An almost globally convergent observer for visual SLAM without persistent excitation

Bowen Yi, Chi Jin, Lei Wang, Guodong Shi and Ian R. Manchester

Abstract—In this paper we propose a novel observer to solve the problem of visual simultaneous localization and mapping (SLAM), only using the information from a single monocular camera and an inertial measurement unit (IMU). The system state evolves on the manifold \( SE(3) \times \mathbb{R}^{3n} \), on which we design dynamic extensions carefully in order to generate an invariant foliation, such that the problem is reformulated into online constant parameter identification. Then, following the recently introduced parameter estimation-based observer (PEBO) and the dynamic regressor extension and mixing (DREM) procedure, we provide a new simple solution. A notable merit is that the proposed observer guarantees almost global asymptotic stability requiring neither persistence of excitation nor uniform complete observability, which, however, are widely adopted in most existing works with guaranteed stability.

I. INTRODUCTION

Simultaneous localization and mapping (SLAM) is a fundamental problem widely studied in the robotics community, as well as in the field of navigation [8, 14]. In SLAM, two main aims are concerned to be accomplished concurrently—mapping an unknown environment, and online estimating the pose, \( i.e. \) attitude and position, of a mobile robot, thus making SLAM an important part for unmanned systems in the absence of absolute positioning systems.

The main approaches to address this problem may generally be classified into two categories. The first one is within the probabilistic and optimization framework, assuming Gaussian noises and processes, and then formulating the problem as maximum likelihood estimation, which generally has nonlinear least squares solutions, \( e.g. \) GraphSLAM [22] and the SLAM++ framework [9]. An alternative is to obtain estimation from an asymptotic convergence viewpoint—known as filtering—by using recursive algorithms. It includes extended Kalman filter (EKF)-SLAM and many modern algorithms, \( e.g. \) FastSLAM, which combines EKF and particle filtering [17]. An essential part of these algorithms is their convergence and consistency analysis, the success of which relies on first-order approximation of systems dynamics [8]. It sometimes yields satisfactory performance, but may have inconsistency issues when starting from a bad initial guess, in particular for large applications, which is caused by small domains of attraction in terms of linearization—invoking high nonlinearity of the associated dynamics. In the last few years, the nonlinear control community shows great interests to SLAM, providing alternatives via nonlinear observer design to address the problem. Indeed, SLAM can be regarded as the problem of state observation of a nonlinear system living on the manifold \( SE(3) \times \mathbb{R}^{3n} \). On the other hand, nonlinear observer for systems on manifolds is a well established topic, with special emphasis to matrix Lie groups [10, 12, 15]. Some very recent papers [5, 14, 21] give several solutions to the robo-centric SLAM problem, \( i.e. \), estimating landmark coordinates in the body-fixed frame, for which the dynamics can be transformed into linear time-varying (LTV) systems, thus avoiding the approximation error from linearization. Then, the Kalman-Bucy filter is applicable to provide \textit{globally} convergent estimation, if the robot movement guarantees uniform complete observability (UCO). It is well known that the UCO of an LTV system is equivalent to a persistency of excitation (PE) condition [20].

In this paper, we propose a new observer-based solution to visual SLAM, a case with only bearing measurement of landmarks available. Similar problems were recently studied in [23], in which the authors introduce a constructive observer design to visual SLAM by lifting to a new symmetry Lie group \( \mathcal{VSLAM}(3) \) in order to make the output function equivariant. Since the system is not strongly differentially observable, in order to be able to achieve asymptotic stability, some PE conditions are required for the robot trajectory. Besides [14, 23], some PE or UCO-type assumptions are also indispensable in some related problems, \( e.g. \), localization using range or direction measurements [6], velocity estimation using normalized measurement [4], and the non-stationary Perspective-n-Point (PnP) problem [7]. Intuitively, these assumptions impose relative motion between the robot and the landmarks, the uniformity of which should hold w.r.t. time. However, such assumptions may not be satisfied in many scenarios, such as, robots stopping in specific tasks, and landmarks appearing in the field of camera only during a finite interval, which validates neither the PE nor UCO conditions. Under these circumstances, the estimates of existing SLAM observers cannot converge to their values; and the observers may even diverge in the presence of measurement noise. Overcoming these problems are one of the motivations of the paper. Our main contributions are:

- **C1** Showing that visual SLAM observer design can be translated into online parameter estimation, and then solved by the recently introduced parameter estimation-based observer (PEBO) [18], guaranteeing invariance and almost global convergence, in contrast to the locality in EKF-SLAM methods;
- **C2** Providing a simple visual SLAM observer, which is...
robust vis-à-vis measurement noise, and enjoys low computation burden;

C3 Removing the PE condition required in some recent results, e.g., [4, 7, 14, 21, 23], the practical importance of which can hardly be overestimated, since it is usually the case in many tasks.

The remainder of the paper is organized as follows. In Section II we will present some preliminaries, notations and the kinematic model used in the paper, as well as giving the mathematical problem formulation. In Section III a novel visual SLAM observer will be designed. It will be followed by some simulation results in Section IV. Finally, the paper is wrapped up by a brief concluding remark. A version with full details of the proofs can be found in [24].

II. Problem Formulation and Preliminaries

A. Notations

We use \( SO(3) \) to represent the special orthogonal group, and \( \mathfrak{so}(3) \) is the associated Lie algebra as the set of skew-symmetric matrices satisfying \( SO(3) = \{ R \in \mathbb{R}^{3 \times 3} | R^T R = I_3, \det(R) = 1 \} \). The unit sphere is denoted as \( S^2 = \{ x \in \mathbb{R}^3 | |x| = 1 \} \). Given \( a \in \mathbb{R}^3 \), we define the operator \( (\cdot)_x \) as

\[
a_x := \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \in \mathfrak{so}(3).
\]

We also consider the special Euclidean group denoted as \( SE(3) = \{ T(R, x) \in \mathbb{R}^{4 \times 4} | R \in SO(3), x \in \mathbb{R}^3 \} \) with

\[
T(R, x) = \begin{bmatrix} R & x \\ 0 & 1 \end{bmatrix}.
\]

The Lie algebra of \( SE(3) \) is defined as

\[
\mathfrak{se}(3) := \left\{ A \in \mathbb{R}^{4 \times 4} \mid A = \begin{bmatrix} \Omega_x & v \\ 0 & 0 \end{bmatrix}, \Omega_x \in \mathfrak{so}(3), v \in \mathbb{R}^3 \right\}.
\]

For any \( x \in \mathbb{R}^3, a \times x \) is the vector cross product, satisfying \( a_x x = a \times x \). We define a wedged mapping \( U^\vee = \begin{bmatrix} \Omega_x & v \\ 0 & 0 \end{bmatrix} \) for a vector \( U := \text{col}(\Omega, v) \in \mathbb{R}^6 \). Given \( A \in \mathbb{R}^{n \times n} \) and \( S \in \mathbb{R}^{n \times n}_+, \) the Frobenius norm is defined as \( \| A \| = \sqrt{\text{tr}(A^T A)} \), and \( S^{\frac{1}{2}} \) and \( \text{adj}(A) \) represent the matrix square root and the adjugate matrix, respectively. We use \( | \cdot | \) to denote Euclidean norms of vectors or its induced matrix norm. For any \( x \in \mathbb{R}^3/(0) \), its projector is defined as \( \Pi_x := I_3 - \frac{1}{|x|^2} x x^T \), which projects a given vector onto the subspace orthogonal of \( x \). \( e_t \) represents exponentially decaying terms with proper dimensions. When clear from context, the arguments and subscripts are omitted. Before closing this subsection, let us recall below the notations of PE and interval excitation (IE). Obviously, IE is significantly weaker than PE, not requiring uniformity in time.

**Definition 1** Given a bounded signal \( \phi : \mathbb{R}_+ \rightarrow \mathbb{R}^n \), it is

- \((T, \delta)-\text{PE}, \) if \( \int_{t_0}^{t_0 + T} \phi(s) \phi^T(s) ds \geq \delta I_n \) for some \( T > 0, \delta > 0 \) and all \( t_0 \geq 0 \).

- \((t_0, t_c, \delta)-\text{IE}, \) if there exist \( t_0 \geq 0 \) and \( t_c \geq 0 \) such that \( \int_{t_0}^{t_0 + t_c} \phi(s) \phi^T(s) ds \geq \delta I_n \) for some \( \delta > 0 \).

B. Kinematic model and visual SLAM problem

The kinematics of a robot with rigid body is given by

\[
\begin{aligned}
\dot{x} &= Rx, \\
\dot{R} &= RO_x .
\end{aligned}
\]

All the definitions of symbols and the spaces where they live in can be found in Nomenclature. We assume that there are \( n \) landmarks appearing in the field view of camera, the coordinates \( z_i \) of which in \( \{ I \} \) are constant, thus satisfying

\[
\dot{z}_i = 0, \quad i \in \mathcal{N} := \{1, \ldots, n\} .
\]

We assume that \( v \) and \( \Omega \) are uniformly bounded for \( t \in [0, +\infty) \), and the kinematic model is forward complete. The time derivative of \( X := T(R, x) \) is

\[
\dot{X} = XU^\vee
\]

with \( U = \text{col}(\Omega, v) \) containing rotational and translational velocities. From some simple geometry identities, the point-type landmarks in \( \{ B \} \) verify \( x^{\mathcal{B}} = R^T (z_i - x) \). In visual SLAM only monocular cameras and IMUs are equipped on robotics, thus only landmark bearings being measurable. Without loss of generality, we assume that the camera frame coincides with \( \{ B \} \). Then, the output is exactly the bearing

\[
y_i = h_i(X, Z) = R^T \frac{z_i - x}{|z_i - x|}, \quad i \in \mathcal{N} .
\]

It is clear that the unit vector \( [z_i - x]/|z_i - x| \in \mathbb{S}^2 \) contains the orientation of the relative vector \( (z_i - x) \). For convenience, we write \( Y := [ y_1 | \ldots | y_n ] \) and \( Z := \text{col}(z_1, \ldots, z_n) \).

**Problem 1.** (visual SLAM observer) Consider the kinematics [2] with the output \( Y \), and assume that \( U \) is measured via IMUs. Design an observer

\[
\dot{\hat{\eta}} = F(\hat{\eta}, Y, U), \quad (\hat{X}, \hat{Z}) = H(\hat{\eta}, Y)
\]

with \( \hat{X} \in SE(3) \) and \( \hat{Z} \in \mathbb{R}^{3n} \), guaranteeing

\[
\lim_{t \rightarrow +\infty} [||\hat{X}(t) - X(t)|| + ||\hat{Z}(t) - Z||] = 0 .
\]
III. MAIN RESULTS

In this section, we will design an almost globally convergent observer to achieve \( \{V\} \), by means of PEBO.

A. Key algebraic identities

The key step in PEBOs is to generate invariant foliations among the system states and dynamic extension [18]. Here we construct a dynamic extension

\[
\dot{X}_e = X_e U^\top, \quad X_e := T(Q, \xi)
\]

(8)

with \( X_e \in SE(3) \), which can be regarded as a virtual robot in \( \{V\} \). We have the following.

Lemma 1 Consider the dynamics (2) and (8). The dynamic extension (8) is forward complete, and there exists a constant matrix

\[
X_e(t) = X_e X, \quad \forall t \geq 0.
\]

(9)

Proof: Defining an error variable \( E(X, X_e) := X_e X^\top \) and calculating its time derivative, we may verify the claim. See [24] for more details.

The above lemma shows that the open-loop dynamic extension (8) and the kinematics (2) admit a linear relationship (9)—more precisely, there is a constant rigid transformation between the frames \( \{I\} \) and \( \{V\} \)—by means of which we reformulate the state estimation of \( X(t) \) into the problem of online constant parameter identification of \( X_e \in SE(3) \). We define the coordinates of all the landmarks \( z_i \) in \( \{V\} \) as

\[
z_i^v := \xi + QR^\top (z_i - x), \quad \forall i \in N.
\]

(10)

A key observation is that the transformation (9) of the ambient from \( \{I\} \) to \( \{V\} \) does not change the relative transformation from a robot to a landmark.

Lemma 2 Consider the dynamics (2) and (8). The landmark coordinates \( z_i^v(t) \) (\( i \in N \)) in \( \{V\} \) are constant, verifying

\[
z_i^v = \xi + Q_c z_i.
\]

(11)

The landmark bearings in \( \{V\} \), defined as

\[
y_i^v := Q^\top \frac{z_i^v - \xi}{z_i^v - \xi}, \quad \forall i \in N,
\]

are measurable, i.e., \( y_i^v(t) \equiv y_i(t), \forall t \geq 0 \).

Proof: Invoking (9) and the definition (10), we have

\[
z_i^v = \xi + Q_c x + (Q_c R)^\top (z_i - x), \quad \forall i \in N
\]

which is clearly constant. About the second claim, we have

\[
y_i^v = Q^\top \frac{Q^\top (z_i - x)}{Q^\top (z_i - x)} = R^\top \frac{z_i - x}{z_i - x} = y_i, \quad \forall i \in N,
\]

where we have used (10) in the first equation.

As shown in Lemma 2, the landmark coordinates \( z_i^v \) in \( \{V\} \) are constant. Despite being quite simple, the above lemmata, show the estimation of time-varying systems state may be translated into online parameter estimation of \( X_e \) and \( z_i^v \). To the best of our knowledge, such a fact has not been used in SLAM observer design before, which, however, provides the possibility to relax significantly the PE or UCO assumptions in existing algorithms. In the sequel, we will show how to design a globally convergent observer.

B. Landmarks Observer in \( \{V\} \)

Now we are able to construct linear regressor equations (LREs) of constant \( z_i^v \). From (8) it yields

\[
Qy_i = \frac{z_i^v - \xi}{z_i^v - \xi} \quad \Rightarrow \quad Qy_i[y_i]^\top (z_i^v - \xi) = z_i^v - \xi.
\]

Noting that \( y(t) \) is a unit vector, we thus obtain the LREs

\[
q_i(t) = \Pi_Q(t)y_i(t), \quad z_i^v,
\]

(13)

by defining measurable signals \( q_i := \Pi_Q y_i, \xi \), where we have used the fact \( y_i(t) \equiv y_i(t) \) introduced in Lemma 2. It is also easy to verify the following equivalence for some \( \delta > 0 \).

\[
\Pi_{Q_{in}} \in (\delta', T) - \text{PE} \iff \int_t^{t+T} \Pi_{Q_{in}} ds \geq \delta I_3, \quad \forall t \geq 0.
\]

The widely used PE or UCO-type assumptions in the existing visual SLAM observers, intuitively, require that all landmarks appear in the view filed of the camera persistently, and the robot keeps moving w.r.t. the landmarks, which in general can hardly be guaranteed in practice. In the following, we will show that with the LREs (13) these restrictive excitation assumptions can be relaxed using the recently introduced DREM technique [1]. We first introduce some linear filters to generate a new LRE. To be precise, for each \( i \in N \) introducing an LTV operator \( \mathcal{H}_i : C^3_{\infty} \rightarrow C^3_{\infty} \) to (13), and then obtaining the extended (E)LRE

\[
\mathcal{H}_i(q_i(t)) = \mathcal{H}_i[\Pi_{i,1}(t)] \mathcal{H}_i[\Pi_{i,2}(t)] \mathcal{H}_i[\Pi_{i,3}(t)] z_i^v
\]

where the symbols \( \Pi_{i,j} (j = 1, 2, 3) \) are used with a slight abuse of notation, to denote the \( j \)-th column of the matrix \( \Pi_{Q(t)}y_i(t) \). Here, we adopt the LTV operator of the form

\[
\mathcal{H}_i(p, t)[\cdot] = \frac{\alpha_i}{p + \alpha_i} \Pi_Q(t)y_i(t)[\cdot]
\]

(14)

with \( \alpha_i > 0 \) and \( p := d/dt \) the differential operator. One of its state space realization is known as Kreisselmeier’s ELRE [11]—written as K-ELRE in the paper—which is given by

\[
q_i = -\alpha_i q_i^v + \alpha_i \Pi_Q y_i, \quad \Phi_i = -\alpha_i \Phi_i + \alpha_i \Pi_Q y_i
\]

(15)

with the system states \( (q_i, \Phi_i) \in \mathbb{R}^3 \times \mathbb{R}^3_{\geq 0} \) for \( i \in N \). It is straightforward to obtain the K-ELRE

\[
q_i(t) = \Phi_i(t) z_i^v + \epsilon_i.
\]

(16)

Then, we mix the regressors (16) to get three decoupled, scalar regressors for each \( i \in N \), that is pre-multiplying the adjugate matrix \( \text{adj}(\Phi_i(t)) \) to the both sides, thus obtaining

\[
Y_{i,j}(t) = \Delta_i(t) z_i^v + \epsilon_i, \quad i \in N, \quad j \in \{1, 2, 3\}
\]

(17)

with the definitions \( Y_{i,j} := \text{adj}(\Phi_i) q_i^v, \quad \Delta_i := \text{det}(\Phi_i) \) and \( Y_i := \text{col}(Y_{1,1}, Y_{1,2}, Y_{1,3}) \). We are in position to present a novel landmark observer in \( \{V\} \), which only requires a strictly weaker condition than PE.
Thus, the system is internally stable.

\[
\chi_i(t) - z_i^v = \exp\left(-\int_0^t \Delta_i(s)^2 ds\right) (\chi_i(0) - z_i^v) = \omega(t)(\chi_i,0 - z_i^v)
\]

which is an LTV dynamics, with \(\Delta_i\) a scalar bounded signal. Its solution is

\[
\chi_i(t) = \chi_i(0) + \int_0^t \Delta_i(s)^2 ds
\]

by noting the solution \(\omega_i(t) = \exp(-\int_0^t \Delta_i(s)^2 ds)\). It yields

\[
\chi_i(t) - \omega_i(t)\hat{\chi}_i,0 = [1 - \omega_i(t)]z_i^v,
\]

in which we underscore that \(\chi_i, \omega_i\) and \(\chi_i,0\) \((t \geq 0)\) are all available in observer design. In this way, we get new scalar LREs \(20\) involving the “integral” information of \(\Delta_i\). Combining \(20\) and \(17\), we obtain new linear regressors

\[
Y_i + k_i^1(\chi_i - \omega_i\hat{\chi}_i,0) = \Delta_i^v z_i^v,
\]

in which \(z_i^v\) are the unknown constant parameters.

By defining the estimation error \(z_i^v = \hat{z}_i^v - z_i^v\), we have \(21\)

\[
\dot{z}_i^v = -\gamma_i(\Delta_i^v)^2 z_i^v.
\]

To verify the first claim, we choose the Lyapunov function \(V = \frac{1}{2} \sum_{i=1}^n (|z_i^v|^2 + |\chi_i - z_i^v|^2 + |\omega_i|^2)\), satisfying

\[
\dot{V} = -\sum_{i=1}^n \left[\gamma_i(\Delta_i^v)^2 |z_i^v|^2 + \Delta_i^v |\chi_i - z_i^v|^2 + \Delta_i^v |\omega_i|^2\right] \leq 0.
\]

Thus, the system is internally stable.

The second claim can be easily verified invoking \(\Delta_i^v\) are scalar signals. Recalling the dynamics of \(\dot{z}_i^v\), the last claim is equivalent to verify \(\Delta_i^v \in \text{PE}\). To simplify presentation, we neglect the index \(i\) of the IE condition from the \(i\)-th landmark. From the IE assumption of \(\Pi_{Q_{yi}}\), there exist \(t_0, t_c, \delta \in \mathbb{R}_+\) such that

\[
\int_{t_0}^{t_0+t_c} \int_{t_0}^{t_0+t_c} \Pi_{Q_{yi}(s)}\Pi_{Q_{yi}(s)} ds \geq \delta I_3,
\]

since \(\Pi_{Q_{yi}}\) is symmetric. Noting \(15\) in the K-ELRE, we can verify

\[
\Phi_i(t_0 + t_c) \geq \int_{t_0}^{t_0+t_c} e^{-\alpha_i(t_0+t_c-s)} \Pi_{Q_{yi}(s)}^T \Pi_{Q_{yi}(s)} ds 
\]

\[
\geq e^{-\alpha_i t_c} \int_{t_0}^{t_0+t_c} \Pi_{Q_{yi}(s)}^T \Pi_{Q_{yi}(s)} ds 
\]

\[
\geq \delta e^{-\alpha_i t_c} I_3,
\]

It implies \(\Delta(t_0 + t_c) \geq \delta_0 = (\delta e^{-\alpha_i t_c})^3\). From the continuity of differential equations, there exist small \(\tau > 0\) and \(\varepsilon > 0\) such that

\[
\int_{t_0+t_c}^{t_0+t_c+\tau} \Delta_i(s) ds \geq \varepsilon \Rightarrow \int_{t_0+t_c}^{t_0+t_c+\tau} \Delta_i^2(s) ds \geq \frac{\varepsilon}{\tau}, \forall t \geq t_c
\]

\[
\Rightarrow 1 - \omega_i \geq 1 - e^{-\frac{\varepsilon}{\tau}}, \forall t \geq t_c
\]

\[
\Rightarrow \Delta_i^v \geq \Delta_i + k_i^1(1 - \omega_i) \in \text{PE}
\]

where we used \(\Delta_i \geq 0\), \(\forall t \geq 0\) and the Cauchy-Schwarz inequality for integrals.

Some remarks are made below about the proposed design.

**R1** The success of the proposed landmark observer relies on the new scalar LREs \(21\)—motivated by \([19]\)—which satisfy the PE condition. It combines two parts, namely, the first contains information in the current “interval” involving the K-ELRE generated by the filter \(15\), and the second one consists of historical data using an integral operation to generate the “state transmission function” \(\omega(t)\). Here, the parameters \(k_i^1\) are adopted to play a role of weighting between them.

\[
\underbrace{Y_i}_{\text{"current interval"}} + \underbrace{k_i^1(\chi_i - \omega_i\hat{\chi}_i,0)}_{\text{historical information}} = \Delta_i^v z_i^v.
\]

**R2** The obtained new LREs \(20\) may be replaced by other integral of \(\Delta_i\), in order to make full use of the IE condition. It, however, may cause parts of internal states in the observer unbounded as time goes to infinity, once \(\Delta_i\) satisfies the PE condition.

### C. Landmark and Pose Observer in \(\mathcal{I}\)

After obtaining the landmarks estimation in \(\mathcal{Y}\), we then need to express it in \(\mathcal{I}\), as well as to estimate the rigid body pose \(X = T(R, x)\). It is widely recognized that the full dynamics \(2\), under the output functions \(12\), is not strongly differentially observable \([13]\)—a notion widely explored in high-gain observers \([3]\). The underlying reason is clear that it is ambiguous to identify the origin of \(\mathcal{I}\) with only body-fixed bearing measurement, and angular/translational velocities, see for example \([23, 25]\) in which the estimation error converges to a quotient manifold rather than an isolated equilibrium. There are generally two technical routes to circumvent such difficulties:

**T1** Assuming the initial conditions of pose states \(X(0) = T(R(0), x(0))\) in the inertial frame \(\mathcal{I}\).

**T2** Incorporating the measurements in \(\mathcal{B}\) of (at least) two vectors, the inertial coordinates of which are known in advance.
In the latter two known vectors provide information to transform the remainder of the task into a rigid body pose estimation problem, which is widely studied in the control literature [2, 15, 16]. However, the former coincides with the “standard” definition of SLAM problems. With this consideration, we will pursue the first route in the sequel. From now on, we assume the initial pose $X(0) = X_0 = T(R_x, x_s) \in SE(3)$ is pre-selected, thus known, in order to “fix” the inertial frame $\{I\}$. It is clear that for $\hat{\rho}$ by choosing particular initial conditions we have

$$\begin{align*}
\hat{\rho}(t) &= x_s, \quad Q(0) = R_x \\
\implies &X(t) = X_e(t), \forall t \geq 0
\end{align*}$$

under ideal circumstance. This, however, is not practically applicable, since the open-loop integral $\int \hat{\rho} dt$ may be problematic yielding error accumulation. In this subsection, we introduce an approach to design a pose observer and robustify the integral operation vis-à-vis measurement noise. Before continuing our design, we define the vectors $r_i := z_{i+1} - z_i, \forall i \in \mathbb{N} \setminus \{n\}$, and make the following assumptions.

**Assumption 1** The origin of $\{B\}$ never coincides with any landmarks for $i \in \mathbb{N}$. There are at least three landmarks $z_i$ such that $r_i \times r_j \neq 0$, for $i \neq j$. These landmarks, without loss of generality, are numbered in the first $n_e \geq 3$.

**Assumption 2** The locomotion of the robot guarantees that $\Pi_{Q(t)q_{0}(t)}(t_{0}, i, t, c_{i}, \delta_{i})$-IE $(i \in \mathbb{N})$ along the trajectory $X(t)$ for all the landmarks to be mapped. Additionally, we assume a moment $T_e = \max_{i \in \{1, \ldots, n_e\}} \{t_{c_{i}}\} > 0$ is known.

**Proposition 2 (Localization)** Consider the kinematics $\Sigma[2]$ under Assumptions $\Sigma[1]$ The pose observer

$$\dot{\hat{\rho}} = \begin{cases}
\hat{\rho}_{j} = \rho_{j}(\tilde{y}_{j} - \hat{\rho}_{j} \tilde{z}_{j}) \\
\hat{Q}_{c} = -(w_{vis})_{i} \hat{Q}_{c}
\end{cases} \quad \begin{align}
\dot{\tilde{z}} &= \hat{R} \tilde{v} + \sum_{j=1}^{n_e} s_{j}(\hat{\rho}_{j} - \hat{Q}_{c} \tilde{z}_{j} - \xi_{j})
\end{align} \quad (23)$$

with parameters $\rho_{j}, k_{j}, s_{j} > 0$ $(j \in \{1, \ldots, n_e\})$, variables $\tilde{v}_{j} = \tilde{z}_{j+1} - \tilde{z}_{j}$, $\hat{R} = \hat{Q}_{c}^T Q$ and $w_{vis} = \sum_{j=1}^{n_e} k_{j} \tilde{v}_{j} \times (\hat{Q}_{c} \tilde{v}_{j})$,

$$\begin{cases}
\tilde{y}_{j} = \sum_{j=1}^{n_e} \Pi_{Q_{y_{j}}} (\xi_{j} - \hat{Q}_{c} \tilde{z}_{j} - \xi_{j}) \Pi_{Q_{y_{j}}} Q(0) R_{x_{s}} \tilde{v}_{j} \\
0,
\end{cases} \quad t \in [0, T_e] \quad t > T_e$$

and the landmark observer consisting of $\Sigma[1]$ and

$$\hat{z}_{i} = \hat{Q}_{c}^T \hat{e}_{i} - \hat{x}, \quad i \in \mathbb{N}$$

achieve the task $\Sigma[7]$ almost globally.

**Proof:** The proof can be found in [24].

To illustrate the proposed design, we give the schematic block diagram in Fig. 4.

**IV. SIMULATIONS**

A robot is simulated from $x(0) = [1, 1, 2] \in \mathbb{R}^3$ and the attitude $R(0) = \begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) & 0 \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) & 0 \end{bmatrix}$, and we assume that the robot stopped at 12s with

$$v(t) = \begin{cases}
[1, 0, 0]^T & t \in [0, 12] \\
0 & t \geq 12
\end{cases}, \quad \Omega = \begin{cases}
[0, 0, -0.4]^T & t \in [0, 12] \\
0 & t \geq 12
\end{cases}$$

It guarantees that all the landmarks satisfy the IE condition with $t_{0} = 0$ and $t_{c} = 12s$. We consider six landmarks and $n_{e} = 3$. The initial conditions in the observer are set as

$$Q(0) = \begin{bmatrix}
\begin{bmatrix}
\alpha_{i} & 0 & 0 \\
0 & \gamma_{i} & 0 \\
0 & 0 & \nu_{i}
\end{bmatrix} & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad Q_{c}(0) = \begin{bmatrix} 0 & 1 \\
1 & 0
\end{bmatrix}$$

$$\hat{x}(0) = 0_{3times1}, \quad \hat{q}_{c}(0) = 0_{3times1}, \quad \Phi_{c}(0) = 0_{3times3}.$$ The observer gains are selected as $\alpha_{i} = 5, \gamma_{i} = 100, k_{i} = 20$ for $i \in \mathbb{N}$, and $\rho_{j} = 1$ for $j = 1, 2, 3$. Simulation was done in Matlab/Simulink, with noise added to measured velocities and bearings, generated by the block “Uniform Random Number”. The pose estimation has a very good performance in Fig. 3 when the trajectory does not guarantee the PE condition for landmarks. The landmark observer in Proposition $\Sigma[1]$ can be used independently for mapping. Here, we compare it to the landmark observer in [14], which requires the system being UCO with sufficient excitation. The observer in [14] studies the case that the landmark coordinates are expressed in $\{B\}$, i.e., $z^{B}_{i} = R^{T} (z_{i} - x), \quad i \in \mathbb{N}$, which are, indeed, time-varying, the simulation results of which are shown in Fig. 4(a) In order to make a fair comparison, we plot the evolution of the norms of the observation errors of $\hat{z}_{i}$ of the proposed design and $z^{B}_{i}$ for the one in [14], since rotation does not affect norms. The proposed landmark observer guarantees the estimation converging to a relatively small neighbourhood of their true values in the absence of the PE condition, and the smallest ultimate error is caused measurement noise, showing good robustness. In Fig. 3 we observe that the estimates from the observer in [14] stop converging at the moment $t_{0} + t_{c} = 12s$ with large errors. It is interesting to note that the estimates diverge from that moment due to the accumulation of noise, which is conspicuous by its absence in our proposed design.
with I3S-CNRS, France, and he appreciates the fruitful discussions with Prof. Tarek Hamel.

**Fig. 3.** Pose $X(t)$ and its estimate $\hat{X}(t)$ without the PE condition

**Fig. 4.** Comparison between the proposed design and [14]

(a) Landmark paths $z_b^i$ and their estimation paths $\hat{z}_b^i$ in $\mathcal{B}$

(b) The landmark estimates $\hat{z}_1$ of the first landmark in the body $\mathcal{B}$

(c) $\tilde{z}_v^i$ from the proposed landmark observer in Prop. 1

(d) $\tilde{z}_B^i$ from the LTV Kalman filter in [14]

V. CONCLUSION

This paper introduces a novel visual SLAM observer design method. A key observation is that the landmarks in the frame of dynamic extension are constant, based on which we are able to get a set of linear regressors, and then transform the problem into online parameter estimation. We extend the PEBO methodology to the manifold $SE(3) \times \mathbb{R}^{3n}$, the unknown “parameters” in our context being the landmark coordinates $z_v^i$ in $\mathcal{V}$ together with the constant relative rigid transformation $X_v$. A simple constructive design is provided, with guaranteed almost global convergence, while significantly relaxing the strong PE or UCO-type conditions required in the existing literature. As future work, it is of practical interest to study the performance limitation from gyro noise and bias, as well as the approach to robustify the proposed observer design.

VI. ACKNOWLEDGEMENT

Partial work has been done when the second author was with I3S-CNRS, France, and he appreciates the fruitful discussions with Prof. Tarek Hamel.

REFERENCES

[1] S. Aranovskiy, A. Bobtsov, R. Ortega and A. Pyrkin, Performance enhancement of parameter estimators via dynamic regressor extension and mixing, *IEEE Trans. on Automatic Control*, vol. 62, pp. 3546–3550, 2016.

[2] S. Berkane, A. Abdessameud and A. Tayebi, Hybrid attitude and gyro-bias observer design on $SO(3)$, *IEEE Trans. on Automatic Control*, vol. 62, pp. 6044–6050, 2017.

[3] G. Besançon (Ed.), *Nonlinear Observers and Applications*, Springer, Berlin, 2007.

[4] E. Bjørne, E. F. Brekke, T. H. Bryne, J. Delaune and T. A. Johansen, Globally stable velocity estimation using normalized velocity measurement, *Int. J. of Robotics Research*, vol. 39, pp. 143–157, 2020.

[5] B. J. Guerreiro, P. Batista, C. Silvestre and P. Oliveira, Globally asymptotically stable sensor-based simultaneous localization and mapping, *IEEE Trans. on Robotics*, vol. 29, pp. 1380–1395, 2013.

[6] T. Hamel and C. Samson, Position estimation from direction or range measurements, *Automatica*, vol. 82, pp. 137–144, 2017.

[7] T. Hamel and C. Samson, Riccati observers for the nonstationary PnP Problem, *IEEE Trans. on Automatic Control*, vol. 63, pp. 726–741, 2018.

[8] S. Huang and G. Dissanayake, Convergence and consistency analysis for extended Kalman filter based SLAM, *IEEE Trans. on Robotics*, vol. 23, pp. 1036–1049, 2007.

[9] V. Ila, L. Polok, M. Solony and P. Svoboda, SLAM++: A highly efficient and temporally scalable incremental SLAM framework, *Int. J. of Robotics Research*, vol. 36, pp. 210–230, 2017.

[10] M. Izadi and A. K. Sanyal, Rigid body pose estimation based on the Lagrange-d’Alembert principle, *Automatica*, vol. 71, pp. 78–88, 2016.

[11] G. Kreisselmeier, Adaptive observers with exponential rate of convergence, *IEEE Trans. on Automatic Control*, vol. 22, pp. 2–8, 1977.

[12] C. Lageman, J. Trumpf and R. Mahony, Gradient-like observers for invariant dynamics on a Lie group, *IEEE Trans. on Automatic Control*, vol. 55, no. 2, pp. 367–377, 2010.

[13] K. W. Lee, W. S. Wijesoma and I. G. Javier, On the observability and observability analysis of SLAM, *Proceeding of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 3569–3574, 2006.

[14] P. Lourenço, P. Batista, P. Oliveira and C. Silvestre, A globally exponentially stable filter for bearing-only simultaneous localization and mapping with monocular vision, *Robotics and Autonomous Systems*, vol. 100, pp. 61–77, 2018.

[15] R. Mahony, T. Hamel and J.-M. Pflimlin, Nonlinear complementary filters on the special orthogonal group, *IEEE Trans. on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[16] A. Moeini and M. Namvar, Global attitude/position estimation using landmark and biased velocity measurements, *IEEE Trans. on Aerospace and Electronic Systems*, vol. 52, pp. 852–862, 2016.

[17] M. Montemerlo and S. Thrun, FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics, *IEEE Trans. on Robotics*, vol. 27, New York, NY, USA, Springer, 2007.

[18] R. Ortega, A. Bobtsov, A. Pyrkin and S. Aranovskiy, A parameter estimation approach to state observation of nonlinear systems, *Systems & Control Letters*, vol. 85, pp. 84–94, 2015.

[19] R. Ortega, A. Bobtsov, N. Nikolaev, J. Schiffer, D. Dochain, Generalized parameter estimation-based observers: Application to power systems and chemical-biological reactors, vol. 129, 109635, 2021.

[20] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness*, Dover Publication, New York, 1989.

[21] F. Tan, W. Lohmiller and J.J. Slotine, Analytical SLAM without linearization, *Int. J. of Robotics Research*, vol. 36, pp. 1554–1578, 2017.

[22] S. Thrun and M. Montemerlo, The graph SLAM algorithm with applications to large-scale mapping of urban structures, *Int. J. of Robotics Research*, vol. 25, pp. 403–429, 2006.

[23] P. van Goor, R. Mahony, T. Hamel and J. Trumpf, Constructive observer design for visual simultaneous localisation and mapping, *Automatica*, vol. 132, 109803, 2021.

[24] B. Yi, C. Jin, L. Wang, G. Shi and I.R. Manchester, An almost globally convergent observer for visual SLAM without persistent excitation, *ArXiv Preprint*, 2021. (arXiv:2104.02966)

[25] D. E. Zlotnik and J. R. Forbes, Gradient-based observer for simultaneous localization and mapping, *IEEE Trans. on Automatic Control*, vol. 63, pp. 4338–4344.