Supplementary Information for Hypothesis-Based Particle Detection for Accurate Nanoparticle Counting and Digital Diagnostics

Neil H. Kim, Xiao-Liu Chu, Joseph B. DeGrandchamp, Matthew R. Foreman

I. Peak Pixel Fraction under Gaussian PSF

For a Gaussian point spread function (PSF) of width σ, the fraction of total signal intensity captured by a single pixel depends on the relative alignment of the PSF center and the pixel grid. The pixel-integrated fraction for a pixel centered at the origin with the PSF centered at offset $(\Delta x, \Delta y)$ is

$$f(\sigma; \Delta x, \Delta y) = f_x(\sigma; \Delta x) f_y(\sigma; \Delta y) \tag{1}$$

where

$$f_s(\sigma; \Delta s) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-0.5}^{0.5} \exp\left[-\frac{(s-\Delta s)^2}{2\sigma^2}\right] ds. \tag{2}$$

For $\sigma = 2$ pixels and constraining the PSF centroid to lie within the domain of the pixel under consideration, numerical evaluation gives

$$\begin{aligned} f_{\min} &\approx 0.0366 \ (3.66\%), \\ f_{\text{avg}} &\approx 0.0378 \ (3.78\%), \\ f_{\max} &\approx 0.0390 \ (3.90\%). \end{aligned}$$

Assuming a particle scattering strength of $I = 20{,}000$, the corresponding expected pixel counts are

$$\begin{aligned} I f_{\min} &\approx 732, \\ I f_{\text{avg}} &\approx 756, \\ I f_{\max} &\approx 780. \end{aligned}$$

At the baseline background level $\lambda_{BG} = 2{,}000$, the peak-to-background ratios are hence

$$\begin{aligned} (I f_{\min})/\lambda_{BG} &\approx 0.366, \\ (I f_{\text{avg}})/\lambda_{BG} &\approx 0.378, \\ (I f_{\max})/\lambda_{BG} &\approx 0.390. \end{aligned}$$

II. Jacobian and Hessian

As described in the main text, we optimize the penalized negative log-likelihood

$$l_p(\boldsymbol{\theta}) = \sum_{i,j} \left[\lambda_{ij} - v_{ij} \log \lambda_{ij}\right] + \alpha \mathcal{P}(\boldsymbol{\theta}), \tag{3}$$

with

$$\lambda_{ij} = \lambda_{BG} + \sum_{k=1}^{N} I_k \, b_k(i) \, a_k(j), \tag{4}$$

The authors are with the Institute for Digital Molecular Analytics and Science, Nanyang Technological University, 59 Nanyang Drive, Singapore 636921, Singapore. M.R. Foreman is also with the School of Electrical and Electronic Engineering Department, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798. email: matthew.foreman@ntu.edu.sg

where $a_k(j) = f_x(\sigma; \Delta x_k^p)$ and $b_k(i) = f_y(\sigma; \Delta y_k^p)$ are the pixel-integrated 1D Gaussian PSF factors in x and y for the k-th particle with position $(\Delta x_k^p, \Delta y_k^p)$ measured from the center of the pixel. The corresponding Jacobian vector and Hessian matrices for the penalized negative log-likelihood are provided to the solver and are explicitly given below.

The Jacobian vector of l_p taken with respect to the parameter vector $\boldsymbol{\theta}$ is defined as

$$\mathbf{J} = \frac{\partial l_p}{\partial \boldsymbol{\theta}} = \left[\frac{\partial l_p}{\partial \lambda_{BG}} \ \ \frac{\partial l_p}{\partial I_1} \ \ \frac{\partial l_p}{\partial x_1^p} \ \ \frac{\partial l_p}{\partial y_1^p} \ \ \cdots \ \ \frac{\partial l_p}{\partial I_n} \ \ \frac{\partial l_p}{\partial x_n^p} \ \ \frac{\partial l_p}{\partial y_n^p}\right], \tag{5}$$

where (x_k^p, y_k^p) is the global position of the kth particle. The partial derivatives can be expressed in the form

$$\frac{\partial l_p}{\partial \lambda_{BG}} = \sum_{i,j} q_{ij}, \tag{6}$$

$$\frac{\partial l_p}{\partial I_k} = \sum_{i,j} q_{ij} \, b_k(i) \, a_k(j), \tag{7}$$

$$\frac{\partial l_p}{\partial x_k^p} = \sum_{i,j} q_{ij} \, I_k \, b_k(i) \, a_k'(j), \tag{8}$$

$$\frac{\partial l_p}{\partial y_k^p} = \sum_{i,j} q_{ij} \, I_k \, b_k'(i) \, a_k(j), \tag{9}$$

where for convenience we define

$$a_k'(j) = \frac{\partial f_x(j; \Delta x_k^p)}{\partial x_k^p}, \qquad b_k'(i) = \frac{\partial f_y(i; \Delta y_k^p)}{\partial y_k^p} \tag{10}$$

and

$$q_{ij} = 1 - \frac{v_{ij}}{\lambda_{ij}}, \qquad r_{ij} = \frac{v_{ij}}{\lambda_{ij}^2}. \tag{11}$$

Similarly, the Hessian matrix of l_p taken with respect to the parameter vector $\boldsymbol{\theta}$ is

$$\mathbb{H} = \frac{\partial^2 l_p}{\partial \boldsymbol{\theta}^2} = \begin{bmatrix} \frac{\partial^2 l_p}{\partial \lambda_{BG}^2} & \frac{\partial^2 l_p}{\partial \lambda_{BG} \partial I_1} & \frac{\partial^2 l_p}{\partial \lambda_{BG} \partial x_1^p} & \cdots \\ \frac{\partial^2 l_p}{\partial I_1 \partial \lambda_{BG}} & \frac{\partial^2 l_p}{\partial I_1^2} & \frac{\partial^2 l_p}{\partial I_1 \partial x_1^p} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{12}$$

for which representative nonzero blocks take the form

$$\frac{\partial^2 l_p}{\partial \lambda_{BG}^2} = \sum_{i,j} r_{ij}, \tag{13}$$

$$\frac{\partial^2 l_p}{\partial I_k^2} = \sum_{i,j} r_{ij} \left[b_k(i) a_k(j)\right]^2, \tag{14}$$

$$\frac{\partial^2 l_p}{\partial I_k \, \partial \lambda_{BG}} = \sum_{i,j} r_{ij} \, b_k(i) a_k(j), \tag{15}$$

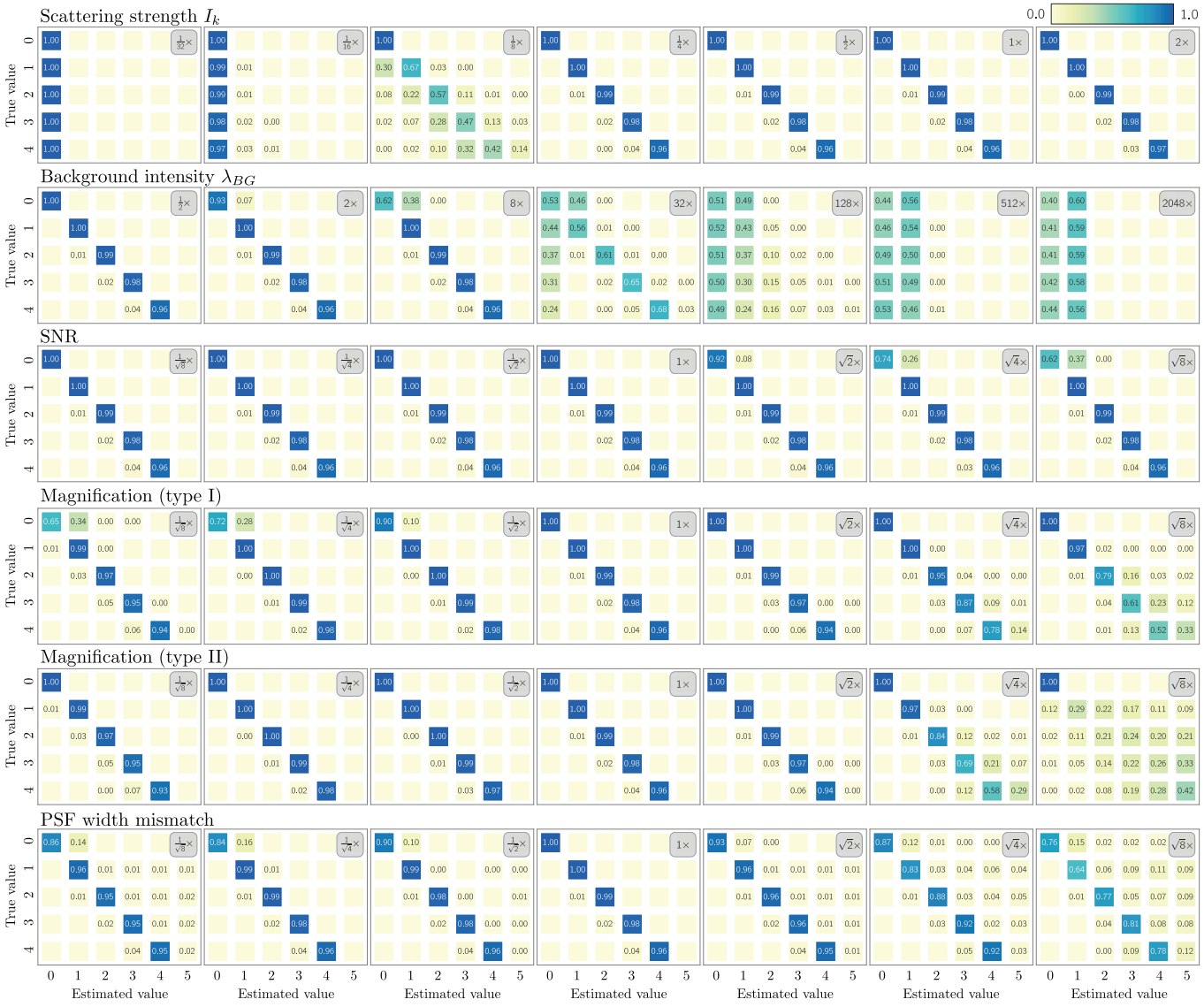

Fig. 1. **Simulated confusion matrices.** Representative confusion matrices compiled from the simulated experiments. Each confusion matrix lists true particle count (rows) versus estimated particle count (columns), aggregated over 10,000 images per condition.

$$\frac{\partial^2 l_p}{\partial x_k^{p2}} = \sum_{i,j}\Big[r_{ij}\,(I_k b_k(i) a'_k(j))^2 + q_{ij}\,I_k\,b_k(i)\,a''_k(j)\Big], \quad (16)$$

$$\frac{\partial^2 l_p}{\partial y_k^{p2}} = \sum_{i,j}\Big[r_{ij}\,(I_k b'_k(i) a_k(j))^2 + q_{ij}\,I_k\,b''_k(i)\,a_k(j)\Big], \quad (17)$$

$$\frac{\partial^2 l_p}{\partial y_k^{p}\,\partial x_k^{p}} = \sum_{i,j} r_{ij}\,I_k^2\,b'_k(i)\,a'_k(j)\,b_k(i)\,a_k(j). \quad (18)$$

where

$$a''_k(j) = \frac{\partial^2 f_x(j;\Delta x_k^p)}{\partial x_k^{p2}}, \qquad b''_k(i) = \frac{\partial^2 f_y(i;\Delta y_k^p)}{\partial y_k^{p2}}. \quad (19)$$

Other entries follow by symmetry. Notably, cross-particle blocks are zero since λ_{ij} is additive with respect to particle index.

The derivatives given above are expressed in terms of true physical parameters (position, scattering intensity etc). In our practical implementation of the algorithm, additional parameter normalization is used to aid convergence. This normalization introduces additional scale factors for the background, intensity, and particle positions. The penalty term contributes additional derivative terms $\alpha\nabla\mathcal{P}$ and $\alpha\nabla^2\mathcal{P}$ (which are trivial to evaluate) only when the particle positions fall outside the region of interest.

III. Confusion Tables

Confusion tables for the accuracy tests described in the main text are shown in Figure 1. As discussed in the main text at very high background levels (128× and above), the confusion matrices show that nearly all predictions fall into the $\hat{N} = 0$ and $\hat{N} = 1$ columns, with the relative ratio between 0 and 1 fluctuating across background levels. Predictions with $\hat{N} \geq 2$ are much less frequent, typically with relative frequency on the order of 10^{-3} at 2048× background. This indicates

that in the extreme high-background regime the estimator functions effectively as a binary classifier, toggling between background-only and single-particle fits. The rare $\hat{N} = 2$ estimates are consistent with the requirement for multiple PSF-like fluctuations within a single image, which occur with far lower probability than a single fluctuation. Shallow local minima in the likelihood surface, or implementation-specific regularization may also contribute to these rare events. Although very infrequent, these cases show that the estimator does not collapse exclusively to 0 or 1 predictions.

IV. Calculation of Overlap Probabilities

To estimate how often close particle pairs occur at $\bar{N} = 1$ in a 100×100 pixel image, we model particle locations as a homogeneous Poisson point process with rate $\rho = \bar{N}/A = 10^{-4}\,\mathrm{px}^{-2}$ where $A = 10^4\,\mathrm{px}^2$. The expected number of unordered pairs of particles separated by less than a radius r is

$$\mu_{\text{pairs}} \approx \frac{1}{2}\,\rho^2 A\,\pi r^2. \tag{20}$$

The probability of at least one such pair appearing in an image is then

$$p_{\text{image}} \approx 1 - e^{-\mu_{\text{pairs}}} \;\approx\; \mu_{\text{pairs}}, \tag{21}$$

where the second approximation is valid for the small values of μ_{pairs} considered here. Noting that overlap requires $N \geq 2$ particles, we normalize by $P(N \geq 2) = 1 - 2e^{-1} \approx 0.264$ to obtain the conditional probability

$$p_{\text{image}|N\geq 2} \approx \frac{p_{\text{image}}}{P(N \geq 2)}. \tag{22}$$

For the baseline case $\sigma = 2$ px, Figure 4 of the main text shows that two particles are essentially always resolved once their separation exceeds $r \approx 2.1\sigma$ (see also Supplementary Figure 2). With $r = 4.2$ px, this gives $\mu_{\text{pairs}} \approx p_{\text{image}} \approx 0.00277$, and

$$p_{\text{image}|N\geq 2} \approx \frac{0.00277}{0.264} \approx 0.0105 \quad (1.05\%).$$

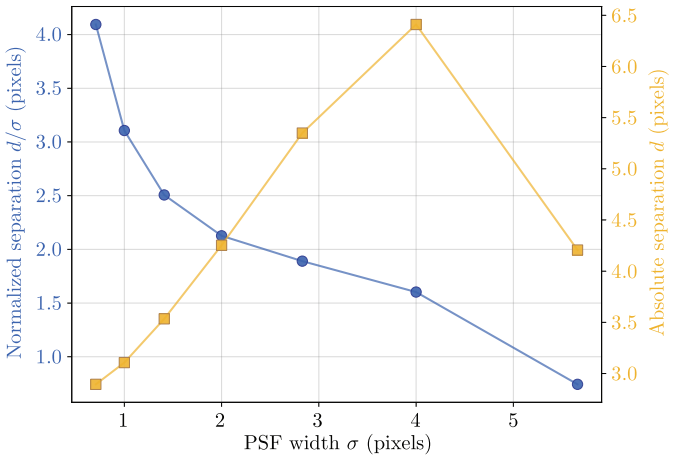

Fig. 2. **Resolvability threshold.** Variation of the normalized and unnormalized particle separations at the resolution limit (defined as the separation at which $\hat{N} = 2$ estimates fell to 50% of the total).

The unconditional value $p_{\text{image}} \approx 0.00277$ (0.277%) represents the overall chance of an overlap in a randomly drawn image, including those with $N = 0$ or 1 where overlaps are impossible. The conditional value $p_{\text{image}|N\geq 2} \approx 0.0105$ (1.05%) instead quantifies the probability of overlap given that at least two particles are present, and is the more relevant measure when assessing resolvability in multi-particle images.

V. DNA Functionalization of Gold and Silver Nanoparticles

Spherical gold nanoparticles (AuNPs) (diameter, 50 nm), spherical silver nanoparticles (AgNPs) (diameter, 50 nm), and gold-shell silica-core nanoparticles (AuSiNPs) (80 nm Si core, 20 nm Au shell) each with 40 kDa polyvinylpyrrolidone (PVP) as capping agent were purchased from nanoComposix (USA). NPs were functionalized with thiolated DNA probes (Table I) (Integrated DNA Technologies, USA) using a modified PVP-assisted conjugation protocol [1], [2]. Probe DNA sequences were chosen semi-empirically for particle stability, with initial designs created using NUPACK Web Application [3], [4] and with subsequent testing at high salt after conjugation. Thiolated DNA was reduced, for use with AgNPs only, by incubation with tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Sigma Aldrich) at a 1:100 DNA:TCEP ratio for 1 hr at room temperature. DNA for use with gold was not treated, as TCEP reduction was found to be unnecessary [5].

For conjugation, PVP-capped NPs were first aliquoted from high concentration stocks (AuNP: 27.9 OD @ 528 nm, *est.* 1.2 nM; AgNP: 134.9 OD @ 419 nm, *est.* 2.7 nM; and Au@SiNP: 37.8 OD @ 666 nm, *est.* 0.15 nM) into DNA Lobind tubes (Eppendorf SE, Germany). The respective thiolated probes were added at a final concentration of 28.5 µM followed by bath sonication of the reaction (10 s) [6] and incubation (10 min) at room temperature. Additional PVP (40 kDa) (Sigma Aldrich) was added to a final concentration of 0.35% (w/v) to further stabilize the particles. A volume of salt-containing buffer, consisting of 20 mM sodium phosphate buffer (PB, pH 8), 200 mM NaBr, and 0. 02% (w/v) sodium dodecyl sulfate (SDS) (Sigma Aldrich), equivalent to the reaction volume was slowly added in a single step. Notably, NaBr was used instead of NaCl to discourage nonspecific adhesion of nucleotides to the gold/silver surface [7]. The reaction mixture was then heated at 60 °C for 2.5 hr with mixing (800 rpm). To remove residual DNA and salts, the resulting conjugated NPs were washed 5 times (or 7 times for viral detection) by centrifugation (18000*g*, 5 min) using 10 mM PB (no NaCl). Conjugated NP concentrations were then approximated using their UV-Vis Absorbance spectra (NanoDrop One Instrument, Thermofisher Scientific, USA) as compared to their reported stock concentrations.

A 188 bp single-stranded DNA sequence ("E-DNA") derived from the Envelope Gene and 200 bp DNA sequence ("N-DNA") derived from the Nucleocapsid Gene of SARS-CoV-2 (NCBI `NC_045512.2`) were designed for use as model DNA targets (see Table I) (Ultramer ssDNA, PAGE purified; Integrated DNA Technologies, USA). We used these two strands at equivalent concentrations to emulate a larger SARS-

TABLE I
ssDNA SEQUENCES FOR TARGET-DNA AND PROBES

ssDNA Strand ID	Sequence (5' to 3')
E-DNA (from SARS-CoV-2 *E* Gene; NCBI NC_045512.2)	agagacaggtacgttaatagttaatagcgtacttcttttttcttgctttcgtggtattctt gctagttacactagccatccttactgcgcttcgattgtgtgcgtactgctgcaatattgtt aacgtgagtcttgtaaaaccttcttttttacgtttactctcgtgttaaaaatctgaattctt ctagag
N-DNA (from SARS-CoV-2 *N* Gene; NCBI NC_045512.2)	actcaacatggcaaggaagaccttaaattccctcgaggacaaggcgttccaattaacacc aatagcagtccagatgaccaaattggctactaccgaagagctaccagacgaattcgtggtg gtgacggtaaaatgaaagatctcagtccaagatggtatttctactacctaggaactgggcc agaagctggacttcccta
PrE1	SH-TTTTTTTTTT-aactattaacgtacctgtctct
PrE2	SH-TTTTTTTTTT-aggatggctagtgtaactag
PrE3	SH-TTTTTTTTTT-ctctagaagaattcagattt
PrN1	SH-TTTTTTTTTT-aatttaaggtcttccttgcc
PrN2	SH-TTTTTTTTTT-gtagccaatttggtcatctg
PrN3	SH-TTTTTTTTTT-tagggaagtccagcttctgg

TABLE II
SUMMARY OF FITTED PARAMETERS FOR NULL COUNT, STANDARD POISSON, AND GPD MODELS. '—' INDICATES A PARAMETER IS NOT APPLICABLE TO THAT MODEL. FOR THE 13.2 PM CASE RESCALED $\bar{N}$ ARE SHOWN IN BRACKETS.

Concentration (pM)	Method	Image size (pixels)	$\bar{N}_{\text{control}}$	ψ_{control}	R^2_{control}	$\bar{N}_{\text{covid}}$	ψ_{covid}	R^2_{covid}
3.3	Null count	50×50	0.335	—	0.989	0.506	—	0.986
6.6	Null count	50×50	0.772	—	0.923	1.185	—	0.775
13.2	Null count	35×35 (50×50)	1.073 (2.191)	—	0.442	0.900 (1.837)	—	0.617
3.3	Poisson	50×50	0.304	—	0.991	0.478	—	0.986
6.6	Poisson	50×50	0.742	—	0.920	1.229	—	0.787
13.2	Poisson	35×35 (50×50)	1.092 (2.228)	—	0.495	0.858 (1.751)	—	0.671
3.3	GPD	50×50	0.334	0.270	1.000	0.501	0.193	0.999
6.6	GPD	50×50	0.770	0.295	0.998	1.184	0.264	0.998
13.2	GPD	35×35 (50×50)	1.087 (2.218)	0.432	0.985	0.909 (1.854)	0.489	0.993

CoV-2 target and refer to them collectively as the "Target-DNA." For Target-DNA detection, particles were functionalized with two probes each (PrE1 and PrN1 for AuSiNPs; PrE2 and PrE3 for AgNPs; PrN2 and PrN3 for AuNPs), which all bind to specific Target-DNA sites and enable target-induced crosslinking. The exact sequences are found in Table I. For the assay, the nanoparticle mixture (in PB), a reaction buffer mixture, and Target-DNA (in TE buffer) at the relevant concentrations were prepared. The three separate solutions were mixed to start the hybridization assay. The reaction was set up such that the following final concentrations were achieved in a 40 µL volume: 10 mM PB, 400 mM NaCl, 2 mM Ethylenediaminetetraacetic acid (EDTA) (Sigma Aldrich), 3% (w/v) dextran (150 kDa) (Sigma Aldrich), 0.05% (w/v) PVP (40 kDa), 0.05% (w/v) SDS, 10 pM Target-DNA ([E-DNA] = [N-DNA] = reported concentration) or TE buffer as control. EDTA was included to chelate undesirable multivalent ions, while dextran was included to promote specific hybridization [8], [9]. Salt and surfactant concentrations were empirically derived for optimal hybridization results (fast and specific). The nanoparticle concentrations were varied as follows:

1) Low: AuSiNPs: 0.3 pM, AuNPs and AgNPs: 1.5 pM each.
2) Medium: AuSiNPs: 0.6 pM, AuNPs and AgNPs: 3 pM each.
3) High: AuSiNPs: 1.2 pM, AuNPs and AgNPs: 6 pM each.

Immediately after mixing, the solution was vortexed and then heated to 50 °C with mixing (800 rpm) for 15 min. The reaction was allowed to cool for 10 min at room temperature before transferring 20 µL to a home-made microscopy chamber consisting of two plasma-etched coverslips sandwiching a parafilm chamber for optical imaging. Notably, single AuSiNPs settle via gravity onto the glass surface, while other particles tend to stay diffuse in solution unless they are clustered. Images were captured in a homebuilt darkfield microscope.

VI. GENERALIZED POISSON DISTRIBUTION FITTING

Table II details the fitted parameters ($\bar{N}$ and ψ where appropriate) and the resulting R^2, for the null count, Poisson, and GPD models across the tested concentration range. For the highest concentration (13.2 pM), as detailed in the main text, the final analysis was performed on reduced 35×35 pixel sub-images. This was motivated by the results in shown in Figure 3, which used a 50×50 sub-image size, in which significant counts at the maximum hypothesis number $n_{\max} = 5$ were seen. Table II includes the raw extracted $\bar{N}$ values and scaled values (in brackets) for this case, so as to facilitate direct comparison.

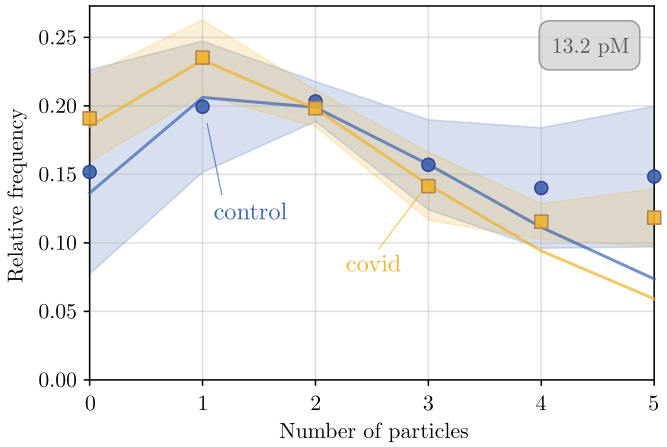

Fig. 3. **Nanoparticle imaging based SARS-CoV-2 assay - high nanoparticle concentration case.** Distribution of estimated particle counts for the high nanoparticle concentration case when 50×50 sub-images were used, for both control (blue markers) and SARS-CoV-2 positive (orange markers) samples, overlaid with GPD fits (solid curves). Error bands correspond to inter-image standard deviations. Significant counts at $\hat{N} = 5$ are evident motivating reduction of the sub-image size.

REFERENCES

[1] J. H. Heo, K.-I. Kim, H. H. Cho, J. W. Lee, B. S. Lee, S. Yoon, K. J. Park, S. Lee, J. Kim, D. Whang, and J. H. Lee, "Ultrastable-stealth large gold nanoparticles with DNA directed biological functionality," *Langmuir*, vol. 31, pp. 13773–13782, 2015. [Online]. Available: https://doi.org/10.1021/acs.langmuir.5b03534

[2] S. Xu, H. Yuan, A. Xu, J. Wang, and L. Wu, "Rapid synthesis of stable and functional conjugates of DNA/gold nanoparticles mediated by tween 80," *Langmuir*, vol. 27, pp. 13629–13634, 2011. [Online]. Available: https://doi.org/10.1021/la203632c

[3] J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks, and N. A. Pierce, "NUPACK: Analysis and design of nucleic acid systems," *Journal of Computational Chemistry*, vol. 32, pp. 170–173, 2011. [Online]. Available: https://doi.org/10.1002/jcc.21596

[4] M. E. Fornace, J. Huang, C. T. Newman, N. J. Porubsky, M. B. Pierce, and N. A. Pierce, "NUPACK: Analysis and Design of Nucleic Acid Structures, Devices, and Systems," *ChemRxiv*, Version 1, 2022, preprint. [Online]. Available: https://doi.org/10.26434/chemrxiv-2022-xv98l

[5] R. Wu, L.-P. Jiang, J.-J. Zhu, and J. Liu, "Effects of Small Molecules on DNA Adsorption by Gold Nanoparticles and a Case Study of Tris(2-carboxyethyl)phosphine (TCEP)," *Langmuir*, vol. 35, pp. 13461–13468, 2019. [Online]. Available: https://doi.org/10.1021/acs.langmuir.9b02652

[6] S. J. Hurst, A. K. R. Lytton-Jean, and C. A. Mirkin, "Maximizing DNA Loading on a Range of Gold Nanoparticle Sizes," *Analytic Chemistry*, vol. 78, pp. 8313–8318, 2006. [Online]. Available: https://doi.org/10.1021/ac0613582

[7] B. Liu, P. Wu, Z. Huang, L. Ma, and J. Liu, "Bromide as a Robust Backfiller on Gold for Precise Control of DNA Conformation and High Stability of Spherical Nucleic Acids," *Journal of the American Chemical Society*, vol. 140, pp. 4499–4502, 2018. [Online]. Available: https://doi.org/10.1021/jacs.8b01510

[8] R. M. Amasino, "Acceleration of nucleic acid hybridization rate by polyethylene glycol," *Analytical Biochemistry*, vol. 152, pp. 304–307, 1986. [Online]. Available: https://doi.org/10.1016/0003-2697(86)90413-6

[9] T. Donnelly, K. Faulds, and D. Graham, "Investigation of Silver Nanoparticle Assembly Following Hybridization with Different Lengths of DNA," *Particle and Particle System Characterization*, vol. 33, pp. 404–411, 2016. [Online]. Available: https://doi.org/10.1002/ppsc.201500238