GLoss: Granger-Regularized Disentanglement

Joshua H. Sylvester, Eric C. Larson, and Mitchell A. Thornton
Darwin Deason Institute for Cybersecurity, Southern Methodist University, Dallas, TX, USA
{jsylvester, mitch, eclarson}@smu.edu

***Abstract*—Disentangling latent representations for multivariate time series remains challenging due to the inherently temporal nature of feature dependencies. Conventional disentanglement approaches such as KL divergence are not designed to account for temporal predictive structure and can be ill-suited for sequential data. We introduce the Granger-Regularized Autoencoder (GLoss-AE), which augments a standard autoencoder with a differentiable Granger-inspired penalty that discourages temporal predictive dependence among latent dimensions. Experiments on a simulated dataset with cluster-structured Granger-causal redundancy show that GLoss consistently reduces inter-latent temporal predictability and yields more compact latent representations, measured by fewer active latent dimensions, while maintaining reconstruction performance comparable to a vanilla autoencoder. These findings demonstrate that suppressing redundant temporal predictability provides a practical and interpretable pathway to time-series disentanglement.**

***Index Terms*—Disentangled learning, Time series analysis, Causality**

I. Introduction

Autoencoders learn compressed latent representations of data, and Variational Autoencoders (VAEs) extend this framework to promote disentangled representations capturing independent factors of variation. However, when applied to sequential data, regularization mechanisms such as KL divergence operate on a per-timestep or distributional basis and are not designed to account for temporal structure, which can exacerbate KL collapse [1]. An effective disentanglement strategy for sequential data should instead prioritize discouraging *temporal predictive dependencies* between latent features. Several VAE-based extensions address sequential disentanglement, including approaches that modify the ELBO to preserve an informative latent structure [1]. Most closely related to our work, Varando et al. [2] incorporate a Granger causality-based loss into an autoencoder to align latent representations with an external index, showing that incentivizing Granger causality between the latent space and an external signal yields more meaningful representations. In contrast, we aim to *discourage* Granger-style predictive dependence *within* the latent space.

Granger causality provides a natural framework for this: given two time series x and y, x Granger-causes y if its past improves prediction of y beyond an autoregressive baseline on y alone [3]. We propose GLoss, a differentiable Granger-inspired penalty that discourages temporal predictive dependence among latent dimensions, and incorporate it into a standard autoencoder to produce the *Granger-Regularized Autoencoder (GLoss-AE)*. We evaluate GLoss-AE against a vanilla reconstruction-only baseline on a simulated multivariate time series dataset with deliberately embedded Granger-causal redundancy, assessing whether GLoss suppresses inter-latent predictability while maintaining reconstruction fidelity.

II. Methodology

A. GLoss-AE: Granger-Regularized Autoencoder

Let $\mathbf{X} \in \mathbb{R}^{T \times K}$ be a length-T multivariate time series with K observed channels. An encoder network maps this input to a sequence of latent representations, $\mathbf{Z} = f_\theta(\mathbf{X}) \in \mathbb{R}^{T \times L}$, where L is the latent dimensionality. A decoder network then reconstructs the input as $\hat{\mathbf{X}} = g_\phi(\mathbf{Z})$. Using these definitions, we introduce the Granger-based loss (GLoss) objective in Equation (1). The scalar parameter $\lambda \geq 0$ controls the strength of a Granger-style regularization term, denoted $\mathcal{G}$, which penalizes temporal predictive dependence among latent features. The quantity $\mathcal{G}(\mathbf{Z})$ represents an aggregate measure of pairwise Granger-style predictive influence computed across all latent dimensions as formalized in Equation (2).

$$\mathcal{L}_{\text{GLoss-AE}}(\mathbf{X}) = \underbrace{\text{MSE}(\mathbf{X}, \hat{\mathbf{X}})}_{\text{Reconstruction}} + \lambda \underbrace{\mathcal{G}(\mathbf{Z})}_{\substack{\text{Granger-style} \\ \text{penalty}}} \tag{1}$$

$$\mathcal{G}(\mathbf{Z}) = \frac{1}{L(L-1)} \sum_{i=1}^{L} \sum_{\substack{j=1 \\ j \neq i}}^{L} g_{i \to j} \tag{2}$$

Each $g_{i \to j}$ is computed by comparing a restricted autoregressive model on $\mathbf{z}_j$ against an unrestricted model that additionally incorporates lagged values of $\mathbf{z}_i$, with both models fit via ridge-stabilized least-squares. To remain fully differentiable, we replace statistical hypothesis tests with a smooth, normalized measure of predictive improvement (Equation (3)) based on the residual sum of squares (RSS) of the restricted and unrestricted models.

$$g_{i \to j} \;=\; \sigma\left(\gamma \, \frac{\text{RSS}_R - \text{RSS}_U}{\text{RSS}_R + \epsilon}\right), \tag{3}$$

Minimizing $\mathcal{L}_{\text{GLoss-AE}}$ thus encourages the model to learn latent representations with reduced temporal predictive dependence between latent dimensions.

B. Dataset, Architecture, and Training

We construct a simulated dataset of multivariate time series consisting of $C = 3$ clusters of $S = 5$ series each ($K = 15$ total channels, $T = 1000$ timesteps), where each cluster is governed by a latent fundamental driver that induces strong intra-cluster Granger-causal dependencies while preserving

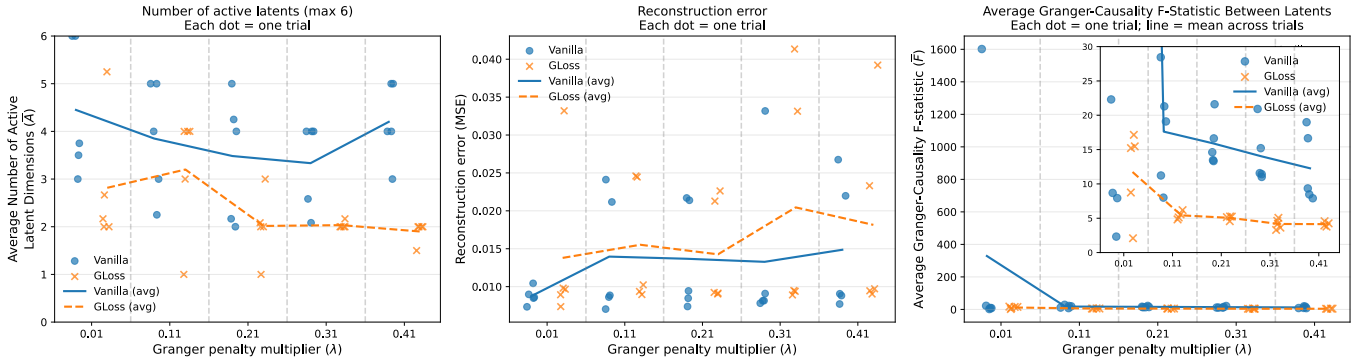

Fig. 1. Scatter plots summarizing experimental results across values of the Granger penalty multiplier λ. Panels show (a) the number of active latent dimensions, (b) reconstruction error measured by mean squared error (MSE), and (c) the average inter-latent Granger-causality F-statistic $\overline{F}$. Each point corresponds to an independent trial, and solid lines denote averages across trials. Lower values in (c) indicate weaker temporal predictive dependence among latent features.

inter-cluster independence. We employ a lightweight convolutional autoencoder with a single 1D convolutional layer (kernel size 12, ReLU) in both encoder and decoder, with $L = 6$ latent filters (double the number of underlying cluster drivers) creating an overcomplete latent space that allows GLoss to suppress redundant temporal structure beyond what the bottleneck alone imposes. Models are trained for up to 100 epochs with early stopping on five independently generated datasets of 120 realizations each (96/12/12 train/val/test split), with the GLoss penalty $\mathcal{G}$ computed at maximum lag $m = 20$ and sigmoid sharpness $\gamma = 3$.

III. Results

We systematically vary the Granger penalty multiplier $\lambda \in \{0.01, 0.11, 0.21, 0.31, 0.41\}$, performing five independent paired trials per value. For each pairing, a vanilla (MSE-only) and GLoss-regularized autoencoder are initialized with identical weights and trained on the same dataset. We evaluate three metrics: (i) reconstruction error (MSE), (ii) average number of active latent dimensions, and (iii) residual Granger-style predictive dependence among latent features, allowing us to assess whether GLoss suppresses inter-latent temporal predictability while maintaining reconstruction fidelity.

We compute the average number of active latent dimensions over the test set $\mathcal{X}_{\text{test}} = \{\mathbf{X}^{(i)}\}_{i=1}^{N}$ as

$$\overline{A} = \frac{1}{N} \sum_{i=1}^{N} \sum_{l=1}^{L} \left[\max_{t \in \{1,\dots,T\}} \left| z_{t,l}^{(i)} \right| > \tau \right], \qquad \tau = 10^{-4}. \quad (4)$$

To quantify residual inter-latent predictive dependence, we perform a Granger causality RSS F-test between each ordered latent pair $(\mathbf{z}_i, \mathbf{z}_j)$, selecting the maximum F-statistic across 20 lags and averaging across all test examples and off-diagonal pairs to produce a summary statistic $\overline{F}$, used here as a descriptive measure of residual predictability.

In Figure 1, we report results across all experiments. In Figure 1(a), the GLoss-AE consistently learns fewer active latent dimensions than the vanilla baseline across all penalty strengths. In Figure 1(b), the vanilla-AE achieves a slightly lower average reconstruction error than the GLoss-AE; however, reconstruction performance remains largely comparable between the two models, indicating that more compact latent representations can be obtained via GLoss regularization without a substantial loss in reconstruction accuracy.

Regarding residual inter-latent predictability in Figure 1(c), GLoss consistently yields lower $\overline{F}$ values than the vanilla baseline across all values of λ, indicating reduced temporal predictive dependence among latent features. Moreover, $\overline{F}$ exhibits substantially lower variability across trials under GLoss, suggesting more consistent convergence behavior.

IV. Conclusions

We proposed GLoss, a differentiable Granger-inspired penalty that discourages temporal predictive dependence among latent dimensions, and demonstrated that incorporating it into a standard autoencoder yields more compact latent representations with reduced inter-latent predictability without incurring a substantial loss in reconstruction accuracy. These findings suggest that suppressing redundant temporal predictability via GLoss is an effective approach for promoting disentanglement in time series representations. Future work will explore GLoss in variational settings and direct comparison against KL-based disentanglement objectives.

References

[1] Y. Li, Z. Chen, D. Zha, M. Du, J. Ni, D. Zhang, H. Chen, and X. Hu, "Towards Learning Disentangled Representations for Time Series," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ser. KDD '22. New York, NY, USA: Association for Computing Machinery, Aug. 2022, pp. 3270–3278. [Online]. Available: https://dl.acm.org/doi/10.1145/3534678.3539140

[2] G. Varando, M.- Fernández-Torres, and G. Camps-Valls, "Learning Granger Causal Feature Representations," in *Climate Change AI*. Climate Change AI, Jul. 2021. [Online]. Available: https://www.climatechange.ai/papers/icml2021/34

[3] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969, publisher: [Wiley, Econometric Society]. [Online]. Available: https://www.jstor.org/stable/1912791