Supplementary Information

Full-Harmonics Phasor Analysis: Unravelling Multi-Exponential Trends in Magnetic Resonance Imaging Data

Wouter M. J. Franssen1,*, Frank J. Vergeldt1, Arjen N. Bader1,2, Herbert van Amerongen1,2, Camilla Terenzi1,*

1Laboratory of Biophysics, Wageningen University & Research, Wageningen, The Netherlands.
2MicroSpectroscopy Centre, Wageningen University & Research, Wageningen, The Netherlands.
1 PROJECTIONS OF PHASOR DATA

Figure S1: (a) 3D phasor plot with axes Re, Im and Re for a series of 1000 simulated bi-exponential decays (orange points) with lifetimes Δt and 10 · Δt, using 20 time points, and relative populations ranging from 0 to 100%. Gaussian noise with a relative amplitude of 0.1 was added in the time domain. The semicircle is shown in blue. (b)-(d) 2D projections plots of data in (a), namely onto Im vs. Re (b), Re2 vs. Re1 (c) and Re2 vs. Im (d) planes.

2 AXIS SELECTION BY THREE LIFETIMES

To define a projection on a plane, we need to define two orthogonal axes. This can be done by choosing three points in -dimensional space, which define a plane (we do not care about the rotation of this plane at the moment). In order to aid the selection of these points, we limit ourselves to a position on the semicircle, which we can calculate. As each position on the semicircle is associated with a lifetime, we essentially need to select three lifetimes to define the projection plane.

Based on the positions of these three lifetimes in -dimensional phasor space, the vectors , and , we can calculate the projection axes in the following way:
• Define $v$ as the difference between the shortest and longest lifetime: $v = \gamma_1 - \gamma_3$
• Normalize $v$ to obtain $\mathbf{axis}_1: \mathbf{axis}_1 = ||v||$
• Define $v_{\perp}$ according to: $v_{\perp} = \gamma_1 - \gamma_2$
• Make $v_{\perp}$ orthogonal to $\mathbf{axis}_1$ by Gram-Schmidt orthogonalization:
  o Calculate the projection of $v_{\perp}$ on $\mathbf{axis}_1$: $\mathbf{err} = v_{\perp} \cdot \mathbf{axis}_1$
  o Subtract this from $v_{\perp}$ to obtain $\mathbf{u}_2$: $\mathbf{u}_2 = v_{\perp} - \mathbf{err}$
  o Normalize $\mathbf{u}_2$ to obtain $\mathbf{axis}_2$: $\mathbf{axis}_2 = ||\mathbf{u}_2||$

These two axes can now be used for the projection. The projection itself is calculated by an inner product. If we take a certain phasor position $P$, we can calculate the value along $\mathbf{axis}_1$ by: $P \cdot \mathbf{axis}_1$. The value along $\mathbf{axis}_2$ is calculated by $P \cdot \mathbf{axis}_2$.

In most cases, it is desirable to rotate $\mathbf{axis}_1$ and $\mathbf{axis}_2$ to make the resulting plot more like a normal, first harmonic phasor plot. For that we request that:

1. All phasor positions are positive in both the $x$ and $y$ direction
2. The semicircle and the $x$-axis should be the limits between which all data should fall in the case of exponential decays (when there is no noise).

In practice this means that we must rotate the axes to make sure that the end points of the semicircle both lie on the positive $x$-axis. In some cases, a flip of the $y$-axis is needed to get the semicircle in the $y > 0$ region.

This rotation is only cosmetic, but does lead to a representation that is closer to what is expected by people acquainted with regular phasor analysis.

An example of projection plane definition by this method is shown in Figure S2.

**Figure S2:** Definition of the optimal projection plane (defined by the green triangle) for the 3D data shown in Figure S1a (here with less noise added). Two of the reference points are the lifetimes of the bi-exponential decay, and the third one is chosen in between those values on the semicircle, so as to maximize the distance between the orange line and the semicircle.
PCA (principal component analysis) can be used to select the projection axes without user input. In the specific case of phasor data, we need a slight adjustment of the normal PCA algorithm, to make sure that what we plot out of it can be interpreted in the same way as regular phasor plots.

PCA selects as projection axis the one where the largest variance occurs. Essentially, PCA spreads out the data points as much as possible, which is also what we want in full harmonics analysis of phasor data.

In most cases, for PCA to work, we must edit the data before we put it in the algorithm. The real and imaginary value of each harmonic must be centered by subtracting its mean, and be scaled to have identical standard deviations. This avoids PCA giving more weight to a variable based on its scaling, instead of its information content.

For phasor data, we still need the mean correction, but not the scaling with the standard deviation. Each variable has the same amount of noise because the noise is assumed to be white, and thus can be considered to be already scaled. This means that absolute variation in, for example, Re1 or Re10 has equal information content. Utilizing PCA in this way makes sure that the semicircle keeps a reasonable shape, i.e. close to the regular first harmonic phasor plot, thereby facilitating the interpretation of the plot.

Although the projection axes are calculated based on the mean-shifted data, the projection of the data to construct the phasor plot is performed on the original data. This makes sure that the phasor information is contained in the positive x and y region of the plot, as in regular phasor analysis, and the plots are directly comparable.

For tri-exponential data, the projection axes established using PCA will lead to identical results as by using three selected lifetimes, as both methods will select the same plane for the projection (see Figure S3).

For bi-exponential data, the separation of the two underlying mono-exponential positions will still be optimized when using PCA, but the distance to the semicircle will be sub-optimal. This is because the semicircle is not included in the data, and is therefore not considered during PCA. An example of this can be found in Figure S3.
Figure S3: Full-harmonics phasor plots of simulated bi-exponential ((a) and (b)) and tri-exponential ((c) and (d)) data including Gaussian noise. (a) and (c): by selection of three lifetimes. (b) and (d): by using PCA.
For MRI data, full harmonics processing can lead to a considerable improvement over the single harmonic case. However, it does not work for FLIM data. The reason for this is the source of the noise and the mathematical description associated with it. In the case of MRI, noise is Gaussian, which means that each point in the time domain, i.e. each image, has the same amount of noise, irrespective of the position within the image, or the amount of signal. Gaussian noise is white, it has the same intensity at each frequency, or Fourier harmonic, and is uncorrelated, which means that the noise in each harmonic is not influenced by that in the neighboring harmonics. These properties mean that linear combinations of different Fourier harmonics lead to the same average noise level as each individual harmonic. Therefore, a full-harmonics approach that increases the signal in an optimized projection will have the same amount of noise as a single-harmonic analysis: in such condition, the signal-to-noise ratio is truly improved.\(^1\)\(^-\)\(^2\)

For FLIM, the situation is different. About \(50\%\) of the noise in this case is thermal, and can be assumed to have the same properties as MRI noise. The other half of the noise is shot noise, also called Poisson noise\(^3\)\(^-\)\(^4\). Shot noise is caused by the discrete nature of photons. Fluorescence is a statistical process, and the photon count in a specific time window (i.e. bin) has some uncertainty to it. The higher the photon count, the higher the absolute uncertainty (although the relative uncertainty is lower, of course). The absolute noise increases with \(\sqrt{n}\), with \(n\) the number of photons. Since the number of expected photons is different for each bin in the measured data, the average shot noise is different for each data point. Shot noise is white, which would seem to indicate that it has the same properties as Gaussian noise. The difference however is that it is correlated: the noise in neighboring harmonics is often of the same sign. This means that a linear combination of different harmonics of the data increases the apparent noise. This makes full-harmonics phasor processing unsuitable for this type of data. An example of this is shown in Figure S4a-b, where a mono-exponential dataset is simulated, including shot noise but not thermal noise. These Figures show that the size of the phasor cloud, and therefore the uncertainty in the data, increases when using full-harmonics processing. Note too, that the shape of the cloud is elongated, which also indicates that there is a correlation between the \(x\) and \(y\)-axes.

When thermal noise also plays a role, the combined effects of both types of noise should be considered (see Figure S4c-d). The thermal noise in these Figures was set to an identical level as the shot noise, i.e. such that the standard deviation along the \(x\)-axis of the first harmonic plot was increased by a factor \(\sqrt{2}\) compared to the data set with only shot noise. This increase in noise is visible when comparing Figure S4c with Figure S4a. In the full-harmonics view, however, no increase in the standard deviation is observed compared to Figure S4b, indicating that the shot noise is dominating. Hence, in this case no improvement in the unmixing accuracy over the first-harmonic representation is obtained, while in cases where the thermal noise is dominating the full-harmonics representation is expected to aid unmixing.
Figure S4: First-harmonic (left column) and full-harmonics (right column) phasor plots of simulated mono-exponential FLIM data including either only shot noise (top), or shot noise and Gaussian Noise (50:50) (bottom). From (a) to (d), the standard deviations along the x-axis are: 0.04, 0.19, 0.06, 0.19, respectively. The data set consist of 20000 simulations, with 100 time bins and 100 photons per simulation. The decay constant was $9.5 \cdot \Delta t$. 
5 ACCURACY OF UNMIXING AND COMPARISON WITH FITTING

Figure S5: Correlation plots for the two lifetimes (left) and respective amplitudes (right) obtained from bi-exponential fit of the data from Figures 2d-f, performed by using the Levenberg–Marquardt algorithm. The data has been reordered in such a way that Lifetime1 > Lifetime2. The spread in fitted lifetimes is very large (true values 6, 2 and 1 \cdot Δt with Δt = 5 ms). As with the mono-exponential fit, only two populations are separated.

Figure S6: Improvement in the unmixing accuracy of full-harmonics processing of simulated bi-exponential data, visualized as correlation plots of the normalized lifetimes using as intensity log-scale the improvement in unmixing accuracy as compared to 1st harmonic processing. Datasets consist of 10 (a) or 100 (b) simulated time-domain points; the time axis is linear, with steps of Δt. The line x = y (white dots) does not represent data, as no unmixing accuracy can be defined if the two lifetimes are identical. The improvement is highest (lowest) when both lifetimes are short (long), as visible in the lower left (upper right) corner.
Figure S7: MRI brain images obtained by unmixing the dataset into CSF and white matter using full-harmonics phasor (left, same data as in Figure 3) or bi-exponential fitting in the time domain (right). For the global fitting, prior selection of pixels was made based on the phasor plots and using two shared lifetimes for each pixel. This resulted in lifetimes of $706 \pm 5$ and $97.5 \pm 0.1$ ms (in the phasor unmixing, $700$ and $97$ ms were used). Pixels masked prior to the analysis are shown in grey. Pixels that were not selected for fitting are plotted as black, as they can be considered to have a null CSF fraction.

6 ESTIMATION OF PARTICLE PORE SIZE

The average pore size of the alumina pellets in system II was calculated based on the *intra*-pore cyclohexane $T_2$, found to be around 6.1 ms (see Figure 3 in main text). By assuming a diamagnetic surface with a surface relaxivity $\rho_2$ of about $10 \ \mu$m/s$^2$, as this yields correct pore size calculation also for other alumina particles$^6$, and spherical pores, we obtain the pore radius $R$ as: $R = 3 \cdot \rho_2 \cdot T_2 \approx 180$ nm$^5$. We note that this estimation is an upper value, as shorter relaxation times could not be accessed with our shortest available echo time (3.2 ms, see main text). We note also that the correction to this calculation due to bulk cyclohexane relaxation, with $T_2 \approx 0.9$ s, proved negligible.

7 REFERENCES

1. McVeigh, E. R.; Henkelman, R. M.; Bronskill, M. J. Noise and filtration in magnetic resonance imaging. *Med Phys.* 1985, 12 (5), 586-591.
2. Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of nuclear magnetic resonance in one and two dimensions*. Clarendon press Oxford: 1987; Vol. 14.
3. Elder, A.; Schlachter, S.; Kaminski, C. F. Theoretical investigation of the photon efficiency in frequency-domain fluorescence lifetime imaging microscopy. *Opt. Soc. Am. A* 2008, 25 (2), 452-462.
4. Luisier, F.; Vonesch, C.; Blu, T.; Unser, M. In Fast Haar-wavelet denoising of multidimensional fluorescence microscopy data, 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE: 2009; pp 310-313.
5. Foley, I. I.; Farooqui, S. A.; Kleinberg, R. L. Effect of Paramagnetic Ions on NMR Relaxation of Fluids at Solid Surfaces. *J. Magn. Reson. A* 1996, 123 (1), 95-104.
6. Terenzi, C.; Sederman, A. J.; Mantle, M. D.; Gladden, L. F. Spatially-resolved 1H NMR relaxation-exchange measurements in heterogeneous media. *J. Magn. Reson.* 2019, 299, 101-108.