Journal of Computational and Applied Mathematics 397 (2021) 113666

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Sensitivity analysis of random linear differential–algebraic equations using system norms

Roland Pulch [a,*], Akil Narayan [b], Tatjana Stykel [c]

[a] Institute of Mathematics and Computer Science, Universität Greifswald, Walther-Rathenau-Straße 47, D-17489 Greifswald, Germany
[b] Scientific Computing and Imaging Institute, Department of Mathematics, University of Utah, 72 Central Campus Dr., Salt Lake City, UT 84112, United States
[c] Institute of Mathematics, Universität Augsburg, Universitätsstraße 14, D-86159 Augsburg, Germany

ARTICLE INFO

Article history:
Received 13 August 2020
Received in revised form 14 May 2021

Keywords:
Differential–algebraic equations
Polynomial chaos
Sensitivity indices
Hardy norm
Balanced truncation
Uncertainty quantification

ABSTRACT

We consider linear dynamical systems composed of differential–algebraic equations (DAEs), where a quantity of interest (QoI) is assigned as output. Physical parameters of a system are modelled as random variables to quantify uncertainty, and we investigate a variance-based sensitivity analysis of the random QoI. Based on expansions via generalised polynomial chaos, the stochastic Galerkin method yields a new deterministic system of DAEs of high dimension. We define sensitivity measures by system norms, i.e., the $\mathcal{H}_\infty$-norm of the transfer function associated with the Galerkin system for different combinations of outputs. To ameliorate the enormous computational effort required to compute norms of high-dimensional systems, we apply balanced truncation, a particular method of model order reduction (MOR), to obtain a low-dimensional linear dynamical system that produces approximations of system norms. MOR of DAEs is more sophisticated in comparison to systems of ordinary differential equations. We show an a priori error bound for the sensitivity measures satisfied by the MOR method. Numerical results are presented for two stochastic models given by DAEs.

1. Introduction

In science and engineering, mathematical modelling often generates dynamical systems consisting of ordinary differential equations (ODEs) or differential-algebraic equations (DAEs). Such dynamical systems may include physical and geometrical parameters describing possible variations in the system. For example, electronic design automation using network approaches yields parametric systems of DAEs as models of electric circuits, see [1–3].

Often uncertainties are inherent in some parameters due to modelling errors, measurement errors or imperfections of an industrial production. In this case, reliable predictions require incorporation of parametric variation in a model. In uncertainty quantification, the most common approach is to replace the parameters by random variables, see [4,5]. Consequently, the solution of a dynamical system becomes a random process.

We consider linear time-invariant control systems of DAEs, also called descriptor systems, which include random parameters. A quantity of interest (QoI) is extracted as a single output. Our aim is to perform a sensitivity analysis of the QoI with respect to the random variables. For this purpose, variance-based or derivative-based sensitivity measures can

* Corresponding author.
E-mail addresses: roland.pulch@uni-greifswald.de (R. Pulch), akil@sci.utah.edu (A. Narayan), stykel@math.uni-augsburg.de (T. Stykel).

https://doi.org/10.1016/j.cam.2021.113666
0377-0427/

be employed, see [6–8]. We apply a global variance-based concept, where the total effect sensitivity indices characterise the significance of the random parameters. Sensitivity analysis of this kind was previously applied to linear dynamical systems in [9,10]. Our approach is based on converting a random linear dynamical system into a larger deterministic linear dynamical system by using the stochastic Galerkin method [11–13]. In [14], a variance-based sensitivity analysis of the QoI was introduced in the case of linear ODEs, where system norms such as the Hankel norm and the $\mathcal{H}_\infty$-norm of the stochastic Galerkin formulations yield sensitivity measures.

In this paper, we extend this approach to the case of linear DAEs. In comparison to ODEs, systems of DAEs exhibit more complicated analytical properties and require modified numerical methods, see [15,16]. Similar to the ODE case [14], we apply the stochastic Galerkin method to the random system of DAEs. Furthermore, we prove a convergence property of the Galerkin approach for partial variances related to the sensitivity measures, which is valid for both ODEs and DAEs. The Galerkin projection is followed by a computation of several $\mathcal{H}_\infty$-norms for the same system of DAEs, but different selections of multiple outputs. In the single-output case, the number of required norms is equal to the number of random parameters.

The numerical computation of $\mathcal{H}_\infty$-norms is typically performed by iterative schemes like bisection methods. Numerical techniques can be found for ODEs in [17,18] and for DAEs in [19,20], see also the references therein. In these techniques, (generalised) eigenvalue problems of size twice as large as the dimension of the dynamical system have to be solved in each iteration step. Similar to the previously investigated ODE case [14], in order to reduce the computational effort for computing $\mathcal{H}_\infty$-norms of the related high-dimensional Galerkin systems, we use model order reduction (MOR). While doing this, we construct low-dimensional systems, whose $\mathcal{H}_\infty$-norms yield approximations to the sensitivity measures. It should, however, be noted that model reduction of DAEs is more involved in comparison to ODEs, see [21] for a survey. Based on the previous works [22,23], we derive a balanced truncation technique for an MOR of the high-dimensional stochastic Galerkin system with many outputs. Consequently, guaranteed error bounds are derived for the sensitivity measures. Finally, we present some results of numerical experiments for two models of electric circuits.

2. Linear differential–algebraic equations

First, we outline some general concepts for linear control systems of DAEs required in the following sections.

2.1. Linear dynamical systems

We consider a linear time-invariant control system

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) \end{aligned} \tag{1}$$

with matrices $A, E \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_{\text{in}}}$, and $C \in \mathbb{R}^{n_{\text{out}}\times n}$. A system of the form (1) is a DAE (not an ODE) system if and only if the matrix E is singular. The input $u : [0,\infty) \to \mathbb{R}^{n_{\text{in}}}$ is supplied and the state is $x : [0,\infty) \to \mathbb{R}^n$. The QoI is given by the output $y : [0,\infty) \to \mathbb{R}^{n_{\text{out}}}$. We examine initial value problems determined by (1) together with the consistent initial value condition $Ex(0) = Ex_0$. Without loss of generality, we assume that $Ex_0 = 0$. If this is not the case, we can proceed similarly to [24] by considering a shifted state $z(t) = x(t) - x_0 e^{-\alpha t}$ with $\alpha > 0$. This approach results in a DAE system with a consistent homogeneous initial condition $Ez(0) = 0$ and a new input $[u(t)^\top, e^{-\alpha t}]^\top$, which belongs to the space $\mathcal{L}^2([0,\infty))$ of square-integrable functions provided $u \in \mathcal{L}^2([0,\infty))$.

Definition 1. The DAE system (1) has a *regular matrix pencil* $\lambda E - A$ if the characteristic polynomial $\chi(\lambda) = \det(\lambda E - A)$ is not identically zero.

The regularity of the matrix pencil $\lambda E - A$ is necessary and sufficient for the well-posedness of the initial value problems. Every regular pencil $\lambda E - A$ can be transformed into a quasi-Weierstrass form

$$E = W \begin{bmatrix} I_{n_1} & 0 \\ 0 & N \end{bmatrix} T, \qquad A = W \begin{bmatrix} J & 0 \\ 0 & I_{n_2} \end{bmatrix} T, \tag{2}$$

where $W, T \in \mathbb{R}^{n\times n}$ are non-singular, $J \in \mathbb{R}^{n_1\times n_1}$, and the matrix $N \in \mathbb{R}^{n_2\times n_2}$ with $n_2 = n - n_1$ is nilpotent, see [16]. The eigenvalues of J are the *finite eigenvalues* of the pencil $\lambda E - A$. The block N corresponds to *eigenvalues at infinity*.

Definition 2. The *index* of the DAE system (1) is the smallest integer ν such that the matrix N in the quasi-Weierstrass form (2) satisfies $N^{\nu-1} \neq 0$ and $N^\nu = 0$.

The index is an important characteristic of DAE systems used in the solvability analysis, numerical methods, and control problems for DAEs. Note that the index is independent of the definition of inputs (matrix B) as well as outputs (matrix C) in the system.

The time-dependent problem (1) with a regular pencil $\lambda E - A$ can be transformed into the frequency domain as follows. Applying the Laplace transform to (1) with $Ex(0) = 0$, we obtain the algebraic equations

$$\begin{aligned} E(s\mathcal{X}(s)) &= A\mathcal{X}(s) + B\mathcal{U}(s), \\ \mathcal{Y}(s) &= C\mathcal{X}(s), \end{aligned} \tag{3}$$

where $\mathcal{U}$, $\mathcal{X}$ and $\mathcal{Y}$ are the Laplace transforms of u, x and y, respectively, and $s \in \mathbb{C}$ is the frequency. Therein, we assume that the Laplace transforms exist for all s satisfying $\mathrm{Re}(s) \geq a$ with some $a \in \mathbb{R}$. Solving the first equation in (3) for $\mathcal{X}(s)$ and substituting it into the second one, we get the input–output relation $\mathcal{Y}(s) = H(s)\mathcal{U}(s)$ with the *transfer function*

$$H(s) = C(sE - A)^{-1}B, \tag{4}$$

which is a matrix-valued rational function defined for all complex numbers that are not a finite eigenvalue of the matrix pencil $\lambda E - A$. Due to the singularity of E, the transfer function H may also have a pole at infinity. Such considerations lead to the following high-level taxonomy of transfer functions.

Definition 3. A transfer function (4) is called *strictly proper*, if the degree of the numerator is smaller than the degree of the denominator in each component. A transfer function (4) is called *proper* if the degree of the numerator does not exceed the degree of the denominator in each component. Otherwise, it is called *improper*.

It should be noted that the properness condition depends not only on the state matrices A and E but also on the input matrix B and the output matrix C. A transfer function of an ODE system (1) with $E = I$ is always strictly proper. Many applications lead to linear DAEs which have proper or even strictly proper transfer functions. For example, if the DAE system can be transformed into an equivalent ODE model with the same input–output relation without feedthrough term in the output equation, then it has a strictly proper transfer function. Furthermore, DAE systems of index $\nu = 1$ have always proper transfer functions.

In order to characterise the properness property for DAEs of index $\nu > 1$, we decompose the transformed input and output matrices into blocks

$$W^{-1}B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad CT^{-1} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

according to (2). Then the transfer function can additively be split into

$$H(s) = H_{\mathrm{SP}}(s) + P(s), \tag{5}$$

where $H_{\mathrm{SP}}(s) = C_1(sI_{n_1} - J)^{-1}B_1$ is strictly proper in view of Definition 3, and $P(s) = C_2(sN - I_{n_2})^{-1}B_2$ is a polynomial matrix-valued function. This function can also be written as

$$P(s) = \sum_{j=0}^{\nu-1} M_j s^j \tag{6}$$

with the matrices $M_j = -C_2N^jB_2$, $j = 0, 1, \ldots, \nu - 1$, where ν is the index of (1). Then the DAE system (1) has a proper transfer function if and only if $M_j = 0$ for $j = 1, \ldots, \nu - 1$. Unfortunately, these conditions are difficult to verify in practice, since the computation of the quasi-Weierstrass form (2) is numerically expensive and may be sensitive to perturbations.

Definition 4. The DAE system (1) is called *stable*, if for every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that any solution of the initial value problem (1) with $u(t) \equiv 0$ and $Ex(0) = Ex_0$ satisfies $\|x(t)\| < \varepsilon$ for all $t \geq 0$ and all consistent initial vectors x_0 with $\|x_0\| \leq \delta$. The DAE system (1) is called *asymptotically stable*, if it is stable and $\lim_{t\to\infty} \|x(t)\| = 0$ for every solution x of (1) with $u(t) \equiv 0$.

The used vector norm $\|\cdot\|$ is arbitrary. Similarly to ODEs, the stability properties of linear DAEs can be characterised in a purely algebraic way. One can show that the DAE system (1) with a regular pencil $\lambda E - A$ is asymptotically stable if and only if all the finite eigenvalues of $\lambda E - A$ lie in the open left complex half-plane, see [25], for example. The asymptotic stability implies that the transfer function (4) is bounded for all s with $\mathrm{Re}(s) \geq a$ including a (small) $a < 0$.

2.2. $\mathcal{H}_\infty$-norm

System's norms are used to measure the size of control systems or, equivalently, the magnitude of the associated transfer functions. We refer to [26, Chapter 5] for a review of various norms for ODE control systems. These norms have been extended to DAE control systems in [27]. The most important norms are the $\mathcal{H}_2$-norm and the $\mathcal{H}_\infty$-norm in the corresponding Hardy spaces. They play a crucial role in robust control and MOR. In this paper, we use the $\mathcal{H}_\infty$-norm defined as follows.

Definition 5. The $\mathcal{H}_\infty$-*norm* of the transfer function H in (4) is defined as

$$\|H\|_{\mathcal{H}_\infty} = \sup_{\omega\in\mathbb{R}} \sigma_{\max}(H(\mathrm{i}\omega)), \tag{7}$$

where $\sigma_{\max}(\cdot)$ denotes the largest singular value of the corresponding matrix and $\mathrm{i} = \sqrt{-1}$.

The supremum in (7) is finite if and only if the transfer function H is proper and does not have finite poles in the closed right complex half-plane. An asymptotically stable DAE system (1) of index $\nu = 1$ guarantees that the supremum is finite. This result immediately follows from the additive decomposition (5).

The $\mathcal{H}_\infty$-norm can also be defined as

$$\|H\|_{\mathcal{H}_\infty} = \sup_{\substack{u\in\mathcal{L}^2([0,\infty))\\ u\neq 0}} \frac{\|y\|_{\mathcal{L}^2([0,\infty))}}{\|u\|_{\mathcal{L}^2([0,\infty))}}, \tag{8}$$

where $\|\cdot\|_{\mathcal{L}^2([0,\infty))}$ denotes the $\mathcal{L}^2$-norm in the Lebesgue space $\mathcal{L}^2([0,\infty))$ of square-integrable vector-valued functions. The right-hand side of (8) represents the operator norm of the input–output mapping in the time domain. Thus, the $\mathcal{H}_\infty$-norm can also be interpreted as an $\mathcal{L}^2$-induced operator norm which characterises the $\mathcal{L}^2$-gain from the input u to the output y. We use, however, the frequency domain representation (7) because it is computable using algorithms from numerical linear algebra.

For the computation of the $\mathcal{H}_\infty$-norm, different numerical methods have been developed [17–20,28,29]. Most of them are based on a relationship between the $\mathcal{H}_\infty$-norm and the finite spectrum of the skew-Hamiltonian/Hamiltonian matrix pencil

$$\lambda \begin{bmatrix} E & 0 \\ 0 & E^\top \end{bmatrix} - \begin{bmatrix} A & \frac{1}{\gamma}BB^\top \\ -\frac{1}{\gamma}C^\top C & -A^\top \end{bmatrix} \tag{9}$$

with $\gamma > 0$, see [17,20] and references therein. In these methods, the $\mathcal{H}_\infty$-norm is determined iteratively by updating γ. In each iteration step, a structured eigenvalue problem for (9) of dimension $2n$ has to be solved until the pencil in (9) has no finite eigenvalues on the imaginary axis. Other methods rely on a connection of the $\mathcal{H}_\infty$-norm to the structured complex stability radius of the corresponding system pencil and are more appropriate for large-scale problems [18,29]. An improved method for computing the $\mathcal{H}_\infty$-norm has been proposed in [19] which is based on the direct optimisation approach. Finally, in [28], a subspace projection method has been presented which relies on the computation of the $\mathcal{H}_\infty$-norm of a reduced transfer function obtained by the successive expansion of the projection subspaces.

3. Random linear differential–algebraic equations

In this section, we introduce the investigated problem and present the stochastic Galerkin method. Our considerations mostly follow the ODE theory in [14].

We consider a parametric DAE system

$$\begin{aligned} E(p)\dot{x}(t,p) &= A(p)x(t,p) + B(p)u(t),\\ y(t,p) &= c(p)^\top x(t,p), \end{aligned} \tag{10}$$

where $E(p), A(p) \in \mathbb{R}^{n\times n}$, $B(p) \in \mathbb{R}^{n\times n_{\mathrm{in}}}$ and $c(p) \in \mathbb{R}^n$ depend on a physical parameter vector $p \in \Pi \subseteq \mathbb{R}^q$. Typically, the input $u : [0,\infty) \to \mathbb{R}^{n_{\mathrm{in}}}$ is independent of physical parameters. Yet, the matrix B, which induces these inputs, can be parameter-dependent. Both the state $x : [0,\infty)\times\Pi \to \mathbb{R}^n$ and the output $y : [0,\infty)\times\Pi \to \mathbb{R}^{n_{\mathrm{out}}}$ depend on the parameters. Note that we have restricted the parametric problem to a single QoI of the system meaning that $n_{\mathrm{out}} = 1$. The sensitivity analysis of Section 4 can also be applied to the case of multiple outputs by investigating each output component separately. Furthermore, we allow the initial condition $E(p)x(0,p) = E(p)x_0(p)$ to be parameter-dependent with a given initial function $x_0 : \Pi \to \mathbb{R}^n$, which is assumed to be consistent for all $p \in \Pi$. Again, by an exponentially decaying shift, we can always attain the homogeneous initial condition $E(p)x(0,p) = 0$. Therefore, this condition will be assumed in what follows.

We assume that the DAE system (10) has the same index and that it is asymptotically stable for all $p \in \Pi$. We also assume that the system matrices in (10) depend continuously on the parameters. In this case, the state x inherits the continuity, due to a classical result of continuous dependence of dynamical system solution maps on parameters. This in turn implies that the QoI y depends also continuously on the parameters.

3.1. Stochastic modelling

In order to ascertain the effect of parametric variation on the output, we model our parameters as mutually independent random variables. Thus, with a complete probability space $(\Omega, \mathcal{F}, P)$, the random variables are $p : \Omega \to \Pi$,

i.e., $\omega \mapsto p(\omega)$. Let a joint probability density function $\rho : \Pi \to \mathbb{R}_{\geq 0}$ be given. If a measurable function $f : \Pi \to \mathbb{R}$ depends on the parameters, then its expected value is

$$\mathbb{E}[f] = \int_{\Pi} f(p)\,\rho(p)\,\mathrm{d}p. \tag{11}$$

The Hilbert space

$$\mathcal{L}^2(\Pi) = \left\{ f : \Pi \to \mathbb{R} \;:\; f \text{ is measurable and } \mathbb{E}[f^2] < \infty \right\} \tag{12}$$

is equipped with the inner product

$$\langle f, g \rangle = \mathbb{E}[fg] = \int_{\Pi} f(p)g(p)\,\rho(p)\,\mathrm{d}p, \qquad f, g \in \mathcal{L}^2(\Pi). \tag{13}$$

The induced norm is given by $\|f\|_{\mathcal{L}^2(\Pi)} = \sqrt{\langle f, f \rangle}$. The stochastic modelling can also be generalised to the case of random variables without densities (e.g., discrete random variables). In such cases, one replaces $\rho(p)\mathrm{d}p$ with the appropriate measure differential. In general, any measure with finite moments can be employed.

We assume that there is an $\mathcal{L}^2(\Pi)$-complete orthonormal system $(\Phi_i)_{i\in\mathbb{N}}$ consisting of polynomials, i.e., for all $i, j \in \mathbb{N}$,

$$\langle \Phi_i, \Phi_j \rangle = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Conditions on ρ that ensure availability of such a system are discussed in [30]. The theory of the generalised polynomial chaos (PC) requires a family of orthonormal polynomials for each random distribution ρ, and most classical probability distributions have such a family, see [5]. The choice of polynomials is not necessary for the general framework we present below, but polynomials are commonly used when the state varies smoothly with respect to the random parameters p.

In our case, where we assume the random parameters $p = (p_1, \ldots, p_q)^\top$ to be independent, then ρ has tensorial structure,

$$\rho(p) = \prod_{\ell=1}^{q} \rho_\ell(p_\ell),$$

where ρ_ℓ is the univariate probability density function for the scalar random variable p_ℓ for each $\ell = 1, \ldots, q$. This allows an explicit construction of the multivariate orthonormal family $\{\Phi_i\}$ using univariate orthogonal polynomials. Let $\left\{\phi_k^{(\ell)}\right\}_{k=0}^{\infty}$ with $\mathrm{degree}(\phi_k^{(\ell)}) = k$ be a family of ρ_ℓ-orthonormal univariate polynomials, i.e.,

$$\int_{\mathbb{R}} \phi_i^{(\ell)}(p_\ell)\phi_k^{(\ell)}(p_\ell)\rho_\ell(p_\ell)\,\mathrm{d}p_\ell = \begin{cases} 1 & \text{for } i = k, \\ 0 & \text{for } i \neq k. \end{cases}$$

Finally, with $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $\beta : \mathbb{N} \to \mathbb{N}_0^q$ denote an enumeration that exhausts $\mathbb{N}_0^q$, so that $\beta(i)$ is a q-dimensional multi-index. Then, the multivariate polynomials can be defined as products

$$\Phi_i(p_1, p_2, \ldots, p_q) = \phi_{\beta_1}^{(1)}(p_1)\phi_{\beta_2}^{(2)}(p_2)\cdots\phi_{\beta_q}^{(q)}(p_q), \tag{14}$$

where $(\beta_1, \ldots, \beta_q) = \beta(i)$. The total degree of the tensorial multivariate polynomial Φ_i is $|\beta(i)|$, i.e., the sum of the degrees of the component univariate polynomials. Multivariate polynomials of a low degree but with large number q of random variables include many constant univariate polynomials in the factorisation (14). Other types of basis functions can also be used in the stochastic approach. However, it is computationally advantageous to use tensorial structures (14) with many constant factors. Let us choose the multi-index ordering β so that polynomials are partially ordered by degree, i.e., polynomials are sorted such that $|\beta(i)| = \mathrm{degree}(\Phi_i) \leq \mathrm{degree}(\Phi_{i+1}) = |\beta(i+1)|$ is satisfied for all $i \geq 1$. Furthermore, assuming positive signs for leading coefficients, it follows that $\Phi_1 \equiv 1$.

Consequently, we expand the state and the output of the DAE system (10) into

$$x(t, p) = \sum_{i=1}^{\infty} v_i(t)\Phi_i(p), \qquad y(t, p) = \sum_{i=1}^{\infty} w_i(t)\Phi_i(p) \tag{15}$$

with coefficient functions $v_i : [0, \infty) \to \mathbb{R}^n$ and $w_i : [0, \infty) \to \mathbb{R}$, respectively. Under mild conditions on ρ, the series (15) converge in the norm of the Hilbert space (12) and pointwise in time, cf. [30]. Since Φ_1 denotes the constant function, the *total variance* of the random QoI has the form

$$V(t) = \mathbb{E}\left[(y(t, \cdot) - \mathbb{E}[y(t, \cdot)])^2\right] = \sum_{i=2}^{\infty} w_i(t)^2 \tag{16}$$

for each $t \geq 0$.

3.2. Stochastic Galerkin method

In a numerical approximation, the infinite sums in (15) must be truncated. We accomplish this by including all polynomials up to a total degree d described by the set of integers

$$\mathcal{I}^d = \{i \in \mathbb{N} \,:\, |\beta(i)| = \text{degree}(\Phi_i) \leq d\}\,. \tag{17}$$

The cardinality of this set is given by $|\mathcal{I}^d| = \frac{(d+q)!}{d!q!} =: m$, see [5, p. 65]. Then the infinite summations in (15) are approximated by

$$x^{(d)}(t,p) = \sum_{i\in\mathcal{I}^d} \hat{v}_i(t)\Phi_i(p), \qquad y^{(d)}(t,p) = \sum_{i\in\mathcal{I}^d} \hat{w}_i(t)\Phi_i(p), \tag{18}$$

respectively, where the coefficient functions $\hat{v}_i$ and $\hat{w}_i$ can be determined by the stochastic Galerkin approach. Inserting the approximations (18) into the DAE system (10) yields the residuals

$$\begin{aligned} R_x(p) &= E(p)\dot{x}^{(d)}(t,p) - A(p)x^{(d)}(t,p) - B(p)u(t),\\ R_y(p) &= y^{(d)}(t,p) - c(p)^\top x^{(d)}(t,p). \end{aligned}$$

The Galerkin approach requires that each component of the residuals is orthogonal to the space spanned by $\{\Phi_i : i \in \mathcal{I}^d\}$ with respect to the inner product (13). Introducing the parameter-dependent arrays

$$\Psi(p) = [\Phi_i(p)\Phi_j(p)]_{i,j=1}^{m,m} \in \mathbb{R}^{m\times m}, \qquad \psi(p) = [\Phi_i(p)]_{i=1}^{m} \in \mathbb{R}^m,$$

we obtain the larger deterministic system of DAEs

$$\begin{aligned} &\hat{E}\dot{\hat{v}}(t) = \hat{A}\hat{v}(t) + \hat{B}u(t), \quad \hat{E}\hat{v}(0) = 0\\ &\hat{w}(t) = \hat{C}\hat{v}(t), \end{aligned} \tag{19}$$

where $\hat{v}(t) = (\hat{v}_1(t)^\top, \ldots, \hat{v}_m(t)^\top)^\top \in \mathbb{R}^{mn}$, $\hat{w}(t) = (\hat{w}_1(t), \ldots, \hat{w}_m(t))^\top \in \mathbb{R}^m$, and the matrices $\hat{E}, \hat{A} \in \mathbb{R}^{mn\times mn}$, $\hat{B} \in \mathbb{R}^{mn\times n_{\text{in}}}$, and $\hat{C} \in \mathbb{R}^{m\times mn}$ are given by

$$\hat{E} = \mathbb{E}[\Psi \otimes E], \quad \hat{A} = \mathbb{E}[\Psi \otimes A], \quad \hat{B} = \mathbb{E}[\psi \otimes B], \quad \hat{C} = \mathbb{E}[\Psi \otimes c^\top]. \tag{20}$$

Here, the symbol $\otimes$ denotes the Kronecker product, and the probabilistic integration (11) is applied componentwise. The system (19) has the state $\hat{v}$, the same input u as in (10), and the output $\hat{w}$. We use homogeneous initial conditions in (19).

If the components of E, A, B and c in (10) are polynomials in the random variables, then the system matrices (20) can be calculated analytically for traditional probability distributions. This property is an advantage in comparison to stochastic collocation techniques, where quadrature errors or sampling errors emerge. In the general case, we cannot, unfortunately, avoid the computation of the Kronecker products in (20). However, if the system matrices in (10) depend affinely on parameters, i.e.,

$$\begin{aligned} E(p) &= E_0 + \sum_{k=1}^{N_E} \eta_k^E(p)E_k, \qquad A(p) = A_0 + \sum_{k=1}^{N_A} \eta_k^A(p)A_k,\\ B(p) &= B_0 + \sum_{k=1}^{N_B} \eta_k^B(p)B_k, \qquad c(p) = c_0 + \sum_{k=1}^{N_c} \eta_k^c(p)c_k, \end{aligned} \tag{21}$$

where $\eta_k^E, \eta_k^A, \eta_k^B, \eta_k^c : \Pi \to \mathbb{R}$ are scalar (nonlinear) coefficient functions, and E_j, A_j, B_j and c_j are parameter-independent. Consequently, the system matrices in (20) can be computed as

$$\begin{aligned} \hat{E} &= (I_m \otimes E_0) + \sum_{k=1}^{N_E} \hat{\Psi}_k^E \otimes E_k, \qquad \hat{A} = (I_m \otimes A_0) + \sum_{k=1}^{N_A} \hat{\Psi}_k^A \otimes A_k,\\ \hat{B} &= (e_1 \otimes B_0) + \sum_{k=1}^{N_B} \hat{\psi}_k^B \otimes B_k, \qquad \hat{C} = (I_m \otimes c_0^\top) + \sum_{k=1}^{N_c} \hat{\Psi}_k^c \otimes c_k^\top, \end{aligned} \tag{22}$$

with the identity matrix $I_m \in \mathbb{R}^{m\times m}$, the unit vector $e_1 \in \mathbb{R}^m$, and the parameter-independent matrices

$$\hat{\Psi}_k^E = \mathbb{E}[\eta_k^E \Psi], \qquad \hat{\Psi}_k^A = \mathbb{E}[\eta_k^A \Psi], \qquad \hat{\psi}_k^B = \mathbb{E}[\eta_k^B \psi], \qquad \hat{\Psi}_k^c = \mathbb{E}[\eta_k^c \Psi].$$

This significantly reduces the computational complexity provided that N_E, N_A, N_B, and N_c are much smaller than the problem dimension n. Note that the full matrices in (20) need not be explicitly formed, and efficient algorithmic implementations would leverage the Kronecker product structure shown above to accelerate computations. In particular, an affine-linear dependence on the parameters is advantageous, where the matrices (21) have the form

$$E(p) = E_0 + \sum_{k=1}^{q} p_k E_k, \qquad A(p) = A_0 + \sum_{k=1}^{q} p_k A_k, \tag{23}$$

etc. Typically, some of the matrices A_k, E_k are zero. The associated matrices $\hat{\Psi}_k^E = \hat{\Psi}_k^A = \mathbb{E}[p_k \Psi]$ are sparse due to the orthogonality of the basis polynomials. The matrices $\hat{A}, \hat{E}$ of the stochastic Galerkin system (19) inherit any sparsity of A_k, E_k and $\hat{\Psi}_k^A, \hat{\Psi}_k^E$ through Kronecker products. Further details on the stochastic Galerkin method for linear dynamical systems can be found in [11–13,31].

We remark that PC in combination with the stochastic Galerkin approach is applicable to systems with a moderate number q of random variables (say, up to one hundred). Problems with a very large number of random parameters result in extremely large systems. Such systems require more customised approaches, but have been successfully treated, for example, in [9] and in [32] for stationary problems.

3.3. DAE aspects of stochastic Galerkin system

Note that the stochastic Galerkin system (19) may be unstable even if the parametric DAE system (10) is asymptotically stable for all $p \in \Pi$, see [33]. Likewise, the matrix $\hat{E}$ in (19) may be non-singular even if $E(p)$ in (10) is singular for all $p \in \Pi$. However, such models are rather seldom in practice. Thus, we henceforth assume that the stochastic Galerkin system (19) is asymptotically stable and it has the same index ν as the DAE system (10). Sufficient conditions ensuring the latter are given in [12,13].

Furthermore, if the DAE system (10) is semi-explicit meaning that $E(p)$ exhibits a partitioning

$$E(p) = \begin{bmatrix} E_{11}(p) & 0 \\ 0 & 0 \end{bmatrix}$$

with a non-singular matrix $E_{11}(p) \in \mathbb{R}^{k \times k}$ for all $p \in \Pi$, then a permutation of equations and unknowns in the stochastic Galerkin system (19) yields an equivalent DAE system with the matrix

$$\hat{E} = \begin{bmatrix} \hat{E}_{11} & 0 \\ 0 & 0 \end{bmatrix}, \tag{24}$$

where $\hat{E}_{11} = \mathbb{E}[\Psi \otimes E_{11}]$. If $E_{11}(p)$ is symmetric and positive definite for all $p \in \Pi$, then the matrix $\hat{E}_{11}$ is also symmetric and positive definite, see [34, Lemma 1]. In this case, the stochastic Galerkin system (19) becomes semi-explicit.

4. Sensitivity analysis

Next, we present a global variance-based sensitivity analysis which is used to measure the influence of the individual random parameters on the QoI.

4.1. Partial variances and total effect sensitivity indices

The Sobol indices provide a set of non-negative real numbers which describe the interaction of each subset of random variables, see [7,8]. Thus, the number of Sobol indices is equal to the number of non-empty subsets of $\{1, \ldots, q\}$, i.e., $2^q - 1$. The Sobol indices yield the total effect sensitivity coefficients which represent a variance-based sensitivity measure.

We employ an equivalent definition of the total effect sensitivity indices, where the PC expansion of a stochastic function, cf. [35], is used. The random output of the DAE system (10) exhibits the PC expansion (15). Let

$$\mathcal{I}_j = \left\{ i \in \mathbb{N} \ : \ \Phi_i \text{ is non-constant in } p_j \right\} \tag{25}$$

be a set of integers for each component p_j, $j = 1, \ldots, q$, of the random vector p. Since Φ_1 is the constant polynomial (degree zero), it holds that $1 \notin \mathcal{I}_j$ for all j. Then the partial variance associated with the jth random variable can be defined via

$$V_j(t) = \sum_{i \in \mathcal{I}_j} w_i(t)^2, \qquad j = 1, \ldots, q, \tag{26}$$

for all $t \geq 0$. It follows from (16) and (26) that $0 \leq V_j(t) \leq V(t)$ for $j = 1, \ldots, q$. We assume that the total variance (16) is positive for all $t > 0$. Then the *total effect sensitivity indices* are defined as

$$S_j(t) = \frac{V_j(t)}{V(t)}, \qquad j = 1, \ldots, q. \tag{27}$$

It holds that $0 \leq S_j(t) \leq 1$ for all $j = 1, \ldots, q$ and all $t > 0$, hence,

$$1 \leq S_1(t) + \cdots + S_q(t) \leq q$$

pointwise in time. Often, the sum of the sensitivity indices is close to the lower bound one.

Restricting the index set $\mathcal{I}_j$ to polynomials up to a total degree d, we obtain the intersection

$$\mathcal{I}_j^d = \mathcal{I}_j \cap \mathcal{I}^d, \qquad j = 1, \ldots, q, \tag{28}$$

cf. (17). Summation over this finite index set generates approximations of the partial variances (26).

4.2. Stochastic Galerkin system with quadratic outputs

The stochastic Galerkin system (19) yields approximations to the sensitivity indices (27) given by

$$\hat{S}_j(t) = \frac{\hat{V}_j(t)}{\hat{V}(t)}, \qquad j = 1, \ldots, q, \tag{29}$$

with

$$\hat{V}_j(t) = \sum_{i \in \mathcal{I}_j^d} \hat{w}_i(t)^2, \qquad \hat{V}(t) = \sum_{i=2}^{m} \hat{w}_i(t)^2 \tag{30}$$

including the index set (28). The following theorem demonstrates that the approximations (30) inherit the convergence of the QoI in the stochastic Galerkin method. The proof is given in the Appendix.

Theorem 1. *Let $y(t,\cdot) \in \mathcal{L}^2(\Pi)$ for all $t \geq 0$. If the stochastic Galerkin method is convergent in the sense*

$$\lim_{d \to \infty} \left\| y(t,\cdot) - \sum_{i \in \mathcal{I}^d} \hat{w}_i(t) \Phi_i(\cdot) \right\|_{\mathcal{L}^2(\Pi)} = 0 \tag{31}$$

for all $t \geq 0$, then the approximations of the total variance and the partial variances in (30) also converge, i.e.,

$$\lim_{d \to \infty} \left| V(t) - \hat{V}(t) \right| = 0,$$
$$\lim_{d \to \infty} \left| V_j(t) - \hat{V}_j(t) \right| = 0, \qquad j = 1, \ldots, q,$$

pointwise for all $t \geq 0$.

The sensitivity indices (27) are time-dependent functions, which are determined by the choice of the input signals u. Alternatively, we can construct a single number as a sensitivity measure for each random variable.

Let $\hat{w}^{(j)}$ be the vector-valued function consisting of the components of $\hat{w}$ in $\mathcal{I}_j^d$ from (28). Then it follows from (30) that

$$\hat{V}_j(t) = \hat{w}^{(j)}(t)^\top \hat{w}^{(j)}(t), \qquad j = 1, \ldots, q. \tag{32}$$

We now construct an output matrix associated with this partial variance. To this end, we consider a diagonal matrix

$$D_j = \operatorname{diag}\left(d_1^{(j)}, \ldots, d_m^{(j)}\right) \quad \text{with} \quad d_i^{(j)} = \begin{cases} 1 & \text{if } i \in \mathcal{I}_j^d \\ 0 & \text{if } i \notin \mathcal{I}_j^d \end{cases}$$

for each $j = 1, \ldots, q$. Then the matrix $\hat{C}_j' = D_j \hat{C}$ contains rows that are zero. It holds that $\operatorname{rank}(\hat{C}_j') \leq m_j = |\mathcal{I}_j^d|$ due to $\operatorname{rank}(D_j) = m_j$. We remove the zero rows from $\hat{C}_j'$ and obtain a condensed matrix $\hat{C}_j \in \mathbb{R}^{m_j \times mn}$. Consequently, (32) implies that

$$\hat{V}_j(t) = \hat{w}^{(j)}(t)^\top \hat{w}^{(j)}(t) = \hat{v}(t)^\top \hat{C}_j'^\top \hat{C}_j' \hat{v}(t) = \hat{v}(t)^\top \hat{C}_j^\top \hat{C}_j \hat{v}(t). \tag{33}$$

The partial variance of the jth random variable can be obtained by a stochastic Galerkin system

$$\begin{aligned} \hat{E}\dot{\hat{v}}(t) &= \hat{A}\hat{v}(t) + \hat{B}u(t), \quad \hat{E}\hat{v}(0) = 0, \\ \hat{V}_j(t) &= \hat{v}(t)^\top \hat{C}_j^\top \hat{C}_j \hat{v}(t), \end{aligned} \tag{34}$$

where a single quadratic output is defined by the symmetric, positive semi-definite matrix $\hat{C}_j^\top \hat{C}_j$ for $j = 1, \ldots, q$. Similarly, the total variance is given as the single quadratic output

$$\hat{V}(t) = \hat{v}(t)^\top \breve{C}^\top \breve{C} \hat{v}(t)$$

from the state in the stochastic Galerkin system (19), where $\breve{C} \in \mathbb{R}^{(m-1)\times mn}$ denotes the matrix $\hat{C}$ without its first row assuming that Φ_1 is the unique constant basis polynomial.

The systems (34) represent linear DAEs with a (single) quadratic output. In [36], MOR of such systems was performed by a transformation to a quadratic-bilinear system in the case of ODEs. However, this transformation is not always feasible in the case of DAEs. Furthermore, in this case, q systems have to be reduced. In [37], a model reduction approach for a control system with a quadratic output specified by a symmetric positive semi-definite matrix was proposed which requires a symmetric decomposition of this matrix like the (pivoted) Cholesky factorisation. In our application, we already possess such a decomposition. Instead of (34), we consider the alternative stochastic Galerkin systems

$$\begin{aligned} \hat{E}\dot{\hat{v}}(t) &= \hat{A}\hat{v}(t) + \hat{B}u(t), \quad \hat{E}\hat{v}(0) = 0, \\ \hat{z}_j(t) &= \hat{C}_j \hat{v}(t) \end{aligned} \tag{35}$$

for $j = 1, \ldots, q$ with multiple outputs. Then the quadratic output of system (34) can be determined as $\hat{V}_j(t) = \hat{z}_j(t)^\top \hat{z}_j(t)$. In Section 5, we will show that reduced-order models of the q systems (35) are obtained by the reduction of a single linear dynamical system.

4.3. System norms and sensitivity measures

Similarly to the ODE case [14], we use the $\mathcal{H}_\infty$-norms of the stochastic Galerkin systems (35) to quantify the sensitivity of the random QoI in the DAE system (10) with single output. In contrast, the $\mathcal{H}_2$-norms and $\mathcal{H}_\infty$-norms of the separate components of the transfer function of (19) were analysed in [38] to specify a sparse representation of the random QoI.

We define the sensitivity measures as follows.

Definition 6. Consider the stochastic Galerkin systems (35) with the transfer functions $\hat{H}_j(s) = \hat{C}_j(s\hat{E} - \hat{A})^{-1}\hat{B}$ for $j = 1, \ldots, q$. Then the *sensitivity coefficients* are defined as

$$\hat{\theta}_j = \|\hat{H}_j\|_{\mathcal{H}_\infty}, \qquad j = 1, \ldots, q. \tag{36}$$

The following lemma provides a relationship between the $\mathcal{H}_\infty$-norms of the transfer functions of the Galerkin system (19) and the modified Galerkin systems (35).

Lemma 1. *Let $\hat{H}$ and $\hat{H}_1, \ldots, \hat{H}_q$ be the transfer functions of the stochastic Galerkin systems (19) and (35), respectively. Then it holds that*

$$\|\hat{H}_j\|_{\mathcal{H}_\infty} \leq \|\hat{H}\|_{\mathcal{H}_\infty}, \qquad j = 1, \ldots, q.$$

Furthermore, the sensitivity measures from Definition 6 can be used to bound the $\mathcal{L}^1$-norm of the partial variances in the time domain.

Lemma 2. *The stochastic Galerkin systems (35) with the initial condition $\hat{E}\hat{v}(0) = 0$ yield approximations to the partial variances satisfying the bounds*

$$\|\hat{V}_j\|_{\mathcal{L}^1([0,\infty))} \leq \hat{\theta}_j^2 \, \|u\|^2_{\mathcal{L}^2([0,\infty))}, \qquad j = 1, \ldots, q.$$

Lemmas 1 and 2 can be proved analogously to the ODE case in [14]. Lemma 2 demonstrates that parameters with small sensitivity coefficients (36) have partial variances close to zero in the time domain. Thus, their variations can be neglected, i.e., the parameters are remodelled as constants. This approach is called screening, see [39], or freezing of insignificant random variables, see [7].

The definition of the total effect sensitivity indices (27) suggest to arrange the relative sensitivity measures

$$\hat{\theta}_j^{\mathrm{rel}} = \frac{\|\hat{H}_j\|_{\mathcal{H}_\infty}}{\|\hat{H}\|_{\mathcal{H}_\infty}}, \qquad j = 1, \ldots, q. \tag{37}$$

Then Lemma 1 implies $0 \leq \hat{\theta}_j^{\mathrm{rel}} \leq 1$. However, since the denominators in (37) are identical for $j = 1, \ldots, q$, we consider just the normalised measures

$$\hat{\theta}_j^{*} = \frac{\hat{\theta}_j}{\hat{\theta}_1 + \cdots + \hat{\theta}_q}. \tag{38}$$

These quantities exhibit the property $0 \leq \hat{\theta}_j^{*} \leq 1$ for $j = 1, \ldots, q$ as well.

5. Model order reduction

The sensitivity measures (36) require the computation of the $\mathcal{H}_\infty$-norms of the large DAE systems (35) obtained by the stochastic Galerkin method. As mentioned in Section 2.2, computing such norms involves solving generalised eigenvalue problems of high dimension and often is a huge computational burden. To reduce the computational effort, we apply model reduction to the DAE systems (35) and use the resulting reduced-order models to determine accurate approximations to the sensitivity measures (36). Note that systems (35) have the same state equation as the stochastic Galerkin system (19), and the outputs of (35) are just selected components of the output of (19). Therefore, we propose first to approximate system (19) by a reduced-order model

$$\begin{aligned} \tilde{E}\dot{\tilde{v}}(t) &= \tilde{A}\tilde{v}(t) + \tilde{B}u(t), \\ \tilde{z}(t) &= \tilde{C}\tilde{v}(t), \end{aligned} \tag{39}$$

with $\tilde{E}, \tilde{A} \in \mathbb{R}^{r\times r}$, $\tilde{B} \in \mathbb{R}^{r\times n_{\mathrm{in}}}$, and $\tilde{C} \in \mathbb{R}^{m\times r}$ with $r \ll nm$. Such a model can be computed by projection

$$\tilde{A} = T_{\mathrm{l}}^\top \hat{A}\, T_{\mathrm{r}}, \qquad \tilde{B} = T_{\mathrm{l}}^\top \hat{B}, \qquad \tilde{C} = \hat{C}\, T_{\mathrm{r}}, \qquad \tilde{E} = T_{\mathrm{l}}^\top \hat{E}\, T_{\mathrm{r}} \tag{40}$$

with the projection matrices $T_{\mathrm{l}}, T_{\mathrm{r}} \in \mathbb{R}^{mn\times r}$. After these projections are identified, we use them for model reduction of systems (35) for all $j = 1, \ldots, q$.

Remark 1. Since the partial variances (26) do not depend on the expected value, the first row of the output matrix $\hat{C}$ in (19) can be removed. However, we keep the complete matrix $\hat{C}$ in model reduction, because the expected value is an important quantity often desired in statistics.

5.1. Balanced truncation

For model reduction of the stochastic Galerkin system (19), we employ a balanced truncation method. This method was first developed for ODE systems [40,41] and then extended to DAEs in [23]. Since the DAE system (19) has many outputs and only a few inputs, we use the balanced truncation approach developed in [22,42] for systems with many inputs or outputs. Here, we present a modification of this approach which requires only one eigenvalue decomposition compared to two eigenvalue decompositions and one singular value decomposition for the method in [22].

Let the stochastic Galerkin system (19) be asymptotically stable. Furthermore, for simplicity, we assume that (19) has a strictly proper transfer function $\hat{H}(s)$. An extension to general systems with possibly improper transfer functions can be done in a manner similar to [22,23]. First, we define a *proper controllability Gramian* $\hat{G}_{\mathrm{c}}$ and a *proper observability Gramian* $\hat{G}_{\mathrm{o}}$ as unique symmetric, positive semi-definite solutions of the projected Lyapunov equations

$$\hat{E}\hat{G}_{\mathrm{c}}\hat{A}^{\top} + \hat{A}\hat{G}_{\mathrm{c}}\hat{E}^{\top} + \hat{P}_{\mathrm{l}}\hat{B}\hat{B}^{\top}\hat{P}_{\mathrm{l}}^{\top} = 0, \quad \hat{G}_{\mathrm{c}} = \hat{P}_{\mathrm{r}}\hat{G}_{\mathrm{c}}\hat{P}_{\mathrm{r}}^{\top}, \tag{41}$$

$$\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{A} + \hat{A}^{\top}\hat{G}_{\mathrm{o}}\hat{E} + \hat{P}_{\mathrm{r}}^{\top}\hat{C}^{\top}\hat{C}\hat{P}_{\mathrm{r}} = 0, \quad \hat{G}_{\mathrm{o}} = \hat{P}_{\mathrm{l}}^{\top}\hat{G}_{\mathrm{o}}\hat{P}_{\mathrm{l}}, \tag{42}$$

where $\hat{P}_{\mathrm{l}}$ and $\hat{P}_{\mathrm{r}}$ are the spectral projectors onto the left and right deflating subspaces of the matrix pencil $\lambda\hat{E} - \hat{A}$ corresponding to the finite eigenvalues. The Gramians can be used to characterise the input and output energy of the system (19) as well as its controllability and observability properties, see [23] for more details.

One can show that the matrix $\hat{G}_{\mathrm{c}}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}$ has non-negative eigenvalues $\lambda_i(\hat{G}_{\mathrm{c}}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E})$, see [23]. We use them to define *proper Hankel singular values*

$$\sigma_i = \sqrt{\lambda_i(\hat{G}_{\mathrm{c}}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E})}, \tag{43}$$

which quantify the contribution of the individual state variables to the energy transfer from the input to the state, and from the state to the output. The DAE system (19) can be transformed into a balanced form such that the Gramians of the transformed system become identical and diagonal matrices with the proper Hankel singular values on the diagonal. Then the state variables corresponding to the low proper Hankel singular values make a negligibly small contribution to the energy transition. Therefore, they can be removed from the system without substantially changing its input–output relationship. This identifies the projection matrices T_{l} and T_{r} defining the reduced-order model in (40). We next describe how these matrices are computed in practice.

5.2. Computation of a reduced-order model

A reduced-order model (39), (40) via balanced truncation can be computed as follows. Using the full-rank Cholesky factorisation of the controllability Gramian $\hat{G}_{\mathrm{c}} = Z_{\mathrm{c}}Z_{\mathrm{c}}^{\top}$ with a full-rank matrix $Z_{\mathrm{c}} \in \mathbb{R}^{mn\times r_{\mathrm{c}}}$, we obtain

$$\sigma_i = \sqrt{\lambda_i(\hat{G}_{\mathrm{c}}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E})} = \sqrt{\lambda_i(Z_{\mathrm{c}}Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E})} = \sqrt{\lambda_i(Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}})}.$$

Then the projection matrices T_{l} and T_{r} can be determined using the dominant subspaces of the symmetric, positive semi-definite matrix $Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}}$ as follows. Let

$$Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}} = [\,W_1,\ W_0\,]\begin{bmatrix}\Lambda_1 & \\ & \Lambda_0\end{bmatrix}[\,W_1,\ W_0\,]^{\top} \tag{44}$$

be the associated eigenvalue decomposition, where $[\,W_1,\ W_0\,]$ is orthogonal,

$$\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_r) \quad \text{and} \quad \Lambda_0 = \mathrm{diag}(\lambda_{r+1}, \ldots, \lambda_{r_{\mathrm{c}}})$$

with decreasingly ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} \geq \cdots \geq \lambda_{r_{\mathrm{c}}}$. Then we determine the projection matrices in (40) as

$$T_{\mathrm{l}} = \hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}}W_1\Lambda_1^{-\frac{3}{4}}, \qquad T_{\mathrm{r}} = Z_{\mathrm{c}}W_1\Lambda_1^{-\frac{1}{4}}. \tag{45}$$

The resulting reduced-order model (39), (40) has a non-singular matrix $\tilde{E}$. It is asymptotically stable and balanced. Indeed, taking into account that the left inverses of T_{l} and T_{r} are given by

$$T_{\mathrm{l}}^{-} = \Lambda_1^{-\frac{1}{4}}W_1^{\top}Z_{\mathrm{c}}^{\top}\hat{E}^{\top}, \qquad T_{\mathrm{r}}^{-} = \Lambda_1^{-\frac{3}{4}}W_1^{\top}Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E},$$

respectively, the Gramians of the reduced-order model (39) have the form

$$T_{\mathrm{r}}^{-}\hat{G}_{\mathrm{c}}(T_{\mathrm{r}}^{-})^{\top} = \Lambda_1^{-\frac{3}{4}}(W_1^{\top}Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}})(Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}}W_1)\Lambda_1^{-\frac{3}{4}} = \Lambda_1^{-\frac{3}{4}}\Lambda_1W_1^{\top}W_1\Lambda_1\Lambda_1^{-\frac{3}{4}} = \Lambda_1^{\frac{1}{2}},$$

$$T_{\mathrm{l}}^{-}\hat{G}_{\mathrm{o}}(T_{\mathrm{l}}^{-})^{\top} = \Lambda_1^{-\frac{1}{4}}(W_1^{\top}Z_{\mathrm{c}}^{\top}\hat{E}^{\top}\hat{G}_{\mathrm{o}}\hat{E}Z_{\mathrm{c}}W_1)\Lambda_1^{-\frac{1}{4}} = \Lambda_1^{\frac{1}{2}}.$$

Moreover, its transfer function $\tilde{H}(s) = \tilde{C}(s\tilde{E} - \tilde{A})^{-1}\tilde{B}$ satisfies the following error bound

$$\left\|\hat{H} - \tilde{H}\right\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=r+1}^{r_c} \sigma_i \tag{46}$$

with the proper Hankel singular values $\sigma_i = \sqrt{\lambda_i}$, see [23] for details.

The most expensive steps in the presented balanced truncation method consist of solving the projected Lyapunov Eqs. (41) and (42). Since the matrix $\hat{B}$ often has only a few columns ($n_{in} \ll nm$), the solution of (41) can usually be well approximated by a low-rank matrix $\hat{G}_c \approx \tilde{Z}_c\tilde{Z}_c^\top$ with $\tilde{Z}_c \in \mathbb{R}^{mn\times\tilde{r}_c}$, $\tilde{r}_c \ll r_c$. The low-rank alternating directions implicit (ADI) method [43] or the extended Krylov subspace method [44] yield such a low-rank approximation. In these methods, the spectral projectors $\hat{P}_r$ and $\hat{P}_l$ are required. For several structured DAE models resulting in circuit simulation, computational fluid dynamics, and constrained mechanical systems, these projectors can be constructed by exploiting the system structure, see [43]. However, for unstructured problems, the projected Lyapunov equation with only small to medium-sized matrices can currently be solved using the Schur–Hammarling method [45].

Next, we briefly discuss how to compute the product matrix $\hat{G}_o\hat{E}Z_c$ required in (44) and (45) with Z_c replaced by $\tilde{Z}_c$. Thus, we proceed to describe the procedure for computing $\hat{G}_o\hat{E}\tilde{Z}_c$. First of all note that the solution of the projected Lyapunov equation (42) can be represented as

$$\hat{G}_o = \frac{1}{2\pi}\int_{-\infty}^{\infty}(-\mathrm{i}\omega\hat{E}^\top - \hat{A}^\top)^{-1}\hat{P}_r^\top\hat{C}^\top\hat{C}\hat{P}_r(\mathrm{i}\omega\hat{E} - \hat{A})^{-1}\,\mathrm{d}\omega.$$

Then we have

$$\hat{G}_o\hat{E}\tilde{Z}_c = \frac{1}{2\pi}\int_{-\infty}^{\infty}(-\mathrm{i}\omega\hat{E}^\top - \hat{A}^\top)^{-1}\hat{P}_r^\top\hat{C}^\top\hat{C}\hat{P}_r(\mathrm{i}\omega\hat{E} - \hat{A})^{-1}\hat{E}\tilde{Z}_c\,\mathrm{d}\omega. \tag{47}$$

This integral can be approximated by a suitable quadrature rule

$$\hat{G}_o\hat{E}\tilde{Z}_c \approx \sum_{j=1}^{\ell}\alpha_j F(\mathrm{i}\omega_j) \tag{48}$$

with the weights α_j, quadrature nodes ω_j, and

$$F(\mathrm{i}\omega_j) = (-\mathrm{i}\omega_j\hat{E}^\top - \hat{A}^\top)^{-1}\hat{P}_r^\top\hat{C}^\top\hat{C}\hat{P}_r(\mathrm{i}\omega_j\hat{E} - \hat{A})^{-1}\hat{E}\tilde{Z}_c \in \mathbb{R}^{mn\times\tilde{r}_c}.$$

The matrix-valued function F can be evaluated efficiently by solving the linear systems

$$(\mathrm{i}\omega_j\hat{E} - \hat{A})Y = \hat{E}\tilde{Z}_c \qquad \text{and} \qquad (-\mathrm{i}\omega_j\hat{E}^\top - \hat{A}^\top)X = \hat{P}_r^\top\hat{C}^\top\hat{C}\hat{P}_rY,$$

if $\tilde{Z}_c$ has a small number of columns. Just one LU decomposition is required to solve the two linear systems for fixed j. Taking into account that

$$\hat{G}_o\hat{E}\tilde{Z}_c = \frac{1}{2\pi}\int_0^{\infty}\left(F(\mathrm{i}\omega) + \overline{F(\mathrm{i}\omega)}\right)\mathrm{d}\omega,$$

the computation of the integral can slightly be simplified.

Remark 2. Note that instead of (48) we can approximate the symmetric matrix $\tilde{Z}_c^\top\hat{E}^\top\hat{G}_o\hat{E}\tilde{Z}_c$ by the quadrature rule as

$$\tilde{Z}_c^\top\hat{E}^\top\hat{G}_o\hat{E}\tilde{Z}_c \approx \sum_{j=1}^{\ell}\alpha_j\left(\hat{C}\hat{P}_r(\mathrm{i}\omega_j\hat{E} - \hat{A})^{-1}\hat{E}\tilde{Z}_c\right)^*\left(\hat{C}\hat{P}_r(\mathrm{i}\omega_j\hat{E} - \hat{A})^{-1}\hat{E}\tilde{Z}_c\right),$$

where only the products $(\mathrm{i}\omega_j\hat{E} - \hat{A})^{-1}\hat{E}\tilde{Z}_c$ need to be computed. However, we still require $\hat{G}_o\hat{E}\tilde{Z}_c$ in (45). Therefore, we prefer to work with (47) from the beginning.

5.3. Error bound for sensitivity measures

For MOR of the Galerkin systems (35), we use the same projection matrices T_l and T_r as for (19) and compute the reduced-order systems with the downsized matrices $\tilde{E}$, $\tilde{A}$, $\tilde{B}$ as in (40) and

$$\tilde{C}_j = \hat{C}_jT_r, \qquad j = 1,\ldots,q.$$

These reduced systems inherit the asymptotic stability and their transfer functions inherit the properness of the full-order models (35). Then we define

$$\tilde{\theta}_j := \|\tilde{H}_j\|_{\mathcal{H}_\infty} \tag{49}$$

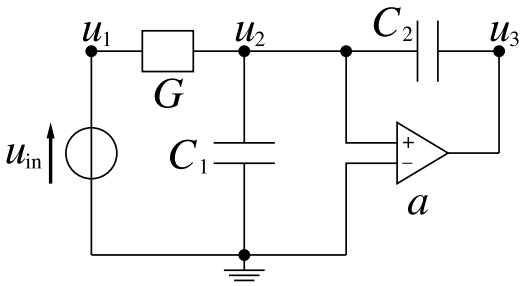

Fig. 1. Electric circuit of the Miller integrator.

with the transfer functions $\tilde{H}_j(s) = \tilde{C}_j(s\tilde{E} - \tilde{A})^{-1}\tilde{B}$ as approximations to the sensitivity coefficients $\hat{\theta}_j$ in (36). The following theorem establishes *a priori* error bounds for these approximations.

Theorem 2. *Consider the Galerkin systems (35) with the transfer functions $\hat{H}_j$ and the reduced-order systems with the transfer functions $\tilde{H}_j$. Let $\hat{\theta}_j$ and $\tilde{\theta}_j$ be the sensitivity measures as defined in (36) and (49), respectively. Then the approximation errors satisfy the bounds*

$$\left|\hat{\theta}_j - \tilde{\theta}_j\right| \leq 2 \sum_{i=r+1}^{r_c} \sigma_i, \qquad j = 1, \ldots, q, \tag{50}$$

where σ_i are the proper Hankel singular values of the stochastic Galerkin system (19).

Proof. Using the definitions of the sensitivity measures in (36), (49), and the reverse triangle inequality of the norms, we obtain

$$\left|\hat{\theta}_j - \tilde{\theta}_j\right| = \left|\|\hat{H}_j\|_{\mathcal{H}_\infty} - \|\tilde{H}_j\|_{\mathcal{H}_\infty}\right| \leq \left\|\hat{H}_j - \tilde{H}_j\right\|_{\mathcal{H}_\infty} \tag{51}$$

for $j = 1, \ldots, q$. The mapping $\hat{H} - \tilde{H}$ represents the transfer function of the (linear) difference system associated with (19), where the same input is supplied to the full-order model and the reduced-order model, while the output is the difference of their outputs. Likewise, the mappings $\hat{H}_j - \tilde{H}_j$ describe the difference systems associated with (35). The outputs of the system (35) are a subset of the outputs of the system (19) due to $\mathcal{I}_j^d \subset \mathcal{I}^d$ for each j. Consequently, the outputs of the difference system for (35) are a subset of the outputs of the difference system for (19). We obtain together with the bound (46) that

$$\left\|\hat{H}_j - \tilde{H}_j\right\|_{\mathcal{H}_\infty} \leq \left\|\hat{H} - \tilde{H}\right\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=r+1}^{r_c} \sigma_i$$

for $j = 1, \ldots, q$. Thus, the error bounds (50) hold. □

Theorem 2 demonstrates that the approximation of the sensitivity measures inherits the error bound of the model reduction by balanced truncation. In (51), the reverse triangle inequality produces pessimistic estimates in some cases. Yet, the derived MOR technique has the advantage that we deal with a single linear dynamical system and linear outputs, whereas the required q-tuples of $\mathcal{H}_\infty$-norms represent more complicated quantities of interest.

6. Illustrative examples

We now apply the sensitivity measures from Section 4 to two test examples.

6.1. Miller integrator

As the first example, we consider the electric circuit of the Miller integrator from [1]. Fig. 1 depicts the circuit diagram. The involved physical parameter vector $p = (C_1, C_2, G, a)^\top$ contains two capacitances, a conductance and an amplification factor. Modified nodal analysis (MNA), see [2], yields a DAE system (10) with the state $x = (u_1, u_2, u_3, \jmath_1, \jmath_2)^\top$, where u_k, $k = 1, 2, 3$, are the node voltages, and $\jmath_k$, $k = 1, 2$, are the branch currents. The involved matrices read as

$$A(p) = \begin{bmatrix} -G & G & 0 & -1 & 0 \\ G & -G & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -a & 1 & 0 & 0 \end{bmatrix}, \quad E(p) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & (C_1 + C_2) & -C_2 & 0 & 0 \\ 0 & -C_2 & C_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

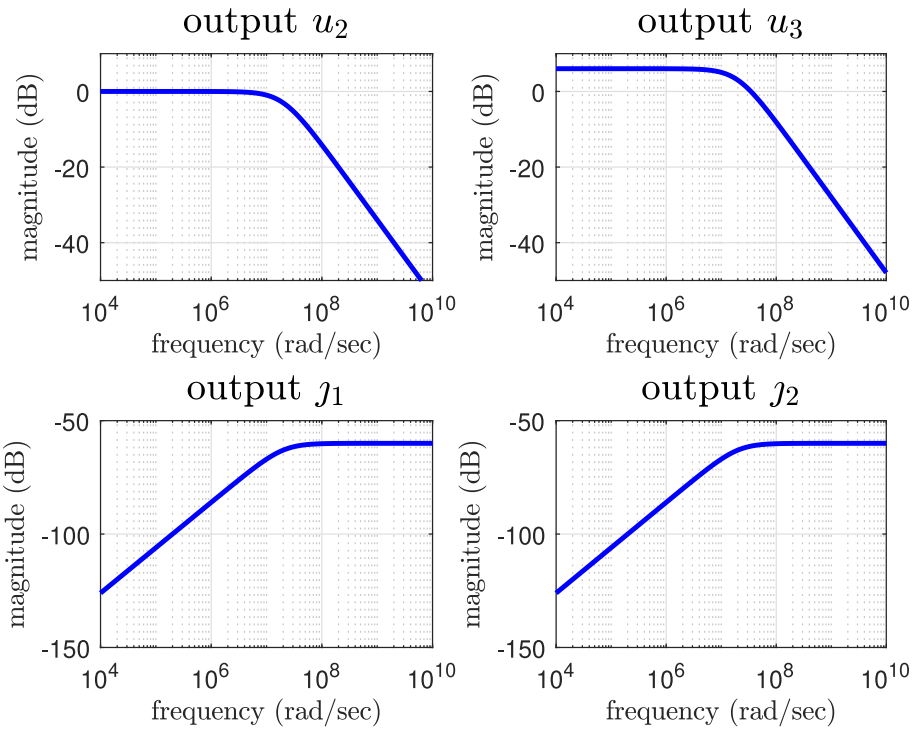

Fig. 2. Bode plots (magnitudes) of the Miller integrator for deterministic parameters and several outputs.

Table 1
$\mathcal{H}_\infty$-norms of the stochastic Galerkin systems (19), sums of the sensitivity measures $\hat{\theta}_j$, and individual sensitivities θ_j^* (in percentage of the sum) for random parameters $p = (p_1, p_2, p_3, p_4) = (C_1, C_2, G, a)$, given separate outputs in the Miller integrator.

Output	$\mathcal{H}_\infty$-norm	$\sum_{j=1}^4 \hat{\theta}_j$	θ_1^* (C_1)	θ_2^* (C_2)	θ_3^* (G)	θ_4^* (a)
u_2	1.0000	0.3873	32.5%	19.8%	14.0%	33.6%
u_3	2.0133	0.8746	30.9%	19.5%	12.4%	37.3%
y_1	0.0010	0.00045	28.1%	17.2%	25.6%	29.1%
y_2	0.0017	0.00090	71.6%	2.3%	20.2%	5.9%

and $B = [0, 0, 0, -1, 0]^\top$. The single input $u = u_{\text{in}}$ is the voltage source, and the single output is the node voltage u_3 which is approximately a multiple of the integral of the input voltage. Nevertheless, we investigate all four variables u_2, u_3, y_1, y_2 as single outputs separately. Fig. 2 shows the magnitude part of the Bode plots for the transfer function associated with each output in the case of constant parameters $C_1 = 10^{-10}$, $C_2 = 5 \cdot 10^{-11}$, $G = 0.001$, $a = 2$.

We take independent uniform probability distributions of the random parameters, which vary 20% around mean values given by the above constant selection. The index of the DAE system is two for all realisations of the random variables. In the PC expansion (15), the multivariate basis polynomials $(\Phi_i)_{i\in\mathbb{N}}$ are products of the univariate Legendre polynomials. We include all basis polynomials up to degree $d = 3$. Hence, the number of basis functions is $m = 35$. The stochastic Galerkin system (19) is a DAE system of dimension $mn = 175$.

Four stochastic Galerkin systems (19) and their associated systems (35) are arranged corresponding to the four separate QoIs. Since these systems are relatively small, we compute their $\mathcal{H}_\infty$-norms (7) without model reduction. Table 1 shows the $\mathcal{H}_\infty$-norms of systems (19) and the sums $\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 + \hat{\theta}_4$ of the sensitivity measures (36) for every separate input. The individual sensitivities with respect to the parameters indicated by percentages of this sum are also presented there. Fig. 3 illustrates the relative sensitivity measures (38). We recognise that the parameters C_1 and a are more important than C_2 and G in the output voltage u_3. If the current y_2 represents the QoI, then the parameter C_1 dominates significantly, while the influence of C_2 and a is negligible.

6.2. Band-pass filter

As the second example, we consider the electric circuit of a band-pass filter shown in Fig. 4. A single input voltage is induced, whereas a single output voltage drops at a load conductance. MNA produces a linear DAE system (10) of

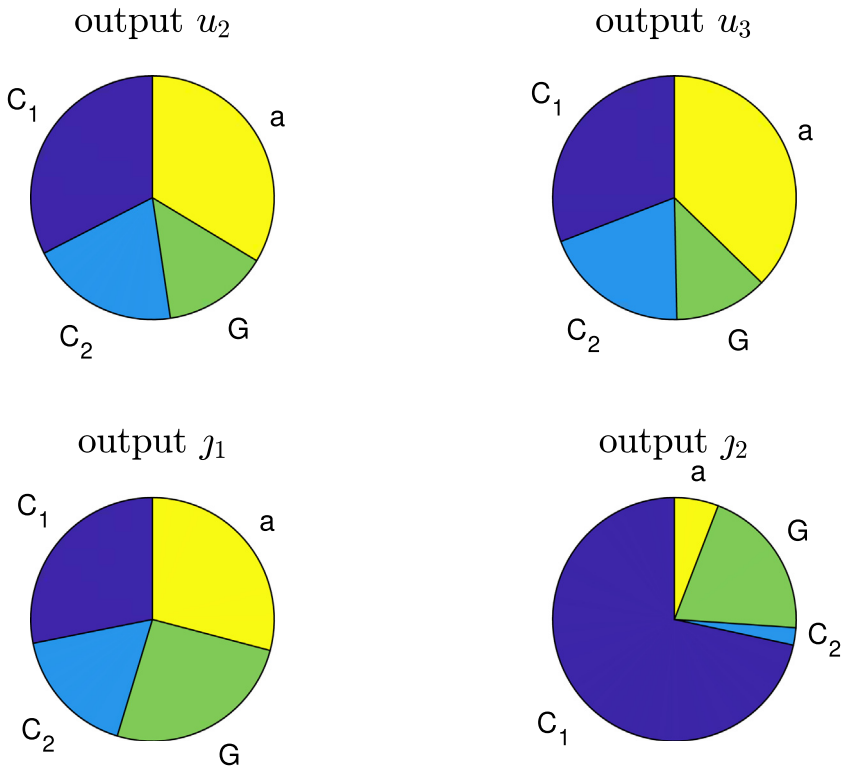

Fig. 3. Sensitivity measures of different outputs in example of the Miller integrator.

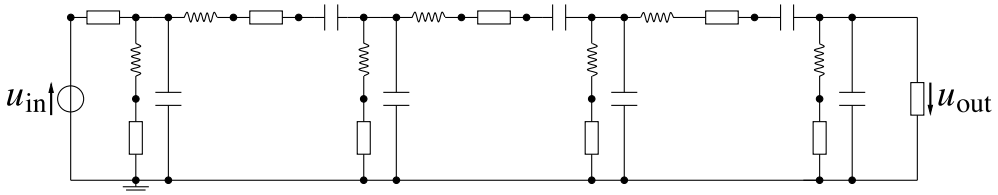

Fig. 4. Electric circuit of a band-pass filter.

Table 2
Physical parameters for a band-pass filter model.

Capacitances	10^{-6}
Inductances	10^{-6}
Conductances	10
Conductance next to input voltage	1
Conductance at output voltage	10^{-3}

dimension $n = 23$. The physical parameter vector consists in 7 capacitances, 7 inductances and 9 conductances ($q = 23$). The system matrices are affine-linear functions of the parameters as in (23). The DAE system has index one for all positive parameters, and its transfer function is strictly proper. Fig. 5 shows the Bode plot of this system for a constant selection of the parameters specified in Table 2.

Now we replace the parameters by independent random variables with uniform probability distributions which vary 10% around the above constant choice. The basis functions are again products of the Legendre polynomials. We include all basis polynomials up to degree $d = 2$. The number of basis functions is $m = 300$, and, hence, the stochastic Galerkin system (19) exhibits the dimension $mn = 6900$. The matrices of the Galerkin system were calculated using the Kronecker products in (22). Table 3 illustrates the sparsity of the computed system matrices. This system inherits the properties

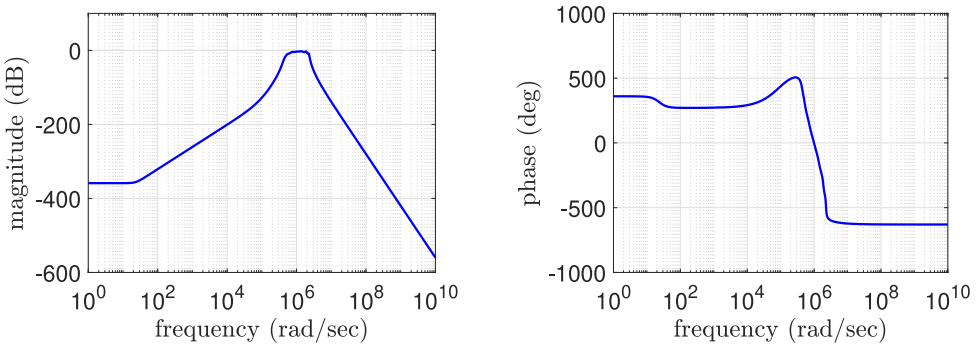

Fig. 5. Bode plot of the band-pass filter model for a constant realisation of the parameters.

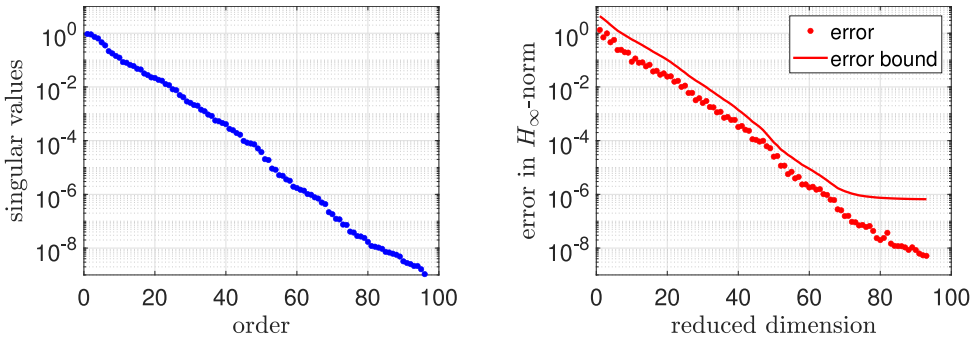

Fig. 6. Proper Hankel singular values (left); approximation errors in the $\mathcal{H}_\infty$-norm and error bound for different reduced dimensions (right).

Table 3
Sparsity of the matrices of the stochastic Galerkin system in the band-pass filter example.

Matrix	# Non-zero entries	Percentage
$\hat{A}$	46,776	0.098%
$\hat{E}$	22,938	0.048%
$\hat{B}$	24	0.348%
$\hat{C}$	300	2.26%

of the original DAE system, i.e., it is of index one, asymptotically stable and has a strictly proper transfer function. This example was also considered in [34,46], where the one-sided Arnoldi algorithm was used for model reduction of the stochastic Galerkin system.

Note that the equations and unknowns of the circuit model can just be rearranged such that the DAE system (10) becomes semi-explicit. Then the associated stochastic Galerkin system is also permuted to obtain a semi-explicit system with a mass matrix of the form (24). With this rearrangement, the number of differential variables is 4200 and the number of algebraic variables is 2700.

For model reduction of the stochastic Galerkin system (19), we apply the balanced truncation method as described in Sections 5.1 and 5.2. For computing the integral (47), we use the Gauss–Kronrod quadrature rule. The dominant $r_{\max} = 93$ proper Hankel singular values are shown in Fig. 6 (left). The approximation errors $\|\hat{H} - \tilde{H}\|_{\mathcal{H}_\infty}$ and error bounds (46) are presented in Fig. 6 (right). In the error bound (46), we substituted all missing singular values σ_i for $i > 93$ by the singular value σ_{93}.

The Galerkin systems (35) were substituted by the reduced-order models of dimension $r = 60$ and their $\mathcal{H}_\infty$-norms were used to compute the values (49) as approximations of the sensitivity measures (36) shown in Fig. 7 (left), where the physical parameters are ordered as in Table 2. The same quantities are also presented in Fig. 9. Furthermore, Fig. 7 (right) demonstrates the sums of the sensitivity measures for each group of the physical parameters. We observe that the conductances are less important than the capacitances and inductances.

For comparison, we perform a transient simulation of the stochastic Galerkin system (19) on the time interval $[0, T]$ with $T = 4 \cdot 10^{-4}$ using the trapezoidal rule as a time integrator. As the input voltage, we supply the chirp signal

$$u_{\text{in}}(t) = \sin\left(2\pi\kappa t^2\right) \qquad \text{with } \kappa = 10^9.$$

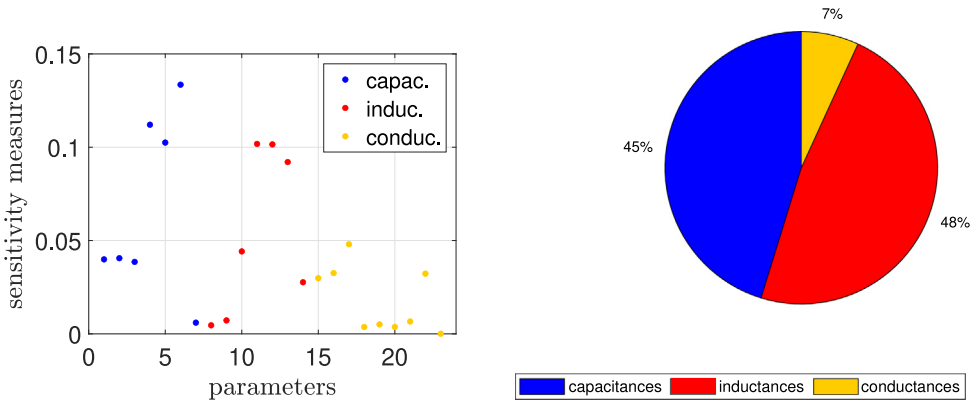

Fig. 7. Sensitivity measures $\hat{\theta}_j$ (left) and merged sensitivity measures for different groups of physical parameters (right) in band-pass filter example.

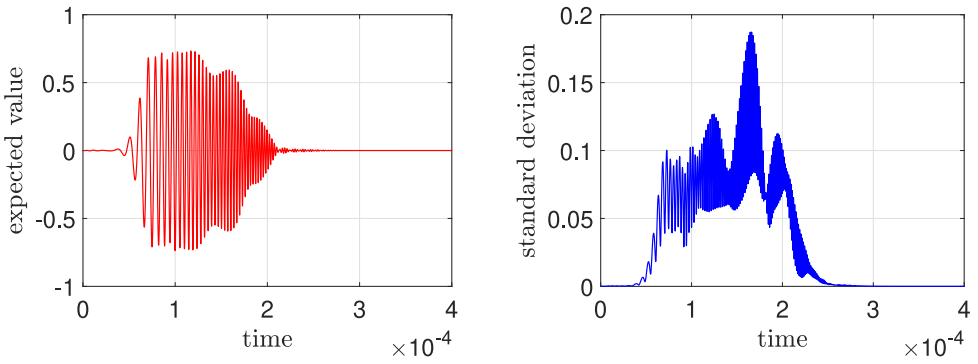

Fig. 8. Expected value (left) and standard deviation (right) of random output voltage in electric circuit of band-pass filter.

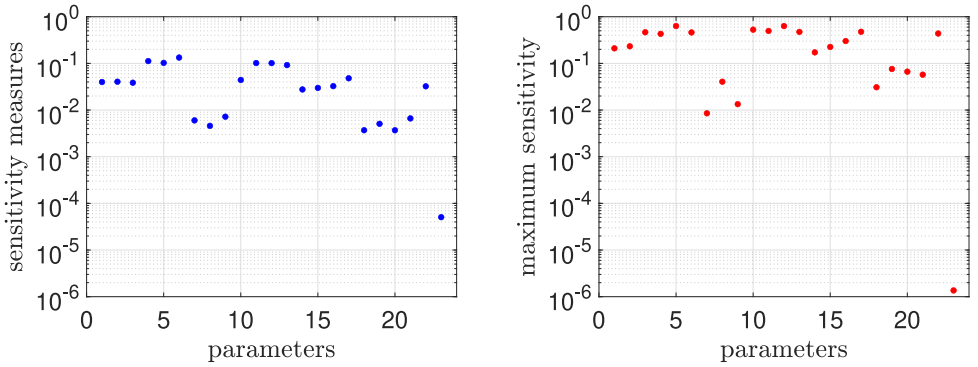

Fig. 9. Sensitivity measures (36) from $\mathcal{H}_\infty$-norms (left) and maximum of total effect sensitivity indices (52) in transient simulation (right) using semi-logarithmic scale for the band-pass filter example.

Fig. 8 depicts the approximations of the expected value as well as the standard deviation associated with the random output voltage. We observe that this electric circuit acts indeed as a band-pass filter, because only intermediate frequencies pass through, while lower and higher frequencies are damped out. The time-dependent total effect sensitivity indices (29) are determined by the outputs of the stochastic Galerkin system (19). The maximum values in time

$$\max_{t\in[0,T]} \hat{S}_j(t), \qquad j = 1, \ldots, q, \tag{52}$$

are shown in Fig. 9 (right). Likewise, Fig. 9 (left) illustrates the same quantities as in Fig. 7 (left) but using a logarithmic scale. The relative positions of the maximum sensitivities (52) agree to the relative positions in the sensitivity measures (36) shown in Fig. 9 (left).

Appendix

Proof of Theorem 1. Let $t \geq 0$ be fixed.
(i) It holds that

$$y(t,p) - \sum_{i \in \mathcal{I}^d} \hat{w}_i(t)\Phi_i(p) = \sum_{i \notin \mathcal{I}^d} w_i(t)\Phi_i(p) + \sum_{i \in \mathcal{I}^d} (w_i(t) - \hat{w}_i(t))\Phi_i(p).$$

The orthonormality of the basis polynomials allows for the application of Parseval's identity. We obtain

$$\left\| y(t,\cdot) - \sum_{i \in \mathcal{I}^d} \hat{w}_i(t)\Phi_i(\cdot) \right\|^2_{\mathcal{L}^2(\Pi)} = \sum_{i \notin \mathcal{I}^d} w_i(t)^2 + \sum_{i \in \mathcal{I}^d} (w_i(t) - \hat{w}_i(t))^2.$$

The first term converges to zero due to the assumption $y(t,\cdot) \in \mathcal{L}^2(\Pi)$. The convergence (31) implies that the second term also converges to zero.

Let $w^d = (w_2, \ldots, w_m)^\top$ and $\hat{w}^d = (\hat{w}_2, \ldots, \hat{w}_m)^\top$ with $m = |\mathcal{I}^d|$. The bound

$$\|w^d(t) - \hat{w}^d(t)\|_2^2 = \sum_{i \in \mathcal{I}^d \setminus \{1\}} (w_i(t) - \hat{w}_i(t))^2 \leq \sum_{i \in \mathcal{I}^d} (w_i(t) - \hat{w}_i(t))^2$$

shows that

$$\lim_{d \to \infty} \|w^d(t) - \hat{w}^d(t)\|_2 = 0 \tag{53}$$

in the Euclidean norm $\|\cdot\|_2$. We apply the reverse triangle inequality

$$\left| \|w^d\|_2 - \|\hat{w}^d\|_2 \right| \leq \|w^d - \hat{w}^d\|_2.$$

The convergence properties and Parseval's identity imply

$$\lim_{d \to \infty} \|\hat{w}^d(t)\|_2 = \lim_{d \to \infty} \|w^d(t)\|_2 = \sqrt{\|y(t,\cdot)\|^2_{\mathcal{L}^2(\Pi)} - w_1(t)^2}. \tag{54}$$

We obtain using the reverse triangle inequality again

$$\begin{aligned} \left| \|w^d\|_2^2 - \|\hat{w}^d\|_2^2 \right| &= \left| \|w^d\|_2 - \|\hat{w}^d\|_2 \right| \cdot \left| \|w^d\|_2 + \|\hat{w}^d\|_2 \right| \\ &\leq \|w^d - \hat{w}^d\|_2 (\|w^d\|_2 + \|\hat{w}^d\|_2). \end{aligned} \tag{55}$$

It follows that

$$\lim_{d \to \infty} \left| \|w^d(t)\|_2^2 - \|\hat{w}^d(t)\|_2^2 \right| = 0, \tag{56}$$

since the upper bound in (55) tends to zero for increasing degree d.

(ii) The approximation of the total variance exhibits the error

$$\begin{aligned} \left| V(t) - \hat{V}(t) \right| &= \left| \sum_{i \notin \mathcal{I}^d} w_i(t)^2 + \sum_{i \in \mathcal{I}^d \setminus \{1\}} w_i(t)^2 - \sum_{i \in \mathcal{I}^d \setminus \{1\}} \hat{w}_i(t)^2 \right| \\ &\leq \sum_{i \notin \mathcal{I}^d} w_i(t)^2 + \left| \|w^d(t)\|_2^2 - \|\hat{w}^d(t)\|_2^2 \right|. \end{aligned}$$

The first term tends to zero, since we assume $y(t,\cdot) \in \mathcal{L}^2(\Pi)$. The second term converges to zero due to (56).

(iii) Let $w^{d,j}$ and $\hat{w}^{d,j}$ be the vectors consisting of the coefficients w_i and $\hat{w}_i$, respectively, for all $i \in \mathcal{I}_j^d$. Furthermore, if $i \in \mathcal{I}_j \setminus \mathcal{I}_j^d$, then it holds that $\text{degree}(\Phi_i) > d$ and thus $i \notin \mathcal{I}^d$. We obtain an estimate for the error with respect to the approximations (30) of the partial variances

$$\begin{aligned} \left| V_j(t) - \hat{V}_j(t) \right| &= \left| \sum_{i \in \mathcal{I}_j \setminus \mathcal{I}_j^d} w_i(t)^2 + \sum_{i \in \mathcal{I}_j^d} w_i(t)^2 - \sum_{i \in \mathcal{I}_j^d} \hat{w}_i(t)^2 \right| \\ &\leq \sum_{i \in \mathcal{I}_j \setminus \mathcal{I}_j^d} w_i(t)^2 + \left| \|w^{d,j}(t)\|_2^2 - \|\hat{w}^{d,j}(t)\|_2^2 \right| \\ &\leq \sum_{i \notin \mathcal{I}^d} w_i(t)^2 + \left| \|w^{d,j}(t)\|_2^2 - \|\hat{w}^{d,j}(t)\|_2^2 \right|. \end{aligned}$$

The first term converges to zero again due to $y(t,\cdot) \in \mathcal{L}^2(\Pi)$. As in (55), we estimate for the second term

$$\begin{aligned} \left| \|w^{d,j}\|_2^2 - \|\hat{w}^{d,j}\|_2^2 \right| &\leq \|w^{d,j} - \hat{w}^{d,j}\|_2 (\|w^{d,j}\|_2 + \|\hat{w}^{d,j}\|_2) \\ &\leq \|w^d - \hat{w}^d\|_2 (\|w^d\|_2 + \|\hat{w}^d\|_2) \end{aligned}$$

uniformly for $j = 1, \ldots, q$. This upper bound also tends to zero due to the convergence (53) and (54). □

References

[1] M. Günther, U. Feldmann, J. ter Maten, Modelling and discretization of circuit problems, in: P.G. Ciarlet (Ed.), Handbook of Numerical Analysis, Vol. 13, Elsevier, 2005, pp. 523–659.
[2] C.W. Ho, A. Ruehli, P. Brennan, The modified nodal approach to network analysis, IEEE Trans. Circuits Syst. 22 (1975) 504–509.
[3] W. Kampowsky, P. Rentrop, W. Schmidt, Classification and numerical simulation of electric circuits, Surv. Math. Ind. 2 (1992) 23–65.
[4] T.J. Sullivan, Introduction to Uncertainty Quantification, Springer International Publishing, 2015.
[5] D. Xiu, Numerical Methods for Stochastic Computations: A Spectral Method Approach, Princeton University Press, 2010.
[6] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis, John Wiley & Sons, Ltd., 2008.
[7] I.M. Sobol, Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp. 1 (1993) 407–414.
[8] I.M. Sobol, S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, Math. Comput. Simulation 79 (2009) 3009–3017.
[9] R. Pulch, E.J.W. ter Maten, F. Augustin, Sensitivity analysis and model order reduction for random linear dynamical systems, Math. Comput. Simulation 111 (2015) 80–95.
[10] T. Soll, R. Pulch, Sample selection based on sensitivity analysis in parameterized model order reduction, J. Comput. Appl. Math. 316 (2017) 271–286.
[11] P. Manfredi, D. De Zutter, D. Vande Ginste, On the relationship between the stochastic Galerkin method and the pseudo-spectral collocation method for linear differential algebraic equations, J. Engrg. Math. 108 (2018) 73–90.
[12] R. Pulch, Polynomial chaos for linear differential algebraic equations with random parameters, Int. J. Uncertain. Quantif. 1 (2011) 223–240.
[13] R. Pulch, Stochastic collocation and stochastic Galerkin methods for linear differential algebraic equations, J. Comput. Appl. Math. 262 (2014) 281–291.
[14] R. Pulch, A. Narayan, Sensitivity analysis of random linear dynamical systems using quadratic outputs, J. Comput. Appl. Math. 387 (2021) 112491.
[15] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, SIAM, Philadelphia, 1996.
[16] E. Hairer, G. Wanner, Solving Ordinary Differential Equations, second ed., in: Stiff and Differential-Algebraic Equations, vol. 2, Springer, Berlin, 1996.
[17] S. Boyd, V. Balakrishnan, P. Kabamba, A bisection method for computing the H_∞ norm of a transfer matrix and related problems, Math. Control Signals Systems 2 (1989) 207–219.
[18] N. Guglielmi, M. Gürbüzbalaban, M.L. Overton, Fast approximation of the H_∞ norm via optimization over spectral value sets, SIAM J. Matrix Anal. Appl. 34 (2013) 709–737.
[19] P. Benner, T. Mitchell, Faster and more accurate computation of the H_∞-norm via optimization, SIAM J. Sci. Comput. 40 (2018) A3609–A3635.
[20] P. Benner, V. Sima, M. Voigt, L_∞-Norm computation for continuous-time descriptor systems using structured matrix pencils, IEEE Trans. Automat. Control 57 (2012) 233–238.
[21] P. Benner, T. Stykel, Model order reduction of differential–algebraic equations: a survey, in: A. Ilchmann, T. Reis (Eds.), Surveys in Differential-Algebraic Equations IV, Differential-Algebraic Equations Forum, Springer, 2017, pp. 107–160.
[22] P. Benner, A. Schneider, Balanced Truncation for Descriptor Systems with Many Terminals, Preprint MPIMD/13-17, Max Planck Institute Magdeburg, 2013, Available from https://www2.mpi-magdeburg.mpg.de/preprints/2013/MPIMD13-17.pdf.
[23] T. Stykel, Gramian-based model reduction for descriptor systems, Math. Control Signals Systems 16 (2004) 297–319.
[24] C. Schröder, M. Voigt, Balanced truncation model reduction with a priori error bounds for LTI systems with nonzero initial value, 2020, Preprint arXiv:2006.02495 [eess.SY].
[25] E. Griepentrog, R. März, Differential-Algebraic Equations and their Numerical Treatment, in: Teubner-Texte zur Mathematik, vol. 88, Teubner, Leipzig, 1986.
[26] A. Antoulas, Approximation of Large-Scale Dynamical Systems, SIAM Publications, 2005.
[27] T. Stykel, On some norms for descriptor systems, IEEE Trans. Automat. Control 51 (2006) 842–847.
[28] N. Aliyev, P. Benner, E. Mengi, P. Schwerdtner, M. Voigt, Large-scale computation of $\mathcal{L}_\infty$-norms by a greedy subspace method, SIAM J. Matrix Anal. Appl. 38 (2017) 1496–1516.
[29] P. Benner, M. Voigt, A structured pseudospectral method for $\mathcal{H}_\infty$-norm computation of large-scale descriptor systems, Math. Control Signals Systems 26 (2014) 303–338.
[30] O. Ernst, A. Mugler, H.-J. Starkloff, E. Ullmann, On the convergence of generalized polynomial chaos expansions, ESAIM Math. Model. Numer. Anal. 46 (2012) 317–339.
[31] R. Pulch, E.J.W. ter Maten, Stochastic Galerkin methods and model order reduction for linear dynamical systems, Int. J. Uncertain. Quantif. 5 (2015) 255–273.
[32] E. Ullmann, H.C. Elman, O.G. Ernst, Efficient iterative solvers for stochastic Galerkin discretizations of log-transformed random diffusion problems, SIAM J. Sci. Comput. 21 (2012) A659–A682.
[33] R. Pulch, F. Augustin, Stability preservation in stochastic Galerkin projections of dynamical systems, SIAM/ASA J. Uncertain. Quantif. 7 (2019) 634–651.
[34] R. Pulch, Stability-preserving model order reduction for linear stochastic Galerkin systems, J. Math. Ind. 9 (10) (2019).
[35] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. Syst. Saf. 93 (2008) 964–979.
[36] R. Pulch, A. Narayan, Balanced truncation for model order reduction of linear dynamical systems with quadratic outputs, SIAM J. Sci. Comput. 41 (2019) A2270–A2295.
[37] R. Van Beeumen, K. Van Nimmen, G. Lombaert, K. Meerbergen, Model reduction for dynamical systems with quadratic output, Internat. J. Numer. Methods Engrg. 91 (2012) 229–248.
[38] R. Pulch, Model order reduction and low-dimensional representations for random linear dynamical systems, Math. Comput. Simulation 144 (2018) 1–20.
[39] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, Environ. Model. Softw. 22 (2007) 1509–1518.
[40] B.C. Moore, Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Trans. Automat. Control AC-26 (1981) 17–32.
[41] M. Safonov, R. Chiang, A Schur method for balanced-truncation model reduction, IEEE Trans. Automat. Control AC-34 (1989) 729–733.

[42] P. Benner, A. Schneider, Balanced truncation model order reduction for LTI systems with many inputs or outputs, in: A. Edelmayer (Ed.), Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010), 5–9 July, 2010, Budapest, Hungary, 2010, pp. 1971–1974.
[43] T. Stykel, Low-rank iterative methods for projected generalized Lyapunov equations, Electron. Trans. Numer. Anal. 30 (2008) 187–202.
[44] T. Stykel, V. Simoncini, Krylov subspace methods for projected Lyapunov equations, Appl. Numer. Math. 62 (2012) 35–50.
[45] T. Stykel, Numerical solution and perturbation theory for generalized Lyapunov equations, Linear Algebra Appl. 349 (2002) 155–185.
[46] R. Pulch, Stability preservation in model order reduction of linear dynamical systems, in: G. Nicosia, V. Romano (Eds.), Scientific Computing in Electrical Engineering SCEE 2018, Mathematics in Industry, Vol. 32, Springer International Publishing, 2020, pp. 277–285.