Faculty of Applied Mathematics

University of Twente
University for Technical and Social Sciences

P.O. Box 217
7500 AE Enschede
The Netherlands
Phone +31-53-4893400
Fax +31-53-4892255
Email: c.h.g.hassing-assink@math.utwente.nl

Memorandum No. 1317

The detection of local changes in (semi-)causal stochastic fields

E.P. Hupkens and A. Bagchi

March 1996

ISSN 0169-2690

The Detection of Local Changes in (Semi)-Causal Stochastic Fields.

Erik P. Hupkens Arunabha Bagchi

March 19 1996

Abstract

Causal and semicausal stochastic fields may be rewritten in a state-space representation. Local changes in these fields may be regarded as time-varying changes in the parameters of this model. Using a Kalman filter bank, the generalized likelihood ratio may be computed if these changes are known to be in a certain set. If only some scaling parameters are unknown, exact expressions for the generalized likelihood ratio may be found.

1991 Mathematics Subject Classification: 60G35, 60G60, 93E10, 93E11, 94A13, 62F03

Keywords: change detection, stochastic fields, Kalman filter.

1 Introduction.

Despite, or may be because of, the enormous speed at which new technologies are developed and applied, faults may still occur in these systems. In fact, the effect of a fault seems to be greater if the technology used is more advanced. Therefore, the necessity of detecting faults is of considerable importance for several applications.

Moreover, if these systems operate in a faulty manner for some time, the damages that result may be disastrous for the system and its environment. This implies that these faults have to be detected as soon as possible after their occurrence.

The detection of changes in dynamic systems received an impulse in the 60's when the Kalman filter was successfully applied for this purpose. Newbold and Ho [1], Mehra and Peschon [2] and Willsky [3] developed the theory of fault detection using the Kalman filter.

Since then the use of Kalman filtering techniques has been applied at a larger scale in several applications. For example, in the introduction of [4], Clark, Frank and Patton mention an impressive number of detection problems that have been dealt with using the Kalman filter.

Classical detection theory focuses on one-dimensional processes, i.e. processes that evolve in time. However, several processes do not have such a natural one-dimensional evolution. The direction of evolution for two-dimensional processes, also known as random fields or stochastic fields, is in general not uniquely determined. A reasonably large class of stochastic fields may be described by one-dimensional processes. We only consider fields that may be described in the form of a stochastic system. For example, the field may be scanned in a column-by-column fashion, where every column may be written as a function of the previous

columns. Although not all fields may be described in such a form, it still covers a large class of stochastic fields [5].

In this paper we use Kalman filtering techniques for the detection of local changes in stochastic fields. Applications may be found, for example, in the detection of flaws on surfaces or object recognition. In an earlier work the detection of global changes in stochastic fields has been analysed [6]. There one single test is performed on the entire field to determine whether or not the structure of the field has changed. The reason for using a local approach here is twofold. Firstly, the changes are local, that is, the part of the field that is affected by the change is relatively small. Therefore, a global statistic may not be affected enough to cause a detection of the change. Secondly, we desire to use as little data as needed to detect a change. If the amount of data on the field is very large, the computational cost for processing all data may be enormous. Using a local approach, these problems may be reduced.

The main difference between the problem presented here and the classical detection problems, consists of the form of the changes. Where changes in standard one-dimensional processes generally appear at a certain time, and remain present for some period (which often lasts forever), in stochastic fields the changes may take rather unstructured forms. Depending on the method of processing the field, the changes may appear at seemingly chaotic instants. Therefore, the parameterization of the changes is a rather important aspect of the detection problem.

The general model for which the algorithms are developed is given in section 2. The actual algorithms and their statistical properties are included in the appendix. In section 3 the structure of the field and the form the changes may take are examined. Finally, section 4 contains some simulations to illustrate the theory.

2 The general model.

Several models have been developed for the description of discrete stochastic fields, mainly for the use in image processing. An overview is given by Jain [7] and Dubes and Jain [8]. One particular model is the one-dimensional state space representation. In this model, the field is divided in equally sized sets, and the sites in each set are gathered in a vector X. The sets are ordered in such a way that the sites in set $k+1$ only depend on the sites in set k. Obviously, this limits our choice of the sets considerably.

Under normal circumstances, the field may be described by some known state space model. Whenever a change in the field occurs, the model is no longer accurate. Two possible strategies exist to detect the occurrence of a change. The first one consists of a goodness-of-fit test that simply tries to detect any deviation from the nominal model (Mehra and Peschon [2]). However, since we assume the changes to be local, the deviations from the nominal model are expected to be rather small. Therefore, this approach may not be very effective in this case.

The second approach consists of parameterizing the changes. In case the possible changes are known, it may be possible to find a model for the changed field. Although the assumption of completely known changes is not very realistic, this approach may be more useful than the goodness-of-fit approach.

Suppose that the field may be described by

$$\begin{aligned} X_{k+1} &= A_k(\theta)X_k + F_k(\theta)W_k + E_k^X(\theta) \\ Y_k &= C_k(\theta)X_k + V_k + E_k^Y(\theta) \end{aligned}$$

The disturbances W_k and V_k are assumed to be zero mean i.i.d. Gaussian variables with covariances $R_W(k)$ and $R_V(k)$, respectively. The vector Y_k denotes the noisy measurement of the field on the kth set. The vectors E_k^X and E_k^Y are deterministic. For notational simplicity, we may sometimes write A_k when we mean $A_k(0)$, and the same holds for other system matrices. Also, the additive terms E_k^X and E_k^Y are assumed to vanish for $\theta = 0$.

The parameter θ represents the change in the field. It may take values in a certain parameter set Θ. The precise structure of this set depends on the changes that may appear. Under the null hypothesis, no change is present in the field. The nominal model of the field coincides on the parameter set with zero.

We now want to test whether a change is present or not. This implies that we test if the null hypothesis

$$H_0 : \theta = 0$$

holds.

Depending on the parameter set Θ we may choose a strategy with the purpose of detecting all deviations from the null hypothesis. In case Θ consists of a finite number of parameters, we may simply calculate the likelihood ratio for each of these parameters, and reject the null hypothesis if the largest of these ratios exceeds a certain threshold (Willsky [3]). Otherwise, a possible solution is to discretize the parameter space and apply the same procedure for the resulting problem. Obviously, this approach introduces some new robustness problems. Kumamaru, Sagara and Söderström [9] used a similar approach. They fixed the parameter such that only one of its components, the so-called monitoring parameter, had to be estimated. For this one-dimensional parameter a discretization of the parameter space may be quite accurate for a relatively low number of parameters. The drawback of this approach is that any change should affect the monitoring parameter, which may not always be the case.

For now, we assume that the parameter set contains a finite number of parameters. For each $\theta \in \Theta$ we may use a Kalman filter to find the log-likelihood ratio (Willsky [10])

$$\begin{aligned}
\ell_n(Y,\theta) &= \log \frac{f_\theta(Y)}{f_0(Y)} \\
&= \log \prod_{k=1}^{n} \frac{f_\theta(Y_k|Y_{k-1},\ldots,Y_1)}{f_0(Y_k|Y_{k-1},\ldots,Y_1)} \\
&= \sum_{k=1}^{n} \log \frac{f_\theta(\gamma_k(\theta))}{f_0(\gamma_k(0))}
\end{aligned}$$

Here $\gamma_k(\theta)$ is the innovation of the process. The algorithms for the computation of all variables and their expectations are given in the appendix. In this case, we may evaluate the LR as

$$\ell_n(Y,\theta) = \frac{1}{2}\sum_{k=1}^{n} \log \frac{|R_k(0)|}{|R_k(\theta)|} + \gamma_k(0)' R_k(0)^{-1}\gamma_k(0) - \gamma_k(\theta)' R_k(\theta)^{-1}\gamma_k(\theta)$$

where R_k denotes the covariance matrix of the innovation. The statistic on which we base our decision is the maximum value of these likelihood ratios, the generalized likelihood ratio (GLR), defined by

$$T_n(Y) = 2 \sup_{\theta\in\Theta} \ell_n(Y,\theta)$$

If $T_n(Y)$ is larger than a certain threshold λ, then the null hypothesis is rejected. Because of the complex nature of the statistic, the choice of the threshold is not that straightforward.

In general, a large number of samples without a change should be processed to determine a threshold value that gives a satisfactory probability of false alarm.

If the number of parameters in Θ is not finite, this approach obviously fails. If all variables are differentiable on Θ, we may approximate the GLR for small values of θ by the score test statistic (Rao [11])

$$S_n(Y) = \mathcal{Z}_n(Y,0)'\Gamma_n(0)^{-1}\mathcal{Z}_n(Y,0)$$

where

$$\mathcal{Z}_n(Y,\theta) = \frac{\partial}{\partial\theta}\ell_n(Y,\theta)$$

is the efficient score, and

$$\begin{aligned}\Gamma_n(\theta) &= \mathbf{E}_\theta \mathcal{Z}_n(Y,\theta)\mathcal{Z}_n(Y,\theta)' \\ &= -\mathbf{E}_\theta \frac{\partial^2}{\partial\theta^2}\ell_n(Y,\theta)\end{aligned}$$

is the Fisher information matrix.

For the computation of this new statistic we do not need to know the parameter θ; it may completely be calculated under the null hypothesis. If the true parameter θ is not close to zero, the score test is no longer a sufficient approximation of the GLR test. To find the value of the GLR test in this case, we may use something like a gradient search algorithm. However, this has the disadvantage that the Kalman filter equations have to be calculated for every iteration in this algorithm. Moreover, the convergence of these algorithms may be very slow so that a large number of iterations is needed to find the MLE.

Other algorithms for the computation of the efficient score are given by Wilson and Kumar [12], Segal and Weinstein [13] and Leland [14].

2.1 A special case.

¿From the algorithms used in the previous section, it may be clear that the changes in the system matrices A, F and C are the main reason for the complexity of the final test statistics. Therefore, we now investigate the simpler problem where these changes are supposed to be known, and the only unknown is a scaling parameter in the additional term. Basically, this implies that we are testing for the presence of a given change with an unknown mean.

Suppose that the field may be described by

$$\begin{aligned}X_{k+1} &= A_k(\theta_0)X_k + F_k(\theta_0)W_{k+1} + c_X E_k^X(\theta_0) \\ Y_k &= C_k(\theta_0)X_k + V_k + c_Y E_k^Y(\theta_0)\end{aligned}$$

where θ_0 is known. The only unknowns therefore are the real numbers c_X and c_Y. Since only the additive terms depend on these unknowns, the covariance $R_k(\theta_0)$ is independent of these constants. Only the mean of the process is affected by these parameters. In fact, we obtain a similar situation as the one dealt with by Willsky [10]. The optimal value of the vector $c = (c_X, c_Y)'$ may be calculated exactly, so that the generalized likelihood ratio also may be found without using any maximization. The GLR is given by

$$T_n(Y) = \sum_{k=1}^{n}(\log\frac{|R_k(0)|}{|R_k(\theta_0)|} + \gamma_k(0)'R_k(0)^{-1}\gamma_k(0) - G_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0)) + Q$$

where the quadratic form Q and the vector G_k are specified in the appendix, together with the algorithms leading to this expression.

Similarly, we may also apply the same procedure in case the parameter θ_0 is unknown, but contained in a finite parameter set Θ. In this situation, the maximum likelihood estimate of c has to be calculated at each time for all possible changes $\theta \in \Theta$. This has the disadvantage that for each parameter a relatively large number of variables has to be stored in memory. In fact, from the algorithms in the appendix it follows that we need approximately three times as much memory space for this problem compared to the problem with known intensity. However, if we use a gradient search algorithm to find the optimal value of the vector c, the computational load is bound to be considerably higher.

3 Detecting local changes in stochastic fields.

¿From now on we use a simplified model for the field. The matrices in the defining equations are assumed to be constant over the field. This implies that, under the null hypothesis, the field may be written as

$$\begin{aligned} X_{k+1} &= AX_k + FW_{k+1} \\ Y_k &= CX_k + V_k \end{aligned}$$

where $X_k \in \mathbb{R}^m$. This state vector may represent a column of the field, part of a column, multiple columns, or just some sites from the field. At the moment we do not specify the exact form of the state. The vectors W_k, V_k and Y_k are denoted as the field noise, measurement noise, and measurements of the field.

Several types of changes may occur in a field.

- A change in the matrix A may be the result of a change in the structure of the field; the relation between the field values at different sites has changed.
- A so-called drift or trend in the field may be modelled by an additive change in the state equation.
- A change in the matrix C may be the result of several events. In most cases, such a change will be accompanied by an extra additive change in the measurement equation.

  Firstly, the measurement device may be defective, thus producing incorrect measurements.

  Secondly, some object may be blocking the view of the measurement device. These objects may vary from clouds when taking satellite pictures, to pieces of dirt when scanning a metal surface with a sonar device.

  Finally, an abrupt change in the field may in some situations be described more efficiently by a change in the C-matrix, possibly together with an additive change in the measurement equation, than by a change in the state equation. This may be explained by the lack of memory in the measurement equation, opposed to the Markovian nature of the state equation. For a fault that is described by a change in the A-matrix, the model is bound to give changed values at the sites in the neighbourhood of the fault. However, if the change is abrupt, the actual field has not changed at these sites, so that the model becomes less accurate.

In the sequel we only consider changes of the third type; a change in the matrix C, accompanied with an additive change in the measurement equation.

There are two possibilities for dealing with local changes. The first one uses a parameterization of the location of the change. This implies that the position of the changed parameters in the system matrices also have to be regarded as (discrete) parameters.

The second method uses local models of the field. It is based on the fact that locally the field may be described by a different model in case a change has occurred.

To compare these methods, we give an example.

Example 1 *Assume we have a model describing a square field (N by N) under normal circumstances. The state X_k is the kth column of the field. Suppose we want to detect an object of known size and form, say a square of size M by M, where M is smaller than N. Using the first method, we would have to find a parametrization of such a change. The matrix C changes as soon as X_k gets covered by the object. The rows of C corresponding to the elements of X_k that are covered now change to zero. Additionally, an extra vector, say E_k, appears in the measurement equation. It is zero everywhere, apart from the elements that correspond to the position of the object. Then we may describe the new measurement as*

$$Y_k = C_k X_k + E_k + V_k$$

The vectors E_k together form a matrix E that is exactly equal to the zero field plus the object. For this new model we may now use the detection method as described in the previous sections. Since we do not know the position of the object, we have to repeat this procedure for all possible positions.

Alternatively, using the second method, we change the way of processing the data. Instead of using the columns of the field as the state, we now use parts of the columns as the state. In this case, we have to process all possible M by M fields that are contained in the original N by N field. For these M by M fields we have the model under the null hypothesis

$$\begin{aligned} X_{j+k+1,i} &= A_i X_{j+k,i} + F_i W_{j+k+1,i} + b_{j+k,i} \\ Y_{j+k,i} &= C_i X_{j+k,i} + V_{j+k,i} + r_{j+k,i} \end{aligned}$$

Here j and i denote the indices of the first column and the first row of the M by M field. The extra terms b and r denote boundary values that appear due to the cut-off of the vectors X_k. The changed measurement equation under the alternative hypothesis is given by

$$Y_{j+k,i} = E_k + V_{j+k,i}$$

The matrix E, consisting of vectors E_k, is nothing else than.the object itself. This implies that the Kalman filter is not needed under the alternative hypothesis; the field may simply be described by the classical signal plus noise model.

Both methods require the same amount of model evaluations; this equals the number of possible locations of the object. For the first method, the model under the null hypothesis requires the evaluation of one Kalman filter. For each possible location, another evaluation of the Kalman filter is needed to obtain the innovation under the alternative hypothesis.

The second method requires an evaluation of the Kalman filter under the null hypothesis for each possible position. The alternative hypothesis however does not need any Kalman filter at all.

Moreover, the first method processes an N by N field for every location, whereas the second method only processes an M by M field ($M < N$). Additionally, if more than one object may be present in the field, the second method is inclined to outperform the first method.

Although the second approach from the previous example appears to outperform the first approach, its implementation may lead to some problems if the changes are allowed to take more complex forms. The boundary values that are required for the evaluation of the model under the null hypothesis may not be independent from the other sites. For example, consider a change that affects two columns of the field, say the $(k-1)$th one and the $(k+1)$th one. Then, to evaluate the model under the null hypothesis, we need the values at the kth column as the initial values for the $(k+1)$th column. However, these values depend on the $(k-1)$th column, so that the statistical behaviour under the null hypothesis is affected.

4 Simulations.

Here we demonstrate the theoretical results we obtained in the previous sections. To illustrate the quality of the tests we plot the probability of detection as a function of the probability of false alarm. These plots are also known as the *receiver operating characteristics* (ROC) [15].

The changes we consider are all of the type as described in the example of the previous section. This implies that some of the rows of the matrix C change to zero, while simultanuously an additive term in the measurement equation appears. If we place all vectors E_k^Y on their positions in the field, we obtain the scratch or damage that we are looking for. The intensity of the scratch is defined as the (assumed constant) value of the elements in E_k^Y.

We use 100 samples of the field without change to obtain the thresholds corresponding to a certain probability of false alarm. An additional 100 samples with change are used to obtain the probabilities of detection. A change is said to be detected if the GLR ($T_k(Y)$) crosses the threshold while the change is ϵ-detectable, that is, the expectation of the log-likelihood ratio has to be larger than ϵ. This extra condition is made to make the statistic more robust. For example, we do not always desire to reject the null hypothesis if only one pixel has a value that deviates from its nominal value. We choose $\epsilon = 0.1$, which implies that almost all changes are detectable.

The original model for the (causal) field is given by

$$\begin{aligned} X_{k,l} &= 0.2X_{k-1,l-1} + 0.4X_{k-1,l} + 0.3X_{k-1,l+1} + W_{k,l} \\ Y_{k,l} &= X_{k,l} + V_{k,l} \end{aligned}$$

where $W_{k,l}$ and $V_{k,l}$ are independent Gaussian noise processes with zero mean and a variance equal to 1. We assume the boundary values to be zero. The corresponding system matrices are given by

$$A = \begin{pmatrix} 0.4 & 0.3 & 0 & \cdots & 0 \\ 0.2 & 0.4 & 0.3 & \ddots & \vdots \\ 0 & 0.2 & 0.4 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0.3 \\ 0 & \cdots & 0 & 0.2 & 0.4 \end{pmatrix}$$

and $F = C = I_{10}$. Note that this field is actually causal in the k direction, so that the state space description becomes very natural. In general, causal fields may be more complex, as was

shown by Jain [16]. However, all causal and semi-causal fields with zero-boundary conditions may be described in a state-space form.

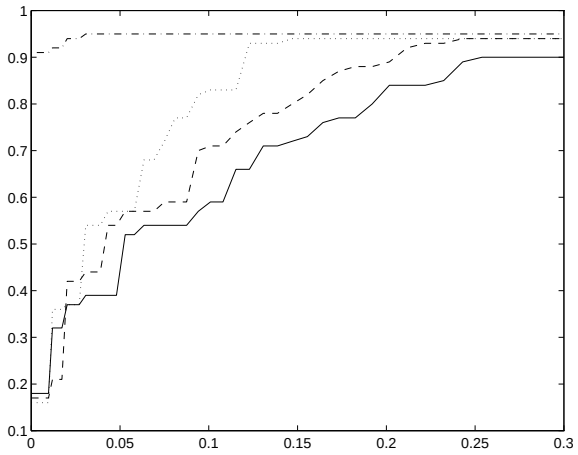

Figure 1: ROC. Parallel scratches. Intensities: 0 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

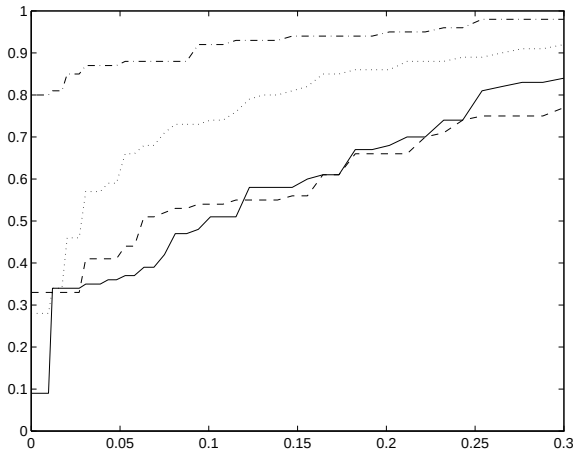

Figure 2: ROC. Orthogonal scratches. Intensities: 0 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

4.1 Known change.

First we assume the change to be known. The parameter space Θ contains one element only,

$$\Theta = \{\theta\}$$

The alternative hypothesis now is simple, so that the LR test does not need a maximization. In this case Wald's inequality for the probability of false alarm also holds;

$$\alpha = \Pr_0[(\exists n)(T_n(Y, \theta) > \lambda)] \leq e^{-\lambda/2}$$

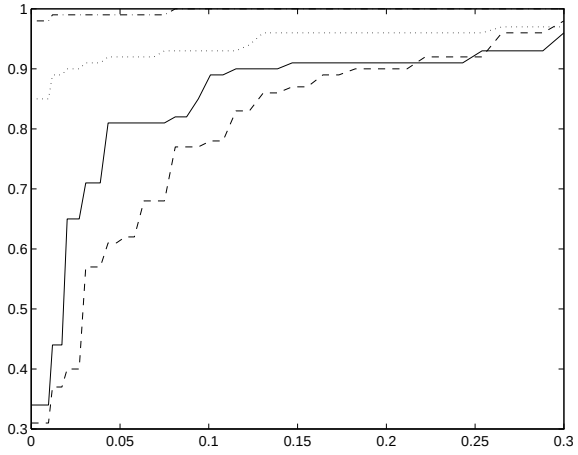

Figure 3: ROC. General damages. Intensities: 0.01 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

However, since this inequality is rather rough, the thresholds that may be obtained from this inequality are too large to be of practical use. Therefore, this inequality is only of interest if we may not use the experimental way of finding thresholds.

We consider three types of changes;

- **Parallel scratches.** In case we know that the only possible change is a certain vertical scratch on the field, the model for this change becomes quite simple. In fact,

  $$E_k^Y = cE\delta_{k-k_c}$$

  where c is the known intensity of the scratch, and E is a known vector consisting of zeroes and ones only. The known number k_c denotes the column where the scratch appears.

  Four different intensities are used, varying from zero to 2. The power of the test is plotted against the size of the test in figure 1.

- **Orthogonal scratches.** In case the only possible change is a horizontal scratch, the model remains quite simple. Denoting k_0 and k_1 as the number of the first and last column in which the scratch appears, we find

  $$E_k^Y = \sum_{t=k_0}^{k_1} cE\delta_{k-t}$$

  where E now contains one 1 only, at a fixed position.

  The same four intensities were used as for the parallel scratches. The ROC is given in figure 2.

- **General damages.** Finally, a combination of the previous changes gives us more general changes. For simplicity, we only consider changes that are constructed of parallel scratches that are connected with each other.

  The corresponding ROC is given in figure 3.

¿From the figures we may observe that a higher value of the intensity improves the detection quality of the test. If the intensity is smaller than 1, this difference is not that clear anymore. Furthermore, the general damages are detected most accurately, which may be explained by the fact that the general damages contain on average more sites than the other scratches. Finally, the parallel scratches are more easily detected than the orthogonal scratches. All sites in a parallel scratch are contained in one column, whereas the sites in the orthogonal scratches are spreaded over a large number of columns. This implies that the relative number of affected sites is larger for parallel scratches, which may explain the higher quality of detection.

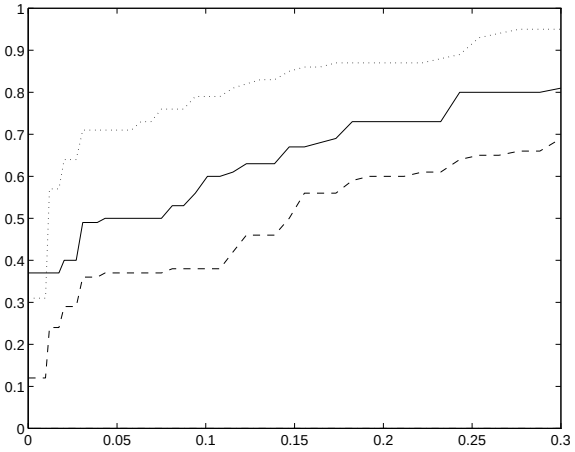

Figure 4: ROC. Unknown intensity, known position and form. Parallel scratches (solid), orthogonal scratches (dashed) and general damages (dotted).

4.2 Unknown intensity.

In case the position and the form of the scratch are known, but the intensity is unknown, we may find some direct results. If we write θ_0 as the change with intensity 1, then the system under the alternative hypothesis may be written as

$$\begin{aligned} X_{k+1} &= AX_k + FW_{k+1} \\ Y_k &= C_k(\theta_0)X_k + V_k + cE_k^Y(\theta_0) \end{aligned}$$

As we may see, this system is of the form as treated in section 2.1. The only difference is that c now is an element of $\mathbb{R}$ in stead of $\mathbb{R}^2$, so that the resulting expectation of the test statistic changes to

$$\mathbf{E}_c T_n(Y) = 1 + \sum_{k=1}^{n} \log \frac{|R_k(0)|}{|R_k(\theta_0)|} - N + \operatorname{tr} R_k(0;\theta_0)R_k(0)^{-1} + c^2 \rho_k(0;\theta_0)' R_k(0)^{-1} \rho_k(0;\theta_0)$$

Applying the theory on a set of samples with intensities varying between -2 and 2, we obtain a ROC as given in figure 4. As before, this figure shows that the best results are obtained for general damages, followed by parallel scratches and the orthogonal scratches again come in last.

As we have seen before, the detection quality improves if the true value of the parameter c is large. Therefore, for a consistent evaluation of these results we have to interpret them

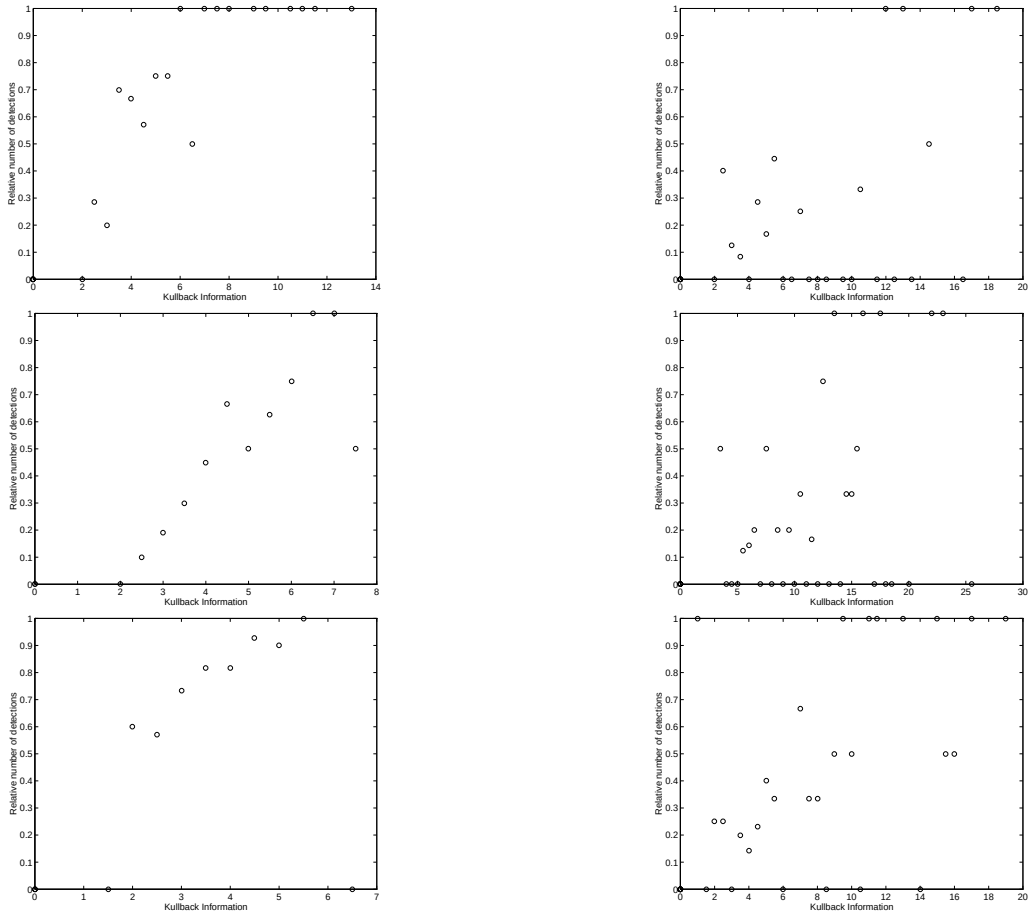

Figure 5: Detection rates as function of Kullback information. Upper left: Parallel scratches of known form. Upper right: Parallel scratches of unknown form. Middle left: Orthogonal scratches of known form. Middle right: Orthogonal scratches of unknown form. Lower left: General scratches of known form, Lower right: General scratches of unknown form.

in an appropriate way. A first attempt may be to consider the detection rate as a function of the intensity. However, since the number of sites in a scratch also varies, this may not be sufficient. Therefore, we use the Kullback information, which is defined as the expectation of the log-likelihood ratio, to indicate the detectability of a scratch. Figure 5 shows the relative number of detections as a function of the Kullback information. The points in the graph indicate the relative number of detections for changes with a Kullback information in an interval of length a half. The values of the Kullback information that are used correspond to $E_c T_N(Y)$, and may be calculated from the algorithms given in the appendix.

As expected, the figure clearly illustrates a correlation between the Kullback information in a change and the quality of detection.

4.3 Unknown change of given intensity.

If the position and the size of the change are unknown, we may use a generalized likelihood ratio test, for which the ROCs are shown in figures 6, 7, and 8. The intensity of the scratch is supposed to be known.

To save computation time we consider only a limited number of parameters for the general scratch case. For the first column of the field we evaluate all possible (parallel) scratches, and store only the 100 changes with the largest likelihood ratios. At every following column, we consider only the 100 previous changes, all possible extensions of these changes, and all possible new (parallel) scratches. Again the 100 most likely changes are stored.

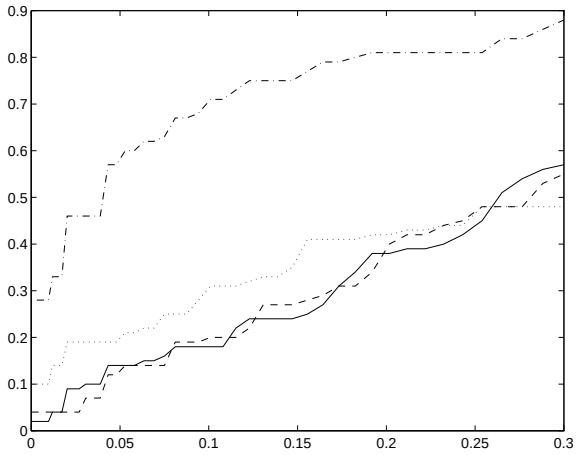

Figure 6: ROC. Unknown parallel scratches. Intensities: 0.01 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

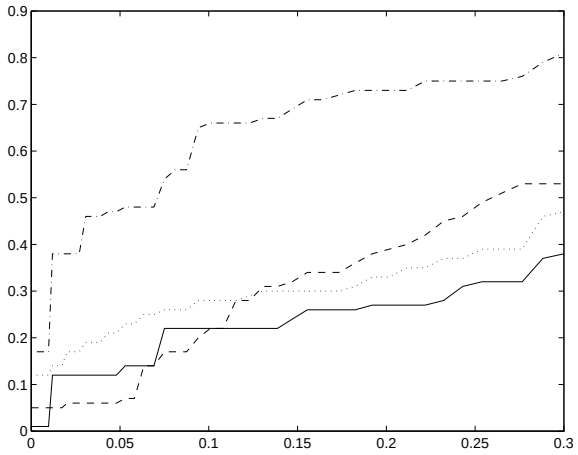

Figure 7: ROC. Unknown orthogonal scratches. Intensities: 0.01 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

¿From the figures we may see that the detection quality for general damages is reasonable, whereas the detection quality for both parallel and orthogonal scratches may only be considered reasonable for an intensity of 2. Note that the difference between parallel and orthogonal scratches is smaller than in the previous situations.

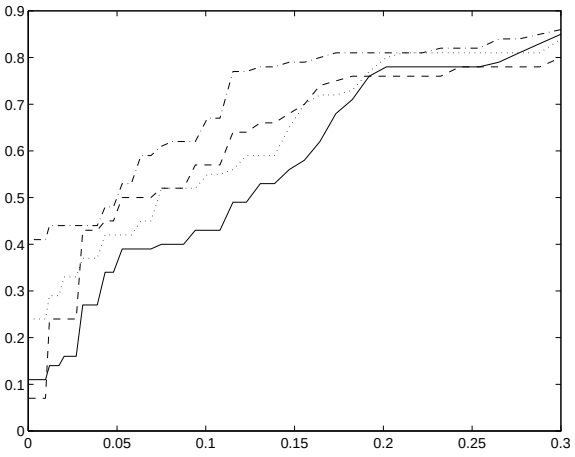

Figure 8: ROC. Unknown general damages. Intensities: 0.01 (solid), 0.5 (dashed), 1 (dotted), 2 (dash-dotted).

4.4 Unknown change of unknown intensity.

Finally, if apart from the position and the size of the scratch, the intensity is also unknown, we may use a combination of the previous methods to detect these changes. For all possible changes, after each iteration, we calculate the most likely value of the intensity of the scratch and its corresponding likelihood ratio. The maximum value of these ratios is compared with a threshold to decide whether or not a change is present.

As before, the value of the threshold is determined empirically. One hundred samples of the field without and one hundred samples of the field with a change are used to obtain a ROC. The resulting plot for these experiments is shown in figure 9. We may see that the detection quality has dropped considerably compared with the previous results. The detection quality for general damages is still superior to the other two classes of scratches. The difference between the orthogonal and the parallel scratches has vanished completely.

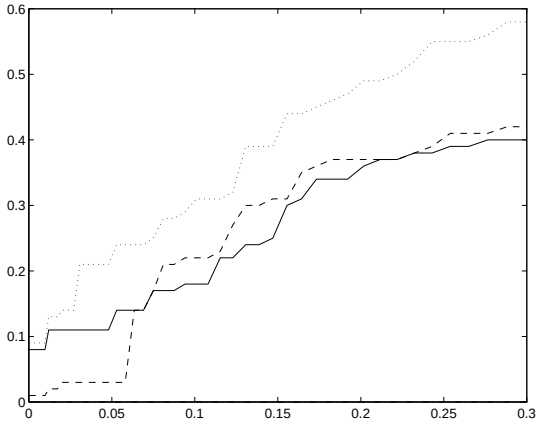

Figure 9: ROC. Unknown intensity, unknown position and form. Parallel scratches (solid), orthogonal scratches (dashed) and general damages (dotted).

As for the scratches of known form, these results are also presented using the Kullback information. Figure 5 illustrates the detection rates as a function of the Kullback information. As might be expected, a correlation between the Kullback information and the detection rate still exists. However, if we compare the figures on the right with the figures on the left, we may see that the correlation is not that clear anymore.

As a final illustration, one particular field has been highlighted. In figure 10 the measured field with a change is shown. The change has an intensity of 1.64 and is positioned as

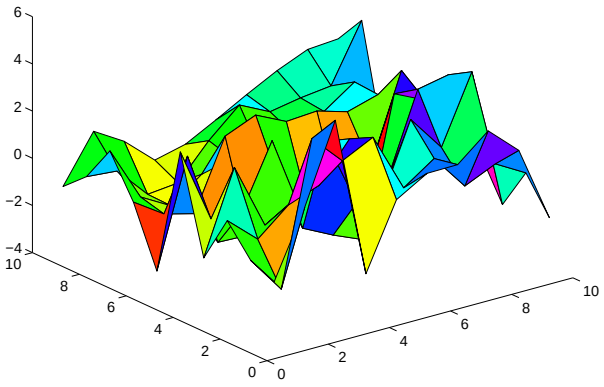

Figure 10: Measurement of a field with a change.

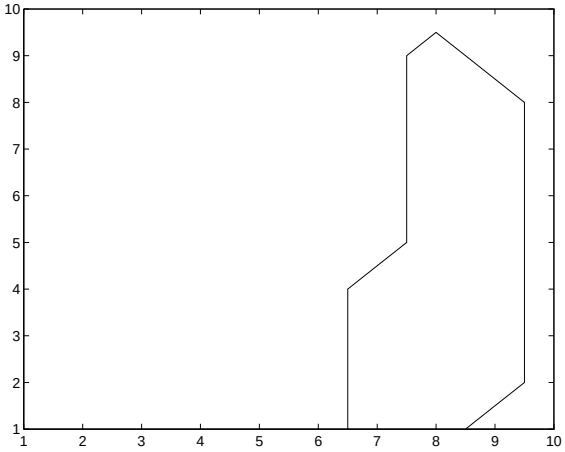

Figure 11: Position of the change.

illustrated in figure 11. The generalized likelihood ratio for this field is plotted in figure 12, together with its expected values, calculated under the assumption that the estimated changes are correct. ¿From the 100 samples without a change we may find that to obtain a false alarm

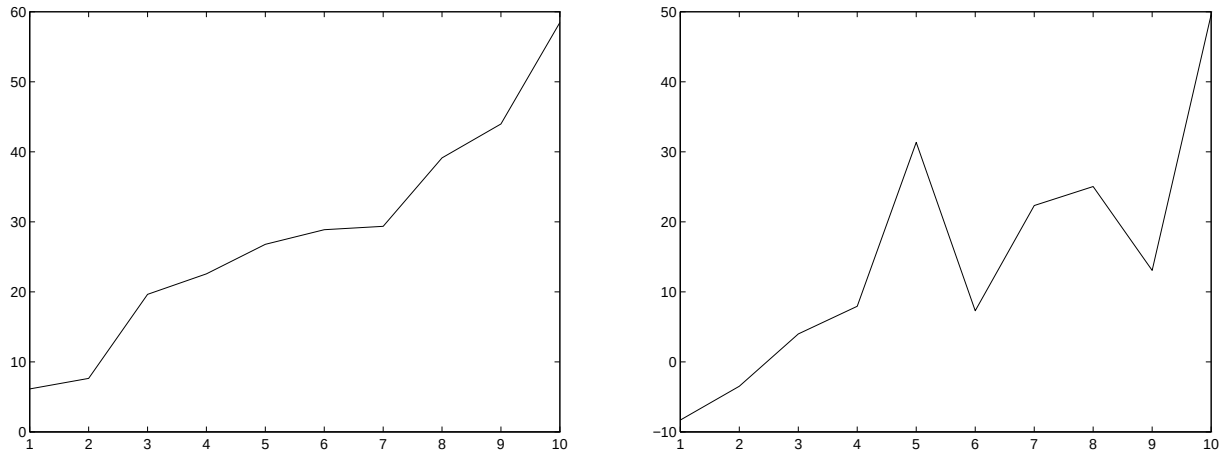

Figure 12: Generalized likelihood ratio (left) and Kullback information (right).

rate of 0.1 the threshold should be 44. To obtain a false alarm rate of 0.3, the threshold may be chosen as 38. ¿From the figure we may see that this second threshold is reached after the 9th column, and the first threshold is reached after the 10th column. However, if we examine the estimates at the 9th and the 10th column we find that only the latter coincides with a

correct detection. In figure 13 the positions of these estimated changes are shown. Their intensities are -1.45 (9th) and 1.61 (10th).

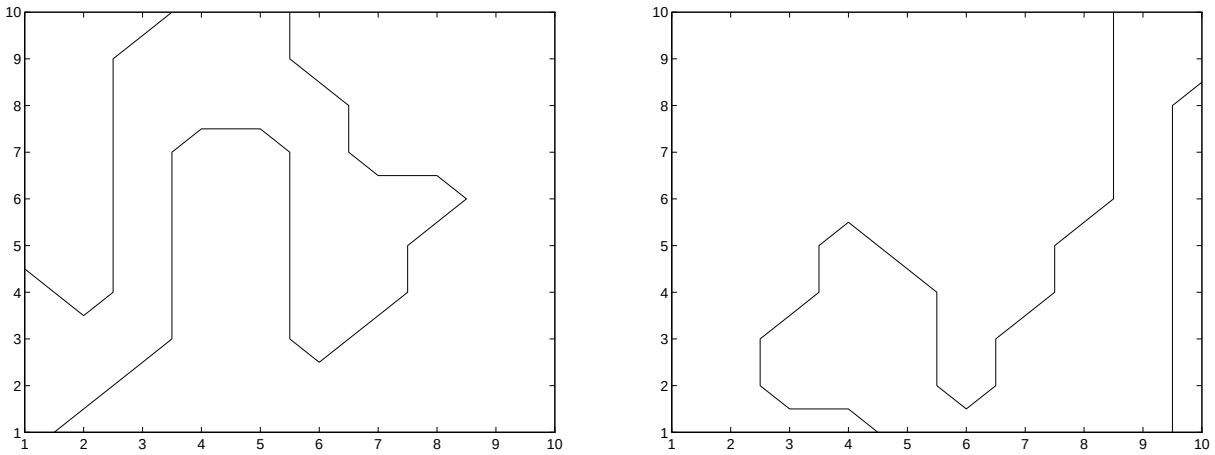

Figure 13: Position of estimated changes after 9th column (left) and 10th column (right).

5 Conclusions.

The problem of detecting local changes in stochastic fields may in some cases be rewritten as a parametric change detection problem in a stochastic system, given in a state space representation. In case the resulting parameter set contains a finite number of parameters, we may use a bank of Kalman filters to generate a test statistic based on the generalized likelihood ratio. A satisfactory value of the threshold used in this test may be obtained in an experimental way. In general, the model is too complicated to find theoretically optimal values.

In case the parameter set is a Euclidean space, the use of a filter bank becomes impossible. If the model satisfies certain smoothness conditions with respect to the parameters, we may use gradient climbing methods to find values for the test statistics. Sometimes it may even be possible to solve the maximum likelihood problem exactly, as in the case of the known change of unknown intensity.

In the simulations, the detection algorithms are tested for several types of changes. The so-called general damages are most easily detected, followed by the so-called parallel scratches and, finally, the orthogonal scratches. Obviously, the more is known about the changes that may appear, the higher the detection quality of the algorithms is. If the amount of available information on the changes decreases, the difference between the three classes also decreases. Furthermore, a positive correlation is clearly present between the Kullback information of the change and the power of the test.

Since the quality of detection deteriorates rapidly if the knowledge about the changes decreases, we do not expect these algorithms to be very efficient in practice. Only if the changes that may appear are well known and easily parameterized, the power of these tests may reach satisfactorily levels.

Another drawback of these algorithms is the computational complexity. Since we need a complete Kalman filter for each possible change, the computational load becomes quite large

when the number of changes increases. Moreover, if the dimension of the state vector becomes large, which is not at all unlikely for image processing applications, even one Kalman filter would be computationally expensive.

Another aspect that may be of interest for future research is the quickest detection problem. Since the test uses the same rejection boundary for all columns of the field, this threshold is inclined to be rather high in order to avoid an excessive number of false alarms. As a result, most of the detections take place at one of the last columns of the field. Using an increasing rejection boundary may therefore improve the quickness of detection.

A Appendix.

A.1 Generation of innovations.

The Kalman filter equations under the null hypothesis are given by

$$\begin{aligned}
\hat{X}_{k+1|k}(0) &= A_k \hat{X}_{k|k}(0) \\
\hat{X}_{k|k}(0) &= \hat{X}_{k|k-1}(0) + K_k \gamma_k(0) \\
\gamma_k(0) &= Y_k - C_k \hat{X}_{k|k-1}(0) \\
P_{k+1|k}(0) &= A_k P_{k|k}(0) A_k' + F_k R_W(k) F_k' \\
R_k(0) &= C_k P_{k|k-1}(0) C_k' + R_V(k) \\
K_k(0) &= P_{k|k-1}(0) C_k' R_k(0)^{-1} \\
P_{k|k}(0) &= (I - K_k(0) C_k) P_{k|k-1}(0)
\end{aligned}$$

Under the alternative hypothesis, the Kalman filter with the correct parameters is given by

$$\begin{aligned}
\hat{X}_{k+1|k}(\theta) &= A_k(\theta) \hat{X}_{k|k}(\theta) + E_k^X(\theta) \\
\hat{X}_{k|k}(\theta) &= \hat{X}_{k|k-1}(\theta) + K_k(\theta) \gamma_k(\theta) \\
\gamma_k(\theta) &= Y_k - C_k(\theta) \hat{X}_{k|k-1}(\theta) - E_k^Y(\theta) \\
P_{k+1|k}(\theta) &= A_k(\theta) P_{k|k}(\theta) A_k(\theta)' \\
&\quad + F_k(\theta) R_W(k) F_k(\theta)' \\
R_k(\theta) &= C_k(\theta) P_{k|k-1}(\theta) C_k(\theta)' + R_V(k) \\
K_k(\theta) &= P_{k|k-1}(\theta) C_k(\theta)' (R_k(\theta))^{-1} \\
P_{k|k}(\theta) &= [I - K_k(\theta) C_k(\theta)] P_{k|k-1}(\theta)
\end{aligned}$$

As we may see, the additive changes $E_k^X(\theta)$ and $E_k^Y(\theta)$ only require some changes in the structure of the Kalman filter.

A.2 Statistical properties of the innovations.

Under the null hypothesis, the innovation $\gamma_k(0)$ is distributed as a Gaussian white noise with zero mean and covariance $R_k(0)$. The alternative innovation $\gamma_k(\theta)$, given that the parameter θ is correct, is distributed as an independent Gaussian variable with zero mean and covariance $R_k(\theta)$. The distribution of $\gamma_k(0)$ becomes more complicated if the null hypothesis does not hold. We have

$$\begin{aligned}
\gamma_k(0) &= Y_k - C_k \hat{X}_{k|k-1}(0) \\
&= Y_k - C_k(\theta) \hat{X}_{k|k-1}(\theta) - E_k^Y(\theta) + C_k(\theta) \hat{X}_{k|k-1}(\theta) + E_k^Y(\theta) - C_k \hat{X}_{k|k-1}(0) \\
&= \gamma_k(\theta) + C_k(\theta) \hat{X}_{k|k-1}(\theta) - C_k \hat{X}_{k|k-1}(0) + E_k^Y(\theta)
\end{aligned}$$

so that we may see that $\gamma_k(0)$ is no longer independent. However, being a linear combination of Gaussian variables, the innovation remains Gaussian.

The expectation may easily be found to be

$$\begin{aligned}
\rho_k(0;\theta) &= \mathbf{E}_\theta \gamma_k(0) \\
&= C_k(\theta) \mu_{k|k-1}(\theta;\theta) - C_k \mu_{k|k-1}(0;\theta) + E_k^Y(\theta)
\end{aligned}$$

where

$$\begin{aligned}
\mu_{k+1|k}(\theta;\theta) &= \mathbf{E}_\theta \hat{X}_{k+1|k}(\theta) \\
&= A_k(\theta) \mu_{k|k-1}(\theta;\theta) + E_k^X(\theta) \\
\mu_{k+1|k}(0;\theta) &= \mathbf{E}_\theta \hat{X}_{k+1|k}(0) \\
&= A_k [\mu_{k|k-1}(0;\theta) + K_k \rho_k(0;\theta)]
\end{aligned}$$

For the covariance, we obtain

$$\begin{aligned} R_k(0;\theta) &= \mathbf{E}_\theta(\gamma_k(0)-\rho_k(0;\theta))(\gamma_k(0)-\rho_k(0;\theta))' \\ &= R_k(\theta)+C_k(\theta)Q_{k|k-1}(\theta,\theta;\theta)C_k(\theta)'-C_k(\theta)Q_{k|k-1}(\theta,0;\theta)C_k' \\ &\quad -C_kQ_{k|k-1}(0,\theta;\theta)C_k(\theta)'+C_kQ_{k|k-1}(0,0;\theta)C_k' \end{aligned}$$

where

$$Q_{k|k-1}(\theta_1,\theta_2;\theta_3)=\mathbf{E}_{\theta_3}(\hat{X}_{k|k-1}(\theta_1)-\mu_{k|k-1}(\theta_1;\theta_3))(\hat{X}_{k|k-1}(\theta_2)-\mu_{k|k-1}(\theta_2;\theta_3))'$$

The complete algorithms for the computation of these matrices follows later. The cross-covariance between the innovations at different times may be calculated as $(i>j)$

$$\begin{aligned} R_{i,j}(0;\theta) &= \mathbf{E}_\theta(\gamma_i(0)-\rho_i(0;\theta))(\gamma_j(0)-\rho_j(0;\theta))' \\ &= C_i(\theta)P_{i,j}(\theta,0;\theta)-C_iP_{i,j}(0,0;\theta) \\ P_{i+1,j}(\theta,0;\theta) &= \mathbf{E}_\theta(\hat{X}_{i+1|i}(0)-\mu_{i+1|i}(\theta;\theta))(\gamma_j(0)-\rho_j(0;\theta))' \\ &= A_i(\theta)P_{i,j}(\theta,0;\theta) \\ P_{j,j}(\theta,0;\theta) &= Q_{j|j-1}(\theta,\theta;\theta)C_j(\theta)'-Q_{j|j-1}(\theta,0;\theta)C_j' \\ P_{i+1,j}(0,0;\theta) &= \mathbf{E}_\theta(\hat{X}_{i+1|i}(0)-\mu_{i+1|i}(0;\theta))(\gamma_j(0)-\rho_j(0;\theta))' \\ &= A_i[P_{i,j}(0,0;\theta)+K_i(0)R_{i,j}(0;\theta)] \\ P_{j,j}(0,0;\theta) &= Q_{j|j-1}(0,\theta;\theta)C_j(\theta)'-Q_{j|j-1}(0,0;\theta)C_j' \end{aligned}$$

As a result, we may find that the new variable

$$\boldsymbol{\gamma}_k(0)=\begin{pmatrix}\gamma_1(0)\\ \vdots \\ \gamma_k(0)\end{pmatrix}$$

is, under the condition that θ is the correct parameter, distributed as a Gaussian variable with mean

$$\boldsymbol{\rho}_k(0;\theta)=\begin{pmatrix}\rho_1(0;\theta)\\ \vdots \\ \rho_k(0;\theta)\end{pmatrix}$$

and covariance matrix

$$\mathbf{R}_k(0;\theta)=\{R_{i,j}(0;\theta)\}_{i,j=1,\ldots,k}$$

Similarly, we may find the expectation and covariance for $\gamma_k(\theta)$ under the null hypothesis. The expectation is given by

$$\begin{aligned} \rho_k(\theta;0) &= \mathbf{E}_0\gamma_k(\theta) \\ &= C_k\mu_{k|k-1}(0;0)-C_k(\theta)\mu_{k|k-1}(\theta;0)-E_k^Y(\theta) \end{aligned}$$

where

$$\begin{aligned} \mu_{k+1|k}(0;0) &= \mathbf{E}_0\hat{X}_{k+1|k}(0) \\ &= A_k\mu_{k|k-1}(0;0) \\ \mu_{k+1|k}(\theta;0) &= \mathbf{E}_0\hat{X}_{k+1|k}(\theta) \\ &= A_k(\theta)[\mu_{k|k-1}(\theta;0)+K_k(\theta)\rho_k(\theta;0)]+E_k^X(\theta) \end{aligned}$$

For the covariance we find

$$\begin{aligned} R_k(\theta;0) &= R_k(0) + C_k Q_{k|k-1}(0,0;0)C_k' - C_k Q_{k|k-1}(0,\theta;0)C_k(\theta)' \\ &\quad -C_k(\theta)Q_{k|k-1}(\theta,0;0)C_k' + C_k(\theta)Q_{k|k-1}(\theta,\theta;0)C_k(\theta)' \end{aligned}$$

The cross-covariance may now be calculated according to $(i > j)$

$$\begin{aligned} R_{i,j}(\theta;0) &= \mathbf{E}_0(\gamma_i(\theta) - \rho_i(\theta;0))(\gamma_j(\theta) - \rho_j(\theta;0))' \\ &= C_i P_{i,j}(0,\theta;0) - C_i(\theta)P_{i,j}(\theta,\theta;0) \\ P_{i+1,j}(0,\theta;0) &= \mathbf{E}_0(\hat{X}_{i+1|i}(0) - \mu_{i+1|i}(0;0))(\gamma_j(\theta) - \rho_j(\theta;0))' \\ &= A_i P_{i,j}(0,\theta;0) \\ P_{j,j}(0,\theta;0) &= Q_{j|j-1}(0,0;0)C_j' - Q_{j|j-1}(0,\theta;0)C_j(\theta)' \\ P_{i+1,j}(\theta,\theta;0) &= \mathbf{E}_0(\hat{X}_{i+1|i}(\theta) - \mu_{i+1|i}(\theta;0))(\gamma_j(\theta) - \rho_j(\theta;0))' \\ &= A_i(\theta)[P_{i,j}(\theta,\theta;0) + K_i(\theta)R_{i,j}(\theta;0)] \\ P_{j,j}(\theta,\theta;0) &= Q_{j|j-1}(\theta,0;0)C_j' - Q_{j|j-1}(\theta,\theta;0)C_j(\theta)' \end{aligned}$$

The variable

$$\boldsymbol{\gamma}_k(\theta) = \begin{pmatrix} \gamma_1(\theta) \\ \vdots \\ \gamma_k(\theta) \end{pmatrix}$$

is under the null hypothesis distributed as a Gaussian variable with mean

$$\boldsymbol{\rho}_k(\theta;0) = \begin{pmatrix} \rho_1(\theta;0) \\ \vdots \\ \rho_k(\theta;0) \end{pmatrix}$$

and covariance matrix

$$\mathbf{R}_k(\theta;0) = \{R_{i,j}(\theta;0)\}_{i,j=1,\ldots,k}$$

A.3 The likelihood ratio.

Expressed in the innovations, the hypotheses may be written as

$$\begin{aligned} H_0 &: \gamma_k(0) \sim N(0, R_k(0)) \\ H_1 &: (\exists\theta \in \Theta\backslash\{0\})\gamma_k(\theta) \sim N(0, R_k(\theta)) \end{aligned}$$

The LR is given by

$$\begin{aligned} \ell_n(Y,\theta) &= \log \frac{f_\theta(Y)}{f_0(Y)} \\ &= \log \prod_{k=1}^{n} \frac{f_\theta(Y_k|Y_{k-1},\ldots,Y_1)}{f_0(Y_k|Y_{k-1},\ldots,Y_1)} \\ &= \sum_{k=1}^{n} \log \frac{f_\theta(\gamma_k(\theta))}{f_0(\gamma_k(0))} \end{aligned}$$

In this case, we may evaluate it as

$$\ell_n(Y,\theta) = \frac{1}{2}\sum_{k=1}^{n} \log \frac{|R_k(0)|}{|R_k(\theta)|} + \gamma_k(0)'R_k(0)^{-1}\gamma_k(0) - \gamma_k(\theta)'R_k(\theta)^{-1}\gamma_k(\theta)$$

Denote $s_k(Y,\theta)$ as twice the increment of the log-likelihood ratio,

$$s_k(Y,\theta) = \log\frac{|R_k(0)|}{|R_k(\theta)|} + \gamma_k(0)'R_k(0)^{-1}\gamma_k(0) - \gamma_k(\theta)'R_k(\theta)^{-1}\gamma_k(\theta)$$

It has expectations

$$\begin{aligned}
\mathbf{E}_0 s_k(Y,\theta) &= \log\frac{|R_k(0)|}{|R_k(\theta)|} + N - \mathrm{tr}\, R_k(\theta;0)R_k(\theta)^{-1} - \rho_k(\theta;0)'R_k(\theta)^{-1}\rho_k(\theta;0)\\
\mathbf{E}_\theta s_k(Y,\theta) &= \log\frac{|R_k(0)|}{|R_k(\theta)|} - N + \mathrm{tr}\, R_k(0;\theta)R_k(0)^{-1} + \rho_k(0;\theta)'R_k(0)^{-1}\rho_k(0;\theta)
\end{aligned}$$

A third possibility is the mismatch. That is, while $*$ is the correct parameter, $\theta \neq *$ is detected. In this case, the expectation of the increment is given by

$$\begin{aligned}
\mathbf{E}_* s_k(Y,\theta) &= \log\frac{|R_k(0)|}{|R_k(\theta)|} + \mathrm{tr}\, R_k(0;*)R_k(0)^{-1} + \rho_k(0;*)'R_k(0)^{-1}\rho_k(0:*)\\
&\quad -\mathrm{tr}\, R_k(\theta;*)R_k(\theta)^{-1} - \rho_k(\theta;*)'R_k(\theta)^{-1}\rho_k(\theta;*)
\end{aligned}$$

with

$$\begin{aligned}
\rho_k(\theta;*) &= \mathbf{E}_*\gamma_k(\theta)\\
&= C_k(*)\mu_{k|k-1}(*;*) - C_k(\theta)\mu_{k|k-1}(\theta;*) + E_k^Y(*) - E_k^Y(\theta)\\
\mu_{k+1|k}(\theta;*) &= \mathbf{E}_*\hat{X}_{k+1|k}(\theta)\\
&= A_k(\theta)[\mu_{k|k-1}(\theta;*) + K_k(\theta)\rho_k(\theta;*)] + E_k^X(\theta)\\
R_k(\theta;*) &= \mathbf{E}_*(\gamma_k(\theta) - \rho_k(\theta;*))(\gamma_k(\theta) - \rho_k(\theta;*))'\\
&= R_k(*) + C_k(*)Q_{k|k-1}(*,*;*)C_k(*)' - C_k(*)Q_{k|k-1}(*,\theta;*)C_k(\theta)'\\
&\quad -C_k(\theta)Q_{k|k-1}(\theta,*;*)C_k(*)' + C_k(\theta)Q_{k|k-1}(\theta,\theta;*)C_k(\theta)'
\end{aligned}$$

A.4 The Q-matrices.

The algorithms for the computation of the Q-matrices are given below. All initial values are zero.

$$\begin{aligned}
Q_{k+1|k}(\theta,\theta;0) &= A_k(\theta)[Q_{k|k-1}(\theta,\theta;0) + (Q_{k|k-1}(\theta,0;0)C_k' - Q_{k|k-1}(\theta,\theta;0)C_k(\theta)')K_k(\theta)'\\
&\quad +K_k(\theta)(C_kQ_{k|k-1}(0,\theta;0) - C_k(\theta)Q_{k|k-1}(\theta,\theta;0)) + K_k(\theta)R_k(\theta;0)K_k(\theta)']A_k(\theta)'\\
Q_{k+1|k}(\theta,0;0) &= A_k(\theta)[(I - K_k(\theta)C_k(\theta))Q_{k|k-1}(\theta,0;0) + K_k(\theta)C_kQ_{k|k-1}(0,0;0)\\
&\quad +K_k(\theta)R_k(0)K_k(0)']A_k'\\
Q_{k+1|k}(0,0;0) &= A_k[Q_{k|k-1}(0,0;0) + K_k(0)R_k(0)K_k(0)']A_k'\\
Q_{k+1|k}(\theta,\theta;\theta) &= A_k(\theta)[Q_{k|k-1}(\theta,\theta;\theta) + K_k(\theta)R_k(\theta)K_k(\theta)']A_k(\theta)'\\
Q_{k+1|k}(0,\theta;\theta) &= A_k[(I - K_k(0)C_k)Q_{k|k-1}(0,\theta;\theta) + K_k(0)C_k(\theta)Q_{k|k-1}(\theta,\theta;\theta)\\
&\quad +K_k(0)R_k(\theta)K_k(\theta)']A_k(\theta)'\\
Q_{k+1|k}(0,0;\theta) &= A_k[Q_{k|k-1}(0,0;\theta) + Q_{k|k-1}(0,\theta;\theta)C_k(\theta)'K_k(0)' - Q_{k|k-1}(0,0;\theta)C_k'K_k(0)'\\
&\quad +K_k(0)C_k(\theta)Q_{k|k-1}(\theta,0;\theta) - K_k(0)C_kQ_{k|k-1}(0,0;\theta) + K_k(0)R_k(0;\theta)K_k(0)']A_k'\\
Q_{k+1|k}(\theta,*;*) &= A_k(\theta)[(I - K_k(\theta)C_k(\theta))Q_{k|k-1}(\theta,*;*) + K_k(\theta)C_k(*)Q_{k|k-1}(*,*;*)\\
&\quad +K_k(\theta)R_k(*)K_k(*)']A_k(*)'\\
Q_{k+1|k}(\theta,\theta;*) &= A_k(\theta)[Q_{k|k-1}(\theta,\theta;*) + K_k(\theta)[C_k(*)Q_{k|k-1}(*,\theta;*) - C_k(\theta)Q_{k|k-1}(\theta,\theta;*)]\\
&\quad +[Q_{k|k-1}(\theta,*;*)C_k(*)' - Q_{k|k-1}(\theta,\theta;*)C_k(\theta)']K_k(\theta)'\\
&\quad +K_k(\theta)R_k(\theta;*)K_k(\theta)']A_k(\theta)'
\end{aligned}$$

A.5 The efficient score and the Fisher information matrix.

The efficient score and the Fisher information matrix may be calculated as

$$\begin{aligned}
\{\mathcal{Z}_n(Y,0)\}_i &= \frac{1}{2}\sum_{k=1}^{n} \mathrm{tr}\, R_k(0)^{-1}[(\gamma_k(0)\gamma_k(0)' R_k(0)^{-1} - I)\frac{\partial R_k(\theta)}{\partial \theta_i}(0) \\
&\quad +2\frac{\partial(\gamma_k(\theta)-\rho_k(\theta))}{\partial\theta_i}(0)\gamma_k(0)'] \\
\{\Gamma_n(0)\}_{ij} &= \frac{1}{2}\sum_{k=1}^{n} \mathrm{tr}\, R_k(0)^{-1}\{[\frac{\partial R_k(\theta)}{\partial\theta_i}(0) - \frac{\partial C_k(\theta)}{\partial\theta_j}P_{k|k-1}C_k' - C_k P_{k|k-1}\frac{\partial C_k(\theta)'}{\partial\theta_j}] \\
&\quad R_k(0)^{-1}\frac{\partial R_k(\theta)}{\partial\theta_j}(0) + 2[\frac{\partial C_k(\theta)}{\partial\theta_i}P_{k|k-1}\frac{\partial C_k(\theta)'}{\partial\theta_j} + C_k\tilde{P}_{k|k-1}(i,j)C_k']\}
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial(\gamma_k(\theta)-\rho_k(\theta))}{\partial\theta_i}(0) &= -\frac{\partial C_k(\theta)}{\partial\theta_i}\hat{X}_{k|k-1}(0) - C_k\frac{\partial}{\partial\theta_i}(\hat{X}_{k|k-1}+\xi_k)(0) - \frac{\partial E_k^Y(\theta)}{\partial\theta_i} \\
\frac{\partial}{\partial\theta_i}(\hat{X}_{k+1|k}+\xi_{k+1})(0) &= \frac{\partial A_k(\theta)}{\partial\theta_i}[\hat{X}_{k|k-1}+K_k(0)\gamma_k(0)] \\
&\quad +A_k[\frac{\partial}{\partial\theta_i}(\hat{X}_{k|k-1}+\xi_k)(0) + \frac{\partial K_k(\theta)}{\partial\theta_i}(0)\gamma_k(0) \\
&\quad +K_k(0)\frac{\partial(\gamma_k(\theta)-\rho_k(\theta))}{\partial\theta_i}(0)] + \frac{\partial E_k^X(\theta)}{\partial\theta_i} \\
\tilde{P}_{k+1|k}(i,j) &= \mathbf{E}_0\frac{\partial(\tilde{X}_{k+1|k}-\xi_{k+1})}{\partial\theta_i}(0)\frac{\partial(\tilde{X}_{k+1|k}-\xi_{k+1})'}{\partial\theta_j}(0) \\
&= \{\frac{\partial A_k(\theta)}{\partial\theta_i}[I-K_k(0)C_k] - A_k[\frac{\partial K_k(\theta)}{\partial\theta_i}(0)C_k + K_k(0)\frac{\partial C_k(\theta)}{\partial\theta_i}]\} \\
&\quad P_{k|k-1}\{\frac{\partial A_k(\theta)}{\partial\theta_j}[I-K_k(0)C_k] - A_k[\frac{\partial K_k(\theta)}{\partial\theta_j}(0)C_k + K_k(0)\frac{\partial C_k(\theta)}{\partial\theta_j}]\}' \\
&\quad +A_k[I-K_k(0)C_k]\tilde{P}_{k|k-1}(i,j)[I-K_k(0)C_k]'A_k' \\
&\quad +(\frac{\partial A_k(\theta)}{\partial\theta_i}K_k(0) + A_k\frac{\partial K_k(\theta)}{\partial\theta_j}(0))R_V(\frac{\partial A_k(\theta)}{\partial\theta_j}K_k(0) + A_k\frac{\partial K_k(\theta)}{\partial\theta_j}(0))' \\
&\quad +\frac{\partial F_k(\theta)}{\partial\theta_i}R_W\frac{\partial F_k(\theta)'}{\partial\theta_j} \\
\frac{\partial R_k(\theta)}{\partial\theta_i}(0) &= \frac{\partial C_k(\theta)}{\partial\theta_i}P_{k|k-1}C_k' + C_k\frac{\partial P_{k|k-1}(\theta)}{\partial\theta_i}(0)C_k' \\
&\quad +C_kP_{k|k-1}\frac{\partial C_k(\theta)}{\partial\theta_i}' \\
\frac{\partial P_{k+1|k}(\theta)}{\partial\theta_i}(0) &= \frac{\partial A_k(\theta)}{\partial\theta_i}P_{k|k}A_k' + A_k\frac{\partial P_{k|k}(\theta)}{\partial\theta_i}(0)A_k' + A_kP_{k|k}\frac{\partial A_k(\theta)}{\partial\theta_i}' \\
&\quad +\frac{\partial F_k(\theta)}{\partial\theta_i}R_W(k)F_k' + F_kR_W(k)\frac{\partial F_k(\theta)}{\partial\theta}' \\
\frac{\partial P_{k|k}(\theta)}{\partial\theta_i}(0) &= -[\frac{\partial K_k(\theta)}{\partial\theta_i}(0)C_k + K_k\frac{\partial C_k(\theta)}{\partial\theta_i}]P_{k|k-1} + [I-K_kC_k]\frac{\partial P_{k|k-1}(\theta)}{\partial\theta_i} \\
\frac{\partial K_k(\theta)}{\partial\theta_i}(0) &= \frac{\partial P_{k|k-1}(\theta)}{\partial\theta_i}C_k'R_k(0)^{-1} + P_{k|k-1}\frac{\partial C_k(\theta)}{\partial\theta_i}'R_k(0)^{-1} \\
&\quad -P_{k|k-1}C_k'R_k(0)^{-1}\frac{\partial R_k(\theta)}{\partial\theta_i}R_k(0)^{-1}
\end{aligned}$$

A.6 The generalized likelihood ratio for the special case.

The derivative of the log-likelihood ratio with respect to $c = (c_X, c_Y)'$ now takes a rather simple form

$$\begin{aligned}\mathcal{Z}_n(Y,c) &= \frac{\partial \ell_n(Y,c)}{\partial c} \\ &= \frac{1}{2}\sum_{k=1}^{n} \gamma_k(c)' R_k(\theta_0)^{-1} \frac{\partial \gamma_k(c)}{\partial c}\end{aligned}$$

Let us write

$$\hat{X}_{k|k-1}(c) = (\, \xi_k^X(\theta_0) \quad \xi_k^Y(\theta_0)\,)\, c + \nu_k(\theta_0)$$

Then

$$\begin{aligned}c_X\xi_{k+1}^X(\theta_0) + c_Y\xi_{k+1}^Y(\theta_0) + \nu_{k+1}(\theta_0) &= A_k(\theta_0)\hat{X}_{k|k}(c) + c_X E_k^X(\theta_0) \\ &= A_k(\theta_0)(\hat{X}_{k|k-1}(c) + K_k(\theta_0)\gamma_k(c)) + c_X E_k^X(\theta_0) \\ &= A_k(\theta_0)[I - K_k(\theta_0)C_k(\theta_0)]\hat{X}_{k|k-1}(c) + A_k(\theta_0)K_k(\theta_0)Y_k \\ &\quad -c_Y A_k(\theta_0)K_k(\theta_0)E_k^Y(\theta_0) + c_X E_k^X(\theta_0) \\ &= A_k(\theta_0)[I - K_k(\theta_0)C_k(\theta_0)](c_X\xi_k^X(\theta_0) + c_Y\xi_k^Y(\theta_0) + \nu_k(\theta_0)) \\ &\quad +A_k(\theta_0)K_k(\theta_0)Y_k - c_Y A_k(\theta_0)K_k(\theta_0)E_k^Y(\theta_0) + c_X E_k^X(\theta_0)\end{aligned}$$

so that we may write

$$\begin{aligned}\xi_{k+1}^X(\theta_0) &= A_k(\theta_0)[I - K_k(\theta_0)C_k(\theta_0)]\xi_k^X(\theta_0) + E_k^X(\theta_0) \\ \xi_{k+1}^Y(\theta_0) &= A_k(\theta_0)[I - K_k(\theta_0)C_k(\theta_0)]\xi_k^Y(\theta_0) - A_k(\theta_0)K_k(\theta_0)E_k^Y(\theta_0) \\ \nu_{k+1}(\theta_0) &= A_k(\theta_0)[I - K_k(\theta_0)C_k(\theta_0)]\nu_k(\theta_0) + A_k(\theta_0)K_k(\theta_0)Y_k\end{aligned}$$

As a result, we may write the innovation as

$$\begin{aligned}\gamma_k(c) &= Y_k - C_k(\theta_0)\nu_k(\theta_0) - c_X C_k(\theta_0)\xi_k^X(\theta_0) - c_Y[C_k(\theta_0)\xi_k^Y(\theta_0) + E_k^Y(\theta_0)] \\ &= G_k(\theta_0) - H_k(\theta_0)c\end{aligned}$$

where

$$\begin{aligned}G_k(\theta_0) &= Y_k - C_k(\theta_0)\nu_k(\theta_0) \\ H_k(\theta_0) &= (\, C_k(\theta_0)\xi_k^X(\theta_0) \quad C_k(\theta_0)\xi_k^Y(\theta_0) + E_k^Y(\theta_0)\,)\end{aligned}$$

Note that $H_k(\theta_0)$ is a deterministic variable.

Substituting the innovation in the derivative of the loglikelihood ratio, and putting the result equal to zero, we obtain

$$\sum_{k=1}^{n}(G_k(\theta_0) - H_k(\theta_0)\hat{c}(n))' R_k(\theta_0)^{-1} H_k(\theta_0) = 0$$

which has a solution

$$\hat{c}(n) = (\sum_{k=1}^{n} H_k(\theta_0)' R_k(\theta_0)^{-1} H_k(\theta_0))^{-1} \sum_{k=1}^{n} H_k(\theta_0)' R_k(\theta_0)^{-1} G_k(\theta_0)$$

This implies that in stead of computing the estimated state $\hat{X}_{k|k-1}(c\theta_0)$, we should compute the three new variables $\xi_k^X(\theta_0)$, $\xi_k^Y(\theta_0)$ and $\nu_k(\theta_0)$. ¿From these variables the optimal value of c and the corresponding maximum value of the LR may be calculated.

The maximum value of twice the LR is

$$\begin{aligned} T_n(Y) &= \sum_{k=1}^{n}(\log\frac{|R_k(0)|}{|R_k(\theta_0)|} + \gamma_k(0)'R_k(0)^{-1}\gamma_k(0) - G_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0)) \\ &\quad + \sum_{k=1}^{n} G_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)(\sum_{k=1}^{n} H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0))^{-1}\sum_{k=1}^{n} H_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0) \end{aligned}$$

For the variable $G_k(\theta_0)$, we have

$$\begin{aligned} \mathbf{E}_c G_k(\theta_0) &= H_k(\theta_0)c \\ \mathbf{cov}_c G_k(\theta_0) &= R_k(\theta_0) \end{aligned}$$

so that

$$\begin{aligned} \mathbf{E}_c\hat{c}(n) &= c \\ \mathbf{cov}_c\hat{c}(n) &= \mathbf{E}_{c\theta_0}(\hat{c}(n)-c)(\hat{c}(n)-c)' \\ &= (\sum_{k=1}^{n} H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0))^{-1} \end{aligned}$$

The estimator is unbiased, and it is consistent if the summation over the quadratic term diverges.

For the expectation of the test statistic under the assumption that c is the correct intensity we may now find

$$\mathbf{E}_c T_n(Y) = 2 + \sum_{k=1}^{n}\log\frac{|R_k(0)|}{|R_k(\theta_0)|} - N + \operatorname{tr} R_k(0;\theta_0)R_k(0)^{-1} + c'\varrho_k(0;\theta_0)'R_k(0)^{-1}\varrho_k(0;\theta_0)c \quad (1)$$

Here $\varrho_k(0;\theta_0)$ is a $M \times 2$ matrix such that

$$\mathbf{E}_c\gamma_k(0) = \varrho_k(0;\theta_0)c$$

which may be easily shown to exist.

$$\begin{aligned} \mathbf{E}_c\gamma_k(0) &= a_k + b_k c \\ &= C_k(\theta_0)\mu_{k|k-1}(c;\theta,c) - C_k\mu_{k|k-1}(0;\theta_0,c) + (\,0 \quad E_k^Y(\theta_0)\,)\,c \\ &= C_k(\theta_0)(f_k + g_k c) - C_k(m_k + n_k c) + (\,0 \quad E_k^Y(\theta_0)\,)\,c \\ a_k &= C_k(\theta_0)f_k - C_k m_k \\ b_k &= C_k(\theta_0)g_k - C_k n_k + (\,0 \quad E_k^Y(\theta_0)\,) \\ \mu_{k+1|k}(c;\theta_0,c) &= f_{k+1} + g_{k+1}c \\ &= A_k(\theta_0)\mu_{k|k-1}(c;\theta_0,c) + (\,E_k^X(\theta_0) \quad 0\,)\,c \\ &= A_k(\theta_0)(f_k + g_k c) + (\,E_k^X(\theta_0) \quad 0\,)\,c \\ f_{k+1} &= A_k(\theta_0)f_k \\ g_{k+1} &= A_k(\theta_0)g_k + (\,E_k^X(\theta_0) \quad 0\,) \\ \mu_{k+1|k}(0;\theta_0,c) &= m_{k+1} + n_{k+1}c \\ &= A_k[\mu_{k|k-1}(0;\theta_0,c) + K_k(0)\mathbf{E}_c\gamma_k(0)] \\ &= A_k[m_k + n_k c + K_k(0)(a_k + b_k c)] \\ m_{k+1} &= A_k(m_k + K_k(0)a_k) \\ &= A_k[(I - K_k(0)C_k)m_k + K_k(0)C_k(\theta_0)f_k] \\ n_{k+1} &= A_k(n_k + K_k(0)b_k) \\ &= A_k[(I - K_k(0)C_k)n_k + K_k(0)C_k(\theta_0)g_k + K_k(0)\,(\,0 \quad E_k^Y(\theta_0)\,)] \end{aligned}$$

Since all variables have zero initial conditions and may only be triggered by $E_k^X(\theta_0)$ and $E_k^Y(\theta_0)$, the only variables that may attain nonzero values are g_k, n_k and b_k, so that indeed

$$\mathbf{E}_c\gamma_k(0) = b_k c = \varrho_k(0;\theta_0)c$$

Now the expectation of the first quadratic form of the GLR is

$$\mathbf{E}_c\gamma_k(0)'R_k(0)^{-1}\gamma_k(0) = \operatorname{tr} R_k(0;\theta_0)R_k(0)^{-1} + c'\varrho_k(0;\theta_0)'R_k(0)^{-1}\varrho_k(0;\theta_0)c$$

Similarly, the expectation for the second quadratic form is

$$\mathbf{E}_c G_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0) = N + c'H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)c$$

Finally, the last quadratic form has expectation

$$\begin{aligned}
& \mathbf{E}_c[\sum_{k=1}^n G_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)(\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0))^{-1}\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0)] \\
= \quad & \mathbf{E}_c\operatorname{tr}[\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}G_k(\theta_0)\sum_{k=1}^n G_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)(\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0))^{-1}] \\
= \quad & \operatorname{tr}[\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)(\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0))^{-1}] + c'\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)c \\
= \quad & 2 + c'\sum_{k=1}^n H_k(\theta_0)'R_k(\theta_0)^{-1}H_k(\theta_0)c
\end{aligned}$$

The resulting expression (1) now follows easily.

Note that the sum of the weighted squares of H_k has to be invertible. In case this assumption does not hold, for example if one of the parameters c_X or c_Y equals zero, the only thing that changes is a reduction of the number 2 that appears in the final equation.

The expectation of the likelihood ratio may now take a different form. Given that the estimate of θ_0 is correct, the expectation of the LR is given by equation (1). Otherwise, assume that $\theta \neq \theta_0$ is the estimate of θ_0. Then we may show that

$$\mathbf{E}_{\theta_0,c}T_n(Y,\theta) = \Lambda_n(\theta;\theta_0) + c'\Gamma_n(\theta;\theta_0)c$$

Note that, although in general the basic inequality

$$\mathbf{E}_* T_n(Y,*) \geq \mathbf{E}_* T_n(Y,\theta)$$

holds for all θ, this no longer is true in this case. The fact that the new statistic $T_n(Y,\theta)$ also contains the estimate of the parameter c, alters the statistical character of this statistic. The estimate of c is no longer deterministic, so that the expectation of the statistic changes.

For the derivation of the expectation, we proceed as follows. We may write

$$G_k(\theta) = \gamma_k(\theta,c) + H_k(\theta)c$$

This implies that its covariance matrix may be obtained from the previous results,

$$\mathbf{cov}_{\,\theta_0,c}G_k(\theta) = R_k(\theta;\theta_0)$$

For the calculation of the expectation, we should use another approach.

$$\begin{aligned}
\mathbf{E}_{\theta_0,c}G_k(\theta) &= \mathbf{E}_{\theta,c}\{G_k(\theta_0) + C_k(\theta_0)\nu_k(\theta_0) - C_k(\theta)\nu_k(\theta)\} \\
&= H_k(\theta_0)c + C_k(\theta_0)\phi_k(\theta_0;\theta_0)c - C_k(\theta)\phi_k(\theta;\theta_0)c \\
&= \varrho_k(\theta;\theta_0)c
\end{aligned}$$

where

$$\begin{aligned}\phi_{k+1}(\theta_0;\theta_0) &= A_k(\theta_0)\phi_k(\theta_0;\theta_0,c) + A_k(\theta_0)K_k(\theta_0)H_k(\theta_0)\\ \phi_{k+1}(\theta;\theta_0) &= A_k(\theta)\phi_k(\theta;\theta_0) + A_k(\theta)K_k(\theta)\varrho_k(\theta;\theta_0)\end{aligned}$$

The fact that this expectation is again linear in c may be shown similarly as in the previous case.

As a result, the following equalities hold.

$$\mathbf{E}_{\theta_0,c} G_k(\theta)' R_k(\theta)^{-1} G_k(\theta) = \operatorname{tr} R_k(\theta;\theta_0) R_k(\theta)^{-1} + c'\varrho_k(\theta;\theta_0)' R_k(\theta)^{-1}\varrho_k(\theta;\theta_0)c$$

$$\begin{aligned}&\mathbf{E}_{\theta_0,c}\sum_{k=1}^n G_k(\theta)' R_k(\theta)^{-1} H_k(\theta)(\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} H_k(\theta))^{-1}\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} G_k(\theta)\\ =\ &\operatorname{tr}\{\sum_{k=1}^n\sum_{l=1}^n H_k(\theta)' R_k(\theta)^{-1} R_{k,l}(\theta;\theta_0) R_l(\theta)^{-1} H_l(\theta)(\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} H_k(\theta))^{-1}\}\\ &+\sum_{k=1}^n c'\varrho_k(\theta;\theta_0)' R_k(\theta)^{-1} H_k(\theta)(\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} H_k(\theta))^{-1}\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1}\varrho_k(\theta;\theta_0)c\end{aligned}$$

The cross-covariance between G_k and G_l is given by (assume $k > l$)

$$\begin{aligned}R_{k,l}(\theta;\theta_0) &= \mathbf{cov}_{\theta_0,c}(G_k(\theta), G_l(\theta))\\ &= \mathbf{cov}_{\theta_0,c}(\gamma_k(\theta,c), \gamma_l(\theta,c))\\ &= C_k(\theta_0)P^{\nu\gamma}_{k,l}(\theta_0) - C_k(\theta)P^{\nu\gamma}_{k,l}(\theta;\theta_0) + C_k(\theta_0)L_{k,l}(\theta_0,\theta_0;\theta_0)C_l(\theta_0)'\\ &\quad -C_k(\theta_0)L_{k,l}(\theta_0,\theta;\theta_0)C_l(\theta)' - C_k(\theta)L_{k,l}(\theta,\theta_0;\theta_0)C_l(\theta_0)' + C_k(\theta)L_{k,l}(\theta,\theta;\theta_0)C_l(\theta)'\end{aligned}$$

Here

$$\begin{aligned}P^{\nu\gamma}_{k+1,l}(\theta_0) &= A_k(\theta_0)P^{\nu\gamma}_{k,l}(\theta_0)\\ P^{\nu\gamma}_{l+1,l}(\theta_0) &= A_l(\theta_0)K_l(\theta_0)R_l(\theta_0)\\ P^{\nu\gamma}_{k+1,l}(\theta;\theta_0) &= A_k(\theta)[I - K_k(\theta)C_k(\theta)]P^{\nu\gamma}_{k,l}(\theta;\theta_0) + A_k(\theta)K_k(\theta)C_k(\theta_0)P^{\nu\gamma}_{k,l}(\theta_0)\\ P^{\nu\gamma}_{l+1,l}(\theta;\theta_0) &= A_l(\theta)K_l(\theta)R_l(\theta_0)\end{aligned}$$

Finally, the cross-covariances between ν_k and ν_l are given by

$$\begin{aligned}L_{k+1,l}(\theta_0,\theta_0;\theta_0) &= A_k(\theta_0)L_{k,l}(\theta_0,\theta_0;\theta_0)\\ L_{k+1,l}(\theta_0,\theta;\theta_0) &= A_k(\theta_0)L_{k,l}(\theta_0,\theta;\theta_0)\\ L_{k+1,l}(\theta,\theta_0;\theta_0) &= A_k(\theta)[I - K_k(\theta)C_k(\theta)]L_{k,l}(\theta,\theta_0;\theta_0) + A_k(\theta)K_k(\theta)C_k(\theta_0)L_{k,l}(\theta_0,\theta_0;\theta_0)\\ L_{k+1,l}(\theta,\theta;\theta_0) &= A_k(\theta)[I - K_k(\theta)C_k(\theta)]L_{k,l}(\theta,\theta;\theta_0) + A_k(\theta)K_k(\theta)C_k(\theta_0)L_{k,l}(\theta_0,\theta;\theta_0)\end{aligned}$$

Finally, gathering these results we may see that

$$\begin{aligned}\Lambda_n(\theta;\theta_0) &= \sum_{k=1}^n\{\log\frac{|R_k(0)|}{|R_k(\theta)|} + \operatorname{tr} R_k(0;\theta_0)R_k(0)^{-1} - \operatorname{tr} R_k(\theta;\theta_0)R_k(\theta)^{-1}\}\\ &\quad +\operatorname{tr}\{\sum_{k=1}^n\sum_{l=1}^n H_k(\theta)' R_k(\theta)^{-1} R_{k,l}(\theta;\theta_0) R_l(\theta)^{-1} H_l(\theta)(\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} H_k(\theta))^{-1}\}\\ \Gamma_n(\theta;\theta_0) &= \sum_{k=1}^n\{\varrho_k(0;\theta_0)' R_k(0)^{-1}\varrho_k(0;\theta_0) - \varrho_k(\theta;\theta_0)' R_k(\theta)^{-1}\varrho_k(\theta;\theta_0)\}\\ &\quad +\sum_{k=1}^n \varrho_k(\theta;\theta_0)' R_k(\theta)^{-1} H_k(\theta)(\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1} H_k(\theta))^{-1}\sum_{k=1}^n H_k(\theta)' R_k(\theta)^{-1}\varrho_k(\theta;\theta_0)\end{aligned}$$

References

[1] P. M. Newbold and Yu-Chi Ho. Detection of changes in the characteristics of a Gauss-Markov process. *IEEE Transactions on Aerospace and Electronic Systems*, 4:707–718, 1968.

[2] R. K. Mehra and J. Peschon. An innovations approach to fault detection and diagnosis in dynamic systems. *Automatica*, 7:637–640, 1971.

[3] Alan S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12:601–611, 1976.

[4] Ron Patton, Paul Frank, and Robert Clark, editors. *Fault Diagnosis in Dynamic Systems: Theory and Applications*. Prentice Hall International Series in Systems and Control Engineering. Prentice Hall, 1989.

[5] Anil K. Jain and Jaswant R. Jain. Partial differential equations and finite difference methods in image processing, part II: Image restoration. *IEEE Transactions on Automatic Control*, 23:817–834, 1978.

[6] Erik P. Hupkens and Arunabha Bagchi. Detection of parametric changes in autoregressive fields. Memorandum 1297, University of Twente, 1995.

[7] Anil K. Jain. Advances in mathematical models for image processing. *Proceedings of the IEEE*, 69:502–528, 1981.

[8] Richard C. Dubes and Anil K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16:131–164, 1989.

[9] K. Kumamaru, S. Sagara, and T. Söderström. Some statistical methods for fault diagnosis for dynamical systems. In Patton et al. [4], chapter 13, pages 439–475.

[10] Alan S. Willsky. Detection of abrupt changes in dynamical systems. In Michèle Basseville and Albert Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, volume 77 of *Lecture Notes in Control and Information Sciences*, pages 27–49. Springer, 1986.

[11] C. R. Rao. *Linear Statistical Inference and Its Applications*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, 1965.

[12] David A. Wilson and Alok Kumar. Derivative computations for the log likelihood function. *IEEE Transactions on Automatic Control*, 27:230–232, 1982.

[13] Mordechai Segal and Ehud Weinstein. A new method for evaluating the log-likelihood gradient (score) of linear dynamic systems. *IEEE Transactions on Automatic Control*, 33:763–766, 1988.

[14] Robert P. Leland. A new formula for the log-likelihood gradient for continuous-time stochastic systems. *IEEE Transactions on Automatic Control*, 40:1295–1300, 1995.

[15] H. Vincent Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, second edition, 1994.

[16] Anil K. Jain. Partial differential equations and finite-difference methods in image processing, part 1: Image representation. *Journal of Optimization Theory and Applications*, 23:65–91, 1977.

Reports recently issued in this series[1]

1280. W. Albers and M. Teulings, *A simple estimator for correlation in incomplete data* (September, 1995)

1281. T. Böhme, H.J. Broersma and H.J. Veldman *Toughness and longest cycles in 2-connected planar graphs* (September, 1995)

1282. W. Albers, G.R.J. Arts and W.C.M. Kallenberg, *Test limits for correlated measurements under prescribed consumer risk* (September, 1995)

1283. I.W.M. Pothof and J. Schut, *Graph-theoretic approach to identifiability in a water distribution network* (October, 1995)

1284. T.S.H. Driessen, *On the core and kernel of the bilateral assignment market* (October, 1995)

1285. T.S.H. Driessen, *Tree enterprises and bankruptcy ventures: a game theoretic similarity due to a graph theoretic proof* (October, 1995)

1286. T.S.H. Driessen, *An alternative game theoretic analysis of a bankruptcy problem from the talmud: the case of the greedy bankruptcy game* (November, 1995)

1287. R.C.W. Strijbos, *Calculation of right matrix fraction descriptions; an algorithm* (November, 1995)

1288. R.C.W. Strijbos, *A polynomial toolbox: implementaion aspects* (November, 1995)

1289. W. Albers and C. Schut, *A survey of rank tests for classes of tail alternatives* (November, 1995)

1290. A.J. van der Schaft and J.M. Schumacher, *The complementary-slackness class of hybrid systems* (November, 1995)

1291. F. Göbel and C. Hoede, *Magic labelings of graphs* (November, 1995)

1292. X. Li, *On graphs with equal clique and independence polynomials* (November, 1995)

1293. S. Brandt and H.J. Veldman, *Degree sums for edges and cycle lengths in graphs* (November, 1995)

1294. E.R. Fledderus and E.W.C. van Groesen, *Deformation of coherent structures* (November, 1995)

1295. W. Albers, W.C.M. Kallenberg and G.D. Otten, *Accurate test limits under nonnormal measurement error* (November, 1995)

1296. P. Wittmann, *Vertex-distinguishing edge-colorings of 2-regular graphs* (November, 1995)

[1] Requests for Memoranda should be addressed to the author(s)

1297. E.P. Hupkens and A. Bagchi, *Detection of parametric changes in autoregressive fields* (November, 1995)

1298. H.J. Broersma and X. Li, *Spanning trees with many or few colors in edge-colored graphs* (November, 1995)

1299. W.C.M. Kallenberg, T. Ledwina, *Data driven smooth tests when the hypothesis is composite* (November, 1995)

1300. J.A. Bondy, H.J. Broersma, J.P.M. van den Heuvel and H.J. Veldman, *Heavy cycles in weighted graphs* (November, 1995)

1301. W.C.A. Maas and M. Dalsmo, *Tracking of a rigid body in $SE(3)$ using singular H_∞ control* (November, 1995)

1302. J.W.Polderman and K. Arent, *Almost optimal adaptive LQ control: SISO case* (November, 1995)

1303. F. Göbel, *On a problem Selfridge* (December, 1995)

1304. A. Loria and H. Nijmeijer, *Output feedback tracking control of Euler-Lagrange systems via bounded controls* (December, 1995)

1305. G. Still, *Generic perturbations of one-parametric symmetric eigenvalue problems* (December, 1995)

1306. C. Hoede, *A comparison of some conditions for non-Hamiltonicity of graphs* (December, 1995)

1307. W.W.E. Wetterling, *Philon's line generalized: An optimization problem from geometry* (December, 1995)

1308. P. Jonker, G. Still and F. Twilt, *Stratifications of classes of matrices: The connected components* (December, 1995)

1309. C. Hoede and X. Li, *Word graphs: The first set* (December, 1995)

1310. H.J. Broersma and X. Li, *Directed path graphs* (January, 1996)

1311. M. Streng, *Load balancing for computational fluid dynamics claculations* (February, 1996)

1312. U. Faigle and W. Kern, *Approximate core allocation for binpacking games* (February, 1996)

1313. C. Hoede, *On the Hamiltonicity of graphs $G(n, k, t)$* (February, 1996)

1314. A. Al-Dhahir, *The Hénon map* (March, 1996)

1315. W. Albers, G.R.J. Arts and W.C.M. Kallenberg, *Test regions using two or more correlated product characteristics* (March, 1996)

1316. I.J.B.F. Adan, E.A. van Doorn, J.A.C. Resing and W.R.W. Scheinhardt, *A queueing system with two different service speeds* (March, 1996)