Multi-Scale Simulation of Complex Systems: A Perspective of Integrating Knowledge and Data

HUANDONG WANG, HUAN YAN, CAN RONG, YUAN YUAN, FENYU JIANG, ZHENYU HAN, HONGJIE SUI, DEPENG JIN, and YONG LI, Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, China

Complex system simulation has been playing an irreplaceable role in understanding, predicting, and controlling diverse complex systems. In the past few decades, the multi-scale simulation technique has drawn increasing attention for its remarkable ability to overcome the challenges of complex system simulation with unknown mechanisms and expensive computational costs. In this survey, we will systematically review the literature on multi-scale simulation of complex systems from the perspective of knowledge and data. Firstly, we will present background knowledge about simulating complex systems and the scales in complex systems. Then, we divide the main objectives of multi-scale modeling and simulation into five categories by considering scenarios with clear scale and scenarios with unclear scale, respectively. After summarizing the general methods for multi-scale simulation based on the clues of knowledge and data, we introduce the adopted methods to achieve different objectives. Finally, we introduce the applications of multi-scale simulation in typical matter systems and social systems.

CCS Concepts: • **Computing methodologies** → **Multiscale systems**; *Machine learning*; • **Applied computing**;

Additional Key Words and Phrases: Complex system, simulation, multiple scales, knowledge, data

1 INTRODUCTION

Complex systems refer to a category of systems with significant complexity, which is a concept with broad boundaries covering numerous fields of science [24, 72, 170, 193]. For example, the human brain composed of a large scale of neurons, the physical system composed of a large scale of particles, and the transport system composed of a large scale of roads and vehicles are all representative complex systems. Their complexity often arises from the existence of the vast number of elements and components within complex systems, and nonlinearity, uncertainty, feedback loops, and emergent properties caused by the complicated interaction and co-evolution among them [17, 96]. At the same time, in the modern age, numerous cutting-edge technology applications, including drug design, weather forecasting, chemical engineering, and socio-economic

Authors' address: Huandong Wang, wanghuanong@tsinghua.edu.cn; Huan Yan, yanhuanthu@gmail.com; Can Rong, rongcan@pku.edu.cn; Yuan Yuan, y-yuan20@mails.tsinghua.edu.cn; Fenyu Jiang, jiangfenyu@tsinghua.edu.cn; Zhenyu Han, hanzy19@mails.tsinghua.edu.cn; Hongjie Sui, huihj21@mails.tsinghua.edu.cn; Depeng Jin, jindp@tsinghua.edu.cn; Yong Li, liyong07@tsinghua.edu.cn, Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, China.

governance, are highly relevant to complex systems, and require effective understanding, predicting, and controlling complex systems. However, the complicated properties within complex systems often prevent us from effectively understanding, predicting, and controlling them. For example, the strong uncertainty of the current states of complex systems makes it difficult to predict their future states accurately. In addition, due to the existence of nonlinearities in the system, the aggregate effect of the interventions we impose on the system becomes no longer a simple summation of the effects of individual interventions, and often is accompanied by the effect of feedback loops and emergent properties, making the control of complex systems extremely difficult.

Fortunately, the rapid development of computing technologies has enabled us to obtain unprecedented computing power. The increasing availability of high-performance computing machines with parallel processors has made it possible to numerically compute the evolution of the massive variables in complex systems and thus effectively simulate them in this way [82]. On the other hand, the rapid development of artificial intelligence (AI) technologies has made data-driven AI-powered simulation a great success, among which the most representative one is the DeepMind's achievements from protein simulation [93] to fluid simulation [163]. Overall, computer simulation has played an irreplaceable role in understanding, predicting, and controlling complex systems by providing excellent insight and foresight, which includes but is not limited to the following:

- **Establish key connections with first-principle simulation:** In many application scenarios of complex systems, we want to establish the connections between variables, structures, and properties in the systems, which is vague in most cases. Under the circumstances, we can simulate the structure and properties from the variables using first-principles simulation, i.e., ab initio simulation, to establish the connection between them. For example, given the amino acid sequence of a protein, we want to predict its function. In this case, first-principles simulation of its structure of protein folding can help us predict its functional properties to a large extent [22]. Or when we want to understand the phase transition process of interdependent complex networks under random attacks, we can also simulate the network connectivity in the process of cascading failure, based on which we can build connections between whether a cascading collapse occurs and the probability of attack [16, 27].
- **Derive future system states with predictive simulation:** Simulation technology, on the other hand, is mostly used to help us derive the future state of the system from the current state to provide foresight of the system, whether through kinetic equations [15, 53] or other artificial intelligence models such as RNN [57, 67, 71] or GCN [117, 142]. For example, by numerical simulation of weather, we can avoid the damage of extreme weather to agriculture. Another example is that by utilizing direct numerical simulation (DNS) to solve Navier–Stokes equations, we can predict fluid dynamics [101], which is helpful in the design of airplanes, vehicles, and engines [127].
- **Eliminate uncertainty with sampling simulation:** When dealing with a complex system with strong uncertainty, the high dimensionality and strong mutual dependency of random variables within the complex system make it hard to analytically understand the system. Take a system of atomic spins described by the three-dimensional Ising model [20] as an example. For the large number of random variables describing the spin directions of atoms in the system, even the simple task of obtaining the probability distribution of the number of atoms with different spin orientations is difficult to be analytically solved. Under this circumstance, sampling simulation, e.g., Monte Carlo Markov simulation [195], can be utilized to generate samples of random variables without requiring analytical solutions to their probability distributions. In this way, we can characterize the system in terms of statistics of these samples, e.g., their mean and variance, thereby eliminating the effects of the strong randomness.

Despite the excellent insight and foresight provided by these methods, how to effectively simulate complex systems is still an open problem with significant challenges, which can be summarized as follows:

- **Unknown mechanisms:** In many cases, the mechanisms of the complex systems that we want to simulate are unclear, incomplete, or difficult to obtain. A direct way to solve this problem is to utilize data-driven methods to fit the unknown mechanisms based on observable data. However, this solution has two limitations. First, the data-driven method struggles with generalization. That is, the performance of the fitted dynamics cannot be guaranteed in the scenario not covered by the observable data [101]. On the other hand, the observable data are also often incomplete and noisy, which makes fitting the mechanisms very difficult.

- **Expensive computational cost:** Even though the dynamics of complex systems are clear and complete, calculating the evolution of extremely high-dimensional state variables in complex systems is usually costly. For example, an accurate direct-numerical simulation of fluid dynamics based on Navier–Stokes equations is impossible at the scale required for solving weather prediction or engine design [101]. How to simulate complex systems with acceptable computational costs is another challenge we have to face.

The technique of multi-scale simulation, which has been investigated for several decades, can help us solve these two challenges. Specifically, the core idea of multi-scale simulation is to utilize the mechanisms corresponding to multiple scales simultaneously to effectively simulate the dynamics of complex systems. The different scales usually have different spatial resolutions, temporal resolutions, or quantities, and thus sometimes have completely different mechanisms [24]. As for the problem of unknown mechanisms, it is trivial that we can complement or derive the macroscale mechanisms based on the simulated microscale dynamics [185]. Moreover, the microscale dynamics can be better modeled with the boundary conditions or background fields [189] obtained based on the simulated macroscale dynamics. Sometimes, the macroscale and microscale mechanisms are both incomplete and tightly coupled to each other, and thus we have to simulate their dynamics together in a concurrent manner [122]. Overall, through the method of multi-scale simulation, we can complete the incomplete mechanisms of a specific scale, and learn or calibrate the parameters of the target scale mechanism, thus overcoming the challenges of unknown mechanisms. As for the expensive computational cost, the dynamics simulated based on macroscale mechanism are more coarse-grained and less accurate, but it is significantly less computationally expensive, since the number of variables that should be handled is greatly reduced. A direct way is to use the macroscale dynamics to approximate the microscale dynamics [176]. Furthermore, the unimportant areas can be simulated and approximated based on macroscale mechanisms, while microscale mechanisms are still used for simulating important areas [109]. Through appropriate message-passing mechanisms between dynamics of different scales, we can simulate the complex system with both computational efficiency and simulation accuracy, thus overcoming the challenges of expensive computational costs.

In this survey, we will systematically review the recent literature on multi-scale simulation of complex systems, which include both matter systems and social systems. Our goal is to conduct an interdisciplinary literature survey to summarize the common paradigms and methodologies of existing work on multi-scale simulation from the perspective of knowledge and data, thereby helping the computer simulation community, machine learning community, transportation society, and others. We hope this survey will help multidisciplinary communities to understand the progress made in existing work and draw useful insights and perspectives from work in other disciplines, as well as potential gaps and opportunities. To provide a comprehensive understanding of interdisciplinary work on multi-scale simulation of complex systems, we provide a novel multi-perspective taxonomy

of existing work in terms of their objectives, general methods, objective-oriented methods, and applications. Furthermore, the taxonomy of methods in our paper is closely organized around the topic of how existing work leverages knowledge and data. To the best of our knowledge, no existing survey has covered existing studies in terms of both matter complex systems and social complex systems. We hope that this paper will help advance progress in the direction of multi-scale simulation of complex systems.

The structure of this paper is as follows. We begin by introducing related reviews of our paper by comparing their main difference. Then, we introduce background about the simulation of complex systems and scales in complex systems. Next, we summarize the main objectives of multi-scale modeling and simulation techniques. Following the introduction of general methods of multi-scale simulation techniques based on the clues of knowledge and data introduced in Section 5, we introduce the mainly adopted methods with different objectives in Section 6. Finally, we introduce the application of multi-scale simulation techniques in different systems, including major matter systems and social systems.

2   RELATED REVIEWS

Paper	Description
[59]	This paper presented the related techniques, tools, and challenges for co-simulation, where a coupled system can be simulated by integrating with multiple simulators.
[147]	This paper summarized the recent developments of finding the collective variables of a system, including the related principles, and the state-of-the-art algorithms.
[99]	This paper studied different data-driven methods to project dynamics of the high-dimensional system into the lower-dimensional space, and further explored the similarity and differences between these methods.
[153]	This paper summarized the recent multi-scale methods that integrate with machine learning, and discussed the mutual benefits with each other.
[5]	This paper surveyed the multi-scale methods in the cases of vehicles and crowds, which mainly focused on the bottom-up derivation of models.

Table 1. Related reviews of complex system modeling and simulation.

There are some comprehensive articles that summarize different techniques, tools, and algorithms in the area of simulation [5, 59, 99, 147, 153]. Gong et al. [59] gave a comprehensive review about co-simulation, where a coupled system can be simulated by the composition of different simulators. The corresponding theory and techniques were surveyed, which are grouped into two types of paradigms: discrete event co-simulation and continuous time co-simulation. Further, this work studied the hybrid co-simulation that establishes the relationship between these two paradigms. Frank et al. [147] focused on the role of collective variables in studying high-dimension dynamical systems. The collective variables can discover meaningful collective phenomena over a long period of time, as well as identify different structures at macroscopic scales. In the system simulation, it is important to find the collective variables of the system to control the simulation. The authors reviewed the relevant theory and algorithms that search for the collective variables from simulation data. Similarly, Klus et al. [99] studied the recent data-driven methods in discovering the collective variables of a system, including data-driven dimension reduction methods and transfer operator approximation methods. Further, they summarized the similarities and differences between these methods, and discussed how extensions designed for a specified method can be applied to other
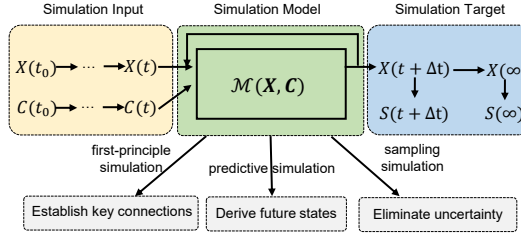
Fig. 1. An illustration of simulation of complex systems, where $X(t)$ is the observable system state variables at time t, $C(t)$ is the controllable condition variable of the system. X and C are the set of system state variables and controllable condition variables from t_0 to t, respectively. $S(t)$ is the target variable that we want to obtain through simulation.

related ones. Peng et al. [153] presented the multi-scale modeling methods based on physics-based knowledge and data-driven machine learning. They discussed how multi-scaling modeling integrates with machine learning, and explained the mutual benefits with each other. Albi et al. [5] surveyed the multi-scale methods in the cases of vehicles and crowds, where the models at the macroscopic scale are derived from individual-based modeling at the microscopic scale. Different from them, our paper gives a comprehensive analysis of multi-scale simulation from different aspects, but the above articles point out a small part of them. To be specific, we identify three interaction ways in the scenarios of clear-scale modeling based on the flow of information between scales, including information flow from microscale to macroscale, from macroscale to microscale, and the mutual interaction between microscale and macroscale. However, [5] only discussed the interaction from microscale to macroscale, and [147] and [99] focused on the interaction from macroscale to microscale. Furthermore, our paper considers the scenarios of multi-scale simulation with unclear scales in complex systems, and two additional objectives of multi-scale simulation, including *discovering unknown scales* and *learning dynamics at the discovered scale*. Therefore, this paper can provide more insightful and valuable information for researchers to understand the multi-scale simulation methods in complex systems.

3 BACKGROUND

To pave the way for reviewing the literature on multi-scale simulation of complex systems, we first introduce background about the concepts of simulation and scales of complex systems, respectively.

3.1 Simulation of complex systems

The goal of the simulation of complex systems is to imitate the operation or behavior of the systems. As shown in Figure 1, by denoting the observable system state variables at time t as $X(t)$ and the controllable condition variables as $C(t)$, a simulation model $\mathcal{M}(X, C)$ is built to mimics the target complex system. Specifically, we define $X = \{X(\tau)\}_{\tau=t_0}^{t}$ and $C = \{C(\tau)\}_{\tau=t_0}^{t}$, which are the set of all historical observable system state variables and controllable condition variables, respectively.

In some systems, the simulation models only take $X(t)$ and $C(t)$ as inputs. It depends on whether the system is Markovian, i.e., whether the future state depends only on the currently observed system state variables. This is true for a large number of physical systems. For example, in a fluid system of gases, given the current velocity field and external forces, the future state, i.e., the future velocity field, can be accurately simulated. However, this is not true for most social systems, such as transportation systems. The essential reason is that the observable state variables are not enough to completely determine the state of the system, and there exist unobservable hidden state variables

in the system. For example, in a transportation system, we can observe the position and speed of each vehicle, but this is not enough to fully determine the future positions and speeds of all vehicles. In addition to the mental states of the vehicle drivers, each vehicle also has a different destination, which is hidden in the system and largely affects the future vehicle behavior. Under these circumstances, the historical sequence of observable state variables should be introduced in the simulation model to estimate the hidden state of the systems.

The simulation target is diverse, and depends on which aspects of the simulation model are required to imitate the operation or behavior of the real complex system. For example, in different simulations, different target variables $S(t)$ are required to be simulated. In addition, sometimes the whole simulated process of system state evolution, i.e., $\{S(t + \Delta t)|\Delta t \in (0, \infty)\}$ is required to imitate the actual complex system, while other times we only focus on the final state of the system $S(\infty)$. At the same time, sometimes we need to simulate the exact value of the system state, i.e., $S = h(X)$, and other times we need to estimate its probability distribution $S = Pr(h(X))$.

3.2   Scales of complex systems

There exist various categories of scales in complex systems. As shown in Figure 2(a), typical scales can be summarized as the following three categories. The first category is the scale of space, of which the definition is diverse in different scientific fields. For example, in physics, from electrons to atoms, further to molecules, and continuum, the corresponding spatial scales range from a few nanometers to a few meters. In biology, it starts from amino acids to polypeptide chains, and finally forms protein molecules. The second category is the scale of time. A representative example is that there exist violent reactions and slow reactions in chemical reaction complex systems, which lead to the formation of fast and slow variables in the system. The last category is the scale of quantity. For example, in society, there exist scales from the individual to the community, and finally to the society. Note that scales can generally be divided into two types. In the first type of scales, the corresponding state is defined over a continuous area, such as the velocity field of a fluid. In the second type of scales, the corresponding state is defined based on a limited number of variables, such as vehicles in a transportation system, and nodes in a complex network. We refer to the former type of scales as the field-based scale and the latter as the particle-based scale.

Despite the typical scales with clear definitions and physical meanings in complex systems, there also exist numerous unclear scales in complex systems. As shown in Figure 2(b), the massive microscale variables in the system form mesoscale variables through complicated combinations and transformations, and further form macroscale variables in the system. These scales often do not correspond to a uniform space size, temporal length, or quantity, but correspond to the emergent structures [81, 112] or laws [186] in the complex system, and play an irreplaceable role in the accurate simulation of the future evolution of the system. In most cases, these scales cannot be effectively extracted based on simple space size, time length, or quantity scale. Under the circumstances, techniques such as manifold learning, and representative learning are required to be utilized to find the hidden scales in complex systems.

4   OBJECTIVES OF MULTI-SCALE MODELING AND SIMULATION

We first summarize the main objectives of existing techniques of multi-scale simulation. As discussed in Section 3.1, there exist clear scales corresponding to different space sizes, temporal lengths, and quantities in complex systems, and numerous unclear scales as well. The clarity of the system's useable scales also results in differences in the objectives of the corresponding multi-scale simulation approaches. Thus, in the following, we will introduce the objectives of multi-scale simulation in the scenario with clear scales and unclear scales, respectively. Without loss of generality, we denote
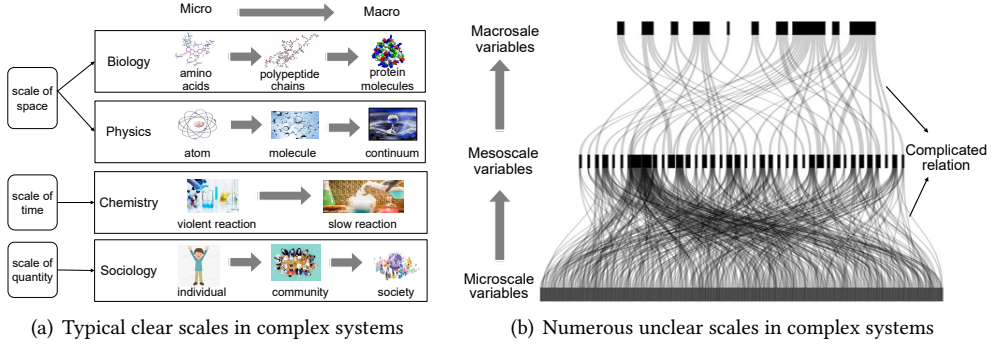
(a) Typical clear scales in complex systems

(b) Numerous unclear scales in complex systems

Fig. 2. An illustration of scales in complex systems.

the microscale system state variables as $X(t)$, and denote the macroscale system state variables as $\xi(t)$ in this section.

4.1 Scenario with clear scale

Although the clear-scale modeling and simulation of complex systems are commonly adopted in many fields, such as physics [47, 51, 95, 117–119, 192], material science [48, 105, 115, 123, 181], chemistry [24, 72], and bioprotein [19, 70, 90]. The scale separation can vary in different fields. In the field of fluid/solid state physics, it is often necessary to focus on atomic and above scales modeled using classical mechanics and microscopic electronic scales modeled using quantum mechanics [24]. In materials science, researchers usually focus on two scales, i.e., continuum strains and atomic movements, due to the challenges posed by the heterogeneity of materials [48]. In the study of biological proteins, different scale separations are done in different research scenarios, such as atomic level in the discovery of new protein structures, and coarse-grained lipids in the study of membrane remodeling [19, 90]. All the above complex system simulations are explicitly divided into multiple scales, and each scale holds a specific spatio-temporal resolution. Consistent with the existing works, in the clear scale scenario of the existing literature, the scales are ordered according to the coarseness of the spatio-temporal granularity.

Specifically, we divide the main objectives of multi-scale simulation into three categories, which correspond to different ways of interaction between different scales. It is worth noting that the methods do not need to be restricted to sequential modeling or concurrent modeling [185]. What matters is whether the flow of information between scales is from macroscale to microscale or vice versa. The introduction of three objectives is as follows.

- **Resolving local fine-grained dynamics (*Objective A*):** In this objective, the local fine-grained microscale dynamics are what the study is interested in. That is, the simulation target depends only on a local subset of the system state variables, i.e., $S(t) = h(X_l(t))$ where $X_l(t) \subset X(t)$. The microscale models are leveraged to resolve the local dynamics $X_l(t)$ in detail, while the macroscale models are utilized to simulate the coarse-grained patterns anywhere else, i.e., $X(t)/X_l(t)$.

- **Completing macroscale mechanisms (*Objective B*):** *Objective B* is the exactly opposite-type of *Objective A*. The macroscale dynamics of systems are what we care about. That is, the simulation target $S(t) = h(\xi(t))$. However, the macroscale mechanism is incomplete. For example, some constitutive information is missing, which is then obtained from microscale models and serves as the key components of macroscale models.

- **Co-simulating microscale and macroscale dynamics (*Objective C*)**. To achieve this objective, there is a mutual interaction between the microscale models and the macroscale models. The dynamics of a specific scale or its derived statistics will be updated through the models of the corresponding scale and then passed to the other. The simulation of the complex system will be a process of iterative computation of the macroscale models and the microscale models and the information interaction between them.

4.2   Scenario with unclear scale

Although classic fields, such as physics, chemistry, and biology, have proposed effective clear-scale models, the useable scales in complex systems are not always clear in complex systems, especially for high-dimensional systems. Here, a fundamental problem is which scales are more appropriate for conducting multi-scale simulations for a particular target system. For example, in a large range of chemical systems, where different reactions compete with each other, it is difficult to study why reactions occur and how to predict and control their behaviors at a clear scale [41, 160]. In ecosystems, species reside and interact with each other, forming complex scales in communities. Learning the dynamics of communities to reveal how ecosystem functions are essential for ecology [83]. Complex systems cannot always be explicitly divided into clear and appropriate scales for different simulation tasks. To deal with these systems, it is necessary to automatically discover appropriate unknown scales in complex systems by reducing the original high-dimensional microscale system state to a low-dimensional macroscale system state, and learn the dynamics at the discovered scale simultaneously. Thus, the objectives of the corresponding multi-scale simulation approaches can be further classified into two categories, of which an illustration is shown in Figure 3 and the corresponding explanations are as follows:

- **Discovering unknown scales (*Objective D*):** Discovering unknown but useful scales is the key to effectively modeling complex systems without clear scales. This objective aims to extract more tractable scales from complex systems. Further, we must ensure the tractability and usefulness of the extracted scales by adding constraints in terms of specific properties, e.g., linearity or preserved relevant information. More specifically, as shown in Figure 3, this objective requires obtaining an encoder $\mathcal{E}$, which maps the microscale state variable X to the discovered macroscale state variable ξ, and a corresponding decoder $\mathcal{D}$ to solve the inverse problem.
- **Learning dynamics at the discovered scale (*Objective E*):** This objective is accompanied by *Objective D*, which aims to learn the effective dynamics $\mathcal{M}'(\xi, C)$ at the discovered target scale. More specifically, it is a variant of *Objective B* in the scenario with unclear scale. Unlike *Objective B*, in this objective, the macroscale mechanism may be completely unknown rather than incomplete. At the same time, constraints can be added to the associated problem of *objective E* to ensure that the macroscale mechanism to be learned has specific properties, e.g., linearity.

By combining *Objective D* and *Objective E*, the high-dimensional state variables of complex systems are able to be reduced to a low-dimensional macroscale state variable system with learned dynamics. Thus, the system is able to be simulated at the discovered low-dimensional scale.

5   GENERAL METHODS OF MULTI-SCALE MODELING AND SIMULATION

Before introducing the multi-scale simulation methods used in achieving each objective in detail, we first summarize the general methods of multi-scale simulation techniques based on whether they are knowledge-driven, data-driven, or data and knowledge jointly driven.
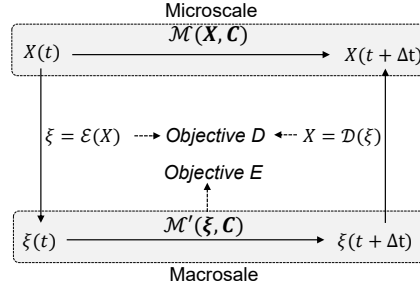
Fig. 3. An illustration of *objective D* and *objective E*.

5.1 Pure knowledge-driven method

Early approaches mainly adopt knowledge-driven methods to model and simulate complex systems. The commonly used methods mainly include kinetic equation models and statistical probability models.

Kinetic equation model. This model concentrates on modeling the evolution of the system variables in the continuous time domain. Specifically, it uses the differential equation of system state variables with respect to time to model and simulate the system, which is able to display the evolution of the complex system over time intuitively. By denoting the system state variable as X, which is usually high-dimensional, a kinetic equation model describing the evolution of the system with respect to time t can be represented by:

$$\frac{dX}{dt} = f(X). \tag{1}$$

For example, the dynamics of a complex network are modeled by [15, 53]:

$$\frac{dx_i}{dt} = F(x_i) + \sum_{(i,j) \in E} G(x_i . x_j), \tag{2}$$

where x_i is the state variable for the node i in the complex network, E is the set of edges of the complex network, and F and G are two functions describing the evolution mechanism of the complex network.

Another example is the fluid dynamics, which can be modeled by the Navier–Stokes equations:

$$\frac{d\boldsymbol{x}}{dt} = - \nabla \cdot (x \otimes x) + \frac{1}{Re} \nabla^2 x - \frac{1}{\rho} \nabla \boldsymbol{p} + f, \tag{3}$$

$$\nabla \cdot \boldsymbol{x} = 0, \tag{4}$$

where $\boldsymbol{x}$ is the flow velocity. Normally, $\boldsymbol{x}$ is a field defined on three-dimensional space. $\boldsymbol{p}$ is the field of pressure.

Kinetic equations can also be utilized to describe complex systems with stochastic dynamics. In this case, stochastic variables are introduced into the system, and derive stochastic differential equations (SDE) as follows:

$$dX = f(X)dt + \sigma(X)dW, \tag{5}$$

where W usually represents a Brownian motion, and thus its derivative represents a Gaussian noise. Specifically, SDE is utilized to model chemical reaction systems [166], biological cell systems [208], and financial systems [134], etc.

Probability statistical model. In some complex systems, the state of the system is not deterministic, but a probabilistic distribution defined on the state space. Under the circumstances, the system

states as well as the system dynamics are modeled based on probabilistic models, in which the most widely used model is the Markov chain (MC) model. Specifically, in quantum mechanics, the system state usually cannot be measured accurately, and thus it is often modeled by a probability density distribution. Furthermore, the SDE in the form of (5) also falls into the category of Markov models. More generally, the continuous-time Markov chain (CTMC) models the system in the continuous time domain, and it describes the system dynamics through the transition density:

$$\mathcal{T}(x, y; \tau) := p(X(t + \tau) = y | X(t) = x), \tag{6}$$

where we can observe that the transition density is independent with the time t, and $X(t + \tau)$ and $X(t)$ can be either defined on the continuous space or the discrete space.

On the other hand, the discrete-time Markov chain (DTMC) models the system in the discretized time domain. In this case, the future state X_t of the system at the time slot t is only dependent on the previous state X_{t-1}, and independent with the system state at earlier time slots, which can be represented by:

$$Pr(X_t | \{X_\tau\}_{\tau < t}) = Pr(X_t | X_{t-1}) =: T(X_{t-1}, X_t) \tag{7}$$

Representative DTMC models include the independent cascade model [162], GLobal Epidemic and Mobility (GLEAM) model [11], etc.

Similar to the kinetic equation model, the Markov chain model also focuses on modeling system evolution over time. However, in some specific applications and scenarios, we do not care about the evolution of complex systems over time. Instead, we focus on the final state of the system $X(\infty)$ or its statistics $S(\infty)$. For example, Hoel et al. [81] only consider one-step state transition of complex systems, where the considered states are referred to as the initial states and the resulting states, respectively. Further, the energy-based model (EBM), e.g., Ising model [144], can be utilized to describe the relation between the system state S and condition variables C. Specifically, by denoting the energy as $E(S, C)$, the probability of S can be represented by $Pr(S) = e^{E(S,C)} / \sum_{S' \in \mathcal{S}} e^{E(S',C)}$, where $\mathcal{S}$ is the set of all possible system state.

Categories	Method	Papers	Considering time	Deterministic/stochastic
Kinetic equation model	ODE	[15, 53]	Yes	Deterministic
	SDE	[166]	Yes	Stochastic
Statistical probability model	CTMM	[148]	Yes	Stochastic
	DTMM	[11, 81, 162]	Yes & No	Stochastic
	EBM	[136, 144]	No	Stochastic

Table 2. Knowledge-based multi-scale modeling and simulating methods

5.2 Pure data-driven method

5.2.1 Data-driven modeling and simulation approaches. Since the recently rapid development of machine learning technologies, and the availability of massive experimental datasets, data-driven methodologies appear as a surrogate to the conventional numerical methods governed by the physical equations in the field of simulation. Due to the computational efficiency and the acceptable simulation accuracy, the data-driven methods are leveraged in biology, chemistry, material science, and fluid dynamics, with the purpose of single-scale or multi-scale modeling. In this section, we prefer to provide an overview of the various machine learning techniques incorporated into the simulation procedure. These approaches mainly include methods such as the Recurrent Neural Network (RNN), the Convolutional Neural Network (CNN), the Graph Neural Network (GNN),

ODENet, and the Gaussian Process (GP), etc. Here we briefly introduce each method and discuss how they benefit the simulation work.

Categories	Scale	Method	Papers
Single-scale simulation	Micro	RNN	[57, 67, 71]
		CNN	[67]
	Macro	ODENet	[85]
Multi-scale simulation	Micro2macro	RNN	[210]
		GP	[158]
	Mixed	GNN	[117]
		ODENet	[197]

Table 3. Data-driven modeling and simulation methods

- **RNN**: RNN is dedicated to capturing the temporal dependencies, by learning the conditional distributions for current and future temporal status based on the past time series. The MixSeq [210] forecasts the macroscopic time series by aggregating a mixture of microscopic temporal sequences predicted by the seq2seq model. The authors assume that several successive microscopic time series follow a mixture of distributions. Hence cluster them into different components before modeling each component with a seq2seq model configured by specific parameters. It is theoretically shown that with the appropriate estimation for each component, the estimation for the macroscopic time series is enhanced. The experimental results show that MixSeq performs well on both synthetic and real data. The AI-ATS algorithm proposed in [71] utilizes LSTM and GRU to learn the optimal time step lengths and step jumps for the discrete simulation of platelet dynamics in shear blood flow. By adaptively choosing the simulating time step size and jump, AI-ATS reduces the unnecessary redundant calculations for efficient computation and attains comparable accuracy performance to the standard algorithm. In [57], the authors replace the micro model in a finite element simulation with an RNN framework for acceleration in materials. By sampling the results generated by the finite element method, they train the RNN model to understand the historic status, which resembles a nonlinear finite element analysis. Experiments reveal the qualification of RNN as a surrogate.
- **CNN**: For the simulations considering spatial characteristics and correlations, the CNN is regarded as a candidate for its capability of describing spatial dependencies. RCNN is proposed in [67] for Computational EM (CEM) to play as a replacement for the Finite Difference Time Domain (FDTD) method. It takes both RNN and CNN to imitate the complex computations along time frames and space grids for electromagnetic field simulation in a microscopic view.
- **GNN**: The GNN is suitable for representing a set of entities, their attributes, and their relationships with each other. The MultiScaleGNN [117] treats the concerned physical domain as an unstructured set of nodes, and edges are the distance between each pair of nodes. By establishing several graphs corresponding to different scales of spatial resolution, and realizing the message passing forwardly and inversely between these graphs, MultiScaleGNN is able to perform as a multi-scale simulator for continuum mechanics, which is tested on the advection problems and the incompressible fluid dynamics.
- **ODENet**: The Ordinary Differential Equations Network (ODENet) [36] is regarded as a replacement of ResNet [74], which depends on the neural network to represent the derivative of variables. Furthermore, it is capable of modeling the time series with irregular time intervals, showing its superiority over RNN. The authors in [197] utilize the ODENet to take place of

the chemical master equations for releasing the burden on both memory and computation costs in the single proliferative compartment model. Moreover, the ODENet executes the model reduction in a gene network with autoregulatory negative feedback through multi-time scales. In [85], the ODENet is considered for modeling both the classical Lokta-Volterra equations in ecology and the chaotic Lorenz equations in atmospheric turbulence. It embeds the hidden dynamics buried within the massive time sequential data by integrating both machine learning algorithms and ODE modeling.

- **GP**: As a typical probabilistic machine learning technique, GP attempts to formulate the observed data with a stochastic process based on Gaussian distribution. When analyzing the stress distribution of a plastic strain band loaded with a transverse tension, the GP method replaces the nested micro-models realized by the finite element approach [158]. It builds a reduction framework for accelerating the computation on a microscale before integrating for macroscopic analysis.

In order to better understand the relationships of the mentioned methods with the simulation scales, we further provide a categorization as listed in Table 3. Though only several representative works are introduced, it indicates the feasibility and appropriateness of utilizing the data-driven methods in the multi-scale simulation. To sum up, both RNN and ODENet are suitable for temporal simulation, while ODENet is able to handle irregular time interval cases. CNN can deal with the simulation task from the spatial perspective. GNN is a wise choice when the concerned simulation problem can be treated as a graph. GP represents the condition when the simulated solution can be fitted by the stochastic distribution. The data-driven methods accelerate the computational process. However, it depends on obtaining the training data as a prerequisite.

5.2.2 Data-driven multi-scale interaction mechanisms. Besides the utilization of data-driven methods in the simulation of single-scale and multi-scale physical problems, the ways in which the data-driven methods bridge and integrate different scales are discussed here. We list some representative works in Table 4 for a clear view.

Category	Scale	Method	Papers
Cross-scale interaction mechanisms	Micro2macro	GP	[151]
		DNN	[182]
	Macro2micro	VAE	[19, 90]
	Mixed	BN	[69]
		PCA	[125]
		RL	[9]

Table 4. Data-driven multi-scale interaction mechanisms

Peirlinck et al. [151] model the heart failure process of pigs from multi-scale perspectives. It propagates the measured uncertainties from the cellular level to the organ level by training a Gaussian process regression. To be specific, it decomposes the heart volume analysis into longitudinal growth monitoring, and correlates such longitudinal growth with the measured cellular level uncertainties via the Gaussian process regression. Wang et al. [182] first design a reduced-order model for generating coarse-grained simulation data in flow dynamics. Then, they take advantage of the Deep Neural Network (DNN) to learn the correlation between the coarse data and the observed fine-scale data. GINN [69] is a hybrid approach that combines deep learning with Probabilistic Graphical Models (PGM). It leverages the Bayesian Network (BN), one of the typical PGMs to

describe the conditional relationships for key variables across multi-scale in the supercapacitor dynamics. With the data generated by BN, a DNN is developed to establish the correlation between inputs and outputs. [9] introduces the machine learning applications in Computational Molecular Design (CMD) from four aspects: property estimation, catalysis, synthesis planning, and design methods. It creatively proposes the usage of Reinforcement Learning (RL) to link the microscopic molecular design with the macroscopic product characteristics. Concretely, the approach selects the candidate design based on a generative model. On passing the decision to a physical simulation system, the current state and reward are produced, which further guides the evolution direction for the RL algorithm. GLIMPS [125] implements the resolution transformation in molecular models based on the Principle Component Analysis (PCA). It captures the structural characteristics of both low-resolution and high-resolution data, and employs the General Linear Model (GLM) to link them. In [19], a dynamic-importance sampling framework is proposed for bridging the macro and micro scales. It explores the phase space of a macro model, then leverages the Variational Autoencoder (VAE) to project the data onto a latent space, and finally matches the simulations on the microscale. Furthermore, it creates a self-healing mechanism for adjusting the parameters of the macro model inspired by the micro-scale simulation. [90] follows a similar technical route as that mentioned in [19] to reveal the dynamics of RAS signaling proteins. The scale bridging is realized by mapping the lipid composition into a latent space, to instantiate the molecular simulations on microscales. There is also a step for refining the macro model parameters based on the feedback.

In summary, the data-driven methods are capable of linking the micro and macro scales. Approaches like GP and BN establish the bridge crossing the scales based on the stochastic distribution or probability theory. PCA and VAE realize the dimensional reduction for matching the feature space from high dimension to low dimension. DNN links the multi-scale simulation via a neural network, while RL observes the current system state and reward on the level of one scale, and provides the action based on the design from another scale. The data-driven methods, mainly supported by machine learning techniques, manage to correlate features across the scales as a result of the ability to construct nonlinear relationships.

5.3 Data and knowledge jointly driven method

In the domain of multi-scale simulation, for the purposes of high efficiency and fidelity, the data-driven and knowledge-driven methods are considered together, taking advantage of both sides.

Micro-scale simulation	Macro-scale simulation	Papers
Data-driven	Knowledge-driven	[151, 187]
Knowledge-driven	Data-driven	[9, 122]
Mixed	Mixed	[108]

Table 5. Data and knowledge jointly driven methods

The characterization and analysis of the heart failure procedure are conducted in [151]. It models the growth of heart by a stretch-driven growth method, and quantifies the uncertainties via machine learning approaches. The measurements are carried out on the hearts of six pigs. With the help of Bayesian inference, the authors predict the uncertainty probability during the cardiac growth process. NH-PINN, which is proposed in [108], converts the multi-scale equation into a homogenized one merging the multi-scale properties. By utilizing the Physics-Informed Neural Network (PINN) for sequentially solving the cell problems and the homogenized equation, NH-PINN manages to improve the accuracy for solving multi-scale problems. In [9], the authors propose a framework

based on RL for molecular design. The current decision is determined by the RL algorithm, while the state and reward are generated through a knowledge-based physical simulation system. To tackle the general 3D problems with arbitrary material and geometric non-linearity, [122] proposes a two-phase framework based on the Deep Material Network (DMN). It first derives the experimental data for training and testing by Direct Numerical Simulation (DNS) tools in the offline stage. Later, it trains the DMN model to learn the relations between the microscale nonlinear properties and the overall macroscale variables, for the purposes of material design and concurrent simulations. In the field of biology, the authors in [187] develop a framework that combines knowledge and data-driven modeling in the multi-scale analysis of the reconstruction of the sheep tibia over a period of 12 months. Specifically, the overall model consists of the Finite Element (FE) analyzer at the macroscopic level, along with one neural network and another series of neural networks performing as the replacements at the microscopic level. The results reveal a comparative performance of the proposed model in accuracy, while a superb performance in efficiency.

The import of the data-driven methods into the physical knowledge dominant simulation is a seamless combination. It not only accelerates the simulation efficiency by learning the relations hidden in the data, but also guarantees the simulation fidelity by obeying the knowledge rules. We believe that the data and knowledge jointly driven simulation method will attract more and more attention, with applications in multiple research areas.

5.4 Method Comparison

In this section, we further discuss the limitations of methods of multi-scale modeling and simulation. To better compare these methods, we further divide the knowledge-driven methods into two categories, i.e., mechanistic models and phenomenological models [1, 43]. The mechanistic model, usually established at micro scales, aims to describe the fundamental principles of the system to facilitate first-principle simulation. Differently, phenomenological models are usually established at macro scales and focus on capturing observed trends, patterns, or behaviors of the system [43]. As the difference compared in Table 6. We can observe that data-driven methods have relatively smaller computational complexity compared with most knowledge-driven methods, but they have higher overfitting risks and lower explainability. On the other hand, phenomenological knowledge-drive models have smaller computational complexity and smaller overfitting risk compared with mechanistic knowledge-drive models, but they have the lowest modeling ability, resulting in an incapacity to capture numerous details of the system. Altogether, by incorporating both knowledge and data, the data and knowledge jointly driven methods aim to simulate the complex system with low computational complexity, low overfitting risk, high explainability, and high modeling ability.

Categories of Methods	Computational Complexity	Overfitting Risk	Explainability	Modeling Ability
Knowledge-driven (mechanistic)	High	Medium	High	High
Knowledge-driven (phenomenological)	Low	Low	High	Low
Data-driven	Low	High	Low	High
Data and knowledge jointly driven	Low	Low	High	High

Table 6. Comparison of different methods of multi-scale modeling and simulation

Categories	Method	Papers	Particle/field based
Domain decomposition	knowledge-driven	[24]	Particle & Field
	data-driven	[121, 141, 179, 181, 198]	Particle & field
Multigrid method	theory-guided	[42, 62]	Field
	technique-introduced	[204]	Field

Table 7. Multi-scale modeling and simulation methods to achieve *Objective A*

6 OBJECTIVE-ORIENTED MULTI-SCALE SIMULATION METHOD

6.1 Methods to achieve Objective A — Resolving local fine-grained dynamics

Simulating microscale dynamics with fine-grained information requires huge computational costs leading to obstacles before utilization. One promising solution is that the fine-grained modeling is only applied to key parts of the system while the coarse-grained modeling is used in the remaining part. The methods to achieve *Objective A*, which resolves the localized region of interests into microscale for the simulation of fine-grained behaviors, are mainly categorized into two subcategories: *domain decomposition-based methods* and *multigrid-based methods*. The domain decomposition based methods segment the regions of numerical calculations, calculate them separately, and then integrate them [185]. The multigrid-based methods, which divide the problem to be solved with different scales of grid splitting, focus on eliminating errors and unnecessary computations by interactions between scales [185]. Recently, data-driven modeling methods, such as deep learning, have been widely adopted to integrate with domain decomposition-based and multigrid-based methods to enhance the modeling capabilities and computational efficiency, which demonstrates the advantages of synergistic modeling of knowledge and data. These works will be highlighted in the following paragraphs.

Domain decomposition based methods. As a classical multi-scale modeling schema, domain decomposition has played a great role in chemistry [24], biology [198], and physics [121] fields. A widely accepted strategy is to use fine-grained accurate models in the local regions of interest and the coarse-grained representation in other regions [48]. CHARMM [24] is a multi-scale molecular simulation platform on which multi-field computational simulation experiments can be performed. The quantum mechanics/molecular mechanics (QM/MM) model, which applies the QM model to the region of interest and the MM model to the remaining region, can achieve precision and efficiency at the same time. Besides, there are many data-driven methods, including variational autoencoder (VAE) [98], generative adversarial networks (GAN) [61], clustering [194] and deep reinforcement learning (DRL) [139], utilized to fit the distribution of microscale configuration based on macroscale quantities [121, 141, 179, 181, 198] in order to handle the heterogeneity of the complex systems. Since recurrent neural networks (RNNs) [138] have the capability of modeling temporal evolution, Wu et al. [189] also model the state changes of microscopic materials affected by history and the macroscale background field using RNNs.

Multigrid-based methods. The multigrid method is an effective computational solution for PDEs (partial differential equations) thanks to its iterative computation and the multi-scale property, which leads to the popular utilization in complex system simulation [137]. In the study of turbulent flow simulation, scale transfer operators implemented through plain aggregation algebraic multigrid method could improve the computational efficiency and accuracy [62]. Yushu et at. [204] use image processing techniques in the multigrid method to improve the computational speed of its modeling multi-scale complex systems. Dai et al. [42] propose that the multi-scale multigrid method is a

high-accuracy solution for PDEs and design a cascadic multigrid method to further enhance the efficiency.

6.2 Methods to achieve Objective B — Completing macroscale mechanisms

Categories	Method	Papers	Particle/field based
Knowledge-driven	FEA	[150, 169]	Fields
	REV	[203]	Field
Data-driven	NN surrogated DFT	[8]	Field
	NN surrogated FEA	[105, 192]	Field
Knowledge and data jointly driven	NN enhanced FEA	[70]	Field
	FEA guided NN	[149]	Field

Table 8. Multi-scale modeling and simulation methods to achieve *Objective B*

As mentioned before, the macroscopic simulation on coarse spatio-temporal granularity usually has the drawback of low accuracy due to the lack of microscopic fine information. In particular, it is often difficult to obtain high-precision solutions when encountering systems with strong heterogeneity. The *Objective B* proposed in this literature brings clean fine-grained information to macroscale modeling to promote the simulation, which can be classified into three subcategories: knowledge-driven, data-driven as well as knowledge and data jointly driven methods.

Knowledge-driven methods. The most commonly used knowledge-driven methods include finite element analysis (FEA) [87] and representative elementary volume (REV) [79]. Sridhar et al. [169] use microdynamics differential equations to model the microscopic local resonance phenomena, which is integrated with the macroscale model for continuum materials. Özdemir et al. [150] utilize FEA to model local microstructural heterogeneities and thermal anisotropy with the continuum macroscale. Yuan et al. [203] introduce discretized eigendeformation fields at different scales to improve the homogenization method, in which the elasticity solutions of unit cells are precomputed and aggregated to higher scales.

Data-driven methods. In data-driven methods, machine learning models have replaced traditional methods with high computational costs. Density functional theory (DFT), which models the quantum mechanism numerically with heavy computation, is replaced by neural networks in [8]. In [105], the neural networks take place the FEA to approximate the surface response and to compute the macroscopic stress and tangent tensor components. Xie et al. [192] utilize graph neural networks, which take the microscale connection of atoms in the crystal and output the material properties, to accelerate the computation.

Knowledge and data jointly driven methods. Recently, physical knowledge and machine learning models have been combined to model complex systems [95]. The deep material network [70] is typical of combining neural networks and REV to model material heterogeneity. The FEA and neural networks are integrated together to model the multi-scale process of bone remodeling. Oishi et al. [149] extract the inherent regularity in a computational mechanics application through machine learning models, and the regularity is then leveraged to design a novel FEA method.

6.3 Methods to achieve Objective C — Co-simulating microscale and macroscale dynamics

In some scenarios, both the computational efficiency and fine granularity of simulation are important in studying the whole system. However, fine-grained simulation will inevitably have a

Categories	Method	Papers	Particle/field Based
Knowledge-driven	Boundary-constraint	[29]	Particle & field
	Overlap interaction	[66, 173]	Field
Data-driven	Sampling-based	[19, 90]	Particle & field
	Fitting-based	[51, 72, 117−119]	Particle & field

Table 9. Multi-scale modeling and simulation methods to achieve *Objective C*

computationally intensive limitation due to the high level of information granularity [44]. Therefore, works [19, 29, 44, 66, 72, 90, 173] develop techniques to exploit multi-scale modeling and specially designed interactions between scales to achieve both fine-grained and efficient simulation. Consistent with the overall review, we grouped this literature into two categories, i.e., knowledge-driven and data-driven.

Knowledge-driven methods. Traditionally, molecular dynamics (MD) is applied as a high-precision simulation method at the atomic level and is used as the basis for fine-grained simulation in many applications. Coarse-grained simulations, on the other hand, usually focus on different spatio-temporal scales in different tasks. Examples include the lipid scale for RAS protein simulations and the continuum scale for heterogeneous material simulations. In order to achieve the numerical integration effect with high accuracy, Tao et al. [173] propose the two scale flow method combining fast and slow variables. For modeling the crack propagation of heterogeneous materials, Chakraborty et al. [29] design a physics-based continuum crack evolution model called concurrent atomistic-continuum computational model, which consists of FEM and MD. The MD is improved through a novel strain-boost hyperdynamics accelerated time marching scheme to bridge the time gap between different scales. In the electron flow simulation, the Navier-Stokes (NS) equations have been widely used in classic continuum scales while the MD is applied in the microscale for more accurate modeling [66]. In [66], Guo et al. divide the simulation into three parts: the near-wall part, the bulk part, and the overlap part. The near-wall part and the bulk part use the NS equations and MD separately. The overlap part models the coupling between the remaining two parts with a stochastic Eulerian-Lagrangian method.

Data-driven methods. There are several works [19, 72, 90] introducing data-driven methods to enhance the coupling of scales in multi-scale modeling, where the main purpose is to more efficiently interact between scales. Bhatia et al. [90] and Helgi et al. [19] adopt variational auto-encoder to embed a latent space for different regions for phase scale in bioprotein simulation and utilize a novel sampling method to select important regions for fine-grained simulation via MD. Recently, Deep Potential [72] leverages a neural network to model the potential surface of atoms in a molecular dynamics system, which accelerates the acquisition of atomic potential energy and achieves the quantum mechanical level of accuracy. With the development of the utilization of GNN in atomic simulation, Lino et al. integrate the idea of multi-scale in simulating the continuum mechanics through specially designed GNN [117], and there are works follow this novelty [51, 118, 119].

6.4 Methods to achieve Objective D — Discover unknown scales

Methods to achieve *Objective D* reduce high-dimensional complex systems to scales that are easier to understand and investigate. We introduce existing methods to solve the two subproblems of *Objective D*, respectively.

Scales with linear properties. High-dimensional complex systems are highly intractable. To address the problem, a series of methods are proposed to obtain simple scales of complex systems. Among them, a typical method is dynamic mode decomposition (DMD) [103, 156, 164, 175], which is

Category	Method	Examples
Scale with linear properties	Dynamic mode decomposition [103, 156, 164, 175]	[68, 77, 165]
Scale with preserved relevant information	Causal emergence	[63, 81]
	Coarse-scale partial differential equations	[14, 107]
	Slow variables	[21, 113, 148, 154, 166, 184]
	Governing equations	[26, 31, 34, 86]
	Renormalization group	[45, 100, 114]

Table 10. Multi-scale modeling and simulation methods to achieve *Objective D*

a data-driven dimensionality reduction algorithm proposed by Peter Schmid [164]. DMD processes time series data by computing a set of dynamic models that are separately associated with a fixed oscillation frequency and decay/growth rate. The linear treatment largely reduces the dimension of complex systems without the loss of accuracy, and DMD models are widely applied in spatiotemporal fluid flow analysis [68, 77, 165].

Scales with preserved relevant information. Researchers seek to identify causal emergence for multi-scale complex systems. For example, Griebenow et al. [63] utilize spectral clustering to find the right scale of a complex network with causal emergence. Researchers also discover that macro scales defined by coarse graining the microscale elements show surprising causal emergence [81]. By defining and evaluating the metric of effective information (EI) both at the micro and macro scales, causal interactions are observed to peak at a macroscale. The coarse scales with higher causalities facilitate investigations into complex systems more effectively. Data-driven models [14, 107] are proposed to identify coarse-scale partial differential equations (PDEs) that describe the system evolution at the macroscopic level. A series methods to identify slow variables in dynamical complex systems [21, 148, 154, 166, 184] are proposed, such as variational approach [148], autoencoders [113, 184], diffusion maps [166]. Meanwhile, governing equations [26, 31] or common principles [34, 86] of the system are extracted. The discovery of governing equations and common principles is essential in data-rich fields that lack well-established theories. Recently, the renormalization group (RG) method has shown great potential to deal with high-dimensional systems [45, 100, 114]. The RG method is intimately related to scale invariance and conformal invariance, symmetries in which a system appears the same at all scales, i.e., so-called self-similarity.

6.5 Methods to achieve Objective E — Learning dynamics at the discovered scale

Methods to achieve *Objective E* solve dynamical complex systems without clear scales by learning effective dynamics. To deal with the challenge of high dimensions, there are two types of methods to learn effective dynamics at a lower-dimensional scale, which can be further utilized to model dynamical systems.

Knowledge-driven methods. Knowledge-driven methods leverage mathematical models to learn dynamical complex systems. The widely used approaches include wavelet-based methods [18, 58, 128, 135], multi-scale homogenization [38, 110, 202], and stochastic methods [49, 84]. For example, the wavelet representation [128], which is an orthogonal multiresolution representation computed by a pyramidal algorithm, can be applied to various scenarios of scale compression. To deal with large spaces, Jacob et al. [49] propose a reduced-order homogenization method combined with the Karhunen−Loeve expansion and stochastic collocation method. Molei et al. [173] also

Category	Method	Examples
Knowledge-driven	Wavelet-based methods	[18, 58, 128, 135]
	Multiscale homogenization	[38, 110, 202]
	Stochastic methods	[38, 110, 202]
Data-driven	Autoencoders	[52, 78, 106, 176, 177]
	Neural ODEs	[36, 131, 132]
	Deep Koopman neural networks	[113]
	Sampling strategies	[19, 32, 46, 60, 129]

Table 11. Multi-scale modeling and simulation methods to achieve *Objective E*

develop a new class of multi-scale integrators for stiff ODEs and SDEs for Hamiltonian systems with hidden dynamics.

Data-driven methods. Data-driven methods leverage machine learning or deep learning models to learn the effective dynamics of complex systems. Vlachas et al. [177] accelerate simulations of molecular systems by learning dynamics at the coarse scale modeled by autoencoders and applying a probabilistic mapping between coarse and fine scales via a mixture density network. Some other studies also use autoencoders to model dynamical systems [52, 106, 176]. Besides, variational autoencoder (VAE) can be utilized to learn latent-space dynamics of high-dimensional systems [78]. This is achieved by reducing complex, nonlinear processes to a single embedding with high fidelity. The performance of long short-term memory networks (LSTMs) [80] and neural ordinary differential equations (NODEs) [36] are explored in learning latent-space representations of dynamical equations [131, 132]. Deep Koopman neural networks [25] and LSTM are utilized to model the dynamics of slow variables and fast variables, respectively [113]. In addition, sampling strategies [19, 32, 46, 60, 129] are proposed for modeling multi-scale dynamics. For example, Harsh et al. [19] present a dynamic-importance sampling approach for adaptive multi-scale simulations. It uses machine learning to dynamically sample the phase space by a macro model using microscale simulations, which enables automatic feedback from the micro to the macro scale.

7 APPLICATION OF MULTI-SCALE SIMULATION IN TYPICAL COMPLEX SYSTEMS

7.1 Matter systems

7.1.1 Atomic and molecular systems. Traditionally, Car-Parrinello molecular dynamics (CPMD) is the widely adopted multi-scale simulation method for calculating the configuration of the atomic and molecular system, where the density functional theory (DFT) [28] deals with the electronic level quantum mechanics and the classic mechanics are modeling with Newton's laws. Even though CPMD will bring very high simulation accuracy, the computation will be very large and the utilization is limited. To achieve both effectiveness and efficiency, CHARMM proposed that the local region of interest could be different from the remaining region in simulating methods, which is called quantum mechanics - molecular mechanics (QM-MM) [24]. Recently, Weinan et al. propose deep potential [72], which introduced deep learning techniques into the complex system simulation, and the effectiveness and efficiency are both obtained at the system level. Following this, GNNs are also introduced to solve the complex problem of turbulence modeling and continuum modeling [51, 117–119].

7.1.2 Biomolecular systems. Multi-scale modeling provides better solutions in many complicated tasks, such as novel protein structure discovery, remodeling of membranes, and biological stress modeling. For example, in [19], which is aimed at achieving simulation with macroscale space size,
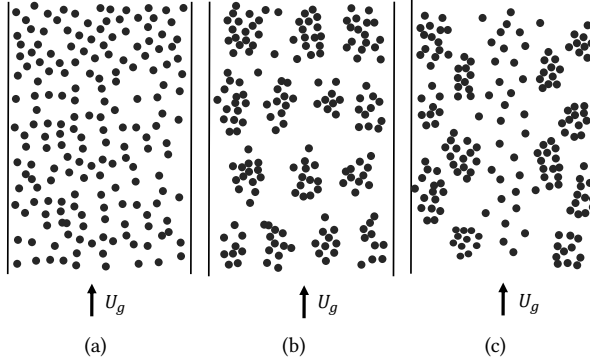
Fig. 4. Three different gas–solid two-phase flow with the same amount of solids and the same gas flow rate U_g. The different mesoscale structures lead to different properties of the system [112].

macroscale temporal length, and microscale precision, the different scales interact with each other via the consistency at the boundary between different scales. Stansfeld et al. design a multi-scale method to simulate the assembly and the interactions of membrane protein/lipid, in which the macroscale is at molecular dynamics while the microscale reaches atomistic resolution [171]. In the collagen modeling application, Masic et al. reveal how the physics of the microscale at the molecular level contributes to the mechanics at the macro behavior level [130].

7.1.3   Chemical systems. Chemical systems are also a category of systems with multi-scale structures. For example, a widely existing complex system in nature and engineering with multi-scale structures is the gas–solid two-phase flow, which is a non-equilibrium system with significant heterogeneity. Specifically, there are two mechanisms in this system, including mechanisms of particles and gas. Then, in a concurrent-up gas–solid two-phase flow as shown in Figure 4, the gas tends to find an upward path with minimum resistance. If the mechanism of gas dominates the systems, which occurs when the flow velocity is particularly large, particles will be separated as much as possible to minimize the resistance to gas, thereby minimizing energy consumption for transporting and suspending particles per unit volume of gas W_{st}. On the other hand, the particles tend to minimize their potential energy. If the mechanism of particles dominates the systems, which occurs when the flow velocity is particularly small, particles will be as concentrated as possible to minimize their potential energy, thereby minimizing the voidage ϵ of the gas–solid flow. However, in most scenarios in nature and engineering, the flow velocity is neither too large nor too small, and neither mechanism is able to dominate the system. Thus, two mechanisms co-exist and compete with each other in the system, which causes a compromise in reaching the optimal states corresponding to the two mechanisms, i.e., minimal W_{st} and minimal ϵ, respectively. During the competition of the two principles of minimum potential energy for particles and minimum resistance to gas, the mesoscale structure is formed, i.e., dense particle clusters surrounded by the dilute broth, which are denoted by dense phase and dilute phase, respectively.

From the microscale, by only observing the interaction between each particle and its surrounding circular area, we can only observe two different mechanisms corresponding to the dense phase and the dilute phase. From the macro scale, only the global state variables cannot fully characterize the system. For example, as we can observe from Figure 4, the three systems have the same global state variables including the same gas flow rate U_g, and the amount of solids. The different mesoscale

particle cluster sizes lead to the difference between Figure 4(a) and (b); The existence of a core-annulus flow further leads to the difference between Figure 4(b) and (c). Thus, the mesoscale structure must be observed and modeled to fully characterize the complex system.

How to model the mesoscale structure of the gas–solid two-phase flow has been studied in numerous existing studies. Li and Kwauk et al. [112] propose the energy-minimization multi-scale model (EMMS) model. In this model, the Gas–solid two-phase flow is described by eight variables, of which the most important variables include the fraction of dense phase f, the voidage of the dense phase ϵ_c, the voidage of the dilute phase ϵ_d, and the diameter of particle clusters d_{cl}, etc. They find that instead of minimizing the energy consumption for transporting and suspending particles per unit volume of gas W_{st} or the voidage ϵ in the system, the energy consumption for the suspension and transportation of particles with respect to unit mass of particles N_{st} is minimized. Further, Ge and Li et al. [56] propose a numerical algorithm to solve this problem. Specifically, this algorithm defines a two-dimensional solution space of (ϵ_c, ϵ_d). For each pair of (ϵ_c, ϵ_d), this algorithm searches for the value of f such that U_{si} calculated based on its definition equals U_{si} calculated based the balance conditions including mass balance, momentum balance, and pressure balance, etc. Specifically, U_{si} represents the slip velocity between two phases. Then, the problem can be solved by traversing all feasible (ϵ_c, ϵ_d) within the solution space and finding the $(\epsilon_c^*, \epsilon_d^*)$ with minimum W_{st}. The EMMS model is further extended to a more general multi-objective variational methodology [111, 140]. Based on this methodology, the target system can be mathematically expressed by the multi-object optimization (MOP) problem. Mo et al. [140] validate the consistency of the solutions obtained based on the EMMS model compared with the solutions obtained based on the MOP model through numerical analysis. What's more, the MOP model can further explore those solutions not found by the EMMS model.

7.2 Social systems

7.2.1 Transportation systems. In the field of transportation, the traffic flow theory is to explore the spatial and temporal patterns of traffic, which has developed several related models. These models can be classified into two groups: microscopic models and macroscopic models. The microscopic models aim to characterize the interactions between the vehicles, while the macroscopic ones focus on the aggregated behaviors of a group of vehicles. Their basis is to investigate the relationship between the density, speed, and flow volume. Specifically, the density ρ is the average number of vehicles passing a unit length of the road segment. The speed v is the average speed of all the vehicles in a road segment. The flow volume q is the average number of vehicles in a road segment in a time slot. Their relationship can be formulated as follows:

$$q = \rho \cdot v. \tag{8}$$

- **Macroscopic traffic models**. Such models focus on the macroscopic characteristics of traffic flow, which regards the traffic flow as the continuous fluid composed of a large amount of vehicles. Based on the theory of fluid mechanics, the researchers derive the partial differential equations for the speed and density to characterize the traffic dynamics. Formally, according to the conservation principle of traffic flow, it satisfies the following continuity equation:

$$\frac{\partial(\rho v)}{\partial x} + \frac{\partial v}{\partial t} = 0, \tag{9}$$

where $\rho(x, t)$ and $v(x, t)$ are the flow density and average speed at location x at time t, respectively. This equation is built upon the traffic fluid mechanics model called Lighthill-Whitham-Richards (LWR) model [116]. Particularly, for the speed $v(x, t)$, the LWR model assumes that the traffic speed and density satisfy the relationship of $v(x, t) = v_e(\rho(x, t))$. Thus, the Equation 9

Categories	Research Problem	Method	Macroscopic	Microscopic	Papers
Spatial	How traffic performance can be improved is unclear (Objective C)	Multi-scale model predictive control approach	Network level perimeter control	Local level signal control	[196]
	Unclear the connectivity relationship within a multi-modal urban transportation network (Objective B)	Multi-modal and multi-scale GIS-T data model	Bus, metro and street transit networks	Bus route segment, metro line, street segments, walking links	[37]
	Unclear changes of future traffic (Objective B)	Hierarchical graph convolution network	Regions	Road segments	[65]
	Unclear changes of future traffic congestion (Objective C)	Cross-scale spatiotemporal GNN	Regions	Road segments	[180]
	Unclear mechanism of the heterogeneous traffic streams (Objective B)	Mean field games theory	Network-level traffic streams	Individual vehicle behavior	[94]
Temporal	Unclear mechanism of driver behavior (Objective A)	A combined CNN-RNN based multi-task learning model	Long time-scale mental cognitive process	Short time-scale driver's physical behaviors	[193]
	Unclear changes of future traffic (Objective B)	Multi-Step dependency relation model	Short time-scale traffic	Long time-scale traffic	[120]
	Unclear mechanism of urban freight operations (Objective A)	Multi-scale agent-based simulation model	Long-term strategic decision	Short-term operational decision	[7]

Table 12. Multi-scale modeling and simulating methods in transportation.

can be rewritten as

$$(v_e + \rho \frac{\partial v_e}{\partial \rho}) \frac{\partial(\rho v)}{\partial x} + \frac{\partial v}{\partial t} = 0. \tag{10}$$

The above equation captures the nonlinear traffic features such as the formation and propagation of shock waves in traffic flow, which facilitates to explore various complex traffic phenomena, e.g., the stop-and-go trend in traffic congestion.

- **Microscopic traffic models**. These models are to learn the moving states of each vehicle in a certain traffic environment, including car-following models and cellular automaton models. The car following models study the dynamic process of each vehicle following the leading vehicle based on the assumption that there is no overtaking phenomenon. Traditionally, their general mathematical expression can be shown as follows,

$$\frac{dv_n(t)}{dt} = f(v_n(t), \Delta v_n(t), \Delta x_n(t)), \tag{11}$$

where $v(t)$ denotes the speed of vehicle n, and $\Delta v_n(t), \Delta x_n(t)$ are the speed difference and displacement difference between vehicle n and its front vehicle. $f(\cdot)$ represents the car following model. To implement $f(\cdot)$, many researchers propose various methods by modeling drivers' behavior of acceleration or deceleration [13, 55, 76, 143, 145, 152, 174, 190, 191, 207]. For example,

Reusche and Pipes [155] proposed the first car following model

$$\frac{dv_n(t + \tau)}{dt} = \lambda \Delta v_n(t), \tag{12}$$

which considers the strength of the driver's response to the stimulus he receives from his driving environment. In this equation, τ is the delayed response time of the vehicle, and λ is the strength coefficient of a driver in response to the stimulus.

The cellular automaton models are discrete in both space and time, thus it is suitable for parallel computation to accelerate the simulation of microscopic traffic flow. The simplest cellular automaton model is one-dimensional where a single road segment is divided into discrete regular grids. Each grid represents a cell with the finite states. At any time, the state of a cell is either empty or occupying a vehicle. Further, the state of a cell at the next time is determined by itself and the state of its neighboring cells at the current time. Since a cell cannot have more than one vehicle, we define the following rule: if the cell ahead is empty, the vehicle can move one step forward; otherwise, it has to keep waiting. According to this rule, all vehicles could update their states at the next time. Based on it, several extended models [30, 40, 54, 75, 97, 172] are proposed to capture more realistic phenomena that occur in traffic flows, which helps to better understand the traffic dynamics and implement the effective traffic control.

Although we can adopt the microscopic traffic models to infer the dynamics of massive microscopic states in transportation systems, it is not easy to derive the macroscopic structure emerged from these states. One important problem is that we cannot accurately realize the scale at which the macroscopic structure will emerge. This is because the accumulated noise from the microscopic scale will prevent us from identifying the macroscopic structure. To solve it, multi-scale modeling is proposed to fully utilize the advantages of microscopic and macroscopic scales for various traffic applications. Through the reviews of recent works, we classify them into two groups in terms of multi-scale properties, i.e., space and time, as listed in Table 12.

- **Spatial multi-scale modeling**. The works belonging to this type adopt different scales of space to design the corresponding models. Authors in [65] first used the spectral clustering method to construct the macro graph of regions from the road network, and then proposed the hierarchical graph convolutional network to forecast the traffic by operating both the micro and macro traffic graphs. Yang et al. [196] proposed a model predictive control based approach to model the interaction between the network level control and the local level control to optimize the performance at the local and the network level as a whole. Chen et al. [37] developed a multi-scale and multi-modal Geographical Information Systems for Transportation (GIS-T) data model to implement the integration and multiple representations of the urban transportation networks that include bus route segments, metro lines, street segments, walking links. Authors in [94] focused on analyzing the heterogeneous traffic streams from connected vehicles and normal vehicles. They proposed a mean field game framework to establish the connection between the individual vehicle behavior and overall macroscopic traffic streams. In this framework, the connected vehicles controlled microscopically can influence and control overall traffic streams. Recently, a cross-scale spatio-temporal GNN [180] was proposed to forecast future traffic congestion, as shown in Figure 5. This approach highlights the advantages of multi-scale simulation methods in intelligent systems. As illustrated in this study, the complex interplay between microscopic and macroscopic dynamics of traffic congestion significantly impacts traffic congestion. Only modeling one of them cannot capture this complexity, resulting in inaccurate prediction performance. To address this challenge, the authors proposed to model the interaction between macro- and micro-dynamics of congestion, effectively addressing *Objective C*. To be specific, they adopted PINN to extract valuable macroscopic information for guiding
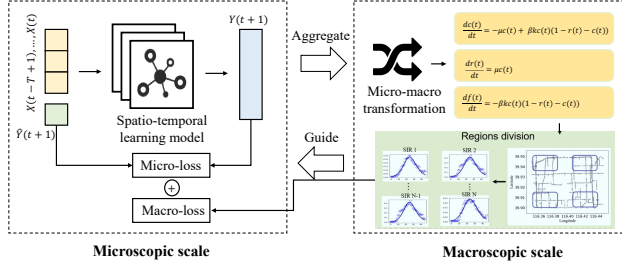
Fig. 5. The cross-scale spatio-temporal graph neural network [180] utilizes multi-scale simulation methods to forecast future traffic congestion (*Objective C*).

the prediction of microscopic traffic congestion. Meanwhile, they designed a micro-macro transformation mechanism to aggregate microscopic states into macroscopic ones in a differentiable manner. This cross-scale modeling significantly improved the prediction performance of traffic congestion.

- **Temporal multi-scale modeling**. This type of methods focuses on exploring the advantages of multi-scale modeling in the time dimension. To be specific, Yang et al. [193] proposed to jointly model the driver behaviors in different time scales for behavior reasoning. In this work, the mental cognitive process is recognized to have longer time-scale patterns, while drivers' physical behavior has fast dynamics. A multi-task framework with the CNN-RNN network is developed to process the task of multi-scale driver behavior recognition. Authors in [120] designed a multi-step dependency relation network to learn the relations between the short-time-scale traffic features and long-time-scale ones. Alho et al. [7] focused on modeling urban freight operations. They proposed a novel agent-based framework integrating the models with different time resolutions, including a long-term strategic decision model and a short-term operational decision model. The former simulates the agent decisions made for one year, and the latter simulates the fast decision-making process. The authors also implemented it on an open-source mobility simulation platform and verified the effectiveness of their proposed approach.

7.2.2 Epidemic contagion systems. Accurately modeling the transmission pattern of the epidemic has been a long-standing research topic for decades, which is attracting more and more public attention with the outbreak of COVID-19. The epidemic dynamics is tightly entangled with human mobility [91, 92, 102, 188], economic development [64, 73, 168], and vaccine distribution [35, 133, 161, 178, 199], forming a complex system. In this survey, we focus on the simulation paradigm of epidemic contagion systems of the recent advances, which can be divided into 3 categories as shown in Table 13.

Categories	Scales	Method	Papers
Single-scale simulation	Macroscopic model	Regression model	[89, 91, 92, 102]
		Compartmental model	[104, 126, 159]
	Microscopic model	Contact Network	[2, 6, 142, 146, 157]
Multi-scale simulation	Multi-scale model	Metapopulation	[11, 12, 23, 33, 39, 188, 205, 206, 209]
		High-order network	[88, 170]

Table 13. Simulation methods of epidemic contagion systems

Regression model. For macroscopic simulation, regression models are widely used. During the first strike of the COVID-19 pandemic, researchers use generalized linear model (GLM) [102], multiplicative exponential model [92], or multilevel Bayesian regression model [91] to investigate how human mobility and control measures affect the early transmission of the epidemic. Specifically, the risk source model proposed in [92] is shown in (13):

$$y_i = c \prod_{j=1}^{m} e^{\beta_j x_{ji}} e^{\sum_{k=1}^{n} \lambda_k I_{ik}}, \tag{13}$$

where y_i is the number of the cumulative (or daily) confirmed cases in prefecture i; x_{1i} is the cumulative population outflow from the source region of the epidemic to prefecture i in different time periods. x_{2i} and x_{3i} are the GDP and the population size of prefecture i, respectively. λ_k is the fixed effect for province k, n is the number of prefectures considered in the analysis, I_{ik} is the dummy for prefecture i, where $I_{ik} = 1$ if $i \in k$, otherwise $I_{ik} = 0$. Finally, c and β_j are the free parameters to be calibrated. From this equation, the author accurately models the macroscopic epidemic dynamics. Similarly, the regression model in [91] and [102] are as shown in (14) and (15) accordingly. All these methods use a simple regression model to depict the aggregated epidemic outcome ($y_{i,t}$ in (14) and $Y(t)$ in (15)) with other variants such as mobility ($M_{i,t}$ in (14) and $M(t - 5)$ in (15)), population (X_i in (14)), the polymerase chain reaction (PCR) test ($IT(t)$ in (15)) and temperature ($T_{i,t}$ in (14)).

$$y_{i,t} = \alpha_{c_t} + X_i \beta + T_{i,t} \theta + C_{s_i,t} \phi + M_{i,t} \gamma_{c_i,t} + \epsilon_{i,t}, \tag{14}$$

$$Y(t) = Y(t - 4) + IT(t) + M(t - 5) + IM(t). \tag{15}$$

Compartmental model. Another type of macroscopic model for epidemic modeling is the famous compartmental model, which can be traced to the 20th century [159]. Classic compartmental models, such as the SEIR model, divide the whole population in a region into different compartments as Susceptible (S), Exposed (E), Infected (I), and Recovered (R) and use the following ordinary differential equations (ODE) to depict the aggregated transmission process:

$$\begin{aligned} dS/dt &= -\beta S I / N, \\ dE/dt &= \beta S I / N - \sigma E, \\ dI/dt &= \sigma E - \gamma I, \\ dR/dt &= \gamma I, \end{aligned} \tag{16}$$

where S, E, I, and R are the susceptible, asymptomatic, infected, and recovered people, respectively. $N = S + E + I + R$ is the total population of the region. For the parameters, β is the infection rate, γ is the recovery rate, and σ is the incubation period of the disease. Specifically, the basic reproductive number of the disease is $R_0 = \beta/\gamma$, which denotes the expected number of cases directly generated by one case in a population where all individuals are susceptible. When $R_0 > 1$, the disease can start to transmit in an exponential pattern. When $R_0 < 1$, the disease will not become an epidemic since the infected people will recover before it can transmit to more people. These compartmental models are macroscopic since the variables are the total number of people in different states, where the heterogeneous contact patterns are replaced by a homogeneous mixing assumption. It limits their expressive power since they can only capture the exponential growth pattern, which only exists during the early beginning of the epidemic and quickly changed due to the introduction of non-pharmaceutical interventions [50, 104, 167].

To extend the expressive power of classic compartmental models, Maier et al. [126] introduce a "containment and quarantine" mechanism to the classic SIR model as follows:

$$
\begin{aligned}
\partial_t S &= -\alpha SI - \kappa_0 S, \\
\partial_t I &= \alpha SI - \beta I - \kappa_0 I - \kappa I, \\
\partial_t R &= \beta I + \kappa_0 S, \\
\partial_t X &= (\kappa + \kappa_0) I,
\end{aligned}
\tag{17}
$$

where α and β are the infection rate and the recovery rate accordingly, and the self-containment rate that affects both susceptible and infected people is κ_0, which assumes the self-containment process will prevent any transmission for these parts of people and directly turn the susceptible and infected people to the new X state (which do not act with any states). Besides, there is also a quarantine rate κ that only influences the infected people. By the new mechanism that protects part of the susceptible people, the proposed model successfully extends the expressive power of the SIR model to capture sub-exponential increase patterns.

Contact network model. For microscopic modeling of the contagion system, most of the research adopts the network approach to represent individual-level contact patterns [2, 6, 142, 146, 157]. Qian et al. [157] use smart card data in urban transit systems to construct individual-level contact networks. Specifically, a classic susceptible-infectious-susceptible (SIS) model is used to model the disease transmission on the network as follows:

$$
p_{i,t} = 1 + p_{i,t-1} \left(q_{i,t} - r \right) - q_{i,t}, \forall i \in V.
\tag{18}
$$

In the above equation, $p_{i,t}$ represents the probability for node i infected at time t, and the recovery rate is noted as r. The variable $q_{i,t}$ represents the probability that node i is in the susceptible state at time t. It depends on all the neighborhood $j \in \mathbb{N}(i)$ that are either in the susceptible state or in the infected state but fail the transmission:

$$
q_{i,t} = \prod_{j \in \mathcal{N}(i)} \left(1 - p_{j,t} + \left(1 - \beta_{i,j} \right) p_{j,t} \right).
\tag{19}
$$

The parameter $\beta_{i,j} = \beta t_{i,j}$ represents the transmission rate between node i and j.

Similarly, Alberto et al. [6] construct a multi-layer network with a household layer, a workplace and community layer, and a school layer to investigate how testing, contact tracing, and household quarantine affect the disease transmission. Abueg et al. [2] simulate the contact network in Washington state by three types of sub-networks: fully connected networks for households, small world networks [183] for workplaces, schools or social circles, and random networks for random interactions. Nishi et al. [146] further investigate how interventions on the structure of contact networks affect the disease transmission. Murphy et al. [142] propose a novel graph convolution network (GCN) model to capture the contagion dynamics in complex contact networks, which is the recent advantage that combines classic epidemiological model with deep learning technology.

From the above discussion, macroscopic models [92, 102, 104] only require the empirical epidemic records at the aggregated level, which can be easily obtained in most parts of the world. However, it also forbids them to explain the underlying mechanism due to the over-aggregation of the model parameters, such as the superspreading phenomenon [3, 124]. Microscopic models greatly improve the expressive power of macroscopic models by introducing individual mobility [6, 142, 157], where the heterogeneous contact patterns explain the emergence of such phenomena. However, it also poses a huge demand for mobility data [10], which are often difficult to collect in low- and middle-income countries and raises significant privacy concerns [4]. This dilemma impels the emergence of multi-scale models that achieve a great balance between the expressive power and data demand of the contagion system.

Metapopulation model. A classic solution is the metapopulation model, where the whole simulation region is divided into several sub-regions, where each sub-region maintains a compartmental model (such as SEIR), where an aggregated mobility network connects these sub-regions. The famous global epidemic and mobility (GLEaM) model [11, 12] considers the multi-scale mobility of short-scale commuting flows and long-range airline traffic on an SEIR model, which is still inspiring recent studies [39, 205]. Wu et al. [188] propose a metapopulation model using aggregated human mobility data across more than 300 prefecture-level cities in mainland China to investigate how domestic and international movement affect the outbreak of the disease, where the number of international outbound and inbound air passengers and the daily number of all domestic outbound and inbound travelers are all considered. Then, according to the mobility network, the sub-regions are connected. The regional mobility data is easier to obtain, which greatly reduces the difficulty of applying the model in real-world scenarios. Zhou et al. [209] use anonymous mobile phone data to construct an individual-level SEIR model in a metapopulation manner, where the city is divided into multiple regions and runs different SEIR models with different regions, where the metapopulation matrix is incorporated.

Furthermore, Chang et al. [33] propose a bipartite network with time-varying edges between census block groups (CBGs) and points of interest (POIs), which represents the residential areas and venues where the transmission may occur. Based on this highly flexible framework, the authors successfully reproduce the superspreading phenomenon and evaluate the disproportionate impact of the epidemic towards disadvantaged racial and socioeconomic groups.

High-order network model. Another problem that lies in the single-scale simulation methods is the ignorance of high-order effects during the disease transmission. For example, the macroscopic compartmental models assume a heterogeneous first-order contact between different compartments, and the microscopic models only consider direct connections between individuals. However, such assumptions disobey real-world observations: the contact between family members is closer than the contact between acquaintances, which leads to a higher probability of infection. To tackle this challenge, researchers propose high-order networks, where simplicial complexes [88] or hypergraphs [170] are introduced to model the transmission patterns in a small group of people. Lacopini et al. [88] assign different infection rates to different motifs to depict the high-order pattern in a social network, where the transition of infection density becomes discontinued and bistable. Similarly, St-Onge et al. [170] also observe such discontinuous transitions with a different method, where they introduce hyperedges (an edge that connects multiple nodes) to model the high-order feature.

To better illustrate how to use multi-scale simulation methods and related theories to analyze real-world complex systems, we further present an example in terms of epidemic contagion systems of MSDNet [201]. In this example, multi-scale simulation methods are employed to predict the spatial spread of infectious diseases on a regional scale, which is mainly based on the mobility flow between different regions. However, only considering mobility flow between regions, which is regarded as the macroscale information in this scenario, fails to incorporate the intricate details of the mobility behaviors of individual users, such as face-to-face encounters. This limitation consequently results in incomplete macroscopic mechanisms derived by employing machine learning approaches to model the spatial spread of epidemics through mobility flow between different regions. Under such circumstances, it becomes instrumental to incorporate microscale information to complete macroscale mechanisms (*Objective B*). In particular, as shown in Figure 6, MSDNet extracts the microscopic feature of the user contact sub-graph of each region through graph pooling methods [200], which is then utilized to modulate key epidemiological parameters in the classical compartment model and also regarded as a "pre-trained" feature vector of the region in the
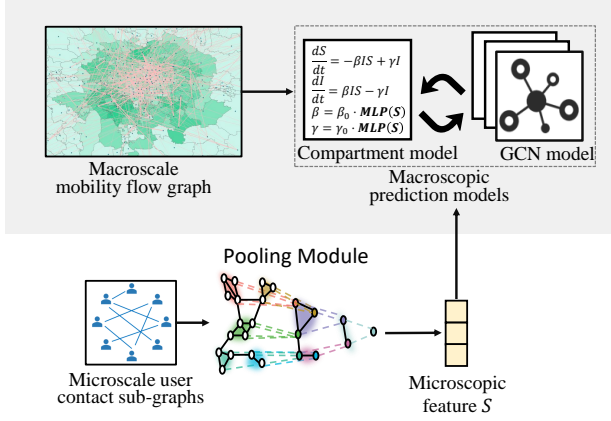
Fig. 6. MSDNet [201] utilizes multi-scale simulation methods to predict the spatial spread of infectious diseases on the regional scale, where microscopic features derived from microscale user contact graph are incorporated to complete macroscale mechanisms (*Objective B*).

macroscale GCN model. The spatial spread of infectious diseases is predicted by collaboratively utilizing the knowledge-driven compartmental model and the data-driven GCN model.

To summarize, going beyond single-scale simulation methods that are stuck in the dilemma of expressive power and data demand, multi-scale simulation uses the metapopulation technique to improve the simulation scale with aggregated level mobility data. It also reveals the underlying high-order effect of disease transmission by considering the contact pattern of a group of people, which brings the simulation of epidemic contagion systems to a new height.

8 OPEN CHALLENGES AND FUTURE DIRECTIONS

A review of the objectives, methods, and applications of multi-scale modeling and simulation for complex systems has been presented in our paper. As we discussed above, unknown mechanisms and expensive computational costs are two key problems in the simulation of complex systems. The former indicates that the mechanism of system evolution is not clear at a specified scale, and the latter refers to the extremely high computational complexity in handling high-dimensional state variables of complex systems. Multi-scale modeling and simulation techniques are regarded as a potential way to solve these problems. To be specific, we can derive the target-scale mechanism by using the knowledge from other scales, which addresses the problem of unclear mechanisms. To deal with the intractable computation, a straightforward solution is to adopt the dimension reduction technique to approximate the massive state variables. As we surveyed in previous sections, many recent works make efforts to develop different multi-scale methods to address these problems. However, we still have to face the following challenges:

- **How to extract the valuable information at the scale with clear mechanisms for deriving target-scale mechanism**. The state variables at the scale with clear mechanisms may be extremely high-dimensional, which results in unacceptable computation complexity. At the same time, there exist complex correlations between mechanisms at different scales, which makes it difficult to explore. To further illustrate this challenge, consider the example of traffic optimization in transportation systems. At the microscopic scale, optimizing traffic involves a number of variables, including individual vehicle speed, acceleration, and lane changes, all of which significantly contribute to the overall dynamics of traffic. However, these state variables

collectively generate exceptionally high-dimensional data, making computational processing prohibitively expensive. Moreover, there are intricate relationships between traffic mechanisms at different scales. For example, the behavior of individual vehicles on roads can influence macroscopic traffic patterns, such as traffic congestion during rush hours. Thus, it is challenging to extract valuable information to derive the target-scale mechanism.

- **How to balance the trade-offs between the computation efficiency and simulation accuracy**. Ideally, if we were to model each state and their interdependencies within a complex system, we could achieve very high simulation accuracy. However, this requires an enormous amount of computational resources, making it practically unattainable in real-world situations. Dimension reduction methods are able to reduce the number of variables that should be handled. They can improve the computation efficiency but may lose some simulation accuracy. A simple way mentioned before is to reduce the dimensions in the unimportant areas. However, it is hard to identify which areas are unimportant in complex systems. VAEs [78] have emerged as a potential solution to this challenge. VAEs have the capability to reduce complex, nonlinear processes to a low-dimensional embedding with high fidelity. However, it introduces uncertainty when determining the appropriate number of dimensions, and increases the overfitting risks. These factors can influence the simulation results. Thus, how to balance the trade-offs between computation efficiency and simulation accuracy is still challenging.

- **How to jointly utilize the transferable information at different scales in system simulation.** In complex systems, the states at different scales interact with each other. Thus, to simulate the states at a specified scale, we not only need to understand the evolution mechanism of this scale, but also need to explore the information from other scales. Meanwhile, the evolution mechanisms at other scales are also affected by the information from this scale. The recently emerged method PINN [95] provides a preliminary attempt toward information transfer across different scales. It integrates known physical laws as constraints during the training process. However, due to its reliance on known physical laws, its capacity for information transfer remains constrained. Thus, it is non-trivial to jointly utilize the transferable information at different scales for simulation.

Based on the above challenges, we discuss the following future possible research directions:

- **Generalized multi-scale simulation framework**. Most existing works focus on specified methods for multi-scale modeling and simulation in their own research areas. These methods are coupled with the knowledge in the specified areas, which makes it hard to generalize to other areas. A potential way is to construct a generalized multi-scale simulation framework. This framework can establish a new paradigm to guide how useful information is transferred across scales.

- **Iterative multi-scale modeling and simulation approaches**. The evolution of state variables at a scale would affect the evolution process at other scales. When simulating a specified scale, we not only consider the influence from other scales, but also explore how this scale acts on other scales. Conducting this iterative interaction allows us to better explore the evolution mechanism, which contributes to the accurate simulation. Thus, developing iterative multi-scale modeling and simulation approaches is a promising direction for future work.

- **Data and knowledge joint driven multi-scale simulation**. In multi-scale simulation, data-driven methods rely on the observation data at the corresponding scale. Their effectiveness may be greatly affected in the scenario where the observation data is not covered. Knowledge-based methods have a relatively strong generalization ability, but it is hard to model the evolution mechanism of a scale with massive high-dimensional state variables. Thus, we

propose to use the advantages of both methods by integrating data and knowledge in multi-scale simulation. Specifically, it integrates the generalization of knowledge-driven methods and the powerful learning ability of data-driven methods, which benefits exploring the complex evolution mechanism at the corresponding scale. Thus, studying data and knowledge joint driven multi-scale simulation is an important tendency for future research.

9 CONCLUSION AND SUMMARY

Multi-scale modeling and simulation techniques are playing an increasingly important role in understanding, predicting, and controlling diverse complex systems, which is essential for numerous cutting-edge technology applications. In this survey, we systematically review the literature on multi-scale simulation of major matter systems and social systems, and provide a novel taxonomy in terms of their objectives, methods, and applications. We hope that this interdisciplinary survey will help researchers make better use of existing multi-scale simulation techniques, and also encourage the development of better multi-scale simulation methods.

ACKNOWLEDGMENTS

REFERENCES

[1] Nicolas Abatzoglou, Esteban Chornet, Khaled Belkacemi, and Ralph P Overend. 1992. Phenomenological kinetics of complex systems: the development of a generalized severity parameter and its application to lignocellulosics fractionation. *Chemical Engineering Science* 47, 5 (1992), 1109–1122.

[2] Matthew Abueg, Robert Hinch, Neo Wu, Luyang Liu, William Probert, Austin Wu, Paul Eastham, Yusef Shafi, Matt Rosencrantz, Michael Dikovsky, et al. 2021. Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. *NPJ digital medicine* 4, 1 (2021), 1–10.

[3] Dillon C Adam, Peng Wu, Jessica Y Wong, Eric HY Lau, Tim K Tsang, Simon Cauchemez, Gabriel M Leung, and Benjamin J Cowling. 2020. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nature Medicine* 26, 11 (2020), 1714–1719.

[4] Norita Ahmad and Preeti Chauhan. 2020. State of data privacy during COVID-19. *Computer* 53, 10 (2020), 119–122.

[5] Giacomo Albi, Nicola Bellomo, Luisa Fermo, S-Y Ha, J Kim, Lorenzo Pareschi, David Poyato, and Juan Soler. 2019. Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives. *Mathematical Models and Methods in Applied Sciences* 29, 10 (2019), 1901–2005.

[6] Alberto Aleta, David Martin-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E Dean, M Elizabeth Halloran, Ira M Longini Jr, Stefano Merler, et al. 2020. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* 4, 9 (2020), 964–971.

[7] André Alho, BK Bhavathrathan, Monique Stinson, Raja Gopalakrishnan, Diem-Trinh Le, and Moshe Ben-Akiva. 2017. A multi-scale agent-based modelling framework for urban freight distribution. *Transportation Research Procedia* 27 (2017), 188–196.

[8] Omar Allam, Byung Woo Cho, Ki Chul Kim, and Seung Soon Jang. 2018. Application of DFT-based machine learning for developing molecular electrode materials in Li-ion batteries. *RSC advances* 8, 69 (2018), 39414–39420.

[9] Abdulelah S Alshehri and Fengqi You. 2022. Machine learning for multiscale modeling in computational molecular design. *Current Opinion in Chemical Engineering* 36 (2022), 100752.

[10] Nafis Anwari, Md Tawkir Ahmed, Md Rakibul Islam, Md Hadiuzzaman, and Shohel Amin. 2021. Exploring the travel behavior changes caused by the COVID-19 crisis: A case study for a developing country. *Transportation Research Interdisciplinary Perspectives* 9 (2021), 100334.

[11] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J Ramasco, and Alessandro Vespignani. 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences* 106, 51 (2009), 21484–21489.

[12] Duygu Balcan, Bruno Gonçalves, Hao Hu, José J Ramasco, Vittoria Colizza, and Alessandro Vespignani. 2010. Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of computational science* 1, 3 (2010), 132–145.

[13] Masako Bando, Katsuya Hasebe, Akihiro Nakayama, Akihiro Shibata, and Yuki Sugiyama. 1995. Dynamical model of traffic congestion and numerical simulation. *Physical review E* 51, 2 (1995), 1035.

[14] Yohai Bar-Sinai, Stephan Hoyer, Jason Hickey, and Michael P Brenner. 2019. Learning data-driven discretizations for partial differential equations. *Proceedings of the National Academy of Sciences* 116, 31 (2019), 15344–15349.

[15] Baruch Barzel and Albert-László Barabási. 2013. Universality in network dynamics. *Nature physics* 9, 10 (2013), 673–681.

[16] Amir Bashan, Yehiel Berezin, Sergey V Buldyrev, and Shlomo Havlin. 2013. The extreme vulnerability of interdependent spatially embedded networks. *Nature Physics* 9, 10 (2013), 667–672.

[17] Federico Battiston, Enrico Amico, Alain Barrat, Ginestra Bianconi, Guilherme Ferraz de Arruda, Benedetta Franceschiello, Iacopo Iacopini, Sonia Kéfi, Vito Latora, Yamir Moreno, et al. 2021. The physics of higher-order interactions in complex systems. *Nature Physics* 17, 10 (2021), 1093–1098.

[18] Gregory Beylkin and Nicholas Coult. 1998. A multiresolution strategy for reduction of elliptic PDEs and eigenvalue problems. *Applied and Computational Harmonic Analysis* 5, 2 (1998), 129–155.

[19] Harsh Bhatia, Timothy S Carpenter, Helgi I Ingólfsson, Gautham Dharuman, Piyush Karande, Shusen Liu, Tomas Oppelstrup, Chris Neale, Felice C Lightstone, Brian Van Essen, et al. 2021. Machine-learning-based dynamic-importance sampling for adaptive multiscale simulations. *Nature Machine Intelligence* 3, 5 (2021), 401–409.

[20] Fabian Böhm, Diego Alonso-Urquijo, Guy Verschaffelt, and Guy Van der Sande. 2022. Noise-injected analog Ising machines enable ultrafast statistical sampling and machine learning. *Nature Communications* 13, 1 (2022), 1–13.

[21] Jason J Bramburger, Daniel Dylewsky, and J Nathan Kutz. 2020. Sparse identification of slow timescale dynamics. *Physical Review E* 102, 2 (2020), 022204.

[22] Emiliano Brini, Carlos Simmerling, and Ken Dill. 2020. Protein storytelling through physics. *Science* 370, 6520 (2020), eaaz3041.

[23] Dirk Brockmann and Dirk Helbing. 2013. The hidden geometry of complex, network-driven contagion phenomena. *science* 342, 6164 (2013), 1337–1342.

[24] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. 2009. CHARMM: the biomolecular simulation program. *Journal of computational chemistry* 30, 10 (2009), 1545–1614.

[25] Steven L Brunton, Marko Budišić, Eurika Kaiser, and J Nathan Kutz. 2021. Modern Koopman theory for dynamical systems. *arXiv preprint arXiv:2102.12086* (2021).

[26] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences* 113, 15 (2016), 3932–3937.

[27] Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. 2010. Catastrophic cascade of failures in interdependent networks. *Nature* 464, 7291 (2010), 1025–1028.

[28] Richard Car and Mark Parrinello. 1985. Unified approach for molecular dynamics and density-functional theory. *Physical review letters* 55, 22 (1985), 2471.

[29] Subhendu Chakraborty and Somnath Ghosh. 2018. Hyperdynamics accelerated concurrent atomistic-continuum model for developing crack propagation models in elastic crystalline materials. *Computational Materials Science* 154 (2018), 212–224.

[30] Partha Chakroborty and Akhilesh Kumar Maurya. 2008. Microscopic analysis of cellular automata based traffic flow models and an improved model. *Transport reviews* 28, 6 (2008), 717–734.

[31] Kathleen Champion, Bethany Lusch, J Nathan Kutz, and Steven L Brunton. 2019. Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences* 116, 45 (2019), 22445–22451.

[32] Kathleen P Champion, Steven L Brunton, and J Nathan Kutz. 2019. Discovery of nonlinear multiscale systems: Sampling strategies and embeddings. *SIAM Journal on Applied Dynamical Systems* 18, 1 (2019), 312–333.

[33] Serina Chang, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2021. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 589, 7840 (2021), 82–87.

[34] Boyuan Chen, Kuang Huang, Sunand Raghupathi, Ishaan Chandratreya, Qiang Du, and Hod Lipson. 2022. Automated discovery of fundamental variables hidden in experimental data. *Nature Computational Science* 2, 7 (2022), 433–442.

[35] Lin Chen, Fengli Xu, Zhenyu Han, Kun Tang, Pan Hui, James Evans, and Yong Li. 2022. Strategic COVID-19 vaccine distribution can simultaneously elevate social utility and equity. *Nature Human Behaviour* (2022), 1–12.

[36] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. 2018. Neural ordinary differential equations. *Advances in neural information processing systems* 31 (2018).

[37] Shaopei Chen, Jianjun Tan, Christophe Claramunt, and Cyril Ray. 2011. Multi-scale and multi-modal GIS-T data model. *Journal of Transport Geography* 19, 1 (2011), 147–161.

[38] Wen Chen and Jacob Fish. 2006. A generalized space–time mathematical homogenization theory for bridging atomistic and continuum scales. *Internat. J. Numer. Methods Engrg.* 67, 2 (2006), 253–271.

[39] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368, 6489 (2020), 395–400.

[40] Debashish Chowdhury, Ludger Santen, and Andreas Schadschneider. 2000. Statistical physics of vehicular traffic and some related systems. *Physics Reports* 329, 4-6 (2000), 199–329.

[41] F Fleming Crim. 2008. Chemical reaction dynamics. *Proceedings of the National Academy of Sciences* 105, 35 (2008), 12647–12648.

[42] Ruxin Dai, Pengpeng Lin, and Jun Zhang. 2019. An EXCMG accelerated multiscale multigrid computation for 3D Poisson equation. *Computers & Mathematics with Applications* 77, 8 (2019), 2051–2060.

[43] Bryan C Daniels and Ilya Nemenman. 2015. Automated adaptive inference of phenomenological dynamical models. *Nature communications* 6, 1 (2015), 8133.

[44] Francesco Di Natale, Harsh Bhatia, Timothy S Carpenter, Chris Neale, Sara Kokkila-Schumacher, Tomas Oppelstrup, Liam Stanton, Xiaohua Zhang, Shiv Sundram, Thomas RW Scogland, et al. 2019. A massively parallel infrastructure for adaptive multiscale simulations: modeling RAS initiation pathway for cancer. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 1–16.

[45] Domenico Di Sante, Matija Medvidović, Alessandro Toschi, Giorgio Sangiovanni, Cesare Franchini, Anirvan M Sengupta, and Andrew J Millis. 2022. Deep Learning the Functional Renormalization Group. *arXiv preprint arXiv:2202.13268* (2022).

[46] Paul Dupuis, Konstantinos Spiliopoulos, and Hui Wang. 2012. Importance sampling for multiscale diffusions. *Multiscale Modeling & Simulation* 10, 1 (2012), 1–27.

[47] Somayeh Faraji, S Alireza Ghasemi, Samare Rostami, Robabe Rasoulkhani, Bastian Schaefer, Stefan Goedecker, and Maximilian Amsler. 2017. High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride. *Physical Review B* 95, 10 (2017), 104105.

[48] Jacob Fish, Gregory J Wagner, and Sinan Keten. 2021. Mesoscopic and multiscale modelling in materials. *Nature materials* 20, 6 (2021), 774–786.

[49] Jacob Fish and Wei Wu. 2011. A nonintrusive stochastic multiscale solver. *Internat. J. Numer. Methods Engrg.* 88, 9 (2011), 862–879.

[50] Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Thomas A Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, et al. 2020. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 584, 7820 (2020), 257–261.

[51] Meire Fortunato, Tobias Pfaff, Peter Wirnsberger, Alexander Pritzel, and Peter Battaglia. 2022. Multiscale meshgraphnets. *arXiv preprint arXiv:2210.00612* (2022).

[52] Kai Fukami, Taichi Nakamura, and Koji Fukagata. 2020. Convolutional neural network based hierarchical autoencoder for nonlinear mode decomposition of fluid field data. *Physics of Fluids* 32, 9 (2020), 095110.

[53] Jianxi Gao, Baruch Barzel, and Albert-László Barabási. 2016. Universal resilience patterns in complex networks. *Nature* 530, 7590 (2016), 307–312.

[54] Kun Gao, Rui Jiang, Shou-Xin Hu, Bing-Hong Wang, and Qing-Song Wu. 2007. Cellular-automaton model with velocity adaptation in the framework of Kerner's three-phase traffic theory. *Physical Review E* 76, 2 (2007), 026105.

[55] HX Ge, HB Zhu, and SQ Dai. 2006. Effect of looking backward on traffic flow in a cooperative driving car following model. *The European Physical Journal B-Condensed Matter and Complex Systems* 54, 4 (2006), 503–507.

[56] Wei Ge and Jinghai Li. 2002. Physical mapping of fluidization regimes—the EMMS approach. *Chemical Engineering Science* 57, 18 (2002), 3993–4004.

[57] F Ghavamian and A Simone. 2019. Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. *Computer Methods in Applied Mechanics and Engineering* 357 (2019), 112594.

[58] AC14927910896 Gilbert. 1998. A comparison of multiresolution and classical one-dimensional homogenization schemes. *Applied and computational harmonic analysis* 5, 1 (1998), 1–35.

[59] Cláudio Gomes, Casper Thule, David Broman, Peter Gorm Larsen, and Hans Vangheluwe. 2018. Co-simulation: a survey. *ACM Computing Surveys (CSUR)* 51, 3 (2018), 1–33.

[60] Xiping Gong, Yumeng Zhang, and Jianhan Chen. 2021. Advanced Sampling Methods for Multiscale Simulation of Disordered Proteins and Dynamic Interactions. *Biomolecules* 11, 10 (2021), 1416.

[61] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (2020), 139–144.

[62] Volker Gravemeier, Michael W Gee, Martin Kronbichler, and Wolfgang A Wall. 2010. An algebraic variational multiscale–multigrid method for large eddy simulation of turbulent flow. *Computer Methods in Applied Mechanics and Engineering* 199, 13-16 (2010), 853–864.

[63] Ross Griebenow, Brennan Klein, and Erik Hoel. 2019. Finding the right scale of a network: efficient identification of causal emergence through spectral clustering. *arXiv preprint arXiv:1908.07565* (2019).

[64] Dabo Guan, Daoping Wang, Stephane Hallegatte, Steven J Davis, Jingwen Huo, Shuping Li, Yangchun Bai, Tianyang Lei, Qianyu Xue, D'Maris Coffman, et al. 2020. Global supply-chain effects of COVID-19 control measures. *Nature human behaviour* 4, 6 (2020), 577–587.

[65] Kan Guo, Yongli Hu, Yanfeng Sun, Sean Qian, Junbin Gao, and Baocai Yin. 2021. Hierarchical Graph Convolution Network for Traffic Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 151–159.

[66] Lin Guo, Shiyi Chen, and Mark O Robbins. 2016. Multi-scale simulation method for electroosmotic flows. *The European Physical Journal Special Topics* 225, 8 (2016), 1551–1582.

[67] Liangshuai Guo, Maokun Li, Shenheng Xu, Fan Yang, and Li Liu. 2021. Electromagnetic Modeling Using an FDTD-Equivalent Recurrent Convolution Neural Network: Accurate Computing on a Deep Learning Framework. *IEEE Antennas and Propagation Magazine* (2021).

[68] Milad Habibi, Scott TM Dawson, and Amirhossein Arzani. 2020. Data-driven pulsatile blood flow physics with dynamic mode decomposition. *Fluids* 5, 3 (2020), 111.

[69] Eric J Hall, Søren Taverniers, Markos A Katsoulakis, and Daniel M Tartakovsky. 2021. Ginns: Graph-informed neural networks for multiscale physics. *J. Comput. Phys.* 433 (2021), 110192.

[70] Ridha Hambli, Houda Katerchi, and Claude-Laurent Benhamou. 2011. Multiscale methodology for bone remodelling simulation using coupled finite element and neural network computation. *Biomechanics and modeling in mechanobiology* 10, 1 (2011), 133–145.

[71] Changnian Han, Peng Zhang, Danny Bluestein, Guojing Cong, and Yuefan Deng. 2021. Artificial intelligence for accelerating time integrations in multiscale modeling. *J. Comput. Phys.* 427 (2021), 110053.

[72] Jiequn Han, Linfeng Zhang, Roberto Car, et al. 2017. Deep potential: A general representation of a many-body potential energy surface. *arXiv preprint arXiv:1707.01478* (2017).

[73] David J Haw, Giovanni Forchini, Patrick Doohan, Paula Christen, Matteo Pianella, Robert Johnson, Sumali Bajaj, Alexandra B Hogan, Peter Winskill, Marisa Miraldo, et al. 2022. Optimizing social and economic activity while containing SARS-CoV-2 transmission using DAEDALUS. *Nature Computational Science* 2, 4 (2022), 223–233.

[74] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[75] Dirk Helbing and Michael Schreckenberg. 1999. Cellular automata simulating experimental properties of traffic flow. *Physical review E* 59, 3 (1999), R2505.

[76] Dirk Helbing and Benno Tilch. 1998. Generalized force model of traffic dynamics. *Physical review E* 58, 1 (1998), 133.

[77] Maziar S Hemati, Matthew O Williams, and Clarence W Rowley. 2014. Dynamic mode decomposition for large and streaming datasets. *Physics of Fluids* 26, 11 (2014), 111701.

[78] Carlos X Hernández, Hannah K Wayment-Steele, Mohammad M Sultan, Brooke E Husic, and Vijay S Pande. 2018. Variational encoding of complex dynamics. *Physical Review E* 97, 6 (2018), 062412.

[79] Rodney Hill. 1963. Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids* 11, 5 (1963), 357–372.

[80] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[81] Erik P Hoel, Larissa Albantakis, and Giulio Tononi. 2013. Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences (PNAS)* 110, 49 (2013), 19790–19795.

[82] John H Holland. 1992. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.*

[83] Jiliang Hu, Daniel R Amor, Matthieu Barbier, Guy Bunin, and Jeff Gore. 2022. Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* 378, 6615 (2022), 85–89.

[84] Nan Hu, Jacob Fish, and Colin McAuliffe. 2017. An adaptive stochastic inverse solver for multiscale characterization of composite materials. *Internat. J. Numer. Methods Engrg.* 109, 12 (2017), 1679–1700.

[85] Pipi Hu, Wuyue Yang, Yi Zhu, and Liu Hong. 2022. Revealing hidden dynamics from time-series data by ODENet. *J. Comput. Phys.* 461 (2022), 111203.

[86] Wenlai Huang, Jinghai Li, and Peter P Edwards. 2018. Mesoscience: exploring the common principle at mesoscales. *National science review* (2018).

[87] Thomas JR Hughes. 2012. *The finite element method: linear static and dynamic finite element analysis.* Courier Corporation.

[88] Iacopo Iacopini, Giovanni Petri, Alain Barrat, and Vito Latora. 2019. Simplicial models of social contagion. *Nature communications* 10, 1 (2019), 1–9.

[89] Stefano Maria Iacus, Carlos Santamaria, Francesco Sermi, Spyros Spyratos, Dario Tarchi, and Michele Vespe. 2020. Human mobility and COVID-19 initial dynamics. *Nonlinear Dynamics* 101, 3 (2020), 1901–1919.

[90] Helgi I Ingólfsson, Chris Neale, Timothy S Carpenter, Rebika Shrestha, Cesar A López, Timothy H Tran, Tomas Oppelstrup, Harsh Bhatia, Liam G Stanton, Xiaohua Zhang, et al. 2022. Machine learning–driven multiscale modeling reveals lipid-dependent dynamics of RAS signaling proteins. *Proceedings of the National Academy of Sciences* 119, 1

(2022), e2113297119.

[91] Sean Jewell, Joseph Futoma, Lauren Hannah, Andrew C Miller, Nicholas J Foti, and Emily B Fox. 2021. It's complicated: Characterizing the time-varying relationship between cell phone mobility and COVID-19 spread in the US. *NPJ digital medicine* 4, 1 (2021), 1–11.

[92] Jayson S Jia, Xin Lu, Yun Yuan, Ge Xu, Jianmin Jia, and Nicholas A Christakis. 2020. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 582, 7812 (2020), 389–394.

[93] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvu-nakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 7873 (2021), 583–589.

[94] Pushkin Kachroo, Shaurya Agarwal, Benedetto Piccoli, and Kaan Özbay. 2017. Multiscale modeling and control architecture for V2X enabled traffic streams. *IEEE Transactions on Vehicular Technology* 66, 6 (2017), 4616–4626.

[95] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. 2021. Physics-informed machine learning. *Nature Reviews Physics* 3, 6 (2021), 422–440.

[96] Eric Karsenti. 2008. Self-organization in cell biology: a brief history. *Nature reviews Molecular cell biology* 9, 3 (2008), 255–262.

[97] Boris S Kerner, Sergey L Klenov, and Dietrich E Wolf. 2002. Cellular automata approach to three-phase traffic theory. *Journal of Physics A: Mathematical and General* 35, 47 (2002), 9971.

[98] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

[99] Stefan Klus, Feliks Nüske, Péter Koltai, Hao Wu, Ioannis Kevrekidis, Christof Schütte, and Frank Noé. 2018. Data-driven model reduction and transfer operator approximation. *Journal of Nonlinear Science* 28, 3 (2018), 985–1010.

[100] Maciej Koch-Janusz and Zohar Ringel. 2018. Mutual information, neural networks and the renormalization group. *Nature Physics* 14, 6 (2018), 578–582.

[101] Dmitrii Kochkov, Jamie A Smith, Ayya Alieva, Qing Wang, Michael P Brenner, and Stephan Hoyer. 2021. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences* 118, 21 (2021), e2101784118.

[102] Moritz UG Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Open COVID-19 Data Working Group†, Louis Du Plessis, Nuno R Faria, Ruoran Li, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368, 6490 (2020), 493–497.

[103] J Nathan Kutz, Steven L Brunton, Bingni W Brunton, and Joshua L Proctor. 2016. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM.

[104] Shengjie Lai, Nick W Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R Floyd, Amy Wesolowski, Mauricio Santillana, Chi Zhang, Xiangjun Du, et al. 2020. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *nature* 585, 7825 (2020), 410–413.

[105] BA Le, Julien Yvonnet, and Q-C He. 2015. Computational homogenization of nonlinear elastic materials using neural networks. *Internat. J. Numer. Methods Engrg.* 104, 12 (2015), 1061–1084.

[106] Kookjin Lee and Kevin T Carlberg. 2020. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *J. Comput. Phys.* 404 (2020), 108973.

[107] Seungjoon Lee, Mahdi Kooshkbaghi, Konstantinos Spiliotis, Constantinos I Siettos, and Ioannis G Kevrekidis. 2020. Coarse-scale PDEs from fine-scale observations via machine learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30, 1 (2020), 013141.

[108] Wing Tat Leung, Guang Lin, and Zecheng Zhang. 2022. NH-PINN: Neural homogenization-based physics-informed neural network for multiscale problems. *J. Comput. Phys.* 470 (2022), 111539.

[109] Michael Levitt and Arieh Warshel. 1975. Computer simulation of protein folding. *Nature* 253, 5494 (1975), 694–698.

[110] Aiqin Li, Renge Li, and Jacob Fish. 2008. Generalized mathematical homogenization: from theory to practice. *Computer methods in applied mechanics and engineering* 197, 41-42 (2008), 3225–3248.

[111] Jinghai Li, Wenlai Huang, Jianhua Chen, Wei Ge, and Chaofeng Hou. 2018. Mesoscience based on the EMMS principle of compromise in competition. *Chemical Engineering Journal* 333 (2018), 327–335.

[112] J Li and M Kwauk. 1994. Particle-Fluid Two-Phase Flow: the Energy-Minimization Multi-Scale Method; Minimization Multi-Scale Method; Metallurg.

[113] Ruikun Li, Huandong Wang, and Yong Li. 2023. Learning slow and fast system dynamics via automatic separation of time scales. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 4380–4390.

[114] Shuo-Hui Li and Lei Wang. 2018. Neural network renormalization group. *Physical review letters* 121, 26 (2018), 260601.

[115] Liang Liang, Minliang Liu, Caitlin Martin, and Wei Sun. 2018. A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. *Journal of The Royal Society Interface* 15, 138 (2018), 20170844.

[116] Michael James Lighthill and Gerald Beresford Whitham. 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229, 1178

(1955), 317–345.

[117] Mario Lino, Chris Cantwell, Anil A Bharath, and Stathi Fotiadis. 2021. Simulating continuum mechanics with multi-scale graph neural networks. *arXiv preprint arXiv:2106.04900* (2021).

[118] Mario Lino, Stati Fotiadis, Anil A Bharath, and Chris Cantwell. 2022. REMuS-GNN: A Rotation-Equivariant Model for Simulating Continuum Dynamics. *arXiv preprint arXiv:2205.07852* (2022).

[119] Mario Lino, Stathi Fotiadis, Anil A Bharath, and Chris Cantwell. 2022. Towards Fast Simulation of Environmental Fluid Mechanics with Multi-Scale Graph Neural Networks. *arXiv preprint arXiv:2205.02637* (2022).

[120] Dachuan Liu, Jin Wang, Shuo Shang, and Peng Han. 2022. Msdr: Multi-step dependency relation networks for spatial temporal forecasting. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1042–1050.

[121] Zeliang Liu, MA Bessa, and Wing Kam Liu. 2016. Self-consistent clustering analysis: an efficient multi-scale scheme for inelastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering* 306 (2016), 319–341.

[122] Zeliang Liu and CT Wu. 2019. Exploring the 3D architectures of deep material network in data-driven multiscale mechanics. *Journal of the Mechanics and Physics of Solids* 127 (2019), 20–46.

[123] Zeliang Liu, CT Wu, and M391298807188754 Koishi. 2019. A deep material network for multiscale topology learning and accelerated nonlinear modeling of heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering* 345 (2019), 1138–1168.

[124] J O Lloyd-Smith, S J Schreiber, P E Kopp, and W M Getz. 2005. Superspreading and the effect of individual variation on disease emergence. *Nature* 438, 7066 (Nov. 2005), 355–359.

[125] Keverne A Louison, Ian L Dryden, and Charles A Laughton. 2021. GLIMPS: A Machine Learning Approach to Resolution Transformation for Multiscale Modeling. *Journal of Chemical Theory and Computation* 17, 12 (2021), 7930–7937.

[126] Benjamin F Maier and Dirk Brockmann. 2020. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* 368, 6492 (2020), 742–746.

[127] Quentin Malé, Gabriel Staffelbach, Olivier Vermorel, Antony Misdariis, Frédéric Ravet, and Thierry Poinsot. 2019. Large eddy simulation of pre-chamber ignition in an internal combustion engine. *Flow, Turbulence and Combustion* 103, 2 (2019), 465–483.

[128] Stephane G Mallat. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence* 11, 7 (1989), 674–693.

[129] Krithika Manohar, Eurika Kaiser, Steven L Brunton, and J Nathan Kutz. 2019. Optimized sampling for multiscale dynamics. *Multiscale Modeling & Simulation* 17, 1 (2019), 117–133.

[130] Admir Masic, Luca Bertinetti, Roman Schuetz, Leonardo Galvis, Nadya Timofeeva, John WC Dunlop, Jong Seto, Markus A Hartmann, and Peter Fratzl. 2011. Observations of multiscale, stress-induced changes of collagen orientation in tendon by polarized Raman spectroscopy. *Biomacromolecules* 12, 11 (2011), 3989–3996.

[131] Romit Maulik, Bethany Lusch, and Prasanna Balaprakash. 2021. Reduced-order modeling of advection-dominated systems with recurrent neural networks and convolutional autoencoders. *Physics of Fluids* 33, 3 (2021), 037106.

[132] Romit Maulik, Arvind Mohan, Bethany Lusch, Sandeep Madireddy, Prasanna Balaprakash, and Daniel Livescu. 2020. Time-series learning of latent-space dynamics for reduced-order model closure. *Physica D: Nonlinear Phenomena* 405 (2020), 132368.

[133] David McAdams, Kaci Kennedy McDade, Osondu Ogbuoji, Matthew Johnson, Siddharth Dixit, and Gavin Yamey. 2020. Incentivising wealthy nations to participate in the COVID-19 Vaccine Global Access Facility (COVAX): a game theory perspective. *BMJ global health* 5, 11 (2020), e003627.

[134] Don L McLeish and Adam W Kolkiewicz. 1997. Fitting diffusion models in finance. *Lecture Notes-Monograph Series* (1997), 327–350.

[135] Shafigh Mehraeen and Jiun-Shyan Chen. 2006. Wavelet Galerkin method in multi-scale homogenization of heterogeneous media. *International journal for numerical methods in engineering* 66, 3 (2006), 381–403.

[136] Pankaj Mehta and David J Schwab. 2014. An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831* (2014).

[137] Christian Miehe and CG2344771 Bayreuther. 2007. On multiscale FE analyses of heterogeneous structures: from homogenization to multigrid solvers. *Internat. J. Numer. Methods Engrg.* 71, 10 (2007), 1135–1180.

[138] Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernockỳ, and Sanjeev Khudanpur. 2010. Recurrent neural network based language model.. In *Interspeech*, Vol. 2. Makuhari, 1045–1048.

[139] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602* (2013).

[140] Yi Mo, Mengjie Du, Wei Ge, and Pingwen Zhang. 2020. Analysis of the energy-minimization multiscale model with multiobjective optimization. *Particuology* 48 (2020), 109–115.

[141] Lukas Mosser, Olivier Dubrule, and Martin J Blunt. 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Physical Review E* 96, 4 (2017), 043309.

[142] Charles Murphy, Edward Laurence, and Antoine Allard. 2021. Deep learning of contagion dynamics on complex networks. *Nature Communications* 12, 1 (2021), 1–11.

[143] Akihiro Nakayama, Yūki Sugiyama, and Katsuya Hasebe. 2001. Effect of looking at the car that follows in an optimal velocity model of traffic flow. *Physical Review E* 65, 1 (2001), 016112.

[144] M Nauenberg and B Nienhuis. 1974. Renormalization-group approach to the solution of general Ising models. *Physical Review Letters* 33, 27 (1974), 1598.

[145] Gordon Frank Newell. 1961. Nonlinear effects in the dynamics of car following. *Operations research* 9, 2 (1961), 209–229.

[146] Akihiro Nishi, George Dewey, Akira Endo, Sophia Neman, Sage K Iwamoto, Michael Y Ni, Yusuke Tsugawa, Georgios Iosifidis, Justin D Smith, and Sean D Young. 2020. Network interventions for managing the COVID-19 pandemic and sustaining economy. *Proceedings of the National Academy of Sciences* 117, 48 (2020), 30285–30294.

[147] Frank Noé and Cecilia Clementi. 2017. Collective variables for the study of long-time kinetics from molecular trajectories: theory and methods. *Current opinion in structural biology* 43 (2017), 141–147.

[148] Frank Noé and Feliks Nuske. 2013. A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Modeling & Simulation* 11, 2 (2013), 635–655.

[149] Atsuya Oishi and Genki Yagawa. 2017. Computational mechanics enhanced by deep learning. *Computer Methods in Applied Mechanics and Engineering* 327 (2017), 327–351.

[150] I Özdemir, WAM Brekelmans, and MGD2376223 Geers. 2008. Computational homogenization for heat conduction in heterogeneous solids. *International journal for numerical methods in engineering* 73, 2 (2008), 185–204.

[151] Mathias Peirlinck, F Sahli Costabal, KL Sack, JS Choy, GS Kassab, JM Guccione, M De Beule, Patrick Segers, and E Kuhl. 2019. Using machine learning to characterize heart failure across the scales. *Biomechanics and modeling in mechanobiology* 18, 6 (2019), 1987–2001.

[152] GH Peng and DH Sun. 2010. A dynamical model of car-following with the consideration of the multiple information of preceding cars. *Physics Letters A* 374, 15-16 (2010), 1694–1698.

[153] Grace CY Peng, Mark Alber, Adrian Buganza Tepole, William R Cannon, Suvranu De, Savador Dura-Bernal, Krishna Garikipati, George Karniadakis, William W Lytton, Paris Perdikaris, et al. 2021. Multiscale modeling meets machine learning: What can we learn? *Archives of Computational Methods in Engineering* 28, 3 (2021), 1017–1037.

[154] Guillermo Pérez-Hernández, Fabian Paul, Toni Giorgino, Gianni De Fabritiis, and Frank Noé. 2013. Identification of slow molecular order parameters for Markov model construction. *The Journal of chemical physics* 139, 1 (2013), 07B604_1.

[155] Louis A Pipes. 1953. An operational analysis of traffic dynamics. *Journal of applied physics* 24, 3 (1953), 274–281.

[156] Joshua L Proctor, Steven L Brunton, and J Nathan Kutz. 2016. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems* 15, 1 (2016), 142–161.

[157] Xinwu Qian, Lijun Sun, and Satish V Ukkusuri. 2021. Scaling of contact networks for epidemic spreading in urban transit systems. *Scientific reports* 11, 1 (2021), 1–12.

[158] IBCM Rocha, Pierre Kerfriden, and FP van der Meer. 2021. On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning. *Journal of Computational Physics: X* 9 (2021), 100083.

[159] Ronald Ross. 1916. An application of the theory of probabilities to the study of a priori pathometry.—Part I. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* 92, 638 (1916), 204–230.

[160] Otto E Rössler. 1976. Chemical turbulence: chaos in a simple reaction-diffusion system. *Zeitschrift für Naturforschung A* 31, 10 (1976), 1168–1172.

[161] Tiziano Rotesi, Paolo Pin, Maria Cucciniello, Amyn A Malik, Elliott E Paintsil, Scott E Bokemper, Kathryn Willebrand, Gregory A Huber, Alessia Melegaro, and Saad B Omer. 2021. National interest may require distributing COVID-19 vaccines to other countries. *Scientific reports* 11, 1 (2021), 1–8.

[162] Kazumi Saito, Ryohei Nakano, and Masahiro Kimura. 2008. Prediction of information diffusion probabilities for independent cascade model. In *International conference on knowledge-based and intelligent information and engineering systems*. 67–75.

[163] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. 2020. Learning to simulate complex physics with graph networks. In *International Conference on Machine Learning*. 8459–8468.

[164] Peter J Schmid. 2010. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* 656 (2010), 5–28.

[165] Peter J Schmid, Larry Li, Matthew P Juniper, and O Pust. 2011. Applications of the dynamic mode decomposition. *Theoretical and Computational Fluid Dynamics* 25, 1 (2011), 249–259.

[166] Amit Singer, Radek Erban, Ioannis G Kevrekidis, and Ronald R Coifman. 2009. Detecting intrinsic slow variables in stochastic dynamical systems by anisotropic diffusion maps. *Proceedings of the National Academy of Sciences (PNAS)* 106, 38 (2009), 16090–16095.

[167] Berber T Snoeijer, Mariska Burger, Shaoxiong Sun, Richard JB Dobson, and Amos A Folarin. 2021. Measuring the effect of Non-Pharmaceutical Interventions (NPIs) on mobility during the COVID-19 pandemic using global mobility data. *NPJ digital medicine* 4, 1 (2021), 1–12.

[168] Alessandro Spelta, Andrea Flori, Francesco Pierri, Giovanni Bonaccorsi, and Fabio Pammolli. 2020. After the lockdown: simulating mobility, public health and economic recovery scenarios. *Scientific reports* 10, 1 (2020), 1–13.

[169] Ashwin Sridhar, Varvara G Kouznetsova, and Marc GD Geers. 2016. Homogenization of locally resonant acoustic metamaterials towards an emergent enriched continuum. *Computational mechanics* 57, 3 (2016), 423–435.

[170] Guillaume St-Onge, Hanlin Sun, Antoine Allard, Laurent Hébert-Dufresne, and Ginestra Bianconi. 2021. Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Physical Review Letters* 127, 15 (2021), 158301.

[171] Phillip J Stansfeld and Mark SP Sansom. 2011. From coarse grained to atomistic: a serial multiscale approach to membrane protein simulations. *Journal of chemical theory and computation* 7, 4 (2011), 1157–1166.

[172] Chen Tao, Jia Bin, Li Xin-Gang, Jiang Rui, and Gao Zi-You. 2008. Synchronized flow in a cellular automaton model with time headway dependent randomization. *Chinese Physics Letters* 25, 8 (2008), 2795.

[173] Molei Tao, Houman Owhadi, and Jerrold E Marsden. 2010. Nonintrusive and structure preserving multiscale integration of stiff ODEs, SDEs, and Hamiltonian systems with hidden slow dynamics via flow averaging. *Multiscale Modeling & Simulation* 8, 4 (2010), 1269–1324.

[174] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical review E* 62, 2 (2000), 1805.

[175] Jonathan H Tu. 2013. *Dynamic mode decomposition: Theory and applications.* Ph. D. Dissertation. Princeton University.

[176] Pantelis R Vlachas, Georgios Arampatzis, Caroline Uhler, and Petros Koumoutsakos. 2022. Multiscale simulations of complex systems by learning their effective dynamics. *Nature Machine Intelligence* 4, 4 (2022), 359–366.

[177] Pantelis R Vlachas, Julija Zavadlav, Matej Praprotnik, and Petros Koumoutsakos. 2021. Accelerated Simulations of Molecular Systems through Learning of Effective Dynamics. *Journal of Chemical Theory and Computation* 18, 1 (2021), 538–549.

[178] Caroline E Wagner, Chadi M Saad-Roy, Sinead E Morris, Rachel E Baker, Michael J Mina, Jeremy Farrar, Edward C Holmes, Oliver G Pybus, Andrea L Graham, Ezekiel J Emanuel, et al. 2021. Vaccine nationalism and the dynamics and control of SARS-CoV-2. *Science* 373, 6562 (2021), 1–8.

[179] Kun Wang and WaiChing Sun. 2019. Meta-modeling game for deriving theory-consistent, microstructure-based traction–separation laws via deep reinforcement learning. *Computer Methods in Applied Mechanics and Engineering* 346 (2019), 216–241.

[180] Mudan Wang, Huan Yan, Huandong Wang, Yong Li, and Depeng Jin. 2023. Contagion Process Guided Cross-scale Spatio-Temporal Graph Neural Network for Traffic Congestion Prediction. In *Proceedings of the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.*

[181] Wujie Wang and Rafael Gómez-Bombarelli. 2019. Coarse-graining auto-encoders for molecular dynamics. *npj Computational Materials* 5, 1 (2019), 1–9.

[182] Yating Wang, Siu Wun Cheung, Eric T Chung, Yalchin Efendiev, and Min Wang. 2020. Deep multiscale model learning. *J. Comput. Phys.* 406 (2020), 109071.

[183] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of 'small-world' networks. *nature* 393, 6684 (1998), 440–442.

[184] Christoph Wehmeyer and Frank Noé. 2018. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *The Journal of chemical physics* 148, 24 (2018), 241703.

[185] E Weinan. 2011. *Principles of multiscale modeling.* Cambridge University Press.

[186] Geoffrey B West, James H Brown, and Brian J Enquist. 1999. The fourth dimension of life: fractal geometry and allometric scaling of organisms. *science* 284, 5420 (1999), 1677–1679.

[187] Chi Wu, Ali Entezari, Keke Zheng, Jianguang Fang, Hala Zreiqat, Grant P Steven, Michael V Swain, and Qing Li. 2021. A machine learning-based multiscale model to predict bone formation in scaffolds. *Nature Computational Science* 1, 8 (2021), 532–541.

[188] Joseph T Wu, Kathy Leung, and Gabriel M Leung. 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* 395, 10225 (2020), 689–697.

[189] Ling Wu, Nanda Gopala Kilingar, Ludovic Noels, et al. 2020. A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths. *Computer Methods in Applied Mechanics and Engineering* 369 (2020), 113234.

[190] Zhao Xiaomei and Gao Ziyou. 2007. The stability analysis of the full velocity and acceleration velocity model. *Physica A: Statistical Mechanics and its Applications* 375, 2 (2007), 679–686.

[191] DF Xie, ZY Gao, and XM Zhao. 2008. Stabilization of traffic flow based on the multiple information of preceding cars. *Communications in Computational Physics* 3, 4 (2008), 899–912.

[192] Tian Xie and Jeffrey C Grossman. 2018. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters* 120, 14 (2018), 145301.

[193] Yang Xing, Chen Lv, Dongpu Cao, and Efstathios Velenis. 2021. Multi-scale driver behavior modeling based on deep spatial-temporal representation for intelligent vehicles. *Transportation research part C: emerging technologies* 130 (2021), 103288.

[194] Rui Xu and Donald Wunsch. 2005. Survey of clustering algorithms. *IEEE Transactions on neural networks* 16, 3 (2005), 645–678.

[195] Kun Yang, Yi-Fan Chen, Georgios Roumpos, Chris Colby, and John Anderson. 2019. High performance Monte Carlo simulation of Ising model on TPU clusters. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 1–15.

[196] Kaidi Yang, Nan Zheng, and Monica Menendez. 2017. Multi-scale perimeter control approach in a connected-vehicle environment. *Transportation research procedia* 23 (2017), 101–120.

[197] Wuyue Yang, Liangrong Peng, Yi Zhu, and Liu Hong. 2020. When machine learning meets multiscale modeling in chemical reactions. *The Journal of Chemical Physics* 153, 9 (2020), 094117.

[198] Zijiang Yang, Xiaolin Li, L Catherine Brinson, Alok N Choudhary, Wei Chen, and Ankit Agrawal. 2018. Microstructural materials design via deep adversarial learning methodology. *Journal of Mechanical Design* 140, 11 (2018).

[199] Yang Ye, Qingpeng Zhang, Xuan Wei, Zhidong Cao, Hsiang-Yu Yuan, and Daniel Dajun Zeng. 2022. Equitable access to COVID-19 vaccines makes a life-saving difference to all countries. *Nature Human Behaviour* 6, 2 (2022), 207–216.

[200] Zhitao Ying, Jiaxuan You, Christopher Morris, Xiang Ren, Will Hamilton, and Jure Leskovec. 2018. Hierarchical graph representation learning with differentiable pooling. *Advances in neural information processing systems* 31 (2018).

[201] Yong Li Yinzhou Tang, Huandong Wang. 2023. Enhancing Spatial Spread Prediction of Infectious Diseases through Integrating Multi-scale Human Mobility Dynamics. In *Proceedings of the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*.

[202] Qing Yu and Jacob Fish. 2002. Multiscale asymptotic homogenization for multiphysics problems with multiple spatial and temporal scales: a coupled thermo-viscoelastic example problem. *International journal of solids and structures* 39, 26 (2002), 6429–6452.

[203] Zheng Yuan and Jacob Fish. 2009. Multiple scale eigendeformation-based reduced order homogenization. *Computer Methods in Applied Mechanics and Engineering* 198, 21-26 (2009), 2016–2038.

[204] Dewen Yushu and Karel Matouš. 2020. The image-based multiscale multigrid solver, preconditioner, and reduced order model. *J. Comput. Phys.* 406 (2020), 109165.

[205] Juanjuan Zhang, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, Stefano Merler, Cécile Viboud, Alessandro Vespignani, et al. 2020. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 368, 6498 (2020), 1481–1486.

[206] Qian Zhang, Kaiyuan Sun, Matteo Chinazzi, Ana Pastore y Piontti, Natalie E Dean, Diana Patricia Rojas, Stefano Merler, Dina Mistry, Piero Poletti, Luca Rossi, et al. 2017. Spread of Zika virus in the Americas. *Proceedings of the national academy of sciences* 114, 22 (2017), E4334–E4343.

[207] Xiao-Mei Zhao and Zi-You Gao. 2005. A new car-following model: full velocity and acceleration difference model. *The European Physical Journal B-Condensed Matter and Complex Systems* 47, 1 (2005), 145–150.

[208] Peijie Zhou, Shuxiong Wang, Tiejun Li, and Qing Nie. 2021. Dissecting transition cells from single-cell transcriptome data through multiscale stochastic dynamics. *Nature communications* 12, 1 (2021), 1–15.

[209] Ying Zhou, Renzhe Xu, Dongsheng Hu, Yang Yue, Qingquan Li, and Jizhe Xia. 2020. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. *The Lancet Digital Health* 2, 8 (2020), e417–e424.

[210] Zhibo Zhu, Ziqi Liu, Ge Jin, Zhiqiang Zhang, Lei Chen, Jun Zhou, and Jianyong Zhou. 2021. MixSeq: Connecting Macroscopic Time Series Forecasting with Microscopic Time Series Data. *Advances in Neural Information Processing Systems* 34 (2021), 12904–12916.