

Parameter estimation for correlated Ornstein-Uhlenbeck time-series

Helmut H. Strey

*Biomedical Engineering Department and Laufer Center for Physical and Quantitative Biology,
Stony Brook University, Stony Brook NY 11794-5281.*

Rajat Kumar

Biomedical Engineering Department, Stony Brook University, Stony Brook NY 11794-5281.

Lilianne Mujica-Parodi

*Biomedical Engineering Department and Laufer Center for Physical and Quantitative Biology,
Stony Brook University, Stony Brook NY 11794-5281. and
Martinus Center, Massachusetts General Hospital and Harvard University, Boston, MA*

(Dated: February 12, 2021)

In this article, we develop a Maximum likelihood (ML) approach to estimate parameters from correlated time traces that originate from coupled Ornstein-Uhlenbeck processes. The most common technique to characterize the correlation between time-series is to calculate the Pearson correlation coefficient. Here we show that for time series with memory (or a characteristic relaxation time), our method gives more reliable results, but also results in coupling coefficients and their uncertainties given the data. We investigate how these uncertainties depend on the number of samples, the relaxation times and sampling time. To validate our analytic results, we performed simulations over a wide range of correlation coefficients both using our maximum likelihood solutions and Markov-Chain Monte-Carlo (MCMC) simulations. We found that both ML and MCMC result in the same parameter estimations. We also found that when analyzing the same data, the ML and MCMC uncertainties are strongly correlated, while ML underestimates the uncertainties by a factor of 1.5 to 3 over a large range of parameters. For large datasets, we can therefore use the less computationally expensive maximum likelihood method to run over the whole dataset, and then we can use MCMC on a few samples to determine the factor by which the ML method underestimates the uncertainties. To illustrate the application of our method, we apply it to time series of brain activation using fMRI measurements of the human default mode network. We show that our method significantly improves the interpretation of multi-subject measurements of correlations between brain regions by providing parameter confidence intervals for individual measurements, which allows for distinguishing between the variance from differences between subjects from variance due to measurement error.

A. Introduction

In many fields of science one is interested in whether two sets of data are linearly correlated. One of the simplest statistical tool for such a case is to calculate the Pearson correlation coefficient ρ . By calculating the Pearson correlation coefficient we assume that the samples are identically and independently drawn from a bivariate Gaussian distribution. Unfortunately, the Pearson coefficient is frequently used for data that falls outside this condition. In particular, when applying Pearson correlations to time-series that may be distributed normally but where adjacent time points are not independent. In this article, we are developing a method to characterize correlations between two time-series that exhibit short-term memory or relaxation. In particular, we will restrict ourselves to the simplest random process that results in a fluctuating time series with a characteristic relaxation time: the Ornstein-Uhlenbeck (OU) process [1]. A OU process with mean zero can be expressed by the following Langevin equation:

$$\dot{x} = -\frac{k}{\gamma}x + \frac{1}{\gamma}\xi(t) \quad (1)$$

where k is the spring constant, γ is the friction coefficient and $\xi(t)$ is a randomly fluctuating force. We and others have recently published maximum likelihood methods to estimate parameters from an OU process time series [2–4], and here we attempt the same for coupled OU processes.

B. Coupled Oscillators

We consider a system of two overdamped oscillators, characterized by a friction coefficient γ and spring constant k , that are coupled by a spring with a spring constant c . The Langevin equations can be written in the following way:

$$\begin{aligned} \dot{x}_1 &= -\frac{k}{\gamma}x_1 - \frac{c}{\gamma}(x_1 - x_2) + \frac{1}{\gamma}\xi_1(t) \\ \dot{x}_2 &= -\frac{k}{\gamma}x_2 - \frac{c}{\gamma}(x_2 - x_1) + \frac{1}{\gamma}\xi_2(t) \end{aligned} \quad (2)$$

This system can be decoupled by the two eigenfunctions $2y_1 = x_1 + x_2$ and $2y_2 = x_1 - x_2$, which results in

$$\begin{aligned} \dot{y}_1 &= -\frac{k}{\gamma}y_1 + \frac{1}{\gamma}\xi_3(t) \\ \dot{y}_2 &= -\frac{k+c}{\gamma}y_2 + \frac{1}{\gamma}\xi_4(t) \end{aligned} \quad (3)$$

where $y_1(t)$ and $y_2(t)$ are two Ornstein-Uhlenbeck processes with different relaxation rates. The equipartition theorem yields $\langle y_1^2 \rangle = A_1 = \frac{k_B T}{k}$ and $\langle y_2^2 \rangle = A_2 = \frac{k_B T}{k+c}$. The other important relationships are: diffusion coefficient $D = \frac{k_B T}{\gamma}$, relaxation times $\tau_1 = \frac{A_1}{D}$ and $\tau_2 = \frac{A_2}{D}$. If the friction coefficients γ are different for x_1 and x_2 one can still decouple the system of differential equations by finding the eigenvectors and eigenvalues [5]. For simplicity we assume that both processes have the same friction coefficient or diffusion coefficient D .

In many fields, the Pearson Correlation Coefficient ρ is used to characterize the correlation between two time-series. The connection between our coupling coefficient c and ρ is:

$$\rho = \frac{\langle x_1 x_2 \rangle}{\sigma_1 \sigma_2} = \frac{\langle (y_1 + y_2)(y_1 - y_2) \rangle}{\sqrt{\langle (y_1 + y_2)^2 \rangle \langle (y_1 - y_2)^2 \rangle}} = \frac{\langle y_1^2 \rangle - \langle y_2^2 \rangle}{\langle y_1^2 \rangle + \langle y_2^2 \rangle} = \frac{c}{2k + c} \quad (4)$$

As expected, for $c = 0$ we obtain $\rho = 0$, and $c = \infty$ results in $\rho = 1$. Often correlation analysis are performed on normalized samples. It therefore makes sense to define a unitless coupling coefficient $C = c/k$ which can be calculated by estimating A_1 and A_2 from the data:

$$C = \frac{A_1 - A_2}{A_2} = \frac{c}{k} \quad (5)$$

We can extend this procedure to negative coupling coefficients. For example, $c = -k$ would result in $\rho = -1$, but at this c (see eq. 3) y_2 becomes unbound. In order to map the normalized coupling coefficient C in the range from $-\infty$ to $+\infty$ to the Pearson correlation coefficient $\rho = -1$ to $\rho = 1$ we suggest the following procedure: From two normalized samples x_1 and x_2 , we calculate $2y_1 = x_1 + x_2$ and $2y_2 = x_1 - x_2$. From y_1 and y_2 , we estimate the amplitudes A_1 and A_2 . If $A_1 > A_2$ then the coupling coefficient $C > 0$ and can be calculated using $C = (A_1 - A_2)/A_2$. If on the other hand $A_1 < A_2$ then the coupling coefficient is negative and $C = (A_1 - A_2)/A_1$. Reversing the process, we can use this procedure to create artificial time-series, by simulating two independent OU processes (eq. 3) to produce two coupled OU processes $x_1 = y_1 + y_2$ and $x_2 = y_1 - y_2$ with a well-defined positive coupling C or ρ . One obtains negative coupling coefficients by using $x_2 = y_2 - y_1$.

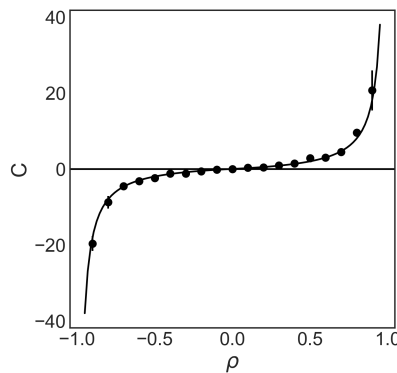


FIG. 1: Relationship between coupling coefficient C and Pearson correlation coefficient ρ . The solid line represents eq. 5, whereas the data points represent the parameter estimation by MCMC simulation on artificial data of length $N = 1000$. The errorbars indicate the distribution of estimated C s.

C. Maximum likelihood solution

In this section we will express the likelihood (see for example [6]) of two correlated time-series x_1 and x_2 in terms of A_1 , A_2 and $B_k \equiv B(\Delta t, D, A_k) = \exp\left(-\frac{D}{A_k} \Delta t\right)$ assuming uniform priors for all parameters.

$$p(\{x_{1,i}(t_i)\}, \{x_{2,i}(t_i)\} | D, A_1, A_2) \propto \frac{1}{\sqrt{2\pi A_1}^N} \frac{1}{\sqrt{(1-B_1^2)}^{(N-1)}} \quad (6)$$

$$\times \exp\left(-\frac{1}{2A_1} \left(x_{1,1}^2 + \sum_{i=1}^{N-1} \frac{x_{1,i+1}^2 - 2x_{1,i+1}x_{1,i}B_1 + x_{1,i}^2B_1^2}{(1-B_1^2)}\right)\right)$$

$$\times \frac{1}{\sqrt{2\pi A_2}^N} \frac{1}{\sqrt{(1-B_2^2)}^{(N-1)}} \quad (7)$$

$$\times \exp\left(-\frac{1}{2A_2} \left(x_{2,1}^2 + \sum_{i=1}^{N-1} \frac{x_{2,i+1}^2 - 2x_{2,i+1}x_{2,i}B_2 + x_{2,i}^2B_2^2}{(1-B_2^2)}\right)\right)$$

In order to find the maximum likelihood it is convenient to take the logarithm of p and then take the derivatives with respect to A_1 , A_2 and D .

$$\begin{aligned} \Phi &= \ln(p(\{x_{1,i}(t_i)\}, \{x_{2,i}(t_i)\} | D, A_1, A_2)) \\ &= C - \frac{N}{2} \ln(A_1) - \frac{N}{2} \ln(A_2) - \frac{N-1}{2} \ln(1-B_1^2) - \frac{N-1}{2} \ln(1-B_2^2) - \frac{1}{2A_1} Q_1(B_1) - \frac{1}{2A_2} Q_2(B_2) \end{aligned} \quad (8)$$

with

$$\begin{aligned} Q_k(B_k) &= x_{k,1}^2 + \sum_{i=1}^{N-1} \frac{x_{k,i+1}^2 - 2x_{k,i+1}x_{k,i}B_k + x_{k,i}^2B_k^2}{(1-B_k^2)} \\ &= \frac{x_{k,1}^2 + x_{k,N}^2}{1-B_k^2} + \frac{1+B_k^2}{1-B_k^2} \sum_{i=2}^{N-1} x_{k,i}^2 - \frac{2B_k}{1-B_k^2} \sum_{i=1}^{N-1} x_{k,i}x_{k,i+1} \end{aligned} \quad (9)$$

with C representing an unimportant constant. $Q_k(B_k)$ reveals the fundamental statistic - the only terms that exclusively contain the data $\{x_{k,i}\}$ with $k = 1, 2$:

$$\begin{aligned} a_{k,EndPoints} &= a_{k,EP} = x_{k,1}^2 + x_{k,N}^2 \\ a_{k,SumSquared} &= a_{k,SS} = \sum_{i=2}^{N-1} x_{k,i}^2 \\ a_{k,Correlation} &= a_{k,C} = \sum_{i=1}^{N-1} x_{k,i}x_{k,i+1} \end{aligned} \quad (10)$$

We find the maximum likelihood parameters $A_{k,max}$ and D_{max} by finding the roots of the first derivatives with respect to the three parameters. The derivative with respect to A_k results in the following two conditions:

$$\left. \frac{\partial}{\partial A_k} \Phi \right|_{A_{k,max}, D_{max}} = -\frac{N}{2A_{k,max}} + \frac{1}{2A_{k,max}^2} Q(B_{k,max}) \quad (11)$$

$$+ \left. \frac{\partial B_k}{\partial A_k} \left(\frac{B_{k,max}(N-1)}{1-B_{k,max}^2} - \frac{1}{2A_{k,max}} \frac{\partial Q(B_k)}{\partial B_k} \right) \right|_{A_{k,max}, D_{max}} = 0 \quad (12)$$

with

$$\frac{\partial}{\partial B_k} Q_k(B_k) = \frac{2}{(1-B_k^2)^2} (B_k a_{k,EP} + 2B_k a_{k,SS} - (1+B_k^2) a_{k,C}) \quad (13)$$

and

$$\frac{\partial B_k}{\partial A_k} = \frac{B_k D \Delta t}{A_k^2} \quad (14)$$

The derivative with respect to D can be written:

$$\left. \frac{\partial}{\partial D} \Phi \right|_{A_{k,max}, D_{max}} = \quad (15)$$

$$\left(\frac{B_1(N-1)}{1-B_1^2} - \frac{1}{2A_1} \frac{\partial}{\partial B_1} Q_1(B_1) \right) \left. \frac{\partial B_1}{\partial D} \right|_{A_{1,max}, D_{max}} \quad (16)$$

$$+ \left(\frac{B_2(N-1)}{1-B_2^2} - \frac{1}{2A_2} \frac{\partial}{\partial B_2} Q_2(B_2) \right) \left. \frac{\partial B_2}{\partial D} \right|_{A_{2,max}, D_{max}} = 0 \quad (17)$$

with

$$\frac{\partial B_k}{\partial D} = -B_k \frac{\Delta t}{A_k} \quad (18)$$

From these three equations we need to find A_1, A_2 and D , which can be done using numerical root finding algorithms that are implemented in standard numerical libraries (e.g. Powell hybrid method implemented in MINPACK [7]). In order to estimate the parameter uncertainties, we need to calculate the Jacobian. Here the only non-zero terms are:

$$\begin{aligned} \frac{\partial^2}{\partial A_k^2} \Phi &= \frac{N}{2A_k^2} - \frac{Q_k(B_k)}{A_k^3} + \frac{N-1}{1-B_k^2} \left(B_k \frac{\partial^2 B_k}{\partial A_k^2} + \left(\frac{\partial B_k}{\partial A_k} \right)^2 \frac{1+B_k^2}{1-B_k^2} \right) - \frac{1}{2A_k} \frac{\partial^2 Q_k(B_k)}{\partial A_k^2} + \frac{1}{A_k^2} \frac{\partial Q_k(B_k)}{\partial A_k} \\ &= -\frac{N}{2A_k^2} + \frac{N-1}{1-B_k^2} \left(B_k \frac{\partial^2 B_k}{\partial A_k^2} + \left(\frac{\partial B_k}{\partial A_k} \right)^2 \frac{1+B_k^2}{1-B_k^2} + \frac{2B_k}{A_k} \frac{\partial B_k}{\partial A_k} \right) - \frac{1}{2A_k} \frac{\partial^2 Q_k(B_k)}{\partial A_k^2} \\ \frac{\partial^2}{\partial A_k \partial D} \Phi &= \frac{1}{2A_k^2} \frac{\partial Q_k(B_k)}{\partial D} - \frac{1}{2A_k} \frac{\partial^2 Q_k(B_k)}{\partial A_k \partial D} + \frac{N-1}{1-B_k^2} \left(B_k \frac{\partial^2 B_k}{\partial A_k \partial D} + \frac{\partial B_k}{\partial A_k} \frac{\partial B_k}{\partial D} \frac{1+B_k^2}{1-B_k^2} \right) \\ \frac{\partial^2}{\partial D^2} \Phi &= \frac{N-1}{1-B_1^2} \left(B_1 \frac{\partial^2 B_1}{\partial D^2} + \left(\frac{\partial B_1}{\partial D} \right)^2 \frac{1+B_1^2}{1-B_1^2} \right) + \frac{N-1}{1-B_2^2} \left(B_2 \frac{\partial^2 B_2}{\partial D^2} + \left(\frac{\partial B_2}{\partial D} \right)^2 \frac{1+B_2^2}{1-B_2^2} \right) \\ &\quad - \frac{1}{2A_1} \frac{\partial^2 Q_1(B_1)}{\partial D^2} - \frac{1}{2A_2} \frac{\partial^2 Q_2(B_2)}{\partial D^2} \end{aligned} \quad (19)$$

with

$$\begin{aligned}
 \frac{\partial^2 B_k}{\partial A_k^2} &= \frac{B_k D \Delta t}{A_k^3} \left(\frac{D \Delta t}{A_k} - 2 \right) \\
 \frac{\partial^2 B_k}{\partial A_k \partial D} &= \frac{B_k \Delta t}{A_k^2} \left(1 - \frac{D \Delta t}{A_k} \right) \\
 \frac{\partial^2 B_k}{\partial D^2} &= B_k \frac{\Delta t^2}{A_k^2} \\
 \frac{\partial^2 Q_k(B_k)}{\partial A_k \partial D} &= \frac{\partial^2 Q_k(B_k)}{\partial B_k^2} \frac{\partial B_k}{\partial D} \frac{\partial B_k}{\partial A_k} + \frac{\partial Q_k(B_k)}{\partial B_k} \frac{\partial^2 B_k}{\partial A_k \partial D} \\
 \frac{\partial^2 Q_k(B_k)}{\partial D^2} &= \frac{\partial^2 Q_k(B_k)}{\partial B_k^2} \left(\frac{\partial B_k}{\partial D} \right)^2 + \frac{\partial Q_k(B_k)}{\partial B_k} \frac{\partial^2 B_k}{\partial D^2} \\
 \frac{\partial^2 Q_k(B_k)}{\partial A_k^2} &= \frac{\partial^2 Q_k(B_k)}{\partial B_k^2} \left(\frac{\partial B_k}{\partial A_k} \right)^2 + \frac{\partial Q_k(B_k)}{\partial B_k} \frac{\partial^2 B_k}{\partial A_k^2} \\
 \frac{\partial^2 Q_k(B_k)}{\partial B_k^2} &= \frac{6B_k + 2}{(1 - B_k^2)^3} (a_{k,EP} + 2a_{k,SS}) - \frac{4B_k^3 + 12B_k}{(1 - B_k^2)^3} a_{k,C}
 \end{aligned} \tag{20}$$

Using these terms, we can now express the Hessian H as:

$$H = \begin{bmatrix} \frac{\partial^2}{\partial A_1^2} \Phi & 0 & \frac{\partial^2}{\partial A_1 \partial D} \Phi \\ 0 & \frac{\partial^2}{\partial A_2^2} \Phi & \frac{\partial^2}{\partial A_2 \partial D} \Phi \\ \frac{\partial^2}{\partial A_1 \partial D} \Phi & \frac{\partial^2}{\partial A_2 \partial D} \Phi & \frac{\partial^2}{\partial D^2} \Phi \end{bmatrix} \tag{21}$$

whose negative inverse determines the parameter's covariance matrix $C = -H^{-1}(A_{1,max}, A_{2,max}, D_{max})$ (see for example [6]). In order to get a better sense of how the parameter estimation errors behave we can calculate the expectation values of the second derivatives considering $N \rightarrow \infty$: $\langle a_{k,EP} \rangle = 2A_k$, $\langle a_{k,SS} \rangle = (N - 2)A_k$, and $\langle a_{k,C} \rangle = (N - 1)A_k B_k$. This results in the following expectation values of the derivatives of $Q_k(B_k)$:

$$\begin{aligned}
 \langle Q_k(B_k) \rangle &= N A_k \\
 \left\langle \frac{\partial}{\partial B_k} Q_k(B_k) \right\rangle &= \frac{2N A_k B_k}{1 - B_k^2} \\
 \left\langle \frac{\partial^2 Q_k(B_k)}{\partial B_k^2} \right\rangle &= -\frac{4N A_k}{(1 - B_k^2)^3} (B_k^4 + 3B_k^2 - 3B_k - 1)
 \end{aligned} \tag{22}$$

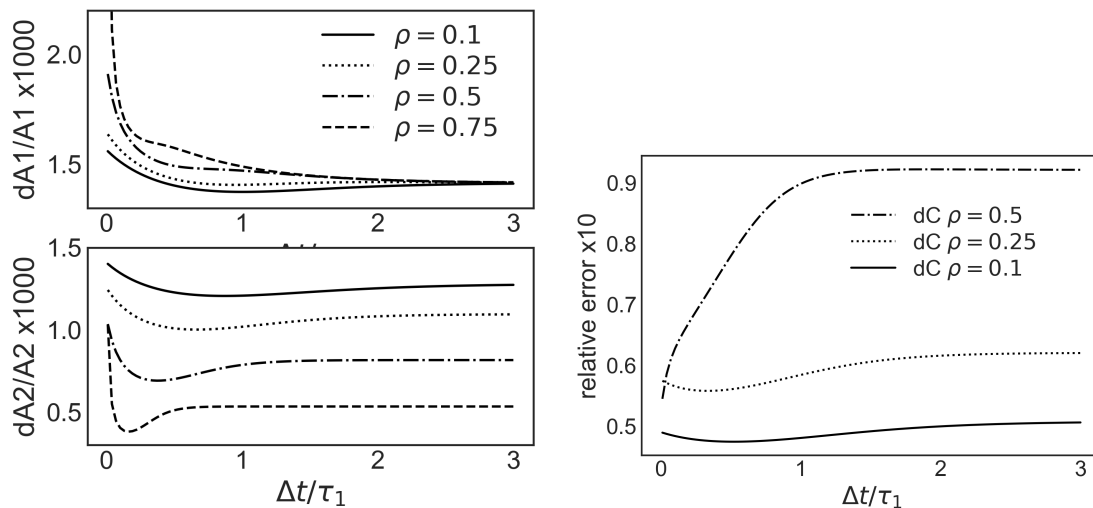
From this, we can express the Hessian using the following second derivatives:

$$\begin{aligned}
 \frac{\partial^2}{\partial A_k^2} \Phi &= -\frac{N}{2A_k^2} - \frac{N}{(1 - B_k^2)^2} \left(\frac{\partial B_k}{\partial A_k} \right)^2 \left(1 + B_k^2 + \frac{6B_k}{1 + B_k} \right) + \frac{2NB_k}{A_k(1 - B_k^2)} \frac{\partial B_k}{\partial A_k} \\
 \frac{\partial^2}{\partial A_k \partial D} \Phi &= \frac{NB_k}{A_k(1 - B_k^2)} \frac{\partial B_k}{\partial D} - \frac{N}{(1 - B_k^2)^2} \frac{\partial B_k}{\partial A_k} \frac{\partial B_k}{\partial D} \left(1 + B_k^2 + \frac{6B_k}{1 + B_k} \right) \\
 \frac{\partial^2}{\partial D^2} \Phi &= -\frac{N}{(1 - B_1^2)^2} \left(\frac{\partial B_1}{\partial D} \right)^2 \left(1 + B_1^2 + \frac{6B_1}{1 + B_1} \right) - \frac{N}{(1 - B_2^2)^2} \left(\frac{\partial B_2}{\partial D} \right)^2 \left(1 + B_2^2 + \frac{6B_2}{1 + B_2} \right)
 \end{aligned} \tag{23}$$

These equations allow us to express how the uncertainties/errors of the parameters depend on the sampling interval $\Delta t/\tau$ given that we measure N samples. In Fig. 2 we plot the relative errors of A_1 , A_2 and C as a function of $\Delta t/\tau_1$ and ρ for large fixed N ($N = 10^6$) using eqs. 23. For example, in Fig.2a we can see that the optimal sampling interval as given by $\Delta t/\tau_1$ is reducing as ρ is increasing. This is most likely due to the fact that the relaxation τ_2 is decreasing as ρ is increasing which will affect the uncertainty of both A_1 and A_2 . We also observe a shift in behavior for larger $\rho > 0.25$ where no minimum $\Delta t/\tau_1$ for A_1 and C exists. This is in contrast to our finding for a single OU process [2] where the relaxation time τ had a minimum of approximately $\Delta t = 0.7\tau$ and dA continuously decreased as $\Delta t/\tau$ decreased.

ρ	A_1 (MCMC)	A_1 (ML)	A_2 (MCMC)	A_2 (ML)	C (MCMC)	C (ML)	dC (MCMC)	dC (ML)
0.1	1.02 ± 0.09	1.00 ± 0.08	0.84 ± 0.07	0.82 ± 0.06	0.24 ± 0.15	0.23 ± 0.14	0.14 ± 0.018	0.068 ± 0.0076
0.25	1.02 ± 0.09	1.00 ± 0.08	0.61 ± 0.04	0.60 ± 0.04	0.70 ± 0.18	0.67 ± 0.17	0.19 ± 0.02	0.09 ± 0.009
0.5	1.02 ± 0.08	1.00 ± 0.08	0.33 ± 0.02	0.33 ± 0.02	2.00 ± 0.29	2.05 ± 0.30	0.31 ± 0.039	0.144 ± 0.013
0.75	1.02 ± 0.08	1.00 ± 0.08	0.145 ± 0.006	0.144 ± 0.005	6.1 ± 0.6	6.0 ± 0.6	0.69 ± 0.08	0.37 ± 0.03
0.9	1.02 ± 0.08	1.00 ± 0.08	0.053 ± 0.002	0.053 ± 0.00003	18.4 ± 1.9	17.97 ± 1.5	1.9 ± 0.24	1.3 ± 0.11

TABLE I: Comparison of parameter estimates using maximum likelihood and Markov-Chain Monte-Carlo methods for different ρ at $D = 1, A_1 = 1, \Delta t = 0.3, N = 1000$



(a) Amplitude errors dA_1 and dA_2 as function of $\Delta t/\tau$ (b) Relative coupling coefficient error dC as function of $\Delta t/\tau$

FIG. 2: Dependence of relative errors on $\Delta t/\tau$ for large N ($N = 10^6$) using eqs. 23

D. Validation by Simulation

In this section, we create artificial correlated time-series employing the method mentioned in B to validate our Maximum-Likelihood method against parameter estimation by Markov-Chain Monte-Carlo (MCMC) methods [8]. Probabilistic calculations using MCMC were carried out in python and Julia, using the anaconda distribution of python [9] with PYMC3 [10] and Julia 1.53 [11] with Turing.jl [12]. For the following table and figures, we created 400 correlated time-series pairs for Pearson Correlation coefficients of 0.9, 0.5, 0.25, and 0.1 for $N = 1000$ at $D = 1, A_1 = 1, \Delta t = 0.3$. Since the problem is symmetric with respect to the sign of the correlation coefficient we don't need to simulate negative correlation coefficients. Fig. 3 shows the probability distributions of A_1 and A_2 for $N = 1000$ and $\rho = 0.5$ estimated by maximum likelihood (ML) and Markov-Chain Monte-Carlo (MCMC). The results of the simulations are summarized in Table I. In terms of uncertainties we find that MCMC correctly estimates parameter uncertainties consistent with the distribution over a large sample, but that the maximum likelihood method underestimates the uncertainty by a constant factor that depends on the degree of correlation ρ or C . In Fig. 4 we show using a scatter plot that the individual uncertainties of MCMC and ML are correlated. This fact is important since MCMC is much more computationally expensive. For large datasets, we can use the maximum likelihood method to run over the whole dataset, and then we can use MCMC on a few samples to determine the factor by which the ML method underestimates the uncertainties. In our example, maximum likelihood underestimates the uncertainties by a factor of 1.5-2 over a wide range of ρ . We believe this underestimation results from the fact that the distributions in Fig. 3 are not Gaussian. The ML method assumes that all posterior distributions are Gaussian and that the width can be determined by the second derivative of the log likelihood. This is clearly not the case. The fact that we underestimate the uncertainties indicates that the curvature of our distribution at the maximum is greater than a corresponding Gaussian distribution.

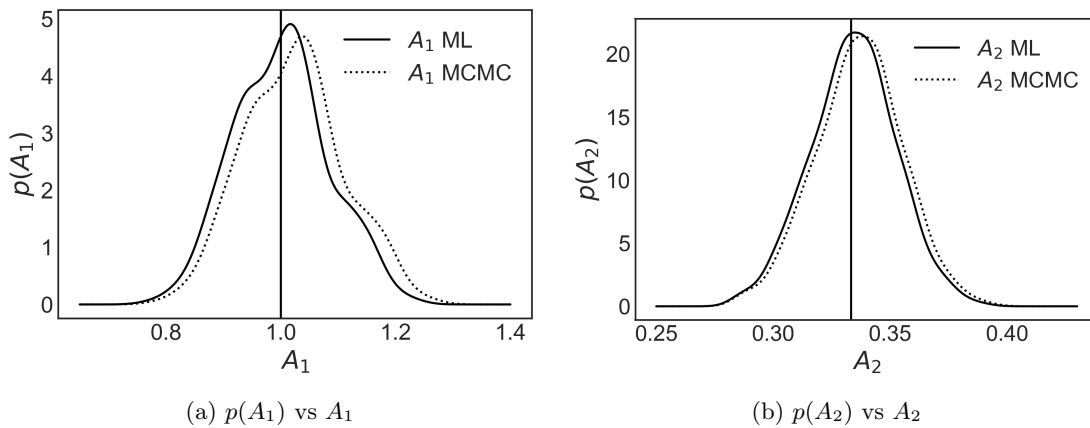


FIG. 3: Comparison of MCMC and maximum likelihood for $\rho = 0.5$ and $D = 1, A_1 = 1, \Delta t = 0.3, N = 1000$. The vertical bar indicates the true expected value

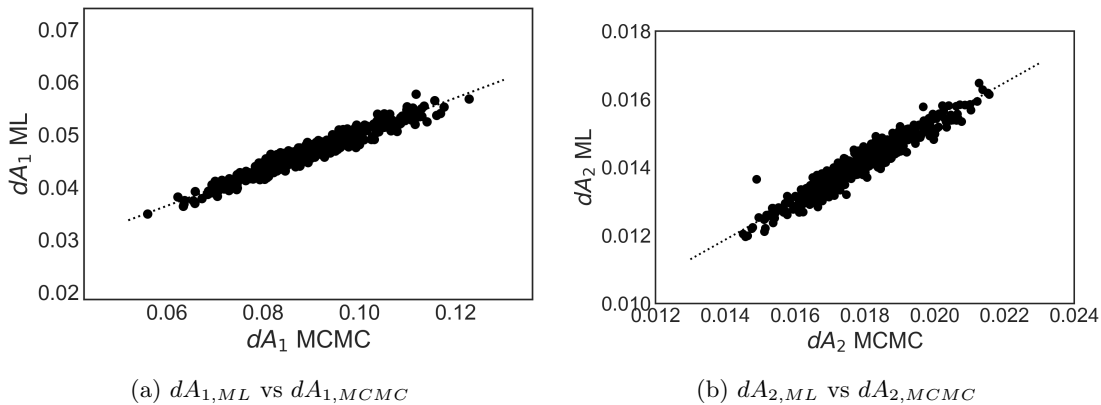


FIG. 4: Correlation between the uncertainties of A estimated by maximum likelihood and MCMC for $\rho = 0.5$ and $D = 1, A_1 = 1, \Delta t = 0.3, N = 1000$. The dotted line represents a linear fit. For this ρ the maximum likelihood underestimates the errors by a factor of 2

E. Application to resting-state functional Magnetic Resonance Imaging

To illustrate the power of our correlation method, we analyzed functional Magnetic Resonance Imaging (fMRI) time-series data from control 7T-fMRI resting-state experiments that were taken from [13]. In particular, we extracted neural activation time-series from four brain regions that make up the default mode network [14] (LLP = left lateral parietal cortex, RLP = right lateral parietal cortex, MPFC = medial prefrontal cortex, PCC = posterior cingulate cortex) that activates when individuals are focused on their internal mental-state processes, such as self-referential processing, interoception etc. and not performing a specific task. Fig. 5 shows the time series of a single subject. Traditionally, correlations between brain regions are evaluated using Pearson correlation coefficients. To ensure that our coupled OU model is applicable, we determined the skewness and kurtosis of all time series to show that they are Gaussian. None of the distributions had an absolute value of skewness/kurtosis of more than 0.5, which is considered safe for applying Gaussian statistics (see for example [15]). Next, we have to show that the time series exhibit an exponentially decaying autocorrelation function. For this, we applied our maximum likelihood method for OU processes to determine the relaxation times for all time series. We found that $\tau_{MPFC} = 4.8 \pm 0.7sec$, $\tau_{PCC} = 4.8 \pm 0.7sec$, $\tau_{RLP} = 4.1 \pm 0.5sec$, $\tau_{LLP} = 3.8 \pm 0.5sec$. The decay-times of all brain regions are in within one or two standard deviations of each other, justifying the assumption made in eq. 3. Fig. 6 shows the coupling coefficients of five subject estimated by our method. Here we show the estimates from MCMC, but the same mean values are obtained by using the maximum likelihood method with uncertainties that are about a factor of 3 smaller than the MCMC estimates (see the Discussion in Simulation section). We chose four of six possible pairings of brain

regions to illustrate the additional information that our method provides over Pearson correlations. By obtaining not only the strength of coupling but also the uncertainties of the value given the data, we can distinguish between variance due to measurement and variance due to difference between subjects. Panel (a) illustrates a case where the uncertainty of the measurements is smaller than the variance of the mean. In this case this is due to an outlier. Here the traditional analysis would have sufficient to detect an outlier. In Panel (b) and (d) the variance of the mean is in between the variances of the individual datapoints and because we know each measurements uncertainty, we can determine how much of the variance of the mean is due to the variance between subjects and due to measurement error. Panel (c) illustrates a case where the uncertainty of the measurement is larger than variance of the mean.

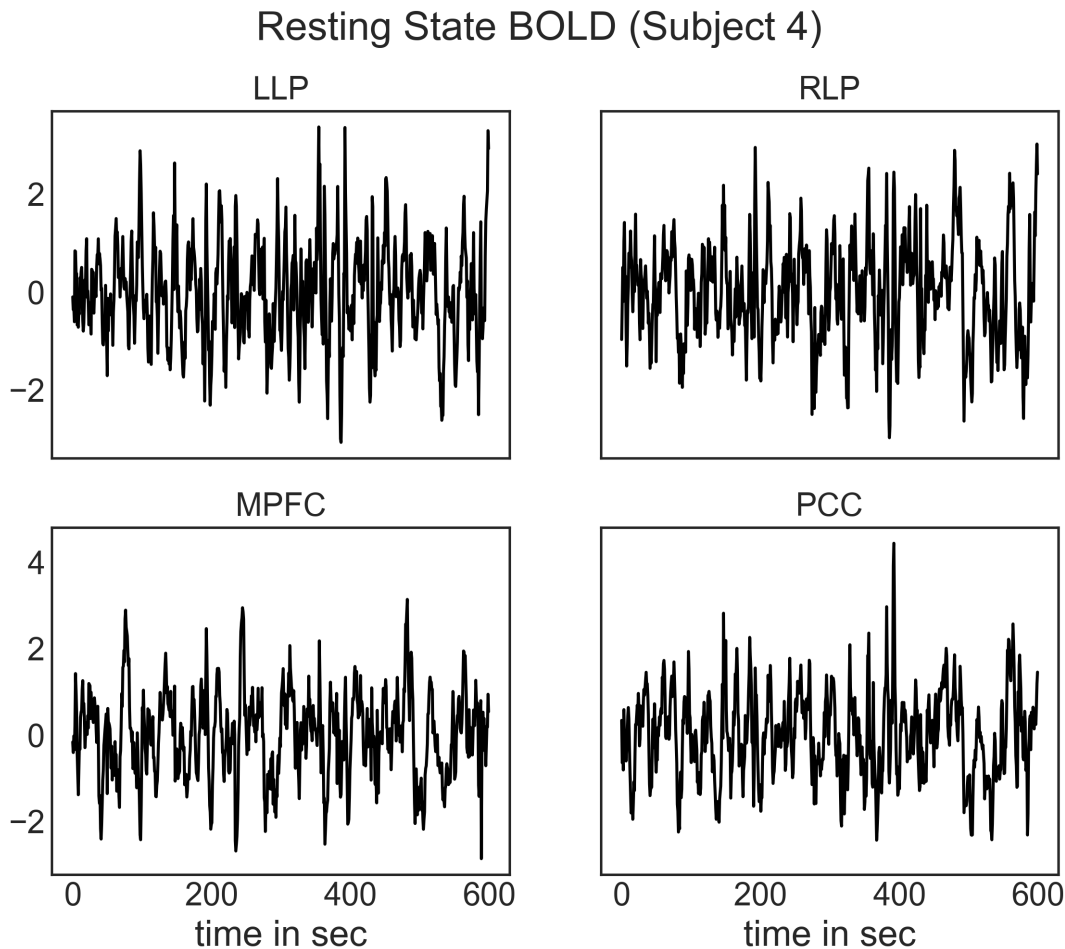


FIG. 5: Time series of fMRI-BOLD signals of four brain regions that constitute the default mode network (LLP = left lateral parietal cortex, RLP = right lateral parietal cortex, MPFC = medial prefrontal cortex, PCC = posterior cingulate cortex)

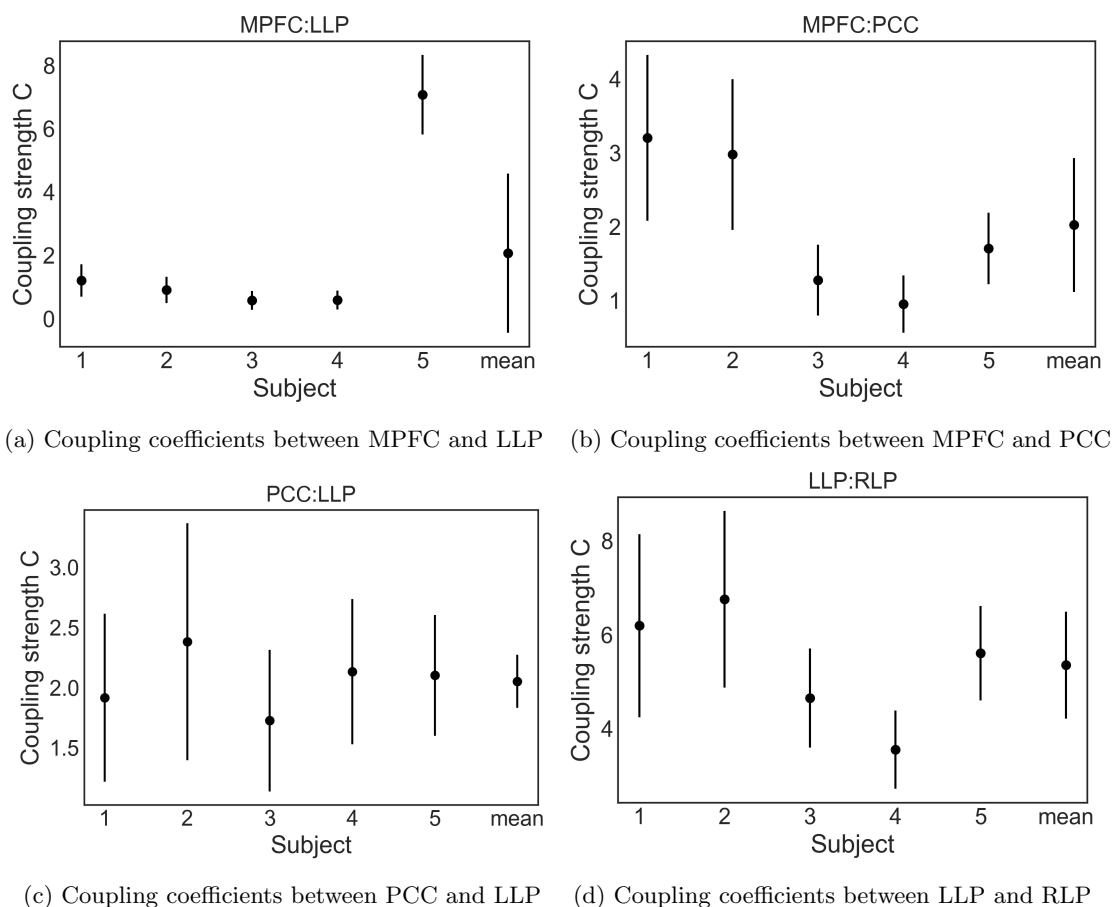


FIG. 6: Coupling coefficients between default mode network brain regions across several subjects and the mean and standard deviation between subjects. For comparison to Pearson correlation coefficients we can use $\rho = C/(2 + C)$ ($\rho = 1/3$ for $C = 1$, $\rho = 0.6$ for $C = 3$, and $\rho = 0.75$ for $C = 6$). The coupling coefficients and their uncertainties were determined by MCMC. We obtained the same mean coupling coefficients using the maximum likelihood method but with uncertainties that are about a factor of 3 smaller than the MCMC estimate.

F. Summary

In the previous sections we developed and validated a maximum likelihood method to estimate the correlation between time series with an intrinsic relaxation time (OU process). As compared to Pearson correlation, our method not only estimates the correlation parameter, but also its confidence interval, which in our opinion is the most useful feature of our method. Knowing the confidence interval of individual measurements allows us to distinguish between measurement error (confidence in a single measurement) and the variance that results from the variance of the sample. In the biological sciences, we often find large biological variance that is manifested in cell-cell variability of genetically identical cells or in the variance in the measurement of brain circuitry between (different subjects) but also within subjects (same subject at different days). In order to characterize these biological variances, biophysical methods have to be developed that are capable of measuring quantities accurately enough as to distinguish between measurement error and biological distributions. Only if the individual error of a measurement is known we can attempt such a deconvolution. We hope that our method will help in achieving this goal.

ACKNOWLEDGMENTS

We wish to acknowledge funding by the the National Institute of Drug Abuse (SBIR Phase 1 & Phase 2 Grant 1R44 DA043277-01) and the National Heart, Lung and Blood Institute (Grant 5U01HL12752202)

-
- [1] G. Uhlenbeck and I. Ornstein, *Phys. Rev.* **36**, 823 (1930).
 - [2] H. H. Strey, *Phys Rev E* **100**, 062142 (2019).
 - [3] S. Bera, S. Paul, R. Singh, D. Ghosh, A. Kundu, A. Banerjee, and R. Adhikari, *Sci Rep* **7**, 41638 (2017).
 - [4] A. Hurn, J. Jeisman, and K. Lindsay, *Journal of Financial Econometrics* **5**, 390–455 (2007).
 - [5] R. Singh, D. Ghosh, and R. Adhikari, *Phys Rev E* **98**, 012136 (2018).
 - [6] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd ed. (Oxford University Press, 2006).
 - [7] J. J. More, B. S. Garbow, and K. E. Hillstrom, 10.2172/6997568.
 - [8] D. Gamerman and H. Lopes, *Markov Chain Monte Carlo*, 2nd ed., Texts in Statistical Science (Chapman & Hall/CRC, Boca Raton, FL, 2006).
 - [9] Anaconda Inc., “Anaconda python,”.
 - [10] J. Salvatier, T. Wiecki, and C. Fonnesbeck, *PeerJ CompSci* **2**, e55 (2016).
 - [11] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *SIAM review* **59**, 65 (2017).
 - [12] H. Ge, K. Xu, and Z. Ghahramani, in *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain* (2018) pp. 1682–1690.
 - [13] J. S. Ide, S. Nedic, K. F. Wong, S. L. Strey, E. A. Lawson, B. C. Dickerson, L. L. Wald, G. La Camera, and L. R. Mujica-Parodi, *Neuroimage* **174**, 35 (2018).
 - [14] M. Raichle, A. MacLeod, A. Snyder, W. Powers, D. Gusnard, and G. Shulman, *PNAS* **98**, 676 (2001).
 - [15] J. Hair, W. Black, B. Babin, and R. Anderson, *Multivariate Data Analysis* (Pearson, 2009).