Discretizing stochastic dynamical systems using Lyapunov equations

Niklas Wahlström, Patrix Axelsson, Fredrik Gustafsson

Division of Automatic Control, Linköping University, Sweden
(e-mail: nikwa@isy.liu.se, axelsson@isy.liu.se, fredrik@isy.liu.se).

Abstract: Stochastic dynamical systems are fundamental in state estimation, system identification and control. System models are often provided in continuous time, while a major part of the applied theory is developed for discrete-time systems. Discretization of continuous-time models is hence fundamental. We present a novel algorithm using a combination of Lyapunov equations and analytical solutions, enabling efficient implementation in software. The proposed method circumvents numerical problems exhibited by standard algorithms in the literature. Both theoretical and simulation results are provided.

Keywords: Optimal sampling, Lyapunov equations, matrix exponential, white noise

1. INTRODUCTION

Dynamical processes in engineering and physics have for a long time successfully been modeled with continuous-time differential equations. In order to account for uncertainties, these equations are usually driven by an unknown stochastic process called process noise. This noise is ideally modeled as completely "white" in order to obtain the Markov property, which is required in recursive Bayesian inference, such as Kalman filtering. However, in order to implement such filtering, the continuous-time model has to be discretized. Such discretization includes solving an integral involving the matrix exponential on the form

$$Q = \int_0^T e^{A\tau} S e^{A^\mathsf{T}\tau} d\tau, \qquad (1)$$

where $A, S, Q \in \mathbb{R}^{n \times n}$.

We propose an algorithm for solving (1) by decomposing the problem into subproblems and then solve these subproblems either analytically or using a combination of Lyapunov and Sylvester equations.

In many practical applications the discrete-time process noise covariance is modeled and tuned directly, rather than discretized from its continuous-time counterpart. However, in certain scenarios the dependency between the discrete-time process noise covariance and the sampling time is important. If the filtering should work on different devices with different sampling frequencies, this dependency should be properly modeled to guarantee the same dynamical behavior of the filter. Further, in data with non-equidistant sampling the process noise covariance has to be rescaled at each time instant. This is often the case in Gaussian process regression which can be described with a state-space model and solved using Kalman filtering (Särkkä et al., 2013).

In the literature there exist different algorithms for computing the integral (1). The probably most well-cited one was presented by Van Loan (1978), which involves computing the matrix exponential for an augmented $2n \times 2n$ matrix followed by a matrix multiplication of two resulting submatrices. This method does not require any

assumption on the model, however the resulting matrix becomes ill-conditioned if the sampling time is large or if the poles of the system are fast. For certain models, (1) can be solved analytically. Rome (1969) presented a direct solution under the assumption that A is diagonalizable. The method requires an eigenvalue decomposition which is not always numerical stable (Higham, 2008) and not all matrices are diagonalizable. Finally, the integral can always be solved numerically using the trapezoidal or the rectangular method.

In this work we present an alternative method for solving (1). This method is based on a Lyapunov equation which characterizes the solution of (1). However, since Lyapunov equations cannot be solved if the system contains integrators (Antoulas, 2005), the problem is decomposed into subproblems where the integrators are treated separately. As will be explained, one set of subproblems cannot be solved using Lyapunov equations, but they do have an analytical solution of (1). Conversely, the remaining set of subproblems do not have a closed form solution of (1), but then the method with Lyapunov equations works fine. The algorithm involves computing the matrix exponential of the $n \times n$ system matrix rather than an augmented $2n \times 2n$ matrix as required by the solution by Van Loan. Furthermore, the proposed algorithm circumvents some numerical problems in the method proposed by Van Loan. Our theoretical algebraic contributions include:

- A Lemma describing the relation between (1) and the aforementioned Lyapunov equation, see Lemma 2.
- A novel extension of this solution which also handles integrators, see Section 4.
- A complete algorithm which solves (1) with complementing numerical properties compared to existing solutions, see Algorithm 3.

The outline of the paper is as follows. In Section 2 the mathematical models are presented and the importance of the discretization method in use is motivated. In Section 3 the discretization using Lyapunov equations is presented together with the main theoretical contributions of the paper. In Section 4 the solutions from the previous two sections will be combined to solve for systems with integrators. In Section 5 a numerical evaluation is performed and in Section 6 the conclusions are summarized and future directions pointed out.

2. MATHEMATICAL PRELIMINARIES

Consider the following Itô stochastic differential equation

$$d\boldsymbol{x}(t) = A\boldsymbol{x}(t)dt + d\boldsymbol{\beta}(t), \tag{2a}$$

where $\boldsymbol{\beta}(t)$ is a Brownian motion with

$$E[d\boldsymbol{\beta}(t)d\boldsymbol{\beta}(t)^{\mathsf{T}}] = Sdt. \tag{2b}$$

The model (2a) is formally equivalent to the stochastic differential equation

$$\frac{d\boldsymbol{x}(t)}{dt} = A\boldsymbol{x}(t) + \boldsymbol{w}(t), \tag{3a}$$

where $\boldsymbol{w}(t)$ is a zero-mean white Gaussian process with

$$E[\boldsymbol{w}(t)\boldsymbol{w}(\tau)^{\mathsf{T}}] = S\delta(t - \tau). \tag{3b}$$

Since $\boldsymbol{w}(t)$ is not square Riemann integrable, the model (3) does not have any mathematical meaning (Jazwinski, 1970). However, we can still intuitively think of it as a stochastic differential equation driven by white noise.

It is important to note that this is just a model of the physical process and cannot be found in reality. For example, white noise has a flat power spectral density requiring infinite power, which is not physically realizable. Nevertheless, using this continuous-time model will lead to sound properties for the equivalent discrete-time model as will be explained later.

By integrating (2a) over the time interval $[t_k, t_{k+1}]$ we can find its discrete-time equivalence as

$$\boldsymbol{x}(t_{k+1}) = \underbrace{e^{AT_k}}_{F_{T_k}} \underbrace{\boldsymbol{x}(t_k)}_{\boldsymbol{x}_k} + \underbrace{\int_{t_k}^{t_{k+1}} e^{A(\tau - t_{k+1})} d\boldsymbol{\beta}(\tau)}_{\boldsymbol{w}_k}, \tag{4}$$

where $T_k = t_{k+1} - t_k$ is the sampling time. This can be stated as a discrete-time stochastic difference equation

$$\boldsymbol{x}_{k+1} = F_{T_k}\boldsymbol{x}_k + \boldsymbol{w}_k. \tag{5a}$$

By following for example Jazwinski (1970), the noise $\boldsymbol{w}_k$ will be zero-mean, white Gaussian

$$E[\boldsymbol{w}_k\boldsymbol{w}_l^{\mathsf{T}}] = Q_{T_k}\delta_{kl}, \tag{5b}$$

where δ_{kl} is the Kronecker delta function and

$$Q_{T_k} = \int_0^{T_k} e^{A\tau} S e^{A^{\mathsf{T}}\tau} d\tau, \tag{6a}$$

which together with the discrete-time system matrix

$$F_{T_k} = e^{AT_k} \tag{6b}$$

completes the discretization procedure.

The integral expression (6a) can be found in multiple textbooks on Kalman filtering (e.g. Bar-Shalom et al. (2001); Grewal and Andrews (2008)) for modeling discrete-time dynamical processes. Nevertheless, the discretization of continuous-time differential equations for filtering applications is often misused. For example, the noise $\boldsymbol{w}(t)$ is commonly assumed to be constant during each sampling interval leading to the following discrete-time noise covariance

$$\bar{Q}_{T_k}^{\mathrm{A}} = \frac{1}{T_k} \underbrace{\left(\int_0^{T_k} e^{A\tau} d\tau \right)}_{\bar{G}_{T_k}} S \underbrace{\left(\int_0^{T_k} e^{A^{\mathsf{T}}\tau} d\tau \right)}_{\bar{G}_{T_k}^{\mathsf{T}}}, \tag{7a}$$

or just rescaling the continuous-time noise covariance with the sampling time

$$\bar{Q}_{T_k}^{\mathrm{B}} = T_k S. \tag{7b}$$

In contrast to the discretization in (6), the assumptions in (7) lead to a dynamical description of the process which depends on the sampling intervals, whereas the actual physical process do not. This can be seen by the property derived in the following example.

Example 2.1. Consider the three time instances t_1, t_2 and t_3. We then have

$$\mathrm{Cov}\Big[\boldsymbol{x}(t_3)\Big|\boldsymbol{x}(t_1)\Big] \tag{8a}$$

$$= \mathrm{Cov}\Big[F_{t_3-t_2}\boldsymbol{x}(t_2) + \boldsymbol{w}_2\Big|\boldsymbol{x}(t_1)\Big] \tag{8b}$$

$$= \mathrm{Cov}\Big[F_{t_3-t_2}\Big(F_{t_2-t_1}\boldsymbol{x}(t_1) + \boldsymbol{w}_1\Big) + \boldsymbol{w}_2\Big|\boldsymbol{x}(t_1)\Big] \tag{8c}$$

$$= \mathrm{Cov}\Big[F_{T_2}\boldsymbol{w}_1 + \boldsymbol{w}_2\Big|\boldsymbol{x}(t_1)\Big] \tag{8d}$$

$$= F_{T_2}Q_{T_1}F_{T_2}^{\mathsf{T}} + Q_{T_2}. \tag{8e}$$

We could also use only one time interval and go from t_1 directly to t_3 with the sampling time $t_3 - t_1 = T_1 + T_2$, which gives

$$\mathrm{Cov}\Big[\boldsymbol{x}(t_3)\Big|\boldsymbol{x}(t_1)\Big] \tag{9a}$$

$$= \mathrm{Cov}\Big[F_{t_3-t_1}\boldsymbol{x}(t_1) + \boldsymbol{w}_1\Big|\boldsymbol{x}(t_1)\Big] \tag{9b}$$

$$= \mathrm{Cov}\Big[\boldsymbol{w}_1\Big|\boldsymbol{x}(t_1)\Big] = Q_{t_3-t_1} = Q_{T_1+T_2}. \tag{9c}$$

This gives the relation

$$Q_{T_1+T_2} = F_{T_2}Q_{T_1}F_{T_2}^{\mathsf{T}} + Q_{T_2}. \tag{10}$$

Indeed, this property is fulfilled for the discretization presented in (6).

Lemma 1. If F_{T_k} and Q_{T_k} are computed as described in (6), then

$$Q_{T_1+T_2} = F_{T_2}Q_{T_1}F_{T_2}^{\mathsf{T}} + Q_{T_2}.$$

Proof.

$$Q_{T_1+T_2} = \int_0^{T_1+T_2} e^{A\tau} S e^{A^{\mathsf{T}}\tau} d\tau$$

$$= \int_0^{T_2} e^{A\tau} S e^{A^{\mathsf{T}}\tau} d\tau + \int_{T_2}^{T_1+T_2} e^{A\tau} S e^{A^{\mathsf{T}}\tau} d\tau$$

$$= \underbrace{\int_0^{T_2} e^{A\tau} S e^{A^{\mathsf{T}}\tau} d\tau}_{Q_2} + \underbrace{e^{AT_2}}_{F_{T_2}} \underbrace{\int_0^{T_1} e^{A^{\mathsf{T}}\tau} S e^{A^{\mathsf{T}}\tau} d\tau}_{Q_{T_1}} \underbrace{e^{A^{\mathsf{T}}T_2}}_{F_{T_2}^{\mathsf{T}}}$$

$$= Q_{T_2} + F_{T_2}Q_{T_1}F_{T_2}^{\mathsf{T}}.$$

With similar calculations we can easily derive the equivalent results for the covariance matrices in (7) and conclude that they do not share this property since

$$\bar{Q}_{T_1+T_2}^{\mathrm{A}} = \bar{Q}_{T_2}^{\mathrm{A}} + F_{T_2}\bar{Q}_{T_1}^{\mathrm{A}}F_{T_2}^{\mathsf{T}} + F_{T_2}\bar{G}_{T_1}S\bar{G}_{T_2}^{\mathsf{T}} + \bar{G}_{T_2}S\bar{G}_{T_1}^{\mathsf{T}}F_{T_2}^{\mathsf{T}}$$

$$\neq \bar{Q}_{T_2}^{\mathrm{A}} + F_{T_2}\bar{Q}_{T_1}^{\mathrm{A}}F_{T_2}^{\mathsf{T}}, \tag{11a}$$

$$\bar{Q}_{T_1+T_2}^{\mathrm{B}} = \bar{Q}_{T_2}^{\mathrm{B}} + \bar{Q}_{T_1}^{\mathrm{B}}$$

$$\neq \bar{Q}_{T_2}^{\mathrm{B}} + F_{T_2}\bar{Q}_{T_1}^{\mathrm{B}}F_{T_2}^{\mathsf{T}}. \tag{11b}$$

Hence by assuming that the underlying continuous-time model is driven by a continuous-time white process the corresponding discrete-time model has the property that the dynamical description does not depend on the sampling intervals, in contrast to other common discretization procedures. We can therefore see (5) and (6) as algebraic relations between A, S, T_k, F_{T_k} and Q_{T_k} fulfilling the property in (10) without deriving it from its continuous-time counterpart.

The main advantage with the alternative expressions in (7) in comparison to (6a) is their ease of calculation

(especially true for (7b)). The remaining part of this work will therefore describe how the integral (6a) can be solved in an efficient manner with good numerical properties.

3. DISCRETIZATION USING LYAPUNOV EQUATIONS

A method for computing the integral (6a) will now be presented. The method will be proposed by requiring the system to be asymptotically stable. Later in this section we will prove that this requirements actually can be relaxed.

3.1 Proposal of solution

If the system is asymptotically stable, i.e. if all eigenvalues of A have negative real part, a stationary covariance will exist and we denote it as

$$\text{Cov}[\boldsymbol{x}(t)] = \text{Cov}[\boldsymbol{x}_k] = P. \tag{12}$$

This covariance satisfies the following two Lyapunov equations for the continuous-time model (2a) and the discrete-time model (5a), respectively

$$0 = AP + PA^\mathsf{T} + S, \tag{13a}$$

$$P = F_{T_k} P F_{T_k}^\mathsf{T} + Q_{T_k}. \tag{13b}$$

which gives a structured way of computing Q_{T_k}, as presented in Algorithm 1.

Algorithm 1 Solution using Lyapunov equation for P

The matrices A and S and the scalar T_k are given. The matrices F_{T_k} and Q_{T_k} in (6) can then be computed as

$$F_{T_k} = e^{AT_k}, \tag{14a}$$

$$Q_{T_k} = P - F_{T_k} P F_{T_k}^\mathsf{T}, \tag{14b}$$

where P is the solution to the Lyapunov equation

$$AP + PA^\mathsf{T} = -S. \tag{14c}$$

This algorithm can also be reformulated such that we do not need to compute P in an intermediate step. By multiplying (14b) with A from the left and with A^T from the right, respectively, we get

$$AQ_{T_k} = AP - F_{T_k} APF_{T_k}^\mathsf{T}, \tag{15a}$$

$$Q_{T_k} A^\mathsf{T} = PA^\mathsf{T} - F_{T_k} PA^\mathsf{T} F_{T_k}^\mathsf{T}, \tag{15b}$$

where the fact that F_{T_k} and A commute has been used since

$$AF_{T_k} = A(I + A + \frac{1}{2}A^2 \dots) = (A + A^2 + \frac{1}{2}A^3 \dots)$$

$$= (I + A + \frac{1}{2}A^2 \dots)A = F_{T_k}A.$$

By adding (15a) and (15b), we get

$$AQ_{T_k} + Q_{T_k}A^\mathsf{T} = AP - F_{T_k} APF_{T_k}^\mathsf{T} + PA^\mathsf{T} - F_{T_k} PA^\mathsf{T} F_{T_k}^\mathsf{T}$$

$$= \underbrace{AP + PA^\mathsf{T}}_{-S} - F_{T_k}(\underbrace{AP + PA^\mathsf{T}}_{-S})F_{T_k}^\mathsf{T}$$

$$= -S + F_{T_k} S F_{T_k}^\mathsf{T}. \tag{16}$$

This gives the following algorithm as presented in Algorithm 2. This algorithm is similar to the solution presented by Axelsson and Gustafsson (2012) derived from a continuous-time differential Lyapunov equation.

From here on we will proceed with Algorithm 2. However, all results (including the final algorithm) can be reformulated to suit Algorithm 1 as well.

Algorithm 2 Solution using Lyapunov equation for Q_{T_k}

The matrices A and S and the scalar T_k are given. The matrices F_{T_k} and Q_{T_k} in (6) can then be computed as

$$F_{T_k} = e^{AT_k} \tag{17a}$$

and Q_{T_k} is the solution to the Lyapunov equation

$$AQ_{T_k} + Q_{T_k} A^\mathsf{T} = -V_{T_k}, \tag{17b}$$

where

$$V_{T_k} = S - F_{T_k} S F_{T_k}^\mathsf{T}. \tag{17c}$$

3.2 Theoretical result

It can now be proven that Algorithm 2 (and consequently also Algorithm 1) gives a solution to (6), provided that the solution of the Lyapunov equation exists and is unique.

Lemma 2. The solution to the integral

$$Q = \int_0^T e^{A\tau} S e^{B\tau} d\tau \tag{18a}$$

satisfies the Sylvester equation

$$AQ + QB = -S + e^{AT} S e^{BT}. \tag{18b}$$

Proof. Start with (18b) and replace Q with the integral (18a). This gives

$$AQ + QB = \int_0^T Ae^{A\tau} S e^{B\tau} d\tau + \int_0^T e^{A\tau} S e^{B\tau} B d\tau$$

$$= \int_0^T \frac{d}{d\tau}[e^{A\tau} S e^{B\tau}] d\tau \tag{19a}$$

$$= e^{A\tau} S e^{B\tau}\Big|_0^T = e^{AT} S e^{BT} - S. \tag{19b}$$

Remark 3. A similar result was presented by Gawronski (2004) in the context of time-limited grammians. However, that result requires $B = A^\mathsf{T}$ and that all eigenvalues of A should have negative real part.

Remark 4. Note that Lemma 2 does not require anything about the matrices A and B. In particular, they do not need to be stable as assumed in (12) and (13). Indeed, the requirements for the Lyapunov equation (17b) to have a unique solution are milder. This is answered by the following proposition, which is given for the more general Sylvester equation.

Proposition 5. The Sylvester equation

$$AP + PB = C \tag{20}$$

has a unique solution P if and only if

$$\lambda_i(A) + \lambda_j(B) \neq 0 \quad \forall i, j. \tag{21}$$

For proof, see for example Antoulas (2005).

For the case where $B = A^\mathsf{T}$ and with the requirement that A is stable, the condition (21) is always fulfilled. By using that observation together with Lemma 2 where $T \to \infty$, we get the following well known results relating the controllability grammian to a Lyapunov equation, which can be found in most textbooks on linear systems, e.g. Rugh (1996).

Corollary 6. If all eigenvalues of A have negative real parts, then for each symmetric matrix S there exists a unique solution of

$$AQ + QA^\mathsf{T} = -S \tag{22a}$$

given by

$$Q = \int_0^\infty e^{A\tau} S e^{A^\mathsf{T}\tau} d\tau. \tag{22b}$$

According to Proposition 5 the integral (6a) cannot be computed using the Lyapunov equation (17b) if A and $-A$ have any common eigenvalues. This is especially the case if the system has integrators, which indeed is common in models intended for Kalman filtering. In the next section we will therefore present a solution which handles such systems as well. With this extension almost all systems of interest will be covered, except for the systems which have at least one pair of non-zero poles mirrored in the imaginary axis.

This extension will be performed by decomposing the problem into subproblems where some of these subproblems still can be solved using parts of the Lyapunov equation (17b), whereas the remaining subproblem can be solved analytically using the integral (6a).

4. SOLUTION FOR SYSTEMS WITH INTEGRATORS

Consider the case when A is block triangular consisting of three blocks

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \tag{23}$$

where

$$\lambda_i(A_{11}) + \lambda_j(A_{11}) \neq 0 \quad \forall i,j, \tag{24a}$$
$$\lambda_i(A_{11}) + \lambda_j(A_{22}) \neq 0 \quad \forall i,j, \tag{24b}$$
$$\lambda_j(A_{22}) = 0 \quad \forall i,j. \tag{24c}$$

In this manner we have partitioned A such that all zero eigenvalues have been placed in A_{22} and all remaining non-zero eigenvalues in A_{11}. Many systems do have such block triangular structure, for example if an observer canonical form has been used, see Example 4.2. If the system does not have that form, an orthogonal transformation can be applied. This transformation can be computed using Schur decomposition and reordering of the eigenvalues (Golub and Van Loan, 1996). This will also affect the covariance matrix S as well as V_{T_k} by considering this transformation as a state transformation, see Appendix A.

4.1 Solution using Lyapunov and Sylvester equations

According to Lemma 2, the solution of the integral (6a) for the block triangular matrix (23) shall obey the following Lyapunov equation

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^\mathsf{T} & Q_{22} \end{bmatrix} + \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^\mathsf{T} & Q_{22} \end{bmatrix}\begin{bmatrix} A_{11}^\mathsf{T} & 0 \\ A_{12}^\mathsf{T} & A_{22}^\mathsf{T} \end{bmatrix} = -\begin{bmatrix} V_{11} & V_{12} \\ V_{12}^\mathsf{T} & V_{22} \end{bmatrix},$$

where Q_{T_k} and V_{T_k} have been partitioned in a similar manner as A. Note that the subscript T_k has been omitted from the submatrices in order to make the notation less cluttered. This gives the following set of Lyapunov and Sylvester equations

$$A_{11}Q_{11} + Q_{11}A_{11}^\mathsf{T} = -V_{11} - A_{12}Q_{12}^\mathsf{T} - Q_{12}A_{12}^\mathsf{T}, \tag{25a}$$
$$A_{11}Q_{12} + Q_{12}A_{22}^\mathsf{T} = -V_{12} - A_{12}Q_{22}, \tag{25b}$$
$$A_{22}Q_{12}^\mathsf{T} + Q_{12}^\mathsf{T}A_{11}^\mathsf{T} = -V_{12}^\mathsf{T} - Q_{22}A_{12}^\mathsf{T}, \tag{25c}$$
$$A_{22}Q_{22} + Q_{22}A_{22}^\mathsf{T} = -V_{22}. \tag{25d}$$

Based on the requirements in (24a) and (24b), Proposition 5 guarantees that Q_{11} and Q_{12} can be solved uniquely using (25a) and (25b) if Q_{22} is known. In contrast, (25d) does not have a unique solution for Q_{22}. Instead, Q_{22} can be solved analytically using the integral (6a). Note that (25c) is just a transposed version of (25b) and does not bring any extra information.

4.2 Analytical solution for the nilpotent part

Due to the block triangular structure of A, the submatrix Q_{22} will only depend on A_{22} and S_{22} via a similar expression as in (6a). By starting from (6a) we have

$$Q_{22} = \begin{bmatrix} 0 & I \end{bmatrix} Q \begin{bmatrix} 0 \\ I \end{bmatrix} \tag{26a}$$

$$= \begin{bmatrix} 0 & I \end{bmatrix} \int_0^{T_k} e^{A\tau} S e^{A^\mathsf{T}\tau} d\tau \begin{bmatrix} 0 \\ I \end{bmatrix} \tag{26b}$$

$$= \int_0^{T_k} \begin{bmatrix} 0 & e^{A_{22}\tau} \end{bmatrix} S \begin{bmatrix} 0 \\ e^{A_{22}^\mathsf{T}\tau} \end{bmatrix} d\tau \tag{26c}$$

$$= \int_0^{T_k} e^{A_{22}\tau} S_{22} e^{A_{22}^\mathsf{T}\tau} d\tau. \tag{26d}$$

Further, since all eigenvalues of A_{22} are zero, the submatrix A_{22} will also be nilpotent (Lancaster and Tismenetsky, 1985) leading to

$$e^{A_{22}\tau} = \sum_{i=0}^{p-1} A_{22}^i \frac{\tau^i}{i!}, \tag{27}$$

where p is the dimension of A_{22}, i.e. the number of integrators in the system. Expression (26d) can then be computed analytically as

$$Q_{22} = \int_0^{T_k} \left(\sum_{i=0}^{p-1} \frac{1}{i!} A_{22}^i \tau^i \right) S_{22} \left(\sum_{j=0}^{p-1} \frac{1}{j!} {A_{22}^j}^\mathsf{T} \tau^j \right) d\tau$$

$$= \sum_{i=0}^{p-1}\sum_{j=0}^{p-1} \frac{1}{i!j!} A_{22}^i S_{22} {A_{22}^j}^\mathsf{T} \int_0^{T_k} \tau^{i+j} d\tau \tag{28a}$$

$$= \sum_{i=0}^{p-1}\sum_{j=0}^{p-1} \frac{T_k^{i+j+1}}{i!j!(i+j+1)} A_{22}^i S_{22} {A_{22}^j}^\mathsf{T}. \tag{28b}$$

This is illustrated with the following example.

Example 4.1. Consider a constant velocity model, which formally can be described on the form

$$\dot{\boldsymbol{x}}(t) = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{A} \boldsymbol{x}(t) + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B} q(t), \quad E[q(t)q(\tau)] = \delta(t-\tau).$$

This system has only zero eigenvalues which gives

$$A = A_{22} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \tag{29a}$$

$$S = S_{22} = E[Bq(Bq)^\mathsf{T}] = \begin{bmatrix} 0 \\ 1 \end{bmatrix} 1 \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{29b}$$

By using this in (26) we get

$$Q_{T_k} = ST_k + SA_{22}^\mathsf{T}\frac{T_k^2}{2} + A_{22}S\frac{T_k^2}{2} + A_{22}SA_{22}^\mathsf{T}\frac{T_k^3}{3}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} T_k + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\frac{T_k^2}{2} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\frac{T_k^2}{2} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\frac{T_k^3}{3}$$

$$= \begin{bmatrix} \dfrac{T_k^3}{3} & \dfrac{T_k^2}{2} \\ \dfrac{T_k^2}{2} & T_k \end{bmatrix}, \tag{29c}$$

which is the same result as given by Grewal and Andrews (2008), but derived in a different way.

4.3 General algorithm

Based on the results in the last section, we can now propose an algorithm for computing the integral (6a), also in the case where A consists of integrators, see Algorithm 3.

Algorithm 3 Proposed algorithm (for systems with arbitrary number of integrators)

The matrices A and S and the scalar T_k are given. The matrices F_{T_k} and Q_{T_k} in (6) can then be computed as

(1) Transform A and S to $\tilde{A}$ and $\tilde{S}$ such that $\tilde{A}$ becomes block triangular

$$U^{-1}AU = \tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \quad U^{-1}SU^{-\mathsf{T}} = \tilde{S} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{12}^{\mathsf{T}} & \tilde{S}_{22} \end{bmatrix},$$

and with all integrators collected in $\tilde{A}_{22}$. This can be done with an orthogonal transformation computed using Schur decomposition and reordering of the eigenvalues.

(2) Compute $\tilde{F}_{T_k} = e^{\tilde{A}T_k}$.

(3) Compute $\tilde{V}_{T_k} = \tilde{S} - \tilde{F}_{T_k}\tilde{S}\tilde{F}_{T_k}^{\mathsf{T}}$.

(4) Compute

$$\tilde{Q}_{T_k} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathsf{T}} & \tilde{Q}_{22} \end{bmatrix}$$

using the following steps:

(a) Compute $\tilde{Q}_{22}$ by evaluating

$$\tilde{Q}_{22} = \sum_{i=0}^{p-1}\sum_{j=0}^{p-1} \frac{T_k^{i+j+1}}{i!j!(i+j+1)} \tilde{A}_{22}^i \tilde{S}_{22}(\tilde{A}_{22}^i)^{\mathsf{T}},$$

where p is the number of integrators.

(b) Compute $\tilde{Q}_{12}$ by solving the Sylvester equation
$$\tilde{A}_{11}\tilde{Q}_{12} + \tilde{Q}_{12}\tilde{A}_{22}^{\mathsf{T}} = -\tilde{V}_{12} - \tilde{A}_{12}\tilde{Q}_{22}. \quad (30)$$

(c) Compute $\tilde{Q}_{11}$ by solving the Lyapunov equation
$$\tilde{A}_{11}\tilde{Q}_{11} + \tilde{Q}_{11}\tilde{A}_{11}^{\mathsf{T}} = -\tilde{V}_{11} - \tilde{A}_{12}\tilde{Q}_{12}^{\mathsf{T}} - \tilde{Q}_{12}\tilde{A}_{12}^{\mathsf{T}}. \quad (31)$$

(5) Transform $\tilde{F}_{T_k}$ and $\tilde{Q}_{T_k}$ back to F_{T_k} and Q_{T_k}
$$F_{T_k} = U\tilde{F}_{T_k}U^{-1}, \quad (32a)$$
$$Q_{T_k} = U\tilde{Q}_{T_k}U^{\mathsf{T}}. \quad (32b)$$

Remark 7. If A does not have any integrators, Algorithm 3 will collapse to the simpler version in Algorithm 2.

Remark 8. In theory, $U^{-1} = U^{\mathsf{T}}$ since U is orthogonal. However, numerical algorithms for computing the Schur decomposition do not make U completely orthogonal. From a numerical point of view it is therefor a benefit to distinguish between U^{-1} and U^{T}.

Remark 9. If $\tilde{A}_{12} = 0$ the coupling in (30) and (31) via $\tilde{Q}_{12}$ and $\tilde{Q}_{22}$ will disappear and they can be solved independently from each other. If this is desired, the transformation in Step 1 can be extended to eliminate $\tilde{A}_{12}$ by solving an addition Sylvester equation (Bavely and Stewart, 1979). However, such transformation is no longer orthogonal and can be arbitrary ill-conditioned if the nonzero eigenvalues are close to zero.

Remark 10. If the system already has a block triangular structure, Step 1 and Step 5 in Algorithm 3 can be omitted. This is the case for the observer canonical form as seen in the following short example.

Example 4.2. Consider a SISO-system of order $n = m + p$ with m non-zero poles and p additional integrators described with a transfer function

$$G(s) = \frac{b_1 s^{m-1} + b_2 s^{m-2} + \cdots + b_{m-1}s + b_m}{s^n + a_1 s^{m-1} + \cdots + a_{m-1}s + a_m} \cdot \frac{1}{s^p}. \quad (33)$$

This system can be described with the observer canonical form (Glad and Ljung, 2000)

$$\dot{x} = \left[\begin{array}{ccc|cccc} -a_1 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ -a_{m-1} & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ -a_m & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{array}\right] x + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix} w, \quad (34a)$$

$$y = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} x, \quad (34b)$$

which can be written more compactly as

$$\dot{x} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] x + Bw, \quad (35a)$$

$$y = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} x. \quad (35b)$$

This system has by construction the desired block triangular structure.

5. NUMERICAL EVALUATION

In this section the numerical properties of the proposed solution will be compared with a standard solution presented by Van Loan (1978) given in Algorithm 4.

Algorithm 4 Van Loan's method

The matrices A and S and the scalar T_k are given. The matrices F_{T_k} and Q_{T_k} in (6) can then be computed as

(1) Compute the matrix exponential of an augmented $2n \times 2n$ matrix

$$\begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix} = e^{HT_k}, \text{ where } H = \begin{bmatrix} A & S \\ 0 & -A^{\mathsf{T}} \end{bmatrix}. \quad (36a)$$

(2) The matrices F_{T_k} and Q_{T_k} are given as

$$F_{T_k} = M_{11}, \qquad Q_{T_k} = M_{12}M_{11}^{\mathsf{T}}. \quad (36b)$$

5.1 Implementation aspects

In both methods MATLAB's built-in function `expm` has been used for computing the matrix exponential. In Step 1 of Algorithm 3 the functions `schur` and `ordschur` have been used for computing the Schur decomposition and the reordering of the eigenvalues. Finally, the Lyapunov and Sylvester equations have been solved using `lyap`.

5.2 Simulation results

In total 100 systems of order $n = 6$ with $m = 4$ stable poles and $p = 2$ additional integrators are randomly generated. Each system is normalized such that the fastest pole is at distance 1 from the imaginary axis, i.e. $\max(|\mathrm{Re}(\lambda_i)|) = 1$. An estimate $\hat{Q}_{T_k}$ is computed using both Algorithm 3 and Algorithm 4 with single precision for different values of the sampling time T_k. Finally, the error

$$\varepsilon = \|\hat{Q}_{T_k} - Q_{T_k}\|_2 / \|Q_{T_k}\|_2$$

is evaluated, where Q_{T_k} is computed using numerical integration of (6a) with double precision, here considered as the true value. The result is presented in Figure 1.
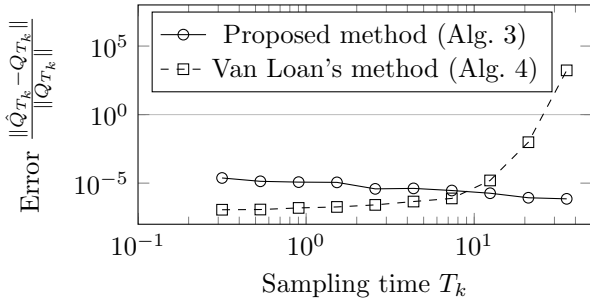
Fig. 1. The performance of the proposed method (Algorithm 3) and Van Loan's method (Algorithm 4).

According to the result the proposed method outperforms the standard method for large T_k. The reason will become clear if we investigate the two methods further. In Algorithm 4, both AT_k and $-A^\mathsf{T}T_k$ are present in the augmented matrix HT_k and the task to compute its matrix exponential (36a) will become ill-conditioned if T_k or $\max(|\mathrm{Re}(\lambda_i)|)$ is large. In fact, the error will grow exponentially with T_k, or the magnitude of work will grow linearly with T_k to keep a certain tolerance (Van Loan, 1978). This issue is not present in the proposed method, which can be seen in its simplified version in Algorithm 1. If T_k is large we have $F_{T_k} = e^{AT_k} \approx 0$ and Q_{T_k} will approach the stationary covariance P according to (14b).

However, for fast sampling the proposed method performers slightly worse. This is especially the case if the system has integrators as well as non-zero poles close to the origin leading to that the Sylvester equation (30) will become ill-conditioned. Consequently, van Loan's method preforms worse when the fastest pole is fast and the sampling is slow, whereas the proposed method preforms worse when the slowest (non-zero) pole is slow and the sampling is fast. The proposed method also has computational complexity advantages since it only needs to compute the matrix exponential of an $n \times n$ matrix rather than of an augmented $2n \times 2n$ matrix as required by van Loan's method.

6. CONCLUSIONS AND FUTURE WORK

An algorithm for computing an integral involving the matrix exponential common in optimal sampling was proposed. The algorithm is based on a Lyapunov equation and is justified with a novel lemma. An extension to systems with integrators was presented. Numerical evaluations showed that the proposed algorithm has advantageous numerical properties slow sampling and fast dynamics in comparison with a standard method in the literature.

Further work includes extending the algorithm further to handle arbitrary matrices, i.e. also matrices with non-zero eigenvalues mirrored in the imaginary axis. Also the numerical properties should be analyzed further and strategies for improving the numerical properties for slow poles should be investigated.

REFERENCES

Antoulas, A.C. (2005). *Approximation of large-scale dynamical systems*. SIAM, Philadelphia, PA, USA.
Axelsson, P. and Gustafsson, F. (2012). Discrete-time solutions to the continuous-time differential Lyapunov equation with applications to Kalman filtering. Technical report, Linköping University, Sweden.
Bar-Shalom, Y., Li, X.R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation:*
Theory Algorithms and Software. John Wiley & Sons, New York, NY, USA.
Bavely, C.A. and Stewart, G.W. (1979). An algorithm for computing reducing subspaces by block diagonalization. *SIAM Journal on Numerical Analysis*, 16(2), 359–367.
Gawronski, W. (2004). *Advanced structural dynamics and active control of structures.* Springer-Verlag, Berlin, Heidelberg, Germany.
Glad, T. and Ljung, L. (2000). *Control theory.* Taylor Francis, New York, NY, USA.
Golub, G.H. and Van Loan, C.F. (1996). *Matrix computations*, volume 3. The Johns Hopkins University Press, Baltimore, MD, USA.
Grewal, M.S. and Andrews, A.P. (2008). *Kalman Filtering. Theory and Practice Using* MATLAB. John Wiley & Sons, Hoboken, NJ, USA, third edition.
Higham, N.J. (2008). *Functions of Matrices – Theory and Computation.* SIAM, Philadelphia, PA, USA.
Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*, volume 64 of *Mathematics in Science and Engineering.* Academic Press, New York, NY, USA.
Lancaster, P. and Tismenetsky, M. (1985). *The theory of matrices*, volume 2. Academic Press, New York, NY, USA.
Rome, H.J. (1969). A direct solution to the linear variance equation of a time-invariant linear system. *IEEE Transactions on Automatic Control*, 14(5), 592–593.
Rugh, W.J. (1996). *Linear System Theory.* Information and System Sciences Series. Prentice Hall Inc., Upper Saddle River, NJ, USA, second edition.
Särkkä, S., Solin, A., and Hartikainen, J. (2013). Spatio-temporal learning via infinite-dimensional Bayesian filtering and smoothing. *IEEE Signal Processing Magazine*, 30(4), 51–61.
Van Loan, C.F. (1978). Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, 23(3), 395–404.

Appendix A. STATE TRANSFORMATION

Consider the following state transformation
$$\boldsymbol{x} = U\tilde{\boldsymbol{x}}. \tag{A.1}$$
By applying (A.1) to the dynamical equation (3a) we get
$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{w} \quad \Rightarrow \tag{A.2a}$$
$$U\dot{\tilde{\boldsymbol{x}}} = AU\tilde{\boldsymbol{x}} + \boldsymbol{w} \quad \Rightarrow \tag{A.2b}$$
$$\dot{\tilde{\boldsymbol{x}}} = U^{-1}AU\tilde{\boldsymbol{x}} + U^{-1}\boldsymbol{w} \quad \Rightarrow \tag{A.2c}$$
$$\dot{\tilde{\boldsymbol{x}}} = \tilde{A}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{w}}. \tag{A.2d}$$
which gives the following transformation of A, S and V
$$\tilde{A} = U^{-1}AU, \tag{A.3a}$$
$$\tilde{S} = E[\tilde{\boldsymbol{w}}\tilde{\boldsymbol{w}}^\mathsf{T}] = E[U^{-1}\boldsymbol{w}(U^{-1}\boldsymbol{w})^\mathsf{T}] = U^{-1}E[\boldsymbol{w}\boldsymbol{w}^\mathsf{T}]U^{-\mathsf{T}}$$
$$= U^{-1}SU^{-\mathsf{T}}. \tag{A.3b}$$
These matrices will then be used to compute $\tilde{F}_{T_k}$ and $\tilde{Q}_{T_k}$ by following Step 2-4 in Algorithm 3. We then have
$$\tilde{\boldsymbol{x}}_{k+1} = \tilde{F}_{T_k}\tilde{\boldsymbol{x}}_k + \tilde{\boldsymbol{w}}_k \quad \Rightarrow \tag{A.4a}$$
$$U^{-1}\boldsymbol{x}_{k+1} = \tilde{F}_{T_k}U^{-1}\boldsymbol{x}_k + \tilde{\boldsymbol{w}}_k \quad \Rightarrow \tag{A.4b}$$
$$\boldsymbol{x}_{k+1} = U\tilde{F}_{T_k}U^{-1}\boldsymbol{x}_k + U\tilde{\boldsymbol{w}}_k \quad \Rightarrow \tag{A.4c}$$
$$\boldsymbol{x}_{k+1} = F_{T_k}\boldsymbol{x}_k + \boldsymbol{w}_k \tag{A.4d}$$
which gives the transformations
$$F_{T_k} = U\tilde{F}_{T_k}U^{-1}, \tag{A.5a}$$
$$Q_{T_k} = E[\boldsymbol{w}_k\boldsymbol{w}_k^\mathsf{T}] = E[U\tilde{\boldsymbol{w}}_k(U\tilde{\boldsymbol{w}}_k)^\mathsf{T}] = UE[\tilde{\boldsymbol{w}}_k\tilde{\boldsymbol{w}}_k^\mathsf{T}]U^\mathsf{T}$$
$$= U\tilde{Q}_{T_k}U^\mathsf{T}. \tag{A.5b}$$
Note that if U is orthogonal, we have $U^{-1} = U^\mathsf{T}$.