Sustainable Diffusion-Based Incentive Mechanism for Generative AI-Driven Digital Twins in Industrial Cyber-Physical Systems

Jinbo Wen , *Graduate Student Member, IEEE*, Jiawen Kang , *Senior Member, IEEE*, Dusit Niyato , *Fellow, IEEE*, Yang Zhang , *Member, IEEE*, and Shiwen Mao , *Fellow, IEEE*

Abstract—Industrial Cyber-Physical Systems (ICPSs) are an integral component of modern manufacturing and industries. By digitizing data throughout product life cycles, Digital Twins (DTs) in ICPSs enable a shift from current industrial infrastructures to intelligent and adaptive infrastructures. Thanks to data process capability, Generative Artificial Intelligence (GenAI) can drive the construction and update of DTs to improve predictive accuracy and prepare for diverse smart manufacturing. However, mechanisms that leverage Industrial Internet of Things (IIoT) devices to share sensing data for DT construction are susceptible to adverse selection problems. In this paper, we first develop a GenAI-driven DT architecture in ICPSs. To address the adverse selection problem caused by information asymmetry, we propose a contract theory model and develop a sustainable diffusion-based soft actor-critic algorithm to identify the optimal feasible contract. Specifically, we leverage dynamic structured pruning techniques to reduce parameter numbers of actor networks, allowing sustainability and efficient implementation of the proposed algorithm. Numerical results demonstrate the effectiveness of the proposed scheme and the algorithm, enabling efficient DT construction and updates to monitor and manage ICPSs.

Index Terms—Contract theory, digital twins (DTs), generative AI, industrial cyber-physical systems, sustainable diffusion models.

I. INTRODUCTION

WITH the advancement of industrial technologies, such as the Industrial Internet of Things (IIoT) and information communication technology, the convergence of physical and cyber spaces gives rise to a new paradigm called Industrial Cyber-Physical Systems (ICPSs) [1]. ICPSs are intricate and intelligent systems that integrate physical and computational components seamlessly, enabling real-time data exchange and decision-making [2]. Predictive maintenance plays an increasingly pivotal role in the sustainability of ICPSs. The emergence of Artificial Intelligence (AI)-based approaches for predictive maintenance and monitoring has become prominent, such as transfer learning [3] and Federated Learning (FL) [4]. To design integrated control and monitoring systems in ICPSs, Digital Twins (DTs) have received significant attention from academia and industry [1]. DTs refer to virtual replicas that cover the life cycle of physical entities. In ICPSs, through simulating the behavior and performance of industrial infrastructures based on real-time sensing data, DTs can offer predictive capabilities and provide insights to eliminate physical mistakes and attacks [1], thus optimizing manufacturing processes and improving the performance of ICPSs. Nonetheless, the establishment and implementation of DTs in ICPSs depend on high-fidelity modeling and strong data interaction, which may meet data issues such as scarcity, bias, and noise.

Generative AI (GenAI) is a branch of AI technology that identifies the structures and patterns from existing data to generate various and original content [5]. The popular class of GenAI applications has emerged from foundation models, such as GPT-3 and Stable Diffusion, which are trained on vast quantities of data by leveraging different learning approaches. For example, ChatGPT is trained on a large corpus of text from diverse sources. This process enables it to acquire knowledge

Jinbo Wen and Yang Zhang are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: jinbo1608@nuaa.edu.cn; yangzhang@nuaa.edu.cn).

Jiawen Kang is with the School of Automation, Guangdong University of Technology, Guangzhou 510006, China (e-mail: kavinkang@gdut.edu.cn).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

of linguistic patterns and structures, thus automatically generating valuable content based on the prompts provided by users. Relying on its incredible data processing and generation capabilities, GenAI technology has great potential to revolutionize various domains. For example, GenAI can drive the progression of modern IoT and enable more adaptive and intelligent IoT applications [5], such as virtual assistants and smart surveillance.

GenAI also provides a novel tool for DT innovation [6], [7]. With the help of GenAI technology, the construction, maintenance, and optimization of DTs can be facilitated. In addition, researchers used GenAI technology to enhance DT emulation, feature abstraction, and decision-making modules [7], driving innovation across diverse applications, especially ICPSs. For instance, in smart manufacturing, GenAI is capable of aiding the creation of adaptive DTs specifically designed for the manufacturing environment, which enhances efficiency in both production scheduling and real-time control. Although GenAI holds the potential to support emerging applications of DTs in ICPSs, there are challenges ahead:

- *Challenge I:* Through extensive literature review, there is no research conducted on the utilization of GenAI to drive DTs within ICPSs. Hence, it is imperative to actively explore GenAI-driven DT architectures in ICPSs, thus improving system performance and sustainability.
- *Challenge II:* GenAI models can synthesize supplementary data based on real-world sensing data for DT construction. However, due to adverse selection problems caused by information asymmetry, IIoT devices may not be willing to share high-quality sensing data for DT construction without reliable incentives [8].
- *Challenge III:* It is a complex and high-dimensional problem to incentivize IIoT devices for DT construction [1], [8]. Generative Diffusion Models (GDMs) as an advanced GenAI model show superior performance in solving high-dimensional decision-making problems [9], however, often at the cost of substantial computational overhead during training [10]. Therefore, it is necessary to develop sustainable GDM-based algorithms to solve high-dimensional optimization problems in ICPSs.

Some studies have been conducted to design incentive mechanisms for DT construction [8], [11]. However, there are no studies directly addressing the incentive mechanism design for DT construction in ICPSs. Moreover, GDMs for solving high-dimensional optimization problems still face substantial computational overhead because of the need for iterative denoising in generating new samples, and no work develops sustainable GDMs for optimal incentive design to mitigate the environmental impact of model training. To address Challenge I, we first design a GenAI-driven DT architecture in ICPSs. To address Challenge II, we propose a contract theory model to motivate IIoT devices to share sensing data for DT construction, where contract theory as an economic tool has been widely used to address incentive design problems with information asymmetry [12]. To address Challenge III, we develop a sustainable diffusion model-based algorithm to find the optimal feasible contract while mitigating the environmental impact for ICPSs. The contributions of this paper are summarized as follows:

- We propose a GenAI-driven DT architecture in ICPSs. In particular, we systematically study how GenAI drives the DT construction pipeline in ICPSs, including the collection of real-time physical data, communications among DTs and between DTs and physical counterparts, DT modeling and maintenance, and DT decision-making, while also presenting the practical applications of these four parts (For Challenge I).
- To effectively alleviate the adverse selection problem caused by information asymmetry, we propose a contract theory model to motivate IIoT devices to participate in sensing data sharing for DT construction. Each IIoT device selects a suitable contract item that matches its type to obtain the highest utility, and the optimal feasible contract is achieved by maximizing the expected utility of the DT server (For Challenge II).
- To achieve sustainability for DT construction within ICPSs, we develop a sustainable diffusion-based Soft Actor-Critic (SAC) algorithm to generate the optimal contract under information asymmetry, where we apply dynamic structured pruning techniques to GDM-based networks for policy exploration in a more computationally efficient and scalable manner. *To the best of our knowledge, this is the first work that leverages sustainable diffusion models for incentive design* (For Challenge III).

The rest of the paper is organized as follows. Section II reviews the related work. In Section III, we introduce the GenAI-driven DT architecture in ICPSs. In Section IV, we propose the contract model to incentivize IIoT devices to share sensing data for DT construction. In Section V, we develop the sustainable diffusion-based SAC algorithm to sustainably generate the optimal contract. In Section VI, we present numerical results to demonstrate the effectiveness and sustainability of the proposed scheme and the algorithm. Section VII concludes the paper.

II. RELATED WORK

In this section, we discuss several related works, including GenAI-driven DTs, incentive mechanism design for DT construction, and structured pruning techniques.

A. Generative AI-Driven Digital Twins

GenAI possesses formidable capabilities to drive DTs from various domains [6], [7], [13], [14]. In [6], the authors explored the potential of GenAI-driven human DTs, including GenAI-enabled data acquisition, communication, data management, digital modeling, and data analytics. The authors in [7] proposed a GenAI-driven DT network architecture to realize intelligent external and internal closed-loop network management, where GenAI models can drive DT status emulation, feature abstraction, and network decision-making. In [13], the authors introduced how GenAI facilitates DT modeling and provided existing implementations of GenAI for DTs in drug discovery and clinical trials. In [14], the authors explored the applications of GenAI models such as transformers and GDMs to empower DTs from several perspectives, including physical-digital modeling, synchronization, and slicing capability. While the integration of
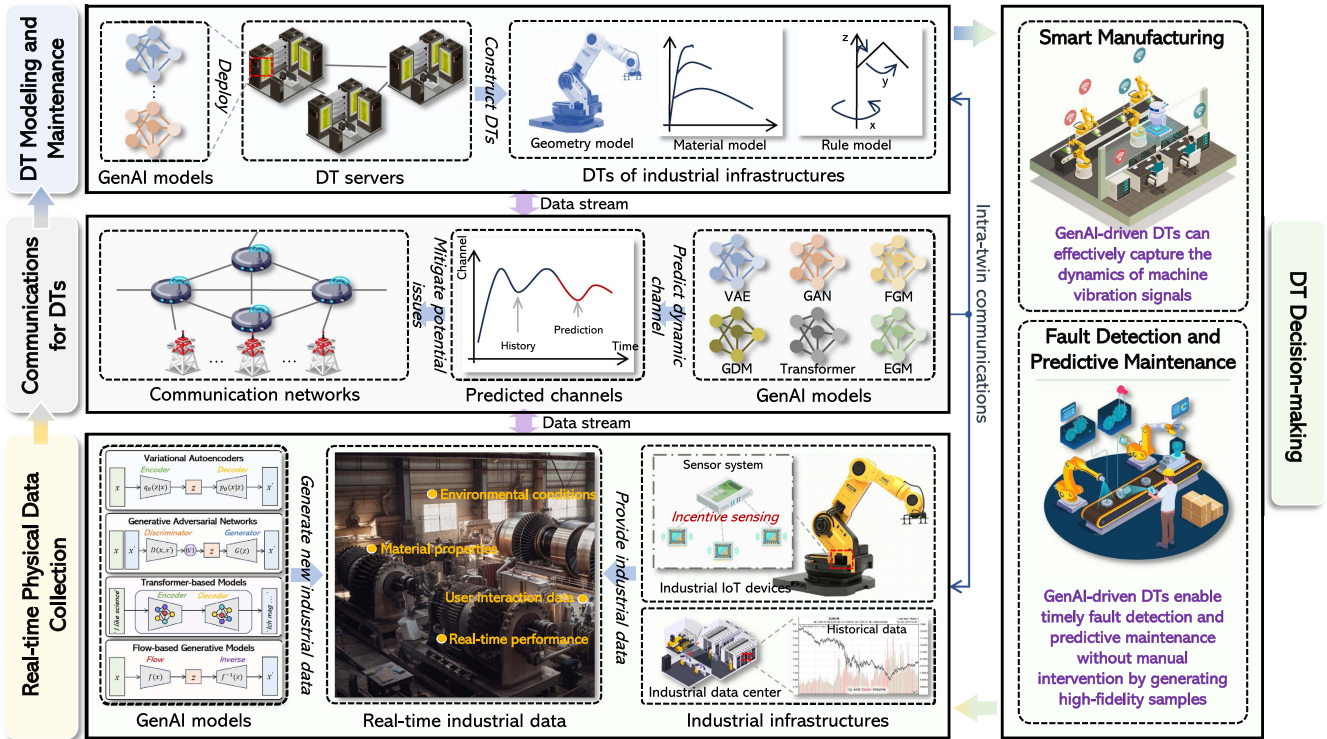
Fig. 1. The architecture of GenAI-driven DTs in ICPSs. We study how GenAI drives the DT construction pipeline, i.e., real-time physical data collection, communications for DTs, DT modeling and maintenance, and DT decision-making.

GenAI and DTs is capable of revolutionizing various sectors, there is no work systematically studying how GenAI can drive DTs in ICPSs.

B. Incentive Mechanism Design for Digital Twin Construction

Since DT construction requires high-quality data and intensive resources, incentive mechanisms play a crucial role in DT construction by promoting data sharing and resource collaboration. A few works have been conducted to design incentive mechanisms for DT construction [8], [11]. In [8], the authors proposed an iterative contract design to motivate IoT devices to share data for DT construction and used a multi-agent reinforcement learning algorithm to solve the formulated contract problem. In [11], the authors proposed a DT edge network framework for flexible and secure DT construction. To efficiently construct DTs, the authors also proposed an iterative double auction-based joint cooperative FL and local model update verification scheme. Incentive mechanism design for ensuring high-quality DT construction is critical, and this topic is still worth investigating, especially in ICPSs. However, none of the existing studies have used GenAI models to discover the optimal incentive design for DT construction.

C. Structured Pruning Techniques

Structured pruning techniques focus on eliminating parameters and substructures from networks [10], which has been used to effectively optimize Deep Reinforcement Learning (DRL)

algorithms. For example, the authors in [15] proposed a tiny multi-agent DRL algorithm to obtain the optimal game solution, which leverages the structured pruning technique to reduce the parameter number of the actor-critic network. The authors in [16] proposed a dynamic structured pruning approach to remove the unimportant neurons of DRL models during the training stage. In [10], the authors proposed a dedicated method for compressing diffusion models by utilizing Taylor expansion. Diffusion model-based DRL algorithms have been proven to be an effective method to solve network optimization problems [9], [17]. This algorithm uses multi-layer perceptions for iterative denoising from the Gaussian noise [9], which consumes a lot of energy. However, there is no research on developing efficient diffusion model-based DRL algorithms for network optimization. Therefore, we, for the first time, apply structured pruning techniques to diffusion model-based DRL algorithms for optimal incentive design, making them more efficient and sustainable.

III. GENERATIVE AI-DRIVEN DIGITAL TWIN ARCHITECTURE

In this section, we introduce the proposed GenAI-driven DT architecture in ICPSs, which systematically studies the role of GenAI on the DT construction pipeline, as shown in Fig. 1.

A. Real-Time Physical Data Collection

Real-time data collection from the physical space is the first step in constructing DTs in ICPSs [18]. The required data, such as historical equipment parameters and real-time operation, are

collected by industrial infrastructures such as IIoT devices and the industrial data center [1], where real-time sensing data allows DTs to continuously mirror the current state of the physical system, and historical data can be used for training predictive models within DTs. However, due to the heterogeneity of diverse sensing sources, some real-world data may be insufficient and confidential [1], affecting the rendering quality of DTs. Fortunately, GenAI supports data generation by capturing data distributions from the actual physical system [5], which can provide ultra-realistic industrial data based on the collected data to improve DT training. For instance, the authors in [14] utilized a Generative Adversarial Network (GAN)-based model to generate a sufficient dataset, which is used to train DT models.

B. Communications for Digital Twins

The key to DT construction is to send the collected data in real time to the server where DTs are deployed [18]. Communications among DTs and between DTs and physical counterparts act as a bridge for data transmission between physical and virtual spaces [6], such as collected data transmission and DT feedback. However, due to the potentially harsh industrial environments, the communication channel conditions in ICPSs may dynamically change [1], causing reliability issues in DT communications. Predictive DT channels provide an effective solution that enables DTs to maintain high fidelity in real-time communications [19]. Fortunately, GenAI can generate synthetic data representing various channel conditions, and the generated data and raw data can be used together to train deep learning models for predicting the DT communication channel dynamics, thus mitigating potential issues such as communication disruptions and ensuring efficient and reliable DT communications. For example, a recent study in [19] proposed a convolutional time-series GAN to generate synthetic data based on original channel data, which can enhance the accuracy of ConvLSTM-based predictive DT channels in dynamic channel conditions.

C. Digital Twin Modeling and Maintenance

Based on collected data, high-fidelity DTs can be constructed by reproducing the geometry and physical properties of industrial infrastructures [18]. These virtual models are maintained on edge servers to ensure real-time updates [18]. However, due to the complicated environments of ICPSs [1], traditional DT modeling methods, such as structural modeling or behavioral modeling [18], rely on accurate simulation parameters [6], making it challenging to continuously guarantee the quality of DTs. Thanks to the robust data generation capability, GenAI can provide supplementary data for DT modeling in ICPSs, including environmental conditions and operational parameters, and even allow DTs to update and refine their knowledge base over time. For instance, the authors in [20] developed a Wasserstein GAN-based DT in industrial asset health monitoring. Besides, the augmented data can preserve almost all patterns from the original data [6], thus protecting private data in ICPSs.

D. Digital Twin Decision-Making

After achieving the virtual-real mapping, DTs can timely analyze real-world data and provide decision-making results to ICPSs [18]. GenAI-driven DTs enable intelligent and reliable applications, such as smart manufacturing [5], intelligent assessment [21], and fault detection and predictive maintenance [22], [23]. For example, in smart manufacturing [5], GenAI-driven DTs can forecast product life in a timely manner and evaluate manufacturing plans based on generated fault diagnosis. Then, manufacturers can react according to the feedback conveyed by DTs. In the intelligent assessment of gear surface degradation [21], by generating synthetic data that represents surface defects, GenAI-driven DTs can recognize different stages and types of gear wear with higher precision. Besides, in fault detection and predictive maintenance [22], by combining GenAI and domain adversarial graph networks, DTs can better diagnose bearing faults by learning domain-invariant features.

IV. CONTRACT MODELING

In ICPSs, IIoT devices equipped with a set of sensors can collect geospatial data from the surrounding environment and send the collected data to DT servers for DT construction [8]. If the data quality is low, the created DTs may not accurately reflect the real-time dynamics of industrial infrastructures. Thus, high-quality sensing data is critical for DT construction. However, due to the high cost of collecting sensing data in ICPSs [1], IIoT devices may not actively provide high-quality sensing data to DT servers, or even maliciously provide harmful data [12], leading to adverse selection problems caused by information asymmetry [8]. To this end, we use contract theory to motivate IIoT devices to provide high-quality sensing data for DT construction, where contract theory as a powerful economy tool aims at addressing information asymmetry [8], [12]. In this paper, we consider a DT server and M IIoT devices for DT construction.

A. Utility of IIoT Devices

We denote $\Psi = \{\psi_k | \psi_k \in \mathbb{N}_+, \ k \in \{1, \ldots, K, \ K \leq M\}\}$ as the set of different sensing and communication levels of IIoT devices [8], where the level is mainly impacted by equipment and algorithms used by IIoT devices for data sensing and transmission [24]. While the sensing and communication levels of IIoT devices are not fully disclosed due to information asymmetry [8], the DT server can leverage statistical distributions of IIoT devices from historical data to group IIoT devices into discrete types [12]. Thus, we consider sensing and communication levels as types and sort the types in ascending order, i.e., $0 < \psi_1 \leq \cdots \leq \psi_K$, where K is the number of types of IIoT devices. Especially, IIoT devices with higher sensing and communication levels are characterized as higher types, which can provide better sensing data. Based on [8], the utility of an IIoT device with type ψ_k to send sensing data with size $\hat{s}_k$ to the DT server is defined as

$$\tilde{u}_k(\hat{s}_k, r_k) = \rho \psi_k r_k - c\hat{s}_k - c_0, \tag{1}$$

where ρ is a pre-defined weight parameter about the incentive, r_k is the received reward associated with the provided data $\hat{s}_k$, c is the unit cost related to the collection, computation, and transmission of sensing data $\hat{s}_k$ [12], and c_0 is an additional cost involving energy consumption [8].

B. Utility of Digital Twin Server

Considering the more high-quality data provided by IIoT devices, the higher satisfaction of the DT server, we adopt the β-fairness function $g(\hat{s}_k)$ to quantify the satisfaction of the DT server [8], given by

$$g(\hat{s}_k) = \frac{1}{1-\beta}\hat{s}_k^{1-\beta}, \tag{2}$$

where $0 \leq \beta < 1$ is a pre-defined constant. Since the DT server just knows the number of IIoT devices and the distribution of each type due to information asymmetry, the overall utility of the DT server is given by

$$U(\hat{\boldsymbol{s}}, \boldsymbol{r}) = M \sum_{k=1}^{K} q_k(\vartheta g(\hat{s}_k) - r_k), \tag{3}$$

where $\vartheta > 0$ is the unit revenue for the satisfaction of the DT server [12], q_k represents the probability that IIoT devices belong to type ψ_k, and $\hat{\boldsymbol{s}} = [\hat{s}_k]_{1 \times K}$ and $\boldsymbol{r} = [r_k]_{1 \times K}$ represent the vectors of high-quality data volume sizes and rewards, respectively.

C. Contract Formulation

To maximize the overall utility, the DT server designs a contract consisting of a group of contract items. The contract is denoted as $\Omega = \{(\hat{s}_k, r_k), k \in \{1, \dots, K\}\}$, where the more $\hat{s}_k$, the higher r_k. To ensure that each IIoT device selects the most suitable contract item to maximize its utility, the feasible contract should satisfy the Individual Rationality (IR) and Incentive Compatibility (IC) constraints [8], [12]:

Definition 1 (Individual Rationality): Type-k IIoT devices gain non-negative utilities by selecting the contract item $(\hat{s}_k, r_k)$ that matches their types, i.e.,

$$\tilde{u}_k(\hat{s}_k, r_k) = \rho\psi_k r_k - c\hat{s}_k - c_0 \geq 0, \ \forall k \in \{1, \dots, K\}, \tag{4}$$

Definition 2 (Incentive Compatibility): Any type of IIoT device prefers to choose the contract item $(\hat{s}_k, r_k)$ designed for it rather than any other contract item $(\hat{s}_n, r_n)$, $n \in \{1, \dots, K\}$, and $n \neq k$, i.e.,

$$\rho\psi_k r_k - c\hat{s}_k \geq \rho\psi_k r_n - c\hat{s}_n, \ \forall k, n \in \{1, \dots, K\}, \ k \neq n. \tag{5}$$

The IR constraints (4) guarantee that the utility of IIoT devices is non-negative, and the IC constraints (5) guarantee that each IIoT device can achieve the highest utility by selecting the optimal contract item designed for its type. The DT server aims to maximize its overall utility. Based on the above IC and IR constraints, the optimization problem is expressed as

$$\max_{\hat{\boldsymbol{s}}, \boldsymbol{r}} U(\hat{\boldsymbol{s}}, \boldsymbol{r})$$

$$\text{s.t. } \tilde{u}_k(\hat{s}_k, r_k) \geq 0, \ \forall k \in \{1, \dots, K\},$$

$$\tilde{u}_k(\hat{s}_k, r_k) \geq \tilde{u}_k(\hat{s}_n, r_n), \ \forall n, k \in \{1, \dots, K\}, \ n \neq k,$$

$$\hat{s}_k \geq 0, r_k \geq 0, \psi_k > 0, \ \forall k \in \{1, \dots, K\}. \tag{6}$$

Remark 1: In the optimization problem (6), there are K IR constraints and $K(K-1)$ IC constraints, which are all non-convex. Thus, it is difficult to directly solve the problem (6). The standard approach is to reduce IR and IC constraints by defining some lemmas and use heuristic and traditional optimization algorithms to solve the derived problem [25]. However, these approaches add difficulty to the problem formulation, which is time-consuming to adjust and prove the corresponding lemmas [8]. Besides, states of IIoT devices and the DT server change dynamically due to the heterogeneity and dynamics of ICPSs [1], and traditional approaches need to be re-designed in practice. Furthermore, these approaches may not be applicable in practical scenarios when the environment lacks complete and accurate information. Thus, traditional mathematical techniques may not effectively solve the problem (6) in practice. DRL algorithms have been widely used to solve optimization problems because they can learn complex policies in high-dimensional spaces, which can find near-optimal solutions to challenging problems where traditional methods might struggle. In particular, GDMs possess a robust generative capability and have been used to enhance DRL algorithms, thus being able to find an optimal policy for dynamic and high-dimensional optimization problems [9]. For efficient implementation in ICPSs, we propose a sustainable GDM-based DRL algorithm to solve (6).

V. SUSTAINABLE DIFFUSION-BASED SOFT ACTOR-CRITIC ALGORITHMS FOR OPTIMAL CONTRACT DESIGN

In this section, we propose the sustainable diffusion-based SAC algorithm to efficiently generate the optimal feasible contract under information asymmetry, where we apply dynamic structured pruning techniques to GDM-based networks to reduce the training cost and carbon emissions. The architecture of the proposed algorithm is shown in Fig. 2.

A. Generative Diffusion Models for Optimal Contract Design

Diffusion models work by corrupting the training data by continuously adding Gaussian noise and then learning to recover the data by reversing this noise process [10]. Through an iterative process of denoising the initial distribution, GDMs can generate the optimal contract $\Omega^0 = (\hat{s}_k^*, r_k^*), k \in \{1, \dots, K\}$. During iterations of T, Gaussian noise is gradually added to the initial contract Ω, emerging a series of contract samples $(\Omega^1, \Omega^2, \dots, \Omega^T)$. We define the environment during the optimal contract design as

$$\boldsymbol{e} \triangleq \{M, K, c, c_0, \vartheta, (q_1, \dots, q_K), (\psi_1, \dots, \psi_K)\}. \tag{7}$$

We denote the contract design policy as $\pi_{\boldsymbol{\theta}}(\Omega|\boldsymbol{e})$ with parameters $\boldsymbol{\theta}$, which is constituted by the GDM-based network, as shown in Fig. 2. Through the reverse process of a conditional
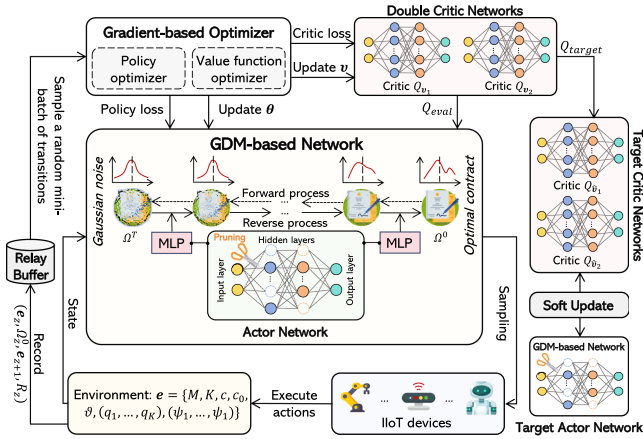
Fig. 2. The proposed algorithm architecture, where we utilize dynamic structured pruning techniques to sparsify the actor networks of the diffusion model. The goal of the algorithm is to design optimal contracts that motivate IIoT devices to provide sensing industrial data, which is one of the data sources for DT construction, as shown in Fig. 1.

diffusion model, the policy $\pi_{\boldsymbol{\theta}}(\Omega|\boldsymbol{e})$ can be expressed as [26]

$$\pi_{\boldsymbol{\theta}}(\Omega|\boldsymbol{e}) = p_{\boldsymbol{\theta}}(\Omega^{0:T}|\boldsymbol{e}) = \mathcal{N}(\Omega^T; \mathbf{0}, \mathbf{I}) \prod_{t=1}^{T} p_{\boldsymbol{\theta}}(\Omega^{t-1}|\Omega^t, \boldsymbol{e}), \quad (8)$$

where the end sample of the reverse chain is the selected contract Ω^0. $p_{\boldsymbol{\theta}}(\Omega^{t-1}|\Omega^t, \boldsymbol{e})$ as a noise prediction model is modeled as a Gaussian distribution $\mathcal{N}(\Omega^{t-1}; \boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t), \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t))$ with the covariance matrix expressed as $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t) = \delta_t \mathbf{I}$ [26], and the mean $\boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t)$ is given by

$$\boldsymbol{\mu}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t) = \frac{1}{\sqrt{\alpha_t}} \left(\Omega^t - \frac{\delta_t}{\sqrt{1 - \bar{\alpha}_t}} \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t) \right), \quad (9)$$

where $\delta_t \in (0, 1)$ is a hyperparameter for model training, $\alpha_t = 1 - \delta_t$, and $\bar{\alpha}_t = \prod_{j=0}^{t} \delta_j$ [9], [26]. $\boldsymbol{\epsilon}_{\boldsymbol{\theta}}$ is a deep model that generates contracts conditioned on the environment $\boldsymbol{e}$, as determined by the policy. We first sample $\Omega^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and then the selected contract can be sampled via the reverse diffusion chain parameterized by $\boldsymbol{\theta}$, given by

$$\Omega^{t-1}|\Omega^t = \frac{\Omega^t}{\sqrt{\alpha_t}} - \frac{\delta_t}{\sqrt{\alpha_t(1 - \bar{\alpha}_t)}} \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\Omega^t, \boldsymbol{e}, t) + \sqrt{\delta_t} \boldsymbol{\epsilon}. \quad (10)$$

To improve the policy $\pi_{\boldsymbol{\theta}}(\Omega|\boldsymbol{e})$, we introduce the Q-function [26], which can guide the reverse diffusion chain to preferentially sample contracts with high values. We build two critic networks $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$, target critic networks $Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$, and target policy $\hat{\pi}_{\hat{\boldsymbol{\theta}}}$. Then, we define the contract quality network as $Q_{\boldsymbol{v}}(\boldsymbol{e}, \Omega) = \min\{Q_{\boldsymbol{v}_1}(\boldsymbol{e}, \Omega), Q_{\boldsymbol{v}_2}(\boldsymbol{e}, \Omega)\}$. Thus, the optimal contract design policy aims at maximizing the expected cumulative reward, expressed as [9]

$$\pi = \arg\max_{\pi_{\boldsymbol{\theta}}} \mathbb{E}\left[\sum_{z=0}^{Z} \gamma^z (R(\boldsymbol{e}_z, \Omega_z) + \varsigma H(\pi_{\boldsymbol{\theta}}(\boldsymbol{e}_z))) \right], \quad (11)$$

where Z denotes the maximum training step, γ denotes the discount factor for future rewards, $R(\boldsymbol{e}_z, \Omega_z)$ represents the

immediate reward upon executing action Ω_z in state $\boldsymbol{e}_z$, ς is the temperature coefficient controlling the strength of the entropy, and $H(\pi_{\boldsymbol{\theta}}(\boldsymbol{e}_z)) = -\pi_{\boldsymbol{\theta}}(\boldsymbol{e}_z) \log \pi_{\boldsymbol{\theta}}(\boldsymbol{e}_z)$ is the entropy of the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{e}_z)$ [27]. To update the Q-function, the optimization of $\boldsymbol{v}_m$ for $m = \{1, 2\}$ aims at minimizing the objective function as [26]

$$\mathbb{E}_{(\boldsymbol{e}_z, \Omega_z, \boldsymbol{e}_{z+1}, R_z) \sim \mathcal{B}_z, \Phi_{z+1}^0 \sim \hat{\pi}_{\hat{\boldsymbol{\theta}}}} \left[\sum_{m=1,2} (R(\boldsymbol{e}_z, \Omega_z) \right.$$
$$\left. + \gamma^z (1 - d_{z+1}) \hat{\pi}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{e}_{z+1}) Q_{\hat{\boldsymbol{v}}}(\boldsymbol{e}_{z+1}) - Q_{\boldsymbol{v}_m}(\boldsymbol{e}_z, \Omega_z))^2 \right], \quad (12)$$

where $\mathcal{B}_z$ is a mini-batch of transitions with a size b sampled from the experience replay memory $\mathcal{D}$ and d_{z+1} is a 0-1 variable that represents the terminated flag [9].

B. Dynamic Structured Pruning

In terms of the network structure, both the policy and critic networks are multilayer perception networks, which consist of an input layer, multiple hidden layers, and an output layer. Considering a L-layer policy network, where parameters in the l-th fully-connected layer are denoted by $\boldsymbol{\theta}^{(l)}$, $l \in \{1, \ldots, L\}$, the output of the l-th layer is expressed as

$$\boldsymbol{h}^{(l)} = f^{(l)} \left(\boldsymbol{\theta}^{(l)} \boldsymbol{h}^{(l-1)} + b^{(l)} \right), \quad (13)$$

where $f^{(l)}$ is the activation function of the l-th layer and $b^{(l)}$ is the deviation at the l-th layer. We focus on pruning the redundant neurons and connected weights of policy networks without affecting the performance, which enhances the training efficiency of the policy network. To this end, we introduce a binary mask $m_i^{(l)}$ to represent the pruning state of each neuron $o_i^{(l)}$ at the l-th layer [16], where $m_i^{(l)} = 0$ indicates that the neuron $o_i^{(l)}$ should be pruned, and $m_i^{(l)} = 1$ indicates that the neuron $o_i^{(l)}$ should be reserved. Based on the binarized mask $\boldsymbol{m}^{(l)}$, the output of the l-th layer is rewritten as

$$\boldsymbol{h}^{(l)} = f^{(l)} \left(\boldsymbol{\theta}^{(l)} \boldsymbol{h}^{(l-1)} \odot \boldsymbol{m}^{(l)} + b^{(l)} \right), \quad (14)$$

where $\odot$ represents the Hadamard product. We use the policy gradient algorithm to update the contract design policy [9]. Specifically, the policy gradient with respect to the policy parameters $\boldsymbol{\theta}$ can be computed as the expectation over $\mathcal{B}_z$. Thus, the gradient is given by

$$\nabla_{\boldsymbol{\theta}_z} J(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{e}_z \sim \mathcal{B}_z}[-\nabla_{\boldsymbol{\theta}_z} \pi_{\boldsymbol{\theta}_z}(\boldsymbol{e}_z) Q_{\boldsymbol{v}_z}(\boldsymbol{e}_z)], \quad (15)$$

where $\boldsymbol{\theta}_z$ and $\boldsymbol{v}_z$ are the policy and critic parameters at the z-th training step, respectively. Thus, the policy parameters $\boldsymbol{\theta}$ are updated by performing gradient descent based on (15), which is expressed as [9], [16]

$$\boldsymbol{\theta}_{z+1}^{(l)} \leftarrow \boldsymbol{\theta}_z^{(l)} - \eta \nabla_{\boldsymbol{h}_z^{(l)} \odot \boldsymbol{m}_z^{(l)}} J(\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}_z^{(l)}} (\boldsymbol{h}_z^{(l)} \odot \boldsymbol{m}_z^{(l)}), \quad (16)$$

where $\eta \in (0, 1]$ is the learning rate of the actor network. For the target networks that have the same network structure as the online networks [9], the parameters of the target policy are also

performed dynamic structured pruning, and their parameters are updated by using a soft update mechanism, given by [26]

$$\hat{\boldsymbol{\theta}}_{z+1} \leftarrow \varepsilon \boldsymbol{\theta}_z + (1-\varepsilon)\hat{\boldsymbol{\theta}}_z,$$
$$\hat{\boldsymbol{v}}_{m,z+1} \leftarrow \varepsilon \boldsymbol{v}_{m,z} + (1-\varepsilon)\hat{\boldsymbol{v}}_{m,z}, \text{ for } m = \{1,2\}, \quad (17)$$

where $\varepsilon \in (0,1]$ is the update rate of the target network.

The pruning threshold plays an important role in the pruning decisions of network parameters [15], [16]. We adopt a dynamic pruning threshold Υ, which is given by [15]

$$\Upsilon_z = \sum_{l=1}^{L} \sum_{i=1}^{I} \phi_i^{(l)} \omega_z,$$
$$\omega_z = \hat{\omega} - \hat{\omega}\left(1 - \frac{z}{N\triangle}\right)^3, \quad (18)$$

where $\phi_i^{(l)}$ and N represent the importance of the i-th neuron of the l-th layer and the total number of pruning steps, respectively. $\triangle$ denotes the pruning frequency, and ω_z and $\hat{\omega}$ represent the current sparsity at the z-th training step and the target sparsity, respectively.

After determining the pruning threshold, neurons ranked below the threshold will be pruned [15], [16], and the binary masks used for pruning are updated, given by [15]

$$m_i^{(l)} = \begin{cases} 1, & \text{if } |m_i^{(l)}\theta_i^{(l)}| \geq \Upsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

Finally, we build a compact policy network according to the redundancy of the sparse policy network. The procedure of the proposed sustainable diffusion model is shown in Algorithm 1. The computational complexity of the proposed algorithm consists of three parts. In the initialization part, the computational complexity is $\mathcal{O}(|\boldsymbol{\theta}| + |\boldsymbol{v}|)$. In the training part, the computational complexity is $\mathcal{O}(Z(T\sum_{l=1}^{L-1}|\boldsymbol{\theta}^{(l)}| + \sum_{l=1}^{L-1}|\boldsymbol{h}^{(l)}|))$ [9], [15]. In the inference part, the computational complexity is $\mathcal{O}(|\hat{\boldsymbol{\theta}}|)$. Thus, the computational complexity of the proposed algorithm is around $\mathcal{O}(|\boldsymbol{\theta}| + |\boldsymbol{v}| + |\hat{\boldsymbol{\theta}}| + Z(T\sum_{l=1}^{L-1}|\boldsymbol{\theta}^{(l)}| + \sum_{l=1}^{L-1}|\boldsymbol{h}^{(l)}|))$, where T is the total number of denosing steps. In the following, we compare the proposed sustainable diffusion-based SAC algorithm with DRL algorithms that are commonly used for optimal contract design to evaluate its performance, i.e., Proximal Policy Optimization (PPO), SAC, diffusion-based SAC, and diffusion-based Deep Deterministic Policy Gradient (DDPG). We simply introduce these four algorithms and present their computational complexities for solving the optimization problem (6). Specifically,

- **PPO:** Update the policy by using a clipped surrogate objective [28]. Its computational complexity is $\mathcal{O}(\sum_{j=1}^{J} o_{j-1}o_j)$ [29], where o_j represents the number of neural units at the j-th layer in the policy network.
- **SAC:** Optimize the stochastic policy in an off-policy way, which maximizes a trade-off between expected return and entropy [27]. Its computational complexity is $\mathcal{O}(\sum_{j=1}^{J} o_{j-1}^Q o_j^Q + \sum_{i=1}^{I} o_{i-1}^{\pi} o_i^{\pi})$ [27], where o_j^Q

Algorithm 1: Sustainable Diffusion Models based on Dynamic Structured Pruning.

Input: Environment e and diffusion parameters.
Output: The optimal contract design Ω^0.

1 ## *Initialize*
2 Initialize relay buffer $\mathcal{D}$, policy network $\pi_{\boldsymbol{\theta}}$, critic networks $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$, target networks $\hat{\pi}_{\hat{\boldsymbol{\theta}}}, Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$, and binary masks $\boldsymbol{m}$.
3 ## *Training*
4 **for** the training step $z = 1$ to Z **do**
5   ### *Generating contracts*
6   Observe environment e_z and initialize a random process $\mathcal{N}$ for contract design exploration.
7   Set Ω_z^T as Gaussian noise and generate contract design Ω_z^0 by (10).
8   Execute Ω_z^0, observe the next environment e_{z+1}, and obtain reward R_z.
9   Store record $(e_z, \Omega_z^0, e_{z+1}, R_z)$ into $\mathcal{D}$.
10   Sample a random mini-batch of transitions $\mathcal{B}_z$ from replay buffer $\mathcal{D}$.
11   Update $Q_{\boldsymbol{v}_1}, Q_{\boldsymbol{v}_2}$ using $\mathcal{B}_z$ by (12).
12   ### *Dynamic pruning and fine-tuning*
13   Compute neuron importance ϕ_z based on [16].
14   Update the policy $\pi_{\boldsymbol{\theta}}$ using $\mathcal{B}_z$ by (16).
15   Update the target networks $\hat{\pi}_{\hat{\boldsymbol{\theta}}}, Q_{\hat{\boldsymbol{v}}_1}, Q_{\hat{\boldsymbol{v}}_2}$ by (17).
16   Compute the dynamic pruning threshold Υ for the policy network parameters $\boldsymbol{\theta}_z$ by (18).
17   Update binary masks $\boldsymbol{m}_z$ for $\boldsymbol{\theta}_z$ by (19).
18   **if** $\phi_i^{(l)} < \Upsilon$ **then**
19   Remove the i-th neuron at the l-th layer with the mask $m_i^{(l)}$ and associated parameters $\boldsymbol{\theta}_z^{(l)}$ from the policy network.
20   **end**
21 **end**
22 Reconstruct the compact policy networks.
23 ## *Inference*
24 Input the environment vector e.
25 Generate the optimal contract design Ω^0 based on the target policy $\hat{\pi}_{\hat{\boldsymbol{\theta}}}$ by (10).
26 **return** $\Omega^0 = \{(\hat{s}_k^*, r_k^*), 1 \leq k \leq K\}$.

---

and o_i^{π} represent the number of neuron units at the j-th layer of $Q_{\boldsymbol{v}}(e, \Omega)$ and the i-th layer of $\pi_{\boldsymbol{\theta}}(\Omega|e)$, respectively.

- **Diffusion-based SAC:** Use diffusion models to enhance SAC algorithms [9]. Its computational complexity is $\mathcal{O}(Z[CV + TC|\boldsymbol{\theta}| + (B+1)(|\boldsymbol{\theta}| + |\boldsymbol{v}|)])$ [9], where C represents the maximum number of collected transitions at each training step, V represents the complexity of interacting with the environment e, and B represents the batch size.
- **Diffusion-based DDPG:** Use diffusion models to enhance DDPG algorithms that learn the policy through the Q-learning function [30]. Its computational complexity is $\mathcal{O}(Z[T|\boldsymbol{\theta}| + |\boldsymbol{v}|])$ [30].
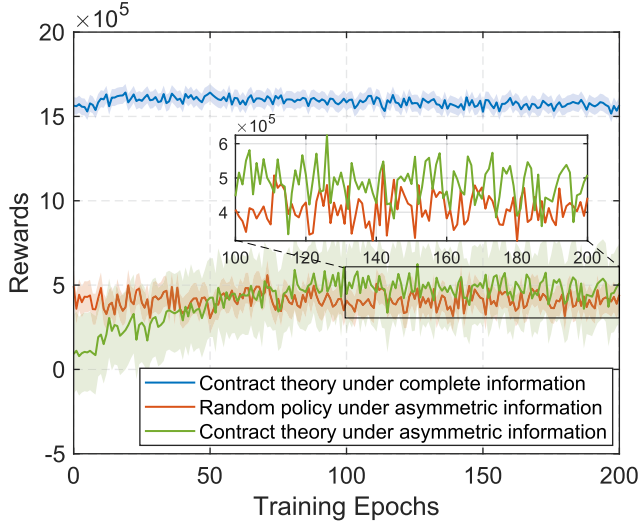
Fig. 3.   Test reward comparison of the proposed scheme with the random scheme under asymmetric information and the contract theory scheme under complete information.
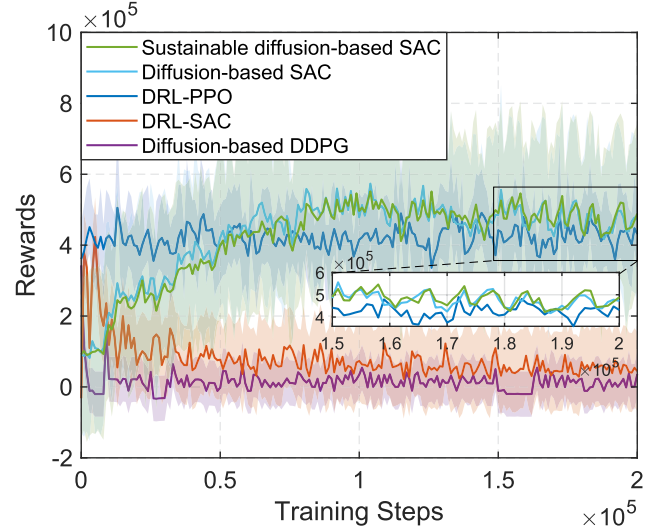

Fig. 4.   Performance comparison of the proposed algorithm with several DRL algorithms in optimal contract design. For the parameter settings of the proposed algorithm, we set the pruning rate to 10%, the diffusion step to 6, the learning rate of actor networks to 2×10^{-7}, and the learning rate of critic networks to 2×10^{-6}.

VI. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed scheme and algorithm. We consider $M = 10$ IIoT devices and divide them into $K = 2$ types [17], [30].

For the settings of experimental parameters, ψ_1 and ψ_2 are randomly sampled within $[50, 100]$ and $[200, 250]$, respectively. q_1 and q_2 are randomly generated following the Dirichlet distribution [5]. Considering the dynamic environment of ICPSs, the unit cost c is randomly sampled within $[25,35]$, and the unit revenue ϑ is randomly sampled within $[10,15]$. In addition, the pre-defined weight parameter ρ is set to 0.6, the additional cost c_0 is set to 0.01, and the pre-defined constant β is set to 0.5. Note that our experiments are conducted using PyTorch with CUDA 12.0 on NVIDIA GeForce RTX 3080 Laptop GPU.

Fig. 3 presents the test reward comparison of the proposed scheme under different schemes and different scenarios. Specifically, we compare the performance of the proposed scheme under asymmetric information and complete information. The contract theory under complete information does not consider IC constraints [12]. Although the performance of this scheme is better than that of the proposed scheme, it is not practical since the environment of complete information is not feasible. Then, we compare the performance of the proposed scheme with the random scheme in which the DT server randomly designs contracts. We can observe that the test reward of the proposed scheme is higher than that of the random scheme. The reason is that the proposed scheme can motivate IIoT devices to select suitable contract items according to their types so that the agent (i.e., the DT server) can obtain more rewards $R(\boldsymbol{e}_z, \Omega_z)$, which indicates that the proposed scheme can effectively mitigate the effect of information asymmetry by utilizing contract theory. Overall, the proposed scheme is effective and reliable.

Fig. 4 shows the performance of the proposed algorithm and other DRL algorithms in optimal contract design. We can

observe that although the proposed algorithm does not converge quickly due to the influence of denoising, it can stabilize the highest final reward compared with other DRL algorithms. The reason is that our algorithm optimizes a stochastic policy in an entropy-augmented reward framework, encouraging exploration and robustness [27]. Moreover, by leveraging diffusion models, our algorithm can generate samples with higher quality by multiple fine-tuning [9], enhancing the sampling accuracy and reducing the effect of uncertainty and noise from the environment. Besides, by pruning unimportant neurons and connected weights of actor networks, the performance of the diffusion-based SAC algorithm can be improved. The reason is that the pruning technique can reduce the complexity of GDM networks and improve model generalization to unseen states.

Fig. 5 illustrates the utility of the DT server and the average utility of IIoT devices under our algorithm and other DRL algorithms, where the average utility of IIoT devices is calculated based on a weighted average of device types. Due to the dynamic environment of ICPSs, we consider that the unit revenue of the DT server satisfaction is dynamically changing, and it is reasonable that the proposed algorithm achieves a higher utility of the DT server rather than the highest. However, the proposed algorithm can achieve the highest average utility of IIoT devices, indicating that the proposed algorithm can generate more reasonable contract items that better incentivize IIoT devices to contribute data for DT construction. In summary, based on the analyses of Figs. 4 and 5, the proposed algorithm can design feasible contracts to achieve the highest utility of IIoT devices under asymmetric information.

Fig. 6 evaluates the environmental impact of the proposed algorithm. We use a Python package named CodeCarbon[1] to

---

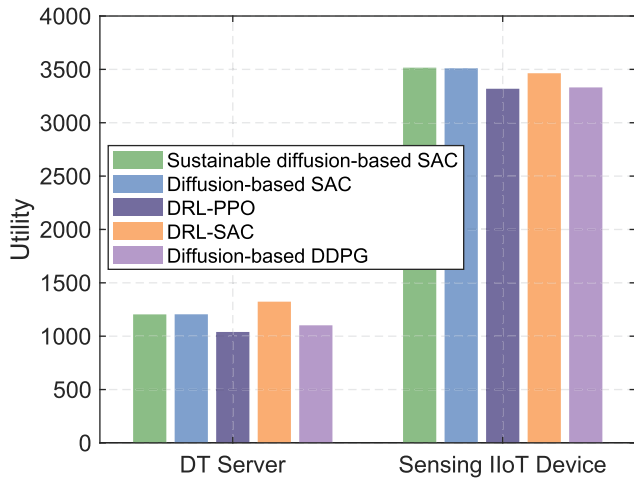[1][Online]. Available: https://github.com/mlco2/codecarbon

Fig. 5. The utility of the DT server and the average utility of IIoT devices under different algorithms.

(a) Diffusion schedules.

(b) Pruning rates.

Fig. 7. Performance analyses of the proposed algorithm, where we evaluate the impacts of different noise schedule strategies and pruning rates on the performance of the algorithm.
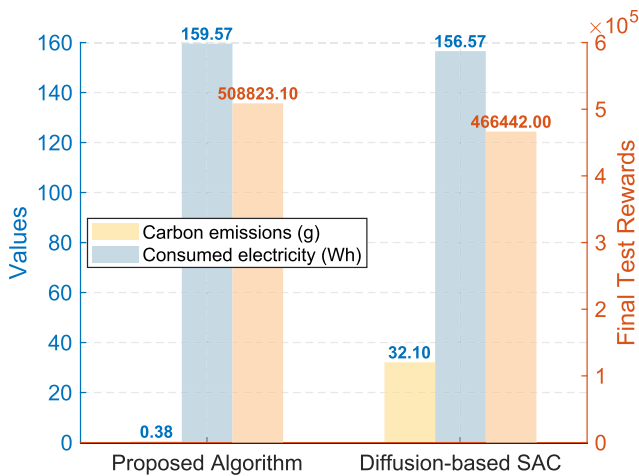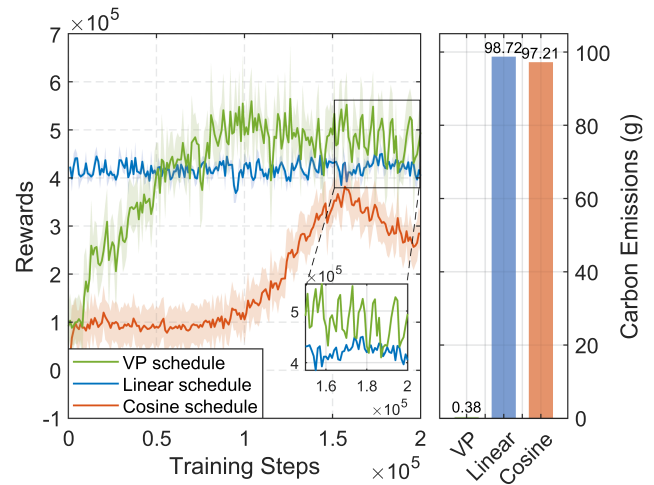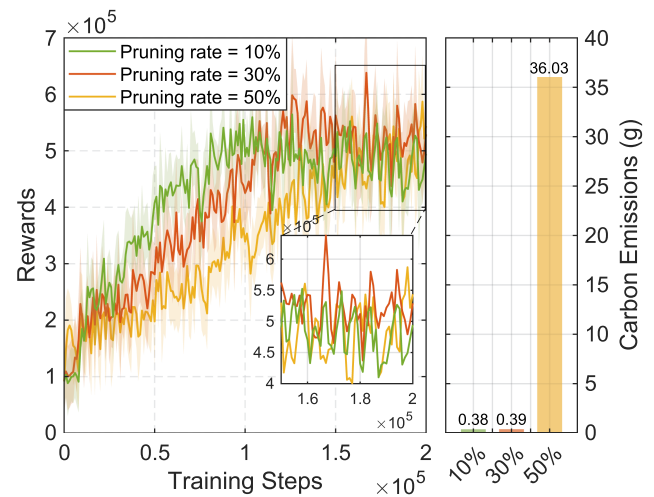

Fig. 6. Comparison of the environmental impacts of the proposed algorithm with the diffusion-based SAC algorithm.

estimate the carbon emissions and electricity consumption of these algorithms in achieving optimal contract design. We set the pruning rate to 10%, which indicates that 10% of channels are pruned in the whole network. Compared with the diffusion-based SAC algorithm, which also performs well in optimal contract design, we can observe that the proposed algorithm not only has better performance, i.e., a higher final test reward, but also produces lower carbon emissions of about 0.38 g during model training for optimal contract design. The reason is that pruning techniques can remove excess neurons that are useless to the performance of actor networks [16], which is beneficial for decreasing model sizes and the need for multiple iterations, thus reducing carbon emissions.

Fig. 7 presents the performance analyses of the proposed sustainable diffusion-based SAC algorithm. First, we evaluate the impact of different noise schedule strategies on the performance of the proposed algorithm, including Variance Proportional (VP), linear, and cosine noise schedule strategies, where

the pruning rate is set to 10%. In diffusion models, the noise schedule strategy determines the amount of noise added to input data at each timestep during the forward diffusion process [31], which affects the quality of generated samples in the reverse diffusion process. As illustrated in Fig. 7(a), we can observe that among the three noise schedule strategies, the proposed algorithm under the VP noise schedule strategy can obtain the highest test reward and produce the lowest carbon emissions, which highlights the superior performance of the VP noise schedule strategy in applying the sustainable diffusion-based SAC algorithm for optimal contract design. Then, we evaluate the impact of different pruning rates on the performance of the proposed algorithm under the VP noise schedule strategy, i.e., 10%, 30%, and 50%. As shown in Fig. 7(b), we can observe that higher pruning rates result in slower convergence of the proposed algorithm but higher final test rewards. The reason is that higher pruning rates may eliminate critical connections that

contribute to the learning ability of actor networks, resulting in slow convergence. Despite the slower convergence, the pruned actor networks often exhibit better generalization by learning more essential features instead of relying on redundant parameters. In addition, when the pruning rate is set to 50%, the carbon emissions generated by model training are the highest, reaching 36.03 g, which indicates that there is an optimal pruning rate to balance the performance and the sustainability of the proposed algorithm. In summary, the above numerical results demonstrate that the proposed algorithm is sustainable and effective.

VII. CONCLUSION

In this paper, we have studied how GenAI empowers DTs in ICPSs. Specifically, we have designed a GenAI-driven DT architecture in ICPSs, which systematically studies how GenAI drives the DT construction pipeline, including real-time physical data collection, communications for DTs, DT modeling and maintenance, and DT decision-making. To motivate IIoT devices to contribute sensing data for GenAI-empowered DT construction, we have proposed a contract theory model under information asymmetry. Furthermore, we have developed a sustainable diffusion-based SAC algorithm to generate the optimal feasible contract, which utilizes dynamic structured pruning techniques to sparsify actor networks of GDMs, allowing efficient implementation of the proposed algorithm in ICPSs. Finally, the numerical results demonstrate that the proposed algorithm outperforms DRL algorithms that are commonly used for optimal contract design. In particular, when the pruning rate is set to 10%, compared with the diffusion-based SAC algorithm, the proposed algorithm can reduce carbon emissions by 99% during model training while ensuring high test rewards. For future work, we will propose a more general and comprehensive architecture of GenAI-empowered DT in ICPSs. In addition, we will focus on the incentive mechanism design among multiple IIoT devices and DT servers and explore combining state-of-the-art techniques with DRL algorithms for optimal contract design.

REFERENCES

[1] K. Zhang, Y. Shi, S. Karnouskos, T. Sauter, H. Fang, and A. W. Colombo, "Advancements in industrial cyber-physical systems: An overview and perspectives," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 716–729, Jan. 2023.

[2] Y. Jiang, S. Wu, R. Ma, M. Liu, H. Luo, and O. Kaynak, "Monitoring and defense of industrial cyber-physical systems under typical attacks: From a systems and control perspective," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 1, pp. 192–207, 2023.

[3] J. Zhang, X. Li, J. Tian, Y. Jiang, H. Luo, and S. Yin, "A variational local weighted deep sub-domain adaptation network for remaining useful life prediction facing cross-domain condition," *Rel. Eng. Syst. Saf.*, vol. 231, 2023, Art. no. 108986. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832022006019

[4] J. Zhang, J. Tian, P. Yan, S. Wu, H. Luo, and S. Yin, "Multi-hop graph pooling adversarial network for cross-domain remaining useful life prediction: A distributed federated learning perspective," *Rel. Eng. Syst. Saf.*, vol. 244, 2024, Art. no. 109950. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832024000255

[5] J. Wen et al., "From generative AI to generative Internet of Things: Fundamentals, framework, and outlooks," *IEEE Internet Things Mag.*, vol. 7, no. 3, pp. 30–37, May 2024.

[6] J. Chen, Y. Shi, C. Yi, H. Du, J. Kang, and D. Niyato, "Generative AI-driven human digital twin in IoT-healthcare: A comprehensive survey," 2024, *arXiv:2401.13699*.

[7] X. Huang, H. Yang, C. Zhou, X. Shen, and W. Zhuang, "When digital twin meets generative AI: Intelligent closed-loop network management," *IEEE Network*, early access, 2024, doi: 10.1109/MNET.2024.3524474.

[8] I. Lotfi, D. Niyato, S. Sun, D. I. Kim, and X. Shen, "Semantic information marketing in the metaverse: A learning-based contract theory framework," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 3, pp. 710–723, Mar. 2024.

[9] H. Du et al., "Diffusion-based reinforcement learning for edge-enabled AI-generated content services," *IEEE Trans. Mobile Comput.*, vol. 23, no. 9, pp. 8902–8918, Sep. 2024.

[10] G. Fang, X. Ma, and X. Wang, "Structural pruning for diffusion models," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2024, vol. 36, pp. 16716–16728.

[11] L. Jiang, H. Zheng, H. Tian, S. Xie, and Y. Zhang, "Cooperative federated learning and model update verification in blockchain-empowered digital twin edge networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11154–11167, Jul. 2022.

[12] J. Kang et al., "Blockchain-empowered federated learning for healthcare metaverses: User-centric incentive mechanism with optimal data freshness," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 348–362, Feb. 2024.

[13] M. Bordukova, N. Makarov, R. Rodriguez-Esteban, F. Schmich, and M. P. Menden, "Generative artificial intelligence empowers digital twins in drug discovery and clinical trials," *Expert Opin. Drug Discov.*, vol. 19, no. 1, pp. 33–42, 2024.

[14] Z. Tao, W. Xu, Y. Huang, X. Wang, and X. You, "Wireless network digital twin for 6G: Generative AI as a key enabler," *IEEE Wireless Commun.*, vol. 31, no. 4, pp. 24–31, Aug. 2024.

[15] J. Kang et al., "Tiny multiagent DRL for twins migration in UAV metaverses: A multileader multifollower Stackelberg game approach," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21021–21036, Jun. 2024.

[16] W. Su, Z. Li, M. Xu, J. Kang, D. Niyato, and S. Xie, "Compressing deep reinforcement learning networks with a dynamic structured pruning method for autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 18017–18030, Dec. 2024.

[17] J. Wen et al., "Diffusion model-based incentive mechanism with prospect theory for edge AIGC services in 6G IoT," *IEEE Internet Things J.*, vol. 11, no. 21, pp. 34187–34201, Nov. 2024.

[18] J. Yu, A. Alhilal, P. Hui, and D. H. Tsang, "Bi-directional digital twin and edge computing in the metaverse," *IEEE Internet Things Mag.*, vol. 7, no. 3, pp. 106–112, May 2024.

[19] G. Liang et al., "Data augmentation for predictive digital twin channel: Learning multi-domain correlations by convolutional timeGAN," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 1, pp. 18–33, Jan. 2024.

[20] W. Hu, T. Wang, and F. Chu, "A Wasserstein generative digital twin model in health monitoring of rotating machines," *Comput. Ind.*, vol. 145, 2023, Art. no. 103807. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361522002032

[21] K. Feng, J. Ji, Y. Zhang, Q. Ni, Z. Liu, and M. Beer, "Digital twin-driven intelligent assessment of gear surface degradation," *Mech. Syst. Signal Process.*, vol. 186, 2023, Art. no. 109896. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327022009645

[22] K. Feng et al., "Digital twin enabled domain adversarial graph networks for bearing fault diagnosis," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 1, pp. 113–122, 2023.

[23] K. Feng, W. A. Smith, R. B. Randall, H. Wu, and Z. Peng, "Vibration-based monitoring and prediction of surface profile change and pitting density in a spur gear wear process," *Mech. Syst. Signal Process.*, vol. 165, 2022, Art. no. 108319. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327021006798

[24] D. Wen, Y. Zhou, X. Li, Y. Shi, K. Huang, and K. B. Letaief, "A survey on integrated sensing, communication, and computation," *IEEE Commun. Surveys Tut.*, early access, Dec. 23, 2024, doi: 10.1109/COMST.2024.3521498.

[25] J. Wen et al., "Task freshness-aware incentive mechanism for vehicle twin migration in vehicular metaverses," in *Proc. IEEE Int. Conf. Metaverse Comput., Netw. Appl.*, 2023, pp. 481–487.

[26] Z. Wang, J. J. Hunt, and M. Zhou, "Diffusion policies as an expressive policy class for offline reinforcement learning," 2022, *arXiv:2208.06193*.

[27] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 1861–1870.

[28] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.

[29] R. Zhang, K. Xiong, Y. Lu, P. Fan, D. W. K. Ng, and K. B. Letaief, "Energy efficiency maximization in RIS-assisted SWIPT networks with RSMA: A PPO-based approach," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 5, pp. 1413–1430, May 2023.

[30] D. Ye et al., "Optimizing AIGC services by prompt engineering and edge computing: A generative diffusion model-based contract theory approach," *IEEE Trans. Veh. Technol.*, early access, Oct. 08, 2024, doi: 10.1109/TVT.2024.3463420.

[31] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *Proc. 38th Int. Conf. Mach. Learn.*, Jul. 2021, vol. 139, pp. 8162–8171. [Online]. Available: https://proceedings.mlr.press/v139/nichol21a.html

Yang Zhang (Member, IEEE) received the B.Eng. and M.Eng. degrees from Beihang University, Beijing, China, in 2008 and 2011, respectively, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore, in 2015. He is currently an Associate Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research focuses on edge computing and multiagent unmanned systems. He is an Editor of IEEE TRANSACTIONS ON MACHINE LEARNING IN COMMUNICATIONS AND NETWORKING.

Jinbo Wen (Graduate Student Member, IEEE) received the B.Eng. degree from the Guangdong University of Technology, Guangzhou, China, in 2023. He is currently working toward the M.S. degree with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests include generative AI, edge intelligence, and metaverse.

Jiawen Kang (Senior Member, IEEE) received the Ph.D. degree from the Guangdong University of Technology, Guangzhou, China, in 2018. He was a Postdoc with Nanyang Technological University, Singapore, from 2018 to 2021. He is currently a Full Professor with the Guangdong University of Technology. His research mainly focuses on blockchain, security, and privacy protection in wireless communications and networking.

Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkuts Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 1999, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2008. He is currently a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests include the areas of sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.

Shiwen Mao (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from Polytechnic University, Brooklyn, NY, USA. He is currently a Professor, Earle C. Williams Eminent Scholar, and Director of the Wireless Engineering Research and Education Center, Auburn University, Auburn, AL, USA. His research interests include wireless networks and multimedia communications. He is a Distinguished Lecturer of the IEEE Communications Society and IEEE Council of RFID. He is the Editor-in-Chief of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, Member-at-Large of IEEE Communications Society Board of Governors, and Vice President of Technical Activities of IEEE Council on Radio Frequency Identification (CRFID). He was the General Chair of IEEE INFOCOM 2022, TPC Chair of IEEE INFOCOM 2018, and TPC Vice-Chair of IEEE GLOBECOM 2022. He was the recipient of the IEEE ComSoc MMTC Outstanding Researcher Award in 2023, SEC 2023 Faculty Achievement Award for Auburn, IEEE ComSoc TC-CSR Distinguished Technical Achievement Award in 2019, Auburn University Creative Research & Scholarship Award in 2018, NSF CAREER Award in 2010, and several IEEE service awards. He was a co-recipient of the 2022 Best Journal Paper Award of IEEE ComSoc eHealth Technical Committee, 2021 Best Paper Award of Elsevier/KeAi Digital Communications and Networks Journal, 2021 IEEE Internet of Things Journal Best Paper Award, 2021 IEEE Communications Society Outstanding Paper Award, IEEE Vehicular Technology Society 2020 Jack Neubauer Memorial Award, 2018 Best Journal Paper Award and 2017 Best Conference Paper Award from IEEE ComSoc MMTC, and 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.