

Proximal Gauss–Newton Method for Box-constrained Parameter Identification of a Nonlinear Railway Suspension System

Enis Chenchene¹, Josef Fuchs², Bernd Luber³, and Kristian Bredies⁴

^{1,4} *Department of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, 8010 Graz, Austria*
enis.chenchene@uni-graz.at
kristian.bredies@uni-graz.at

^{2,3} *Virtual Vehicle Research GmbH, Inffeldgasse 21a, 8010 Graz, Austria*
josef.fuchs@v2c2.at
bernd.luber@v2c2.at

ABSTRACT

The identification of railway vehicle components' characteristics from measured data is a challenging task with compelling applications in health monitoring, fault detection, and system prognosis. Usually, though, such systems are highly nonlinear, and naive identification techniques may lead to unstable methods and inaccurate results. In this paper, we show that these issues can be easily tackled with the recently introduced proximal Gauss–Newton method, which we employ to identify the parameters of a railway nonlinear suspension system. In the proposed model, the parameters are subject to safety bounds in form of box constraints, which allows preventing nonphysical solutions. The suspension system we consider is highly nonlinear due to the presence of an airspring in the secondary suspension, which we introduce in a simplified Berg model. Numerical examples, featuring data corrupted by various noise levels, demonstrate the accuracy and efficiency of our proposed method. Comparisons with state-of-the-art approaches are also provided.

1. INTRODUCTION

Currently, the maintenance process of railway vehicles and their components is mainly based on fixed inspection intervals, which periodically freezes operations and leads to corresponding high costs. For this reason, the development of reliable approaches to identify and isolate component faults in real-time is gaining increasing interest in the railway industry. One of the most popular approaches for health monitoring and system prognosis consists in describing the system via its mathematical model and deduc-

ing the model parameters from the system response, e.g., measured accelerations. Large deviations from nominal components' characteristics would indeed suggest a possible component fault. In the literature, this is often referred to as the *parameter identification method* in *model-based* approaches for Fault Detection and Isolation (FDI) (Strano & Terzo, 2019), which only represents a small portion of the techniques currently under research, see, e.g., (Bruni, Goodall, Mei, & Tsunashima, 2007) for an extensive review.

The parameter identification problem for linear dynamical systems is often solved with recursive least-square methods and Kalman filter (KF) approaches (e.g., Extended and Unscented KF, Cubature KF), which can be employed for identification problems considering the model characteristics as additional system states. For linear suspension models, KF-based identification techniques are predominant and have been proven to be effective, see for instance (Zoljic-Beglerovic, Stettinger, Luber, & Horn, 2018; Zoljic-Beglerovic, Luber, Stettinger, Müller, & Horn, 2020) and the references therein. However, modern suspension systems, which include, e.g., airsprings, are better described via highly nonlinear models, with possibly complex architectures (Berg, 1999; Mazzola & Berg, 2014). The additional source of nonlinearity can cause KF approaches to suffer from severe issues, such as unstable and quickly divergent behaviors, poor linearization and/or erratic behaviors (Xin-Chun & Cheng-Jun, 2013). Contrarily to KF-based approaches, the Gauss–Newton method, a popular iterative method for multibody dynamics identification (Eich-Soellner & Führer, 1998), can be easily generalized to include constraints on the parameters. This is indeed a crucial feature that can be leveraged to design stable methods with robust convergence guarantees, as we will detail in Section 3.2. The main computational drawback of considering complex (i.e., nonlinear) constraints is that at each

Enis Chenchene et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://doi.org/10.36001/IJPHM.2024.v15i2.3919>

iteration one needs to solve a quadratic problem, which is usually tackled via sophisticated active-set strategies. In this paper, we offer a simple and robust alternative based on the so-called *proximal Gauss–Newton* method (pGN), originally introduced by Salzo and Villa in (Salzo & Villa, 2012).

1.1. Contributions and highlights

In the following, we outline our contributions for convenience to the reader. We propose pGN for parameter identification for railway vehicle suspension systems as a simple yet robust alternative to standard approaches in the field. In this context, we demonstrate its efficiency by addressing a compelling parameter identification problem for a quarter railway vehicle featuring a highly nonlinear Berg’s airsprung model in the secondary suspension (Berg, 1999; Mazzola & Berg, 2014), and a linear spring/damper system in the primary suspension. In particular, we give an answer to the following research question: Can we identify the parameters of the given model, relying solely on the knowledge of vertical track irregularities as the system excitation and the corresponding simulated vertical accelerations of both the bogie and the carbody? The answer is positive in a sense that is clarified in Section 5.

To underscore the strengths of the proposed approach, let us highlight some of its features. First, pGN stands out for its simplicity and robustness. It allows us to tackle box constraints in parameter identification for dynamical systems by composing the standard Gauss–Newton step with a possibly inexact projection onto the feasible set, cf. (Gonçalves & Menezes, 2020, Theorem 3), which we compute with a low-cost first-order scheme, as in Algorithm 1 in Section 3.

It is also worth emphasizing that we tackle the parameter identification problem for a quarter railway vehicle equipped with an airsprung in the secondary suspension, a modern suspension technology gaining increasing popularity in the railway industry, yet remaining largely unexplored in parameter identification. In this specific use case, despite the presence of this highly nonlinear suspension system, the proposed methodology is capable of swiftly and accurately identifying primary suspension characteristics, even in scenarios with elevated levels of noise.

Eventually, when compared against conventional active-set approaches equipped with state-of-the-art quadratic solvers, the proposed method achieves comparable overall performance, while not relying on any black-box solver.

1.2. Related works

In the remainder of the introduction, we provide an overview of main related works devoted to parameter identification for model-based approaches, highlighting the

similarities and the differences between the state-of-the-art and the proposed approach.

Many classical monographs, such as (Bard, 1974; Eich-Soellner & Führer, 1998), include chapters on parameter identification for constrained, even nonlinear, vehicle suspension models. Standard approaches encompass Gauss–Newton and projected-gradient-type methods. For instance, (Eich-Soellner & Führer, 1998, Chapter 7) is devoted to a parameter identification problem for a suspension system with nonlinear damping for a truck model. The identification model is tackled with a Gauss–Newton method via Karush–Kuhn–Tucker (KKT) conditions. Here, inequality constraints are avoided, as they would lead to much more delicate active-set strategies. As we mentioned, inequality constraints can be tackled either via projection or via penalty methods, which are described in detail in (Bard, 1974). Such penalty methods, also known as barrier-like approaches, are implemented, e.g., in (Grupp & Kortüm, 1993), for the identification of parameters for nonlinear suspension system of a truck model with equality and inequality constraints.

In (Ding, Pan, & Chen, 2012), the authors propose a Levenberg–Marquardt trust region method to estimate the unknown parameters, in which the second-order sensitivity analysis is performed using the adjoint method. However, the methodology is not applied to railway vehicle suspension systems. Meanwhile, (Callejo & de Jalón, 2015) proposes the use of automatic differentiation to compute the gradients of the objective function. The proposed model in (Callejo & de Jalón, 2015) is nonlinear in particular due to the piece-wise linearity of the damper. In Section 4.2, the authors also include an interesting discussion about parameter constraints, and their relevance in the vehicle parameter identification. From an optimization viewpoint, the main difference with the present paper is that in (Callejo & de Jalón, 2015) the authors tackle the constrained identification model via a projected-gradient descent method, which, while enjoying global convergence guarantees, is known to be significantly slower than Gauss–Newton-type methods.

In (Serban & Freeman, 2001), the authors employ a Gauss–Newton method for the optimization of a spring-damper system of a 14-body model of the U.S. Army’s High Mobility Multipurpose Wheeled Vehicle. However, no constraints are considered in the identification example. In (Vyasarayani, Uchida, Carvalho, & McPhee, 2012), the authors apply the homotopy technique to the problem of parameter identification to mitigate the fact that Gauss–Newton does sometimes converge to local minimizers instead of global minimizers. Applications to a vehicle suspension system with nonlinearities are shown as well. In (Puel, Bourgeteau, & Aubry, 2013), a nonlinear time-

dependent model describing a rubber bushing's mechanical behavior is presented and the parameters associated with this model are then identified from experimental tests using an adjoint state formulation of the identification problem. We eventually refer to (Kraft, Puel, Aubry, & Funschilling, 2016) for an application of the adjoint state approach to the nonlinear vehicle-track system, where a cubic nonlinearity is introduced in the primary suspension.

Considering the recent developments, the parameter identification literature in multibody systems has yet to recognize the underexplored potential of the proximal Gauss-Newton algorithmic framework. Our contribution seeks to fill this gap by presenting a compelling application to the identification of parameters for a highly nonlinear quarter vehicle model, featuring an airspring model in the secondary suspension.

2. DYNAMICAL SYSTEM

In our application, we consider railway vehicles equipped with airsprings in the secondary suspension. To begin, we provide an overview of airsprings and the various modeling approaches.

2.1. Airsprings

Airsprings are attracting increasing interest in the railway industry. They first appeared in the 1960s and were considered somewhat of a novelty at the time but, nowadays, air suspension is a standard fitting for passenger vehicles. The mathematical modeling of airsprings is an ongoing challenge that found one of its first accurate descriptions with the Berg model introduced in the late 1990s in (Berg, 1999). According to Berg (see Figure 1), the vertical force generated by the airspring is given by the following law:

$$\begin{cases} F = k_e z + k_v(z - w) + \frac{z}{z_2 + z} F_{f,max}, \\ M w'' = k_v(z - w) - C |w'|^\beta \text{sign}(w'), \quad \beta = 1.8, \end{cases} \quad (1)$$

where z is the displacement, k_e , k_v are elastic and viscous parameters respectively, and M is the air mass that due to high air acceleration cannot be neglected. The third term on the right hand-side of the first identity is a friction contribution, which is defined up to two input parameters $F_{f,max}$ and z_2 , i.e., the maximum friction force and the friction displacement respectively, see, e.g., (Sayyaadi & Shokouhi, 2009; Berg, 1997) for more details.

A seminal airspring model appeared even earlier in the works of Nishimura (Matsumiya, Nishioka, Nishimura, & Suzuki, 1969; Oda & Nishimura, 1969), and an interesting nonlinear variant has been analyzed more recently in (Mazzola & Berg, 2014). The latter can be described with the following equations of motion, which shall be compared

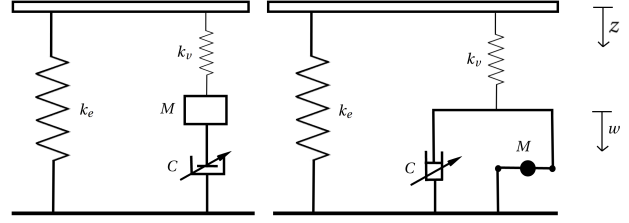


Figure 1. Left: the airspring model presented in (Moheyeldin et al., 2018). Right: the airspring model presented in (Berg, 1999) without the friction component.

with Figure 1:

$$\begin{cases} F = k_e z + k_v(z - w), \\ M w'' = k_v(z - w) - C (w')^2 \text{sign}(w'). \end{cases} \quad (2)$$

All the parameters in Eq. (2) have analogous meanings to those in Eq. (1) and are specified for instance in (Moheyeldin, Abd-El-Tawwab, El-gwwad, & Salem, 2018). Note that Eq. (2) can also be understood as a simplified setting of the Berg model in Eq. (1) with no friction term and $\beta = 2$. For a more detailed discussion on different airspring models we refer the interested reader to (Mazzola & Berg, 2014) and the references therein.

In this work, we consider a model in-between Eq. (2) and Eq. (1), that reads as

$$\begin{cases} F = k_e z + k_v(z - w), \\ M w'' = k_v(z - w) - C |w'|^\beta \text{sign}(w'). \end{cases} \quad (3)$$

Note that we neglect the friction term, which, however, is usually most pronounced for horizontal motion (Berg, 1999; Sayyaadi & Shokouhi, 2009), which we do not consider.

2.2. Quarter vehicle model

We integrate the nonlinear airspring with elastic and viscous contributions according to the simplified Berg model described in Eq. (3) with a linear spring/damper system in a two-mass oscillator configuration that instantiates our quarter railway vehicle model. This model represents only one railway vehicle wheel excited by track irregularities modeled by an input function u , the so-called single-point excitation. Note in particular that the vertical displacements of the wheel are thus assumed to be equal to the irregularities of the track. In this configuration, the mass m_2 is one-eighth of the car-body mass and accompanying masses, while m_1 is one-fourth of the bogie and accompanying masses. The linear spring/damper system, with stiffness and damping coefficients k_1 and d_1 , respectively, constitute the primary suspension, while the airspring model in Eq. (3) represents the secondary.

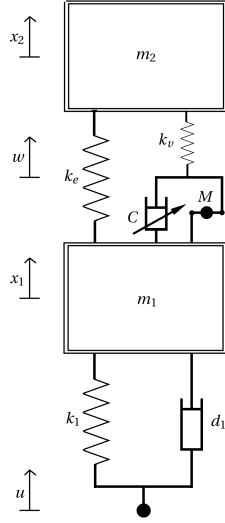


Figure 2. Quarter vehicle model with Berg's airspring model without friction as secondary suspension system.

The equations of motion of the quarter vehicle model depicted in Figure 2 read as

$$\begin{cases} m_2 x_2'' + k_e(x_2 - x_1) + k_v(x_2 - w) = 0, \\ M w'' - k_v(x_2 - w) + C f(w' - x_1') = 0, \\ m_1 x_1'' - k_e(x_2 - x_1) - C f(w' - x_1') + d_1(x_1' - u') \\ \quad + k_1(x_1 - u) = 0, \end{cases} \quad (4)$$

where we set $f(z) := |z|^\beta \text{sign}(z)$ for all $z \in \mathbb{R}$. The latter is a second-order nonlinear differential system that depends on five model parameters: three parameters that define the airspring, namely (k_e, k_v, C) , and two parameters corresponding to the primary suspension, namely (d_1, k_1) .

We reformulate Eq. (4) as a first-order dynamical system in the phase space, resulting in

$$\begin{cases} m_2 v_2' + k_e(x_2 - x_1) + k_v(x_2 - w) = 0, \\ M v_w' - k_v(x_2 - w) + C f(v_w - v_1) = 0, \\ m_1 v_1' - k_e(x_2 - x_1) - C f(v_w - v_1) + d_1(v_1 - u') \\ \quad + k_1(x_1 - u) = 0, \\ x_1' = v_1, \quad x_2' = v_2, \quad w' = v_w. \end{cases} \quad (5)$$

We shall assume that if all the parameters are positive and the external excitation function u is regular enough, given initial states, the dynamical system Eq. (5) admits a unique global solution on a reference time-range, say $\mathcal{T} = [0, T]$, where $T > 0$. In the remainder of this paper, we consider resting initial conditions, namely $x_1(0) = x_2(0) = v_1(0) = v_2(0) = w(0) = v_w(0) = 0$.

The system in Eq. (5) can be discretized with finite differences. Specifically, facing a highly oscillatory behavior, we use a semi-implicit Euler scheme on a uniform grid of $m+1$

points and step $h > 0$, i.e., $\{0, t_1, \dots, t_m\}$ with $t_i = hi$ for all $i \in \{0, \dots, m\}$, and $t_m = T$, thus getting

$$\begin{cases} v_2^{i+1} = v_2^i - \frac{h}{m_2} (k_e(x_2^i - x_1^i) + k_v(x_2^i - w^i)), \\ v_w^{i+1} = v_w^i - \frac{h}{M} (-k_v(x_2^i - w^i) + C f(v_w^i - v_1^i)), \\ v_1^{i+1} = v_1^i - \frac{h}{m_1} (-k_e(x_2^i - x_1^i) - C f(v_w^i - v_1^i) \\ \quad + d_1(v_1^i - (u')^i) + k_1(x_1^i - u^i)), \\ x_1^{i+1} = x_1^i + h v_1^{i+1}, \quad x_2^{i+1} = x_2^i + h v_2^{i+1}, \\ w^{i+1} = w^i + h v_w^{i+1}, \end{cases} \quad (6)$$

for all $i \in \{0, \dots, m-1\}$, with resting initial conditions. The discrete derivatives of $\mathbf{v}_1 := (v_1^0, \dots, v_1^m)$ and $\mathbf{v}_2 := (v_2^0, \dots, v_2^m)$, denoted by \mathbf{v}'_1 and \mathbf{v}'_2 , yield two vectors in \mathbb{R}^m defined by

$$(\mathbf{v}'_1)^i = \frac{(v_1^i)^{i+1} - (v_1^i)^i}{h}, \quad \text{and} \quad (\mathbf{v}'_2)^i = \frac{(v_2^i)^{i+1} - (v_2^i)^i}{h},$$

for all $i \in \{0, \dots, m-1\}$.

3. METHODOLOGY

Once our model in Eq. (6) is fixed, we can now introduce our identification methodology in the presence of safety bounds in form of box constraints.

3.1. Identification model

We denote by θ the vector containing the parameters' values we aim to identify, specifically:

$$\theta := (k_e, k_v, C, k_1, d_1) \in \mathbb{R}_+^5, \quad (7)$$

fix a time step $h > 0$ and consider $\mathbf{x}_{1,\theta}, \mathbf{x}_{2,\theta}, \mathbf{v}_{1,\theta}, \mathbf{v}_{2,\theta} \in \mathbb{R}^{m+1}$ defined through Eq. (6) with model parameters specified by θ , and masses m_1, m_2 and M as given in Table 1. Given some simulated accelerations $\mathbf{a}_\ell = (a_\ell^0, \dots, a_\ell^{m-1}) \in \mathbb{R}^m$ for $\ell \in \{1, 2\}$ of the bogie and the carbody respectively, we shall identify the model parameters solving the following nonlinear constrained optimization problem

$$\min_{\theta \in \mathcal{B}} \|\mathbf{v}'_{1,\theta} - \mathbf{a}_1\|^2 + \|\mathbf{v}'_{2,\theta} - \mathbf{a}_2\|^2, \quad (8)$$

where \mathcal{B} is a set in \mathbb{R}^5 that puts a-priori bounds on the parameters. We set

$$\mathcal{B} := [\tau, 2k_{e,n}] \times [\tau, 2k_{v,n}] \times [\tau, 2C_n] \times [\tau, 2k_{1,n}] \times [\tau, 2d_{1,n}], \quad (9)$$

where $\tau := 10^{-8}$ is a small threshold and $k_{e,n}, k_{v,n}, C_n, k_{1,n}, d_{1,n}$ are the nominal values of k_e, k_v, C, k_1, d_1 , respectively, which are taken from (Berg, 1999) and listed in Table 1. Note that we are imposing that each parameter is positive and upper bounded with twice their nominal values,

Table 1. Nominal model parameters for Eq. (4).

Sym.	Definition	Value (10 ⁻³)	Unit
k_e	elastic airspring stiffness	250	N/m
k_v	viscous airspring stiffness	420	N/m
C	damping stiffness of airspring	11.508	$N(s/m)^{1.8}$
k_1	primary elastic stiffness	282	N/m
d_1	primary damping stiffness	21.9	Ns/m
m_1	mass of quarter bogie	0.7725	kg
m_2	mass of eighth carbody	5.6875	kg
M	mass of air in the airspring	0.218	kg

and that the values of m_1 , m_2 and M are fixed.

Problem (8) can be reformulated in the standard root-finding framework introducing the nonlinear functions $F_1: \mathbb{R}^5 \rightarrow \mathbb{R}^m$ and $F_2: \mathbb{R}^5 \rightarrow \mathbb{R}^m$ defined by

$$F_1(\theta) := \mathbf{v}'_{1,\theta} - \mathbf{a}_1, \quad \text{and} \quad F_2(\theta) := \mathbf{v}'_{2,\theta} - \mathbf{a}_2,$$

Combining F_1 and F_2 to a single function $F: \mathbb{R}^5 \rightarrow \mathbb{R}^{2m}$, we can reformulate Pbl. (8) to the standard form

$$\min_{\theta \in \mathcal{B}} \|F(\theta)\|^2. \quad (10)$$

Note that Pbl. (10) is a particular case of (\mathcal{P}) in (Salzo & Villa, 2012) with $J(\theta) = \mathbf{I}_{\mathcal{B}}(\theta)$, where $\mathbf{I}_{\mathcal{B}}$ is the indicator function of the set \mathcal{B} , i.e., $\mathbf{I}_{\mathcal{B}}(\theta) = +\infty$ if $\theta \notin \mathcal{B}$ and $\mathbf{I}_{\mathcal{B}}(\theta) = 0$ else.

3.2. Optimization scheme

As we described in Section 1.2, to solve the parameter identification problem when the number of parameters is less than the mesh size, it is customary to employ the Gauss–Newton method, which is an iterative method that reads as

$$\theta^{k+1} = \theta^k - \mathbf{H}_k^{-1} \mathbf{J}_k^* F(\theta^k), \quad \text{for } k \in \mathbb{N}, \quad \theta^0 \in \mathbb{R}^5, \quad (11)$$

where, for every $k \in \mathbb{N}$, $\mathbf{H}_k := \mathbf{J}_k^* \mathbf{J}_k$, $\mathbf{J}_k := F'(\theta^k)$ is the Jacobian of F at θ^k and \mathbf{J}_k^* denotes its transpose. Here, we assume that \mathbf{J}_k has full rank, so that \mathbf{H}_k is symmetric positive definite and hence invertible. Note that for every $\theta \in \mathbb{R}^5$ such that the integration of Eq. (6) is feasible with model parameters given by θ , the Jacobian of F at θ can be explicitly obtained with a standard sensitivity analysis, see for instance (Eich-Soellner & Führer, 1998, Chapter 7).

However, a naive implementation of the Gauss–Newton method according to Eq. (11) can lead to serious issues along the iterations as non-physical negative stiffness values or very high coefficients can be reached, both of which can make the integration of Eq. (6), hence the method itself, highly unstable. pGN allows to overcome this issue with a *backward* correction step as follows

$$\theta^{k+1} = \text{prox}_{\psi}^{\mathbf{H}_k} \left(\theta^k - \mathbf{H}_k^{-1} \mathbf{J}_k^* F(\theta^k) \right), \quad \text{for } k \in \mathbb{N}, \quad \theta^0 \in \mathbb{R}^5, \quad (12)$$

where $\psi(\theta) := \mathbf{I}_{\mathcal{B}}(\theta)$, \mathcal{B} being the constraint set defined in Eq. (9). The so-called *proximity operator* $\text{prox}_{\psi}^{\mathbf{H}_k}$ can be defined by means of a minimization problem as follows. Given $\theta \in \mathbb{R}^p$ with $p \in \mathbb{N}$, a p by p symmetric positive definite operator \mathbf{M} , and a proper, convex, lower semicontinuous function ψ , $\text{prox}_{\psi}^{\mathbf{M}}$ is defined by

$$\text{prox}_{\psi}^{\mathbf{M}}(\theta) := \arg \min_{\tilde{\theta} \in \mathbb{R}^p} \psi(\tilde{\theta}) + \frac{1}{2} \|\tilde{\theta} - \theta\|_{\mathbf{M}}^2, \quad (13)$$

where $\|\theta\|_{\mathbf{M}}^2 := \theta^* \mathbf{M} \theta$ for all $\theta \in \mathbb{R}^p$. Note, indeed, that the optimization problem in the right hand-side of Eq. (13) always admits a unique solution, and, thus, $\text{prox}_{\psi}^{\mathbf{M}}$ is a well defined (nonlinear) function on \mathbb{R}^p . If $\psi = \mathbf{I}_{\mathcal{B}}$ and $\mathbf{M} = \mathbf{H}_k$ it is easy to see that Eq. (13) is equivalent to

$$\text{prox}_{\psi}^{\mathbf{H}_k}(\theta) = \arg \min_{\tilde{\theta} \in \mathcal{B}} \|\tilde{\theta} - \theta\|_{\mathbf{H}_k}^2 = \arg \min_{\tilde{\theta} \in \mathcal{B}} \|\mathbf{J}_k \tilde{\theta} - \mathbf{J}_k \theta\|^2, \quad (14)$$

or, in other words, $\text{prox}_{\psi}^{\mathbf{H}_k}$ is the projection onto the set \mathcal{B} with respect to the metric \mathbf{H}_k , which we denote by $\text{P}_{\mathcal{B}}^{\mathbf{H}_k}$. If $\mathbf{H}_k = I$, we will omit the superscript writing $\text{P}_{\mathcal{B}}$ instead of $\text{P}_{\mathcal{B}}^I$. Note that Eq. (14) defines a non-trivial quadratic problem, which in general can be computed only approximately with an inner procedure. Of course, if θ already belongs to \mathcal{B} , no inner iterations are required, which is asymptotically the case when the optimal solution to Pbl. (10) belongs to the interior of \mathcal{B} .

The convergence guarantees in (Salzo & Villa, 2012) are of local type, and they are comparable to those obtained for the classical Gauss–Newton method. In particular, under suitable regularity assumptions on the Jacobian of F , we can expect linear convergence in the general case, and quadratic convergence for zero residual problems, i.e., whenever

$$\|\mathbf{F}(\tilde{\theta})\| = 0, \quad \text{when } \tilde{\theta} \text{ solves Pbl. (10)}. \quad (15)$$

Note that Eq. (15) in our context means that if $\tilde{\theta}$ is the optimal solution to Pbl. (10), the reconstructions $\mathbf{v}'_{h,\tilde{\theta}}$ fit perfectly the simulated accelerations \mathbf{a}_h for $h = 1, 2$, which, however, is unrealistic as \mathbf{a}_h is typically subject to noise. In fact, in our experiments, we observe a linear convergence behavior, even though the regularity assumptions on the Jacobian of F are not necessarily met, cf. Section 4.2.

3.3. Active-set methods: two different approaches

It is important to note that our numerical method closely aligns with classical approaches to solve Pbl. (10). In fact, using the definition of the proximity operator and the fact that for all $\theta \in \mathbb{R}^5$ it holds $\|\theta\|_{\mathbf{H}_k} = \|\mathbf{J}_k \theta\|$, we can see that

Eq. (12) can be expressed as

$$\begin{aligned}\theta^{k+1} &= \arg \min_{\theta \in \mathbb{R}^5} \mathbf{I}_{\mathcal{B}}(\theta) + \frac{1}{2} \|\theta - \theta^k + (\mathbf{J}_k^* \mathbf{J}_k)^{-1} \mathbf{J}_k^* \mathbf{F}(\theta^k)\|_{\mathbf{H}_k}^2 \\ &= \arg \min_{\theta \in \mathcal{B}} \frac{1}{2} \|\theta - \theta^k + (\mathbf{J}_k^* \mathbf{J}_k)^{-1} \mathbf{J}_k^* \mathbf{F}(\theta^k)\|_{\mathbf{H}_k}^2 \\ &= \arg \min_{\theta \in \mathcal{B}} \frac{1}{2} \|\mathbf{J}_k(\theta - \theta^k) + \mathbf{J}_k (\mathbf{J}_k^* \mathbf{J}_k)^{-1} \mathbf{J}_k^* \mathbf{F}(\theta^k)\|^2.\end{aligned}$$

Now, using that $\mathbf{J}_k (\mathbf{J}_k^* \mathbf{J}_k)^{-1} \mathbf{J}_k^*$ is simply the projection operator onto the image of \mathbf{J}_k , we can conclude, in case \mathbf{J}_k has full rank, that

$$\theta^{k+1} = \arg \min_{\theta \in \mathcal{B}} \frac{1}{2} \|\mathbf{J}_k(\theta - \theta^k) + \mathbf{F}(\theta^k)\|^2, \quad (16)$$

which is a box-constrained quadratic problem (Salzo & Villa, 2012) that can be solved with any efficient quadratic solver. This is usually done in practice via active-set strategies, cf. Section 1.2, which can benefit from *warm* and *hot start* capabilities. One of the most efficient implementations of warm and hot started active-set methods is contained in the open-source software qpOASES (Ferreau, Kirches, Potschka, Bock, & Diehl, 2014).

Our computational methodology is based on a different approach. First, we compute the usual Gauss–Newton iteration in Eq. (11) using a simple generic solver¹, and, then, use a cheap first-order method to project the obtained solution onto the feasible set \mathcal{B} . Note that if the projection is computed exactly up to numerical tolerances, then the two iterations would coincide. We compare the numerical performances of these two approaches in Section 4.

Computing the projection While the method is quite robust to errors in the computation of the projections, the choice of the inner procedure to compute the projection does indeed affect the overall performance of the algorithm (Salzo & Villa, 2012; Gonçalves & Menezes, 2020). In our specific instance, we noticed that a generic projected-gradient descent method provides excellent performances. Specifically, at iteration $k \in \mathbb{N}$, our sub-routine that computes the projection of $\tilde{\theta}^{k+1}$ (obtained computing the Gauss–Newton iteration in Eq. (11)) onto \mathcal{B} with respect to the \mathbf{H}_k norm consists of alternating a *forward* step, i.e., evaluating $\theta \mapsto \theta - \sigma_k \mathbf{H}_k(\theta - \tilde{\theta}^{k+1})$ with a suitable $\sigma_k > 0$, and a *backward* step, i.e., a projection (with respect to the ℓ^2 norm) onto \mathcal{B} . The latter can be computed by projecting every component to the corresponding interval. Eventually, our optimization method reads as in Algorithm 1.

Data: The measured accelerations \mathbf{a}_1 and \mathbf{a}_2 .

Result: $\tilde{\theta} = \lim_k \theta^k$ optimal solution to Pbl. (10).

Initialize: $\theta^0 \in \mathcal{B}$;

for $k = 0, 1, 2, \dots$ **do**

 Compute \mathbf{J}_k with sensitivity analysis;

$\mathbf{H}_k \leftarrow \mathbf{J}_k^* \mathbf{J}_k$;

 Solve for s^{k+1} : $\mathbf{H}_k s^{k+1} + \mathbf{J}_k^* \mathbf{F}(\theta^k) = 0$;

$\tilde{\theta}^{k+1} \leftarrow \theta^k + s^{k+1}$;

$\tilde{\theta}_{0,k+1} \leftarrow \tilde{\theta}^{k+1}$;

$\sigma_k \leftarrow \frac{1}{2} \|\mathbf{J}_k\|^2$;

for $n = 0, 1, 2, \dots$ **do**

$\tilde{\theta}_{n+1,k+1} \leftarrow \mathbf{P}_{\mathcal{B}}(\tilde{\theta}_{n,k+1} - \sigma_k \mathbf{H}_k(\tilde{\theta}_{n,k+1} - \tilde{\theta}^{k+1}))$;

$\theta^{k+1} = \lim_n \tilde{\theta}_{n,k+1}$;

Algorithm 1: Proximal Gauss–Newton method to solve Pbl. (10), with a projected-gradient descent as a inner procedure to compute the projection onto \mathcal{B} with respect to the metric \mathbf{H}_k .

4. NUMERICAL EXPERIMENTS

We performed computational experiments to test the proposed method to solve Pbl. (8). All computations were carried out in Python on a Laptop with 16 GB RAM and an Intel Core i7-1255U CPU@1,7-4,7GHz.

4.1. Simulated data generation

In this section, we describe the generation of synthetic railway irregularity profiles and corresponding accelerations. To simulate realistic conditions, we account for noise arising from measurement errors and structured noise introduced through a downsampling procedure. Specifically, railway irregularity data, i.e., u in Eq. (4), is generated synthetically with

$$u(t) := \mathcal{F}(t) \left(\sum_{j=1}^r \xi_j \sin\left(\frac{2\pi\nu}{\lambda_j} t\right) + \eta_j \cos\left(\frac{2\pi\nu}{\lambda_j} t\right) \right), \quad (17)$$

where $\mathcal{F}: \mathbb{R} \rightarrow \mathbb{R}$ is a *fade-in* function to guarantee resting initial conditions, which satisfies $\mathcal{F}(0) = \mathcal{F}'(0) = 0$ and $\mathcal{F}(t) = 1$ for all $t \geq 1$, e.g.,

$$\mathcal{F}(t) := (1 - \min\{\max\{1.2 - t, 0\}, 1\})^2,$$

$r = 500$ is the signal complexity, ξ_1, \dots, ξ_r are numbers sampled from a Gaussian distribution with zero mean and variance $2 \cdot 10^{-4}$, and η_1, \dots, η_r are defined by $\eta_j := \tilde{\eta}_j - \sum_{i=1}^r \tilde{\eta}_i$, where $\tilde{\eta}_1, \dots, \tilde{\eta}_r$ are again sampled from a Gaussian distribution with zero mean and variance $2 \cdot 10^{-4}$. In this way, $u(0) = 0$, and, due to the fade-in function, also $u'(0) = 0$. In our experiments, we picked $\nu = 10$ and $\lambda_j, \tilde{\lambda}_j$ sampled from a Gaussian distribution with mean 11 and variance 3, in such a way that the resulting signals always simulate railway irregularities of average wavelength 11 *m*, and average amplitude of 3 *cm* for a railway vehicle moving at 10 *m/s*.

¹In our numerical experiments, we have used `numpy.linalg.solve` in Python.

Eventually, we sample the continuous function $u: \mathbb{R} \rightarrow \mathbb{R}$ on a uniform time mesh with $\tilde{m} = 10^5$ nodes and final time $T = 10$ s.

If the underlying model parameters, i.e., k_e, k_v, C, k_1, d_1 , are specified by some $\tilde{\theta} \in \mathbb{R}^5$, we generate the accelerations \mathbf{a}_1 and \mathbf{a}_2 synthetically by solving Eq. (6) with data u, u' as above, model parameters $\tilde{\theta}$, and masses m_1, M, m_2 specified in Table 1. To mimic *measured* data, we:

- Apply white noise of variance $\sigma A(\mathbf{a}_1)$ and $\sigma A(\mathbf{a}_2)$ to \mathbf{a}_1 and \mathbf{a}_2 , respectively, where, for $\ell = 1, 2$, $A(\mathbf{a}_\ell)$ denotes the maximal amplitude of \mathbf{a}_ℓ , i.e., $A(\mathbf{a}_\ell) := \max \mathbf{a}_\ell - \min \mathbf{a}_\ell$,
- *Downsample* $\mathbf{a}_1, \mathbf{a}_2, u, u'$ from the uniform time mesh with $\tilde{m} = 10^5$ nodes to a uniform time mesh with $m = 10^3$ nodes, i.e., we consider a sampling rate of 100 Hz. To do so, we compute the mean of the values for every chunk of 10^2 nodes.

We emphasize that compared to the generated synthetic data, experimental data would be further corrupted by more structured *process error*. Indeed, while the one-wheel approximation can provide useful insights and simplify the proposed identification approach, it fails to capture the full range of dynamic interactions and structural complexities of a railway vehicle bogie. Experimental data would therefore also be influenced by 3D interactions between wheels, axles, and the body of the vehicle, among other effects. For this reason, our synthetic data is only a first approximation of actual experimental data, as it includes only noise from sensor error (up to 20%) and noise arising from the down-sampling procedure described in Section 4.1. Note that the latter can be understood as a first approximation of process error. Altogether, exploring the effect of actual experimental data is a fundamental avenue for future research, which can be addressed with the proposed methodology.

4.2. Experiments

To simulate a potential faulty scenario, the underlying model parameters are assumed to be equal to the nominal parameters, with two exceptions: k_1 is set to be 20% higher and k_e is set to be 20% lower than the corresponding nominal values. We refer to these parameters as the *ground-truth* values and collect them in a vector $\tilde{\theta}$. Once the model parameters are fixed, we perform the following set of experiments.

1. Robustness to noise. We consider five levels of noise ranging from $\sigma = 0\%$ to $\sigma = 20\%$. Here, the former is an ideal zero noise case, and the latter is a very noisy scenario where the data has been corrupted with a noise of variance of magnitude equal to 20% of the amplitude of the signal. For each noise level σ , following Section 4.1, we gen-

erate railway irregularity data and simulated measured accelerations 20 times with a white noise of level σ . In each case, we run Algorithm 1 with initialization $\theta^0 = \theta_n$, where $\theta_n \in \mathbb{R}^5$ is the vector of nominal model parameters. We stop the outer and inner loops as soon as $\|\theta^{k+1} - \theta^k\| \leq 10^{-5}$ or $k = K$ with $K = 500$, and $\|\tilde{\theta}_{n+1,k} - \tilde{\theta}_{n,k}\| \leq 10^{-5}$ or $n = 10000$, respectively. For comparison, on the same data, we also employ a state-of-the-art active-set approach solving the quadratic subproblem, i.e., Eq. (16), using the quadratic solver qpOASES (Ferreau et al., 2014) (GN-QP). To conclude, we repeat the same experiment but with $\tilde{\theta} = \theta_n$, i.e., we simulate no actual fault in the system components.

If the noise increases, the initialization θ^0 might be out of the basin of attraction of the method, see Theorem 1 in (Salzo & Villa, 2012), and, therefore, the method could cycle or behave wildly, yet yielding always feasible iterates, i.e., with $\theta^k \in \mathcal{B}$. We will therefore say that the method *fails* if the maximum amount of iterations is reached. We also count the number of iterations in which the projection subroutine onto \mathcal{B} performs at least one iteration (PA), or, in other words, the number of k such that $\tilde{\theta}^{k+1}$ does *not* belong to \mathcal{B} . This allows us to keep track of the number of times that Algorithm 1 activates the aforementioned *backward* correction step. Additionally, we compute mean and standard deviation of the reconstructed parameters (Reconstructed parameters) at the final iteration and their Relative Error (RE parameters) with respect to the ground-truth values, that means, the value $RE(p) := \frac{|p - \tilde{p}|}{\tilde{p}}$ for each parameter p either k_e, k_v, C, k_1 or d_1 , where \tilde{p} is the corresponding ground-truth value. Eventually, we compute the eigenvalues of the positive definite matrix $\mathbf{H}(\tilde{\theta}) := \mathbf{F}'(\tilde{\theta})^* \mathbf{F}'(\tilde{\theta})$, where $\mathbf{F}'(\tilde{\theta})$ is the Jacobian of \mathbf{F} at $\tilde{\theta}$, $\tilde{\theta}$ being the vector of ground-truth parameters. The condition number of this matrix provides us with two interconnected properties. On the one hand, it is related to the *identifiability* of the system² according to (Grewal & Glover, 1976). On the other hand, together with the noise level, it influences the diameter of the basin of attraction of the method, see (Salzo & Villa, 2012, Theorem 1). All the results are summarized in Tables 2, 3, 4, and 5. Note that the reconstructed parameters and the relative errors have been rounded to the second and the first decimal digit, respectively. In Figure 3, we show how the reconstructed acceleration fits the simulated data, and in Figure 4 we display a bar chart with the eigenvalues of $\mathbf{H}(\tilde{\theta})$ and of $\mathbf{H}(\theta_n)$.

2. Comparison with active-set methods. We perform a time comparison of a state-of-the-art active-set approach and the proposed pGN approach according to Algorithm 1. In the former, at each iteration, we solve Eq. (16) using

²Note that, due to the presence of noise, $\tilde{\theta}$ is only an approximation of the actual minimizer, but in low-noise regimes, this can be a good approximation.

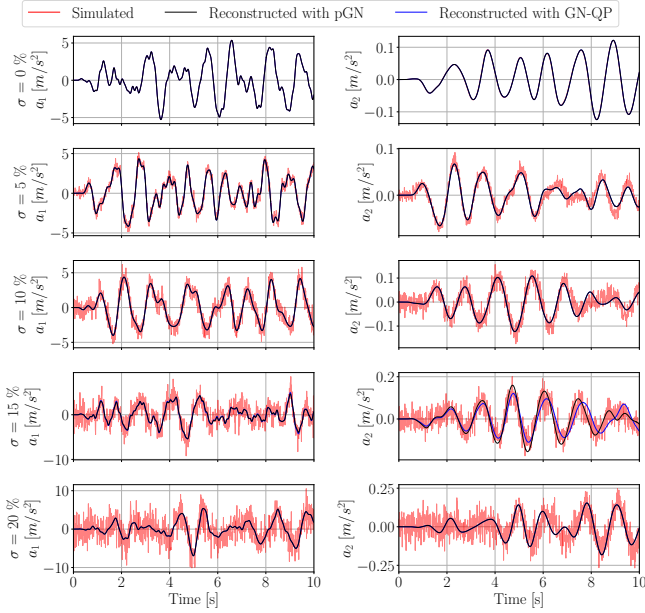


Figure 3. Simulated and reconstructed estimates of the accelerations a_1 and a_2 under several noisy regimes. Confer Section 4.2 point 1 for a detailed description.

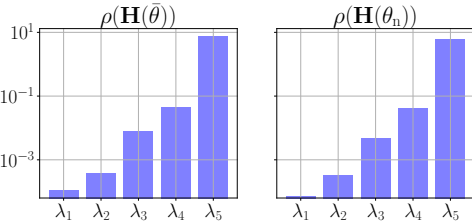


Figure 4. Eigenvalues of the approximate Hessian $\mathbf{H}(\theta) := \mathbf{F}'(\theta) * \mathbf{F}'(\theta)$, for $\theta = \bar{\theta}$ the vector of ground-truth (faulty) parameters and $\theta = \theta_n$ the vector of nominal parameters.

the state-of-the-art quadratic solver qpOASES (Ferreau et al., 2014). For the latter, we consider the same inner stopping criteria that we employed in our first experiment. For the two methods, we set the maximum number of outer iterations to $k = 50$. We generate railway irregularity data and synthetic measured accelerations according to Section 4.1. We corrupt data with an error level of $\sigma = 5\%$ and run the two identification methods. If one of these two *fail*, i.e., does not reach the prescribed accuracy within the maximum number of iterations allowed, we corrupt data once again and repeat the experiment until both identification methods succeed. We do so 30 times. In each case, we measure: *final time*, i.e., the time employed for the method to perform 50 iterations, and *sub-routine time*, namely, the total time employed without considering the computation of the sensitivity matrices, i.e., the operators \mathbf{J}_k in our notation. Means and standard deviations of our results are summarized in Table 6. The average distance to minimizer as a

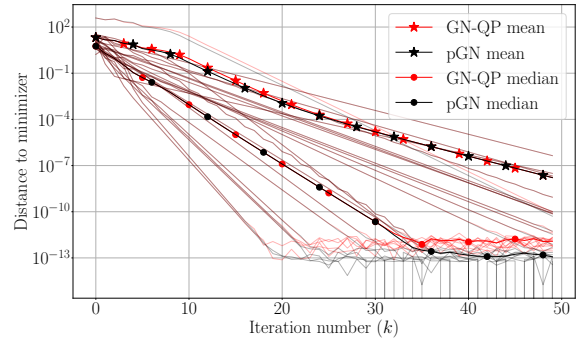


Figure 5. Distance to minimizer as a function of the iteration number. Aggregate results of 30 independent runs in a comparison between pGN according to Algorithm 1 and the Gauss–Newton method with the qpOASES subroutine (GN–QP) to solve Eq. (16). Confer Section 4.2 point 2 for a detailed description.

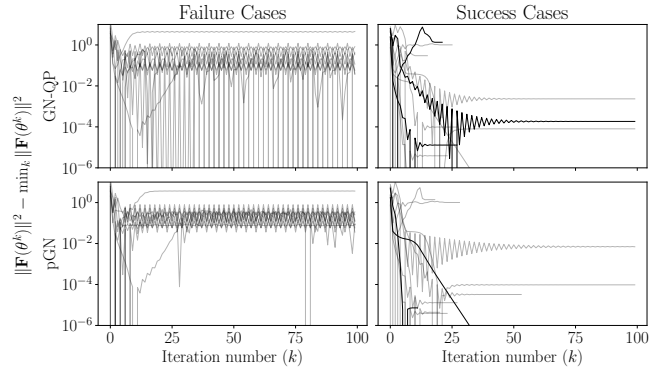


Figure 6. For $\sigma = 10\%$, objective values across iterations for failure and success cases using Algorithm 1 and the quadratic solver qpOASES to solve Eq. (16) at each iteration $k \in \mathbb{N}$. Five distinct convergence behaviors are highlighted with thick lines, cf. Section 4.3 for further comments.

function of the iteration is shown in Figure 5.

4.3. Results

We begin our discussion with the outcomes of our initial experiment reported in Tables 2 and 3. As expected, under the ideal scenario of zero noise, we observe that the residual of the pGN method converges to the prescribed tolerance after only a few iterations (approximately 7) with highly accurate estimations of the ground-truth parameters. Note, however, that the relative error is not zero due to the downsampling. The situation changes when we introduce a noise level of 5%. In this scenario, we detect increased uncertainty in estimating the airspring's characteristics, particularly for parameter C , associated with the nonlinear damping component in the airspring. Nonetheless, this increased uncertainty does not compromise the

Table 2. Faulty case, results of experiment in point 1 in Section 4.2 utilizing pGN to approximate Eq. (16).

Noise level	Info	Reconstructed parameters ($\times 10^{-3}$)	RE parameters (%)
$\sigma = 0\%$	Success: 20 Failures: 0 Iter.: 7 (0) PA: 0 (0)	k_e : 200.39 (0.0) k_v : 424.23 (0.0) C: 10.83 (0.0) k_1 : 330.88 (0.0) d_1 : 21.81 (0.0)	k_e : 0.2 (0.0) k_v : 1.0 (0.0) C: 5.9 (0.0) k_1 : 2.2 (0.0) d_1 : 0.4 (0.0)
$\sigma = 5\%$	Success: 13 Failures: 7 Iter.: 214 (229) PA: 24 (106)	k_e : 205.42 (21.1) k_v : 435.26 (20.86) C: 10.2 (20.86) k_1 : 329.5 (3.36) d_1 : 21.8 (0.25)	k_e : 9.4 (5.5) k_v : 4.2 (4.5) C: 38.8 (19.4) k_1 : 2.6 (1.0) d_1 : 1.1 (0.6)
$\sigma = 10\%$	Success: 18 Failures: 2 Iter.: 73 (143) PA: 2 (4)	k_e : 191.85 (52.71) k_v : 445.18 (75.02) C: 14.01 (75.02) k_1 : 335.65 (10.15) d_1 : 21.86 (0.77)	k_e : 22.1 (14.9) k_v : 8.8 (16.6) C: 51.9 (30.9) k_1 : 2.4 (1.9) d_1 : 2.7 (2.3)
$\sigma = 15\%$	Success: 10 Failures: 10 Iter.: 262 (237) PA: 131 (194)	k_e : 161.77 (87.83) k_v : 605.98 (194.79) C: 18.51 (194.79) k_1 : 331.97 (17.26) d_1 : 21.95 (1.51)	k_e : 42.0 (23.0) k_v : 48.5 (41.9) C: 84.0 (29.2) k_1 : 4.7 (2.8) d_1 : 6.5 (2.3)
$\sigma = 20\%$	Success: 10 Failures: 10 Iter.: 261 (239) PA: 224 (221)	k_e : 256.37 (149.45) k_v : 336.0 (411.51) C: 12.21 (411.51) k_1 : 321.26 (15.09) d_1 : 20.85 (2.17)	k_e : 68.6 (40.9) k_v : 100.0 (0.0) C: 77.1 (36.7) k_1 : 6.3 (2.5) d_1 : 10.0 (4.7)

estimation accuracy of the elastic component k_e characteristics of the airspring, which exhibit an average accuracy of about 90% in all considered cases. Even as the noise level rises further, causing the estimation of C to deteriorate rapidly, it is noteworthy that primary components k_1 and d_1 still yield acceptable estimations with an average accuracy of 90%, even at a noise level of $\sigma = 20\%$. In particular, the reconstructed accelerations demonstrate a robust fit to the simulated ones, even in regimes with high levels of noise, as illustrated in Figure 3.

One potential reason for the increased loss of accuracy at higher noise levels can be deduced from Figure 4, which displays the eigenvalues of the approximate Hessian $\mathbf{H}(\theta) := \mathbf{F}'(\theta)^* \mathbf{F}'(\theta)$ for $\theta = \bar{\theta}$ (the ground-truth parameters) and $\theta = \theta_n$ (the nominal parameters). We observe in particular that the condition numbers of these matrices are on the order of 10^5 , with smallest eigenvalues on the order of 10^{-4} . This suggests an inherent *identifiability* issue (according to (Grewal & Glover, 1976)) in the considered system, or, in other words, that solving Pbl. (8) is intrinsically difficult and obtaining inaccurate estimations of some system parameters (C in this case) might be unavoidable. Note as well that according to Theorem 1 in (Salzo & Villa, 2012), high condition numbers of these approximate Hessians at the minimizer (together with the noise level) lead to smaller basins of attraction of the method, making pGN (as well as any other methodology based on Gauss–Newton such

Table 3. Faulty case, results of experiment in point 1 in Section 4.2 utilizing qpOASES to approximate Eq. (16).

Noise level	Info	Reconstructed parameters ($\times 10^{-3}$)	RE parameters (%)
$\sigma = 0\%$	Success: 20 Failures: 0 Iter.: 7 (0) PA: 0 (0)	k_e : 200.39 (0.0) k_v : 424.23 (0.0) C: 10.83 (0.0) k_1 : 330.88 (0.0) d_1 : 21.81 (0.0)	k_e : 0.2 (0.0) k_v : 1.0 (0.0) C: 5.9 (0.0) k_1 : 2.2 (0.0) d_1 : 0.4 (0.0)
$\sigma = 5\%$	Success: 13 Failures: 7 Iter.: 214 (229) PA: 0 (0)	k_e : 205.42 (21.1) k_v : 435.26 (20.86) C: 10.2 (20.86) k_1 : 329.5 (3.36) d_1 : 21.8 (0.25)	k_e : 9.4 (5.5) k_v : 4.2 (4.5) C: 38.8 (19.4) k_1 : 2.6 (1.0) d_1 : 1.1 (0.6)
$\sigma = 10\%$	Success: 18 Failures: 2 Iter.: 73 (143) PA: 0 (0)	k_e : 193.1 (53.03) k_v : 445.1 (75.01) C: 14.01 (75.01) k_1 : 335.41 (10.09) d_1 : 21.85 (0.77)	k_e : 22.4 (14.6) k_v : 8.8 (16.6) C: 51.9 (30.9) k_1 : 2.4 (2.0) d_1 : 2.6 (2.3)
$\sigma = 15\%$	Success: 7 Failures: 13 Iter.: 335 (223) PA: 0 (0)	k_e : 205.66 (87.09) k_v : 659.95 (199.27) C: 16.57 (199.27) k_1 : 323.35 (16.49) d_1 : 21.14 (1.41)	k_e : 42.1 (11.5) k_v : 59.6 (44.2) C: 77.1 (32.5) k_1 : 6.1 (2.5) d_1 : 6.9 (2.5)
$\sigma = 20\%$	Success: 9 Failures: 11 Iter.: 283 (239) PA: 0 (0)	k_e : 237.81 (139.57) k_v : 466.67 (417.4) C: 9.8 (417.4) k_1 : 322.85 (15.38) d_1 : 20.95 (2.15)	k_e : 63.3 (34.8) k_v : 100.0 (0.0) C: 85.4 (28.4) k_1 : 5.9 (2.5) d_1 : 10.1 (3.6)

as any traditional approach that solves the inner quadratic problem with specific quadratic solvers like qpOASES) systematically fail for large noise levels, as we can also observe in Tables 4 and 5. It is important to note, though, that noise levels greater than 10% are unrealistic in practical scenarios, where sensor errors are usually of the order of 5%. Nonetheless, we tested the method in these extreme cases to demonstrate the limits of our approach.

The typical convergence behavior of the method is illustrated in Figure 6. We observe in particular that success and failure regimes for a noise level of $\sigma = 10\%$ can be quickly distinguished. In failure cases, the iterates exhibit high oscillations for all $k \in \mathbb{N}$. Conversely, in success cases, the oscillations tend to vanish rapidly. Thus, failing regimes can be identified *on the fly* during the identification process, which represents a particularly appealing feature for applications to FDI. In the same figure, we also observe that success regimes can display a variety of behaviors. Some cases show the method oscillating for more than 25 iterations before settling into the basin of attraction of an isolated minimum, while in other cases, the method converges to a solution in fewer than 10 iterations. Although failure and success cases are easily distinguishable visually, efficient automatic classification requires dedicated criteria depending on the specific case study and the level of noise considered. One approach to handling failing cases is to repeat the identification process with different data. Naturally, persis-

Table 4. Non faulty case, results of experiment in point 1 in Section 4.2 utilizing pGN to approximate Eq. (16).

Noise level	Info	Reconstructed parameters ($\times 10^{-3}$)	RE parameters (%)
$\sigma = 0\%$	Success: 20	k_e : 254.38 (0.0)	k_e : 1.8 (0.0)
	Failures: 0	k_v : 422.52 (0.0)	k_v : 0.6 (0.0)
	Iter.: 6 (0)	C: 10.45 (0.0)	C: 9.2 (0.0)
	PA: 0 (0)	k_1 : 274.35 (0.0)	k_1 : 2.7 (0.0)
		d_1 : 21.7 (0.0)	d_1 : 0.9 (0.0)
$\sigma = 5\%$	Success: 20	k_e : 251.34 (29.03)	k_e : 9.5 (6.7)
	Failures: 0	k_v : 421.54 (21.36)	k_v : 4.1 (3.0)
	Iter.: 54 (67)	C: 12.85 (21.36)	C: 37.1 (31.3)
	PA: 1 (4)	k_1 : 276.99 (4.86)	k_1 : 2.1 (1.4)
		d_1 : 21.74 (0.31)	d_1 : 1.3 (0.9)
$\sigma = 10\%$	Success: 10	k_e : 250.27 (28.95)	k_e : 8.7 (7.6)
	Failures: 10	k_v : 599.27 (215.34)	k_v : 51.4 (42.5)
	Iter.: 279 (225)	C: 14.95 (215.34)	C: 61.4 (36.5)
	PA: 190 (216)	k_1 : 277.79 (6.46)	k_1 : 2.3 (1.5)
		d_1 : 21.65 (0.38)	d_1 : 1.4 (1.6)
$\sigma = 15\%$	Success: 9	k_e : 212.29 (100.42)	k_e : 33.3 (27.0)
	Failures: 11	k_v : 646.88 (293.87)	k_v : 80.1 (37.3)
	Iter.: 284 (238)	C: 14.86 (293.87)	C: 89.5 (21.5)
	PA: 238 (230)	k_1 : 283.18 (11.34)	k_1 : 3.1 (2.6)
		d_1 : 22.04 (0.44)	d_1 : 1.8 (1.1)
$\sigma = 20\%$	Success: 4	k_e : 292.67 (86.03)	k_e : 27.0 (27.3)
	Failures: 16	k_v : 301.56 (178.82)	k_v : 33.8 (38.2)
	Iter.: 410 (181)	C: 11.51 (178.82)	C: 100.0 (0.0)
	PA: 378 (183)	k_1 : 272.22 (11.01)	k_1 : 3.5 (3.8)
		d_1 : 21.16 (0.74)	d_1 : 4.2 (2.3)

tent and repetitive failures in the context of FDI can also be indicative. These might suggest a major component fault, where the nominal parameters consistently fail to fall into a basin of attraction of the method.

It is interesting that according to Tables 4 and 5 the same behavior occurs even if the ground-truth coincides with the nominal parameters and the initialization of the method, no matter the methodology utilized to approximate the quadratic subroutine. This is due to the fact that, due to downsampling (process error), the ground-truth data does not necessarily coincide with the actual minimizer.

Concerning our second experiment, Table 6 reveals only a slight acceleration in computation when employing a state-of-the-art quadratic solver to address Eq. (16), achieving a speedup of approximately 10% at a sub-routine level. However, this efficiency gain does not translate into an overall time reduction, as the main computational bottleneck remains the computation of the sensitivity matrix, a task present in both scenarios. From Figure 5, we also observe that, even if the projection is not computed up to numerical tolerances, overall, the distance to the minimizer (computed letting the algorithm run for 250 iterations) decreases with similar behavior. Further, the convergence rate seems to be linear even though the regularity assumptions in (Salzo & Villa, 2012) are not necessarily met. Notably, it appears that inexact computations do not impact

Table 5. Non faulty case, results of experiment in point 1 in Section 4.2 utilizing qpOASES to approximate Eq. (16).

Noise level	Info	Reconstructed parameters ($\times 10^{-3}$)	RE parameters (%)
$\sigma = 0\%$	Success: 20	k_e : 254.38 (0.0)	k_e : 1.8 (0.0)
	Failures: 0	k_v : 422.52 (0.0)	k_v : 0.6 (0.0)
	Iter.: 6 (0)	C: 10.45 (0.0)	C: 9.2 (0.0)
	PA: 0 (0)	k_1 : 274.35 (0.0)	k_1 : 2.7 (0.0)
		d_1 : 21.7 (0.0)	d_1 : 0.9 (0.0)
$\sigma = 5\%$	Success: 20	k_e : 251.6 (28.67)	k_e : 9.3 (6.7)
	Failures: 0	k_v : 421.6 (20.92)	k_v : 4.1 (2.9)
	Iter.: 54 (67)	C: 12.85 (20.92)	C: 37.1 (31.3)
	PA: 0 (0)	k_1 : 276.95 (4.81)	k_1 : 2.0 (1.4)
		d_1 : 21.74 (0.31)	d_1 : 1.3 (0.9)
$\sigma = 10\%$	Success: 11	k_e : 250.93 (26.54)	k_e : 8.4 (6.5)
	Failures: 9	k_v : 621.48 (216.28)	k_v : 55.8 (42.9)
	Iter.: 251 (226)	C: 13.52 (216.28)	C: 65.2 (35.9)
	PA: 0 (0)	k_1 : 276.9 (6.23)	k_1 : 2.4 (1.5)
		d_1 : 21.69 (0.41)	d_1 : 1.4 (1.5)
$\sigma = 15\%$	Success: 9	k_e : 241.68 (102.24)	k_e : 31.7 (26.1)
	Failures: 11	k_v : 588.67 (227.33)	k_v : 50.5 (44.7)
	Iter.: 285 (237)	C: 12.31 (227.33)	C: 89.6 (21.4)
	PA: 0 (0)	k_1 : 277.61 (16.39)	k_1 : 4.8 (3.6)
		d_1 : 22.01 (0.43)	d_1 : 1.7 (1.1)
$\sigma = 20\%$	Success: 7	k_e : 233.83 (102.77)	k_e : 33.1 (25.2)
	Failures: 13	k_v : 527.2 (200.93)	k_v : 35.4 (41.1)
	Iter.: 355 (221)	C: 3.29 (200.93)	C: 100.0 (0.0)
	PA: 0 (0)	k_1 : 278.61 (12.74)	k_1 : 3.5 (3.1)
		d_1 : 22.04 (0.7)	d_1 : 2.6 (2.0)

Table 6. Time comparison between the pGN method and the Gauss–Newton method with qpOASES subroutine to solve Eq. (16). Confer Section 4.2 point 2 for a detailed description.

	mean (std)	median	unit
Final time pGN	1.274835 (0.839708)	1.555865	s
Final time GN–QP	1.259149 (0.816052)	1.552521	s
Sub-routine time pGN	0.016892 (0.014701)	0.017929	s
Sub-routine time GN–QP	0.013765 (0.009359)	0.015717	s

this linear rate of convergence.

5. DISCUSSION

Gauss–Newton-type methods are a widely used choice in the context of parameter identification in dynamical systems. In this paper, we proposed an efficient alternative which can easily handle the presence of physical boundaries. Altogether, the results allow us to conclude that the research question in Section 1.1 can be answered positively, but they also show the limitations of the approach. Indeed, our method consistently delivers satisfactory results in low-noise regimes and demonstrates remarkable robustness to noise for primary suspension characteristics. However, it is also important to acknowledge that the identification of parameters C and k_v could be significantly affected by noise, especially if the Hessian of the objective at the isolated minimizer is ill-conditioned. These issues, which

seem related to the local convergence behavior of Gauss–Newton-type methods, can be addressed with additional computational overhead by employing suitable *globalization* strategies in the spirit of (Gonçalves & Menezes, 2020). Note, however, that for pGN, these were established only under quite restrictive assumptions, and in the context of parameter identification and FDI they could significantly impact performances.


6. CONCLUSIONS

While quite efficient but arguably involved approaches to quickly solve constrained parameter identification problems are already well established in the literature, the pGN method outlined in Algorithm 1 emerges as a competitive alternative, with a simple implementation that does not rely on any black-box software to run efficiently. Additionally, the method exhibits the important feature of generating sequences of solution estimates that remain within physical boundaries, ensuring operational stability also when data is corrupted with high noise.

We showcased the effectiveness of the pGN method through a compelling application to parameter identification of a quarter railway vehicle's model characteristics, particularly in the presence of an airspring in the secondary suspension system. The results showcased the robust performance of the proposed methodology, achieving an accuracy of over 90% for the primary suspension system. However, due to identifiability issues, airspring parameter identification was more sensitive to noise, yet without compromising overall performance. In conclusion, we expect that pGN will prove quite useful in the community, with a high potential also for larger models with more nonlinear components.

Moving forward, we plan to explore the performance of our identification methodology under different forms of noise, further investigate the identifiability of Pbl. (8), and develop a more sensitive approach to identify nonlinearities in suspension systems.

ACKNOWLEDGMENTS

This work has been supported by the TraDE-OPT project which received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 861137 .

This work was written at Virtual Vehicle Research GmbH in Graz and partially funded within the COMET K2 Competence Centers for Excellent Technologies from the Austrian Federal Ministry for Climate Action (BMK), the Austrian Federal Ministry for Labour and Economy (BMAW), the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promo-

tion Agency (FFG) has been authorised for the programme management.

REFERENCES

- Bard, Y. (1974). *Nonlinear parameter estimation*. Academic Press.
- Berg, M. (1997). A model for rubber springs in the dynamic analysis of rail vehicles. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 211(2), 95–108.
- Berg, M. (1999). A three-dimensional airspring model with friction and orifice damping. *Vehicle System Dynamics*, 33(sup1), 528–539.
- Bruni, S., Goodall, R., Mei, T. X., & Tsunashima, H. (2007). Control and monitoring for railway vehicle dynamics. *Vehicle System Dynamics*, 45(7-8), 743–779.
- Callejo, A., & de Jalón, J. G. (2015). Vehicle suspension identification via algorithmic computation of state and design sensitivities. *Journal of Mechanical Design*, 137(2).
- Ding, J., Pan, Z., & Chen, L. (2012). Parameter identification of multibody systems based on second order sensitivity analysis. *International Journal of Non-Linear Mechanics*, 47(10), 1105–1110.
- Eich-Soellner, E., & Führer, C. (1998). *Numerical methods in multibody dynamics* (1st ed.). Vieweg+Teubner Verlag Wiesbaden.
- Ferreau, H., Kirches, C., Potschka, A., Bock, H., & Diehl, M. (2014). qpOASES: A parametric active-set algorithm for quadratic programming. *Mathematical Programming Computation*, 6(4), 327–363.
- Gonçalves, M., & Menezes, T. (2020). Gauss–Newton methods with approximate projections for solving constrained nonlinear least squares problems. *Journal of Complexity*, 58, 101459.
- Grewal, M., & Glover, K. (1976). Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 21(6), 833–837.
- Grupp, F., & Kortüm, W. (1993). Parameter identification of nonlinear descriptor systems. In *Advanced multibody system dynamics* (pp. 457–462). Springer.
- Kraft, S., Puel, G., Aubry, D., & Funfschilling, C. (2016). Parameter identification of multi-body railway vehicle models – application of the adjoint state approach. *Mechanical Systems and Signal Processing*, 80, 517–532.
- Matsumiya, S., Nishioka, S., Nishimura, S., & Suzuki, M. (1969). On the diaphragm air spring “Sumiride”. *Sumitomo Search*, 2, 86–92.
- Mazzola, L., & Berg, M. (2014). Secondary suspension of railway vehicles-air spring modelling: Performance and critical issues. *Proceedings of the Institution of*

- Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 228(3), 225–241.
- Moheyeldein, M., Abd-El-Tawwab, A. M., El-gwwad, K. A., & Salem, M. (2018). An analytical study of the performance indices of air spring suspensions over the passive suspension. *Beni Suef University Journal of Basic and Applied Sciences*, 7(4), 525–534.
- Oda, N., & Nishimura, S. (1969). Vibration of air suspension bogies and their design. *JSME International Journal Series B Fluids and Thermal Engineering*, 13, 43–50.
- Puel, G., Bourgeteau, B., & Aubry, D. (2013). Parameter identification of nonlinear time-dependent rubber bushings models towards their integration in multi-body simulations of a vehicle chassis. *Mechanical Systems and Signal Processing*, 36(2), 354–369.
- Salzo, S., & Villa, S. (2012). Convergence analysis of a proximal Gauss–Newton method. *Computational Optimization and Applications*, 53(2), 557–589.
- Sayyaadi, H., & Shokouhi, N. (2009). A new model in rail-vehicles dynamics considering nonlinear suspension components behavior. *International Journal of Mechanical Sciences*, 51(3), 222–232.
- Serban, R., & Freeman, J. S. (2001). Identification and identifiability of unknown parameters in multibody dynamic systems. *Multibody System Dynamics*, 5(4), 335–350.
- Strano, S., & Terzo, M. (2019). Review on model-based methods for on-board condition monitoring in railway vehicle dynamics. *Advances in Mechanical Engineering*, 11(2).
- Vyasarayani, C., Uchida, T., Carvalho, A., & McPhee, J. (2012). Parameter identification in dynamic systems using the homotopy optimization approach. *Lecture Notes in Control and Information Sciences*, 418, 129–145.
- Xin-Chun, Z., & Cheng-Jun, G. (2013). Cubature Kalman filters: Derivation and extension. *Chinese Physics B*, 22(12), 128401.
- Zoljic-Beglerovic, S., Luber, B., Stettinger, G., Müller, G., & Horn, M. (2020). Parameter identification for railway suspension systems using cubature Kalman filter. In *Advances in dynamics of vehicles on roads and tracks* (pp. 128–132). Springer International Publishing.
- Zoljic-Beglerovic, S., Stettinger, G., Luber, B., & Horn, M. (2018). Railway suspension system fault diagnosis using cubature Kalman filter techniques. *IFAC-PapersOnLine*, 51(24), 1330–1335. (10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2018)