Computing the stochastic H^∞-norm by a Newton iteration

Tobias Damm[1] and Peter Benner[2] and Jan Hauth[3]

Abstract— The stochastic H^∞-norm is defined as the L^2-induced norm of the input-output operator of a stochastic linear system. Like the deterministic H^∞-norm it is characterised by a version of the bounded real lemma, but without a frequency domain description or a Hamiltonian condition. Therefore, we base its computation on a parametrised algebraic Riccati-type matrix equation and a Newton iteration. For large dimensions, our algorithm outperforms LMI-methods.

I. INTRODUCTION

The H^∞-norm is a fundamental concept for asymptotically stable deterministic linear time invariant systems. It is equal to the input/output norm of a system both in the frequency and the time domain. It is used in robustness analysis and serves as a performance index in H^∞ control. In model order reduction, it is an important measure for the quality of the approximation. There are very efficient algorithms for the computation of the H^∞-norm, which are based on a Hamiltonian characterization. The most widely used among these was described in [1], [2], but recent progress has also been made e.g. in [3], [4], [5], [6].

A stochastic version of the H^∞-norm was introduced by Hinrichsen and Pritchard in [7]. It has a similar range of applications as its deterministic counterpart, but its numerical computation has hardly been considered in the literature. A major obstacle in transferring ideas and algorithms from the deterministic case is the lack of a suitable frequency domain interpretation or a Hamiltonian characterization in the stochastic setup.

But according to the stochastic bounded real lemma, [7], the stochastic H^∞-norm can be characterized as the minimum (or infimum) value of a minimization problem with constraints in terms of linear matrix inequalities (LMIs). Thus its computation lies in the scope of general purpose LMI-solvers based on semidefinite programming techniques. To our knowledge, this is the only approach readily available so far. However, these methods turn out to be quite expensive for large dimensions.

As an alternative, we present an algorithm based on a Riccati characterization. Again by the stochastic bounded real lemma, the norm is the infimum of all $\gamma > 0$ for which a given parametrized Riccati equation has a stabilizing

solution. We check the solvability of the Riccati equation by a Newton iteration. To reduce the complexity, the linear equations in each Newton step are solved by a preconditioned Krylov subspace iteration. Numerical experiments confirm that the asymptotic complexity of our algorithm is much lower than that of LMI-solvers.

Our investigations were motivated by our recent work [8], [9] on model order reduction for stochastic systems. To quantify the approximation error, the stochastic H^∞-norm needs to be computed for large systems.

The paper is structured as follows. In Section 2 we introduce stochastic systems, define the stochastic H^∞-norm and state the stochastic bounded real lemma. We also provide a new version of the non-strict bounded real lemma and give some new bounds for the stabilizing solution, which are proven in Appendix B. In Section 3 we describe our basic algorithm and discuss ways to make all the steps fast. In Section 4 we report on numerical experiments. In particular, we compare our algorithm with LMI-solvers. To keep the notational burden low, we confine ourselves to the case, where only one multiplicative noise term affects the state vector. Our results can easily be extended to more general situations which we hint at in Appendix A.

II. THE STOCHASTIC H^∞-NORM

We consider stochastic linear systems of the form

$$dx = (Ax + Bu)\, dt + Nx\, dw\,, \quad y = Cx + Du\,, \quad (1)$$

where $A, N \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, $D \in \mathbb{R}^{p \times m}$, and $w = (w(t))_{t \in \mathbb{R}_+}$ is a zero mean real Wiener process on a probability space $(\Omega, \mathcal{F}, \mu)$ with respect to an increasing family $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ (e.g. [10], [11]). Let $L^2_w(\mathbb{R}_+, \mathbb{R}^q)$ denote the corresponding space of non-anticipating stochastic processes v with values in $\mathbb{R}^q$ and norm

$$\|v(\cdot)\|^2_{L^2_w} := \mathcal{E}\left(\int_0^\infty \|v(t)\|^2 dt\right) < \infty,$$

where $\mathcal{E}$ denotes expectation. For initial data $x(0) = x_0$ and input $u \in L^2_w(\mathbb{R}_+, \mathbb{R}^m)$ we denote the solution and the output of (1) by $x(t, x_0, u)$ and $y(t, x_0, u)$, respectively.

Definition 1: System (1) is called *asymptotically mean-square-stable*, if

$$\mathcal{E}(\|x(t, x_0, 0)\|^2) \stackrel{t \to \infty}{\longrightarrow} 0\,,$$

for all initial conditions x_0. In this case, for simplicity, we also call the pair (A, N) asymptotically mean-square stable. If (A, N) is asymptotically mean-square stable, then (1) defines an input-output operator $\mathbb{L} : u \mapsto y$ from $L^2_w(\mathbb{R}_+, \mathbb{R}^m)$

[1]Tobias Damm is with the Department of Mathematics, University of Kaiserslautern, 67653 Kaiserslautern, Germany damm@mathematik.uni-kl.de

[2]Peter Benner is with the Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany benner@mpi-magdeburg.mpg.de

[3]Jan Hauth is with the Fraunhofer Institute for Industrial Mathematics (ITWM), 67663 Kaiserslautern, Germany jan.hauth@itwm.fraunhofer.de

to $L^2_w(\mathbb{R}_+, \mathbb{R}^p)$ via $u \mapsto y(\cdot, 0, u)$, see [7]. By $\|\mathbb{L}\|$ we denote the induced operator norm,

$$\|\mathbb{L}\| = \sup_{\|u\|_{L^2_w} = 1} \|y(\cdot, 0, u)\|_{L^2_w} , \qquad (2)$$

which is an analogue of the deterministic H^∞-norm. Following [7], we call it the *stochastic H^∞-norm* of system (1).

A. The stochastic bounded real lemma

The norm (2) can be characterized by the stochastic bounded real lemma. To this end, we define the quadratic (Riccati-type) mapping $\mathcal{R}_\gamma : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, which depends on the parameter $\gamma > \|D\|_2$, by

$$\mathcal{R}_\gamma(X) = A^T X + XA + N^T XN - C^T C$$
$$- (B^T X - D^T C)^T (\gamma^2 I - D^T D)^{-1}(B^T X - D^T C) .$$

Its Fréchet derivative at some $X \in \mathbb{R}^{n \times n}$ is the linear mapping $(\mathcal{R}_\gamma)'_X : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ given by

$$(\mathcal{R}_\gamma)'_X(\Delta) = A_X^T \Delta + \Delta A_X + N^T \Delta N , \qquad (3)$$

where $A_X = A - B(\gamma^2 I - D^T D)^{-1}(B^T X - D^T C)$. Writing $\mathcal{L}_A : X \mapsto A^T X + XA$ and $\Pi_N : X \mapsto N^T XN$, we have

$$(\mathcal{R}_\gamma)'_X(\Delta) = \mathcal{L}_{A_X}(\Delta) + \Pi_N(\Delta) . \qquad (4)$$

The pair (A, N) is asymptotically mean-square stable if and only if $\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_- = \{\lambda \in \mathbb{C} \mid \Re\lambda < 0\}$, cf. [12].

Theorem 2: [7] Let $\sigma(\mathcal{L}_A + \Pi_N) \subset \mathbb{C}_-$. For $\gamma > \|D\|_2$, the following are equivalent.

(i) $\|\mathbb{L}\| < \gamma$.

(ii) *There exists a negative definite solution $X < 0$ to the linear matrix inequality*

$$\begin{bmatrix} (\mathcal{L}_A + \Pi_N)(X) - C^T C & XB - C^T D \\ B^T X - D^T C & \gamma^2 I - D^T D \end{bmatrix} > 0. \quad (5)$$

(iii) *There exists a negative definite solution $X < 0$ to the strict Riccati inequality $\mathcal{R}_\gamma(X) > 0$.*

(iv) *There exists a solution $X \leq 0$ to the Riccati equation $\mathcal{R}_\gamma(X) = 0$, such that $\sigma((\mathcal{R}_\gamma)'_X) \subset \mathbb{C}_-$.*

Remark 3: A solution of the Riccati equation $\mathcal{R}_\gamma(X) = 0$ with $\sigma((\mathcal{R}_\gamma)'_X) \subset \mathbb{C}_-$ is called a *stabilizing solution*. If it exists, then it is the uniquely defined largest solution of the inequality $\mathcal{R}_\gamma(X) \geq 0$, see [12]. We will write $X_+(\gamma)$ for this solution. By Theorem 2, the norm $\|L\|$ is the infimum of all γ such that $\mathcal{R}_\gamma(X) = 0$ possesses a stabilizing solution,

$$\|\mathbb{L}\| = \inf \left\{ \gamma > \|D\|_2 \mid \exists X < 0 : \begin{array}{l} \mathcal{R}_\gamma(X) = 0 \text{ and} \\ \sigma((\mathcal{R}_\gamma)'_X) \subset \mathbb{C}_- \end{array} \right\} . \tag{6}$$

Under a controllability assumption we can also give a nonstrict version of Theorem 2 for asymptotically mean-square stable systems. This is useful to understand the limit case when γ approaches $\|\mathbb{L}\|$. We define the controllability Gramian P of system (1) as the solution of

$$AP + PA^T + NPN^T = -BB^T . \qquad (7)$$

If the system is stable, then P is nonnegative definite, $P \geq 0$.

Corollary 4: Assume that (A, N) is asymptotically mean-square stable and $P > 0$ in (7). For $\gamma > \|D\|_2$, the following are equivalent.

(i) $\|\mathbb{L}\| \leq \gamma$.

(ii) *There exists a solution $X \leq 0$ to the linear matrix inequality*

$$\begin{bmatrix} (\mathcal{L}_A + \Pi_N)(X) - C^T C & XB - C^T D \\ B^T X - D^T C & \gamma^2 I - D^T D \end{bmatrix} \geq 0. \quad (8)$$

(iii) *There exists a solution $X \leq 0$ to the Riccati equation $\mathcal{R}_\gamma(X) = 0$.*

Moreover, if $\|\mathbb{L}\| = \gamma$, then $\mathcal{R}_\gamma(X) = 0$ has a largest solution $X = X_+(\gamma)$, for which $0 \in \sigma((\mathcal{R}_\gamma)'_X) \subset \mathbb{C}_- \cup i\mathbb{R}$.

This result is slightly stronger than [13, Proposition 9.6] or [12, Corollary 5.3.14], where it was shown that (i) implies (iii) if (A, B) is controllable. In the appendix we give a new simplified proof, which can also be modified to obtain lower bounds for solutions of (8) as follows.

B. Inequalities for solutions of the Riccati equation

Lemma 5: Assume that (A, N) is asymptotically mean-square stable, and $\gamma > \|D\|_2$. Let $P^\dagger \geq 0$ be the Moore-Penrose inverse of P from (7). If $X \leq 0$ satisfies (8), then

$$0 \leq \text{trace}(-B^T XB) \leq m^2 \gamma^2 \|B^T P^\dagger B\|_2 . \qquad (9)$$

Note that $\text{trace}(-B^T XB)$ is monotonically decreasing. Hence, if (9) is violated for some $X \leq 0$ and $\tilde{X} \leq X$, then $\tilde{X}$ cannot solve (8). This bound is very easy to check. Alternatively, we may compare with solutions of Riccati equations from deterministic control. Let $\mathcal{R}_\gamma^{\text{det}}$ denote the counterpart of $\mathcal{R}_\gamma$ with $N = 0$, i.e.

$$\mathcal{R}_\gamma^{\text{det}}(X) = \mathcal{R}_\gamma(X) - N^T XN .$$

Lemma 6: Assume that (A, N) is asymptotically mean-square stable, and $\gamma_1 \geq \gamma > \|\mathbb{L}\|$.
Then the Riccati equation from the deterministic case

$$\mathcal{R}_{\gamma_1}^{\text{det}}(X) = 0 \qquad (10)$$

possesses a smallest solution $X_-^{\gamma_1} \leq 0$, and $X_-^{\gamma_1} \leq X$ for all solutions X of (8).

III. COMPUTATION OF THE STOCHASTIC H^∞-NORM

To exploit the characterization (6), we need a method to check, whether the Riccati equation $\mathcal{R}_\gamma(X) = 0$ possesses a stabilizing solution. Given the Fréchet derivative of $R_\gamma(X)$ displayed in (3), it is natural to apply Newton's method to solve the stochastic algebraic Riccati equation from part (iv) of Theorem 2. The following result was proven in [13].

Theorem 7: Let (A, N) be mean-square stable and assume that $\gamma > \|\mathbb{L}\|$. Consider the Newton iteration

$$X_{k+1} = X_k - (\mathcal{R}_\gamma)'^{-1}_{X_k}(\mathcal{R}_\gamma(X_k)) , \qquad (11)$$

where we assume $\sigma((\mathcal{R}_\gamma)'_{X_0}) \subset \mathbb{C}_-$. Then the sequence X_k converges to X_+, and for all $k \geq 1$ it holds that

$$\sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-, \ \mathcal{R}_\gamma(X_k) \leq 0, \text{ and } X_k \geq X_{k+1} . \quad (12)$$

Moreover, under the given assumptions $X_0 = 0$ is a suitable initial guess. This actually follows from [12, Corollary 5.3.5], but the argument there is not really taylored for our purpose here. Therefore a more direct proof (within the framework of [12]) is given in the appendix.

Lemma 8: Let (A, N) *be mean-square stable and assume that* $\gamma > \|\mathbb{L}\|$. *Then* $\sigma((\mathcal{R}_\gamma)'_0) \subset \mathbb{C}_-$.

For a given $\gamma > \|D\|_2$, we can check whether $\gamma > \|\mathbb{L}\|$ by running the Newton iteration (11) starting from $X_0 = 0$. If all iterates are stabilizing, and the sequence converges with a given level of tolerance, then we conclude that $\gamma \geq \|\mathbb{L}\|$.

Conversely, if $\gamma < \|\mathbb{L}\|$, then either $\sigma((\mathcal{R}_\gamma)'_{X_k}) \not\subset \mathbb{C}_-$ for some k, or the sequence X_k is monotonically decreasing and unbounded. In principle, the latter case can be detected by Lemma 6. It suffices to compute X^{γ_1} for some $\gamma_1 > \|\mathbb{L}\|$ and to test whether $X_k \geq X_-^{\gamma_1}$. If X_k is unbounded, then this condition will be violated for some iterate indicating, that $\gamma < \|\mathbb{L}\|$. In practice, however, we never observed that this bound (or the easier condition of Lemma 5) was violated earlier than the stability condition.

Thus, if for some k the condition $\sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-$ is violated or the iteration takes more than a fixed number of steps, then we conclude that $\gamma \leq \|\mathbb{L}\|$.

Using bisection, we can thus compute $\|\mathbb{L}\|$ up to a given precision, which in theory can be arbitrarily small, in practice, of course, is restricted by roundoff effects.

A. The basic algorithm

We summarize this approach as our *basic algorithm*.

Algorithm 9 *Computation of the stochastic H^∞-norm*

1: Choose $\gamma_0 < \|\mathbb{L}\| < \gamma_1$, $k_{\max}$, tol
2: **repeat**
3:     Set $\gamma = \frac{\gamma_0 + \gamma_1}{2}$, $X_0 = 0$
4:     **repeat**
5:       **if** $\sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-$ **then**
6:         $X_{k+1} = X_k - (\mathcal{R}_\gamma)'^{-1}_{X_k}(\mathcal{R}(X_k))$
7:       **end if**
8:     **until** convergence **or** $k = k_{\max}$ **or** $\sigma((\mathcal{R}_\gamma)'_{X_k}) \not\subset \mathbb{C}_-$
9:     **if** convergence **then**
10:       $\gamma_1 = \gamma$,
11:     **else**
12:       $\gamma_0 = \gamma$
13:     **end if**
14: **until** $\gamma_1 - \gamma_0 < $ tol

The stability test in line 5 and the solution of the linear system in line 6 are central issues. Both concern the generalized Lyapunov mapping $\mathcal{R}'_{X_k}$. A naive implementation with general purpose eigenvalue and linear system solvers, respectively, would result in an overall complexity of about $\mathcal{O}(n^6)$. The complexity of semidefinite programming methods is slightly smaller: For Lyapunov and Riccati equations, [14] reports a worst case estimate of $\mathcal{O}(n^{5.5})$ and an average of $\mathcal{O}(n^4)$. Our numerical experiments suggest that similar estimates hold for the computation of the stochastic H^∞-norm. It is, however, well known that standard Lyapunov

equations of the form $\mathcal{L}_{A_{X_k}}(X) = Y$ can be solved in $\mathcal{O}(n^3)$ operations, using e.g. the Bartels-Stewart algorithm, [15]. Exploiting this in iterative approaches, we can bring down the complexity of Algorithm 9 to $\mathcal{O}(n^3)$. This will be explained briefly in the following two subsections. Moreover, we suggest a way to choose γ_0 and γ_1 in line 1.

In the numerical experiments, we will show that our algorithm outperforms general purpose LMI methods.

B. The stability test

The condition $\sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-$ in line 5 holds if and only if $\sigma(A_{X_k}) \subset \mathbb{C}_-$ and $\rho(\mathcal{L}_{A_{X_k}}^{-1} \Pi_N) < 1$, where ρ denotes the spectral radius, [12, Theorem 3.6.1]. Hence, we can first check, whether $\sigma(A_{X_k}) \subset \mathbb{C}_-$ and then apply the power method to compute the spectral radius ρ of $\mathcal{L}_{A_{X_k}}^{-1} \Pi_N$. Note that the mapping $-\mathcal{L}_{A_{X_k}}^{-1} \Pi_N$ is nonnegative, in the sense that it maps the cone of nonnegative definite matrices to itself, see [12]. Hence, the iterative scheme

$$P_0 = I, \ P_{k+1} = -\mathcal{L}_{A_{X_k}}^{-1} \Pi_N(P_k), \ \rho_k = \frac{\text{trace}(P_k P_{k+1})}{\text{trace}(P_k P_k)}$$

produces a sequence of nonnegative definite matrices P_k which generically converge to the dominant eigenvector. In the limit we have $P_{k+1} \approx \rho P_k$, i.e. $\rho_k \overset{k \to \infty}{\to} \rho$.

C. The generalized Lyapunov equation

By (4), in the Newton step in line 6, the generalized Lyapunov equation

$$A_{X_k}^T \Delta + \Delta A_{X_k} + N^T \Delta N = -\mathcal{R}_\gamma(X_k) \qquad (13)$$

has to be solved for Δ to obtain $X_{k+1} = X_k + \Delta$. Equations of this type have been studied e.g. in [16].

Here $\Delta = \Delta^T \in \mathbb{R}^{n \times n}$ satisfies the fixed point equation

$$\Delta = -\mathcal{L}_{A_{X_k}}^{-1} (\Pi_N(\Delta) + \mathcal{R}_\gamma(X_k)).$$

The condition $\sigma((\mathcal{R}_\gamma)'_{X_k}) \subset \mathbb{C}_-$ implies $\rho(\mathcal{L}_{A_{X_k}}^{-1} \Pi_N) < 1$, where ρ denotes the spectral radius. Hence the iteration

$$\Delta_{j+1} = -\mathcal{L}_{A_{X_k}}^{-1} (\Pi_N(\Delta_j) + \mathcal{R}_\gamma(X_k))$$

is convergent. In each step this iteration only requires the solution of a standard Lyapunov equation at a cost in $\mathcal{O}(n^3)$. The speed of convergence can be improved by using a Krylov subspace approach like `gmres` or `bicgstab`. For details see [16]. More recently, also low-rank techniques have been considered in [17], [18], [19].

D. Choosing γ_0 and γ_1

For the bisection it is useful to find suitable upper and lower bounds for $\|\mathbb{L}\|$. Let $G(s) = C(sI - A)^{-1}B + D$ be the transfer function of the deterministic system obtained from (1) by replacing N with zero. The H^∞-norm $\|G\|_{H^\infty}$ equals the input-output norm of this deterministic system. Then from Theorem 2 we have $\|G\|_{H^\infty} \leq \|\mathbb{L}\|$, because the inequality (5) for a given matrix $X < 0$ implies that the corresponding linear matrix inequality with $N = 0$ holds for the same X. Hence, if $\gamma > \|\mathbb{L}\|$, then $\gamma > \|G\|_{H^\infty}$. So,

we choose $\gamma_0 = \|G\|_{H^\infty}$ and try $\gamma_1 = 2\gamma_0$. If the Newton iteration does not converge for γ_1, then we replace γ_0 by $2\gamma_0$ and repeat the previous step, until we have $\gamma_1 > \|\mathbb{L}\|$.

IV. Numerical experiments

The experiments were carried out on a 2016 MacBook Pro with a 3.3 GHz Intel Core i7 processor and 16 GB Memory running OS X 10.12.4 using MATLAB® version R2017a.

A. Random systems

We first consider random data (A, N, B, C) produced by `randn`. The matrix A is made stable by mirroring the unstable eigenvalues at $i\mathbb{R}$. Then the spectral radius ρ of $\mathcal{L}_A^{-1}\Pi_N$ is estimated as described in subsection III-B and an update of N is obtained by multiplication with $(2\rho+1)^{-1/2}$. Thus (A, N) is guaranteed to be mean-square stable. We compute the stochastic H^∞-norm by our algorithm and compare it with the results obtained by both the MATLAB-function `mincx`, and Sedumi [20] via the toolbox YALMIP, [21] (see Appendix C). For each dimension, 20 random examples are considered. In most of our tests, the computed H^∞-norm coincided for all three methods roughly up to the chosen relative precision `tol=1e-08`. Among the 120 examples for dimension $n = 10, 20, 40, 60, 80, 100$, there were only 2 cases, where `mincx` did not converge, and two more cases where the results of Algorithm 9 and `mincx` differed between 1% and 4%. However, in 13 cases (not analyzed further), the result computed by Sedumi differed by magnitudes from the other two. The averaged computing times are given in Table I and Figure 1. Note that `mincx` is a built-in function, and therefore outperforms generic MATLAB functions for small n, as it saves the time the interpreter needs to execute the code and to manage the memory requirements. For high-level implementations of our method and of Sedumi, we would expect to observe a speed-up also for smaller dimensions. For larger n the asymptotic complexity becomes relevant. The regression lines in Figure 1 have slopes $s = 2.7$ for Algorithm 9, $s = 6.4$ for `mincx` and $s = 5.9$ for Sedumi, indicating that the respective complexities are roughly $\mathcal{O}(n^s)$.

TABLE I
AVERAGED COMPUTING TIMES (IN SEC) FOR RANDOM SYSTEMS.

n	10	20	40	60	80	100	200	400
`mincx`	0.03	0.3	10	141	867	3635	-	-
Sedumi	1.1	1.1	11	107	577	2550	-	-
Alg. 9	1.2	2.4	5	15	23	50	280	1600

B. A heat transfer problem

This stochastic modification of a heat transfer problem described in [8] was also discussed in [9]. On the unit square $\Omega = [0,1]^2$, the heat equation $T_t = \Delta T$ for $T = T(t,x)$ is given with Dirichlet condition $T = u_j$, $j = 1, 2, 3$, on three of the boundary edges and a stochastic Robin condition $n \cdot \nabla T = (1/2 + \dot{w})T$ on the fourth edge (where $\dot{w}$ stands for white noise). We measure the average value $y(t) = \int_\Omega T(t,x)\,dx$.
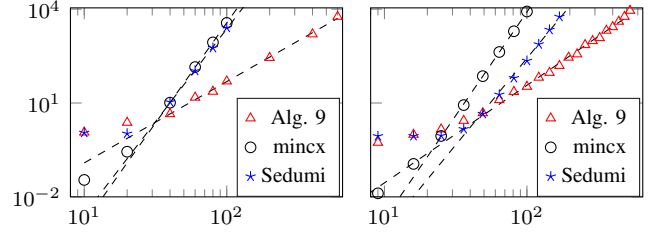

Fig. 1. Computing times for random examples as in Table I (left) and the heat transfer problem (right)
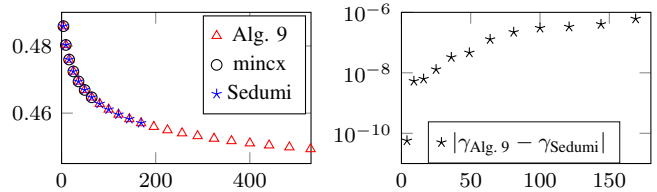

Fig. 2. Computed norms for discretized heat equation (left) and their deviation for Algorithm 9 and Sedumi (right)

A standard 5-point finite difference discretization on a $k \times k$ grid leads to a modified Poisson matrix $A \in \mathbb{R}^{n \times n}$ with $n = k^2$ and corresponding matrices $N \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 3}$ $C = \frac{1}{n}[1, \ldots, 1] \in \mathbb{R}^{n \times n}$. The H^∞-norm of this discretization $\|\mathbb{L}_n\|$ approximates the induced input/output norm of the partial differential equation. Figure 1 (see also Table II) displays the computing times (in seconds) for $n = 4, 9, 16, \ldots, 529$ and the regression lines with slopes $s = 3.1$ for Algorithm 9, $s = 6.6$ for `mincx`, and $s = 5.5$ for Sedumi. Figure 2 shows the computed norms (which nearly coincide for all methods) and their differences exemplarily for Algorithm 9 and Sedumi.

TABLE II
COMPUTING TIMES (IN SEC) FOR DISCRETIZED HEAT EQUATION.

n	16	49	100	144	196	256	400
`mincx`	0.1	72	8418	-	-	-	-
Sedumi	0.9	4.5	223	2191	13219	-	-
Alg. 9	1	4.8	33	93	285	720	2699

Again, we observe that Algorithm 9 allows to treat larger dimensions than the LMI-solver. However, the computing times for our algorithm also grow fairly fast. As an alternative to bisection one might consider extrapolating the spectral radii $\rho(\gamma) = \rho\left((\mathcal{R}_\gamma)'_{X_+(\gamma)}\right)$ which are computed in the course of the process for $\gamma > \|\mathbb{L}\|$, or perhaps the spectral abscissae $\alpha(\gamma) = \max \Re\sigma\left((\mathcal{R}_\gamma)'_{X_+(\gamma)}\right)$. Then the norm $\|\mathbb{L}\|$ is given as the value of γ, where $\rho(\gamma) = 1$, or $\alpha(\gamma) = 0$. Unfortunately, the slopes of ρ and α are very steep as γ approaches $\|\mathbb{L}\|$. Thus, an extrapolation does not seem promising. The behaviour is visualized for the heat equation system with $n = 25$ in Figure 3.

V. Conclusions

We have suggested an algorithm to compute the stochastic H^∞-norm. It builds upon several ideas developed in the
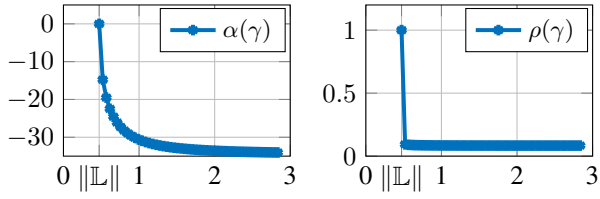
Fig. 3. Spectral abscissa $\alpha(\gamma)$ and spectral radius $\rho(\gamma)$ of $X_+(\gamma)$ close to the critical value $\gamma = \|\mathbb{L}_{25}\| = 0.47241$.

literature, and is the first algorithm, whose complexity is considerably smaller than that of a general purpose LMI-solver. We chose to present the algorithm for the simplest case of just one multiplicative noise term, which however can easily be generalized to the class described in Appendix A. Already in the simple case, the stochastic H^∞-norm is much harder to compute than the H^∞-norm of a deterministic system, and the computing times are still very high. We see it as a challenge to come up with a faster method.

The note also contains some extensions of known results with new proofs, like the nonstrict stochastic bounded real lemma and lower bounds for Riccati solutions.

APPENDIX

A. Generalization

System (1) can be generalized in a straight-forward manner to the case of multiple noise terms at the state and the input (see e.g. [12]). Then our system takes the form

$$dx = (Ax + Bu)\,dt + \sum_{j=1}^{\nu}(N_{x,j}x + N_{u,j}u)\,dw_j \quad (14)$$

$$y = Cx + Du\,, \quad (15)$$

where $N_{x,j} \in \mathbb{R}^{n \times n}$, $N_{u,j} \in \mathbb{R}^{n \times m}$ and the w_j are independent Wiener processes. Here, $\mathcal{R}_\gamma$ is given by

$$\begin{aligned}
\mathcal{R}_\gamma(X) &= P(X) - S(X)^T Q_\gamma(X)^{-1} S(X)\,, \text{ where} \\
P(X) &= A^T X + XA + \sum_{j=1}^{\nu} N_{x,j}^T X N_{x,j} - C^T C\,, \\
S(X) &= B^T X + \sum_{j=1}^{\nu} N_{u,j}^T X N_{x,j} - C^T D\,, \\
Q_\gamma(X) &= \sum_{j=1}^{\nu} N_{u,j}^T X N_{u,j} + \gamma^2 I - D^T D\,.
\end{aligned}$$

Our basic algorithm and all our considerations carry over to this case literally. Only the expressions for $\mathcal{R}_\gamma$ and $(\mathcal{R}_\gamma)'_X$ become more technical.

B. Proofs

Proofs are provided here for results which are new (Lemmas 5 and 6), or partly new (Corollary 4), or – in the case of Lemma 8 – whose proof is not easy to access. Note that the novelty of Corollary 4 lies in its usage of the controllability Gramian P. This makes the statement more general and leads to a significant simplification of the proof compared to [13]. As in the main text, we write $\mathcal{L}_A : X \mapsto A^T X + XA$ and $\Pi_N : X \mapsto N^T XN$. On the space of symmetric matrices we consider the scalar product $\langle X, Y \rangle = \operatorname{trace} XY$, and note that the corresponding adjoint operators are

$$\mathcal{L}_A^* : X \mapsto AX + XA^T \quad \text{and} \quad \Pi_N^* : X \mapsto NXN^T\,.$$

Further general facts on Riccati- and Lyapunov-type operators are cited from [12].

Proof of Corollary 4: (iii)$\Rightarrow$(ii) follows from the definiteness criterion via the Schur-complement.

(ii)$\Rightarrow$(iii): If (ii) holds, then $\mathcal{R}_\gamma(X) \geq 0$, and by [13] there exists a solution $X_+ \leq 0$ to the equation $\mathcal{R}_\gamma(X) = 0$.

(ii)$\Rightarrow$(i): If (5) holds and we replace C and D by $C_\varepsilon = \left[\begin{smallmatrix} C \\ \varepsilon I \end{smallmatrix}\right]$ and $D_\varepsilon = \left[\begin{smallmatrix} D \\ 0 \end{smallmatrix}\right]$, then we get

$$\begin{bmatrix} A^T X + XA + N^T XN - C_\varepsilon^T C_\varepsilon & XB - C_\varepsilon^T D_\varepsilon \\ B^T X - D_\varepsilon^T C_\varepsilon & \gamma^2 I - D_\varepsilon^T D_\varepsilon \end{bmatrix} > 0\,.$$

This implies $\|\mathbb{L}_\varepsilon\| < \gamma$ for the corresponding modified input-output operator. By $\|\mathbb{L}_\varepsilon\| \overset{\varepsilon \to 0}{\to} \|\mathbb{L}\|$, we obtain $\|\mathbb{L}\| \leq \gamma$.

(i)$\Rightarrow$(ii): If (i) holds, then $\|\mathbb{L}\| < \gamma + \frac{1}{k}$ for all $k \in \mathbb{N}$, $k > 0$. Hence there exist stabilizing solutions $X_k \leq 0$ of $\mathcal{R}_{\gamma+\frac{1}{k}}(X) = 0$. Moreover, X_k is the largest solution of (8) with γ replaced by $\gamma + \frac{1}{k}$. Hence it follows that $X_{k+1} \leq X_k$ for all k. If the X_k are bounded below, then the sequence (X_k) converges and the limit satisfies the nonstrict linear matrix inequality in (iii). Thus it suffices to show boundedness. We assume that the sequence is not bounded, i.e. $\|X_k\| \to \infty$ for $k \to \infty$. Consider the normalized sequence $\tilde{X}_k = \frac{X_k}{\|X_k\|}$, which – by Bolzano-Weierstrass – has a convergent subsequence $\tilde{X}_{k_j}$ with limit $\tilde{X} \neq 0$. Then

$$\begin{aligned}
0 &\leq \frac{1}{\|X_{k_j}\|}\begin{bmatrix} (\mathcal{L}_A + \Pi_N)(X_{k_j}) - C^T C & X_{k_j}B - C^T D \\ B^T X_{k_j} - D^T C & \gamma^2 I - D^T D \end{bmatrix} \\
&\overset{j \to \infty}{\to} \begin{bmatrix} A^T \tilde{X} + \tilde{X}A + N^T \tilde{X}N & \tilde{X}B \\ B^T \tilde{X} & 0 \end{bmatrix} \geq 0\,,
\end{aligned}$$

implying $B^T \tilde{X} = 0$ and $0 \neq A^T \tilde{X} + \tilde{X}A + N^T \tilde{X}N \geq 0$. Since, by assumption, $P > 0$, we obtain

$$\begin{aligned}
0 &> \operatorname{trace}\left(P(A^T \tilde{X} + \tilde{X}A + N^T \tilde{X}N)\right) \\
&= \operatorname{trace}\left((AP + PA^T + NPN^T)\tilde{X}\right) \\
&= -\operatorname{trace} BB^T \tilde{X} = 0
\end{aligned}$$

which is a contradiction.

Thus, $\mathcal{R}_\gamma(X) = 0$ has a solution X_∞, which is the limit of the largest and stabilizing solutions X_k of $\mathcal{R}_{\gamma+\frac{1}{k}}(X) \geq 0$. Thus X_∞ is the largest solution of $\mathcal{R}_\gamma(X) = 0$ and $\sigma(\mathcal{R}_\gamma)'_{X_\infty} \subset \mathbb{C}_- \cup i\mathbb{R}$. If $\gamma = \|\mathbb{L}\|$ then $\sigma(\mathcal{R}_\gamma)'_{X_\infty} \cap i\mathbb{R} \neq \emptyset$ and [12, Theorem 3.2.3] yields that $0 \in \sigma(\mathcal{R}_\gamma)'_{X_\infty}$. ∎

Proof of Lemma 5:

We have $P = -(\mathcal{L}_A + \Pi_N)^{-*}(BB^T)$.

In the following consider an arbitrary matrix $X \leq 0$, $X \neq 0$, satisfying $(\mathcal{L}_A + \Pi_N)(X) = Y \geq 0$. Then

$$\begin{aligned}
m\|B^T XB\|_2 &\geq |\operatorname{trace}(B^T XB)| \\
&= \langle(\mathcal{L}_A + \Pi_N)^{-1}(Y), -BB^T\rangle = \langle Y, P\rangle\,.
\end{aligned}$$

There exists a vector $u \in \mathbb{R}^m$ with $\|u\|_2 = 1$ and $\langle Y, Buu^T B^T\rangle = u^T B^T XBu = -\|B^T XB\|_2$. Moreover

$$\langle Y, Buu^T B^T\rangle \leq \langle Y, BB^T\rangle \leq \alpha_*\langle Y, P\rangle \quad (16)$$

for $\alpha_* = \|B^T P^\dagger B\|_2$. To see this, note that the image of B is contained in the image of P. Hence there exists a unitary U, such that

$$\alpha P - BB^T = U \begin{bmatrix} \alpha P_1 - B_1 B_1^T & 0 \\ 0 & 0 \end{bmatrix} U^T, \quad \det P_1 \neq 0.$$

The largest zero of $\chi(\alpha) = \det(\alpha P_1 - B_1 B_1^T)$ is

$$\alpha_* = \|P^{-1/2} B_1\|_2^2 = \|B^T P^\dagger B\|_2^2 .$$

For $\alpha \geq \alpha_*$, we have $\alpha P - BB^T \geq 0$ which proves (16). We set $\mu(X) = \langle Y, P \rangle = |\operatorname{trace}(B^T X B)|$. Let now X satisfy (8). With the given data and $\eta > 0$ this implies

$$0 \leq \begin{bmatrix} Bu \\ \eta u \end{bmatrix}^T \begin{bmatrix} Y & XB \\ B^T X & \gamma^2 I \end{bmatrix} \begin{bmatrix} Bu \\ \eta u \end{bmatrix}$$
$$= u^T B^T Y B u + 2 \eta u^T B^T X B u + \gamma^2 \eta^2$$
$$\leq \alpha_* \mu(X) - \frac{2}{m} \mu(X) \eta + \gamma^2 \eta^2$$
$$= \gamma^2 \left(\eta - \frac{\mu(X)}{m \gamma^2} \right)^2 - \frac{\mu(X)^2}{m^2 \gamma^2} + \mu(X) \alpha_* .$$

If we assume $\mu(X) > m^2 \gamma^2 \alpha_*$, then the right hand is negative for $\eta = \frac{\mu(X)}{m \gamma^2}$, which is a contradiction. Hence, we have $|\operatorname{trace}(B^T X B)| \leq m^2 \gamma^2 \|B^T P^\dagger B\|_2$. ∎

Proof of Lemma 6: Note that $\mathcal{R}_{\gamma_1}^{\det}(X) \geq \mathcal{R}_\gamma(X)$ if $X \leq 0$ and $\gamma \leq \gamma_1$ and thus every solution of $\mathcal{R}_{\gamma_1}(X) > 0$ also satisfies $\mathcal{R}_{\gamma_1}^{\det}(X) > 0$. Hence $\mathcal{R}_{\gamma_1}^{\det}(X) = 0$ possesses a stabilizing solution, and, consequently, also an anti-stabilizing solution X_-, which is the smallest solution of $\mathcal{R}_{\gamma_1}^{\det}(X) \geq 0$. Thus also $X_- \leq X$ for every solution X of $\mathcal{R}_{\gamma_1}(X) \geq 0$. ∎

Proof of Lemma 8: We exploit the concavity of $\mathcal{R}_\gamma$ and the resolvent positivity of $(\mathcal{R}_\gamma)_0'$, see [12]. If $\|\mathbb{L}\| \leq \gamma$, then there exists $X \leq 0$ such that, by concavity,

$$0 = \mathcal{R}_\gamma(X) \leq \mathcal{R}_\gamma(0) + (\mathcal{R}_\gamma)_0'(X) . \quad (17)$$

Assume that $\sigma\left((\mathcal{R}_\gamma)_0'\right) \not\subset \mathbb{C}_-$. Then by [12, Theorem 3.2.3] there exists $H \geq 0$, $\lambda \geq 0$, such that $(\mathcal{R}_\gamma)_0'(H) = \lambda H$. Taking the scalar product of inequality (17) with H, we get

$$0 \leq \langle \mathcal{R}_\gamma(0), H \rangle + \lambda \langle X, H \rangle \leq 0 .$$

It follows that $\mathcal{R}_\gamma(0) H = 0$, which implies $D^T C H = 0$ and thus $A_0 H = AH$. But then $(\mathcal{L}_A + \Pi_N)^*(H) = \lambda H$ in contradiction to the stability of (A, N). ∎

C. Usage of LMI-solvers in MATLAB

The solver `mincx` was used as in the following listing.

```
setlmis([])
X = lmivar(1,[n,1]);g = lmivar(1,[1,1]);
lmiterm([1 1 1 X],N',N);
lmiterm([1 1 1 X],A',1,'s');
lmiterm([1 1 1 0],C'*C);
lmiterm([1 1 2 X],1,B,'s');
lmiterm([1 2 2 g],-1,1);
lmisys = getlmis;
c = mat2dec(lmisys,zeros(n),1);
options = [tol,0,0,0,1];
copt = mincx(lmisys,c,options);
gamma = sqrt(copt)
```

Sedumi was called from the toolbox YALMIP, [21].

```
P=sdpvar(n,n);gamma=sdpvar(1);
F=[P<=0,[A'*P+P*A+N'*P*N-C'*C,P*B-C'*D;
   B'*P-D'*C gamma*eye(size(B,2))-D'*D]>=0];
ops = sdpsettings('solver','sedumi',
 'sedumi.eps',tol,'verbose',0);
optimize(F,gamma,ops);
gamma=sqrt(value(gamma))
```

REFERENCES

[1] S. Boyd and V. Balakrishnan, "A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L^∞-norm," *Syst. Control Lett.*, vol. 15, no. 1, pp. 1–7, 1990.

[2] N. A. Bruinsma and M. Steinbuch, "A fast algorithm to compute the H^∞-norm of a transfer function matrix," *Syst. Control Lett.*, vol. 14, pp. 287–293, 1990.

[3] N. Guglielmi, M. Gürbüzbalan, and M. L. Overton, "Fast approximation of the H_∞-norm via optimization over spectral value sets," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 2, pp. 709–737, 2016.

[4] P. Benner and M. Voigt, "A structured pseudospectral method for $\mathcal{H}_\infty$-norm computation of large-scale descriptor systems," *Math. Control Signals Syst.*, vol. 26, no. 2, pp. 303–338, 2014.

[5] M. A. Freitag, A. Spence, and P. Van Dooren, "Calculating the H_∞-norm using the implicit determinant method," *Linear Algebra Appl.*, vol. 35, no. 2, pp. 619–635, 2014.

[6] N. Aliyev, P. Benner, E. Mengi, P. Schwerdtner, and M. Voigt, "Large-scale computation of $\mathcal{L}_\infty$-norms by a greedy subspace method," arXiv:1705.10086, 2017.

[7] D. Hinrichsen and A. J. Pritchard, "Stochastic H_∞," *SIAM J. Control Optim.*, vol. 36, no. 5, pp. 1504–1538, 1998.

[8] P. Benner and T. Damm, "Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems," *SIAM J. Control Optim.*, vol. 49, no. 2, pp. 686–711, 2011.

[9] P. Benner, T. Damm, and Y. Rocio Rodriguez Cruz, "Dual pairs of generalized Lyapunov inequalities and balanced truncation of stochastic linear systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 782–791, 2017.

[10] L. Arnold, *Stochastic Differential Equations: Theory and Applications. Translation.* New York etc.: John Wiley and Sons Inc., 1974.

[11] B. Oeksendal, *Stochastic Differential Equations.* Springer, 1998.

[12] T. Damm, *Rational Matrix Equations in Stochastic Control*, ser. Lecture Notes in Control and Information Sciences. Springer, 2004, no. 297.

[13] T. Damm and D. Hinrichsen, "Newton's method for a rational matrix equation occuring in stochastic control," *Linear Algebra Appl.*, vol. 332–334, pp. 81–109, 2001.

[14] L. Vandenberghe and S. Boyd, "A primal-dual potential reduction method for problems involving matrix inequalities," *Mathematical Programming*, vol. 69, no. 1, pp. 205–236, Jul. 1995.

[15] R. H. Bartels and G. W. Stewart, "Algorithm 432: The solution of the matrix equation $AX + XB = C$," *Comm. ACM*, vol. 15, no. 9, pp. 820–826, 1972.

[16] T. Damm, "Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations," *Numer. Lin. Alg. Appl.*, vol. 15, no. 9, pp. 853–871, 2008.

[17] P. Benner and T. Breiten, "Low rank methods for a class of generalized Lyapunov equations and related issues." *Numer. Math.*, vol. 124, no. 3, pp. 441–470, 2013.

[18] S. Shank, V. Simoncini, and D. Szyld, "Efficient low-rank solutions of generalized Lyapunov equations," *Numer. Math.*, vol. 134, pp. 327–342, 2016.

[19] D. Kressner and P. Sirković, "Greedy low-rank methods for solving general linear matrix equations," ANCHP, MATHICSE, EPF Lausanne, Switzerland, Technical report, 2014.

[20] J. F. Sturm, "Using SeDuMi 1. 02, a MATLAB toolbox for optimization over symmetric cones." *Optim. Methods Softw.*, vol. 11-12, no. 1-4, pp. 625–653, 1999.

[21] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.