Supporting Information

for

Photothermal Microscopy beyond Intensity Detection: Exploiting Spatial Signal Distributions for Enhanced Sensitivity

Shu-hei Urashima, Tomotaro Namba, Ryoji Kusaka*

Nuclear Science and Engineering Center, Japan Atomic Energy Agency (JAEA), Tokai, Ibaraki 319-1195, Japan.

S1. Sample and measurement cell

Ponceau 4R (Marubeni; 65 wt% purity with 35 wt% NaCl) was used as a model nonfluorescent solute. Pure water served as the solvent. A 1 mM solution was initially prepared and subsequently diluted to obtain lower concentrations. The molar absorptivity of aqueous Ponceau 4R is 18800 M^{-1} cm^{-1}.[1] Due to its negligibly small fluorescent yield, the excitation energy is predominantly relaxed via a thermal relaxation process.

The solutions were introduced into a microchannel (Institute of Microchemical Technology; 500 μm deep and 500 μm wide) made of fused silica. The thickness of the fused silica substrate of the microchannel was 700 μm. The solution was replaced via tubes connected to the microchannel, while the microchannel remained fixed on the microscope stage.

S2. Optical configuration

The optical configuration of the photothermal reflectivity setup is illustrated in Figure S1. The pump and the probe beams were generated using turnkey lasers (Edmund Optics; product numbers 35-072 and 35-069, respectively), which were continuous wave lasers. The output power for these lasers was adjusted to 50 mW (pump) and 10 mW (probe) by controlling diode current.

The probe beam was initially enlarged using a lens pair to approximately match the pupil of a microscopic objective lens (OL) described later. It then passed through a polarized beam splitter (PBS). As the probe beam was horizontally polarized, it primarily transmitted through the PBS, and its polarization was converted to circular polarization by a quarter-waveplate (WP).

The probe beam was coaxially overlapped with the pump beam using a dichroic mirror (DM; cut-on wavelength: 425 nm), and focused onto the glass/aqueous interface with an OL (Olympus; LUCPlanFL N 60×, numerical aperture (NA): 0.70). The OL was piezo-controlled to maintain a consistent focal plane throughout the measurements. The autofocusing algorithm is described in Section S3. The probe beam reflected at the interface was collected by the same OL and its polarization converted to vertical by the WP. The reflected probe beam was then reflected by the PBS towards the detection path. The probe beam was further divided into two by a non-polarizing beam splitter (BS; reflection : transmission = 90 : 10). Ninety percent of the probe beam was directed into a lock-in camera (Heliotis; C4M, pixel pitch: 12 μm), which selectively recorded the periodic modulation of the probe beam profile, synchronized with the pump beam's chopping frequency. Two bandpass filters (BPF; Thorlabs; FBH405-10, center wavelength: 405 nm, full width of half maximum (FWHM) of the bandpass: 10 nm, optical density: >5) were placed to block stray pump light. The remaining 10% was focused onto a scientific complementary metal oxide semiconductor (sCMOS) camera using an imaging lens to monitor the microscopic image, which was used for autofocusing.

For the pump beam, a lens pair was used to enlarge it and match its diameter with the pupil of the OL. A mechanical optical chopper was placed at the focal point of the lens pair. The chopping frequency was 349 Hz, identical to that used in our previous intensity-based study. Although the photothermal signal increases at lower modulation frequencies, the frequency of 349 Hz provided the best signal-to-noise ratio because the noise level also depends on the frequency. The pump beam was introduced to the glass/aqueous interface along with the probe beam. The probing volume corresponds to the focal volume of the pump beam. For order-of-magnitude estimation, the pump beam focal volume was approximated using a commonly

employed cylindrical model, with a radius equal to the focal spot size (464 nm) and a height equal to twice the Rayleigh length (1270 nm), as determined by the NA of the OL and the wavelength. This yields a probing volume of 1.7×10^{-18} m^3 (1.7×10^{-15} L).

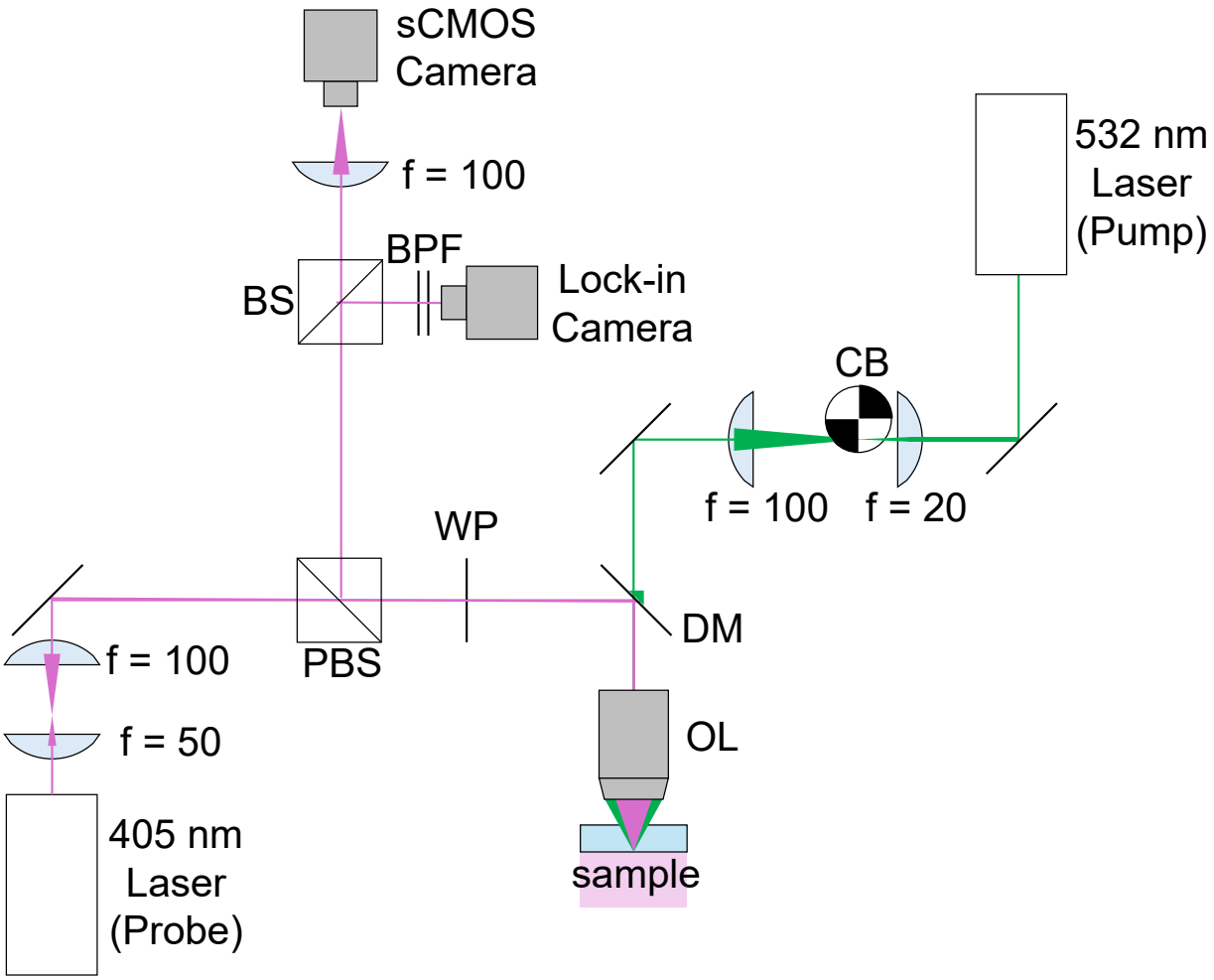

Figure S1. Optical layout of the photothermal reflectivity spectrometer utilizing a lock-in camera. Components included: PBS (polarized beam splitter), WP (quarter-wave plate), DM (dichroic mirror), OL (microscope objective lens), CB (optical chopper blade), BS (nonpolarized beam splitter), BPF (bandpass filter), and sCMOS (scientific complementary metal oxide semiconductor) camera. "f" represents the focal length of the lens (in mm).

S3. Measurement procedure

As described in the main text, the lock-in camera recorded the photothermal modulation of the probe beam pattern with each cycle of the 349 Hz chopping of the pump beam. Consecutive images of 64 cycles were recorded and transmitted to a computer at once, taking approximately 7–8 s. In practice, it was necessary to adjust the focus of the microscope OL and to measure a baseline image before acquiring the photothermal modulation. To adjust the focus, the OL was scanned using the piezo controller with a step size of 200 nm, and the focus was fixed when the reflected probe beam was most tightly focused on the sCMOS camera. Subsequently, because the lock-in camera's baseline depended on pixels and average light intensity, the baseline was measured by blocking the pump beam for 64 cycles × 4 repetitions, and the average was used as the baseline. With the same focus and baseline, 64 cycles × 32 sets of the photothermal images were acquired. To minimize vibration from the shutter opening, a 1-minute wait was implemented before acquiring the photothermal image. This measurement cycle (autofocus, baseline acquisition, and photothermal signal acquisition) was repeated 8 times for each concentration, resulting in a total of 64 × 256 images per concentration. The data was collected over 5 days. The concentrations measured were 0, 0.3, 0.7, 1, 3, 10, 30, and 1000 nM for Days 1–4, and 0, 0.1, 0.5, 1, 5, and 10 nM for Day 5.

While the solution concentration was generally increased sequentially from the lowest value (0 nM), additional control measurements were performed to assess potential long-term temporal drift within a single experimental day. On Day 1, the measurement at 0 nM was repeated after completion of the measurement at the highest concentration (1000 nM). Similarly, on Day 2, measurements started at 0.3 nM, and both 0 and 0.3 nM were remeasured after the 1000 nM measurement. Analogous measurement sequences were employed on Days 3 and 4,

with starting concentrations of 0.7 and 1 nM, respectively. All datasets acquired both at the beginning and at the end of each day (e.g., the repeated 0 nM measurements on Day 1) were included in the subsequent quantitative analysis. By varying the starting concentration across days and remeasuring identical concentrations at different time points within a day, gradual temporal changes in experimental conditions can be assessed independently of concentration-dependent effects.

S4. Heat conduction simulation

To estimate the extent to which the pump and probe foci can be spatially mismatched, heat conduction induced by the photothermal effect was simulated. The details of the calculation are described in our previous study.[1] Briefly, the heat conduction equation

$$\frac{\partial}{\partial t}T(t,r,z) = \frac{k}{\rho c}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{\partial^2}{\partial z^2}\right]T(t,r,z) + \frac{Q}{\rho c}\ , \qquad (1)$$

was numerically solved. Here, T, k, ρ, and c represent the temperature, the thermal conductivity, the density, and the specific heat capacity, respectively. Q represents the heat generated by light absorption per unit time. The variables t, r, and z denote time, radial distance, and height in cylindrical coordinates, respectively. Equation (1) assumes no dependence on the azimuthal angle. The heat source term Q is expressed as

$$Q(r,z,t) = \frac{4P\varepsilon C}{\pi^2\omega_{\mathrm{pump}}^2(z)}\exp\left(-\frac{2r^2}{\omega_{\mathrm{pump}}^2(z)}\right)s(t), \qquad (2)$$

where P, ε, and C, denote the pump laser power, molar absorptivity, and concentration, respectively. The function $s(t)$ is a step function that alternates between 0 and 1 every 1.43 ms, corresponding to a modulation frequency of 349 Hz. $\omega_{\mathrm{pump}}\ (z)$ represents the pump laser radius at position z, which is described as

$$\omega_{\mathrm{pump}}^2(\mathrm{z}) = \omega_0^2\left(1 + \frac{z^2}{z_\mathrm{r}^2}\right) . \qquad (3)$$

ω_0 and z_r are the radius of the focus and Rayleigh length, respectively, and are given by:

$$\omega_0 = \frac{0.61\lambda}{\mathrm{NA}} \qquad (4)$$

$$z_\mathrm{r} = \frac{\pi\omega_0^2}{\lambda}, \qquad (5)$$

where NA and λ are the numerical aperture of the lens and the wavelength of the light, respectively. From NA = 0.70 and λ = 532 nm, ω_0 and z_r of the pump laser are calculated to be 464 nm and 1270 nm, respectively. For comparison, the focal radius of the probe beam (wavelength: 405 nm) is calculated to be 353 nm. The physical parameters *k*, *ρ*, and *c* were set to 0.610 W/m·K,[2] 997 kg/m^3,[3] and 4182 J/kg·K[4] for the aqueous phase, respectively, and 1.27 W/m·K,[5] 2200 kg/m^3,[6, 7] and 700 J/kg·K[8] for the glass phase, respectively. The molar absorptivity *ε* was 1880 m^2/mol (18800 and 0 M^{-1} cm^{-1}) for the aqueous phase and 0 for the glass phase, respectively. Eq. (1) was solved using the finite element method, and the mesh size and time step used in the calculation were 0.5 μm and 0.5 μs, respectively. The initial temperature was set to 25 °C throughout the system. The laser focus was assumed to be positioned 1.5 μm (*ca.* the Rayleigh length) into the aqueous phase from the glass/aqueous interface with a numerical aperture (NA) of 0.70 to account for possible imperfections in optical alignment. Boiling and convection of water were not considered.

Figure S2 shows the calculated temperature distribution at $t = 1.43$ms. Although the temperature at the focal spot increases locally, it decreases rapidly with distance in the transverse direction. The temperature difference from the initial temperature (25 °C) becomes less than half at a distance of approximately 1 μm from the focus. Therefore, to obtain a sufficient change in

reflectivity, the allowable spatial mismatch between the pump and probe beams is expected to be less than 500 nm.

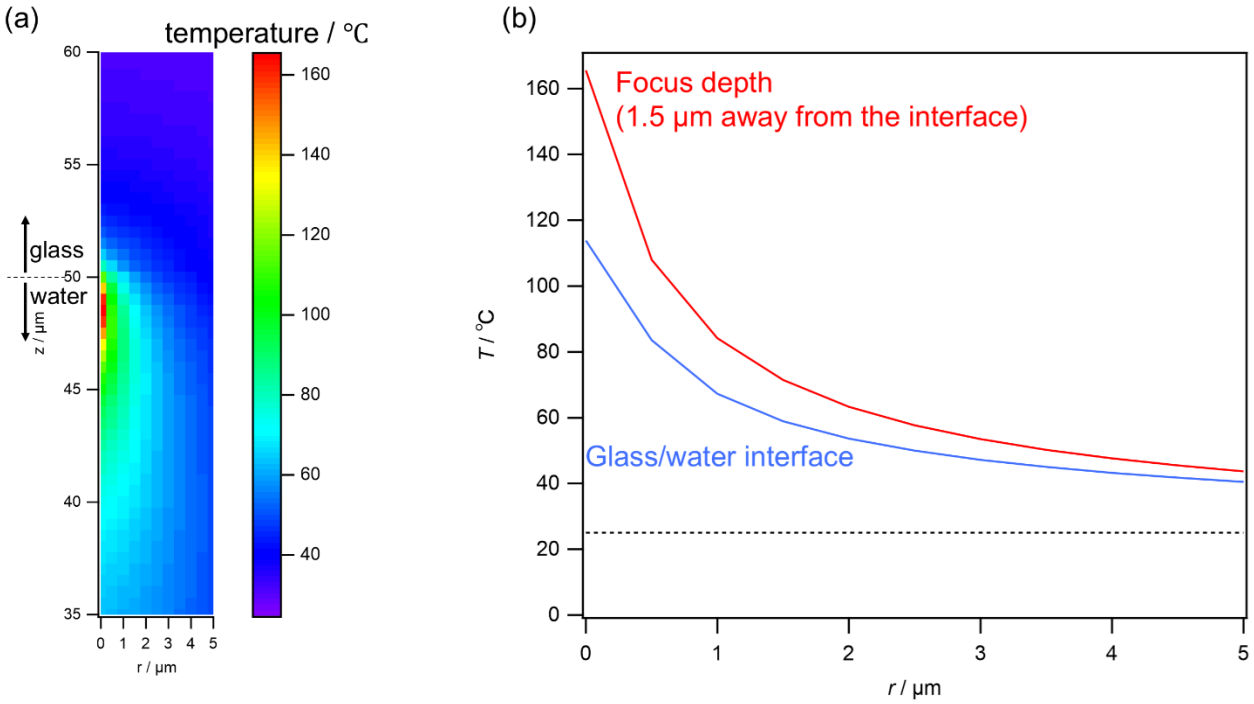

Figure S2. Thermal distribution calculated for a glass/aqueous interface pumped with 50 mW of 532 nm light for 1.4 ms (half cycle at 349 Hz). The concentration of Ponceau 4R and the ambient temperature were assumed to be 1 mM and 25 °C, respectively. (a) Two-dimensional spatial map of the temperature distribution. The horizontal axis (r) represents the transverse distance from the focal point, while the vertical axis (z) represents the longitudinal position, with z = 50 µm defined as the glass/aqueous interface. (b) Temperature distributions at fixed z positions, shown for z = 50 µm and z = 48.5 µm.

S5. Preprocessing of the photothermal lock-in images

For discussion and deep-learning recognition, the photothermal modulation of the pattern was preprocessed as follows. First, the images were Fourier transformed and filtered by multiplying a filter function that eliminates a constant term and high-frequency components. The

filtered images were then inverse-Fourier transformed to restore the signal. Because the high-frequency components were removed, the restored image was further downscaled by averaging 4 × 4 pixels into 1 pixel. While the lock-in camera outputs two images, I and Q, which are phase-shifted in 90° as is typical lock-in amplifier outputs X and Y, the strong pattern appeared in either I or Q in our experimental conditions, likely due to a slight difference between the actual chopping frequency and the reference frequency processed in the lock-in camera. To correct this accidental phase shift, the smoothed images for I and Q were phase-rotated so that the center of the concentric ring pattern aligned with 0°. Specifically, the rotated I image has a positive value at its center, while the rotated Q image approaches zero. In this study, only the rotated I image was used for discussion and analysis. Examples of the raw data, its Fourier transformation, the filter function, and the processed image are shown in Figure S2. The preprocessed image shown was scaled to a range from -0.5 to +0.5.

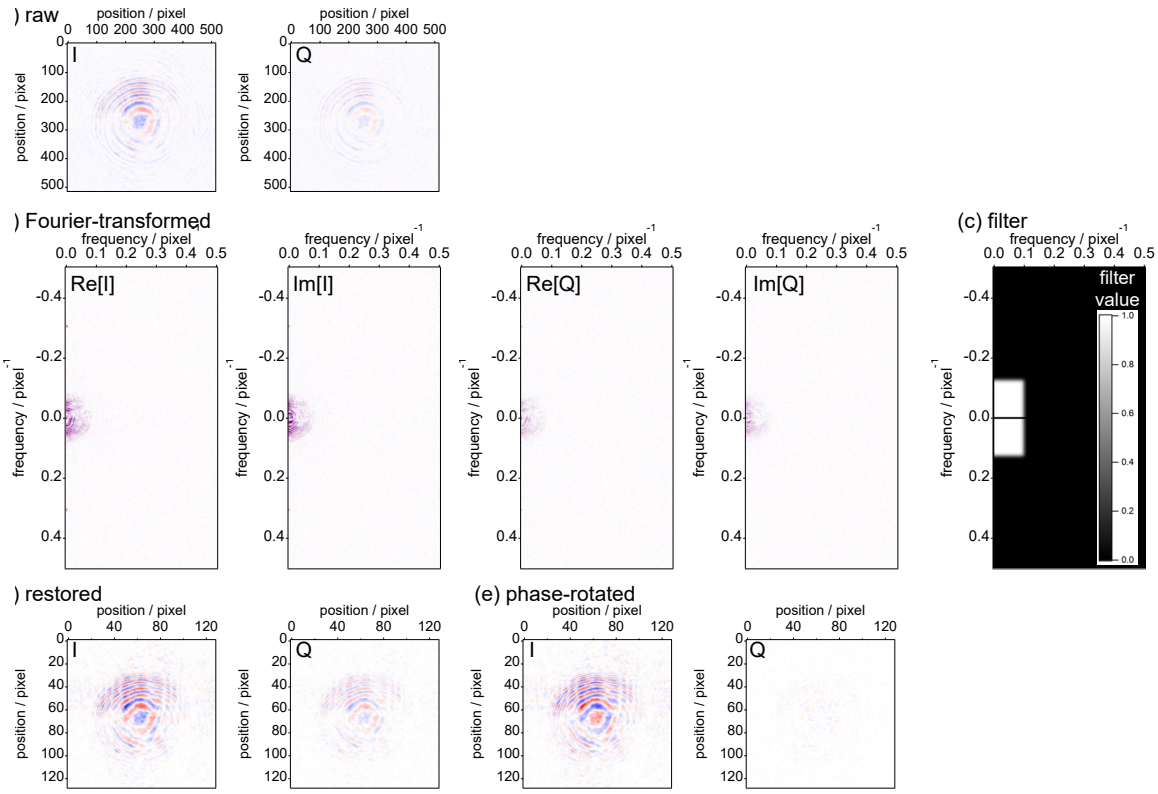

Figure S3. Representative lock-in images. (a) raw image (pixel pitch: 12 μm), (b) Fourier-transformed image, (c) filter function, (d) restored image (pixel pitch: 48 μm), (e) phase-rotated image (pixel pitch: 48 μm).

S6. The representative lock-in images and their position-based analysis

While the main text presents images only for 0 and 1000 nM, images obtained at other concentrations are shown in Figure S3. As shown in the right panels, the positive (red) area shrank and shifted slightly toward the upper left at concentrations ≥ 100 nM.

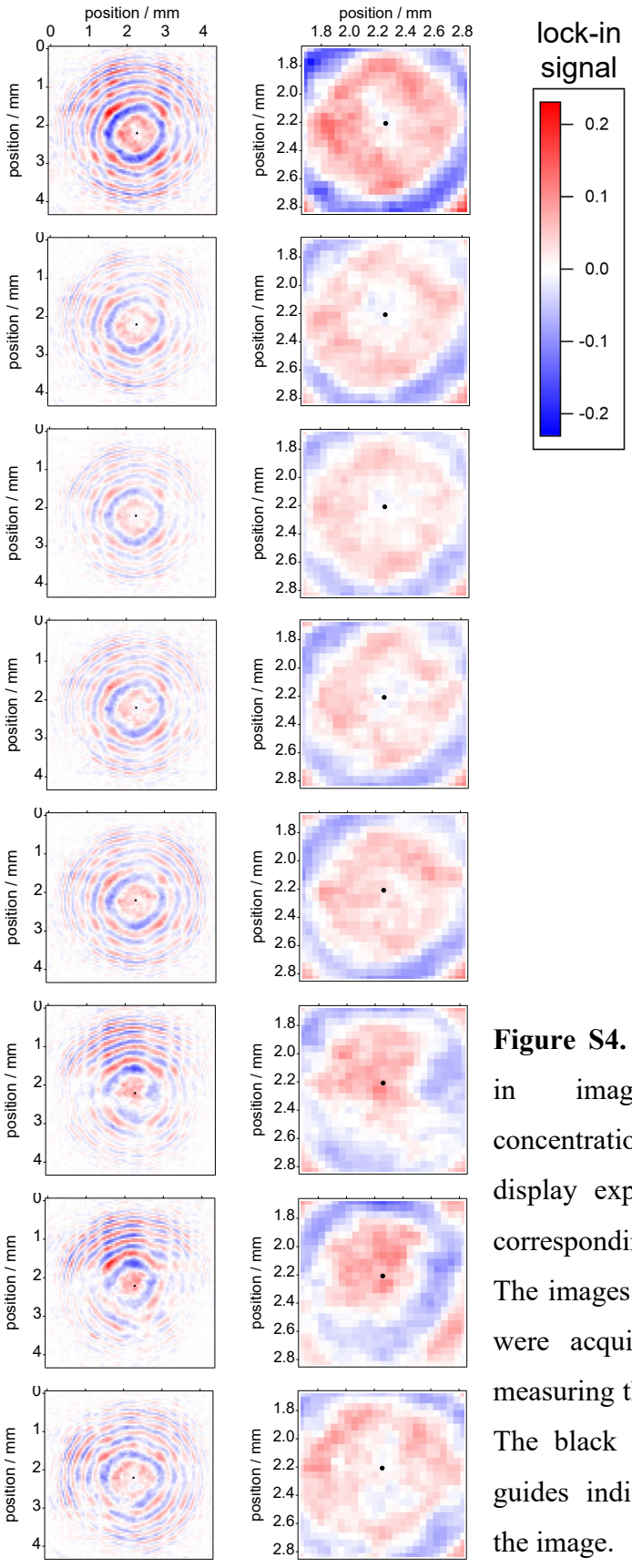

Figure S4. Representative lock-in images at various concentrations. The right panels display expanded views of the corresponding images on the left. The images in the bottom panels were acquired at 0 nM *after* measuring the 1000 nM solution. The black dots serve as visual guides indicating the center of the image.

Figure S4 presents 64 consecutive lock-in images acquired in a single acquisition of a 1000 nM solution. Although the intensity fluctuated substantially from image to image, the pattern shape remained consistent, implying that the pattern recognition-based analysis is more robust to noise than intensity-based methods.

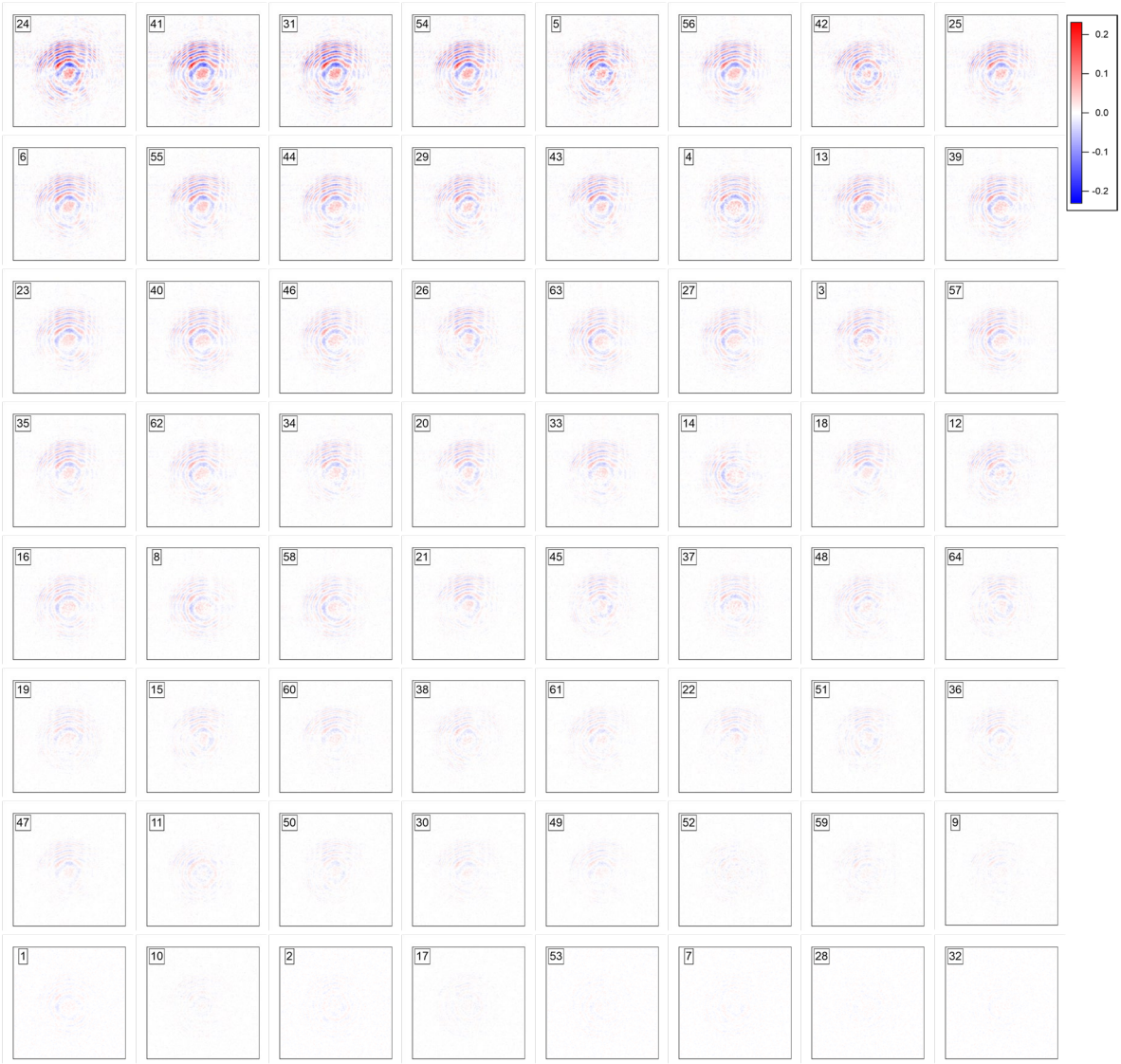

Figure S5. Representative lock-in images consecutively observed in a single acquisition, ordered by standard deviation (*i.e.*, from the strongest image) from left to right and from top to bottom. The numbers at the top left corner represent the frame number within the 64 images. The randomness of the frame number indicates that the intensity fluctuates significantly over time. Ticks and labels are omitted for clarity but are identical to those in Figure S4.

As the ring patterns shifted toward the upper left at higher concentration (Figures 2 and S3), the center position of the red patterns was analyzed for randomly selected datasets acquired at 0 nM and 1000 nM (Figure S5). It was determined as the point that minimized the standard deviation of the signal intensity measured along the circumferences of concentric circles with varying radii. However, the center position exhibited considerable variability, making it difficult to clearly distinguish even between the 0 nM and 1000 nM solutions. Notably, the center positions for 0 nM solutions measured at the beginning and the end of a single day overlapped within experimental error, indicating the observed shift is not attributable to temporal drift in the experimental setup.

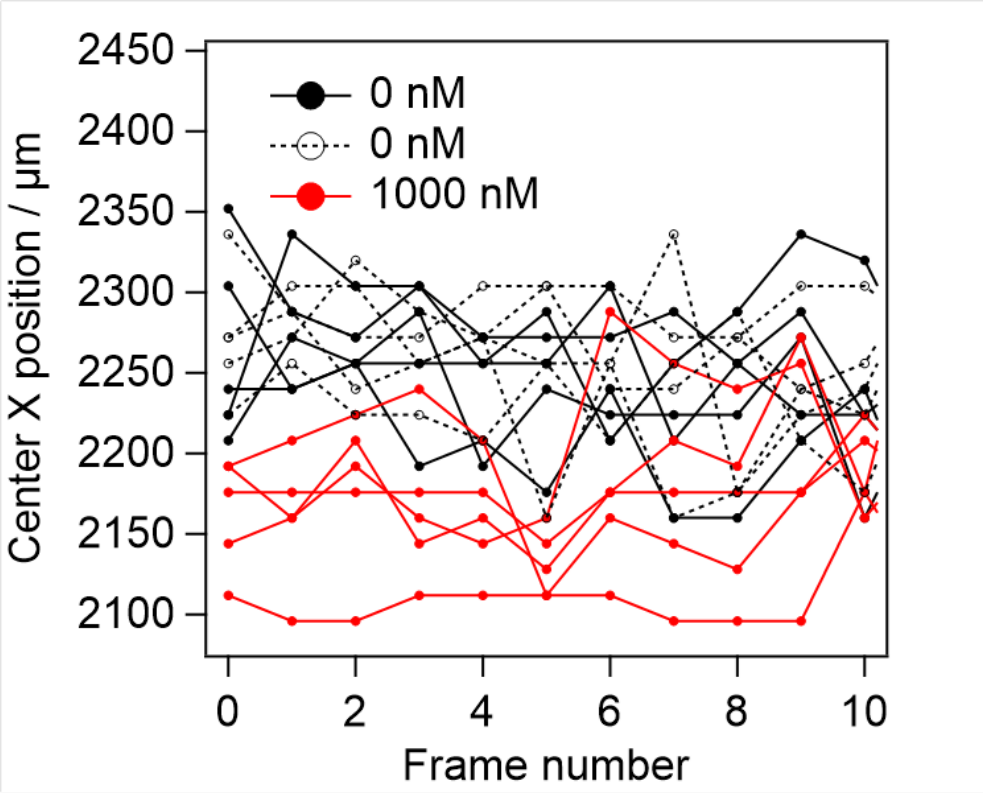

Figure S6. Center position of red patterns in the X direction (smaller values correspond to the left side of the image) for randomly selected images obtained for 0 nM (black) and 1000 nM (red) solutions. For the 0 nM data, filled circles represent measurements taken at the beginning of the day, while open circles represent those taken at the end of the day.

S7. Deep-learning analysis

Lock-in images obtained at concentrations of 0–30 nM measured on Days 1–4 and 0–10 nM on Day 5 were used for model training. As a general rule, images acquired on the same experimental day as the concentration estimation were excluded from the training set to avoid trivial correlations arising from day-specific experimental conditions. An exception was made for the highest-concentration data (30 nM) from the test day (*i.e.*, the day on which concentration estimation was performed), which were included in the training set as a coarse alignment reference to mitigate inter-day domain shift; this reference was not treated differently in the training procedure, but provides a consistent high-signal anchor that facilitates learning of day-dependent offsets without constraining concentration scaling at lower levels, complementing the intra-day drift control described in Section S3. After the 4-pixel downscaling described in Section S4, a 90 × 90 pixel region containing the concentric ring pattern was extracted to reduce the computational cost. Because the signal intensity fluctuated substantially even within each set of 64 consecutively acquired images—and in some cases the photothermal modulation was negligibly weak (Figure S4)—each set of 64 images was treated as a single dataset. To construct the input for the learning model, the 10 images exhibiting the strongest photothermal modulation within each dataset were selected and averaged.

During preliminary tests, we evaluated several input representations that retained the original signal intensity. However, we found that when intensity information was included, the model tended to rely predominantly on overall signal strength rather than on the spatial deformation and displacement of the pattern, which are central to the present detection concept. This behavior led to reduced robustness against day-to-day variations and poorer generalization performance. To encourage the model to focus on the spatial features of the photothermal signal

rather than its absolute intensity, the averaged images were therefore binarized to ±1 according to the sign of the lock-in signal at each pixel.

Given that both the shape and position of the observed pattern are expected to correlate with solute concentration (Figure 2), the deep learning model was designed to capture both local spatial features and their global arrangement. The architecture consists of a convolutional neural network (CNN) followed by a fully connected regression pathway. The input to the model was a single-channel lock-in image (90 × 90 pixels), preprocessed as described above. To emphasize spatial deformation and displacement of the photothermal pattern, edge information was incorporated by applying Scharr operators to generate horizontal and vertical gradient maps. These gradient components were combined with the original image to form a three-channel representation with a size of 90 × 90 × 3. The CNN backbone comprised four convolutional layers with decreasing numbers of filters (16, 8, 4, and 2), each followed by a leaky-ReLU activation and 2 × 2 max-pooling. This structure progressively reduced the spatial resolution while extracting hierarchical spatial features. After the final pooling operation, the feature maps had a size of 5 × 5 × 2, which were then flattened and passed through two fully connected layers (32 and 16 units). Dropout and L2 regularization were applied to both convolutional and dense layers to suppress overfitting and improve generalization. The model was trained in a supervised regression framework, where the target variable was the solute concentration transformed using a log(1 + concentration) representation to stabilize learning across a wide dynamic range. The Huber loss function ($\delta = 1.0$) was employed to provide robustness against occasional outliers, and model parameters were optimized using the Adam optimizer with a learning rate of 1×10^{-3}. All models were constructed and trained within the Keras framework.[9] For each training run, the available dataset was randomly divided into training and validation subsets, with 10% of the data

reserved for validation. Training was monitored using the validation mean absolute error, and early stopping was applied to terminate training when no improvement was observed for 25 consecutive epochs, with the model parameters restored to those corresponding to the best validation performance. To reduce sensitivity to random initialization and stochastic optimization effects, the training procedure was repeated using ten different random seeds. Each trained model produced an independent concentration estimate, and the final reported value was obtained by averaging the outputs of these ten models, thereby improving robustness and reducing variance in the predicted concentrations.

References

1 S. Urashima, and R. Kusaka. *Analyst* 2025, ***150***, 819.

2 F. J. Dietz, J. J. De Groot, and E. U. Franck. *Bunsenges. Phys. Chem.* 2010, ***85*** (11), 1005.

3 G. S. Kell. *J. Chem. Eng. Data* 2002, ***20*** (1), 97.

4 D. C. Ginnings, and G. T. Furukawa. *J. Am. Chem. Soc.* 1953, ***75*** (3), 522.

5 E. H. Ratcliffe. *Brit. J. Appl. Phys.* 1959, ***10*** (1), 22.

6 H. Ye, X. Li, and C. Jiang. *Int. J. Appl. Glass Sci.* 2022, ***13*** (4), 664.

7 B. T. Poe, C. Romano, and G. Henderson. *J. Non Cryst. Solids* 2004, ***341*** (1-3), 162.

8 S. Belgacem, and H. Galai. Cement Refractory Bricks Characteristics: The Importance of Mineralogical Quantification in the Evaluation of the Refractory Bricks Corrosion. In *Refractory Materials: Characteristics, Properties and Uses*, Bryant, C. Ed.; Nova Science Pub Inc, 2018.

9 F. Chollet et.al., *Keras*; 2015. https://keras.io