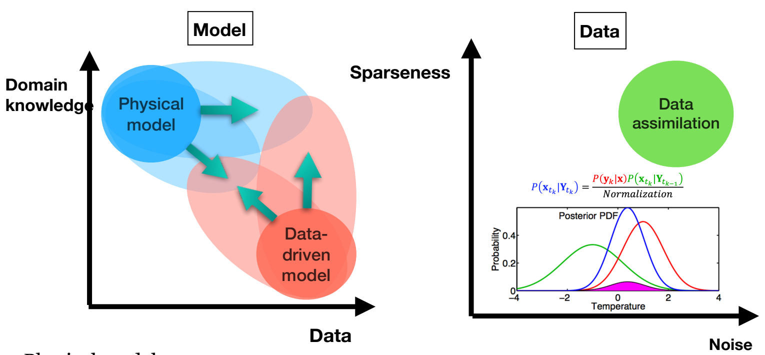
CGKN: A Deep Learning Digital Twins Framework for Stochastic Modeling, Forecast, and Data Assimilation

Chuanqi Chen, **Zhongrui Wang**, Nan Chen, Jin-long Wu
Department of Mechanical Engineering & Department of Mathematics, University of Wisconsin-Madison

Motivation

For **complex dynamical systems** (nonlinear, chaotic, multi-scale, turbulent, intermittent, non-Gaussian), to:

1. Make predictions (models)
2. Use observations to improve predictions (data assimilation; DA)

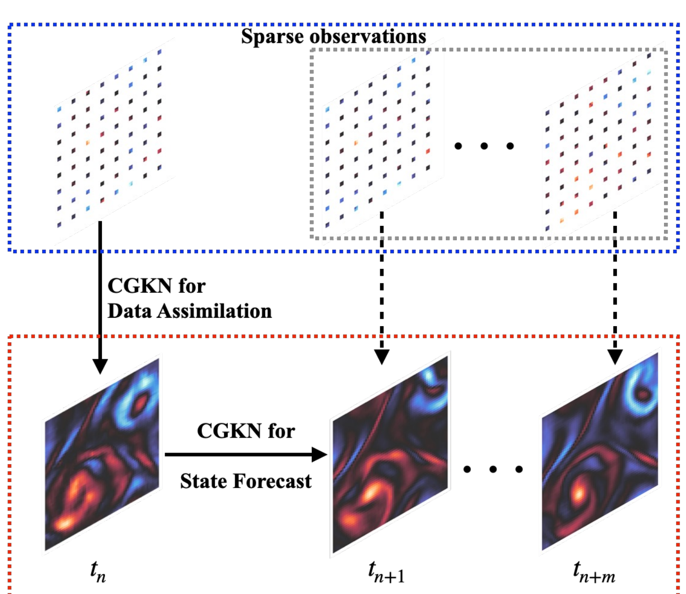


- Physical model:
  - Based on governing equations derived from first principles (interpretable)
  - May require strong assumptions; Usually computationally expensive (e.g., NWP)
- Data-driven model:
  - Works with governing equations unknown (but lack of interpretability)
  - Computational efficient; Needs a large amount of data (may be sparse and noisy in reality)
- Data assimilation is especially useful when data is sparse and noisy
  - Combing data with existing models, DA can recover complete data, with reduced uncertainty
  - As new observations become available, DA can utilize this information to improve real-time predictions

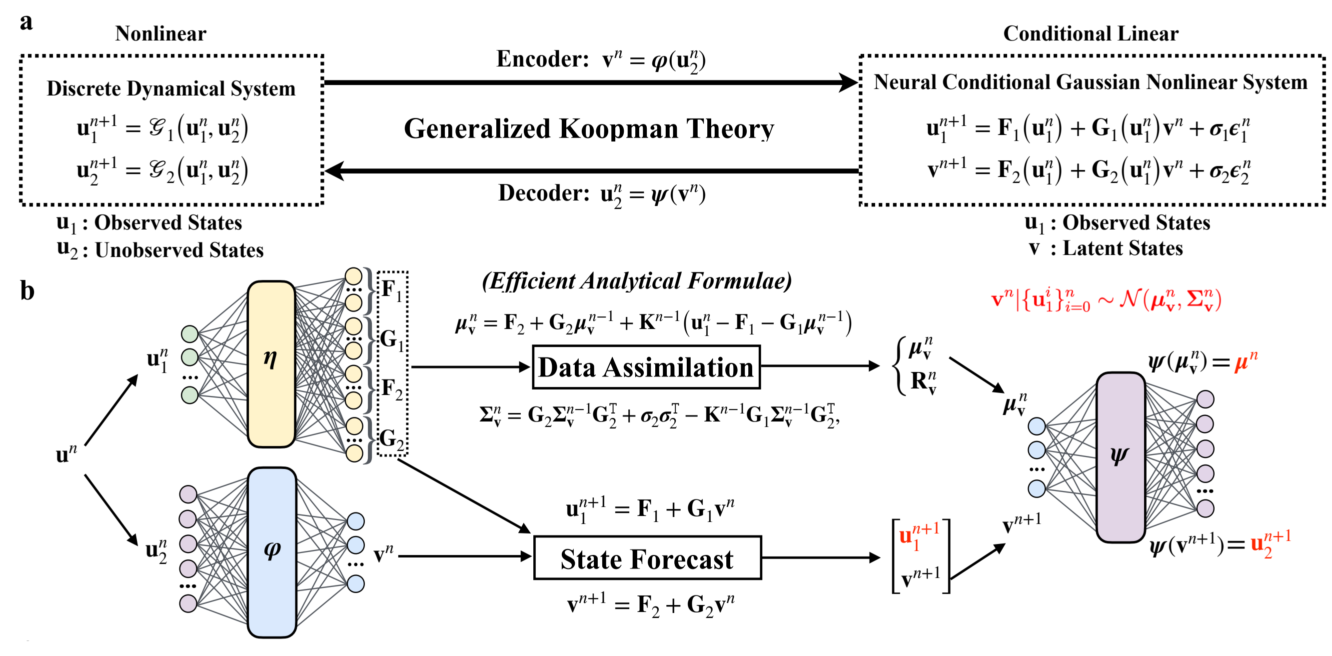
CGKN: Overview

Conditional Gaussian Ksoopman Network (CGKN):

A unified framework of SciML and DA, to learn surrogate models that performs efficient prediction and DA for nonlinear partially observed dynamical systems.



CGKN: Methodology



Autoencoder loss | Forecast loss of u_1 and u_2 | Forecast loss of v | DA loss

$$L(\theta_\varphi, \theta_\psi, \theta_\eta) := \lambda_{AE}L_{AE}(\theta_\varphi, \theta_\psi) + \lambda_u L_u(\theta_\varphi, \theta_\psi, \theta_\eta) + \lambda_v L_v(\theta_\varphi, \theta_\eta) + \lambda_{DA}L_{DA}(\theta_\psi, \theta_\eta)$$

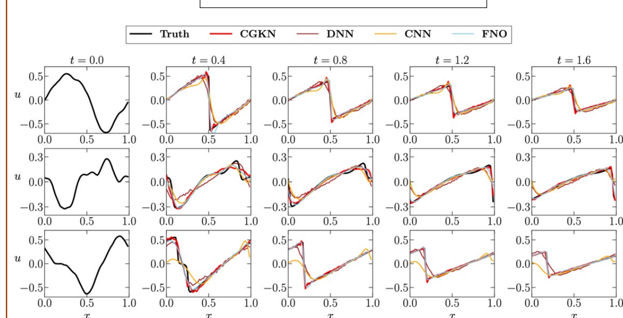
CGKN for VBE

Viscous Burgers' equation (shock waves)

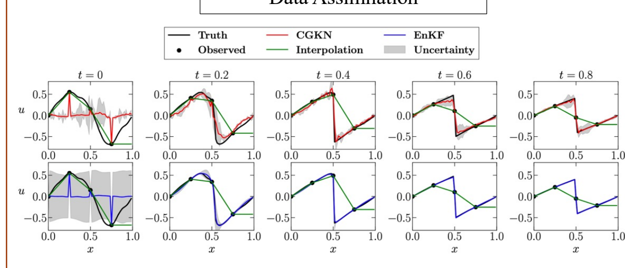
$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} + \nu\frac{\partial^2 u}{\partial x^2}, \quad x \in [0, L_x] \quad t \in [0, L_t]$$

- Observed states: **4** out of **64** (uniformly distributed); unobserved states: 60 out of 64
- Encoder $\varphi : \mathbb{R}^{60} \mapsto \mathbb{R}^{10}$, decoder $\psi : \mathbb{R}^{10} \mapsto \mathbb{R}^{60}$

State Forecast



Data Assimilation



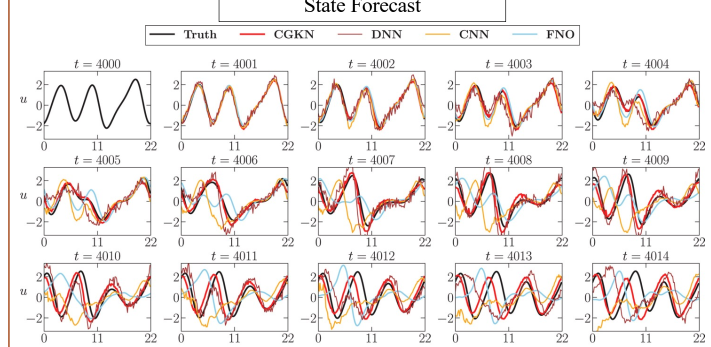
CGKN for KSE

Kuramoto-Sivashinsky Equation (1-D Chaotic PDE)

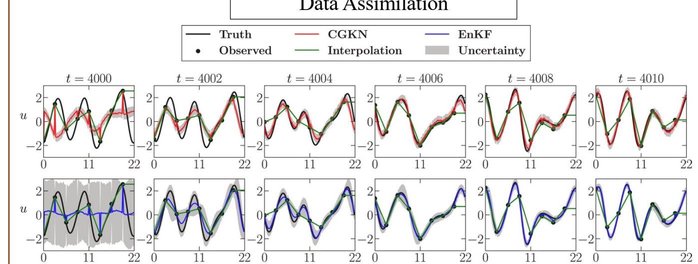
$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4}, \quad x \in [0, L_x] \quad t \in [0, L_t]$$

- Observed states: **8** out of **128** (uniformly distributed); unobserved states: 120 out of 128; with observational noise ~ $\mathcal{N}(0, 0.2^2)$
- Encoder $\varphi : \mathbb{R}^{120} \mapsto \mathbb{R}^{12}$ decoder $\psi : \mathbb{R}^{12} \mapsto \mathbb{R}^{120}$

State Forecast



Data Assimilation



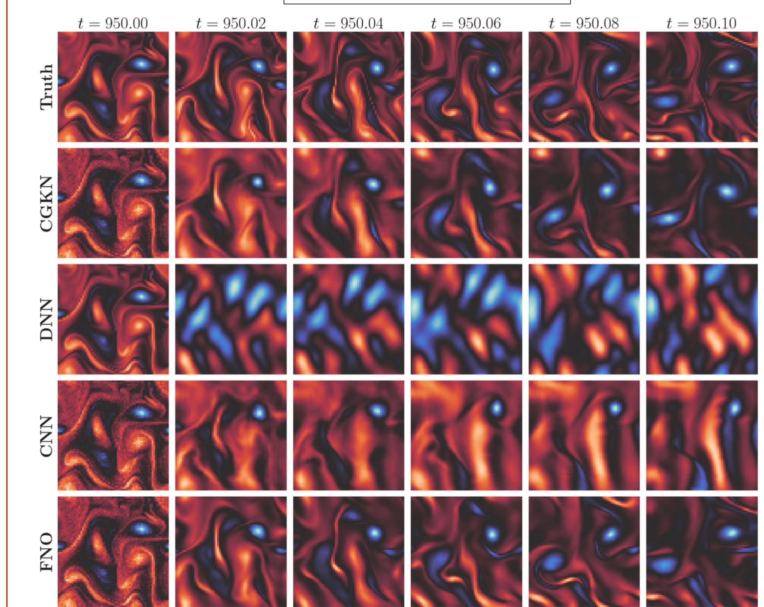
CGKN for NSE

Navier-Stokes Equations (2-D turbulent PDE)

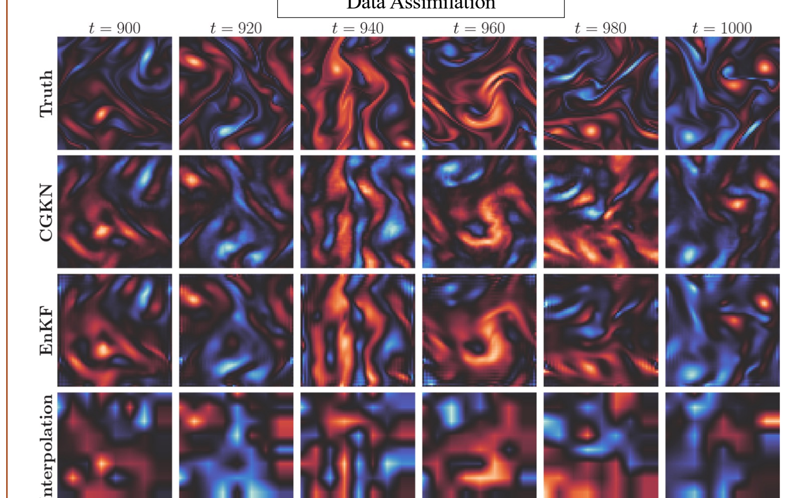
$$\frac{\partial \mathbf{u}}{\partial t} = -\mathbf{u}\cdot\nabla\mathbf{u} + \nu\nabla^2\mathbf{u} - \frac{1}{\rho}\nabla p + \mathbf{f}, \quad \mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^2$$
$$\nabla\cdot\mathbf{u} = 0, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^2, \quad t \in [0, L_t]$$

- Observed states: **8x8 out of 64x64** (uniformly distributed); unobserved states: 120 out of 128; observational noise ~$\mathcal{N}(0, 2.5^2)$
- Encoder $\varphi : \mathbb{R}^{64\times64} \mapsto \mathbb{R}^{16\times16}$ decoder $\psi : \mathbb{R}^{16\times l6} \mapsto \mathbb{R}^{64\times64}$

State Forecast



Data Assimilation



Summary of Numerical Results

MSEs; one-step error for forecast/time averaged error for DA

Methods \ Examples	Viscous Burgers Equation		Kuramoto–Sivashinsky Equation		Navier–Stokes Equations	
	Forecast Error	DA Error	Forecast Error	DA Error	Forecast Error	DA Error
CGKN	7.5683e-04	7.5037e-04	1.1042e-02	2.4927e-02	1.9754e+01	6.0940e+01
EnKF	—	5.8125e-04	—	2.4882e-02	—	6.9010e+01
Interpolation	—	1.3514e-02	—	4.3097e-01	—	1.2844e+02
DNN	6.4816e-03	—	4.7332e-02	—	1.0936e+02	—
CNN	2.3727e-03	—	2.6111e-02	—	3.0600e+01	—
FNO	3.9715e-04	—	5.4859e-03	—	1.7129e+01	—

Computational efficiency compared to EnKF:
~ 600 times faster for VBE; ~ 125 times faster for KSE; ~ 300 times faster for NSE;

A **discrete-time CGKN** is developed to learn surrogate models for **efficient state forecasting and DA for high-dimensional, partially observed, complex dynamical systems**. CGKN leverages **Koopman embedding** to construct latent variables representing unobserved states, whose dynamics are **conditionally linear given the observed states**. This structure yields a **conditional Gaussian system with closed-form DA formulae**, which are embedded into the learning process as inductive bias, resulting in **a unified framework that integrates scientific machine learning (SciML) with DA**. Beyond DA, the CGKN framework exemplifies how SciML models can be designed to seamlessly interface with outer-loop applications such as design optimization, inverse problems, and optimal control.

Chuanqi Chen, Zhongrui Wang, Nan Chen, and Jin-Long Wu. Modeling partially observed nonlinear dynamical systems and efficient data assimilation via discrete-time conditional Gaussian Koopman network.. Computer Methods in Applied Mechanics and Engineering, 2025.