Signal Fluctuation Sensitivity: An Improved Metric for Optimizing Detection of Resting-State fMRI Networks

Daniel J. DeDora[1†], Sanja Nedic[1†], Pratha Katti[1], Shafique Arnab[1], Lawrence L. Wald[2,3,4], Atsushi Takahashi[5], Koene R. A. Van Dijk[2,6], Helmut H. Strey[1] and Lilianne R. Mujica-Parodi[1,2,3,5*]

[1] Department of Biomedical Engineering, Stony Brook University School of Medicine, Stony Brook, NY, USA, [2] Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA, [3] Department of Radiology, Harvard Medical School, Boston, MA, USA, [4] Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA, USA, [5] McGovern Institute for Brain Research at MIT, Massachusetts Institute of Technology, Boston, MA, USA, [6] Department of Psychology, Center for Brain Science, Harvard University, Cambridge, MA, USA

Task-free connectivity analyses have emerged as a powerful tool in functional neuroimaging. Because the cross-correlations that underlie connectivity measures are sensitive to distortion of time-series, here we used a novel dynamic phantom to provide a ground truth for dynamic fidelity between *blood oxygen level dependent* (BOLD)-like inputs and fMRI outputs. We found that the *de facto* quality-metric for task-free fMRI, *temporal signal to noise ratio* (tSNR), correlated inversely with dynamic fidelity; thus, studies optimized for tSNR actually produced time-series that showed the greatest distortion of signal dynamics. Instead, the phantom showed that dynamic fidelity is reasonably approximated by a measure that, unlike tSNR, dissociates signal dynamics from scanner artifact. We then tested this measure, *signal fluctuation sensitivity* (SFS), against human resting-state data. As predicted by the phantom, SFS—and not tSNR—is associated with enhanced sensitivity to both local and long-range connectivity within the brain's default mode network.

Keywords: Functional MRI, signal fluctuation sensitivity, resting state connectivity, temporal signal to noise ratio, dynamic phantom, fidelity

INTRODUCTION

Unprecedented investment in functional neuroimaging has ushered in a new era of brain research, in which fMRI's original role in mapping the areas of the brain most "active" under a task, now includes task-free characterization of brain connections and circuits. This evolution implies a fundamental—and yet largely unacknowledged—shift in how we understand *signal* vs. *noise*.

During fMRI's first decade, researchers almost exclusively used stimulus presentation to evoke *blood oxygen level dependent* (*BOLD*) activity in subjects. To identify the relationship between different brain regions and their functional roles, tasks included one or more experimental conditions (*tasks*), as well as a baseline measure absent of stimuli (*rest*). FMRI time-series were then fitted to the expected hemodynamic shape for each condition (*canonical hemodynamic*

response function, or *HRF*). Once fitted, trials for each condition were averaged and used to statistically compare hemodynamic amplitudes for each condition (*contrasts*) across subjects. Contrasts that met statistical thresholds were then represented as activity, producing activation maps. Importantly, the fitting, averaging, and subtraction approach used to analyze task-based data was designed to distinguish between time-series fluctuations originating from two sources. On the one hand, it amplified *fluctuations of interest* corresponding to task-based activation and that therefore were correlated with the experiment's design matrix. On the other hand, it suppressed *fluctuations of nuisance* corresponding to (scanner, physiological) artifact and that therefore were independent of the experiment's design matrix.

In the late 1990's several influential papers (Biswal et al., 1997; Greicius et al., 2003; Salvador et al., 2005) showed for the first time that the brain showed strong and reliable correlations between fMRI time-series even in the absence of a well-defined task (*resting-state connectivity*); more recently, the relationship between correlation-derived networks, and the neuronal events that underlie them, has been identified using fMRI acquired simultaneously with electrophysiological recordings of local field potentials (Logothetis et al., 2012). The fMRI community quickly responded, and today task-free *connectivity* studies— which map connections between brain nodes as defined by correlations between their time-series—comprise over 20% of human neuroimaging studies published every year (Bandettini, 2014). Connectivity analyses include not only those obtained by correlations with a pre-defined region (*seed-based*) but also those that describe graph-theoretic features of the functional connectome (*complex network analyses*) (Bullmore and Sporns, 2009). Together, connectivity studies have contributed a wealth of new human brain data on aging (Damoiseaux et al., 2008), psychiatric (Greicius et al., 2007) and neurological (Bettus et al., 2009) disorders, and injury (Mayer et al., 2011). Resting-state fMRI protocols are easily standardized, require minimal patient compliance, and permit exploratory analyses; as such, they would appear to be well positioned for both clinical neurodiagnostics as well as large-scale international bio-repositories established for epidemiological research.

However, the transition from *activation maps* to *connections between nodes* not only produced a conceptual shift with respect to the role of functional neuroimaging, but also increased dependence upon time-series power spectra (see Section Materials and Methods). The standard measure for establishing the quality of task-based data has been the *contrast-to-noise ratio (CNR)*, defined as the contrast (mean activation level acquired during task minus the mean activation level acquired during rest) divided by the standard deviation of the time-series (Bandettini and Cox, 2000). For task-free designs however, CNR cannot be computed, and thus normally is replaced by the *temporal signal-to-noise ratio* (*tSNR*), defined as the mean of the time-series divided by its standard deviation (Kruger et al., 2001). Intuitively, both CNR and tSNR compare the amplitude of a signal against a background of undesired physiological, thermal, and scanner noise present in all fMRI studies. This manner of conceptualizing what is "signal" vs. what is "noise" makes perfect sense within the context of activation

maps, in which a task activates the brain reliably more under one condition (signal) than another (noise)(Murphy et al., 2007). However, for task-free analyses, the "baseline" fluctuations themselves also include the "signal." Thus, for most task-free analyses, tSNR would appear to do exactly the opposite of what one would wish, as it penalizes sensitivity to the fluctuations (i.e., the standard deviation of the time-series) upon which experimental results are also based. Indeed, several recent studies have reported little correspondence between resting-state tSNR and the detection of stable functional networks (Smith et al., 2013; Welvaert and Rosseel, 2013; Gonzalez-Castillo et al., 2014; Molloy et al., 2014).

For task-free analyses, rather than relegate time-series fluctuations to the category of noise as per tSNR, we want to— as with task-based analyses—functionally distinguish between fluctuations of interest that are *neurobiologically significant* (e.g., emanating from BOLD signal consequent to neuronal response) from fluctuations of nuisance that are *neurobiologically insignificant* (e.g., physiological, scanner, and motion artifact). The dissociation between the two can be characterized by *signal fluctuation sensitivity (SFS)*, which we define at a single-voxel level as:

$$SFS_{voxel} = \frac{\mu_{ROI}}{<\mu_{global}>} \times \frac{\sigma_{ROI}}{<\sigma_{nuisance}>} \tag{1}$$

In the first term, the numerator consists of the mean signal (μ) of a time-series acquired from a voxel in the *region of interest* (*ROI*). For the denominator, we average over all voxel-specific signal for the entire brain (*global*). The first term ensures that SFS decreases for regions with signal drop out, while remaining unit-less (as with tSNR). In the second factor, the numerator consists of the standard deviation (σ) of the time-series acquired from the voxel of interest. For the denominator, we average over all voxel-specific σ from a region in which BOLD signals are not expected, but in which physiological, scanner, and motion artifacts are still present (*nuisance*). Prior work suggests that time-series obtained from cerebrospinal fluid (CSF) meet criteria for the *nuisance* denominator (Wald, 2012). SFS for a region of interest is then computed by averaging voxel-specific SFS values over all voxels in the region ($SFS_{ROI} = <SFS_{voxel}>_{ROI}$). In order to more easily compare SFS with tSNR, we scale them comparably by multiplying SFS values by 100.

We define *dynamic fidelity* as the degree to which fMRI accurately captures true BOLD fluctuations. In order to test our hypothesis that SFS should reflect dynamic fidelity of time-series more accurately that tSNR, we first needed to know the "ground truth" for those fluctuations. To access that ground truth, we designed and constructed a *dynamic phantom*, which provides user-controlled—and thus known—dynamic BOLD-like inputs to which fMRI-derived outputs can be compared. We then tested the impact of dynamic fidelity, as defined by our phantom, in predicting detection sensitivity to functional connectivity in human data across three different sets of acquisition parameters, chosen to represent a breadth of realistic optimization strategies utilized within the neuroimaging field for human connectivity studies.

MATERIALS AND METHODS

Design and Prototyping of Dynamic Phantom

Summary of Strategy

We designed a dynamic phantom that is fully automated, capable of producing complex waveforms detected by fMRI, and contains no paramagnetic materials. The basis for the signal is the fact that the magnetic susceptibility of agarose gels is concentration dependent (Olsrud et al., 2008), in which higher concentrations produce lower fMRI signal. By varying the concentration of agarose gel present within a voxel over time, the dynamic phantom produces changes in T2* that can be tuned to amplitudes typically seen with BOLD in humans; the phantom can be programmed to simulate both task-based and resting-state BOLD-like signals. With known inputs, the relationship between

the signal produced (by the dynamic phantom) and the signal detected (by the fMRI scanner) can be rigorously quantified as a measure of *dynamic fidelity*.

The dynamic phantom is composed of calibrated agarose gels housed within two concentric cylinders. The outer cylinder contains a *baseline* gel, while the inner cylinder is longitudinally divided with both (i) a baseline gel matching the outer cylinder and (ii) an *active* gel with slightly lower concentration of agarose (**Figure 1**). The longitudinally divided inner cylinder produces dynamic fMRI signal via rotation about its long axis. We developed a novel fMRI-compatible pneumatic motor to drive rotation of the inner cylinder, while the outer cylinder remains motionless. The position of the inner cylinder is continuously monitored with a fiber optic feedback system, and the device is operated from the fMRI control room with a microcontroller. Compressed air drives rotation of the inner cylinder, and
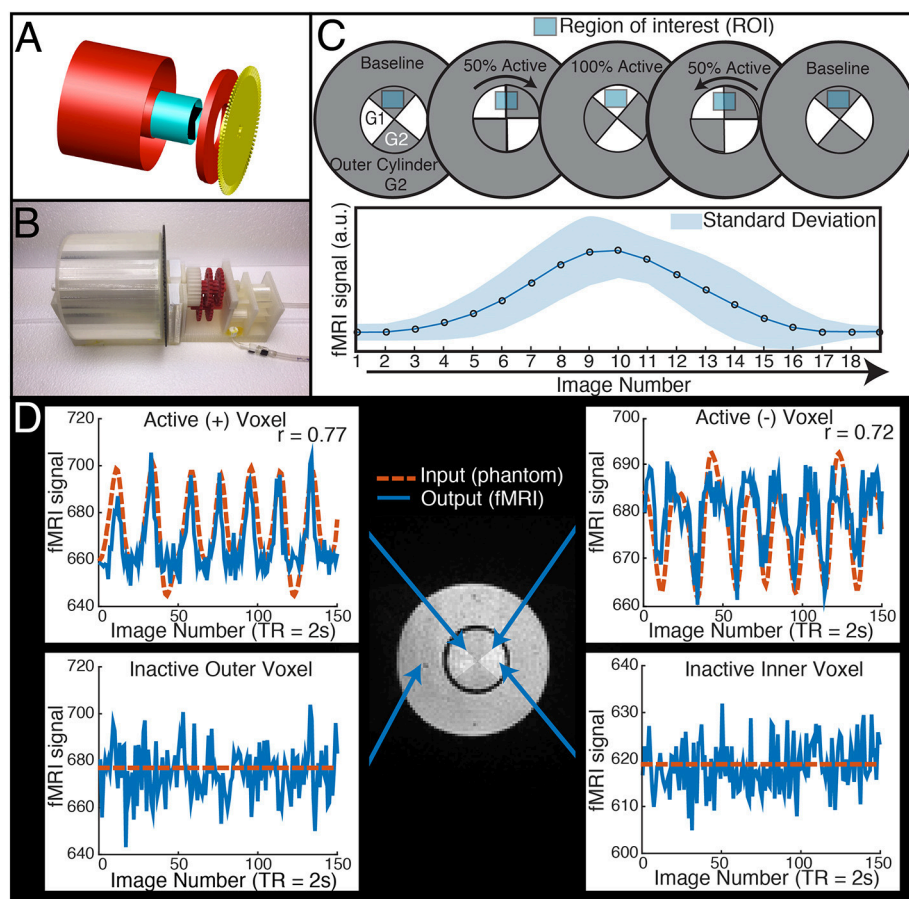

FIGURE 1 | The dynamic phantom produces tightly controlled changes in functional MRI signal, establishing a ground truth for quantifying dynamic fidelity of scanner outputs to signal inputs. (A,B) The dynamic phantom uses concentric cylinders filled with agarose gels. The inner cylinder is coupled to an MRI-compatible pneumatic motor and fiber optic feedback system. **(C)** The inner cylinder is longitudinally compartmentalized into four chambers. One of two calibrated agarose gels with different concentrations is contained in each; the gels are in direct contact. The outer cylinder contains a single agarose gel. Because magnetic susceptibility changes as a function of agarose concentration, precisely timed rotation of the inner cylinder between images creates a "gradient" effect, in which different proportions of each agarose compartment pass through—and are averaged over—a region of interest. Motion across the "gradient" thus is capable of producing smooth dynamic changes in fMRI signal (bottom panel of **C**). **(D)** The top two panels demonstrate "active" voxels within the inner cylinder of the phantom along the gel-gel interfaces; these voxels exhibit strong input-output fidelity. The bottom two panels show that the inactive outer cylinder and inactive inner cylinder voxels are indistinguishable. For validation of phantom performance, a simple event-related design is pictured in D. During the phantom scanning for SFS experiments, the phantom utilized a more complex input mimicking human resting-state fluctuations (**Figure 4A**).

monitoring of the phantom occurs through plastic fiber optic cables, which run between the scanner and control room. Thus, the dynamic phantom is comprised of two main systems: (1) the *scanned phantom*, consisting of two concentric cylinders and supports, a plastic gearbox, tubing, and fiber optic cables, and (2) the *control unit*, consisting of a microcontroller, compressor, and circuit board. The description of the design will be broken down within these two systems and their interface.

System 1: Scanned Phantom
Phantom Housing
We used AutoCAD (AutoDesk, Inc) to design a cylinder-within-cylinder phantom (**Figure 1**). The inner cylinder contains four compartments, divided longitudinally. All custom phantom parts were printed with a Makerbot 3D printer with non-pigmented polylactic acid filament (Makerbot, Inc, Brooklyn, NY). The volume of the outer cylinder was 600 mL, while the volume of the inner cylinder was 150 mL.

Agarose Gels and Materials Justification
We chose agarose as a contrast medium because of its relative ease of use, flexibility in preparation, and physical stability. The use of agarose in phantom construction has been validated throughout the literature, and it is shown to be homogenous with respect to MR relaxation properties (Christoffersson et al., 1991; Olsrud et al., 2008). The outer cylinder was filled uniformly with 2.27% agarose. The inner cylinder was filled with 2.21 and 2.27% agarose gels (**Figure 1C**). No dividing materials were used, i.e., the gels were in direct contact. We used a "baseline" agarose gel concentration approximately matching previous agarose phantom development (Olsrud et al., 2008). We then calibrated the "active" agarose gel concentration by empirically measuring fMRI signal intensity at 3T for agarose concentrations between 1.5 and 3.0%, and chose the concentration that produced an approximate 1% signal change in a 3T 12 channel MRI during a simulated event-related design. Gels were degassed with a vacuum chamber.

Interface between System I and System II
Control and Automation
To achieve automated rotation of the inner cylinder, we designed and fabricated a fully MRI-compatible pneumatic motor system. The motor consists of a compressor, valves, manifold, tubes, dual fans, and a gearbox. An air compressor is placed in the control room of the MRI center; input pressure is set to 40 pounds per square inch at 1.9 cubic feet per minute. Plastic tubing guides the compressed air through a splitter and into two Arduino controlled solenoid valves (SparkFun Electronics, Boulder CO). Compressed air leaves the two independent valves and is guided through two tubes into the scanner bed. The compressed air is released from the pairs of tubes via pneumatic connectors, resulting in high velocity airflow. Depending on which valve is open, this airflow powers one of two fans; these fans are coupled to a gearbox and spin in opposing directions. The dual fan setup allows the gearbox to be driven in either direction and also allows precise braking. The rapid rotation of the fans is stepped down and torque is increased via five 3:1 compound gears, resulting

in a step down ratio of 243:1. The gearbox ultimately interfaces with the inner cylinder and optical interruption disk to produce pneumatically controlled rotation. The outer cylinder does not rotate.

Fiber Optic Feedback
We designed a fiber-optic feedback system using plastic fiber-optic cables, an LED light source, a photodiode, and an interrupter disc. An Arduino microcontroller powers a high-powered 10 mm LED (SparkFun Electronics, Boulder CO), which is coupled with a 1.5 mm diameter fiber optic cable (Thor Labs, NJ). The first cable guides light from the LED source within control room to the scanner bed through a waveguide. The fiber optic cables are positioned opposite each other and spaced 5 mm apart, such that as the inner cylinder rotates, the interrupter disk (3 mm thickness) will intermittently block light transmission between the two cables. The interrupter disc has 60 teeth, corresponding to ~6° of rotation per interrupt. The second fiber optic cable receives light and is fed back to a photo-diode on the microcontroller. As the interrupter disc spins, the photo-diode receives differential intensity readings. The microcontroller then displays the interruption count as a live feed at each TR. We calibrated the phantom to traverse an average of one interrupt per image. Prior to each fMRI scan, the device performs a self-calibrating procedure to ensure optimal position encoding regardless of ambient light.

System II: Control Unit
Arduino Microcontroller and fMRI Communication
TR signals are sent to the Arduino via USB input from the MRI scanner. To properly calibrate the phantom rotation and avoid motion artifacts in regions of interest, we ran a simple EPI acquisition (*TR* = 2s, *TE* = 30 ms, 25 slices) in which the phantom began rotation just after the start of each TR, and examined each slice for motion artifacts. We found that motion artifacts occurred when the phantom rotated during or before a slice was acquired, whereas slices acquired before the phantom rotates within a *TR* contained no noticeable artifact. Because the phantom rotates in plane with the image, no material leaves or enters the imaging slice; this feature avoids potential spin history artifacts. Therefore, if the phantom is programmed to begin rotation towards the end of a *TR* (after a sufficient number of slices have been acquired) and to stop rotation *just before* the next TR, motion artifacts are mitigated (see **Figure 1** to illustrate, and **Figure 2** for a representative time-series acquired with the dynamic phantom). Empirical testing with this design indicated that the phantom should begin rotation 650 ms prior to each TR, and stop ~100 ms before the TR. Thus, for *TR* = 2 s, the dynamic phantom begins rotation at 1500 ms and ends at 1900 ms; for *TR* = 1080 ms, rotation begins at ~600 ms and ends at ~980 ms; for *TR* = 802 ms, rotation begins at ~300 ms, and ends at ~700 ms. This strategy produced minimal motion artifacts in images of the center of the phantom, where inactive inner cylinder voxels (which experience motion) and inactive outer cylinder voxels (which experience no motion) showed no significant differences in standard deviation ($p = 0.89$, rank sum test).

Arduino Software

The dynamic phantom is controlled with an Arduino Mega (www.Arduino.cc). We developed all software in-house. The phantom can operate in three distinct modes: (1) *stimulus-driven* (for simulation of task-designs), (2) *guided-mode* (for simulation of resting-state), and (3) *static*. For this experiment, the phantom utilized *guided mode*, for which the user preprograms the interruption destination for each image. This allowed for the production of specific time-series, such as a pink-noise time-series equivalent to those produced by resting-state fMRI.

Using the Dynamic Phantom to Test SFS and tSNR

Acquisition Parameters

We scanned the dynamic phantom in three separate MRI scanners. Detailed scan parameters are listed in **Table 1**. The three scanners utilized in this phantom study represent the following: (i) a 3T Siemens MRI with 32-channel head-coil (McGovern Institute for Brain Research, Massachusetts Institute of Technology), (ii) a 3T Siemens MRI with 64-channel head-coil (Human Connectome Scanner—Martinos Center for Biomedical Engineering, Massachusetts General Hospital), and (iii) a 7T Siemens MRI with 32-channel head-coil (Martinos Center for Biomedical Engineering, Massachusetts General Hospital). For

each scanner, we tested three sampling rates, representing typical time-resolution for fMRI studies ($TR = 2000$–2010 ms), increased time-resolution acquired for the Human Connectome Project ($TR = 1010$–1080 ms), and ultra-fast imaging paradigms ($TR = 802$–824 ms). Thus, we performed a factorial study (three scanners and three sampling rates each) with the dynamic phantom, for a total of nine scans (**Table 1**), each 10 min long. For both 3T scanners, we performed standard shimming; due to dramatically increased susceptibility artifacts at 7T, we utilized a partial shim centered on the inner cylinder of the phantom. Visual inspection of the resulting images, as well as correlations between the dynamic phantom inputs and fMRI outputs, confirmed data quality.

Statistical Analyses

While most human fMRI data undergoes significant preprocessing, for the dynamic phantom we used raw data after implementing only voxel-wise trend removal (linear and quadratic) to remove scanner drift, and no further temporal preprocessing, in order to characterize dynamic fidelity as transparently as possible. For the *region of interest* (*ROI*) fluctuations, we extracted the average time-series from the four quadrants of the inner cylinder (corresponding to the four chambers, with respect to the initial position of the phantom) with an automated masking procedure using MATLAB software
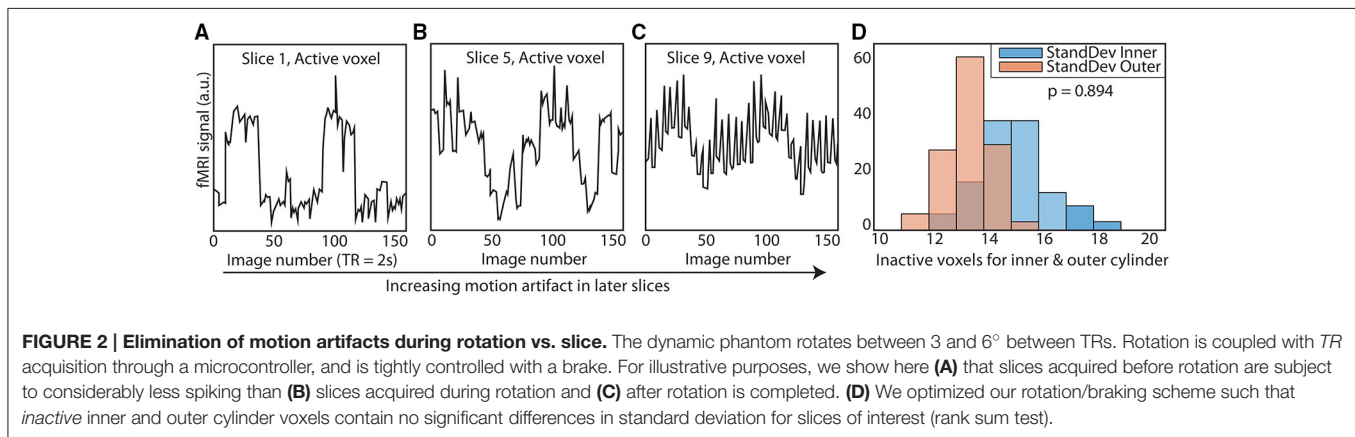

FIGURE 2 | Elimination of motion artifacts during rotation vs. slice. The dynamic phantom rotates between 3 and 6° between TRs. Rotation is coupled with *TR* acquisition through a microcontroller, and is tightly controlled with a brake. For illustrative purposes, we show here **(A)** that slices acquired before rotation are subject to considerably less spiking than **(B)** slices acquired during rotation and **(C)** after rotation is completed. **(D)** We optimized our rotation/braking scheme such that *inactive* inner and outer cylinder voxels contain no significant differences in standard deviation for slices of interest (rank sum test).

TABLE 1 | Acquisition parameters for the nine dynamic phantom scans.

Scan	Main field	Head coil	*TR* (ms)	*TE* (ms)	SMS	iPAT	Flip angle (°)	Bandwidth (Hz/Px)	Resolution (mm)	Slice Gap (mm)	Slices
1	3T	32 Ch	2000	30	3	2	75	1860	2 × 2 × 2	0.2	69
2	3T	32 Ch	1080	30	4	2	60	1860	2 × 2 × 2	0.2	60
3	3T	32 Ch	802	30	5	2	33	1860	2 × 2 × 2	0.2	55
4	3T	64 Ch	2000	30	2	2	85	2840	2 × 2 × 2	0	62
5	3T	64 Ch	1080	30	4	2	60	2840	2 × 2 × 2	0	68
6	3T	64 Ch	824	30	5	2	55	2840	2 × 2 × 2	0	65
7	7T	32 Ch	2010	20	2	2	33	2264	2 × 2 × 1.5	0	86
8	7T	32 Ch	1010	20	4	2	55	2264	2 × 2 × 1.5	0	84
9	7T	32 Ch	802	20	5	2	33	2368	2 × 2 × 1.5	0	85

We tested three scanners at three TRs with the dynamic phantom. Highlighted rows indicate scans for which we collected corresponding human data, in which Scan 1 corresponds to Acquisition A, Scan 5 corresponds to Acquisition B, and Scan 9 corresponds to Acquisition C. All sequences utilized interleaved acquisition.

developed in-house. We repeated this for six slices positioned in the center of the phantom ($n = 24$ time-series per scan). For the *nuisance* fluctuations, we extracted the time-series from the outer cylinder of the phantom, which does not activate. We then computed quadrant-wise SFS based on the definition:

$$SFS_{dynamic\ phantom\ quadrant} = \frac{< \mu_{inner\ cylinder} >_{quadrant}}{< \mu_{global} >} \times \frac{< \sigma_{inner\ cylinder} >_{quadrant}}{< \sigma_{outer\ cylinder} >} \quad (2)$$

In the first term, the numerator consists of the mean signal (μ) of an averaged time-series over each of the four dynamic phantom quadrants (*quadrant*). For the denominator, we average over signal for the entire phantom (*global*). The first term ensures that SFS decreases for regions with signal drop out, while remaining unit-less (as with tSNR). In the second term, the numerator consists of the mean standard deviation (σ) of an averaged time-series over each of the four dynamic phantom quadrants. For the denominator, we average over σ from a region in which signals are not expected, but in which physiological, scanner, and motion artifacts are still present. In this case, we use the outer cylinder, which is static. In order to avoid biasing values for standard deviation due to differences in the number of voxels between inner quadrants and outer cylinder, we averaged time-series in the outer compartment over the same number of voxels used to average time-series in each of the inner quadrants. We computed standard deviations for each of these inner quadrant-sized (39 voxels) averaged time-series, and then averaged across those standard deviations to produce the standard deviation for the entire outer cylinder (i.e., the denominator of the second factor). In order to more easily compare SFS with tSNR, we scale them comparably by multiplying SFS values by 100. TSNR was computed as the mean for the averaged time-series over each of the four dynamic phantom quadrants, divided by its standard deviation (after detrending). Dynamic fidelity was computed as the correlation between inputs (dynamic phantom user-defined function) and outputs (fMRI time-series). We then computed the correlation between fidelity and both SFS and tSNR for each of the 24 time-series per scan.

Human Scanning

Acquisition

In an effort to represent a wide variety of task-free scanning paradigms, we analyzed three sets of human data ($N = 12$ subjects each) collected with the same acquisition parameters utilized for the phantom studies, but using only the time-resolutions previously optimized for each study (**Table 1**). Thus, *Acquisition A* refers to the 3T MRI with 32-channel head-coil and a $TR = 2000$ ms; *Acquisition B* refers to the 3T MRI with a 64-channel head-coil and a $TR = 1080$ ms, and *Acquisition C* represents the 7T MRI with a 32-channel head-coil with a $TR = 802$ ms. *Acquisition A* lasted 5 min, while *Acquisition B* (originally 6.2 min) and *Acquisition C* (originally 10 min) data were truncated to match this duration. Anterior to posterior phase encoding and interleaved acquisition were used in all scans. For *Acquisition A*, we acquired whole-brain T1-weighted

structural volumes using a conventional MPRAGE sequence with the following parameters: $TR = 2530$ ms, $TE = 3.39$ ms, $TI = 1100$ ms, flip angle $= 7°$, voxel size $= 1 \times 1 \times 1.3$ mm. Conventional B0 field maps derived from phase differences between gradient echo images acquired at $TR = 4.22$ and 6.68 ms were also acquired ($TR = 584$ ms, flip angle $= 55°$, voxel size $= 2 \times 2 \times 2$ mm, slice gap $= 0.2$ mm, 69 slices). For *Acquisition B*, we also acquired whole-brain T1-weighted structural volumes using a conventional MPRAGE sequence with $TR = 2530$ ms, $TE = 1.15$ ms, $TI = 1100$ ms, flip angle $= 7°$, 1 mm isotropic voxel size. For *Acquisiton C*, we acquired whole-brain T1-weighted structural volumes using a multi-echo MPRAGE (MEMPRAGE) sequence with four echoes and the following protocol parameters: $TR = 2530$ ms, $TE1 = 1.61$ ms, $TE2 = 3.47$ ms, $TE3 = 5.33$ ms, $TE4 = 7.19$ ms, $TI = 1100$ ms, flip angle $= 7°$, 1 mm isotropic voxel size. Conventional B0 field maps derived from phase differences between gradient echo images acquired at $TE = 4.60$ and 5.62 ms were also acquired ($TR = 723$ ms, flip angle $= 36°$, voxel size $= 1.7 \times 1.7 \times 1.5$ mm, 89 slices).

All human subject studies were approved by the Institutional Review Boards of institutions at which subjects were tested (Massachusetts Institute of Technology for *Acquisition A*, Massachusetts General Hospital for *Acquisitions B* and *C*). All subjects were healthy adults, had full capacity, and provided informed consent.

All subjects were age matched ($\mu_A = 25.6 \pm 3.7$; $\mu_B = 23.3 \pm 4.2$; $\mu_C = 25.6 \pm 3.4$, $p = 0.35$, Kruskall-Wallis test). There were no significant differences in motion across the three groups (maximum absolute translation $p = 0.60$, maximum absolute rotation $p = 0.96$, mean root mean square (RMS) motion $p = 0.10$, maximum RMS motion $p = 0.27$, Kruskall-Wallis test). All participants were instructed to lie quietly with eyes open in the scanner, orienting to a fixation cross, without moving for the duration of the scan. We removed the first 10 s of data for all datasets.

Preprocessing

We followed the standard SPM 8 pipeline for realignment, co-registration to a structural image, and normalization to Montreal Neurological Institute (MNI) space. Co-registered structural images were segmented into probabilistic maps of gray matter, white matter, and CSF using SPM's New Segment tool. Where noted, we utilized a 4-mm (2 voxel) FWHM Gaussian smoothing kernel. As per standard practice for fMRI analyses, we performed slice time correction only on *Acquisition A* data, since the 2000 ms sampling rate was considerably slower than those of the other two scanners. We performed field map correction on *Acquisitions A and C* (distortion correction scheme was performed on *Acquisition B* immediately following image acquisition). Scrubbing was performed to remove the influence of motion, with scan-to-scan global signal deviation from the mean >3 and scan-to-scan composite motion >0.5 mm as thresholds for removal (Power et al., 2012). The mean percentage of data points removed between all three groups was 1.97%, with no subjects having more than 9% of data scrubbed. To assess the impact of spatial smoothing, we computed all of our measures on

both unsmoothed and smoothed data, both of which underwent each of the other preprocessing steps listed here.

Computation of SFS, tSNR, ALFF, ReHo and Long-Range (mPFC-PCC) Connectivity

We used MATLAB to compute voxel-wise SFS according to:

$$SFS_{voxel} = \frac{\mu_{default\ mode\ network\ ROI}}{<\mu_{global}>} \times \frac{\sigma_{default\ mode\ network\ ROI}}{<\sigma_{cerebrospinal\ fluid}>} \quad (3)$$

In the first term, the numerator consists of the mean signal (μ) of a time-series acquired from a voxel in the *region of interest* (*ROI*) in the default mode network, as defined below. For the denominator, we average over all voxel-specific signal for the entire brain (*global*). The first term ensures that SFS decreases for regions with signal drop out, while remaining unit-less (as with tSNR). In the second term, the numerator consists of the standard deviation (σ) of a time-series acquired from the voxel of interest in the default mode network, as defined below. For the denominator, we average over all voxel-specific σ from a region in which BOLD signals are not expected, but in which physiological, scanner, and motion artifacts are still present (*nuisance*). Prior work suggests that time-series obtained from cerebrospinal fluid (CSF) meet criteria for the *nuisance* denominator (Wald, 2012). SFS for a region of interest is then computed by averaging voxel-specific SFS values over all voxels in the region ($SFS_{ROI} = < SFS_{voxel} >_{ROI}$). We additionally computed voxel-wise tSNR as the mean for each voxel's time-series divided by its standard deviation after detrending. In order to more easily compare SFS with tSNR, we scale them comparably by multiplying SFS values by 100.

For SFS, standard deviations of the cerebrospinal fluid voxels (*nuisance fluctuations*) were computed using an eroded probabilistic map of CSF (SPM8 segmented map of CSF thresholded at 70%), to ensure minimal contributions from neural sources. To avoid distorting time-series dynamics by averaging them, standard deviations were computed for each voxel in the nuisance ROI, with voxel-based values averaged for the ROI. Mean global signal included the entire brain (conjunction of gray matter, white matter, and cerebrospinal fluid, thresholded at 70%). Mean values and standard deviations for each voxel were acquired before confound correction, but after SPM8 preprocessing and scrubbing.

Prior to functional connectivity analyses, we performed further regression of nuisance variables (confound-correction). This included detrending, regression of mean CSF and white matter signals (white matter map thresholded at 70%), and regression of six motion parameters from the realignment step. Finally, we performed temporal band-pass filtering in the 0.01–0.1 Hz range using 5th order Butterworth filter.

Both *amplitude of low frequency fluctuations* (*ALFF*) and *local synchronization of neighboring voxels* (*regional homogeneity* or ReHo: 27-voxel KCC-ReHo) were computed from confound-corrected data, using the REST toolbox (Song et al., 2011). Resulting subject-specific voxel-wise ReHo and ALFF maps were standardized by dividing each voxel's value by the mean value of the whole brain.

To test whether SFS or tSNR were predictive of these established resting-state measures, we computed within-subject correlations between (i) SFS and ALFF, (ii) SFS and ReHo, (iii) tSNR and ALFF, and (iv) tSNR and ReHo for voxels belonging to the well-established *default mode network* (DMN) regions: *medial prefrontal cortex* (*mPFC*), *posterior cingulate cortex* (*PCC*), and *left and right lateral parietal cortices* (*LLP* and *RLP*). These regions were defined as 10-mm radius spheres centered on previously established coordinates (Fox et al., 2005), intersected with an SPM8 brain mask to ensure only brain voxels were included (Fox et al., 2005). For the extraction of ROI-based SFS and tSNR values, we used the four aforementioned DMN masks, as well as a probabilistic gray matter mask from SPM8 ($P >$ 50%). We obtained subcortical ROI masks from bilateral regions included in FSL Harvard-Oxford subcortical atlas (thresholded at 50%).

As a measure of long-range mPFC-PCC connectivity strength we used Fisher-z transformed correlations coefficients between mean time-series extracted for mPFC and PCC. To test whether SFS or tSNR were predictive of the mPFC-PCC connection strength, we computed between-subject correlations ($N = 36$) between the minimum SFS or tSNR for each mPFC-PCC pair and connectivity strength. The decision to use the SFS or tSNR value for the system as a whole based upon the minimum value is intuitively based upon the intuition that for networks that include two or more nodes, signal for the network as a whole can only be as strong as that of its weakest node. However, this intuition is not completely accurate; it is only a better solution than the next easiest option, which is to take the mean.

Estimating the Impact of Noise on Correlations

While some types of analyses (ReHo) are calculated from a single node, for other types of analyses (e.g., long-range connectivity, dynamic causal modeling) it may be desirable to optimize over the system/circuit as a whole. In order to do so, it is necessary to calculate the "mutual" tSNR or SFS of a multiple-node system, for which each node may have different values. While one option would be to calculate the tSNR or SFS for each node, and then average between them, it turns out that this approach underestimates the impact of noise on the statistics.

To address the general question of how noise affects the Pearson correlation coefficient $r(x,y)$, we start with its definition:

$$r(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) \quad (4)$$

Thus:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \text{ and } s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}, \quad (5)$$

and analogously for $\bar{y}$ and s_y.

Let us assume that both datasets x and y consist of correlated data and uncorrelated noise. We can therefore write:

$$x_i = x_{Ci} + x_{Ni} \quad (6)$$

If we define SNR_x as x_{Ci}/x_{Ni} then,

$$x_i = x_{Ci}\left(1 + \frac{1}{SNR_x}\right) \tag{7}$$

At the same time s_x calculates in the following way, assuming that x_C and x_N are uncorrelated:

$$s_x^2 = s_{Cx}^2 + s_{Nx}^2 = s_{Cx}^2 + \frac{s_{Cx}^2}{SNR_x^2} = s_{Cx}^2\left(1 + \frac{1}{SNR_x^2}\right) \tag{8}$$

Therefore, if $r(x_C, y_C) = r_C$, and we substitute $x_i - \bar{x} = (x_{Ci} + x_{Ni} - \bar{x}_{Ci} - \bar{x}_{Ni})$ (and equivalently for the y part) in the definition of r above, then:

$$r(x, y) = \frac{r_C s_{Cx} s_{Cy}}{s_x s_y} = \frac{r_C}{\sqrt{1 + \frac{1}{SNR_x^2}}\sqrt{1 + \frac{1}{SNR_y^2}}} \tag{9}$$

We can illustrate the practical impact of noise on measured correlations between time-series by assuming the existence of two signals with perfect correlation ($r_C = 1$), each of which is subjected to different levels of noise (provided by tSNR values that match the variation within the literature: 4.42–280; Welvaert and Rosseel, 2013). As shown by Equations (4–9), if Node 1 has SNR of 4.42, and Node 2 has SNR of 280, the r-value will actually decrease from 1 to 0.975. On the other hand, averaging the two tSNR values provides an adjusted r-value of ∼1 (0.999951). Obviously, our approach in taking the lowest of the tSNR values for all nodes is also inaccurate, overestimating the impact of the noise to $r = 0.951$. However, for the purpose of optimization, it makes more sense to err on the side of being conservative, and thus we take the lower value (overestimating the impact of noise) rather than the average (underestimating the impact of noise).

Unsurprisingly, even without considering the compensatory fitting and averaging steps typically employed in contrast-based analyses, correlations are more sensitive to distortions of the frequency spectrum than are traditional contrast-based analyses. The purpose of dynamic fidelity, and therefore also of SFS, is to preserve the frequency spectrum. In order to do so, we require that time-series be linearly amplified. Let us assume that the measured BOLD signal has undergone amplification and also

assume that this amplification is linear over some range of T_2^*s but that it is non-linear over the edges of the linear range (a sigmoidal shape for example). Since the T_2^*s vary over the brain regions of interest some voxels will be amplified linearly and some will not. Even if there were a perfect correlation between two brain regions, such a distortion would reduce the correlation, whereas for task-based designs this distortion would not be as much of a problem since only the difference between contrasts is important. To illustrate this, we can use a pink-noise power law time-series modeled upon our resting-state data (**Figure 3A**). We then transform the data using a sigmoidal curve to simulate a non-ideal amplifier as shown in **Figure 3B**. The resulting transformation (**Figure 3C**) lowers the correlation with the original time-series, from $r = 1$ to $r = 0.76$.

The typical rule of thumb for optimizing fMRI is to set TE such that the signal amplifies at the center of the linear range. This is consistent with our aims, since **Figures 3A,C** demonstrates that scanning in the nonlinear ends of the range will distort the time-series dynamics. Optimizing for SFS puts one at the center of the linear range (**Figure 3B**), where responses are maximized. However, optimizing over tSNR will always place one in the upper nonlinear location (**Figure 3B**), since it is the point at which the amplitude is highest and fluctuations are minimized.

RESULTS

Our dynamic phantom exploits the fact that the magnetic susceptibility of agarose gel is concentration-dependent; thus, varying the concentration of agarose gel present within a voxel over time produces changes in T2* that we experimentally tuned to amplitudes (∼1%) typically observed with BOLD (Olsrud et al., 2008). The dynamic phantom is constructed from two concentric cylinders coupled with a pneumatic motor and fiber optic feedback system (**Figure 1A**). The outer cylinder contains a "baseline" agarose gel (2.27% w/w), while the inner cylinder is longitudinally divided with both (i) a baseline gel matching the outer cylinder and (ii) an "active" gel with slightly lower concentration of agarose (2.21% w/w), which produces slightly greater fMRI signal than the baseline gel. The longitudinally divided inner cylinder rotates about its long axis via a novel
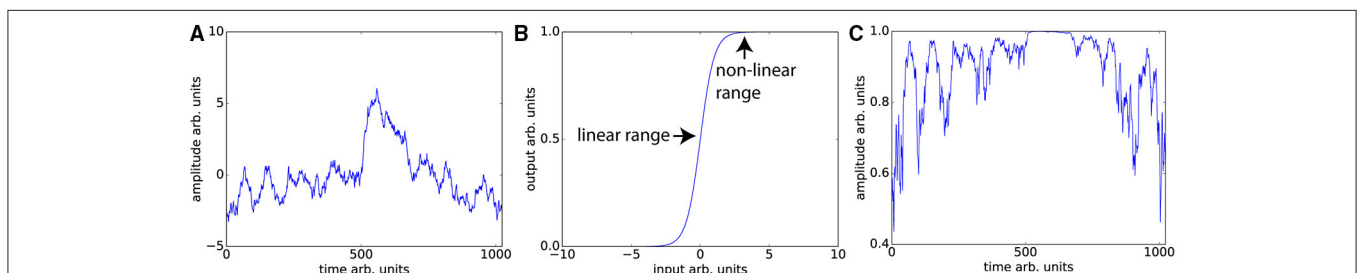
FIGURE 3 | Nonlinear amplifier distorts power spectra and reduces detection of correlation. (A) Simulated data set whose power spectrum obeys a power law with mean zero. **(B)** Characteristic function of the non-linear amplifier. To generate this curve we used a scaled version of the logistic function $1/(1 + \exp(-2x))$. If the mean of the data is around 0, this amplifier is perfectly linear, but when the mean shifts beyond the linear range, this amplifier will distort the data as shown in **(C)** Data set in **(A)** has undergone non-linear amplification by shifting the mean up by three and transforming the data using the characteristic function in **(B)**. These two data sets (original and transformed) maintain the power law but their Pearson correlation is reduced from $r = 1$ to $r = 0.76$.

MRI-compatible pneumatic motor to drive rotation of the inner cylinder (**Figures 1A,B**). By averaging time-series across a *region of interest* (ROI) that, over time, includes different proportions of the two concentrations, we were able to reproduce the effect of a concentration gradient in producing smooth fMRI time-series (**Figure 1C**). The dynamic phantom receives image acquisition signals from the MRI, and rotates only between image acquisitions to avoid motion artifacts. As the dynamic phantom rotates, position is monitored continuously through a fiber-optic feedback system. As the fMRI acquires each image, the dynamic phantom reads out its position, which serves as a 1:1 "input" for input-output mapping. The dynamic phantom can be programmed to produce fMRI signals that follow any dynamic input, including those expected for task-generated event-related and block designs, as well as the resting-state (e.g., pink-noise) fluctuations used to obtain the results below. As shown in **Figure 1D**, the dynamic phantom produces tightly controlled changes in fMRI signal without motion artifacts, and therefore can provide a ground-truth upon which to establish the degree to which SFS and tSNR reflect fMRI's dynamic fidelity to the original BOLD signal.

Dynamic Phantom Assessment of SFS vs. tSNR in Predicting Dynamic Integrity

We programmed the dynamic phantom to mimic resting-state oscillations observed in human fMRI (Van Den Heuvel et al., 2008) (**Figure 4A**), and scanned the dynamic phantom under three different sets of acquisition parameters. *Acquisition A* represents what would normally be considered to be the standard for typical resting-state studies, using a 3T scanner with 32-channel head coil and 2000 ms temporal-resolution (TR). *Acquisition B* uses a set of parameters that were specifically designed for resting-state connectivity analyses as part of the *Human Connectome Project*. These include a 3T scanner that increases the temporal-resolution to 1080 ms in order to achieve greater sensitivity to fluctuation dynamics; to compensate for signal loss associated with accelerated scanning, *Acquisition B* uses a custom-built 64-channel head coil. *Acquisition C* pushes even further than *Acquisition B* in optimizing over temporal resolution (802 ms). *Acquisition C* retains the 32-channel head coil, but compensates for signal loss associated with accelerated scanning by increasing the field strength to 7T. In each scanner, we scanned the dynamic phantom for 10 min under each acquisition paradigm optimized for human studies, as well as at two other TRs comparable to those previously optimized for the other two scanners. Thus, in total we performed nine scans (three scanners × three TRs each) with the dynamic phantom; scanners and scan parameters used for each session are provided in **Table 1**. For human data, we also acquired T1-weighted structural images and B0 field maps for correction of EPI data.

We then computed dynamic fidelity, SFS, and tSNR on raw data acquired from the dynamic phantom. Standard deviations were computed after voxel-wise removal of linear and quadratic trends. The dynamic phantom is longitudinally divided into four chambers, and rotates about the long axis orthogonally to the

FIGURE 4 | Dynamic phantom results show dynamic fidelity positively correlates with signal fluctuation sensitivity (SFS) and negatively correlates with classical temporal signal to noise ratio (tSNR). (A) To accurately mimic human resting-state fluctuations in the dynamic phantom, we utilized a complex pink-noise waveform as shown by the dotted line. The 10-min input function originated from our previous neuroimaging data and was subsequently programmed into the phantom. The dynamic phantom inputs are derived from position tracking during rotation. A representative output fMRI signal is superimposed (*fMRI Output* axis), as acquired under *Acquisition B:* 3T magnet, 64 Channel head-coil, at *TR* = 1080 ms (see **Table 1**). This waveform input was used for all nine phantom fMRI scans. **(B)** Input-output fidelity was positively correlated with SFS (median *r* = 0.67, see **Table 2**) and negatively correlated with tSNR (median *r* = −0.63, see **Table 2**). Groups presented here match the scanning parameters presented in the subsequent human data: acquisition A is a 3 Tesla magnet with a 32-channel headcoil (*TR* = 2000 ms), acquisition B is a 3 Tesla magnet with a 64-channel headcoil (*TR* = 1080 ms), and acquisition C is a 7 Tesla magnet with a 32-channel head coil (*TR* = 802 ms). **Table 1** provides detailed acquisition parameters for each scan, while **Table 2** provides detailed results from all nine dynamic phantom scans.

main field. For our *region of interest*, we extracted the average time-series from each the four quadrants of the inner cylinder with an automated masking procedure, and repeated this for six slices positioned in the center of the dynamic phantom (*n* = 24 time-series per scan). *Dynamic fidelity* was defined as the correlation between user-defined *dynamic inputs*, provided by the phantom rotation, and *dynamic outputs* acquired from the scanner in the region of interest. To compute the *nuisance* term within SFS (analogous to CSF in humans), we extracted fluctuations acquired from the outer cylinder in these six slices, which includes only inactive voxels.

Dynamic fidelity *directly* correlated with SFS for each of the nine scans (**Figure 4B**; **Table 2**; median *r* = 0.67) and *inversely* correlated with tSNR for each of the nine scans (**Figure 4B**; **Table 2**; median *r* = −0.63). Thus, when the scanner was most sensitive in capturing dynamic inputs, SFS was maximized while tSNR was minimized, and vice-versa. Researchers typically optimize acquisition parameters for fMRI connectivity studies by

TABLE 2 | Dynamic fidelity in all nine dynamic phantom scans was positively correlated with SFS and negatively correlated with tSNR.

	TR (ms)	3T, 32 Ch	3T, 64 Ch	7T, 32 Ch
CORRELATION WITH DYNAMIC FIDELITY				
	~2000	0.51 ($p = 0.010$)	0.69 ($p = 2.1 \times 10^{-4}$)	0.54 ($p = 6.6 \times 10^{-3}$)
SFS	~1080	0.71 ($p = 1.1 \times 10^{-4}$)	0.76 ($p = 1.8 \times 10^{-5}$)	0.69 ($p = 1.9 \times 10^{-4}$)
	~802	0.67 ($p = 3.7 \times 10^{-4}$)	0.49 ($p = 0.014$)	0.63 ($p = 9.0 \times 10^{-4}$)
	~2000	−0.67 ($p = 3.2 \times 10^{-4}$)	−0.64 ($p = 7.0 \times 10^{-4}$)	−0.44 ($p = 0.030$)
tSNR	~1080	−0.87 ($p = 4.6 \times 10^{-8}$)	−0.72 ($p = 8.3 \times 10^{-5}$)	−0.58 ($p = 3.0 \times 10^{-3}$)
	~802	−0.63 ($p = 1.1 \times 10^{-3}$)	−0.27* ($p = 0.20$)	−0.54 ($p = 6.2 \times 10^{-3}$)

*n.s

trying to maximize tSNR. Yet doing so would appear to produce the greatest amount of distortion for the BOLD fluctuations upon which connectivity results are based. Thus, we tested the implications of our dynamic phantom results for human connectivity studies.

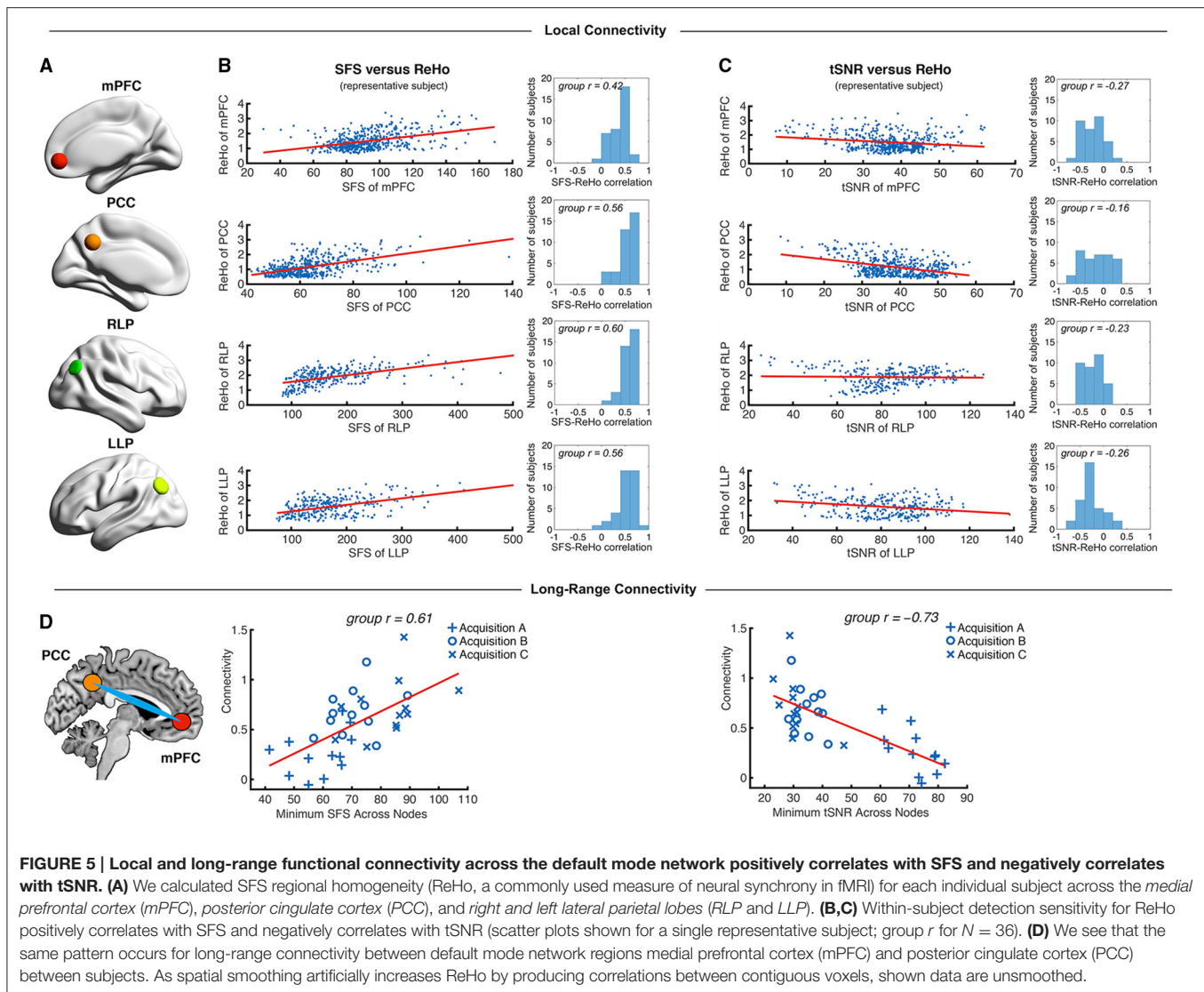## Human Subjects Assessment of SFS vs. tSNR in Detecting Resting-State Connectivity

We calculated SFS and tSNR in human neuroimaging data acquired using *Acquisitions A*, *B*, and *C* (restricting our analyses to the *TR* originally, and independently, optimized for each scanner), and assessed the utility of each in predicting detection sensitivity to resting-state network features. Human data were preprocessed according to standard methods, including the SPM8 preprocessing pipeline; to gauge the impact of spatial smoothing, we calculated all values both with and without this step. After preprocessing, we used MATLAB to compute SFS and tSNR as per Equation (1). For resting-state connectivity, we computed three commonly used measures. The first was the between-voxel measure of local connectivity, *regional homogeneity* (*ReHo*; Zang et al., 2004). The second was the within-voxel *amplitude of low-frequency fluctuations* (*ALFF*; Zang et al., 2007), which is thought to underlie resting-state connectivity (Biswal et al., 1995). The third was *long-range connectivity* between two nodes of the default model network (Raichle, 2015): the *medial prefrontal cortex* and the *posterior cingulate cortex* (mPFC-PCC). DMN regions were defined as 10 mm radius spheres centered upon previously established coordinates (Fox et al., 2005). Long-range connectivity forms the basis for graph theoretic/complex network analyses (Bullmore and Sporns, 2009) used within the fMRI field.

To test the degree to which SFS and tSNR were sensitive to well-established resting-state features, we computed correlations between SFS and ReHo, ALFF, and long-range connectivity; as well as tSNR and ReHo, ALFF, and long-range connectivity. ReHo and ALFF were computed for voxels within the well-established default mode network, comprised of the *medial prefrontal cortex, posterior cingulate cortex*, and *bilateral parietal cortices* (**Figure 5A**). Long-range connectivity focused upon the two-node MPFC-PCC connection, which we found to be reliable across subjects within our dataset (33 out of 36 subjects

showed significant positive correlation between mean time-series from these two regions). For networks that include two or more nodes, we used the minimum SFS or tSNR for each mPFC-PCC pair (as justified in the Materials and Methods Section).

We first tested whether SFS and tSNR would predict local connectivity (ReHo) at a single-subject level. Without smoothing, region-specific correlations within the *default mode network* showed robust positive relationships between SFS and ReHo (median $r = 0.53$: 96% $p < 0.05$, 95% $p < 0.01$, 94% $p < 0.001$; by acquisition set: $r_A = 0.54$, $r_B = 0.51$, $r_C = 0.54$; see **Figure 5B** for median across subjects and default mode network regions). In contrast, the correlation between tSNR and ReHo was either non-significant or significant but negative within most subjects' default mode network regions (median $r = -0.24$: 80% $p < 0.05$, 76% $p < 0.01$, 68% $p < 0.001$—even using the most liberal threshold of $p < 0.05$, only 11% of all correlations were positive between tSNR and ReHo; by acquisition set: $r_A = -0.42$, $r_B = -0.20$, $r_C = -0.06$; see **Figure 5C** for median across subjects and default mode network regions). Smoothing only magnified this effect. After smoothing, SFS positively correlated with ReHo (median $r = 0.68$: 98% $p < 0.05$, 98% $p < 0.01$, 98% $p < 0.001$; by acquisition set: $r_A = 0.69$, $r_B = 0.72$, $r_C = 0.55$) and tSNR negatively correlated with ReHo (median $r = -0.62$: 97% $p < 0.05$, 97% $p < 0.01$, 97% $p < 0.001$; by acquisition set: $r_A = -0.72$, $r_B = -0.60$, $r_C = -0.57$). The same relationship was preserved when looking at the entire group ($N = 36$). The correlation between SFS and ReHo was $r = 0.39$, $p = 0.02$ (removing two outliers two $SD \pm$ the mean: $r = 0.52$, $p = 0.002$), while the correlation between tSNR and ReHo was $r = -0.04$, $p = 0.83$ (removing the same two outliers: $r = -0.03$, $p = 0.88$).

While ReHo is a measure of between-voxel local connectivity, ALFF is a single-voxel measure that estimates the total power of the low frequency component of an fMRI signal. Thus, we expected the relationship between ALFF and SFS (both single voxel measures) to be even more robust than the relationship between SFS and ReHo. Indeed, SFS strongly correlated with ALFF (median $r = 0.82$, all p's < 0.001; by acquisition set: $r_A = 0.90$, $r_B = 0.71$, $r_C = 0.77$), whereas tSNR was negatively correlated with ALFF (median $r = -0.70$, all p's < 0.001; by acquisition set: $r_A = -0.82$, $r_B = -0.65$, $r_C = -0.58$). Again, smoothing magnified this effect for both SFS (median
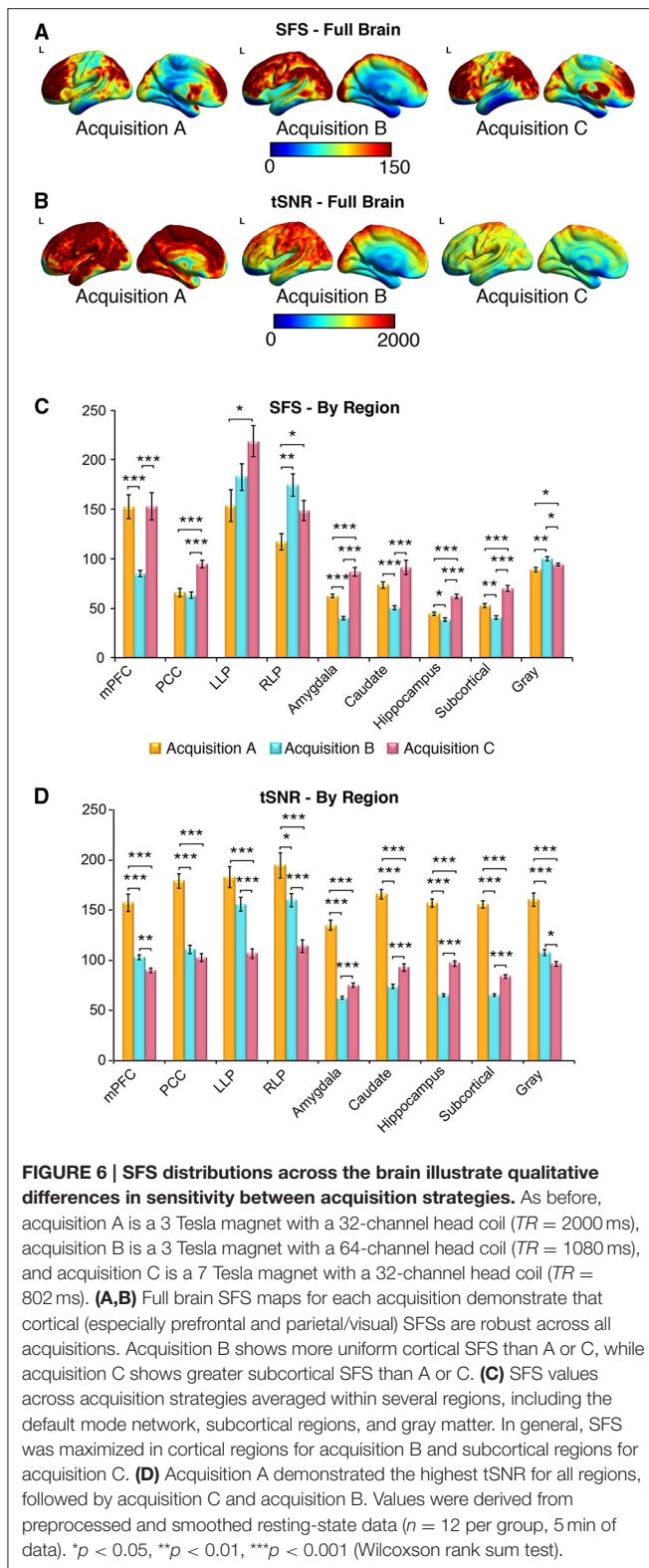
FIGURE 5 | Local and long-range functional connectivity across the default mode network positively correlates with SFS and negatively correlates with tSNR. (A) We calculated SFS regional homogeneity (ReHo, a commonly used measure of neural synchrony in fMRI) for each individual subject across the *medial prefrontal cortex* (*mPFC*), *posterior cingulate cortex* (*PCC*), and *right and left lateral parietal lobes* (*RLP* and *LLP*). **(B,C)** Within-subject detection sensitivity for ReHo positively correlates with SFS and negatively correlates with tSNR (scatter plots shown for a single representative subject; group *r* for N = 36). **(D)** We see that the same pattern occurs for long-range connectivity between default mode network regions medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC) between subjects. As spatial smoothing artificially increases ReHo by producing correlations between contiguous voxels, shown data are unsmoothed.

$r = 0.93$, all p's $<< 0.001$; by acquisition set: $r_A = 0.94$, $r_B = 0.92$, $r_C = 0.93$), and tSNR (median $r = -0.84$, all p's $<< 0.001$; by acquisition set: $r_A = -0.86$, $r_B = -0.83$, $r_C = -0.84$). Again, the same relationship was preserved when looking at the entire group (N = 36). The correlation between SFS and ALFF was $r = 0.74$, $p = 0.03 \times 10^{-7}$, while the correlation between tSNR and ALFF was $r = -0.19$, $p = 0.26$.

As a measure of long-range connectivity, we tested SFS and tSNR against the default mode network's MPFC-PCC connection (Fisher-z normalized) across our three datasets (N = 36). Consistent with the other connectivity measures, SFS positively correlated with MPFC-PCC connectivity ($r_{A,B,C} = 0.61$, $p = 8.65 \times 10^{-5}$) and tSNR negatively correlated with MPFC-PCC connectivity ($r_{A,B,C} = -0.73$, $p = 4.46 \times 10^{-7}$) (**Figure 5D**). As with previous measures, smoothing did not qualitatively change the results for either SFS ($r_{A,B,C} = 0.40$, $p = 0.015$) or tSNR ($r_{A,B,C} = -0.70$, $p = 1.67 \times 10^{-6}$).

SFS and tSNR Values between Acquisition Sets

The purpose of sensitivity metrics, for any measurement, is to provide accurate feedback by which parameters can be tuned to optimize performance, as well as to aid in the interpretation and artifact-correction of results. Our three representative acquisition strategies illustrate clearly the practical importance of using SFS rather than tSNR when optimizing fMRI studies for task-free analyses, and therefore dynamic fidelity. We compared SFS and tSNR values between acquisition paradigms for the default mode network, subcortical regions critical to the emotion and reward circuits, and global gray matter. Because human studies normally utilize smoothing, and because our previous analyses (above) showed comparable results for smoothed and unsmoothed results, **Figure 6** results include only 4-mm smoothed data. To directly compare SFS and tSNR values between acquisition sets, we extracted average values from the four regions of the *default mode network* and three subcortical regions (*amygdala,*

FIGURE 6 | SFS distributions across the brain illustrate qualitative differences in sensitivity between acquisition strategies. As before, acquisition A is a 3 Tesla magnet with a 32-channel head coil (*TR* = 2000 ms), acquisition B is a 3 Tesla magnet with a 64-channel head coil (*TR* = 1080 ms), and acquisition C is a 7 Tesla magnet with a 32-channel head coil (*TR* = 802 ms). **(A,B)** Full brain SFS maps for each acquisition demonstrate that cortical (especially prefrontal and parietal/visual) SFSs are robust across all acquisitions. Acquisition B shows more uniform cortical SFS than A or C, while acquisition C shows greater subcortical SFS than A or C. **(C)** SFS values across acquisition strategies averaged within several regions, including the default mode network, subcortical regions, and gray matter. In general, SFS was maximized in cortical regions for acquisition B and subcortical regions for acquisition C. **(D)** Acquisition A demonstrated the highest tSNR for all regions, followed by acquisition C and acquisition B. Values were derived from preprocessed and smoothed resting-state data (*n* = 12 per group, 5 min of data). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ (Wilcoxson rank sum test).

caudate, hippocampus), as well as *average subcortical* and *average gray matter*. Each subcortical region was defined from FSL Harvard-Oxford Subcortical Atlas and *average subcortical* includes *bilateral accumbens, amygdala, caudate, hippocampus,*

pallidum, putamen, and *thalamus.* The *gray matter* mask was defined as SPM's probabilistic gray matter map thresholded at $P > 0.5$.

As shown in **Figures 6A–D**, SFS identifies advantages for dynamic fidelity in increasing temporal resolution, as well as the costs and benefits associated with increasing the number of head-coil channels vs. field strength in order to recover signal loss from accelerated acquisition. In general, the ultra-dense head-coil strategy employed by *Acquisition B* optimizes dynamic fidelity in cortical regions, whereas the ultra-high-field strategy employed by *Acquisition C* optimizes dynamic fidelity in subcortical regions. TSNR provides a very different story: showing the greatest stability in *Acquisition A*, diminished performance in *Acquisition B*, and the worst performance in *Acquisition C*. Which strategy is ideal, for any particular study, therefore depends critically upon the scientific questions to be asked: not only with respect to the regions of interest implicated, but also the types of analyses to be performed.

DISCUSSION

Functional neuroimaging has ushered in a new era of brain research, in which task-free fluctuations play an increasingly large role. As such, we need to reconsider whether fMRI optimization paradigms that rely solely on maximizing stability might actually be leading us astray, by failing to functionally dissociate fluctuations underlying signal vs. those underlying noise. Here we propose a new measure—SFS—that distinguishes between neurobiologically-relevant fluctuations of interest, and nuisance fluctuations due to physiological or scanner artifact. We demonstrate that SFS positively—and tSNR negatively—correlates with dynamic fidelity in a dynamic phantom, as well as with the detection power of local and long-range functional networks in humans, across three sets of representative acquisition parameters independently optimized for human fMRI studies.

Our design of the dynamic phantom was motivated by the need to rigorously test the fidelity of fMRI time-series to its known dynamic inputs, which would be otherwise impossible using either a static phantom or human data. While we could have simulated input-output fidelity in the presence of physiological and scanner noise, models can be susceptible to bias and often over-simplify the complexities of fMRI noise (Renvall and Hari, 2009; Erhardt et al., 2012). The empirical approach defined here captures actual scanner noise, and thus is more accurate in evaluating the utility of SFS to human neuroimaging. One of the most challenging aspects of our phantom's design from an engineering standpoint was the need to avoid motion artifacts for a machine with rotating elements. While the phantom's inner cylinder is programmed to move only between the scanner's intermittent radio frequency pulses, residual motion artifact would have a devastating impact on SFS values, because it would preferentially affect σ_{ROI} from the inner (rotating) compartment, while bypassing $\sigma_{nuisance}$ from the outer (static) compartment. The fact that the slices

from which we acquired data did not show characteristic motion artifact (clearly visible during slices that, for testing, were deliberately acquired during motion), that SFS values for the phantom correlated with dynamic fidelity (as shown in **Figure 2**, motion artifact corrupts fidelity), and that the relationships observed for dynamic fidelity were supported by human connectivity data, provides assurance with respect to the integrity of the phantom's design. Future validation of SFS would benefit from a biological ground truth for measurement of dynamic fidelity, using simultaneous inputs recorded from local field potentials and their associated hemodynamic responses, combined with outputs obtained from fMRI time-series.

For dynamic analyses, the structure of the SFS equation reflects the equally important need to optimize over signal amplitude (provided by the mean) and signal change (provided by the standard deviation). In our datasets, we found that the relationship between SFS vs. tSNR and long-range connectivity was driven primarily by the standard deviation component of each equation (including only the standard deviation ratio component of the SFS equation, correlations were $r_{A,B,C} = 0.68$, $p = 4.34 \times 10^{-6}$ for $N = 36$; unsmoothed), while the mean signal component (including only the mean amplitude of signal or the ratio of the mean amplitude to global amplitude) showed no statistical relationship to long-range connectivity strength ($r_{A,B,C}$ $4 = -0.18$, $p = 0.31$; $r_{A,B,C} = 0.02$, $p = 0.90$). However, our fMRI data had minimal signal drop out in regions of interest, which is not always the case. While the mean amplitude of the signal did not play a role in evaluating our data set, nevertheless this term of the SFS equation should be retained in order to avoid assigning high SFS to areas of the brain that show signal loss.

In developing SFS for humans, one important decision is the optimal location for the acquisition of *nuisance* fluctuations. We chose *cerebrospinal fluid*, rather than surrounding air, white matter, or whole brain, because time-series from the cerebrospinal fluid contain the greatest proportion of nuisance variance of the three brain compartments (Wald, 2012), including motion, scanner variance, and some physiological effects. Moreover, unlike white matter (Gawryluk et al., 2014) and the global signal, the eroded CSF masks used here are unlikely to contain neurobiologically-relevant fluctuations of interest.

In extending our phantom results to the brain, we faced the problem of what to look for as a measure of validation, since we lacked the phantom's advantage of known inputs. Thus, we used highly conservative and well-replicated connections in order to evaluate detection sensitivity for resting-state data: a measure of local-connectivity (*regional homogeneity – ReHo Zang et al., 2004*), a single voxel measure of resting-state signals (*amplitude of low frequency fluctuations – ALFF Zang et al., 2007*), and the *long-range connection* between two nodes within the default mode network (*medial prefrontal cortex* and *posterior cingulate cortex*; Raichle, 2015). Both ALFF and ReHo have been widely used to study resting-state brain activity, with clinical applications to Parkinson's disease (Wu et al., 2009), Alzheimer's disease (Liu et al., 2008), and psychiatric illnesses (Han et al.,

2011). Likewise, identification of the default mode network via long-range connectivity represents a fundamental finding in neuroscience (Raichle, 2015), with direct implications for neurodevelopment and aging. Although the correlations for SFS and tSNR were both significant but of opposite sign, it is critical to note that they are not trivially inverses of one another. From a theoretical perspective, SFS and tSNR differ fundamentally in their dissociation of fluctuations of interest vs. those of nuisance. From a practical perspective, **Figure 6** demonstrates that the two measures provide qualitatively different mapping of optimization over the brain. Thus, we demonstrate that, by optimizing for *dynamic fidelity* rather than the current standard of *dynamic stability*, SFS can have direct practical applications for markedly increasing detection sensitivity of clinical neuroimaging results.

Although we have emphasized the application of SFS to correlational analyses due to their increasing prevalence within the field, it is important to note that other types of task-free analyses will also be much better served by optimization to SFS than tSNR. This category of analyses includes those based upon power spectra and complexity (e.g., *ALFF*, *power spectrum scale invariance*, *entropic analyses*, *spectral dynamic causal modeling*), which also are more highly sensitive to subtle dynamic features of the time-series than are traditional contrast-based analyses[1]. Consistent with these approaches, one additional area for future exploration is whether the principles underlying SFS can be applied not only for optimization, but also to identify neural activity for task-free paradigms. The ability to dissociate signal fluctuations from noise fluctuations may be fruitful when the aim is to map the strength and location of task-free brain responses, rather than their connectivity.

Finally, while we have focused here on presenting limitations of CNR and tSNR, we wish to emphasize that both may still be useful and accurate measures in answering particular questions. Temporal SNR is a measure of signal stability that is proportional to field strength, voxel size, and sampling rate (Kruger et al., 2001; Triantafyllou et al., 2005); thus, in static phantoms, tSNR can be used to quantify and minimize scanner-related noise. If the primary aim of a study is to show contrast between two conditions then CNR, and not SFS, is correct. For task-free analyses, however, CNR is not directly measurable; thus, classical tSNR is normally cited as a surrogate (Van Dijk et al., 2012; Smith et al., 2013). It is in this case that using tSNR as guide will minimize, rather than maximize, detection sensitivity. As with so many zero-sum decisions in fMRI acquisition, it is important to realize that we optimize over one parameter at the expense of the other. Therefore, just as tuning of acquisition parameters benefits enormously from knowing *a priori* the region of interest to be

---

[1]To illustrate the impact of signal fidelity on spectral methods, we calculated the power spectrum scale invariance (after detrending, with a full frequency range from 0 to the Nyquist limit of 0.5*1/TR) for time-series acquired with the dynamic phantom. As with connectivity analyses, SFS—but not tSNR—correlated with accuracy: the percentage difference (i.e., "error"), between "true" power-law slope β for the phantom and the "measured" β values acquired from fMRI, decreased with higher SFS ($r = -0.67$; $p < 0.05$ for 9/9 scans) and increased with tSNR ($r = 0.66$; $p < 0.05$ for 8/9 scans).

targeted, knowing *a priori* the type of analysis to be performed will permit researchers to decide whether to optimize for stability (tSNR) or for dynamics (SFS).

AUTHOR CONTRIBUTIONS

DD, LM, and HS designed the dynamic phantom. DD, PK, and SA built and programmed the dynamic phantom. DD, LW, AT, and LM tested the dynamic phantom. LM, SN, HS, DD, LW, and AT developed the SFS algorithm. KV and LM contributed data to test the SFS algorithm. LM, DD, SN, HS, LW, AT, and KV wrote the paper.

REFERENCES

Bandettini, P. A. (2014). Neuronal or hemodynamic? Grappling with the functional MRI signal. *Brain Connect* 4, 487–498. doi: 10.1089/brain.2014.0288

Bandettini, P. A., and Cox, R. W. (2000). Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. *Magn. Reson. Med.* 43, 540–548. doi: 10.1002/(SICI)1522-2594(200004)43:4<540::AID-MRM8>3.0.CO;2-R

Bettus, G., Guedj, E., Joyeux, F., Confort-Gouny, S., Soulier, E., Laguitton, V., et al. (2009). Decreased basal fMRI functional connectivity in epileptogenic networks and contralateral compensatory mechanisms. *Hum. Brain Mapp.* 30, 1580–1591. doi: 10.1002/hbm.20625

Biswal, B. B., Van Kylen, J., and Hyde, J. S. (1997). Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. *NMR Biomed.* 10, 165–170.

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575

Christoffersson, J. O., Olsson, L. E., and Sjoberg, S. (1991). Nickel-doped agarose gel phantoms in MR imaging. *Acta Radiol.* 32, 426–431. doi: 10.1177/028418519103200519

Damoiseaux, J. S., Beckmann, C. F., Arigita, E. J., Barkhof, F., Scheltens, P., Stam, C. J., et al. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cereb. Cortex* 18, 1856–1864. doi: 10.1093/cercor/bhm207

Erhardt, E. B., Allen, E. A., Wei, Y., Eichele, T., and Calhoun, V. D. (2012). SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. *Neuroimage* 59, 4160–4167. doi: 10.1016/j.neuroimage.2011.11.088

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102

Gawryluk, J. R., Mazerolle, E. L., and D'arcy, R. C. (2014). Does functional MRI detect activation in white matter? A review of emerging evidence, issues, and future directions. *Front Neurosci* 8:239. doi: 10.3389/fnins.2014.00239

Gonzalez-Castillo, J., Handwerker, D. A., Robinson, M. E., Hoy, C. W., Buchanan, L. C., Saad, Z. S., et al. (2014). The spatial structure of resting state connectivity stability on the scale of minutes. *Front. Neurosci.* 8:138. doi: 10.3389/fnins.2014.00138

Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., et al. (2007). Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiatry* 62, 429–437. doi: 10.1016/j.biopsych.2006.09.020

Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258. doi: 10.1073/pnas.0135058100

Han, Y., Wang, J., Zhao, Z., Min, B., Lu, J., Li, K., et al. (2011). Frequency-dependent changes in the amplitude of low-frequency fluctuations in amnestic mild cognitive impairment: a resting-state fMRI study. *Neuroimage* 55, 287–295. doi: 10.1016/j.neuroimage.2010.11.059

Kruger, G., Kastrup, A., and Glover, G. H. (2001). Neuroimaging at 1.5 T and 3.0 T: comparison of oxygenation-sensitive magnetic resonance imaging. *Magn. Reson. Med.* 45, 595–604. doi: 10.1002/mrm.1081

Liu, Y., Wang, K., Yu, C., He, Y., Zhou, Y., Liang, M., et al. (2008). Regional homogeneity, functional connectivity and imaging markers of Alzheimer's disease: a review of resting-state fMRI studies. *Neuropsychologia* 46, 1648–1656. doi: 10.1016/j.neuropsychologia.2008.01.027

Logothetis, N. K., Eschenko, O., Murayama, Y., Augath, M., Steudel, T., Evrard, H. C., et al. (2012). Hippocampal-cortical interaction during periods of subcortical silence. *Nature* 491, 547–553. doi: 10.1038/nature11618

Mayer, A. R., Mannell, M. V., Ling, J., Gasparovic, C., and Yeo, R. A. (2011). Functional connectivity in mild traumatic brain injury. *Hum. Brain Mapp.* 32, 1825–1835. doi: 10.1002/hbm.21151

Molloy, E. K., Meyerand, M. E., and 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

Murphy, K., Bodurka, J., and Bandettini, P. A. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *Neuroimage* 34, 565–574. doi: 10.1016/j.neuroimage.2006.09.032

Olsrud, J., Nilsson, A., Mannfolk, P., Waites, A., and Stahlberg, F. (2008). A two-compartment gel phantom for optimization and quality assurance in clinical BOLD fMRI. *Magn. Reson. Imaging* 26, 279–286. doi: 10.1016/j.mri.2007.06.010

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annu. Rev. Neurosci.* 38, 433–447. doi: 10.1146/annurev-neuro-071013-014030

Renvall, V., and Hari, R. (2009). Transients may occur in functional magnetic resonance imaging without physiological basis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20510–20514. doi: 10.1073/pnas.0911265106

Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D., and Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* 15, 1332–1342. doi: 10.1093/cercor/bhi016

Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., et al. (2013). Resting-state fMRI in the Human Connectome Project. *Neuroimage* 80, 144–168. doi: 10.1016/j.neuroimage.2013.05.039

Song, X. W., Dong, Z. Y., Long, X. Y., Li, S. F., Zuo, X. N., Zhu, C. Z., et al. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS ONE* 6:e25031. doi: 10.1371/journal.pone.0025031

Triantafyllou, C., Hoge, R. D., Krueger, G., Wiggins, C. J., Potthast, A., Wiggins, G. C., et al. (2005). Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. *Neuroimage* 26, 243–250. doi: 10.1016/j.neuroimage.2005.01.007

Van Den Heuvel, M. P., Stam, C. J., Boersma, M., and Hulshoff Pol, H. E. (2008). Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage* 43, 528–539. doi: 10.1016/j.neuroimage.2008.08.010

Van Dijk, K. R., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438. doi: 10.1016/j.neuroimage.2011.07.044

Wald, L. L. (2012). The future of acquisition speed, coverage, sensitivity, and resolution. *Neuroimage* 62, 1221–1229. doi: 10.1016/j.neuroimage.2012.02.077

Welvaert, M., and Rosseel, Y. (2013). On the definition of signal-to-noise ratio and contrast-to-noise ratio for FMRI data. *PLoS ONE* 8:e77089. doi: 10.1371/journal.pone.0077089

Wu, T., Long, X., Zang, Y., Wang, L., Hallett, M., Li, K., et al. (2009). Regional homogeneity changes in patients with Parkinson's disease. *Hum. Brain Mapp.* 30, 1502–1510. doi: 10.1002/hbm.20622

Zang, Y.-F., He, Y., Zhu, C.-Z., Cao, Q.-J., Sui, M.-Q., Liang, M., et al. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 29, 83–91. doi: 10.1016/j.braindev.2006.07.002

Zang, Y., Jiang, T., Lu, Y., He, Y., and Tian, L. (2004). Regional homogeneity approach to fMRI data analysis. *Neuroimage* 22, 394–400. doi: 10.1016/j.neuroimage.2003.12.030