A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

OFFPRINT

Overarching framework for data-based modelling

BJÖRN SCHELTER, MALENKA MADER, WOLFGANG MADER, LINDA SOMMERLADE, BETTINA PLATT, YING-CHENG LAI, CELSO GREBOGI and MARCO THIEL

EPL, **105** (2014) 30004

epl

A Letters Journal Exploring the Frontiers of Physics

AN INVITATION TO SUBMIT YOUR WORK

www.epljournal.org

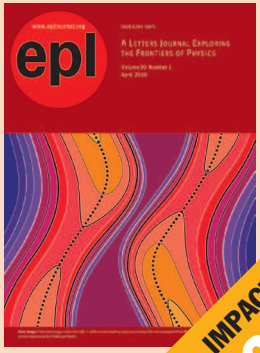
The Editorial Board invites you to submit your letters to EPL

EPL is a leading international journal publishing original, high-quality Letters in all areas of physics, ranging from condensed matter topics and interdisciplinary research to astrophysics, geophysics, plasma and fusion sciences, including those with application potential.

The high profile of the journal combined with the excellent scientific quality of the articles continue to ensure EPL is an essential resource for its worldwide audience. EPL offers authors global visibility and a great opportunity to share their work with others across the whole of the physics community.

Run by active scientists, for scientists

EPL is reviewed by scientists for scientists, to serve and support the international scientific community. The Editorial Board is a team of active research scientists with an expert understanding of the needs of both authors and researchers.



IMPACT FACTOR 2.753*

*As ranked by ISI 2010

IMPACT FACTOR
2.753*
* As listed in the ISI® 2010 Science Citation Index Journal Citation Reports

OVER
500 000
full text downloads in 2010

30 DAYS
average receipt to online publication in 2010

16 961
citations in 2010
37% increase from 2007

"We've had a very positive experience with EPL, and not only on this occasion. The fact that one can identify an appropriate editor, and the editor is an active scientist in the field, makes a huge difference."

Dr. Ivar Martinv
Los Alamos National Laboratory, USA

Six good reasons to publish with EPL

We want to work with you to help gain recognition for your high-quality work through worldwide visibility and high citations.

1. **Quality** – The 40+ Co-Editors, who are experts in their fields, oversee the entire peer-review process, from selection of the referees to making all final acceptance decisions
2. **Impact Factor** – The 2010 Impact Factor is 2.753; your work will be in the right place to be cited by your peers
3. **Speed of processing** – We aim to provide you with a quick and efficient service; the median time from acceptance to online publication is 30 days
4. **High visibility** – All articles are free to read for 30 days from online publication date
5. **International reach** – Over 2,000 institutions have access to EPL, enabling your work to be read by your peers in 100 countries
6. **Open Access** – Articles are offered open access for a one-off author payment

Details on preparing, submitting and tracking the progress of your manuscript from submission to acceptance are available on the EPL submission website **www.epletters.net**.

If you would like further information about our author service or EPL in general, please visit **www.epljournal.org** or e-mail us at **info@epljournal.org**.

EPL is published in partnership with:

European Physical Society | Società Italiana di Fisica | EDP Sciences | IOP Publishing

A Letters Journal Exploring the Frontiers of Physics

February 2014

EPL, **105** (2014) 30004
doi: 10.1209/0295-5075/105/30004

www.epljournal.org

Overarching framework for data-based modelling

Björn Schelter[1,2], Malenka Mader[2,3], Wolfgang Mader[2,3], Linda Sommerlade[1,2,4], Bettina Platt[5], Ying-Cheng Lai[1,6], Celso Grebogi[1,4] and Marco Thiel[1]

[1] *ICSMB, University of Aberdeen - Aberdeen, UK*
[2] *Freiburg Center of Data Analysis and Modeling, University of Freiburg - Freiburg, Germany*
[3] *University Medical Center Freiburg - Freiburg, Germany*
[4] *Freiburg Institute for Advanced Studies, University of Freiburg - Freiburg, Germany*
[5] *IMS, University of Aberdeen - Aberdeen, UK*
[6] *School of Electrical, Computer and Energy Engineering, Arizona State University - Tempe, AZ, USA*

received 19 October 2013; accepted in final form 23 January 2014
published online 20 February 2014

PACS 05.45.-a – Nonlinear dynamics and chaos
PACS 05.10.-a – Computational methods in statistical physics and nonlinear dynamics
PACS 05.45.Tp – Time series analysis

Abstract – One of the main modelling paradigms for complex physical systems are networks. When estimating the network structure from measured signals, typically several assumptions such as stationarity are made in the estimation process. Violating these assumptions renders standard analysis techniques fruitless. We here propose a framework to estimate the network structure from measurements of arbitrary non-linear, non-stationary, stochastic processes. To this end, we propose a rigorous mathematical theory that underlies this framework. Based on this theory, we present a highly efficient algorithm and the corresponding statistics that are immediately sensibly applicable to measured signals. We demonstrate its performance in a simulation study. In experiments of transitions between vigilance stages in rodents, we infer small network structures with complex, time-dependent interactions; this suggests biomarkers for such transitions, the key to understand and diagnose numerous diseases such as dementia. We argue that the suggested framework combines features that other approaches followed so far lack.

Introduction. – Recent years have seen a large increase in the availability of data. In fact, increasing amounts of data play a key role in physics, such as for the Large Hadron Collider (CERN) and the Square Kilometre Array (South Africa), but also in other areas such as in biology, *e.g.* genomic data, and data mining in the social sciences. Dealing with these data sets efficiently determines the success of the projects. This necessity to better understand and analyse data has led to an outburst of research into advanced methods of data analysis. Especially when dealing with complex data sets these algorithms have to fulfill certain fundamental requirements: i) they need to deal with truly multivariate data, *i.e.* they must distinguish between direct and indirect influences, ii) they have to account for various concurrent noise sources, iii) they need to addresses both linear and nonlinear systems, iv) provide results for each sampling point, v) and estimate the strengths of the directed interactions. Finally, vi) they need to provide a rigorous statistical framework to allow their evaluation and vii) be numerically efficient. A multitude of algorithms has been developed to address these extremely challenging requirements, but until now none can address them simultaneously.

This is partly due to the fact, that a rigorous mathematical framework, *i.e.* a theory of a suitable highly versatile class of mathematical models to comprise all of these features, was still lacking. In this letter, we provide a mathematical theory which encompasses a model class that is versatile enough to describe systems of this general nature. We provide the mathematical framework to fit models of this class to measured data and finally present an algorithm which automises this process. This algorithm delivers a new, statistically rigorous view of a large class of physical systems, and can contribute to a substantially deeper understanding of their dynamics and interactions.

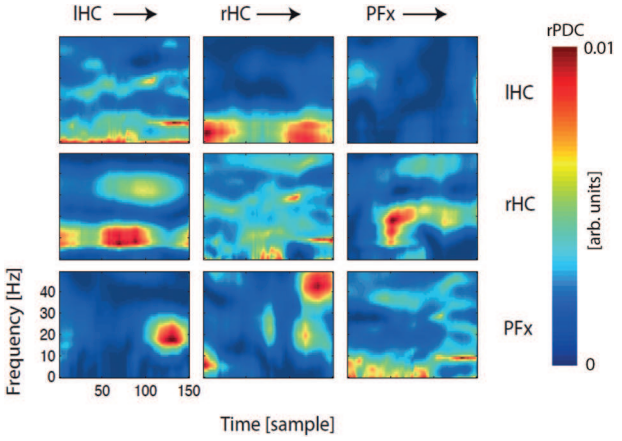


Fig. 1: (Colour on-line) rPDC analysis applied to murine EEG data recorded bilaterally above the hippocampus (left: lHC, right: rHC), and prefrontal cortex (Pfx) during a transition from REM to NREM sleep. Diagonal: logarithm of the spectra colour-coded with respect to frequency and time in samples —sampling rate 199 Hz. Off-diagonal, time-resolved rPDC analysis; colour coding indicates strength of coupling.

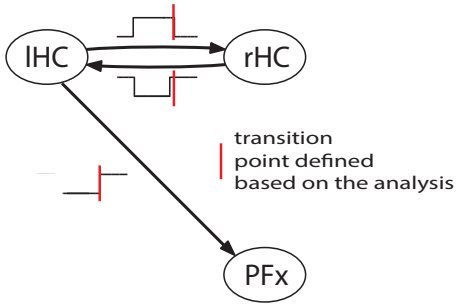


Fig. 2: (Colour on-line) Graph obtained from rPDC analysis applied to murine EEG data recorded bilaterally above the hippocampus (left: lHC, right: rHC), and prefrontal cortex (Pfx) during a transition from REM to NREM sleep. The red vertical line marks the transition point that has been determined based on the rPDC analysis with a precision of 5 ms.

Before we introduce the theory and the algorithm, we briefly demonstrate their performance analysing experimental electroencephalogram (EEG) signals, which are non-linear, non-stationary, and stochastic (fig. 1) [1]. We obtained these EEG signals from a mouse during a transition from slow-wave (NREM) to rapid eye movement (REM) sleep under freely moving conditions with a wireless device (sampling rate: 199 Hz). Electrodes were located in a prefrontal position, *i.e.* above the right prefrontal cortex, and in two parietal locations, *i.e.* above the left and right hippocampus [2,3]. Using our algorithm, that applies renormalised partial directed coherence (rPDC), as introduced below, we analyse the strengths of the directed interactions with respect to time and frequency among the brain regions. From fig. 1 the time-resolved directed network structure (fig. 2) can be inferred for each sampling point. This leads to a temporal resolution of approximately 5 ms, limited only by the sampling rate.

In the prominent alpha/beta range, *i.e.* $\sim$ 10–20 Hz a statistically significant influence from left hippocampus (lHC) to prefrontal cortex (PFx) occurs at sampling point $\sim$ 110. Interestingly, the interaction lHC to right hippocampus (rHC) drops prior to this transition, while there is an increase in the opposite direction (both statistically significant). This finding might serve as a biomarker for NREM/REM transitions[1]. Note that based on these observations alone, we can define the transition with a temporal resolution of one sampling point, which corresponds to approximately 5 ms resolution in this case (cf. network in fig. 2); to our knowledge such a resolution has not been achieved before[2]. An interpretation of the results will be given elsewhere.

Although in some applications the network structure can be determined by analysing meta information, in other applications the network structure needs to be inferred from observations of the dynamics it exhibits. There has been extensive work on exploiting multivariate, linear stochastic modelling to detect, for instance, neural interactions and to infer the structures of small-scale networks and the nodal interaction patterns [4–8]. Multivariate non-linear deterministic as well as stochastic approaches called dynamic causal modelling have been introduced [9]. Non-parametric, non-linear multivariate approaches have only recently been suggested for small networks [10]. Despite these efforts, a general framework that allows inferring the network structure from measured multivariate signals is still lacking.

In this letter, we develop a theory for the most general analysis and data-based modelling of measured signals from non-linear, non-stationary, stochastic systems. We demonstrate that these systems can be represented in a *general non-stationary state space model*. This, in turn, suggests the development of a numerically efficient algorithm based on the expectation maximisation optimisation approach applying the dual Kalman filter. Based on this algorithm we derive the frequency domain measure of renormalised partial directed coherence to quantify the direction and strength of network connections. A rigorous statistical evaluation procedure guarantees the applicability to measured signals. A simulation study motivated by typical scenarios faced in experiments complements our letter.

Methods: theory and estimation. – As the *first step*, we develop the *underlying theory* by representing non-linear systems in the framework of state space models. In experiments, a time continuous multivariate dynamical

[1]Fragmentation of the REM sleep is an indication of the presence of Alzheimer's disease.

[2]Temporal precision in a (pre-)clinical environment is typically in the order of 4–5 s.

process $Z(t)$ can only be observed as a multivariate time discrete sampled signal,

$$Y(t_i) = g(Z(t_i), v) + \eta(t_i), \tag{1}$$

where $g(\cdot)$ denotes some observation function with parameter set v; $\eta(t)$ is Gaussian-distributed independent measurement noise with a given variance. The sampling points are denoted by $t_i = i\delta t$ for the sampling time δt. Assuming a linear observation function, we obtain

$$\begin{aligned} \mathrm{d}Z(t) &= f(Z(t), w)\,\mathrm{d}t + \mathrm{d}W(t), & (2)\\ Y(t_i) &= CZ(t_i) + \eta(t_i), \quad \eta(t_i) \sim \mathcal{N}(0, \Gamma) & (3) \end{aligned}$$

for the observed dynamical system, where now C represents the linear observation matrix, $f(\cdot)$ some potentially unknown dynamical system with parameters w, Gaussian-distributed observational noise $\eta(t_i)$, and $\mathrm{d}W(t)$ the increment of a Wiener process. This model separates the observed dynamics into the hidden dynamical equation (2) and the observation equation (3).

We emphasise that the multivariate stochastic dynamical system, eq. (2), is key to our flexible overarching framework. It allows us to model systems whose dynamics is known, which is often the case in first-principle modelling in physics; it, however, also provides powerful means to model unknown dynamics in data-based modelling such as brain dynamics. It also accounts for exogenous input; to make this more explicit we could use

$$\mathrm{d}Z(t) = f(Z(t), U(t), w)\mathrm{d}t + \mathrm{d}W(t), \tag{4}$$

in which $U(t)$ would model the exogenous input. This modification is not needed as $U(t)$ can always be accounted for by augmenting $Z(t)$. By adjusting $W(t)$, our flexible framework incorporates stochastic as well as deterministic, linear as well as non-linear, stationary as well as non-stationary dynamics. Models for effective connectivity such as deterministic or stochastic dynamic causal modelling [9] would thus be a special case of our general framework. To keep the calculations clearer we assumed a linear observation function. The framework is, however, applicable to non-linear observation functions as well.

Formally, eq. (2) can be solved by integration,

$$Z(t+\Delta t) = Z(t) + \underbrace{\int_t^{t+\Delta t} f(Z(\tau), w)\mathrm{d}\tau}_{=B(t)Z(t)} + \int_t^{t+\Delta t} \mathrm{d}W(\tau)$$

with some function $B(t)$ that will be determined below. The Ito integral

$$\int_t^{t+\Delta t} \mathrm{d}W(\tau) = \sqrt{\Delta t}\epsilon(t), \quad \epsilon(t) \sim \mathcal{N}(0, \Sigma) \tag{5}$$

further simplifies the hidden dynamical equation (2). Note that the integration time Δt and the sampling time δt do not need to be identical. In the following we assume, however, that $\Delta t = \delta t$. The more general case in which $n\Delta t = \delta t$ with $n \in \mathbb{N}$ being a positive integer can be coped with by modifying the observation equation accordingly; an observation is only made every $n\Delta t$ integration steps [11].

Thus, we obtain the following model:

$$\begin{aligned} Z(i) &= A(i)Z(i-1) + \xi(i), \quad \xi(i) \sim \mathcal{N}(0, \tilde{\Sigma}), & (6)\\ Y(i) &= CZ(i) + \rho(i), \quad \rho(i) \sim \mathcal{N}(0, \tilde{\Gamma}) & (7) \end{aligned}$$

for some appropriately chosen variances $\tilde{\Sigma}$ and $\tilde{\Gamma}$ that will be optimally determined in the estimation process. Time $t = i\delta t = i\Delta t$ is characterised by its non-negative integer time i in these equations. The time-dependent $A(i)$ is equal to $\mathbf{1}_n + B(i)$, where $\mathbf{1}_n$ denotes the n-dimensional identity matrix.

Determining the matrix $A(i)$ from measured signals in the most general case without any further assumptions is not possible. A reasonable assumption regarding a sensible integration time Δt, however, is that the parameter matrix $A(i)$ should change more slowly than the (stochastic) dynamics itself. This ensures a separation of time scales. For such a case, we can augment the model to the over-arching *state space model*,

$$\begin{aligned} a(i) &= a(i-1) + \zeta(i), \quad \zeta(i) \sim \mathcal{N}(0, \tilde{\Omega}), & (8)\\ Z(i) &= A(i)Z(i-1) + \xi(i), \quad \xi(i) \sim \mathcal{N}(0, \tilde{\Sigma}), & (9)\\ Y(i) &= CZ(i) + \rho(i), \quad \rho(i) \sim \mathcal{N}(0, \tilde{\Gamma}). & (10) \end{aligned}$$

that will reflect the subsequent analyses. The $a(i)$ are the matrix entries of $A(i)$ rearranged into a vector. Although the augmented parameter equation for $a(i)$ (eq. (8)) appears to describe a non-stationary process, it is only a restriction on the parameter variations; the time discrete version of a bounded derivative

$$a(i) - a(i-1) = \zeta(i), \tag{11}$$

where the boundedness originates from the finite variance of the stochastic variable $\zeta(i)$. Large parameter changes are, thus, extremely unlikely.

The extent to which the past values influence the present is determined by the parameters $a(i)$ or the matrix $A(i)$, respectively. These *causal* influences can be represented as directed edges in a network, in which the nodes represent the processes. In this sense the matrix $A(i)$ contains the interactions between the components of the original process $Z(t)$ modelled via $f(Z(t), w)$, the information about the network structure is contained in this matrix as well. As we do not make any assumptions about the origin of $Z(t)$, it can model the sensor as well as the source space equally well.

In several applications, the dynamics $f(Z(t), w)$ actually depends on previous time steps as well. In networks, influences with a certain delay are typically relevant. This can be accounted for in state space modelling by including

higher orders, *i.e.* including previous time steps in eq. (9),

$$Z(i) = \left[\sum_{i'=1}^{p} A_{i'}(i) Z(i-i')\right] + \xi(i), \quad \xi(i) \sim \mathcal{N}(0, \tilde{\Sigma}), \tag{12}$$

up to some maximum time lag p. This maximum lag p can be determined relying on *a priori* knowledge or based on model selection criteria such as Akaike's information criterion. Rewriting this higher-order process as a first-order process by introducing $\hat{Z}(i) = (Z(i), Z(i-1), \ldots, Z(i-p+1))'$ leads to

$$\hat{Z}(i) = \hat{A}(i)\hat{Z}(i-1) + \hat{\xi}(i), \qquad \hat{\xi}(i) \sim \mathcal{N}(0, \hat{\Sigma}). \tag{13}$$

The matrix $\hat{A}(i)$ thereby assumes a specific structure:

$$\hat{A}(i) = \begin{pmatrix} A_1(i) & A_2(i) & \ldots & A_p(i) \\ \mathbf{1}_n & \mathbf{0}_n & \ldots & \mathbf{0}_n \\ \mathbf{0}_n & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{14}$$

Replacing eq. (9) by eq. (13) leaves the state space model unaltered.

As the *second step*, we present the *numerically efficient algorithm* to determine the parameters in the state space model. In the above form of eqs. (8)–(10) the trajectory of the $Z(i)$ and the $a(i)$ can be determined purely based on the observations $Y(i)$. To this end, the so-called extended or unscented Kalman filters, here optimally the dual Kalman filter can be proven to be highly efficient estimators [12–14]. These filters, however, rely on the knowledge of the exact dynamics, *i.e.* on a precise knowledge of C, $\tilde{\Omega}$, $\tilde{\Sigma}$, and $\tilde{\Gamma}$. As these are typically not known, they need to be estimated. The *expectation maximisation algorithm* addresses this challenge [3,13,15]. Roughly speaking, it is a recursive maximum likelihood estimator that approximates the incomplete data likelihood applying the dual Kalman filter. In the estimation step the likelihood is calculated based on a (potentially wrong) assumption about the dynamics. The maximisation step optimises the parameters C, $\tilde{\Omega}$, $\tilde{\Sigma}$, and $\tilde{\Gamma}$; thereby, yielding a better approximation of the dynamics. Applying these steps iteratively ensures convergence to the best estimator of the underlying dynamical process $Z(i)$ as well as the parameters $a(i)$; "best estimator" here refers to the one that minimises the squared distance between the true values and their estimates [14].

It is worth noting that the estimation procedure is optimal under certain assumptions such as the Gaussianity of the noise. If such assumptions are violated, better estimators might exist. We refer the reader to the extensive literature for such modifications, *e.g.* [13]. The numerical complexity of the algorithm enables a sensible application in real time on standard laptop computers for a reasonably small number of nodes of the network.

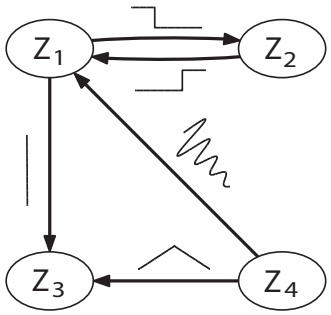


Fig. 3: Network structure as simulated based on eq. (16). The interaction between processes Z_1 and Z_2 is switched, it is constant from Z_1 to Z_3, triangularly shaped for Z_4 to Z_1, and an exponentially decaying oscillation from Z_4 to Z_1, cf. footnote [3].

As the *third step*, we derive the *renormalised partial directed coherence* (rPDC) [16],

$$\lambda_{u \leftarrow v}(\omega) = X_{uv}(\omega)' \left(V_{uv}(\omega)\right)^{-1} X_{uv}(\omega), \tag{15}$$

a frequency domain measure for Granger causality (see [16] and references therein) that quantifies the direction and strength of network connections. It is the normalised squared Fourier transform of the (u, v) component of parameter matrices $A_{i'}(i)$, $X_{uv}(\omega) = \left[\mathcal{R}\left(\mathcal{FT}(A_{i',uv})\right), \mathcal{I}\left(\mathcal{FT}(A_{i',uv})\right)\right]'$ with $\mathcal{R}$ the real part and $\mathcal{I}$ the imaginary part, and N the number of data points analysed. The normalisation by $\left(V_{uv}(\omega)\right)^{-1}$ is given by the inverse of N-times the covariance matrix

$$V_{uv}(\omega) = \sum_{i'_1, i'_2 = 1}^{p} \operatorname{cov}\left(A_{i'_1,uv}\,,\, A_{i'_2,uv}\right) \cdot \begin{pmatrix} \cos\omega i'_1 \cos\omega i'_2 & \cos\omega i'_1 \sin\omega i'_2 \\ \sin\omega i'_1 \cos\omega i'_2 & \sin\omega i'_1 \sin\omega i'_2 \end{pmatrix}$$

of the estimates of X. The covariance matrix of the estimated parameters $\operatorname{cov}\left(A_{i'_1,uv}, A_{i'_2,uv}\right)$ is determined in the expectation maximisation algorithm [17].

As the *fourth step*, we introduce the *rigorous statistical evaluation procedure*. We propose a bootstrap-based framework [17]:

1) Estimate the time-dependent parameter matrix $\hat{A}(i)$ by the state space model, eqs. (8)–(10). Then, estimate the rPDC $\hat{\lambda}$ based on the parameters according to eq. (15).

2) Using $\hat{A}(i)$ and the estimated parameters C, $\tilde{\Omega}$, $\tilde{\Sigma}$, and $\tilde{\Gamma}$ of the dynamics, generate m parametric bootstrap realisations of the process according to the non-stationary state space model in eqs. (8)–(10), *i.e.* simulate the model using the estimated parameters.

3) Analogously to step 1, estimate the parameter matrices $\{\widehat{A}^r(i)\}_{r=1,\ldots,m}$ for all m bootstrap realisations. Estimate the rPDC of the m bootstrap realisations, $\{\hat{\lambda}_r\}_{r=1,\ldots,m}$, for each link and at each frequency of interest, respectively.

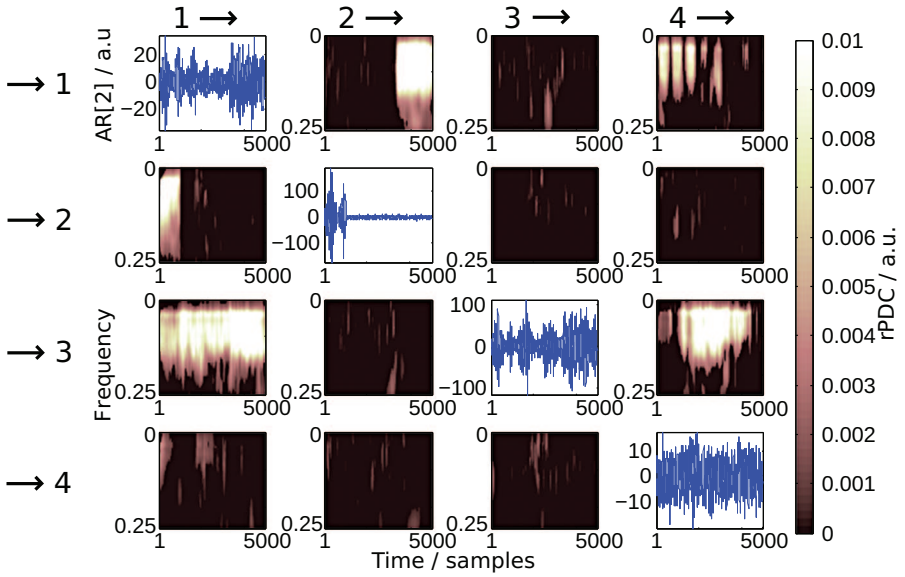


Fig. 4: (Colour on-line) Time-resolved renormalised partial directed coherence (rPDC) of the four-dimensional non-linear non-stationary system of eq. (16). For this example we used the parameters given in footnote [3]. It was estimated assuming a model order $p = 10$ and dimension $n = 4$. Non-significant (5%-significance level) values are in black. On the diagonal the raw data are shown for each node. Off-diagonal plots show time-resolved rPDC values colour-coded for frequencies between 0 and 0.25 (in inverse sampling time) and for all 5000 data points. Switching behaviour in the coupling ($2 \to 1$ and $1 \to 2$) as well as gradually changing coupling strengths ($4 \to 1$, an exponentially decaying sinusoidal, and $4 \to 3$, with increasing and after 2500 samples decreasing coupling strength) are revealed by the rPDC. The corresponding graph is shown in fig. 3.

4) Calculate the standard error σ of bootstrap rPDC-values $\{\hat{\lambda}_r\}_{r=1,\ldots,m}$. This is a measure of natural fluctuation of the rPDC-values of the original system with parameters obtained in step 1).

5) Determine the confidence intervals around the estimated rPDC values $\hat{\lambda}$ of the original process as the Gaussian quantiles at a given confidence level α, *i.e.* $[\hat{\lambda} - k_\alpha\sigma, \hat{\lambda} + k_\alpha\sigma]$ with k_α being the α-quantiles of the standard Gaussian distribution. For $\alpha = 0.95$ this yields $[\hat{\lambda} - 1.96\sigma, \hat{\lambda} + 1.96\sigma]$. We note here that the assumption of a Gaussian distribution for the rPDC value is an approximation [16]; if needed the exact distribution is a possibly non-central χ^2-distribution [16].

6) If the confidence interval includes 0, the rPDC values are considered to be non-significant.

We emphasise that the framework we introduced above does not make any assumptions about the dimensionality of the problem. We restrict ourselves to a 4-dimensional model in the following simulation study as an illustrative example.

Simulation study. – We now perform a *simulation study* to demonstrate and illustrate the performance of our general framework, as applied to murine vigilance stages (fig. 1). To this end, we simulate an $n = 4$ dimensional second-order non-stationary system given by (cf. eq. (12))

$$Z(i) = \sum_{i'=1}^{2} A_{i'}(i)Z(i-i') + \xi(i), \quad \xi(i) \sim \mathcal{N}(0, \mathbf{1}_4) \quad (16)$$

with additional Gaussian observational noise with variance 1; the parameters and network structure are depicted in fig. 3 and are given in footnote[3]. Parameters are estimated based on the expectation maximisation algorithm applying the dual Kalman filter (eqs. (8)–(10)). We then estimate the renormalised partial directed coherence according to eq. (15); rPDC is shown in fig. 4 at frequencies ω between 0 and 0.25 (in inverse sampling time) and for all 5000 simulated data points. We perform statistical analysis as described above. Non-significant values are black. The estimated network structure solely based on the measurements reflects the true simulated network structure (fig. 3). The statistical analysis based on bootstrapping enables us to infer the correct network structure for various time-dependent coupling strengths and various signal-to-observational noise ratios.

We emphasise that our general framework is not restricted to situations as presented in the above example. It can be applied to non-linear, non-stationary, noisy systems in general, see also the "Remarks" section below.

[3]Parameters for the example: $A_{1,11}(i) = 1.3$, $A_{1,22}(i) = 1.6$, $A_{1,33}(i) = 1.5$, $A_{1,44}(i) = 1.7$, $A_2(i) = -0.8 \cdot \mathbf{1}_4$, $A_{1,12}(i) = 0.7$ if $i > 3333$, $A_{1,21}(i) = 0.7$ if $i \leq 1000$, $A_{1,14}(i) = \exp(-i/2500) \cdot \sin(0.005i)$, $A_{1,31}(i) = 0.5$, $A_{1,34}(i) = 0.8i/2500$ if $i \leq 2500$, $A_{1,34}(i) = 2 - i/2500$ if $i > 2500$, and $A_{i',kl}(i) = 0$ for all remaining parameters and the remaining times, respectively.

The EEG example of fig. 1 demonstrates the performance of the approach to non-linear, non-stationary, stochastic systems. In a forthcoming study we will investigate the robustness of results we identified in this experiment. We will investigate its role for diseases such as dementia.

Conclusion. – We have shown that a wide class of models that generate time series can be interpreted as a non-linear, non-stationary state space model. This led to an overarching framework for data-based modelling that for the first time allows us i) to deal with truly multivariate data, *i.e.* it distinguishes between direct and indirect influences, ii) to account for various concurrent noise sources, iii) to address both linear and non-linear systems, iv) to provide results for each sampling point, v) to estimate the strengths of the directed interactions, vi) to provide a rigorous statistical framework, and vii) to analyse complex systems numerically efficiently. As demonstrated numerically, novel insights into processes can be gained through this method. We particularly emphasise that this overarching framework can provide justification for several analyses that have been preformed retrospectively; the theory can be extended as well. We discuss this briefly in the following remarks as a detailed description would be beyond the scope of this manuscript.

Remarks. – a) In several publications including [16], it has been shown numerically that ordinary, stationary, linear Granger-causality–based approaches can elucidate the network structure for non-linear systems such as the Rössler system or the stochastic van der Pol system. The, so far, missing mathematical theory underlying these observations is included in the theory introduced in this manuscript. Averaging the time-resolved coefficients estimated in the state space model yields the stationary coefficients that characterise ordinary Granger-causality analysis.

b) By fitting a time varying autoregressive model of higher orders to the signals, the spectral content of the signals can be obtained as well. This enables tracking the dominant frequencies over time.

c) In the derivations above, we limited ourselves to a specific observation function. This is not a limitation of the technique *per se* as the overarching framework can be adapted to more general observation functions; we used the linear observation function here to keep the notations simpler.

d) State space modelling is robust against few isolated outliers in the data. This is due to the robustness of the Kalman filter. This robustness can further be improved by applying the Huber norm rather than the L2 norm in the estimation process.

e) For larger networks the time-resolved rPDC approach suffers from the challenge that too many parameters are non-zero, though non-significant inherent in maximum likelihood L2-based estimation; this could potentially be overcome by L1 regularisation as, for instance, in the least absolute shrinkage and selection operator (LASSO) [18] estimation.

f) Compressive sensing [19] opens the possiblity to extend the time-resolved rPDC approach to sums of various non-linear functions of $Z(t)$, *i.e.* explicitly accounting for non-linearities in the dynamics. This becomes particularly relevant in cases in which the approximation that the $a(i)$ in eq. (8) change on a slow time scale is invalid.

REFERENCES

[1] Celka P. and Colditz P., *IEEE Biomed. Eng.*, **49** (2002) 556.

[2] Jyoti A., Plano A., Riedel G. and Platt B., *J. Alzheimer's Dis.*, **22** (2010) 873.

[3] Sommerlade L., Thiel M., Platt B., Plano A., Riedel G., Grebogi C., Timmer J. and Schelter B., *J. Neurosci. Methods*, **203** (2012) 173.

[4] Kamiński M., Ding M., Truccolo W. A. and Bressler S. L., *Biol. Cybern.*, **85** (2001) 145.

[5] Zhu L., Lai Y.-C., Hoppensteadt F. and He J., *Neural Comput.*, **15** (2003) 2359.

[6] Zhu L., Lai Y.-C., Hoppensteadt F. and He J., *Math. Biosci. Eng.*, **2** (2005) 1.

[7] Wang X., Chen Y.-H., Bressler S. L. and Ding M., *Int. J. Neural Syst.*, **17** (2007) 71.

[8] Wen X.-T., Yao L., Liu Y.-J. and Ding M., *J. Neurosci.*, **32** (2012) 1284.

[9] Daunizeau J., Stephan K. E. and Friston K. J., *Neuroimage*, **62** (2012) 464.

[10] Nawrath J., Romano M., Thiel M., Kiss I., Wickramasinghe M., Timmer J., Kurths J. and Schelter B., *Phys. Rev. Lett.*, **104** (2010) 038701.

[11] Lenz M., Musso M., Linke Y., Tüscher O., Timmer J., Weiller C. and Schelter B., *Physiol. Meas.*, **32** (2011) 1752.

[12] Honerkamp J., *Stochastic Dynamical Systems* (VCH, New York) 1993.

[13] Shumway R. and Stoffer D., *Time Series Analysis and its Application* (Springer, New York) 2000.

[14] Honerkamp J., *Statistical Physics,* 3rd edition (Springer) 2012.

[15] Killmann M., Sommerlade L., Mader W., Timmer J. and Schelter B., *Biomed. Tech. (Berl.)*, **57** (2012) 387.

[16] Schelter B., Timmer J. and Eichler M., *J. Neurosci. Methods*, **179** (2009) 121.

[17] Efron B. and Tibshirani R., *An Introduction to the Bootstrap* (Chapman & Hall, New York) 1998.

[18] Tibshirani R., *J. R. Stat. Soc. B*, **58** (1996) 267.

[19] Wang W.-X., Lai Y.-C., Grebogi C. and Ye J., *Phys. Rev. X*, **1** (2011) 021021.