2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)
June 28-30, 2023. Seattle, Washington, USA

Geometry-agnostic Melt-pool Homogenization of Laser Powder Bed Fusion through Reinforcement Learning

Bumsoo Park[1] and Sandipan Mishra[2]

Abstract— **This work explores the feasibility of a geometry-agnostic laser power control strategy for laser powder bed fusion (L-PBF) using reinforcement learning. The controller is designed to anticipate and compensate geometry-induced process inhomogeneities, as well respond to in-process uncertainty through feedback control. To train the reinforcement learning controller, first a reduced-order simulation model is fit to experimental data. Then, the optimal control strategy is found through reinforcement learning on this reduced-order model. After the training, we demonstrate that the learned control strategy can reduce up to 55% of the error 2-norm and 59% of the standard deviation with respect to a given reference value. Moreover, the learned control strategy is applicable to novel build geometries without any additional tuning, or modification of the controller, in which we find that the controller attenuated 2-norm error by 62% and variation levels by 60% when deployed on a new (test) geometry, presenting the efficacy of the proposed controller. Finally, the experimental validation of the algorithm in a 'playback' setting resulted in a 24% reduction of both 2-norm error and variation levels, highlighting its potential in an industrial L-PPBF system.**

I. INTRODUCTION

Additive manufacturing (AM) technologies have seen a dramatic growth over the past decade [1] due to the increase in demand of rapid prototyping processes and systems capable of producing complex structures. In particular, metal AM processes such as laser powder bed fusion (L-PBF) are now finding wider application in the aerospace, automotive, and medical industries. However, quality control of the produced parts yet remains an open research challenge and thus research attempts are being made to address various aspects of the quality control problem [2]–[4]. One of the keys to quality control is to monitor the process and be able to properly compensate for variations in measurements through appropriate process control strategies.

For L-PBF systems, melt pool control is of particular interest, as the part quality is strongly correlated with the melt pool behavior [5]; with homogeneous melt-pool properties being desirable over fluctuating and changing melt pools. Consequently, the control problem in L-PBF is often stated as a melt pool regulation problem: the goal is to compensate for unexpected deviations in the melt pool geometry measurements, ultimately to reduce effects such as overheating in acute corners, or dross formation [6]. The major factors that affect the behavior of the melt pool are geometric features of the scan layer, e.g. sharp corners or narrow areas, process parameters, e.g. material type or scan velocity, and in-layer variation due to process noise, e.g. spatter, [7].

The majority of prior work on geometric feature or process parameter based control incorporates *layer-wise* feedforward control strategies, typically using models that predict the geometry-dependent behavior of the melt-pool. For instance, a number of studies focused on the demonstration of residual heat-based models [8], [9] or meltpool cross-sectional area dynamics models to design feedforward controllers [10]. Data-driven models have also shown potential in the context of control [11]–[13], through machine-learning (ML) techniques such as Gaussian process regression (GPR). Other studies directly determine the feedforward laser control signal without the use of an explicit model, by employing methods such as iterative learning control [14]–[16] to compensate for geometric effects throughout the build part. These studies altogether show that geometric effects tend to be addressable by determining the controller a-priori based on the build geometry or scan parameters.

Studies that address in-situ deviation of the melt pool indicator measurements focus on in-layer feedback control strategies because a reactive control must be addressed in real-time. However, due to the high demand on controller response time (typically in the 2-5 kHz range), feedback control strategies have been less popular in L-PBF systems compared to feedforward or layer-to-layer control strategies. The few studies that have been conducted to date [17]–[19] typically use simple feedback strategies based on PID control. While these studies show the reduction of unexpected deviations in the measurements through feedback control, such control strategies cannot explicitly address the issue of *geometry-induced melt pool inhomogeneities*.

The development of an effective L-PBF control strategy must therefore incorporate both feedforward and feedback capabilities to anticipate geometry-related effects on the melt pool and behave correspondingly, while properly responding to in-situ variation in measurements. Model-based control strategies are challenging to design for L-PBF and similar AM processes due to the complex phenomena involved [7], facilitating recent advances in data-driven/machine learning (ML) techniques [20], [21] such as reinforcement learning RL [22], because of their ability to deal with strongly nonlinear and complex dynamics. Nonetheless, the development of a control strategy that simultaneously accounts for geometric effects and in-situ process variations in L-PBF remains an open research problem.

With the motivation to design an effective feedback control strategy that can account for both geometric effects (in a

[1,2]Department of Mechanical, Aerospace, and Nuclear Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180. parkb5@rpi.edu, mishrs2@rpi.edu

feedforward manner) and process variations (through feedback), this study proposes an RL-based approach to combine geometric information with feedback control by encoding both information into the controller states. Additionally, because data-driven methods rely on substantial training data and are time-intensive, we employ a reduced-order model for the training to alleviate extensive experimentation and training efforts required for typical data-hungry ML methods. This also eliminates safety-related issues, as it is undesirable for the controller to take random actions during the training on the physical system. The key contributions of this work include the following:

1) The development and identification of a reduced-order model that captures the geometric effects in the process, which can then be used for training an RL algorithm.
2) Design of a control strategy that incorporates geometric information (in a feedforward manner) along with feedback control in L-PBF, that is applicable to novel geometries without further tuning or modification once learned.
3) Experimental demonstration of the learned optimal laser power profile in an L-PBF system in a 'playback' setting.

II. L-PBF CONTROL PROBLEM FORMULATION

A. L-PBF system

The development of the models and control strategies in this work are based on the open architecture L-PBF system described in [14] (Fig. 1). This system is equipped with a $400W$ laser for actuation, and a coaxial near-IR (NIR) camera-based setup for the measurements, similar to the setup in [17], [23]. The images are acquired at a rate of 2kHz, with a size of 64×64. The power commands are synchronized with the camera loop, and thus the power can be updated every $500\mu s$.

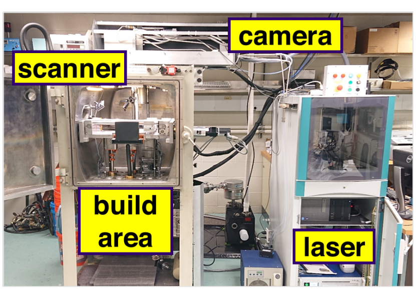

Fig. 1: L-PBF system used in this study [14]. All data for model identification was acquired from this machine and experiments were performed on this same system.

B. Problem formulation

Fig. 2 shows the general melt pool homogenization control problem addressed in this paper. The objective of this work is to design an L-PBF feedback control strategy that can anticipate and properly respond to geometric effects, in the context of the system described in the previous subsection. Information regarding the laser scan path is available in the scan file a-priori, while the near-IR camera provides melt-pool measurements at each time step. Thus, the controller must be able to appropriately incorporate this information to determine the laser power profile to respond to systemic as well as unanticipated behavior.

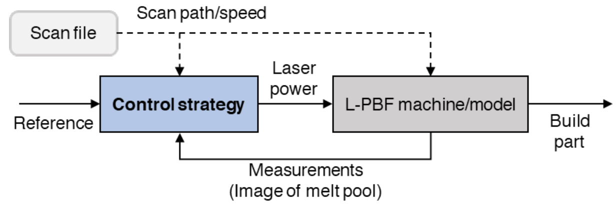

Fig. 2: Geometry-agnostic feedback control problem for L-PBF. The goal is to find a feedback control strategy that can incorporate geometric information. Given the melt-pool size and subsequent position as measurements at each time-step, the controller must select an appropriate power value.

III. PRELIMINARIES

Because of the complex dynamics of the process and the necessity of incorporating layer-wise geometry information, we aim to find the optimal control strategy through a reinforcement learning (RL) approach, instead of model-based feedback-feedforward design. Next, we provide the preliminaries for the formulation of a general RL problem and details of the specific RL algorithm used in this study.

A. Markov Decision Process

The formulation of a RL problem is based on the definition of a Markov Decision Process (MDP), which is a representation of the system in terms of states $S_t \in \mathcal{S}$, actions $A_t \in \mathcal{A}$, rewards $R_t \in \mathcal{R}$, and transition probabilities $Pr(S_{t+1}|S_t, A_t)$. Here, $\mathcal{S}$ provides a formal representation of the observations, $\mathcal{A}$ denotes the set of possible actions, $\mathcal{R}$ provides an assessment of the current circumstance, and $P(S_{t+1}|S_t, A_t)$ is the probability of transitioning from S_t to S_{t+1} taking action A_t.

B. Reinforcement learning

With all elements of the MDP defined, the control strategy, commonly referred to as the *policy* π, is defined as a function that maps the states $\mathcal{S}$ to actions $\mathcal{A}$, i.e., $\pi : \mathcal{S} \rightarrow \mathcal{A}$, in which the optimal policy is found through interaction with the system. The goal of RL is to find a policy such that maximizes the cumulative future rewards at each time-step. Two representative methods for finding the optimal policy in model free RL are value-based and policy-based [22], where the former focuses on learning an optimal value function to indirectly derive the optimal policy, and the latter focuses on directly learning the optimal policy through a parametrized function.

To implement a computationally efficient algorithm for feedback control, we use a value-based algorithm in this study, known as Q-learning [24] (tabular form). The value function with respect to a given policy π is defined as the expectation of cumulative future rewards G_t given a state s, $V^\pi(s)$ (state value function), or state-action pair (s, a),

$Q^{\pi}(s, a)$ (state-action value function). These functions can be written as:

$$V^{\pi}(s) \equiv \mathbb{E}_{\pi}\left[G_t \mid S_t = s\right] \quad (1)$$

$$Q^{\pi}(s, a) \equiv \mathbb{E}_{\pi}\left[G_t \mid S_t = s, A_t = a\right]. \quad (2)$$

The optimal value function is then found through the Bellman optimality equation [25], as shown below.

$$Q(s,a) \leftarrow Q(s,a) + \alpha\Big[r + \gamma\ max_a Q(s', a) - Q(s,a).\Big] \quad (3)$$

Here, α is the learning rate and γ is the discount factor that quantifies the importance of future rewards.

Remark. Although RL algorithms that employ function approximators (e.g. Deep Q Networks (DQN)) can also be used, the scope of this study focuses on real-time implementation of the RL scheme. Hence, a computationally efficient RL algorithm (tabular Q-learning) was used in this study.

IV. HEURISTIC SPATIO-TEMPORAL MODEL

We next define the model used for training the RL algorithm. Using the experimental system (or a high fidelity model) is not feasible for training the RL because of its data-hungry nature (i.e., a large number of trials ($10^4 - 10^6$) are often needed to train the RL scheme). Thus we present a reduced-order model that can capture the key characteristic behaviors observed in the melt pool dynamics for the training and validation of the RL.

A. Measurements and spatio-temporal registration

Because the images acquired from the L-PBF are multi-dimensional, to enable a single input single output (SISO) representation of the process we extracted a signature from the image that is indicative of melt pool size. We denote this melt pool size indicator of an image at time t as m_t.

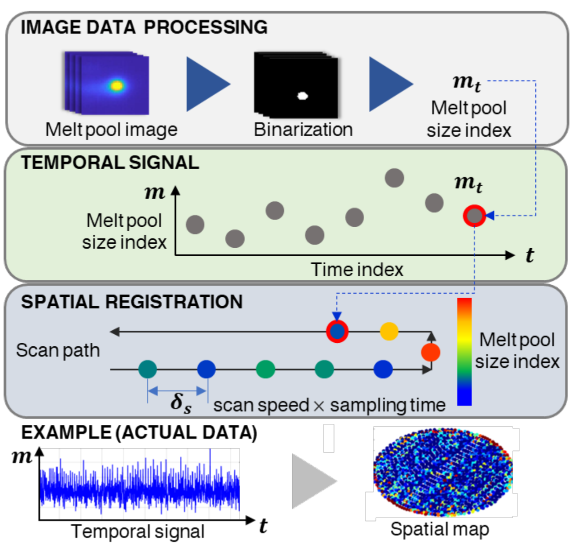

Fig. 3: Spatio-temporal registration of the measurements. Features (melt pool size index) are extracted from the image, providing a temporal signal. This signal is spatially mapped based on the scan pattern and sampling time.

The location of each measurement at time t was estimated based on the nominal scan pattern, assuming constant scan velocity and ideal trajectory tracking. We used the scan speed and sampling time (1/camera frame rate) to interpolate the individual location of each measurement. As a result, we obtain a transformation of the temporal signal $m(t)$ into a spatial map $m(x_t, y_t)$ (Fig. 3).

B. Physics-informed reduced-order spatio-temporal model for RL

Based on the spatio-temporal mapping of the measurements described above, we first create a suitable physics-informed spatio-temporal model structure, inspired by [26]. This model structure is based on the analytical temperature solution of the heat equation with Gaussian heat source moving at constant velocity over an infinite plate. The resulting solution temperature field exponentially decays over time and with increasing distance from the source location. Hence, the reduced-order spatio-temporal model was parameterized as

$$\hat{m}_{t+1} = \sum_{j=0}^{M}(m_{t-j} \cdot e^{-\lambda_d \Delta d_{tj}^{\ 2}} \cdot e^{-\lambda_t \Delta t_{tj}}) + f(p_t, v), \quad (4)$$

where $\hat{m}_{t+1}$ is the prediction at time $t+1$. The prediction considered the current measurement and $M - 1$ previous points along the scan path, where each point was regarded as a Gaussian heat source exponentially decaying over time. m_{t-j} represents the previous measurements, and λ_d and λ_t are model parameters to be identified from experimental data. Δd_{tj} and Δt_{tj} denote the difference in distance and time between the current and j^{th} points, respectively. $f(p_t, v)$ is a function mapping the laser power p_t to a scalar value, which represents the effect of the laser power. Note that while $f(\cdot, v)$ is dependent on the laser scan speed v, typically scan speed v is set apriori; and hence only the power was considered in $f(\cdot)$.

The model parameters λ_d, λ_t, and the function $f(\cdot)$ were identified from experimental data. λ_d and λ_t were found such that a mean square error was was minimized, and $f(\cdot)$ was found through linear regression based on experimental data. We evaluated the model through comparison with experimental data (Fig. 4). Due to high noise levels in experimental data, we average the spatial data over 10 layers. Fig. 4 shows that the model is capable of replicating geometric effects such as overheating due to acute turnarounds along the edges.

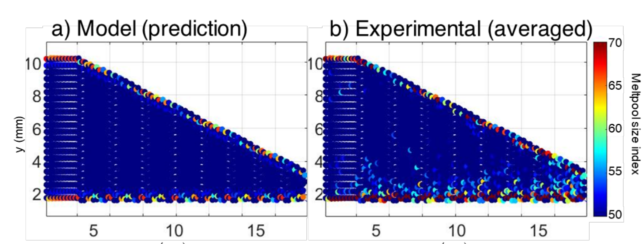

Fig. 4: Validation of developed model. Due to the noise levels in the data, the model is validated with respect to the averaged value over 10 layers from an experimental dataset. The location of points that overheat due to turnarounds from the model coincide with the points from the experiment.

V. MDP FORMULATION

As mentioned in Section III, an appropriate MDP must be defined to determine the optimal policy through RL. For real-time implementation, we prescribe an MDP with a *discrete and finite* state and action spaces; on which the tabular Q-learning RL algorithm is constructed.

State definition. The states were designed to incorporate both the positional information at time $t+1$ and the measurement at time t (Fig. 5) to allow the controller to (1) respond to current measurements, (2) while predicting the behavior of the subsequent point. Two parameters k_1 and k_2 were used to represent each piece of information. $k_1 \in \{0, 1, 2, 3\}$ was assigned a discrete value based on the range of error $m_{err} = m_{ref} - m$. $k_2 \in \{0, 1\}$ was assigned a discrete value based on the Euclidean distance to the subsequent point to represent acute turnaround points along the scan path. Turnaround points, i.e., points that have a subsequent point closer than the sampling distance δ_s (defined as the scan velocity divided by the camera frame rate) are assigned 1, and 0, otherwise (5).

$$k_1 = \begin{cases} 1 & \text{if } \|m(x_t, y_t) - m(x_{t+1}, y_{t+1})\| < \delta_s, \\ 0 & \text{if } \|m(x_t, y_t) - m(x_{t+1}, y_{t+1})\| = \delta_s. \end{cases} \quad (5)$$

The state is then defined as the Cartesian product of the two parameters, i.e., $s = [k_1, k_2]$, resulting in a total of 8 states. (Fig. 5).

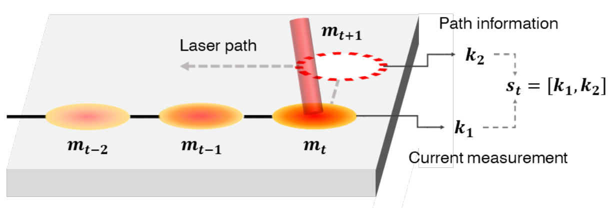

Fig. 5: State definition for MDP. k_1 represents the range of m_{err}, and k_2 represents the proximity of the subsequent point to the current point.

Action definition. Next, the actions were defined as a set of discrete power values with respect to the open loop power p_{OL}. 16 values ($p_{OL}-50, p_{OL}-45, \ldots, p_{OL}+30$), were used as the set of allowable actions. The increment was chosen as $5W$, mainly due to the fact that the difference in effect of the laser becomes relatively insignificant if $|p_{\epsilon_1} - p_{\epsilon_2}| < 5, \forall \epsilon_1 \neq \epsilon_2$.

Reward construction. Finally, we constructed the reward function to guide the policy towards a strategy to minimize error (6). The reward was constructed as a piece-wise linear function, with varying slopes to emphasize the discouragement of overheating over undermelting. The coefficients $l_1, \ldots, l_{11} \in \mathbb{R}$ and ranges were empirically tuned.

$$r(t) = \begin{cases} -l_1\{(m_{ref} + l_2) - m\} + l_3 & \text{if } m_{err} \leq -4, \\ -l_4|m_{err}| + l_5 & \text{if } |m_{err}| < 4, \\ -l_6\{(m_{ref} + l_7) - m\} + l_8 & \text{if } 4 \leq m_{err} < 8, \\ -l_9\{(m_{ref} + l_{10}) - m\} + l_{11} & \text{if } m_{err} \geq 8. \end{cases} \quad (6)$$

VI. RESULTS AND DISCUSSION

With the simulation model (to be used for training the RL) and the MDP defined above, we trained the RL algorithm on a single geometry shown in Fig. 6 (a). We first demonstrate the performance of the learned policy in the training geometry, and further show that the learned policy is applicable to novel geometries without further tuning or modification. We then 'play back' this policy on the actual experimental system to demonstrate its performance.

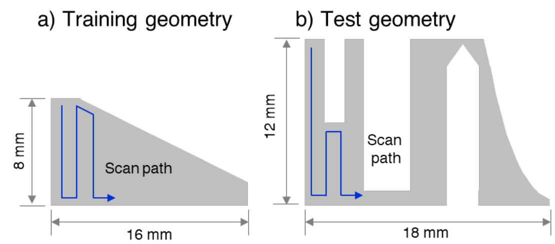

Fig. 6: Part geometry used for RL training and testing. (a) A triangular geometry was used, with hatch spacing set as 0.1 mm. (b) A relatively more complex geometry was used for deployment performance evaluation and experimental validation. Hatch spacing and scan pattern was identical to that of the train geometry.

A. Training results

The RL algorithm was trained for 200 iterations, and the geometry shown in Fig. 6 (a) was used for the training, where an entire layer is considered a single iteration (episode). The hatch spacing for the geometry was $0.1mm$ and open loop power was $250W$, with a scan speed of $800mm/s$. The open loop power was heuristically determined from prior test results, such that adequate performance was guaranteed. The discount factor $\gamma = 0.7$, the learning rate $\alpha = 0.2$, which were empirically tuned.

Next, to validate the performance of the algorithm on the training geometry, we compared the power profile, melt pool indicator measurements, and the 2-norm error with respect to m_{ref}, for the open-loop and RL-controlled case (Fig. 7). Each result (measurements and power) is shown as color coded points proportional to the value at each location. The RL agent learned to effectively lower the power at the turnaround points, resulting in a more uniform measurement map. Through this strategy, the RL agent reduced up to 55% of the melt pool signal variation[1] and 59% of the 2-norm error.

Comparison with PI control. To compare the performance of RL with respect to other feedback algorithms, we compared the results from RL to that of a PI controller (Fig. 7). Because the PI controller is reactive, the controller is unable to eliminate the overheating around the turnaround points, thus we omit the PI control in further analyses.

[1]Variation is defined as standard deviation; $\sqrt{\frac{1}{N}\sum|m_i - \mu_m|^2}$

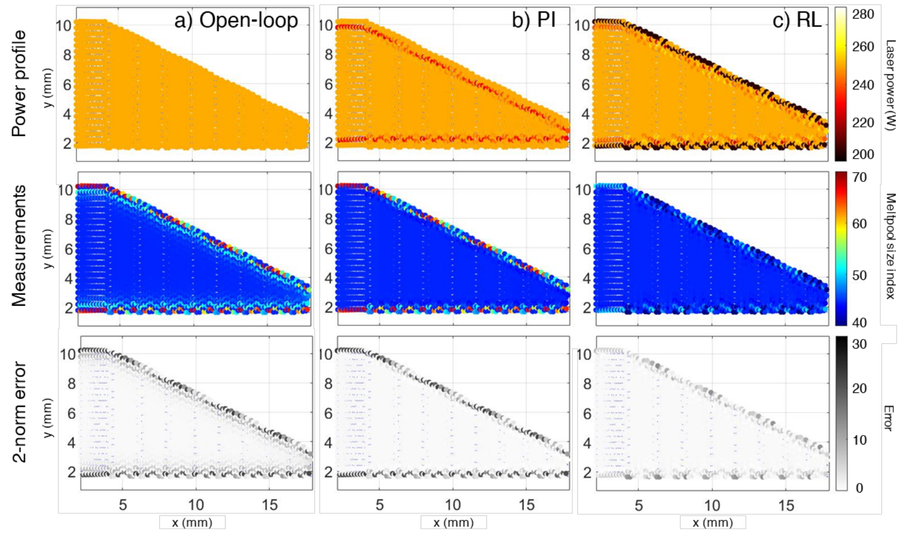

Fig. 7: Comparison of the performance of the proposed RL algorithm against open loop and PI control. The reference meltpool signal was set to $m_{ref} = 44$, based on appropriate calibrated melt pool size. Power profiles, measurements, and 2-norm errors are compared with each other. The PI control attenuated the noise by 3% and reduced the 2-norm error by 7%, whereas the RL attenuated the variation by 55% and reduced the 2-norm error by 59%. We note also that the PI control is unable to compensate for overheating due to turnarounds.

B. Deployment of learned policy

To further show that the learned policy is applicable to novel geometries/scan-paths without further tuning, we deployed the learned policy in a novel build geometry. This test geometry, shown in Fig. 6 (b), was designed to have a more complex structure. The scan pattern, hatch spacing, and scan parameters were identical to that of the train geometry.

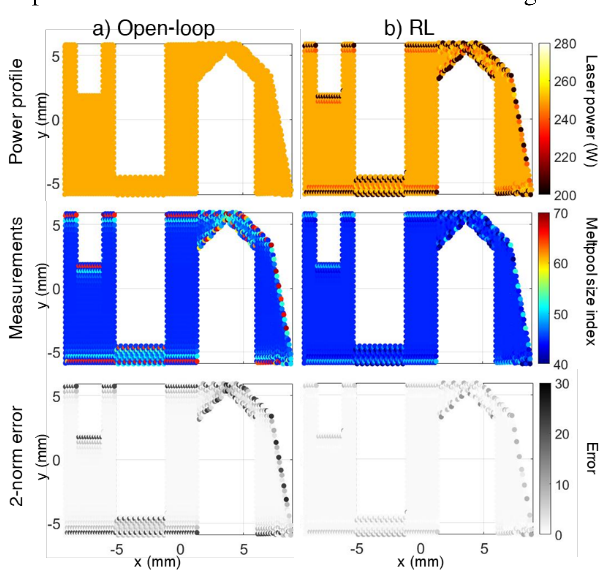

Fig. 8: Deployment of learned policy. Open-loop results are compared to RL results, similar to Fig. 7. Results are shown for $m_{ref} = 44$. The RL attenuated the variance by 60% and reduced the 2-norm error by 62%.

We compared the power profiles, measurements, and 2-norm errors of the RL against open loop (Fig. 8). The results show that even for a novel geometry, the RL controller effectively lowered the power at the turnaround points. Additionally, the controller attenuated the local heating induced by the narrow channels through feedback, also demonstrating feedback capabilities. The RL reduced up to 60% of the variation and 62% of the 2-norm error, showing that the learned policy is applicable to unforeseen geometries without parameter tuning.

Remark. Note that the actions are bounded (Sec.V), and thus the output of the system is expected to be bounded. Moreover, the system is open-loop stable, hence additional stability analysis was omitted from the scope of this study.

C. Experimental validation through playback

Finally, we experimentally validated the proposed RL algorithm by 'replaying' the power profiles on the actual system. The same geometry from the deployment scenario (Fig. 6 (b)) was used in the experiment, hence the power profile from Fig. 8 was applied in a feedforward manner.

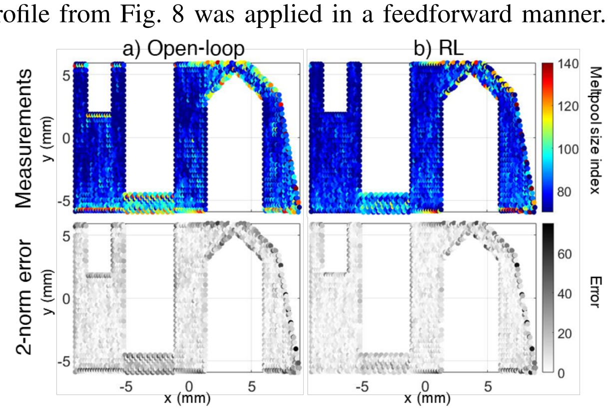

Fig. 9: Experimental validation in 'play back' setting. Open-loop results are compared to RL results, with the same power profile from Fig. 8. The applied power profile resulted in a 24% reduction of both the 2-norm error and variation.

Because the same power profile from Fig. 8 was applied, here we visualize the measurements and 2-norm error (Fig. 9). Similar to the results from the simulation, we the reduction of power around the edges result in mitigation of overheating points. Due to the high level of noise in the actual system, we averaged the measurements across 10 layers (as demonstrated in Fig. 4). The 2-norm error and variation reduction for the experimental case was both 24%, suggesting that the power profile determined offline was effective in terms of the measurement homogenization for the experimental setup. Note that the measurement range is different from that of Fig. 4 due to changes in the camera settings; and thus the reference value ($m_{ref} = 77$) was recalculated by fitting the model again from the open loop data of the experimental validation geometry, and converting the previous reference value ($m_{ref} = 44$).

VII. CONCLUSIONS AND FUTURE WORK

In this study, we designed and investigated an RL-based control strategy for an L-PBF process, that is capable of both responding to geometry features and in-situ process measurements. The demonstration of the algorithm in a simulator showed that the 2-norm error can be reduced up to 60% and variance up to 62%. Furthermore, we found that the trained policy is applicable to novel geometries without further tuning or modification. Finally, the proposed algorithm reduced the 2-norm error and variation in an actual system by 24% in a 'play back' setting, well demonstrating the potential of the proposed RL-based control in L-PBF systems.

Future work will address the following: Although the control strategy is capable of feedback, due to the lack of noise in the simulation model such capabilities were not investigated to full extent in this study. Hence the algorithm will be implemented and tested on an L-PBF testbed to demonstrate real-time feedback capabilities.

ACKNOWLEDGMENT

This work was supported by NSF Data-Driven Cyberphysical Systems Award #1645648 and by the State of New York ESD/NYSTAR program.

REFERENCES

[1] O. Abdulhameed, A. Al-Ahmari, W. Ameen, and S. H. Mian, "Additive manufacturing: Challenges, trends, and applications," *Advances in Mechanical Engineering*, vol. 11, no. 2, p. 1687814018822880, 2019.
[2] B. Zhang, Y. Li, and Q. Bai, "Defect formation mechanisms in selective laser melting: a review," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 3, pp. 515–527, 2017.
[3] D. Wang, Y. Yang, Z. Yi, and X. Su, "Research on the fabricating quality optimization of the overhanging surface in slm process," *The International Journal of Advanced Manufacturing Technology*, vol. 65, no. 9-12, pp. 1471–1484, 2013.
[4] J. C. Fox, S. P. Moylan, and B. M. Lane, "Effect of process parameters on the surface roughness of overhanging structures in laser powder bed fusion additive manufacturing," *Procedia Cirp*, vol. 45, pp. 131–134, 2016.
[5] W. E. King, A. T. Anderson, R. M. Ferencz, N. E. Hodge, C. Kamath, S. A. Khairallah, and A. M. Rubenchik, "Laser powder bed fusion additive manufacturing of metals: physics, computational, and materials challenges," *Applied Physics Reviews*, vol. 2, no. 4, p. 041304, 2015. [Online]. Available: https://e-reports-ext.llnl.gov/pdf/800252.pdf http://aip.scitation.org/doi/10.1063/1.4937809
[6] T. Craeghs, S. Clijsters, E. Yasa, F. Bechmann, S. Berumen, and J.-P. Kruth, "Determination of geometrical factors in layerwise laser melting using optical process monitoring," *Optics and Lasers in Engineering*, vol. 49, no. 12, pp. 1440–1446, 2011.
[7] S. R. Narasimharaju, W. Zeng, T. L. See, Z. Zhu, P. Scott, X. Jiang, and S. Lou, "A comprehensive review on laser powder bed fusion of steels: Processing, microstructure, defects and control methods, mechanical properties, current challenges and future trends," *Journal of Manufacturing Processes*, vol. 75, pp. 375–414, 2022.
[8] H. Yeung, B. Lane, and J. Fox, "Part geometry and conduction-based laser power control for powder bed fusion additive manufacturing," *Additive Manufacturing*, vol. 30, 2019.
[9] H. Yeung and B. Lane, "A residual heat compensation based scan strategy for powder bed fusion additive manufacturing," *Manufacturing Letters*, vol. 25, pp. 56–59, 2020. [Online]. Available: https://doi.org/10.1016/j.mfglet.2020.07.005
[10] Q. Wang *et al.*, "Model-based feedforward control of laser powder bed fusion additive manufacturing," *Additive Manufacturing*, vol. 31, 2020.
[11] H. Yeung, Z. Yang, and L. Yan, "A meltpool prediction based scan strategy for powder bed fusion additive manufacturing," *Additive Manufacturing*, vol. 35, 2020.
[12] Y. Ren and Q. Wang, "Gaussian-process based modeling and optimal control of melt-pool geometry in laser powder bed fusion," *Journal of Intelligent Manufacturing*, pp. 1–18, 2021.
[13] A. Shkoruta, B. Park, and S. Mishra, "An empirical model for feedforward control of laser powder bed fusion," *arXiv preprint arXiv:2201.09978*, 2022.
[14] A. Shkoruta, W. Caynoski, S. Mishra, and S. Rock, "Iterative learning control for power profile shaping in selective laser melting," in *IEEE International Conference on Automation Science and Engineering*, vol. 2019-August, 2019, pp. 655–660.
[15] X. Wang, C. S. Lough, D. A. Bristow, R. G. Landers, and E. C. Kinzel, "A layer-to-layer control-oriented model for selective laser melting," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 481–486.
[16] X. Wang, R. G. Landers, and D. A. Bristow, "Spatial transformation of a layer-to-layer control model for selective laser melting," in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 2886–2891.
[17] J.-P. Kruth, P. Mercelis, J. Van Vaerenbergh, and T. Craeghs, "Feedback control of selective laser melting," in *Proceedings of the 3rd international conference on advanced research in virtual and rapid prototyping*. Taylor & Francis Ltd, 2007, pp. 521–527.
[18] V. Renken, A. V. Freyberg, K. Schünemann, F. Pastors, and A. Fischer, "In-process closed-loop control for stabilising the melt pool temperature in selective laser melting," *Progress in Additive Manufacturing*, no. 0123456789, 2019. [Online]. Available: https://doi.org/10.1007/s40964-019-00083-9
[19] A. Shkoruta, S. Mishra, and S. J. Rock, "Real-Time Image-Based Feedback Control of Laser Powder Bed Fusion," *ASME Letters in Dynamic Systems and Control*, vol. 2, no. 2, 07 2021, 021001. [Online]. Available: https://doi.org/10.1115/1.4051588
[20] P. Wang, Y. Yang, and N. S. Moghaddam, "Process modeling in laser powder bed fusion towards defect detection and quality control via machine learning: The state-of-the-art and research challenges," *Journal of Manufacturing Processes*, vol. 73, pp. 961–984, 2022.
[21] F. Ogoke and A. B. Farimani, "Thermal control of laser powder bed fusion using deep reinforcement learning," *Additive Manufacturing*, vol. 46, p. 102033, 2021.
[22] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT Press, 2018.
[23] A. Gökhan Demir, C. De Giorgi, and B. Previtali, "Design and implementation of a multisensor coaxial monitoring system with correction strategies for selective laser melting of a maraging steel," *Journal of Manufacturing Science and Engineering*, vol. 140, no. 4, 2018.
[24] C. J. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, pp. 279–92, 1992.
[25] R. Bellman, "A markovian decision process," *Journal of mathematics and mechanics*, pp. 679–684, 1957.
[26] T. Eagar, N. Tsai, *et al.*, "Temperature fields produced by traveling distributed heat sources," *Welding journal*, vol. 62, no. 12, pp. 346–355, 1983.