BOSON^{-1}: Understanding and Enabling Physically-Robust Photonic Inverse Design with Adaptive Variation-Aware Subspace Optimization

Pingchuan Ma[1], Zhengqi Gao[2], Amir Begovic[3], Meng Zhang[3], Haoyu Yang[4], Haoxing Ren[4],
Rena Huang[3], Duane S. Boning[2], Jiaqi Gu[1]
[1]Arizona State University, [2]MIT, [3]RPI, [4]Nvidia
{pingchua, jiaqigu}@asu.edu

Abstract— **Nanophotonic device design aims to optimize photonic structures to meet specific requirements across various applications. Inverse design has unlocked non-intuitive, high-dimensional design spaces, enabling the discovery of compact, high-performance device topologies beyond traditional heuristic or analytic methods. The adjoint method, which calculates analytical gradients for all design variables using just two electromagnetic simulations, enables efficient navigation of this complex space. However, many inverse-designed structures, while numerically plausible, are difficult to fabricate and highly sensitive to physical variations, limiting their practical use. The discrete material distributions with numerous local-optimal structures also pose significant optimization challenges, often causing gradient-based methods to converge on suboptimal designs. In this work, we formulate inverse design as a fabrication-restricted, discrete, probabilistic optimization problem and introduce BOSON^{-1}, an end-to-end, adaptive, variation-aware subspace optimization framework to address the challenges of manufacturability, robustness, and optimizability. We explicitly consider the fabrication process and differentiably optimize the design in the fabricable subspace. To overcome optimization difficulty, we propose dense target-enhanced gradient flows to mitigate misleading local optima and introduce a conditional subspace optimization strategy to create high-dimensional tunnels to escape local optima. Furthermore, we significantly reduce the prohibitive runtime associated with optimizing across exponential variation samples through an adaptive sampling-based robust optimization method, ensuring both efficiency and variation robustness. On three representative photonic device benchmarks, our proposed inverse design methodology BOSON^{-1} delivers fabricable structures and achieves the best convergence and performance under realistic variations, outperforming prior arts with 74.3% post-fabrication performance. We open-source our codes at link.**

Fig. 1: Inverse design often yields non-fabricable devices. Optimization difficulty leads to suboptimal designs.

I. INTRODUCTION

Integrated photonics has shown a wide range of applications in computing, communication, and sensing. Currently, many photonic devices are manually architected by tuning a few design parameters via inefficient trial and error, which relies heavily on expert knowledge and time-consuming simulations. In contrast, inverse design requires minimal physical prior knowledge and opens up non-intuitive, high-dimensional design spaces, making it possible to discover highly efficient and compact device designs [8], [13]. The adjoint method-based inverse design [8] is particularly powerful for its ability to compute analytical gradients of an objective with respect to high-dimensional design variables using only two simulations.
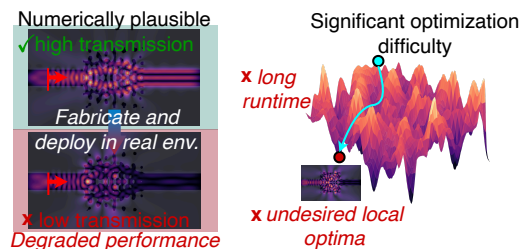
While the adjoint inverse design can produce numerically plausible designs, a significant gap exists between pre-fab and post-fab performance. As illustrated in Fig. 1, the inverse-optimized design exhibits high performance, but the tiny structures within the design pattern are **non-manufacturable**, leading to severe post-fabrication performance degradation. Additionally, **fabrication variations** during lithography and etching, along with **operational variations** such as temperature drift, introduce robustness concerns that further compromise the performance of inverse-designed devices. Previous approaches have attempted to address manufacturability issues by controlling the minimum feature size (MFS) during optimization through heuristic methods, such as blurring or adding curvature penalties, to eliminate non-fabricable structures and mitigate post-fabrication performance drops [1], [4], [6], [9], [12], [16]–[19], [23]. To enhance fabrication robustness, prior work models variations as uniform erosion and dilation of the device geometry and simultaneously optimizes objectives under different variation corner cases. However, this over-simplified method only marginally improves robustness as it fails to capture actual variations accurately. When considering multiple variation effects, exhaustive Monte Carlo sampling of all corner cases induces **exponential simulation cost**. Beyond fabricability, solving this **discrete high-dimensional stochastic optimization problem is particularly challenging**. As the adjoint optimizer is only driven by a single objective sparsely defined in the device output port, it leads to a **poor objective landscape** shown in Fig. 1, making the optimization highly sensitive to initialization and prone to getting trapped in unreasonable suboptimal solutions.

To address the photonic device inverse design challenge, which requires both fabricability and robustness, we formulate this task as a fabrication-restricted, robust stochastic optimization problem. We propose a novel inverse design framework, BOSON^{-1}, which enables effective optimization directly within the fabricable subspace with full variation awareness, ensuring efficient inverse design toward robust device structures. Our main contributions are as follows:

- We provide a comprehensive analysis of the fabricability and optimization challenges in inverse design and introduce BOSON^{-1}, an adaptive, variation-aware inverse design framework for physically robust photonic devices.
- **Fabrication-Aware Subspace Optimization**: BOSON^{-1} integrates differentiable fabrication modeling into adjoint optimization to ensure devices in the fabricable subspace.
- **Loss Landscape Reshaping**: BOSON^{-1} largely reduces the specious local optima by introducing auxiliary dense objectives, enhancing gradient flow and improving optimization.
- **Local Optima Escaping**: We introduce a light-concentrated initialization method with conditional subspace relaxation, facilitating escaping local optima toward better solutions.
- **Linear-Cost Variation-aware Optimization**: We propose an adaptive variation optimization method based on novel axial corner sampling and worse-case optimization, reducing simulation costs from exponential to linear.
- On three photonic device benchmarks, our BOSON^{-1} achieves 74.3% performance enhancement on average compared to previous art, enabling variation-robust photonic inverse design with high efficiency.

II. PRELIMINARY

We will briefly illustrate the basic concepts in fabrication and variations and review previous work focusing on improving fabricability and robustness.

A. Fabrication and Operation Variations

The binary photonic device pattern from inverse design often deviates from the real device due to the limited resolution in fabrication and the non-ideality in actual operation, e.g., temperature drift. We decoupled this complication process into the following three steps as shown in Fig. 2.
Lithography. During photolithography, light shines through the mask to transfer the design pattern onto the wafer, as shown in step 1 of Fig. 2. However, due to diffraction, the light intensity distribution on the wafer differs from the intended design, often leading to the loss of small features. Moreover, lithography in practice induces variations in focus, defocus, and exposure dose, causing further structure discrepancy from the desired one. In this work, we utilize a differentiable lithography model [22] based on the Hopkins diffraction equation to predict the three corner cases of the post-lithography pattern using the inverse-designed mask.
Etching. Following lithography, the etching process is applied to the exposed wafer, removing material from the exposed regions to form the device structures. We model this process

as a binarization projection applied to the continuous post-lithography pattern using a specific threshold, η. In practice, etching introduces spatial variations across the wafer. In this work, we model these spatially varying etching effects by characterizing η as a random field, utilizing the expansion optimal linear estimation (EOLE) method [15].
Operation Variations. The last step in Fig. 2 refers to the variation during practical operation that will cause the permittivity to drift from its nominal value. In this work, we consider the temperature-caused permittivity drift. Since we use air as the void (cladding) and silicon as the solid to build the photonic devices, we model the temperature-dependent Si permittivity by $\epsilon_{Si}(t) = (3.48 + 1.8 \times 10^{-4} \cdot (t - 300))^2$ [10]

B. Prior Fabrication-Robust Photonic Inverse Design Methods

Minimum feature size (MFS) control is a widely used technique to enhance fabricability. Designs can be intentionally parametrized to form smooth gating structures [1] that meet the MFS requirements of the foundry. However, this method is not applicable to complex 2-D patterns. Low-pass filters or blurring have been applied to remove high-frequency features from pre-fabrication patterns [12]. This method is only an approximation of the lithography process and cannot fully guarantee performance after actual lithography and etching. To account for etching variations, previous work has optimized both the nominal design and the dilated/eroded versions to achieve robust gating structures against over-etching and under-etching [1], [7], [20]. However, the uniform etching threshold used in this method oversimplifies the actual etching effects, which limits their robustness in real-world fabrication scenarios. Moreover, the exhaustive corner sweeping-based method [1] results in exponential simulation costs, making it unscalable for handling more complex variation models.

III. ROBUST PHOTONIC INVERSE DESIGN WITH ADAPTIVE VARIATION-AWARE SUBSPACE OPTIMIZATION

We first give a formulation of robust inverse design problem and analyze its fabrication constraints and robust optimization challenges. To solve those challenges, BOSON^{-1} proposes to directly optimize the design pattern in **a low-dimensional manufacturable subspace** and enable efficient and robust optimization via gradient-enhanced landscape reshaping, conditional subspace optimization, and adaptive variation sampling.

A. Problem Formulation

Our **goal** is to optimize the design variables $\theta \in \mathbb{R}^N$ so that the corresponding binary design pattern ϵ can obtain the maximum expected post-fabrication figure-of-merits (FoM)

under variations. We formulate this task as a **constrained, discrete, stochastic** optimization problem as follows,

$$\theta^* = \underset{\theta \in \Theta}{\mathrm{argmax}} \quad \mathbb{E}_{T,\eta,L} F(\epsilon(\theta)|\lambda_c),$$

$$\text{s.t.} \quad \tilde{\tilde{\rho}}' = (\mathcal{T}_t \circ \mathcal{E}_\eta \circ \mathcal{L}_l \circ \mathcal{P})(\theta)$$

$$\epsilon = \epsilon_v + (\epsilon_s - \epsilon_v) \cdot \tilde{\tilde{\rho}}'$$

$$\mathcal{P} : \theta \in \mathbb{R}^N \to \rho \in \{0,1\}^{N^x \times N^y} \qquad (1)$$

$$\mathcal{L}_l : \rho \in \{0,1\}^{N^x \times N^y} \to \bar{\rho} \in [0,1]^{N^x \times N^y}$$

$$\mathcal{E}_\eta : \bar{\rho} \in [0,1]^{N^x \times N^y} \to \tilde{\bar{\rho}} \in \{0,1\}^{N^x \times N^y}$$

$$\mathcal{T}_t : \tilde{\bar{\rho}} \in \{0,1\}^{N^x \times N^y} \to \tilde{\tilde{\rho}}' \in \{0, \alpha_t\}^{N^x \times N^y}$$

where θ are latent design variables that encode the topology of the device pattern ϵ. The compound mapping function from θ to ϵ includes the following transformations. (1) $\mathcal{P}$: it maps θ to a Boolean pattern ρ that indicates the probability of solid material ϵ_s and void space ϵ_v. In this work, we adopt a popular level set method [21] to parameterize the topology. (2) $\mathcal{L}_l$: the lithography model maps ρ to a non-binary post-litho pattern $\bar{\rho}$; (3) $\mathcal{E}_\eta$: the etching model maps the post-litho pattern to a binarized post-etching pattern $\tilde{\bar{\rho}}$; and (4) $\mathcal{T}_t$: it represents the actual operational condition that can potentially change the device permittivity.

The target is not to maximize the objective $F(\cdot)$ for central wavelength λ_c at a single nominal device design point, but the **expected performance over random variations during fabrication and operation** T, η, and L. $L \in \{l_{min}, l_{norm}, l_{max}\}$ refers to different lithography corners due to defocusing. η is the random spatially varying etching threshold field that leads to pattern distortion. T represents the undesired permittivity drift during actual device operation, e.g., temperature drift.

B. Understanding the Fabricability Constraint and Optimization Difficulty of Photonic Inverse Design

Based on the above complex formulation, we thoroughly analyze its *subspace optimization property and pinpoint sources of optimization difficulty*, which motivates our proposed BOSON^{-1} methodology.

1) Fabrication Constraints: We decouple the fabrication into two cascaded processes: lithography $\mathcal{L}_l$ and etching $\mathcal{E}_\eta$. **Lithography**. As shown in Fig. 2(a), fine-grained voids or solids are blurred and removed, and edges with shape curvature are smoothed since the feature sizes are smaller than the diffraction limit. This is the main mechanism that restricts the device patterns to the low-dimensional manufacturable subspace. **Etching**. As shown in Fig. 2(b), etching is modeled as a binarization projection of the post-litho patterns. Small features can hardly survive under the non-ideal etching threshold, e.g., under-etching removes small holes in Fig. 2(b).

Both processes have non-ideal variations in practice, which complicates the inverse design problem as a stochastic optimization over different fabrication conditions.

2) Optimization Difficulty: A common observation of photonic inverse design is that the adjoint optimization often converges to **sub-optimal or even unreasonable design patterns**. We attribute this phenomenon to the following reasons.

❶ The traditional single objective, e.g., transmission, leads to poor properties of the loss landscape, where **numerous undesired yet sharp local optima/saddle points** hinder effective exploration. ❷ The traditional **design objective is overly sparse**, i.e., transmission through a small monitor, which causes optimization instability due to vanishing gradients when no lights pass through the monitor. ❸ Due to the binary nature of the optimization, the **highly discrete** non-convex problem leads to high **sensitivity to initialization**. ❹ Stochastic optimization requires optimizing over the entire variation distribution. Naive Monte Carlo sampling or corner sweeping method has **a prohibitive cost and attention distraction** and is thus unstable, inefficient, and leads to sub-optimal robust design.

To tackle these difficulties, several questions have to be answered first:

❶ *What defines a good loss landscape for adjoint gradient-based photonic inverse design?* Finding a high-quality design becomes exceedingly challenging with a poorly structured loss landscape. In Fig. 3, we illustrate this using an optical isolator as an example, where the objective is to minimize transmission contrast ratio $\frac{E_{bwd}}{E_{fwd}}$ by maximizing forward transmission and minimizing backward transmission. This objective leads to a poor loss landscape where the optimizers are easily stuck at specious designs. For example, from a randomly initialized pattern, it converges to a low contrast result by reducing the backward transmission, even when the forward transmission is unreasonably low. Even with a good heuristic initialization, it converges to a sub-optimal point with unsatisfying E_{fwd}. The objective fails to guide the optimization to a desired space. Hence, it is essential to **reshape the loss landscape with fewer sharp yet undesired local optima**.

❷ *How can we ensure a richer gradient flow?* As shown in Fig. 3, the gradient flow is weak when the optimization objective is based solely on the power collected at the output port, widely adopted in photonic inverse design. Drawing an analogy to the gradient vanishing problem in neural networks, the optimization stagnates when all neurons are dead, e.g., due to ReLU activations. In the case of photonic inverse design, without careful initialization, strong local reflection and radiation occur easily in early optimization stages, leading to no light passing through the output monitor and thus vanishing gradients. This, in turn, traps the optimization in sub-optimal or unreasonable design points. We conclude that inverse design based on sparse objective is **inherently unstable from the gradient's perspective**. To mitigate this, inspired by dense supervision in neural network knowledge distillation [14], we are compelled to **provide *dense supervision* with auxiliary objectives to ensure a richer gradient flow, equivalently smoothing out the loss landscape**, preventing the optimization from stagnating in unsuitable patterns.

❸ *How can we efficiently sample the variation distribution for robust optimization?* Monte Carlo sampling is the common method to estimate the behavior of a distribution. As a heuristic, hardware designers tend to use min-max sampling to sweep all variation corners. However, as shown in Fig. 3, the **exponentially many variation corner cases make it unscalable** for
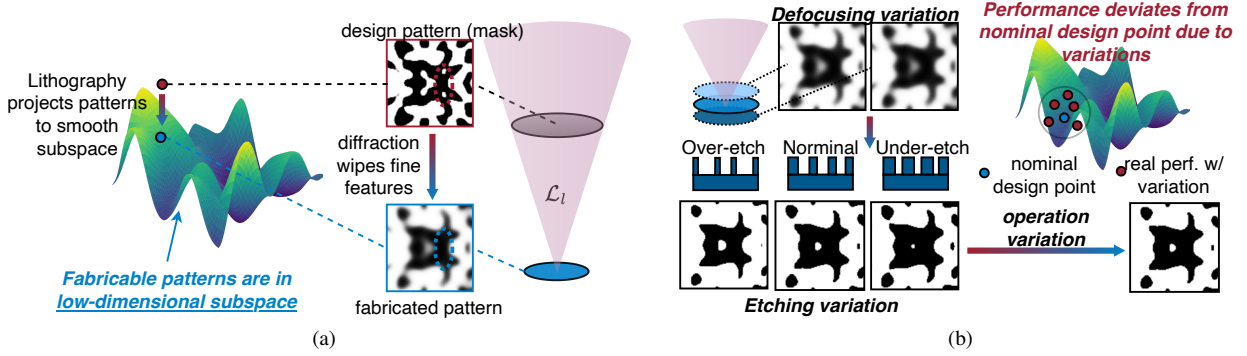
Fig. 2: (a) Lithography and etching during fabrication restricts manufacturable patterns in the subspace. (b) Fabrication variations (defocusing and etching) and operation variations cause performance deviation.
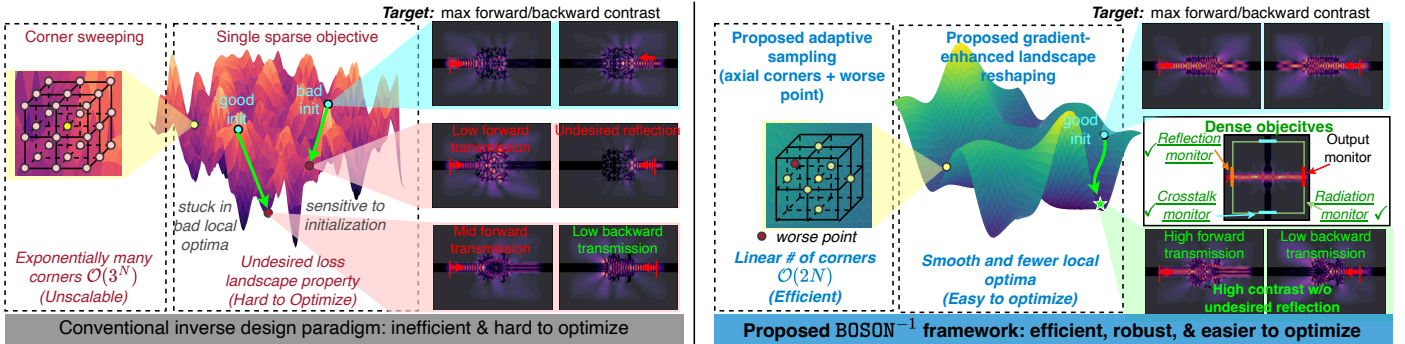

Fig. 3: Our proposed BOSON^{-1} framework enables efficient, robust optimization with better convergence and optimality.

this simulation-in-the-loop inverse design. To address this, it is crucial to draw samples adaptively, ideally with linear cost, to **balance optimization efficiency and variation robustness**. ❹ *How to initialize the design?* The discrete nature of the binary pattern design problem makes it sensitive to initialization. A randomly initialized pattern tends to scatter lights to different paths, which biases the converged pattern to an undesired design with many fabrication-unfriendly small structures. Besides, this can lead to a physically unstable local resonance state where the optimizer can hardly escape. To solve this problem, it is most effective if a **good heuristic initialization** can be found. Furthermore, it is preferred to reduce the initialization sensitivity by having a smooth loss landscape with effective optimization strategies to **escape from local optima**.

C. Fabrication-Restricted Subspace Optimization Toward Guaranteed Manufacturability

To minimize the post-fab performance gap in conventional methods using heuristic MFS control and a separate inverse lithography mask correction step [3], [5], BOSON^{-1} enables *end-to-end inverse design with explicit fabrication and variation modeling* in the optimization loop.

We adopt a Hopkins diffraction-based lithography modeling [22] $\mathcal{L}(\cdot)$ and a gradient-estimated etching modeling during device construction before evaluating the adjoint gradient to enable differentiable fab-aware topology optimization. This *re-parametrization* trick enforces the generated pattern to be aware of realistic smoothing and distortion from the actual

fabrication modeling, such that the optimization trajectory is restricted to the manufacturable subspace. Compared to mask correction, which tries to match the freely optimized pattern at the post-processing stage, our **subspace optimization method guarantees fabricability and eliminates the performance gap due to topology discrepancy** after correction.

D. Gradient-Enhanced Subspace Adjoint Optimization Toward Better Optimality

We solve the optimization difficulty of photonic inverse design from three aspects: (1) improve the loss landscape property, (2) better method to escape local optima, and (3) better initialization.

1) Gradient-Enhanced Objective Landscape Reshaping via Auxiliary Constraints: As discussed in Section III-B2, a sparse objective leads to numerous local optima that hinder optimization. Hence, we guide the optimization with auxiliary constraints $F_i(\epsilon|\lambda_c) \leq C_i$ and relax the inequality constraints as additional penalty terms in the objective as follows,

$$obj = F(\epsilon|\lambda_c) + \sum_i w_i \cdot [F_i(\epsilon|\lambda_c) - C_i]_+, \quad (2)$$

where $(\cdot)_+$ adds penalty only when constraints are violated. For the isolator design as an example, we encourage forward transmission higher than 80%, reflection less than 10%, backward radiation higher than 90%, etc. At the early stage, the penalty terms calculated on extra power monitors, as shown in Fig. 3, provide dense supervision and enhanced gradient flow that push the design towards a desired region. Bad and sharp local optima
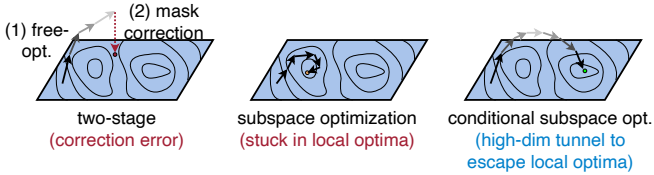
Fig. 4: High-dimensional tunnel for local minima escaping.

in the original loss landscape disappear in the reshaped design space. Moreover, gradient vanishing problems when using a small output port monitor are resolved as the added auxiliary monitors provide rich gradients for optimization. As the penalty terms disappear once constraints are satisfied, the optimization is driven by the main objective, eventually converging to a high-FoM and feasible design.

2) Subspace Local Optima Escaping via Relaxed High-Dimensional Tunneling: The main reason that the variation-aware optimization is easier to get stuck in the local optima is that the optimization is restricted in a low-dimensional fabricable subspace which is harder for the optimizer to escape sharp local optima. Furthermore, the gradients backpropagated through the lithography model are hurdled as the model filters out small features, which cause gradient vanishing on small structures. So instead of solely relying on the gradient from projected design pattern $\tilde{\rho}' = (\mathcal{T}_t \circ \mathcal{E}_\eta \circ \mathcal{L}_l \circ \mathcal{P})(\theta)$, we add a high-dimensional tunnel outside the fabricable subspace through pattern ρ as shown in Fig. 4, making the optimization objective a weighted sum of fabrication-aware and ideal term,

$$obj = p \cdot \mathbb{E}_{T,\eta,L}\left[\sum_i w_i F_i\left(\widetilde{\tilde{\epsilon}'}(\theta)|\lambda_c\right)\right] + (1-p) \cdot \sum_i w_i F_i\left(\epsilon(\theta)|\lambda_c\right). \quad (3)$$

p will gradually increase to 1 to ensure variation awareness and guarantee fabricability. With this technique, the subspace optimization can easily escape from the local optima.

3) Optical Path Concentrated Initialization: A well-chosen initialization helps prevent the optimization process from getting trapped in suboptimal points. In BOSON^{-1}, rather than randomly initializing the design parameters θ, which can lead to undesired initial optical behavior or even significant reflection that hinders optimization as discussed in Section III-B2, we initialize the design region to simple yet effective geometry with concentrated optical paths to ensure enough light passing through monitors with strong gradients.

E. Adaptive Sampling Strategy for Efficiency and Robustness

The number of corners increases exponentially with the number of parameters characterizing the fabrication error, which makes the prior exhaustive sampling method unscalable when considering multiple error sources. For example, in our error modeling, there are 3 lithography corners, 3 temperature corners, and 3 global etching thresholds if we consider a simpler etching model. In total, there are $3^3 = 27$ corners, requiring 54 simulations in one iteration. This motivates us to adopt an adaptive sampling strategy. Besides the nominal corner C_n, we sample 6 more corners, namely, C_L, C_T and C_η in which C_L represents the minimum and maximum corner of the lithography model and use the nominal value for the

TABLE I: Main result, the FoM for bending and crossing is the transmission efficiency, the higher the better; the FoM for isolator is the isolation contrast, the lower the better

Benchmark	Model	Fwd & bwd transmission	Avg FoM
Crossing	Density	N/A	0.916 → 0.0487
	InvFabCor-M-3	N/A	0.953 → 0.7
	BOSON^{-1}	**N/A**	**0.967**
	avg improvement: 60%		
Bending	Density	N/A	0.996 → 0.0141
	InvFabCor-M-3	N/A	0.935 → 0.691
	BOSON^{-1}	**N/A**	**0.982**
	avg improvement: 63%		
Optical isolator	Density	[0.676, 3.53e-06] → [0.0204, 0.0757]	4.89E-06 → 3.71
	InvFabCor-M-3	[0.152, 2.361e-4] → [0.0228, 0.0275]	0.00156 → 0.528
	BOSON^{-1}	**[0.8275, 0.0022]**	**0.00262**
	avg improvement: 100%		
total avg improvement: 74.3%			

temperature and global etching threshold η and so does the rest two as shown in Fig. 3. Based on this **axial sampling strategy**, we **reduce the complexity from exponential to linear**. However, these sampled corners still cannot model the spatially varying etching field. As such, inspired by the robust optimization on neural network [2], [11], we sample an additional worst corner C_{worst} by one-step gradient ascend on the temperature and the weight of the etching field basis. Combining axial corners and worst corners, we can ensure both efficiency and robustness.

IV. EVALUATION RESULTS

A. Evaluation Setup

a) Benchmarks: We adopt three representative photonic devices to demonstrate the effectiveness and efficiency of BOSON^{-1}: (1) a waveguide bending, where light is steered by $90°$; (2) a waveguide crossing, where light propagates through intersecting waveguides without crosstalk; and (3) an optical isolator, where light is converted from TM_1 mode to TM_3 mode in forward propagation with high transmission efficiency while the backward light is isolated via radiation.

b) Objectives: The optimization objectives are transmission efficiency for the bending and the crossing $\frac{E_{out}}{E_{in}}$. For the optical isolator, the objective is the isolation contrast $\frac{E_{bwd}}{E_{fwd}}$ where E_{bwd} and E_{fwd} refer to the transmission efficiency in backward and forward propagation.

B. Main Results

Benchmark Notations. Before presenting the results, we define the baseline notations. 'LS' and 'Density' denote level set or density-based methods for parameterizing the design, with BOSON^{-1} using level set by default. 'InvFabCor' refers to a two-stage process: first, levelset is used to achieve a high-performance design ρ^*, followed by mask optimization to match the post-fabrication pattern. '-M' indicates MFS control is included, while '-#' specifies the number of lithography corner to match. '-eff' signifies that the objective is transmission efficiency, not isolation contrast used in optical isolator. The arrow ($\rightarrow$) in the tables shows performance transitioning from the optimized result to the evaluated real-world performance.

To evaluate the performance of different design patterns produced by various optimization methods, we use Monte Carlo sampling, where lithography corners, random η fields, and
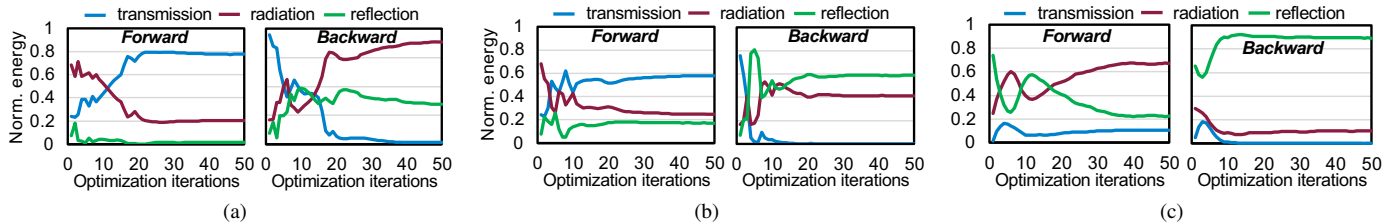
Fig. 5: Fabrication-aware optimization trajectories of optical isolator with forward/backward transmission, radiation, and reflection. No variation is added. (a) Proposed method (light-concentrated initialization, dense objectives). (b) Light-concentrated initialization with contrast as the single sparse objective. (c) Random initialization with contrast as the single sparse objective.

TABLE II: Ablation study of BOSON^{-1}

model	[fwd, bwd]	contrast ↓	degradation
BOSON^{-1}	**[0.8275, 0.0022]**	**2.62E-03**	**N/A**
- loss landscape reshaping	[0.4382, 0.0023]	5.41E-03	52%
- subspace relax	[0.8066, 0.0025]	3.25E-03	19%
exhaustive sample	[0.8395, 0.0113]	1.36E-02	81%
random init	[0.0610, 0.0356]	7.04E-01	100%

temperature are treated as random variables. We measured the average of 20 samples for each case under uniform distribution, and the result is shown in Table I. In all three benchmarks, traditional density-based inverse design produces numerically plausible results. However, without MFS control, these designs contain too many fine features that are non-fabricable, resulting in poor post-fabrication performance. While inverse fabrication correction finds high-performance designs near the fabricable space, performance degrades sharply after projection into the fabricable subspace, highlighting the need for direct optimization within this subspace. In the optical isolator benchmark, inverse fabrication correction failed to produce a viable device due to poor initialization. In contrast, BOSON^{-1} achieves the best performance across all benchmarks by directly optimizing the fabricable subspace, using well-chosen initialization, and applying dense supervision.

C. Ablation Study and Discussion

In the ablations study and discussion section, all the experiment results are obtained on optimizing the most challenging benchmark, the optical isolator, where various effects, such as reflection, resonance, and radiation, are involved.

a) Ablation Study: Table II compares our method with its variants, each lacking a key technique. Without loss landscape reshaping and dense supervision, contrast performance decreases by 52%, and more critically, the forward efficiency is severely compromised. Omitting subspace relaxation results in a 19% reduction in contrast, replacing adaptive sampling with exhaustive sampling leads to a dramatic 81% decrease in contrast, and most critically if a random initialization is used, the device becomes invalid.

b) Loss Landscape Reshaping and Dense Objectives: As shown in Fig. 5, without loss landscape reshaping, random initialization causes forward transmission efficiency to stagnate due to vanishing gradients. While the backward efficiency is low, giving a seemingly good contrast ratio, this is mainly due to unwanted reflection. Even with better initialization, the
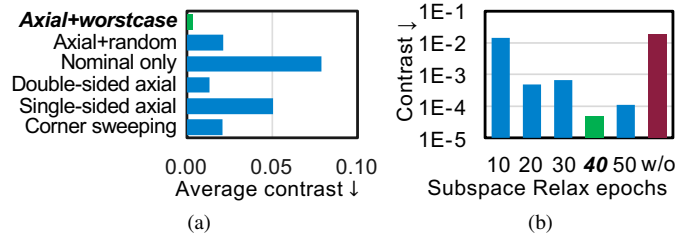

Fig. 6: (a) Comparison between different sampling strategies, Axial + worst case shows best average contrast; (b) Comparison between different relaxation epochs and no relaxation.

TABLE III: Compare different methods for optical isolator

model	Fwd & bwd transmission	Avg FoM
Density	[0.625, 4.14e-4]→[0.419, 0.298]	3.7e-4→0.763
Density-M	[0.405, 0.275]→[0.0883, 0.737]	0.674→9.83
LS	[0.0547, 0.00137]→[0.0126, 0.0339]	0.0252→2.91
LS-M	[0.722, 0.00140]→[0.604, 0.400]	0.00208→0.677
InvFabCor-1	[0.0547, 0.00138]→[0.0398, 0.0163]	0.0252→0.563
InvFabCor-3	[0.0547, 0.00138]→[0.0335, 0.0162]	0.0252→0.627
InvFabCor-M-1	[0.722, 0.00150]→[0.684, 0.0201]	0.00208→0.0306
InvFabCor-M-3	[0.722, 0.00150]→[0.687, 0.0179]	0.00208→0.0271
InvFabCor-M-3-eff	[0.990, 0.669]→[0.891, 0.674]	0.676→0.756
BOSON^{-1}	**[0.828, 0.0022]**	**0.00262**

poor loss landscape results in low backward efficiency and suboptimal forward efficiency. In contrast, reshaping the loss landscape and using dense objectives with multiple penalties guide the design to a high-performance solution with strong isolation contrast and good functionality.

c) Subspace Relaxation: Figure 6(b) shows the optimized contrast under different epochs of relaxation and the one without subspace relaxation, from which it is clear that subspace relaxation can greatly improve the contrast ratio by around 400 times. Note that the hyperparameter is searched on the nominal corner without variation.

d) Adaptive Sampling: We evaluate exhaustive sampling ($\mathcal{O}(3^N)$), single-sided axial sampling ($\mathcal{O}(N)$), and axial sampling ($\mathcal{O}(2N)$) in Fig. 6(a). Single-sided axial performs poorly due to its asymmetry, while axial sampling surprisingly outperforms exhaustive sampling by avoiding low-probability corner cases. Without accounting for variations (Nominal-only), performance drops significantly. Axial sampling with one worst-case corner boosts performance by 6× compared to two random samples with the same simulation cost.

e) Compare Different Methods: Table III compares different methods with good initialization. Our BOSON^{-1} outper-

forms the strongest baseline (two-stage variation-aware optimization with mask correction) with one order of magnitude better FoM, even with fabrication variations.

V. Conclusion

In this work, we propose BOSON^{-1}, a photonic device inverse design framework that addresses the manufacturing constraints by directly optimizing the device patterns in the fabricable subspace. Various optimization techniques have been integrated to reshape the lass landscape and mitigate the optimization difficulty for more efficient exploration toward variation-robust devices. We evaluate BOSON^{-1} across three photonic device inverse design tasks, demonstrating an average performance improvement of 74.3% over prior approaches.

References

[1] Mingkun Chen, Jiaqi Jiang, and Jonathan A. Fan. Design space reparameterization enforces hard geometric constraints in inverse-designed nanophotonic devices. *ACS Photonics*, 7(11):3141–3151, 2020.

[2] Pierre Foret, Ariel Kleiner, Hossein Mobahi, and Behnam Neyshabur. Sharpness-aware minimization for efficiently improving generalization. In *International Conference on Learning Representations*, 2021.

[3] Jhih-Rong Gao, Xiaoqing Xu, Bei Yu, and David Z Pan. Mosaic: Mask optimizing solution with process window aware inverse correction. In *Proceedings of the 51st Annual Design Automation Conference*, pages 1–6, 2014.

[4] Erez Gershnabel, Mingkun Chen, Chenkai Mao, Evan W Wang, Philippe Lalanne, and Jonathan A Fan. Reparameterization approach to gradient-based inverse design of three-dimensional nanophotonic devices. *ACS Photonics*, 10(4):815–823, 2022.

[5] Dusan Gostimirovic, Dan-Xia Xu, Odile Liboiron-Ladouceur, and Yuri Grinberg. Deep learning-based prediction of fabrication-process-induced structural variations in nanophotonic devices. *ACS Photonics*, 9(8):2623–2633, 2022.

[6] Alec M Hammond, Ardavan Oskooi, Mo Chen, Zin Lin, Steven G Johnson, and Stephen E Ralph. High-performance hybrid time/frequency-domain topology optimization for large-scale photonics inverse design. *Optics Express*, 30(3):4467–4491, 2022.

[7] Alec M Hammond, Ardavan Oskooi, Steven G Johnson, and Stephen E Ralph. Photonic topology optimization with semiconductor-foundry design-rule constraints. *Optics Express*, 29(15):23916–23938, 2021.

[8] Tyler W. Hughes, Momchil Minkov, Ian A. D. Williamson, and Shanhui Fan. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photonics*, 2018.

[9] Erfan Khoram, Xiaoping Qian, Ming Yuan, and Zongfu Yu. Controlling the minimal feature sizes in adjoint optimization of nanophotonic devices using b-spline surfaces. *Opt. Express*, 28(5):7060–7069, Mar 2020.

[10] J. Komma, C. Schwarz, G. Hofmann, D. Heinert, and R. Nawrodt. Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures. *Applied Physics Letters*, 101(4):041905, July 2012.

[11] Jungmin Kwon, Jeongseop Kim, Hyunseo Park, and In Kwon Choi. Asam: Adaptive sharpness-aware minimization for scale-invariant learning of deep neural networks, 18–24 Jul 2021.

[12] Simei Mao, Lirong Cheng, Houyu Chen, Xuanyi Liu, Zihan Geng, Qian Li, and Hongyan Fu. Multi-task topology optimization of photonic devices in low-dimensional fourier domain via deep learning. *Nanophotonics*, 12(5):1007–1018, 2023.

[13] Momchil Minkov, Ian AD Williamson, Lucio C Andreani, Dario Gerace, Beicheng Lou, Alex Y Song, Tyler W Hughes, and Shanhui Fan. Inverse design of photonic crystals through automatic differentiation. *Acs Photonics*, 7(7):1729–1741, 2020.

[14] Adriana Romero, Nicolas Ballas, Samira Ebrahimi Kahou, Antoine Chassang, Carlo Gatta, and Yoshua Bengio. Fitnets: Hints for thin deep nets, 2015.

[15] M. Schevenels, B.S. Lazarov, and O. Sigmund. Robust topology optimization accounting for spatially varying manufacturing errors. *Computer Methods in Applied Mechanics and Engineering*, 200(49):3613–3627, 2011.

[16] Mattias Schevenels, Boyan Stefanov Lazarov, and Ole Sigmund. Robust topology optimization accounting for spatially varying manufacturing errors. *Computer Methods in Applied Mechanics and Engineering*, 200(49-52):3613–3627, 2011.

[17] Mattias Schevenels, Boyan Stefanov Lazarov, and Ole Sigmund. Robust topology optimization accounting for spatially varying manufacturing errors. *Computer Methods in Applied Mechanics and Engineering*, 200(49-52):3613–3627, 2011.

[18] Martin F. Schubert, Alfred K. C. Cheung, Ian A. D. Williamson, Aleksandra Spyra, and David H. Alexander. Inverse design of photonic devices with strict foundry fabrication constraints. *ACS Photonics*, 9(7):2327–2336, 2022.

[19] Evan W Wang, David Sell, Thaibao Phan, and Jonathan A Fan. Robust design of topology-optimized metasurfaces. *Optical Materials Express*, 9(2):469–482, 2019.

[20] Fengwen Wang, Jakob S Jensen, and Ole Sigmund. Robust topology optimization of photonic crystal waveguides with tailored dispersion properties. *JOSA B*, 28(3):387–397, 2011.

[21] Michael Yu Wang, Xiaoming Wang, and Dongming Guo. A level set method for structural topology optimization. *Computer Methods in Applied Mechanics and Engineering*, 192(1):227–246, 2003.

[22] Haoyu Yang and Haoxing Ren. ILILT: Implicit learning of inverse lithography technologies. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 56319–56331. PMLR, 21–27 Jul 2024.

[23] You Zhou, Yixuan Shao, Chenkai Mao, and Jonathan A Fan. Inverse-designed metasurfaces with facile fabrication parameters. *Journal of Optics*, 26(5):055101, 2024.