On Adaptive Optimization of Filter Performance Based on Markov Representation for Output Prediction Error

Hong Son Hoang and Rémy Baraille

Abstract—This paper addresses the problem of how one can improve the performance of a non-optimal filter. First the theoretical question on dynamical representation for a given time correlated random process is studied. It will be demonstrated that for a wide class of random processes, having a canonical form, there exists a dynamical system equivalent in the sense that its output has the same covariance function. It is shown that the dynamical approach is more effective for simulating and estimating a Markov and non-Markovian random processes, computationally is less demanding, especially with increasing of the dimension of simulated processes. Numerical examples and estimation problems in low dimensional systems are given to illustrate the advantages of the approach. A very useful application of the proposed approach is shown for the problem of state estimation in very high dimensional systems. Here a modified filter for data assimilation in an oceanic numerical model is presented which is proved to be very efficient due to introducing a simple Markovian structure for the output prediction error process and adaptive tuning some parameters of the Markov equation.

Keywords—Statistical simulation, canonical form, dynamical system, Markov and non-Markovian processes, data assimilation.

I. INTRODUCTION

THIS paper addresses the problem of how one can improve the performance of a non-optimal filter by introducing a new stochastic difference equation (SDE) for the output prediction error (or simply the innovation) process and adaptively estimating some parameters of this SDE to minimize the prediction error. It is commonly well known fact, truly optimal filter like a Kalman filter (KF) does not exist in practice when one applies it to solving engineering problems. As a consequence, the innovation process (IP) resulting from the application of the KF, does not form a white process and it remains time correlated. The question we are interested in is how one can improve the performance of the filter using a more complicated model for the innovation sequence ? For example, if we are given, as supposed in many studies [1],[2],[3], that the observation noise is time correlated and is modelled by some Markovian process, the optimal in minimum mean square (MMS) filter can be written out and the resulting IP satisfies the relation which is very close to some SDE (Markovian). In [3] an approximate approach to this problem has been studied by introducing first a structure

of the filter. Next a sub-optimal MMS filter (i.e. an unbiased minimum variance filter in a chosen class of filters) is obtained and studied. However, as seen from the algorithms for the filters in [1],[2],[3], their application requires still too much information on the statistics of the entering random processes. Other difficulties are related to expensive computational cost and impractical implementation, especially for very high dimensional systems.

In the present paper first we are studying the theoretical question on which process could potentially be represented by some SDE (Markovian or no Markovian). It will be shown that for the class of RPs having CovF with separable variables, it is possible to write out all the parameters of the SDE (including the statistics of the entering RPs). By this way one can well approximate covariance functions (CovFs), with the objective to better generate the samples or to produce their estimates on the basis of available observations. We will restrict our attention to the class of RPs having CovF with separable variables (for details, see section 2). Note that this class of RPs, in some sense, is equivalent to the class of RPs which can be represented in the canonical form (CF) [7]. Namely, let $\zeta(t)$ be an p-dimensional zero-mean vector RP. Given its covariance $K_\zeta(t, \tau)$, we are interested in solving two following problems : The first is to generate a zero-mean RP $\zeta_a(t)$ whose covariance matches the given covariance $K_\zeta(t, \tau)$, either exactly or approximately. The second quantity of interest is a low-rank approximation to the covariance. That is, one is interested in computing, for a given number of components, such representation that fits best the covariance $K_\zeta(t, \tau)$.

The paper is organized as follows. In section 2 the CF theory is presented briefly. In fact the CF theory is closely related to differential eigenvalue/eigenvector problems. To avoid solving this type of problems, one of practical widely used algorithms for computing the coefficients and coordinate functions of the CF based on LU factorization is given in this section (see [7]). Section 3 summarizes the procedures for construction of dynamical systems (DSs) for Markov and non-Markovian RPs conditioned that their CovFs are separable and known. Computational complexity of two approaches CF and DS is clearly seen from examining simple numerical examples in section 4. The performances of CF and DS approaches are compared in section 5 by simulation studies. Here two particular problems, one is a simulation of samples for an RP, another is related to filtering problem, are considered. It is shown that applying the algorithms CF and DS for solving these two problems offers a superior performance of the DS approach

Dr. H. S. Hoang is with Research the Department of The French Naval Oceanographic and Hydrographic Service (SHOM/HOM), 42 Avenue Gaspard Coriolis, 31100 Toulouse, France (phone: +33 5 67 76 68 23, fax: +33 5 61 40 52 58, Email: hhoang@shom.fr).

Dr. R. Baraille is with Research the Department of The French Naval Oceanographic and Hydrographic Service (SHOM/HOM), 42 Avenue Gaspard Coriolis, 31100 Toulouse, France (Email: remy.baraille@shom.fr)

World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
Vol:6, No:11, 2012

with respect to the CF approach. Application of this approach to improving the performance of the very high dimensional filter for data assimilation in an oceanic numerical model is also presented. It will be shown that by introducing a class of Markovian SDE for approximating a time correlated output prediction error and adaptive minimizing the error variance, it is possible to significantly improve the filter performance at low computational cost. The conclusions are given in section 6.

II. CANONICAL FORM OF RANDOM PROCESS

A. Eigen-structure of covariance and CF

Let (Ω, A, P) be some probability space, $L_2 = L_2(\Omega, A, P)$ be a Hilbert space of real variables induced by the inner product $< ., . >$, i.e. if $\zeta \in L_2$ then $E\zeta^2 < \infty$ and $< \zeta, \eta >= E(\zeta, \eta), \forall \zeta, \eta \in L_2$. Then the distance between two random variables ζ and η is $d(\zeta, \eta) = [E(\zeta - \eta)^2]^{1/2}$.

Let $\zeta(t)$ be a process with $E[\zeta^2(t)] < \infty$. Consider the collection of real-valued RPs $\zeta(t)$ for which $\int_0^T E[\zeta^2(t)]dt < \infty$. Stochastic processes in this set are verified to constitute a linear vector space denoted as $L_2[0,T]$ with the inner product $< \zeta(t), \eta(t) >_L = E[< \zeta(t), \eta(t) >_t], < \zeta(t), \eta(t) >_t := \int_0^T \zeta(t)\eta(t)dt$.

One of the most interesting results of the theory of RPs is that the normed vector space for processes previously defined is *separable*. Consequently, there exists a complete (and orthonormal) set $\varphi_i(t), i = 1, 2, ...$ of deterministic (non-random) functions which constitutes a basis such that

$$\zeta(t) = \sum_{k=1}^{\infty} \xi_k \varphi_k(t), \xi_k = < \zeta(t), \varphi_k(t) >_t . \quad (1)$$

More precisely the following result holds

Lemma 1: (Karhunen-Loève theorem [5])

A measurable continuous in quadratic mean RP defined over a probability space (Ω, A, P) can be represented in the form (1) (for $E\zeta(t) = 0$). In (1) ξ_k is a sequence of uncorrelated random variables, $E(\xi_k^2) = \lambda_k$, where λ_k and $\varphi_k(t)$ are the eigenvalue and eigenvector of the CovF $K_\zeta(t, \tau)$,

$$\int_{[0,T]} K_\zeta(t, \tau)\varphi_k(\tau)d\tau = \lambda_k \varphi_k(t),$$
$$K_\zeta(t, \tau) = E[\zeta(t)\zeta(\tau)] = \sum_k \lambda_k \varphi_k(t)\varphi_k(\tau). \quad (2)$$

According to [7], the CF of $\zeta(t)$ is called any representation for $\zeta(t)$ in the form of its mathematical expectation and the sum of mutually uncorrelated elementary RPs $Y_k(t) = v_k x_k(t)$ where v_k is a random variable, $x_k(t)$ is a deterministic function,

$$\zeta(t) = \bar{\zeta}(t) + \sum_{k=1}^{n} v_k x_k(t), \bar{\zeta}(t) = E[\zeta(t)] \quad (3)$$

In (3) v_k are the coefficients of CF, $x_k(t)$ - coordinate functions. Thus the n-truncated approximation of (1) $\zeta_n(t) = \sum_{k=1}^{n} \xi_k \varphi_k(t)$ is an CF for $\zeta(t)$.

Representation of the RP in the form of CF is a very convenient method for performing different operations with random functions, especially for linear random functions. In the CF only its coefficients are random variables.

B. One practical method

Let $\{v_k\}$ be a set of random variables, $E(v_i) = 0, E(v_i v_j) = \sigma_i^2 \delta_{ij}$, where δ_{ij} be the Kronecker symbol. Suppose we would like to approximate the process $\zeta(t)$ by $\zeta(t) \approx \zeta_n(t) \approx \sum_{k=1}^{n} v_k x_k(t)$.

By solving $min\ arg\ _{x_k(t)}\ E[\zeta(t) - \sum_{k=1}^{n} v_k x_k(t)]^2$ one obtains $x_k^0(t) = E[\zeta(t)v_k]/\sigma_k^2$. It remains to determine the variables v_k. Generally speaking, v_k should be such that the sum must be convergent as $n \rightarrow \infty$ (see Lemma 1 for ξ_k). To obtain a more attractive algorithm, let $v_1^n := (v_1, ..., v_n)^T$, $A = |a_{ij}|_{i,j=1}^{n}$, $y_1^n := [\zeta(t_1), ..., \zeta(t_n)]^T$. Then we want to find v_1^n as $v_1^n = Ay_1^n$. The coefficients a_{ij} must be determined as a solution of $AK_\zeta(n)A^T = D$, $D := diag(\sigma_2^1, ..., \sigma_2^n), K_\zeta(n) := |K_\zeta(t_i, t_j)|_{i,j=1}^{n}$. One great disadvantage of this choice is that the system for A must be solved each time when we want to involve more values of ζ. A recursive algorithm can be obtained by applying the LDU algorithm [8] by assuming $a_{ii} = 1, a_{ij} = 0, j > i$. We have

Lemma 2: [7] Consider the RP $\zeta(t)$. Then the optimal in mean square CF $\zeta^0(t)$ is given by

$$v_1 = \zeta_1, \sigma_1^2 = K_\zeta(t_1, t_1), x_1^0(t) = \frac{1}{\sigma_1^2} K_\zeta(t, t_1),$$
$$v_k = \zeta_k - \sum_{j=1}^{k-1} v_j x_j^0(t_k),$$
$$\sigma_k^2 = K_\zeta(t_k, t_k) - \sum_{j=1}^{k-1} \sigma_j^2 [x_j^0(t_k)]^2,$$
$$x_k^0(t) = \frac{1}{\sigma_k^2}[K_\zeta(t, t_k) - \sum_{j=1}^{k-1} \sigma_j^2 x_j^0(t)x_j^0(t_k)],$$
$$P(t) = E[\zeta(t) - \sum_{k=1}^{n} v_k x_k^0(t)]^2 =$$
$$K_\zeta(t, t) - \sum_{k=1}^{n} \sigma_k^2 [x_k^0(t)]^2.$$

C. CF and RP with separable variables

The fact that the RP $\zeta(t)$ has a CF implies that its CovF has the following form

$$K_\zeta(t, \tau) = E[\sum_{i,j} v_i v_j x_i(t)x_j(\tau)] = \sum_{i=1}^{n} \sigma_i^2 x_i(t)x_j(\tau) \quad (4)$$

For $K_i(t) := x_i(t), K_i(\tau) := \sigma_i^2 x_i(\tau)$ we have

$$K_\zeta(t, \tau) = \sum_{i=1}^{n} K_i(t)K_i(\tau) \quad (5)$$

Thus the CovF $K_\zeta(t, \tau)$ as a function of t and τ is a function of separable variables. The RPs with CovF of the form (5) is investigated in [6]. We shall call (5) an CovF with separable variables. Thus when the RP $\zeta(t)$ can be represented in the CF, it has the CovF belonging to the class of CovF with separable variables. We note that all the results related to the CFs, CovFs of the scalar RPs can be extended to the vector RPs [7].

III. DYNAMICAL SYSTEM (DS) APPROACH

It turns out that the RP $\zeta(t)$ having an CF, can be generated as an output of some DS. The DS approach can serve as an efficient tool for simulating samples of the considered RP as well as for solving the estimation problems with correlated noises as studied in [1].

World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
Vol:6, No:11, 2012

A. DS for a non-Markovian RP

Theorem 1 ([6]) *Let* $\{w_i^{(1)}\}$ *be an* n_1-*dimensional RP with zero mean and CF*

$$Q_1(i,j) = K_1^{(1)}(i)P_1^{(1)}(j) \tag{6}$$

where $K_1^{(1)}(i)$ *is non-singular. Then (6) is necessary and sufficient for* $\{w_i^{(1)}\}$ *to be generated as output of*

$$w_{i+1}^{(1)} = C_i^{(1)}w_i^{(1)} + w_i^{(10)}, \tag{7}$$

$\{w_i^{(10)}\}$ *is a white RP uncorrelated with* $w_0^{(1)}$,

$$E(w_i^{(10)}) = 0, E[w_i^{(10)}(w_i^{(10)})^T] = Q_{10}(i)\delta_{ij} \tag{8}$$

For a given (6), the matrix $C_i^{(1)}$ *in (8) is determined by*

$$C_i^{(1)} = K_1^{(1)}(i+1)[K_1^{(1)}(i)]^{-1} \tag{9}$$

and $Q_{10}(i) = K_1^{(1)}(i+1)\{P_1^{(1)}(i+1)[(K_1^{(1),T}(i+1))]^{-1} - P_1^{(1)}(i)[K_1^{(1),T}(i+1)]^{-1}\}K_1^{(1),T}(i+1)$, $K_1^{(1)}(0) = I$ *is the unit matrix.*

Theorem 2 ([6]) *Let* $\zeta_i := w_i^{(2)}$ *be an* n_2-*dimensional RP with zero mean and covariance*

$$K_{w^{(2)}}(i,j) = Q_2(i,j) := \sum_{l=1}^{2} K_l^{(2)}(i)P_l^{(2)}(j), \tag{10}$$

where $K_1^{(2)}(i), P_1^{(2)}(i)$ *are* $(n_2 \times n_2)$ *matrices,* $K_1^{(2)}(i)$ *is non-singular;* $K_2^{(2)}(i), P_2^{(2)}(i)$ - $(n_2 \times n_1)$ *matrices. Let* $w_i^{(1)}$ *be an RP being a solution of (7) and*

$$K_2^{(2)}(i)P_1^{(1)}(i) = P_2^{(2),T}(i)K_1^{(1),T}(i) \tag{11}$$

Then (10) is necessary and together with (11) are sufficient for $w_i^{(2)}$ *to be presented as*

$$w_{i+1}^{(2)} = C_i^{(2)}w_i^{(2)} + D_i^{(2)}w_i^{(1)} + w_i^{(20)} \tag{12}$$

$C_i^{(2)}, D_i^{(2)}$ *are the matrices of dimensions* $(n_2 \times n_2)$ *and* $(n_2 \times n_1)$ *respectively,* $w_0^{(1)}, w_0^{(2)}, w_i^{(10)}, w_i^{(20)}$ *are of zero mean and uncorrelated,*

$$E[w_i^{(k0)}w_j^{(l0),T}] = W_{kl}(i)\delta_{ij},$$
$$K_1^{(2)}(0) = I, K_2^{(2)}(0) = 0.$$

For a given (10), the matrices $C_i^{(2)}, C_i^{(1)}, D_i^{(2)}$ *and the RPs* $w_i^{(10)}, w_i^{(20)}$ *are determined by*

$$C_i^{(2)} = K_1^{(2)}(i+1)[K_1^{(2)}(i)]^{-1},$$
$$D_i^{(2)} = -C_i^{(2)}K_2^{(2)}(i)[K_1^{(1)}(i)]^{-1} + K_2^{(2)}(i+1)[K_1^{(1)}(i)]^{-1},$$
$$C_i^{(1)} = K_1^{(1)}(i+1)[K_1^{(1)}(i)]^{-1},$$
$$W_{22}(i) = K_1^{(2)}(i+1)\{P_1^{(2)}(i+1)[(K_1^{(2),T}(i+1))]^{-1} - P_1^{(2)}(i)[K_1^{(2),T}(i+1)]^{-1}\}K_1^{(2),T}(i+1)i,$$
$$W_{11}(i) = K_1^{(1)}(i+1)\{P_1^{(1)}(i+1)[(K_1^{(1),T}(i+1))]^{-1} - P_1^{(1)}(i)[K_1^{(1),T}(i+1)]^{-1}\}K_1^{(1),T}(i+1),$$
$$W_{21}(i) = K_1^{(2)}(i+1)\{P_1^{(1)}(i+1)[(K_1^{(1),T}(i+1))]^{-1} - P_2^{(1)}(i)[K_1^{(1),T}(i+1)]^{-1}\}K_1^{(1),T}(i+1).$$

IV. COMPARISON BETWEEN CF AND DS APPROACHES: EXAMPLES

A. Stationary process

1) Application of Lemma 1: Consider the RP with zero mean and CovF

$$K_\zeta(t,\tau) = K_\zeta(t-\tau) = c^2 e^{-\beta|t-\tau|}, t, \tau \in [0,T] \tag{13}$$

Generally speaking, it is impossible to find analytically the system of eigenvalues $\{\lambda_k\}$ and eigenvectors $\{x_k(t)\}$ of $K_\zeta(t,\tau)$. However, in this particular case one can prove [7]

$$\zeta(t) = \sum_k v_k x_k(t), x_k(t) = \frac{2}{T+\lambda_k}sin[\omega_k(t-\frac{T}{2}) + \frac{k\pi}{2}],$$
$$\sigma_k^2 = \frac{c^2}{2}\lambda_k(T+\lambda_k), \lambda_k = c^2\frac{2\beta}{\beta^2+\omega_k^2},$$

where ω_k are the positive roots of the equation, $tan\omega T = -\frac{2\beta\omega}{\beta^2-\omega^2}$.

2) Application of Lemma 2: Suppose it is impossible to obtain the system of eigenvalues and eigenvectors of $K_\zeta(t,\tau)$. To approximate the process $\zeta(t)$, let us first construct the optimal in mean square CF based on the values of $\zeta(t)$ at the points $t_i = (i-1)\Delta t$. For $\rho := e^{-\beta\Delta t}$, we have

$$v_1 = \zeta(t_1), \sigma_1^2 = c^2, x_1^0(t) = e^{-\beta|t|},$$
$$\sigma_2^2 = c^2(1-\rho^2), x_2^0(t) = \frac{1}{1-\rho^2}[e^{-\beta|t-\Delta t|} - \rho e^{-\beta|t|}],$$
$$\sigma_3^2 = c^2 - \sum_{k=1}^{2}\sigma_k^2[x_k^0(t_3)]^2,$$
$$x_3^0(t) = \frac{1}{\sigma_3^2}[e^{-\beta|t-2\Delta t|} - \sum_{k=1}^{2}\sigma_k^2 x_k^0(t)x_k^0(t_3)^2],$$

where $x_1^0(t_3) = e^{-2\beta\Delta t} = \rho^2$, $x_2^0(t_3) = \frac{1}{1-\rho^2}[e^{-\beta\Delta t} - \rho e^{-\beta 2\Delta t}] = \rho$. The coefficients σ_k^2 and coordinate functions $x_k^0(t)$ can be computed in the similar manner for $k = 4, 5, ...$ The above algorithm is written out here to see its complexity when dealing with an arbitrary RP. In this example, there exist more compact formulas

$$\sigma_k^2 = c(1-\rho^2), x_k^0(t) = $$
$$\frac{e^{-\beta|t-(k-1)\Delta t|} - \rho e^{t-\beta|t-(k-2)\Delta t|}}{1-\rho^2}. \tag{14}$$

3) Application of Theorem 1: It is easy to check that (13) is equivalent to (6) subject to

$$t \geq \tau : K_1(t) = e^{-\beta t}, P_1(\tau) = c^2 e^{\beta\tau},$$
$$t < \tau : K_1(t) = e^{\beta t}, P_1(\tau) = c^2 e^{-\beta\tau} \tag{15}$$

This choice ensures $K_1(0) = 1$. Consider the case $t \geq \tau$. From Theorem 1,

$$C_i^{(1)} = \rho = e^{-\beta(t_{i+1}-t_i)},$$
$$Q_{10}(i) = c^2(1 - e^{-\beta\Delta t}) = c^2(1-\rho^2). \tag{16}$$

The recursive equation for simulating $\zeta_i = \zeta(t_i)$ is

$$\zeta_{i+1} = \rho\zeta_i + w_i, \tag{17}$$

w_i is an uncorrelated sequence of zero mean and variance $Q_{10}(i)$ defined in (16).

World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
Vol:6, No:11, 2012

B. Non-stationary random process

Consider the non-stationary RP with zero mean and covariance

$$K_\zeta(t,\tau) = c^2 e^{-\beta_t t} e^{-\beta_\tau \tau} = K_1(t)P_1(\tau),$$
$$K_1(t) = e^{-\beta_t t}, P_1(\tau) = c^2 e^{-\beta_\tau \tau}, \quad (18)$$

which ensures $K_1(0) = 1$.

1) Application of Lemma 2: The optimal in mean square CF based on the values of $\zeta(t)$ at the points $t_i = (i-1)\Delta t, i = 1, ..., n$ is expressed by ($\rho(t) := e^{-\beta_t \Delta t}$),

$$v_1 = \zeta(t_1), \sigma_1^2 = c^2, x_1^0(t) = \frac{e^{-\beta_t t}}{c^2},$$
$$\sigma_2^2 = (c^2 - 1)\rho^2(t_1), x_2^0(t) = \frac{(c^2+1)\rho(t_2)e^{-\beta_t t}}{\rho^2(t_1)}$$
$$etc \quad (19)$$

For this simple example, write out here the formulas for $\sigma_k, x_k^0(t)$ for all $k = 3, 4, ...$ is a hard task, not to say they may not be realizable.

2) Application of Theorem 1: From Theorem 1 it follows $Q_1(0,0) = c^2$ and

$$w_{i+1}^{(1)} = C_i w_i^{(1)}, C_i = e^{-\alpha_i \Delta t},$$
$$\alpha_i = i\Delta\beta_i + \beta_i, \Delta\beta_i := \beta_{i+1} - \beta_i, \beta_i := \beta_{t_i}. \quad (20)$$

C. Non-Markovian random process

Consider the RP $\zeta(t)$

$$\zeta(t) = \sum_{k=1}^{2} \xi_k f_k(t), \quad (21)$$

$$\xi = (\xi_1, \xi_2)^T, E(\xi) = 0, E(\xi\xi^T) = \Xi,$$
$$\Xi = [\Xi_{ij}], i, j = 1, 2,$$
$$\Xi_{11} = \Xi_{22} = 2, \Xi_{12} = \Xi_{21} = 1,$$
$$f_1(t) = \frac{1}{1+t}, f_2(t) = \frac{1}{1+t^2}.$$

This process $\zeta(t)$ has the covariance

$$K_\zeta(t,\tau) = f_1(t)[2f_1(\tau) + f_2(\tau)] + f_2(t)[2f_2(\tau) + f_1(\tau)] \quad (22)$$

1) Application of Lemma 2: For $t_i = (i-1)\Delta t$, application of Lemma 2 subject to (22) yields

$$\sigma_1^2 = 6, x_1^0(t) = 3[1/(1+t) + 1/(1+t^2)],$$
$$\sigma_2^2 = K_\zeta(t_2, t_2) - \sigma_1^2[x_1^0(t_2)]^2,$$
$$K_\zeta(t_2, t_2) = 2[1/(1+\Delta t^2) + 1/(1+\Delta t^2)^2 + 1/((1+\Delta t)(1+\Delta t^2))],$$
$$x_2^0(t) = [K_\zeta(t, t_2) - 4x_1^0(t)x_1^0(t_2)]/\sigma_2^2,$$
$$x_1^0(t_2) = 3[1/1 + \Delta t) + 1/(1+\Delta t^2)]...,$$
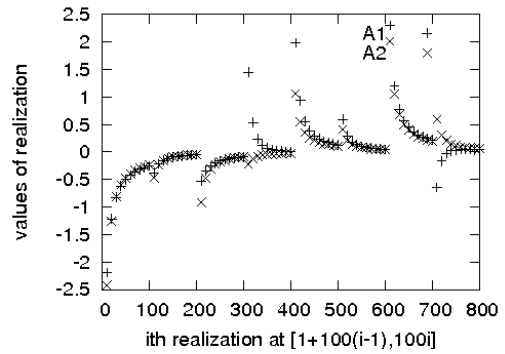$$\quad (23)$$

Fig.1. Realizations produced by A1 and A2. The realizations $4^{th}, 5^{th}, 8^{th}$ of A2 are far too much from the true A1.

2) Application of Theorem 2: To make the choice of $K_1^{(1)}(t), P_1^{(1)}(t)$ easier (since $K(t,\tau) = K_1^{(1)}(t)P_1^{(1)}(\tau)$ must be a covariance function and satisfy (11)), represent (21) in the form

$$K_\zeta(t,\tau) = \sum_{i=1}^{2} K_i^{(2)}(t)P_i^{(2)}(\tau), \quad (24)$$

$$K_i^{(2)}(t) = P_i^{(2)}(t) = x_i(t), i = 1, 2,$$
$$x_1(t) = a_{11}/(1+t) + a_{21}/(1+t^2),$$
$$x_2(t) = a_{22}/(1+t^2),$$
$$a_{11} = 1,414, a_{12} = 0, a_{21} = 0,707, a_{22} = 1.225.$$

The representation (24) is obtained by applying the Cholesky algorithm [8] to Ξ which results in $\Xi = AA^T$. From Theorem 3 one obtains the equations for simulating of $\zeta(t) := w_i^{(2)}$, noticing that by the choice $K_1^{(1)}(t) = P_1^{(1)} = 1$ the condition (11) is automatically satisfied,

$$w_{i+1}^{(2)} = C_i^{(2)} w_i^{(2)} + D_i^{(2)} w_i^{(1)}, \quad (25)$$

$$C_i^{(2)} = \frac{x_1(i+1)}{x_1(i)}, D_i^{(2)} = -C_i^{(2)} x_2(i) + x_2(i+1),$$
$$x_1(i) := \frac{a_{11}}{(i-1)\Delta t} + \frac{a_{21}}{1+[(i-1)\Delta t]^2},$$
$$x_2(i) := \frac{a_{22}}{1+[(i-1)\Delta t]^2},$$
$$w_1^{(1)}(i+1) = w_1^{(1)}(i), i = 1, 2, ...,$$
$$E[w_0^{(2)} w_0^{(1)}] = 0, Q_1(0,0) = 1, Q_2(0,0) = 6.$$

It is seen that $\zeta(t)$ is not Markov since $w_i^{(1)} = w_0^{(1)}, \forall i$ is not an uncorrelated sequence.

V. SIMULATION STUDIES

A. Simulation of process (21)

Three following algorithms will be applied to simulate the realizations of (21):

(i) Using $\zeta_1(t_i) = \sum_{k=1}^{n} \xi_k f_k^0(t_i), n = 2, t_i = (i-1)\Delta i, i = 1, ..., 100; \Delta = 0.1$. Concretely, we simulate $\xi = Av$ where A is given in (24). As to $v = (v_1, v_2)^T$, it is normally distributed with zero mean and the unit covariance (Algorithm 1 - A1). These realizations can serve as "true" realizations (the variance of v_2 is $\sigma_2^2 = 0.00328$ hence it is enough to use $n = 2$).

(ii) Algorithm in Lemma 2: $\zeta_2(t_i) = \sum_{i=1}^{n} \epsilon_i x_i^0(t_i), \epsilon_i = \sigma_i v_i$, subject to covariance function (22). Mention that the

World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
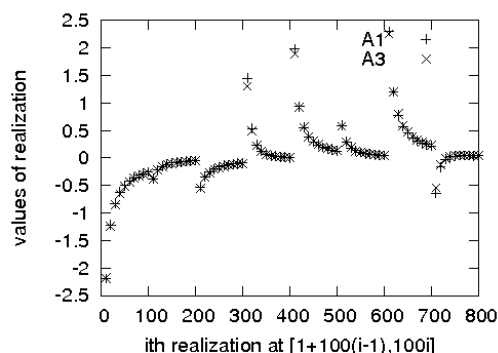Vol:6, No:11, 2012

Fig.2. Realizations produced by A1 and A3 : All realizations of the A3 are close to that of A1.

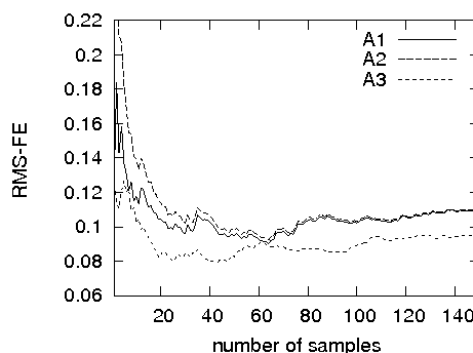
Fig.3. Ensemble averaged RMSs of the FE produced by A1, A2 and A3. The prediction is made over $\Delta t = 1$. The algorithm A3 has produced the better estimates.

realizations of $v = (v_1, v_2)^T$ are taken to be that obtained in (i) (Algorithm 2 - A2).

(iii) Algorithm (24)-(25) (DS approach). Here the realizations of $(w_0^{(1)}, w_0^{(2)})^T$ are taken as $(w_0^1, w_0^2)^T = v$, v is given in (i) (Algorithm 3 - A3).

The simulation results are shown partly in Figs 1-2. Here 8 realizations produced by each of three algorithms A1, A2 and A3 are shown. All three algorithms have the same initial realization $\zeta_k(1), k = 1, 2, 3$. One sees from Fig. 2 that the algorithms A1, A3 have produced almost the same realizations whereas the realizations 4th, 5th, 8th of A2 in Fig. 1 are far away from that of A1, especially at the beginning.

B. Estimation of random process (21)

Suppose the observations are given according to

$$z(t_{io}) = \zeta(t_{io}) + \nu(t_{io}), io = 1, 2, ...,$$
$$t_{io} = 1 + (io - 1)\Delta T, \Delta T = 10\delta t, \delta t = 0.1, \quad (26)$$

with $\zeta(t_{io})$ being given by (21) and $\nu(t_{io})$ is an uncorrelated sequence of observational error with zero mean and variance $\sigma_\nu^2 = 1$. Assume further that $\nu(t_{io})$ and ξ in (21) is uncorrelated. The problem is to estimate the value of $\zeta(t_i)$. ΔT symbolizes the interval between two observations (in hours, days).

1) Filtering problem: (i.1) Algorithm A1. From (21)(26) one has

$$z(t_{io}) = H(t_{io})\theta + \nu(t_{io}), io = 1, 2, ...,$$
$$H(t_{io}) = [f_1(t_{io}), f_2(t_{io})], \theta := (\theta_1, \theta_2)^T,$$
$$E(\theta) = 0, E(\theta\theta^T) = \Xi.$$

The optimal in minimum mean square (MMS) estimation procedure can be written out for θ.

(i.2) Algorithm A2. For the model (23) the algorithm remains the same as A1 with the differences

$$H(t_{io}) = [x_1^0(t_{io}), x_2^0(t_{io})], P(t_1) = diag[\sigma_1^2, \sigma_2^2].$$

(i.3) Algorithm A3. Consider the DS (25) along with the observation system. Introducing $y(t_i) = (w_i^{(2)}, w_i^{(1)})^T$ leads to the filtering problem in state-space form

$$y(t_{io+1}) = C_{io}^e y(t_{io}), y(t_1) := (w_1^{(2)}, w_1^{(1)})^T,$$
$$C^e(t_{io}) = |c_{kl}^e|_{k,l=1}^2, c_{11}^e = C_{t_{io}}^{(1)}, c_{12}^e = D_{t_{io}}^{(1)},$$

$$c_{21}^e = 0, c_{22}^e = 1,$$
$$z(t_{io}) = H^e y(t_{io}) + \nu(t_{io}), H^e = [I, 0], io = 1, 2, ...,$$

from which one can write out the MMS filter.

2) Numerical results: In Fig. 3 we show the ensemble averaged RMSs of the filtered error (FE) produced by three algorithms A1, A2, A3. It is undoubted that the DS approach (A3) produces the best estimates. The algorithm A1 is slightly better than A2 especially at the beginning, but in general they behave in the same way. Noticing that all three algorithms have filtered well the observation noise since its variance is equal 1. It means that the algorithms reduce about 90 % noise level in the estimates.

C. Application to oceanic data assimilation

The objective of oceanic data assimilation is to estimate the ocean state and to produce its best forecast for the period of interest (10 days, for example) using a numerical model (NM) and available observations. Last years the satellite sea surface height (SSH) is one of the most important sources of observations. Due to very high dimension of the numerical model and large set of observations (orders of 10^{6-7} and 10^{4-5} respectively), non-linearities of NM ... only approximate filters can be applied. Moreover, in the context of satellite SSH data, we are given only along-track observations which are irregular in space and in time. In this situation, estimation of the velocity requires to interpolate the IP over the domain of interest which introduces inevitable errors in the innovation vector. As a consequence, the resulting "innovation" sequence $\zeta(t)$ is rather time-correlated even given noise-free observations.

In order to improve the performance of the filter, we assume the hypothesis that the innovation forms a correlated sequence and, more concretely, it assumes a Markov model. Thus the innovation sequence is described by a SDE. Mention that even in the case when the observation noise satisfies a Markov SDE, this structure for the innovation does not follow in the optimal MMS filter (for example, see [1],[2],[3] ...). Table I displays the estimation errors produced by two filters, one is the filter based on principle of conservation of the potential velocity (denoted as CHF - Cooper-Haines Filter, see [4]) and the other (denoted as MIF - Markov Innovation Filter) is based on the

World Academy of Science, Engineering and Technology
International Journal of Physical and Mathematical Sciences
Vol:6, No:11, 2012

TABLE I
TIME AVERAGE RMS PREDICTION ERRORS

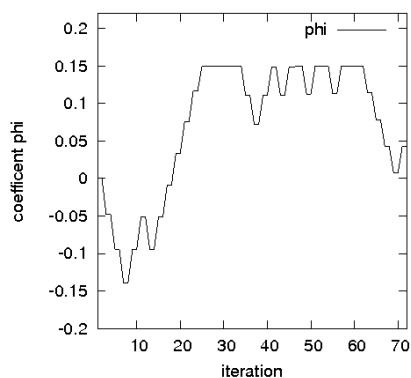| Filter	SSH	RME-PE	RMS-FE
CHF	6.12 cm	7.39 cm/s	6.13 cm/s
MIF	5.47 cm	6.91 cm/s	5.8 cm/s
Reduction	11 %	6 %	5.7 %

Fig.4. Tuning coefficient ϕ during assimilation

assumption that the SSH innovation is not white (i.e. $\zeta(t+1) \neq \nu(t)$) and rather it is a Markov process, i.e. $\zeta(t+1) = \phi\zeta(t) + \nu(t)$. Here for simplicity, ϕ is assumed to be an unknown constant, $\nu(t)$ is a white noise sequence. The parameter ϕ is updated during assimilation to minimize the variance of $\nu(t)$. The experiment has been carried out using the noise-free SSH data (3 years) and the NM Micom [4] to model the circulation in North Atlantic. In Table I, RMS-PE signifies Root Mean Squares of the Prediction Error, RMS-FE - Root Mean Squares of the Filtered Error. It is seen that by introducing a simple structure of Markov process for the SSH innovation, one can reduce 11 % of the PE for the SSH variable and about 6% for the velocity error. Fig. 4 shows how the parameter ϕ changes during adaptation process. The value initial $\phi = 0$ corresponds to the hypothesis that the innovation process is white as in the CHF.

VI. CONCLUSION

In this paper we have attempted to give a broad-based review of what we consider the most important aspects of the approach for the construction of a DS model for a RP given its CovF with separation of variables. We have tried to emphasize the possibility to apply this approach to the specific applications like simulation of RPs as well as state estimation problem. The advantages of this approach compared with the traditional approach known as CF representation are clearly seen from a number of examples and experiments presented. This approach has been also applied successfully in solving the practical engineering problem known as satellite data assimilation with the oceanic numerical model. The obtained results in this experiment are encouraging since the Markov structure is introduced directly in the representation for the innovation process regardless of the noise-free of the observations.

REFERENCES

[1] Bryson A.E. and Henrikson L.J. (1968) Estimation Using Sampled Data Containing Sequentially Correlated Noise. *Journal of Spacecraft*, Vol. 5, No. 6, pp. 662-665.
[2] T. Kailath (1968) An Innovations Approach to Least-Squares Estimation, Pt. I: Linear Filtering in Additive Noise, *IEEE Trans. Autom. Contr.*, 13(6), pp. 646-655.
[3] Hoang H.S. and Baraille R. (2011) Approximation Approach to Linear Filtering Problem with Correlated Noise. *Engineering and Technology*, 59, pp. 298-305.
[4] Hoang H.S. and Baraille R. (2012) On Gain Initialization and Optimization of Reduced-Order Adaptive Filter. *Journal IAENG IJAM*, 42:1, pp. 19-33.
[5] Loève, M. (1978) *Probability theory*. Vol. II, 4th ed. Graduate Texts in Mathematics. 46. Springer-Verlag
[6] Nguyen T.L. and Hoang H.S. (1982) On one model of non-Markovian random process and its application to optimal estimation theory. *Automat. Remote Contr.*, 43 (1982), 8 (1), pp. 1021-1032.
[7] Pugachev V.S. and Sinitsyn I.N. (1987). *Stochastic systems : theory and applications*, John Wiley and Sons.
[8] Golub G.H. and van Loan C.F. (1996). *Matrix Computations*, Cambridge University Press.