Research Article

Data-Driven, Parameterized Reduced-order Models for Predicting Distortion in Metal 3D Printing

Indu Kant Deo[1], Youngsoo Choi[2], Saad A. Khairallah[2], Alexandre Reikher[2], Maria Strantza[2]

1. Department of Mechanical Engineering, University of British Columbia, Canada; 2. Lawrence Livermore National Laboratory, Livermore, United States

In Laser Powder Bed Fusion (LPBF), the applied laser energy produces high thermal gradients that lead to unacceptable final part distortion. Accurate distortion prediction is essential for optimizing the 3D printing process and manufacturing a part that meets geometric accuracy requirements. This study introduces data-driven parameterized reduced-order models (ROMs) to predict distortion in LPBF across various machine process settings. We propose a ROM framework that combines Proper Orthogonal Decomposition (POD) with Gaussian Process Regression (GPR) and compare its performance against a deep-learning based parameterized graph convolutional autoencoder (GCA). The POD-GPR model demonstrates high accuracy, predicting distortions within $\pm 0.001\mathrm{mm}$, and delivers a computational speed-up of approximately 1800x.

1. Introduction

LPBF is a popular metal additive manufacturing technique that has gained significant attention in recent years due to its ability to fabricate complex geometries with high precision. In LPBF, a thin layer of metal powder is deposited on a build platform, and a high-energy laser selectively melts and fuses the powder particles together to form a solid layer. This process is repeated layer by layer, with each 2D layer fusing to the previous one, ultimately constructing fully dense 3D components[1].

The repeated melting and solidification cycles in LPBF lead to significant thermal gradients, resulting in notable distortion in the as-built part. This distortion can compromise the dimensional accuracy and structural integrity of the final component, which is a critical requirement in many applications[2]. To address this challenge, currently, the additive manufacturing community relies on

 doi.org/10.32388/OS5U6O

a trial-and-error method, which involves conducting numerous experiments or simulations that are time-consuming and expensive. The approach is a distortion compensation technique, which involves pre-distorting the part design in such a way that upon printing, the final built shape matches the intended geometry[3]. The problem is predicting quickly the pre-distorted part geometry is not easy.

The amount of distortion for a given geometry depends on various machine settings, such as scan speed, laser power, and dwell time[4]. Accurately predicting the distortion based on these parameters is crucial for effective distortion compensation and process optimization[5]. High-fidelity finite element models offer a cost-effective alternative, enabling repeated trials without the need to physically build the parts[6]. However, developing accurate physics-based models for LPBF distortion is a complex task due to the non-linear dependence of distortion on the process and complex part geometries: one pre-distortion solution that fits all problems does not exist[7].

This study aims to develop parameterized data-driven reduced-order models (ROMs) for accurately predicting distortion in the LPBF process under various machine settings[8][9][10]. Specifically, it employs a combination of POD and Gaussian Process Regression to create a ROM, which is then compared with a parameterized graph convolutional autoencoder for distortion prediction. The POD-GPR ROM achieves a distortion prediction accuracy within ± 0.001 mm, and offers a computational speed-up over the high fidelity model of nearly 1800 times. This significant improvement highlights the model's potential for enabling rapid and precise distortion predictions, which is critical for optimizing LPBF processes. The ability to efficiently predict distortion not only reduces the reliance on costly and time-consuming experimental trials but also enhances the overall process control, making this approach highly valuable for industrial applications[11].

2. Methods

2.1. LPBF simulation data

In this study, we analyze data generated from Laser Powder Bed Fusion (LPBF) simulations conducted using ANSYS®Additive Suite. The dataset is parameterized based on the dwell time dt, which represents the time interval required for the laser or heat source to revisit a specific location to deposit a subsequent material layer[12]. The impact of this interlayer dwell time becomes more pronounced as structures grow in size and complexity, significantly affecting the thermal history and geometric

distortion outcomes. Specifically, we performed simulations on a cylindrical geometry with dwell times $dt \in [20, 80]s$, sampled at intervals of $5s$. The computational mesh used in these simulations comprised $N_h = 77,151$ nodes, and each simulation covered 34 layers of metal deposition, yielding $N_t = 34$ time steps per simulation. On average, each simulation required approximately 2 hours of computation on 112 cores of an Intel(R) Xeon(R) CLX-8276L processor. From this extensive dataset, we selected N_μ samples for training, with the remaining samples designated for validation and testing. Let $\mu^{(i)} = (dt^{(i)})$ represent the i-th parameter in the training set. The primary quantity of interest is the final distortion field, $\mathbf{u}$. We denote the snapshot vector of distortion data at time-step n for parameter $\mu^{(i)}$ as $\mathbf{u}_n^{(i)} \in \mathbb{R}^{N_h}$, and the corresponding data matrix for all nodes and time steps as $\mathbf{U}^{(i)} = [\mathbf{u}_0^{(i)}, \ldots, \mathbf{u}_{N_t}^{(i)}] \in \mathbb{R}^{N_h \times N_t}$. By aggregating all these matrices, we form a third-order tensor $\mathbf{U} = [\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N_\mu)}] \in \mathbb{R}^{N_\mu \times N_h \times N_t}$, which constitutes the training dataset utilized in this work. A visual representation of this dataset is provided in Figure 1.

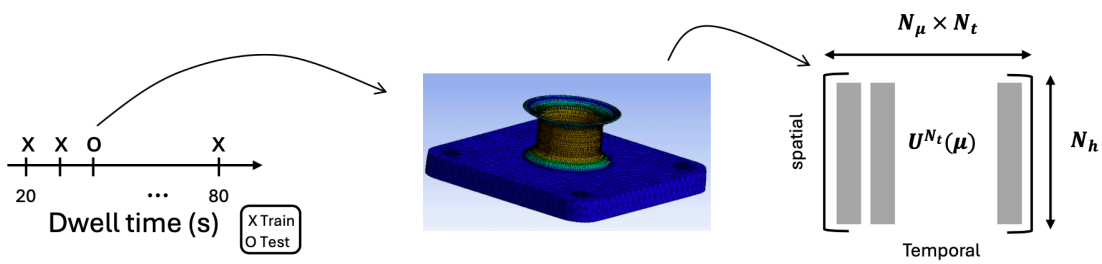

Figure 1. Schematic of dataset generation. LPBF simulations were generated for parameter dwell time. Distortion data was extracted for each simulation and arranged into a training snapshot matrix.

2.2. *POD-GPR*

POD-GPR is a parameterized, data-driven reduced order modeling method which consists of two main features: 1) POD[13] to learn a linear spatial compression of data to a latent space; 2) GPR[14] to map the POD coefficients of final distortion layer to a given parameter value.

POD is a dimensionality reduction technique widely used in the analysis of complex systems to identify dominant patterns or modes. Given a set of snapshots $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_{N_t}] - \mathbf{U}_{ref} \in \mathbb{R}^{N_h \times N_t}$, where each snapshot $\mathbf{U}_i \in \mathbb{R}^{N_h}$ represents the state of the system at a particular time or parameter setting, the goal of POD is to find a set of orthonormal basis vectors $\{\phi_k\}_{k=1}^{r}$ that capture the most energetic features of the data. This is achieved by solving the

eigenvalue problem associated with the covariance matrix $\mathbf{C} = \mathbf{U}\mathbf{U}^T \in \mathbb{R}^{N_h \times N_h}$. The eigenvectors corresponding to the largest eigenvalues provide the POD modes, which minimize the reconstruction error in a least-squares sense. Mathematically, this can be expressed as:

$$\min_{\phi_k} \sum_{i=1}^{N_t} \left\| \mathbf{U}_i - \sum_{k=1}^{r} \langle \mathbf{U}_i, \phi_k \rangle \phi_k \right\|^2,$$

where $\langle \mathbf{U}_i, \phi_k \rangle$ denotes the projection of $\mathbf{U}_i$ onto the k-th POD mode. The reduced representation of the data is then given by the projection coefficients $\mathbf{a}_i = [\langle \mathbf{U}_i, \phi_1 \rangle, \ldots, \langle \mathbf{U}_i, \phi_r \rangle]^T$, enabling efficient analysis and computation in a lower-dimensional subspace. We select the first r modes for which the total energy is greater than 99.99%. The energy stored in the first r modes is given by:

$$E_r = \frac{\sum_{j=1}^{r} \sigma_j^2}{\sum_{j=1}^{N} \sigma_j^2},$$

where σ is the eigenvalue.

GPR is a powerful non-parametric method for modeling complex functions, particularly in cases where uncertainty quantification is important. In this work, we employ GPR to learn r independent Gaussian process regressions, each mapping the final time POD coefficients to their corresponding parameter μ. Let $\mathbf{a}(\mu) = [a_1(\mu), \ldots, a_r(\mu)]^T$ denote the vector of POD coefficients at the final time for the parameter μ. For each coefficient $a_j(\mu)$, where $j \in \{1, \ldots, r\}$, we model it as a Gaussian process:

$$a_j(\mu) \sim \mathcal{GP}(m_j(\mu), k_j(\mu, \mu')).$$

where $m_j(\mu)$ is the mean function and $k_j(\mu, \mu')$ is the covariance function (kernel) associated with the j-th POD coefficient. In our approach, we use a constant mean function for $m_j(\mu)$ and a radial basis function (RBF) kernel for $k_j(\mu, \mu')$, which is defined as:

$$k_j(\mu, \mu') = \sigma_j^2 \exp\left(-\frac{\|\mu - \mu'\|^2}{2\ell_j^2} \right),$$

where σ_j^2 is the signal variance, and ℓ_j is the length-scale parameter.

Given training data $\mathcal{D} = \{(\mu^{(i)}, \mathbf{a}^{(i)})\}_{i=1}^{N_\mu}$, the posterior distribution for each $a_j(\mu^*)$ at a new test point μ^* can be computed as:

$$a_j(\mu^*) \mid \mathcal{D}, \mu^* \sim \mathcal{N}(\hat{m}_j(\mu^*), \hat{k}_j(\mu^*, \mu^*)),$$

where $\hat{m}_j(\mu^*)$ and $\hat{k}_j(\mu^*, \mu^*)$ are the posterior mean and variance, respectively. This approach allows us to efficiently predict the POD coefficients for unseen parameters μ, capturing the underlying

uncertainty in the process.

2.3. *Parameterized graph-convolution autoencoder*

We employ a graph convolutional autoencoder (GCA) [15][16][17] to map a graph-based representation of the distortion field into a latent space. The encoder, $\mathcal{E}$, applies graph convolution layers to extract features and reduce dimensionality, resulting in a latent representation $\hat{\mathbf{U}} = \mathcal{E}(\mathbf{X}; \Theta_e)$, where $\mathbf{X}$ denotes the input graph and Θ_e the encoder parameters. Simultaneously, we train a fully connected neural network (FCNN), $\mathcal{F}$, that maps the dwell time dt, to this same latent space, $\hat{\mathbf{U}}^{(p)} = \mathcal{F}(dt; \Theta_f)$, with Θ_f representing the network parameters. This setup ensures that the latent space encodes both the geometrical and operational characteristics influencing the distortion. The decoder, $\mathcal{D}$, then reconstructs the distortion field from the latent space, $\hat{\mathbf{X}} = \mathcal{D}(\hat{\mathbf{U}}; \Theta_d)$, aiming to minimize the reconstruction loss $\mathcal{L}_{rec} = \|\mathbf{X} - \hat{\mathbf{X}}\|^2$. Figure 2 depicts the architecture of parameterized graph convolutional autoencoder.

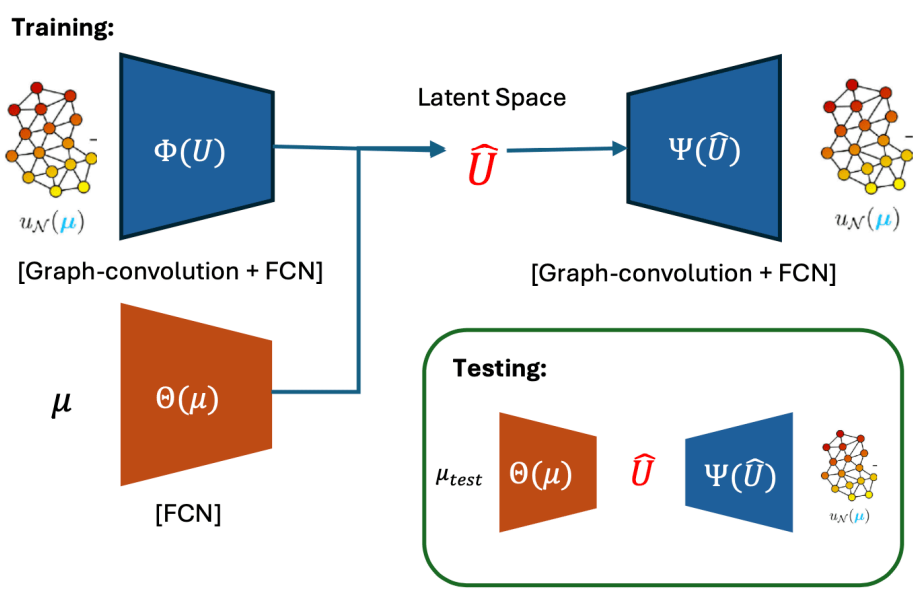

Figure 2. Illustration of parameterized graph convolutional autoencoder architecture

The overall training objective also includes a parameter consistency loss $\mathcal{L}_{param} = \|\hat{\mathbf{U}} - \hat{\mathbf{U}}^{(p)}\|^2$, ensuring that the FCNN's output aligns with the encoder's latent space. The mathematical formulation of our training objective combines these losses, optimizing Θ_e, Θ_f, and Θ_d through:

doi.org/10.32388/OS5U6O

$$\min_{\Theta_e,\Theta_f,\Theta_d} \mathcal{L}_{rec}(\mathbf{X},\hat{\mathbf{X}}) + \lambda \mathcal{L}_{param}(\hat{\mathbf{U}},\hat{\mathbf{U}}^{(p)}),$$

where λ is a regularization parameter balancing the two loss components. For the present work, we have used $\lambda = 0.5$.

3. Results

In this study, the training dataset consisted of dwell times $dt = \{20, 25, 35, 40, 50, 55, 65, 70, 80\}$ seconds. The present study aims to test the performance of data-driven models in a scarce data regime. Validation was conducted using dwell times $dt = \{30, 60\}$ seconds, while testing was performed on $dt = \{45, 75\}$ seconds. In particular, for the POD-GPR model, there is no separate validation set; both validation and test sets are combined into a single test set. The current selection of validation and test sets focuses on evaluating the performance of these models within the interpolation range of the parameters. In future work, we plan to extend the parameter space to include the extrapolation range, enabling a more comprehensive assessment of the models' generalization capabilities.

Our implementation of POD successfully preserved 99.99% of the variance with 129 modes. Subsequently, 129 independent GPRs were trained. The performance of the first four POD coefficients corresponding to the test dwell times is visually represented in the Figure 3.

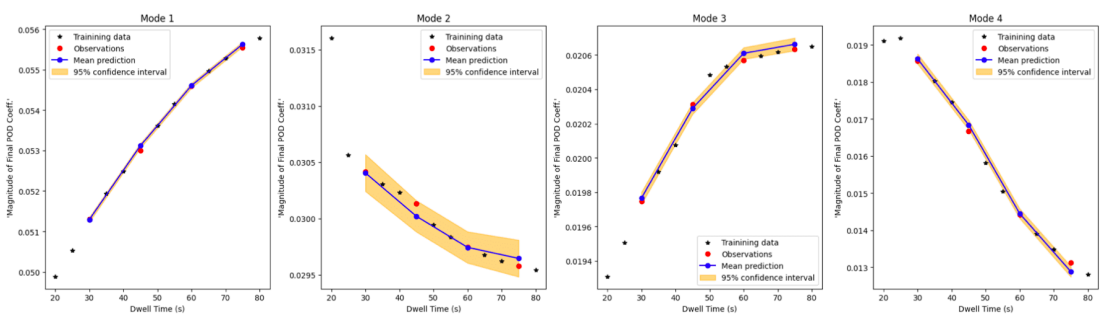

Figure 3. Graphical representation of the first four POD coefficients of final layer predicted by the GPR model for various dwell times in the test set ($dt = \{30, 45, 60, 75\}$ seconds) along with the 95% confidence interval.

The industry standard for accuracy in additive manufacturing processes is within ± 0.1 mm. Our POD-GPR model significantly exceeds this requirement, achieving an accuracy of ± 0.001 mm for the

doi.org/10.32388/OS5U6O

maximum displacement value, showcasing an excellent agreement with the finite element simulations. Notably, the runtime for the POD-GPR is approximately 4 seconds, providing a computational speed-up of about 1800 times compared to traditional finite-element methods.

The parameterized GCA was trained within a denoising autoencoder framework, employing early stopping with a patience of 50 epochs and cosine annealing warm restarts[18] for learning rate adjustment to optimize training and mitigate overfitting. The AdamW optimizer[19] was used for parameter updates. Despite setting the latent space dimension to 12 for a detailed yet compact data representation, the GCA showed tendencies of overfitting, attributed to the limited dataset size of only nine training points. This limited dataset impaired the model's generalization capabilities, particularly noticeable in test performance for dwell times of 45s and 75s. Figure 4 contrasts displacement predictions from the POD-GPR and GCA models against finite element simulations. The figure and results underscore the POD-GPR model's superior performance, highlighting its greater accuracy and effectiveness in distortion prediction for additive manufacturing.

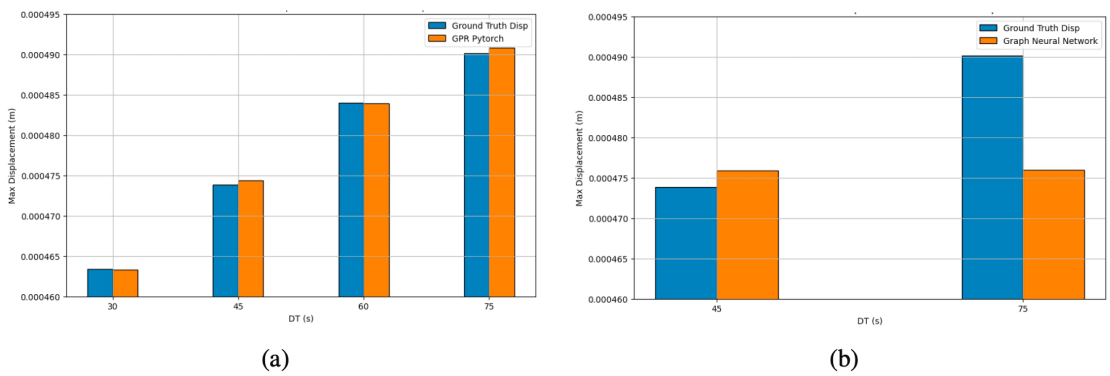

Figure 4. (a) Maximum displacement predictions from the POD-GPR model compared with ground truth from finite element analysis, (b) Predictions from the parameterized GCA model

4. Conclusion

This study highlights the POD-GPR model's exceptional accuracy and computational efficiency in distortion prediction, achieving accuracies within ± 0.001 mm and a 1800-fold speed improvement, demonstrating its suitability for engineering applications. While the parameterized GCA model faces challenges in generalizing due to a limited dataset, its versatility in adapting to different geometries

indicates significant potential for broader uses. Future work will focus on enhancing the GCA model with an enlarged dataset and exploring advanced non-linear methods such as weak-LaSDI[20].

Notes

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

Acknowledgments

I. K. Deo, Y. Choi, S. A. Khairallah were supported by Laboratory Directed Research and Development (22-SI-007) Program. Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344. IM release number: LLNL-CONF-869106. The authors extend their sincere gratitude to Gabe Guss for their invaluable assistance in conducting experiments and fabricating the 3D-printed cylindrical geometry.

References

1. ^*Gibson I, Rosen D, Stucker B. Additive Manufacturing Technologies. Springer; 2015. p. 107-145.*
2. ^*DebRoy T, et al. Additive manufacturing of metallic components - process, structure and properties. Progress in Materials Science. 92:112–224, 2018.*
3. ^*Vayre B, Vignat F, Villeneuve F (2012). "Metallic additive manufacturing: state-of-the-art review and prospects". Mechanics & Industry. 13 (2): 89–96.*
4. ^*King WE, et al. (2015). "Laser powder bed fusion additive manufacturing of metals; physics, computational, and materials challenges". Applied Physics Reviews. 2 (4): 041304.*
5. ^*Paudel BJ, Deng H, To AC (2023). "A physics-based data-driven distortion compensation model for sintered binder jet parts considering size effects". Additive Manufacturing. 68: 103517.*
6. ^*Dunbar AJ, Denlinger ER, Gouge MF, Michaleris P (2016). "Experimental validation of finite element modeling for laser powder bed fusion deformation". Additive Manufacturing. 12: 108–120.*
7. ^*Yadollahi A, Shamsaei N (2017). "Additive manufacturing of fatigue resistant materials: Challenges and opportunities". International Journal of Fatigue. 98: 14--31.*
8. ^*Fries WD, He X, Choi Y (2022). "LaSDI: Parametric latent space dynamics identification". Computer Methods in Applied Mechanics and Engineering. 399: 115436.*

9. ^Dong W, Paudel BJ, Deng H, Garner S, To AC (2024). "Data-driven distortion compensation for laser powder bed fusion process using Gaussian process regression and inherent strain method". Materials & Design. Page 113063.
10. ^Brown AL, Chin EB, Choi Y, Khairallah SA, McKeown JT. A Data-Driven, Non-Linear, Parameterized Reduced Order Model of Metal 3D Printing. Lawrence Livermore National Laboratory (LLNL), Livermore, CA (United States); 2023.
11. ^Parry L, Ashcroft IA, Wildman RD (2016). "Understanding the effect of laser scan strategy on residual stress in selective laser melting through thermo-mechanical simulation". Additive Manufacturing. **12**: 1–15.
12. ^Foster BK, Beese AM, Keist JS, McHale ET, Palmer TA (2017). "Impact of interlayer dwell time on microstructure and mechanical properties of nickel and titanium alloys". Metallurgical and Materials Transactions A. **48**: 4411–4422.
13. ^Berkooz G, Holmes P, Lumley JL (1993). "The proper orthogonal decomposition in the analysis of turbulent flows". Annual review of fluid mechanics. **25** (1): 539–575.
14. ^Schulz E, Speekenbrink M, Krause A (2018). "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions". Journal of Mathematical Psychology. **85**: 1–16.
15. ^Pichi F, Moya B, Hesthaven JS (2024). "A graph convolutional autoencoder approach to model order reduction for parametrized PDEs". Journal of Computational Physics. **501**: 112762.
16. ^Lee J, Gan C, Yang Z, Nabian MA, Zeng J, et al. Virtual Foundry Graphnet for Metal Sintering Deformation Prediction. arXiv preprint arXiv:2404.11753. 2024.
17. ^Gao R, Deo IK, Jaiman RK (2024). "A finite element-inspired hypergraph neural network: Application to fluid dynamics simulations". Journal of Computational Physics. **504**: 112866.
18. ^Liu Z. "Super convergence cosine annealing with warm-up learning rate." In: CAIBDA 2022; 2nd International Conference on Artificial Intelligence, Big Data and Algorithms. VDE; 2022. p. 1-7.
19. ^Loshchilov I, Hutter F, et al. (2017). "Fixing weight decay regularization in adam". arXiv preprint arXiv:1711.05101. **5**.
20. ^Tran A, He X, Messenger DA, Choi Y, Bortz DM (2024). "Weak-form latent space dynamics identification". Computer Methods in Applied Mechanics and Engineering. **427**: 116998.

Declarations

Funding: I. K. Deo, Y. Choi, S. A. Khairallah were supported by Laboratory Directed Research and Development (22-SI-007) Program. Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344. IM release number: LLNL-CONF-869106.

Potential competing interests: No potential competing interests to declare.