



Resting-State fMRI Dynamics and Null Models: Perspectives, Sampling Variability, and Simulations

Robyn L. Miller^{1*}, Anees Abrol^{1,2}, Tulay Adali³, Yuri Levin-Schwarz³ and Vince D. Calhoun^{1,2}

¹ The Mind Research Network, Albuquerque, NM, United States, ² Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, United States, ³ Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, United States

OPEN ACCESS

Edited by:

Amir Shmuel,
McGill University, Canada

Reviewed by:

J. Bruce Morton,
University of Western Ontario, Canada
Baxter P. Rogers,
Vanderbilt University, United States

*Correspondence:

Robyn L. Miller
rmiller@mrn.org

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 24 January 2018

Accepted: 20 July 2018

Published: 06 September 2018

Citation:

Miller RL, Abrol A, Adali T,
Levin-Schwarz Y and Calhoun VD
(2018) Resting-State fMRI Dynamics
and Null Models: Perspectives,
Sampling Variability, and Simulations.
Front. Neurosci. 12:551.
doi: 10.3389/fnins.2018.00551

Studies of resting state functional MRI (rs-fMRI) are increasingly focused on “dynamics”, or on those properties of brain activation that manifest and vary on timescales shorter than the scan’s full duration. This shift in focus has led to a flurry of interest in developing hypothesis testing frameworks and null models applicable to the dynamical setting. Thus far however, these efforts have been weakened by a number of crucial shortcomings that are outlined and discussed in this article. We focus here on aspects of recently proposed null models that, we argue, are poorly formulated relative to the hypotheses they are designed to test, i.e., their potential role in separating functionally relevant BOLD signal dynamics from noise or intermittent background and maintenance type processes is limited by factors that are fundamental rather than merely quantitative or parametric. In this short position paper, we emphasize that (1) serious care must be exercised in building null models for rs-fMRI dynamics from distributionally stationary univariate or multivariate timeseries, i.e., timeseries whose values are each independently drawn from one pre-specified probability distribution; and (2) measures such as kurtosis that quantify over-concentration of observed values in the far tails of some reference distribution may not be particularly suitable for capturing signal features most plausibly contributing to functionally relevant brain dynamics. Other metrics targeted, for example, at capturing the type of epochal signal variation that is often viewed as a signature of brain responsiveness to stimuli or experimental tasks, could play a more scientifically clarifying role. As we learn more about the phenomenon of functionally relevant brain dynamics and its imaging correlates, scientifically meaningful null hypotheses and well-tuned null models will naturally emerge. We also revisit the important concept of distributional stationarity, discuss how it manifests within realizations vs. across multiple realizations, and provide guidance on the benefits and limitations of employing this type of stationarity in modeling the absence of functionally relevant temporal dynamics in resting state fMRI. We hope that the discussions herein are useful, and promote thoughtful consideration of these important issues.

Keywords: fMRI, fMRI methods, dynamic connectivity, functional network connectivity, brain dynamics

INTRODUCTION

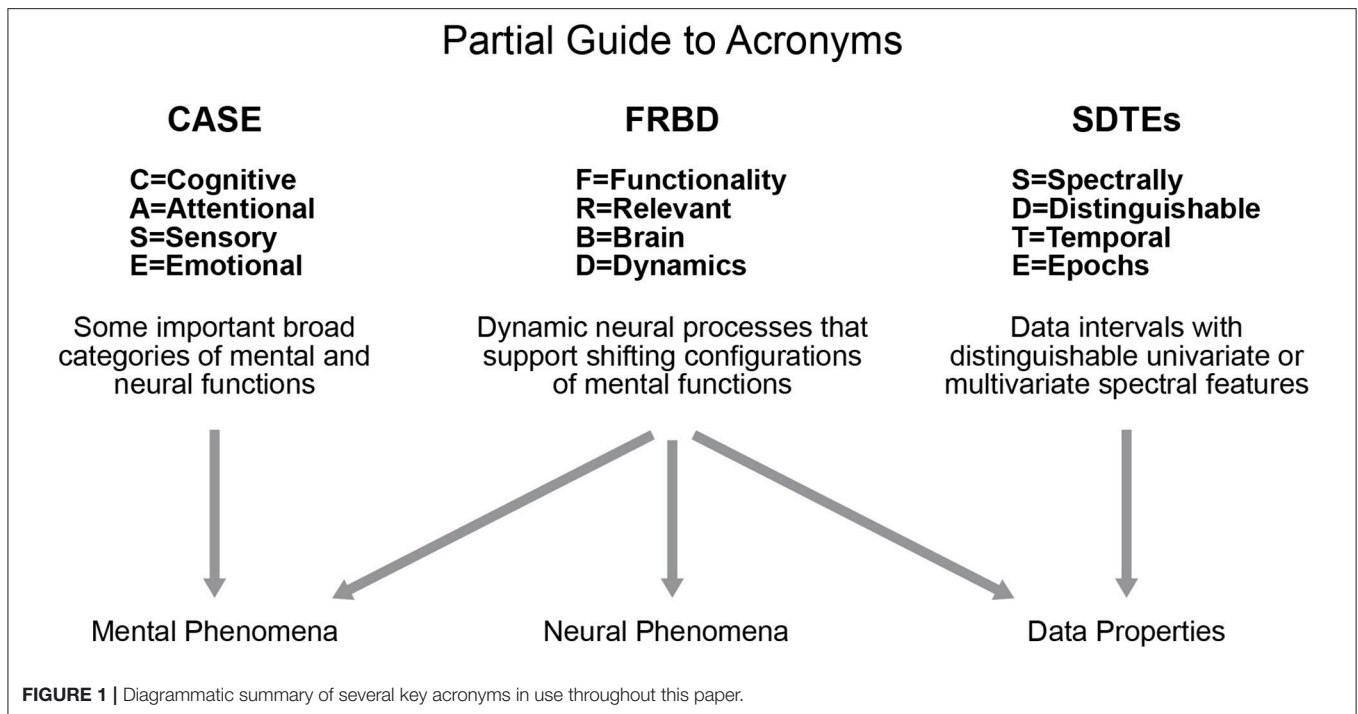
Studies of blood oxygenation-level dependent (BOLD) resting state functional magnetic resonance imaging (rs-fMRI) have been increasingly focused on properties of functional activation that manifest and vary on timescales shorter than the full duration of the scan (Hutchison et al., 2013). Such an approach is a more natural way to analyze such data as we know the brain is a highly dynamic organ. In resting-state studies, we do not have the benefit of external indicators of behavior other than the narrow case of subject wakefulness, which can be studied by using simultaneous electroencephalogram (EEG) recordings (Wu et al., 2010; Tagliazucchi and Laufs, 2015; Allen et al., 2018). In the case of resting data, we are seeking evidence in the scan itself of shifts in a subject's cognitive focus, emotional state, attention or consciousness level: e.g., we are attempting to locate the temporal and correlative signatures of complex internal ("ecologically authentic") sequences of mental tasks. The growing focus on shorter timescale analysis and examining these functionally relevant brain dynamics (FRBD) in rs-fMRI has spawned a number of efforts (Hutchison et al., 2012, 2013; Jones et al., 2012; Calhoun et al., 2014; Preti et al., 2016) to identify variations (univariate, multivariate and relational/correlative) in network behavior that arise from reconfigurations of the subject's *cognitive, attentional, sensory, emotional* (CASE) states (see **Figure 1**, partial guide to acronyms). These shifts can be conscious or non-conscious, and targets of investigation can be expanded to include temporal fluctuations that shape and maintain crucial neural circuitry that indirectly supports effective brain function.

Testing the statistical significance of the dynamic (i.e., time-varying) measures capturing these CASE shifts in such shorter timescale analyses assumes great importance because of the noisy (inherent physiological and artifactual confounds) nature of BOLD rs-fMRI data. While it would be highly useful to replicate the behavior of "noiseless" BOLD data to construct appropriate null simulations, the absence of a baseline, i.e., ground truth for resting state, makes this step extremely challenging. A null distribution of the test-statistic capturing the phenomenon of interest must therefore be approximated using multiple, independent "surrogate" realizations of the empirical data. The surrogate data realizations ideally retain all statistical properties of the empirical data other than the phenomenon of interest, and hence give a meaningful null to validate statistical significance of the metric capturing the phenomenon of interest. Notably, non-parametric null models based on phase randomization (PR) and/or vector auto-regression (VAR) approaches (Chang and Glover, 2010; Cribben et al., 2013) have been widely used to seek evidence for presence of non-stationarities (discussed in section Statistical Stationarity, Gaussianity, and rs-Brain Dynamics). These models allow us to comment on the Gaussian, stationary, and linear properties of the studied data. The work on null models and hypothesis testing (e.g., Zalesky et al., 2014; Zalesky and Breakspear, 2015; Hindriks et al., 2016; Laumann et al., 2016; Shakil et al., 2016; Shi et al., 2016; Abrol et al., 2017; Liégeois et al., 2017) frameworks for shorter timescale analysis of network behavior and dynamics has produced some preliminary

(although possibly contrasting) insights but limitations remain. One major limitation of using these (phase randomized or vector autoregressive) non-parametric null models as also pointed in a few of these studies (Zalesky and Breakspear, 2015; Shakil et al., 2016) is that they can also be rejected due to presence of non-linearities. Because of this, rejection of null hypotheses stemming from such approaches in their current form do not allow us to conclude the presence/absence of non-stationarities. More importantly, the fact that values within a signal are not inconsistent with values drawn from some fixed pre-specified distribution, i.e., that the signal is not provably distributionally non-stationary does not rule out the presence of FRBD¹⁴. Given the dynamic nature of the human brain, the most interesting hypotheses about fMRI dynamics would not focus on *whether* they exist, but rather on how they might manifest differentially over different temporal, spatial, and functional scales.

The best we can hope for right now is for researchers to focus on what signal properties they are seeking to quantify, and why they believe that a strong presence of these properties should be taken as evidence of FRBD (or of nuisance factors that are not easily separable from FRBD at current levels of measurement resolution). It can be argued that analyzing univariate or multivariate timeseries variations that are most plausibly connected with actual shifts in mental functioning (e.g., temporal epochs exhibiting profound changes in magnitudes and spectrum) will ultimately yield more scientifically clarifying information about observed resting state brain dynamics than approaches based on quantifying the improbability of observed values with respect to some fixed reference distribution. As we will be discussing in this paper, the presence of spectrally distinguishable temporal epochs (SDTEs) (see **Figure 1**) at both univariate and multivariate scales (where they play a larger role in connectivity) is one potential form of evidence for rs-FRBD that might warrant more attention. There are many features and timescales over which functionally relevant temporal variations might manifest, and most prospective metrics will present some combination of over-sensitivity to irrelevant features and blindness to important features and/or timescales. Development of valid metrics for functionally relevant brain dynamics presents thus presents a major challenge in the field (as also discussed in Hindriks et al., 2016). Such a metric would essentially, with some degree of specificity, rise in the presence of those univariate or multivariate timeseries variations that are most plausibly connected with actual shifts in mental functioning. The factors that obstruct development of powerful, valid metrics also present serious challenges to the development of valid, scientifically useful null models. In studies where the natural null hypothesis is effectively that "measured dynamic signal properties do not reflect functionally relevant dynamic neural processes," an appropriate null space of network timecourses would have to lack variations consistent with actual shifts in mental functioning, for example, (1) there must be no task, experimental or ecological condition whose signature presents as a type of epochal variation¹ generically observable in this null

¹The term "epoch" has been used in this paper to refer to a duration, time period, and not necessarily of repetitive nature.



space, and (2) the timeseries features that occur most rarely in this space must be exactly those that are most strongly consistent with a brain undergoing shifts in CASE state. Another, more general, hidden risk when using multi-parameter simulation-based null models is ensuring that the resulting distributions of key test statistics are not influenced by auxiliary model parameters unrelated to those explicitly connected with the null hypothesis.

In what follows, we further detail these points and address them, in particular, as they relate to the null model and test metric used in a recent paper (Laumann et al., 2016). Our concerns are applicable well beyond the scope of Laumann et al. (2016) but this recent paper motivates us to discuss some of the inherent challenges of dynamic fMRI analyses, and helps illustrate the key pitfalls of employing naive null models in efforts to detect the presence of ubiquitous, complex and poorly understood phenomena such as functionally-relevant brain dynamics.

As this position paper was written in partial dialogue with a recent published paper (Laumann et al., 2016) addressing the validity of dynamic functional network connectivity, we provide a very brief summary of some salient details from that paper to facilitate the reader's engagement with the discussion that follows. Using a null space of windowed connectivity matrices computed from spectrally and covariance constrained Gaussians (SCC Gaussians), the authors employ multivariate kurtosis as a test statistic to test the null hypothesis of "no dynamics". Among other findings, they report (1) a strong relationship between motion (mean frame displacement) and kurtosis; and (2) that the clusters formed by time-varying connectivity matrices from the null space are nearly indistinguishable from those

computed from empirical data. Also, while the focus of Laumann et al. (2016) was predominantly on connectivity dynamics, connectivity-based findings do carry much broader implications about the validity of shorter timescale dynamics of fMRI voxel and network timecourses as well (if connectivity is not dynamically changing on some timescale, then the underlying univariate signals are not dynamically changing in any ways that validly affect their measured correlative relationships etc.).

We hope our discussion clarifies the implications of Laumann et al. (2016), while also conveying the challenges of the project its authors undertook. It must also be noted that for remainder of the manuscript we shall be focusing on discussing appropriateness of null models in the context of epochal temporal variations (key contributors to estimated FC measures and potential signatures of FRBD) as the specific phenomenon of interest, and not in context of testing for scan-length statistical stationarity (unless otherwise specified).

MODELING FUNCTIONALLY RELEVANT BRAIN DYNAMICS VS. MODELING THEIR ABSENCE (DATA MODELS VS. NULL MODELS)

The relationship between well-constructed models of observed data and null models for a given phenomenon depends largely on the relationship between observed/observable data and the phenomenon being investigated. In situations such as those presented by the study of resting-state functionally relevant brain

dynamics (rs-FRBD) using human subject fMRI data, where the null hypothesis is that a certain phenomenon is not present but the empirical data being modeled (in this case real rs-fMRI data) happens to be data in which this phenomenon is continually present, then the goals of building a null model for the phenomenon and accurately modeling the data can diverge significantly. Though there are many cases where highly accurate models of the data are also appropriate models of a setting in which the null hypothesis applies, e.g., they are also good null models. In the fMRI setting, for example, a high-fidelity model of motion-censored data might be useful for flagging future scans that warrant examination for motion contamination.

As discussed earlier, one of the challenges in the field of brain dynamics is that it is difficult to generate a true/valid null model as the phenomena of interest are rather poorly defined. Thus, modeling this data with high fidelity is not going to be the same as producing a null model for brain dynamics. In fact, a valid null simulation model of multivariate signals lacking features corresponding to rs-FRBD would by necessity diverge from actual scans observed in living people. Depending on what turn out to be the most reliable timeseries indicators of CASE-driven brain dynamics, it is possible that there could be a valid null model that exhibits some similarities with the observed data (or output from good models of that data). However, a model built on the null hypothesis of no brain dynamics would by necessity only rarely create realizations that look similar to real rs-fMRI multivariate timeseries.

Statistical Stationarity, Gaussianity, and rs-Brain Dynamics

The statistical or distributional stationarity of a generating process is defined through the invariance of its joint probability distribution across any number of samples, to any time shift, and as such, with real data, it can be only inferred using multiple realizations of the finite-time empirical timeseries. A process is stationary if the statistical characterization computed over N realizations of any k -tuple of timepoints of length j $\{t_{j_1}, \dots, t_{j_k}\}$ and their τ -translates $\{t_{j_1} + \tau, \dots, t_{j_k} + \tau\}$ converge to identical values as $N \rightarrow \infty$. A practical way to infer stationarity is by estimating a finite set of moments as they are easier to compute than a full probability distribution. Of course, it is important to note that any definition of stationarity depends on the interval over which it is evaluated. For a given interval over which stationarity holds, it is quite possible for strongly distinguishing spectral characteristics to manifest over shorter sub-intervals as we demonstrate next. Matching white Gaussian noise to a template spectrum actually produces every possible signal with the given time-averaged spectral content given by the target. Statistical stationarity without other explicit constraints on the process does not imply that individual realizations of the univariate or multivariate timeseries (e.g., for fMRI these are individual subjects) are not featuring pronounced temporal epochs (see **Figure 2**). Even white Gaussian noise, for example, matched to some empirically-valid band-limited spectrum—a common statistical tactic that was also a step in the null model proposed in Laumann et al. (2016)—can be markedly epochal within typical individual realizations. This general phenomenon

Statistical Stationary Does not Ensure Absence of Evident Temporal Epochs

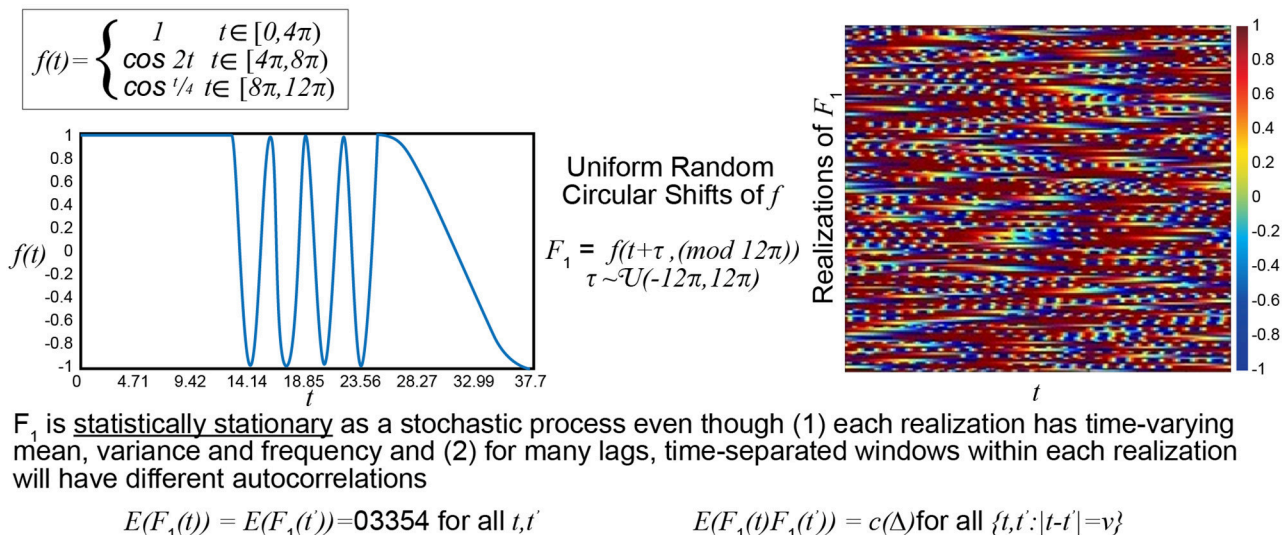


FIGURE 2 | The function f is a highly stylized example of a signal with distinguishable temporal epochs. The first, second, and final third have different means, variances and characteristic frequencies. The stochastic process F_1 , however, whose realizations are obtained through uniform random circular shifts of f , is statistically stationary in that that the statistical summaries assessed at distinct timepoints over large numbers of realizations are the same. All realizations have spectral and epochally clear variations, which would be reflective of FRBD, but as a stochastic process the collection of phase-shifted versions of f are statistically stationary.

is mostly clearly visually evident in cases like the one presented in **Figure 3** where Gaussian white noise is spectrally matched to a narrow-band spectral template.

The simulated null model of Laumann et al. (2016) is built from spectrally and covariance constrained multivariate Gaussian processes (SCC Gaussians). The use of an SCC Gaussian timeseries as a null model for functionally-relevant brain dynamics (FRBD) rests on an implicit assumption that (in the case of a 95% confidence level) fewer than 5% of the signals in this space exhibit characteristics consistent with FRBD. The use of fMRI as an imaging modality also rests on

an implicit assumption that some significant proportion of the data being recorded reflects functionally relevant brain dynamics. It would be hard, for example, to justify studying even scan-length (static) spectral and relational/connectivity characteristics of timeseries whose fluctuations are believed to be largely artifactual. The functioning human brain during any state of wakefulness is (hopefully inarguably) continually engaged in myriad temporally-varying combinations of cognitive, sensorimotor, attentional, emotional, planning, imagination, and memory-related tasks. Many of these functions are in use even during sleep. Thus, any null model of multivariate timeseries

Spectrally Matching Gaussian Noise to a Fixed Narrow-Band Target Spectrum Does not Ensure Absence of Evident Temporal Epochs

White Noise Spectrally Matched to the Spectrum of $S(t)$:
60 Realizations (Displayed as 3 Sets of 20 Time series)

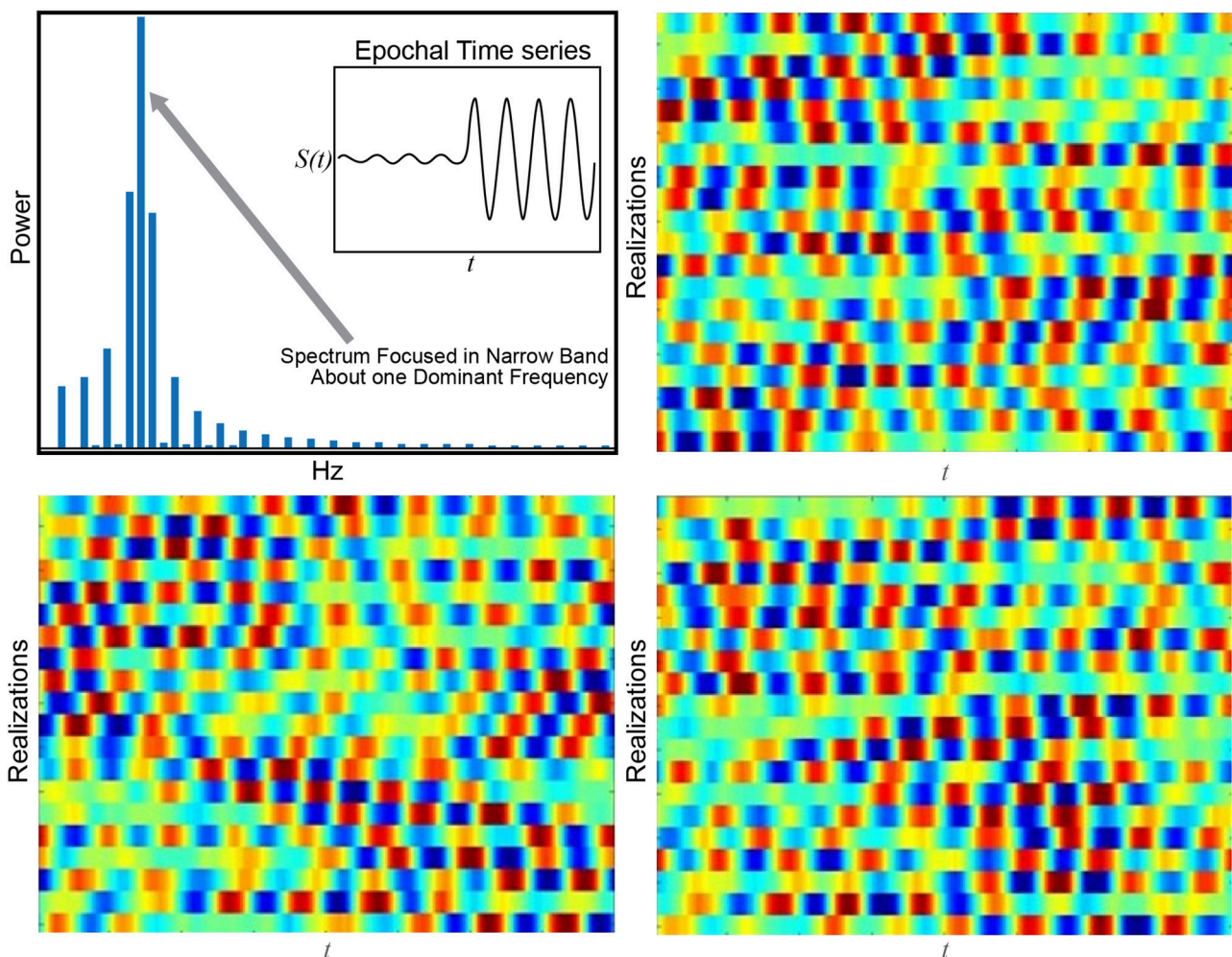


FIGURE 3 | The signal $S(t)$ consisting of a low-amplitude 0.08 Hz segment followed by a high-amplitude 0.08 Hz segment is one manifestation of a signal with a narrow-band spectrum (shown top left) focused at 0.08 Hz. Matching 60 timeseries of Gaussian white noise to this spectrum yields (in sets of 20) the timeseries shown in the other three panels of this figure. It is evident that spectral-matching of Gaussian noise to a generic template spectrum can contain discernible temporal epochs with any given realization. Thus, such a model cannot be a good null model for dynamic connectivity as it will contain the very dynamics we are interested in studying.

whose characteristics are highly consistent with empirically observed fMRI-based brain measurements has little utility (notably in the context of functionally relevant brain dynamics), since the phenomenon that it is testing for is ubiquitous rather than rare. The space of SCC multivariate Gaussians replicates real fMRI network timeseries with sufficient fidelity to induce broad consistency in measurable characteristics between the simulated data and the empirical data it was modeled upon (Liégeois et al., 2017). Moreover, there is no *a priori* reason to believe that aberrant or “tail” phenomena in this space should be more strongly associated with functionally-relevant brain dynamics than with measurement noise, motion or other artifacts, e.g., the sort of features that might warrant examining a scan for possible removal rather than positioning it as an exemplar of functionally-relevant resting state brain dynamics.

Multivariate Signals Lacking Plausible Markers of rs-Brain Dynamics (Valid Null Models for Brain Dynamics)

The development of valid null models for rs-FRBD is substantially hindered by a dearth of fMRI recordings from living subjects under conditions that all but preclude the ongoing CASE shifts unavoidably present even in sleeping or mentally impaired subjects. Resting state fMRI data is generally recorded under conditions in which functionally relevant brain dynamics ought to be continually present. Thus, the measurable features of empirically observed multivariate fMRI network timeseries are intractably “contaminated” from the standpoint of parameterizing a null space in which signal properties reflecting rs-FRBD are ensured to be rare.

Another challenge for hypothesis testing of rs-FRBD resides in identifying quantifiable signal features for which every upward increment of the associated measure unambiguously yields stronger evidence for the presence of rs-FRBD. Without this property, observations from the distributional tails of the measure are simply improbable, but not necessarily in ways that relate to FRBD. Thus, the first-layer of challenges is posed by our limited understanding of the signal properties whose variation through time reflect some shift in one of the brain’s myriad comingling functions. These are often amplified by non-monotonic relationships between those properties and the neural functions they putatively reflect. One example of such a property is kurtosis (see Technical Supplement), a higher-order statistical moment that has been employed (Laumann et al., 2016) to gauge the (presumably function-relevant) “dynamic-ness” of simulated and empirical fMRI signals. Univariate kurtosis captures the “peakedness” of a unimodal distribution; it rises with the number and magnitude of observations in a sample that would be outliers if the underlying process was stationary Gaussian, i.e., a Gaussian with constant mean and variance. Modestly elevated kurtosis might well reflect some unusually strong or active brain dynamics – this would have to be demonstrated, but is not implausible. However, extremely large kurtosis values occur when a sample contains numerous wildly large (in magnitude) observations. A brain recording with these characteristics is more likely to suggest noise, motion or other nuisance factors than anything connected

with actual rs-FRBD occurring during the scan (see **Figure 5**). More generally though, the effort to statistically validate the presence of a phenomenon that is almost axiomatically continual in any valid recording seems misguided. The more interesting hypotheses about fMRI dynamics would not focus on *whether* they exist, but rather on how they might manifest differentially over different timescales, spatial scales and functional scales.

MEASURING FUNCTIONALLY RELEVANT BRAIN DYNAMICS VS. IDENTIFYING OUTLYING OBSERVATIONS

A valid metric of brain dynamics should rise monotonically with the strength of the signal features that are, at our current level of scientific knowledge, widely believed to have associations with CASE or task-driven shifts in brain function. The metric should also be as blind as possible to signal features believed to represent nuisance factors. The problem of quantifying signal features that have a high likelihood of representing evolving CASE states is admittedly very difficult. Every procedure will be biased by assumptions whose validity cannot be ensured based on the current state of knowledge. The best we can hope for right now is for researchers to be clear about what signal properties they are seeking to quantify, and why they believe that a strong presence of these properties should be taken as evidence of FRBD (or of nuisance factors that are not easily separable from FRBD at current levels of measurement resolution).

We show in this section that a measure based on kurtosis, while sensitive to outliers, is not an ideal metric to capture brain dynamics and it is quite easy to show that kurtosis can be more sensitive to very rare outliers than it will be to more prevalent FRBD. We show in both stylized examples but also in real data that kurtosis preferentially captures signal features likely to arise from measurement disruptions (e.g., motion), while suppressing evidence of more extended spectral epochs within network timeseries. We also propose a new metric, Φ , which we believe shows some of the desired properties.

Univariate and Multivariate Kurtosis Under Stationary Gaussianity

Univariate excess or normalized kurtosis, the fourth statistical moment rescaled by squared variance and centered by subtraction of 3 has a well-understood distribution under the Gaussianity assumption that is applicable to Laumann et al. (2016) and here. Under this assumption, there is a closed-form transformation (dependent on the sample size, n) that converts observed excess kurtosis $b = \frac{\mathbb{E}((X-\mu_X)^4)}{\mathbb{E}((X-\mu_X)^2)^2} - 3$ into an unbiased estimator of true kurtosis, $\mathcal{B}_n$ (see Technical Supplement) that is distributed as a standard normal $\mathcal{N}(0, 1)$ (Cain et al., 2016) and hence now has altered limits. The use of $\mathcal{B}_n$ enables statistical evaluation of departures from stationary Gaussianity. When we refer to values of univariate kurtosis, these are values of the unbiased estimator $\mathcal{B}_n$. Samples that yield elevated values of $\mathcal{B}_n$ (say, $\mathcal{B}_n > 1.96$, so that $p < 0.05$) contain high-magnitude observations that are too numerous and/or too extreme for the

sample to have even a 5% chance of having been generated by a stationary Gaussian process.

There is a similar transform for Mardia's multivariate kurtosis (m.v. kurtosis), with a similar interpretation. The unbiased estimator $B_{n,p}$ (dependent on sample size, n , and vector length, p , see Technical Supplement) for Mardia's multivariate kurtosis (Mardia, 1970) is:

$$B_{n,p} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_i)^T C^{-1} (x_i - \mu_i)$$

where C^{-1} is the $n \times n$ inverse covariance matrix of the time-indexed p -vectors $\{x_1, x_2, \dots, x_n\}$.

Wavelet-Based Metric of Spectrally-Distinguishable Temporal Epochs

We briefly introduce a novel metric $\Phi: \mathbb{R}^T \rightarrow \mathbb{R}$ that explicitly captures the presence of spectrally distinguishable temporal epochs in a timeseries (see **Figure 4** and the Technical Supplement). The metric has a natural multivariate extension $\tilde{\Phi}: \mathbb{R}^{N \times T} \rightarrow \mathbb{R}$ that evaluates spectrally distinguishable temporal epochs in multivariate timeseries. Φ is not a primary focus of this short paper, but it plays a role in the discussion that follows because it provides a more targeted measurement than, for example, kurtosis, of timeseries characteristics that could form FRBD.

Epochal Stationarity and Kurtosis

The presence of spectrally distinguishable temporal epochs across realizations, i.e., multiple subjects' connectivity characteristics in individual or multivariate network TCs is one reasonable potential form of evidence for rs-FRBD. Although it is also possible that this type of phenomenon could arise from nuisance factors, epochal behavior has structure that makes it less likely to be sourced dominantly in nuisance factors such as motion, measurement noise or physiological rhythms. For ease of exposition we will introduce some terminology, *epochal stationarity* (resp. *epochal non-stationarity*), to indicate the absence (resp. presence) of signal variation that presents *within* a given empirical observation as a sequence of one or more spectrally distinguishable temporal epochs. Kurtosis, which can help identify the presence of outliers in Gaussian data, has been proposed as a metric to detect FRBD, however kurtosis is highly susceptible to unstructured amplitude variations. Moreover, as an outlier metric, kurtosis has greater sensitivity to sharp, transient, high-amplitude anomalous intervals than to signals with amplitude and frequency variation on more functionally relevant timescales (see **Figure 5**). In fact, the properties leading to elevated kurtosis are sometimes more present in an epochally stationary signal than in an epochally non-stationary signal (see **Figure 6**), i.e., one that is stationary except within a given duration. This is not to say that measuring epochal non-stationarity is straightforward. There are many features and timescales on which the non-stationarity might be exhibited, and most prospective metrics will present some combination of

over-sensitivity to irrelevant features and blindness to important features and/or timescales. We are currently working on a flexible, tunable approach to capturing the kind of epochally structured frequency domain variation that promises to provide valid evidence for brain dynamics after careful evaluation of sensitivity to nuisance factors.

Empirical Data and Simulation Regimes

A set of network timecourses from a clinical rs-fMRI study on which dynamic functional network connectivity (FNC) results have already been published (Allen et al., 2014), and five simulation regimes modeled on that data are employed to explore and illustrate the role of Gaussianity and statistical stationarity as well as, spectral and covariance stationarity in modeling rs-FRBD (and/or its absence). It is important to note that we refer to stationarity in the true sense as statistical (non)stationarity in what follows. We use the terms "spectral (non)stationarity" and "covariance (non)stationarity" to refer to other definitions used including (Laumann et al., 2016) that invoke the concept of (non)stationarity through the analysis of a single realization either in the spectral domain (for the former definition) or using covariance function (for the latter definition). The approaches are lightly outlined here, with more details provided the Technical Supplement.

Real Data

We used previously published (Allen et al., 2014) network timecourse data from a large multisite clinical resting-state fMRI study. These timecourses (314 subjects, 47 networks, 158 timepoints), subsequently filtered for frequencies at most 0.08 Hz and z-scored, are referred to below as "Real Data" (Allen et al., 2014) (see **Figure 7**; top left). The average power in each frequency bin in $[0.003, 0.08]$ Hz for all network TCs for all subjects is denoted $\bar{P}_{data}$. The average cross-network covariance matrix for all subjects is denoted $\bar{C}_{data}$. Every simulation regime described below consists of 1000 simulated subjects, each characterized by a set of 47, length-158 timeseries.

SCC Gaussians: Statistically Stationarity Without Constraint on SDTEs

Following (Laumann et al., 2016), each simulated subject in the SCC Gaussian regime is a multivariate timeseries resulting from the projection of a 47×158 matrix of low-pass filtered white noise spectrally matched to $\bar{P}_{data}$ onto the eigenspace of $\bar{C}_{data}$ (see **Figure 7**; top middle).

Covariance-Dynamic SCC Gaussians: Piecewise Stationary, Two Distinct Covariance Regimes, No Constraint on SDTEs (CD-SCC Gaussian)

This regime introduces explicit covariance non-stationarity. Each CD-SCC Gaussian subject starts as a 47×158 matrix consisting of low-pass filtered white noise spectrally matched to $\bar{P}_{data}$, which is then divided into three windows, determined by a middle window of randomly chosen length between 40 and 60 TRs. The middle window is projected onto the eigenspace of $\bar{C}_{data}$, which has structure and is strongly connected, while the first and final windows are projected onto the eigenspace of C_{weak} , a covariance

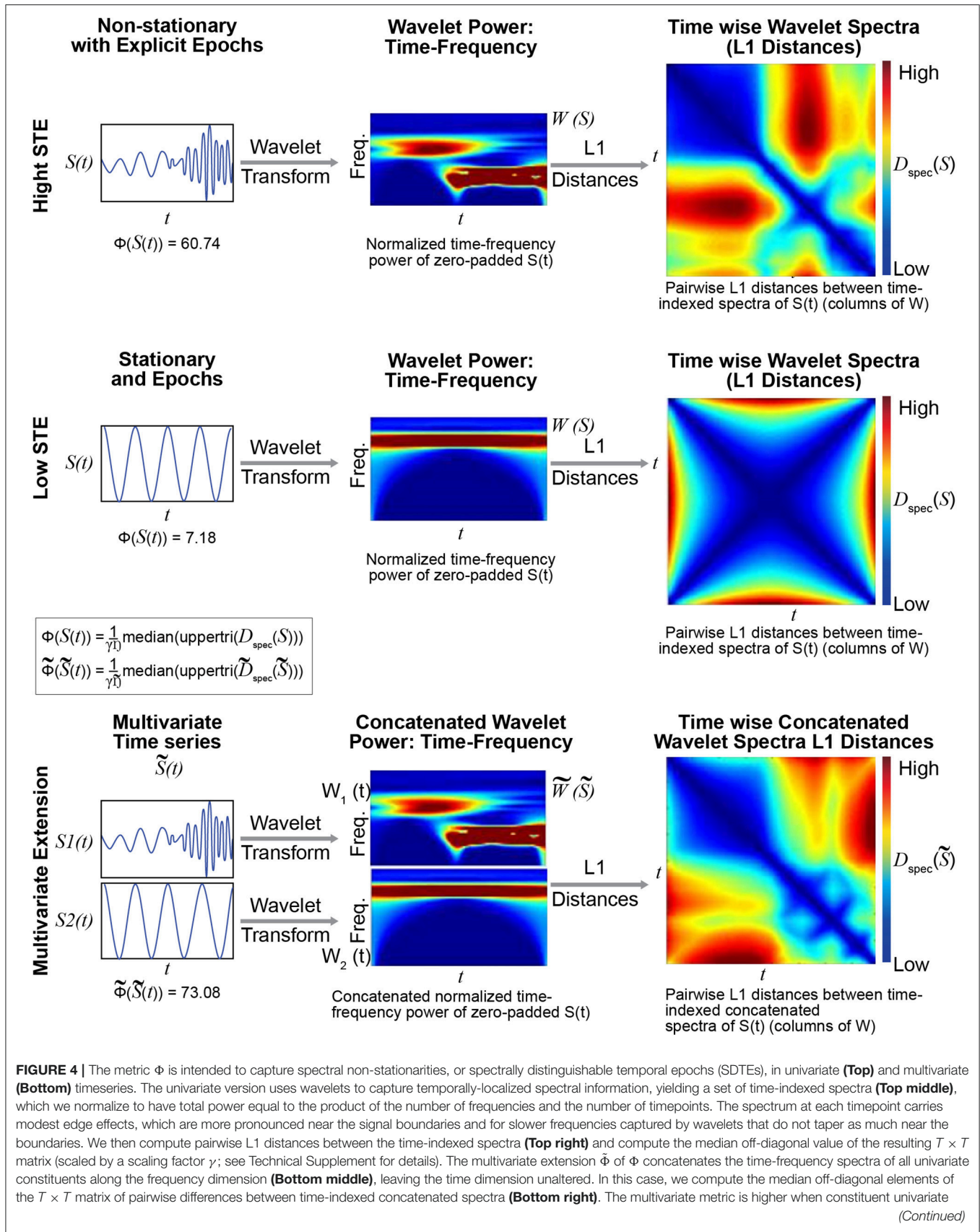


FIGURE 4 | time-frequency spectra exhibit their largest within-signal spectral differences in mutually distinct temporal intervals. The case shown here does not illustrate the role of differential intervals of SDTEs among constituent timeseries. The multivariate extension of this metric was successful in capturing slight changes in frequencies in several consecutive short term epochs too (Supplementary Figure 1). Finally, in all cases, the results were highly similar for both L1 (Manhattan) and L2 (Euclidean) distances (multivariate Φ outperformed multivariate kurtosis for both distance measures).

matrix reflecting very weak unstructured network connectivity (see **Figure 7**; bottom left).

SCC Gaussians With Noise: Statistically Stationary With a Single Spike Randomly Inserted Into a Small Proportion of Network Timeseries (“Noisy SCC Gaussian”)

This regime introduces extremely sparse, high-amplitude noise to the SCC Gaussian setting. Each Noisy SCC Gaussian subject starts as an SCC Gaussian subject, i.e., as a 47×158 matrix consisting of low-pass filtered white noise spectrally matched to $\bar{P}_{data}$ then covariance matched to $\bar{C}_{data}$. Of the 47 timeseries in this matrix, between 3 and 15 are selected at random to carry a single high frequency spike centered at some randomly selected timepoint. The entire multivariate timeseries contains between 3 and 15 of these noise artifacts, with at most one in any given univariate timeseries (see **Figure 7**; top right).

Spectrally and Statistically Non-stationary: Explicitly Non-stationary Both Statistically and Epochally (“SS Non-stationary”)

This regime introduces simulated task-responsiveness to the SCC Gaussian setting. Each simulated SS Non-stationary subject starts as an SCC Gaussian subject (a 47×158 multivariate timeseries produced by subjecting white noise to spectral and covariance constraints exhibited by the real data). The 47 networks for each subject are divided into at most 17 task-positive networks (TPNs), at most 11 task-negative networks (TNNs), with the remaining $19 - 29$ networks designated as non-responders (NRs). Following the same procedure employed for the CD-SCC Gaussian regime, multivariate timeseries in the SS Non-stationary regime are divided into three windows determined by a middle window of randomly-chosen length between 40 and 60 TRs. The hypothetical task takes place during the middle window, in which (relative to the first and final window) the selected TPNs exhibit faster, higher amplitude behavior, the selected TNNs exhibit slower, lower amplitude behavior and the NRs exhibit no change in behavior (see **Figure 7**; bottom middle).

Covariance-Dynamic Spectrally and Statistically Non-stationary: Explicitly Non-stationary Both Statistically and Epochally With Two Distinct Covariance Regimes (“CD-SS Non-stationary”)

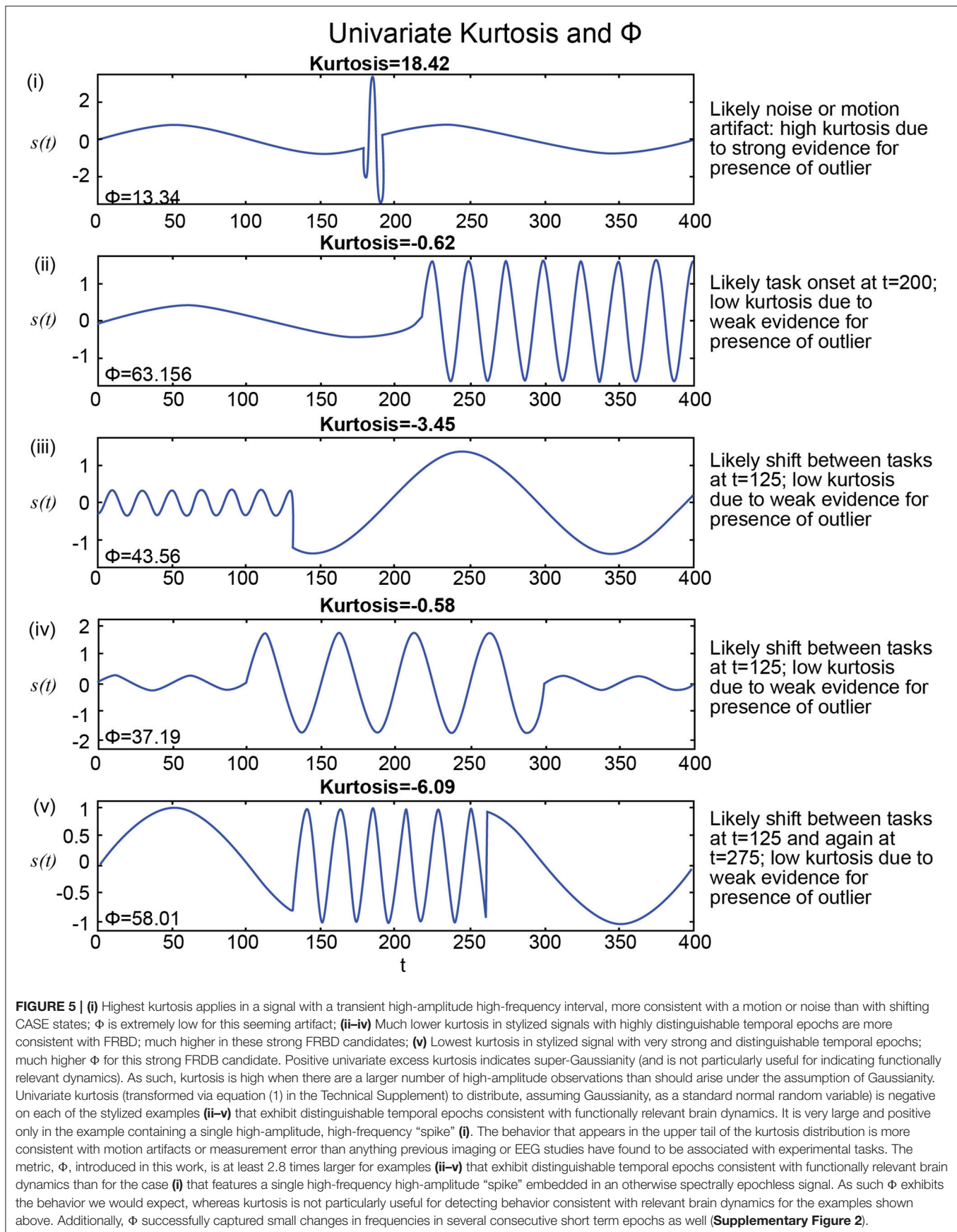
This regime introduces explicit covariance non-stationarity to the SS Gaussian setting. Each simulated CD-SS Non-stationary subject starts as an SS Non-stationary subject (see immediately above). In this regime, however, the temporally task-responsive middle window is additionally subjected to explicitly different covariance constraints than the task-free first and final windows. Following the procedure from the CD-SCC Gaussian, the middle window of CD-SS Non-stationary subjects is projected onto

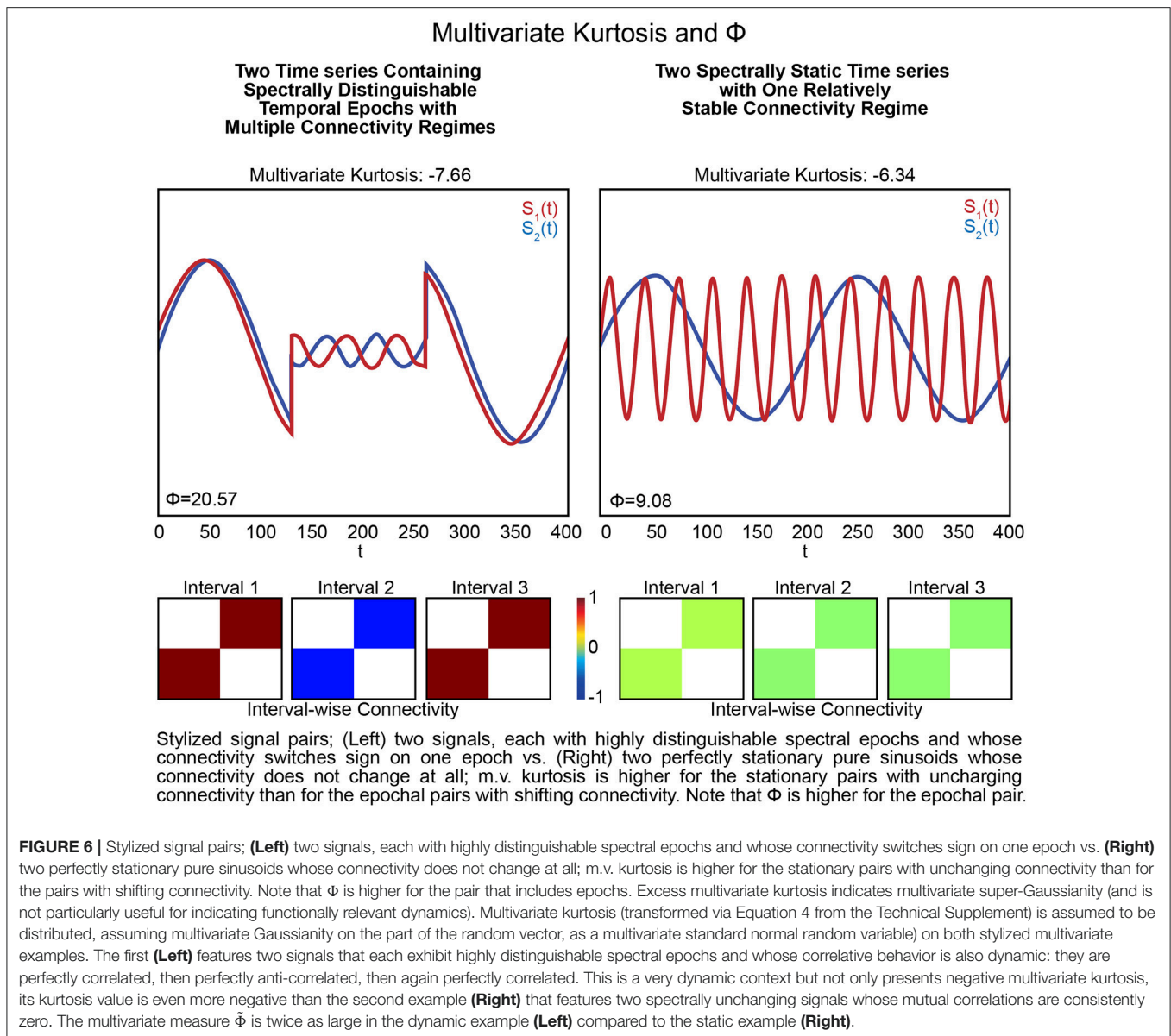
the eigenspace of $\bar{C}_{data}$, while the first and final windows are projected onto the eigenspace of C_{weak} (see **Figure 7**; bottom right).

Kurtosis Is Overly Sensitive to Noise Artifacts and Very Rarely Identifies Explicit Spectral, Statistical, and Covariance Non-stationarities

As mentioned earlier, the utility of kurtosis as a measure of functionally relevant brain dynamics is mitigated by its highly-tuned sensitivity to spikes and outliers in the data. We saw this earlier in stylized univariate timeseries (see **Figure 5**). The issue is equally evident (see **Figure 8**) in more complex multivariate simulations involving different degrees of non-stationarity, exhibited in different ways. Employing m.v. kurtosis as a measure of FRBD implies that timeseries on which it assumes values large enough to provide significant evidence ($p < 0.05$) against the null hypothesis of stationary multivariate Gaussianity are those in which the features associated with FRBD are most markedly present. More specifically, these features increase monotonically with m.v. kurtosis. In **Figure 8** however, we see that that the only the simulation regime in which a non-negligible percentage (93%) of realizations is identified as exhibiting FRBD using m.v. kurtosis is the regime featuring a handful of spikes in an otherwise stationary 47×158 multivariate Gaussian. Only 5.3% of realizations from the explicitly covariance-dynamic regime are identified using the m.v. kurtosis metric as exhibiting FRBD, and for the other non-stationary regimes the percentage of realizations identified as exhibiting FRBD is less than a tenth of a percent. Similarly, the actual data recorded from subjects undergoing continual CASE-shifts exhibits no evidence by the m.v. kurtosis criterion of having arisen from a source in which functionally relevant brain dynamics are present. This clearly highlights the limitations of kurtosis as an indicator of FRBD. Elevated kurtosis indicates the sample contains too many points that are too extreme in magnitude under an assumption of stationary Gaussianity; it is less effective at capturing the dynamic changes in temporal behavior (including covariance) that are more plausible markers of FRBD and are richly present in most realizations of even stationary multivariate Gaussian processes.

Conversely, as shown in **Figure 9**, the proposed metric (defined in **Figure 4** and the Technical Supplement) responds in a more reasonable way to those features of real and simulated multivariate timeseries that have strong likelihood of reflecting FRBD vs. those that are simply aberrant in some other way. Results show that, in contrast to multivariate kurtosis, the lightly contaminated SCC Gaussians in terms of $\tilde{\Phi}$ exhibit significantly less evidence of reflecting FRBD than either of the SS Non-stationary regimes (**Figure 9**, row 2, columns 5 and 6) and real data is statistically indistinguishable from

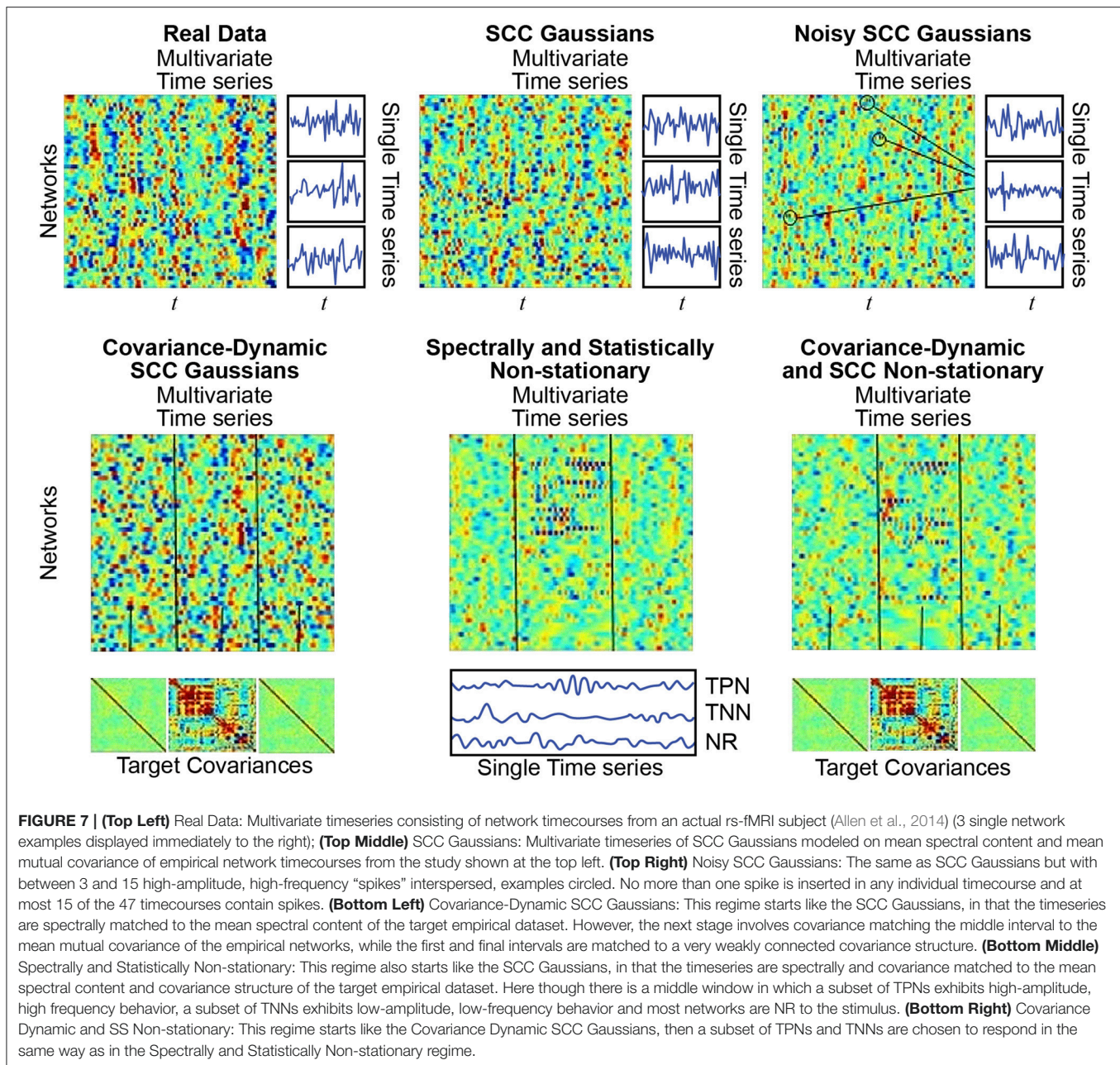




the covariance-dynamic SCC Gaussian regime (Figure 9, row 1, column 4). So $\tilde{\Phi}$ is putting regimes exhibiting different kinds of temporally epochal behavior—including the epochal behavior seen in statistically stationary Gaussian processes—in what seems a plausible ordering with respect to dynamism: SCC Gaussians $\preceq$ Real Data $\approx$ Covariance-Dynamic SCC Gaussians $\preceq$ Noise-Contaminated SCC Gaussians $\preceq$ SS Non-stationary $\preceq$ Covariance-Dynamic SS Non-stationary (where curly binary relations indicate ordinal evidence of potentially relevant multivariate epochal behavior as measured by $\tilde{\Phi}$ (see legend for more details).

This plausible ordering with respect to the t -statistic values from the analysis is much more meaningful for the Φ metric than multivariate kurtosis. The spectra for the SCC Gaussians are more uniformly spread than the real data (please note that it is only the “average spectral power” that is matched to

the real data) which is potentially one reason why the SCC Gaussians exhibit less evidence of dynamics than the real data. On the other hand, the multiple high frequency spikes add to temporal variation in the spectral content of the SCC Gaussians to which they are introduced, therefore resulting in false, elevated values of the multivariate metric in the modified (noisy) SCC Gaussian regime. Finally, both of the non-stationary classes are, by construction, explicitly non-stationary both statistically and epochally; hence, they are expected to exhibit greater evidence of dynamics. However less evidence of dynamics in real data than these two particular classes suggests that real data is not as non-stationary in nature as these explicitly non-stationary classes. In any case, the preliminary advantages of using a raw measure of this type are clearly evident from the meaningful patterns conveyed by the pairwise t -tests on metric values from the different regimes (Figure 9). Still, a lot more developmental and validation work



is required to identify an exhaustive statistical measure most appropriate for the entire possible range/type of spectral and amplitudinal variations (or controlling ability for different classes of noise) in the resting state time-courses, and that validates as being statistically significant.

CHOOSING A TEST METRIC THAT YIELDS DIFFERENT DISTRIBUTIONAL TAILS FOR NULL AND ALTERNATIVE HYPOTHESIS

A key challenge in the use of simulated data for hypothesis testing is identification of context-appropriate null generating models

and test statistics that, in combination, allow for meaningful inferences.

Specifically, the phenomenon being investigated should not be generically present (should have low probability of occurring) in data from the generating model, and (as also discussed in Hindriks et al., 2016) the test statistic should have highly specific, strictly monotonic sensitivity to the investigative target phenomenon (values deeper in the relevant tail of the test statistic distribution should explicitly correspond with increased presence of that phenomenon). The field of dynamic connectivity in particular has struggled with this as there are multiple threads of research proposing use of different models of dynamic behavior including covariance dynamics (Hutchison et al., 2013; Allen

Percentage of Multivariate Observations Found to be Non-stationary or Dynamic by the Multivariate Kurtosis Metric

Multivariate Kurtosis (Percentage of observations whose z-transformed multivariate kurtosis is elevated enough to offer significant ($\alpha=0.05$) evidence against the observation being a stationary multivariate gaussian)

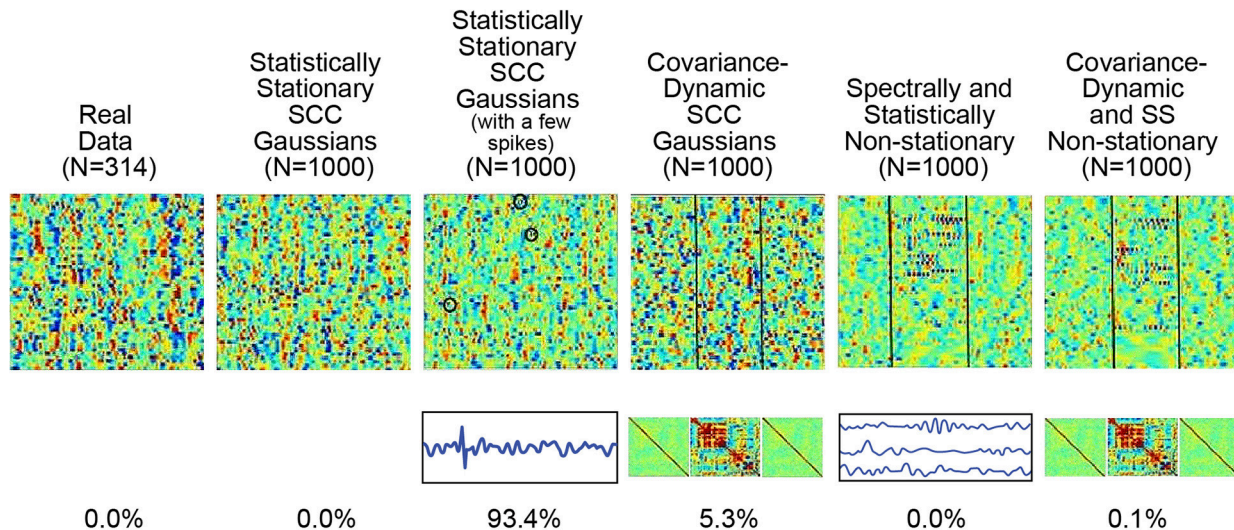


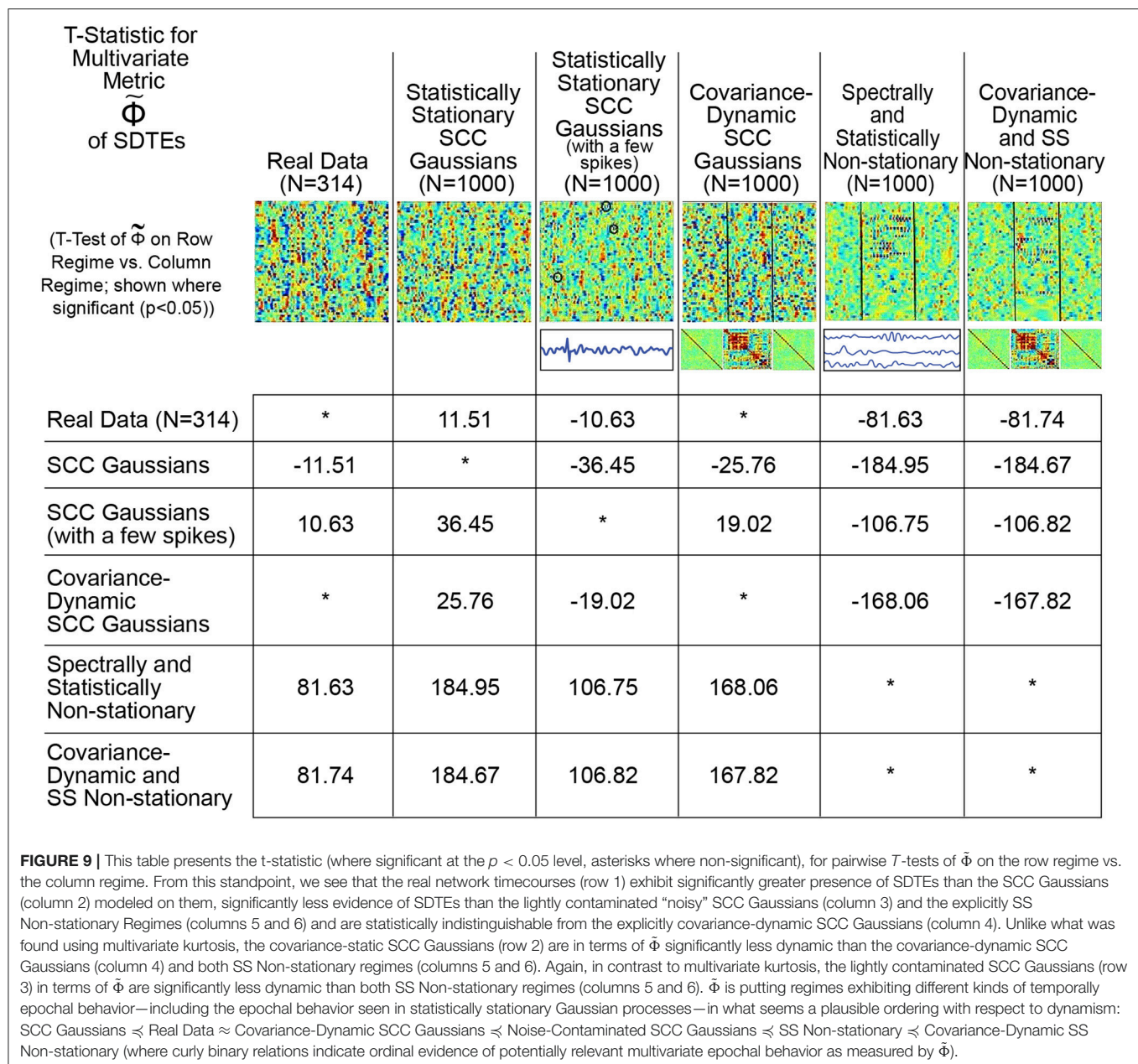
FIGURE 8 | The percentage of multivariate timeseries from each indicated simulation regime (and the real rs-fMRI study on which the simulated data was modeled) that present significant evidence ($p < 0.05$) of having been generated by some process that is not a stationary multivariate Gaussian. The explicitly implemented spectral and statistical non-stationarities of the two SS Non-stationary regimes (columns 5 and 6) are effectively never found to exhibit significant evidence against being generated by stationary multivariate Gaussians. In the two explicitly covariance-non-stationary regimes (columns 4 and 6), one of which is also spectrally and statistically non-stationary variations (column 6), at most 5% of the 1,000 simulated subjects—each of which exhibits the explicit non-stationarity—are identified as unlikely to have arisen from a stationary multivariate Gaussian process. Like the SCC Gaussian simulations (column 2), the SS Non-stationary simulations (columns 5 and 6) and the covariance-non-stationary simulations (columns 4 and 6), empirical observations from real subjects (column 1) in whom the phenomenon of interest (FRBD) is ubiquitous are not statistically distinguishable from realizations of a stationary multivariate Gaussian process. The only regime that multivariate kurtosis reliably distinguishes from realizations of a stationary multivariate Gaussian process is the case of SCC Gaussians in which a single high-amplitude, high-frequency spike is inserted into between 3 and 15 of the 47 univariate timeseries from the multivariate observation (column 3). This regime is basically just a lightly contaminated version of the SCC Gaussian regime (column 2) and of all of the simulation regimes exhibits the least evidence of functionally relevant brain dynamics. The behavior underlying upper-tail observations of multivariate kurtosis looks more like scan contamination than anything task-paradigm fMRI studies suggest would be strongly associated with FRBD.

et al., 2014) and oscillatory dynamics (Golos et al., 2015). In this context, it is very important to apply any proposed null model only within the narrow conditions under which it is able to reject a particular hypothesis about brain dynamics. For example, as seen earlier and also shown in Liégeois et al. (2017), a null model based on stationary oscillatory dynamics may well produce data exhibiting dynamically-changing covariances with high probability, something that would impair the model's ability to function as a proxy for the null hypothesis in studies of dynamic connectivity (Hindriks et al., 2016). In addition to that, even models of covariance dynamics have been shown to be limited to a rather particular set of parameters, which we do not yet fully understand how to set for the human brain (Shakil et al., 2016). There is also evidence that the brain functions as a non-linear dynamical system (He, 2013). The result is that many of the null models that have been proposed are making strong assumptions about brain dynamics which, while having some justification, are not able to rule out dynamics of a different sort. We

provide some additional discussion on this important point below.

Whole-Brain Windowed Connectivity States and Occupancies From Statistically Stationary and Explicitly Non-stationary Null Models Strongly Resemble Each Other

As we have discussed previously, creating a valid null model for rs-FRBD is difficult without explicitly understanding the properties that multivariate network timeseries might exhibit in response to complex CASE variations. One approach, suggested recently by Laumann et al. (2016), employs a space of low-pass filtered multivariate white noise, spectrally matched to the average spectrum of empirical timecourse data and then projected onto the eigenspace of empirically observed mean network covariance. As discussed in earlier sections, this is a space of timeseries explicitly modeled on real data recorded from a material in which the phenomenon of interest (i.e.,



functionally relevant brain dynamics) is continually present. It is therefore not a space in which univariate or multivariate temporal behavior plausibly sourced in CASE variations is vanishingly rare. Which is to say, it is not a useful null space for the identification of rs-FRBD. Moreover, it is a space of signals whose time-varying behavior is jointly determined by all simulation parameters and assumptions: the auxiliary spectral and covariance constraints as well as the Gaussianity assumption and the primary assumption of statistical stationarity. Above and beyond the problematic success of this simulation model in replicating empirical data recorded under circumstances in which the phenomenon of interest is uninterrupted and continual, this simulation regime's value as a null model is further undermined by the unexamined role of auxiliary parameters

in shaping key measures and distributions. This can be seen in **Figure 10**, rows 2 and 3, where we break the model's core assumption of statistical stationarity, a property, arguably incorrectly, associated by the authors with an absence of FRBD, without discernibly disrupting either the clusters formed by short-timescale FNCs or the average cluster occupancy rates (see the Technical Supplement for a brief background on this approach).

The distributional tails of occupancy rates for each connectivity state in the SCC Gaussian and the SS Non-stationary regimes have significant overlap (see **Figure 11**) and as such this model can quite easily rule both for and against dynamic connectivity at the same time, an obvious flaw in the approach (i.e., evaluating statistical significance

of measures of a phenomenon of interest generically present in a model chosen as the null as well as its alternative). This example illustrates the difficulties of building hypothesis-testing frameworks for phenomena whose distinguishing quantifiable characteristics are not well-understood. If, in contrast to Laumann et al. (2016), one realizes that the space within which one is working contains the very dynamics that one is trying to rule out (a point subsequently made by Liégeois et al., 2017) the conclusions that can be made are unconvincing and uninteresting.

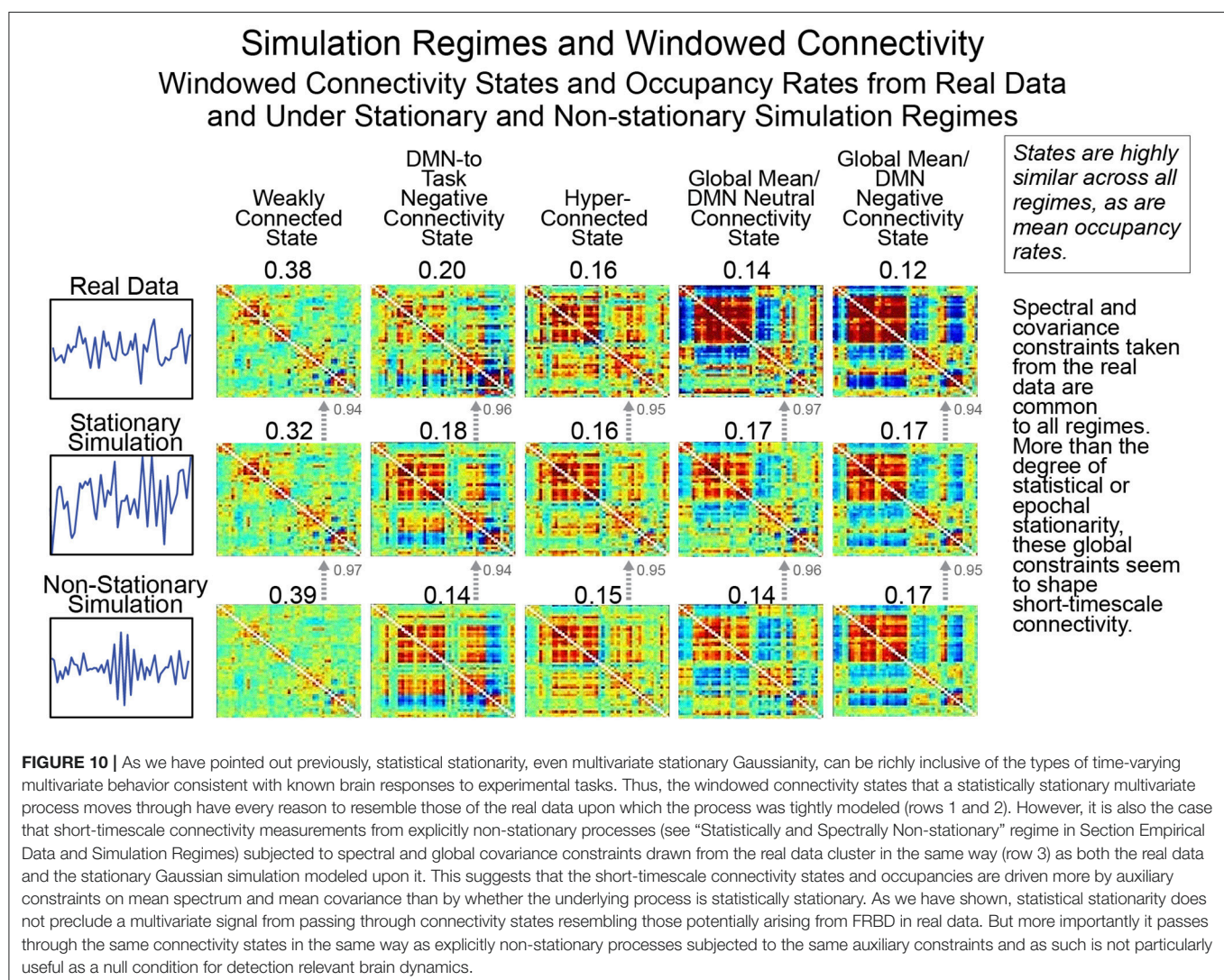
DISCUSSION

In this position paper, we have attempted to clarify and develop some of the important issues related to dynamic connectivity within the resting brain. Our focus has mostly centered around the use of metrics to detect possible dynamic behavior and also the creation of appropriate null models of functionally relevant brain dynamics, including but not limited to dynamic connectivity. We have discussed crucial limitations

hindering some existing proposed models, demonstrated how a metric such as the one used in this paper could effectively detect potential properties of functionally relevant brain dynamics, and have, we hope, provided the context for a rich ongoing discussion of where the field should head from here. In the remainder of this paper, we highlight a few high-level questions that we hope make this point clear.

Are Functional Brain Dynamics Rare?

One of the key points we emphasized relates to the creation of a null simulation model. A useful null simulation model should be built on assumptions complementary to those that would apply if the phenomenon of interest were present, and then combined with a test statistic that is sufficiently sensitive to the null assumptions that it very rarely achieves extreme values when they are absent. Due to these concerns (as illustrated in the paper earlier), the null model in Laumann et al. (2016) does not seem to be the right null model to evaluate functionally relevant brain dynamics with time-varying connectivity methods. A particular



Simulation Regimes and Lack of Hypothesis-Separating Distributional Tails Overlapping Tails in the Occupancy Distributions of Windowed Connectivity States Under Stationary and Non-stationary Simulation Regimes

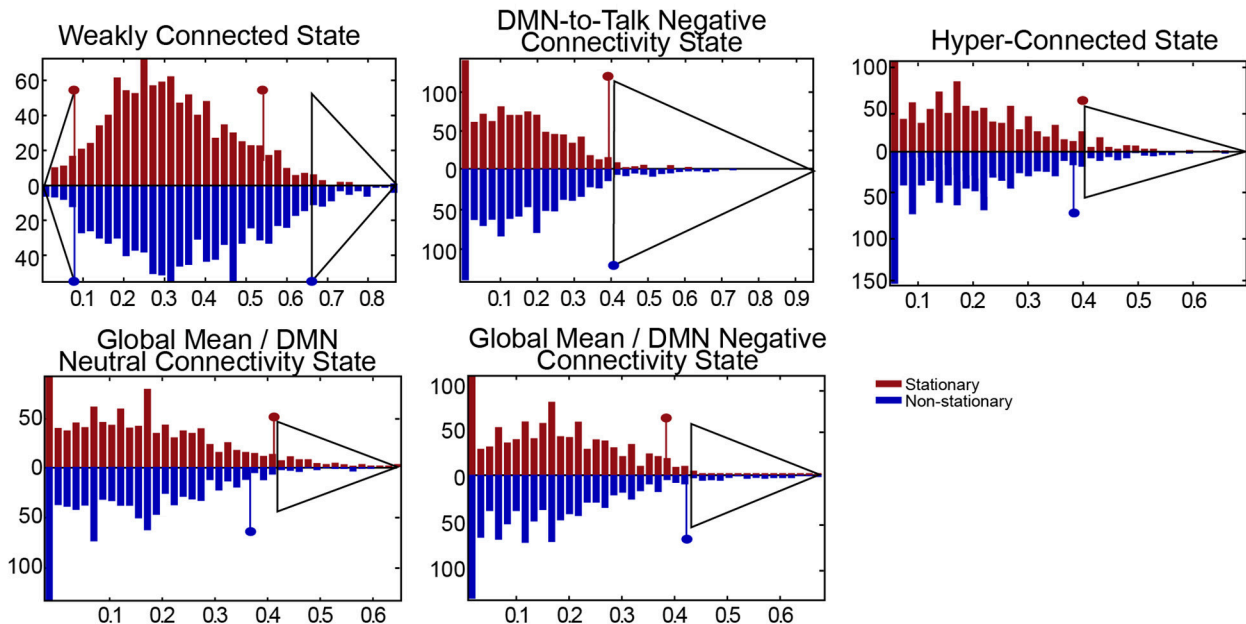


FIGURE 11 | For a combination of null model and test metric to allow for meaningful inferences, the null model should exhibit evidence of the phenomenon being investigated (and believed to be captured through the test metric) extremely rarely. SCC Gaussians arguably present evidence of FRBD as ubiquitously as the real fMRI data does. It is also not clear a priori that tail phenomena in SCC Gaussians should have an especially strong association to FRBD rather than, for example, to motion artifacts or measurement noise. Moreover, the measures evaluated on a null model are only helpful in identifying behavior of interest if the distributional tails of the measure do not overlap with those of the same measure evaluated on a model that violates assumptions of the original null hypothesis. The upper tails of the distributions of occupancy rates on the windowed connectivity states from SCC Gaussians (red) and the SS Non-stationary simulation regime (blue) have substantial overlap (semi-transparent gray triangles indicate the overlapping part of the tails for each state). Insufficient clarity on the unique, distinguishing features of the phenomenon being studied and of the models/measures being employed neutralizes the measure's role in hypothesis testing as the same observation can then present significant evidence against multiple, mutually contradictory null hypotheses. In terms of occupancy rates of windowed connectivity states, we see here that a newly scanned subject who spends 65% of their time in state 2 exhibits significant evidence against one statistically stationary null hypothesis and also against an explicitly non-stationary null hypothesis.

and carefully constructed null model was proposed in Hindriks et al. (2016), a study that evaluated statistical significance of variance as a test statistic to assess time-varying connectivity and emphasized importance of appropriate surrogate data testing in time-varying studies. The distributional tails of this test statistic on their static null simulation model matched the tails of variance computed on a wide range of explicitly dynamic variants of the model in which parameters were meaningfully changing with time. As also clearly pointed out by authors in this study (Hindriks et al., 2016), such observations would simply suggest lack of sensitivity of this metric to separate out real fMRI data from stationary null model directly and would not suggest absence of brain dynamics in the fMRI data. Additionally, partially contrasting results have indeed been reported using the same null model (or null model retaining similar properties of the data) on the same metric (Laumann et al., 2016) or other time-varying measures factoring in this metric (Hindriks et al., 2016). These studies have provided evidence that the stationary, linear and Gaussian model does not capture all the

properties of the fMRI data; hence, the study of functional brain dynamics presents challenges to the utility of this entire type of modeling (to fit or study fMRI data). Moreover, we know that functional brain dynamics are constantly present in conscious human beings, so the phenomenon is anything but rare empirically.

Are Functional Brain Dynamics Separable From Nuisance Factors and Background Brain Rhythms?

The null model for a phenomenon necessarily produces, with very low probability, values of some test statistic that are consistent with the phenomenon which is being tested. This kind of model makes sense primarily when there is a very clear understanding of the range of values a particular test statistic will assume in the presence of the hypothesized phenomenon. So although in this short paper we are focusing primarily on the insufficiency of null models built on many

realizations of some statistically stationary multivariate process modeled on empirical timecourse spectra and covariance, the larger problem is really that our present understanding cannot rule out the possibility of non-trivial intersection between signal features foreseeably connected with brain dynamics and those arising from nuisance factors and background brain rhythms. Even a spectrally pure signal, e.g., a single-frequency sinusoid, features amplitude changes that could be consistent with the ebb and flow of a network's contributions to temporally varying CASE demands. Certainly pure sinusoids and other epochally stationary signals (whose phase-randomized stochastic analogs are statistically stationary) could provide evidence of network responsiveness in certain experimental task paradigms, e.g., those involving repetitive motor or sensory tasks. Since most temporal changes in a signal, including raw amplitude changes, could plausibly be correlated with some complex sequence of CASE conditions, the present state of knowledge makes it difficult to construct null models that can claim to yield, almost-exclusively, timeseries (multivariate or univariate) lacking features prospectively associated with brain dynamics.

This is true even when the null model is narrow. Indeed, it can be difficult to ensure that the test statistic being assessed does not have distributional tails roughly matched to the distributional tails of that statistic on a similarly narrow but explicitly "dynamic" model. In that case, we can easily conceive of examples where we are in the position of having an empirical measurement of the test statistic that simultaneously leads to rejection of one static null hypothesis and various related dynamic null hypotheses. In such a case, the desired test has been rendered essentially useless. Due to factors delineated earlier in this paper, statistically stationary null models are not necessarily well suited for evaluating functionally relevant brain dynamics in rs-fMRI. We expect functionally relevant temporal variations in brain activation to be constantly occurring throughout the experiment, and as the phenomenon of FRBD is better understood, improved null models will naturally emerge. We hope that the discussions herein are useful and promote thoughtful consideration of these important issues.

CONCLUSIONS

To summarize, serious and continuing investigation of dynamic multivariate brain activation patterns (including dynamic connectivity) is scientifically important and central to many core open questions in brain science. The time-varying measurements provided by BOLD fMRI currently play a vital supporting role in this overall project. We have discussed some of the limitations of existing null models and metrics for capturing dynamics, and provide initial evaluations of a new wavelet-based metric to demonstrate advantages of exploring more targeted measurements of time-series characteristics that form FRBD. While the evaluated metric appears to provide sensible results in a number of simulated scenarios, it still needs to be tuned and validated for an exhaustive range of spectral/amplitude

variations and different noise classes. Moreover, a broader framework will ultimately be necessary to not only locate evidence of FRBD *per se* in univariate and multivariate brain data, but to also identify specific timepoints at which signals, signal-pairs and arbitrary signal n -tuples yield evidence that an underlying functional shift was underway. Finally, we urge caution in the development of null models in the context of dynamic connectivity. Especially for studies in which subjects are not engaging in a common, narrow experimental task, the relevant features, temporal and spatial/functional scales are not yet well-understood. Specific well-defined questions about how particular signal features evolve on a range of spatial and temporal scales could produce more useful and testable hypotheses about how the brain signals we measure relate to high-level processes by which the brain organizes, directs and rotates through some of its central tasks: e.g., cognition, sense-making, generative thinking, memory-formation, memory-retrieval and emotion regulation, among others.

Finally, we would like to strongly emphasize that this paper is not an argument against the use of formal hypothesis testing in investigations of resting state dynamics. Nor is it an argument in favor of null spaces whose constituents have no recognizable relationship to actual brain data. In addition to tests for group differences or regression analyses (two cases in which there is always a well-defined null hypothesis), we are arguing in favor of narrower hypotheses that attach narrowly defined, non-ubiquitous neural/mental phenomena to a very narrow domain of signal characteristics. This will allow for null models that are, up to a limited set of signal properties, modeled on real fMRI data.

Following the choice of a narrow, non-ubiquitous target neural/mental phenomenon, we would advocate identifying some metric or test statistic λ on the data under which there is good reason to believe in the existence of an interval I (or region if λ is multi-dimensional) outside of which λ is both highly specific to and roughly monotonic in the presence (strength or abundance) of the target neural/mental phenomenon. This allows the construction of a useful null space built from simulated fMRI data that is as realistic as possible subject to the constraint that its elements induce values of λ outside of I at a rate consistent with the level of the target phenomenon an investigation is concerned with: very rarely if the target is present at levels of interest immediately beyond I , or less rarely if λ really is roughly monotonic and the target is only present at levels of interest as λ gets much further from I .

We would like to close by re-emphasizing our strong belief that hypothesis testing is crucial to advancing knowledge. The purpose of this paper is simply to highlight that in relatively young observational sciences the range of hypotheses that are rigorously testable at any given time might not extend to some questions of central importance to researchers. Most importantly, the set of rigorously testable hypotheses is constantly growing via ongoing vibrant interplay between more exploratory studies and

studies that leverage the existing set of testable hypotheses for scientific gains.

ETHICS STATEMENT

This study was determined by the University of New Mexico IRB not to be a human subjects study since the data is anonymized with no access to links to any identifying information.

AUTHOR CONTRIBUTIONS

RM made significant contributions to the conception and writing of the work, conducted the analyses presented. AA contributed to writing and analyses. YL-S contributed to writing. TA

contributed to writing. VC contributed to conception, analyses and writing and supported the work with grant funds.

FUNDING

This work was supported by National Institutes of Health (NIH) via a COBRE grant P20GM103472, R01 grants R01EB005846, 1R01EB006841, 1R01DA040487 and REB020407, and National Science Foundation (NSF) grants 1539067 and 1631838.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00551/full#supplementary-material>

REFERENCES

- Abrol, A., Damaraju, E., Miller, R. L., Stephen, J. M., Claus, E. D., Mayer, A. R., et al. (2017). Replicability of time-varying connectivity patterns in large resting state fMRI samples. *Neuroimage* 163, 160–176. doi: 10.1016/j.neuroimage.2017.09.020
- Allen, E., Damaraju, E., Eichele, T., Wu, L., and Calhoun, V. D. (2018). EEG signatures of dynamic functional network connectivity states. *Brain Topogr.* 31, 101–116. doi: 10.1007/s10548-017-0546-2
- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352
- Cain, M. K., Zhang, Z., and Yuan, K. H. (2016). Univariate and multivariate skewness and kurtosis for measuring nonnormality: prevalence, influence and estimation. *Behav. Res. Methods* 49, 1716–1735. doi: 10.3758/s13428-016-0814-1
- Calhoun, V. D., Miller, R., Pearson, G., and Adali, T. (2014). The chroconnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* 84, 262–274. doi: 10.1016/j.neuron.2014.10.015
- Chang, C., and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50, 81–98. doi: 10.1016/j.neuroimage.2009.12.011
- Cribben, I., Wager, T. D., and Lindquist, M. A. (2013). Detecting functional connectivity change points for single-subject fMRI data. *Front. Comput. Neurosci.* 7:143. doi: 10.3389/fncom.2013.00143
- Golos, M., Jirsa, V., and Dauce, E. (2015). Multistability in large scale models of brain activity. *PLoS Comput. Biol.* 11:e1004644. doi: 10.1371/journal.pcbi.1004644
- He, B. J. (2013). Spontaneous and task-evoked brain activity negatively interact. *J. Neurosci.* 33, 4672–4682. doi: 10.1523/JNEUROSCI.2922-12.2013
- Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., et al. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *Neuroimage* 127, 242–256. doi: 10.1016/j.neuroimage.2015.11.055
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promises, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079
- Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., and Menon, R. S. (2012). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum. Brain Mapp.* 34, 2154–2177. doi: 10.1002/hbm.22058
- Jones, D. T., Vemuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., et al. (2012). Non-stationarity in the “resting brain’s” modular architecture. *PLoS ONE* 7:e39731. doi: 10.1371/journal.pone.0039731
- Laumann, T. O., Snyder, A. Z., Mitra, A., Gordon, E. M., Gratton, C., Adeyemo, B., et al. (2016). On the stability of BOLD fMRI correlations. *Cereb. Cortex* 27, 4719–4732. doi: 10.1093/cercor/bhw265
- Liégeois, R., Laumann, T. O., Snyder, A. Z., Zhou, J., and Yeo, B. T. T. (2017). Interpreting temporal fluctuations in resting-state functional connectivity MRI. *bioRxiv.* 163, 437–455. doi: 10.1016/j.neuroimage.2017.09.012
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530. doi: 10.1093/biomet/57.3.519
- Preti, M. G., Bolton, T. A., and Van De Ville, D. (2016). The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* 160, 41–54. doi: 10.1016/j.neuroimage.2016.12.061
- Shakil, S., Lee, C. H., and Keilholz, S. D. (2016). Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states. *Neuroimage* 133, 111–128. doi: 10.1016/j.neuroimage.2016.02.074
- Shi, Z., Rogers, B. P., Chen, L. M., Morgan, V. L., Mishra, A., Wilkes, D. M., et al. (2016). Realistic models of apparent dynamic changes in resting-state connectivity in somatosensory cortex. *Hum. Brain Mapp.* 37, 3897–3910. doi: 10.1002/hbm.23284
- Tagliazucchi, E., and Laufs, H. (2015). Multimodal imaging of dynamic functional connectivity. *Front. Neurol.* 6:10. doi: 10.3389/fneur.2015.00010
- Wu, L., Eichele, T., and Calhoun, V. D. (2010). Reactivity of hemodynamic responses and functional connectivity to different states of alpha synchrony: a concurrent EEG-fMRI study. *Neuroimage* 52, 1252–1260. doi: 10.1016/j.neuroimage.2010.05.053
- Zalesky, A., and Breakspear, M. (2015). Towards a statistical test for functional connectivity dynamics. *Neuroimage* 114, 466–470. doi: 10.1016/j.neuroimage.2015.03.047
- Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., and Breakspear, M. (2014). Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10341–10346. doi: 10.1073/pnas.1400181111

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Miller, Abrol, Adali, Levin-Schwarz and Calhoun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.