

VALINOR: a lightweight leg inertial odometry for humanoid robots

Arnaud Demont* , Mehdi Benallegue , and Abdelaziz Benallegue 

Abstract: This article describes the preparation procedure for publication in the International Journal of Control, Automation, and Systems (IJCAS), and this template applies both for initial submission and the final camera-ready manuscript of the paper. When authors submit their work for review, it is necessary to follow these instructions. The abstract should not exceed 300 words for regular papers or 75 words for technical notes and correspondence without equations, references, and footnotes.

Keywords: Legged robots, proprioceptive odometry, state estimation, tilt estimation.

1. INTRODUCTION

Humanoid robot navigation in real-world environments remains a core challenge in robotics. For robots to be deployed at scale in society and industry, whether in healthcare facilities or warehouses, they must operate robustly in dynamic, unstructured, and human-centric environments [1]. This requires strong guarantees on their ability to plan motion, and execute actions in real time. In comparison to other legged robots, the control of humanoids is particularly difficult due to their complex underactuated dynamics controlled through contact with the environment [2]. Their control must therefore be as theoretically grounded as possible, in order to ensure safe, predictable, and certifiable behavior [3].

In the meantime, methods for motion generation continue to gain in complexity and capability. Notably, Model Predictive Control (MPC) [4–6], reinforcement learning-based controllers [7, 8], and foundation models [9, 10] are becoming more prevalent due to their ability to handle long-horizon planning, complex dynamics, and to their versatility. However, these advances come at the cost of increased computational complexity, running them in real-time on resource-limited platforms like drones or embedded systems thus becomes challenging [11–14]. Since the control part is getting heavier, one practical solution is to reduce the computational load of other parts of the pipeline, for example the state estimation part.

With this direction of lightweight estimation, our work also aims at proving that the estimation accuracy can be improved, not necessarily by adding new sensors, but by focusing on improving the estimation of key variables, which impact the overall estimation. Our works thus leverage only the commonly readily available proprioceptive

sensors on the robot, namely an IMU, joint encoders, and force sensors at the contacts. Using solely proprioceptive sensors provides the observability of only the tilt of the robot [15]. The latter is essential for the robot's stabilization loop, especially for humanoid robots [16]. Beyond this, it directly impacts odometry performances, since even a slight error on the tilt estimate can lead to a non-negligible error on the translation estimate over long walks. We therefore show that we can improve proprioceptive odometry by exploiting a very accurate tilt estimate. This attitude estimation is made by a lightweight complementary filter, with mathematical guarantees on the estimation convergence, aligning with our goal of lightweight yet accurate and reliable state estimation.

To ensure a reliable control, the pose of the floating base must be estimated at very high frequency. To this end, one commonly performs inertial odometry **CITE** by integrating the signals provided by IMUs, composed of an accelerometer and a gyrometer, that respectively measure the angular velocity and the linear acceleration (including the gravitational acceleration) of the sensor in the world frame, expressed in the frame of the sensor. However, using only these sensors, the position, linear velocity and the yaw in the world frame are non-observable. They will therefore be subject to drift at low frequency due to the integration of the sensor noises. A widely adopted solution to deal with the drift inherent in the inertial odometry is to combine it with a legged odometry **Citer**, that uses the contacts that the robot successively creates with the environment as fixed points in the environment, constraining the kinematics **bloesch2013state**. Using the joint encoders, the legged odometry provides an estimate of the position at low frequency and prevents the drift. However,

A. Demont, M. Benallegue and A. Benallegue are with the CNRS-AIST JRL (Joint Robotics Laboratory), IRL, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8560 Japan.

A. Demont and A. Benallegue are also with Université Paris-Saclay, 3 rue Joliot Curie, Bâtiment Breguet, 91190 Gif-sur-Yvette, France, and Laboratoire d'Ingénierie des Systèmes de Versailles, 10-12 avenue de l'Europe, 78140 Vélizy, France.

e-mails: arnaud.demont@aist.go.jp, mehdi.benallegue@aist.go.jp, abdelaziz.benallegue@uvsq.fr.

* Corresponding author.

this method relies on the assumption that there is no slippage between the robot and the environment at the contact. In addition, it highly depends on the accuracy of the estimated pose of the contact at the time it is considered as fixed, and thus it depends on the compliance of the robot, the contact detection framework, etc. This combination of the inertial and legged odometry methods is the most common example of *dead reckoning* [ros-ton1991deadReckoning](#) in the literature, a term that more generally refers to any approach based on proprioceptive sensors. Dead reckoning is a satisfactory solution for tasks that don't require a perfect knowledge of the position and yaw in the world frame, for example for some tele-operation tasks and stabilization. The latter notably relies on the estimation of the tilt of the robot, which is partially observable using the accelerometer measurement. Work has therefore been done to improve this estimation. At first obtained by considering the linear acceleration negligible with respect to the gravitational acceleration [mahony2008nonlinear](#), it also took benefit from the legged odometry, the constraint on the contact allowing for a better dissociation between linear motions and the orientation [Rephrase?](#). Further works focused on [Continue](#) or with some enhanced methods [benallegue2017tilt](#). In our work, we propose to enrich our legged odometry with the orientation of contacts. These orientations are integrated to our attitude estimator to correct the estimated yaw of the robot at low frequency, similarly to their position. legged odometry using contact orientations: [ramadoss2022wholebodyKineLieKF](#)

However, knowing these absolute kinematics is mandatory for other tasks, notably when an autonomous behaviour of the robot is desired.

To make the absolute pose of the robot observable, exteroceptive sensors, like LiDARs, cameras, etc., can be added to the estimation. The information provided by these sensors is extremely valuable, but is available only at low frequency and require a high computation capability. Also, similarly to the legged odometry failing under the slippage of contacts or with bad contact estimations, these methods can provide outliers. Known reasons can be a low illumination of the environments, low-textured environments, occlusion of the sensors, etc. The legged odometry and exteroceptive measurements therefore have complementary advantages and cons, the challenge addressed by most of the recent estimation methods is to improve the fusion between both.

To cite: Invariant Smoother for Legged Robot State Estimation With Dynamic Contact Event Information (they ignore contacts with too high velocity?)

1.1. Motivation

State estimation remains a cornerstone of robust control, providing the essential feedback required for planning and control to operate effectively. Recognizing this

trade-off, we propose an approach focused on reducing the runtime demands of the state estimator, while preserving high estimation accuracy. This way, more processing time can be allocated to the increasingly demanding control algorithms without compromising the overall performance of the system.

With this direction of lightweight estimation, our work also aims at proving that the estimation accuracy can be improved, not necessarily by adding new sensors, but by focusing on improving the estimation of key variables, which impact the overall estimation. Our works thus leverages only the commonly readily available proprioceptive sensors on the robot, namely an IMU, joint encoders, and force sensors at the contacts. Using solely proprioceptive sensors provides the observability of only the tilt of the robot [\[15\]](#). The latter is essential for the robot's stabilization loop, especially for humanoid robots [\[16\]](#). Beyond this, it directly impacts odometry performances, since even a slight error on the tilt estimate can lead to a non-negligible error on the translation estimate over long walks, more noticeably along the vertical axis [Demont](#). We therefore show that we can improve proprioceptive odometry by exploiting a very accurate tilt estimate. This attitude estimation is made by a lightweight complementary filter, with mathematical guarantees on the estimation convergence, aligning with our goal of lightweight yet accurate state estimation.

A close work, by [Ken Masuya and Tomomichi Sugihara](#), addressed lightweight dead reckoning for biped robots using a complementary filter. However, their approach of the matter is different, especially the key variable whose estimation they aim to improve. Indeed, rather than focusing on the impact of the tilt estimate, their work focuses on wisely fusing position estimates coming from Leg odometry and Inertial odometry. The attitude estimate they leverage actually comes from an external estimator.

The attitude estimate they leverage comes from an IMU-based external estimator

The closest work to the proposed method is note that other works have addressed the problem of dead reckoning for humanoid robots based on complementary filters [Dead reckoning for biped robots that suffers less from foot contact condition based on anchoring pivot estimation](#). However, the proof of their filter's convergence is not provided,

Table 1. The caption must be shown before the table.

	Target of the complementary filter	Estimation of the remaining pose component
Ken Masuya and Tomomichi Sugihara	Position coming from IMU and Leg odometry	Attitude, using IMU
Ours	Tilt coming from IMU and Leg information	Position and Yaw coming from Leg odometry

they don't test on real experiments, and they don't focus on orientations: no focus on tilt and on the impact of the yaw.

As one can imagine, a slight error on the estimation of the yaw of the robot can lead to non negligible errors on the translation estimation. Therefore, estimating it as best as possible is a major objective in state estimation. The use of exteroceptive sensors is a direct solution to this issue, however they still don't offer a perfect odometry even with the most advanced methods (ex factor graphs with vilens), notably in some scenarios (ex: no loop closure, long tunnels for lidar, low-textured / low illumination environments, etc.). We propose a new orientation sensor that offers a very accurate and reliable measurement, even in challenging scenarios mentioned previously (eg: low-textured / low illumination environments), and at higher frequency (is it really much higher?). This sensor relies on a very lightweight algorithm, and can thus be used along other computationally expensive algorithms (ex: SLAM). In this article, we use this new sensor in pair with a highly accurate tilt estimation as inputs for a basic legged odometry. The purpose of this is to show how improved the legged odometry is with our orientation measurement even in a naive estimation framework unaware of information coupling, etc., to highlight the interest to include it to more sophisticated ones (that most of the time just add combine the legged odometry with other sensors in the most effective way possible).

LOSELY COUPLED (don't forget that we are loosely coupled) TIGHTLY COUPLED

While working on a basic legged odometry implementation, we noticed that the main source of error in the estimation was the yaw, which is highly subject to drifts while turning. Over long translations, the erroneous yaw leads to drastic errors on the absolute position estimation. We aim to correct this by using a novel orientation measurement. CONTINUER. We propose a legged odometry based on a very accurate tilt estimation, allowing for a good contact estimation along translations, and on a very accurate yaw orientation.

Possible use case: spatial exploration

1.2. Contributions

- Presentation of axis-agnostic orientation combinations.
- Lightweight combination of Leg Odometry with a highly accurate tilt estimate.
- Experimental evaluation of the Tilt Observer.
- The code of the Tilt Observer is open source, as well as a ROS wrapper.

2. PRELIMINARIES

2.1. General notations

- The general notation for kinematic variables is ${}^1\mathcal{O}_2$, expressing the kinematics of the frame 2 in the frame 1. To simplify the notation, kinematics in the world frame are written without the \mathcal{W} symbol: ${}^{\mathcal{W}}\mathcal{O}_2 = \mathcal{O}_2$.
- We define Ω the function that associates the rotation matrix to the corresponding rotation vector:

$$\text{SO}(3) \rightarrow \mathbb{R}^3 \quad (1)$$

$$\Omega: \mathbf{R} \mapsto \mathbf{v}, \quad \text{s.t.} \quad \exp(\text{S}(\mathbf{v})) = \mathbf{R} \text{ and } |\mathbf{v}| \leq \pi \quad (2)$$

- $[\cdot]_{\times}$ is the skew-symmetric operator.
- Define here the tilt?
- Define vec
- The world frame and the robot's IMU frame are denoted \mathcal{W} and \mathcal{I} , respectively. The frame associated with the i -th contact is denoted \mathcal{C}_i . The anchor point, defined in Section 4, is denoted \mathcal{A} .
- We define n_c the number of current contacts set with the environment.

3. AXIS-AGNOSTIC ORIENTATION COMBINATIONS

In this Section, we introduce the concept of *axis-agnostic* orientation representation. This answers the need for a representation that would allow for the combination of orientations, without axis ambiguities.

3.1. Tilt and yaw representation

3.2. Rotation matrix to yaw agnostic Useless here?

Rotation matrix to yaw axis agnostic:

$$\theta = \text{atan2}(\mathbf{v}_x \cdot \tilde{\mathbf{v}}_y - \mathbf{v}_y \cdot \tilde{\mathbf{v}}_x, \mathbf{v}_x \cdot \tilde{\mathbf{v}}_x + \mathbf{v}_y \cdot \tilde{\mathbf{v}}_y) \quad (3)$$

where $\mathbf{v} = \begin{bmatrix} R_{3,2} \\ -R_{3,1} \\ 0 \end{bmatrix}$ is the invariant horizontal vector projected onto the XY-plane $\in \mathbb{R}^2$ and $\tilde{\mathbf{v}} = \mathbf{R}_{2 \times 2} \cdot \mathbf{v}$

3.3. Merge tilt + yaw axis agnostic

Let us consider two rotation matrices \mathbf{R}_1 and \mathbf{R}_2 , corresponding to the orientations of the frames 1 and 2. The objective is to merge the *tilt* of the frame 1 with the *yaw* of the frame 2.

First, we compute the projection of *tilt*₁ in the frame 2:

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \mathbf{R}_2 \cdot \text{tilt}_1 \quad (4)$$

This allows us to compute the invariant orthogonal vector \mathbf{m} :

$$\mathbf{m} = \begin{cases} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, & \text{if } \left\| \begin{bmatrix} v_x \\ v_y \end{bmatrix} \right\| \approx 0 \\ \frac{1}{\sqrt{v_x^2 + v_y^2}} \begin{bmatrix} v_y \\ -v_x \\ 0 \end{bmatrix}, & \text{otherwise} \end{cases} \quad (5)$$

We then project \mathbf{m} in the frame 2:

$$\mathbf{m}_2 = \mathbf{R}_2^T \mathbf{m}. \quad (6)$$

We finally obtain a rotation matrix \mathbf{R} , resulting from the combination of the tilt of the frame 1 and of the yaw of the frame 2:

$$\mathbf{R} = \begin{pmatrix} \frac{\mathbf{m} \times \mathbf{e}_z}{\|\mathbf{m} \times \mathbf{e}_z\|} & \frac{\mathbf{e}_z \times \mathbf{m} \times \mathbf{e}_z}{\|\mathbf{m} \times \mathbf{e}_z\|} & \mathbf{e}_z \end{pmatrix} \cdot \begin{pmatrix} \frac{\mathbf{m}_2 \times \mathbf{tilt}_1}{\|\mathbf{m}_2 \times \mathbf{tilt}_1\|} & \frac{\mathbf{tilt}_1 \times \mathbf{m}_2 \times \mathbf{tilt}_1}{\|\mathbf{m}_2 \times \mathbf{tilt}_1\|} & \mathbf{tilt}_1 \end{pmatrix}^T \quad (7)$$

4. DEFINITION OF THE ANCHOR POINT

define rest frames. I already talk about the fixed contacts right after so adapt it. J'essaie de n'avoir qu'un repere pour les contacts, donc pas de repos We define here the notion of anchor point, since it will be used extensively in the following sections. The anchor point, denoted \mathcal{A} , is a point attached to the robot, and which is considered to have a zero velocity in the world frame. As will be explained in Section sec, this frame is especially important, since it provides a measurement of the robot's IMU velocity in the world frame. In the case of legged robots, it is common to assume that contacts with the environment are fixed. Any point of a contact surface would thus be usable as an anchor point. However, in order to pick a unique point that remains valid as contacts are created and broken, and which respects as best as possible the zero-velocity criterion, we compute its kinematics through a weighted average of the kinematics of the current contact reference frames. This is shown in Figure 1.

The weighting coefficients are defined such that weaker contacts, which are prone to violate Coulomb's inequality and thus to slip, contribute less to the anchor point's kinematics computation. Given mg the weight of the robot, we first compute $u_i = \frac{F_{i,z}}{\sqrt{F_{i,z}^2 + F_{i,x}^2 + \varepsilon mg}}$ the ratio of the normal force to the tangential force¹ at the contact i . The contact's weighting coefficient λ_i is then:

$$\lambda_i = \frac{u_i}{\sum_{j=1}^{n_c} u_j} \quad (8)$$

¹The term ε , arbitrarily small, allows to deal with the case where the tangential force is zero.

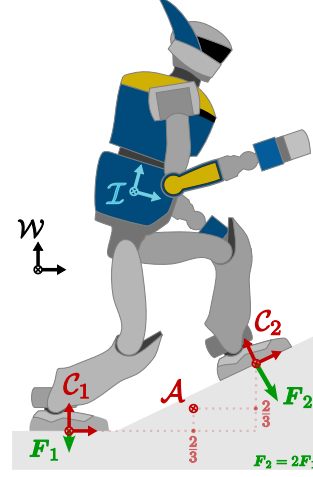


Fig. 1. Illustration of the anchor point position computation and of the reference frames used in VALINOR. \mathcal{W} : world frame; \mathcal{I} : IMU frame; \mathcal{C}_i : frame of the i -th contact. For clarity, only the weighted average of the contact positions is shown. In this example, $F_2 = 2F_1$ and so $\lambda_2 = \frac{2}{3}$.

It is important to note that although the anchor point is defined to have zero velocity in the world, its pose may still evolve over time. In discrete time, it has zero velocity? Maybe such a definition helps getting rid of the additional frame which coincides with it.

Based on this definition, we give the position and linear velocity of the anchor point in the IMU's frame, and its position in the world frame:

$${}^{\mathcal{I}}\mathbf{p}_{\mathcal{A}} = \sum_i^{n_c} \lambda_i {}^{\mathcal{I}}\mathbf{p}_{\mathcal{C}_i}, \quad (9)$$

$${}^{\mathcal{I}}\dot{\mathbf{p}}_{\mathcal{A}} = \sum_i^{n_c} \lambda_i {}^{\mathcal{I}}\dot{\mathbf{p}}_{\mathcal{C}_i}, \quad (10)$$

$$\mathbf{p}_{\mathcal{A}} = \sum_i^{n_c} \lambda_i \mathbf{p}_{\mathcal{C}_i}^*, \quad (11)$$

with ${}^{\mathcal{I}}\mathbf{p}_{\mathcal{C}_i}$ the position and ${}^{\mathcal{I}}\dot{\mathbf{p}}_{\mathcal{C}_i}$ the linear velocity of the i -th contact in the IMU's frame, which are directly provided by the robot's joint encoders. $\mathbf{p}_{\mathcal{C}_i}^*$ is the constant reference position of the i -th contact, which was its position at the instant it was created. What about p star and R star for the reference pose of the contacts? So we make clear they are different. But when should I talk about the reference pose of contacts? before or after? Right above i talk about the fixed contact assumption. Remove \mathcal{A}' ?

TODO

5. TILT OBSERVER WITH PROOF OF CONVERGENCE

The proposed estimator relies on a highly accurate estimate of the IMU's tilt provided by a complementary filter,

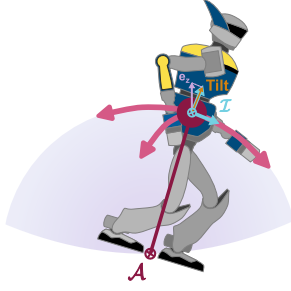


Fig. 2. Tilt

introduced in a previous work [citer](#), which we will call *Tilt Observer*.

5.1. Definition of the State Variables

The Tilt Observer is able to provide estimates of the following two variables:

- $v_{I,l} \triangleq R_I^T v_I$ the linear velocity of the IMU's frame in the world frame, expressed in the frame of the IMU.
- $R_I^T e_z$ the tilt of the IMU.

We thus define the state variables:

$$x_1 \triangleq v_{I,l}, x_1 \in \mathbb{R}^3, \quad (12)$$

$$x_2 \triangleq R_I^T e_z, x_2 \in \mathbb{S}^2. \quad (13)$$

The set $\mathbb{S}^2 \subset \mathbb{R}^3$ is the unit sphere centered at the origin, and defined as

$$\mathbb{S}^2 = \{x \in \mathbb{R}^3 \mid \|x\| = 1\} \quad (14)$$

This figure is uncited

5.2. Definition of the Measurements

The measurements required by the Tilt Observer are:

- y_g the signal of the IMU's gyrometer.
- y_a the signal of the IMU's accelerometer.
- y_v a measurement of $v_{I,l}$.

Since no sensor provides a direct measurement of $v_{I,l}$, we obtain y_v from an intermediate zero-velocity pseudo-measurement of the anchor point \mathcal{A} in the world, giving:

$$y_v = -[y_g]_{\times}^T p_{\mathcal{A}} - \dot{p}_{\mathcal{A}} \quad (15)$$

${}^I p_{\mathcal{A}}$ and $\dot{p}_{\mathcal{A}}$ are the position and linear velocity of the anchor point in the IMU's frame, obtained using (9) and (10).

5.3. Definition of the Filter

As explained in [Cite](#), the state dynamics of our system can be written:

$$\dot{x}_1 = -[y_g]_{\times} x_1 - g_0 x_2 + y_a, \quad (16)$$

$$\dot{x}_2 = -[y_g]_{\times} x_2. \quad (17)$$

We can then write a complementary filter which uses the system's dynamics as a feed-forward, and corrects the estimated state using the velocity measurement y_v :

$$\dot{\hat{x}}_1 = -[y_g]_{\times} \hat{x}_1 - g_0 \hat{x}'_2 + y_a + \alpha_1 (y_v - \hat{x}_1), \quad (18)$$

$$\dot{\hat{x}}'_2 = -[y_g]_{\times} \hat{x}_2 - \frac{\alpha_2}{g_0} (y_v - \hat{x}_1), \quad (19)$$

$$\dot{\hat{x}}_2 = -[y_g - \gamma[\hat{x}_2]_{\times} \hat{x}'_2]_{\times} \hat{x}_2. \quad (20)$$

α_1 , α_2 and γ are positive scalar gains. \hat{x}_1 and \hat{x}_2 are estimates of x_1 and x_2 , respectively.

The particularity of the Tilt Observer, in comparison to similar complementary filters (e.g. [Martin2016](#)), is the use of an intermediate estimate of the tilt, denoted \hat{x}'_2 . This intermediate variable allows for a fast convergence of \hat{x}_2 , while making sure it remains on the unit sphere.

5.4. Advantages of the Tilt Observer

The use of a complementary filter for the Tilt Observer presents notable strengths in comparison to other methods like the commonly used Kalman Filter. First, it allows us to work in the frequency domain. This is particularly suitable for our model since the assumption of fixed contacts used to obtain the velocity measurement y_v is more valid in low frequency than in high frequency. In (16), we thus use the IMU measurements for the high frequency variation of \hat{x}_1 , and y_v for its low frequency variation. Second, one iteration of the filter only consists in computing three equations, and is thus extremely computationally cheap, as will be shown in [Section](#). Finally, the formulation as a complementary filter allows to conduct a convergence analysis of the estimation error, providing strong mathematical guarantees on the estimator's performances. Especially here, it has been shown in [Cite](#) that:

- The dynamics of the estimation error is autonomous, and thus does not depend on the state.
- The intermediate estimator $\{\hat{x}_1, \hat{x}'_2\}$ is *globally exponentially stable*, with respect to the origin $(0,0)$.
- The full estimator is *almost globally asymptotically stable*, and locally *exponentially stable*.

6. LEG ODOMETRY

While the tilt of the IMU's frame in the world frame is estimated by the Tilt Observer, its position and yaw are obtained using Leg odometry. Once a contact i is created, its pose is obtained by forward kinematics from the current IMU's frame pose and the robot's joint encoders. This initial pose, which we call the contact's reference pose $\{p_{C_i}^*, R_{C_i}^*\}$, is then considered constant, allowing to recover the pose of the IMU's frame in the world frame. With the proposed pipeline, we thus leverage both the

accuracy and mathematical guarantees provided by the Tilt Observer, and the robustness to drift provided by the Leg odometry. Similarly to the computation of the anchor point in Section 4, the contribution of contacts to the Leg odometry is weighted to trust more contacts which are the least prone to slippage.

What about using the force ratio even for the correction of contacts?

6.1. Averaging of orientations and merging with tilt

To obtain the yaw from contact information, we compute the weighted average between the reference orientation of the two most reliable contacts. In our case, we assume they are the contacts with the highest measured reaction force, as explained in Section 4. For each contact i , we can write:

$$\mathbf{R}_{\mathcal{L},i} = \mathbf{R}_{C_i}^* \mathbf{R}_{\mathcal{L}} \quad (21)$$

We then compute the average $\mathbf{R}_{\mathcal{L},\text{avg}}$ between both contacts using the formalism defined by SO(3) the Lie group of rotation matrices:

$$\tilde{\mathbf{R}} = \mathbf{R}_{\mathcal{L},1}^T \mathbf{R}_{\mathcal{L},2} \quad (22)$$

$$\mathbf{R}_{\mathcal{L},\text{avg}} = \mathbf{R}_{\mathcal{L},1} \exp(\lambda_2 \text{vec}(\log(\tilde{\mathbf{R}}))) \quad (23)$$

\exp and \log are the *exponential* and *logarithm* maps of SO(3). We note that for small angles, we can use the approximation:

$$\log(\tilde{\mathbf{R}}) = \frac{1}{2}(\tilde{\mathbf{R}} - \tilde{\mathbf{R}}^T), \quad (24)$$

Once the average orientation has been computed, we

Check if Log or log. In the implementation, i don't do $R - R^T$

6.2. Averaging of positions

We use the updated orientation.

$$\mathbf{p}_{\mathcal{L}} = \mathbf{p}_A - \mathbf{R}_{\mathcal{L}}^T \mathbf{p}_A \quad (25)$$

7. EXPERIMENTAL EVALUATION

The proposed estimator has been evaluated across two experimental scenarios on two different humanoid robots²:

- a walk on a flat ground over about 18 meters with the robot RHP Friends. This experiment was repeated 5 times, for a total distance of about 90 meters.
- a multi-contact motion over about 2 meters with the robot HRP-5P. This motion involved an additional contact at the robot's left hand, and contacts on tilted obstacles. This experiments was repeated 4 times for a total distance of about 8 meters.

²we used the dataset built to evaluate the Kinetics Observer in [cite](#)

We compare VALINOR, our proposed estimator, with the Right-Invariant EKF (RI-EKF) by [Hartley2020RIEKF](#). For a fair comparison, both estimators received the same contact state information, obtained from a Schmitt Trigger applied to force sensor data (with thresholds set to 10% and 15% of the robot's weight). Ground truth pose is obtained from a motion capture system (OptiTrack, 16 Prime^X 13 cameras), and ground truth velocity is computed via finite differences and filtered with a zero-phase low-pass filter.

The odometry performance of VALINOR and the RI-EKF is evaluated against the ground truth trajectory using the Relative Error (RE), as defined in [Zhang2018QuantitativeTrajectoryEvaluation](#). RE evaluates the estimation drift over segments of fixed distances. It is particularly relevant for proprioceptive odometry, as it is not affected by global drift, which is unavoidable without exteroceptive sensors. RE therefore provides valuable and easily interpretable insight into the expected local accuracy of the estimator over a given distance.

We compute the Relative Error for lateral (x, y) and vertical (z) translations separately, since these components are generally influenced by different factors. Accuracy in lateral translation would mainly rely on that of local displacements and yaw estimation, whereas accuracy in vertical translations would be affected by the accuracy of the estimation of the vertical local displacement, of the tilt, and the reliability of the contact height initialization. Similarly, angular errors in tilt and yaw are also evaluated independently.

Finally, we also assess the estimation of the IMU's linear velocity in the world. This velocity is expressed in the frame of the IMU such that it is not affected by errors in the orientation estimate. The velocity error is also separated into lateral and vertical components.

7.1. Multi-contact

We evaluated the performances of VALINOR over a multi-contact motion, implying

7.2. Walk on flat floor

Show table and plot of pose and vel. Focus on tilt.

We can see from Table 3 that the yaw estimation made by the RI-EKF is slightly better than that of VALINOR. We explain this by the fact this orientation comes only from

7.3. Computation times comparison

To prove the computation speed performance of VALINOR, we compared its average computation time per iteration over the four multi-contact experiments, to that of the RI-EKF. In average, an iteration of VALINOR was computed in 2.547 μs , against 19.315 μs for the RI-EKF. The proposed method is thus more than 7.5 times faster

3D Trajectories with Start/End Markers

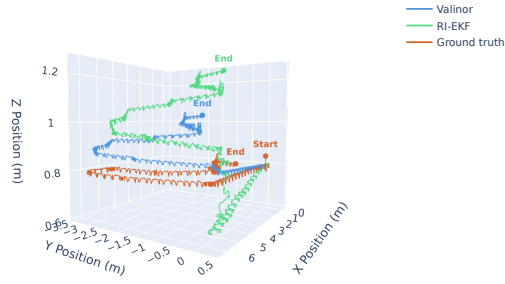


Fig. 3. Tilt

than the state-of-the-art estimator, while presenting superior or equivalent accuracy on the target variables.

8. CONCLUSION

Opening: include a correction of the contact pose references. Gyrometer and accelerometer biases.

APPENDIX A

The author(s) can insert an appendix with a meaningful title here.

DECLARATIONS

Conflict of Interest

Always applicable and includes interests of a financial or personal nature. For example, “The authors declare that

there is no competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.”

Authors’ Contributions

If there is one more author, please ensure that all authors’ contributions are individually mentioned with their full names.

Funding

It should be provided in the Declarations, separate from the Acknowledgements. If any of these declarations listed are not relevant to the content of your submission, please state that this declaration is “Not Applicable”.

REFERENCES

- [1] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake. “Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot.” *Autonomous Robots*, vol. 40, 07 2015.
- [2] T. Sugihara and M. Morisawa. “A survey: dynamics of humanoid robots.” *Advanced Robotics*, vol. 34, no. 21-22, pp. 1338–1352, 2020.
- [3] B. Weng, G. A. Castillo, W. Zhang, and A. Hereid. “On safety testing, validation, and characterization with scenario-sampling: A case study of legged robots.” *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5179–5186, 2022.
- [4] S. Katayama, M. Murooka, and Y. T. and. “Model predictive control of legged and humanoid robots: models and algorithms.” *Advanced Robotics*, vol. 37, no. 5, pp. 298–315, 2023.

Table 2. Mean and standard deviation (in parentheses) of errors computed during multi-contact motions. The 0.3 m Relative Error is represented. The best results for each metric are highlighted in bold.

	RE Translation [m]		RE Orientation [°]		Linear velocity [m.s ⁻¹]	
	Lateral { <i>x</i> , <i>y</i> }	Vertical <i>z</i>	Tilt {roll, pitch}	Yaw	Lateral { <i>x</i> , <i>y</i> }	Vertical <i>z</i>
VALINOR (Proposed)	0.007 (0.005)	0.002 (0.003)	0.23 (0.17)	0.40 (0.32)	0.018 (0.018)	0.006 (0.010)
RI-EKF [18]	0.012 (0.016)	0.004 (0.007)	0.57 (1.02)	0.47 (0.65)	0.015 (0.015)	0.006 (0.009)

Table 3. Mean and standard deviation (in parentheses) of errors computed during the flat odometry. The 1 m Relative Error is represented. The best results for each metric are highlighted in bold.

	RE Translation [m]		RE Orientation [°]		Linear velocity [m.s ⁻¹]	
	Lateral { <i>x</i> , <i>y</i> }	Vertical <i>z</i>	Tilt {roll, pitch}	Yaw	Lateral { <i>x</i> , <i>y</i> }	Vertical <i>z</i>
VALINOR (Proposed)	0.031 (0.021)	0.030 (0.046)	0.94 (1.97)	1.15 (1.18)	0.018 (0.018)	0.006 (0.010)
RI-EKF [18]	0.046 (0.056)	0.026 (0.051)	1.00 (1.84)	0.96 (0.92)	0.015 (0.015)	0.006 (0.009)

- [5] E. Dantec, M. Naveau, P. Fernbach, N. Villa, G. Saurel, O. Stasse, M. Taix, and N. Mansard. “Whole-body model predictive control for biped locomotion on a torque-controlled humanoid robot.” In *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pp. 638–644. Nov 2022. ISSN 2164-0580.
- [6] A. Dallard, M. Benallegue, N. Scianca, F. Kanehiro, and A. Kheddar. “Robust Bipedal Walking with Closed-Loop MPC: Adios Stabilizers.”, February 2024. URL <https://hal.science/hal-04147602>. Media material :<https://youtu.be/hfg7dTpuqu4>.
- [7] J. Peters, S. Vijayakumar, and S. Schaal. “Reinforcement learning for humanoid robotics.” *Proceedings of the third IEEE-RAS international conference on humanoid robots*, pp. 1–20, 01 2003.
- [8] Z. Li, X. B. Peng, P. Abbeel, S. Levine, G. Berseth, and K. Sreenath. “Reinforcement learning for versatile, dynamic, and robust bipedal locomotion control.” *The International Journal of Robotics Research*, vol. 44, no. 5, pp. 840–888, 2025.
- [9] NVIDIA, :, J. Bjorck, F. Castañeda, N. Cherniadev, X. Da, R. Ding, L. J. Fan, Y. Fang, D. Fox, F. Hu, S. Huang, J. Jang, Z. Jiang, J. Kautz, K. Kundalia, L. Lao, Z. Li, Z. Lin, K. Lin, G. Liu, E. Llontop, L. Magne, A. Mandekar, A. Narayan, S. Nasiriany, S. Reed, Y. L. Tan, G. Wang, Z. Wang, J. Wang, Q. Wang, J. Xiang, Y. Xie, Y. Xu, Z. Xu, S. Ye, Z. Yu, A. Zhang, H. Zhang, Y. Zhao, R. Zheng, and Y. Zhu. “Gr00t n1: An open foundation model for generalist humanoid robots.”, 2025. URL <https://arxiv.org/abs/2503.14734>.
- [10] K. Kawaharazuka, T. Matsushima, A. Gambardella, J. Guo, C. Paxton, and A. Zeng. “Real-world robot applications of foundation models: A review.” *Advanced Robotics*, vol. 38, no. 18, pp. 1232–1254, 2024.
- [11] M. N. Zeilinger, D. M. Raimondo, A. Domahidi, M. Morari, and C. N. Jones. “On real-time robust model predictive control.” *Automatica*, vol. 50, no. 3, pp. 683–694, 2014. ISSN 0005-1098.
- [12] R. Findeisen and F. Allgöwer. “Computational delay in nonlinear model predictive control.” *IFAC Proceedings Volumes*, vol. 37, no. 1, pp. 427–432, 2004.
- [13] P. Thodoroff, W. Li, and N. D. Lawrence. “Benchmarking real-time reinforcement learning.” In *NeurIPS 2021 Workshop on Pre-registration in Machine Learning*, volume 181 of *Proceedings of Machine Learning Research*, pp. 26–41. PMLR, 13 Dec 2022.
- [14] R. Firoozi, J. Tucker, S. Tian, A. Majumdar, J. Sun, W. Liu, Y. Zhu, S. Song, A. Kapoor, K. Hausman, B. Ichter, D. Driess, J. Wu, C. Lu, and M. Schwager. “Foundation models in robotics: Applications, challenges, and the future.” *The International Journal of Robotics Research*, vol. 44, no. 5, pp. 701–739, 2025.
- [15] M. Bloesch, M. Hutter, M. A. Hoepflinger, S. Leutenegger, C. Gehring, C. D. Remy, and R. Siegwart. “State estimation for legged robots-consistent fusion of leg kinematics and IMU.” *Robotics*, vol. 17, pp. 17–24, 2013.
- [16] M. Benallegue, R. Cisneros, A. Benallegue, Y. Chitour, M. Morisawa, and F. Kanehiro. “Lyapunov-Stable Orientation Estimator for Humanoid Robots.” *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6371–6378, October 2020.
- [17] A. Demont, M. Benallegue, A. Benallegue, P. Gergondet, A. Dallard, R. Cisneros, M. Murooka, and F. Kanehiro. “The Kinetics Observer: A Tightly Coupled Estimator for Legged Robots.”, December 2024. URL <https://hal.science/hal-04616647>. Working paper or preprint.
- [18] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle. “Contact-aided invariant extended kalman filtering for robot state estimation.” *The International Journal of Robotics Research*, vol. 39, no. 4, pp. 402–430, 2020.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.