Data-driven Parameters Tuning for Predictive Performance Improvement of Wire Bonder Multi-body Model

X. Cheng [1], A. Di Bucchianico [2], N. Javanmardi [3], M. de Jong [3], E. L. Diget [4] †, C. Please [5], D. Lahaye [6], Q. Peng [7], C. Reisch [8], and D. Sclosa [9],

[1] *Wageningen University and Research, The Netherlands*
[2] *Eindhoven University of Technology, The Netherlands*
[3] *Groningen University, The Netherlands*
[4] *University of Southern Denmark, Denmark*
[5] *Oxford University, United Kingdom*
[6] *Delft University of Technology, The Netherlands*
[7] *Leiden University, The Netherlands*
[8] *Technische Universität Braunschweig, Germany*
[9] *Vrije University Amsterdam, The Netherlands*

(*Communicated to* MIIR *on 1 May 2024*)

Study Group: ESGI170, Study Group Mathematics with Industry, University of Groningen, January 30 – February 03, 2023.

Communicated by: Julian Koellermeier, Pieter Tibboel, Stephan Trenn.

Industrial Partner: ASMPT, https://alsi.semi.asmpt.com/

Presenter: ASMPT

Team Members: X. Cheng, Wageningen University and Research, The Netherlands; A. Di Bucchianico, Eindhoven University of Technology, The Netherlands; N. Javanmardi, Groningen University, The Netherlands; M. de Jong, Groningen University, The Netherlands; E. L. Diget, University of Southern Denmark, Denmark; C. Please, Oxford University, United Kingdom; D. Lahaye, Delft University of Technology, The Netherlands; Q. Peng, Leiden University, The Netherlands; C. Reisch, Technische Universität Braunschweig, Germany; D. Sclosa, Vrije University Amsterdam, The Netherlands;

Industrial Sector: Electronics.

Key Words: frequency domain, optimisation, sensitivity analysis

† Corresponding Author: `eld@mmmi.sdu.dk`

Summary

This report describes work performed during SWI 2023 at the University of Groningen in relation with Problem 1 posed by the company ASMPT.

ASMPT makes a very large number of different machines for manufacturing of electronic devices. They have detailed simulation software of one of these machines and they compare the results of this with physical experimental results. There is a significant difference between the simulated and measured data, and it is the goal of this work to study how to estimate the parameters in the simulation model using the experimentally measured frequency response.

First, two toy models are studied to understand the challenges of parameter estimation in the frequency domain. Later, optimization methods are applied. Several different approaches of reducing the dimensionality of the parameter space are explored, including determining the parameter sensitivity. A suggestion for increasing the detail of the model, specifically related to the machine base, is also outlined.

In the summary, we supply a discussion of the key insights we gained during the week.

1 Introduction

With the increase in demand of electrical components the manufacturing has to be increasingly efficient and reliable. ASMPT is a world spanning company that produces as very large number of different machines for rapid manufacturing of electronic devices. Among these are wire bonding machines, shown in Figure 1, used in chip manufacturing. Wire bonding is the process of installing thin wires that connect internal chip areas to external breakout pads, see Figure 2.

Crucial to the efficient operation of these machines is that they can operate at high speed and hence they require control systems that will ensure vibrations generated in the system do not degrade the quality of the operations. To enable this measurements are taken where the machine is actively vibrated and its frequency response determined. In addition a detailed simulations of the vibrations are generated based on the CAD design of the machine. These simulations can then be used to enable high quality control of the vibrations to be made.

2 Problem Description

The problem of interest is how to make the computational simulations replicate, as best as possible, the measurements of the frequency response. To do this it is necessary to identify the values for the numerous parameters in the model used for the simulations. Some of these parameters can be measured by careful procedures (such as the mass of any particular part) but other parameters cannot be measured independently and must be identified by fitting the results of the model to the data. This report discusses procedures for performing such fitting.

Figure 1. An ASMPT wire bonder machine.

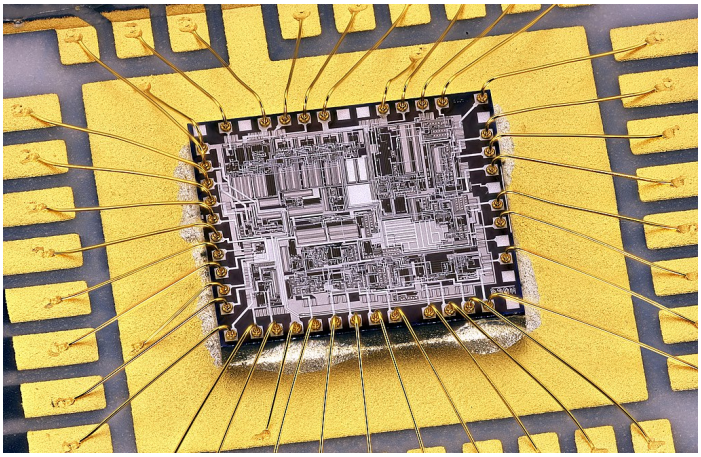

Figure 2. "Gold wire ball-bonded on a silicon die" by Mister rf, licensed under CC BY-SA 4.0.

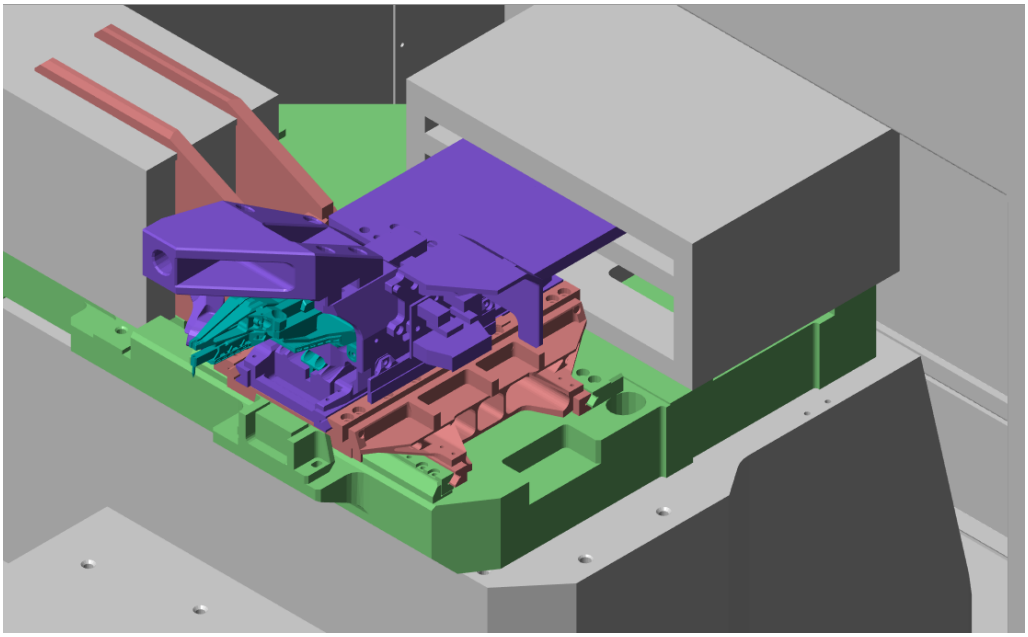

Figure 3. Details of the CAD design.

An image of the CAD model can be seen in Figure 3. Here the large grey lower region is the chassis, the green is the base, the red is the x-stage, which can move horizontally in the x-direction, the purple is the y-stage, which can move horizontally in the y-direction, and the blue is the z-stage (also called the θ-stage), which is a rotating arm that performs the wire bonding. In addition to these main features the system has some small flexibility

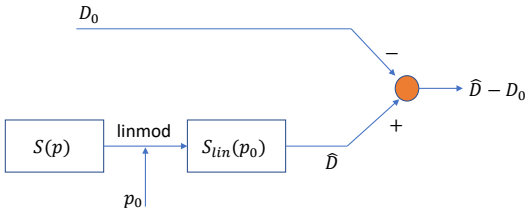

Figure 4. Outline of the data and simulations used in this report.

in the joints and guides that enable motion of each of the stages, out of their specific x-, y-, or z-plane. This flexibility is described by "parasitic elements" in the model. Some of these parasitic phenomena can be quantified by measurements, however there is great uncertainty in these measurements and, in addition some parasitic behaviour cannot be quantified independently in any easy manner.

The overall system that we are interested in can be summarised in Figure 4. Here the measured data from experiments is given by D_0, while $S(p)$ denotes the full nonlinear numerical model built in Simulink with the Simscape Multibody toolbox. The MATLAB program `linmod` is applied to the $S(p)$ to produce a linearised model $S_{\text{lin}}(p)$ which can then generate $\widehat{D}$, the computed frequency response of the linearised system. Finally some norm of the error between the simulated and measured data, $\widehat{D} - D_0$ will be exploited to determine how well a specific set of parameters p_0 fits the measurements so that optimal parameters values can be identified.

3 Approach

This report sets out the approaches that were considered for finding the optimal set of parameters to make the data from the simulations fit the measured data. The full model is very complicated and so we will start by considering two highly simplified models in order to gain some insight into the difficulties that might arise in fitting parameters to such models of oscillations. We will then look at possible cost functions that might be used to assess the quality of the fit between the simulated and measured data. Subsequently we focus our attention on how to seek optimal parameters conditions when the number of parameters is very large, as this is a serious barrier to using conventional methods of optimisation.

4 Modelling

In this section two toy models of the full dynamic model are investigated, and an alternative model for the machine base is considered.

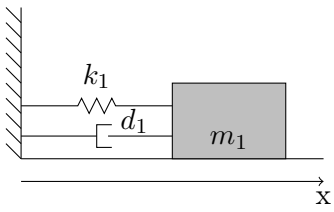

Figure 5. Diagram of simple one-mass oscillator.

4.1 Toy Models

We consider two very simple models of mass-spring-damper systems to see what type of behaviour occurs and how parameters might be fit. Additional details and code for these models can be found online[1].

4.1.1 *One-Mass Oscillator*

A single point mass attached to a stationary structure by a single spring and damper is studied. The mass can be subjected to some external force, and the position, $x(t)$, of the mass can be observed. This simple model is demonstrated in Figure 5.

The mass m_1 is assumed to be attached to a reference frame with a spring with spring constant k_1 and damper with damping constant d_1. The mass will thus oscillate horizontally around its equilibrium point. We assume that the mass can be determined a-priori using a scale. This allows us to reduce the design problem from three to two design variables and to consider the scaled stiffness constant k_1/m_1 and the scaled damping constant d_1/m_1.

The equation of motion is

$$m_1\ddot{x}(t) + d_1\dot{x} + k_1x(t) = F(t), \tag{4.1}$$

where $F(t)$ is an external force. Initial conditions complete the problem.

In the following, we study the frequency response of the system. The response is the ratio of the position to the forcing, and depends on the parameter values m_1, k_1 and d_1.

The state-space representation of the mechanical system allows to determine the transfer function. We therefore consider this representation here. The state-space representation is written in terms of the state vector $X(t)$ defined as

$$X(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix}. \tag{4.2}$$

The output and control are written as $y(t)$ and $u(t)$, respectively, where $y(t) = x(t)$, $u(t) = F(t)$. Unlike in more realistic applications, the equation governing the motion of the single point-mass is a linear ordinary differential equation. This allows us to recover the classical results obtained using Laplace transforms. We furthermore consider to have a single control variable only. This control parameter u has no part in the output. The

[1] `https://github.com/ziolai/software/blob/master/swi-groningen2023.ipynb`

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613

state-space representation then reads

$$\dot{X} = AX + Bu, \tag{4.3}$$

$$y = CX + Du, \tag{4.4}$$

where

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 \\ -k_1/m_1 & -d_1/m_1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ C &= \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned} \tag{4.5}$$

where $u = F(t)$ is the external force exciting the system equally at all frequencies with an amplitude F_0.

The observability of the system depends on the rank of the matrix

$$\begin{bmatrix} C \\ CA \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1, \end{pmatrix} \tag{4.6}$$

which in this case is equal to 2, and therefore the system is fully observable and controllable.

The transfer function of the system is defined in the Laplace domain with the variable s, given by

$$H(s) = \frac{1}{m_1\, s^2 + d_1\, s + k_1}. \tag{4.7}$$

After setting $s = j\,\omega$, where ω is the frequency, we obtain the transfer function in the frequency domain

$$H(\omega) = \frac{1}{-m_1\, \omega^2 + j\, d_1\, \omega + k_1}\,. \tag{4.8}$$

A numerical example using the parameter values of the z-stage of the wire bonder machine highlights the insight we may gain from the transfer function. The mass of the z-stage was assumed as $m_1 = 0.1363$ kg, the spring constant and damping with respect to the connection to the y-stage was assumed as $k_1 = 1.4429$ Nm/rad and $d_1 = 0.0031$ Nms/rad. Regarding the z-stage as a single mass-spring-damper system, we get the frequency response displayed in Figure 6.

The graph in the frequency domain shows three characteristics. First, for small frequencies we observe the limit case for $\omega \to 0$. This limit allows due to $\log(|H(\omega)|) = \log(1/k_1)$ to determine the spring constant k_1.

Second, in the limit $\omega \to \infty$ we have $H(\omega) = -1/(m_1\omega^2)$. Thus in this limit $\log(|H(\omega)|) = \log(1/(m_1)) - 2\log(\omega)$. This limit allows to determine the value of m_1, and we note the slope is 40 dB per decade.

Finally, the value of the damping coefficient d_1 can be determined from the value of $H(\omega)$ at the resonant frequency $\omega_0 = \sqrt{\frac{k_1}{m_1}}$, where

$$H(\omega_0) = \frac{1}{jd_1\sqrt{k_1/m_1}}. \tag{4.9}$$

The Bode plot for the amplitude therefore allows to determine all three parameters of the single mass-oscillator.

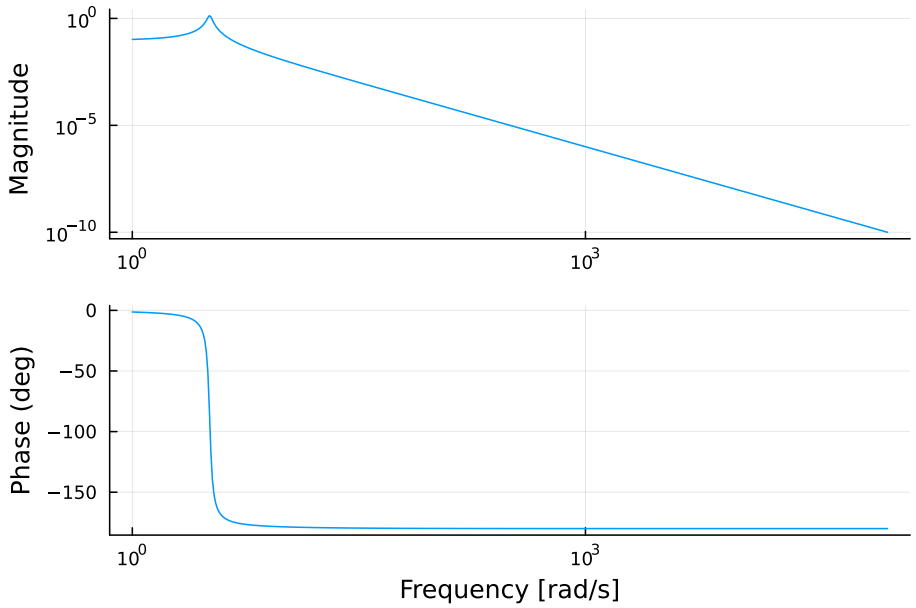

Figure 6. Bode diagram of simple one-mass oscillator using the parameter values of the z-stage.

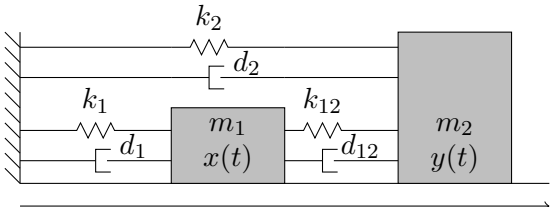

Figure 7. Diagram of simple two-mass oscillator.

Even though the transfer function is linear in forcing F_0, it is non-linear in the parameters k_1, d_1 and m_1. The optimization problem we wish to solve is thus non-linear in these parameters.

4.1.2 *Two-Mass Oscillator*

The second toy model has a single point mass attached to a stationary structure by a single spring and damper, which is subjected to some external force. However, we also consider that there is a second point mass which is connected by springs and dampers, both to the first mass and also to the stationary structure. The second mass and its connections are an attempt to model the parasitic behaviour that occurs in the real system. Hence our only observations are on the position, $x(t)$ of the first mass, and the response is the ratio of the motion of this first mass to its forcing.

We want to investigate the influence of adding more spring-dampers to a two masses system. Therefore, we compare the behavior of the system in Figure 7 with the sys-

tem behavior without the spring and damper connecting mass m_2 with the stationary structure.

The dynamical system reads

$$M \begin{pmatrix} \ddot{x}(t) \\ \ddot{y}(t) \end{pmatrix} + D \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} + K \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} F(t) \\ G(t) \end{pmatrix}, \tag{4.10}$$

with the diagonal mass matrix

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix},$$

the stiffness matrix

$$K = \begin{pmatrix} k_1 + k_{12} & -k_{12} \\ -k_{12} & k_2 + k_{12} \end{pmatrix},$$

and the damping matrix

$$D = \begin{pmatrix} d_1 + d_{12} & -d_{12} \\ -d_{12} & d_2 + d_{21} \end{pmatrix}.$$

We regard the spring k_2 and damper d_2 to be parasitic. In the following comparison we investigate the cases $k_2 = d_2 = 0$ and $k_2, d_2 > 0$.

Transferring the ordinary differential equation in the frequency domain gives the representation

$$F(\omega) = Z(\omega)\, X(\omega)\,, \tag{4.11}$$

where $Z(\omega)$ is the impedance matrix given by $Z(\omega) = -\omega^2 M - j\,\omega D + K$. Solving for the unknown displacement can thus be done as $X(\omega) = H(\omega)F(\omega)$, where

$$H(\omega) = Z^{-1}(\omega) \tag{4.12}$$

is the transfer function.

For our system, we can see that the influence of k_2 and d_2 affects the whole system behavior. The transfer function then reads

$$H(\omega) = \frac{1}{\det(Z(\omega))} \begin{pmatrix} H_{11} & -(j\omega d_{12} - k_{12}) \\ -(j\omega d_{12} - k_{12}) & H_{22} \end{pmatrix}, \tag{4.13}$$

where $\det(Z)$ depends on d_2 and k_2 in almost all coefficients of the monomials of ω and the diagonal elements are

$$H_{11} = -\omega^2 m_2 - j\omega(d_{12} + d_2) + k_{12} + k_2, \tag{4.14}$$

and

$$H_{22} = -\omega^2 m_1 - j\omega(d_1 + d_{12}) + k_1 + k_{12}. \tag{4.15}$$

The change in the system behavior is displayed in Figure 8, where the external force acts on mass m_1, and Figure 9, where the external force acts on mass m_2. The parameter values are chosen accordingly to the y-stage and the z-stage of the wire bonder machine.

While the change in the system response by a force on mass m_1 is negligible, the dynamics of mass m_1 on a force acting on mass m_2 are larger. Figure 9 shows, for the model with parasitic connections of mass m_2 to the stationary structure, an oscillating

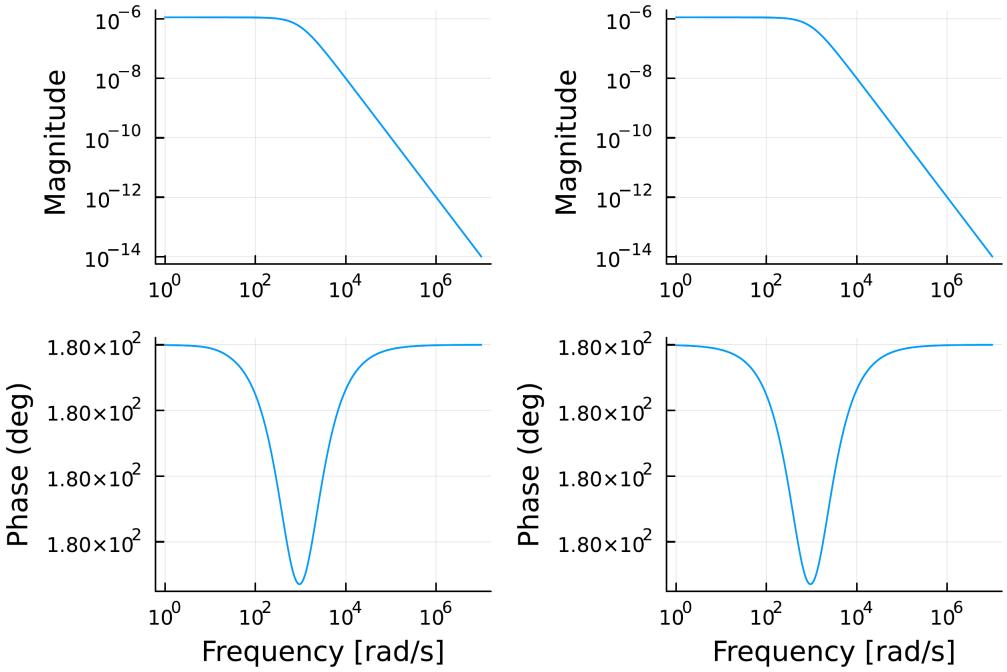

Figure 8. Bode plot of the two-mass oscillator. Systems response of mass m_1 on a force on mass m_1. Left: Additional connection with $k_2, d_2 \neq 0$ to the stationary structure. Right: No connection with $k_2, d_2 = 0$.

behavior for large frequencies. The decay of the magnitude is smaller for the model including the additional parasitic spring and damper.

The two changes in the slope of the magnitude in Figure 9 roughly correspond to the resonance frequencies of the single one-mass-system for the z-stage, see Figure 6, and the y-stage.

4.2 Alternative Model of the Chassis

The chassis of the machine is currently described in the model by a single rigid body at a fixed position with appropriate mass and inertial parameters. This is reasonable for any part of the machine where the speed of wave propagation in the part is sufficient that it moves quasi-statically. However, for the chassis, this might not be an adequate approximation since it has dimensions of around 1 m and shear waves travel at around 3 km/s, so that frequencies of over 1 kHz may create waves that are as short as the part. A more complex model of this could be created, using the finite element method (FEM) for example, however here we describe an alternative approach that retains the simplicity and number of parameters currently exploited in the rigid body model but accounts for wave propagation.

We consider an idealised base consisting of a uniform block of elastic material of height L. On top of the block the other parts of the machine are placed and these dictate a

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613

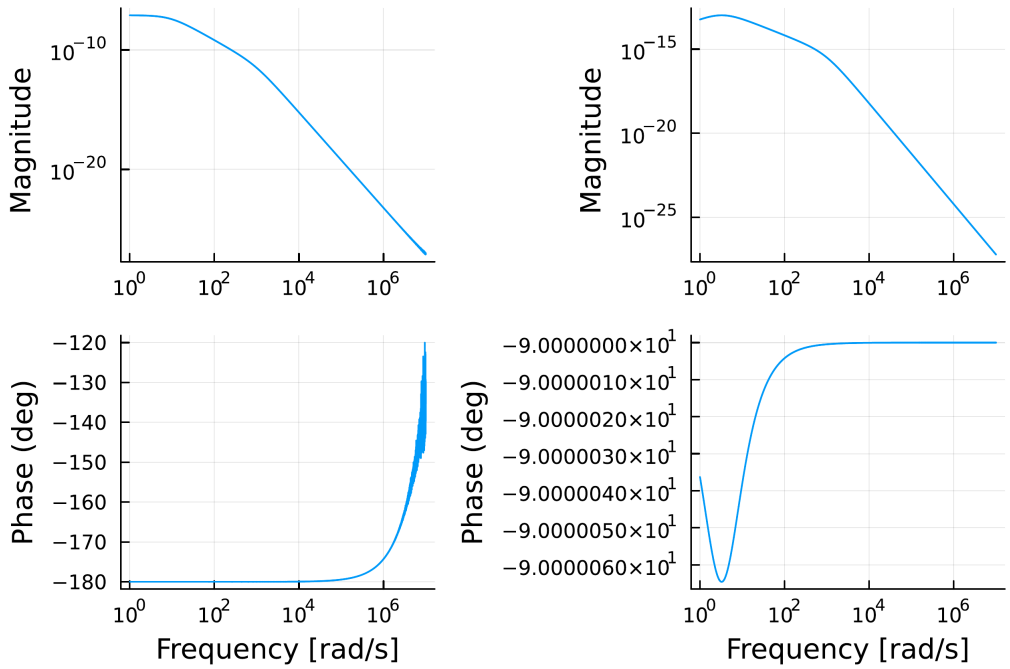

Figure 9. Bode plot of the two-mass oscillator. Systems response of mass m_1 on a force on mass m_2. Left: Additional connection with $k_2, d_2 \neq 0$ to the stationary structure. Right: No connection with $k_2, d_2 = 0$.

known displacement in the horizontal plane $x(t)$, $y(t)$, and vertical displacement $z(t)$. We are interested in determining the response by quantifying how the resulting horizontal and vertical forces acting on this surface depend on the displacements.

Our simplification will be that the motion of the block is uniform at each height z and that the horizontal motion (u, v) induces shear waves and the vertical motion, w, induces compressive waves. The bottom of the base is typically on legs and these are quite weak so for the purposes of modelling it is adequate to assume the bottom surface is stress-free. We now give details of the behaviour in the x direction but the other two directions are similar.

The governing equation is

$$\rho \frac{\partial^2 u}{\partial t^2} = \mu \left(\frac{\partial^2 u}{\partial z^2} - d_s \frac{\partial u}{\partial t} \right), \tag{4.16}$$

where ρ is the material density, μ its shear modulus and d_s the shear damping coefficient. Note the equation for v is identical, and for w is the same except μ is replaced by $\lambda + 2mu$, where λ is the first Lame coefficient, and d_s replaced by the compressive dissipation d_c. Boundary conditions for the problem are that

$$u(t, 0) = x(t) \qquad \text{and} \qquad \mu \frac{\partial u}{\partial z}(t, L) = 0. \tag{4.17}$$

If we move to frequency space, by either taking Fourier transforms or simply considering

$x(t) = x_\omega \exp(-j\omega)$ with $u(t) = u_\omega(z)\exp(-j\omega)$ we can find that the solution is given by

$$u_\omega(t) = x_\omega \frac{\exp\left((2L/c)\sqrt{\omega^2 - jd_s\omega}\right)}{1 + \exp\left((2L/c)\sqrt{\omega^2 - jd_s\omega}\right)}, \tag{4.18}$$

where $c = \sqrt{\mu/\rho}$ is the shear wave speed.

Finally we are interested in the relation between the displacement and the stress on the top surface, so that we can incorporate it into the more general modelling framework. Here we find

$$\begin{aligned} \text{stress} &= \mu \frac{\partial u}{\partial z} \\ &= x_\omega \left((2L/c)\sqrt{\omega^2 - jd_s\omega}\right) \\ &\quad \cdot \frac{1 - \exp\left((2L/c)\sqrt{\omega^2 - jd_s\omega}\right)}{1 + \exp\left((2L/c)\sqrt{\omega^2 - jd_s\omega}\right)}. \end{aligned} \tag{4.19}$$

We can therefore interpret this expression as giving a transfer function in the form

$$-\sqrt{\frac{k}{m(\omega^2 - jd_s\omega)}} \coth\left(\sqrt{mk}\sqrt{\omega^2 - jd_s\omega}\right), \tag{4.20}$$

where we have introduced a mass, m, and a stiffness, k. This puts the transfer function in the same form as that for a single mass connected with stiffness and damping where we would expect

$$\frac{1}{-m\omega^2 + jd\omega + k}, \tag{4.21}$$

and observe that the two agree, with the frequency behaviour proportion to $1/(m\omega^2)$, when the frequency is low in the shear wave and high in the single mass.

Implementation of this model of the chassis may not be straightforward in SimScape. An alternative approach would be to consider the chassis as a small number of individual masses linked by springs and dampers that represent the elastic behaviour of the region. Such a simplified model would go some way to describing the wave-like behaviour in the region.

5 Measured Data

The system has been measured using three sources of excitation and three senors of the motion with one of each on the x-stage, the y-stage and the z- (or θ-) stage. After analysis this measured data consists of nine complex numbers (the xx, yy, zz, xy, xz and yz representing the response, where mn corresponds to response in direction m due to excitation in n) at each frequency. We denote the measured data by $Y_{\text{mes}.xx}$ and the simulated data from models by $Y_{\text{sim}.xx}(\theta)$ where θ is the vector of model parameters. Experiment data was supplied by the project owner.

The data can be presented in the form of Bode plots where the magnitude (measured typically in dB) of each response and the phase (measured typically in degrees, and

restricted to the range $[-180; 180]$) of each response is plotted against the logarithm of the frequency.

6 Optimization Methods

We considered the full model of the wire bonder as developed by ASMPT and explored approaches to finding the parameter values that would fit the given data to the computations of the full model. A first issue to address was what cost function should be used to quantify the fit between the data and the computations. We discuss this is in Section 6.1. Next we explore approaches that will allow optimal parameter values to be estimated when the number of parameters is very large. Here we take two directions. The first is to look at the sensitivity of the cost function to the entire parameter set. Results for this are discussed in Section 6.2. The second approach was to do a limited optimisation where part of the parameter set was considered to be fixed at its nominal value while a smaller set was considered using an optimisation process. Three examples of this are presented. Each example optimises different parts of the parameter space. The first example studies the inertial parameters of the system, the second example we vary the parasitic parameters for the complete system, and in the third example we study the parasitic parameters in relation with the z-stage. These three examples are presented in Section 6.3, 6.4, and 6.5, respectively.

In our examples we have concentrated on local optimisation methods but this problem does require methods that will seek the global optimum amongst many local optima. A discussion and overview of current state of derivative free global optimisation methods can be found in [2] and [3].

6.1 Cost Function

The cost function is a crucial part of an optimization problem. Multiple error measures can be considered.

One strategy is to concentrate on just the magnitude of the complex function and only consider a small subset (e.g. just one) of the various responses. Hence one possible cost function is:

$$C_{\mathrm{mag}}(\theta) = \frac{1}{2} 20 \log_{10} \left(\frac{1}{N} \sum^{N} \left(|Y_{\mathrm{mes}}| - |Y_{\mathrm{sim}}(\theta)| \right)^2 \right) \text{ dB}, \tag{6.1}$$

where N is the number of frequency bins. This function has the property that it is dominated by those frequencies with large responses. For example the measured data shows approx 100 dB variation across the expected frequency range so any fitting will ignore large parts of the curve.

An alternative, that addresses the large variations, is to make the cost function directly applicable to the typical response curve where we seek to minimize the distance between the measured and simulated curves. This corresponds to taking

$$C_{\mathrm{mag}}(\theta) = \left(\frac{1}{N} \sum^{N} (20 \log_{10} |Y_{\mathrm{mes}}| - 20 \log_{10} |Y_{\mathrm{sim}}(\theta)|)^2 \right)^{1/2} \text{ dB}. \tag{6.2}$$

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613

A similar account of large variations can be made by normalising each simulated response by the magnitude of the corresponding measured response. It is then possible to simply consider the magnitude of the simulated response

$$C_{\text{mag}}(\theta) = \frac{1}{N}\sum^{N}\left|\frac{Y_{\text{mes}}}{|Y_{\text{mes}}|} - \frac{Y_{\text{sim}}(\theta)}{|Y_{\text{mes}}|}\right|. \tag{6.3}$$

Note that this last cost function has the property that it does not significantly penalise simulations that greatly under-predict (in the sense of dB) the magnitude of response of the system but does penalise over-prediction of the magnitude.

The issue of the nine different responses can be addressed in a number of ways but the vast difference in the size of the various responses (e.g. $Y_{\text{mes},xx}$ is more than 50 dB smaller than $Y_{\text{mes},zz}$) makes it preferable to extend the ideas in (6.2) and (6.3) but to consider a sum across the nine different responses, perhaps with some weighting if certain directions are perceived as being of greater physical importance.

It is worth noting here that the norms used in the expressions for the cost functions have been assumed to be the L^2-norm. However, there are benefits to considering the L^1-norm.

A general optimisation problem minimizing a cost function with respect to the parameters subject to bounds and constraints looks like:

$$\begin{aligned} \min_{\theta} \quad & C_{\text{mag}}(\theta), \\ \text{s.t.} \quad & \underline{\theta} \leq \theta \leq \bar{\theta}, \\ & m_i > 0, \\ & I_i \succ 0, \end{aligned} \tag{6.4}$$

where $\underline{\theta}$ and $\bar{\theta}$ are the lower and upper parameter bound, respectively, m_i is the ith stage mass, and I_i its corresponding inertia matrix around the center of mass. The constraints on the inertial parameters of each rigid body in a robotics arm are studied in [8]; the so-called physical feasibility. We advise the company partner to incorporate this into a solution.

6.2 Sensitivity Analysis

To reduce the dimensionality of the problem, we study which parameters affect the cost function the most. Therefore, we conduct a sensitivity analysis on the parameters, which provide an insight on how the change of parameter values affect the value of the cost function. This is performed by changing one parameter at a time, comparing with the experiment data, and computing variation to cost function. Each parameter has been varied in the range of $50\% - 150\%$ around its nominal value. For this sensitivity analysis we have used the cost function given in (6.1), but these results could equally be generated for the other cost functions described earlier.

In summary the important parameters appear to be the inertial parameters, especially masses of x-stage and y-stage, while damper and spring parameters and masses of chassis do not influence the cost much. Note that besides physical constraints on mass and inertia,

the rigid body center of mass (CoM) could also be considered so that it remains in the convex hull of the rigid body to be realistic.

The specific sensitivity procedure is described in Algorithm 1.

Algorithm 1 Sensitivity Analysis

Require: nominal values of parameters $\theta_0 \in \mathbb{R}^{n_\theta}$;
range $[r_l, r_u]$, step size r_Δ
Ensure: cost function matrix $\mathcal{C}$
$\theta \leftarrow \theta_0$
for do $i = 1, 2, ..., n_\theta$ ▷ traverse all parameters
$j \leftarrow 1$, $r \leftarrow r_l + jr_\Delta$
while $r \neq 1$ and $r \leq r_u$ **do**
$\theta \leftarrow \theta_0$
$\theta(i) \leftarrow \theta(i)r$
Compute $C_{\text{mag}}(\theta)$ of the cost function (6.1).
$\mathcal{C}_{ij} \leftarrow C_{\text{mag}}(\theta)$
$j \leftarrow j + 1$, $r = r_l + jr_\Delta$
end while
end for

In this algorithm, we set $r_i = 0.5$, $r_u = 1.5$, and $r_\Delta = 0.1$. In each loop, only one parameter is changing, while the others are fixed at their initial values. The values of the cost function are computed with different parameter values, which is illustrated in Figure 10. We can also evaluate the cost function in each direction x, y, and z, when changing the parameter values. Then, Figure 11 shows the results for the cost function for each direction.

Combining the data obtained from Algorithm 1 and the Figures 10, 11, 12, we have the following observations.

- There are two significant parameters:
  - The mass of the x- and the y-stage.
- There are other parameters that affect the cost function:
  - The center of mass and inertia for the x-, the y-, and the z-stage, and the mass of the z-stage.
- The rest of parameters do not affect the cost function; the parasitic spring and dampers have little effect on the overall cost.

6.3 Experiment 1: Varying the Inertial Parameters

Sensitivity analysis suggests a prominent role of the masses in the model, which is reflected in a significant role in optimization. We will allow masses to deviate, perhaps unrealistically, up to ±20% from the initial parameter values. Quite surprisingly, we see that a better overall fit is obtained for masses close to the maximum deviation allowed.

The experiment details are as follows:

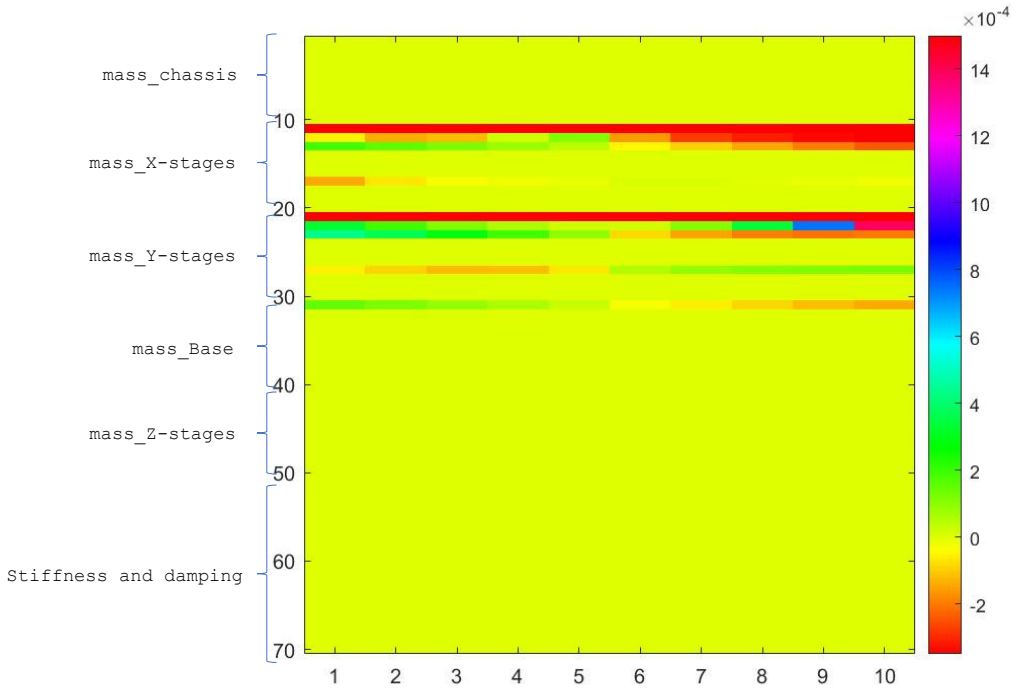

Figure 10. Value of the cost function (6.1) as the 70 parameters vary in the range 50% to 150%.

- Frequency-range: 200–2000 Hz
- Parameters varied (17 total):
  - Inertial parameters of the x-stage; mass, center of mass, inertia.
  - Parasitic torsional spring between the base and x-stage.
  - Kinematic position of the joint between the base and the x-stage.
- Maximum deviation allowed: $\pm 20\%$ for masses, $\pm 50\%$ for all other parameters
- Optimization method: `fmincon` with numerically computed derivatives
- We restrict to fitting the Y_{xx}-response of the x-stage

Overall, the optimization algorithm leads to a better fit. In particular, we obtain a better fit of the 0.245 dip, see Figure 13.

However, we notice that the one optimized parameter, namely the mass of the x-stage, coincides with the maximum value allowed. That is, a better fit has been obtained by increasing the x-stage mass by 20%.

This shows that allowing masses to change leads to a better fit, but probably to an unrealistic model. We suggest some possible causes: First, the "real" masses might not correspond to the global minimum of the cost function; second, large deviations might have brought us on a slope that minimizes the cost function but leads to a region of non-physical parameters (see Section 6.1 for a possible solution); third, the algorithm in use might not be suitable for the problem at hand.

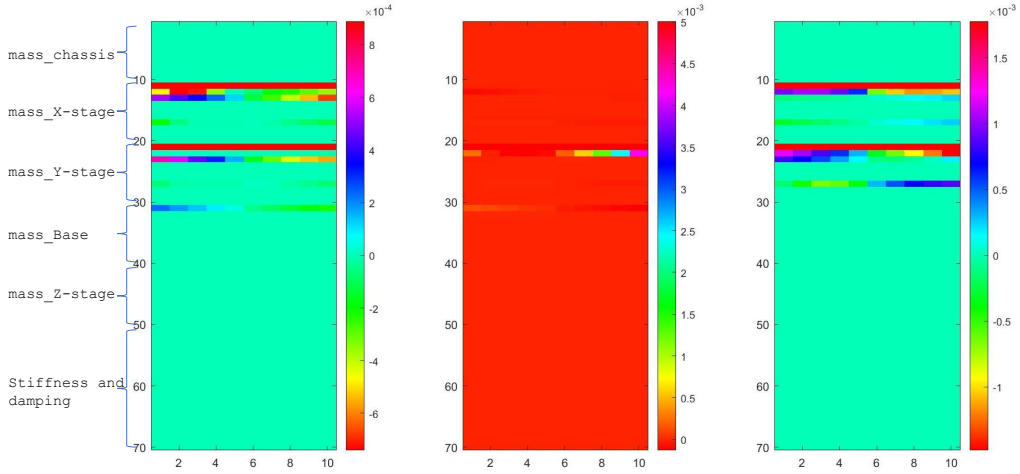

Figure 11. Values of the cost functions computed for x-, y-, and z-directions as the 70 parameters vary in the range 50% to 150%.

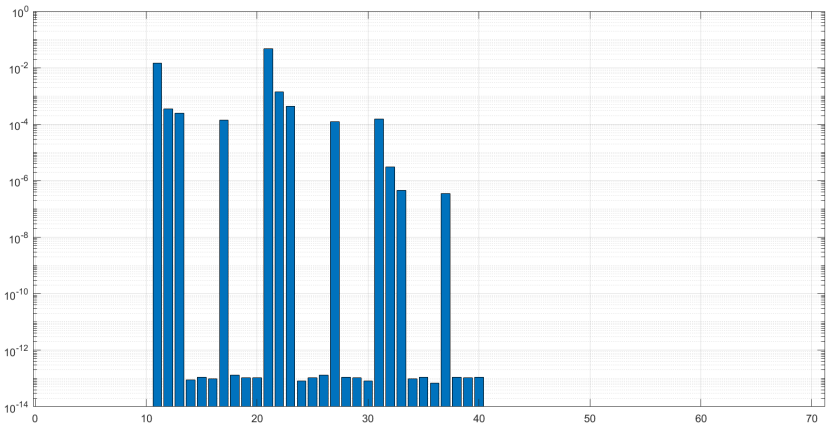

Figure 12. The maximum change in the magnitude of the cost function vs. parameter when the parameters are changed individually.

6.4 Experiment 2: Varying the Parasitic parameters

As the masses are quite accurately measured and the inertial parameters can be well estimated from CAD drawings the full model was explored with these parameters set at their nominal values. The parasitic behaviour of the system is much less well understood and hence in this section we explore how the simulations fit to the data when only parasitic effects are altered. At the same frequency range we vary the parasitic parameters. These are the details:

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613

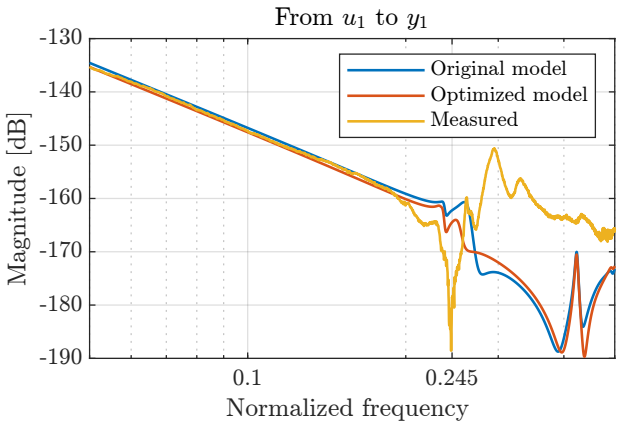

Figure 13. Optimizing the mass gives a better fit of the 0.245 dip. The frequency range was normalized by dividing with the Nyquist frequency.

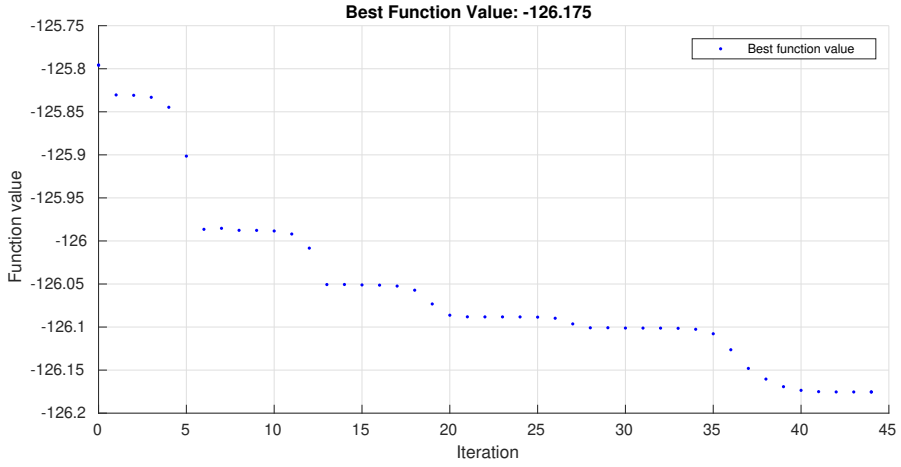

Figure 14. Decrease of the cost-functional during the optimization process. The convergence process exhibits various plateaus.

- Frequency-range: 200–2000 Hz
- Parameters varied: Parasitic parameters
- Parameter bounds: $\theta > 0$
- Optimization method: `fmincon` with numerically computed derivatives
- Objective: Minimize diagonal costs (related to Y_{xx}, Y_{yy}, Y_{zz})

In contrast to the case of masses, the optimization process here goes through many plateaus, see Figure 14, taking a lot of time. The result is an overall better fit in the relevant range 200–2000 Hz, see Figure 15.

Notice that one parameter, the parasitic torsional damping between the base and the x-stage, is 40 times its original value, see Figure 16. This does not come as a huge

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613

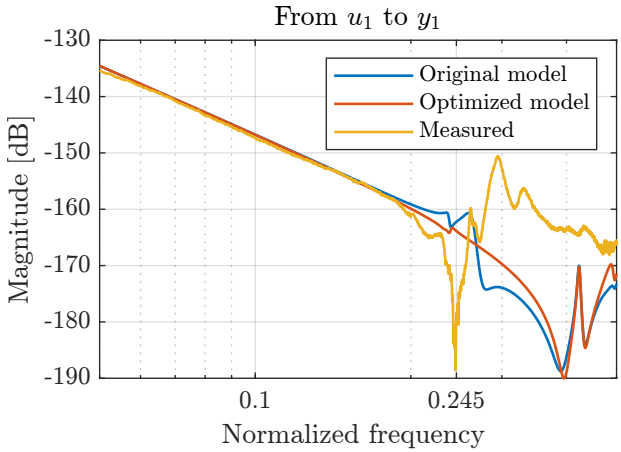

Figure 15. F_x-response of x-stage with optimized parasitic parameters and original parameters. The frequency range was normalized by dividing with the Nyquist frequency.

surprise. Indeed, sensitivity analysis already shows that modifying these parameters in the 50%–150% range has little to no effect on the system.

6.5 Experiment 3: Exploring the z-Stage

Instead of optimizing the complete cost of all three axes, in this simulation we only optimize the behaviour in the z-axis. According to the project owner, the measurement of the z-axis is more precise in the whole range from 1–4000 Hz than the other two axes. We optimize 12 parameters characterising the parasitic joints between the three stages, see Figure 17. The details are as follows:

- Frequency-range: 1–1000 Hz
- Parameters varied: Parasitic parameters
- Parameter bounds: $\theta > 0$
- Optimization method: `fmincon` with numerically computed derivatives
- Objective: Minimize z-stage cost (related to Y_{zz})

Unsurprisingly, it can be seen that the two last parameters, the spring and damper between the y- and the z-stage, change the most as a result of the optimization; around 30%, see Figure 17. The fit itself can be seen on Figure 18, where the fit of the bump at ~ 12 Hz is better with the new parameters compared with the nominal model.

7 Some Ideas on Data Analysis

In this section, we discuss some issues related to inaccurate data, which can generate incorrect models. Considering these factors may help to produce rich datasets, especially in low frequencies. Noticeably, discovering good datasets is important in system identification.

- The effect of the noise on the training of models should be considered. In studies like

https://doi.org/10.33774/miir-2024-f3zf3 **ORCID:** https://orcid.org/0000-0001-6743-5613 Content not peer-reviewed by MIIR.

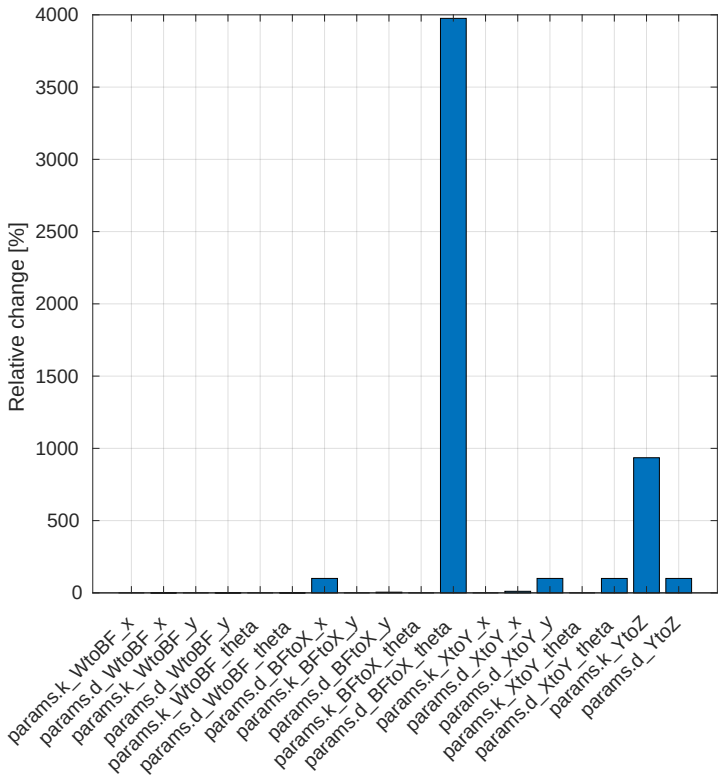

Figure 16. Absolute relative difference in percentage for the parasitic spring and damper coefficients before and after optimisation.

[9], [4], the effect of noisy data on the performance of the models trained on them is shown. However, real-world datasets often contain a significant fraction of noisy labels and uncertainties. In [4], the authors suggest robust training methods to learn from noisy-labeled data.

- To identify a model for a system it is important to check ‘sufficient excitation’ conditions. Persistent Excitation (PE) typically results from sufficiently rich reference inputs, and parameter convergence is achieved only in the presence of PE. In [1], [6] and [7], more information about PE conditions can be found.
- As studied in [5], Principal component analysis (PCA) is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss.

8 Discussion and Recommendations

The key insights in to the behaviour of the simulation model and how predictions might be fitted to experimental measurements are as follows. The sensitivity of each parameter depends on the frequency range being considered, i.e. dominant parameters change with respect to different frequency ranges. In particular we observe that the low frequencies are dominated by stiffness parameters, at near resonance frequencies the damping parameters

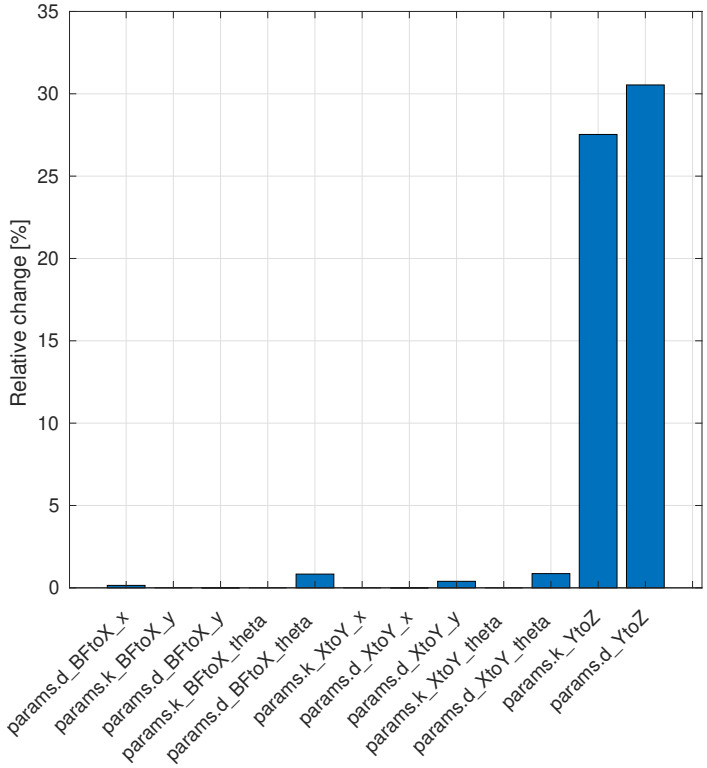

Figure 17. Absolute relative change in percentage for the parasitic spring and damper coefficients after optimization.

dominate, and at the high frequencies the inertial parameters dominate. Exploring a very limited range of frequencies may allow improved estimation of parameters for starting values for subsequent fitting to the large range of relevant frequencies.

The choice of the cost function has a great impact on the performance of the optimisation. Finding a global minimum is plagued by local minima, even in (very simple) toy models. The scale of the discrepancy (measured in the L^2 norm) in z-direction is e^{50} orders of magnitude higher than those of the $x-$ and y-directions and needs to be accounted for carefully.

From our preliminary explorations we make the following recommendations. It may be beneficial to extend the model to more carefully account for shear waves in the chassis. This is because in the relevant frequency range the chassis may be larger than the shear wave length and hence not well described by a single point mass. As a preliminary scoping of parameters values it may be beneficial to focus on the local optimization of parameters rather than global. It may also help to divide the frequency domain into a number of smaller ranges and consider initially optimising in each range. Some parameters can be kept constant, for example some damping parameters at low frequencies. The results of

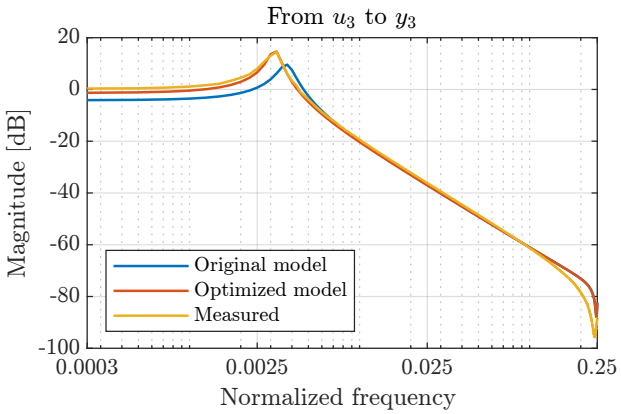

Figure 18. Frequency response of the z-stage. The frequency range was normalized by dividing with the Nyquist frequency.

the sensitivity test results might guide which parameters to consider and which to hold constant in the different frequencies ranges.

The sensitivity analysis in this report shows that the inertial parameters of the three stages dominate the cost function. The problem with this is that these are typically values are quite accurately known from independent measurements; these essentially create the nominal model. Instead, the optimisation could initially only be performed on the parasitic parts of the model. More work is needed to look at the sensitivities that occur for different cost functions. It is necessary to adopt strategies that can overcome the curse of dimensionality.

We suggest the following might be appropriate methods:

- Perform a sensitivity analysis for different frequency ranges separately,
- Optimize considering *only* the parasitic parameters as variables,
- Apply physical constraints to the inertial parameters of the rigid bodies including considering them as constant,
- Apply model order reduction, methodology.

Because of the vast range of responses in the x, y, and θ directions these should either be normalised or their logarithm taken when considering what cost function to use. Because of the need to find global optimal, state-of-the-art global optimization software, e.g. `Py-BOBYQA`, should be explored.

References

[1] Sergio Bittanti, Marco C Campi, and Lei Guo. Persistence of excitation properties for the identification of time-varying systems. In *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No. 00CH37187)*, volume 1, pages 680–684. IEEE, 2000.

[2] Coralia Cartis, Jan Fiala, Benjamin Marteau, and Lindon Roberts. Improving the flexibility and robustness of model-based derivative-free optimization solvers. *ACM Transactions on Mathematical Software (TOMS)*, 45(3):1–41, 2019.

[3] Coralia Cartis, Lindon Roberts, and Oliver Sheridan-Methven. Escaping local minima with local derivative-free methods: a numerical investigation. *Optimization*, 71(8):2343–2373, 2022.

[4] Oussama Dhifallah and Yue Lu. On the inherent regularization effects of noise injection during training. In *International Conference on Machine Learning*, pages 2665–2675. PMLR, 2021.

[5] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[6] Ti-Chung Lee, Ying Tan, and Dragan Nešić. Stability and persistent excitation in signal sets. *IEEE Transactions on Automatic Control*, 60(5):1188–1203, 2014.

[7] Yongduan Song, Kai Zhao, and Miroslav Krstic. Adaptive control with exponential regulation in the absence of persistent excitation. *IEEE Transactions on Automatic Control*, 62(5):2589–2596, 2016.

[8] Cristóvão D. Sousa and Rui Cortesão. Physical feasibility of robot base inertial parameter identification: A linear matrix inequality approach. *The International Journal of Robotics Research*, 33:931–944, 5 2014.

[9] Yihao Xue, Kyle Whitecross, and Baharan Mirzasoleiman. The final ascent: When bigger models generalize worse on noisy-labeled data. *arXiv preprint arXiv:2208.08003*, 2022.