Systems & Control Letters 178 (2023) 105568

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Second-order-optimal filtering on $SE(2) \times \mathbb{R}^2$ for the Chaplygin sleigh

Damiano Rigo [a,*], Nicola Sansonetto [b], Riccardo Muradore [a]

[a] Department of Engineering for Innovation Medicine, University of Verona, Strada le Grazie 15, Verona, 37134, Italy
[b] Department of Computer Science, University of Verona, Strada le Grazie 15, Verona, 37134, Italy

ARTICLE INFO

Article history:
Received 5 April 2022
Received in revised form 10 February 2023
Accepted 24 May 2023
Available online xxxx

Keywords:
Nonlinear systems
Filtering
Nonholonomic constraints
Lie groups

ABSTRACT

In the last decades many filters have been proposed to solve the problem of pose estimation that naturally arises in robotics, both for manipulators and mobile robots. The dynamics of mobile robots can be mathematically described by exploiting the theory of Lie groups embedding holonomic and nonholonomic constraints. In this paper we design a second-order-optimal filter for the so-called Chaplygin sleigh, that is a mechanical system with a nonholonomic constraint. In particular we examine the importance to know the exact dynamic equations, and to exploit the underline Lie group structure of the system. Moreover, we investigate the conditions that ensure the preservation of the nonholonomic constraint by properly choosing the affine connection which guarantees that the orthogonal velocity is equal to zero. In this work the sensing system consists of a GPS-like configuration (Global Positioning System) obtained by using two antennas attached to the planar rigid body and an INS-like unit (Inertial Navigation System) to measure the velocities.

1. Introduction

The problem of estimating the pose and velocity is crucial in mechanical and robotic applications since in many cases the available sensors do not provide their direct measurements. Moreover, measurements are affected by noise that it is necessary to filter out. Accurate estimations are also important to effectively solve the tracking problem.

In the last decades many observers and filters have been proposed in the literature. The most famous approach is the Kalman filter [1], that uses a recursive algorithm to produce estimates of state variables by updating a joint probability distribution at each timeframe. The Kalman filter is optimal only for linear systems with Gaussian noises, however, generalizations aiming at extending it to nonlinear systems, such as the extended Kalman filter (EKF), have been proposed [2]. The EKF computes at each step the linear approximation of the dynamic and measurement equations. One way to overcome the Gaussian assumption is by using the unscented Kalman filter [3] and the Particle filter [4] that approximate the probability density function with a certain number of sampling points and updates their values according to the past state and current measurements.

Another generalization to nonlinear mechanical systems can be made by solving a minimum-energy optimization problem generating progressive realizable approximations of the minimum-energy functional as done in [5]. The solution is then obtained by differentiating the boundary conditions of the associated optimal control problem. Starting from this work, the authors of [6–8] proposed a *second-order-optimal minimum-energy filter* carrying on the idea of exploiting the Lie group structure of the phase space of a mechanical system to minimize the square of the estimation error (energy). The main theoretical result in [8] is given for a generic Lie group, and the authors specify that in the case of a (dynamic) mechanical system this group is the tangent bundle of a smaller Lie group. It is called *second-order*-optimal in the sense that it is a truncation of the exact solution that would be an infinite dimensional system. The filter takes the form of a gradient observer coupled with a Riccati-like differential equation that updates its gain (similarly to the standard Kalman filter).

If nonholonomic constraints, that naturally arise in planning and tracking problems, are taken into account, further geometric adaptations are needed. An example of a nonholonomic system is the so-called Chaplygin sleigh (see e.g. [9,10]), that models a platform, supported on two points and on a blade. The blade is free to rotate, but it cannot slide in the orthogonal direction.

In this paper we extend the second-order-optimal minimum energy filter of [8,11], designed for Lie groups, to the nonholonomic realm, and specifically to the Chaplygin sleigh system. We stress that this work is not a simple application of the previous results since the presence of a nonholonomic constraint changes significantly the dynamic behavior and the geometric structures; precisely, it restricts the space of admissible velocities and it does not allow to describe the phase space as the tangent bundle (of a Lie group in this specific case). Thus in the filter design particular attention has to be given to the use of Hamel's coordinates, and

* Corresponding author.
E-mail address: damiano.rigo@univr.it (D. Rigo).

https://doi.org/10.1016/j.sysconle.2023.105568
0167-6911/

(left-trivialized) dynamics that put in evidence the Lie group structure of the state space. In this paper we will focus on the finite-horizon estimation problem on $\mathrm{SE}(2)\times\mathbb{R}^2$. This formulation will then be the starting point for studying steady-state stability and robustness in the future.

The paper is organized as follows. In Section 2 we report the notations used in this paper. Section 3 is divided into two parts, in Section 3.1 we describe the Chaplygin sleigh dynamics while in Section 3.2 we recall the structure of the $\mathrm{SE}(2)\times\mathbb{R}^2$ group and its Lie algebra. The explicit filter for the nonholonomic case constructed considering the Cartan–Schouten (0)-connection is presented in Section 4. In Section 5 we investigate the conditions to design a filter that satisfies the nonholonomic constraint, whose implementation could be relevant for control purposes. Section 6 is devoted to some numerical simulations and comparison of the two filters, and conclusions are drawn in Section 7.

2. Notation

Notation on the structure of Lie groups adapted to the case of the Chaplygin sleigh that will be studied in this paper. For a detailed discussion on Lie groups see, e.g., [12–14].

g	element of the Lie group $G = \mathrm{SE}(2)$
$V = (\omega, v)$	element of the Lie group $\mathbb{R}^2$
$\bar{g} = (g, V)$	element of the Lie group $\bar{G} = G \times \mathbb{R}^2$
$\boldsymbol{x} = (\theta, x, y, \omega, v)$	vector representation of an element in $\bar{G}$
$\mathfrak{g} = \mathfrak{se}(2)$	Lie algebra of G
$\bar{\mathfrak{g}} = \mathfrak{g} \times \mathbb{R}^2$	Lie algebra of $\bar{G}$
η^g	element of the Lie algebra $\mathfrak{g}$
$\eta^V = (\eta^\omega, \eta^v)$	element of the Lie algebra $\mathbb{R}^2$
$\eta^{\bar{g}} = (\eta^g, \eta^V)$	element of the Lie algebra $\bar{\mathfrak{g}}$
$\widehat{\ } : \mathbb{R}^3 \to \mathfrak{se}(2)$	Lie algebra isomorphism
$\eta^{\boldsymbol{x}} = (\eta^\theta, \eta^x, \eta^y, \eta^\omega, \eta^v)$	element of the Lie algebra $\mathbb{R}^5$
$T_eL_{\bar{g}}(\eta^{\boldsymbol{x}}) = (\theta', x', y', \omega', v')$	vector form of an element in $T_{\bar{g}}\bar{G}$
$\mathrm{ad}_X : \mathbb{R}^5 \to \mathbb{R}^5$	matrix representation of the adjoint operator $\mathrm{ad}_X : \bar{\mathfrak{g}} \to \bar{\mathfrak{g}}$
$T_eL_{\bar{g}} : \mathbb{R}^5 \to \mathbb{R}^5$	matrix representation of the tangent map $T_eL_{\bar{g}} : \bar{\mathfrak{g}} \to T_{\bar{g}}\bar{G}$

3. The Chaplygin sleigh and its geometric structure

3.1. The Chaplygin sleigh

The Chaplygin sleigh is a nonholonomic system that models a planar rigid body supported at three points, two of which slide freely while the third is a blade at distance a from the center of mass and that cannot move perpendicularly. Let $\Sigma_b = \{e_1^b, e_2^b\}$ be the right-handed body frame attached to the Chaplygin sleigh centered at the contact point between the blade and the ground with the e_1^b-axis aligned with the blade, and let $\Sigma_i = \{e_x^i, e_y^i\}$ be an inertial frame (also called space frame) fixed in space as shown in Fig. 1.

The configuration space is $\mathrm{SE}(2)$ with coordinates $q = (\theta, x, y)$, where (x, y) denote the position of the contact point of the blade in the plane, and θ is the angle that the blade forms with the horizontal axis e_x^i. The velocity components, also named quasi or Hamel's velocities, with respect to the body frame Σ_b are $(\omega, v, v_\perp)$, where ω is the angular velocity, v and $v_\perp$ are the (linear) velocities of the body along the e_1^b and e_2^b axes, respectively. The nonholonomic constraint imposes that the orthogonal component $v_\perp$ of the velocity with respect to the blade vanishes,

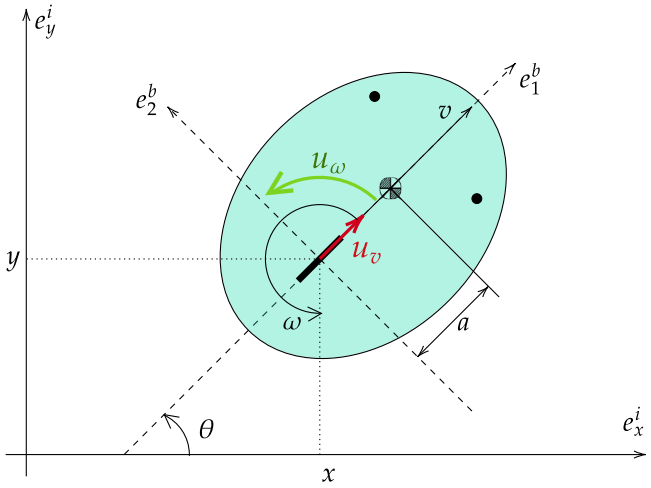

Fig. 1. Planar rigid body where the blade is indicated with a thicker segment and the two passive supporting wheels as •.

namely $v_\perp = 0$. In the inertial frame, where $(\dot\theta, \dot x, \dot y)$ are the components of the velocity, the constraint reads as

$$\dot x \sin\theta - \dot y \cos\theta = 0. \tag{1}$$

The velocities $(\dot\theta, \dot x, \dot y)$ and quasi-velocities $(\omega, v, v_\perp)$ are related by

$$\dot x = v\cos\theta, \quad \dot y = v\sin\theta, \quad \dot\theta = \omega. \tag{2}$$

The nonholonomic constraint (1) defines a constant rank-2 distribution $\mathcal{D}$ on the configuration space locally generated by

$$X_v = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}, \quad X_\omega = \frac{\partial}{\partial\theta}, \tag{3}$$

called constrained manifold. The state space, that is the constrained manifold, can be therefore identified with $\mathrm{SE}(2)\times\mathbb{R}^2$ and parameterized by $(\theta, x, y, \omega, v)$.

The center of mass (x_c, y_c) is settled at distance a from the contact point (x, y) according to the equations

$$(x_c, y_c) = (x + a\cos\theta, y + a\sin\theta). \tag{4}$$

We denote by J the inertia along the axis passing to the center of mass and orthogonal to the plane and with m the mass of the rigid body. The control inputs $\boldsymbol{u}(t) = (u_\omega(t), u_v(t))$ are functions of time that act respectively as a torque applied around the contact point and a force applied to the center of the body frame along e_1^b as shown in Fig. 1.

In order to put in evidence the geometric structure of the dynamics, it is useful to consider Hamel's approach to the equations of motion of a nonholonomic system (see [15,16]). Hamel's equations with external input are then

$$\begin{aligned} m\dot v &= ma\omega^2 + u_v, \\ (J + ma^2)\dot\omega &= -ma v\omega + u_\omega. \end{aligned} \tag{5}$$

Eqs. (5) together with (2) define the motion of the Chaplygin sleigh with external forces. We stress the fact that the constrained manifold $\mathcal{D}$ can be identified with $\mathrm{SE}(2)\times\mathbb{R}^2$, and therefore it can be endowed with a Lie group structure.

Since we assume that our model is not perfectly accurate, we add a model error that consists on an additive term that takes into account unmodeled dynamics and uncertainty on the parameters' values. According to [8], it only affects the evolution of the velocity of the system and not the kinematics. This error is denoted by $\boldsymbol{\xi}(t) = (\xi_\omega(t), \xi_v(t))^T$ and is modeled as a Gaussian

white noise with zero mean and diagonal and positive definite variance Σ.

The controlled Chaplygin sleigh equations are then (see e.g. [15])

$$\begin{cases} \dot{\theta}(t) &= \omega(t) \\ \dot{x}(t) &= v(t)\cos\theta(t) \\ \dot{y}(t) &= v(t)\sin\theta(t) \\ \dot{\omega}(t) &= -\frac{ma}{J+ma^2}\omega(t)v(t) + \frac{1}{J+ma^2}u_\omega(t) + \xi_\omega(t) \\ \dot{v}(t) &= a\omega(t)^2 + \frac{1}{m}u_v(t) + \xi_v(t) \end{cases} \tag{6}$$

or, in compact form,

$$\dot{\boldsymbol{x}}(t) = f_{\boldsymbol{x}}(\boldsymbol{x}(t)) + f_{\boldsymbol{u}}(\boldsymbol{u}(t)) + (0,0,0,\xi_\omega(t),\xi_v(t))^T \tag{7}$$

where $\boldsymbol{x}(t) = (\theta(t), x(t), y(t), \omega(t), v(t))^T$, and $f_{\boldsymbol{x}} : \mathbb{R}^5 \to \mathbb{R}^5$, $f_{\boldsymbol{u}} : \mathbb{R}^2 \to \mathbb{R}^5$ are two functions given by

$$f_{\boldsymbol{x}}(\boldsymbol{x}(t)) = \begin{bmatrix} \omega(t) \\ v(t)\cos\theta(t) \\ v(t)\sin\theta(t) \\ -\frac{ma}{J+ma^2}\omega(t)v(t) \\ a\omega(t)^2 \end{bmatrix}, \quad f_{\boldsymbol{u}}(\boldsymbol{u}(t)) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{J+ma^2}u_\omega(t) \\ \frac{1}{m}u_v(t) \end{bmatrix}. \tag{8}$$

Let $\boldsymbol{y}(t_k) \in \mathbb{R}^p$ be the measurement vector at time t_k, where $\{t_k\}_{k\in\mathbb{N}}$ is a sequence of sampling instants such that $t_k = kT_s$ with T_s the sampling time. This sequence represents the set of instants when the measurements are available. The measurement equation is given by

$$\boldsymbol{y}(t_k) = h(\boldsymbol{x}(t_k)) + \boldsymbol{\varepsilon}(t_k) \tag{9}$$

where

$$h(\boldsymbol{x}(t_k)) = \begin{bmatrix} x(t_k) + \ell\,\cos(\theta(t_k)) \\ y(t_k) + \ell\,\sin(\theta(t_k)) \\ x(t_k) - \ell\,\cos(\theta(t_k)) \\ y(t_k) - \ell\,\sin(\theta(t_k)) \\ \omega(t_k) \\ v(t_k) \end{bmatrix}, \tag{10}$$

represents a GPS-like system that provides the position of two antennas on the rigid body located at distance ℓ to the origin of the body frame Σ_b (see Fig. 2), and an INS-like system that measures the angular velocity ω and the linear velocity v along the e_1^b axis. The measurement noise $\boldsymbol{\varepsilon}$ is modeled as a Gaussian white noise with zero mean and diagonal and positive definite variance Λ.

Remark 1. This hypothesis of Gaussianity for the model and measurement errors is not necessary for the filter since it considers the errors as unknown deterministic functions of time. We chose it only for the simulations.

3.2. The $\mathrm{SE}(2)\times\mathbb{R}^2$ structure

In this paper we identify elements of groups and algebras with their matrix representations. An element $g \in \mathrm{SE}(2)$ is represented by a 3×3 matrix

$$g = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix}. \tag{11}$$

Given the Lie algebra $\mathfrak{se}(2)$ of $\mathrm{SE}(2)$, we introduce the Lie algebra isomorphism $^\wedge : \mathbb{R}^3 \to \mathfrak{se}(2)$

$$\Omega = \begin{bmatrix} \eta^\theta \\ \eta^x \\ \eta^y \end{bmatrix} \longmapsto \Omega^\wedge = \eta^g = \begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix} \tag{12}$$

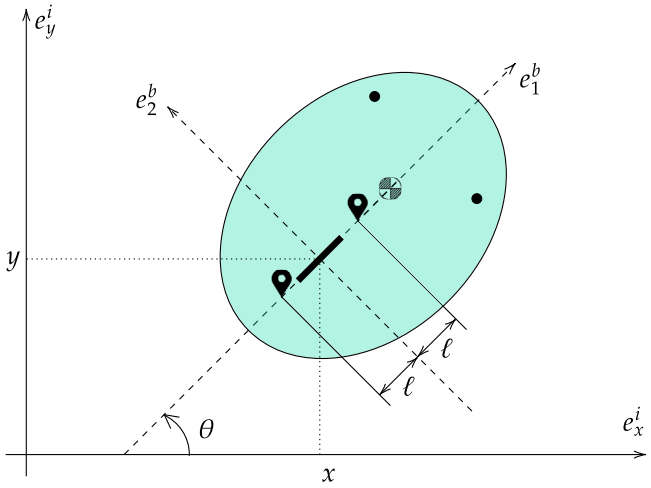

Fig. 2. Planar rigid body with two antennas.

from the Lie algebra $(\mathbb{R}^3, \star)$ to the matrix Lie algebra $(\mathfrak{se}(2), [\cdot,\cdot])$, where $\star : \mathbb{R}^3\times\mathbb{R}^3 \to \mathbb{R}^3$ is the Lie bracket operation defined as

$$\begin{bmatrix} \eta^{\theta_1} \\ \eta^{x_1} \\ \eta^{y_1} \end{bmatrix} \star \begin{bmatrix} \eta^{\theta_2} \\ \eta^{x_2} \\ \eta^{y_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \eta^{y_1}\eta^{\theta_2} - \eta^{\theta_1}\eta^{y_2} \\ \eta^{\theta_1}\eta^{x_2} - \eta^{x_1}\eta^{\theta_2} \end{bmatrix}, \tag{13}$$

and $[\cdot,\cdot]$ is the usual matrix commutator (see, e.g., [14]).

The tangent bundle $\mathrm{TSE}(2)$ is isomorphic to $\mathrm{SE}(2)\times\mathfrak{se}(2)$ via left translation; since we impose the nonholonomic constraint $v_\perp = 0$, a sub-bundle of it describes the admissible velocities. The velocity pair $V = (\omega, v)^T$ is an element of $\mathbb{R}^2$ that is a Lie group with respect to the sum operation, with abelian Lie algebra $\mathbb{R}^2$. The elements of such Lie algebra are denoted by $\eta^V = (\eta^\omega, \eta^v)$.

An element $\bar{g} = (g, V) \in \bar{G} := SE(2)\times\mathbb{R}^2$ can be represented by a 6×6 matrix

$$\bar{g} = \begin{bmatrix} g & 0_{3\times 2} & 0_{3\times 1} \\ 0_{2\times 3} & I_{2\times 2} & V \\ 0_{1\times 3} & 0_{1\times 2} & 1 \end{bmatrix}, \tag{14}$$

and the group operation is defined by

$$(g, V)\cdot(f, W) = (gf, V + W) \tag{15}$$

with unit element $e = (I_{3\times 3}, 0_{2\times 1})$ and inverse $(g,V)^{-1} = (g^{-1}, -V)$. The Lie algebra $\bar{\mathfrak{g}} = \mathfrak{se}(2)\times\mathbb{R}^2$ of $\bar{G}$ can be identified, up to Lie algebra isomorphism, with $\mathbb{R}^5$. We exploit this isomorphism to work on $\mathbb{R}^5$ and to report the operators that appear in the optimal filter equations in their matrix forms.

Let $\eta^{\bar{g}} = (\eta^g, \eta^V)$ be an element of $\bar{\mathfrak{g}}$, the matrix form of the $\mathrm{ad}_{\eta^{\bar{g}}}$ operator is given by the 5×5 matrix

$$\mathrm{ad}_{\eta^{\bar{g}}} = \begin{bmatrix} 0 & 0 & 0 & 0_{1\times 2} \\ \eta^y & 0 & -\eta^\theta & 0_{1\times 2} \\ -\eta^x & \eta^\theta & 0 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix}. \tag{16}$$

Finally, the matrix representation of the tangent map is

$$T_e L_{\bar{g}} = \begin{bmatrix} 1 & 0 & 0 & 0_{1\times 2} \\ 0 & \cos\theta & -\sin\theta & 0_{1\times 2} \\ 0 & \sin\theta & \cos\theta & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & I_{2\times 2} \end{bmatrix} \tag{17}$$

whose dual map matrix representation $T_e L_{\bar{g}}^*$ satisfies $T_e L_{\bar{g}}^* = (T_e L_{\bar{g}})^T$.

4. The second-order-optimal filter

We now report the main contribution of this paper, that is the design of the second-order-optimal filter for the Chaplygin sleigh presented in Section 3. The general formulation of this filter, [8], exploits the dynamic programming principle in order to provide the Hamilton–Jacobi–Bellman equation for the value function. The advantage of this method is that it is possible to compute the filter online and preserve the optimality. The non-trivial aspect of applying the general equation in a particular mechanical system is that the solution must be approximated since the analytical one would be an infinite dimensional system.

In the following we will denote with $\widehat{\cdot}$ an estimated quantity. We rewrite Eq. (7) using the geometric structure outlined in Section 3.2 as

$$\dot{\bar{g}}(t) = \bar{g}(t)(\lambda(\bar{g}(t), \boldsymbol{u}(t)) + B\bar{\boldsymbol{\xi}}(t)), \quad \bar{g}(t_0) = \bar{g}_0, \tag{18}$$

where $\bar{g}(t) = (g(t), V(t)) \in \bar{G}$ is the state, $\boldsymbol{u}(t) \in \mathbb{R}^2$ is the input, $\bar{\boldsymbol{\xi}}(t) = (\bar{\xi}_\omega(t), \bar{\xi}_v(t))^T$ is a Gaussian white noise with zero mean and unit variance, $B : \mathbb{R}^2 \to \bar{\mathfrak{g}}$ is a linear map with matrix representation $B \in \mathbb{R}^{5\times 2}$

$$B = \begin{bmatrix} 0_{3\times 2} \\ B_2 \end{bmatrix}, B_2 = \begin{bmatrix} b_\omega & 0 \\ 0 & b_v \end{bmatrix} \tag{19}$$

such that $\boldsymbol{\xi}(t) = B_2\bar{\boldsymbol{\xi}}(t)$ is a zero mean white Gaussian noise with variance $\Sigma = B^TB$, and $\lambda : G \times \mathbb{R}^2 \to \bar{\mathfrak{g}}$ is the left-trivialized dynamics (21). From now on, we often drop the explicit dependence on time of the signals from our notation. The function λ can be rewritten splitting it into its SE(2) and $\mathbb{R}^2$ components obtaining

$$g^{-1}\dot{g} = \lambda_g{}^\wedge, \quad \dot{V} = \lambda_V + \begin{bmatrix} \xi_\omega \\ \xi_v \end{bmatrix} \tag{20}$$

where $\lambda = (\lambda_g{}^\wedge, \lambda_V) \in \bar{\mathfrak{g}}$. In particular, the expressions for λ_g and λ_V are

$$\lambda_g(\bar{g}, u) = \begin{bmatrix} \omega \\ v \\ 0 \end{bmatrix}, \quad \lambda_V(\bar{g}, u) = \begin{bmatrix} -\frac{ma}{J+ma^2}\omega v + \frac{1}{J+ma^2}u_\omega \\ a\omega^2 + \frac{1}{m}u_v \end{bmatrix}. \tag{21}$$

The measurement Eq. (9) can be then rewritten as

$$\boldsymbol{y}(t_k) = h(\bar{g}(t_k)) + D\bar{\boldsymbol{\varepsilon}}(t_k) \tag{22}$$

where $h : \bar{G} \to \mathbb{R}^6$ is the output map (10), $\bar{\boldsymbol{\varepsilon}} \in \mathbb{R}^6$ is the measurement error and $D : \mathbb{R}^6 \to \mathbb{R}^6$ is an invertible linear map with the property that $\boldsymbol{\varepsilon} = D\bar{\boldsymbol{\varepsilon}}$ (i.e., $\Lambda = D^TD$, $\bar{\boldsymbol{\varepsilon}}$ is a Gaussian white noise with zero mean and unit variance).

Given the input $\boldsymbol{u}$ and the measurement output $\boldsymbol{y}$, the second-order-optimal filter presented in [8] gives the best estimate of the state $\bar{g}(\cdot)$ which minimizes the cost functional

$$J(\boldsymbol{\xi}, \boldsymbol{\varepsilon}, \bar{g}_0; t, t_0) := m(\bar{g}_0, t, t_0) + \int_{t_0}^{t} l(\boldsymbol{\xi}(\tau), \boldsymbol{\varepsilon}(\tau), t, \tau)\mathrm{d}\tau, \tag{23}$$

where $l : \mathbb{R}^5 \times \mathbb{R}^6 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the incremental cost defined by

$$l(\boldsymbol{\xi}, \boldsymbol{\varepsilon}, t, \tau) := \frac{1}{2}e^{-\alpha(t-\tau)}(\mathcal{R}(\boldsymbol{\xi}) + \mathcal{Q}(\boldsymbol{\varepsilon})), \tag{24}$$

with $\alpha \in \mathbb{R}$, $\alpha \geq 0$ the forgetting factor, and

$$\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}, \quad \boldsymbol{\xi} \mapsto \boldsymbol{\xi}^T R \boldsymbol{\xi}$$
$$\mathcal{Q} : \mathbb{R}^6 \to \mathbb{R}, \quad \boldsymbol{\varepsilon} \mapsto \boldsymbol{\varepsilon}^T Q \boldsymbol{\varepsilon}$$

two quadratic forms with matrix representations $R = \mathrm{diag}\,(r_\omega, r_v)$ and $Q = \mathrm{diag}\,(q_1, \ldots, q_6)$. The function $m : \bar{G} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the initial cost

$$m(\bar{g}_0, t, t_0) := \frac{1}{2}e^{-\alpha(t-t_0)}m_0(\bar{g}_0) \tag{25}$$

with known initial data

$$m_0(\bar{g}) = \frac{1}{2}\|I - \bar{g}^{-1}(t)\bar{g}_0\|_F^2 \tag{26}$$

where $\|\cdot\|_F^2$ stands for the Frobenius norm.

In the simulations we will consider $\alpha = 0$, $R = I_{2\times 2}$, $B_2 = \mathrm{diag}\,(b_\omega, b_v)$, $D = \mathrm{diag}\,(d_1, \ldots, d_6)$, $Q = I_{6\times 6}$. We define the weighted output error $\widetilde{\boldsymbol{y}}$ as

$$\widetilde{\boldsymbol{y}} = \left[\mathrm{diag}\left(\frac{q_1}{d_1^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_5}{d_5^2}, \frac{q_6}{d_6^2}\right)\right](\boldsymbol{y} - \widehat{\boldsymbol{y}}) \in \mathbb{R}^6. \tag{27}$$

To design the filter it is necessary to choose an affine connection on the state space. A left-invariant affine connection ∇ on $\bar{G}$ is characterized by its bilinear connection function $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ through the identity $\nabla_{\bar{g}X}(\bar{g}Y) = \bar{g}\omega(X, Y)$ for all $X, Y \in \bar{\mathfrak{g}}$ (see, e.g., [17,18]). Thus $\nabla_X Y \in \bar{\mathfrak{g}}$, and, since this map is $\mathbb{R}$-linear, it is a multiplication in $\bar{\mathfrak{g}}$. We choose skew-symmetric connection function of the form $\nabla_X Y = \lambda[X, Y]$, $\lambda \in \mathbb{R}$ (see e.g. [19,20]). The choices $\lambda = 0, \frac{1}{2}, 1$ define the $(-)$, (0), $(+)$ connections, that have negative, null and positive torsion, respectively. In this section we use the Cartan–Schouten (0)-connection given by $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$ [7].

The following proposition is an extension of the theorems in [8,11] to the nonholonomic case, and represents the second-order-optimal filter tailored for the Chaplygin sleigh case. The operators and functions are given with respect to the Lie group $\mathbb{R}^5$ instead of $\bar{\mathfrak{g}}$.

Proposition 1. *The second-order-optimal filter for the nonholonomic dynamic system (6) with measurement Eq. (10) and with respect to the cost functional (23)–(25) is given by*

$$\widehat{g}^{-1}\dot{\widehat{g}} = (\widehat{\omega}, \widehat{v}, 0)^\wedge + (K_{11}r^g + K_{12}r^V)^\wedge \tag{28}$$

$$\dot{\widehat{V}} = \begin{bmatrix} -\frac{ma}{J+ma^2}\widehat{\omega}\widehat{v} + \frac{u_\omega}{J+ma^2} \\ a\widehat{\omega}^2 + \frac{u_v}{m} \end{bmatrix} + (K_{21}r^g + K_{22}r^V) \tag{29}$$

where the residual r is

$$r = \begin{bmatrix} r^g \\ r^V \end{bmatrix}^T = \begin{bmatrix} -(\widetilde{y}_1 - \widetilde{y}_3)\ell\,\sin\widehat{\theta} + (\widetilde{y}_2 - \widetilde{y}_4)\ell\,\cos\widehat{\theta} \\ (\widetilde{y}_1 + \widetilde{y}_3)\cos\widehat{\theta} + (\widetilde{y}_2 + \widetilde{y}_4)\sin\widehat{\theta} \\ -(\widetilde{y}_1 + \widetilde{y}_3)\sin\widehat{\theta} + (\widetilde{y}_2 + \widetilde{y}_4)\cos\widehat{\theta} \\ \widetilde{y}_5 \\ \widetilde{y}_6 \end{bmatrix}^T. \tag{30}$$

The optimal gain $K = (K_{11}, K_{12}; K_{21}, K_{22}) : (\mathbb{R}^5)^* \to \mathbb{R}^5$ (with $K_{11} \in \mathbb{R}^{3\times 3}$, $K_{12} \in \mathbb{R}^{3\times 2}$, $K_{21} \in \mathbb{R}^{2\times 3}$ and $K_{22} \in \mathbb{R}^{2\times 2}$) is the solution of the perturbed matrix Riccati differential equation

$$\begin{aligned} \dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K, r)K - KW(K, r)^T \end{aligned} \tag{31}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \widehat{\omega} & 0 & 1 \\ \widehat{v} & -\widehat{\omega} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{ma}{J+ma^2}\widehat{v} & -\frac{ma}{J+ma^2}\widehat{\omega} \\ 0 & 0 & 0 & 2a\widehat{\omega} & 0 \end{bmatrix}, \tag{32}$$

$$E = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0 & 0 \\ a_{2,1} & 2 & 0 & 0 & 0 \\ a_{3,1} & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{33}$$

with

$$
\begin{aligned}
a_{1,1} =&(\widetilde{y}_1-\widetilde{y}_3)\ell\,\cos\widehat{\theta}+(\widetilde{y}_2-\widetilde{y}_4)\ell\,\sin\widehat{\theta}+2\ell^2,\\
a_{1,2} =&-\frac{1}{2}(\widetilde{y}_2+\widetilde{y}_4)\cos\widehat{\theta}+\frac{1}{2}(\widetilde{y}_1+\widetilde{y}_3)\sin\widehat{\theta},\\
a_{1,3} =&+\frac{1}{2}(\widetilde{y}_2+\widetilde{y}_4)\sin\widehat{\theta}+\frac{1}{2}(\widetilde{y}_1+\widetilde{y}_3)\cos\widehat{\theta},\\
a_{2,1} =&+\frac{1}{2}(\widetilde{y}_2+\widetilde{y}_4)\cos\widehat{\theta}-\frac{1}{2}(\widetilde{y}_1+\widetilde{y}_3)\sin\widehat{\theta},\\
a_{3,1} =&-\frac{1}{2}(\widetilde{y}_2+\widetilde{y}_4)\sin\widehat{\theta}-\frac{1}{2}(\widetilde{y}_1+\widetilde{y}_3)\cos\widehat{\theta},
\end{aligned}
$$

and

$$
W(K,r)=\begin{bmatrix}\frac{1}{2}\mathrm{ad}_{(K_{11}r^g+K_{12}r^V)^\wedge} & 0_{3\times2}\\ 0_{2\times3} & 0_{2\times2}\end{bmatrix}. \tag{34}
$$

The initial conditions for the Eqs. (28)–(29) and (31) are given by

$$
\widehat{\bar{g}}(t_0)=\bar{g}(t_0) \tag{35}
$$

$$
K(t_0)=I_{5\times5}. \tag{36}
$$

Proof. The proof is split into a few sections for the sake of readability. In what follows we will use $\eta^{\boldsymbol{x}}=(\eta^\theta,\eta^x,\eta^y,\eta^\omega,\eta^v)^T\in\mathbb{R}^5$ to indicate the vector form of an element of the Lie algebra $\bar{\mathfrak{g}}$ and with $T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}})=(\theta',x',y',\omega',v')^T$ the vector form of an element of the tangent space $T_{\widehat{\bar{g}}}\bar{G}$.

4.1. Computation of r

According to [8,11], the residual $r\in(\mathbb{R}^5)^*$ is given by

$$
r(\widehat{\bar{g}})=T_eL^*_{\widehat{\bar{g}}}\left[\left((D^{-1})^*\circ Q\circ D^{-1}(\boldsymbol{y}-h(\widehat{\bar{g}}))\right)\circ \mathrm{d}h(\widehat{\bar{g}})\right]. \tag{37}
$$

This term considers the difference between the real measurements and the estimated ones. Through this operator the estimation errors, that belong to $\mathbb{R}^6$, are mapped onto the dual of the Lie algebra $\bar{\mathfrak{g}}$, that is isomorphic to $\mathbb{R}^5$. From the definition of the matrices D and Q it follows that

$$
(D^{-1})^*\circ Q\circ D^{-1}(\boldsymbol{y}-h(\widehat{\bar{g}}))=(\boldsymbol{y}-\widehat{\boldsymbol{y}})^T. \tag{38}
$$

Given $T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}})\in\mathbb{R}^5$, the differential of h in $\widehat{\bar{g}}$ applied to $T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}})$ is

$$
\begin{aligned}
&\mathrm{d}h(\widehat{\bar{g}})(T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}}))\\
&=\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\widehat{x}(s)+\ell\,\cos\widehat{\theta}(s)\\ \widehat{y}(s)+\ell\,\sin\widehat{\theta}(s)\\ \widehat{x}(s)-\ell\,\cos\widehat{\theta}(s)\\ \widehat{y}(s)-\ell\,\sin\widehat{\theta}(s)\\ \widehat{\omega}(s)\\ \widehat{v}(s)\end{bmatrix}=\begin{bmatrix}\widehat{x}'-\ell\,\widehat{\theta}'\sin\widehat{\theta}\\ \widehat{y}'+\ell\,\widehat{\theta}'\cos\widehat{\theta}\\ \widehat{x}'+\ell\,\widehat{\theta}'\sin\widehat{\theta}\\ \widehat{y}'-\ell\,\widehat{\theta}'\cos\widehat{\theta}\\ \widehat{\omega}'\\ \widehat{v}'\end{bmatrix},
\end{aligned} \tag{39}
$$

and we can write the operator $\mathrm{d}h(\widehat{\bar{g}})$ as

$$
\mathrm{d}h(\widehat{\bar{g}})=\begin{bmatrix}-\ell\,\sin\widehat{\theta} & 1 & 0 & 0 & 0\\ +\ell\,\cos\widehat{\theta} & 0 & 1 & 0 & 0\\ +\ell\,\sin\widehat{\theta} & 1 & 0 & 0 & 0\\ -\ell\,\cos\widehat{\theta} & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}. \tag{40}
$$

Using (40) and (38) we obtain

$$
\begin{aligned}
&\left((D^{-1})^*\circ Q\circ D^{-1}(\boldsymbol{y}-h(\widehat{\bar{g}}))\right)\circ \mathrm{d}h(\widehat{\bar{g}})\\
&=\begin{bmatrix}-(\widetilde{y}_1-\widetilde{y}_3)\ell\,\sin\widehat{\theta}+(\widetilde{y}_2-\widetilde{y}_4)\ell\,\cos\widehat{\theta}\\ \widetilde{y}_1+\widetilde{y}_3\\ \widetilde{y}_2+\widetilde{y}_4\\ \widetilde{y}_5\\ \widetilde{y}_6\end{bmatrix}^T.
\end{aligned} \tag{41}
$$

Evaluating $T_eL^*_{\widehat{\bar{g}}}$ on (41) we finally get (30).

4.2. Computation of A

The expression for the operator $A:\mathbb{R}^5\to\mathbb{R}^5$ is

$$
A=\mathrm{d}_1\lambda(\widehat{\bar{g}},u)\circ T_eL_{\widehat{\bar{g}}}-\mathrm{ad}_{\lambda(\widehat{\bar{g}},u)}-T_{\lambda(\widehat{\bar{g}},u)}. \tag{42}
$$

The operator A represents the coefficient of the linear part of the Riccati equation (31). The first term $\mathrm{d}_1\lambda(\widehat{\bar{g}},u)\circ T_eL_{\widehat{\bar{g}}}$ is the differential of the left-trivialized dynamics with respect to the group elements, $\mathrm{ad}_{\lambda(\widehat{\bar{g}},u)}$ is the adjoint operator with respect to λ, the last term $T_{\lambda(\widehat{\bar{g}},u)}$ is the torsion and takes care of the choice of the connection function adopted.

Given $T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}})\in\mathbb{R}^5$, the differential of λ is

$$
\begin{aligned}
\mathrm{d}_1\lambda(\widehat{\bar{g}},u)(T_eL_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}}))&=\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\lambda_g(s)\\ \lambda_V(s)\end{bmatrix}\\
&=\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\widehat{\omega}(s)\\ \widehat{v}(s)\\ 0\\ -\frac{ma\widehat{\omega}(s)\widehat{v}(s)}{J+ma^2}+\frac{u_\omega}{J+ma^2}\\ a\widehat{\omega}^2(s)+\frac{u_v}{m}\end{bmatrix}=\begin{bmatrix}\widehat{\omega}'\\ \widehat{v}'\\ 0\\ -\frac{ma(\widehat{\omega}\widehat{v}'+\widehat{v}\widehat{\omega}')}{J+ma^2}\\ 2a\widehat{\omega}\widehat{\omega}'\end{bmatrix}
\end{aligned} \tag{43}
$$

and thus

$$
\mathrm{d}_1\lambda(\widehat{\bar{g}},u)\circ T_eL_{\widehat{\bar{g}}}=\begin{bmatrix}0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -\frac{ma\widehat{v}}{J+ma^2} & -\frac{ma\widehat{\omega}}{J+ma^2}\\ 0 & 0 & 0 & 2a\widehat{\omega} & 0\end{bmatrix}. \tag{44}
$$

The adjoint matrix representation (16) implies

$$
\mathrm{ad}_{\lambda(\widehat{\bar{g}},u)}=\begin{bmatrix}0 & 0 & 0 & 0_{1\times2}\\ 0 & 0 & -\widehat{\omega} & 0_{1\times2}\\ -\widehat{v} & \widehat{\omega} & 0 & 0_{1\times2}\\ 0_{2\times1} & 0_{2\times1} & 0_{2\times1} & 0_{2\times2}\end{bmatrix}. \tag{45}
$$

Since we consider the Cartan–Schouten (0)-connection form $\omega^{(0)}=\frac{1}{2}\mathrm{ad}$, the torsion function $T_{\lambda(\widehat{\bar{g}},u)}$ vanishes (see [21]), thus, in matrix form it is given by

$$
T_{\lambda(\widehat{\bar{g}},u)}=\begin{bmatrix}0_{3\times3} & 0_{3\times2}\\ 0_{2\times3} & 0_{2\times2}\end{bmatrix}. \tag{46}
$$

Using (44), (45) and (46) we obtain (32).

4.3. Computation of E

The operator $E:\mathbb{R}^5\to(\mathbb{R}^5)^*$ takes the form

$$
\begin{aligned}
E=-T_eL^*_{\widehat{\bar{g}}}\circ\Big[&\left((D^{-1})^*\circ Q\circ D^{-1}(\boldsymbol{y}-h(\widehat{\bar{g}}))\right)^{T_{\widehat{\bar{g}}}\bar{g}}\\
&\circ\mathrm{Hess}h(\widehat{\bar{g}})-(\mathrm{d}h(\widehat{\bar{g}}))^*\circ(D^{-1})^*\\
&\circ Q\circ D^{-1}\circ\mathrm{d}h(\widehat{\bar{g}})\Big]\circ T_eL_{\widehat{\bar{g}}},
\end{aligned} \tag{47}
$$

and represents the second-order term of the Riccati equation (31). It is worth highlighting that E does not depend on the choice of the connection function. From (40) and the definitions of the matrices D and Q we can find the composition

$$\begin{aligned}&(\mathrm{d}h(\widehat{\bar{g}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h(\widehat{\bar{g}}) \\ &= \mathrm{diag}\left(2\ell^2, 2, 2, 1, 1\right).\end{aligned} \tag{48}$$

The function

$$\left((D^{-1})^* \circ Q \circ D^{-1}(\boldsymbol{y} - h(\widehat{\bar{g}}))\right) : \mathbb{R}^6 \to (\mathbb{R}^6)^*$$

in lifted through the exponential functor $(\cdot)^{T_{\widehat{\bar{g}}}\bar{G}}$ to the linear map

$$\left((D^{-1})^* \circ Q \circ D^{-1}(\boldsymbol{y} - h(\widehat{\bar{g}}))\right)^{T_{\widehat{\bar{g}}}\bar{G}} : \mathcal{L}(T_{\widehat{\bar{g}}}\bar{g}, \mathbb{R}^6) \to \mathcal{L}(T_{\widehat{\bar{g}}}\bar{g}, (\mathbb{R}^6)^*)$$

defined as

$$\left((D^{-1})^* \circ Q \circ D^{-1}(\boldsymbol{y} - h(\widehat{\bar{g}}))\right)^{T_{\widehat{\bar{g}}}(\xi)} = \left((D^{-1})^* \circ Q \circ D^{-1}(\boldsymbol{y} - h(\widehat{\bar{g}}))\right) \circ \xi.$$

Let $T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1}) = (\widehat{\theta}_1', \widehat{x}_1', \widehat{y}_1', \widehat{\omega}_1', \widehat{v}_1')^T$, $T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2}) = (\widehat{\theta}_2', \widehat{x}_2', \widehat{y}_2', \widehat{\omega}_2', \widehat{v}_2')^T \in T_{\widehat{\bar{g}}}\bar{G}$ be two vector fields, then the Hessian matrix is defined by

$$\begin{aligned}&\mathrm{Hess}\, h(\widehat{\bar{g}})(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1}))(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2})) = \\ &\qquad \mathrm{d}(\mathrm{d}h(\widehat{\bar{g}})(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2})))(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1})) \\ &\qquad - \mathrm{d}h(\widehat{\bar{g}})(\nabla_{T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1})}(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2}))),\end{aligned} \tag{49}$$

from the choice of Cartan–Schouten (0)-connection, we get

$$\nabla_{T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_1}}(T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_2}) = \frac{1}{2} T_e L_{\widehat{\bar{g}}}(\mathrm{ad}_{\eta^{\boldsymbol{x}_1}}\eta^{\boldsymbol{x}_2}). \tag{50}$$

The Hessian evaluated in $T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_1}$ and $T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_2}$ is therefore

$$\begin{aligned}&\mathrm{Hess}h(\widehat{\bar{g}})(T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_1})(T_e L_{\widehat{\bar{g}}}\eta^{\boldsymbol{x}_2}) \\ &\qquad = \begin{bmatrix} -\widehat{\theta}_1'\widehat{\theta}_2'\ell\, \cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_1'\widehat{\theta}_2'\ell\, \sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\, \cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\, \sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \end{bmatrix}.\end{aligned} \tag{51}$$

From (38) and (51) it follows that

$$\begin{aligned}&\left((D^{-1})^* \circ Q \circ D^{-1}(\boldsymbol{y} - h(\widehat{\bar{g}}))\right)^{T_{\widehat{\bar{g}}\bar{g}}} \circ \mathrm{Hess}h(\widehat{\bar{g}}) = \\ &\begin{bmatrix} a_{1,1} & -\frac{1}{2}(\widetilde{y}_2 + \widetilde{y}_4) & \frac{1}{2}(\widetilde{y}_1 + \widetilde{y}_3) & 0_{1\times 2} \\ \frac{1}{2}(\widetilde{y}_2 + \widetilde{y}_4) & 0 & 0 & 0_{1\times 2} \\ -\frac{1}{2}(\widetilde{y}_1 + \widetilde{y}_3) & 0 & 0 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix},\end{aligned} \tag{52}$$

$$a_{1,1} = -\widetilde{y}_1\ell\, \cos\widehat{\theta} - \widetilde{y}_2\ell\, \sin\widehat{\theta} + \widetilde{y}_3\ell\, \cos\widehat{\theta} + \widetilde{y}_4\ell\, \sin\widehat{\theta}.$$

In conclusion, combining (48) and (52) with $T_e L_{\widehat{\bar{g}}}^*$ and $T_e L_{\widehat{\bar{g}}}$, the matrix (33) is obtained.

4.4. Computation of W

The operator W represents the second-order term of the Riccati equation (31) that depends on the choice of the connection function. From the adjoint matrix form (16) and the Cartan–Schouten (0)-connection (see, e.g., [6,7]), we have

$$\begin{aligned}W(K, r) &= \frac{1}{2}\mathrm{ad}_{((K_{11}r^g + K_{12}r^V)^\wedge,\ (K_{21}r^g + K_{22}r^V)^\wedge)} \\ &= \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^V)^\wedge} & 0_{3\times 2} \\ 0_{2\times 3} & 0_{2\times 2} \end{bmatrix}.\end{aligned} \tag{53}$$

4.5. Initial condition

The initial condition for the filter is given by

$$\widehat{\bar{g}}_0 = \arg\min_{\bar{g}\in\bar{G}} m_0(\bar{g}), \tag{54}$$

while the initial condition for the gain is $K(t_0) = X_0^{-1}$ where the operators $X_0 : \bar{\mathfrak{g}} \to \bar{\mathfrak{g}}^*$ satisfies

$$X_0 = T_e L_{\widehat{\bar{g}}_0}^* \circ \mathrm{Hess}\, m_0(\widehat{\bar{g}}) \circ T_e L_{\widehat{\bar{g}}_0}. \tag{55}$$

From the definition of m_0 in (26) it is easy to see that $\widehat{\bar{g}}(t_0) = \bar{g}(t_0)$. Since $\mathrm{d}\, m_0(\widehat{\bar{g}})(\nabla_{T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1})}(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2}))) = 0_{5\times 5}$ we have that

$$\begin{aligned}&\mathrm{Hess}\, m_0(\widehat{\bar{g}})(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1}))(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2})) = \\ &\qquad \mathrm{d}(\mathrm{d}\, m_0(\widehat{\bar{g}})(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_2})))(T_e L_{\widehat{\bar{g}}}(\eta^{\boldsymbol{x}_1})).\end{aligned} \tag{56}$$

Thus, the Hessian and the X_0 operator are the identity matrix, and so is $K(t_0)$.

This computation ends the proof. □

Remark 2. If one considers the case where only GPS measurements (the first four lines in (10)) are available, then the residual r and the matrix E become:

$$r = \begin{bmatrix} -(\widetilde{y}_1 - \widetilde{y}_3)\ell\, \sin\widehat{\theta} + (\widetilde{y}_2 - \widetilde{y}_4)\ell\, \cos\widehat{\theta} \\ (\widetilde{y}_1 + \widetilde{y}_3)\cos\widehat{\theta} + (\widetilde{y}_2 + \widetilde{y}_4)\sin\widehat{\theta} \\ -(\widetilde{y}_1 + \widetilde{y}_3)\sin\widehat{\theta} + (\widetilde{y}_2 + \widetilde{y}_4)\cos\widehat{\theta} \\ 0_{2\times 1} \end{bmatrix}^T, \tag{57}$$

$$E = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0_{1\times 2} \\ a_{2,1} & 2 & 0 & 0_{1\times 2} \\ a_{3,1} & 0 & 2 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix}. \tag{58}$$

The state is still observable using a computation similar to [22]. ◇

5. The second-order-optimal filter with nonholonomic constraint

The aim of this section is to investigate under which conditions the filter preserves the nonholonomic constraint (1), that the formulation in Section 4 cannot guarantee. The preservation of the nonholonomic constraint produces more feasible trajectories that will be more suitable for control implementations.

We first observe that, if the dynamics does not present model errors, the nonholonomic information is enclosed by the left trivialized dynamics, in particular by its "group" part

$$\lambda_g(\bar{g}, u) = \begin{bmatrix} \omega \\ v \\ 0 \end{bmatrix}, \tag{59}$$

where the last entry is equal to 0, because it represents $v_\perp$, that is the orthogonal projection of the linear velocity with respect to the first axis direction. In order to preserve the constraint, it is necessary to "force" the filter to keep such value equal to 0 in the right-hand side of the Eqs. (28)–(29). It is not possible to act directly on the filter parameters in order to impose the third component of the residual (30) be always 0, thus, it is necessary to operate on the gain operator K structure whose dynamics is governed by the Riccati equation (31).

Whenever the third row and column of the operator K in (31) are 0, the product KEK has the third row and column equal to 0, whatever the components of E. Moreover, since the system has only model errors that apply to the dynamics that model the evolution of the velocity and not to the kinematics, the

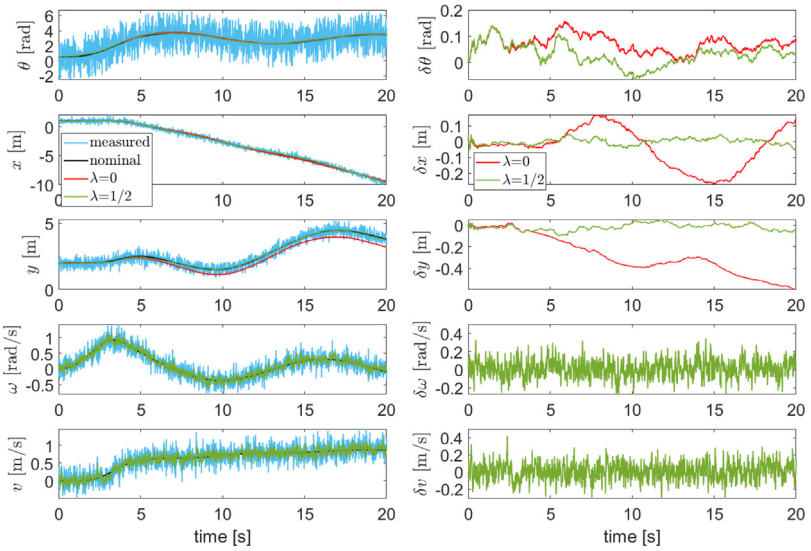

Fig. 3. Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with GPS and INS in the case $T_s = 10$ ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

term $BR^{-1}B^T$ has non-null components only on the last 2×2 submatrix. The analysis can be therefore limited to the study of the operators A and W and on the choice of the connection function. Continuing working with skew-symmetric connection functions (see e.g. [19,20]), a choice of the connection could be $\nabla_X Y = \lambda[X, Y]$ with $\lambda \in \mathbb{R}$. A generic version of the operator A provided in (32) suggests, for our purposes, to set $\lambda = 0$, which corresponds to the Cartan Schouten (-)-connection $\omega^{(-)} = 0$ and produces

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{ma}{J+ma^2}\widehat{v} & -\frac{ma}{J+ma^2}\widehat{\omega} \\ 0 & 0 & 0 & 2a\widehat{\omega} & 0 \end{bmatrix}. \tag{60}$$

With this choice, the operator W is represented by

$$W = 0_{5\times 5}, \tag{61}$$

the null 5×5 matrix, and then the products $W(K, r)K$ and $K^T W(K, r)$ have the third rows and columns equal to 0. The new matrix representation for the operator E becomes:

$$E = \begin{bmatrix} a_{1,1} & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{62}$$

with

$$a_{1,1} = (\widetilde{y}_1 - \widetilde{y}_3)\ell \cos\widehat{\theta} + (\widetilde{y}_2 - \widetilde{y}_4)\ell \sin\widehat{\theta} + 2\ell^2.$$

The preservation of the nonholonomic constraint is then granted by choosing the initial conditions of the operator K as

$$K(t_0) = \begin{bmatrix} 1_{2\times 2} & 0_{2\times 1} & 0_{2\times 2} \\ 0_{1\times 2} & 0_{1\times 1} & 0_{1\times 2} \\ 0_{2\times 2} & 0_{2\times 1} & 1_{2\times 2} \end{bmatrix}. \tag{63}$$

6. Simulations

In this section we show how the second-order-optimal filter applied to the Chaplygin sleigh works in simulated scenarios. We set $J = 6.125\ \mathrm{kg\,m^2}$, $m = 125$ kg, $a = 0.3$ m and $\ell = 0.2$ m. We consider the control inputs $u_\omega(t) = 10\sin((1/2)t)$ and $u_v(t) = (1/2)\sin((1/5)t)$. The total simulation time is $T = 20$ s. The initial conditions are $\theta(0) = \pi/6$ rad, $x(0) = 1$ m, $y(0) = 2$ m, $\omega(0) = 0$ rad/s and $v(0) = 0$ m/s. For the matrix D we choose $d_1 = d_2 = d_3 = d_4 = 0.4$, $d_5 = d_6 = 0.2$, while for the matrix B_2 we choose $b_\omega = b_v = 10$. To solve all the differential equations we used a Runge–Kutta 4th order method with steps $T_s = 10$ ms and $T_s = 1$ ms.

Figs. 3 and 4 show how the filters work with a sample time of 10 ms in the cases of Cartan–Schouten (-)-connection ($\lambda = 0$) and Cartan–Schouten (0)-connection ($\lambda = 1/2$) using both GPS and INS devices and only GPS. The estimation errors are calculated with respect the nominal values

$$\delta\theta \triangleq \theta - \widehat{\theta}, \quad \delta x \triangleq x - \widehat{x}, \quad \dots.$$

Figs. 5 and 6 show the same trajectories and errors in the case of a sample time of 1 ms. Tables 1 and 2 compare the mean, standard deviation and root mean square values of the errors obtained by applying the second-order filter in the GPS + INS and GPS cases, with both connection functions, using $T_s = 10$ ms and $T_s = 1$ ms, respectively.

The filter designed on the Chaplygin sleigh using the Cartan–Schouten (0)-connection provides a good estimation both for the pose and the velocity, even if the two antennas do not measure directly the position of the contact point of the blade nor the orientation. The filter designed using the Cartan–Schouten (-)-connection has less accurate estimations but preserves the nonholonomic constraint. With no INS measurements, the accuracy of the velocities for the Cartan–Schouten (-)-connection gets worse, but it allows to gain precision on the positions. Instead, for the Cartan–Schouten (0)-connection the addition of the INS measurements improves all the estimations. In both cases, a finer sampling of measurements greatly improves the accuracy of the estimates as can be seen by comparing the $T_s = 10$ ms and $T_s = 1$ ms cases.

7. Conclusions

In this paper we designed a second-order-optimal filter for the Chaplygin sleigh, which is a rigid body with a nonholonomic

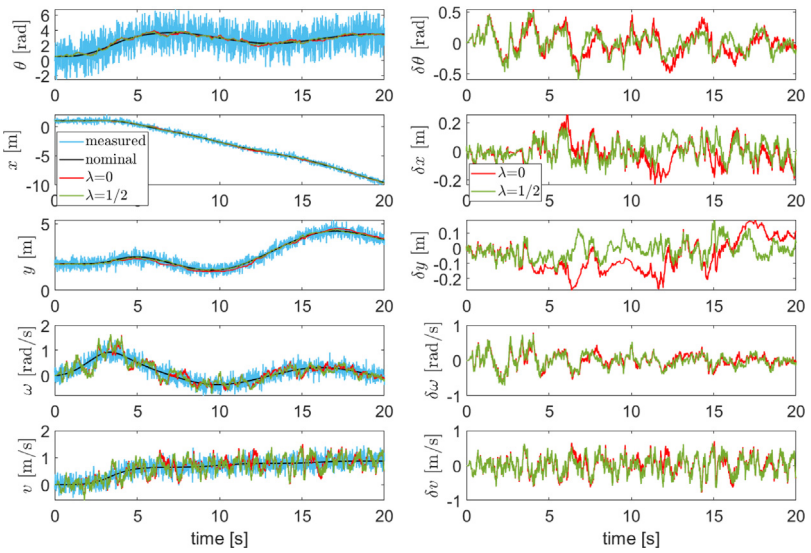

Fig. 4. Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with only GPS in the case $T_s = 10$ ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

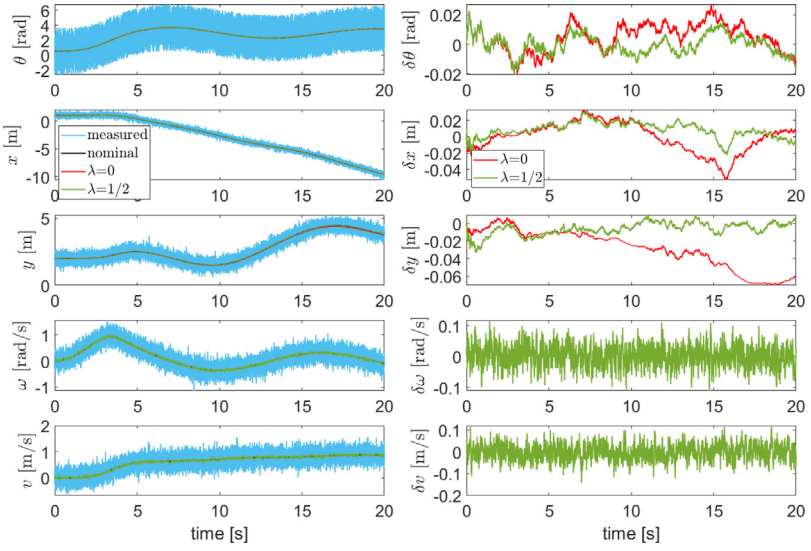

Fig. 5. Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with GPS and INS in the case $T_s = 1$ ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Mean, standard deviation and root mean square values of the errors for $T_s = 10$ ms for the cases $\lambda = 1/2$ and $\lambda = 0$ (in bracket).

	GPS + INS			GPS		
	μ	σ	rms	μ	σ	rms
$\delta\theta$ [mrad]	28.7	41.1	50.1	−2.1	180.5	180.4
	(71.1)	(32.7)	(78.2)	(−6.8)	(194.7)	(194.8)
δx [mm]	−0.3	19.2	19.2	4.9	66.8	67.0
	(−37.8)	(122.6)	(128.3)	(−13.8)	(77.9)	(79.1)
δy [mm]	−20.5	30.9	37.1	−11.6	59.3	60.4
	(−281.0)	(177.6)	(332.3)	(−51.9)	(103.2)	(115.5)
$\delta\omega$ [mm/rad]	12.3	94.2	95.0	−2.5	208.8	208.8
	(12.3)	(94.2)	(95.0)	(4.4)	(203.8)	(203.8)
δv [mm/rad]	4.9	99.4	99.5	0.0	249.5	249.4
	(4.8)	(99.4)	(99.4)	(15.4)	(251.9)	(252.3)

constraint. This filter exploits the theory of Lie groups and provides good estimations for both pose and velocity. We studied the conditions for which the filter preserves the nonholonomic constraint and the effects of the choice of different connection

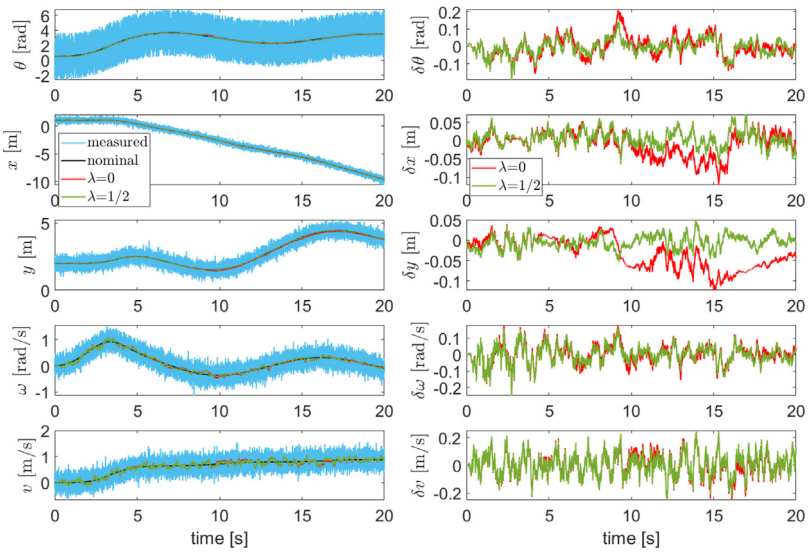

Fig. 6. Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with only GPS in the case $T_s = 1$ ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Mean, standard deviation and root mean square values of the errors for $T_s = 1$ ms for the cases $\lambda = 1/2$ and $\lambda = 0$ (in bracket).

	GPS + INS			GPS		
	μ	σ	rms	μ	σ	rms
$\delta\theta$ [mrad]	−0.1	6.6	6.6	−6.1	44.4	44.8
	(4.9)	(9.2)	(10.4)	(−4.0)	(57.0)	(57.1)
δx [mm]	8.3	9.5	12.6	5.5	20.9	21.7
	(−0.4)	(18.0)	(18.0)	(−8.8)	(31.9)	(33.1)
δy [mm]	−6.1	7.3	9.5	−5.1	16.3	17.1
	(−28.6)	(22.0)	(36.1)	(−33.5)	(36.6)	(49.6)
$\delta\omega$ [mm/rad]	1.5	31.2	31.2	−1.1	54.2	54.2
	(1.5)	(31.2)	(31.2)	(−0.4)	(57.1)	(57.1)
δv [mm/rad]	−1.8	33.9	33.9	2.8	77.1	77.2
	(−1.9)	(33.9)	(33.9)	(1.0)	(80.0)	(80.0)

functions. To design the filter it is of particular importance the left trivialization map, which arises in the study of Lie groups dynamics and includes the information of the nonholonomic constraint. In the future we will study a generic formulation of the filter for nonholonomic systems that behave as the Chaplygin sleigh, for example nonholonomic systems with symmetries that satisfy the so-called vertical symmetry condition [23,24] and we will tackle the convergence of the filter by studying the stability of the differential equations for the state and the gain operator for $t \to \infty$.

CRediT authorship contribution statement

Damiano Rigo: Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Nicola Sansonetto:** Methodology, Writing – review and editing, Supervision. **Riccardo Muradore:** Conceptualization, Writing – review and editing, Supervision, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Damiano Rigo reports financial support was provided by "Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR)" (Project "Dipartimento di Eccellenza 2018-2022"). Nicola Sansonetto reports financial support was provided by INDAM ("Progetto Giovani Ricercatori-GNFM").

Data availability

No data was used for the research described in the article.

Acknowledgments

This work has been partially supported by the project "Dipartimento di Eccellenza 2018-2022" of the "Ministero dell'Istruzione, dell'Università e della Ricerca"' (MIUR). Nicola Sansonetto also thanks GNFM-INDAM for the support.

References

[1] R.E. Kalman, A new approach to linear filtering and prediction problems, ASME–J. Basic Eng. 82 (1) (1960) 35–45.
[2] B.D. Anderson, J.B. Moore, Optimal Filtering, Courier Corporation, 2012.
[3] S.J. Julier, J.K. Uhlmann, Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations, in: American Control Conference, Vol. 2, 2002, pp. 887–892.
[4] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, IEEE Trans. Signal Process. 50 (2) (2002) 174–188.
[5] R. Mortensen, Maximum-likelihood recursive nonlinear filtering, J. Optim. Theory Appl. 2 (6) (1968) 386–394.
[6] A. Saccon, A.P. Aguiar, J. Hauser, Lie group projection operator approach: Optimal control on TSO(3), in: 2011 50th IEEE Conference on Decision and Control and European Control Conference, IEEE, 2011, pp. 6973–6978.
[7] A. Saccon, J. Hauser, A.P. Aguiar, Optimal control on Lie groups: The projection operator approach, IEEE Trans. Automat. Control 58 (9) (2013) 2230–2245.

[8] A. Saccon, J. Trumpf, R. Mahony, A.P. Aguiar, Second-order-optimal minimum-energy filters on Lie groups, IEEE Trans. Automat. Control 61 (10) (2015) 2906–2919.
[9] J.M. Osborne, D.V. Zenkov, Steering the Chaplygin sleigh by a moving mass, in: IEEE Conference on Decision and Control, 2005, pp. 1114–1118.
[10] N. Sansonetto, M. Zoppello, On the trajectory generation of the hydrodynamic Chaplygin sleigh, IEEE Control Syst. Lett. 4 (4) (2020) 922–927.
[11] D. Rigo, C. Segala, N. Sansonetto, R. Muradore, Second-order-optimal filter on Lie groups for planar rigid bodies, IEEE Trans. Automat. Control 67 (9) (2022) 4971–4977.
[12] F. Bullo, A.D. Lewis, Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems, Vol. 49, Springer, 2019.
[13] V.S. Varadarajan, Lie Groups, Lie Algebras, and their Representations, Vol. 102, Springer Science & Business Media, 2013.
[14] J.E. Marsden, T.S. Ratiu, Introduction to mechanics and symmetry, Phys. Today 48 (12) (1995) 65.
[15] A.M. Bloch, Nonholonomic Mechanics and Control, Vol. 24, Springer, 2015.
[16] A.M. Bloch, J.E. Marsden, D.V. Zenkov, Quasivelocities and symmetries in non-holonomic systems, Dyn. Syst. 24 (2) (2009) 187–222.
[17] K. Nomizu, Invariant affine connections on homogeneous spaces, Amer. J. Math. 76 (1) (1954) 33–65.
[18] J.Q. Gallier, J. Quaintance, Differential Geometry and Lie Groups, Vol. 12, Springer, 2020.
[19] A. Cogliati, P. Mastrolia, Cartan, Schouten and the search for connection, Hist. Math. 45 (1) (2018) 39–74.
[20] M.M. Postnikov, Geometry VI: Riemannian Geometry, Vol. 91, Springer Science & Business Media, 2013.
[21] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, Vol. 2, Wiley and Sons, 1969.
[22] C. Spagnol, R. Muradore, M. Assom, A. Beghi, R. Frezza, Trajectory reconstruction by integration of GPS and a swarm of MEMS accelerometers: model and analysis of observability, in: IEEE Conference on Intelligent Transportation Systems, 2004, pp. 64–69.
[23] P. Balseiro, The Jacobiator of nonholonomic systems and the geometry of reduced nonholonomic brackets, Arch. Ration. Mech. Anal. 214 (2) (2014) 453–501.
[24] P. Balseiro, N. Sansonetto, First integrals and symmetries of nonholonomic systems, Arch. Ration. Mech. Anal. (2022) 1–47.