14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

AN ENTROPY-BASED FORMULATION OF BAYESIAN NETWORK FOR LARGE-SCALE DIGITAL TWINS

ZIHENG GENG[1] AND MINGHUI CHENG[1,2]

[1]Department of Civil & Architectural Engineering
University of Miami, Coral Gables, FL 33146, USA
ziheng.geng@miami.edu

[2]School of Architecture
University of Miami, Coral Gables, FL 33146, USA
minghui.cheng@miami.edu

Key words: Entropy, Uncertainty Quantification, Bayesian Network, System Digital Twin.

Abstract. Bayesian networks (BNs) have emerged as a fundamental mathematical tool for developing system digital twins (SDTs). Despite substantial advancements, the scalability of SDTs remains constrained by the increasing storage and computational demands associated with large BNs. To address this limitation, this paper proposes an entropy-based formulation to improve the computational efficiency of BNs while maintaining predictive accuracy. The validity of entropy in capturing node uncertainty is first demonstrated through two representative BN structures: (1) in the chain structure, entropy change decreases with distance from observations; and (2) in the common effect structure, entropy change increases as additional observations become available. Then, analytic expressions are derived from a two-node BN structure. It shows that key statistical parameters, including entropy, mean, and variance, can effectively model uncertainty propagation between connected nodes within BNs, achieving an average RMSE of 0.027 across 100 random BNs. The formulation is further generalized to the common effect structure, yielding an average RMSE of 0.082 for entropy prediction. By reducing BN storage requirements under varying network complexities, the proposed formulation offers a promising pathway for enabling scalable and efficient SDTs for complex infrastructure systems.

1 INTRODUCTION

With rapid advancements in artificial intelligence (AI), Internet of Things (IoT), and computational power, digital twin (DT) has emerged as a promising and transformative tool for effective management of infrastructure systems. By definition, DTs are virtual representations of physical assets or systems, wherein digital entities are dynamically and continuously synchronized with their physical counterparts through bidirectional information exchange [1]. Specifically, the digital replicas are updated through integrating muti-source data from the real world to mirror the real-time condition and state of the physical assets. The updated virtual representations can provide insights into informed decision-making for the operation, management, and maintenance of physical systems. This closed-loop feedback between the physical and virtual spaces holds the potential to enhance the performance of infrastructure systems and significantly reduce their life-cycle costs.

Recently, the concept of system digital twin (SDT) was proposed by Cheng et al. [2-4] to develop DTs from the perspective of complex systems. The SDT framework involves three phases: (a) construction of a knowledge graph (KG) to capture the statistical correlations and interdependencies across various levels of the complex system; (b) transformation of the KG into a Bayesian network (BN) by fitting conditional probability tables (CPTs) to quantify probabilistic dependencies; and (c) conversion of the BN into an SDT by specifying data and performance nodes to enable automated Bayesian inference. The SDT framework is demonstrated via a case study for monitoring the risk of bridge network in Miami-Dade County. This county-scale SDT interlinks three systems (bridge, traffic, and river) for probabilistic modeling and risk updating. However, the SDT faces significant computational challenges: the SDT comprises 6,478 nodes and 10,584 edges, requiring over 256 GB of memory to store its CPTs even with a coarse discretization. Such demands severely limit the scalability and practical applications of SDTs in large-scale infrastructure systems.

To address this challenge, this paper proposes a novel entropy-based formulation as an efficient surrogate for BNs. The formulation is designed to quantify uncertainty propagation within BNs in a computationally tractable manner while preserving predictive accuracy. The rationale for using entropy as an indicator of uncertainty propagation is first demonstrated through two representative BN topologies: the chain structure and the common effect structure. Subsequently, analytic expressions are derived from a simple two-node BN to capture the relationship between the entropies of connected nodes, thereby replacing the computationally intensive probabilistic updates via CPTs. Finally, the proposed formation is extended to a complex BN topology, demonstrating strong generalization capability and robust performance in modeling uncertainty propagation within BN.

2 CONCEPTS OF BAYESIAN NETWORK

BNs are probabilistic graphical models that represent a set of random variables and their conditional dependencies via a directed acyclic graph [5]. They have emerged as key enablers of SDTs due to their alignment with the core objectives of SDTs. Specifically, SDTs are expected to fulfill the following functions: (a) modeling complex correlations and interdependencies within infrastructure systems; (b) integrating heterogeneous, multi-source data from the physical world; and (c) enabling probabilistic inference and predictive analytics to support informed decision-making. BNs are well-suited to support these functions through their unique capabilities: (a) explicit representation of statistical dependencies via a graphical structure; (b) seamless integration of diverse data sources by linking them to relevant nodes in the network; (c) incorporation of observed evidence to infer the posterior probability of variables of interest.

The development of a BN entails two primary steps: specifying the network structure and estimating the CPTs. The former step establishes the statistical dependencies among variables, while the latter step quantifies the probabilistic relationships between connected nodes. The network structure can be obtained through expert knowledge, empirical data analysis, or a combination of both. To identify an optimal network topology, structure learning algorithms, including constraint-based, score-based, or hybrid approaches [6-8], are commonly used. After the structure is established, a dataset, derived from filed observations or numerical simulations, is required to estimate the marginal and CPTs of BN. Then, parameter learning is conducted using methods such as expectation–maximization and maximum likelihood estimation [9,10]. With the constructed BN, Bayesian inference can be performed to update the posterior distributions of variables of interest given partial or complete evidence.

3 FORMULATE UNCERTAINTY PROPAGATION WITHIN BN USING ENTROPY-BASED ANALYTIC EXPRESSIONS

While BN has proven effective in the development and implementation of SDTs, it faces a significant computational challenge as system size and complexity scale, i.e., the exponential growth in storage demands associated with CPTs. To address this challenge, this study proposes an entropy-based formulation as a computationally efficient surrogate for BN. By deriving and fitting analytical expressions that capture uncertainty propagation between connected nodes, the proposed formulation aims to improve computational efficiency while preserving predictive accuracy.

3.1 Quantify node uncertainty via entropy

Entropy is a quantitative metric that characterizes the degree of disorder, randomness, and uncertainty within a system [11]. Originating from classical thermodynamics and later formalized in information theory, the concept of entropy has been broadly adopted across a wide range of fields. For a discrete random variable X that assumes N distinct states $\{x_1, x_2, \ldots, x_N\}$, the entropy of X, denoted as H_X is defined as [12]:

$$H_X = -\sum_{i=1}^{N} p(x_i) \log_N p(x_i) \tag{1}$$

where $p(x_i)$ represents the probability of occurrence of state x_i, satisfying $0 \leq p(x_i) \leq 1$ and $\sum_{i=1}^{N} p(x_i) = 1$. The base of the logarithmic function is set to N, the total number of discrete states, to normalize the entropy to the range from 0 to 1.

In the context of BN, entropy can serve as a measure for quantifying the uncertainty associated with marginal and conditional probability distributions. To evaluate the validity of entropy in capturing node uncertainty within BN, several Bayesian updating scenarios are constructed. These scenarios are designed to investigate how uncertainty propagates across BN and how entropy of individual nodes changes in response to new evidence. Herein, two typical BN topologies are selected: the chain structure and the common effect structure. The chain structure consists of a linear sequence of nodes, where each node is dependent on a single parent and influences a single child, forming a unidirectional path of dependencies. In contrast, the common effect structure involves a single child node that is dependent on multiple parent nodes, capturing the joint influence of multiple upstream variables.

For the chain structure, a BN comprising a linear sequence of ten nodes is constructed, as shown in Figure 1(a). Specifically, node 1 functions as the root node and is assigned a marginal probability distribution, while nodes 2 through 10 are assigned with conditional probability distributions. Each node is discretized into two distinct states to ensure computational tractability. A Bayesian updating scenario is selected to examine how uncertainty propagates through the network. In this scenario, new evidence is introduced at node 1, after which Bayesian inference is conducted to update the posterior probability distributions of the downstream nodes in the chain. To systematically investigate the uncertainty propagation under various parameter configurations, a total of 100 BNs are created by randomly sampling the CPTs for the nodes.

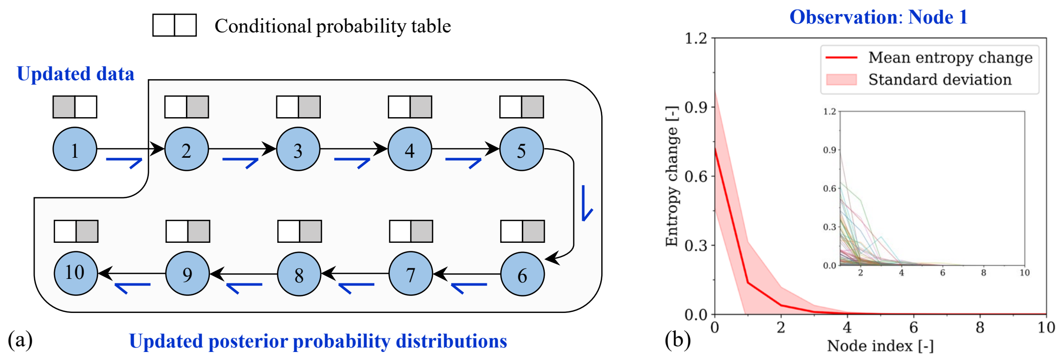

Figure 1: Bayesian updating in a chain-structured BN: (a) Bayesian updating using data observed at root node; (b) Entropy change of nodes following Bayesian updating.

The results are post-processed as follows: for each node, the entropy is calculated before (initial entropy) and after Bayesian inference (updated entropy). The entropy change is then determined as the difference between the two values. The above procedures are conducted across all 100 BNs, yielding individual entropy change curves in the inset of Figure 1(b). These curves are aggregated to calculate the mean entropy change and its standard deviation at each node, as depicted in Figure 1(b). The figure reveals a clear trend: the entropy change decreases sharply as the distance from the observed node increases, indicating that the impact of new observations weakens as it propagates along the chain-structured network. Notably, the entropy change of nodes 5 to 10 approaches zero, implying that the updates at data nodes have a negligible effect on the uncertainties of these distant nodes.

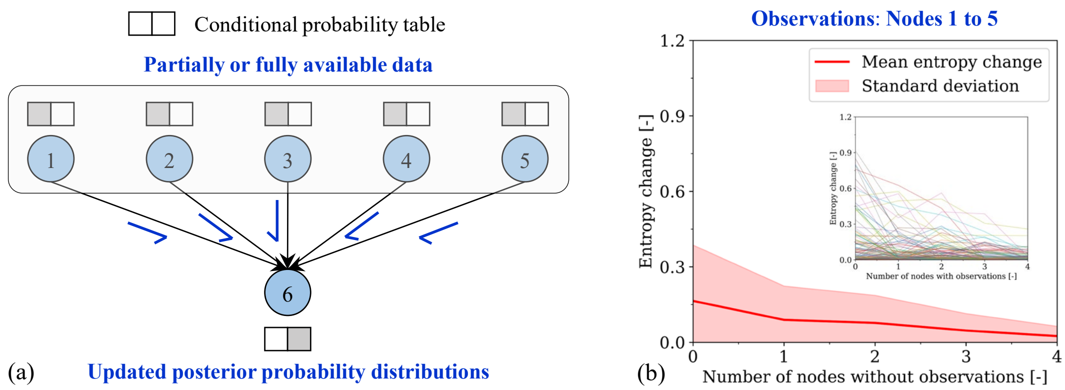

Figure 2: Bayesian updating in BN with a common effect structure: (a) Bayesian updating using partially or fully available data from parent nodes; (b) Entropy change of child node following Bayesian updating.

For the common effect structure, a BN comprising six nodes is constructed, where node 6 functions as the child node influenced by five parent nodes (nodes 1 to 5), as shown in Figure 2(a). Nodes 1 to 5 are assigned marginal probability distributions, whereas node 6 is equipped with a conditional probability distribution. Consistent with the previous experiment, each node is discretized into two states and a total of 100 BNs are generated with randomly sampled CPTs.

To investigate uncertainty propagation across the network, several Bayesian updating scenarios are created where partially or fully observed data from parent nodes are incorporated to perform Bayesian inference thereby updating the posterior probability distribution of child node.

The results are post-processed as follows: for each scenario, the entropy change of node 6 is computed to examine how uncertainty of child node evolves in response to various numbers of observations from parent nodes. This yields entropy change curves under different observation conditions, as shown in the inset of Figure 2(b). These individual curves are aggregated to derive the mean entropy change and its standard deviation under various number of node observations, as depicted in Figure 2(b). The figure reveals an obvious trend: the entropy change of the child node increases as more parent nodes are observed. Specifically, when fewer than three parent nodes are observed, the entropy exhibits minimal change, indicating that limited evidence has a modest effect on reducing uncertainty. In contrast, when four or five parent nodes are observed, the entropy change increases substantially, suggesting a significant reduction in the uncertainty of the child node.

3.2 Derive analytic expressions for entropy-based uncertainty propagation

The findings from the preceding subsection demonstrate the rationale for selecting entropy as the central metric for quantifying uncertainty in the probability distributions within BNs. To complement entropy and provide a comprehensive characterization of probability distributions, additional statistical features are also extracted, including measures of central tendency (e.g., mean, median, and mode), dispersion (e.g., variance and standard deviation), asymmetry (e.g., skewness), and tailedness (e.g., kurtosis). With these descriptors, the uncertainty propagation between connected nodes can be modeled by formulating entropy-based analytic expressions that capture statistical relationships across nodes. For example, the analytic expressions linking two nodes can be represented as follows:

$$H_B = f_1(H_A, \mu_A, \sigma_A^2) + \varepsilon \tag{2}$$

$$\mu_B = f_2(H_A, \mu_A, \sigma_A^2) + \varepsilon \tag{3}$$

$$\sigma_B^2 = f_2(H_A, \mu_A, \sigma_A^2) + \varepsilon \tag{4}$$

where μ and σ^2 denote the mean and variance, respectively. ε represents the residual error. This formulation enables the transformation of uncertainty propagation in BN from computationally intensive CPT-based updates to efficient analytic expressions, as illustrated in Figure 3.

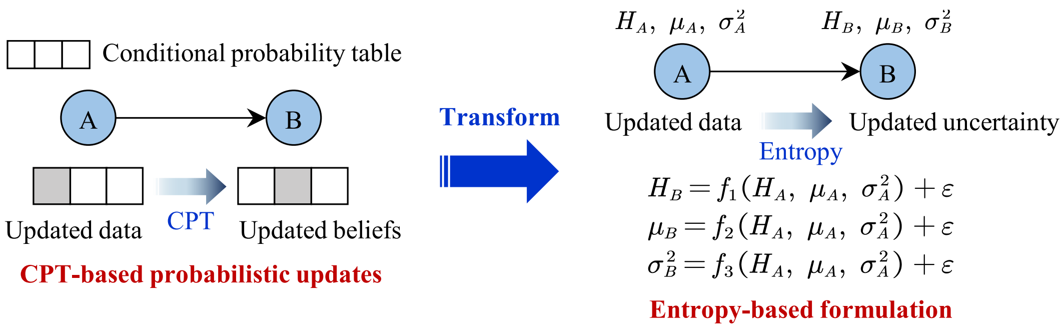

Figure 3: Transformation of uncertainty propagation in BN from computationally intensive CPT-based updates to efficient entropy-based analytic expressions.

To derive the functional forms of the analytic expressions, a simple BN with two nodes is constructed, in which node B is the child of node A. Both nodes are discretized into five distinct states. A marginal probability distribution is assigned to node A, while a conditional probability distribution is specified for node B. To simulate various evidence scenarios, node A is assigned a range of prior probability distributions. For each scenario, Bayesian inference is performed to compute the posterior probability distribution of node B. This procedure yields a dataset comprising paired samples of prior distributions for node A and the corresponding posterior distributions for node B. From these distributions, statistical features are extracted to serve as the basis for fitting analytic expressions that capture the uncertainty propagation within BN. Note that all statistical features are normalized to the range of 0 to 1.

Next, regression techniques are employed to derive functional relationships between the statistical features of prior and of posterior distributions. A range of candidate functional forms are explored to capture these relationships, including linear, polynomial, and nonlinear functions (e.g., trigonometric, exponential, and logarithmic forms). To achieve an optimal balance between model complexity and predictive accuracy, Akaike information criterion (AIC) [13] is utilized to identify and exclude redundant terms that contribute negligibly to the model performance. The quality of the fitted expressions is assessed using standard metrics of goodness-of-fit, including the coefficient of determination (R^2) and the relative mean squared error (RMSE).

To derive generalized analytic expressions for uncertainty propagation, a total of 100 BNs are generated, each with a randomly sampled CPT. The fitting results demonstrate that three statistical features, i.e., entropy, mean, and variance, can effectively capture the uncertainty propagation within BNs. The mean (μ_X) and variance (σ_X^2) of a discrete random variable X are defined as below:

$$\mu_X = \sum_{i=1}^{\mathrm{N}} x_i \cdot p(x_i) \tag{5}$$

$$\sigma_X^2 = \sum_{i=1}^{\mathrm{N}} (x_i - \mu_X)^2 \cdot p(x_i) \tag{6}$$

where x_{max} and x_{min} represent the maximum and minimum values of the variable state space, respectively. The derived optimal functional forms for predicting the entropy, mean, and variance of the child node B, based on those features of the parent node A, are as follows:

$$H_B = a \cdot H_A + b \cdot \mu_A + c \cdot \sigma_A^2 + d \cdot {\mu_A}^3 + e \cdot \mu_A \cdot \sigma_A^2 + f \cdot \exp(g \cdot \mu_A + h \cdot \sigma_A^2) + i \tag{7}$$

$$\mu_B = a \cdot \mu_A + b \cdot \sigma_A^2 + c \cdot {\mu_A}^2 + d \cdot {\mu_A}^3 + e \cdot \mu_A \cdot \sigma_A^2 + f \cdot \exp(g \cdot H_A + h \cdot \mu_A + i \cdot \sigma_A^2) + j \tag{8}$$

$$\sigma_B^2 = a \cdot \mu_A + b \cdot {\mu_A}^2 + c \cdot {\mu_A}^3 + d \cdot H_A \cdot \sigma_A^2 + e \cdot \exp(f \cdot H_A + g \cdot {\mu_A}^3 + h \cdot \sigma_A^6) + i \cdot \log(\mu_A + j) + k \tag{9}$$

Across all 100 BNs, the fitted analytic expressions exhibit robust predictive performance, with average R^2 values exceeding 0.81 and average RMSE values below 0.03. Fitting results for a representative BN using the proposed analytic expressions are depicted in Figure 4.

To evaluate the effectiveness of the proposed formulation across various levels of discretization, BNs with various numbers of discrete states are constructed. For each specified number of states, 100 BNs are generated with randomly sampled CPTs. Then, the analytic expressions are fitted individually to each BN. The resulting performance metrics, including R^2 and RMSE, are averaged over 100 networks, summarized in Tables 1-3. Generally, the proposed formulation exhibits robust performance across a wide range of discretization levels. Even in the case of nodes discretized into seven states, the average R^2 for entropy prediction reaches 0.74 and the average RMSE remains below 0.03. Additionally, a consistent trend is observed across the fitting of entropy, mean, and variance: as the number of discrete states increases, the predictive accuracy of the formulation tends to decline. This decline can be attributed to the increasing complexity of the CPTs. As the number of discrete states grows, the conditional relationships between nodes become more intricate, requiring a larger number of parameters to accurately capture the probabilistic dependencies.

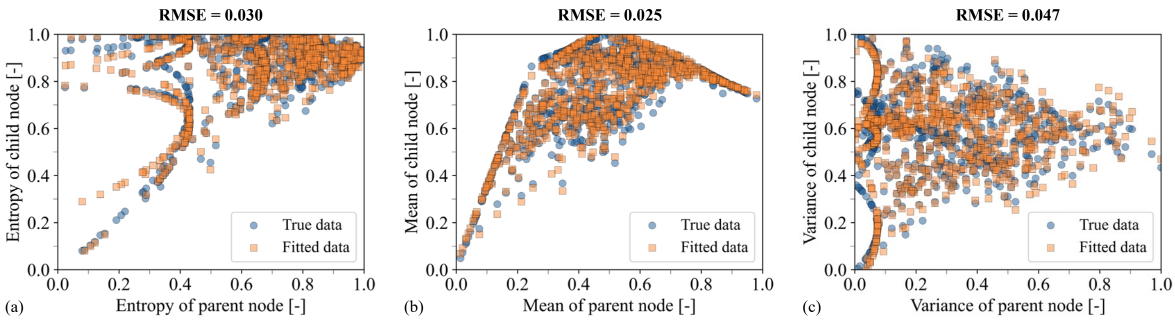

Figure 4: Fitting results for a representative BN using the proposed formulation: (a) entropy; (b) mean, and (c) variance.

Table 1: Fitting performance for entropy across various numbers of discrete states.

Number of discrete states	2	3	4	5	6	7
Average R^2	1.000	0.984	0.875	0.817	0.759	0.740
Average RMSE	0.000	0.005	0.025	0.027	0.024	0.023

Table 2: Fitting performance for mean across various numbers of discrete states.

Number of discrete states	2	3	4	5	6	7
Average R^2	1.000	1.000	0.911	0.845	0.748	0.700
Average RMSE	0.000	0.000	0.048	0.060	0.075	0.077

Table 3: Fitting performance for variance across various numbers of discrete states.

Number of discrete states	2	3	4	5	6	7
Average R^2	1.000	0.995	0.875	0.802	0.677	0.678
Average RMSE	0.000	0.009	0.060	0.066	0.083	0.082

4 GENERALIZE FORMULATION TO COMPLEX BN TOPOLOGIES

The above results demonstrate the robust performance of the proposed formulation in capturing probabilistic dependencies within a two-node BN. Building upon this foundation, this section extends the formulation to a complex BN topology: the common effect structure. The common effect structure describes a configuration in which multiple parent nodes jointly influence a single child node. This topology is frequently used in large-scale SDTs to model scenarios where an outcome results from the interaction of several contributing factors, for example, using data from adjacent traffic monitoring sites to estimate the failure consequences of a bridge [14,15]. The common effect structure is a primary contributor to the high memory demand of BNs, as the number of parameters in the CPTs grows exponentially with the number of parent nodes. Consequently, this structure is selected to illustrate the capability of the proposed formulation to improve computational efficiency.

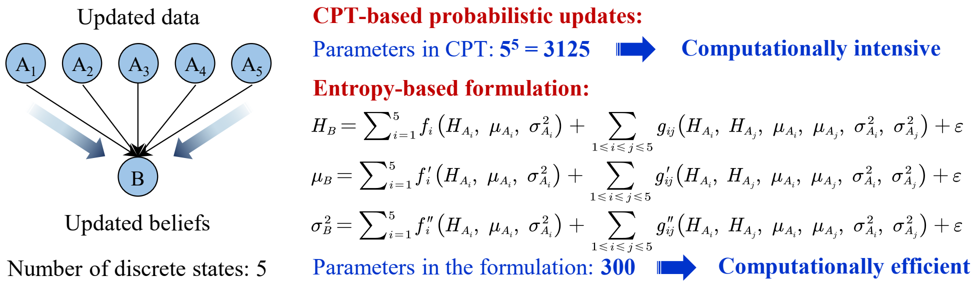

Figure 5: Comparison between CPT-based probabilistic updates and the proposed entropy-based formulation in a common effect structure.

Herein, a BN with six nodes is constructed to instantiate the common effect structure, wherein five parent nodes are connected to a single child node, as shown in Figure 5. Each node is discretized into five distinct states. Accordingly, the marginal probability distribution of each parent node has five parameters, while the conditional probability distribution of the child node has $5^5 = 3125$ parameters, leading to substantial memory demands. The proposed formulation can significantly alleviate this issue. Specifically, the entropy, mean, and variance of the child node can be obtained from analytic expressions that utilize the corresponding features of the parent nodes as inputs.

Unlike the two-node BN, where the child node is influenced by a single parent, the modeling of a common effect structure needs to capture both the individual contributions of each parent node and the interactions among them. As such, the entropy, mean, and variance of the child node are formulated as a combination of individual contribution terms and interaction terms:

$$H_B = \sum_{i=1}^{5} f_i\left(H_{A_i}, \mu_{A_i}, \sigma_{A_i}^2\right) + \sum_{1\le i\le j\le 5} g_{ij}(H_{A_i}, H_{A_j}, \mu_{A_i}, \mu_{A_j}, \sigma_{A_i}^2, \sigma_{A_j}^2) + \varepsilon \quad (10)$$

$$\mu_B = \sum_{i=1}^{5} f_i'\left(H_{A_i}, \mu_{A_i}, \sigma_{A_i}^2\right) + \sum_{1\le i\le j\le 5} g_{ij}'(H_{A_i}, H_{A_j}, \mu_{A_i}, \mu_{A_j}, \sigma_{A_i}^2, \sigma_{A_j}^2) + \varepsilon \quad (11)$$

$$\sigma_B^2 = \sum_{i=1}^{5} f_i''\left(H_{A_i}, \mu_{A_i}, \sigma_{A_i}^2\right) + \sum_{1\le i\le j\le 5} g_{ij}''(H_{A_i}, H_{A_j}, \mu_{A_i}, \mu_{A_j}, \sigma_{A_i}^2, \sigma_{A_j}^2) + \varepsilon \quad (12)$$

where $f_i(\cdot)$, $f_i'(\cdot)$, and $f_i''(\cdot)$ denote the individual contributions of each parent node to the entropy, mean, and variance of the child node, respectively. Their functional forms are detailed in Equations 7-9. $g_{ij}(\cdot)$, $g_{ij}'(\cdot)$, and $g_{ij}''(\cdot)$ represent the contributions of interactions between different parent nodes to the entropy, mean, and variance of the child node, respectively. For entropy fitting, interaction terms are constructed from the products of the entropies, means, and variances of all pairs of parent node. For the fitting of mean and variance, additional interaction terms are included, encompassing the products between the entropy of one parent node and the mean or variance of another, as well as the products between the means and variances of different parent nodes.

To evaluate the generalizability and robustness of the proposed formulation across diverse parameter configurations, 100 BNs are created with randomly sampled CPTs. For each BN, a set of prior probability distributions is assigned to parent nodes. Bayesian inference is then conducted to obtain the posterior probability distribution of the child node. This yields a dataset consisting of paired samples of prior distributions for the five parent nodes and the corresponding posterior distributions for the child node. Subsequently, the entropy, mean, and variance are extracted from each probability distribution. Regression analysis is then performed using the functional forms defined in Equations 10-12 to derive analytic expressions that capture uncertainty propagation within the common effect structure. Generally, the proposed formulation demonstrates strong predictive capability in capturing the statistical features of the child node. Across 100 BNs, the average RMSE values for entropy, mean, and variance predictions are 0.082, 0.059, and 0.062, respectively. Figures 6-8 show several representative fitting results for the entropy, mean, and variance of the child node, respectively.

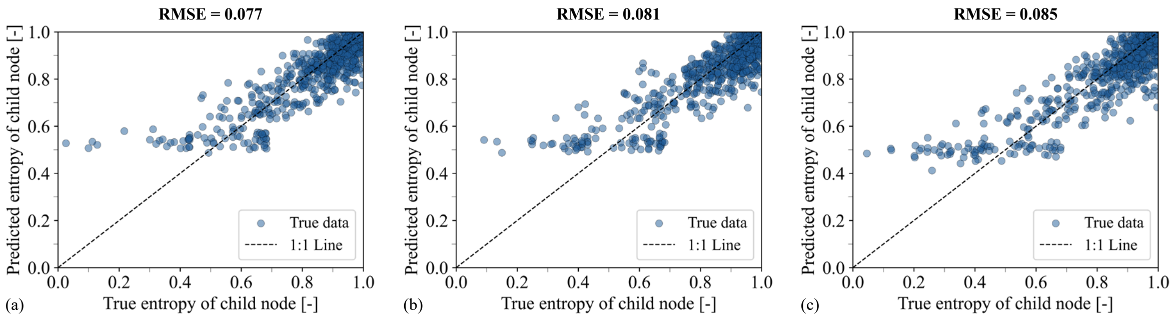

Figure 6: Representative fitting results for the entropy of the child node within the common effect structure using the proposed formulation.

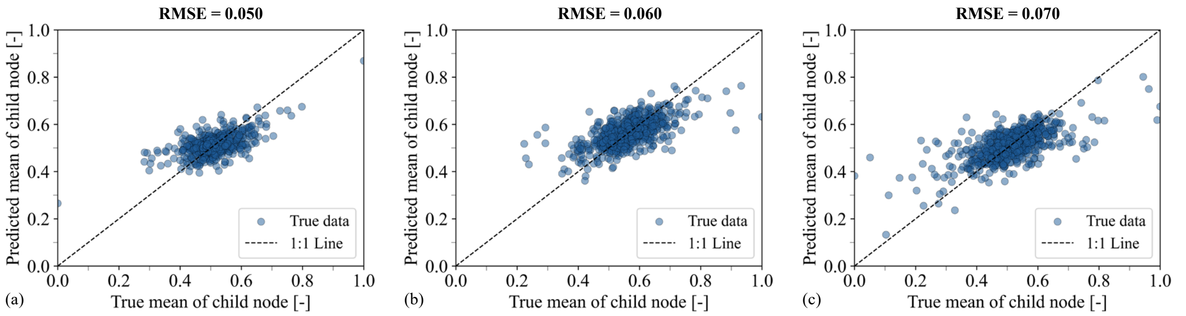

Figure 7: Representative fitting results for the mean of the child node within the common effect structure using the proposed formulation.

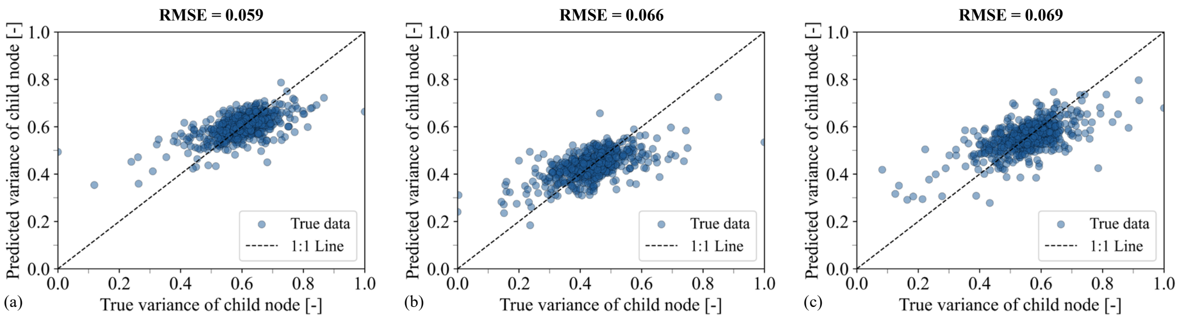

Figure 8: Representative fitting results for the variance of the child node within the common effect structure using the proposed formulation.

The improvement in storage efficiency of the proposed formulation compared to CPTs is demonstrated through two scenarios: (a) five parent nodes are connected to a single child node, wherein the number of discrete states per node varies from 3 to 8; and (b) the number of parent nodes in the common effect structure varies from 3 to 8, with each node having five discrete states. For the first scenario, as the number of discrete states increases, the number of parameters in the CPTs grows exponentially, leading to an explosive rise in storage demand, as summarized in Table 4. In contrast, the number of parameters in the proposed formulation remains constant at 300 across all discretization levels, highlighting its storage efficiency in fine discretization granularity. For the second scenario, the number of parameters in CPTs increases exponentially with the number of parent nodes. Particularly, when eight parent nodes are connected to one child node, the parameter count in CPTs becomes prohibitively large. By comparison, the number of parameters in the proposed formulation increases linearly with the number of parent nodes, demonstrating its superior scalability for BN modeling.

Table 4: Comparison of storage requirements between CPTs and the proposed formulation under varying numbers of node discrete states in a common effect structure (five parent nodes connected to one child node).

Number of **discrete states**	3	4	5	6	7	8
Parameters in CPTs	$3^5 = 243$	$4^5 = 1024$	$5^5 = 3125$	$6^5 = 7776$	$7^5 = 16{,}807$	$8^5 = 32{,}768$
Parameters in formulation	300	300	300	300	300	300

Table 5: Comparison of storage requirements between CPTs and the proposed formulation under varying numbers of parent nodes in a common effect structure (each node has five discrete states).

Number of **parent nodes**	3	4	5	6	7	8
Parameters in CPTs	$5^3 = 125$	$5^4 = 625$	$5^5 = 3125$	$5^6 = 15{,}625$	$5^7 = 78{,}125$	$5^8 = 390{,}625$
Parameters in formulation	135	210	300	405	525	660

5 CONCLUSIONS

This paper proposes an entropy-based formulation for Bayesian networks (BNs) to tackle the storage and computational challenges in modeling large-scale system digital twins (SDTs). The formulation characterizes the probability distributions of network nodes using three statistical features: entropy, mean, and variance. By deriving and fitting analytic expressions that capture the probabilistic dependencies between connected nodes, the proposed formulation can improve computational efficiency of BNs while preserving predictive accuracy. The main conclusions are summarized as follows:

- The analytic expressions are derived from a simple two-node BN structure. The fitting results show that the proposed formulation effectively captures the entropy of the child node, achieving average R^2 values exceeding 0.81 and mean RMSE values below 0.03 across 100 randomly sampled BNs.
- The proposed formulation is further extended to model the common effect structure, where five parent nodes are connected to one child node. The results show that the proposed formulation achieves an average RMSE of 0.082 in entropy prediction across 100 BNs, demonstrating its robust generalization capability to complex network topologies.
- The proposed formulation significantly improves storage efficiency of BNs under varying network complexities. Specifically, the number of parameters in the proposed formulation remains constant across various discretization levels and increases linearly with the number of parent nodes, in contrast to the exponential growth in conditional probability tables (CPTs).

REFERENCES

[1] VanDerHorn, E., & Mahadevan, S. Digital Twin: Generalization, characterization and implementation. *Decis Support Syst* (2021) **145**: 113524.
[2] Cheng, M., and Gao, H. O. *Data-driven life-cycle risk assessment of bridge networks using Bayesian network*. In Life-Cycle of Structures and Infrastructure Systems, pp. 1953-1960. CRC Press, (2023).
[3] Cheng, M., Shah, S. M., Nanni, A., & Gao, H. O. Automated knowledge graphs for complex systems (AutoGraCS): Applications to management of bridge networks. *Resilient Cities Struct*, (2024) **3.4**: 95-106.
[4] Geng, Z., Zhang, C., Jiang, Y., Pugliese, D., & Cheng, M. Integrating multi-source data for life-cycle risk assessment of bridge networks: a system digital twin framework. *J Infrastruct Preserv Resil*, (2025) **6.1**: 9.
[5] Pourret, O., Na, P., & Marcot, B. *Bayesian networks: a practical guide to applications*. John Wiley & Sons, (2008).
[6] Schulte, O., Frigo, G., Greiner, R., Luo, W., & Khosravi, H. *A new hybrid method for Bayesian network learning with dependency constraints*. In 2009 IEEE Symposium on Computational Intelligence and Data Mining, pp. 53-60. IEEE, (2009).
[7] Scutari, M., Graafland, C. E., & Gutiérrez, J. M. Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. *Int J Approx Reason* (2019) **115**: 235-253.
[8] Tsamardinos, I., Brown, L. E., & Aliferis, C. F. The max-min hill-climbing Bayesian network structure learning algorithm. *Mach learn* (2006) **65**: 31-78.
[9] Laitila, P., & Virtanen, K. Improving construction of conditional probability tables for ranked nodes in Bayesian networks. *IEEE Trans Knowl Data Eng* (2016) **28.7**: 1691-1705.
[10] Petitjean, F., Buntine, W., Webb, G. I., & Zaidi, N. Accurate parameter estimation for

Bayesian network classifiers using hierarchical Dirichlet processes. *Mach learn* (2018) **107.8**: 1303-1331.
[11] Zurek, W. H. Algorithmic randomness and physical entropy. *Phys Rev A* (1989) **40.8**: 4731.
[12] Namdari, A., & Li, Z. A review of entropy measures for uncertainty quantification of stochastic processes. *Adv Mech Eng* (2019) **11.6**: 1687814019857350.
[13] Vrieze, S. I. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol methods* (2012) **17.2**: 228.
[14] Dong, Y., Frangopol, D. M., & Saydam, D. Time - variant sustainability assessment of seismically vulnerable bridges subjected to multiple hazards. *Earthq Engng Struct Dyn* (2013) **42.10**: 1451-1467.
[15] Cheng, M., & Frangopol, D. M. Efficient scenario analysis for optimal adaptation of bridge networks under deep uncertainties through knowledge transfer. *Struct Saf* (2023) **100**: 102278.