# Improving OOD prediction by disentangling ODE parameters from dynamics


Stathi Fotiadis[1], Mario Lino[2], Chris Cantwell[2], Anil Bharath[1]

[1]Department of Bioengineering, Imperial College London [2]Department of Aeronautics, Imperial College London


Imperial College London

## MOTIVATION

Deep networks are becoming increasingly more accurate in modeling dynamical systems, but generalization is still elusive even in simple cases.

For example, predicting the trajectory of a swinging pendulum with unseen length is far from trivial for NNs.

## DISENTANGLEMENT IN VAEs

Disentangling the latent space of Variational Autoencoders[1] aims to capture the factors of variation in the data with distinct latents.

It can help to capture out-of-distribution characteristics in image modelling[2].

Previous attempts in disentanglement for dynamical systems have focused on the unsupervised setting[3].

## ODE PARAMETERS AS FACTORS OF VARIATION

Treating ODEs parameters as factors of variation allows employing latent disentanglement using ground truth values from simulated data.

**Simple pendulum:** $\ddot{\theta} + \frac{g}{\ell} \sin\theta = 0$

**Lotka-Volterra:** $\dot{x} = \alpha x - \beta x y$, $\dot{y} = \delta x y - \gamma y$

**3-body system:** $\bar{m}_i \frac{d\vec{v}_i}{dt} = K_1 \sum_j \frac{\bar{m}_i \bar{m}_j}{\bar{r}_{ij}^3} \overrightarrow{r_{ij}}$, $\frac{d\vec{x}_i}{d\bar{t}} = K_2 \vec{v}_i, i \in 1,2,3$

## MODELS

VAE-SD: a loss term is added that promotes latent disentanglement (green part). G=Identity

VAE-SSD: SD+linearly scaling of latent variables by min/max values of ODE params $\xi_i$

$\mathcal{G}\left(\mu_{z_i}\right) = \mu_{z_i} \cdot (\max(\xi_i) - \min(\xi_i)) + \min(\xi_i)$

MLP-SD:'disentangled' deterministic autoencoder

## DISENTANGLED VAE LOSS FUNCTION

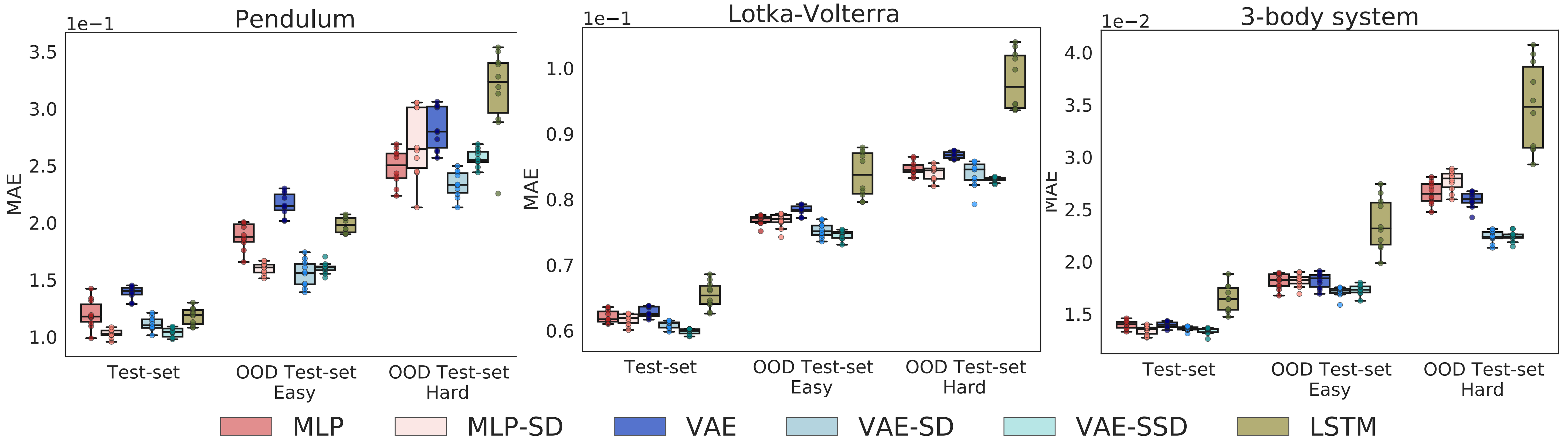
$$\mathcal{L}_{\phi,\theta}(\boldsymbol{x}) =$$

$$\mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x}^{(i)})}\left[\frac{1}{\gamma}\left\|\boldsymbol{x}^{(i)} - \boldsymbol{\mu}_x(\boldsymbol{z};\theta)\right\|_1\right] + d\log\gamma +$$

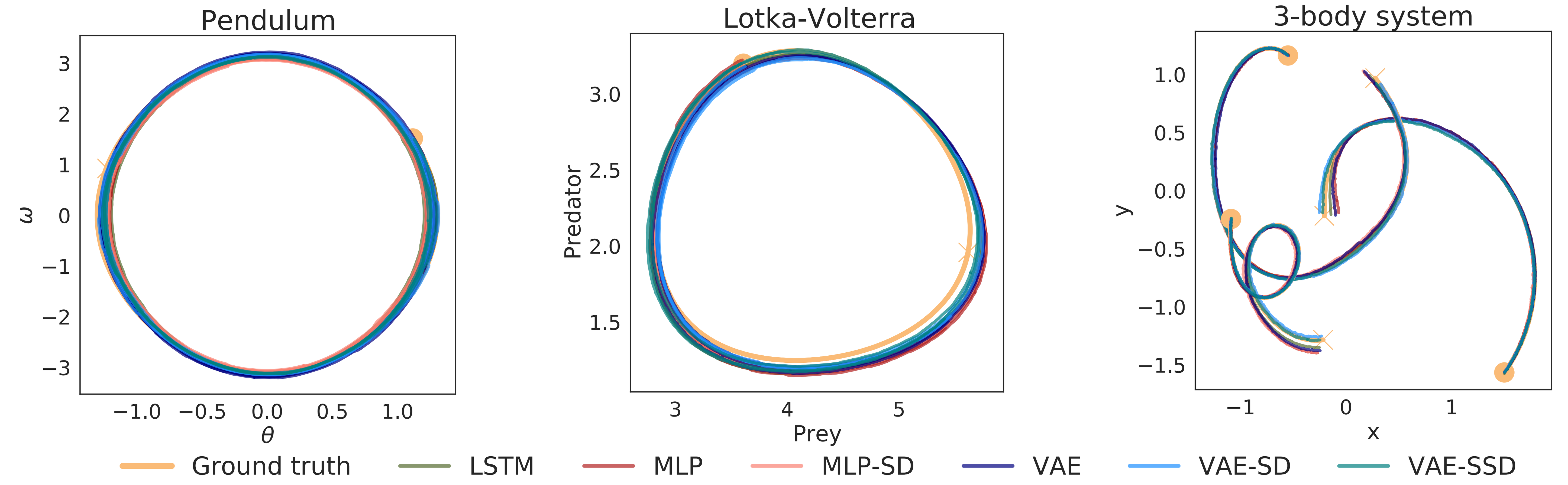
$$\left\|\boldsymbol{\sigma}_z\left(\boldsymbol{x}^{(i)};\phi\right)\right\|_2^2 - \log\left|\operatorname{diag}\left[\boldsymbol{\sigma}_z\left(\boldsymbol{x}^{(i)};\phi\right)\right]^2\right| + \left\|\boldsymbol{\mu}_z\left(\boldsymbol{x}^{(i)};\phi\right)\right\|_2^2 +$$

$$\delta\left\|\boldsymbol{\xi}^{(i)} - \mathcal{G}\left(\boldsymbol{\mu}_{z_{1:k}}\left(\boldsymbol{x}^{(i)};\phi\right)\right)\right\|_1$$

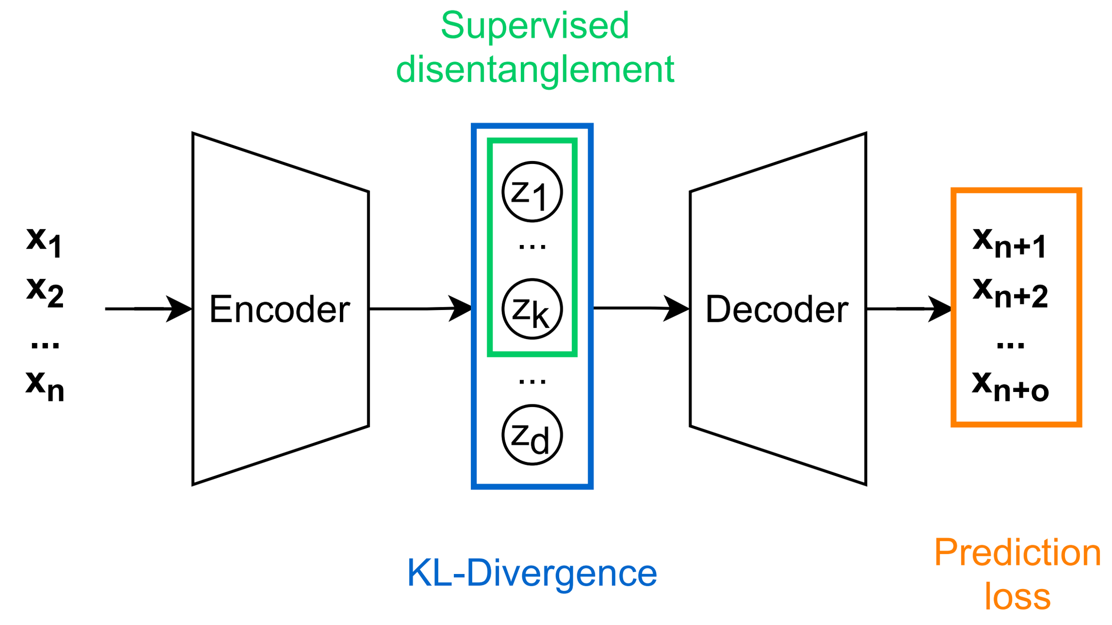
## RESULTS: MAE AFTER 200 TIMESTEPS




## SAMPLE TRAJECTORIES




## DISENTANGLED VAE MODEL SCHEMATIC




## CONCLUSIONS

Supervised disentanglement allows VAEs to efficiently extrapolate to ODE parameter spaces that were not present during training.

Disentanglement regularizes the encoder, enforces an implicit hierarchy in the latent space and enables conditioning of the decoder.

Disentanglement in deterministic autoencoders does not yield equally consistent improvements indicating that using the extra parameter information is not a straightforward task and requires further exploration for practical uses.

## REFERENCES


[1] Kingma, D. P. and Welling, M. Auto-encoding variational bayes. In International Conference on Learning Representations, ICLR, 2014

[2] Locatello, F., Tschannen, M., Bauer, S., Ratsch, G., Schoelkopf, B. and Bachem, O. Disentangling Factors of Variation Using Few Labels. 2019.

[3] Miladinovic, ., Gondal, M. W., Schokopf, B., Buhmann, J. M. and Bauer, S. Disentangled State Space Representations., 2019.