New Fast Kalman filter method

Hojat Ghorbanidehno*, Hee Sun Lee*

1. Introduction

Data assimilation methods combine dynamical models of a system with typically noisy observations to obtain estimates of the state of the system with time. When these dynamical and observation models are linear, the Kalman filter (KF) algorithm gives the best estimate in a least square sense (Kalman, 1960). However, for the large number of unknowns of such large systems, KF methods are impractical, because updating the huge covariance matrix is computationally expensive. Low rank approximation methods have been devised to overcome this problem; such methods assume a low rank of the covariance matrix, which they approximate by smaller, full-rank matrices. The approximation error is small for smooth functions, but may be larger for more complex physical problems, leading to filter divergence and inaccurate state estimates.

We present a new algorithm that utilizes the exact covariance matrix avoiding the approximation error (we call it SpecKF for now), and does so in a computationally efficient way, even for systems with large numbers of unknowns. The computational speed-up of the SpecKF is achieved by updating small cross-covariance matrices instead of large covariance matrices. The benefit can be considerable, especially in large systems, because the computational complexity of SpecKF scales linearly with the number of unknown, as opposed to the KF that scales quadratically. We investigate the accuracy and performance of the SpecKF for a diffusion case with random perturbations. Our results show that the SpecKF provides high accuracy at a smaller computational cost than the Ensemble Kalman Filter (which is a very popular KF variant).

The Kalman filtering process has two steps: (i) the *prediction* of the state at time t_{k+1} using the forward model and the current estimate at time t_k, and (ii) the *update* of the predicted state at time t_{k+1} based on newly available measurements. Consider a state estimate vector $s_k \in R^m$ and a vector of noisy measurements $z_k \in R^n$ at time t_k. The Kalman Filtering process is given in the following recurrence (f: forecast, a: analysis):

Prediction/Forecast:

$$s^f_{k+1} = F_{k+1} s^a_k \tag{1a}$$

$$P^f_{k+1} = F_{k+1} P^a_k F^T_{k+1} + Q_{k+1} \tag{1b}$$

Update/Analysis:

$$K_{k+1} = P_{k+1} H^T_{k+1} \left(H_{k+1} P^f_{k+1} H^T_{k+1} + R_{k+1} \right)^{-1} \tag{1c}$$

$$s^a_{k+1} = s^f_{k+1} + K_{t+1} \left(y_{k+1} - H_{k+1} s^f_{k+1} \right) \tag{1d}$$

$$P^a_{k+1} = P^f_{k+1} - K_{k+1} H_{k+1} P^f_{k+1} \tag{1e}$$

The SpecKF updates cross-covariances instead of covariance matrices(P^f and P^a), and makes this possible for arbitrary forward models by means of a forward model approximation. This approximation brings about a small and controllable approximation error. The computational cost of estimation with our method is reduced from $O(nm^2)$ to $O(n^2 m)$ (n: number of measurements, m: number of unknown), which can be a dramatic decrease, especially for large systems where $n \ll m$. As a result, our method can handle a large number of unknowns in a computationally efficient way, both in terms of storage and computational time.

2. Derivation of method

The implementation of the KF algorithm for large scale systems becomes infeasible, because of the high storage cost and computational cost of the matrix operations outlined in Equations (1b) to (1e), involving several multiplications of $m \times m$ matrices when updating the covariance matrix P. In our method, we modifying the recurrence matrix operations such that only the $m \times n$ cross-covariance matrices (such as $PH^T = P^a H^T$

* *Email addresses:* hojjatgh@stanford.edu (Hojat Ghorbanidehno), hslee88@stanford.edu (Hee Sun Lee)

and $P^F = PF^T H^T$) are stored and updated at every time step. This modification takes advantage of the Kalman Filter recurrence to calculate products of big matrices like PH^T without explicitly having to construct and multiply the individual matrices. This is achieved by multiplying equations (1b) and (1e) by appropriate Jacobians to obtain:

$$P^f_{k+1} H^T = F P^a_k F^T H^T + Q H^T \tag{2}$$

$$P^a_{k+1} H^T = P^f_{k+1} H^T - K_{k+1} H P^f_{k+1} H^T \tag{3}$$

However, after the above modifications, equation (2) still involves the product $PF^T H^T$, which is computationally intractable for large m. Expanding this term and using equations (1b) and (1e), we obtain:

$$\begin{aligned} P^a_{k+1} F^T H^T &= (I - K_{k+1}) P^f_{k+1} F^T H^T \\ &= (I - K_{k+1} H)(F P^a_k F^T F^T H^T + Q F^T H^T) \end{aligned} \tag{4}$$

It becomes apparent that while equations (2) and (3) can be used to update the cross-covariances using the recurrence rather than direct matrix-vector multiplication, the term $FPF^T F^T H^T$ negates this benefit as it involves terms that are equally expensive to compute. To overcome this, we proceed by discretizing the forward model F in time with a Taylor series:

$$F - I = (F)_t dt + O(dt^2) \tag{5}$$

where $(F)_t$ is the Jacobian of F with respect to t. Using the above, we obtain the following:

$$PF^T F^T H^T = -P^a_k H^T + 2 P^a_k F^T_{k+1} H^T_{k+1} + P((F - I)^2)^T H^T \tag{6}$$

where $(F - I)^2 \sim O(((F_k)_t dt)^2)$. As a result, for a small dt between data assimilations in the Kalman Filter, the error introduced by neglecting the last term of equation (6) is expected to be small. However, since this approximation is used only once within the algorithm (i.e. the full forward operator F is used in all other calculations except in the expansion of equation (6)), depending on the non-linearity of the problem, the error introduced by this forward model approximation will likely not affect the results of the filter significantly.

Using the above and defining the new matrices $T_{k+1} = P^f_{k+1} H^T$, $P^H_{k+1} = P^a_{k+1} H^T$, and $P^F_{k+1} = P^a_{k+1} F^T H^T$, the SpecKF algorithm can be summarized by the following recurrence:

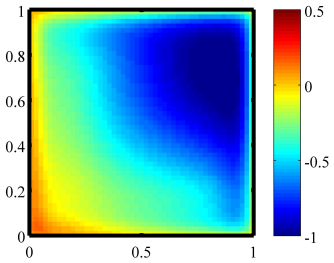


Figure 1: Purturbed solution after 1000 steps

$$s^f_{k+1} = F_{k+1} s^a_k \tag{7a}$$

$$T_{k+1} = F_{k+1} P^F_k + Q H^T \tag{7b}$$

$$K_{k+1} = T_{k+1} (H T_{k+1} + R)^{-1} \tag{7c}$$

$$s^a_{k+1} = s^f_{k+1} + K_{k+1} (z_{k+1} - H s^f_{k+1}) \tag{7d}$$

$$\begin{aligned} P^F_{k+1} &= (I - K_{k+1} H)(-F_{k+1} P^H_k \\ &\quad + 2 F_{k+1} P^F_k + Q F^T_{k+1} H^T_{k+1}) + O((\Delta_t)^2) \end{aligned} \tag{7e}$$

$$P^H_{k+1} = (I - K_{k+1} H) T_{k+1} \tag{7f}$$

3. Validation

In order to validate the SpecKF implementation, we compare it to the results obtained from the original KF algorithm; the difference between the two indicates how significant is the combined error of the forward model approximation and the low-rank approximation of Q, which are the two sources of error in the SpecKF algorithm. The benchmark problem we use is the linear diffusion equation in a two-dimensional unit domain, $[0, 1] \times [0, 1]$, where the state (pressure) is contaminated with random noise:

$$\frac{\partial \phi}{\partial t} = D \nabla^2 \phi + noise \tag{8}$$

In equation (8), D represents the diffusion coefficient which is assumed to be spatially and temporally constant. Discretizing equation (8) using explicit Euler in time and central difference in space, we obtain:

$$\begin{aligned} \phi^{(k+1)}_{i,j} &= \phi^{(k)}_{i,j} + \frac{dt}{(dx)^2} \left(\phi^{(k)}_{i+1,j} - 2\phi^{(k)}_{i,j} + \phi^{(k)}_{i-1,j} \right) + \\ &\quad \frac{dt}{(dy)^2} \left(\phi^{(k)}_{i,j+1} - 2\phi^{(k)}_{i,j} + \phi^{(k)}_{i,j-1} \right) + \sqrt{10 dt}\, w^{(k)}_{i,j} \end{aligned} \tag{9}$$

where w is noise that represents various small and erratic sources and sinks over the domain with zero mean and covariance matrix of $\sqrt{10dtQ}$. In this problem Q is selected from the Gaussian family of covariance functions where $Q_{i,j} = exp(-(\frac{x_i - x_j}{L})^2)$. Equation 9 is solved for a 51×51 grid resulting in $m = 2401$ (excluding the boundaries), with a discretization $dx = dy = 0.02$, with Dirichlet boundary conditions at all boundaries and initial $\phi = 0$ throughout the domain except at boundaries. For time discretization, a $dt = 10^{-5}$ is chosen, to satisfy the numerical stability requirement $dt \leq 4 \times 10^{-4}$. There are 20 point observations uniformly distributed in the unit domain, with observation error covariance $I_{20\times20}$. Figure 1 shows the distribution of ϕ after 1000 time steps, indicating the distortion of an otherwise smooth field due to the noise added.

In addition to the comparison to the KF results for validation of the SpecKF, the diffusion problem is used to conduct error analysis and evaluate the computational efficiency of the SpecKF compared to both the KF and the EnKF.

Figure 2 shows the estimates obtained with SpecKF at time step 1000. Comparison to Figure 1 indicates that the SpecKF results provide high accuracy.

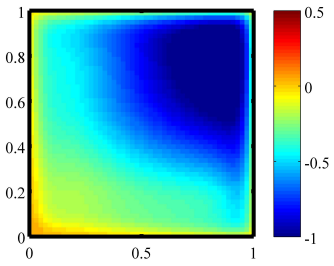
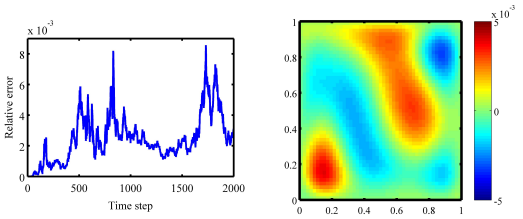


Figure 2: SpecKF estimation at time step 1000.

Comparison to the KF results (Figure 3) shows that the difference between the estimations obtained by SpecKF and KF is negligible, with the absolute difference $\frac{\|\phi_{SpecKF} - \phi_{KF}\|}{\|\phi_{KF}\|}$ in the order of 10^{-3}, where the true state varies from -1.5 to 0.5. This shows that SpecKF is consistent with KF and the error of the forward model approximation in this case is small. As noted previously, in this case the time step being used is rather small, representing an optimistic case for data assimilation. Nevertheless, the good agreement of SpecKF with KF under these conditions validates the SpecKF implementation. It should be noted that the KF estimate is the optimal estimate for a linear problem with given set of noisy measurements. As shown in Figure 3 the overall error added is maintained at low levels, since at all times



(a) (b)

Figure 3: a) Relative difference between SpecKF and KF estimates with time b) Absolute difference between SpecKF and KF estimations at time step 1000

the time step taken is small (for numerical stability purposes as outlined previously), and since the error of the forward model approximation is second order in time. The contribution of the low-rank approximation of the Q matrix is also included in this error. It should be mentioned that in this problem the SpecKF provides estimation with the same accuracy as the KF method, while the computational cost of specKF is much smaller than the computational cost of KF method, specially for the cases with large number of unknowns (Figure 4).

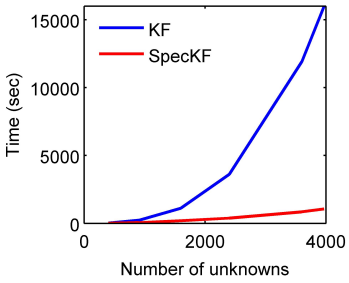


Figure 4: Comparison of time taken by full KF and SpecKF methods for different number of unknowns

Finally, we compare the results of the SpecKF for the diffusion problem with EnKF an alternative fast Kalman Filter algorithms in terms of accuracy of estimation. The results of each method are compared to the original KF algorithm. First, we compare the SpecKF results to those obtained by the Ensemble Kalman Filter, using the same 20 noisy measurements. Three different ensemble sizes were evaluated, all larger than the number of measurements, in order to allow the EnKF to converge. As shown in Figure 5, increasing the ensemble size reduced the relative error of the EnKF, however in all three cases the EnKF difference from the KF was more pronounced compared to the SpecKF. Since the ensemble sizes were greater than the number of mea-

surements, the computational cost of the EnKF in all three cases was also greater than that of the SpecKF.

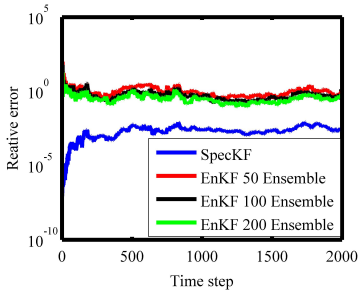


Figure 5: Comparison between relative error of SpecKF estimation and EnKF estimation(with 3 different ensemble sizes) respect to KF estimation

3.1. Approximation of covariance matrix

The SpecKF algorithm gives the posterior covariance only at measurement points. However, we may need to know the uncertainty of the estimates at locations that are not known a priori. We can obtain an estimate of the full matrix P by assuming that the posterior covariance has a low-rank representation:

$$P \approx USU^T$$

U is a known, preselected $m \times n$ matrix, i.e., we assume that P can be approximated by a low-rank matrix with rank equal to the number of measurements.

Then, the cross-covariance HPH^T, which is available in the SpecKF, can be written as:

$$\begin{aligned} HPH^T &\approx HUSU^TH^T = (HU)S(HU)^T \\ &= (HUS^{1/2})(HUS^{1/2})^T \end{aligned}$$

Then, we perform a Cholesky decomposition of the known cross-covariance $HPH^T = VV^T$. This gives:

$$V \approx HUS^{1/2}$$

Since V, H and U is known we can obtain an expression for $S^{1/2}$:

$$S^{1/2} = (HU)^{-1}V$$

This step assumes that HU is non-singular. H is a fat matrix with n rows, while U is a thin matrix with n columns. For HU to be non-singular, we need to assume that $\text{span}(U) \cap \text{null}(H) = \emptyset$. Geometrically, this means that none of the "important" components of P are in the null space of the observation matrix H, that is, if there are components of P (= span(U)) that are not observed, they cannot be reconstructed by this method. With this assumption, we can obtain S and get the desired approximation of P:

$$P \approx USU^T$$

4. Application: CO_2 monitoring

The second benchmark application evaluates the performance of the SpecKF for the non-linear problem of CO_2 transport in a two dimensional homogeneous domain. The accuracy of the SpecKF is compared with that of the EnKF. The forward model used to simulate the injection of CO_2 in the subsurface is TOUGH2, a multi-dimensionsional, multi-phase, multi-component non-isothermal numerical simulator for flow and transport in porous media (Pruess, 1991; Pruess et al., 1999). Initially the forward simulation is run for 225 days to provide the true data for pressure and CO_2 saturation, which are the two state variables we are interested in tracking with the Kalman Filter. From this data, 25 saturations and 9 pressures ($n = 34$) are collected every 15 days at locations indicated in Figure 6. These measurements are then contaminated with Gaussian error. Using these measurements, the Kalman Filter algorithms are used to estimate the pressure and saturation for the whole domain, resulting in $2 \times 2025 = 4050$ unknowns ($m = 4050$). The assimilation of data begins on day 30 of the true experiment. The state of the system at this time would be unknown in practice. For this reason, we initialize the Kalman Filter from a wrong initial condition corresponding to an approximate knowledge of the location of the CO_2 based on the location of the injection wells. The objective of the Kalman Filter is to predict the states at subsequent times, based on these erroneous initial conditions, and the noisy measurements collected. Both filters are initiated with the same erroneous initial conditions and the same set of noisy measurements is assimilated every 15 days. Results are compared to the true state at different times. We did not compare our results with KF, because implementation of KF method for this problem was very expensive.

Figure 7 compares the true state with the estimates obtained by the SpecKF and the EnKF with same computational cost. The SpecKF after the first data assimilation step does not give an accurate estimate. However, as more measurements are assimilated the estimation improves drastically such that the SpecKF estimate at 90 days greatly resembles the true state. The relative error of the SpecKF and EnKF is compared in Figure 9, which shows both methods approximate the true solution with small relative error. However, when the noise

in measurements is small compared to the noise in the forward model, the accuracy of the EnKF decreases as shown in Figure 8. In contrast, the accuracy of SpecKF increases, indicating that SpecKF would be more robust than EnKF for a wider range of measurement errors. While we recognize that the above performance comparison may be specific to the application at hand, non-linearities, parameterization of the filters etc, our results indicate that the SpecKF can be a more reliable and faster estimation method than the EnKF, especially for non-linear problems with large numbers of measurements.

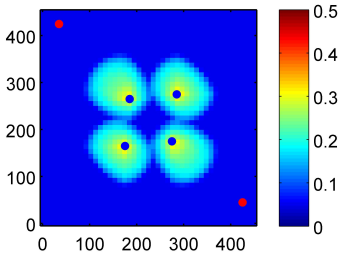


Figure 6: CO_2 saturation distribution after 15 days. Red dots: production wells and blue dots: injection wells.

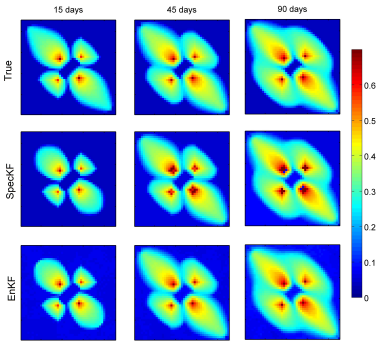


Figure 7: Comparison of true saturation with SpecKF and EnKF estimations. First row) True saturation, second row) SpecKF estimation and third row) EnKF estimation for saturation with same computational cost

5. Conclusion

This project presented the Spectral Kalman Filter, a new Kalman Filter variant that can be effectively used for data assimilation in non-linear dynamical systems with a large number of unknowns. The computational cost of the SpecKF is dramatically reduced compared to that of the original Extended Kalman Filter, as it scales

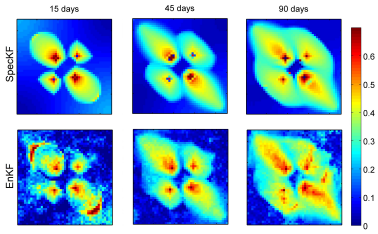


Figure 8: Comparison of true saturation with SpecKF and EnKF estimations for the case with small error. First row) SpecKF estimation and second row) EnKF estimation for saturation with same computational cost.

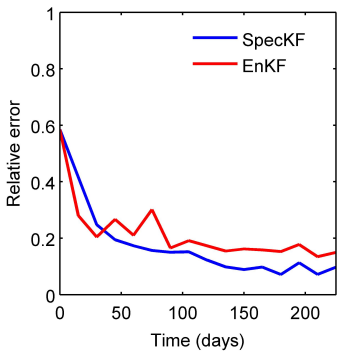


Figure 9: Relative error of the EnKF estimation and the SpecKF estimation compared to the true solution.

linearly (as opposed to quadratically) with the number of unknowns and makes the estimation of states in excess of thousands of unknowns feasible.

The SpecKF algorithm was validated for a linear diffusion problem, and it was shown that for a small time step between data assimilations the SpecKF gives results nearly as accurate as the full Kalman Filter. However, this method does not gives us error covariance matrix, but we showed that the full covariance matrix can be obtained by using a post processing step based on a low rank approximation of the covariance matrix. Since, this approximation is a post processing step, it does not affect the estimation with SpecKF.

References

Kalman, R., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82, 35–45.

Pruess, K., 1991. TOUGH2: A general-purpose numerical simulator for multiphase fluid and heat flow. NASA STI/Recon Technical Report N 92.

Pruess, K., Oldenburg, C., Moridis, G., 1999. Tough2 user's guide, version 2.0. Lawrence Berkeley National Laboratory, Berkeley .