
Differentiable Models for Control of Complex Physical Systems: A Case Study in Laser Pulse Shaping

Denis Ilia

Deutsches Elektronen-Synchrotron DESY
University of Hamburg
denis.ilia@desy.de

Nihat Ay

Institute for Data Science Foundations,
Hamburg University of Technology
nihat.ay@tuhh.de

Ingmar Hartl

Deutsches Elektronen-Synchrotron DESY
ingmar.hartl@desy.de

Wolfgang Hillert

Institute for Experimental Physics,
University of Hamburg
wolfgang.hillert@uni-hamburg.de

Henrik Tünnemann

Deutsches Elektronen-Synchrotron DESY
henrik.tuennermann@desy.de

Abstract

Precise control of the temporal profile of laser pulses is critical for many scientific and industrial applications. Achieving this via spectral shaping is particularly challenging in complex systems due to nonlinear effects, input fluctuations, and hardware imperfections. Conventional approaches—such as manual iterative tuning, precomputed spectral settings, or evolutionary optimization—are often insufficient or scale poorly with system complexity. We introduce a differentiable physics-based framework for spectral pulse shaping that embeds a physical model of the laser within a gradient-based optimization loop. This approach enables rapid system identification and control, accurately capturing complex laser dynamics while optimizing control inputs to achieve target temporal pulse shapes. We demonstrate the method by shaping near-infrared pulses as a proof of concept for photoinjector laser systems, highlighting its generality and potential for high-dimensional control of complex physical systems.

1 Introduction

The ability to control the temporal intensity profile of laser pulses is crucial for both research and industrial applications [3, 9, 10, 13]. A wide range of techniques have been developed, typically classified as temporal- or spectral-domain methods. Temporal-domain methods include direct modulation of pulse intensity using fast devices such as electro-optic modulators, as well as pulse stacking schemes that combine delayed replicas [11]. While effective, their resolution is limited either by device response times (typically 100 ps–1 ns [7]) or by the fixed nature of the stacking geometry.

Here, we focus on shaping pulses in the 100 fs–100 ps range using spectral-domain techniques. These methods exploit the Fourier relationship between time and frequency: pulse spectrum is dispersed onto a spatial plane, where amplitude and/or phase can be modulated with programmable devices such as acousto-optic modulators (AOMs) or spatial light modulators (SLMs) [12].

In principle, control settings could be calculated directly from the Fourier transform of the desired temporal profile, assuming the input pulse is well-characterized and propagation is linear. In practice, however, input fluctuations, nonlinear propagation effects, and device imperfections create complex, system-dependent relationships between the control parameters and the resulting pulse shape.

From a control perspective, each spectral element—whether a “pixel” in an SLM or a “channel” in an AOM—represents a tunable parameter. Spectral shaping thus involves simultaneously adjusting thousands of parameters, framing it as a high-dimensional optimization problem. Conventional strategies—either static precomputed settings or manual iterative tuning—are typically insufficient, as they cannot cope efficiently with system drift, nonlinearities, or hardware imperfections.

Traditional control approaches struggle at this scale: evolutionary algorithms are sample-inefficient, classical control cannot handle thousands of coupled parameters, and analytical solutions break down under nonlinear propagation effects.

Our approach addresses these challenges by representing the physical system in a differentiable framework, enabling gradient-based system identification and optimization. In a two-step loop, experimental data first calibrate the system-specific internal parameters to ensure the model matches the real system, then all control parameters are updated via gradients to approach the target pulse. Alternating these steps allows rapid, sample-efficient adaptation to drift, nonlinearities, and device imperfections.

We demonstrate this method by shaping near-infrared (NIR) pulses as a first step toward UV pulse shaping in a photoinjector laser [6]. This system delivers UV pulses to free-electron laser (FEL) facilities, where precisely shaped pulses directly impact electron bunch quality and overall facility performance [4, 5]. Although the demonstration is currently in NIR, the approach is fully transferable to UV pulses and will be part of future work.

2 Differentiable Physics-based Model

The physics of laser systems can be described by well-established mathematical models, and recent advances in differentiable programming[2] allow these models to be implemented directly within modern machine learning frameworks.

We developed a differentiable simulation framework for ultrafast pulse propagation in fibers and laser systems, implemented in PyTorch with GPU acceleration and autograd support. The framework works with Pulse objects, which provide access to both time- and frequency-domain representations of an optical pulse. The design follows PyTorch’s nn.Module abstraction: each module represents a physical component or process—such as a fiber, amplifier, spectral filter, or dispersive element—and is encapsulated as a reusable, differentiable block. Physical parameters are exposed as nn.Parameter objects, enabling gradient-based optimization and system identification. Modules act on Pulse objects and return updated Pulse instances, ensuring composable simulations.

The simulation framework operates on discretized pulse representations with configurable grid points (typically 2^{10} to 2^{14} points) and time windows that together determine spectral resolution and memory requirements. This discretization forms the foundation for all physical operations within the framework. Different physical processes are implemented according to their underlying mathematics. Fiber propagation uses the split-step Fourier method (SSFM) applied to the nonlinear Schrödinger equation [1], with the number of integration steps per element adjustable to balance accuracy and computational cost. Active media such as amplifiers are described by rate equations in steady-state approximation [8]. Additional modules implement analytic transformations such as dispersion, group delay, and spectral filtering, which modify the pulse according to known physics without requiring full numerical simulation. Spectral shaping operates by element-wise multiplication of the pulse spectrum with a complex-valued tensor of matching dimensionality. This simple operation creates the high-dimensional control problem discussed earlier, as each spectral grid point becomes an independent control parameter capable of modifying both amplitude and phase.

Together, these physics-based components provide an accurate representation of ultrafast laser systems while remaining fully compatible with gradient-based optimization. The framework supports both single- and batched-pulse simulations, enabling exploration and optimization of multiple configurations in parallel. Complex laser systems can be constructed by chaining modules into a sequential model, as illustrated in Figure 1. This modular architecture facilitates integration with

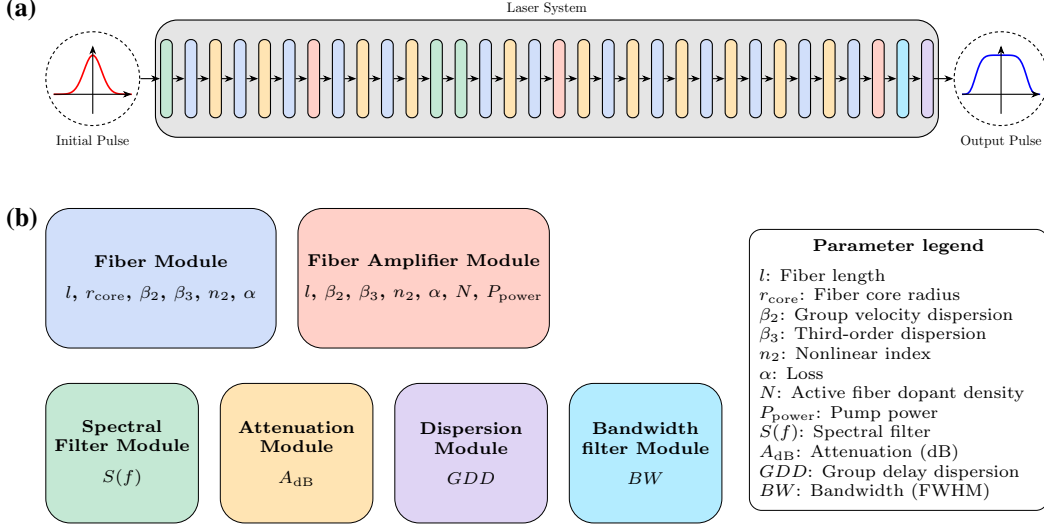


Figure 1: The laser system (a) is constructed by chaining together individual differentiable modules, each representing a physical component or process. The system processes an initial pulse sequentially through the optical chain to produce the output pulse. (b) shows the available module types with their learnable parameters (detailed in legend). Each module is implemented as a PyTorch nn.Module with parameters exposed as nn.Parameter objects, enabling gradient computation through the entire physical system. Color coding in (a) corresponds to module types in (b).

existing ML pipelines while lowering the barrier for researchers to adapt or extend the framework for new applications.

3 Control problem definition and optimization

The task of pulse shaping can be formalized as an optimization problem in which the control inputs $u \in \mathbb{C}^N$ are applied to the real experimental system, which we denote by $F(\cdot)$, to achieve the target pulse y^* . The experiment returns a measured pulse

$$y_{\text{exp}} = F(u) + \varepsilon,$$

where ε captures measurement noise and unmodeled effects. The ideal control objective is therefore

$$u^* = \arg \min_{u \in \mathcal{U}} \|F(u) - y^*\|_2^2,$$

with \mathcal{U} encoding hardware constraints such as finite dynamic range and bandwidth limitations. Because direct optimization of F is impractical—measurements are costly and sparse, and u is high-dimensional—we introduce a differentiable physical model

$$y_{\text{model}} = f(u, \theta)$$

that depends on control inputs u and internal state parameters $\theta \in \mathbb{R}^M$. The differentiable nature of f enables efficient gradient-based optimization of both control inputs and model parameters. However, even with a high-fidelity model, two key challenges remain. First, while rough estimates of the physical parameters are available (e.g., fiber lengths, dispersion coefficients, nonlinear refractive index, pump power), small deviations can show cumulative effects, producing significantly different outputs. Second, unmodeled effects prevent the model from faithfully capturing the behavior of the physical system across the entire control input space, limiting its accuracy to the local operating regime.

To address these challenges, we propose an adaptive approach that alternates between two phases:

System identification (model fitting): Using a sliding window buffer containing the most recent B experimental pairs $\{(u^{(i)}, y_{\text{exp}}^{(i)})\}$, we estimate parameters θ by

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^B \left\| f(u^{(i)}, \theta) - y_{\text{exp}}^{(i)} \right\|_2^2.$$

The buffer size B is chosen to balance model accuracy with memory usage and computation time. Forward and backward passes over the buffered measurements are parallelized, improving scalability until all operations can eventually be fit on a single GPU.

Control optimization: Using the updated model parameters $\hat{\theta}$, we optimize control inputs by

$$u^{(k+1)} = \arg \min_{u \in \mathcal{U}} \|f(u, \hat{\theta}) - y^*\|_2^2 + \lambda R(u),$$

where $R(u)$ enforces additional physics informed constraints and λ is a regularization weight.

The procedure begins by initializing $u^{(0)}$, chosen as the Fourier transform of the target pulse. At each iteration, a measurement $(u^{(k)}, y_{\text{exp}}^{(k)})$ is obtained from the physical system and stored in a sliding window buffer containing the most recent B samples. Using the data in this buffer we perform system identification to update the internal model parameters. Control optimization follows generating the updated control inputs $u^{(k+1)}$. Both routines exploit gradient-based optimization through the differentiable model f .

A detailed description of the algorithm is provided in Appendix A.

4 Experimental Validation

We demonstrate our adaptive pulse shaping approach on a photoinjector laser, where precise temporal control over the pulse intensity profile is crucial for improving electron beam quality. The control vector $u \in \mathbb{C}^{6219}$ is defined as the complex-valued filter applied by the spatial light modulator (SLM), located in the front-end of the laser system. While the final output of the system is a UV pulse, the shaping target is the intermediate IR pulse prior to frequency conversion. This choice serves to establish a proof of principle for the method. For optimization, u is interpolated onto the simulation grid of 2^{11} points spanning a 240 ps temporal window; after optimization, the resulting filter is interpolated back to the SLM control grid. Pulse measurements y_{exp} are obtained via cross-correlation with 0.2 ps resolution, with each acquisition requiring approximately 3-5 minutes due to the scanning process. For more details on the laser system and its modeling, see Appendix B.

Our target specifications focus on generating flat-top pulses in the 10-20 ps duration range, which are optimal for generation of high-quality electron beams in FEL[4]. To demonstrate the versatility of our approach, we evaluate performance on three distinct target shapes of similar temporal duration: a flat-top pulse (our primary application target), a triangular pulse, and an arbitrary composite shape.

Figure 2 presents the experimental results. The adaptive shaping procedure, which ran for 30 iterations with a system identification buffer size of $B=5$, achieved substantial improvement over the initial condition. The optimized pulse profiles closely match all three target shapes, with a relative L_2 error of less than 2%.

5 Discussion

5.1 Convergence considerations

The gradient-based optimization employed in both system identification and control optimization is susceptible to local minima due to the nonconvex nature of the problem. However, global convergence is not strictly necessary—reaching a solution that meets performance specifications is often sufficient in practice. This requires two conditions: (i) the system identification step must yield a model that accurately represents local system behavior near the current control input, and (ii) the initial condition $u^{(0)}$ must lie within the basin of attraction of a satisfactory local optimum.

The adaptive design of our approach addresses the first condition by continuously updating model parameters using recent measurements, thereby maintaining local validity as optimization progresses. The second condition is more challenging to guarantee. We initialize $u^{(0)}$ using the Fourier transform

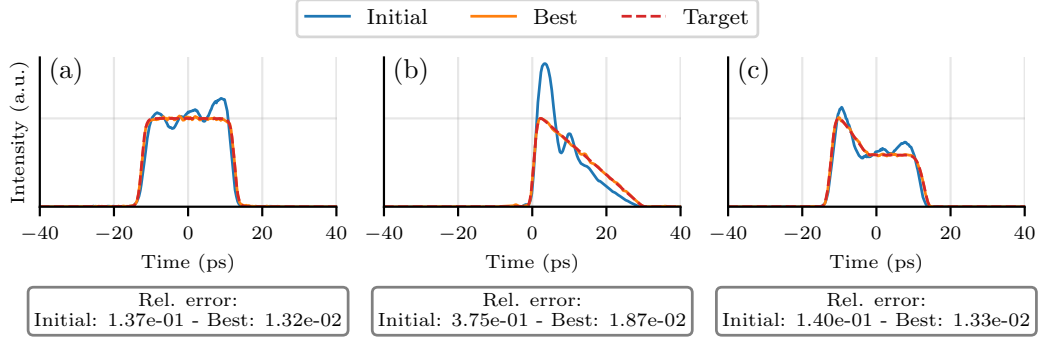


Figure 2: Experimental results of adaptive pulse shaping. Panels (a), (b), and (c) show the shaping of a flat-top, triangle, and arbitrary composite pulse, respectively. In blue we highlight the initial pulse profiles, in orange the profiles after optimization and in red the target profiles. Relative errors are reported as the relative L_2 error between the measured and target pulses for each case.

of the target pulse, providing a physics-informed starting point that aligns with the expected system response under near-linear propagation.

Convergence behavior depends critically on optimizer choice and hyperparameters, including learning rates, regularization weights, and buffer size B . These parameters govern the trade-off between model adaptation speed and control update stability; poor settings can lead to slow convergence, oscillations, or stagnation. Stable performance requires allowing both routines sufficient iterations to reach approximate convergence at each step. We initially tuned hyperparameters empirically using a digital twin of the physical system, with final adjustments during experimental runs. Once identified, appropriate hyperparameters enable robust performance across different target pulses, as demonstrated by the consistent results in Figure 2.

5.2 Transferability to UV shaping

While our experimental demonstration focuses on shaping NIR pulses, the approach is fully compatible with UV pulse shaping. The differentiable model can be extended to include frequency conversion stages by incorporating additional modules that simulate these nonlinear processes. The control optimization framework remains unchanged, as it operates on the same principles of gradient-based adaptation. The primary challenge in UV shaping lies in the strong nonlinearities introduced during frequency conversion, which can make the control optimization landscape more complex and sensitive to initial conditions. Fourier-based initialization becomes less effective in this regime, necessitating alternative strategies. Initial UV shaping experiments confirm that the approach converges successfully when provided with an appropriate $u^{(0)}$. In these experiments, we extended the control optimization at the first iteration, performing a longer optimization using the approximate model to compute the initial condition. Alternative strategies such as machine learning methods trained on synthetic data from the approximate model could potentially provide better initial conditions more efficiently by learning an inverse mapping from target pulses to control inputs.

6 Conclusions

We introduced a differentiable, physics-based framework for adaptive control of complex laser systems, achieving precise pulse shaping and robust adaptation to drift, nonlinearities, and hardware imperfections. The alternating system identification and control optimization enables efficient tuning of thousands of parameters. Modular and physics-agnostic, the framework generalizes to other high-dimensional physical systems and can integrate with neural network accelerators or reinforcement learning. This approach paves the way for robust, scalable control of sophisticated scientific and industrial systems.

Acknowledgments and Disclosure of Funding

This work was supported by the Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) under the OPAL-FEL project. We acknowledge DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities and required hardware. We acknowledge European XFEL in Schenefeld, Germany, for providing the laser facilities used in this work.

We would like to thank Yujiao Jiang, James Good, Chen Li for helping characterizing the laser system, which greatly assisted the modeling effort. We also thank Christoph Mahnke for fruitful discussions regarding the photoinjector laser system.

References

- [1] G. Agrawal. Chapter 2 - Pulse Propagation in Fibers. In G. Agrawal, editor, *Nonlinear Fiber Optics (Fifth Edition)*, Optics and Photonics, pages 27–56. Academic Press, Boston, fifth edition, 2013. doi: <https://doi.org/10.1016/B978-0-12-397023-7.00002-4>. URL <https://www.sciencedirect.com/science/article/pii/B9780123970237000024>. ISSN: 15575837.
- [2] M. Blondel and V. Roulet. The Elements of Differentiable Programming, July 2024. URL <http://arxiv.org/abs/2403.14606>. arXiv:2403.14606 [cs].
- [3] M. B. Hild, M. Safaeisadeh, and A. Lindinger. Spatial and temporal laser pulse shaping for two color excitation. *Optik*, 250:168293, Jan. 2022. ISSN 0030-4026. doi: 10.1016/j.ijleo.2021.168293. URL <https://www.sciencedirect.com/science/article/pii/S003040262101809X>.
- [4] M. Krasilnikov, F. Stephan, G. Asova, H.-J. Grabosch, M. Groß, L. Hakobyan, I. Isaev, Y. Ivanisenko, L. Jachmann, M. Khojyan, G. Klemz, W. Köhler, M. Mahgoub, D. Maityutin, M. Nozdrin, A. Oppelt, M. Otevrel, B. Petrosyan, S. Rimjaem, A. Shapovalov, G. Vashchenko, S. Weidinger, R. Wenndorff, K. Flöttmann, M. Hoffmann, S. Lederer, H. Schlarb, S. Schreiber, I. Templin, I. Will, V. Paramonov, and D. Richter. Experimentally minimized beam emittance from an SLS-band photoinjector. *Physical Review Special Topics - Accelerators and Beams*, 15(10):100701, Oct. 2012. doi: 10.1103/PhysRevSTAB.15.100701. URL <https://link.aps.org/doi/10.1103/PhysRevSTAB.15.100701>. Publisher: American Physical Society.
- [5] O. J. Luiten, S. B. Van Der Geer, M. J. De Loos, F. B. Kiewiet, and M. J. Van Der Wiel. How to Realize Uniform Three-Dimensional Ellipsoidal Electron Bunches. *Physical Review Letters*, 93(9):094802, Aug. 2004. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.93.094802. URL <https://link.aps.org/doi/10.1103/PhysRevLett.93.094802>.
- [6] Mahnke, Christoph, Li, Chen, Ahmed, Areeb, Akcaalan, Oender, Brinker, Frank, Chen, Ye, Good, James, Grosse-Wortmann, Uwe, Guetg, Marc, Ilia, Denis, Kschuev, Nick, Mohr, Christian, Panuganti, Harsha, Pressacco, Federico, Schaper, Lucas, Schulz, Sebastian, Schreiber, Siegfried, Tavakol, Hamed, Tünnermann, Henrik, Vidoli, Caterina, Vogt, Matthias, Winkelmann, Lutz, and Hartl, Ingmar. Novel Photocathode Lasers for the Hard- and Soft-X-ray Free Electron Lasers EuXFEL and FLASH. *EPJ Web Conf.*, 307:04001, 2024. doi: 10.1051/epjconf/202430704001. URL <https://doi.org/10.1051/epjconf/202430704001>.
- [7] P. Oliveira, S. Addis, J. Gay, K. Ertel, M. Galimberti, and I. Musgrave. Control of temporal shape of nanosecond long lasers using feedback loops. *Optics Express*, 27(5):6607, Mar. 2019. ISSN 1094-4087. doi: 10.1364/OE.27.006607. URL <https://opg.optica.org/abstract.cfm?URI=oe-27-5-6607>.
- [8] R. Paschotta, J. Nilsson, A. Tropper, and D. Hanna. Ytterbium-doped fiber amplifiers. *IEEE Journal of Quantum Electronics*, 33(7):1049–1056, July 1997. ISSN 1558-1713. doi: 10.1109/3.594865. URL <https://ieeexplore.ieee.org/document/594865/?arnumber=594865>. Conference Name: IEEE Journal of Quantum Electronics.

- [9] M. Schmidt, K. Cvecek, J. Duflou, F. Vollertsen, C. B. Arnold, and M. J. Matthews. Dynamic beam shaping—Improving laser materials processing via feature synchronous energy coupling. *CIRP Annals*, 73(2):533–559, 2024. ISSN 0007-8506. doi: <https://doi.org/10.1016/j.cirp.2024.05.005>. URL <https://www.sciencedirect.com/science/article/pii/S0007850624001215>.
- [10] S.-H. Shim and M. T. Zanni. How to turn your pump–probe instrument into a multidimensional spectrometer: 2D IR and Vis spectroscopies via pulse shaping. *Physical Chemistry Chemical Physics*, 11(5):748–761, 2009. doi: [10.1039/B813817F](https://pubs.rsc.org/en/content/articlelanding/2009/cp/b813817f). URL <https://pubs.rsc.org/en/content/articlelanding/2009/cp/b813817f>. Publisher: Royal Society of Chemistry.
- [11] C. W. Siders, J. L. W. Siders, A. J. Taylor, S.-G. Park, and A. M. Weiner. Efficient high-energy pulse-train generation using a 2ⁿ-pulse michelson interferometer. *Applied Optics*, 37(22):5302, Aug. 1998. ISSN 0003-6935, 1539-4522. doi: [10.1364/AO.37.005302](https://doi.org/10.1364/AO.37.005302). URL <https://opg.optica.org/abstract.cfm?URI=ao-37-22-5302>.
- [12] A. M. Weiner. Ultrafast optical pulse shaping: A tutorial review. *Optics Communications*, 284(15):3669–3692, July 2011. ISSN 00304018. doi: [10.1016/j.optcom.2011.03.084](https://doi.org/10.1016/j.optcom.2011.03.084). URL <https://linkinghub.elsevier.com/retrieve/pii/S0030401811003750>.
- [13] A. Yogo, K. Mima, N. Iwata, S. Tosaki, A. Morace, Y. Arikawa, S. Fujioka, T. Johzaki, Y. Sentoku, H. Nishimura, A. Sagisaka, K. Matsuo, N. Kamitsukasa, S. Kojima, H. Nagatomo, M. Nakai, H. Shiraga, M. Murakami, S. Tokita, J. Kawanaka, N. Miyanaga, K. Yamanoi, T. Norimatsu, H. Sakagami, S. V. Bulanov, K. Kondo, and H. Azechi. Boosting laser-ion acceleration with multi-picosecond pulses. *Scientific Reports*, 7(1):42451, Feb. 2017. ISSN 2045-2322. doi: [10.1038/srep42451](https://doi.org/10.1038/srep42451). URL <https://www.nature.com/articles/srep42451>. Publisher: Nature Publishing Group.

A Adaptive Pulse Shaping Algorithm

The adaptive pulse shaping procedure alternates between system identification and control optimization. Both subroutines are based on gradient methods and are here implemented by Adam optimizers with a global learning rate schedule. The control parameters are initialized using the Fourier transform of the target pulse amplitude, providing a physics-informed starting point that assumes linear propagation. The system identification step uses a sliding buffer of the most recent B measurements.

Algorithm 1 Adaptive Pulse Shaping

- 1: **Input:** Initial model $\theta^{(0)}$, target y^* , buffer size B , control constraints \mathcal{U} , max iterations n_{iter}
 - 2: Initialize control $u^{(0)} = \mathcal{F}(\sqrt{y^*})$, measurement buffer $\mathcal{B} = \emptyset$
 - 3: **for** $k = 0, 1, 2, \dots, n_{\text{iter}} - 1$ **do**
 - 4: Acquire measurement $y_{\text{exp}}^{(k)} = F(u^{(k)})$
 - 5: Update buffer $\mathcal{B} \leftarrow \mathcal{B} \cup \{(u^{(k)}, y_{\text{exp}}^{(k)})\}$, keep most recent B pairs
 - 6: $\theta^{(k+1)} \leftarrow \text{SYSTEMIDENTIFICATION}(\mathcal{B}, \theta^{(k)})$
 - 7: $u^{(k+1)} \leftarrow \text{CONTROLOPTIMIZATION}(y^*, \theta^{(k+1)}, u^{(k)})$
 - 8: **if** $\|f(u^{(k+1)}, \theta^{(k+1)}) - y^*\|_2 < \varepsilon_{\text{global}}$ **then break**
 - 9: **end for**
 - 10: **Output:** Optimized control u^* and model parameters θ^*
-

Algorithm 2 Subroutine: SYSTEMIDENTIFICATION($\mathcal{B}, \theta_{\text{init}}$)

```
1: Input: Buffer  $\mathcal{B}$ , initial parameters  $\theta_{\text{init}}$ , max iterations  $n_{\text{sys\_id}}$ , tolerance  $\varepsilon_\theta$ 
2: Initialize  $\theta \leftarrow \theta_{\text{init}}$ , Adam optimizer  $\text{Adam}_\theta$ 
3: for  $j = 0, 1, \dots, n_{\text{sys\_id}} - 1$  do
4:   Compute loss  $\mathcal{L}_\theta = \frac{1}{|\mathcal{B}|} \sum_{(u_i, y_i) \in \mathcal{B}} \|f(u_i, \theta) - y_i\|_2^2$ 
5:   Compute gradient  $g_\theta \leftarrow \nabla_\theta \mathcal{L}_\theta$ 
6:   Update  $\theta \leftarrow \text{Adam}_\theta(\theta, g_\theta, \eta_\theta)$ 
7:   if  $\|g_\theta\|_2 < \varepsilon_\theta$  then break {Early stopping}
8: end for
9: return  $\theta$ 
```

Algorithm 3 Subroutine: CONTROLOPTIMIZATION($y^*, \theta, u_{\text{init}}$)

```
1: Input: Target  $y^*$ , model parameters  $\theta$ , initial control  $u_{\text{init}}$ , max iterations  $n_{\text{optim}}$ , tolerance  $\varepsilon_u$ 
2: Initialize  $u \leftarrow u_{\text{init}}$ , Adam optimizer  $\text{Adam}_u$ 
3: for  $j = 0, 1, \dots, n_{\text{optim}} - 1$  do
4:   Compute loss  $\mathcal{L}_u = \|f(u, \theta) - y^*\|_2^2 + \lambda R(u)$ 
5:   Compute gradient  $g_u \leftarrow \nabla_u \mathcal{L}_u$ 
6:   Update  $u \leftarrow \text{Adam}_u(u, g_u, \eta_u)$ 
7:   Project onto constraints  $u \leftarrow \text{Project}_{\mathcal{U}}(u)$ 
8:   if  $\|g_u\|_2 < \varepsilon_u$  then break {Early stopping}
9: end for
10: return  $u$ 
```

The global learning rate scheduler updates η_θ and η_u based on cosine annealing with warmup. The regularization term $R(u)$ penalizes filter magnitude away from the spectral center, promoting physically plausible shaping profiles. It is defined as

$$R(u) = \sum_{i=1}^N |u_i|^2 \left[1 - \exp \left(- \left(\frac{\omega_i - \omega_{\text{center}}}{\Delta\omega} \right)^8 \right) \right],$$

where ω_i denote the discrete frequencies associated with each control parameter, ω_{center} is the central frequency of the pulse spectrum, and $\Delta\omega$ represents the effective spectral bandwidth.

Based on the system's physical characteristics, we set $\Delta\omega = 5 \times 10^{11}$ Hz (corresponding to approximately 2 nm at 1030 nm). This regularization suppresses excessive modulation at the spectral edges—where shaping has limited physical impact and higher sensitivity to noise—while maintaining flexibility near the spectral region of interest. The control solutions are always constrained by the defined hardware limits \mathcal{U} .

The variable $\varepsilon_{\text{global}}$ acts as the primary stopping condition for the entire adaptive pulse shaping procedure, while the individual subroutines, SYSTEMIDENTIFICATION and CONTROLOPTIMIZATION, have their own internal stopping criteria (ε_θ and ε_u) to determine if they've found a good solution.

B Photoinjector Laser

The front-end of the laser system is composed of a Yb-fiber mode-locked oscillator, a passive fiber stretcher, a fiber-coupled spatial light modulator (SLM), and Yb-fiber amplifiers. This setup generates a pulse train at a center wavelength of 1030 nm. The output is then fed into a Yb:YAG power amplifier. After compression, the pulse is up-converted to a 257.5 nm wavelength via two stages of second harmonic generation using LBO and BBO crystals. A separate Yb-fiber amplifier generates 150 fs, 1030 nm pulses to characterize the power amplifier output using a second harmonic generation (SHG) cross-correlator.

Our differentiable model replicates the photoinjector laser system from oscillator output to the cross-correlation measurement point. The layout of the modeled laser is shown in Figure 1(a).

The input pulse in the model is defined as a transform-limited Gaussian pulse with a duration of 600 fs. While this is not necessarily representative of the oscillator pulse, it provides a convenient starting point, and the system identification algorithm optimizes it through a learnable spectral filter module.

Similarly, at the input of the SLM, we account for wavelength-dependent imperfections and potential nonlinearities by introducing a learnable transfer function, which is also optimized during system identification.

For the solid-state power amplifier, the complex gain dynamics and spectral filtering are simplified to a static Gaussian bandpass filter with an initial bandwidth of 2 nm. This captures the dominant spectral narrowing effect while significantly reducing computational complexity.

In all cases, the model leverages learnable components to adaptively correct for system imperfections, ensuring an accurate and flexible representation of the physical laser system.

Details about the photoinjector system can be found in [6].