

MP-PCA denoising of fMRI time-series data can lead to artificial activation “spreading”

Francisca F. Fernandes^a, Jonas L. Olesen^{b,c}, Sune N. Jespersen^{b,c}, Noam Shemesh^{a,*}

^a Champalimaud Research, Champalimaud Foundation, Lisbon, Portugal

^b Center of Functionally Integrative Neuroscience (CFIN) and MINDLab, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

^c Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark



ARTICLE INFO

Keywords:

MP-PCA denoising
preclinical
fMRI

ABSTRACT

MP-PCA denoising has become the method of choice for denoising MRI data since it provides an objective threshold to separate the signal components from unwanted thermal noise components. In rodents, thermal noise in the coils is an important source of noise that can reduce the accuracy of activation mapping in fMRI. Further confounding this problem, vendor data often contains zero-filling and other post-processing steps that may violate MP-PCA assumptions. Here, we develop an approach to denoise vendor data and assess activation “spreading” caused by MP-PCA denoising in rodent task-based fMRI data. Data was obtained from $N = 3$ mice using conventional multislice and ultrafast fMRI acquisitions (1 s and 50 ms temporal resolution, respectively), using a visual stimulation paradigm. MP-PCA denoising produced SNR gains of 64% and 39%, and Fourier Spectral Amplitude (FSA) increases in BOLD maps of 9% and 7% for multislice and ultrafast data, respectively, when using a small [2 2] denoising window. Larger windows provided higher SNR and FSA gains with increased spatial extent of activation that may or may not represent real activation. Simulations showed that MP-PCA denoising can incur activation “spreading” with increased false positive rate and smoother functional maps due to local “bleeding” of principal components, and that the optimal denoising window for improved specificity of functional mapping, based on Dice score calculations, depends on the data’s tSNR and functional CNR. This “spreading” effect applies also to another recently proposed low-rank denoising method (NORDIC), although to a lesser degree. Our results bode well for enhancing spatial and/or temporal resolution in future fMRI work, while taking into account the sensitivity/specification trade-offs of low-rank denoising methods.

1. Introduction

Preclinical functional MRI (fMRI) offers unique opportunities for bridging noninvasive human fMRI findings with their underlying biological origins (Jonckers et al., 2015; Zerbi, 2022; Gozzi and Zerbi, 2023). Great advantages are offered by the rodents’ genetic flexibility in terms of disease models, cell-type specificity (e.g. optogenetics (Lim et al., 2013)), and by the ability to validate noninvasive MRI metrics with gold-standard methods such as histology, electrophysiology, or multimodal simultaneous recordings (Tian et al., 2009) and imaging (Dana et al., 2019). Ultrafast fMRI acquisitions are emerging as novel means for better characterization of fMRI responses and for better representation of neural information flow (Yu et al., 2014; Lee et al., 2019; Gil et al., 2021; Jung et al., 2021). However, the demand for ever higher spatiotemporal resolution necessarily involves a decrease in signal-to-noise ratio (SNR), which in turn hinders the accuracy and precision of functional activation estimation. Indeed, noise (Aja-Fernández and Vegas-

Sánchez-Ferrero, 2016; St-Jean et al., 2020), as well as the relatively low contrast-to-noise ratio (CNR) (Welvaert and Rosseel, 2013) in fMRI (Hirano et al., 2011), is one of the main factors that limits the spatial and/or temporal resolution of MRI data (Liu, 2016). Even though this can be in part ameliorated by using cryogenic coils (Gil et al., 2021; Ratering et al., 2008; Nunes et al., 2021), data denoising in preclinical fMRI remains critical.

Different denoising approaches have been developed throughout the years for fMRI. Total variation minimization (Rudin et al., 1992; Knoll et al., 2011) and non-local means (Coupe et al., 2008; Manjón et al., 2008; Manjón et al., 2010; Kafali et al., 2018) have proven useful to reduce noise in different MRI applications. For fMRI in particular, spatial and temporal filtering (Smith and Brady, 1997; Triantafyllou et al., 2006; Molloy et al., 2014) and the addition of noise regressors to general linear model (GLM) fitting (GLMdenoise) (Kay et al., 2013) have enabled improved activation mapping. However, despite the usefulness of these techniques, they can be limited

* Corresponding author at: Champalimaud Research, Champalimaud Foundation, Av. Brasilia 1400-038, Lisbon, Portugal.
E-mail address: noam.shemesh@neuro.fchampalimaud.org (N. Shemesh).

either by a loss in spatial resolution (smoothing, blurring of anatomical details and introduction of partial volume effects) that leads to compromised accuracy in further quantitative analyses, or by the need of subjective user input, which can limit the reproducibility between operators.

Principal component analysis (PCA)-based approaches (Pai et al., 2011; Manjón et al., 2013; Manjón et al., 2015; Sonderer and Chen, 2018) have been proposed to address these issues. Most of the signal-related variance in redundant data is typically contained in just a few principal components while the noise is evenly spread over all components. This allows the reduction of thermal noise without requiring training data or accurate brain region segmentation. Although these approaches have exhibited very good potential for increasing BOLD CNR in fMRI data (Behzadi et al., 2007; Caballero-Gaudes and Reynolds, 2017), they require a subjective and often empirically defined (Manjón et al., 2013) user-input threshold for determining which principal components to keep, and which components to reject as “noise”. The lack of objective methods for defining these thresholds can lead to inconsistency between different sites, and, if wrong thresholds are applied, important signal components can be removed.

Marchenko-Pastur PCA (MP-PCA) denoising was recently proposed to overcome these limitations (J. Veraart et al., 2016; J. Veraart et al., 2016). Using random matrix theory principles developed early on by Marchenko and Pastur (Marčenko and Pastur, 1967), MP-PCA provides an objective data-driven (model-independent) threshold to identify noise components based on asymptotic mathematical properties of the eigenspectrum of random covariance matrices. In the large matrix size limit, the noise contribution to the histogram of PCA eigenvalues universally follows the MP distribution (Marčenko and Pastur, 1967). Therefore, by leveraging MRI data redundancy, MP-PCA denoising is able to estimate the local noise level and target its suppression in a deterministic way. Moreover, MP-PCA is a selective denoising technique, as it removes – ideally – only thermal noise and not any other structured non-white noise that arises from respiratory, cardiac or spontaneous neuronal activity (physiological noise), motion, or artifacts, thereby preserving fine anatomical detail and even enabling a better identification of such nuisance components. Currently, MP-PCA is the most widespread method for noise suppression in diffusion MRI (dMRI) (J. Veraart et al., 2016; Adanyeguh et al., 2018; McKinnon et al., 2018; Grussu et al., 2020) and it has also been recently successfully applied to MRI relaxometry data (Grussu et al., 2020; Does et al., 2019) and Magnetic Resonance Spectroscopy (Mosso et al., 2022). MP-PCA denoising was recently also proposed for fMRI, both for task-based fMRI (language mapping in humans (Ades-Aron et al., 2021)) as well as for enhanced temporal SNR (tSNR) and network mapping in human and rat resting-state fMRI data (Adhikari et al., 2019; Diao et al., 2021).

The original MP-PCA work assumes infinite matrices and Gaussian noise. In the more practical finite matrix case, deviations from the asymptotic MP-PCA expression may incur “tails” near the right edge of the MP distribution. As a way to solve this problem, the NOise Reduction with DIstribution Corrected (NORDIC) PCA approach was recently developed, initially for dMRI (Moeller et al., 2021) and more recently for fMRI (Vizioli et al., 2021; Dowdle et al., 2023), showing improvements in key metrics of functional mapping while avoiding image blurring or decrease of spatial precision. Like the MP-PCA denoising method, NORDIC is a low-rank patch-based denoising method that works directly on the spectrum of the Casorati matrix with the purpose of suppressing thermal noise. However, NORDIC numerically estimates the threshold via a Monte Carlo simulation with finite random noise matrices, whose variance matches the experimentally measured thermal noise, to generate the sample average for the largest PCA eigenvalue. Moreover, NORDIC performs signal and noise scaling to ensure a zero-mean and spatially identical noise prior to denoising. A further complication can arise from images reconstructed by the vendor, which may not be ideally suited for the immediate application of MP-PCA denoising. In particular, partial Fourier reconstructions (Cuppen et al., 1986) and acquisi-

tion along gradient ramps (e.g., EPI) commonly used in fMRI to achieve faster acquisition rates, require extrapolation or zero-filling (ZF) and/or regridding in the phase encoding and readout directions of k-space, respectively, which generate spatial correlations that change the noise statistics, and violate the i.i.d. noise characteristics on which MP-PCA assumptions rely. Lastly, the projection to a local low-rank approximation of the data upon MP-PCA denoising may involve local “bleeding” of active signal components that result in “spreading” of fMRI activation patterns.

Here, we develop an approach to denoise preclinical fMRI data based on vendor reconstructions (that include reconstruction steps that correlate noise between voxels) and more generally assess the spatial specificity upon MP-PCA denoising in preclinical fMRI acquisitions in terms of activation extent and potential “spreading”. Indeed, MP-PCA has not yet been explored in preclinical task-based fMRI, which may represent a missed opportunity given the predominance of thermal noise in rodent coils (Ratering et al., 2008). To that end, we acquired conventional multislice “slow” and more challenging ultrafast (temporal resolution = 50 ms) fMRI data from the mouse brain while performing a visual stimulation task eliciting strong neuronal and BOLD responses (Grubb and Thompson, 2003; Niell and Stryker, 2008; Wang et al., 2010; Nirajan et al., 2016; Nirajan et al., 2017; Dinh et al., 2021). Moreover, we use a more permissive MP-PCA threshold than in the original work (J. Veraart et al., 2016), that takes into account the tails at the edge of the MP distribution arising from the finite matrix sizes, thus enabling the preservation of more signal components. We validate the performance using simulations with a known ground truth (GT), and finally, we extend our conclusions to the more recent NORDIC PCA denoising approach. Our findings identify denoising-driven artifactual activation spreading and provide means to avoid it based on SNR and CNR, which can be important for future studies seeking to harness denoising in fMRI.

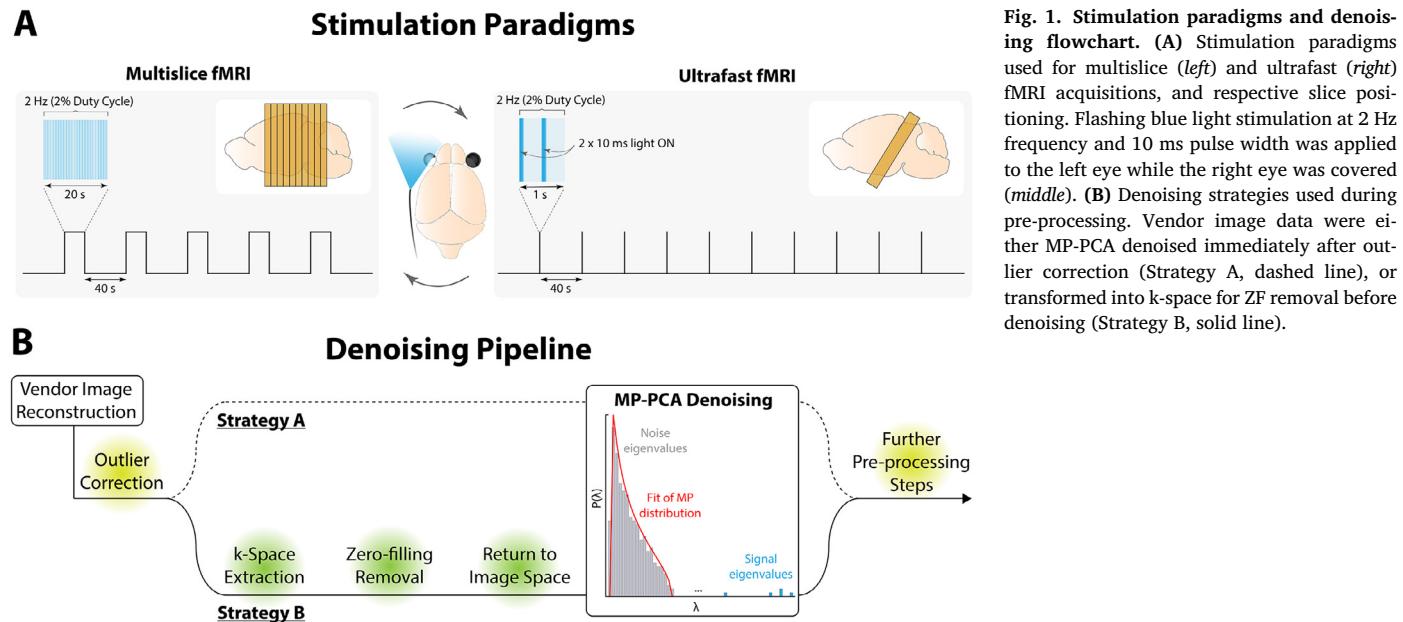
2. Methods

All animal experiments were preapproved by the institutional and national authorities and carried out in accordance to European Directive 2010/63.

2.1. Animal preparation

Adult C57BL/6 male mice ($N = 3$) weighing 25 ± 1 g and aged 8.0 ± 0.3 weeks were used in this study (animal weights and ages are reported as mean \pm SD). Animals were reared in a temperature-controlled room and held under a 12 h/12 h light dark regime with ad libitum access to food and water. Anesthesia was induced with 5% isoflurane (Vetflurane®, Virbac, France) mixed with oxygen-enriched (28%) medical air. Animals were then weighed, moved to the animal bed (Bruker BioSpin, Germany) and isoflurane was reduced to 2–3%. Ear and bite bars were fixed and eye drops (Bepanthen® Eye Drops, Bayer AG, Germany) were applied to prevent eye dryness. A 0.4 mg/kg subcutaneous bolus of medetomidine (Dormilan®, Vetpharma Animal Health S.L., Spain; 1 mg/ml, diluted 1:10 in saline) was injected ~6.5 min after isoflurane induction, followed by a gradual discontinuation of isoflurane to 0% over the next 10 min. After this time period, a continuous infusion of medetomidine (0.8 mg/kg/h) was initiated and maintained until the end of the experiment (Adamczak et al., 2010).

Breathing rate and rectal temperature were monitored throughout the sessions using a pillow sensor and an optic fiber probe (SA Instruments, Inc., USA), respectively. A warm-water recirculating pad was used for body temperature control (maintaining temperatures between 35.0–36.8 °C). At the end of experiments, a subcutaneous injection of atipamezole (Antidorm, Vetpharma Animal Health S.L., Spain; 5 mg/ml, diluted 1:10 in saline) at 2 mg/kg was given to all animals to reverse the medetomidine effects.



2.2. MRI protocol

Animals were imaged using a 9.4 T BioSpec MRI scanner (Bruker, Karlsruhe, Germany) equipped with an AVANCE III HD console, producing isotropic pulsed field gradients of up to 660 mT/m (120 μ s rise time), an 86 mm-ID quadrature resonator for RF transmission and a 10 mm loop surface coil for signal reception, thus avoiding coil-driven spatial correlation of the noise due to multiple-channel receiver coil (e.g. multichannel cryoprobe) combinations. The scanner runs ParaVision 6.0.1 software (Bruker, Ettlingen, Germany).

Following routine adjustments for center frequency, RF calibration, acquisition of B_0 maps and automatic shimming, anatomical images were acquired using a T_2 -weighted Turbo RARE sequence in the coronal, sagittal and oblique planes (TR/TE = 2000/40 ms, FOV = 20×16 mm², in-plane resolution = 80×80 μ m², RARE factor = 5, slice thickness = 0.5 mm, number of slices = 20 (if coronal or oblique) or 23 (if sagittal), $t_{\text{acq}} = 1 \text{ min } 18 \text{ s}$). An extra acquisition with 3 oblique slices (TR/TE = 2000/40 ms, FOV = 16×12.35 mm², in-plane resolution = 64×62 μ m², RARE factor = 5, slice thickness = 1 mm, $t_{\text{acq}} = 1 \text{ min } 18 \text{ s}$) was also performed for later coregistration of ultrafast fMRI scans.

Functional imaging began ~40 min after isoflurane induction (Pradier et al., 2021). Conventional multislice fMRI acquisitions were performed using a GE-EPI sequence: TR/TE = 1000/15 ms, flip angle = 60°, FOV = 16×12 mm², in-plane resolution = 145×145 μ m², partial Fourier factor in the PE direction = 0.8, slice thickness = 0.5 mm, number of slices = 10, repetitions = 340, $t_{\text{acq}} = 5 \text{ min } 40 \text{ s}$. Ultrafast fMRI acquisitions with a single oblique slice capturing the entire visual pathway (Gil et al., 2021) (GE-EPI, TR/TE = 50/17.5 ms, flip angle = 15°, FOV = 16×12.35 mm², resolution = 167×167 μ m², partial Fourier factor along the PE direction = 0.8, slice thickness = 1 mm, repetitions = 9000, $t_{\text{acq}} = 7 \text{ min } 30 \text{ s}$, dummy scans = 5240 to ensure the coil reaches a steady temperature during these intensive acquisitions) were also performed.

2.3. Visual stimulation

As depicted in Fig. 1A, left eye monocular flashing stimulation (frequency = 2 Hz; pulse width = 10 ms) was performed using a blue (wavelength = 470 nm) LED connected to an optical fiber that extended to the interior of the scanner (at ~1 cm from the eye). The right eye was covered with a patch. The stimulation paradigms consisted of five blocks of 20 s stimulation and 40 s rest for multislice fMRI, and ten cycles of 1 s

Fig. 1. Stimulation paradigms and denoising flowchart. (A) Stimulation paradigms used for multislice (left) and ultrafast (right) fMRI acquisitions, and respective slice positioning. Flashing blue light stimulation at 2 Hz frequency and 10 ms pulse width was applied to the left eye while the right eye was covered (middle). (B) Denoising strategies used during pre-processing. Vendor image data were either MP-PCA denoised immediately after outlier correction (Strategy A, dashed line), or transformed into k-space for ZF removal before denoising (Strategy B, solid line).

stimulation and 40 s rest for ultrafast fMRI. Both acquisitions were alternated and repeated 2–4 times per animal, and separated by a resting period of at least 6 min to avoid habituation.

2.4. Data analysis

All datasets were analyzed in MATLAB® (MathWorks, USA).

Data was first corrected for outliers by manually selecting time points whose average brain signal strongly deviated ($\pm 2-3$ SD) from its 2nd order polynomial trend, and estimating new voxel values at those time points using piecewise cubic interpolation from the signal at the remaining time points. This correction was performed independently for each slice. Only < 0.2% of datapoints/scan were corrected in multislice fMRI data, whereas no outliers were found in ultrafast data.

As shown in the diagram of Fig. 1B, after outlier correction, data was denoised using two different strategies. In Strategy A, data was MP-PCA denoised in its magnitude form using 15 different spatial sliding windows in case of multislice fMRI (size varying between [2 2], [5 5], [10 10], [20 20] and [40 40] in the row and columns dimensions, and between 1, 5 and 10 in the slice dimension) and 6 in case of ultrafast fMRI ([2 2], [5 5], [20 20], [40 40], [60 60] and [74 96], where the last one corresponds to the entire image matrix). In Strategy B, data was first transformed in its complex form into k-space for ZF removal. After returning to image space, complex images were MP-PCA denoised at the slice level using 8 different sliding windows ([2 2], [3 3], [4 4], [5 5], [10 10], [15 15], [20 20] and [25 25]), and then converted to magnitude images. Data patches used for denoising were maximally overlapping (i.e., the window slid voxel-by-voxel) and were equally averaged after denoising. All fMRI repetitions (340 in case of multislice fMRI and 9000 in case of ultrafast fMRI) were used to build the matrices for denoising in both strategies. Moreover, we used a slightly improved version of the algorithm suggested by J. Veraart et al. (2016) that has shown better results in the estimation of the number of significant signal components (<https://github.com/sunenj/MP-PCA-Denoising>) (Olesen et al., 2022). Particularly, instead of increasing the number of signal components p from 0 until

$$\frac{1}{(M' - p)N'} \sum_{i=p+1}^{M'} \lambda_i \geq \frac{\lambda_{p+1} - \lambda_{M'}}{4} \frac{1}{\sqrt{N'(M' - p)}} \quad (1)$$

we used

$$\frac{1}{(M' - p)(N' - p)} \sum_{i=p+1}^{M'} \lambda_i \geq \frac{\lambda_{p+1} - \lambda_{M'}}{4} \frac{1}{\sqrt{N' M'}} \quad (2)$$

where M is the number of image repetitions, N is the total number of voxels within the sliding window, λ are the PCA eigenvalues, $M' = \min(M, N)$ and $N' = \max(M, N)$.

Further pre-processing included slice-timing correction (only for multislice data), motion correction, coregistration, normalization to the Allen Reference Atlas (Lein et al., 2007) and smoothing with a 3D isotropic Gaussian kernel with FWHM = 145 μm (if multislice fMRI) or 167 μm (if ultrafast fMRI). These steps were performed using SPM12 tools.

Image residuals were calculated after MP-PCA denoising to evaluate the algorithm performance, by subtracting the original undenoised images from the denoised images. The residuals were then divided by the map of estimated noise standard deviation σ (which is an output of the MP-PCA denoising procedure, obtained by voxel-wise averaging of the σ values estimated for each patch). The normality of residuals was assessed by histogram analysis and normal distribution fitting. Once normality was confirmed, the variance of this distribution fit then allowed to compute the percentage of estimated noise variance σ^2 explained by the residuals (equivalent to $\sigma_{\text{residuals}}^2/\sigma^2 \times 100$). The SNR gain obtained by the denoising was also calculated at this stage, by using the expected variance of residuals σ_r^2 obtained by truncating Gaussian noise components only (i.e., $\sigma_r^2 = \frac{(M'-P)(N'-P)}{M'N'}\sigma^2$), and computing

$$\begin{aligned} \text{SNR}_{\text{gain}}(\%) &= \left(\sqrt{\frac{\sigma^2}{\sigma^2 - \sigma_r^2}} - 1 \right) \times 100 \\ &= \left(\sqrt{\frac{M'N'}{M'N' - (M' - P)(N' - P)}} - 1 \right) \times 100 \end{aligned} \quad (3)$$

where P is the number of retained “signal” components given as an output of the method (obtained after voxel-wise averaging of the P values estimated for each patch).

To avoid hemodynamic response function (HRF) and GLM-related a-priori assumptions, a data-driven Fourier approach was performed for BOLD mapping (Gil et al., 2021; Nunes et al., 2019), where the Fourier Spectral Amplitude (FSA) at the paradigm’s fundamental frequency (i.e., the first harmonic) was mapped voxelwise to detect activated areas. Single-voxel time-courses were detrended and standardized to z-score prior to fast Fourier transform computation. In case of ultrafast data, FSA values at the two following harmonics of the paradigm (i.e., the second and third harmonics) were also extracted and summed with the FSA at the fundamental frequency before mapping. Phase information was used to separate positive from negative responses. Specifically, the phase difference $\Delta\Phi$ between the data and the paradigm was computed and averaged at the chosen harmonics for each voxel. Values of $|\Delta\Phi| < \pi/2$ translated into positive responses, whereas $\pi/2 < |\Delta\Phi| < \pi$ translated into negative responses. Individual maps were then thresholded with an FSA threshold of 0.3 (multislice fMRI) or 0.08 (ultrafast fMRI), whereas group maps (computed from the voxel-timecourses averaged between a total of $n = 9$ multislice or ultrafast scans) were thresholded with an FSA threshold of 0.45 (multislice fMRI) or 0.09 (ultrafast fMRI). These thresholds were manually chosen so that they were above the noise level of the Fourier spectrum of undenoised data (determined at higher and non-stimulation frequencies of the spectrum and from different brain regions) and were approximately equivalent to $p < 0.05$ (individual maps) or $p < 0.01$ (group maps) in undenoised data if using a *t*-test to statistically compare FSA values at the paradigm’s chosen harmonics with the FSA values obtained at higher frequencies (half) of the spectrum. A minimum cluster size of 10 (multislice fMRI) or 8 (ultrafast fMRI) voxels was also imposed, based on the minimum size required to detect activation in the smallest regions of the visual pathway (i.e., the lateral posterior and geniculate nuclei of the thalamus). To quantify improvements from MP-PCA denoising on BOLD mapping, the average FSA in activated voxels (i.e., the voxels above the threshold in the BOLD map obtained from undenoised data) was extracted from each map and percentage increases relative to the undenoised data results were quantified. The number of activated voxels in each map was also calculated to obtain the increase

in spatial extent of activation upon denoising. To understand if different functional CNR (fCNR) and temporal SNR (tSNR) values could explain the variations in the increase of FSA values between individual datasets, we run a Pearson’s correlation to assess the relationship between those variables for each sliding window. The average FSA extracted from activated voxels of undenoised data BOLD maps was used as an indicator for fCNR. The tSNR values were defined as the ratio of the average temporal signal over its standard deviation (averaged across the brain) and were calculated from a resting period in the middle of the undenoised data acquisitions.

2.5. Simulations with MP-PCA denoising

To better investigate changes in spatial extent of activation after denoising, we simulated ultrafast fMRI data with an a-priori known GT. A double-gamma HRF was generated for every voxel of four different bilateral ROIs of the visual pathway (primary visual cortex (V1), superior colliculi (SC), lateral posterior (LP) and lateral geniculate (LGN) nuclei of the thalamus), with slight changes in latency and amplitude per ROI and voxel, and convolved with the ultrafast fMRI paradigm to simulate BOLD responses. In particular, the BOLD responses were built so that they would peak randomly within the [1.7, 2.3] s interval after stimulus onset in voxels of the left hemisphere and within the [3.7, 4.3] s interval in voxels of the right hemisphere. Moreover, the percent signal change of the BOLD signal was randomly distributed between 0.75–1.25% in V1 and LP and 1.75–2.25% in the SC and LGN. The temporal resolution was set to 50 ms. Gaussian white noise was then added to GT data so that the average tSNR in the brain = 6.7, i.e., similar to the average values observed in real data. On the contrary to real data, no structured non-white noise (e.g., physiological noise) was added to the images.

Simulated complex data were MP-PCA denoised using 8 different sliding windows ([2 2], [3 3], [4 4], [5 5], [10 10], [15 15], [20 20] and [25 25]) and maximally-overlapping patches. All repetitions (9000) were used to build the matrices for denoising. Residual maps were again calculated by subtracting the undenoised (noisy) images from the denoised images and dividing the result by the map of estimated noise σ . Error maps were generated by subtracting the noise-free GT data from undenoised or denoised images and dividing the result by the standard deviation σ_s of the added noise. To characterize the strength and the pattern of the temporal correlation of neighboring voxels in the brain upon denoising, the noise kernel was estimated for each denoising window using LayNii (Huber et al., 2021) and fitted with a 2D isotropic Gaussian distribution to obtain its FWHM.

BOLD Fourier maps were calculated and thresholded in the same way as for real data (FSA threshold = 0.08 and minimum cluster size = 8). Percentage increases of the average FSA in GT activated voxels relative to the undenoised data results were quantified. The increase in spatial extent of activation relative to the GT data was also calculated. To further quantify the performance of MP-PCA denoising in terms of spatial extent of activation and determine the optimal denoising window for this dataset, the Dice similarity coefficient was calculated. Moreover, the sensitivity and false positive rate (FPR) were quantitatively evaluated as in the study of Fang et al. (2016). Specifically, the sensitivity was quantified as the number of true positive voxels over the number of true positive and false negative voxels, and the FPR as the ratio between the number of false positive voxels and the number of false positive plus true negative voxels within the first to fifth pixel perimeter layers of the GT activation volume inside the brain mask (FPR₁ to FPR₅).

To study the performance of this denoising method in more extreme conditions, four supplementary simulations were also performed: two with more (tSNR = 3) or less (tSNR = 15) noise, and two with smaller or higher values of BOLD percent signal change, respectively, one with 0.10–0.50% changes in V1 and LP and 0.30–0.70% in the SC and LGN, and another with 2.75–3.25% changes in V1 and LP and 3.75–4.25% in the SC and LGN.

To verify if different FSA thresholds would provide different optimal window results for each simulation, the area under the ROC curve (AUC) was estimated for each denoising window after quantifying the TPR and FPR of BOLD Fourier maps thresholded with 25 different FSA thresholds between 0 and 1.41 (largest FSA value found in every simulation), specifically: 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.10, 0.13, 0.19, 0.25, 0.32, 0.39, 0.46, 0.53, 0.64, 0.75, 0.86, 0.97, 1.08, 1.19, 1.30 and 1.41. The best threshold, i.e., the threshold that minimized the Euclidean distance to the optimal classifier (FPR = 0 and TPR = 1), was also determined for each window.

2.6. Simulations with NORDIC PCA denoising

To assess if NORDIC PCA denoising also causes activation “spreading” in BOLD maps, the first simulated ultrafast fMRI dataset (with tSNR = 6.7) was denoised with the NORDIC PCA approach using 5 different sliding windows ([2 2], [5 5], [10 10], [25 25] and [40 40]), maximally overlapping patches, and with and without the phase-stabilization correction approach (low-pass filter width = 10) described by Moeller et al. (2021). As in the simulations with MP-PCA denoising, the noise kernel was estimated for each denoising window. The FWHM of the noise kernel was then obtained by estimating the area of the FWHM level set (using cubic interpolation to estimate the FWHM of the noise kernel in any given direction) and calculating the radius of the circle with the same area. BOLD Fourier maps were calculated and thresholded as in the previous simulations.

2.7. General linear model analysis of functional data

Given that the majority of BOLD fMRI publications employ general linear model (GLM) for data analysis, we computed BOLD maps generated by GLM analysis for the acquired multislice fMRI datasets and the ultrafast fMRI simulations with tSNR = 6.7 and 0.75–2.25% BOLD changes to test the generalizability of our results. All GLM analyses were performed in SPM12. For both types of data, the experimental regressor of the design matrix was generated by convolution of a double gamma HRF with the respective paradigm, so that the expected BOLD response would peak at 3 s after stimulus onset. Motion correction parameters were used as nuisance regressors for multislice fMRI data. A high-pass filter with cutoff at 80 s (for multislice fMRI data) or 61 s (for the ultrafast fMRI simulation) was incorporated into the GLM to remove slow signal drifts. Moreover, a first-order autoregressive model (Friston et al., 2002) (for multislice fMRI data) or the FAST model implemented in SPM12 (Bollmann et al., 2018; Corbin et al., 2018; Olszowy et al., 2019) (for ultrafast data; using 18 components as implemented by default in the software) were used to model temporal correlations and pre-whiten the data before estimating GLM parameters.

Resulting *t*-value maps were thresholded with a minimum *t*-value of 1.65 (equivalent to $p < 0.05$ in undenoised data) and a minimum cluster size of 10 (multislice fMRI) or 8 (ultrafast fMRI simulation) voxels. A fixed-effect group analysis was also run using the $n = 9$ multislice fMRI datasets. In this case, the resulting *t*-value maps were thresholded with a minimum *t*-value of 3.09 (equivalent to $p < 0.001$ in undenoised data) and a minimum cluster size of 10 voxels.

3. Results

3.1. Conventional multislice and ultrafast fMRI provide artifact-free images and robust activation maps in the mouse visual pathway

Before application of MP-PCA denoising, we assessed the quality of the acquired multislice and ultrafast datasets. Figs. 2A and 3A show a single GE-EPI image obtained from a representative multislice and ultrafast fMRI acquisition, respectively, revealing good brain contrast and no remarkable artifacts. The temporal SNR in the brain was 13.1 ± 0.7 in

multislice datasets and 6.7 ± 0.4 in ultrafast datasets (both reported as mean \pm SD). Moreover, Fourier maps obtained from both data types (Figs. 2D and 3D) show robust contralateral (right) BOLD responses along the entire mouse visual pathway (V1, SC, LP and LGN) upon monocular flashing stimulation.

3.2. MP-PCA denoising of vendor reconstructed multislice fMRI images requires large sliding windows ($N \geq 250$) for good noise removal performance and improved functional activation mapping

Figure 2B presents representative MP-PCA denoised images with 5 different sliding windows on vendor reconstructed multislice images (Strategy A). Upon denoising, noise levels were approximately at the same level as the noise observed in undenoised data (Fig. 2A) when a small sliding window ([5 5 1], i.e., $N = 25$, where N is the total number of voxels within the sliding window) was used. Noise levels were clearly reduced when larger sliding windows ([20 20 1], [40 40 1], [20 20 5] and [10 10 10], i.e., $N \geq 400$) were employed. Residual images (Fig. 2C) do not contain edge effects or exhibit specific anatomical features and are thus considered to carry more of the estimated noised variance when larger windows were used. This effect becomes clearer when inspecting the plots portraying quantitative values of SNR gain and estimated noise variance in the residuals (Fig. 2H,I): in this particular scan, these were limited to 33% and 20%, respectively, when denoising was performed at the slice level (i.e. sliding window size = 1 in the slice dimension) and $N \leq 100$. The values increased to 64% and 35%, respectively, when using 3D windows (and $N \leq 125$), but only reached their maximal levels (>126% and >62%) when $N \geq 250$ (2D or 3D windows). Similar trends were observed in the BOLD Fourier maps: whereas the increase of FSA in activated voxels (Fig. 2J) was limited to 6% for smaller windows (2D windows or $N \leq 125$), gains of up to 17% were reached with the larger windows (3D windows and $N \geq 250$), accompanied by a greater volume of activation (Fig. 2E,K). As shown on the Fourier difference maps (Fig. 2F), MP-PCA with larger windows highlighted activation in regions with BOLD effect, while leaving areas with only noise contribution approximately intact. Consistent with these results, single-voxel time-courses obtained when using a small denoising window ([5 5 1]) show almost identical profiles as those obtained from undenoised data (Fig. 2G), whereas larger windows ([20 20 1] and [20 20 5]) clearly decrease the signal variation in the time-courses without affecting the functional changes obtained upon stimulation.

3.3. MP-PCA denoising of vendor reconstructed ultrafast fMRI images does not improve functional activation mapping even with the largest possible sliding window

When Strategy A was applied to ultrafast fMRI data, good noise removal performance was only achieved with [60 60] and [74 96] windows, as shown in Fig. 3. Specifically, in this particular scan, windows [2 2], [5 5], [20 20] and [40 40] ($N \leq 1600$) only provided a maximum SNR gain of 16% (Fig. 3H), with image residuals only containing a maximum of 10% of the estimated noise variance (Fig. 3I) and showing some structured appearance (Fig. 3C). When a larger window of [60 60] ($N = 3600$) was employed, image residuals started to contain more of the estimated noise variance (51%) and did not show anatomical detail (Fig. 3C), and an SNR gain of 81% could be achieved. These values further increased to 86% and 137%, respectively, when the largest possible denoising window ([74 96], corresponding to the entire image matrix) was used (Fig. 3H,I). These improvements did not, however, translate into significant BOLD Fourier map enhancements. FSA values in activated regions remained in the same range as the values obtained from undenoised data (Fig. 3E,F,J), with slightly more spatially extensive activation in the larger ($N \geq 3600$) windows (Fig. 3E,K). Moreover, single-voxel time-courses remained nearly identical to their undenoised counterparts (Fig. 3G).

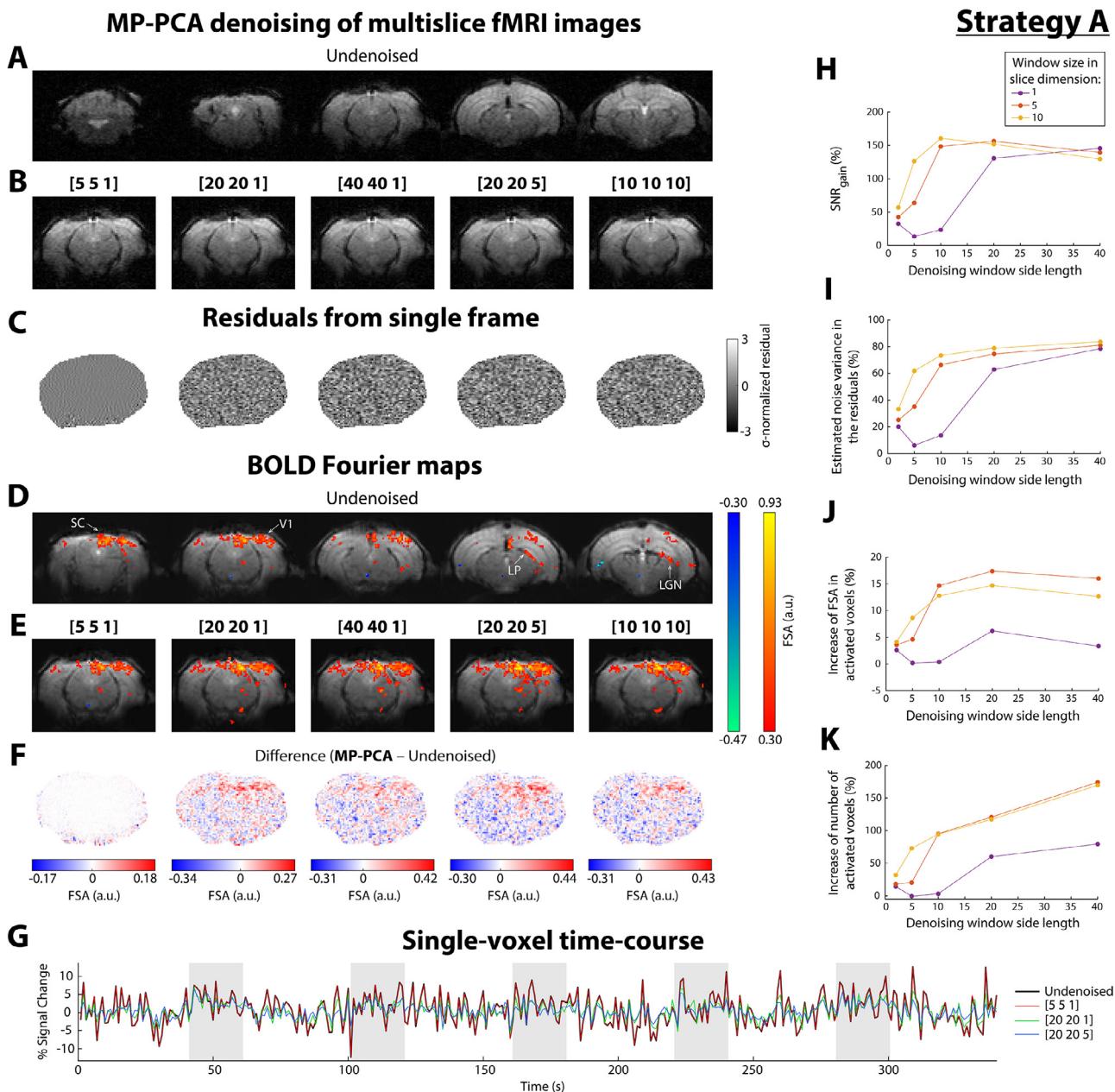


Fig. 2. MP-PCA denoising of multislice fMRI data reconstructed by vendor software (Strategy A). (A) Single GE-EPI image obtained from a representative multislice fMRI acquisition, before (5 out of 10 slices shown) and (B) after (1 out of 10 slices shown) MP-PCA denoising with 5 different sliding windows ([5 5 1], [20 20 1], [40 40 1], [20 20 5] and [10 10 10]) immediately after outlier correction. (C) Map of σ -normalized residuals of a single frame obtained after MP-PCA denoising with these 5 windows. (D) BOLD Fourier maps obtained before and (E) after MP-PCA denoising with these 5 windows. Maps are thresholded with a minimum FSA at paradigm's fundamental frequency = 0.3 and a minimum cluster size = 10. (F) Difference between the functional maps shown in (E) and (D). (G) Single-voxel detrended time-courses before (black line) and after MP-PCA denoising with 3 different sliding windows ([5 5 1] in red, [20 20 1] in green and [20 20 5] in blue). Gray areas represent the periods of visual stimulation. (H) Average brain SNR gain, (I) percentage of estimated noise variance explained by the residuals, (J) percentage increase of FSA in activated voxels, and (K) increase in spatial extent of activation, obtained after MP-PCA denoising with 15 different sliding windows. In particular, sliding window size varied between [2 2], [5 5], [10 10], [20 20] and [40 40] in the row and columns dimensions, and between 1 (in purple), 5 (in orange) and 10 (in yellow) in the slice dimension.

3.4. The performance of the MP-PCA denoising algorithm improves when ZF is removed from k-space

Vendor k-space data can contain inherent ZF in typical EPI settings (e.g. due to regridding and/or partial Fourier reconstruction), which generate inter-voxel correlations in image space that perturb the entire eigenspectrum, reduce MP distribution fitting and entail additional “apparent” signal components in the PCA domain. However, these zeros can be removed prior to denoising to mitigate spatial correlations. Video S1

shows GE-EPI images from representative multislice and ultrafast fMRI scans after ZF removal from k-space and after application of MP-PCA denoising at the slice level with 8 different sliding windows (Strategy B). The movies reveal no apparent motion besides breathing or any type of artifact, neither before or after denoising. Moreover, fine features of the images observed in undenoised data are preserved after MP-PCA denoising.

In contrast to Strategy A, multislice images denoised with Strategy B already show highly enhanced SNR with the smallest possible window

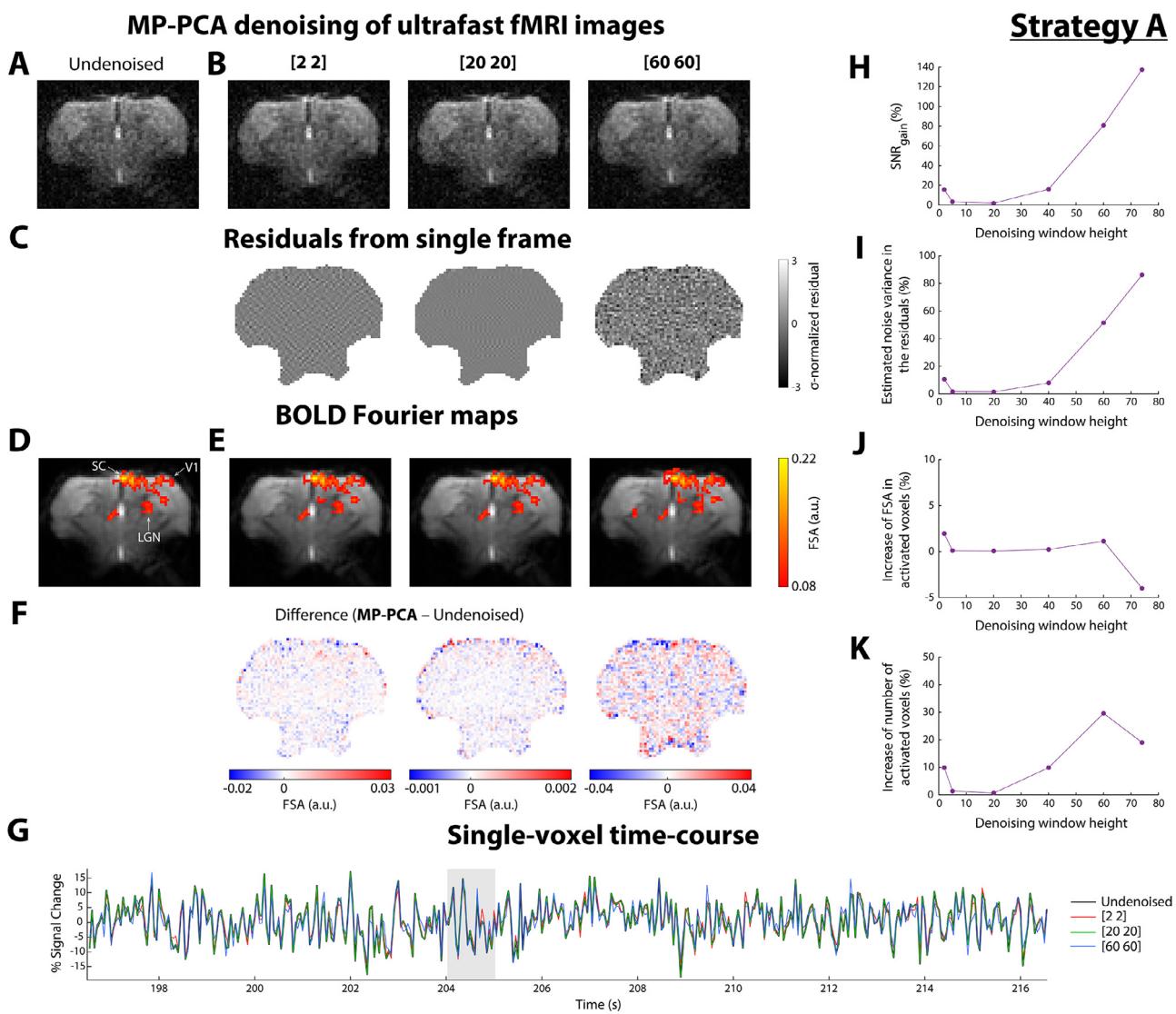


Fig. 3. MP-PCA denoising of ultrafast fMRI data reconstructed by vendor software (Strategy A). (A) GE-EPI image obtained from a representative ultrafast fMRI acquisition, before and (B) after MP-PCA denoising with 3 different sliding windows ([2 2], [20 20] and [60 60]) immediately after outlier correction. (C) Map of σ -normalized residuals of a single frame obtained after MP-PCA denoising with these 3 windows. (D) BOLD Fourier maps obtained before and (E) after MP-PCA denoising with these 3 windows. Maps are thresholded with a minimum sum of FSA at paradigm's fundamental frequency and two following harmonics = 0.08 and a minimum cluster size = 8. (F) Difference between the functional maps shown in (E) and (D). (G) Single-voxel detrended time-courses (only repetitions 3931 to 4331 are shown) before (black line) and after MP-PCA denoising with 3 different sliding windows ([2 2] in red, [20 20] in green and [60 60] in blue). The gray area represents the fifth period of visual stimulation. (H) Average brain SNR gain, (I) percentage of estimated noise variance explained by the residuals, (J) percentage increase of FSA in activated voxels, and (K) increase in spatial extent of activation, obtained after MP-PCA denoising with 6 different sliding windows.

([2 2]), which is further increased with progressively larger windows (Video S1A and Fig. 4A). Specifically, the average SNR gain increased from 64% to 134% for sliding windows between [2 2] and [5 5] ($N \leq 25$), respectively, and reached its maximum at 459% with a sliding window of [25 25] ($N = 625$), with relatively small deviations observed between individual scans (Fig. 4D). As with Strategy A, residual maps do not contain recognizable anatomic features for all tested sliding windows (Fig. 4B). Moreover, σ -normalized residuals are well approximated by a zero-centered normal distribution and have lower variance than unity, i.e., MP-PCA residuals have lower variance than the estimated noise variance, suggesting that the technique suppressed local signal fluctuations originating from thermal noise. Quantitatively, the variance of the residuals ranged from 45% (with the [2 2] window) to 96% (with the [25 25] window) of the estimated noise variance, with very small deviations registered between individual scans (Fig. 4E). Moreover, it closely matched the expected variance of residuals σ_r^2 obtained by truncating Gaussian noise components only (Fig. S1B,C). Single-voxel time-

courses exhibited in Fig. 4C show attenuation of spurious signal fluctuations present in undenoised multislice data with denoising, with larger windows decreasing more the variations, without disturbing the signal changes observed during visual stimulation.

Similarly to what was observed in multislice data and contrary to Strategy A, the application of denoising Strategy B on ultrafast data produced images with highly enhanced SNR already with the [2 2] ($N = 4$) window (Video S1B and Fig. 5A). Particularly, MP-PCA denoising achieved an average SNR gain of 39% (Fig. 5D) and image residuals already contained on average 33% of the estimated noise variance (Fig. 5E) when using this window. However, unlike multislice data results, this effect faded out with gradually larger windows and only came back with the [20 20] ($N = 400$) window (Video S1B, Fig. 5A,D,E), with SNR gain values and percentage of estimated noise variance present in the residuals reaching a maximum of 73% and 48%, respectively, when the largest tested window ([25 25], i.e., $N = 625$) was used. Single-voxel time-courses obtained when using a [5 5] denoising window are

Strategy B

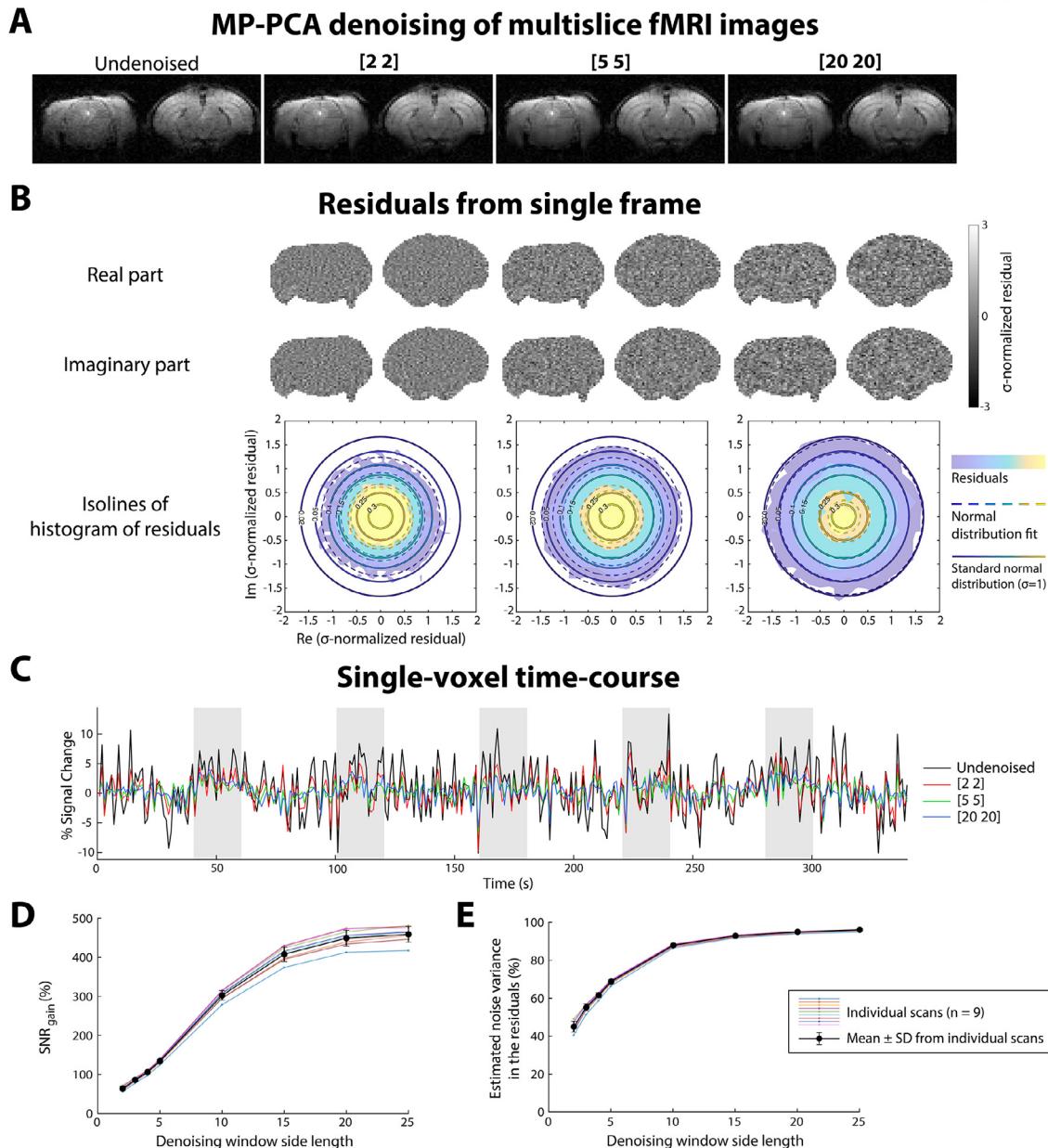


Fig. 4. Residuals, single-voxel time-courses and SNR gain of multislice fMRI data after MP-PCA denoising (Strategy B). (A) GE-EPI image of a representative multislice fMRI acquisition (2 out of 10 slices shown), before and after MP-PCA denoising at the slice level with 3 different sliding windows ([2 2], [5 5] and [20 20]). ZF was removed from k-space prior to data denoising. (B) Maps of real (top) and imaginary (middle) part of σ -normalized residuals of a single frame obtained after MP-PCA denoising with these 3 windows, together with filled contour plots containing isolines of the histogram of those residuals (bottom). Dashed lines represent the approximated normal distribution for the residuals whereas solid lines represent the standard normal distribution (unitary variance). (C) Single-voxel detrended time-courses before (black line) and after MP-PCA denoising with 3 different sliding windows ([2 2] in red, [5 5] in green and [20 20] in blue). Gray areas represent the periods of visual stimulation. (D) Average brain SNR gain across 8 different sliding windows and $n = 9$ different scans (represented by each color plot) of multislice fMRI data. The black dots with error bars and black lines represent the mean \pm SD results obtained from the individual scans. (E) Percentage of estimated noise variance explained by the residuals across the same sliding windows and scans.

very close to undenoised data (Fig. 5C), whereas smaller and larger windows ([2 2] and [20 20], respectively) slightly decreased the signal variation in the time-courses without affecting the functional changes obtained upon stimulation. Residuals show lack of anatomical structure and are well fitted by a zero-centered normal distribution (Fig. 5B). As with multislice data, the variance of residuals (Fig. S1I) followed the trend expected from MP-PCA theory (Fig. S1H).

Other MP-PCA denoising output metrics, such as the estimated noise variance σ^2 , varied between individual scans and showed a convex trend

with increasing window size in both types of data (Fig. S1A,G), although much more accentuated in the ultrafast fMRI datasets. For a [2 2] sliding windows, the estimated σ^2 was higher in regions with smaller signal values (i.e., regions with large blood vessels) both in multislice (Fig. S2B) and ultrafast (Fig. S2E) data. For larger sliding windows ([5 5] and [20 20] in Fig. S2B,E), σ^2 maps were less resolved and therefore smoother in both multislice and ultrafast data. Moreover, the sum of noise eigenvalues (Fig. S1D,J), the number of retained “signal” components P (Fig. S1E,K) and the number of eliminated “noise” components $M' - P$ (Fig.

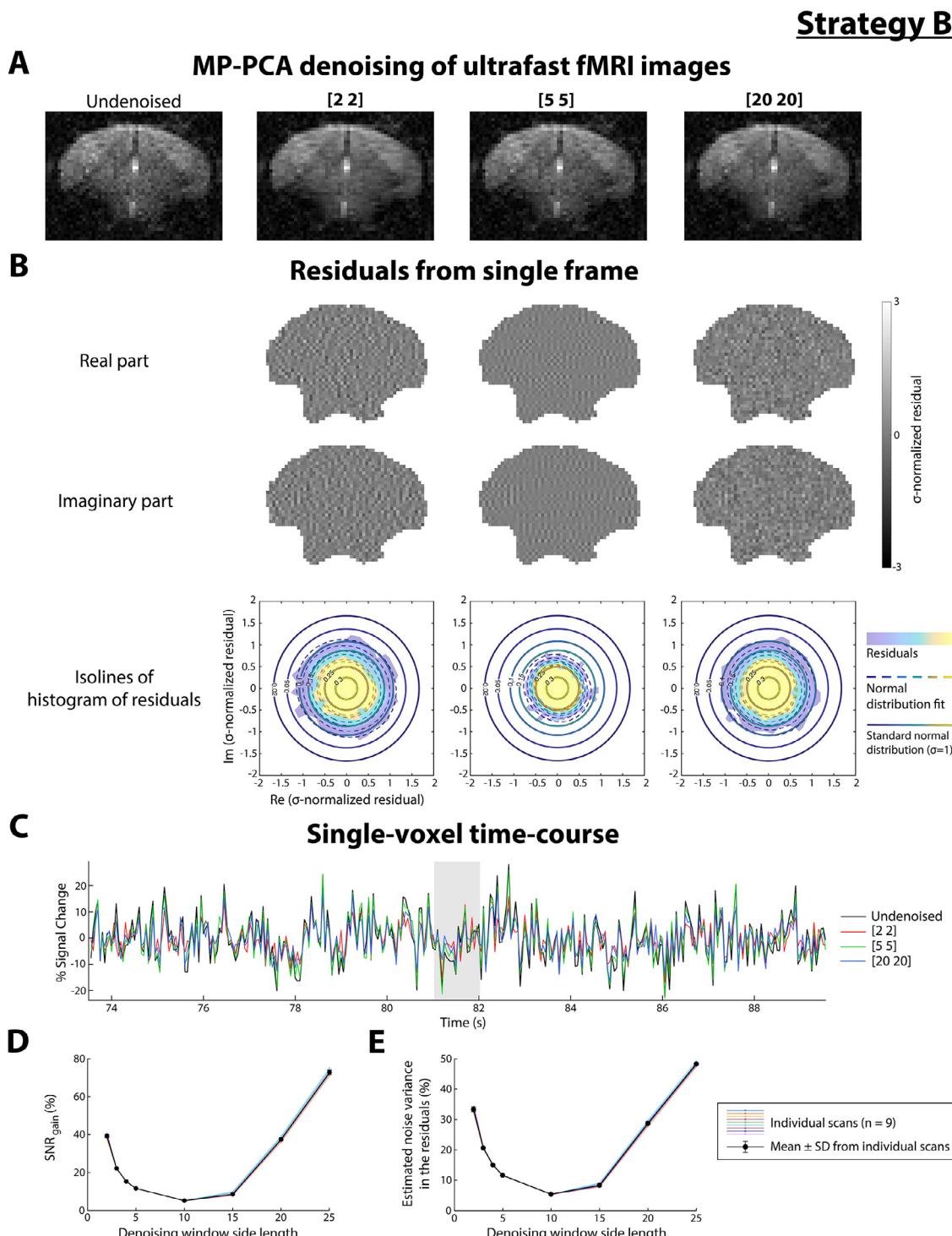


Fig. 5. Residuals, single-voxel time-courses and SNR gain of ultrafast fMRI data after MP-PCA denoising (Strategy B). (A) GE-EPI image of a representative ultrafast fMRI acquisition, before and after MP-PCA denoising with 3 different sliding windows ([2 2], [5 5] and [20 20]). ZF was removed from k-space prior to data denoising. (B) Maps of real (top) and imaginary (middle) part of σ -normalized residuals of a single frame obtained after MP-PCA denoising with these 3 windows, together with filled contour plots containing isolines of the histogram of those residuals (bottom). Dashed lines represent the approximated normal distribution for the residuals whereas solid lines represent the standard normal distribution (unitary variance). (C) Single-voxel detrended time-courses (only repetitions 1471 to 1791 are shown) before (black line) and after MP-PCA denoising with 3 different sliding windows ([2 2] in red, [5 5] in green and [20 20] in blue). The gray area represents the second period of visual stimulation. (D) Average brain SNR gain across 8 different sliding windows and $n = 9$ different scans (represented by each color plot) of ultrafast fMRI data. The black dots with error bars and black lines represent the mean \pm SD results obtained from the individual scans. (E) Percentage of estimated noise variance explained by the residuals across the same sliding windows and scans.

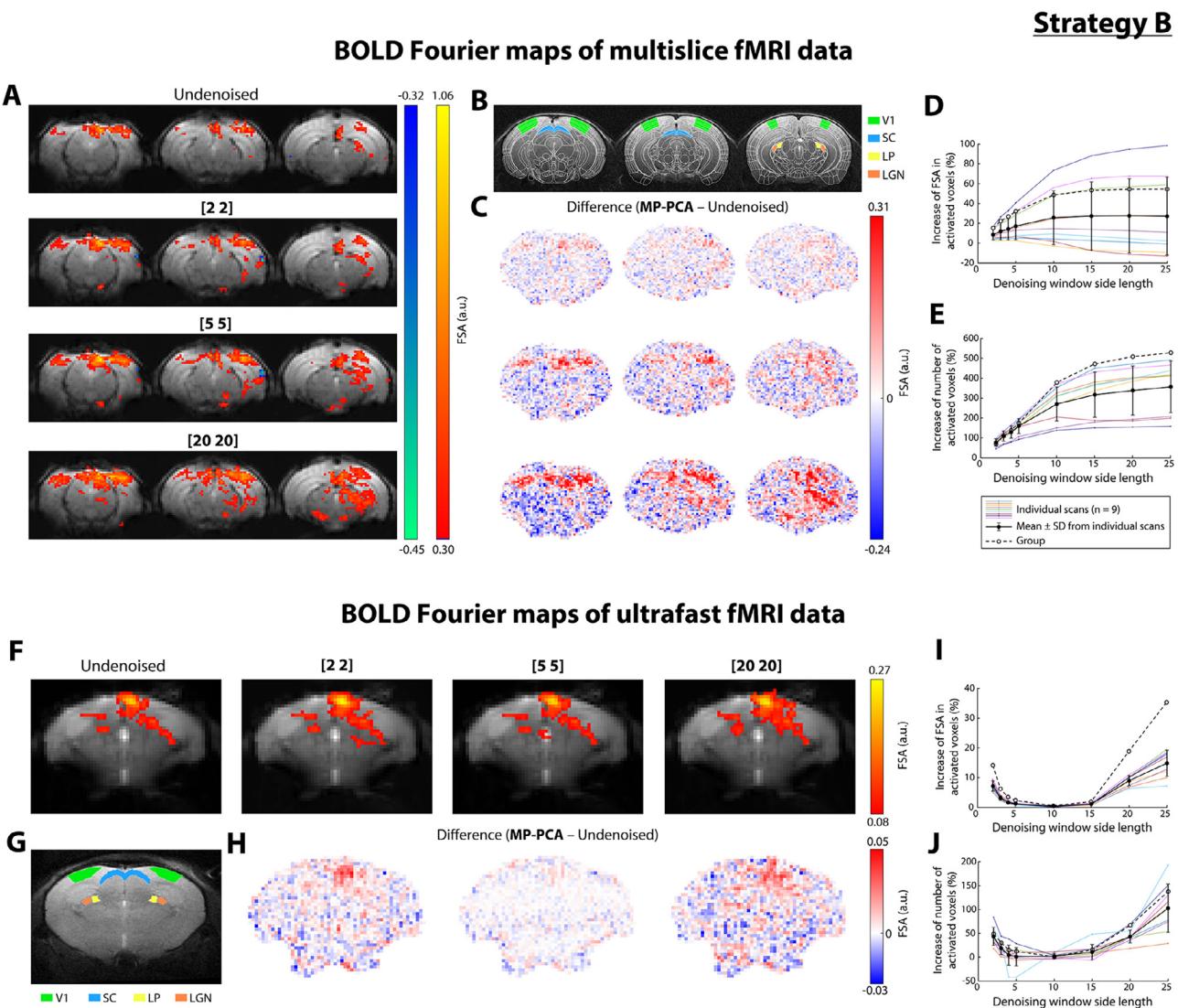


Fig. 6. Individual BOLD Fourier maps of multislice and ultrafast fMRI data after MP-PCA denoising (Strategy B). (A) BOLD Fourier maps of a representative multislice fMRI acquisition (3 out of 10 slices shown) obtained before and after MP-PCA denoising at the slice level with 3 different sliding windows ([2 2], [5 5] and [20 20]). ZF was removed from k-space prior to data denoising. Maps are thresholded with a minimum FSA at paradigm's fundamental frequency = 0.3 and a minimum cluster size = 10. (B) ROIs of the mouse visual pathway from the Allen Reference Atlas delineated on anatomical images. (C) Difference between the functional maps shown in (A). (D) Percentage increase of FSA at paradigm's fundamental frequency in activated voxels relative to the undenoised data results across 8 different sliding windows and $n = 9$ different scans (represented by each color plot) of multislice fMRI data. (E) Percentage increase of number of activated voxels in the BOLD maps relative to the undenoised data results across the same sliding windows and scans. The empty dots and dashed lines in (D) and (E) are the results from the group BOLD maps. (F-H) Same as (A-C), respectively, but for a representative ultrafast fMRI scan. Maps are thresholded with a minimum sum of FSA at paradigm's fundamental frequency and two following harmonics = 0.08 and a minimum cluster size = 8. (I-J) Same as (D-E), respectively, but for $n = 9$ different scans of ultrafast fMRI data.

S1F,L) monotonically increase with larger window sizes (until the [20 20] sliding window). For a [2 2] window, the number of signal components P was higher in regions closer to the receiver coil (i.e., with higher signal values) in multislice data (Fig. S2C), and practically homogeneous across the brain in ultrafast data (Fig. S2F). As with σ^2 , P maps were less resolved and smoother in both types of data when using larger windows.

3.5. Larger denoising windows enhance sensitivity to functional activation and increase spatial extent of activation

Figure 6 shows the BOLD Fourier maps and respective difference maps when using Strategy B and 3 different sliding windows. Consistent with the results above, functional maps already show a clear increase in FSA values in visual areas when a sliding window as small as [2 2]

is employed (Fig. 6A,F). Specifically, FSA values in activated regions increased on average 9% in multislice data and 7% in ultrafast data (Fig. 6D,I) for the [2 2] window. In multislice data, these values further increased with progressively larger windows (Fig. 6A,C), reaching a maximum average increase of 28% when a [20 20] window was used (Fig. 6D). On the contrary, in ultrafast data, the FSA values in activated regions began to decrease with larger windows, reaching the undenoised data value range with a [5 5] window, and only started increasing again with the [20 20] window (Fig. 6F,H), reaching a maximum average increase of 15% when a [25 25] window was used (Fig. 6I).

Interestingly, the increase in sensitivity to functional activation provided by larger windows was accompanied by an increase in the number of activated voxels, as clearly seen in Fig. 6A,F. Specifically, the increase in the number of activated voxels reaches average values as high as 270% with a [10 10] sliding window in multislice data, and of 100%

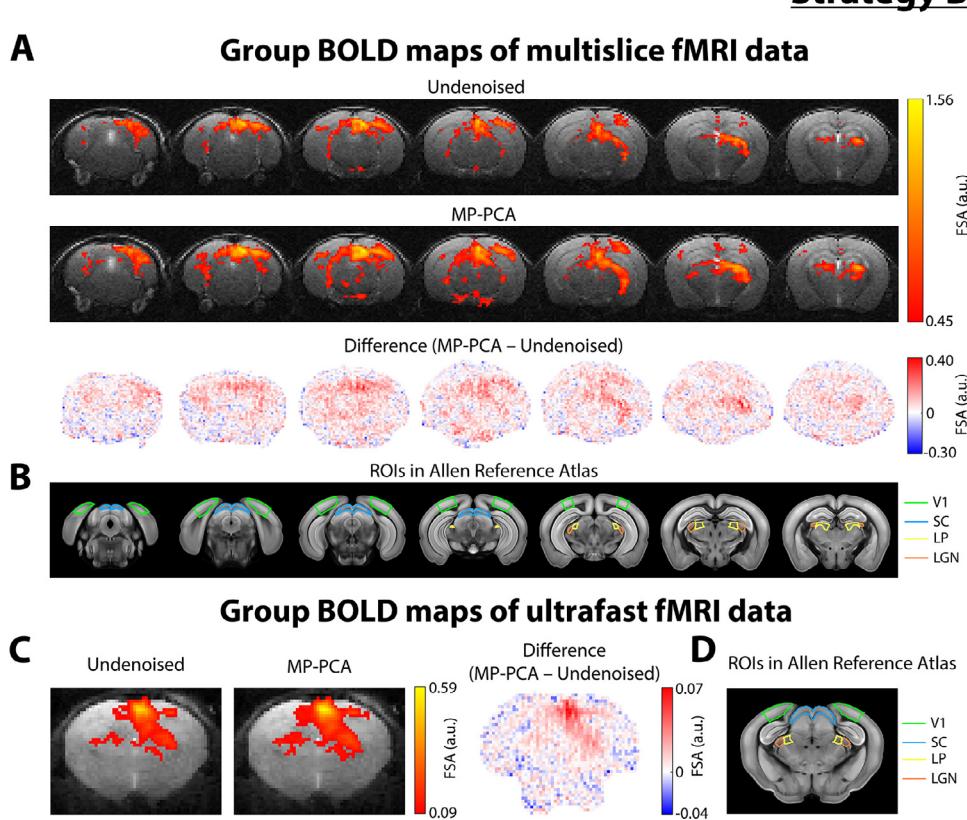


Fig. 7. Group BOLD Fourier maps of multislice and ultrafast fMRI data after MP-PCA denoising (Strategy B). (A) Group BOLD Fourier maps ($n = 9$ scans) of multislice fMRI data (7 out of 10 slices shown) obtained before (top) and after (middle) MP-PCA denoising at the slice level with a sliding window = [2 2], and respective difference map (bottom). Maps are thresholded with a minimum FSA at paradigm's fundamental frequency = 0.45 and a minimum cluster size = 10, and displayed above anatomical images. (B) ROIs of the mouse visual pathway delineated on the Allen Reference Atlas. (C-D) Same as (A-B), respectively, but for ultrafast fMRI data ($n = 9$ scans). Maps are thresholded with a minimum sum of FSA at paradigm's fundamental frequency and two following harmonics = 0.09 and a minimum cluster size = 8.

with a [25 25] sliding window in ultrafast data (Fig. 6E,J). This increase in spatial extent of activation “spreads out” of the regions that were activated in the maps obtained from undenoised data (Fig. 6C,H) and out of the ROIs of the visual pathway (Fig. 6B,G). Moreover, it can be seen from the plots of Fig. 6D,I that the increase of FSA in activated voxels starts to strongly vary between individual scans when larger windows ($N \geq 100$ in multislice data and $N \geq 400$ in ultrafast data) are employed. In some individual multislice scans, FSA values actually started to deviate from the average increasing pattern and to decrease when using windows equal or larger than [10 10] ($N \geq 100$).

A Pearson’s correlation was performed to test the relationship between the average FSA extracted from activated voxels of undenoised data BOLD maps (indicator of fCNR) and the values of FSA increase for each sliding window (Fig. S3A,B). A high correlation between these variables was observed: on average $r(7) = 0.83$ with $p < 0.008$ for all windows in multislice fMRI data, and $r(7) = 0.78$, $p < 0.03$ for all windows in ultrafast fMRI data except the [10 10] and [15 15] ($p > 0.15$), where the FSA increase was the lowest (Fig. 6I). On the contrary, we found a weak correlation between initial tSNR and FSA increase values for each sliding window, with an average $r(7) = -0.25$ ($p > 0.46$) in multislice data and $r(7) = -0.17$ ($p > 0.21$) in ultrafast data (Fig. S3C,D).

3.6. Group analysis shows higher activation to stimulation confined to regions of the visual pathway upon MP-PCA denoising with a [2 2] sliding window

Individual data MP-PCA denoised with a sliding window of [2 2] and Strategy B were averaged to produce the group BOLD maps shown in Fig. 7. Both in multislice (Fig. 7A) and ultrafast (Fig. 7C) group data results, MP-PCA highlighted activation in areas that were activated in the maps obtained from undenoised data, without substantially increasing the spatial extent of activation. Specifically, FSA values in activated voxels increased 15% in multislice data and 14% in ultrafast data in comparison to undenoised data results (Fig. 6D,I). These increases were

larger than the average increases of 9% and 7% obtained in individual BOLD maps, respectively. The increase in the volume of activation remained, however, approximately at the same level as the average increase observed in individual maps (Fig. 6E,J). This remained true for gradually larger windows up to [5 5] in multislice data and [15 15] in ultrafast data. For larger windows, the increase in the number of activated voxels in the group maps became significantly greater than the average increase observed in individual maps.

3.7. Magnitude versus complex data denoising produce similar results

The fact that MP-PCA denoising was applied to magnitude data in Strategy A and to complex data in Strategy B raises the question of whether the gains demonstrated with Strategy B are not only due to ZF removal from k-space prior to denoising but also partially related to the differences in the nature of the data that is being denoised. To address this question, we repeated Strategy A immediately after outlier correction, but using complex data. Very similar results to the magnitude denoising (Figs. 2 and 3) were obtained (Figs. S4 and S5). We also note in passing that reintroducing the zeros to the denoised k-space when using Strategy B restores the original resolution and has no impact on subsequent analyses (Fig. S6).

3.8. Simulations show that MP-PCA denoising incurs activation “spreading” and smoother functional maps

As shown in Fig. 8A, the Fourier map of GT simulated ultrafast fMRI data reveals a clear activation in the four different bilateral ROIs of the visual pathway (V1, SC, LP and LGN), with slightly higher FSA values found in the right hemisphere where the BOLD response peaked later. When Gaussian white noise was added so that the average tSNR in the brain = 6.7 (Fig. 8B), FSA values in activated regions markedly decreased, and LGN and LP disappeared from the thresholded map. After applying MP-PCA denoising with a [2 2] sliding window on simulated

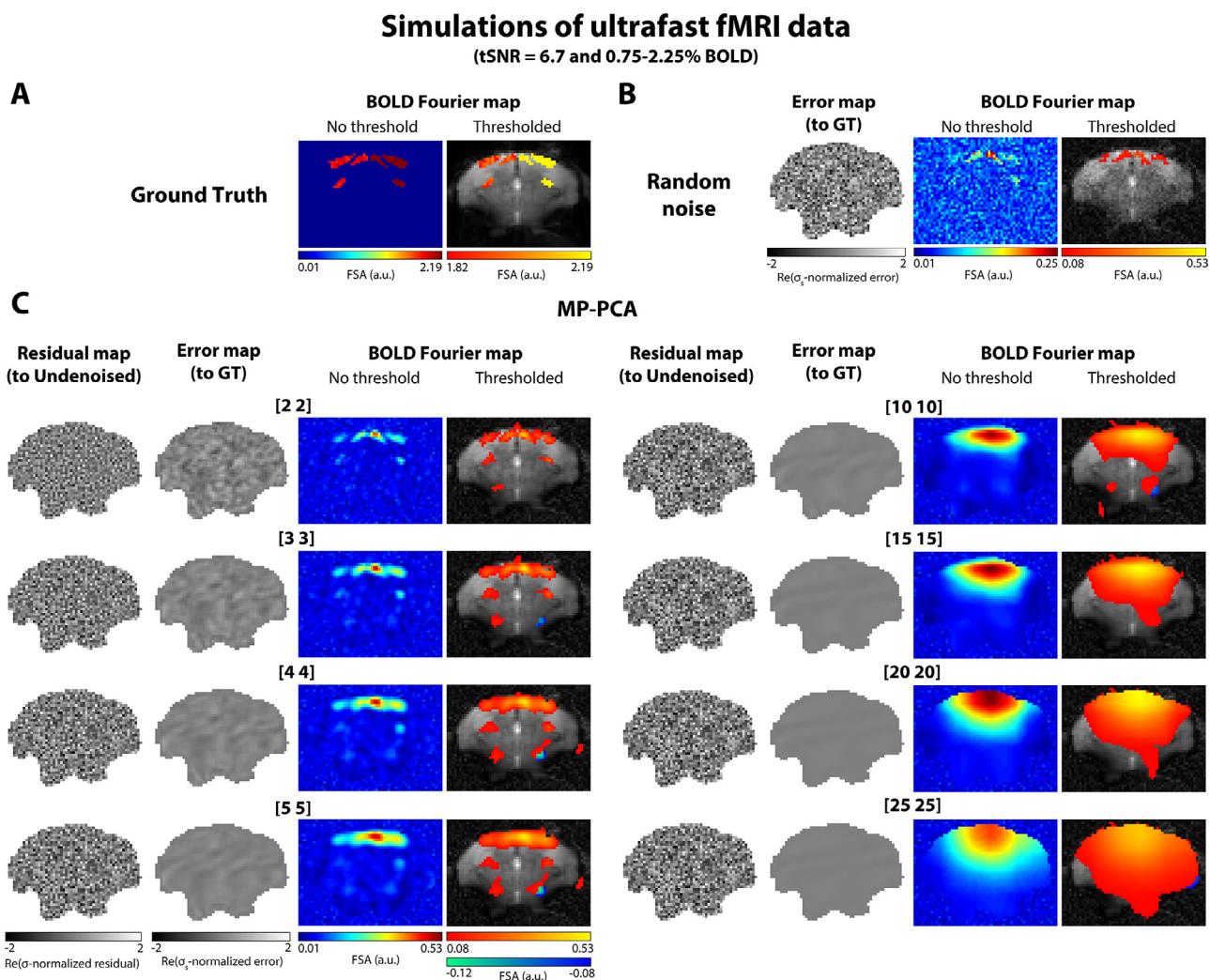


Fig. 8. Residual, error and BOLD maps of simulated ultrafast fMRI data (with tSNR = 6.7 and 0.75–2.25% BOLD changes) after MP-PCA denoising. (A) BOLD Fourier map of ultrafast fMRI data simulated without any source of noise, with (right) and without (left) threshold. (B) Same as (A) after addition of Gaussian white noise so that the average tSNR in the brain = 6.7. A map of σ_s -normalized error of a single frame is displayed on the left. (C) Maps of σ_s -normalized residuals and σ_s -normalized error of a single frame (left) and BOLD Fourier maps with (right) and without (middle) threshold upon MP-PCA denoising with 8 different sliding windows. Maps are thresholded with a minimum sum of FSA at paradigm's fundamental frequency and two following harmonics = 0.08 and a minimum cluster size = 8. A representative image of the data is shown below the maps.

noisy data (Fig. 8C), the map shows a clear distinction between activated and non-activated regions and all the ROIs of the pathway pass the FSA threshold. Moreover, denoised images (on which statistical maps are overlaid) show a strong noise reduction, and error maps show lack of anatomy, indicating signal-preservation. However, when larger denoising windows are employed ([3 3] to [25 25]), although error maps show a smaller difference between denoised and GT images and FSA values in activated regions increase, functional maps become smoother and activation starts to “spread out” to neighboring voxels that were not activated in GT data. This effect is also visible in the single-voxel time-courses (Fig. S7). Although an activated voxel in the right V1 (voxel 1 in Fig. S7A) regained part of its initial fCNR (Fig. S7B) with denoising (Fig. S7D), a voxel close to V1 and SC that was not activated in GT data (voxel 2 in Fig. S7A) also started to exhibit an artifactual response to stimulation when a [10 10] sliding window was used (Fig. S7E). A similar effect also occurred with another voxel even further away from activated regions (voxel 3 in Fig. S7A), which also started to show an activation profile with the [20 20] denoising window (Fig. S7F). After estimating the average temporal correlation of neighboring voxels through noise kernel calculation of brain image data (Fig. S8), we found that the functional point spread function (PSF) follows a Gaussian shape isotropically

across space upon denoising (Fig. S8B) and that its FWHM increases approximately linearly with denoising window side length (Fig. S8D).

3.9. Activation “spreading” patterns are consistent if only the center voxel of each patch is kept during denoising

To understand if the observed increase in spatial extent of activation could be partially related to equal averaging in data patches after denoising, we repeated MP-PCA denoising on the simulated data (with tSNR = 6.7), but kept only the center voxel of the patch at each slide (Fig. S9). Despite slightly lower blurring, activation “spreading” patterns and smoothing of functional maps upon larger denoising windows are still visible on the maps (Fig. S9C).

3.10. Optimal MP-PCA sliding window for improved specificity of functional mapping depends on data tSNR and fCNR

While the [2 2] sliding window improved accuracy of functional measurements without generating false BOLD responses in non-activated regions in the simulated ultrafast datasets with tSNR = 6.7, different results were obtained for varying noise levels. When a tSNR of

3 was imposed (Fig. S10), no voxels passed the threshold (Fig. S10B). Although the right V1 and SC started to stand out from the rest of the brain in the BOLD map when a [2 2] sliding window was used (Fig. S10C), the left V1 and SC only showed clear activation when a [4 4] sliding window was employed. LGN and LP do not pass the threshold in these maps, except when activation has already spread to the entire upper part of the brain (with a [20 20] window). In contrast, for higher tSNR of 15 (Fig. S11) the BOLD map of undenoised data reflected activity in nearly the entire visual pathway (Fig. S11B), and activation spreading to non-activated regions was already visible when a sliding window as small as [2 2] was used (Fig. S11C).

The quantification of these results is summarized in Fig. 9. The FSA increase in GT activated voxels (Fig. 9A) and the increase in spatial extent of activation (Fig. 9B) grow with larger windows at different rates, depending on the tSNR. Moreover, the Dice similarity coefficient (Fig. 9C) calculated between each noisy or MP-PCA map and the GT BOLD map shows that the optimal sliding window for the most accurate spatial extent of activation depends on the tSNR: a maximum score of 0.51 was obtained for the [4 4] window when tSNR = 3, 0.76 for the [2 2] window when tSNR = 6.7, and 0.83 for undenoised data when tSNR = 15 (i.e., when the SNR is that high, MP-PCA denoising only produces activation spreading). Those optimal windows allow a sensitivity of at least 0.70 in the maps (specifically, 0.70 for the [4 4] window when tSNR = 3, 0.91 for the [2 2] window when tSNR = 6.7, and 0.75 when data is not denoised when tSNR = 15) (Fig. 9D), and to limit the FPR to values of 0.37 in the 1-pixel perimeter layer of the GT activation volume (Fig. 9E), 0.10 in the 2-pixel perimeter layer (Fig. 9F), and values lower than 0.05 in the 3- to 5-pixel perimeter layers (Fig. 9G-I). Although larger denoising windows provide higher sensitivity, they also lead to higher FPR values.

Clearly these effects also depend on the fCNR (amplitude of the BOLD response). Similar trends were observed for datasets that were simulated with smaller (Fig. S12) or higher (Fig. S13) values of percent signal changes of the BOLD signal. Even though the Fourier map of GT data was exactly the same for all conditions (Figs. 8A, S12A and S13A), large differences began to appear once noise was added (Figs. 8B, S12B and S13B). Moreover, although error maps and denoised images (Figs. 8C, S12C and S13C) suggest that MP-PCA had the same performance in terms of removal of estimated noise, when smaller percent signal changes were imposed, activation spreading only became obvious when windows larger than [5 5] were employed (Fig. S12C); when BOLD responses were higher, a [2 2] denoising window already revealed a notable amount of false positives (Fig. S13C). Indeed, the plots with the Dice score obtained for each simulation and denoising window (Fig. 10C) reveal that the optimal window size for improved mapping specificity varies with the amplitude of the BOLD response: a maximum score of 0.49 was obtained for the [3 3] window when BOLD amplitudes were distributed in the 0.10–0.70% range, and a maximum score of 0.93 was obtained for the undenoised data when responses were in the 2.75–4.25% range. Again, these were accompanied by reasonable levels of sensitivity (at least 0.51, Fig. 10D) and limited FPR (0.26 for FPR₁, 0.07 for FPR₂, and 0.02 for FPR₃, and 0 for FPR₄ and FPR₅, Fig. 10E-I). As with the results shown for simulations with different tSNR levels, different values of fCNR resulted in different values of FSA (Fig. 10A) and spatial extent (Fig. 10B) increase for each of the denoising windows tested. Moreover, the larger the fCNR, the larger the growth rate of those variables. When the fCNR was the lowest (shown in purple in Fig. 10A), it was possible to see a decreasing pattern of FSA increase values for windows larger than [10 10].

When performing a threshold-independent analysis of the maps based on AUC values (Fig. S14B), the optimal windows for the best balance between sensitivity and specificity were [5 5] when tSNR = 3 (with AUC = 0.88) and [2 2] when tSNR = 6.7 (with AUC = 0.98). When tSNR = 15, a maximum AUC value of 0.99 was obtained for undenoised data. When data were simulated with smaller (0.10–0.70%) signal changes, the optimal window was [5 5] (with AUC = 0.81); and

when data were simulated with higher (2.75–4.25%) signal changes, a maximum AUC value of ≈ 1 was obtained for undenoised data. Therefore, the optimal windows based on AUC values for each simulation (indicated by the asterisks on top of the graphs shown in Fig. S14) coincided or were very close in size to the windows providing the highest Dice score when imposing an FSA threshold of 0.08 (Fig. S14A). Moreover, the best threshold obtained for those optimal windows (Fig. S14C) was also 0.08, except when MP-PCA denoising data simulated with a tSNR of 15 or with lower fCNR, in which cases the best threshold was 0.07.

3.11. NORDIC PCA denoising also incurs activation “spreading” but less smoothed BOLD maps

Figure 11 shows the Fourier maps obtained after applying NORDIC PCA denoising with 5 different windows to the simulated ultrafast fMRI dataset with a tSNR = 6.7 and 0.75–2.25% BOLD changes. When the phase-stabilization correction was not performed, activation started to “spread out” for windows larger than [5 5] (Fig. 11C). The bilateral LGN and LP did not appear activated with any of the sliding windows tested, except when the entire brain was shown as activated (e.g. with a [40 40] window). Moreover, FSA values in GT activated voxels increased a maximum of 68% in comparison to undenoised data results when a [25 25] window was used (Fig. 11E).

Phase-stabilization increased the maximum FSA gain to 95% when a [10 10] window was used (Fig. 11E), reduced the amount of spreading (Fig. 11D,F), especially when using the [25 25] and [40 40] windows, and allowed to get activation in both LGN and LP with a [5 5] sliding window; still, activation spreading was incurred for large windows. Nevertheless, it is noteworthy that NORDIC PCA denoising did not produce smooth maps as in MP-PCA denoising (Fig. 8C). Noise kernels estimated after NORDIC PCA denoising (Fig. S8C) indicated a slight increase in spatiotemporal correlations when compared to undenoised data, although limited to a much smaller neighborhood and FWHM (Fig. S8D) than those obtained with MP-PCA denoising (Fig. S8B). Moreover, contrary to MP-PCA, the functional PSFs are not isotropic across space, nor seem to follow a Gaussian shape in any direction.

3.12. Activation “spreading” patterns occur also for GLM analyses

To test whether our FSA-based conclusions generalize to more common GLM data analysis approaches, we performed a GLM analysis to the acquired multislice fMRI datasets and one of the ultrafast fMRI simulations. Fig. 12A,C displays the GLM analysis results for the multislice fMRI acquisition shown in Fig. 6A,C after MP-PCA denoising. Despite the higher detection of spurious negative responses using GLM analysis (Fig. 12A) when compared to Fourier analysis results (Fig. 6A), the increase in sensitivity to functional activation provided by larger windows was consistent between the methods (Figs. 6C and 12C). Specifically, *t*-values in activated regions increased on average 5% in multislice data when using a [2 2] sliding window, reaching a maximum average increase of 14% when a [10 10] window was used (Fig. 12D). Again, the increase of *t*-values in activated voxels started to strongly vary between individual scans when larger windows ($N \geq 100$) were employed (Fig. 12D). In one of the scans, *t*-values actually started to decrease when a window as small as [2 2] was used. Moreover, the monotonic increase in the number of activated voxels provided by larger windows observed in BOLD Fourier maps (Fig. 6A) was also visible in the BOLD GLM maps (Fig. 12A). Specifically, the spatial extent of activation on these maps increased on average 127% when using a [10 10] window, reaching a maximum average 157% increase when the largest sliding window ([25 25]) was employed (Fig. 12E). Therefore, despite growing at slightly lower rates when using a GLM analysis, the “spreading” patterns were consistent between both methods.

In agreement with these results, the GLM analysis obtained for the simulated ultrafast data after MP-PCA denoising (Fig. 12F-H) also shows

Simulations of ultrafast fMRI data with different tSNR

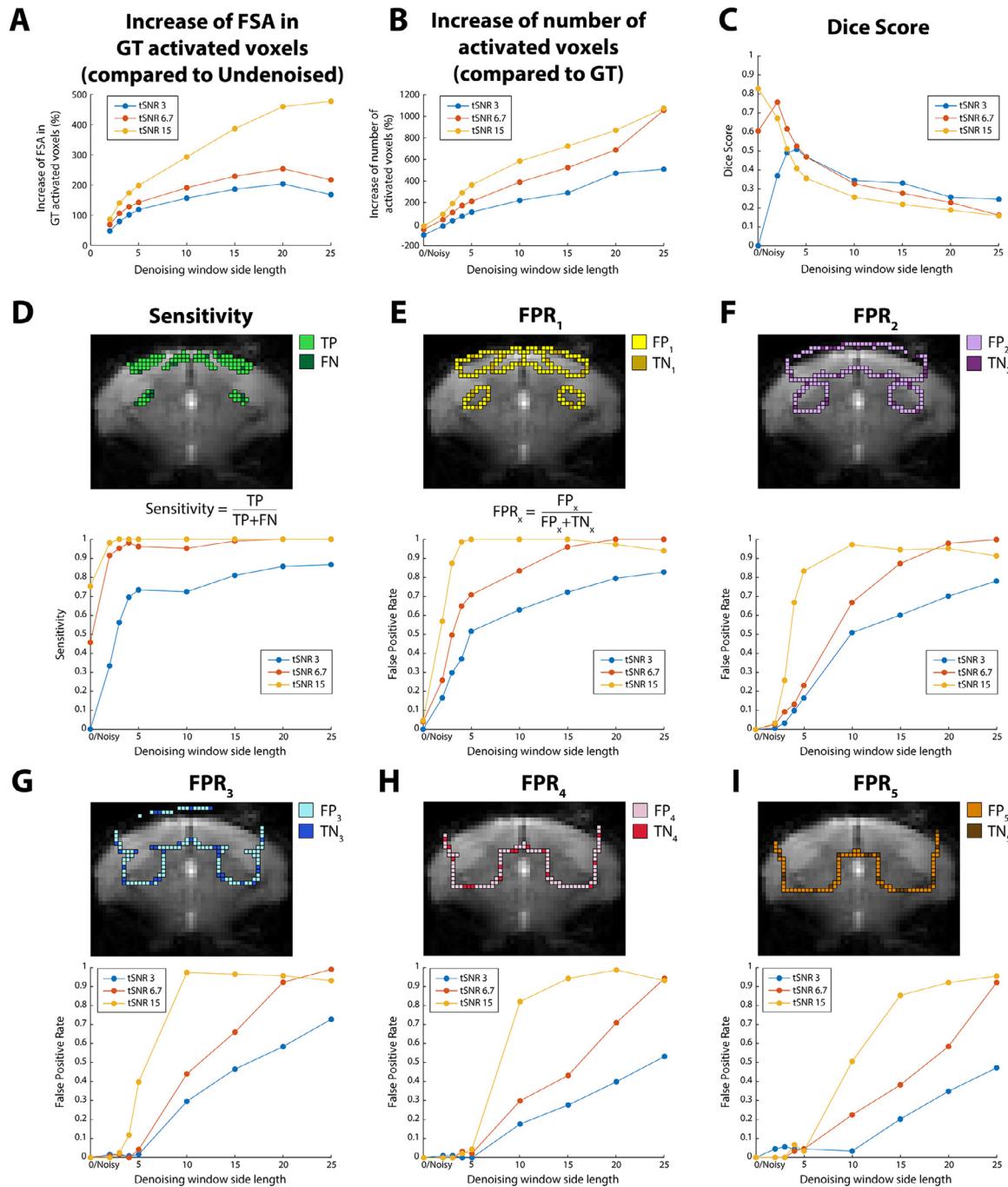


Fig. 9. Key metrics of BOLD maps of simulated ultrafast fMRI data (with 3 different tSNRs) after MP-PCA denoising. (A) Percentage increase of FSA at paradigm's fundamental frequency and two following harmonics in GT activated voxels (i.e., the voxels above the threshold in the BOLD map obtained from GT data) relative to the undenoised data results across 8 different sliding windows and 3 different tSNRs (3 in blue, 6.7 in orange, 15 in yellow) of simulated ultrafast fMRI data. (B) Percentage increase of number of activated voxels in the BOLD maps relative to the GT data results across different sliding windows and tSNRs. (C) Dice score of activation maps across different sliding windows and tSNRs. (D) (Top) The sensitivity of activation maps is defined as the ratio between the number of true positive (TP) voxels and the number of true positive and false negative (FN) voxels. (Bottom) Sensitivity of maps across different sliding windows and tSNRs. (E-I) (Top) The FPR is defined as the number of false positive (FP) voxels over the number of false positive and true negative (TN) voxels within the 1- to 5-pixel perimeter layers of the GT activation volume inside the brain mask (FPR₁ to FPR₅). (Bottom) FPR₁ to FPR₅ of maps across different sliding windows and tSNRs.

large similarities with the Fourier analysis results (Fig. 8): slightly higher *t*-values found in the right hemisphere in the GT data map (Fig. 12F), the disappearance of the left LGN and LP from the thresholded map when Gaussian white noise was added to the images (Fig. 12G), the clearer distinction between activated and non-activated regions in BOLD maps

after denoising with a [2 2] sliding window (Fig. 12H), and, although at a slightly higher rate than Fourier analysis, the “spreading” pattern of activation with larger denoising windows (Fig. 12H-J).

Finally, GLM results obtained for NORDIC PCA denoised data (with phase-stabilization correction) were also very consistent with the

Simulations of ultrafast fMRI data with different % BOLD change

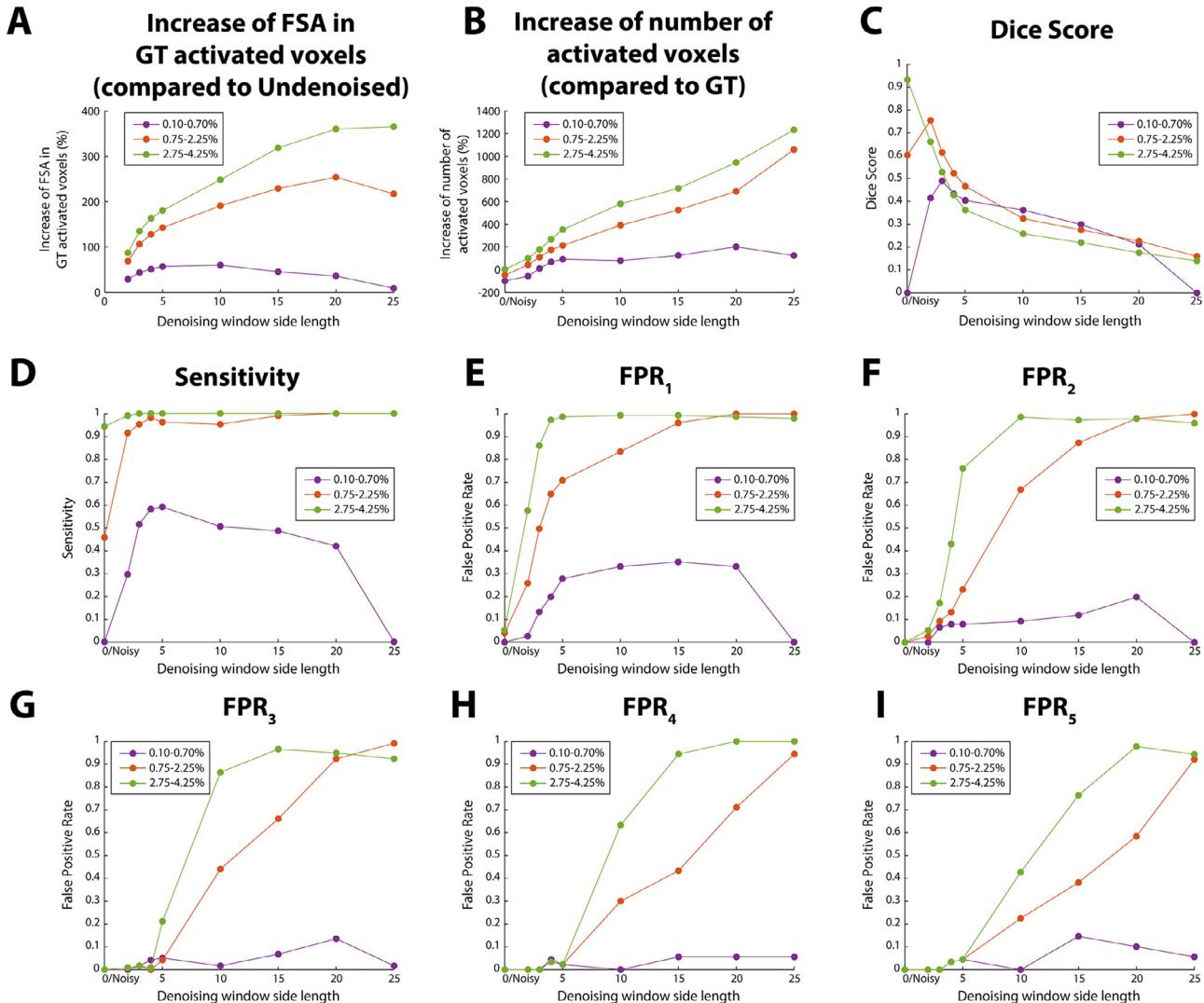


Fig. 10. Key metrics of BOLD maps of simulated ultrafast fMRI data (with 3 different intervals of % BOLD change) after MP-PCA denoising. (A) Percentage increase of FSA at paradigm's fundamental frequency and two following harmonics in GT activated voxels (i.e., the voxels above the threshold in the BOLD map obtained from GT data) relative to the undenoised data results across 8 different sliding windows and 3 different intervals of % BOLD change (0.10–0.70% in purple, 0.75–2.25% in orange, 2.75–4.25% in green) of simulated ultrafast fMRI data. (B) Percentage increase of number of activated voxels in the BOLD maps relative to the GT data results, and (C) Dice score, (D) Sensitivity and (E–I) FPR₁ to FPR₅ of activation maps across different sliding windows and intervals of % BOLD change.

Fourier analysis results (Fig. 11) both in terms of sensitivity to functional activation (Fig. 12L) and in increasing the spatial extent of activation (Fig. 12M) with window size, showing activation “spreading” for windows larger than [2 2]. Again, GLM BOLD maps from NORDIC appeared less smoothed than the maps obtained after MP-PCA denoising (Fig. 12K).

4. Discussion

MP-PCA denoising has proved highly useful for thermal noise removal in MR data containing a redundancy (J. Veraart et al., 2016; Adanyeguh et al., 2018; McKinnon et al., 2018; Grussu et al., 2020; Does et al., 2019). In the context of fMRI, previous work has shown that the application of MP-PCA denoising improved task-based and resting-state mapping (Ades-Aron et al., 2021; Adhikari et al., 2019; Diao et al., 2021), but potential sensitivity/specificity trade-offs and potential pitfalls from using data reconstructed by the vendor – especially in pre-clinical fMRI – were not thoroughly investigated yet. We thus sought to

investigate the advantages and potential limitations of MP-PCA denoising in fMRI, in particular in rodent preclinical MRI. Our findings suggest that significant gains in sensitivity can be obtained from MP-PCA denoising in such data, but that spatial specificity can be compromised, since activation may “spread” in space, and that this activation spreading will depend on the tSNR, fCNR, and on the denoising window size. There are many instances where sensitivity would be preferred even if the exact location of activity is somewhat compromised, and MP-PCA can greatly benefit these types of investigations; however, studies focusing on very specific locations of activity should use MP-PCA with caution, to avoid misinterpretation of activated areas. Below, these aspects are discussed in greater detail.

4.1. Reducing spatial correlations promotes better MP-PCA denoising

We found that MP-PCA applied directly to preclinical vendor reconstructed fMRI data – typically acquired with EPI that requires regridding (Bruder et al., 1992) and other operations that can produce spatial

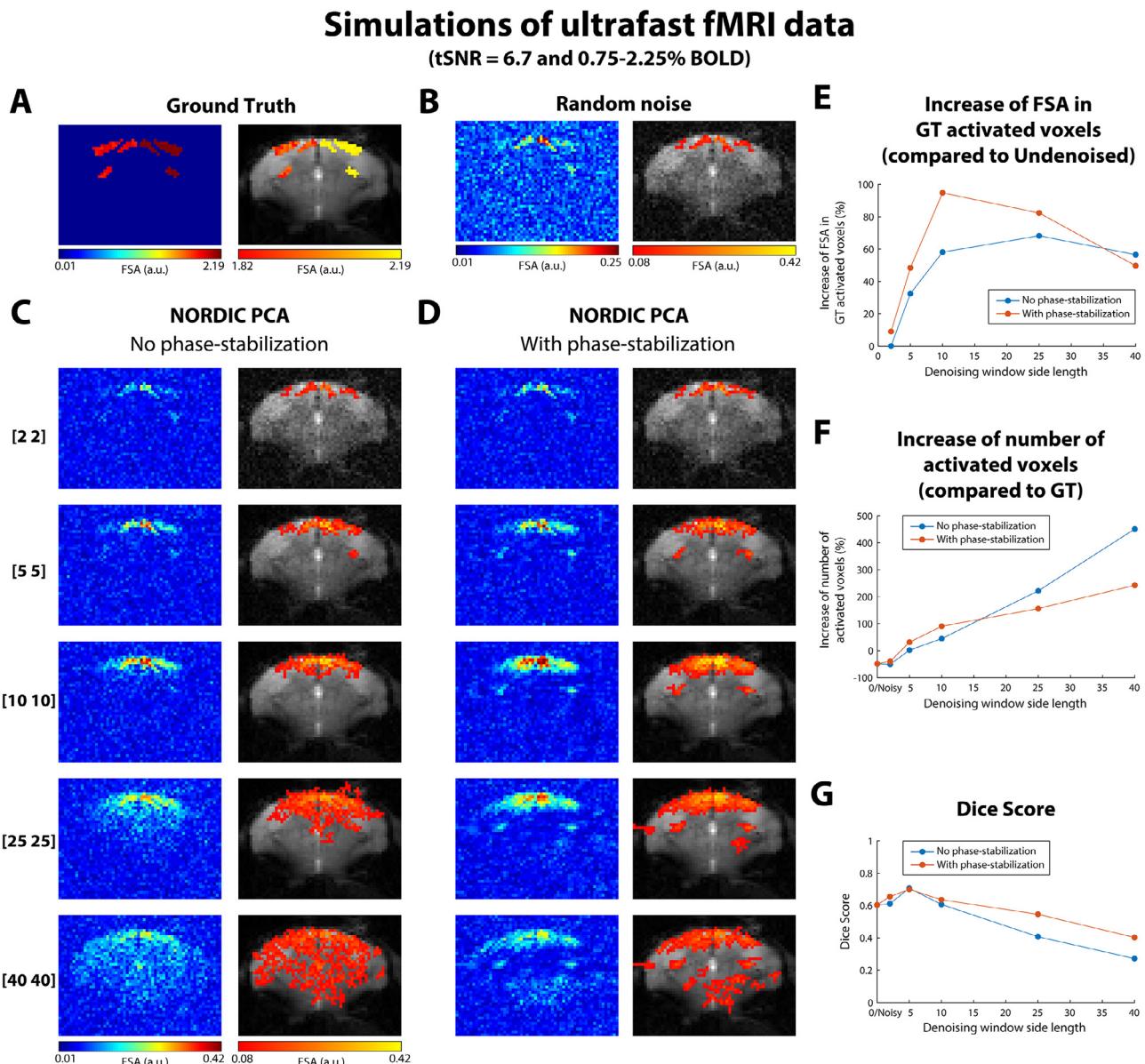


Fig. 11. BOLD maps of simulated ultrafast fMRI data (tSNR = 6.7) after NORDIC PCA denoising. (A) BOLD Fourier map of ultrafast fMRI data simulated without any source of noise, with (right) and without (left) threshold. (B) Same as (A) after addition of Gaussian white noise so that the average tSNR in the brain = 6.7. (C) BOLD Fourier maps with (right) and without (left) threshold upon NORDIC PCA denoising with 5 different sliding windows and no phase-stabilization correction. Maps are thresholded with a minimum sum of FSA at paradigm's fundamental frequency and two following harmonics = 0.08 and a minimum cluster size = 8. A representative image of the data is shown below the maps. (D) Same as (C) but after NORDIC PCA denoising with phase-stabilization correction. (E) Percentage increase of FSA at paradigm's fundamental frequency and two following harmonics in GT activated voxels (i.e., the voxels above the threshold in the BOLD map obtained from GT data) relative to the undenoised data results across 5 different sliding windows. Blue datapoints: no phase-stabilization correction prior to NORDIC denoising. Orange datapoints: phase-stabilization correction applied before NORDIC denoising. (F) Percentage increase of number of activated voxels in the BOLD maps relative to the GT data results across different sliding windows. (G) Dice score of activation maps across different sliding windows.

correlations – (Strategy A, Figs. 2 and 3) provides only marginal improvements in SNR and nearly no improvements in functional mapping, and additionally requires very large windows. Removal of ZF before the denoising step (at the expense of greater phase wrapping in image domain; our Strategy B) produced much better denoising fidelity with more reasonable window sizes, although residual inter-voxel correlations still exist. Therefore, reduction of reconstruction-driven spatial correlations, whether arising from coil combination or reconstruction steps, would be beneficial for application of MP-PCA, as has also been suggested in previous work (J. Veraart et al., 2016). Sequences avoiding spatial interpolation could also be considered in the future to mitigate such effects (Lemberskiy et al., 2019; Lengerskiy et al., 2020). Still, residuals of our

EPI data did not show anatomical details and their histogram approximated the zero-centered normal distribution (Figs. 4B and 5B), suggesting that MP-PCA denoising conservatively suppressed mainly thermal (Gaussian) noise. Moreover, the variance of the residuals was systematically lower than the estimated noise variance (Figs. 4E and 5E), in agreement with previous studies (J. Veraart et al., 2016; J. Veraart et al., 2016) and as expected, since data variability due to noise can only be partially canceled by MP-PCA denoising, due to the noise corruption of the retained P components. We also note that the complex/magnitude nature of the data was not critical here, in line with Gudbjartsson and Patz (1995) who have shown that for SNR larger than two, the Rician noise distribution approximates to a Gaussian distribution. Here, both

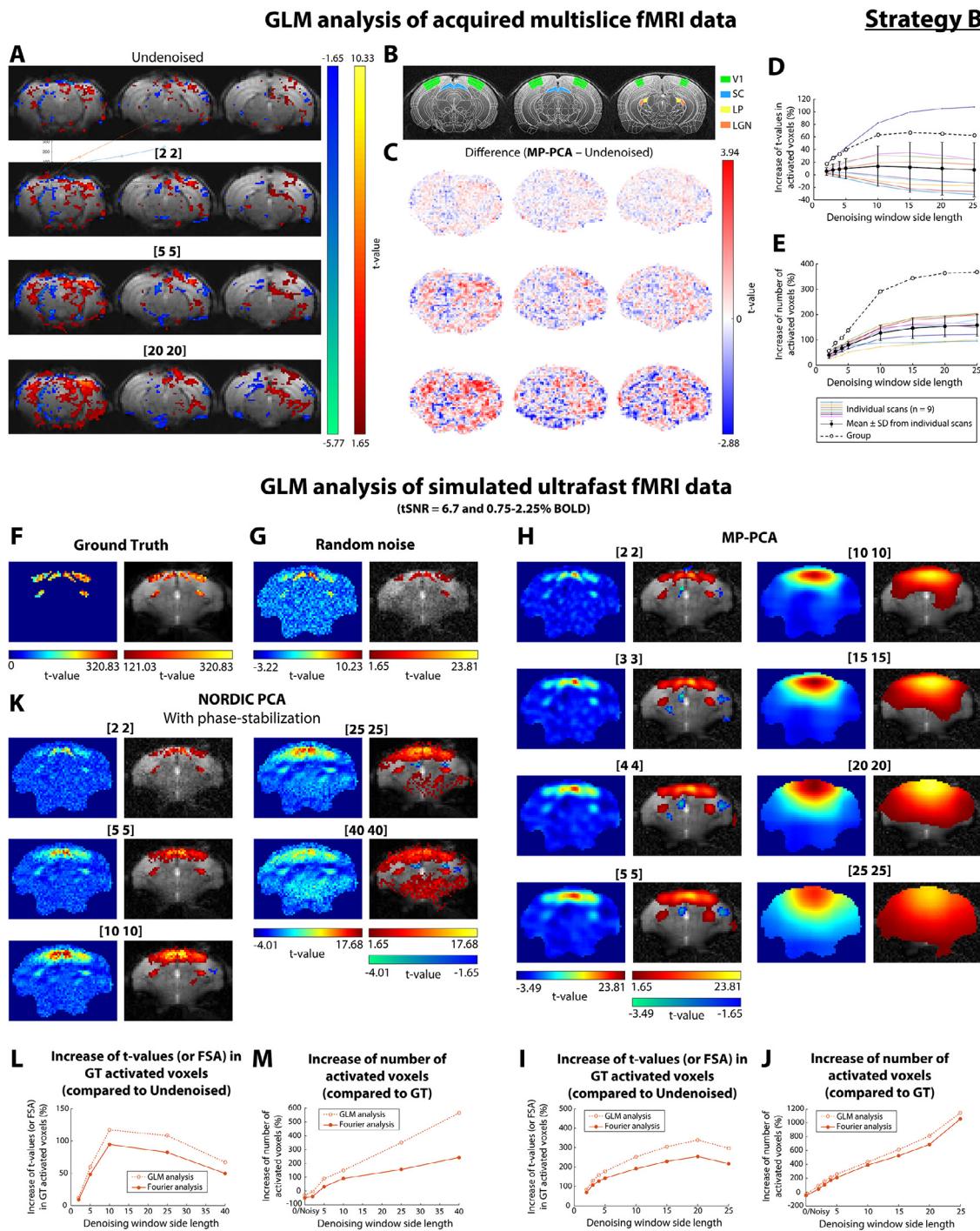


Fig. 12. Individual GLM BOLD maps of acquired multislice and simulated ultrafast fMRI data after MP-PCA denoising. (A) BOLD maps of a representative multislice fMRI acquisition (3 out of 10 slices shown) obtained before and after MP-PCA denoising at the slice level with 3 different sliding windows ([2 2], [5 5] and [20 20]) and generated by GLM analysis. ZF was removed from k-space prior to data denoising. Maps are thresholded with a minimum t -value = 1.65 and a minimum cluster size = 10. (B) ROIs of the mouse visual pathway from the Allen Reference Atlas delineated on anatomical images. (C) Difference between the functional maps shown in (A). (D) Percentage increase of t -values in activated voxels relative to the undenoised data results across 8 different sliding windows and $n = 9$ different scans (represented by each color plot) of multislice fMRI data. (E) Percentage increase of number of activated voxels in the BOLD maps relative to the undenoised data results across the same sliding windows and scans. The empty dots and dashed lines in (D) and (E) are the results from the group BOLD maps. (F) BOLD map of ultrafast fMRI data (with 0.75–2.25% BOLD changes) simulated without any source of noise, with (right) and without (left) threshold, generated by GLM analysis. (G) Same as (F) after addition of Gaussian white noise so that the average tSNR in the brain = 6.7. (H) Same as (G) after MP-PCA denoising with 8 different sliding windows. A representative image of the data is shown below the maps. Maps are thresholded with a minimum t -value = 1.65 and a minimum cluster size = 8. (I) Percentage increase of t -values (or FSA in case of Fourier analysis results) in GT activated voxels relative to the undenoised data results and (J) Percentage increase of number of activated voxels in the BOLD maps relative to the GT data results across 8 different MP-PCA denoising windows in this simulated ultrafast fMRI dataset. Empty dots with dashed lines and filled dots with solid lines represent the variations observed in maps generated using GLM and Fourier analysis, respectively. (K-M) Same as (H-J), respectively, but for NORDIC PCA denoising (with phase-stabilization) using 5 different sliding windows.

multislice and ultrafast fMRI data were characterized by a sufficiently large SNR to avoid the violation of the Gaussian assumption and performance degradation when denoising magnitude data. However, we emphasize that this result is particular to these datasets, therefore future studies should carry out an SNR analysis before performing MP-PCA denoising of magnitude data, and that phase-stabilization (Moeller et al., 2021) may improve the MP-PCA performance. Denoising of real-valued data is also a possibility, reducing potential systematic bias due to the noise floor in magnitude data (Eichner et al., 2015; Fan et al., 2020). We also note that including an outlier correction step before MP-PCA denoising might influence its performance. However, here only a very small percentage of datapoints was corrected in each scan. Specifically, we did not find any outlier in ultrafast data and, in multislice data, a maximum of 2 outliers/slice/scan (out of 340 time points) were found. Therefore, we do not expect major influences of outlier correction on the performance of denoising and subsequent analyses on our datasets.

4.2. Sliding window size considerations

Another important factor in MP-PCA denoising is the sliding window size. Previous work (J. Veraart et al., 2016) recommended using a window size similar or larger than the number of measurements, i.e., $N \gtrsim M$, for improved performance of MP-PCA denoising on dMRI data. However, in fMRI experiments, when high temporal resolutions are required, (1) a multislice type of acquisition may not always be possible, limiting the maximum denoising window size to the entire 2D matrix size, and (2) many more measurements may need to be acquired to obtain sufficient temporal data. Under these conditions, as in our ultrafast fMRI data (where $N_{\max} = 7104$ and $M = 9000$), it may not be possible to meet this recommendation.

Moreover, MP-PCA assumes that noise is spatially uniform within the sliding window for a precise estimation of the noise threshold. Here, we used a single-channel loop coil for signal reception, thus avoiding coil combination-driven spatial correlation of the noise (Dietrich et al., 2008; Aja-Fernández and Tristán-Vega, 2012; Aja-Fernández et al., 2013). However, even if we could assume that the noise level in our data is constant among all elements of the window and increase N to be closer to M , larger patches can become increasingly heterogeneous, risking different HRFs in the patch, thereby deviating low-rank assumptions (P will increase) and making denoising less efficient. Diao et al. (2021) recommended choosing a window within which BOLD fluctuations are likely to be correlated, thus avoiding the removal of genuine BOLD fluctuations. Hence, smaller windows may be favored when possible.

Interestingly, different sliding window size dependences were observed between multislice and ultrafast datasets in terms of SNR gain and estimated noise variance in the residuals. Whereas these metrics monotonically increased with larger window sizes in multislice data (Fig. 4D,E), they showed a U-shaped behavior in ultrafast data, only starting to increase for windows with $N > 100$ (Fig. 5D,E). The convex appearance of these plots is mostly due to the larger difference between the Casorati matrix dimensions, i.e., the higher value of M used in ultrafast data (9000) in comparison to only 340 repetitions in multislice data for the same values of N , and to the number of “signal” components P required for each type of data. In both data types the sum of

noise eigenvalues ($\sum_{i=P+1}^{M'} \lambda_i$), the number of retained “signal” components P and the number of eliminated “noise” components $M' - P$ monotonically increased with larger window sizes (until the [20 20] sliding window) (Fig. S1D-F,J-L). However, the estimated noise variance σ_r^2 , the expected variance of residuals σ_r^2 and the expected SNR gain derived from the MP-PCA denoising theory also strongly depend on the ratios obtained from the sizes of the original and denoised Casorati matrices. In particular, the decreasing trend of $1/(M' - P)(N' - P)$ counterbalances the increasing trend of the sum of noise eigenvalues (Fig. S1D,J) with larger window sizes and confers a convex shape to the estimated noise variance of both multislice and ultrafast data (Fig. S1A,G).

Whereas this is fully compensated by the monotonically increasing trend of $(M' - P)(N' - P)/(M'N')$ with larger window size in multislice data when estimating the expected variance of residuals σ_r^2 (Fig. S1B), in ultrafast data this ratio is actually convex, thereby giving an accentuated convex shape to the expected variance of residuals (Fig. S1H) and consequently to the estimated SNR gain and the estimated noise variance in the residuals shown in Fig. 5D,E.

When analysing the variation of the estimated σ^2 values across the brain (Fig. S2B,E), we noted that the estimated σ^2 maps showed anatomical features when small sliding windows were employed in both types of data, probably reflecting some physiological noise interference. For larger windows, σ^2 maps became less resolved and smoother due to the larger coverage of these windows. Therefore, when a [20 20] sliding window was used, σ^2 maps did not show any type of anatomical features anymore, but showed slightly lower values on the right and left extremes of the brain, probably due to the stronger influence of lower SNR values from outside of the brain.

Moreover, we noticed some ringing-like structured artifacts in the residual maps when data was denoised with its under-performing windows, i.e., those that were not able to remove much of the estimated noise and therefore resulted in less SNR gain (e.g., window [5 5 1] in Fig. 2C, windows [2 2] and [20 20] in Fig. 3C, and window [5 5] in Fig. 5B), and when window sizes were very small (e.g., window [2 2] in Figs. 4B and 5B). In all of these cases, there was a very small number of components available to explain thermal noise (Fig. S1F,L), either because the number of components required to be retained as “signal” was high (Fig. S1K) or simply because the total number of components was small ($M' \leq 16$). It is possible that in these cases the fitting performance is reduced and the method is not able to properly estimate the right edge of the MP distribution and extract purely random noise, therefore eliminating some signal, which can be expressed as some residual structure in the images. Therefore, the MP theorem might have some difficulty to hold when very small windows are used.

Finally, it is worth mentioning that most MP-PCA or NORDIC PCA denoising studies (J. Veraart et al., 2016; Ades-Aron et al., 2021; Vizioli et al., 2021; Dowdle et al., 2023) were performed on human data, which is usually acquired with isotropic voxels in 3D, and used isotropic 3D patches for denoising. When resolution is isotropic, a 3D patch has a smaller spatial extent (i.e., a smaller maximum radius to the center of the patch) than a 2D patch with the same N , and can therefore achieve a smaller “spreading footprint” upon denoising. Whereas isotropic 3D patches can be adequate for such datasets, in anisotropic resolution data (such as our multislice data), a 3D isotropic window will be highly heterogeneous in the slice dimension with respect to different HRFs and can thus induce a larger volume of artifactual activation. Using anisotropic 3D patches is also possible, but in our multislice fMRI data, the slice dimension would have to be reduced to a point where the first two dimensions would already be close to the size of 2D patches with equal N . Thus, we only presented the spreading effect for 2D patches. However, we note in passing that activation spreading would also occur in the third window dimension if 3D patches had been used.

4.3. Sensitivity enhancement vs. activation “spreading”

We have shown that larger windows ($N \geq 100$) allow a better MP distribution fitting, higher noise removal performance and higher FSA gains than small windows ($N \leq 25$) using Strategy B (Fig. 6D,I). These represent increasing sensitivity towards activation. However, a concomitant increase in spatial extent of activation spreading around the activated regions was also noted in all cases (Fig. 6C,H), representing a degree of decreased “specificity” in the spatial sense. In addition, a smoothing effect of the activation patterns (rather than of the underlying images, i.e., even if the residuals relative to undenoised data were apparently random (Fig. 8C)) was observed.

Our simulations (Figs. 8 and S7) corroborated that these effects are expected, since MP-PCA denoising uses the P signal components of the

eigenspectrum to reconstruct all voxels present in the patch. Components that carry fMRI activation patterns then “bleed” to other areas, whose extent depends on the window size used, as well as the tSNR and fCNR. Moreover, we confirmed that the observed “spreading” effect is not related to averaging of overlapping voxels from multiple patches after denoising (Fig. S9). This spreading effect can be difficult to detect using commonly used residual or temporal autocorrelation analyses, suggesting that simulations may assist in determining the extent of the phenomenon given particular tSNR and fCNR. In line with these concerns, Kay (2022) has recently drawn attention to the possibility of introducing bias when exploring new denoising methods. In particular, dimensionality reduction using PCA yielded variance reduction at the cost of introducing bias into the data: the stronger the dimensionality reduction, the more the variance reduction but also the more systematic deviation from the GT being introduced, in agreement with our results. This bias can lead to incorrect inferences and should therefore be carefully considered when choosing denoising methods, depending on whether sensitivity or spatial specificity are the critical factors in any particular study.

4.4. Generalization to other low-rank approaches

To verify that these effects are not specific only to MP-PCA, our simulations were extended to the more recent NORDIC PCA denoising (Moeller et al., 2021; Vizioli et al., 2021; Dowdle et al., 2023), which has been reported to improve accuracy of functional maps without compromising spatial precision, global image smoothness or functional PSF. Although this method performs better with 3D isotropic windows with $N \approx 11 \cdot M$, which was impossible with our simulated data (and which may pose a practical limitation for some types of datasets), we show that sliding windows as small as $N = 100$ already produce the “spreading” effects but with much less smoothing than the MP approach. Indeed, NORDIC PCA denoising advantageously generated much narrower PSFs (Fig. S8). Additionally, spatial correlations induced by coil combination are partly alleviated by NORDIC’s mapping of spatially varying noise to spatially identical noise through g-factor normalization (Moeller et al., 2021). Since MP-PCA was initially and more widely used for diffusion MRI, it is possible that its thresholding approach might not be optimal for fMRI data (though it has indeed been applied for such data as well (Ades-Aron et al., 2021)) where the number of signal components may differ substantially from dMRI. However, it should be noted that NORDIC uses Monte Carlo simulations to assess the MP-PCA cutoff, which requires some knowledge of the true noise variance in the data, which (a) may not be easy to infer (St-Jean et al., 2020; Bazin et al., 2019) and (b) may be spatially varying due to sample properties. Hence, the two methods offer some trade-offs between objective thresholding, smoothing, and accounting for spatial correlations. Further studies directly comparing these methods may shed more light into specific situations in which either would be beneficial over the other. Moreover, Henriques et al., 2023 recently developed new PCA thresholding approaches to denoise spatially correlated noise in diffusion MRI data that have proven to be more robust and efficient than the standard thresholding approach of the MP-PCA denoising method. Together, these studies suggest that an optimized selection of the PCA threshold will probably assist in some of these fMRI activation spreading effects.

4.5. tSNR and fCNR considerations

Simulations with three distinct levels of tSNR and BOLD response amplitudes showed that the optimal MP-PCA sliding window size for improved specificity of activation mapping (i.e., providing the highest Dice score) varies with data tSNR (Fig. 9C) and fCNR (Fig. 10C), with lower tSNR/fCNR data requiring larger windows (Figs. S10 and S12) and higher tSNR/fCNR data not even needing denoising (Figs. S11 and S13). We note that finding the optimal denoising windows based on Dice score calculations is a threshold-dependent approach, since the Dice score is

obtained from maps thresholded with a specific FSA threshold. However, a brief analysis of the ROC curve obtained after applying different FSA thresholds to the Fourier BOLD maps of simulated data (Fig. S14) showed that the windows maximizing the area under the ROC curve coincided or at least were very close in size to the windows providing the highest Dice score when using the FSA threshold of 0.08, and that this was indeed the most adequate threshold for these datasets, resulting in the same conclusion that the lower the tSNR/fCNR the larger the window required for denoising. This conclusion could explain in part the variations in BOLD maps’ metrics observed between each type of acquisition that was performed in our fMRI experiments (multislice data has higher tSNR than ultrafast data) and between their respective individual scans, which possibly have different percent signal changes of BOLD response due to intra- and inter-mouse variability in task sensitivity (Fig. 6D,E,I,J). In fact, similarly to what was observed between individual scans in our fMRI experiments (Fig. 6D,I), different values of fCNR in the simulations resulted in different values of FSA increase for each of the denoising windows tested (Fig. 10A). Moreover, when the fCNR was the lowest (purple dots in Fig. 10A), it was actually possible to see a decreasing pattern of FSA increase values for windows larger than [10 10], similar to the one reported for some individual scans of multislice fMRI data (Fig. 6D). Given the strong indications for a relationship between fCNR and FSA increase for each denoising window in our acquired data, we tested it through Pearson’s correlation and indeed found a high correlation between these variables for both multislice and ultrafast data (Fig. S3A,B). On the contrary, given the small variation of tSNR between individual scans of the same type ($CV = 5\%$ in multislice data and $CV = 6\%$ in ultrafast data), we did not expect tSNR to have an influence on the different FSA increase values obtained between individual scans. Nevertheless, we also tested this hypothesis and found a weak negative correlation between initial tSNR and FSA increase values for each sliding window (Fig. S3C,D). Moreover, according to our simulations (Fig. 9A), in case there would be a relationship between these variables, tSNR would have a positive correlation with FSA increase values, not negative.

Although only variations of tSNR and fCNR have been explored in the simulations, as well as the effect of ZF removal in real data, we do not exclude the possibility that other factors may also affect the efficiency of denoising, the rate of activation spreading and therefore the optimal MP-PCA sliding window size. Structured non-white noise (e.g., physiological noise) can have an impact on low-rank denoising performance, as these additional fluctuations will entail additional signal components in the PCA domain and make data less sparse in the PCA domain. One example of this effect is the fact that even though we set the same FOV, spatial and temporal resolution, repetitions, paradigm, tSNR and fCNR in the simulations that we used in the acquired ultrafast fMRI datasets, there were different spreading behaviors between their results (Fig. 6I,J versus Figs. 9A,B and 10A,B), probably due to the presence of physiological noise in real data. The simulations’ goal was not to replicate exactly every experimental scenario, but rather to show that the activation spreading is expected even under completely controlled conditions, reason by which any type of structured noise was not included in these datasets. Moreover, we note that the local bleeding effect reported in this work may not only involve “active” principal components but also any other retained P signal components. Therefore, physiological noise can also be locally spread. However, the identification and quantification of such an effect would require other targeted analyses that are out of the scope of this work. We do not expect, however, an increase in the level of physiological noise upon denoising, as these methods are only focused on the reduction of thermal noise and therefore are only able at most to make other components easier to detect, not increase their absolute level.

Another factor that can affect the capability to detect BOLD responses and therefore the rate of activation spreading and the optimal MP-PCA window size is the stimulation paradigm. Although different paradigms were used in the multislice and ultrafast fMRI experiments

in this work, it is hard to conclude how the paradigm difference could have contributed to the differences observed between datasets, as the temporal resolution and the number of harmonics inspected were also different between these datasets. We note, however, that for the same tSNR, fCNR and temporal resolution, a longer stimulation period will in general produce responses that are more easy to detect through Fourier or GLM analyses, therefore requiring a lower degree of denoising (i.e., smaller windows), and that paradigms with very different stimulation and rest durations will strongly benefit from a Fourier analysis that inspects more harmonics (rather than the fundamental frequency only) of the paradigm's spectrum.

4.6. FSA vs. GLM analyses

Another aspect that warrants further discussion is the method chosen for activation mapping. Here, we mainly used the data-driven Fourier spectral analysis method since it takes advantage of the periodicity of the paradigm to detect active brain areas without having to assume or have a-priori knowledge of the HRFs (Gil et al., 2021; Nunes et al., 2019). Moreover, this method requires much less computational time than conventional GLM analyses, especially when analyzing data with a very large number of repetitions such as our ultrafast fMRI data. Additionally, GLM analyses can be limited by HRF assumptions (e.g. typically it assumes same HRF for every voxel), which can be far from accurate (Aguirre et al., 1998; Handwerker et al., 2004). Although this does not pose a major problem for block design paradigms with long stimulation periods due to BOLD response saturation, when the blocks are very short, as in our ultrafast fMRI data, the use of a single HRF shape can have a major impact in the detection of activated areas.

Despite the differences between these methods, their estimated metrics (FSA and GLM-driven *t*-values) similarly depend on the noise level of the data: the FSA values due to the previous z-score normalization of the data, and the *t*-values due to the GLM canonical standard error. Under the MP-PCA assumption that absolute signal values are not affected by denoising, it was therefore expected that both of these measures would have a similar trend to SNR gain with larger window sizes. We note, however, that for some denoising windows and scans the FSA and *t*-values on activated voxels actually decreased, which we attribute to the smoothing effect caused by denoising on the absolute signal values. Although in theory the signal values should not be affected by denoising, by performing simulations we found that GT non-activated voxels can exhibit artifactual responses to stimulation with larger denoising windows (as shown in the single-voxel time-courses of Fig. S7E,F). Just as the smoothing effect adds artifactual activity to non-activated voxels, it is also likely that it can reduce activity in GT activated voxels, thus counteracting the increase in FSA/*t*-values generated by thermal noise reduction. Moreover, the artifactual responses generated in GT non-activated voxels keep the same shape and timings/frequencies as the neighboring GT activated voxels. Therefore, even though Fourier or GLM-based methods differ in the way they test if a voxel is active, both of them should be able to detect this type of spurious activation.

As expected, the spatial extent of activation of BOLD maps produced by GLM analysis of multislice (Fig. 12A,E) and simulated (Fig. 12H,J) data increased with larger denoising windows. Therefore, we confirm the generalizability of our FSA-based conclusions to general fMRI studies employing conventional GLM analyses. We note that, due to component truncation, low-rank denoising methods alter the dimensionality of the data, resulting in fewer degrees of freedom for the voxel time-courses that are crucial in deriving statistical maps. Therefore, denoising can have an impact on the value and interpretation of statistical maps of functional activation. Given that the estimation of the effective degrees of freedom of an fMRI time-series is still unresolved, not only in the field of denoising (Patel and Bullmore, 2016) but for fMRI in general (e.g., when using high temporal resolution sequences) (Kruggel et al., 2002; Monti, 2011; Chen et al., 2019), we did not attempt to directly address this issue here. Therefore, we avoided statistical *p*-value thresh-

olds and used FSA thresholds instead (and *t*-value thresholds in case of GLM analysis) to simply provide a measure of activation relative to noise. Nevertheless, as a reference, we report the equivalent *p*-value (0.05) for undenoised data. Moreover, to best account for changes in degrees of freedom with increasing denoising window size and achieve valid statistical inferences of the noise model fit parameters, we incorporated temporal filtering of long-range correlations and pre-whitening into the GLM. By using an autoregressive model or the FAST model implemented in SPM12 (which has shown to improve pre-whitening performance in time-series acquired with sub-second TR (Bollmann et al., 2018; Corbin et al., 2018; Olszowy et al., 2019)) to model temporal correlations, we avoided overestimation of *t*-scores and an inflated FPR. We note, however, that the observed increase in spatial extent of activation and false positive rate with larger window sizes may be slightly overestimated due to using the same FSA threshold across all window sizes and not accounting for the changes in degrees of freedom with different window sizes. To achieve the same statistical significance across the datasets, different FSA thresholds should be employed. Indeed, although the chosen FSA threshold of 0.08 has proven to be the most adequate for the optimal denoising windows (i.e., the ones maximizing the AUC) found for each simulation, we also show that, in general, the larger the denoising window, the higher the FSA threshold required to minimize the distance to the optimal classifier in the ROC curve (Fig. S14C). This could explain in part why the AUC did not considerably drop by using larger window sizes in some simulations as opposed to the Dice score.

Finally, data normalization to z-score using measurements of mean and standard deviation of the signal prior to Fourier spectrum calculation may not have been the ideal approach to compare MP-PCA denoised data results, as it assumes data distribution shape is identical between conditions. This may not be true for different denoising windows, as the amount of thermal (Gaussian) noise removed is different between each window and therefore some conditions might show a distribution that is less Gaussian. Although we do not anticipate large deviations from the observed trends of FSA values with increasing window size when using a different analysis methodology (there is always some retention of Gaussian noise in the data), future studies should consider other ways to compare MP-PCA denoising windows in terms of functional activation mapping without data distribution assumptions, such as a simple DC offset removal prior to fast Fourier transform computation and a non-parametric test to statistically compare FSA values at the paradigm's first harmonics with FSA values at higher frequencies of the spectrum.

5. Conclusions

We investigated MP-PCA denoising for task-based fMRI data series, especially in the context of preclinical fMRI. We find that using vendor-reconstructed EPI data reduces denoising performance due to spatial correlations and show a workaround that mitigates the issues. Although MP-PCA and NORDIC denoising provide a higher sensitivity to BOLD responses, we also find significant activation "spreading" and increases in FPR due to the local bleeding of active signal components which depend strongly on data tSNR, fCNR and denoising window size, which should be considered before denoising. Our results bode well for enhancing spatial and/or temporal resolution in future fMRI work, while taking into account these sensitivity/specification trade-offs of low-rank denoising methods.

Data and code availability

All data used in this study is available upon request. MP-PCA denoising code is available on GitHub at <https://github.com/sunenj/MP-PCA-Denoising>.

Declaration of competing interests

None.

Credit authorship contribution statement

Francisca F. Fernandes: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Jonas L. Olesen:** Methodology, Software, Writing – review & editing. **Sune N. Jespersen:** Methodology, Software, Writing – review & editing, Supervision, Funding acquisition. **Noam Shemesh:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Acknowledgements

The authors thank Rita Gil for support with ultrafast data acquisition and Dr. Cristina Chavarriás for advice on simulations. This study was funded in part by the European Research Council (ERC) (agreement No. 679058) and by Champalimaud Foundation. The authors acknowledge the vivarium of the Champalimaud centre for the Unknown, a facility of CONGENTO which is a research infrastructure co-financed by Lisboa Regional Operational Programme ([Lisboa2020](#)), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF) and Fundação para a Ciência e Tecnologia (Portugal) under the project [LISBOA-01-0145-FEDER-022170](#).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2023.120118](#).

References

- Adamczak, J.M., Farr, T.D., Seehafer, J.U., Kalthoff, D., Hoehn, M., 2010. High field BOLD response to forepaw stimulation in the mouse. *Neuroimage* 51 (2), 704–712. doi:[10.1016/j.neuroimage.2010.02.083](#), Jun.
- Adanyeguh, I.M., et al., 2018. Autosomal dominant cerebellar ataxias: imaging biomarkers with high effect sizes. *Neuroimage Clin* 19, 858–867. doi:[10.1016/j.nicl.2018.06.011](#), Jun.
- Ades-Aron, B., et al., 2021. Improved Task-based Functional MRI language mapping in patients with brain tumors through marchenko-pastur principal component analysis denoising. *Radiology* 298 (2), 365–373. doi:[10.1148/radiol.2020200822](#), Feb.
- Adhikari, B.M., et al., 2019. A resting state fMRI analysis pipeline for pooling inference across diverse cohorts: an ENIGMA rs-fMRI protocol. *Brain Imag. Behav.* 13 (5), 1453–1467. doi:[10.1007/s11682-018-9941-x](#), Oct.
- Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. *Neuroimage* 8 (4), 360–369. doi:[10.1006/nimg.1998.0369](#), Nov.
- Aja-Fernández, S., Tristán-Vega, A., 2012. Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction. *Magn. Reson. Med.* 67 (2), 580–585. doi:[10.1002/mrm.23020](#), Feb.
- Aja-Fernández, S., Vegas-Sánchez-Ferrero, G., 2016. Statistical Analysis of Noise in MRI: Modeling, Filtering and Estimation, 1st ed. Springer International Publishing doi:[10.1007/978-3-319-39934-8](#).
- Aja-Fernández, S., Brion, V., Tristán-Vega, A., 2013. Effective noise estimation and filtering from correlated multiple-coil MR data. *Magn. Reson. Imag.* 31 (2), 272–285. doi:[10.1016/j.mri.2012.07.006](#), Feb.
- Bazin, P.L., Alkemade, A., van der Zwaag, W., Caan, M., Mulder, M., Forstmann, B.U., 2019. Denoising High-Field Multi-Dimensional MRI With Local Complex PCA. *Front Neurosci* 13, 1066. doi:[10.3389/fnins.2019.01066](#), Oct.
- Behzadi, Y., Restom, K., Liu, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37 (1), 90–101. doi:[10.1016/j.neuroimage.2007.04.042](#), Aug.
- Bollmann, S., Puckett, A.M., Cunnington, R., Barth, M., 2018. Serial correlations in single-subject fMRI with sub-second TR. *Neuroimage* 166, 152–166. doi:[10.1016/j.neuroimage.2017.10.043](#), Feb.
- Bruder, H., Fischer, H., Reinfelder, H.-E., Schmitt, F., 1992. Image reconstruction for echo planar imaging with non-equispaced k-space sampling. *Magn. Reson. Med.* 23 (2), 311–323. doi:[10.1002/mrm.1910230211](#), Feb.
- Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. *Neuroimage* 154, 128–149. doi:[10.1016/j.neuroimage.2016.12.018](#), Jul.
- Chen, J.E., Polimeni, J.R., Bollmann, S., Glover, G.H., 2019. On the analysis of rapidly sampled fMRI data. *Neuroimage* 188, 807–820. doi:[10.1016/j.neuroimage.2019.02.008](#), Mar.
- Corbin, N., Todd, N., Friston, K.J., Callaghan, M.F., 2018. Accurate modeling of temporal correlations in rapidly sampled fMRI time series. *Hum. Brain Mapp.* 39 (10), 3884–3897. doi:[10.1002/hbm.24218](#), Oct.
- Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C., 2008. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans. Med. Imaging* 27 (4), 425–441. doi:[10.1109/tmi.2007.906087](#), Apr.
- Cuppen, J.J.M., Groen, J.P., Konijn, J., 1986. Magnetic resonance fast Fourier imaging. *Med. Phys.* 13 (2), 248–253. doi:[10.1111/1.595905](#), Mar.
- Dana, H., et al., 2019. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* 16 (7), 649–657. doi:[10.1038/s41592-019-0435-6](#), Jun.
- Diao, Y., Yin, T., Gruetter, R., Jelescu, I.O., 2021. PIRACY: an optimized pipeline for functional connectivity analysis in the rat brain. *Front. Neurosci.* 15, 602170. doi:[10.3389/fnins.2021.602170](#), Mar.
- Dietrich, O., Raya, J.G., Reeder, S.B., Ingrisch, M., Reiser, M.F., Schoenberg, S.O., 2008. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. *Magn. Reson. Imag.* 26 (6), 754–762. doi:[10.1016/j.mri.2008.02.001](#), Jul.
- Dinh, T.N.A., Jung, W.B., Shim, H.J., Kim, S.G., 2021. Characteristics of fMRI responses to visual stimulation in anesthetized vs. awake mice. *Neuroimage* 226, 117542. doi:[10.1016/j.neuroimage.2020.117542](#), Feb.
- Does, M.D., et al., 2019. Evaluation of principal component analysis image denoising on multi-exponential MRI relaxometry. *Magn. Reson. Med.* 81 (6), 3503–3514. doi:[10.1002/mrm.27658](#), Jun.
- Dowdle, L.T., et al., 2023. Evaluating increases in sensitivity from NORDIC for diverse fMRI acquisition strategies. *Neuroimage* 270, 119949. doi:[10.1016/j.neuroimage.2023.119949](#), Apr.
- Eichner, C., et al., 2015. Real diffusion-weighted MRI enabling true signal averaging and increased diffusion contrast. *Neuroimage* 122, 373–384. doi:[10.1016/j.neuroimage.2015.07.074](#), Nov.
- Fan, Q., et al., 2020. Axon diameter index estimation independent of fiber orientation distribution using high-gradient diffusion MRI. *Neuroimage* 222, 117197. doi:[10.1016/j.neuroimage.2020.117197](#), Nov.
- Fang, Z., van Le, N., Choy, M.K., Lee, J.H., 2016. High spatial resolution compressed sensing (HSPARSE) functional MRI. *Magn. Reson. Med.* 76 (2), 440–455. doi:[10.1002/mrm.25854](#), Aug.
- Friston, K.J., Glaser, D.E., Henson, R.N.A., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* 16 (2), 484–512. doi:[10.1006/nimg.2002.1091](#).
- Gil, R., Fernandes, F.F., Shemesh, N., 2021. Neuroplasticity-driven timing modulations revealed by ultrafast functional magnetic resonance imaging. *Neuroimage* 225, 117446. doi:[10.1016/j.neuroimage.2020.117446](#), Jan.
- Gozzi, A., Zerbini, V., 2023. Modeling brain dysconnectivity in rodents. *Biol. Psychiatry* 93 (5), 419–429. doi:[10.1016/j.biopsych.2022.09.008](#), Mar.
- Grubbs, M.S., Thompson, I.D., 2003. Quantitative characterization of visual response properties in the mouse dorsal lateral geniculate nucleus. *J. Neurophysiol.* 90 (6), 3594–3607. doi:[10.1152/jn.00699.2003](#), Dec.
- Grussu, F., et al., 2020. Multi-parametric quantitative in vivo spinal cord MRI with unified signal readout and image denoising. *Neuroimage* 217, 116884. doi:[10.1016/j.neuroimage.2020.116884](#), Aug.
- Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* 34 (6), 910–914. doi:[10.1002/mrm.1910340618](#).
- Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21 (4), 1639–1651. doi:[10.1016/j.neuroimage.2003.11.029](#), Apr.
- R.N. Henrique, A. Ianus, L. Novello, J. Jovicich, S.N. Jespersen, and N. Shemesh, "Efficient PCA denoising of spatially correlated MRI data," *bioRxiv*, p. 2023.03.29.534707, Mar. 2023, doi: [10.1101/2023.03.29.534707](#).
- Hirano, Y., Stefanovic, B., Silva, A.C., 2011. Spatiotemporal evolution of the functional magnetic resonance imaging response to ultrashort stimuli. *J. Neurosci.* 31 (4), 1440–1447. doi:[10.1523/jneurosci.3986-10.2011](#), Jan.
- Huber, L.R., et al., 2021. LayNii: a software suite for layer-fMRI. *Neuroimage* 237, 118091. doi:[10.1016/j.neuroimage.2021.118091](#), Aug.
- Jonckers, E., Shah, D., Hamida, J., Verhoye, M., van der Linden, A., 2015. The power of using functional fMRI on small rodents to study brain pharmacology and disease. *Front. Pharmacol.* 6, 231. doi:[10.3389/fphar.2015.00231](#), Oct.
- Jung, W.B., Im, G.H., Jiang, H., Kim, S.-G., 2021. Early fMRI responses to somatosensory and optogenetic stimulation reflect neural information flow. *Proc. Natl. Acad. Sci. U. S. A.* 118 (11), e023265118, Mar.
- Kafali, S.G., Çukur, T., Saritas, E.U., 2018. Phase-correcting non-local means filtering for diffusion-weighted imaging of the spinal cord. *Magn. Reson. Med.* 80 (3), 1020–1035. doi:[10.1002/mrm.27105](#), Sep.
- Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., Wandell, B.A., 2013. GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Front Neurosci* 7, 247. doi:[10.3389/fnins.2013.00247](#), Dec.
- Kay, K., 2022. The risk of bias in denoising methods: examples from neuroimaging. *PLoS ONE* 17 (7), e0270895. doi:[10.1371/journal.pone.0270895](#), Jul.
- Knoll, F., Bredies, K., Pock, T., Stollberger, R., 2011. Second order total generalized variation (TGV) for MRI. *Magn. Reson. Med.* 65 (2), 480–491. doi:[10.1002/mrm.22595](#), Feb.
- Kruggel, F., Péligrini-Issac, M., Benali, H., 2002. Estimating the effective degrees of freedom in univariate multiple regression analysis. *Med. Image Anal.* 6 (1), 63–75. doi:[10.1016/s1361-8415\(01\)00052-4](#), Mar.
- Lee, H.L., Li, Z., Coulson, E.J., Chuang, K.H., 2019. Ultrafast fMRI of the rodent brain using simultaneous multi-slice EPI. *Neuroimage* 195, 48–58. doi:[10.1016/j.neuroimage.2019.03.045](#), Jul.

- Lein, E.S., et al., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445 (7124), 168–176. doi:[10.1038/nature05453](https://doi.org/10.1038/nature05453), Jan.
- Lemberskiy, G., Baete, S., Veraart, J., Shepherd, T.M., Fieremans, E., Novikov, D.S., 2019. Achieving sub-mm clinical diffusion MRI resolution by removing noise during reconstruction using random matrix theory. in: *Proc. Intl. Soc. Mag. Reson. Med.* 27, 0770.
- Lengerskiy, G., Baete, S., Veraart, J., Shepherd, T.M., Fieremans, E., Novikov, D.S., 2020. MRI below the noise floor. *Proc. Intl. Soc. Mag. Reson. Med.* 28, 3451.
- Lim, D.H., LeDue, J., Mohajerani, M.H., Vanni, M.P., Murphy, T.H., 2013. Optogenetic approaches for functional mouse brain mapping. *Front. Neurosci.* 7, 54. doi:[10.3389/fnins.2013.00054](https://doi.org/10.3389/fnins.2013.00054), Apr.
- Liu, T.T., 2016. Noise contributions to the fMRI signal: an overview. *Neuroimage* 143, 141–151. doi:[10.1016/j.neuroimage.2016.09.008](https://doi.org/10.1016/j.neuroimage.2016.09.008), Dec.
- Manjón, J.V., Carbonell-Caballero, J., Lull, J.J., García-Martí, G., Martí-Bonmatí, L., Robles, M., 2008. MRI denoising using non-local means. *Med. Image Anal.* 12 (4), 514–523. doi:[10.1016/j.media.2008.02.004](https://doi.org/10.1016/j.media.2008.02.004), Aug.
- Manjón, J.V., Coupé, P., Martí-Bonmatí, L., Collins, D.L., Robles, M., 2010. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magnet. Resonanc. Imag.* 31 (1), 192–203. doi:[10.1002/jmri.22003](https://doi.org/10.1002/jmri.22003), Jan.
- Manjón, J.V., Coupé, P., Concha, L., Buades, A., Collins, D.L., Robles, M., 2013. Diffusion weighted image denoising using overcomplete local PCA. *PLoS ONE* 8 (9), e73021. doi:[10.1371/journal.pone.0073021](https://doi.org/10.1371/journal.pone.0073021), Sep.
- Manjón, J.V., Coupé, P., Buades, A., 2015. MRI noise estimation and denoising using non-local PCA. *Med. Image Anal.* 22 (1), 35–47. doi:[10.1016/j.media.2015.01.004](https://doi.org/10.1016/j.media.2015.01.004), May.
- Marčenko, V.A., Pastur, L.A., 1967. Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sbornik* 1 (4), 457–483. doi:[10.1070/SM1967V001N04ABEH001994](https://doi.org/10.1070/SM1967V001N04ABEH001994), Apr.
- McKinnon, E.T., Helpert, J.A., Jensen, J.H., 2018. Modeling white matter microstructure with fiber ball imaging. *Neuroimage* 176, 11–21. doi:[10.1016/j.neuroimage.2018.04.025](https://doi.org/10.1016/j.neuroimage.2018.04.025), Aug.
- Moeller, S., et al., 2021. NOise reduction with distribution corrected (NORDIC) PCA in dMRI with complex-valued parameter-free locally low-rank processing. *Neuroimage* 226, 117539. doi:[10.1016/j.neuroimage.2020.117539](https://doi.org/10.1016/j.neuroimage.2020.117539), Feb.
- Molloi, E.K., Meyerand, M.E., Birn, R.M., 2014. The influence of spatial resolution and smoothing on the detectability of resting-state and task fMRI. *Neuroimage* 86, 221–230. doi:[10.1016/j.neuroimage.2013.09.001](https://doi.org/10.1016/j.neuroimage.2013.09.001), Feb.
- Monti, M.M., 2011. Statistical Analysis of fMRI time-series: a critical review of the GLM approach. *Front. Hum. Neurosci.* 5, 28. doi:[10.3389/fnhum.2011.00028](https://doi.org/10.3389/fnhum.2011.00028), Mar.
- Mosso, J., Simicic, D., Şimşek, K., Kreis, R., Cudalbu, C., Jelescu, I.O., 2022. MP-PCA denoising for diffusion MRS data: promises and pitfalls. *Neuroimage* 263, 119634. doi:[10.1016/j.neuroimage.2022.119634](https://doi.org/10.1016/j.neuroimage.2022.119634), Nov.
- Niell, C.M., Stryker, M.P., 2008. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* 28 (30), 7520–7536. doi:[10.1523/jneurosci.0623-08.2008](https://doi.org/10.1523/jneurosci.0623-08.2008), Jul.
- Niranjan, A., Christie, I.N., Solomon, S.G., Wells, J.A., Lythgoe, M.F., 2016. fMRI mapping of the visual system in the mouse brain with interleaved snapshot GE-EPI. *Neuroimage* 139, 337–345. doi:[10.1016/j.neuroimage.2016.06.015](https://doi.org/10.1016/j.neuroimage.2016.06.015), Oct.
- Niranjan, A., Siow, B., Lythgoe, M., Wells, J., 2017. High temporal resolution BOLD responses to visual stimuli measured in the mouse superior colliculus. *Matters (Zur)* 3 (2). doi:[10.19185/matters.201701000001](https://doi.org/10.19185/matters.201701000001), Feb.
- Nunes, D., Ianus, A., Shemesh, N., 2019. Layer-specific connectivity revealed by diffusion-weighted functional MRI in the rat thalamocortical pathway. *Neuroimage* 184, 646–657. doi:[10.1016/j.neuroimage.2018.09.050](https://doi.org/10.1016/j.neuroimage.2018.09.050), Jan.
- Nunes, D., Gil, R., Shemesh, N., 2021. A rapid-onset diffusion functional MRI signal reflects neuromorphological coupling dynamics. *Neuroimage* 231, 117862. doi:[10.1016/j.neuroimage.2021.117862](https://doi.org/10.1016/j.neuroimage.2021.117862), May.
- Olesen, J.L., Ianus, A., Østergaard, L., Shemesh, N., Jespersen, S.N., 2022. Tensor denoising of multidimensional MRI data. *Magn. Reson. Med.* 1–13. doi:[10.1002/mrm.29478](https://doi.org/10.1002/mrm.29478), Oct.
- Olszowy, W., Aston, J., Rue, C., Williams, G.B., 2019. Accurate autocorrelation modeling substantially improves fMRI reliability. *Nat Commun* 10, 1220. doi:[10.1038/s41467-019-09230-w](https://doi.org/10.1038/s41467-019-09230-w), 2019Mar.
- Pai, V.M., Rapacchi, S., Kellman, P., Croisille, P., Wen, H., 2011. PCATMIP: enhancing signal intensity in diffusion-weighted magnetic resonance imaging. *Magn. Reson. Med.* 65 (6), 1611–1619. doi:[10.1002/mrm.22748](https://doi.org/10.1002/mrm.22748), Jun.
- Patel, A.X., Bullmore, E.T., 2016. A wavelet-based estimator of the degrees of freedom in denoised fMRI time series for probabilistic testing of functional connectivity and brain graphs. *Neuroimage* 142, 14–26. doi:[10.1016/j.neuroimage.2015.04.052](https://doi.org/10.1016/j.neuroimage.2015.04.052), Nov.
- Pradier, B., et al., 2021. Combined resting state-fMRI and calcium recordings show stable brain states for task-induced fMRI in mice under combined ISO/MED anesthesia. *Neuroimage* 245, 118626. doi:[10.1016/j.neuroimage.2021.118626](https://doi.org/10.1016/j.neuroimage.2021.118626), Dec.
- Ratering, D., Baltes, C., Nordmeyer-Massner, J., Marek, D., Rudin, M., 2008. Performance of a 200-MHz cryogenic RF probe designed for MRI and MRS of the murine brain. *Magn. Reson. Med.* 59 (6), 1440–1447. doi:[10.1002/mrm.21629](https://doi.org/10.1002/mrm.21629), Jun.
- Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60 (1–4), 259–268. doi:[10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F), Nov.
- Smith, S.M., Brady, J.M., 1997. SUSAN - a new approach to low level image processing. *Int. J. Comput. Vis.* 23 (1), 45–78. doi:[10.1023/A:1007963824710](https://doi.org/10.1023/A:1007963824710).
- Sonderer, C.M., Chen, N.-K., 2018. Improving the accuracy, quality, and signal-to-noise ratio of mri parametric mapping using rician bias correction and parametric-contrast-matched principal component analysis (PCM-PCA). *Yale J. Biol. Med.* 91 (3), 214 Sep. Accessed: Jan. 20, 2022. [Online]. Available: [/pmc/articles/PMC6153621/](https://PMC6153621/).
- St-Jean, S., de Luca, A., Tax, C.M.W., Viergever, M.A., Leemans, A., 2020. Automated characterization of noise distributions in diffusion MRI data. *Med. Image Anal.* 65, 101758. doi:[10.1016/j.media.2020.101758](https://doi.org/10.1016/j.media.2020.101758), Oct.
- Tian, L., et al., 2009. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* 6 (12), 875–881. doi:[10.1038/nmeth.1398](https://doi.org/10.1038/nmeth.1398), Dec.
- Triantafyllou, C., Hoge, R.D., Wald, L.L., 2006. Effect of spatial smoothing on physiological noise in high-resolution fMRI. *Neuroimage* 32 (2), 551–557. doi:[10.1016/j.neuroimage.2006.04.182](https://doi.org/10.1016/j.neuroimage.2006.04.182), Aug.
- Veraart, J., Fieremans, E., Novikov, D.S., 2016a. Diffusion MRI noise mapping using random matrix theory. *Magn. Reson. Med.* 76 (5), 1582–1593. doi:[10.1002/mrm.26059](https://doi.org/10.1002/mrm.26059), Nov.
- Veraart, J., Novikov, D.S., Christiaens, D., Ades-aron, B., Sijbers, J., Fieremans, E., 2016b. Denoising of diffusion MRI using random matrix theory. *Neuroimage* 142, 394–406. doi:[10.1016/j.neuroimage.2016.08.016](https://doi.org/10.1016/j.neuroimage.2016.08.016), Nov.
- Vizioli, L., et al., 2021. Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging. *Nat. Commun.* 12 (1), 5181. doi:[10.1038/s41467-021-25431-8](https://doi.org/10.1038/s41467-021-25431-8), Aug.
- Wang, L., Sarnai, R., Rangarajan, K., Liu, X., Cang, J., 2010. Visual receptive field properties of neurons in the superficial superior colliculus of the mouse. *J. Neurosci.* 30 (49), 16573–16584. doi:[10.1523/jneurosci.3305-10.2010](https://doi.org/10.1523/jneurosci.3305-10.2010), Dec.
- Welvaert, M., Rosseel, Y., 2013. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI Data. *PLoS ONE* 8 (11), e77089. doi:[10.1371/journal.pone.0077089](https://doi.org/10.1371/journal.pone.0077089), Nov.
- Yu, X., Qian, C., Chen, D., Dodd, S.J., Koretsky, A.P., 2014. Deciphering laminar-specific neural inputs with line-scanning fMRI. *Nat. Methods* 11 (1), 55–58. doi:[10.1038/nmeth.2730](https://doi.org/10.1038/nmeth.2730), Jan.
- Zerbi, V., 2022. On the use of preclinical imaging to explore the principles of brain function in rodent models and their relevance for illnesses of the human mind. *Transl. Psychiatry* 12 (1), 155. doi:[10.1038/s41398-022-01924-y](https://doi.org/10.1038/s41398-022-01924-y), Apr.