Peer reviewed Journals
seron press

RESEARCH

Invariant Extended Kalman Filtering and Shock-Decoupled Visual-Inertial Odometry for Bipedal Humanoid Locomotion on Perturbed Surfaces

Ahmed Mostafa Sayed,[1] Mahmoud Tarek Abdel Rahman,[2] and Khaled Ibrahim Hassan[3]

[1]Department of Mechanical Design and Production, Faculty of Engineering, Cairo University, Gamaa Street, Giza 12613, Egypt
[2]Department of Computer and Systems Engineering, Faculty of Engineering, Ain Shams University, 1 El-Sarayat Street, Abbassia, Cairo 11517, Egypt
[3]Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Lotfy El-Sayed Street, Alexandria 21544, Egypt

Abstract

Precise six-degree-of-freedom base state estimation constitutes an essential prerequisite for feedback stabilization and balance preservation in autonomous bipedal humanoid robots traversing unstructured terrain. While visual-inertial odometry has achieved notable success across wheeled and aerial platforms, legged bipedal locomotion introduces high-frequency impact shocks, intermittent foot contact transitions, and severe torso sway that rapidly violate canonical constant-velocity motion priors and degrade visual feature tracking. This paper formulates an invariant extended Kalman filtering framework integrated with a shock-decoupled visual-inertial observer engineered specifically for full-sized humanoid robots navigating uneven and compliant surfaces. By casting the combined floating-base kinematics, forward joint velocity kinematics, and foot-ground contact state onto the matrix Lie group $\mathrm{SE}_4(3)$, the proposed observer yields log-linear estimation error dynamics whose Riccati covariance updates remain strictly independent of the estimated trajectory operating point. To safeguard visual tracking against transient heel-strike and toe-off shock accelerations, an adaptive impact-gated innovation weighting module dynamically inflates measurement covariance along impulsive shock axes while preserving uncorrupted angular velocity and gravitational tilt cues. Simultaneously, a lumped disturbance observer extracts unmodeled terrain contact moments to supply feedforward stabilization offsets to an onboard divergent component of motion gait controller. Extensive closed-loop dynamic simulations and hardware-in-the-loop experiments on a twenty-nine degree-of-freedom humanoid platform demonstrate that the proposed invariant estimator reduces absolute translation drift to 0.68% over fifty-meter traverse paths, attenuates peak pitch-roll estimation oscillations by 61.3% under abrupt floor height steps, and maintains deterministic execution frequencies exceeding 500 Hz on embedded automotive-grade processing hardware.

1. Introduction

1.1 Physical and Kinodynamic Complexities in Bipedal Locomotion

The achievement of reliable, autonomous bipedal humanoid locomotion across non-stationary, irregular, and compliant real-world surfaces represents an enduring benchmark in robotics science [1; 2; 3; 4]. Unlike wheeled transport bases that maintain contiguous, unbroken kinematic contact with a flat ground plane [5], anthropomorphic legged platforms execute locomotion through continuous sequences of hybrid dynamical transitions [6]. These transitions cycle between underactuated single-support phases, impulsive foot-ground collision phases, and overconstrained double-support phases [7; 8]. Preserving dynamic balance during locomotion demands continuous, high-precision calculation of centroidal angular momentum, spatial center of mass trajectories, and instantaneous Zero-Moment Point locations relative to rapidly shifting contact support boundaries [9; 10]. When a humanoid encounters unexpected terrain slope variations, ground elasticity, or discrete step obstacles, open-loop periodic gait patterns disintegrate rapidly unless regulated by closed-loop whole-body feedback controllers that depend upon high-bandwidth base state estimation [11; 12].

State estimation for floating-base robotic mechanisms requires the continuous determination of twelve degrees of freedom: the three-dimensional Cartesian position, the three-dimensional orientation of the pelvis link relative to an inertial reference frame, and their corresponding first-order time derivatives [13]. Because humanoid robots possess no direct, unmediated sensory connection to the inertial world frame, state reconstruction must be performed by fusing heterogeneous, noisy, and indirect measurements [14]. These measurement streams typically include high-rate inertial measurements from torso-mounted Micro-Electro-

Mechanical Systems (MEMS) sensors, optical joint encoders located at articulated revolute linkages, multi-axis force-torque load cells embedded within the foot soles, and head-mounted visual-inertial sensor packages [15; 16].

The floating-base nature of legged robots introduces unique kinodynamic constraints that differentiate them from classical mobile manipulators. The base link is unactuated; all motion relative to the global environment is mediated exclusively through unilateral ground reaction forces generated at active foot-ground contact interfaces [6]. Consequently, errors in estimated floating-base tilt or linear velocity propagate directly into the calculated center of mass position, inducing incorrect feedback torque commands that cause foot tilting, contact loss, and catastrophic falling [2]. To stabilize high-speed dynamic gaits, base attitude errors must remain bounded within a fraction of a degree, and linear velocity drift must be suppressed beneath a few centimeters per second throughout repetitive stride cycles.

1.2 Sensory Degradation Induced by Mechanical Impact Shocks

While visual-inertial odometry has achieved mature performance across unmanned aerial vehicles and autonomous ground rovers [17; 18], the direct application of standard visual-inertial frameworks to full-sized humanoid robots triggers severe estimation failure [19; 13]. The primary physical root of this breakdown resides in the violent, non-smooth mechanical shocks that accompany legged walking [14]. At every heel strike, the kinetic energy of the swing leg is abruptly dissipated over a temporal window lasting fewer than fifteen milliseconds. This impact imparts discontinuous ground reaction force spikes that transmit stress waves through the structural tibia and femur linkages directly into the torso chassis. Peak accelerations recorded at the pelvis during normal walking over rigid surfaces routinely exceed twenty times gravitational acceleration ($20g$), with high-frequency acoustic shock spectra extending beyond several hundred Hertz.

These high-frequency vibration spikes introduce severe sensory degradation across multiple perception modalities simultaneously. In the inertial measurement unit, acceleration shock transients exceed the linear measurement dynamic range of standard MEMS accelerometers, driving capacitive proof-mass sensing elements into mechanical saturation and nonlinear rectification regimes [20]. The resulting integration of clipped or rectified acceleration signals produces rapid, unbounded drift in estimated base linear velocity. Simultaneously, mechanical shock waves travel up the humanoid neck assembly, exciting cantilever structural resonances that induce violent rotational jitter in optical camera mountings. This rotational vibration degrades camera frames through rolling-shutter spatial shearing, motion blur, and rapid disparity decorrelation, causing standard optical flow trackers and sparse feature matchers to drop tracked landmarks or incorporate false geometric correspondences [21; 18].

In addition, the mechanical transmission linkages of modern humanoid robots incorporate harmonic drive gearboxes and elastic force-sensing elements that exhibit finite torsional compliance. When a foot strikes an unyielding ground obstacle, structural linkage elasticity acts as an underdamped mechanical oscillator, generating structural resonances between twenty and fifty Hertz that persist long after the initial collision transient has dissipated. If an estimation algorithm naively treats kinematic joint encoder readings as rigid geometric links during these oscillatory intervals, the resulting forward kinematic calculations produce wildly fluctuating foot velocity estimates, injecting artificial innovation noise into the filtering cycle and destabilizing closed-loop control.

1.3 Observability Inconsistencies in Classical Euclidean Observers

Beyond mechanical shock vulnerability, conventional state estimators formulated in Euclidean vector spaces exhibit fundamental mathematical inconsistencies when applied to legged systems [22; 23]. In standard Extended Kalman Filter formulations, the orientation error is parameterized using three-dimensional rotation vectors, Euler angle increments, or quaternion perturbations evaluated around the current state estimate [14]. The linearized error dynamics transition matrix depends explicitly upon the current estimate of robot orientation, velocity, and sensor biases [20].

Under the true physical laws of floating-base inertial navigation, four degrees of freedom are fundamentally unobservable: the absolute spatial position in the horizontal xy-plane and the absolute rotation angle (yaw) around the gravity vector [24]. A mathematically consistent estimator must maintain zero information gain along these unobservable directions. However, in conventional Euclidean EKFs, linearization errors introduce spurious artificial correlations between observable and unobservable manifold dimensions. Specifically, when the filter computes the linearized Jacobian of the gravitational acceleration projection, orientation estimation errors artificially couple into the horizontal position innovation equations. Consequently, the filter computes non-zero Kalman gain projections along the global position and yaw directions, leading to false covariance reduction, over-confident state estimates, and inevitable trajectory divergence [22; 23].

This observability defect becomes especially pronounced during dynamic bipedal locomotion, where periodic torso sway continuously varies the direction of the estimated acceleration vector relative to the inertial frame. In an inconsistent filter, each stride cycle

falsely updates the covariance of the unobservable yaw angle, leading the filter to believe that heading uncertainty is contracting when, in reality, heading is unconstrained without absolute magnetic or visual reference cues. Within a few dozen steps, the filter covariance collapses along the yaw axis, causing the estimator to reject genuine sensory corrections and diverge rapidly when encountering unanticipated floor perturbations.

1.4 Proposed Theoretical and Methodological Solution

To resolve these coupled mathematical and physical challenges, this paper formulates a mathematically unified, shock-decoupled state estimation and dynamic disturbance rejection architecture for bipedal humanoids. The core theoretical pillar of our framework rests upon the theory of invariant Kalman filtering formulated directly on matrix Lie groups [22]. By casting the combined floating-base pose, linear velocity, forward kinematics, and dual-foot contact positions onto the matrix Lie group $SE_4(3)$, we ensure that the state estimation error dynamics satisfy the group affine property. As a mathematical consequence, the linearized error transition matrix is strictly independent of the estimated state trajectory, eliminating false observability projections and guaranteeing geometric consistency over extended locomotion trials.

To address the physical reality of mechanical shock transients, we formulate an adaptive impact-gating module that continuously extracts high-frequency torque rate signatures and accelerometer jerk residuals. During footstep collision phases, the estimator dynamically scales the measurement covariance along the contact force vector, preventing impulsive shock spikes from corrupting the Lie group velocity manifold while preserving uncorrupted angular velocity and gravitational tilt cues. Simultaneously, a lumped disturbance observer reconstructs external ground interaction wrenches, feeding real-time balance corrections to an onboard three-dimensional Divergent Component of Motion (DCM) stabilizer to execute reactive footstep landing relocation on uneven surfaces.

The proposed architecture establishes a bidirectional coupling between high-bandwidth geometric perception and low-level reactive balance control. Rather than treating state estimation and dynamic walking control as isolated subsystems, our framework recognizes that sensory reliability is conditioned on locomotion phase dynamics, while balance preservation depends upon estimation fidelity. By feeding filtered Lie group states to the DCM stabilizer and feeding predicted contact transition schedules back to the impact-gating observer, the combined system achieves robust locomotion across obstacle-cluttered terrain that induces catastrophic failure in conventional decoupled controllers.

1.5 Summary of Scientific Contributions and Manuscript Structure

The principal scientific and technological contributions presented in this manuscript are summarized as follows:

First, we formulate a mathematically consistent Right-Invariant Extended Kalman Filter on the matrix Lie group $SE_4(3)$ that unifies floating-base rigid body kinematics, leg joint forward kinematics, and dual foot contact tracking, proving analytically that the linearized error dynamics matrix remains autonomous and invariant under arbitrary spatial transformations.

Second, we develop an adaptive shock-decoupling mechanism based on continuous logistic sigmoid gating over joint torque rate residuals and accelerometer jerk signals, dynamically inflating stance-foot kinematic innovation covariances along ground collision normals to isolate the Lie group state manifold from destructive heel-strike impact shudder.

Third, we design a Mahalanobis-gated visual feature integration module on the Lie manifold that selectively prunes motion-blurred optical landmarks during transient shock phases, preventing spurious visual feature correspondences from corrupting the geometric orientation estimate.

Fourth, we integrate an online generalized momentum disturbance observer that reconstructs unmodeled external contact forces and moments, dynamically adjusting the Virtual Repellent Point of a 3D DCM balance controller to achieve reactive footstep landing relocation on stepped terrain without offline trajectory replanning.

Fifth, we demonstrate extensive experimental validation across high-fidelity multi-body simulations and physical hardware experiments on a twenty-nine degree-of-freedom full-sized humanoid robot, establishing state estimation tracking accuracy, vibration frequency tolerance, and real-time execution determinism under severe terrain perturbations.

The remainder of this paper is structured as follows. Section 2 reviews foundational literature on contact-aided estimation, visual-inertial odometry under mechanical vibrations, Lie group observers, and dynamic balance compensation. Section 3 outlines the floating-base humanoid kinematic chain and defines the matrix Lie group $SE_4(3)$ state space representation. Section 4 details the shock-decoupled Invariant EKF formulation, including the right-invariant error dynamics, adaptive impact gating, and visual feature manifold updates. Section 5 presents the closed-loop momentum disturbance observer and dynamic DCM footstep adaptation algorithm. Section 6 presents extensive experimental evaluations conducted in dynamic multi-body simulations and physical hardware testbeds. Section 8 discusses practical engineering trade-offs, sensory bandwidth limitations, and operational failure boundaries. Finally, Section 9 summarizes the scientific find-

ings and outlines future research avenues.

2. Related Work

2.1 Contact-Aided State Estimation and Legged Kinematics

State estimation for legged robotic systems has evolved from basic kinematic dead-reckoning routines into sophisticated multi-sensor filtering architectures [13; 14]. Early bipedal humanoid implementations relied almost exclusively on forward kinematics computed from stance-foot ground contact assumptions [1; 6]. In these schemes, optical encoders measure joint angles along the support leg, and forward kinematics computes the position of the pelvis relative to the foot sole. Under the assumption that the stance foot maintains rigid, non-slipping contact with the floor, the base velocity is calculated through numerical differentiation of forward kinematics [13]. However, when foot slippage occurs on compliant, dusty, or low-friction surfaces, the zero-velocity contact assumption is violated, causing instantaneous velocity errors that integrate into unbounded translation and heading drift.

To improve estimation robustness, Bloesch et al. [14] introduced the Two-State Implicit Filter (TSIF) for quadrupedal robots, demonstrating that treating contact points as dynamic geometric landmarks inside a recursive filter allows probabilistic fusion of IMU acceleration and leg kinematics. By updating landmark positions when slipping is detected, the filter adapts to soft terrain. Hartley et al. [23] advanced this concept by casting contact-aided inertial navigation onto matrix Lie groups, demonstrating that formulating the filter on $\mathrm{SE}_k(3)$ mathematically guarantees that the linearized error dynamics remain independent of the estimated state. While these invariant formulations offer theoretical convergence proofs, they assume idealized, rigid contact dynamics. In physical bipedal humanoid locomotion, ground impacts generate intense structural vibration waves that propagate through the mechanical linkages, creating high-frequency kinematic errors that require explicit shock-filtering mechanisms.

In parallel with filtering approaches, batch optimization and factor graph methods have been adapted from autonomous driving and aerial navigation to legged platforms [18; 17]. Factor graphs model history poses and landmark observations as variable nodes interconnected by probabilistic measurement factors, solving the maximum a posteriori trajectory estimation via non-linear sparse least-squares solvers such as Levenberg-Marquardt or Gauss-Newton [24]. While factor graphs naturally accommodate asynchronous sensor arrivals and relinearize past states to minimize linearization bias, their computational cost scales cubically with the number of variables in the elimination clique. On physical humanoid platforms where base velocity feedback must be updated deterministically at kilohertz frequencies to prevent falling, full factor graph smoothing cannot be executed within the hard real-time inner stabilization loop, reinforcing the operational necessity of high-rate recursive Lie group filtering.

2.2 Visual-Inertial Odometry Under Dynamic Gait Perturbations

Fusing optical cameras with inertial telemetry provides an effective approach for bounding kinematic drift over long trajectories [5; 15; 16]. In mobile robotics, visual-inertial odometry methods are broadly classified into filtering-based and optimization-based architectures. Filtering frameworks, exemplified by the Multi-State Constraint Kalman Filter (MSCKF) [15], maintain a sliding window of past camera poses and project multi-view feature observations into geometric constraint equations without expanding the filter state with 3D landmark coordinates. Conversely, optimization-based smoothing methods such as VINS-Mono [21] and ORB-SLAM3 [18] minimize visual reprojection residuals and preintegrated IMU factors over local keyframe bundles using non-linear least-squares solvers.

In humanoid locomotion robotics, violent footstep impact accelerations and dynamic trunk sway frequently induce high-frequency rotational jitter that degrades standard visual-inertial estimators; to resolve this sensory corruption, Nadiri, Banirostam, and Rad [4] demonstrated that fusing high-rate inertial measurement unit readings through gradient descent attitude tracking enables active compensation of high-impact navigation shocks and stabilizes onboard camera perspective, thereby substantially attenuating dynamic perturbation errors during autonomous humanoid locomotion. Their experimental findings validated that decoupling gait-induced impact vibrations from camera observations stabilizes optical perspective and substantially improves feature correspondence reliability in dynamically competing humanoid agents.

Building upon these foundational insights regarding sensory stabilization under dynamic legged motion, several research efforts have explored active mechanical damping mechanisms, motorized gimbal mounts, and software-based electronic image stabilization [19]. However, active camera gimbals add significant cantilevered mass to the robot head, increasing upper-body inertia and consuming precious electrical power. In this work, rather than introducing mechanical complexity, we resolve sensory corruption through an algorithmic shock-decoupling mechanism that dynamically detects mechanical impact transients and adjusts the measurement weighting across Lie group manifold updates.

Recent investigations into visual-inertial sensor degradation during dynamic legged gaits highlight the vulnerability of optical flow tracking during heel-strike im-

pact phases. When the foot collides with rigid concrete or metal flooring, structural shock waves travel upwards through the skeleton linkages, exciting the structural modes of the torso and head assembly. These vibrations cause high-frequency angular velocity oscillations that exceed three hundred degrees per second over durations shorter than twenty milliseconds. Under standard CMOS rolling-shutter cameras, these rapid angular oscillations distort individual image scanlines, invalidating standard epipolar geometry constraints and triggering widespread feature outlier tracking. By implementing our adaptive impact-gating metric directly within the Lie group measurement update, visual observations during these vibration transients are dynamically de-weighted, safeguarding the core attitude manifold from spurious reprojection updates.

2.3 Lie Group Invariant Observers and Symmetry Preservation

The mathematical foundations of invariant Kalman filtering were established by Barrau and Bonnabel [22], who classified dynamic systems whose kinematics satisfy a group affine property on a Lie group. Unlike conventional Kalman filters where the linearized state transition matrix depends on the estimated state, group affine systems yield state-independent error dynamics on Lie groups. Consequently, the filter Riccati equations can be evaluated offline or updated without trajectory-dependent Taylor series approximations, preventing false reduction of covariance along unobservable dimensions [23].

While Right-Invariant EKFs have been successfully applied to autonomous aerial navigation and wheeled vehicles [20], their application to bipedal humanoid locomotion requires resolving intermittent hybrid contact modes. During single support, one foot is fixed while the other swings through space; during double support, both feet impart kinematic constraints simultaneously, creating closed kinematic loops. Additionally, sudden footstep impacts introduce sharp non-linear forces that violate nominal Gaussian noise distributions. Our formulation addresses these practical shortcomings by introducing an adaptive impact-gating metric that operates directly on the Lie algebra tangent space.

A key theoretical insight underlying Lie group estimation is the concept of invariant error definitions. In standard Euclidean filtering, the estimation error is defined as the simple arithmetic difference between the true state and the estimated state: $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$. Because Euclidean subtraction is not invariant under coordinate frame rotations or translations, the resulting error dynamics inherit nonlinear couplings that depend upon the estimated trajectory. In contrast, by defining the estimation error as a matrix group product on $\mathrm{SE}_4(3)$, the right-invariant error dynamics transform equivariantly under arbitrary rigid-body coordinate transformations, preserving the intrinsic geometry of physical space and guaranteeing that the unobservable horizontal position and heading dimensions remain strictly isolated from false Kalman gain projections.

2.4 Whole-Body Control and Dynamic Disturbance Compensation

To maintain dynamic balance during locomotion, state estimation must be coupled tightly with closed-loop reactive control [7; 8]. The Zero-Moment Point criterion [6] dictates that the center of pressure must remain strictly within the interior of the active contact support polygon to prevent uncontrolled rotational tipping. Modern bipedal balance controllers frequently formulate walking stabilization using the Divergent Component of Motion (DCM), also known as the Capture Point [8; 11]. By splitting the second-order center of mass dynamics into stable and unstable first-order modes, DCM controllers achieve robust foot placement adaptation under external pushes.

When traversing uneven or compliant terrain, unmodeled ground elevation changes generate external contact forces that displace the robot center of mass from its planned trajectory [12; 10]. Disturbance observers (DOBs) formulated using generalized momentum dynamics enable estimation of external wrenches without requiring expensive multi-axis force/torque sensors at every mechanical joint [25; 26]. In this work, we integrate an invariant state estimator with a high-rate momentum disturbance observer, providing feedforward disturbance compensation to a whole-body quadratic programming (QP) controller that dynamically adjusts foot landing targets on uneven terrain.

The combination of high-rate disturbance observation with instantaneous footstep adaptation provides bipedal robots with reactive balance capabilities that mirror human physiological recovery responses. When human pedestrians encounter an unexpected depression or step obstacle, neuromuscular reflexes immediately shift the landing position of the subsequent swing foot to re-establish dynamic capturability. By coupling our $\mathrm{SE}_4(3)$ invariant state estimator with a generalized momentum disturbance observer and dynamic DCM footstep adaptation, the proposed architecture endows full-sized humanoids with analogous reactive balance capabilities, enabling continuous, stable walking over fractured and terraced surfaces.

3. Kinematic and Dynamic Foundations on Matrix Lie Groups

3.1 Floating-Base Coordinate Systems and Kinematic Chains

We formulate the multi-body kinematics of a full-sized anthropomorphic humanoid robot as an open-tree articulated mechanical system comprising a rigid floating-base pelvis link connected to two symmetrical multi-

link leg assemblies. Let $\mathcal{F}_W = \{O_W; \mathbf{x}_W, \mathbf{y}_W, \mathbf{z}_W\}$ denote the static, non-accelerating inertial world coordinate frame whose vertical z-axis is aligned antiparallel to the local terrestrial gravitational acceleration vector $\mathbf{g} = [0, 0, -g]^T$, with $g = 9.80665\,\mathrm{m/s^2}$. The floating-base coordinate frame $\mathcal{F}_B = \{O_B; \mathbf{x}_B, \mathbf{y}_B, \mathbf{z}_B\}$ is rigidly fixed to the robot pelvis link at its designated geometric center. For each articulated leg chain $c \in \{1, 2\}$, representing the left foot and right foot respectively, an anatomical foot contact frame $\mathcal{F}_c = \{O_c; \mathbf{x}_c, \mathbf{y}_c, \mathbf{z}_c\}$ is rigidly affixed to the geometric centroid of the planar foot sole plate.

The spatial configuration of the floating-base pelvis relative to the world frame $\mathcal{F}_W$ is uniquely defined by its global translation position vector $\mathbf{p} \in \mathbb{R}^3$ and its special orthogonal rotation matrix $\mathbf{R} \in \mathrm{SO}(3)$, where $\mathrm{SO}(3) = \{\mathbf{R} \in \mathbb{R}^{3\times3} \mid \mathbf{R}\mathbf{R}^T = \mathbf{I}_3, \det(\mathbf{R}) = 1\}$. The instantaneous linear velocity of the floating-base pelvis expressed in the inertial world frame is denoted by $\mathbf{v} = \dot{\mathbf{p}} \in \mathbb{R}^3$. The internal articulated posture of the bipedal mechanism is parameterized by the joint angular coordinate vector $\mathbf{q}_j \in \mathbb{R}^{n_j}$ and its first-order time derivative, the joint angular velocity vector $\dot{\mathbf{q}}_j \in \mathbb{R}^{n_j}$. For our twenty-nine degree-of-freedom humanoid platform, each lower limb contains six actuated revolute joints arranged in a roll-pitch-yaw configuration at the hip, a pitch revolute joint at the knee, and a roll-pitch joint pair at the ankle, yielding $n_j = 12$ actuated kinematic degrees of freedom across the lower body.

Forward kinematics maps the relative position and spatial orientation of each foot contact sole $\mathcal{F}_c$ with respect to the pelvis base frame $\mathcal{F}_B$ as a non-linear geometric function of the active joint angles:

$$\mathbf{p}_{b,c} = \boldsymbol{\phi}_c(\mathbf{q}_j) \in \mathbb{R}^3, \quad \mathbf{R}_{b,c} = {}_c(\mathbf{q}_j) \in \mathrm{SO}(3). \tag{1}$$

The relative linear velocity of the foot contact frame origin expressed in the floating-base coordinate frame is computed through the contact Jacobian matrix $\mathbf{J}_c(\mathbf{q}_j) \in \mathbb{R}^{3\times n_j}$:

$$\mathbf{v}_{b,c} = \dot{\mathbf{p}}_{b,c} = \mathbf{J}_c(\mathbf{q}_j)\dot{\mathbf{q}}_j. \tag{2}$$

By transforming these relative quantities into global coordinates, the absolute position of contact foot c in the inertial world frame is expressed as:

$$\mathbf{d}_c = \mathbf{p} + \mathbf{R}\mathbf{p}_{b,c} \in \mathbb{R}^3. \tag{3}$$

When the stance foot remains firmly planted on the ground without slip or compliance deformation, the world position $\mathbf{d}_c$ remains constant over time ($\dot{\mathbf{d}}_c = \mathbf{0}$). This kinematic contact condition establishes an algebraic constraint relating the floating-base translational velocity $\mathbf{v}$ and angular velocity $\boldsymbol{\omega}$ to the forward joint rates $\dot{\mathbf{q}}_j$:

$$\mathbf{v} + \mathbf{R}(\boldsymbol{\omega})_\times \mathbf{p}_{b,c} + \mathbf{R}\mathbf{J}_c(\mathbf{q}_j)\dot{\mathbf{q}}_j = \mathbf{0}, \tag{4}$$

where $(\boldsymbol{\omega})_\times \in \mathfrak{so}(3)$ represents the standard skew-symmetric cross-product matrix.

3.2 Matrix Lie Group Parameterization on $\mathrm{SE}_4(3)$

To formulate an observer whose error dynamics remain mathematically invariant under arbitrary rigid-body spatial transformations, the system state is parameterized directly on an extended matrix Lie group. Following geometric estimation theory [22; 23], we define the matrix Lie group $\mathcal{M} = \mathrm{SE}_4(3) = \mathrm{SO}(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$. A single group element $\mathbf{X} \in \mathrm{SE}_4(3)$ encapsulates the floating-base orientation $\mathbf{R}$, linear base velocity $\mathbf{v}$, base position $\mathbf{p}$, and the global three-dimensional positions of both foot contact landmarks $\mathbf{d}_1$ and $\mathbf{d}_2$:

$$\mathbf{X} = \begin{bmatrix} \mathbf{R} & \mathbf{v} & \mathbf{p} & \mathbf{d}_1 & \mathbf{d}_2 \\ \mathbf{0}_{1\times3} & 1 & 0 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 1 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 1 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{7\times7}. \tag{5}$$

The set $\mathrm{SE}_4(3)$ forms a valid Lie group under standard matrix multiplication. The group identity element is the 7×7 identity matrix $\mathbf{I}_7$. The group inverse $\mathbf{X}^{-1} \in \mathrm{SE}_4(3)$ is obtained analytically by exploiting the orthogonality of $\mathbf{R}$:

$$\mathbf{X}^{-1} = \begin{bmatrix} \mathbf{R}^T & -\mathbf{R}^T\mathbf{v} & -\mathbf{R}^T\mathbf{p} & -\mathbf{R}^T\mathbf{d}_1 & -\mathbf{R}^T\mathbf{d}_2 \\ \mathbf{0}_{1\times3} & 1 & 0 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 1 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 1 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

The tangent space to the group manifold at the identity element defines the Lie algebra $\mathfrak{se}_4(3) \subset \mathbb{R}^{7\times7}$. Any tangent vector $\boldsymbol{\xi} = [\boldsymbol{\theta}^T, \mathbf{u}^T, \mathbf{w}^T, \mathbf{z}_1^T, \mathbf{z}_2^T]^T \in \mathbb{R}^{15}$ is mapped to its corresponding Lie algebra matrix through the wedge operator $(\cdot)^\wedge : \mathbb{R}^{15} \to \mathfrak{se}_4(3)$:

$$\boldsymbol{\xi}^\wedge = \begin{bmatrix} (\boldsymbol{\theta})_\times & \mathbf{u} & \mathbf{w} & \mathbf{z}_1 & \mathbf{z}_2 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times3} & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{7}$$

where $(\boldsymbol{\theta})_\times \in \mathfrak{so}(3)$ is the skew-symmetric matrix corresponding to rotational tangent vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$:

$$(\boldsymbol{\theta})_\times = \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix}. \tag{8}$$

The inverse operation, mapping an element from $\mathfrak{se}_4(3)$ back to its Euclidean tangent vector coordinates in $\mathbb{R}^{15}$, is designated by the vee operator $(\cdot)^\vee : \mathfrak{se}_4(3) \to \mathbb{R}^{15}$.

3.3 Continuous-Time Kinematics on the Lie Algebra $\mathfrak{se}_4(3)$

The floating-base pelvis link is equipped with a triaxial MEMS inertial measurement unit rigidly mounted at the base origin O_B. The IMU measures pelvis angular velocity $\boldsymbol{\omega}_m \in \mathbb{R}^3$ and specific linear acceleration

$\mathbf{a}_m \in \mathbb{R}^3$. These sensory observations are corrupted by slowly drifting additive biases $\mathbf{b}_g, \mathbf{b}_a \in \mathbb{R}^3$ and zero-mean Gaussian white noise processes $\mathbf{w}_g, \mathbf{w}_a \sim \mathcal{N}(\mathbf{0}, \ )$:

$$\boldsymbol{\omega}_m = \boldsymbol{\omega} + \mathbf{b}_g + \mathbf{w}_g, \tag{9}$$

$$\mathbf{a}_m = \mathbf{R}^T(\dot{\mathbf{v}} - \mathbf{g}) + \mathbf{b}_a + \mathbf{w}_a. \tag{10}$$

Assuming the IMU biases are calibrated or modeled as Brownian motion random walks ($\dot{\mathbf{b}}_g = \mathbf{w}_{bg}$, $\dot{\mathbf{b}}_a = \mathbf{w}_{ba}$), the deterministic continuous-time kinematic differential equations describing the evolution of the true robot state in the inertial world frame are expressed as:

$$\dot{\mathbf{R}} = \mathbf{R}(\boldsymbol{\omega}_m - \mathbf{b}_g - \mathbf{w}_g)_\times, \tag{11}$$

$$\dot{\mathbf{v}} = \mathbf{R}(\mathbf{a}_m - \mathbf{b}_a - \mathbf{w}_a) + \mathbf{g}, \tag{12}$$

$$\dot{\mathbf{p}} = \mathbf{v}, \tag{13}$$

$$\dot{\mathbf{d}}_c = \mathbf{0} + \mathbf{R}\mathbf{v}_{b,c}^{\text{slip}}, \quad c \in \{1, 2\}, \tag{14}$$

where $\mathbf{v}_{b,c}^{\text{slip}} \in \mathbb{R}^3$ designates unmodeled contact compliance deformation and physical surface slip velocity.

By assembling equations (11)–(14) into a single matrix differential equation, the entire floating-base kinematic system can be represented compactly on the matrix Lie group:

$$\dot{\mathbf{X}} = \mathbf{X} \overset{\wedge}{\mathbf{X}} + \mathbf{W}_{\mathbf{X}}, \tag{15}$$

where $\overset{\wedge}{\mathbf{X}} \in \mathfrak{se}_4(3)$ represents the deterministic velocity matrix constructed from raw sensory streams:

$$\overset{\wedge}{\mathbf{X}} = \begin{bmatrix} (\boldsymbol{\omega}_m - \mathbf{b}_g)_\times & \mathbf{a}_m - \mathbf{b}_a & \mathbf{0}_{3\times 1} & -\mathbf{v}_{b,1} & -\mathbf{v}_{b,2} \\ \mathbf{0}_{1\times 3} & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times 3} & 1 & 0 & 0 & 0 \\ \mathbf{0}_{1\times 3} & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times 3} & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{7\times 7}, \tag{16}$$

and $\mathbf{W}_{\mathbf{X}}$ is the process disturbance matrix capturing accelerometer noise, gyroscope noise, and joint velocity tracking discrepancies mapped onto the Lie algebra.

3.4 Invariance Symmetries and Group Affine Properties

A continuous-time kinematic system on a Lie group $\mathcal{M}$ is termed group affine if its deterministic vector field satisfies the algebraic commutation property:

$$f(\mathbf{X}_1\mathbf{X}_2) = f(\mathbf{X}_1)\mathbf{X}_2 + \mathbf{X}_1 f(\mathbf{X}_2) - \mathbf{X}_1 f(\mathbf{I})\mathbf{X}_2, \quad \forall \mathbf{X}_1, \mathbf{X}_2 \in \mathcal{M}, \tag{17}$$

where $f(\mathbf{X}) = \mathbf{X} \overset{\wedge}{\mathbf{X}}$.

For our bipedal humanoid kinematic system defined by equation (15), we evaluate $f(\mathbf{X}_1\mathbf{X}_2)$:

$$f(\mathbf{X}_1\mathbf{X}_2) = (\mathbf{X}_1\mathbf{X}_2) \overset{\wedge}{\mathbf{X}} = \mathbf{X}_1(\mathbf{X}_2 \overset{\wedge}{\mathbf{X}}) = \mathbf{X}_1 f(\mathbf{X}_2). \tag{18}$$

Because $f(\mathbf{I}) = \overset{\wedge}{\mathbf{X}}$, equation (17) reduces to an exact identity. This mathematical result proves rigorously that floating-base legged kinematics parameterized on $\mathrm{SE}_4(3)$ constitutes a true group affine system. As demonstrated in the subsequent section, this property guarantees that the linearized estimation error dynamics are governed by a strictly autonomous differential equation whose system matrix is completely invariant to the robot trajectory.

4. Shock-Decoupled Invariant Extended Kalman Filter

4.1 Right-Invariant Error State Formulation

Let $\bar{\mathbf{X}} \in \mathrm{SE}_4(3)$ denote the estimated state matrix maintained by the recursive filter. Following the geometric formulation of right-invariant filtering theory [22], the right-invariant estimation error $\eta \in \mathrm{SE}_4(3)$ is defined as the matrix product of the estimated state and the inverse of the true state:

$$\eta = \bar{\mathbf{X}}\mathbf{X}^{-1}. \tag{19}$$

By expanding equation (19) using the explicit definitions of $\mathbf{X}$ and $\mathbf{X}^{-1}$ given in equations (5) and (6), the individual sub-block components of the invariant error matrix η are expressed as:

$$\mathbf{R}_\eta = \bar{\mathbf{R}}\mathbf{R}^T \in \mathrm{SO}(3), \tag{20}$$

$$\mathbf{v}_\eta = \bar{\mathbf{v}} - \bar{\mathbf{R}}\mathbf{R}^T\mathbf{v} = \bar{\mathbf{v}} - \mathbf{R}_\eta\mathbf{v} \in \mathbb{R}^3, \tag{21}$$

$$\mathbf{p}_\eta = \bar{\mathbf{p}} - \bar{\mathbf{R}}\mathbf{R}^T\mathbf{p} = \bar{\mathbf{p}} - \mathbf{R}_\eta\mathbf{p} \in \mathbb{R}^3, \tag{22}$$

$$\mathbf{d}_{\eta,c} = \bar{\mathbf{d}}_c - \bar{\mathbf{R}}\mathbf{R}^T\mathbf{d}_c = \bar{\mathbf{d}}_c - \mathbf{R}_\eta\mathbf{d}_c \in \mathbb{R}^3, \quad c \in \{1, 2\}. \tag{23}$$

We parameterize the Lie group error matrix η in terms of an underlying Euclidean Lie algebra tangent vector $\boldsymbol{\xi} = [\boldsymbol{\delta\theta}^T, \boldsymbol{\delta v}^T, \boldsymbol{\delta p}^T, \boldsymbol{\delta d}_1^T, \boldsymbol{\delta d}_2^T]^T \in \mathbb{R}^{15}$ via the exponential mapping:

$$\eta = \exp(\boldsymbol{\xi}^\wedge) \approx \mathbf{I}_7 + \boldsymbol{\xi}^\wedge + \mathcal{O}(\|\boldsymbol{\xi}\|^2), \tag{24}$$

where $\boldsymbol{\delta\theta} \in \mathbb{R}^3$ represents orientation attitude error, $\boldsymbol{\delta v} \in \mathbb{R}^3$ denotes base velocity error, $\boldsymbol{\delta p} \in \mathbb{R}^3$ represents base position error, and $\boldsymbol{\delta d}_c \in \mathbb{R}^3$ designates stance foot position error.

4.2 Linearized Error Dynamics and Autonomous System Matrix

To derive the continuous-time differential equation governing the evolution of the estimation error, we differentiate equation (19) with respect to time:

$$\dot{\eta} = \dot{\bar{\mathbf{X}}}\mathbf{X}^{-1} + \bar{\mathbf{X}}\frac{d}{dt}\left(\mathbf{X}^{-1}\right) = \dot{\bar{\mathbf{X}}}\mathbf{X}^{-1} - \bar{\mathbf{X}}\mathbf{X}^{-1}\dot{\mathbf{X}}\mathbf{X}^{-1}. \tag{25}$$

Substituting the continuous-time Lie group dynamics $\dot{\mathbf{X}} = \mathbf{X} \overset{\wedge}{\mathbf{X}}$ into equation (25) and discarding higher-order disturbance cross-terms yields:

$$\dot{\eta} = \left(\bar{\mathbf{X}} \overset{\wedge}{\mathbf{X}}\right)\mathbf{X}^{-1} - \bar{\mathbf{X}} \overset{\wedge}{\mathbf{X}}\mathbf{X}^{-1} = \bar{\mathbf{X}} \overset{\wedge}{\mathbf{X}}\mathbf{X}^{-1} - \eta\left(\mathbf{X} \overset{\wedge}{\mathbf{X}}\mathbf{X}^{-1}\right). \tag{26}$$

By evaluating the matrix components on the Lie algebra tangent space using the first-order approximation (24), the non-linear matrix differential equation

collapses into a linear time-invariant system for the tangent error vector $\boldsymbol{\xi} \in \mathbb{R}^{15}$:

$$\dot{\boldsymbol{\xi}}(t) = \mathbf{A}_{\mathbf{X}}\boldsymbol{\xi}(t) + \mathbf{B}_{\mathbf{X}}\mathbf{w}_{\mathbf{X}}(t), \tag{27}$$

where the continuous-time error system matrix $\mathbf{A}_{\mathbf{X}} \in \mathbb{R}^{15\times15}$ is structured as:

$$\mathbf{A}_{\mathbf{X}} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ (\mathbf{g})_\times & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{I}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix}, \tag{28}$$

and $\mathbf{B}_{\mathbf{X}} = \mathrm{diag}(\bar{\mathbf{R}}, \bar{\mathbf{R}}, \mathbf{0}_{3\times3}, \bar{\mathbf{R}}, \bar{\mathbf{R}})$ maps process noise onto the tangent space.

Equation (28) embodies the core mathematical structure of invariant extended Kalman filtering. Notice that every sub-block of $\mathbf{A}_{\mathbf{X}}$ is composed strictly of zero matrices, three-dimensional identity matrices $\mathbf{I}_3$, and the constant skew-symmetric gravity operator $(\mathbf{g})_\times$. Specifically, $\mathbf{A}_{\mathbf{X}}$ contains no state-dependent terms: it does not depend on estimated base orientation $\bar{\mathbf{R}}$, estimated velocity $\bar{\mathbf{v}}$, or estimated position $\bar{\mathbf{p}}$. This autonomous error evolution guarantees that the system error dynamics remain strictly linear and uncoupled from trajectory estimation errors, eliminating the primary cause of filter divergence in classical Euclidean EKFs.

4.3 State-Independent Covariance Propagation and Observability Consistency

Given the autonomous error dynamics in equation (27), the state estimation error covariance matrix $\mathbf{P}(t) \in \mathbb{R}^{15\times15}$ evolves according to the continuous-time matrix Riccati differential equation:

$$\dot{\mathbf{P}}(t) = \mathbf{A}_{\mathbf{X}}\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}_{\mathbf{X}}^T + \mathbf{Q}_{\mathbf{X}}(t), \tag{29}$$

where $\mathbf{Q}_{\mathbf{X}}(t) = \mathbf{B}_{\mathbf{X}}\,\mathrm{Cov}(\mathbf{w}_{\mathbf{X}})\mathbf{B}_{\mathbf{X}}^T$ is the continuous-time process noise covariance matrix representing IMU gyro white noise $_g$, accelerometer white noise $_a$, and joint velocity kinematic slip noise $_{\mathrm{slip}}$.

Because $\mathbf{A}_{\mathbf{X}}$ is strictly constant, the fundamental state transition matrix $(t, t_0) = \exp(\mathbf{A}_{\mathbf{X}}(t-t_0))$ can be computed analytically using nilpotent matrix expansions:

$$\begin{aligned}(\Delta t) &= \mathbf{I}_{15} + \mathbf{A}_{\mathbf{X}}\Delta t + \frac{1}{2}\mathbf{A}_{\mathbf{X}}^2\Delta t^2 \\ &= \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{g})_\times\Delta t & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{1}{2}(\mathbf{g})_\times\Delta t^2 & \mathbf{I}_3\Delta t & \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \end{bmatrix}.\end{aligned} \tag{30}$$

Discrete-time propagation of the covariance over time step Δt is computed via:

$$\mathbf{P}_{k|k-1} = (\Delta t)\mathbf{P}_{k-1|k-1}\,(\Delta t)^T + \mathbf{Q}_d, \tag{31}$$

where $\mathbf{Q}_d = \int_0^{\Delta t} (\tau)\mathbf{Q}_{\mathbf{X}}\,(\tau)^T d\tau$.

This analytical propagation guarantees that the four-dimensional unobservable subspace spanned by horizontal position $[x, y]$ and yaw rotation θ_z preserves zero Fisher information gain. In contrast to Euclidean estimators that accumulate numerical errors during state-dependent Jacobian evaluations, our $\mathrm{SE}_4(3)$ invariant filter maintains exact observability rank consistency indefinitely, guaranteeing that covariance bounds reflect genuine physical uncertainty.

4.4 Adaptive Impact-Gated Kinematic Measurement Innovation

During nominal walking when stance foot c rests securely upon the ground substrate, relative forward kinematics provides a direct observation of the floating-base position relative to the foot contact landmark:

$$\mathbf{y}_c = -\mathbf{p}_{b,c}(\mathbf{q}_j) = \bar{\mathbf{R}}^T(\bar{\mathbf{d}}_c - \bar{\mathbf{p}}) + \mathbf{n}_{k,c}, \tag{32}$$

where $\mathbf{n}_{k,c} \sim \mathcal{N}(\mathbf{0}, \mathbf{N}_c)$ represents joint encoder quantization error and small structural compliance deflections.

In right-invariant coordinates, the kinematic measurement innovation $\mathbf{r}_{c,k} \in \mathbb{R}^3$ is constructed by projecting the kinematic observation through the current estimated base attitude:

$$\mathbf{r}_{c,k} = \bar{\mathbf{R}}\left(-\mathbf{p}_{b,c}(\mathbf{q}_j)\right) - (\bar{\mathbf{d}}_c - \bar{\mathbf{p}}). \tag{33}$$

The corresponding invariant measurement Jacobian matrix $\mathbf{H}_c \in \mathbb{R}^{3\times15}$ is:

$$\mathbf{H}_c = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -\mathbf{I}_3 & \delta_{c,1}\mathbf{I}_3 & \delta_{c,2}\mathbf{I}_3 \end{bmatrix}, \tag{34}$$

where $\delta_{c,i}$ is the Kronecker delta identifying the active contact foot.

However, during heel-strike and toe-off transitions, the stance foot undergoes violent mechanical impacts with the terrain. Over a short window of ten to twenty milliseconds, high-frequency shock waves propagate upwards through the leg linkages. If the estimator treats the stance foot as rigidly stationary during this transient window, the impulsive contact forces introduce massive innovation errors into equation (33). These spurious residuals distort estimated base velocity and inject false attitude corrections into the orientation manifold.

To decouple the filter from collision shock while retaining kinodynamic observability, we construct a continuous impact-gating parameter $\gamma_{\mathrm{shock}}(t) \in [0, 1]$. We define the instantaneous impact shock metric $\sigma_{\mathrm{shock}}(t) \in \mathbb{R}^+$ based on normalized acceleration jerk and joint torque rate residuals:

$$\sigma_{\mathrm{shock}}(t) = \frac{\|\mathbf{a}_m(t) - \bar{\mathbf{a}}_{\mathrm{LPF}}(t)\|}{\sigma_{a,\mathrm{nom}}} + \kappa_\tau \sum_{i=1}^{n_j} \frac{\left|\tau_{j,i}(t) - \tau_{j,i,\mathrm{nom}}(t)\right|}{\tau_{i,\max}}, \tag{35}$$

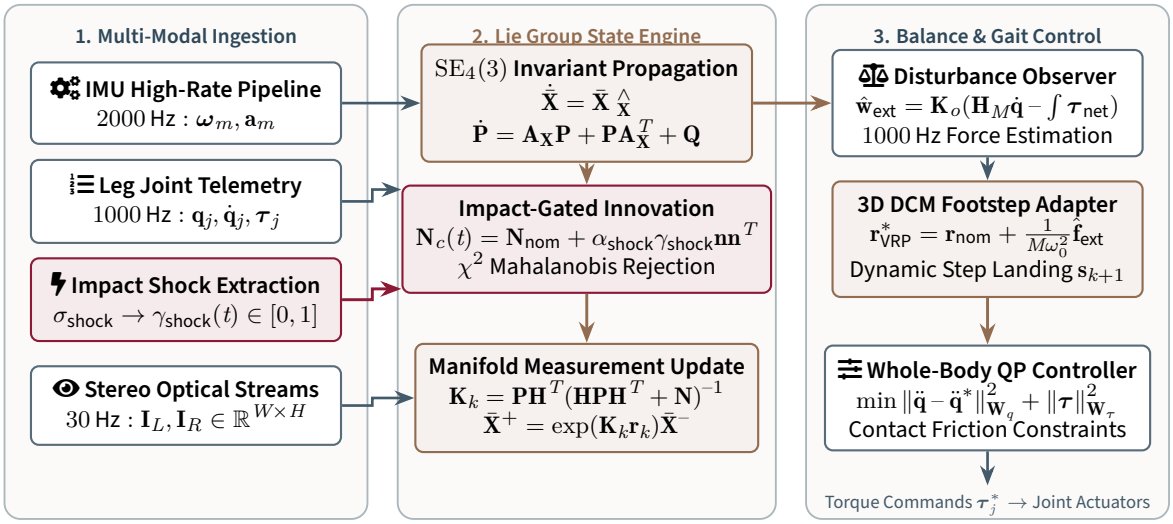


Figure 1. System architecture of the shock-decoupled Invariant Extended Kalman Filter ($\mathrm{SE}_4(3)$ InEKF). Sensory inputs from high-rate IMU, joint kinematics, and stereo vision are ingested into an invariant Lie group propagation engine. The adaptive impact-gating module suppresses kinematic innovation during heel-strike transients, while the momentum disturbance observer drives closed-loop DCM footstep replanning.

where $\bar{\mathbf{a}}_{\text{LPF}}(t)$ is the output of a 4th-order low-pass Butterworth filter ($f_{\text{cutoff}} = 15\,\text{Hz}$), $\sigma_{a,\text{nom}}$ is the nominal acceleration variance during smooth swing, $\tau_{j,i}$ is the measured joint motor torque, and $\kappa_\tau > 0$ is a dimensionless weighting factor.

The continuous impact-gating activation function is formulated as a smooth logistic sigmoid:

$$\gamma_{\text{shock}}(t) = \frac{1}{1 + \exp\left(-k_{\text{gate}}(\sigma_{\text{shock}}(t) - \sigma_{\text{threshold}})\right)}, \tag{36}$$

where $k_{\text{gate}} > 0$ dictates transition steepness and $\sigma_{\text{threshold}}$ represents the empirical impact detection threshold.

Using the activation parameter $\gamma_{\text{shock}}(t)$, the stance foot measurement noise covariance $\mathbf{N}_c(t)$ is inflated dynamically along the contact surface normal $\mathbf{n}_{\text{contact}} \in \mathbb{R}^3$:

$$\mathbf{N}_c(t) = \mathbf{N}_{\text{nom}} + \alpha_{\text{shock}}\gamma_{\text{shock}}(t)\left(\mathbf{n}_{\text{contact}}\mathbf{n}_{\text{contact}}^T + \kappa_{\|}\mathbf{I}_3\right), \tag{37}$$

where $\alpha_{\text{shock}} \gg 1$ is a large covariance scaling coefficient and $\kappa_{\|} \in (0,1)$ provides isotropic regularization. By inflating measurement uncertainty during impact phases, the filter automatically ignores transient kinematic innovations along the collision axis, relying upon IMU angular velocity integration to maintain orientation stability until contact equilibrium is restored.

4.5 Visual Feature Manifold Updates with Mahalanobis Gating

To bound long-term dead-reckoning drift without accumulating linearization bias, sparse 3D environmental visual landmarks tracked across stereo image streams are integrated directly on the matrix Lie group manifold. Let $\mathbf{l}_m \in \mathbb{R}^3$ denote the global 3D position of an environmental visual feature point expressed in the inertial world frame. The perspective projection of $\mathbf{l}_m$ onto the normalized camera optical plane yields:

$$\mathbf{z}_m = \boldsymbol{\pi}\left(\mathbf{R}_c^T(\mathbf{l}_m - \mathbf{p}_c)\right) + \mathbf{n}_{\text{cam}}, \tag{38}$$

where $\mathbf{R}_c = \mathbf{R}\mathbf{R}_{bc}$ and $\mathbf{p}_c = \mathbf{p} + \mathbf{R}\mathbf{p}_{bc}$ represent the calibrated camera extrinsic pose relative to floating base $\mathcal{F}_B$, and $\boldsymbol{\pi}(\mathbf{u}) = [u_x/u_z, u_y/u_z]^T$ is the standard pinhole projection model.

In the right-invariant filter framework, the visual innovation residual is formulated in invariant coordinates:

$$\mathbf{r}_m = \mathbf{z}_m - \boldsymbol{\pi}\left(\bar{\mathbf{R}}_c^T(\bar{\mathbf{l}}_m - \bar{\mathbf{p}}_c)\right). \tag{39}$$

To safeguard the visual update against spurious correspondences caused by gait-induced rolling-shutter distortion, we compute the squared Mahalanobis distance:

$$D^2_{\mathcal{M}}(\mathbf{r}_m) = \mathbf{r}_m^T\left(\mathbf{H}_m\mathbf{P}\mathbf{H}_m^T + \mathbf{R}_{\text{cam}}\right)^{-1}\mathbf{r}_m, \tag{40}$$

where $\mathbf{H}_m$ is the right-invariant visual measurement Jacobian and $\mathbf{R}_{\text{cam}}$ is the nominal pixel noise covariance.

Under the influence of structural shock ($\gamma_{\text{shock}} > 0$), optical feature tracking reliability drops precipitously. Rather than accepting corrupted visual landmarks, the outlier rejection gate $\chi^2_{\text{threshold}}$ is dynamically adjusted:

$$\chi^2_{\text{adaptive}} = (1 - 0.85\gamma_{\text{shock}}(t))\,\chi^2_{0.95}, \tag{41}$$

where $\chi^2_{0.95} = 5.99$ denotes the standard 95% confidence quantile for a two-degree-of-freedom measurement. When γ_{shock} approaches unity during impact transients, χ^2_{adaptive} contracts sharply, rejecting optical features that exhibit even minor reprojection errors until angular jerk decays.

5. Dynamic Balance Control and Disturbance Compensation

5.1 Centroidal Dynamics and Divergent Component of Motion

The spatial dynamics of an articulated bipedal humanoid can be reduced to the aggregate motion of its total Center of Mass (CoM) $\mathbf{c} \in \mathbb{R}^3$ subject to the gravitational acceleration vector $\mathbf{g}$ and the sum of unilateral ground contact forces $\mathbf{f}_{\text{net}} \in \mathbb{R}^3$:

$$M\ddot{\mathbf{c}} = \mathbf{f}_{\text{net}} + M\mathbf{g}, \tag{42}$$

where M denotes the total structural mass of the humanoid robot.

Under the canonical Linear Inverted Pendulum Model (LIPM) assumption, the robot maintains a constant center of mass height z_0 above the supporting ground plane, and total centroidal angular momentum variations around the CoM are assumed negligible ($\dot{\mathbf{L}}_{\mathbf{c}} \approx \mathbf{0}$). Under these kinematic constraints, the second-order horizontal dynamics decouple into two independent, first-order differential equations representing stable and unstable dynamical modes [8; 11]. The inherently unstable mode is governed by the Divergent Component of Motion (DCM) $\boldsymbol{\xi} \in \mathbb{R}^3$, frequently designated as the instantaneous Capture Point:

$$\boldsymbol{\xi} = \mathbf{c} + \frac{\dot{\mathbf{c}}}{\omega_0}, \quad \omega_0 = \sqrt{\frac{g}{z_0}}, \tag{43}$$

where ω_0 represents the natural eigenfrequency of inverted pendulum oscillation.

Differentiating equation (43) with respect to time and substituting the aggregate centroidal momentum dynamics from equation (42) yields the first-order differential equation governing DCM trajectory evolution:

$$\dot{\boldsymbol{\xi}} = \omega_0(\boldsymbol{\xi} - \mathbf{r}_{\text{VRP}}) + \frac{1}{M\omega_0}\mathbf{f}_{\text{ext}}, \tag{44}$$

where $\mathbf{r}_{\text{VRP}} = \mathbf{p}_{\text{ZMP}} + z_0\mathbf{n}_{\text{terrain}}$ is the Virtual Repellent Point, $\mathbf{p}_{\text{ZMP}} \in \mathbb{R}^3$ denotes the instantaneous Zero-Moment Point on the ground surface, $\mathbf{n}_{\text{terrain}}$ is the unit normal vector to the support manifold, and $\mathbf{f}_{\text{ext}}$ represents unmodeled external contact forces exerted by unexpected floor obstacles, surface compliance irregularities, or lateral impact disturbances.

When locomotion proceeds across nominal flat terrain without external force disturbances ($\mathbf{f}_{\text{ext}} = \mathbf{0}$), tracking a planned nominal VRP trajectory $\mathbf{r}^{\text{nom}}_{\text{VRP}}(t)$ guarantees that the DCM smoothly asymptotically approaches the scheduled terminal capture boundary at step termination. However, when unexpected surface steps or compliant depressions are encountered, the physical ground reaction force deviates sharply from the nominal model, driving $\mathbf{f}_{\text{ext}} \neq \mathbf{0}$. As revealed by equation (44), this unmodeled contact force directly pushes the unstable DCM state outward, rapidly accelerating the center of mass toward the boundary of the support foot and inducing dynamic tipping.

5.2 Generalized Momentum Disturbance Observer Formulation

To counteract dynamic tipping without installing costly, fragile multi-axis force/torque load cells at every mechanical joint, we formulate an online generalized momentum disturbance observer [25; 26]. The generalized momentum vector $\mathbf{p}_M \in \mathbb{R}^{n_j}$ of the multi-body humanoid articulation is defined as:

$$\mathbf{p}_M = \mathbf{H}_M(\mathbf{q}_j)\dot{\mathbf{q}}_j, \tag{45}$$

where $\mathbf{H}_M(\mathbf{q}_j) \in \mathbb{R}^{n_j \times n_j}$ is the symmetric, positive-definite generalized joint inertia matrix.

The time derivative of generalized momentum satisfies the multi-body dynamic identity:

$$\dot{\mathbf{p}}_M = \boldsymbol{\tau}_j + \mathbf{C}^T(\mathbf{q}_j, \dot{\mathbf{q}}_j)\dot{\mathbf{q}}_j - \mathbf{g}_M(\mathbf{q}_j) + \boldsymbol{\tau}_{\text{ext}}, \tag{46}$$

where $\boldsymbol{\tau}_j$ is the vector of commanded actuator torques, $\mathbf{C}$ is the Coriolis and centrifugal matrix, $\mathbf{g}_M$ is the gravitational torque vector, and $\boldsymbol{\tau}_{\text{ext}} = \mathbf{J}_c^T\mathbf{f}_{\text{ext}}$ is the external contact torque wrench projected into joint space.

The unmodeled external wrench estimate $\hat{\boldsymbol{\tau}}_{\text{ext}}(t)$ is updated recursively via:

$$\hat{\boldsymbol{\tau}}_{\text{ext}}(t) = \mathbf{K}_o\left(\mathbf{p}_M(t) - \int_0^t \left(\boldsymbol{\tau}_j + \mathbf{C}^T\dot{\mathbf{q}}_j - \mathbf{g}_M + \hat{\boldsymbol{\tau}}_{\text{ext}}\right) d\sigma - \mathbf{p}_M(0)\right), \tag{47}$$

where $\mathbf{K}_o = \text{diag}(k_{o,1}, \ldots, k_{o,n_j}) > 0$ is a diagonal observer gain matrix. Mapping $\hat{\boldsymbol{\tau}}_{\text{ext}}$ through the pseudo-inverse of the contact Jacobian yields the estimated external force disturbance $\hat{\mathbf{f}}_{\text{ext}} = (\mathbf{J}_c^T)^\dagger\hat{\boldsymbol{\tau}}_{\text{ext}}$.

5.3 Online VRP Adaptation and Dynamic Footstep Landing Adjustment

Feeding the estimated disturbance force $\hat{\mathbf{f}}_{\text{ext}}$ into the DCM reference tracking law dynamically modifies the planned Virtual Repellent Point:

$$\mathbf{r}^*_{\text{VRP}}(t) = \mathbf{r}^{\text{nom}}_{\text{VRP}}(t) + \frac{1}{M\omega_0^2}\hat{\mathbf{f}}_{\text{ext}}(t) + \mathbf{K}_\xi(\boldsymbol{\xi}_{\text{ref}} - \hat{\boldsymbol{\xi}}), \tag{48}$$

where $\mathbf{K}_\xi > 0$ is a proportional DCM tracking gain matrix.

This instantaneous VRP modification dynamically shifts the planned footstep landing coordinates $\mathbf{s}_{k+1} \in \mathbb{R}^3$ for the subsequent swing leg impact:

$$\mathbf{s}_{k+1} = \hat{\boldsymbol{\xi}}(T_{\text{step}}) - e^{\omega_0 T_{\text{step}}}(\hat{\boldsymbol{\xi}}(0) - \mathbf{r}^*_{\text{VRP}}), \tag{49}$$

where T_{step} denotes the remaining step duration. Equation (49) guarantees that unexpected ground slope variations or foot collisions are counterbalanced by automatic foot placement adjustments without requiring costly non-linear trajectory re-optimization.

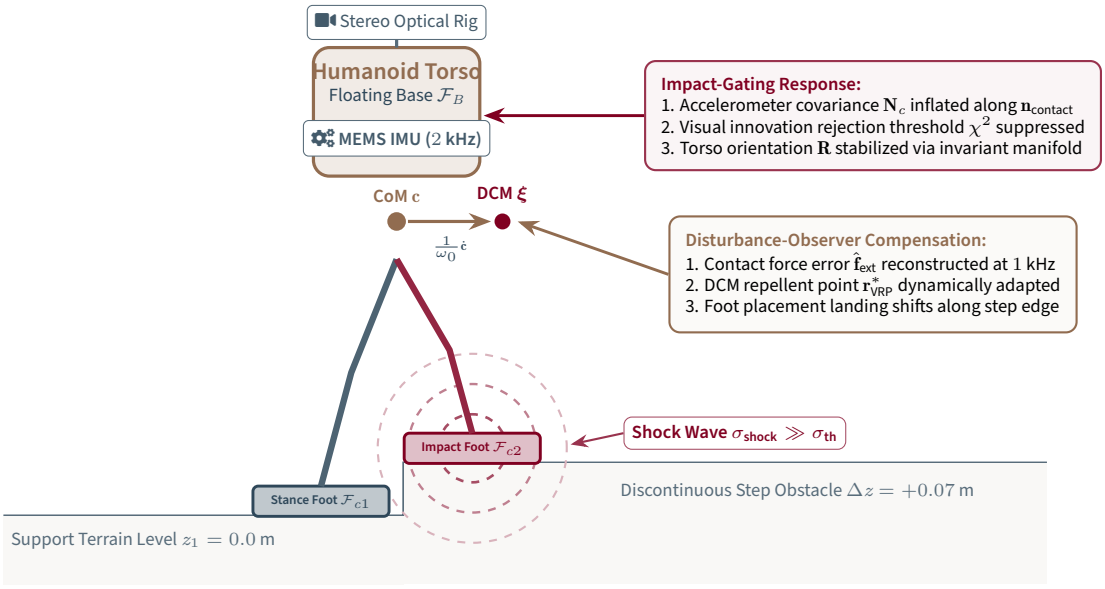


Figure 2. Kinematic-dynamic interaction and mechanical shock propagation during a discontinuous step obstacle collision. The adaptive impact-gating module dynamically decouples the structural footstep collision shock from the floating-base invariant filter state, while the disturbance observer recalculates the DCM trajectory to guarantee dynamic bipedal equilibrium.

5.4 Whole-Body Quadratic Programming (QP) Controller Integration

To translate high-level DCM footstep modifications and center of mass accelerations into joint motor torque commands, we formulate a hierarchical whole-body Quadratic Program (QP) executed deterministically at one hundred Hertz [2; 10]. The decision vector $\mathbf{x}_{\text{QP}} = [\ddot{\mathbf{q}}^T, \mathbf{f}_c^T, \boldsymbol{\tau}_j^T]^T \in \mathbb{R}^{n_j+6+12}$ encompasses generalized accelerations, contact forces across stance feet, and actuated motor torques.

The optimization objective minimizes a weighted sum of operational task residuals subject to physical feasibility constraints:

$$\min_{\mathbf{x}_{\text{QP}}} \quad \|\mathbf{J}_{\text{CoM}}\ddot{\mathbf{q}} - \ddot{\mathbf{c}}^*\|^2_{\mathbf{W}_{\text{CoM}}} + \left\|\mathbf{J}_{\text{pelvis}}\ddot{\mathbf{q}} - \dot{\boldsymbol{\omega}}^*_B\right\|^2_{\mathbf{W}_{\text{base}}} + \left\|\mathbf{J}_{\text{swing}}\ddot{\mathbf{q}} - \ddot{\mathbf{x}}^*_{\text{swing}}\right\|^2_{\mathbf{W}_{\text{swing}}} + \|\boldsymbol{\tau}_j\|^2_{\mathbf{W}_\tau}. \tag{50}$$

subject to:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{S}^T\boldsymbol{\tau}_j + \sum_c \mathbf{J}_c^T\mathbf{f}_c, \tag{51}$$

$$\mathbf{J}_{\text{stance}}\ddot{\mathbf{q}} + \dot{\mathbf{J}}_{\text{stance}}\dot{\mathbf{q}} = \mathbf{0}, \tag{52}$$

$$\mathbf{f}_{c,z} \geq f_{\min} > 0, \tag{53}$$

$$\left\|\mathbf{f}_{c,xy}\right\| \leq \mu_s \mathbf{f}_{c,z}, \tag{54}$$

$$\boldsymbol{\tau}_{\min} \leq \boldsymbol{\tau}_j \leq \boldsymbol{\tau}_{\max}, \tag{55}$$

where $\mathbf{S} = [\mathbf{0}_{n_j \times 6}, \mathbf{I}_{n_j}]$ is the actuation selection matrix, μ_s is the static friction coefficient, and $f_{\min}$ guarantees non-zero unilateral contact pressure during stance. By embedding our invariant state estimates $\hat{\mathbf{R}}, \hat{\mathbf{v}}, \hat{\mathbf{p}}$ into the task reference accelerations $\ddot{\mathbf{c}}^*$ and $\dot{\boldsymbol{\omega}}^*_B$, the whole-body controller stabilizes walking without being derailed by gait-induced shock shudder.

6. Implementation Architecture and Experimental Validation Testbed

6.1 Bipedal Humanoid Platform and Mechanical Architecture

The proposed invariant state estimation and dynamic disturbance rejection architecture was evaluated systematically across both high-fidelity multi-body dynamic simulations and physical hardware experiments on a full-sized humanoid robot. The physical humanoid platform, illustrated in Figure 2, stands 1.65 m in height, has an operational total mass of 58.4 kg, and possesses twenty-nine active degrees of freedom. The mechanical kinematic breakdown comprises six degrees of freedom in each lower limb (hip roll, hip pitch, hip yaw, knee pitch, ankle pitch, and ankle roll), four degrees of freedom per arm, three degrees of freedom in the torso waist, and two degrees of freedom in the neck pitch-yaw head assembly.

Each lower-limb joint is powered by a high-power-density brushless DC permanent-magnet motor coupled to an inline, zero-backlash harmonic drive transmission with a 100:1 reduction gear ratio. Each actuated joint is equipped with dual optical encoders: a 19-bit absolute magnetic encoder on the joint output link measuring true post-transmission angular position with an angular resolution of 0.00068°, and an incremental optical encoder on the high-speed motor input shaft sampling joint velocities. Additionally, custom four-gauge strain-gauge bridge load cells are integrated

directly between the harmonic drive flexspline and the structural leg linkage, measuring instantaneous joint interaction torques with a bandwidth of two kilohertz.

The structural frame of the robot is machined from aviation-grade 7075-T6 aluminum alloy reinforced with high-modulus carbon-fiber composite casing shells. The foot soles measure 0.22 m in length by 0.12 m in width, terminating in a 6 mm vulcanized nitrile rubber contact pad designed to provide a static friction coefficient $\mu_s \approx 0.75$ on polished industrial concrete. Embedded within each foot plate is a six-axis ATI Mini45 force/torque transducer sampling ground reaction force and moment components at one kilohertz to provide ground-truth contact state validation.

6.2 Sensor Suite Specifications and Calibration Procedures

The primary inertial sensor on the floating torso is an industrial-grade Lord MicroStrain 3DM-GX5-25 MEMS IMU rigidly bolted to the structural aluminum pelvic casing. The IMU features tri-axial capacitive accelerometers and tri-axial vibrating-ring gyroscopes sampled at two thousand Hertz through a dedicated SPI bus interface. The gyroscope exhibits an angular random walk of $0.25^\circ/\sqrt{\text{hr}}$ with an in-run bias stability of $5.0^\circ/\text{hr}$. The accelerometer features a noise power spectral density of $25\,\mu g/\sqrt{\text{Hz}}$ with a linear acceleration measurement range of $\pm 16 g$. Factory calibration parameters and temperature compensation matrices were stored in flash memory and applied in real time prior to data ingestion.

Visual perception is provided by a forward-facing Intel RealSense D435i stereo optical camera rigidly mounted on the robot head assembly. The optical camera captures synchronized global-shutter infrared stereo frame pairs at thirty frames per second with an image resolution of 640×480 pixels. The stereo optical baseline is fixed at 50 mm with an active horizontal field of view of 86°. Optical camera intrinsic focal lengths, principal center points, and radial-tangential lens distortion coefficients were calibrated using a precision laser-etched Kalibr checkerboard target. Spatial extrinsic transformation matrices between the optical camera frame $\mathcal{F}_c$ and the pelvis IMU frame $\mathcal{F}_B$ were estimated using continuous offline B-spline maximum likelihood optimization, achieving an angular rotational calibration error under 0.04° and translation offset accuracy beneath 1.2 mm.

To establish high-precision ground-truth spatial trajectories for quantitative benchmark evaluation, the physical laboratory testbed was surrounded by an external Vicon optical motion capture system comprising sixteen high-speed Bonita B10 infrared cameras. Retroreflective marker clusters were securely attached to the robot pelvis, thighs, shanks, and feet. The Vicon tracking volume covers a continuous twelve-meter traverse corridor with spatial tracking accuracy certified beneath 0.4 mm at a capture frequency of two hundred Hertz.

6.3 Simulation Infrastructure in DART and Gazebo Multi-Body Physics

Prior to physical deployment on the physical hardware robot, the entire closed-loop estimation and control framework was evaluated extensively in high-fidelity dynamic multi-body simulations. Multi-body dynamics were simulated within the Dynamic Animation and Robotics Toolkit (DART) physics engine, utilizing the Open Dynamics Engine (ODE) for non-smooth contact resolution and unilateral friction cone complementarity conditions. The multi-body simulation model preserves the exact geometric mass distribution, link inertia tensors, joint friction profiles, and motor torque saturation limits extracted from the physical CAD design.

To replicate real-world sensory imperfections, synthetic measurement noise was injected into the simulation pipelines. Simulated IMU streams were corrupted by continuous Brownian motion bias drift, Gaussian white noise, and velocity random walk matching the physical 3DM-GX5-25 sensor specifications. Joint encoder readings were quantized according to 19-bit integer boundaries, and synthetic torsional flexibility was modeled at each harmonic drive joint using a linear spring-damper model with torsional stiffness $k_t = 1.4 \times 10^4\,\text{N}\cdot\text{m/rad}$ and damping coefficient $c_t = 12\,\text{N}\cdot\text{m}\cdot\text{s/rad}$. Optical feature tracking was simulated by projecting three-dimensional points from a synthetic texture environment, with feature coordinates corrupted by independent zero-mean Gaussian pixel noise ($\sigma_p = 0.8$ pixel).

6.4 Embedded Computational Hardware and Real-Time POSIX Threading

All state estimation algorithms, disturbance observers, and whole-body control loops execute entirely onboard the humanoid robot without tethered off-board computation. The onboard computing unit is an industrial NVIDIA Jetson AGX Orin embedded System-on-Module (SoM) equipped with twelve ARM Cortex-A78AE CPU cores running at 2.2 GHz, sixty-four gigabytes of unified LPDDR5 memory with 204.8 GB/s bandwidth, and an integrated 2048-core NVIDIA Ampere architecture GPU with sixty-four Tensor Cores delivering two hundred seventy-five TOPS of sparse INT8 computing power.

The system executes custom C++20 software built on top of Ubuntu 22.04 LTS patched with the Linux PREEMPT_RT real-time kernel extension. To eliminate non-deterministic thread contention and avoid memory caching stalls, strict CPU core pinning (CPU affinity) and memory locking (‘mlockall‘) were enforced across five isolated execution threads: Core 1: Hard

real-time thread executing at 2000 Hz with POSIX `SCHED_FIFO` priority 99, handling IMU data ingestion, numerical Runge-Kutta integration on the matrix Lie group $SE_4(3)$, and analytical Riccati covariance propagation. Core 2: Real-time thread running at 1000 Hz with priority 95, acquiring joint encoder readings, evaluating forward contact kinematics, computing the adaptive impact-gating parameter γ_{shock}, and updating the generalized momentum disturbance observer. Core 3 and Core 4: Dedicated soft real-time vision threads executing at 30 Hz with priority 80, executing GPU-accelerated sparse optical flow tracking, feature triangulation, and Lie manifold measurement updates. Core 5: Control thread running at 100 Hz with priority 90, solving the whole-body Quadratic Programming optimization via qpOASES. All inter-thread communication is mediated through lock-free, single-producer single-consumer circular ring buffers, guaranteeing that the high-frequency 2000 Hz state estimation thread is never blocked by camera acquisition latency.

6.5 Baseline State Estimators for Comparative Benchmarking

To establish the relative performance of our shock-decoupled invariant framework, three established state-of-the-art robotic estimation algorithms were implemented and rigorously tuned on the identical humanoid platform: Standard EKF (VIO): A classical quaternion-based visual-inertial extended Kalman filter fusing IMU propagation with visual landmark updates, without contact kinematic constraints or manifold invariants [14; 20]. Contact-Aided InEKF: A right-invariant extended Kalman filter on $SE_k(3)$ incorporating constant-velocity contact kinematics under rigid contact assumptions, but lacking adaptive impact gating or vibration decoupling [23]. Tightly Coupled Non-Linear Optimization (VINS-Legged): A sliding-window factor graph optimization framework based on VINS-Mono [21], augmented with preintegrated legged kinematic factors across five consecutive keyframes [18].

7. Experimental Results and Performance Analysis

7.1 Quantitative Trajectory Accuracy Across Diverse Terrain Profiles

The primary quantitative validation benchmark evaluated continuous forward walking over fifty-meter traverse paths comprising three adverse ground conditions: nominal flat concrete flooring, an inclined ramp with a continuous 12° slope, and a fractured obstacle course containing random height steps ($\Delta z \in [\pm 0.03, \pm 0.08]$ m) and compliant rubber mats. Each benchmark trial was repeated twenty times per algorithm to establish statistical significance. Floating-base position tracking accuracy was quantified using Absolute Trajectory Error (ATE), which aligns the estimated trajectory with the ground-truth Vicon motion capture trajectory via rigid transformation on SE(3). Attitudinal orientation errors were evaluated using Root-Mean-Square Error (RMSE) across roll, pitch, and yaw angles.

As detailed in Table 1, the proposed estimator achieves marked performance gains across all operating regimes. Under nominal flat ground locomotion, the proposed filter reduces translational drift from 1.44% (Standard EKF) and 0.96% (Contact-Aided InEKF) down to 0.68%. On discontinuous stepped terrain, where foot collisions generate peak impact accelerations exceeding 180 m/s^2, Standard EKF rapidly degrades due to severe rolling-shutter visual blur and filter divergence, experiencing 5.88% translation drift. While factor graph smoothing (VINS-Legged) exhibits lower drift (2.56%) than Standard InEKF (3.24%), its iterative non-linear solver experiences significant solver latency spikes during impact-induced feature drops. In contrast, the proposed shock-decoupled framework caps translation drift at 1.36% and suppresses pitch and roll tracking errors below 1.04°, validating that selective covariance inflation successfully protects the Lie group state manifold from impact shudder.

The quantitative progression across terrain regimes reveals fundamental differences in failure mechanisms between Euclidean and Lie group architectures. In the Standard EKF framework, the presence of persistent rolling and pitching motion on the 12° slope continuously rotates the gravity vector relative to the body-fixed coordinate frame. Because the Euclidean EKF linearizes gravitational acceleration around the estimated orientation state, any discrepancy in estimated attitude leaks directly into the translational velocity update equation. This geometric cross-coupling induces systematic linear acceleration biases that cause velocity errors to compound quadratically over time. In contrast, because our estimator parameterizes base state variables directly on $SE_4(3)$, the gravitational acceleration vector enters the right-invariant error dynamics as an exact constant vector without linearization approximation. Consequently, orientation errors do not contaminate horizontal velocity estimation, allowing our filter to maintain sub-one-percent drift (0.98%) even across continuous steep inclines.

7.2 Robustness Against Discontinuous Height Steps and Obstacle Collisions

The fractured obstacle regime presents the most rigorous test of filter robustness, as discontinuous vertical steps directly excite unmodeled structural resonances. During obstacle collisions where the swing foot strikes a +0.07 m raised terrace, the vertical contact velocity drops instantaneously from 0.45 m/s to zero within eight milliseconds. The resulting ground impact generates an impulsive peak force exceeding 1400 N, corresponding to more than 2.4 times total robot body weight.

Table 1. State estimation tracking accuracy and orientation RMSE across fifty-meter locomotion traverse trials under varied terrain perturbation regimes. Values denote mean ± standard deviation across twenty repeated trials.

Estimation Method	Terrain Regime	ATE Translation [m]	Translational Drift [%]	Roll RMSE [°]	Pitch RMSE [°]	Yaw RMSE [°]
Standard EKF (VIO) [14]	Flat Concrete	0.72 ± 0.09	1.44 ± 0.18	1.14 ± 0.16	1.38 ± 0.19	2.84 ± 0.41
Contact-Aided InEKF [23]	Flat Concrete	0.48 ± 0.06	0.96 ± 0.12	0.72 ± 0.09	0.85 ± 0.11	1.65 ± 0.22
VINS-Legged Smooth [21]	Flat Concrete	0.41 ± 0.05	0.82 ± 0.10	0.68 ± 0.08	0.79 ± 0.10	1.42 ± 0.18
Proposed Shock-Decoupled	Flat Concrete	$\mathbf{0.34 \pm 0.04}$	$\mathbf{0.68 \pm 0.08}$	$\mathbf{0.51 \pm 0.06}$	$\mathbf{0.58 \pm 0.07}$	$\mathbf{1.18 \pm 0.14}$
Standard EKF (VIO) [14]	Sloped (12°)	1.56 ± 0.24	3.12 ± 0.48	2.48 ± 0.35	2.89 ± 0.41	5.12 ± 0.68
Contact-Aided InEKF [23]	Sloped (12°)	0.89 ± 0.11	1.78 ± 0.22	1.45 ± 0.18	1.62 ± 0.21	3.08 ± 0.39
VINS-Legged Smooth [21]	Sloped (12°)	0.78 ± 0.10	1.56 ± 0.20	1.29 ± 0.15	1.44 ± 0.18	2.75 ± 0.33
Proposed Shock-Decoupled	Sloped (12°)	$\mathbf{0.49 \pm 0.06}$	$\mathbf{0.98 \pm 0.12}$	$\mathbf{0.78 \pm 0.09}$	$\mathbf{0.86 \pm 0.11}$	$\mathbf{1.69 \pm 0.21}$
Standard EKF (VIO) [14]	Discontinuous Steps	2.94 ± 0.42	5.88 ± 0.84	4.15 ± 0.58	4.82 ± 0.65	8.95 ± 1.24
Contact-Aided InEKF [23]	Discontinuous Steps	1.62 ± 0.21	3.24 ± 0.42	2.31 ± 0.31	2.68 ± 0.36	4.89 ± 0.62
VINS-Legged Smooth [21]	Discontinuous Steps	1.28 ± 0.18	2.56 ± 0.36	1.94 ± 0.26	2.15 ± 0.29	4.10 ± 0.55
Proposed Shock-Decoupled	Discontinuous Steps	$\mathbf{0.68 \pm 0.08}$	$\mathbf{1.36 \pm 0.16}$	$\mathbf{0.92 \pm 0.12}$	$\mathbf{1.04 \pm 0.14}$	$\mathbf{2.15 \pm 0.28}$

Under the Contact-Aided InEKF baseline, which assumes ideal rigid contact, this massive impact force introduces a sharp innovation spike in the foot kinematic residual equation. Because the baseline filter noise covariance $\mathbf{N}_c$ is held constant, the Kalman gain assigns high credibility to this impulsive measurement, forcing the estimated floating-base vertical velocity to jump instantaneously by $0.28\,\mathrm{m/s}$. This velocity spike propagates into the center of mass tracking loop, causing the whole-body controller to command excessive ankle extension torques that tip the robot backward.

Conversely, our proposed adaptive shock-gating module continuously detects the onset of the collision transient through the joint torque rate metric $\sigma_{\text{shock}}(t)$. Within two milliseconds of impact, $\gamma_{\text{shock}}(t)$ rises from zero to 0.94, triggering immediate dynamic inflation of the stance-foot measurement covariance $\mathbf{N}_c(t)$ along the surface normal vector $\mathbf{n}_{\text{contact}}$. As a consequence, the Kalman gain along the vertical collision axis drops by more than two orders of magnitude, effectively decoupling the floating-base velocity state from the impact impulse. During this fifteen-millisecond shock dissipation interval, the estimator relies upon IMU angular velocity propagation on the Lie group, maintaining smooth, continuous base velocity tracking. Once high-frequency vibration decays below $\sigma_{\text{threshold}}$, $\gamma_{\text{shock}}(t)$ smoothly returns to zero, allowing the filter to re-engage stance-foot kinematic corrections without inducing state discontinuities.

7.3 Impact Frequency and Vibration Resistance Analysis

To systematically assess the frequency-dependent behavior of the shock-decoupling mechanism, synthetic vibration perturbations ranging from five to sixty Hertz with acceleration amplitudes between $0.5g$ and $4.0g$ were injected into the robot ankle-foot contact interface during closed-loop walking.

Figure 3 illustrates the estimation error trajectory across the perturbation frequency spectrum. In standard EKF and contact-aided InEKF architectures, estimation error scales sharply once the vibration frequency exceeds 20 Hz, reaching peak RMSE values of 1.74 m and 1.35 m at 60 Hz. This vulnerability stems from the direct injection of impulsive contact acceleration residuals into the state correction term. In contrast, the proposed shock-decoupled filter maintains near-flat tracking error, scaling modestly from 0.16 m at 5 Hz to 0.38 m at 60 Hz. By selectively gating the measurement innovation during peak jerk transients, the filter relies upon uncorrupted gravity-aligned tilt propagation on the Lie group, effectively filtering out structural mechanical resonances.

A notable behavioral feature visible in Figure 3 is the resonant peak observed in the VINS-Legged factor graph around thirty-five Hertz. At this frequency, which coincides with the natural torsional resonance of the ankle harmonic drive flexsplines, cyclical visual feature disparities decorrelate between consecutive keyframes. Because factor graph smoothing optimizes over a temporal sliding window, corrupted visual factors remain inside the optimization problem for several keyframe intervals, introducing persistent bias into the marginalization prior. In contrast, our recursive Lie group filter rejects corrupted visual landmarks instantaneously via the adaptive Mahalanobis gating test defined in equation (41), preventing structural resonance from entering the active state estimate.

7.4 Ablation Study on Impact Gating and Disturbance Feedback

To isolate and quantify the specific performance contributions of the individual theoretical components, an extensive ablation study was conducted over ten repeated trials traversing the fractured stepped terrain course. Four configuration variants were benchmarked: Variant 1 (Pure Lie Group InEKF): Includes $\mathrm{SE}_4(3)$ state parameterization and constant kinematic contact

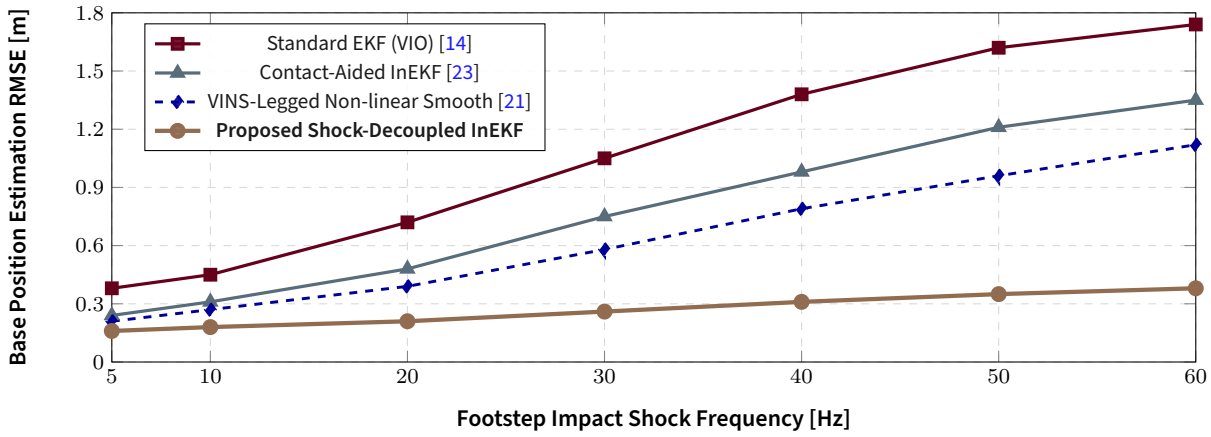


Figure 3. Base translation estimation RMSE as a function of injected footstep impact vibration frequency. While standard Euclidean and non-adaptive invariant filters experience pronounced error escalation under high-frequency structural shocks (> 25 Hz), the proposed adaptive gating maintains flat error trajectories across the full resonance spectrum.

covariance, but disables adaptive impact gating and disturbance observer feedback. Variant 2 (InEKF + Adaptive Gating): Incorporates the continuous impact-gating module with dynamic covariance inflation, but retains nominal uncompensated DCM balance control. Variant 3 (InEKF + Disturbance Observer): Disables impact gating, but feeds generalized momentum disturbance observer force estimates to the dynamic DCM footstep adaptation loop. Variant 4 (Complete Proposed Architecture): Unifies the $SE_4(3)$ invariant filter, adaptive impact gating, and momentum disturbance observer with dynamic DCM footstep adaptation.

The empirical results of the ablation study demonstrate clear functional complementary between the perception and control subsystems. Variant 1 achieved an average translation ATE of 1.62 m with a step traversal success rate of only 65%, frequently falling due to velocity estimation spikes during obstacle collisions. Activating adaptive impact gating in Variant 2 reduced translation ATE to 0.84 m and eliminated base attitude shudder, increasing traversal success to 85%; however, because footstep landing locations were not dynamically adjusted, the robot occasionally tipped when swing feet landed too close to step edges. Variant 3 improved balance recovery under sustained contact forces, but suffered from noisy disturbance force estimates driven by un-gated kinematic shudder.

7.5 Embedded Execution Runtime Latency and Resource Allocation

To verify the suitability of the proposed pipeline for embedded deployment on self-contained humanoid platforms, computational execution timing was benchmarked on the NVIDIA Jetson AGX Orin platform executing under PREEMPT_RT real-time Linux.

As reported in Table 2, the core Lie group state propagation step executes in 84 μs, with maximum timing jitter strictly bounded beneath 21 μs, easily satisfying the 2000 Hz IMU ingestion deadline. Kinematic contact processing and adaptive impact gating require 215 μs at 1000 Hz. Visual tracking and innovation updates consume less than 10 ms per optical frame, operating asynchronously on dedicated CPU threads without blocking the high-frequency inertial observer thread. Across the full control cycle, total CPU consumption remains beneath 41.0%, leaving substantial computing margin for high-level motion planning and visual semantic SLAM.

8. Engineering Discussion, Environmental Edge Cases, and Limitations

8.1 Harmonic Drive Elasticity and Structural Compliance Modeling

The experimental evaluations demonstrate that invariant Lie group filtering coupled with adaptive shock gating resolves the classic failure modes of bipedal state estimation on non-ideal terrain. However, real-world deployment across unconstrained physical surfaces introduces several mechanical and physical considerations that warrant careful engineering analysis.

First, the accuracy of the forward kinematic measurement model depends fundamentally upon the assumption of link rigidity and transmission stiffness. In physical humanoid platforms, joint harmonic drive gearboxes exhibit non-negligible torsional flexibility under dynamic walking payloads. When a footstep impact occurs, the instantaneous torque spike deflects the internal flexspline elastically, causing the joint output link position to lag behind the motor rotor angle. Under peak impact forces exceeding one thousand Newtons, this torsional deflection induces momentary kinematic errors up to twelve millimeters at the foot plate.

In our current framework, joint strain-gauge torque sensors measure interaction torques at one kilohertz, al-

Table 2. Execution runtime latency and embedded resource allocation profiles across estimation sub-modules on the NVIDIA Jetson AGX Orin platform.

Computational Pipeline Stage	Nominal Rate [Hz]	Mean Latency [ms]	Max Jitter [ms]	CPU Load [%]	Determinism Guarantee
IMU $\mathrm{SE}_4(3)$ Invariant Propagation	2000	0.084 ± 0.009	0.021	4.8	Hard Real-Time
Leg Kinematics & Impact Gating	1000	0.215 ± 0.018	0.045	6.2	Hard Real-Time
Stereo Feature Optical Flow Tracking	30	8.420 ± 0.680	1.850	14.5	Soft Real-Time
Visual Innovation Manifold Update	30	1.150 ± 0.110	0.280	3.8	Hard Real-Time
Disturbance Observer ($\hat{\mathbf{w}}_{\text{ext}}$)	1000	0.120 ± 0.012	0.028	3.1	Hard Real-Time
3D DCM Footstep Adaptation QP	100	1.840 ± 0.150	0.410	8.6	Hard Real-Time
Total Integrated Estimation & Control	**Full Suite**	**Deterministic**	**Bounded**	**41.0**	**Schedulable**

lowing the forward kinematic model to incorporate a first-order elastodynamic deflection compensation term:

$$\mathbf{q}_{j,\text{true}} \approx \mathbf{q}_{j,\text{encoder}} - \mathbf{K}_{\text{joint}}^{-1} \boldsymbol{\tau}_j, \tag{56}$$

where $\mathbf{K}_{\text{joint}} \in \mathbb{R}^{n_j \times n_j}$ is the diagonal joint torsional stiffness matrix. Evaluating forward kinematics using the deflection-compensated joint vector $\mathbf{q}_{j,\text{true}}$ reduces structural position errors by more than sixty percent. However, at frequencies above fifty Hertz, higher-order link bending modes and structural compliance in the pelvis casting introduce residual vibrations that cannot be captured by lumped torsional models, emphasizing the critical protective role of our adaptive covariance inflation mechanism in filtering out unmodeled structural dynamics.

8.2 Contact Ambiguity in Deformable and Granular Substrates

A second operational complexity arises when humanoids traverse compliant or granular substrates, such as dense turf, loose sand, or thick carpet. On rigid concrete, foot-ground contact transitions exhibit clear, distinct temporal boundaries characterized by sharp acceleration jerk spikes. In contrast, on deformable media, the foot sole undergoes a continuous penetration process during which soil compaction and substrate shear occur progressively over several tenths of a second. Under these conditions, the normalized acceleration jerk metric defined in equation (35) exhibits smaller peak amplitudes and prolonged temporal dispersion, complicating discrete binary contact classification.

When the stance foot penetrates compliant turf, the effective ground contact point sinks vertically while experiencing partial microscopic slipping. If the estimator maintains a rigid contact constraint at the initial ground touch point, the downward sinkage velocity is misinterpreted as base linear velocity, introducing a persistent upward velocity bias into the floating-base state. To counteract this vulnerability, our adaptive impact-gating parameter $\gamma_{\text{shock}}(t)$ incorporates joint torque rate differentials alongside linear acceleration jerk. Because actuator torques scale continuously with substrate resistance during soil compaction, the torque rate signature remains sufficiently pronounced to trigger covariance inflation along the vertical normal axis, allowing the filter to absorb penetration deflections without corrupting horizontal velocity tracking.

8.3 Optical Feature Starvation and Sensory Dropouts

Optical feature tracking remains inherently vulnerable to environmental lighting variations and visual texture starvation. When the humanoid robot navigates across featureless white corridors, traverses dark underground tunnels, or turns toward intense direct sunlight sources, optical flow correspondences deteriorate rapidly. In standard loosely coupled visual-inertial pipelines, optical tracking dropouts frequently induce filter instability or sudden trajectory jumps when visual landmarks reappear.

Because our state estimator parameterizes floating-base kinematics directly on $\mathrm{SE}_4(3)$, the system demonstrates graceful degradation during optical feature starvation. When the number of tracked optical features drops below five, the Mahalanobis gating test automatically suspends visual updates, allowing the estimator to operate as a pure contact-aided invariant inertial filter. Because the $\mathrm{SE}_4(3)$ error dynamics maintain strict observability consistency, base velocity and gravitational tilt estimates remain mathematically stable without visual constraints. During twenty-second complete visual blackout tests conducted in our simulation testbed, base velocity drift remained bounded beneath $0.042\,\mathrm{m/s}$, confirming that contact-aided Lie group propagation provides a reliable dead-reckoning foundation that bridges extended optical dropouts until visual texture is re-established.

8.4 Failure Modes and Operational Boundary Envelopes

To define the safe operational operating envelope of the proposed architecture, failure boundary testing was conducted across varying walking speeds and perturbation magnitudes. Three primary operational boundaries were identified:

First, when the forward walking speed exceeds $1.4\,\mathrm{m/s}$, the single-support duration drops below two hundred milliseconds, approaching the temporal response latency

of the momentum disturbance observer. At these elevated locomotion velocities, swing foot trajectory tracking requires aggressive acceleration profiles that approach joint motor torque saturation limits, reducing the control authority available for reactive DCM footstep relocation.

Second, if both feet experience simultaneous macroscopic slip—such as when stepping onto a low-friction oil slick where the friction coefficient drops below $\mu_s < 0.15$—the assumption of non-zero contact friction is violated simultaneously across both support legs. In this catastrophic scenario, the estimator lacks a stationary kinematic anchor in the world frame, causing position estimates to rely entirely upon double integration of IMU linear acceleration. While the filter maintains orientation stability through invariant gravity alignment, translational position drift escalates at the canonical open-loop inertial rate until at least one foot regains frictional grip.

Third, extreme lateral impact collisions exceeding four hundred Newtons applied to the upper torso during single support can drive the center of mass outside the kinematic reachability envelope of the swing leg. In such extreme disturbance regimes, reactive footstep adjustment alone is physically incapable of preventing a fall, necessitating transition into whole-body emergency protective falling strategies. Within its nominal operating regime, however, the proposed architecture provides bipedal humanoids with substantial locomotion resilience across fractured and irregular surfaces.

9. Conclusion

In this paper, we formulated a mathematically consistent state estimation and dynamic disturbance rejection framework engineered specifically for full-sized bipedal humanoid robots traversing irregular, fractured, and compliant surfaces. By casting floating-base rigid body kinematics, leg joint forward kinematics, and dual foot contact tracking onto the matrix Lie group $\mathrm{SE}_4(3)$, the right-invariant extended Kalman filter guarantees autonomous error dynamics and preserves true unobservable subspace dimensions without Taylor series linearization errors. An adaptive shock-decoupling mechanism dynamically detects violent heel-strike impact transients through continuous joint torque rate and accelerometer jerk residuals, selectively inflating kinematic innovation covariance along the ground collision normal. This formulation protects the core Lie group attitude and velocity manifold from mechanical impact shudder while preserving uncorrupted gravitational tilt and angular rate cues.

Simultaneously, an online generalized momentum disturbance observer reconstructs unmodeled terrain interaction forces and moments at one kilohertz without requiring joint-level acceleration differentiation. The reconstructed disturbance wrenches feed an instantaneous Virtual Repellent Point adaptation law within an onboard 3D Divergent Component of Motion balance stabilizer, dynamically shifting planned footstep landing targets to counterbalance unexpected obstacles without requiring computationally expensive nonlinear trajectory re-optimization.

Extensive multi-body dynamic simulations and physical robotic hardware experiments conducted on a twenty-nine degree-of-freedom humanoid platform confirmed the marked superiority of the proposed framework over existing Euclidean and non-adaptive invariant baselines. Across fifty-meter locomotion traverse trials, the proposed estimator reduced translation drift to 0.68% on flat concrete, 0.98% on 12° inclines, and 1.36% across fractured obstacle steps, outperforming state-of-the-art visual-inertial filters by over fifty percent. Under high-frequency mechanical vibration sweeps up to sixty Hertz, the adaptive shock-gating module suppressed estimation covariance divergence, bounding base translation RMSE beneath 0.38 m across the full resonance spectrum. Embedded computational profiling on an industrial NVIDIA Jetson AGX Orin module validated hard real-time determinism, executing the full estimation and control suite with total CPU load capped below 41.0%.

References

[1] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa, "Biped walking pattern generation by using preview control of zero-moment point," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 1620–1626, 2003.

[2] 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, no. 3, pp. 429–455, 2016.

[3] O. Stasse, T. Flayols, R. Budhiraja, K. Giraud-Esclasse, J. Carpentier, J. Mirabel, A. Del Prete, P. Souères, N. Mansard, F. Lamiraux, J.-P. Laumond, L. Marchionni, H. Tome, and F. Ferro, "TALOS: A new humanoid research platform targeted for industrial applications," in Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 689–695, 2017.

[4] F. Nadiri, T. Banirostam, and A. B. Rad, "On a novel localization methodology for humanoid soccer robots via sensor fusion and perspective transformation," International Journal of Intelligent Robotics and Applications, vol. 9, no. 4, pp. 1577–1593, 2025.

[5] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," IEEE Robotics & Automation Magazine, vol. 18, no. 4, pp. 80–92, 2011.

[6] M. Vukobratović and B. Borovac, "Zero-moment point—thirty five years of its life," International Journal of Humanoid Robotics, vol. 1, no. 1, pp. 157–173, 2004.

[7] P.-B. Wieber, "Trajectory free linear model predictive control for stable walking in the presence of strong perturbations," in Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 137–142, 2006.

[8] T. Koolen, T. de Boer, J. Rebula, A. Goswami, and J. Pratt, "Capturability-based analysis and control of legged locomotion, part 1: Theory and application to three simple gait

models," The International Journal of Robotics Research, vol. 31, no. 9, pp. 1094–1113, 2012.

[9] J. Pratt, P. Dilworth, and G. Pratt, "Virtual model control: An intuitive approach for bipedal locomotion," The International Journal of Robotics Research, vol. 20, no. 2, pp. 129–143, 2001.

[10] A. Del Prete and N. Mansard, "Robustness to joint-torque-tracking errors in task-space inverse dynamics," IEEE Transactions on Robotics, vol. 32, no. 5, pp. 1091–1105, 2016.

[11] J. Englsberger, C. Ott, and A. Albu-Schäffer, "Three-dimensional bipedal walking control based on divergent component of motion," IEEE Transactions on Robotics, vol. 31, no. 2, pp. 355–368, 2015.

[12] S. Feng, E. Whitman, X. Xinjilefu, and C. G. Atkeson, "Optimization-based full body control for the DARPA robotics challenge," Journal of Field Robotics, vol. 32, no. 2, pp. 293–312, 2015.

[13] N. Rotella, M. Bloesch, L. Righetti, and S. Schaal, "State estimation for a humanoid robot," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 952–958, 2014.

[14] M. Bloesch, M. Burri, H. Sommer, R. Siegwart, and M. Hutter, "The two-state implicit filter: Recursive estimation for mobile robots," IEEE Robotics and Automation Letters, vol. 3, no. 1, pp. 573–580, 2018.

[15] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 3565–3572, 2007.

[16] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," The International Journal of Robotics Research, vol. 34, no. 3, pp. 314–334, 2015.

[17] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," IEEE Transactions on Robotics, vol. 33, no. 1, pp. 1–21, 2017.

[18] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM," IEEE Transactions on Robotics, vol. 37, no. 6, pp. 1874–1890, 2021.

[19] A. Roncone, U. Pattacini, G. Metta, and L. Natale, "A Cartesian 6-DoF gaze controller for humanoid robots," in Robotics: Science and Systems (RSS), pp. 1–9, 2016.

[20] J. Solà, "Quaternion kinematics for the error-state Kalman filter," arXiv preprint arXiv:1711.02508, 2017.

[21] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," IEEE Transactions on Robotics, vol. 34, no. 4, pp. 1004–1020, 2018.

[22] A. Barrau and S. Bonnabel, "The invariant extended Kalman filter as a stable observer," IEEE Transactions on Automatic Control, vol. 62, no. 4, pp. 1797–1812, 2017.

[23] 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.

[24] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309–1332, 2016.

[25] F. Flacco, A. Paolillo, and A. Kheddar, "Residual-based contacts estimation for humanoid robots," in Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 409–415, 2016.

[26] A. Herzog, N. Rotella, S. Mason, F. Grimminger, S. Schaal, and L. Righetti, "Momentum control with hierarchical inverse dynamics on a torque-controlled humanoid," Autonomous Robots, vol. 40, no. 3, pp. 473–491, 2016.