Hardware-in-the-Loop Nonlinear Photonic Computing via Noise-Aware Digital Twins and Lagrangian Optimization

Hua Tu

1. Abstract

Nonlinear optical systems offer a compelling physical substrate for analog computing due to their intrinsic ultrafast and non-Gaussian responses. However, training such hardware directly remains challenging because real optical processes are non-differentiable, noisy, and constrained by physical limits. In this work, we present a **hardware-in-the-loop nonlinear photonic computing framework** based on **second-harmonic generation (SHG)**, enabled by a **differentiable noise-aware digital twin** and **constrained optimization**.

I independently developed a first-principles SHG forward model with a learned differentiable surrogate that captures both deterministic nonlinear responses and device-specific variations. This digital twin is integrated into a **Physics-Aware Training (PAT)** pipeline, where forward propagation is performed on real hardware while gradients are computed through the surrogate model. To ensure physically valid actuation, we incorporate **Lagrangian constraints** that explicitly enforce hardware-safe bounds during optimization.

Using this hybrid framework, we demonstrate end-to-end trainable nonlinear photonic computation on a vowel recognition task, achieving **85% classification accuracy**. Our results show that nonlinear photonic hardware can be treated as a programmable computational layer when combined with differentiable modeling and constraint-aware training. This approach provides a general pathway for optimizing non-differentiable photonic systems and extends naturally to other nonlinear and quantum optical platforms.

2. System Architecture

2.1 Input Encoding: Spectral Amplitude Modulation via DMD

Input data are encoded as **spectral amplitude vectors** rather than spatial patterns. Each input specifies amplitude weights over a predefined wavelength grid, which are mapped onto a **binary digital micromirror device (DMD)** through wavelength-to-space dispersion. Individual micromirrors thus modulate distinct spectral components.

This spectral encoding provides a stable interface between numerical data and nonlinear optical hardware, while avoiding sensitivity to spatial mode variations.

2.2 First-Principles Second-Harmonic Generation Modeling and Dataset Generation

The spectrally shaped fundamental light is injected into a nonlinear medium, where **second-harmonic generation** implements a fixed quadratic transformation governed by χ^2 light–matter interaction. The resulting second-harmonic spectrum depends nonlinearly on the input amplitudes and constitutes the physical forward mapping for a neural network framework.

A **first-principles SHG based on coupled-wave equations is developed in this study** as a **physics-informed baseline**. Experimentally measured input–output pairs are collected across the operating parameter space and serve as training data for a digital surrogate that reflects real hardware behavior.

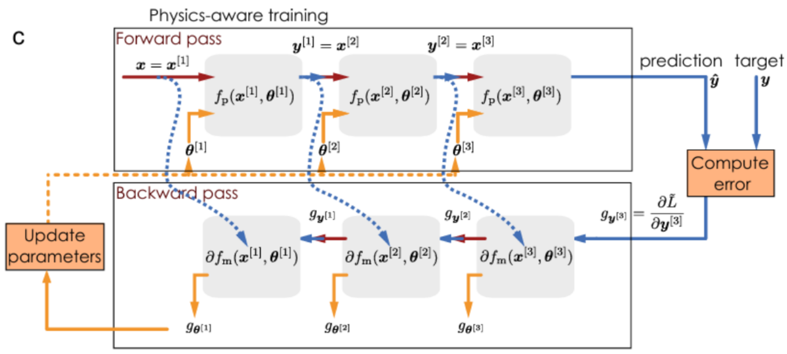
2.3 Noise-Aware Digital Twin Training for Backpropagation

Because the physical SHG process is non-differentiable and noisy, direct backpropagation through hardware is infeasible. To address this, we train a **noise-aware digital twin** on experimentally acquired SHG datasets. The digital twin captures both the deterministic nonlinear response and stochastic hardware-induced deviations.

Once trained, the digital twin is used exclusively in the **backward pass** to provide surrogate gradients for optimization.

2.4 Physics-Aware Training (PAT) Combining SHG Forward and Digital Twin Backpropagation

Figure 1. Physics-Aware Training (PAT) strategy, where forward propagation is executed on the five-layer physical SHG and gradients are computed through the digital twin.



This hybrid scheme preserves fidelity to real hardware dynamics while enabling efficient gradient-based learning. Optimization is performed under explicit constraints to enforce physically valid actuation.

II. Experimental Workflow

The experimental workflow is organized into modular stages covering optical encoding, SHG spectral acquisition, digital-twin training, and physics-aware optimization. Implementation details are omitted to emphasize system-level methodology.

Input patterns are applied to a DMD through **wavelength-to-space dispersion,** and the resulting SHG spectra are measured using a calibrated spectrometer. The raw 2D spectral images are reduced to 1D spectra and further compressed to fixed-dimensional outputs for learning.

Measured **SHG spectra** are normalized and compressed from **448 dimensions to a 50-dimensional representation** via interpolation and averaging, enabling efficient downstream learning while preserving spectral structure.

Experimental control and data acquisition are implemented via a custom **Python-GPIB** interface; implementation details are omitted for clarity.

Wavelength-to-pixel mappings for both the DMD and spectrometer are obtained through **prior calibration** and reused throughout all experiments.

3. SHG-Based Physical Forward Model and Digital Twin Training

3.1 Physics-inspired SHG Forward Operator

To emulate the experimental second-harmonic generation (SHG) nonlinear transformation, I construct a physics-inspired forward operator in which output spectral components arise from **pairwise interactions of input spectral amplitudes**. Conceptually, this mapping follows the frequency-mixing relation

$$\frac{1}{\lambda_3} = \frac{1}{\lambda_1} + \frac{1}{\lambda_2}$$

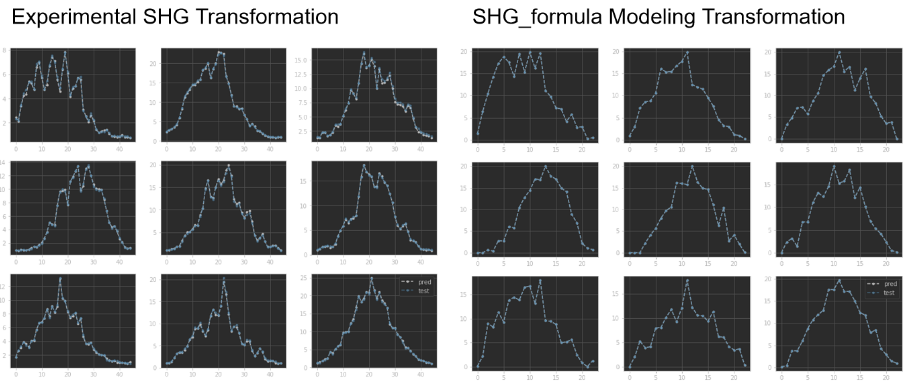
For each discrete output channel, corresponding input pairs are identified on a predefined wavelength grid, and the output intensity is computed as the weighted sum of pairwise amplitude products.

This SHG-based forward layer serves as a **structured nonlinear operator** that captures the dominant interaction Of physical system, while remaining compatible with efficient numerical evaluation and integration into a hardware-in-the-loop neural network training.

3.2 Dataset Generation

Using identical spectral-amplitude inputs, I compare the responses of the physical SHG hardware and the SHG-inspired forward operator.

Figure 2. Comparison between the SHG hardware response and the SHG_formula forward layer under identical spectral-amplitude inputs, showing consistent nonlinear trends and overall spectral structure.



The comparison demonstrates that the forward operator reproduces the characteristic nonlinear mixing behavior observed in experiment at the level required for system-level modeling. While fine-scale spectral features differ due to experimental filtering and phase-matching effects, the overall response structure and nonlinearity are well aligned, validating the operator as a suitable abstraction for downstream learning and analysis.

The resulting paired input–output data are used to construct training datasets for the digital twin described below.

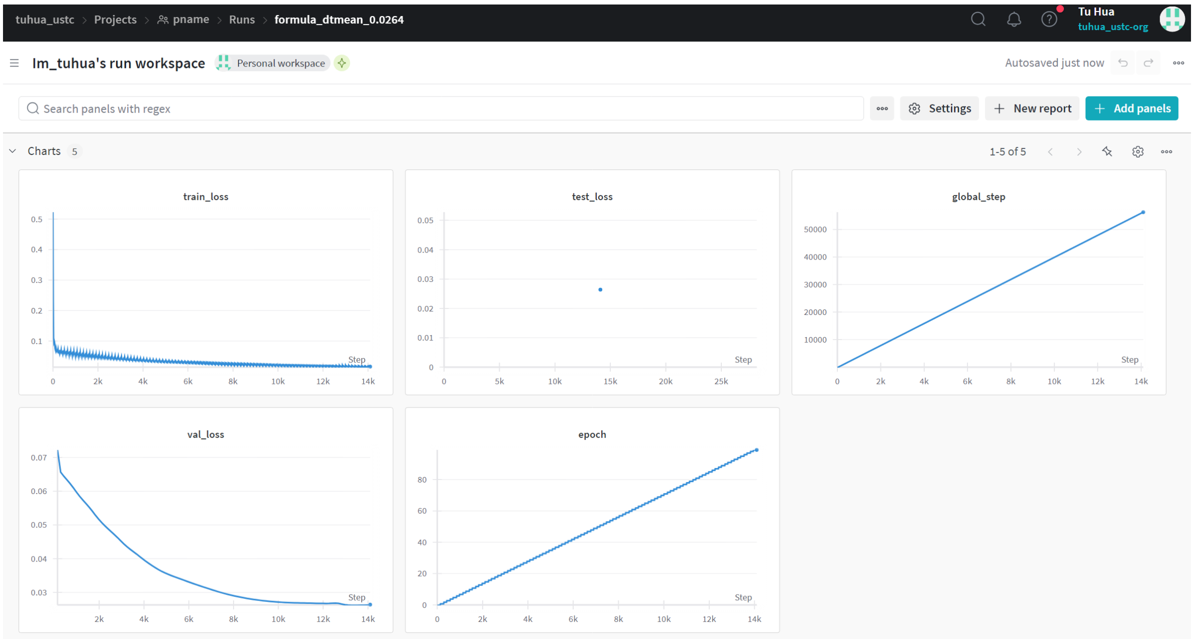
3.3 Noise-Aware Digital Twin Training

Why we need a differentiable noise-aware SHG model?

Direct backpropagation through the physical SHG process is infeasible due to measurement noise, hardware variability, and the absence of analytical gradients. To enable **gradient-based optimization**, I introduced a noise-aware digital twin trained on experimentally acquired SHG datasets.

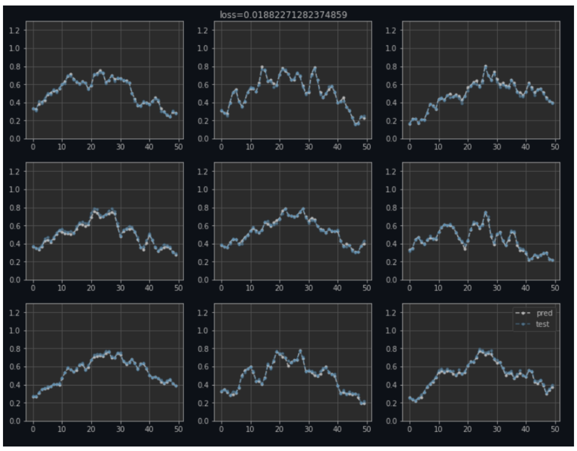
The digital twin is designed to learn two complementary aspects of the hardware response:
(i) the **deterministic nonlinear mapping** implemented by the SHG process, and
(ii) stochastic deviations arising from **experimental noise** and device-specific fluctuations.

Figure 3. shows representative training and validation loss curves, indicating stable convergence without overfitting. Once trained, the digital twin is used **exclusively in the backward pass** to provide surrogate gradients, while all forward propagation during training is executed on the physical hardware. This separation preserves fidelity to real system dynamics while enabling efficient optimization within the Physics-Aware Training framework.



Comparison between experimentally measured SHG spectral outputs (solid lines) and digital twin predictions (dashed lines) under identical spectral-amplitude inputs. The digital twin accurately captures both the overall nonlinear spectral structure and local amplitude variations across the operating regime, providing reliable surrogate gradients for Physics-Aware Training. Minor deviations reflect stochastic measurement noise and device-specific fluctuations inherent to the physical system.

Figure 4. Digital twin fitting results for the SHG-based nonlinear transformation.



4. Constraint-Aware Optimization in Physics-Aware Training

4.1 Constraint-Aware Optimization

A central challenge in Physics-Aware Training (PAT) for nonlinear photonic hardware is enforcing **physically valid actuation during optimization**. In the SHG-based system, both input spectral amplitudes and internal learnable parameters are constrained by DMD modulation limits, optical power budgets, and nonlinear saturation. Unconstrained gradient updates can therefore drive the system into physically inaccessible or unstable regimes.

To address this, I incorporate **constraint-aware optimization** into the PAT framework using **Lagrangian penalty terms**. Instead of hard clipping, which introduces non-differentiability and disrupts gradient flow, soft penalties enforce bounded actuation while preserving smooth optimization dynamics, ensuring informative gradients near constraint boundaries.

Constraints are applied **at each SHG layer** to regulate both propagated spectral amplitudes and learnable parameters, with a **stronger penalty at the first layer** to suppress early-stage amplification and prevent error accumulation across the network.

Empirically, constraint-aware optimization is essential for stable Physics-Aware Training. Without constraints, optimization frequently diverges or converges to physically invalid solutions, even with accurate digital-twin gradients. With constraints enabled, training remains stable and converges reliably under hardware-in-the-loop settings.

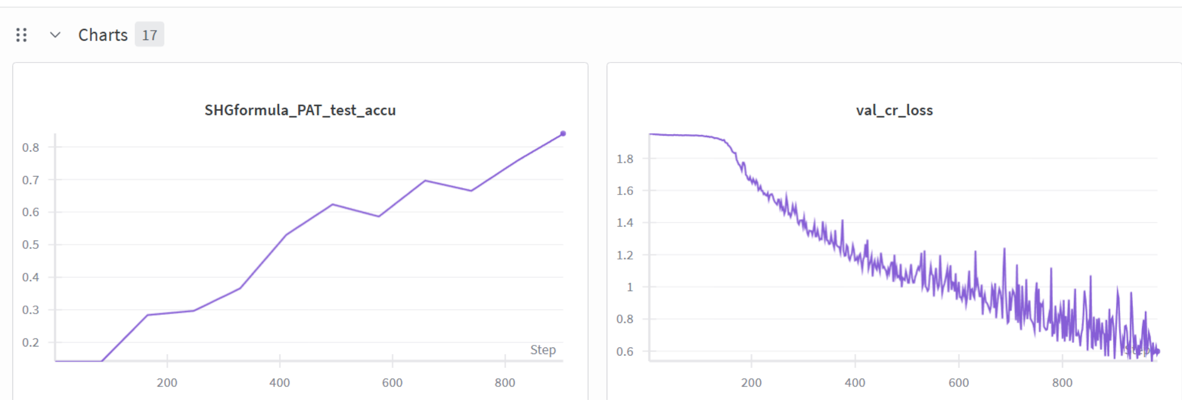
4.2 Physics-Aware Training Results on Vowel Recognition

To validate the Physics-Aware Training (PAT) under physical constraints, I apply my developed digital twins and the framework to a vowel recognition task. Input samples are encoded as spectral-amplitude vectors, propagated through an SHG-based nonlinear layer, and trained end-to-end using constraint-aware optimization.

Training follows the PAT paradigm: forward propagation is executed through the SHG-based nonlinear transformation, while gradients are computed via the noise-aware digital twin.

Due to computational constraints, training is conducted for up to 600 epochs when using the SHG-based forward operator. Both training and validation losses exhibit a clear downward trend, indicating successful task learning, with moderate fluctuations attributable to the sharp spectral features of the SHG nonlinearity. Crucially, constraint-aware optimization prevents divergence and ensures convergence to physically valid solutions across repeated runs.

Figure 5. Under this training framework, the network achieves a **test accuracy of 85%** on the vowel recognition task. This result demonstrates that SHG-based optical nonlinearities, when combined with physics-aware digital twins and constraint-aware optimization, can function as effective and trainable computational layers for practical classification tasks.



5. Conclusion

This work demonstrates that nonlinear photonic hardware, despite being non-differentiable, noisy, and physically constrained, can be **systematically trained as a programmable computational substrate** when embedded in a Physics-Aware Training framework.

By combining **hardware-in-the-loop forward propagation**, a **noise-aware differentiable digital twin**, and **constraint-aware optimization**, I establish an end-to-end training pipeline that respects physical limits while retaining the efficiency of gradient-based learning. Using second-harmonic generation (SHG) as a representative χ^2 nonlinear process, I show that experimentally measured optical nonlinearities can be abstracted into structured forward operators and paired with learned surrogates that provide reliable gradients for optimization.

Crucially, the introduction of **Lagrangian constraint enforcement** is shown to be essential for stable training. Rather than treating constraints as implementation details, this work elevates them to a core algorithmic component, ensuring physically valid actuation and preventing divergence in both hardware-in-the-loop and surrogate-based training. Under these constraints, the system achieves **85% accuracy on a vowel recognition task**, validating the feasibility of nonlinear photonic computation for practical learning problems.

Beyond the specific SHG platform, the framework presented here is **model-agnostic and extensible**. The separation of physical forward execution and differentiable backward approximation naturally generalizes to **other nonlinear and quantum optical systems where direct differentiation is impossible**. As such, this work provides a principled pathway toward training complex photonic and quantum hardware using modern learning paradigms, positioning physical dynamics not as obstacles, but as computational resources.