



# Method for assessing the spatiotemporal resolution of structured illumination microscopy (SIM)

ABDERRAHIM BOUALAM AND CHRISTOPHER J. ROWLANDS\*

*Department of Bioengineering, Imperial College London, South Kensington, London, SW7 2BP, UK*

*\*c.rowlands@imperial.ac.uk*

**Abstract:** A method is proposed for assessing the temporal resolution of structured illumination microscopy (SIM), by tracking the amplitude of different spatial frequency components over time, and comparing them to a temporally-oscillating ground-truth. This method is used to gain insight into the performance limits of SIM, along with alternative reconstruction techniques (termed ‘rolling SIM’) that claim to improve temporal resolution. Results show that the temporal resolution of SIM varies considerably between low and high spatial frequencies, and that, despite being used in several high profile papers and commercial microscope software, rolling SIM provides no increase in temporal resolution over conventional SIM.

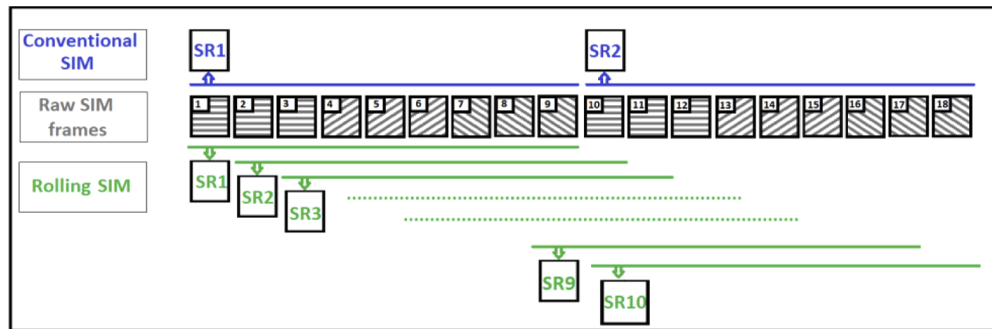
Published by The Optical Society under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

## 1. Introduction

In wide-field microscopy, the sample is illuminated with uniform intensity, and the spatial resolution of the observed fluorescence distribution is limited by the microscope’s pass band – the microscope functions as a low-pass filter, attenuating higher spatial frequencies until they cannot be observed and cannot contribute to the image (see the observable region in Fig. S1). This inability to resolve high-spatial-frequency components is known as the diffraction limit. Super-resolution structured illumination microscopy (SIM) is a widely-used super-resolution optical microscopy technique which enables up to two-fold improvement in lateral resolution [1], over and above this diffraction limit. Instead of illuminating the sample uniformly, a series of interference patterns at different phases and angles is used [1,2], as shown in Fig. 1. This high-spatial-frequency illumination mixes with the spatial frequency content of the fluorophores in the sample, producing sum and difference frequencies which are then low-pass filtered by the microscope’s pass band. High spatial frequencies which were previously beyond the diffraction limit can then be inferred. Since at least three different angles and phases are needed to adequately cover the frequency domain, the reconstruction of a super-resolved image typically requires a data-set of at least nine frames [3,4]. The reconstruction is achieved computationally, using a suitable algorithm [5–7].

During SIM acquisition, the fluorescence emission distribution consists of the sample fluorophore distribution multiplied by the illumination intensity pattern, which in the Fourier domain consists of the convolution of the Fourier transform of the sample fluorophore distribution with three spots (one central spot and two outer spots located symmetrically about the center). It is these outermost spots which ‘alias’ the high-frequency information from beyond the edge of the microscope’s pass band back inside it, allowing its recovery.

Conventional SIM reconstruction is carried out using the successive frames independently, which means that the first SIM image is reconstructed using the first nine frames (1 to 9), the second one using the next nine frames (10 to 18) and so on (see Fig. 1, bottom). One might therefore intuitively expect that, since SIM requires nine frames to reconstruct one super-resolved

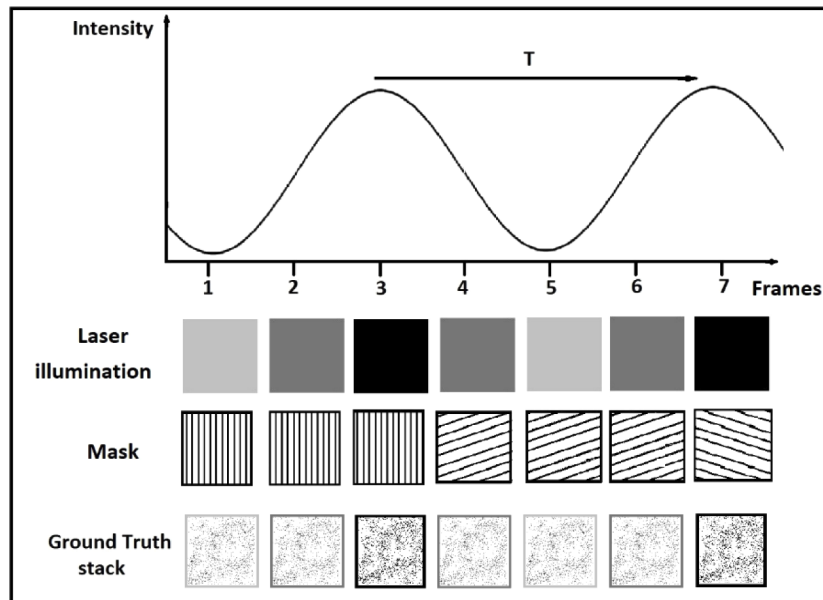


**Fig. 1.** Principle of conventional (top) and rolling SIM (bottom) reconstructions. In conventional SIM reconstruction, the reconstruction of the second super-resolution (SR) image uses frames 10-18 whereas rolling SIM reconstruction uses frames 2-10. With the same data set, rolling SIM can reconstruct eight additional SR images using frames 2-10, 3-11, 4-12 etc. up until 10-18.

image, its temporal resolution should be nine times worse than conventional microscopy. It is for this reason that techniques such as ‘rolling SIM’ reconstruction have been proposed, in which SIM frames are ‘re-used’ in subsequent reconstructions in an attempt to increase the temporal resolution. Using these methods, instead of using frames 10 to 18 to create the next super-resolved image, frames 4-12 or even frames 2-10 are used, creating more super-resolved images using the same data, theoretically increasing the temporal resolution (see Fig. 1, top). This approach has been either proposed or used in several high-profile papers [8–10], and has even been included in commercial software, such as the BURST Mode employed in the software package used on the Zeiss Elyra 7 super-resolution microscope. Nevertheless, while there have been attempts to demonstrate the performance advantages of rolling SIM previously [8], a rigorous assessment of the performance across a wide range of spatial and temporal frequencies has not been performed; this is necessary in order to fairly determine whether rolling SIM truly has the benefits stated.

There has been a consistent interest in the temporal resolution of SIM almost since its inception. In the original SIM implementation of Gustafsson, the striped illumination patterns were generated by moving and rotating a physical grating [3]. However, the imaging speed of this technique was limited to around 1.4 s per reconstructed frame [11]. Several attempts have been made to boost the imaging speed of SIM since then [11–14], notably the implementation of Heintzmann [11], which uses a spatial light modulator (SLM) instead of a physical grating. SLMs are devices which can create arbitrary, computer-controlled illumination patterns by modifying the phase or intensity distribution of the incident light field. The necessary rotations and translations of the illumination pattern can be achieved faster by reprogramming the SLM than by physically rotating or translating a material grating, resulting in this case in a super-resolved frame rate of 7.6 fps per reconstructed 2D slice. Nevertheless, the fastest implementation of SIM to date has been achieved by Curd *et al.* [14,15] in which the reconstruction of the high-resolution SIM images is performed optically instead of using a reconstruction algorithm. This technique can achieve 3D imaging at frame rates that exceed 100fps, which corresponds to a temporal resolution below 10ms. Other attempts have been made to increase the speed of SIM without the need to improve the hardware optical design. Instead, they focus on using new reconstruction algorithms such as the rolling SIM approach mentioned earlier, or by reducing the number of frames required for the reconstruction [16–19]. Notably, Ströhl and Kaminski, building on work from Ingaramo *et al.*, performed SIM reconstruction using a joint Richardson–Lucy (jRL) deconvolution algorithm [16,20,21] which only requires the acquisition of 3 frames (at 3 different stripes orientations, without any phase shift), allowing the temporal resolution of traditional SIM to be doubled [16].

In this paper, the temporal resolution of low and high spatial frequency information in SIM will be assessed, by exploiting the linearity property of the Fourier Transform: any uniform variation in the intensity profile in the real domain should result in the exact same variation in the amplitude/modulus of the whole Fourier spectrum (i.e. this variation should be the same for every spatial frequency). The experiment is performed *in silico* by modulating the intensity of each frame with a known sine wave and taking the normalized root-mean-squared-error between this known sine wave and the temporal variation of each spatial frequency component in the reconstructed super-resolution images. By taking this error over a number of oscillation periods, it is possible to assess the temporal resolution for different spatial frequency components independently (see Fig. 2).



**Fig. 2.** Illustration of how the SIM frames are generated. The simulated laser illumination varies sinusoidally over time, and so therefore does the average intensity of the ground-truth. The resulting SIM frames are then generated by multiplying the ground-truth frames with the intensity-modulated masks.

## 2. Simulation details

A microscopy sample with high spatial frequency support was simulated by randomly assigning 1% of pixels in a blank image to an intensity of 1 (in the case of simulations with no noise) or the stated maximum signal intensity (in the case of simulations with noise). The intensity of this image was modulated over time by an offset sine wave  $G_{T_m} = 0.75 + [\sin(2\pi t/T_m)]/4$  where  $t$  is the time,  $T_m$  is the modulation period, and  $G_{T_m}$  is the ground-truth temporal modulation.  $G_{T_m}$  oscillates between 0.5 and 1, in order to avoid artefacts caused by frames with zero intensity were it to oscillate between 0 and 1. An illustration of this process can be seen in Fig. 2.

Raw SIM frames were created from this stack by multiplying the frames with the required intensity-modulated sinusoidal illumination profile, followed by simulation of a microscope imaging process by convolution with a simulated detection point-spread function (PSF). The simulated PSF was obtained by taking the inverse Fourier Transform of the 2D autocorrelation of a circular aperture, which corresponds to the simulated optical transfer function (OTF) of a microscope [22]. The cutoff frequency of this OTF was set to 128 pixels (which corresponds

to a cutoff period of 4 pixels in the real domain). The period of the sinusoidal illumination was set to 4.75 pixels in the real domain, a value slightly higher than the cutoff period so that the illumination patterns are not filtered by the OTF. At this point, Poisson noise was added if required, by drawing samples from a Poisson distribution for each pixel, with the expected value of the Poisson distribution equal to the pixel value. Finally, high-resolution (HR) SIM images were created following the interleaving reconstruction approach, i.e.: the first HR image was reconstructed using frames 1 to 9, the second using frames 2-10 etc. (note that every ninth frame of this sequence corresponds to a conventional SIM frame). Simulations were repeated with 512 different temporal modulation periods, uniformly distributed between  $T_m=1$  and  $T_m=512$  frames. Image stack simulations and reconstructions were performed using SIMply, a SIM toolbox written in MATLAB. SIMply is an implementation of the original Gustafsson reconstruction method (3), so results are applicable to many, if not all other SIM reconstruction codes. All code used to produce data in this paper is freely available online on the Rowlands Lab website.

### 3. Method for deriving the spatiotemporal resolution of SIM

The spatial variation of light intensity in an image sequence  $H(x, y, t)$  can be decomposed into a finite spectrum of spatial frequencies  $(f_x, f_y)$  by using the Fourier transform, as shown in Eq. (1). In the 2D Fourier domain, each pixel represents a unique spatial frequency with a given orientation and phase, and the modulus of the intensity  $|I(f_x, f_y)|$  is proportional to the contribution of this frequency to the spectrum. The further the pixel is from the centre of the spectrum, the greater the modulus of the spatial frequency it represents. As stated previously, because of the linearity property of the Fourier transform, any uniform variation in the intensity of the image affects the amplitude of the whole Fourier spectrum in the same way (i.e. the intensities of low and high spatial frequencies should experience the same variation). By plotting the magnitude of each spatial frequency and comparing to a ‘ground-truth’ (consisting of a plot of the laser intensity over time) it is possible to assess the temporal resolution, i.e. the ability for the reconstructed image to replicate different temporal frequencies resulting from the externally-modulated signal.

$$\begin{aligned} I(f_x, f_y, t) &= FFT_{2D}(H(x, y, t)) \\ &= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} H(x, y, t) \exp \left[ \frac{-2\pi i \cdot (f_x \cdot x + f_y \cdot y)}{N} \right] \end{aligned} \quad (1)$$

where  $N$  is the number of pixels in each direction. For each modulation period  $T_m$ , we compute the error in the Fourier domain, by taking the difference between the normalized Fourier spectrum magnitude  $\widetilde{I_{T_m}(f_x, f_y, t)} = (I_{T_m} - \overline{I_{T_m}}) / \sigma_{I_{T_m}}$  and the ground-truth  $G_{T_m}$  (see Eq. (2));  $\overline{I_{T_m}}$  and  $\sigma_{I_{T_m}}$  are the mean and the standard deviation respectively of  $I_{T_m}(f_x, f_y, t)$ .

$$Error_{T_m}(f_x, f_y, t) = \widetilde{I_{T_m}(f_x, f_y, t)} - G_{T_m}(t) \quad (2)$$

Afterwards, we compute the root-mean square error  $RMSE_{T_m}(f_x, f_y)$  for each  $f_x$  and  $f_y$  over all  $t$ :

$$RMSE_{T_m}(f_x, f_y) = \sqrt{\frac{\sum_{t=0}^{t=\max} Error_{T_m}(f_x, f_y, t)^2}{n_t}} \quad (3)$$

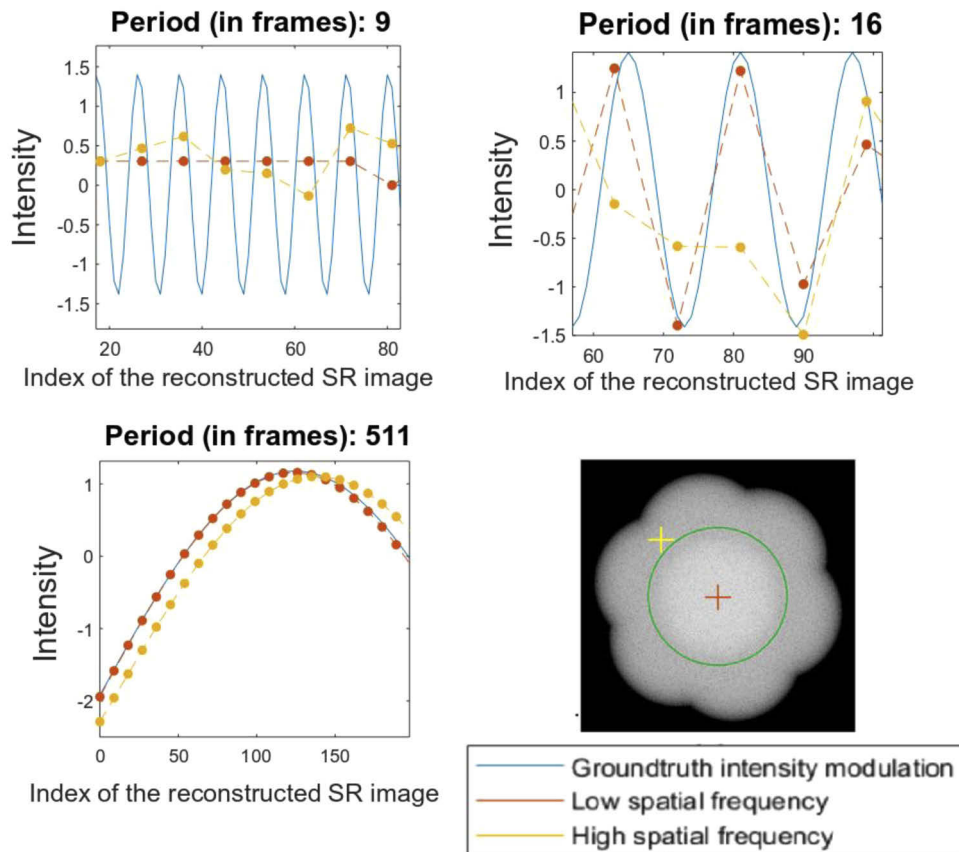
where  $n_t$  is the number of time steps. We repeat this process for every modulation period  $T_m$ . Finally, in order to be able to visualize the RMSE as a function of both the spatial frequency magnitude  $f_s = \sqrt{f_x^2 + f_y^2}$  and the modulation periods, we compute the radial average of the

RMSE for each modulation period  $T_m$ :

$$\overline{RMSE}(T_m, f_s) = \frac{1}{\text{card}(C_r)} \sum_{P \in C_r} RMSE_{T_m}(P) \quad (4)$$

with  $C_r$  being the set of pixels located within the range  $r$  to  $r+1$  from the centre of the image and  $\text{card}(C_r)$  being the cardinality, or number of elements of this set (i.e. number of pixels in the range  $r$  to  $r+1$ ).

Ordinarily, the normalized RMSE should vary between zero (a perfect reproduction of the ground truth temporal modulation) and one (the reconstruction is pure noise and only correlates with the ground truth by chance), but in some cases it can range as high as two, when the temporal modulation is reproduced, but with a phase shift of  $\pi$ . It is for this reason that, in some cases, the RMSE might be considered an overly-strict measure of the ability of a system to resolve a particular temporal frequency due to the sensitivity to lags between  $I_{T_m}(f_x, f_y, t)$  and  $G_{T_m}(t)$  which, in some cases, may not be important. These lags can lead to an apparent increase in the value of the RMSE, even though the temporal profile might appear very similar to the ground-truth (e.g. Figure 3, bottom-left). To suppress this dependency of the RMSE on the time-lag, we also used



**Fig. 3.** Plotting the intensity of a high-frequency versus a low-frequency component for example modulation periods of 9, 16 and 511 frames. Inset bottom right is an illustration of the Fourier spectrum of a reconstructed SIM image, with the locations of the low (red cross) and high (yellow cross) spatial frequencies which are tracked. The green ring demarcates the widefield diffraction limit.



an alternative metric, in which the error is defined not as the difference between  $I_{T_m}(\widetilde{f_x, f_y}, t)$  and  $G_{T_m}(t)$ , but rather assessed by the dot product  $\dot{X}$  between the discrete Fourier transforms in time of these spectra. This permits quantification of their similarities without the effect of phase:

$$\dot{X} = \sum_{t=0}^{N-1} |FFT_{1D}(I_{T_m}(\widetilde{f_x, f_y}, t))|^2 \cdot |FFT_{1D}(G_{T_m}(t))|^2 \quad (5)$$

## 4. Results and discussion

### 4.1. Comparison of the temporal resolution of two spatial frequencies

To illustrate the method for assessing temporal resolution as a function of spatial frequency, two representative low and high spatial frequencies can be plotted, to show how they can achieve different temporal resolutions. For each modulation period, the 2D Fourier Transform of each reconstructed super-resolved image was computed; the normalized temporal intensity profile of an arbitrarily randomly-selected high and low spatial frequency component (see yellow and red crosses in Fig. 3, bottom-right) were plotted against the intensity profile of the ground-truth.

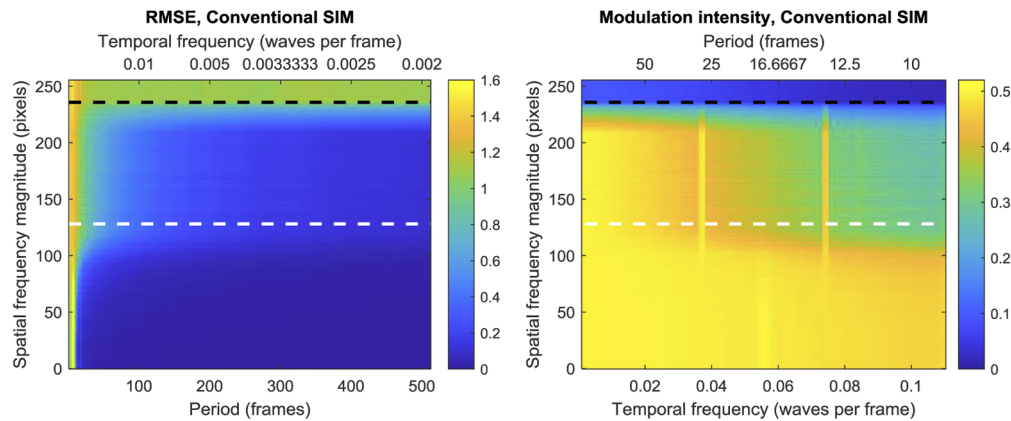
As stated previously, to quantify the differences between those profiles, the normalized root-mean square error (RMSE) between the intensity profile of low and high spatial frequencies and that of the ground-truth were computed. For these isolated examples several things can be observed. Firstly, for very long periods the low frequency pixel seems to track the ground truth accurately, but the high-spatial-frequency appears to have a noticeable ‘phase lag’ – the functional form of the intensity modulation appears correct, but with a small time delay relative to the ground truth (and the low-spatial-frequency component). Conversely, for a modulation with a period of 9 frames, both low-spatial-frequency and high-spatial-frequency results are essentially random. This is consistent with the fact that SIM reconstruction acts as a temporal low-pass filter, and in this case the period of the oscillation is equal to the filter window and therefore cannot be resolved.

The results for a 16 frame temporal period are the most telling; they suggest that the temporal resolution of high spatial frequencies is worse than for low spatial frequencies (a visual representation of this data can be seen in Fig. S2). For any given experiment however, the relevant frequencies will be very different; some experiments require enhancement of the spatial frequencies just above the widefield OTF, whereas others will depend more heavily on the very high spatial frequency information just below the SIM OTF. We must therefore assess the performance fairly, across many spatial frequencies simultaneously.

### 4.2. Estimating temporal resolution as a function of spatial frequency magnitude

To assess all spatial frequencies simultaneously, for each modulation period, an image stack is produced containing the magnitude of the FFT for each spatial frequency. For each frame, the root-mean-squared error between the normalized ground truth and the intensity over time of each spatial frequency is taken. Spatial frequencies with the same modulus are then averaged together (see Fig. S3); the mean RMSE can be plotted as a function of the radius (in pixels) from the origin, and the period of the modulation (see Fig. 4 left). Readers should note that the SIM reconstructions provide a little under 2× increase in spatial resolution due to the 4.75 pixel fringe period mentioned previously; the limit of the SIM OTF is therefore 236 pixels (not 256), corresponding to a 1.92× increase in spatial resolution.

Figure 4 highlights a number of trends. Firstly, it is clear that as the period of the modulation increases, the RMSE is improved for all spatial frequencies. This is largely expected, but even for very large periods, the high-frequency components have a larger error than the low-frequency components, albeit to a lesser degree. This is likely due to the fact that these frequencies are, on average, sampled less regularly than lower spatial frequencies; low spatial frequencies appear in every frame, but super-resolved frequencies by definition lie outside the OTF boundary and are



**Fig. 4.** Left: RMSE between the average intensity of different Fourier spatial frequencies (y axis) of the reconstructed image and the temporal intensity modulation of the raw SIM frames for different modulation periods (x axis). The intensity of different spatial frequency components is averaged along a certain radius in the Fourier domain (y -axis). The bigger the radius, the higher the spatial frequency. The white and black horizontal dashed lines highlight the theoretical diffraction limit and the limit of the SIM OTF respectively. Right: Normalized temporal frequency magnitude at the temporal oscillation frequency, as a function of temporal frequency and spatial frequency magnitude. Note that the X axis is in monotonically-increasing units of temporal frequency, because the temporal periods are constrained to an even integer divisor of 512 (the number of simulated timepoints).

only sampled when the required SIM pattern aliases them into the pass-band of the microscope. This reflects another feature of the data, which is that for all periods, the RMSE is worse for higher spatial frequencies, with a transition approximately where the spatial frequency magnitude corresponds to the period of the illuminating SIM fringes.

As noted previously, the use of normalized RMSE as a metric for assessing temporal resolution could be overly-strict in some cases, and an alternative metric which just measures whether a temporal frequency modulus can be resolved (but not necessarily with the correct phase) may be acceptable for certain niche applications. Nevertheless, the results (see Fig. 4, right) show that, even under this more generous condition, the conclusions are broadly similar. The temporal resolution at low temporal modulation frequencies is preserved for both low and high spatial frequencies, but as the temporal frequency increases, high spatial frequencies degrade more rapidly than low spatial frequencies. This demonstrates that any ‘phase lag’ of the reconstructed temporal profiles explains only a small part (if any) of the observed difference between the temporal resolution of high and low spatial frequencies.

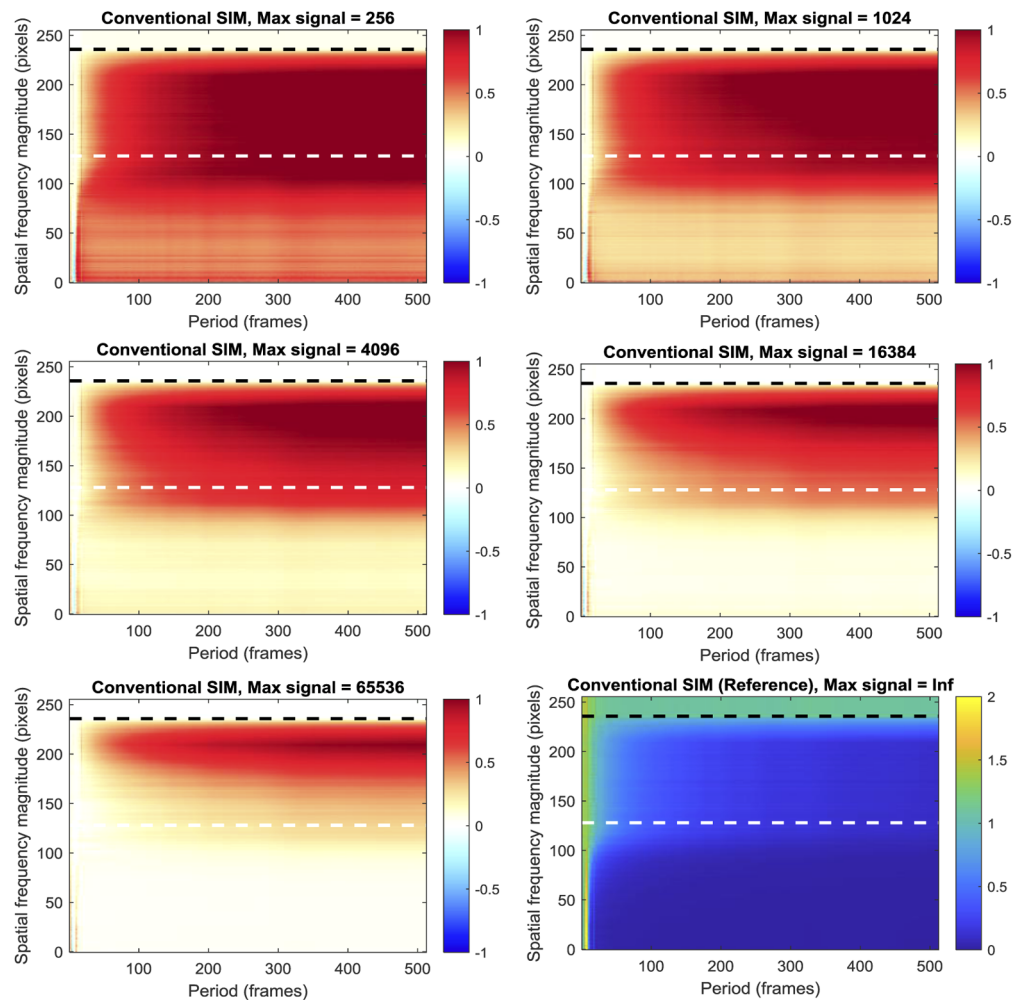
While the data supports the conclusion that SIM reconstruction performance is degraded for low temporal periods and high spatial frequencies, the reader must decide for themselves whether their application is sensitive to these phenomena.

#### 4.3. Effect of noise

While it is enlightening to examine noise-free SIM reconstruction, in reality all experiments have noise, and it is important to consider whether this affects the results obtained above. For this analysis, a purely Poisson noise model is used; this represents a best-case scenario, in which camera noise and any excess amplification noise is negligible. The noisy data is simulated by setting the intensity of the isolated point-sources to a specified value of ‘Max signal’ *before* applying the SIM illumination pattern and blurring using the simulated microscope OTF; the

noise is then incorporated by drawing values from a Poisson distribution for each pixel in the resulting blurred image. The data are then reconstructed in an identical manner to before.

The results (seen in Fig. 5) show that the predominant effect of noise is to degrade the performance of the SIM reconstruction at longer periods; performance at shorter periods is already quite poor, so the additional noise does not degrade the reconstruction to the same degree. More notable however is a clear difference between the low ('widefield') spatial frequencies and the high ('super-resolved') spatial frequencies. Low spatial frequency performance is degraded by a broadly constant value, regardless of the period of the temporal oscillation, but the high spatial frequencies are reconstructed proportionally poorly at longer temporal oscillation periods, compared to the reference (zero noise) reconstruction. Notably, this phenomenon doesn't affect all 'super-resolved' spatial frequencies equally (unlike the 'widefield' spatial frequencies) – higher 'super-resolved' spatial frequencies are degraded proportionately more.



**Fig. 5.** Noise sensitivity of SIM reconstruction, as a function of spatial frequency magnitude and temporal oscillation period. All data presented as RMSE relative to the noise-free reference presented as RMSE relative to the noise-free reference image, bottom right. Note the different scale on the color bars for the reference data relative to the noise data. The white and black horizontal dashed lines highlight the theoretical diffraction limit and the limit of the SIM OTF respectively.



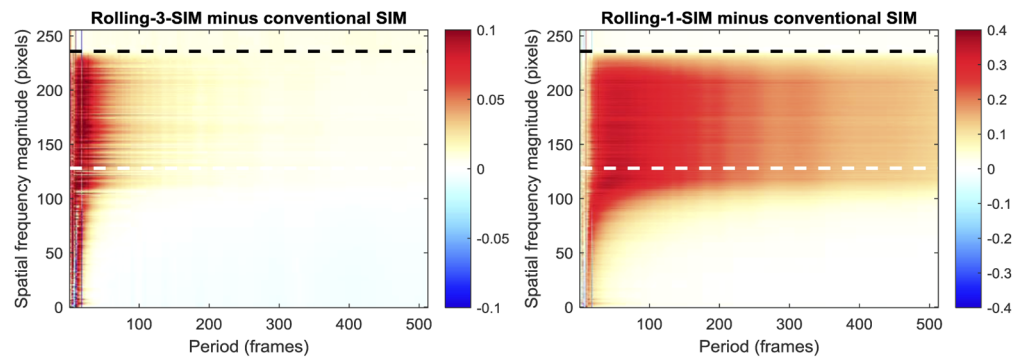
This has implications for SIM imaging of dynamic events. Good super-resolution imaging normally requires bright images, but as the temporal frequency content of the dataset increases, the benefit of a high signal-to-noise ratio is lost; simultaneous reconstruction of both high spatial and high temporal frequencies is poor regardless of signal-to-noise ratio. Counterintuitively, for moderate temporal frequencies, the photon budget of the system may mean that better temporal resolution can actually be achieved by imaging *more slowly*, but with better signal-to-noise ratio.

#### 4.4. Effectiveness of Rolling SIM reconstruction

As noted previously, one approach to increasing the temporal resolution of SIM is to perform a ‘rolling’ reconstruction, in which data from previous super-resolved frames are ‘re-used’. Nevertheless, it is important to assess whether this approach truly increases the temporal resolution. Importantly, it is logical to assert that, in order to be considered successful, rolling SIM should increase the temporal resolution of the *super-resolved* spatial frequency components (because if conventional widefield resolution were acceptable for a given application, there would be little point in using SIM).

Two versions of rolling SIM will be explored; in ‘rolling-3-SIM’, a SIM reconstruction is performed every three frames, ensuring that all frames with the same angle but different phases are taken sequentially (i.e. frames 1-9, 4-12, 7-15 and so on, but not frames 2-10). In ‘rolling-1-SIM’ the SIM reconstruction is performed every frame, accepting that in two thirds of cases, one set of three phases will be split between the last and first frames of the reconstructed set. One might therefore expect that the temporal resolution of rolling-1-SIM would be prone to temporal artefacts.

Figure 6 highlights the differences between the RMSE values obtained with rolling SIM and conventional SIM. It should be noted that, because rolling SIM attempts to reconstruct additional datapoints between those that would normally be taken by conventional SIM, there are necessarily more comparisons between the reconstructed frames and the ground truth for rolling SIM than conventional SIM. Since the errors are averaged together, this should not have a significant effect given a sufficiently large number of reconstructed frames, but readers should be aware that the number of comparisons is not the same for each algorithm.



**Fig. 6.** Difference between the temporal resolution of Rolling-3-SIM and conventional SIM, and of Rolling-1-SIM and conventional SIM, obtained by subtracting the RMSE for conventional SIM from that of Rolling SIM. The white and black horizontal dashed lines highlight the theoretical diffraction limit and the limit of the SIM OTF respectively.

For slowly-varying temporal frequencies, the performance of rolling-3-SIM shows a subtle advantage at lower spatial frequencies, but an equally subtle disadvantage at higher spatial frequencies. As stated above, because the super-resolved spatial frequencies are the important ones, this data suggests that any performance advantage over conventional SIM reconstruction

is, at best, negligible for these temporal periods. At low temporal periods however (where improved temporal resolution is most necessary) there is a pronounced degradation in the RMSE performance, leading to the conclusion that rolling-3-SIM should not be employed in order to increase a SIM microscope's temporal resolution.

As an aside, this result may seem counterintuitive. Since rolling-3-SIM consists of three interleaved conventional SIM reconstructions, it is reasonable to expect that its performance should be no worse than conventional SIM. This assertion ignores a subtle assumption, which is that the mean brightness of the interleaved SIM reconstructions is the same for all three; this is not guaranteed to be true, and violation of this assumption can lead to poor RMSE performance. In conventional SIM, when the normalization is applied, the average brightness will be compensated, but when rolling-3-SIM is performed, the normalization might cause the first interleaved reconstruction to be too dim and the second to be too bright, for example. The result is a poor RMSE. Unfortunately, this cannot be compensated for in a real SIM application, since unlike in these tests, the underlying temporal profile is unknown. A user might observe a pronounced high-temporal-frequency signal, but without further knowledge of the true underlying temporal profile, there would be no way to distinguish the artefact from a real signal.

As predicted, the performance of rolling-1-SIM is clearly worse than rolling-3-SIM, most likely due to reconstruction artefacts, owing to the previously-identified split phases in two-thirds of the reconstructions. Notably, there does not even appear to be a small benefit at low spatial frequencies (in the manner of rolling-3-SIM), which means this reconstruction method is not recommended even in the rare cases when the low spatial frequencies are of interest in a super-resolution image.

Readers should note that the colour ranges of Fig. 6, *left* and Fig. 6, *right* are not the same and that the performance difference between rolling-1-SIM and rolling-3-SIM is more pronounced than may appear. For readers who would like to compare the two plots on the same colour axes, Fig. S4 is another version of Fig. 6 where the colour ranges are set to the range -0.4 to 0.4.

## 5. Conclusion

The effective spatial resolution and information content achievable by SIM and other super-resolution techniques have been subject to intense scrutiny in the last decade [23–25]. Though super-resolution methods open the door to the investigation of biological processes with unprecedented detail, they also significantly increase the technical complexity of microscopes and image reconstruction techniques. Hence, in order to acquire and interpret high-quality super-resolution data, it is essential to understand the specific limitations and distortions imposed by a given microscopy technique. Every microscope, even the simplest one, performs a spatial mapping of the sample structure onto an image, and biological interpretation relies on understanding this mapping. In this work, we addressed this mapping from a temporal perspective, on which interpreting biological significance relies just as much.

In summary, a method for assessing the spatiotemporal resolution of optical super-resolution microscopy has been developed and has been used to investigate the spatiotemporal resolution properties of Structured Illumination Microscopy. Notably, it was found that the temporal resolution of the system is not the same for both the low and high spatial frequencies; the low-frequency information (that would ordinarily be gathered by a widefield microscope) has superior (and more consistent) temporal resolution, whereas the super-resolved spatial frequencies are much more prone to temporal artefacts. It is important to be clear that, because of its higher sampling rate, rolling-3-SIM *does* allow the tracking of intensity changes in wide-field features at a higher temporal resolution than conventional SIM. However, super resolved features clearly do not benefit from this enhancement in temporal resolution. The effect of noise on temporal resolution is very pronounced, and very high signal-to-noise ratios are needed to resolve temporal periods comparable with the 9-frame period of a SIM measurement. Furthermore, it was found that the

interleaved reconstruction of rolling SIM does not support any increase in temporal resolution of the super-resolved information, and in fact results in a small (but largely insignificant in the case of rolling-3-SIM) reduction in temporal resolution for these spatial frequency components. Alternative methods are therefore recommended for increasing temporal resolution, including boosting the imaging speed [11–16], as well as reduction of the number of raw SIM frames to be acquired [16,20,21].

**Funding.** Imperial College London, Excellence Fund for Frontier Research; Wellcome Trust (212490/Z/18/Z); Engineering and Physical Sciences Research Council (EP/S016538/1); Biotechnology and Biological Sciences Research Council (BB/T011947/1).

**Acknowledgements.** CJR acknowledges support from the Wellcome Trust (212490/Z/18/Z), EPSRC (EP/S016538/1), BBSRC (BB/T011947/1) and the Imperial College Excellence Fund for Frontier Research. AB is grateful for a PhD studentship from the Imperial College Department of Bioengineering. Both authors are indebted to Debora Machado Andrade Schubert for proofreading and commenting on the manuscript.

**Disclosures.** The authors declare no conflicts of interest.

See [Supplement 1](#) for supporting content.

## References

1. J. -A. Conchello, M. G. L. Gustafsson, D. A. Agard, J. W. Sedat, C. J. Cogswell, A. G. Tescher, and T. Wilson, “Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination,” *Three-Dimensional and Multidimensional Microscopy: Image Acquisition Processing VII 2000*, p. 141–150. (2000)
2. Y. Wu and H. Shroff, “Faster, sharper, and deeper: structured illumination microscopy for biological imaging,” *Nat. Methods* **15**(12), 1011–1019 (2018).
3. M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *J. Microsc.* **198**(2), 82–87 (2000).
4. J. Demmerle, C. Innocent, A. J. North, G. Ball, M. Muller, E. Miron, A. Matsuda, I. M. Dobbie, Y. Markaki, and L. Schermelleh, “Strategic and practical guidelines for successful structured illumination microscopy,” *Nat. Protoc.* **12**(5), 988–1010 (2017).
5. A. Lal, C. Y. Shan, and P. Xi, “Structured illumination microscopy image reconstruction algorithm,” *IEEE J. Sel. Top. Quant.* **12**, 6803414 (2016).
6. M. Muller, V. Monkemoller, S. Hennig, W. Hubner, and T. Huser, “IMAGE J Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ,” *Nat. Commun.* **7**(1), 10980 (2016).
7. V. Perez, B. J. Chang, and E. H. Stelzer, “Optimal 2D-SIM reconstruction by two filtering steps with Richardson-Lucy deconvolution,” *Sci. Rep.* **6**(1), 37149 (2016).
8. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, and L. Chen, “Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy,” *Nat. Biotechnol.* **36**(5):451 (2018)
9. Y. Ma, D. Li, Z. J. Smith, D. Li, and K. Chu, “Structured illumination microscopy with interleaved reconstruction (SIMILR),” *J. Biophotonics* **11**(2), e201700090 (2018).
10. Y. Guo, D. Li, S. Zhang, Y. Yang, J. J. Liu, X. Wang, C. Liu, D. E. Milkie, R. P. Moore, U. S. Tulu, D. P. Kiehart, J. Hu, J. Lippincott-Schwartz, E. Betzig, and D. Li, “Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales,” *Cell* **175**(5), 1430–1442 (2018).
11. R. Forster, H. W. Lu-Walther, A. Jost, M. Kielhorn, K. Wicker, and R. Heintzmann, “Simple structured illumination microscope setup with high acquisition speed by using a spatial light modulator,” *Opt. Express* **22**(17), 20663 (2014).
12. D. Dan, M. Lei, B. Yao, W. Wang, M. Winterhalder, A. Zumbusch, Y. Qi, L. Xia, S. Yan, Y. Yang, P. Gao, T. Ye, and W. Zhao, “DMD-based LED-illumination super-resolution and optical sectioning microscopy,” *Sci. Rep.* **3**(1), 1116 (2013).
13. A. Markwirth, M. Lachetta, V. Monkemoller, R. Heintzmann, W. Hubner, T. Huser, and M. Muller, “Video-rate multi-color structured illumination microscopy with simultaneous real-time reconstruction,” *Nat. Commun.* **10**(1), 4315 (2019).
14. A. Curd, A. Cleasby, K. Makowska, A. York, H. Shroff, and M. Peckham, “Construction of an instant structured illumination microscope,” *Methods* **88**, 37–47 (2015).
15. F. Ströhl and C. F. Kaminski, “Frontiers in structured illumination microscopy,” *Optica* **3**(6), 667 (2016).
16. F. Ströhl and C. F. Kaminski, “Speed limits of structured illumination microscopy,” *Opt. Lett.* **42**(13), 2511 (2017).
17. J. T. Frohn, H. F. Knapp, and A. Stemmer, “True optical resolution beyond the Rayleigh limit achieved by standing wave illumination,” *Proc. Natl. Acad. Sci. U. S. A.* **97**(13), 7232–7236 (2000).
18. R. Heintzmann, “Saturated patterned excitation microscopy with two-dimensional excitation patterns,” *Micron* **34**(6-7), 283–291 (2003).
19. F. Orieux, E. Sepulveda, V. Lorient, B. Dubertret, and J. C. Olivo-Marin, “Bayesian estimation for optimized structured illumination microscopy,” *IEEE Trans Image Process.* **21**(2), 601–614 (2012).

20. M. Ingaramo, A. G. York, E. Hoogendoorn, M. Postma, H. Shroff, and G. H. Patterson, "Richardson-Lucy deconvolution as a general tool for combining images with complementary strengths," *ChemPhysChem* **15**(4), 794–800 (2014).
21. F. Ströhl and C. F. Kaminski, "A joint Richardson-Lucy deconvolution algorithm for the reconstruction of multifocal structured illumination microscopy data," *Methods Appl. Fluoresc.* **3**(1), 014002 (2015).
22. G. D. Boreman, *Modulation Transfer Function in Optical and Electro-optical Systems* (Society of Photo Optical, 2001)
23. D Li, L Shao, BC Chen, X Zhang, M Zhang, B Moses, DE Milkie, JR Beach, JA Hammer 3rd, M Pasham, T Kirchhausen, MA Baird, MW Davidson, P Xu, and E Betzig, "Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics," *Science* **349**(6251), aab3500 (2015).
24. S. J. Sahl, F. Balzarotti, J. Keller-Findeisen, M. Leutenegger, V. Westphal, A. Egner, F. Lavoie-Cardinal, A. Chmyrov, T. Grotjohann, and S. Jakobs, "Comment on "Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics","" *Science* **352**(6285), 527 (2016).
25. R. P. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grunwald, S. Stallinga, and B. Rieger, "Measuring image resolution in optical nanoscopy," *Nat. Methods* **10**(6), 557–562 (2013).