Reduced order models for diffusion systems

Bharath Bhikkaji and Torsten Söderström

Department of Systems and Control, Information Technology,
Uppsala University, P O Box 27, SE 75121, Uppsala, Sweden,
email ts@SysCon.uu.se, bb@SysCon.uu.se

August 15, 2000

Abstract

Mathematical models for diffusion processes like heat propagation, dispersion of pollutants etc., are normally partial differential equations which involve certain unknown parameters. To use these mathematical models as the substitutes of the true system, one has to determine these parameters.

Partial differential equations (PDE) of the form

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{L}u(x,t) \tag{1}$$

where $\mathcal{L}$ is a linear differential (spatial) operator, describe infinite dimensional dynamical systems. To compute a numerical solution for such partial differential equations, one has to approximate the underlying system by a finite order one. By using this finite order approximation, one then computes an approximate numerical solution for the PDE.

We consider a simple case of heat propagation in a homogeneous wall. The resulting partial differential equation, which is of the form (1), is approximated by finite order models by using certain existing numerical techniques like Galerkin and Collocation etc. These reduced order models are used to estimate the unknown parameters involved in the PDE, by using the well developed tools of system identification.

In this paper we concentrate more on the model reduction aspects of the problem. In particular, we examine the model order reduction capabilities of the Chebyshev polynomial methods used for solving partial differential equation.

Keywords: Diffusion, State-space model, system identification, Collocation method, Chebyshev polynomials, Galerkin method.

1 Introduction

Diffusion systems have been studied with a lot of interest, [1], [4], due to its historical and practical importance. A lot of physical process like heat propagation in a medium, flow of chemicals across the concentration gradient in a chemical battery and dispersion of pollutants in the atmosphere are examples of diffusion phenomena. Systems which involve diffusion phenomena are called diffusion systems. Diffusion is normally modeled by a partial differential equation (PDE), which may involve some unknown parameters. In order to model such a system one has to determine these parameters.

System identification is an experimental approach to model dynamic systems [7]. An experiment is performed by exciting the system with an input signal and recording the corresponding system output. A model is then fitted for the recorded input-output data. A key problem in system identification is to select a model structure, within which a good model has to be found for the given input-output data. Fitting a model within a given structure (parameter estimation) is a simpler problem in most cases. Normally one uses physical insight and prior knowledge about the system when selecting a model structure. If no prior knowledge or physical insight is available, then one selects a model structure which is flexible and have been successful in the past. Models of the former type are called grey box models and latter are called black box models. Unlike black box models grey box models are strongly application dependent.

In this paper, we consider the simple case of heat propagation in a homogeneous wall. We use this simple case of heat propagation and illustrate the use of system identification in modeling diffusion systems. The PDE that results from this simple case of heat diffusion is reduced to approximate finite order models by using Chebyshev polynomial based numerical methods, and these reduced order models are used to estimate the unknown parameters involved in the PDE.

In this paper we focus on the model reduction part rather than the parameter estimation part. In particular, we look at the model reduction capabilities of the Chebyshev polynomial methods used for solving the PDE.

This paper is classified into five sections. In section 2 we define and extensively discuss the problem to be solved and also discuss our approach to solve the problem. In section 3 we derive the true infinite order dynamics of the system, while in section 4 we construct some approximate finite order models and discuss in detail about them.

2 A case study: the heat diffusion equation

2.1 Statement of the problem

Consider a homogeneous wall of thickness d units. Let x be the coordinate across the wall, T_e the temperature on the exterior side, T_i the temperature on the

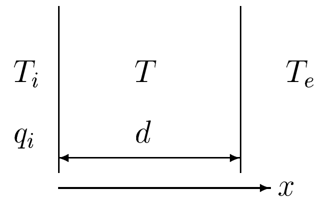

Figure 1: Heat diffusion through a homogeneous wall.

interior side and q_i the supplied heat on the interior side, cf Figure 1.

Let $T = T(x,t)$ denote the temperature at position x and at time t. Then the heat flow and the temperature are described by

$$\frac{\partial T(x,t)}{\partial t} = \alpha \frac{\partial^2 T(x,t)}{\partial x^2} \tag{2}$$

with boundary constraints

$$q_i(t) = -\kappa \frac{\partial T(x,t)}{\partial x} |_{x=0}, \tag{3}$$

$$T_e(t) = T(d,t) \tag{4}$$

For simplicity we consider the initial condition

$$T(x,0) = 0. \tag{5}$$

In the model (2)-(4) the diffusion coefficient α and and the thermal conductivity κ are the unknown parameters.

Needless to say, both the case selected for study and its geometry are simple. While the geometry is chosen for technical ease, the simple case is aimed at getting insight into more difficult systems for future work.

Consider the wall as a dynamic system with input $u(t) = (q_i(t), T_e(t))^T$ and output $y(t) = T_i(t)$.

The estimation problem to be considered is the following.

Given: Discrete-time data $T_e(h), q_i(h), T_i(h), T_e(2h), q_i(2h), \ldots, T_i(Nh)$.
Find: Parameter estimates of α and κ.

It is apparent that the above problem is an example of grey box modeling. Here we consider an output error approach, [7].

We fit a model

$$\mathcal{M}(\theta) : y^m(t,\theta) = G(q^{-1},\theta)u(t) \tag{6}$$

for the given data, where an estimate of $\theta = (\alpha, \kappa)^T$ is determined as the minimizer of

$$V(\theta) = \frac{1}{N}\sum_{j=0}^{N-1}[y(jh) - y^m(jh, \theta)]^2 \tag{7}$$

As mentioned earlier, the key problem is to determine the model structure $\mathcal{M}(\theta)$. The parameter θ can be determined by any standard non-linear search method.

Any stable numerical scheme which computes the solution of a PDE, is a model reduction procedure which approximates the true infinite order dynamical system (PDE) by a finite order dynamical system. The order of the approximate dynamical system determines the degree of approximation: The higher the order, the better the approximation. Normally one would prefer to have a low model order in order to reduce the computational complexity without compromising too much on the accuracy.

We first look at the dynamics of the true system and then later proceed to construction of approximate models.

3 The system dynamics

The defined system is linear and has two inputs, q_i and T_e, and one output, T_i. Interestingly enough, one can rather easily derive the transfer function of the system, which will give an input-output description of the dynamics, [6]. For convenience we repeat the derivation here. To do this, first take the Laplace transform of (2) which gives

$$sT(x, s) = \alpha\frac{d^2}{dx^2}T(x, s) \tag{8}$$

where $T(x, s)$ is the Laplace transform of the temperature $T(x, t)$. Regarding s as a fixed parameter, (8) is a standard ordinary differential equation. It is found that the general solution to (8) can be written as

$$T(x, s) = A(s)\cosh(x\sqrt{\frac{s}{\alpha}}) + B(s)\sinh(x\sqrt{\frac{s}{\alpha}}) \tag{9}$$

Further, taking Laplace transforms on (3),(4) and evaluating the constants $A(s)$ and $B(s)$ gives the following expression for temperature at the interior side.

$$T_i(s) = \underbrace{\frac{\tanh(d\sqrt{\frac{s}{\alpha}})}{\kappa\sqrt{\frac{s}{\alpha}}}}_{G_1(s)} q_i(s) + \underbrace{\frac{1}{\cosh(d\sqrt{\frac{s}{\alpha}})}}_{G_2(s)} T_e(s). \tag{10}$$

Note that both $G_1(s)$ and $G_2(s)$ are transcendental functions and have poles at

$$p_n = -\frac{\pi^2}{d^2}(n - \frac{1}{2})^2\alpha, \quad n = 1, 2 \ldots . \tag{11}$$

It is also worth noting that $G_1(s)$ and $G_2(s)$ are both of low pass character, see Figures 2 and 3. This property enables us to choose models approximating (10) depending on the frequency contents of the input signals. Since the transfer functions $G_1(s)$ and $G_2(s)$ are of finite magnitude, $T_i(t)$ depends continuously on the input data $q_i(t)$ and $T_e(t)$.

Since the true system (10) is of infinite order, a model large enough but of finite order is to be fitted. Normally one would like to choose a model order, which renders the estimated parameters $\hat{\theta}$ close to the true parameter θ_0.

Recall that we use an output error approach to estimate the parameters, that is the parameter estimate $\hat{\theta}$ is the minimizer of

$$V(\theta) = \frac{1}{N}\sum_{t=0}^{N-1}[y(t) - y^m(t, \theta)]^2. \tag{12}$$

If N is large enough then

$$V(\theta) \approx E\{[y(t) - y^m(t, \theta)]^2\}. \tag{13}$$

where the system output is of the form

$$y(t) = G(p, \theta_0)u(t) \tag{14}$$

and the model output is given by

$$y^m(t, \theta) = \hat{G}^n(p, \theta)u(t), \tag{15}$$

and $\hat{G}^n(p, \theta)$ is an finite order approximation of the infinite order transfer function $G(p, \theta)$. For any good approximate model $\hat{G}^n(p, \theta)$ of $G(p, \theta)$, we must have

$$G(i\omega, \theta) \approx \hat{G}^n(i\omega, \theta). \tag{16}$$

in any frequency region of relevance. To see how the model structure determination and model order selection are important in the problem, we do the following calculations, with the assumption that $\hat{\theta} \approx \theta$.

Let

$$V(\theta) = E\{[y(t) - y^m(t, \theta)]^2\}. \tag{17}$$

Since $\hat{\theta} \approx \theta_0$ and $\hat{\theta}$ minimizes $V(\theta)$, a Taylor approximation of $\frac{\partial V}{\partial \theta}|_{\theta=\hat{\theta}}$, would give

$$0 = \frac{\partial V}{\partial \theta}|_{\theta=\hat{\theta}} \approx \frac{\partial V}{\partial \theta}|_{\theta=\theta_0} + \frac{\partial^2 V}{\partial \theta^2}|_{\theta=\theta_0}(\hat{\theta} - \theta_0), \tag{18}$$

which implies that

$$(\hat{\theta} - \theta_0) \approx -[\frac{\partial^2 V}{\partial \theta^2}|_{\theta=\theta_0}]^{-1}(\frac{\partial V}{\partial \theta}|_{\theta=\theta_0}). \tag{19}$$

Note that

$$\frac{\partial V}{\partial \theta}|_{\theta=\theta_0} = 2E\{[-\frac{\partial \hat{G}^n}{\partial \theta}|_{\theta=\theta_0} u(t)][(G(p,\theta_0) - G^n(p,\hat{\theta}_0))u(t)]\} \tag{20}$$

$$\frac{\partial^2 V}{\partial \theta^2}|_{\theta=\theta_0} \approx 2E\{[-\frac{\partial \hat{G}^n}{\partial \theta} u(t)][-\frac{\partial \hat{G}^n}{\partial \theta} u(t)|^T_{\theta=\theta_0}]\}, \tag{21}$$

which gives an expression for the bias. If one assumes $\frac{\partial G^n}{\partial \theta}$ to be bounded, then the bias decreases with increasing model order n at the rate of convergence of G^n to G. It is for this reason we focus in this report on how well $\hat{G}^n(p,\theta)$ approximates $G(p,\theta_0)$ for a number of numerical PDE solvers.

4 Approximate Models

Before going into the details of constructing approximate models, we first discuss briefly about Chebyshev polynomials

4.1 Chebyshev Polynomials

Chebyshev polynomials are defined by the recurrence relation

$$C_{i+1}(x) = 2xC_i(x) - C_{i-1}(x) \quad i \geq 1 \tag{22}$$

with $C_0(x) = 1$ and $C_1(x) = x$.

Alternatively, Chebyshev polynomials can also be defined as

$$C_i(x) = \cos(i \cos^{-1}(x)) \quad i \geq 0 \tag{23}$$

Lemma 1. *The definitions (22) and (23) of the Chebyshev polynomials are equivalent.*

Proof. Standard and given in in several works, see [5] as an example

Let $\mathcal{P}$ be a an infinite dimensional space which is spanned by $\{C_j(x)\}_{j=0}^{\infty}$ and having a well defined inner product

$$\langle u, v \rangle = \int_{-1}^{1} \frac{u(x)v(x)}{\sqrt{1-x^2}} dx \tag{24}$$

where $u(x), v(x) \in \mathcal{P}$.

Let $\mathcal{P}_n$ be the vector space of polynomials of degree less than or equal to n, with inner product as defined in (24).

Lemma 2. *The set* $\left\{\sqrt{\frac{2}{\pi d_i}}C_i(x)\right\}_{i=0}^{n}$, *where*

$$d_0 = 2$$
$$d_i = 1 \quad i \geq 1, \tag{25}$$

forms an orthonormal basis for the space $\mathcal{P}_n$.

Proof. Standard. For completeness, a proof is included in the appendix.

Let

$$x_j = \cos(\frac{\pi j}{n+1}) \quad j = 0, 1, \ldots, n+1. \tag{26}$$

Note that

$$|C_{n+1}(x_j)| = 1 \quad j = 0, 1, \ldots, n+1. \tag{27}$$
$$C_j(x_k) = C_k(x_j) = \cos(\frac{\pi jk}{n+1}) \quad j, k = 0, 1, \ldots, n+1. \tag{28}$$

The following lemma is a very useful tool for dealing with Chebyshev polynomials.

Lemma 3. *Let* $h(x)$ *be a polynomial of degree* $2n+1$. *Then it holds that*

$$\int_{-1}^{1} \frac{h(x)}{\sqrt{1-x^2}}dx = \frac{\pi}{n+1}\sum_{k=0}^{n+1}\frac{h(x_k)}{c_k} \tag{29}$$

where

$$c_0 = c_{n+1} = 2$$
$$c_k = 1 \quad k = 1, \ldots, n \tag{30}$$

and the points x_k *are given by (26).*

Proof. See [2], [5].

Remark: Lemma 3 is normally referred to as **Gauss Lobatoo quadrature formula**.

We will also need the derivatives of Chebyshev polynomials, particularly at the points x_j, as they will come very handy while constructing the approximate models.

Straightforward differentiation of (23) leads to

$$C_j^{'}(x) = \frac{j \sin(j \cos^{-1}(x))}{\sqrt{1-x^2}}, \quad x^2 \neq 1 \tag{31}$$

Hence, by using L'Hospital's rule we get

$$C_j^{'}(x_{n+1}) = (-1)^{j-1} j^2 \tag{32}$$

$$j = 0, \ldots, n+1$$

$$C_j^{'}(x_0) = j^2 \tag{33}$$

where x_k are as defined in (26).

4.2 Galerkin method

In this section, we dicuss the Galerkin method by applying by applying it to our problem (2) - (5). For a more general discussion on Galerkin, please refer to [3].

In Galerkin method, one assumes that the solution $T(x,t)$ of the PDE (2) - (4) belongs to an infinite dimensional function space, with a well defined inner product. In our case let us assume that the solution belongs to the space $\mathcal{P}$, defined in the previous subsection, with the inner product (24).

A Galerkin approximation of the solution $T(x,t)$ of (2) - (5) is a truncated series

$$T_n(x,t) = \sum_{j=0}^{n} a_j(t)\phi_j(x), \tag{34}$$

where the time independent functions $\phi_j(x)$ are polynomials of degree less than or equal to n and the set $\{\phi_j(x)\}_{j=0}^{n}$ is assumed to linearly independent. It is also assumed that each $\phi_j(x)$ satisfies the boundary conditions (3) and (4). Thus, $T_n(x,t)$ necessaryly satisfies the boundary conditions. The expansion coefficients $a_j(t)$ are determined by assuming that $T_n(x,t)$ is an orthogonal projection of $T(x,t)$ into the space (subspace) spanned by $\{\phi_j(x)\}_{j=0}^{n}$, or analytically

$$\langle \frac{\partial T_n(x,t)}{\partial t} - \alpha \frac{\partial^2 T_n(x,t)}{\partial x^2}, \phi_j(x) \rangle = 0 \qquad j = 0, 1, \ldots, n. \tag{35}$$

At a first look Galerkin method may look simple, but is not so simple, since each $\phi_j(x)$ is defined only in the interval (recall that Span($\{\phi_j(x)\}_{j=0}^{n}) \subset \mathcal{P}$) $[-1,1]$, and each $\phi_j(x)$ must satify the boundary conditions (3) and (4). Note that the boundary conditions are time-varying functions but the $\phi_j(x)$ are time independent. Both these problems can be circumvented by the following transformations. Let

$$z = \frac{x - \frac{d}{2}}{\frac{d}{2}} \tag{36}$$

which leads to the following transformation on (2)-(4).

$$\frac{\partial T(z,t)}{\partial t} = \frac{4\alpha}{d^2}\frac{\partial^2 T(z,t)}{\partial z^2} \tag{37}$$

with boundary constraints

$$q_i(t) = -\frac{2\kappa}{d}\frac{\partial T(z,t)}{\partial z}_{|z=-1}, \tag{38}$$

$$T_e(t) = T(1,t). \tag{39}$$

Now, let

$$U(z,t) \triangleq T(z,t) - \frac{(z+1)^2}{4}T_e(t) + \frac{d}{2\kappa}(z-1)q_i(t) \tag{40}$$

which implies that

$$U_{zz}(z,t) = T_{zz}(z,t) - \frac{1}{2}T_e(t) \tag{41}$$

$$\dot{U}(z,t) = \dot{T}(z,t) - \frac{(z+1)^2}{4}\dot{T}_e(t) + \frac{d}{2\kappa}(z-1)\dot{q}_i(t) \tag{42}$$

Using (37) we get a new PDE in $U(z,t)$:

$$\frac{\partial U(z,t)}{\partial t} = \frac{4\alpha}{d^2}\frac{\partial^2 U(z,t)}{\partial x^2} + F(z,t) \tag{43}$$

where

$$F(z,t) = F_1(z)T_e(t) + F_{21}(z)\dot{T}_e(t) + F_{22}(z)\dot{q}_i(t) \tag{44}$$

$$F_1(z,t) = \frac{2\alpha}{d^2} \tag{45}$$

$$F_{21}(z) = -\frac{(z+1)^2}{4} \tag{46}$$

$$F_{22}(z) = \frac{d}{2\kappa}(z-1) \tag{47}$$

with boundary conditions

$$\frac{\partial U(z,t)}{\partial z}_{|z=-1} = 0, \tag{48}$$

$$U(1,t) = 0. \tag{49}$$

Thus, we have a new problem with zero boundary conditions. It is not unresonable to find basis functions that satisfy these new boundary conditions. For example, let

$$\phi_j(z) = C_j(z) - (z-1)(-1)^{j-1}j^2 - 1 \quad j = 2, 3, \dots, n \tag{50}$$

where $C_j(z)$ is the j^{th} Chebyshev polynomial, see (22). Note that $\phi_j(z)$ is a polynomial of degree j. From (32), we can see that

$$\phi_j^{'}(-1) \;=\; C_j^{'}(-1) - (-1)^{j-1}j^2 = 0 \tag{51}$$

and from (23)

$$\phi_j(1) \;=\; C_j(1) - 1 = 0, \tag{52}$$

and hence $\{\phi_j(x)\}_{j=2}^n$ satisfies the boundary conditions (48) and (49). Note that, since

$$\begin{aligned} \phi_0(z) &= 0 \quad \text{and} && (53)\\ \phi_1(z) &= z - (z-1) - 1 = 0 && (54) \end{aligned}$$

we have not included them in (50).

From now on for the sake of convenience we refer to the variable z in (43) - (49) as x. The approximate solution for the PDE (43) - (49) is a truncated series

$$U_n(x,t) = \sum_{j=0}^{n} a_j(t)\phi_j(x), \tag{55}$$

as in (34), and each $\phi_j(x)$ satisfy the boundary conditions (48) and (49).

The expansion coefficients $\{a_j(t)\}_{j=0}^n$ are determined by assuming that $U_n(x,t)$ is an orthogonal projection of $U(x,t)$ onto the space (subspace) spanned by $\{\phi_j(x)\}_{j=0}^n$, that is,

$$\langle \frac{\partial U_n(x,t)}{\partial t} - \alpha \frac{\partial^2 U_n(x,t)}{\partial x^2} - F(x,t), \phi_k(x)\rangle = 0 \qquad k = 0, 1, \dots, n, \tag{56}$$

which implies that

$$\begin{aligned} \sum_{j=0}^{n} \dot{a}_j(t)\langle \phi_j(x), \phi_k(x)\rangle &= \alpha \sum_{j=0}^{n} a_j(t)\langle \phi_j^{''}(x), \phi_k(x)\rangle + \langle F(x,t), \phi_k(x)\rangle \\ & \text{for} \quad k = 0, 1, \dots, n. && (57) \end{aligned}$$

When (57) is written as a system of differential equations we get

$$\begin{aligned} \dot{a}(t) &= \alpha \mathcal{A}^{-1}\mathcal{S}a(t) + \mathcal{A}^{-1}\mathcal{F}_1 u + \mathcal{A}^{-1}\mathcal{F}_2 \dot{u} && (58)\\ a(0) &= 0 \end{aligned}$$

where

$$\begin{aligned}
a(t) &= (a_0(t), \dots, a_n(t))^T \\
\mathcal{A}_{jk} &= \langle \phi_j(x), \phi_k(x) \rangle \qquad j,k = 0,1,\dots,n & (59)\\
\mathcal{S}_{jk} &= \langle \phi_k''(x), \phi_j(x) \rangle \qquad j,k = 0,1,\dots,n & (60)\\
\mathcal{F}_{1k} &= [\langle F_1(x), \phi_k(x) \rangle, 0] \qquad k = 0,1,\dots,n & (61)\\
\mathcal{F}_{2k} &= [\langle F_{21}(x), \phi_k(x) \rangle, \langle F_{22}(x), \phi_k(x) \rangle] \qquad k = 0,1,\dots,n & (62)\\
u &= [T_e(t), q_i(t)]^T & (63)
\end{aligned}$$

Since

$$U(-1,t) = \sum_{j=0}^{n} a_j(t)\phi_j(-1), \tag{64}$$

using (40), we see that the desired output of the system, $T_i(t)$, is given by

$$T_i(t) = T(-1,t) = \sum_{j=0}^{n} a_j(t)\phi_j(-1) - \frac{d}{\kappa} q_i(t). \tag{65}$$

Rewriting (58) and (65) in a state space like form, we get

$$\begin{aligned}
\dot{a} &= Aa + B_1 u + B_2 \dot{u} & (66)\\
T_i &= Ca + Du, & (67)
\end{aligned}$$

where

$$\begin{aligned}
A &= \alpha \mathcal{A}^{-1} \mathcal{S} & (68)\\
B_1 &= \mathcal{A}^{-1} \mathcal{F}_1 & (69)\\
B_2 &= \mathcal{A}^{-1} \mathcal{F}_2 & (70)\\
C &= (\phi_0(-1), \dots, \phi_n(-1)) & (71)\\
D &= [0, -\frac{d}{\kappa}] & (72)
\end{aligned}$$

Though (66) and (67) are not in the standard state space form (due to the presence of $\dot{u}$ in (66)), one can still write them in the desired standard state space form with $T_e(t)$ and $q_i(t)$ as inputs, by using the following transformation. Set

$$z = a - B_2 u \tag{73}$$

which implies

$$\begin{aligned}
\dot{z} &= \dot{a} - B_2 \dot{u} \\
&= Aa + B_1 u \\
&= A(z + B_2 u) + B_1 u \\
&= Az + (B_1 + AB_2) u & (74)
\end{aligned}$$

and

$$\begin{aligned} T_i &= Ca + Du \\ &= Cz + (D + CB_2)\,u \end{aligned} \tag{75}$$

Thus we have rewritten (66) and (67) into the standard form state space form given by (74) and (75).

Rewriting (74) and (75) in the standard input-output form we have

$$Y(s) = [G_n^1(s) \;\; G_n^2(s)]\, U(s) \tag{76}$$

where

$$Y(s) = T_i(s) \tag{77}$$

$$G_n^1(s) = \bar{D}_1 + C(sI - A)^{-1}\bar{B}_1 \tag{78}$$

$$G_n^2(s) = \bar{D}_2 + C(sI - A)^{-1}\bar{B}_2 \tag{79}$$

$$U(s) = [q_i(s), T_e(s)]^T \tag{80}$$

The quantities $q_i(s)$, $T_e(s)$ and $T_i(s)$ are the Laplace transforms of $q_i(t)$, $T_e(t)$ and $T_i(t)$ respectively, and the subscript n in $G_n^1(s)$ and $G_n^2(s)$ denotes the order of approximation. Further $\bar{B}_1$, $\bar{B}_2$ are the column vectors of $B_1 + AB_2$, and $\bar{D}_1$, $\bar{D}_2$ are the column vectors of $D + CB_2$.

Figures 2 and 3 show the true transfer functions $G_1(s)$ and $G_2(s)$, respectively, along with their approximations (78) and (79) for different model orders n. As one can see from the Figures 2 and 3 at low frequencies the true transfer functions are well approximated, but at high frequencies the approximations detoriate, more so in the case of G_1. One can consider (78) and (79) as a good approximate model, if the input signal are of low frequency characteristics.

Since the approximate system converges to the true system as $n \to \infty$, one would expect the approximate system poles to be close to the true system poles. Note that (11) are the true system poles and the eigenvalues of the matrix A are the approximate system poles. Alternatively one can also use

$$p_k^{'} = \frac{p_k}{\alpha} = -\frac{\pi^2}{d^2}(k - \frac{1}{2})^2, \quad k = 1, 2 \ldots . \tag{81}$$

as the true system poles and the eigenvalues of the matrix

$$A_0 = \frac{A}{\alpha} \tag{82}$$

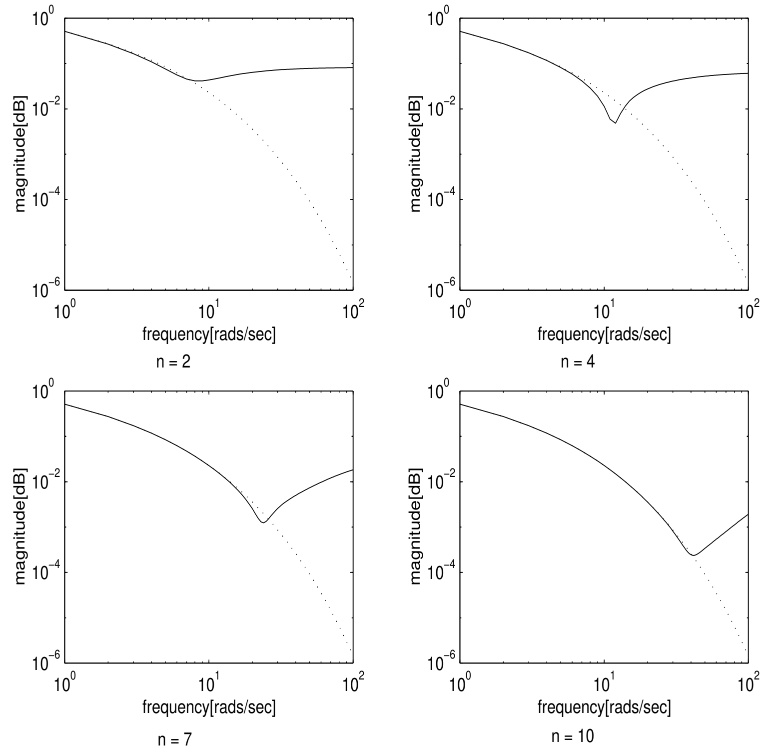

Figure 2: The Galerkin approximation of $G_1(i\omega)$ for some different model orders($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK. The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

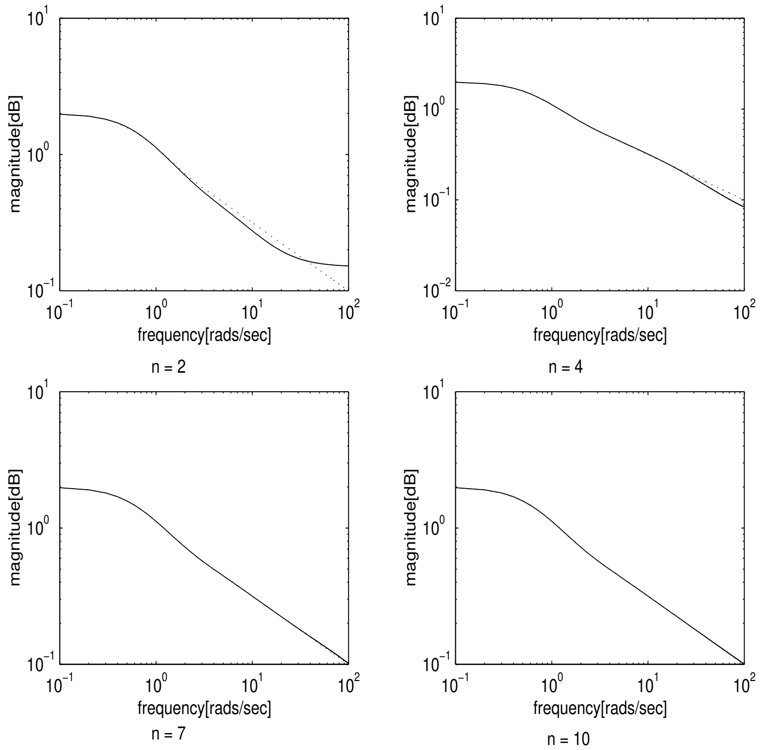

Figure 3: The Gelarkin approximation of $G_2(i\omega)$ for some different model orders($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK . The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

as the approximate system poles, in order to remove the effect of the parameter α. Since the poles close to the origin have more effect on the low frequency region than the poles that are quite away from the origin, and from Figures 2 and 3, one would expect the poles close to the origin to be approximated better than the poles away from the origin.

Figure 4 shows the relative error in the first five eigenvalues. By this we mean

$$\mid \frac{p_k^{'} - \hat{p}_k}{p_k^{'}} \mid \quad \text{for } k = 1, 2, \ldots, 5, \tag{83}$$

where $p_k^{'}$ is as defined in (81) and $\hat{p}_k$ is the approximate system pole or the eigenvalues of the matrix A_0 in (82). One can observe that for a given model order, the relative error is minimum for the pole with index $k = 1$, and the relative error increases as the index k increases to the model order n. Though in Figure 4, we have plotted only for the values $k = 1, 2, \ldots, 5$, the inference we have drawn is found to hold for higher values of k and n, and hence, we conclude that the error in the eigenvalues increases with k or alternatively the poles are approximated in a decreasing order.

Since the poles are approximated in a decreasing order, one would expect the error in the transfer functions, i.e, $\mid G_i^n(i\omega) - G_i(i\omega) \mid$ where $i = 1, 2$, to be much smaller in low frequeny region compared to the high frequency region. Possibly, the error $\mid G_i^n(i\omega) - G_i(i\omega) \mid$ for $i = 1, 2$, happens to be an increasing function of the frequency with an asymptotic upper bound at $\omega = \infty$. This hypothesis is supported by the numerical findings for higher frequencies as shown in the Figures 5 and 6.

To complete our analysis on Galerkin, we estimate the rate of convergence of our transfer functions. To estimate the rate of convergence we assume that, the error in the transfer functions is of the form

$$\mid G_i^n(i\omega) - G_i(i\omega) \mid = \frac{C_i(\omega)}{n^{a_i}} \quad \text{for all } \omega, \tag{84}$$

where $C(\omega)$ can be interpreted as a constant which depends only on ω and not on the model order n. Equation (84) implies that

$$\log \mid G_i^n(i\omega) - G_i(i\omega) \mid = -a_i \log(n) + \log(C_i(\omega)). \tag{85}$$

Note that, if the assumption (84) were to be true then, as (85) suggests, a $\log \mid G_i^n(i\omega) - G_i(i\omega) \mid$ vs $\log(n)$ plot must be a straight line with slope $-a_i$ and intercept $\log(C_i(\omega))$, for a given ω.

Figures 7, 8 and 9 show that the assumption, (84) is at least valid in the high frequency region. Also note that $C_i(\omega)$ becomes almost independent of ω at the high frequencies. In the low frequency region, the relation seems to be more complicated than (84). Nevertheless, since the error, $\mid G_i^n(i\omega) - G_i(i\omega) \mid$, in

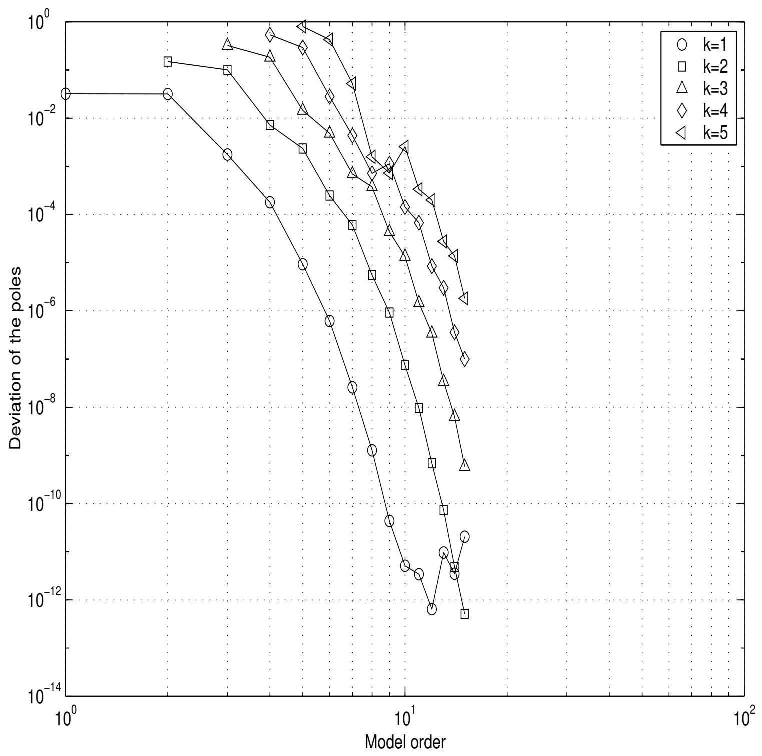

Figure 4: Deviation of the eigenvalues of the Galerkin approximate system from the true system poles, for the first five poles close to the origin.

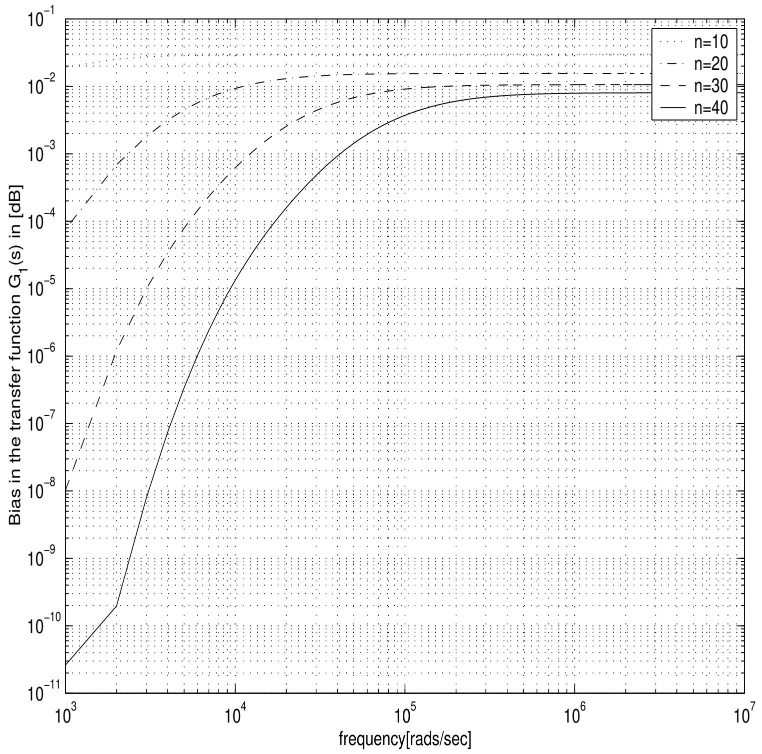

Figure 5: Absolute value of the deviation of the true transfer function $G_1(i\frac{\omega}{\alpha})$ from the Galerkin approximated transfer functions $G_n^1(i\frac{\omega}{\alpha})$ for different model orders $(n = 10,\ 20,\ 30,\ 40)$.

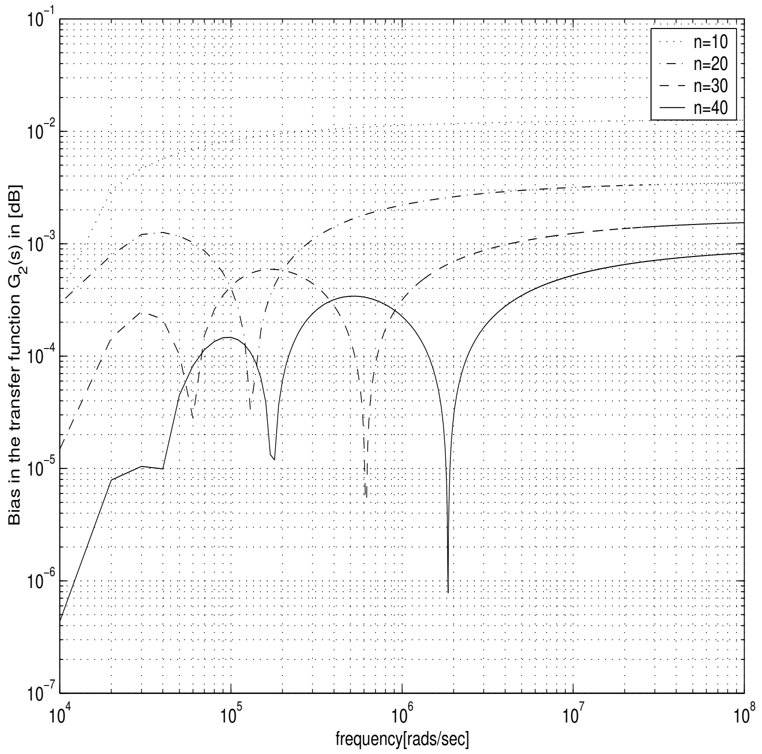

Figure 6: Absolute value of the deviation of the true transfer function $G_2(i\frac{\omega}{\alpha})$ from the Galerkin approximated transfer functions $G_n^2(i\frac{\omega}{\alpha})$ for different model orders ($n = 10,\ 20,\ 30,\ 40$).

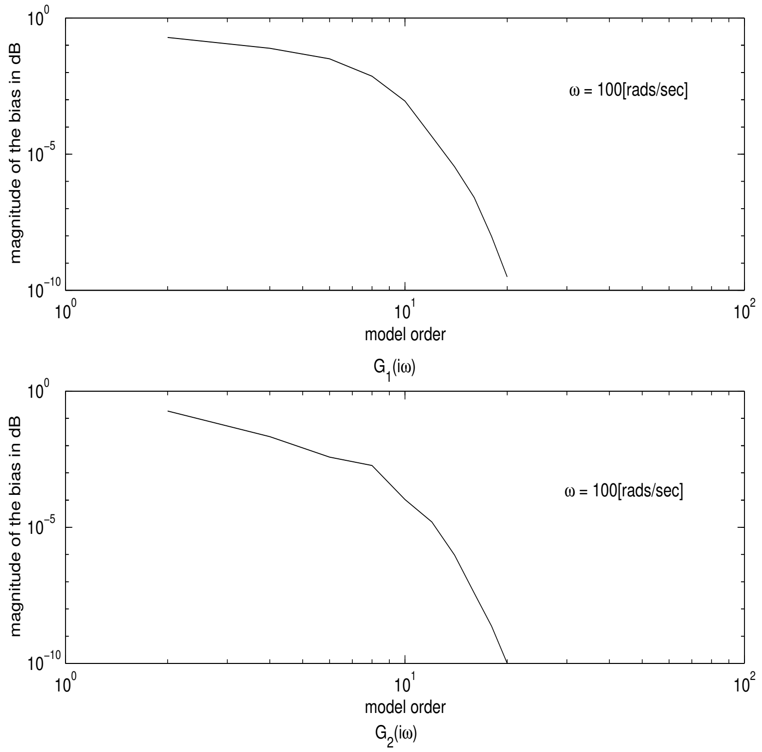

Figure 7: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 100$.

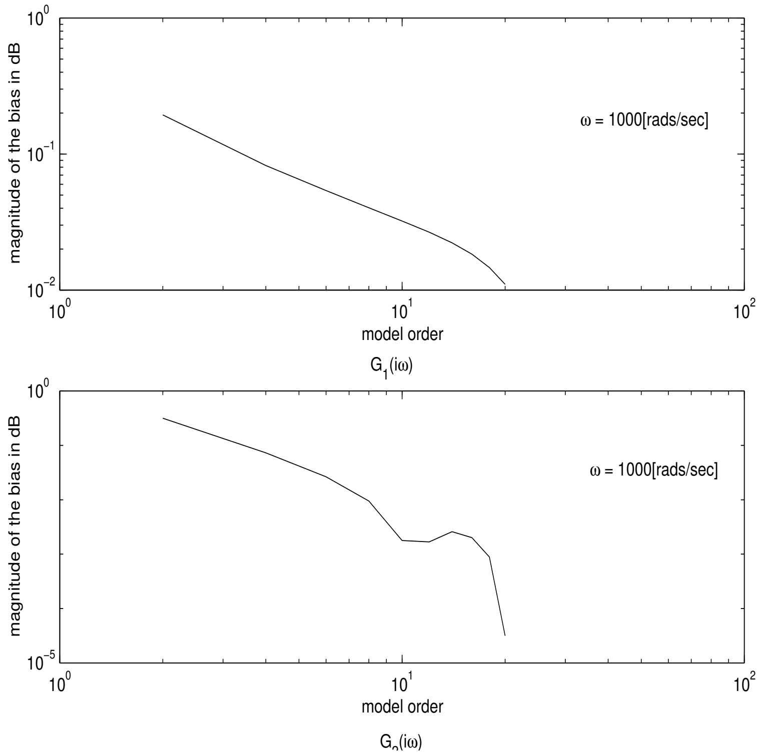

Figure 8: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 10000$.

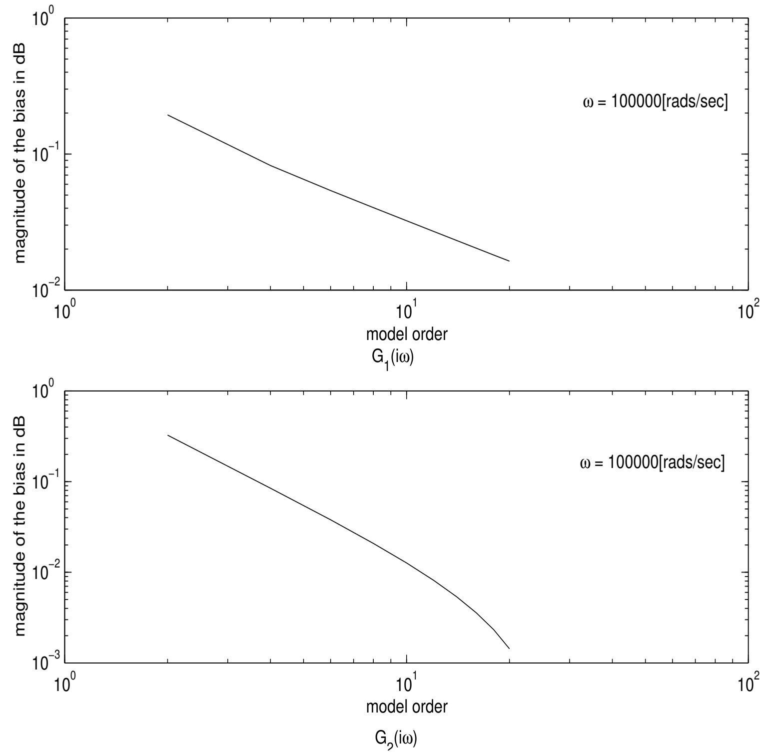

Figure 9: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 1000000$.

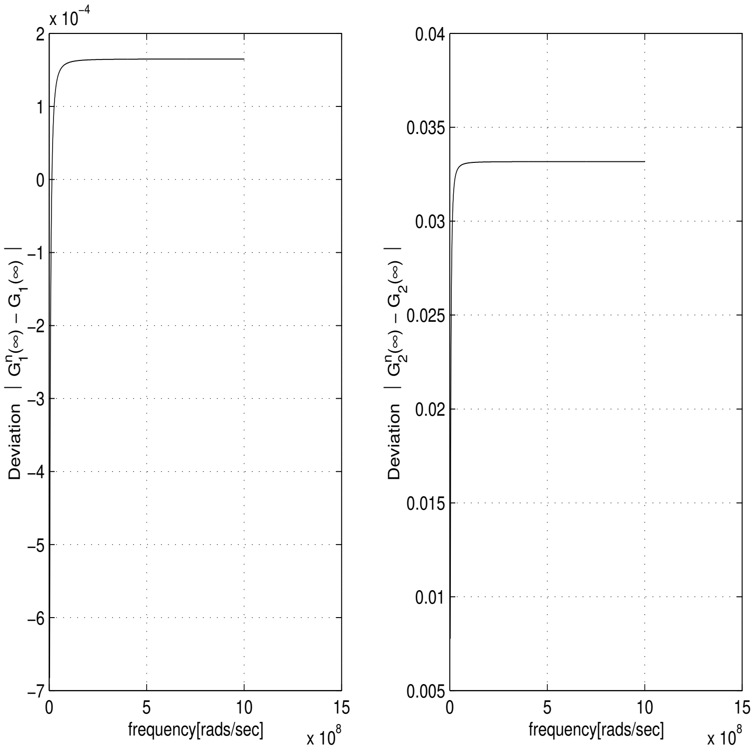

Figure 10: Deviation of the Galerkin approximate transfer function from the true transfer function at high freq uencies for $n = 100$.

transfer functions is an increasing function of the frequency with an asymptotic upper bound, a worst case secenario of the form

$$| G_i^n(i\omega) - G_i(i\omega) | \leq \frac{C_i}{n^{a_i}} \quad \text{for all } \omega, \text{ and } i = 1, 2 \tag{86}$$

can be obtained by estimating the values of C_i and a_i, for $i = 1, 2$, in the high frequency region. Numerical experiments suggest the following values $a_1 = 1$, $a_2 = 2$, $C_1 = 3.317$ and $C_2 = 1.65$. Numerical values of a_1 and a_2 can also verified from the Figure 9 and the values of C_1 and C_2 can be seen in the Figure 10.

Note that, even though the Galerkin method looks quite natural in its approach, it is a very laborious procedure. Even to get the state space equations, one has to play around with both the basis functions and the given problem to make them compatible with each other. In the coming subsections we look at certain approximate schemes that are much simpler than the Galerkin method, but based on the same theme.

4.3 Tau method

A popular method that is simple to use and is extensively used,[2], is called Tau method. Tau method is closely related to the Galerkin method but does not impose any conditons on the basis functions.

In Tau method we seek a solution of the form (34), but unlike the Galerkin method, the basis functions need not satisfy the boundary conditions. Unlike the Galerkin method, where the solution is projected onto the space spanned by $\{\phi_j\}_{j=0}^{n}$, in the Tau method the solution is projected onto the space spanned by $\{\phi_j\}_{j=0}^{n-2}$, i.e,

$$\langle \frac{\partial T_n(x,t)}{\partial t} - \alpha \frac{\partial^2 T_n(x,t)}{\partial x^2}, \phi_j(x) \rangle = 0 \qquad j = 0, 1, \ldots, n-2. \tag{87}$$

and the boundary conditions are imposed by letting

$$a_{n-1}(t) = \frac{-\frac{q_i(t)}{\kappa} - \sum_{j=0}^{n-2} a_j(t)\phi_j'(0) - a_n(t)\phi_n'(0)}{\phi_{n-1}'(0)} \tag{88}$$

$$a_n(t) = \frac{T_e(t) - \sum_{j=0}^{n-1} a_j(t)\phi_j(d)}{\phi_n(d)} \tag{89}$$

Hence we have $n-2$ differential equations (87) and 2 algebraic equations (88) and (89).

Note that (87) can also be written as

$$\sum_{j=0}^{n} \dot{a}_j(t) \langle \phi_j(x), \phi_k(x) \rangle = \alpha \sum_{j=0}^{n} a_j(t) \langle \phi_j''(x), \phi_k(x) \rangle \; k = 0, 1, \ldots, n-2. \tag{90}$$

Since our goal is to use Chebyshev polynomials, we use the rescaled problem (37) - (39). Let

$$T(x,t) = \sum_{j=0}^{n} a_j(t)C_j(x) \tag{91}$$

be the desired approximate solution, where $C_j(x)$, $j = 0,1,2,\ldots,n$ are the Chebyshev polynomials. Replacing $\phi_j(x)$ by $C_j(x)$, α by $\frac{4\alpha}{d^2}$ and κ by $\frac{2\kappa}{d}$ in the equations (88), (89) and (90), and using the fact that Chebyshev polynomials are orthogonal, see Lemma 2, we get

$$\dot{a}_k(t) = \frac{4\alpha}{d^2}\frac{d_k}{\pi}\sum_{j=0}^{n} a_j(t)\langle C_j''(x), C_k(x)\rangle \quad k = 0,1,\ldots,n-2 \tag{92}$$

$$a_{n-1}(t) = \frac{-\frac{q_i(t)d}{2\kappa} - \sum_{j=0}^{n-2} a_j(t)C_j'(-1) - a_n(t)C_n'(-1)}{C_{n-1}'(-1)} \tag{93}$$

$$a_n(t) = \frac{T_e(t) - \sum_{j=0}^{n-1} a_j(t)C_j(1)}{C_n(1)}. \tag{94}$$

Solving for $a_n(t)$ and $a_{n-1}(t)$ from (93) and (94), and using (28) and (32) we further get

$$\begin{aligned} a_{n-1}(t) &= \frac{T_e(t)n^2}{2n^2-2n+1} + \sum_{j=0}^{n-2}\frac{(-1)^{j+n}j^2 - n^2}{2n^2-2n+1}a_j(t) \\ &\quad + \frac{\frac{q_i(t)d}{2\kappa}(-1)^{n-1}}{2n^2-2n+1} \end{aligned} \tag{95}$$

$$\begin{aligned} a_n(t) &= \frac{T_e(t)(n-1)^2}{2n^2-2n+1} + \sum_{j=0}^{n-2}\frac{(-1)^{j+n-1}j^2 - (n-1)^2}{2n^2-2n+1}a_j(t) \\ &\quad + \frac{\frac{q_i(t)d}{2\kappa}(-1)^{n-2}}{2n^2-2n+1} \end{aligned} \tag{96}$$

Plugging (95) and (96) into (92) gives

$$\begin{aligned} \dot{a}_k(t) &= \frac{4\alpha}{d^2}T_e(t)\left(\frac{(n-1)^2 z_{1k} + n^2 z_{2k}}{2n^2-2n+1}\right) \\ &\quad + \frac{4\alpha}{d^2}\sum_{j=0}^{n-2} a_j(t)\left(\frac{d_k}{\pi}\langle C_j''(x), C_k(x)\rangle - \frac{w_{1j}z_{1k} + w_{2j}z_{2k}}{2n^2-2n+1}\right) \\ &\quad + \frac{4\alpha}{d^2}\frac{q_i(t)d}{2\kappa}\left(\frac{(-1)^{n-2}z_{1k}}{2n^2-2n+1} + \frac{(-1)^{n-1}z_{2k}}{2n^2-2n+1}\right) \end{aligned} \tag{97}$$

where

$$
\begin{aligned}
z_{1k} &= \frac{d_k}{\pi}\langle C_n''(x), C_k(x)\rangle \quad k = 0, 1, \ldots, n-2 && (98)\\
z_{2k} &= \frac{d_k}{\pi}\langle C_{n-1}''(x), C_k(x)\rangle \quad k = 0, 1, \ldots, n-2 && (99)\\
w_{1j} &= \frac{(-1)^{j+n-1}j^2 - (n-1)^2}{2n^2 - 2n + 1} \quad j = 0, \ldots, n-2 && (100)\\
w_{2j} &= \frac{(-1)^{j+n}j^2 - n^2}{2n^2 - 2n + 1} \quad j = 0, \ldots, n-2. && (101)
\end{aligned}
$$

Note that the scalar products involved in (97), (98) and (99) involve polynomials whose product is of degree less than $2n-1$. Hence one can use Gauss Lobatoo formula, (29), to evaluate them.

The desired output is

$$
T_i(t) = T_n(-1, t) = \sum_{j=0}^{n} a_j(t) C_j(-1), \tag{102}
$$

which using (95), (96) and (28), can be rewritten as

$$
\begin{aligned}
T_i(t) &= \frac{(-1)^{n-1}(2n-1)}{2n^2 - 2n + 1} T_e(t)\\
&\quad + \sum_{j=0}^{n-2}\left((-1)^j + \frac{2j^2(-1)^{j-1} - (-1)^{n-1}(2n-1)}{2n^2 - 2n + 1}\right) a_j(t)\\
&\quad + \frac{d}{\kappa(2n^2 - 2n + 1)} q_i(t)
\end{aligned} \tag{103}
$$

Setting $a = (a_1(t), \ldots, a_{n-1}(t))^T$ and rewriting (97) and (103) in the state space form we get

$$
\begin{aligned}
\dot{x} &= Ax + Bu\\
y &= Cx + Du
\end{aligned} \tag{104}
$$

where

$$a_{kj} = \frac{4\alpha}{d^2}\sum_{j=0}^{n-2}\left(\frac{d_k}{\pi}\langle C_j''(x), C_k(x)\rangle - \frac{w_{1j}z_{1k}+w_{2j}z_{2k}}{2n^2-2n+1}\right) \tag{105}$$

$$j,k = 1,\dots,n-2$$

$$b_{k1} = -\frac{2\alpha}{d}\frac{1}{\kappa}\left(\frac{(-1)^{n-2}z_{1k}}{2n^2-2n+1}+\frac{(-1)^{n-1}z_{2k}}{2n^2-2n+1}\right) \quad k=1,\dots,n-2 \tag{106}$$

$$b_{k2} = \frac{4\alpha}{d^2}\left(\frac{(n-1)^2z_{1k}+n^2z_{2k}}{2n^2-2n+1}\right) \quad k=1,\dots,n-2 \tag{107}$$

$$c_{1k} = \left((-1)^k+\frac{2k^2(-1)^{k-1}-(-1)^{n-1}(2n-1)}{2n^2-2n+1}\right) \quad k=1,\dots,n-2 \tag{108}$$

$$D = \left(\frac{(-1)^{n-1}(2n-1)}{2n^2-2n+1}, \frac{d}{\kappa(2n^2-2n+1)}\right). \tag{109}$$

As we did in the previous section, we rewrite (104) in the standard input output form by taking the Laplace transforms and plot the approximate transfer functions along with their respective true transfer functions.

Figures 11 and 12 show the true transfer functions along with their approximations. As before, we can see that the transfer functions are well approximated at low frequencies and the approximation deteriorates as the frequency increases. From Figures 13, 14 and 15 we can see that, convergence is very rapid at low frequencies and stablises to a constant rate of the type (86) as the frequency increases. We purposefully avoid an eigenvalue analysis here, since the results are almost similar to the ones obtained in the previous section. Figures 16 and 17 show the deviation of the approximate transfer functions from the true ones. One can see that the error in the transfer function, for a given n, is an increasing function of the frequency (for high ω) and converges to an asymptotic upper bound.

As in the previous section we can obtain an expression for the rate of convergence. In this case we are more fortunate, as we have a very simple expression for the D vector in (104).

Note that

$$\mid G_i^n(s) - G_i(s) \mid \;\leq\; \mid G_i^n(\infty) - G_i(\infty), \mid \quad \text{for all } s \text{ and } i=1,2. \tag{110}$$

Since

$$G_i(\infty) = 0 \quad \text{for } i=1,2 \tag{111}$$

and

$$G_i^n(\infty) = D_i \quad \text{for } i=1,2, \tag{112}$$

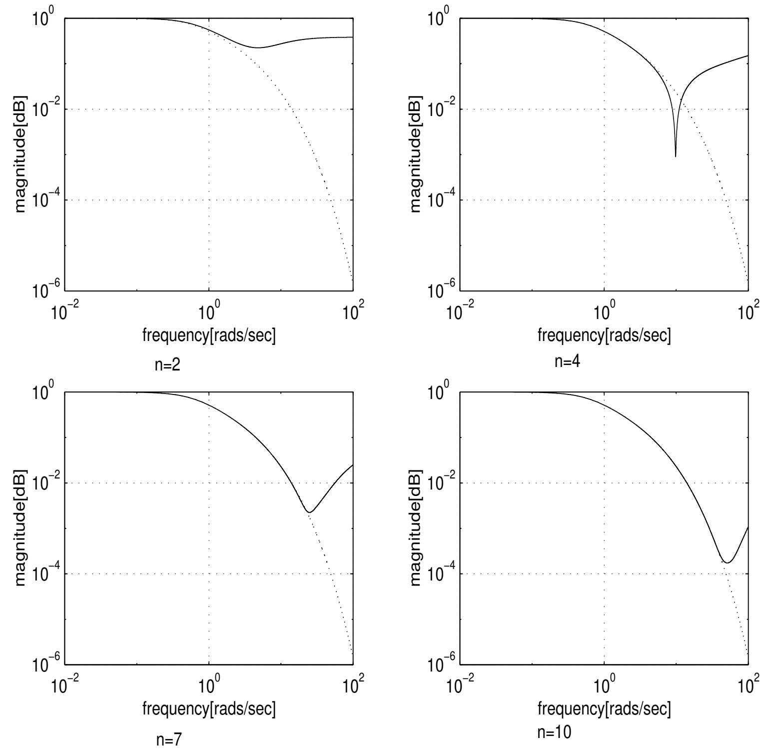

Figure 11: The Tau approximation of $G_1(i\omega)$ for some different model orders($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK. The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

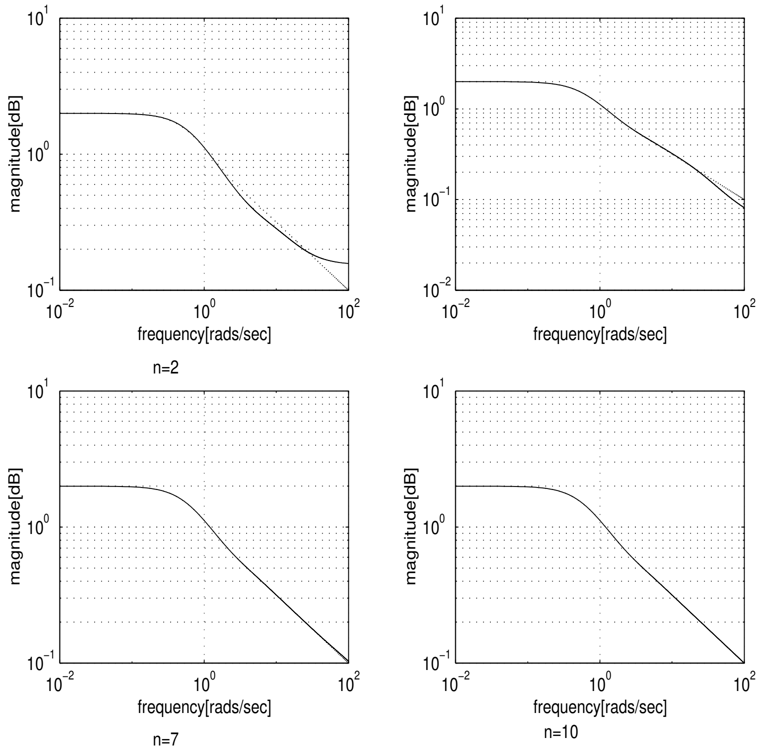

Figure 12: The Tau approximation of $G_2(i\omega)$ for some different model orders($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK . The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

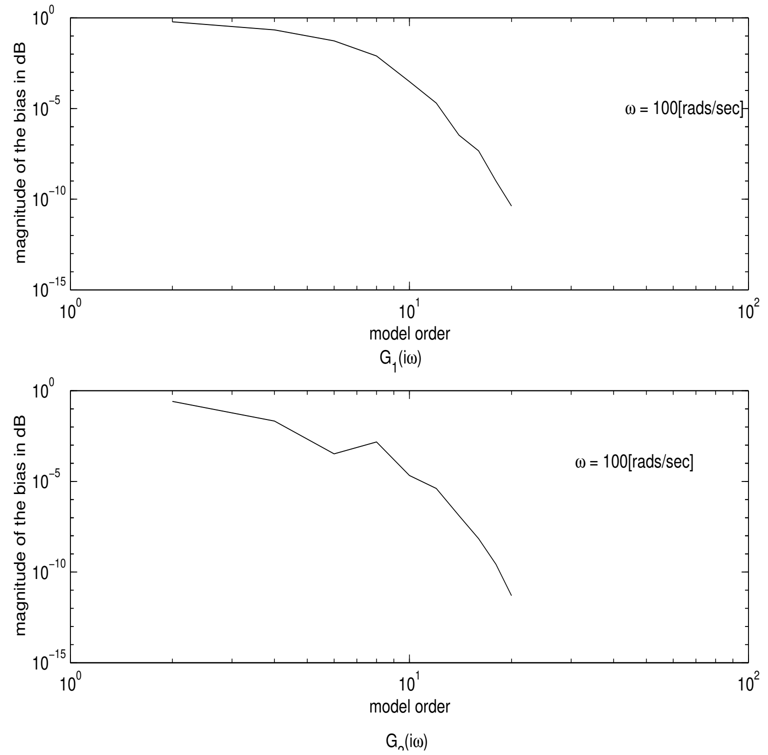

Figure 13: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 100$.

from (109) we have

$$\begin{aligned} \mid G_1^n(i\omega) - G_1(i\omega) \mid &\leq \frac{(2n-1)}{2n^2 - 2n + 1} \quad \text{for all } \omega, && (113) \\ &\approx \frac{1}{n} \quad \text{for large n} && (114) \end{aligned}$$

and

$$\begin{aligned} \mid G_2^n(i\omega) - G_2(i\omega) \mid &\leq \frac{d}{\kappa(2n^2 - 2n + 1)} \quad \text{for all } \omega. && (115) \\ \frac{\kappa}{d} \mid G_2^n(i\omega) - G_2(i\omega) \mid &\approx \frac{1}{2n^2} \quad \text{for large } n && (116) \end{aligned}$$

Thus we can see that from (114) and (116), $G_1^n(s)$ converges $G_1(s)$ at a rate faster than $O(\frac{1}{n})$ and $G_2^n(s)$ converges $G_2(s)$ at a rate faster than $O(\frac{1}{n^2})$. Note that the straight lines in Figure 15 have slopes that are almost equal to the value predicted by (114) and (116) and the constants predicted by (114) and (116) are almost close the asymptotes seen in Figure 18.

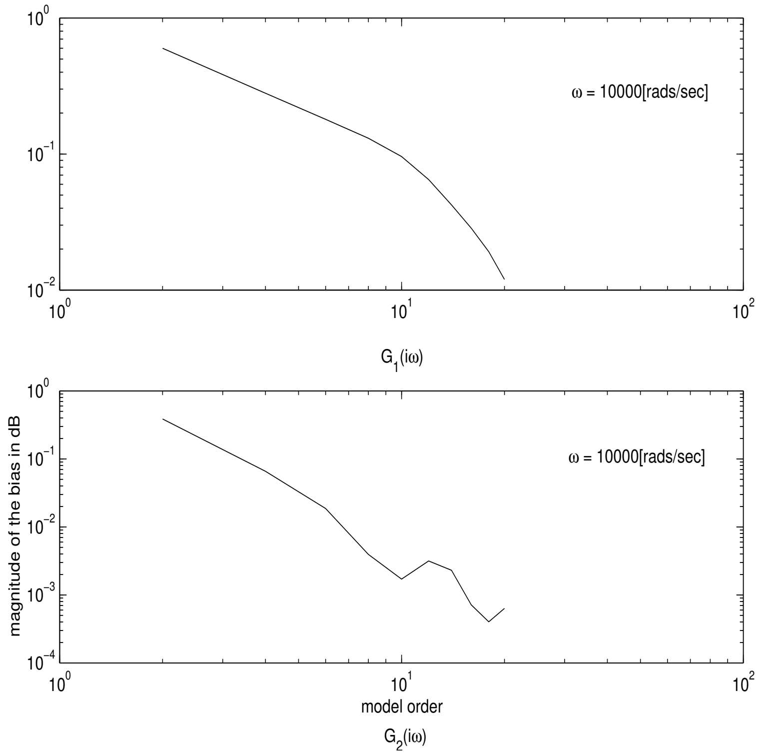

Figure 14: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 10000$.

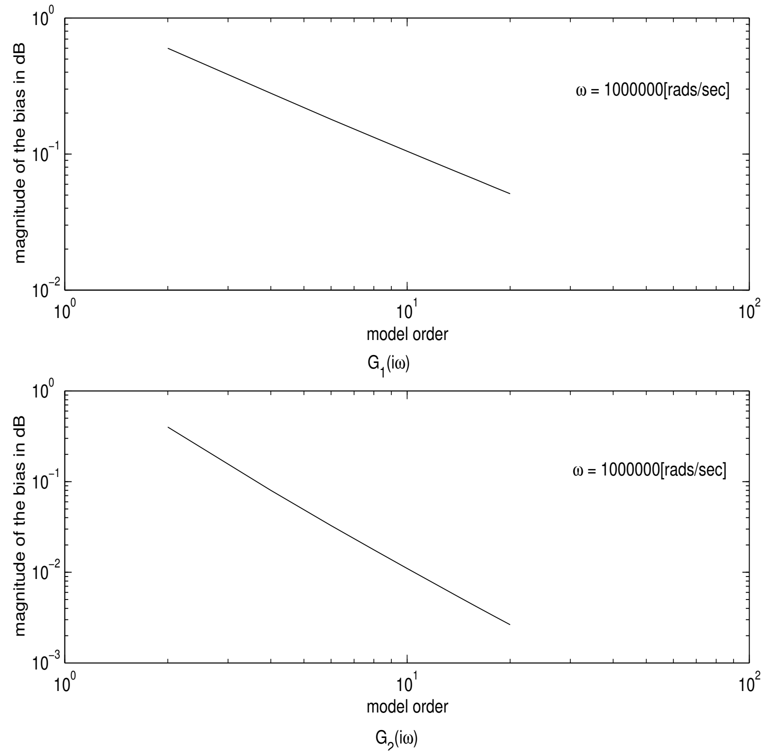

Figure 15: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \dots, 20$ at $\omega = 1000000$.

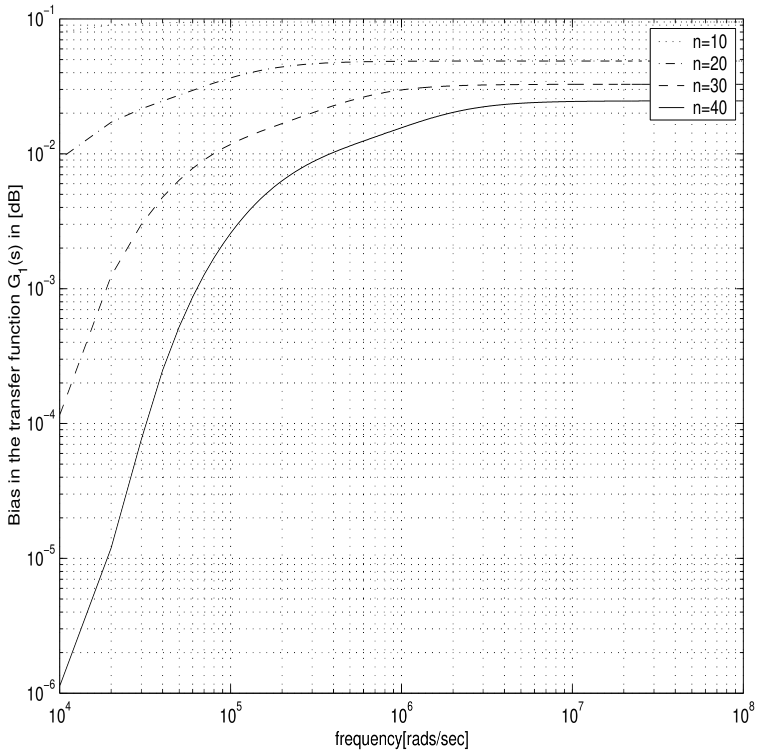

Figure 16: Absolute value of the deviation of the true transfer function $G_1(i\frac{\omega}{\alpha})$ from the Tau approximated transfer functions $G_n^1(i\frac{\omega}{\alpha})$ for different model orders ($n = 10,\ 20,\ 30,\ 40$).

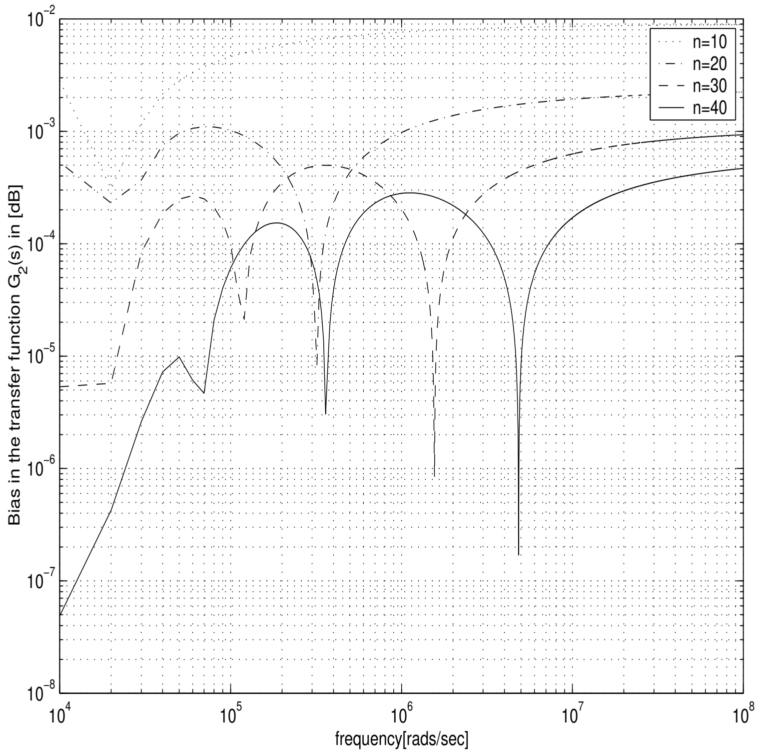

Figure 17: Absolute value of the deviation of the true transfer function $G_2(i\frac{\omega}{\alpha})$ from the Tau approximated transfer functions $G_n^2(i\frac{\omega}{\alpha})$ for different model orders ($n = 10,\ 20,\ 30,\ 40$).

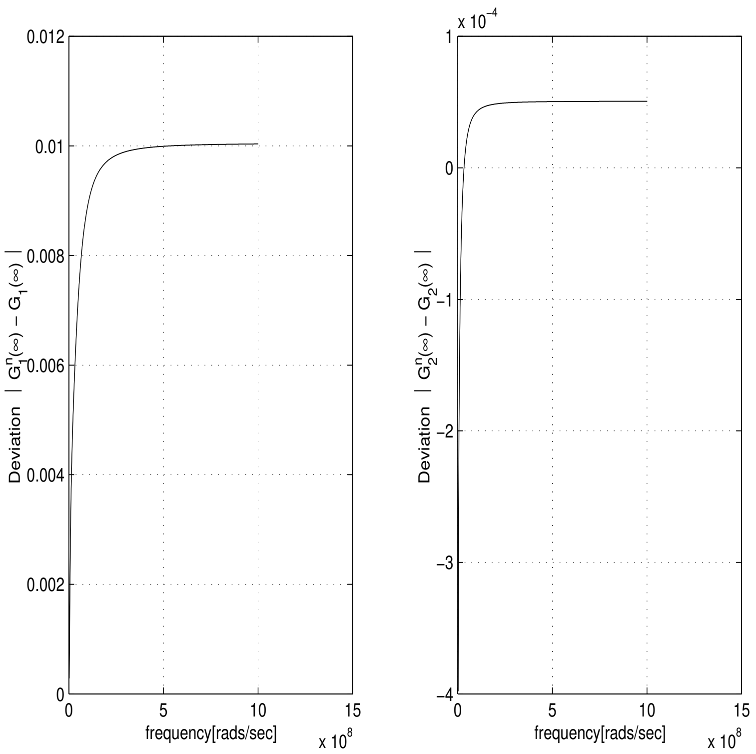

Figure 18: Deviation of the Tau approximate transfer function from the true transfer function at high frequencies for $n = 100$. Note that the factor $\frac{\kappa}{d}$ which appears in $G_2^n(\infty)$ has been normalised to 1.

4.4 Collocation method

There are many types of collocation, but here we use a collocation that is the closest to the Chebychev Galerkin and Chebychev Tau methods we have mentioned above. For other types of collocation methods we refer to [3],[8].

Let $f(x)$ be a function such that

$$| \int_{-1}^{1} \frac{| f(x) |}{\sqrt{1-x^2}} dx | < \infty \tag{117}$$

Then the Galerkin projection of $f(x)$ in the space $\mathcal{P}_{n+1}$, the space of polynomials of degree less that or equal to $n+1$ with inner product (24), is given by

$$f_{n+1}(x) = \sum_{j=0}^{n+1} a_j C_j(x) \tag{118}$$

where

$$a_j = \frac{2}{\pi d_j} \int_{-1}^{1} \frac{f(x) C_j(x)}{\sqrt{1-x^2}} dx \tag{119}$$

and d_j as defined in Lemma 2.

By applying Gauss Lobatoo formula, (29), to (119), we have

$$a_j \approx \frac{2}{(n+1)d_j} \sum_{k=0}^{n+1} \frac{f(x_k) C_j(x_k)}{c_k} \quad j = 0, \ldots, n+1. \tag{120}$$

where x_k and c_k are as defined in (26) and (30), respectively. Replacing for a_j in (118) we have

$$f_{n+1}(x) = \sum_{j=0}^{n+1} f(x_j) l_j(x) \tag{121}$$

where

$$l_j(x) = \frac{2}{(n+1)c_j} \sum_{k=0}^{n+1} \frac{C_k(x) C_k(x_j)}{d_k}. \tag{122}$$

Thus we have a new projection scheme (121) with $l_j(x)$ as our basis functions. Note that this projection in to the space $\mathcal{P}_{n+1}$ is not exactly but approximately orthogonal.

Note that if $f(x)$ is a polynomial of degree less than or equal to n, then the product $f(x)C_j(x)$ is polynomial of degree less than $2n+1$ for $j = 0, \ldots, n+1$, hence Gauss Lobatoo formula, (29), is exact in (120).

Since Chebyshev polynomials are defined only in the interval $[-1,1]$, $l_j(x)$ are also defined on $[-1,1]$, hence we have to use the rescaled problem (37) - (39).

Let $T_n(x,t)$ be the approximate solution of (37) - (39), i.e.,

$$T_n(x,t) = \sum_{j=0}^{n+1} T_n(x_j,t) l_j(x) \tag{123}$$

Note that in this method the expansion coefficients are actually point values of the approximate solution.

In a general collocation method the state space equations are obtained by forcing the approximate solution to satisfy the PDE at a finite number of points called collocation points. In our case, the approximate solution (123) is forced to satisfy the PDE (37) - (39) at x_j, $j = 0,1,2,\ldots,n+1$, where x_j are as defined in (26). Thus

$$\begin{aligned}
\frac{\partial T_n(x,t)}{\partial t}\,|_{x=x_j} &= \frac{4\alpha}{d^2}\sum_{k=0}^{n+1} T_n(x_k,t) l_k''(x)_{|x=x_j} \quad j=1,\ldots,n && (124)\\
q_i &= -\frac{2\kappa}{d}\sum_{k=0}^{n+1} T_n(x_k,t) l_k'(x_{n+1}) && (125)\\
T_e(t) &= T_n(x_0,t) && (126)
\end{aligned}$$

Equation (125) implies that the desired output $T_i(t)$, which is equal to $T_n(x_{n+1},t)$ is given by

$$T_i(t) = T_n(x_{n+1},t) = \frac{-\frac{q_i d}{2\kappa} - \sum_{k=1}^{n} T_n(x_k,t) l_k'(x_{n+1}) - T_e(t) l_0'(x_{n+1})}{l_{n+1}'(x_{n+1})} \tag{127}$$

Plugging (127) into (124) we get

$$\begin{aligned}
\dot{T}_n(x_j,t) &= \frac{-2\alpha q_i(t) l_{n+1}''(x_j)}{d\kappa l_{n+1}'(x_{n+1})} + \frac{4\alpha}{d^2}\sum_{k=1}^{n} T_n(x_k,t)\left(l_k''(x_j) - \frac{l_{n+1}''(x_j) l_k'(x_{n+1})}{l_{n+1}'(x_{n+1})}\right)\\
&\quad + T_e(t)\frac{4\alpha}{d^2}\left(l_0''(x_j) - \frac{l_{n+1}''(x_j) l_0'(x_{n+1})}{l_{n+1}'(x_{n+1})}\right)
\end{aligned} \tag{128}$$

Setting $x = (T_n(x_1),\ldots,T_n(x_n))$ and rewriting the system of equations (125), (126) and (128) in the standard state space form gives

$$\begin{aligned}
\dot{x} &= Ax + Bu\\
y &= Cx + Du
\end{aligned} \tag{129}$$

where

$$a_{ij} = \frac{4\alpha}{d^2}\left(l_j^{''}(x_i) - \frac{l_{n+1}^{''}(x_i)l_j^{'}(x_{n+1})}{l_{n+1}^{'}(x_{n+1})}\right) \qquad i,j = 1,\ldots,n-1 \tag{130}$$

$$b_{j1} = \frac{-2\alpha}{d\kappa}\frac{l_{n+1}^{''}(x_j)}{l_{n+1}^{'}(x_{n+1})} \qquad j = 1,\ldots,n-1 \tag{131}$$

$$b_{j2} = \frac{4\alpha}{d^2}\left(l_0^{''}(x_j) - \frac{l_{n+1}^{''}(x_j)l_0^{'}(x_{n+1})}{l_{n+1}^{'}(x_{n+1})}\right) \qquad j = 1,\ldots,n-1 \tag{132}$$

$$C = \left(-\frac{l_1^{'}(x_{n+1})}{l_{n+1}^{'}(x_{n+1})},\ldots,-\frac{l_n^{'}(x_{n+1})}{l_{n+1}^{'}(x_{n+1})}\right) \tag{133}$$

$$D = \left(\frac{l_0^{'}(x_{n+1})}{l_{n+1}^{'}(x_{n+1})},-\frac{d}{2\kappa l_{n+1}^{'}(x_{n+1})}\right) \tag{134}$$

By rewriting (129) in the input output form, as we have done in the previous subsections, we would get the approximate transfer functions $G_1^n(s)$ and $G_2^n(s)$ for the Chebyshev collocation. Figures 19 and 20 show the approximate transfer functions along with their respective true transfer functions. As before one can see that $G_1^n(s)$ is not well approximated at high frequencies. One can see from the Figures 21 and 22, that the error in the transfer functions is an increasing function of the frequency (for a large ω) and has an upper bound. Thus the behaviour is almost similar to the approximate models in the previous sections. Hence, as we did in the previous sections, a worst case analysis for the rate of convergence can be done by using the D matrix.

Note that, form (122), we have

$$l_{n+1}^{'}(x) = \frac{2}{(n+1)c_{n+1}}\sum_{k=0}^{n+1}\frac{C_k^{'}(x)C_k(x_{n+1})}{d_k} \quad \text{and} \tag{135}$$

$$l_0^{'}(x) = \frac{2}{(n+1)c_0}\sum_{k=0}^{n+1}\frac{C_k^{'}(x)C_k(x_0)}{d_k} \tag{136}$$

from (23), (30) and (32) we know that

$$c_0 = c_{n+1} = 2, \tag{137}$$

$$C_k^{'}(x_{n+1}) = (-1)^{k-1}k^2 \tag{138}$$

$$C_k(x_{n+1}) = (-1)^k \quad \text{and} \tag{139}$$

$$C_k(x_0) = 1. \tag{140}$$

Hence

$$| l'_{n+1}(x_{n+1}) | = \frac{1}{n+1}\sum_{k=1}^{n+1} k^2 \tag{141}$$

$$| l'_0(x_{n+1}) | = | \frac{1}{n+1}\sum_{k=1}^{n+1}(-1)^{k-1}k^2 | . \tag{142}$$

From (134), (141) and (142) we can see that

$$| D_1 | = | \frac{\sum_{k=1}^{n+1}(-1)^{k-1}k^2}{\sum_{k=1}^{n+1} k^2} | . \tag{143}$$

$$| D_2 | = \frac{d(n+1)}{2\kappa\sum_{k=1}^{n+1} k^2}. \tag{144}$$

One can show that, for a large n

$$| \sum_{k=1}^{n+1}(-1)^{k-1}k^2 | \approx \frac{n^2}{2}. \tag{145}$$

A proof for (145) is provided in the appendix. Since

$$\sum_{k=1}^{n+1} k^2 = \frac{(n+1)(n+2)(2n+3)}{6}, \tag{146}$$

for a large n we have

$$| D_1 | \approx \frac{6n^2}{2(n+1)(n+2)(2n+3)} \tag{147}$$

and

$$| D_2 | = \frac{3d}{\kappa(n+2)(2n+3)}, \tag{148}$$

which implies that

$$| D_1 | \approx \frac{3}{2n} \tag{149}$$

$$\frac{\kappa}{d} | D_2 | \approx \frac{3}{2n^2}. \tag{150}$$

Hence $G_1^n(s)$ converges at a rate faster than $O(\frac{1}{n})$ and $G_2^n(s)$ converges at a rate faster than $O(\frac{1}{n^2})$. It can also be seen in the Figures 23, 24 and 25 that the constants predicted by (149) and (150) are almost close to the ones in Figure 26.

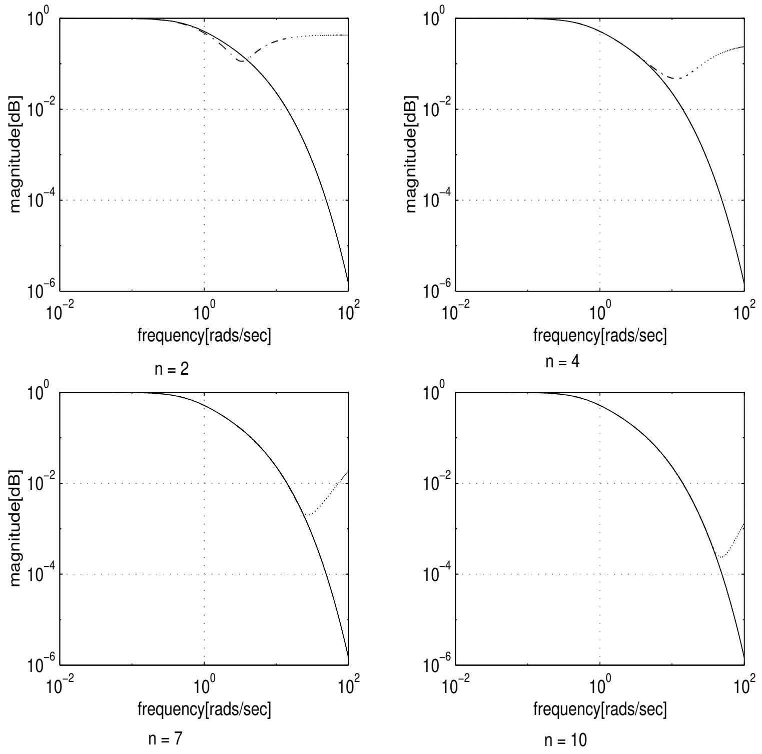

Figure 19: The Chebyshev collocation approximation of $G_1(i\omega)$ for some different model orders ($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK. The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

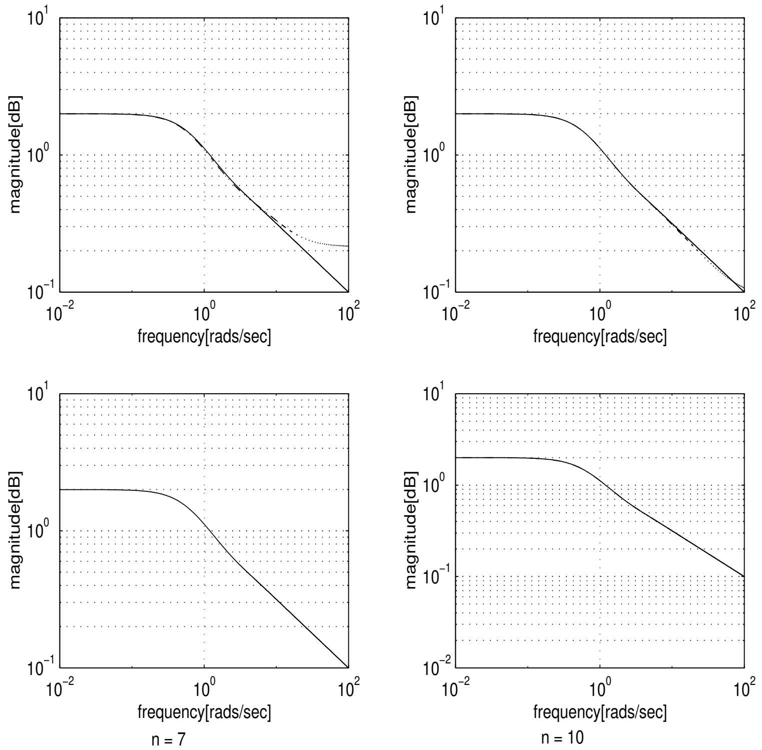

Figure 20: The Chebyshev collocation approximation of $G_2(i\omega)$ for some different model orders($n = 2,\ 4\ 7,\ 10$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK . The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

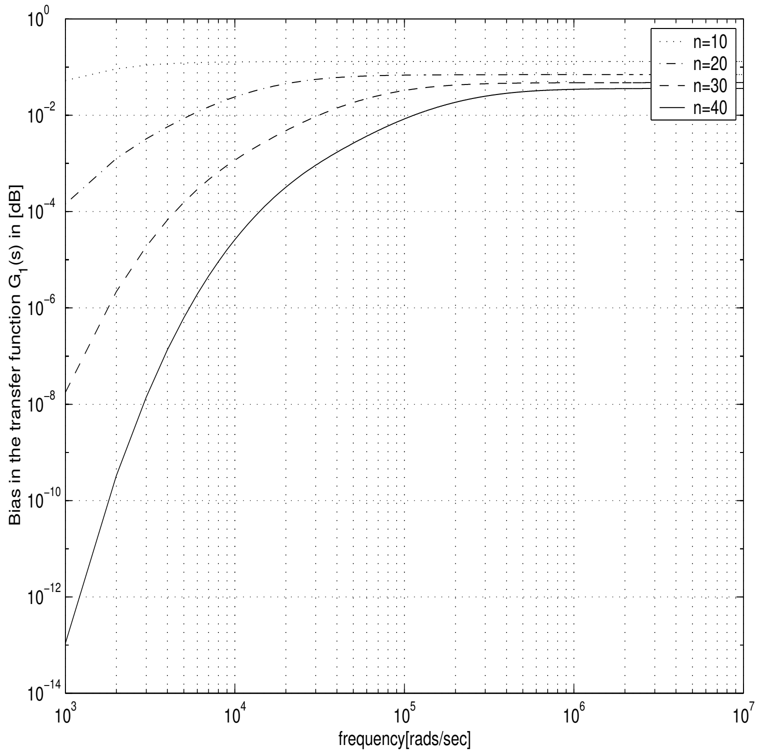

Figure 21: Absolute value of the deviation of the true transfer function $G_1(i\frac{\omega}{\alpha})$ from the Chebyshev collocation approximated transfer function $G_n^1(i\frac{\omega}{\alpha})$ for different model orders ($n = 10,\ 20,\ 30,\ 40$).

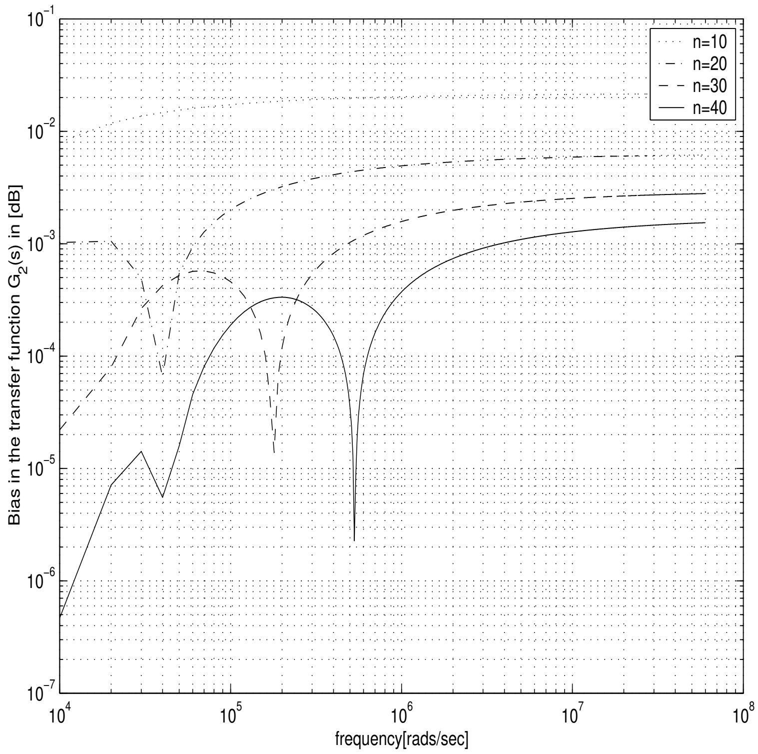

Figure 22: Absolute value of the deviation of the true transfer function $G_2(i\frac{\omega}{\alpha})$ from the Chebyshev collocation approximated transfer function $G_n^2(i\frac{\omega}{\alpha})$ for different model orders ($n = 10,\ 20,\ 30,\ 40$).

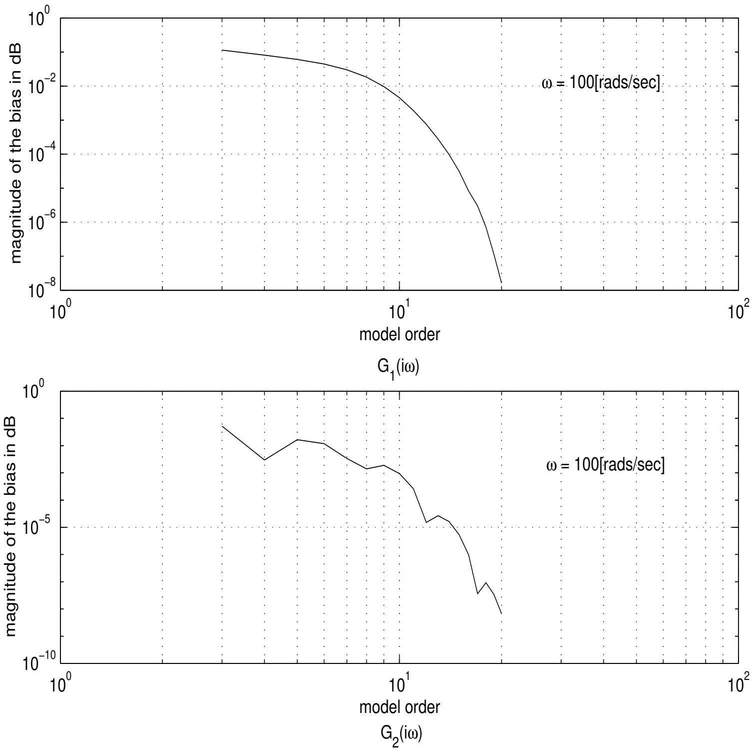

Figure 23: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 100$.

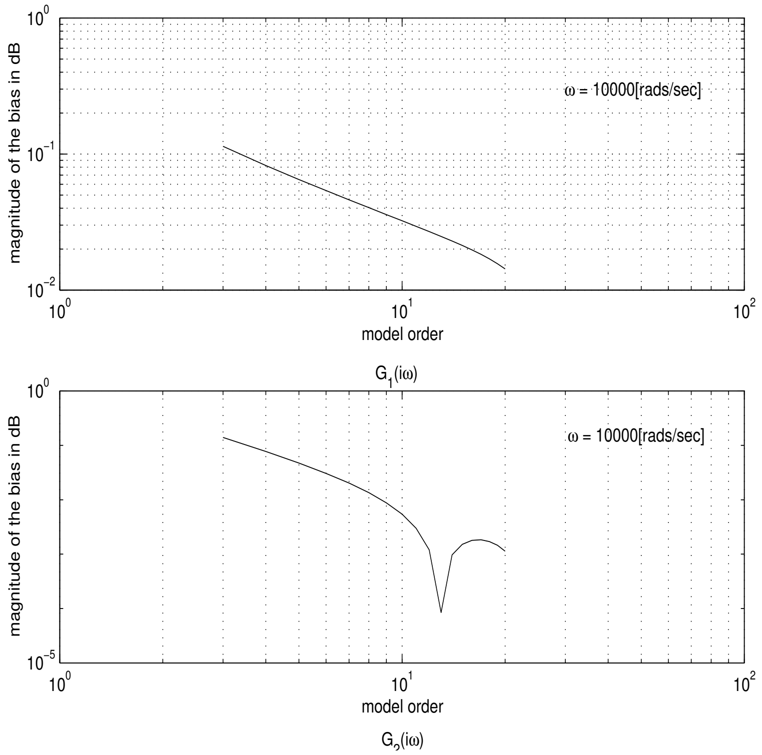

Figure 24: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 10000$.

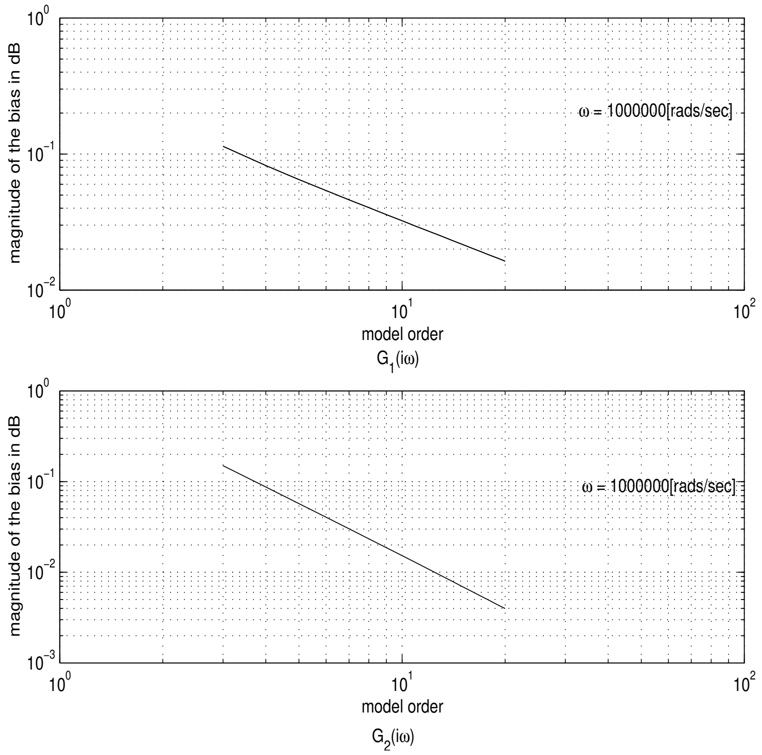

Figure 25: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form the true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 1000000$.

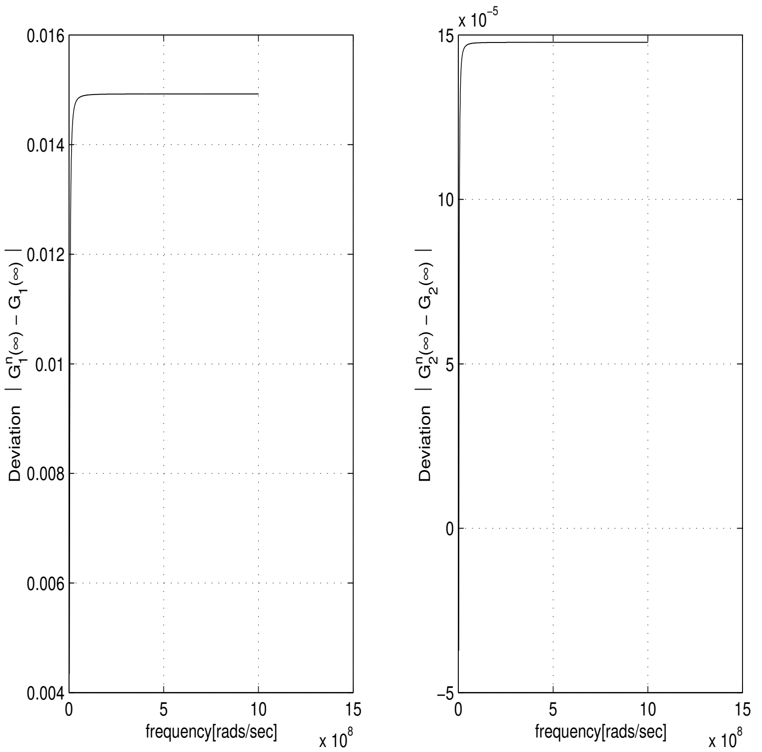

Figure 26: Deviation of the Chebyshev collocation approximate transfer function from the true transfer function at high frequencies for $n = 100$. Note that the factor $\frac{\kappa}{d}$ which appears in $G_2^n(\infty)$ has been normalised to 1.

4.5 Chebyshev interpolation

This method was recently proposed by Gottlieb, see [2]. This method runs close to collocation method mentioned in the previous section, with a subtle difference which will become apparent as we proceed

Let

$$g_j(x) = \frac{2}{(n+1)c_j} \sum_{k=0}^{n+1} \frac{C_k(x)C_k(x_j)}{c_k} \tag{151}$$

where c_j and x_j are as defined in (30) and (26), respectively.

The approximate solution of the PDE with $g_j(x)$ as the basis functions is

$$T_n(x,t) = \sum_{j=0}^{n+1} a_j(t) g_j(x). \tag{152}$$

One can see that the only difference between $g_j(x)$, (151), and $l_j(x)$ defined in (122) is that, in $l_j(x)$ the numerator $C_k(x)C_k(x_j)$ is divided by d_k, but in $g_j(x)$ the numerator $C_k(x)C_k(x_j)$ is divided by c_k. This small difference makes life easier while numerically computing the approximate model. We state the following lemma, which would enable us to see how this factor c_k could make such a big difference when computing the solution.

Lemma 4. *For the polynomials $g_j(x)$ as defined in (151), it holds,*

$$g_j(x_k) = \delta_{jk} \quad j,k = 0,\dots,n+1 \tag{153}$$

where x_k is given by (26), and δ_{jk} is the Kronecker delta.

Proof. Refer to [5] for a proof.

Observe from (152) that

$$T_n(x_k,t) = \sum_{j=0}^{n+1} a_j(t) g_j(x_k). \tag{154}$$

Since $g_j(x_k) = \delta_{jk}$ from Lemma 4, we can see that (154) implies

$$T_n(x_k,t) = a_k(t). \tag{155}$$

Hence the approximate solution (152) interpolates the solution of the PDE at the collocation points x_k. This justifies the name Chebyshev interpolation.

As earlier, we use the scaled problem (38) - (91) instead of the original problem and we force that the approximate solution (152) to satisfy the PDE at the collocation points x_j, which gives

$$\begin{aligned}
\frac{\partial T_n(x,t)}{\partial t}\,|_{x=x_j} &= \frac{4\alpha}{d^2}\sum_{k=0}^{n+1} T_n(x_k,t)g_k^{''}(x)_{|x=x_j} \quad j=1,\dots,n && (156)\\
q_i(t) &= -\frac{2\kappa}{d}\sum_{k=0}^{n+1} T_n(x_k,t)g_k^{'}(x_{n+1}) && (157)\\
T_e(t) &= T_n(x_0,t) && (158)
\end{aligned}$$

Equation (157) implies

$$T_n(x_{n+1},t) = \frac{-\frac{q_i d}{2\kappa} - \sum_{k=1}^{n} T_n(x_k,t)g_k^{'}(x_{n+1}) - T_e(t)g_0^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})} \tag{159}$$

Note that $T_i(t) = T_n(x_{n+1},t)$. Plugging (159) into (156) we have

$$\begin{aligned}
\dot{T}_n(x_j,t) = \frac{-2\alpha q_i(t)g_{n+1}^{''}(x_j)}{d\kappa g_{n+1}^{'}(x_{n+1})} + \frac{4\alpha}{d^2}\sum_{k=1}^{n} T_n(x_k,t)\left(g_k^{''}(x_j) - \frac{g_{n+1}^{''}(x_j)g_k^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})}\right)\\
+\, T_e(t)\frac{4\alpha}{d^2}\left(g_0^{''}(x_j) - \frac{g_{n+1}^{''}(x_j)g_0^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})}\right) && (160)
\end{aligned}$$

Setting $x = (T_n(x_1),\dots,T_n(x_{n-1}))$ we can write the system of equations (90), (36) and (38) in the standard state space form

$$\begin{aligned}
\dot{x} &= Ax + Bu\\
y &= Cx + Du
\end{aligned} \tag{161}$$

where

$$\begin{aligned}
a_{ij} &= \frac{4\alpha}{d^2}\left(g_j^{''}(x_i) - \frac{g_{n+1}^{''}(x_i)g_j^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})}\right) \quad i,j=1,\dots,n && (162)\\
b_{j1} &= \frac{4\alpha}{d^2}\left(g_0^{''}(x_j) - \frac{g_{n+1}^{''}(x_j)g_0^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})}\right) \quad j=1,\dots,n && (163)\\
b_{j2} &= \frac{-2\alpha}{d\kappa}\frac{g_{n+1}^{''}(x_j)}{g_{n+1}^{'}(x_{n+1})} \quad j=1,\dots,n && (164)\\
C &= \left(-\frac{g_1^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})},\dots,-\frac{g_n^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})}\right) && (165)\\
D &= \left(-\frac{g_0^{'}(x_{n+1})}{g_{n+1}^{'}(x_{n+1})},-\frac{d}{2\kappa g_{n+1}^{'}(x_{n+1})}\right) && (166)
\end{aligned}$$

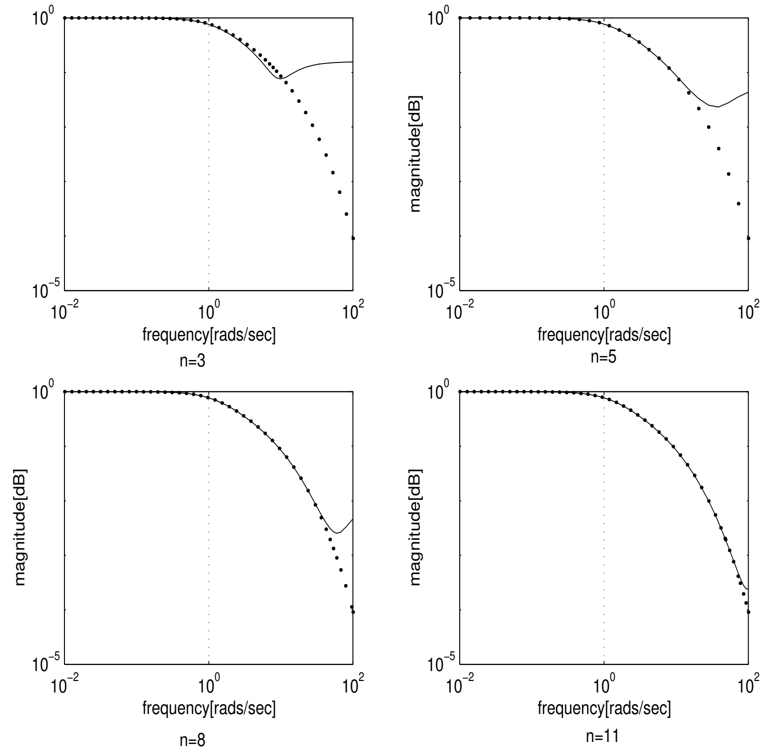

Figure 27: The Chebyshev interpolation approximation of $G_1(i\omega)$ for some different model orders($n = 3,\ 5,\ 8,\ 11$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK . The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

Rewriting (161) in the standard input output form and plotting the approximate transfer functions $G_1^n(s)$ and $G_2^n(s)$ along with their respective true transfer functions, Figure 27 and 28, we see that all the low frequency points are well approximated in both $G_1(s)$ and $G_2(s)$, but once again the high frequency components are not so well approximated in the case of $G_1(s)$, particularly for low model orders.

Figures 29, 30 and 31 suggest that conclusions similar to Galerkin and Tau holds for Chebyshev interpolation also. The rate of convergence for Chebyshev interpolation has been extensively dealt with in [5], where it has been derived that both the approximate transfer functions converge at a rate faster $O(\frac{1}{n^2})$.

5 Concluding discussion

We conclude this paper by comparing Galerkin, Tau, Chebyshev collocation and Chebyshev interpolation. All the above schemes seem to approximate the true transfer functions quite well at low frequencies. Even in the high frequency

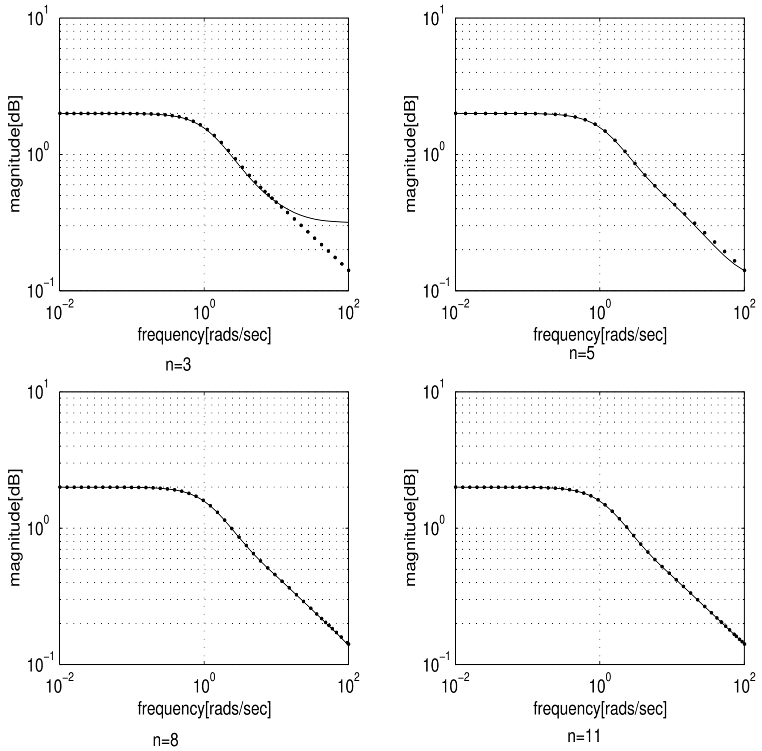

Figure 28: The Chebyshev interpolation approximation of $G_2(i\omega)$ for different-model orders ($n = 3,\ 5,\ 8,\ 11$) with $\alpha = 2$ m/s^2 and $\kappa = 1$ J/msK. The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line.

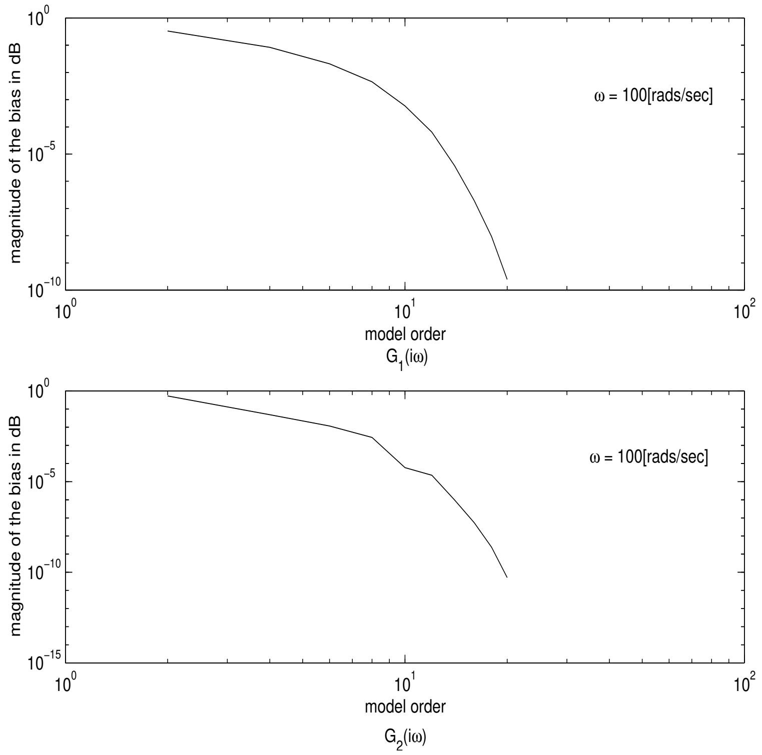

Figure 29: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \dots, 20$ at $\omega = 100$.

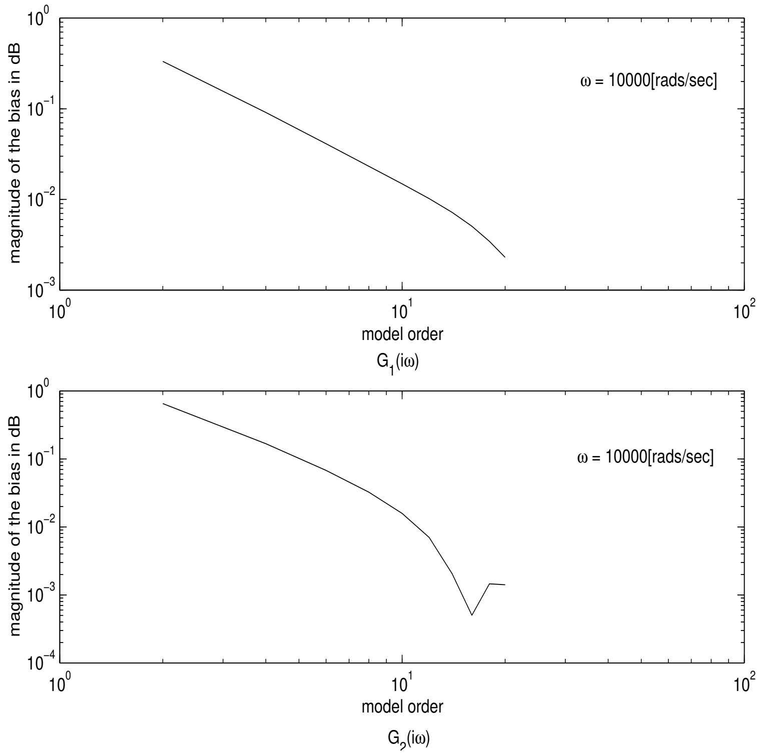

Figure 30: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \ldots, 20$ at $\omega = 10000$.

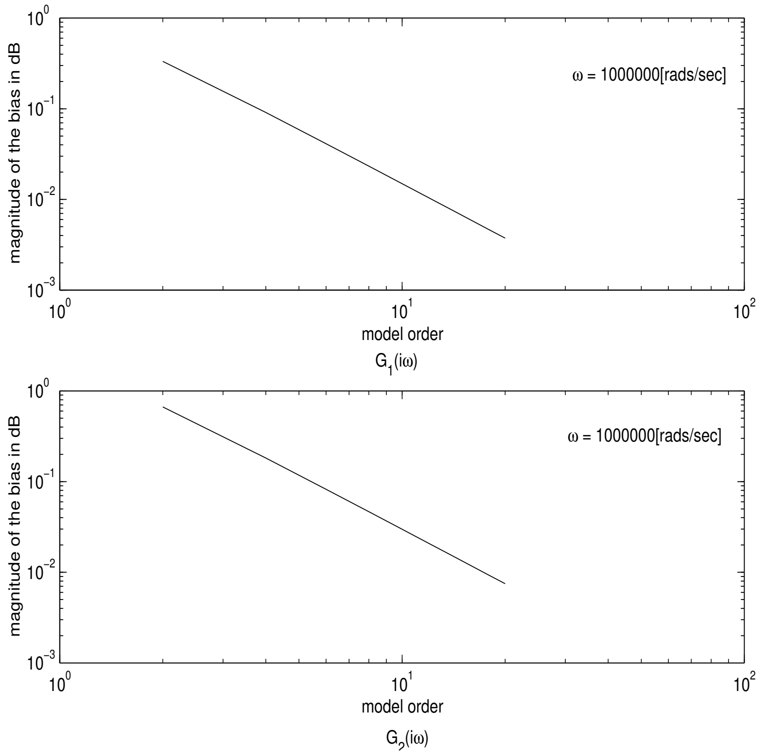

Figure 31: Deviations of the approximate transfer functions $G_i^n(i\omega)$ form t he true transfer functions $G_i(i\omega)$, $i = 1, 2$, for different models orders, $n = 2, 3, \dots, 20$ at $\omega = 1000000$.

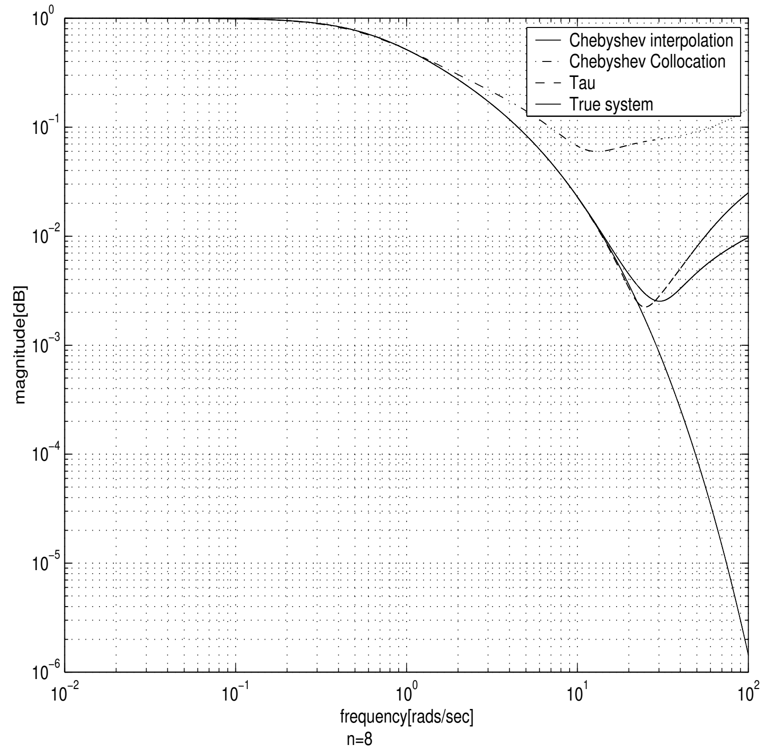

Figure 32: Tau, Chebyshev collocation and Chebyshev interpolation approximation for $G_1(s)$ for the model order $n = 8$.

region $G_2(s)$ is approximated quite well by all the schemes, but all the schemes differ significantly in approximating $G_1(s)$. Clearly in all the schemes, the error increases with increasing frequency and has an upper bound at $\omega = \infty$. If one looks at the worst case rate of convergence for all the schemes then, clearly Chebyshev interpolation is the best as it converges at a rate faster than $O(\frac{1}{n^2})$ for both the transfer functions as opposed to the other cases where $G_1^n(s)$ converges a rate faster then $O(\frac{1}{n})$. If one looks at the Figures 6, 17 and 22 more carefully, one can note that at certain frequency points there is a dip in the error, but such frequency points are very few to be worried about. Hence in general, it is better to use Chebyshev interpolation as it converges at a rate faster than $O(\frac{1}{n^2})$ for both the transfer functions at all the frequency points.

A Proof of Lemma 2

Recall the trigonometric integral

$$\int_0^\pi \cos n\theta \cos m\theta \, d\theta = d_n \frac{\pi}{2} \delta_{mn} \qquad (167)$$

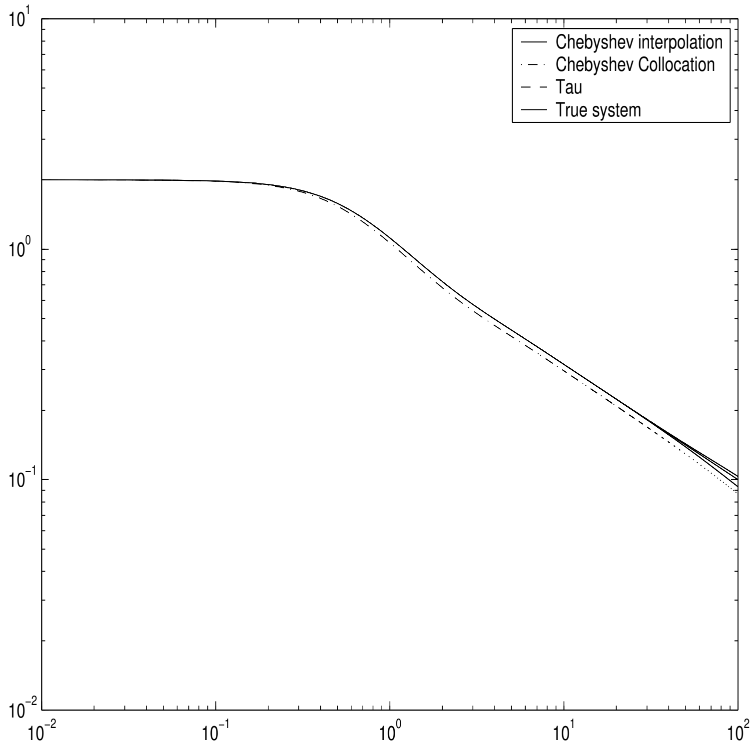

Figure 33: Tau, Chebyshev collocation and Chebyshev interpolation approximation for $G_2(s)$ for the model order $n = 8$.

where

$$d_0 = 2, \quad d_n = 2 \;\; (n \neq 0) \tag{168}$$

and δ_{mn} is the Kronecker delta.

With the usual change of variable, $x = \cos\theta$, and use of (23), (167) becomes

$$\int_{-1}^{1} \cos\left(n\cos^{-1}(x)\right)\cos\left(m\cos^{-1}(x)\right)\frac{dx}{\sqrt{1-x^2}} = d_n\delta_{mn}.$$

which can be rewritten using (23) and (24) as

$$\langle C_n(x), C_m(x)\rangle = \int_{-1}^{1}\frac{C_n(x)C_m(x)}{\sqrt{1-x^2}}dx = d_m\delta_{mn}. \tag{169}$$

Hence the set $\left\{\sqrt{\frac{2}{\pi d_i}}C_i(x)\right\}_{i=0}^{n+1}$, is a set of $n+1$ orthonormal polynomials, and therefore orthogonal basis for the space $\mathcal{P}_{n+1}$. ■

B Proof of the equation (145)

Let $S_n = \sum_{k=1}^{n}(-1)^{k-1}k^2$ and consider the following cases
Case 1: Let n be even, that is let $n = 2m$, where m is any positive integer. Then

$$\begin{aligned}
S_{2m} &= \sum_{k=1}^{2m}(-1)^{k-1}k^2 \\
&= \sum_{k=1}^{m}[(2k-1)^2 - (2k)^2] \\
&= \sum_{k=1}^{m}[4k^2 - 4k + 1 - 4k^2] \\
&= -4\frac{m(m+1)}{2} + m \\
&= -m(2m+1) \\
&= -\frac{n}{2}(n+1).
\end{aligned} \tag{170}$$

Hence for a large n we have

$$\mid S_n \mid \;\approx\; \frac{n^2}{2}. \tag{171}$$

Case 2: Let n be odd, that is let $n = 2m + 1$, where m is any positive integer, then

$$\begin{aligned} S_{2m+1} &= S_{2m} + (2m+1)^2 \\ &= -m(2m+1) + 4m^2 + 4m + 1 \\ &= 2m^2 + 3m + 1 \\ &= \frac{(n-1)^2}{2} + \frac{3(n-1)}{2} + 1 \\ &= \frac{n^2}{2} + \frac{n}{2}. \end{aligned} \tag{172}$$

Hence for a large n we have

$$\mid S_n \mid \approx \frac{n^2}{2}. \tag{173}$$

Hence (145) is proved. ∎

References

[1] H. S. Carlslaw and J. C. Jaeger. *Operational Methods In Applied Mathematics.* Oxford University Press, London, 1947.

[2] D. Gottlieb. *Lecture notes on Chebyshev collocation methods.* 1999. Uppsala University.

[3] D. Gottlieb and A.Orzag. *Numerical Analysis of Spectral Methods: Theory and Applications.* SIAM, Philadelphia, 1977.

[4] K. Kunisch. Numerical Methods for Parameter Estimation Problems. In *Inverse Problems in Diffusion Process.* SIAM, Philadelphia, 1995.

[5] T. Söderström and B. Bhikkaji. Reduced order models for diffusion systems via collocation methods. August 2000. Report 2000-018. Department of Information Technology, Uppsala University.

[6] T. Söderström and S. Remle. Parameter estimation for diffusion models. To appear in Proc of 12th IFAC Symposium on System Identification, Santa Barbara, CA, 21-23 June 2000.

[7] T. Söderström and P. Stoica. *System Identification.* Prentice Hall International, Hemel Hempstead, UK, 1989.

[8] J. Villadsen and M. L. Michelsen. *Solution of Differential Equation Models by Polynomial Approximation.* Prentice Hall, Englewood Cliffs, N. J., 1978.