reporting its marginal distribution as computed from the estimated model

preparatory skills form an interconnected subsystem organized around *goal setting* and *self-confidence*.

For comparison, we also learned a BN directly from the discretized data by applying standard discrete structure learning without the latent Gaussian step. In this case, only three connections among psychological traits were identified, and, as in the latent MMHC model, no directed edges from sport or lifestyle variables to psychological traits were detected. Compared to the latent MMHC network, the discretized model yields a substantially sparser structure, with many relationships failing to achieve stable bootstrap support.

This behavior is consistent with the edge stability analysis in Fig. 5, which shows that discretization leads to more aggressive edge elimination in this dataset. While the discrete model does not contradict the main qualitative findings, it yields a less detailed representation of the dependency structure among psychological traits, potentially obscuring weaker yet consistently supported relationships.

Sensitivity analysis

To quantify the relative importance of predictors for key downstream variables, we computed variance-based Sobol indices using the exact algorithm introduced in [39]. Indices represent the proportion of variance in the output explained by each input. Table 7

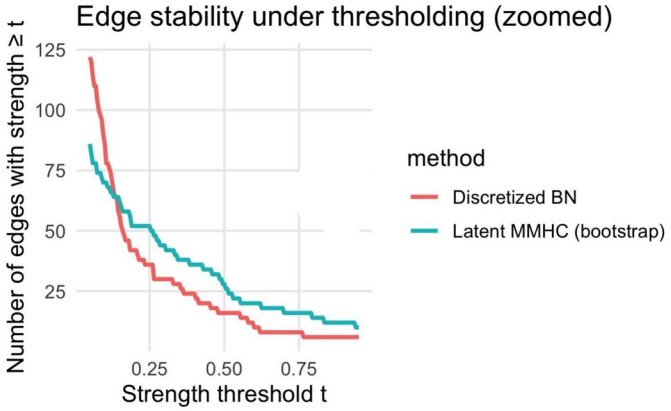

Fig. 5 Number of edges with bootstrap selection frequency at least *t* for the discretized Bayesian network and the latent MMHC approach

Table 7 Sobol indices (in %) for four selected outputs

Input	Worry	Self-confidence	Match preparation	Self-esteem
Neuroticism	6.95	0.00	0.00	0.00
Agreeableness	0.16	0.65	0.19	2.58
Conscientiousness	0.38	0.00	0.00	0.00
Extraversion	2.85	12.88	3.86	61.41
Openness	0.05	0.00	0.00	0.00
Self-talk	0.23	5.64	55.59	1.92
Goal setting	3.81	61.86	38.26	22.97
Self-confidence	3.73	–	31.27	45.56
Emotional arousal	14.55	49.05	14.72	38.49
Worry	–	1.85	0.56	1.41
Concentration disruption	24.47	0.00	0.00	0.00
Mental practice	1.34	37.35	35.75	12.99
Match preparation	0.83	20.28	–	7.01
Self-esteem	1.04	25.14	5.60	–

Entries correspond to the variance component attributed to each input variable; zeros indicate negligible contributions

reports the results for four outputs of particular interest: *Worry*, *Self-confidence*, *Match preparation*, and *Self-esteem*. All demographic and background variables not reported in Table 7 have Sobol indices equal to zero, as the learned Bayesian network contains no directed paths from these variables to the outcomes considered.

The decomposition highlights distinct drivers for each construct. *Worry* is dominated by concentration disruption, with further contributions from emotional arousal and neuroticism. *Self-confidence* is primarily shaped by goal setting and emotional arousal, with additional influence from mental practice. *Match preparation* receives balanced contributions from goal setting and mental practice, while self-talk appears almost entirely explained by preparation. Finally, *Self-esteem* reflects both personality (extraversion) and skills (self-confidence, emotional arousal), suggesting that athletes' global self-worth arises from an interaction of dispositional and trainable factors. Overall, these results confirm the layered structure identified in the network and quantify the extent to which each factor accounts for variability in key outcomes.

Conditional probabilities and belief dynamics

While Sobol indices identify the most influential predictors of a target, they do not reveal the shape or direction of these effects. In a BN framework, this information is captured by conditional probabilities, which describe how beliefs about an outcome update when evidence on a precursor is observed. To illustrate, we focus on *Match preparation*, one of the key downstream skills identified in the network. The corresponding conditional probabilities are reported in Table 8, while results for other outputs are deferred to the Appendix.

The table shows the baseline distribution of *Match preparation* alongside the conditional distributions when each of the five Sobol-relevant predictors (*Self-talk*, *Goal setting*, *Self-confidence*, *Emotional arousal*, and *Mental practice*) is fixed to one of its ordinal levels. Clear monotonic patterns emerge. For *Self-talk* and *Self-confidence*, very

Table 8 Conditional distributions of *Match preparation* (columns; categories One–Six) given key precursors (rows), compared to the baseline distribution

	One	Two	Three	Four	Five	Six
Baseline (no evidence)						
Baseline	0.03	0.21	0.27	0.36	0.06	0.07
Given self-talk						
One	0.32	0.64	0.01	0.01	0.01	0.01
Two	0.06	0.47	0.37	0.09	0.00	0.00
Three	0.02	0.20	0.29	0.41	0.07	0.00
Four	0.02	0.12	0.25	0.47	0.07	0.07
Five	0.00	0.05	0.21	0.42	0.11	0.21
Six	0.00	0.00	0.00	0.42	0.14	0.42
Given goal setting						
One	0.17	0.33	0.33	0.17	0.00	0.00
Two	0.10	0.32	0.32	0.23	0.00	0.03
Three	0.02	0.32	0.30	0.23	0.09	0.02
Four	0.00	0.13	0.24	0.56	0.04	0.04
Five	0.00	0.00	0.26	0.37	0.16	0.21
Six	0.00	0.10	0.10	0.40	0.10	0.30
Given self-confidence						
One	0.20	0.36	0.29	0.13	0.01	0.01
Two	0.09	0.29	0.30	0.25	0.03	0.03
Three	0.04	0.33	0.27	0.28	0.04	0.03
Four	0.01	0.14	0.28	0.49	0.06	0.02
Five	0.01	0.19	0.31	0.33	0.08	0.09
Six	0.00	0.07	0.10	0.32	0.17	0.33
Given emotional arousal						
One	0.11	0.38	0.31	0.15	0.03	0.02
Two	0.07	0.31	0.28	0.24	0.04	0.04
Three	0.04	0.23	0.29	0.35	0.05	0.04
Four	0.01	0.16	0.27	0.44	0.06	0.06
Five	0.01	0.13	0.22	0.39	0.10	0.16
Six	0.00	0.08	0.19	0.38	0.13	0.22
Given mental practice						
One	0.11	0.22	0.55	0.11	0.00	0.00
Two	0.06	0.44	0.23	0.17	0.02	0.08
Three	0.02	0.13	0.31	0.44	0.09	0.02
Four	0.00	0.03	0.26	0.60	0.03	0.08
Five	0.00	0.14	0.00	0.28	0.28	0.28
Six	0.01	0.01	0.01	0.32	0.32	0.32

Table 9 Definition of illustrative scenarios used for joint reasoning

Scenario	Definition (fixed evidence)
High motivation & arousal	GOAL SETTING = Six, EMOTIONAL AROUSAL = Six
Low motivation & arousal	GOAL SETTING = One, EMOTIONAL AROUSAL = One
Skills boost	SELF TALK = Five, MENTAL PRACTICE = Five
Rehearsal focus	MENTAL PRACTICE = Six, GOAL SETTING = Four
Resilient extrovert	EXTRAVERSION = Five, NEUROTICISM = Two
Vulnerable attentional	NEUROTICISM = Five, CONCENTRATION DISRUPTION = Five

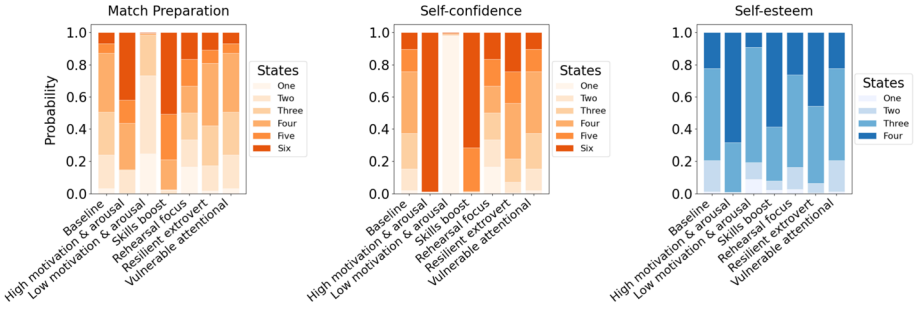

Fig. 6 Scenario-conditional distributions of *Match preparation*, *Self-confidence*, and *Self-esteem*. Bars show the baseline and six illustrative athlete profiles defined in Table 9

low levels (One–Two) concentrate mass on poor preparation, while higher levels (Four–Six) shift probability toward strong preparation. *Goal setting* and *Mental practice* display a similar gradient, with their higher levels producing the sharpest increases in the top preparation states. *Emotional arousal* has a more gradual effect, but higher levels still reduce the probability of poor preparation and increase the likelihood of mid-to-high preparation.

Taken together, these conditional distributions complement the variance-based analysis by clarifying not only which skills matter most, but also how their improvement reshapes the distribution of preparation outcomes. More generally, the additional results reported in the Appendix confirm that this pattern holds across other psychological targets, underscoring the value of conditional probability analysis as a complement to variance decomposition.

Joint scenario reasoning

While one-at-a-time conditioning clarifies the marginal impact of individual traits, practitioners are often interested in the combined effect of multiple skills or dispositions. BN allow for such *joint scenario reasoning*, where evidence is set on several predictors simultaneously and the resulting outcome distributions are compared to the baseline. This approach provides a decision-oriented complement to variance-based sensitivity analysis.

Table 9 summarizes six illustrative profiles of volleyball athletes considered in our analysis. These scenarios were constructed by fixing combinations of psychological traits suggested by the network structure and sensitivity analysis, and represent distinct motivational and cognitive styles.

Figure 6 displays the scenario-conditional distributions for three downstream outcomes of particular interest: *Match preparation*, *Self-confidence*, and *Self-esteem*. Each

bar summarizes the probability distribution over ordinal categories, with the baseline shown alongside the six scenarios.

Clear contrasts emerge across profiles. Athletes in the *High motivation & arousal* scenario are strongly shifted toward the top categories of both self-confidence and self-esteem, with match preparation also redistributed toward higher readiness. Conversely, the *Low motivation & arousal* scenario concentrates mass in the lowest states of self-confidence and preparation, underscoring the joint importance of goal orientation and arousal regulation. The *Skills boost* and *Rehearsal focus* profiles also increase the likelihood of stronger preparation, albeit through different combinations of mental skills. Finally, the *Resilient extrovert* profile improves both preparation and confidence, while the *Vulnerable attentional* scenario closely resembles the baseline, reflecting limited incremental impact when neuroticism and concentration disruption are both high.

Conclusions

This paper introduced *latent MMHC*, a new hybrid algorithm for learning BNs with mixed-type variables. By combining a latent Gaussian copula representation with a constraint-based skeleton and a constrained score-based refinement, the method extends existing copula-based learners to return a single oriented DAG. A bootstrap-aggregated variant further enhances stability. Together, these advances provide a general framework for extracting interpretable graphical structures from heterogeneous data, bridging a gap between purely discrete and Gaussian approaches.

Through a comprehensive simulation study, we showed that latent MMHC outperforms established competitors across a range of conditions. In particular, the SEM-based scoring variants reduced structural Hamming distance and improved edge recall, while maintaining high specificity and acceptable runtimes. These results confirm that a hybrid approach offers tangible benefits over constraint-only procedures, especially in orienting edges reliably under finite samples.

We then applied the method to a dataset of 164 female volleyball athletes, combining demographic information with standardized psychological measures. Treating questionnaire responses as continuous in the learning phase and then discretizing back to ordinal scales for interpretation proved effective: the procedure captured nuanced dependencies during structure estimation while yielding a final model that was interpretable in terms of the original psychometric categories. This approach illustrates a principled workflow for sports psychology datasets, where constructs are inherently latent continuous traits but are commonly measured on ordinal scales.

The learned network of psychological attributes revealed a layered organization: Big Five traits, and in particular neuroticism and extraversion, appeared upstream of coping-related constructs; goal setting and self-confidence formed a tightly linked motivational core; and emotional arousal connected motivation to anxiety-related outcomes. Sobol indices and conditional probability analyses quantified how shifts in key skills, such as improving goal setting or mental practice, propagate through the system to enhance match preparation, self-confidence, and ultimately self-esteem.

From an applied perspective, these findings underscore the potential of BNs as practical tools for coaches, sport psychologists, and athlete support staff. By modeling conditional relations explicitly, BNs allow practitioners to move beyond single-skill assessments and to reason about how improvements in one area may ripple through

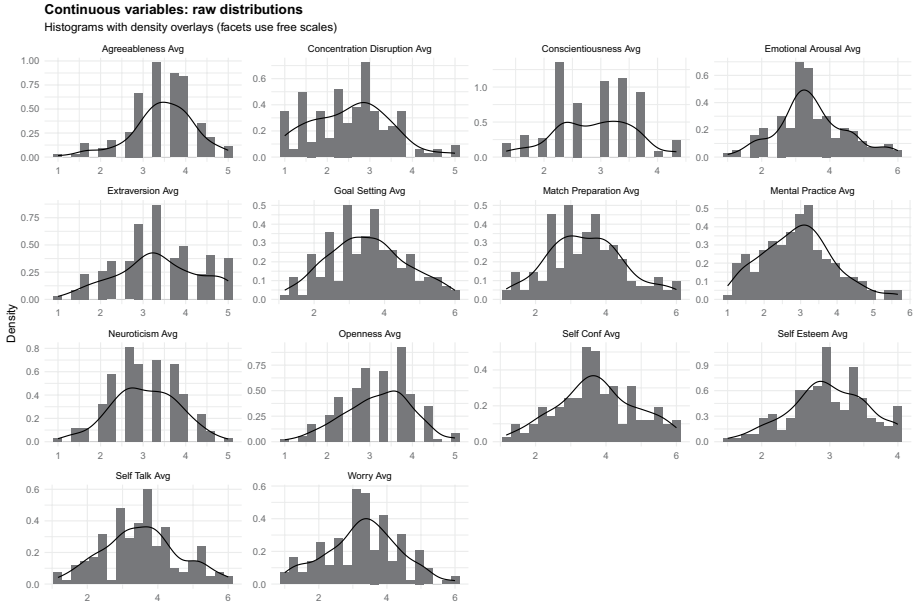

Fig. 7 Empirical distributions of the continuous psychometric scores (averaged questionnaire responses) used in the volleyball application. Each panel shows a histogram with a smoothed density overlay

others. Scenario analyses further provide an intuitive interface for exploring "what-if" questions, making BNs a promising complement to traditional profiling methods.

More broadly, the study illustrates how routinely collected psychological data can be transformed into decision support tools. Coaches can use such models to identify leverage points in mental skills training, to tailor interventions to individual profiles, and to anticipate indirect effects that may otherwise be overlooked. Incorporating probabilistic reasoning into athlete development thus has the potential to enhance both the personalization and effectiveness of training programs.

Future work may extend this line of research by integrating longitudinal data, linking psychological profiles with performance outcomes, and exploring causal inference under additional assumptions. Nonetheless, the present results demonstrate that BNs, learned via hybrid copula-based methods, offer a transparent and actionable framework for modeling the complex interplay of psychological traits in competitive sport.

A Raw distributions of continuous psychometric scores

This appendix reports the empirical distributions of the continuous psychological variables used in the volleyball application. As described in Sect. "Modeling psychological traits of volleyball athletes", these variables are obtained by averaging responses to multi-item Likert-type questionnaires (Big Five, IPPS-48, and self-esteem), a standard practice in psychometrics to summarize latent traits.

Although treated as continuous during the structure-learning phase, the resulting scores retain clear signatures of their ordinal measurement origin. In particular, the distributions exhibit bounded support, asymmetry, and visible heaping at round values, reflecting both scale limits and response patterns common in questionnaire data. These features are typical of psychological constructs measured on finite ordinal scales and

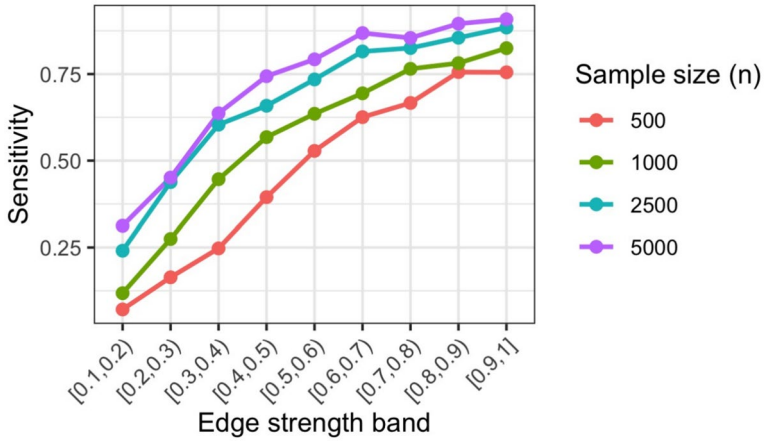

Fig. 8 Sensitivity of latent MMHC ($\alpha = 0.01$) stratified by edge-strength band for dense graphs ($p = 30$, $s = 5$). Each curve corresponds to a different sample size, averaged over 50 replications

motivate modeling approaches that accommodate heterogeneous data types while preserving interpretability (Fig. 7).

B Additional simulation results

B.1 Edge-strength–stratified sensitivity

To further interpret the sensitivity results, we examine how edge recovery depends on the strength of the underlying dependencies in the data-generating model. We focus on the most complex simulation setting ($p = 30$, $s = 5$), which yields denser graphs with a larger number of true edges and therefore allows for a more reliable assessment of sensitivity patterns.

For each sample size $n \in \{500, 1000, 2500, 5000\}$, we generated 50 random DAGs as described in Sect. "Simulation study". Each true edge was assigned a random weight with absolute value in [0.1, 1]. Edges were grouped into bands according to their absolute generating strength, and sensitivity was computed within each band as the proportion of edges correctly recovered by latent MMHC ($\alpha = 0.01$), averaged over replications.

Figure 8 shows that sensitivity increases monotonically with edge strength: strong dependencies are recovered with high probability even at moderate sample sizes, while missed edges are concentrated among weaker relationships. Increasing n mainly improves recovery in the intermediate-strength bands. These results confirm that lower overall sensitivity in dense regimes is driven primarily by weak edges, rather than failures to recover strong relationships.

B.2 Effect of discretization on edge recovery

To assess whether limited recovery of mixed-type edges could be attributed to discretization rather than to the learning algorithm itself, we conducted an additional simulation comparing the proposed mixed-type method with discretization-based baselines. Continuous variables were discretized using quantile binning into $K \in \{2, \ldots, 6\}$ categories, and structure learning was performed on the resulting fully discrete data. Figure 9 reports the difference in sensitivity between the two approaches by edge type and sample size.

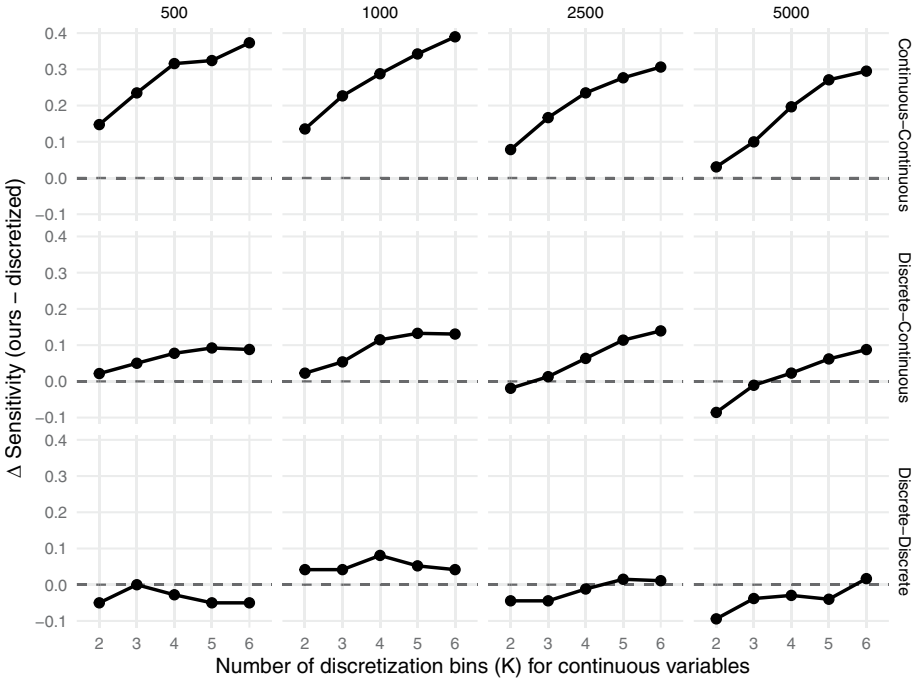

Fig. 9 Difference in sensitivity between the proposed mixed-type method and discretization-based baselines, Δ Sensitivity = $\text{Sensitivity}_{\text{ours}} - \text{Sensitivity}_{\text{discretized}}$, as a function of the number of discretization bins *K*. Rows correspond to edge-type categories and columns to sample size *n*

C Additional conditional probability inferences

See Tables 10, 11 and 12.

Table 10 Conditional distributions of *Worry* (columns; categories One–Six) given its main precursors (rows), compared to the baseline distribution

	One	Two	Three	Four	Five	Six
Baseline (no evidence)						
Baseline	0.07	0.16	0.29	0.34	0.10	0.04
Given neuroticism						
One	0.12	0.30	0.23	0.12	0.12	0.12
Two	0.07	0.28	0.25	0.34	0.03	0.03
Three	0.07	0.12	0.37	0.36	0.07	0.02
Four	0.07	0.12	0.23	0.35	0.18	0.05
Five	0.13	0.13	0.19	0.18	0.25	0.13
Given concentration disruption						
One	0.23	0.14	0.35	0.15	0.07	0.07
Two	0.09	0.29	0.25	0.29	0.06	0.02
Three	0.02	0.07	0.26	0.52	0.10	0.02
Four	0.04	0.07	0.45	0.20	0.18	0.06
Five	0.15	0.15	0.15	0.15	0.25	0.15
Six	0.17	0.17	0.17	0.17	0.17	0.17

Table 11 Conditional distributions of *Self-confidence* (columns; categories One–Six) given its main precursors (rows), compared to the baseline distribution

	One	Two	Three	Four	Five	Six
Baseline (no evidence)						
Baseline	0.02	0.13	0.22	0.38	0.14	0.10
Given extraversion						
One	0.08	0.34	0.30	0.23	0.01	0.05
Two	0.05	0.22	0.29	0.34	0.07	0.04
Three	0.01	0.14	0.23	0.40	0.13	0.08
Four	0.01	0.08	0.18	0.43	0.19	0.10
Five	0.00	0.07	0.15	0.34	0.19	0.24
Given self-talk						
One	0.06	0.26	0.33	0.22	0.10	0.03
Two	0.03	0.18	0.28	0.33	0.14	0.04
Three	0.02	0.14	0.22	0.41	0.14	0.08
Four	0.02	0.12	0.20	0.41	0.14	0.11
Five	0.01	0.10	0.17	0.38	0.15	0.19
Six	0.00	0.08	0.14	0.33	0.16	0.30
Given goal setting						
One	0.33	0.65	0.00	0.00	0.00	0.00
Two	0.03	0.32	0.45	0.16	0.03	0.00
Three	0.00	0.12	0.30	0.44	0.14	0.00
Four	0.00	0.05	0.11	0.63	0.13	0.07
Five	0.00	0.00	0.16	0.11	0.42	0.31
Six	0.00	0.00	0.00	0.20	0.10	0.69
Given emotional arousal						
One	0.32	0.32	0.32	0.01	0.01	0.01
Two	0.08	0.35	0.31	0.23	0.00	0.04
Three	0.00	0.18	0.32	0.39	0.09	0.04
Four	0.00	0.04	0.13	0.58	0.19	0.06
Five	0.00	0.00	0.13	0.19	0.37	0.31
Six	0.00	0.00	0.00	0.11	0.22	0.66
Given mental practice						
One	0.00	0.33	0.22	0.11	0.33	0.00
Two	0.04	0.25	0.38	0.17	0.13	0.02
Three	0.02	0.07	0.22	0.53	0.11	0.05
Four	0.00	0.05	0.05	0.63	0.13	0.13
Five	0.00	0.00	0.00	0.00	0.28	0.70
Six	0.01	0.01	0.01	0.01	0.01	0.96
Given match preparation						
One	0.12	0.40	0.30	0.13	0.04	0.01
Two	0.03	0.19	0.35	0.26	0.13	0.03
Three	0.02	0.15	0.22	0.40	0.16	0.04
Four	0.01	0.09	0.17	0.51	0.13	0.09
Five	0.00	0.06	0.13	0.35	0.17	0.29
Six	0.00	0.06	0.11	0.14	0.18	0.51
Given self-esteem						
One	0.09	0.02	0.21	0.16	0.31	0.20
Two	0.01	0.22	0.30	0.41	0.02	0.04
Three	0.03	0.16	0.23	0.41	0.12	0.05
Four	0.00	0.00	0.12	0.30	0.28	0.30

Table 12 Conditional distributions of *Self-esteem* (columns; categories One–Four) given its main precursors (rows), compared to the baseline distribution

	One	Two	Three	Four
Baseline (no evidence)				
Baseline	0.01	0.19	0.57	0.22
Given extraversion				
One	0.15	0.55	0.15	0.15
Two	0.03	0.36	0.59	0.03
Three	0.00	0.24	0.65	0.12
Four	0.00	0.06	0.54	0.40
Five	0.00	0.06	0.48	0.46
Given goal setting				
One	0.02	0.27	0.68	0.03
Two	0.01	0.28	0.63	0.08
Three	0.01	0.23	0.60	0.16
Four	0.01	0.17	0.58	0.24
Five	0.02	0.09	0.45	0.44
Six	0.02	0.09	0.36	0.53
Given self-confidence				
One	0.05	0.15	0.75	0.05
Two	0.00	0.32	0.68	0.00
Three	0.01	0.27	0.60	0.12
Four	0.00	0.21	0.61	0.17
Five	0.02	0.02	0.51	0.45
Six	0.02	0.07	0.27	0.64
Given emotional arousal				
One	0.03	0.24	0.65	0.07
Two	0.03	0.33	0.58	0.06
Three	0.01	0.23	0.63	0.13
Four	0.01	0.15	0.58	0.27
Five	0.00	0.07	0.44	0.49
Six	0.00	0.03	0.39	0.58

Author contributions
M.I. collected and curated the data. All authors contributed to data exploration and to the design of the methodology. M.L. implemented the methodology. All authors wrote and reviewed the manuscript and approved the final version.

Funding
D.J.L and M.L. were partially funded by PID2023-153222OB-I00 granted by MCIU/AEI/10.13039/501100011033/FEDER, UE.

Data availability
Due to privacy restrictions and data-use agreements, the underlying dataset cannot be shared.

Declarations

Competing interests
The authors declare no competing interests.

Received: 26 September 2025 / Accepted: 7 April 2026
Published online: 29 April 2026

References

1. Ayranci M, Aydin MK. The complex interplay between psychological factors and sports performance: a systematic review and meta-analysis. PLoS One. 2025;20(8):0330862.
2. Lochbaum M, Zanatta T, Kirschling D, May E. The profile of moods states and athletic performance: a meta-analysis of published studies. Eur J Investig Health Psychol Educ. 2021. https://doi.org/10.3390/ejihpe11010005.
3. Allen MS, Laborde S. The role of personality in sport and physical activity. Curr Dir Psychol Sci. 2014;23(6):460–5.
4. Laborde S, Allen MS, Katschak K, Mattonet K, Lachner N. Trait personality in sport and exercise psychology: a mapping review and research agenda. Int J Sport Exerc Psychol. 2020;18(6):701–16.

5. McGarry T. Applied and theoretical perspectives of performance analysis in sport: scientific issues and challenges. Int J Perform Anal Sport. 2009;9(1):128–40.
6. Fabbricatore R, Iannario M, Romano R, Vistocco D. Component-based structural equation modeling for the assessment of psycho-social aspects and performance of athletes: measurement and evaluation of swimmers. AStA Adv Stat Anal. 2023;107(1):343–67.
7. Iannario M, Romano R, Vistocco D. Dyadic analysis for multi-block data in sport surveys analytics. Ann Oper Res. 2023;325(1):701–14.
8. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques. London: MIT Press; 2009.
9. Briganti G, Scutari M, McNally RJ. A tutorial on bayesian networks for psychopathology researchers. Psychol Methods. 2023;28(4):947.
10. Fuster-Parra P, García-Mas A, Ponseti F, Palou P, Cruz J. A Bayesian network to discover relationships between negative features in sport: a case study of teen players. Qual Quant. 2014;48(3):1473–91.
11. Ponseti FJ, Almeida PL, Lameiras J, Martins B, Olmedilla A, López-Walle J, et al. Self-determined motivation and competitive anxiety in athletes/students: a probabilistic study using Bayesian networks. Front Psychol. 2019;10:1947.
12. Constantinou AC, Fenton NE, Neil M. pi-football: a Bayesian network model for forecasting Association Football match outcomes. Knowl Based Syst. 2012;36:322–39.
13. D'Urso P, De Giovanni L, Vitale V. A Bayesian network to analyse basketball players' performances: a multivariate copula-based approach. Ann Oper Res. 2023;325(1):419–40.
14. Zhou JQ, Liu Y. Probability prediction of groundstroke stances among male professional tennis players using a tree-augmented Bayesian network. Int J Perform Anal Sport. 2024;24(5):403–15.
15. Beuzen T, Marshall L, Splinter KD. A comparison of methods for discretizing continuous variables in Bayesian networks. Environ Model Softw. 2018;108:61–6.
16. Nojavan F, Qian SS, Stow CA. Comparative analysis of discretization methods in Bayesian networks. Environ Model Softw. 2017;87:64–71.
17. Hoff P. Extending the rank likelihood for semiparametric copula estimation. Ann Appl Stat. 2007;1:265–83.
18. Cui R, Groot P, Heskes T. Copula PC algorithm for causal discovery from mixed data. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2016;377–392. Springer
19. Cai Z, Xi D, Zhu X, Li R. Causal discoveries for high dimensional mixed data. Stat Med. 2022;41(24):4924–40.
20. Castelletti F. Learning Bayesian networks: a copula approach for mixed-type data. Psychometrika. 2024;89(2):658–86.
21. Tsamardinos I, Brown LE, Aliferis CF. The max-min hill-climbing Bayesian network structure learning algorithm. Mach Learn. 2006;65(1):31–78.
22. Caravagna G, Ramazzotti D. Learning the structure of Bayesian networks via the bootstrap. Neurocomputing. 2021;448:48–59.
23. McCrae RR, Costa PT. The five-factor theory of personality. In: John OP, Robins RW, Pervin LA, editors. Handbook of Personality: Theory and Research. New York, NY: Guilford Press; 2008. p. 159–81.
24. Robazza C, Bortoli L, Gramaccioni G. L'inventario psicologico della prestazione sportiva (ipps-48). Giornale Italiano di Psicologia dello Sport. 2009;4:14–20.
25. Rosenberg M. Society and the adolescent self-image. Princeton, NJ: Princeton University Press; 1965.
26. Andersson SA, Madigan D, Perlman MD. A characterization of Markov equivalence classes for acyclic digraphs. Ann Stat. 1997;25(2):505–41.
27. Colombo D, Maathuis MH. Order-independent constraint-based causal structure learning. J Mach Learn Res. 2014;15(1):3741–82.
28. Scutari M. Learning bayesian networks with the bnlearn R package. J Stat Softw. 2010;35:1–22.
29. Scutari M, Graafland CE, Gutiérrez JM. Who learns better bayesian network structures: accuracy and speed of structure learning algorithms. Int J Approx Reason. 2019;115:235–53.
30. Kalisch M, Bühlman P. Estimating high-dimensional directed acyclic graphs with the PC-algorithm. Journal of Machine Learning Research 2007;8(3).
31. Harris N, Drton M. PC algorithm for nonparanormal graphical models. J Mach Learn Res. 2013;14(1):3365–83.
32. Raghu VK, Poon A, Benos PV. Evaluation of causal structure learning methods on mixed data types. In: Proceedings of 2018 ACM SIGKDD Workshop on Causal Discovery, 2018;48–65. PMLR
33. Talvitie T, Eggeling R, Koivisto M. Learning bayesian networks with local structure, mixed variables, and exact algorithms. Int J Approx Reason. 2019;115:69–95.
34. Liu H, Lafferty J, Wasserman L. The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. Journal of Machine Learning Research 2009;10(10).
35. Dobra A, Lenkoski A. Copula Gaussian graphical models and their application to modeling functional disability data. The Annals of Applied Statistics. 2011;2A:969–93.
36. Kalisch M, Mächler M, Colombo D, Maathuis MH, Bühlmann P. Causal inference using graphical models with the R package pcalg. J Stat Softw. 2012;47:1–26.
37. Leonelli M, Ramanathan R, Wilkerson RL. Sensitivity and robustness analysis in bayesian networks with the bnmonitor R package. Knowl-Based Syst. 2023;278:110882.
38. Leonelli M. bnRep: a repository of Bayesian networks from the academic literature. Neurocomputing. 2025;624:129502.
39. Ballester-Ripoll R, Leonelli M. Computing sobol indices in probabilistic graphical models. Reliab Eng Syst Saf. 2022;225:108573.

Publisher's Note