This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Supplementary Material: Phase-Pair Holographic Inversion for OCT Digital Phantom Synthesis

Justin Namuk Kim

Department of Biomedical Engineering, Case Western Reserve University
justinkim1217@gmail.com

This supplement reports post-review diagnostic analyses requested by the reviewers. These analyses were not part of the primary all-120 evaluation and should be interpreted as sensitivity checks. Except where stated otherwise, the diagnostic set contains one B-scan from each of 40 acquisition series: the scan labeled "frame 250" in the dataset filename, which is the middle acquisition index among frames 50, 250, and 450. This label is not an iteration count. The analyses characterize the numerical sensitivity of regularized alternating projections (RAP) and the phase-pair encoding; they do not establish unique recovery of the acquired phase or tissue microstructure. Because only displayed magnitudes are available, they cannot identify the precise generating mechanism of visible artifacts or validate their physical localization. MS-SSIM is reported throughout; higher values are better.

Figure S1 reports organizer-hosted rendering. Sections S1–S4 are local-model diagnostics addressing, respectively, phase initialization, lateral regularization, finite-resolution overlap, and amplitude constraints. Throughout, M denotes normalized target field magnitude, while A and L denote the axial and lateral scanner operators.

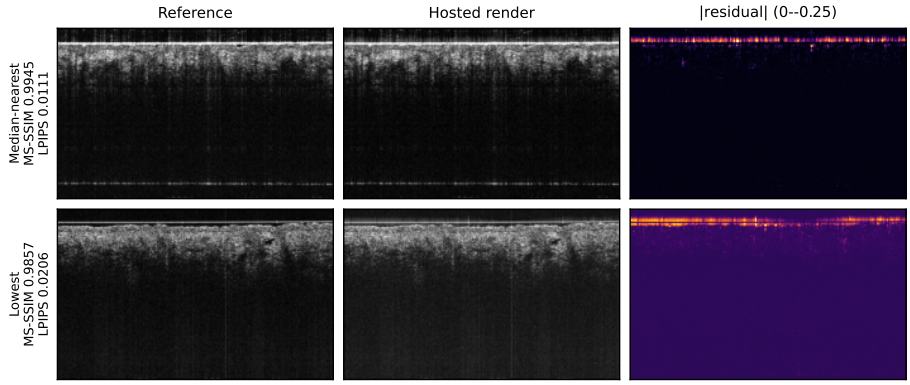


Fig. S1. Organizer-hosted outputs of the final 200-iteration RAP phase-pair configuration for two of the 120 evaluation B-scans, selected by structural MS-SSIM as nearest to the median (top row) and lowest (bottom row). Columns show the displayed reference, hosted render, and absolute display-domain residual; residuals use a common 0–0.25 scale. Row labels report structural MS-SSIM (higher is better) and LPIPS (lower is better). The reproduced deep horizontal band is an axial/spectral artifact rather than reflectors at its displayed depth: matching it demonstrates display-domain fidelity but physically incorrect localization. Its generating mechanism is not identified.

S1 Phase Initialization

The exploratory initialization study used every third series in the sorted 40-scan diagnostic set ($n = 14$). At every checkpoint, the 51 dB target-magnitude inversion, $(\alpha_A, \alpha_L) = (0.02, 0.05)$, and $\beta = 1$ were fixed; recovered coefficients were phase-pair encoded, rendered with the local implementation of the published forward model, and scored against the displayed target.

We compared three deterministic initial phases while keeping target magnitude M unchanged. *Zero phase* sets every initial phase to zero and serves as the standardized empirical baseline. The *axial analytic-signal phase* takes only the angle of $M + \mathrm{i}\mathcal{H}_z[M]$ along each A-scan; it discards that analytic signal's altered magnitude and restores M. The *minimum-phase heuristic* uses a different relation: under the transform convention used here and a true minimum-phase model, phase is obtained from the negative Hilbert transform of log magnitude. It therefore transforms $\log M$, not M, and uses that transform directly as the initial phase:

$$\begin{aligned} \theta_0 &= 0, & F_0^{(0)} &= M, \\ \theta_{\mathrm{H}} &= \arg(M + \mathrm{i}\mathcal{H}_z[M]), & F_0^{(\mathrm{H})} &= M\exp(\mathrm{i}\theta_{\mathrm{H}}), \\ \theta_{\mathrm{MP}} &= -\mathcal{H}_z[\log(\max\{M, 10^{-12}\})], & F_0^{(\mathrm{MP})} &= M\exp(\mathrm{i}\theta_{\mathrm{MP}}). \end{aligned}$$

Here $\mathcal{H}_z$ is the discrete real Hilbert transform along the sampled axial dimension, implemented by the FFT-based analytic-signal transform without added padding. Zero phase is not asserted to be physical, and neither Hilbert construction recovers measured OCT phase: the displayed B-scan is a magnitude image, not an interferometric quadrature or data known to satisfy the minimum-phase assumptions. They are sensitivity-test initializers only.

At checkpoint t, the dimensionless inverse–forward magnitude residual before encoding was

$$\varepsilon_t = (N_z N_x)^{-1} \left\| \, |AC_t L^{\mathsf{T}}| - M \, \right\|_1,$$

where lower values indicate closer agreement with M under the local scanner model.

Table S1. Initialization sensitivity on 14 frame-250 B-scans (every third sorted acquisition series). All runs fix the 51 dB inversion, $(\alpha_A, \alpha_L) = (0.02, 0.05)$, $\beta = 1$, phase-pair encoding, and local rendering. Columns give completed RAP iterations. Top: mean MS-SSIM (higher is better). Bottom: mean dimensionless inverse–forward residual $10^3\varepsilon_t$ before encoding (lower is better). Bold marks the best mean within each metric block and column; 200 iterations is a fixed budget, not a convergence designation.

	Completed RAP iterations				
Initialization	0	10	25	50	200
Mean local-rendered MS-SSIM					
Zero phase (θ_0)	**0.97951**	0.98373	0.99044	0.99315	**0.99429**
Axial analytic-signal phase (θ_{H})	0.97611	**0.98846**	**0.99233**	**0.99328**	0.99419
Minimum-phase heuristic (θ_{MP})	0.96745	0.98341	0.99086	0.99308	0.99405
Mean dimensionless $10^3\varepsilon_t$					
Zero phase (θ_0)	**2.431**	1.605	0.987	0.778	0.628
Axial analytic-signal phase (θ_{H})	2.733	**1.167**	**0.871**	**0.748**	**0.626**
Minimum-phase heuristic (θ_{MP})	4.242	1.467	0.956	0.777	0.630

At 10, 25, and 50 iterations, the analytic-signal start improved mean MS-SSIM over zero phase and was higher on 14/14, 14/14, and 11/14 scans, respectively. At 200 iterations it was higher on 4/14 scans, while the zero-phase mean was larger by 0.00010. No analytic-signal checkpoint through iteration 50 reached the same scan's zero-phase 200-iteration score. The analytic-signal start had lower ε_t than zero phase on all 14 scans at iterations 10, 25, and 50 and on 13/14 at iteration 200. All three mean residuals still decreased from iteration 199 to 200. The score and residual therefore agree that the axial analytic-signal phase accelerates early local-model progress. They do not show physical phase recovery, convergence at 200, or a validated fidelity-preserving reduction of the iteration budget.

S2 Lateral Regularization

The lateral scanner operator has singular values from 0.0126 to 2.869. A matched-regularization spectrum diagnostic used $(\alpha_A, \alpha_L) = (0.02, 0.05)$ before encoding and compared zero phase with 200 RAP iterations. Across the 40 scans, the mean within-scan fraction of squared coefficient energy in the 104 modes with $\sigma < 0.05$ was 36.75% at zero phase and 16.46% after phase selection. This supports greater weak-mode reliance under zero phase.

The historical branch comparison changed α_A, phase selection, and encoding together, so it cannot identify which factor explains the different α_L values. A targeted matched factorial therefore fixed $\alpha_A = 0.02$ and, at each tested α_L, crossed zero versus 200 RAP iterations with single-scatterer versus phase-pair encoding. An exhaustive two-dimensional α_A–α_L grid was not performed.

Both single-scatterer rows benefited from stronger damping, whereas both phase-pair rows favored weaker damping (Table S2). Thus the sensitivity direction follows encoding more strongly than phase selection in this diagnostic. The analytic derivation shows that the single shifted scatterer retains a first-order broadband phase error while the pair cancels it; the sweep is consistent with stronger damping compensating errors in the single-scatterer representation, but does not prove that this term causes the score trend. The historical $\alpha_L = 0.20$ remains an empirical setting, not a demonstrated unique optimum. The final value 0.05 was fixed before the all-120 evaluation and was not retuned from this post-review analysis.

Table S2. Mean local-rendered MS-SSIM on 40 frame-250 B-scans, one per acquisition series (higher is better). At each dimensionless α_L, zero phase versus 200-iteration RAP is crossed with single-scatterer versus phase-pair encoding; the 51 dB inversion, $\alpha_A = 0.02$, $\beta = 1$, and all other settings are fixed. Bold marks the within-row maximum: the single-scatterer rows favor stronger tested damping and the phase-pair rows weaker damping, but the sweep does not identify a unique optimum. † The final $\alpha_L = 0.05$ was fixed before the all-120 evaluation, not selected from this sweep.

	Lateral regularization α_L				
Configuration	0.02	0.05†	0.10	0.20	0.40
Zero phase, single scatterer	0.72934	0.82712	0.90034	0.94258	**0.95631**
RAP, 200 iter., single scatterer	0.91951	0.95509	0.97065	**0.97802**	0.97798
Zero phase, phase pair	**0.98337**	0.98033	0.97733	0.97460	0.97238
RAP, 200 iter., phase pair	**0.99530**	0.99443	0.99248	0.98913	0.98420

S3 Finite-Resolution Overlap

At 6 μm sampling, the approximately 20 μm challenge beam width spans about 3.3 lateral samples. Width conventions differ slightly in the local model: its 10 μm Gaussian radius enters the amplitude kernel as $\exp(-\Delta r^2/w_0^2)$, whose sampled lateral amplitude-response FWHM is 17.15 μm, or 2.86 pitches. The truncated kernel has nonzero weights at five lateral grid positions for an interior A-scan. Two allocated file rows per coefficient therefore give 5.72 nominal row slots across the lateral FWHM, close to the reviewer's estimate of six. The axial amplitude-response FWHM is 13.52 μm, or 2.25 pitches; the rectangular axial–lateral footprint is $2.25 \times 2.86 = 6.44$ coefficient-site areas, corresponding to 12.88 nominal row slots. These are geometric allocation measures, not physical scatterer counts.

One RAP inverse–forward step can be written as

$$\widehat{F}_t = (AA^+_{\alpha_A})F_t(LL^+_{\alpha_L})^{\mathsf{T}}.$$

For either operator $B \in \{A, L\}$, the field-side factor $BB^+_{\alpha_B}$ scales left singular mode r by

$$\rho_{B,r} = \frac{\sigma^2_{B,r}}{\sigma^2_{B,r} + \alpha^2_B},$$

so joint axial–lateral mode (i, j) receives gain $\rho_{A,i}\rho_{L,j}$. This is soft shrinkage: weak scanner modes are attenuated rather than generally excluded. Columns of A are axial point-spread functions formed by the Hann-windowed spectrum and encode coherence-gated depth coupling; L encodes Gaussian beam overlap between adjacent A-scans. The target-magnitude reset is samplewise, but its phase comes from this coupled, filtered field. RAP is therefore neither an explicit Markov/sliding-window prior nor a uniqueness result for magnitude-only data.

Both members of a phase pair share their lateral coordinate and beam weight, and the low-reflectivity linear model superposes their responses with those of neighboring pairs through the same operators. Pair encoding remains a coefficientwise algebraic step after phase selection; it does not jointly fit a tissue-realistic scatterer population. Table S3 quantifies the resulting response widths and coefficient participation.

For output pixel (z, x), let $p_{z'x'} = |A_{zz'}C_{z'x'}L_{xx'}|$ be the magnitude of the modeled complex contribution from coefficient site (z', x'). We computed

$$N_{\text{eff}}(z, x) = \frac{\left(\sum_{z',x'} p_{z'x'}\right)^2}{\sum_{z',x'} p^2_{z'x'}}.$$

Within each scan, pixelwise values were averaged with $|(ACL^{\mathsf{T}})_{zx}|$ as output-field-magnitude weights and then averaged equally across the 40 scans. The mean was 9.91 effective coefficient sites per rendered pixel. This quantity is not doubled into a row count because the two phase-pair weights are generally unequal. The FWHM footprint uses a hard geometric cutoff, whereas the participation ratio includes response tails. Both are model diagnostics, not biological scatterer counts.

The deep horizontal band in the main-paper qualitative figure is treated as an axial/spectral artifact, not as tissue at its displayed depth. These diagnostics cannot distinguish spectral ghosting from saturation-related ringing or another transfer mechanism. Reproducing the band demonstrates display-domain agreement but also exposes the method's physically incorrect localization of artifact energy.

Table S3. Local-model finite-resolution overlap. FWHMs and nominal phase-pair row allocations are deterministic geometry quantities; pair rows are encoding allocations, not scatterer counts. Mean weighted N_{eff} is the output-field-magnitude-weighted participation ratio, computed per scan from 51 dB, 200-iteration RAP coefficient fields at $(\alpha_A, \alpha_L) = (0.02, 0.05)$ and $\beta = 1$, then averaged equally over 40 frame-250 B-scans. All entries quantify modeled spatial mixing, not biological scatterer density.

Quantity	Value
Lateral amplitude FWHM	17.15 μm (2.86 samples)
Axial amplitude FWHM	13.52 μm (2.25 samples)
Axial–lateral FWHM footprint	6.44 coefficient-site areas
Nominal phase-pair rows across lateral FWHM	5.72
Nominal phase-pair rows in axial–lateral footprint	12.88
Mean weighted N_{eff}	9.91 coefficient sites/pixel

S4 Amplitude Constraints

All comparisons used the same coefficient field for each of 40 selected frame-250 scans: 200-iteration RAP with 51 dB target-magnitude inversion, $(\alpha_A, \alpha_L) = (0.02, 0.05)$, and $\beta = 1$, evaluated with the local implementation of the published forward model. Each variant was finally rescaled to a maximum reflection amplitude of 10^{-3}.

The continuous-weight phase-pair encoder deterministically maps 131,072 coefficients to 262,144 pair rows; 37,856 zero-energy rows complete the 300,000-row submission. Under its fixed carrier-equivalent geometry, $d_1 = d_0 - \text{sign}(d_0)\lambda/2$, equal within-pair weights satisfy the zero-first-moment condition only at $|\phi| = \pi$. The equal-within-pair control assigned $a/2$ to both members at d_0 and d_1; amplitudes were therefore equal within a pair but still varied across coefficients.

For each phantom, the N-level control formed N positive log-spaced levels from the amplitude quantile at probability $10^{-4} = 0.0001$ (the 0.01 percentile) to the maximum positive continuous-weight reference amplitude, added zero, assigned each amplitude to the nearest level in linear amplitude, and then applied the common global rescaling. Thus "N levels" below means N positive levels plus zero, and the alphabet was adaptive per phantom rather than fixed across the dataset. The one-global-amplitude control instead approximated the continuous-weight reference amplitudes by integer multiplicities of one nonzero amplitude at d_0 and d_1, with counts adjusted to use exactly 300,000 rows; it did not optimize positions.

Table S4 reports the resulting mean MS-SSIM. These local post-review diagnostics characterize only the tested encodings. The adaptive alphabets and multiplicity construction are not biological priors, and none of the scores is an upper bound on an optimized equal- or discrete-amplitude method. Such a method would require additional scatterers, optimized positions, or joint discrete amplitude–position optimization.

Table S4. Mean local-model MS-SSIM for encoding controls on 40 frame-250 B-scans, one per acquisition series. The reference and all controls share the same 51 dB, 200-iteration RAP coefficient field at $(\alpha_A, \alpha_L) = (0.02, 0.05)$ and $\beta = 1$, followed by a maximum reflection amplitude of 10^{-3} (higher is better; bold marks the best tested mean). An N-level control uses N positive per-phantom log-spaced levels plus zero; equal-within-pair amplitudes still vary across coefficients. The one-global-amplitude control uses integer row multiplicities without position optimization. These tested scores are not upper bounds on optimized amplitude-constrained encodings.

Encoding	Mean MS-SSIM
Continuous-weight phase pair (reference)	**0.99443**
Phase pair, 256 adaptive amplitude levels	0.98828
Phase pair, 64 adaptive amplitude levels	0.93569
Phase pair, 16 adaptive amplitude levels	0.75930
Phase pair, 4 adaptive amplitude levels	0.54734
Phase pair, 2 adaptive amplitude levels	0.37095
Phase pair, equal within-pair amplitudes	0.95497
Single scatterer per coefficient	0.95509
One global amplitude, 300,000 rows	0.45068

Generative-AI assistance. The disclosure in the main manuscript also applies to this supplement's diagnostic code and document preparation.