ANGULAR SPECTRUM PROPAGATION MODELING FOR LIGHTWEIGHT IMAGE DEHAZING

Anonymous authors
Paper under double-blind review

ABSTRACT

Many dehazing networks primarily process spatial-domain features, while explicit modelling of haze-related spectral phase and amplitude remains comparatively less explored. We introduce Angular Spectrum Propagation Modeling (ASPM), a lightweight frequency-domain front-end inspired by the angular spectrum method in physical optics. ASPM treats a hazy image as a complex-valued virtual wave field and performs learned, per-channel spectral filtering through an input-conditioned frequency-phase parameter field and phase maps. Designed as an architecture-agnostic module within dehazing, ASPM provides a physics-*inspired* structured prior—not a claim of literal coherent-wave physics—that supports restoration even at minimal parameter scales. In a 100-epoch component ablation, the full model scores 1.30 dB higher than the tested learnable-complex-filter variant; this evidence is specific to the evaluated setup. Our resulting model, ASPMDehazeNet-S, is compact (0.04M parameters) and achieves 26.26 dB PSNR on SOTS-Indoor at 4.2 ms (over 230 FPS) on an NVIDIA RTX 4090. We evaluate on synthetic benchmarks (SOTS-Indoor, SOTS-Outdoor, RS-Haze) and two real-world benchmarks (NH-Haze, Dense-Haze): a larger configuration (ASPMDehazeNet-L, 0.1M) reaches 32.97 dB on SOTS-Indoor. Together with the protocol-qualified comparisons in the paper, these results characterize ASPM under an extreme parameter budget; accuracy and deployment latency are reported separately.

1 INTRODUCTION

Image dehazing aims to recover clear scene radiance from observations degraded by participating media such as haze or fog. This problem is physically rooted in light transport: scattering introduces veiling luminance components while wavelength-dependent attenuation reduces contrast in a depth-dependent manner (Narasimhan & Nayar, 2002; 2003). A large body of classical work addresses this by estimating the transmission map and atmospheric light under physically-motivated statistical priors, including the dark channel prior (He et al., 2009), color attenuation prior (Zhu et al., 2015), non-local haze-line formulation (Berman et al., 2016), and multi-scale fusion strategies (Ancuti & Ancuti, 2013). Learning-based methods subsequently became dominant, progressing from early CNN pipelines (Cai et al., 2016; Ren et al., 2016a) and compact end-to-end networks (Li et al., 2017) to attention-based and Transformer architectures that substantially improve restoration quality (Liu et al., 2019; Qin et al., 2020; Dong et al., 2020; Wu et al., 2021; Song et al., 2023). Higher-capacity Transformer models can achieve strong benchmark accuracy, while compact dehazing networks span a broad accuracy–cost range. This work asks whether an explicit spectral structure is useful at an extreme parameter budget; it does not assume that compact networks without this structure are generally unable to use their parameters effectively.

Concretely, the classical scattering model $I = J \cdot t + A \cdot (1 - t)$ (Narasimhan & Nayar, 2002; 2003) attenuates the clear scene J's high-frequency detail via the transmission map t while the airlight A adds a smooth, low-frequency component, so dehazing is at its core a *frequency-selective restoration* problem (boosting attenuated high frequencies relative to the added low-frequency haze), with Mie/Rayleigh scattering making the rate channel-specific (Goodman, 1968). Spatial-domain networks can also model broad context, and compact methods such as LKD-t demonstrate that sub-million CNNs can achieve high SOTS accuracy. ASPM tests a different design point: FFT supplies global frequency interactions, while an input-conditioned per-channel response is constrained to

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

an angular-spectrum-inspired analytic form with fixed effective RGB constants. In the reported component ablation (Sec. 4.5), the full model scores 1.30 dB above the tested learnable-complex-filter variant. This single result motivates further investigation but does not establish a general ranking between constrained and unconstrained filters.

ASPM also avoids a stored $H \times W \times 3$ distance tensor or free complex mask. Its compact ViT predicts $\mathbf{d}(I_{\text{in}})$ from a global per-channel term and a $16\times$ downsampled residual, making $\mathbf{d}$ a runtime activation rather than $3HW$ stored parameters; a free complex RGB mask at 256×256 would alone require 393,216 real coefficients. The unit-modulus analytic H_c and low-dimensional radial gain provide global spectral correction without full-resolution learned filters. ASPM is therefore *parameter-light rather than computation-free*: FFT/IFFT still costs $\mathcal{O}(HW \log HW)$ and requires separate latency reporting.

The angular spectrum method (ASM) is the canonical framework for how a wave field's spatial frequencies accumulate phase during propagation (Goodman, 1968). We stress up front that we use this as a *mathematical* inductive bias, not a claim that hazy images are literal coherent wave fields (Sec. 5). Inspired by this, we introduce **Angular Spectrum Propagation Modeling (ASPM)**[1], a differentiable lightweight module representing the hazy image as a *virtual* complex-valued wave field and applying learned frequency-phase operators to produce a spectrally corrected representation.

ASPM serves as an architecture-agnostic frequency-domain front-end that can be prepended to different lightweight dehazing backbones without architectural modification. We stress that this generality is *within* the dehazing task: ASPM is designed to complement, not replace, intensity-domain dehazing backbones, and we make no claim of task-agnostic restoration. Within ASPM, a lightweight Vision Transformer predicts a phase map and an input-conditioned frequency-phase parameter field from the hazy input, while a data-driven frequency-domain gain module—derived from empirical analysis of spectral attenuation patterns—corrects spectral amplitude within the same spectral branch. Building on this design, we construct **ASPMDehazeNet-S**, a lightweight end-to-end model with only $\sim$0.04M parameters. Under standard PyTorch (FP32) execution without hardware acceleration, it processes benchmark images in **4.2 ms** on an NVIDIA RTX 4090. A larger configuration (**ASPMDehazeNet-L**, 0.1M) reaches 32.97 dB on SOTS-Indoor. These models probe restoration under an extreme parameter budget; their accuracy is compared in Table 1 and deployment latency in Sec. 4.8.

The main contributions of this work are as follows:

- We propose **ASPM**, a differentiable angular-spectrum-inspired front-end that introduces angular-spectrum-inspired spectral correction as a learned structural prior for image dehazing, complementing intensity-domain processing without simulating atmospheric scattering physics.
- We propose a **data-driven frequency-domain gain module** that empirically characterizes the spectral attenuation ratio between paired hazy and clean images, identifies a parametric gain curve capturing the dominant compensation pattern, and embeds it as a transmission-guided learned module within the ASPM pipeline.
- In the reported 100-epoch component ablation, the full model exceeds the tested learnable-complex-filter variant by 1.30 dB. This result is specific to the evaluated setup; the current experiments do not establish a general advantage over all unconstrained filters or capacity-matched alternatives.

2 Related Work

Pixel-Intensity Dehazing. Traditional methods rely on hand-crafted priors like DCP (He et al., 2009) and CAP (Zhu et al., 2015), physically interpretable but fragile under complex lighting. Learning-based approaches progressed from early CNNs (Cai et al., 2016; Ren et al., 2016b) to attention/Transformer models such as DehazeFormer (Song et al., 2023), Restormer (Zamir et al., 2022), DeHamer (Guo et al., 2022), and MB-TaylorFormer (Qiu et al., 2023), alongside efficient CNNs such as gUNet (Song et al., 2022), DEA-Net (Chen et al., 2024), ConvIR (Cui et al., 2024),

[1]The angular spectrum method shares the abbreviation ASM with the atmospheric scattering model, the standard physical degradation model in the dehazing literature. We use ASPM to avoid ambiguity.

LKD-t (Luo et al., 2023), and SLRNet (Qu et al., 2026). These methods operate primarily on pixel-domain features; LKD-t (0.634M) and SLRNet (0.097M) also show that strong compact dehazing models can lie below one million parameters. ASPM explores a complementary frequency-domain inductive bias at approximately 0.1M parameters.

Lightweight and Physics-Guided Models. Models like AOD-Net (Li et al., 2017), MITNet (Shen et al., 2023), and TCL-Net (Tang & Lou, 2024) use reformulated scattering equations or lightweight attention, while physics-guided designs (Zhang & Patel, 2018; Zheng et al., 2023) embed scattering models into network branches. Real-world/physics-driven methods inject stronger priors: PSD (Chen et al., 2021), D4 (Yang et al., 2022), and RIDCP (Wu et al., 2023) fine-tune with physical losses, decompose transmission by depth, or inject a generative codebook prior, respectively. These primarily operate in the intensity domain or rely on auxiliary/generative networks; ASPM instead studies a structured spectral prior without large attention or generative backbones.

Frequency-Domain Learning. FFC (Chi et al., 2020) and FNO (Li et al., 2021) use Fourier transforms for global dependencies but as spectral *sub-modules* inside larger backbones; Fourmer (Zhou et al., 2023), SFHformer (Jiang et al., 2024), and FSDGN (Yu et al., 2022) (2.73M) learn generic Fourier/amplitude corrections for restoration or dehazing. These approaches generally use free-form or generic Fourier transforms. ASPM instead predicts its input-conditioned spectral phase mask through an angular-spectrum-inspired analytic form with fixed effective wavelength constants, providing a narrower structural bias with a compact predictor.

3 METHOD

ASPM fixes the analytic form of H_c (Eq. 3), its numerical clamp, effective RGB constants, and the amplitude-gain form $A(\rho)$ (Eq. 4), while learning $d_c(I_{\text{in}}; u, v)$, phase maps $\hat{\phi}^c$, and gain coefficients. It therefore predicts a compact dynamic spectral mask rather than storing a full-resolution one; the LCF ablation tests this structural parameterization.

3.1 ANGULAR SPECTRUM PROPAGATION MODULE (ASPM)

A diagram of the ASPM module's internal structure (shared FFT $\rightarrow$ H_c multiplication $\rightarrow$ IFFT pipeline for both the CNN-ASPM and ViT-ASPM variants) is given in Appendix C, Fig. 4, alongside the full forward-pass pseudocode (Algorithm 1).

3.1.1 VIRTUAL WAVE FIELD CONSTRUCTION.

For each color channel $c \in \{R, G, B\}$, we lift the hazy input $I_{\text{in}} \in \mathbb{R}^{H\times W\times 3}$ into a complex-valued virtual wave field:

$$I_{01}^c = \text{clip}\left(\frac{I_{\text{in}}^c + 1}{2}, 0, 1\right), \qquad U_0^c(x, y) = \sqrt{\max(I_{01}^c(x, y), 10^{-6})} \cdot e^{j\hat{\phi}^c(x,y)}, \tag{1}$$

where $\hat{\phi}^c$ is predicted by a lightweight phase network. The phase is learned, not measured; U_0^c is a mathematical inductive bias rather than a physically measured coherent field.

3.1.2 INPUT-CONDITIONED EFFECTIVE PROPAGATION-DISTANCE PARAMETER FIELD.

Rather than storing a global propagation-distance map, ASPM predicts from each input image a three-channel parameter field:

$$\mathbf{d}(I_{\text{in}}) = [d_R, d_G, d_B] \in \mathbb{R}^{H\times W\times 3}. \tag{2}$$

The ViT combines a global per-channel term with a low-resolution residual, then upsamples their sum. Thus $\mathbf{d}$ is an image-dependent activation, not a global matrix with $3HW$ stored parameters. Its entries index the Fourier grid when constructing H_c, so $d_c(I_{\text{in}}; f_x, f_y)$ parameterizes a dynamic frequency-phase mask rather than physical local propagation. The evaluated model fixes the effective RGB constants to $[0.7890,\ 0.5907,\ 0.4241]$; with frequencies in cycles per pixel, they are structural scales rather than optical wavelengths.

3.1.3 DIFFERENTIABLE SPECTRAL PHASE OPERATOR.

The differentiable spectral operator is the mathematical backbone of our module. In the evaluated configuration, U_0^c is reflect-padded by eight pixels on each side before the 2-D FFT and cropped back after the inverse FFT; the optional Tukey window is disabled. The core spectral operator is

$$H_c(I_{\text{in}}; f_x, f_y) = \exp\left(-j \cdot 2\pi d_c(I_{\text{in}}; f_x, f_y)\sqrt{\max\left(0, \tfrac{1}{\lambda_c^2} - f_x^2 - f_y^2\right)}\right), \tag{3}$$

where f_x, f_y are the normalized spatial-frequency coordinates (cycles per pixel), $j = \sqrt{-1}$, d_c is an input-conditioned effective propagation-distance parameter indexed by the Fourier grid, and λ_c is a fixed effective channel constant. For the evaluated frequency grid ($|f_x|, |f_y| \leq 0.5$ cycles/pixel) and fixed constants $\lambda_c \leq 0.789$, the radicand is positive at every bin, so the clamp is inactive. It remains a numerical safeguard for other grid or constant choices; no evanescent-component suppression is claimed. The operator applies a frequency-dependent phase correction. The spectral branch returns $J^c = \left|\mathcal{F}^{-1}\{\mathcal{U}_0^c \cdot H_c \cdot A_\theta^c(\rho)\}\right|$, where $A_\theta^c(\rho)$ is the learned amplitude gain defined in Sec. 3.2 and ρ is the frequency radius. The Mix head subsequently fuses J with the input to predict the final image.

All operations are differentiable, so reconstruction gradients train the phase and $\mathbf{d}$ predictors. Since $|H_c| \equiv 1$, H_c changes phase only; FreqEnhance compensates amplitude (Sec. 3.2).

3.1.4 TWO ASPM VARIANTS AND BACKBONE INTEGRATION.

CNN-ASPM uses a lightweight CNN for phase estimation, parameter-matched to the target backbone ($\sim$0.002M with AOD-Net, $\sim$0.009M with DehazeNet), and can be prepended to any lightweight dehazing backbone without architectural changes, trained end-to-end (Table 4); Sec. 4.4 discusses the accompanying *AOD+AOD* capacity-matched control.

ViT-ASPM replaces the CNN parameter estimator with a patch-based ViT encoder (Dosovitskiy et al., 2021), which jointly predicts the phase map and the dynamic $\mathbf{d}$ field. ViT-ASPM is used only in the full ASPMDehazeNet; Table 3 reports the CNN-ASPM→ViT-ASPM upgrade as a separate ablation step.

3.2 FREQUENCY-DOMAIN ENHANCEMENT FROM SYNTHETIC AND REAL HAZE STATISTICS

From paired hazy/clean images, we compute the radial PSD ratio $R_{\text{PSD}}(\rho) = P_{\text{clean}}(\rho)/P_{\text{hazy}}(\rho)$ and its amplitude-domain counterpart $A_{\text{stat}}(\rho) = \sqrt{R_{\text{PSD}}(\rho)}$. These statistics exhibit a smooth shelf-shaped trend on synthetic and paired real haze (Fig. 1), motivating the amplitude-gain curve

$$A(\rho) = A_0 + (A_\infty - A_0)\left(1 - e^{-(\rho/\rho_0)^q}\right), \tag{4}$$

where A_0, A_∞ are amplitude asymptotes, ρ_0 is the transition frequency, and q controls its steepness. FreqEnhance predicts the per-channel amplitude gain $A_\theta^c(\rho)$ from a haze-strength descriptor. The displayed figures plot power ratios; the corresponding amplitude-domain statistic is their square root. These measurements motivate the functional form, not the learned parameter values: no PSD-ratio value is directly multiplied with a complex spectrum.

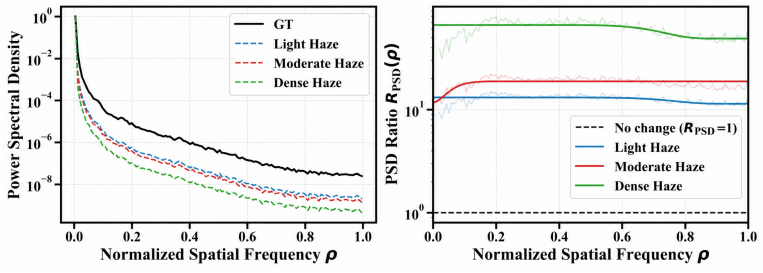
To make the amplitude gain haze-adaptive, we derive a scalar haze-strength descriptor $s = 1 - t$ from a DCP transmission prior refined by a lightweight MLP. DCP supplies only this global scalar, not a pixel-wise dehazing solution. The parameters $\mathbf{a} = [A_0, \Delta A, \rho_0, q]^\top$ are predicted per channel through sigmoid-linear maps of s. The resulting amplitude gain is applied with the angular-spectrum operator:

$$\hat{\mathbf{U}}^c(f_x, f_y) = \hat{\mathbf{U}}_0^c(f_x, f_y) \cdot H_c(f_x, f_y) \cdot A_\theta^c(\rho), \tag{5}$$

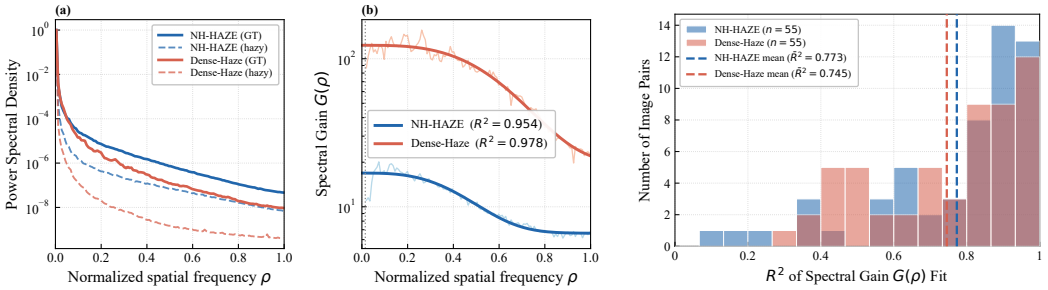
so phase correction (H_c) and learned amplitude compensation (A_θ^c) act in one pass. Full parameter-prediction equations are in Appendix D.

3.3 ASPMDEHAZENET: END-TO-END ARCHITECTURE

Figure 2 shows ASPMDehazeNet-S (0.04M): one ASPM block, a lightweight Mix head, and a light output-refinement head. A ViT predicts phase maps and the input-conditioned $\mathbf{d}$ field; fixed effective RGB constants convert the latter into the dynamic phase mask, and FreqEnhance supplies



(a) Synthetic haze (light/moderate/dense). **(left)** Normalized radial PSD of GT and hazy images. **(right)** Fitted power ratio $R_{\mathrm{PSD}}(\rho)$; the dashed line at R_{PSD}=1 marks no change in power.



(b) Real-haze PSD ratios $R_{\mathrm{PSD}}(\rho)$ and fitted curves.

(c) R^2 distribution (real haze).

Figure 1: **PSD-ratio statistics on synthetic and real haze.** The panels plot power ratios $R_{\mathrm{PSD}} = P_{\mathrm{clean}}/P_{\mathrm{hazy}}$, not amplitude gains. The amplitude-domain statistic is $A_{\mathrm{stat}} = \sqrt{R_{\mathrm{PSD}}}$; its shelf-shaped trend motivates the form in Eq. 4. Neither plotted ratio is applied directly as the complex-spectrum multiplier.

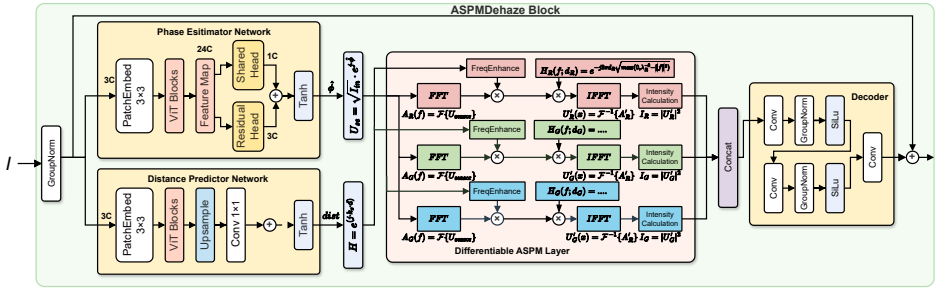


Figure 2: Overview of ASPMDehazeNet. The network utilizes a single ASPMDehaze block for frequency-domain correction and a Mix head for final image reconstruction. Total parameters: ∼0.04M

amplitude compensation. The Mix head fuses the spectral magnitude with the input and predicts the residual. ASPMDehazeNet-L (0.1M) is a larger configuration, not a width-only scaling: relative to S it changes the ViT width/depth and weight sharing, adds a second ASPM layer, and enables the multiscale atmospheric module with a larger output-refinement head. Thus S/L accuracy differences cannot be attributed to parameter count alone. Appendix B summarizes the configuration differences; Appendix C gives the shared forward-pass structure.

3.4 Training Objective

The released ASPM-S/L configurations use a weighted objective rather than an ℓ_1-only loss:

$$\mathcal{L} = \sum_k \lambda_k \mathcal{L}_k, \tag{6}$$

Table 1: Quantitative comparison of dehazing methods on SOTS and RS-Haze benchmarks. **Bold**: best value among methods evaluated under our protocol for the corresponding benchmark; cited ‡ values are excluded. ASPM rows are set apart by horizontal rules; the remaining learned methods are ordered by parameter count (DCP is shown first as a prior-based reference). ‡ marks values cited from the original paper, not reproduced under our training protocol. LKD-t and SLRNet are retrained rows, and only metrics independently re-evaluated in this study are reported.

Method	Params (M)	MACs (G)	Time (ms)	SOTS-Indoor		SOTS-Outdoor		RS-Haze	
				PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DCP (He et al., 2009)	–	–	–	16.62	0.818	19.13	0.815	17.86	0.734
AOD-Net (Li et al., 2017)	0.002	0.114	0.825	19.71	0.821	24.14	0.920	24.90	0.830
MSCNN (Ren et al., 2016b)	0.008	–	–	19.84	0.833	22.06	0.908	22.80	0.823
DehazeNet (Cai et al., 2016)	0.009	0.581	1.996	22.32	0.872	24.75	0.927	23.16	0.816
ASPMDehazeNet-S (Ours)	0.04	0.116	4.237	26.26	0.928	27.32	0.931	30.27	0.933
SLRNet (reproduced) (Qu et al., 2026)	0.097	–	–	–	–	28.44	–	–	–
ASPMDehazeNet-L (Ours)	∼0.1	4.937	39.193	32.97	0.983	33.77	0.980	34.58	0.964
GFN (Ren et al., 2018)	0.499	–	–	22.30	0.880	21.55	0.844	29.24	0.910
LKD-t (Luo et al., 2023)	0.634	6.340	–	36.84	–	36.68	–	–	–
DehazeFormer-T‡ (Song et al., 2023)	0.686	6.658	–	35.15	0.989	33.71	0.982	39.11	0.968

where the terms and weights used in both configurations are listed in Appendix A. The pixelwise ℓ_1 term is the primary reconstruction term; auxiliary image-statistic and ASPM-field regularizers are also enabled. No auxiliary supervision is applied to intermediate restored images. For the cross-backbone experiments, the original backbone training objectives are retained.

4 Experiments

4.1 Experimental Setup

Datasets. RESIDE-ITS contains 13,990 indoor image pairs. The ASPM-S/L configurations split these pairs by clean-image ID into 80% training and 20% validation subsets (11,192/2,798 pairs with the current dataset count). LKD-t and SLRNet were retrained on the same ITS training split. The table reports only re-evaluated baseline metrics: LKD-t SOTS-Indoor/Outdoor PSNR and SLRNet SOTS-Outdoor PSNR; the other cells are unavailable in the current run record. Evaluation uses three synthetic benchmarks: **SOTS-Indoor** (500 images) and **SOTS-Outdoor** (500 images) (Li et al., 2018), and **RS-Haze** (Song et al., 2023) (500 challenging high-resolution pairs).

Implementation. ASPM-S and ASPM-L use their respective released configurations: AdamW with initial learning rate 10^{-4}, weight decay 10^{-4}, five warm-up epochs and cosine decay over 300 epochs; batch sizes are 16 and 8, respectively. Training uses random 256×256 crops and horizontal flips. Both models use the weighted objective in Sec. 3.4 and Appendix A. For the FFT branch, inputs are reflect-padded by 8 pixels and cropped back after inverse transformation; the optional Tukey window is disabled. The reported LKD-t and SLRNet values use the same evaluation implementation for their respective SOTS sets listed above. The shared training split and metric implementation do not imply identical architecture-specific optimization choices.

Metrics. We report PSNR (dB, ↑) and SSIM (↑) on all three benchmarks.

4.2 Comparison with Lightweight Methods

Table 1 compares three synthetic benchmarks. For a controlled training-data comparison, we retrained LKD-t and SLRNet on the same 80% ITS training split as ASPM. We report LKD-t SOTS-Indoor/Outdoor PSNR and SLRNet SOTS-Outdoor PSNR using the same metric implementation; SSIM and other benchmark cells are unavailable from these re-runs and remain blank. Their PSNRs are 36.84/36.68 dB (LKD-t, Indoor/Outdoor) and 28.58 dB (SLRNet, Outdoor), each 0.20 dB below the corresponding value reported in its source paper. The source-paper values are not substituted into our table; the difference is disclosed as a reproduction gap, and metrics not independently re-evaluated are left blank. The shared ITS training data controls the dataset split, not necessarily every architecture-specific optimization choice. LKD-t achieves higher PSNR than ASPMDehazeNet-L

Table 2: Real-world haze evaluation on NH-Haze and Dense-Haze. Upper block: synthetic-to-real generalization from RESIDE training without fine-tuning (our own protocol). Lower block: recent methods, cited from their original papers under their own training settings, for context. **Bold**: best among synthetic-only models in the upper block; lower-block results are cited under different protocols and are not ranked. ‡: cited, not reproduced under our protocol.

		NH-Haze		Dense-Haze	
Method	Params	PSNR	SSIM	PSNR	SSIM
AOD-Net	0.002	14.99	0.580	13.42	0.569
DehazeNet	0.009	15.50	0.555	15.85	0.440
GridDehazeNet	0.958	16.84	0.560	17.70	0.459
ASPM-S (Ours)	0.04	16.91	0.660	16.68	0.466
ASPM-L (Ours)	0.1	20.18	0.776	19.25	0.638
MIFSN‡ (Wang et al., 2025)	13.43	20.33	0.805	17.70	0.637
HazeTrendNet‡ (Wang & Fan, 2025)	6.17	22.14	0.812	17.68	0.655

while using 0.634M parameters, so ASPM does not claim the best accuracy throughout the sub-million-parameter regime. In the near-capacity SOTS-Outdoor comparison (SLRNet: 0.097M; ASPM-L: $\sim$0.1M), ASPM-L scores 5.19 dB higher than our SLRNet re-run. ASPMDehazeNet-S reaches 26.26 dB on SOTS-Indoor, exceeding DehazeNet by 3.94 dB among the early compact baselines reported here. For context, DehazeFormer-T (0.686M, 6.658G MACs) reports 35.15/0.989, 33.71/0.982, and 39.11/0.968 on SOTS-Indoor, SOTS-Outdoor, and RS-Haze, respectively. These cited DehazeFormer-T values follow their authors' protocols and are not a controlled baseline; our comparative claims are restricted to methods evaluated under our protocol.

4.3 Real-World Generalization

We report results on two paired real-haze benchmarks, NH-Haze (Ancuti et al., 2020) and Dense-Haze (Ancuti et al., 2019), using models trained only on synthetic RESIDE data without fine-tuning. In the synthetic-only block of Table 2, ASPMDehazeNet-S improves over DehazeNet on both benchmarks (16.91/0.660 NH-Haze, 16.68/0.466 Dense-Haze), and ASPMDehazeNet-L improves further (20.18/0.776, 19.25/0.638). The cited MIFSN NH-Haze PSNR (20.33) is 0.15 dB higher than ASPM-L, so we do not claim the best result across all reported lightweight methods. NH-Haze and Dense-Haze were collected under controlled conditions rather than as unconstrained in-the-wild data. Moreover, the same two datasets informed the gain-curve design/validation (Sec. 3.2), so these are data-informed evaluations on paired real-haze benchmarks, not independent evidence of unrestricted real-world generalization.

4.4 Cross-Backbone Generalization and Capacity Control

We prepend ASPM to existing backbones (AOD-Net, DehazeNet); full results are in Table 4 (Appendix E). Against *AOD+AOD* ($\sim$0.004M, two stacked AOD-Net modules, capacity-matched to ASPM+AOD-Net), AOD+AOD fails to close the gap—showing ASPM beats this naive capacity-matched alternative, though not that no equal-capacity design could. ASPM also transfers to **DehazeNet** with a consistent 1.10 dB gain.

4.5 Ablation Study

For the ablation study, all models were trained for a reduced number of 100 epochs to analyze relative contributions efficiently.

Exploratory Real-Haze Ablation. Appendix F reports a separate NH-Haze FreqEnhance ablation. Its training strategy differs from the main protocol and the available record does not fully specify the real-style data and test-image separation. We therefore report it as exploratory and do not use it as independent evidence of transfer.

Table 3: Ablation study on SOTS-Indoor. **Group A**: internal network variant comparison (CNN vs. ViT). **Group B**: effectiveness of the physical wave-propagation prior. **Group C**: frequency domain enhancement ablation.

Configuration	PSNR-In	SSIM-In
Group A: Internal network variant		
CNN-ASPM + decoder	22.34	0.876
ViT-ASPM + decoder (= ASPMDehazeNet)	**24.19**	**0.918**
Group B: Spectral Component Ablation		
FFT + FreqEnhance (w/o ASPM)	20.13	0.783
LCF + FreqEnhance (learnable complex filter)	22.89	0.882
Full ASPMDehazeNet	**24.19**	**0.918**
Group C: Frequency Domain Enhancement Ablation		
w/o FreqEnhance	22.11	0.891
w/ Simple NN Freq. Enhancement	22.83	0.892
Full ASPMDehazeNet	**24.19**	**0.918**

Group A: Internal Backbone Variant. Replacing a CNN-based ASPM with our **ViT-based** architecture yields a 1.85 dB improvement (Table 3), suggesting ViT's global receptive field better estimates the input-conditioned phase and frequency-phase parameter fields.

Group B: Spectral Component Ablation. *FFT + FreqEnhance* (amplitude-only, bypassing H_c) scores 20.13 dB, 4.06 dB below the full model (24.19 dB). Replacing H_c with the tested unconstrained *learnable complex filter* (LCF) scores 22.89 dB, 1.30 dB below the full model in this 100-epoch ablation. Because the table does not report a parameter-matching analysis or repeated-run uncertainty, this result is suggestive evidence for the tested structured operator, not a general or capacity-controlled conclusion.

Group C: Frequency Enhancement Design. In this 100-epoch ablation, the tested simple-NN gain variant scores 1.36 dB below the parametric-gain configuration. This single result is suggestive for the tested setup; without parameter matching or repeated runs, it does not establish a general advantage over unconstrained gain predictors.

4.6 Qualitative Analysis

Visual Comparison. Figure 3 shows selected SOTS examples for AOD-Net (Li et al., 2017), DehazeNet (Cai et al., 2016), and our models. On these examples, ASPM outputs show less residual haze than the two displayed compact baselines. This qualitative comparison is limited to the methods and scenes shown; it does not establish visual superiority over the recent methods in Table 1.

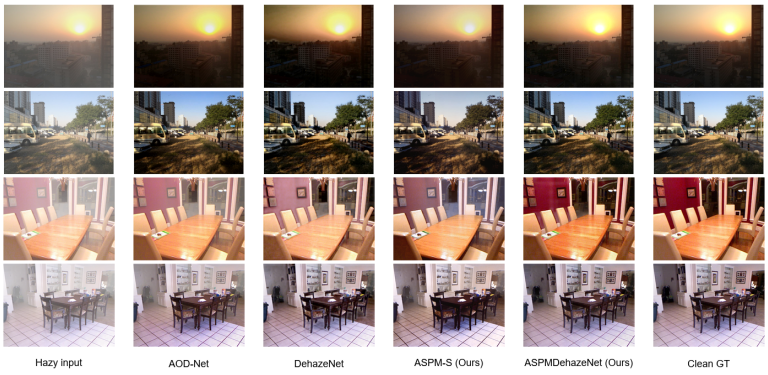


Figure 3: Selected visual examples on RESIDE-SOTS comparing ASPMDehazeNet with the displayed earlier lightweight baselines.

4.7 Operator Diagnostics

ASPM exposes image-conditioned internal fields after training: $\hat{\phi}^c$ and $\mathbf{d}$ can be visualized to diagnose how the predictor modulates the dynamic spectral mask. Under impulse/edge inputs, the operator produces non-local responses characteristic of the implemented Fourier-domain filter. These visualizations are diagnostics of the implemented operator, not evidence that $\mathbf{d}$ is a physical depth map or that the input images are coherent wave fields (Sec. 5; full probes in Appendix G).

4.8 Computational Efficiency

Complexity and Speed. ASPM's FFT/multiply/IFFT steps cost $\mathcal{O}(HW \log HW)$ per channel, comparable to one convolutional layer at typical resolutions. On an RTX 4090 at 620×460 (plain PyTorch FP32), **ASPMDehazeNet-S** runs at **4.24 ms** (>230 FPS)—modestly slower than AOD-Net (0.8 ms) at similar MACs due to FFT/IFFT's global memory access, but well beyond real-time—and **ASPMDehazeNet-L** runs at **39.2 ms** ($\sim$25 FPS). The implementation is GPU-optimized (single-core CPU: $\sim$0.5 FPS), so we scope the real-time claim to GPU execution; GPU-embedded platforms (Jetson, CUDA FFT) are left to future work.

Parameter accounting. The reported ASPMDehazeNet-S checkpoint contains 40,728 trainable parameters: 37,245 in the ASPM ViT, phase/mixing, and gain-prediction block, plus 3,483 in the output-refinement head. The multiscale atmospheric module is disabled in ASPM-S. The effective distance field $\mathbf{d}$ and transfer function H_c are generated at runtime and contribute no stored full-resolution parameters. This accounting explains the small model footprint but not the resolution-dependent FFT latency or activation memory, which are reported separately above.

5 Discussion

Physical status: what ASPM is and is not. As introduced in Sec. 1 and Sec. 3.1, the ASPM-haze connection is *mathematical, not physical*: natural images do not satisfy coherent wave equations, and H_c does not simulate atmospheric scattering. Its effectiveness instead stems from being a narrow inductive bias matching haze's spectral-attenuation statistics, applied to a differentiable field where $\sqrt{I^c_{\text{in}}}$ anchors observable energy and $\hat{\phi}^c$ supplies correction freedom. **Limitations:** (i) Dense uniform haze may reduce the structural gain. (ii) NH-Haze/Dense-Haze transfer is not fully independent, since the same data informed the gain-curve fit (Sec. 3.2), and neither captures unconstrained real-world haze. (iii) AOD+AOD (Sec. 4.4) beats one naive capacity-matched control, not that none could match it; FFT cost also scales as $O(HW \log HW)$, raising latency at >4K.

6 Conclusion

We presented ASPM, a differentiable angular-spectrum-inspired frequency-domain prior for lightweight dehazing. It uses a compact predictor to generate an input-conditioned effective propagation-distance parameter field and constrains the resulting dynamic spectral mask to an analytic angular-spectrum form, rather than storing a free complex-valued filter. In one 100-epoch ablation, the tested structured operator scored 1.30 dB above the tested LCF variant; this comparison is not established as capacity-matched or statistically robust. ASPMDehazeNet-S (0.04M, 26.26 dB) and -L (0.1M, 32.97 dB) provide compact reference points on the reported benchmarks. The LKD-t re-run bounds the claim: ASPM is an approximately 0.1M structured-prior design point, not the accuracy leader across sub-million models.

AI use statement

In this work, we used generative AI tools (a large language model, Claude) to assist with LaTeX formatting and template migration when preparing this manuscript, and to help draft and revise portions of the prose (e.g., clarifying the scope of physical claims and tightening ablation descriptions) based on reviewer feedback from prior submissions of this work. We did not use generative AI tools to design the proposed method, to generate or fabricate any experimental results, or to write the core technical derivations. All AI-assisted text and formatting changes were reviewed by the authors,

who take full responsibility for the final content of this work, including all text, claims, and reported results.

Reproducibility statement

The ASPM-S/L configuration files specify the data split, optimizer, learning-rate schedule, batch size, crop size, training duration, loss weights, and FFT padding used in our runs. Section 4.1 summarizes these settings; Appendix A lists the objective weights, while Eqs. 3 and 4 and Algorithm 1 specify the operator. LKD-t and SLRNet were retrained on the same ITS training split, although architecture-specific optimization details are not fully standardized. The NH-Haze/Dense-Haze gain analysis used these paired datasets, so their results are not independent tests of the gain prior. We evaluate on SOTS-Indoor/Outdoor, RS-Haze, NH-Haze, and Dense-Haze; SOTS-Mix is not used in this revision.

References

Codruta O. Ancuti, Cosmin Ancuti, Mateu Sbert, and Radu Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *IEEE International Conference on Image Processing (ICIP)*, pp. 1014–1018. IEEE, 2019.

Codruta O. Ancuti, Cosmin Ancuti, and Radu Timofte. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *CVPR Workshops*, pp. 444–445, 2020.

Codruta Orniana Ancuti and Cosmin Ancuti. Single image dehazing by multi-scale fusion. *IEEE Transactions on Image Processing*, 22(8):3271–3282, 2013.

Dana Berman, Shai Avidan, et al. Non-local image dehazing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1674–1682, 2016.

Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE transactions on image processing*, 25(11):5187–5198, 2016.

Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. Psd: Principled synthetic-to-real dehazing guided by physical priors. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7176–7185. IEEE, 2021.

Zixuan Chen, Zewei He, and Zhe-Ming Lu. Dea-net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Transactions on Image Processing*, 33: 1002–1015, 2024.

Lu Chi, Borui Jiang, and Yadong Mu. Fast fourier convolution. *Advances in Neural Information Processing Systems*, 33:4479–4488, 2020.

Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Revitalizing convolutional network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 9423–9438, 2024.

Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2157–2167, 2020.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021. URL `https://openreview.net/forum?id=YicbFdNTTy`.

Joseph W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1968.

Chunle Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5802–5810. IEEE, 2022.

Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1956–1963. IEEE, 2009.

Xingyu Jiang, Xiuhui Zhang, Ning Gao, and Yue Deng. When fast fourier transform meets transformer for image restoration. In *Computer Vision - ECCV 2024*, pp. 381–402. Springer Nature Switzerland, 2024.

Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pp. 4770–4778, 2017.

Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE transactions on image processing*, 28 (1):492–505, 2018.

Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=c8P9NQVtmnO.

Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7314–7323, 2019.

Pinjun Luo, Guoqiang Xiao, Xinbo Gao, and Song Wu. Lkd-net: Large kernel convolution network for single image dehazing. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1601–1606, 2023. doi: 10.1109/ICME55011.2023.00276.

Srinivasa G Narasimhan and Shree K Nayar. Vision and the atmosphere. *International journal of computer vision*, 48(3):233–254, 2002.

Srinivasa G. Narasimhan and Shree K. Nayar. Contrast restoration of weather degraded images. *IEEE transactions on pattern analysis and machine intelligence*, 25(6):713–724, 2003.

Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 11908–11915, 2020.

Yuwei Qiu, Kaihao Zhang, Chenxi Wang, Wenhan Luo, Hongdong Li, and Zhi Jin. Mb-taylorformer: Multi-branch efficient transformer expanded by taylor formula for image dehazing. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12756–12767. IEEE, 2023.

Guanheng Qu, Fan Jiang, and Jiangming Liu. Slrnet: Super lightweight residual network for real-time image dehazing. In *Proceedings of the International Conference on Intelligent Computing*, pp. 218–231, 2026. doi: 10.65286/icic.v22i1.71074.

Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *European conference on computer vision*, pp. 154–169. Springer, 2016a.

Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, pp. 154–169, 2016b.

Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang. Gated fusion network for single image dehazing. In *CVPR*, pp. 3253–3261, 2018.

Hao Shen, Zhong-Qiu Zhao, Yulun Zhang, and Zhao Zhang. Mutual information-driven triple interaction network for efficient image dehazing. In *Proceedings of the 31st ACM international conference on multimedia*, pp. 7–16, 2023.

Y Song, Y Zhou, H Qian, and X Du. Rethinking performance gains in image dehazing networks. *arXiv preprint arXiv:2209.11448*, 2022.

Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941, 2023.

Cheng Tang and Wenqi Lou. Tcl-net: A lightweight and efficient dehazing network with frequency-domain fusion and multi-angle attention. In *Proceedings of the Asian Conference on Computer Vision*, pp. 566–582, 2024.

Chen Wang and Yuanyuan Fan. Hazetrendnet: Single-image dehazing via haze-concentration-trend guidance. *Electronics Letters*, 61(1):e70396, 2025. doi: 10.1049/ell2.70396.

Zhibo Wang, Jia Jia, and Jeongik Min. Multiscale implicit frequency selective network for single-image dehazing. *ETRI Journal*, 47(4):695–706, 2025. doi: 10.4218/etrij.2024-0146.

Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10551–10560, 2021.

Rui-Qi Wu, Zheng-Peng Duan, Chun-Le Guo, Zhi Chai, and Chongyi Li. Ridcp: Revitalizing real image dehazing via high-quality codebook priors. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22282–22291. IEEE, 2023.

Yang Yang, Chaoyue Wang, Risheng Liu, Lin Zhang, Xiaojie Guo, and Dacheng Tao. Self-augmented unpaired image dehazing via density and depth decomposition. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2027–2036. IEEE, 2022.

Hu Yu, Naishan Zheng, Man Zhou, Jie Huang, Zeyu Xiao, and Feng Zhao. Frequency and spatial dual guidance for image dehazing. In *European conference on computer vision*, pp. 181–198. Springer, 2022.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728–5739, 2022.

He Zhang and Vishal M Patel. Densely connected pyramid dehazing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3194–3203, 2018.

Yu Zheng, Jiahui Zhan, Shengfeng He, Junyu Dong, and Yong Du. Curricular contrastive regularization for physics-aware single image dehazing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5785–5794, 2023.

M Zhou, J Huang, CL Guo, and C Li. Fourmer: An efficient global modeling paradigm for image restoration. In *International Conference on Machine Learning (ICML)*, volume 202, pp. 42589–42601, 2023.

Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE transactions on image processing*, 24(11):3522–3533, 2015.

A Training Loss Weights

The ASPM-S and ASPM-L configurations use the same weights for the training objective. The perceptual and distance-regularization weights are zero (and therefore inactive); all other listed terms are enabled. The training loss is implemented as the weighted sum in Sec. 3.4:

Term	Weight	Term	Weight
ℓ_1 reconstruction	1.0	MSE	0.2
SSIM	0.05	Color	0.08
Luma	0.05	Exposure	0.03
Edge	0.02	Saturation	0.02
Depth smoothness	5×10^{-4}	Phase TV	2×10^{-4}
Perceptual	0	Distance regularization	0

B ASPM-S AND ASPM-L CONFIGURATION DIFFERENCES

ASPM-S uses embedding width 32, two ViT backbone blocks, one mixing block, one ASPM layer with shared weights, and disables the multiscale atmospheric module. ASPM-L uses width 40, three backbone blocks, two mixing blocks, two ASPM layers with unshared weights, enables multiscale atmospheric correction at scales $\{2, 4, 8\}$, and uses a wider output-refinement head. Both use the same effective RGB constants, weighted loss, 8-pixel reflect padding, and training schedule summarized in Sec. 4.1. This is a configuration comparison, not an isolated parameter-scaling experiment.

C ASPM FORWARD-PASS ALGORITHM

Algorithm 1 details one ASPM block. ASPM-S applies one block; ASPM-L applies two independent blocks in sequence. The outer model then applies its configured atmospheric and output-refinement heads as summarized below.

Algorithm 1 ASPM Forward Pass

Require: Hazy input $I_{\text{in}} \in \mathbb{R}^{H\times W\times 3}$; fixed effective λ_c, predictors $\theta_{\text{phase}}, \theta_d$, $\{a_i, b_i\}_{i=0}^{3}$
Ensure: Spectrally corrected output I_{out}

$\hat{\phi}^c \leftarrow \mathcal{N}_{\text{phase}}(I_{\text{in}}; \theta_{\text{phase}})$ *// per-channel phase map*
$\mathbf{d} \leftarrow \mathcal{N}_d(I_{\text{in}}; \theta_d)$ *// input-conditioned effective distance-parameter field*
$\mathbf{F} \leftarrow \text{Feature}(I_{\text{in}}; \theta_{\text{phase}})$ *// Latent features from phase net*
Compute DCP transmission prior t_{prior} from I_{in}
$t \leftarrow \text{clip}(t_{\text{prior}} + \alpha_t \tanh(f_\theta(\mathbf{F})),\, t_{\min}, t_{\max})$; $s \leftarrow 1 - t$
Predict $(A_0, A_\infty, \rho_0, q)^c$ per channel via the sigmoid-linear amplitude-gain map
for all $c \in \{R, G, B\}$ **do**
    $I_{01}^c \leftarrow \text{clip}((I_{\text{in}}^c + 1)/2, 0, 1)$; $U_0^c \leftarrow \sqrt{\max(I_{01}^c, 10^{-6})} \cdot e^{j\hat{\phi}^c}$ *// normalized virtual wave field*
    Reflect-pad U_0^c by 8 pixels; no window is applied in the reported configurations
    $\hat{\mathcal{U}}_0^c \leftarrow \mathcal{F}\{U_0^c\}$ *// 2-D FFT*
    Compute $H_c(I_{\text{in}}; f_x, f_y)$ using $d_c(I_{\text{in}}; f_x, f_y), \lambda_c$
    Compute $A_\theta^c(\rho)$ for each (f_x, f_y)
    $\hat{\mathcal{U}}_{\text{out}}^c \leftarrow \hat{\mathcal{U}}_0^c \cdot H_c \cdot A_\theta^c(\rho)$ *// phase operator + amplitude gain*
    $J^c \leftarrow \text{Crop}(|\mathcal{F}^{-1}\{\hat{\mathcal{U}}_{\text{out}}^c\}|)$; $J^c \leftarrow 2\,\text{clip}(J^c, 0, 1) - 1$ *// crop padding and map to* $[-1, 1]$ *before Mix*
end for
$I_{\text{out}} \leftarrow \text{clip}(I_{\text{in}} + \alpha\, \text{Mix}(I_{\text{in}}, J))$
If enabled, apply the configured atmospheric-correction module to I_{out}
If enabled, apply the output-refinement module conditioned on I_{in}, then clip to the configured range
return I_{out}

D FREQUENCY-ENHANCEMENT DETAILS

This appendix expands the condensed description of the haze-adaptive amplitude-gain module in the main paper. We derive a scalar haze-strength descriptor $s = 1 - t$ from the input, where a global transmission prior t_{prior} is first estimated using the dark channel prior (DCP) (He et al., 2009) and then refined by a lightweight residual head:

$$\Delta t = \alpha_t \tanh\big(f_\theta(\mathbf{F})\big), \tag{7}$$

where $\mathbf{F}$ are the latent features from the ASPM parameter estimator, $f_\theta(\cdot)$ is a small MLP, and α_t bounds the residual range. The final transmission and haze strength are

$$t = \text{clip}(t_{\text{prior}} + \Delta t,\ t_{\min},\ t_{\max}), \qquad s = 1 - t. \tag{8}$$

The parameters $\mathbf{a} = [A_0, \Delta A, \rho_0, q]^\top$ are predicted per channel through sigmoid-linear mappings of s:

$$\mathbf{a} = \mathbf{a}_{\min} + (\mathbf{a}_{\max} - \mathbf{a}_{\min}) \cdot \sigma(\mathbf{u}s + \mathbf{v}), \tag{9}$$

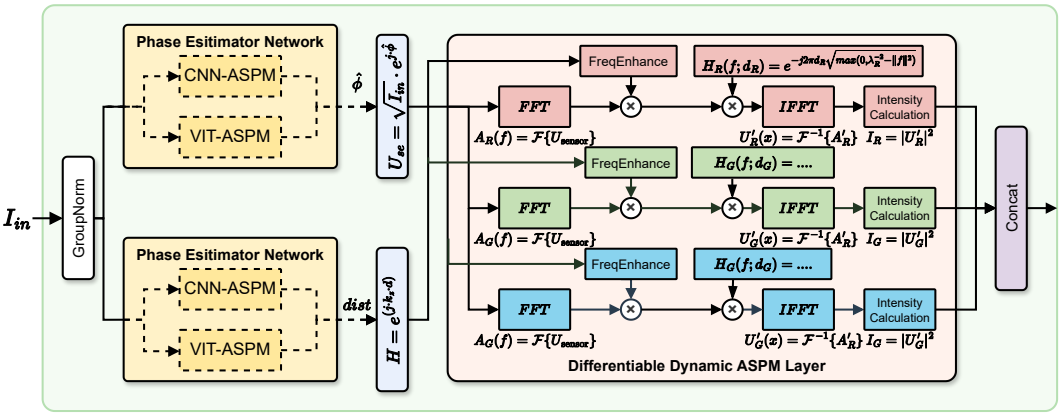


Figure 4: Structure of the ASPM module for a single channel c (referenced from Sec. 3.1). Two internal network variants (CNN-ASPM and ViT-ASPM) share the same propagation pipeline (FFT $\rightarrow H_c$ multiplication $\rightarrow$ IFFT); they differ only in how the phase map $\hat{\phi}^c$ and the input-conditioned effective distance-parameter field $\mathbf{d}$ are predicted. The latter parameterizes the Fourier-domain phase mask and is not a global per-channel matrix.

where $\mathbf{u}, \mathbf{v}$ are learnable weights and $A_\infty = \mathrm{clip}(A_0 + \Delta A, A_{\min}, A_{\max})$. The resulting A_θ^c multiplies the complex spectrum, so it is an amplitude gain, not the PSD ratio used in the descriptive analysis.

E Cross-Backbone Generalization: Full Results

Table 4 reports the full per-configuration results summarized in Sec. 4.4: ASPM prepended to AOD-Net and DehazeNet, compared against the parameter-matched *AOD+AOD* capacity-control baseline. All experiments use CNN-ASPM (without ViT) as a confounding factor.

Table 4: ASPM prepended to multiple lightweight backbones on SOTS. *AOD+AOD* is a parameter-matched capacity baseline.

Configuration	Params (M)	PSNR-In	SSIM-In	PSNR-Out
AOD-Net	0.002	19.71	0.821	24.14
AOD+AOD *(color balancing)*	∼0.004	19.14	0.796	23.65
ASPM+AOD-Net	∼0.004	20.27	0.853	24.83
DehazeNet	0.009	22.32	0.867	24.75
ASPM+DehazeNet	∼0.018	23.42	0.907	25.51

F Real-Haze Ablation

Table 5 reports a completed real-haze ablation on NH-Haze for ASPM-S, comparing FreqEnhance settings under a separate training strategy tailored to this ablation study (adaptive haze-strength estimation with mixed synthetic/real-style pretraining). The real-style data sources and the split protecting NH-Haze evaluation images are not fully documented in the available record. These values are therefore reported as exploratory and are not treated as independent validation or compared directly with Tables 1 and 2.

Within this separate ablation, the full configuration has the highest reported score. Because the real-style training data and the split protecting NH-Haze evaluation images are not fully documented in the available record, this ordering is descriptive only and is not interpreted as independent real-haze validation.

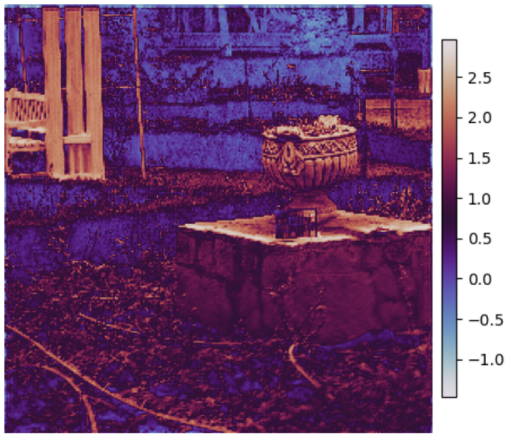
Table 5: Exploratory NH-Haze ablation on ASPM-S using a separate training strategy; real-style training data and evaluation-image separation are not fully documented.

Configuration	PSNR	SSIM
Full ASPMDehazeNet-S (adaptive + mixed pretrain)	**15.53**	**0.609**
w/o FreqEnhance	15.49	0.580
w/ Simple NN Freq. Enhancement	14.69	0.537

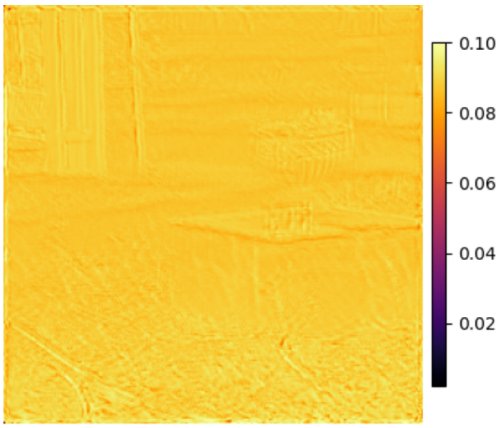
G Operator Diagnostics

This appendix visualizes learned fields and the responses of the implemented operator. These diagnostics do not establish physical wave propagation or causal improvement in dehazing quality.

Analysis of Physical Inductive Bias. To probe whether the constructed operator enables non-local energy redistribution, we evaluate ASPM's response to impulse and edge inputs. As illustrated in Fig. 6, ASPM exhibits wave-optical response patterns, such as PSF spreading and diffraction-like oscillations, that are expected from a nonlocal Fourier-domain filter. The figure is a qualitative response visualization; it does not show that these responses reverse atmospheric scattering or outperform a generic convolutional layer on restoration quality. The hazy images are not treated as coherent wave fields (Sec. 5).



(a) Learned phase maps $\hat{\phi}^c$ for channels.



(b) Input-conditioned effective distance-parameter fields $\mathbf{d}$.

Figure 5: **Input-Conditioned Fields.** (a) Predicted phase maps $\hat{\phi}^c$; (b) predicted effective distance-parameter fields $\mathbf{d}$. In the implementation, $\mathbf{d}$ parameterizes the dynamic Fourier-domain phase mask; its visual spatial texture is not interpreted as a physical local propagation map.

We do not report the earlier HFER, edge-width, or ringing-score comparisons in this revision: the available materials do not specify their calculation and sampling protocols sufficiently for independent interpretation. The response visualization above is retained as a qualitative description of the operator, not as an image-restoration metric.

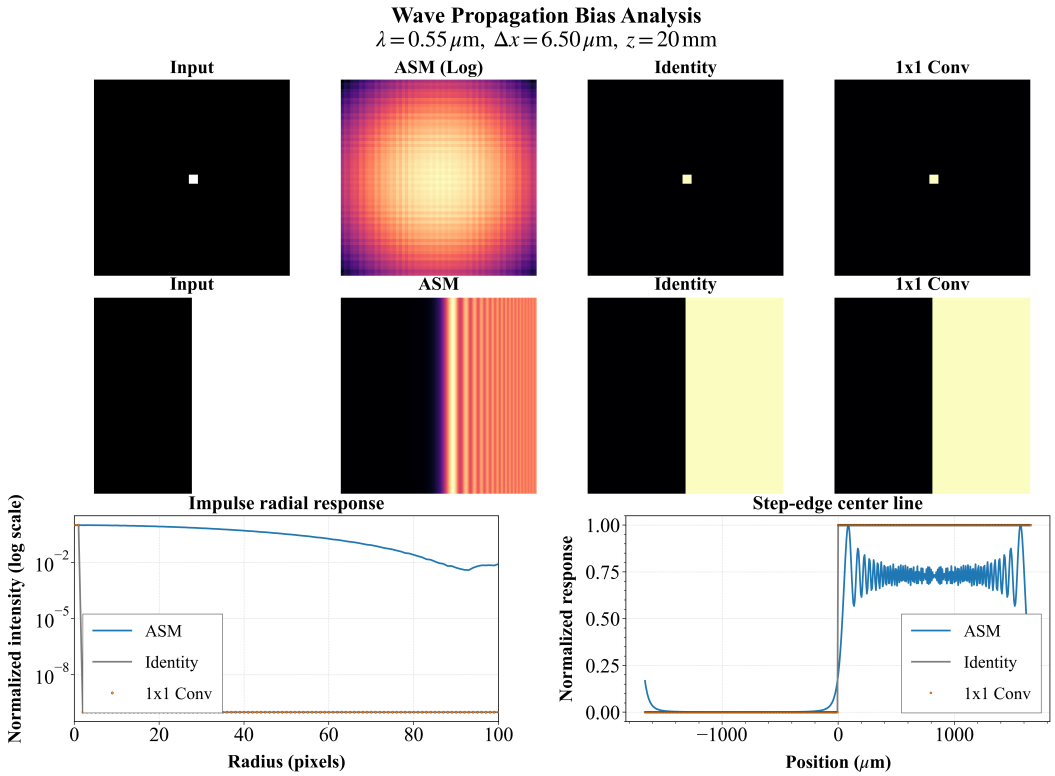


Figure 6: Qualitative response of the implemented ASPM spectral operator to the displayed impulse/edge probes.