

On the statistical performance of Granger-causal connectivity estimators

Koichi Sameshima · Daniel Y. Takahashi · Luiz A. Baccalá

Received: 23 January 2015 / Accepted: 23 March 2015 / Published online: 22 April 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In this article, we extend the statistical detection performance evaluation of linear connectivity from Sameshima et al. (in: Slezak et al. (eds.) *Lecture Notes in Computer Science*, 2014) via brand new Monte Carlo simulations of three widely used toy models under different data record lengths for a classic time domain multivariate Granger causality test, information partial directed coherence, information directed transfer function, and include conditional multivariate Granger causality whose behaviour was found to be *anomalous*.

Keywords Partial directed coherence · Directed transfer function · Granger causality · Null hypothesis test performance · Conditional multivariate Granger causality

1 Introduction

This paper compares the statistical performance of linear connectivity detection [1] using four popular neural connectivity estimators. In addition to the classic Granger causality test (GCT) from [2], we employ our recently derived rigorous results [3, 4] about the asymptotic

behaviour of information PDC (*i*PDC) and information DTF (*i*DTF) [5] that, respectively, generalize partial directed coherence (PDC) [6] and directed transfer function (DTF) [7] which correctly describe coupling effect size issues. A fourth method was included in this extended version of [8] and consists of the proposal put forward by [9] (*c*MVGC) for detecting conditional Granger causality between time series pairs and is applied here using their published MVGC package. There have been many recent papers [10–14] aimed at comparing contending connectivity estimation procedures. In fact, almost every new connectivity estimation procedure sports some form of appraisal by counting the number of correct detection decisions. What sets the present effort apart, as emphasized by using the word *statistical*, is that we focus on methods that have rigorous theoretically derived asymptotic detection criteria.

In the comparisons, we used Monte Carlo simulations of three widely used toy models from the literature and verified the performance of null connectivity hypothesis rejection as a function of data record length, K . To complement the study, we also computed false positive (FP) and false negative (FN) test rates for each estimator alternative. In the MVGC package case, false detection rates were computed with and without author-recommended corrections [9].

2 Methods and results

2.1 Monte Carlo simulations

Following our recently proposed information PDC and information DTF [5], and their corresponding rigorous asymptotic statistics (see [3] and [4] for details), we first

K. Sameshima (✉)
Radiology & Oncology Department, Faculdade de Medicina,
University of São Paulo, São Paulo, SP 01246-903, Brazil
e-mail: ksameshi@usp.br

D. Y. Takahashi
Psychology Department, Neuroscience Institute, Princeton
University, Princeton, NJ, USA

L. A. Baccalá
Department of Telecommunications and Control, Escola
Politécnica, University of São Paulo, São Paulo, SP, Brazil

gauged their statistical performance against that of the well-established time-domain GCT test [2]. For added comparison here, we added results from conditional connectivity detection obtained via the MVGC package [9]. Monte Carlo simulations were performed in the MATLAB environment using its normally distributed pseudorandom number generator to simulate systems with uncorrelated zero mean and unit variance innovation noise as model

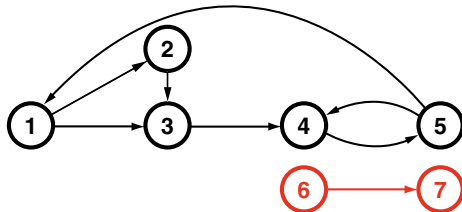


Fig. 1 Diagram depicting the essential elements of Model 1 represented by Eq. 1 from [16]. The elements from x_1 to x_5 establish closed-loop connections, with short and long connected paths, while x_6 and x_7 are part of a completely separate substructure, i.e. disconnected from $\{x_1, \dots, x_5\}$, but sharing a common frequency of oscillation

inputs. To test the performance of the latter four connectivity estimators, for each toy model and at each data record length, we selected values of $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$ repeating 1000 simulations for each case. For each simulation, a burn-in set of 5000 initial data points were discarded to eliminate possible transients before selecting the K value of interest. We used the Nuttall-Strand algorithm for multivariate autoregressive (MAR) model estimation and the Akaike information criterion (AIC) for model order selection [15] for GCT, i PDC and i DTF, while the Levinson–Wiggins–Robinson solution of the multivariate Yule-Walker equations was used as default for c MVGC [9]. Detection threshold was set in compliance to $\alpha = 1\%$. For i PDC and i DTF, p values were computed at 32 uniformly separated normalized frequency points covering the whole interval with a connection being deemed detected for a given pair of structures if its p value resulted to be less than α for some frequency within the interval. This connectivity decision criterion is somewhat lax and tends to overestimate the presence of connectivity for i PDC and i DTF. In particular for i PDC, one should

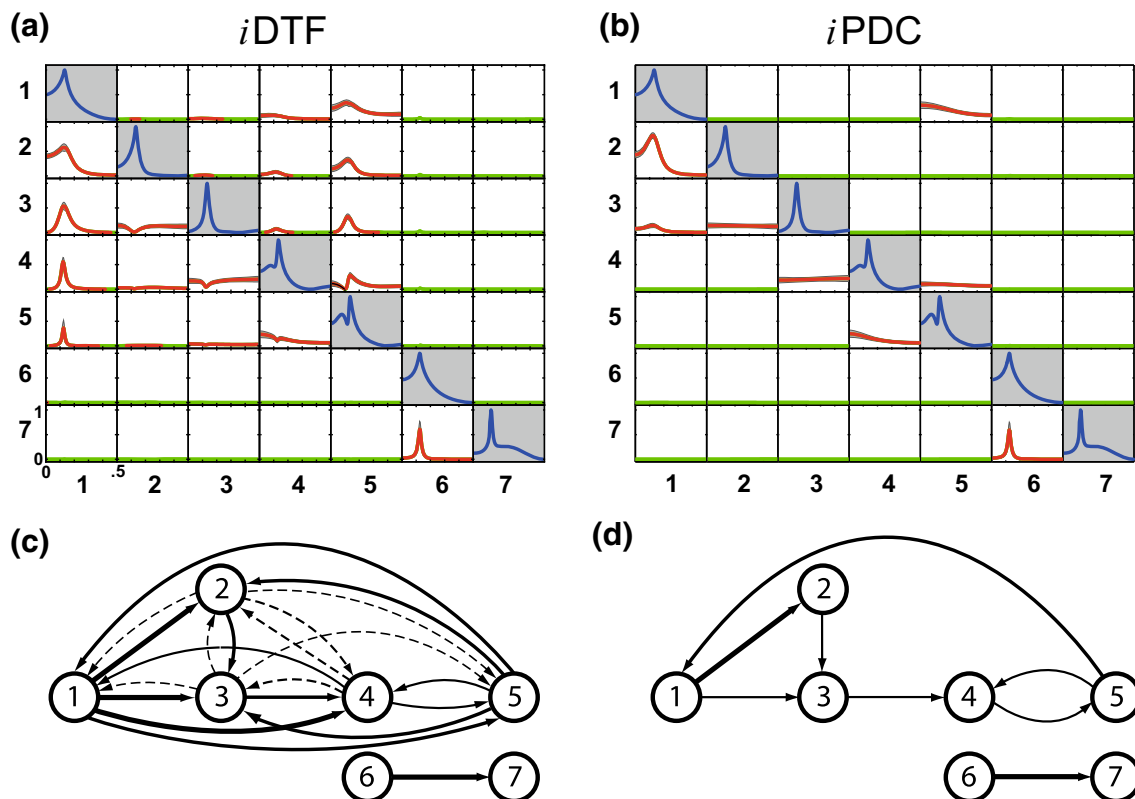


Fig. 2 A single-trial results of **a** i DTF and **b** i PDC estimations obtained using a data simulation of Model 1 with $K = 2000$ points. In both subfigures, **a**, **b**, the main diagonal subplots with *gray background* contain power spectra, while each off-diagonal subplot represents i DTF or i PDC measure in the frequency domain with variables in *columns* representing the sources and in *rows* the target structures, in which significant measure is drawn in *red lines* at

$\alpha = 1\%$, and in *green lines* otherwise. **c** Note that, as theoretically expected, according to i DTF estimation, all nodes of $\{x_1, x_2, x_3, x_4, x_5\}$ set can reach one another, **d** while i PDC correctly exposes, similar to GCT, the immediate adjacent node's connectivity pattern. See further discussion about the contrast between i DTF and i PDC in [17]. (Color figure online)

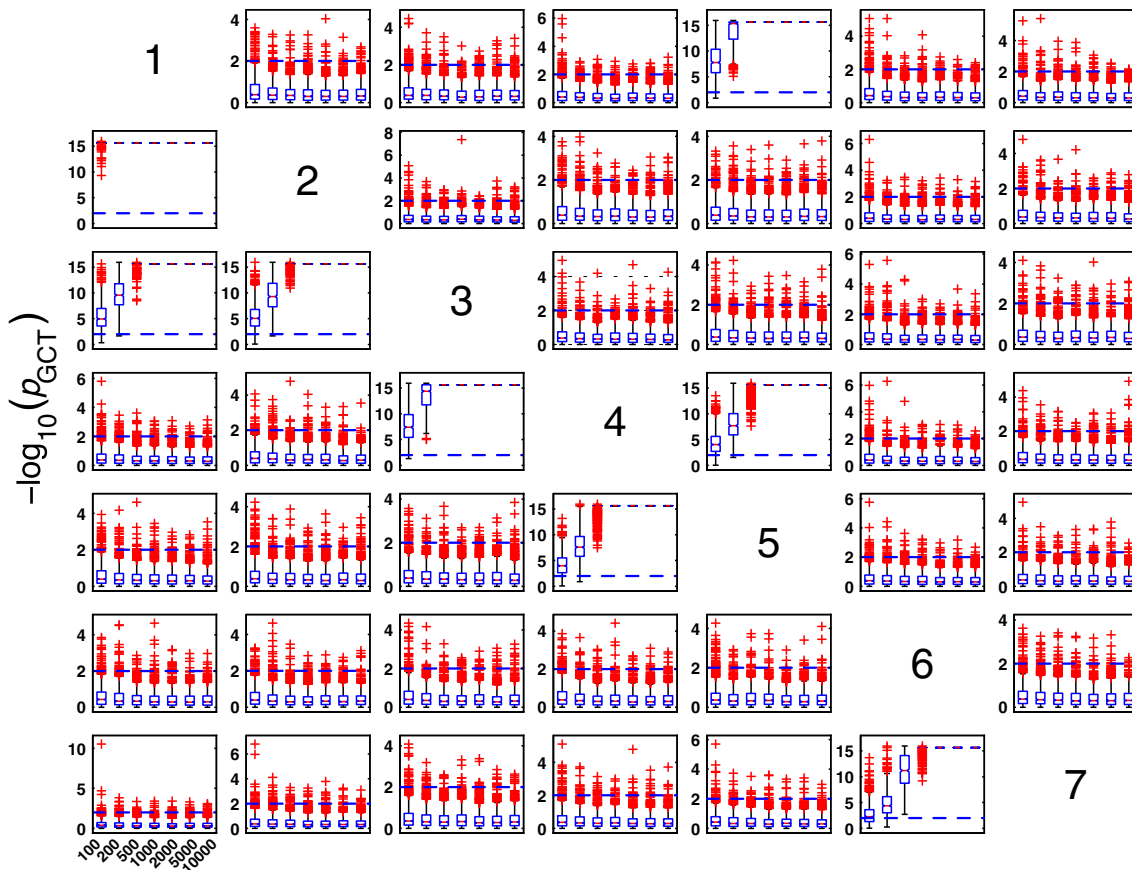


Fig. 3 The patterns (in this and all the figures of similar kind that follow) containing subplots with variables in columns representing the sources and the target structures in rows. Each subplot possesses boxplots of the distribution of GCT $-\log_{10}(p \text{ value})$ for 1000 Monte Carlo simulations over different record lengths $K = \{100, 200, 500,$

$1000, 2000, 5000, 10000\}$ along the x -axis of each subplot. Since $\alpha = 0.01$, values above 2 (*dashed-line*) indicate rejection of the null hypothesis of connectivity absence. *Red crosses* indicate p value distribution outliers, and those above *dashed-line* represent false positives (FPs) for nonexistent connections. (Color figure online)

expect connectivity detection more often than GCT, i.e. more FPs are likely.

The reader may access our open MATLAB codes for GCT and for both $iPDC$ and $iDTF$ asymptotic statistics used in this study at <http://www.lcs.poli.usp.br/~baccala/BIHExtension2014/>.

The Web site, furthermore, contains the datasets of the employed simulation results and a copy of the exact version of the MVGC package used in the present comparisons [9]. This allows full reader accessing disclosure of the data/procedures with the possibility of cross-checking and replaying all results. Additional graphs and results are available there and may be consulted for details; only the overall representative behaviour is summed up here.

Next we describe the toy models and the allied simulations results.

2.2 Model 1: Closed-loop model

It is an $\{N = 7\}$ -variable model, borrowed from [16] (Fig. 1), with two completely disconnected substructures,

$\{x_1, x_2, x_3, x_4, x_5\}$ and $\{x_6, x_7\}$, which share a common frequency of oscillation. The set of descriptive equations is

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) \\ \quad + 0.5x_5(t-2) + w_1(t) \\ x_2(t) = -0.5x_1(t-1) + w_2(t) \\ x_3(t) = 0.2x_1(t-1) + 0.4x_2(t-2) + w_3(t) \\ x_4(t) = -0.5x_3(t-1) + 0.25\sqrt{2}x_4(t-1) \\ \quad + 0.25\sqrt{2}x_5(t-1) + w_4(t) \\ x_5(t) = -0.25\sqrt{2}x_4(t-1) + 0.25\sqrt{2}x_5(t-1) + w_5(t) \\ x_6(t) = 0.95\sqrt{2}x_6(t-1) - 0.9025x_6(t-2) + w_6(t) \\ x_7(t) = -0.1x_6(t-2) + w_7(t), \end{cases} \tag{1}$$

where w_i stand for uncorrelated $\mathcal{N}(0, 1)$ Gaussian innovations.

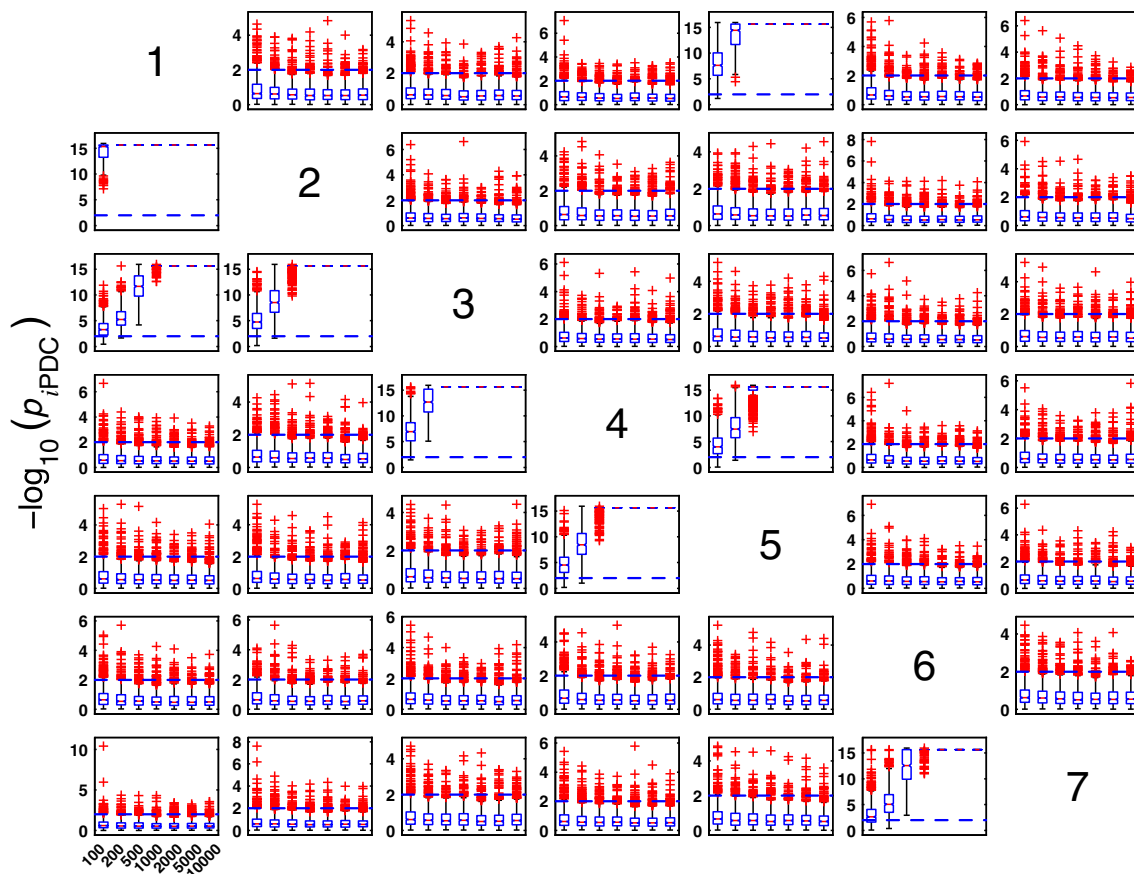


Fig. 4 Model 1 boxplot performance summary of *iPDC* asymptotics. Most outliers for absent connections are above the α threshold decision line

Results from a single-trial example of *iDTF* and *iPDC* connectivity estimations in the frequency domain are depicted in Fig. 2a, b, respectively, with significant values, at $\alpha = 0.01$, represented by red solid lines. The corresponding connectivity graph diagrams are contained in Fig. 2c, d, where arrow thickness represents estimate magnitude. Note that *iPDC* reflects adjacent connections, Fig. 2b, d, while *iDTF*, Fig. 2a, c, represents graph reachability aspects of the directed structure [17, 18]. The notion of reachability refers to the net influence from a time series onto another through various signal pathways, i.e. it measures how much of one series ends up influencing another.

2.2.1 Granger causality test for Model 1

The boxplots of $-\log_{10}(p \text{ value})$ in Fig. 3 summarize GCT performance for Model 1 and $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$ data record lengths. As expected, for $K > 200$, it properly detects connectivity between adjacent structures with zero observed FNs for all pairs of existing connections.

2.2.2 *iPDC* performance for Model 1

Figure 4 summarizes the asymptotic *iPDC* statistical performances for the same data and record lengths as for GCT in Fig. 3 with similar performance (Figs. 5, 6). Closer comparison on identical trials for each estimator leads to Fig. 7 depicting *iPDC* versus GCT performance ($K = 2000$), further revealing a pattern of consistently higher FP values for *iPDC* expectedly resulting from how the test was performed with *iPDC* decision dictated by a single maximum frequency above threshold. In Fig. 7, the average slopes are above 45° consistent with the larger number of FPs for *iPDC*.

At this point, one should note that for trial-by-trial comparisons between methods only those against GCT are present for the sake of conciseness. Pairwise behaviour for other pairs of methods is easy to infer. GCT’s choice as a reference was dictated by its canonical behaviour in terms of the expected performance in the Neyman–Pearson hypothesis testing framework. In the Web site, it is possible to use available routines to examine the results that apply to the comparison between other pairs of methods.

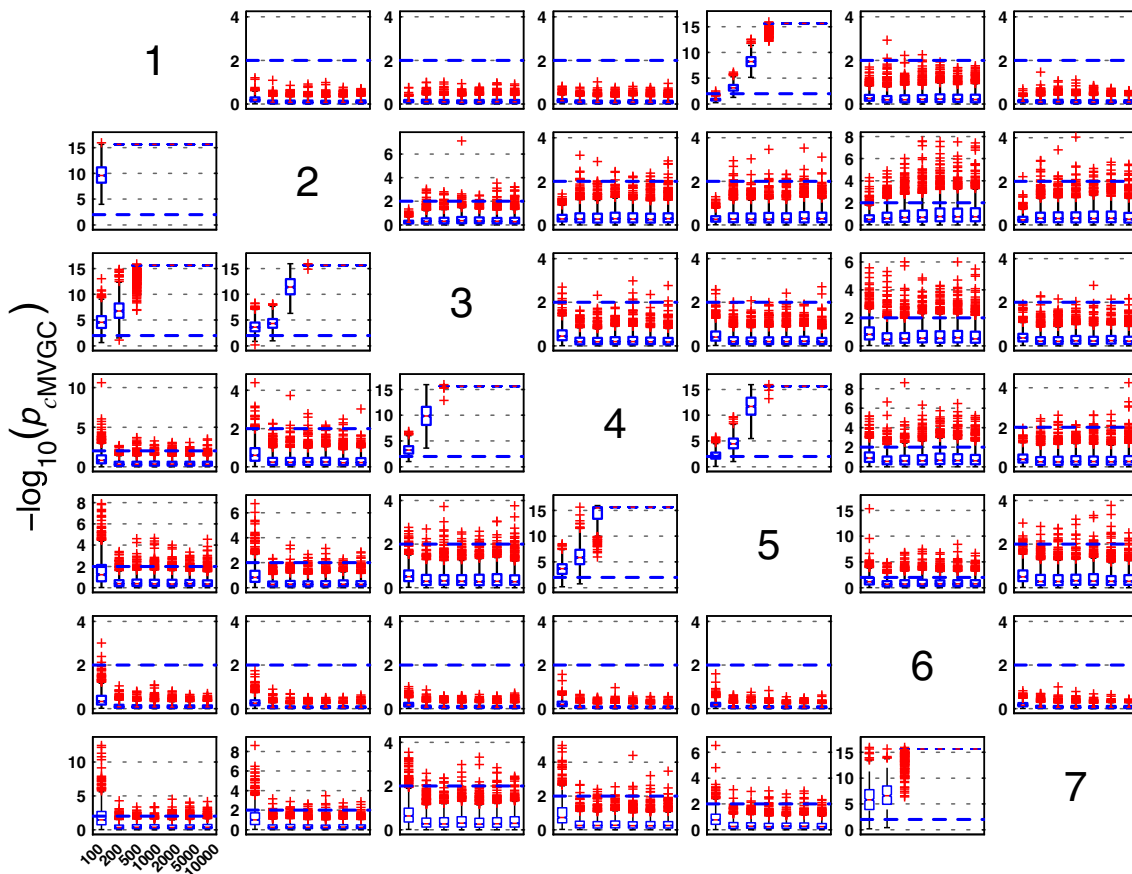


Fig. 5 Model 1 boxplot performance summary of *cMVGC* asymptotics. Note that *red cross* outliers for the connections into x_1 and x_6 (see the *first* and *sixth* rows of subplots' layouts) are consistently

below $-\log_{10}(p \text{ value}) = 2$, and those from x_6 consistently above for all K , something that is reflected in Fig. 8. (Color figure online)

2.2.3 *cMVGC* behaviour for Model 1

Figure 5 summarizes the performance of pairwise conditional MVGC in the form of boxplots. They asymptotically capture the structure of Fig. 1 despite differences compared to GCT and *iPDC*. These differences are easier to appreciate on the trial-by-trial comparison with respect to GCT (Fig. 8), which shows that *cMVGC*'s FP rates are sometimes well below the imposed $\alpha = 1\%$ and even become more extreme after authors' recommended corrections [9] ($K = 2000$). Note how point distributions in Fig. 8 hardly ever cluster round the 45° line for connections reaching the x_1 and x_6 oscillators. For connections leaving x_6 , the pattern is reversed. It is this failure to meet the preset $\alpha = 1\%$ irrespective of which connection is under consideration, which we call *anomalous* here.

2.2.4 *iDTF* performance for Model 1

Figure 6 summarizes the performance of the asymptotic statistics for *iDTF*. The boxplots clearly show that for larger sample sizes, *iDTF* correctly detects the reachability structure shown in Fig. 2c. Note that the weakest, and in

this case, the farthest connection ($x_2 \rightarrow x_1$) requires longer record lengths for proper detection.

2.3 Model 2: Five-variable model

Model 2 introduced by [6] is graphically represented in Fig. 9 with its corresponding set of defining equations:

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) + w_1(t) \\ x_2(t) = 0.5x_1(t-2) + w_2(t) \\ x_3(t) = -0.4x_1(t-3) + w_3(t) \\ x_4(t) = -0.5x_1(t-2) + 0.25\sqrt{2}x_4(t-1) \\ \quad + 0.25\sqrt{2}x_5(t-1) + w_4(t) \\ x_5(t) = -0.25\sqrt{2}x_4(t-1) + 0.25\sqrt{2}x_5(t-1) + w_5(t), \end{cases} \tag{2}$$

where w_i , as before, stand for uncorrelated Gaussian innovations. Computations were performed for $K = \{100, 200, 500, 1000, 2000\}$ long records over 1000 Monte Carlo repetitions.

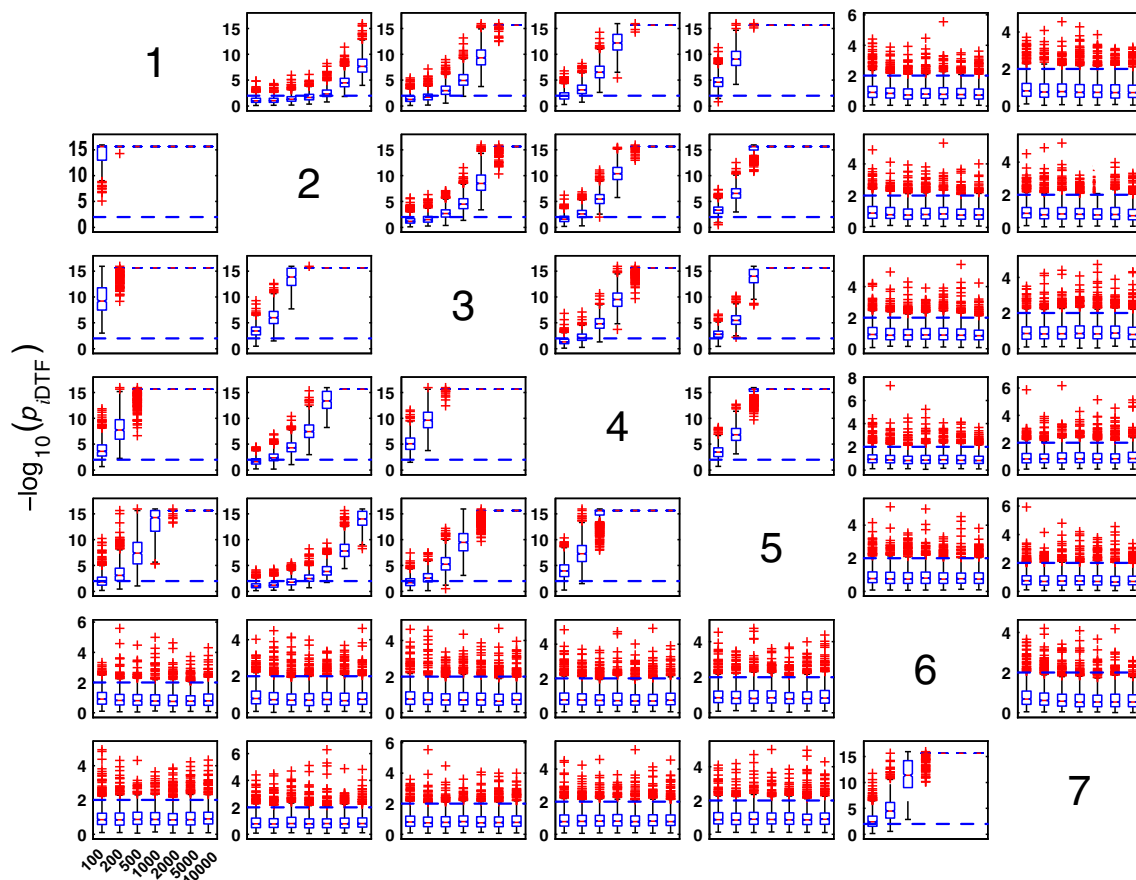


Fig. 6 Model 1 boxplot performance summary of *iDTF* asymptotics. Note that every node of $\{x_1, x_2, x_3, x_4, x_5\}$ set can directionally reach one another, as depicted in Fig. 2c. Note also that FN rates decrease consistently for larger K

2.3.1 *GCT performance*

As before, Model 2 also shows that GCT’s performance improves with the increased record length (Fig. 10). At $K = 200$, GCT already performs well with FN rate below 5 %, reaching overall FN rates below 2 % for $K = 2000$.

2.3.2 *iPDC performance*

For Model 2, FN rates are practically negligible when $K > 200$ for all measures of GCT, *iPDC*, and *cMVGC* (See Figs. 10–12). Overall, the pattern of *iPDC* performance is similar to that of GCT’s. Yet *iPDC*’s FP rates are slightly higher than GCT’s. For example, performance for $K = 2000$ is between 2.7 and 5.6 % (Fig. 13).

2.3.3 *cMVGC asymptotic behaviour for Model 2*

cMVGC performance for $K = \{100, 200, 500, 1000, 2000\}$ is shown in Fig. 12. When taken with respect to GCT (Fig. 14), FPs are consistently lower than GCT’s for $K = 2000$ and, as in the case of the previous model, it does

not conform to a preset $\alpha = 1\%$ for FP rates. This is also easy to appreciate for other values of K in Fig. 12 as most outliers (red crosses) are below the $-\log_{10}(p) = 2$ line for nonexistent connections.

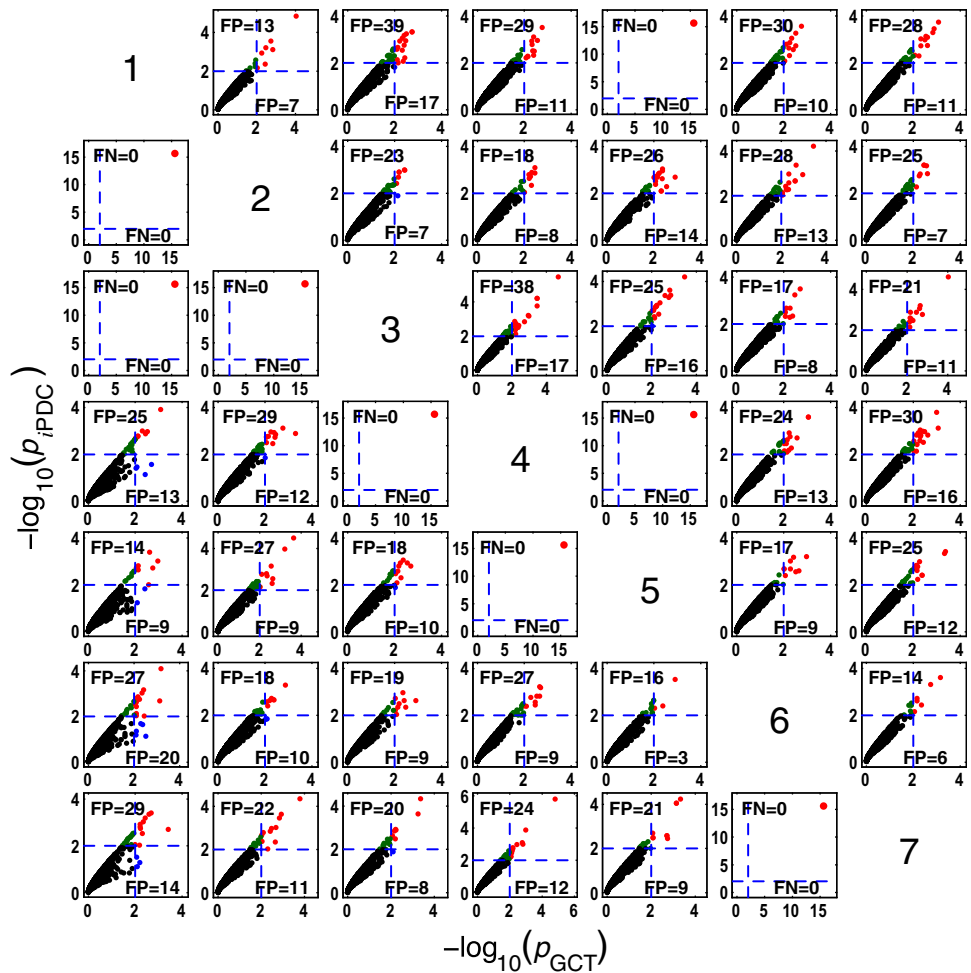
Taking GCT as a reference, trial-by-trial comparisons of *iPDC* and *cMVGC*, respectively, confirm the pattern of higher FP for the former compared to a pattern of FP, below 1 %, for *cMVGC* with or without correction (See Figs. 13, 14). This is also suggestive of possible problems encountered in how the MVGC package handles the FP rate, which may be fortuitously benign to MVGC in this example, but does not represent the general case, since it does not hold for Model 1.

2.4 Model 3: Modified five-var model

To further probe the statistical behaviours of GCT, *iPDC* and *cMVGC*, we simulated the modified five-channel toy Model 3, originally introduced in [6], under the formulation variant proposed by [19] and reproduced here for reference (Fig. 15).

The corresponding set of equations is

Fig. 7 GCT and *i*PDC connectivity comparative detection performance ($K = 2000$) where *i*PDC $-\log_{10}(p \text{ value})$ for each one of the 1000 simulations is plotted against its GCT's $-\log_{10}(p \text{ value})$. Results for all connections are clustered slightly above the 45° line with $b = 1.1157 \pm 0.0283$ (median = 1.1216; minimum = 1.0198; and maximum = 1.1486). The number of FP and FN detections over the 1000 simulations are also shown for *i*PDC (top left) and GCT (bottom right) for each subplot



$$\begin{cases}
 x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) \\
 \quad + e_1(t) + a_1e_6(t) + b_1e_7(t-1) + c_1e_7(t-2) \\
 x_2(t) = 0.5x_1(t-2) \\
 \quad + e_2(t) + a_2e_6(t) + b_2e_7(t-1) + c_2e_7(t-2) \\
 x_3(t) = -0.4x_1(t-3) \\
 \quad + e_3(t) + a_3e_6(t) + b_3e_7(t-1) + c_3e_7(t-2) \\
 x_4(t) = -0.5x_1(t-2) + 0.25\sqrt{2}x_4(t-1) + 0.25\sqrt{2}x_5(t-1) \\
 \quad + e_4(t) + a_4e_6(t) + b_4e_7(t-1) + c_4e_7(t-2) \\
 x_5(t) = -0.25\sqrt{2}x_4(t-1) + 0.25\sqrt{2}x_5(t-1) \\
 \quad + e_5(t) + a_5e_6(t) + b_5e_7(t-1) + c_5e_7(t-2),
 \end{cases} \tag{3}$$

additionally containing the large exogenous input $e_6(t)$ and the latent variable $e_7(t)$. In the simulations, $e_i(t)$ were uncorrelated zero mean unit variance Gaussian innovation noises, and the parameters were chosen as $a_i \sim U(0, 1)$, $b_i = 2$ and $c_i = 5, i = 1, \dots, 5$ according to [19].

The proposal in [19] of introducing exogenous/latent variables is an interesting idea which allows investigating the influence of large common additive noise sources on the performance of GCT, *i*PDC and *c*MVGC. Here, to assess the impairment that the extra exogenous/latent variables possibly inflict on null-hypothesis testing, we repeated the procedure not just under the same conditions of [19], but also using a broader range of data record sizes: $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$.

2.4.1 GCT performance in the presence of exogenous noise, Model 3

The GCT performance for Model 3 can be appreciated in Fig. 16. When compared with Model 2, GCT's performance deteriorates in the presence of exogenous noises. Interestingly, its performance with respect to detecting existing connections increases with longer data records, while in the absence of connections, the FP rates increase sharply especially for the $K = 10000$ case. For $K = 500$,

Fig. 8 *c*MVGC versus GCT performance ($K = 2000$) clusters along lines, with high $b = 0.8329 \pm 0.5741$ coefficient spread (median = 0.8417; minimum = 0.1762; and maximum = 2.3323), confirming *c*MVGC’s abnormal behaviour. Connections out of x_6 are consistently clustered above the 45° line, which contrasts with those reaching the x_1 and x_6 consistently with low *c*MVGC FP values (*top left*) (except for $x_5 \rightarrow x_1$). *c*MVGC abnormality is apparent as FP values differ much from 10 as should happen for $\alpha = 1\%$. *c*MVGC-corrected FPs (separated by *slash*) do not improve the situation

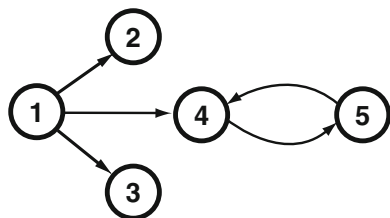
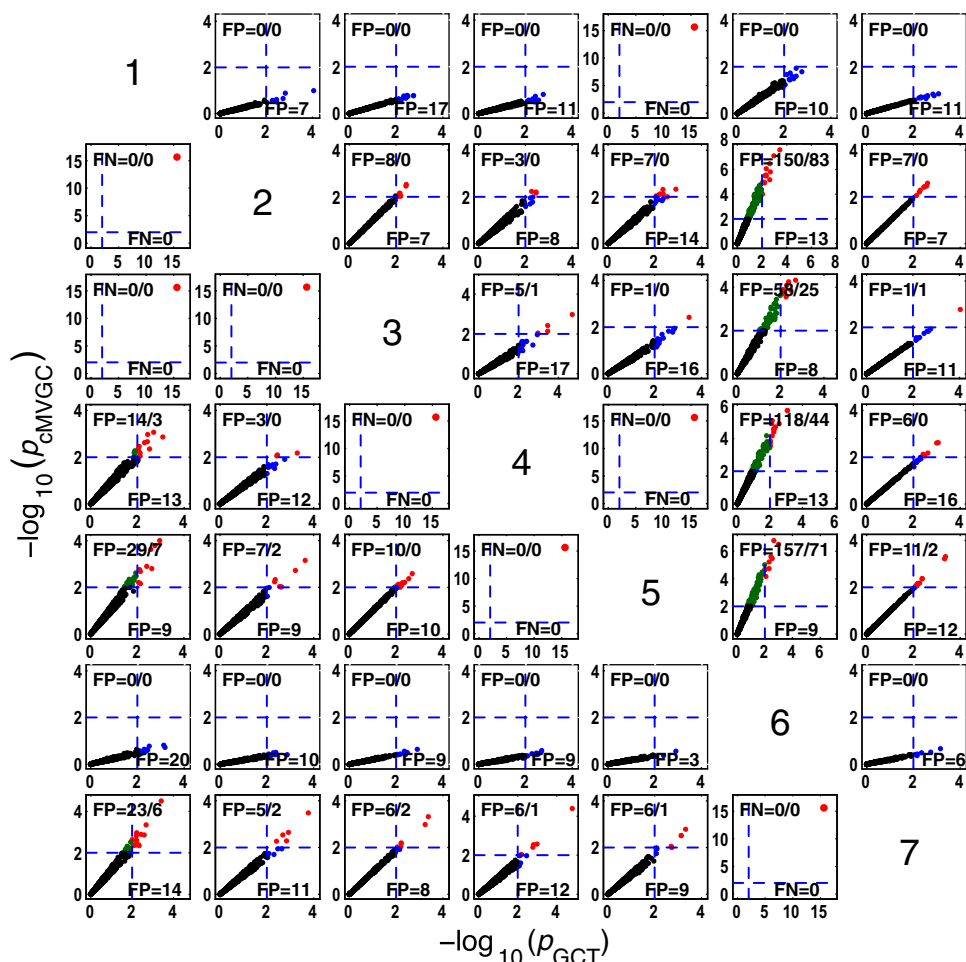


Fig. 9 Diagram depicting the essential elements of Model 2 introduced by [6]

the overall FP rates are between 2.3 and 7.9 % with a median of 3.8 %. At $K = 10000$, the latter rates grow to a range between 20.8 and 40.4 % with the median value of 26.6 %. FN rates are negligible.

2.4.2 *i*PCD performance in the presence of exogenous noise

*i*PCD performance in detecting connectivity is similar to GCT’s (See Figs. 16, 17). As noted before, *i*PCD tends to

have higher FP rates compared with GCT due possibly to the chosen frequency domain detection criterion of using a single-frequency with significant p value as indicative of a valid connection. Overall, FP rates range between 6.7 and 11.7 % (median 8.5 %) at $K = 100$ increasing to the range (30.8, 49.6 %) range (median 40.1 %) at $K = 10000$.

2.4.3 *c*MVGC performance for Model 3

Here (Fig. 18) the qualitative behaviour is the same as for the other estimators. However, as for Model 1, false decision rates are out of control,—sometimes, much below GCT’s, and sometimes, way above it, irrespective of corrections which fail to restore Neyman–Pearson expected behaviour. Again taking GCT as reference, Fig. 19 shows the similarity of *i*PCD’s result to GCT’s with the same pattern of larger FP values well above $\alpha=1\%$. The corresponding results for *c*MVGC compared with GCT portray a bias towards lower *c*MVGC FPs (Fig. 20).

Fig. 10 GCT performance for Model 2 and $K = \{100, 200, 500, 1000, 2000\}$ data record lengths

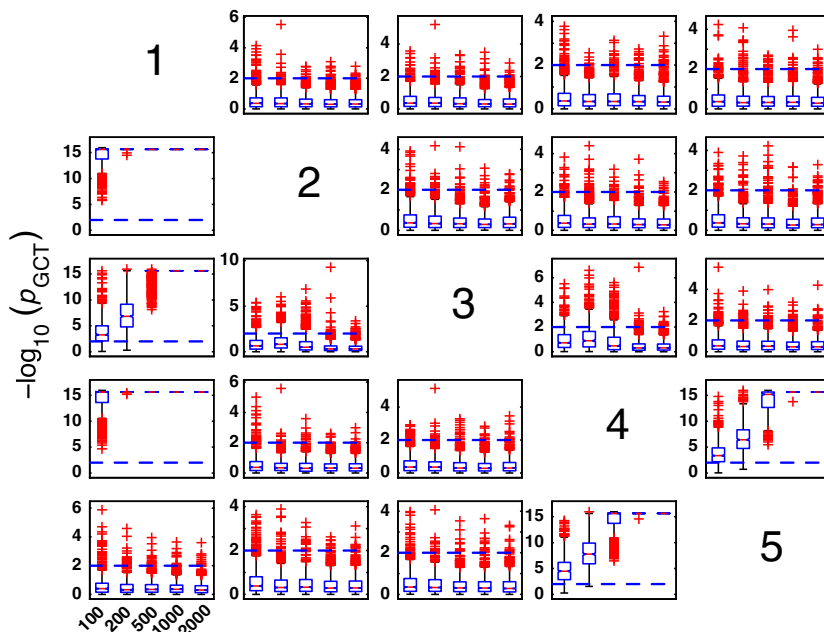
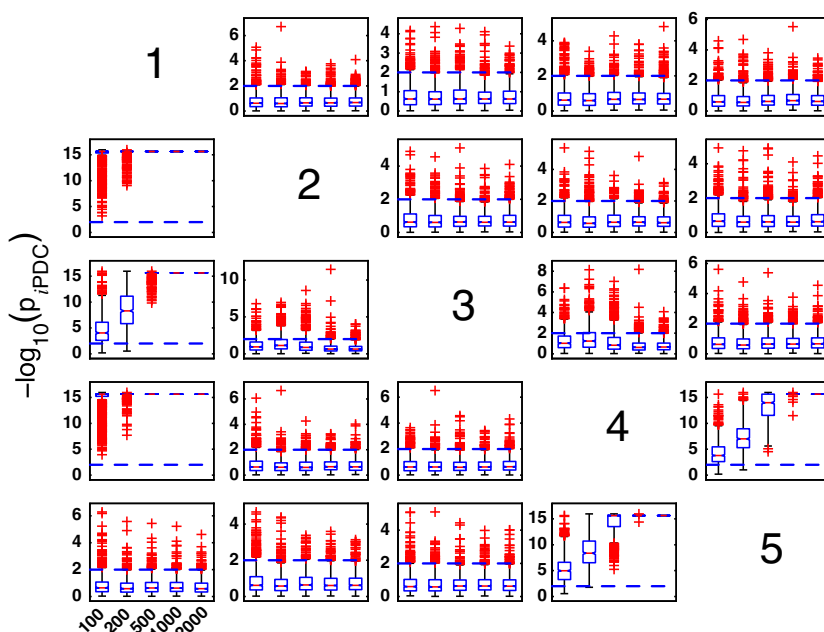


Fig. 11 *i*PDG performance for Model 2 and $K = \{100, 200, 500, 1000, 2000\}$ data record lengths



3 Discussion

This study presents simulation evidence about the performances of statistical connectivity tests: two in time domain and using two new frequency domain measures.

One should remind the reader that the frequency domain tests, *i*DTF and *i*PDG, measure different aspects of connectivity and are not immediately comparable as discussed at length in [17, 18]. This contrasts with GCT, *i*PDG and *c*MVGC which are geared towards describing the same aspect of connectivity between adjacent structures [17].

Among the tests in the latter class, GCT proved to be the one most in accord with the expected Neyman–Pearson behaviour in the sense that observed FP rates are in accord with the preset value of α justifying its employment as reference in the trial-by-trial comparisons between methods as summed up herein.

Qualitatively *i*PDG closely mirrors GCT behaviour, and predictably produces higher FP rates as a consequence of how *i*PDG connectivity was detected by deeming just one frequency above threshold as significant. Whereas one may conceivably improve on how to employ *i*PDG for testing,

Fig. 12 *c*MVGC performance for Model 2 and $K = \{100, 200, 500, 1000, 2000\}$ data record lengths

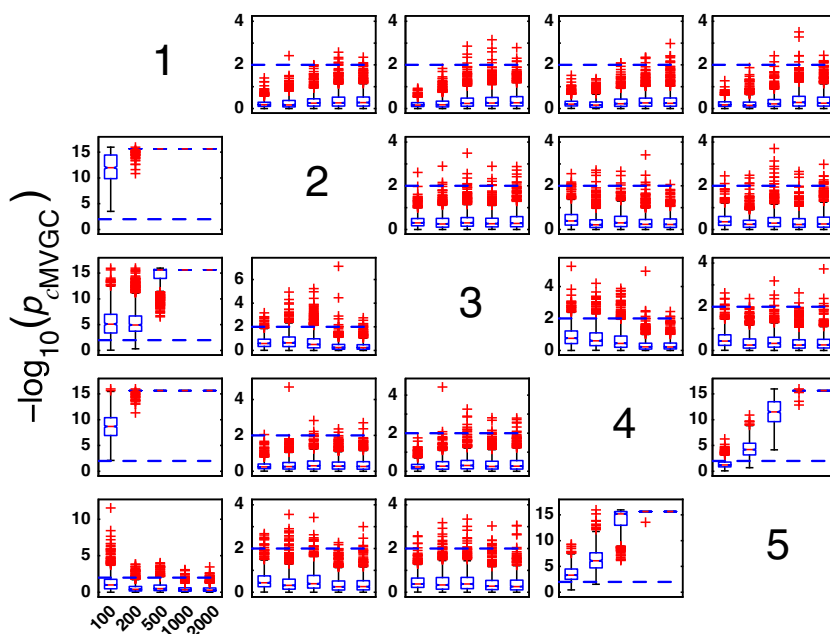
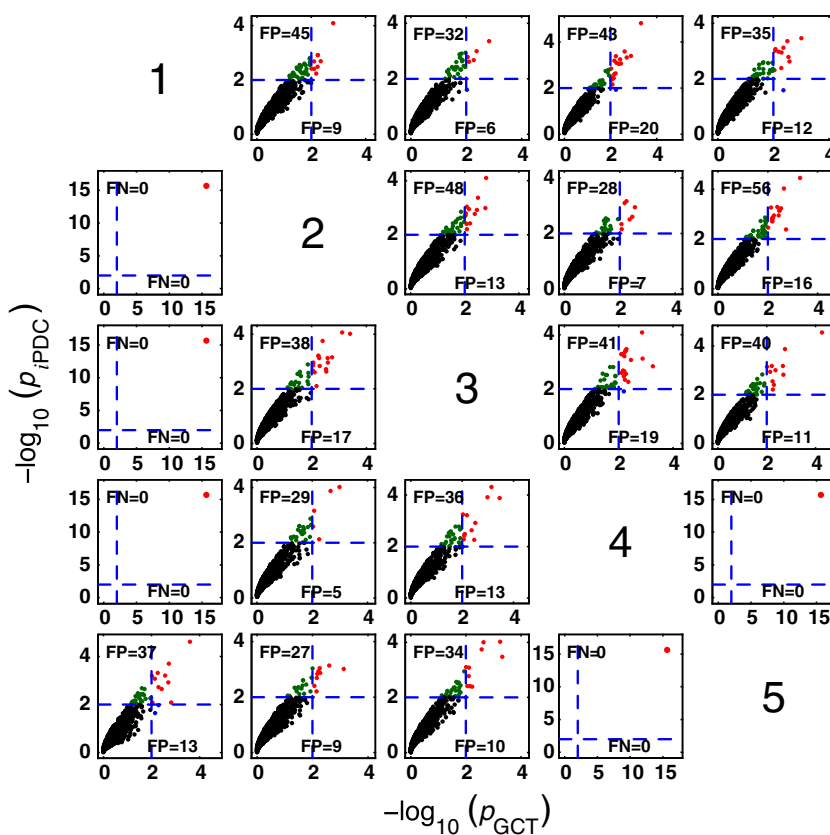


Fig. 13 Comparative performance between GCT and *i*PDC detection performances at $K = 2000$ time samples for Model 2



its use is recommended when there is frequency content of physiological interest.

Added for comparison, *c*MVGC detection proved to be biased towards a reduction of the FP rates in many cases. By contrast, examination of its behaviour for other K (available in more detail from our Web site) suggests that,

for small K , it tends to miss existing connections more often than the other methods.

Perhaps more striking and more important, however, in the sense of Neyman–Pearson detection for α compliance, is that procedures are usually constructed to impart control over FP decisions, which according to the present

Fig. 14 Comparative performance between GCT and cMVGC detection performances at $K = 2000$ time samples for Model 2

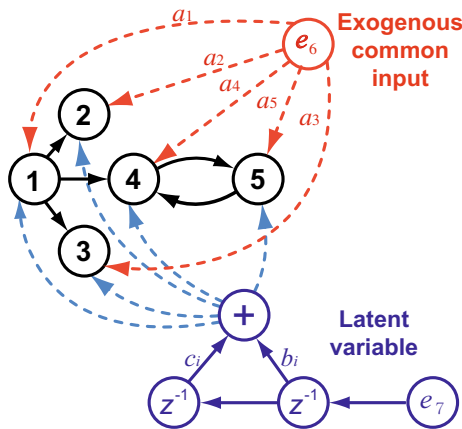
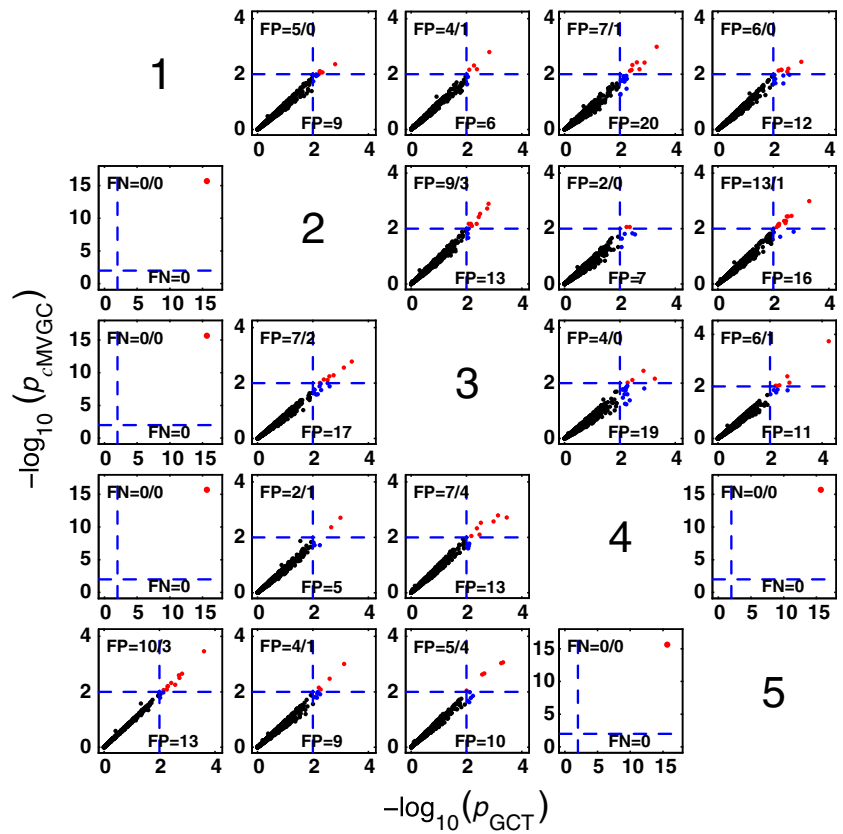


Fig. 15 Diagram depicting the essential elements of model introduced by [19] modified from Model 2 [6]. For each simulation, the parameters a_i were chosen randomly from a uniform distribution in $[0, 1]$ interval, and all $b_i = 2$ and $c_i = 5$, while the innovations, e_i , were drawn from random variables with $\mathcal{N}\{0, 1\}$

observations, is a condition that fails to be met by the cMVGC implementation from [9] which was used here without modification. It is also important to note that employing author-recommended decision corrections [9] usually aggravates matters. It is this lack of compliance to Neyman–Pearson criteria that we termed *anomalous*.

Whether this happens due to an eventual software glitch, or reflects a more fundamental issue, is unknown. One should note that on many instances, cMVGC produced fewer FPs, something good in itself. This apparent quality is counterbalanced by much worse performance for some links, as in Model 1, in sharp contrast to other methods whose results attain the prescribed α and are balanced for all connections to within the attainable accuracies of the Monte Carlo simulations.

Based on its good asymptotic control of FP observations, it is fair to suggest that, at least provisionally, GCT, as proposed by [2], be taken as a gold standard for detecting connectivity between adjacent structures and that iPDC and cMVGC should be used taking into account adequate forewarning of their present observed limitations.

The present Monte Carlo simulations showed good large sample fit and robustness for Models 1 and 2. In the presence of large exogenous/latent variables (Model 3), we observed poor performance for large samples possibly due to the poor performance of the MAR model estimation algorithms under low signal-to-noise ratio regardless of the statistical procedure ($K > 5000$). In this regard, Model 3 deserves the special comment that its comparatively worse performance is not surprising since, strictly speaking, it violates the usual assumptions behind the development of all the test detection procedures discussed herein.

Fig. 16 GCT performance on Model 3 and $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$ data record lengths

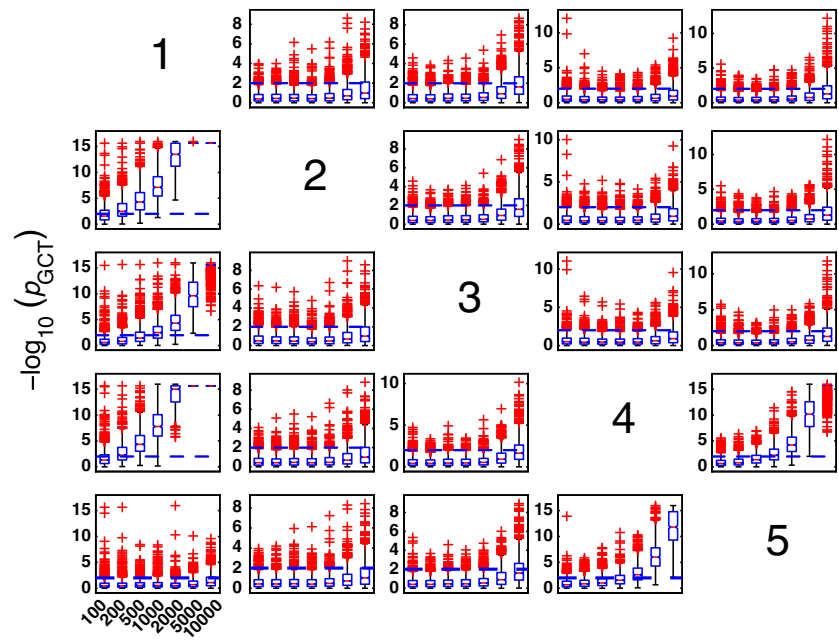
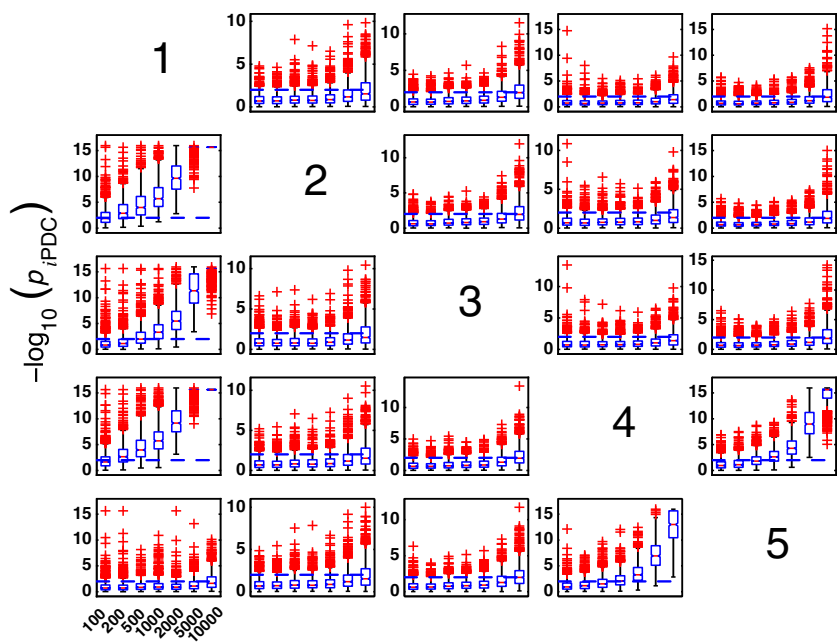


Fig. 17 *i*PDC performance on Model 3 and $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$ data record lengths



Finally, we propose that the present methodology represents the seed of a potential tool for systematically comparing connectivity estimators. The reason for this is twofold: (a) the framework provides a standardized approach whereby comparisons can be made systematically and (b) may be used even in the absence of formally

rigorous statistical criteria, i.e. even if only ad hoc decision rules are available and is therefore not restricted to methods with theoretically well-established detection criteria. We have future plans to include bootstrap-based connectivity detection schemes under the same standardized framework for comparison purposes.

Fig. 18 c MVGC performance on Model 3 and $K = \{100, 200, 500, 1000, 2000, 5000, 10000\}$ data record lengths

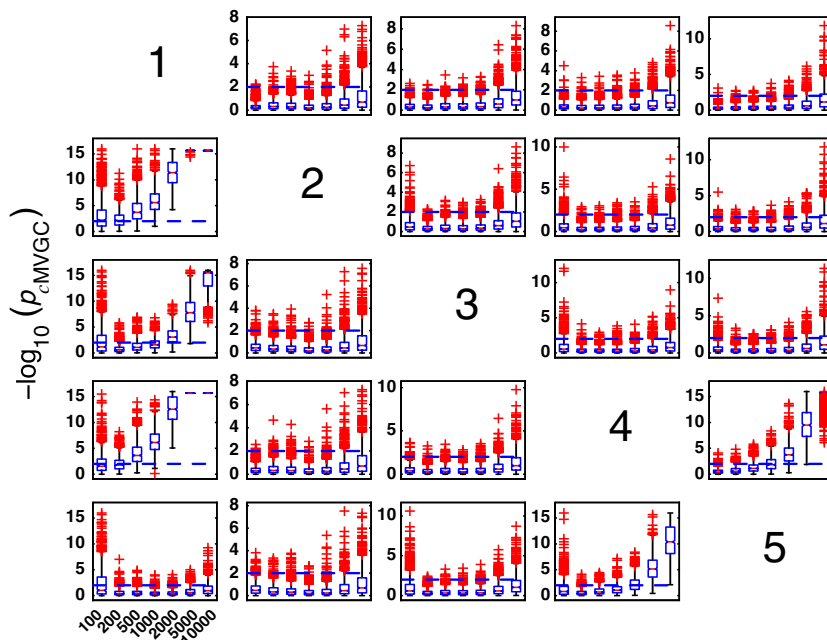


Fig. 19 Comparative performance between GCT and i PDC detection performances at $K = 2000$ time samples for Model 3

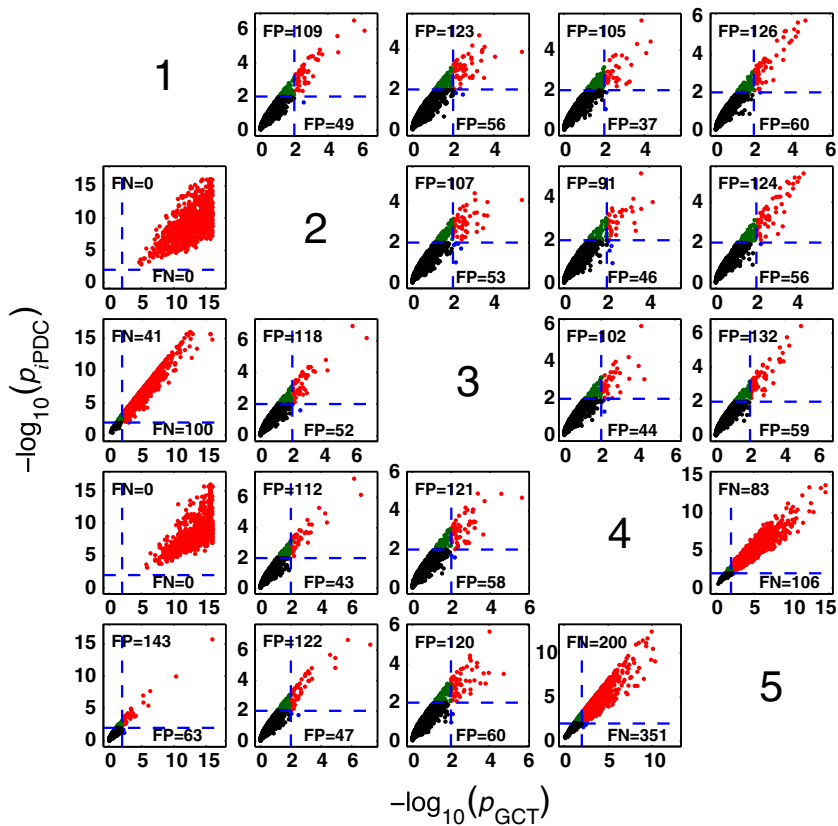
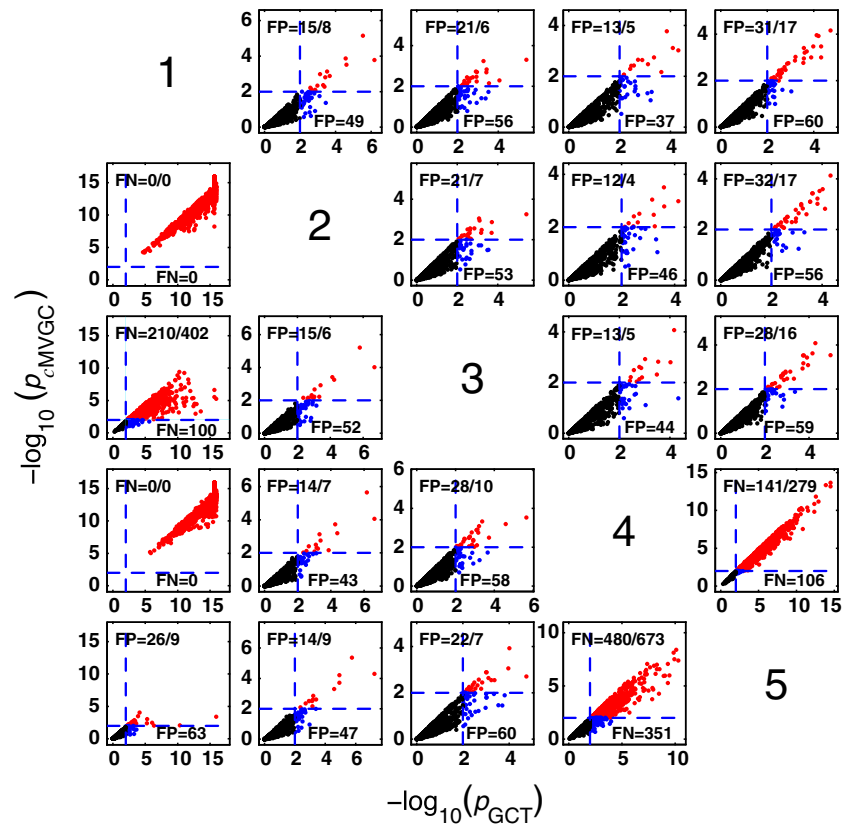


Fig. 20 Comparative performance between GCT and cMVGC detection performances at $K = 2000$ time samples for Model 3



Acknowledgments CNPq Grant 307163/2013-0 to L.A.B. and CNPq 309381/2012-6 and FAPESP 2014/12907-3 Grants to K.S. are also gratefully acknowledged, and the authors offer thanks to NAPNA - Núcleo de Neurociência Aplicada from the University of São Paulo. Part of this work was carried out during FAPESP Grant 2005/56464-9 (CInAPCe).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Baccalá LA, Sameshima K (2014) Brain connectivity. In: Sameshima K, Baccalá LA (eds) *Methods in brain connectivity inference through multivariate time series analysis*. CRC Press, Boca Raton, pp 1–9
- Lütkepohl H (2005) *New introduction to multiple time series analysis*. Springer, New York
- Baccalá L, De Brito C, Takahashi D, Sameshima K (2013) Unified asymptotic theory for all partial directed coherence forms. *Philos Trans R Soc A* 371:1–13
- Baccalá LA, Takahashi DY, Sameshima K (2015) Consolidating a link centered neural connectivity framework with directed transfer function asymptotic. arXiv: q-bio.nc/1166340
- Takahashi D, Baccalá L, Sameshima K (2010) Information theoretic interpretation of frequency domain connectivity measures. *Biol Cybern* 103:463–469
- Baccalá LA, Sameshima K (2001) Partial directed coherence: a new concept in neural structure determination. *Biol Cybern* 84:463–474
- Kamiński M, Blinowska KJ (1991) A new method of the description of the information flow in brain structures. *Biol Cybern* 65:203–210
- Sameshima K, Takahashi DY, Baccalá LA (2014) On the statistical performance of connectivity estimators in the frequency domain. In: Slezak D, Tan AH, Peters JF, Schwabe L (eds) *Lecture Notes in Computer Science*. Springer, Heidelberg, pp 412–423
- Barnett L, Seth AK (2014) The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *J Neurosci Methods* 223:50–68
- Haufe S, Nikulin VV, Müller KR, Nolte G (2013) A critical assessment of connectivity measures for EEG data: a simulation study. *NeuroImage* 64:120–133
- Wu MH, Frye RE, Zouridakis G (2011) A comparison of multivariate causality based measures of effective connectivity. *Comput Biol Med* 41:1132–1141
- Florin E, Gross J, Pfeifer J, Fink GR, Timmermann L (2011) Reliability of multivariate causality measures for neural data. *J Neurosci Methods* 198:344–358
- Fasoula A, Attal Y, Schwartz D (2013) Comparative performance evaluation of data-driven causality measures applied to brain networks. *J Neurosci Methods* 215:170–189
- Astolfi L, Cincotti F, Mattia D, Marciani MG, de Baccala LA, Vico Fallani F, Salinari S, Ursino M, Zavaglia M, Ding L, Edgar JC, Miller GA, He B, Babiloni F (2007) Comparison of different

- cortical connectivity estimators for high-resolution EEG recordings. *Hum Brain Mapp* 28:143–157
15. Marple SL Jr (1987) *Digital spectral analysis: with applications*. Prentice Hall, Englewood Cliffs
 16. Baccalá LA, Sameshima K (2001) Overcoming the limitations of correlation analysis for many simultaneously processed neural structures. *Prog Brain Res Adv Neural Popul Coding* 130:33–47
 17. Baccalá LA, Sameshima K (2014) Causality and Influentiability: The need for distinct neural connectivity concepts. In: Slezak D, Tan AH, Peters JF, Schwabe L (eds) *Lecture Notes in Computer Science*. Springer, Heidelberg, pp 424–435
 18. Baccalá LA, Sameshima K (2014) Multivariate time series brain connectivity: a sum up. In: Sameshima K, Baccalá LA (eds) *Methods in brain connectivity inference through multivariate time series analysis*. CRC Press, Boca Raton, pp 245–251
 19. Guo S, Wu J, Ding M, Feng J (2008) Uncovering interactions in the frequency domain. *PLoS Comput Biol* 4:e1000087

Koichi Sameshima is a graduate in Electrical Engineering and M.D. from the University of São Paulo. He was trained in cognitive neuroscience, brain electrophysiology, and multivariate time series analysis at the University of São Paulo and the University of California at San Francisco. His research interests centre around the studies of neural plasticity, cognitive function, and information processing aspects of mammalian brain through behavioural, electrophysiological, and computational neuroscience protocols. To functionally characterize collective multichannel neural activity and correlate to animal or human behaviour, normal and pathological, he

has also been seeking and developing robust and clinically useful methods and measures for brain dynamics staging, brain connectivity inferences, etc.

Daniel Y. Takahashi holds the B.Sc. degree in Applied and Computational Mathematics (at the University of São Paulo), an M.D. (the University of São Paulo), and a Ph.D. in Bioinformatics (at the University of São Paulo). His main interest is to better understand how animal behaviours emerge from the interaction between brain, body, environment, and other animals. To achieve such understanding, he has been working on developing new methods for statistical inference of dynamic interactions and on developing new animal models of social vocal communication.

Luiz A. Baccalá after majoring in Electrical Engineering and Physics at the University of São Paulo (1983/1984) went on to obtain his M.Sc. by studying the time series evolution of bacterial resistance to antibiotics in a nosocomial environment at the same University (1991). He has since been involved in statistical signal processing and analysis and obtained his Ph.D. from the University of Pennsylvania (1995) by proposing new statistical methods of communication channel identification and equalization. Since his return to his alma mater, he has taught courses in applied stochastic processes and advanced graduate-level statistical signal processing courses that include wavelet analysis and spectral estimation. His current research interests focus on the investigation of multivariate time series methods for neural connectivity inference and on problems of inverse source determination using arrays of sensors that include fMRI imaging and multi-electrode EEG processing.