FAST RECURSIVE RECONSTRUCTION OF LARGE TIME VARYING MULTIDIMENSIONAL FIELDS

Amir Asif[1] *José M. F. Moura*[2]

[1]Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Faculty of Electronics Engineering, Topi 23460, Swabi, Pakistan

[2]Carnegie Mellon University, Department of Electrical and Computer Engineering, 5000 Forbes Avenue, Pittsburgh, PA15213, USA

ABSTRACT

We develop computationally fast and storage efficient implementations for the Kalman-Bucy filter (KBf) for data assimilation problems with large time varying multidimensional fields. We refer to them as the block KBf (bKBf) and the localized block KBf (lbKBf). For fields defined on a 2D lattice of linear dimension I, the bKBf reduces the computational complexity of the KBf by $O(I)$. The lbKBf saves further on computations by a factor of I and decreases the storage requirements by $O(I)$. We illustrate the lbKBf in assimilating satellite measurements in physical oceanography, presenting simulations for an equatorial beta plane.

1. INTRODUCTION

In this paper, we develop computationally practical and storage efficient recursive signal processing algorithms for multidimensional physical systems, modeled by partial differential equations (pde). We are motivated by problems of reconstructing image fields in physical oceanography like ocean velocity or sea surface height by coupling the underlying models describing the dynamics of the field with the sparse measurements. In Physical oceanography these problems are referred to as data assimilation, [1].

A major challenge in reconstructing *time* and *spatially* dependent fields is the large dimensionality associated with these problems. This curse of dimensionality precludes the direct application of recursive filtering algorithms like the Kalman-Bucy filter (KBf) to reasonably sized domains. We develop a computationally efficient implementation of the KBf, referred to as the block KBf (bKBf). It takes advantage of two unique aspects associated with these problems. First, the fields that we reconstruct are described by models obtained from the discretization of pdes (dpde). The dpde models are localized. This is reflected in the block banded structure of the system matrices, with blocks themselves banded and sparse. Second, the measurements are sparse. We are concerned with undersample scanned data, like altimetry data collected by an orbiting satellite.

For fields defined on a 2D lattice of linear dimensions I, the bKBf reduces the computational complexity of the KBf by a factor of I. To obtain further savings, we approximate the *error field* associated with the bKBf as a first order Gauss Markov random field (GMrf). The resulting KBf implementation is called the local block KBf (lbKBf) which reduces by an additional factor of I, the computational complexity of the bKBf. The storage requirements are $O(I^3)$, reduced from $O(I^4)$ for the KBf.

We apply the lbKBf to estimate the ocean circulation fields in physical oceanography. The physics underlying ocean circulation are captured by nonlinear differential equations, which under a standard set of assumptions are of the Navier Stokes (NS) type, [2]. Remote sensing image data of the Earth's surface, acquired from either spacecraft platforms or satellites, is used to calculate maps for the sea surface height. We combine the sparse satellite measurement with the model dynamics to estimate the ocean circulation. We generate a video depicting the circulation fields over time.

The paper is structured as follows. In section 2, we discuss the Dpde models. In section 3, we describe the multiple scanned measurements. In section 4, we derive the lbKBf which is later used for data assimilation in the equatorial beta plane. Finally, in section 6, we conclude the paper.

2. STATE MODELS: DISCRETIZED PDES

We discretize linear models of the type

$$\frac{\partial \phi}{\partial t} = \mathcal{L}\phi + g, \quad \mathcal{L} = \sum_{n_1,n_2,n_3} \frac{\partial^{n_1}}{\partial x^{n_1}} \frac{\partial^{n_2}}{\partial y^{n_2}} \frac{\partial^{n_3}}{\partial z^{n_3}} a_{n_1 n_2 n_3}(x,y,z) \quad (1)$$

where $\mathcal{L}$ is a pde linear operator and g is the forcing term which may include random effects.

In physical sciences, such models arise in various contexts. First, for an inviscid, isentropic, shallow water motion on an equatorial beta plane (the plane within latitudes of 30°), the ocean circulation is modeled by coupled pdes, which are linear and of the type (1). See [2] for details. For more general nonlinear models, we decouple the field in a deterministic and a random component. The model nonlinearities are taken care of by the deterministic component, which represents the underlying trend. Given initial and boundary conditions, the trend is constructed by integrating a deterministic nonlinear version of the model equations. The random component represents the natural variability of the field away from the deterministic component. A set of stochastic pdes obtained by linearizing the model equation, is coupled with the available data to construct the random component. These stochastic pdes are similar to (1).

In our data assimilation algorithms, we assume linear models of type (1). For nonlinear systems, we use the decoupled approach just described to linearize the random

0-8186-7919-0/97 $10.00

component of the modeling equations. In the latter case, the linear models are coupled with the measurements to estimate via the local block KBf (lbKBf) the correction needed in the deterministic predictions.

Discretizing (1) results in dpde models of the type

$$\Psi^{(k+1)} = \mathcal{A}\Psi^{(k)} + \mathcal{C}W^{(k)} \tag{2}$$

where Ψ is obtained by stacking in lexicographic order the field vectors on top of each other. Similarly, the input vector W is obtained by stacking the forcing term g values. See, for example, [3], where the linear pdes modeling the equatorial beta region are discretized by the Lax-Friedrich method. The vector $W^{(k)}$ is modeled as a zero mean, white Gauss noise with covariance

$$\mathcal{E}[W^{(k)}W^{(l)}] = Q\delta_{kl}$$

where $\mathcal{E}$ is the expectation operator. Equation (2) is the dpde model. We refer to it as the state equation. Borrowing terminology from Systems Theory, we call $\mathcal{A}$ the state matrix and $\mathcal{C}$ the input matrix which are both block banded, block Toeplitz, and sparse. The block bandwidth of the state matrix $\mathcal{A}$ depends on the order of the spatial differential operator $\mathcal{L}$ and on the numerical discretizing schemes. For pdes with up to second order partial derivatives in space and using a first or second order finite difference scheme, the state matrix $\mathcal{A}$ is block tridiagonal; with up to fourth order spatial differential operators, $\mathcal{A}$ is block pentadiagonal, and so on for higher orders. In the following discussion, we assume that the maximum order of the spatial derivatives in $\mathcal{L}$ is 2. The state matrix $\mathcal{A}$ is, therefore, block tridiagonal with the lower diagonal block entries given by B_1, the diagonal entries by B_2, and the upper diagonal entries by B_3. The input matrix $\mathcal{C}$ is assumed diagonal with blocks C.

Applications in the physical sciences usually deal with coupled fields. For multiple fields, the structure of the modeling equation, (2), remains essentially the same, except the dimensions of the field vector Ψ increase accordingly. Although the results presented below are fairly general, in our discussion, we restrict ourselves to a scalar field defined on a 2D grid with dimensions $I \times J$.

3. SCANNED MEASUREMENTS

In physical oceanography, instrumentation on board orbiting satellites measure, for example, sea surface height (SSH) or sea surface temperature (SST). We assume that during each scan the field characteristics do not change significantly. In other words, in a single pass, we assume that the satellite scan is instantaneous. On the other hand, over successive scans, the field may change significantly. The satellite scans are analogous to the scanning in the video monitors. In interlaced video displays, images are scanned every other row at a time, say, the even numbered rows, field 1, first, and then the remaining odd numbered rows, field 2. At each time instant, the observations are at spatially separated points along a single row of the tow dimensional image. Similarly, at each time instant, an orbiting satellite scans a single row of the physical field.

Assume that the satellite scans instantly N rows, say rows $m, \ldots, p$, of the field Ψ. Alternatively, we may have more than one satellite. In the latter case, the rows are isolated and far apart. Assume that N rows, say $m, n, ..., p$ with $m < n ... < p$, are scanned. The measurement model is given by

$$\begin{aligned} \mathcal{Y}^{(k)} &= \mathcal{H}^{(k)}\Psi^{(k)} + \Xi^{(k)} \\ &= \begin{bmatrix} \Theta_m^{(k)}\Psi_m^{(k)} \\ \vdots \\ \Theta_p^{(k)}\Psi_p^{(k)} \end{bmatrix} + \begin{bmatrix} \xi_m^{(k)} \\ \vdots \\ \xi_p^{(k)} \end{bmatrix}. \end{aligned} \tag{3}$$

The associated covariances are

$$\mathcal{E}[\xi_m^{(k)}\xi_p^{(l)}] = R_{mp}\delta_{mp}\delta_{kl}.$$

4. LOCALIZED BLOCK KBF

To reconstruct the field, we assimilate the measurements of section 3 to the dynamical dpde model, (2), via Kalman-Bucy filtering (KBf). Direct application of the KBf is computationally prohibitive. In the block KBf (bKBf), we partition all vector and matrix quantities in the filter according to the block structure of $\mathcal{A}$ and of the other model matrices. We partition, for example, the error covariance matrix $\mathcal{P}$ into blocks $\{P_{i_1 i_2}\}$'s which denote the cross-covariance between rows i_1 and i_2 of the error field, i.e.,

$$P_{i_1 i_2} = \mathcal{E}[(\Psi_{i_1} - \widehat{\Psi}_{i_1})(\Psi_{i_2} - \widehat{\Psi}_{i_2})^T].$$

We expand the filter equations into these partitioned quantities and exploit the sparse structure of the measurements. The predictor equations use the tridiagonal structure of the state matrix $\mathcal{A}$. The bKBf for multiscan measurements is:

Predictor update: $1 \le i \le I$

$$\widehat{\Psi}_i(k+1|k) = \sum_{\tau=1}^{3} B_\tau \widehat{\Psi}_{i+\tau-2}(k|k) \tag{4}$$

Predictor covariance update: $1 \le i_1, i_2 \le I, I$

$$\begin{aligned} P_{i_1 i_2}(k+1|k) &= \sum_{\tau_1,\tau_2=1}^{3} B_{\tau_1} P_{i+\tau_1-2\ i+\tau_2-2} B_{\tau_2}(k|k) \\ &+ CQC^T\delta_{i_1 i_2} \end{aligned} \tag{5}$$

Filter update: $1 \le i \le I$

$$\begin{aligned} \widehat{\Psi}_i(k+1|k+1) &= \widehat{\Psi}_i(k+1|k) + \\ &\chi^{(k+1)}\Gamma^{(k+1)} \begin{bmatrix} \mathcal{Y}_m(k+1) - \Theta_m\widehat{\Psi}_m(k+1\,|\,k) \\ \vdots \\ \mathcal{Y}_p(k+1) - \Theta_p\widehat{\Psi}_p(k+1\,|\,k) \end{bmatrix} \end{aligned} \tag{6}$$

Filter covariance update: $1 \le i_1, i_2 \le I, I$

$$\begin{aligned} P_{i_1 i_2}(k+1\,|\,k+1) &= P_{i_1 i_2}(k+1\,|\,k) + \\ &\chi^{(k+1)}\,\Gamma^{(k+1)} \begin{bmatrix} \Theta_m\, P_{m i_2}(k+1\,|\,k) \\ \vdots \\ \Theta_p\, P_{p i_2}(k+1\,|\,k) \end{bmatrix} \end{aligned} \tag{7}$$

where $\chi^{(k+1)}$ is a $J \times NJ$ matrix and $\Gamma^{(k+1)}$ a matrix of order NJ, given by

$$\chi^{(k+1)} = \left[\; P_{i_1 m}\Theta_m^T \ldots \qquad P_{i_1 p}\Theta_p^T \; \right] \tag{8}$$

$$\Gamma^{(k+1)} = \begin{bmatrix} \Theta_m P_{mm}\Theta_m^T + R_{mm} & \ldots & \Theta_m P_{mp}\Theta_p^T + R_{mp} \\ \vdots & \ddots & \vdots \\ \Theta_p P_{pm}\Theta_m^T + R_{pm} & \ldots & \Theta_p P_{pp}\Theta_p^T + R_{pp} \end{bmatrix}^{-1} \tag{9}$$

with $\{P_{i_1 i_2}\}$'s on the right hand side measured at $(k+1\,|\,k)$. The bKBf reduces over KBf the computational requirements by a factor of I. To improve further, we approximate the error field, $\epsilon = \Psi - \widehat{\Psi}$, at each iteration as a Gauss Markov random field. We call the *approximate* implementation as the local bKBf (lbKBf). Using results in [4], the blocks $\{P_{i_1 i_2}\}$'s are related by

$$P_{i_1 i_2} = \begin{cases} \prod_{\tau=i_2-2}^{i_1} (P_{\tau+1\tau+1}^{-1} P_{\tau+1\tau})^T P_{i_2 i_2-1}^T & \text{for } 1 \le i_1 \le (i_2-2) \\ P_{i_1 i_1-1} \prod_{\tau=i_1-2}^{i_2} (P_{\tau+1\tau+1}^{-1} P_{\tau+1\tau}) & \text{for } (i_2+2) \le i_1 \le I. \end{cases}$$

In the lbKBf we update only $\{P_{ii}\}$'s and $\{P_{ii+1}\}$'s. Any other blocks $\{P_{i_1 i_2}\}$'s if required may then be obtained from blocks $\{P_{ii}\}$'s and $\{P_{ii+1}\}$'s using the above equation.

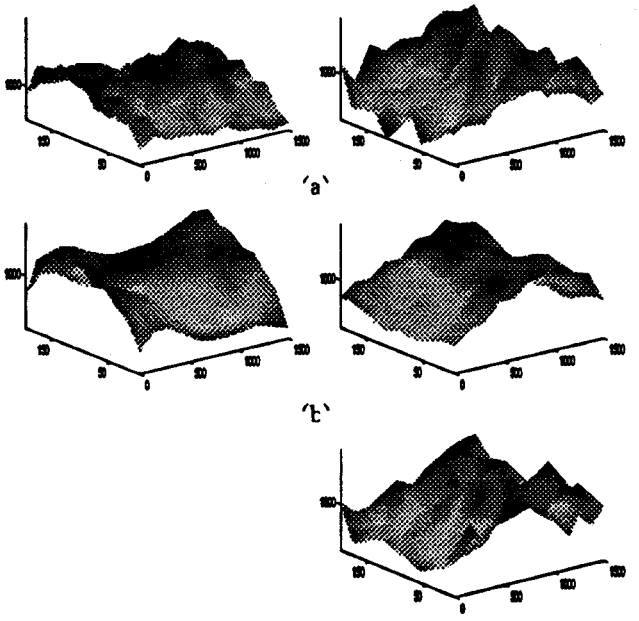

Fig. 1: Comparison between mesh plots of the sea surface height obtained from: (a) the nature run, (b) the deterministic run, and (c) the data assimilated run. The plots on the left hand side are at the start of data assimilation and those on the right, are 7 days and 9 hours later.

4.1. Computation Count

Although the appearance of bKBf is rather formidable, the bKBf is computationally much simpler than the standard implementation of KBf. The reason is because bKBf deals with $I \times I$ blocks rather than $I^2 \times I^2$ blocks. The two major matrix operations involved in the algorithm are multiplication and inversion, which for an $I \times I$ matrix are both of order $O(I^3)$. For a lattice of size $10^3 \times 10^3$, instead of the inversion and multiplication of a $10^6 \times 10^6$ error covariance matrix $\mathcal{P}$, we need to invert and multiply 10^6 matrices of $10^3 \times 10^3$, which represents a reduction of 10^3 in the operation count respectively. Further, reduction is provided in the lbKBf. In the lbKBf, instead of updating and storing the error covariance matrix $\mathcal{P}$ which for the lattice size considered earlier is an $I^2 \times I^2$ matrix, we update I blocks P_{ii} and $I-1$ blocks P_{ii+1} which are all $I \times I$ matrices. This reduces overall the computations by $O(I^2)$, i.e., a factor of 10^6 for the lattice size considered earlier.

4.2. Storage

Since in the lbKBf, instead of storing the error covariance matrix $\mathcal{P}$, we store I blocks P_{ii} and $I-1$ blocks P_{ii+1}, the storage requirements are also decreased from $O(I^4)$ to $O((2I-1)I^2)$, a reduction by a factor of I over the conventional KBf.

5. EXPERIMENT

A twin experiment for the equatorial channel version of the Pacific ocean modeled by the Beta-plane approximation, is performed. The fields that we reconstruct, are sea surface height (SSH) and the velocity components (u,v). The ocean basin is rectangular with dimensions 1500 km in the east-west direction and 200 km in the north-south direction. The resolution in the east-west direction, Δx, is 50km (1/2° longitude) and in the north-south direction, Δy, is 20 km (1/6°l atitude). The time step, Δt, is 120 s, which satisfies the stability bounds. The ocean depth, H, is 1 km. The specific volume of the sea water is 9.7582×10^{-4} $m^3 kg^{-1}$. The model is run by an east-west surface wind stress X and a north-south wind stress Y,

$$X = -X_m \sin\frac{\pi x}{a}\cos\frac{\pi y}{b}, \qquad Y = Y_m \cos\frac{\pi x}{a}\sin\frac{\pi y}{b} \tag{10}$$

where X_m and Y_m are $1 N/m^2$, a is 1500 km and b 200 km.

Before assimilating data, we ran two versions of the Beta-plane model over a period of 140 days (100,000 iterations). One version simulates the actual conditions and is called the "nature run". We add to the wind stresses (X, Y) white Gaussian noise corresponding to 12 dBs of signal to noise ratio (SNR). The other version, models deterministic

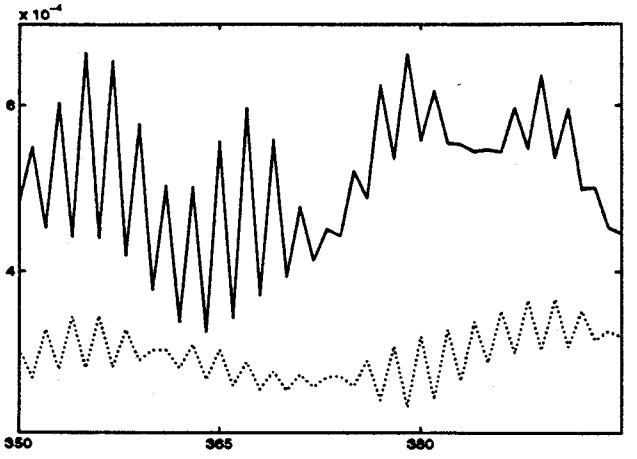

Fig. 2: Relative mean square error for sea surface height data assimilation as a function of model time in hours. Solid curves are with no data being assimilated. Dashed curves are with assimilation of sea surface height.

predictions and is called the "deterministic" setup. It propagates simply (X, Y). Fig. 1 shows the "nature" and the "deterministic" SSHs at the end of 140 days run.

We consider two satellites whose orbits are offset in space by half the width of the ocean domain in consideration. The satellites measure the SSH at spatially separated points along the rows of the ocean domain. Each satellite follows the scan pattern of Topex/Poseidon [5]. The first satellite begins its scans from row 1 and scans every fifth row, i.e., it measures SSH along rows 1, 6, 11, and so on till the entire ocean basin is covered. In the second orbit, rows 2, 7, 12, and the corresponding higher numbered rows are scanned. Similarly, for the next orbits. For a grid of dimensions $I \times J$, the second satellite starts from rows $J/2+1$ and follows the non-recurrent scan pattern explained above for the first satellite. The ocean domain that we consider has dimensions 31 x 11. Row 16 is, therefore, scanned first by the second satellite. We now assimilate the SSHs obtained from these satellites into the deterministic model explained in section 2.

The experiment presented above was performed to emphasize the usefulness of data assimilation and the improvement it offers over predictions regarding ocean circulation solely based on integrating the OGCM with assumed initial conditions, boundary conditions, and external forcings. We compare first the subjective quality achieved by data assimilation techniques and later comment on resulting mean square error (MSE).

5.1. Qualitative Comparison

The SSH assimilated with the satellite scanned data are better estimates of the actual state conditions, in our case the real world run, than the fields predicted by the OGCM referred to as the deterministic fields. A visual comparison of fig. 1 illustrates this point. The deterministic fields do not capture much of the finer details and of the high frequency features such as troughs and crests. In fact, the deterministic fields are a low pass version of the real world SSH. The SSH's produced by the assimilation algorithm improve on the frequency contents of the reconstructed fields. The assimilated fields are visually closer to the actual fields.

5.2. Quantitative Comparison

The MSE plots presented in fig. 2 illustrate that the data assimilation run (dashed curve) has smaller MSE than the deterministic run with no data assimilation. In fact, the errors are about half the errors that those in the "deterministic" run.

6. SUMMARY

The paper develops a computationally fast and a storage efficient implementation of the KBf for filtering and prediction of time and space dependent signals. The resulting algorithm, the lbKBf, simplifies the computations by $O(I^2)$ and reduces the storage by $O(I)$ for a grid of size $I \times I$. We apply the lbKBf to estimate ocean circulation fields specifically the sea surface height in Physical Oceanography. In our experiment, we assimilated data from two different satellites flying in the same orbit with no offset in time but separated in space by half the ocean domain. The satellites follow Topex/Poseidon scanning pattern. The fields reconstructed by assimilating satellite data via the lbKBf based algorithm provide better estimates of the actual conditions than the fields obtained from numerical simulation of the physical model. The latter is the same as the predictions provided by the OGCMs. A study based on the MSE illustrates quantitatively this gain.

REFERENCES

[1] M. Ghil and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," *Advances in Geophysics*, vol. 33, pp. 141–265, 1991.

[2] A. E. Gill, *Atmosphere-Ocean Dynamics*, Vol. 30 of *Int. Geophysics Series*. NY: Academic Press, 1982.

[3] A. Asif and J. Moura, "Assimilation of Satellite Data in Beta Plane Ocean Circulaton Models," in *Proceedings ICASSP 95*, Vol. V, (Detroit, MI), 1995, pp. 2789–92.

[4] A. Asif, *Fast Recursive Algorithms for Large Time Varying Multidimensional Fields*. PhD thesis, Carnegie Mellon Univ., Pittsburgh, PA, 1996.

[5] C. S. Ruf, S. J. Keihm, and M. A. Janssen, "Topex/poseidon microwave radiometer (tmr): I. instrument description and antenna temperature calibration," *IEEE Transactions on Geoscience and Remote sensing*, vol. 33(1), pp. 125–137, 1995.