Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics Letters A 333 (2004) 261–268
PHYSICS LETTERS A
www.elsevier.com/locate/pla

GARCH modelling of covariance in dynamical estimation of inverse solutions

Andreas Galka [a,b,*], Okito Yamashita [c], Tohru Ozaki [b]

[a] *Institute of Experimental and Applied Physics, University of Kiel, 24098 Kiel, Germany*
[b] *Institute of Statistical Mathematics (ISM), Minami-Azabu 4-6-7, Tokyo 106-8569, Japan*
[c] *ATR Computational Neuroscience Laboratories, Hikaridai 2-2-2, Kyoto 619-0288, Japan*

Received 13 August 2004; accepted 8 October 2004

Available online 29 October 2004

Communicated by C.R. Doering

Abstract

The problem of estimating unobserved states of spatially extended dynamical systems poses an inverse problem, which can be solved approximately by a recently developed variant of Kalman filtering; in order to provide the model of the dynamics with more flexibility with respect to space and time, we suggest to combine the concept of GARCH modelling of covariance, well known in econometrics, with Kalman filtering. We formulate this algorithm for spatiotemporal systems governed by stochastic diffusion equations and demonstrate its feasibility by presenting a numerical simulation designed to imitate the situation of the generation of electroencephalographic recordings by the human cortex.

PACS: 02.30.Zz; 05.45.Tp; 95.75.Wx; 02.50.-r

Keywords: Multivariate time series; State space modelling; Inverse problem; Kalman filtering; GARCH

1. Introduction

In many fields of science spatially extended systems are studied which evolve in time according to some possibly complicated dynamics. It is a typical situation that the relevant state variables of such systems cannot be observed directly, but only through an observation function; in many cases this function performs a (possibly non-linear) projection of the high-dimensional true state space of the system onto an observation space of much lower dimension. The task of retrieving estimates of the true states from such observations represents a typical inverse problem. Due to the absence of a simple invertible relationship between state and observation such problems are ill-posed, i.e., the solutions ("inverse solutions") are unstable and ambiguous. As examples of fields where such problems typically arise, we mention brain

* Corresponding author.
E-mail address: galka@physik.uni-kiel.de (A. Galka).

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2004.10.045

research [1], seismic tomography [2], oceanography [3], quantum scattering [4–6], electrical impedance tomography [7,8], microstructure analysis [9] and lossy data compression [10]; more examples could be added easily.

The literature on methods for estimating approximative solutions of these inverse problems is vast; many approaches fall into the category of constrained least-squares methods employing Tikhonov regularisation [11,12]. In the case of the inverse problem of estimating the current density field generating the human electroencephalogram (EEG) this approach has become known as "low-resolution electromagnetic tomography" (LORETA) [13]. Generally these algorithms are instantaneous, i.e., they are applied independently to the data obtained at different points of time; however, it is possible to introduce an additional temporal smoothness constraint [14,15], which, from a dynamical systems perspective, corresponds to a random-walk model for the evolution of the unobserved states. This approach has recently been generalised to a general spatiotemporal dynamical model [16].

In this Letter we would like to discuss and extend an alternative approach which is based on Kalman filtering. The Kalman filter is the natural tool for estimating unobserved states in dynamical systems [17]; nevertheless, only a small number of applications of Kalman filtering to inverse problems have been reported so far [18–20]. The application of Kalman filtering to spatiotemporal systems poses considerable challenges; through discretisation the spatial dimension can be cast into a high-dimensional state vector, but the dynamical model governing the dynamics of this state vector will depend on a very large number of unknown parameters. If various approximations are imposed, such as assuming homogeneity and time-invariance of parameters, this number can be reduced considerably, such that the remaining parameters can be estimated by maximum likelihood methods [16,20]; however, in many cases such strong approximations are inappropriate with respect to the actual physics of the underlying system. This applies in particular to approximations relating to the covariance matrix of the noise driving the dynamics within a spatiotemporal dynamical model.

There exists an elegant approach for providing more freedom to the covariance matrix in the context of state estimation, which is known as the "generalised autoregressive conditional heteroskedasticity" (GARCH) model, originally introduced in econometrics as a model for time-varying volatility (and recently honoured with the Nobel Prize in Economic Sciences 2003) [21,22]. The idea of applying a Kalman filter to GARCH models was proposed by Harvey et al. [23]; here we take a different point of view by regarding the combination of a Kalman filter with a covariance model of GARCH type as a new variant (in fact, generalisation) of Kalman filtering, instead of just as an example for the application of Kalman filtering. Again, this new generalised Kalman filter can be applied to a wide variety of dynamical models and data sets. Since covariance matrices represent a core element of the Kalman filter iteration, the incorporation of the GARCH model into the filter constitutes a major modification of the standard Kalman filter. Numerical results will be shown for a simulation study imitating the situation given with the inverse problem of EEG generation.

2. Problem formulation

Assume that the unobserved true states of a spatiotemporal system are given by a vector field $\boldsymbol{x}(\boldsymbol{r}, t)$, where $\boldsymbol{r}$ and t denote space and time, respectively; $\boldsymbol{x}$ denotes the local $n_{\boldsymbol{x}}$-dimensional state vector. Space and time shall be discretised into grid points $\boldsymbol{r}_v$, $v = 1, \ldots, N_{\boldsymbol{r}}$, and time points t_k, $k = 1, \ldots, N_t$; for simplicity let the sampling points of discretised space and time be denoted by v and k, respectively, instead of $\boldsymbol{r}_v$ and t_k. Then

$$\mathbf{X}(k) = \left(\boldsymbol{x}(1,k)^{\dagger}, \ldots, \boldsymbol{x}(N_{\boldsymbol{r}},k)^{\dagger}\right)^{\dagger},$$

denotes the global state vector of the system at time k. The dimension of $\mathbf{X}$ is given by $N_{\mathbf{X}} = n_{\boldsymbol{x}} N_{\boldsymbol{r}}$. A vector of $N_{\mathbf{Y}}$ scalar measurements at time k shall be given by $\mathbf{Y}(k) = (y_1(k), \ldots, y_{N_{\mathbf{Y}}}(k))^{\dagger}$. We assume the unfavourable case $N_{\mathbf{Y}} \ll N_{\mathbf{X}}$. In this Letter we anticipate that the observation equation is approximately linear:

$$\mathbf{Y}(k) = \mathsf{K}\mathbf{X}(k) + \boldsymbol{\epsilon}(k), \tag{1}$$

where $\boldsymbol{\epsilon}(k)$ denotes a $N_{\mathbf{Y}}$-dimensional vector of observational noise, which we assume to be white and Gaussian with zero mean and covariance matrix C_{ϵ}.

The $N_{\mathbf{Y}} \times N_{\mathbf{X}}$ observation matrix K is assumed to be known and of full rank.

We will make the assumption that C_ϵ has the simplest possible structure, namely

$$\mathsf{C}_\epsilon = \sigma_\epsilon^2 \mathsf{I}_{N_{\mathbf{Y}}}, \tag{2}$$

where $\mathsf{I}_{N_{\mathbf{Y}}}$ denotes the $N_{\mathbf{Y}} \times N_{\mathbf{Y}}$ identity matrix, i.e., we assume that the observation noise is uncorrelated between all pairs of sensors and of equal variance for all sensors. These assumptions may be relaxed in future work.

Since direct inversion of Eq. (1) is impossible, additional constraints have been employed. It has been suggested that among all states $\mathbf{X}$ which are consistent with the observations $\mathbf{Y}$ (for a given time point k), the state with maximum spatial smoothness should be chosen [13]. In order to quantify the smoothness of a given state, the following definitions are useful.

For each grid point v let $\mathcal{N}(v)$ denote the set of nearest neighbours; in a rectangular grid there will be six nearest neighbours for each grid point, except for points directly at the borders of the system. Let N denote a $N_{\boldsymbol{r}} \times N_{\boldsymbol{r}}$ matrix having $\mathsf{N}_{vv'} = 1$ if $v' \in \mathcal{N}(v)$ and 0 otherwise. Then a discrete spatial Laplacian operator can be defined by

$$\mathsf{L}_{N_{\mathbf{X}}} = \left(\mathsf{I}_{N_{\boldsymbol{r}}} - \frac{1}{\kappa}\mathsf{N}\right) \otimes \mathsf{I}_{n_{\boldsymbol{x}}}, \tag{3}$$

here the symbol $\otimes$ denotes Kronecker multiplication of matrices. For κ a value of 6 is to be expected, but there may be reasons for choosing slightly different values. Multiplication of $\mathsf{L}_{N_{\mathbf{X}}}$ with $\mathbf{X}$ represents a discrete spatial derivative of second order.

In the regularised LORETA approach the inverse solution is obtained by minimizing the objective function

$$E(\mathbf{X}) = \left\|(\mathbf{Y} - \mathsf{K}\mathbf{X})\right\|^2 + \lambda^2 \|\mathsf{L}_{N_{\mathbf{X}}}\mathbf{X}\|^2, \tag{4}$$

i.e., a weighted sum of the observation fitting error and of a term measuring non-smoothness by the norm of the Laplacian of the state vector. The hyperparameter λ expresses the balance between fitting of observations and the smoothness constraint; a non-zero value for λ provides regularisation for the solution [13]. A closed form for the solution of this minimisation task can be derived [16]; it has to be applied independently for each time point k.

3. Spatiotemporal models

It has to be regarded as a weakness of the LORETA approach that for estimating the inverse solution for time point k no results from previous measurements $\mathbf{Y}(k')$, $k' < k$ and from the corresponding inverse solutions are employed. The idea of incorporating such additional information into the estimation process leads to dynamical inverse solutions [20].

For the purpose of estimation of dynamical inverse solutions a model for the dynamics of the system is needed. If detailed models can be deduced from basic principles, they will be extremely valuable, but in many cases (e.g., brain research) this is not the case. As a first-order approximation under the assumption of linearity, a general-purpose ansatz is given by the stochastic partial differential equation

$$\frac{\partial^p \boldsymbol{x}(\boldsymbol{r},t)}{\partial t^p} = b\frac{\partial^2 \boldsymbol{x}(\boldsymbol{r},t)}{\partial \boldsymbol{r}^2} + \boldsymbol{\eta}(\boldsymbol{r},t), \tag{5}$$

where p is a small positive integer (typical choices are $p = 1$ or $p = 2$), and $\boldsymbol{\eta}(t)$ denotes an integrable dynamical noise term. After discretisation (with respect to time and space), for the case $p = 2$ this equation approximately corresponds to

$$\begin{aligned}\boldsymbol{x}(v,k) = {} & \mathsf{a}_1 \boldsymbol{x}(v,k-1) + \mathsf{a}_2 \boldsymbol{x}(v,k-2) \\ & + \mathsf{b}_1\left[\mathsf{L}_{N_{\mathbf{X}}}\mathbf{X}(k-1)\right]_v + \boldsymbol{\eta}(v,k),\end{aligned} \tag{6}$$

where $\mathsf{L}_{N_{\mathbf{X}}}$ has been defined in Eq. (3), $[\mathbf{X}]_v$ denotes the $n_{\boldsymbol{x}}$-dimensional vector composed of those elements within $\mathbf{X}$ corresponding to grid point v, and, in the simplest case, the autoregressive parameter matrices are given by

$$\mathsf{a}_1 = a_1 \mathsf{I}_{n_{\boldsymbol{x}}}, \qquad \mathsf{a}_2 = a_2 \mathsf{I}_{n_{\boldsymbol{x}}} \quad \text{with}$$
$$a_1 = 2 \quad \text{and} \quad a_2 = -1,$$

also b_1 can be modelled in the form $\mathsf{b}_1 = b_1 \mathsf{I}_{n_{\boldsymbol{x}}}$. We see that the local model at grid point v becomes an autoregressive model of second order (AR(2)) with an additional first-order term contributed by the discrete spatial Laplacian, i.e., involving the neighbours of grid point v. Globally this model corresponds to a multivariate AR model

$$\mathbf{X}(k) = \mathsf{A}_1 \mathbf{X}(k-1) + \mathsf{A}_2 \mathbf{X}(k-2) + \boldsymbol{H}(k), \tag{7}$$

where

$$\mathsf{A}_1 = a_1 \mathsf{I}_{N_{\mathbf{X}}} + b_1 \mathsf{L}_{N_{\mathbf{X}}} \quad \text{and} \quad \mathsf{A}_2 = a_2 \mathsf{I}_{N_{\mathbf{X}}}. \tag{8}$$

Here most elements of the $N_{\mathbf{X}} \times N_{\mathbf{X}}$ parameter matrices A_1 and A_2 are zero, and the dynamical noise term $\boldsymbol{H}(k)$ is a $N_{\mathbf{X}}$-dimensional vector, which we assume to be white and Gaussian with zero mean and covariance matrix C_η. Generically it is not possible to assume that C_η is diagonal, but by analogy with the LORETA method the choice

$$\mathsf{C}_\eta = \sigma_\eta^2 \left(\mathsf{L}_{N_{\mathbf{X}}}^\dagger \mathsf{L}_{N_{\mathbf{X}}}\right)^{-1}, \tag{9}$$

provides a reasonable first approximation [16]; since $\mathsf{L}_{N_{\mathbf{X}}}$ and A_i commute (compare Eq. (8)), this ansatz is equivalent to assuming that if in Eq. (7) $\mathbf{X}(k)$ is replaced by $\tilde{\mathbf{X}}(k) = \mathsf{L}_{N_{\mathbf{X}}} \mathbf{X}(k)$, the corresponding covariance matrix $\mathsf{C}_{\tilde{\eta}}$ will be diagonal. Since $\mathsf{L}_{N_{\mathbf{X}}}$ is a differentiation operator, we call this transformation "spatial whitening" and replace $\mathbf{X}(k)$ by $\tilde{\mathbf{X}}(k)$ from now on; for convenience the tilde will be omitted.

4. Spatiotemporal Kalman filtering

In principle, based on Eqs. (1) (where we replace K by $\tilde{\mathsf{K}} = \mathsf{K}\mathsf{L}_{N_{\mathbf{X}}}^{-1}$, but again omit the tilde) and (7) standard Kalman filtering could be employed, in order to obtain estimates for $\mathbf{X}(k)$; but for large $N_{\mathbf{X}}$ the computational time and memory consumption would become prohibitively large. In [20] a spatiotemporal Kalman filter was introduced, which is based on the idea of replacing the intractable $N_{\mathbf{X}}$-dimensional filtering problem by $N_{\boldsymbol{r}}$ coupled tractable $n_{\boldsymbol{x}}$-dimensional filtering problems; this step requires diagonality of C_η, which explains the need for the spatial whitening transformation discussed in the previous section.

In [20] the spatiotemporal Kalman filter has been developed within the context of the inverse problem of EEG generation; here we will summarise the main steps in generalised notation. In the following, by writing $\hat{x}(k_1 \mid k_2)$ we denote an estimate of the value of some quantity x computed at time k_1, based on all observations obtained until time $k_2 \leqslant k_1$. First, in order to confine the autoregressive model Eq. (6) to first order $\boldsymbol{x}(v,k)$ is replaced by

$$\mathbf{x}(v,k) = \left(\boldsymbol{x}(v,k)^\dagger, \boldsymbol{x}(v,k-1)^\dagger\right)^\dagger,$$

new local parameter matrices are defined by

$$\mathsf{a} = \begin{pmatrix} \mathsf{a}_1 & \mathsf{a}_2 \\ \mathsf{I}_{n_{\boldsymbol{x}}} & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{b} = \begin{pmatrix} \mathsf{b}_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{10}$$

then for each v the local state prediction is given by (compare Eq. (6))

$$\hat{\mathbf{x}}(v,k \mid k-1) = \mathsf{a}\hat{\mathbf{x}}(v,k-1 \mid k-1) + \mathsf{b} \begin{pmatrix} [\mathsf{L}_{N_{\mathbf{X}}} \hat{\mathbf{X}}(k-1 \mid k-1)]_v \\ \mathbf{0} \end{pmatrix} \tag{11}$$

and the corresponding local prediction error covariance matrix can be approximated by

$$\mathsf{p}(v,k \mid k-1) = \mathsf{a}\mathsf{p}(v,k-1 \mid k-1)\mathsf{a}^\dagger + \mathsf{c}_\eta, \tag{12}$$

where the $2n_{\boldsymbol{x}} \times 2n_{\boldsymbol{x}}$ local dynamical noise covariance matrix c_η is given by

$$\mathsf{c}_\eta = \begin{pmatrix} \sigma_\eta^2 \mathsf{I}_{n_{\boldsymbol{x}}} & 0 \\ 0 & 0 \end{pmatrix}. \tag{13}$$

The global state prediction for all grid points is given by

$$\hat{\mathbf{X}}(k \mid k-1) = \left(\left[\hat{\mathbf{x}}(1,k \mid k-1)^\dagger\right]_{1:n_{\boldsymbol{x}}}, \ldots, \left[\hat{\mathbf{x}}(N_{\mathbf{X}},k \mid k-1)^\dagger\right]_{1:n_{\boldsymbol{x}}}\right)^\dagger, \tag{14}$$

where $[\mathbf{x}]_{1:n}$ denotes the vector formed by the first n elements of vector $\mathbf{x}$. The observation prediction vector follows as

$$\hat{\mathbf{Y}}(k \mid k-1) = \mathsf{K}\hat{\mathbf{X}}(k \mid k-1). \tag{15}$$

The actual observation at time k is $\mathbf{Y}(k)$, and the observation prediction error results as

$$\Delta\mathbf{Y}(k) = \mathbf{Y}(k) - \hat{\mathbf{Y}}(k \mid k-1). \tag{16}$$

The corresponding observation prediction error covariance matrix can be approximated by

$$\mathsf{R}(k \mid k-1) = \sum_{v=1}^{N_{\mathbf{X}}} \mathsf{Q}(v)\mathsf{p}(v,k \mid k-1)\mathsf{Q}(v)^\dagger + \sigma_\epsilon^2 \mathsf{I}_{n_c}, \tag{17}$$

where the matrix $\mathsf{Q}(v)$ is defined by

$$\mathsf{Q}(v) = \begin{pmatrix} 0 & \ldots & 0 \\ [\mathsf{K}]_v & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}, \tag{18}$$

here $[\mathsf{K}]_v$ denotes those $n_{\boldsymbol{x}}$ columns from K which correspond to the vth grid point, and the array of zeros on the right has the size $N_{\mathbf{Y}} \times n_{\boldsymbol{x}}$. The Kalman gain matrix for grid point v follows as

$$\mathsf{g}(v,k) = \mathsf{p}(v,k \mid k-1)\mathsf{Q}(v)^\dagger \mathsf{R}(k \mid k-1)^{-1}, \tag{19}$$

and finally the local state estimation and the corresponding local estimation error covariance matrix are given by

$$\hat{\mathbf{x}}(v,k \mid k) = \hat{\mathbf{x}}(v,k \mid k-1) + \mathsf{g}(v,k)\Delta\mathbf{Y}(k) \tag{20}$$

and

$$\mathsf{p}(v,k \mid k) = \big(1 - \mathsf{g}(v,k)\mathsf{Q}(v)\big)\mathsf{p}(v,k \mid k-1), \tag{21}$$

respectively.

Formally, this algorithm results from combining the standard linear Kalman filter, applied locally to each grid point, with the additional idea of treating the contributions of the neighbours as exogeneous inputs, which requires certain extensions of the standard equations. A more pronounced step beyond the standard linear Kalman filter will be introduced in the next section.

5. GARCH modelling of covariance

Application of the Kalman filter, as described in the previous section, anticipates that estimates of the parameters $a_1, a_2, b, \sigma_\epsilon^2$ and σ_η^2 are available; in [20] it has been described, how such estimates can be obtained from given data by likelihood maximisation. This is feasible due to the implicit homogeneity and stationarity assumptions, according to which the values for these parameters do not depend on v and k, i.e., on spatial location and time. If for certain parameters these assumptions are relaxed or dropped, improved inverse solutions can be obtained, but the parameter estimation step becomes considerably more difficult, since the number of parameters to be estimated will assume much larger values.

As an alternative we suggest to include the dynamical noise variance σ_η^2 into the local state, thereby making it dependent on space and time:

$$\boldsymbol{\xi}(v,k) = \big(\boldsymbol{x}(v,k), \boldsymbol{x}(v,k-1), \sigma_\eta^2(v,k)\big), \tag{22}$$

it has been known for long that an appropriate choice of σ_η^2 is particularly important in Kalman filtering, whereas the other parameters are less critical. The dynamics of $\sigma_\eta^2(v,k)$ could be modelled again by a stochastic autoregression, as it is sometimes done in "stochastic volatility models" in econometrics [24]; however, this would render the filtering process considerably more difficult. The central idea of GARCH modelling is to employ the prediction error obtained at the *previous* time step as an estimate of the stochastic shocks driving the dynamics of the time-dependent covariance. Since for inverse problems these prediction errors (in state space, i.e., at the local grid points) are not directly available, we suggest to propagate the information contained in the observation prediction errors back into state space by using the corresponding Kalman gain, i.e., using the previous value of the local state prediction error $\mathsf{g}(v,k-1)\Delta\mathbf{Y}(k-1)$ instead; this approach represents the consistent application of the basic idea of GARCH modelling to the state-space case. Among numerous possible choices for the GARCH dynamics, we choose a variant based on logarithms [25], given by

$$\begin{aligned}\log\sigma_\eta^2(v,k) &= \log\sigma_c^2 + \alpha\log\sigma_\eta^2(v,k-1) \\ &\quad + \gamma\log\sum_{i=1}^{n_{\boldsymbol{x}}}\big[\mathsf{g}(v,k-1)\Delta\mathbf{Y}(k-1)\big]_i^2,\end{aligned} \tag{23}$$

the summation over (squared) vector components (labelled by i) represents the simplest possibility to adapt a vector-valued state dynamics to a scalar covariance dynamics. This equation is inserted into the Kalman filter after Eq. (19); during the next iteration, the resulting new values of $\sigma_\eta^2(v,k)$ for all v are used in Eq. (12).

6. Simulation example

We will now present results for the estimation of inverse solutions for a simulated spatiotemporal system which imitates the typical situation given for the generation of the human EEG. The advantage of employing a simulation is that the unobserved true sources are precisely known, such that it becomes easier to evaluate the performance of different estimators.

A model cortex is discretised into $N_{\mathbf{X}} = 3433$ grid points, using an average probabilistic brain atlas [26], and at each grid point a time-dependent 3-dimensional ($n_{\boldsymbol{x}} = 3$) current density vector is assumed to represent the local state. A highly simplified dynamics is implemented for these states by employing Eq. (6) with $a_1 = 0.7$, $a_2 = 0$, $b_1 = 0.3$ and $\sigma_\eta^2 = 0$; however, for two spherical regions within cortex, one being lo-

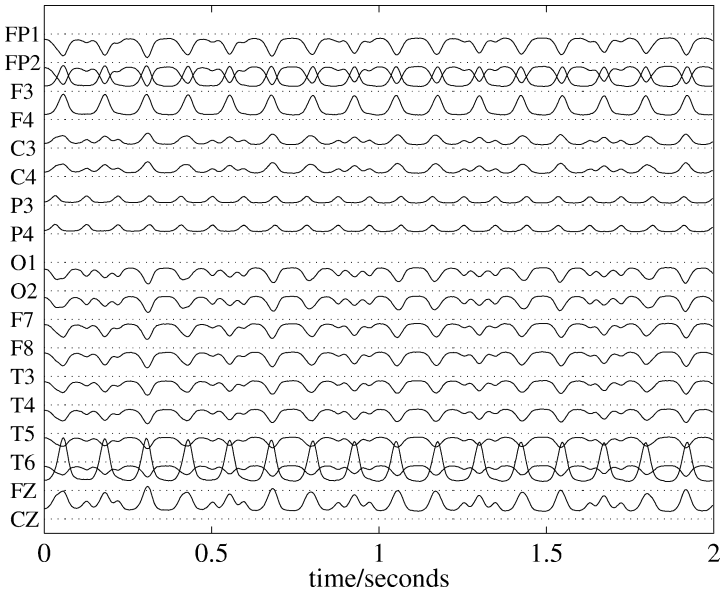

Fig. 1. Simulated EEG recording for 18 standard electrodes according to the 10–20 system (PZ has been omitted); electrode abbreviations are given on the vertical axis. The EEG potential is measured in arbitrary units versus average reference of 19 electrodes (including PZ), time is measured in seconds, assuming a sampling rate of the simulated dynamics of 256 Hz.

cated in frontal brain and one in occipital brain, the parameter a_1 is assumed to depend explicitly on time, according to

$$a_1(t) = a_c\big(1 + a_s \sin(2\pi f t)\big), \tag{24}$$

slightly different values are chosen for a_c, a_x, f and b_1 for both regions (see [20] for further details). As a consequence of this oscillation of the first local autoregressive parameter the dynamics becomes periodically unstable within these two regions (which happens for $a_1(t) + b_1 > 1$), and the corresponding activation spreads out into the cortical grid, where it gradually damps out. These spatially extended oscillations are intended to imitate rhythmic activity, like alpha activity, which is know to be typical of cortical dynamics. All local current vectors are pointing into the vertical direction.

By using the appropriate observation matrix K (known as *lead field matrix*), which can be derived for this situation from basic electromagnetic theory [27], simulated 19-dimensional EEG observations (corresponding to 19 electrodes attached to the surface of the head according to the standard clinical 10-20 system) can be generated; when transformed to average-reference derivation, the chosen data has a standard deviation of 94.09 units, averaged over all electrodes. A small observational noise component of 1% size (with respect to standard deviations) is already included in this value. The simulated EEG time series (of $N_t = 512$ points length, corresponding to two seconds of simulated time for the chosen sampling frequency) is shown in Fig. 1. As can be seen, they clearly display two oscillations with different frequencies, corresponding to the two oscillating source regions, but in a quite blurred fashion.

From the resulting data set inverse solutions are computed by

- an instantaneous method (LORETA) [13];
- spatiotemporal Kalman filtering with the simplest possible dynamical model; i.e., AR(1) with constant homogeneous covariance;
- spatiotemporal Kalman filtering with an AR(2) model and GARCH modelling of covariances;
- spatiotemporal Kalman filtering with the perfect model, i.e., the model employed in generating the simulated data.

For two selected grid points, one of which, denoted by “OGP”, was chosen out of the frontal re-

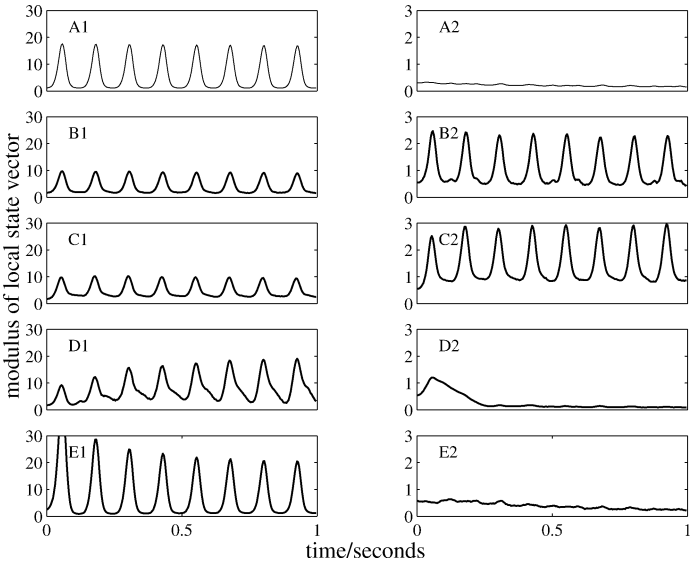

Fig. 2. True and reconstructed local states (inverse solutions) for the first half of the data shown in Fig. 1, for two grid points chosen from an average brain model. Modulus of local current density is shown vs. time; note the different vertical scale for left and right columns of figures. Panels A1, A2 display true local states according to simulation, panels B1, B2 display state estimates from an instantaneous method, panels C1, C2 display estimates from Kalman filtering with the simplest possible dynamical model, panels D1, D2 display estimates from Kalman filtering with GARCH modelling of covariances, and panels E1, E2 display estimates from Kalman filtering with the perfect model. See text for details.

gion displaying high-amplitude oscillations, while the other, denoted by "NOGP", was chosen from a non-oscillating region, the estimated time series of the inverse solutions (modulus of local current density vector) are shown in Fig. 2; for comparison, the true evolution of local states is also shown (top panels). In the figure the inverse solutions are shown only for the first second (256 points) of the data displayed in Fig. 1. The parameters in Eq. (23) were chosen as $\sigma_c^2 = 10$, $\alpha = 0.85$, $\gamma = 0.2$, while the parameters in the dynamical models were estimated by maximisation of likelihood (see [20] for details).

From the figure it can be seen that the instantaneous solution correctly retrieves the oscillation in OGP, but underestimates its true amplitude by a factor of approximately 2; in NOGP spurious oscillations are found. Both these errors are typical for the well-known tendency of instantaneous inverse solutions to produce "blurred" solutions. The dynamical inverse solution with an AR(1) model and constant homogeneous covariance seems not to offer clearly perceptible improvements over the instantaneous solution in this respect, although it can be shown to be superior with respect to likelihood maximisation [20]; we remark that the same holds true for an AR(2) model (results not shown). The dynamical inverse solution with GARCH modelling of covariances starts at the beginning of the data with similar estimates as also the other inverse solutions, but then after a transient of approximately 0.25–0.5 seconds arrives at much better estimates: the oscillation amplitude for OGP is retrieved much better (although the wave shape is somewhat distorted), and no spurious oscillation is induced for NOGP. The initial transient was to be expected for the GARCH dynamics, Eq. (23). Finally, for comparison also the case of the perfect model is shown in the figure, i.e., the case of providing the spatiotemporal Kalman filter with the exact model which had been used for generating the data. It can be seen that, after an initial transient, for both grid points very good estimates are achieved. Of course, in real applications the perfect model will never be available; nevertheless this result is not trivial, since also in this case the Kalman filter has to estimate $3 \times 3433 = 10299$ unobserved state variables from only 18 measurements. It has be-

come obvious that the identification of an appropriate model forms a crucial precondition for obtaining inverse solutions of improved quality.

7. Conclusion

In this Letter we have discussed and extended a general framework for obtaining inverse solutions for spatially extended dynamical systems through a recently introduced new variant of Kalman filtering, and we have proposed to link this methodology with the concept of GARCH modelling of covariance, which, to the best of our knowledge, had previously not been employed in the context of dynamical inverse problems.

Through a simulation study, which extends the results of a previous paper [20], we have demonstrated the feasibility of this new approach and its potential for providing considerably improved inverse solutions. Since our interest in the field of inverse problems was stimulated by the task of source analysis using EEG time series, we have chosen a simulation which is based on this particular situation, but we expect that the ideas underlying spatiotemporal Kalman filtering could be useful in a variety of problems arising in present-day scientific research.

We have only been able to present the core ideas and first results, while many questions still remain open. First, we are not yet able to propose a systematic approach for choosing the GARCH parameters σ_c^2, α and γ in Eq. (23). It is clear that they need to be chosen properly, otherwise the GARCH effect will not be initiated. So far we have not succeeded to employ the maximum-likelihood method for this task, although in principle this should be possible. A similar remark applies to the estimation of variances for inverse solutions, i.e., the estimation of error intervals; while for the spatiotemporal Kalman filter with homogeneous covariance it was shown to be possible to obtain credible estimates of the error (as demonstrated in [20]), this has not yet been achieved for the generalisation to GARCH modelling of covariance. These issues remain to be addressed by future work.

References

[1] S. Baillet, J. Mosher, R. Leahy, IEEE Signal Proc. Mag. 18 (6) (2001) 14.
[2] J. Trampert, Inverse Problems 14 (1998) 371.
[3] D.L.T. Anderson, J. Sheinbaum, K. Haines, Rep. Prog. Phys. 59 (1996) 1209.
[4] G. Vollmer, H. Krüger, Phys. Lett. A 28 (1968) 165.
[5] K. Chadan, P.C. Sabatier, Inverse Problems in Quantum Scattering Theory, second ed., Springer, Berlin, 1989.
[6] D.K. Saldin, A. Seubert, K. Heinz, Phys. Rev. Lett. 88 (2002) 115507.
[7] J.P. Kaipio, P.A. Karjalainen, E. Somersalo, M. Vauhkonen, Ann. N.Y. Acad. Sci. 873 (1999) 430.
[8] S. Ciulli, M.K. Pidcock, C. Sebu, Phys. Lett. A 325 (2004) 253.
[9] M.G. Rozman, M. Utz, Phys. Rev. E 63 (2001) 066701.
[10] T. Murayama, Phys. Rev. E 69 (2004) 035105.
[11] A. Neumaier, SIAM Rev. 40 (1998) 636.
[12] P.C. Hansen, B.H. Jacobsen, K. Mosegaard, Methods and Applications of Inversion, Lecture Notes in Earth Science, vol. 92, Springer, Berlin, 2000.
[13] R.D. Pascual-Marqui, C.M. Michel, D. Lehmann, Int. J. Psychophysiol. 18 (1994) 49.
[14] S. Baillet, L. Garnero, IEEE Trans. Biomed. Engrg. 44 (1997) 374.
[15] U. Schmitt, A.K. Louis, C. Wolters, M. Vauhkonen, Inverse Problems 18 (2002) 659.
[16] O. Yamashita, A. Galka, T. Ozaki, R. Biscay, P.A. Valdés-Sosa, Human Brain Mapping 21 (2004) 221.
[17] M.S. Grewal, A.P. Andrews, Kalman Filtering: Theory and Practice Using MATLAB, Wiley, New York, 2001.
[18] D. Baroudi, J. Kaipio, E. Somersalo, Inverse Problems 14 (1998) 799.
[19] E. Somersalo, A. Voutilainen, J.P. Kaipio, Inverse Problems 19 (2003) 1047.
[20] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, P.A. Valdés-Sosa, NeuroImage 23 (2004) 435.
[21] R.F. Engle, Econometrica 50 (1982) 987.
[22] T. Bollerslev, J. Econometrics 31 (1986) 307.
[23] A. Harvey, E. Ruiz, E. Sentana, J. Econometrics 52 (1992) 129.
[24] N. Shephard, Statistical aspects of ARCH and stochastic volatility, in: D.R. Cox, D.V. Hinkley, O.E. Barndorff-Nielsen (Eds.), Time Series Models in Econometrics, Finance and Other Fields, Chapman & Hall, London, 1996, pp. 1–67.
[25] D.B. Nelson, Econometrica 59 (1991) 347.
[26] J.C. Mazziotta, A. Toga, A.C. Evans, P. Fox, J. Lancaster, NeuroImage 2 (1995) 89.
[27] J.J. Riera, M.E. Fuentes, P.A. Valdés, Y. Ohárriz, Inverse Problems 14 (1998) 1009.