Localization Microscopy with Atomic Precision[*]

Yuqin Duan, et.al.

(Dated: January 1, 2025)

CONTENTS

---

[*] A footnote to the article title

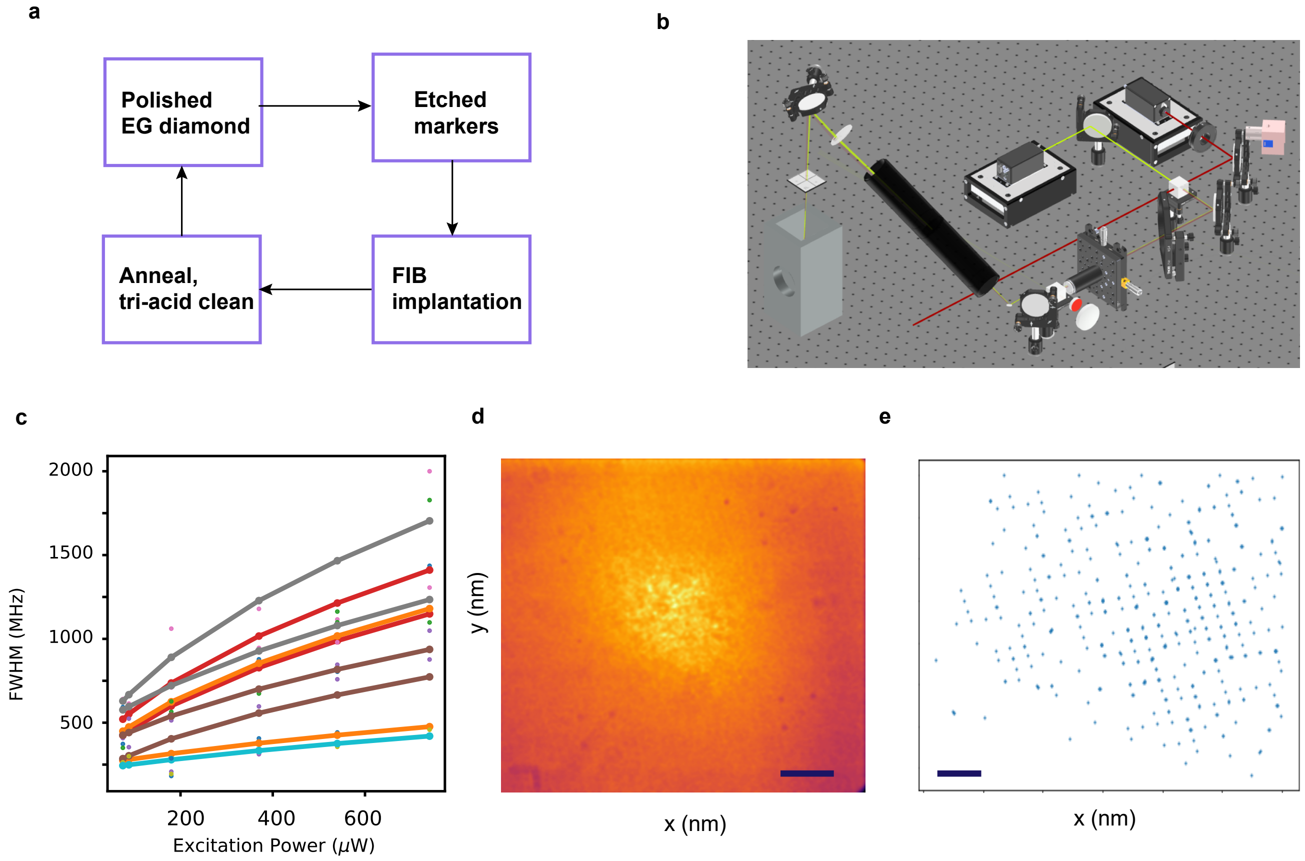


FIG. 1. **a**, Flow chart of sample fabrication process. **b**, CAD model of our custom-built SMLM setup. **c**, Power broadening of the ZPL linewidth. Each color corresponds to a different emitter operating at different $\omega_i/2\pi$. **d** Diffraction-limited EMCCD image showing unresolved emitter clusters. Scale bar: 10 μm and **e** Super-resolved clusters' localization with a period around 400 nm achieved using widefield PLE. scale bar of 1.5 μm

I. MATERIALS AND METHOD

A. Sample preparation

The overview of sample preparation is shown in Fig. 1a. The sample preparation workflow is adapted from Wan *et al.* [1]. We start with a polished homoepitaxial (EG) diamond substrate, onto which QR code markers are etched to serve as precise location references. The fabrication process, depicted in Fig. 1b, involves the following key steps: (i) growth of a silicon nitride (Si_3N_4) mask on the diamond substrate; (ii) electron beam lithography to write the QR pattern; (iii) cold development to accurately reveal the pattern; and (iv) dry etching to transfer the QR pattern from the Si_3N_4 mask into the diamond. A 1 nm thick gold layer is then deposited via sputtering to provide conductive pathways for subsequent ion implantation. Focused ion beam (FIB) implantation of Si^{++} ions into a 1 μm patterned array on the diamond surface is performed using the Raith Velion FIB system. The process is completed with annealing and tri-acid cleaning, which activate the silicon vacancy (SiV) centers and remove residual impurities, ensuring high-quality quantum emitters integrated within the diamond lattice.

B. SMLM

We use a homebuilt cryogenic widefield SMLM microscope, depicted in Fig. 1(b). In contrast to a conventional confocal setup, our widefield approach incorporates an additional telescope in a 4f arrangement that relays the excitation beam to the objective's back aperture. This design ensures uniform illumination across the entire field of view, enabling simultaneous imaging of multiple single emitters.

To validate the functionality of widefield PLE, we fabricated a sample with SiV centers arranged in a 400 μm array. Figure 1d presents diffraction-limited optical images of this sample. Given that the diffraction limit is calculated as

$$\sigma_0 = \frac{\lambda}{2\,\mathrm{NA}} = 295 \text{ nm}, \tag{1}$$

resolving individual clusters using conventional methods is inefficient. By applying widefield PLE, we acquire a series of EMCCD images at each frequency sweep, denoted as $I(x, y, \omega)$. We then perform region grouping by determining the centroid of each bright region using the following equation:

$$\mathbf{c} = \frac{1}{V}\left(\sum_{i=1}^{V} x_i, \sum_{i=1}^{V} y_i, \sum_{i=1}^{V} \omega_i\right), \tag{2}$$

where V represents the number of voxels in the region. Figure 1e displays $\sum_\omega \mathbf{c}$, demonstrating that each cluster is resolved below the diffraction limit.

SMLM using widefield PLE is limited by spectral resolution when the spontaneous emission signals of SiV centers overlap. This overlap can result from either homogeneous broadening or power broadening. Homogeneous broadening is constrained by the emitters' intrinsic spontaneous emission rates, whereas power broadening can be more readily optimized. To address power broadening, we incorporated a half-wave plate into the resonant laser path and measured its power immediately before it enters the objective. Figure 1c illustrates the PLE linewidth as a function of resonant laser power. The power broadening is described by the equation:

$$\Delta\omega/2\pi = \Gamma_0 + \Gamma\sqrt{1 + \frac{P}{P_s}}, \tag{3}$$

where P is the resonant laser power, Δ_ω is the measured full-width half maximum (FWHM) of PLE spectrum, Γ_0 represents spectral diffusion, P_s is the saturation power corresponding to the rate of optical cycles, and Γ is the linewidth in the limit of vanishing power [2]. From our fitting, we achieve a minimum linewidth of 192 MHz, which aligns with the lifetime-limited broadening.

II. DIGITAL TWIN LOCALIZATION

To achieve Ångström-level precision and accuracy in localizing a single emitter, we leverage the unique properties of color-center emitters, which differ significantly from biological specimens. Unlike biological samples, color-center emitters are immobile optical fixed-dipoles. The polarization of these dipoles influences the imaging PSF, introducing asymmetry. This asymmetry can cause the intensity peak to deviate substantially from the true center of the dipole. Moreover, physical imperfections within the system, if misrepresented or overlooked, can further introduce accuracy errors.

To eliminate system bias and accurately determine the limits of solid-state imaging systems, we have developed a digital twin that meticulously models the entire optical imaging setup. A digital twin is a virtual representation of a physical object or process, continuously updated with real-time data from its physical counterpart. Unlike traditional digital simulations, a digital twin incorporates comprehensive environmental factors and provides live feedback from the physical system. This integration bridges the gap between physical experiments and mathematical models, enabling rapid and precise digital testing that would otherwise be time-consuming and costly in experimental settings.

Digital twins have seen widespread application across various industries, including smart manufacturing, power plants, smart cities, and even planetary modeling [3]. In the context of fundamental laboratory experiments, particularly those requiring extended durations and minimal error budgets, constructing a digital twin is essential. It facilitates the monitoring and optimization of experimental conditions in real time, ensuring high fidelity between the physical and virtual systems.

The schematic representation of our digital twin is illusstrated in Fig. 2. This schematic encompasses the wave equation governing fixed-dipole emission and the parameterization of the physical apparatus, including the system drift over time. To ensure the accuracy and reliability of the digital twin, we benchmarked it using simulation data and subsequently validated its performance through experiments conducted on SiVs. The figure of merit is localization accuracy Δr and localization precision σ. Δr measured the error between fitting position and its ground truth location, and σ is the standard error and location estimation. In this section we include an error budget of all calibrations involved in details.

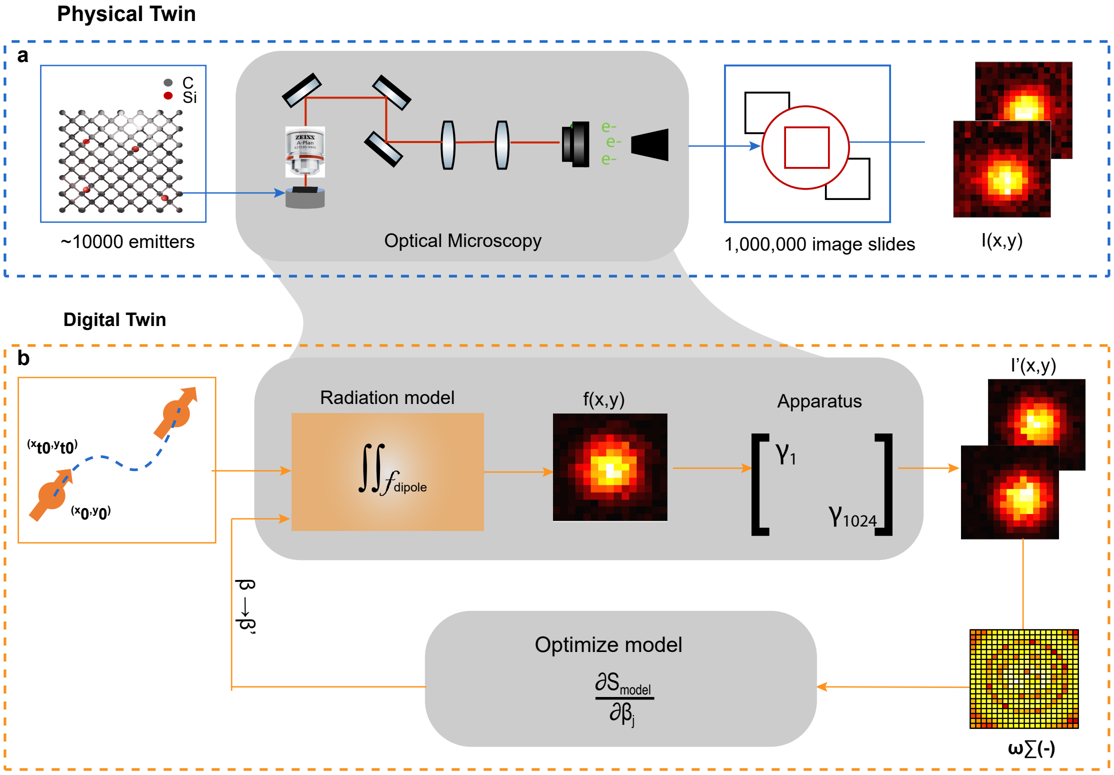


FIG. 2. Introduction to Digital Twin.**a** is the physical system of the fluorescence microscopy. We collect emitter's PSF from the camera as $I(x, y)$. **b** is the digital twin (DT) system of the physical system in **a**. Input is emitter's real location at time t is (x_{t0}, y_{t0}). The DT box includes a radiation model that analytically generate the farfield intensity pattern and the apparatus matrix derived from the physical twin. The output PSF intensity is $II(x, y)$. The parameters of the DT are tuned by minimizing the modeling residual so that the DT perform image generation accurately in the presence of physical noise.

A. Dipole radiation model

The PSF model used in this work follows the approach outlined in [4, 5]. It is derived from Snell's law and encompasses the dipole emitter's electric field, its propagation through the optical layers, and the subsequent intensity distribution in the image plane. Moreover, the model incorporates vertical astigmatism, which breaks the PSF symmetry above and below the focal plane, thereby allowing an estimate of the defocus.

This setup is equivalent to an air objective that is corrected for the presence of a cryostat window. We define the coordinate system such that the the z-axis coincides with the optical axis and the xy-plane with the focal plane. A dipole emitter located at the focal plane is described by its position (x, y, z) and orientation (θ, ϕ), where θ and ϕ are the inclination and azimuthal angles, respectively. The resulting PSF derived from the model is denoted as $E(x, y)$.

To understand the contribution of each parameter to localization accuracy (Δr), Fig. 3 provides a summary of fitting errors. Fig. 3a illustrates the fitting of a dipole emission profile using a simplified 2D Gaussian model, where the localization error (Δr) saturates at approximately 13 nm. In contrast, when employing a full dipole model, the accuracy improves with an increasing number of collected photons, ultimately achieving sub-Ångström. Fig.3b presents an error map that highlights the effects of polarization mismatch on localization accuracy.Fig. 3c and Fig. 3d explore the impact of z-axis variation and system defocus, respectively, on Δr. Notably, the localization error (Δr) exhibits a linear relationship with z-axis error. This linearity arises because the z-component contributes directly to the phase of the electric field $E(x, y)$, thereby influencing the overall localization accuracy.

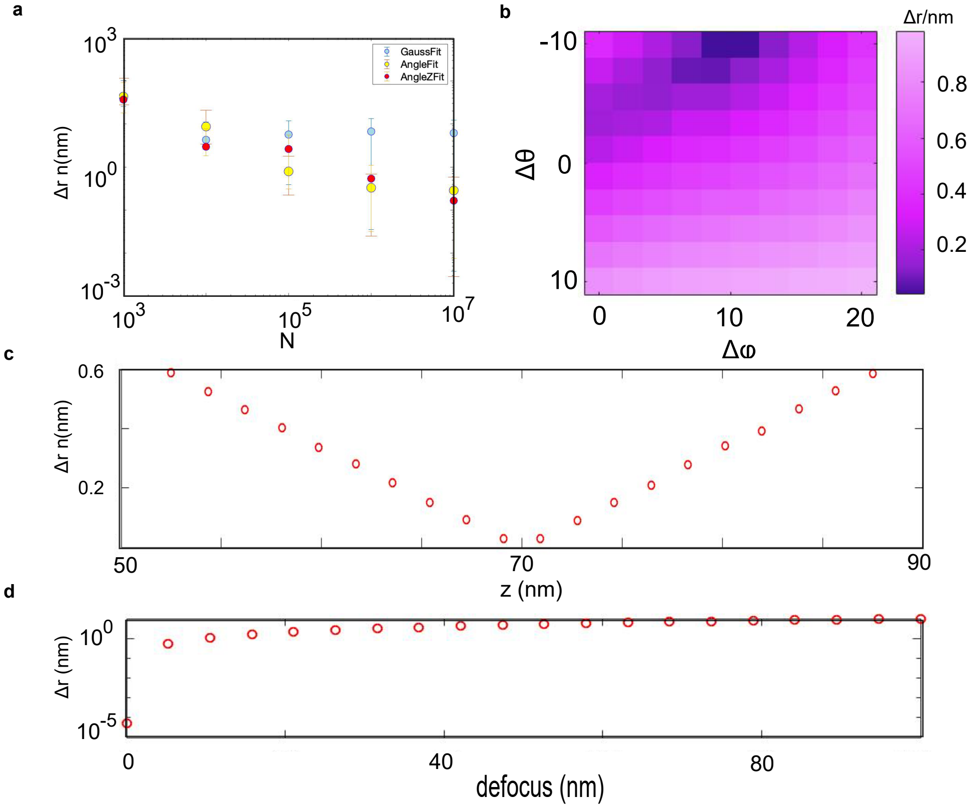


FIG. 3. **a** Localization accuracy (Δr) for different localization models: 2D Gaussian (blue), polarization-inclusive model (yellow), and model incorporating both polarization and z-axis information (red) across various photon counts (N). **b**, A color map of Δr with polarization mismatching in both θ and ϕ. **c** and **d** illustrate the impact of z-axis variation and system defocus, respectively, on the fitting error (Δr)

B. Physical system calibration

With a comprehensive radiation model in place, the next step is to minimize localization errors by calibrating the physical hardware. The schematic of the home-built widefield SMLM system is shown in Fig. 1b. Emission from the sample is collected by an objective and captured by an emCCD camera. Our primary goals are to characterize and correct optical aberrations and to map out the noise distribution across the entire field of view.

1. *Optical aberration*

Aberrations arising from imperfect objectives or lenses can significantly increase localization errors [6]. To quantify and correct these distortions, we used a reference sample featuring a 1x1 μm etched grid (Fig. 5a). By comparing the measured grid pattern with its expected coordinates, we can derive both radial and tangential distortions, modeled as follows:

$$x_d = x \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 xy + p_2(r^2 + 2x^2) \tag{4}$$

$$y_d = y \cdot (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y^2) + 2p_2 xy \tag{5}$$

where (x, y) are the normalized undistorted coordinates and (x_d, y_d) are the distorted coordinates, k_1, k_2, k_3 are the first-, second-, and third-order radial distortion coefficients, and p_1, p_2 are first- and second- order tangential distortion coefficients.

To perform the calibration, we use a standard Python package that automatically fits the distortion parameters to the measured grid. The resulting coefficients, listed in Table II B 1, enable us to apply a correction transform in

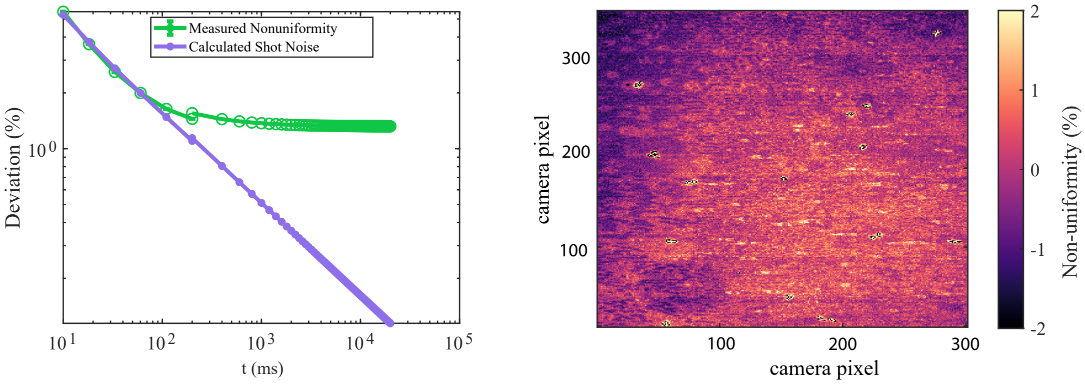


FIG. 4. Interpixel photoresponse non-uniformity (PRNU) characterization. **a**, Average of pixel deviation with increasing of integration time t. The purple line is the shot noise scale or pixel deviation that scales proportional to $\sqrt{N}$, while the green line is the measured deviation that saturates when PRNU becomes the dominant noise soruce. **b**,Inter-pixel EMCCD PRNU maps across 301x401 pixels calibrated as explained in the text.

post-processing or real-time imaging workflows. This correction ensures that the image data align accurately with the physical sample coordinates, thereby reducing localization errors and improving the overall reliability of our SMLM measurements.

k_1	$3.46\text{x}10^{-13}$
k_2	$1.15\text{x}10^{-21}$
k_3	$1.90\text{x}10^{-31}$
p_1	$5.12\text{x}10^{-11}$
p_2	$3.01\text{x}10^{-11}$

TABLE I. Caption

2. Interpixel photoresponse non-uniformity

In our widefield SMLM setup, the output signal is recorded by an electron-multiplying CCD (emCCD) camera. Photons incident on the sensor are first converted to electrons via the quantum efficiency (QE) of each pixel, then electronically amplified, and finally mapped into digital units. Three major noise sources affecting the PSF model of the camera are: (1) interpixel photoresponse non-uniformity (PRNU), (2) dark noise, and (3) readout noise.

PRNU refers to pixel-to-pixel variability in the camera's QE, which leads to local differences in brightness and can introduce systematic localization errors if uncorrected. Previous studies [7] reported that inhomogeneities in PRNU can be comparable to the Gaussian width of the PSF. Hence, accurately accounting for PRNU is crucial for high-precision localization.

To experimentally measure the PRNU of our emCCD, we illuminate the camera with a uniform parallel light source. We acquire 100 images with the light source on then capture another 100 images with the light off to measure the background. The response of each pixel was measured as the average signal (light on) minus the average background (light off). Fig. 4a shows the measured standard deviation of pixel responses (green curve) at different integration times t compared against simulated shot-noise-limited behavior (purple curve). We observe that the standard deviation between pixels is approximately 2.2%. Fig. 4(b) spresents the spatial map of this PRNU, highlighting systematic variations across the sensor.

In all subsequent PSF analyses, we incorporate this PRNU map to correct for pixel-to-pixel response variations. This step ensures that residual camera non-uniformities do not degrade the accuracy of our localization measurements.

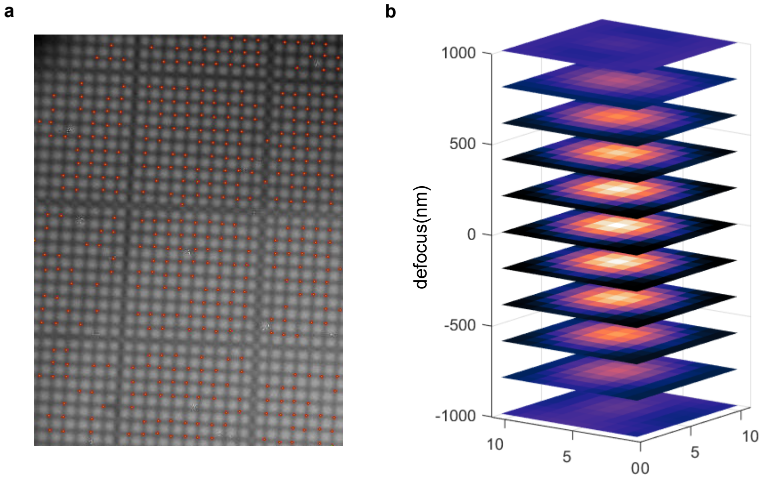


FIG. 5. Spatial aberration and astigmatism. **a**, White light fluorescence image of the calibration sample. The red dots indicate the fitted grid indices used for extracting aberration parameters. **b** PSF at different defocus levels.

3. *Dark and readout noise*

Next we model the noise model of the camera. After photon-electron conversion, the total variance of the signal is modeled as [8]:

$$\sigma_{tot}^2 = A^2M^2F^2(\sigma_{signal}^2 + \sigma_D^2)t + A^2\sigma_{readout}^2 \tag{6}$$

where A is the general video gain, M is the electron amplification, F is the excess noise factor.

To characterize the dark noise and readout noise, we first turn off all light sources ($\sigma_{signal} = 0$), and set emGain = 1500 to match with our experimental setting. We collected 800 images and Fig. 6a shows the resulting variance of each pixel. Here $\sigma_{tot} = A^2\sigma_{readout}$ with mean $\sigma_{readout} = 146.43$ image intensity. Next, we sweep across different acquisition time $t = [500, 1000, 1500, 2000]$ ms. For each time step, we acquired 5 batches of 800 sample images and repeat the variance calculation σ_{ij} within each batch. Subsequently, we fitted Eq. 6 to σ_{ij} using gradient method, and confirm the fitting result by R-square. Fig. 6b shows R-square distribution across 1024 x 1024 pixels and Fig. 6c shows σ_{ij} over t from six selective pixels. The result shows 99.06% of pixels fitting are above 0.99.

C. Digital Twin Generation

By building the digital twin for the physical SMLM system, we are able to examine localization accuracy and precision through the digital twin before hours of experiments. We defined a dipole located at (x_0, y_0) with polarization (θ_0, ϕ_0) and generated its PSF from the method described above. We fitted the PSF from the ideal dipole emission model with fitting parameters (x, y, θ, ϕ) and weighted the residual according to the inverse of pixel intensity. We compared the localization accuracy $(\Delta x, \Delta y)$ and localization precision $(\delta x, \delta y)$ with the radiation model and simplified gaussian model in different photon counts(N) conditions.

Fig. 7 shows the results of localization accuracy and precision. The first row shows the result when $N = 10^4$. (a) is the resulting PSF generated by the digital twin. The scale bar represents simulated camera units. We monte carlo 10^4 times of this process. (b) shows the fitting localization and (c) shows the histogram of the localization results. Green represents using the dipole model for the fit and purple represents using the Gaussian model. Both histograms follow a Gaussian distribution, with mean representing Δr and linewidth representing δr. The result shows by using the dipole model we achieved $\Delta r \pm \delta r = 0.57 \pm 1.11$ (nm) while purple is off by (3.96 ± 1.45) nm. Compare the bottom row when we group every 100 images so the resulting collection photon $N = 10^6$. The fitting localization distribution becomes more compact following the scaling law of $1/\sqrt{N}$. The result fitting of green is 0.007 ± 0.3 Åwhile purple is 3.53 ± 0.05 nm. With increasing N, not only does the localization precision improve, but also the fitting bias is reduced. This is because with better signal-noise-ratio (SNR), the phase retrival from PSF intensity at focal plane is more sensitive to polarization angle,

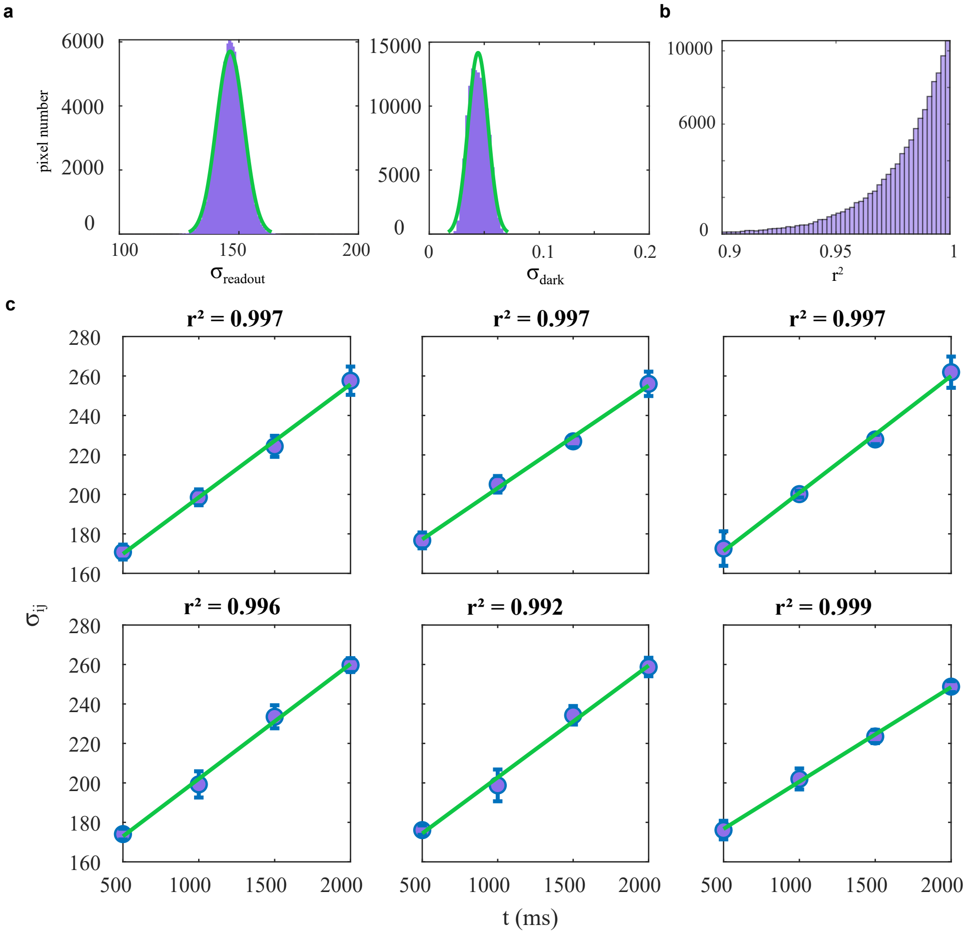


FIG. 6. Readout noise and dark noise.**a**, Histogram of pixel intensity when light is off. Readout noise is measured at very short integration time $t = 0.1$ ms. **b** is the histogram of dark noise fitting confidence r^2 around all pixels. **c** are examples of pixel intensity at different integration time. The green line is fitted by Eq. 6

1. Drift correction

The Kalman filter operates on a state-space model. The state transition model is defined as:

$$x_t = \mathbf{A}x_{t-1} + \mathbf{B}u_t + w_t \tag{7}$$

where x_t is the hidden state vector at time t, $\mathbf{A}$ is the state transition matrix, $\mathbf{B}$ and u_t are the control input matrix and vector, w_t is the process noise, assumed to be Gaussian with covariance Q. The observation model is defined as:

$$y_t = \mathbf{C}x_t + v_t \tag{8}$$

where y_t is the observation vector at time t, $\mathbf{C}$ is the observation matrix and v_t is the observation noise. The Kalman filter recursively performs two main steps. First is the prediction step to predict the next state and next covariance:

$$\hat{x}_{t|t-1} = \mathbf{A}\hat{x}_{t-1|t-1} + \mathbf{B}u_t \tag{9}$$

$$P_{t|t-1} = \mathbf{A}P_{t-1|t-1}\mathbf{A}^T + Q \tag{10}$$

Then second is an update step to minimize the covariance of the whole travel path. It involves computing the Kalman gain:

$$\mathbf{K}_t = P_{t|t-1}\mathbf{C}^T(\mathbf{C}P_{t|t-1}\mathbf{C}^T + \mathbf{R})^{-1} \tag{11}$$

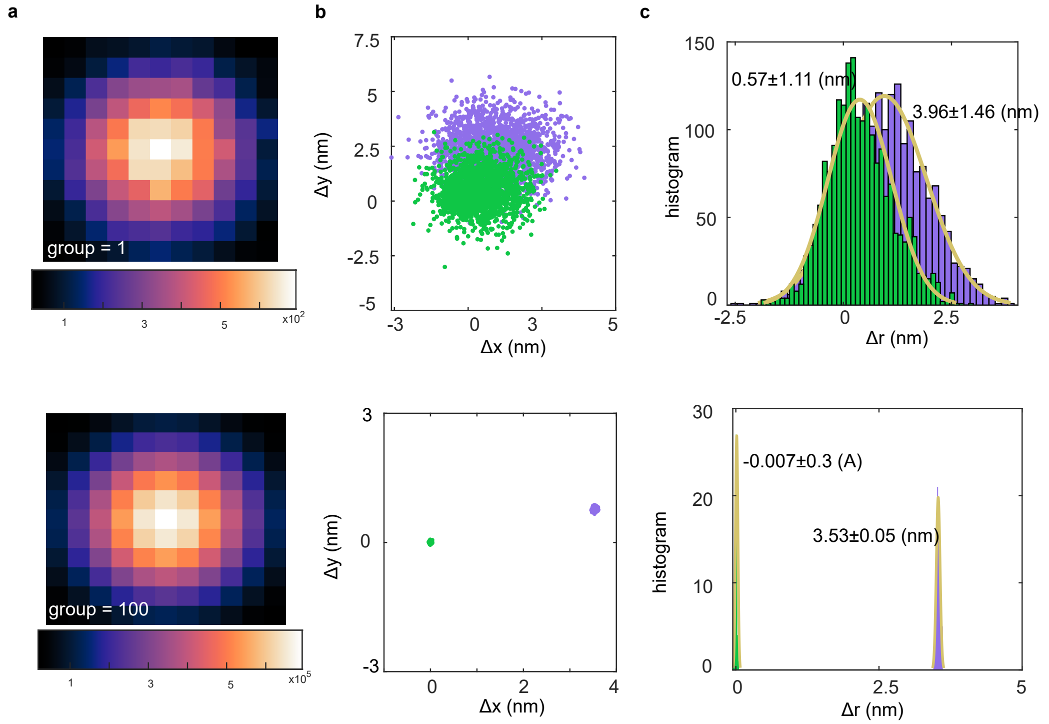


FIG. 7. Digital twin simulation of Silicon Vacancy (SiV) emitter localization accuracy and precision. Repetitive localization of simulated emitter emissions is performed to evaluate fitting methodologies. **a** Point Spread Function (PSF) obtained from a single frame with $N = 10^4$ photons and the aggregated PSF from 100 combined frames. **b** Individual emitter localizations generated through Monte Carlo simulations. **c** Histogram of localization density, where green indicates fitting results using the dipole model and purple indicates fitting results using a 2D Gaussian model.

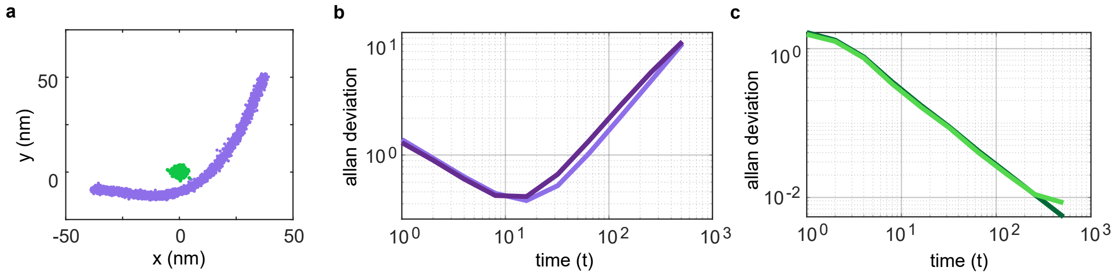


FIG. 8. Drift-induced localization error and correction via digital twin. **a** Localization trajectories in the (x, y) plane for each frame. Purple markers represent raw fitting results exhibiting drift, while green markers indicate drift-corrected fitting results achieved through the digital twin. **b** Allan variance analysis of localization error before drift correction. **c** Allan variance analysis after applying drift correction.

then updating the state estimator and covariance with observation:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \mathbf{K}_t(y_t - \mathbf{C}\hat{x}_{t|t-1}) \tag{12}$$

$$P_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{C})P_{t|t-1} \tag{13}$$

We further use the digital twin to test drift conditions. As in majority of SMLM experiment, with long-integration measurement, the systematic drift consequence is not negligible. Here we define the beginning position of the emitter is at (x_{t0}, y_{t0}) and end position (x_t, y_t), where $x_t = x_{t0} + 0.14t$ and $y_t = y_{t0} + 0.1t^2 + 0.15t$, t ranges from 1 to 700 seconds. Fig. 8(a) purple dots represent direct fitting results for each time frame. Without drift correction,

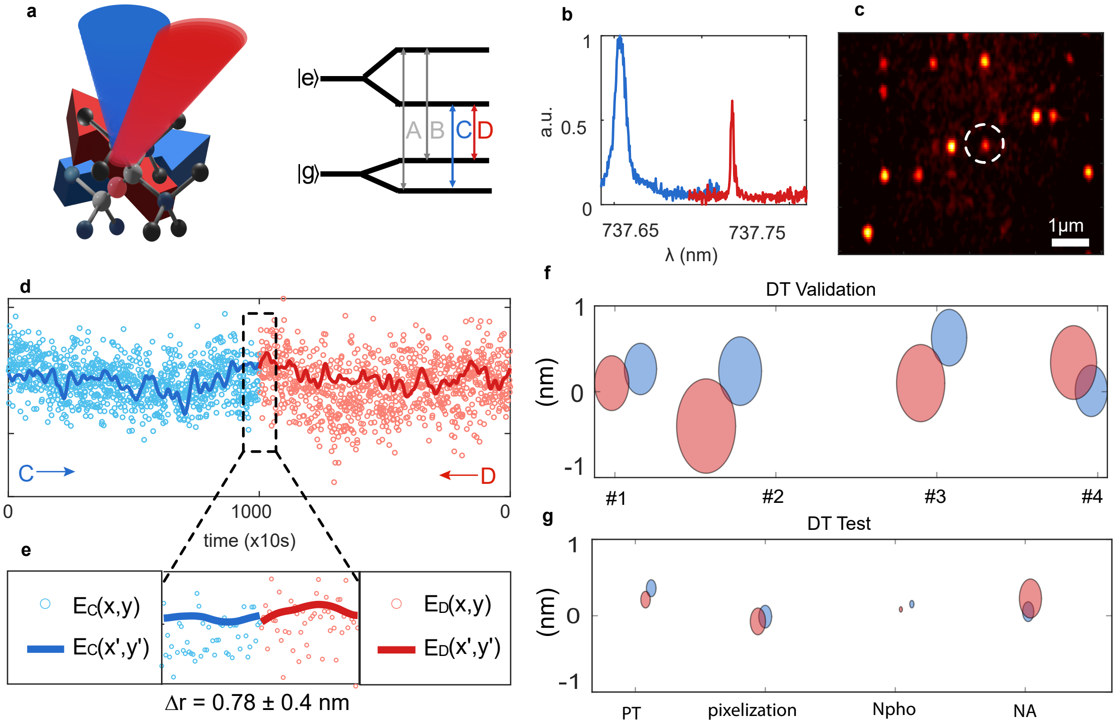


FIG. 9. Digital twin experimental validation. **a** shows SiV atomic structure and electronic band structure of four ZPL transitions (A,B,C,D). The emission polarization at C,D transition are marked in blue and red respectively. **b** is the wide-field PLE image projecting in spatial. **c** is the spectrum PLE of the single emitter highlighted in **b**. Blue and red corresponds to C, D transition. The markers in **d** are emitter's fitting location (x_t, y_t) at C D transition and the lines are digital twin predicted localization (x'_t, y'_t). **e** zooms in the end frame and the digital twin is able to distinguish two different polarization with localization accuracy of $\Delta r \pm \delta r = 0.78 \pm 0.4$ nm. **f** summarizes four single emitters localization result with same experimental condition. **g** summarizes four digital twin test at different physical conditions.

localization error is $\Delta r \pm \delta r = 10 \pm 10(\text{nm})$. To eliminate drift but also maximize SNR, we applied allan deviation test. Allan deviation is used to analyze the stability of time series data and is particularly useful in identifying and quantifying drift and noise characteristics in localization microscopy. It is defined as:

$$\sigma_x^2(\tau) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} (\bar{x}_{i+1} - \bar{x}_i)^2 \tag{14}$$

where $\bar{x}_i$ is the average position over the i-the time interval of duration τ, and M is the total number of such intervals. Fig. 8(b) shows the allan deviation of $\delta_t x$ and $\delta_t y$ from purple dots in (a). δ_t reaches a floor when $t \sim 10$s, implying the maximum group number before drift dominates the error is at $t \sim 10$s.

We performed the same analysis for all three emitters (see Supplementary Information). The correlation coefficient between their predicted paths is 0.99, and the covariance is 0.01 nm^2, implying that their movements are strongly linearly related, likely due to common environmental factors affecting the sample stage. However, the magnitude of their individual displacements remains small after drift correction, confirming the

By reaching its best SNR, we then conducted drift correction through Kalman filter.

The digital twin approach is widely applicable across microscopy and sets a new standard for imaging analysis. These results will be built upon to advance various applications, including quantum sensing and single-photon detection, where subatomic resolution enables a new frontier in spatial resolution.

D. Digital Twin Test

To verify diamond color centers location at subdiffraction resolution is very challenging. Commonly used FRET and TEM, but diamond struggles because FRET varies inversely with the sixth power of the distance, requiring emitters distance < 10 nm, and TEM requires to have a very thin diamond to due to its wide bandgap. Therefore we need to find alternative method to verify the digital twin system with some reference. Here we look at SiV's unique feature: their electronic structure, shown in Fig.9(a), consists of two-fold spin-degenerate orbital doublet ground and excited states, resulting in four lines of the zero-phonon line (ZPL) at zero magnetic field and cryogenic temperatures. Specifically, the C and D transitions, marked in Fig. 9(a), correspond to two different polarizations, one in the [1,1,0] orientation and the other in the [1,1,1] orientation [9]. It implies that these two transitions have different PSF but share the same (x, y). We use this fact to test the digital twin system described above.

To address the C and D transitions, we examined the wide-field image in the low-dose (35 ions/μm) implantation region. We first sweep the frequency from 406. to 406. THz with a step of 25 GHz and record the wide-field image at each frequency. Fig.9b displays the projection of all images into x-y domain. The sparsity of the image implies that the mean number of emitters in this region is approximately one. Fig.9c shows the spectrum PLE of the targeted emitter circled in white in b, with blue and red lines indicating the C and D transitions at 737 nm.

We then integrated 10,000 image frames with an exposure time of 1 second per frame at both frequencies sequentially and used the digital twin to fit the emitter's position. Fig. 9d depicts the emitter's positions $(x, y)_C$ and $(x, y)_D$, marked by blue and red dots, integrated every 10 frames. The blue and red lines represent the corrected positions $(x', y')_C$ and $(x', y')_D$ after applying the digital twin drift correction.

Fig. 9e zooms in on the final frame for the C and D transitions. The fitting accuracy and precision result in atomic resolution, with $\Delta r \pm \delta r = (x, y)_C - (x, y)_D = 0.78 \pm 0.4$ nm. Although this is off from the ground truth where $\Delta r = 0$, the system has reached its shot noise limit and aligns with the digital twin's prediction (as detailed in the supplementary materials). To potentially eliminate this bias, one could use a higher-resolution camera, an immersive lens, or a higher quantum efficiency material to improve the signal-to-noise ratio (SNR).

III. THREE EMITTERS LOCALIZATION

Fig. 10 summarizes the localization performance of three emitters from Fig. **??**, focusing on the evaluation of drift-induced errors. We applied Equations 7-13 of the Kalman filter to correct for drift. In Fig. 10a, we compares the averaged PSFs over 8000 frames with and without drift correction. The PSF without drift correction (purple) exhibits a clear asymmetry, whereas the PSF after drift correction (green) is highly symmetric. Fig. 10b displays the (x, y) fitting results before and after drift correction. Different colors are used to represent various groupings and mappings of emitters, as illustrated in **Fig. 10c**. The Allan variance analysis, derived from Eq. 14, is also shown in **Fig. 10c**. Here, the purple line represents the raw fitting data, and the green line indicates the drift-corrected data. Without drift correction, the localization variance reaches a floor of $\sigma = 1.8$ nm at an integration time of $t = 20$ s. The Kalman filter effectively tracks the emitters' drift, as demonstrated in **Fig. 10d**, aligning the localization accuracy with the shot noise limit.

IV. DIGIT

A. Bayesian Framework

The bayesian statistics is defined as:

$$p(A|B) = p(B|A)\frac{p(A)}{p(B)} \tag{15}$$

where $p(A)$ is the prior probability of A, which is the naive guess of emitter localization of equally distributed over lattice point,

$$p(A) = \sum_{n \in Z} p_0 \delta(x - n) \tag{16}$$

$p(B|A)$ is the likelihood function, which can be interpreted from digital twin fitting PSF, $p(B|A) = f(\mu_x) = \frac{1}{2\pi\sigma^2} e^{-(\frac{\mu_x - \mu_0}{\sqrt{2}\sigma})^2}$. For sufficiently narrow Gaussian, $\mu_0 \to x$, the real location of the atom. Mathematically this is,

$$p(B|A) = p(\mu_x|x) = \frac{1}{2\pi\sigma^2} e^{-(\frac{\mu_x - x}{\sqrt{2}\sigma})^2} \tag{17}$$

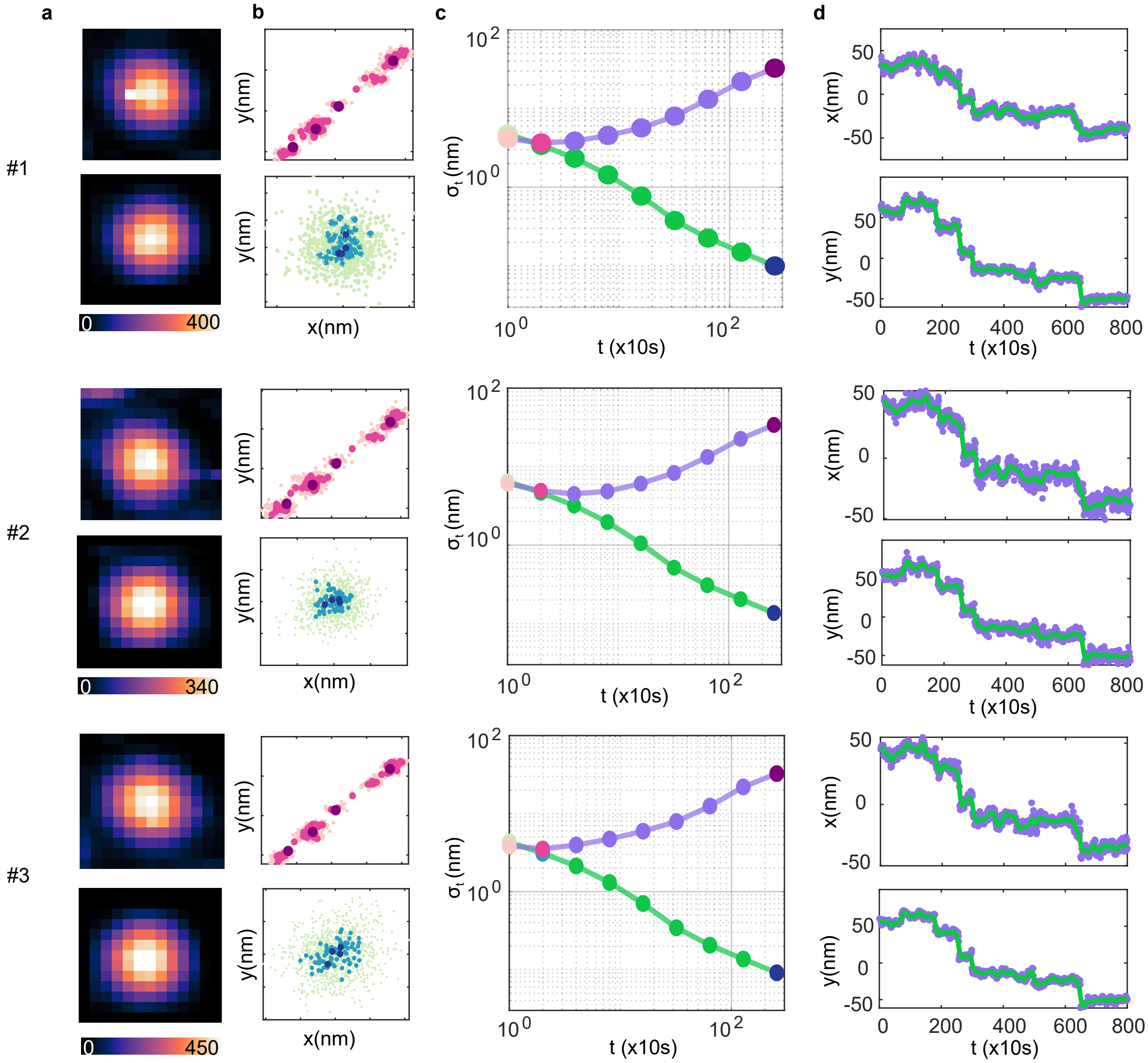


FIG. 10. Evaluation of localization performance for three Emitters. **a** Averaged PSF from 8000 frames, comparing results with (green) and without (purple) drift correction. The color bar represents intensity in camera units. **b** (x, y) localization results categorized into different groups, illustrating the positions of emitters before and after drift correction. **c** Allan variance analysis of the localization data presented in **b**, comparing raw (purple) and drift-corrected (green) results. The analysis highlights the reduction in localization variance achieved through drift correction. **d** Trajectories (x_t, y_t) of emitters, where purple markers denote raw localization data and green lines represent the predicted paths from the Kalman filter.

With prior knowledge, for example, by taking many similar measurements on a quantum dot nanoparticle, it is possible to determine σ. It is also possible to estimate σ using Fisher information and Cramer-Rao bound. The probability of evidence $P(B)$ using the law of total probability,

$$p(B) = \sum_i p(B|A_i)P(A_i) \tag{18}$$

. Substitute Eq. 16 and 17,

$$p(B) = p_0 \sum_{n \in Z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\frac{\mu_x - n}{\sqrt{2}\sigma})^2} \tag{19}$$

Therefore, the posterior $p(A|B)$ as in the main text of $P(x)$ is

$$p(A|B) = P(x) = \frac{e^{-(\frac{\mu_x - x}{\sqrt{2}\sigma})^2}}{\sum_{n\in Z} e^{-(\frac{\mu_x - n}{\sqrt{2}\sigma})^2}} \tag{20}$$

This basically gives a discrete Gaussian (numerator) that takes on the same value as the continuous Gaussian but is normalized differently. Intuitively, if $\sigma \gg a$, where a is lattice constant, $\int dx \to \sum_{n\in Z}$. But if $\sigma \ll a$, the summation and the integral differ significantly.

For an emitter pair of each emitter is at location x_1 and x_2 respectively, and distance r apart, following the same Bayesian statistic,

$$p(r|\mu_{x_1}, \mu_{x_2}) = \sum p(r|x_1, x_2) p(x_1|\mu_{x_1}) p(x_2|\mu_{x_2}) \tag{21}$$

$p(r|x_1, x_2)$ is the conditional probability depending on values of x_1 and x_2. Substitute Eq. 20 we get:

$$p(r|\mu_{x_1}, \mu_{x_2}) = \frac{1}{Z_1 Z_2} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{4\sigma^2}(\mu_{x_1} - \mu_{x_2} - r)^2} \sum e^{-\frac{1}{\sigma^2}[(x_1 - \mu_{x_1}) + \frac{\mu_{x_2} - \mu_{x_2} - r}{2}]^2} \tag{22}$$

where Z_1 and Z_2 are normalized prefactors so that $p(r|\mu_{x_1}, \mu_{x_2}) = 1$.

Adding this concept to the digital twin to reflect the real-world, we consider a two-dimensional lattice localization. In experiments, the full alignment of the diamond crystal lattice is inherently unknown. Therefore here we further consider diamond lattice localization. The model formulation is following: first we define lattice position with a rotation matrix $R(\theta) = \left(\begin{smallmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{smallmatrix}\right)$ and lattice offset $U = (u_x, u_y)$. Given lattice points (m_i, n_i), the emitters' position after rotation and applying the offset is calculated as

$$\mu_i = R(\theta) \cdot (m_i \mathbf{a}_1 + n_i \mathbf{a}_2) + U \tag{23}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the lattice base vectors. The goal is to find lattice rotation angle and offset, we use maximum likelihood estimator (MLE) to optimize it. The likelihood function for a single measurement is given by:

$$L(\theta, U, x_i, y_i) = \frac{1}{2\pi\sigma^2} e^{-\frac{||\mu_i - (x_i, y_i)||^2}{2\sigma^2}} \tag{24}$$

The total log-likelihood for all measurements K of all emitters M is:

$$\log L(\theta, U) = \sum_{i=1}^{M} \sum_{j=1}^{N} \left(-\frac{\|\mu_{Dij} - \mu_{ij}\|^2}{2\sigma^2} \right) - NM \log\left(2\pi\sigma^2\right)$$

Both analytical and numerical implementation of this DIGIT protocol can be found through our online code repository.

B. Simulation result

V. WIDEFIELD DIGIT

The flowchart of widefield DIGIT is described in the Methods. After widefield PLE is conducted at the targeted area referenced by the QR code, we applied a Gaussian filter as a low-pass filter before detecting image intensity group peaks to roughly localize emitters. The Gaussian filter is applied to the image projection $I'(x, y)$ by convolution. Mathematically it is defined as:

$$I''(x, y) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} I'(u, v) \cdot G(x - u, y - v) \tag{25}$$

where $G(x, y)$ is the Gaussian filter. To localize cluster's position so to collect each individual emitter's resonance frequency, we applied a fast-peak find algorithm. To make it more robust, the software interface also takes customized low-pass or high-pass ratio to justify how many clusters get detected as shown in Fig. 12b.

After identifying clusters' location, we integrate image intensity at the location over different frequencies to identify each emitter's resonance frequency. Some example of PLE spectrum is shown in Fig. 11. To find its resonance, we

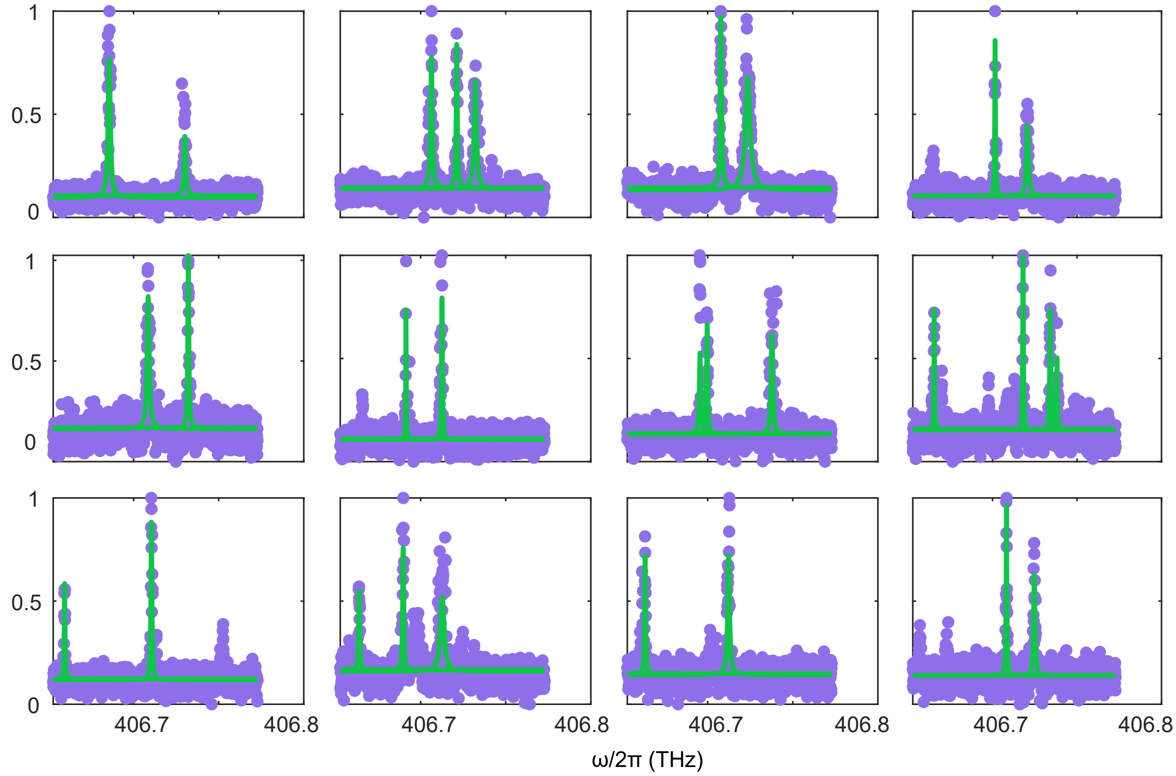


FIG. 11. PLE spectrum of selected clusters. Shown are the PLE spectra from 12 distinct clusters, with green lines representing the fits based on Eq. 26.

developed a customized function called lorentzianNPeak to accomodate multiple resonances at one cluster. Each individual peak is fit by:

$$f(\omega) = \frac{1}{\pi} \cdot \frac{\frac{\Gamma}{2}}{(\omega-\omega_0)^2 + \left(\frac{\Gamma}{2}\right)^2} \tag{26}$$

where ω_0 is the center peak as the emitter's resonance frequency and Γ is the full width at half maximum. The fitting results are shown in in Fig. 11.

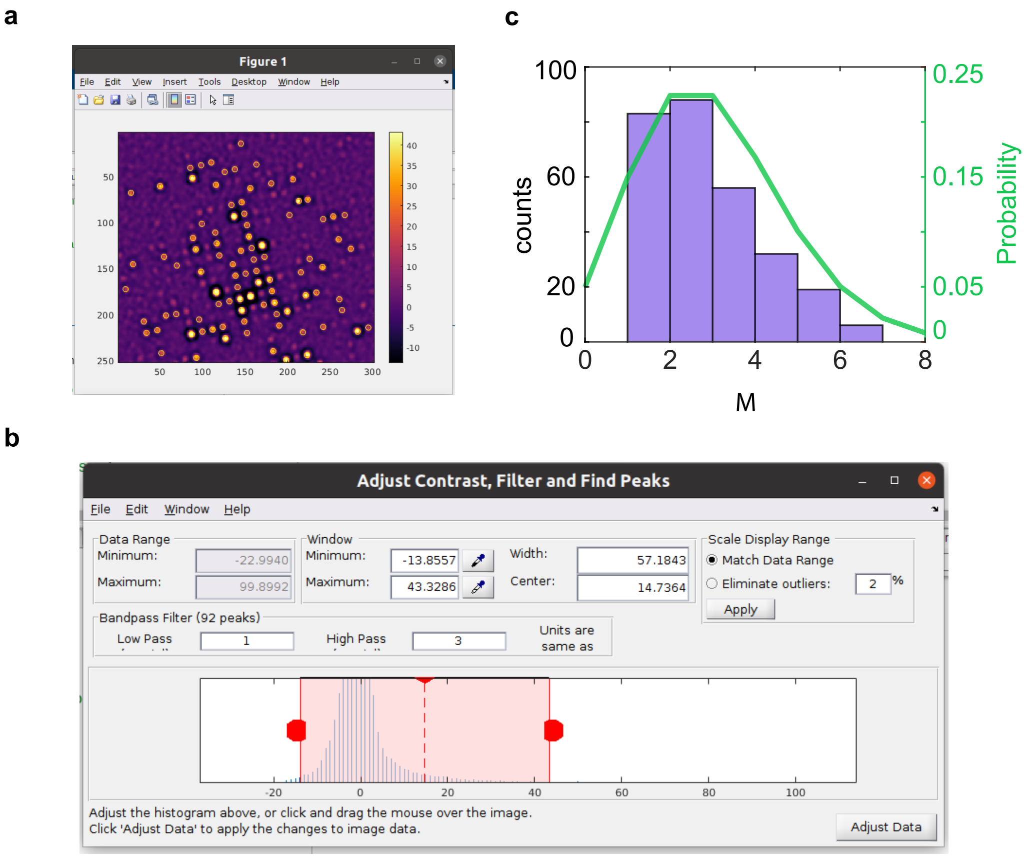
Applying Lorentzian fitting to all PLE spectra using Eq. 26, we generated a histogram of emitter counts (M) for each cluster, as shown in Fig. 11c. The histogram was fitted with a Poisson distribution, resulting in a mean emitter count of $M = 3$ within the area of interest. However, the fitting process appears to underestimate the number of emitters in regions with high emitter counts. This underestimation suggests that our spectral resolution is limited, potentially hindering the accurate characterization of clusters with multiple emitters.

VI. THEORY PROPOSAL OF OPTICALLY RESOLVED LATTICE STRUCTURE

DIGIT with atomic localization shows great potential to identify emitters' electronic orbital structure. Further, for a new material, whether there is a lattice, what lattice constant is can potentially achieved by optical measurement. By solving it, we can perform DIGIT with no prior knowledge needed, further extend to broaded community.

Assume there are M emitters in a space of L (Fig. 13a), then there are $\frac{M(M-1)}{2}$ pairs of distance vector. Similar to the lattice problem in computer science, by measuring all pairs information with an error σ, can we find the biggest periodicity that agrees with the measurement. This periodicity is also the crystal lattice length. Fig. 13 shows the in silicone of the proposal. First we apply a two-dimensional correlation function as:

$$C(\Delta x, \Delta y) = \frac{1}{N}\sum_{i=1}^{N} f(x,y) f(x_i + \Delta x, y_i + \Delta y) \tag{27}$$



Data Range
Minimum: -22.9940
Maximum: 99.8992

Window
Minimum: -13.8557
Maximum: 43.3286
Width: 57.1843
Center: 14.7364

Scale Display Range
● Match Data Range
○ Eliminate outliers: 2 %

Bandpass Filter (92 peaks)
Low Pass 1
High Pass 3
Units are same as

Adjust the histogram above, or click and drag the mouse over the image.
Click 'Adjust Data' to apply the changes to image data.

FIG. 12. Widefield PLE analysis interface. **a** and **b** display the widefield analysis interface with options to customize the low (σ_L) and high (σ_H) pass filter bandwidth thresholds, respectively. **c** Histogram of the parameter M, which follows a fitted Poisson distribution with a mean value of 3.

where $C(\Delta x, \Delta y)$ is the correlation function and $f(x, y)$ is emitters' localization (Fig. 13b). Fig. 13c shows the result of autocorrelation projected in one dimension. It's hard to tell the difference between the periodic and the random case and also challenging to extract periodicity directly. In order to solve that, we applied Fourier transform on Eq. 27, the result is $F(C) = \sum_{-\infty}^{\infty} C(\tau) e^{-j2\pi f \tau}$, where f is the frequency. Physically, this is the power spectrum density (PSD) and the result is shown in Fig. 13d. The left one clearly shows a distinguish peak at $f = 1/a$ while the random scenario's the most distinguish peak happens at $f = 0$ indicating lack of pattern.

To further evaluate the capability of our localization technique in resolving lattice structures, we analyzed the resolvability of the lattice constant (a) in relation to emitter localization uncertainty (σ) and emitter position density. Fig. 13e presents a comprehensive color map illustrating the conditions under which the lattice constant can be accurately resolved. The map delineates regions based on varying levels of localization precision and emitter density, indicating that lower localization uncertainties and higher emitter densities significantly enhance the ability to resolve the underlying periodic lattice. Specifically, the color gradient transitions from regions of poor resolvability (higher σ and lower density) to regions of excellent resolvability (lower σ and higher density), underscoring the interplay between these parameters in achieving atomic-scale imaging.

Complementing this analysis, Fig. 13f depicts the Cramér-Rao Lower Bound (CRLB) for the estimation of the lattice constant (a). The CRLB provides a theoretical benchmark for the minimum achievable variance in the estimation of a.

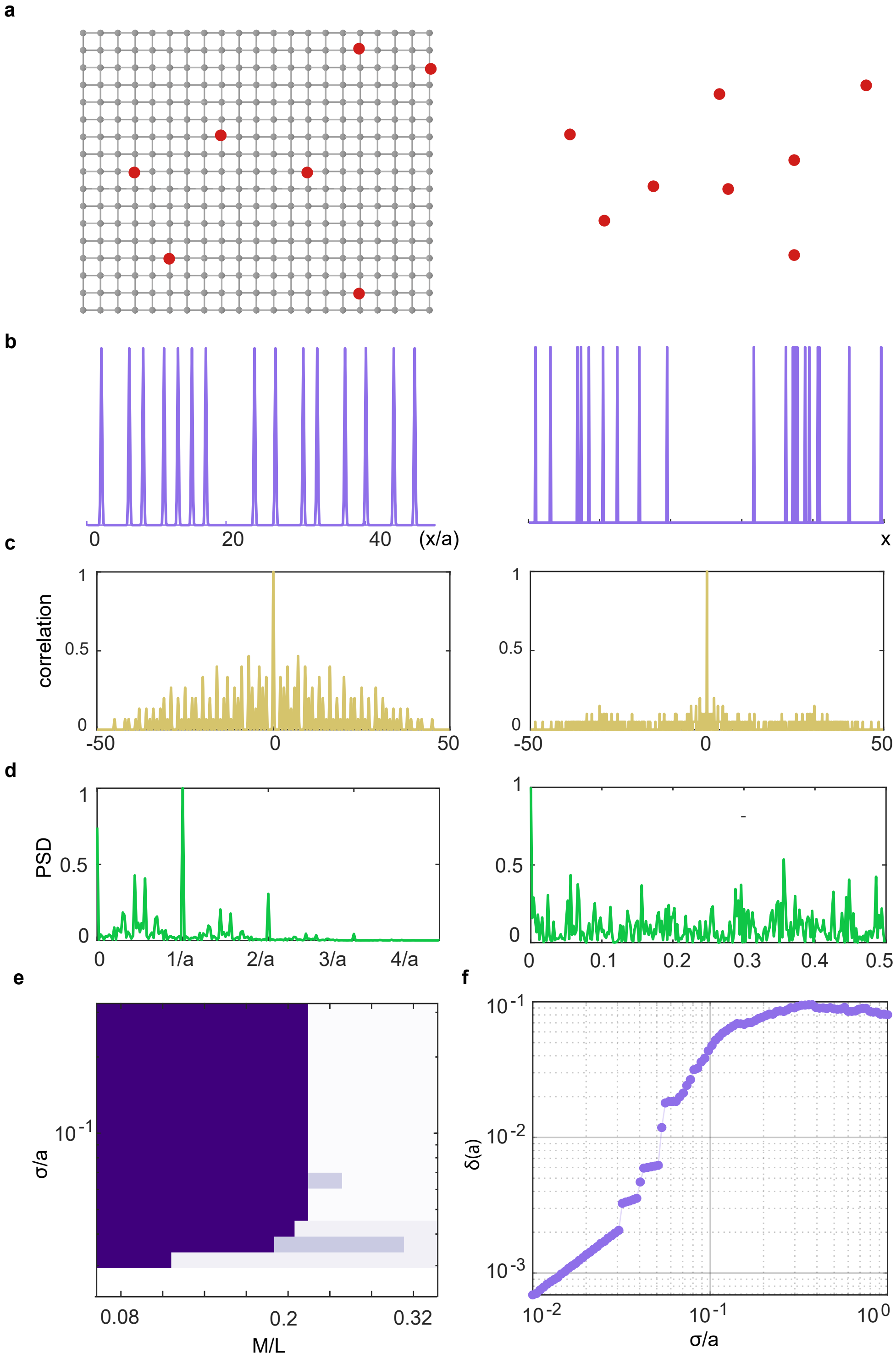


FIG. 13. Optically resolved lattice structures. The left panel represents a periodic lattice structure, whereas the right panel depicts a random emitter distribution. **a** Schematic illustration of the underlying emitter arrangements, comparing periodic and random configurations. **b** One-dimensional localization of each emitter with corresponding uncertainty (σ), demonstrating the precision of emitter positioning. **c** Auto-correlation functions calculated using Eq. 27 for both periodic and random emitter distributions, highlighting characteristic patterns. **d** Power Spectral Density (PSD) functions for periodic and random emitter arrangements, illustrating distinct spectral features. **e** Color map indicating the resolvability of the lattice constant (a) as a function of emitter localization uncertainty (σ) and emitter position density. This map delineates regions where the lattice constant can be accurately resolved. **f** Cramér-Rao Lower Bound (CRLB) for the estimation of the lattice constant (a), providing theoretical limits on the precision of lattice constant measurements.

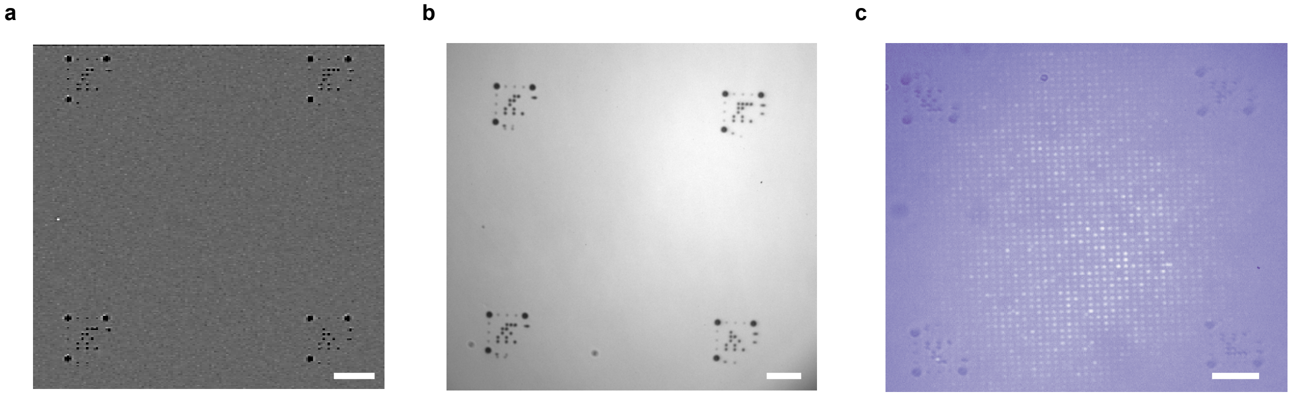


FIG. 14. Fabricated QR marker imaging from **a**, Scanning Electron Microscope (SEM) **b**, white-light and **c**, fluorescent microscopy

---

[1] N. H. Wan, S. Mouradian, and D. Englund, Two-dimensional photonic crystal slab nanocavities on bulk single-crystal diamond, Applied Physics Letters **112** (2018).

[2] A. Dréau, M. Lesik, L. Rondin, P. Spinicelli, O. Arcizet, J.-F. Roch, and V. Jacques, Avoiding power broadening in optically detected magnetic resonance of single nv defects for enhanced dc magnetic field sensitivity, Phys. Rev. B **84**, 195204 (2011).

[3] Z. Lv, S. Xie, Y. Li, M. S. Hossain, and A. El Saddik, Building the metaverse using digital twins at all scales, states, and relations, Virtual Reality & Intelligent Hardware **4**, 459 (2022).

[4] F. Hinterer, M. C. Schneider, S. Hubmer, M. López-Martinez, P. Zelger, A. Jesacher, R. Ramlau, and G. J. Schütz, Robust and bias-free localization of individual fixed dipole emitters achieving the cramér rao bound for applications in cryo-single molecule localization microscopy, PloS one **17**, e0263500 (2022).

[5] P. N. Petrov, Y. Shechtman, and W. Moerner, Measurement-based estimation of global pupil functions in 3d localization microscopy, Optics express **25**, 7945 (2017).

[6] K. L. Blythe, E. J. Titus, and K. A. Willets, Objective-induced point spread function aberrations and their impact on super-resolution microscopy, Analytical Chemistry **87**, 6419 (2015), pMID: 26011175, https://doi.org/10.1021/acs.analchem.5b01848.

[7] A. Pertsinidis, Y. Zhang, and S. Chu, Subnanometre single-molecule localization, registration and distance measurements, Nature **466**, 647 (2010).

[8] M. S. Robbins and B. J. Hadwen, The noise performance of electron multiplying charge-coupled devices, IEEE transactions on electron devices **50**, 1227 (2003).

[9] L. J. Rogers, K. D. Jahnke, M. W. Doherty, A. Dietrich, L. P. McGuinness, C. Müller, T. Teraji, H. Sumiya, J. Isoya, N. B. Manson, and F. Jelezko, Electronic structure of the negatively charged silicon-vacancy center in diamond, Phys. Rev. B **89**, 235101 (2014).

VII. EXTENDED DATA FIGURE

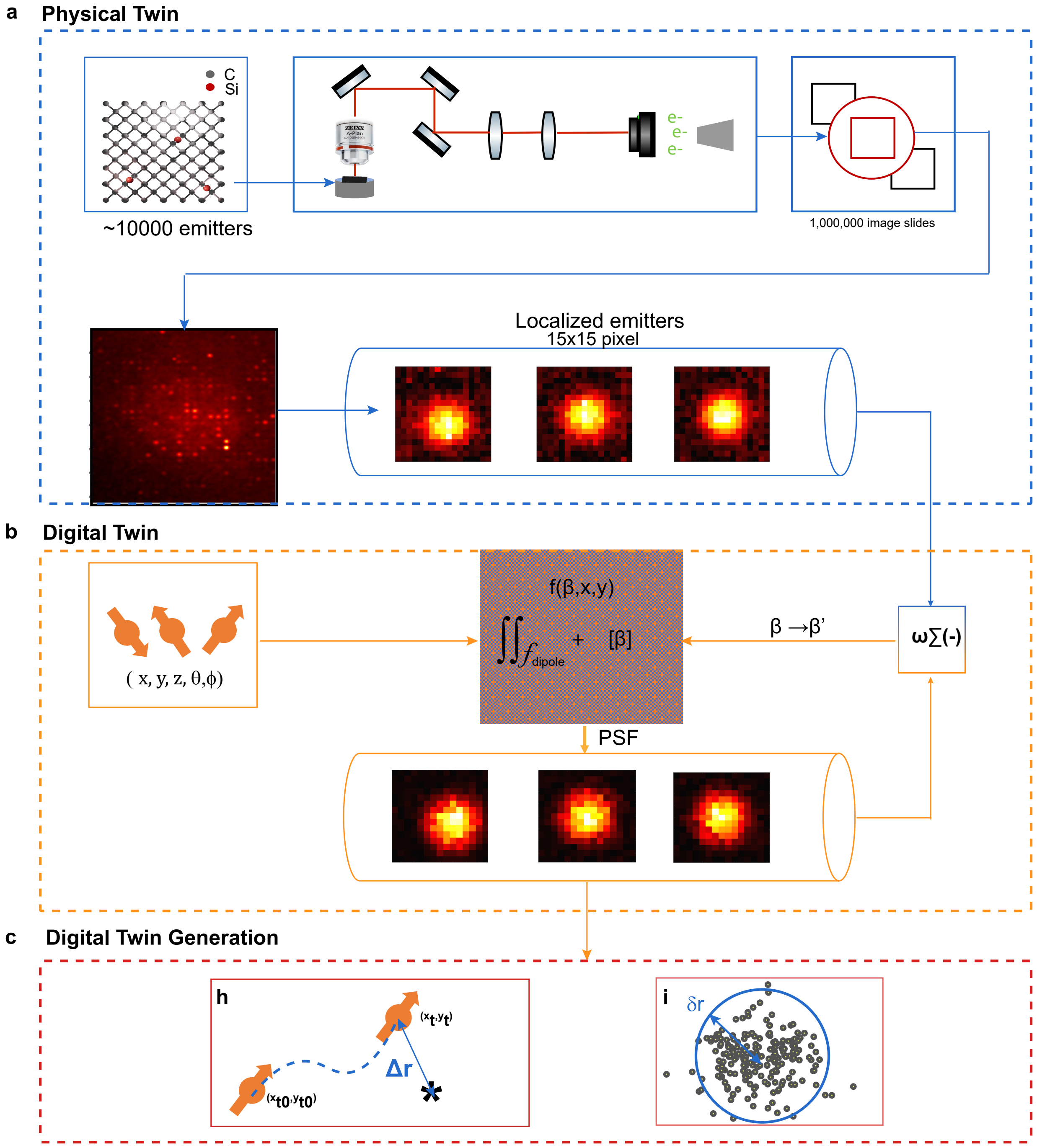


FIG. 15. Digital-physical twin system. **a**, Physical twin of SMLM setup. **b**, Digital twin based on the dipole radiation model, accurately reflects the physical performance generated from **a**. **c** Digital twin generation creates a virtual environment for examining localization experiments.

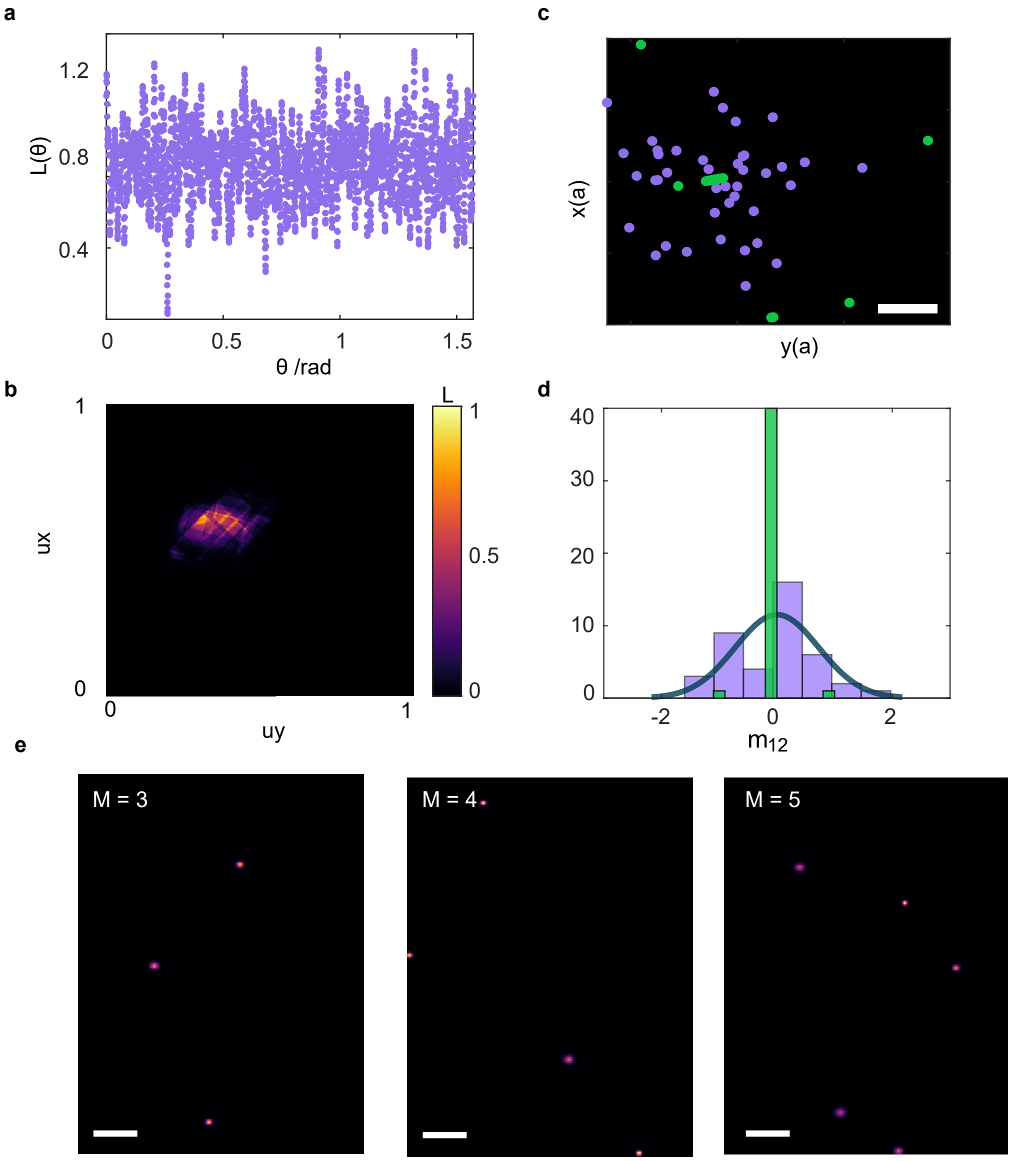


FIG. 16. DIGIT mapping emitters location on lattice points. **a**, Likelihood distribution of lattice spatial rotation θ. **b**, Likelihood map of lattice offset position. **c** Localization of Emitter 1 (x_1, y_1) by SMLM (Purple) and DIGIT (green). **d**, Histogram of localization uncertainty, directly showing the advantage of incorporating discrete lattice prior. **e**. Reconstructed localization for $M = 3$, 4, and 5.

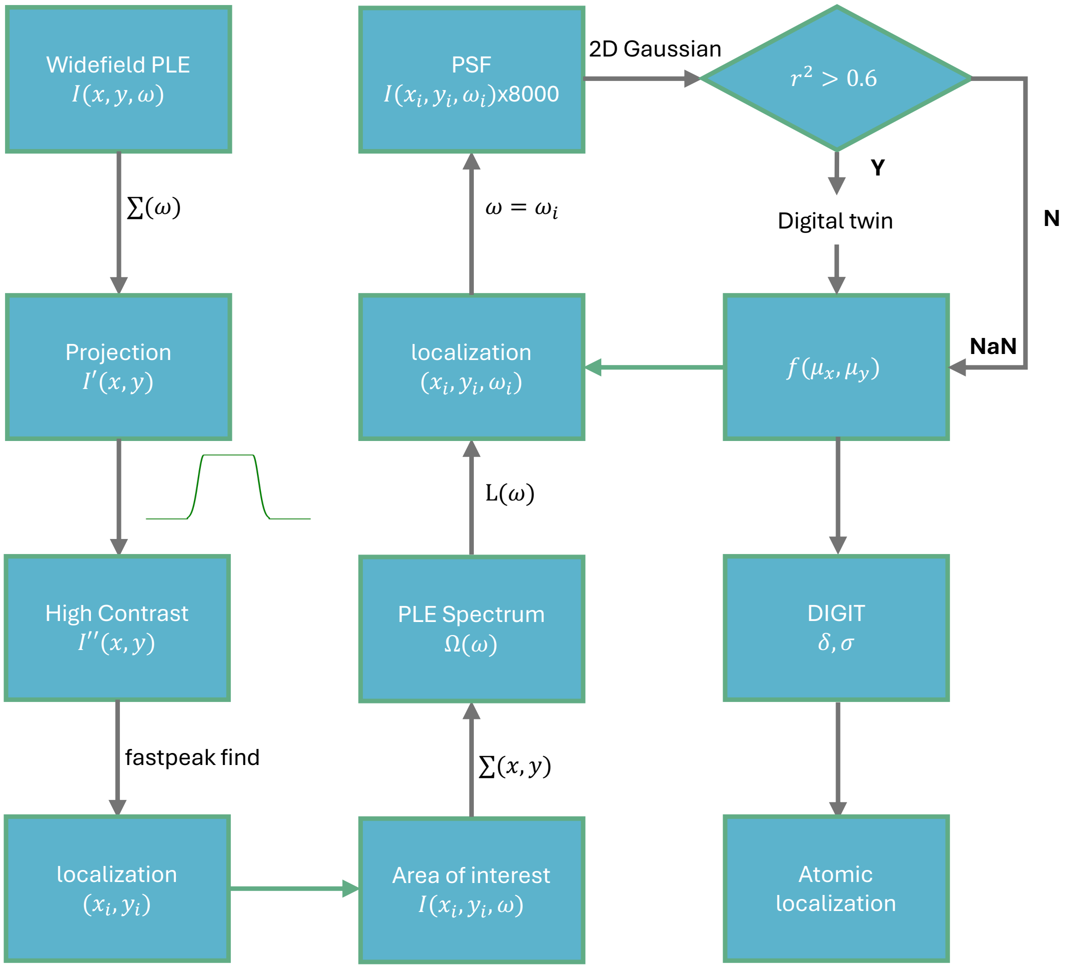


FIG. 17. Widefield DIGIT data analysis flowchart.