OPTICA

Property-driven analysis of glasses for data storage via femtosecond laser writing

Takashi Lawson, Rokas Drevinskas, Masaaki Sakakura, Charles E. Whittaker, Austin Donnelly, Benn Thomsen, and Richard Black*

Microsoft Research, 198 Cambridge Science Park, Milton Road, Cambridge CB4 0AB, UK
**Richard.Black@microsoft.com*

Received 11 February 2026; revised 18 March 2026; accepted 19 March 2026; published 2 April 2026

The growing demand for high-density, long-term data storage has intensified interest in three-dimensional optical storage based on laser-written structures in glass. We evaluate a diverse set of commercial glasses to determine how intrinsic material properties influence the formation of multibit phase voxels. By analyzing voxel quality and write efficiency, we develop a materials-screening framework that links storage performance to measurable glass properties. Thermal diffusivity emerges as the property most strongly correlated with quality ($r = 0.78$). The resulting performance landscape identifies a Pareto front in which Schott K10 exhibits the highest write efficiency (0.115 bit nJ^{-1} voxel^{-1}), while Borofloat 33 provides the highest quality (1.76 bit voxel^{-1}). Overall, this work demonstrates a generalizable methodology for selecting glasses for advanced laser-processing applications, highlighting thermal diffusivity as a key factor for achieving precise photonic structuring. © 2026 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

https://doi.org/10.1364/OPTICA.592661

1. INTRODUCTION

The exponential growth of global data generation, driven by advances in artificial intelligence and digital communication, has created an urgent need for high-density, long-term data storage solutions. Conventional magnetic and solid-state storage technologies, while effective for short- to medium-term applications, face limitations in terms of longevity, environmental stability, and scalability [1]. Optical data storage has emerged as a promising alternative, offering the potential for high-capacity, durable, and energy-efficient archival systems [2–7]. Among various optical storage media, glass has gathered significant attention due to its exceptional physical and chemical stability, resistance to environmental degradation, and capacity for multidimensional data encoding [8–10]. Its inherent durability and transparency render it an ideal candidate for long-term data preservation [11], particularly in harsh or inaccessible environments. Furthermore, the ability to encode data in three dimensions within the bulk of the glass enables storage capacities beyond those achievable with traditional planar media [9,12].

In practice, the implementation of glass-based optical storage systems presents several challenges. Writing volume pixels (voxels) into glass typically involves the use of ultrafast lasers to induce localized structural modifications, resulting in isotropic or anisotropic refractive index changes [13–18]. This process demands precise control over laser parameters, including pulse energy, repetition rate, and focusing conditions, as well as a deep understanding of the material properties of the glass. Similarly, accurate data retrieval requires optical readout techniques capable of resolving sub-micron features with high fidelity. To optimize the write efficiency and reliability of glass-based optical storage, it is essential to elucidate the complex interplay between laser writing parameters (processing), material properties, and storage performance (function). A comprehensive understanding of these relationships will support the rational selection and design of glass compositions, facilitate the development of predictive models [19,20], and enable adaptive writing strategies [21], ultimately advancing scalable, high-performance optical data storage technologies.

We assess the performance of glasses for optical data storage using two key metrics: quality (bit voxel^{-1}) and write efficiency (bit nJ^{-1} voxel^{-1}). Quality is defined as the number of user bits reliably stored in a single voxel, determined using a low-density parity-check (LDPC) error correction code [22]. Write efficiency quantifies the number of user bits written per voxel per nanojoule and reflects both the pulse energy and the achieved quality. It directly influences the energy consumption and write throughput of the storage system, with implications for multibeam scalability. For example, doubling the write efficiency means that the laser source can be split into twice as many data writing beams [12].

This work presents the first systematic evaluation of 30 commercial glass compositions for femtosecond laser direct writing of isotropic phase voxels, using a whole-system approach grounded in these two metrics. The metrics are derived from an end-to-end experimental pipeline that encompasses laser writing, optical readout, and LDPC-based error correction, enabling a rigorous and comprehensive assessment of storage performance. In our

most extensive tests, 7 billion voxels were written into each single 2 mm thick glass sample and processed through the entire pipeline. The study focuses exclusively on isotropic phase encoding; polarization-based schemes and anisotropic structures are beyond its scope. In contrast to prior investigations that examined isolated aspects of the writing process or material response [8,17,23–29], this study integrates all stages of the storage workflow to provide a holistic evaluation. By linking performance metrics to intrinsic material properties, this study reveals the key factors governing data quality and write efficiency in femtosecond laser-based optical storage.

2. METHODS

The glasses were tested using an end-to-end system previously described [12], with software modifications to both the writing and reading subsystems to accommodate variations in refractive index. Specifically, the collar correction firmware was updated to automatically adjust the objective lens collar position based on the refractive index of the glass, in addition to the writing or reading depth. The influence of higher-order aberrations and dispersion from refractive index variations was deemed negligible relative to the linear spherical aberration. In addition, the z-stage movement during both writing and reading was scaled to account for refractive index dependent changes in the depth of focus.

A. Write

In brief, data were written using single 400 fs pulses generated from the second harmonic (516 nm) of an amplified femtosecond laser (Satsuma HP3, Amplitude Systems, 10 MHz). The pulse duration was verified using an autocorrelator (Carpe, APE GmbH). The laser train was attenuated by a tunable attenuator and directed through a quartz acousto-optic modulator (AOM) (G&H, I-M110-2C10B6-3-GH26) for amplitude modulation. The AOM was modulated with an RF signal (frequency-locked to the laser) that encoded the data in the first-order diffracted beam. The modulated laser pulses were collimated and directed to a polygon scanner (Novanta, SA24) with 24 facets spinning at 17,000 rpm, to provide fast-axis scanning. The scanned laser pulses were directed through a custom-made f-theta scan lens ($f = 63$ mm) followed by a relay lens. Finally, the laser pulses were focused into the glass sample using a high numerical aperture (NA) objective lens (Olympus LUCPLFLN40X, NA = 0.6). The emission generated during voxel formation was collected by the same objective lens and directed to a CMOS sensor (Grasshopper3 GS3-U3-41C6M-C) for closed-loop feedback control. The CMOS sensor shutter (integration time) was adjusted for each glass to capture a sufficiently strong emission signal, thereby reducing the impact of noise on the feedback loop. Integration times ranged from 7 to 68 ms. A notch filter was used to block the scattered laser light.

The glass samples were mounted on an xy stage (PI, V-551.7D and 551.4D) for precise lateral positioning during the writing process. Writing depth within the glass was controlled by translating the objective lens along the z axis. To compensate for spherical aberrations, a closed-loop energy control system was employed alongside a custom motorized actuator that adjusted the objective's correction collar (x_c) based on the refractive index of the glass (n) and the target writing depth (z). The energy control system monitored photo-induced emission during voxel formation and dynamically adjusted the pulse energy to maintain consistent voxel modification strength. This feedback mechanism helped counteract depth-dependent focal distortions, ensuring reliable encoding across the full writing volume. Additionally, the closed-loop energy control compensated for variations in voxel modification strength caused by fluctuations in laser output power and differences in polygon facet reflectivity [12].

An xy voxel pitch of 0.5×0.7 μm was used for all tests. The y direction corresponded to the fast axis of the polygon scanner, while the x direction corresponded to the slow axis controlled by the xy stage.

For Stage 1 (Section 3.A), each voxel encoded two bits using a heuristically defined four-level amplitude modulation (symbols A–D) [12], controlled by the AOM. Symbol D corresponded to the highest energy (emission setpoint), with symbols C, B, and A representing progressively lower energies.

B. Read

In brief, data were read using a custom-made Zernike phase contrast microscope [30] fitted with a LED light source (Thorlabs Solis-445C, $\lambda = 445$ nm) and sCMOS camera (Hamamatsu ORCA Flash4 v3.0). An illumination objective (Thorlabs MY20X-804) with an annular amplitude mask was used to illuminate the sample. A high NA objective (Olympus LUCPLFLN40XPH, NA = 0.6) with a phase ring was used to image the sample. The glass sample was mounted on a motorized xyz stage for precise lateral positioning during the readout process. A track (vertical stack of sectors) was read by translating the stage in z in a single pass to capture a series of images at different depths. An autofocus (AF) algorithm was used to determine the optimal focal plane for each sector. To compensate for spherical aberrations during readout, the objective's correction collar (x_c) was adjusted based on the refractive index of the glass (n) and the focal depth (f). For glasses with lower refractive indices, a simple proportional relationship ($x_c \propto nf$) was sufficient to compensate for spherical aberrations. However, for higher-index glasses ($n > 1.51$), this approximation proved inadequate. In these cases, an empirical correction was applied by optimizing the AF signal across the full depth of the glass, resulting in improved quality. The glasses requiring empirical correction included Schott N-SK2, Schott N-BAF10, Schott N-KZFS4, Sumita K-PBK40, and Sumita K-CD300.

C. Decode

Images captured during the readout process were processed using a pipeline that included preprocessing, symbol inference, symbol-to-bit mapping, and error correction [12]. Preprocessing involved finding the sector boundaries by using a small amount of known data written at the edge of the sector. This known voxel pattern provided an easily detectable signature to localize sector boundaries. Symbol inference converted the stack of images for each sector into a 2D array of symbol probabilities for every voxel. Two inference methods were used depending on the stage of the workflow.

Symbol inference for stage 1 of the workflow (initial encodings) was performed by thresholding the measured voxel intensities to classify the corresponding symbols. The symbol error rate (SER) was computed by comparing decoded results against the ground truth symbols.

Symbol inference for stage 2 of the workflow (symbol placement analysis) was performed using a convolutional neural network (CNN) model trained to classify voxel symbols from the readout images. The CNN applied successive convolutions, nonlinear activations, and downsampling operations to extract features relevant to symbol classification from the input image at different resolutions.

The resulting symbol-probability maps are converted to bit-probability maps according to the amplitude modulation scheme used during writing. Finally, LDPC error correction coding is applied to correct remaining errors within each sector. The procedure for selecting the LDPC code rate (defined as the ratio of useful bits to total bits) follows the methodology established in prior work [12].

The CNN model was re-trained for each glass and each emission setpoint, using 18 tracks for training and 3 tracks for validation. For a 258 layer sample, 18 tracks corresponded to 4644 sectors or 128,397,312 voxels of training data.

3. RESULTS

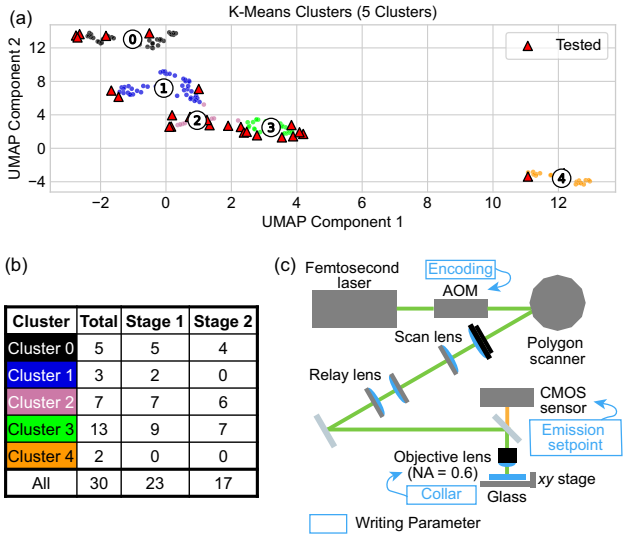
Our investigation focused on undoped base glasses to isolate intrinsic material properties, which are fundamentally governed by the structural motifs and bonding configurations within the amorphous network. This approach ensured that the observed trends in writing behavior and performance could be attributed to inherent glass characteristics rather than extrinsic factors such as dopant-induced modifications. Although dopant incorporation can modulate the optical response and enhance data storage performance, it was deliberately excluded from the present study to maintain a clear link between material properties and storage metrics.

To narrow the number of samples tested, we first verified that the key material property ranges offered by major glass manufacturers (e.g., Schott, Sumita, Hoya, AGC, and Ohara) were broadly comparable. This enabled the selection of representative glasses without compromising coverage of the relevant optical, thermal, and mechanical property space.

Clustering analysis was then performed on the Schott optical glass catalogue with 154 glasses, which provided a sufficiently dense and well-characterized dataset for dimensionality reduction and grouping. Uniform manifold approximation and projection (UMAP) [31] was applied to ten material properties—including refractive index (n), density (ρ), Young's modulus (Y), glass transition temperature (T_g), thermal diffusivity (D_{th}), and coefficient of thermal expansion (α)—followed by k-means clustering to identify glass clusters with similar intrinsic properties [Fig. 1(a)].

The number of clusters was determined empirically by evaluating the separation and compactness of the grouped data in the reduced UMAP space. Five clusters provided a balance between capturing meaningful variation in intrinsic glass properties and avoiding over-segmentation. A subset of 18 glasses was selected from across these clusters to ensure coverage of the full range of intrinsic properties [shown as triangles in Fig. 1(a)]. To further broaden the scope of the study, additional commercial glasses from Sumita, AGC, and Corning were included, resulting in a total of 30 glasses for evaluation (see Supplement 1 for the full list of glasses).

The selected glasses were evaluated using a two-stage experimental pipeline, which was designed such that each stage could be completed using a single glass sample.



Cluster	Total	Stage 1	Stage 2
Cluster 0	5	5	4
Cluster 1	3	2	0
Cluster 2	7	7	6
Cluster 3	13	9	7
Cluster 4	2	0	0
All	30	23	17

Fig. 1. (a) UMAP of the dimensionally reduced Schott optical glass dataset, overlaid with the subset of experimentally tested glasses (red). K-means clustering was used to identify similar glasses in the dimensionally reduced dataset (five clusters). The selected glasses span a broad range of material properties, including refractive index (n), density (ρ), Young's modulus (Y), glass transition temperature (T_g), thermal conductivity (λ), and coefficient of thermal expansion (α). (b) Summary table showing the number of tested glasses grouped by cluster, along with counts of glasses that successfully completed Stages 1 and 2 of the experimental workflow. Stage 1 involved establishing an initial multibit encoding scheme by writing a single-layer of voxels, while Stage 2 focused on optimizing symbol placement and assessing storage performance during multilayer writing. (c) Schematic of the laser writing system. Optimized parameters for each glass included amplitude encoding (via AOM), emission setpoint (CMOS sensor), and collar correction (objective lens). AOM, acousto-optic modulator; NA, numerical aperture; and CMOS, complementary metal oxide semiconductor.

- **Stage 1** served as a rapid screening step to establish an initial multibit encoding scheme for each glass type using single-layer samples (Section 3.A).
- **Stage 2** focused on optimizing this encoding scheme and assessing storage performance by writing multiple layers to fill the full 2 mm glass thickness in the most promising glasses (Section 3.B).

This workflow was developed based on extensive optimization using Borofloat 33 glass as a foundation, in which terabytes of data have been written to refine multibit encoding schemes and laser parameters for inscribing distinct phase states within individual voxels [12]. The equipment and methodologies established through this process provided a robust starting point for the broader screening effort. Leveraging these tools, we tailored the writing parameters to each glass type to determine the optimal symbol number per voxel and the corresponding pulse energy required to inscribe each phase-modified state, thereby maximizing information density. The number of tested glasses and those that successfully completed each stage are summarized in Fig. 1(b).

An xy voxel pitch of 0.5×0.7 μm was used for all tests, where the y direction corresponded to the fast axis of a polygon scanner, and the x direction corresponded to the slow axis controlled by an xy stage. The selected areal pitch, based on optimization established in prior work [12], ensured a fair comparison across all glass types, with the y-pitch intentionally set larger than the

x-pitch to mitigate thermal accumulation effects. We note that for a particular glass type, the voxel pitch can be further optimized to enhance storage performance, as demonstrated in prior work [12].

A closed-loop energy control system maintained consistent voxel modification strength by monitoring photoinduced emission during voxel formation [12,32]. Specifically, a CMOS sensor detected this emission, collected back through the objective lens, and provided feedback to dynamically adjust the AOM amplitude, thereby modulating the pulse energy delivered to the glass until the target emission setpoint was reached. This emission setpoint served as a proxy for voxel modification strength and corresponded to the maximum pulse energy required to inscribe the highest-energy symbol in the multibit encoding scheme.

The specific laser writing parameters that were optimized for each glass are shown in Fig. 1(c) and include the pulse energy (via the emission setpoint) [12], the collar correction of the objective lens to account for the refractive index of the sample, and the amplitude encoding scheme of the acousto-optic modulator (AOM). Section 2.A provides further details of the writing method.

A. Stage 1: Modification Thresholds and Initial Encodings

Stage 1 established an initial multibit encoding scheme by defining the dynamic range of writing energies required to produce distinct voxel states. The modification threshold for each glass was defined as the minimum pulse energy required to induce a detectable isotropic change in refractive index. The lower bound was set near this threshold, while the upper bound corresponded to the energy at which stable, high-contrast modifications were achieved without causing damage. These bounds defined the energy window for symbol encoding and informed subsequent optimization steps. To characterize these parameters, a single layer of voxels was written at varying pulse energies 100 μm beneath the glass surface. Each glass sample contained 14,376,960 voxels organized into 520 sectors, with each sector sized to fit within the field of view of the read subsystem.

The emission setpoint was varied from 500 to 6500, and the corresponding pulse energies were measured using a calibrated photodiode. Figure 2(a) shows the typical linear relationship between pulse energy and emission counts for Borofloat 33, with high-energy deviations from this linearity attributed to thermal accumulation. Specifically, at high pulse energies the thermal energy imparted by each laser pulse is insufficiently dissipated into the surrounding volume before subsequent pulses arrive, leading to localized thermal accumulation. This results in a preheating effect that influences the thermal state of the next voxel to be written, as well as a postheating effect that alters the cooling dynamics of the previously written voxel.

The modification threshold was determined by extrapolating the linear region to zero emission, thereby avoiding the nonlinear behavior associated with thermal accumulation at higher emission setpoints. For Borofloat 33, this yielded a threshold of 14.3 nJ. Repeatability of the threshold measurements was confirmed by writing multiple single-layer samples of Borofloat 33 on different days, yielding a standard deviation of 0.2 nJ (see Supplement 1).

Corresponding symbol error rate (SER) analysis [Fig. 2(b)] decoded 40 sectors per emission setpoint, each containing randomly generated data. Boxplots for each emission setpoint show the distribution of SERs across the 40 sectors, with the median

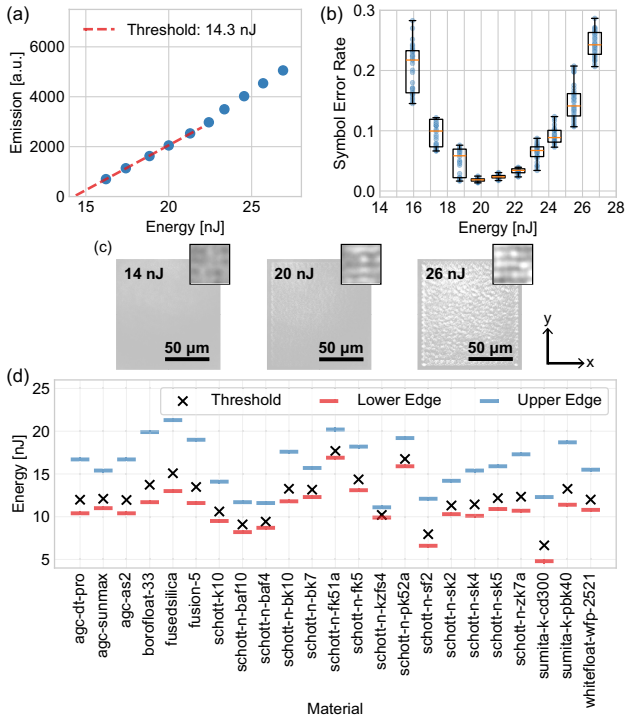


Fig. 2. (a) Energy versus emission counts for Borofloat 33. Extrapolating the trend back to zero emission counts gives the modification threshold energy for the glass (14.3 nJ). The measurements correspond to data writing at a depth of 100 μm, a xy voxel pitch of 0.5 × 0.7 μm and a pulse duration of 400 fs. (b) Energy versus symbol error rate (SER) for Borofloat 33, by decoding 40 sectors of randomly generated data per emission setpoint. The minimum SER was obtained by writing with an energy of 19.8 nJ. Boxplots for each emission setpoint show the distribution of SERs across the 40 sectors, with the median indicated by the horizontal line within each box. (c) Phase contrast images of sectors written at various energies, showing weak contrast at the modification threshold (14.3 nJ) and stronger contrast at higher energies. Scale bar = 50 μm. The insets show a magnified view of approximately 5 × 4 voxels. (d) Summary of the modulation range for each glass, based on the symbol energies to store 2 bits per voxel. The modification thresholds obtained using the method described in (a) are indicated by the black crosses. The upper edge was based on the emission setpoint used to obtain the lowest SER, as exemplified in (b). The lower edge was defined by allocating 75% of the span between the modification threshold and the SER-optimized energy, with the remaining 25% positioned below the modification threshold.

SER indicated by the horizontal line within each box. As the pulse energy approaches the modification threshold, the median SER increases as the voxel contrast becomes too weak to reliably distinguish between symbols. At higher pulse energies, the median SER increases due to thermal accumulation effects and/or the formation of larger modification regions, which lead to cross-talk between adjacent voxels and a reduced ability to distinguish between phase states. In addition to this upwards shift in the median SER, the distribution of SERs also increases at higher energies, indicating greater variability in how well symbols can be separated across different sectors. This effect was confirmed by writing sectors with a larger voxel separation along the fast axis for writing (y) to reduce heat accumulation and/or more effectively separate adjacent voxels. This adjustment led to a measurable reduction in SER at higher pulse energies (see Supplement 1). The minimum SER for Borofloat 33 was obtained at 19.8 nJ. Phase contrast images

[Fig. 2(c)] of 96.0 × 100.8 μm sectors (192 × 144 voxels) written at fixed energies demonstrate: no modification at 14 nJ threshold, clear contrast at 20 nJ, and thermal damage at 26 nJ.

This analysis was repeated using different emission setpoints until a clear minimum in SER was identified for each glass. The modification threshold and the SER-optimized energy were then used to define the modulation range for each glass. The modulation edges were calculated by allocating 75% of the energy range between the modification threshold and the SER-optimized energy. The resulting bounds are summarized in Fig. 2(d) and were used as inputs for the symbol placement optimization described in Section 3.B. The modification thresholds varied significantly between glasses, ranging from 6.6 nJ (Sumita K-CD300) to 17.7 nJ (Schott N-FK51A). Moreover, some glasses have a wider energy range for modulation (e.g., Borofloat 33), while others have a narrow range (e.g., Schott N-KZFS4). As shown in Section 3.C, a negative correlation ($r = -0.68$) was observed between the modulation energy range and glass density (ρ), suggesting that higher-density glasses tend to restrict the available energy range. Correlations with the coefficient of thermal expansion (α, $r = -0.61$) and thermal diffusivity (D_{th}, $r = 0.55$) were also observed, suggesting that thermally induced stresses and heat dissipation may influence the energy range required for accurate data writing. The size of this energy range is important, as it determines how well the symbols can be distinguished from each other, and therefore how reliably they can be decoded. For example, a wider energy range could indicate the potential for more symbols to be used, which could increase the storage density. In addition, a broader range improves robustness against pulse energy fluctuations during writing.

The following glasses exhibited no detectable change in photoinduced emission as pulse energy increased up to 32 nJ: Schott LF5G19 (Cluster 3), Schott F2HT (Cluster 4), Schott N-SF6 (Cluster 1), Schott SF11 (Cluster 3), and Sumita K-LCV93 (Cluster 3). Although voxels can be written in these glasses, the lack of variation in emission with pulse energy renders them unsuitable for reliable data storage using a closed-loop energy control system; they were therefore excluded from further analysis. Sumita K-LCV161 (Cluster 4) and K-PSFN166 (Cluster 3) failed to decode using multibit (four-symbol) writing at all tested emission setpoints with 400 fs pulses. Even at the lowest setpoints, sector images revealed pronounced thermal damage, and voxels could not be reliably distinguished (see Supplement 1). Lowering the emission setpoint compresses the modulation energy range, reducing symbol separation and producing closely spaced phase states. These states exhibit intrinsically low optical contrast, limiting resolvability within the read subsystem. Consequently, while these glasses may support single-bit encoding, their suitability for multibit encoding appears limited without substantial improvements in the read subsystem's ability to discriminate small phase differences or adjustments to pulse duration on the write subsystem to modify energy deposition dynamics. These glasses were therefore excluded from further analysis.

It should be noted that neither of the two Cluster 4 glasses progressed to Stage 2. While this observation may indicate that Cluster 4 compositions could present challenges for multibit femtosecond laser writing, the sample size is too limited to support a general conclusion. Additional data will be required to determine whether this trend is representative of the broader cluster.

B. Stage 2: Symbol Placement Optimization and Quality Measurement

Equidistant symbol placement across the modulation energy range does not necessarily yield the best quality [12]. This section presents the optimal encoding schemes for multilayer writing on the 23 glasses where a modulation range was defined in Stage 1 (Section 3.A). To optimize symbol placement, data were recorded at five emission setpoints around the minimum SER point identified in Stage 1, employing 31-level amplitude modulation to encode nearly five bits per voxel.

Voxels were written to fill the top 1 mm in the glass, using either 129 layers with 7 μm z-spacing or 65 layers with 14 μm z-spacing. For the most promising glasses, writing was extended to the full 2 mm thickness of the platter (258 layers at 7 μm z-spacing). The 258 layer samples contained 6847,856,640 voxels organized into 5247,680 sectors or 960 tracks. These samples were read and decoded using machine learning (ML).

A vertical stack of sectors (a "track") was read in a single pass by continuously translating the stage in z during readout (Section 2.B). Samples were subsequently decoded using a convolutional neural network (CNN) to enhance symbol classification by compensating for noise and inter-voxel crosstalk beyond the capabilities of simple thresholding [12]. ML decoding requires extensive training data for robust performance. The substantial increase in stored data within multilayer samples, compared to a single-layer in Stage 1, provided a much larger dataset for training and validating the ML decoder, thereby enhancing its classification accuracy. At each setpoint, sectors containing randomly generated data were written (192 tracks per setpoint), and 24 representative tracks were read and decoded for symbol placement evaluation. Randomized data ensured that the ML decoder generalized effectively to unseen patterns rather than memorizing training sequences. Section 2.C provides further details of the ML decoding method.

Figures 3(a)–3(d) show Corning Fusion 5 as a representative example because it exhibits typical trends clearly and demonstrates performance that is relatively central in terms of quality and write efficiency, making it well suited for illustrating the optimization process.

ML decoder predictions for Corning Fusion 5 [Fig. 3(a)] show a flat response below the modification threshold (0.540), predictable modulation at intermediate levels, and flattening due to thermal accumulation at higher modulations (blue shading: 1σ standard deviation). Wide predictable modulation ranges yield well-separated symbols, while narrow ranges cause overlapping symbols prone to misclassification. The optimal placement of symbols was determined based on the ML decoder predictions. For each glass, the optimal placement of three, four, and five symbols was determined. Figure 3(b) shows the optimal placement of four symbols for Corning Fusion 5.

Based on the optimal encoding schemes, LDPC error correction analysis was then performed to determine the quality and write efficiency of each glass. The LDPC algorithm jointly optimized two parameters:

- **Code rate**, defined as the ratio of user data bits to total encoded bits, where the remaining bits are used for error correction.
- **Sector recovery rate**, defined as the proportion of data sectors that were successfully decoded and corrected.

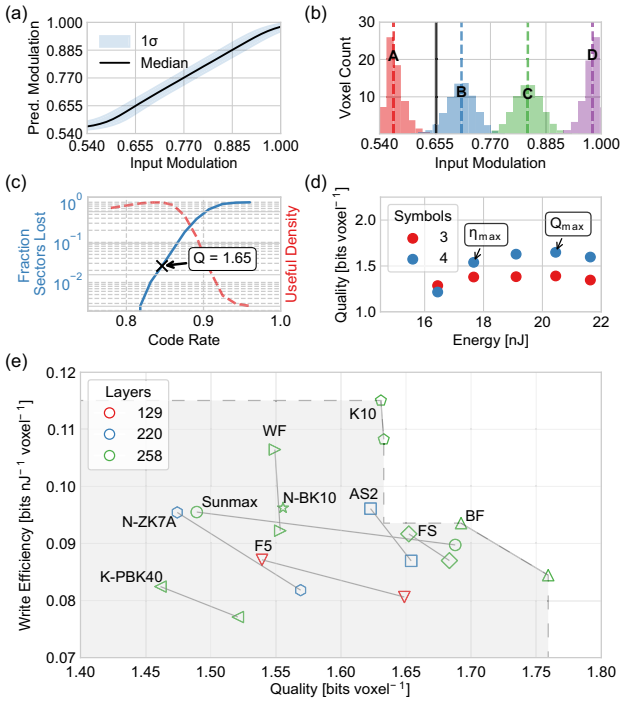


Fig. 3. (a) Example response of a voxel channel shown as predicted versus input modulation (Corning Fusion 5). (b) Voxel histogram of (a) showing the symbol classification of voxels by input modulation for four symbols. The dashed lines indicate the modulation that should be used to write each symbol. The solid black line indicates the modification threshold. (c) Error correction (LDPC) analysis on (b), showing the optimal code rate that maximizes the recoverable information stored and the useful density of the glass. The quality is defined at this point as the code rate, multiplied by the recovery rate (1—fraction sectors lost), multiplied by the number of bits written per voxel. The useful density of the glass as a function of code rate is shown in red. (d) Example quality variation as a function of the average writing energy and the number of symbols, to identify the optimal writing conditions (Corning Fusion 5). The highest quality (Q_{max}) and write efficiency (η_{max}) are annotated. (e) Best measured quality and write efficiency for the 10 top-performing glasses, evaluated over a 1 mm thickness (129 layers) or a 2 mm thickness (258 layers). The AGC AS2 and Schott N-ZK7A samples were found to be 1.7 mm thick, and so the quality was measured over this thickness (220 layers). Note that the writing energy that maximizes quality does not necessarily coincide with the energy that maximizes write efficiency. Where these optima occur at different energies, the connecting line between the two points indicates the operating range for that glass. N-BK10 is the exception, as both optima occur at the same writing energy, resulting in a single point.

These parameters were combined to derive the quality metric (in bit voxel^{-1}) that reflects the effective number of user bits reliably stored per voxel. Figure 3(c) shows the results with four symbols for Corning Fusion 5. Lower code rates reduce sector loss but decrease user bits per voxel; higher rates increase sector loss, reducing recoverable information. The optimal code rate maximizing recoverable information and storage density is marked with a black cross. The corresponding quality (bit voxel^{-1}) is equal to the optimal code rate, multiplied by the sector recovery rate (1—fraction sectors lost) at this code rate, multiplied by the bits per voxel (three symbols: 1.5 bit, four symbols: 2 bit). For Corning Fusion 5, the quality is 1.65 bit voxel^{-1} with four symbols at a code rate of 0.85 (sector recovery rate of 0.97).

Figure 3(d) shows an example of how the quality varies as a function of the writing energy (highest energy symbol) and the number of symbols used. At low writing energies, the symbols are spaced too closely together to be reliably distinguished, resulting in a lower quality. This manifests as a reduction in the number of symbols that can be used reliably, as shown by the transition from four symbols (blue) to three symbols (red) at 16.4 nJ. At high energies, thermal accumulation and/or the formation of larger modification regions cause cross-talk between adjacent voxels, which also reduces the quality. The optimal energy for maximizing quality is 20.5 nJ, yielding 1.65 bit voxel^{-1} with four symbols for Corning Fusion 5. Across all tested glasses, the optimal encoding consistently involved three or four symbols; five-symbol encoding was never selected due to insufficient separation between adjacent symbol states.

Write efficiency was calculated by dividing the quality by the pulse energy used to write the highest energy symbol. To account for depth-dependent variations mitigated by the closed-loop energy control system, the pulse energy was averaged over all layers. The separate optimum for write efficiency (17.7 nJ) for Corning Fusion 5 is also indicted in Fig. 3(d).

Optimization results for the top 10 performing glasses are summarized in Fig. 3(e). The best measured quality and write efficiency is shown over a 1 mm thickness (129 layers), and for selected glasses, extended to 2 mm thickness (258 layers) to evaluate deeper writing performance. In some cases, the sample thickness was below 2 mm, limiting the number of layers to 220 (e.g., AGC AS2, Schott N-ZK7A). Borofloat 33 and AGC AS2 were among the glasses selected for 220/258-layer writes based on their favorable performance at 129 layers. A direct comparison between 129-layer and 220/258-layer results for these glasses is provided (see Supplement 1). In general, quality at greater depths decreases due to stronger spherical aberrations during readout, which arise from increased wavefront distortion as light propagates deeper through the glass. This degradation is primarily a limitation of the read process rather than the write process, as the same reduction in quality with focal depth is observed when the sample is read in its original orientation and when inverted for upside-down readout [12]. Additionally, voxel modifications attenuate the illumination source, reducing light transmission and thereby increasing noise and diminishing readout fidelity when more layers are written.

The writing energy that maximizes data quality does not necessarily coincide with the energy that maximizes write efficiency, as the latter is inherently dependent on the pulse energy used during writing. The selection of an operating point for each glass is therefore an economic decision, influenced by the relative cost of energy and the number of available writing beams. For instance, in the case of Borofloat 33, an optimal operating range lies between 15.7 and 20.5 nJ. Although lower voxel quality reduces per-platter capacity and per-beamline throughput, higher write efficiency enables more effective use of the available laser power. As a glass with higher write efficiency allows the laser beam to be split into a larger number of writing beams, the total write throughput of the system can increase even if individual beamline quality is slightly lower. However, if the objective is to build a single-beamline system with a flexible laser power budget, then the material with the highest voxel quality should be selected.

Figure 3(e) also illustrates the Pareto front of storage performance across all tested glasses (the boundary of the grey box), highlighting those that achieve optimal trade-offs between quality and writing efficiency. Glasses such as Schott K10 and Borofloat 33

lie on this front, indicating that no other tested samples simultaneously outperform them in both metrics. AGC Sunmax, AGC AS2, and fused silica (FS) also approach the Pareto front, demonstrating competitive performance. Their position establishes them as dominant solutions within the tested set and as benchmarks for future material development targeting archival optical data storage.

Glasses with the highest thermal expansion coefficients in the tested set (exceeding approximately 1.0×10^{-5} °C^{-1}, e.g., Schott N-PK52A, Schott N-FK51A, Sumita K-CD300) exhibited a pronounced tendency to fracture during multilayer writing. This was particularly evident when increasing the layer count from 65 layers at a 14 μm z-pitch to 129 layers at a 7 μm z-pitch, where substantial local volume changes during voxel formation generated cumulative stress that is hypothesized to have driven fracture. Such glasses represent a poor choice for optical data storage, as their limited tolerance to stress directly constrains achievable storage capacity. These glasses were not part of the top 10 performers and are absent from Fig. 3(e).

Six glasses, Schott N-BAF4 (Cluster 1), Schott N-KZFS4 (Cluster 2), Schott N-SF2 (Cluster 1), Schott N-BAF10 (Cluster 0), Sumita K-CD300 (Cluster 3), and AGC DT-Pro (Cluster 3), were evaluated in Stage 2 but could not be decoded when a 129 layer sample was written with multibit encoding. For these glasses, decoding failure is attributable to either (i) a narrow energy modulation range identified in Stage 1, which limits the achievable separation between symbols and reduces the signal-to-noise ratio required for reliable multilevel readout, or (ii) a refractive index above roughly 1.65, which indicates that we may be approaching the limits of the objective collar correction for mitigating spherical aberration effects during both writing and/or reading. As this study focuses specifically on multibit encoding performance, glasses that did not yield a decodable multibit sample under multilayer writing conditions were not tested further.

It should be noted that the absolute values of quality and write efficiency reported here for Borofloat 33 are lower than those previously reported ($Q = 1.84$ bit voxel^{-1}, $\eta_W = 0.113$ bit nJ^{-1} voxel^{-1}) [12]. This difference reflects the smaller training set of tracks used for ML decoding in this study (18 tracks versus 200 tracks), constrained by the need to assess a large number of glasses within practical timeframes. In addition, the reported metrics were derived by selecting the optimal number of symbols from 31-level amplitude modulation data rather than by direct writing with the optimal amplitude encoding. These factors likely led to a slight underestimation of performance; details of direct writing with optimal amplitude encoding for selected glasses are provided in Supplement 1. Nevertheless, for comparative analysis across glasses, the relative performance trends remain valid.

C. Material Properties and Storage Performance

To understand the relationship between material properties and storage performance, we calculated the correlation coefficient (r) between the measured storage metrics and the material properties of the glasses [Fig. 4(a)]. The corresponding scatter plots are provided in Supplement 1. As each voxel is formed by a single femtosecond pulse, the dominant physical processes (i.e., multiphoton ionization, free-electron generation, and picosecond-scale heat flow) occur before any substantial changes in the local material properties can develop. Consequently, voxel formation is expected to be influenced primarily by the response of the unmodified material, and the properties used in our analysis correspond to this initial state.

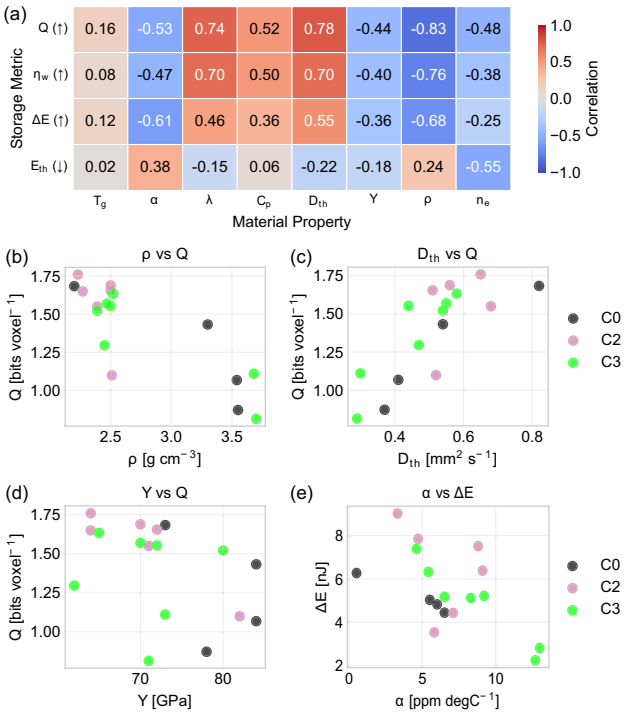


Fig. 4. (a) Correlation between material properties and storage metrics. Positive correlations are shown in red and negative correlations are shown in blue. The arrow next to the storage metric indicates the direction of positive performance changes. Material properties: glass transition temperature (T_g, °C), coefficient of thermal expansion (α, ppm K^{-1}), thermal conductivity (λ, W m^{-1} K^{-1}), isobaric specific heat capacity (C_p, J g^{-1} K^{-1}), thermal diffusivity (D_{th}, mm^2 s^{-1}), Young's modulus (Y, GPa), density (ρ, g cm^{-3}), and refractive index at 546 nm (n_e). Storage metrics: best quality (Q, bits voxel^{-1}), best write efficiency (η_w, bits nJ^{-1} voxel^{-1}), modulation energy range (ΔE, nJ), and modification threshold (E_{th}, nJ). (b) The relationship between density (ρ) and the best quality (Q). The clusters assigned from k-means clustering in Fig. 1(a) are indicated by different colors. No glasses from Cluster 4 advanced to Stage 2 of testing. Glasses from Cluster 1 are not shown because their quality was far below 1.0 bit voxel^{-1}, making them unsuitable for multibit encoding. (c) The relationship between the thermal diffusivity (D_{th}) and the best quality (Q). (d) The relationship between Young's modulus (Y) and the best quality (Q). (e) The relationship between the coefficient of thermal expansion (α) and the modulation energy range (ΔE).

The quality metric (Q) exhibits strong correlations with glass density (ρ, $r = -0.83$), thermal diffusivity (D_{th}, $r = 0.78$), and thermal conductivity (λ, $r = 0.74$). Scatter plots of the quality against density and thermal diffusivity are shown in Figs. 4(b) and 4(c). Thermal diffusivity is defined as $\lambda/(C_p\rho)$, where C_p denotes the specific heat capacity. These relationships indicate that glasses with lower density and higher thermal conductivity promote efficient heat dissipation during writing, thereby reducing thermal accumulation and mitigating cross-talk between neighboring voxels. In contrast, quality showed only a weak correlation with Young's modulus (Y, $r = -0.44$), suggesting that mechanical properties play a lesser role than thermal properties [Fig. 4(d)]. The negative trend implies that greater mechanical flexibility, associated with a lower Young's modulus, facilitates local glass network reconfiguration during and after femtosecond laser exposure. This flexibility reduces stress accumulation and may enable more precise

phase modulation and improved voxel uniformity in softer glasses. Nevertheless, thermal properties remain the dominant factors governing data quality.

Similar trends were observed for the write efficiency (η_w), which also showed strong correlations with the density ($r = -0.76$), thermal diffusivity ($r = 0.70$), and thermal conductivity ($r = 0.70$). This outcome is expected because write efficiency is directly proportional to quality. Collectively, these findings highlight the critical importance of thermal management in achieving high-quality, energy-efficient data storage in glass.

The modulation energy range (ΔE), defined as the difference between the average energy of the highest-energy symbol (E_{avg}) and the modification threshold (E_{th}), exhibited three notable correlations with material properties. A moderate negative correlation with the glass density (ρ, $r = -0.68$) indicates that denser glasses provide narrower energy windows for symbol modulation. Similarly, the coefficient of thermal expansion (α, $r = -0.61$) showed a negative correlation [Fig. 4(e)], suggesting that glasses with higher expansion coefficients experience greater thermally induced volume changes and network stresses, which constrain the range of stable modifications. In contrast, the thermal diffusivity (D_{th}, $r = 0.55$) displayed a moderate positive correlation, implying that glasses with higher diffusivity support wider modulation ranges because efficient heat dissipation mitigates thermal accumulation effects.

The modification threshold energy (E_{th}) shows a moderate negative correlation with the refractive index (n_e at 546 nm, $r = -0.55$), whereas correlations with other material properties are weak ($|r| < 0.4$). This trend reflects enhanced light-matter interaction in higher-index glasses, where increased refractive index indicates greater polarizability, amplifying the material's response to the femtosecond laser field and promoting nonlinear absorption mechanisms such as multiphoton ionization [15,32]. These processes lower the energy required for permanent structural modification. Additionally, high-index materials typically exhibit stronger nonlinear behavior [17,33], resulting in pronounced Kerr self-focusing. The associated optical confinement further contributes to the reduction in E_{th} observed in high-index glasses.

Although thermal diffusivity emerges as the strongest correlating material property in our analysis, it is not necessarily the most fundamental descriptor of the laser-glass interaction. Other material properties, such as viscosity and structural relaxation behaviour, are also expected to influence how a glass responds to ultrafast excitation [29]. However, these quantities are strongly temperature dependent, making it difficult to correlate them with the storage metrics in Fig. 4(a) without knowledge of the absolute temperature history in the photoexcited region. Thermal diffusivity, by contrast, is widely tabulated and consistently measured across manufacturers, making it a practical and reliable screening parameter for the diverse set of glasses examined here. While our findings do not imply that thermal diffusivity is the mechanistic driver of modification, its strong empirical correlation suggests that thermal transport processes play a key role in voxel formation by governing how quickly absorbed energy is dissipated and how much heat accumulates during writing. In this context, the absence of any meaningful correlation with the glass transition temperature (T_g) provides the first indication that the ability to store multiple bits within a voxel is not dominated by viscous flow or structural-relaxation processes, but instead by the rate at which heat can be removed from the modified region during ultrafast laser exposure. Future work incorporating less commonly reported material properties may provide deeper insight into the underlying physics.

Figures 4(b)–4(e) also differentiate glasses by their cluster membership derived from the UMAP analysis [Fig. 1(a)]. Notably, multibit voxels could not be successfully written, read, and decoded in glasses from Clusters 1 and 4. While these glasses may still support single-bit encoding, this was not explored in the present study. The limited performance suggests that these clusters may share structural or compositional features that hinder multibit encoding. In contrast, glasses from Clusters 0, 2, and 3 exhibited a broad distribution of performance, with no single cluster exclusively associated with high-performing samples. A re-clustering exercise focused on thermal properties within Clusters 0, 2, and 3 could provide further insight into the material characteristics that enable high-quality, energy-efficient writing. Continued investigation is warranted to elucidate the factors that distinguish high-performing glasses from the broader dataset.

4. CONCLUSION

In this study, we systematically evaluated 30 commercial glass compositions for femtosecond laser direct writing of isotropic phase voxels, with the goal of identifying key material properties that govern data storage fidelity and efficiency. By analyzing the quality and write efficiency across a diverse set of glasses, we elucidated the interdependencies between laser writing parameters, intrinsic material properties, and overall storage performance.

Three principal findings emerged from our analysis:

1. Thermal properties dominate over mechanical properties in governing quality and write efficiency.
2. Thermal diffusivity, a composite metric incorporating thermal conductivity, specific heat capacity, and density, emerges as a critical factor. High thermal diffusivity enables efficient heat dissipation, minimizing thermal accumulation during voxel writing.
3. While a higher refractive index lowers the glass modification threshold energy, it does not necessarily improve quality or write efficiency. Instead, both storage metrics correlate negatively with refractive index, indicating that lower-index glasses may be more suitable for multibit data writing.

In summary, glasses with lower density and higher thermal conductivity and diffusivity exhibited superior data storage performance, whereas those with high density and thermal expansion showed restricted modulation ranges and poor multibit capability. These findings underscore the importance of selecting compositions with favorable thermal properties for high quality, energy-efficient writing of phase voxels. Although the present study used a fixed pulse duration of 400 fs, the strong correlation between thermal diffusivity and quality indicates that the identified material trends are robust under the temporal conditions explored here. Nonetheless, further storage performance gains may be achievable by systematically varying pulse duration and investigating temporal optimization strategies for individual glasses. Looking forward, compositional tuning and machine learning-guided material design offer promising pathways to accelerate discovery of glass systems tailored for next-generation optical data storage.

Funding. Microsoft Research UK.

Acknowledgment. The authors acknowledge the Microsoft Project Silica Team for providing the end-to-end infrastructure that enabled this study. Technical contributions from James Allison, Michael Capp, and Oliver Snowdon are gratefully recognized, as is the constructive feedback on the paper provided by James Clegg and Iain Christie. The authors acknowledge the supply of a small number of commercial glass samples from Sumita Optical Glass Europe GmbH and AGC Incorporated without cost.

Disclosures. TL: Microsoft Corporation (E); RD, MS, CEW, AD, BT and RB: Microsoft Corporation (E,P).

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. P. Anderson, E. Aranas, Y. Assaf, *et al.*, "Project silica: towards sustainable cloud archival storage in glass," ACM Trans. Storage **21**, 1–31 (2024).
2. D. A. Parthenopoulos and P. M. Rentzepis, "Three-dimensional optical storage memory," Science **245**, 843–845 (1989).
3. Q. Dai, M. Ouyang, W. Yuan, *et al.*, "Encoding random hot spots of a volume gold nanorod assembly for ultralow energy memory," Adv. Mater. **29**, 1701918 (2017).
4. Q. Zhang, Z. Xia, Y.-B. Cheng, *et al.*, "High-capacity optical long data memory based on enhanced Young's modulus in nanoplasmonic hybrid glass composites," Nat. Commun. **9**, 1183 (2018).
5. M. Zhao, J. Wen, Q. Hu, *et al.*, "A 3D nanoscale optical disk memory with petabit capacity," Nature **626**, 772–778 (2024).
6. L. Gao, Q. Zhang, R. A. Evans, *et al.*, "4D ultra-high-density long data storage supported by a solid-state optically active polymeric material with high thermal stability," Adv. Opt. Mater. **9**, 2100487 (2021).
7. Z. Hu, X. Huang, Z. Yang, *et al.*, "Reversible 3D optical data storage and information encryption in photo-modulated transparent glass medium," Light Sci. Appl. **10**, 140 (2021).
8. I. Ktafi, J. Kong, M. Cavillon, *et al.*, "A new approach toward extreme thermal stability of femtosecond laser induced modifications in glasses," Laser Photonics Rev. **19**, 2401086 (2025).
9. H. Wang, Y. Lei, L. Wang, *et al.*, "100-layer error-free 5D optical data storage by ultrafast laser nanostructuring in glass," Laser Photonics Rev. **16**, 2100563 (2022).
10. G. Cheng, Y. Wang, J. D. White, *et al.*, "Demonstration of high-density three-dimensional storage in fused silica by femtosecond laser pulses," J. Appl. Phys. **94**, 1304–1307 (2003).
11. J. Zhang, M. Gecevicius, M. Beresna, *et al.*, "Seemingly unlimited lifetime data storage in nanostructured glass," Phys. Rev. Lett. **112**, 033901 (2014).
12. J. Allison, P. Anderson, E. Aranas, *et al.*, "Laser writing in glass for dense, fast and efficient archival data storage," Nature **650**, 606–612 (2026).
13. E. N. Glezer, M. Milosavljevic, L. Huang, *et al.*, "Three-dimensional optical storage inside transparent materials," Opt. Lett. **21**, 2023–2025 (1996).
14. R. Taylor, C. Hnatovsky, and E. Simova, "Applications of femtosecond laser induced self-organized planar nanocracks inside fused silica glass," Laser Photonics Rev. **2**, 26–46 (2008).
15. R. Gattass and E. Mazur, "Femtosecond laser micromachining in transparent materials," Nat. Photonics **2**, 219–225 (2008).
16. Y. Shimotsuma, M. Sakakura, P. G. Kazansky, *et al.*, "Ultrafast manipulation of self-assembled form birefringence in glass," Adv. Mater. **22**, 4039–4043 (2010).
17. D. Tan, K. N. Sharafudeen, Y. Yue, *et al.*, "Femtosecond laser induced phenomena in transparent solid materials: fundamentals and applications," Prog. Mater. Sci. **76**, 154–228 (2016).
18. M. Sakakura, Y. Lei, L. Wang, *et al.*, "Ultralow-loss geometric phase and polarization shaping by ultrafast laser writing in silica glass," Light Sci. Appl. **9**, 15 (2020).
19. M. L. Bødker, M. Bauchy, T. Du, *et al.*, "Predicting glass structure by physics-informed machine learning," npj Comput. Mater. **8**, 192 (2022).
20. H. Liu, Z. Fu, K. Yang, *et al.*, "Machine learning for glass science and engineering: a review," J. Non-Crystalline Solids **557**, 119419 (2021).
21. P. S. Salter and M. J. Booth, "Adaptive optics in laser processing," Light Sci. Appl. **8**, 110 (2019).
22. D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University, 2004).
23. Y. Lei, H. Wang, L. Skuja, *et al.*, "Ultrafast laser writing in different types of silica glass," Laser Photonics Rev. **17**, 2200978 (2023).
24. Y. Wang, S. Wei, M. R. Cicconi, *et al.*, "Femtosecond laser direct writing in SiO_2-Al_2O_3 binary glasses and thermal stability of type II permanent modifications," J. Am. Ceram. Soc. **103**, 4286–4294 (2020).
25. S. Shin, "Review of high-precision femtosecond laser materials processing for fabricating microstructures: effects of laser parameters on processing quality, ablation efficiency, and microhole shape," J. Laser Appl. **36**, 041201 (2024).
26. A. Butkutė and L. Jonušauskas, "3D manufacturing of glass microstructures using femtosecond laser," Micromachines **12**, 499 (2021).
27. T. Asai, Y. Shimotsuma, T. Kurita, *et al.*, "Systematic control of structural changes in GeO_2 glass induced by femtosecond laser direct writing," J. Am. Ceram. Soc. **98**, 1471–1477 (2015).
28. Y. Shimotsuma, S. Kubota, A. Murata, *et al.*, "Tunability of form birefringence induced by femtosecond laser irradiation in anion-doped silica glass," J. Am. Ceram. Soc. **100**, 3912–3919 (2017).
29. N. Shchedrina, M. Sosa, M. Cavillon, *et al.*, "Properties, mechanisms, and perspectives of ultrafast laser modifications in silicate glass volume," Int. J. Appl. Glass Sci. **17**, e16706 (2026).
30. F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," Phys. D **9**, 686–698 (1942).
31. L. McInnes, J. Healy, N. Saul, *et al.*, "UMAP: uniform manifold approximation and projection," J. Open Source Softw. **3**, 861 (2018).
32. P. Jürgens, C. L. Garcia, P. Balling, *et al.*, "Diagnostics of fs laser-induced plasmas in solid dielectrics," Laser Photonics Rev. **18**, 2301114 (2024).
33. R. C. Miller, "Optical second harmonic generation in piezoelectric crystals," Appl. Phys. Lett. **5**, 17–19 (1964).