Appendix of MM' 22 paper "LVI-ExC: A Target-free LiDAR-Visual-Inertial Extrinsic Calibration Framework"

1 SUMMARY OF RELATED WORK

In order to help understand the characteristics of existing related work, we summarize their main features and shortcomings in Table 1 according to the types of the sensor suites to be calibrated and the categories of the calibration methods.

2 3D ON-MANIFOLD OPERATORS

As mentioned in the main body, Exp denotes the mapping of an element in $\mathfrak{so}(3)/\mathfrak{se}(3)$ to the special orthogonal/Euclidean group $\mathbb{SO}(3)/\ \mathbb{SE}(3)$, which conforms to,

$$\mathrm{Exp}: \mathfrak{so}(3) \ni \boldsymbol{\phi} \to \exp(\lfloor \boldsymbol{\phi} \rfloor_\times) \in \mathbb{SO}(3) \tag{1}$$

$$\mathfrak{se}(3) \ni \boldsymbol{\xi} \to \exp(\lfloor \boldsymbol{\xi} \rfloor_\times) \in \mathbb{SE}(3) \tag{2}$$

$$\mathrm{Log}: \mathbb{SO}(3) \ni \boldsymbol{R} \to \log(\boldsymbol{R})^\vee \in \mathfrak{so}(3) \tag{3}$$

$$\mathbb{SE}(3) \ni \boldsymbol{T} \to \log(\boldsymbol{T})^\vee \in \mathfrak{se}(3). \tag{4}$$

Given $\boldsymbol{\phi} = [\phi_x, \phi_y, \phi_z]^T$, $\boldsymbol{\rho} \in \mathbb{R}^{3\times1}$ and $\boldsymbol{\xi} = [\boldsymbol{\phi}^T, \boldsymbol{\rho}^T]^T$, the operator $\lfloor \cdot \rfloor_\times$ produces the matrices of the associated vectors as,

$$\lfloor \boldsymbol{\phi} \rfloor_\times = \begin{bmatrix} 0 & -\phi_z & \phi_y \\ \phi_z & 0 & -\phi_x \\ -\phi_y & \phi_x & 0 \end{bmatrix} \text{ and } \lfloor \boldsymbol{\xi} \rfloor_\times = \begin{bmatrix} \lfloor \boldsymbol{\phi} \rfloor_\times & \boldsymbol{\rho} \\ \boldsymbol{0}^T & 0 \end{bmatrix}. \tag{5}$$

Its reverse operator is denoted by $\vee$, with $\lfloor \boldsymbol{\omega} \rfloor_\times^\vee = \boldsymbol{\omega}$.

3 IMU PRE-INTEGRATION

The raw acceleration and angular velocity readings of the IMU come from its local coordinate system at any given time. The physical measurement model of an IMU can be written as,

$$\tilde{\boldsymbol{a}}_b^b = \boldsymbol{R}_b^{wT}(\boldsymbol{a}_b^w - \boldsymbol{g}^w) + {}^a\boldsymbol{b} + {}^a\boldsymbol{n} \tag{6}$$

$$\tilde{\boldsymbol{\omega}}_b^b = \boldsymbol{\omega}_b^b + {}^\omega\boldsymbol{b} + {}^\omega\boldsymbol{n}, \tag{7}$$

where $\tilde{\boldsymbol{a}}_b^b$ and $\tilde{\boldsymbol{\omega}}_b^b$ denote the measurements of the accelerometer and the gyroscope, respectively. ${}^a\boldsymbol{b}$ and ${}^\omega\boldsymbol{b}$ are the biases of the acceleration and the angular velocity, while ${}^a\boldsymbol{n}$ and ${}^\omega\boldsymbol{n}$ represent the measurement noise of the accelerometer and the gyroscope respectively.

If the initial state of the carrier in the world system is known, we can predict its state at time t theoretically according to the well-known Newton's law via integrating the IMU readings. However, when we update an estimated historical state, all associated states must be re-integrated, which is extremely time-consuming and impractical. To cope with this problem, we resort to the "IMU pre-integration" [4] technology, which skillfully establishes the correlation between the relative motion and the raw IMU data.

According to Forster's results [4], the relationship among the IMU pre-integration, true states and noises conforms to,

$$\begin{aligned} \Delta\tilde{\boldsymbol{R}}_{b_j}^{b_i} &\approx \boldsymbol{R}_{b_i}^{wT}\boldsymbol{R}_{b_j}^{w}\mathrm{Exp}(\delta\boldsymbol{\phi}_{b_j}^{b_i}) \\ \Delta\tilde{\boldsymbol{v}}_{b_j}^{b_i} &\approx \boldsymbol{R}_{b_i}^{wT}(\boldsymbol{v}_{b_j}^{w} - \boldsymbol{v}_{b_i}^{w} - \boldsymbol{g}^w\Delta t_{ij}) + \delta\boldsymbol{v}_{b_j}^{b_i} \\ \Delta\tilde{\boldsymbol{p}}_{b_j}^{b_i} &\approx \boldsymbol{R}_{b_i}^{wT}(\boldsymbol{p}_{b_j}^{w} - \boldsymbol{p}_{b_i}^{w} - \boldsymbol{v}_{b_i}^{w}\Delta t_{ij} - \frac{1}{2}\boldsymbol{g}^w\Delta t_{ij}{}^2) + \delta\boldsymbol{p}_{b_j}^{b_i}, \end{aligned} \tag{8}$$

where $\boldsymbol{p} \in \mathbb{R}^3$, $\boldsymbol{v} \in \mathbb{R}^3$, and $\boldsymbol{R} \in \mathbb{R}^{3\times3}$ denote the position, the velocity, and the rotation respectively; $\delta\boldsymbol{\phi}_{b_i}^{b_j}$, $\delta\boldsymbol{v}_{b_i}^{b_j}$ and $\delta\boldsymbol{p}_{b_i}^{b_j}$ are the

Table 1: Traits of existing extrinsic calibration studies.

Sensor Suite	References	Category	Main Features	Shortcomings
Camera-IMU	Mirzaei and Roumeliotis [16], Kelly and Sukhatme [8], Fleps *et al.* [3], Furgale *et al.* [5], Rehder and Siegwart [19] , Kim *et al.* [9]	Target-based	Estimating camera pose with a checkerboard. Fusing IMU and camera data via Kalman filter or joint optimization.	The checkerboard limits the carrier's motion. The initial values of the extrinsics rely on prior knowledge.
LiDAR-IMU	Lv *et al.* [14], Gentil *et al.* [12]	Target-free	Constructing point to plane constraints from natural scenes.	Necessary initialization is missing.
LiDAR-Camera	Zhou *et al.* [22], Guindel *et al.* [6], Pusztai and Hajder [18], Chen *et al.* [2], Bai *et al.* [1], Kümmerle and Kühner [11]	Target-based	Designing auxiliary calibration objects and extracting geometric features from co-visible area.	Poor flexibility, tedious production of calibration objects, limited to offline calibration.
	Ishikawa *et al.* [7], Park *et al.* [17], Kim and Park [20], Koo *et al.* [10], Taylor and Nieto [21], Zhu *et al.* [23]	Target-free	Estimating trajectories separately and employing hand-eye calibration.	Odometry with a single sensor is not accurate enough.

Gaussian noises of the pre-integration measurements; Δt_{ij} is the time difference between the i-th and the j-th IMU readings; and the pre-integrations can be computed from the raw IMU data via,

$$\Delta\tilde{\boldsymbol{R}}_{b_j}^{b_i} = \prod_{k=i}^{j-1} \mathrm{Exp}((\tilde{\boldsymbol{\omega}}_{b_k}^{b_k} - {}^{\omega}\boldsymbol{b}_{b_i})\Delta t)$$

$$\Delta\tilde{\boldsymbol{v}}_{b_j}^{b_i} = \sum_{k=i}^{j-1} [\Delta\tilde{\boldsymbol{R}}_{b_k}^{b_i}(\tilde{\boldsymbol{a}}_{b_k}^{b_k} - {}^{a}\boldsymbol{b}_{b_i})\Delta t]$$

$$\Delta\tilde{\boldsymbol{p}}_{b_j}^{b_i} = \sum_{k=i}^{j-1} [\Delta\tilde{\boldsymbol{v}}_{b_k}^{b_i}\Delta t + \frac{1}{2}\Delta\tilde{\boldsymbol{R}}_{b_k}^{b_i} \cdot (\tilde{\boldsymbol{a}}_{b_k}^{b_k} - {}^{a}\boldsymbol{b}_{b_i})\Delta t^2]. \quad (9)$$

4 BUNDLE ADJUSTMENT

After the constraints among different sensors are established, we can jointly optimize all the variables via resorting to common mathematical tools. Those variables include the extrinsics ($\boldsymbol{R}_c^b$, $\boldsymbol{p}_c^b$, $\boldsymbol{R}_l^b$ and $\boldsymbol{p}_l^b$), the time differences (τ_c and τ_l), the gravity vector in $\mathscr{F}_b$ at the starting time ($\boldsymbol{g}^{b_0}$), the IMU biases (${}^{a}\boldsymbol{b}$ and ${}^{\omega}\boldsymbol{b}$), the inverse depths of all the 3D visual points ($\boldsymbol{d}$), the surfel parameters ($\boldsymbol{\Pi}$), and the control points of the trajectories (${}^{R}\boldsymbol{c}$ and ${}^{p}\boldsymbol{c}$), resulting in a compact vector,

$$\boldsymbol{x} = [{\boldsymbol{R}_c^b}^T, {\boldsymbol{p}_c^b}^T, {\boldsymbol{R}_l^b}^T, {\boldsymbol{p}_l^b}^T, \tau_c, \tau_l, {\boldsymbol{g}^{b_0}}^T, {}^{a}\boldsymbol{b}^T, {}^{\omega}\boldsymbol{b}^T, \boldsymbol{d}^T, \boldsymbol{\Pi}^T, {}^{R}\boldsymbol{c}^T, {}^{p}\boldsymbol{c}^T]^T.$$

With the error terms defined above, the concrete loss function can be formulated as,

$$\begin{aligned} F(\boldsymbol{x}) = &\sum_{m,n} \|{}^{L}\boldsymbol{e}_{m,n}\|_{\Lambda_L} + \sum_{i,j,k} \|{}^{V}\boldsymbol{e}_{i,j,k}\|_{\Lambda_V} \quad (10) \\ &+ \sum_{n,i,k} \|{}^{V}\boldsymbol{e}_{n,i,k}\|_{\Lambda_V} + \sum \|{}^{I}\boldsymbol{e}_a\|_{\Lambda_a} + \sum \|{}^{I}\boldsymbol{e}_\omega\|_{\Lambda_\omega}, \end{aligned}$$

in which $\|\boldsymbol{e}_\alpha\|_{\Lambda_\alpha} = \frac{1}{2}\boldsymbol{e}_\alpha^T \boldsymbol{\Lambda}_\alpha^{-1} \boldsymbol{e}_\alpha, \alpha \in \{L, V, a, \omega\}$; $\boldsymbol{\Lambda}_L$, $\boldsymbol{\Lambda}_V$, $\boldsymbol{\Lambda}_a$, and $\boldsymbol{\Lambda}_\omega$ are the measuring covariances of the LiDAR, the camera, the accelerator, and the gyroscope, respectively. At last, the optimization objective is defined as,

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} F(\boldsymbol{x}). \quad (11)$$

To find the optimal solution, we start from the estimated initial values and resort to the Levenberg-Marquardt algorithm [13, 15] to solve the problem. Specifically, suppose that all the elements of ${}^{L}\boldsymbol{e}_{m,n}$, ${}^{V}\boldsymbol{e}_{i,j,k}$, ${}^{I}\boldsymbol{e}_a$, and ${}^{I}\boldsymbol{e}_\omega$ are rearranged into a stacked function vector $\boldsymbol{f}(\boldsymbol{x})$. Denote $\boldsymbol{H}$ and $\boldsymbol{\delta}$ by,

$$\boldsymbol{H} = \boldsymbol{J}^T \boldsymbol{\Lambda}^{-1} \boldsymbol{J} \quad (12)$$

$$\boldsymbol{\delta} = \boldsymbol{J}^T \boldsymbol{\Lambda}^{-1} \boldsymbol{f}, \quad (13)$$

where $\boldsymbol{J} = \frac{d\boldsymbol{f}(\boldsymbol{x})}{d\boldsymbol{x}}$, $\boldsymbol{\Lambda}$ is the concatenated covariance matrix, and $\boldsymbol{f}$ is the error of the current iteration, respectively. The compact variable $\boldsymbol{x}$ can be updated via,

$$\boldsymbol{x} \leftarrow \boldsymbol{x} \ominus (\boldsymbol{H} + \gamma \boldsymbol{I})^{-1} \boldsymbol{\delta}, \quad (14)$$

where γ is the damping coefficient of the current iteration, $\boldsymbol{I}$ is the identity matrix which has the same dimension as $\boldsymbol{H}$, and $\ominus$ means the minus operation on the manifold.

REFERENCES

[1] Zixuan Bai, Guang Jiang, and Ailing Xu. 2020. LiDAR-camera calibration using line correspondences. *Sensors* 20, 21 (2020), 6319.

[2] Shoubin Chen, Jingbin Liu, Xinlian Liang, Shuming Zhang, Juha Hyyppä, and Ruizhi Chen. 2020. A novel calibration method between a camera and a 3D LiDAR with infrared images. In *Proceedings of IEEE International Conference on Robotics and Automation*. Paris, France, 4963–4969.

[3] Michael Fleps, Elmar Mair, Oliver Ruepp, Michael Suppa, and Darius Burschka. 2011. Optimization based IMU camera calibration. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Francisco, CA, USA, 3297–3304.

[4] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza. 2017. On-manifold preintegration for real-time visual-inertial odometry. *IEEE Transactions on Robotics* 33, 1 (2017), 1–21.

[5] Paul Furgale, Timothy D. Barfoot, and Gabe Sibley. 2012. Continuous-time batch estimation using temporal basis functions. In *Proceedings of IEEE International Conference on Robotics and Automation*. Saint Paul, MN, USA, 2088–2095.

[6] Carlos Guindel, Jorge Beltrán, David Martín, and Fernando García. 2017. Automatic extrinsic calibration for lidar-stereo vehicle sensor setups. In *Proceedings of IEEE International Conference on Intelligent Transportation Systems*. Yokohama, Japan, 1–6.

[7] Ryoichi Ishikawa, Takeshi Oishi, and Katsushi Ikeuchi. 2018. LiDAR and camera calibration using motions estimated by sensor fusion odometry. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. Madrid, Spain, 7342–7349.

[8] Jonathan Kelly and Gaurav S. Sukhatme. 2009. Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration. In *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*. Daejeon, Korea (South), 360–368.

[9] Dongshin Kim, Seunghak Shin, and In So Kweon. 2018. On-Line initialization and extrinsic calibration of an inertial navigation system with a relative preintegration method on manifold. *IEEE Transactions on Automation Science and Engineering* 15, 3 (2018), 1272–1285.

[10] Gunhee Koo, Jaehyeon Kang, Bumchul Jang, and Nakju Doh. 2020. Analytic plane covariances construction for precise planarity-based extrinsic calibration of camera and LiDAR. In *Proceedings of IEEE International Conference on Robotics and Automation*. Paris, France, 6042–6048.

[11] Julius Kümmerle and Tilman Kühner. 2020. Unified intrinsic and extrinsic camera and LiDAR calibration under uncertainties. In *Proceedings of IEEE International Conference on Robotics and Automation*. Paris, France, 6028–6034.

[12] Cedric Le Gentil, Teresa Vidal-Calleja, and Shoudong Huang. 2018. 3D LiDAR-IMU calibration based on upsampled preintegrated measurements for motion distortion correction. In *Proceedings of IEEE International Conference on Robotics and Automation*. Brisbane, QLD, Australia, 2149–2155.

[13] Kenneth Levenberg. 1944. A method for the solution of certain problems in least square. *Quarterly Applied Mathematics* 2, 2 (1944), 164–168.

[14] Jiajun Lv, Jinhong Xu, Kewei Hu, Yong Liu, and Xingxing Zuo. 2020. Targetless calibration of LiDAR-IMU system based on continuous-time batch estimation. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. Las Vegas, NV, USA, 9968–9975.

[15] Donald W. Marquardt. 1963. An algorithm for least-squares estimation of non-linear parameter. *Society for Industrial and Applied Mathematics* 11, 2 (1963), 431–441.

[16] F.M. Mirzaei and S.I. Roumeliotis. 2008. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE Transactions on Robotics* 24, 5 (Oct. 2008), 1143–1156.

[17] Chanoh Park, Peyman Moghadam, Soohwan Kim, Sridha Sridharan, and Clinton Fookes. 2020. Spatiotemporal camera-LiDAR calibration: A targetless and structureless approach. *IEEE Robotics and Automation Letters* 5, 2 (2020), 1556–1563.

[18] Zoltan Pusztai and Levente Hajder. 2017. Accurate calibration of LiDAR-camera systems using ordinary boxes. In *Proceedings of IEEE/CVF International Conference on Computer Vision Workshop*. Venice, Italy, 394–402.

[19] Joern Rehder and Roland Siegwart. 2017. Camera/IMU calibration revisited. *IEEE Sensors Journal* 17, 11 (2017), 3257–3268.

[20] Eung su Kim and Soon-Yong Park. 2019. Extrinsic calibration between camera and LiDAR sensors by matching multiple 3D planes. *Sensors* 20, 1 (2019), 52.

[21] Zachary Taylor and Juan Nieto. 2016. Motion-Based calibration of multimodal sensor extrinsics and timing offset estimation. *IEEE Transactions on Robotics* 32, 5 (2016), 1215–1229.

[22] Lipu Zhou, Zimo Li, and Michael Kaess. 2018. Automatic extrinsic calibration of a camera and a 3D LiDAR using line and plane correspondences. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. Madrid, Spain, 5562–5569.

[23] Yufeng Zhu, Chenghui Li, and Yubo Zhang. 2020. Online camera-LiDAR calibration with sensor semantic information. In *Proceedings of IEEE International Conference on Robotics and Automation*. Paris, France, 4970–4976.