⚙ EPJ Nonlinear Biomedical Physics
a SpringerOpen Journal

RESEARCH   **Open Access**

Brain connectivity extended and expanded

Obrad Kasum[1,2], Edin Dolicanin[1,2], Aleksandar Perovic[1,2] and Aleksandar Jovanovic[1,2*]

* Correspondence:
aljosha.jovanovich@gmail.com
[1]GIS-Group for Intelligent Systems,
Faculty of Math, University of
Belgrade, Studen trg 16, 11000
Belgrade, Serbia
[2]State University of Novi Pazar, Novi
Pazar, Serbia

Abstract

Background: The article is focused on the brain connectivity extensions and expansions, with the introductory elements in this section.

Method: In Causality measures and brain connectivity models, the necessary, basic properties demanded in the problem are summerized, which is followed by short introduction to Granger causality, Geweke developments, PDC, DTF measures, and short reflections on computation and comparison of measures.

Results: Analyzing model semantic stability, certain criteria are mandatory, formulated in preservation/coherence properties. In the sequel, a shorter addition to earlier critical presentation of brain connectivity measures, together with their computation and comparison is given, with special attention to Partial Directed Coherence, PDC and Directed Transfer Function, DTF, complementing earlier exposed errors in the treatment of these highly renowned authors and promoters of these broadly applied connectivity measures. Somewhat more general complementary methods are introduced in brain connectivity modeling in order to reach faithful and more realistic models of brain connectivity; this approach is applicable to the extraction of common information in multiple signals, when those are masked by, or embedded in noise and are elusive for the connectivity measures in current use; the methods applied are: Partial Linear Dependence and the method of recognition of (small) features in images contaminated with noise. Results are well illustrated with earlier published experiments of renowned authors, together with experimental material illustrating method extension and expansion in time.

Conclusion: Critical findings, mainly addressing the connectivity model stability, together with the positive effects of method extension with weak connectivity are summarized.

Keywords: Granger causality; Brain connectivity measure; PDC; DTF; Measure comparison; Weak brain connectivity; Partial (Linear) Dependence; Small object recognition

Background

Granger's method (some extended application in [1-9]), [10-12], has been in the focus of extensive research in neuroscience, expanded in various developments. We list some of the standardized connectivity measure terminology [13-20], mentioning Granger – Geweke counterpart measure couples, contrary to Bacala - Sameshima [18] concept of causality measure counterpart. These renowned leading authors refer to "proper frequency domain counterparts to Granger causality". Our correction is based on Geweke fundamental relation between temporal and frequency domain causality measures. We shortly focus our attention to the earlier analyzed measure comparisons, adding important argumentation.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 2 of 35

We introduce methods of Partial (Linear) Dependence – PLD and (image) small object recognition in order to deal with the weak brain connectivity- connectivity elusive or undetectable by the connectivity measures or heavily masked by noise, hopefully extending standard methods. These methods are applicable to both, frequency distributions, spectral like objects, and to frequency-time distributions, e.g. spectrograms, which expand the time point to dynamics-in-time view. This article is partly extension of our work [4], from which we reproduce fragments necessary for developments and discussion here.

In some circumstances brain processes might exhibit behavior similar to stochastic systems or fluid dynamics. No matter how much such analogies and similarities might be fruitful, we better keep some reserve for subjecting the brain to either statisticians or plumbers. We should not forget that brain is a highly complex information processing system, with reach information flow between large number of co-processing points, which is our basic initial hypothesis, better: *axiom-1.* Then, obviously, the dynamics of connectivity patterns has essential role, which includes connectivity patterns and their time switching as well.

With broader application domains which include neurology research and practice, expanding the sophistication of involved models which already operate with connectivity arguments in the most sensitive segments, strongly influencing expert's decision making, the demand for careful critical reinvestigation of theory and application has become continuously necessary. Establishing of neurological disorders, psychological evaluations, highly confident polygraphy are all of crucial significance for the subjects involved. Finally, we witnessed on a recent conference, an expert's elaboration of evaluations concerning level of patient brain damage after a stroke, consciously lost according to the contemporary criteria, with bad prognosis and consequential termination reasoning and planning. We *know that we do not know* the circumstances so well in order to produce categorical conclusions in such matters.

First, we observe, that in the contemporary connectivity modeling some procedures, computations and estimates need increased care in order to lead towards correct conclusions. It is shown that the neighborhood of zero is of accented importance in such evaluations and that unification of values with the difference below zero thresholds is necessary as the first step in computation and comparison evaluation. Harmonization of thresholds corresponding to measures involved in comparisons is an open issue requiring mathematically reasonable solutions. Computational stability is a general demand everywhere. Varying fundamental parameters in small neighborhoods of elsewhere published and established values, we perform detailed analysis of semantic stability of the deduced published exemplary connectivity models. When we face computational instability, if we deal with models of the real world, it immediately generates semantic instability and often a singularity. In this context, if connectivity graphs essentially change when computational differences of arguments occur within computational zero, then this is immediately reflected semantically as proportionally unstable maps of brain connectivity structures. This is not acceptable in any interpretation of experimental data and questions applicability of the brain connectivity models. We extract some characteristic examples involving wrong logic and those based on reasoning with insufficient care and precision.

Analysis of the operators involved - used in the measure computation and comparison by renowned authors, proved that the spectral maximum selected as the representative

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 3 of 35

invariant for both DTF and PDC measures before their comparison is not justified and might lead to the not well founded or invalid conclusions. Suggestions for the improvements of used simplification operators in computation of measures and for complementary comparisons are given. Second, there is an attitude with respect to connectivity present at large, where the thicker connectivity arrows are proportionally more important than the thinner ones, with discrimination made following corresponding signal intensity-energy level. If we remember the above stated axiom, we have to become more sensitive towards connectivity concept in general and rephrase importance criterion, or rather erase it completely. In the information processes, a weaker energy process can be much more important than those at higher energies. Also, short or ultra-short messages might precede hierarchically the longer transports. Consequently, when building a brain connectivity model, we have to take in consideration all discernible data related to connectivity.

Even more, we should accept an extra hypothesis, *axiom-2*: there are processes related to connectivity, which are indiscernible or hardly discernible from noise, with unknown importance for the brain functioning models established by connectivity measures in the current use.

Connectivity as currently performed and understood is omitting the essential temporal dimension. The conclusive connectivity graphs, the aims in connectivity estimations, are to be replaced or rather expanded in time dimension, since brain dynamic changes can be essential at the millisecond scale. This is for us the *axiom-3*, which has to be respected. Even for short events, these graphs change or massively change in time and it is necessary to integrate their time dynamics into the model that should make sense, strongly analogous to the relation between individual spectrum and time-spectrum, spectrogram. The former makes sense only when there is no intrinsic dynamics, thus, in highly stationary processes only. This does not apply to e.g. music: one cannot be aware of a melody, nor detect it applying single spectrum.

Finally, let us note that connectivity graphs present in recent research reports are usually structures with highly limited number of nodes, essentially a lot under currently achieved resolution of signal acquisition sensors (EEG, MEG, mixes). With earlier listed simplifications we are offered highly reduced graphs as brain connectivity models, which leads to oversimplified understanding of brain functioning. Certainly, in very short time, what we depict with 6 node or restricted 20 node graph today, with stronger currents as essential, thus with up to 20 directed graph links, when expanded with real weaker connectivity within currently achieved resolution, with added connectivity time dynamics in the range of millisecond resolution, will be represented with hundreds of nodes and higher number of more realistic links, probably easily reaching 2^{16} or 2^{20} links. Such combinatorial explosion is realistic. The forthcoming large size models would have to be generated, inspected, analyzed, compared, classified and monitored automatically, offering synthesis in higher abstraction synthetic invariants to experts. Who is modelling Internet functional or dysfunctional connectivity with up to 20 links? Brain is much more complex than the Internet. We are approaching the end of the connectivity modelling golden age, end of simple functionality explanations, we better be prepared for the change and work on it.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 4 of 35

Method

We present connectivity measure terminology (expanded in the Appendix), then our enhancements applicable in weak brain connectivity, where the standard connectivity measures remain undecidable, followed by examples with published measure comparisons with extended scrutiny and examples of applications of methods focused on weak brain connectivity, ending with conclusions integrated into a Proposition. Signals and software used in experiments shown here are available at our web: http://www.gisss.-math.rs/.

Causality measures, brain connectivity models

Initial points

After detailed inspection of arguments involved in analysis and comparison of certain mostly used connectivity measures in the current literature, we propose inclusion of the following points when building connectivity models:

I1. *connectivity estimation separated from other properties of interest, e.g. "connectivity strength";*

I2. *beside directed connectivity, separate treatment of bare connectivity – with no direction indication, when direction is more difficult to determine, and as a correctness test in the graphs deduced;*

I3. *more precise calculations and aggregations;*

I4. *scrutiny of involved operators;*

I5. *appropriate changes in the calculation and comparison procedures resulting in the more precise modeling of the connectivity structures and properties;*

I6. *special attention to the thresholds involved and related numeric zero which is basic for all other conclusions;*

I7. *stability analysis; stability of computations and model in the neighborhood of zero; stability wrt. all involved parameters.*

I8. *model stability in time;*

I9. *differential connectivity: inspection of deduced connectivity models by comparison to the structures deduced by other faithful connectivity measures and methods;*

I10. *harmonization of basic parameters of involved measures;*

I11. *proper definition of the reduction level (rounding filtered or "negligible" contents);*

I12. *connectivity graphs time expansion;*

I13. *alternative or additional approaches in model integration;*

Granger causality, Geweke developments, PDC, DTF

All details on the method are available in the cited and other literature. All definitions and elements are briefly given in Appendix. When we have three variables $x(t)$, $w(t)$ and $y(t)$, if the value of $x(t+1)$ can be determined better using past values of all the three, rather than using only x and w, then it is said that the variable y Granger causes x, or G-causes x. Here w is a parametric variable or a set of variables.

In the bivariate case, G-causality is expressed using linear autoregressive mode

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 5 of 35

$$x(t) = \sum_{j=1}^{p} a_{11}(j)x(t-j) + \sum_{j=1}^{p} a_{12}(j)y(t-j) + E_1(t)$$

$$y(t) = \sum_{j=1}^{p} a_{21}(j)x(t-j) + \sum_{j=1}^{p} a_{22}(j)y(t-j) + E_2(t), \tag{1}$$

where p is the order of linear model and E_i are the prediction errors. The model consists of the linear recursive and the stochastic component. Thus, if coefficients of y in the first equation of (1) are not all zero, we say that y G-causes x; similarly for variable y. The multivariate formulation was exploited more by Granger followers, Geweke [13,14] and others (e.g. [15]) rewriting (1)

$$\boldsymbol{x}(t) = \sum_{j=1}^{p} \boldsymbol{A}(j)\boldsymbol{x}(t-j) + \boldsymbol{E}(t), \tag{2}$$

where $\boldsymbol{x}(t) = (x_1(t), ..., x_n(t))$ is a vector of variables, $\boldsymbol{A}(j)$, $j = 1, ..., p$ coefficient matrix defining variable contributions at step $t - j$, $\boldsymbol{E}(t)$ are prediction errors. The condition on this model is that the covariances of variables are stationary, which is not always easy to assess.

With other contributions, Geweke introduced spectral form of connectivity measures $I_{j \to i}^2(\lambda)$, G-causality measure from channel j to i at frequency λ (now G- should be doubled) as well as a set of other suitable measures, which were all popular among the followers. He introduced conditional causality; we mention here his linear causality $F_{y \to x}$ of y to x; in frequency domain he introduced the measure of linear causality at a given frequency $f_{y \to x}(\lambda)$, which was followed by other similar or very similar concepts, among which the directed transfer function $DTF_{ij}(\lambda)$, and partial directed coherence $PDC_{ij}(\lambda)$, measuring connectivity from channel j to i at frequency λ gained major attention and application. After numerous analysis and comparisons of these two measures e.g. [17,18], later in [21] authors of PDC defined information PDC and DTF, aiming at measuring the information flow between nodes j and i at frequency λ, the measures $iPDC_{ij}(\lambda)$, $iDTF_{ij}(\lambda)$. They state a theorem in [21] with nine equivalent conditions characterizing absence of connectivity between two nodes j and i, of which we reproduce conditions 4- 6:

0) nodes j and i are not connected; $\qquad\qquad$ (3)
a) $iPDC_{ij}(\lambda) = 0, \forall \lambda \in [-\pi, \pi)$;
b) $iDTF_{ij}(\lambda) = 0, \forall \lambda \in [-\pi, \pi)$;
c) $f_{y \to x}(\lambda) = 0, \forall \lambda \in [-\pi, \pi)$

The theorem is for two var case. For the general case, authors announced soon publishing. Otherwise, we note that all important conclusions in their earlier papers, especially [18] are reaffirmed again in [21].

Computation and comparison of measures

Certain normalizations are often necessary before measure comparisons, when we estimate their difference at a point or on a subset of a common domain [4], e.g. for normalized measures, for compatibility estimation we could define

$$mc(\mu, \nu, \xi) = |\mu(\xi) - \nu(\xi)|,$$
$$mc^*(\mu, \nu, D) = \int_D |\mu(\xi) - \nu(\xi)| d\xi. \tag{4}$$

The measure comparison provides similarity degree - a metrics in a suitable space of measures. After analysis of published measure comparisons we noticed presence of certain operators which we expose here. Measuring similarity of measures (involving other operators [4]): quite generally (observing the contemporary practice needs) we define similarity for measure μ and ν by a scheme

$$Sim(\mu, \nu, i, j, D) = P(sim(N_1(\mu, i, j, D), N_2(\nu, i, j, D))), \tag{5}$$

where N_k are normalization operators, sim a basic similarity, P external operator (e.g. posterior grading), i, j is the graph link from j to i, D parameter-set for N_k; e.g. mc and mc^* measures are the special cases of (5). Choosing the operators properly would contribute to estimation quality; the opposite will generate erroneous reasoning. Computational stability implying semantic stability, depending on involved operators is demanding, before deriving any conclusions. Comparisons in three predicates: basic connectivity between two nodes, directed connectivity, connectivity-grading, should be desagregated- done separately. Previous considerations should include time dynamics which is omitted here while staying closer to the existing practice in the treatment of standard brain connectivity.

Results and discussion

Preservation/coherence properties

Measures satisfy preservation properties, e.g. monotony, cardinal monotony, translation invariance, some additivity, approximations. The following semantic stability criteria STC are mandatory.

P1. *substructure invariance, i.e. restriction of a measure to a substructure should not change its range; thus, measure values on the intersection of substructures remains coherent.*

P2. *Structural stability; measure computation and comparisons/similarity estimates should be invariant to some degree of fluctuations of the involved operators (here P, sim, N_k). These conditions should secure measure stability in extended, repeated and similar experiments.*

P3. *measure comparison* should be stable in all involved parameters.

P4. *continuity in models - similarity measures*: small must remain small and similar has to remain similar in measure comparisons. The small difference of argument implies bounded difference of the result; this applies to predicate *connected*, with small shift of argument.

Structural properties of measures are determined on small objects - in the zero neighborhoods, whence the measure zero ideal is of key importance, which is the reason to list separately the zero-axioms, ZAx, for either a single considered measure or for a set of compared measures:

ZAx:

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 7 of 35

Z0. *substructure partitioning invariance (measure restriction to a substructure remains coherent);*

Z1. *fluctuations of operators involved in measure computation and comparison need be tolerant (continuity);*

Z2. *in similar circumstances numeric zero (significance threshold) should be stable quantity (to allow comparability of results);*

Z3. *comparison of a set of measures needs prior unification (argumentation necessary) of their zero thresholds (for otherwise, what is zero for one is not zero for other measures; consequently, the measure values which are identical for one measure, are discerned as different by other measures; that must cause problems);*

Z4. *in similar circumstances grading should be stable quantity;*

Z5. *measure values which are different by ≤ numeric-zero should remain identical in any posterior computation/grading, if applied (this is in accordance with the prior congruence on the ideal of zero measure sets);*

Z6. *values in any posterior computation/grading (if applied) should differ by no less than numeric-zero and grades should be of unified diameter; in this way values in posterior grading range are harmonics of numeric-zero;*

Z7. *final grading as a (small) finite projection of normalized range* [0, 1] *needs some conceptual harmonization with the standard additivity of measures; this step should involve fuzzification;*

Z8. *grading should be acceptable by various aspects present in the interpretation of related experimental practice (that means that the picture obtained using a projection/grading of* [0, 1] *range should not semantically be distant from the original picture – based on the* [0, 1] *range with, e.g. a sort of continuous grading);*

Z9. *Connectivity graphs time expansion; it is necessary to introduce time dynamics in these observations.*

Mathematical principles must be respected for consistency preservation; computations in repeated and similar experiments have to be comparable and stable. Measure computation and comparisons are complex, consisting of steps, some of which usually do not commute, demanding care and justification.

Enhancements

We will present two methods applicable to the weak brain connectivity, the case when a set of signals share a common information, which is either hardly detectable or even undetectable by direct observation or the connectivity measures in current use. The first is based on rather pragmatic property, partial linear dependence, PLD: for the set of functions (signals or vectors) $S = \{f_i : i \in I\}$ with the common domain, the set of restrictions to a subdomain is linearly dependent (wrt. some usual scalar product), while the complementary restrictions are linearly independent. PLD could be used to extract the common information easier. Then we might be looking for the maximal PLD sets, corresponding to different common information. If we generalize this slightly we come to PD in case when we have dependence which is nonlinear. The second approach exploits the methods originally introduced to eliminate clutter in radar images and to enhance small objects in images and is applicable to both signals and images. Since

spectrograms are a sort of images, we can apply those methods to the spectrograms, spectrogram composites, or somewhat more general objects of the similar kind. Both methods are applicable to the noise contaminated signals.

PLD - PD

If applied to a given frequency, a reduced size set of frequencies, or a known frequency band, this method can supply good answers with not really complex calculations. Similar can be the case for the frequency distributions (spectra, composites, connectivity measures, spectrograms) - parts containing frequencies with poor signal to noise ratio, especially when multiple spectra or spectrograms are available.

Alternatively, if we start with independent sequences, feeding all of them with relatively small magnitude process, we should be able to establish the threshold level from the lower side, i.e. when the shared information becomes perceptible. By a MS – a modulation system we designate the usual meanings of signal modulation, i.e. coding or fusion of information process with some (set of) base function (carrier). In technical practice, an MS can be of any usual sort as AM, FM, PCM, BFH, some of their meaningful combinations, or generalized technically and mathematically. Thus,

$$MS : (F, S) \rightarrow H,$$

where F is a subset of B - a system of base functions, while $S = \{g_1, g_2, ..., g_n\}$ is a set of information contents, H the fusion output, all components of F, S, H are time functions, in practice - finite sequences. In simplest case F, S, H are all singletons. Obviously, a brain connectivity path might accommodate broader activities, inclusive lower frequency and high frequency information patterns. For two functions f_1 and f_2 we say that they are independent wrt. causality measure μ, if

$$\mu(f_1, f_2) = 0. \tag{6}$$

In practice, for experimental f_1, f_2, that would be

$$\mu(f_1, f_2) < \varepsilon \text{ for all } \varepsilon > 0,$$

down to the numeric (noise or statistical) threshold. If we inject/modulate a sequence G of information sequences, into a couple of μ - independent f_1 and f_2, resulting in f_1' and f_2' and if

$$\|G\| < \min(\|f_1\|, \|f_2\|) \tag{7}$$

for a suitable norm, where the norm of G could be e.g.

$$\|G\| = \sup\{\|g\| : g \in S\} \tag{8}$$

we can well have

$$\mu\left(f_1', f_2'\right) \approx 0,$$

while f_1' and f_2' share a vector of information sequences G. The case becomes more complicated if f_1' and f_2' are modulated using different MS's, or have different delays involved in modulation of a set G, or if we deal with spectral features with non-constant frequencies, or if modulation processes involve some headers - protocols. Let us pretend that the PD is PLD, thus simplifying expressions. Similarly, we introduce the local P(L)D, as the (linear) dependence in a subset of a set of all coordinates, in our examples

in the frequency distribution form (e.g. power spectra or composite spectra). A set G of time functions/sequences is locally (linearly) dependent at frequency λ ($\in SR$ – a frequency domain [0,Nf] (Niquest frequency) for sampling rate SR)

$$\pi(G, \lambda) = \Pi\{F(g, \lambda) : g \in G\} > 0, \tag{9}$$

where $F(g, \lambda)$ is e.g. the λ-th coordinate of Fourier (power) spectrum of g. Then for Fourier spectrograms of G in time interval T, define

$$\pi_s(G, \lambda, T) = \Pi\{S(g, \lambda, T) : g \in G\} > 0, \tag{10}$$

where $S(g, \lambda, T)$ is time integral of $F(g, \lambda)$ in the epoch (scrolling time interval) T, i.e. the integral of the time trace of the spectral line at λ:

$$S(g, \lambda, T) = \int_T F(d, \lambda) dt. \tag{11}$$

The condition (9) expresses that all spectra of elements of G have a non zero λ co-ordinate (or its time trace integral). For a fragment Λ of the spectral domain SR, select

$$\pi(G, \Lambda) = \int_\Lambda \pi(G, \lambda) d\lambda, \tag{12}$$

the restriction to Λ of the product of all power spectra of elements of G, and $\pi_s(G,\Lambda,T)$ integrating $\pi_s(G, \lambda, T)$ over Λ likewise:

$$\pi_s(G, \Lambda) = \int_\Lambda \pi_s(G, \lambda, T) d\lambda. \tag{13}$$

Besides we define simple quotient measures for power spectra, energy density indices for λ,

$$ED(g, \lambda) = \frac{F(g, \lambda)}{\displaystyle\int_{SR} F(g, \lambda) d\lambda} \,, \tag{14}$$

the energy at the frequency λ relative total spectral energy and the similar index for spectral neighborhood of λ, i.e. for Λ subset of SR (the spectral resolution - the set of all frequencies in the spectrum). Thus, with

$$ED(g, \lambda, \Lambda) = \frac{F(g, \lambda)}{\displaystyle\int_\Lambda F(g, \lambda) d\lambda}, \tag{15}$$

where Λ is a subset of SR, some neighborhood of λ. Clearly, the more prominent (globally) the spectral line at λ, the higher the first index; the more prominent spectral line at λ locally (within Λ), the higher the second index. Similarly for products for $g \in G$, we define $\Pi ED(G, \lambda)$, and $\Pi ED(G, \lambda, \Lambda)$

$$\Pi ED(G, \lambda, \Lambda) = \frac{\pi(G, \lambda)}{\pi(G, \Lambda)}, \tag{16}$$

$$\Pi ED(G, \lambda) = \frac{\pi(G, \lambda)}{\pi(G, SR)}, \tag{17}$$

and for spectrogram-like composites

$$\Pi EDS(G, \lambda, T) = \frac{\pi_s(G, \lambda, T)}{\pi_s(G, \Lambda, T)},$$ (18)

with $\Pi EDS(G, T) = \Pi EDS(G, SR, T)$.

In practice, all above integrals become finite sums, we might wish to use different indices in different occasions, which is why we distinguish them all. Obviously, we can have situations with the global index negligible while the local index is perceptible. Obviously, we can easily extend the above definitions to include some modulation fluctuations, or we could rephrase the concepts in order to fit better specific needs. In our practice that means that we can search for sets, the subsets of E, signals - electrode network measurements, which contain the same frequency component and select the optimal subsets in P(E). Clearly, the larger set of signals shares common information, the easier should be its extraction. However, with E at 300 now days, growing larger in time, the search for suitable, or larger G's through P(E), having already 2^{298} elements, is quite a task, without a specializing guidance it will miss all optimums. With some previous knowledge on involved functionality, or starting with rather small sets – as seeds and expanding them we might be in a position to learn how to enlarge initial seeds. After preprocessing of connectivity for selectivity for certain application, we can use (linear) dependent channels to enhance the periodic component present in all members of a set G, which might be near or even below noise threshold in all inspected signals, as illustrated with the experiments below. The advantage is in the following property: the local processes contain independent components behaving quite randomly, the noise behaves randomly and random components will be zero flushed by the above criteria, in either composite spectra or composite spectrograms. Even without knowing at which frequencies interesting periodic patterns might be expected, the above method provides a high resolution spectral and spectrogram scanning. If there are artifacts which are characteristic for certain frequency bands, in case when the searched information is out of these bands with sufficient frequency separation, we might be able to localize and extract even the features embedded in the noise. Thus, e.g. in the composite spectra and spectrograms, the first index $ED(g, \lambda)$ might easily converge to the numeric zero, while the second index $ED(g, \Lambda)$, for certain spectral neighborhoods Λ of λ can locally amplify the hidden information, exposing it to perception. The same is even more obvious for spectrograms and products (ΠED, ΠEDS) indices, where we might exploit further properties of spectrogram (composite's) features. When G is modulated by certain MS's, we can still separate the carriers if known or well estimated, even within the same procedure, as above. The above procedure could be extended involving specific sorts of comb like filters, unions of the narrow band filters, which could enhance weak spectral components.

Application of image processing methods

A variety of problems in image processing is related to the contour – object detection, extraction and recognition. This we encountered in cytology preparations, variety of optical images, radar images, spectrogram features [4], mixing and concatenating processing methods.

Small object recognition

We have developed procedures for small object recognition and filtering by size based on the intensity discrimination (intensity of considered pixels) and by specific sort of

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 11 of 35

image spectroscopy. Here we shortly present an alternative method for the efficient recognition of smaller, dot-like objects, with diameter < 10 pixels. Method applies to both matrices and vectors, thus to both images and signals/spectrograms. Spectral features which are stable and narrow in frequency are examples of such sort of vectors. The method is an improved Tomasi, Shi, Kanade procedure (see e.g. [22]) for the extraction of characteristic features from a bitmap. It is robust and proved efficient, possessing all highly desirable properties. As an input we use a simple monochrome (0 = white, 255 = black) bitmap (matrix) A of a fixed format, (here in 400×400 pixel resolution). The components of A, signal amplitude values, or e.g. spectrogram intensities are denoted by $A(x, y)$, where x indicates the corresponding row and y indicates the column. Spatial x -wise and y -wise differences I_x and I_y are defined as follows:

$$I_x = \frac{\partial A(x, y)}{\partial x}, \quad I_y = \frac{\partial A(x, y)}{\partial y}. \tag{19}$$

The matrix G of sums of spatial square differences is given with

$$G = \sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix}, \tag{20}$$

where $\omega_x = \omega_y$ is the width of integration window (with optimal values between 2 and 4), p_x and p_y are such that the formula (20) is defined. Rewriting G more compactly as

$$G = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{21}$$

computing the eigenvalues

$$\lambda_{1,2} = \frac{a+d}{2} \pm \frac{\sqrt{(a-d)^2 + 4bc}}{2}. \tag{22}$$

Define

$$\lambda(x, y) = \min(\lambda_1(x, y), \lambda_2(x, y)) \tag{23}$$

for inner pixels; for given lower threshold T_{min} and parameter A_{max} (here 255) set

$$\lambda_{max} = \max\{\lambda(x, y) | (x, y) \text{ is an inner pixel}\} \tag{24}$$

and define the extraction matrix by

$$E(x, y) = \begin{cases} \frac{A_{max}}{\lambda_{max}} \lambda(x, y), & \frac{A_{max}}{\lambda_{max}} \lambda(x, y) > T_{min} \\ 0, & \frac{A_{max}}{\lambda_{max}} \lambda(x, y) < T_{min} \end{cases}. \tag{25}$$

If two images or spectrograms are available (two consecutive shots or two significantly linearly independent channels) we obtain a solution in even harder case for automatic extraction. Let B and C be two images where every pixel is contaminated with noise which has a normal Gaussian distribution, in which stationary signal is injected, objects at coordinates $(x_1, y_1), \dots (x_{10}, y_{10})$, all with intensity e.g. m (within [0, 255] interval) and fluctuation parameter p; we generate the new binary image A in e.g. two steps (or by some other efficient method):

$$A(x,y) = \text{abs}(B(x,y) - C(x,y))$$
$$\text{If } A(x,y) < p \text{ then } A(x,y) = 255 \text{ else } A(x,y) = 0; \tag{26}$$

This simple discrimination reduces random noise significantly and exhibits signals together with residual noise. Performing procedure (19) thru (26), we generate filtered image with extracted signals. The method is adaptable, using two parameter optimization (minimax): minimal integral surface of detected objects, then maximization of the number of small objects.

Small object recognition using Kalman filter banks

Alternative method for the detection/extraction of small features is based on a bank of Kalman filters. After the construction of the initial sequence of images Z_k, the bank of one-dimensional simplified Kalman filters (see e.g. [23]) is defined using the iterative procedure as follows:

$$K_k(x,y) = \frac{P_{k-1}(x,y) + Q}{P_{k-1}(x,y) + Q + R}$$
$$\hat{X}_k(x,y) = \hat{X}_{k-1}(x,y) + K_k(x,y)\big(Z_k(x,y) - \hat{X}_{k-1}(x,y)\big) \tag{27}$$
$$P_k(x,y) = (1 - K_k(x,y))(P_{k-1}(x,y) + Q).$$

Initially: $P_0(x,y) = \hat{X}_0(x,y) = 0, Q = 1, R = 100$, where Q is the covariance of the noise in the target signal, R is the covariance of noise of the measurement. We put (depending on problem dynamics): the output filtered image in k-th iteration is the matrix $\hat{X}_k$, the last of which is input in the procedure described by equations (19) to (26), finally generating the image with extracted objects. This method shows that the signal level could remain unknown if we approximately know statistical parameters of noise and statistics of measured signal to some extent. In our basic case we know that signal mean is somewhere between 0 and 255 and that it is contaminated with noise with unknown sigma.

Examples
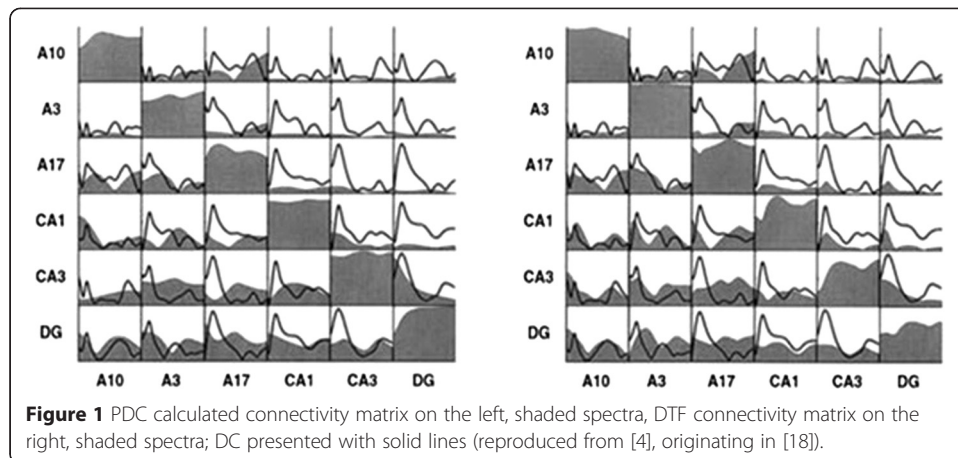#### Connectivity measure evaluations

In this section we focus to the final result of the connectivity measure computations – the brain connectivity directed graphs, as the main model representing brain connectivity patterns. Due to various technological and methodological limitations, contemporary mapping of brain activity using electroencephalography and magneto encephalography operates with a few hundreds of brain signals, thus, close to mega links. No doubt, this resolution will be continuously increasing, down to a few millimeters per electrode and better, all in 3D, increasing proportionally the cardinality of connectivity graphs, as discussed earlier. In a graph we define orbits of individual nodes: the k-*th* orbit of a node a will consist of nodes whose distance via a directed path from node a is k (separate for both in/out paths). It is assumed that the connectivity graphs exhibit direct connections of processes which are directed. This was ambition of all scientists who proposed the connectivity measures in brain analysis; this is expectation of all scientists interpreting their experiments with computation of the connectivity measures. We just add that there might be scenarios where bare connectivity is decidable, while directed is hard to resolve.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 13 of 35

We rather briefly analyze some important published examples of connectivity measure computations and measure comparisons. These measures are commonly used to determine brain connectivity patterns. We will exhibit erroneous or misleading conclusions in modeling of brain connectivity which is of crucial importance for experimental scientists in this area. In [4] we presented critical observations concerning key examples from [18,24], which were originally used in the argumentation in comparison of two major brain connectivity measures, PDC and DTF, showing essential superiority of PDC. We will reconsider our example 6 from [4], where corresponding calculations and comparisons were exposed to our detailed investigation. The reason for this reinspection is in the following. Our methodology used in this example was hypothetical to a small extent: investigating parametric stability of connectivity models following from argumentation of Baccala − Sameshima, we borrowed statistical significance values for PDC from other published results of the same and other authors, while omitting their asymptotic estimation published in [25], because the main theorem there was without proof and a small fraction of the text was unclear, probably with mistake, which in the absence of the proof was not easy to overcome. Since that time, we noticed that numerous renowned researchers use the results from [25] asserting that the PDC statistical significance is solved with results there published. We start with our original example from [4], appending arguments concerning this asymptotic estimation of PDC from [25]. It was multiply shown in [18,24] for the simulated models that PDC exactly determines the structural connectivity graphs of directly connected processes, while DTF is rather imprecise, mixing the direct connections with transitive influences, thus reducing use of DTF to the first orbits. The results for PDC on examples with synthetic models are impressive. Conclusions in [18,24] on the two measure comparisons using neurological data are quite the same: PDC exactly describes the direct structural connectivity, while DTF has undetermined degree of imprecision in description of direct structural connectivity, however, profoundly respecting D. Adams Axiom, with the antecedent regularly fulfilled:

$$(\forall x)(\forall y)(Thresh(x,y) \Rightarrow Conn(x,y)). \tag{28}$$

Example 1. (partly shortened example 6 from [24]). Analyzing structural stability, in order to emphasize importance of all steps in measure computations and comparison procedure, we discuss in more details a crucial example of PDC/DTF computation and comparison, using real neurological data, by Sameshima − Baccala, which was used in quality estimation of PDC over DTF. Analysis of an experiment focused on two shortly separated time slices: [8,10] s and [13,15] s, with frequency range [0,48] Hz, detailed in [18] (exhibiting structural connectivity changes which *supported our demand to expand connectivity models in time dimension*). Recordings were made synchronously at CA1, CA3, A3, A10, A17, and DG electrodes (having a common substructure with another of their experiments). We reproduce some of their findings/diagrams in order to be able to present our analysis. The first time slice representation with mutual interactions of recorded signals- structures for both PDC and DTF is given on Figure 1 (the same way of presentation is rather frequent in the related literature), depicting classical coherence with solid lines; shaded spectra correspond to respectively PDC and DTF calculations.

Figure 1 PDC calculated connectivity matrix on the left, shaded spectra, DTF connectivity matrix on the right, shaded spectra; DC presented with solid lines (reproduced from [4], originating in [18]).
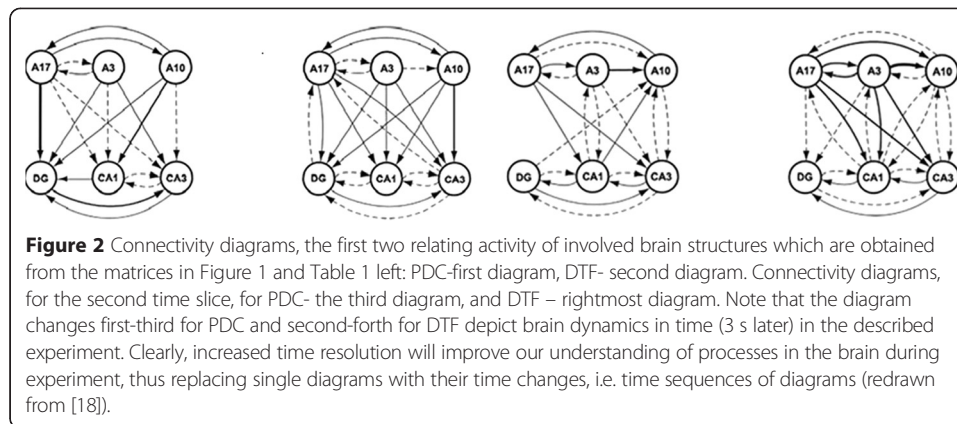
The authors chose here common 0.20 zero-threshold (which is very high: 20% of the normalized range, or four to five times greater than for PDC in their previous examples). They introduced some simplification: not comparing the measures at each respective frequency, these matrices in Figure 1 were used to determine the spectral maximum for calculated PDC and DTF values for all frequencies, for each individual link, which is presented in the Table 1 left side. The matrix in the Table 1 left was used subsequently to integrate the connectivity diagrams for both PDC and DTF, for the first time slice of the experiment, shown as first two graphs in Figure 2. The similar kind of spectral distribution matrices as in Figure 1 was published for second time slice (interval [13,15] s), from which the table of maximums for both PDC and DTF was derived similarly, which is presented in Table 1 right. This matrix of maximums was subsequently used to generate connectivity diagrams for the second time slice, shown as

Table 1 Left side table corresponds to the first time slice of the experiment – related to Figure 1; each matrix coordinate has on top the PDC spectral maximum from the shadow spectrum at corresponding coordinates in Figure 1 left, below the spectral maximum for DTF - similarly obtained from Figure 1 right; the connectivity links are sorted column vise, i.e. in the first column are A10 links towards the areas defined as row names

	A10	A3	A17	CA1	CA3	DG		A10	A3	A17	CA1	CA3	DG
A10		0.18	0.53	0.03	0.06	0.09	**A10**		0.77	0.32	0.49	0.33	0.31
		0.25	0.58	0.06	0.10	0.09			0.82	0.76	0.32	0.22	0.19
A3	0.09		0.27	0.06	0.04	0.07	**A3**	0.09		0.49	0.32	0.06	0.14
	0.13		0.28	0.14	0.12	0.17		0.28		0.79	0.22	0.07	0.09
A17	0.41	0.41		0.11	0.12	0.10	**A17**	0.44	0.36		0.11	0.07	0.11
	0.43	0.47		0.17	0.23	0.23		0.37	0.68		0.34	0.18	0.13
CA1	0.66	0.39	0.23		0.32	0.11	**CA1**	0.14	0.29	0.53		0.44	0.25
	0.49	0.53	0.57		0.32	0.38		0.20	0.65	0.66		0.41	0.28
CA3	0.27	0.46	0.33	0.39		0.67	**CA3**	0.24	0.26	0.54	0.28		0.52
	0.61	0.50	0.51	0.29		0.48		0.27	0.65	0.72	0.22		0.28
DG	0.44	0.44	0.47	0.47	0.41		**DG**	0.10	0.09	0.19	0.53	0.30	
	0.58	0.59	0.35	0.35	0.31			0.15	0.36	0.39	0.57	0.49	
	A10	**A3**	**A17**	**CA1**	**CA3**	**DG**		**A10**	**A3**	**A17**	**CA1**	**CA3**	**DG**

The right side table is obtained in the similar way from the spectral matrices corresponding to the second time slice of the experiment (data reproduced from [4], originating in [18]).

Figure 2 Connectivity diagrams, the first two relating activity of involved brain structures which are obtained from the matrices in Figure 1 and Table 1 left: PDC-first diagram, DTF- second diagram. Connectivity diagrams, for the second time slice, for PDC- the third diagram, and DTF – rightmost diagram. Note that the diagram changes first-third for PDC and second-forth for DTF depict brain dynamics in time (3 s later) in the described experiment. Clearly, increased time resolution will improve our understanding of processes in the brain during experiment, thus replacing single diagrams with their time changes, i.e. time sequences of diagrams (redrawn from [18]).
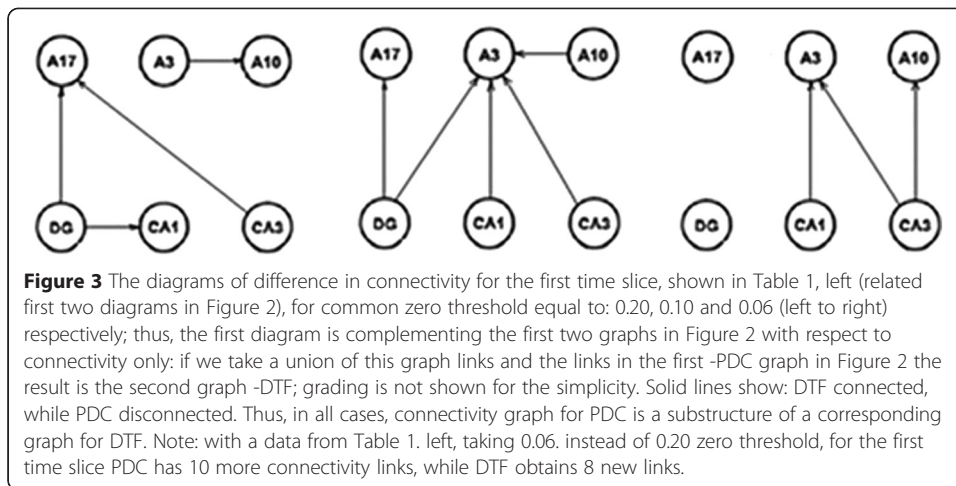
third and fourth graph in Figure 2. Graphs in Figure 2 depict together connectivity patterns and degree (power intensity at maximum) of each connectivity link by arrows in five different degrees (blank for zero and dashed, thin, thicker, thick), in the normalized [0,1] range partitioned into five values, each 0.2 in diameter. Thus, with zero ≤ 0.2, spectral maxima were extracted for each calculated signal pair, projecting - grading the obtained value into the corresponding connectivity degree for each of PDC, DTF, finally considering their difference in connectivity degree to draw the conclusions on PDC/ DTF performance (analysis of connectivity diagrams differences) – Figure 2.

In terms of comparisons/similarity of measures as in [18], we can reconstruct here applied procedure (similarly in numerous other studies, which is partly implicit), as the following sequence of steps (*):

(*) 1. *set* common *zero = 0.2;*
2. *apply N_1 operator: provides PDC power maximum, for all frequencies, for a given pair of input nodes;*
3. *apply N_2 operator: provides DTF power maximum for a given pair of input nodes, for all frequencies;*
4. *apply P operator (the same operator P) for both PDC and for DTF were applied as projections (the five value grading, corresponding to connectivity degree, after calculation of spectral maxima);*
5. *Difference of the graded maxima is exhibited as visualized difference –a pair of connectivity graphs depicting all pairs of signals in the respective time slices, 2 s each.*

First, in concordance with the structural stability conditions which we stated above (on the intersection of two substructures measure is common; in repeated measurements (here, experiments) measure fluctuations must remain tolerable, i.e. obtained values coherent), we will show how rather slight variations of zero threshold, borrowed from the similar experiments cited above and presented in the cited articles, influence connectivity estimates in the same example. Thus, ranging zero threshold thru {0.2, 0.1, 0.06} (thus, including values from other experiments), we obtain three different connectivity difference patterns, for each time slice of this experiment. Rather than comparing all connectivity degrees, we restrict our comparisons to a single quality: *the existence of connectivity only* – shown in the graphs in Figures 3 and 4 as the differences at zero

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 16 of 35

Figure 3 The diagrams of difference in connectivity for the first time slice, shown in Table 1, left (related first two diagrams in Figure 2), for common zero threshold equal to: 0.20, 0.10 and 0.06 (left to right) respectively; thus, the first diagram is complementing the first two graphs in Figure 2 with respect to connectivity only: if we take a union of this graph links and the links in the first -PDC graph in Figure 2 the result is the second graph -DTF; grading is not shown for the simplicity. Solid lines show: DTF connected, while PDC disconnected. Thus, in all cases, connectivity graph for PDC is a substructure of a corresponding graph for DTF. Note: with a data from Table 1. left, taking 0.06. instead of 0.20 zero threshold, for the first time slice PDC has 10 more connectivity links, while DTF obtains 8 new links.

level which is essential. Complete connectivity difference diagrams (including connectivity degrees as usual) are easily regenerated according to the related grading: five grades if zero = 0.2; ten grades if zero = 0.1; 33 grades when zero = 0.06.

We notice immediately that small changes in zero-threshold have substantial consequences in the changes of connectivity structures and their differences. Stability analysis is mandatory whenever we have serious synthesis, i.e. when we organize and map experimental data into higher level structures with semantic significance. The brain connectivity graphs are of high importance and their stability is mandatory. Second, in order to reduce or overcome some of listed problems, we shall make/suggest some changes in the measure comparison sequence, while maintaining the original procedure as much as possible:

(**) 1. *Varying common zero as done in Figures* 3 *and* 4;
2. *apply* N_1 *operator: provides PDC power maximum, for all frequencies, for a given pair of inputs;*
3. *apply* N_2 *operator: provides DTF power maximum, for all frequencies, for a given pair of inputs;*
4. *perform zero- ideal congruence for PDC; i.e. identify the corresponding values from previous step whose difference ≤ zero;*

Figure 4 The diagrams of difference in connectivity for the second time slice of the experiment as shown in Table 1, right (related to third and fourth graph in Figure 2), when common zero threshold is equal to: 0.20, 0.10 and 0.06 respectively (e.g. the first diagram is complementing graphs three and four in Figure 2 with respect to connectivity only, grading not shown for simplicity). Solid lines: DTF connected, while PDC disconnected; dashed lines – opposite. In the first two cases connectivity graph for PDC is not a substructure of a corresponding graph for DTF; note (Table 1, right) the increase of connectivity links for the second time slice, resulting from reduced threshold, which is similar to the first time slice.

5. *perform zero- ideal congruence for DTF;*
6. *perform zero-ideal congruence for PDC and DTF corresponding values;*
7. *generate the graph of connectivity difference;*
8. *apply P operator (the same projector operator P) for both PDC and for DTF, on their respectivegraphs (optional).*

Clearly, in (**) we have two updates of the original procedure (*):

- *zero-threshold:* common as in (*), varying over values which were present in the above mentioned similar, related experiments.
- *zero unification* - performed prior to grading, consequently, avoiding that the small (difference) becomes bigger or big, just because ranges of measures are replaced (simplified) by coarser than original smooth [0,1]-range;

In Table 2 we calculated coordinate vise differences (obtained from Table 1) for corresponding $N_{1,2}$ normalized values for PDC and DTF for both time slices of the experiment. The calculated differences in Table 2 are used in the corrected comparison sequence (**) in the zero – unification step, in order to generate more appropriate diagrams of PDC/DTF connectivity difference, which are presented in Figure 5 (for the first time slice of the experiment, for the three different zero-threshold values) and Figure 6 (for the second time slice of the experiment).
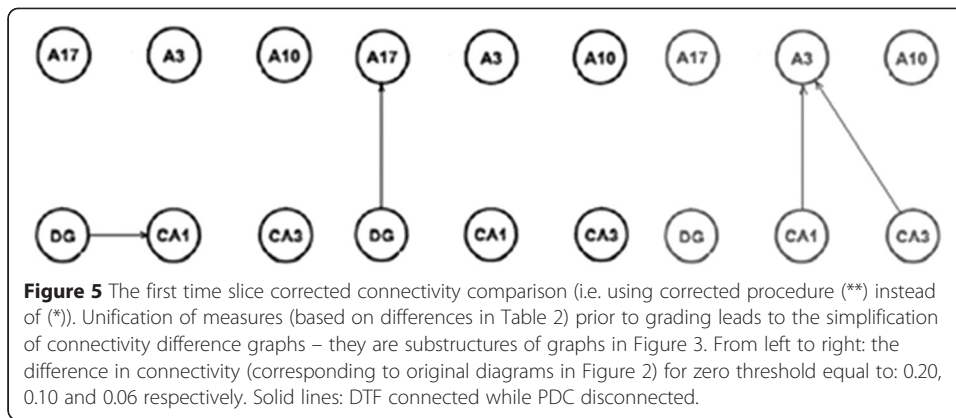
After the above basic convergence of the two measure comparison, we should not omit the following divergence, basically maintaining the same sort of procedure, just introducing the slight variations in the same kind of argumentation.

Third, as mentioned above we did not essentially depart from measure computations and comparison deduced by Sameshima and collaborators in the cited papers. However, we have to notice that the Z3 is violated in the above analysis and resulting graphs, in the following sense: zero thresholds (with the large difference) are unified to the max of the two without proper argumentation. Strictly: the measures have to be independently computed for each node, generating corresponding connectivity graphs. These computations have to be performed independently for each measure, using the corresponding significance level for the zero threshold, without any common zero harmonization. Finally, the agreement of the two measures is presented with the two graphs, to obtain the combined connectivity difference = measure comparison graph. If

Table 2 The difference of PDC power maximum and DTF power maximum coordinate vise

	A10	A3	A17	CA1	CA3	DG		A10	A3	A17	CA1	CA3	DG
A10		−0.07	−0.05	−0.03	−0.04	0.0	**A10**		−0.05	−0.44	0.17	0.11	0.12
A3	−0.04		−0.01	−0.08	−0.08	−0.10	**A3**	−0.19		−0.30	0.10	−0.01	0.05
A17	−0.02	−0.06		−0.06	−0.11	−0.13	**A17**	0.07	−0.32		−0.23	−0.01	−0.02
CA1	0.15	−0.14	−0.34		0.0	−0.27	**CA1**	−0.06	−0.36	−0.13		0.03	−0.03
CA3	−0.34	−0.04	−0.18	0.10		0.19	**CA3**	−0.03	−0.39	−0.18	0.06		0.24
DG	−0.14	−0.15	0.45	0.12	0.10		**DG**	−0.05	−0.27	−0.20	−0.04	−0.19	
	A10	**A3**	**A17**	**CA1**	**CA3**	**DG**		**A10**	**A3**	**A17**	**CA1**	**CA3**	**DG**

Left side corresponds to the first time slice (obtained from Table 1), while the right side corresponds to the second time slice of the experiment (obtained in the same way from Table 1, right side); each matrix coordinate is the difference of the values of PDF and DTF at the same coordinate, as exhibited in Table 1; the connectivity links are sorted column vise, i.e. in the first column are A10 links towards the areas defined as row names.
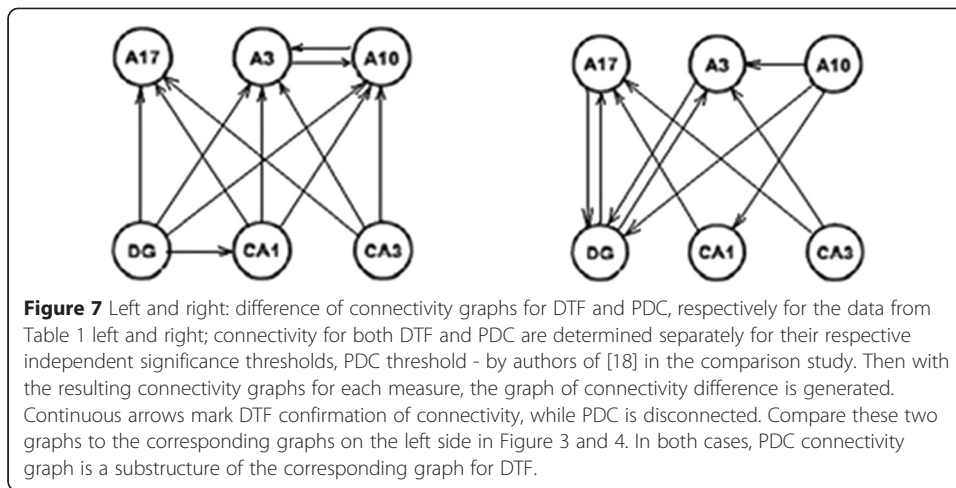
Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 18 of 35

Figure 5 The first time slice corrected connectivity comparison (i.e. using corrected procedure (**) instead of (*)). Unification of measures (based on differences in Table 2) prior to grading leads to the simplification of connectivity difference graphs – they are substructures of graphs in Figure 3. From left to right: the difference in connectivity (corresponding to original diagrams in Figure 2) for zero threshold equal to: 0.20, 0.10 and 0.06 respectively. Solid lines: DTF connected while PDC disconnected.

we strictly follow the procedure from [18] corrected to (***) and the arguments related to statistical significance, with values 0.2 for PDC taken from [18] as above, and known value for normalized DTF (0.0045), then we would get results differing much more. In this case the procedure corresponding to (*) is corrected as follows

(***) 1. *Set zero separately for each of {PDC, DTF};*
2. *apply N_1 operator: provides PDC power maximum, for all frequencies, for a given pair of inputs;*
3. *apply N_2 operator: provides DTF power maximum, for all frequencies, for a given pair of inputs;*
4. *perform zero- ideal congruence for PDC;*
5. *perform zero- ideal congruence for DTF;*
6. *generate the graph of connectivity difference.*

For instance, just for the matrices in Table 1, we have to conclude, for the connectivity only with degree of connectivity omitted, that there are numerous other links differing the resulting graphs. The strict graphs of differential connectivity only for Table 1 respecting (***), we present in Figure 7. Comparing these with the first graphs in Figures 3 and 4, respectively, we note gigantic discrepancy. The same should be performed for other thresholds in these examples. Even, when the measures have identical value, but between the two thresholds one measure will indicate connectivity, the other will deny it; example: in the first matrix

Figure 6 The second time slice corrected connectivity comparison (i.e. using corrected procedure (**) instead of (*)). Unification of measures prior to grading leads to the graphs of connectivity difference, which are substructures of graphs in Figure 4. From left to right, the difference in connectivity (corresponding to original diagrams in Figure 2) for common zero threshold equal to: 0.20, 0.10 and 0.06 respectively. Solid lines: DTF connected while PDC disconnected.

Figure 7 Left and right: difference of connectivity graphs for DTF and PDC, respectively for the data from Table 1 left and right; connectivity for both DTF and PDC are determined separately for their respective independent significance thresholds, PDC threshold - by authors of [18] in the comparison study. Then with the resulting connectivity graphs for each measure, the graph of connectivity difference is generated. Continuous arrows mark DTF confirmation of connectivity, while PDC is disconnected. Compare these two graphs to the corresponding graphs on the left side in Figure 3 and 4. In both cases, PDC connectivity graph is a substructure of the corresponding graph for DTF.

$$0 < \text{DTF–}threshold\ DTF_{DG \to A10}\ = numerically\ \text{PDC}_{DG \to A10} = \text{PDC–threshold};$$

Hence, even when the analyzed systems are tuned so that the compared measures measure all the links approximately identically, when we have largely departing zero thresholds for the involved measures, we can obtain easily arbitrarily large number of links which are zero for one and non-zero for the measure with lower threshold. Moreover, when the measure values are in the opposite order, i.e. when the one with lower threshold is smaller than the other with the larger threshold, but both measure values being between the thresholds, the measure with smaller value of threshold will indicate connectivity, while the other with the larger value will deny it. Example: in the second matrix (in Table 1.)
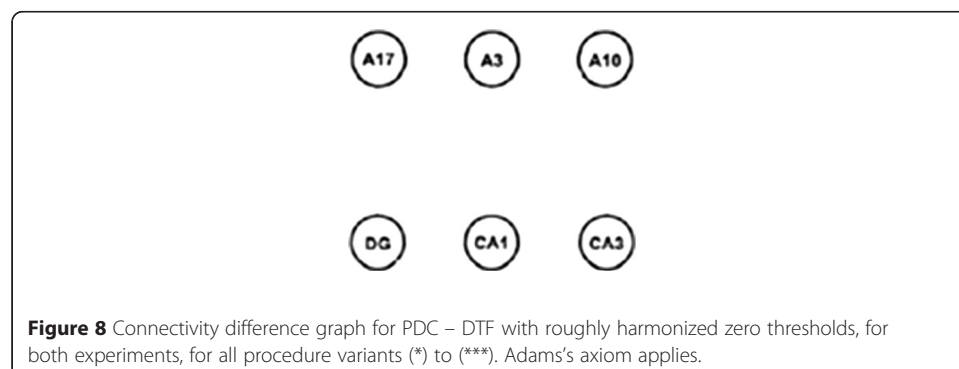
$$0 < threshold\ DTF_{DG \to A3} < numerically\ \text{PDC}_{DG \to A3} = PDC\text{–}threshold,$$

which is completely paradoxal. Obviously, arbitrary choice of zero threshold can distort mathematics and generate paradoxes. Obviously, harmonizing the thresholds (reducing their difference) will reduce or eliminate listed problems. And obviously this cannot be done at will. This is why careful prior investigation of connectivity data related to Z3 and the method is mandatory, or otherwise we remain in the alchemist morasses.

Our fundamental concepts, models and comprehension cannot depend on the sample rate. We will conclude this stability investigation with the following observations. Let us shortly comment the threshold values for PDC published in [25] as obtained from asymptotic studies. The statistical significance for PDC for VAR processes of infinite order, with the estimates for appropriate approximations under certain conditions are established. Theoretical part is without proofs, but there are examples which should illustrate the theoretical achievements. The authors offer under hypothesized conditions the threshold distributions for respectively 20, 200 and 2000 samples, which are bounded respectively by 0.2, 0.15 and 0.01 (for 0.15: the non-constant distribution, with 80% of frequency range below 0.1). The conditions are rather general, so that the above

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 20 of 35

threshold bounds are quite generally applicable, with the last case regularly prevailing. Consequently, regularly used samplings easily provide basis for the application of the last distribution threshold. The first two values corresponding to 20 and 200 data samples have been included in the above elaboration, as the first two threshold values, the first as provided by authors, the second borrowed from similar experiments. With the experiment frequency domain, for 2 s time we have not less than 200 samples. With contemporary usual sampling for the studied intervals one should have not less 2000 samples, which corresponds to the final listed threshold, 0.01. When this threshold is applied, we can say that the thresholds for both measures are roughly harmonized, one for PDC being (still) double of threshold for normalized DTF, which we stated to be the basic condition prior to measure comparison. However, in this case, if we look at the data provided by authors, the connectivity difference graph is identic for both time slices to the graph in Figure 8, applying either original procedure with the PDC threshold adjusted to 0.01 or its corrections. Thus the (**) and (***) corrections are losing importance. Namely, every link gets unified and Adams's axiom applies. Conclusively, in this case, for the data provided with this illustrative experiment, the two measures have no difference in connectivity. Unfortunately, in this highly realistic case as affirmed in [25], all nodes are connected. All that (strongly) implies that PDC advantages essentially or completely vanish with the reasonably realistic increase of order – number of samples followed with measure thresholds harmonization. In this way the original analysis and comparison performed, together with fundamental conclusions, in [17,18] and other publications of these top rated authors, eroded to a complete annihilation and method destabilization, though exclusively based on the data and theory provided by authors, thus strengthening the previous findings published in [4]. Concluding this investigation of measure comparison, we note that our choice of thresholds for PDC in stability study [4], using the borrowed values from similar published experiments, proved to be completely consistent with the values finally delivered in [25], which strongly supported our procedure. (For other challenges with DTF see [26-30]).

Clearly, either the argumentation shown here was unknowable to the authors of [18] or they purposely selected the huge bias in the compared measures thresholds, in order to optimize the targeting conclusions.

Figure 8 Connectivity difference graph for PDC – DTF with roughly harmonized zero thresholds, for both experiments, for all procedure variants (*) to (***). Adams's axiom applies.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 21 of 35

### Experiments with PLD and small object recognition

Here we briefly present potential of the methods, starting with PLD application examples, following with application of small object recognition methods. Some of the work was developed in [5].

Example 2. In the neuro acoustic experiments, the first shown is the simple example where PLD ~ 0 for two spectrograms. The signals containing calibration, external stimulation at respectively 1 kHz and 3 kHz, the two of each, are shown in Figure 9, left side, together with corresponding signal power spectra, from two signals, with direct reading of stimulation tunes intensities, on the right side in Figure 9. Clearly, the power spectra of signals, for the signals of the same tune should be linearly dependent in the stimulated frequency, while the power spectra of signals containing different stimulation tunes would remain independent in the large neighborhoods of the stimulated tunes. Thus, we get here
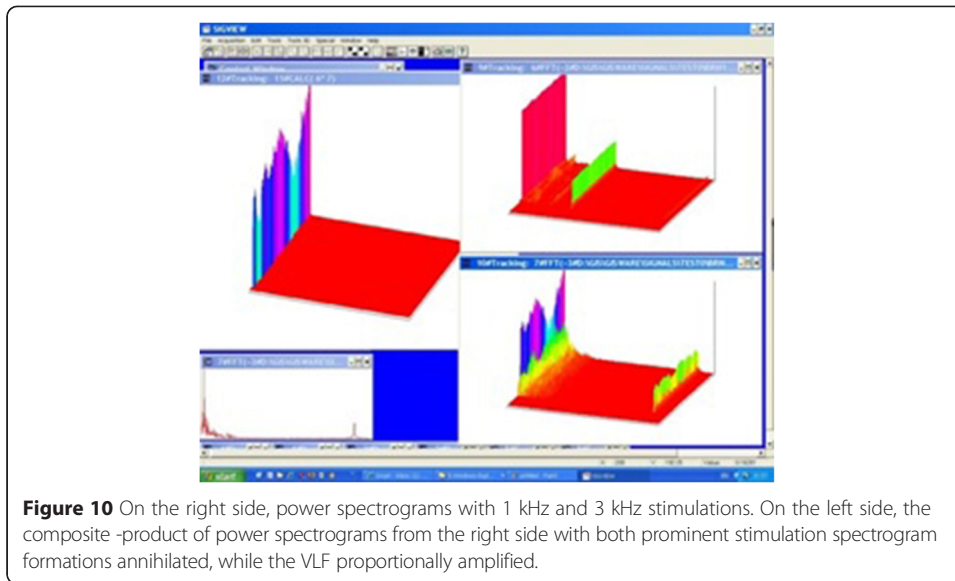
$\pi(\{sig1, sig2\}, 1\ kHz) > 0$, $\pi(\{sig3, sig4\}, 3\ kHz) > 0$, while $\pi(\{sig1/2, sig3\}, \lambda) = 0$ and $\pi(\{sig1/2, sig4\}, \lambda) = 0$,

for large interval Λ of λ's, and similar results for the πs index. This is well shown in Figure 10, right side, top-down, presenting power spectrograms of sig1 and sig3, with 1 kHz and 3 kHz stimulations respectively. On the left side of Figure 10, the composite - product spectrogram is shown (over the whole domain), exhibiting amplified very low frequency - VLF structure, while 1 and 3 kHz structures are mutually annihilated, as linearly independent; here, for $\Lambda = SR\backslash[0, 52\ Hz]$, $\pi(G, \Lambda) \sim 0$ if G contains signals with different stimulation (e.g. {sig1, sig3}, {sig1, sig4}, the same with sig2) else, it is large. The indices *EDS(g, λ, T)*, *EDS(g, Λ, T)*, *ΠEDS(G, λ, T)*, and *ΠEDS(G, Λ, T)* will resolve very well this situation (in the values {0, large number of λ's}), as well.
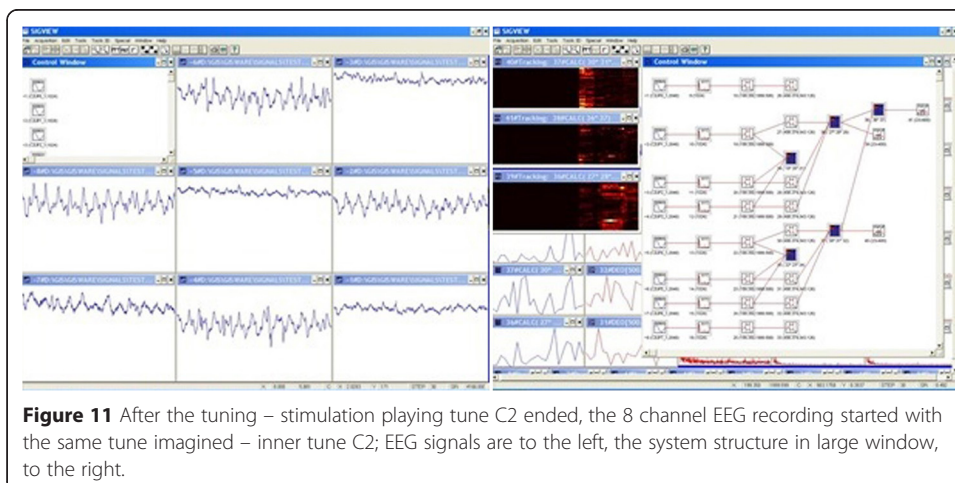
Example 3. PLD has been applied in the experiments with imagined – inner tunes. We briefly comment some of that work, where we filtered - combing the imagined tunes. We filtered - combed signals with imagined tunes together with signals with stimulated tunes, with nice results as well. In Figure 11, left side, we have 8 EEG signals with the imagined tune C2 which were recorded after the externally played C2 - the tuning, ended. On the right we have the system with elements corresponding to their power spectra and a number of PLD components and indices, both local and global.
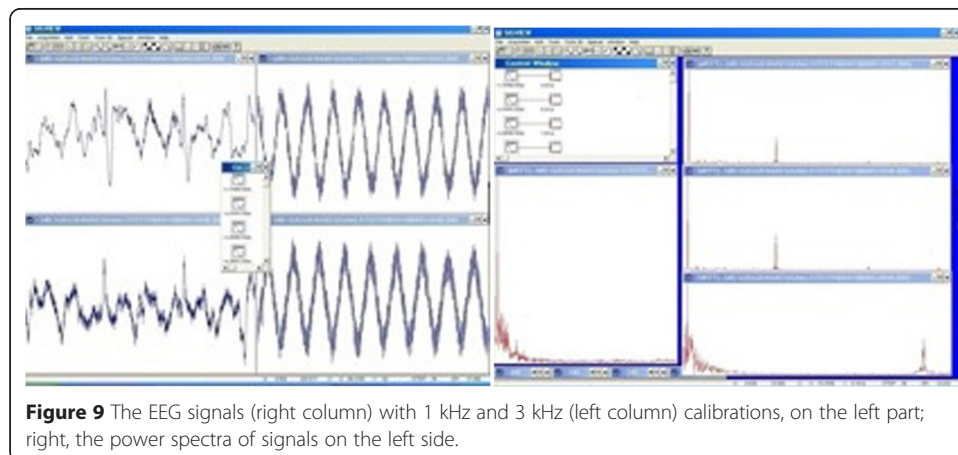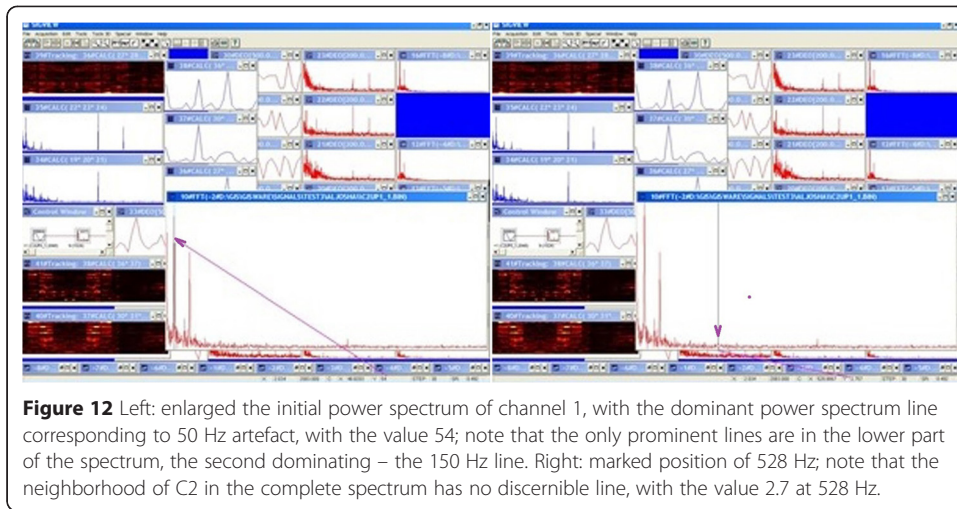

Figure 9 The EEG signals (right column) with 1 kHz and 3 kHz (left column) calibrations, on the left part; right, the power spectra of signals on the left side.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 22 of 35

Figure 10 On the right side, power spectrograms with 1 kHz and 3 kHz stimulations. On the left side, the composite -product of power spectrograms from the right side with both prominent stimulation spectrogram formations annihilated, while the VLF proportionally amplified.
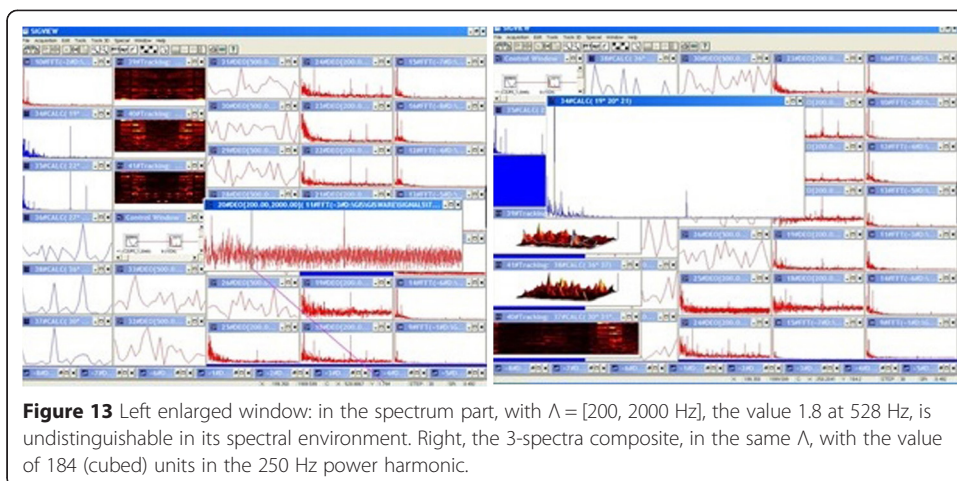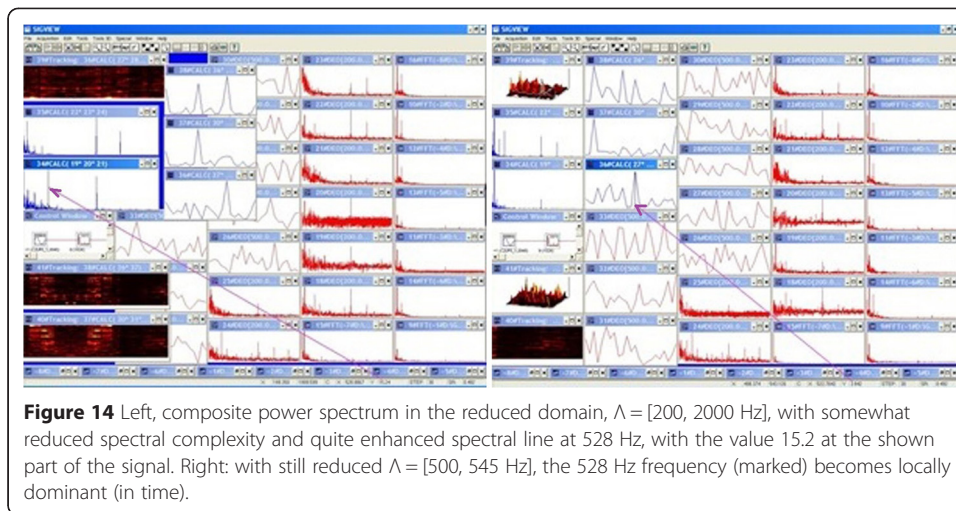
The Figure 12, in the enlarged windows shows major energy distribution in the LF part of spectrum of one channel; relative magnitude of the 50 Hz line has the value 54, while the frequency 528 Hz has the value of 2.7 and is globally and locally indiscernible and embedded in the spectral environment, with a pro mile fraction of spectral energy. In Figure 13 magnified neighborhood of C2 frequency shows no discernible spectral line, while the composite spectra magnify (globally) artifacts, the 50 Hz and its harmonics. Similar holds for spectrograms of individual signals. Power spectra of these signals exhibit some artifacts (50 Hz multiples and some other isolated HF), while the traces of the imagined – inner tune are indistinguishable from the noise level. Switching to PLD indices, multiple spectral dot products reduce overall spectral randomness – the coordinates with random fluctuations are mutually linearly independent, so their

Figure 11 After the tuning – stimulation playing tune C2 ended, the 8 channel EEG recording started with the same tune imagined – inner tune C2; EEG signals are to the left, the system structure in large window, to the right.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 23 of 35

Figure 12 Left: enlarged the initial power spectrum of channel 1, with the dominant power spectrum line corresponding to 50 Hz artefact, with the value 54; note that the only prominent lines are in the lower part of the spectrum, the second dominating – the 150 Hz line. Right: marked position of 528 Hz; note that the neighborhood of C2 in the complete spectrum has no discernible line, with the value 2.7 at 528 Hz.
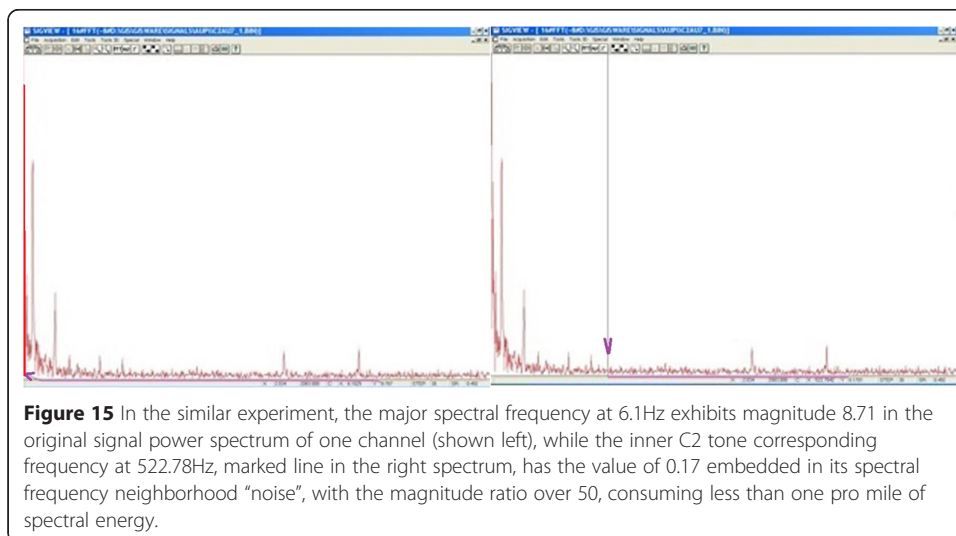
products converge to zero, while the coordinates with the linearly dependent values are enhanced relatively to the noise threshold and become locally discernible or prominent; even when their integral contribution to the overall composite spectral power is very small. Experiments in Figures 14, 15, 16, 17 and 18, in the reduced domain SR to Λ = [200, 2000 Hz], we have the major artifact spectral frequency at 250Hz with value 184 (energy**3), while the inner tone C2 shows 15 units in the composite spectrogram. The power spectrograms time*frequency*intensity matrix S is basically exponentiated by 3: S**3- corresponding to the first 3 EEG channel power spectra coordinate wise product in time, πs({sig1, sig2, sig3}, SR, T) over the 5 s time interval T will be multiply enlarged compared to initial spectrograms. However the ΠED({sig1, sig2, sig3}, C2, [500, 545 Hz]) and its time integration ΠEDS over the selected interval T and Λ = [500, 545 Hz] are becoming dominating high (relative T, Λ), allowing clear C2 recognition in the expected frequency interval (arrow marked composite spectrum max on the right); at the same time the composite energy ratio of restriction to Λ (local) and global – composite spectrum energy over SR is nearly zero, i.e. negligible. At the same time, taking the whole set of signals G = E erases C2: $\Pi ED(G$, C2, [500, 545 Hz]$)$ ~0, similarly
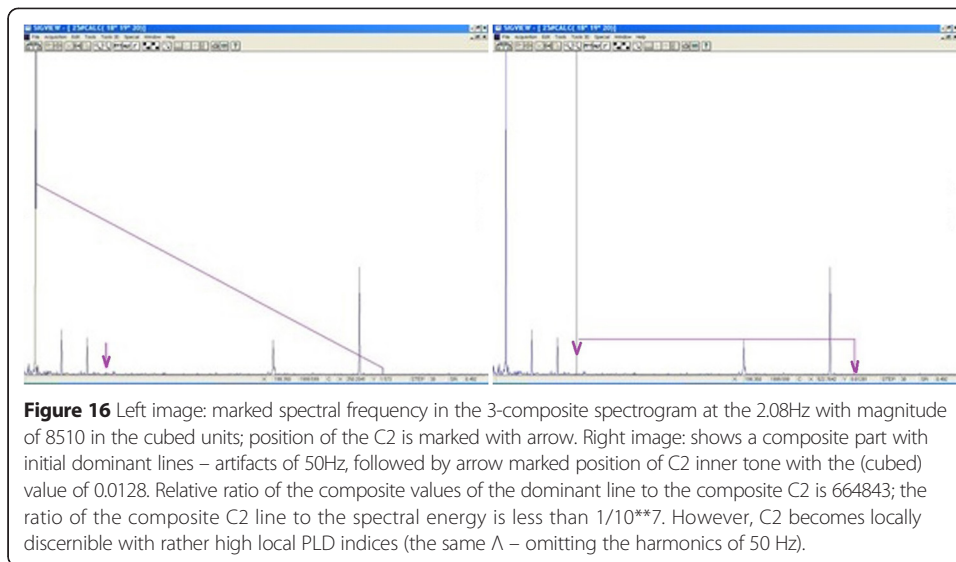

Figure 13 Left enlarged window: in the spectrum part, with Λ = [200, 2000 Hz], the value 1.8 at 528 Hz, is undistinguishable in its spectral environment. Right, the 3-spectra composite, in the same Λ, with the value of 184 (cubed) units in the 250 Hz power harmonic.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 24 of 35

Figure 14 Left, composite power spectrum in the reduced domain, Λ = [200, 2000 Hz], with somewhat reduced spectral complexity and quite enhanced spectral line at 528 Hz, with the value 15.2 at the shown part of the signal. Right: with still reduced Λ = [500, 545 Hz], the 528 Hz frequency (marked) becomes locally dominant (in time).

ΠEDS show no trace of C2, confirming it is not present in all elements of *G*. Rather similar situation we have in other experiments shown on the following figures, supporting the PLD effectiveness in case when there is a common small or invisible information in multiple signal spectra, and more so, for spectrograms. The sets of signals which do not share "common information" are erased. Similar conclusions should work in more general cases.

Example 4. In the following synthetic example we have introduced several dots (useful signals) with the amplitude $a = 120$ and we have contaminated the image with random and the cloudlike noise. The left hand side image in Figure 19 shows bitmap with the random contamination of signal – dots. The right hand side image of the same figure shows the resulting bitmap after the application of the procedure for the noise reduction: initially setting $A_{max} = 255$, $T_{min} = 124$, the extraction procedure yields image shown in Figure 19 right. Somewhat different situation we have in Figure 20. An application of the method of
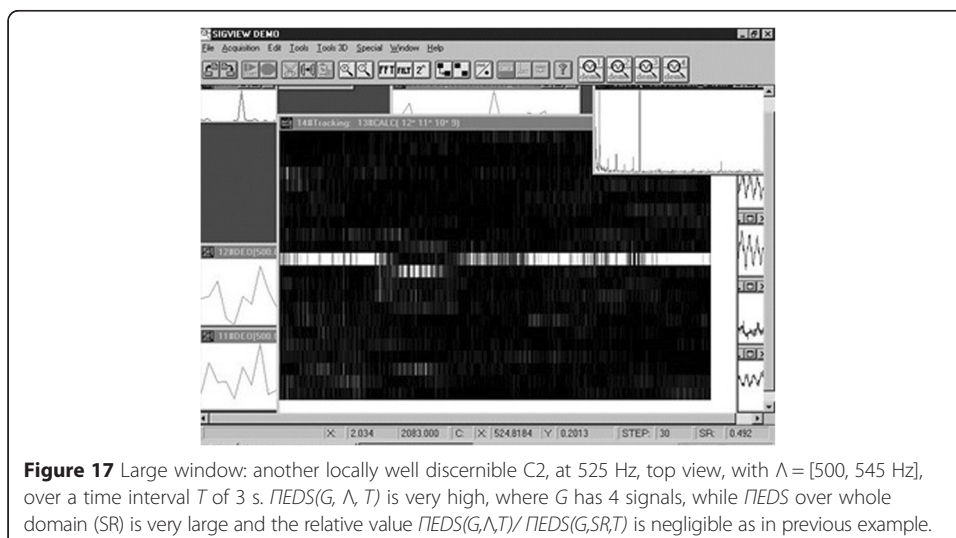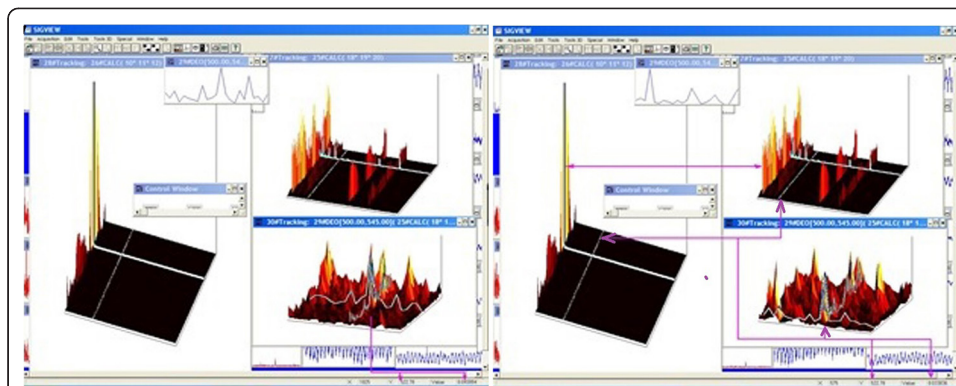

Figure 15 In the similar experiment, the major spectral frequency at 6.1Hz exhibits magnitude 8.71 in the original signal power spectrum of one channel (shown left), while the inner C2 tone corresponding frequency at 522.78Hz, marked line in the right spectrum, has the value of 0.17 embedded in its spectral frequency neighborhood "noise", with the magnitude ratio over 50, consuming less than one pro mile of spectral energy.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 25 of 35

Figure 16 Left image: marked spectral frequency in the 3-composite spectrogram at the 2.08Hz with magnitude of 8510 in the cubed units; position of the C2 is marked with arrow. Right image: shows a composite part with initial dominant lines – artifacts of 50Hz, followed by arrow marked position of C2 inner tone with the (cubed) value of 0.0128. Relative ratio of the composite values of the dominant line to the composite C2 is 664843; the ratio of the composite C2 line to the spectral energy is less than 1/10**7. However, C2 becomes locally discernible with rather high local PLD indices (the same Λ – omitting the harmonics of 50 Hz).

small feature extraction with signal embedded in the Gaussian noise, with one source and the two independent sources, are shown on Figures 21 and 22 (respectively), verifies the problem approach with the method of small object recognition.

Example 5. An application of Kalman filters in small object extraction. In the experiment shown, the initial sequence of images Z_k of the size 200×200 pixels is generated as follows. First, in each image we have introduced the noise by $Z_k(x, y)$ = randn(0, 90); Here " randn " generates random numbers in the interval $[0, 255]$ using Gaussian distribution with $\mu = 0$ and $\sigma = 90$. Then, in each image 10 objects are injected (useful signal) at the same positions, each of them of the size around 10 pixels, with random (Gaussian) fluctuation in intensity around mean value (here 120). After the construction of the initial sequence of images, the bank of $200 \times 200 = 40000$ one-dimensional simplified Kalman filters is defined using the iterative procedure as defined above. The result of Kalman filtering is passed as an input to the initial procedure for small object recognition, finally resulting in the image with extracted objects which is shown in Figure 23. We can notice

Figure 17 Large window: another locally well discernible C2, at 525 Hz, top view, with $\Lambda = [500, 545$ Hz], over a time interval T of 3 s. $\Pi EDS(G, \Lambda, T)$ is very high, where G has 4 signals, while ΠEDS over whole domain (SR) is very large and the relative value $\Pi EDS(G,\Lambda,T)/ \Pi EDS(G,SR,T)$ is negligible as in previous example.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4
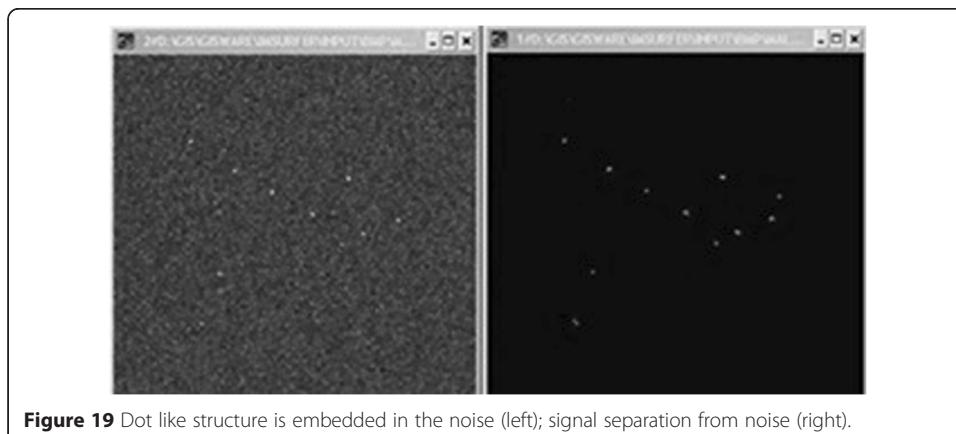
Page 26 of 35

Figure 18 Another example with search for inner C2. In both images, on the left side enlarged are composite power distributions, with LF very magnified and the cross measurements active, showing the value of max at VLF and the 522.8 Hz indistinguishable transversal profile, along the time axes. In the top right diagrams, we see the Λ = SR/Initial VLF, showing some structures arising from 0 level, while the temporal feature corresponding to C2 is still negligible at the cross line, within the 0 floor, lower arrow in the right-right top structure. The C2 temporally lasting features emerged in the reduced Λ = [500, 545 Hz], as seen in both images in lower right windows (arrowed features), the 3-composite spectrogram structures. The proportions of PLD indices are similar to previous cases.

that the minor small objects reshaping is present in the result, with the whole pattern preservation. Further improvements and corrections are possible. The method of small object recognition originally developed for marine radar object tracking, works with vectors too. It is applicable for automatic extraction of signals which are embedded in noise and are imperceptible, in spectra and spectrograms as well, like PLD, especially in case when we can provide at least two sources which are sufficiently linearly independent. The performance constraint to small object – those within 10 pix in diameter is quite generally easily met with spectra, spectrograms, composites and frequency distributions like e.g. connectivity measure and its time trace.

Conclusions

Consistently and consequently with the initial points and preservation properties and the presented examples and argumentation, the following more or less compact conclusion on connectivity measure computation, comparison, comprising methods for the
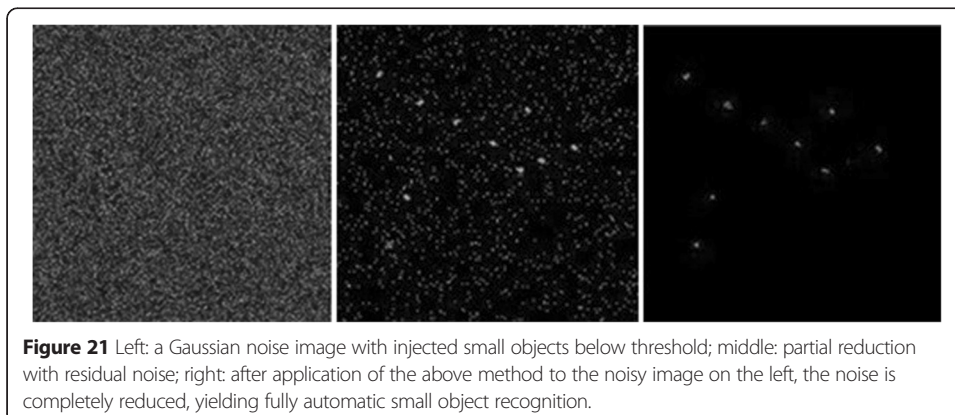

Figure 19 Dot like structure is embedded in the noise (left); signal separation from noise (right).
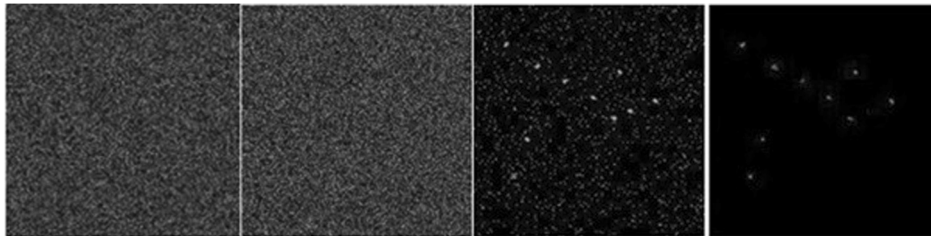
Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 27 of 35

Figure 20 The left hand side image shows similar example to the Figure 19 contaminated with nonhomogeneous noise containing aggregations and granular elements similar in size and intensity to the signal. The right hand side image shows results of the reduction of noise: some new dots belonging to noise cannot be distinguished from the signal – top and low right. Note that the amplitude of the target signal is lower than the chosen lower threshold.

weak connectivity, incorporating remarks from [4] in order to group together the complete set is formulated as

Proposition 1

1. *Separation of different properties.* We proposed here to detach and separately investigate connectivity from connectivity degree. We propose further to distinguish between directed connectivity and non-directed connectivity. There are different situations in which it could be possible to establish the last without solving directed connectivity, especially in the case of weak connectivity.
2. *Differentiated properties could* be investigated with different methods.
3. *Partial linear dependence and method of small object recognition.* They can determine graph substructures with the shared information, which is their contribution in case when the connectivity measures do not resolve the issue for whatever reason. They can be used to generate time expansion of connectivity structures, exposing model's time dynamics.

Figure 21 Left: a Gaussian noise image with injected small objects below threshold; middle: partial reduction with residual noise; right: after application of the above method to the noisy image on the left, the noise is completely reduced, yielding fully automatic small object recognition.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 28 of 35

Figure 22 Two Gaussian noise images with the injected small objects below threshold, following with the signals well extracted and the noise completely reduced – rightmost image.

4. *The PLD indices* might be of special interest if there is any prior knowledge on where the masked information, as a frequency pulse or trigonometric polynomial components might be.

5. *Covering noise.* Both enhancing methods might resolve the common information in case when it is indiscernibly embedded in noise or e.g. spectral neighborhood.

6. *Time delay.* Both methods might uncover substructures with common information with tracking in time, resolving possibly present time delays.

7. *Comparison/computational sequence.* The corrected comparisons of DTF and PDC, for connectivity only, as performed above with published data, ultimately show very reduced differences of two measures for zero-threshold received from the analyzed and similar published experiments (most importantly above the zero-threshold, exhibiting connected structures versus those which are not), thus confirming that if analyzed with computation and comparison procedure corrections proposed here, the connectivity structures are much less different than it was demonstrated in [17,18], as presented in the graphs to the left, Figures 3 and 4, versus Figures 5 and 6 (left graphs) with original *common* zero threshold. However, strictly performed original comparison method opens room for large connectivity difference, and large connectivity model oscillations (Figures 3, 4, 5, 6, 7 and 8). This is even more accented with the asymptotic study by authors, where 0.01 is very reasonable PDC threshold. This zero threshold eliminates the difference of measures in the offered examples with the author's comparison methodology, against original aims and involves other undesirable properties.
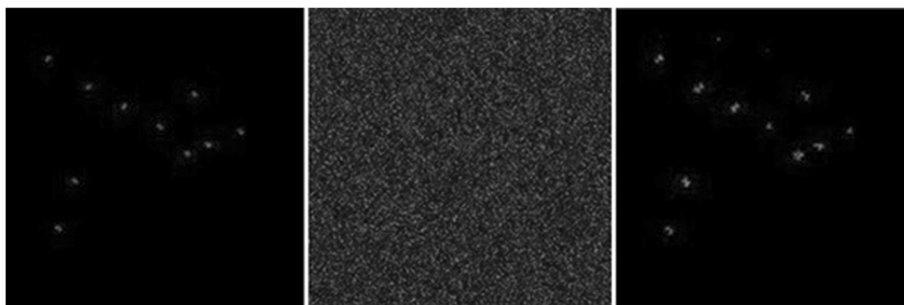
Figure 23 Left: injected signal; middle: the image on the left injected in the Gaussian noise - contamination; right: the result of processing after 44th iteration and application of the method described by formulae 19-26, providing complete object extraction.

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 29 of 35

8. *The instability of PDC – DTF comparisons* is most significantly due to the possibly different values of statistical significance of PDC, ultimately corresponding to different sampling.

9. *Aggregation prior to comparison; functionally related frequencies.* The above analysis was performed, maintaining strictly the reasons and methodology performed by authors of the original analysis [17,18]. Here we have to stress that performed as it is and with our interventions in the original evaluations as well, PDC and DTF measure comparison was not performed directly on the results of these measures computations, thus, comparing directly $DTF_{ij}(\lambda)$ and $PDC_{ij}(\lambda)$, (for relevant $\lambda's$) - the results of measurements at each couple of nodes for each frequency in the frequency domain, but, instead, as in the cited articles with the original comparisons, the measurement of differences of these two connectivity measures was performed on their synthetic representations - their prior "normalizations" - aggregations, obtained as the

$$\max\{\mu(i,j,\lambda) \mid \lambda \in rng(Sp)\}$$

(where $rng(Sp)$ is the effective spectral - frequency range) and, in the original, on their further coarser projections.

10. *Essential departure from connectivity comparison analysis.* In this way, in comparison of these measures the authors had substantial departure from original connectivity measurement computations for PDC and DTF which cannot be accepted without detailed further argumentation.

11. *Comparison of connectivity in unrelated frequencies.* If the parts of frequency domain are related to different processes which are (completely) unrelated, for example, if one spectral band is responsible for movement detection in BCI, while the other is manifested in the deep sleep, then, depending on the application, either can be taken as representative, but most often we will not take such individual maxima of both as representing quantity in the functional connectivity analysis; however, on the restrictions such procedure can be completely reasonable. If we look closely at the corresponding coordinates in the distribution matrices in Figure 1, for the compared measures we will find examples of frequency maxima distant in the frequency domain or even in the opposite sides of frequency distributions (e.g. (5,1) – first column fifth row; then (4,2), (5,2) or (6,5)).

12. *Essential abandoning of the frequency domain in PDC, DTF measure computation and comparison.* Obviously, we are approaching the question: when we have advantage of frequency measures over the temporal domain measures. In the comparison of DTF and PDC via their aggregations as explained above, much simpler insight is obtained; supplementary argumentation is necessary: for the aggregation choice, stability estimation complemented with comparison of DTF and PDC counterparts, for which we would propose their G-inverses DTF and PDC, as we introduced them.

13. *Zero thresholds and connectedness.* Maintaining original or corrected computational sequence in measure comparison, note that DTF computation and PDC computation are performed independently. Consequently, each of these two measures computations should apply the corresponding appropriate zero threshold, thus determining zero-DTF and zero-PDC independently for each of the

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 30 of 35

computations, with independent connectedness conclusion for each measure for each pair of nodes. Then, the connectivity graphs would be statistically correct. However, if the two zeroes differ substantially, that might cause paradoxal results in comparisons. Possible zero threshold harmonization - unification would be highly desirable, as in e.g. [18] and other cited articles, but it must be properly derived.
  Our fundamental concepts, models and comprehension cannot depend on the sample rate.

14. *Aggregations over frequency domain.* Note that frequency distributions exhibited in Figure 1 for PDC and DTF are somewhere identical, somewhere similar/ proportional and somewhere hardly related at all - as the consequence of different nature of these two measures (which is established by other numerous elaborations). The same is true for the spectral parts above zero threshold.
  Consequently, if the comparison of the two measures connectivity graphs was performed at individual frequencies or narrow frequency bands, the resulting graphs of differences in connectivity would be more fateful; they would be similar to those presented at certain frequencies, but would differ much more on the whole frequency domains. Obviously, connectivity at certain frequency or provably related frequencies is sufficient connectivity criterion; such criterion is valid to establish that compared measures behave consistently or diverge.

15. *Brain dynamics and connectivity measures.* Spectral time distributions – Spectrogram like instead of spectral distributions are necessary to depict brain dynamics. In the cited articles, dynamic spectral behavior is nowhere mentioned in measure comparison considerations, but it is modestly present in some examples of brain connectivity modeling – illustrating PDC applicability to the analysis focused on specific event - details in [17,18]. Trend change: in [27] authors recently started using matrices of spectrogram distributions instead of matrix distributions as in Figure 1. This is gaining popularity.

16. *Spectral stability analysis.* Comparison of PDC and DTF as in here analyzed articles, shows no concerns related to frequency distribution stability /spectral dynamics and comparison results. It is clear that comparisons based on individual frequency distributions are essentially insufficient, except in proved stationary spectra, and that local time history of frequency distributions – spectrograms, need to be used instead. Brain is not a static machine with a single step instruction execution.

17. *Characterization theorem for (dis)connectedness* [21]. Here we have simply sensitive play of quantifiers. By contraposition of the statement of the characterization theorem, involving information PDC and DTF as cited above, we obtain equivalence of the following conditions
  o) the nodes j, i are connected;
  a') $\exists \lambda (\lambda \in [-\pi, \pi] \wedge i\mathrm{PDC}_{ij}(\lambda) \neq 0)$;
  b') $\exists \lambda (\lambda \in [-\pi, \pi] \wedge i\mathrm{DTF}_{ij}(\lambda) \neq 0)$;
  c') $\exists \lambda (\lambda \in [-\pi, \pi] \wedge f_{j \rightarrow i}(\lambda) \neq 0)$;
  and similarly with other conditions in the list.
  Observe conditions a') and b').

18. *Note that λ is independently existentially quantified above.* That would suggest that iPDC and iDTF simultaneously confirm the existence of connectivity from j to i.

However they might do it in totally unrelated frequencies, which could make that equivalence meaningless, similarly as discussed in 9. This is all wrong.

19. *The equivalence of* a') *and* b') *clearly contradicts the nonequivalence of PDC and DTF,* which is extensively verified in the cited very detailed analysis of Sameshima and collaborators, since these are the special cases of iPDC and iDTF. However, the statement of the theorem is true for the two var case only, when the orbits are reduced to 1st orbits only. In this case cumulative influence reduces to the direct influence, with no transitive nodes. This generates a limit for the theorem generalizations.

20. *Zero threshold in iPDC and iDTF.* Authors in [21] do not mention zero thresholds at all. As shown multiply above, in practice it has to be determined. Again, as in detail discussed above, note that the same problems are equally present here. E.g. computationally we could easily have

$$0 < i\mathrm{DTF}_{ij}(\lambda) = i\mathrm{PDC}_{ij}(\lambda) = 0.$$

Nobody will like that.

21. *Recent DTF based connectivity graphs with simplified orbits.* In the recent publications and conference reports of research teams using DTF as connectivity measure [27-29], presenting even rather complex brain connectivity graphs involving rather numerous nodes, majority of graphs contain practically only 1st orbits, which is the case when deficiencies of DTF are significantly masked since cascade connectivity is hidden, graphs are not faithful, departing seriously from reality.

22. *Both DTF and PDC measures are not applicable in real time applications* like Brain Computer Interfaces – BCI, where the will generated patterns in brain signals are recognized and classified by a number of direct methods. Some of methods related to weak connectivity are applicable in real time.

23. *The DTF based connectivity diagrams where the zero is chosen arbitrarily high or much higher than the established zero threshold and where connectivity is restricted to a single narrow band*, intentionally reduce the number of really connected connectivity links by large amount, offering highly distorted facts that are established by DTF. The similar holds for synthetic spectrogram connectivity matrices. If the methodology of [18] and [21] was used, one could not deduce less than 10 times more connected nodes in the *"memory"* task and the *"cognitive"* task using DTF, which is strongly inconsistent with the presented connectivity diagrams – factual proofs, which the DTF authors derived from the supplied matrices. In all experiments DTF converges towards Adams Axiom.

24. *The DTF has been making a number of serious problems since its invention.* The authors have been continuously making efforts to solve the problems inventing newer modifications of DTF, adding additional measures, or applying arbitrary restrictions to their connectivity measure in order to reach connectivity diagrams which should look more faithful. Hardly had they succeeded in these intentions.

Clearly, without careful mathematical consideration and argumentation connectivity graphs, in here cited and many others published articles are of shaken fatefulness and need supplementary corroboration. Connectivity measures are different enough that

the question of their logical coherence is appropriate. This is elaborated through measure comparisons. Here the published comparison of DTF and PDC measures is discussed in some detail as an illustrative example, giving enough material for this issue to be more carefully investigated. As verified on a number of nontrivial synthetic systems, connectivity conclusions by DTF are not well founded, while PDC has good capacity in precise structural description, confirming PDC superiority to DTF measure. Quite often PDC ≤ DTF, but it does not hold generally, hence PDC is not a general refinement of DTF and these two measures are essentially different, especially if compared frequency – pointwise. When applied to real neurologic data with the original methodology, the methods seem to be highly semantically unstable generating large model structural oscillations with possible PDC threshold variation. Quite generally on published data, when thresholds are roughly harmonized PDC-DTF connectivity differences vanish, opposite to the conclusions published in [18]. Comparison after frequency aggregations leads to wrong conclusions on functional connectivity, unless appropriate modulators are involved. The zero threshold harmonization when comparing measures is a difficult and challenging issue which ought's to be solved properly, prior to measure computations and comparisons in general.

Two methods, the Partial Linear Dependence, PLD and the method of small object recognition are added for enhanced connectivity problem treatment, comprising time expansions, with examples of their contribution in cases when the shared information is masked or embedded in noise. The number of innovative alternative approaches is growing; aiming to overcome certain difficulties they are successfully applied in demanding applications e.g. [31,32].

Appendix

Here we list some terminology, mostly introduced by Geweke and followers, with definition of mostly used and prominent brain connectivity measures.

Geweke [13] defined spectral form of G-causality, which from (2) by Fourier transform, gives:

$$A(\lambda)x(\lambda) = E(\lambda), \tag{30}$$

$$A(\lambda) = -\sum_{j=0}^{p} A(j)e^{-2i\pi\lambda j}, \tag{31}$$

with $A(0) = -I$ giving for x

$$x(\lambda) = A^{-1}(\lambda)E(\lambda) = H(\lambda)E(\lambda), \tag{32}$$

where H is a transfer matrix of the system. In the bivariate case or with two blocks of variables, G-causality measure from channel j to i at frequency λ, Geweke [13] defined by

$$I_{j \to i}^2 = \left| H_{ij}(\lambda) \right|^2 = |a_{ij}(\lambda)|^2 |A(\lambda)|^{-2}. \tag{33}$$

He introduced conditional causality and a number of measures; his linear causality of y to x is defined as

$$F_{y \to x} = \ln(|\Sigma_1|/|\Sigma_2|), \tag{34}$$

where $\Sigma_1 = \text{var}(\varepsilon_1)$, $\Sigma_2 = \text{var}(E_1(t))$, with similar expressions for vector variables. In

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 33 of 35

frequency domain he introduced the measure of linear causality at a given frequency. Stated for two variables or two blocks of variables:

$$f_{y \to x}(\lambda) = \ln\left(|S_{xx}(\lambda)| |H_{xx}(\lambda)\Sigma_2(\lambda)H_{xx}^*(\lambda)|^{-1} \right). \tag{35}$$

Here, $H_{xx}^*(\lambda)$ is the Hermitian transpose of $H_{xx}(\lambda)$, $|\cdot|$ is matrix determinant and $S_{xx}(\lambda)$ is the upper left block of the spectral density matrix $S(\lambda)$ written as

$$S(\lambda) = \begin{bmatrix} S_{xx}(\lambda) & S_{yx}^*(\lambda) \\ S_{yx}(\lambda) & S_{yy}(\lambda) \end{bmatrix} = H(\lambda)\Sigma_2(\lambda)H^*(\lambda), \ H(\lambda) = \begin{bmatrix} H_{xx}(\lambda) & H_{xy}(\lambda) \\ H_{yx}(\lambda) & H_{yy}(\lambda) \end{bmatrix}.$$

We mention improvement of Geweke causality measures, proposed in [19], using matrix partitioning, providing corrections of Geweke conditional measures not suffering of deficits of the original Geweke measures – occasional negative values and peaks believed to be artifacts. Later, also in frequency domain, Kaminski and Blinowska [15] introduced an adaptation of Granger - Geweke causality measure to m variables, which they called Directed Transfer Function (DTF), with

$$\text{DTF}_{ij}(\lambda) = |H_{ij}(\lambda)| \left(\sum_{k=1}^{n} |H_{ik}|^2 \right)^{-\frac{1}{2}}, \tag{36}$$

measuring causality from j to i at frequency λ; before, they use the same expression as in (32) for the non-normalized DTF definition. Authors of DTF multiply claimed that DTF is superior over Granger's measure in causality application to the brain connectivity problems, but with accumulation of experience with DTF, Kaminski et al [16] propose use of additional connectivity measure DC together with DTF, in order to reach direct connectivity between nodes i and j in frequency domain; DC is defined by

$$\text{DC}_{ij}(\lambda) = \sigma_{jj}H_{ij}(\lambda) \left(\sum_{k=1}^{n} \sigma_{kk}^2 |H_{ik}(\lambda)|^2 \right)^{-1/2}, \tag{37}$$

where σ_{kl} $(k, l = 1, ..., n)$ are components of the covariance matrix Σ_2. DC measure was earlier considered by Sameshima and Baccala, e.g. [17], and other authors. Baccala and Sameshima [18] introduced a normalized measure called Partial Directed Coherence (PDC), measuring direct influence of channel j to channel i at frequency λ, with

$$\text{PDC}_{ij}(\lambda) = \pi_{ij}(\lambda) = A_{ij}(\lambda) \left(\boldsymbol{a}_j^*(\lambda)\boldsymbol{a}_j(\lambda) \right)^{-1/2}, \tag{38}$$

where $\boldsymbol{a}_j$ is the j-th column in $A(\lambda)$ and $\boldsymbol{a}_j^*$ is Hermitian transpose of $\boldsymbol{a}_j$. Among certain further generalizations, we mention here $i\text{PDC}_{ij}(\lambda)$, the *information* PDC, intended to measure information flow between nodes j and i (in the sense of Information theory) by Sameshima and collaborators [21], which is obtained from (37) with the expansion by a factor

$$i\text{PDC}_{ij}(\lambda) = \bar{A}_{ij}(\lambda)\sigma_{ii}^{-1/2} \left(\bar{a}_j^*(\lambda)\Sigma_{\boldsymbol{w}}^{-1}\bar{a}_j(\lambda) \right)^{-1/2}, \tag{39}$$

where $\Sigma_{\boldsymbol{w}} = \mathbb{E}(\boldsymbol{w}(n)\boldsymbol{w}^T(n))$ is a positive definite covariance matrix of the so called zero mean wide stationary process $\boldsymbol{w}(n)$. With the same intention for DTF, they define *information* DTF as following,

Kasum *et al. EPJ Nonlinear Biomedical Physics* (2015) 3:4

Page 34 of 35

$$i\text{DTF}_{ij}(\lambda) = \bar{H}_{ij}(\lambda)\rho_{jj}^{1/2}\left(\bar{h}_j^*(\lambda)\boldsymbol{\Sigma}_{\boldsymbol{w}}^{-1}\bar{h}_j(\lambda)\right)^{-1/2}, \tag{40}$$

where ρ_{jj} is the variance of the so called partialized innovation process $\zeta_j(n)$ defined by $\zeta_j(n) = w_j(n) - \mathbb{E}\big(w_j(n)/\{w_l(n) : l \neq j\}\big)$; obviously, as generalizations they have respectively PDC and DTF as their special cases.

Granger inverse. In [18] authors claim that PDC is the proper counterpart of Granger causality measure in frequency domain. We observe first the fundamental agreement of the two approaches given by Geweke:

$$F_{\boldsymbol{y} \to \boldsymbol{x}} = (2\pi)^{-1} \int\limits_{-\infty}^{+\infty} f_{\boldsymbol{y} \to \boldsymbol{x}}(\lambda) d\lambda. \tag{41}$$

We will use it as a definition of the counterpart measure when only one of the measures is defined. For a couple of measures $F_{x,y}$ and $f_{x,y}$, with parameter vector x, y we say that they are G-counterparts or G-inverse (G for Geweke) over domain $D \boxtimes \mathbb{R}$ ($\mathbb{R}$ is the set of real numbers), if they satisfy

$$F_{xy} = c \int\limits_{D} f_{x,y}(\lambda) d\lambda = c \int\limits_{\mathbb{R}} \chi_D(\lambda) f_{x,y}(\lambda) d\lambda, \tag{42}$$

where χ_D is the characteristic function of the set D, which is slightly more general than (41), omitting D when $D = \mathbb{R}$. Thus, if one of the counterparts is given, the other can be calculated using (42). Specially, substituting DTF_{ij} and PDC_{ij} for f in (41)/(42), we obtain their proper (time domain) counterparts, the G-inverses, which we designate by $\textbf{DTF}_{ij}$ and $\textbf{PDC}_{ij}$. This would complement the claim of Baccala – Sameshima that PDC is the proper inverse of Granger causality measure, with the proper solution for proper inverse. The G-inverse can be defined more generally, i.e. relative more general aggregation operator than integral form present in (42), thus realizing G-inverse relative arbitrary aggregation operator.

References
1. Kroger JK, Elliott L, Wong TN, Lakey J, Dang H, George J. Detecting mental commands in high frequency EEG: Faster brain-machine interfaces. In: Proc. of the 2006 Biomedical Engineering Society Annual Meeting, Chicago, 2006.
2. Watkins C, Kroger J, Kwong N, Elliott L, George J. Exploring high-frequency EEG as a faster medium of brain-machine communication. In: Proceedings Institute of Biological Engineering 2006 Annual Meeting, Tucson.
3. Jovanović A, Perović A. Brain computer interfaces - some technical remarks. Int J Bioelectromagn. 2007;9(3):191–203. http://www.ijbem.org/volume9/number3/090311.pdf.

4. Jovanovic A, Perovic A, Borovcanin M. Brain connectivity measures: computations and comparisons. EPJ Nonlinear Biomed Phys. 2013;1:2. www.epjnonlinearbiomedphys.com/content/1/1/2.
5. Perović A, Dordević Z, Paškota M, Takači A, Jovanović A. Automatic recognition of features in spectrograms based on some image analysis methods. Acta Polytechnica Hungarica. 2013;10:2.
6. Jovanovic A, Kasum O, Peric N, Perovic A. Enhancing microscopic imaging for better object and structural detection, insight and classification. In: Mendez-Vilas A, editor. Microscopy: advances in scientific research and education, FORMATEX Microscopy series N6, vol. 2. 2014.
7. Liu L, Arfanakis K, Ioannides A. Visual field influences functional connectivity pattern in a face affect recognition task. Int J Bioelectromagnetism. 2007;9:4.
8. Aoyama A, Honda S, Takeda T. Magnetoencephalographic study of auditory feature analysis associated with visually based prediction. Int J Bioelectromagn. 2009;11(3):144–8. http://www.ijbem.org/volume11/number3/1103008.pdf.
9. Grierson M. Composing with brainwaves: Minimal trial P300b recognition as an indication of subjective preference for the control of a musical instrument. Proceedings of the ICMC, Belfast. 2008.
10. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica. 1969;37:424.
11. Granger CWJ. Testing for causality: a personal viewpoint. J Econ Dyn Contr. 1980;2:329.
12. Granger CWJ, Morris MJ. Time series modelling and interpretation. J R Stat Soc Ser A. 1976;139:246.
13. Geweke J. Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc. 1982;77:304.
14. Geweke J. Measures of conditional linear dependence and feedback between time series. J Am Stat Assoc. 1984;79:907.
15. Kaminski M, Blinowska K. A new method of the description of the information flow in the brain structures. Biol Cybern. 1991;65:203.
16. Kaminski M, Ding M, Truccolo W, Bressler S. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern. 2001;85:145.
17. Sameshima K, Baccala LA. Using partial directed coherence to describe a neuronal assembly interactions. J Neurosci Meth. 1999;94:93.
18. Baccala L, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern. 2001;84:463.
19. Chen Y, Bressler SL, Ding M. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. J Neurosci Methods. 2006;150:228.
20. Schelter B, Winterhalder M, Eichler M, Peifer M, Hellwig B, Guschlbauer B, et al. Testing for directed influences among neural signals using partial directed coherence. J Neurosci Meth. 2005;152:210.
21. Takahashi DS, Baccala LA, Sameshima K. Information theoretic interpretation of frequency domain connectivity measures. Biol Cybern. 2010;103:463–9. 28.
22. Shi J, Tomasi C. Good Features to Track, preprint, www.ai.mit.edu/courses/6.891/handouts/shi94good.pdf.
23. Welch G, Bishop G. An Introduction to the Kalman Filter. Chapter Hill: University of North Carolina at Chapter Hill; 2004. TR 95-014, April 5.
24. Baccala L, Sameshima K. Overcoming the limitations of correlation analysis for many simultaneously processed neural structures, Chapter 3, M.A.L. In: Nicolelis ed. Progress in brain research, vol 130, Elsevier Sc; 2001. p. 1–15.
25. Takahashi DY, Baccalá LA, Sameshima K. Partial directed coherence asymptotics for VAR processes of infinite order. Int J Bioelectromagn. 2008;10(1):31–6. http://www.ijbem.org/volume10/number1/100105.pdf.
26. Blinowska K. Review of the methods of determination of directed connectivity from multichannel data. Med Biol Eng Comput. 2011;49:521. doi: 10.1007/s11517-011-0739-x.
27. Blinowska K, Kus R, Kaminski M, Janiszewska J. Transmission of brain activity during cognitive task. Brain Topogr. 2010;23:205. doi: 10.1007/s10548-010-0137-y.
28. Brzezicka A, Kaminski M, Kaminski J, Blinowska K. Information transfer during a transitive reasoning task. Brain Topogr. 2011;24:1. doi: 10.1007/s10548-010-0158-6.
29. Kus R, Blinowska K, Kaminski M, Basinska-Starzycka A. Transmission of information during continuous attention test. Acta Neurobiol Exp. 2008;68:103.
30. Blinowska K. Methods for localization of time-frequency specific activity and estimation of information transfer in brain. Int J Bioelectromagn. 2008;10(1):2–16. www.ijbem.org.
31. Dhamala M, Rangarajan G, Ding M. Estimating Granger causality from Fourier and wavelet transforms of time series data. Phys Rev Lett. 2008;100(018701):1.
32. Singh H, Li Q, Hines E, Stocks N. Classification and feature extraction strategies for multi channel multi trial BCI data. Int J Bioelectromagn. 2007;9(4):233. www.ijbem.org.