

# Journal of Materials Chemistry C

Materials for optical, magnetic and electronic devices

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Stockdale, J. Bible, E. Mattingly and O. T. Mefford, *J. Mater. Chem. C*, 2026, DOI: 10.1039/D6TC00751A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Cite this: DOI: 00.0000/xxxxxxxxxx

# Accounting for Non-Stationary Variability in Magnetic Particle Spectroscopy<sup>†</sup>

Arabella Stockdale,<sup>\*a</sup> Joe Bible,<sup>b‡</sup> Eli Mattingly,<sup>c</sup> and Olin Thompson Mefford<sup>a</sup>Received Date  
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Magnetic particle spectroscopy (MPS) is an emerging analytical technique with potential for rapid, accessible, and affordable point of care biomarker detection based on magnetic signal fluctuations. Accurate interpretation of MPS signals requires both colloidal stability of nanoparticles in aqueous and complex media and statistical models that appropriately represent experimental variability. Conventional least squares (LS) approaches assume stationary and homoscedastic noise, assumptions which break down in experimental settings due to inherent temporal and within sample variability. This study investigates the signal variability in 22 nm iron oxide nanoparticles (IONPs) coated with 20 kDa poly(ethylene glycol). We evaluate the validity of the standard linear model and reveal that the least squares approach fails to account for inherent temporal and within sample variability. Repeated measurements under identical treatment conditions reveal substantial structured fluctuations consistent with a non-stationary measurement process. Transformations commonly used to normalize mass loading effects, including harmonic ratios, redistribute variance without resolving the underlying variability. A hierarchical mixed effects model is implemented to partition within-run and temporal sources of variation and to provide confidence estimates that reflect the observed variance structure across observation levels. These findings establish a statistical framework for representing intrinsic variability in MPS measurements and support more rigorous interpretation of harmonic signals. Explicit modeling of hierarchical uncertainty is an important step towards quantitative, reproducible MPS analysis and translation of MPS for sensitive biomarker detection.

Characterizing the magnetization and relaxation dynamics of iron oxide nanoparticles (IONPs) ensembles remains a central challenge in developing magnetic nanoparticles for various biomedical applications such as magnetic resonance imaging (MRI),<sup>1,2</sup> magnetic particle imaging (MPI),<sup>3-5</sup> and liquid phase biomarker detection.<sup>6-9</sup> Magnetic particle spectroscopy (MPS) has emerged as a rapid and accessible characterization technique to probe dynamic magnetic tracers through driving IONPs into their nonlinear regime under and alternating magnetic field.<sup>10</sup> The resulting harmonic spectra encode specific properties such as the viscosity of the solution, binding apparatus of chemical and biological analytes, and hydrodynamic diameter.<sup>8</sup> However, the complex interplay between aggregation, interparticle interactions, and colloidal stability, often introduces signal variability that is not captured by simplified analytic models. Importantly, this variability is not purely Gaussian noise and thus reflects a temporally evolving system that challenges commonly used statistical meth-

ods. Although often overlooked for simplicity, accurately modeling this variation is a necessary component to the clinical translation of this technology. Without acknowledging these complications, false positives or unreproducible data can occur.

Signal quality in MPS is often reported using metrics such as the signal-to-noise ratio (SNR) or harmonic decay ratio (e.g.,  $A_5/A_3$ ) with uncertainty typically estimated using the standard error of the mean (see Eq. 2).<sup>11-13</sup> The standard error asymptotically shrinks with increasing measurement acquisition making the estimates of uncertainty artificially small and result in unrealistically narrow confidence intervals. These approaches implicitly assume stationary means and statistically equivalent observations, neglecting between-sample variation that may overstate reproducibility. In practice, MPS datasets frequently include multiple aliquots, repeated runs, and day to day measurements making the definition of an independent replicate nontrivial. Treating these as independent observations collapses distinct sources of variability into a single estimate and overstates reproducibility. As a result, conventional uncertainty metrics can give the appearance of a well defined expected mean even when substantial structured variability persists in the data.

Magnetic particle spectroscopy (MPS) is widely used for biosensing applications due to its exceptional sensitivity to

<sup>a</sup> a. Department of Materials Science and Engineering, Clemson University, Clemson, SC 29634, USA E-mail: arabelh@clemson.edu

<sup>b</sup> b. School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA

<sup>c</sup> Independent Researcher



changes in nanoparticle media, surface chemistry, and interfacial dynamics. Ligand binding and particle mobility can induce subtle measurable shifts in the harmonics spectra enabling the detection of a variety of biomarkers.<sup>6–9,14,15</sup> To support these applications, surface coatings such as poly(ethylene glycol) (PEG) are widely used to enhance colloidal stability and provide chemical handles for target moieties.<sup>11,16–18</sup> MPS is inherently sensitive to minute changes in nanoparticle dynamics, providing sensing capabilities for subtle perturbations in magnetic behavior. While this enables powerful bio sensing applications, it is highly prone to measurement variability in external factors such as temperature, aggregation, and instrumentation even under tightly controlled experimental conditions. Such variability may mimic true biological binding events that could lead to false positives or indistinguishable results in clinical applications.<sup>19,20</sup>

Aggregation, local viscosity changes, and medium induced fluctuations have been shown to cause non-constant frequency scaling, while colloidal stabilization can nonlinearly affect relaxation behavior, hydrodynamics, and electrophoretic mobilities.<sup>6,21</sup> Importantly, these fluctuations can be hidden by the intrinsic nature of scale in magnetic measurements. When harmonic data for grossly different samples are visualized such is the case in calibration dilutions, fluctuations may appear suppressed due to logarithmic scaling despite suffering from the same variability problems on local signal interpretation as shown in Fig. 3.

These observations highlight the need for a robust statistical model that explicitly accounts for known causes of variation in MPS harmonic measurements. A common approach for analyzing harmonic data is using least squares (LS) regression or standard linear models that implicitly assume stationary and homoscedastic noise. Such methods offer practical estimates of harmonic behavior and are frequently adapted in the MPS literature.<sup>7,9,11,14,22</sup> However, these assumptions are poorly aligned with non-stationary systems which fail to account for within-sample and temporal variability. As a result, uncertainty is systematically underestimated, raising the risk of overly confident interpretations when random noise, system drift, or sudden fluctuations in the signal occur.

In MPS measurements, LS based models collapse temporal and within sample variation into a single error term, thus leading to systematically underestimated uncertainty. This produces artificially narrow confidence intervals that mask real signal fluctuations and create a false impression of precision. This limitation reflects the known behavior inherent to MPS and MPI devices where signal distortions arise from a combination of factors, such as background feed-through, system drift, random noise, drive field fluctuations, and abrupt fluctuations of unidentified origin.<sup>23,24</sup> While there will always be mechanical improvements to limit variability, such occurrences cannot be fully eliminated in practice. Consequently, robust statistical modeling is required to accurately represent and interpret MPS in data in non ideal experimental systems.

Recognizing this gap, the present study treats within and between sample variability as an inherent property of this material system, to establish a statistically rigorous framework for evaluating MPS measurements. We systematically examine signal re-

producibility in PEG coated iron oxide nanoparticles by quantifying both aliquot-level (within sample,  $\sigma_{W_i,j}$  and day-to-day variation/temporal,  $\sigma_{B_i}$ ) for liquid phase 20 kDa PEG-coated 22 nm iron oxide nanoparticles. We demonstrate the necessity of mixed effect models that explicitly account for the aforementioned random effects, providing an accurate and statistically defensible description of MPS data. This framework enables meaningful uncertainty quantification and reproducible interpretation in non-ideal experimental setups. By looking closely into the origin and persistence of these perturbations, this work advances the reliability of MPS measurements, introduces a new standard for analysis protocol, and establishes a statistically sound foundation for the clinical translation in biomarker detection, bio imaging, and targeted therapeutic applications.

## Results and Discussion

### Structural and colloidal characterization

Iron oxide nanoparticles with a mean core diameter of 22 nm were synthesized and functionalized with a 20 kDa nitroDOPA terminated poly(ethylene glycol) (PEG) ligand for this study. Standard characterizations including transmission electron microscopy (TEM) and X-ray diffraction (XRD) were employed to determine the size and magnetite phase of this material. Transmission electron microscopy (TEM) confirmed the formation of monodisperse, spherical iron oxide particles with a core size range of  $21.6 \pm 2.7$  nm (see Fig. 1 c).

X-ray diffraction patterns were indexed to magnetite confirming the crystalline phase (see Fig. 1 d). positions and profile changes across the series. The diffraction patterns are indexed to the cubic inverse spinel structure of  $\text{Fe}_3\text{O}_4$  (JCPDS 85-1436). The full width at half maximum value of the (311) reflection as determined from a Gaussian fitting was input to the Debye Scherrer equation, as described in Equation 5. The crystallite size was estimated to be 14.2 nm. This value is slightly smaller than the average particle size obtained from TEM ( $21.6 \pm 2.7$  nm). This discrepancy is expected as XRD provides the size of diffracting crystalline domains, whereas TEM measures the overall particle size; in this case, it suggests multiple domains within a single particle.

Applying a polymer ligand is well established method to increase colloidal stability of nanoparticles in water.<sup>18,25,26</sup> This molecular weight (20 kDa) was chosen due to its high steric repulsion and stabilization properties.<sup>16,18</sup> The PEG coating is anchored to the iron oxide surface via nitroDOPA, a catechol amine, forming strong chelation interactions that are known to provide stable surface functionalization under aqueous conditions.<sup>16,17,26</sup> Dynamic light scattering (DLS) was therefore used to assess colloidal stability in DI water over 24 hours. Figures 1 a and b show the DLS intensity average and Z-average distributions for PEG coated IONPs in DI water respectively. The hydrodynamic diameter decreased slightly from 96.1 to 90.4 nm over 24 hours with narrow intensity average distributions and no emergence of large aggregates. This indicates that gross colloidal instability or irreversible aggregation is not present, and any variability in MPS measurements likely originate from mechanisms beyond simple



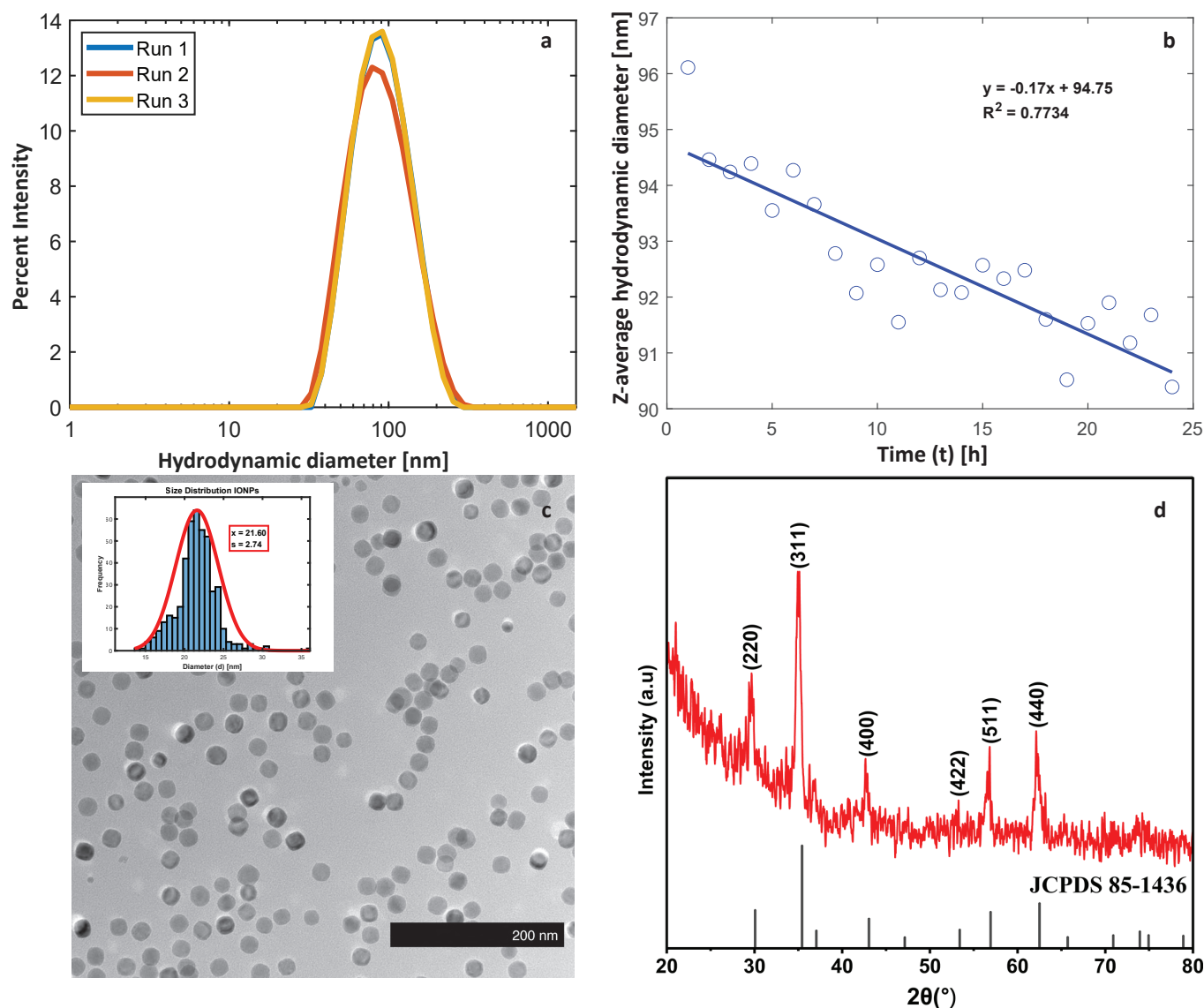


Fig. 1 The colloidal dynamics of the particles measured through DLS show a lack of aggregate formation over the course of twenty four hours (b). Intensity distribution data (a) show a mean hydrodynamic diameter of 118.1 [nm] with a polydispersity average index (PDI) of 0.18 [-]. Transmission electron microscopy images (c) show monodisperse spherical particles. A normal distribution fit to the particles shows a mean of 21.60 nm with a standard deviation of 2.74 nm. The crystalline phase of the synthesized IONPs was confirmed with X-ray diffraction (d), with diffraction peaks consistent with key peaks for Magnetite.

sedimentation within the timescale of a day. Together, these results suggest that ligand binding and colloidal stability are maintained over the course of measurements across multiple days.

### Temporal variability in MPS measurements

Quantitative analysis of MPS data often relies on three main harmonic parameters: the mass normalized third harmonic  $A_3$ , the ratio of the 5th to 3rd harmonic  $A_5/A_3$ , and the number of detectable harmonics above the noise floor  $A_k$ .<sup>11–13</sup> To evaluate signal reproducibility while minimizing operator bias, a statistically rigorous MPS measurement protocol was implemented. Measurements were conducted on three independently prepared aliquots per sample over five days, with random measurement order and treatment. Despite consistent handling and identical concentra-

tion, substantial variability well above the noise floor was observed in the third harmonic amplitude ( $A_3$ ) and the harmonic ratio ( $A_5/A_3$ ) both between sample runs and across acquisition days. Specifically, abrupt within-run fluctuations are observed in the amplitude data, whereas these features are less apparent in the ratio data. This apparent normalization of the harmonic amplitudes does not mitigate systemic uncertainty; instead, it amplifies within-run observation noise ( $\epsilon_{ijk}$ ) relative to between-day and between-run variation. Computing the ratio removes location information while rescaling the random variables. Consequently,  $\epsilon_{ijk}$  cannot be directly compared between amplitude and ratio representations because the original scale and location is not preserved. However, the relative contribution of  $\epsilon_{ijk}$  to the overall variance structure remains interpretable. Visualization of the



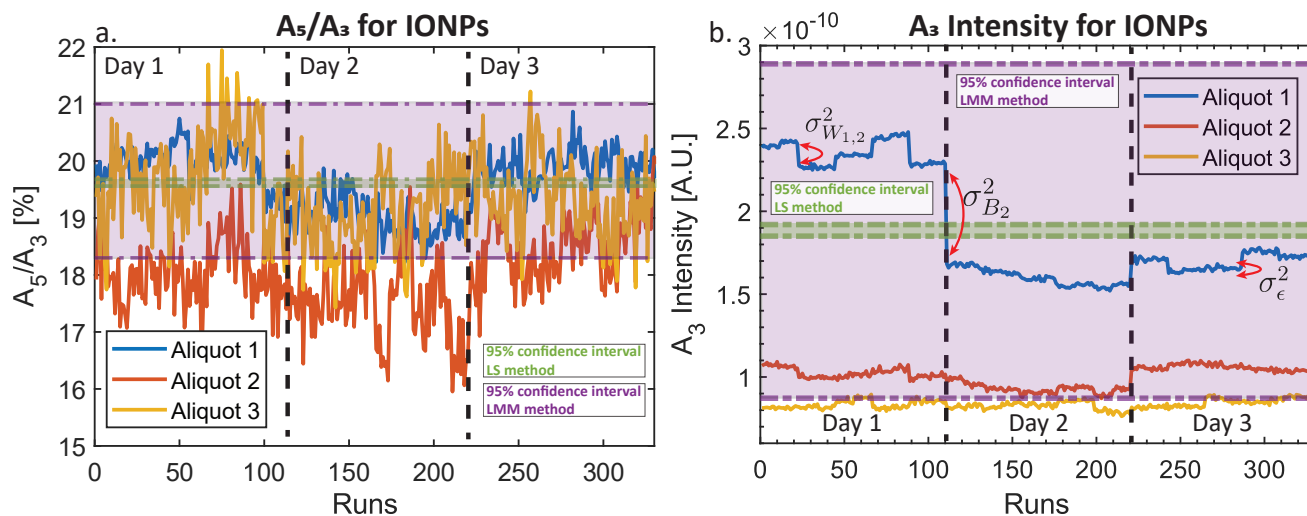


Fig. 2 Fluctuations in the (a) harmonic ratio  $A_5/A_3$  and (b) harmonic amplitude  $A_3$  over three days. Shown are the harmonic ratios and background subtracted amplitudes of 20 kDa 22 nm IONPs aliquots obtained on Day 1 measured over days 1 - 3. The 95% confidence intervals using the least squares and generalized mixed effect model are shown for both data sets.

raw harmonic amplitudes therefore reveals sources of variation evident which are otherwise masked when ratios are used.

To illustrate the consequences of these assumptions, least-squares (LS) regression was applied to the third harmonic ( $A_3$ ) and harmonic ratio ( $A_5/A_3$ ) to estimate the mean and corresponding 95% confidence intervals. Three independently prepared aliquots of 20 kDa 22 nm IONPs were measured randomly over three days to evaluate within-run and between-day variability. When 95% confidence intervals for mean third harmonic,  $\bar{A}_3$ , were calculated by treating each day as an independent sample, the removal of temporal structure resulted in non overlapping confidence intervals. The resulting LS confidence intervals are summarized in Table 1. This approach concludes that measurements of the same aliquot collected on different days are statistically different.

When all three days are pooled into a single dataset, the resulting LS confidence intervals [ $1.85 \cdot 10^{-10} - 1.92 \cdot 10^{-10}$  A.U.] shown in Figure 2 b, remain unrealistically narrow. In fact the resulting intervals never contains a single data point for the series. Consequently, all measured values fall well outside the 95% confidence bounds, indicating a systemic underestimation of uncertainty and complete misrepresentation of the data. This failure arises because LS regression collapses within aliquot and day to day variability random effects into a single residual error term thus producing artificially narrow confidence intervals. Critically, this behavior persists even when data appear well behaved on logarithmic scales with well known commercial agent, Synomag™-D 70 nm lot: 54525104-04 (see Fig 3). This underscores the necessity of a model that attributes variation to random effects such as temporal variability.

To properly account for this variability, a linear mixed effects model was applied (see Eq. 3) where  $\mu$  is the mean of the tar-

geted harmonic,  $B_i$  represents day to day fluctuations,  $W_{ij}$  captures between-run within-day variation and  $\epsilon_{ijk}$  denotes residual measurement noise. The random effect structure should be selected based on known or potential sources of variation as a function of the physical workflow of the observation process. For example, variability arising from day to day changes (e.g., ambient temperature, sample aging, or instrument drift) should be modeled at the day level capturing valuable timescales for degradation studies or biomarker uptake. Additionally, variability introduced within a measurement session (e.g., local heating, sample repositioning, and sonication induced rearrangement of nanoparticle clusters) may be captured at the run level. In practice, researchers should identify candidate sources of variability and determine whether they occur at predictable intervals, then assign random effects accordingly (see Section Statistical modeling).

Unlike LS regression, the mixed effects model produces confidence intervals that accurately summarize the uncertainty in our estimates based on the observed data. Applying this model to the dataset of aliquot one across all days shown in Fig 2 illustrates the lack of confidence in the estimated value of the harmonic data. With this method, the individual sources of temporal variability are accounted for and result in all harmonic data falling within the LMM confidence intervals. Assigning defined jumps in the data to random effects  $B_i$  and  $W_{ij}$  with the LMM method, yield a 95% confidence bound [ $8.72 \cdot 10^{-11} - 2.89 \cdot 10^{-10}$ ] as highlighted in Fig 2 b for aliquot one across all days. This method well encompasses the amplitude data in contrast to the LS method.

It should be noted that the power of the LMM method arises from its ability to accommodate temporal and between-run fluctuations. If the dataset were collapsed by day as shown in Table 1, the mixed effect model simplifies to a LS model where the random effect terms go to zero such that all variability is attributed



Table 1 Least squares confidence intervals for aliquot one 20 kDa IONPs

Day	95% CI [A.U.]	Sample Estimates Mean [A.U.]
1	$[2.34 \cdot 10^{-10} - 2.36 \cdot 10^{-10}]$	$2.35 \cdot 10^{-10}$
2	$[1.59 \cdot 10^{-10} - 1.61 \cdot 10^{-10}]$	$1.60 \cdot 10^{-10}$
3	$[1.69 \cdot 10^{-10} - 1.71 \cdot 10^{-10}]$	$1.70 \cdot 10^{-10}$

to residual measurement noise. While the mixed effects model is general and can, in principle, be applied to other nanoparticle formulations and MPS platforms, the present study evaluates a single well characterized IONP system. Extension to particles with different core sizes, coatings, or interaction effects may yield different variance structures and should be validated in future studies.

Alternative strategies such as harmonic ratio analysis or observation of harmonic phase are commonly employed to mitigate mass loading effects.<sup>7,9,14</sup> In the present data, such techniques did not reduce overall variability but instead redistributed it, amplifying within-run stochastic noise relative to between-run and between-day variation. Although amplitude measurements exhibit abrupt fluctuations, these changes occur at predictable intervals and can be ascribed to identifiable aspects of the observation process that are naturally accommodated within a hierarchical random effects framework. In contrast, ratio operation introduces a nonlinear combination of two random variables spanning an order of magnitude. This operation removes information about the absolute signal magnitude, which reflects factors such as particle mass, interactions and system sensitivity, while rescaling variance contributions. As a result, direct comparison of the noise structure within a linear mixed effects model becomes more complex. The apparent compactness of the ratio distribution therefore reflects a scaling effect rather than an intrinsic reduction in signal variability (see Fig 2 a.) When mass is independently known, modeling amplitude data provides a more direct representation of drift, trends, and structured fluctuations across observation intervals. Ratios may instead be viewed as additional derived observations that require explicit modeling of their constituent harmonic relationships.

Prior work has demonstrated empirical correlations between harmonic amplitudes to reduce quantification uncertainty.<sup>19</sup> While such relationships improve calibration, they do not account for intrinsic signal variability arising from non-stationary particle dynamics. Importantly, the observed variance in amplitude measurements reflects fundamental characteristics of MPS which will vary instrument to instrument. Spatial sensitivity of detection coils and temperature dependent nanoparticle dynamics introduce hierarchical sources of uncertainty that persist even in optimized systems. As the demand for measurement sensitivity increases, statistical frameworks that explicitly model hierarchical variability become essential for resolving subtle signal perturbations necessary for advanced magnetic characterization and bio diagnostic applications.

### Potential sources of variability

Across multiple days, runs, and aliquots, our PEG coated IONPs exhibited non negligible temporal and within sample variability

in the measured harmonic parameters ( $A_3^*$ ,  $A_5/A_3$ ), far exceeding the instrument noise floor (see Fig 3 b). Confidence intervals obtained via standard LS methods fail to accommodate these extra sources of variation and systematically underestimate uncertainty. The operation of computing the harmonic ratio, while commonly used in the field, propagated noise and masked the attributions of underlying variances observed in the amplitude data. While the phase data could pose as a more stationary mass independent metric, this falls out of the scope of this project and is suggested as future work. We hypothesize that a portion of the observed variance arises from reversible interacting nanostructures (e.g. dimers or small chains) dynamically created during MPS acquisition. Similar field induced clustering has been extensively shown in literature.<sup>3,11,18,27-29</sup> These structures would transiently modulate Brownian relaxation times on millisecond to second timescales, producing instantaneous changes in harmonic amplitudes.<sup>30</sup> This mechanism is consistent with the colloidal stability confirmed through dynamic light scattering (DLS) (Fig 1 b) and non-stationary fluctuations within a single run.

Mixed-effects modeling provides a statistically rigorous method to partition variance and correctly represent uncertainty in MPS measurements. While this paints a clear picture of the data, this method does not inherently improve sample distinguishability, highlighting a limitation of diagnostic applications. Future improvements may involve instrumentation adjustments to control extraneous variation (e.g., feedback controlled drive fields, improved sample assembly), or the use of grossly different materials in which the threshold of positive signal in the case of disease testing would vastly out scale the variance. For example, nanoparticles with highly different hydrodynamic sizes, surface coatings, or morphology would be expected to significantly influence the magnetic response.<sup>27,28,31</sup> These considerations are also relevant for smaller superparamagnetic iron oxide nanoparticles (SPIONs), where reduced core size impacts signal characteristics and potentially interactions and variance structure. While these effects may differ from the present system, the mixed effects framework remains applicable for quantifying variability when the relevant sources and timescales are properly defined.

In this study, a 20 kDa PEG coating was selected to provide a stable, well dispersed system; however, varying PEG molecular weights may further alter hydrodynamic size, interparticle interactions, and relaxation dynamics, therefore impacting both signal magnitude, harmonic decay, and variability. This highlights that material driven differences can outweigh measurement variability as illustrated in Fig 3, where separation is possible even in the presence of large fluctuations. Together, these findings highlight the need for both statistical rigor and careful experimental design in reliable MPS instrumentation.

## Experimental Methods

### Statistical modeling

We propose using a linear mixed effect model (LMM) for estimating mean harmonics. A typical least squares (LS) analysis handles observed variation including day to day, between-run, and within-run noise into a single source and averages the deviations



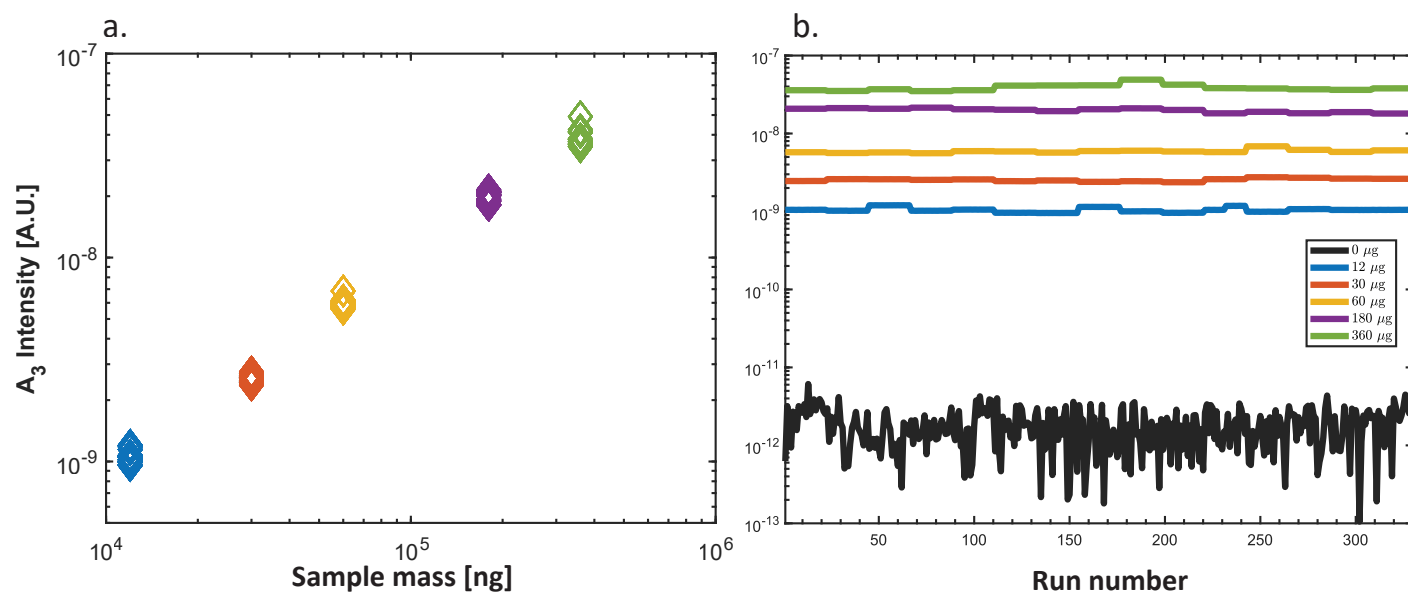


Fig. 3 Synomag-D 70 nm dilution series plotted as signal intensity of the third harmonic  $A_3$  vs (a) iron mass and (b) run number.

over all the observations. The LMM, on the other hand, assumes specific independent error terms for each source of variation such that when the day to day and between-run variation are large compared to the within-run noise, the LMM will give a more realistic, albeit wider, estimate of uncertainty. When day to day and between-run variations are small compared to the noise, the two methods will give very similar results.

In a standard LS approach, we assume that the perturbations we observe in harmonic signals are independent with a mean of 0 and common variance about the true mean  $\mu$ . Typically, normality is assumed although approximations can typically be obtained of non-normal data. We would define

$$\mathcal{Y}_{ijk} = \mu + \varepsilon_{ijk}, \text{ where} \quad (1)$$

$$\varepsilon_{ijk} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{\varepsilon_{LS}}^2)$$

where the indices  $i$ ,  $j$  and  $k$  correspond to the days, individual runs (within-days) and single observations (within-run nested in days),  $\mu$  is the mean of the harmonic and  $\varepsilon_{ijk}$  is the independent observed deviations. The maximum likelihood estimator of  $\mu$  i.e.  $\hat{\mu}$  is obtained (for the balanced case) as  $\hat{\mu}_{LS} = \frac{\sum_i \sum_j \sum_k \mathcal{Y}_{ijk}}{IJK}$  where  $I$ ,  $J$  and  $K$  are the number of days, runs and observations respectively. The standard error of an estimate  $\hat{\mu}_{LS}$  is

$$\widehat{SE}_{LS} = \sqrt{\frac{S_{\varepsilon_{LS}}^2}{IJK}} \quad (2)$$

The limitation of Equation 2 is that when between-days and between-runs variability dominate relative to the more frequently observed within signal variation, the large denominator mutes these effects. As a result, the least squares confidence interval in Figure 2 b is overly narrow and does not contain any of the observation specific harmonics.

With the LMM we assume

$$\mathcal{Y}_{ijk} = \mu + B_i + W_{ij} + \varepsilon_{ijk}, \text{ where}$$

$$B_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_B^2)$$

$$W_{ij} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_W^2)$$

$$\varepsilon_{ijk} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{\varepsilon_{LMM}}^2) \quad (3)$$

The  $B_i$ 's are the day to day observations, the  $W_{ij}$ 's describe the run to run variation and the  $\varepsilon_{ijk}$ 's correspond to the between signal perturbations with a balanced design  $\hat{\mu}_{LS} = \hat{\mu}_{LMM}$ , however

$$\text{Var}(\hat{\mu}_{LMM}) = \frac{\sigma_B^2}{I} + \frac{\sigma_W^2}{IJ} + \frac{\sigma_{\varepsilon_{LMM}}^2}{IJK} \quad (4)$$

where  $\sigma_{\varepsilon_{LMM}}^2 < \sigma_{\varepsilon_{LS}}^2$ . Even though the noise variance in the LMM is smaller, in LS the day to day and run to run variation can only be reduced by replicating the entire experiment. With non stationary observations, these terms end up dominating the estimate of the standard error.

### Synthesis of PEG IONPs

Iron oxide nanoparticles (IONPs) of 22nm diameter were synthesized using a thermal decomposition of iron(III) acetylacetonate, following a modified procedure described by Sun and Zeng.<sup>32</sup> Methoxy terminated polyethylene glycol (PEG, 20 kDa,) purchased from JenKem was functionalized using a previously described procedure with nitroDOPA : a catechol amine with high binding affinity to the nanoparticle surface.<sup>18,26</sup> The modified PEG was attached to the IONPs using a ligand exchange to create an aqueous dispersion following a previous protocol.<sup>17</sup>



### Particle characterization

The hydrodynamic size, degree of clustering, and polydispersity were characterized using dynamic light scattering (DLS, Malvern Zetasizer Nano ZS) using a 633 nm wavelength laser at 25 °C. Water dispersed samples were prepared by filtering with a 5  $\mu\text{m}$  Teflon syringe filter into disposable plastic cuvettes prior to measurement.

Size analysis of the magnetic core was obtained using transmission electron microscopy (TEM, Hitachi H7830). The samples were prepared by drop casting diluted hexane dispersions onto carbon-coated copper grids and dried under ambient conditions. Image analysis of approximately 300 particles was performed using ImageJ®.

Iron concentration was determined by inductively coupled plasma mass spectroscopy (ICP-MS) following nitric acid digestion. Two microliters of each sample were dried in a 500 °C oven over night to remove organic matter, then digested in 5 mL of trace metals nitric acid at 90 °C for one hour. The sample was then dispersed in 5 mL of 2 % nitric acid, sonicated, filtered with a 0.2  $\mu\text{m}$  Teflon syringe filter, and quantified using calibration curves derived from iron standards.

X-ray diffraction (XRD) was performed using a Rigaku Smart-Lab powder X-ray diffractometer to determine the phase and composition. Measurements were carried out using Cu  $K\alpha$  radiation ( $\lambda = 1.54\text{\AA}$ ). Samples were prepared by drop wise deposition of hexane dispersions onto a circular glass pane, forming uniform dry films. Diffraction patterns were collected over the  $2\theta$  range of 15 to 80 degrees with a scan rate of 5 degrees/minute and a step size of 0.02 degrees.

To improve the signal to noise ratio, a smoothing function was applied using a Savitzky-Golay filter in Origin. The smoothing was used only to assist peak visualization and fitting does not affect peak broadening. Peak positions and full width at half maximum (FWHM) values were obtained by fitting the diffraction peaks with a Gaussian function in Origin. The (311) reflection was selected due to its high intensity and minimal overlap with neighboring peaks. The crystallite size in nanometers was estimated using the Debye Scherrer equation

$$d = \frac{K\lambda}{\beta \cos \theta} \quad (5)$$

where  $K = 0.9$ ,  $\lambda = 1.54\text{\AA}$ ,  $\beta$  is the FWHM in radians, and  $\theta$  is the Bragg angle. No explicit correction for instrumental broadening was applied; therefore, the calculated crystallite sizes should be considered approximate.

### Magnetic measurements

Magnetic particle spectroscopy (MPS) measurements were performed on a custom-built instrument based off an open source assembly.<sup>22</sup> Samples were exposed to an alternating magnetic field with an of 10 mT at a frequency of 23.8 kHz. The induced voltage signal was sampled at 2 MHz using a National Instruments (Austin, TX, USA) USB 6363 data acquisition card (DAQ). Each run consisted of 22 measurements, with each measurement lasting 40 ms. From these measurements, 22 fast Fourier transform

(FFT) harmonic spectra were generated. After 11 measurements within a run, the sample is moved well outside the bore and a new baseline is acquired. The individual data points analyzed in this work are the FFT values corresponding to the frequency for the  $n$ th harmonic.

For each sample type, three aliquots containing 60  $\mu\text{L}$  of substance, were prepared in 0.6 mL microcentrifuge Eppendorf tubes. Each observation consisted of five sequential runs. Between runs, the motor was homed to null any accrued motor position offset and the baseline was reset between excitation pulses. This procedure yielded a total of 110 harmonic spectra per observation.

Prior to measurement, samples were sonicated in a bath sonicator at 25 °C for 30 seconds to ensure thermal equilibrium and minimize settling effects before entering the bore. The series of samples measured each day were input in a randomized sequence. The instrument noise floor was defined as  $A_k = \mu_k + 10\sigma_k$  where  $A_k$  is the amplitude of the  $k_{th}$  harmonic,  $\mu_k$  is the mean of the  $k_{th}$  harmonic and  $\sigma_k$  is the standard deviation obtained from 10 blank runs, respectively, following the guidance of the International Union of Pure and Applied Chemistry (IUPAC).

To adequately sample potential sources of variation, a five day measurement protocol was conducted. On day 1, three aliquots from the 20 kDa PEG IONP sample were prepared to serve as the initial aliquot reference. On subsequent days, three aliquots were freshly pipetted and measured in a randomized sequence alongside Day 1 references to evaluate within sample and temporal variation. This allowed for analysis of aliquot-to-aliquot variation between two series of a given sample measured on the same day. Additionally, the same aliquots from day 1 were remeasured daily to assess instrument drift.

### Conclusions

This work demonstrates that commonly applied statistical approaches in magnetic particle spectroscopy do not fully represent experimental data acquired under realistic, non-stationary conditions. Measurements of synthesized 20 kDa PEG-coated 22 nm iron oxide nanoparticles exhibited substantial variability within aliquots and across days despite controlled preparation and measurement protocols, motivating the implementation of a hierarchical mixed effects framework to partition variance and represent uncertainty across observation levels. Least squares modeling and standard error based confidence intervals systematically underestimated this variability. Commonly employed transformations such as harmonic ratios propagated noise without resolving its underlying structure, complicating direct interpretation of signal uncertainty. We further show that measurement ranges spanning logarithmic scales, such as in dilution series, can visually suppress structured fluctuations that dominate within-run stochastic noise in diagnostically relevant signal levels. In contrast, mixed effects modeling captures within and temporal variability, providing confidence intervals that reflect the hierarchical nature of the measurement process and enable meaningful comparisons where traditional approaches are insufficient. These findings emphasize the importance of statistical frameworks that explicitly model hierarchical variability rather than relying on



normalization or scaling transformations. Such rigor is essential for reliable interpretation of MPS signal and advancing magnetic characterization. Establishing methods that accurately represent experimental uncertainty is a necessary step towards translating MPS for disease diagnostics, where subtle fluctuations in magnetic signal form the basis of biomarker detection.

## Author contributions

Arabella Stockdale: data curation, conceptualization, funding acquisition, investigation, methodology

Joe Bible: formal analysis, conceptualization, methodology

Eli Mattingly: data curation, investigation

Olin Thompson Mefford: Conceptualization, funding acquisition, methodology, project administration

## Conflicts of interest

There are no conflicts to declare.

## Data availability

Data for this article, including images, raw magnetic data, ICP concentrations, are available at Zenodo at <https://doi.org/10.5281/zenodo.18914692>

## Acknowledgments

The authors would like to thank David White, Prof. Thomas Crawford, Prof. Brian Powell, and Dr. Benjamin Fellows for their many influential conversations. This material is based upon work supported by the NSF Graduate Research Fellowship Program under Grant No. 2139915. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. A.R.S. acknowledges funding from the Beckman Scholars Program (award year 2021, host institution was Clemson University).

## Notes and references

For the reference section, the style file `rsc.bst` can be used to generate the correct reference style.<sup>§</sup>

1. Citations should appear here in the format A. Name, B. Name and C. Name, *Journal Title*, 2000, **35**, 3523;
2. A. Name, B. Name and C. Name, *Journal Title*, 2000, **35**, 3523.

...

We encourage the citation of primary research over review articles, where appropriate, in order to give credit to those who first reported a finding. Find out more about our commitments to the principles of San Francisco Declaration on Research Assessment (DORA).

- 1 H. Shokrollahi, *Mater Sci Eng C Mater Biol Appl*, 2013, **33**, 4485–97.
- 2 F. Li, D. Zhi, Y. Luo, J. Zhang, X. Nan, Y. Zhang, W. Zhou, B. Qiu, L. Wen and G. Liang, *Nanoscale*, 2016, **8**, 12826–33.
- 3 Z. W. Tay, S. Savliwala, D. W. Hensley, K. L. B. Fung, C. Colson, B. D. Fellows, X. Zhou, Q. Huynh, Y. Lu, B. Zheng, P. Chan-

- drasekharan, S. M. Rivera-Jimenez, C. M. Rinaldi-Ramos and S. M. Conolly, *Small Methods*, 2021, **5**, 2100796.
- 4 B. Gleich and J. Weizenecker, *Nature*, 2005, **435**, 1214–1217.
- 5 O. C. Sehl, J. J. Gevaert, K. P. Melo, N. N. Knier and P. J. Foster, *Tomography*, 2020, **6**, 315–324.
- 6 D. Jyoti, S. W. Gordon-Wylie, D. B. Reeves, K. D. Paulsen and J. B. Weaver, *Sensors (Basel)*, 2022, **22**, 123–130.
- 7 E. L. Rösch, R. Sack, M. S. Chowdhury, F. Wolgast, M. Zaborski, F. Ludwig, M. Schilling, T. Viereck, U. Rand and A. Lak, *ChemBioChem*, 2024, **n/a**, e202400251.
- 8 K. Wu, J. Liu, R. Saha, D. Su, V. D. Krishna, M. C. J. Cheeran and J.-P. Wang, *ACS Applied Materials & Interfaces*, 2020, **12**, 13686–13697.
- 9 K. Wu, V. K. Chugh, V. D. Krishna, A. Di Girolamo, Y. A. Wang, R. Saha, S. Liang, M. C. J. Cheeran and J.-P. Wang, *ACS Applied Materials & Interfaces*, 2021, **13**, 44136–44146.
- 10 S. Biederer, T. Knopp, T. F. Sattel, K. Lüdtke-Buzug, B. Gleich, J. Weizenecker, J. Borgert and T. M. Buzug, *Journal of Physics D: Applied Physics*, 2009, **42**, 205007.
- 11 A. Remmo, D. Eberbeck, O. Kosch, L. Kampen, A. Ludwig, A.-N. Egler-Kemmerer, F. Wiekhorst and N. Loewa, *ACS Applied Nano Materials*, 2024, **7**, 24315–24324.
- 12 K. Wu, D. Su, R. Saha, J. Liu, V. K. Chugh and J.-P. Wang, *ACS Applied Nano Materials*, 2020, **3**, 4972–4989.
- 13 N. Löwa, D. Gutkelch, E.-A. Welge, R. Welz, F. Meier, A. Baki, R. Bleul, T. Klein and F. Wiekhorst, *Nanomaterials*, 2020, **10**, 2277.
- 14 P. Vogel, M. A. Rückert, B. Friedrich, R. Tietze, S. Lyer, T. Kampf, T. Hennig, L. Dölken, C. Alexiou and V. C. Behr, *Nature Communications*, 2022, **13**, 7230.
- 15 X. Zhang, D. B. Reeves, I. M. Perreard, W. C. Kett, K. E. Griswold, B. Gimi and J. B. Weaver, *Biosens Bioelectron*, 2013, **50**, 441–6.
- 16 K. Davis, B. Qi, M. Witmer, C. L. Kitchens, B. A. Powell and O. T. Mefford, *Langmuir*, 2014, **30**, 10918–10925.
- 17 A. Lassenberger, O. Bixner, T. Gruenewald, H. Lichtenegger, R. Zirbs and E. Reimhult, *Langmuir*, 2016, **32**, 4259–69.
- 18 S. L. Saville, B. Qi, J. Baker, R. Stone, R. E. Camley, K. L. Livesey, L. Ye, T. M. Crawford and O. T. Mefford, *J Colloid Interface Sci*, 2014, **424**, 141–51.
- 19 N. Löwa, M. Seidel, P. Radon and F. Wiekhorst, *Journal of Magnetism and Magnetic Materials*, 2017, **427**, 133–138.
- 20 L. Moor, S. Scheibler, L. Gerken, K. Scheffler, F. Thieben, T. Knopp, I. K. Herrmann and F. H. Starsich, *Nanoscale*, 2022, **14**, 7163–7173.
- 21 A. Lak, Y. Wang, P. J. Kolbeck, C. Pauer, M. S. Chowdhury, M. Cassani, F. Ludwig, T. Viereck, F. Selbach, P. Tinnfeld, M. Schilling, T. Liedl, J. Tavacoli and J. Lipfert, *Nanoscale*, 2024, **16**, 7678–7689.
- 22 E. Mattingly, A. C. Barksdale, M. Śliwiak, J. Chacon-Caldera, E. E. Mason and L. L. Wald, *Review of Scientific Instruments*, 2024, **95**, 063706.
- 23 H. Paysen, O. Kosch, J. Wells, N. Loewa and F. Wiekhorst, *Physics in Medicine & Biology*, 2020, **65**, 235031.



- 24 F. H. Niebel, J. Chacon-Caldera, E. Mattingly, A. C. Barksdale, M. Sliwiak and L. L. Wald, *International Journal on Magnetic Particle Imaging*, 2024, **10**, 1–4.
- 25 J. N. Anker and O. T. Mefford, *Biomedical Applications of Magnetic Particles*, CRC Press, 2020.
- 26 E. Amstad, T. Gillich, I. Bilecka, M. Textor and E. Reimhult, *Nano Letters*, 2009, **9**, 4042–4048.
- 27 Z. Zhao and C. Rinaldi, *Physics in Medicine & Biology*, 2020, **65**, 185013.
- 28 S. K. Avugadda, S. Wickramasinghe, D. Niculaes, M. Ju, A. Lak, N. Silvestri, S. Nitti, I. Roy, A. C. S. Samia and T. Pellegrino, *Nanomaterials*, 2020, **11**, 62.
- 29 A. V. Makela, M. A. Schott, C. S. Madsen, E. M. Greeson and C. H. Contag, *Nano Letters*, 2022, **22**, 4630–4639.
- 30 Y. Sun, Z. Du, H. Zhang, H. Wang, T. Sasayama and T. Yoshida, *Nanoscale*, 2023, **15**, 16089–16102.
- 31 Z. Zhao, N. Garraud, D. P. Arnold and C. Rinaldi, *Physics in Medicine & Biology*, 2020, **65**, 025014.
- 32 S. Sun and H. Zeng, *Journal of the American Chemical Society*, 2002, **124**, 8204–8205.



## Data Availability

Data for this article, including images, raw magnetic data, ICP concentrations, are available at Zenodo at <https://doi.org/10.5281/zenodo.18914692>

