Map-supported Positioning Enables In-service Condition Monitoring of Railway Tracks

Michael Roth, Benjamin Baasch, Patrik Havrila, Jörn Groos

Institute of Transportation Systems,
German Aerospace Center (DLR)
Braunschweig, Germany
Email: m.roth@dlr.de

Abstract—We demonstrate how positioning concepts enable in-service condition monitoring of railway tracks. Specifically, it is shown that accurate georeferencing of monitoring data can be achieved by sensor fusion of GNSS and IMU measurements with a map of the railway network. Because such georeferencing is an offline positioning problem, a two-stage approach that operates on batches of data is developed: First, path hypotheses are estimated from the GNSS data and the railway map. Second, a nonlinear Rauch-Tung-Striebel smoother provides on-track positions and speeds in path coordinates given each path hypothesis and the IMU and GNSS data.

The developed methods are an essential part of a track condition monitoring system developed at DLR. The positioning results are used for the track-dependent analysis of axle-box-acceleration data. Accordingly, all results shown in this paper have been obtained on real data collected in the harbor railway network of Braunschweig, Germany.

I. INTRODUCTION

Maintenance costs for railway infrastructure operators are high. The DLR Institute of Transportation Systems works on infrastructure monitoring systems that can reduce these costs by facilitating condition-based maintenance [1], [2]. One application example is the monitoring of railway tracks via the analysis of axle-box accelerations (ABA) recorded by sensors on in-service rail vehicles (specifically: switchers (US), shunters (UK)) [3]. In order to associate anomalies in the ABA data with defects in the track, position information is required. The task of generating accurate position stamps with a high spatial resolution is here referred to as georeferencing and the main subject of this paper.

Georeferencing comes with several challenges. Railway regulations permit only little interaction with in-service vehicles, such that the applicable sensors are limited. For example, access to odometry signals is difficult on most rail vehicles. Furthermore, sensors systems for retrofitting in-service vehicles are restricted to low-cost hardware to be economically viable. The comparably expensive ABA sensors leave little budget for positioning sensors. Therefore, only a minimum set-up consisting of a global navigation satellite system (GNSS) receiver and an inertial measurement unit (IMU) is used here. Low-cost GNSS receivers provide sampling rates of $1 \ldots 5\,\mathrm{Hz}$, which is too low for accurate georeferencing. At a speed of $100\,\mathrm{km/h}$, for instance, the distance traveled between two sampling instances is $28\,\mathrm{m}$ at $1\,\mathrm{Hz}$. A typical IMU sampling rate of $100\,\mathrm{Hz}$ translates to a spatial resolution of $0.28\,\mathrm{m}$

instead. So, sensor fusion of the GNSS and IMU data is required to obtain position stamps with a sufficiently high sampling rate for the position-dependent analysis of ABA data. Low-cost IMUs often exhibit large calibration errors that need to be taken into account, in addition to the permanent vibrations due to the large Diesel engine. Typical areas of operation, i.e., marshalling yards, industrial and harbor railways, often challenge the GNSS reception. Satellite visibility is compromised next to, e.g., buildings, cranes, container stacks, and in underpasses. Restrictions on the antenna placement challenge the GNSS reception further.

However, there are also simplifying aspects compared to other positioning problems. First, the analysis of monitoring data is often carried out after all data has been collected. Hence, the georeferencing is an offline problem and the fusion of map, GNSS, and IMU data can be performed by smoothing rather than filtering algorithms [4], [5]. Moreover, this allows for data pre-processing to correct time stamps and systematic sensor errors, and to sort out irrelevant data. Second, rail vehicles are bound to move on tracks that are connected only at switches. Track-constrained motion models can be used to exploit this fact in a sensor fusion algorithm. Here, map information is required, which can be obtained from, e.g., construction drawings or aerial photographs. The digital representation of such a map allows for different tests to, e.g., see if the tracks are properly connected.

The relevant literature related to this paper can be split into condition monitoring and positioning papers. Examples of the former include [6] and [7], that both describe advanced acceleration data analysis but merely mention GPS for geo-referencing. The DLR works [1]–[3] describe georeferencing challenges and solutions in more detail, and are here complemented with a description of the currently implemented georeferencing system. The rail vehicle positioning literature is focused on online applications with a range of on-board sensors. An early example of map-supported positioning using turn rate and speed measurements is [8]. Positioning with an eddy current sensor and a map is described in [9]. Magnetometer measurements are explored in [10]. Most recent approaches rely on GNSS data in some form, as seen in the survey [11]. Examples include [12], which claims high accuracy through tightly coupled GNSS/IMU integration [13] with maps; [14], which discusses the potential of map-
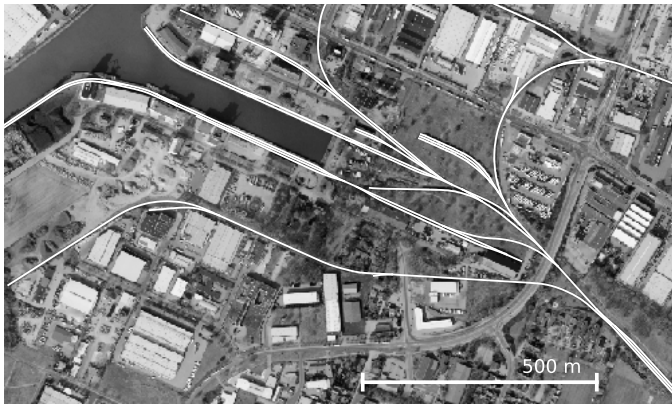
Fig. 1. Braunschweig harbor area with a part of the railway network. Aerial photograph: Stadt Braunschweig (Abteilung Geoinformation).

based multi-sensor positioning for safety critical applications; [15], which employs Rao-Blackwellized particle filters [4] for loosely coupled GNSS/IMU integration; and the DLR contributions [16], [17] with approaches to handle multiple track hypotheses. Aspects of path-constrained positioning, also beyond the rail, are discussed in [18] and [19].

In contrast to the aforementioned positioning resources, this paper focuses on an offline positioning algorithm that combines map information with GNSS and IMU data. Our system employs batches of data to first determine likely hypotheses of the driven path in a network. Based on these, a Rauch-Tung-Striebel smoother [4], [5] with a path-constrained motion model estimates on-track positions and speeds. With the georeferencing application and the limited sensor setup in mind, a simple and pragmatic approach with a solid theoretical sensor fusion foundation is developed. The development has been carried out on large data sets collected on in-service vehicles in the Braunschweig harbor railway, which provides direct proof of the developed concepts.

The outline of the paper is as follows. Sec. II introduces our multi-sensor system, the testbed, and the collected data. Sec. III describes the developed algorithms. Experimental results on real data are shown in Sec. IV. Concluding remarks are given in Sec. V.

II. The Test Site, Sensors, and Data

DLR collaborates with the Braunschweig harbor[1] as rail infrastructure operator and prospective user of the developed condition monitoring system. The harbor railway network comprises more than 15 km of tracks with a connection to the national rail network and is used as test site for the developments of this paper. A segment of the network is shown in Fig. 1.

In the harbor, two switchers arrange the transfer of goods between water and rail. Both are equipped with modular multi-sensor systems developed at DLR. The systems collect monitoring data from ABA sensors and positioning data from

a GNSS receiver and an IMU. Due to the modular system architecture, further sensors could be included in principle. The minimal GNSS and IMU set-up is here selected for its low cost and simple deployment without interfaces to safety critical parts of the rail vehicles. The data management is performed by the Robot Operating System[2] on a Linux platform.

The ABA data are provided by an analog acceleration sensor connected to a high-end analog-to-digital converter. For experimental reasons, a rate of $20\,\mathrm{kHz}$ is used for recording, which motivates for georeferences at high sampling rates. The employed GNSS receiver is a u-Blox NEO-M8N[3] with a roof-mounted antenna. Its sampling rate is set to a desired value $5\,\mathrm{Hz}$ but varies often. Besides horizontal position and speed measurements, a large number of parameters are provided by the receiver. Also accuracy information is included, though without concise specifications. Similar GNSS receivers could also provide raw data (pseudo-range, phase, and Doppler for each satellite in view [13]). For simplicity reasons, we choose to employ the internally calculated position and speed data though. The employed IMU is an Xsens MTi[4] that provides 3D accelerations, turn rates, and magnetic field strengths. Currently, only the acceleration data are used, although the use of turn rates can be viable [8], [16], [17]. The IMU frequency of $100\,\mathrm{Hz}$ determines the sampling rate of the georeferencing system. Newer versions of the DLR multi-sensor system (used in follow-up projects) are equipped with less expensive IMU hardware.

In addition to the sensor data, a highly accurate digital map of the harbor railway is used. Within the Application Platform for Intelligent Mobility (AIM) research project [20], the map has been produced from measurement data and accurate aerial photographs provided by the city of Braunschweig. The map comprises labeled line segments for each railway track, with new track identifiers after every switch. Track connections are encoded via common track start or end points.

All the presented work has been implemented in Python, so as to make the tools easily distributable without the need to purchase extra licenses. Hence, established Python modules for handling large data sets or graph theoretic methods could be used.

III. The Proposed Positioning Approach

The following paragraphs describe the implemented positioning system. With the georeferencing application in mind, we suggest a two-stage offline approach that separates finding the driven path from a smoothing problem in path coordinates. While having a solid theoretical foundation, the outlined methods are simple to implement and simple to adjust for alternative sensor configurations.

A. Separation into journeys

The analysis of ABA data yields useful results only when the vehicle is moving. Therefore, the collected data is divided

---

[1]Hafenbetriebsgesellschaft Braunschweig mbH, http://www.braunschweig-hafen.de/

[2]ROS, http://www.ros.org/

[3]https://www.u-blox.com/
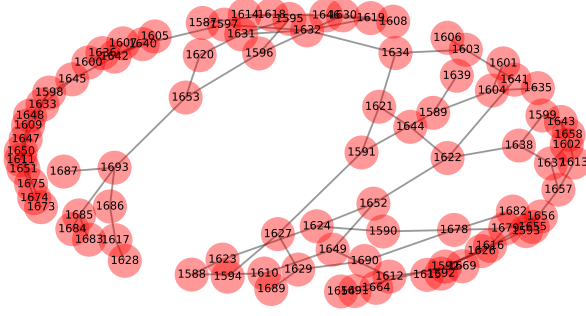
[4]https://www.xsens.com/

Fig. 2. The graph of the railway network with track identifiers as node labels.

into journeys that comprise the motion between two stops. Vehicle motion can be easily detected from the GNSS speed or IMU data. Thresholds for minimum journey lengths (e.g., $>10\,\text{s}$) and minimum speeds (e.g., $>1\,\text{m/s}$) ensure that irrelevant stop-and-go motion is excluded, e.g., while loading bulk cargo to attached wagons.

The intuitive separation into journeys comes with several advantages. Journeys start and end with zero speed, and within a journey the speed does not change sign. The driving direction (forward or backward facing) can be determined from the signs of the longitudinal acceleration data at the beginning of a journey. The data recorded between journeys can be employed to determine the start and end tracks and positions. Without any motion between journeys, the start position of journey i is the end position of journey $i-1$. The recorded data in standstill can be used for sensor calibration, i.e., bias correction of the acceleration data. Finally, the track-bound motion within a journey can be used to find the driven path, i.e., a sequence of tracks.

B. A graph of the railway network

The digital railway map of Sec. II comprises information about the N_t tracks of the network, e.g., track identifiers t_i and lengths $L_{\text{t},i}$ with $i = 1, \ldots, N_\text{t}$. The connection information enters in the form of shared start or end points and can be used to construct a mathematical graph $\mathcal{G}$ [21] with one vertex per track. Edges of $\mathcal{G}$ are included only for allowed transitions at track switches, which can be determined from the angles of the connected line segments. Fig. 2 illustrates the graph of our test railway network with 89 tracks. The node labels are the track identifiers t_i.

Positions on track t are determined by a scalar distance $d_\text{t} \in [0, L_\text{t}]$. The connection type of adjacent tracks, e.g., that $d_{\text{t},i} = L_{\text{t},i}$ on t_i corresponds to $d_{\text{t},j} = 0$ on t_j, can be encoded in the edge weights of $\mathcal{G}$. This is relevant to create path coordinate frames from the railway network graph.

Our Python implementation uses the NetworkX package[5] for representing $\mathcal{G}$, which provides several useful functions.

C. Local Cartesian, track, and path coordinates

Arbitrary horizontal positions are expressed in a local Cartesian coordinate frame and denoted as $p = (x, y)$. For the conversion of the GNSS latitude and longitude, the Universal Transverse Mercator projection (UTM, grid zone 32N) is used. However, the origin is shifted such that $p = (0, 0)$ corresponds to the position of the vehicle depot in the harbor area.

On-track distances d_t can be easily mapped to positions p by finding the corresponding line segment of the track t. Let $p = \phi_\text{t}(t, d_\text{t})$ denote the function for this conversion from track coordinates to Cartesian coordinates.

For arbitrary p there are no exact (t, d_t) in general. However, orthogonal projections can be used to find the closest track coordinates. Such simple map-matching is denoted by $(t, d_\text{t}) = \psi_\text{t}(p)$.

Journeys typically comprise motion on a sequence of tracks. Therefore, we introduce the concept of a path P which comprises an ordered sequence of tracks $[t_i]$, $i = 1, \ldots, M_\text{P}$, visited during a journey. Valid paths must comply with the structure of $\mathcal{G}$. That is, for t_{i-1} and t_i that are connected at $d_{\text{t},i} = 0$, t_i and t_{i+1} must be connected at $d_{\text{t},i} = L_{\text{t},i}$. To simplify the handling of such information, track orientations $o_{\text{t},i} \in \{-1, 1\}$ with respect to the path are introduced. The length L of a path P follows from the sum of track lengths:

$$L = \sum_{i=1}^{M_\text{P}} L_{\text{t},i}. \tag{1}$$

Similar to track coordinates (t, d_t), we can introduce path coordinates (P, d) that describe positions on the path P with a scalar variable $d \in [0, L]$. Using the track lengths $L_{\text{t},i}$ and the orientations $o_{\text{t},i}$ of P, distances d can be easily mapped to (t, d_t) and subsequently to Cartesian coordinates. Let

$$p = \phi(P, d) \tag{2}$$

denote the function for this conversion. Furthermore, let

$$d = \psi(p, P) \tag{3}$$

denote the path distance obtained by projecting p onto the tracks of P, expressed in the path coordinates of P.

D. Time considerations

A main objective of the georeferencing is to obtain regularly sampled position stamps with a higher sampling rate than the GNSS frequency, which is varying around an average of $5\,\text{Hz}$. This is achieved by running algorithms that are based on discrete-time state-space models with the IMU sampling rate of $100\,\text{Hz}$. Let $k = 0, \ldots, L$ be the time index that describes the times of a journey in steps of $T = 10\,\text{ms}$. The GNSS and IMU time stamps are quantized to appear as multiples of T. Furthermore, delayed IMU measurements that sporadically arrive in batches are re-distributed before processing.

Clearly, GNSS and IMU data are not available for all k due to the low GNSS rate and sporadic inconsistencies in the IMU data. Hence, we introduce the index sets S_GNSS and S_IMU that contain the k with available GNSS and IMU data, respectively.

E. Generation of path hypotheses

A central idea of our approach is to first estimate likely path hypotheses from the GNSS data, and second, to perform state estimation in path coordinates given each path hypothesis. Hence, the objective is to estimate a path P and a sequence of states $x_{0:L} = \{x_0, x_1, \ldots, x_L\}$ from a sequence of measurements $y_{1:L}$. Even without specifying $x_{0:L}$ and $y_{1:L}$, we can assert that the Bayesian solution [4] amounts to finding the conditional mixed point mass and probability density function

$$p(P, x_{0:L} \,|\, y_{1:L}) = p(P \,|\, y_{1:L}) p(x_{0:L} \,|\, P, y_{1:L}). \quad (4)$$

Apparently, the problem can be split into finding the conditional path probabilities $p(P \,|\, y_{1:L})$ and a smoothing problem [4] given the paths, $p(x_{0:L} \,|\, P, y_{1:L})$.

Now, the point mass function $p(P \,|\, y_{1:L})$ is not simple to compute. After all, it would require assessing all possible paths in the network. Instead, we pursue the following pragmatic approach based on the GNSS position measurements $\bar{p}_k$, $k \in S_{\mathsf{GNSS}}$. With the standstill data between journeys at hand, likely start and end tracks can be estimated easily. A list of path hypotheses P_i, $i = 1, \ldots, N_P$, between these track pairs can be queried using the network graph $\mathcal{G}$. For each P_i, path-matched Cartesian positions can be obtained using (2) and (3),

$$\bar{p}'_k = \phi\big(P_i, \psi(\bar{p}_k, P_i)\big). \quad (5)$$

The above positions on the path P_i allow for computing projection errors $\bar{p}'_k - \bar{p}_k$ from the GNSS data, which can be assessed to find the best P_i. The intuition behind this method is that the GNSS position sequences, although corrupted by local errors, still contain enough information about the path. Metrics for assessing the projection errors are subject to closer investigation. Important is that they are robust with respect to spurious large errors (unlike mean squared errors) but also take into account the entire error sequence (unlike median errors).

F. Filtering and smoothing in path coordinates

The following paragraphs describe the sensor fusion of GNSS and IMU data given a valid path P. By working in path coordinates, the horizontal positioning is reduced to estimating scalar distances d_k. This exploits the track-bound vehicle motion and reduces the number of unknowns to be estimated. By working on journeys, we furthermore know whether the vehicle is moving forward or backward and that the speed s_k does not change sign.

The employed Kalman filter (KF) framework is based on discrete-time state-space models. The state vector is given by

$$x_k = \big(d_k, \quad s_k\big) \quad (6)$$

and comprises the distance and speed in path coordinates. The temporal evolution of x_k is well described by a constant velocity model

$$x_{k+1} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} v_k = F x_k + G v_k, \quad (7)$$

which is a common choice for positioning problems. The process noise v_k is characterized by

$$v_k \sim \mathcal{N}(\bar{a}_k, Q), \quad k \in S_{\mathsf{IMU}}, \quad (8)$$

and used to include the longitudinal acceleration measurements $\bar{a}_k$. Hence, the IMU data are interpreted as noisy model input rather than measurements. For $k \notin S_{\mathsf{IMU}}$ we assume $\bar{a}_k = 0$. Depending on the driving direction of the rail vehicle, the sign of all $\bar{a}_k$ is adjusted such that positive values increase s_k.

From the above model follows a KF time update [4], [5]

$$\hat{x}_{k+1|k} = F \hat{x}_{k|k} + G \bar{a}_k, \quad (9a)$$
$$P_{k+1|k} = F P_{k|k} F^{\mathrm{T}} + G Q G^{\mathrm{T}} \quad (9b)$$

that provides a predicted state estimate and its error covariance matrix based on the filtering results of time k. The speed predictions $\hat{s}_{k+1|k}$ are not constrained to be positive, which would violate the journey assumptions. If necessary, the results are therefore adjusted according to

$$\hat{s}_{k+1|k} = \max(\hat{s}_{k+1|k}, 0). \quad (10)$$

The measurement vector y_k contains the GNSS position and speed measurements for $k \in S_{\mathsf{GNSS}}$. The measurement equation is formally given by

$$y_k = \big(\bar{p}_k, \quad \bar{s}_k\big) = h(x_k) + e_k, \quad (11)$$

where e_k describes the measurement noise to be specified. Whereas the speed is linearly related to x_k, the position measurements exhibit a nonlinear relation. Accordingly, a nonlinear KF update must be performed [22] for the latter.

We propose separate KF measurement updates for $\bar{p}_k$ and $\bar{s}_k$. This provides the means to select which measurements to process depending on the GNSS accuracy indications, and retains a linear KF update for $\bar{s}_k$. The theory requires uncorrelated measurement errors for separate updates [4], which appears in conflict with the fact that $\bar{p}_k$ and $\bar{s}_k$ stem from the same device. However, the speed measurements are derived from the Doppler shifts in the carrier satellite signals whereas the positions are computed from the corresponding pseudo-ranges [13]. Hence, explicit dependencies are difficult to model and the assumption of zero correlation is justified.

Because the GNSS speed derived from Doppler shifts is typically very accurate, the measurements $\bar{s}_k$ are processed first. Their relation to the state vector is given by

$$\bar{s}_k = \begin{bmatrix} 0, 1 \end{bmatrix} x_k + \tilde{s}_k = H^{\mathsf{s}} x_k + \tilde{s}_k. \quad (12)$$

The noise $\tilde{s}_k$ in the speed measurements is modeled by

$$\tilde{s}_k \sim \mathcal{N}(0, R^{\mathsf{s}}_k), \quad (13)$$

where R^{s}_k is determined by the accuracy parameter provided by the receiver, increased by a tuning factor.

The Kalman gain and the updated state estimate and covariance are given by

$$K_k^{\mathsf{s}} = P_{k|k-1}(H^{\mathsf{s}})^{\mathrm{T}} / \left(H^{\mathsf{s}} P_{k|k-1}(H^{\mathsf{s}})^{\mathrm{T}} + R_k^{\mathsf{s}} \right), \tag{14a}$$

$$\hat{x}_{k|k}^{\mathsf{s}} = \hat{x}_{k|k-1} + K_k^{\mathsf{s}} \left(\bar{\mathsf{s}}_k - H^{\mathsf{s}} \hat{x}_{k|k-1} \right), \tag{14b}$$

$$P_{k|k}^{\mathsf{s}} = P_{k|k-1} - K_k^{\mathsf{s}} \left(H^{\mathsf{s}} P_{k|k-1}(H^{\mathsf{s}})^{\mathrm{T}} + R_k^{\mathsf{s}} \right) (K_k^{\mathsf{s}})^{\mathrm{T}}. \tag{14c}$$

For absent (or rejected) speed measurements, the prediction results are maintained, i.e., $\hat{x}_{k|k}^{\mathsf{s}} = \hat{x}_{k|k-1}$ and $P_{k|k}^{\mathsf{s}} = P_{k|k-1}$.

The measurement equation for the GNSS position is

$$\bar{\mathsf{p}}_k = \phi(\mathsf{P}, \mathsf{d}_k) + \tilde{\mathsf{p}}_k = h^{\mathsf{p}}(x_k) + \tilde{\mathsf{p}}_k \tag{15}$$

and contains the nonlinear mapping from path to Cartesian coordinates. Again, the noise is assumed Gaussian with

$$\tilde{\mathsf{p}}_k \sim \mathcal{N}(0, R_k^{\mathsf{p}}) \tag{16}$$

and R_k^{p} as increased accuracy parameter provided by the GNSS receiver.

Because of the nonlinear relation in (15), a nonlinear KF measurement update is required [22]. We employ an extended Kalman filter (EKF) correction with a numerically computed Jacobian matrix of $\phi(\mathsf{P}, \mathsf{d}_k)$ that is column-wise given by

$$H_k^{\mathsf{p}}(:,1) = \frac{1}{\Delta\mathsf{d}} \left(\phi(\mathsf{P}, \hat{\mathsf{d}}_{k|k}^{\mathsf{s}} + \tfrac{\Delta\mathsf{d}}{2}) - \phi(\mathsf{P}, \hat{\mathsf{d}}_{k|k}^{\mathsf{s}} - \tfrac{\Delta\mathsf{d}}{2}) \right),$$
$$H_k^{\mathsf{p}}(:,2) = \begin{pmatrix} 0 & 0 \end{pmatrix} \tag{17}$$

and determined by a step length $\Delta\mathsf{d}$. The EKF update is completed by

$$K_k^{\mathsf{p}} = P_{k|k}^{\mathsf{s}}(H_k^{\mathsf{p}})^{\mathrm{T}} \left(H_k^{\mathsf{p}} P_{k|k}^{\mathsf{s}}(H_k^{\mathsf{p}})^{\mathrm{T}} + R_k^{\mathsf{p}} \right)^{-1}, \tag{18a}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k}^{\mathsf{s}} + K_k^{\mathsf{p}} \left(\bar{\mathsf{p}}_k - \phi(\mathsf{P}, \hat{\mathsf{d}}_{k|k}^{\mathsf{s}}) \right), \tag{18b}$$

$$P_{k|k} = P_{k|k}^{\mathsf{s}} - K_k^{\mathsf{p}} \left(H_k^{\mathsf{p}} P_{k|k}^{\mathsf{s}}(H_k^{\mathsf{p}})^{\mathrm{T}} + R_k^{\mathsf{p}} \right) (K_k^{\mathsf{p}})^{\mathrm{T}}. \tag{18c}$$

Again, previous results are kept for absent or rejected measurements $\bar{\mathsf{p}}_k$, i.e., $\hat{x}_{k|k} = \hat{x}_{k|k}^{\mathsf{s}}$ and $P_{k|k} = P_{k|k}^{\mathsf{s}}$. Furthermore, the speed is adjusted similar to (10) after (14) and (18).

There are alternatives to the above EKF correction (18). Based on experiments, it could be shown that the choice of KF variant is secondary. Unscented KF and EKF with divided differences [22] show similar performance in this specific example.

The KF equations (9), (14), and (18) are initialized with a state estimate $\hat{x}_{0|0}$ and its covariance $P_{0|0}$, computed from the measurements prior to and at $k = 0$, and iterated for all $k = 0, \ldots, L$. The filtering and prediction results form the basis for the following smoothing iteration.

Because of the separate estimation of the driven path, the above filtering stage is fundamentally different from other approaches that jointly estimate track identifiers and on-track positions. The explicit treatment of path hypotheses allows for the application of KF to achieve the actual fusion of sensor and map data. With the typically low number of likely path hypotheses, the computational costs are low compared to, e.g., the particle methods of [15] that require a large number of particles (joint track and state hypotheses in essence) to work

well. A multiple KF approach for online applications with an online management of path hypotheses, an idea that [18] implemented for road networks, appears promising and is subject to future work.

The fact that we operate in an offline-setting allows for the improvement of the filtering results via smoothing. The smoothing extension to the KF is called the Rauch-Tung-Striebel (RTS) smoother [4], [5] and amounts to a backward iteration for $k = L - 1, \ldots, 0$. The update equations

$$\Lambda_k = P_{k|k} F^{\mathrm{T}} P_{k+1|k}^{-1}, \tag{19a}$$

$$\hat{x}_{k|L} = \hat{x}_{k|k} + \Lambda_k (\hat{x}_{k+1|L} - \hat{x}_{k+1|k}), \tag{19b}$$

$$P_{k|L} = P_{k|k} + \Lambda_k \left(P_{k+1|L} - P_{k+1|k} \right) \Lambda_k^{\mathrm{T}} \tag{19c}$$

exhibit structural similarities with the measurment updates (14) and (18). However, previously obtained prediction, filtering, and smoothing results are combined without re-processing any measurements y_k. The RTS iteration is typically initialized with the filtering results at $k = L$. However, we include the accurately determined final position from the standstill data and the zero-speed knowledge here.

A major advantage of the RTS iteration is that possible jumps in the filtering results at the times $k \in S_{\mathsf{GNSS}}$, both in $\hat{x}_{k|k}$ and $P_{k|k}$, are smoothed out. This is especially useful for longer periods of absent GNSS data during which only time updates (9) are carried out. Here, $P_{k|k} = P_{k|k-1}$ increases until the next available $k \in S_{\mathsf{GNSS}}$. The difference between $\hat{x}_{k|k}$ and $\hat{x}_{k|k-1}$ is larger due to the large difference between the measurements and the predicted outputs. The results of (19) exhibit much more consistent results in general. An important requirement for a working smoother is that the motion model (7) reflects the vehicle motion accurately, though.

IV. Experimental Results

The here presented experiments have been carried out on data collected in the Braunschweig harbor railway since September 2015. In the period until June 2017, more than 6000 journeys longer than 10 seconds were recorded. This amounts to more than 140 hours of journey motion. The journeys have a median and mean length of 80 and 50 seconds, respectively. The longest journey is over 18 minutes long. The deployed systems are still in use and recording continuously.

Due to space constraints, results are presented for two representative journeys.

A. On the GNSS and IMU measurements

We discussed the discrepancies in sampling time between the GNSS and IMU data. Here, we show examples that demonstrate their error characteristics.

Fig. 3 illustrates the GNSS positions of one journey from south-east to north-west. The coloring illustrates the respective GNSS speeds. The positions appear close to the driven tracks most of the times, but show larger deviations while passing under a highway bridge in the lower right corner of the figure. Fig. 4 provides a closer look at this segment. The compromised
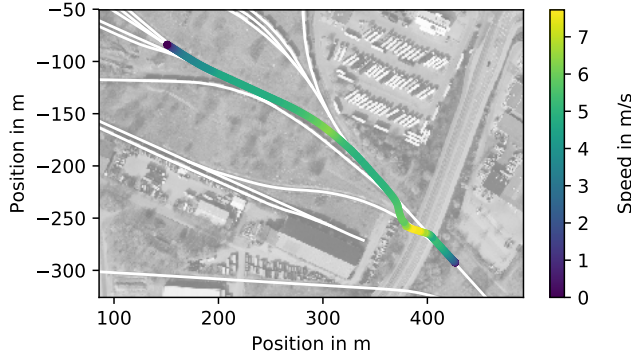
Fig. 3. The GNSS data of one journey, starting in the south-east. The GNSS positions do not lie on tracks exactly and exhibit larger errors while passing the highway bridge in the lower right of the figure. The GNSS speed is included via line coloring and exhibits unlikely deviations next to the bridge.
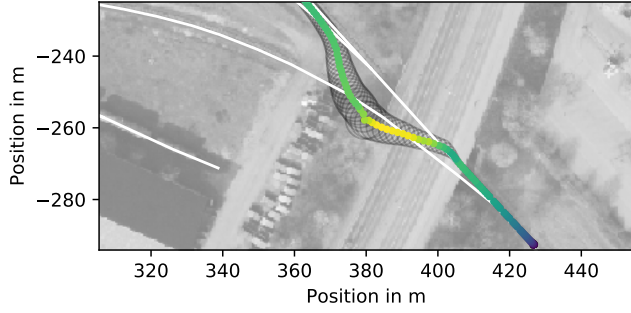

Fig. 4. A zoom of the data in Fig. 3 highlights the errors next to the bridge. The black circles illustrate the optimistic GNSS uncertainty as provided by the receiver.

satellite view leads to large GNSS errors. Clearly, simple map matching to the closest tracks leads to inconsistent results here. Also illustrated is a 95 percent uncertainty circle derived from the horizontal position accuracy parameter provided by the receiver. Although the receiver is over-confident, the qualitative information is still valuable and used to determine the position measurement covariance R_k^{p} used in our sensor fusion algorithm. Furthermore, the speed measurements exhibit false accelerations after entering the underpass.

Speed measurements for a different journey are illustrated in Fig. 5. Specifically, we show the GNSS speed with uncertainty intervals and the integrated longitudinal acceleration as functions of time. The GNSS speed exhibits a lower sampling rate and spurious larger errors, e.g., when going through an underpass after the 250 second mark. Similar to the GNSS position, the uncertainty information provided by the receiver is optimistic. The dead reckoned IMU speed is available at a much higher rate, but drifts with time and heavily depends on the IMU bias. After all, a constant acceleration bias results in linearly and quadratically increasing speed and position errors, respectively. Clearly, fusion of IMU and GNSS data is required.
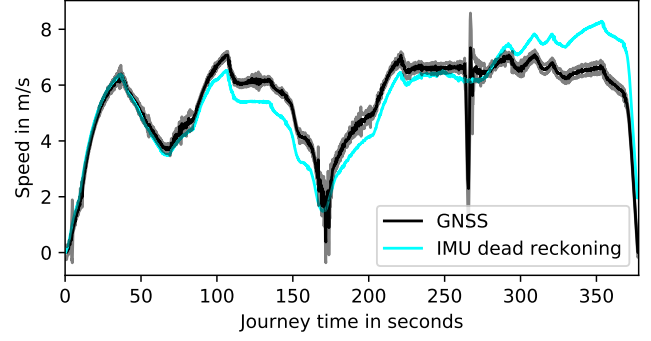

Fig. 5. GNSS speed measurements and the integrated longitudinal IMU acceleration. The former exhibits spurious large errors, the latter drifts with time due to bias errors.
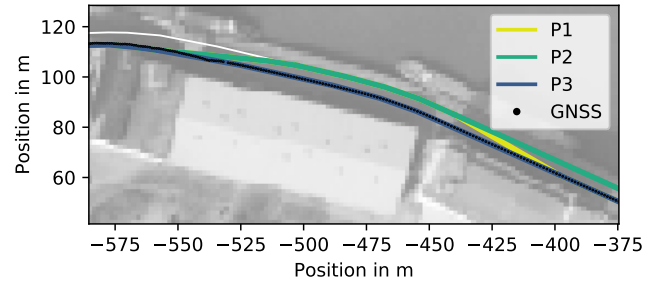

Fig. 6. Path hypotheses for one journey (zoom). The paths partially occlude another because of shared tracks. The GNSS positions lie close to P_3.
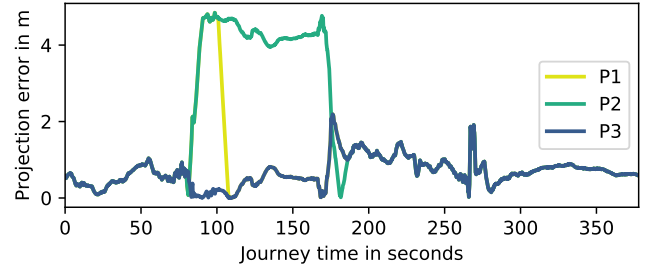

Fig. 7. Projection errors for the path hypotheses of Fig. 6. Hypothesis P_3 shows the best performance.

B. On finding paths

We here illustrate the method to determine the driven path from a sequence of GNSS positions, described in Sec. III-E. Fig. 6 shows a segment of the railway network with three path hypotheses (P_1, P_2, P_3, partially occluded by another). Also shown are the GNSS measurements that are very close to P_3.

Fig. 7 shows the corresponding projection errors between the measured and path-matched GNSS positions. Again, there is occlusion because the hypotheses share tracks. Nevertheless, it can be seen that P_3 clearly outperforms its competitors. It should be noted that several tracks with lower projection errors could exist, but that they have been excluded because they do
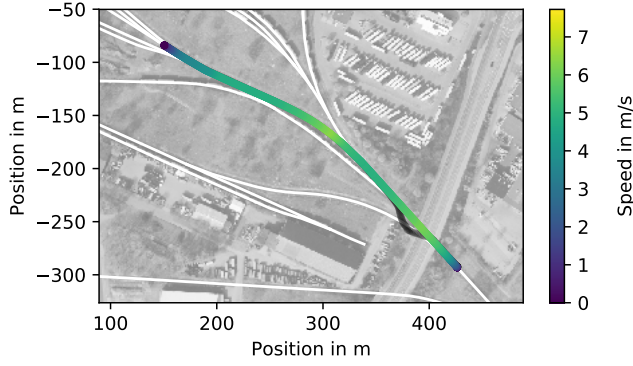
Fig. 8. RTS smoothing results for the journey of Fig. 3. The large errors in speed and position next to the bridge have been eliminated. The original GNSS positions are included in black.
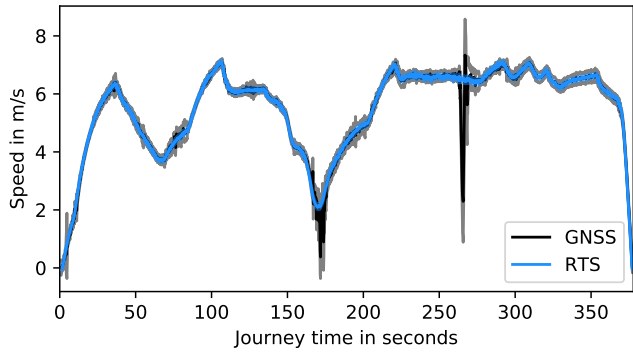

Fig. 9. RTS smoothing speed estimates including uncertainty intervals for the journey of Fig. 5. Also included are the GNSS speed measurements.
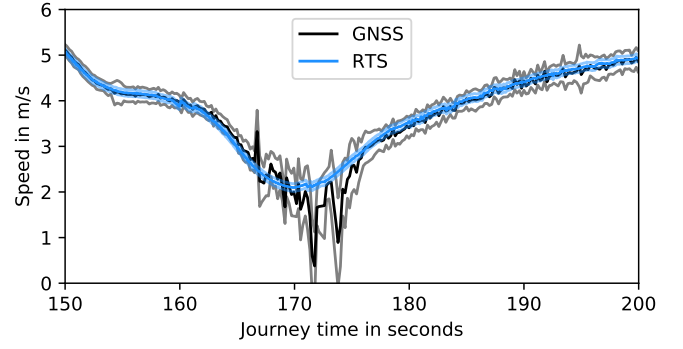

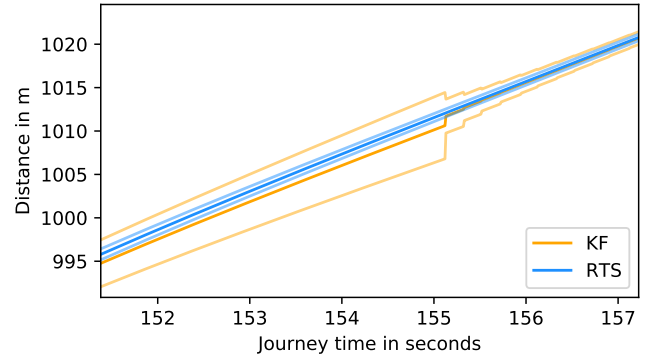Fig. 10. A zoom of Fig. 9 shows the smoothed speed in comparison to the noise-corrupted GNSS data.

Fig. 11. Comparison of KF and RTS results for the distance in path coordinates. To highlight the smoothing advantages in the absence of measurements, no GNSS data were processed between 140 and 155 seconds.

not belong to a valid path between the start and end tracks. One such example is the left branch under the bridge in Fig. 4.

C. Filtering and smoothing results

The RTS smoothing results for the journey of Fig. 3 are shown in Fig. 8. Clearly, the GNSS errors next to the bridge (Fig. 4) are corrected in terms of speed and position. It should be noted that the spatial resolution of position stamps has been increased due to the sampling rate of $100\,\mathrm{Hz}$.

The RTS smoothing speed estimates for the journey of Fig. 5 are given in Fig. 9. The smooth signal without the spurious errors of the GNSS speed reflects the vehicle motion very well. The robustness towards intermittent GNSS disturbances is further highlighted in the zoomed segment of Fig. 10.

A comparison of the KF and RTS results is given in Fig. 11. Illustrated are the distance estimates $\hat{\mathsf{d}}_{k|k}$ and $\hat{\mathsf{d}}_{k|L}$ obtained by filtering and smoothing with their respective uncertainty intervals. In order to highlight the smoothing advantages, the GNSS data between 140 and 155 seconds have been removed. Accordingly, the KF results show increasing uncertainties until the 155 second mark and a subsequent jump in the state estimate. The RTS results no longer exhibit these jumps and provide a more consistent picture.

V. CONCLUDING REMARKS

It has been shown that accurate track-selective georeferencing of monitoring data can be achieved by combining map information with IMU and GNSS data. Specifically, we have presented a two-stage approach which first determines likely path hypotheses from the GNSS positions, and second, calculates RTS-smoothed position and speed estimates in path coordinates. All developments have been carried out on real data collected in the Braunschweig harbor railway, with the purpose to provide position information for the monitoring data of a track condition monitoring system developed at DLR.

Future work includes the adaptation of the proposed algorithms to real-time positioning, based on online management of path hypotheses and multiple path-constrained Kalman filters. In order to challenge our algorithms in a larger testbed, a joint research project with the Hamburg Port Authority as operator of a larger industrial railway network has been initiated in 2017. From a positioning viewpoint, prospective research directions include the combination of further sensor data that is readily available with map information. That includes the up to now unused IMU data (turn rates, lateral and vertical accelerations) as well as the GNSS raw data (pseudo ranges, phase, Doppler).

REFERENCES

[1] R. Schenkendorf, J. C. Groos, and L. Johannes, "Strengthening the Rail Mode of Transport by Condition Based Preventive Maintenance," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 964–969, Jan. 2015.

[2] R. Schenkendorf and J. C. Groos, "Global Sensitivity Analysis applied to Model Inversion Problems: A Contribution to Rail Condition Monitoring," *International Journal of Prognostics and Health Management*, vol. 6, no. 19, May 2015.

[3] J. C. Groos, P. Havrila, and L. A. Schubert, "In-service railway track condition monitoring by analysis of axle box accelerations for small to mid-size infrastructure operators," in *WCCM 2017 - 1st World Congress on Condition Monitoring 2017*, London, UK, Jun. 2017.

[4] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge, UK: Cambridge University Press, 2013.

[5] M. Roth, *Advanced Kalman Filtering Approaches to Bayesian State Estimation*. Linköping Studies in Science and Technology. Dissertations. Linköping: Linköping University Electronic Press, 2017, vol. 1832.

[6] M. Molodova, Z. Li, A. Núñez, and R. Dollevoet, "Automatic Detection of Squats in Railway Infrastructure," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 1980–1990, Oct. 2014.

[7] G. Lederman, S. Chen, J. Garrett, J. Kovačević, H. Y. Noh, and J. Bielak, "Track-monitoring from the dynamic response of an operational train," *Mechanical Systems and Signal Processing*, vol. 87, no. Part A, pp. 1–16, Mar. 2017.

[8] S. S. Saab, "A map matching approach for train positioning. I. Development and analysis," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 467–475, Mar. 2000.

[9] S. Hensel, C. Hasberg, and C. Stiller, "Probabilistic Rail Vehicle Localization With Eddy Current Sensors in Topological Maps," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1525–1536, Dec. 2011.

[10] O. Heirich, B. Siebler, and E. Hedberg, "Study of Train-Side Passive Magnetic Measurements with Applications to Train Localization," *Journal of Sensors*, vol. 2017, Article ID 8073982, 2017.

[11] J. Marais, J. Beugin, and M. Berbineau, "A Survey of GNSS-Based Research and Developments for the European Railway Signaling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2602–2618, Oct. 2017.

[12] K. M. Betts, T. J. Mitchell, D. L. Reed, S. Sloat, D. P. Stranghoener, and J. D. Wetherbee, "Development and operational testing of a sub-meter Positive Train Location system," in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, May 2014, pp. 452–461.

[13] P. D. Groves, *Principles of GNSS, inertial, and multi-sensor integrated navigation systems*, 2nd ed. Boston: Artech House, 2013.

[14] M. Lauer and D. Stein, "A Train Localization Algorithm for Train Protection Systems of the Future," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 970–979, Apr. 2015.

[15] O. Heirich, "Bayesian Train Localization with Particle Filter, Loosely Coupled GNSS, IMU, and a Track Map," *Journal of Sensors*, vol. 2016, Article ID 2672640, 2016.

[16] K. Gerlach and C. Rahmig, "Multi-hypothesis based map-matching algorithm for precise train positioning," in *12th International Conference on Information Fusion*, Jul. 2009, pp. 1363–1369.

[17] K. Lüddecke and C. Rahmig, "Evaluating multiple GNSS data in a multi-hypothesis based map-matching algorithm for train positioning," in *2011 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2011, pp. 1037–1042.

[18] M. Ulmke and W. Koch, "Road-map assisted ground moving target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 4, pp. 1264–1274, Oct. 2006.

[19] C. Hasberg, S. Hensel, and C. Stiller, "Simultaneous Localization and Mapping for Path-Constrained Motion," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 541–552, Jun. 2012.

[20] S. Knake-Langhorst, K. Gimm, T. Frankiewicz, and F. Köster, "Test Site AIM – Toolbox and Enabler for Applied Research and Development in Traffic and Mobility," *Transportation Research Procedia*, Transport Research Arena TRA2016, vol. 14, pp. 2197–2206, Jan. 2016.

[21] A. Bondy and M. R. Murty, *Graph Theory*, Graduate Texts in Mathematics. London: Springer, 2008.

[22] M. Roth, G. Hendeby, and F. Gustafsson, "Nonlinear Kalman Filters Explained: A Tutorial on Moment Computations and Sigma Point Methods," *Journal of Advances in Information Fusion*, vol. 11, no. 1, pp. 47–70, 2016.