Accepted manuscript

Laser Wakefield Accelerator modelling with Variational Neural Networks

M.J.V. Streeter,[1,*] C. Colgan,[2] C.C. Cobo,[3] C. Arran,[3] E.E. Los,[2] R. Watt,[2] N. Bourgeois,[4] L. Calvin,[1] J. Carderelli,[5] N. Cavanagh,[1] S.J.D. Dann,[4] R. Fitzgarrald,[5] E. Gerstmayr,[2] A.S. Joglekar,[5,6] B. Kettle,[2] P. McKenna,[7] C.D. Murphy,[3] Z. Najmudin,[2] P. Parsons,[4] Q. Qian,[5] P.P. Rajeev,[4] C.P. Ridgers,[3] D.R. Symes,[4] A.G.R. Thomas,[5] G. Sarri,[1] and S.P.D. Mangles[2]

[1]*School of Mathematics and Physics, Queen's University Belfast, BT7 1NN, Belfast UK*
[2]*The John Adams Institute for Accelerator Science, Imperial College London, London, SW7 2AZ, UK*
[3]*York Plasma Institute, School of Physics, Engineering and Technology, University of York, York YO10 5DD, UK*
[4]*Central Laser Facility, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK*
[5]*Gérard Mourou Center for Ultrafast Optical Science, University of Michigan, Ann Arbor, MI 48109-2099, USA*
[6]*Ergodic LLC, San Francisco, CA 94117, USA*
[7]*Department of Physics, SUPA, University of Strathclyde, Glasgow G4 0NG, UK*

(Dated: December 7, 2022)

A machine learning model was created to predict the electron spectrum generated by a GeV-class laser wakefield accelerator. The model was constructed from variational convolutional neural networks which mapped the results of secondary laser and plasma diagnostics to the generated electron spectrum. An ensemble of trained networks was used to predict the electron spectrum and to provide an estimation of the uncertainty on that prediction. It is anticipated that this approach will be useful for inferring the electron spectrum prior undergoing any process which can alter or destroy the beam. In addition, the model provides insight into the scaling of electron beam properties due to stochastic fluctuations in the laser energy and plasma electron density.

INTRODUCTION

Laser wakefield accelerators (LWFAs) generate multi-GeV electron beams from cm-scale plasma channels using $\sim 100\,\mathrm{TW}$ laser pulses [1–6]. The extreme acceleration gradients of LWFA, coupled with their relative accessibility, has led to widespread research and pursuit of several applications, such as compact light sources [7–10], generation of bright γ-ray [11] and ultra-relativistic positron beams [12], and for future particle colliders [13]. Also, the combination of GeV electron beams and high intensity laser pulses allows for the study of fundamental physics such as strong-field quantum electrodynamics [14–17].

In LWFA, the non-linear laser pulse evolution [18, 19] and its effect on the injection and acceleration processes [20–23] are highly sensitive to initial conditions and can lead to significant shot-to-shot variation of the electron beam properties [24, 25]. Recent work on high-stability laser systems and plasma sources has demonstrated improved stability, with the observation of few-percent variation in electron beam energy and charge over 24-hours of continuous operation [26]. Long-term high-repetition rate operation has opened up the possibility of using machine learning techniques to model the sources of electron beam variation and to use closed-loop algorithms to optimise performance [26–31].

For applications such as the study of radiation reaction, knowledge of the pre-interaction electron beam properties are required to make precise measurements of any changes of these properties and thereby infer the validity of theoretical models [32–34]. The destructive nature of the measurements necessitates predictable LWFA performance through either: improved stability; preserving part of the spectrum as a reference [33]; or by developing models capable of producing the electron beam properties from a given shot. In general, the ability to make predictions of the outputs from plasma accelerators will be advantageous to many of their applications.

Previous work in developing machine learning models for LWFAs has demonstrated prediction of scalar metrics of the electron beam, such as total charge or peak energy [29–31, 35]. However, many applications will require the prediction of vector properties, such as the spectrum or the longitudinal phase space, for which neural networks provide a convenient framework. A densely connected neural network (DNN) is made of densely connected layers, in which every input is the weighted sum of all of the outputs of the previous layer, with the individual weights as free parameters of the model. A non-linear activation function then (e.g. a sigmoid function) takes the weighted sum plus a bias value (another free model parameters) as its argument and returns an output value. An alternative to deeply connected layers is a convolutional layer, which performs convolutions between the input vector and a set of kernels. Networks using these layers, known as convolutional neural networks (CNN) have been shown to be better suited to learning meaningful features from natural signals [36]. Further improvement to the predictive power of neural networks has been seen when including stochasticity in the outputs of individual nodes, in an architecture known as variational neural networks (VNN) [37].

In conventional accelerators, Emma *et al.* [38] demonstrated training of a DNN to produce synthetic diagnostic outputs that matched the measured outputs for a new unseen dataset. CNNs have been used to predict x-ray

10.1017/hpl.2022.47

properties from the post-undulator electron beam spectrum [39], while ensembles of DNNs have also been used to predict the electron beam longitudinal phase space and current profile from non-destructive bending radiation measurements[40]. In this work, we report on the training of an ensemble of VNNs to model the LWFA-generated electron spectrum using secondary diagnostics of the laser and plasma conditions. The LWFA ensemble was trained using a subset of experimental measurements of the electron spectrum with the remainder used for model validation. Each individual VNN in the ensemble was trained with a different subset of the training data, so that the ensemble provided both a mean prediction and an estimate of its uncertainty. The model also reveals the extent to which the measurements obtained from the available diagnostics are predictive of the accelerator performance, and which parameters have the strongest influence.

EXPERIMENTAL METHODS AND RESULTS

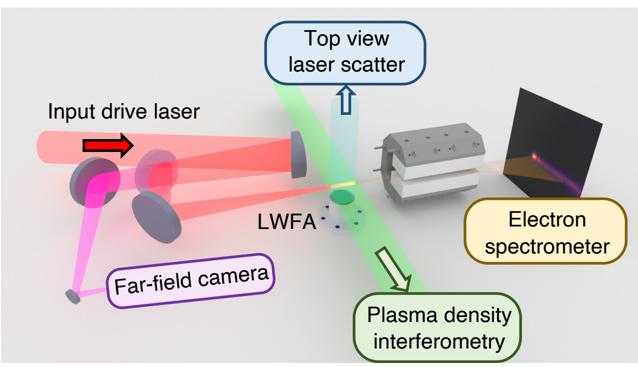

FIG. 1. An illustration of the experimental setup (not to scale). The primary laser focus was aligned to the front edge of a supersonic gas jet emitted from a 15 mm diameter nozzle positioned 10 mm below the laser pulse propagation axis. The input laser energy, was measured by integrating the signal on a near-field camera before the compressor, which was cross calibrated with an energy meter and adjusted for the 60% compressor throughput. The scattered laser signal was observed from above by an optical camera, and the plasma channel electron density profile was measured using interferometry with a transverse short-pulse probe laser. The small ($\lesssim 0.1\%$) transmission of the focusing laser pulse through a dielectric mirror was directed onto a CCD camera to obtain an on-shot far-field image. Electron beams from the LWFA were deflected by a magnetic dipole onto two lanex screens (only the first is shown here) which were used to determine the electron spectrum in the range of $0.3 < E < 2.5\,\mathrm{GeV}$.

The experiment was performed using the Gemini laser system at the Central Laser Facility in the UK, (see figure 1 for details). Laser pulses with an energy of $E_L = (6.6 \pm 0.5)\,\mathrm{J}$ and a pulse duration of $\approx 50\,\mathrm{fs}$ were used to drive a GeV-scale LWFA. The pulses were focused with an $f/40$ off-axis parabolic mirror to a spot size of $(50 \pm 2) \times (45 \pm 2)\,\mu\mathrm{m}$ in the horizontal (polarisation) and vertical planes respectively, giving a peak intensity of $(5.5 \pm 0.5) \times 10^{18}\,\mathrm{Wcm}^{-2}$. The focus was aligned to a gas jet which was composed of a mixture of 2% nitrogen and 98% helium, enabling ionisation injection [41–44]. The gas jet had an average electron density of $(1.00 \pm 0.07) \times 10^{18}\,\mathrm{cm}^{-3}$ over a 17 mm length.

The LWFA generated electron energy spectrum $\mathrm{d}W/\mathrm{d}E$ was measured using the spectrometer scintillator screen images, which were energy-calibrated by numerical tracking of electron trajectories in the magnetic field. The interferometry and top view cameras were used to extract the electron density profile, $n_e(z)$ and the laser scatter profile, $S_L(z)$, respectively, where z is the laser propagation axis. A 2D Gaussian fit was performed on the far-field image to obtain six parameters: the peak fluence I_0; the centroids x_0 and y_0; the major and minor RMS spot widths σ_a and σ_b; and the angle of the major axis of the ellipse with x axis θ. Due to the aberrations and clips caused by this beam-line, this far-field is not an exact replica of the main laser focus, but is representative of the shot-to-shot focal spot fluctuations.

The experimental results for this analysis were taken from an investigation of radiation reaction, in which a second counter-propagating laser pulse used to collide with the LWFA electron beam. For training and validating our predictive tool, we wish to only use shots where the laser pulse did not significantly overlap with the electron beam, so that the electron spectrum was not affected. For successful collisions, a gamma-beam was generated via the inverse Compton scattering interaction and was diagnosed spatially with a CsI scintillator array [16] imaged onto a 1024×1024 pixel CCD.

Due to the shot-to-shot variation in the electron beam position, most shots did not result in a significant collision, providing a large number of null shots for model training and testing. The brightness of the signal on the gamma detector was used to provide an approximate metric of the collision intensity. The 99.99^{th} percentile pixel value of the background subtracted CCD image was taken as the peak of the the gamma signal C_γ. The highest value of this metric was $C_\gamma = 4380$, whereas the median value was $C_\gamma = 12$. From analysis of the collision statistics, a value of $C_\gamma \leq 100$ was estimated to result from collision with a peak normalised vector potential of $a_0 < 1.4$. For 1 GeV electrons, this would result in a $< 1\%$ energy loss [14], approximately equal to the resolution of the spectrometer. Therefore, this value was taken as a threshold for null shots, for which the electron beam is unaffected by the collision. The experimental data was taken during a 5-hour period with a total of 779 shots. Model training and validation datasets were taken from shots for which $C_\gamma \leq 100$, with 90% (570 shots) used for training and 10% (75 shots) reserved for model validation.

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

NEURAL NETWORK ARCHITECTURE AND TRAINING

The measurements of $n_e(z)$, $S_L(z)$ and $\mathrm{d}W/\mathrm{d}E$ were stored as one-dimensional vectors of lengths 310, 100 and 200 respectively. Although each of these signals are composed of at least a hundred values, the variations over the full dataset is limited, and so in principle only a few parameters are required for each to encode these variations. An appropriate decoder would be able to generate a good approximation to the measured signals from this reduced set of parameters, which are called *latent space* variables. In this work, variational autoencoders (VAE)[45, 46] incorporating convolutional and densely connected layers, were trained as illustrated in figure 2. By using a bottleneck of only a few nodes, the VAEs were trained to find an optimal latent space representation which of the data which allowed the decoder to reconstruct the measured signals.

The trained encoders for $n_e(z)$ and $S_L(z)$ were used to encode their respective measurements to their latent space representation, which were then combined with measurements of the laser far-field and the laser energy to create the inputs for the predictive model. A VNN, which we call the *translator* network, takes those inputs and returns values which are passed to the trained electron spectra decoder to generate the predicted spectrum. The translator was trained to the learn the correlation between the reduced input set and the latent variables of the electron spectra decoder, as illustrated in figure 3.

For the variational layers, two parameters are calculated for each node which represent the expectation value μ_m and standard deviation σ_m. During training, values were sampled from Gaussian distributions given by these parameters, $\mathcal{N}(\mu_m, \sigma_m)$, such that the latent values for a given input set, x_m, would vary according to σ_m.

The training loss function used was [45],

$$\begin{aligned}\mathcal{L}_T &= \mathcal{L}_{MSE} - \beta D_{KL} \\ \mathcal{L}_T &= \frac{1}{N}\sum_{n=0}^{N}\left[W(E_n) - W_R(E_n)\right]^2 - \beta D_{KL}\ , \end{aligned} \tag{1}$$

where $D_{KL} = \sum_{m=0}^{M}(1 + \log(\sigma_m) - \mu_m^2 - \sigma_m)/(2M)$, is the Kullback-Leibler (KL) divergence and $\mathcal{L}_{MSE}$ is the mean squared error (MSE), and M is the total number of input sets in a given training iteration. The same loss function was used to train each VAE and also the final translator VNN, with the MSE taken between the predicted and measured diagnostic output ($n_e(z)$, $S_L(z)$ or $\mathrm{d}W/\mathrm{d}E$). The β parameter was used to scale the relative importance of the regularisation, following the beta-VAE approach [45]. During model validation, only the mean weights for the the variational layers were used and the D_{KL} term from equation (1) was omitted. Every node of the neural networks used the leaky rectified linear unit (leaky-ReLU)[47] activation function with $\alpha = 0.3$, which exhibited superior learning performance in comparison to sigmoid and hyperbolic tan functions, as well as leaky-ReLU with other values of α.

For the diagnostic VAEs, the number of latent parameters was chosen to be the minimum which give high-fidelity reconstructions, with the β parameter manually tuned to ensure that the distribution of each latent parameter for the training datasets was close to a standard normal distribution ($\mathcal{N}(0, 1)$). One latent space parameter was directly set as the average of the input signal (normalised by the training dataset). This parameter was then used to scale the decoder output and ensured that one of the latent space variables represents the amplitude of the signal, aiding interpretation of the trained networks. Once the VAEs were trained, the weights were frozen during the translator training process.

The translator is a densely connected neural network with a variational last layer. The translator VNN architecture (number of nodes and number of layers) and the value of β was optimised using a genetic algorithm. During this process the training data was divided in two, with 50% of the data used to train each network, and the other 50% used to calculate the test loss. This ensured that the validation dataset was kept purely for validation of the final model performance and not used in any tuning of the predictive model. The optimal architecture for the translator network, shown in figure 3, is comprised of three densely connected layers, with a final variational layer with five outputs.

In order to quantify the uncertainty in the model predictions, 100 translator VNNs were trained, each using a randomly selected 50% sample of the training dataset. The prediction of each of these models can then be used to obtain an average prediction, while the variation between model predictions is indicative of the random uncertainty and the finite size of the training data. In particular, the random sub-sampling affects the predictive quality in regions where the training data is sparse, typically at the extremes of the input parameters, resulting in a larger uncertainty in those regions.

The parameters for the trained VAEs and translator networks are summarised in table I. Each autoencoder was trained for 1000 iterations with a batch size of 64. The translator network was trained in three stages with 200, 400 and 300 iterations performed at 10, 4 and 1 times the final β value to balance reconstruction fidelity with latent space smoothness [46]. The training processes were all performed using the Adam optimiser [48], with a learning rate of 10^{-3}, which was found to converge well.

LWFA PREDICTION RESULTS

The measured electron spectra from the validation dataset are shown in figure 4a, along with the reconstruc-

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

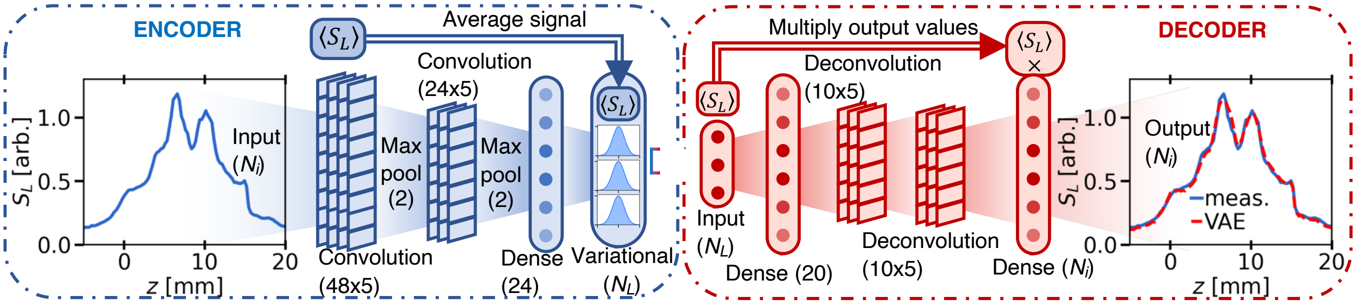

FIG. 2. Variational autoencoder (VAE) architecture for determining latent space representation of diagnostics. The type and dimension of each layer is indicated in the labels. The inset plots show an example laser scatter signal S_L and the approximation returned by the VAE. The input (and output) size N_i is equal to the data binning of the results for each individual diagnostic. Max pooling was used at the output of each convolution layer, which combined neighbouring output pairs and returned only the maximum of each pair. The average signal, in this case $\langle S_L \rangle$, was passed as an additional latent space parameter for the encoder and was used the scale the output of the decoder. The autoencoder structure was the same for each diagnostic, except for the size of the latent space.

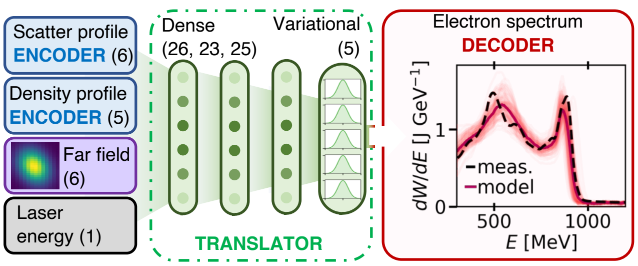

FIG. 3. A diagram of the translator network architecture. Shown inset is an example measurement from the experimental data (black), with the mean prediction of the LWFA model ensemble (red). Also shown are the individual predictions of each sub-model in blue.

Model	N_i	N_L	β	Validation $\mathcal{L}_{MSE}$
Density profile	310	4+1	2×10^{-3}	1.7×10^{-3}
Scattering profile	100	5+1	10^{-3}	2.3×10^{-3}
Electron spectra	200	4+1	2×10^{-3}	1.1×10^{-2}
LWFA single	18	5	* 5×10^{-4}	$(7.3 \pm 0.5) \times 10^{-2}$
LWFA ensemble	18	5	*5×10^{-4}	5.7×10^{-2}

TABLE I. Summary of autoencoder parameters used for each diagnostic and for the translator model. *For the LWFA translator models, the value of β varied from high to low during the training, with the final value given in the table. The training time for each autoencoder was 10 minutes and training of the 100 translator networks took a total of 3 hours, using an Intel Xeon Gold 6130 CPU at 2.1 GHz with 32 Gb of RAM. The analysis and model training was performed on the CLF Data Analysis as a Service (CDAS) [49]. The neural networks were built using the Keras API (https://keras.io).

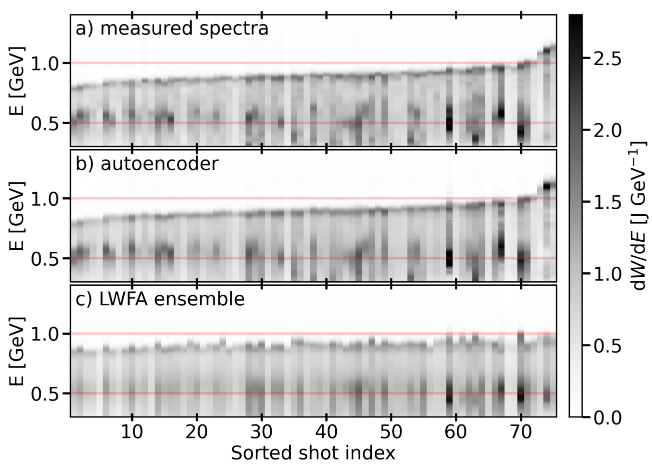

FIG. 4. a) Measured electron spectra and reproduced electron spectra using b) the trained variational autoencoder and c) the mean prediction of the ensemble of LWFA models. The individual shots are sorted by cut-off energy, determined as the highest energy for which the spectra exceed a threshold value.

tions by the electron spectra VAE figure 4b and the average of the LWFA model ensemble predictions figure 4c. The electron spectra VAE had an MSE of 0.011, and shows a good qualitative and quantitative reproduction of the measured electron spectra. This indicates that the five parameters of the latent space, in combination with the structures learnt by the decoder, are sufficient to accurately generate the set of observations from the validation dataset. In other words, the five latent parameters are sufficient to generate the full variability of electron beams for this experimental setup. The question is then whether the secondary diagnostics are sufficient to determine the correct latent variables for each shot and thereby give an accurate prediction of the electron spectrum. The mean prediction of the LWFA model ensemble had an MSE of 0.057 and shows a similar trend in cutoff energy as the data, except for the few high and low energy outliers. By comparison, a naive prediction that all measured spectra are equal to the the average spectrum from the training dataset gives an MSE value

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

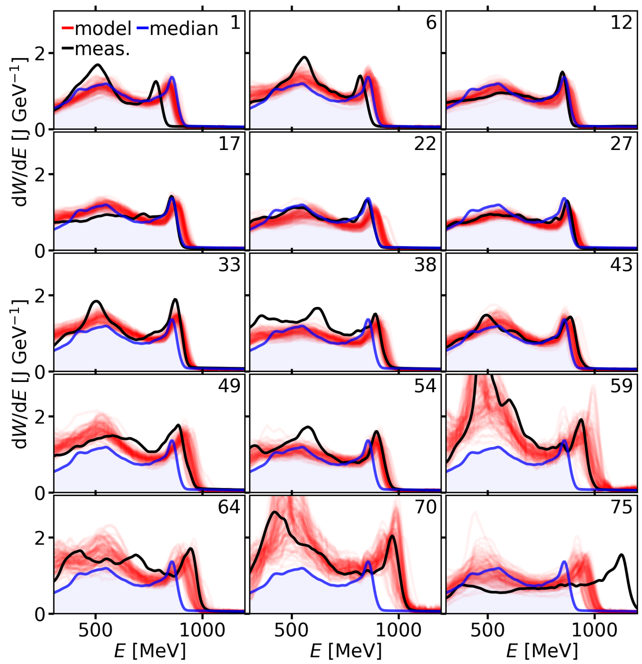

FIG. 5. Individual shots selected at equally spaced intervals of the sorted shot index from figure 4. The measured spectra (black) are shown alongside the predictions of each LWFA model from the trained ensemble (red) and a individual spectrum measurement closest to the median of the training data (blue). The sorted shot index is shown in the top right of each panel.

of 0.11, indicating that the LWFA model has a significant predictive capability.

Individual predictions of each model of the LWFA ensemble, along with the corresponding measured electron spectra are shown in figure 5. The variation in model predictions for a given shot is indicative of the uncertainty, due to the random sub-sampling of the training data and the stochastic training process. For a large region of the parameter space, the LWFA model predictions show a good agreement with the measurements, with large discrepancies occurring for the outliers in terms of cutoff energy. These shots also exhibit the largest variation in predictions between individual models within the ensemble. The total electron beam energy is reasonably accurately predicted with relative RMS error of 12% for the entire validation dataset, compared to the relative beam energy RMS variation of 30%.

The relative influence of each input parameter to the LWFA model can be seen by varying each one in turn and measuring the effect on the resultant spectra as shown in figure 6. The plasma density parameters have a relatively modest effect on the electron spectrum, indicating that the shot-to-shot variation of the plasma density profile is not the dominant contributor to the electron spectrum variation. Variations of the laser energy and the scattering profile are more significant, having the greatest effect on the generated electron spectra. The spatio-temporal distribution of the laser pulse is only indirectly diagnosed from the far-field diagnostic and the effect on the scattering profile, and is known to have a large influence on the accelerated electrons [26, 28, 29]. Including additional laser diagnostics, such as measurement of the spatial phase profile [26, 30], should enable higher fidelity predictions.

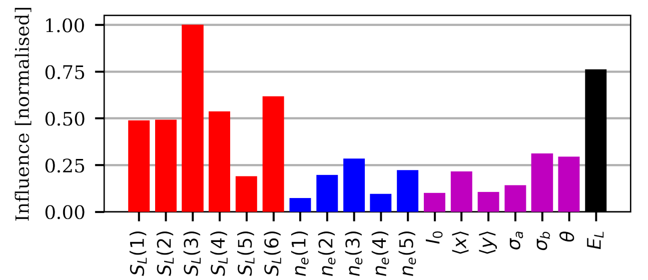

FIG. 6. The relative influence of the translator VNN input parameters on the predicted electron spectra. Each parameter is set to the mean value of the training dataset and then varied over ±3 standard deviations in 11 steps, with the variation in the spectrum quantified by the average RMS change to the spectrum. The n^{th} latent space parameters for the scattering and density profile encoders are labelled $S_L(n)$ and $n_e(n)$ respectively. $S_L(6)$ and $n_e(5)$ are proportional to the average laser scattering signal and plasma electron density respectively.

Although many of the input parameters are not straightforward to interpret physically, i.e. those which are the latent space of the autoencoders, the laser energy is a physically important parameter in LWFA. In practice, the inputs for the LWFA models are not independent of one another, as characterised by calculating the Pearson correlation coefficients for the training dataset. This reveals relatively strong correlations between the laser energy and several other parameters, especially $S_L(3)$, $S_L(4)$, $S_L(6)$, $n_e(4)$, $n_e(5)$ and I_0 which had correlation coefficients ranging from $r = 0.31$ to $r = 0.55$. The trained LWFA model is then able to show what effect laser energy fluctuations have on the electron spectrum by varying each parameter proportionally according to their correlation coefficients with laser energy E_L, as shown in figure 7a. As the laser energy increases, the peak electron energy is relatively constant, while the overall charge increases. The total electron beam charge Q_B is plotted as a function of laser energy in figure 7b, for both the raw data and the LWFA model predictions. The model prediction shows an approximately linear increase with laser energy with the equation $Q_B[\mathrm{nC}] = 0.48E_L[\mathrm{J}] - 2.1$.

The scaling parameters $S_L(6)$ and $n_e(5)$ are also easy to interpret, as they the average scattering signal and electron density respectively (normalised to the mean and variance over the training data set). The effect of $n_e(5)$

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

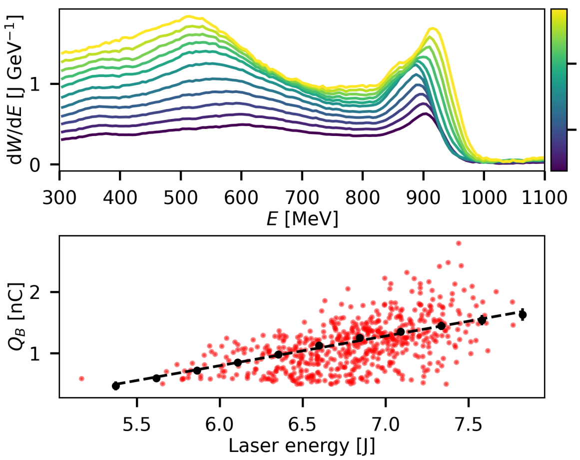

FIG. 7. The model predicted effect of varying the laser energy on a) the predicted electron spectra and b) the total electron beam charge. The data for each shot in the training data (red) is shown in b), overlaid from the values calculated from the predicted spectra of the LWFA model (black points) with a linear fit (black dashed line).

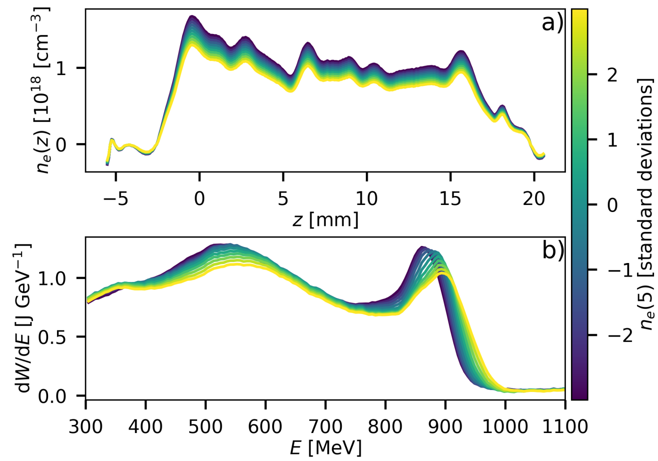

FIG. 8. The effect of changing $n_e(5)$ on a) the electron density profile and b) the predicted electron spectrum. All other latent space parameters are kept fixed at zero (i.e. their average values from the training dataset) while $n_e(5)$ is varied over the range of ± 3 standard deviations in the training dataset.

on the electron density profile and the predicted electron spectrum are shown in figure 8. The average plasma electron density varied by 4% over the training data set, as illustrated by the small perturbations to the density profile observed in figure 8a. A more significant effect is seen on the electron spectra in figure 8b, with the peak energy shifting higher as the average density drops, as expected for a dephasing limited LWFA [50, 51]. The effect on the spectrum is much smaller than that seen to be caused by the laser energy variation in figure 7. This indicates that the level of natural variations of the plasma electron density in this dataset was sufficiently low that it was not a dominant contributor to the shot-to-shot variations in the electron spectra.

The other latent parameters generated by VAEs do not have straightforward physical interpretations and only have meaning in combination with the trained encoders. In order to gain some insight into their physical meaning, the effect of changing each parameter can be observed on the corresponding diagnostic output, as well as on the predicted electron spectrum. An example is shown in figure 9, where the effect of varying $S_L(3)$, the most dominant input parameter to the translator VNN, is shown.

Figure 9a shows that positive $S_L(3)$ correlates with an increased laser scattering peak at the entrance to the gas jet ($z = 0$) and for the last half of the plasma, while suppressing the signal for $1 > z > 7$ mm. This also results in an increased predicted total charge as well as an increased predicted maximum electron energy (see figure 9b, a clearly beneficial effect for many applications. The scattered laser intensity is associated with Raman side-scattering and wavebreaking radiation, generated as

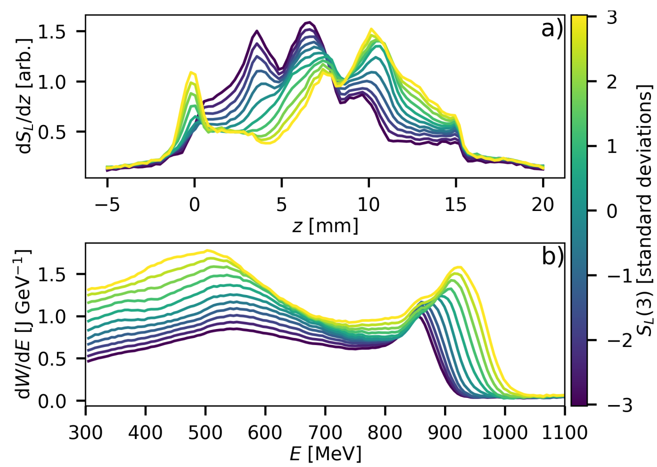

FIG. 9. The effect of changing $S_L(3)$ on a) the laser scattering profile and b) the predicted electron spectrum. All other latent space parameters are kept fixed at zero (i.e. their average values from the training dataset) while $S_L(3)$ is varied over the range of ± 3 standard deviations of $S_L(3)$ in the training dataset.

the laser self-guides and self-compresses to a high peak intensity in the plasma channel [52, 53]. Therefore, the increase of this scattering signal seen in figure 9a indicates an increased possibility for the injection of electrons into the plasma wakefield at $z = 0$ mm, while maintaining a high amplitude plasma wave for $z > 7$ mm, resulting in the enhanced electron spectrum predicted in figure 9b.

CONCLUSION

In conclusion, we have constructed and trained a predictive model for a laser wakefield accelerator, capable

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

of predicting the electron spectrum for a given shot, based on secondary diagnostics of the laser and plasma conditions. The model is constructed from separately trained variational convolutional autoencoders, with a variational neural network used to map a reduced parameter set to the latent space of an electron spectra decoder. An ensemble of models were trained on sub-sets of the training data, with the range of model predictions providing an estimate of the uncertainty. The predictive model ensemble performs better than the naive assumption that the electron spectrum is constant, and so has utility in estimating the electron spectrum in the case of destructive processes, such as radiation reaction. The model fidelity is most likely limited by the lack of on-shot spatio-temporal information about the laser pulse, which is known to have a strong influence on the accelerated electron beam [26]. It is expected that this technique can be improved by including additional diagnostics of the laser spatial and spectral phase, and by increasing the size of the training dataset, especially for reducing the prediction error for the outliers. Further diagnostics of the laser-plasma interaction, such as spectrally resolving the scattering signal, may also provide additional information to improve the prediction accuracy. Neural networks of this kind could be an important tool for understanding the performance sensitivities of plasma accelerators, and also in providing synthetic diagnostics for applications of their electron beams and secondary sources.

This work was supported by UK STFC ST/V001639/1, UK EPSRC EP/V049577/1 and Horizon 2020 funding under European Research Council (ERC) Grant Agreement No. 682399. M.J.V.S. acknowledges support from the Royal Society URF-R1221874. A.G.R.T and A.S.J acknowledge support from the US DOE grant DE-SC0016804.

* m.streeter@qub.ac.uk

[1] W. P. Leemans, B. Nagler, A. J. Gonsalves, C. Tóth, K. Nakamura, C. G. R. Geddes, E. Esarey, C. B. Schroeder, and S. M. Hooker, Nature Physics **2**, 696 (2006).
[2] S. Kneip, S. R. Nagel, S. F. Martins, S. P. D. Mangles, C. Bellei, O. Chekhlov, R. J. Clarke, N. Delerue, E. J. Divall, G. Doucas, K. Ertel, F. Fiuza, R. Fonseca, P. Foster, S. J. Hawkes, C. J. Hooker, K. Krushelnick, W. B. Mori, C. A. J. Palmer, K. Ta Phuoc, P. P. Rajeev, J. Schreiber, M. J. V. Streeter, D. Urner, J. Vieira, L. O. Silva, and Z. Najmudin, Physical Review Letters **103**, 035002 (2009).
[3] C. E. Clayton, J. E. Ralph, F. Albert, R. A. Fonseca, S. H. Glenzer, C. Joshi, W. Lu, K. A. Marsh, S. F. Martins, W. B. Mori, A. Pak, F. S. Tsung, B. B. Pollock, J. S. Ross, L. O. Silva, and D. H. Froula, Physical Review Letters **105**, 105003 (2010).
[4] X. Wang, R. Zgadzaj, N. Fazel, Z. Li, S. A. Yi, X. Zhang, W. Henderson, Y.-Y. Chang, R. Korzekwa, H.-E. Tsai, C.-H. Pai, H. Quevedo, G. Dyer, E. Gaul, M. Martinez, A. C. Bernstein, T. Borger, M. Spinks, M. Donovan, V. Khudik, G. Shvets, T. Ditmire, and M. C. Downer, Nature Communications **4**, 1988 (2013).
[5] W. P. Leemans, A. J. Gonsalves, H.-S. Mao, K. Nakamura, C. Benedetti, C. B. Schroeder, C. Tóth, J. Daniels, D. E. Mittelberger, S. S. Bulanov, J.-L. Vay, C. G. R. Geddes, and E. Esarey, Physical Review Letters **113**, 245002 (2014).
[6] A. J. Gonsalves, K. Nakamura, J. Daniels, C. Benedetti, C. Pieronek, T. C. H. de Raadt, S. Steinke, J. H. Bin, S. S. Bulanov, J. van Tilborg, C. G. R. Geddes, C. B. Schroeder, C. Tóth, E. Esarey, K. Swanson, L. Fan-Chiang, G. Bagdasarov, N. Bobrova, V. Gasilov, G. Korn, P. Sasorov, and W. P. Leemans, Physical Review Letters **122**, 084801 (2019).
[7] H. P. Schlenvoigt, K. Haupt, A. Debus, F. Budde, O. Jäckel, S. Pfotenhauer, H. Schwoerer, E. Rohwer, J. G. Gallacher, E. Brunetti, R. P. Shanks, S. M. Wiggins, and D. A. Jaroszynski, Nature Physics **4**, 130 (2008).
[8] A. R. Maier, A. Meseck, S. Reiche, C. B. Schroeder, T. Seggebrock, and F. Grüner, Physical Review X **2**, 031019 (2012).
[9] F. Albert and A. G. Thomas, "Applications of laser wakefield accelerator-based light sources," (2016).
[10] W. Wang, K. Feng, L. Ke, C. Yu, Y. Xu, R. Qi, Y. Chen, Z. Qin, Z. Zhang, M. Fang, J. Liu, K. Jiang, H. Wang, C. Wang, X. Yang, F. Wu, Y. Leng, J. Liu, R. Li, and Z. Xu, Nature **595**, 516 (2021).
[11] G. Sarri, W. Schumaker, A. Di Piazza, M. Vargas, B. Dromey, M. E. Dieckmann, V. Chvykov, A. Maksimchuk, V. Yanovsky, Z. H. He, B. X. Hou, J. A. Nees, A. G. Thomas, C. H. Keitel, M. Zepf, and K. Krushelnick, Physical Review Letters **110**, 255002 (2013).
[12] G. Sarri, D. J. Corvan, W. Schumaker, J. M. Cole, A. Di Piazza, H. Ahmed, C. Harvey, C. H. Keitel, K. Krushelnick, S. P. D. Mangles, Z. Najmudin, D. Symes, A. G. R. Thomas, M. Yeung, Z. Zhao, and M. Zepf, Physical Review Letters **113**, 224801 (2014).
[13] B. Cros and P. Muggli, arXiv (2019).
[14] A. G. R. Thomas, C. P. Ridgers, S. S. Bulanov, B. J. Griffin, and S. P. D. Mangles, Physical Review X **2**, 041004 (2012).
[15] T. G. Blackburn, C. P. Ridgers, J. G. Kirk, and A. R. Bell, Physical Review Letters **112**, 015001 (2014).
[16] J. M. Cole, K. T. Behm, E. Gerstmayr, T. G. Blackburn, J. C. Wood, C. D. Baird, M. J. Duff, C. Harvey, A. Ilderton, A. S. Joglekar, K. Krushelnick, S. Kuschel, M. Marklund, P. McKenna, C. D. Murphy, K. Poder, C. P. Ridgers, G. M. Samarin, G. Sarri, D. R. Symes, A. G. R. Thomas, J. Warwick, M. Zepf, Z. Najmudin, and S. P. D. Mangles, Physical Review X **8**, 011020 (2018).
[17] K. Poder, M. Tamburini, G. Sarri, A. Di Piazza, S. Kuschel, C. D. Baird, K. Behm, S. Bohlen, J. M. Cole, D. J. Corvan, M. Duff, E. Gerstmayr, C. H. Keitel, K. Krushelnick, S. P. D. Mangles, P. McKenna, C. D. Murphy, Z. Najmudin, C. P. Ridgers, G. M. Samarin, D. R. Symes, A. G. R. Thomas, J. Warwick, and M. Zepf, Physical Review X **8**, 031004 (2018).
[18] A. G. R. Thomas, Z. Najmudin, S. P. D. Mangles, C. D. Murphy, A. E. Dangor, C. Kamperidis, K. L. Lancaster,

W. B. Mori, P. A. Norreys, W. Rozmus, and K. Krushelnick, Physical Review Letters **98**, 095004 (2007).

[19] M. J. V. Streeter, S. Kneip, M. S. Bloom, R. A. Bendoyro, O. Chekhlov, A. E. Dangor, A. Döpp, C. J. Hooker, J. Holloway, J. Jiang, N. C. Lopes, H. Nakamura, P. A. Norreys, C. A. J. Palmer, P. P. Rajeev, J. Schreiber, D. R. Symes, M. Wing, S. P. D. Mangles, and Z. Najmudin, Physical Review Letters **120**, 254801 (2018).

[20] C. Xia, J. Liu, W. Wang, H. Lu, W. Cheng, A. Deng, W. Li, H. Zhang, X. Liang, Y. Leng, X. Lu, C. Wang, J. Wang, K. Nakajima, R. Li, and Z. Xu, Physics of Plasmas **18**, 113101 (2011).

[21] A. Sävert, S. P. D. Mangles, M. Schnell, E. Siminos, J. M. Cole, M. Leier, M. Reuter, M. B. Schwab, M. Möller, K. Poder, O. Jäckel, G. G. Paulus, C. Spielmann, S. Skupin, Z. Najmudin, and M. C. Kaluza, Physical Review Letters **115**, 055002 (2015).

[22] J. Wang, J. Feng, C. Zhu, Y. Li, Y. He, D. Li, J. Tan, J. Ma, and L. Chen, Plasma Physics and Controlled Fusion **60**, 034004 (2018).

[23] C. J. Zhang, Y. Wan, B. Guo, J. F. Hua, C.-H. Pai, F. Li, J. Zhang, Y. Ma, Y. P. Wu, X. L. Xu, W. B. Mori, H.-H. Chu, J. Wang, W. Lu, and C. Joshi, Plasma Physics and Controlled Fusion **60**, 044013 (2018).

[24] J. Osterhoff, A. Popp, Z. Major, B. Marx, T. Rowlands-Rees, M. Fuchs, M. Geissler, R. Hörlein, B. Hidding, S. Becker, E. Peralta, U. Schramm, F. Grüner, D. Habs, F. Krausz, S. Hooker, and S. Karsch, Physical Review Letters **101**, 085002 (2008).

[25] N. A. M. Hafz, T. M. Jeong, I. W. Choi, S. K. Lee, K. H. Pae, V. V. Kulagin, J. H. Sung, T. J. Yu, K.-h. Hong, T. Hosokai, J. R. Cary, D.-k. Ko, and J. Lee, Nature Photonics **2**, 571 (2008).

[26] A. R. Maier, N. M. Delbos, T. Eichner, L. Hübner, S. Jalas, L. Jeppe, S. W. Jolly, M. Kirchen, V. Leroux, P. Messner, M. Schnepp, M. Trunk, P. A. Walker, C. Werle, and P. Winkler, Physical Review X **10**, 031039 (2020).

[27] Z.-H. He, B. Hou, V. Lebailly, J. A. Nees, K. Krushelnick, and A. G. R. Thomas, Nature communications **6**, 7156 (2015).

[28] S. J. D. Dann, C. D. Baird, N. Bourgeois, O. Chekhlov, S. Eardley, C. D. Gregory, J.-N. Gruse, J. Hah, D. Hazra, S. J. Hawkes, C. J. Hooker, K. Krushelnick, S. P. D. Mangles, V. A. Marshall, C. D. Murphy, Z. Najmudin, J. A. Nees, J. Osterhoff, B. Parry, P. Pourmoussavi, S. V. Rahul, P. P. Rajeev, S. Rozario, J. D. E. Scott, R. A. Smith, E. Springate, Y. Tang, S. Tata, A. G. R. Thomas, C. Thornton, D. R. Symes, and M. J. V. Streeter, Physical Review Accelerators and Beams **22**, 041303 (2019).

[29] R. J. Shalloo, S. J. Dann, J. N. Gruse, C. I. Underwood, A. F. Antoine, C. Arran, M. Backhouse, C. D. Baird, M. D. Balcazar, N. Bourgeois, J. A. Cardarelli, P. Hatfield, J. Kang, K. Krushelnick, S. P. Mangles, C. D. Murphy, N. Lu, J. Osterhoff, K. Põder, P. P. Rajeev, C. P. Ridgers, S. Rozario, M. P. Selwood, A. J. Shahani, D. R. Symes, A. G. Thomas, C. Thornton, Z. Najmudin, and M. J. Streeter, Nature Communications **11**, 1 (2020).

[30] S. Jalas, M. Kirchen, P. Messner, P. Winkler, L. Hübner, J. Dirkwinkel, M. Schnepp, R. Lehe, and A. R. Maier, Physical Review Letters **126**, 104801 (2021).

[31] M. Kirchen, S. Jalas, P. Messner, P. Winkler, T. Eichner, L. Hübner, T. Hülsenbusch, L. Jeppe, T. Parikh, M. Schnepp, and A. R. Maier, Physical Review Letters **126**, 174801 (2021).

[32] G. M. Samarin, M. Zepf, and G. Sarri, Journal of Modern Optics **65**, 1362 (2018).

[33] C. D. Baird, C. D. Murphy, T. G. Blackburn, A. Ilderton, S. P. Mangles, M. Marklund, and C. P. Ridgers, New Journal of Physics **21**, 053030 (2019).

[34] C. Arran, J. M. Cole, E. Gerstmayr, T. G. Blackburn, S. P. Mangles, and C. P. Ridgers, Plasma Physics and Controlled Fusion **61**, 074009 (2019).

[35] J. Lin, Q. Qian, J. Murphy, A. Hsu, A. Hero, Y. Ma, A. G. R. Thomas, and K. Krushelnick, Physics of Plasmas **28**, 083102 (2021).

[36] W. Rawat and Z. Wang, Neural Computation **29**, 2352 (2017).

[37] A. Kristiadi, M. Hein, and P. Hennig, in *Proceedings of the 37th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 119, edited by H. D. III and A. Singh (PMLR, 2020) pp. 5436–5446.

[38] C. Emma, A. Edelen, M. J. Hogan, B. O'Shea, G. White, and V. Yakimenko, Physical Review Accelerators and Beams **21**, 112802 (2018).

[39] X. Ren, A. Edelen, A. Lutman, G. Marcus, T. Maxwell, and D. Ratner, Physical Review Accelerators and Beams **23**, 040701 (2020).

[40] A. Hanuka, C. Emma, T. Maxwell, A. S. Fisher, B. Jacobson, M. J. Hogan, and Z. Huang, Scientific Reports **11**, 1 (2021).

[41] T. P. Rowlands-Rees, C. Kamperidis, S. Kneip, A. J. Gonsalves, S. P. D. Mangles, J. G. Gallacher, E. Brunetti, T. Ibbotson, C. D. Murphy, P. S. Foster, M. J. V. Streeter, F. Budde, P. A. Norreys, D. A. Jaroszynski, K. Krushelnick, Z. Najmudin, and S. M. Hooker, Physical Review Letters **100**, 105005 (2008).

[42] A. Pak, K. A. Marsh, S. F. Martins, W. Lu, W. B. Mori, and C. Joshi, Physical Review Letters **104**, 025003 (2010).

[43] C. McGuffey, A. G. R. Thomas, W. Schumaker, T. Matsuoka, V. Chvykov, F. J. Dollar, G. Kalintchenko, V. Yanovsky, A. Maksimchuk, K. Krushelnick, V. Y. Bychenkov, I. V. Glazyrin, and a. V. Karpeev, Physical Review Letters **104**, 025004 (2010).

[44] M. Chen, E. Esarey, C. B. Schroeder, C. G. R. Geddes, and W. P. Leemans, Physics of Plasmas **19**, 033101 (2012).

[45] D. P. Kingma and M. Welling, 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings (2014), 10.48550/arxiv.1312.6114.

[46] C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner, arXiv (2018), 10.48550/arxiv.1804.03599.

[47] B. Xu, N. Wang, T. Chen, and M. Li, arXiv (2015), 10.48550/arxiv.1505.00853.

[48] D. P. Kingma and J. Ba, 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings (2014), 10.48550/arxiv.1412.6980.

[49] F. Barnsley, B. Matthews, T. Griffin, J. W. A. Salt, D. Ross, Cătălin, Condurache, A. Dibbo, and S. Y. Crompton (2016).

[50] W. Lu, M. Tzoufras, C. Joshi, F. Tsung, W. Mori, J. Vieira, R. Fonseca, and L. Silva, Physical Review Special Topics - Accelerators and Beams **10**, 061301 (2007).

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press

[51] M. S. Bloom, M. J. V. Streeter, S. Kneip, R. A. Bendoyro, O. Cheklov, J. M. Cole, A. Döpp, C. J. Hooker, J. Holloway, J. Jiang, N. C. Lopes, H. Nakamura, P. A. Norreys, P. P. Rajeev, D. R. Symes, J. Schreiber, J. C. Wood, M. Wing, Z. Najmudin, and S. P. D. Mangles, Physical Review Accelerators and Beams **23**, 061301 (2020).

[52] A. G. Thomas, S. P. D. Mangles, Z. Najmudin, M. C. Kaluza, C. D. Murphy, and K. Krushelnick, Physical Review Letters **98**, 54802 (2007).

[53] T. Matsuoka, C. McGuffey, P. G. Cummings, Y. Horovitz, F. Dollar, V. Chvykov, G. Kalintchenko, P. Rousseau, V. Yanovsky, S. S. Bulanov, A. G. R. Thomas, A. Maksimchuk, and K. Krushelnick, Physical Review Letters **105**, 034801 (2010).

https://doi.org/10.1017/hpl.2022.47 Published online by Cambridge University Press