

# SUPPLEMENTARY MATERIAL: RACE: REAL-TIME ADAPTIVE CAMERA-INTRINSICS ESTIMATION VIA CONTROL THEORY

**Anonymous authors**

Paper under double-blind review

## 1 SUPPLEMENTARY

This supplementary document presents an extended robustness and generalization analysis of RACE and empirically evaluates its limitations.

### CONTENTS

<b>1</b>	<b>Supplementary</b>	<b>1</b>
1.1	Ablation Study . . . . .	1
1.1.1	Initial-offset . . . . .	2
1.1.2	Thermal Drift . . . . .	5
1.1.3	Plateau Drift . . . . .	5
1.1.4	Measurement Noise . . . . .	5
1.1.5	Combined disturbance robustness . . . . .	5
1.2	Computational Cost . . . . .	5
1.3	Limitations and Future Work . . . . .	18
1.3.1	TartanAir Dataset . . . . .	18

In Section 1.1, we isolate the impact of (a) initial-offset magnitude (Section 1.1.1), (b) continuous thermal drift (Section 1.1.2), (c) sudden parameter jumps or plateau drift (Section 1.1.3), (d) measurement noise (Section 1.1.4), and (e) the simultaneous combination of all disturbances (Section 1.1.5) on convergence speed and reprojection/mapping error. Section 1.2 analyzes the per-frame computational cost required to maintain real-time true intrinsic values. Finally, in Section 1.3, we detail the current approach’s limitations derived from our empirical results and discuss potential remedies, including criteria for pausing or adjusting parameter updates.

### 1.1 ABLATION STUDY

We selected the EuRoC MAV dataset for our ablation study because of its long, continuous sequences, such as MH\_01, V1\_01, and V2\_01, which can effectively showcase RACE’s robustness to diverse real-time calibration disturbances. We plot both convergence trajectories and reprojection (mapping) errors for each of these sequences to support our conclusions. In every ablation study, the baseline initialization applies a uniform 25% offset to all intrinsic parameters.

Note that the black dotted line represents the ground truth intrinsic parameter values; under thermal drift 1.1.2, plateau drift 1.1.3 and combined disturbances 1.1.5, the actual parameters themselves follow the sinusoid or drifted path.

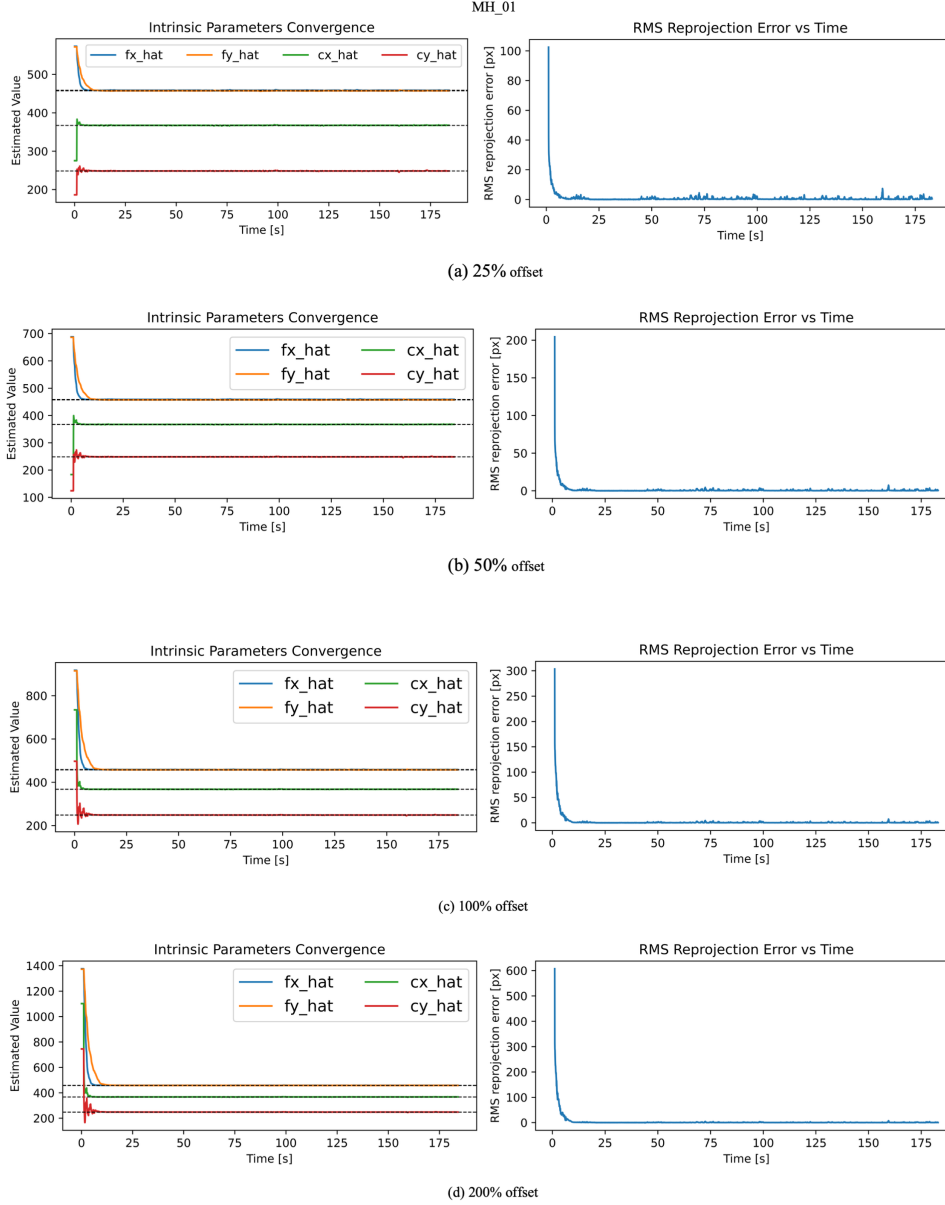


Figure 1: Convergence of Intrinsic parameters and RMS reprojection error for the MH\_01 sequences under initial intrinsic-parameter offsets of (a) 25%, (b) 50%, (c) 100 and (d) 200%. In Intrinsic parameters convergence plots, each colored curve tracks error over time (in seconds), with the ground-truth intrinsics indicated by black dashed lines. Even with a 200% initial\_offset, RACE drives error below 1% within 15s.

### 1.1.1 INITIAL-OFFSET

In this section we provide extended figures and quantitative results for the convergence of RACE under large initial miscalibrations. While the main paper (Appendix C.3) summarizes representative runs, here we include complete plots for each sequence (MH\_01 -1, V1\_01-2, and V2\_01-3) and each offset level ({25%, 50%, 100%, 200%}).

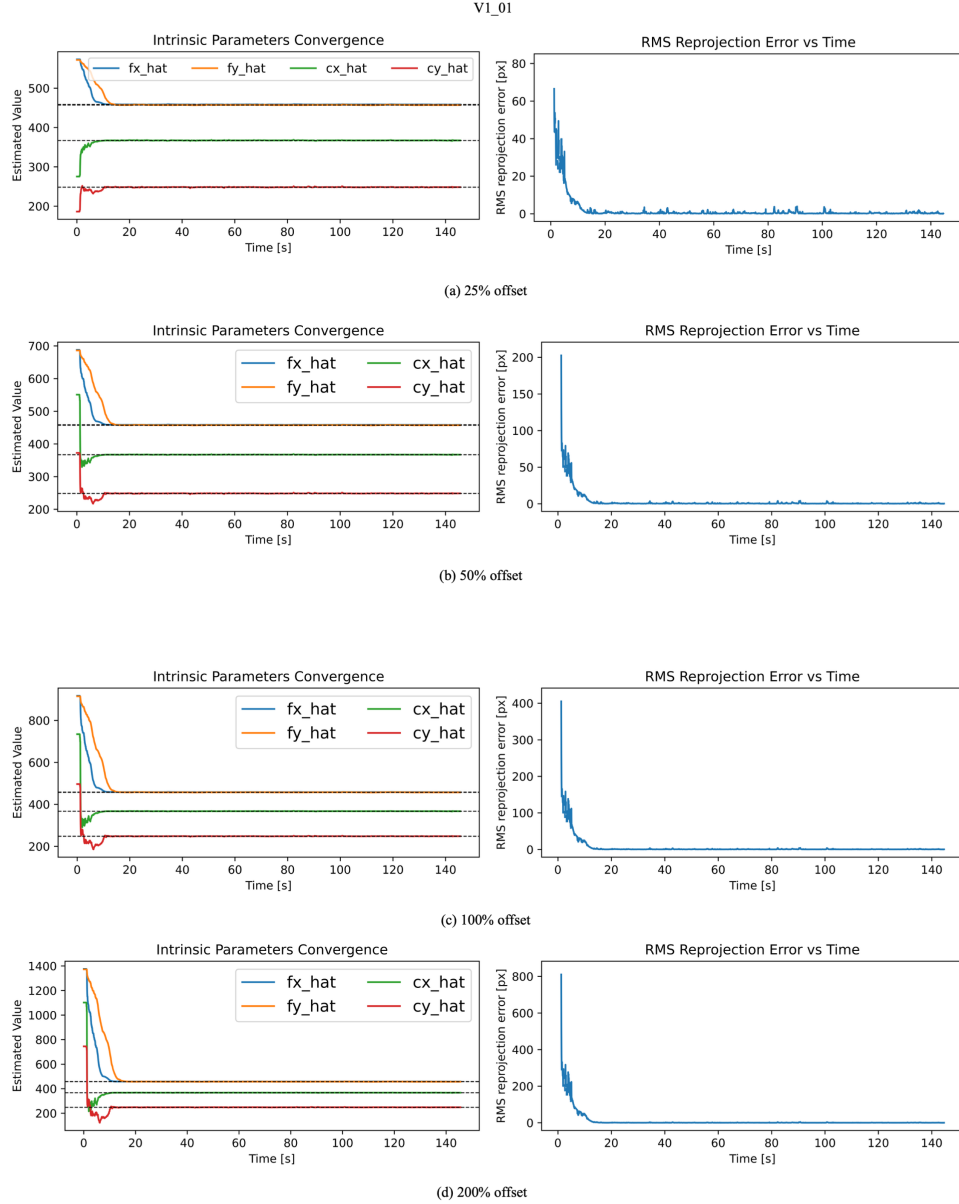


Figure 2: Convergence of Intrinsic parameters and RMS reprojection error for the V1.01 sequences under initial intrinsic-parameter offsets of (a) 25%, (b) 50%, (c) 100 and (d) 200%. In Intrinsic parameters convergence plots, each colored curve tracks error over time (in seconds), with the ground-truth intrinsics indicated by black dashed lines. Even with a 200% initial\_offset, RACE drives error below 1% within 15s.

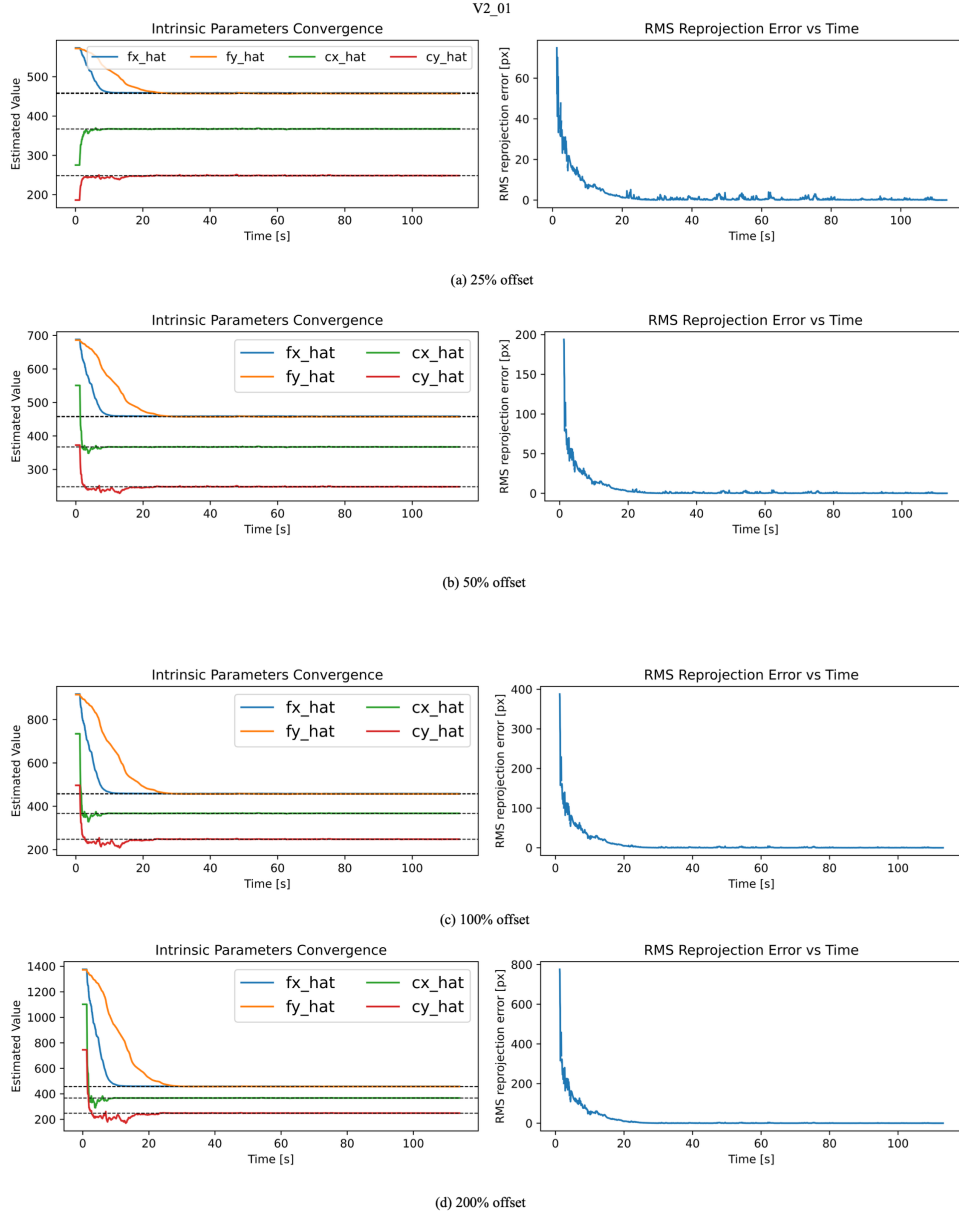


Figure 3: Convergence of Intrinsic parameters and RMS reprojection error for the V2.01 sequences under initial intrinsic-parameter offsets of (a) 25%, (b) 50%, (c) 100 and (d) 200%. In Intrinsic parameters convergence plots, each colored curve tracks error over time (in seconds), with the ground-truth intrinsics indicated by black dashed lines. Even with a 200% initial\_offset, RACE drives error below 1% within 30s.



### 1.1.2 THERMAL DRIFT

Continuous operation of imaging sensors generates heat, causing gradual variation in camera intrinsics. To simulate this effect, we perturb the default 25% offset initialization with a slow sinusoidal drift:  $\theta_t = \theta_0 \cdot (1 + A_{\text{therm}} \sin(2\pi t/T_{\text{therm}}))$ , where  $\theta_0$  is the ground truth intrinsic parameters. We consider a range of amplitude  $A_{\text{therm}} = \{10, 20\} \%$  and  $T_{\text{therm}} = 10\text{s}$  as the drift period.

We apply this thermal perturbation for 50s, see Figures 4, 5, 6 and observe that RACE continuously tracks the drifting intrinsics, maintaining sub-pixel RMS reprojection error across all amplitudes. The parameter estimates adapt smoothly without oscillation or divergence, empirically validating our estimator’s theoretical stability and bounded-error guarantees.

### 1.1.3 PLATEAU DRIFT

A plateau drift refers to a persistent shift in camera parameters occurring over a fixed time interval. In real-world settings, such drift can arise from hardware reconfigurations, lens adjustments, or mechanical shocks. In contrast to slow-varying thermal drift, plateau drift introduces an abrupt and sustained bias, often leading to significant errors in parameter estimation or failure in downstream operations. This type of perturbation is particularly challenging for calibration algorithms that assume gradual and minor variations in intrinsic parameters. To emulate this, we apply step changes of  $\{5\%, 10\%, 20\%\}$  to all intrinsics at  $t = \{30, 50, 80\}\text{s}$ , on top of a 25% initial offset. As shown in Figures 7, 8, 9, each jump produces a spike in reprojection (mapping) error; crucially, RACE re-converges to the new true values within a few frames, demonstrating per-frame adaptability and resilience to abrupt calibration shocks.

### 1.1.4 MEASUREMENT NOISE

To evaluate the effect of continuous pixel-level noise, we inject zero-mean Gaussian perturbations with standard deviations  $\{3, 5\}\text{px}$  into the detected feature coordinates on real EuRoC sequences. As shown in Figure 10, 11, 12, the reprojection (mapping) error exhibits only minor fluctuations and remains uniformly bounded throughout each trial, even under sustained noise injection. In the absence of measurement noise ( $n = 0$ ), RACE converges asymptotically to the true intrinsics, achieving sub-pixel RMS error.

### 1.1.5 COMBINED DISTURBANCE ROBUSTNESS

To evaluate the complete robustness of RACE under real-world conditions, we simultaneously inject multiple sources of disturbance: (i) 25% initial intrinsic offset, (ii) additive continuous zero-mean pixel noise with 0.5 px, (iii) thermal drift of 10% peak amplitude with a 10 s period, and (iv) a plateau drift of 5% applied to all intrinsics at  $t = 30, 60, 90\text{s}$ , see Figures 13, 15, 15. This worst-case scenario emulates concurrent calibration errors that often co-occur in long-duration robotic operations. Remarkably, RACE remains stable, and these findings confirm that our estimator tolerates each disturbance in isolation and generalizes seamlessly to complex, nonlinear interactions among multiple error sources, an essential capability for truly autonomous, long-horizon deployments.

## 1.2 COMPUTATIONAL COST

A key advantage of RACE is that it delivers guaranteed stability, robustness, and convergence of the intrinsic parameters with minimal runtime overhead. Across a diverse set of public benchmark sequences, we measure an average processing time of 8.53 ms per frame on a single CPU core.

As shown in Figure 16, compute time correlates closely with the number of active feature tracks, most frames execute in approximately under 20ms. We observe occasional spikes when the pipeline re-initializes lost tracks or re-detects landmarks, which can exceed 20 ms but do not compromise estimator stability or convergence. To avoid skewing the average with rare but expensive feature-pipeline resets, we compute this mean over all frames whose processing time is below 20ms, treating anything above that threshold as an outlier. Excluding those outliers highlights RACE’s real-time viability for long-horizon robotic perception.

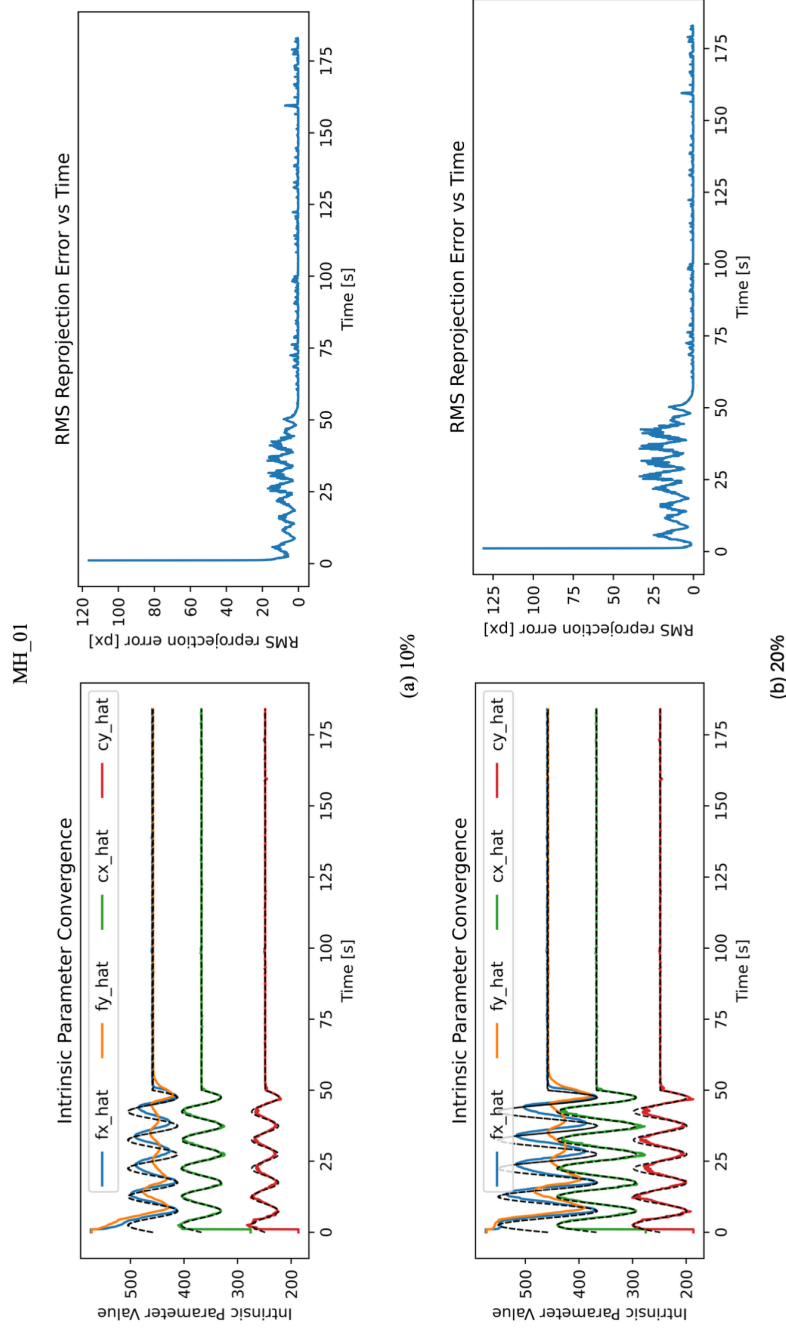


Figure 4: Thermal Drift Tracking on EuRoC MAV (MH\_01). We apply a sinusoidal thermal perturbation of amplitude  $A_{\text{therm}} \in \{5\%, 10\%, 20\%\}$  and period  $T_{\text{therm}} = 10\text{ s}$  to all intrinsics, in addition to default 25% offset initialization. RACE’s estimated intrinsics (solid lines) continuously follow the true drift, yielding sub-pixel RMS error (bottom subplot) throughout the 50s experiment. **Note:** the true intrinsic parameters vary according to the sinusoidal drift model 1.1.2.

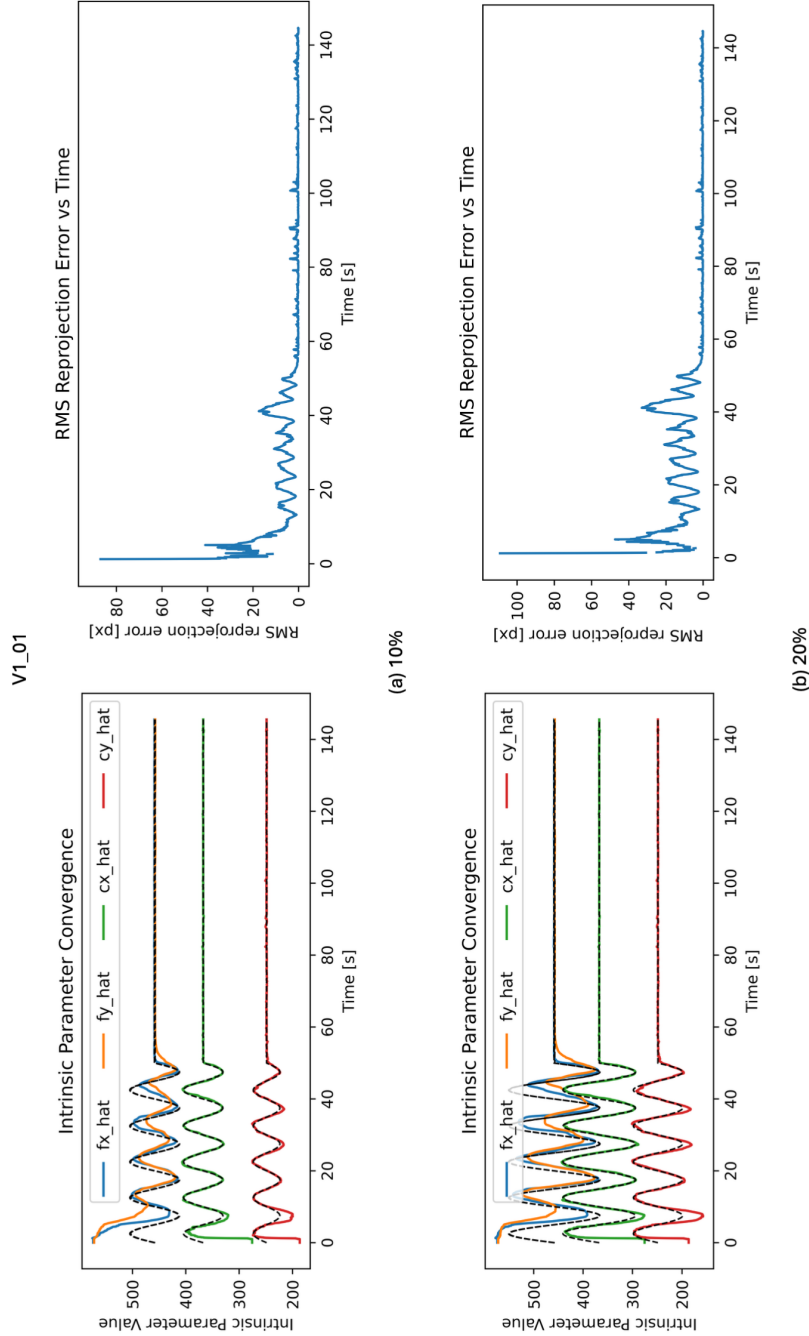


Figure 5: Thermal Drift Tracking on EuRoC MAV (V1\_01). We apply a sinusoidal thermal perturbation of amplitude  $A_{\text{therm}} \in \{5\%, 10\%, 20\%\}$  and period  $T_{\text{therm}} = 10\text{ s}$  to all intrinsics in addition to default 25% offset initialization. RACE’s estimated intrinsics (solid lines) continuously follow the true drift, yielding sub-pixel RMS error (bottom subplot) throughout the 50s experiment. **Note:** the true intrinsic parameters vary according to the sinusoidal drift model 1.1.2.

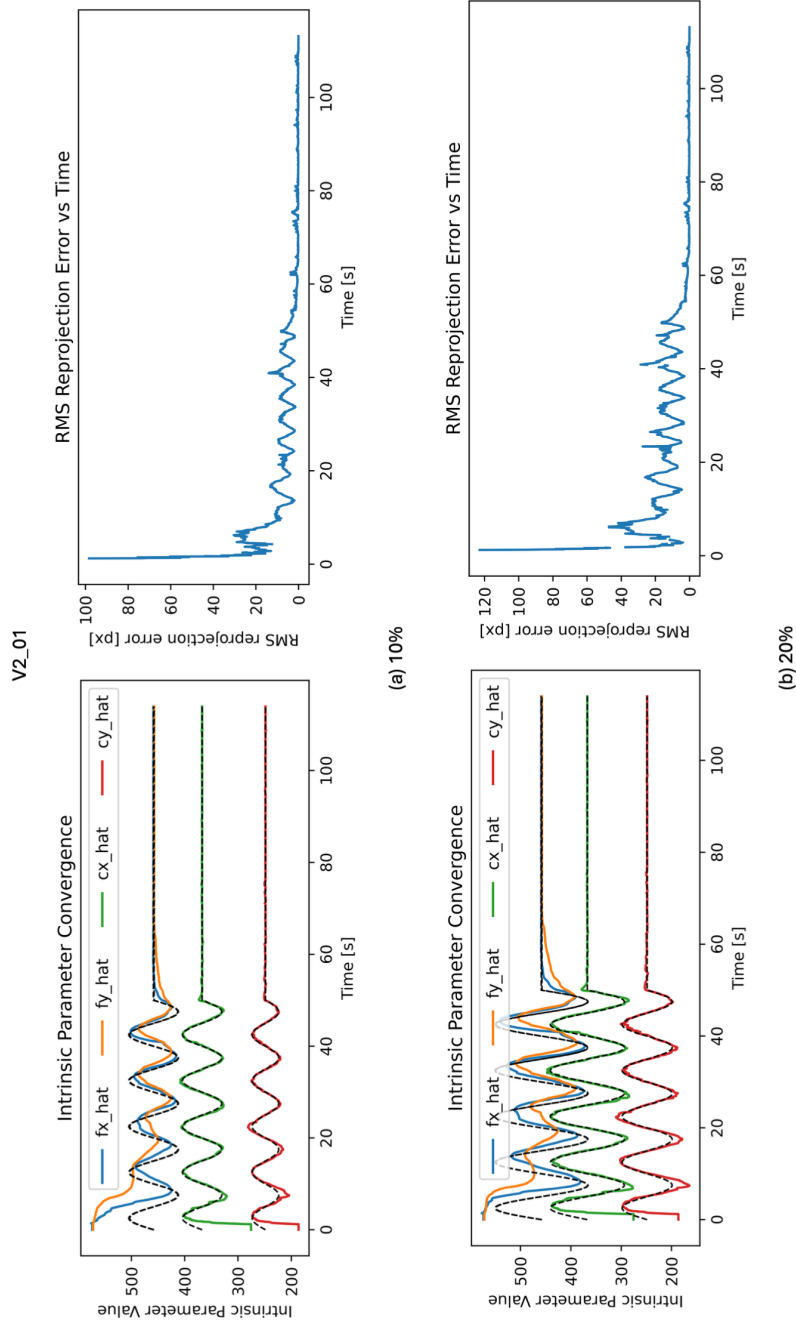
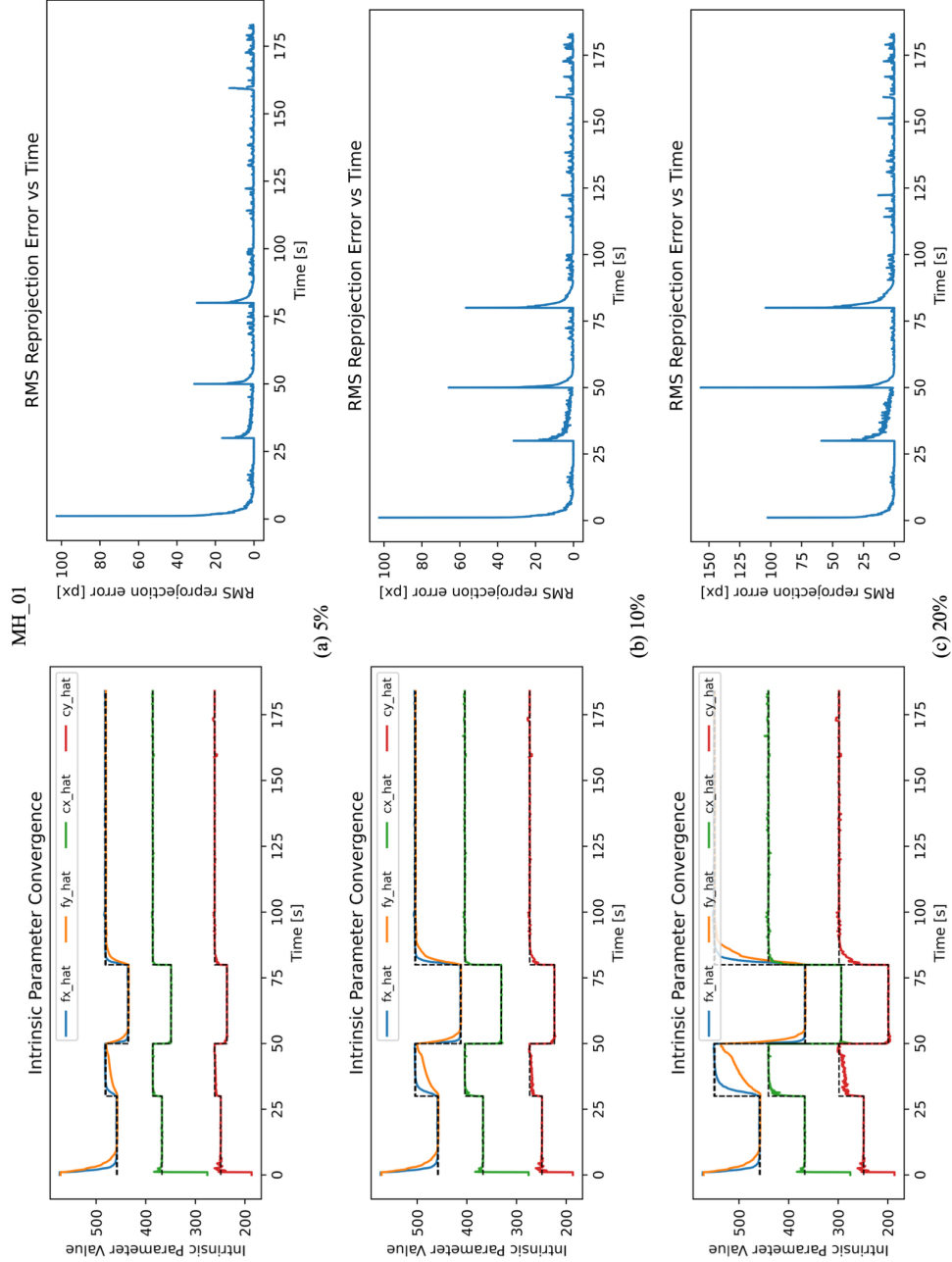


Figure 6: Thermal Drift Tracking on EuRoC MAV (V2\_01). We apply a sinusoidal thermal perturbation of amplitude  $A_{\text{therm}} \in \{5\%, 10\%, 20\%\}$  and period  $T_{\text{therm}} = 10\text{ s}$  to all intrinsics in addition to default 25% offset initialization. RACE’s estimated intrinsics (solid lines) continuously follow the true drift, yielding sub-pixel RMS error (bottom subplot) throughout the 50s experiment. **Note:** the true intrinsic parameters vary according to the sinusoidal drift model 1.1.2.

432



434

480 Figure 7: Plateau drift Recovery on EuRoC MAV (MH.01). At  $t = \{30, 50, 80\}$  s we introduce  
 481 step shifts of  $\{5\%, 10\%, 20\%\}$  in all intrinsics. Each jump produces a sharp spike in ground  
 482 truth intrinsic parameters and reprojection (error) error, but RACE reconverges to the new true values  
 483 within 5-10 frames. This per-frame adaptability underscores its robustness to abrupt calibration  
 484 shocks, a scenario where gradual methods fail catastrophically. **Note:** the true intrinsic parameters  
 485 vary according to the Plateau drift model 1.1.3.

486

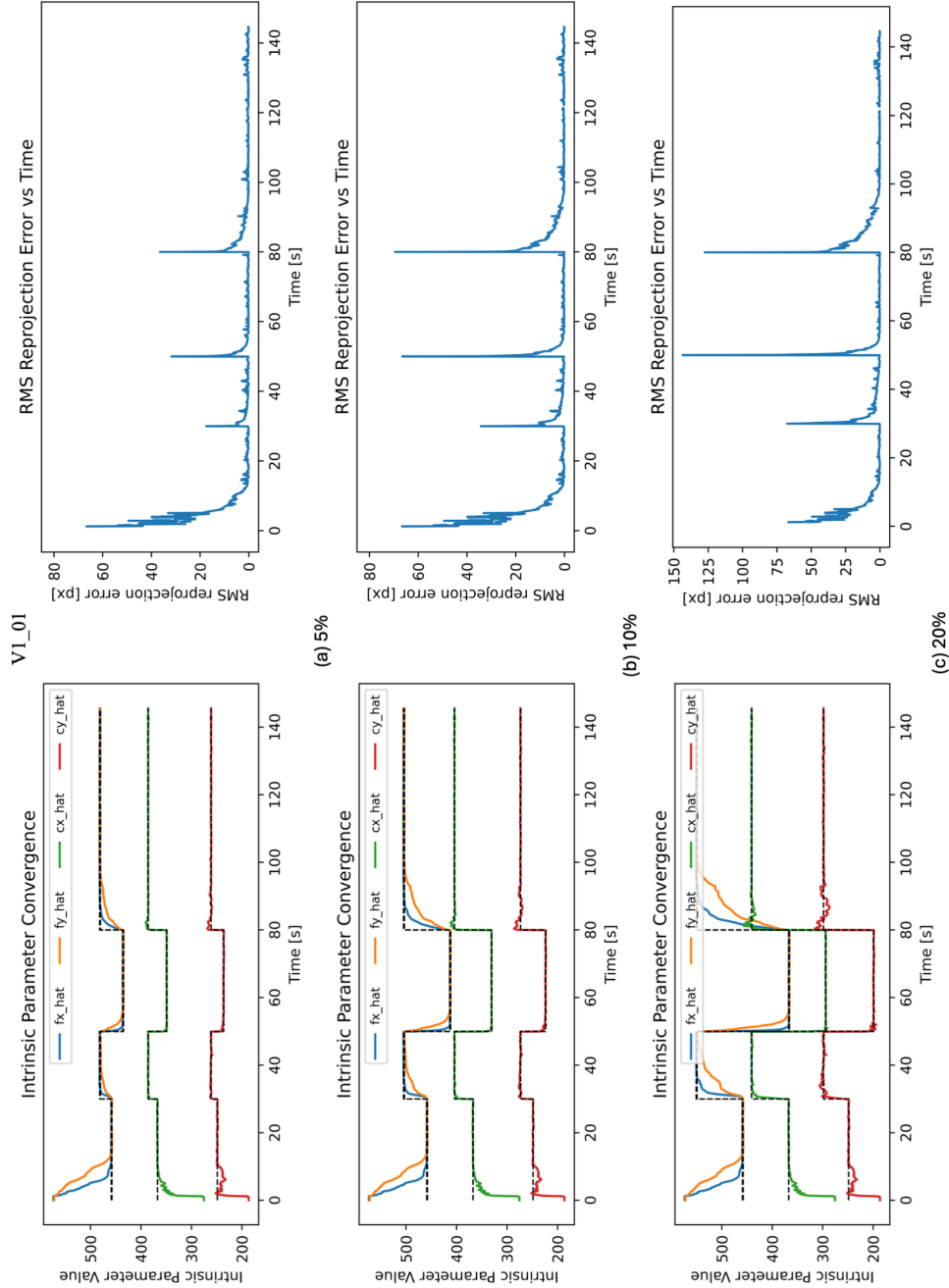
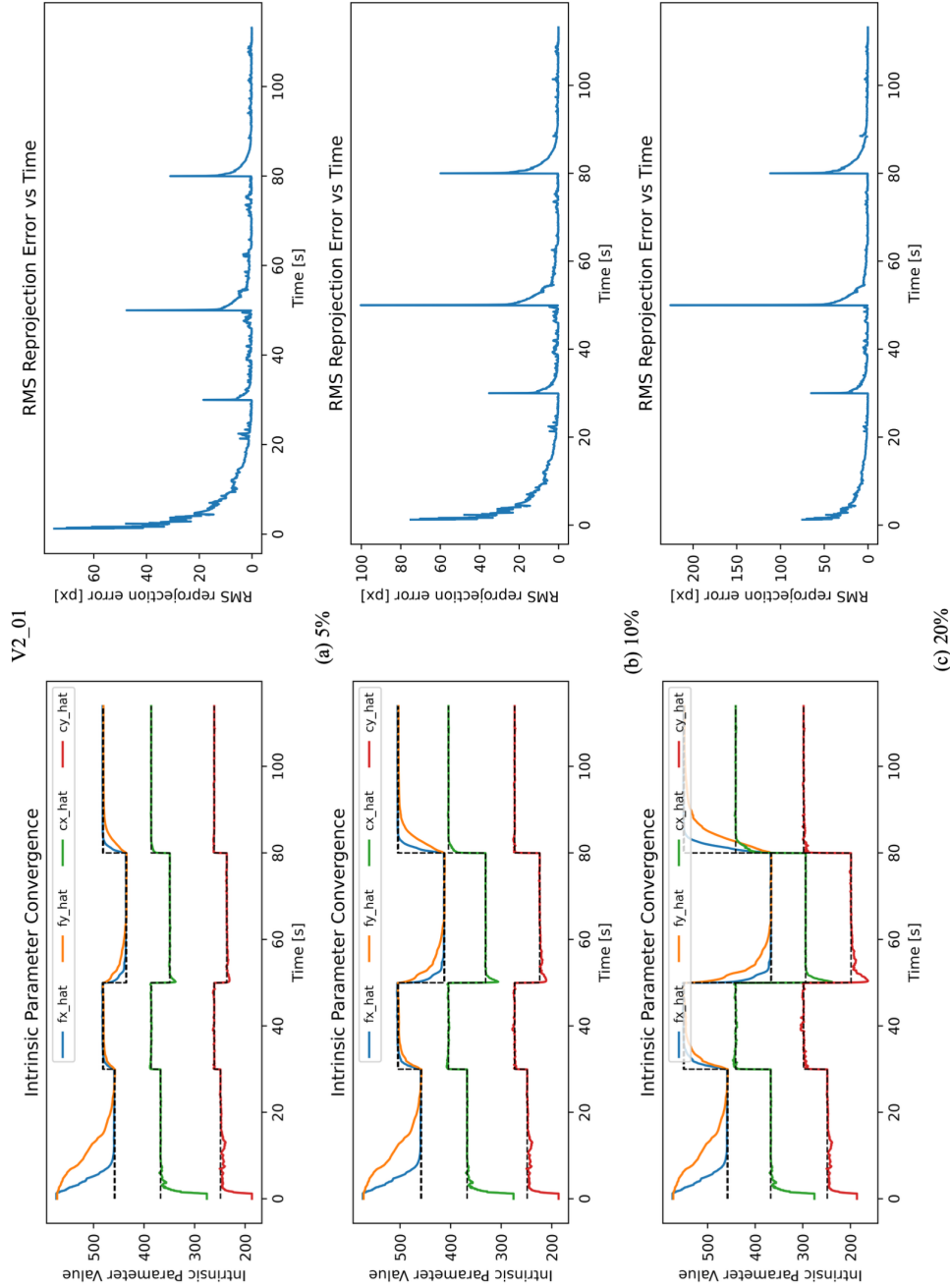


Figure 8: Plateau drift Recovery on EuRoC MAV (V1\_01). At  $t = \{30, 50, 80\}$  s we introduce step shifts of  $\{5\%, 10\%, 20\%\}$  in all intrinsics. Each jump produces a sharp spike in ground truth intrinsic parameters and reprojection (error) error, but RACE reconverges to the new true values within 5-10 frames. This per-frame adaptability underscores its robustness to abrupt calibration shocks, a scenario where gradual methods fail catastrophically. **Note:** the true intrinsic parameters vary according to the Plateau drift model 1.1.3.

540



587

588

589

590

591

592

593

Figure 9: Plateau drift Recovery on EuRoC MAV (V2\_01). At  $t = \{30, 50, 80\}$  s we introduce step shifts of  $\{5\%, 10\%, 20\%\}$  in all intrinsics. Each jump produces a sharp spike in ground truth intrinsic parameters and reprojection (error) error, but RACE reconverges to the new true values within 5-10 frames. This per-frame adaptability underscores its robustness to abrupt calibration shocks, a scenario where gradual methods fail catastrophically. **Note:** the true intrinsic parameters vary according to the Plateau drift model 1.1.3.

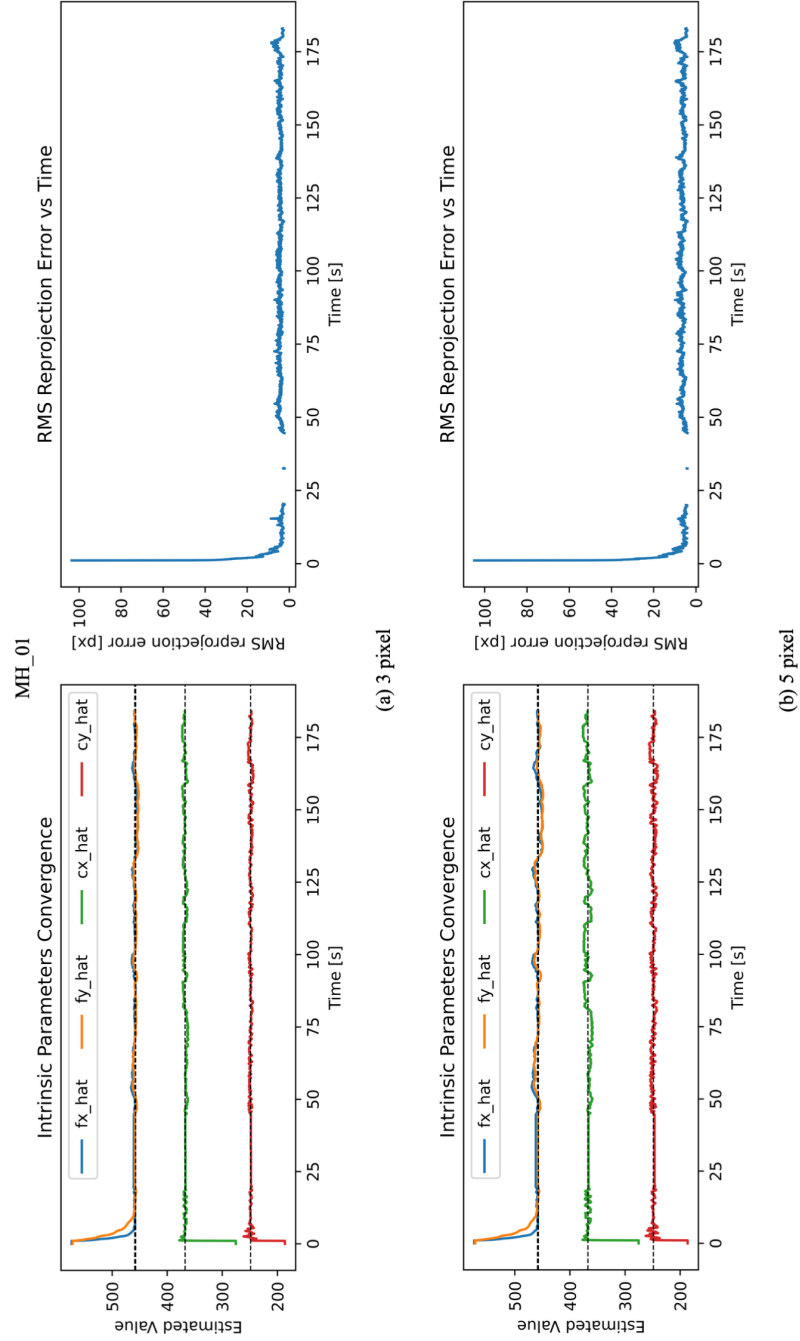


Figure 10: Measurement-Noise Robustness on EuRoC MAV (MH\_01). We inject continuous zero-mean Gaussian noise with  $\{3, 5\}$  px into feature coordinates. Despite continuous noise, RACE keeps RMS error fluctuations within bounded pixel values and never diverges. The increase in the ultimate error bound matches our Theorem 2, confirming quantitative robustness to measurement perturbations.



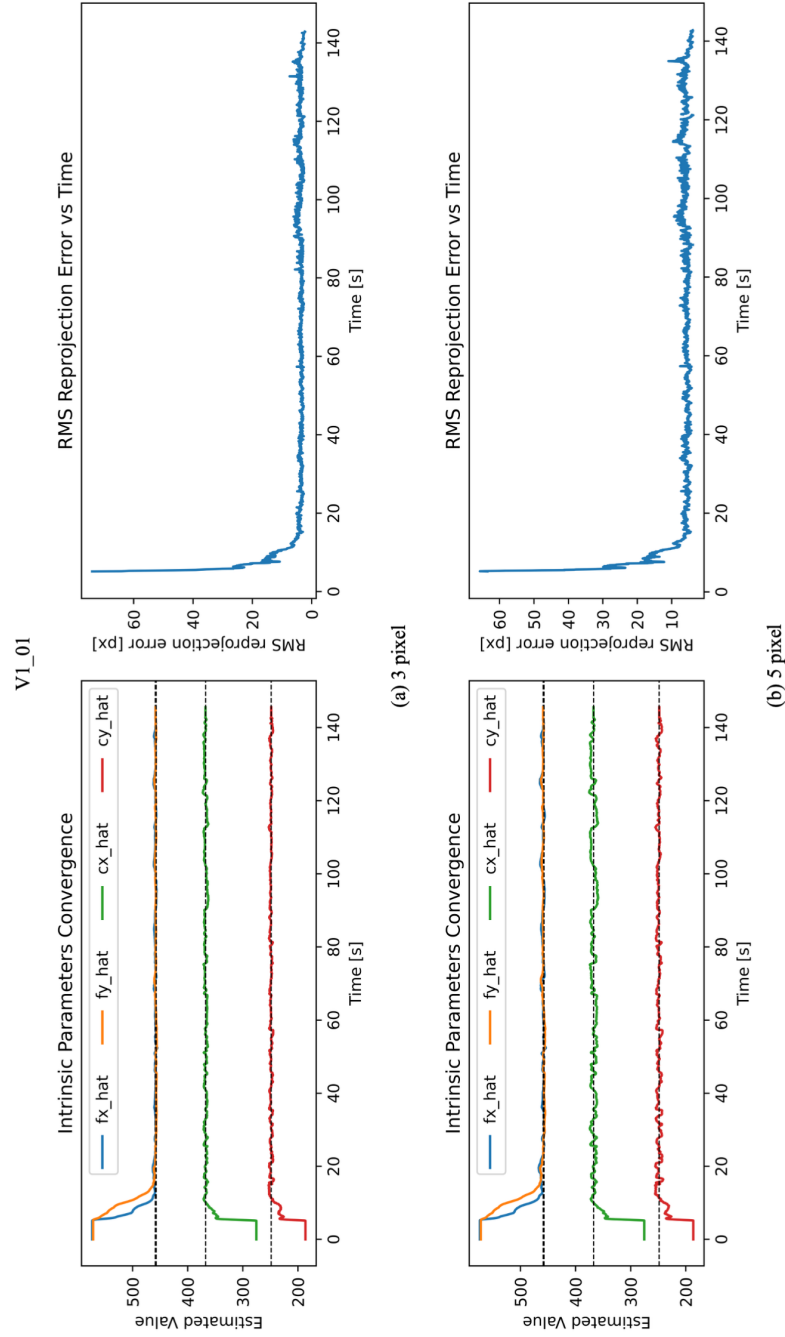


Figure 11: Measurement-Noise Robustness on EuRoC MAV (V1\_01). We inject continuous zero-mean Gaussian noise with  $\{3, 5\}$  px into feature coordinates. Despite continuous noise, RACE keeps RMS error fluctuations within bounded pixel values and never diverges. The increase in the ultimate error bound matches our Theorem 2, confirming quantitative robustness to measurement perturbations.

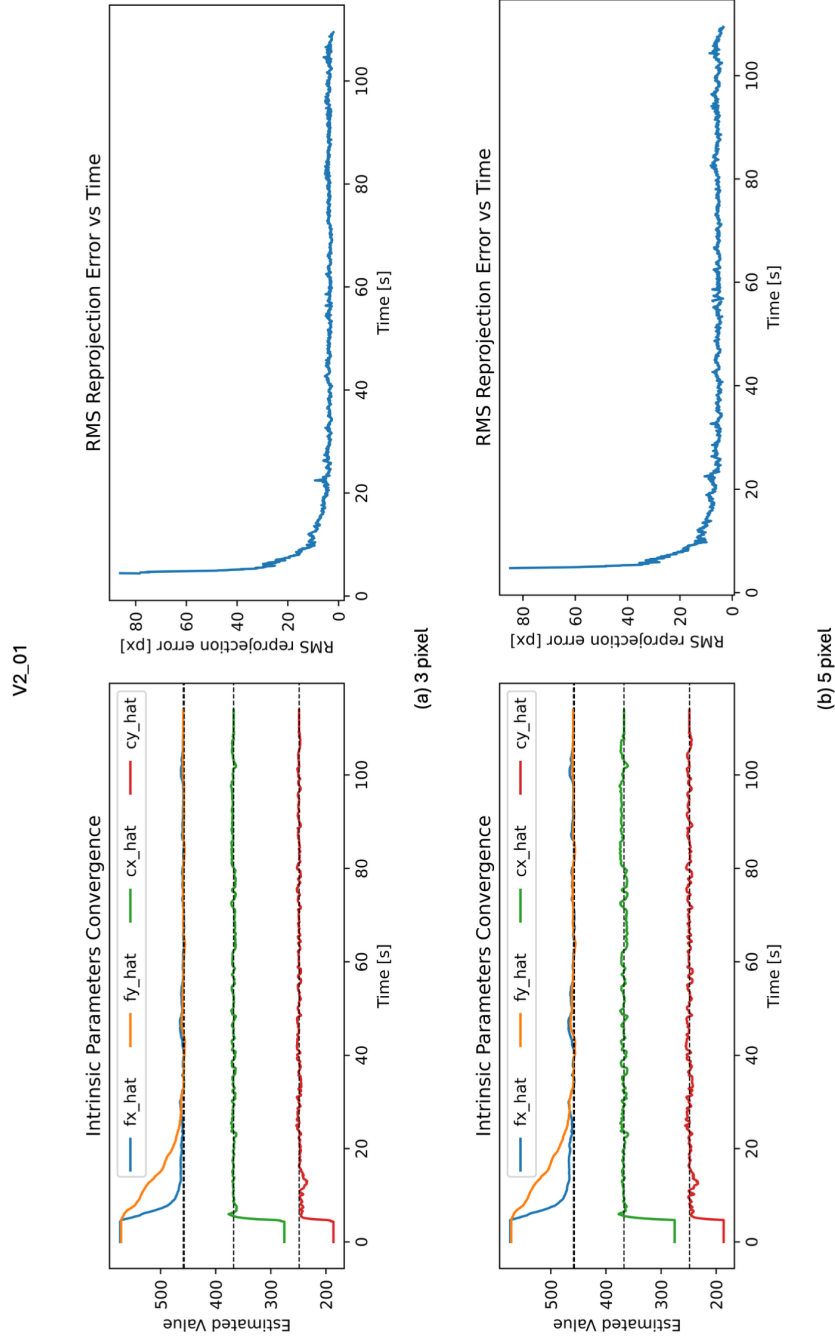


Figure 12: Measurement-Noise Robustness on EuRoC MAV (V2\_01). We inject continuous zero-mean Gaussian noise with  $\{3, 5\}$  px into feature coordinates. Despite continuous noise, RACE keeps RMS error fluctuations within bounded pixel values and never diverges. The increase in the ultimate error bound matches our Theorem 2, confirming quantitative robustness to measurement perturbations.

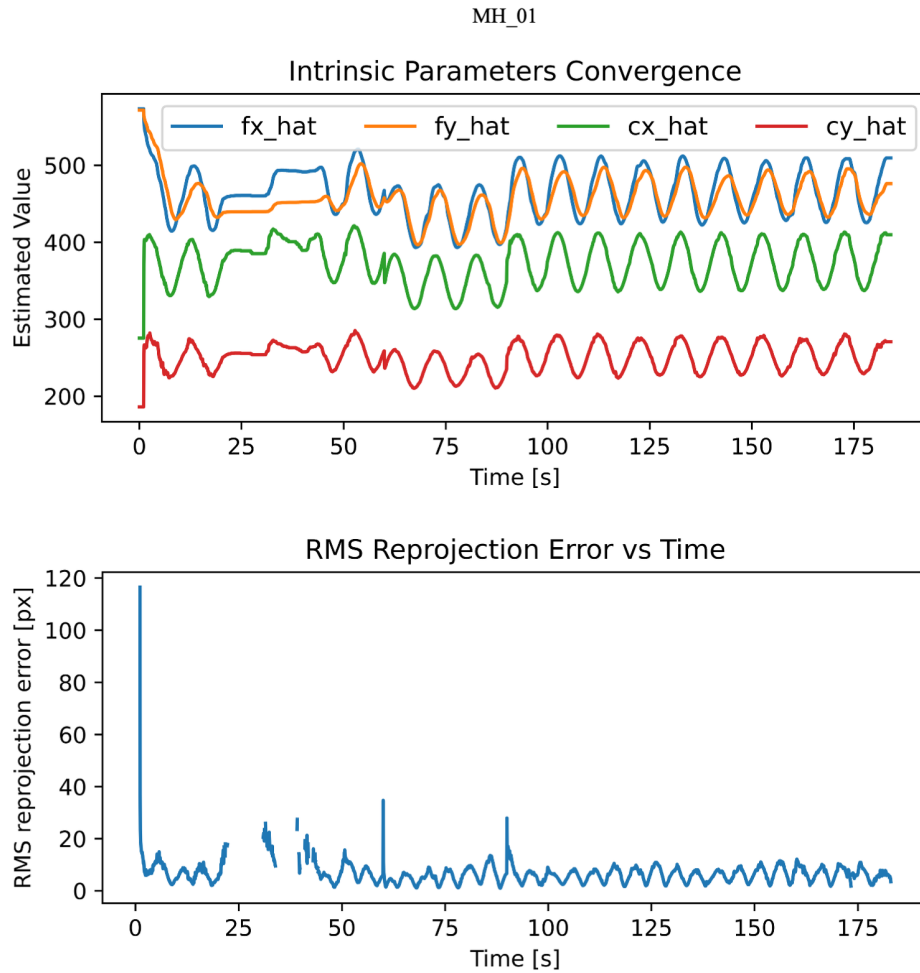


Figure 13: Combined Stress Test on EuRoC MAV (MH\_01). Simultaneously applied disturbances include a 25% initial offset, 0.5 px Gaussian noise, 10% thermal drift (10 s period), and 5 % plateau drift at 30/60/90 s. RACE remains stable under this worst-case compound scenario, gracefully tracking the time-varying true intrinsic parameters and maintaining stable RMS error. **Note:** the true intrinsic parameters vary according to the drift added but are not shown in the plots for visual clarity.

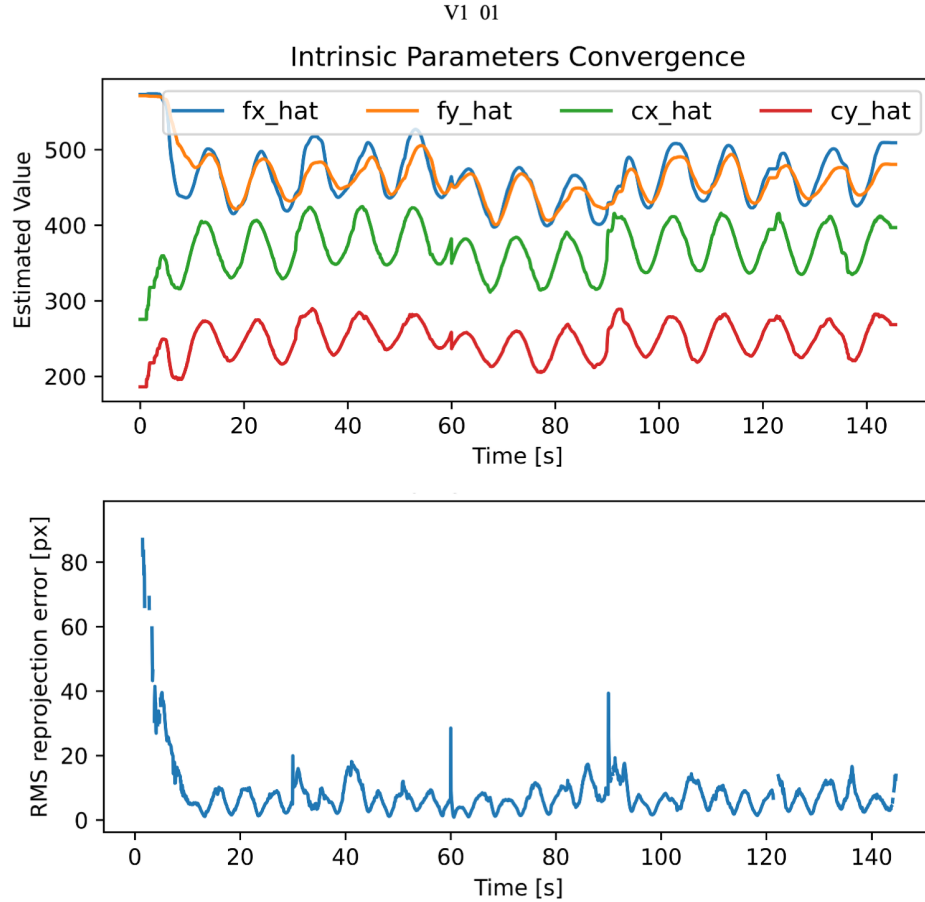


Figure 14: Combined Stress Test on EuRoC MAV (V1.01). Simultaneously applied disturbances include a 25% initial offset, 0.5 px Gaussian noise, 10% thermal drift (10 s period), and 5 % plateau drift at 30/60/90 s. RACE remains stable under this worst-case compound scenario, gracefully tracking the time-varying true intrinsic parameters and maintaining stable RMS error. **Note:** the true intrinsic parameters vary according to the drift added but are not shown in the plots for visual clarity.

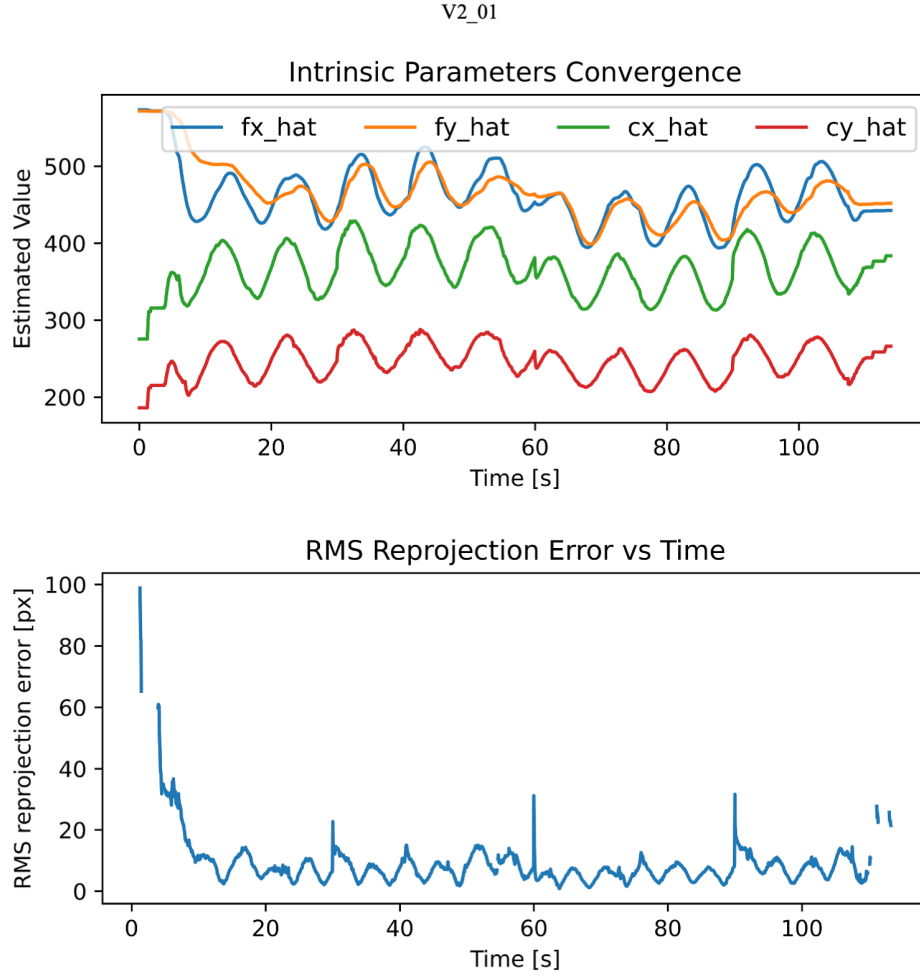


Figure 15: Combined Stress Test on EuRoC MAV (V2\_01). Simultaneously applied disturbances include a 25% initial offset, 0.5 px Gaussian noise, 10% thermal drift (10 s period), and 5 % plateau drift at 30/60/90 s. RACE remains stable under this worst-case compound scenario, gracefully tracking the time-varying true intrinsic parameters and maintaining stable RMS error. **Note:** the true intrinsic parameters vary according to the drift added but are not shown in the plots for visual clarity.

### 1.3 LIMITATIONS AND FUTURE WORK

Despite its strong theoretical guarantees and empirical robustness, RACE has several limitations that warrant further investigation:

- **Dependence on feature tracking.** RACE relies on a steady stream of well-distributed image features. In low-texture or pure-rotation scenarios, feature loss can degrade the regressor conditioning and slow convergence or, in extreme cases, lead to temporary loss of calibration updates.
- **Persistent-excitation requirement.** Our stability and convergence guarantees rely on the assumption that the regressor  $\Phi_t$  is persistently exciting. In practice, motion sequences that lack sufficient diversity (e.g., slow planar movement) may violate this assumption, resulting in degraded parameter identifiability. In these scenarios, we can monitor the PE and perform parameter updates only when the system observes sufficient feature richness and motion diversity.
- **Synchronization effects.** Our model assumes a camera and perfectly time-synchronized measurements. However, the presence of timestamp jitter can introduce a model mismatch that is not currently addressed.
- **Computational spikes from feature re-initialization.** Although the core Jacobian update is lightweight, occasional spikes occur when the feature pipeline re-detects landmarks. Embedding a low-overhead, dedicated tracker or GPU-offloading could further stabilize worst-case latency.
- **No end-to-end SLAM integration.** We have validated RACE in isolation; its interaction with full SLAM or VIO pipelines (e.g. loop closure) remains to be explored. Integrating RACE as an online module within a visual SLAM is a promising direction.
- **Single-camera focus.** While the method generalizes to stereo and RGB-D rigs in principle, we have only demonstrated monocular experiments. Multi-camera validation and real-world depth-scale calibration are future extensions.

Addressing these limitations will further enhance RACE’s applicability to a broader range of challenging real-world scenarios.

#### 1.3.1 TARTANAIR DATASET

In contrast to our real-world benchmarks, RACE exhibits degraded performance on the TartanAir synthetic sequences (see Section 4), motivating a closer examination of its failure modes. We analyze three representative data sequences, MH\_000, MH\_001, and MH\_002, and identify the exact frames where reprojection error spikes (see Figures 17, 18, 19). In each case, calibration deteriorates when the image stream lacks rich, well-distributed features, particularly in low illumination (MH\_000) and foggy (MH\_002) environments, causing the regressor  $\Phi_t$  to become ill-conditioned.

This behavior aligns with our theoretical requirement for persistent excitation: without sufficient feature diversity, the parameter update law cannot reliably recover true intrinsics. To mitigate this, one can (i) pause updates when feature quality or regressor condition falls below a threshold or (ii) trigger active motion excitation small, information-rich camera maneuvers to restore excitation and resume stable convergence. Incorporating these strategies into RACE is an essential direction for future work to ensure robust calibration even under challenging visual conditions.

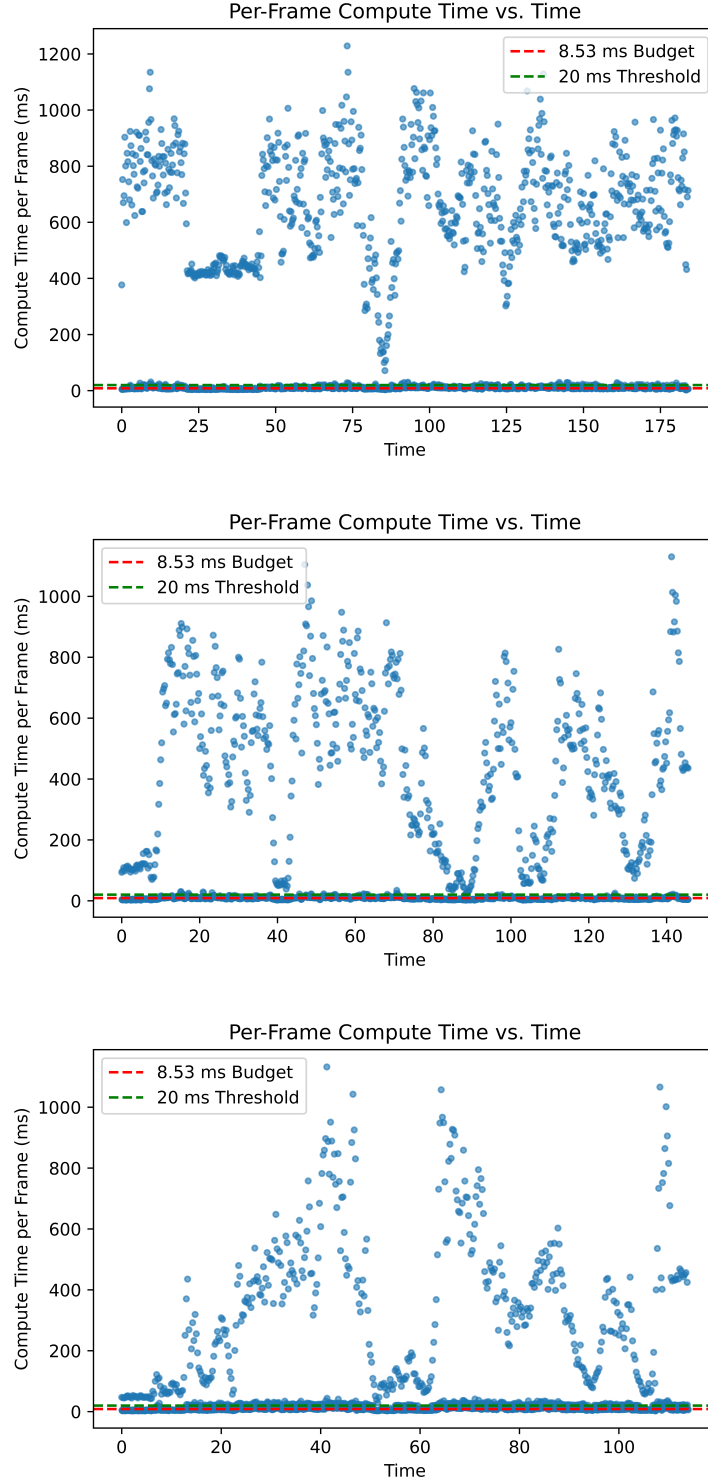


Figure 16: Compute Time vs. Time (EuRoC MAV). Scatter plot of per-frame compute time (ms) versus time, with the 8ms real-time budget shown as an orange dashed line. Most frames cluster between 5-20 ms, while occasional spikes above 20ms (green) occur during feature re-initialization or landmark re-detection. The majority of the computation power is spent on Jacobian  $\phi_t$  formation and parameter update  $\hat{\theta}_t$ .

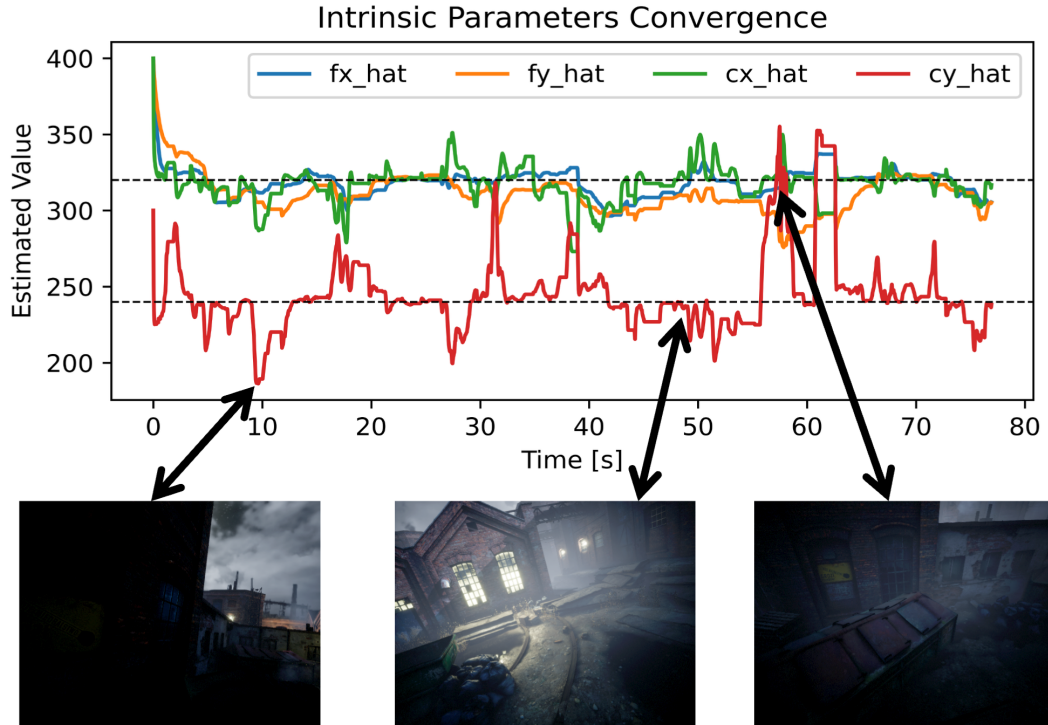


Figure 17: Failure Modes on TartanAir Sequences (MH\_000). The plot highlights the frames where the estimation of intrinsic parameters degrade, resulting in spikes. Prominent error spikes occur in frames with sparse, low-contrast features under poor illumination (leftmost and rightmost frames). In these regions, the regressor  $\Phi_t$  becomes ill-conditioned, stalling parameter updates and causing transient divergence. By contrast, the middle frame corresponds to a frame with richer feature content and exhibits stable convergence.



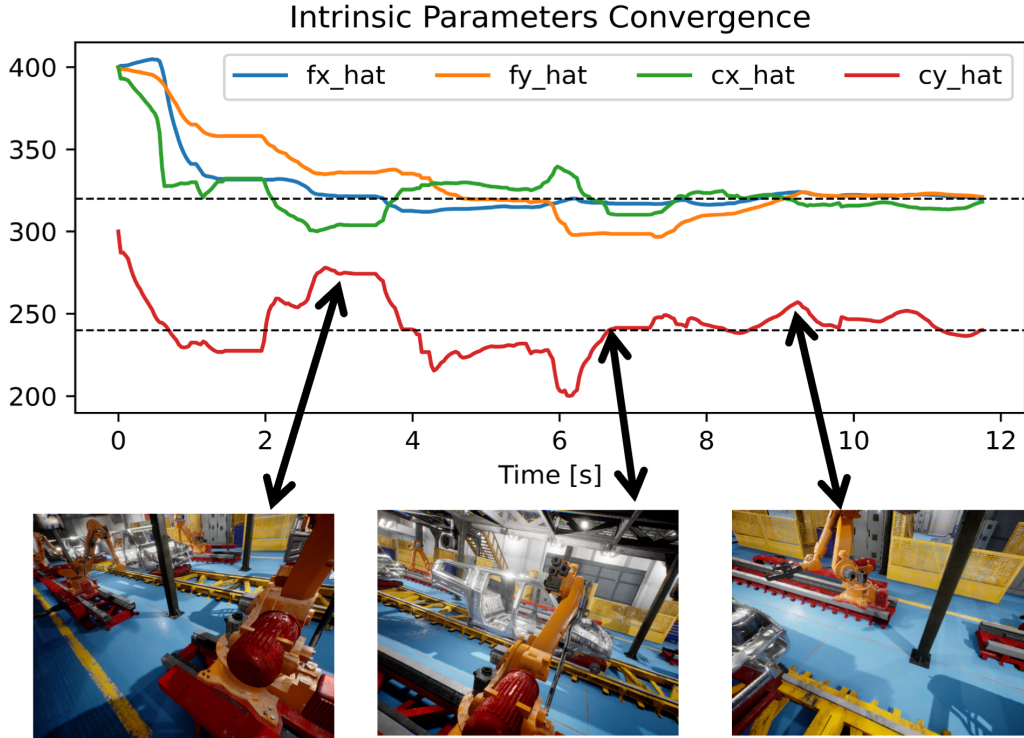


Figure 18: Failure Modes on TartanAir Sequences (MH.001). The plot highlights the frames where the estimation of intrinsic parameters degrade, resulting in spikes. Prominent error spikes occur in frames with sparse, low-contrast features under poor illumination (leftmost and rightmost frame). In these regions, the regressor  $\Phi_t$  becomes ill-conditioned, stalling parameter updates and causing transient divergence. By contrast, the middle frame corresponds to a frame with richer feature content and exhibits stable convergence.

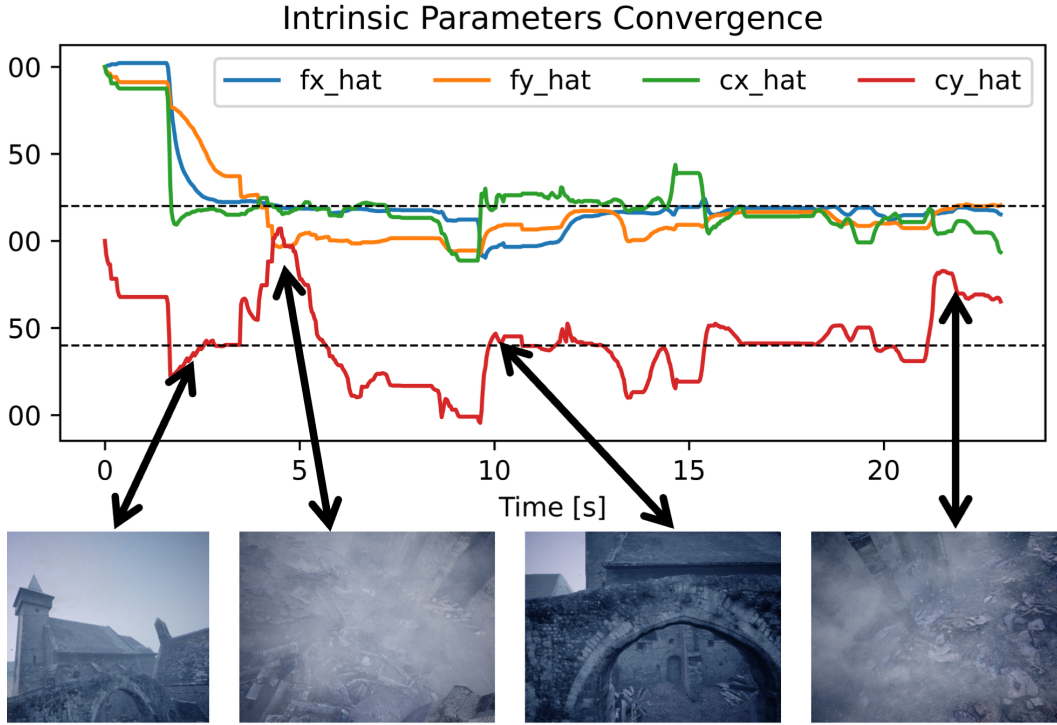


Figure 19: Failure Modes on TartanAir Sequences (MH\_002). The plot highlights frames where intrinsic parameter estimation degrades, resulting in prominent error spikes. Frames are discussed from left to right. Significant spikes are observed in frames with sparse, low-contrast features under poor illumination (second and rightmost frames). In these regions, the regressor  $\Phi_t$  becomes ill-conditioned, hindering parameter updates and causing transient divergence. In contrast, the leftmost and third frames contain richer visual features and demonstrate stable convergence behavior