Event-based Sparse-view Background-Oriented Schlieren Tomography

Xinyu Zhou, Shihao Hu, Peiqi Duan*, Chao Xu, and Boxin Shi*

Peking University
{zhouxiny, duanqi0001, shiboxin}@pku.edu.cn, xuchao@cis.pku.edu.cn

Abstract. Background-oriented schlieren (BOS) tomography reconstructs 3D flow density fields from refraction-induced distortions observed from one or more viewpoints. Capturing high-speed flows with frame-based BOS typically requires high-speed cameras and bright illumination to compensate for short exposure times. In this paper, we propose an event-based BOS tomography approach that reconstructs time-varying 4D density fields from event streams. Leveraging the high temporal resolution and high dynamic range of event cameras, the proposed approach enables high-speed airflow reconstruction under ambient lighting. We represent the spatiotemporal density field as a neural implicit field and render BOS observations via refractive ray tracing. We utilize physics-informed regularization to improve reconstruction under sparse views. Experiments on simulated and real data in single-view and orthogonal dual-view setups demonstrate accurate reconstructions.

Keywords: Background-oriented schlieren tomography · Event camera · Neural implicit field · Physics-informed regularization

1 Introduction

Visualizing flow in transparent media is important to many scientific and engineering applications, including aerodynamic testing [7, 20, 53] and diagnostics of heating, ventilation, and air conditioning (HVAC) systems [1, 42]. Most optical techniques for examining such transparent media rely on imaging refractive-index variations within the flow. Among them, *background-oriented schlieren* (BOS) has received substantial attention because of its simple, low-cost setup. In BOS, a camera images a textured background through the medium (Fig. 1(a)); spatial refractive-index variations deflect light rays and induce apparent distortions of the background pattern in the recorded images [44, 47–49]. These

* Corresponding authors.
Xinyu Zhou and Chao Xu are with State Key Laboratory of General Artificial Intelligence, School of Intelligence Science and Technology, Peking University. Peiqi Duan, Shihao Hu, and Boxin Shi are with State Key Laboratory of Multimedia Information Processing and National Engineering Research Center of Visual Technology, School of Computer Science, Peking University.

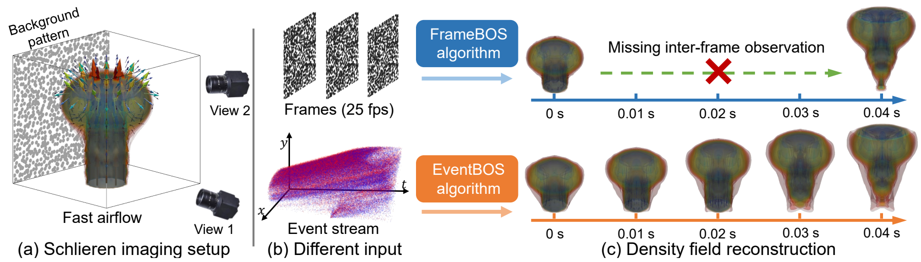


Fig. 1: Event-based versus frame-based BOS tomography. (a) Cameras look through the airflow toward a textured background. (b) Different input: low-frame-rate frame sequences versus asynchronous event signals. (c) Our method reconstructs fast airflow from event streams, whereas low-frame-rate frame sequences yield reconstructions only at discrete timestamps.

refraction effects are then recovered by estimating a 2D displacement field between a "deflected" image and an undistorted reference image [44]. Recovering the underlying 3D density field further requires tomographic reconstruction from BOS observations acquired from one or more viewpoints. This inverse problem is often ill-posed and generally requires sufficient viewpoints to mitigate spatial ambiguities [47].

An important application of BOS is the measurement of high-speed fluid phenomena, including shock-wave transients [52] and thermal plumes or jets [3, 58]. Capturing such dynamics typically requires high-speed cameras, imposing substantial demands on data bandwidth for transmission and processing—especially when synchronized multi-view acquisition is required for 3D tomographic reconstruction. In addition, BOS is often operated with a small aperture (large f-number) to increase depth of field [36, 47]; maintaining signal quality at short exposure times typically requires bright illumination. Together, these factors limit real-world applications of frame-based BOS (FrameBOS), particularly for time-resolved 3D tomographic reconstruction.

Event cameras [13, 32] are neuromorphic sensors that asynchronously detect per-pixel brightness changes with microsecond-level timestamps. They offer advantages such as high temporal resolution, high dynamic range, and low bandwidth requirements, making them promising for high-speed BOS measurements [51]. Moreover, events provide a direct measurement of temporal contrast variations caused by refractive distortions, enabling a live preview of schlieren phenomena during acquisition.

Existing event-based BOS (EventBOS) studies have primarily focused on qualitative visualization and 2D displacement estimation [33, 34, 51], and EventBOS tomography remains unexplored. Event cameras encode radiance changes rather than absolute intensities, which departs from the standard FrameBOS formulation. Under constant illumination, the displacement can only be computed between two nearby timestamps [51], whereas FrameBOS typically computes displacement fields from paired reference and deflected frames. Consequently, FrameBOS tomographic pipelines cannot be directly applied to EventBOS. Prior work [33, 34] has used pulsed illumination to induce dense event streams that can be integrated to recover FrameBOS-like images. However, this strategy may diminish the data efficiency of event sensing [34]. Therefore, *how to effectively exploit the unique characteristics of event cameras for BOS tomography remains an open question.*

EventBOS tomography poses two key challenges. First, event streams in BOS scenes are sparse and noisy, making it difficult to reliably estimate 2D displacements from events using existing optical-flow methods [51]. Second, deploying dense multi-view event-camera arrays is often impractical due to increased hardware cost and calibration complexity, requiring sparse-view tomography. In this paper, we propose an event-based BOS tomography method that reconstructs time-varying 4D density fields from event streams. Compared to low-frame-rate frame inputs (Fig. 1(b)), our method recovers high-speed airflow dynamics (Fig. 1(c)). To avoid explicitly estimating 2D displacement, we represent the unknown spatiotemporal density as a neural implicit field, and optimize it by matching measured event signals to BOS predictions rendered via differentiable refractive ray tracing. Motivated by recent physics-informed BOS tomography methods [10, 55], we incorporate physics-informed regularization to improve reconstruction accuracy under sparse-view measurements. Specifically, we jointly estimate auxiliary flow variables and penalize violations of fluid-dynamics priors. We evaluate both single-view and orthogonal dual-view settings, and demonstrate effectiveness across both sparse-view regimes. The key contributions of our work are summarized as follows:

- We present the first EventBOS tomography pipeline that reconstructs 4D density fields from event streams.
- We formulate a differentiable EventBOS inverse-rendering objective by combining refractive ray tracing with the event generation model, enabling direct supervision from sparse events.
- We demonstrate that the proposed method enables high-speed airflow reconstruction (outperforming 1000 fps frame-based baselines), and validate its effectiveness on both simulated and real data.

2 Related Work

2.1 Background-Oriented Schlieren

Schlieren photography [48] visualizes refractive flows by exploiting spatial variations in refractive index. Unlike tracer-based techniques such as particle image velocimetry (PIV) [2, 45], it does not require seeding the flow with particles or dyes. BOS is a widely used schlieren variant in which a camera images a textured background through the working fluid. Flow-induced refraction is then recovered by estimating the displacement field between a "deflected" image and an undistorted reference image. Numerous deflection-sensing algorithms have been proposed, including cross-correlation [11, 43], iterative least squares [40], optical flow [4], and Fourier demodulation [60]. While BOS often uses artificial patterns [44], Hargather *et al.* [22] showed that natural scenes (*e.g.*, forests) can provide sufficient texture, enabling BOS measurements over a large field of view.

Most BOS techniques focus on visualization. Recently, several works have extended BOS toward quantitative flow analysis by exploiting spatiotemporal coherence in refractive patterns. Xue *et al.* [63] observed that intensity variations induced by moving refractive structures remain locally consistent over small space–time volumes, enabling optical-flow estimation to track the projected motion of the refractive medium. Similarly, Settles *et al.* [50] used streak-schlieren analysis (kymography) to extract fluid flow velocity.

BOS tomography seeks to reconstruct the underlying density (or refractive-index) field from BOS measurements that encode line-of-sight integrated refraction effects. Early BOS tomography methods focus on axisymmetric flows, where the problem reduces to a 2D reconstruction that can be solved using the Abel transform [57, 58, 62]. For general non-axisymmetric 3D flows, however, single-view observations are inherently ambiguous, and multi-view acquisition is typically required to resolve spatial ambiguities. Atcheson *et al.* [5], for example, demonstrated instantaneous reconstructions of the hot plume above a candle flame using an array of 16 machine-vision cameras. In practice, BOS tomography often employs on the order of 6 to 30 views [47], and cameras are often arranged along a semicircle with equal angular spacing [17, 18]. Algebraic reconstruction algorithms are the most widely used methods, which discretize the volume and measurement equations to form a linear system [9, 19, 29, 38]. Motivated by recent advances in inverse neural rendering, several works have explored tomographic reconstruction using neural implicit representations [30, 55, 64]. By combining neural implicit representations with physics-informed regularization, Teh *et al.* [55] demonstrated steady airflow reconstruction from single-view BOS measurements. Our method follows this physics-regularized neural BOS tomography line, but differs by reconstructing time-varying 4D density fields from event signals rather than from frame-based BOS observations.

2.2 Event-based Fluid Flow Measurement

Compared with frame-based sensors, event cameras offer high temporal resolution, low data redundancy, high dynamic range, and low power consumption. These characteristics make them well-suited for flow visualization and measurement under rapid dynamics and challenging illumination. Recent studies have applied event cameras to flow visualization, including particle-based velocimetry [8, 59, 61] and BOS [33, 34, 51]. Borer *et al.* [8] used three synchronized event cameras for 3D particle tracking in a medium-sized wind tunnel, and Wang *et al.* [59] proposed a stereo event-camera approach for reconstructing 3D velocity fields. For BOS, Shiba *et al.* [51] investigated event-based BOS under continuous illumination and estimated 2D displacement fields using both events and frames. Unlike FrameBOS that estimates displacement between the reference image and the deflected image, event cameras respond only to temporal intensity changes. Lyu *et al.* [33] used pulsed illumination to reconstruct FrameBOS-like frames

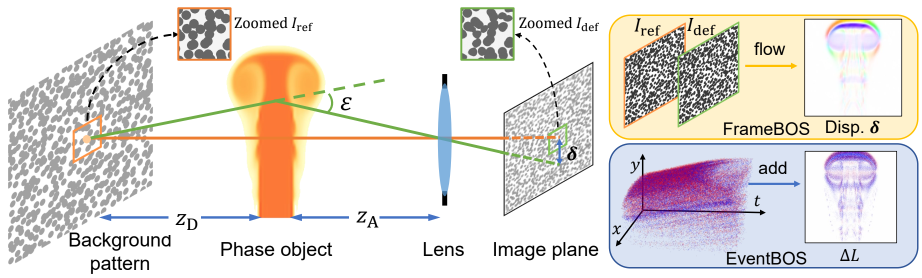


Fig. 2: Sketch of BOS imaging. The image-plane displacement $\boldsymbol{\delta}$ is determined by the deflection angle $\boldsymbol{\varepsilon}$ and the imaging geometry. FrameBOS typically estimates the displacement field to infer deflection, whereas in EventBOS, accumulated events encode scene log-brightness changes ΔL.

from events, enabling displacement estimation via standard cross-correlation as in FrameBOS. However, this strategy may diminish the bandwidth advantage of event sensing. Despite these advances, event-based BOS tomography for general 3D flows remains unexplored.

3 Event-based Schlieren Formulation

This section presents the event-based schlieren forward model. We first derive the BOS imaging model that maps a 3D refractive-index field to image-plane displacements, and then introduce the event generation model.

3.1 BOS Imaging Formulation

In BOS, spatial variations in refractive index bend light rays, producing apparent distortions of the background pattern, as shown in Fig. 2. The flow region is commonly referred to as the *phase object* in schlieren imaging. The medium is characterized by its refractive index n, which for gases depends approximately linearly on mass density ρ via the Gladstone–Dale relation:

$$n = 1 + G\rho \quad \Rightarrow \quad \nabla n = G \nabla \rho, \tag{1}$$

where G is the Gladstone–Dale constant [16]. For air at visible wavelengths, $G \approx 2.26 \times 10^{-4}\,\mathrm{m}^3/\mathrm{kg}$. A light ray can be parameterized by its 3D position $\mathbf{x} \in \mathbb{R}^3$ and its optical momentum $\mathbf{v} \in \mathbb{R}^3$. We define the optical momentum as $\mathbf{v}(s) \triangleq n(\mathbf{x}(s))\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}$, where s denotes the geometric arc-length parameter along the ray. With this definition, ray propagation through an inhomogeneous medium follows Hamilton's equations [26]:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \frac{\mathbf{v}(s)}{n(\mathbf{x}(s))}, \quad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}s} = \nabla n(\mathbf{x}(s)). \tag{2}$$

Integrating along the ray yields the deflection angle

$$\boldsymbol{\varepsilon} = \frac{1}{n_0} \int \nabla_{\perp} n\big(\mathbf{x}(s)\big)\, \mathrm{d}s, \tag{3}$$

where n_0 denotes the ambient refractive index, and $\nabla_{\perp}$ denotes the component of the refractive-index gradient perpendicular to the principal ray direction. The deflection angle is related to the image-plane displacement through the BOS imaging geometry [44, 47, 49]. As shown in Fig. 2, let z_{D} denote the background-to-phase-object distance, z_{A} the phase-object-to-lens distance, and z_{B} the background-to-lens distance. With pixel pitch ψ_{pix} and focal length f, the magnification is $M = \frac{f}{z_{\mathrm{B}}-f}$. Assuming refractive-index gradients are negligible outside the phase object, the image-plane displacement $\boldsymbol{\delta}$ is

$$\boldsymbol{\delta} = \frac{M z_{\mathrm{D}}}{\psi_{\mathrm{pix}}}\, \boldsymbol{\varepsilon}. \tag{4}$$

The forward model thus maps a 3D refractive-index field to a 2D displacement field on the image plane. Let $I_{\mathrm{ref}}(\boldsymbol{\xi})$ be the reference background image and $I_{\mathrm{def}}(\boldsymbol{\xi}, t)$ the deflected image at time t. Under brightness constancy, the reference and deflected images are related by

$$I_{\mathrm{def}}(\boldsymbol{\xi}, t) = I_{\mathrm{ref}}(\boldsymbol{\xi} + \boldsymbol{\delta}(\boldsymbol{\xi}, t))\,. \tag{5}$$

3.2 Event Formation Model

We define the logarithmic brightness $L(\boldsymbol{\xi}, t) \triangleq \log I_{\mathrm{def}}(\boldsymbol{\xi}, t)$. An event $e_k = (\boldsymbol{\xi}_k, t_k, p_k)$ is triggered when the change in L at pixel $\boldsymbol{\xi}_k$ exceeds the contrast threshold [14, 32]. We denote the positive and negative thresholds by C^{+} and C^{-}, and write $C_{p_k} \in \{C^{+}, C^{-}\}$ for the threshold associated with polarity p_k:

$$L(\boldsymbol{\xi}_k, t_k) - L(\boldsymbol{\xi}_k, t_k - \Delta t_k) = p_k C_{p_k}. \tag{6}$$

where $p_k \in \{1, -1\}$ is the polarity and $t_k - \Delta t_k$ is the timestamp of the preceding event at pixel $\boldsymbol{\xi}_k$. For any two time instants $t_1 < t_2$, events accumulated over $(t_1, t_2]$ provide a discrete approximation of the log-brightness change. Define the event set $\mathcal{E}_{(t_1,t_2]} \triangleq \{e_k \mid t_1 < t_k \leq t_2\}$ and the accumulated event image:

$$E(\boldsymbol{\xi}; t_1, t_2) \triangleq \sum_{e_k \in \mathcal{E}_{(t_1,t_2]}} p_k\, C_{p_k}\, \delta_{\mathrm{D}}(\boldsymbol{\xi} - \boldsymbol{\xi}_k), \tag{7}$$

where the Dirac delta function δ_{D} selects the pixel $\boldsymbol{\xi}$. Under this model, the log-brightness difference is given by

$$\Delta L(\boldsymbol{\xi}; t_1, t_2) = L(\boldsymbol{\xi}, t_2) - L(\boldsymbol{\xi}, t_1) \approx E(\boldsymbol{\xi}; t_1, t_2). \tag{8}$$

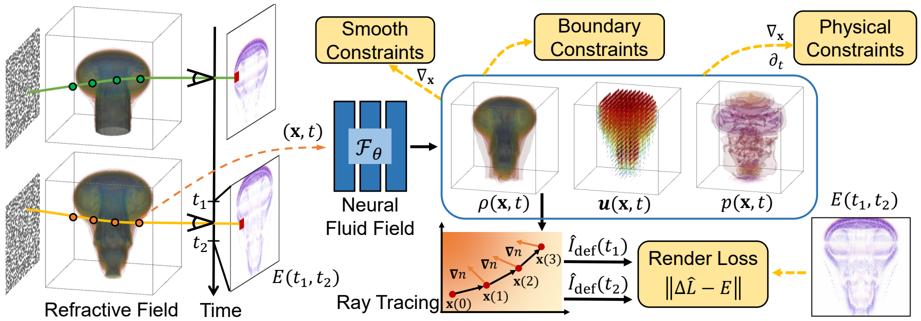


Fig. 3: Method overview. We model the dynamic fluid field with a coordinate-MLP. Given two nearby timestamps, we predict the brightness change via refractive ray tracing. Physical constraints are incorporated to regularize sparse-view reconstruction.

4 Event-based Density Field Reconstruction

To solve the inverse problem in event-based BOS tomography, we adopt an analysis-by-synthesis framework. We represent the density field with a coordinate-based MLP and optimize its parameters to match the observed event streams. Fig. 3 offers an overview of our method. The following sections detail the neural representation and the corresponding optimization procedure.

4.1 Neural Fluid Field Representation

Inspired by recent neural BOS tomography reconstruction methods [23, 30, 55], we represent the fluid scene as continuous space–time fields using one coordinate-based MLP:

$$\mathcal{F}_\theta : (\mathbf{x}, t) \mapsto (\rho, \mathbf{u}, p), \tag{9}$$

where ρ is the density, $\mathbf{u}$ the fluid velocity, and p the hydrostatic-pressure-corrected pressure. Given a 4D coordinate $(\mathbf{x}, t)$, the network outputs the corresponding physical quantities at that space–time location. The predicted density ρ is converted to refractive index n via the Gladstone–Dale relation (Eq. (1)), which is then used by the BOS forward model to compute ray deflections. Our method additionally requires the reference background image $I_{\mathrm{ref}}(\boldsymbol{\xi})$. In practice, $I_{\mathrm{ref}}(\boldsymbol{\xi})$ can be captured with a frame-based camera under the same viewpoint [51], or obtained via brightness modulation that allows an event camera to measure dense scene irradiance (*e.g.*, using active lighting [21] or an aperture shutter [6]). For any time instant t within the observation window, the deflected image $\widehat{I}_{\mathrm{def}}(\boldsymbol{\xi}, t)$ is synthesized by the BOS forward model in Sec. 3.1, yielding the corresponding log-brightness $\widehat{L}(\boldsymbol{\xi}, t)$. The auxiliary outputs $\mathbf{u}$ and p are introduced for physics-informed regularization. Direct supervision of these auxiliary variables is typically unavailable in BOS. We therefore estimate $\mathbf{u}$ and p implicitly during optimization, using them only as auxiliary variables that regularize the density reconstruction through the physics-informed constraints.

4.2 Loss Functions

We optimize the neural fluid field by minimizing a weighted sum of four terms:

$$\mathcal{L} = \lambda_{\text{render}}\mathcal{L}_{\text{render}} + \lambda_{\text{bnd}}\mathcal{L}_{\text{bnd}} + \lambda_{\text{tv}}\mathcal{L}_{\text{tv}} + \lambda_{\text{phy}}\mathcal{L}_{\text{phy}}. \tag{10}$$

These terms enforce (i) consistency with the observed event streams, (ii) boundary conditions needed to recover absolute density, (iii) spatial smoothness via total variation, and (iv) physical plausibility of thermal-convection dynamics.

Render loss. For a time window $(t_1, t_2]$, the accumulated event image $E(\boldsymbol{\xi}; t_1, t_2)$ provides a discrete approximation to the log-brightness change. We optimize a mean-squared error (MSE) loss that aligns the rendered log-brightness difference with the measured event accumulation:

$$\mathcal{L}_{\text{render}} = \mathbb{E}_{(\boldsymbol{\xi}, t_1, t_2)}\left[\left\| \left(\widehat{L}(\boldsymbol{\xi}, t_2) - \widehat{L}(\boldsymbol{\xi}, t_1)\right) - E(\boldsymbol{\xi}; t_1, t_2)\right\|_2^2\right], \tag{11}$$

where $\widehat{L}(\boldsymbol{\xi}, t)$ is the rendered log-brightness at time t.

Boundary loss. BOS observations encode integrated refractive-index gradients and therefore cannot directly determine absolute density. In our method, the absolute density is calibrated by an ambient boundary condition. Specifically, we assume part of the volume boundary remains undisturbed ambient air with known density ρ_0, which fixes the density-offset ambiguity. We implement this assumption by encouraging ρ to match ρ_0 near $\partial\Omega$:

$$\mathcal{L}_{\text{bnd}}^{\text{amb}} = \mathbb{E}_{(\mathbf{x},t)}\left[\|\rho(\mathbf{x}, t) - \rho_0\|_2^2\right]. \tag{12}$$

Optionally, when inlet priors are available, we supervise the inlet region Ω_{in}:

$$\mathcal{L}_{\text{bnd}}^{\text{in}} = \mathbb{E}_{(\mathbf{x},t)}\left[\|\rho(\mathbf{x}, t) - \rho_{\text{in}}(\mathbf{x}, t)\|_2^2 + \alpha_{\mathbf{u}}\|\mathbf{u}(\mathbf{x}, t) - \mathbf{u}_{\text{in}}(\mathbf{x}, t)\|_2^2\right], \tag{13}$$

which can improve accuracy in sparse-view settings. The full boundary loss is

$$\mathcal{L}_{\text{bnd}} = \mathcal{L}_{\text{bnd}}^{\text{amb}} + w_{\text{in}}\,\mathcal{L}_{\text{bnd}}^{\text{in}}, \quad w_{\text{in}} \in \{0, 1\}, \tag{14}$$

where w_{in} indicates whether inlet supervision is enabled.

Total-variation regularization. To encourage spatial smoothness of the reconstructed density, we apply an ℓ_1 total-variation penalty on the density gradients:

$$\mathcal{L}_{\text{tv}} = \mathbb{E}_{(\mathbf{x},t)}\left[\|\nabla_{\mathbf{x}}\rho(\mathbf{x}, t)\|_1\right]. \tag{15}$$

Physics-informed regularization. Prior work has integrated BOS tomography with a PINN framework [10, 37, 55]. Here we focus on thermal convection and regularize the reconstruction with simplified Navier–Stokes equations coupled with a heat-transfer equation. We assume low-Mach, nearly isobaric thermal convection in air and obtain temperature from density via the ideal-gas relation

$T(\rho) = \frac{P_0}{R\rho}$ under the assumed constant ambient pressure P_0. We penalize the residuals of mass conservation, momentum conservation, and heat transfer:

$$r_{\text{mass}} \triangleq \partial_t \rho + \nabla \cdot (\rho \mathbf{u}), \tag{16}$$

$$r_{\text{mom}} \triangleq \rho\Big(\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u}\Big) + \nabla p - \mu \nabla^2 \mathbf{u} - (\rho - \rho_0)\mathbf{g}, \tag{17}$$

$$r_{\text{heat}} \triangleq \partial_t T(\rho) + \mathbf{u}\cdot\nabla T(\rho) - \kappa \nabla^2 T(\rho), \tag{18}$$

where μ is the dynamic viscosity, κ is the thermal diffusivity, and $\mathbf{g}$ is gravity. We non-dimensionalize the governing equations and construct the physics loss from the resulting dimensionless residuals, which mitigates scale imbalance across terms and improves training stability. The physics loss is

$$\mathcal{L}_{\text{phy}} = \mathbb{E}_{(\mathbf{x},t)} \left[\|r_{\text{mass}}\|_2^2 + \|r_{\text{mom}}\|_2^2 + \|r_{\text{heat}}\|_2^2 \right], \tag{19}$$

with all derivatives computed via numerical differentiation [23, 31].

4.3 Optimization

Quasi-linear ray-tracing approximation. Accurate BOS forward rendering requires non-linear ray tracing through the phase-object region. In principle, Hamilton's ray equations (Eq. (2)) can be integrated by iteratively updating the ray state and querying $n(\mathbf{x})$ and $\nabla n(\mathbf{x})$ along the trajectory. In our setting, however, $n(\mathbf{x})$ is predicted by a neural field, so dense numerical integration would require many network queries and is therefore expensive. Following previous works [25, 55], we adopt a quasi-linear approximation that evaluates the neural field only along a straight-line proxy within the phase object. Let $t \in [t_n, t_f]$ denote the segment inside the phase object, with entry state $(\mathbf{x}_0, \mathbf{v}_0)$ at $t = t_n$, and approximate the path as $\bar{\mathbf{x}}(t) = \mathbf{x}_0 + (t - t_n)\mathbf{v}_0$. We query n and ∇n along $\bar{\mathbf{x}}(t)$ and update the ray direction and ray position by

$$\mathbf{v}(t) \approx \mathbf{v}_0 + \int_{t_n}^{t} \nabla n\big(\bar{\mathbf{x}}(\tau)\big)\, \mathrm{d}\tau, \quad t \in [t_n, t_f], \tag{20}$$

$$\mathbf{x}(t) \approx \mathbf{x}_0 + \int_{t_n}^{t} \frac{\mathbf{v}(\tau)}{n\big(\bar{\mathbf{x}}(\tau)\big)}\, \mathrm{d}\tau, \quad t \in [t_n, t_f], \tag{21}$$

and the exit state is given by $(\mathbf{x}(t_f), \mathbf{v}(t_f))$.

During training, we implement this tracing in a memory-bounded form. We first query the MLP on a sampled refractive-index grid and trace rays on this grid. Following adjoint nonlinear ray tracing [56], we compute the gradient of the rendering loss with respect to the sampled refractive field using an adjoint-gradient formulation, avoiding storage of full ray histories.

Implementation details. We model the neural fluid fields with a coordinate-based MLP using sinusoidal positional encodings [35, 54] and a sigmoid linear unit (SiLU) activation [24] to encourage smoother gradients. The network has 8 layers with 256 hidden units per layer, and we use 7 spatial and 9 temporal frequency bands for the encoding. We adopt a coarse-to-fine training schedule via positional-encoding annealing [41] across both spatial and temporal frequencies, which empirically improves spatiotemporal smoothness and stabilizes optimization. For optimization, we use the Adam optimizer [28] with a learning rate of 5×10^{-4} and train for 80K iterations per scene. In our experiments, we set λ_{render}, λ_{phy}, λ_{bnd}, and λ_{tv} to 500, 100, 5, and 1, respectively. In each iteration, we sample 16,384 pixels and use 160 steps per ray for ray tracing. Training uses about 12 GB GPU memory and takes about 5 h on one RTX 3090 GPU.

5 Experiments

In this section, we first construct a synthetic dataset and benchmark our method on simulated data (Sec. 5.1), evaluating both density reconstruction (Sec. 5.2) and the predicted optical flow (Sec. 5.3). We then evaluate on real-captured data under both single-view and dual-view setups (Sec. 5.4).

5.1 Synthetic Dataset

We construct a synthetic dataset that mirrors the full sensing pipeline: (i) CFD-based flow simulation, (ii) BOS image synthesis using our forward optical model, and (iii) event generation from the synthesized images. Flow fields are simulated in OpenFOAM [27,39], focusing on thermal-convection scenarios consistent with our real experiments, including hot-air-gun plumes and cold/hot inflow configurations. We use the transient buoyancy-aware solver `buoyantPimpleFoam` with a large-eddy simulation (LES) turbulence model to capture turbulent plume structures. The simulation outputs are recorded at a time step of 5×10^{-4} s to support high-speed flows. We assume an ambient temperature of 20°C and adopt the ideal-gas model. The physical domain size is $0.2\,\text{m} \times 0.2\,\text{m} \times 0.2\,\text{m}$.

To synthesize BOS observations, we match the optical configuration to the real capture setup [51]. We use a pixel size of 5 μm and a 35 mm lens; the camera is positioned 1.7 m from the phase-object center and 3.3 m from the background. The background pattern is a randomly generated dot pattern. Ray tracing within the phase-object region is performed using fourth-order Runge–Kutta (RK4) integration with a step size of 1 mm. We then generate event streams from the synthesized BOS image sequence using the ESIM simulator [15, 46], with a contrast threshold of 0.15 for both positive and negative events.

5.2 Density Field Evaluation

We evaluate the accuracy of the recovered density field on the synthetic dataset. For evaluation, we query the continuous reconstruction on an 80^3 grid at 1000 Hz for approximately 1 s sequences, corresponding to about 1000 timestamps. Reconstruction quality is quantified using root mean squared error (RMSE) and

Table 1: Quantitative results of density field reconstruction on the synthetic dataset. ↑/↓ indicates higher/lower is better throughout this paper. The best performances are highlighted in **bold**.

		Heated air jet		Hot airflow		Heated air puff		Cold airflow	
		RMSE↓	PCC↑	RMSE↓	PCC↑	RMSE↓	PCC↑	RMSE↓	PCC↑
1 view	Frame (25 fps)	0.0243	0.912	0.0247	0.913	0.0361	0.716	0.0267	0.898
	Frame (1000 fps)	0.0184	0.946	0.0234	0.921	0.0154	0.825	0.0251	0.913
	Ours	**0.0183**	**0.947**	**0.0229**	**0.928**	**0.0148**	**0.852**	**0.0236**	**0.924**
2 views	NeRIF* [23] (1000 fps)	0.0209	0.937	0.0221	0.929	0.0133	0.884	0.0341	0.826
	Frame (25 fps)	0.0216	0.929	0.0199	0.940	0.0317	0.769	0.0194	0.944
	Frame (1000 fps)	0.0172	0.955	0.0152	0.967	0.0129	0.899	0.0143	0.971
	Ours	**0.0149**	**0.967**	**0.0144**	**0.971**	**0.0126**	**0.917**	**0.0141**	**0.972**

the Pearson correlation coefficient (PCC), which measures linear correlation between the predicted and ground-truth densities. We consider two sparse-view acquisition setups: a single-view setting and a dual-view setting with orthogonal viewpoints. To highlight the temporal advantage of event sensing in high-speed scenarios, we additionally replace the event-stream input with frame sequences at 25 fps and 1000 fps. For the dual-view setup, we compare against the neural BOS tomography baseline NeRIF [23]. Since NeRIF reconstructs the density field at a single time instant, we extend it to a spatiotemporal (4D) variant by feeding the time coordinate t into the network, and we denote this extension as NeRIF*. For a fair comparison, NeRIF* uses the same MLP architecture, training iterations, and sampling strategy as our method.

Quantitative results are reported in Tab. 1. Our method outperforms the frame-based baseline with both 25 and 1000 fps inputs in single-view and dual-view settings, demonstrating the advantage of event sensing for high-speed BOS. NeRIF* performs worse overall, highlighting the importance of physical constraints to stabilize sparse-view tomography. We also observe that for near-axisymmetric sequences (*Heated air jet* and *Heated air puff*), adding an orthogonal view yields a smaller improvement than for asymmetric flows.

Fig. 4 presents qualitative comparisons. With only two views, our method reconstructs the density field with high fidelity. Using a single view still yields reasonable results but exhibits depth ambiguity. Compared with frame-based inputs, event-based reconstruction yields better noise suppression because events concentrate on regions with refractive distortions. Unlike event streams, 25 fps frames undersample high-speed schlieren dynamics, forcing the model to infer the evolution during inter-frame dead time. Fig. 5 visualizes a temporally continuous reconstruction result. With 25 fps video, supervision is available only at sparse instants (*e.g.*, t=0 and t=0.04 s), leading to nearly interpolative behavior between frames. In contrast, event streams provide dense temporal supervision, enabling more faithful recovery of time-varying density dynamics.

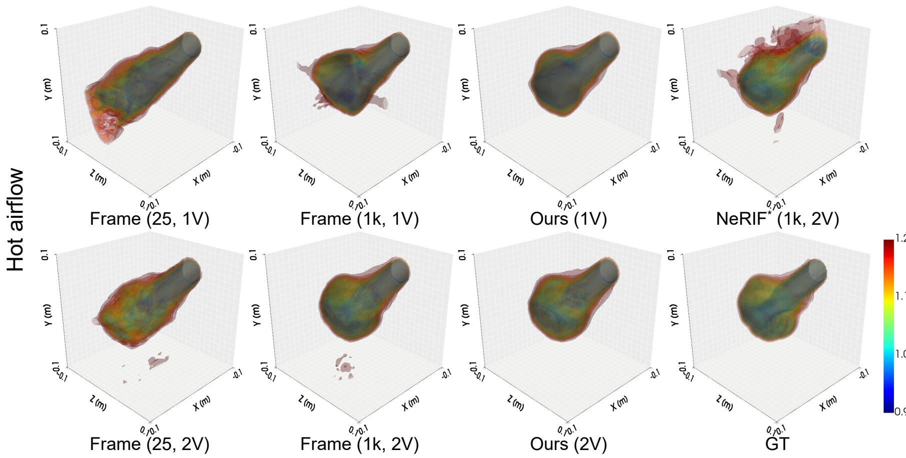


Fig. 4: Qualitative comparison of density reconstruction on synthetic data. We compare reconstructions obtained from event streams (ours) and from frame sequences at 25 fps and 1000 fps. Labels indicate the input frame rate (25 or 1k) and the number of views (1V/2V); 1k denotes 1000 fps.

Table 2: Ablation study on the *Hot airflow* sequence (2 views).

	W/o Bnd	W/o Rend	W/o {In,Phy}	W/o In	W/o Phy	W/o TV	Ours
RMSE↓	0.526	0.0539	0.0242	0.0209	0.0171	0.0152	**0.0144**
PCC↑	0.672	0.547	0.915	0.949	0.958	0.967	**0.971**

Ablation studies. We conduct an ablation study on the *Hot airflow* sequence under a dual-view setup and report the results in Tab. 2. Here, "W/o" denotes removing the corresponding component (In: inlet supervision, Phy: physics-informed regularization, TV: total variation, Bnd: boundary loss, Rend: rendering loss). The drops in "W/o Phy" and "W/o Bnd" show the benefit of physics-informed regularization and the necessity of boundary constraints for recovering absolute density. Removing any component degrades performance, and the full model achieves the best results, indicating that these terms are complementary and jointly contribute to accurate density reconstruction.

5.3 Optical Flow Evaluation

Our method reconstructs the underlying 4D density field directly from event streams, bypassing explicit estimation of the image-plane displacement field. To compare with the event-based schlieren optical-flow method EBOS [51], we reproject the reconstructed density via the BOS forward model and obtain optical flow in the camera plane. Specifically, we perform ray tracing to synthesize BOS displacements at two consecutive instants and convert their difference into

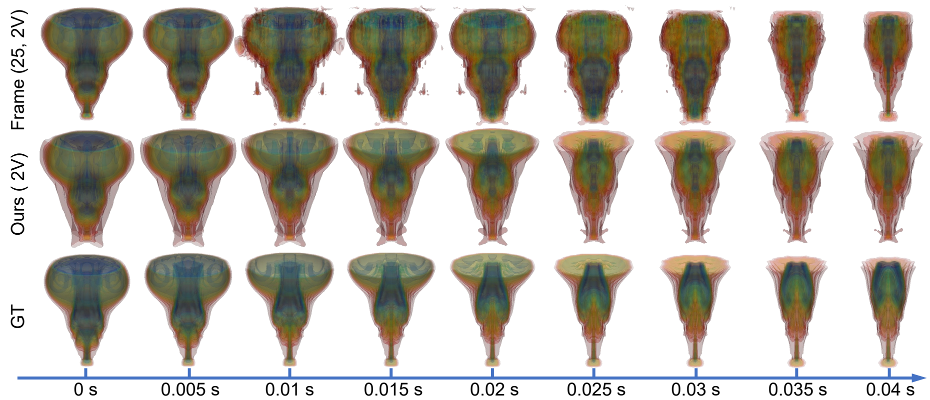


Fig. 5: Temporally continuous density reconstruction on a synthetic sequence.

Table 3: Quantitative results of optical flow estimation.

	Heated air jet			Hot airflow			Heated air puff		
	AEE↓	%Out↓	AE↓	AEE↓	%Out↓	AE↓	AEE↓	%Out↓	AE↓
Farnebäck's [12]	0.081	0.00	4.07	0.062	0.00	3.05	0.040	0.00	1.98
EBOS [51]	0.467	7.76	23.05	0.660	20.35	29.24	0.739	25.51	31.48
Ours	0.102	0.00	5.16	0.075	0.00	3.77	0.044	0.00	2.21

a pixel-level flow field. Ground-truth flow is computed by applying the same procedure to the ground-truth density.

On the synthetic dataset, we evaluate flow between adjacent instants at 200 fps. As an additional reference, we compute optical flow on the synthesized image pairs using the classical Farnebäck method [12]. Following EBOS, we report three metrics: average endpoint error (AEE), the percentage of pixels with AEE $>$ 1 px ("% Out"), and angular error (AE, measured in degrees). All metrics are computed over pixels with at least one event. Quantitative results are summarized in Tab. 3. Our reprojection-based flow substantially outperforms EBOS and is comparable to Farnebäck. Notably, our method achieves "% Out" = 0 across all sequences, indicating that it avoids large pixel-level flow outliers.

Fig. 6 presents qualitative comparisons. Our results closely match the ground truth, whereas EBOS exhibits noticeable errors, especially in regions with few events. Moreover, EBOS requires synchronized intensity frames from a co-located conventional camera, whereas our method requires only a reference background image. Without such synchronized frames, EBOS would likely degrade further.

5.4 Evaluation on Real Data

We further evaluate our method on real-captured event-based BOS data. We first test on the single-view dataset of [51], which provides synchronized events

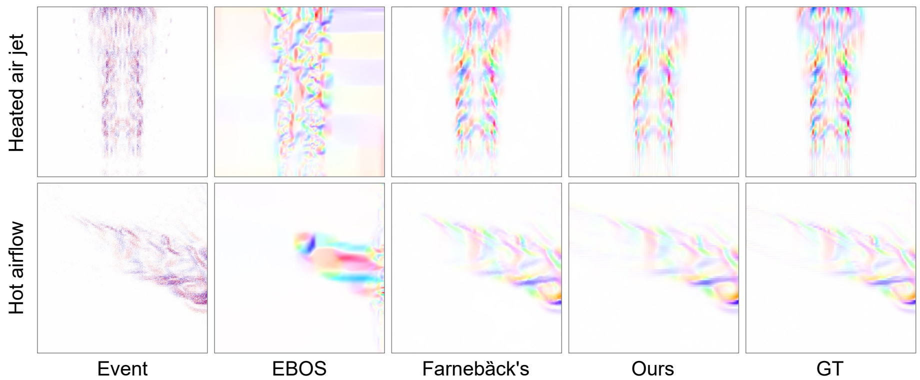


Fig. 6: Qualitative comparison of flow estimation. We compare our method to the event-based schlieren optical-flow method EBOS [51] and the classical frame-based Farnebäck method [12].

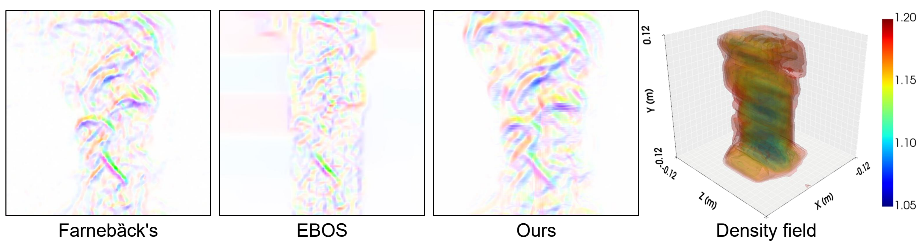


Fig. 7: Results on real-captured data from [51] (*Hot plate 1*). We show density reconstruction from events and the corresponding optical-flow comparison.

and intensity frames. We use the reference image captured by the co-located frame camera as the background. Fig. 7 shows our reconstruction on the *Hot plate 1* sequence. The single-view reconstruction contains noticeable noise but still captures the overall plume structure and temporal evolution. Similar to Sec. 5.3, we compare optical flow derived from our reconstructed density field with EBOS [51] and the classical Farnebäck method [12]. As shown in Fig. 7, our flow results align more closely with Farnebäck. Using Farnebäck as a proxy ground truth, our method achieves AEE $= 0.096$, %Out $= 0.191$, and AE $= 5.22$, outperforming EBOS (AEE $= 0.487$, %Out $= 12.215$, AE $= 24.12$).

To evaluate performance under a dual-view setup and to validate our method without auxiliary intensity frames, we also capture our own two-view real data. The background reference is obtained from dense events triggered by toggling the illumination [6, 21]. Fig. 8 presents reconstructions of a thermal convection plume generated by an electric heater. To verify reconstruction consistency, we use the reconstructed density to predict the log-brightness change $\Delta\widehat{L}$ between two consecutive timestamps and compare it against the event observation. The

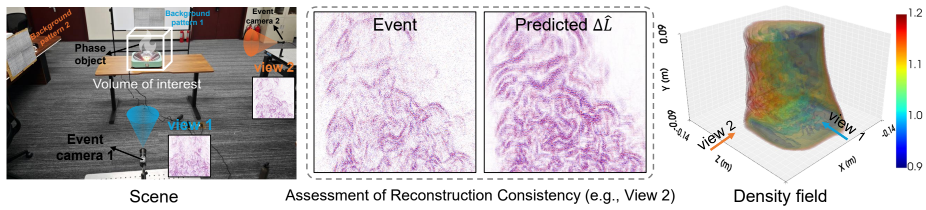


Fig. 8: Results on our real-captured dual-view data. We assess reconstruction consistency by comparing the predicted brightness change $\Delta\widehat{L}$ with event signals.

predicted $\Delta\widehat{L}$ matches the measured event signal well and is robust to event noise, supporting the fidelity of our reconstructions on real data.

6 Conclusion

We propose the first EventBOS tomography pipeline that reconstructs time-varying 4D density fields from event streams, enabling high-speed, continuous airflow reconstruction. Our method represents the dynamic fluid flow with a neural implicit representation and incorporates physics-informed regularization to improve reconstruction under sparse views. Experiments on both simulated and real data demonstrate accurate reconstructions in single-view and orthogonal dual-view settings. Our method outperforms 1000 fps frame-based baselines and existing event-based BOS optical-flow methods.

Limitations. Because event cameras respond to brightness changes, near-steady flows may produce very sparse event signals, making reliable measurement challenging. The absolute density estimate depends on the ambient boundary condition and the known value ρ_0; if this boundary is unavailable or ρ_0 is inaccurate, the reconstructed absolute values may be biased. Moreover, our current formulation is tailored to thermal convection in air and does not directly generalize to other scenarios (*e.g.*, combustion or gas leaks) without adapting the scene-specific priors and governing physical models.

Acknowledgements

This work is supported by the National Key R&D Program of China (Grant No. 2025YFA1016700), National Natural Science Foundation of China (Grant No. 62402014 and 62276007), Beijing Natural Science Foundation (Grant No. L233024), Beijing Municipal Science & Technology Commission, Administrative Commission of Zhongguancun Science Park (Grant No. Z241100003524012), and Beijing High Innovation Plan (Contract No. 202604841443). Peiqi Duan was supported by China National Postdoctoral Program for Innovative Talents (Grant No. BX20230010) and China Postdoctoral Science Foundation (Grant No. 2023M740076).

References

1. Abedi, M., Jazizadeh, F., Huang, B., Battaglia, F.: Smart hvac systems—adjustable airflow direction. In: Proc. of Workshop of the European Group for Intelligent Computing in Engineering (2018)
2. Adrian, R.J., Westerweel, J.: Particle image velocimetry. No. 30 in Cambridge Aerospace Series, Cambridge university press (2011)
3. Akamine, M., Teramoto, S., Okamoto, K.: Microphones and three-dimensional background-oriented schlieren measurements of an ideally expanded supersonic jet. In: Proc. of AIAA SciTech 2024 Forum (2024)
4. Atcheson, B., Heidrich, W., Ihrke, I.: An evaluation of optical flow algorithms for background oriented schlieren imaging. Experiments in Fluids **46**(3), 467–476 (2009)
5. Atcheson, B., Ihrke, I., Heidrich, W., Tevs, A., Bradley, D., Magnor, M., Seidel, H.P.: Time-resolved 3d capture of non-stationary gas flows. ACM transactions on graphics (TOG) **27**(5), 1–9 (2008)
6. Bao, Y., Sun, L., Ma, Y., Wang, K.: Temporal-mapping photography for event cameras. In: Proc. of European Conference on Computer Vision (2024)
7. Bathel, B.F., Weisberger, J.M., Ripley, W.H., Jones, S.B.: Preparations for tomographic background-oriented schlieren at the 31-inch mach 10 wind tunnel. In: Proc. of AIAA Aviation 2022 Forum (2022)
8. Borer, D., Delbruck, T., Rösgen, T.: Three-dimensional particle tracking velocimetry using dynamic vision sensors. Experiments in Fluids **58**(12), 165 (2017)
9. Cai, H., Song, Y., Ji, Y., Li, Z., He, A.: Direct background-oriented schlieren tomography using radial basis functions. Optics Express **30**(11), 19100–19120 (2022)
10. Cai, S., Wang, Z., Fuest, F., Jeon, Y.J., Gray, C., Karniadakis, G.E.: Flow over an espresso cup: inferring 3D velocity and pressure fields from tomographic background oriented schlieren via physics-informed neural networks. Journal of Fluid Mechanics **915** (2021)
11. Dalziel, S.á., Hughes, G.O., Sutherland, B.R.: Whole-field density measurements by 'synthetic schlieren'. Experiments in Fluids **28**(4), 322–335 (2000)
12. Farnebäck, G.: Two-frame motion estimation based on polynomial expansion. In: Proc. of Scandinavian conference on Image analysis (2003)
13. Finateu, T., Niwa, A., Matolin, D., Tsuchimoto, K., Mascheroni, A., Reynaud, E., Mostafalu, P., Brady, F., Chotard, L., LeGoff, F., et al.: A 1280× 720 back-illuminated stacked temporal contrast event-based vision sensor with 4.86 μm pixels, 1.066 geps readout, programmable event-rate controller and compressive data-formatting pipeline. In: Proc. of IEEE International Solid-State Circuits Conference (2020)
14. Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A.J., Conradt, J., Daniilidis, K., et al.: Event-based vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(1), 154–180 (2020)
15. Gehrig, D., Gehrig, M., Hidalgo-Carrió, J., Scaramuzza, D.: Video to events: Recycling video datasets for event cameras. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020)
16. Gladstone, J.H., Dale, T.P.: Researches on the refraction, dispersion, and sensitiveness of liquids. Philosophical Transactions of the Royal Society of London **153**, 317–343 (1863)

17. Grauer, S.J., Mohri, K., Yu, T., Liu, H., Cai, W.: Volumetric emission tomography for combustion processes. Progress in Energy and Combustion Science **94**, 101024 (2023)
18. Grauer, S.J., Steinberg, A.M.: Fast and robust volumetric refractive index measurement by unified background-oriented schlieren tomography. Experiments in Fluids **61**(3), 80 (2020)
19. Grauer, S.J., Unterberger, A., Rittler, A., Daun, K.J., Kempf, A.M., Mohri, K.: Instantaneous 3D flame imaging by background-oriented schlieren tomography. Combustion and Flame **196**, 284–299 (2018)
20. Gupta, A., Miller, R., Bell, K., Langner, D., Agrawal, A.K.: Quantitative measurements in the exhaust flow of a rotating detonation combustor using rainbow schlieren deflectometry. Journal of Engineering for Gas Turbines and Power **146**(12), 121004 (2024)
21. Han, J., Asano, Y., Shi, B., Zheng, Y., Sato, I.: High-fidelity event-radiance recovery via transient event frequency. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)
22. Hargather, M.J., Settles, G.S.: Natural-background-oriented schlieren imaging. Experiments in Fluids **48**(1), 59–68 (2010)
23. He, Y., Zheng, Y., Xu, S., Liu, C., Peng, D., Liu, Y., Cai, W.: Neural refractive index field: Unlocking the potential of background-oriented schlieren tomography in volumetric flow visualization. Physics of Fluids **37**(1) (2025)
24. Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415 (2016)
25. Hu, W., Yang, L., Zhang, Y., Wang, P., Li, J.: Reconstruction refinement of hybrid background-oriented schlieren tomography. Physics of Fluids **36**(2), 023618 (2024)
26. Ihrke, I., Ziegler, G., Tevs, A., Theobalt, C., Magnor, M.A., Seidel, H.: Eikonal rendering: efficient light transport in refractive objects. ACM Transactions on Graphics **26**(3), 59 (2007)
27. Jasak, H.: Openfoam: Open source cfd in research and industry. International Journal of Naval Architecture and Ocean Engineering **1**(2), 89–94 (2009)
28. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
29. Léon, O., Donjat, D., Olchewsky, F., Desse, J.M., Nicolas, F., Champagnat, F.: Three-dimensional density field of a screeching under-expanded jet in helical mode using multi-view digital holographic interferometry. Journal of Fluid Mechanics **947**, A36 (2022)
30. Li, J., Meng, X., Xiong, Y., Jia, T., Pan, C., Wang, J.: Neural deflection field for sparse-view tomographic background oriented schlieren. Physics of Fluids **36**(12) (2024)
31. Li, Z., Müller, T., Evans, A., Taylor, R.H., Unberath, M., Liu, M.Y., Lin, C.H.: Neuralangelo: High-fidelity neural surface reconstruction. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)
32. Lichtsteiner, P., Posch, C., Delbruck, T.: A 128$\times$ 128 120 db 15 μs latency asynchronous temporal contrast vision sensor. IEEE Journal of Solid-State Circuits **43**(2), 566–576 (2008)
33. Lyu, Z., Cai, W., Liu, Y.: Event-triggered background-oriented schlieren: high-frequency visualization of a heated jet flow. Optics Letters **49**(10), 2565–2568 (2024)
34. Lyu, Z., Cai, W., Liu, Y.: An event-triggered background-oriented schlieren technique combined with dynamic projection using dynamic mirror device. Measurement Science and Technology **35**(10), 105302 (2024)

35. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In: Proc. of European Conference on Computer Vision (2020)
36. Molnar, J.P., LaLonde, E.J., Combs, C.S., Léon, O., Donjat, D., Grauer, S.J.: Forward and inverse modeling of depth-of-field effects in background-oriented schlieren. AIAA Journal **62**(11), 4316–4329 (2024)
37. Molnar, J.P., Venkatakrishnan, L., Schmidt, B.E., Sipkens, T.A., Grauer, S.J.: Estimating density, velocity, and pressure fields in supersonic flows using physics-informed BOS. Experiments in Fluids **64**(1), 14 (2023)
38. Nicolas, F., Todoroff, V., Plyer, A., Le Besnerais, G., Donjat, D., Micheli, F., Champagnat, F., Cornic, P., Le Sant, Y.: A direct approach for instantaneous 3D density field reconstruction from background-oriented schlieren (BOS) measurements. Experiments in Fluids **57**(1), 13 (2016)
39. OpenCFD Ltd: Openfoam. `https://www.openfoam.com/`, accessed: 2026-06-25
40. Pan, B., Asundi, A., Xie, H., Gao, J.: Digital image correlation using iterative least squares and pointwise least squares for displacement field and strain field measurements. Optics and Lasers in Engineering **47**(7-8), 865–874 (2009)
41. Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021)
42. Patankar, S.V.: Airflow and cooling in a data center. Journal of Heat Transfer **132**(7), 073001 (2010)
43. Raffel, M., Richard, H., Meier, G.: On the applicability of background oriented optical tomography for large scale aerodynamic investigations. Experiments in Fluids **28**(5), 477–481 (2000)
44. Raffel, M.: Background-oriented schlieren (BOS) techniques. Experiments in Fluids **56**(3), 60 (2015)
45. Raffel, M., Willert, C.E., Scarano, F., Kähler, C.J., Wereley, S.T., Kompenhans, J.: Particle image velocimetry: a practical guide. Springer (2018)
46. Rebecq, H., Gehrig, D., Scaramuzza, D.: ESIM: an open event camera simulator. In: Proc. of Conference on Robot Learning (2018)
47. Schmidt, B.E., Bathel, B.F., Grauer, S.J., Hargather, M.J., Heineck, J.T., Raffel, M.: Twenty-five years of background-oriented schlieren: advances and novel applications. AIAA journal **63**(12), 5028–5058 (2025)
48. Settles, G.S.: Schlieren and shadowgraph techniques: visualizing phenomena in transparent media. Springer (2001)
49. Settles, G.S., Hargather, M.J.: A review of recent developments in schlieren and shadowgraph techniques. Measurement Science and Technology **28**(4), 042001 (2017)
50. Settles, G.S., Liberzon, A.: Schlieren and bos velocimetry of a round turbulent helium jet in air. Optics and Lasers in Engineering **156**, 107104 (2022)
51. Shiba, S., Hamann, F., Aoki, Y., Gallego, G.: Event-based background-oriented schlieren. IEEE Transactions on Pattern Analysis and Machine Intelligence **46**(4), 2011–2026 (2023)
52. Sommersel, O., Bjerketvedt, D., Christensen, S., Krest, O., Vaagsaether, K.: Application of background oriented schlieren for quantitative measurements of shock waves from explosions. Shock Waves **18**(4), 291–297 (2008)
53. Takahashi, H., Oki, J., Hirotani, T., Taguchi, H.: Tomographic reconstruction of hypersonic aircraft flowfields in large-scale wind tunnel experiments. AIAA Journal **62**(3), 869–881 (2024)

54. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. In: Advances in Neural Information Processing Systems (2020)
55. Teh, A., Ali, W.H., Rapp, J., Mansour, H.: Indoor airflow imaging using physics-informed schlieren tomography. In: Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2025)
56. Teh, A., O'Toole, M., Gkioulekas, I.: Adjoint nonlinear ray tracing. ACM Transactions on Graphics **41**(4), 1–13 (2022)
57. Venkatakrishnan, L., Meier, G.: Density measurements using the background oriented schlieren technique. Experiments in Fluids **37**(2), 237–247 (2004)
58. Venkatakrishnan, L.: Density measurements in an axisymmetric underexpanded jet by background-oriented schlieren technique. AIAA journal **43**(7), 1574–1579 (2005)
59. Wang, Y., Idoughi, R., Heidrich, W.: Stereo event-based particle tracking velocimetry for 3d fluid flow reconstruction. In: Proc. of European Conference on Computer Vision (2020)
60. Wildeman, S.: Real-time quantitative schlieren imaging by fast fourier demodulation of a checkered backdrop. Experiments in Fluids **59**(6), 97 (2018)
61. Willert, C.E., Klinner, J.: Event-based imaging velocimetry: an assessment of event-based cameras for the measurement of fluid flows. Experiments in Fluids **63**(6), 101 (2022)
62. Xiong, Y., Kaufmann, T., Noiray, N.: Towards robust bos measurements for axisymmetric flows. Experiments in Fluids **61**(8), 178 (2020)
63. Xue, T., Rubinstein, M., Wadhwa, N., Levin, A., Durand, F., Freeman, W.T.: Refraction wiggles for measuring fluid depth and velocity from video. In: Proc. of European Conference on Computer Vision (2014)
64. Zhao, B., Levis, A., Connor, L., Srinivasan, P.P., Bouman, K.L.: Single view refractive index tomography with neural fields. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024)

Event-based Sparse-view Background-Oriented Schlieren Tomography

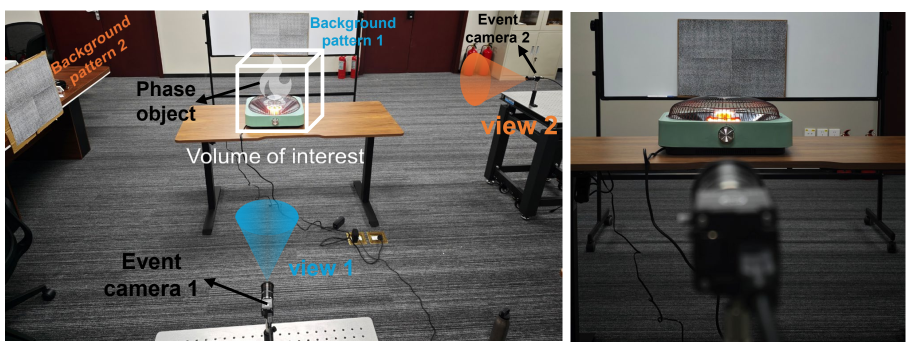
Supplemental Material

Xinyu Zhou, Shihao Hu, Peiqi Duan*, Chao Xu, and Boxin Shi*

Peking University
{zhouxiny, duanqi0001, shiboxin}@pku.edu.cn, xuchao@cis.pku.edu.cn

This supplementary material provides additional real-data acquisition and implementation details, followed by extended experimental analyses. We also provide a supplementary video that introduces the BOS optical setup, and presents the reconstructed 4D density fields on both synthetic and real data.

7 Real Data Acquisition Details



Scene View 1

Fig. 9: Real-data acquisition setup.

As described in the main paper, to validate event-based BOS tomography without frame-camera assistance, we build a dual-view event-camera setup to capture thermal convection phenomena. As illustrated in Fig. 9, the two cameras are placed horizontally with approximately orthogonal viewing directions, and

* Corresponding authors.
Xinyu Zhou and Chao Xu are with State Key Laboratory of General Artificial Intelligence, School of Intelligence Science and Technology, Peking University. Peiqi Duan, Shihao Hu, and Boxin Shi are with State Key Laboratory of Multimedia Information Processing and National Engineering Research Center of Visual Technology, School of Computer Science, Peking University.

a random-dot background is positioned behind the volume of interest for each view. Both cameras are Prophesee EVK4 event cameras equipped with IMX636 sensors and 35 mm lenses operated at an aperture of $f/12$. The distance from each camera to the phase object is approximately 1.7 m, and the distance to the background is approximately 3.2 m. The intrinsic and extrinsic parameters of the two cameras are calibrated in advance using a blinking checkerboard [2,6]. Before each recording, we obtain the background patterns by switching the illumination on and off [1,3]. In practice, this simple procedure provides sufficiently accurate background references for airflow reconstruction. We capture thermal convection above several heat sources, including an electric heater, a hair dryer, and a candle flame. Each recording lasts approximately 10 s.

8 Additional Implementation Details

Sampling strategy. In each training iteration, we sample two adjacent timestamps to form a time window. We draw these timestamps uniformly at random, and since we target high-speed flow scenarios, we constrain the maximum time window to 0.02 s. For each time window, we sample pixel locations from the accumulated event image using a mixture of nonzero pixels (where events occurred) and zero-event pixels for negative sampling [5]. We set the negative sampling ratio to 0.5, *i.e.*, half of the sampled pixels come from zero-event locations.

Real-data loss configuration. The render loss requires the event thresholds of the camera to be known *a priori* or calibrated beforehand. In real cameras, however, accurately determining these threshold values is sometimes difficult. We therefore also consider the following normalized form:

$$\mathcal{L}_{\text{render}}^{\text{norm}} = \mathbb{E}_{(\boldsymbol{\xi},t_1,t_2)}\left[\left\|\frac{\Delta\widehat{L}(\boldsymbol{\xi};t_1,t_2)}{\left\|\Delta\widehat{L}(\boldsymbol{\xi};t_1,t_2)\right\|_1} - \frac{E(\boldsymbol{\xi};t_1,t_2)}{\left\|E(\boldsymbol{\xi};t_1,t_2)\right\|_1}\right\|_2^2\right], \tag{22}$$

which also performs well in practice. For the real data used in our experiments, the position, size, and temperature of the heat sources are not accurately calibrated. Therefore, we do not apply the inlet loss to real data.

9 Additional Experimental Analyses

This section provides additional experiments that further examine the robustness and operating range of the proposed method. Unless otherwise specified, these experiments are conducted on the *Hot airflow* synthetic sequence.

9.1 Velocity Visualization

We visualize the auxiliary velocity field to illustrate the behavior of the physics-informed regularization beyond the density reconstruction. Fig. 10 shows the

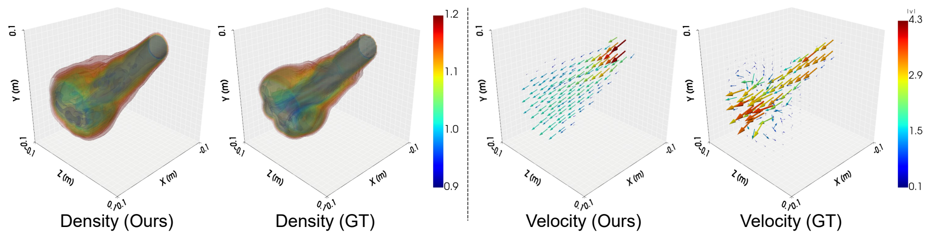


Fig. 10: Reconstructed density and velocity fields produced by our method, together with the ground truth. The density field is visualized using isosurfaces. The velocity field is visualized using arrow glyphs, where arrow orientations indicate flow directions, and arrow lengths and colors encode velocity magnitudes.

reconstructed density and the implicitly estimated velocity field together with the ground truth. Because BOS does not directly supervise velocity, the estimated velocity still differs noticeably from the ground truth; nevertheless, it captures the main active regions and the overall flow direction. This qualitative behavior is consistent with the ablation in Table 2 of the main paper, where removing the physics-informed regularization increases the density RMSE from 0.0144 to 0.0171 and decreases PCC from 0.971 to 0.958. It supports our use of the auxiliary flow variables as regularization for density recovery.

9.2 Additional Ablations

We conduct additional ablation analyses on the *Hot airflow* sequence under a dual-view setup and report the results in Tab. 4. Each “W/o” variant removes one physics residual from the full model. The SIREN [7] and Instant-NGP [4] variants replace the default coordinate-MLP representation while keeping the remaining setting unchanged. Removing any physics residual reduces reconstruction quality, indicating that the mass, momentum, and heat constraints provide complementary regularization. The default coordinate-MLP representation also outperforms the tested SIREN and Instant-NGP variants.

Table 4: Additional ablation study of physics residuals and neural representations on the *Hot airflow* sequence (2 views).

	W/o Mass	W/o Mom.	W/o Heat	SIREN	Instant-NGP	Ours
RMSE↓	0.0179	0.0164	0.0169	0.0163	0.0207	0.0144
PCC↑	0.958	0.962	0.959	0.961	0.942	0.971

9.3 View-count and Camera-layout Analysis

We next analyze how view count and camera layout influence sparse-view reconstruction. Our standard two-view setting uses an approximately 90° angular baseline. We further evaluate 4/8-view configurations with equally spaced views on a semicircle and two-view layouts with 30°, 60°, and 90° baselines, and report the results in Tab. 5. Additional views provide modest but consistent gains: 4V and 8V reach RMSE/PCC of 0.0134/0.974 and 0.0123/0.978, respectively. Among the tested two-view layouts, the orthogonal 90° configuration performs best, suggesting that a near-orthogonal arrangement most effectively reduces spatial ambiguity in the sparse-view setting.

Table 5: Sparse-view reconstruction with different view counts and camera layouts.

	2V (30°)	2V (60°)	2V (90°)	4V	8V
RMSE↓	0.0179	0.0155	0.0144	0.0134	0.0123
PCC↑	0.955	0.966	0.971	0.974	0.978

9.4 Quasi-linear Ray-tracing Stress Test

Our ray tracing uses a quasi-linear approximation, where rays are traced with a straight-line proxy to reduce the cost of dense nonlinear ray integration. Compared with dense Hamilton integration, this approximation greatly reduces training time while preserving accuracy in our experiments. To evaluate its operating range under strong refractive-index gradients, we resimulate the data by scaling the Gladstone–Dale constant to $4G$ and $10G$. Quasi-linear tracing remains accurate at $4G$ (RMSE/PCC = 0.0142/0.972), indicating a broad operating range, but fails at $10G$ (0.0403/0.801). Dense Hamilton integration recovers 0.0148/0.969 at $10G$, but increases training time from 5 h to 20 h.

9.5 Robustness to Event Noise

We evaluate the sensitivity of our method to event noise. Specifically, we consider two common sources of event noise: contrast-threshold uncertainty and hot pixels. For threshold uncertainty, we add per-pixel Gaussian perturbations with standard deviation 0.02 to the contrast thresholds, obtaining RMSE/PCC of 0.0146/0.970, close to the original setting. In preprocessing, hot pixels are detected by abnormally high event counts and removed before reconstruction. To simulate the effect of this removal, we randomly mask event pixels in the input; masking either 10 or 1000 event pixels produces nearly unchanged results. These results indicate that our method is robust to the tested event-noise perturbations.

9.6 Reference-background Sensitivity

As in standard BOS, our method requires a reference background image and assumes that this reference is geometrically aligned with the event-camera measurements in the image plane. Since BOS displacements are usually subtle, small errors in the background-gradient location can bias the inferred refraction. We therefore evaluate the sensitivity to both reference-image brightness noise and spatial misalignment. Adding Gaussian noise with standard deviation 0.03 to I_{ref} has little impact, yielding RMSE/PCC of 0.0148/0.969. We then shift I_{ref} to test alignment errors: a 0.2 px shift still gives effective reconstruction with RMSE/PCC of 0.0154/0.966, whereas a larger 0.5 px shift fails in our stress test. These results indicate that the method is robust to moderate reference brightness noise, but accurate reference alignment is important for reliable BOS reconstruction.

9.7 Scope of Thermal-convection Prior

The physical constraints used for regularization should be consistent with the governing physics of the observed target fluid. We further examine this issue by testing the thermal-convection constraints on a flow outside the target regime of our main experiments. We simulate a CH_4 combustion scene using the same synthetic-data pipeline and treat it as an out-of-scope case for the thermal-convection assumptions. This experiment is intended to quantify the effect of applying mismatched physical constraints rather than to claim that the same constraints are universally applicable. On this combustion scene, applying the mismatched thermal-convection constraints reduces the reconstruction PCC from 0.950 to 0.934. The result suggests that the event-BOS rendering supervision can still provide useful reconstruction cues, while the physics-informed regularizer should be adapted to the governing physics of the observed flow.

References

1. Bao, Y., Sun, L., Ma, Y., Wang, K.: Temporal-mapping photography for event cameras. In: Proc. of European Conference on Computer Vision (2024)
2. Dominguez-Morales, M.J., Jimenez-Fernandez, A., Jimenez-Moreno, G., Conde, C., Cabello, E., Linares-Barranco, A.: Bio-inspired stereo vision calibration for dynamic vision sensors. IEEE Access **7**, 138415–138425 (2019)
3. Han, J., Asano, Y., Shi, B., Zheng, Y., Sato, I.: High-fidelity event-radiance recovery via transient event frequency. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)
4. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics **41**(4), 102:1–102:15 (2022)
5. Rudnev, V., Elgharib, M., Theobalt, C., Golyanik, V.: EventNeRF: Neural radiance fields from a single colour event camera. In: Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)

6. Salah, M., Ayyad, A., Humais, M., Gehrig, D., Abusafieh, A., Seneviratne, L., Scaramuzza, D., Zweiri, Y.: E-calib: A fast, robust, and accurate calibration toolbox for event cameras. IEEE Transactions on Image Processing **33**, 3977–3990 (2024)
7. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. In: Advances in Neural Information Processing Systems (2020)