Pareto Optimization and Tuning of a Laser Wakefield Accelerator

F. Irshad,[1] C. Eberle,[1] F. M. Foerster[●],[1] K. v. Grafenstein[●],[1] F. Haberstroh,[1] E. Travac,[1] N. Weisse,[1]
S. Karsch[●],[1,2] and A. Döpp[●][1,2,*]

[1]*Fakultät für Physik, Ludwig-Maximilian-Universität München, Am Coulombwall 1, 85748 Garching, Germany*
[2]*Max Planck Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, Garching 85748, Germany*

Optimization of accelerator performance parameters is limited by numerous trade-offs, and finding the appropriate balance between optimization goals for an unknown system is challenging to achieve. Here, we show that multiobjective Bayesian optimization can map the solution space of a laser wakefield accelerator (LWFA) in a very sample-efficient way. We observe that there exists a wide range of Pareto-optimal solutions that trade beam energy versus charge at similar laser-to-beam efficiency. Moreover, many applications such as light sources require particle beams at certain target energies. We demonstrate accurate energy tuning of the LWFA from 150 to 400 MeV via the simultaneous adjustment of eight parameters. To further advance this use case, we propose an inverse model that allows a user to specify desired beam parameters. Trained on the forward Gaussian process model, the inverse model generates input parameter value ranges within which the desired setting is likely to be reached. The method reveals different strategies for accelerator tuning and is expected to drastically facilitate the operation of LWFAs in the near future.

Laser wakefield acceleration (LWFA) [1,2] and related radiation sources [3] are an emerging technology with potentially broad applications in science, industry, and medicine [4]. Over the past decade, there has been significant progress with regard to both quality and stability of the accelerated electron beams [5] and recently, first experiments have demonstrated that beam parameters are sufficient to drive free electron lasers [6,7]. Much of this progress can be attributed to improved performance of laser systems and better targets, in particular the use of controlled injection methods such as shock-front injection or ionization-induced injection [8–11]. Nonetheless, the role of expert human operators in reaching optimal performance cannot be overstated. Their manual optimization typically relies on sequential line scans, using a combination of intuition and experience to determine optimal parameters for, e.g., target position and laser pulse duration. However, the continuously increasing complexity of laser-plasma experiments has made reaching optimal performance in a reproducible manner increasingly difficult.

Machine learning techniques offer powerful tools to address this challenge. A particularly popular method for

the optimization of laser-plasma accelerators is Bayesian optimization (BO) [12–14]. BO is a global optimization method based on searching optima in a probabilistic surrogate model that is updated iteratively as the experiment progresses [15]. It is extremely sample efficient, meaning it converges to the optimum with relatively few measurements. This makes it very suitable for the optimization of laser-plasma accelerators, where measurements are often acquired at a relatively low acquisition rate. The underlying probabilistic model commonly referred to as a surrogate model is usually chosen as a Gaussian process (GP) [16,17]. This is a nonparametric probabilistic model that assumes prior knowledge about possible relations between parameters and objectives. Within each iteration of the BO, a GP model is fitted to the current observations. This model is used to estimate a good position for the next measurement and the resulting observation is then appended to the data. The objective of optimizing a laser-plasma accelerator typically consists of maximizing one or more metrics that are combined into a single scalar value using empirically determined weights [13,14]. In a previous publication [18] we studied different objective functions such as combinations of mean electron energy, energy bandwidth and charge. There we found that *a priori* definition of objective weights does often not lead to the desired outcome. Instead, we showed that Pareto optimization via expected improvement of the hypervolume occupied by all objectives in the output space can efficiently explore the trade-offs between objectives and be used to choose adequate Pareto-optimal solutions *a posteriori*. The latter are all the best possible combinations of the

*Contact author: a.doepp@physik.uni-muenchen.de

(a) Experiment

Diagnostics

3-D Output Space $\vec{y} = \begin{pmatrix} Q \\ \bar{E} \\ \sigma_E \end{pmatrix}$

(b) Probabilistic Modeling

Measurements

8-D Input Space

$\vec{x} = \begin{pmatrix} x_{jet} \\ z_{jet} \\ x_{blade} \\ z_{blade} \\ p \\ \phi^{(2)} \\ \phi^{(3)} \\ \phi^{(4)} \end{pmatrix}$
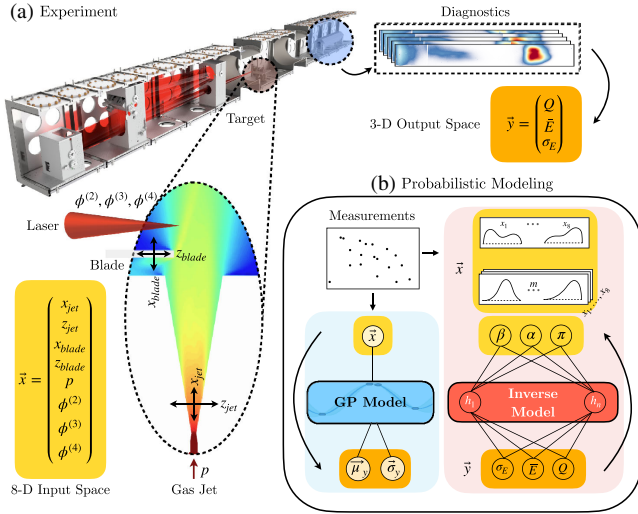
FIG. 1. (a) The Experimental setup for a laser wakefield accelerator with shock-front injection. The controllable input laser parameters for the optimization are the second, third, and fourth orders of dispersion ϕ of the laser. Regarding the plasma parameters, the jet and blade were moved in (x, z) position and the backing pressure p of the nozzle was also changed. This makes a total of eight controllable input parameters that were given to the Bayesian optimization algorithm. (b) During multiobjective optimization a Gaussian process (GP) model is trained based on the measurements to predict outputs y given an input x. The inverse model instead takes a desired output y and predicts a distribution of inputs x that are likely to produce these results.

objectives that are achievable, which form the so-called Pareto front [19]. This technique has shown promise in optimizing the parameters of an accelerator but has been restricted to using only surrogate models based on simulations [20].

In this Letter, we present the first experimental results on such a Pareto optimization of a laser wakefield accelerator. We also present a novel technique to tune the input parameters to result in electron beams desired by the user. Experiments were performed with the ATLAS laser system at the Centre for Advanced Laser Applications in Garching, Germany and the experimental setup is shown in Fig. 1. During the experiment the laser delivered 30 fs pulses at a central wavelength of 800 nm with an estimated energy of (4 ± 0.8) J energy within the 46 µm waist. The resulting peak intensity is (7.5 ± 0.4) 10^{18} W/cm^2, corresponding to a normalized vector potential of $a_0 = 1.8$.

The target consisted of a 7-mm-long supersonic gas jet using either pure hydrogen or a mixture containing a nitrogen dopant. The jet is mounted on a hexapod for three-dimensional positioning. Moving the gas jet along the laser focus (z axis) changes the peak intensity at the target entrance, while movement in the vertical (x) direction simultaneously changes the density and cross-section of the gas jet that the laser interacts with. A silicon wafer is mounted on an (x, z) motorized stage to locally perturb the

supersonic gas flow, leading to the formation of a shock front. At this shock front, the gas density rapidly drops, resulting in an expansion of the plasma wakefield that facilitates the injection of electrons [9]. The wafer can be moved along the gas jet to control the point of injection. This can be used to tune the energy of the electron beams since the acceleration length can be reduced or increased by moving the blade in or out of the gas jet, respectively. The blade is also motorized vertically over the gas jet to result in different shapes of the shock [21]. The gas density can be changed by changing the backing pressure of the nozzle and three different orders of dispersion of the laser pulse are controlled through an acousto-optic programmable dispersive filter. Tango controls is used to change the longitudinal gas jet position, gas jet height, blade position, blade height, gas pressure, and the 3 orders of dispersion of the laser in an automated fashion [22,23]. A dipole magnet spectrometer is used to disperse the electron beam onto several charge-calibrated scintillating screens [24], which are imaged and stitched together to yield the complete electron energy spectrum. More details on the experimental setup can be found in Supplemental Material [25].

From the energy spectrum we determine the total beam charge Q, median energy $\bar{E}$, and energy spread σ_E. These parameters uniquely define a normally distributed energy spectrum and in their three-dimensional objective space $\mathcal{Y}$, their trade-offs span the Pareto "surface" $\mathcal{P}$. However, the spectral distribution of a laser-accelerated electron beam may considerably differ from a normal distribution and because of this, metrics such as the mean energy may not be characteristic of peaks in the spectral distribution. As an example, when a gas mixture is used, the resulting shock-assisted ionization injection [30] produces multiple electron bunches with low-energy tails in the spectrum as well as a high-energy peak. This skews the metrics defined over the whole spectrum and thus requires a method to separate the individual bunches within a single spectrum. To this end, we used a Gaussian mixture model (GMM) [12] which is discussed in detail in Supplemental Material [25,31]. This postprocessing step allows us to isolate specific bunch contributions in real time, avoiding contamination and enabling precise statistical analysis of each target electron beam bunch (an example for GMM segmented spectra is shown in Fig. 2).

In our initial optimization effort, we focused on enhancing the bunch's beam charge and average energy, while aiming to reduce the energy spread. This task was undertaken through multiobjective optimization (MOO), utilizing a subset of four input parameters (z_{jet}, x_{jet}, z_{blade}, x_{blade}) from a total of eight available. Detailed discussions of these preliminary results are provided in Supplemental Material [25], showcasing the effectiveness of our technique. Following the successful demonstration of the algorithm's convergence and effectiveness, we expanded our approach to include the full set of eight dimensions. In this
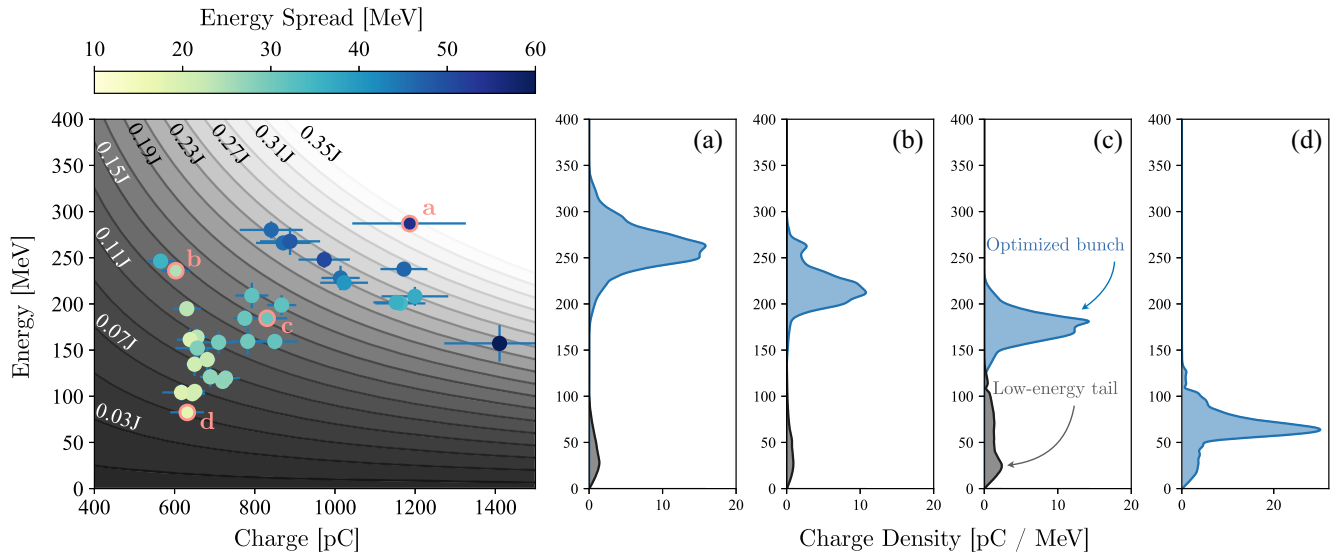
FIG. 2. Identified Pareto-optimal configurations of the laser wakefield accelerator for an eight-dimensional Pareto optimization of beam energy, charge, and energy spread. The energy spread is color-coded while the errorbars depict the standard deviation of charge and energy. Regions of equal beam energy, and hence, accelerator efficiency, are shaded in gray. Note that the 8D optimization resulted in higher efficiencies of the laser wakefield accelerator since it was able to also find better positions of the four new parameters that were introduced in this optimization task. Panels (a)–(d) show representative spectra (similar beam parameters to average) for the respectively marked points *a–d* from the Pareto front. The grayed out part of the spectrum is the low energy tail isolated through via the Gaussian mixture model discussed previously.

comprehensive 8D optimization, we maintained the original objectives but incorporated additional variables, namely the gas pressure and the three orders of laser dispersion, to further refine our results. Moreover, in this optimization we allow the optimizer control over the number of shots rather than taking a fixed number of shots at each position. This is enabled by using a multiobjective multifidelity optimization technique called Trust-MOMF [32] (for details, see Supplemental Material [25]). This allows the algorithm to dynamically change the number of shots it needs to take in different regions of the input parameter space. The minimum and the maximum number of shots is determined by the user and in this case was kept to 3 and 15, respectively. The total number of positions taken within this optimization run was 95 including 10 random initial positions. The resulting Pareto front contained 35 points. The results are shown in Fig. 2 and when compared with the results of the 4D optimization shown in Sec. III B of Supplemental Material, we see that the introduction of the four additional parameters results in increased efficiency of the LWFA: the efficiency increased from about 4% to 6.5% on average, with up to 8.25% for one configuration. In this instance, we can also see the same behavior where we have to reduce the beam energy to yield low bandwidth beams. Additionally, in our case the efficiency needs to be decreased and less number of electrons injected to result in beams with lower bandwidth. The four panels on the right side in Fig. 2 are representative spectra of four selected points. As one can see, a single multiobjective optimization can yield different

electron beams optimized at different energies. A user could then *a posteriori* select the solution from among these that is required for a particular application. Also shown in the left panel of Fig. 2 as error bars are the standard deviations of the charge and the energy due to shot-to-shot fluctuations that we aimed to reduce with the next optimization task. Multiobjective optimization can also be applied to different objectives, such as a desired target energy or stability. These additional results can be found in Supplemental Material [25].

Once the Pareto front is established to a sufficient degree, one can proceed to use the model for different goals, leveraging its predictive capabilities to cater to changing user preferences dynamically without running new optimizations from the start. This approach is advantageous compared to starting with single-objective optimization, as the goals are now informed by a model that is less biased toward a particular outcome. By separating the multiobjective exploration and single-objective exploitation phases, we present an efficient methodology for addressing both phases. The GP model built during the Pareto optimization can be reused to serve different user preferences more dynamically, such as energy tuning or targeting specific beam parameters. This is achieved by exploiting specific solutions using an *a posteriori* scalarization of the objectives. Here, we use the upper confidence bound acquisition (UCB) function, which considers both the probability of yielding beams at a particular target energy and the uncertainty of our prediction. By minimizing the
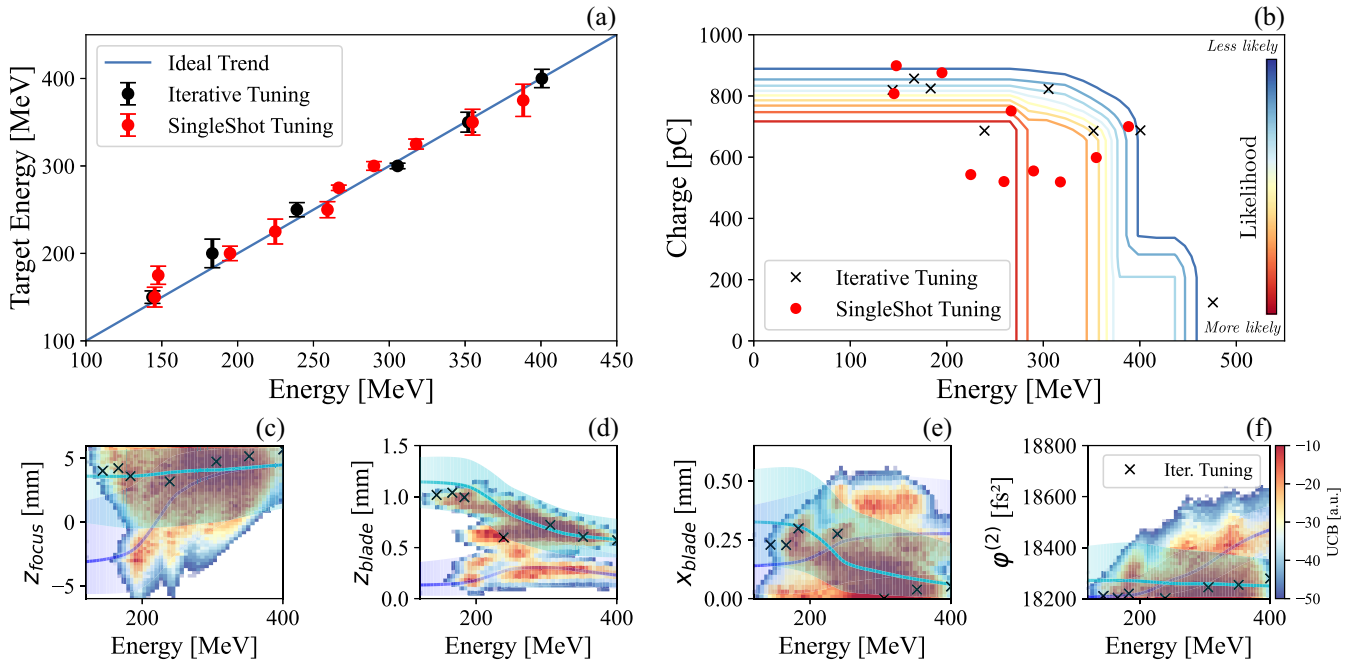
FIG. 3. (a) Energy tuning of the LWFA. The ideal tuning line is shown in blue while the experimentally attained iterative (single shot) tuning points are shown in black (red). The error bars show the standard error of the mean. It can be seen that the optimizer is able to tune the mean energy of the electron beam in the region between 150 to 400 MeV. (b) Probabilistic Pareto front. This shows the types of beams that are accessible to the user with a certain likelihood. Colors represent the likelihood of attaining a particular value. Also shown within the plot are both sets of tuning points. Note that the straight lines are a result of the domination concept, where a solution is considered optimal if no other solution can improve one objective without sacrificing another at a given likelihood. (c)–(f) Maps of the UCB acquisition function for four of eight input parameters and different target energies. Overlaid are the two most likely tuning curves derived from the mixture density network represented by the median (solid line) and 95% confidence interval (shaded region) of the associated mixture distribution. Transparency encodes their respective likelihood.

UCB function using a gradient optimizer [33], we obtain a new eight-dimensional configuration of the laser Wakefield accelerator that is most likely to generate electron beams at the user-defined target energy.

To demonstrate the effectiveness of this approach, we conducted iterative tuning by taking ten shots at each suggested position, appending the new data to the GP's training data, and repeating this process 3 times for each target energy. The results, referred to as iterative tuning points. Additionally, we performed single-shot tuning without appending any more data. The results are shown in Fig. 3, where the iterative (single-shot) scan ranged from 150 to 400 MeV in 50 MeV (25 MeV) increments. The results clearly demonstrate that our method allows for accurate energy tuning over an octave in energy. The successful single-shot tuning demonstrates the model's efficiency in rapid accelerator control, while iterative tuning showcases its adaptability to long-term performance drifts. Indeed, the concurrent movement across an eight-dimensional parameter space exceeds the capability of even the most experienced human operators. After tuning in the appropriate range, the limits of the model were tested by asking it to achieve a target energy of 600 MeV. There were some shots that resulted in beams close to 600 MeV but on

average the beam energy was 470 MeV as shown in Fig. 3(b). Furthermore, it is remarkable that this was achieved using a model that was not particularly optimized for this task; the GP was based on a multiobjective optimization with a target energy of 250 MeV.

A more general approach to tuning the energy or the other objectives can be obtained by inverting the input and output domains. In this technique, the objectives (charge, energy, bandwidth) become the input parameters and the eight different parameters that are controlled become the outputs. A model is then trained from the output objective space $\mathcal{Y}$ to the input parameter space $\mathcal{X}$ which we call an inverse model. This perspective makes it easier for the user to formulate the problem of tuning. The user requests particular beam parameters and the model then outputs the required adjustments on the controlled parameters to achieve that particular beam.

Inverting the forward model $\mathcal{P}[\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\varphi}]$, which maps from the input space to the output space with parameters $\boldsymbol{\varphi}$, is challenging due to the difference in dimensionality. The forward mapping is typically not one-to-one, meaning that multiple input configurations can lead to the same output. Consequently, a unique inverse function does not exist. To address this issue, we propose a probabilistic approach that

learns a set-valued inverse mapping $f^{-1}: \mathcal{Y} \to \mathcal{X}^m$. In this formulation, each desired beam property y is associated with a set of possible accelerator configurations $x_1, \ldots, x_m$, rather than a single input. This allows us to capture the inherent ambiguity in the inverse problem and to identify multiple strategies for achieving a given beam property.

To model such a function while staying within a probabilistic framework we propose the use of a mixture density network (MDN) [34], where the probability distribution over inputs $\{x_1, \ldots, x_m\}$ for a given objective y is given by a multimodal mixture distribution. Specifically, as the input space is bounded, the eight dimensional unit hypercube $[0, 1]^8$ after normalization of coordinates, we chose a mixture of beta stochastic processes

$$\mathcal{P}[x|y; \pi, \alpha, \beta] = \sum_{i=1}^{m} \pi_i(y) \text{beta}[x|\alpha_i(y), \beta_i(y)], \quad (1)$$

where $\alpha_i(y)$, $\beta_i(y)$ are the shape functions, and $\pi(y)$ is the mixture weight function, defining the local mixture at y and learned by the MDN. The model is trained by minimizing the negative log likelihood $\mathcal{H}[\mathcal{D}; \pi, \alpha, \beta] = -\sum_{x,y \in D} \log \mathcal{P}[x|y; \pi, \alpha, \beta]$ of the data set $\mathcal{D}$. Additionally, a regularization term is added to ensure consistency in expectation with the GP model (see Supplemental Material, Sec. IV).

Using such an inverse model, a user can then tune specific output objectives for specific applications. In particular, the model can be queried with a tuning range to obtain a probability distribution over all possible tuning curves. In Fig. 3(c) through 3(f), we show the two most probable tuning curves out of five for changing the median energy from 150 to 450 MeV. These probabilistic tuning curves show how the inputs must be changed in order to obtain a certain output energy. It needs to be stressed that at each E the distribution over input parameters Eq. (1) is such that sampling from this distribution and running the sampled values through the forward model, on average, yields the correct energy. Each of the tuning curve represents a certain strategy to obtain a desired electron beam. One strategy that the inverse model correctly identifies is moving the blade further into the gas jet to reduce the electron beam median energy while compensating the focus position by a similar amount. Furthermore, parameters such as blade height and dispersion are adjusted subtly. It also suggests an alternative strategy where the jet is moved along the focus position and second order dispersion is increased, instead of relying on the blade. Our results show that GP modeling and inverse mixture models can not only perform well, but they can also inform human operators about adequate strategies to change accelerator parameters.

In conclusion, we have demonstrated the successful application of multiobjective Bayesian optimization to map the solution space of a laser wakefield accelerator in a sample-efficient manner. The method reveals a wide range of Pareto-optimal solutions that trade beam energy and charge at similar laser-to-beam efficiency. Moreover, we have shown accurate energy tuning of the LWFA from 150 to 400 MeV by simultaneously adjusting eight parameters. The proposed inverse model, trained on the forward Gaussian process model, generates input parameter ranges likely to yield the desired beam settings. This approach uncovers different strategies for accelerator tuning and is expected to significantly facilitate the operation of LWFAs. These results represent an important step toward the systematic optimization and control of laser-plasma accelerators, paving the way for their future applications in science, industry, and medicine.

[1] E. Esarey, C. B. Schroeder, and W. P. Leemans, Physics of laser-driven plasma-based electron accelerators, Rev. Mod. Phys. **81**, 1229 (2009).

[2] V. Malka, Laser plasma accelerators, Phys. Plasmas **19**, 055501 (2012).

[3] S. Corde, K. T. Phuoc, G. Lambert, R. Fitour, V. Malka, A. Rousse, A. Beck, and E. Lefebvre, Femtosecond x rays from laser-plasma accelerators, Rev. Mod. Phys. **85**, 1 (2013).

[4] F. Albert and A. G. Thomas, Applications of laser wakefield accelerator-based light sources, Plasma Phys. Control. Fusion **58**, 103001 (2016).

[5] A. R. Maier, N. M. Delbos, T. Eichner, L. Hübner, S. Jalas, L. Jeppe, S. W. Jolly, M. Kirchen, V. Leroux, P. Messner *et al.*, Decoding sources of energy variability in a laser-plasma accelerator, Phys. Rev. X **10**, 031039 (2020).

[6] M. Labat, J. C. Cabadağ, A. Ghaith, A. Irman, A. Berlioux, P. Berteaud, F. Blache, S. Bock, F. Bouvet, F. Briquez *et al.*, Seeded free-electron laser driven by a compact laser plasma accelerator, Nat. Photonics **17**, 1 (2022).

[7] W. Wang, K. Feng, L. Ke, C. Yu, Y. Xu, R. Qi, Y. Chen, Z. Qin, Z. Zhang, M. Fang *et al.*, Free-electron lasing at 27 nanometres based on a laser wakefield accelerator, Nature (London) **595**, 516 (2021).

[8] 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, Self-guided laser wakefield

acceleration beyond 1 GeV using ionization-induced injection, Phys. Rev. Lett. **105,** 105003 (2010).

[9] K. Schmid, A. Buck, C. M. 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, Phys. Rev. ST Accel. Beams **13,** 091301 (2010).

[10] J. Wenz, A. Döpp, K. Khrennikov, S. Schindler, M. Gilljohann, H. Ding, J. Götzfried, A. Buck, J. Xu, M. Heigoldt *et al.*, Dual-energy electron beams from a compact laser-driven accelerator, Nat. Photonics **13,** 263 (2019).

[11] J. Götzfried, A. Döpp, M. F. Gilljohann, F. M. Foerster, H. Ding, S. Schindler, G. Schilling, A. Buck, L. Veisz, and S. Karsch, Physics of high-charge electron beams in laser-plasma wakefields, Phys. Rev. X **10,** 041015 (2020).

[12] A. Döpp, C. Eberle, S. Howard, F. Irshad, J. Lin, and M. Streeter, Data-driven science and machine learning methods in laser-plasma physics, High Power Laser Sci. Eng. **11,** 55 (2023).

[13] R. J. Shalloo *et al.*, Automation and control of laser wakefield accelerators using Bayesian optimization, Nat. Commun. **11,** 6355 (2020).

[14] S. Jalas, M. Kirchen, P. Messner, P. Winkler, L. Hübner, J. Dirkwinkel, M. Schnepp, R. Lehe, and A. R. Maier, Bayesian optimization of a laser-plasma accelerator, Phys. Rev. Lett. **126,** 104801 (2021).

[15] C. E. Rasmussen, C. K. Williams *et al.*, *Gaussian Processes for Machine Learning* (Springer, New York, 2006), Vol. 1.

[16] C. Williams and C. Rasmussen, Gaussian processes for regression, Adv. Neural Inf. Process. Syst. **8,** 514 (1995).

[17] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, Botorch: A framework for efficient Monte-Carlo Bayesian optimization, Adv. Neural Inf. Process. Syst. **33,** 21524 (2020).

[18] F. Irshad, S. Karsch, and A. Döpp, Multi-objective and multi-fidelity Bayesian optimization of laser-plasma acceleration, Phys. Rev. Res. **5,** 013063 (2023).

[19] Note that the GP model provides a probabilistic surrogate of the objective function landscape and thus, instead of point estimates, it inherently yields a probabilistic delineation of the potential location of the Pareto front that will improve during optimization.

[20] R. Roussel, A. Hanuka, and A. Edelen, Multiobjective Bayesian optimization for online accelerator tuning, Phys. Rev. Accel. Beams **24,** 062801 (2021).

[21] E. Guillaume, Control of electron injection and acceleration in laser-wakefield accelerators, Ph.D. thesis, École Polytechnique, 2015.

[22] N. Weiße, L. Doyle, J. Gebhard, F. Balling, F. Schweiger, F. Haberstroh, L. D. Geulig, J. Lin, F. Irshad, J. Esslinger *et al.*, Tango controls and data pipeline for petawatt laser experiments, High Power Laser Sci. Eng. **11,** 1 (2022).

[23] TANGO Controls Collaboration, Tango controls, https://www.tango-controls.org (2024).

[24] T. Kurz, J. P. Couperus, J. M. Krämer, H. Ding, S. Kuschel, A. Köhler, O. Zarini, D. Hollatz, D. Schinkel, R. D'Arcy, J. P. Schwinkendorf, J. Osterhoff, A. Irman, U. Schramm, and S. Karsch, Calibration and cross-laboratory implementation of scintillating screens for electron bunch charge determination, Rev. Sci. Instrum. **89,** 093303 (2018).

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.133.085001, which includes Refs. [26–29], for additional information about the experimental methods, optimization techniques and additional results.

[26] M. Kirchen, S. Jalas, P. Messner, P. Winkler, T. Eichner, L. Hübner, T. Hülsenbusch, L. Jeppe, T. Parikh, M. Schnepp, and A. R. Maier, Optimal beam loading in a laser-plasma accelerator, Phys. Rev. Lett. **126,** 174801 (2021).

[27] F. Foerster, A. Döpp, F. Haberstroh, K. v. Grafenstein, D. Campbell, Y.-Y. Chang, S. Corde, J. C. Cabadağ, A. Debus, M. Gilljohann *et al.*, Stable and high-quality electron beams from staged laser and plasma wakefield accelerators, Phys. Rev. X **12,** 041016 (2022).

[28] E. Jang, S. Gu, and B. Poole, Categorical reparameterization with gumbel-softmax, arXiv:1611.01144.

[29] D. P. Kingma and M. Welling, Auto-encoding variational bayes, arXiv:1312.6114v1.

[30] C. Thaury, E. Guillaume, A. Lifschitz, K. Ta Phuoc, M. Hansson, G. Grittani, J. Gautier, J.-P. Goddet, A. Tafzi, O. Lundh *et al.*, Shock assisted ionization injection in laser-plasma accelerators, Sci. Rep. **5,** 16310 (2015).

[31] A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc. **39,** 1 (1977).

[32] F. Irshad, S. Karsch, and A. Döpp, Leveraging trust for joint multi-objective and multi-fidelity optimization, Mach. Learn. **5,** 015056 (2024).

[33] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput. **16,** 1190 (1995).

[34] C. M. Bishop, Mixture density networks, Technical Report No. NCRG/94/004, Aston University, 1994.