Supplementary for Polarimetric Full-Wavefield Coherent Lidar

This supplementary document provides complementary material that supports and extends the main paper. Section 1 details the reconstruction framework, including the derivation of the coherent polarimetric measurement model, the estimation of the heterodyne noise variance, and the definition of polarimetric descriptors. Section 2 presents our polarimetric calibration procedure, including the validation of the unitary approximation for the system polarization response, a global-phase-invariant calibration formulated in the Mueller domain, and a performance evaluation of the calibrated model with a varying number of measurements. Section 3 investigates polarization speckle through modeling and experiments, including how speckle statistics vary with field of view and illumination spot size. Section 4 describes the Mueller matrix representation derived from the recovered Jones measurements, supported by physics-based simulations and experimental validation. Section 5 provides additional comparisons against baseline reconstruction algorithms and discusses practical limitations.

1. RECONSTRUCTION

A. Derivation of Statistical Measurement Model

This section provides a complete derivation of the statistical measurement model used in our reconstruction. Starting from the polarimetric coherent measurement model in Eq. (3), we separate the received baseband field into a deterministic coherent return driven by the modulated transmitted waveform and a zero-mean random component that captures scattering and receiver noise. Under the strong-LO approximation and independence assumptions [1], we show that the two-channel measurement follows a complex Gaussian distribution $\mathbf{E}_{\mathrm{rx}}(t) \sim \mathcal{CN}(\boldsymbol{\mu}(t), \boldsymbol{\Sigma}(t))$. We then derive the closed-form expressions for the mean $\boldsymbol{\mu}(t)$ and covariance $\boldsymbol{\Sigma}(t)$ in Eq. (6).

Polarimetric coherent measurement model. Let $\mathbf{E}_{\mathrm{tx}}(t) \in \mathbb{C}^2$ and $\mathbf{E}_{\mathrm{rx}}(t) \in \mathbb{C}^2$ denote the transmitted and received complex Jones vectors in two orthogonal polarization channels. Following Eq. (3) in the main paper, the coherently detected complex baseband signal can be written as

$$\mathbf{E}_{\mathrm{rx}}(t) = \kappa\, \mathbf{E}_{\mathrm{LO}}^{\dagger}(t) \Big(\underbrace{\sum_k \mathbf{J}_k\, \mathbf{E}_{\mathrm{tx}}(t-\tau_k)\, e^{j2\pi\nu_k t}}_{\triangleq\, \mathbf{E}_{\mathrm{coh}}(t)} + \mathbf{E}_{\mathrm{rand}}(t) \Big) + \boldsymbol{\eta}(t), \tag{S1}$$

where $\mathbf{E}_{\mathrm{LO}}(t)$ is the local-oscillator field, κ is a heterodyne mixing constant, $(\cdot)^{\dagger}$ denotes Hermitian transpose, $\mathbf{J}_k \in \mathbb{C}^{2\times 2}$ is the polarimetric impulse response at delay bin k, $\boldsymbol{\eta}(t) \in \mathbb{C}^2$ models receiver noise (electronic + shot noise after coherent detection), τ_k denotes the round-trip propagation delay, and ν_k denotes the corresponding Doppler frequency shift. The coherent sum is defined as $\mathbf{E}_{\mathrm{coh}}(t)$. The term $\mathbf{E}_{\mathrm{rand}}(t) \in \mathbb{C}^2$ aggregates random scattering contributions that are not phase-stable with respect to the modulated coherent return.

Assumptions. We adopt the standard strong-LO approximation: the LO is treated as deterministic (given time t), and the detection is linear in the signal field [1]. We further assume that $\mathbf{E}_{\mathrm{rand}}(t)$ and $\boldsymbol{\eta}(t)$ are independent, zero-mean, complex random vectors:

$$\mathbb{E}[\mathbf{E}_{\mathrm{rand}}(t)] = \mathbf{0}, \qquad \mathbb{E}[\boldsymbol{\eta}(t)] = \mathbf{0}, \qquad \mathbb{E}[\mathbf{E}_{\mathrm{rand}}(t)\, \boldsymbol{\eta}^{\dagger}(t)] = \mathbf{0}. \tag{S2}$$

Mean of $\mathbf{E}_{\mathrm{rx}}(t)$. Taking expectation of Eq. (S1) and using Eq. (S2) yields

$$\begin{aligned} \boldsymbol{\mu}(t) \triangleq \mathbb{E}[\mathbf{E}_{\mathrm{rx}}(t)] &= \kappa\, \mathbf{E}_{\mathrm{LO}}^{\dagger}(t)\, \mathbb{E}[\mathbf{E}_{\mathrm{coh}}(t) + \mathbf{E}_{\mathrm{rand}}(t)] + \mathbb{E}[\boldsymbol{\eta}(t)] \\ &= \kappa\, \mathbf{E}_{\mathrm{LO}}^{\dagger}(t)\, \mathbf{E}_{\mathrm{coh}}(t) = \kappa\, \mathbf{E}_{\mathrm{LO}}^{\dagger}(t) \sum_k \mathbf{J}_k\, \mathbf{E}_{\mathrm{tx}}(t-\tau_k)\, e^{j2\pi\nu_k t}. \end{aligned} \tag{S3}$$

Covariance of $\mathbf{E}_{\mathrm{rx}}(t)$. Define the zero-mean fluctuation $\delta\mathbf{E}_{\mathrm{rx}}(t) = \mathbf{E}_{\mathrm{rx}}(t) - \boldsymbol{\mu}(t)$. From Eq. (S1)–Eq. (S3),

$$\delta\mathbf{E}_{\mathrm{rx}}(t) = \kappa\, \mathbf{E}_{\mathrm{LO}}^{\dagger}(t)\, \mathbf{E}_{\mathrm{rand}}(t) + \boldsymbol{\eta}(t). \tag{S4}$$

Hence the covariance (for circular complex vectors) is

$$\boldsymbol{\Sigma}(t) \triangleq \mathbb{E}\Big[\delta\mathbf{E}_{\mathrm{rx}}(t)\,\delta\mathbf{E}_{\mathrm{rx}}^{\dagger}(t)\Big] = |\kappa|^2\,\|\mathbf{E}_{\mathrm{LO}}(t)\|_2^2\,\underbrace{\mathbb{E}\Big[\mathbf{E}_{\mathrm{rand}}(t)\mathbf{E}_{\mathrm{rand}}^{\dagger}(t)\Big]}_{\triangleq\,\mathbf{C}_{\mathrm{rand}}(t)} + \underbrace{\mathbb{E}\Big[\boldsymbol{\eta}(t)\boldsymbol{\eta}^{\dagger}(t)\Big]}_{\triangleq\,\mathbf{C}_{\eta}(t)}. \tag{S5}$$

The random-scattering term $\mathbf{E}_{\mathrm{rand}}(t)$ models the contribution of unpolarized light. A standard statistical characterization of unpolarized light is that its coherency matrix is proportional to the identity:

$$\mathbf{C}_{\mathrm{rand}}(t) = \lambda_2(t)\,\mathbf{I}_2, \tag{S6}$$

where the scalar $\lambda_2(t) \geq 0$ is the unpolarized power in each polarization channel. This definition is consistent with the eigendecomposition of the coherency matrix in Eq. (4): when the received field is modeled as a coherent rank-1 component plus an isotropic unpolarized term, $\lambda_2(t)$ corresponds to the smaller eigenvalue of $\mathbf{C}_{\mathrm{out}}(t)$ and therefore quantifies the strength of the unpolarized contribution. Similarly, the heterodyne receiver noise is well-approximated as independent and identically distributed across the two detected polarization channels [2]:

$$\mathbf{C}_{\eta}(t) = \sigma_{\mathrm{het}}^2(t)\,\mathbf{I}_2. \tag{S7}$$

Substituting Eq. (S6)–Eq. (S7) into Eq. (S5) gives

$$\boldsymbol{\Sigma}(t) = \Big(\sigma_{\mathrm{het}}^2(t) + |\kappa|^2\,\|\mathbf{E}_{\mathrm{LO}}(t)\|_2^2\,\lambda_2(t)\Big)\mathbf{I}_2 = \big(\sigma_{\mathrm{het}}^2(t) + \beta(t)\big)\mathbf{I}_2, \tag{S8}$$

where we let

$$\beta(t) = |\kappa|^2\,\|\mathbf{E}_{\mathrm{LO}}(t)\|_2^2\,\lambda_2(t). \tag{S9}$$

Complex-Gaussian process. Eq. (S1) is an affine transform of independent circular complex Gaussian terms, and the sum of independent Gaussians is Gaussian; therefore, the measurement follows

$$\mathbf{E}_{\mathrm{rx}}(t) \sim \mathcal{CN}\big(\boldsymbol{\mu}(t), \boldsymbol{\Sigma}(t)\big), \tag{S10}$$

where $\boldsymbol{\mu}(t)$ is given by Eq. (S3), and $\boldsymbol{\Sigma}(t)$ is given by Eq. (S8).

B. Estimation of Heterodyne Noise Variance

To disentangle the scene-dependent incoherent power $\beta(t)$ from the heterodyne noise variance in $\boldsymbol{\Sigma}(t)$, we estimate a global heterodyne noise variance σ_{het}^2 from a depth region where no scene signal returns and the measurements are dominated by receiver noise. Let $\mathcal{P} = \{X, Y\}$ denote the two polarization channels. For each polarization pair $(p, q) \in \mathcal{P} \times \mathcal{P}$, we compute the discrete cross-correlation between the received channel $E_{\mathrm{rx},p}(t)$ and the known transmitted waveform $E_{\mathrm{tx},q}(t)$,

$$g_{pq}[n] \triangleq \sum_t E_{\mathrm{rx},p}(t)\,E_{\mathrm{tx},q}^{*}(t-n), \tag{S11}$$

and extract a length-D segment $\{g_{pq}[n_0+k]\}_{k=0}^{D-1}$ aligned with the delay index k used in reconstruction. Because the effective overlap between the shifted waveforms varies with lag, we normalize the correlation energy by the overlap energy of the transmitted channel at the corresponding lag. For each delay index k, we define

$$\hat{\sigma}_{pq}^2(k) \triangleq \frac{\big|g_{pq}[n_0+k]\big|^2}{\mathcal{E}_{\mathrm{tx},q}(k)}, \qquad \mathcal{E}_{\mathrm{tx},q}(k) \triangleq \sum_{t\in\mathcal{D}(k)} \big|E_{\mathrm{tx},q}(t)\big|^2, \tag{S12}$$

where $\mathcal{D}(k)$ denotes the set of samples that overlap for the lag corresponding to delay k. We then aggregate the four polarization-pair estimates using the median to obtain a per-delay noise curve,

$$\hat{\sigma}^2(k) \triangleq \mathrm{median}_{(p,q)\in\mathcal{P}\times\mathcal{P}}\,\hat{\sigma}_{pq}^2(k). \tag{S13}$$

Finally, we estimate the global heterodyne noise variance as the median of $\hat{\sigma}^2(k)$ over a depth region $\mathcal{K}$ where no scene return is expected:

$$\hat{\sigma}_{\mathrm{het}}^2 \triangleq \mathrm{median}_{k\in\mathcal{K}}\,\hat{\sigma}^2(k). \tag{S14}$$

This procedure yields a stable estimate of the noise floor while remaining insensitive to occasional outliers and residual weak returns.

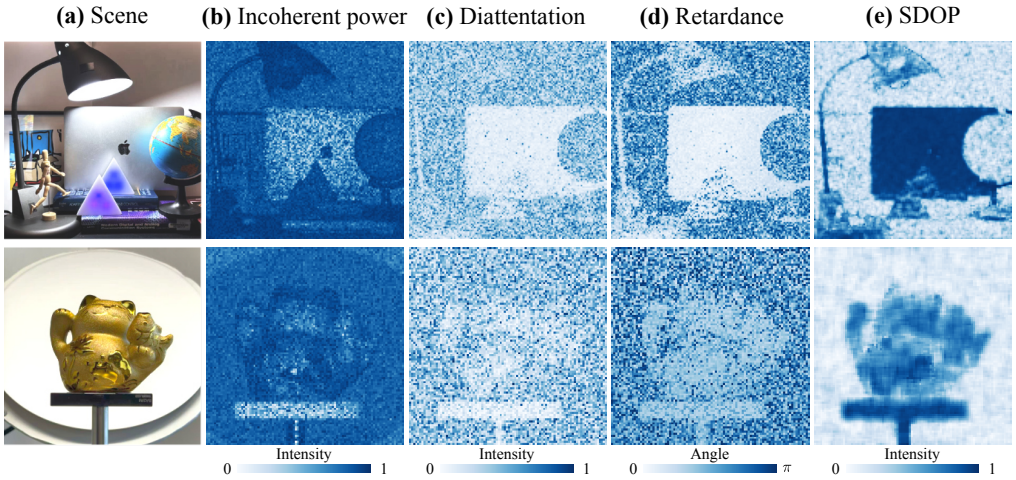

Fig. S1. Per-pixel and spatial polarimetric descriptors. (a) Scene images. **(b)** Incoherent power $\beta(x, y)$, indicating residual returns not explained by a single coherent surface Jones response. **(c–d)** Per-pixel diattenuation and retardance derived from the recovered surface Jones matrices $\mathbf{J}_{\text{surf}}(x, y)$, where diattenuation and retardance encode amplitude imbalance and phase delay, respectively, between orthogonal polarization channels, thereby revealing material-dependent polarimetric contrast. **(e)** Spatial degree of polarization (SDoP) computed from locally averaged Mueller matrices, summarizing the spatial consistency of the recovered polarization; metallic and dielectric surfaces exhibit clearly different SDoP patterns.

C. Per-pixel and Spatial Polarimetric Descriptors

After decoupling the surface response via Eq. (9), we obtain the per-pixel surface-reflection Jones matrix $\mathbf{J}_{\text{surf}}(x, y)$, where (x, y) indexes the image pixel location, together with the incoherent power $\beta(x, y)$ estimated by our statistical model. As illustrated in Fig. S1, these outputs enable both (1) *per-pixel* physically interpretable polarization descriptors and (2) *spatially aggregated* polarimetric statistics that summarize polarization speckle and micro-geometry over a region.

Incoherent power β. In our complex-Gaussian measurement model, the received field is decomposed into a coherent component explained by the modulated waveform and the estimated Jones response, and a residual component that captures multiply scattered light, background illumination, and other unmodeled fluctuations. We represent this residual as an additive, polarization-isotropic covariance term and estimate its power as the incoherent power $\beta(x, y)$. Intuitively, β increases in regions where the return exhibits reduced second-order coherence, and remains low when the received field is well-explained by a single coherent Jones response. We use β both to assess model fit and to measure the incoherent power.

Per-pixel diattenuation and retardance. Given the recovered surface Jones matrix $\mathbf{J}_{\text{surf}}(x, y) \in \mathbb{C}^{2\times 2}$, we extract per-pixel polarimetric descriptors via the polar decomposition

$$\mathbf{J}_{\text{surf}} = \mathbf{U}\mathbf{H}, \tag{S15}$$

where $\mathbf{U}$ is unitary and $\mathbf{H}$ is Hermitian positive semidefinite. This decomposition separates phase-only polarization evolution ($\mathbf{U}$) from amplitude anisotropy ($\mathbf{H}$).

Diattenuation. Let $\sigma_1 \geq \sigma_2 \geq 0$ denote the eigenvalues of $\mathbf{H}$. We define the diattenuation as

$$D \triangleq \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2} \in [0, 1], \tag{S16}$$

which quantifies the relative imbalance between the two principal polarization attenuation axes. This definition is monotonically equivalent to the conventional diattenuation defined from $\mathbf{J}^\dagger\mathbf{J}$ [3, 4].

Retardance. Retardance δ measures the relative phase delay between orthogonal polarization channels, capturing surface-induced birefringence independently of global phase. It is determined from the unitary factor $\mathbf{U}$ through a phase-invariant trace relation,

$$\delta \triangleq 2 \arccos\left(\frac{|\text{tr}(\mathbf{U})|}{2}\right) \in [0, \pi], \tag{S17}$$

which is invariant to any global phase multiplying $\mathbf{U}$.

Spatially aggregated polarimetric statistics SDoP. Beyond per-pixel polarimetric descriptors, the recovered surface response also exhibits spatial fluctuations caused by micro-geometry and coherent interference, leading to polarization speckle. To summarize the local spatial polarimetric structure, we compute the *spatial degree of polarization* (SDoP) from the spatially averaged Mueller matrix. Given a region of interest, we first map each recovered Jones matrix $\mathbf{J}_{\text{surf}}(x, y)$ to its corresponding Mueller matrix $\mathbf{M}(x, y) \in \mathbb{R}^{4\times4}$ using the standard Jones-to-Mueller mapping $\mathbf{M} = \mathcal{M}(\mathbf{J})$ [4]. Importantly, this mapping is invariant to the Jones global phase: $\mathcal{M}(e^{j\phi}\mathbf{J}) = \mathcal{M}(\mathbf{J})$, which makes the following spatial statistic robust to interferometric phase drift. We then form the spatially averaged Mueller matrix

$$\bar{\mathbf{M}} \triangleq \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \mathbf{M}(x, y), \tag{S18}$$

where Ω denotes the ROI and $|\Omega|$ is the number of pixels. From $\bar{\mathbf{M}}$, we compute SDoP as

$$\text{SDoP} \triangleq \sqrt{\frac{\|\bar{\mathbf{M}}\|_F^2 - \bar{M}_{00}^2}{3\,\bar{M}_{00}^2}}, \qquad \bar{M}_{00} > 0, \tag{S19}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\bar{M}_{00}$ is the $(0, 0)$ entry of $\bar{\mathbf{M}}$. This normalization yields SDoP $\in [0, 1]$ in the ideal case: SDoP $= 1$ for a pure polarimetric response, and it decreases as spatial averaging mixes polarization states within the ROI. We note that spatial averaging is essential in this construction. Since each per-pixel Mueller matrix $\mathbf{M}(x, y)$ is derived from a single Jones matrix, it represents a fully polarized response and therefore yields SDoP $= 1$ in the absence of spatial averaging.

Fig. S1 provides a qualitative view of the recovered descriptors discussed above. The incoherent power β aggregates the residual, incoherent return that is not explained by a single coherent surface Jones response, including contributions induced by depolarization and multiple scattering. The remaining polarimetric descriptors characterize complementary aspects of the recovered polarimetric response either per-pixel or over a local spatial region.

In particular, retardance and SDoP discriminate between metallic and dielectric regions. In the first scene, the tape on the lamp is difficult to isolate from the scene image due to the similar albedo, yet it becomes clearly visible in the retardance and SDoP maps. In the second scene, the metallic cat exhibits distinct retardance and SDoP patterns relative to the dielectric background, consistent with the different polarization behavior of conductive versus non-conductive surfaces. Overall, the combination of β and polarimetric descriptors provides richer cues for surface understanding.

D. Stability of the Jones/β Decomposition under Anisotropic Covariance

Our reconstruction model assumes that the residual random term and the receiver noise can be approximated by a scalar covariance, $\Sigma(t) = \big(\sigma_{\text{het}}^2 + \beta(t)\big) I_2$, where σ_{het}^2 denotes the heterodyne noise variance and $\beta(t)$ represents the incoherent power. This model is exact when the residual random term and receiver noise are polarization-isotropic. In the presence of anisotropic depolarized covariance or polarization-dependent noise, this scalar-covariance model becomes an approximation. We therefore evaluate the stability of the Jones/β decomposition under controlled covariance mismatch.

Synthetic experiment setup. To isolate the effect of covariance mismatch, we generate the deterministic coherent return using our depth–Doppler–Jones forward model, which corresponds to the coherent component measured by the modem under the homodyne detection scheme. We then perturb only the residual covariance away from a scalar multiple of the identity. Specifically, we introduce anisotropic depolarized covariance and quantify the perturbation using the covariance anisotropy level $a(\Sigma_{\text{true}}) \in [0, 1]$:

$$a(\Sigma_{\text{true}}) = \frac{\lambda_{\max}(\Sigma_{\text{true}}) - \lambda_{\min}(\Sigma_{\text{true}})}{\lambda_{\max}(\Sigma_{\text{true}}) + \lambda_{\min}(\Sigma_{\text{true}})}, \tag{S20}$$

where Σ_{true} denotes the true 2×2 residual covariance matrix used to generate the synthetic noise, and $\lambda_{\max}(\Sigma_{\text{true}})$ and $\lambda_{\min}(\Sigma_{\text{true}})$ are its largest and smallest eigenvalues, respectively.

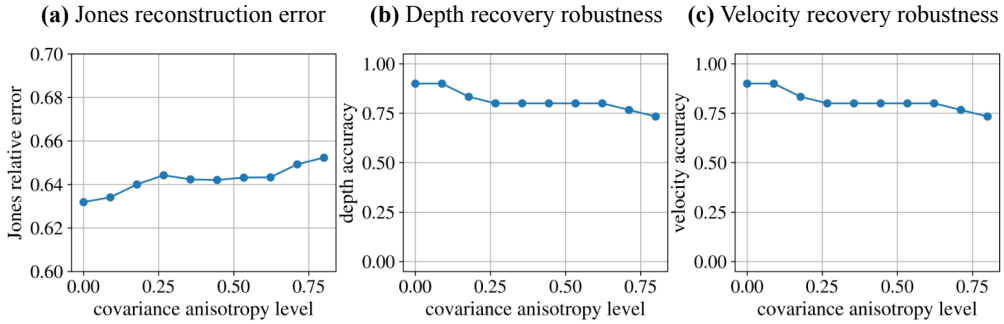

Fig. S2. Stability of the Jones/β decomposition under covariance anisotropy. (a) The Jones relative error varies only within about 3% as the covariance anisotropy level increases from 0 to 0.8, indicating gradual degradation rather than abrupt failure. (b–c) Depth and velocity recovery remain above 70% accuracy even at the strongest tested anisotropy level.

Results. As shown in Fig. S2, the decomposition remains relatively stable as the covariance anisotropy level increases. Across the tested anisotropy range from 0 to 0.8, the Jones relative error varies only within about 3%, while the depth and velocity accuracies decrease gradually but remain above 70% even at the strongest tested anisotropy level. These results indicate graceful degradation rather than breakdown under anisotropic covariance mismatch.

Reason for stability. The observed stability comes from three aspects of the reconstruction. First, as shown in Eq. (8) of the main paper, the optimization objective regularizes the depth–velocity Jones response volume with a sparsity-promoting term, suppressing spurious responses and improving the stability of the inversion.

Second, even when the true residual covariance is anisotropic, its average power can still be represented by the scalar variance term. Specifically, the true covariance can be decomposed into an isotropic component and a zero-trace anisotropic residual:

$$\Sigma_{\text{true}} = \underbrace{\frac{\operatorname{tr}(\Sigma_{\text{true}})}{2} I_2}_{\text{isotropic component}} + \underbrace{\Delta}_{\text{anisotropic residual}}, \qquad \operatorname{tr}(\Delta) = 0. \tag{S21}$$

Our scalar covariance model captures the isotropic component through β, while only the anisotropic residual Δ remains unmodeled. Therefore, covariance anisotropy introduces a gradual model mismatch rather than an abrupt failure, causing the recovered Jones response to degrade smoothly as the anisotropy increases.

Third, depth and velocity estimation depend primarily on the location of the dominant response in the recovered depth–Doppler Jones volume, rather than on the exact value of every Jones matrix element. Specifically, we form the Frobenius-norm energy map

$$E(k,m) = \|J_{k,m}\|_F^2, \qquad E_d(k) = \sum_m E(k,m), \tag{S22}$$

and estimate the depth and velocity indices as

$$k^\star = \arg\max_k E_d(k), \qquad m^\star = \arg\max_m E(k^\star, m). \tag{S23}$$

Therefore, moderate covariance mismatch may bias the recovered Jones matrix values, but depth and velocity remain stable as long as the dominant depth–Doppler peak is preserved.

E. Empirical Stability of Joint Reconstruction

The joint reconstruction of depth, velocity, Jones response, and incoherent power β is a non-convex inverse problem, so trade-offs and local minima may exist in principle. To evaluate its practical stability, we repeat the same reconstruction on a captured measurement from a single scanned point on a dielectric surface using different random initializations. For all trials, the measured waveform, depth–Doppler reconstruction grid, hyperparameters, and optimization settings are fixed; only the random initialization is varied.

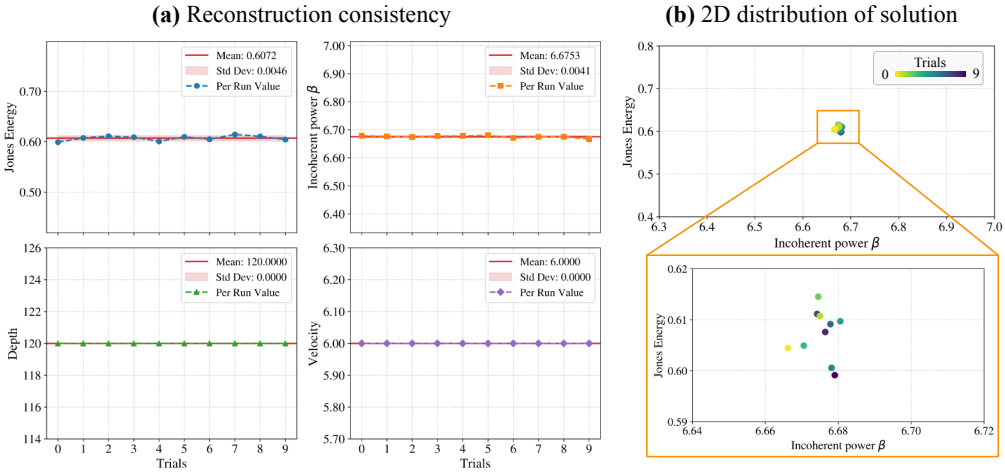

Fig. S3. Empirical stability analysis of the joint reconstruction under random initialization. **(a)** Reconstruction consistency across repeated runs with different random initializations. From top left to bottom right, we show the recovered Jones energy, incoherent power β, depth, and velocity for each trial, together with the mean and standard deviation across runs. All recovered quantities remain highly consistent, indicating reliable convergence. **(b)** Two-dimensional distribution of the recovered solutions in the Jones-energy–β plane. The recovered values from different trials are tightly clustered in a small region, showing that repeated runs converge to highly similar solutions rather than spreading across distinct modes.

As shown in Fig. S3(a), all recovered quantities, including Jones energy, incoherent power β, depth, and velocity, remain highly consistent across random initializations. The mean and standard deviation across trials show only small run-to-run variation, indicating reliable convergence under this measurement setting. Fig. S3(b) further visualizes the recovered solutions in the Jones-energy–β plane. The solutions from different trials are tightly clustered in a small region, suggesting that the optimizer converges to highly similar solutions rather than spreading across distinct local modes. This analysis provides empirical evidence of stability for the tested captured measurement, but it is not a theoretical guarantee that local minima or stronger ambiguities cannot occur in all scenes, SNR levels, or scattering conditions.

F. Implementation Details

Optimization settings. We use Adam with learning rate 10^{-2} and weight decay 10^{-4}. Each per-pixel solve runs for 100 iterations using all available time samples. We clip the gradient norm to 1.0 for stability. The group-sparsity weight is set to $\lambda_{\text{reg}} = 0.6$ unless otherwise stated. Jones volumes are initialized with small complex Gaussian noise (standard deviation 10^{-5} for both real and imaginary parts).

FFT acceleration. To avoid recomputing Doppler-modulated transmitted waveforms during optimization, we precompute the FFT of the transmitted signal for all Doppler bins once per pixel. We use an FFT length $N_{\text{fft}} = T + P - 1$ to implement linear convolution, where T is the number of time samples and P is the number of delay bins. This reduces the per-iteration cost to a small number of FFTs and batched complex matrix multiplications.

2. POLARIMETRIC CALIBRATION

A. Experimental Validation of the Unitary Approximation for $\mathbf{J}_{\text{tx}}$ and $\mathbf{J}_{\text{rx}}$

The Jones matrices $\mathbf{J}_{\text{tx}}$ and $\mathbf{J}_{\text{rx}}$ are complex 2×2 matrices. Directly optimizing the entries of these matrices introduces redundant degrees of freedom and makes the joint reconstruction ill-posed. We therefore impose a physically motivated constraint in our optimization because polarization evolution in single-mode fibers and reciprocal passive optics is dominated by birefringent mixing rather than diattenuation. Such transformations can be represented by *unitary* Jones matrices up to a global phase. In this section, we provide an empirical justification that the reconstructed system

paths are well-approximated by this unitary constraint, which motivates the parameterization in Eq. (13) of the main text.

Preliminaries. With a rotatable linear polarizer $\mathbf{J}_{\text{pol}}(\theta)$ inserted in the calibration path (Fig. 4(a) in the main paper), the recovered round-trip Jones matrix is

$$\mathbf{J}^{\star}(\theta) = \gamma\, \mathbf{J}_{\text{rx}}\, \mathbf{J}_{\text{pol}}(\theta)\, \mathbf{J}_{\text{surf}}\, \mathbf{J}_{\text{pol}}(\theta)\, \mathbf{J}_{\text{tx}}, \tag{S24}$$

where $\gamma \in \mathbb{R}_{+}$ absorbs polarization-independent system gain.

Lemma 1 (rank-1 structure of Jones matrix makes the output polarization state input-independent). Let the inserted linear polarizer be $\mathbf{J}_{\text{pol}}(\theta) = \mathbf{d}(\theta)\mathbf{d}(\theta)^{\mathsf{T}}$ with $\mathbf{d}(\theta) = [\cos\theta,\ \sin\theta]^{\mathsf{T}}$. Then the effect of the forward and backward propagation through the polarizer collapses to a single, scaled polarizer:

$$\mathbf{J}_{\text{pol}}(\theta)\, \mathbf{J}_{\text{surf}}\, \mathbf{J}_{\text{pol}}(\theta) = \mathbf{d}(\theta)\big(\mathbf{d}(\theta)^{\mathsf{T}} \mathbf{J}_{\text{surf}} \mathbf{d}(\theta)\big)\mathbf{d}(\theta)^{\mathsf{T}} = s(\theta)\, \mathbf{J}_{\text{pol}}(\theta), \tag{S25}$$

where $s(\theta) \in \mathbb{C}$ is a scalar. Substituting Eq. (S25) into Eq. (S24) gives

$$\mathbf{J}^{\star}(\theta) = \gamma\, s(\theta)\, \mathbf{J}_{\text{rx}}\, \mathbf{J}_{\text{pol}}(\theta)\, \mathbf{J}_{\text{tx}}. \tag{S26}$$

Because $\mathbf{J}_{\text{pol}}(\theta)$ is an outer product, it has rank one for all θ. Since matrix multiplication cannot increase rank, $\operatorname{rank}\big(\mathbf{J}^{\star}(\theta)\big) \le 1$ and $\mathbf{J}^{\star}(\theta)$ is rank-1 whenever $s(\theta) \neq 0$.

This rank-1 structure implies that the output polarization state does not depend on the incident polarization. Indeed, for any incident Jones vector $\mathbf{E}_{\text{in}}$,

$$\begin{aligned} \mathbf{E}_{\text{out}}(\theta) &= \mathbf{J}^{\star}(\theta)\mathbf{E}_{\text{in}} = \gamma s(\theta)\, \mathbf{J}_{\text{rx}} \mathbf{d}(\theta)\, \mathbf{d}(\theta)^{\mathsf{T}} \mathbf{J}_{\text{tx}} \mathbf{E}_{\text{in}} \\ &= a(\theta, \mathbf{E}_{\text{in}})\, \mathbf{J}_{\text{rx}} \mathbf{d}(\theta), \end{aligned} \tag{S27}$$

where $a(\theta, \mathbf{E}_{\text{in}}) \triangleq \gamma s(\theta)\, \mathbf{d}(\theta)^{\mathsf{T}} \mathbf{J}_{\text{tx}} \mathbf{E}_{\text{in}}$ is a scalar. Therefore, after normalization, $\mathbf{E}_{\text{out}}(\theta)$ is fully determined by $\mathbf{J}_{\text{rx}}\mathbf{d}(\theta)$. This establishes that the Stokes trajectory observed on the Poincaré sphere under a polarizer sweep reflects an intrinsic property of $\mathbf{J}_{\text{rx}}$ (rather than a particular choice of $\mathbf{E}_{\text{in}}$), providing a necessary condition for the subsequent unitary-model validation.

Lemma 2 (rigid rotation on the Poincaré sphere $\Leftrightarrow$ unitary Jones operator). For a fully polarized field, define the normalized Stokes vector $\mathbf{s}(\mathbf{E}) \in \mathbb{R}^3$ on the Poincaré sphere. A Jones transform $\mathbf{E}' = \mathbf{U}\mathbf{E}$ induces a Mueller transform on Stokes vectors. If $\mathbf{U}$ is unitary (up to a global phase), then the induced Mueller matrix is a pure retarder and acts as a rigid rotation $\mathbf{s}(\mathbf{E}') = \mathbf{R}\,\mathbf{s}(\mathbf{E})$ with $\mathbf{R} \in \mathrm{SO}(3)$. Conversely, a polarization transform that preserves the degree of polarization and is explainable by a single $\mathbf{R} \in \mathrm{SO}(3)$ corresponds to a unitary Jones operator. Therefore, observing a rigid rotation between input and output Stokes trajectories is evidence that the underlying Jones operator is unitary.

Implication for $\mathbf{J}_{\text{rx}}$ and $\mathbf{J}_{\text{tx}}$. By Lemma 1, sweeping θ produces a set of output states proportional to $\mathbf{J}_{\text{rx}}\mathbf{d}(\theta)$ whose normalization is independent of the illumination. For an ideal input sweep, a rotating linear polarizer traces an equatorial great circle on the Poincaré sphere (i.e., $S_3 = 0$ with varying azimuth). As shown in Fig. 4(b) of the main paper, the recovered output Stokes samples remain well confined to a single plane; fitting this plane reveals that the locus is still a circle rather than a distorted curve. This geometric invariance indicates a rigid transformation on the sphere, consistent with a pure rotation $\mathbf{R} \in \mathrm{SO}(3)$ as characterized in Lemma 2. We therefore model $\mathbf{J}_{\text{rx}}$ as unitary up to a scalar gain absorbed into γ. The same argument applies to $\mathbf{J}_{\text{tx}}$ under the same physical conditions (i.e., given a single-mode fiber and reciprocal passive optics), and in practice, our joint calibration fits measurements from the transmitted and received paths using this unitary constraint with low residual error.

Resulting parameterization. Motivated by the above, we parameterize $\mathbf{J}_{\text{tx}}$ and $\mathbf{J}_{\text{rx}}$ using the unitary form in Eq. (13) of the main paper, which enforces $\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}$ by construction and reduces the number of calibration variables to three real parameters for each transmitted and received path.

Quantification of residual polarization-dependent loss. To further quantify the validity of the polarization-loss-free approximation, we measure the polarization-dependent loss (PDL) of the cascaded optical path using the isolated internal round-trip reflection as a reference response, as shown in Fig. S4. The optical modem transmits a dual-polarization random-modulated sequence containing 65,600 complex Jones samples, which spans a broad set of input polarization

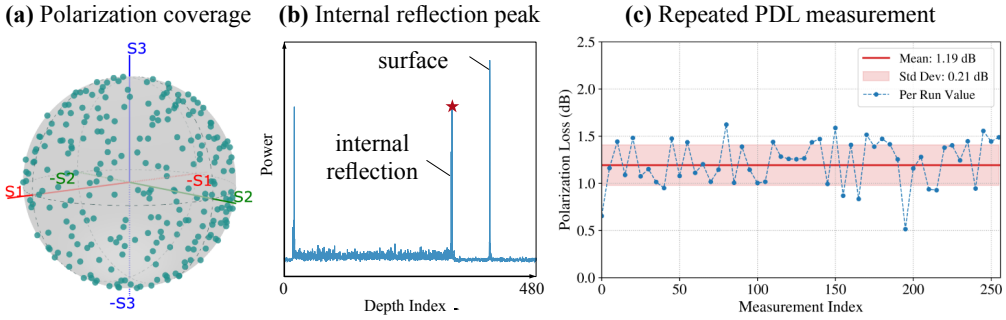

Fig. S4. Experimental quantification of the residual polarization-dependent loss of the cascaded optical path. (a) The transmitted Jones vector sequence spans a broad set of input polarization states on the Poincaré sphere, providing diverse excitation for probing the system polarization response. (b) Correlating the received waveform with the known transmitted waveform yields a time-resolved response, where the dominant internal round-trip echo is clearly separated from the later surface return in depth. (c) Effective PDL measured from the recovered internal-echo Jones matrix over 256 repeated measurements. The mean PDL is approximately 1.19 dB, with a standard deviation of 0.21 dB, indicating that the cascaded path is not perfectly polarization-loss-free but exhibits only modest residual diattenuation.

states. We correlate the received waveform with the known transmitted waveform to isolate the dominant internal echo in depth, reconstruct the corresponding round-trip Jones matrix, and compute its PDL from the singular-value ratio. Across 256 repeated measurements under the same configuration, the measured residual PDL of the cascaded optical path is approximately 1.19 ± 0.21 dB, which is small relative to the dominant polarization mixing and is consistent with the unitary approximation used in our calibration model.

B. Global-Phase-Invariant Calibration in Mueller Domain

Global phase ambiguity in repeated coherent measurements. In our coherent receiver, each Jones response $\mathbf{J}^\star(\theta_n)$ is estimated from a measurement acquired at a different time. Due to interferometric phase drift between the local oscillator and the signal (e.g., path-length fluctuations, laser frequency noise, and LO–signal phase wander), the recovered Jones response is typically corrupted by an unknown global phase:

$$\mathbf{J}^\star(\theta_n) = e^{j\phi_n}\,\widetilde{\mathbf{J}}(\theta_n), \tag{S28}$$

where $\phi_n \in \mathbb{R}$ varies across measurements and $\widetilde{\mathbf{J}}(\theta_n)$ denotes the underlying physical response. This ambiguity keeps the relative amplitudes and relative phase between the two orthogonal polarization components unchanged. However, if we define the calibration loss directly in the Jones domain, the unknown ϕ_n induces large, artificial discrepancies between $\mathbf{J}(\theta_n;\boldsymbol{\Theta})$ and $\mathbf{J}^\star(\theta_n)$ and can prevent convergence.

Figure S5 empirically visualizes this ambiguity under repeated measurements with identical geometry. As shown in Fig. S5(a), each complex entry of the reconstructed Jones matrix traces an approximately circular locus in the complex plane across 256 trials, consistent with a multiplicative global phase factor $e^{j\phi_n}$ that rotates all elements by the same angle. In contrast, Fig. S5(b) shows that phase-invariant polarimetric quantities extracted from $\mathbf{J}^\star$—including the singular values (hence diattenuation), as well as retardance and orientation—remain tightly clustered across trials, indicating that the underlying physical polarization response $\tilde{\mathbf{J}}(\theta_n)$ is stable and only the global phase varies.

Mueller mapping removes global phase by construction. Polarization in our measurements is summarized by second-order field moments captured in the coherency matrix. Since a global phase factor does not change these moments, the Mueller representation induced by a Jones matrix is inherently invariant to the global phase of $\mathbf{J}$. We therefore use Mueller-domain objectives for calibration to avoid phase ambiguity. Let $\mathbf{C}_{\text{in}} = \mathbb{E}[\mathbf{E}_{\text{in}}\mathbf{E}_{\text{in}}^\dagger]$ be the input coherency matrix. A Jones transformation produces

$$\mathbf{C}_{\text{out}} = \mathbb{E}[\mathbf{E}_{\text{out}}\mathbf{E}_{\text{out}}^\dagger] = \mathbf{J}\,\mathbf{C}_{\text{in}}\,\mathbf{J}^\dagger. \tag{S29}$$

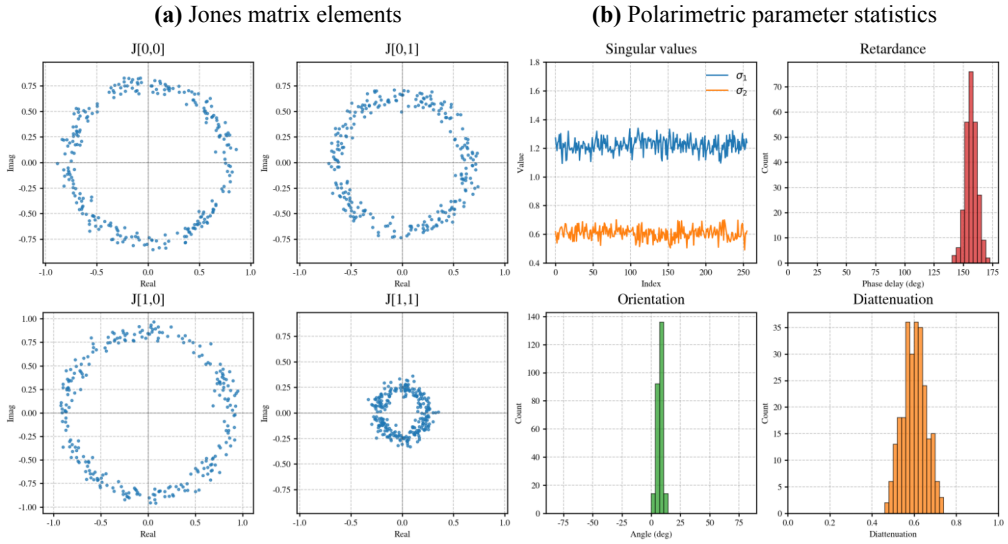

Fig. S5. Global phase drift across repeated coherent reconstructions. We repeat the same acquisition 256 times at a single scan point on a dielectric surface under identical geometry, and reconstruct the corresponding surface Jones matrices $\mathbf{J}$ for each repeat. **(a)** Each complex Jones element traces an approximately circular locus in the complex plane, indicating that trial-to-trial variation is dominated by an unknown global phase factor. **(b)** In contrast, phase-invariant polarimetric descriptors (Jones matrix singular values, retardance, orientation, and diattenuation) remain tightly clustered across repetitions, showing that the underlying polarization response is stable.

If the Jones matrix is multiplied by a global phase $e^{j\phi}$, then

$$(e^{j\phi}\mathbf{J})\,\mathbf{C}_{\text{in}}\,(e^{j\phi}\mathbf{J})^{\dagger} = e^{j\phi}e^{-j\phi}\,\mathbf{J}\,\mathbf{C}_{\text{in}}\,\mathbf{J}^{\dagger} = \mathbf{J}\,\mathbf{C}_{\text{in}}\,\mathbf{J}^{\dagger}, \tag{S30}$$

so the output coherency matrix is invariant to the global phase.

The Stokes vector provides a real-valued parametrization of the coherency matrix, collecting its projections onto the Pauli basis and encoding the same second-order polarization information. The Mueller matrix $\mathcal{M}(\mathbf{J})$ is the corresponding real 4×4 linear operator that maps input Stokes vectors to output Stokes vectors: $\mathbf{s}_{\text{out}} = \mathcal{M}(\mathbf{J})\mathbf{s}_{\text{in}}$. Equivalently, the standard Jones-to-Mueller construction can be written as

$$\mathcal{M}(\mathbf{J}) = \mathbf{A}\big(\mathbf{J} \otimes \mathbf{J}^{*}\big)\mathbf{A}^{-1}, \tag{S31}$$

where $\otimes$ denotes the Kronecker product, $(\cdot)^{*}$ denotes element-wise complex conjugation and $\mathbf{A}$ is a fixed change-of-basis matrix between the Pauli basis and the Stokes basis:

$$A = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix}. \tag{S32}$$

Under a global phase, $\mathbf{J}' = e^{j\phi}\mathbf{J}$ yields

$$\mathbf{J}' \otimes \mathbf{J}'^{*} = (e^{j\phi}\mathbf{J}) \otimes (e^{-j\phi}\mathbf{J}^{*}) = \mathbf{J} \otimes \mathbf{J}^{*}, \quad \Rightarrow \quad \mathcal{M}(\mathbf{J}') = \mathcal{M}(\mathbf{J}), \tag{S33}$$

which matches the coherency argument above.

Implication for the calibration objective. Because $\mathcal{M}(\mathbf{J})$ is invariant to the measurement-dependent global phase factors in Eq. (S28), defining the calibration loss in the Mueller domain eliminates this ambiguity. This justifies the objective in Eq. (14), where we minimize discrepancies between predicted and measured responses in Mueller space, thereby fitting only physically observable polarimetric behavior while remaining agnostic to interferometric phase drift.

Table S1. Evaluation of calibration angle number on reconstruction accuracy. We reduce the number of linear-polarizer calibration angles (columns 1–2) while keeping the same held-out polarizer test angles (column 3), and report reconstruction error as Mueller-matrix MSE and Stokes-vector MSE (columns 4–5).

Number of Calibration Angle	Calibration Angles (deg)	Test Angles (deg)	Mueller MSE	Stokes MSE (S1,S2,S3)
19	[0–180]/10	[5–185]/10	0.0040	0.0044
10	[0–180]/20	[5–185]/10	0.0043	0.0048
7	[0–70]/10	[5–185]/10	0.0049	0.0058
4	[20,60,130,170]	[5–185]/10	0.0050	0.0063
3	[0,80,160]	[5–185]/10	0.0076	0.0142
2	[0,80]	[5–185]/10	0.0075	0.0112

Note: Stokes metrics are computed with input polarization $E_{\text{in}} = [1, j]$. The last two rows with gray background indicate settings that lead to failed or unstable reconstructions.

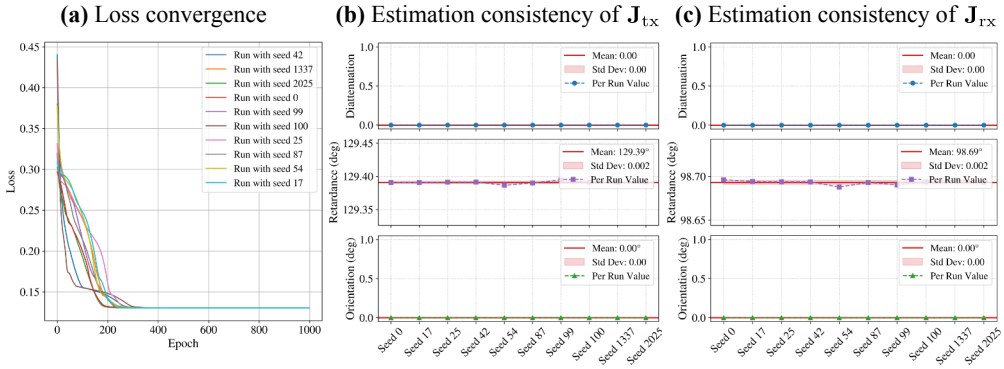

Fig. S6. Repeatability of polarimetric calibration under random initialization. We evaluate the reproducibility of the calibration optimization by running the same measurement set with different random seeds. **(a)** Loss convergence across trials: all runs converge to the same optimum. **(b–c)** Estimation consistency of $\mathbf{J}_{\text{tx}}$ and $\mathbf{J}_{\text{rx}}$: we compute diattenuation, retardance, and orientation from the estimated Jones matrices and observe nearly identical values across runs, with extremely small standard deviations.

C. Evaluation of Calibration Model Optimization

Evaluation of calibration efficiency. We progressively reduce the number of linear-polarizer (LP) calibration angles from 19 down to 2 and measure the impact on reconstruction accuracy. For each setting, we keep the test protocol fixed: we reconstruct $\mathcal{M}(\mathbf{J})$ and the corresponding Stokes vectors on the same held-out set of 19 LP angles uniformly spanning 5°–185° (10° increments), and compare them to analytical ground truth from the ideal LP model (with a fixed input $\mathbf{E}_{\text{in}} = [1, j]^{\mathsf{T}}$). Table S1 reports the mean-squared error (MSE) in both the Mueller and Stokes domains. Reducing the calibration set from 19 to 4 angles yields only a marginal increase in error, whereas using 3 or 2 angles causes a noticeable rise in Stokes error, suggesting that four calibration angles are sufficient for near-optimal performance in our setup.

Evaluation of calibration repeatability. To evaluate the reproducibility of our polarimetric calibration, we rerun the optimization 10 times on the same measurement set with different random initializations. As shown in Fig. S6(a), all trials converge to the same final loss value, indicating a stable optimum. Moreover, the estimated transmit- and receive-path Jones matrices are highly consistent across runs: Fig. S6(b–c) reports the diattenuation, retardance, and orientation extracted from $\mathbf{J}_{\text{tx}}$ and $\mathbf{J}_{\text{rx}}$, respectively, showing nearly identical per-run values with extremely small standard deviations. Together, these results demonstrate that our calibration procedure is repeatable and insensitive to random initialization.

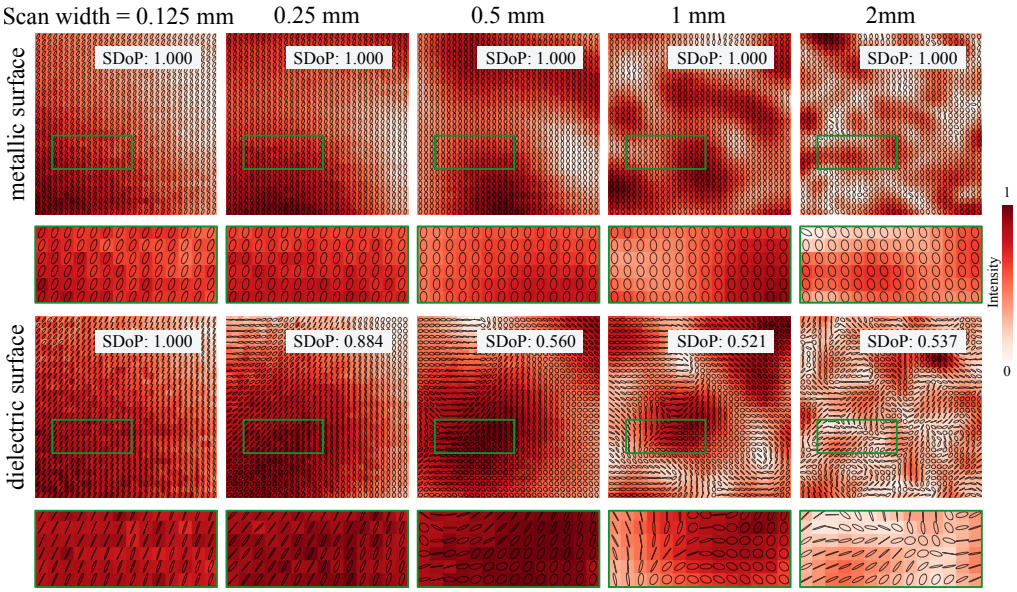

Fig. S7. Polarization and intensity speckle under increasing FOV (dielectric vs. metallic). For each FOV, we show intensity-speckle maps with polarization ellipse overlays for a metallic surface (top) and a dielectric surface (bottom). While both materials exhibit qualitatively similar intensity speckle as the FOV increases, their polarization behavior diverges: dielectric targets show increasingly randomized polarization speckle with decreasing SDoP, whereas metallic targets maintain a spatially coherent polarization transformation with SDoP remaining close to 1.

D. Implementation Details

All calibration parameters are optimized using the Adam optimizer with a learning rate of 10^{-2}. We run the optimization for 2000 iterations using full-batch updates over all calibration angles. All angular parameters are initialized from uniform random distributions over their valid ranges, while phase parameters are initialized uniformly over $[0, 2\pi)$. Amplitude-related parameters are initialized with small random values.

3. POLARIZATION SPECKLE

A. Intensity and Polarization Speckle under Varying Field of View

We compare intensity speckle and polarization speckle as the scanned field of view (FOV) increases, using two representative planar targets: a dielectric surface and a metallic surface. The FOV is swept by enlarging the square scan region from side length 0.125 mm to 2 mm. The laser spot size is 2 mm. For each FOV, we visualize the spatial distribution of intensity and polarimetric response and report the corresponding SDoP in Fig. S7.

Intensity speckle. Across both dielectric and metallic targets, the intensity patterns exhibit comparable qualitative changes with FOV: enlarging the scan region reveals increasingly diverse spatial interference structures, leading to visually richer intensity-speckle textures. This indicates that the basic coherent speckle formation mechanism is present in both materials under our imaging geometry.

Polarization speckle. In contrast, the behavior of polarization speckle shows a strong dependence on surface material. For dielectric surfaces, the spatial variation of reconstructed polarization states closely follows the intensity speckle pattern. As the FOV increases, polarization speckle grains become finer and more spatially randomized, resulting in a gradual reduction of the SDoP. This behavior is consistent with our micro-facet model, where locally varying micro-facet orientations and multiple scattering paths produce polarization responses that progressively decorrelate as the FOV grows and more independent scatterers are included. This indicates that dielectric polarization responses are dominated by locally varying micro-facet orientations and

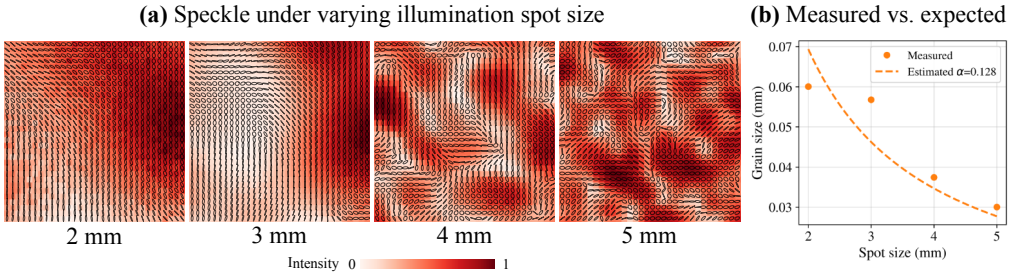

Fig. S8. Polarization speckle patterns under varying illumination spot sizes. (a) The polarization ellipse maps are reconstructed from the recovered Jones matrices for a fixed dielectric surface while varying the effective illumination aperture. Larger spot sizes lead to finer spatial variations in the polarization field, reflecting changes in the polarization speckle grain size. **(b)** Comparison between the measured polarization-speckle grain size and the expected inverse-aperture scaling across spot sizes, demonstrating consistency with the theoretical prediction.

multiple scattering paths, whose contributions decorrelate as more independent scatterers are included.

For metallic surfaces, however, the polarization speckle exhibits markedly different behavior. Despite increasingly fine intensity speckle with larger FOV, the reconstructed polarization states remain highly spatially coherent. The polarization ellipses exhibit consistent orientation and ellipticity across the field, indicating that the polarization response is governed by a surface-dominated reflection that preserves a stable polarization transformation. As a result, the SDoP remains close to unity across all tested FOVs, showing little sensitivity to changes in speckle grain size.

Implications. These observations support two conclusions that align with the main paper: (1) the distinct polarization-speckle behavior provides a robust cue for discriminating metallic versus dielectric surfaces, even when intensity-speckle trends are similar; and (2) for dielectric materials, polarization-speckle statistics (summarized by SDoP) provide a meaningful descriptor that can be further correlated with surface roughness, which we exploit in the main text in our quantitative analysis.

B. Polarization Speckle Size under Varying Illumination Spot Size

We experimentally study how the spatial scale of polarization speckle varies with the illumination spot size on a dielectric surface. The goal of this experiment is to validate the physical relationship between the effective aperture size and the resulting polarization speckle grain size under coherent illumination.

Experimental configuration. The experiment is conducted on a planar dielectric target with a smooth macroscopic geometry. To control the illumination spot size, we insert an adjustable aperture in front of the collimating lens, thereby varying the effective system aperture D_{eff} while keeping all other optical components and acquisition parameters fixed. This procedure changes the transverse extent of the beam incident on the surface, resulting in spot diameters of approximately 2 mm, 3 mm, 4 mm, and 5 mm, respectively. For each spot size, we raster-scan a fixed spatial region on the surface and reconstruct the complex Jones matrix at each spatial location. Polarization ellipses are then visualized to reveal the spatial structure of the recovered polarization field.

Observed polarization speckle patterns. Representative polarization ellipse maps for different spot sizes are shown in Fig. S8. When the illumination spot is small, the polarization field varies smoothly across space, indicating a relatively large polarization speckle grain size. As the spot size increases, the polarization ellipses exhibit increasingly rapid spatial fluctuations, corresponding to finer polarization speckle structures.

The observed behavior is consistent with the classical speckle scaling law for coherent imaging systems,

$$\Delta x \;=\; \alpha\,\frac{\lambda z}{D_{\text{eff}}}, \tag{S34}$$

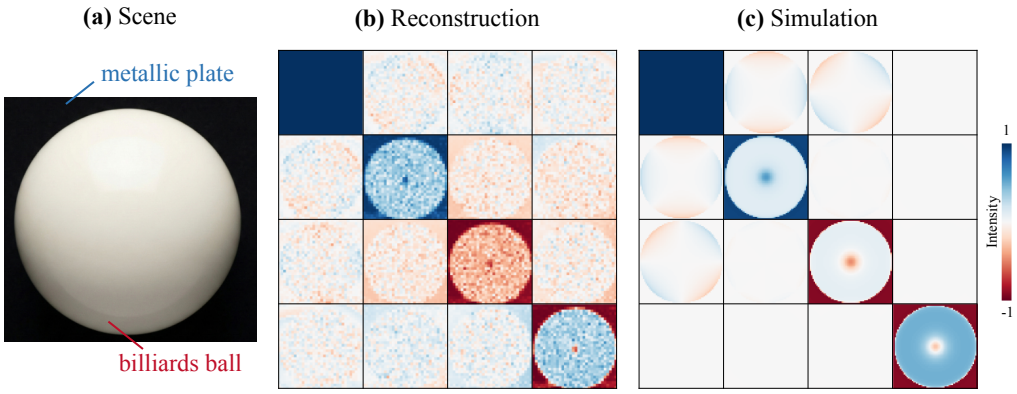

Fig. S9. Reconstructed versus simulated Mueller matrix. (a) Scene containing a dielectric billiard ball and a metallic plate. **(b)** Mueller matrix components reconstructed from experimental measurements. **(c)** Simulated Mueller matrix components under the same scene configuration. The reconstruction closely matches the simulation: the metallic surface and specular reflection observed at the ball center exhibit the expected diagonal signature $[1, 1, -1, -1]$, while off-diagonal terms (e.g., M_{01}, M_{02}, M_{10}, M_{20}) show surface-normal-dependent sign reversals (red/blue) that are consistently reproduced in the reconstruction.

where Δx denotes the characteristic speckle grain size, z is the propagation distance, λ is the wavelength, and D_{eff} is the effective aperture diameter, and α is a dimensionless proportionality constant that absorbs the dependence on the chosen grain-size definition. To estimate Δx from measurements, we compute the spatial autocorrelation of the recovered Jones matrices, and then perform radial averaging to obtain the radial autocorrelation $C(r)$. We report the speckle grain size using the conventional $1/e$ criterion [5], i.e., we take Δx to be the smallest radius r such that $C(r) = e^{-1}$. Fitting the measured grain sizes across spot sizes with the model in Eq. (S34) yields $\alpha = 0.128$, and the resulting prediction is consistent with the measured trend (Fig. S8b).

4. MUELLER REPRESENTATION

Challenges in Mueller reconstruction from coherent Jones measurements. Our coherent imaging system transmits and receives dual-channel, polarization-encoded Jones vector sequences, enabling the surface response to be directly reconstructed as a complex-valued Jones matrix $\mathbf{J}_{\text{surf}}(x, y)$ at each scanned spatial location. From this per-pixel Jones representation, a Mueller matrix can be obtained via the standard Jones-to-Mueller mapping. However, since a Jones matrix characterizes only the fully polarized component of light, the derived Mueller matrix is inherently incomplete on a per-pixel basis: depolarized contributions are not explicitly represented and therefore cannot be reliably recovered at a single pixel. In addition, coherent interference and spatially varying micro-geometry introduce strong polarization speckle, which manifests as spatial noise in per-pixel Mueller estimates. For regions dominated by fully polarized reflections, our method can reconstruct Mueller matrices accurately (see Fig. 6 in the main paper).

Spatial averaging for complete Mueller representation. To mitigate both incompleteness and speckle-induced variability, we adopt spatial averaging in the Mueller domain over local neighborhoods. This differs from conventional Mueller polarimetry that relies on temporal averaging—e.g., sequentially rotating waveplates or polarizers to synthesize multiple polarization states over time. In our scanned setting, spatial diversity provides an alternative set of independent measurements: averaging Mueller matrices from nearby pixels suppresses coherent speckle fluctuations while effectively accumulating the depolarized contribution over the region. As a result, spatial averaging yields a more stable and practically more complete Mueller representation for coherent measurements.

Polarimetric time-of-flight Mueller simulation. To provide a physics-based reference, we simulate the expected Mueller matrix under the same coaxial configuration using the model proposed by Baek *et al.* [6]. This model predicts the polarimetric time-of-flight response of a coaxial system by decomposing the scene reflectance into two components: (1) a surface-reflection term that models polarization changes induced by specular microfacet reflection via Fresnel reflection, and

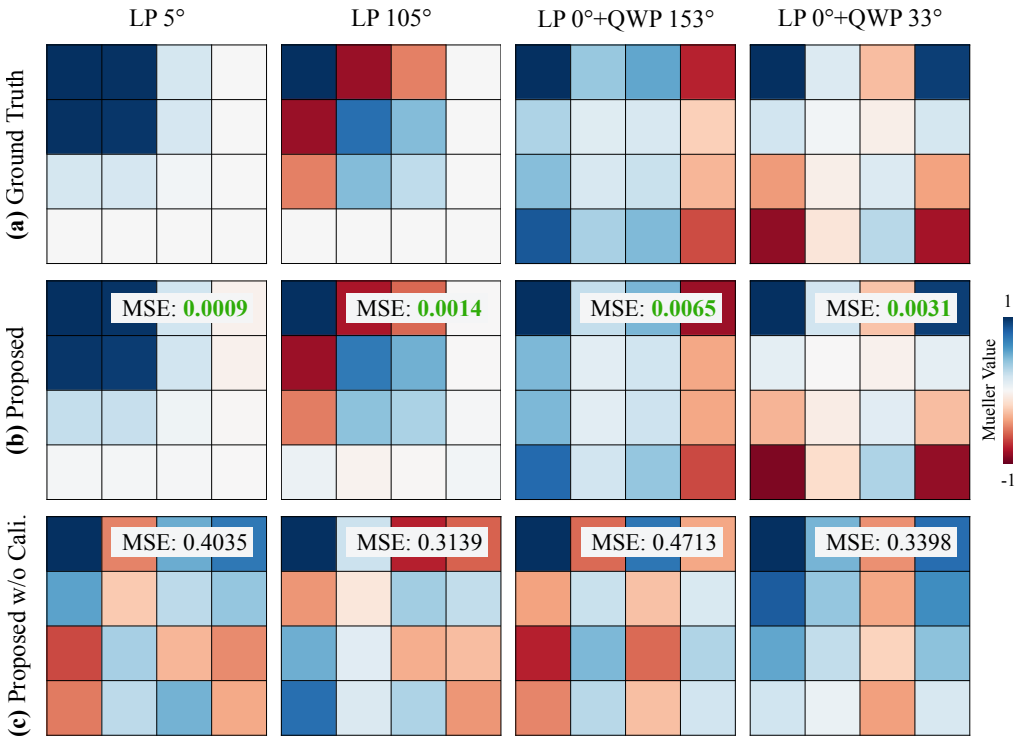

Fig. S10. Polarization reconstruction with vs. without calibration. (a) Ground-truth Mueller matrices for different input polarization states, including linear and elliptical polarization. **(b)** Reconstructions using the proposed method with polarization calibration, achieving consistently low MSE across all inputs. **(c)** Reconstructions using the proposed method but without polarization calibration, resulting in substantially larger errors.

(2) a sub-surface reflection term that accounts for transmission into the material followed by scattering-induced depolarization and delayed re-emission. We use this simulation as a reference for evaluating our reconstructed Mueller matrices.

Experimental validation. Fig. S9 shows a representative example consisting of a white dielectric billiard ball placed in front of a metallic plate. We scan the scene on a 128×128 grid, recover per-pixel Jones matrices, convert them to Mueller matrices, and then apply spatial averaging to obtain a 32×32 Mueller representation. The metallic plate exhibits a predominantly fully polarized response and is therefore reconstructed with high fidelity across Mueller components. For the billiard ball, the center region is dominated by strong specular reflection and also agrees well with the simulation. As shown in Fig. S9(c), Mueller elements M_{01}, M_{02}, M_{10}, and M_{20} exhibit sign reversals (red and blue in the figure) resulting from surface-normal-dependent Fresnel polarization under the global coordinate frame. The reconstructed Mueller matrices capture the same angularly varying sign patterns, although with noise due to coherent speckle.

Overall, this experiment demonstrates that the spatially averaged Mueller representation derived from the coherent Jones measurements can recover physically meaningful polarimetric structure, while substantially reducing speckle artifacts and alleviating limitations of the coherent-only Jones matrix description.

5. RESULTS

A. Polarization reconstruction with vs. without calibration

Fig. S10 compares polarimetric reconstructions obtained by the proposed pipeline *with* and *without* polarization calibration under multiple input states, including both linear and elliptical polarization. Without calibration, the recovered Mueller matrices are strongly corrupted by system-induced polarization distortions from internal optical interfaces (e.g., fiber propagation, the circulator, and other reciprocal optics), leading to large MSE and poor agreement with

the ground truth. In contrast, our polarimetric calibration explicitly removes these system effects, enabling accurate recovery of the scene-dependent polarization response and faithfully reproducing the expected linear and elliptical signatures with consistently low reconstruction error.

B. Comparison of depth and velocity reconstruction across algorithms

Fig. 7 in the main paper compares depth and radial velocity reconstruction results obtained using a matched-filter baseline, the full-wavefield lidar (FWL) method [7], and the proposed approach. The matched-filter baseline estimates depth by locating peaks in the cross-correlation between the received and transmitted signals. However, in dynamic scenes, both propagation delay and Doppler-induced frequency shifts affect the correlation response. As a result, direct cross-correlation cannot disentangle depth delay from motion-induced phase modulation, leading to inaccurate depth estimates for moving objects and preventing explicit velocity recovery.

The FWL method [7] addresses this limitation by explicitly modeling Doppler shifts through a convolutional optimization over joint delay and velocity bins. While this formulation enables simultaneous depth and velocity reconstruction, it relies on time-domain grouped convolutions with a large number of parameters, resulting in substantial computational cost. As indicated by the reported runtimes in the main paper, FWL is significantly slower than the proposed method. Moreover, both depth and velocity estimates become unreliable in regions with weak signal returns, such as specular surface patches with low effective backscatter. As illustrated in Fig.7 of the main paper, these regions are dominated by noise-like artifacts in both the reconstructed depth and velocity maps.

In contrast, the proposed method formulates reconstruction as a statistically grounded negative log-likelihood optimization over a joint depth–velocity Jones response. By evaluating the forward model using FFT-based linear convolution and precomputing Doppler-modulated transmitted spectra, the optimization is performed efficiently in the frequency domain. This design achieves a nearly threefold wall-clock speedup over FWL when run on the same hardware (an Intel Xeon Gold 5120 CPU and an NVIDIA TITAN RTX GPU). In addition, the likelihood-based formulation explicitly accounts for incoherent components, enabling more robust estimation in low-SNR regions. As a result, the proposed method produces depth and velocity reconstructions with fewer artifacts compared to FWL.

Velocity ground truth generation by high-speed camera. We estimate the ground-truth optical-axis velocity of the spinning disk shown in Fig. 7 by the following. First, we measure the disk's angular velocity (ω_p) with a high-speed camera. A strip of tape is attached radially to the disk to introduce photometric texture. We then capture a video at a sufficiently high frame rate (10kfps, SPAD512 [8]) to localize the angular position of this strip over time. We compute a temporal Fourier transform and estimate ω_p from the fundamental frequency peak of the resulting spectrum, obtained by summing magnitudes across the selected pixel region.

Next, we compute a 3D point map of the scene using a high-quality depth map together with the known deflections of the scanning galvanometer mirrors. The disk is segmented as a connected component via flood fill. Knowing the 3D point locations $\vec{r}$ of the disk relative to the center of projection of the scanning mirrors, we estimate the disk plane normal $\hat{n}_p$ by least-squares plane fitting. The disk center $\vec{o}_p$ is recovered from perimeter sample pairs of the segmented point cloud using a chord-and-perpendicular-bisector construction.

Finally, for each point in the velocity map, we calculate the tangential velocity of the spinning disk $\vec{v}_t = (\vec{r} - \vec{o}_p) \times \omega \hat{n}_p$ and take its component along the scanning axis.

C. Quantitative Performance under Varying Depth, Velocity, and SNR

We further evaluate the reconstruction accuracy under different target depths, target velocities, and signal-to-noise ratios (SNRs). The effective matched-filter SNR is computed as the ratio between the signal power at the reconstructed target depth bin and the noise-floor power measured from an empty depth region. We vary the laser power to obtain effective SNRs of 11.64, 10.51, 6.70, 5.38, 3.23, and 0.22 dB.

For the depth experiment, we scan a planar target along a line under different laser powers. Since the target is planar, its depth profile across the scan follows a physical plane-induced depth model determined by the galvanometer scan angle. We fit this model to the highest-SNR scan and use the fitted planar profile as the ground-truth depth. Reconstruction error is then computed by comparing the estimated depth at each scan position with this fitted ground truth. As shown in Fig. S11(a), depth reconstruction remains accurate across the tested depth range

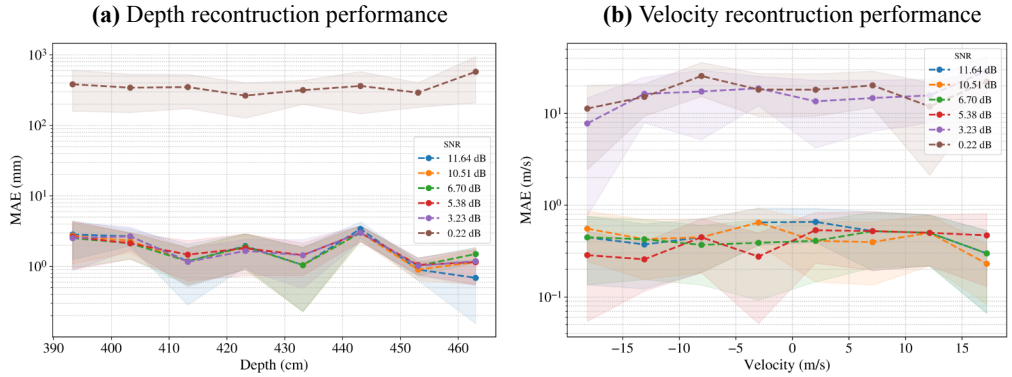

Fig. S11. Quantitative reconstruction performance under different SNR levels. (a) Depth MAE evaluated across target depths, with ground truth obtained by fitting a physical planar-depth model to a high-SNR planar-target scan. **(b)** Velocity MAE evaluated across target velocities, with ground truth calibrated using a high-speed camera measurement of the rotating disk. Depth reconstruction achieves an average MAE of about 2 mm down to 0.22 dB SNR, while velocity reconstruction achieves an average MAE of about 0.48 m/s for SNRs above 3.23 dB.

above the operating SNR threshold, with millimeter-level MAE. At the lowest SNR, the target return becomes too weak relative to the noise floor and residual background, leading to a sharp increase in error.

For the velocity experiment, we scan a high-speed rotating disk to generate controlled radial velocities over the target surface. The ground-truth velocity is calibrated using a high-speed camera measurement of the disk rotation, following the procedure described in Sec. 5.B. As shown in Fig. S11(b), velocity reconstruction remains stable for SNRs above approximately 3.23 dB, with an average absolute error of about 0.48 m/s across the tested velocity range. Below this SNR level, the Doppler-dependent phase information is no longer reliable, and the reconstruction may select spurious Doppler responses, causing the velocity MAE to increase sharply.

D. Additional Polarimetric Imaging-through-Scattering Results

We evaluate polarimetric imaging through scattering media under a wider range of scattering strengths. We consider two scattering regimes: thin diffuser layers with varying grit levels (Thorlabs DG20-1500, DG10-600, DG20-220, and DG10-120) and thicker semi-translucent PMMA acrylic slabs with varying thicknesses (1, 1.5, 2, and 3 mm). For both regimes, the target is a dielectric "T" attached to a metallic plate, and the measurement is performed in a reflective geometry, so light traverses the scattering medium twice.

We quantify scattering strength using the transmittance, defined as the ratio of collected power with versus without the scattering medium under the same illumination and aperture-matched collection geometry. For the diffuser layers, this transmittance ranges from 45.71% to 3.12%. For the PMMA slabs, the measured transmittance ranges from 36.36% to 0.08%, corresponding to optical thicknesses from 1.01 to 4.76.

To evaluate material recovery, we apply the same two-component clustering model to either intensity features or Mueller-matrix features, segmenting each pixel into metallic-background or dielectric-target regions. We compute PSNR against a binary "T" mask obtained from a reference scan without the scattering medium. As shown in Fig. S12, polarization-based clustering achieves higher PSNR than intensity-based clustering across both diffuser and PMMA conditions. The advantage decreases as scattering becomes stronger, indicating that the method has a finite operating range rather than arbitrary recovery through strongly scattering media.

E. Additional Simultaneous Depth, Velocity, and Polarimetric Reconstruction

Material-aware scene sensing is useful in robotic inspection, sorting, and manipulation, where decisions may depend not only on object geometry and motion, but also on material composition. Conventional lidar outputs such as depth and intensity can localize objects and recover scene geometry, but they do not directly indicate whether a region is metallic or plastic.

To illustrate this capability, we capture a tabletop inspection scene containing a metallic plate, a plastic plate, a metallic can with partially removed plastic packaging, a spinning hemisphere,

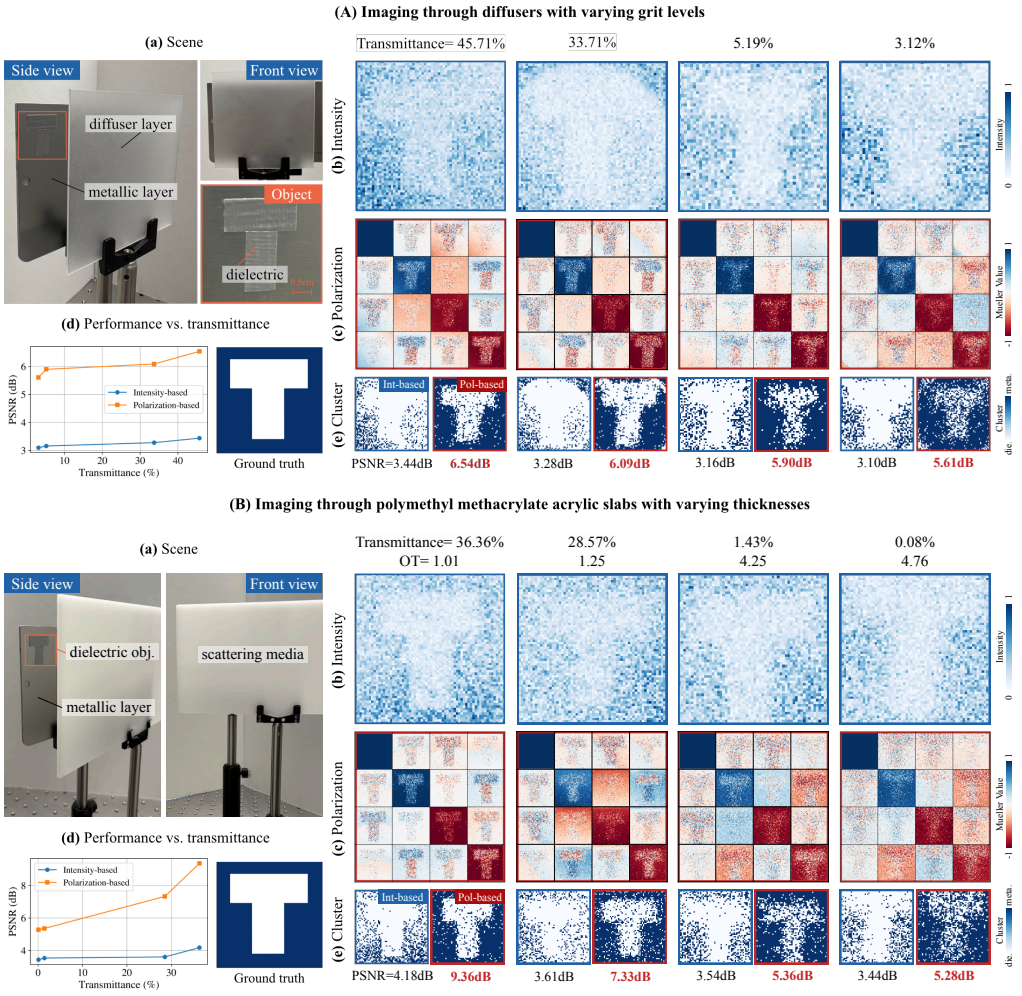

Fig. S12. Polarimetric imaging through scattering media with varying scattering strength. **(A)** Imaging through thin diffuser layers with different grit levels (Thorlabs DG20-1500, DG10-600, DG20-220, and DG10-120). **(B)** Imaging through polymethyl methacrylate acrylic slabs with different thicknesses (1, 1.5, 2, and 3 mm). In both experiments, a dielectric "T" target is attached to a metallic plate and imaged in a reflective imaging-through-scattering geometry. Scattering strength is quantified by the transmittance for both media and, additionally, by optical thickness (OT) for the PMMA slabs. For each scattering medium, **(a)** shows the experimental configuration, **(b)** shows intensity reconstructions, **(c)** shows reconstructed Mueller-matrix responses, **(d)** reports PSNR as a function of measured in-aperture transmittance, and **(e)** compares material clustering using intensity-based and polarization-based features. The PSNR is computed against the binary "T" ground truth. Across both diffuser and PMMA media, polarization-based clustering consistently outperforms intensity-based clustering, while performance degrades as scattering becomes stronger.

and a figurine. As shown in Fig. S13(b), the depth map recovers the spatial layout of the objects, while the Doppler velocity map in Fig. S13(d) identifies the moving hemisphere. In contrast, the polarization map in Fig. S13(c) provides complementary material-sensitive information: it distinguishes the metallic plate from the plastic plate and separates the plastic packaging from the exposed metallic region of the can.

F. Discussion of Limitations

Limitations of coaxial systems for polarization reconstruction. Coaxial transmit–receive geometries impose an inherent angular selectivity on the measured return. Since the receiver collects light predominantly within a narrow cone around the optical axis, the signal is dominated by micro-

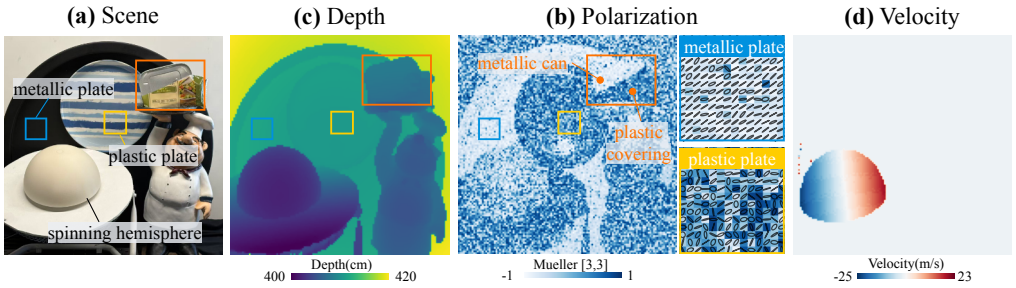

Fig. S13. Simultaneous depth, velocity, and polarimetric reconstruction. (a) Experimental scene containing a metallic plate, a plastic plate, a metallic can that is partially covered by a plastic wrap, a spinning hemisphere, and a figurine, providing both static material variation and dynamic motion. **(b)** Reconstructed depth map, which recovers the 3D layout of the scene. **(c)** Reconstructed polarization map, visualized by the Mueller M_{33} component and local polarization ellipses. The polarimetric response clearly distinguishes the metallic and plastic plates, and also reveals the difference between the metallic and plastic regions of the can, which could not be distinguished from the depth map alone. **(d)** Reconstructed Doppler velocity map, which captures the motion of the spinning hemisphere.

facets whose specular directions fall inside this cone. When the macroscopic surface normal is not aligned with the optical axis, only a small subset of microfacets contributes to the measurement, reducing the collected energy and yielding an incomplete sampling of the surface reflectance.

This constraint also limits polarization observability. The recovered Jones response primarily reflects Fresnel-induced polarization effects under near-specular reflections from the contributing microfacets. Specular components from other microfacets are not collected, even though they may carry distinct polarization signatures; consequently, part of the Fresnel-dependent polarimetric information is missing in a coaxial measurement.

Finally, off-specular and multiply scattered components are only partially captured and cannot be uniquely attributed to a single deterministic Jones response, which is a fundamental limitation of any coaxial (and near-coaxial) systems. While our statistical model absorbs their energy through the incoherent term, the recovered polarimetric quantities should be interpreted as effective responses of the subset of microfacets visible to the coaxial aperture. Extending the sensing geometry (e.g., non-coaxial or multi-view collection) to cover a wider range of reflection angles is a natural direction for future work.

Acquisition time and point-scanning overhead. Although the optical communication module supports extremely high symbol rates, enabling microsecond-level effective exposure times per spatial location, the overall acquisition time is dominated by data readout and storage. In the current implementation, each point measurement incurs a latency of roughly one second due to the hardware readout, buffering, and disk I/O, making point-by-point scanning the primary bottleneck for acquisition speed. As a result, the total capture time scales linearly with the number of scanned spatial locations, despite the negligible optical dwell time (approximately one microsecond). Future improvements could substantially reduce this overhead by adopting parallel acquisition pipelines, streaming-based readout, or more efficient memory and buffer management, thereby better matching the high-throughput capability of the optical front end of the coherent optical modem.

REFERENCES

1. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics, 2-Volume Set* (John Wiley & Sons, 2019).
2. G. P. Agrawal, *Fiber-Optic Communication Systems* (John Wiley & Sons, 2012).
3. D. H. Goldstein, *Polarized Light* (CRC Press, 2017).
4. R. Chipman, W. S. T. Lam, and G. Young, *Polarized Light and Optical Systems* (CRC Press, 2018).
5. J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications* (Roberts and Company Publishers, 2007).
6. S.-H. Baek and F. Heide, "All-photon polarimetric time-of-flight imaging," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, (2022), pp. 17876–17885.

7. P. Mirdehghan, B. Buscaino, M. Wu, *et al.*, "Coherent optical modems for full-wavefield LiDAR," in *Proc. SIGGRAPH Asia,* (2024).
8. AXIOM OPTICS, "Spad512 single-photon avalanche diode array," https://www.axiomoptics.com/products/spad512s/ (2023).