

Dirichlet Splatting: Differentiable Rendering for Wave-Based Inverse Problems

XINGYU CHEN, University of California San Diego, USA
WUQIONG ZHAO, University of California San Diego, USA
XINYU ZHANG, University of California San Diego, USA
TZU-MAO LI, University of California San Diego, USA

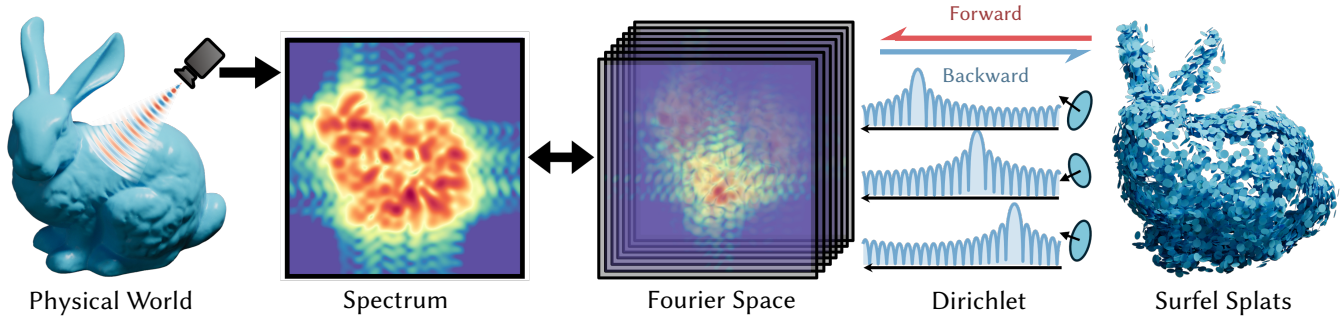


Fig. 1. *Dirichlet Splatting* renders coherent wave-based measurements with a primitive whose spectral footprint is the physically exact Dirichlet kernel of a finite-window discrete Fourier transform. A coherent appearance model makes each surflet behave like a real surface element rather than an isotropic point reflector, while a tailored sliding Frank-Wolfe solver recovers reflector geometry from cold starts where Gaussian-splat baselines diverge.

We present *Dirichlet Splatting*, a physically grounded differentiable rendering method for inverse problems with wave-based coherent sensors. Unlike optical cameras, these sensors observe a scene in a band-limited Fourier space whose continuous point response is complex-valued and oscillatory. Our key observation is that the point spread function of a reflector in this Fourier space is the *Dirichlet kernel*, not a Gaussian. We instantiate each splat as a Dirichlet kernel in closed form, model each primitive as a planar surflet rendered through a *surflet-aware Dirichlet splat* whose complex amplitude is set by wave propagation physics, and introduce a sliding Frank-Wolfe adaptation tailored to the rugged, non-convex landscape that the Dirichlet oscillation induces. On terahertz tomography, Dirichlet Splatting recovers reflector geometry with an order of magnitude fewer parameters, faster convergence, and higher accuracy than 3D Gaussian Splatting.

CCS Concepts: • **Computing methodologies** → **Rendering**; **Reconstruction**; **Computational photography**.

Additional Key Words and Phrases: differentiable rendering, inverse rendering, coherent imaging, Dirichlet kernel, terahertz imaging, radar, splatting

ACM Reference Format:

Xingyu Chen, Wuqiong Zhao, Xinyu Zhang, and Tzu-Mao Li. 2026. Dirichlet Splatting: Differentiable Rendering for Wave-Based Inverse Problems. *ACM Trans. Graph.* 45, 6, Article 199 (December 2026), 19 pages. <https://doi.org/10.1145/3842559>

Authors' Contact Information: Xingyu Chen, University of California San Diego, La Jolla, California, USA, xic063@ucsd.edu; Wuqiong Zhao, University of California San Diego, La Jolla, California, USA, wqzhao@ucsd.edu; Xinyu Zhang, University of California San Diego, La Jolla, California, USA, xyzhang@ucsd.edu; Tzu-Mao Li, University of California San Diego, La Jolla, California, USA, tzli@ucsd.edu.



This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 1557-7368/2026/12-ART199
<https://doi.org/10.1145/3842559>

1 Introduction

Differentiable rendering has become a foundational tool for inverse problems in graphics and sensing. By making the forward image-formation pipeline differentiable end-to-end, it turns scene reconstruction into gradient-based optimization, recovering geometry, materials, lighting, and motion directly from sensor measurements [Kerbl et al. 2023; Mildenhall et al. 2020]. Nevertheless, *non-visible-spectrum* sensors, including terahertz (THz) tomographs, sonars, and radar, equally demand differentiable rendering to solve their inverse problems.

3D Gaussian Splatting (3DGS) [Kerbl et al. 2023] renders scenes via differentiable rasterization of anisotropic Gaussian primitives. Recent work extends 3DGS to wave-based sensors via phase-coherent combination of splats [Wen et al. 2026; Yang et al. 2025]. However, this port carries a structural mismatch, as the Gaussian primitive, designed to approximate optical point spread, cannot represent the complex-valued, oscillatory point response that coherent wave measurement produces.

Optical imaging interpolates a smooth radiance field through a near-Gaussian aperture, making a Gaussian splat a faithful local approximation of how light from a small region accumulates in a pixel. Wave-based coherent imaging is structurally different. A sparse scene of reflectors drives an analytic, continuous physical operator, and the sensor records its values on a discrete band-limited grid only at the readout stage of acquisition. The natural target of an inverse solver is therefore the continuous physics, namely the positions, orientations, and amplitudes of the reflectors that drive this operator, rather than the post-processed grid it produces.

The point spread function (PSF) of this operator is complex-valued and oscillatory (right inset), with sidelobes carrying analytic information about the reflector geometry and a phase that governs coherent interference between reflectors [Harris 1978]. Existing ports of 3DGS approximate the operator by a smooth real envelope on the discrete grid. The approximation collapses both features and reduces the inverse problem to a fit on a sampled output rather than a differentiation of the underlying physics.

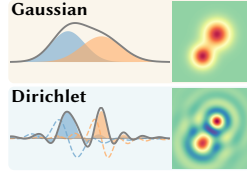
We present *Dirichlet Splatting*, a differentiable rendering method for inverse problems with wave-based coherent sensors. *First*, the PSF of a reflector in this band-limited Fourier space is a *Dirichlet kernel* [Harris 1978]. We therefore instantiate each splat as a Dirichlet kernel evaluated in $O(1)$ closed form. *Second*, we model each primitive as a planar *surf*el and render it through a *surf*el-aware *Dirichlet splat* whose complex amplitude is set by wave propagation physics, yielding a finite-area surface element rather than an isotropic point reflector. However, the very oscillation that makes the Dirichlet kernel physically faithful induces a rugged, non-convex loss landscape. Generic first-order optimizers such as Adam encounter spurious gradient nulls in every sidelobe and stall before recovering even simple multi-reflector scenes from cold starts. *Third*, we adapt *Sliding Frank–Wolfe* [Bredies and Pikkariainen 2013; Denoyelle et al. 2020] to the Dirichlet landscape. Our adaptation, *Dirichlet Sliding Frank–Wolfe* (DSFW), replaces the standard kernel-translation step with variable projection over surfel parameters, residual dual certificates that drive hard replacement of low-utility surfels, and periodic low-resolution coupled Levenberg–Marquardt correction.

In summary, this paper makes the following contributions:

- (i) A *Dirichlet splat primitive* whose spectral footprint is the exact point spread function in band-limited Fourier space (§5.2.1).
- (ii) A *surf*el-aware coherent appearance model that makes each splat behave as a real surface element rather than an isotropic point reflector (§5.2).
- (iii) A sliding Frank–Wolfe adaptation tailored to the rugged Dirichlet loss landscape (§6).
- (iv) A real-data evaluation on sub-THz tomography, with ablations of the surfel modulation and the solver components, and convergence analysis (§7).

2 Related Work

Differentiable rendering and splatting primitives. Differentiable rendering enables gradient-based optimization of scene parameters from image observations, with NeRF-family volumetric fields [Baron et al. 2022; Mildenhall et al. 2020; Müller et al. 2022] and differentiable rasterization [Laine et al. 2020] as representative families. 3DGS [Kerbl et al. 2023] introduced a real-time alternative based on projected 3D Gaussian primitives, since extended to dynamic scenes [Luiten et al. 2024] and surface reconstruction [Huang et al. 2024b]. A common thread across these methods is that the choice of kernel (i.e., Gaussian) is motivated by mathematical convenience rather than by the physics of image formation. This is reasonable for optical cameras with approximately Gaussian PSFs but structurally



inappropriate for sensors whose image formation is a finite-window discrete Fourier transform (DFT).

In contrast, we preserve the splatting paradigm and its efficiency while replacing the kernel with the physically exact Dirichlet response of the underlying DFT. A parallel line of work targets computer-generated holography (CGH) by splatting complex-valued primitives directly onto a wavefront [Chao et al. 2025; Choi et al. 2025; Zhan et al. 2026]. These methods use complex-valued Gaussians to capture coherent superposition (e.g., diffraction, interference) and demonstrate state-of-the-art CGH quality. However, equipping a Gaussian with a complex amplitude changes *what* the primitive carries, not the shape of the envelope it splats. The envelope itself remains a smooth, monotonic Gaussian, which cannot reproduce the periodic sidelobe nulls or the phase law of a windowed DFT response (§3). The kernel mismatch is therefore structural rather than parametric, and persists regardless of how rich the amplitude representation becomes. Our Dirichlet primitive instead replaces the envelope itself with the analytic windowed-DFT response, embedding coherent sidelobe and phase structure into the primitive rather than leaving them as residuals for the optimizer to absorb.

Wave-optical rendering and appearance models. A separate line of work makes the light transport itself wave-optical. Generalized ray formulations carry phase, coherence, and diffraction through the rendering integral and reduce to geometric optics in the incoherent limit [Steinberg et al. 2024; Steinberg and Yan 2021], while wave-optical appearance models replace the geometric-optics BRDF with a diffraction-aware response derived from surface microgeometry [Holzschuch and Pacanowski 2017; Stam 1999]. These methods target passive imaging, where a scene is lit by ambient or area sources and the sensor integrates a partially coherent field over a finite aperture. We instead target active coherent sensing, where the sensor illuminates the scene with a known waveform and records the complex baseband return, so that the measurement is a finite-window DFT whose point response is the Dirichlet kernel in closed form (§3). A wave-optical renderer describes how a field arrives at an aperture, while our appearance model describes what a surfel contributes to a Fourier-space bin after that field has been mixed, sampled, and transformed.

Differentiable coherent sensing. Simulation of coherent wave-based sensors has traditionally relied on ray-tracing methods that propagate electromagnetic or acoustic waves geometrically and accumulate path losses, reflections, and scattering [Auer et al. 2016; Chen et al. 2025; Hirsenkorn et al. 2017; Scheiner et al. 2020]. Such simulators produce realistic time-domain signals but are computationally expensive and difficult to differentiate. Learning-based radar synthesis instead trains generative models [Chen and Zhang 2023; Ouaknine et al. 2021; Wheeler et al. 2017] to bypass physical simulation entirely, at the cost of physical grounding and gradient-based inversion. AutoMS and FlowForm use differentiable RF simulation to reshape mmWave coverage by optimizing passive metasurfaces [Ma et al. 2024; Zhao et al. 2026]. A more direct route is to differentiate through the full waveform pipeline (e.g., signal generation, mixing, sampling, fast Fourier transform (FFT)) with automatic differentiation [Chen et al. 2026]. This is physically exact but materializes the full time-domain signal and its FFT gradient at $O(MN \log MN)$ per

step. Recent neural-scene approaches such as DART [Huang et al. 2024a] and Radar Fields [Borts et al. 2024] adapt NeRF-style representations to radar by combining implicit geometry with an explicit or learned sensor model, but they either operate on detection-level data (i.e., discarding spectral information) or approximate the radar forward model with a neural network rather than its analytic DFT response. We instead render directly in the spectral domain using closed-form Dirichlet responses, encoding the known DFT physics into the rendering kernel itself rather than learning it from data.

Sparse recovery and variable projection. That the DFT of a finite-length sinusoid is a Dirichlet kernel [Harris 1978; Oppenheim et al. 1999] and the related windowed-FFT super-resolution practice [Harris 1978] provide the building blocks for our renderer. Off-the-grid sparse recovery contributes sliding Frank–Wolfe with variable projection [Denoyelle et al. 2020; Golub and Pereyra 1973; Jaggi 2013] and subspace methods such as MUSIC [Schmidt 1986] or ESPRIT [Roy and Kailath 1989] bypass DFT resolution by exploiting signal-subspace structure but do not provide gradients to a renderer. Our optimizer combines sliding Frank–Wolfe, variable projection, and residual dual certificates inside a differentiable splatting pipeline, a combination that, to the best of our knowledge, has not previously been applied to differentiable rendering.

3 Background and Motivation

Wave-based coherent sensors form images by transmitting a known waveform, mixing the received echoes against it, and applying a finite-length DFT along each of L measurement axes to assemble a complex spectral tensor (Appendix A). The axis count L is modality-dependent (e.g., $L = 1$ for coherent spectroscopy, $L = 2$ for synthetic-aperture imaging, $L = 3$ for range/Doppler/azimuth sensing). The PSF-shaping step in the FFT-based imaging pipeline we target is a finite-window DFT. Each spectral bin records a complex amplitude, and contributions from nearby reflectors interfere through their phase before any intensity is observed. Fig. 2 sketches the source (a coherent bistatic array), the targets (planar surfels with normal and velocity), and the resulting Dirichlet PSF in the spectral tensor.

A single point reflector at continuous bin offset Δ from a spectral bin center contributes the exact analytic *Dirichlet kernel*

$$\kappa_N(\Delta) = e^{-j\pi \frac{N-1}{N} \Delta} \frac{\sin(\pi \Delta)}{\sin(\pi \Delta / N)}, \quad (1)$$

where N is the length of the DFT window applied by the sensor along the corresponding measurement axis (i.e., the number of samples the hardware coherently integrates before transforming). Eq. (1) is the exact DFT response of a windowed complex exponential, derived in Appendix A.1. The kernel is complex-valued, oscillatory, and periodic, with sidelobes carrying $\sim 10\%$ of the total spectral energy for a rectangular window [Harris 1978] and a phase that governs coherent interference between reflectors. The sidelobe ratio and phase law are fixed by the window length N and require no fitting.

A Gaussian splat (even with a complex amplitude) decays monotonically and has no zero crossings, reproducing neither the periodic sidelobe nulls nor the phase law that govern coherent interference between reflectors at sub-bin separation. The resulting modeling

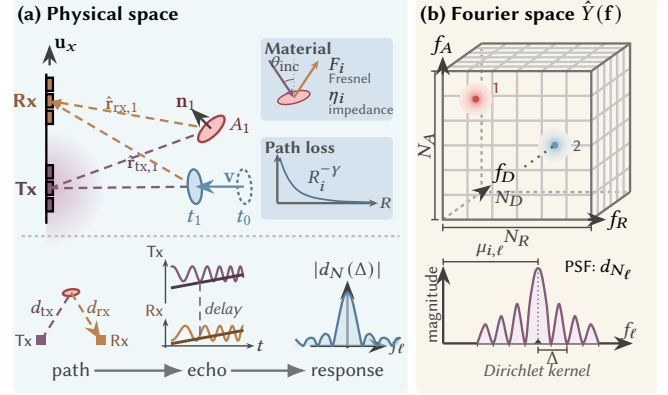


Fig. 2. **The Dirichlet forward model maps physical space to Fourier space.** (a) A coherent bistatic array images a static and a moving surfel; the bottom row traces the signal chain (path, echo, response) along one measurement axis. (b) Each surfel becomes a Dirichlet PSF in the spectral tensor $\hat{Y}(\mathbf{f})$; motion shifts Surfel 2 off the $f_D = 0$ face. The inset shows the 1D kernel $d_N(\Delta)$.

error dominates any reconstruction of coherent data. The reflector-domain sparsity of wave-based scenes (most volume is empty, with specular returns concentrating energy on isolated surfaces) keeps a system-rank-equal Dirichlet representation tractable, and adopting κ_N as the rendering primitive matches the windowed-DFT measurement physics exactly [Çetin and Karl 2001].

4 Overview

Dirichlet Splatting performs differentiable scene reconstruction from active coherent measurements (THz, acoustic, or mmWave). A known illuminator emits a band-limited waveform, and the sensor records the complex echo as an L -axis spectral tensor whose axes are modality-specific (e.g., depth alone for tomographic sensing; range, Doppler, and azimuth for transceiver-array sensing; frequency alone for coherent spectroscopy). The scene is modeled as oriented planar surfels that scatter the incident waveform once before reaching the sensor. Unlike the emissive primitives of standard 3D Gaussian Splatting (3DGS), our surfels are reflectors whose appearance is governed by complex reflectivity, bistatic foreshortening, and a material-dependent Fresnel response. Each surfel contributes a 3D Dirichlet kernel, the exact point-spread function of the windowed DFT along each axis, placed directly in the spectral domain rather than in the primal image domain of standard splatting.

The renderer’s projection, surfel-aware splats, and coherent compositing (§5) feed DSFW (§6), an optimizer built on variable projection and residual dual certificates.

5 Dirichlet Splatting

We render a scene of S planar surfels into a coherent L -axis spectral tensor in three logical steps. Each surfel carries a small parameter set Θ_i (§5.1). The renderer projects it to a continuous Fourier coordinate $\mu_{i,\ell}$ on each measurement axis ℓ via a sensor-specific map (§5.3), evaluates a surfel-modulated Dirichlet kernel whose width is set by the FFT length (§5.2), and coherently sums across all surfels (§5.3.2).

More specifically, the full forward model is

$$\hat{Y}(\mathbf{f}) = \sum_{i=1}^S \underbrace{\alpha_i \cdot \sqrt{A_i} V_i F_i R_i^{-\gamma}}_{\text{amplitude}} \cdot \underbrace{\prod_{\ell=1}^L d_{N_\ell}(\mathbf{f}_\ell - \mu_{i,\ell})}_{\text{kernel}}, \quad (2)$$

$\text{splat}_i(\mathbf{f})$

a coherent sum of per-surfel splats $\text{splat}_i(\mathbf{f})$, each with a complex amplitude (set by surfel physics) modulated by a Dirichlet point spread (fixed by the per-axis FFT lengths N_ℓ). The product form across ℓ is the exact separable DFT response of a single point reflector (Appendix A.4). We evaluate the spectral tensor at the continuous Fourier coordinate $\mathbf{f} = (f_1, \dots, f_L)$, where ℓ indexes the measurement axes. The measured bins lie at integer coordinates of $\mathbf{f}$. Each kernel factor uses the fractional offset $f_\ell - \mu_{i,\ell}$ between a bin and a surfel's continuous center. In the $L = 3$ running example, $\mathbf{f} = (f_R, f_D, f_A)$ for range, Doppler, and azimuth. The quantities $\alpha_i, A_i, V_i, F_i, R_i, \gamma$ are respectively the complex reflectivity, oriented-disk area, bistatic foreshortening, Fresnel response, sensor-to-surfel range, and path-loss exponent of surfel i (§5.1, §5.2). The per-axis Fourier center $\mu_{i,\ell}$ is the projection of $(\mathbf{p}_i, \mathbf{v}_i)$ onto axis ℓ (§5.3).

5.1 Primitive Parameterization

We represent the scene as a collection of planar surfels, following the geometric primitive of 2DGS [Huang et al. 2024b] but replacing its appearance model in full (§5.2).

Each of the S surfels carries the parameter set

$$\Theta_i = \{\mathbf{p}_i \in \mathbb{R}^3, \mathbf{n}_i \in \mathbb{S}^2, A_i \in \mathbb{R}_{>0}, \alpha_i \in \mathbb{C}, \mathbf{v}_i \in \mathbb{R}^3, \eta_i\}, \quad (3)$$

where $\mathbf{p}_i$ is the surfel center, $\mathbf{n}_i$ its outward normal, A_i the area of the oriented disk it represents, α_i a complex reflectivity $\mathbf{v}_i$ a per-surfel translational velocity that drives Doppler-axis behavior (§5.3; $\mathbf{0}$ for static scenes), and $\eta_i \in \mathbb{C}$ the wave impedance of the surfel material, entering the Fresnel response (§5.2). The geometric triple $(\mathbf{p}_i, \mathbf{n}_i, A_i)$ describes a small oriented disk. The appearance pair (α_i, η_i) describes the scene's coherent reradiation, distinct from 3DGS's spherical-harmonic (SH) colors. α_i is a scalar complex coefficient that scales the whole splat and carries the surfel's magnitude and phase. η_i enters only through the Fresnel factor $F_i \equiv F(\mathbf{n}_i, \theta_{\text{inc},i}, \eta_i)$ of Eq. (2), which controls the angular variation of that magnitude (Eq. (5)). A Dirichlet surfel carries 13 parameters versus 58 for a 3DGS primitive ($\sim 4.5\times$ fewer; Tab. 1).

5.2 Surfel-Aware Dirichlet Splat

The per-surfel splat $\text{splat}_i(\mathbf{f})$ in Eq. (2) is the complex contribution of surfel i to $\hat{Y}(\mathbf{f})$. Surfel attributes enter only as a pre-multiplicative complex amplitude on the kernel product, cleanly separating the sensor model (the kernel κ_{N_ℓ} , evaluated through the unit-peak form $d_{N_\ell} = \kappa_{N_\ell}/N_\ell$) from the scene model $(\alpha_i, A_i, V_i, F_i)$. The Fourier coordinate $\mu_{i,\ell}$ is defined per axis in §5.3, Eq. (6).

5.2.1 Complex Dirichlet kernel. Each surfel's energy spreads across the Fourier-domain bins through the Dirichlet kernel of §3, the exact point-spread function of an N -point DFT. We parameterize it in *bin units*: for a surfel with continuous center μ evaluated at integer bin k , the bin offset is $\Delta = k - \mu$. We adopt the analytic Dirichlet form of Eq. (1) (§3) as the per-measurement-axis kernel (one factor per axis

Table 1. **Primitive parameter comparison.** Counts are degrees of freedom: a unit quaternion is three, a unit normal two, and the complex amplitude is solved in closed form (§6.2). Most of the reduction comes from appearance: a coherent sensor uses one complex amplitude, whereas a camera uses a directional radiance field.

Attribute	3D Gaussian Splat		Dirichlet Surfel (ours)	
	expression	#par	expression	#par
Position	$\mathbf{p} \in \mathbb{R}^3$	3	$\mathbf{p} \in \mathbb{R}^3$	3
Orientation	unit quaternion $\mathbf{q} \in \mathbb{H}$	3	unit normal $\mathbf{n} \in \mathbb{S}^2$	2
Extent	scale $\mathbf{s} \in \mathbb{R}_{>0}^3$	3	area $A \in \mathbb{R}_{>0}$	1
Velocity	—	0	$\mathbf{v} \in \mathbb{R}^3$	3
Appearance	SH coefficients, deg. $L = 3$	48	complex amplitude $\alpha \in \mathbb{C}$	2
Opacity/material	opacity $o \in [0, 1]$	1	impedance $\eta \in \mathbb{C}$	2
Kernel shape	Gaussian (from $\mathbf{q}, \mathbf{s}$)	0	fixed by the aperture	0
Compositing	alpha blending	—	coherent additive sum	—
Total / primitive		58		13

ℓ). It admits an $O(1)$ closed form per sample, with a Taylor expansion for numerical stability near the removable singularities at $\Delta \in N\mathbb{Z}$ where the denominator vanishes (Appendix A.1; analytic gradient and its singularity handling in Appendix A.2). The splat product uses the unit-peak variant $d_N(\Delta) = \kappa_N(\Delta)/N$, with $d_N(0) = 1$. For sensors that apply a tapering window before the DFT, d_N generalizes to the corresponding window's frequency response (Appendix A.3).

5.2.2 Coherent appearance model. Each surfel reradiates a complex amplitude with magnitude from cell area, foreshortening, Fresnel reflectivity, and free-space spreading, and phase from round-trip propagation. Surfels combine coherently at the sensor and are initialized from a random or preprocessed point cloud.

Amplitude factors. Cell area enters as $\sqrt{A_i}$ so that $|\text{splat}_i|^2$ contributes detected power proportional to A_i . Foreshortening $V_i \equiv V(\mathbf{n}_i; \mathbf{t}_x, \mathbf{r}_x) \in [0, 1]$ is the bistatic geometric attenuation

$$V_i = \max(\hat{\mathbf{r}}_{\text{tx},i} \cdot \mathbf{n}_i, 0) \max(\hat{\mathbf{r}}_{\text{rx},i} \cdot \mathbf{n}_i, 0), \quad (4)$$

where $\hat{\mathbf{r}}_{\text{tx},i}, \hat{\mathbf{r}}_{\text{rx},i}$ are the unit-vector directions from transmitter and receiver to surfel i (§5.3). This factor is continuous and sub-differentiable in $\mathbf{n}_i$, vanishing whenever the surfel is back-facing. Fresnel reflectivity $F_i \equiv F(\mathbf{n}_i, \theta_{\text{inc},i}, \eta_i)$ is the closed-form angle-dependent response at a boundary of wave impedance η_i . Let $\theta_{\text{inc},i}$ be the angle between the incident direction and $\mathbf{n}_i$, and let η_0 be the wave impedance of the surrounding medium. For transverse-electric polarization at a planar boundary,

$$F(\mathbf{n}_i, \theta_{\text{inc},i}, \eta_i) = \frac{\eta_i \cos \theta_{\text{inc},i} - \eta_0 \cos \theta_{t,i}}{\eta_i \cos \theta_{\text{inc},i} + \eta_0 \cos \theta_{t,i}}, \quad (5)$$

where the transmitted angle $\theta_{t,i}$ follows from Snell's law, which for non-magnetic media reads $\sin \theta_{t,i} / \sin \theta_{\text{inc},i} = \eta_i / \eta_0$. Intuitively, η_i fixes the impedance mismatch at the boundary and therefore the reflected fraction, while the two cosines make that fraction vary with incidence angle. A perfect conductor is the limit $|\eta_i| \rightarrow 0$, where $F \rightarrow -1$ at every angle. This is the regime of the metal targets in §7.7. The transverse-magnetic coefficient and the derivation are in Appendix A.7. Path-loss exponent $\gamma \in [1, 2]$ accounts for one-way or two-way free-space spreading.

5.3 Spectral Rendering

Eq. (2) leaves two components of the forward model unspecified: (i) a projection μ_i that supplies each surfel's Fourier coordinates along every axis, and (ii) a coherent compositing rule that sums the per-surfel splats into the spectral tensor.

5.3.1 Projection. The role of $\mu_{\ell,i}$ mirrors the affine projection Jacobian in 3DGS. Where 3DGS pushes a 3D covariance into a 2D image-space covariance, we push a surfel's 3D position and velocity into Fourier-axis offsets, one formula per axis ℓ . We instantiate the three formulas for the $L = 3$ running example below, with $\ell \in \{R, D, A\}$ for range, Doppler, and azimuth (Fig. 2).

Let $\mathbf{o}$ denote the sensor origin and $\mathbf{u}_x$ the array axis. For each surfel i , we compute range $R_i = \|\mathbf{p}_i - \mathbf{o}\|$, the line-of-sight direction $\hat{\mathbf{r}}_i = (\mathbf{p}_i - \mathbf{o})/R_i$, radial velocity $v_{r,i} = \langle \mathbf{v}_i, \hat{\mathbf{r}}_i \rangle$, and azimuth $\sin \theta_i = \langle \hat{\mathbf{r}}_i, \mathbf{u}_x \rangle$. The continuous Fourier coordinates are then

$$\mu_{R,i} = \frac{N_R}{F_s} \frac{2K}{c} R_i, \quad \mu_{D,i} = \frac{2N_D T_c}{\lambda} v_{r,i}, \quad \mu_{A,i} = N_A \frac{\delta_{\text{el}}}{\lambda} \sin \theta_i. \quad (6)$$

The three coefficients map (i) round-trip delay $2R_i/c$ to range bins, (ii) inter-pulse phase advance $4\pi v_{r,i} T_c / \lambda$ over N_D pulses to Doppler bins, and (iii) inter-element steering phase $2\pi \delta_{\text{el}} \sin \theta_i / \lambda$ over N_A elements to azimuth bins [Richards 2014]. Here N_R, N_D, N_A are the per-axis FFT lengths and λ the carrier wavelength. The remaining constants ($F_s, K, T_c, \delta_{\text{el}}$) are sensor-side acquisition parameters whose mapping is detailed in Appendix B.

Doppler effects. The radial component $v_{r,i}$ enters Eq. (6) as the Doppler-axis offset, and the same backward pass returns closed-form Jacobians in both $\mathbf{p}_i$ and $\mathbf{v}_i$ (Appendix A.2), without the temporal extension layer of 3DGS-style dynamic representations [Luiten et al. 2024; Yang et al. 2024]. The construction is exact under the standard coherent-sensing regime (stop-and-hop, constant radial velocity, negligible range walk, Doppler unambiguity). Validity bounds are in Appendix A.6.

5.3.2 Coherent compositing. The complex spectral tensor is Eq. (2) evaluated at the integer bin grid. For the $L = 3$ running example, $\hat{Y}[p, k, q] = \sum_{i=1}^S \text{splat}_i(p, k, q)$ and the observable power image is $\hat{H}[p, k, q] = |\hat{Y}[p, k, q]|^2$. Coherent signals superpose as complex phasors before the detector squares the result, matching the true physics and correctly modeling the constructive and destructive interference between reflectors separated by a few resolution cells. Compositing is purely additive (i.e., no depth order, no opacity), making the renderer purely data-parallel. When inter-surfel phase is not modeled, the renderer optionally composites magnitude-squared splats via the Fejér kernel (Appendix A.1).

6 Optimization

The Dirichlet kernel of §5.2.1 is faithful in the forward direction but oscillatory. Its sidelobes propagate into the loss surface, yielding a rugged, non-convex landscape on which first-order optimizers stall.

We collect the surfel parameters of §5 into a center matrix $C \in \mathbb{R}^{S \times 3}$ and a complex coefficient vector $\mathbf{b} = (\alpha_1, \dots, \alpha_S)^T \in \mathbb{C}^S$ that absorbs the per-splat amplitude product of Eq. (2). Let Ω index the measured spectral bins (so $|\Omega| = N_R N_D N_A$ in the 3D case). With $A(C) \in \mathbb{C}^{|\Omega| \times S}$ the matrix whose i -th column is the spectral

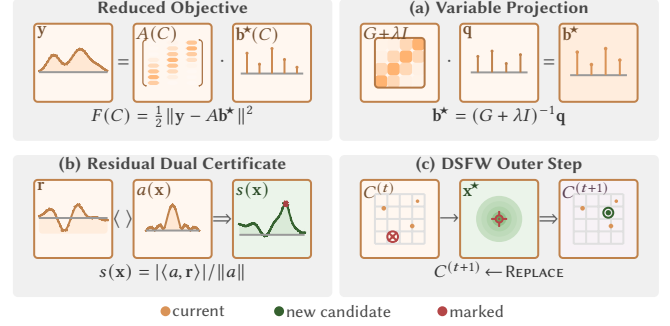


Fig. 3. **DSFW overview.** DSFW maintains a surfel set C by iterating three steps over the reduced objective $F(C)$ (top-left). (a) *Variable Projection* eliminates the linear coefficients $\mathbf{b}^*$ in closed form. (b) The *residual dual certificate* $s(\mathbf{x})$ peaks where adding a splat most reduces the residual, bypassing sidelobe minima. (c) The outer step *hard-replaces* the lowest-utility surfel with $\arg \max_{\mathbf{x}} s(\mathbf{x})$. Icons depict the complex spectral tensors $\mathbf{y}, \mathbf{r}$ as magnitude envelopes, $A(C)$ as a dictionary matrix with one Dirichlet splat per column, and $a(\mathbf{x})$ as the Dirichlet kernel of Fig. 2.

footprint of surfel i and $\mathbf{y} \in \mathbb{C}^{|\Omega|}$ the measured spectral image, the inverse problem is

$$\min_{C, \mathbf{b}} \mathcal{L}(C, \mathbf{b}) = \frac{1}{|\Omega|} \|A(C)\mathbf{b} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{b}\|_2^2. \quad (7)$$

For reporting and for generic first-order baselines we additionally use a log-power Charbonnier loss that handles the 40–60 dB dynamic range typical of coherent measurements [Richards 2014]:

$$\mathcal{L}_{\text{img}} = \frac{1}{|\Omega|} \sum_{k \in \Omega} \sqrt{(\log(\hat{H}_k + \epsilon) - \log(H_k^{\text{obs}} + \epsilon))^2 + \epsilon_c^2}, \quad (8)$$

with $\hat{H}_k = \| [A(C)\mathbf{b}]_k \|^2 + \sigma^2$ the rendered power (σ^2 the noise floor), $H_k^{\text{obs}} = |y_k|^2$ the measured power, Ω the set of evaluated bins, and constants $\epsilon = 10^{-6}$, $\epsilon_c = 10^{-3}$. Our Dirichlet Sliding-Frank-Wolfe (DSFW) algorithm minimizes Eq. (7) (closed-form in b). Eq. (8) is the outer evaluation criterion and the loss used for AdamW baselines.

6.1 Optimization Challenges

Three properties of $\mathcal{L}$ make Eq. (7) hard for plain gradient descent.

Sidelobe minima. A rectangular window gives $\kappa_N - 13$ dB sidelobes that place a gradient null at every sidelobe peak. A center initialized one main-lobe width from the true peak descends into the nearest sidelobe and stalls, and plain AdamW from a cold start fails on the benchmarks of §7.3 and §7.4.

Amplitude shrinkage. Eq. (7) couples centers and amplitudes: reducing $|\alpha_i|$ lowers the residual by trading positional accuracy for coefficient magnitude, leaving small-amplitude splats parked at sidelobe minima rather than at true scatterers.

Inter-atom coupling. Splats within a kernel width interfere coherently, coupling their center gradients. Adam's diagonal preconditioning misses the off-diagonal blocks of $\mathcal{H} = \text{Re}(J_c^H J_c) / |\Omega|$, so two coupled splats cannot make the joint move that resolves their interference.

DSFW addresses these three challenges in turn (Fig. 3). *Variable projection* (§6.2) eliminates $\mathbf{b}$ analytically and leaves only the centers C to optimize, removing amplitude shrinkage from the search. The *residual dual certificate* (§6.3) globally scores how much a new splat at each grid point would reduce the residual, allowing the search to move beyond local sidelobe minima. DSFW (§6.4) combines both mechanisms by replacing the lowest-utility splat with the certificate maximizer and periodically taking a Gauss–Newton step to account for inter-splat coupling.

6.2 Variable Projection

Our forward model is linear in the splat amplitudes $\mathbf{b}$, since surfels superpose coherently as a sum of complex amplitudes (Eq. (2)). This linearity enables both the closed-form coefficient solve developed in this subsection and the SFW outer loop of §6.4. 3DGS lacks this property because alpha compositing renders the image as a non-linear function of the per-Gaussian opacity and color [Kerbl et al. 2023], which precludes a closed-form solve. The obstruction lies in the compositing rule rather than the kernel. Gaussian variants with a linear sum-of-splats forward model (e.g., GSRF [Yang et al. 2025], WRF-GS+ [Wen et al. 2026]) admit the same machinery, which we evaluate as GS+SFW, the Gaussian cell of the 2×2 kernel-optimizer ablation in §7.6.

For fixed centers C , Eq. (7) is a complex ridge regression with closed-form minimizer

$$\begin{aligned} \mathbf{b}^*(C) &= (G(C) + \lambda I)^{-1} q(C), \\ G(C) &= \frac{1}{|\Omega|} A(C)^H A(C), \quad q(C) = \frac{1}{|\Omega|} A(C)^H \mathbf{y}. \end{aligned} \quad (9)$$

Substituting back yields the variable-projected reduced objective [Golub and Pereyra 1973; O’Leary and Rust 2013],

$$F(C) = \mathcal{L}(C, \mathbf{b}^*(C)), \quad (10)$$

which depends only on the center variables. Eliminating $\mathbf{b}$ removes amplitude shrinkage as a degree of freedom and yields a better-conditioned nonlinear search at each step.

6.3 Residual Dual Certificates

With $\mathbf{b}^*(C)$ fixed by Eq. (9), the *residual dual certificate* [Denoyelle et al. 2020; Jaggi 2013] provides a global signal for where to move centers next:

$$s(\mathbf{x}) = \left| \langle \mathbf{a}(\mathbf{x}), \mathbf{y} - A(C)\mathbf{b}^*(C) \rangle \right| / \|\mathbf{a}(\mathbf{x})\|, \quad (11)$$

where $\mathbf{a}(\mathbf{x})$ is the spectral footprint a hypothetical surfel placed at position $\mathbf{x}$ would produce. Intuitively, peaks of s identify locations where adding a splat most reduces the residual. Equivalently, s is the dual variable of the off-the-grid sparse recovery problem of which Eq. (7) is the primal.

Global signal vs. local gradient. Unlike a local gradient, $s(\mathbf{x})$ is evaluated everywhere on a candidate grid, and therefore bypasses the sidelobe minima of §6.1. Its cost is dominated by candidate-voxel correlations and grows quickly with grid resolution. A fused kernel reduces certificate evaluation from a per-experiment to a per-step cost (Appendix C.4).

6.4 Dirichlet Sliding Frank–Wolfe

The certificate of §6.3 drives a global, sparse-recovery update that bypasses the sidelobe landscape entirely. DSFW is *inspired* by the sliding Frank–Wolfe (SFW) template [Denoyelle et al. 2020] for off-the-grid spike recovery on the BLASSO problem, but deviates in one important way. Classical SFW *adds* a new splat at $\arg \max_{\mathbf{x}} s(\mathbf{x})$ and then merges nearby splats, growing the support monotonically. DSFW instead operates at a fixed budget of S splats and *replaces* the lowest-utility splat with the certificate’s argmax. The fixed budget reflects the physics of coherent imaging. Aperture and bandwidth bound the number of resolvable scatterers, and unbounded growth would fit sidelobes and noise as spurious splats. We set S slightly above a back-projection sparsity estimate (§7.4). The trade-off is that the classical SFW guarantees [Denoyelle et al. 2020] no longer apply, since they require both the BLASSO non-degenerate-source condition and unconstrained splat addition.

Each DSFW outer step performs four operations. (i) Solve Eq. (9) for $\mathbf{b}^*(C)$. (ii) Evaluate the certificate Eq. (11) on a coarse candidate grid. (iii) *Hard-replace* the surfel of lowest utility with $\arg \max_{\mathbf{x}} s(\mathbf{x})$, where utility ranks splats by amplitude and by a coherent-cancellation efficiency that down-weights splats already representable by their neighbors. (iv) Optionally refine with a Levenberg–Marquardt sub-step on the spatial residual. Algorithm 1 specifies the loop, schedules, and optional steps.

Utility score. Step (iii) must select a splat for replacement. Amplitude alone is insufficient because neighboring splats may reproduce a large-amplitude splat without increasing the residual, whereas an isolated small-amplitude splat may be indispensable. We therefore score each surfel j by

$$u_j = \frac{\frac{1}{2} \frac{|b_j^*|}{\max_k |b_k^*|} + \frac{1}{2} \text{eff}_j}{\|a_j\|_2}, \quad \text{eff}_j = \frac{\|P_{A_{\setminus j}}^\perp a_j\|_2}{\|a_j\|_2} \in [0, 1], \quad (12)$$

where $\mathbf{a}_j$ is the spectral footprint of surfel j and $P_{A_{\setminus j}}^\perp$ projects onto the orthogonal complement of $\text{span}\{\mathbf{a}_k\}_{k \neq j}$. Intuitively, the first term measures relative amplitude. The *coherent-cancellation efficiency* measures how much of the surfel’s footprint lies outside the span of the other splats. A value of eff_j near zero means that the other splats already span this footprint. Replacing such a surfel changes the residual little and reallocates its center to the certificate maximum.

Initialization. Cold-start runs draw the S centers uniformly at random from the candidate volume, which is the box spanned by the sensor’s range gate and its azimuth field of view, and give every center the same initial amplitude. Where a modality supplies a cheap signal-processing estimate, the centers may instead be seeded from back-projection peaks or from FFT-domain thresholding, as the multi-view fit of §7.5 and the printed-object fit of §7.7 do. Per-experiment candidate grids and budgets are listed in Appendix B.4.

DSFW carries no convergence guarantee, but recovers ground-truth geometry from random initialization on every benchmark in §7.3.

Algorithm 1 DSFW outer loop.

Require: measurement y , initial centers C_0 , kernel order N , ridge weight λ , iteration budget T , low-resolution schedule $\mathcal{T}_{lr}$, candidate grid $\mathcal{X}$

Ensure: refined centers C , coefficients b^*

```

1:  $C \leftarrow C_0$ 
2: for  $t = 1, \dots, T$  do
3:    $b^* \leftarrow \arg \min_b \|y - A(C)b\|^2 + \lambda \|b\|^2$   $\triangleright$  VarPro, Eq. (9)
4:    $s(x) \leftarrow$  certificate on  $\mathcal{X}$  from residual  $r = y - A(C)b^*$   $\triangleright$ 
   Eq. (11)
5:    $u_j \leftarrow \frac{1}{2} \left| b_j^* \right| / \max_k |b_k^*| + \frac{1}{2} \text{eff}_j$  for each  $j$   $\triangleright$  Eq. (12)
6:    $j^\dagger \leftarrow \arg \min_j u_j$ 
7:    $c_{j^\dagger} \leftarrow \arg \max_{x \in \mathcal{X}} s(x)$   $\triangleright$  hard replacement
8:   (optional) refine  $C$  by a Levenberg–Marquardt sub-step on
   the spatial residual
9:   if  $t \in \mathcal{T}_{lr}$  then
10:    propose  $\Delta C$  from a coupled low-resolution LM step and
    accept it if the full-resolution loss decreases
11:   end if
12: end for
13: return  $C, b^*$ 

```

Furthermore, DSFW augments the outer step with a periodic Gauss–Newton correction and four small stabilizers (window matching, kernel-continuation curriculum, pairwise repulsion, coarse-to-fine schedule on N). The refinements and experiment settings are described in Appendix B. With these corrections, DSFW achieves the lowest final loss among the optimizers we evaluate (§7.3).

7 Results and Application

We evaluate Dirichlet Splatting on controlled synthetic benchmarks (§§7.2–7.6) and on a real hardware application (§7.7). Synthetic experiments use a DFT setup with $N = 256$ range bins, $M = 64$ slow-time bins, and $A = 8$ array elements unless noted. Ground-truth signal synthesis and per-experiment protocols are in Appendix C. AdamW hyperparameters and DSFW defaults are listed in Appendix B.4.

7.1 Baselines

We compare against published splatters and classical/physics references. Internal ablations are deferred to §7.6. **3DGS** [Kerbl et al. 2023] is the canonical Gaussian splatter; **GSRF** [Yang et al. 2025] and **WRF-GS+** [Wen et al. 2026] are the state of the art in coherent Gaussian splatting for radio-frequency imaging; **back-projection** is the matched-filter coherent adjoint; and **WaveformAD** differentiates through the full synthesis $\rightarrow \text{FFT} \rightarrow |\cdot|^2$ pipeline with autograd, serving as a physics oracle that bounds attainable accuracy. All comparisons share scene parameters, learning rate, truncation radius, splat budget, and loss function unless noted. §7.3 fits targets rendered by the same operator, which isolates the solver from measurement-model error. §7.5 shares that operator but scores geometry against a closed mesh. The other experiments score against independent targets, namely the analytic FFT of a point reflector

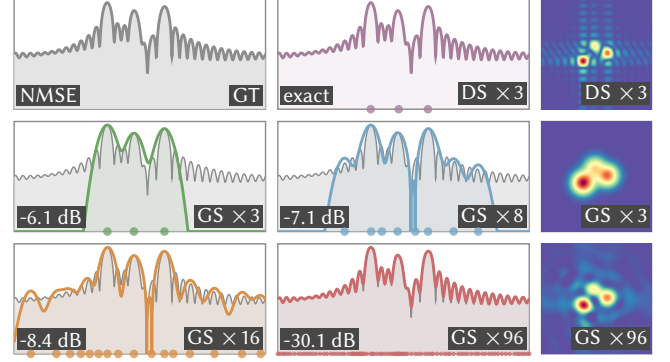


Fig. 4. Forward model fidelity. Three Dirichlet primitives reproduce the sidelobe-rich coherent range profile of three complex reflectors exactly (top), whereas a fixed-width Gaussian representation still needs 96 splats (32× more) to reach -30.1 dB magnitude normalized mean squared error (NMSE). Right: the corresponding 2D Dirichlet interference pattern in linear scale.

(§7.2), raw range profiles from a standard FMCW chain (§7.4), and measured captures (§7.7).

7.2 Forward Model Validation

The first experiment tests the physical claim of §5.2.1. The Dirichlet kernel is the PSF of the finite-window DFT. A Gaussian splat is at best a main-lobe approximation that discards the sidelobe energy carrying coherent phase information.

Setup. We place a single complex reflector at fractional sub-bin offsets $\delta \in \{0.0, 0.1, \dots, 0.5\}$ on a finite-window grid and compare each method’s rendered response against the corresponding FFT-domain ground truth. We sweep over rectangular, Hann, and Kaiser ($\beta = 14$) windows to verify that the windowed kernel form (Appendix A.3) tracks arbitrary tapers, and over 1D, 2D, and 3D grids to verify separability. The experiment isolates the kernel choice cleanly, with no optimization, no scene complexity, and only forward evaluation against an exact analytic reference.

Results. Fig. 4 shows the 1D comparison on a three-reflector range profile with unequal amplitudes. Three Dirichlet primitives reproduce the sidelobe-rich coherent waveform directly, while a Gaussian representation must spend many splats to approximate the same dB profile. The right column shows the corresponding 2D Dirichlet interference pattern on the same three-reflector configuration in linear scale. With only three Gaussian splats, the fit captures the main peaks but misses the sidelobe structure. Increasing the budget improves the envelope only gradually, implying the Gaussian mismatch is a representation issue, not just an optimizer issue. Quantitatively, the three Dirichlet primitives are exact up to the numerical floor (-120 dB NMSE), while Gaussian controls improve only slowly with splat count: 3, 8, and 16 splats remain between -6.1 and -8.4 dB NMSE, and even 96 splats reach only -30.1 dB.

Doppler validation. The same kernel extends to the Doppler axis. Each surfel carries a constant velocity that the rendering operator

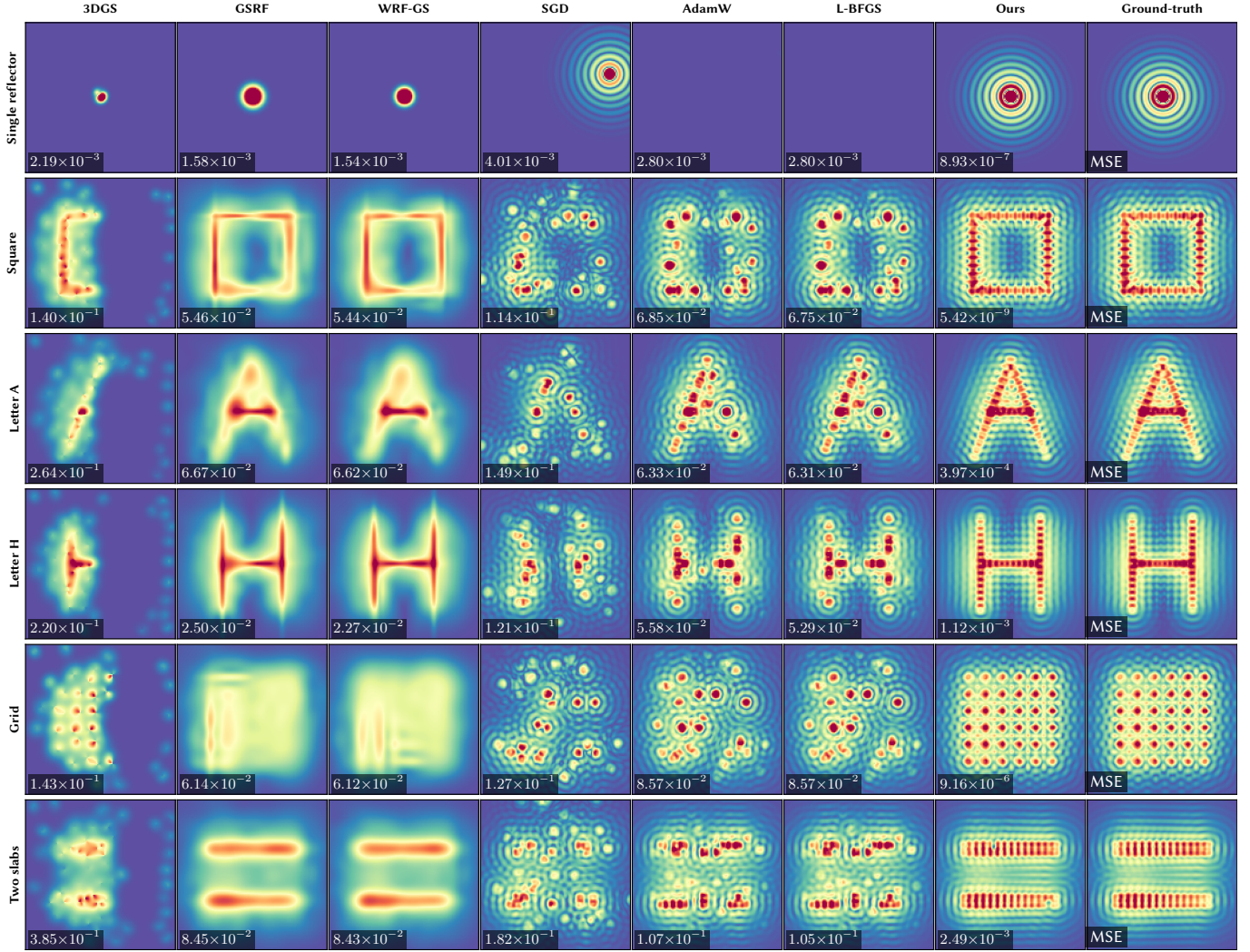


Fig. 5. **Inverse fitting in signal space (§7.3).** Methods fit Dirichlet splats directly to a 2D Fourier-space target tensor, with no aperture and no measurement model. Rows are interpretable signal-space targets (single reflector, square, letters A/H, grid, two slabs). Columns show a 3DGS-style no-phase baseline, GSRF, WRF-GS+, SGD, AdamW, an AdamW+L-BFGS quasi-Newton control, DSW (ours), and the ground-truth scene, the noise-free Dirichlet-splat scene the targets were rendered from. Each panel shows the magnitude of the complex spectral tensor $\hat{Y}(f)$ of Eq. (2) over the two Fourier axes of the target. All methods share the same target tensor, cold initialization, splat budget, and loss function. Corresponding complex spatial MSE numbers are reported in §7.3. Only DSW recovers the sidelobe field. The Gaussian columns reconstruct the silhouette alone.

of §5.3 projects into the slow-time spectrum. We validate this projection on a $96 \times 64 \times 8$ frequency-modulated continuous-wave (FMCW) cube containing three reflectors moving at -6.0 , 0.0 , and 7.5 m/s, with at most 0.051 range bins of range walk over the dwell. The joint position–velocity Dirichlet model matches the finite-window FFT oracle to the numerical floor (-120 dB complex NMSE) and picks the correct velocity for every reflector from a 0.05 m/s search grid. Clamping velocity to zero on the same scene degrades the fit to -1.11 dB, which isolates the contribution of the Doppler parameter. The experiment isolates the constant-velocity parameterization. Dynamic-scene reconstruction remains outside its scope. Acquisition constants and the search protocol are in Appendix C.1.

7.3 Inverse Fitting in Signal Space

This experiment tests the optimization claim of §6. We scope the experiment to the *signal level*, fitting Dirichlet splats directly to a 2D Fourier-space target tensor with no aperture and no measurement model. The setup isolates the solver from any imaging error. The full inverse imaging problem (raw measurements $\rightarrow$ geometry) is studied in §7.4. We compare DSW against Dirichlet-AdamW and stochastic gradient descent (SGD) optimizer controls and against GSRF [Yang et al. 2025] and WRF-GS+ [Wen et al. 2026] (§7.1). Optimizer-only and kernel-only ablations are deferred to §7.6.

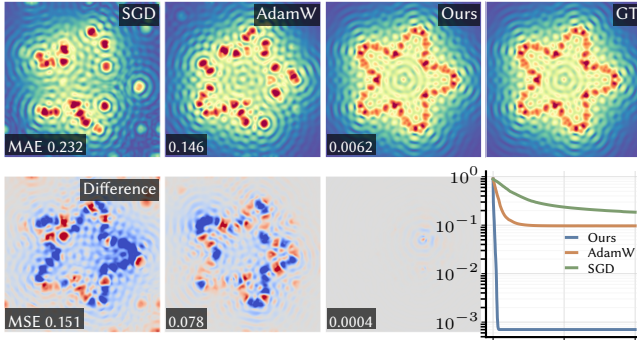


Fig. 6. **Cold-start convergence on a structured-star scene.** DSW (ours) recovers the reflector field, while Dirichlet-AdamW stalls in sidelobe minima and SGD remains farther from the ground-truth scene under the same iteration budget. Top: magnitude reconstructions, labelled with mean absolute error (MAE). Bottom: complex spatial mean squared error (MSE) trajectories and signed magnitude-difference maps against the ground-truth scene, labelled with MSE.

Setup. All targets in this section are synthetic 2D Fourier-space tensors of known Dirichlet-splat scenes, with methods fitting Dirichlet splats directly to them. The convergence and basin studies use sparse-splat scenes: the random $K = 10$ 2D scene, the structured $K = 50$ star scene, and the measured $K = 5$ basin sweep. The result grid (Fig. 5) displays the same signal-space pipeline on interpretable 128×128 -bin spectral targets. Splats are cold-initialized uniformly at random with no scene prior, and splat budget and loss follow §7.1.

Results. Fig. 6 shows a representative cold-start case. DSW recovers the reflector field and drives down the complex spatial loss, Dirichlet-AdamW lowers it into a mismatched sidelobe arrangement, and SGD lags farther under the same budget. Across a 33×33 grid of cold-start offsets over 8 bins, DSW recovers all reflector positions to within 0.1 bin on 77.4% of trials, against 0.07% for Dirichlet-AdamW, which stays trapped in the sidelobe neighborhoods of §6.1 (heat map in Appendix C.2). Fig. 5 extends the pattern to recognizable Fourier-domain targets. Over the 7-scene benchmark, DSW drives mean complex MSE from 3.32×10^{-1} at initialization to 6.71×10^{-4} . By contrast, Dirichlet-AdamW and SGD end at 6.06×10^{-2} and 1.12×10^{-1} , respectively. The two-stage AdamW+L-BFGS quasi-Newton control reaches 5.97×10^{-2} after a 300-step AdamW warmup and 50 L-BFGS polish iterations, giving only a small local refinement over AdamW and remaining two orders of magnitude above DSW. The 3DGS-style no-phase baseline ends at 1.92×10^{-1} . Its panels are visibly suppressed on one side of several letters and slabs because it renders only a positive real Gaussian radiance field. On those halves the complex Dirichlet target has strong imaginary components or negative real projections, so the least-squares optimum drives many splats toward low opacity rather than matching the displayed magnitude. GSRF [Yang et al. 2025] and WRF-GS+ [Wen et al. 2026] end at 4.43×10^{-2} and 4.38×10^{-2} . The Gaussian baselines reconstruct smooth silhouettes but do not match the Dirichlet sidelobe field. Cross-scene PSNR, magnitude NMSE, and SSIM are reported in Appendix C.2.

7.4 Inverse Imaging from Raw Signals

§7.3 fits splats to signal-space targets and isolates the solver from the imaging measurement model. We now evaluate the full inverse imaging problem: the input is a raw complex multi-aperture scan of a target with known material response. The output is a dense geometric support estimate, scored against ground-truth occupancy.

Setup. The inverse-imaging grid uses 7 dense targets shown in Fig. 7. The simulator renders each target into raw complex range profiles using a 96×96 aperture, 4 GHz bandwidth, and a 56×56 candidate grid. The simulator follows the standard FMCW signal chain, synthesizing the chirp echo from the target occupancy and material response, dechirping it against the transmitted waveform, and taking the FFT of the beat signal. Our model fits a Dirichlet support field directly to these raw profiles, with eikonal, total variation, and binarization regularization. Among baselines from §7.1, 3DGS, GSRF, and WRF-GS+ are supervised on the back-projection image. **GS+SFW**, our coherent Gaussian baseline sharing DSW’s forward model and SFW solver but with a Gaussian kernel, is supervised on the raw range profiles. All methods decode into a common occupancy space.

Results. Reconstruction grids and per-scene support-domain NMSE are shown in Fig. 7. Ours reaches the lowest mean support NMSE across 7 targets (0.136), against 0.294 (GS+SFW), 0.387 (back-projection), 0.389 (3DGS), 0.391 (WRF-GS+), and 0.406 (image-domain GSRF). Raw-channel fitting improves GSRF without closing the gap: its raw NMSE stays an order of magnitude higher (9.12×10^{-3} vs. 7.74×10^{-4} for ours), with wider support-NMSE gaps on the grid (0.943 vs. 0.219) and scissors (0.395 vs. 0.296). GS+SFW is competitive on simple stroke-like shapes (e.g., letters A/T/H), which is consistent with broad Gaussian fields being effective for smooth connected support but less reliable for separated or thin multi-component geometry. Cross-scene support NMSE, PSNR, SSIM, IoU, and Dice are reported in Appendix C.2.

7.5 3D Reconstruction

We evaluate Dirichlet Splatting on a 3D target, fitting one global surfel set jointly across apertures that view it from different directions.

Setup. The target is a 60 mm Stanford bunny. We sample its closed mesh into 60,000 oriented surfels and render each view with the forward model of §5. Each view uses a 128×128 virtual aperture at 2.5 mm pitch and 350 mm standoff. The sensor uses a 120 GHz carrier, 15 GHz bandwidth, and a 320 mm aperture, yielding 1.4 mm resolution across the target. The fit runs against the concatenated raw complex range profiles of all views under complex mean squared error. We concatenate the raw profiles of the V views into a single tensor. The foreshortening factor V_i of Eq. (4) provides per-view visibility, so multi-view fitting reuses the forward model without a separate visibility solver. We seed surfel centers from peaks of a joint back-projection volume using the optional initializer of §6.4.

Metrics. We measure (i) the median distance from each recovered surfel to the ground-truth surface, (ii) F-score and recall at 2 mm, (iii) symmetric Chamfer distance, (iv) NMSE on the raw profiles, and

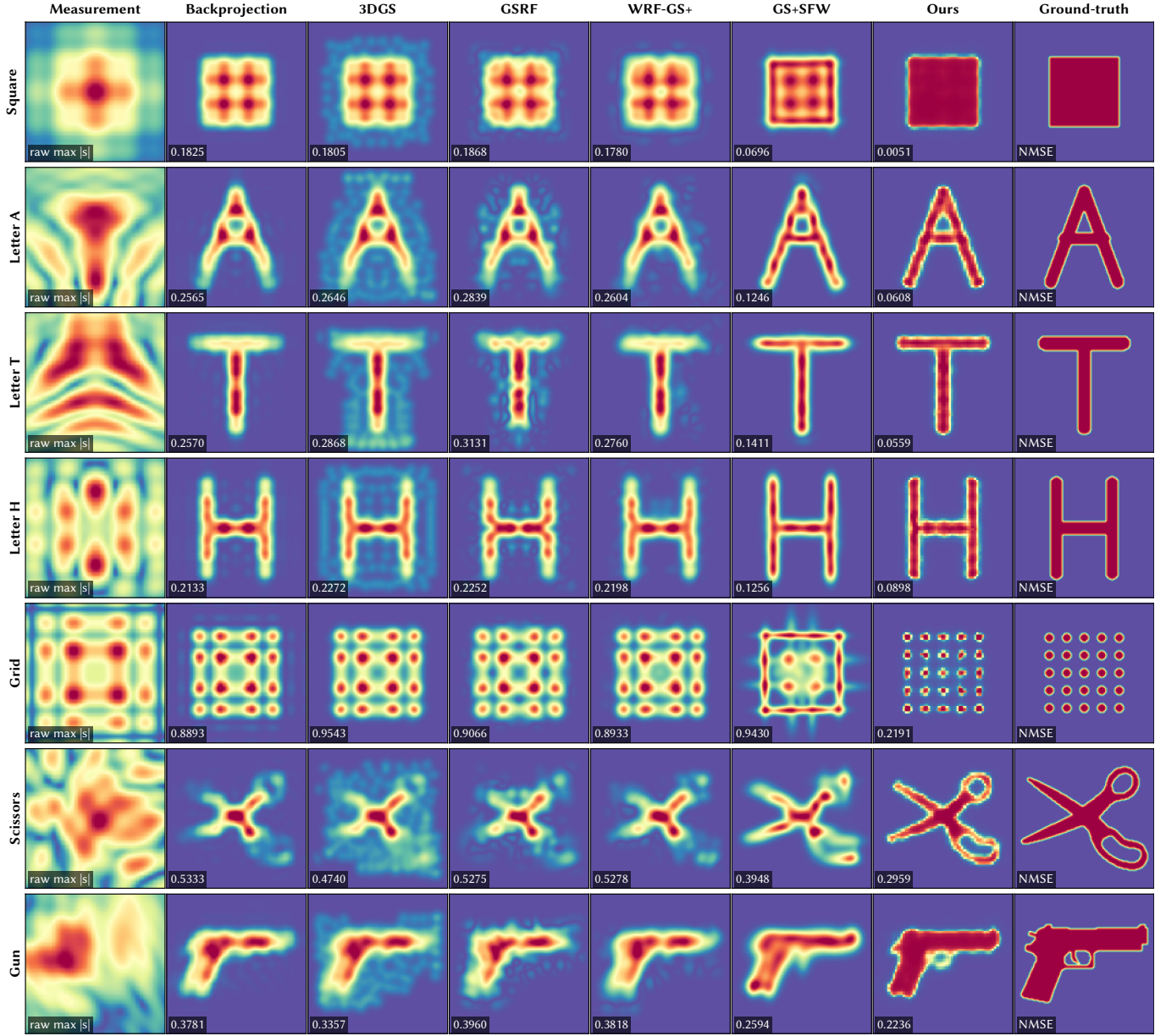


Fig. 7. **Inverse imaging from raw signals (§7.4).** Each method fits raw complex multi-aperture range profiles of a target with known material response and decodes a dense geometric support estimate; reconstruction-panel labels report support-domain NMSE against the ground-truth support, meaning the binary occupancy map the simulator rendered from, with lower values better. Ours and GS+SFW are supervised on the raw range profiles and the remaining learned baselines on the back-projection image decoded from those same profiles. All learned methods share the cold initialization and the splat budget S . Rows are seven dense planar targets: a square, letters A/T/H, a reflector grid, a scissors silhouette, and a binarized pistol image. The leftmost column, *Measurement*, is a measurement-domain sanity check (the simulator output collapsed to the per-aperture maximum range-profile magnitude), plotted in aperture coordinates rather than object coordinates, and is excluded from the support-domain NMSE labels. The remaining columns show back-projection, 3DGS, GSRF, WRF-GS+, GS+SFW, the raw-profile Dirichlet reconstruction (ours), and the ground-truth support. Only the raw-profile Dirichlet column keeps thin and separated components intact. The Gaussian columns merge them into a single blob.

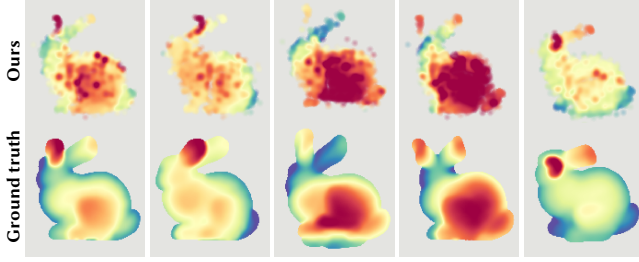


Fig. 8. **3D reconstruction of the Stanford bunny (§7.5).** We jointly fit one global surfel set to the raw complex range profiles of five apertures clustered within 25° of the front axis and render it from the corresponding view directions. Color encodes depth along the view axis using the colormap of Figs. 5 and 7. Each column uses the ground-truth depth range for direct comparison. Our method recovers the silhouette and ears from all five views.

(v) the fraction of the ground-truth surface that at least one view observes. Scores restricted to that fraction are marked *recoverable*. For a matched-budget comparison, we apply the same non-maximum suppression to the brightest back-projection peaks and retain the same number of points as our surviving surfels. We evaluate both point sets with the same metrics. Multi-view numbers are the mean $\pm$ half-range over three seeds.

Results. Fig. 8 shows the reconstruction from the five fitted view directions. Three orthogonal views increase the observed fraction from 0.35 to 0.596. They reach a full-surface recall@2mm of 0.439 ± 0.011 , a median surface distance of 1.67 ± 0.07 mm, and a raw-profile NMSE of 0.203 ± 0.012 . No single view reaches this recall. Three views clustered within 25° observe a more overlapping surface region. Their raw-profile NMSE decreases to 0.069 ± 0.012 , and their recoverable F-score@2mm increases to 0.647 ± 0.047 , at the cost of lower surface coverage. Five clustered views reduce the median surface distance to 1.49 ± 0.21 mm. At a matched point budget on the three-view cluster, our reconstruction reaches a median surface distance of 1.51 mm, compared with 1.92 mm for back-projection. Its recoverable recall@2mm is 0.750 versus 0.530, and its symmetric Chamfer distance is 3.87 mm versus 5.17 mm. Back-projection superposes energy without enforcing cross-view consistency. Defocus caustics from individual views therefore remain off the surface. It also produces an energy volume rather than a model of the raw signal.

7.6 Ablations

We isolate the two main contributions, the Dirichlet kernel and the DSFW solver, through a 2×2 factorial over kernel (Gaussian, Dirichlet) and optimizer (AdamW, SFW), evaluated on a five-reflector cold-start benchmark over 10 seeds at a shared splat budget and initialization, counting a trial successful when every matched center lands within 0.1 bin of the ground-truth scene. The proposed pair succeeds on 100.0% of trials at a final complex spatial MSE of 5.32×10^{-9} . **(i)** Holding the optimizer at SFW and replacing the Dirichlet kernel with a Gaussian drops the success rate to 10.0% and raises the MSE from 5.32×10^{-9} to 6.71×10^{-3} , a gap of more than $10^6 \times$, and reintroduces the ring artifacts of §7.4. Certificate-driven global search alone cannot compensate for a kernel that discards sidelobe energy. **(ii)** Holding the kernel at Dirichlet and replacing

SFW with AdamW reduces the cold-start success rate from 100.0% to 0.0% and raises the MSE to 1.18×10^{-2} , confirming that the correct kernel is necessary but not sufficient on the rugged landscape of §6.1. Replacing both components with the standard Gaussian–AdamW pairing yields a 0.0% success rate and an MSE of 8.54×10^{-3} . Appendix C.3 reports center RMSE.

7.7 Application: Sub-THz Imaging

Sub-THz imaging penetrates optically opaque materials at quasi-optical resolution [Willis and Wilson 2013] and is exactly the finite-aperture-diffraction regime that motivated the Dirichlet kernel. We evaluate the full pipeline on real sub-THz measurements.

Setup. We reuse the raw-signal fitting protocol of §7.4 on measurements from a custom-built sub-THz software-defined radio (120 GHz center, 1.23 GHz bandwidth, 8192 samples per chirp). We evaluate three metal letter scans: T, H, and Z. We call a measured capture a *raw scan*, as distinct from the synthetic ground-truth support of §7.4. The T scan uses a 37×49 2.5 mm raster; H and Z use 37×37 rasters. All use 50 coherent captures per aperture, a 20° Tx/Rx bistatic angle, and a 15 cm focus distance. We compare a region-of-interest power projection (ROI Power), bistatic back-projection, and GS+SFW from §7.4. Without a pixel-accurate ground-truth mask, we report label-free contrast, support compactness, and raw-profile fit residuals.

Results. Fig. 9 and the three-scan baseline audit show distinct failure modes. ROI Power is a strong aperture-coordinate energy map (average p_{99}/p_{50} contrast 44.8, robust SNR 30.5), yet not a geometric inversion: at a 35%-peak threshold it occupies only 9.6% of pixels across 10.3 disjoint components. Back-projection brings in the bistatic geometry at the cost of diffraction sidelobes: with p_{99}/p_{50} falling to 3.71 and 35%-peak support spreading over 26.2% of pixels in 69.0 components. GS+SFW reaches mean normalized loss 1.70 (1.95/1.42/1.72 for T/H/Z) on a compact support (10.7% area, 8.3 components) of broad blobs that miss the measured finite-window sidelobe structure.

3D printed object. To evaluate the same pipeline on a real object with curvature, we 3D-print the target of §7.5 at 120 mm, coat it with conductive ink, and stand it on a narrow mount inside an absorber-lined enclosure (Fig. 10a). The same transceiver scans the target confocally over a 121×121 aperture at 1.25 mm pitch across a 150 mm square window. We sweep 4.8 GHz at a 150 mm focus and coherently average 32 chirps per aperture position. One range cell spans 31.25 mm. We register eight rasters and retain five after discarding three with broadband interference. Two rasters use 2.5 mm pitch. We also remove aperture positions dominated by the mount return, which is stronger than the body return in the same range cell. The target is a continuous surface, so we fit a dense field of 4991 Dirichlet atoms whose centers stay on a fixed lattice, with the same kernel, forward model, and loss. The resulting fit reaches 0.482 NMSE against the stacked raw complex range profiles.

Fig. 10b shows the reconstruction. At a 35% threshold over the full scan window, 93% of the recovered support lies in its largest connected component. In comparison, the measured range-gated energy and back-projection contain 34 and 6 components, respectively. Rendering the fitted field with the measurement operator yields a

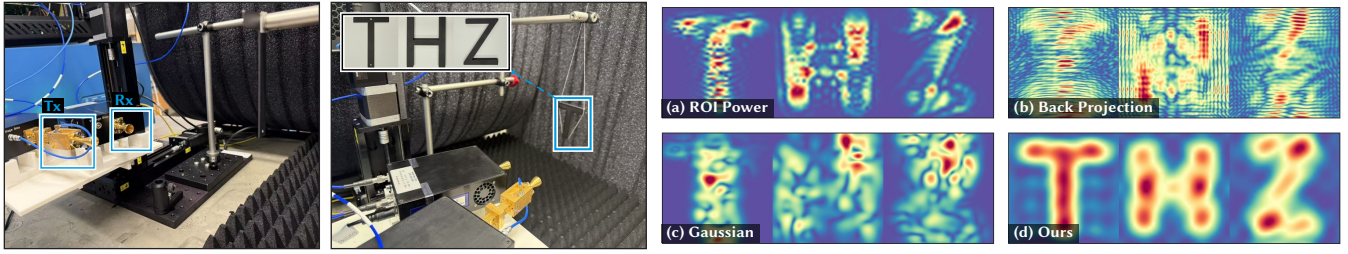


Fig. 9. **Real sub-THz reconstruction.** The imaged scene consists of three metal letters T, H, and Z suspended in front of a sub-THz transceiver pair (left: setup photograph and a zoomed view of the hanging target). Panels (a)–(d) compare four reconstructions on the same measurement: (a) region of interest power, (b) back projection, (c) Gaussian, and (d) ours. All four methods decode the same raw scan on a common pixel grid and color scale. Our reconstruction preserves the three separated letters while reproducing the measured finite-window sidelobes.

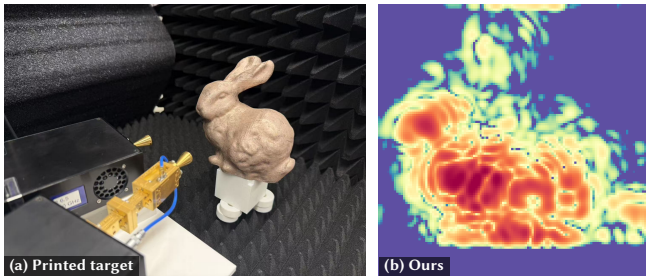


Fig. 10. **Reconstruction of a 3D-printed object (\$7.7).** (a) The metallized target inside the absorber-lined enclosure. (b) Our reconstruction over the full scan window, rendered with the measurement operator and displayed using the colormap of Fig. 9 with gamma correction.

dB-domain correlation of 0.74 with the raw-scan image over the object. Two instrument limits bound the result: the transceiver is modeled as an ideal confocal pair, and the aperture resolves 2.5 mm across the target against a range cell of 31.25 mm.

7.8 Wavefront Sensing

A finite pupil produces the same response in visible light, allowing the same primitive to represent a wavefront. We simulate off-axis holograms of four coherent beacons observed through a continuous 5 mm square pupil at 532 nm under Poisson and read noise. After sideband demodulation, each method fits the beacon positions, complex amplitudes, and twelve shared Zernike modes to the recovered field using the same initialization and parameter count. We generate the reference field by Gauss–Legendre quadrature over the continuous aperture, so the Dirichlet atoms approximate rather than generate the measurements. Decreasing the beacon separation creates overlapping sidelobe lattices and tests whether the methods can resolve them. At one main-lobe FWHM of separation, the Dirichlet fit reaches 0.021 waves of residual RMS wavefront error over fifty trials, compared with 0.051 for a diffraction-matched Gaussian and 0.110 for back-projection. Across the tested separations, it remains within two standard errors of an oracle using the continuous-aperture response. At separations down to half an FWHM, only the Dirichlet fit and the oracle remain below $\lambda/30$. The Gaussian fit requires four times this separation to meet the same threshold. At one FWHM, the Gaussian and Dirichlet fits fail in 52% and 16% of trials, respectively.

8 Limitations and Future Work

Indirect propagation. Multipath and inter-object scattering are modeled as virtual surfels at the total path length, which the renderer reproduces in intensity. The optimizer treats them as independent reflectors and lacks the method-of-images constraint tying them to the primary geometry, so the recovered support contains ghosts at correct path lengths but incorrect physical locations.

Occlusion. The visibility term in Eq. (4) handles back-facing surfels but not opaque self-occlusion between surfels. In partially transmissive modalities (THz, mmWave, synthetic-aperture sonar), occluded contributions cancel through coherent phase rather than a discrete visibility test, and the omission is benign for our experiments.

Hardware calibration. The forward model assumes known aperture positions and a calibrated transceiver response. The aperture extent and sampling pitch determine the kernel. Position errors produce wavelength-scale phase errors. An uncalibrated antenna pattern produces per-aperture amplitude errors that the fit absorbs into α_i . Bandwidth determines the range resolution. Joint calibration and reconstruction is a promising direction for future work.

DSFW failure modes. DSFW has two calibration-related failure modes. (i) Gross mis-scaling of A_i pushes the inner ridge to absorb the error into $|b_i|$, demoting correctly placed surfels in the utility ranking. (ii) A mis-calibrated noise floor σ^2 shifts the certificate magnitude, causing DSFW to either fit noise as splats (under-calibrated) or stop short of the true sparsity (over-calibrated).

9 Conclusion

We present *Dirichlet Splatting*, a differentiable renderer for inverse problems with wave-based coherent sensors. It replaces the Gaussian footprint inherited from 3D Gaussian Splatting with the analytic Dirichlet point-spread function and represents each primitive as a surfel whose complex amplitude follows wave propagation. Dirichlet Sliding Frank–Wolfe addresses the rugged sidelobe landscape induced by this oscillatory kernel. We believe Dirichlet Splatting is a step toward physically grounded differentiable rendering for radar, sonar, and other coherent sensing modalities.

Acknowledgments

We appreciate the insightful comments and feedback from the anonymous reviewers and shepherd. This work is partially supported by the NSF grants 2238839, 2341952, 2403124, and 2408393, and gifts from Adobe, Google, Activision, and Meshy.

References

- Stefan Auer, Richard Bamler, and Peter Reinartz. 2016. RaySAR – 3D SAR Simulator: Now Open Source. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (Beijing, China). IEEE, New York, NY, USA, 6730–6733. doi:10.1109/IGARSS.2016.7730757
- Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. 2022. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (New Orleans, LA, USA). IEEE, New York, NY, USA, 5460–5469. doi:10.1109/CVPR52688.2022.00539
- David Borts, Erich Liang, Tim Broedermann, Andrea Ramazzina, Stefanie Walz, Edoardo Palladin, Jipeng Sun, David Brueggemann, Christos Sakaridis, Luc Van Gool, Mario Bijelic, and Felix Heide. 2024. Radar Fields: Frequency-Space Neural Scene Representations for FMCW Radar. In *ACM SIGGRAPH 2024 Conference Papers* (Denver, CO, USA) (*SIGGRAPH '24*). Association for Computing Machinery, New York, NY, USA, Article 130, 10 pages. doi:10.1145/3641519.3657510
- Kristian Bredies and Hanna Katriina Pikkarainen. 2013. Inverse Problems in Spaces of Measures. *ESAIM: Control Optim. Calc. Var.* 19, 1 (2013), 190–218. doi:10.1051/cocv/2011205
- Müjdat Çetin and William C. Karl. 2001. Feature-Enhanced Synthetic Aperture Radar Image Formation Based on Nonquadratic Regularization. *IEEE Trans. Image Process.* 10, 4 (April 2001), 623–631. doi:10.1109/83.913596
- Brian Chao, Jacqueline Yang, Suyeon Choi, Manu Gopakumar, Ryota Koiso, and Gordon Wetzstein. 2025. Random-phase Wave Splatting of Translucent Primitives for Computer-generated Holography. arXiv:2508.17480 [cs.GR] <https://arxiv.org/abs/2508.17480>
- Xingyu Chen, Zihao Feng, Kun Qian, and Xinyu Zhang. 2025. Radio Frequency Ray Tracing with Neural Object Representation for Enhanced RF Modeling. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, New York, NY, USA, 21339–21348.
- Xingyu Chen and Xinyu Zhang. 2023. RF Genesis: Zero-Shot Generalization of mmWave Sensing through Simulation-Based Data Synthesis and Generative Diffusion Models. In *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems (Istanbul, Türkiye) (SenSys '23)*. Association for Computing Machinery, New York, NY, USA, 28–42. doi:10.1145/3625687.3625798
- Xingyu Chen, Xinyu Zhang, Kai Zheng, Xinmin Fang, Tzu-Mao Li, Chris Xiaoxuan Lu, and Zhengxiong Li. 2026. Physically Accurate Differentiable Inverse Rendering for Radio Frequency Digital Twin. In *Proceedings of the 32nd Annual International Conference on Mobile Computing and Networking (MobiCom)*. ACM, Austin, TX, USA. doi:10.1145/3795866.3796866
- Suyeon Choi, Brian Chao, Jacqueline Yang, Manu Gopakumar, and Gordon Wetzstein. 2025. Gaussian Wave Splatting for Computer-Generated Holography. *ACM Trans. Graph.* 44, 4, Article 57 (July 2025), 13 pages. doi:10.1145/3731163
- Quentin Denoyelle, Vincent Duval, Gabriel Peyré, and Emmanuel Soubies. 2020. The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy. *Inverse Problems* 36, 1 (2020), 014001. doi:10.1088/1361-6420/ab2a29
- Gene H. Golub and Victor Pereyra. 1973. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. *SIAM J. Numerical Analysis* 10, 2 (1973), 413–432. doi:10.1137/0710036
- Fredric J Harris. 1978. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. *Proc. IEEE* 66, 1 (Jan. 1978), 51–83. doi:10.1109/PROC.1978.10837
- Nils Hirsenkorn, Paul Subkowski, Timo Hanke, Alexander Schaermann, Andreas Rauch, Ralph Raschhofer, and Erwin Biehl. 2017. A Ray Launching Approach for Modeling an FMCW Radar System. In *18th International Radar Symposium (IRS)* (Prague, Czech Republic). IEEE, New York, NY, USA, 1–10. doi:10.23919/IRS.2017.8008120
- Nicolas Holzschuch and Romain Pacanowski. 2017. A Two-Scale Microfacet Reflectance Model Combining Reflection and Diffraction. *ACM Trans. Graph.* 36, 4, Article 66 (July 2017), 12 pages. doi:10.1145/3072959.3073621
- Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2024b. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. In *ACM SIGGRAPH 2024 Conference Papers* (Denver, CO, USA) (*SIGGRAPH '24*). Association for Computing Machinery, New York, NY, USA, Article 32, 11 pages. doi:10.1145/3641519.3657428
- Tianshu Huang, John Miller, Akarsh Prabhakara, Tao Jin, Tarana Laroia, Zico Kolter, and Anthony Rowe. 2024a. DART: Implicit Doppler Tomography for Radar Novel View Synthesis. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Seattle, WA, USA). IEEE, New York, NY, USA, 24118–24129. doi:10.1109/CVPR52733.2024.02277
- Martin Jaggi. 2013. Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization. In *Proceedings of the 30th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 28)*, Sanjoy Dasgupta and David McAllester (Eds.). PMLR, Atlanta, GA, USA, 427–435. <https://proceedings.mlr.press/v28/jaggi13.html>
- Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Trans. Graph.* 42, 4, Article 139 (July 2023), 14 pages. doi:10.1145/3592433
- Samuli Laine, Janne Hellsten, Tero Karras, Yeongho Seol, Jaakko Lehtinen, and Timo Aila. 2020. Modular primitives for high-performance differentiable rendering. *ACM Trans. Graph.* 39, 6, Article 194 (Nov. 2020), 14 pages. doi:10.1145/3414685.3417861
- Jonathon Luiten, Georgios Kopanas, Bastian Leibe, and Deva Ramanan. 2024. Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis. In *2024 International Conference on 3D Vision (3DV)* (Davos, Switzerland). IEEE, New York, NY, USA, 800–809. doi:10.1109/3DV62453.2024.00044
- Ruichun Ma, Shicheng Zheng, Hao Pan, Lili Qiu, Xingyu Chen, Liangyu Liu, Yihong Liu, Wenjun Hu, and Ju Ren. 2024. AutoMS: Automated Service for mmWave Coverage Optimization using Low-cost Metasurfaces. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking* (Washington D.C., DC, USA) (*MobiCom '24*). Association for Computing Machinery, New York, NY, USA, 62–76. doi:10.1145/3636534.3649347
- Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *Computer Vision – ECCV 2020* (Glasgow, United Kingdom) (*Lecture Notes in Computer Science, Vol. 12346*). Springer, Cham, Switzerland, 405–421. doi:10.1007/978-3-030-58452-8_24
- Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.* 41, 4, Article 102 (July 2022), 15 pages. doi:10.1145/3528223.3530127
- Dianne P. O'Leary and Bert W. Rust. 2013. Variable Projection for Nonlinear Least Squares Problems. *Comput. Optim. Appl.* 54, 3 (2013), 579–593. doi:10.1007/s10589-012-9492-9
- Alan V. Oppenheim, Ronald W. Schaffer, and John R. Buck. 1999. *Discrete-Time Signal Processing* (2nd ed.). Prentice Hall, Upper Saddle River, NJ, USA.
- Arthur Ouaknine, Alasdair Newson, Julien Rebut, Florence Tupin, and Patrick Pérez. 2021. CARRADA Dataset: Camera and Automotive Radar with Range- Angle-Doppler Annotations. In *25th International Conference on Pattern Recognition (ICPR)* (Milan, Italy). IEEE, New York, NY, USA, 5068–5075. doi:10.1109/ICPR48806.2021.9413181
- Mark A. Richards. 2014. *Fundamentals of Radar Signal Processing* (2nd ed.). McGraw-Hill, New York, NY, USA.
- Richard Roy and Thomas Kailath. 1989. ESPRIT—Estimation of Signal Parameters via Rotational Invariance Techniques. *IEEE Trans. Acoust., Speech, Signal Process.* 37, 7 (July 1989), 984–995. doi:10.1109/29.32276
- Nicolas Scheiner, Florian Kraus, Fangyin Wei, Buu Phan, Fahim Mannan, Nils Appenrodt, Werner Ritter, Jürgen Dickmann, Klaus Dietmayer, Bernhard Sick, and Felix Heide. 2020. Seeing Around Street Corners: Non-Line-of-Sight Detection and Tracking In-the-Wild Using Doppler Radar. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, New York, NY, USA, 2065–2074. doi:10.1109/CVPR42600.2020.00214
- Ralph O. Schmidt. 1986. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Trans. Antennas Propag.* 34, 3 (March 1986), 276–280. doi:10.1109/TAP.1986.1143830
- Jos Stam. 1999. Diffraction Shaders. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '99)*. ACM Press/Addison-Wesley Publishing Co., USA, 101–110. doi:10.1145/311535.311546
- Shlomi Steinberg, Ravi Ramamoorthi, Benedikt Bitterli, Eugene d'Eon, Ling-Qi Yan, and Matt Pharr. 2024. A Generalized Ray Formulation For Wave-Optical Light Transport. *ACM Trans. Graph.* 43, 6, Article 241 (Nov. 2024), 15 pages. doi:10.1145/3687902
- Shlomi Steinberg and Ling-Qi Yan. 2021. A Generic Framework for Physical Light Transport. *ACM Trans. Graph.* 40, 4, Article 139 (July 2021), 20 pages. doi:10.1145/3450626.3459791
- Chaozheng Wen, Jingwen Tong, Yingdong Hu, Zehong Lin, and Jun Zhang. 2026. Neural Representation for Wireless Radiation Field Reconstruction: A 3D Gaussian Splatting Approach. *IEEE Trans. Wireless Commun.* 25 (2026), 7490–7504. doi:10.1109/TWC.2025.3631663
- Tim A. Wheeler, Martin Holder, Hermann Winner, and Mykel J. Kochenderfer. 2017. Deep Stochastic Radar Models. In *IEEE Intelligent Vehicles Symposium (IV)* (Los Angeles, CA, USA). IEEE, New York, NY, USA, 47–53. doi:10.1109/IVS.2017.7995697
- Karl D. Willis and Andrew D. Wilson. 2013. InfraStructs: fabricating information inside physical objects for imaging in the terahertz region. *ACM Trans. Graph.* 32, 4, Article 138 (July 2013), 10 pages. doi:10.1145/2461912.2461936
- Kang Yang, Gaofeng Dong, Sijie Ji, Wan Du, and Mani Srivastava. 2025. GSRF: Complex-Valued 3D Gaussian Splatting for Efficient Radio-Frequency Data Synthesis. In *Advances in Neural Information Processing Systems* (San Diego, CA, USA), D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (Eds.), Vol. 38. Curran Associates, Inc., Red Hook, NY, USA, 63076–63105. https://proceedings.neurips.cc/paper_files/paper/2025/file/5b5618e7d061748267d74478b7c5b1ab-Paper-Conference.pdf
- Ziyi Yang, Xinyu Gao, Wen Zhou, Shaohui Jiao, Yuqing Zhang, and Xiaogang Jin. 2024. Deformable 3D Gaussians for High-Fidelity Monocular Dynamic Scene Reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Seattle, WA, USA). IEEE, New York, NY, USA, 20331–20341. doi:10.1109/CVPR52733.

2024.01922

Yicheng Zhan, Dong-Ha Shin, Seung-Hwan Baek, and Kaan Aksit. 2026. Complex-Valued Holographic Radiance Fields. *ACM Trans. Graph.* 45, 3, Article 31 (April 2026), 16 pages. doi:10.1145/3804450

Wuqiong Zhao, Baicheng Chen, Kai Zheng, Xingyu Chen, Wenqian Zhang, and Xinyu Zhang. 2026. FlowForm: Scalable Passive Metasurface Network for mmWave Coverage Expansion. In *Proceedings of the ACM SIGCOMM 2026 Conference* (Denver, CO, USA) (SIGCOMM '26). Association for Computing Machinery, New York, NY, USA, 564–581. doi:10.1145/3789240.3829102

A Mathematical Derivations

A.1 DFT response of a windowed exponential

THEOREM 1 (DFT RESPONSE OF A WINDOWED COMPLEX EXPONENTIAL). *Let $x[n] = \alpha w[n] \exp(j2\pi v n)$ for $n = 0, \dots, N-1$, where $w[n]$ is a window function. The N -point DFT is*

$$X[k] = \alpha W_N\left(v - \frac{k}{N}\right), \quad W_N(\delta) = \sum_{n=0}^{N-1} w[n] \exp(j2\pi \delta n). \quad (13)$$

For the rectangular window $w[n] \equiv 1$, this reduces to

$$X[k] = \alpha e^{j\pi(N-1)\delta} \frac{\sin(\pi N \delta)}{\sin(\pi \delta)}, \quad \delta = v - \frac{k}{N}, \quad (14)$$

Its magnitude is the absolute value of the Dirichlet response:

$$|X[k]| = |\alpha| |D_N(\delta)|, \quad D_N(\delta) = \frac{\sin(\pi N \delta)}{\sin(\pi \delta)}. \quad (15)$$

PROOF. Substitution into the DFT definition gives

$$X[k] = \alpha \sum_{n=0}^{N-1} w[n] e^{j2\pi(v - k/N)n}.$$

For $w[n] = 1$, summing the geometric series yields

$$\sum_{n=0}^{N-1} e^{j2\pi \delta n} = \frac{1 - e^{j2\pi N \delta}}{1 - e^{j2\pi \delta}} = e^{j\pi(N-1)\delta} \frac{\sin(\pi N \delta)}{\sin(\pi \delta)}.$$

At integer δ , the response follows by continuity. $\square$

The bin offset used in the renderer is $\Delta = k - \mu$, with $\mu = Nv$, so $\delta = -\Delta/N$. Thus

$$X[k] = \alpha e^{-j\pi \frac{N-1}{N} \Delta} \frac{\sin(\pi \Delta)}{\sin(\pi \Delta/N)} = \alpha \kappa_N(\Delta). \quad (16)$$

The power response $|X[k]|^2 = |\alpha|^2 |\kappa_N(\Delta)|^2$ is proportional to the Fejér kernel.

A.2 Kernel derivative

For the complex kernel $\kappa_N(\Delta) = e^{-j\pi \frac{N-1}{N} \Delta} \frac{\sin(\pi \Delta)}{\sin(\pi \Delta/N)}$, the derivative with respect to the bin offset Δ is

$$\frac{\partial \kappa_N(\Delta)}{\partial \Delta} = \kappa_N(\Delta) \left(-j\pi \frac{N-1}{N} + \pi \cot(\pi \Delta) - \frac{\pi}{N} \cot\left(\frac{\pi}{N} \Delta\right) \right). \quad (17)$$

The product rule gives the phase term, and differentiating the sine ratio gives the two cotangent terms. At integer offsets, the factored expression must be evaluated by a limit. An equivalent expression without singular factors follows directly from the finite sum:

$$\kappa'_N(\Delta) = -\frac{j2\pi}{N} \sum_{n=0}^{N-1} n e^{-j2\pi n \Delta/N}. \quad (18)$$

Appendix B.3 gives the limiting forms.

Geometry derivatives. Let a_i denote the amplitude factor multiplying the separable kernel product in Eq. (2), and let $\Delta_{\ell,i} = k_\ell - \mu_{\ell,i}$. For a scene parameter ξ ,

$$\begin{aligned} \partial_\xi \text{splat}_i &= (\partial_\xi a_i) \prod_\ell d_{N_\ell}(\Delta_{\ell,i}) \\ &\quad - a_i \sum_\ell (\partial_\xi \mu_{\ell,i}) d'_{N_\ell}(\Delta_{\ell,i}) \prod_{h \neq \ell} d_{N_h}(\Delta_{h,i}). \end{aligned} \quad (19)$$

The first term differentiates the amplitude; the second moves the spectral footprint. Writing $\hat{\mathbf{r}} = \hat{\mathbf{r}}_i$ and $R = R_i$, the projection in Eq. (6) gives

$$\begin{aligned} \nabla_{\mathbf{p}_i} \mu_{R,i} &= \frac{2KN_R}{cF_s} \hat{\mathbf{r}}, \\ \nabla_{\mathbf{p}_i} \mu_{D,i} &= \frac{2N_D T_c}{\lambda R} (I - \hat{\mathbf{r}} \hat{\mathbf{r}}^\top) \mathbf{v}_i, \\ \nabla_{\mathbf{v}_i} \mu_{D,i} &= \frac{2N_D T_c}{\lambda} \hat{\mathbf{r}}, \\ \nabla_{\mathbf{p}_i} \mu_{A,i} &= \frac{N_A \delta_{\text{el}}}{\lambda R} (I - \hat{\mathbf{r}} \hat{\mathbf{r}}^\top) \mathbf{u}_x. \end{aligned} \quad (20)$$

The range and azimuth coordinates are independent of velocity.

A.3 Windowed kernels

A taper changes the point response from κ_N to

$$\kappa_N^w(\Delta) = \sum_{n=0}^{N-1} w[n] e^{-j2\pi n \Delta/N}. \quad (21)$$

This finite sum gives the response at fractional offsets for any sampled window, including Hann, Hamming, and Kaiser windows. The renderer uses the sensor's window; the main-lobe width and side-lobe attenuation depend on that choice [Harris 1978].

A.4 Separable spectral response

PROPOSITION 1 (SEPARABLE DIRICHLET STRUCTURE). *For a single point reflector that produces a separable complex exponential of the form $x[m, n, a] = \alpha \exp(j2\pi(v_D m + v_R n + v_A a))$, the 3D DFT (with optional separable windows $w_D[m]$, $w_R[n]$, $w_A[a]$) factorizes as*

$$Y[p, k, q] = \alpha W_M\left(v_D - \frac{p}{M}\right) W_N\left(v_R - \frac{k}{N}\right) W_A\left(v_A - \frac{q}{A}\right), \quad (22)$$

where each W is a 1D window spectrum as in Eq. (13). For rectangular windows, each factor is (up to phase) a Dirichlet kernel, and the point response is therefore a separable product of Dirichlet kernels across the three Fourier axes.

PROOF. The signal and window separate over m , n , and a , so the windowed DFT factorizes into three sums of the form Eq. (13). $\square$

We evaluate the point response as a product of three 1D kernels.

A.5 FMCW separable measurement model

For FMCW acquisition [Richards 2014], a baseband chirp of duration T , slope $K = B/T$, and carrier f_c , dechirped against the transmit reference under the stop-and-hop approximation, returns a complex exponential at the beat frequency $f_b = 2KR/c$ for a reflector at range R . Across M chirps with repetition T_c , target motion at radial velocity v_r induces a slow-time Doppler $f_D = 2v_r/\lambda$ with $\lambda = c/f_c$. Using the azimuth sign convention of Eq. (6), an A -element uniform

Table 2. **Notation.** Symbols used throughout the paper, grouped by role. Per-surfel quantities carry index $i \in \{1, \dots, S\}$; λ denotes the carrier wavelength in physical contexts and the ridge weight in optimization contexts, disambiguated by usage.

<i>Scene and sensor</i>	
$\mathbf{p}_i, \mathbf{n}_i, A_i, \alpha_i, \mathbf{v}_i, \eta_i$	Surfel position, normal, cell area, complex amplitude, velocity, wave impedance; collected in Θ_i
$S, C \in \mathbb{R}^{S \times 3}$	Splat budget; stack of all surfel centers
$\mathbf{o}, R_i, \hat{\mathbf{r}}_i$	Sensor origin; range and line-of-sight unit vector to surfel i
V_i, F_i, γ	Bistatic foreshortening, Fresnel factor, path-loss exponent
$F_s, K, T_c, \lambda, \delta_{\text{el}}$	ADC rate, chirp slope, chirp repetition interval, carrier wavelength, array element spacing
<i>Fourier domain and kernels</i>	
$\ell \in \{R, D, A\}, N_R, N_D, N_A$	Axis index and per-axis FFT lengths (range / Doppler / azimuth)
$\ \Omega\ = N_R N_D N_A$	Total measurement dimension
$\mu_{\ell, i}$	Continuous Fourier coordinate of surfel i on axis ℓ (Eq. (6))
$\Delta = k - \mu$	Continuous bin offset (bin units)
$\kappa_N(\Delta), d_N = \kappa_N / N$	Complex DFT kernel ($\kappa_N(0) = N$) and its unit-peak form
κ_N^w, ϕ_N	Windowed and power (Fejér) kernel variants
<i>Forward model and inverse problem</i>	
$\text{splat}_i(\mathbf{f})$	Per-surfel spectral splat (Eq. (2))
$\hat{\mathbf{Y}}, \hat{\mathbf{H}} = \hat{\mathbf{Y}} ^2$	Rendered complex spectral cube and power image
$\mathbf{y}, H^{\text{obs}}$	Measured complex cube and measured power
$A(C) \in \mathbb{C}^{\ \Omega\ \times S}, \mathbf{b}$	Dictionary matrix (column i = surfel footprint) and complex coefficient vector
$\mathbf{b}^*(C), F(C)$	Closed-form ridge solution and variable-projected objective (Eqs. (9) and (10))
G, q	Gram matrix $A^H A / \ \Omega\ $ and right-hand side $A^H \mathbf{y} / \ \Omega\ $
<i>DSFW certificate and search</i>	
$s(\mathbf{x})$	Residual dual certificate field (Eq. (11))
$u_j, \text{eff}_j \in [0, 1]$	Per-surfel utility and coherent-cancellation efficiency (Eq. (12))
r	Per-axis truncation radius for splat support (default 8 bins)

linear array with spacing d has a far-field steering phase across antennas is $\exp(j2\pi a u(\theta))$ with $u(\theta) = (d/\lambda) \sin \theta$. The sampled dechirped signal across (slow-time, fast-time, array) is therefore separable:

$$x[m, n, a] = \alpha \exp(j2\pi(f_D m T_c + f_b n T_s + a u(\theta))), \quad (23)$$

which matches the canonical form of Proposition 1 with $v_D = f_D T_c$, $v_R = f_b T_s$, $v_A = u(\theta)$. The Fourier-processing pipeline therefore reduces to evaluating one windowed DFT per axis, whose closed form is the Dirichlet kernel of Eqs. (15)–(16).

A.6 Validity of the separable model

Equation (23) assumes stop-and-hop acquisition, constant radial velocity, and negligible range walk under the far-field array model above. We also give the interval in which velocity can be recovered without aliasing.

(A1) *Stop-and-hop.* The target is treated as stationary during a single chirp (fast-time), with motion accumulating only between chirps (slow-time). This drops a within-chirp contribution $\exp(j2\pi f_D n T_s)$ that is negligible whenever

$$|v_r| \ll \frac{c}{2f_c N_R T_s}. \quad (24)$$

(A2) *Constant radial velocity.* The slow-time phase is linear only for constant radial velocity over the coherent processing interval

$T_{\text{CPI}} = N_D T_c$. Radial acceleration a_r adds a quadratic phase proportional to $a_r (m T_c)^2 / \lambda$, producing a Fresnel-type response. This contribution is negligible when

$$|a_r| T_{\text{CPI}}^2 \ll \lambda. \quad (25)$$

Acceleration outside this regime requires a different Doppler response.

(A3) *Negligible range walk.* A target with non-zero radial velocity traverses $\Delta R_{\text{walk}} = v_r T_{\text{CPI}}$ during the CPI, which couples range and Doppler once ΔR_{walk} exceeds one range bin $\Delta R_{\text{bin}} = c/(2B)$ (sweep bandwidth B). Separability of $\prod_{\ell} d_{N_{\ell}}(\cdot)$ in Eq. (2) therefore requires

$$|v_r| T_{\text{CPI}} \ll \frac{c}{2B}. \quad (26)$$

(A4) *Doppler unambiguity.* The slow-time sampling imposes the Nyquist bound $|f_D| < 1/(2T_c)$. Radial velocities are therefore unambiguous only on

$$|v_r| < \frac{\lambda}{4T_c}. \quad (27)$$

Velocities separated by integer multiples of $\lambda/(2T_c)$ produce the same sampled Doppler response. This ambiguity affects velocity recovery, not the separability of the forward response.

A.7 Fresnel response at a planar boundary

Setup. A surfel is a locally planar boundary between the surrounding medium, of wave impedance $\eta_0 = \sqrt{\mu_0/\epsilon_0}$, and the surfel

material, of wave impedance $\eta_i = \sqrt{\mu_i/\epsilon_i}$. A plane wave arrives at angle θ_{inc} measured from the surfel normal $\mathbf{n}_i$ and excites a transmitted wave at angle θ_t . Matching the tangential electric and magnetic fields across the boundary gives the classical two-equation system whose solution is the reflection coefficient below.

Transverse-electric coefficient. When the incident electric field is perpendicular to the plane of incidence, tangential-field matching yields

$$F_{\text{TE}} = \frac{\eta_i \cos \theta_{\text{inc}} - \eta_0 \cos \theta_t}{\eta_i \cos \theta_{\text{inc}} + \eta_0 \cos \theta_t}, \quad (28)$$

which is Eq. (5). The transmitted angle follows from Snell's law, $\epsilon_0^{1/2} \sin \theta_{\text{inc}} = \epsilon_i^{1/2} \sin \theta_t$, which for non-magnetic media ($\mu_i = \mu_0$) simplifies to $\sin \theta_t / \sin \theta_{\text{inc}} = \eta_i / \eta_0$.

Transverse-magnetic coefficient. When the incident magnetic field is perpendicular to the plane of incidence, the same matching argument with the roles of the two fields exchanged gives

$$F_{\text{TM}} = \frac{\eta_i \cos \theta_t - \eta_0 \cos \theta_{\text{inc}}}{\eta_i \cos \theta_t + \eta_0 \cos \theta_{\text{inc}}}. \quad (29)$$

Our sub-THz transceiver is linearly polarized and the targets are oriented so that the transverse-electric branch dominates, so the renderer evaluates Eq. (28).

Two limits. For a perfect conductor, $|\eta_i| \rightarrow 0$ and Eq. (28) gives $F \rightarrow -1$ at every incidence angle, so the angular dependence collapses and α_i alone determines the splat amplitude. In the dielectric case, where η_i is comparable to η_0 , $|F|$ varies over the aperture by a factor that grows with the angular extent of the scan, and fitting α_i alone cannot absorb that variation because α_i is constant across the aperture.

Identifiability. The measurement constrains only the product $\alpha_i F_i$. Separating the two parameters requires an aperture spanning enough incidence angles to reveal the angular variation of F . Under a narrow aperture the pair is degenerate up to a scale, and we hold η_i at its material prior. The reported sub-THz experiments use metal targets in the $|\eta_i| \rightarrow 0$ regime, so impedance is not estimated.

B Implementation Details

B.1 Sensor parameters

The projection in Eq. (6) uses sensor constants read from the device datasheet or programmed waveform profile. Table 3 maps these constants to the two front-ends used in this paper. None are learned.

Bin-scale consistency. Substituting the FMCW identities

$$F_s = KT \cdot (N_R/N_R \text{ samples per chirp}), \quad \Delta R_{\text{bin}} = c/(2B)$$

reduces $\partial \mu_R / \partial R = (2K/F_s)(N_R/c)$ to $N_R/\Delta R_{\text{bin}}$, giving a unit bin shift for one range-bin displacement. The Doppler scale $2N_D T_c / \lambda$ maps the velocity-bin width $\lambda/(2T_c N_D)$ to one bin; the azimuth scale $N_A \delta_{\text{el}} / \lambda$ does the same on the spatial axis. Porting the renderer therefore requires only the device constants ($2K/cF_s$, $2T_c/\lambda$, $\delta_{\text{el}}/\lambda$). The remaining forward model is shared across modalities.

B.2 Renderer truncation and noise floor

Truncation. We truncate each splat's contribution to a radius $r = 8$ bins per axis, which reduces the per-step cost from $O(S \prod_{\ell=1}^L N_\ell)$ to $O(S r^L)$ and makes the renderer scale linearly in surfel count rather than in spectral-tensor volume. For the rectangular window with $N \in [64, 256]$, $r = 8$ captures $> 99\%$ of the kernel energy. Tapered windows (Hann, Hamming) are even more concentrated, and the same radius remains safe.

Noise floor. We add a learned or calibrated Gaussian noise floor σ^2 to the power image: $\hat{H} = |\hat{Y}|^2 + \sigma^2$. Section 6 describes the loss design, including log-power supervision for the 40–60 dB dynamic range of coherent measurements.

Matrix-free accumulation. Materializing $A(C) \in \mathbb{C}^{|\Omega| \times S}$ for Eq. (9) exceeds 1 GiB in single-precision complex for our 3D settings. We instead stream voxel chunks and accumulate $G = A^H A$ and $q = A^H y$. Since G is Hermitian, we store its upper triangle, reducing the working set to $O(\text{chunk} \cdot S)$. A fused CUDA kernel combines footprint evaluation, conjugate multiplication, and reduction without intermediate writes. The streamed path is numerically equivalent to the dense reference and reduces peak memory and per-step time by over an order of magnitude (Appendix C.4).

B.3 Training stabilizers and numerical safeguards

Periodic coupled correction. To account for the off-diagonal coherent coupling in §6, every $T_{\text{LM}} = 40$ outer steps we therefore take a global Gauss–Newton step on a downsampled spectral target, using the VarPro-reduced center Jacobian $J_{\text{red}} = (I - P_A)J_c$ [Golub and Pereyra 1973] with a trust-radius clip and a full-resolution acceptance check. The periodic correction implements the low-resolution acceptance sub-step of Algorithm 1.

Window matching. The analytic kernel κ_N^w uses the same taper as the sensor, so that the rendered point response matches the deployed measurement chain to within numerical precision.

Kaiser-to-target kernel-continuation curriculum. Training begins with a smooth Kaiser kernel and anneals the shape parameter toward the deployed taper, mirroring the multi-resolution spectral-estimation practice of windowed-FFT super-resolution [Harris 1978]. This restores sidelobe contrast while avoiding early sidelobe peaks.

Pairwise repulsion penalty. A short-range inter-splat penalty prevents centers from collapsing under coherent cancellation.

Coarse-to-fine schedule on kernel order N . Short windows are used early (broad mainlobes, weak sidelobes) and the full N near convergence, so that initial certificate peaks are well separated and final accuracy is set by the deployed kernel.

Kernel evaluation near integer offsets. The derivative formula $\partial \kappa_N / \partial \Delta$ of Appendix A.2 contains two cotangent terms that individually diverge at integer bin offsets but whose combination remains finite. Near $\Delta \approx 0$,

$$\lim_{\Delta \rightarrow 0} \left[\pi \cot(\pi \Delta) - \frac{\pi}{N} \cot\left(\frac{\pi}{N} \Delta\right) \right] = 0,$$

Table 3. **Mapping from device configuration to the constants in Eq. (6).** The mmWave column corresponds to the FMCW Doppler-validation experiment of Appendix C.1. The sub-THz column corresponds to the THz application of §7.7.

Constant	mmWave FMCW radar	Sub-THz SDR
f_c (carrier frequency)	77 GHz (waveform-profile f_{start})	120 GHz (LO synthesizer setpoint)
$\lambda = c/f_c$	3.90 mm	2.50 mm
B (sweep bandwidth)	250 MHz (frequency-slope $\times T$)	1.23 GHz (programmed sweep span)
T (chirp duration)	user-set fast-time window	8192 T_s (acquisition window)
$K = B/T$ (chirp slope)	derived; sets range-bin scale via $2K/c$	derived from B and T
T_c (chirp repetition interval)	60 μs (chirp-cycle-time register)	N/A (single-shot acquisition)
F_s (ADC sampling rate)	ADC sampling-rate register	SDR-card sample-rate setting
N_R	ADC samples per chirp	8192 (fast-time samples)
N_D	chirps per frame	captures per frame
N_A	$N_{\text{tx}}N_{\text{rx}}$ (virtual ULA)	raster-scan grid size
δ_{el} (array element spacing)	$\lambda/2$ (cascade EVM)	raster step (mechanical)

Table 4. **DSFW hyperparameters per experiment.** T : outer iterations; λ : VarPro ridge weight; r : truncation radius in bins; T_{lr} : low-resolution sub-step interval; β_0 : initial Kaiser shape.

Experiment	T	λ	r	T_{lr}	β_0
2D synthetic (128×128, $S = 50$)	200	10^{-4}	8	every 20	14
3D synthetic (64×256×256)	300	10^{-4}	4	every 25	14
THz planar target	400	10^{-3}	8	every 20	10
mmWave automotive cube	200	10^{-4}	4	every 25	14

We switch to a Taylor branch when $|\Delta| < \epsilon$, with $\epsilon = 10^{-4}$ throughout:

$$\pi \cot(\pi\Delta) \approx \frac{1}{\Delta} - \frac{\pi^2\Delta}{3} + O(\Delta^3),$$

$$\frac{\pi}{N} \cot\left(\frac{\pi\Delta}{N}\right) \approx \frac{1}{\Delta} - \frac{\pi^2\Delta}{3N^2} + O(\Delta^3).$$

The difference $-\frac{\pi^2\Delta}{3}(1 - 1/N^2) + O(\Delta^3)$ is well-behaved at the origin. The same ϵ -regularization is applied near every Δ that is a non-zero multiple of N , where $\sin(\pi\Delta/N) = 0$. This branch switch is the only floating-point safeguard in the forward and backward kernel evaluations.

B.4 Hyperparameters and Reproducibility

Table 4 lists the experiment settings, fixed across seeds. Only the seed and initial centers C_0 vary.

Baseline optimizer settings. The first-order controls run for 300 steps under the paper preset. Dirichlet-AdamW and SGD use center/coefficient learning rates $8 \times 10^{-3}/2.5 \times 10^{-2}$, with AdamW weight decay 10^{-4} and SGD momentum 0.9. The 3DGS-style, GSRF, and WRF-GS+ baselines use AdamW with learning rate 0.03 and weight decay 10^{-4} . The quasi-Newton control follows 300 AdamW steps with 50 L-BFGS iterations (learning rate 0.8, history 20, strong-Wolfe line search). AdamW uses the PyTorch default moment coefficients (0.9, 0.999).

All experiments run on a single NVIDIA RTX A6000 with PyTorch 2.1 and CUDA 12.1. Metrics average 10 seeds, controlling the uniform initial centers C_0 and tie-breaking among low-utility candidates.

Table 5. **Candidate grid, range gate, and splat budget.** Each comparison shares budget S . The raw-signal budget is capped by both the listed cap and the scene reflector count. The sub-THz candidate pitch is $0.75\times$ the beam sigma, upsampled $2\times$.

Experiment	Candidate grid	Range gate	S
Signal-space fitting (§7.3)	128×128	full axis	50
Basin sweep (§7.3)	128×128	full axis	5
Raw-signal imaging (§7.4)	56×56	full axis	≤ 192
Sub-THz letters (§7.7)	beam-derived pitch	ROI gate	512

Candidate volume and initialization. Table 5 lists the candidate grids and splat budgets. The candidate grid spans the sensor’s range gate and azimuth field of view. Each comparison shares this grid and budget S , with cold-start centers drawn uniformly without replacement. The multi-view fit of §7.5 and the printed-object fit of §7.7 seed their centers from back-projection peaks instead.

C Additional experiments

C.1 Doppler validation

We validate the Doppler parameter under the constant-radial-velocity, no-range-walk assumptions of Appendix A.6. Acceleration, tracking, occlusion changes, and dynamic topology are outside this validation’s scope.

Protocol. We synthesize a three-reflector range-Doppler-azimuth cube with $N_R = 96$, $N_D = 64$, and $N_A = 8$. The FMCW constants are carrier frequency $f_c = 77$ GHz, chirp interval $T_c = 60 \mu\text{s}$, and bandwidth $B = 250$ MHz, giving a velocity-bin spacing of 0.507 m/s. The three targets have radial velocities -6.0 , 0.0 , and 7.5 m/s, with maximum range walk of 0.051 range bins across the coherent processing interval. We compare three rows: (i) the finite-window FFT synthesis oracle; (ii) a joint Dirichlet position-and-velocity model that searches the radial velocity coordinate on a 0.05 m/s grid over ± 8.0 m/s, 321 candidates, and refits complex amplitudes by least squares; and (iii) a static position-only Dirichlet control with the same range and azimuth support but Doppler center fixed to zero.

The joint velocity model matches the FFT oracle to the numerical floor, while clamping velocity to zero on the same scene degrades

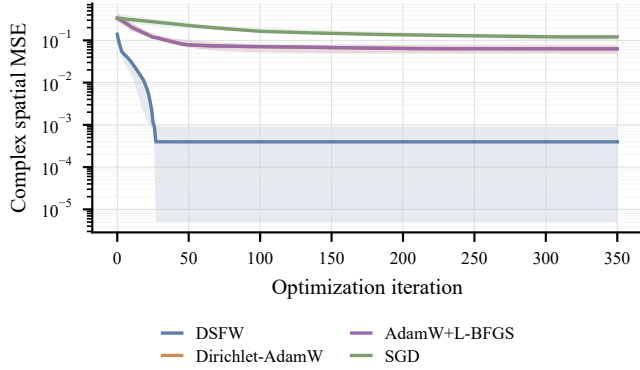


Fig. 11. **Cross-scene convergence trajectories.** Median complex spatial MSE with interquartile shading over the seven signal-space scenes in Fig. 5.

Table 6. **Doppler validation metrics.** The joint velocity model recovers the moving-reflector cube, while the static control cannot explain Doppler-shifted energy.

Method	Velocity RMSE (m/s)↓	Complex NMSE (dB)↓
FFT reference	0.000	-120.00
Joint Dirichlet velocity	0.000	-120.00
Static position-only	5.545	-1.11

the fit by more than two orders of magnitude in complex NMSE. This isolates the contribution of the Doppler parameter itself, holding range and angle fixed.

Randomized sweep. We repeat the Doppler validation on eight randomized three-reflector scenes with off-bin moving-target velocities ($1.25 \leq |v_r| \leq 8.0$ m/s) and complex Gaussian noise at 40, 30, 20, and 10 dB SNR. We use the same range and azimuth support with a coarser 0.25 m/s velocity-search step. The joint model obtains 0.074 m/s mean velocity RMSE at all four SNRs. Its mean complex NMSE changes from -11.97 dB at 40 dB SNR to -8.14 dB at 10 dB SNR. The static control remains near 0 dB, giving static-minus-joint NMSE gaps of 11.97 dB and 8.13 dB at those two endpoints. This experiment tests the Doppler parameter across randomized velocities and noise levels. It does not evaluate dynamic-scene reconstruction.

C.2 Cross-scene quantitative summaries

Tables report mean $\pm$ standard deviation across the displayed scenes in the signal-space fitting and raw-signal imaging experiments. PSNR and SSIM are computed on the displayed reconstruction panels.

Convergence suite. To complement the representative loss trace in the main paper, Fig. 11 aggregates the recorded histories over the seven scenes in Fig. 5.

Basin of attraction. Fig. 12 visualizes the two-dimensional cold-start success map summarized in §7.3.

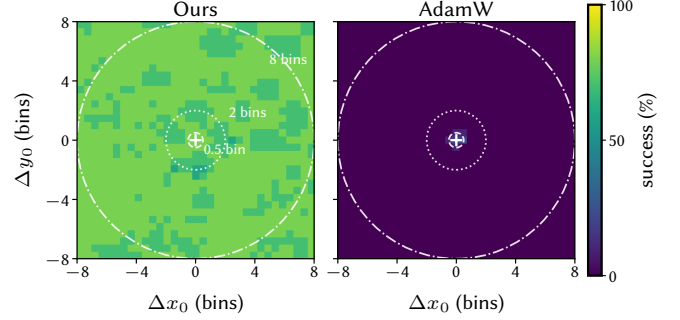


Fig. 12. **Basin of attraction.** Two-dimensional cold-start success over initial position offset on $K = 5$ reflector scenes, averaged over ten random seeds. The certificate update gives DSFW a broad attraction basin across the tested eight-bin range (77.4% aggregate success), while Dirichlet-AdamW succeeds on only 0.07% of trials.

C.3 Per-component ablations

§7.6 reports the 2×2 kernel-optimizer factorial. Tab. 11 reports a rerun under the paper preset (10 seeds, 128×128 grid, 5 splats), which adds the center RMSE column.

Table 11. **Kernel-and-optimizer factorial rerun.** The Dirichlet kernel and SFW/VarPro optimizer are both required for reliable cold-start recovery under the E4 paper preset.

Method	Success (%)	MSE↓	Center RMSE (bins)↓
Gaussian + AdamW	0.0	8.54×10^{-3}	36.310
Gaussian + SFW	10.0	6.71×10^{-3}	0.187
Dirichlet + AdamW	0.0	1.18×10^{-2}	41.630
Dirichlet + DSFW	100.0	5.32×10^{-9}	0.001

At MSE 1.18×10^{-2} Dirichlet+AdamW trails Gaussian+AdamW (8.54×10^{-3}), because the Dirichlet kernel's sub-bin phase and side-lobe structure produce a more rugged landscape than the Gaussian's monotonic decay. The kernel and solver are therefore co-designed. The $10^6 \times$ MSE gap between Dirichlet+AdamW and Dirichlet+DSFW measures the contribution of the solver given a fixed Dirichlet kernel.

C.4 Runtime and memory

Protocol. We measure wall-clock time and peak GPU allocation for the fused forward renderer, varying the number of primitives S on standard 2D and 3D signal grids.

Backward pass. At 16,384 primitives on an RTX 5080, the Dirichlet primitive costs 1.2 to $1.6 \times$ a Gaussian in the forward pass (3.2 versus 2.1 ms) and $1.08 \times$ in the backward pass (25.1 versus 23.2 ms). A DSFW step takes 50 to 180 ms, compared with 1 to 9 ms for AdamW. However, DSFW reaches within 10 to 30 steps a loss that AdamW does not reach within 2000 steps. The mean total reconstruction time for §7.4 is 16.4 s.

The 256^3 memory values are dominated by the complex output volume. The small increase with primitive count comes from the primitive parameter buffers and temporary shader state.

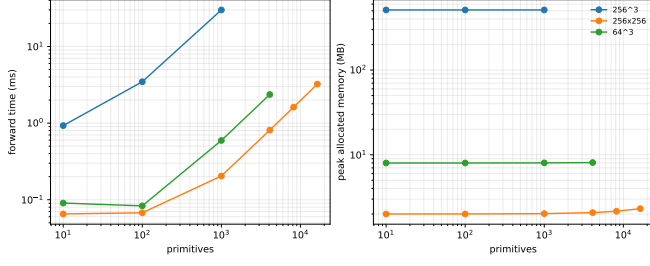


Fig. 13. **Forward render runtime and memory scaling.** Forward render time scales with primitive count, while peak allocation is dominated by the output grid size for these fused-renderer settings.

Table 7. **Forward runtime and memory.** Steady-state fused Dirichlet forward render time and peak GPU allocation.

Grid	Primitives	Time (ms)	Peak MB
256×256	10	0.065	2.0
256×256	100	0.068	2.0
256×256	1000	0.204	2.0
256×256	4096	0.810	2.1
256×256	8192	1.615	2.2
256×256	16384	3.236	2.3
64 ³	10	0.090	8.0
64 ³	100	0.083	8.0
64 ³	1000	0.594	8.0
64 ³	4096	2.363	8.1
256 ³	10	0.928	512.0
256 ³	100	3.466	512.0
256 ³	1000	29.984	512.0

Table 8. **Cross-scene inverse signal-space statistics.** Mean $\pm$ standard deviation over the seven signal-space scenes in Fig. 5. PSNR and SSIM are computed on the displayed magnitude fields against the reference target.

Method	Complex MSE↓	Mag. NMSE (dB)↓	PSNR↑	SSIM↑
3DGS	$1.92 \times 10^{-1} \pm 1.18 \times 10^{-1}$	-3.09 ± 0.59	17.44 ± 4.94	0.396 ± 0.154
GSRF	$4.43 \times 10^{-2} \pm 3.02 \times 10^{-2}$	-10.90 ± 3.75	25.26 ± 3.71	0.654 ± 0.135
WRF-GS+	$4.38 \times 10^{-2} \pm 3.04 \times 10^{-2}$	-11.01 ± 3.74	25.37 ± 3.68	0.650 ± 0.126
SGD	$1.12 \times 10^{-1} \pm 5.57 \times 10^{-2}$	-3.98 ± 2.51	18.34 ± 3.12	0.542 ± 0.108
Dirichlet-AdamW	$6.06 \times 10^{-2} \pm 3.32 \times 10^{-2}$	-6.62 ± 3.40	20.98 ± 3.36	0.712 ± 0.128
AdamW+L-BFGS	$5.97 \times 10^{-2} \pm 3.27 \times 10^{-2}$	-6.69 ± 3.44	21.04 ± 3.35	0.717 ± 0.131
Ours	$6.71 \times 10^{-4} \pm 9.06 \times 10^{-4}$	-38.54 ± 18.53	52.89 ± 18.69	0.998 ± 0.003

Table 9. **Cross-scene convergence statistics.** Mean $\pm$ standard deviation over the seven signal-space scenes in Fig. 5. ρ is the empirical exponential loss-decay rate fitted to each recorded trace. The threshold column reports the mean iteration among scenes that reach complex MSE 10^{-3} , followed by the reach count.

Method	ρ ↑	Iter. to 10^{-3} ↓	Final MSE↓
DSFW	0.790 ± 1.525	18.6 ± 11.1 (5/7)	$6.71 \times 10^{-4} \pm 9.06 \times 10^{-4}$
Dirichlet-AdamW	0.064 ± 0.010	– (0/7)	$6.06 \times 10^{-2} \pm 3.32 \times 10^{-2}$
AdamW+L-BFGS	0.063 ± 0.010	– (0/7)	$5.97 \times 10^{-2} \pm 3.27 \times 10^{-2}$
SGD	0.012 ± 0.005	– (0/7)	$1.12 \times 10^{-1} \pm 5.57 \times 10^{-2}$

Table 10. **Cross-scene raw-signal inverse-imaging statistics.** Mean $\pm$ standard deviation over the seven displayed planar targets in Fig. 7. PSNR and SSIM are computed on object-plane reconstruction panels against the reference support.

Method	Support NMSE↓	PSNR↑	SSIM↑	IoU↑	Dice↑
Back-projection	0.387 ± 0.252	13.25 ± 2.02	0.601 ± 0.106	0.612 ± 0.211	0.740 ± 0.176
3DGS	0.389 ± 0.266	13.23 ± 1.69	0.392 ± 0.088	0.628 ± 0.207	0.753 ± 0.171
GSRF	0.406 ± 0.248	12.98 ± 1.82	0.554 ± 0.045	0.590 ± 0.204	0.723 ± 0.171
WRF-GS+	0.391 ± 0.250	13.19 ± 1.94	0.602 ± 0.048	0.610 ± 0.208	0.739 ± 0.173
GS+SFW	0.294 ± 0.306	15.38 ± 3.00	0.683 ± 0.130	0.740 ± 0.269	0.823 ± 0.208
Ours	0.136 ± 0.109	19.41 ± 5.51	0.810 ± 0.061	0.801 ± 0.150	0.883 ± 0.095