Hindawi
Structural Control and Health Monitoring
Volume 2023, Article ID 8855542, 18 pages
https://doi.org/10.1155/2023/8855542

Research Article

A Model-Based Bayesian Inference Approach for On-Board Monitoring of Rail Roughness Profiles: Application on Field Measurement Data of the Swiss Federal Railways Network

Charikleia D. Stoura,[1] **Vasilis K. Dertimanis**,[1] **Cyprien Hoelzl**,[1] **Claudia Kossmann**,[2] **Alfredo Cigada**,[3] **and Eleni N. Chatzi**[1]

[1]*Institute of Structural Engineering, Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Stefano-Franscini-Platz 5, Zürich 8093, Switzerland*
[2]*Infrastructure–Facilities & Technology-Track, Swiss Federal Railways (SBB AG), Hilfikerstrasse 3, Bern 3000, Switzerland*
[3]*Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 34, Milan 20156, Italy*

Correspondence should be addressed to Charikleia D. Stoura; charikleia.stoura@ibk.baug.ethz.ch

Received 14 August 2023; Revised 6 November 2023; Accepted 8 December 2023; Published 29 December 2023

Academic Editor: Yong Xia

According to the International Union of Railways, railway networks count more than one million kilometers of tracks worldwide, a number that is to rise further as the goal is to promote rail transportation as a sustainable means to face the challenge of increased mobility. However, such a vast expansion further necessitates efficient and reliable infrastructure monitoring schemes able to guarantee the quality and safety of rail transportation. Traditional monitoring approaches, relying on visual inspection and portable measuring devices, cannot rise to the task as they do not allow for continuous inspection of extended portions of rail infrastructure. Therefore, mobile monitoring methodologies based on dedicated diagnostic vehicles have emerged as an alternative. Despite revolutionizing traditional monitoring methods, such vehicles are usually expensive and can only operate under the suspension of regular rail service. In this work, we propose an alternative approach for mobile sensing of railway infrastructure based on on-board monitoring data collected from low-cost vibration sensors, e.g., accelerometers, which can be mounted on in-service trains. Specifically, we focus on identifying the roughness profile of the tracks and propose a fusion of reduced-order vehicle models with a Bayesian inference approach for joint input-state estimation. To enhance the inference, we opt for a prior updating of the vehicle model parameters on the basis of an unscented Kalman filter and available measurements from a diagnostic vehicle. The key contributions of this work are (i) the consideration of the dynamic interaction between trains and tracks, which is usually ignored in rail roughness estimation, (ii) the adoption of reduced train vehicle models that decrease the computational effort of the identification task, (iii) the updating of the vehicle parameters to account for inconsistencies in the model used, and (iv) the application of the proposed methodology to actual acceleration measurements collected from a diagnostic vehicle of the Swiss Federal Railways network.

1. Introduction

To guarantee the quality and safety of rail transportation, monitoring and assessing the condition of rail infrastructure is imperative. Focusing on the track system, continuous monitoring can prohibit the development of faults at an early stage, alleviating damage to rail and vehicle components, rolling noise, and passenger discomfort due to excessive vibrations, and ensuring rail transport safety. Such monitoring traditionally relies on visual inspection or nondestructive evaluation techniques that employ portable measuring devices, such as hand-held trolleys (Figure 1(a)), that perform ultrasonic inspection and eddy-current tests, or gauges that measure the geometry of tracks [1]. Despite reflecting a high degree of fidelity, these techniques entail high costs and effort, while they cannot cover extended

portions of rail infrastructure on a frequent basis. As a result, mobile sensing approaches have been explored as an alternative [4].

Recently used mobile sensing solutions comprise track recording vehicles (TRVs) equipped with optical and inertial sensors, such as laser scanners and high-speed cameras, collecting geometric data by periodically traveling on specific sections of a railway network. Such vehicles gather information about the track geometry along the longitudinal and alignment levels, track twists, and track gradients [5]. The gezogenes Diagnosefahrzeug (gDfZ) vehicle of the Swiss Federal Railways (SBB) comprises such a recording vehicle (Figure 1(b)) [2, 6]. The gDfZ uses laser scanners to assess rail roughness and employs acceleration sensors placed at different locations of the vehicle to capture vibration acceleration data [7]. While such diagnostic vehicles provide valuable insights into track properties, including roughness and isolated defects, and allow for a more dense coverage of the railway network compared to traditional monitoring approaches, they come with some drawbacks. They typically operate at lower speeds than regular trains, especially high-speed trains, to ensure accurate scanning of the roughness profile [8]. Therefore, they can only operate while regular rail service is suspended, hindering the continuous supply of data regarding the condition of tracks. Moreover, these diagnostic vehicles are typically costly due to their advanced sensor equipment, such as laser profilometers, and require frequent maintenance and updates of their measuring systems [8].

A more recent mobile sensing direction involves fitting in-service trains with simple and low-cost sensing systems, most commonly micro-electromechanical system (MEMS) acceleration sensors [3]. This approach enables the continuous gathering of data from railway tracks in a cost-effective manner. This enhances track monitoring over time, allowing tracking potential degradation and planning scheduled maintenance. SBB has deployed solutions in this direction, with the roll-out of the ICN passenger train (Figure 1(c)) [3], which is equipped with accelerometers in various locations and a pair of tensiometric wheels to measure contact forces [4]. The objective is to create a comprehensive monitoring platform by integrating the high-quality, diverse data collected by the dedicated gDfZ diagnostic vehicle with the extensive vibration data collected by the ICN passenger train. The ICN train regularly traverses different parts of the SBB network, offering an abundance of data thatfacilitates continuous track monitoring.

Likely, the most straightforward method to determine rail roughness profiles based on vibration data is yielded via double integration of acceleration measurements from different vehicle parts [9, 10]. However, integrating the measured signals often leads to low accuracy due to integration errors relating to noise content and associated drifts at low frequencies. In alleviating this issue, different signal filtering techniques have alternatively been used to infer roughness profiles or detect and assess defects in track systems. Lee et al. [11] used axle box and bogie vibration acceleration data and used a mixed filtering approach, including Kalman, bandpass, and compensation filters, to estimate the underlying rail roughness profile. A different approach, proposed by O'Brien et al. [12], estimated roughness profiles based on bogie vertical accelerations and angular velocities using a cross-entropy optimization technique. The common element of these identification approaches is that they neglect the interaction between vehicles and tracks, i.e., they ignore the physics of the system while solving the inverse problem.

Recent studies have moved toward the incorporation of the dynamic interaction between rail vehicles and tracks into the identification task. Dertimanis et al. [13] simulated a simple 6 degree-of-freedom (DOF) vehicle running on a specified roughness profile and applied a Bayesian inference approach, i.e., a dual Kalman filter (DKF) [14], to recover the employed roughness profile based on simulated acceleration data from the vehicle's front wheel. This work forms an initiating study towards incorporating coupled dynamics in roughness identification approaches. However, it relies on a simple vehicle-track configuration that cannot demonstrate the method's applicability to real train-track systems. Later, Li et al. [15] used an adaptive Kalman filter to estimate the roughness of rails using a three-dimensional (3D) vehicle model. This method demonstrated very good results for the identification of rail roughness; however, the acceleration and velocity of all DOFs of the system's body and bogies were used in the analysis to ensure sufficient quality in the identification. In addition, the proposed method was verified only in simulated data derived from the same system as the one that was used for the inverse analysis [15]; this procedure, sometimes referred to as an "inverse crime," [16] somewhat facilitates the inference task. The approaches of Dertimanis et al. [13] and Li et al. [15] reconstruct rail roughness profiles on the basis of measured accelerations. Other studies examine the possibility of estimating roughness profiles based on further types of data, e.g., strain measurements collected from strain gauges and angular velocities measured via gyroscopes [17, 18]. In this work, we do not further elaborate on such studies, as this research exploits accelerations as a form of low cost, easily measurable data for developing an on-board monitoring (OBM) methodology for reconstructing roughness profiles.

To this end, we propose a hybrid approach, fusing a physics-based model with monitoring data, in order to identify rail roughness profiles from acceleration data collected from the axle boxes of instrumented train vehicles. We here account for the physics underpinning the coupled train-track interaction system [19, 20] and perform the identification of rail roughness based on reduced-order railroad vehicle models and a recursive Bayesian inference method [14]. Specifically, the vehicles are modeled as multibody systems, where different bodies are connected with springs and dampers. The contact between the vehicles and the rails assumes an elastic normal contact model [21]. The vehicle models are first reduced and then employed in the identification task via their state-space representation. The identification of roughness relies on a dual implementation of the Kalman filter, the DKF [14], which estimates the input and states of the vehicle system in a sequential manner. Before this identification task, which can be assumed to be also executed in near real time, an updating of the interaction model is carried out. Specifically, assuming the availability of

(a) (b) (c)

FIGURE 1: Track monitoring methods: (a) hand-push trolley for direct inspection of tracks [1], (b) diagnostic vehicle of SBB network for testing rail infrastructure [2], and (c) revenue train of SBB equipped with accelerometers [3].

a prior measurement of a part of the roughness profile, possibly from a previous diagnostic measurement, we employ an unscented Kalman filter (UKF) [22] to accomplish a joint state-parameter estimation, which allows us to calibrate certain parameters of the vehicle model that affect the identification of the input. Although this implementation is inspired by the joint input-state-parameter estimation scheme proposed by Dertimanis et al. [23], the tasks of state-parameter and input-state estimation are herein executed asynchronously, capitalizing on the availability of different types of information: (i) geometric data from diagnostic trains (even if infrequent) and (ii) accelerations from in-service OBM trains. The goal of the proposed approach is to make use of measurement acceleration data from diagnostic or in-service trains equipped with low-cost acceleration sensors. Therefore, for our initial field application, we utilize readily available actual OBM data from the gDfZ diagnostic vehicle of SBB. This implementation attests to the feasibility of the proposed approach for practical applications employing also in-service trains.

The remainder of this manuscript is organized as follows: Section 2 describes a generic railroad vehicle model excited by rail roughness, which is then used to formulate the inverse problem of estimating the states and input of the vehicle. Then, Section 3 describes two alternative schemes to reduce the order of the multibody vehicle model. Section 4 demonstrates the proposed Bayesian inference framework for estimating rail roughness profiles in tracks. Section 5 presents verifications and validations on numerically simulated and on-board measured vibration data, respectively, proving the applicability of the proposed scheme to real-world problems, and, lastly, Section 6 summarizes the presented methodology.

2. Problem Formulation

With respect to Figure 2, the vertical dynamics of a single railroad vehicle traveling along a fixed track can be described by a second-order vector differential equation of motion (EOM) of the form

$$\mathbf{m}\ddot{\mathbf{u}}(t) + \mathbf{c}\dot{\mathbf{u}}(t) + \mathbf{k}\mathbf{u}(t) = \mathbf{W}\mathbf{p}(t), \tag{1}$$

where $\mathbf{m}$, $\mathbf{c}$, and $\mathbf{k}$ are the $[N \times N]$ mass, non-proportional damping, and stiffness matrices, respectively, $\mathbf{u}$ is the $[N \times 1]$ vector of displacements and rotations of the vehicle's DOFs (refer to Sections 3.1 and 3.2 and Table 1), and $\mathbf{p}(t)$ is the vector of contact forces between the vehicle and the track. The latter is applied to the vehicle via the contact direction matrix $\mathbf{W}$ and follows an elastic normal contact model [21], i.e.,

$$\mathbf{p}(t) = k\mathbf{r}(x = vt), \tag{2}$$

where k is the contact stiffness between the wheels and the rails and $\mathbf{r}(x)$ is the vector of the roughness profile of the rails at the contact points with the wheels. The roughness profile changes in space with respect to the location of the vehicle x, which is a function of the running speed v and time t.

By forming the $[2N \times 1]$ state vector

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{u}(t) \\ \dot{\mathbf{u}}(t) \end{bmatrix}, \tag{3}$$

and by assuming the availability of vibration acceleration response measurements, equation (1) can be transformed into state-space as

$$\dot{\mathbf{x}}(t) = \mathbf{A}_c\mathbf{x}(t) + \mathbf{B}_c\mathbf{p}(t), \tag{4a}$$

$$\mathbf{y}(t) = \mathbf{C}_c\mathbf{x}(t) + \mathbf{D}_c\mathbf{p}(t), \tag{4b}$$

in which

$$\begin{aligned} \mathbf{A}_c &= \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{m}^{-1}\mathbf{k} & -\mathbf{m}^{-1}\mathbf{c} \end{bmatrix}, \mathbf{B}_c = \begin{bmatrix} \mathbf{0} \\ \mathbf{m}^{-1}\mathbf{W} \end{bmatrix}, \\ \mathbf{C}_c &= \left[-\mathbf{S}\mathbf{m}^{-1}\mathbf{k} - \mathbf{S}\mathbf{m}^{-1}\mathbf{c}\right], \\ \mathbf{D}_c &= \mathbf{S}\mathbf{m}^{-1}\mathbf{W}, \end{aligned} \tag{5}$$

are the state, input, output, and feedforward matrices, respectively, and $\mathbf{S}$ is a DOF selection matrix of appropriate dimensions.

Under this setting, the aim of the current study is the estimation of the rail roughness using noise-corrupted vibration acceleration data and Bayesian inference. Our approach is built on the following assumptions:

Table 1: DOFs of the multibody railroad vehicle model of Figure 2.

DOFs	Car body	Bogies ($i = 1 - 2$)	Wheelsets ($i = 1 - 4$)
Vertical displacement	z^c	z^{ti}	z^{wi}
Lateral displacement	y^c	y^{ti}	y^{wi}
Yawing rotation	ψ^c	ψ^{ti}	ψ^{wi}
Rolling rotation	ϕ^c	ϕ^{ti}	ϕ^{wi}
Pitching rotation	θ^c	θ^{ti}	θ^{wi}

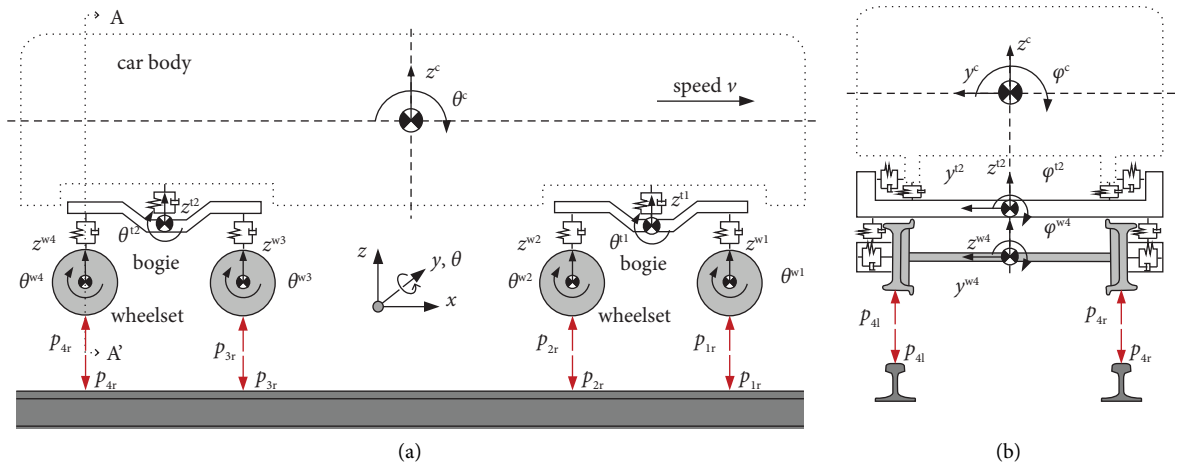

Figure 2: A 3D model of the vertical dynamics of a railroad vehicle, indicating the dominant DOFs and contact forces: (a) pitch-bounce view and (b) roll-bounce view (section A-A').

(1) The contact force modeling assumes the track system as a rough surface, comprising the rail roughness profile, which introduces an external excitation to the vehicle system.

(2) Each set of wheels runs on the same roughness profile, i.e., the left/right wheel of each wheelset traverses the same left/right roughness profile.

It is further noted that full-order vehicle models typically entail a large number of DOFs. The induced computational cost renders such models inappropriate for online estimation, without necessarily adding value to the inference task. To this end, prior to the implementation of the Bayesian observer, model-order reduction (MOR) is applied, using eigenvalue-based modal truncation and substructuring. MOR also allows for eliminating unstable modes and ensuring that the vehicle system used for the inverse analysis is controllable and observable. This is outlined in the next section.

3. Model Reduction

3.1. Eigenvalue Analysis. MOR based on eigenvalue analysis is a well-known approach for reducing the size of complicated models. It relies on a decomposition of the EOM of the system in modal equations and, subsequently, truncation of modes. Different from common systems, the train vehicle adopted herein is a multibody, not proportionally damped system, while zero-mass elements corresponding to springs and dampers of the suspension system of the model may exist (depending on the complexity of the train model). Therefore, the eigenvalue decomposition is performed based on a factorization of the state-space system matrix of the vehicle as [24]

$$\mathbf{A}_c = \mathbf{V}\mathbf{L}\mathbf{V}^{-1}, \tag{6}$$

where $\mathbf{L} = \mathrm{diag}(\mathbf{\Lambda}, \mathbf{\Lambda}^*)$ is a diagonal matrix consisting of the complex conjugate eigenvalues of the system and $\mathbf{V}$ is the eigenvector matrix written as

$$\mathbf{V} = \begin{bmatrix} \mathbf{\Psi} & \mathbf{\Psi}^* \\ \mathbf{\Psi}\mathbf{\Lambda} & \mathbf{\Psi}^*\mathbf{\Lambda}^* \end{bmatrix}. \tag{7}$$

Matrix $\mathbf{V}$ contains the complex conjugate eigenmode pairs, where $\square^*$ denotes the complex conjugate and $\mathbf{\Psi}$ is a matrix containing the modal vectors in the physical space. Since the system is not proportionally damped, $\mathbf{\Psi}$ is a complex matrix.

Having defined the modal basis of the system, the state-space system of equation (2) assumes the form

$$\dot{\zeta}(t) = \mathbf{A}_m\zeta(t) + \mathbf{B}_m\mathbf{p}(t), \tag{8a}$$

$$\mathbf{y}(t) = \mathbf{C}_m\zeta(t) + \mathbf{D}_m\mathbf{p}(t), \tag{8b}$$

where $\zeta(t)$ is the modal state vector. The state and measurement matrices, initially defined by equation (5), become

$$\mathbf{A}_m = \mathbf{V}^{-1}\mathbf{A}_c\mathbf{V}, \mathbf{B}_m = \mathbf{V}^{-1}\mathbf{B}_c, \mathbf{C}_m = \mathbf{C}_c\mathbf{V}, \quad \mathbf{D}_m = \mathbf{D}_c. \tag{9}$$

To reduce the size of the model, only a subset $\tilde{\zeta}(t)$ of $\zeta(t)$ is retained, while the rest of the modes are truncated. The modes that remain are chosen based on a dominance

measure D_r proposed by Föllinger et al. [25]. This measure represents the significance of each mode on the transfer function of the system, and it is expressed as

$$D_r = \left| \frac{c_{r,i} b_{r,j}}{\lambda_r} \right|, \tag{10}$$

where the coefficient $c_{r,i}$ is the i-th element of the r-th column of $\mathbf{C}_m$ matrix connecting the i-th output to the r-th mode of the system. Accordingly, the coefficient $b_{r,j}$ is the j-th element of the r-th row of $\mathbf{B}_m$ matrix connecting the j-th input to the r-th mode. Lastly, λ_r is the eigenvalue of each mode. Eventually, the reduced-order system consists of a reduced number of modes encapsulated in $\widetilde{\mathbf{\Lambda}}$ and $\widetilde{\mathbf{\Psi}}$, representing the matrices of the selected eigenvalues and eigenvectors, respectively. Accordingly, the state-space matrices of equation (9) are transformed into $\widetilde{\mathbf{A}}_m$, $\widetilde{\mathbf{B}}_m$, and $\widetilde{\mathbf{C}}_m$. The feedforward matrix $\mathbf{D}_m$ remains unchanged as it is not associated with the truncated state vector.

3.2. Substructuring. MOR based on substructuring allows for component-based monitoring, i.e., monitoring only a subsystem (substructure) of the entire assembly [26]. The specification of the substructure to be modeled depends on the point of application of the applied input as well as the position of the deployed sensors, where the output is measured. Following such an approach, the number of model parameters that need to be specified is significantly decreased. For example, in the case of roughness identification on the basis of axle box acceleration (ABA) measurements from the vehicle, the objective is to identify the force input on the vehicle wheels (in contact with the tracks). In this case, the monitored substructure consists of the respective wheelsets (where the input is applied) and axle boxes (where the measurements are collected from).

To this end, in view of Figure 3, the vehicle model is partitioned into two substructures: the upper part substructure, consisting of the car body and bogies, and the lower part substructure, consisting of the wheelsets and axle boxes. Subsequently, equation (1) can be written as

$$\begin{bmatrix} \mathbf{m}^u & \mathbf{0} \\ \mathbf{0} & \mathbf{m}^l \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{u}}^u(t) \\ \ddot{\mathbf{u}}^l(t) \end{bmatrix} + \begin{bmatrix} \mathbf{c}^u & \mathbf{c}^{\mathrm{ul}} \\ \mathbf{c}^{\mathrm{lu}} & \mathbf{c}^l \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}}^u(t) \\ \dot{\mathbf{u}}^l(t) \end{bmatrix} + \begin{bmatrix} \mathbf{k}^u & \mathbf{k}^{\mathrm{ul}} \\ \mathbf{k}^{\mathrm{lu}} & \mathbf{k}^l \end{bmatrix} \begin{bmatrix} \mathbf{u}^u(t) \\ \mathbf{u}^l(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^l \end{bmatrix} \mathbf{p}(t), \tag{11}$$

where the superscript $\square^u$ denotes the upper substructure and $\square^l$ the lower substructure, or in other words, the wheelset-axle box system. Superscripts $\square^{\mathrm{ul}}$ and $\square^{\mathrm{lu}}$ indicate the matrices that connect the upper to the lower substructure of the vehicle. To create independent, non-overlapping substructures, equation (11) can be rewritten as

$$\begin{bmatrix} \mathbf{m}^u & \mathbf{0} \\ \mathbf{0} & \mathbf{m}^l \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{u}}^{\mathrm{u}}(t) \\ \ddot{\mathbf{u}}^l(t) \end{bmatrix} + \begin{bmatrix} \mathbf{c}^u & \mathbf{0} \\ \mathbf{0} & \mathbf{c}^l \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}}^u(t) \\ \dot{\mathbf{u}}^l(t) \end{bmatrix} + \begin{bmatrix} \mathbf{k}^u & \mathbf{0} \\ \mathbf{0} & \mathbf{k}^l \end{bmatrix} \begin{bmatrix} \mathbf{u}^u(t) \\ \mathbf{u}^l(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^l \end{bmatrix} \mathbf{p}(t) - \begin{bmatrix} \mathbf{c}^{\mathrm{ul}} & \mathbf{0} \\ \mathbf{0} & \mathbf{c}^{\mathrm{lu}} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}}^l(t) \\ \dot{\mathbf{u}}^u(t) \end{bmatrix} - \begin{bmatrix} \mathbf{k}^{\mathrm{ul}} & \mathbf{0} \\ \mathbf{0} & \mathbf{k}^{\mathrm{wu}} \end{bmatrix} \begin{bmatrix} \mathbf{u}^l(t) \\ \mathbf{u}^u(t) \end{bmatrix}, \tag{12}$$

where the last two terms of the right-hand side of equation (12) denote the internal forces between different substructures. Subsequently, the substructure of interest, i.e., the lower part, is isolated as

$$\mathbf{m}^l \ddot{\mathbf{u}}^l(t) + \mathbf{c}^l \dot{\mathbf{u}}^l(t) + \mathbf{k}^l \mathbf{u}^l(t) = \mathbf{W}^l \mathbf{p}(t) - \mathbf{c}^{\mathrm{lu}} \dot{\mathbf{u}}^u(t) - \mathbf{k}^{\mathrm{lu}} \mathbf{u}^u(t), \tag{13}$$

and it can be written as

$$\mathbf{m}^l \ddot{\mathbf{u}}^l(t) + \mathbf{c}^l \dot{\mathbf{u}}^l(t) + \mathbf{k}^l \mathbf{u}^l(t) = \mathbf{W}^l \mathbf{p}(t) + \mathbf{g}(t), \tag{14}$$

where $\mathbf{W}^l \mathbf{p}(t)$ represents the external forces and $\mathbf{g}(t)$ represents the internal forces acting on the substructure. Equation (14) can then be expanded as

$$\begin{bmatrix} \mathbf{m}^w & \mathbf{0} \\ \mathbf{0} & \mathbf{m}^a \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{u}}^w(t) \\ \ddot{\mathbf{u}}^a(t) \end{bmatrix} + \begin{bmatrix} \mathbf{c}^w & \mathbf{c}^{\mathrm{wa}} \\ \mathbf{c}^{\mathrm{aw}} & \mathbf{c}^a \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}}^w(t) \\ \dot{\mathbf{u}}^a(t) \end{bmatrix} + \begin{bmatrix} \mathbf{k}^w & \mathbf{k}^{\mathrm{wa}} \\ \mathbf{k}^{\mathrm{aw}} & \mathbf{k}^a \end{bmatrix} \begin{bmatrix} \mathbf{u}^w(t) \\ \mathbf{u}^a(t) \end{bmatrix} = \begin{bmatrix} \mathbf{W}^w \mathbf{p}(t) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{g}^a(t) \end{bmatrix}, \tag{15}$$

where the superscript $\square^w$ denotes the wheelsets and $\square^a$ indicates the axle boxes connecting the lower part of the vehicle to the upper part. Thus, the axle boxes can be considered as the boundary DOFs, and the wheelsets as the internal ones. Consequently, the contact force vector $\mathbf{p}(t)$ acts only on the wheelsets and corresponds to external forces. At the same time, $\mathbf{g}^a(t)$ is an internal force vector acting on the axle boxes, and is typically orders of magnitude smaller than the contact force vector [20].

This substructuring of the system allows for isolating components from further pieces whose structural parameters are either unknown or suffer uncertainties in their definition. Equation (15) can be further reduced following the eigenmode analysis described in Section 3.1; thus, the

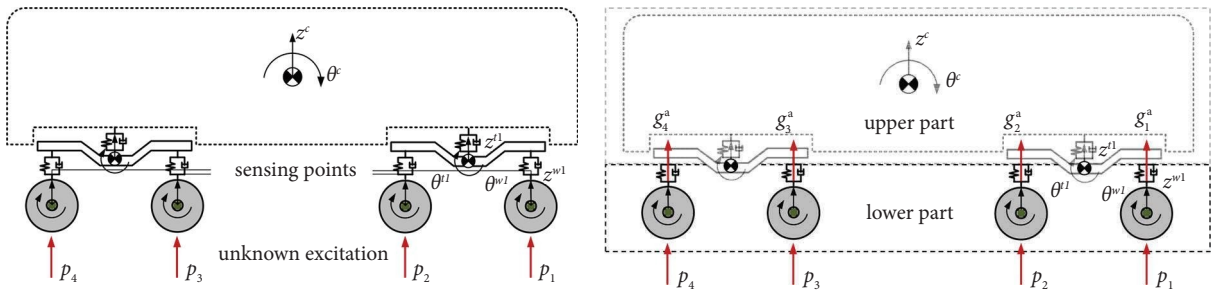

FIGURE 3: Schematic representation of (a) the full-order vehicle model and (b) the substructure-based reduced-order model for joint input-state estimation.

state-space system assumes the form of equations (8a) and (8b). To this end, potential rigid body motions should be considered [26].

The substructure-based system of equation (15) indicates that data are collected from sensors on the axle boxes of the railroad vehicle. Alternatively, if data are collected from sensors mounted on other parts of the vehicle, e.g., the bogies, these components need to also be considered in the substructure of interest.

4. Rail Roughness Identification

4.1. Discrete-Time Equations. Since the acquired data arrives in digital format, the state-space system of equations (8a) and (8b) needs to be discretized at an appropriate sampling rate f_s. In doing so, the Tustin approximation [27] is herein applied. By notating as $H_a(s)$ and $H_d(z)$ the associated input-output transfer functions of the analog and digital systems, respectively, it follows that

$$H_d(z) = H_a(s)|_{s=(2/T_s)\cdot(z-1/z+1)}, \tag{16}$$

where $T_s = 1/f_s$. By applying equation (16) to equations (8a) and (8b), the digital state-space model of the system is written as

$$\widetilde{\zeta}_{k+1} = \widetilde{\mathbf{A}}_d\widetilde{\zeta}_k + \widetilde{\mathbf{B}}_d\mathbf{p}_k, \tag{17a}$$

$$\mathbf{y}_k = \widetilde{\mathbf{C}}_d\widetilde{\zeta}_k + \widetilde{\mathbf{D}}_d\mathbf{p}_k, \tag{17b}$$

in which

$$\widetilde{\mathbf{A}}_d = \left(\mathbf{I} - \frac{f_s}{2}\mathbf{A}_m\right)^{-1}\left(\mathbf{I} + \frac{f_s}{2}\mathbf{A}_m\right), \widetilde{\mathbf{B}}_d = \sqrt{\frac{f_s}{2}}\left(\mathbf{I} - \frac{f_s}{2}\mathbf{A}_m\right)^{-1}\mathbf{B}_m,$$
$$\widetilde{\mathbf{C}}_d = \sqrt{\frac{f_s}{2}}\mathbf{C}_m\left(\mathbf{I} - \frac{f_s}{2}\mathbf{A}_m\right)^{-1}, \widetilde{\mathbf{D}}_d = \mathbf{D}_m + \frac{1}{2\sqrt{f_s}}\mathbf{C}_m\widetilde{\mathbf{B}}_d. \tag{18}$$

Since the identification of rail roughness relies on a Bayesian inference approach, zero-mean Gaussian white noise is superimposed to the state and output equations as

$$\widetilde{\zeta}_{k+1} = \widetilde{\mathbf{A}}_d\widetilde{\zeta}_k + \widetilde{\mathbf{B}}_d\mathbf{p}_k + \mathbf{w}_k, \tag{19a}$$

$$\mathbf{y}_k = \widetilde{\mathbf{C}}_d\widetilde{\zeta}_k + \mathbf{D}_d\mathbf{p}_k + \mathbf{r}_k, \tag{19b}$$

where $\mathbf{w}_k$ and $\mathbf{r}_k$ denote the discrete-time process and measurement noise terms, with known covariance matrices $\mathbf{Q}_w = \mathbb{E}[\mathbf{w}_k\mathbf{w}_k^T]$ and $\mathbf{Q}_r = \mathbb{E}[\mathbf{r}_k\mathbf{r}_k^T]$, respectively. The stochastic discrete-time state-space representation of the system, as described by equations (19a) and (19b), is henceforth implemented to the reconstruction of the roughness profile.

4.2. Input Estimation via a Bayesian Inference Approach. The identification of rail roughness relies on estimating the contact forces acting on the vehicle, assuming that a state-space representation of the vehicle is known. In this case, the DKF for joint input-state estimation can directly be employed for estimating the unknown contact forces. However, certain parameter values of the assumed vehicle model may significantly deviate from the real parameter values of the actual railroad vehicle used to collect acceleration data. The contact stiffness between the wheels and the rails comprises such a parameter, which, on the one hand, is typically unknown and, on the other hand, can significantly affect the identification task. This problem implies that this unknown parameter should also be inferred.

One way to achieve such a task is to opt for joint input-state-parameter estimation by first estimating the input via a Kalman filter, assuming known states and parameters, and then augmenting the parameters with the state vector and applying a UKF for state-parameter estimation [23]. This approach works well when certain parameters of the system $\mathbf{A}_c$ or output $\mathbf{C}_c$ matrices are unknown or partially known. However, in this case, this approach of joint or synchronous estimation is not applicable, since the unknown contact stiffness parameter and input quantities appear in a bivariate product (equation (2)), rendering their individual identification task an ill-posed one.

Under this setting, the applied methodology for rail roughness profile identification depends on the quality of the assumed vehicle model and the prior availability of data. To

this end, this study proposes different classes of Bayesian observers, illustrated in Figure 4 and discussed in the following sections.

4.2.1. Case I: "Perfectly" Known Vehicle Model. The simplest scenario calls for the availability of a vehicle model that resembles the dynamics of the actual measurement vehicle to a great extent. It is noted that model errors may still be present; however, these can be effectively handled by the process and measurement noise terms of equations (19a) and (19b). In this case, the roughness profile of the rails is estimated by a double implementation of the Kalman filter that allows for concurrent identification of the states and inputs of the reduced-order vehicle system. This recursive Bayesian inference approach constitutes the DKF proposed by Eftekhar Azam et al. [14]. This method adopts a random walk model for the temporal evolution of the unknown input, i.e.,

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{v}_k, \tag{20}$$

where $\mathbf{v}_k$ is a discrete-time input noise term with covariance matrix $\mathbf{Q}_v = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T]$. According to Eftekhar Azam et al. [14], assuming that an estimate of the state $\widetilde{\zeta}_k$ at time t_k is available, a dual state-space equation can be formulated as

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{v}_k, \tag{21a}$$

$$\mathbf{y}_k = \widetilde{\mathbf{C}}_d \widetilde{\zeta}_k + \mathbf{D}_d \mathbf{p}_k + \mathbf{r}_k, \tag{21b}$$

where the state $\widetilde{\zeta}_k$ plays the role of the "known input." This way the implementation of a typical Kalman filter allows for the estimation of $\mathbf{p}_k$ through equations (21a) and (21b). Then, the estimated input $\mathbf{p}_k$ is employed in equations (19a) and (19b), and a successive Kalman filter is applied for the estimation of $\widetilde{\zeta}_{k+1}$.

Algorithm 1 shows the detailed steps of the DKF for the estimation of the unknown states and inputs. Eftekhar Azam et al. [14] demonstrated that the accuracy of the identification task via the DKF relies on three parameters: the process noise covariance $\mathbf{Q}_w$, the input noise covariance $\mathbf{Q}_v$, and the measurement noise covariance $\mathbf{Q}_r$. The process noise covariance matrix $\mathbf{Q}_w$ represents the accuracy of the adopted physical model, $\mathbf{Q}_v$ depends on the noise of the input, and $\mathbf{Q}_r$ depends on the accuracy of the measurement instruments used. $\mathbf{Q}_w$ and $\mathbf{Q}_r$ can be specified for a given model and measurement device. On the other hand, $\mathbf{Q}_v$ is typically adopted as the tuning parameter. Therefore, an L-curve is used to tune the input noise and optimize the performance of the adopted filter [28].

4.2.2. Case II: Imperfect Vehicle Model and Part of Input Available. When a small portion of the roughness profile is known, e.g., from former measurement campaigns, a model calibration task may be realized prior to the implementation of the DKF. Such an example constitutes measurements of the longitudinal level of tracks from a diagnostic vehicle equipped with laser scanners. In this case, a prior inference of the uncertain contact stiffness parameter between the vehicle model and the track can be accomplished. Accordingly, the calibrated vehicle model is used to identify the rail roughness profile, as described in Algorithm 1. Model calibration is herein realized via the UKF [22, 29] for joint state-parameter estimation. To this end, the state vector is augmented with the parameter vector as

$$\overline{\mathbf{x}} = \begin{bmatrix} \widetilde{\zeta} \\ k \end{bmatrix}, \tag{22}$$

where k corresponds to the uncertain stiffness parameter. The corresponding covariance matrix $\mathbf{Q}_z$ includes the discrete-time process and parameter noise terms. Thus, the UKF only solves for one single state vector, the augmented vector $\overline{\mathbf{x}}$.

The UKF models the augmented state vector as a Gaussian random variable with a distribution approximated by a set of deterministic points termed the sigma points [29]. Sigma points capture the prior mean and covariance of the state and, when propagated through the nonlinear function, provide an improved posterior estimate of the transformed state, which constitutes the unscented transformation [22]. The spread of the Sigma points around the mean state value is controlled by α and κ parameters. The spread of the Sigma points is proportional to a, which is usually a small positive value, and proportional to the square root of κ, which is often set to zero. A third parameter β affects the weights of the propagated states when calculating the state and measurement covariance. Parameter β incorporates prior knowledge of the distribution of the states, which for Gaussian distribution is optimally set as $\beta = 2$. The calculation of the weights for the mean of the predicted measurements is as follows:

$$W_m^0 = 1 - \frac{N}{\alpha^2 (N + \kappa)}, \tag{23a}$$

$$W_m^i = \frac{1}{2\alpha^2 (N + \kappa)}, i = 1, \ldots, 2N, \tag{23b}$$

where N is the number of states. Accordingly, the calculation of the weights for the covariance of the predicted measurements is as follows:

$$W_c^0 = \left(2 - \alpha^2 + \beta^2\right) - \frac{N}{\alpha^2 (N + \kappa)}, \tag{24a}$$

$$W_c^i = \frac{1}{2\alpha^2 (N + \kappa)}, i = 1, \ldots, 2N. \tag{24b}$$

Algorithm 2 demonstrates in detail the successive steps of the UKF.

4.2.3. Case III: Imperfect Vehicle Model and Part of Filtered Input Available. In case a filtered portion of the input, rather than the input itself, is known, the UKF cannot directly be applied for the estimation of the unknown parameter. For example, as shown later in Section 5.3, a portion of the roughness profile is scanned with laser scanners, but the measurements need to be filtered in certain wavelengths before being used to calibrate the contact

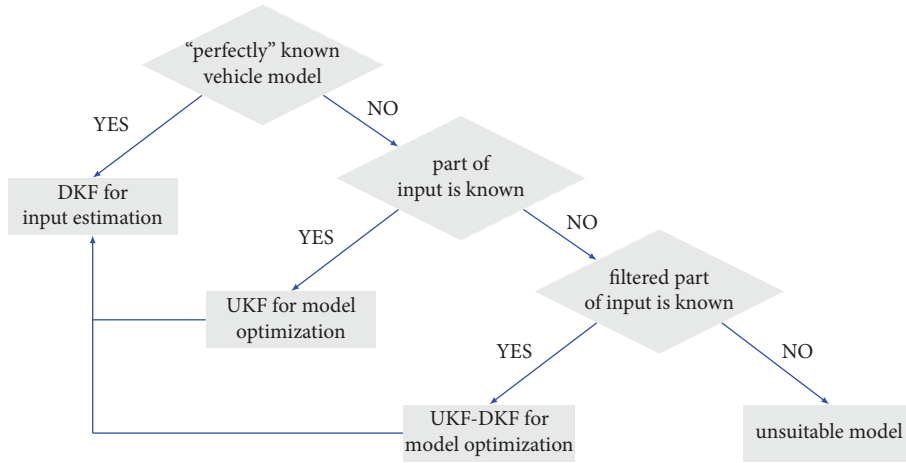

Figure 4: Flowchart for the selection of the appropriate Bayesian framework for the estimation of rail roughness profiles in tracks.

Initialize
Set the initial values for the state vector mean $\widetilde{\zeta}_0$ and the input vector mean $\mathbf{p}_0$. Set the initial values for the respective covariance matrices $\mathbf{P}_0^{\widetilde{\zeta}}$ and $\mathbf{P}_0^p$. Define the noise parameters $\mathbf{Q}_w$, $\mathbf{Q}_v$, and $\mathbf{Q}_r$.
Update input and state vectors based on observations
(1) Calculate the Kalman gain for the input: $\mathbf{K}_k^p = \mathbf{P}_k^{p-}\mathbf{D}_d^T\left(\mathbf{D}_d\mathbf{P}_k^{p-}\mathbf{D}_d^T + \mathbf{Q}_r\right)^{-1}$
(2) Update the input mean and covariance:
$\mathbf{p}_k = \mathbf{p}_k^- + \mathbf{K}_k^p\left(\mathbf{y}_k - \mathbf{C}_d\widetilde{\zeta}_k^- - \mathbf{D}_d\mathbf{p}_k^-\right)$
$\mathbf{P}_k^p = \mathbf{P}_k^{p-} - \mathbf{K}_k^p\mathbf{D}_d\mathbf{P}_k^{p-}$
(3) Calculate the Kalman gain for the state vector: $\mathbf{K}_k^{\widetilde{\zeta}} = \mathbf{P}_k^{\widetilde{\zeta}-}\mathbf{C}_d^T\left(\mathbf{C}_d\mathbf{P}_k^{\widetilde{\zeta}-}\mathbf{C}_d^T + \mathbf{Q}_r\right)^{-1}$
(4) Update the state vector mean and covariance:
$\widetilde{\zeta}_k = \widetilde{\zeta}_k^- + \mathbf{K}_k^{\widetilde{\zeta}}\left(\mathbf{y}_k - \mathbf{C}_d\widetilde{\zeta}_k^- - \mathbf{D}_d\mathbf{p}_k\right)$
$\mathbf{P}_k^{\widetilde{\zeta}} = \mathbf{P}_k^{\widetilde{\zeta}-} + \mathbf{K}_k^{\widetilde{\zeta}}\mathbf{C}_d\mathbf{P}_k^{\widetilde{\zeta}-}$
Predict the input mean and covariance and state vector mean and covariance
(1) Predict the input mean and covariance at step $k+1$:
$\mathbf{p}_{k+1}^- = \mathbf{p}_k$
$\mathbf{P}_{k+1}^{p-} = \mathbf{P}_k^p + \mathbf{Q}_v$
(2) Predict the state vector mean and covariance at step $k+1$:
$\widetilde{\zeta}_{k+1}^- = \mathbf{A}_d\widetilde{\zeta}_k + \mathbf{B}_d\mathbf{p}_k$
$\mathbf{P}_{k+1}^{\widetilde{\zeta}-} = \mathbf{A}_d\mathbf{P}_k^{\widetilde{\zeta}}\mathbf{A}_d^T + \mathbf{Q}_w$

Algorithm 1: The DKF for joint input-stateestimation [14].

stiffness parameter. Therefore, this section demonstrates a sequential Bayesian inference approach that first estimates the input to the system based on a Kalman filter, then filters the obtained input to appropriate wavelengths, and finally estimates the unknown parameter via a UKF for state-parameter estimation. Algorithm 3 demonstrates in detail this sequential Bayesian inference scheme. Specifically, for the first step (Algorithm 3), assuming known the state vector $\mathbf{x}_k^-$ as predicted in the previous time step, the input to the vehicle, i.e., the contact force $\mathbf{p}_k$, is estimated based on equations (21a) and (21b). Then, in the second step (Algorithm 3), the contact force is written as the product of the contact stiffness k and the roughness profile $\mathbf{r}_k(x)$ (equation (2)). Subsequently, the roughness profile is filtered and gives $\mathbf{r}_k^{\text{filt}}$. For the implementation of the UKF in the last step, the joint input vector that includes the contact forces estimated in the first step and the filtered roughness estimated in the second step is written as

$$\overline{\mathbf{u}}_k = \begin{bmatrix} \mathbf{p}_k \\ \mathbf{r}_k^{\text{filt}} \end{bmatrix}. \tag{25}$$

Thus, in the third step (Algorithm 3), a UKF implementation estimates the augmented state-parameter vector $\overline{\mathbf{x}}_k$ (equation (22)) assuming known the joint input vector $\overline{\mathbf{u}}_k$ and considering as output vector the measured accelerations and the filtered roughness profile recorded by the laser scanners. After estimating the unknown stiffness parameter k, assuming known a small filtered portion of the roughness profile of a railway network, the DKF (Algorithm 1) can be applied to identify the roughness profile of the entire railway network.

Initialize

(1) Set the initial values for the augmented state vector mean $\overline{\mathbf{x}}_0$ and for the respective covariance matrix $\mathbf{P}_0^{\overline{x}}$. Define the noise parameters $\mathbf{Q}_z$ and $\mathbf{Q}_r$.

(2) Set the parameters for the UKF, α, β, and κ, and estimate the weights for the mean and covariance of the predicted measurements: W_m^0, W_m^i, W_c^0, and W_c^i for, $i = 1, \ldots, 2N$ where N is the number of states.

Update the augmented state vector based on observations:

(1) Calculate the Sigma points: $\mathbf{X}_k^{0-} = \overline{\mathbf{x}}_k^-, \mathbf{X}_k^{i-} = \overline{\mathbf{x}}_k^- + \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}-}}\right)_i$ for $i = 1, \ldots, N$, $\mathbf{X}_k^{i-} = \overline{\mathbf{x}}_k^- - \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}-}}\right)_i$ for $i = N+1, \ldots, 2N$ where $c = \alpha^2(N + \kappa)$

(2) Propagate the Sigma points through the output equation: $\widehat{\mathbf{y}}_k^- = \mathbf{h}\left(\mathbf{X}_k^-, \mathbf{p}_k\right)$

(3) Calculate the output mean and covariance:

$$\overline{\mathbf{y}}_k = \sum_{i=0}^{2N} W_m^i \widehat{\mathbf{y}}_k^{i-}$$

$$\mathbf{P}_k^y = \sum_{i=0}^{2N} W_c^i \left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)\left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)^T + \mathbf{Q}_r$$

(4) Calculate the cross-covariance between the states and output: $\mathbf{P}_k^{\overline{x}y} = (1/2c)\sum_{i=1}^{2N} W_c^i \left(\mathbf{X}_k^{i-} - \overline{\mathbf{x}}_k^-\right)\left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)^T$

(5) Calculate the Kalman gain for the states: $\mathbf{K}_k^{\overline{x}} = \mathbf{P}_k^{\overline{x}y}\left(\mathbf{P}_k^y\right)^{-1}$

(6) Update the state mean and covariance:

$$\overline{\mathbf{x}}_k = \overline{\mathbf{x}}_k^- + \mathbf{K}_k^{\overline{x}}\left(\mathbf{y}_k - \overline{\mathbf{y}}_k\right)$$

$$\mathbf{P}_k^{\overline{x}} = \mathbf{P}_k^{\overline{x}-} - \mathbf{K}_k^{\overline{x}}\mathbf{P}_k^y\left(\mathbf{K}_k^{\overline{x}}\right)^T$$

Predict the augmented state vector mean and covariance for the next step:

(1) Calculate the Sigma points: $\mathbf{X}_k^0 = \overline{\mathbf{x}}_k, \mathbf{X}_k^i = \overline{\mathbf{x}}_k + \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}}}\right)_i$ for $i = 1, \ldots, N$, $\mathbf{X}_k^i = \overline{\mathbf{x}}_k - \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}}}\right)_i$ for $i = N+1, \ldots, 2N$

(2) Propagate the Sigma points through the state equation: $\widehat{\mathbf{x}}_k = \mathbf{f}\left(\mathbf{X}_k, \mathbf{p}_k\right)$

(3) Predict the state mean and covariance at step $k+1$:

$$\overline{\mathbf{x}}_{k+1}^- = \sum_{i=0}^{2N} W_m^i \widehat{\mathbf{x}}_k^i$$

$$\mathbf{P}_{k+1}^{\overline{x}} = \sum_{i=0}^{2N} W_c^i \left(\widehat{\mathbf{x}}_k^i - \overline{\mathbf{x}}_{k+1}^-\right)\left(\widehat{\mathbf{x}}_k^i - \overline{\mathbf{x}}_{k+1}^-\right)^T + \mathbf{Q}_z$$

Algorithm 2: The UKF for joint state-parameter estimation [22].

5. Applications on Simulated and Field Data

5.1. Application on Simulated Data from a 3D Vehicle Traversing Tracks on Solid Ground. The first numerical application concerns a simple 3D train vehicle model traversing tracks comprising random vertical roughness profiles for both left and right rails along the longitudinal direction. This example aims at demonstrating the efficacy of the DKF in identifying rail roughness profiles on the basis of the sole availability of acceleration data collected from the axle boxes of traversing vehicles. In this case, the vehicle is modeled as a multibody assembly consisting of seven rigid bodies: one car body, two bogies, and four wheelsets connected with linear springs and dashpots [30], as illustrated in Figure 2. Each body has five DOFs: two translational in the vertical z and lateral y directions and three rotational, namely, yawing ψ, rolling ϕ, and pitching θ. Table 1 summarizes the DOFs of the employed train vehicle model. The properties of the vehicle are derived from the study of Zeng and Dimitrakopoulos [31]. The track is modeled as a rough surface, and the assumed roughness profile is considered a stationary stochastic process with the spectral representation method [32]. Different roughness profiles are assumed for the left and right rails.

First, simulated acceleration data are generated via forward modeling of the full vehicle-track system. The vehicle runs on straight rails at 200 km/h, a typical operating speed for high-speed trains. The properties of the profile are defined according to the German spectra for high-speed railways [33]. This case study assumes the following parameters for the involved roughness profile: $A_v = 4.032 \cdot 10^{-7}\,\mathrm{m}^2 \cdot \mathrm{rad/m}$, $\Omega_c = 0.8246$ rad/m, and $\Omega_r = 0.0206$ rad/m [33]. The corresponding wavelength range is 0.5 m to 80 m. Data are collected from the axle boxes (placed on the left and right sides of each wheelset) at a sampling frequency of $f_s = 1000$ Hz. Acceleration measurements are utilized as displacement data are generally more challenging to measure. The acceleration measurements are finally contaminated with 5% Gaussian noise in order to simulate sensor noise contamination.

Before moving to the identification task, the train model is reduced. In this case, the model-order reduction follows the eigenvalue analysis of Section 3.1. The modes retained are decided based on a dominance measure for each mode and each input/output to the system [25]. Eventually, the total number of modes after truncation is 18. The retained modes frequency range is from 1 Hz to 8.4 Hz. Decreasing the number of modes to less than 18 further enhances the computational efficiency at the cost of the estimation accuracy for the roughness profile.

The rail roughness is estimated using a DKF (Section 4.2.1 and Algorithm 1) with a bilinear transform assumed for the discretization of the system and a sampling period $T_s = 0.001$ s. Assuming an accurate train model, the covariance matrix of the process noise is set to $\mathbf{Q}_w = 10^{-8} \cdot \mathbf{I}_1$, where $\mathbf{I}_1 \in \mathbb{R}^{18\times 18}$ is an identity matrix with dimension equal to the number of states of the reduced system. The measurement noise covariance matrix is $\mathbf{Q}_r = 10^{-1} \cdot \mathbf{I}_2$, where $\mathbf{I}_2 \in \mathbb{R}^{8\times 8}$ is an identity matrix with dimension equal to the number of measurements. The estimation of the covariance matrix of the input noise

Initialize
(1) Set the initial values for the state vector mean $\mathbf{x}_0$, the augmented state vector mean $\overline{\mathbf{x}}_0$ (equation (22)), and the input vector mean $\mathbf{p}_0$. Set the initial values for the respective covariance matrices $\mathbf{P}_0^x$, $\mathbf{P}_0^{\overline{x}}$, and $\mathbf{P}_0^p$. Define the noise parameters $\mathbf{Q}_z$, $\mathbf{Q}_v$, and $\mathbf{Q}_r$.
(2) Set the parameters for the UKF, α, β, and κ, and estimate the weights for the mean and covariance of the predicted measurements: W_m^0, W_m^i, W_c^0, and W_c^i for $i = 1, \ldots, 2N$ where N is the number of states.
First step: Update input (Kalman filter)
(1) Calculate the Kalman gain for the input: $\mathbf{K}_k^p = \mathbf{P}_k^{p-}\mathbf{D}^T\left(\mathbf{D}\mathbf{P}_k^{p-}\mathbf{D}^T + \mathbf{Q}_r\right)^{-1}$
(2) Update the input mean and covariance:
$\mathbf{p}_k = \mathbf{p}_k^- + \mathbf{K}_k^p\left(\mathbf{y}_k - \mathbf{C}\mathbf{x}_k^- - \mathbf{D}\mathbf{p}_k^-\right)$
$\mathbf{P}_k^p = \mathbf{P}_k^{p-} - \mathbf{K}_k^p\mathbf{D}\mathbf{P}_k^{p-}$
Second step: Filter estimated input to the wavelengths of Eurocode
(1) Write the estimated input as the product of the unknown parameter k and roughness profile $\mathbf{r}_k$ (equation (2))
(2) Filter the roughness profile to D0–D2: $\mathbf{r}_k^{\text{filt}}$
(3) Formulate the joint input vector $\overline{\mathbf{u}}_k$ assuming known the contact force $\mathbf{p}_k$ and filtered roughness $\mathbf{r}_k^{\text{filt}}$: $\overline{\mathbf{u}}_k = \begin{bmatrix} \mathbf{p}_k \\ \mathbf{r}_k^{\text{filt}} \end{bmatrix}$ (equation (25))
Third step: Update state-parameter vector (UKF)
(1) Calculate the Sigma points: $\mathbf{X}_k^{0-} = \overline{\mathbf{x}}_k^-, \mathbf{X}_k^{i-} = \overline{\mathbf{x}}_k^- + \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}-}}\right)_i$ for $i = 1, \ldots, N, \mathbf{X}_k^{i-} = \overline{\mathbf{x}}_k^- - \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}-}}\right)_i$ for $i = N+1, \ldots, 2N$ where $c = \alpha^2(N + \kappa)$
(2) Propagate the Sigma points through the output equation: $\widehat{\mathbf{y}}_k^- = \mathbf{h}\left(\mathbf{X}_k^-, \overline{\mathbf{u}}_k\right)$
(3) Calculate the output mean and covariance:
$\overline{\mathbf{y}}_k = \sum_{i=0}^{2N} W_m^i \widehat{\mathbf{y}}_k^{i-}$
$\mathbf{P}_k^y = \sum_{i=0}^{2N} W_c^i\left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)\left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)^T + \mathbf{Q}_r$
(4) Calculate the cross-covariance between the states and output: $\mathbf{P}_k^{\overline{x}y} = (1/2c)\sum_{i=1}^{2N} W_c^i\left(\mathbf{X}_k^{i-} - \overline{\mathbf{x}}_k^-\right)\left(\widehat{\mathbf{y}}_k^{i-} - \overline{\mathbf{y}}_k\right)^T$
(5) Calculate the Kalman gain for the states: $\mathbf{K}_k^{\overline{x}} = \mathbf{P}_k^{\overline{x}y}\left(\mathbf{P}_k^y\right)^{-1}$
(6) Update the state mean and covariance:
$\overline{\mathbf{x}}_k = \overline{\mathbf{x}}_k^- + \mathbf{K}_k^{\overline{x}}\left(\mathbf{y}_k - \overline{\mathbf{y}}_k\right)$
$\mathbf{P}_k^{\overline{x}} = \mathbf{P}_k^{\overline{x}-} - \mathbf{K}_k^{\overline{x}}\mathbf{P}_k^y\left(\mathbf{K}_k^{\overline{x}}\right)^T$
Predict the input and state mean and covariance for the next step:
(1) Predict the input mean and covariance at step $k+1$:
$\mathbf{p}_{k+1} = \mathbf{p}_k$
$\mathbf{P}_{k+1}^{p-} = \mathbf{P}_k^p + \mathbf{Q}_v$
(2) Calculate the Sigma points: $\mathbf{X}_k^0 = \overline{\mathbf{x}}_k, \mathbf{X}_k^i = \overline{\mathbf{x}}_k + \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}}}\right)_i$ for $i = 1, \ldots, N, \mathbf{X}_k^i = \overline{\mathbf{x}}_k - \left(\sqrt{c\overline{\mathbf{P}}_k^{\overline{x}}}\right)_i$ for $i = N+1, \ldots, 2N$
(3) Propagate the Sigma points through the state equation: $\widehat{\mathbf{x}}_k = \mathbf{f}\left(\mathbf{X}_k, \overline{\mathbf{u}}_k\right)$
(4) Predict the state mean and covariance at step $k+1$:
$\overline{\mathbf{x}}_{k+1}^- = \sum_{i=0}^{2N} W_m^i \widehat{\mathbf{x}}_k^i$
$\mathbf{P}_{k+1}^{\overline{x}} = \sum_{i=0}^{2N} W_c^i\left(\widehat{\mathbf{x}}_k^i - \overline{\mathbf{x}}_{k+1}^-\right)\left(\widehat{\mathbf{x}}_k^i - \overline{\mathbf{x}}_{k+1}^-\right)^T + \mathbf{Q}_z$

Algorithm 3: A sequential Bayesian inference scheme for contact stiffness parameter optimization.

follows an L-curve analysis [28], where the input noise is considered to be the regularization parameter. This study calibrates the input noise based on the L-curve of the acceleration measured on the left axle box of the first wheelset, as demonstrated in Figure 5. Figure 5 plots the mean squared error (MSE) of the novelty term of the Kalman filter $\sum_{k=1}^{N_s}\|\mathbf{y}_k - \widetilde{\mathbf{C}}_d\widetilde{\zeta}_k^- - \widetilde{\mathbf{D}}_d\mathbf{p}_k\|_2^2/N_s$ against the corresponding values of the covariance of the input noise Q_v. In the innovation term, $\widetilde{\zeta}_k^-$ is the predicted value of the state vector at step k and N_s is the total number of steps. The plot of Figure 5 shows that the value $\mathbf{Q}_v = 10^{-4} \cdot \mathbf{I}_3$, where $\mathbf{I}_3 \in \mathbb{R}^{8\times 8}$ is an identity matrix with dimension equal to the number of inputs of the system, should be chosen as the covariance of the input noise for the DKF, as this is the inflection point of the L-curve minimizing both the MSE for the innovation term and the corresponding Q_v.

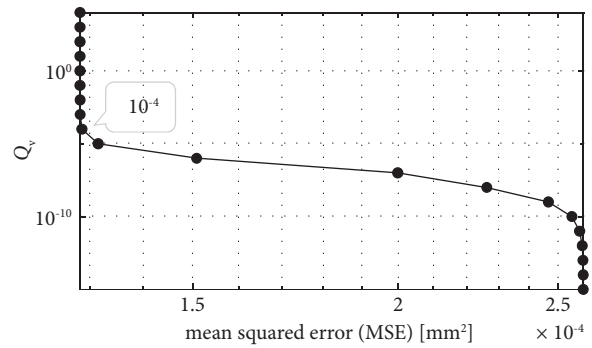

Figure 5: L-curve of the acceleration of the left axle box of the first wheelset: input noise Q_v against the mean squared error of the innovation term of the DKF.

Figure 6 demonstrates the estimated roughness profiles of the left (Figures 6(a) and 6(b)) and right (Figures 6(c) and 6(d)) rails of the track system and their corresponding one-sided power spectral density (PSD) using Welch's averaged, modified periodogram method assuming a Hamming window. The identification task returns a very good

estimate of the reference (used in the forward analysis) roughness profiles, despite the use of a reduced-order model. It should be further mentioned that using the full-order model would not be advisable in this case since the identification task should avoid the use of the same forward and inverse models in order to alleviate an inverse crime [16].

To check the efficacy of the proposed methodology to reconstruct the reference roughness profile, this study adopts a fit value based on the normalized root mean square error (NRMSE) between the reference and estimated roughness profiles. The fit between the profiles is given as

$$\text{fit}_{\text{NRMSE}} = 100\left(1 - \frac{\left\|\mathbf{r} - \mathbf{r}_{\text{est}}\right\|}{\left\|\mathbf{r} - \bar{r}\right\|}\right), \tag{26}$$

where $\mathbf{r}$ is the reference roughness profile, $\bar{r}$ is the mean value of the reference profile, and $\mathbf{r}_{\text{est}}$ is the estimated roughness profile. The NRMSE fits of the time histories for the roughness profiles of the left and right rails are 91.3% and 90.3%, respectively. A cosine similarity criterion is also used to measure the correlation between the estimated and reference roughness profiles. Cosine similarity can measure the shape similarity of data series vectors based on their angles but cannot capture variations in their magnitude. The cosine similarities between the estimated and reference profiles for the left and right wheels are, respectively, 0.99 and 0.98, indicating a very good agreement. When performing the inverse analysis utilizing the full train model, the results are similar to those of the reduced model, so they are omitted from Figure 6 for clarity.

5.2. Application on Simulated Data from an SBB Diagnostic Vehicle Traversing a Measured Roughness Profile. This section employs a realistic railroad vehicle model of the SBB network running on a roughness profile measured by SBB. Both vehicle and track systems are modeled in SIMPACK software [34]. The adopted vehicle is the diagnostic (gDfZ) vehicle of SBB, consisting of one car body, two bogies, and four wheelsets with their respective axle boxes. For the car body, its torsional stiffness around the longitudinal axis is also considered. The primary suspension system consists of coil springs, axle bushing elements, and primary dampers. The secondary suspension system includes springs, vertical, lateral, and yaw dampers, as well as a detailed modeling of the antiroll bar and push-pull rods. The contact between the wheels of the diagnostic train and the track system follows an elastic normal contact model (Hertzian contact) with stiffness determined based on the modulus of elasticity and Poisson's ratio. Figure 7 illustrates the configuration of the diagnostic vehicle. The assumed roughness profile is measured by SBB and corresponds to wavelengths between 0.5 m and 70 m. This vehicle-track model simulated in SIMPACK software is used to produce acceleration measurements of the axle boxes at a sampling frequency $f_s = 1000$ Hz. The acceleration data are then used in the identification task.

The identification task, i.e., the inverse modeling of the system, is realized in MATLAB software [35]. To this end, the first step is to export the state-space model of the diagnostic vehicle and the acceleration measurements from the axle boxes from SIMPACK to MATLAB. The vehicle model, which is simulated in SIMPACK software, includes nonlinear elements, e.g., nonlinear dampers and a nonlinear contact stiffness model. The proposed Bayesian inference scheme assumes linear state-space matrices. To this end, the state-space model is linearized around the initial equilibrium state of the vehicle. This is implemented in SIMPACK by running an online static solver that performs Newton equilibrium, i.e., it brings the vehicle model to a state of zero acceleration [36]. Once static equilibrium is achieved, a linearized state-space model of the vehicle is exported, which is then used for the identification task realized in MATLAB. The total number of states of the exported linearized system is 144, and the number of inputs is eight, corresponding to the eight wheels of the four wheelsets of the vehicle. The number of outputs is four, corresponding to the four accelerometers mounted on the diagnostic train—two on the two axle boxes of the first wheelset and another two on the axle boxes of the last wheelset. To reduce the computational effort of the identification task, we first reduce the model based on the eigenvalue analysis of Section 3.1 and by estimating the dominance measure of each mode according to equation (10). The total number of modes retained is 23, including 15 complex conjugate modes and 8 modes with unit damping (corresponding to damping elements of the model). The frequency content of the retained complex conjugate modes varies between 5 Hz and 189 Hz, which is higher than that of the simple 3D vehicle model of Section 5.1. This is attributed to the increased complexity of the diagnostic vehicle model, which also includes higher frequency components. Thus, the total number of states in the state-space representation of the system is 38, which is less than one-third of the initial number of states (144). Finally, the exported acceleration measurements are contaminated with 5% Gaussian white noise representing the actual measurement noise in real systems. Note that in this case, the acceleration data used as output for the identification task are exported from SIMPACK software using the detailed full-order vehicle model simulated therein. On the other hand, the vehicle model used for the realization of the identification scheme is a linearized, reduced-order vehicle model, thus different from the one used for the production of measured data.

The estimation of the roughness profiles follows the DKF (Section 4.2.1 and Algorithm 1) with a bilinear transformation assumed for the discretization of the state-space system. The covariance matrix of the process noise is set to $\mathbf{Q}_w = 10^{-8} \cdot \mathbf{I}_1$, the measurement noise covariance matrix is $\mathbf{Q}_r = 10^{-1} \cdot \mathbf{I}_2$, and the covariance matrix of the input noise is estimated as $\mathbf{Q}_v = 10^{-8} \cdot \mathbf{I}_3$, based on an L-curve analysis. Figure 8 plots the estimated rail roughness profiles (Figures 8(a) and 8(c)) and the corresponding PSDs (Figures 8(b) and 8(d)) when using the reduced-order model and the full train model. For comparison, Figure 8 also shows the roughness profile used in the forward analysis to generate measurement data, referred to in the following as the *reference roughness profile* (ground truth). For this case study, the roughness profile estimated with the reduced-order vehicle model is slightly closer to the reference

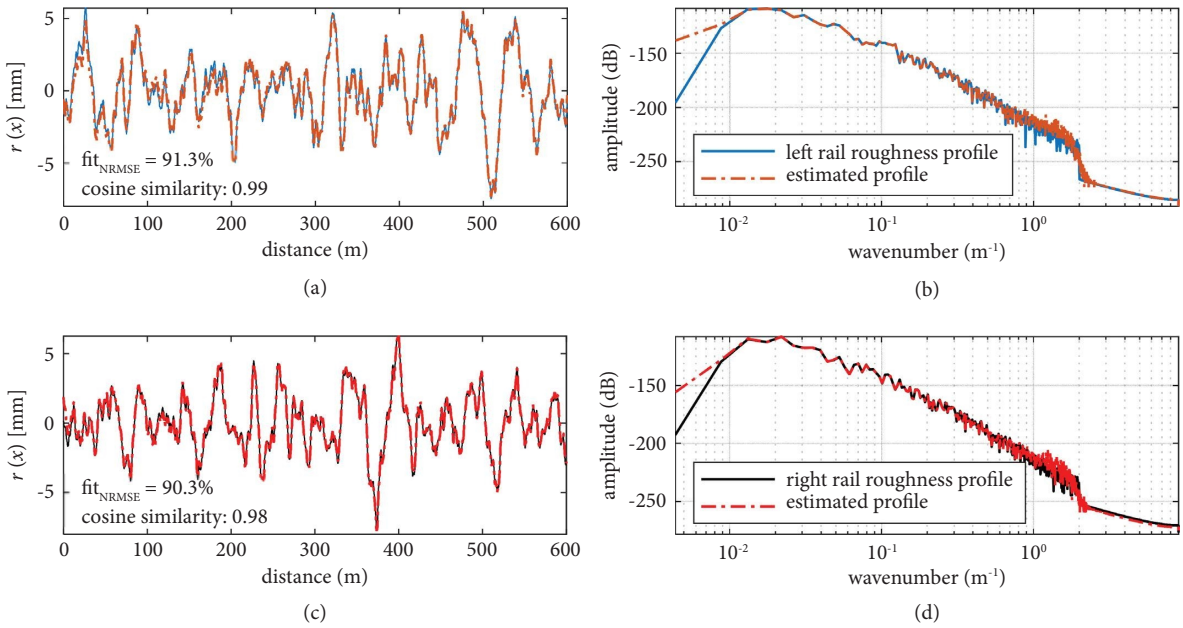

FIGURE 6: Profiles (left) and corresponding one-sided PSDs (right) of the rail roughness used for the forward analysis (blue and black) and of the estimated via DKF (orange and red) rail roughness of the (a, b) left and (c, d) right rails using the reduced via eigenvalue analysis vehicle model.

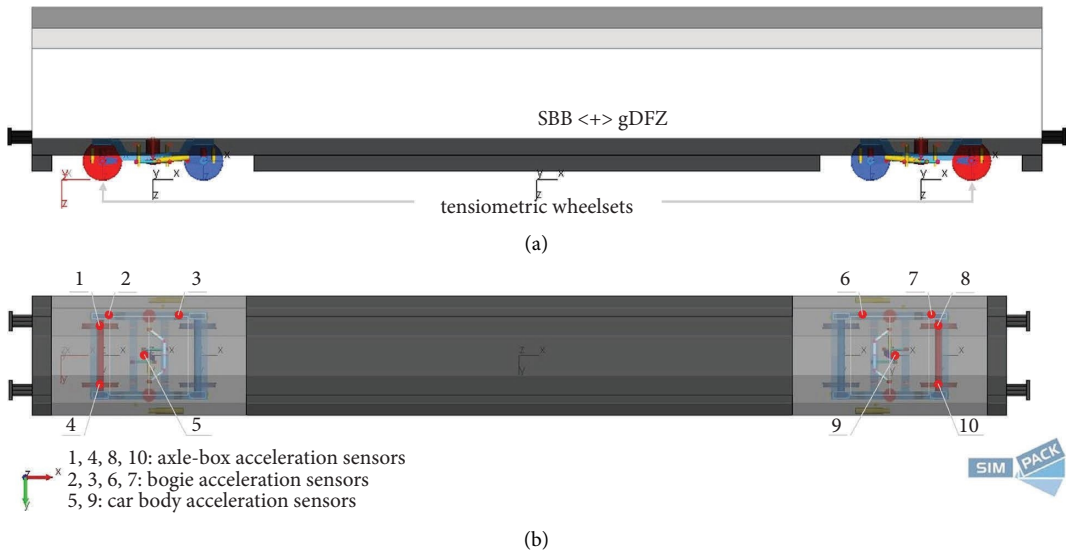

FIGURE 7: SBB diagnostic vehicle (gDfZ) and sensor configuration: (a) side view and (b) plan view (implemented in SIMPACK software) [34].

roughness profile compared to that estimated with the full vehicle model. This is confirmed by the NRMSE fit between the estimated and reference roughness profiles, illustrated in Table 2, which in the case of the eigenvalue-based reduced-order model is 71.6% while for the full model this is 64.7%. The corresponding cosine similarity values are 0.96 and 0.95, respectively. This shows that the shape of the estimated roughness profile is very close to that of the reference roughness profile in both cases. On the other hand, in the PSD of the roughness profile estimated based on the reduced-order model (Figures 8(b) and 8(d)), small peaks appear in the frequency band of 0.5–2 m^{-1} (Figure 8(d)), which do not appear when the full-order vehicle model is employed for the inverse analysis. These peaks denote that the reduced-order model misses some modes that dampen its response in the band of 0.5–2 m^{-1}. Should more modes be considered, the size of the model would considerably increase, close to the size of the full-order vehicle model, increasing the computational effort of the analysis.

To reduce the order of the system, we also employ the substructure-based approach of Section 3.2. In this case, the wheelset-axle box system is separated from the upper part of

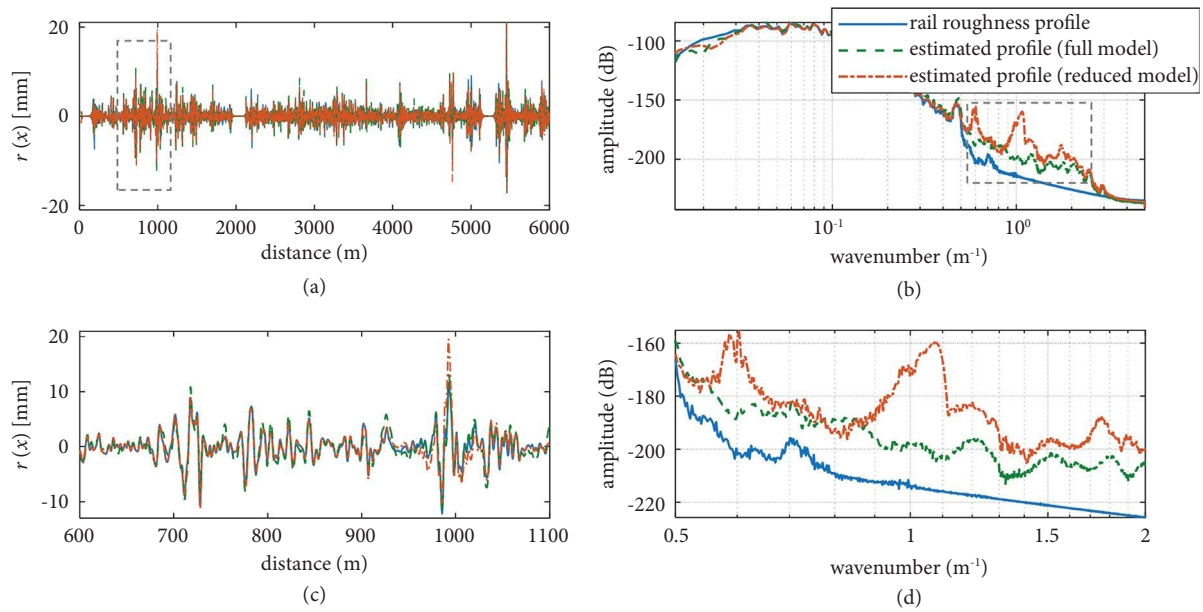

FIGURE 8: (a) Profiles and (b) corresponding one-sided PSDs of the reference rail roughness (blue) and of the rail roughness estimated via DKF with the full-order train model (green) and with the eigenvalue-based reduced-order train model (orange), (c) zoom-in of (a) between 600 m and 1100 m, and (d) zoom-in of (b) between 0.5 m^{-1} and 2 m^{-1}.

TABLE 2: NRMSE fit and cosine similarity between the reference roughness profile and the estimated roughness profiles, and corresponding time cost ratios.

Train model	NRMSE fit (%)	Cosine similarity	Time cost ratio
Full train model	64.7	0.95	1.0
Reduced model: eigenvalue analysis	71.6	0.96	7.6
Reduced model: substructure-based	64.7	0.95	4.0
Reduced model: substructure-based (first wheelset)	64.7	0.95	14.0

the train (car body and bogies), and internal forces arise at the connection with the upper part, as in equation (15). Figure 3 illustrates the retained bodies. The total number of states retained is 64. Figure 9 illustrates the identified profile of the rails based on the substructure-based reduced model of the train. For comparison, it also shows the estimated roughness profile when the full-order model is used. The identified profile and its PSD show a good agreement with the reference roughness profile (Figure 9), while they are exactly the same as those estimated with the full-order vehicle model. This is also confirmed by the NRMSE fit and cosine similarity between the estimated and reference roughness profiles, demonstrated in Table 2, which are the same in the case of the full and reduced-order via substructuring vehicle models. In addition, opposite to the analysis with the aid of the eigenvalue-based reduced model (Figure 8), no peak appears in the frequency band between 0.5 and 2 m^{-1}. To summarize, as illustrated in Figures 8 and 9 and Table 2, for this system and measured acceleration data, the eigenvalue-based reduced model returns a somewhat more accurate roughness profile, while the substructure-based reduced model gives a better estimate of the frequency content of the identified signal. However, note that in all cases (full- and reduced-order vehicle models), a mismatch is present in the PSD of the estimated roughness profile with respect to the reference one for wavenumbers between 0.5 m^{-1} and 3 m^{-1}. This is possibly attributed to the assumed model for the temporal evolution of the unknown input, which is chosen to be a random walk model (Section 4.2.1). Alternative latent force models, such as those employing a Gaussian process assumption [37–40], could be explored to potentially improve the performance of the identification task; however, this lies outside the scope of this study.

As illustrated in the substructure-based model of Figure 3, each wheelset-axle box system appears to be independent of the other. Thus, the adopted model for the identification of rail roughness profiles is further reduced to solely the first wheelset-axle box of the vehicle. According to Table 2, the identification results are exactly the same as those in the case that all four wheelset-axle box systems are employed, but the time cost of the analysis significantly decreases. Table 2 summarizes the NRMSE fit and cosine similarity values between the reference roughness profile and the roughness profiles estimated with the full- and reduced-order train models, as well as the computational cost of each analysis. The NRMSE fit and cosine similarity values of the estimated roughness profiles are comparable in all cases. In addition, the computational effort for the reduced vehicle models is significantly lower than that of the

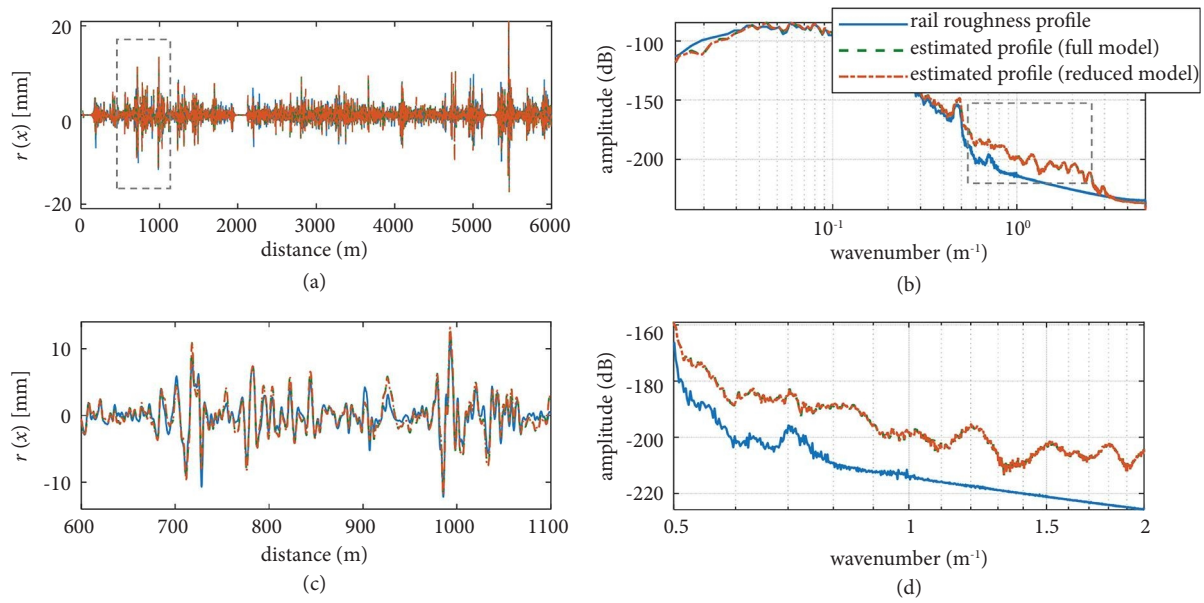

FIGURE 9: (a) Profiles and (b) corresponding one-sided PSDs of the reference rail roughness (blue) and of the rail roughness estimated via DKF with the full-order train model (green) and with the substructure-based reduced-order train model (orange), (c) zoom-in of (a) between 600 m and 1100 m, and (d) zoom-in of (b) between 0.5 m^{-1} and 2 m^{-1}.

full model (a higher number indicates higher efficiency compared to the full model), demonstrating the importance of reduced-order models in inverse problems, especially when online schemes are of interest.

Before moving to the last application of this study, which employs a set of real-world field OBM acceleration data, this section examines the case of uncertain model parameters. This is particularly important for real-world applications like the one demonstrated in Section 5.3, as specific structural parameters of the train, e.g., the contact stiffness between the tracks and the wheels, are usually approximated based on assumptions and engineering judgment. This section focuses on the optimization of the contact stiffness parameter k, as this highly affects the identification of the roughness profile, as demonstrated in equation (2). Specifically, to determine the value of k, it assumes that part of the input, i.e., of the roughness profile, is known and applies a UKF for state-parameter estimation (Section 4.2.2 and Algorithm 2). To demonstrate the parameter updating scheme, an artificial error is introduced to the state-space model used for the identification task. Specifically, a "false" value of the contact stiffness parameter $k_{\text{"false"}}$ is assumed, which is double the original contact stiffness parameter k of the train model. The parameters used for the application of the UKF are $\alpha = 1$, $\beta = 2$, and $\kappa = 0$. As shown in Figure 10, applying the UKF, the ratio of the estimated stiffness parameter k_{UKF} to the assumed "false" stiffness parameter $k_{\text{"false"}}$ (black line) quickly drops and converges to 0.5, i.e., to the ratio of the original stiffness parameter k to the "false" stiffness parameter $k_{\text{"false"}}$ (red line). Thus, the UKF can adequately estimate the actual stiffness parameter of the model, considering that part of the input is known. The ratio between k_{UKF} and $k_{\text{"false"}}$ fluctuates around 0.5 as the estimated parameter k_{UKF} is considered a random variable that is augmented with the state vector for the application of the UKF for the task of joint state-parameter estimation. Certain peaks that occur at around 800 m (Figure 10) are attributed to localized peaks at the measured output, which cause local instabilities during the identification task. Note that this application makes use of simulated acceleration data collected by a forward analysis of the vehicle-track system in SIMPACK software; thus, numerical peaks may appear in the collected output data.

5.3. Application on Field OBM Data from an SBB Diagnostic Vehicle Traversing the SBB Network. The previous applications on simulated data successfully proved the efficacy of the proposed methodology in identifying rail roughness profiles. However, such applications are still biased as the simulated data are generated from models that are then used (in a reduced form) in the identification task. At the same time, external noise or other unpredicted disturbances are not taken into account. Therefore, the proposed methodology is here also applied to real-world OBM data collected by the gDfZ diagnostic vehicle of SBB. The diagnostic vehicle is equipped with three-axial accelerometers on the axle boxes of the first and fourth wheelsets (four in total)—the first and fourth wheelsets are tensiometric, i.e., they can measure contact forces—one-axis accelerometers on the bogie frames (four in total), two-axial accelerometers on the vehicle's car body (two, one in the front and one in the back), two sensors to measure moisture in rails, laser scanners to scan rail roughness profiles, and analog to digital converters. The accelerometers measure data at a sampling frequency $f_s = 24000$ Hz, and the laser scanners have a spatial measurement frequency of 0.25 m^{-1}. Figure 7 illustrates the location of sensors on the vehicle. This section uses the vertical acceleration data recorded by the sensors installed

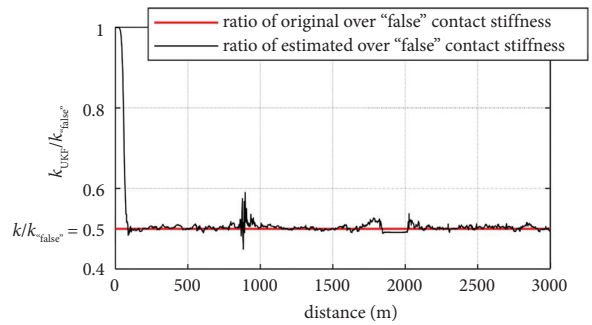

FIGURE 10: Ratio between the estimated via UKF contact stiffness parameter k_{UKF} and the initial "false" contact stiffness parameter $k_{"false"}$ of the vehicle model (black line) versus the ratio between the original stiffness parameter k and $k_{"false"}$ (red line).

on the axle boxes of the diagnostic vehicle and the rail roughness profiles measured by the laser scanners. The measured roughness profiles are considered the ground truth to examine the accuracy of the identified roughness profile via acceleration data. The diagnostic vehicle traverses more than 1000 km of the SBB network at frequent intervals, but a smaller part is considered herein for better visualization of the obtained results.

The identification scheme is implemented in MATLAB software with the aid of the reduced state-space models employed in Section 5.2. Specifically, the eigenvalue-based reduced vehicle model and the substructure-based reduced model that only consider the vehicle's first wheelset are used for this application. On the other hand, the OBM data are filtered and down-sampled to avoid noisy results, low-frequency drifts, and aliasing. The band-pass filter applied has a low cut-off frequency of 0.2 Hz and an upper cut-off frequency of 40 Hz. The data are downsampled to 1000 Hz.

Prior to estimating the rail roughness profile, the contact stiffness parameter needs to be calibrated. This is achieved by assuming a known portion of the roughness profile as measured by the laser scanners. However, the roughness profile measured by the scanners needs to be filtered to the wavelengths specified by Eurocode, i.e., D0: 0.5 m to 3 m, D1: 3 m to 25 m, and D2: 25 m to 70 m. Therefore, in this case, a filtered portion of the input rather than the input itself is known; thus, the sequential Bayesian inference scheme presented in Section 4.2.3 and Algorithm 3 is implemented to estimate an optimized contact stiffness parameter k_{UKF}.

Figure 11(a) demonstrates the ratio between the optimized contact stiffness parameter k_{UKF} and the initial contact stiffness parameter of the vehicle model $k_{initial}$. For the UKF, the filter parameters are $\alpha = 1$, $\beta = 2$, and $\kappa = 0$. This ratio appears to fluctuate around 1.2. Thus, to estimate the actual contact stiffness parameter to be used in the identification of the rail roughness profile, the contact stiffness of the initial vehicle model, directly exported from SIMPACK software, is multiplied by 1.2. Figure 11(b) shows the rail roughness profile filtered to the wavelengths of Eurocode estimated according to the OBM acceleration data collected from the axle box sensors of the diagnostic vehicle (gDfZ) against the filtered roughness profile recorded by the laser scanners. Note that Figure 11 assumes the substructure-based reduced model of the vehicle that only considers the first wheelset. Moreover, for clarity, only a portion of 8.0 km is demonstrated. The rail roughness profile estimated via OBM acceleration data is in very good agreement with the roughness profile measured by the laser scanners. Figure 11(d), which shows a zoom-in of Figure 11(b) between 7000 m and 7600 m, further attests to the good agreement between the two profiles. Figure 11(c) illustrates the PSDs of the roughness profiles and shows that the PSD of the estimated roughness profile is very close to that of the roughness profile recorded via laser scanners. Some discrepancies that appear at very low wavelengths can be attributed to the inevitable shortcomings of modeling approximations. As a comparison, Figure 11 also shows the obtained filtered roughness profile when integrating twice the OBM acceleration data. This is a common approach to obtain the roughness profiles of rails [9, 10] that is characterized by simplicity and ease of implementation. Although double integration returns acceptable results, the profile estimated via DKF after model updating shows a better agreement with the roughness profile measured by the laser scanners (considered here the ground truth). Specifically, the NRMSE fit of the roughness profiles estimated with the proposed UKF-DKF and with double integration with respect to the reference roughness profile measured by the laser scanners are, respectively, 66.4% and 56.5%. Note that the NRMSE fit of the roughness profile estimated via UKF-DKF before updating the contact stiffness parameter is 64.1%. The better performance of the proposed methodology is attributed to its model-based character that considers the dynamics of the adopted vehicle in the identification task. In addition, the employment of the UKF tunes uncertain parameters of the vehicle that could affect the identification task, such as the contact stiffness parameter between the rails and the wheels, thus increasing the accuracy of the identified input. On the other hand, the cosine similarity of the estimated via UKF-DKF roughness profile with respect to the reference roughness profile is 0.94, against 0.93 when double integration is used. Both values show that the shape of the estimated profile is very close to the reference one. However, cosine similarity does not capture variations in the amplitude of the estimated profiles.

Lastly, Figure 12 demonstrates the estimated roughness profile and the corresponding PSD employing the eigenvalue-based reduced-order model of the vehicle. The estimated roughness profile shows an NRMSE fit of 47.9% and a cosine similarity of 0.85 with respect to that recorded from the laser scanners. As expected, according to the application of Section 5.2, the corresponding PSD shows some peaks in the frequency band between $0.5\,m^{-1}$ and $2\,m^{-1}$. The substructure-based reduced order vehicle model appears to capture better the dynamics of the diagnostic vehicle, thus achieving a higher match with respect to the recorded roughness profile. This is because this MOR method retains the physics of the full-order model.

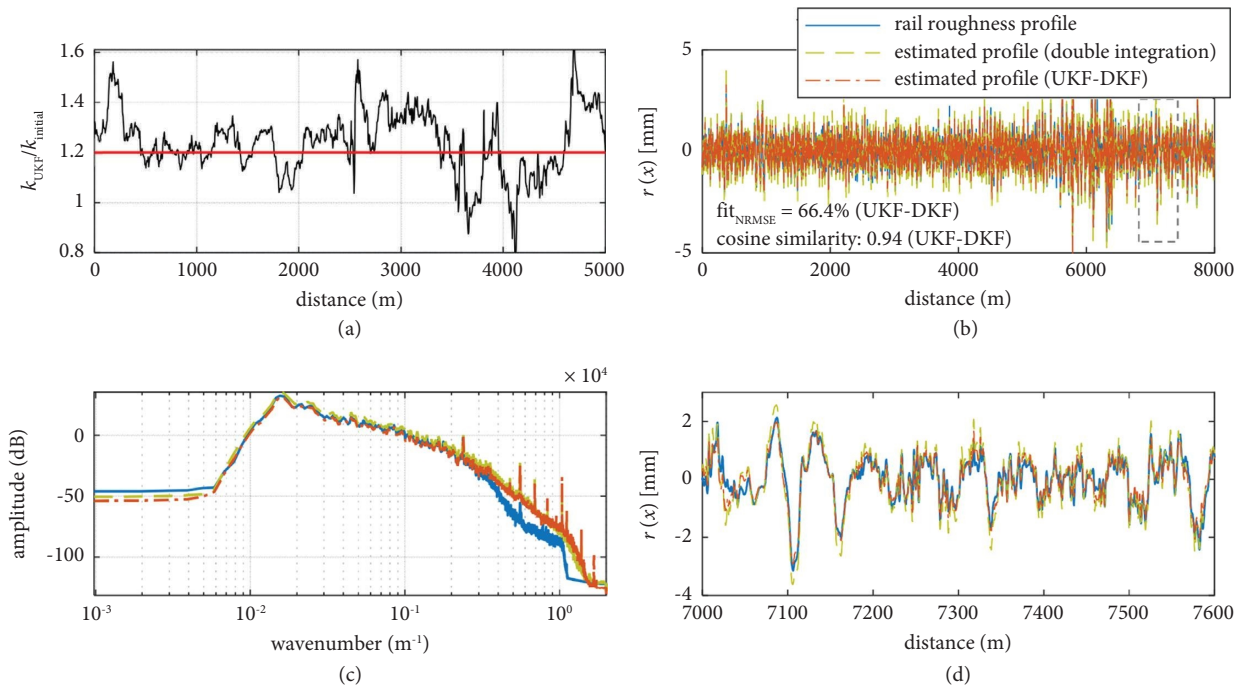

FIGURE 11: (a) Ratio between the estimated k_{UKF} and initial $k_{initial}$ contact stiffness parameters, (b) profiles and (c) corresponding one-sided PSDs of the measured via laser scanners (reference) (blue) and of the estimated via double integration of the OBM measurements (olive green) rail roughness and of the rail roughness estimated via UKF-DKF assuming a substructure-based reduced-order vehicle model (orange), and (d) zoom-in of (b) between 7000 m and 7600 m. All profiles and PSDs correspond to wavelengths between 0.5 m and 70 m (D0–D2 of Eurocode).

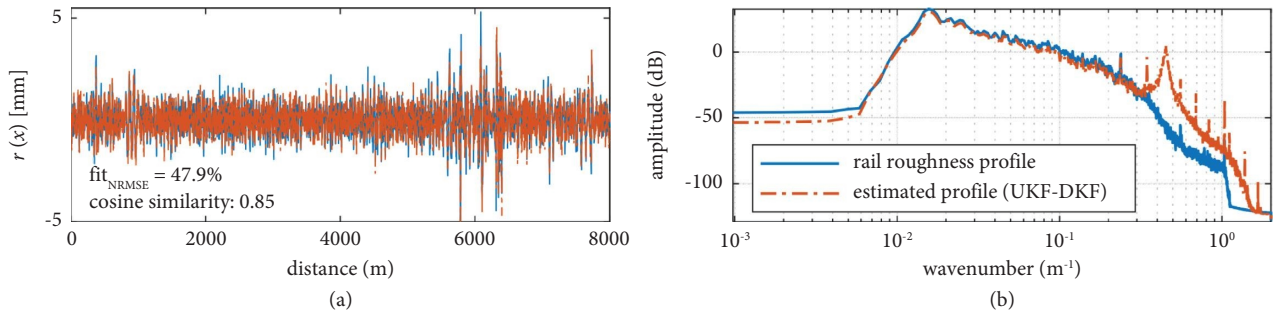

FIGURE 12: (a) Profiles and (b) corresponding one-sided PSDs of the measured via laser scanners (reference) (blue) rail roughness and of the rail roughness estimated via UKF-DKF assuming an eigenvalue-based reduced-order vehicle model (orange). All profiles and PSDs correspond to wavelengths between 0.5 m and 70 m (D0–D2 of Eurocode).

6. Conclusions

This study proposes an indirect approach to estimate rail roughness profiles based on OBM data extracted from traversing trains. A distinct characteristic of the proposed methodology is that it accounts for the dynamic interaction between trains and tracks, identifying roughness profiles via a model-based Bayesian inference method. To this end, first, the adopted vehicle model is reduced in order to avoid excessive computational effort during the identification task. The reduction of the model is here attempted via two alternate schemes; an eigenvalue analysis approach that decomposes the vehicle EOM into independent modal coordinates and retains a set of modes that can efficiently describe the dynamics of the system, and a substructure-based reduction method that partitions the vehicle system into independent, non-overlapping structures. Following, to increase the accuracy of the employed reduced vehicle model and, thus, the accuracy of the identification task, this study updates uncertain parameters of the model via a nonlinear Bayesian filtering algorithm, namely, a UKF, which performs joint state and parameter identification. This is implemented on a portion of the network where the input (roughness) has been measured via a diagnostic vehicle. Finally, using the updated reduced vehicle model, the identification of the roughness profile, i.e., the input to this algorithmic setup, is achieved via a dual implementation of the Kalman filter that first estimates the input and then, assuming known input,

estimates the system's state vector. Note that the roughness identification task is implemented on portions of the network where OBM acceleration data are available.

To assess the efficacy of the proposed methodology, this paper examines three case studies. The first assumes simulated data generated by a simple 3D vehicle running on roughness profiles, which are assumed to be different for the left and right rails. For the inverse problem, i.e., the rail roughness identification task, the reduction of the vehicle is performed based on an eigenvalue analysis. The identified roughness profiles of both rails agree very well with the roughness profiles used for the forward analysis. In addition, the identification results assuming a reduced vehicle model match very well those obtained when the full vehicle model is considered. The second case study examines a real diagnostic vehicle of SBB that runs on tracks of the SBB network. This application still employs simulated data extracted from a run of the vehicle on a measured track profile of the SBB network with the aid of SIMPACK software. The reduction of the vehicle model follows both eigenvalue analysis and substructuring. The reduced-order models return results close to those of the full-order vehicle model, achieving, at the same time, higher computational efficiency and paving the way for online applications. Lastly, the third case study uses actual OBM data measured from the axle boxes of the gDfZ diagnostic vehicle of SBB when that runs across the SBB network. In this case, the identification of rail roughness follows three approaches: estimation of the roughness profile based on acceleration data from the axle boxes via a UKF for parameter update and then a DKF for state-input estimation, double integration of the collected acceleration data, and direct measurement of the roughness profiles of the tracks via laser scanners. Assuming that direct measurements via laser scanners constitute the ground truth, the proposed UKF-DKF approach returns more accurate results compared to double integration when a substructure-based reduced model of the diagnostic vehicle is adopted. On the other hand, when using an eigenvalue-based reduced model of the vehicle, the agreement between the estimated and reference (ground truth) roughness profiles decreases. Thus, for realistic, multibody train systems, the substructure-based reduction method may return a better representation of the dynamics of the full train model.

The proposed methodology has so far been applied to data from a diagnostic vehicle of SBB, whose parameters are carefully measured and then calibrated. The ultimate goal is the development of a comprehensive methodology that is also applicable to in-service trains, which come with more uncertainties compared to dedicated diagnostic vehicles. Employing in-service trains that frequently traverse the railway network allows for continuous monitoring of the state of tracks and, thus, timely identification of damage or progression of faults. Towards this direction, a distinct advantage of the proposed scheme lies in the consideration of the physics of the underlying problem, i.e., the adoption of a physical model. Considering the physics of the problem increases the identification accuracy while enabling the utilization of measurements from other parts of the vehicle. In addition, it allows the fusion of data from multiple sources or the fusion of different types of data, i.e., accelerations and strains. Another key feature is the adoption of a UKF for state-parameter estimation that allows for updating of uncertain parameters that can possibly affect the identification task, i.e., the contact stiffness between the tracks and the wheels or the mass of the vehicle in the case of in-service trains. The proposed approach is herein applied to the identification of longitudinal roughness profiles in the vertical direction but can also be extended to the lateral direction.

Data Availability

The data used to support this study may be available from the corresponding author upon request and only after permission from the Swiss Federal Railways (SBB).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Stavros Niarchos Foundation through the ETH Zürich Foundation and the ETH Zürich Postdoctoral Fellowship Scheme. This work was also supported by the Swiss Federal Railways (SBB) as part of the ETH Mobility Initiative program, under project On board Monitoring for Integrated Systems Understanding and Management Improvement in Railways (OMISM).

References

[1] Railway-News, "Measurement and inspection services-accuracy at first sight," 2019, https://railway-news.com/measurement-and-inspection-services-goldschmidt-group/.

[2] Forum 3 Rails, "Voiture diagnostic infrastructure CFF," 2018, https://forum.3rails.fr/t/voiture-diagnostic-infrastructure-cff/12905.

[3] SBB CFF FFS, "Our Trains: ICN," 2023, https://www.sbb.ch/en/travel-information/services-on-train/our-trains/icn.html.

[4] C. Hoelzl, V. Dertimanis, M. Landgraf, L. Ancu, M. Zurkirchen, and E. Chatzi, "On-board monitoring for smart assessment of railway infrastructure: a systematic review," in *Advanced Structural Sensing and Monitoring Systems*, pp. 223–259, Elsevier, Amsterdam, Netherlands, 2022.

[5] E. Berggren, A. Nissen, and B. Paulsson, "Track deflection and stiffness measurements from a track recording car," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 228, no. 6, pp. 570–580, 2014.

[6] SBB CFF FFS, "Fahrwegmanagement mit swissTAMP," 2016, https://bahninfrastruktur.sbb.ch/de/produkte-dienstleistungen/bahninformatiksysteme/anlagenmanagement/swisstamp.html.

[7] C. Hoelzl, G. Arcieri, L. Ancu et al., "Fusing expert knowledge with monitoring data for condition assessment of railway welds," *Sensors*, vol. 23, no. 5, p. 2672, 2023.

[8] Y. Yang, Z. Wang, K. Shi, H. Xu, and Y. Wu, "State-of-the-art of vehicle-based methods for detecting various properties of highway bridges and railway tracks," *International Journal of*

Structural Stability and Dynamics, vol. 20, no. 13, Article ID 2041004, 2020.

[9] Y. Ren, J. Keenahan, and E. O'Brien, "A two-stage direct integration approach to find the railway track profile using in-service trains," in *Proceedings of the Civil Engineering Research in Ireland 2020*, Cork, Ireland, November 2020.

[10] X. Wei, F. Liu, and L. Jia, "Urban rail track condition monitoring based on in-service vehicle acceleration measurements," *Measurement*, vol. 80, pp. 217–228, 2016.

[11] J. Lee, S. Choi, S. Kim, C. Park, and Y. Kim, "A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 3, pp. 749–758, 2012.

[12] E. O'Brien, C. Bowe, P. Quirke, and D. Cantero, "Determination of longitudinal profile of railway track using vehicle-based inertial readings," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 231, no. 5, pp. 518–534, 2017.

[13] V. Dertimanis, M. Zimmermann, F. Corman, and E. Chatzi, "On-board monitoring of rail roughness via axle box accelerations of revenue trains with uncertain dynamics," in *Model Validation and Uncertainty Quantification*, pp. 167–171, Springer, Berlin, Germany, 2020.

[14] S. Eftekhar Azam, E. Chatzi, and C. Papadimitriou, "A dual Kalman filter approach for state estimation via output-only acceleration measurements," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 866–886, 2015.

[15] C. Li, J. He, X. Wu, and B. Ke, "Research on inverse method of vertical track irregularities based on Adaptive Kalman Filter," 2022.

[16] E. Shimron, J. Tamir, K. Wang, and M. Lustig, "Implicit data crimes: machine learning bias arising from misuse of public data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 119, no. 13, 2022.

[17] S. Muñoz, P. Urda, and J. Escalona, "Experimental measurement of track irregularities using a scaled track recording vehicle and Kalman filtering techniques," *Mechanical Systems and Signal Processing*, vol. 169, Article ID 108625, 2022.

[18] P. Weston, C. Roberts, G. Yeo, and E. Stewart, "Perspectives on railway track geometry condition monitoring from in-service railway vehicles," *Vehicle System Dynamics*, vol. 53, no. 7, pp. 1063–1091, 2015.

[19] C. Stoura and E. Dimitrakopoulos, "Additional damping effect on bridges because of vehicle-bridge interaction," *Journal of Sound and Vibration*, vol. 476, Article ID 115294, 2020.

[20] C. Stoura and E. Dimitrakopoulos, "A Modified Bridge System method to characterize and decouple vehicle-bridge interaction," *Mechanical Journal*, vol. 231, no. 9, pp. 3825–3845, 2020.

[21] K. Johnson, *Contact Mechanics*, Cambridge University Press, Cambridge, UK, 1987.

[22] J. Julier and J. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Proceedings of the Signal processing, sensor fusion, and target recognition VI*, pp. 182–193, Orlando, FL, USA, July 1997.

[23] V. Dertimanis, E. Chatzi, S. Eftekhar Azam, and C. Papadimitriou, "Input-state-parameter estimation of structural systems from limited output information," *Mechanical Systems and Signal Processing*, vol. 126, pp. 711–746, 2019.

[24] J. Cara, "Computing the modal mass from the state space model in combined experimental-operational modal analysis," *Journal of Sound and Vibration*, vol. 370, pp. 94–110, 2016.

[25] O. Föllinger, U. Konigorski, B. Lohmann, G. Roppenecker, and A. Trächtler, *Regelungstechnik*, VDE-Verlag, Berlin, Germany, 11th edition, 2013.

[26] K. Tatsis, V. Dertimanis, C. Papadimitriou, E. Lourens, and E. Chatzi, "A general substructure-based framework for input-state estimation using limited output measurements," *Mechanical Systems and Signal Processing*, vol. 150, Article ID 107223, 2021.

[27] G. Franklin, J. Powell, and M. Workman, *Digital Control of Dynamic Systems*, Addison-Wesley, Menlo Park, CA, USA, 1998.

[28] P. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Review*, vol. 34, no. 4, pp. 561–580, 1992.

[29] K. Tatsis, V. Dertimanis, and E. Chatzi, "Sequential bayesian inference for uncertain nonlinear dynamic systems: a tutorial," *Journal of Structural Dynamics*, vol. 1, pp. 236–262, 2022.

[30] Q. Zeng, C. Stoura, and E. Dimitrakopoulos, "A localized Lagrange multipliers approach for the problem of vehicle-bridge-interaction," *Engineering Structures*, vol. 168, pp. 82–92, 2018.

[31] Q. Zeng and E. Dimitrakopoulos, "Vehicle-bridge interaction analysis modeling derailment during earthquakes," *Nonlinear Dynamics*, vol. 93, no. 4, pp. 2315–2337, 2018.

[32] Y. Yang, J. Yau, and Y. Wu, *Vehicle-Bridge Interaction Dynamics: With Applications to High-Speed Railways*, World Scientific, Singapore, 2004.

[33] W. Guo, H. Xia, G. De Roeck, and K. Liu, "Integral model for train-track-bridge interaction on the Sesia viaduct: dynamic simulation and critical assessment," *Computers & Structures*, vol. 112-113, pp. 205–216, 2012.

[34] Dassault Systémes, "Simpack," 2021, https://www.3ds.com/products-services/simulia/products/simpack/.

[35] MathWorks, "Matlab," 2021, https://ch.mathworks.com/de/products/matlab.html.

[36] Dassault Systémes, "SIMULIA user assistance 2021," 2021, https://www.3ds.com/products-services/simulia/products/simpack/.

[37] S. Vettori, E. Di Lorenzo, B. Peeters, and E. Chatzi, "Assessment of alternative covariance functions for joint input-state estimation via Gaussian Process latent force models in structural dynamics," 2023, https://arxiv.org/abs/2306.16302.

[38] J. Zou, E. Lourens, and A. Cicirello, "Virtual sensing of subsoil strain response in monopile-based offshore wind turbines via Gaussian process latent force models," *Mechanical Systems and Signal Processing*, vol. 200, Article ID 110488, 2023.

[39] T. Rogers, K. Worden, and E. Cross, "On the application of Gaussian process latent force models for joint input-state-parameter estimation: with a view to Bayesian operational identification," *Mechanical Systems and Signal Processing*, vol. 140, Article ID 106580, 2020.

[40] R. Nayek, S. Chakraborty, and S. Narasimhan, "A Gaussian process latent force model for joint input-state estimation in linear structural systems," *Mechanical Systems and Signal Processing*, vol. 128, pp. 497–530, 2019.