

Supplementary Material of Stochastic Neural Simulator for Generalizing Dynamical Systems Across Environments

CONTENTS

1 Proof of Proposition 1	2
2 Implementation Details	2
A Initial Encoder	2
B Context-informed Encoder	3
C Latent ODE Function and ODE Solver	3
D State Decoder	3
E Dimensions of Environment-specific variable	4
3 Details of The Baseline Models	4
A The Implementation of CoDA	4
B The Implementation of LEADS	5
4 Experiment Settings	5
A Lotka-Volterra (LV)	5
B Glycolytic-Oscillator (GO)	5
C Gray-Scott (GS)	6
D Navier-Stokes (NS)	7
E Heat diffusion equation (HEAT)	7
5 Additional Results	8
A Performance Evaluation Results	8
B Discovering Governing Law of Dynamical Systems	8
C Additional Visualizations	8

1. PROOF OF PROPOSITION 1

In the CoNDP, given \mathbf{u}^e and \mathbf{z}_{t_0} , the evolved latent state \mathbf{z}_{t_i} can be seen as a deterministic function dominated by t_i and \mathbf{z}_{t_0} . Thus, we have the following Lemma to assist in proving our propositions.

Lemma 1. *The evolved latent state \mathbf{z}_{t_i} can be seen as a deterministic function $\mathcal{F}(t_i)$ for a given fixed \mathbf{u}^e and \mathbf{z}_{t_0} .*

Proof. We have the fact that

$$\mathbf{z}_{t_i} = \mathbf{z}_{t_0} + \int_{t_0}^{t_i} f_{\{\theta^e, \theta^e\}}(\mathbf{z}_\tau, \tau) d\tau,$$

where, $\theta^e = \text{con}_{\varphi_m}(\mathbf{u}^e)$. Note that only \mathbf{u}^e (global control variable) and \mathbf{z}_{t_0} (initial hidden states) are uncertain, while the rest are deterministic. Therefore, when both \mathbf{u}^e and \mathbf{z}_{t_0} are fixed, the above equation can be regarded as a deterministic function controlled by t_i , i.e., $\mathcal{F}(t_i)$. \square

Proposition 1. *xc CoNDP satisfies the exchangeability and consistency conditions.*

Proof. First, we prove that CoNDP satisfies the exchangeability condition. Following Lemma 1, for any given \mathbf{u}^e and \mathbf{z}_{t_0} , any permutation π of $\{1, 2, \dots, T\}$ on $t_{1:T}$ would automatically act on $\mathcal{F}_{\pi(1:T)}$, i.e.,

$$\mathcal{F}_{\pi(1:T)} = \mathcal{F}(t_{\pi(1:T)}),$$

where $\pi(1 : T) = (\pi(1), \pi(2), \dots, \pi(T))$. And then the permutation π would act on the condition predictive distribution, i.e., $p(x_{\pi(1:T)}^e | \mathbf{z}_{t_0}, \mathbf{u}^e, t_{\pi(1:T)}) = \mathcal{N}(\mu_{x_{\pi(1:T)}}, \sigma_{x_{\pi(1:T)}}^2)$, where $\mu_{x_{\pi(1:T)}}, \sigma_{x_{\pi(1:T)}}^2 = \text{dec}_{\varphi_d}(\mathbf{z}_{t_{\pi(1:T)}})$. Therefore, the exchangeability condition is guaranteed.

Then, we prove that CoNDP satisfies the consistency condition. Based on Lemma 1, we can write the joint distribution similar to NPs [1] or NDPs [2] as follows

$$\rho_{t_{1:T}}(\mathbf{X}_{t_{1:T}}) = \int p(\mathcal{F}) \prod_{i=1}^T p(\mathbf{X}_{t_i} | \mathcal{F}(t_i)) d\mathcal{F}.$$

Since the probability density function of any \mathbf{X}_{t_i} depends only on the corresponding t_i , integrating out any subset of $\mathbf{X}_{t_{1:T}}$ gives the joint distribution of the remaining random variables in the sequence as

$$\begin{aligned} & \int \rho_{t_{1:T}}(\mathbf{X}_{t_{1:T}}) d\mathbf{X}_{t_{n+1:T}} \\ &= \int \int p(\mathcal{F}) \prod_{i=1}^T p(\mathbf{X}_i | \mathcal{F}(t_i)) d\mathcal{F} d\mathbf{X}_{t_{n+1:T}} \\ &= \int p(\mathcal{F}) \prod_{i=1}^n p(\mathbf{X}_i | \mathcal{F}(t_i)) d\mathcal{F} \\ &= \rho_{t_{1:n}}(\mathbf{X}_{t_{1:n}}). \end{aligned}$$

Therefore, consistency is also guaranteed. \square

2. IMPLEMENTATION DETAILS

We now introduce the implementation details in our model.

A. Initial Encoder

The initial encoder maps the initial state variable $x_{t_0}^e$ into the latent state \mathbf{z}_{t_0} . We utilize a multi-layer perceptron (MLP) for the LV and GO systems, and convolutional neural networks (CNN) with 3×3 kernels for the GS, NS and HEAT systems. The number of hidden layers, dimensions/channels and activations are listed in Table S1.

Table S1. Neural architectures of the initial encoder.

Systems	Hidden Layers	Dimensions/Channels
LV	3	64,32,16
GO	3	64,32,16
GS	4	32,64,32,16
NS	4	32,64,32,32
HEAT	4	32,64,32,16

B. Context-informed Encoder

The context-informed encoder performs representation learning for all observations in context. We use a MLP to encode the context for the LV and GO systems, and convolutional neural networks (CNN) for the GS, NS and HEAT systems. For both LV and GO systems, the encoding is conducted directly on the context sequence. For GS, NS and HEAT systems, each observation is first fed into a CNN for dimension reduction, the CNN network is shared across all environments. The details are listed in Tables S2 and S3.

Table S2. Neural architectures of the context-informed encoder (MLP part).

Systems	Hidden Layers	Dimensions
LV	2	64,64
GO	2	32,64
GS	2	32,64,64
NS	3	32,64,64
HEAT	3	32,64,64

Table S3. Neural architectures of the context-informed encoder (CNN part).

Systems	Hidden Layers	Kernel	Channels
GS	4	3×3	64,64,32,16
NS	4	3×3	64,64,32,16
HEAT	4	3×3	32,128,64,16

C. Latent ODE Function and ODE Solver

In the process of solving the initial value problem, we need to model the latent ODE function $f_{\{\theta^e, \theta^c\}}$. We use a MLP to build the function $f_{\{\theta^e, \theta^c\}}$, where the weights of the first layer is set to be θ^e tuned per-environment and others as θ^c shared across environments. The number of layers in function f and the dimension of the evolving latent state z_t are listed in Table S4.

For solving the initial value problem, we use fourth-order Runge-Kutta method [3] from the torchdiffeq¹ package as the ODE solver and use the Adjoint method described in [4] for backpropagation, which reduces the memory cost to a constant.

D. State Decoder

The state decoder calculates the mean and variance values for state variable at each time points, i.e., $\mu_{x_{t_j}}, \sigma_{x_{t_j}}^2 = dec_{\varphi_d}(z_{t_j}), 1 \leq j \leq T$, by equally splitting the output of a neural network into two

¹<https://github.com/rtqichen/torchdiffeq>

Table S4. Neural architectures of the latent ODE network.

Systems	Hidden Layers	Dimensions
LV	2	32,64
GO	2	32,64
GS	2	32,64
NS	3	32,64,64
HEAT	2	32,64

halves. The first half is the mean and the second half is the variance which should be non-negative. Therefore, the output dimension or channel is set to be 2 times of the system state dimension.

For the LV and GO systems, we use a MLP to calculate the prediction distribution. For the GS, NS and HEAT systems, we use transposed Convolutional Neural Networks and batch normalisation for decoding. The number of layers, channels/dimensions and activations are listed in Table S6.

Table S5. Neural architectures of the state decoder.

Systems	Hidden Layers	Dimensions/Channels	Activation
LV	4	64,64,64,64	Swish [5]
GO	4	64,64,64,64	Swish
GS	4	64,64,64,64	Swish
NS	4	64,64,64,64	Swish
HEAT	4	32,128,64,32	ReLU

E. Dimensions of Environment-specific variable

The environment-specific variable ξ^e is utilized as the input for the hyper network to compute the coefficients of the attention layers. It has a relatively low dimensionality compared to the attention parameters, allowing for efficient compression of the parameter size involved in the fine-tuning process. Below are the dimensions of ξ^e for each system.

Table S6. Dimensions of Environment-specific variable.

Systems	Dimensions
LV	4
GO	10
GS	6
NS	8
HEAT	8

3. DETAILS OF THE BASELINE MODELS

We now introduce the implementation details of the baseline models.

A. The Implementation of CoDA

Consistent with the architecture reported in the respective published paper. We implement the model with the following architectures:

- LV,GO:4-layer MLPs with hidden layers of width 64.
- GS: 4-layer ConvNet with 64-channel hidden layers, and 3×3 convolution kernels
- NS:Fourier Neural Operator [6] with 4 spectral convolution layers. 12 frequency modes and hidden layers with width 10.

B. The Implementation of LEADS

In accordance with the architecture detailed in the respective published paper, we implement the model using the following architectures.

- LV,GO:4-layer MLPs with hidden layers of width 64.
- GS,NS:4-layer ConvNet with 64-channel hidden layers, and 3×3 convolution kernels

4. EXPERIMENT SETTINGS

Here, we introduce the detailed experiment settings in all dynamical systems.

A. Lotka-Volterra (LV)

Lotka-Volterra (LV, [7]) system describes the interaction between a prey-predator pair in an ecosystem, formalized into the following ODE:

$$\begin{aligned}\frac{dx}{dt} &= \alpha x - \beta xy \\ \frac{dy}{dt} &= \delta xy - \gamma y,\end{aligned}$$

where x, y are the quantities of the prey and the predator respectively, and $\alpha, \beta, \delta, \gamma$ define how two species interact.

We generate trajectories on a temporal grid with $\Delta t = 0.1$ and the temporal horizon of 20. For training, we sample $N_{tr}^{tra} = 10$ initial conditions from a uniform distribution $p(\mathbf{x}_{t_0}) = \text{Unif}([1, 3]^2)$ to produce the training set. Model is tuned on $N_{tu}^{tra} = 10$ trajectories under the new environment and then employed to testing on $N_{te}^{tra} = 100$ trajectories. We consider $N_{tr}^{env} = 9$ environments for training. The ranges of each parameter for training, tuning, and testing are listed in Table S7.

Table S7. Ranges for parameters in the LV system. Note that the tuning/testing environments are sampled from the complementary set of the training environments.

	Tuning/Testing	Training
α	$[0.3, 0.7]$	$[0, 0.3) \cup (0.7, 1]$
β	$[0.55, 0.95]$	$[0.25, 0.55) \cup (0.95, 1.25]$
δ	$[0.55, 0.95]$	$[0.25, 0.55) \cup (0.95, 1.25]$
γ	$[0.3, 0.7]$	$[0, 0.3) \cup (0.7, 1]$

B. Glycolytic-Oscillator (GO)

Glycolytic-Oscillator (GO, [8]) describes yeast glycolysis dynamics with the following ODEs:

$$\begin{aligned}\frac{dS_1}{dt} &= J_0 - \frac{k_1 S_1 S_6}{1 + (1/K_1^q) S_6^q} \\ \frac{dS_2}{dt} &= 2 \frac{k_1 S_1 S_6}{1 + (1/K_1^q) S_6^q} - k_2 S_2 (N - S_5) - k_6 S_2 S_5 \\ \frac{dS_3}{dt} &= k_2 S_2 (N - S_5) - k_3 S_3 (A - S_6) \\ \frac{dS_4}{dt} &= k_3 S_3 (A - S_6) - k_4 S_4 S_5 - \kappa (S_4 - S_7)\end{aligned}$$

$$\begin{aligned}\frac{dS_5}{dt} &= k_2 S_2(N - S_5) - k_4 S_4 S_5 - k_6 S_2 S_5 \\ \frac{dS_6}{dt} &= -2 \frac{k_1 S_1 S_6}{1 + (1/K_1^q) S_6^q} + 2k_3 S_3(A - S_6) - k_5 S_6 \\ \frac{dS_7}{dt} &= \psi \kappa (S_4 - S_7) - k S_7,\end{aligned}$$

where $S_1, S_2, S_3, S_4, S_5, S_6, S_7$ represent the concentrations of 7 biochemical species.

We generate trajectories on a temporal grid with $\Delta t = 0.05$ and the temporal horizon of 1. For training, we sample $N_{tr}^{tra} = 10$ initial conditions from a uniform distribution $p(x_{t_0})$ in Table 2 in [8] to produce the training set. Model is tuned on $N_{tu}^{tra} = 10$ trajectories under the new environment and then employed to testing on $N_{te}^{tra} = 150$ trajectories. We consider $N_{tr}^{env} = 9$ environments for training. The ranges of each parameter for training, tuning, and testing are listed in Table S8.

Table S8. Ranges for parameters in the GO system. Note that the tuning/testing environments are sampled from the complementary set of the training environments.

	Tuning/Testing	Training
J_0	[2.4, 2.6]	[2.2, 2.4) \cup (2.6, 2.8]
k_1	[95, 105]	[85, 95) \cup (105, 115]
k_2	[5.5, 6.5]	[4.5, 5.5) \cup (6.5, 7.5]
k_3	[15, 17]	[13, 15) \cup (17, 19]
k_4	[95, 105]	[85, 95) \cup (105, 115]
k_5	[1.23, 1.33]	[1.13, 1.23) \cup (1.33, 1.43]
k_6	[11.5, 12.5]	[10.5, 11.5) \cup (12.5, 13.5]
K_1	[0.65, 0.85]	[0.45, 0.65) \cup (0.85, 1.05]
q	[3.8, 4.2]	[3.4, 3.8) \cup (4.2, 4.6]
N	[0.8, 1.2]	[0.4, 0.8) \cup (1.2, 1.6]
A	[3.8, 4.2]	[3.4, 3.8) \cup (4.2, 4.6]
κ	[12, 14]	[10, 12) \cup (14, 16]
ψ	[0.08, 0.12]	[0.04, 0.08) \cup (0.12, 0.16]
k	[1.7, 1.9]	[1.5, 1.7) \cup (1.9, 2.1]

C. Gray-Scott (GS)

The PDE describes a reaction-diffusion system with complex spatiotemporal patterns through the following 2D PDEs [9]:

$$\begin{aligned}\frac{\partial u}{\partial t} &= D_u \Delta u - uv^2 + F(1 - u) \\ \frac{\partial v}{\partial t} &= D_v \Delta v + uv^2 - (F + k)v,\end{aligned}$$

where u, v represent the concentrations of two chemical components. D_u, D_v denote the diffusion coefficients for u and v respectively, and F, k are the reaction parameters.

We generate trajectories on a temporal grid with $\Delta t = 40$ and the temporal horizon of 400 in a 2D spatial domain of dimensions 32×32 with spatial resolution $\Delta s = 2$. For training, we sample $N_{tr}^{tra} = 10$ to produce the training set. We define initial conditions $(u_0, v_0) \sim p(x_{t_0})$ by uniformly sampling three two-by-two squares which trigger the reactions. $(u_0, v_0) = (1 - \epsilon, \epsilon)$ with $\epsilon = 0.05$ inside the squares and $(u_0, v_0) = (0, 1)$ outside the squares. Model is tuned on $N_{tu}^{tra} = 10$ trajectories under the new environment and then employed to testing on $N_{te}^{tra} = 200$ trajectories. We consider $N_{tr}^{env} = 9$ environments for training. The ranges of each parameter for training, tuning, and testing are listed in Table S9.

Table S9. Ranges for parameters in the GS system. Note that the tuning/testing environments are sampled from the complementary set of the training environments.

	Tuning/Testing	Training
D_u	$[0.1, 0.2]$	$[0, 0.1) \cup (0.2, 0.3]$
D_v	$[0.1, 0.2]$	$[0, 0.1) \cup (0.2, 0.3]$
F	$[0.03, 0.039]$	$[0.0225, 0.03) \cup (0.039, 0.0435]$
k	$[0.058, 0.062]$	$[0.056, 0.058) \cup (0.062, 0.064]$

D. Navier-Stokes (NS)

Navier-Stokes (NS, [10]) system describes the dynamics of incompressible flows with the 2D PDEs:

$$\begin{aligned}\frac{\partial w}{\partial t} &= -v\nabla w + \nu\Delta w + f \text{ where } w = \nabla \times v \\ \nabla v &= 0,\end{aligned}$$

where v is the velocity field, $w = \nabla \times v$ is the vorticity, ν is the viscosity and f is the constant forcing term.

We generate trajectories on a temporal grid with $\Delta t = 1$ and the Temporal Horizon of 10 in a 2D spatial domain of dimensions 32×32 with spatial resolution $\Delta s = 1$. For training, we sample $N_{tr}^{tra} = 10$ initial conditions from $p(x_{t_0})$ as in [6] to produce the training set. Model is tuned on $N_{tu}^{tra} = 10$ trajectories under the new environment and then employed to testing on $N_{te}^{tra} = 100$ trajectories. We consider $N_{tr}^{env} = 15$ environments for training. The ranges of each parameter for for training, tuning, and testing are listed in Table S10.

Table S10. Ranges for parameters in the NS system. Note that the tuning/testing environments are sampled from the complementary set of the training environments.

	Tuning/Testing	Training
ν	$[0.8, 1.2]$	$[0.6, 0.8) \cup (1.2, 1.4]$

E. Heat diffusion equation (HEAT)

The Heat diffusion equation (HEAT, [11]) describes how heat diffuses through a given region following:

$$\frac{\partial u}{\partial t} = \alpha\Delta u,$$

where u is the thermal energy and α is the thermal diffusivity of the medium.

We generate trajectories on a temporal grid with $\Delta t = 1$ and the Temporal Horizon of 20. For training, we sample $N_{tr}^{tra} = 10$ initial conditions to produce the training set. We define intial conditions $u_0 \sim p(x_{t_0})$ bu uniformly sampling three two-by-two squares which trigger the reactions. $u_0 = 1$ inside the squares and $u_0 = 0$ outside the squares. Model is tuned on $N_{tu}^{tra} = 10$ trajectories under the new environment and then employed to testing on $N_{te}^{tra} = 100$ trajectories. We consider $N_{tr}^{env} = 5$ environments for training. The ranges of each parameter for for training, tuning, and testing are listed in Table S11.

Table S11. Ranges for parameters in the HEAT system. Note that the tuning/testing environments are sampled from the complementary set of the training environments.

	Tuning/Testing	Training
α	$[0.4, 0.6]$	$[0.2, 0.4) \cup (0.6, 0.8]$

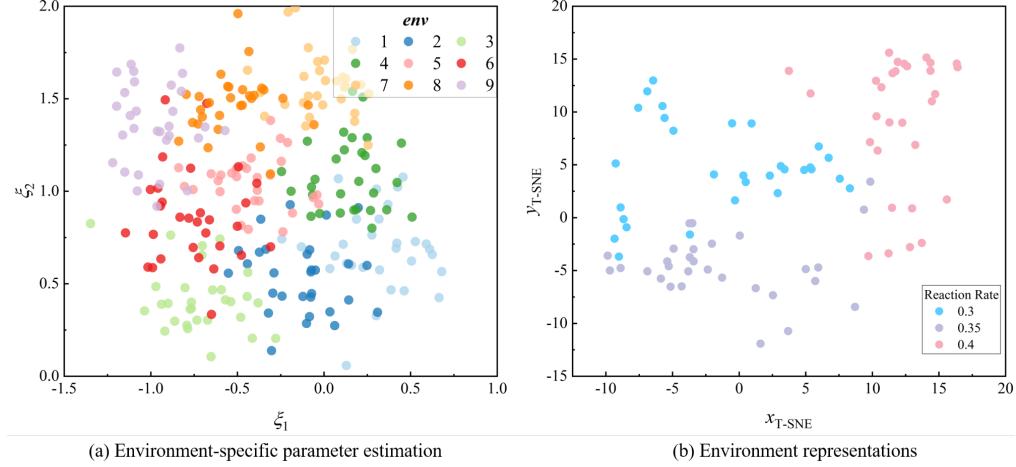


Fig. S1. T-SNE visualizations of the environment-specific parameter ξ^e from GS system (a) and the environment representations u^e also from GS system (b).

5. ADDITIONAL RESULTS

A. Performance Evaluation Results

Performance evaluation under different sparsity levels for all systems is shown in Fig. S2. We also visualize the predictive results of the LEADS, CoDA, and CoNDP on the GS and NS systems, as shown in Figs. S3 and S4. We consider a new testing trajectory under the inductive setting with the new environment.

B. Discovering Governing Law of Dynamical Systems

The observed data is generated with the following ODEs:

$$\begin{aligned}\frac{dx}{dt} &= 0.5x - 0.75xy, \\ \frac{dy}{dt} &= 0.75xy - 0.5y.\end{aligned}$$

The discovered governing laws for the LV system under various sparse levels by the LEADS, CoDA and CoNDP are listed in Table S12.

C. Additional Visualizations

We visualize the environment-specific parameter ξ^e and the environment representations u^e from GS system in S1.

REFERENCES

1. M. Garnelo, J. Schwarz, D. Rosenbaum, *et al.*, “Neural processes,” in *ICML Workshop on Theoretical Foundations and Applications of Deep Generative Models*, (2018).
2. A. Norcliffe, C. Bodnar, B. Day, *et al.*, “Neural ode processes,” in *International Conference on Learning Representations*, (2020).
3. M. Schober, S. Särkkä, and P. Hennig, “A probabilistic model for the numerical solution of initial value problems,” *Stat. Comput.* **29**, 99–122 (2019).
4. R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, “Neural ordinary differential equations,” *Adv. neural information processing systems* **31** (2018).
5. P. Ramachandran, B. Zoph, and Q. V. Le, “Searching for activation functions,” arXiv preprint arXiv:1710.05941 (2017).
6. Z. Li, N. Kovachki, K. Azizzadenesheli, *et al.*, “Fourier neural operator for parametric partial differential equations,” arXiv preprint arXiv:2010.08895 (2020).
7. A. J. Lotka, *Elements of physical biology* (Williams & Wilkins, 1925).

8. B. C. Daniels and I. Nemenman, "Efficient inference of parsimonious phenomenological models of cellular dynamics using s-systems and alternating regression," *PloS one* **10**, e0119821 (2015).
9. J. E. Pearson, "Complex patterns in a simple system," *Science* **261**, 189–192 (1993).
10. G. Stokes, "On the effect of the internal friction of fluids on the motion of pendulums," *Transactions Camb. Philos. Soc.* **9**, 8 (1851).
11. D. V. Widder, *The heat equation*, vol. 67 (Academic Press, 1976).

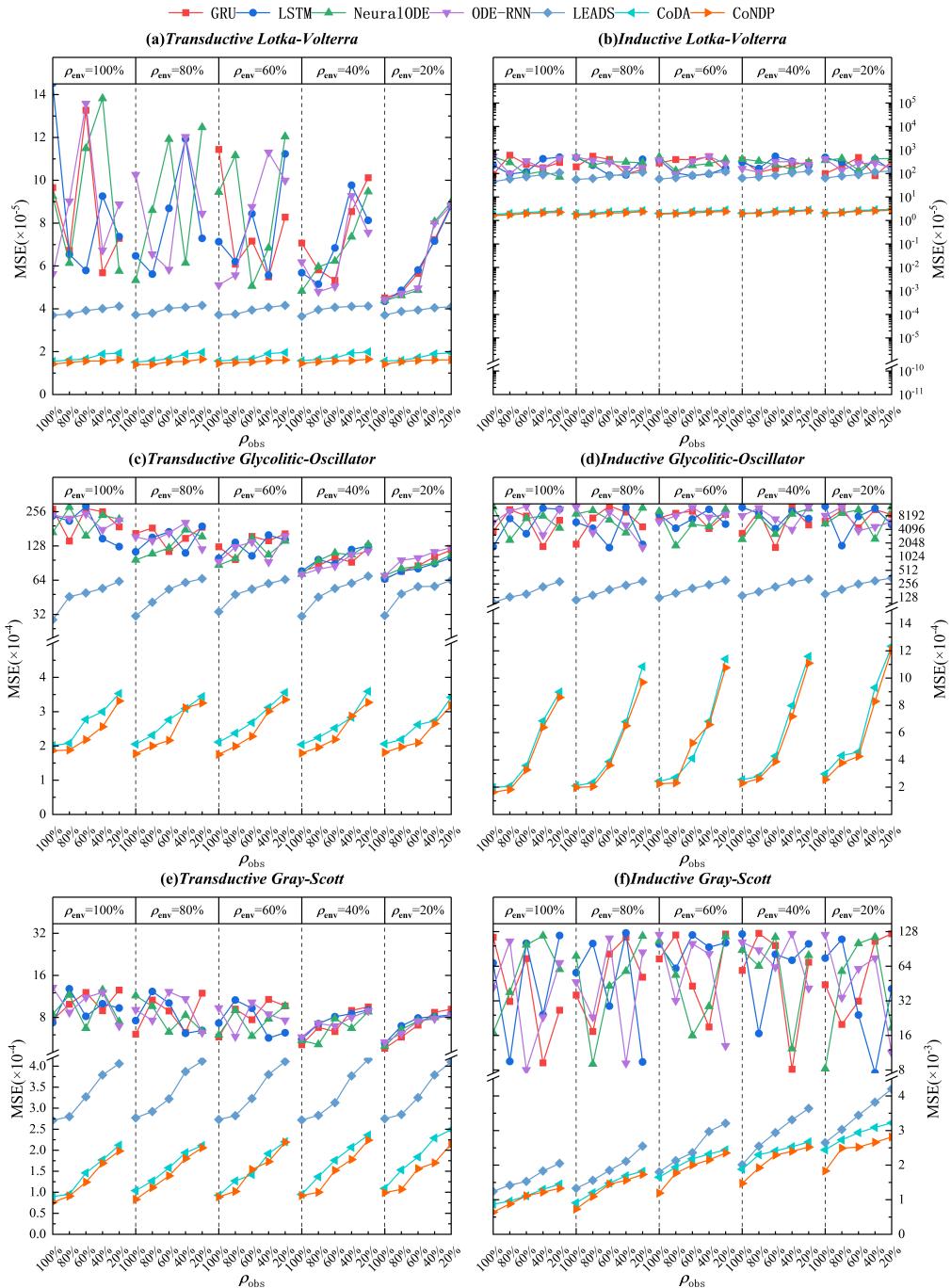


Fig. S2. Comparison of prediction errors under different sparsity levels for LV (a-b), GO (c-d), and GS (e-f) systems. We see that the sparse scenarios do indeed weaken model performance. Sufficient training data enables the models to learn the commonalities and differences among environments, thereby improving performance. And, the sparsity of testing data makes it difficult to capture features for new environments. Our CoNDP can achieve optimal results compared with common neural simulators and state-of-the-art cross-environmental models.

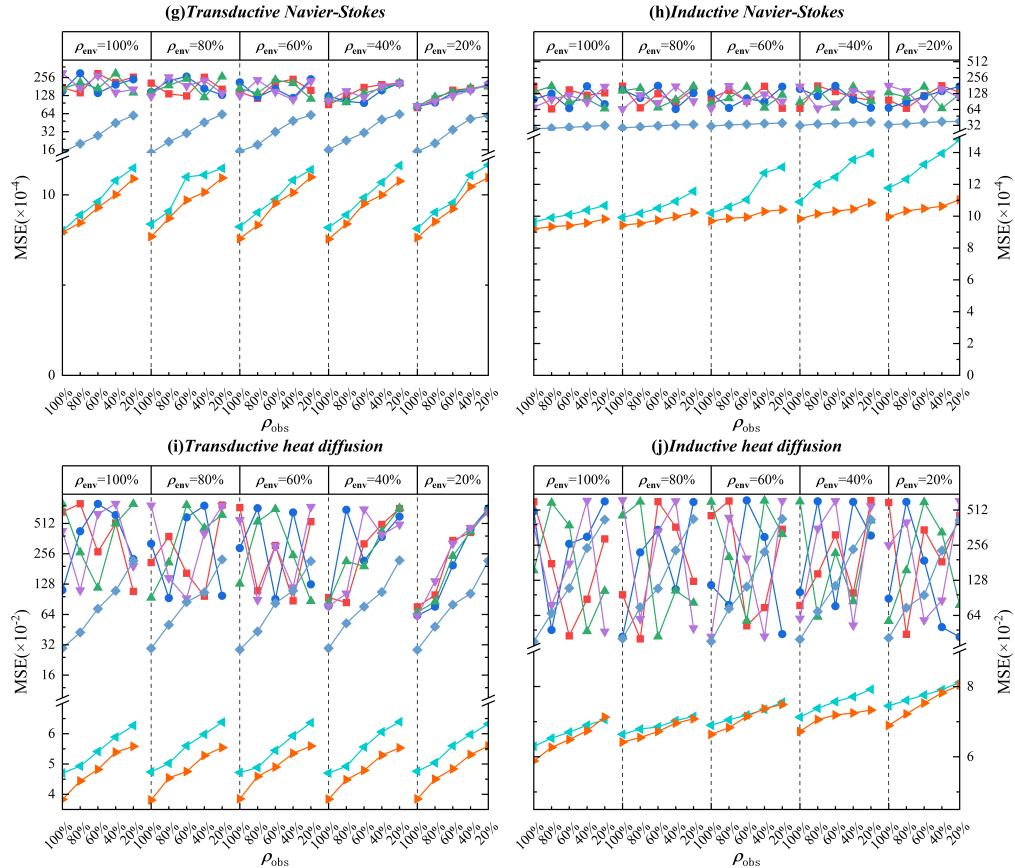


Fig. S2. Comparison of prediction errors under different sparsity levels for NS (g-h) and HEAT (i-j) systems. We see that the sparse scenarios do indeed weaken model performance. Sufficient training data enables the models to learn the commonalities and differences among environments, thereby improving performance. And, the sparsity of testing data makes it difficult to capture features for new environments. Our CoNDP can achieve optimal results compared with common neural simulators and state-of-the-art cross-environmental models.

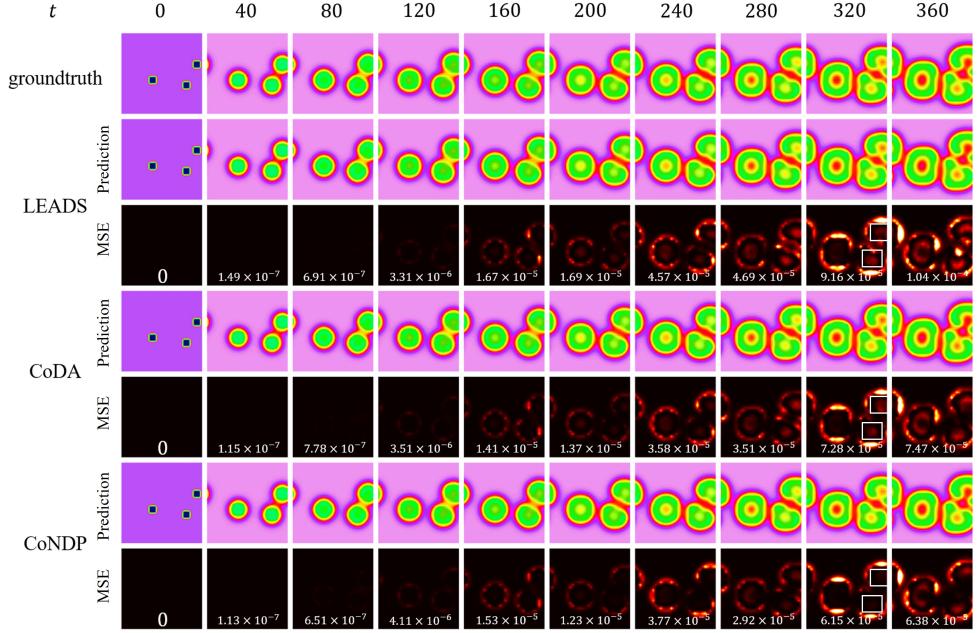


Fig. S3. Adaptation to new GS system with $(D_u, D_v, F, k) = (0.18, 0.11, 0.035, 0.059)$. The ground truth trajectory, the predictive trajectories, and the Mean Square Error between prediction and ground truth per frame for LEADS, CoDA, and CoNDP are shown in the figure. We see that our CoNDP has more accurate predictions and lower errors than other models.

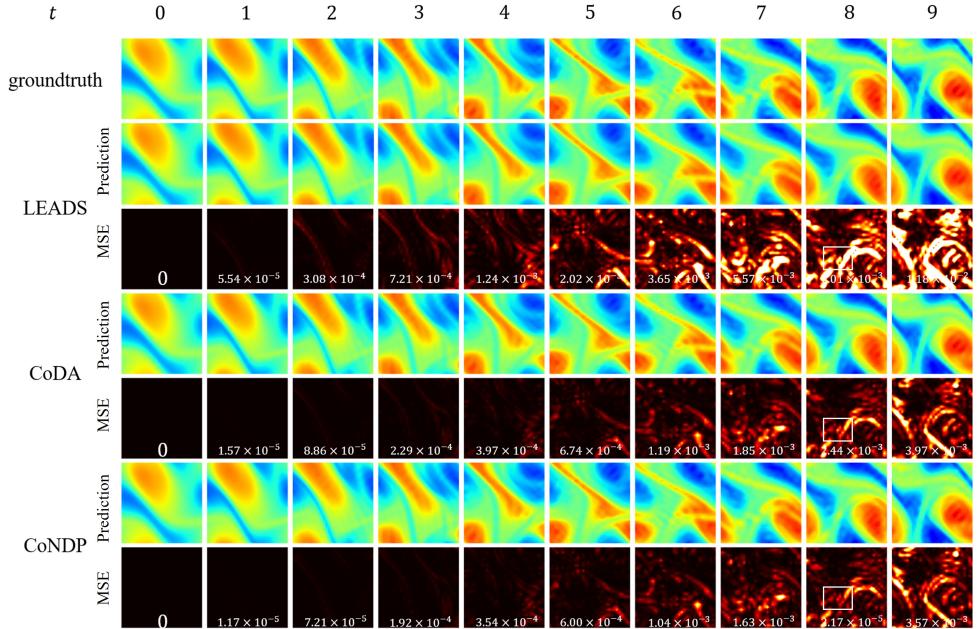


Fig. S4. Adaptation to new NS system with $\nu = 1.05 \times 10^{-4}$. The ground truth trajectory, the predictive trajectories, and the Mean Square Error between prediction and ground truth per frame for LEADS, CoDA, and CoNDP are shown in the figure. We see that our CoNDP has more accurate predictions and lower errors than other models.

Table S12. The discovered governing laws for the LV system under various sparse levels by the LEADS, CoDA, and CoNDP. The dominant terms in ground truth are colored. We see that the discovered equations all contain dominant terms, but as the scene becomes sparser, more and more redundant terms appear in the discovered equations. Compared to the LEADS and CoDA, our CoNDP has relatively fewer redundant terms and their corresponding coefficients are also relatively small.

ρ_{test}	ρ_{train}	LEADS	CoDA	CoNDP
100%	100%	$\begin{aligned} \frac{dx}{dt} = & 0.013 + \textcolor{blue}{0.48x} + 0.026y \\ & + 0.008x^2 + 0.012y^2 - \textcolor{red}{0.71xy} \\ & + 0.006 \sin x + 0.01 \sin y \\ & - 0.043 \cos x + 0.011 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.216 + 0.019x - \textcolor{red}{0.519y} \\ & - 0.156x^2 + \textcolor{blue}{0.761xy} \\ & - 0.023 \sin x - 0.019 \sin y \\ & + 0.021 \cos x - 0.048 \cos y \end{aligned}$ <p>(19 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & 0.0004 + \textcolor{blue}{0.501x} - 0.001y \\ & - 0.001x^2 - 0.001y^2 - \textcolor{red}{0.746xy} \\ & + 0.001 \sin y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.001 - \textcolor{red}{0.499y} \\ & + 0.0005x^2 + \textcolor{blue}{0.751xy} \\ & + 0.002 \sin y \end{aligned}$ <p>(12 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & \textcolor{blue}{0.502x} \\ & + 0.001x^2 - \textcolor{red}{0.749xy} \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -\textcolor{red}{0.499y} \\ & + \textcolor{blue}{0.75xy} \\ & + 0.002 \sin y \end{aligned}$ <p>(6 terms in total)</p>
100%	60%	$\begin{aligned} \frac{dx}{dt} = & -0.02 + \textcolor{blue}{0.52x} + 0.035y \\ & + 0.028x^2 + 0.017y^2 - \textcolor{red}{0.735xy} \\ & - 0.0219 \sin y + 0.024 \cos x - 0.024 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.0055 - 0.035x - \textcolor{red}{0.512y} \\ & + 0.051x^2 + 0.0111y^2 + \textcolor{blue}{0.819xy} \\ & - 0.054 \sin x - 0.037 \sin y \\ & + 0.011 \cos x \end{aligned}$ <p>(18 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.032 + \textcolor{blue}{0.514x} + 0.039y \\ & - 0.007x^2 + 0.011y^2 - \textcolor{red}{0.724xy} \\ & - 0.02 \sin y \\ & + 0.025 \cos x - 0.02 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.004 - 0.048x - \textcolor{red}{0.512y} \\ & + 0.044x^2 + \textcolor{blue}{0.81xy} \\ & - 0.023 \sin x - 0.024 \sin y \\ & + 0.01 \cos x \end{aligned}$ <p>(17 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.006 + \textcolor{blue}{0.512x} + 0.0008y \\ & + 0.012x^2 - 0.006y^2 - \textcolor{red}{0.756xy} \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.0034 - 0.007x - \textcolor{red}{0.496y} \\ & - 0.003y^2 + \textcolor{blue}{0.758xy} \\ & - 0.01 \sin x \end{aligned}$ <p>(12 terms in total)</p>
100%	20%	$\begin{aligned} \frac{dx}{dt} = & -0.097 + \textcolor{blue}{0.52x} + 0.03y \\ & + 0.084x^2 + 0.026y^2 - \textcolor{red}{0.714xy} \\ & - 0.02 \sin x \\ & + 0.025 \cos x - 0.04 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.014 - 0.065x - \textcolor{red}{0.579y} \\ & + 0.0732x^2 + 0.012y^2 + \textcolor{blue}{0.84xy} \\ & - 0.06 \sin x - 0.07 \sin y \\ & + 0.1 \cos y \end{aligned}$ <p>(18 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.075 + \textcolor{blue}{0.44x} + 0.068y \\ & + 0.052x^2 - 0.047y^2 - \textcolor{red}{0.77xy} \\ & - 0.026 \sin x + 0.07 \sin y \\ & - 0.076 \cos x + 0.043 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.04 - \textcolor{red}{0.486y} \\ & - 0.08x^2 + 0.023y^2 + \textcolor{blue}{0.799xy} \\ & - 0.013 \sin x \\ & + 0.064 \cos x - 0.086 \cos y \end{aligned}$ <p>(18 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & 0.017 + \textcolor{blue}{0.48x} + 0.0316y \\ & - \textcolor{red}{0.729xy} \\ & - 0.01 \sin y \\ & - 0.063 \cos x + 0.01 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.025 + 0.013x - \textcolor{red}{0.5y} \\ & + \textcolor{blue}{0.77xy} \\ & - 0.037 \sin x + 0.013 \sin y \\ & + 0.021 \cos x - 0.084 \cos y \end{aligned}$ <p>(15 terms in total)</p>
20%	20%	$\begin{aligned} \frac{dx}{dt} = & -0.087 + \textcolor{blue}{0.476x} + 0.1y \\ & + 0.0805x^2 - 0.569y^2 - \textcolor{red}{0.768xy} \\ & - 0.07 \sin x + 0.04 \sin y \\ & - 0.084 \cos x + 0.064 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.033 + 0.035x - \textcolor{red}{0.476y} \\ & - 0.0414x^2 + 0.02y^2 + \textcolor{blue}{0.74xy} \\ & - 0.033 \sin x \\ & + 0.1 \cos x - 0.041 \cos y \end{aligned}$ <p>(19 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.058 + \textcolor{blue}{0.439x} + 0.0474y \\ & + 0.0383x^2 - 0.021y^2 - \textcolor{red}{0.789xy} \\ & - 0.027 \sin x + 0.0437 \sin y \\ & - 0.1 \cos x + 0.045 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.021 - 0.0167x - \textcolor{red}{0.491y} \\ & - 0.096x^2 + 0.05y^2 + \textcolor{blue}{0.78xy} \\ & - 0.02 \sin x + 0.02 \sin y \\ & + 0.058 \cos x - 0.08 \cos y \end{aligned}$ <p>(20 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.02 + \textcolor{blue}{0.48x} + 0.056y \\ & + 0.022x^2 - 0.014y^2 - \textcolor{red}{0.772xy} \\ & - 0.0217 \sin x \\ & - 0.096 \cos x + 0.043 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.0165 - 0.0164x - \textcolor{red}{0.505y} \\ & - 0.036x^2 + \textcolor{blue}{0.757xy} \\ & - 0.049 \sin x - 0.098 \cos y \end{aligned}$ <p>(16 terms in total)</p>
20%	12%	$\begin{aligned} \frac{dx}{dt} = & 0.024 + \textcolor{blue}{0.38x} - 0.06y \\ & + 0.042x^2 - 0.019y^2 - \textcolor{red}{0.83xy} \\ & + 0.055 \sin x + 0.069 \sin y \\ & + 0.065 \cos x + 0.11 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.055 + 0.0298x - \textcolor{red}{0.45y} \\ & - 0.0536x^2 - 0.117y^2 + \textcolor{blue}{0.827xy} \\ & - 0.13 \sin x + 0.033 \sin y \\ & - 0.082 \cos x - 0.806 \cos y \end{aligned}$ <p>(20 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & -0.0088 + \textcolor{blue}{0.41x} - 0.046y \\ & + 0.0454x^2 - 0.036y^2 - \textcolor{red}{0.81xy} \\ & + 0.033 \sin x + 0.037 \sin y \\ & + 0.0598 \cos x + 0.1 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.039 + 0.06x - \textcolor{red}{0.46y} \\ & - 0.05x^2 - 0.11y^2 + \textcolor{blue}{0.808xy} \\ & - 0.1398 \sin x \\ & - 0.0969 \cos x - 0.075 \cos y \end{math> (19 terms in total) $	$\begin{aligned} \frac{dx}{dt} = & 0.033 + \textcolor{blue}{0.36x} \\ & + 0.01x^2 - 0.01y^2 - \textcolor{red}{0.84xy} \\ & + 0.01 \sin x + 0.04 \sin y \\ & + 0.021 \cos x + 0.04 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & -0.027 + 0.0315x - \textcolor{red}{0.48y} \\ & - 0.036x^2 - 0.17y^2 + \textcolor{blue}{0.8xy} \\ & - 0.08 \sin x - 0.013 \sin y \\ & - 0.077 \cos x - 0.061 \cos y \end{math> (19 terms in total) $
20%	4%	$\begin{aligned} \frac{dx}{dt} = & 0.131 + \textcolor{blue}{0.596x} + 0.117y \\ & + 0.094x^2 - 0.162y^2 - \textcolor{red}{0.91xy} \\ & - 0.144 \sin x + 0.388 \sin y \\ & + 0.0447 \cos x - 0.4 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.148 + 0.213x - \textcolor{red}{0.217y} \\ & + 0.473x^2 - 0.164y^2 + \textcolor{blue}{0.8xy} \\ & + 0.274 \sin x - 0.164 \sin y \\ & + 0.277 \cos x - 0.247 \cos y \end{aligned}$ <p>(20 terms in total)</p>	$\begin{aligned} \frac{dx}{dt} = & 0.125 + \textcolor{blue}{0.59x} + 0.12y \\ & + 0.065x^2 - 0.13y^2 - \textcolor{red}{0.81xy} \\ & - 0.11 \sin x + 0.39 \sin y \\ & + 0.048 \cos x - 0.34 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.148 + 0.172x - \textcolor{red}{0.236y} \\ & + 0.43x^2 - 0.17y^2 + \textcolor{blue}{0.89xy} \\ & + 0.277 \sin x - 0.212 \sin y \\ & + 0.235 \cos x - 0.241 \cos y \end{math> (20 terms in total) $	$\begin{aligned} \frac{dx}{dt} = & -0.085 + \textcolor{blue}{0.233x} + 0.0225y \\ & + 0.188x^2 - 0.148y^2 - \textcolor{red}{0.704xy} \\ & - 0.035 \sin x - 0.085 \sin y \\ & - 0.11 \cos x + 0.144 \cos y \end{aligned}$ $\begin{aligned} \frac{dy}{dt} = & 0.144 - 0.035x - \textcolor{red}{0.416y} \\ & + 0.077x^2 + 0.459y^2 + \textcolor{blue}{0.786xy} \\ & + 0.14 \sin x - 0.189 \sin y \\ & + 0.138 \cos x \end{math> (19 terms in total) $