Jacobian Regularizer-based Neural Granger Causality

Wanqi Zhou, Shuanghao Bai, Shujian Yu, Qibin Zhao, Badong Chen*

{zwq785915792, baishuanghao}@stu.xjtu.edu.cn; yusj9011@gmail.com;
qibin.zhao@riken.jp; chenbd@mail.xjtu,edu.cn

Background

Motivation and Challenge

- How to propose a new framework without designing separate predictive models like previous neural approaches?
- How to design good causal representation and causal constraints for causal discovery when using a single nerual network (NN) with shared hidden layer?

Key Contributions

- To our best knowledge, this is the first work to harness a **single NN model** with shared hidden layers for multivariate Granger causality analysis.
- We propose a novel neural network framework to learn Granger causality by incorporating an **input-output Jacobian regularizer** in the training objective.
- Our method can not only obtain the summary Granger causality but also the full-time Granger causality.
- Extensive experiments show our method can outperform state-of-the-art baselines.

Preliminaries

Jacobian Regularizer

Definition We regularize the L_1 ***norm*** or squared ***Frobienus norm*** of the input-output Jacobian matrix:

$$\mathbf{J} = \left[\frac{\partial \mathbf{f}}{\partial x_1^{t-\tau}} \cdots \frac{\partial \mathbf{f}}{\partial x_D^{t-1}} \right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1^{t-\tau}} \cdots \frac{\partial f_1}{\partial x_1^{t-1}} \cdots \frac{\partial f_1}{\partial x_D^{t-\tau}} \cdots \frac{\partial f_1}{\partial x_D^{t-1}} \\ \vdots \\ \frac{\partial f_D}{\partial x_1^{t-\tau}} \cdots \frac{\partial f_D}{\partial x_1^{t-1}} \cdots \frac{\partial f_D}{\partial x_D^{t-\tau}} \cdots \frac{\partial f_D}{\partial x_D^{t-1}} \end{bmatrix},$$

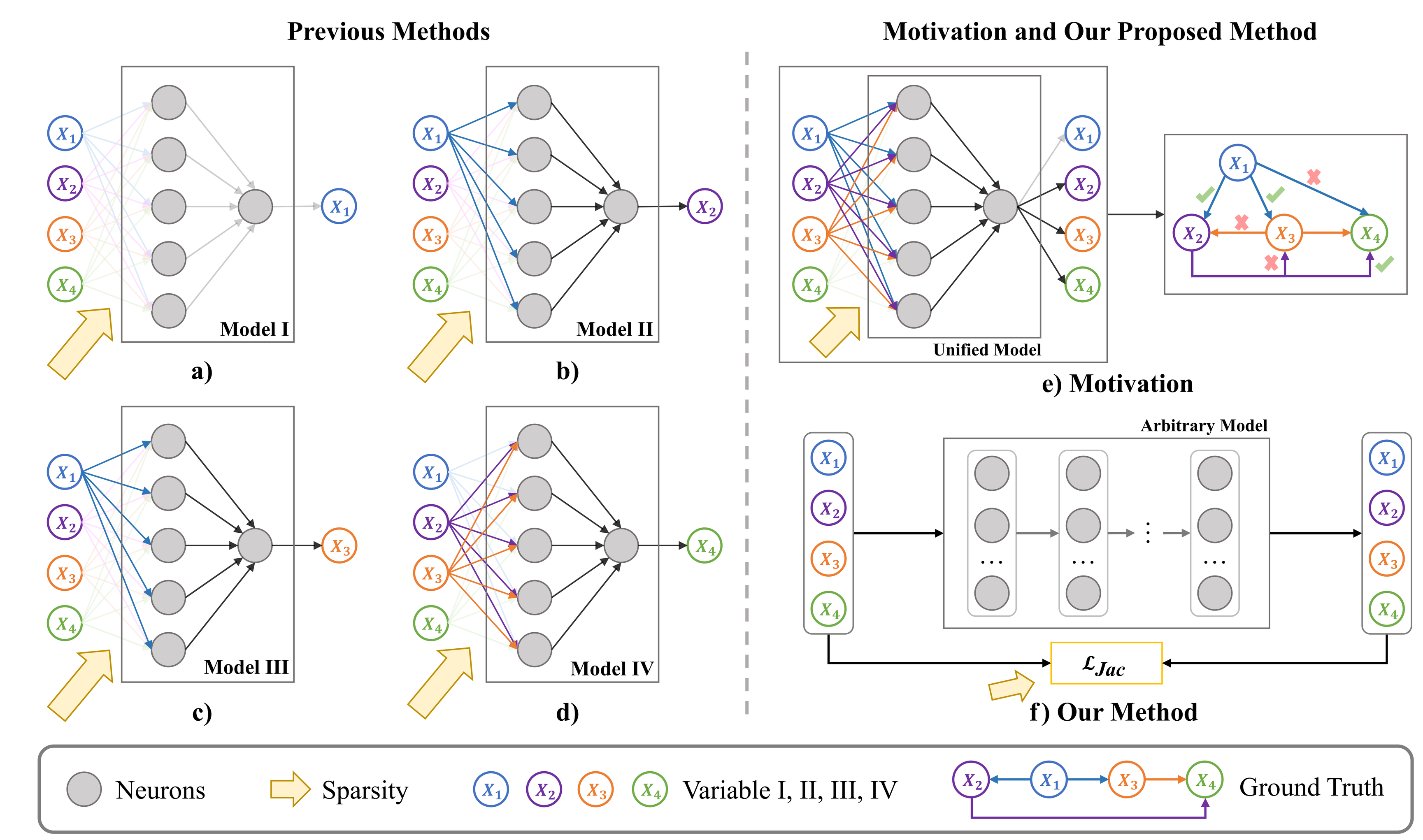
$$\|\mathbf{J}\|_1 = \sum_{i,j} \left| \frac{\partial f_i}{\partial x_j} \right|$$

L_1 norm

$$\|J(\boldsymbol{x})\|_F^2 = \mathrm{Tr}(JJ^\mathrm{T}) = \sum_{\{e\}} eJJ^\mathrm{T}e^\mathrm{T} = \sum_{\{e\}} \left[\frac{\partial (e \cdot z)}{\partial x} \right]^2,$$

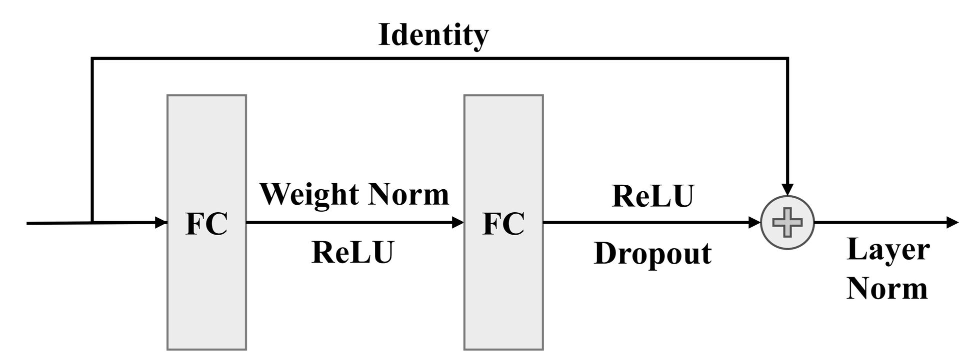
Frobienus norm

Approach



Construct a Time Series Forecasting Neural Network

Residual MLP-based Model



Incorporate Input-output Jacobian Matrix Regularizer During Training

Loss Function

$$\|J(\boldsymbol{x})\|_F^2 \approx \frac{1}{n_{\text{proj}}} \sum_{\mu=1}^{n_{\text{proj}}} \left[\frac{\partial (\hat{\boldsymbol{v}}^\mu \cdot \boldsymbol{z})}{\partial \boldsymbol{x}} \right]^2,$$

$\hat{\boldsymbol{v}}^\mu$ *random vector*

S^{D-1} *unit sphere*

n_{proj} *random projection*

$$\min_{\mathbf{W}} \sum_{t=\tau}^{T} (x_t - f(\mathbf{x}_{<t}))^2 + \lambda \|J(\mathbf{x}_{<t})\|_1,$$

JRNGC_L1

$$\min_{\mathbf{W}} \sum_{t=\tau}^{T} (x_t - f(\mathbf{x}_{<t}))^2 + \lambda \|J(\mathbf{x}_{<t})\|_F^2,$$

JRNGC_F

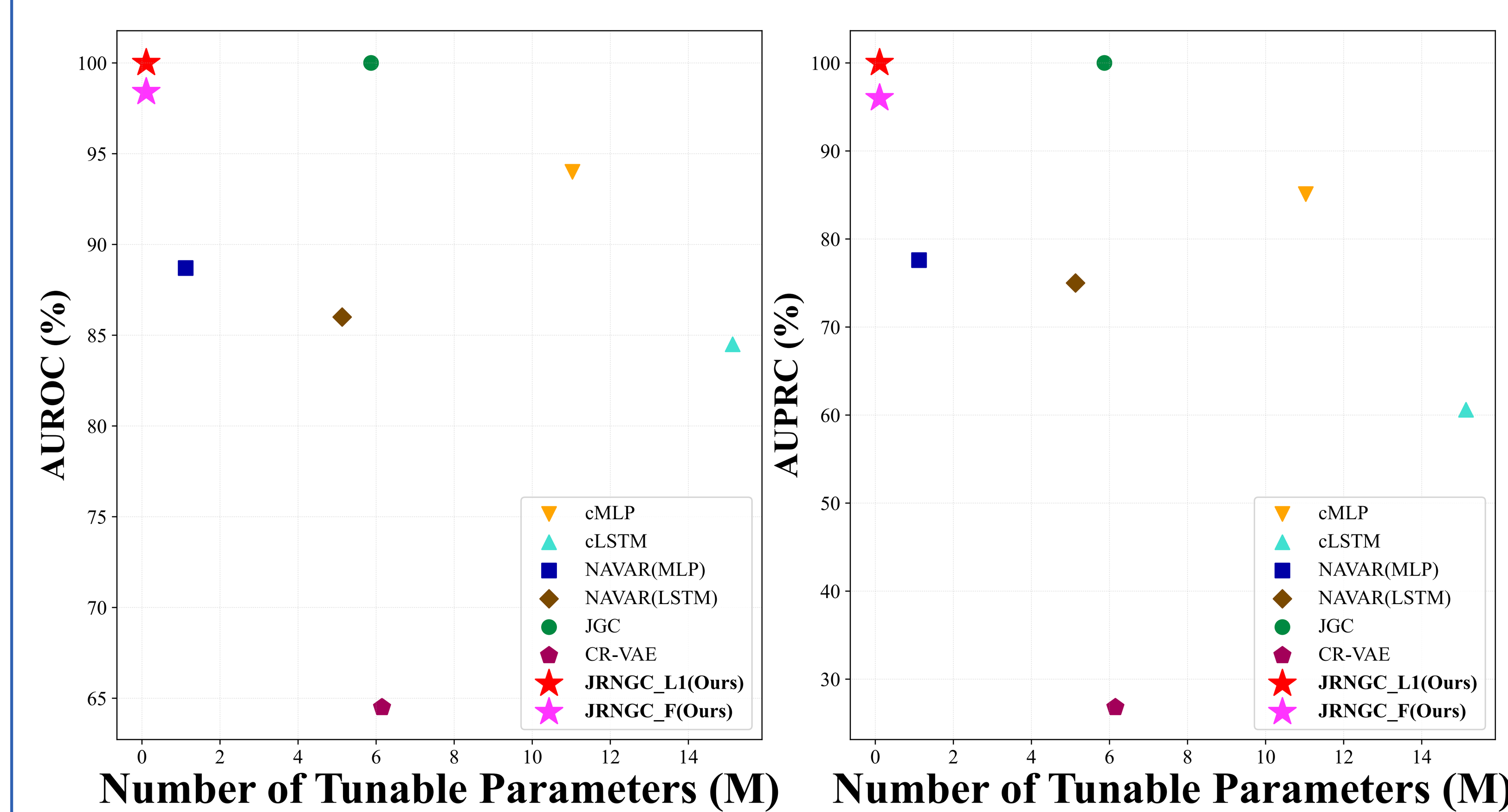
Let W be the set of parameters to be optimized for the neural network f

Experimental Results

Extensive experiments show that our JRNGC method achieves SOTA. More details can be found in https://arxiv.org/pdf/2405.08779

Table 5. Comparative performance on CausalTime benchmark datasets. We highlight the best and the second best in bold and with underlining, respectively.

Methods	AUROC			AUPRC		
	AQI	Traffic	Medical	AQI	Traffic	Medical
GC	0.4538± 0.0377	0.4191± 0.0310	0.5737± 0.0338	0.6347± 0.0158	0.2789± 0.0018	0.4213± 0.0281
SVAR	0.6225± 0.0406	0.6329± 0.0047	0.7130± 0.0188	0.7903± 0.0175	0.5845± 0.0021	0.6774± 0.0358
N.NTS	0.5729± 0.0229	<u>0.6329± 0.0335</u>	0.5019± 0.0682	0.7100± 0.0228	0.5770± 0.0542	0.4567± 0.0162
PCMCI	0.5272± 0.0744	0.5422± 0.0737	0.6991± 0.0111	0.6734± 0.0372	0.3474± 0.0581	0.5082± 0.0177
Rhino	0.6700± 0.0983	0.6274± 0.0185	0.6520± 0.0212	0.7593± 0.0755	0.3772± 0.0093	0.4897± 0.0321
CUTS	0.6013±0.0038	0.6238±0.0179	0.3739±0.0297	0.5096±0.0362	0.1525±0.0226	0.1537±0.0039
CUTS+	<u>0.8928±0.0213</u>	0.6175±0.0752	**0.8202 ±0.0173**	<u>0.7983±0.0875</u>	**0.6367 ±0.1197**	0.5481±0.1349
NGC	0.7172±0.0076	0.6032±0.0056	0.5744±0.0096	0.7177±0.0069	0.3583±0.0495	0.4637±0.0121
NGM	0.6728±0.0164	0.4660±0.0144	0.5551±0.0154	0.4786±0.0196	0.2826±0.0098	0.4697±0.0166
LCCM	0.8565±0.0653	0.5545±0.0254	<u>0.8013±0.0218</u>	**0.9260 ±0.0246**	0.5907±0.0475	**0.7554 ±0.0235**
eSRU	0.8229±0.0317	0.5987±0.0192	0.7559±0.0365	0.7223±0.0317	0.4886±0.0338	<u>0.7352±0.0600</u>
SCGL	0.4915±0.0476	0.5927±0.0553	0.5019±0.0224	0.3584±0.0281	0.4544±0.0315	0.4833±0.0185
TCDF	0.4148±0.0207	0.5029±0.0041	0.6329±0.0384	0.6527±0.0087	0.3637±0.0048	0.5544±0.0313
JRNGC-F (ours)	**0.9279 ±0.0011**	**0.7294 ±0.0046**	0.7540±0.0040	0.7828±0.0020	<u>0.5940±0.0067</u>	0.7261±0.0016



Scan the QR code for contacting us:

WeChat of the Author   Code   Paper