INVERTIBLE SURROGATE MODELS: JOINT SURROGATE MODELLING AND RECONSTRUCTION OF LASER-WAKEFIELD ACCELERATION BY INVERTIBLE NEURAL NETWORKS

Friedrich Bethke[1], **Richard Pausch**[1], **Patrick Stiller**[1],
Alexander Debus[1], **Michael Bussmann**[1,2], **Nico Hoffmann**[1]

[1] Helmholtz-Zentrum Dresden - Rossendorf, Bautzner Landstrasse 400, 01328 Dresden, Germany
[2] CASUS - Center for Advanced Systems Understanding, Untermarkt 20, 02826 Görlitz, Germany
{f.bethke, n.hoffmann}@hzdr.de

ABSTRACT

Invertible neural networks are a recent technique in machine learning promising neural network architectures that can be run in forward and reverse mode. In this paper, we will be introducing invertible surrogate models that approximate complex forward simulation of the physics involved in laser plasma accelerators: iLWFA. The bijective design of the surrogate model also provides all means for reconstruction of experimentally acquired diagnostics. The quality of our invertible laser wakefield acceleration network will be verified on a large set of numerical LWFA simulations.

1 INTRODUCTION

Laser plasma accelerators are a new technique to accelerate charged particles, providing acceleration gradients orders of magnitude higher than with conventional accelerators. One of the most established techniques is the so-called laser wakefield acceleration (LWFA) (Tajima & Dawson, 1979; Esarey et al., 2009). A short and intense laser pulse drives a plasma wave inside which electrons can be accelerated. Especially in the so-called blowout or bubble regime (Pukhov & Meyer-ter Vehn, 2002; Geddes et al., 2004; Faure et al., 2004; Mangles et al., 2004), quasi-monoenergetic electron bunches with pulse duration and extent in the range of a few fs and $\mu\mathrm{m}$ can be generated which are not achievable by conventional accelerators. One of the most commonly used methods to inject electrons into the plasma accelerator is downramp injection (Schmid et al., 2010; Buck et al., 2013; Swanson et al., 2017), during which the plasma wake transitions from a high- to a low-density region, thereby expanding the plasma cavity and, if the laser is strong enough, injecting electrons that are further accelerated. This injection method can provide significantly better beam quality compared to other injection methods (Barber et al., 2017). The injection process depends both on the laser evolution as well as on the plasma dynamics, which are nonlinearly coupled. During the consecutive acceleration, the beam quality is influenced by the complex interplay between beam and wake, called beam loading Couperus et al. (2017). Making predictions for the injection process and acceleration thus require modeling with large-scale particle-in-cell simulations (Martinez de la Ossa et al., 2017). A surrogate model for these simulations is highly desireable since the simulations are highly computationally demanding even for highly-optimized particle-in-cell codesBurau et al. (2010), A fast prediction of electron parameters of interest could accelerate the applicability of down-ramp injection methods even further. Main contributions of this paper is the introduction of invertible surrogate modelling applied to accelerator physics. The proposed invertible LWFA surrogate model (iLWFA) approximates the non-linear forward simulation of LWFA with focus on down-ramp injection. Furthermore, the choice of architecture allows us to jointly solve the inverse problem, i.e. reconstruction of simulation parameters based on energy spectra.

2 RELATED WORKS

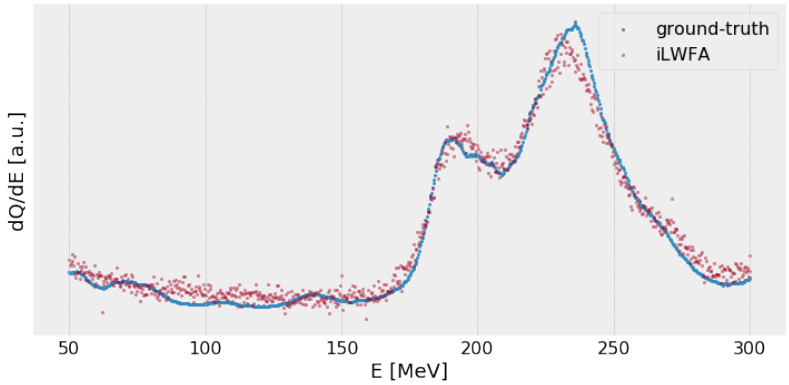


Figure 1: The prediction of the iLWFA network (red) on validation data very closely approximates the energy spectrum of numerical particle-in-cell code (blue) in a fraction of time. The spectrum shows the final charge of the electrons per unit energy $\mathrm{d}Q/\mathrm{d}E \hat{=} \Psi$ over energy E. For this particular parameter set p, the most electrons end up to have an energy between $180\,\mathrm{MeV}$ and $250\,\mathrm{MeV}$.

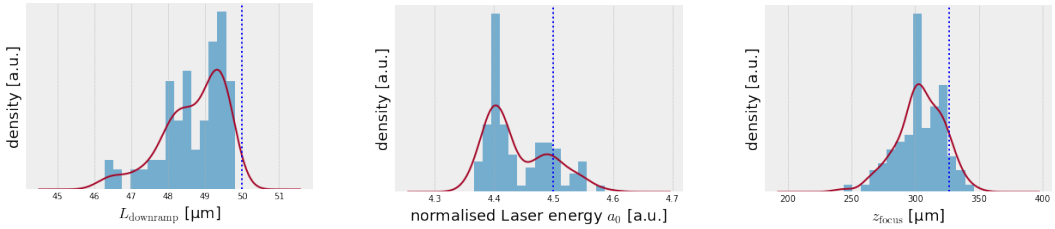
Computationally expensive simulations are necessary for simulation of complex physical processes. Scanning the input parameters of physical simulations significantly increases the runtime of the simulations, since the simulation has to be cold-started over and over again. Surrogate models which are able to learn the relationship among the input parameters and the simulation output could significantly lower the computational complexity and enables more extensive analysis of the underlying physical system without restarting the simulation every time.

Neural networks are a valuable choice as surrogate models due to their ability as universal function approximators (Hornik et al., 1989). Certain established approaches rely on the multilayer perceptron architecture (Raissi et al., 2017; Sirignano & Spiliopoulos, 2017; Stiller et al., 2020) while more sophisticated approaches such as generative adversial networks, autoencoders (Xie et al., 2018; Kim et al., 2019) and graph neural networks are also used (Sanchez-Gonzalez et al., 2020). However, all these approaches are designed as surrogates for the forward process of the underlying simulation, so that the inverse mapping must be represented by a separate model (e.g. (Raissi & Karniadakis, 2017), (Raissi, 2018)).

Invertible neural networks are a promising candidate as surrogate models through there ability to learn the forward and inverse mapping simultaneously. Padmanabha & Zabaras (2020) applied a conditional invertible neural network as a surrgogate model for the estimation of a non-Gaussian permeability field in multiphase flows. In this paper, we will be learning an invertible conditional mapping between simulation parameters and -outcomes by adapting the architecture of (Ardizzone et al., 2018), which is more challenging to train but promises a more accurate approximation of the forward pass (simulation) and the inverse process (reconstruction).

3 METHOD

A surrogate model of laser wakefield acceleration can be seen as mapping $g(p) = \Psi(E)$ from n_p simulation parameters $p \in \mathbb{R}^{n_p}$ to energy spectrum $\frac{\mathrm{d}Q}{\mathrm{d}E} \hat{=} \Psi = [\Psi_1, \Psi_2, ..., \Psi_{n_E}]$. The spectrum Ψ is discretized for n_E energy bins $E_i \in [40; 300]\ \mathrm{MeV}$ and represents the total charge of all electrons in that specific energy bin E_i. In experiment, parameters p are typically not directly measurable but provide important insights about the state of the system rendering the need for reconstruction $f(\cdot)$ of parameters p given the energy spectrum Ψ, i.e. $f(\Psi) = p$. The reconstruction can be seen as (ill-posed) inverse process of our surrogate model which, however, might not be injective due to loss of information as the energy spectrum might not represent the full state of the system. This means that certain parameters p_j, p_k might map to the same energy spectrum Ψ, i.e. $g(p_j) = g(p_k) = \Psi$. This ambiguity implies that any neural network $n \approx f$ would either pick any of the feasible



(a) (b) (c)

Figure 2: The posterior predictive distribution of iLWFA on validation data is close to the ground-truth (blue). The method is not restricted to certain family of distributions as well as number of modes as (a)-(c) show. The two modes of (b) encode information about the ambiguity when reconstructing the laser parameter from a single energy spectrum.

parameters or return the average of both, depending on the choice of architecture and objective function. We approach this challenge by invertible neural networks (Ardizzone et al., 2018) that allow us to recover the full posterior predictive distribution of p given the observed energy spectrum Ψ, i.e. $\pi_{\mathcal{P}}(p|\Psi, z) = f(\Psi, z)$. The modes of this posterior predictive distribution $\pi_{\mathcal{P}}(p|\Psi, z)$ correspond to all parameter configurations explaining the observed energy spectrum Ψ.

The main idea of an invertible surrogate model for Laser-Wakefield acceleration is that some neural networks transform a simple multivariate base distribution $z \sim \pi_{\mathcal{Z}} = N(0, I)$ with zero mean and identity matrix I into a complex posterior distribution of parameters $\pi_{\mathcal{P}}$ subject to observation Ψ in terms of a non-linear function $f(\Psi, z)$. The posterior can be derived by the change-of-variables approach as of

$$\pi_{\mathcal{P}} = \pi_{\mathcal{Z}} \left| det \left(\frac{\partial f(\Psi, z|\theta)}{\partial [\Psi^T, z^T]} \right) \right| \tag{1}$$

given parameters θ of $f(\cdot)$. This architecture allows us to resolve ambiguous mappings since $f(\cdot)$ is trained such taht it transforms different parts of our base distribution $\pi_{\mathcal{Z}}$ to each mode of the data distribution while any simpler data-driven approach would either pick one mode or return the average value of all corresponding parameters. It has been shown that this choice of architecture is able to return the true posterior distribution (Ardizzone et al., 2018).

The function $f(\cdot)$ is defined by a composition of L affine coupling transforms (Dinh et al., 2016). The input $u^{(l)}$ to the l-th coupling transform is split into two equally sized halves $u^{(l)} = [u_1^{(l)}, u_2^{(l)}]$ and gets transformed into $v^{(l)} = [v_1, v_2]$ according to

$$\begin{aligned} v_1^{(l)} &= u_1^{(l)} \cdot exp(s^{(l)}(u_2^{(l)})) + t^{(l)}(u_2^{(l)}) \\ v_2^{(l)} &= u_2^{(l)} \cdot exp(s^{(l)}(v_1^{(l)})) + t^{(l)}(v_1^{(l)}) \end{aligned} \tag{2}$$

for any block $1 \leq l \leq L$ of coupling transform and fully-connected neural networks $s^{(l)}, t^{(l)}$. This design can be seen as a modified version of the GLOW approach (Kingma & Dhariwal, 2018). Finally, we need to introduce some zero-padding to the co-domain of $f : \mathbb{R}^{M_E+M_z} \rightarrow \mathbb{R}^{M_p+M_0}$ to learn a bijective mapping, i.e. $[\Psi, z] = u^{(1)} \in \mathbb{R}^{M_E+M_z}$ and $[p, 0] = v^{(L)} \in \mathbb{R}^{M_p+M_0}$ with $M_p + M_0 = M_E + M_z$. M_E denotes the number of energy bins of our energy spectra Ψ while M_p is the number of parameters (here: 3), M_z is the size of our latent variable z and M_0 the resulting zero-padding. This choice of coupling transform is easily invertible and leads to a triangular Jacobian structure such that the Jacobian determinant of eqn. 1 can be computed computationally inexpensively (Dinh et al., 2016). Additionally, the affine coupling transforms also enable fast inference times of the forward direction (surrogate model), i.e. $g(p) = INN(p, 0) = [\Psi, z]$ as well as the reverse direction (reconstruction).

The bidirectional training of the invertible surrogate model is carried out by minimizing our objective function L with respect to the parameters θ of all subnetworks $s^{(l)}, t^{(l)}$,

$$L = L_\Psi + \lambda_p L_p + \lambda_z L_z \tag{3}$$

with L_Ψ being the supervised forward loss, L_p is the supervised reconstruction loss and L_z preserves the prior distribution of the latent variable z. The Lagrange multipliers λ_p, λ_z scale each term of our

objective function accordingly. The supervised forward loss,

$$L_\Psi = \sum_{i=0}^{N} ||\Psi_i - \widehat{\Psi}_i||_1 \tag{4}$$

enforces similarity of the approximated forward simulation $[\widehat{\Psi}_i, \widehat{z}] = f(p_i, 0)$ with ground-truth energy spectrum Ψ_i of parameter p_i. The reconstruction loss L_p reads

$$L_p = \sum_{i=0}^{N} (||p_i - \widehat{p}_i||_2^2 + ||\widehat{0}_i||_2^2) \tag{5}$$

with $[\widehat{p}_i, \widehat{0}_i] = g(\Psi_i, z)$. All terms are evaluated at at N samples of the training set. The latter term of L_p enforces that the invertible network is not encoding any information into the *predicted* zero-padding $\widehat{0}$ of g. Finally, the predicted latent variable $\widehat{z}$ is regularized by maximum mean discrepancy (MMD) as of (Ardizzone et al., 2018),

$$L_z = MMD(f(p, 0), \pi_\Psi \pi_z) \tag{6}$$

to enforce normality of $\widehat{z}$ and independence to the predicted $\widehat{\Psi}$.

4 Results & Discussion

All 2.7 Terabytes of training data were generated by large scale PIConGPU (Bussmann et al., 2013; Burau et al., 2010) simulations run by a jupyter-based scheduler (Rudat, 2019). In the following the 3 input parameters to the PIConGPU simulations are: the laser's normalized field strength a_0, the laser's focus position z_{focus} with $z_{\text{focus}} = 105\,\mu\text{m}$ representing a focus position at the density downramp, and the length of the downramp density transition L_{downramp} from a density $2 \cdot 10^{19}\,\text{cm}^{-3}$ to $1.1 \cdot 10^{19}\,\text{cm}^{-3}$. Weights λ of our objective function were subject to a hyperparameter optimisation. The best architecture consists of latent distribution π_z with $M_z = 20$ and a stack of 10 coupling transforms. Each transform consists of two fully-connected neural networks with four hidden layers amounting to $2.9M$ parameters for that transform. The total number of parameter of the invertible surrogate model is $\approx 29M$.

4.1 Surrogate model

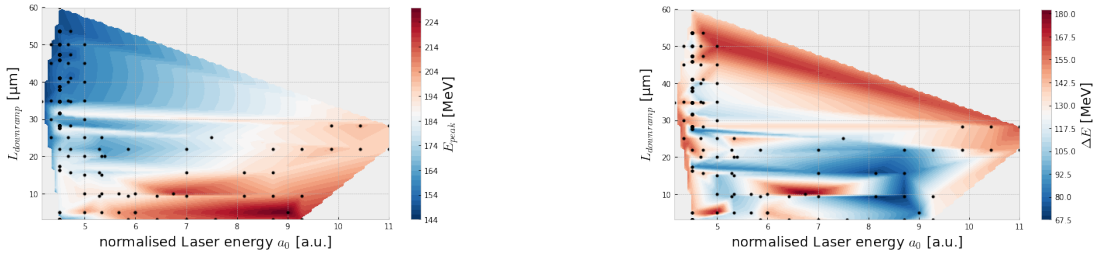
The performance of our iLWFA model was assessed by randomly splitting our dataset into training and validation data. We found that the approximation of the forward pass yields a mean squared error of $MSE < 0.007$ meaning that the energy spectra were recovered reasonable well (see e.g. Fig. 1). Furthermore, the similarity in shape of the reconstructed energy spectra with respect to groundtruth data was quantified by structured-similarity index $SSIM \approx 0.86$, i.e. the recovered energy spectra closely resemble the ground-truth data. The reconstruction of simulation parameters given an energy spectrum relates to solving an inverse problem. iLWFA allows us to solve this inverse problem by querying the *invertible* neural network in reverse mode. A representative reconstruction of the posterior predictive distribtuion on validation data emphasizes the strengths of INNs for learning complex posterior distributions (see Fig. 2). The median relative errors for reconstruction of simulation parameters from energy spectra is rather low (see table 1) meaning that the invertible network is learning features for joint forward simulation as well as parameter reconstruction. The good reconstruction performance on a_0 could be explained by strong contribution of that parameter to the reconstructed energy spectra (see Fig. 4).

4.2 Direct physical application of the surrogate model

Various beam parameters need to be tuned simultaneously to achieve a certain LWFA downramp injection for a targeted application. For examples narrow energy bandwidth ΔE around a high peak energy E_{peak} is required for operating Thomson scattering light sources Jochmann et al. (2013); Krämer et al. (2018) or compact optical FELs Steiniger et al. (2019). On the other hand, using these electron beams as driver in a compact plasma wakefield accelerator Kurz et al. (2021), requires high-current beams, thus a high peak energy is needed but the requirements on ΔE are less strict.

Table 1: The relative error of the inverse pass are proportional to the contribution of each parameter to the energy spectrum.

RELATIVE ERROR	TRAINING	VALIDATION
a_0	0.2%	0.3%
L_{downramp}	0.6%	3.7%
z_{focus}	2.5%	8.2%



(a) (b)

Figure 3: The surrogate model approximates the forward pass quite smoothly resulting in smooth transitions of our derived quantities peak energy E_{peak} with full-width half-maximum ΔE. Hereby we are able to identify that highest peak energy can be reached with $a_0 \approx 9$ and a short downramp length L_{downramp} at a narrow energy bandwidth $\Delta E = 82\,\text{MeV}$.

5 Conclusions

Laser wakefield acceleration is an established compact accelerator method promising significantly higher acceleration gradients than conventional particle accelerators. The numerical simulation of the involved complex physics requires jointly solving kinematic- and Maxwell's equation using particle-in-cell method. Main contribution of this work is an invertible surrogate model, iLWFA, that is approximating the forward pass of a full LWFA simulation while providing all means for reconstruction of experimentally acquired quantities. Another benefit of ML-driven surrogate models is the differentiability of the model allowing us to infer physical knowledge from the incorporated non-linear mapping. An evaluation on large set of numerical LWFA simulations emphasizes the benefits of this approach but also outlines future work for making the reconstruction more robust.

REFERENCES

Lynton Ardizzone, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W. Pellegrini, Ralf S. Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. Analyzing Inverse Problems with Invertible Neural Networks. In *Proceedings of the 7th International Conference on Learning Representations*, number i, pp. 1–20, 2018. URL http://arxiv.org/abs/1808.04730.

S. K. Barber, J. van Tilborg, C. B. Schroeder, R. Lehe, H.-E. Tsai, K. K. Swanson, S. Steinke, K. Nakamura, C. G. R. Geddes, C. Benedetti, E. Esarey, and W. P. Leemans. Measured Emittance Dependence on the Injection Method in Laser Plasma Accelerators. *Physical Review Letters*, 119(10):104801, sep 2017. ISSN 0031-9007. doi: 10.1103/PhysRevLett.119.104801. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.104801.

A. Buck, J. Wenz, J. Xu, K. Khrennikov, K. Schmid, M. Heigoldt, J. M. Mikhailova, M. Geissler, B. Shen, F. Krausz, S. Karsch, and L. Veisz. Shock-Front Injector for High-Quality Laser-Plasma Acceleration. *Physical Review Letters*, 110(18):185006, may 2013. ISSN 0031-9007. doi: 10.1103/PhysRevLett.110.185006. URL https://link.aps.org/doi/10.1103/PhysRevLett.110.185006.

Heiko Burau, Renée Widera, Wolfgang Hönig, Guido Juckeland, Alexander Debus, Thomas Kluge, Ulrich Schramm, Tomas E. Cowan, Roland Sauerbrey, and Michael Bussmann. PIConGPU: A fully relativistic particle-in-cell code for a GPU cluster. *IEEE Transactions on Plasma Science*, 38(10 PART 2):2831–2839, 2010. ISSN 0093-3813. doi: 10.1109/TPS.2010.2064310.

M. Bussmann, H. Burau, T. E. Cowan, A. Debus, A. Huebl, G. Juckeland, T. Kluge, W. E. Nagel, R. Pausch, F. Schmitt, U. Schramm, J. Schuchart, and R. Widera. Radiative signatures of the relativistic Kelvin-Helmholtz instability. In *SC '13 Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 5–1 – 5–12, 2013. ISBN 978-1-4503-2378-9. doi: 10.1145/2503210.2504564. URL http://dl.acm.org/citation.cfm?doid=2503210.2504564.

J.P. Couperus, R. Pausch, A. Köhler, O. Zarini, J.M. Krämer, M. Garten, A. Huebl, R. Gebhardt, U. Helbig, S. Bock, K. Zeil, A. Debus, M. Bussmann, U. Schramm, and A. Irman. Demonstration of a beam loaded nanocoulomb-class laser wakefield accelerator. *Nature Communications*, 8:487, dec 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-00592-7. URL http://www.nature.com/articles/s41467-017-00592-7.

Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using Real NVP. 2016. URL http://arxiv.org/abs/1605.08803.

Eric Esarey, C. B. Schroeder, and W. P. Leemans. Physics of laser-driven plasma-based electron accelerators. *Reviews of Modern Physics*, 81(3):1229–1285, 2009. ISSN 0034-6861. doi: 10.1103/RevModPhys.81.1229.

J. Faure, Y. Glinec, A. Pukhov, S. Kiselev, S. Gordienko, E. Lefebvre, J.-P. Rousseau, F Burgy, and V. Malka. A laser-plasma accelerator producing monoenergetic electron beams. *Nature*, 431(7008):541–544, sep 2004. ISSN 0028-0836. doi: 10.1038/nature02963. URL http://www.nature.com/articles/nature02963.

C G R Geddes, Cs. Toth, J. van Tilborg, E Esarey, C B Schroeder, D Bruhwiler, C Nieter, J Cary, and W P Leemans. High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding. *Nature*, 431(7008):538–541, sep 2004. ISSN 0028-0836. doi: 10.1038/nature02900. URL http://www.nature.com/articles/nature02900.

Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. ISSN 0893-6080. doi: https://doi.org/10.1016/0893-6080(89)90020-8. URL https://www.sciencedirect.com/science/article/pii/0893608089900208.

A. Jochmann, A. Irman, M. Bussmann, J. P. Couperus, T. E. Cowan, A. D. Debus, M. Kuntzsch, K. W D Ledingham, U. Lehnert, R. Sauerbrey, H. P. Schlenvoigt, D. Seipt, Th Stöhlker, D. B. Thorn, S. Trotsenko, A. Wagner, and U. Schramm. High Resolution Energy-Angle Correlation

Measurement of Hard X Rays from Laser-Thomson Backscattering. *Physical Review Letters*, 111(11):114803, sep 2013. ISSN 0031-9007. doi: 10.1103/PhysRevLett.111.114803. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.114803.

Byungsoo Kim, Vinicius Azevedo, Nils Thuerey, Theodore Kim, Markus Gross, and Barbara Solenthaler. Deep fluids: A generative network for parameterized fluid simulations. *Computer Graphics Forum*, 38:59–70, 05 2019. doi: 10.1111/cgf.13619.

Diederik P. Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. 2018.

J.M. Krämer, A. Jochmann, M. Budde, M. Bussmann, J.P. P. Couperus, T.E. E. Cowan, A. Debus, A. Köhler, M. Kuntzsch, A. Laso García, U. Lehnert, P. Michel, R. Pausch, O. Zarini, U. Schramm, and A. Irman. Making spectral shape measurements in inverse Compton scattering a tool for advanced diagnostic applications. *Scientific Reports*, 8(1):1398, 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-19546-0. URL http://www.nature.com/articles/s41598-018-19546-0.

T. Kurz, T. Heinemann, M.F. F. Gilljohann, Y.Y. Y. Chang, J.P. Couperus Cabadağ, A. Debus, O. Kononenko, R. Pausch, S. Schöbel, R.W. W. Assmann, M. Bussmann, H. Ding, J. Götzfried, A. Köhler, G. Raj, S. Schindler, K. Steiniger, O. Zarini, S. Corde, A. Döpp, B. Hidding, S. Karsch, U. Schramm, A. Martinez De La Ossa, A. Irman, J. P. Couperus Cabadağ, A. Debus, O. Kononenko, R. Pausch, S. Schöbel, R.W. W. Assmann, M. Bussmann, H. Ding, J. Götzfried, A. Köhler, G. Raj, S. Schindler, K. Steiniger, O. Zarini, S. Corde, A. Döpp, B. Hidding, S. Karsch, U. Schramm, A. Martinez de la Ossa, and A. Irman. Demonstration of a compact plasma accelerator powered by laser-accelerated electron beams. *Nature Communications - accepted*, 2021. URL http://arxiv.org/abs/1909.06676.

S P D Mangles, C D Murphy, Z Najmudin, a G R Thomas, J L Collier, a E Dangor, E J Divall, P S Foster, J G Gallacher, C J Hooker, D a Jaroszynski, a J Langley, W B Mori, P a Norreys, F S Tsung, R Viskup, B R Walton, and K Krushelnick. Monoenergetic beams of relativistic electrons from intense laser-plasma interactions. *Nature*, 431(7008):535–538, sep 2004. ISSN 0028-0836. doi: 10.1038/nature02939. URL http://www.nature.com/articles/nature02939.

A. Martinez de la Ossa, Z. Hu, M. J. V. Streeter, T. J. Mehrling, O. Kononenko, B. Sheeran, and J. Osterhoff. Optimizing density down-ramp injection for beam-driven plasma wakefield accelerators. *Physical Review Accelerators and Beams*, 20(9):091301, sep 2017. ISSN 2469-9888. doi: 10.1103/PhysRevAccelBeams.20.091301. URL https://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.091301.

Govinda Anantha Padmanabha and Nicholas Zabaras. Solving inverse problems using conditional invertible neural networks. 2020. doi: 10.1016/j.jcp.2021.110194.

A. Pukhov and J. Meyer-ter Vehn. Laser wake field acceleration: The highly non-linear broken-wave regime. *Applied Physics B: Lasers and Optics*, 74(4-5):355–361, 2002. ISSN 0946-2171. doi: 10.1007/s003400200795.

Maziar Raissi. Deep hidden physics models: Deep learning of nonlinear partial differential equations, 2018.

Maziar Raissi and George Em Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. 2017. doi: 10.1016/j.jcp.2017.11.039.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. *arXiv preprint arXiv:1711.10561*, 2017.

Sophie Rudat. Laser Wakefield Acceleration Simulation as a Service, 2019. URL https://doi.org/10.5281/zenodo.3529741.

Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter W. Battaglia. Learning to simulate complex physics with graph networks, 2020.

K. Schmid, A. Buck, C. M.S. Sears, J. M. Mikhailova, R. Tautz, D. Herrmann, M. Geissler, F. Krausz, and L. Veisz. Density-transition based electron injector for laser driven wakefield accelerators. *Physical Review Special Topics - Accelerators and Beams*, 13(9):1–5, 2010. ISSN 10984402. doi: 10.1103/PhysRevSTAB.13.091301.

Justin Sirignano and Konstantinos Spiliopoulos. Dgm: A deep learning algorithm for solving partial differential equations. 2017. doi: 10.1016/j.jcp.2018.08.029.

Klaus Steiniger, Daniel Albach, Michael Bussmann, Markus Loeser, Richard Pausch, Fabian Röser, Ulrich Schramm, Mathias Siebold, and Alexander Debus. Building an optical free-electron laser in the traveling-wave thomson-scattering geometry. *Frontiers in Physics*, 6(JAN):155, jan 2019. ISSN 2296-424X. doi: 10.3389/fphy.2018.00155. URL https://www.frontiersin.org/article/10.3389/fphy.2018.00155/full.

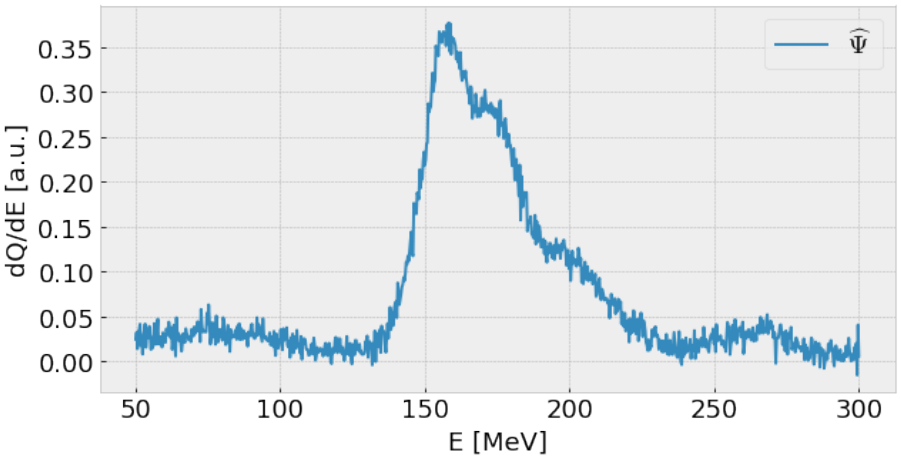
Patrick Stiller, Friedrich Bethke, Maximilian Böhme, Richard Pausch, Sunna Torge, Alexander Debus, Jan Vorberger, Michael Bussmann, and Nico Hoffmann. Large-scale Neural Solvers for Partial Differential Equations. *arXiv preprint arXiv:2009.03730*, 2020.

K. K. Swanson, H.-E. Tsai, S. K. Barber, R. Lehe, H.-S. Mao, S. Steinke, J. van Tilborg, K. Nakamura, C. G. R. Geddes, C. B. Schroeder, E. Esarey, and W. P. Leemans. Control of tunable, monoenergetic laser-plasma-accelerated electron beams using a shock-induced density downramp injector. *Physical Review Accelerators and Beams*, 20(5):051301, may 2017. ISSN 2469-9888. doi: 10.1103/PhysRevAccelBeams.20.051301. URL http://link.aps.org/doi/10.1103/PhysRevAccelBeams.20.051301.

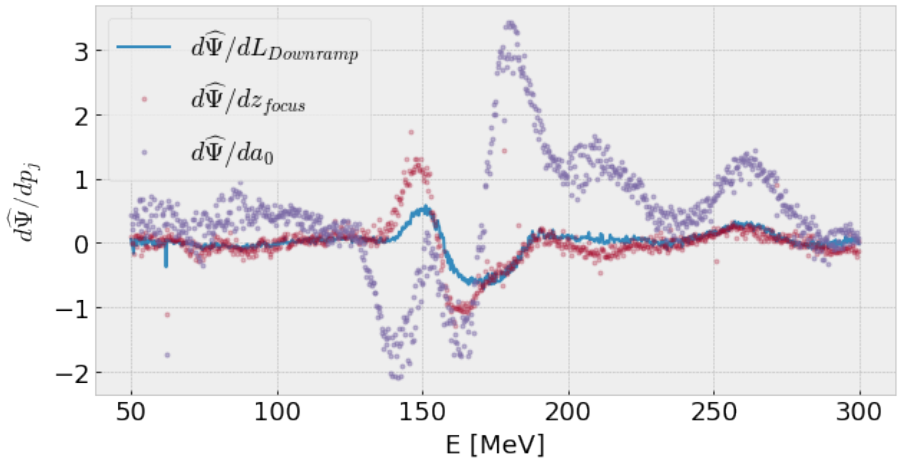
T. Tajima and J. M. Dawson. Laser Electron Accelerator. *Physical Review Letters*, 43(4):267–270, 1979. ISSN 0031-9007. doi: 10.1103/PhysRevLett.43.267. URL https://link.aps.org/doi/10.1103/PhysRevLett.43.267.

You Xie, Erik Franz, Mengyu Chu, and Nils Thuerey. tempoGAN: A Temporally Coherent, Volumetric GAN for Super-resolution Fluid Flow. *ACM Transactions on Graphics (TOG)*, 37(4):95, 2018.

6 SUPPLEMENTARY MATERIAL



(a)



(b)

Figure 4: Figure (a) shows the predicted energy spectrum $\widehat{\Psi}$ at a_0=5, z_{focus}=1e-4, L_{downramp}=3.5e-5. Another advantage of iLWFA is that we are able to unveil the local contribution of certain input parameters p_j with $p_0 = a_0$, $p_1 = z_{\text{focus}}$, $p_2 = L_{\text{downramp}}$ and $j \leq 2$ to $\widehat{\Psi}$. A significant and widespread contribution of e.g. a_0 implies that it might also be easier to reconstruct that parameter since the semantic region is large respectively (b). Additionally, this analysis might also stimulate further physics research on the non-linear contribution of simulation parameter to simulation outcome.