Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Materials: Non-uniform sampling in pulse dipolar spectroscopy by EPR: the redistribution of noise and the optimization of data acquisition

Anna G. Matveeva[a,b], Victoria N. Syryamina[c], Vyacheslav M. Nekrasov[b,c], Michael K. Bowman*[d,e]

[a.] *Institute of Solid State Chemistry and Mechanochemistry of the Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia.*
[b.] *Novosibirsk State University, 630090 Novosibirsk, Russia.*
[c.] *Voevodsky Institute of Chemical Kinetics and Combustion Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia.*
[d.] *N. N. Vorozhtsov Novosibirsk Institute of Organic Chemistry of the Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia.*
[e.] *Department of Chemistry & Biochemistry, The University of Alabama, Tuscaloosa, AL 35487, USA.*

Contents:

S.1 Propagation of random noise

A semilog plot, Figure S1, of the error propagation function $g(r)$ from Fig. 1B, C shows that random noise causes the largest errors in the recovered distance distribution at the shortest distances and steadily falls off throughout the window of valid distances and beyond. The curves obtained by the DeerAnalysis and MeTA approaches are consistent; the curves keep crossing each other. The rapid fluctuations in the two curves are caused by the limited number of nine realizations of random noise used to construct the curves. These minor fluctuations are also seen in Figure 6 between the black and white lines indicating the area that is essentially free of sampling and truncation noise. Figure S1 shows that Tikhonov regularization and MeTA are equally effective at recovering the distance distribution from the dipolar trace and are equally affected by random noise.

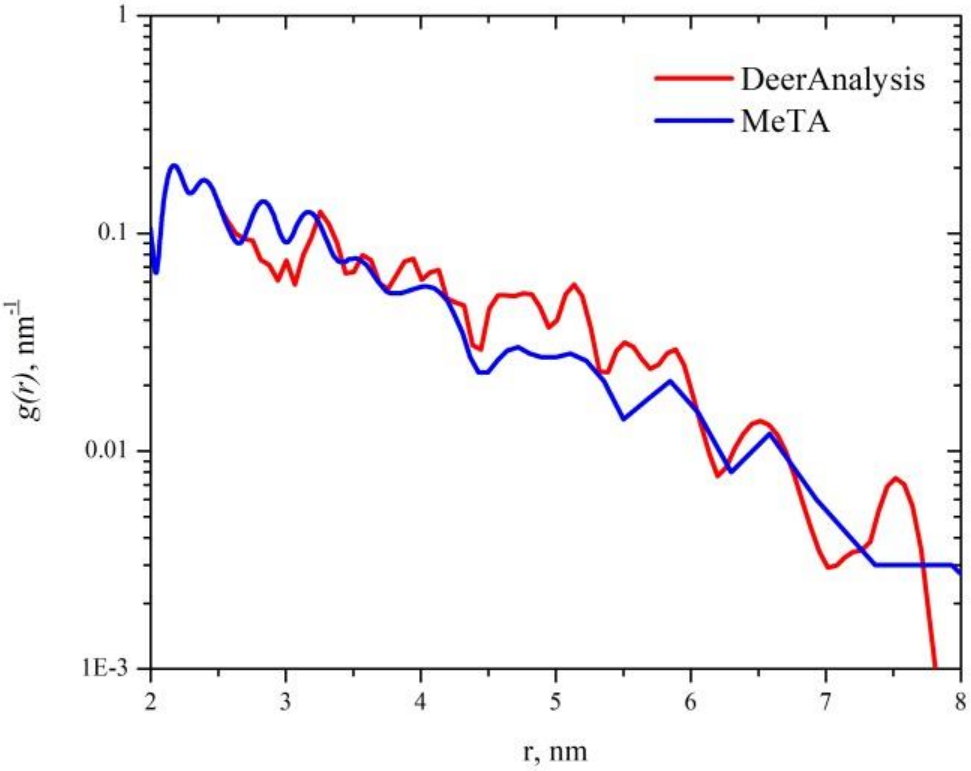

Figure S1 Semilogarithmic plot comparing the error propagation functions $g(r)$ from Figure 1B, C for random noise.

S.2 Scaling of the dipolar function

According to eq. (1), the kernel in the Fredholm equation scales very simply: a change of $r \longrightarrow n\,r$ results in $T \longrightarrow n^3\,T$, which simultaneously results in increasing the $P(r)$ width n times. At the same time, the dipolar trace V(n^3T) keeps its shape at the appropriate scaling time axis, Figure S2A, B.

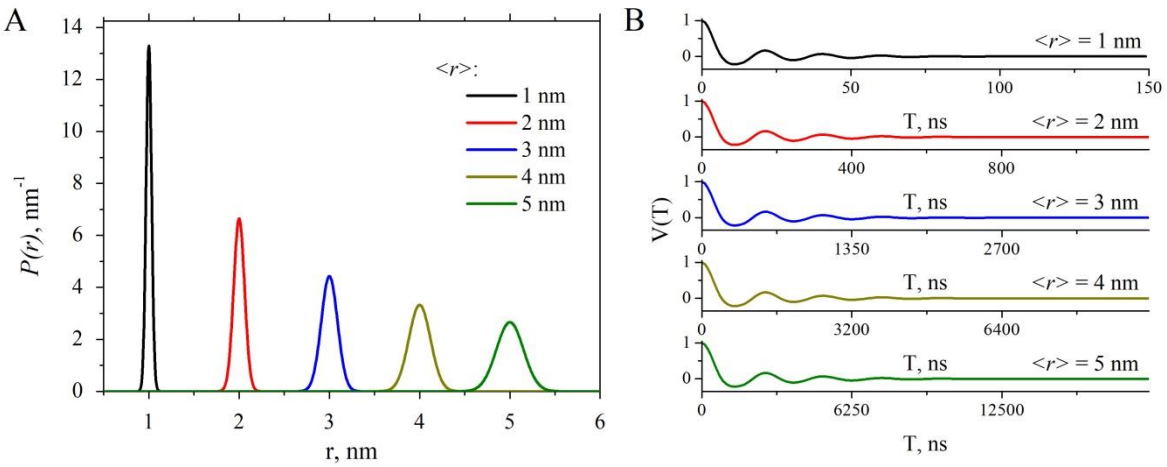

Figure S2 **A**) The mono-modal distance distribution function $P(r)$ with $\delta r/\langle r\rangle = 0.1$ **B**) The dipolar traces for the $P(r)$ using the same color code. The time axis was scaled by n^3 to demonstrate that the shape of V(T) remains constant with $r \longrightarrow n\,r$ scaling.

S.3 Propagation of random noise: the impact of noise amplitude, length of the dipolar trace, and sampling step

Four parameters affect the shape of $g(r)$: the total experimental time T_{all}, the length of the dipolar trace T_{trace}, the noise level σ, and the sampling step dT. We independently tested the impact of the noise amplitude σ and the sampling step dT on $g(r)$ using MeTA.

- dT = *const* = 32 ns, T_{trace} = *const*, with σ varied (T_{all} varies automatically), Figure S3A. The amplitude of $g(r)$ grows linearly, Figure S3B, while its shape does not change.

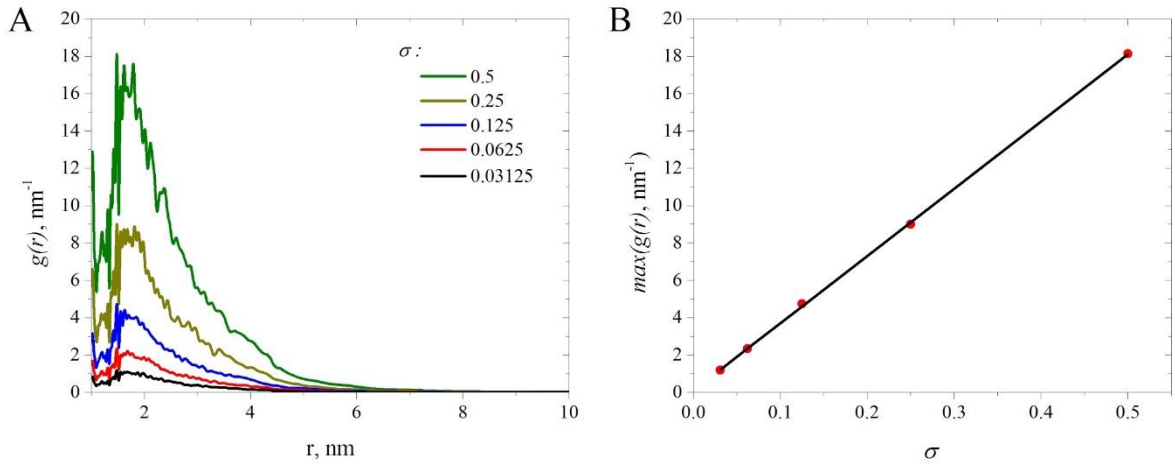

Figure S3 **A**) The error propagation function $g(r)$ for different random noise amplitudes σ at a fixed sampling step dT = 32 ns. **B**) The maximum of the $g(r)$ function at different noise amplitudes σ (circles) and the linear regression line.

- σ = *const* = 0.0625, T_{trace} = *const*, with *dT* varied (T_{all} varies automatically), Figure S4A. As *dT* increases, the amplitude of $g(r)$ decreases, Figure S4B, and the maximum of $g(r)$ shifts to longer distances, Figure S4C. The shift obeys the power law $r\left(max\left(g(r)\right)\right) \sim dT^{1/3}$ and corresponds to the minimum distance r_{min} that can be recovered at a given sampling step *dT*.

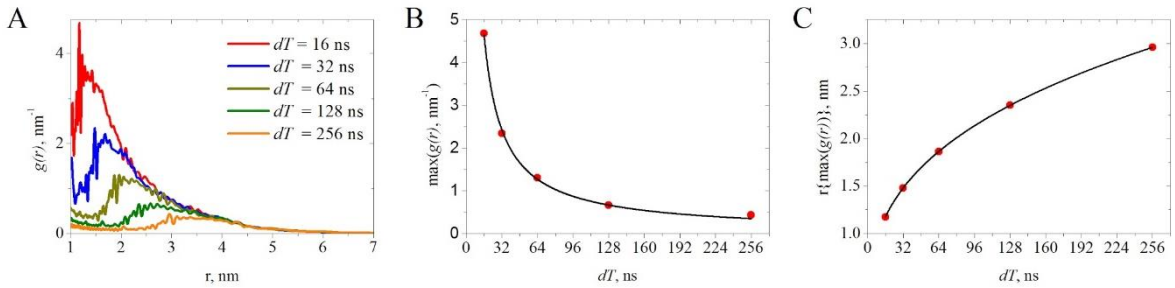

Figure S4 **A**) The error propagation function $g(r)$ at fixed noise amplitude σ for different sampling step *dT*. **B**) The maximum of $g(r)$ at different sampling steps (circles); and **C**) The distance of the maximum of $g(r)$ at different sampling steps (circles) and the regression line for $dT^{1/3}$ (line).

- σ = *const* = 0.0625, *dT* = *const* = 64 ns, T_{trace} varied (T_{all} varies automatically), Figure S5. The error propagation function is almost independent of T_{trace}, except at distances > 3-5 nm that lie beyond the window of valid distances for the shorter traces (0.756 - 1.26 μs). The decrease is not unexpected. Truncation of the signal results in information loss about the dipolar trace at long distances and in redistribution or distortion of the distance distribution. But, noise at long times carries information about the frequency spectrum of the noise and how it is distributed in $g(r)$ function. We estimated a point r_{cr} on each g(r), where the slope changes. Figure S5C shows a strong linear dependence of r_{cr} on T_{trace}.

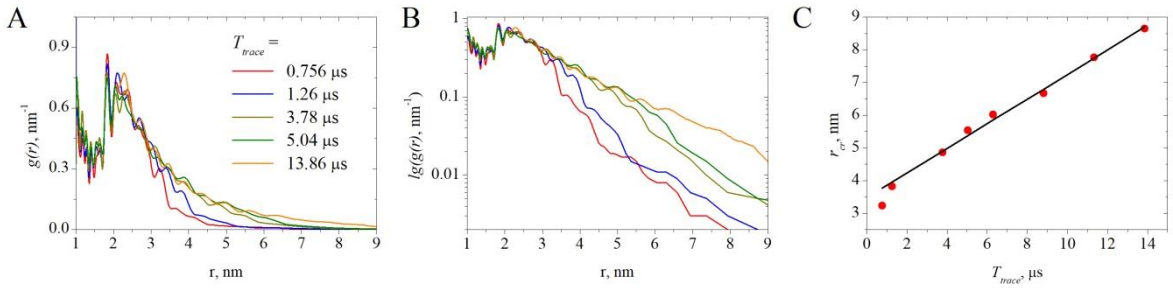

Figure S5 The random error function $g(r)$ at fixed noise amplitude σ and sampling step *dT* and different length of the dipolar trace T_{trace}.

S.4 Scaling of the distortions of *f(r)* for different truncation lengths

Here we continue examining the propagation of truncation noise into errors in the recovered $f(r)$ function. We use a unimodal distribution function with <r> = 3 nm and vary $\delta r/r$ from 0.01 to 0.1 for several different T_{trace}= n/ω_D. Uncertainties in the position r_{max} of the $f(r)$ maximum and its amplitude $f(r_{max})$, normalized to its maximal value at different truncation times, are shown in Figure S6A, B. The inaccuracy in determining <r> of the peak of the distribution is almost independent of the $\delta r/\langle r\rangle$ value and less than 3% when $T_{trace} \geq 2\ T_D$. However, the rate at which the width convergence to its "true" value does depend on $\delta r/\langle r\rangle$, Figure S6B. The narrower the peak, the larger T_{trace} should be. The family of curves in Fig. S6 B scales with a coefficient k for the abscissa, Figure S6C. That scaling factor depends linearly on $\delta r/\langle r\rangle$, Figure S6D, and k enables us to predict the length of the dipolar trace required to recover $f(r)$ free from truncation noise.

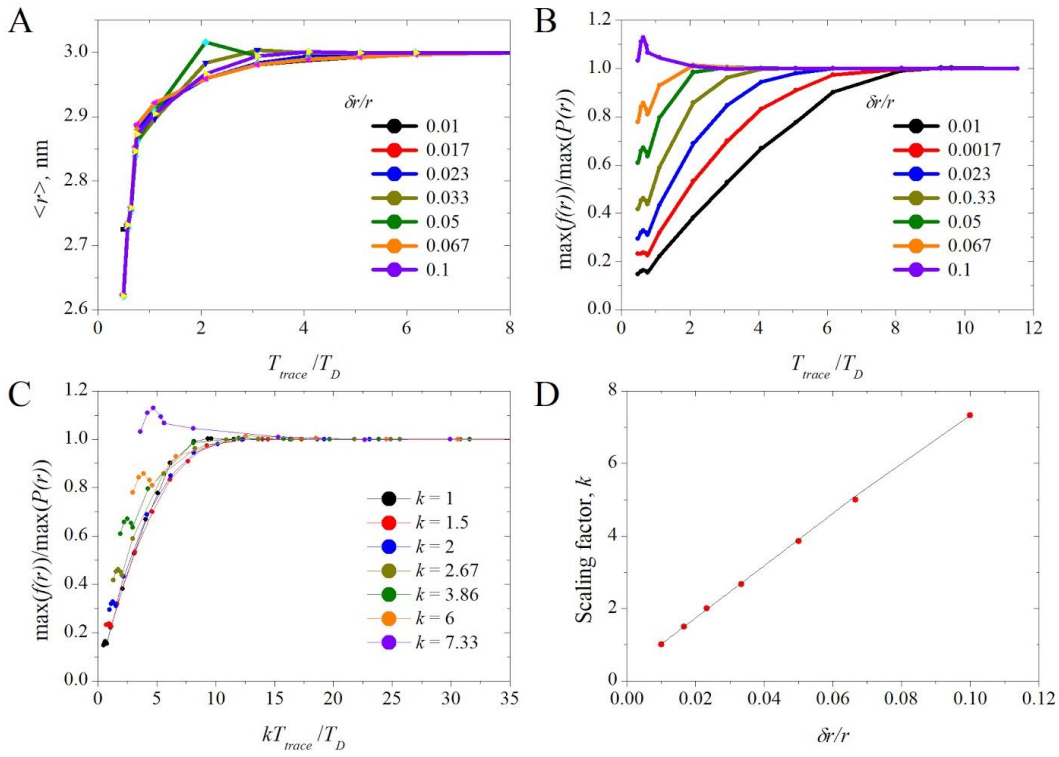

Figure S6 **A**) Position r_{max} of the $f(r)$ maximum for different truncation times. **B**) Amplitude of $f(r_{max})$, normalized to its undistorted value, for different truncation times. **C**) Amplitude $f(r_{max})$, normalized to its undistorted value, with the x-axis scaled by a factor *k*. **D**) The scaling factor *k* for different values of $\delta r/\langle r\rangle$ (points) and the linear regression fit (line).

S.5 Optimization of sampling noise in the *eNUS* scheme

We examined different sampling schemes with T_{trace} = 5.04 μs by varying *NPT* at fixed b = 0.2 (a and dT_1 vary automatically). For each *NPT*, Figure S7, we calculate the mean square error (*mse*) between $f(r)$ and $P(r)$ for different $\langle r \rangle$ (at fixed $\delta r/\langle r \rangle$. The optimal sampling scheme at each $\langle r \rangle$ corresponds to $dT_1 \approx dT_{opt} = \pi \langle r \rangle^3$, at least near b = 0.2.

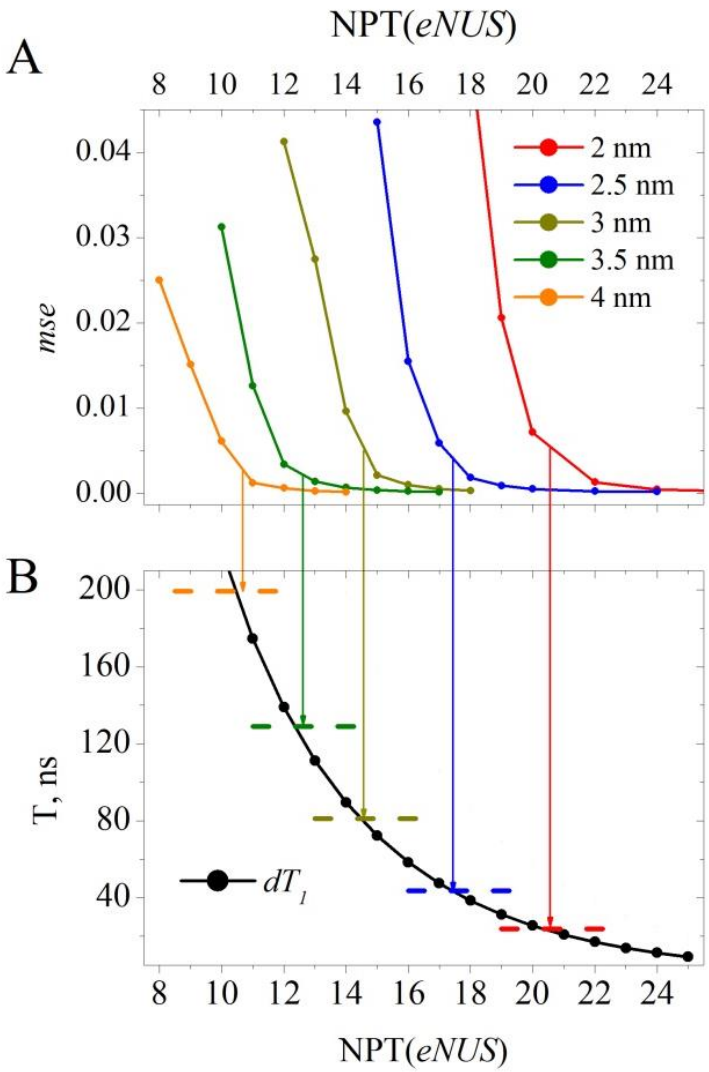

Figure S7 Optimization of sampling noise for the *eNUS*. **A**) The mean square error (*mse*) between $f(r)$ and $P(r)$ for different $\langle r \rangle$ and number of points *NPT* with T_{trace} = 5.04 μs. **B**) Lengths of the initial sampling steps dT_1 for each *NPT*. The horizontal lines indicate the optimal sampling step dT_{opt} for each $\langle r \rangle$, with the same color code as in **A**). The vertical arrows reveal the coincidence between dT_1 and dT_{opt}.

S.6 Calculation of noise redistribution in the *NUA* scheme

The *NUA* experiments emphasize measurements at the beginning of the dipolar trace. Shots per point *SPP* depend on the position *i* of each experimental point in the dipolar trace as $SPP_i \propto q^i$. Keeping the total measurement time equal for two experiments means that the total number of shots are equal: if the *US* measurement requires *n* scans, then the *NUA* starting with the same SPP_0 requires n_x:

$$\begin{aligned} NT \times n &= \sum_{i=0}^{NT-1} q^i n_x \\ &= \frac{1-q^{NT-1}}{1-q} \times n_x \\ \Rightarrow n_x &= \frac{NT \times n(1-q)}{1-q^{NT-1}} \end{aligned}$$

which means that the number of shots at the *i*-th point of *NUA* is

$$n_i = \frac{NT \times n(1-q)q^i}{1-q^{NT-1}}.$$

Further, if the noise in the US is σ_0, then noise in *NUA* depends on *i*. Knowing that

$$\begin{aligned} \sigma_0 &= \frac{const}{\sqrt{n}} \\ \sigma_i &= \frac{const}{\sqrt{n_i}} \end{aligned}$$

giving

$$\begin{aligned} \sigma_i &= \frac{\sigma_0\sqrt{n}}{\sqrt{n_i}} \\ &= \sigma_0 \sqrt{\frac{(1-q^{NT-1})}{NT(1-q)q^i}} \end{aligned} \tag{S1}$$

S.7 Calculation of noise redistribution for *US*, *eNUS* and *NUA*

The 2D plot in Figure 6 was built from calculations on a grid, which were then interpolated onto a finer grid for a smoother plot. For *US*, *dT* varied from 16 to 368 ns with a 32 ns time step, for *eNUS*, *NPT* = 158, 105, 60, 50, 40, 30, 28, 26, 24, 22, 20, 18, 16, 15, 14, and for *NUA* the *q*-value was varied between 1 and 0.9 with a 0.01 step.

S.8 Comparison of *eNUS* with *US*

To make it easier to compare the relative performance of *eNUS* and *US* with random noise, slices with the same 'open window' were extracted from Figure 6A and B. Their ratios were taken and used to construct Figure S8 which compares their normalized relative performance with random noise for the same 'open window'. Blue shows where *eNUS* has better performance and Red where *US* is better. However, as Figure 6 shows, random noise is much more intense for both sampling schemes at shorter distances than it is at longer distances.

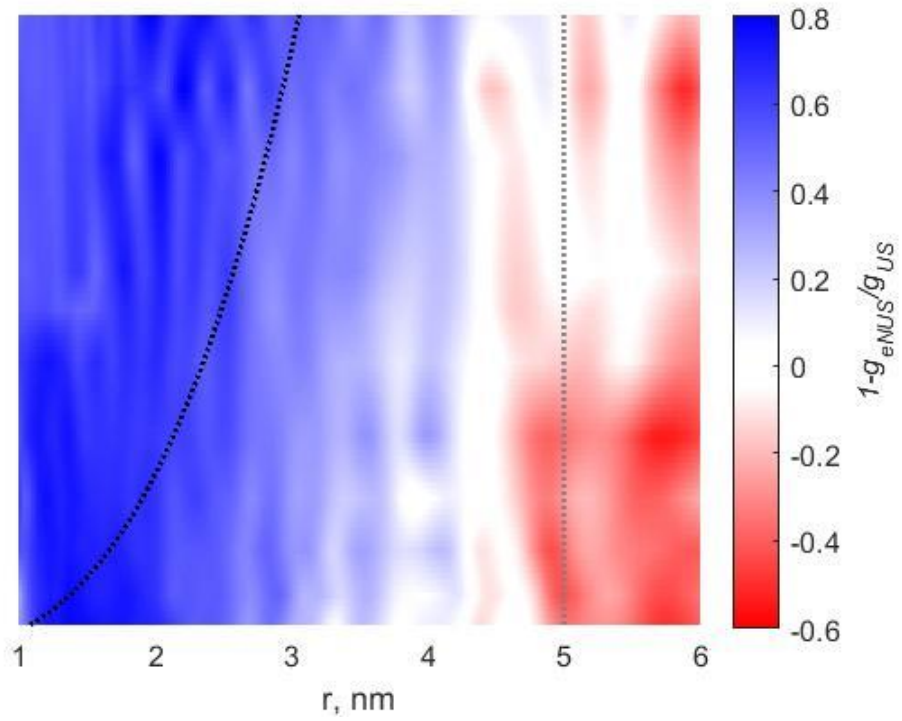

Figure S8 Normalized comparison of the relative performance of eNUS versus US for the same 'open window' indicated by the dotted lines. Calculated from data in the data plotted in Figure 6A and B. The top slice corresponds to the Black and Red curves in Figure 7.