

# A DETERMINISTIC DISCRETISATION-STEP UPPER BOUND FOR STATE ESTIMATION VIA CLARK TRANSFORMATIONS

W. P. MALCOLM , R. J. ELLIOTT, AND J. VAN DER HOEK

*Received 11 November 2003 and in revised form 2 June 2004*

We consider the numerical stability of discretisation schemes for continuous-time state estimation filters. The dynamical systems we consider model the indirect observation of a continuous-time Markov chain. Two candidate observation models are studied. These models are (a) the observation of the state through a Brownian motion, and (b) the observation of the state through a Poisson process. It is shown that for robust filters (via Clark's transformation), one can ensure nonnegative estimated probabilities by choosing a maximum grid step to be no greater than a given bound. The importance of this result is that one can choose an *a priori* grid step maximum ensuring nonnegative estimated probabilities. In contrast, no such upper bound is available for the standard approximation schemes. Further, this upper bound also applies to the corresponding robust smoothing scheme, in turn ensuring stability for smoothed state estimates.

## 1. Introduction

In much of the literature concerning stochastic numerics, for continuous-time filters, the main emphasis is placed upon minimising errors in estimation, for example, see [9].

However, there are indeed other equally important criteria concerning the implementation of continuous-time filters. One example is the issue of numerical stability; in particular, there is a well-known flaw in the Euler-Maruyama scheme applied to the Wonham filter, that is, the estimated probabilities can be negative (see [9, page 448]). Despite negative probabilities being meaningless in state estimation, this particular problem has received little attention in the literature.

In this paper, we show that one can guarantee nonnegative state estimation probabilities by using the so-called "robust" filter due to Clark and making a judicious choice for the maximum subinterval in a discretisation partition. It is shown that there exists a simple deterministic upper bound for the maximum time step (in a discretisation), ensuring nonnegative probabilities with the robust filters due to Clark. It is also shown that no such bounds exist for the more standard discretisation schemes, such as the Euler-Maruyama and the Milstein. The robust filter ideas of Clark are also considered here in

the context of Markov-modulated Poisson process observations. In this scenario, one can construct a robust filter; however, the meaning of continuous dependence in the space of sample paths relies on the Skorokhod metric, as the observations are cadlag and belong to the space  $D(0, \infty)$ . Continuous dependence in this sense is established for a robust filter driven by Poisson observations.

## 2. Model dynamics

Two observation models are considered, each describing an indirect observation of a continuous-time Markov chain, whose dynamics we now describe. Initially all processes are defined on the fixed probability space  $(\Omega, \mathcal{F}, P)$ .

**2.1. State process dynamics.** Suppose a state process  $X = \{X_t, 0 \leq t\}$  is a finite-state time-homogeneous Markov chain evolving in continuous time. Without loss of generality, we can take the state space of  $X$  as  $\mathcal{L} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} \subseteq \mathbb{R}^n$ , where  $\mathbf{e}_i$  denotes a column vector in  $\mathbb{R}^n$  with unity in the  $i$ th position and zero elsewhere. The dynamics for this process are

$$X_t = X_0 + \int_0^t A X_u du + M_t, \quad (2.1)$$

where  $M$  is a  $(P, \sigma\{X_u, 0 \leq u \leq t\})$ -martingale and  $A$  is an  $n \times n$  rate matrix.

## 2.2. Observation process dynamics

*2.2.1. Observation through a Brownian motion.* We suppose that the process  $X$  is not observed directly, rather, we observe a scalar-valued process

$$y_t = \int_0^t \langle X_u, \mathbf{g} \rangle du + W_t. \quad (2.2)$$

Here  $W$  is a standard Wiener process and  $\mathbf{g} = (\langle \mathbf{g}, \mathbf{e}_1 \rangle, \dots, \langle \mathbf{g}, \mathbf{e}_n \rangle)' \in \mathbb{R}^n$  is a vector of the so-called drift coefficients, or levels for the Markov chain.

*2.2.2. Observation through a Poisson process.* We suppose that the process  $X$  is not observed directly, rather, we observe a scalar-valued univariate Poisson process with intensity model

$$\lambda_t = \langle X_t, \boldsymbol{\lambda} \rangle. \quad (2.3)$$

Here  $\boldsymbol{\lambda} = (\langle \boldsymbol{\lambda}, \mathbf{e}_1 \rangle, \dots, \langle \boldsymbol{\lambda}, \mathbf{e}_n \rangle)' \in \mathbb{R}_+^n$ . The dynamics for  $N$  have the form

$$N_t = \int_0^t \langle X_u, \boldsymbol{\lambda} \rangle du + V_t. \quad (2.4)$$

Here the process  $V$  is a  $(P, \sigma\{N_u, 0 \leq u \leq t\})$ -martingale.

*Remark 2.1.* The observation processes  $y$  and  $N$  are each scalar-valued. However, the results in this paper are routinely extended to vector-valued models.

*Remark 2.2.* Equation (2.4) can also be interpreted as a counting measure. For example, suppose  $A_t \triangleq (0, t]$  and the sequence  $\{\tau_\ell\}_{\ell \geq 1}$  is a sequence of jump epochs for a Poisson process. Then

$$N_t = \text{card}\{n \mid \tau_n \in A_t\} \tag{2.5}$$

exhibits the interpretation of a Poisson process as a random counting measure; see, for example, [8, 10].

**2.3. Reference probability.** The filters we consider in this paper are in the form of dynamics for unnormalised probabilities. Such filters can be computed with reference probability techniques and Girsanov’s theorem, or versions of Girsanov’s theorem. Central to this approach is the abstract form of Bayes’ rule.

*Notation.* Write  $\mathcal{Y}_t$  for either information in  $\sigma\{y_u \mid 0 \leq u \leq t\}$  or  $\sigma\{N_u \mid 0 \leq u \leq t\}$ . Suppose  $y = \{y_u, 0 \leq u \leq t\}$  is a process and we wish to estimate  $E[y_t \mid \mathcal{Y}_t]$ . Using a form of Bayes’ rule [2],

$$E[y_t \mid \mathcal{Y}_t] = \frac{E^\dagger[\Lambda_t y_t \mid \mathcal{Y}_t]}{E^\dagger[\Lambda_t \mid \mathcal{Y}_t]}. \tag{2.6}$$

Here  $E^\dagger[\cdot]$  denotes expectation under a reference measure  $P^\dagger$  and  $\Lambda$  denotes a Radon-Nikodym derivative  $dP/dP^\dagger$ . Further details on the reference probability methods can be found in [2, 3]. Finally, suppose we consider the observation model given at (2.2). Then, using the numerator in (2.6), we write

$$q_t \triangleq E^\dagger[\Lambda_t X_t \mid \sigma\{y_u \mid 0 \leq u \leq t\}] \in \mathbb{R}^n. \tag{2.7}$$

**3. State estimation filters**

Here we recall some state estimation filters whose stability we wish to investigate.

**3.1. Filters for  $X$  observed through a Wiener process**

**THEOREM 3.1** (Wonham, 1965). *Suppose the process  $X$  satisfies dynamics given by (2.1) and a process  $y$  satisfies the dynamics at (2.2).*

*With  $q_t \triangleq E^\dagger[\Lambda_t X_t \mid \mathcal{Y}_t]$ ,*

$$q_t = q_0 + \int_0^t A q_u du + \int_0^t \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} q_u dy_u. \tag{3.1}$$

To determine the corresponding normalised probability for the dynamics at (3.1), one computes, for example,

$$p(X_t = \mathbf{e}_i \mid \mathcal{Y}_t) = \frac{\langle q_t, \mathbf{e}_i \rangle}{\langle q_t, \mathbf{1} \rangle}. \tag{3.2}$$

In [1], it was shown that the process  $q$  satisfying the dynamics at (3.1) could be transformed to a new process whose dynamics do not involve stochastic integration. The importance of this result cannot be understated, as it eliminates the numerical difficulties concerning the approximation of stochastic integrals.

*Definition 3.2.* Define a matrix-valued stochastic process  $\Phi \in \mathbb{R}^{n \times n}$ , where

$$\Phi_t = \text{diag}\{\phi_t^1, \phi_t^2, \dots, \phi_t^n\} \tag{3.3}$$

with  $\phi_t^i = \exp(\langle \mathbf{g}, \mathbf{e}_i \rangle y_t - (1/2) |\langle \mathbf{g}, \mathbf{e}_i \rangle|^2 t)$ .

**THEOREM 3.3** (Clark [1]). Write  $\bar{q}_t \triangleq \Phi_t^{-1} q_t$ . The process  $\bar{q}$  satisfies the linear ordinary differential equation

$$\frac{d\bar{q}_t}{dt} = \Phi_t^{-1} A \Phi_t \bar{q}_t, \quad \bar{q}_0 = q_0. \tag{3.4}$$

Conversely, the process  $\Phi \bar{q}$  satisfies (3.1) when  $\bar{q}$  satisfies (3.4).

**LEMMA 3.4.** The quantity

$$\pi_t(X) \triangleq \frac{\Phi_t \bar{q}_t}{\langle \Phi_t \bar{q}_t, \mathbf{1} \rangle} \tag{3.5}$$

defines a locally Lipschitz continuous version of the expectation  $E[X_t | \mathcal{Y}_t]$ .

**Lemma 3.4** is established in [1, 7].

### 3.2. Filters for $X$ observed through a Poisson process

**THEOREM 3.5.** Suppose the process  $X$  satisfies dynamics given by (2.1). Suppose a Poisson process  $N$  is observed whose intensity model has the form

$$\lambda_t = \langle X_t, \boldsymbol{\lambda} \rangle = \sum_{i=1}^n 1_{\{X_t = \mathbf{e}_i\}} \langle \boldsymbol{\lambda}, \mathbf{e}_i \rangle. \tag{3.6}$$

With  $q_t \triangleq E^\dagger[\Lambda_t X_t | \mathcal{Y}_t]$ ,

$$q_t = q_0 + \int_0^t A q_u du + \int_0^t \text{diag}\{\langle \boldsymbol{\lambda}, \mathbf{e}_i \rangle - 1\} q_u - (dN_u - du). \tag{3.7}$$

#### 3.2.1. A robust filter for Poisson observations

*Definition 3.6.* Define a matrix-valued stochastic process  $\Gamma \in \mathbb{R}^{n \times n}$ , where

$$\Gamma_t = \text{diag}\{\gamma_t^1, \gamma_t^2, \dots, \gamma_t^n\} \tag{3.8}$$

with  $\gamma_t^i = \exp((1 - \langle \boldsymbol{\lambda}, \mathbf{e}_i \rangle)t) \langle \boldsymbol{\lambda}, \mathbf{e}_i \rangle^{N_t}$ ,  $i = 1, \dots, n$ .

**THEOREM 3.7.** Write  $\bar{q}_t \triangleq \Gamma_t^{-1} q_t$ . The process  $\bar{q}$  satisfies the linear ordinary differential equation

$$\frac{d\bar{q}_t}{dt} = \Gamma_t^{-1} A \Gamma_t \bar{q}_t, \quad \bar{q}_0 = q_0. \tag{3.9}$$

Conversely, the process  $\Gamma \bar{q}$  satisfies (3.7) when  $\bar{q}$  satisfies (3.9).

The process  $\Gamma \bar{q}$  satisfies (3.7).

**Theorem 3.7** was established in [11].

**LEMMA 3.8.** The quantity

$$\pi_t(X) \triangleq \frac{\Gamma_t \bar{q}_t}{\langle \Gamma_t \bar{q}_t, \mathbf{1} \rangle} \tag{3.10}$$

defines a Skorokhod continuous version of the expectation  $E[X_t \mid \mathcal{Y}_t]$ .

*Proof of Lemma 3.8.* Suppose

$$\begin{aligned} N(\omega_1) &= \{N_t(\omega_1), 0 \leq t \leq T\}, \\ N(\omega_2) &= \{N_t(\omega_2), 0 \leq t \leq T\} \end{aligned} \tag{3.11}$$

are two counting process observation paths. The distance between the two counting process paths will be defined in terms of the Skorokhod metric:

$$d(N(\omega_1), N(\omega_2)) \triangleq \inf_{\lambda} \left\{ \sup_{0 \leq t \leq T} |\lambda(t) - t| \vee \sup_{0 \leq t \leq T} |N_t(\omega_1) - N_{\lambda(t)}(\omega_2)| \right\}. \tag{3.12}$$

Here, the infimum is taken over the set of increasing functions  $\lambda : [0, T] \rightarrow [0, T]$  such that  $\lambda(0) = 0$  and  $\lambda(T) = T$ . That is, each  $\lambda$  gives a time change on  $[0, T]$ . Clearly for counting processes, when  $d(N(\omega_1), N(\omega_2)) < 1$ , the two processes  $N(\omega_1), N(\omega_2)$  have the same number of jumps on  $[0, T]$ . Suppose this is the case and suppose the jumps of  $N(\omega_1)$  occur at times  $T_i, 1 \leq i \leq k$ , and that those of  $N(\omega_2)$  occur at  $S_i, i \leq i \leq k$ . Then

$$d(N(\omega_1), N(\omega_2)) = \max_{1 \leq i \leq k} |T_i - S_i|. \tag{3.13}$$

Now

$$\begin{aligned} \bar{q}_t(\omega_1) &= \bar{q}_0 + \int_0^t \Phi_u^{-1}(\omega_1) A \Phi_u(\omega_1) \bar{q}_u(\omega_1) du, \\ \bar{q}_t(\omega_2) &= \bar{q}_0 + \int_0^t \Phi_u^{-1}(\omega_2) A \Phi_u(\omega_2) \bar{q}_u(\omega_2) du. \end{aligned} \tag{3.14}$$

$\Phi_u(\omega_1) = \Phi_u(\omega_2)$  except where  $N_u(\omega_1) \neq N_u(\omega_2)$ . Therefore, it follows that

$$\begin{aligned} \|\bar{q}_t(\omega_1) - \bar{q}_t(\omega_2)\| &\leq C \int_0^t \|\bar{q}_u(\omega_1) - \bar{q}_u(\omega_2)\| du + D \sum_{1 \leq i \leq k} |T_i - S_i| \\ &\leq C \int_0^t \|\bar{q}_u(\omega_1) - \bar{q}_u(\omega_2)\| du + D k d(N(\omega_1), N(\omega_2)). \end{aligned} \tag{3.15}$$

Letting

$$\begin{aligned} \phi(t) &= \max_{0 \leq u \leq t} \|\bar{q}_u(\omega_1) - \bar{q}_u(\omega_2)\|, \\ \phi(t) &\leq C \int_0^t \phi(u) du + Dkd(N(\omega_1), N(\omega_2)), \end{aligned} \tag{3.16}$$

and using Gronwall's inequality, we have that

$$\sup_{0 \leq t \leq T} \|\bar{q}_t(\omega_1) - \bar{q}_t(\omega_2)\| \leq Kd(N(\omega_1), N(\omega_2)). \tag{3.17}$$

□

#### 4. Discretisation schemes

For all time discretisations, we will consider a partition, on the interval  $[0, T]$ , and write

$$\Pi^{(K)} = \{0 = t_0, t_1, \dots, t_K = T\}. \tag{4.1}$$

Here the partition is strict, that is,  $t_0 < t_1 < \dots$ .

To denote the mesh of the partition, we write

$$\|\Pi^{(K)}\| = \max_{1 \leq k \leq K} \{t_k - t_{k-1}\}. \tag{4.2}$$

For brevity, we will use the notation  $\xi_k \stackrel{\Delta}{=} \xi_{t_k}$ , where  $\xi_k$  denotes a process  $\xi$  at a time point  $t_k$ .

**4.1. Observation through a Brownian motion.** The discrete-time recursions given here are standard. These schemes can be developed by approximating stochastic Taylor series expansions; for example, see [9].

(1) *The Euler-Maruyama scheme:*

$$q_k = [\mathbf{I} + \Delta A + \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} (y_k - y_{k-1})] q_{k-1}. \tag{4.3}$$

(2) *The Milstein scheme:*

$$\begin{aligned} q_k &= [\mathbf{I} + \Delta A + \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} (y_k - y_{k-1})] q_{k-1} \\ &\quad + \left[ \frac{1}{2} \left( (y_k - y_{k-1})^2 - \Delta \right) \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle^2\} \right] q_{k-1}. \end{aligned} \tag{4.4}$$

(3) *Order-1 strong Taylor scheme:*

$$\begin{aligned} q_k &= [\mathbf{I} + \Delta A + \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} (y_k - y_{k-1})] q_{k-1} \\ &\quad + \left[ \frac{1}{2} \left( (y_k - y_{k-1})^2 - \Delta \right) \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle^2\} \right. \\ &\quad + \frac{1}{2} A^2 \Delta^2 + \frac{1}{2} (A \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} + \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle\} A) (y_k - y_{k-1}) \Delta \\ &\quad \left. + \frac{1}{6} \text{diag}\{\langle \mathbf{g}, \mathbf{e}_i \rangle^3\} \left( (y_k - y_{k-1})^3 - 3(y_k - y_{k-1}) \Delta \right) \right] q_{k-1}. \end{aligned} \tag{4.5}$$

(4) *Robust discretisation schemes* (see [1, 7]):

$$q_k = \Phi_k \Phi_{k-1}^{-1} [\mathbf{I} + \Delta A] q_{k-1}. \tag{4.6}$$

*Remark 4.1.* Note that in each of the approximate recursions (4.3), (4.4), and (4.5), the difference  $y_k - y_{k-1}$  appears explicitly. However, in the robust recursion at (4.6), this difference appears as an argument of the exponentials in the matrix product  $\Phi_k \Phi_{k-1}^{-1}$ .

**4.2. Observation through a Poisson process**

(1) *The Euler-Maruyama scheme:*

$$q_k = q_{k-1} + A q_{k-1} \Delta + \text{diag}\{\langle \boldsymbol{\lambda}, \mathbf{e}_i \rangle - 1\} q_{k-1} ((N_k - N_{k-1}) - \Delta). \tag{4.7}$$

(2) *Robust discretisation schemes* (see [11]):

$$q_k = \Gamma_k \Gamma_{k-1}^{-1} [\mathbf{I} + \Delta A] q_{k-1}. \tag{4.8}$$

**5. Discretisation limits**

*Definition 5.1.* A numerical implementation of dynamics to compute the estimated unnormalised probability  $q_k$ , either for an observation of the process  $X$  through a Brownian motion, or a Poisson process, is said to be stable on  $\Pi^{(K)}$  if for each  $i \in \{1, 2, \dots, n\}$  and for each  $k \in \{1, 2, \dots, K\}$ , the following inequality holds:

$$\langle q_k, \mathbf{e}_i \rangle \geq 0. \tag{5.1}$$

**5.1. Observation through a Brownian motion**

**THEOREM 5.2.** *The robust time-discretised dynamics at (4.6) are stable on a partition  $\Pi^{(K)}$ , provided the following inequality is satisfied:*

$$\|\Pi^K\| = \max_k (t_k - t_{k-1}) \leq \frac{1}{\max\{|a_{(i,i)}|\}}. \tag{5.2}$$

*Proof of Theorem 5.2.* Consider the  $i$ th component of the vector  $q_k$ . Without loss of generality, we take  $\langle q_{k-1}, \mathbf{e}_i \rangle \geq 0$  for each  $i$ . Recalling the dynamics at (4.6), we see that

$$\begin{aligned} \langle q_k, \mathbf{e}_i \rangle &= \langle \Phi_k \Phi_{k-1}^{-1} [\mathbf{I} + \Delta_t A] \langle q_{k-1}, \mathbf{e}_i \rangle, \mathbf{e}_i \rangle \\ &= \xi_{k,k-1}^i \langle q_{k-1}, \mathbf{e}_i \rangle - \xi_{k,k-1}^i \Delta_t^i |a_{(i,i)}| \langle q_{k-1}, \mathbf{e}_i \rangle + \sum_{\substack{i=1 \\ i \neq j}}^n \xi_{k,k-1}^i \Delta_t^i a_{(i,j)} \langle q_{k-1}, \mathbf{e}_i \rangle. \end{aligned} \tag{5.3}$$

Here  $\xi_{k,k-1}^i = \exp(g_i(y_k - y_{k-1}) - (1/2)|g_i|^2 \Delta_t^i)$ . The stability condition given in Definition 5.1 requires that the left-hand side of (5.3) remain nonnegative, that is,

$$\xi_{k,k-1}^i \langle q_{k-1}, \mathbf{e}_i \rangle - \xi_{k,k-1}^i \Delta_t^i |a_{(i,i)}| \langle q_{k-1}, \mathbf{e}_i \rangle + \sum_{\substack{i=1 \\ i \neq j}}^n \xi_{k,k-1}^i \Delta_t^i a_{(i,j)} \langle q_{k-1}, \mathbf{e}_i \rangle \geq 0. \tag{5.4}$$

Simplifying this inequality, we get

$$(1 - \Delta_t^i |a_{(i,i)}|) \langle q_{k-1}, \mathbf{e}_i \rangle + \Delta_t^i \sum_{\substack{i=1 \\ i \neq j}}^n a_{(i,j)} \langle q_{k-1}, \mathbf{e}_i \rangle \geq 0. \tag{5.5}$$

Since the off-diagonal elements of the matrix  $A$  are always nonnegative, the term concerning these elements in (5.5) is always nonnegative, that is,

$$\Delta_t^i \sum_{\substack{i=1 \\ i \neq j}}^n a_{(i,j)} \langle q_{k-1}, \mathbf{e}_i \rangle \geq 0. \tag{5.6}$$

So, to ensure the inequality at (5.5) is satisfied, we need only choose  $\Delta_t^i$  such that the quantity  $(1 - \Delta_t^i |a_{(i,i)}|)$  is nonnegative, that is,

$$\Delta_t^i \leq \frac{1}{|a_{(i,i)}|}. \tag{5.7}$$

The corresponding global upper limit is, therefore,

$$\Delta_t \leq \frac{1}{\max_i \{|a_{(i,i)}|\}}. \tag{5.8}$$

□

*Remark 5.3.* It is interesting to note that the bound given by [Theorem 5.2](#) does not depend upon the parameters  $g_1, \dots, g_n$  and depends only on those elements along the main diagonal of the matrix  $A$ .

To emphasise the value of this result, consider a similar calculation for the corresponding Euler-Maruyama scheme given at (4.3). By imposing the same stability demand and carrying out calculation such as those above, we get an inconclusive result, that is,

$$(1 - \Delta_t^i |a_{(i,i)}|) \langle q_{k-1}, \mathbf{e}_i \rangle + \Delta_t^i \sum_{\substack{i=1 \\ i \neq j}}^n a_{(i,j)} \langle q_{k-1}, \mathbf{e}_i \rangle + g^i \langle q_k, \mathbf{e}_i \rangle (y_k - y_{k-1}) \geq 0. \tag{5.9}$$

Here there is simply no choice of  $\Delta_t^i$  one can make to ensure that inequality (5.9) is satisfied, as the left-hand side of this inequality is stochastic, depending both upon the magnitude and sign of the difference  $y_k - y_{k-1}$ . Moreover, carrying out the same calculations for the Milstein and higher-order schemes also results in stochastic inequalities involving the difference  $y_k - y_{k-1}$ . In contrast, the upper bound given by [Theorem 5.2](#) is deterministic and therefore holds for any observation sample path.

*Remark 5.4.* The robust Wonham filter can be extended to a robust smoother using the ideas first introduced in [4, 11, 12]. For these smoothers, one computes a backward recursion very similar to the recursion at (4.6). Smoothed estimates are obtained by combining forward and backward recursions. It can also be shown, that the stability



in [Definition 5.1](#) holds for the (robust) smoothed state estimates. Again, there is no such stability for the corresponding nonrobust discretisation of smoothing schemes. Further, in [\[5\]](#), a discretization-step upper bound is obtained for  $M$ -ary detection filters.

**5.2. Observation through a Poisson process**

For the models with Poisson observations, one can apply the Euler scheme or the robust discretisation. In contrast to the Wonham filter, the stochastic integration in the filter at [\(3.7\)](#) is an integral against a process of bounded variation. Further, if  $N_k - N_{k-1} \neq 0$ , then  $N_k - N_{k-1} = 1$ , provided the discretisation is chosen so at most one jump can occur in any subinterval of time.

**THEOREM 5.5** (Poisson process models). *For the robust discretisation [\(4.8\)](#) and for any set of nonnegative Poisson intensities  $\{\lambda_1, \dots, \lambda_n\}$ , the stability given by [Definition 5.1](#) is guaranteed  $P$ -a.s. by choosing a maximum grid step such that*

$$\|\Pi^K\| = \max_k (t_k - t_{k-1}) \leq \frac{1}{\max_i \{ |a_{(i,i)}| \}}. \tag{5.10}$$

The proof of [Theorem 5.5](#) is very similar to the proof of [Theorem 5.2](#) and so is omitted. To emphasise the value of the upper bound in [Theorem 5.5](#), consider again a similar calculation for the corresponding Euler discretisation. The result of this calculation is

$$\max_k (t_k - t_{k-1}) \leq \frac{1}{\max_i \{ |a_{(i,i)}| + \langle \lambda, e_i \rangle - 1 \}}. \tag{5.11}$$

While the inequality at [\(5.11\)](#) is not stochastic, it does depend upon the parameters  $\lambda_i$ . Further, it is strictly less than the upper bound at [\(5.10\)](#). What this means is that the robust discretisation will tolerate a ‘‘coarser’’ partition. This might be of advantage when considering reductions in computation.

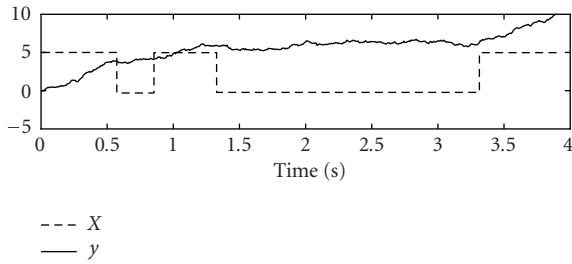
The filter for Poisson observations given at [\(3.7\)](#) is in some ways quite distinct to the Wonham filter. For example, in between jump events in the observation process, it is essentially a parabolic partial differential equation. Suppose, for example, that the first jump time is  $\tau_1(\omega)$ . Then on the interval  $(0, \tau_1(\omega))$ , the filter dynamics are

$$q_t = q_0 + \int_0^t \{A - \text{diag}\{\langle \lambda, e_i \rangle - 1\}\} q_u du. \tag{5.12}$$

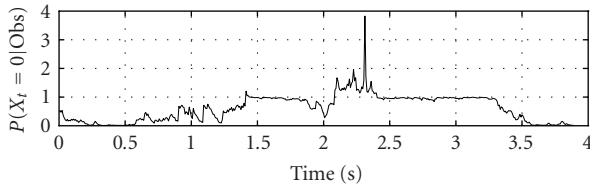
This admits the explicit solution

$$q_t = \exp((A - \text{diag}\{\langle \lambda, e_i \rangle - 1\})t) q_0 \quad \text{on } (0, \tau_1(\omega)). \tag{5.13}$$

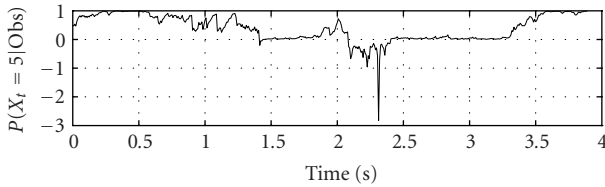
In general, the matrices  $A$  and  $\text{diag}\{\langle \lambda, e_i \rangle - 1\}$  do not commute, so the dynamics at [\(5.13\)](#) cannot be further simplified. To implement the dynamics at [\(5.13\)](#) requires computing



(a)



(b)



(c)

Figure 6.1. Euler-Maruyama approximation to the Wonham filter.

the matrix exponential which is not trivial [13]. To avoid this matrix exponential, one might consider a first-order approximation, that is,

$$q_t = \{I + \Delta(A - \text{diag}\{\langle \lambda, e_i \rangle - 1\})\} q_0. \tag{5.14}$$

However, these dynamics can result in negative probabilities, as is shown in the examples below.

### 6. Examples

The simulation studies here include two examples, each illustrating the benefits of using the discrete-time recursions based upon the robust filters. In the first example, we

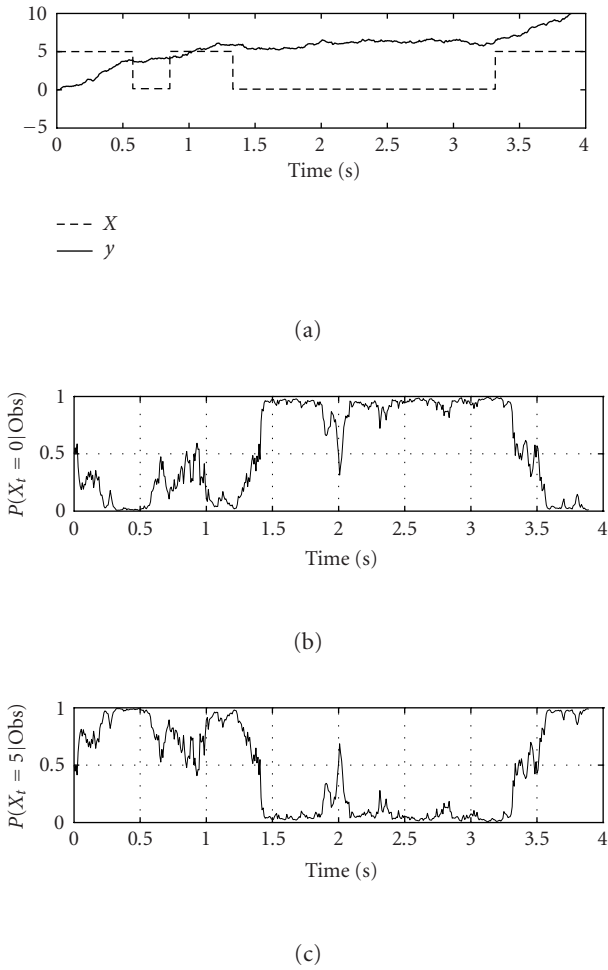


Figure 6.2. Robust approximation to the Wonham filter.

consider the robust Wonham filter and, in particular, an example studied in [9] (see pages 447-448). The model parameters considered are

$$\begin{aligned}
 A &= \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix}, \\
 \langle \mathbf{g}, \mathbf{e}_1 \rangle &= 0, \\
 \langle \mathbf{g}, \mathbf{e}_2 \rangle &= 5.
 \end{aligned}
 \tag{6.1}$$

For this study, a regular discretisation of  $[0,4]$  was used with a time step  $\Delta = 2^{-7}$ . The plots given in Figure 6.1 show realisation of the state and observation processes and the estimated probabilities computed by using the Euler-Maruyama approximation to

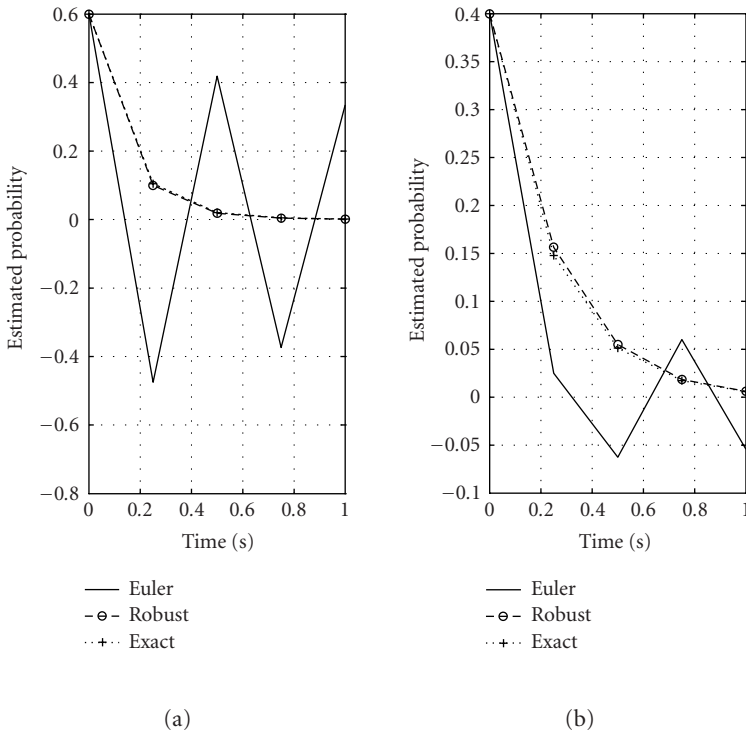


Figure 6.3. Various approximations to the Poisson process filter in between jump events : (a)  $\text{Prob}(X_t = e_1)$  and (b)  $\text{Prob}(X_t = e_2)$ . Here  $\Delta = 0.25$ .

the Wonham filter at (3.1). It is clear from these plots that not only is the estimation performance very poor, but also it has also produced negative probabilities.

In Figure 6.2, we show the same state and observation process realisation, but in this case, the estimated filter probabilities have been computed using the robust recursion at (4.6).

In our second simulation study, we consider the filter driven by Poisson observations for the particular scenario described by (5.13). For this example, the two Poisson intensities used were

$$\langle \lambda, e_1 \rangle = 8, \quad \langle \lambda, e_2 \rangle = 5. \tag{6.2}$$

and the rate matrix  $A$  was again as above. The plots in Figure 6.3 show the computed probabilities for three schemes: the Euler-Maruyama scheme, the Robust scheme, and the matrix exponential computed by a scaling and squaring algorithm with a Pade approximation [6]. Here the time step was coarse, set at  $\Delta = 0.25$ . The results show that the Euler-Maruyama scheme produced negative probabilities. In contrast, the robust scheme produced positive probabilities and these estimates are in excellent agreement with the exact

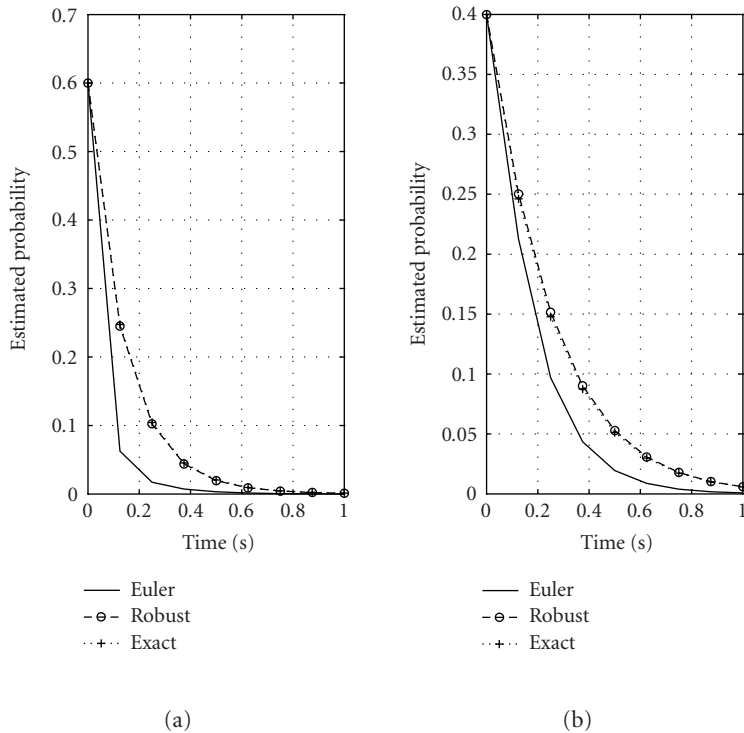


Figure 6.4. Various approximations to the Poisson process filter in between jump events: (a)  $\text{Prob}(X_t = \mathbf{e}_1)$  and (b)  $\text{Prob}(X_t = \mathbf{e}_2)$ . Here  $\Delta = 0.125$ .

scheme. Similar calculations are repeated in Figure 6.4, but with a finer time step, that is,  $\Delta = 0.125$ . In this scenario, the robust recursion again has given far better performance than the Euler-Maruyama scheme.

## References

- [1] J. M. C. Clark, *The design of robust approximations to the stochastic differential equations of nonlinear filtering*, Communication Systems and Random Process Theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977), NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., no. 25, Sijthoff & Noordhoff, Alphen aan den Rijn, 1978, pp. 721–734.
- [2] R. J. Elliott, *Stochastic Calculus and Applications*, Applications of Mathematics, vol. 18, Springer-Verlag, New York, 1982.
- [3] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models*, Applications of Mathematics, vol. 29, Springer-Verlag, New York, 1995.
- [4] R. J. Elliott and W. P. Malcolm, *General smoothing formulae for Markov modulated Poisson process observations*, to appear in IEEE Trans. Automat. Control.
- [5] ———, *Robust M-ary detection filters and smoothers for continuous-time jump Markov systems*, IEEE Trans. Automat. Control **49** (2004), no. 7, 1046–1055.
- [6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins Series in the Mathematical Sciences, vol. 3, Johns Hopkins University Press, Maryland, 1989.

- [7] M. R. James, V. Krishnamurthy, and F. Le Gland, *Time discretization of continuous-time filters and smoothers for HMM parameter estimation*, IEEE Trans. Inform. Theory **42** (1996), no. 2, 593–605.
- [8] O. Kallenberg, *Random Measures*, Akademie-Verlag, Berlin, 1976.
- [9] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics, vol. 23, Springer-Verlag, Berlin, 1992.
- [10] G. Last and A. Brandt, *Marked Point Processes on the Real Line*, Probability and Its Applications, Springer-Verlag, New York, 1995.
- [11] W. P. Malcolm, *Robust filtering and estimation with Poisson observations*, Ph.D. thesis, The Australian National University, Canberra, Australia, 1999.
- [12] W. P. Malcolm and R. J. Elliott, *A general smoothing equation for Poisson observations*, Proc. 38th IEEE Conference on Decision and Control (Arizona), 1999, pp. 4106–4110.
- [13] C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Rev. **20** (1978), no. 4, 801–836.

W. P. Malcolm: Haskayne School of Business, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

*E-mail address:* [malcolmw@ucalgary.ca](mailto:malcolmw@ucalgary.ca)

R. J. Elliott: Haskayne School of Business, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

*E-mail address:* [robert.elliott@haskayne.ucalgary.ca](mailto:robert.elliott@haskayne.ucalgary.ca)

J. van der Hoek: School of Mathematical Sciences, The University of Adelaide, SA 5005, Australia

*E-mail address:* [jvanderh@maths.adelaide.edu.au](mailto:jvanderh@maths.adelaide.edu.au)