GAR: Generalized Autoregression for Multi-Fidelity Fusion

Yuxin Wang#, Zheng Xing#
Wei. W. Xing*

Abstract

Due to the simplicity, robustness, accuracy, and tractability of classic autoregression (AR), we proposed generalized autoregression (GAR) by using tensor formulation and latent features. Furthermore, we prove the autokrigeability theorem based on *GAR* in the multi-fidelity case and develop *CIGAR*, a simplified GAR with the exact predictive mean accuracy with and much faster computation. Our main contributions are listed as below:

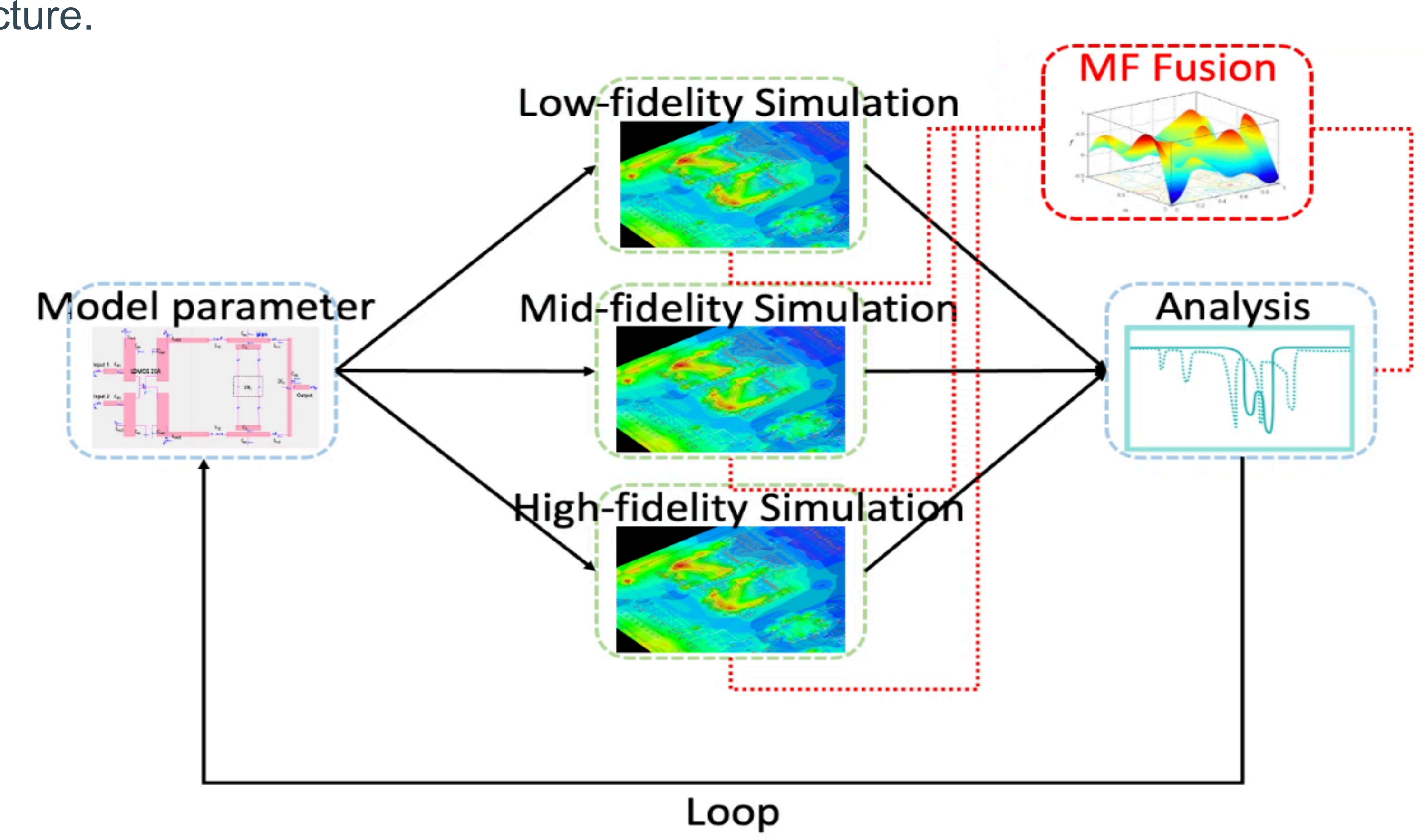
- Generalization of AR to GAR for arbitrary non-structured high-dimensional outputs.
- Generalization of AR and the proposed GAR for non-subset problems with a closed-form solution
- Revealing the autokrigeability [1] for the multi-fidelity fusion within an AR structure and deriving conditional independent GAR (CIGAR) with the exact accuracy in posterior mean predictions.

Introduction

In many scientific research and engineering applications, where repeated simulations of complex systems are conducted. To reduce the expensive cost of generating training examples, it has become a promising approach to combine the results of low-fidelity (fast but inaccurate) and high-fidelity (slow but accurate) simulations. Despite the success of many state-of-the-art (SOTA) approaches, they normally assume that,

- the output dimension is the same and aligned across all fidelities;
- the high-fidelity samples' corresponding inputs form a subset of the low-fidelity inputs;
- the output dimension is small, which is not practical for scientific computing where the dimension can be 1 million (for a 100*100*100 spatial-temporal field).

To resolve these challenges, instead of proposing another ad-hoc model with pre-processing and simplification we propose generalized autoregression (GAR), which can deal with arbitrary high-dimensional problems without the subset multi-fidelity data structure.



Circuit design optimisation as an example

Contributions

1. Vector Output AR

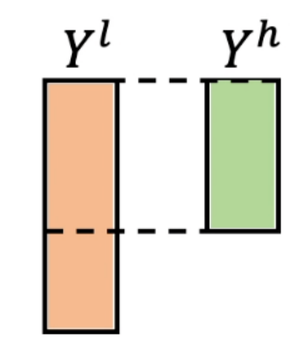
$$f^h(\mathbf{x}) = \rho f^l(\mathbf{x}) + f^r(\mathbf{x}),$$

$$\mathbf{F}^h(\mathbf{x}) = \mathbf{F}^l(\mathbf{x}) \times_1 \mathbf{W}_1, \dots, \times_M \mathbf{W}_M + \mathbf{F}^r(\mathbf{x}),$$

Tensor GP

Tensor matrix product

Tensor GP

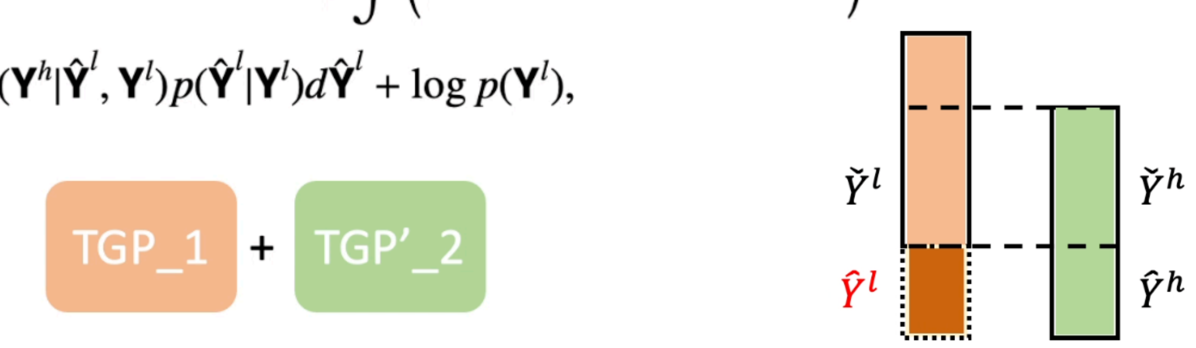


In our paper, we prove that after generalising the model to the tensor gaussian process, it can also be decomposed as the low-fidelity model and residual model.

2. Non-subset Decomposition

$$\log p(\mathbf{Y}^l, \mathbf{Y}^h) = \log \int p(\mathbf{Y}^l, \mathbf{Y}^h, \hat{\mathbf{Y}}^l) d\hat{\mathbf{Y}}^l = \log \int \left(p(\mathbf{Y}^h | \hat{\mathbf{Y}}^l, \mathbf{Y}^l) p(\hat{\mathbf{Y}}^l | \mathbf{Y}^l) p(\mathbf{Y}^l) \right) d\hat{\mathbf{Y}}^l$$

$$= \log \int p(\mathbf{Y}^h | \hat{\mathbf{Y}}^l, \mathbf{Y}^l) p(\hat{\mathbf{Y}}^l | \mathbf{Y}^l) d\hat{\mathbf{Y}}^l + \log p(\mathbf{Y}^l),$$



The restriction of subset structure is releasable, since we prove that the integral has closed-form solution.

3. Autokrigeability of AR and CIGAR

Autokrigeability[1] also holds in AR

$$\mathbf{S}^h_m = \mathbf{I}, \mathbf{S}^l_m = \mathbf{I}$$

$$\mathbf{Z}^l(\mathbf{x}, \mathbf{x}') \sim \mathcal{TGP}\left(\mathbf{0}, k^l(\mathbf{x}, \mathbf{x}'), \mathbf{S}^l_1, \dots, \mathbf{S}^l_M\right), \mathbf{Z}^r(\mathbf{x}, \mathbf{x}') \sim \mathcal{TGP}\left(\mathbf{0}, k^r(\mathbf{x}, \mathbf{x}'), \mathbf{S}^r_1, \dots, \mathbf{S}^r_M\right),$$

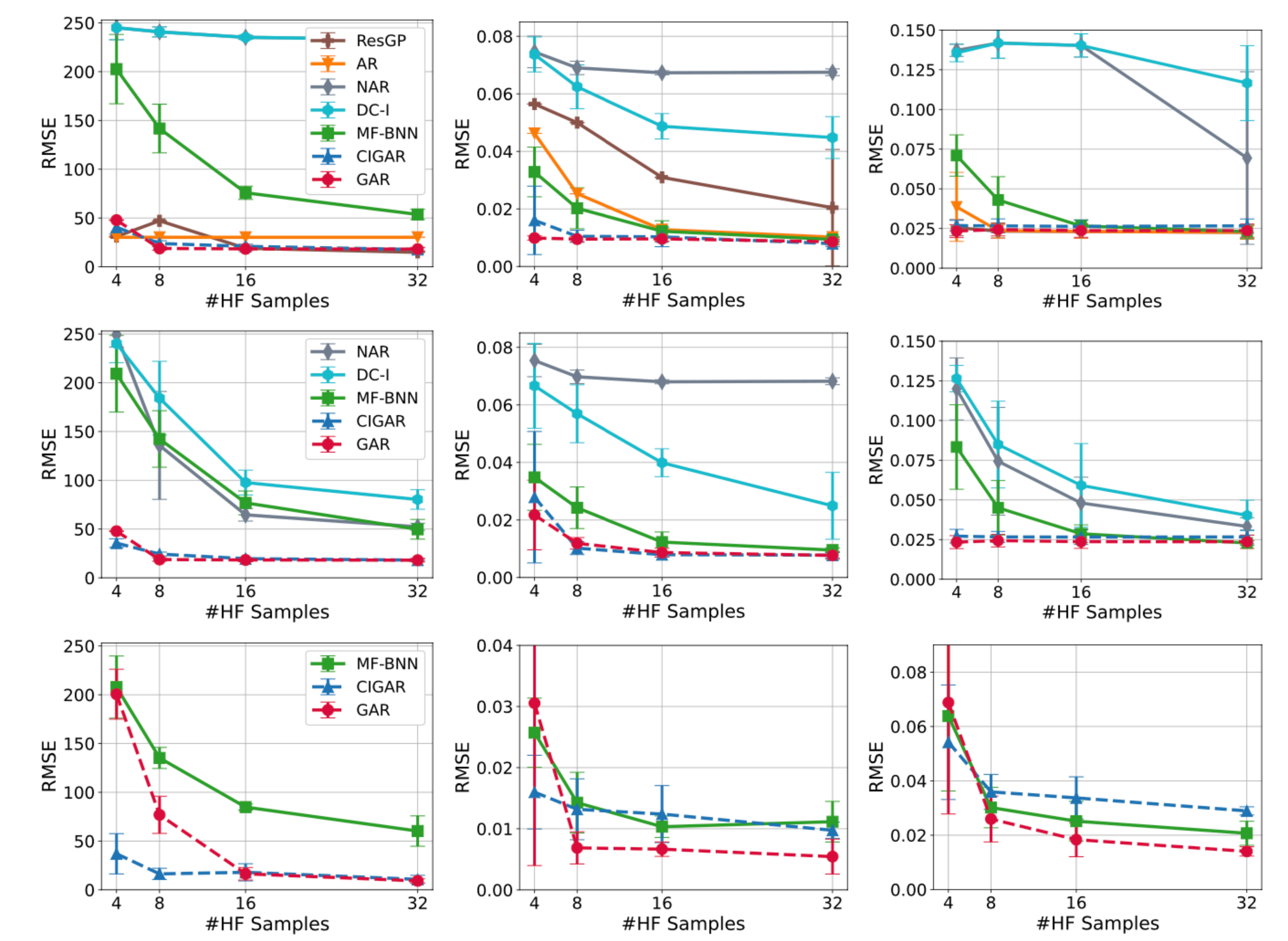
Conditional Independent GAR

$$\mathbf{W}_m^T \mathbf{W}_m = \mathbf{I}$$

$$\mathcal{O}\left(\textstyle\sum_i \sum_{m=1}^M (d^i_m)^3 + (N^i)^3\right) \Rightarrow \mathcal{O}\left(\textstyle\sum_i (N^i)^3\right)$$

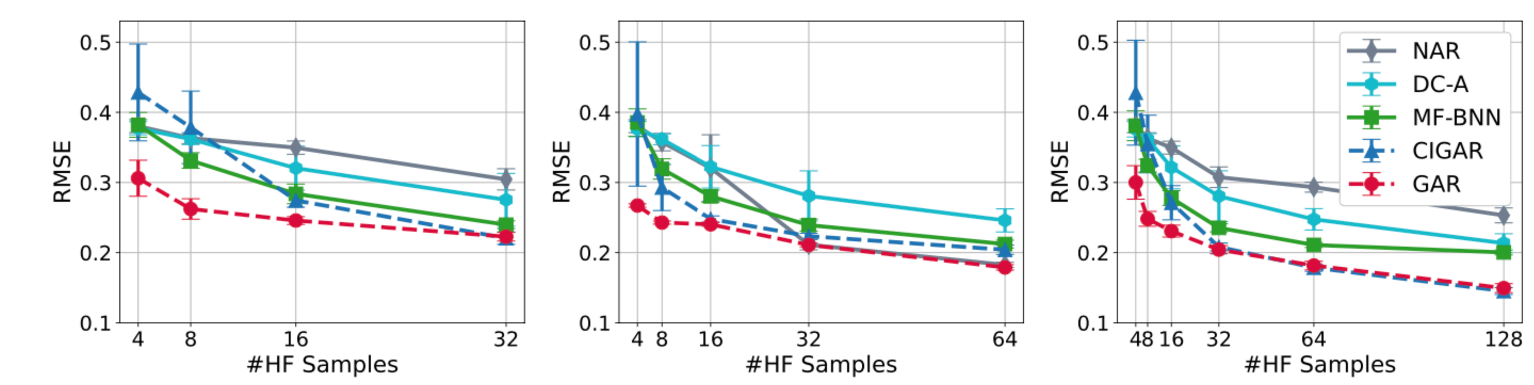
Since the predictive mean function does not depend on any output covariance matrixes, which reassemble the autokrigeability (no knowledge transfer in noiseless cases for mean predictions) based on the GAR framework. We can introduce a conditional independent output correlation and orthogonal weight matrices to reduce the computational complexity from $O((N^i d^i)^3)$ further down to $O((N^i)^3)$.

Empirical Results



(a) Poisson's equation (b) Burger's (c) Heat

RMSE against an increasing number of high-fidelity training samples for Poisson's, Burger's, and heat equations with aligned outputs (top row), non-aligned outputs (middle row), and non-subset data (bottom row).



RMSE for Optimal topology structure dataset with low-fidelity training sample number fixed to {32,64,128}

More Details

Code: https://github.com/zen-xingle/ML_gp
Email: wayne.xingle@gmail.com
WYXtt_2011@163.com
Zheng.xing@rock-chips.com

References

Mauricio A. Alvarez, Lorenzo Rosasco, and Neil D. Lawrence. Kernels for Vector-Valued Functions: A Review.

Shandian Zhe, Wei Xing, and Robert M. Kirby. Scalable High-Order Gaussian Process Regression. In The 22nd International Conference on Artificial Intelligence and Statistics, pages 2611–2620. PMLR.

Wackernagel, Hans. Multivariate geostatistics: an introduction with applications. Springer Sciences and Business Media, 2003