2736 Vol. 51, No. 10 / 15 May 2026 / Optics Letters Letter

Optics Letters

# Autofocus in digital holography with time reversal and depth-dependent sampling

Ye Liu,[1] Bing-Zhong Wang,[2] Edmund Lam,[3] and Haiyan Ou[1,2,*]

[1] Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen 518000, China
[2] School of Physics, University of Electronic Science and Technology of China, Chengdu 611731, China
[3] Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China
*ouhaiyan@uestc.edu.cn

Received 27 January 2026; revised 19 April 2026; accepted 20 April 2026; posted 21 April 2026; published 4 May 2026

**Accurate object localization is critical for robust autofocus in digital holography (DH). Here, we present a time-reversal (TR) symmetry-enabled autofocus strategy integrated with depth-dependent randomized sparse sampling. Our approach is grounded in the low-rank structure of the TR operator. This inherent property, arising from the dynamical constraints of the wave equation, provides a physical foundation for sparse sampling without compromising subspace separation. By using independent, uniformly distributed test point sets at each axial plane, the method effectively decorrelates sampling-induced clutter and acts as an approximately unbiased Monte Carlo estimator. This analytical, training-free solution preserves the inherent orthogonality of signal and noise subspaces, ensuring high-fidelity localization even at ultra-low sampling rates (e.g., 0.25*N*). Simulations and experiments demonstrate that the proposed method achieves high-resolution autofocus with significant computational savings, paving the way for real-time dynamic autofocus in diverse sensing environments.** 

https://doi.org/10.1364/OL.591504

Digital holography (DH) is a non-invasive three-dimensional (3D) imaging technique that encodes the complex wavefront of an object into a two-dimensional (2D) hologram [1]. Since the development of digital holography (combining holography with digital recording/reconstruction), DH has found broad applications in biomedical imaging, 3D shape/deformation measurement, particle field analysis, quantitative phase imaging, and material characterization, among others [2,3].

In DH, hologram recording is a forward problem while reconstruction is an inverse problem; the core challenge of the latter—reconstructing an unknown image from a hologram—lies in accurately localizing the object, as imprecise localization causes blurred or distorted reconstructions. To address this issue, various autofocus algorithms have been developed, including sharpness-metric-based approaches (e.g., amplitude/entropy/gradient criteria) [4–6], edge-sparsity/structure-based criteria [7–9], frequency-domain methods [10], and data-driven approaches such as deep learning [11–13].

In this work, we propose an autofocus strategy based on time-reversal (TR) symmetry and depth-dependent randomized sparse sampling. The key motivation is physical: wave propagation governed by Maxwell's equations exhibits TR symmetry in lossless reciprocal media, which is the core mechanism underpinning TR Mirror (TRM) focusing [14]. TRM enables high-resolution wavefront focusing and imaging in complex media [15,16]. Notably, the holographic forward model and numerical back-propagation naturally form a TR pair, which directly leverages this reciprocity-driven TR capability. By leveraging TRM focusing and subspace separation, we have previously demonstrated that TR-based methods (e.g., Time Reversal-Multiple Signal Classification, TR-MUSIC) can provide high-resolution axial localization [17,18]. These preliminary studies highlight two critical observations: (1) the autofocus inverse problem in DH admits a sparsity prior normally and (2) holographic datasets are characterized by significant inherent redundancy. The core challenge thus lies in how to eliminate redundant data while effectively leveraging the sparsity prior inherent to this inverse problem. We tackle this key issue by leveraging depth-dependent randomized sparse sampling, which concurrently reduces holographic data redundancy and exploits the sparsity prior inherent to the autofocus inverse problem.

Optical scanning holography (OSH), a representative DH acquisition scheme, is adopted here [19,20]. Figure 1 depicts its setup: a laser beam is split by beam splitter (BS1) and follows two paths: one traverses mirror M1, pupil $p_1(x, y) = 1$, and lens L2; the other is frequency-modulated via an acousto-optic frequency shifter (AOFS), then passes mirror M2, pupil $p_2(x, y) = \delta(x, y)$, and lens L3. The two beams recombine at BS2 and scan the object via an X-Y scanning mirror. The signal is detected by a photodetector (PD) and demodulated into a complex hologram.

For a 3D object discretized into $N$ axial layers, the generated hologram can be expressed as:

$$g(x, y) \approx \sum_{i=1}^{N} \left( |O(x, y; z_i)|^2 \otimes h(x, y; z_i) \right), \tag{1}$$

where $O(x, y; z_i)$ is the complex amplitude of the $i$-th layer, $h(x, y; z)$ is the spatial impulse response, $\otimes$ denotes convolution, and $z_i$ is the axial distance of the $i$-th layer. $G = \mathcal{F}\{g(x, y)\}$ is the Fourier transform of the hologram. To reconstruct the image at the $l$-th layer, prior knowledge of $z_l$ is required and the

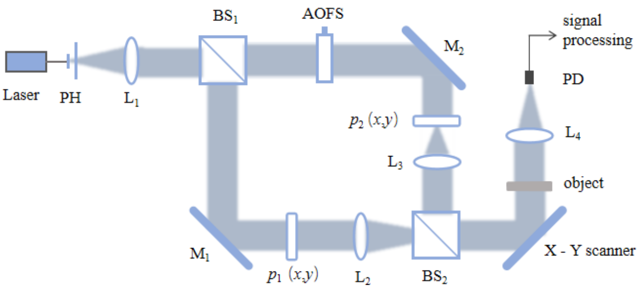



**Fig. 1.** Schematic diagram of the optical scanning holography (OSH) system. AOFS, acousto-optic frequency shifter; BS, beam splitter; L, lens; M, mirror; PD, photodetector; PH, pin hole; $p(x, y)$, pupil.

reconstructed image is given by:

$$I_{\mathrm{out}}(x, y; z_l) = |O(x, y; z_l)|^2 \otimes h(x, y; z_l) \otimes h^*(x, y; z_l) + \sum_{i \neq l} |O(x, y; z_i)|^2 \otimes h(x, y; z_i) \otimes h^*(x, y; z_l), \quad (2)$$

where $*$ denotes conjugation. The first term corresponds to the focused contribution of the target layer, while the second term represents the defocus noise from other layers.

In a lossless homogeneous medium, solutions to Maxwell's equations remain valid under time-reversal ($t \to -t$), a symmetry that gives rise to wave propagation reciprocity, the foundational mechanism for refocusing energy to its origin [21]. A TRM—the physical realization of TR focusing—relies on this reciprocity: it uses a transceiver array to record the outgoing wavefield from a source, time-reverses it, and retransmits it [22]. Specifically, this TRM core (an array of receivers) aligns naturally with digital holography: in DH, every CCD/CMOS pixel functions as a miniature receiver, which captures object wavefront information via interferometric recording. This inherent "micro-receiver array" structure means DH can natively leverage the TRM focusing mechanism. To quantify this TRM-based focusing mechanism mathematically, we define the corresponding TR matrix as follows:

$$T_{\mathrm{TR}} = \mathcal{F}^{-1}\{GG^H\}, \quad (3)$$

where $H$ denotes conjugate transpose. Performing singular value decomposition (SVD) on this TR matrix, we obtain:

$$T_{\mathrm{TR}} = U \Sigma U^H, \quad (4)$$

where $U = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]$ is a unitary matrix whose columns are orthonormal singular vectors (eigenvectors of the Hermitian matrix $T_{\mathrm{TR}}$), and $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_N)$ contains the singular values in descending order.

The dominant singular vectors span the signal subspace (encoding target-related wavefield information), while the remaining vectors form the noise subspace (containing background and measurement noise). This distinct separation of subspaces provides a foundational framework for high-resolution localization methods—following the TR-MUSIC idea, we leverage this subspace orthogonality to compute pseudo-spectra along two separable dimensions $x$ and $y$ [17]:

$$P_x(X_p, z_i) = \frac{\|u_1(X_p, z_i)\|^2}{\sum_{m=M+1}^{N} \left|u_x(m)^T u_1^*(X_p, z_i)\right|^2}, \quad (5)$$

$$P_y(X_p, z_i) = \frac{\|u_2(X_p, z_i)\|^2}{\sum_{m=M+1}^{N} \left|u_y^*(m)^T u_2(X_p, z_i)\right|^2}, \quad (6)$$

where $X_p$ represents the horizontal position of the test sampling point, and $M$ is the dimension of the signal subspace. The combined location is therefore $P(X_p, z_i) = P_x(X_p, z_i) \cdot P_y(X_p, z_i)$. The axial localization metric at depth $z_i$ is obtained by summing pseudo-spectrum values over spatial test positions:

$$P(z_i) = \sum_{p=1}^{V} P(X_p, z_i), \quad (7)$$

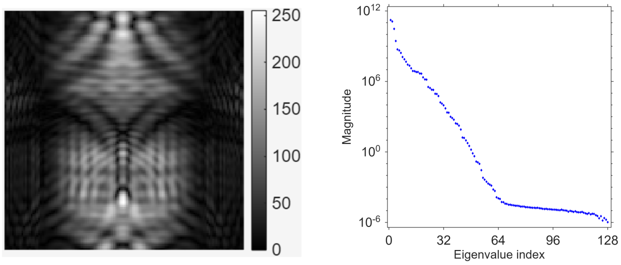
where $V$ represents the number of sampling points per reconstruction plane $z_i$. It can be deduced from Eq. (7) that TR-MUSIC inherently relies on dense, point-by-point scanning to localize objects, imposing substantial computational overhead. To alleviate this burden, our earlier efforts explored conventional scanning strategies including diagonal/Lissajous scanning, which delivered initial yet limited improvements: diagonal scanning restricts test points to a single line, severely undersampling off-diagonal spatial responses [17]; while denser, Lissajous scanning suffers from periodic clustering and non-uniform spatial coverage, causing some regions to be over-weighted and others under-represented. This distribution bias distorts the pseudo-spectrum $P(z_i)$ and introduces spurious peaks, particularly under noisy or multi-depth interference conditions [18].

Building on our prior observation that the autofocus inverse problem exhibits a sparsity prior and holographic datasets are highly redundant, we address these limitations by adopting randomized sparse sampling—a strategy that selects a small subset of test points with approximately uniform spatial probability, offering two advantages: (I) From a numerical-integration perspective, uniformly random sampling provides an approximately unbiased Monte Carlo estimator of the full-grid sum for $P(z_i)$ [23]: when properly scaled by the full-grid-to-sampled-point ratio, while preserving statistical consistency under sparse depth sampling; (II) From a compressed-sensing perspective, random sampling reduces correlation between sampling patterns and signal structure [24], enabling efficient probing of the TR pseudo-spectrum at sparse points without reconstructing the full field.

Depth-dependent artifacts occur when sampling is shared across depths. We address this by employing **independent random sampling sets at different** $z_i$: $\Omega_i \neq \Omega_j \ (i \neq j)$, where $\Omega_i$ denotes the sampling set at depth $z_i$ containing $V$ unique points drawn without replacement from the $N \times N$ grid. The axial metric in Eq. (7) can therefore be computed by summing probabilities over the depth-dependent sparse set:

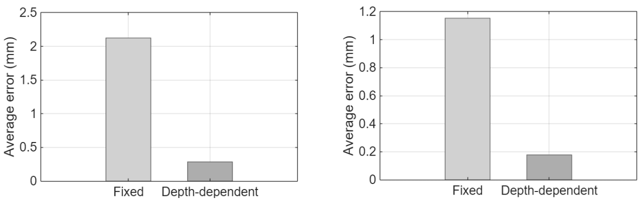
$$P(z_i) = \sum_{(x_p, y_p) \in \Omega_i} P((x_p, y_p), z_i). \quad (8)$$

This design thus ensures uniform spatial coverage, mitigates sampling bias, and maintains low computational costs. While prior compressed sensing achieves image reconstruction at 0.24% compression [25], our autofocus only requires axial localization—not full imaging—enabling even lower sampling. We therefore adopt a random sampling rate of 0.25$N$, yielding 0.2% compression for 128 × 128 images and 0.05% for larger sizes such as 500 × 500, effectively balancing efficiency with focusing stability.




(a) (b)

**Fig. 2.** Simulated hologram (a), and eigenvalues of the TR matrix (b).




(a) (b)

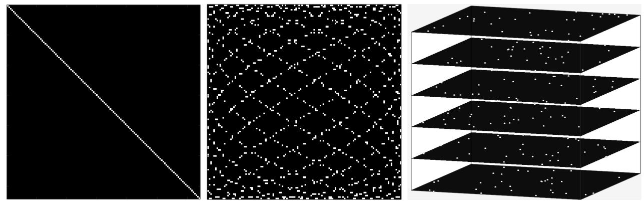
**Fig. 3.** Localization error analysis: fixed vs. depth-dependent sampling at $z_1$ (a) and $z_2$ (b).

To systematically quantify these performance gains and validate the robustness of our approach, simulations were performed on an AMD Ryzen 7 7800X3D processor (4.20 GHz, 32 GB RAM) with a laser wavelength of 632.8 nm. A composite object with two axial layers at $z_1 = 20$ mm and $z_2 = 30$ mm was used, with each layer containing 128 × 128 grayscale patterns. The corresponding hologram is shown in Fig. 2(a), and the rapid decay of eigenvalues in Fig. 2(b) confirms the sparsity of the holographic field.

To evaluate the robustness of our sampling strategy, we compare the localization accuracy of fixed versus depth-dependent randomized sampling across different axial planes. As shown in Fig. 3, fixed sampling yields significant localization bias and inconsistent performance across depths. In contrast, our depth-dependent strategy effectively suppresses depth-induced artifacts, reducing the average error by over 80% at both $z_1$ and $z_2$. This confirms the superior robustness and accuracy of the randomized sampling scheme in diverse imaging planes.

Five autofocus strategies were evaluated over the axial range 15–35 mm with a step size of 0.05 mm: edge detection (ED), Tanimoto coefficient (TC), diagonal scanning (DS, V = N = 128), Lissajous scanning (LS, V = 10$N$ = 1280), and the proposed randomized sparse sampling (RS, V = 0.25$N$ = 32). These abbreviations are used consistently throughout (Fig. 4). Figure 4 illustrates the sampling masks of DS, LS, and the proposed RS strategies, where points denote test positions. Unlike DS and LS, the RS mask varies independently across depth planes.

Table 1 and Fig. 5 summarize the comparison results. LS gives the narrowest axial response (FWHM = 0.06 mm), while the proposed RS achieves a comparable resolution (FWHM = 0.07 mm) with substantially lower computation time—96% faster than LS and 69% faster than DS. By contrast, DS shows a broader response (FWHM = 0.29 mm), and ED/TC yield much wider peaks with less accurate axial localization. In addition, RS achieves the better peak-to-sidelobe ratio (PSLR) and peak-



**(a)** DS **(b)** LS **(c)** RS

**Fig. 4.** Sampling masks for different strategies (points denote test positions).

**Table 1. Simulation Performance Comparison**

| Method | ED | TC | DS | LS | RS |
|---|---|---|---|---|---|
| Time (s) | 6.7 | 5.4 | 12.5 | 96.4 | 3.8 |
| $z_1$ (mm) | 20.32 | 20.65 | 20.01 | 20 | 20 |
| $z_2$ (mm) | 30.44 | 29.50 | 29.98 | 30 | 30 |
| Sampling points | / | / | 128 | 1280 | 32 |
| FWHM (mm) | 2.17 | 3.29 | 0.29 | 0.06 | 0.07 |
| PBR (dB) | 0.85 | / | 13.68 | 88.59 | 309.30 |
| PSLR (dB) | −9.67 | / | 8.22 | 18.93 | 27.60 |

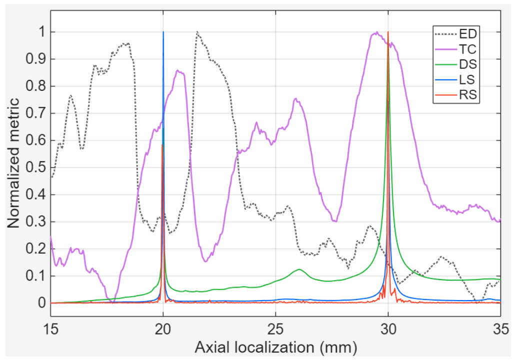



**Fig. 5.** Simulated autofocus results with different methods.

to-background ratio (PBR) values. These results indicate that RS preserves the sharp TR-MUSIC localization behavior while greatly reducing computation.

To validate the method experimentally, we used a 500 × 500-pixel hologram of a dual-layer object at $z_1$ = 87 cm and $z_2$ = 107 cm. The sampling settings were V = $N$ = 500 for DS, V = 10$N$ = 5000 for LS, and V = 0.25$N$ = 125 for RS, with autofocus performed over 80–115 cm at a step size of 0.175 cm.

Figure 6 displays the hologram (a) and the eigenvalues of the TR matrix (b). The rapid decay of the eigenvalues directly evidences the sparsity of the experimental holographic signal. Figure 7 and Table 2 summarize the experimental autofocus results. LS gives the sharpest axial response (FWHM = 0.24 cm), while RS achieves comparable resolution (FWHM = 0.25 cm) with 97% lower computation time (58.3 s vs. 2239.7 s), which is further reduced to 0.56 s via CPU parallelization. DS exhibits a broader response (FWHM = 0.31 cm), and ED/TC yield much wider profiles (FWHM = 6.07 cm and 12.07 cm, respectively). Moreover, the relatively high PSLR and PBR values of RS indicate effective sidelobe suppression and strong peak prominence over the background. Focusing errors for LS and RS are below 0.05 cm for both layers, compared to 0.11 cm for DS and 0.79–3.65 cm for ED/TC. These results confirm the su-

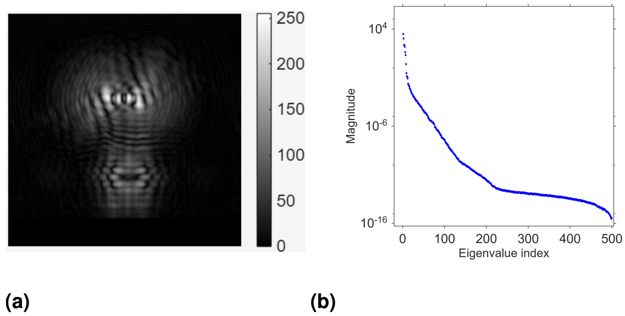



**Fig. 6.** (a) Hologram and (b) eigenvalues of the TR matrix.

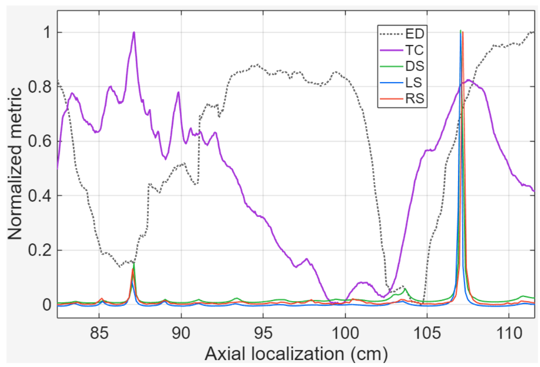



**Fig. 7.** Experimental autofocus results with different methods.

**Table 2. Experimental Performance Comparison**

| Method | ED | TC | DS | LS | RS |
|---|---|---|---|---|---|
| Time (s) | 14.2 | 14.9 | 257.5 | 2239.7 | 58.3 |
| $z_1$ (cm) | 87.83 | 85.65 | 87.11 | 86.96 | 86.96 |
| $z_2$ (cm) | 107.66 | 110.65 | 107.04 | 106.97 | 107.04 |
| Sampling points | / | / | 500 | 5000 | 125 |
| FWHM (cm) | 6.07 | 12.07 | 0.31 | 0.24 | 0.25 |
| PBR (dB) | 1.61 | 2.17 | 118.53 | 226.62 | 170.49 |
| PSLR (dB) | −0.42 | 0.47 | 25.64 | 34.18 | 32.77 |

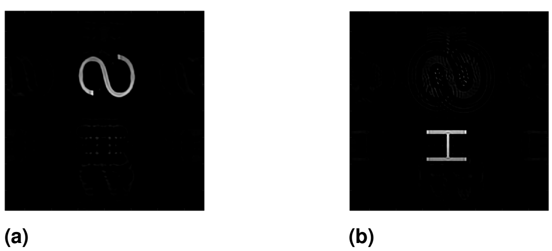



**Fig. 8.** Reconstructed images at (a) $z_1 = 86.96$ cm, and (b) $z_2 = 107.04$ cm.

perior axial localization accuracy and reliability of TR-based approaches.

The corresponding reconstructed sections are shown in Figs. 8(a) and 8(b), respectively. Together with the preceding performance comparisons, these reconstructed results confirm the feasibility of our TR symmetry-enabled autofocus approach.

In conclusion, we have presented a physics-guided sparse TR-MUSIC framework for high-fidelity holographic autofocus. By leveraging the intrinsic dynamical constraints of the wave equation, we reveal the low-rank structure of the TR matrix as the fundamental physical basis for sparse sampling. Our depth-dependent randomization strategy effectively suppresses coherent artifacts by decorrelating sampling noise along the axial direction, ensuring that subspace orthogonality remains robust even at a $0.25N$ sampling rate. Unlike iterative compressed sensing or data-driven deep learning, this analytical solution is training-free and offers full physical interpretability. Experimental results confirm that its inherent parallelism enables sub-second processing (0.56 s) on a standard multi-core CPU, while future GPU acceleration holds significant potential for real-time 3D sensing of biological samples and other complex targets in diverse dynamic environments.

**Funding.** Sichuan Science and Technology Program (2026NSFSC0398).

**Acknowledgment.** During manuscript preparation, the authors used ChatGPT to enhance text clarity and fluency. All content was subsequently reviewed and edited, and the authors assume full responsibility for the final publication.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying this study are available from the corresponding author upon reasonable request.

## REFERENCES

1. B. Javidi, A. Carnicer, A. Anand, *et al.*, Opt. Express **29**, 35078 (2021).
2. M. Fratz, T. Seyler, A. Bertz, *et al.*, Light: Adv. Manuf. **2**, 283 (2021).
3. J. Rosen, S. Alford, B. Allan, *et al.*, Appl. Phys. B: Lasers Opt. **130**, 166 (2024).
4. F. Dubois, C. Schockaert, N. Callens, *et al.*, Opt. Express **14**, 5895 (2006).
5. M. Liebling and M. Unser, J. Opt. Soc. Am. A **21**, 2424 (2004).
6. J. Liu, Y. Liu, H. Ou, *et al.*, Opt. Lett. **50**, 5342 (2025).
7. Y. Zhang, H. Wang, Y. Wu, *et al.*, Opt. Lett. **42**, 3824 (2017).
8. Z. Ren, N. Chen, and E. Y. Lam, Opt. Lett. **42**, 1720 (2017).
9. Z. Ren, E. Y. Lam, and J. Zhao, Opt. Express **28**, 14712 (2020).
10. S. Oh, C.-Y. Hwang, I. K. Jeong, *et al.*, Opt. Express **22**, 28926 (2014).
11. Y. Rivenson, Y. Zhang, H. Günaydın, *et al.*, Light. Sci. Appl. **7**, 17141 (2018).
12. T. Zeng, Y. Zhu, and E. Y. Lam, Opt. Express **29**, 40572 (2021).
13. J. Kim, Y. Kim, H. S. Lee, *et al.*, Nat. Commun. **16**, 4840 (2025).
14. M. Fink, J. Phys. D: Appl. Phys. **26**, 1333 (1993).
15. G. Lerosey, J. de Rosny, A. Tourin, *et al.*, Science **315**, 1120 (2007).
16. A. P. Mosk, A. Lagendijk, G. Lerosey, *et al.*, Nat. Photonics **6**, 283 (2012).
17. H. Ou, Y. Wu, E. Y. Lam, *et al.*, Opt. Express **26**, 3756 (2018).
18. J. Liu, H. Ou, H. Wang, *et al.*, Chin. Opt. Lett. **22**, 080501 (2024).
19. B. D. Duncan and T.-C. Poon, J. Opt. Soc. Am. A **9**, 229 (1992).
20. T.-C. Poon, J. Opt. Soc. Korea **13**, 406 (2009).
21. M. Fink, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 555 (1992).
22. M. Fink, J. Physics: Conf. Ser. **118**, 012001 (2008).
23. W. K. T. Cho and Y. Y. Liu, Phys. A: Stat. Mech. Appl. **506**, 170 (2018).
24. D. L. Donoho, IEEE Trans. Inf. Theory **52**, 1289 (2006).
25. V. Kravets and A. Stern, Opt. Lasers Eng. **171**, 107818 (2023).