



**HAL**  
open science

## Generatives models approach on TFM data based on variational inference for multi-fidelity data generation

Gerardo Granados, Giorgia Colombera, Filippo Gatti, Roberto Miorelli,  
Didier Clouteau

### ► To cite this version:

Gerardo Granados, Giorgia Colombera, Filippo Gatti, Roberto Miorelli, Didier Clouteau. Generatives models approach on TFM data based on variational inference for multi-fidelity data generation. IUTAM Symposium, Oct 2022, Paris, France. cea-04410156

**HAL Id: cea-04410156**

**<https://cea.hal.science/cea-04410156>**

Submitted on 22 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Generatives models approach on TFM data based on variational inference for multi-fidelity data generation

Gerardo E. Granados<sup>\*#</sup>, Giorgia Colombera<sup>#</sup>, Filippo Gatti<sup>#</sup>, Roberto Miorelli<sup>\*</sup>, Didier Clouteau<sup>#</sup>

<sup>\*</sup> Université Paris-Saclay,  
CEA, List,  
F-91120, Palaiseau, France  
gerardo.granados@cea.fr  
roberto.miorelli@cea.fr

<sup>#</sup> Université Paris-Saclay  
Laboratoire de Mécanique Paris-Saclay  
CentraleSupélec, ENS Paris-Saclay, CNRS  
Gif-sur-Yvette, France  
giorgia.colombera@centralesupelec.fr  
filippo.gatti@centralesupelec.fr  
didier.clouteau@centralesupelec.fr

### Abstract

Fast and reliable numerical models are commonly used in nondestructive testing and evaluation (NDT&E) domain in order to design inspection procedure, design probes, interpret acquisitions, etc. In ultrasound testing (UT) full matrix capture (FMC) acquisitions are often post-processed via the total focus method (TFM) imaging technique to facilitate detection, classification and characterization tasks based on the analysis of amplitude signal maps (i.e., mono-channel “images”). In this context, the use of simulated data for training machine learning (ML) algorithm to be applied on experimental data has shown to be a promising schema for NDT&E applications [2] in view of its possible application to help to decision and automatic diagnostic tools.

Despite the very good accuracy obtained by the most advanced numerical simulation tools, the impact of measure uncertainties and experimental noise remain the biggest hundreds to overcome in order to unleash the prediction power of the most advanced deep learning algorithms. That is, the use of realistic (i.e., as closer as possible to experimental measures) simulated data is of paramount importance to develop accurate suitable for this task since the data samples of a given experiment are virtually unlimited.

The presented approach aims at developing deep-learning meta-model to generate realistic NDT&E results. A conditional generative adversarial network (c-GAN) is trained on a TFM database, composed of experimental and simulated images respectively. A reduced-order latent manifold  $z$  is inferred via an auto-encoder architecture, trained in an adversarial way. An encoder  $F$  is trained to link the latent manifold space  $\mathcal{Z}$  with the  $\mathcal{X}$  space (e.g: TFM image data-set), as shown on equation (1).  $G$  generates a  $(c, z_g)$  conditional sample  $x_g$ , where  $c$  is a class (simulation or experimental TFM image), and  $z_g$  contains all the information of  $x$  projected on  $\mathcal{Z}$ . Additionally, the discriminator  $D_x(x, x_g)$  is a neural network who learns the differences between real  $x$  and fake images  $x_g$ . Likewise,  $D_z$  judges if  $z_g = F(x)$  and  $z$  are equally distributed. The generator uses the scores form  $D_x$  as a loss to learn to generate realistic images, while the encoder use  $D_z$  to learn to generate  $z_g$  from  $x$ . Additional consideration on the losses are done to ensure bijective  $F$  and  $G$ .

$$z_g = F(x) \quad X_g = G(z, c) \quad (1)$$

Where,

$$x \in \mathcal{X}; \mathcal{X} = \mathbb{R}^D \quad z \in \mathcal{Z}; \mathcal{Z} = \mathbb{R}^d \quad c \in \mathcal{C}; \mathcal{C} = \mathbb{R}^C \quad \text{with } D \gg d \quad \text{and } C = 2$$

$$z \in \mathcal{Z}; \mathcal{Z} = \mathbb{R}^d \quad \text{with } z \sim \mathcal{N}(\mathbf{I}, \mathbf{0})$$

The ML model learns the differences between simulation and experimental data from TFM images in a non-supervised frame. The objective is to generate hybrid data spanning different degrees of fidelity, up to the highest one corresponding to the experimental samples.

The regression on the reduced space conditionally generates samples with different degrees of fidelity. In our network, the hybrid generation starts with a simulated TFM input image as content to then add the realistic style, similarly to [1][3]. A consistent addition of noise and variations on the hybrid data is expected.

The possibility to create unlimited data is now extended to realistic hybrid data, which is indistinguishable from reality since it contains the style of the experimental data set. Furthermore, the inferred latent space contains the information of the parameters (labels) of the simulated data-set. This space is explored to condition the hybrid data-set generation when creating new instances of TFM realistic images labeled by the corresponding parameters.

The present work evaluates the translation from a numerical input into a higher fidelity output by transferring experimental style to simulated data. The hybrid and experimental data similarities are measured to validate the newly generated images. Future work aims to use the hybrid data to enhance the data generation to be used for training classification and regression models in view of decision support systems.

## References

- [1] Donahue, J.; Simonyan, K.: Large scale adversarial representation learning in Advances in Neural Information Processing Systems 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), article No.: 946, Pages 10542–10552.
- [2] Miorelli, R.; Kulakovskiy, A. ; Chapuis, B. C; D’Almeida, O. ; Mesnil, O.: Supervised learning strategy for classification and regression tasks applied to aeronautical structural health monitoring problems Ultrasonics vol. 113, p. 106372, May 2021, doi: 10.1016/j.ultras.2021.106372.
- [3] Zhou, Y.;Gu, K.;Huang, T.: Unsupervised Representation Adversarial Learning Network: From Reconstruction to Generation Proceedings of the International Joint Conference on Neural Networks, 2019, vol. 2019-July, doi: 10.1109/IJCNN.2019.8852395.