Pair-Relative Partial Equivariance for Rotation-Robust Keypoint Descriptor Matching

— Supplementary Material —

This supplement collects analyses that support the main paper but do not fit its page budget. Section A develops the full theory (the closed-form optimal steerer and the non-homomorphism theorem). Section B reports the matched-budget per-benchmark absolute scores for the three-way comparison. Section C details the AIMS [31] evaluation protocol. Section D delimits the scope under DeDoDe-G [9] joint training. Section E lists the HPatches [1] training, evaluation, and test splits. Section F shows dataset-by-dataset qualitative comparisons. Throughout, **PaSteer-free** is our free-form Fourier steerer and *exact* is the Steerers [5] representation retrained at our budget on a frozen DeDoDe-B descriptor.

A. Why relaxation helps: theory

This section gives the theory behind PaSteer in full, expanding the brief summary in Sec. 3.1. The argument is short and has one moving part: for a *fixed*, non-equivariant descriptor, the best possible linear steerer at each angle is itself generally not a group representation, so an exact representation steerer is provably suboptimal and a relaxed steerer can do strictly better.

The best steerer at a single angle. Fix the descriptor f and one rotation angle θ. The steerer that best transports descriptors from the upright frame into the θ-rotated frame is the linear map minimizing the average squared steering error,

$$\rho^\star(\theta) = \operatorname*{arg\,min}_{M \in \mathbb{R}^{D\times D}} \mathbb{E}_I \big\| f(R_\theta I) - M f(I) \big\|^2. \tag{7}$$

This is an ordinary ridge least-squares problem, so it has the closed-form solution

$$\begin{aligned} \rho^\star(\theta) &= C(\theta)\,(G + \lambda I)^{-1}, \\ C(\theta) &= \mathbb{E}\big[f(R_\theta I) f(I)^\top\big], \\ G &= \mathbb{E}\big[f(I) f(I)^\top\big]. \end{aligned} \tag{8}$$

Here G is the covariance of the descriptor and $C(\theta)$ is the cross-covariance between upright and rotated descriptors. We call $\rho^\star$ the *per-angle optimal steerer*: it is the best any linear steerer can do at angle θ, but it estimates a separate full $D \times D$ matrix for each angle and therefore serves as an oracle reference rather than a practical compact model.

The ceiling is not a representation. The exact steerer implicitly assumes that the angle-wise transport maps chain together into a representation, $\rho(\theta + \phi) = \rho(\theta)\rho(\phi)$. This holds only under a restrictive second-order condition. Write $L(\rho) = \mathbb{E}_{\theta\sim q}\, \mathbb{E}_I \, \|f(R_\theta I) - \rho(\theta) f(I)\|^2$ for the aggregate steering loss.

Theorem 1 (The per-angle optimum is generally not a representation)**.** *Assume $\lambda \to 0^+$ and G is invertible, so that $\rho^\star(\theta) = C(\theta)G^{-1}$. Then $\rho^\star$ satisfies $\rho^\star(\theta + \phi) = \rho^\star(\theta)\rho^\star(\phi)$ for all θ, ϕ if and only if*

$$C(\theta + \phi) = C(\theta) G^{-1} C(\phi). \tag{9}$$

Thus, the per-angle optimum is a representation only when the angle-wise descriptor correlations compose consistently. Frozen upright-trained descriptors generally violate this condition, so their optimal angle-wise steerers need not form a single exact $\mathrm{SO}(2)$ action. Because $\rho^\star$ is the angle-wise least-squares optimum, it lower-bounds the loss of any exact representation, with a strict gap whenever the least-squares optimum is unique and the violation occurs on angles of positive q-measure. This is the load-bearing claim of the paper.

Two caveats. First, failure of the *composition law* is the property we rely on. A separate, weaker fact is that $\rho^\star$ need not be *orthogonal*; but an $\mathrm{SO}(2)$ representation need not be orthogonal either, unless one additionally asks for a unitary representation. We therefore report both metrics and lean on the composition violation. Second, at finite ridge $\lambda > 0$ the solution has $\rho^\star(0) = G(G + \lambda I)^{-1} \neq I$. This is a numerical regularization effect rather than a property of f, and the zero-angle error vanishes in the unregularized limit $\lambda \to 0^+$. We therefore state the theorem in the clean population limit and use λ only to stabilize the estimator.

From ceiling to method. $\rho^\star$ is an oracle per-angle matrix; the practical question is how closely a compact steerer can approach it. PaSteer-free approximates $\rho^\star$ with a handful of harmonics, while PaSteer-soft (Eq. (3)) trades a little of this gap for an exact upright identity $\tilde{\rho}(0) = I$ and graceful extrapolation beyond the training angles. Section 4.2 verifies all of this on real DeDoDe-B [9] and DeDoDe-G descriptors: the per-angle optimum violates the composition law, a free-form steerer nearly attains the optimal floor, and a trained exact steerer falls short, with the gap widening on the larger descriptor, as the theorem predicts.

B. Matched-budget per-benchmark absolutes

Controlled-comparison convention. Unless stated otherwise, a controlled comparison means that the descriptor, detector, keypoints, matcher, data split, training budget, and

steering	upright	C4	SO2
Matched-budget controlled comparison			
dual softmax (none)	49.8	12.4	11.4
exact expm (ours, HPatches)	49.8	28.5	19.4
PaSteer-soft (r=12, ours)	49.8	37.5	24.4
PaSteer-free (ours)	49.0	**43.1**	**33.0**
Released reference models, not controlled			
exact Setting-A (release)	49.8	39.5	31.3
Steerers SO2-Spread-B	49.7	49.0	47.5

Table 10. **MegaDepth-1500 [19] matched-budget absolutes.** Pose AUC@5° on a frozen DeDoDe-B [9]. The top block is the matched-budget controlled comparison (exact < PaSteer-soft < PaSteer-free); the bottom block lists released models as absolute references only.

steering	AP
Matched-budget controlled comparison	
unsteered (dual softmax)	25.5
exact expm (ours, HPatches)	37.1
PaSteer-soft (r=12, ours)	44.4
PaSteer-free (ours)	**51.2**
Released reference models, not controlled	
exact Setting-A (release, MegaDepth)	46.3
Steerers SO2-Freq1-B	57.7

Table 11. **AIMS [31] matched-budget absolutes.** Average precision on the FMAP evaluation subset, frozen DeDoDe-B [9], with identical candidate set and positives across rows. The top block is the matched-budget controlled comparison (same ordering as Roto-360 [17]); the bottom block lists released models as absolute references only.

evaluation protocol are fixed, and only the steerer changes. Released or published models are included only as absolute-performance references and are not used for mechanism-level claims.

Tables 10 and 11 show the MegaDepth-1500 [19] and AIMS [31] per-benchmark absolute scores behind the matched-budget three-way comparison (exact, PaSteer-soft, PaSteer-free) of Sec. 4.1, each trained on the same HPatches [1] sequences over a frozen DeDoDe-B [9] descriptor; the Roto-360 [17] table is in the main paper (Tab. 4). These complement the main comparison by confirming the same ordering, exact < PaSteer-soft < PaSteer-free, on per-benchmark pose AUC and average precision. Following the convention above, the top block of each table is the controlled comparison and the bottom block lists released models as absolute references only, so the mechanism-level claim should be read from the matched-budget block. The released exact steerer is MegaDepth-trained (hence in-domain on MegaDepth-1500) and the released joint Steerers [5] additionally differ in steerer family and, on AIMS, use a non-public Procrustes protocol (Sec. C).

B.1. Additional full-scale absolute results and protocol notes

Tables 12 and 13 show full-scale Setting-A controlled comparisons of the exact and relaxed steerers on IMC-PT [13] and the AIMS [31] full-FMAP set, on a frozen DeDoDe-B [9]. The AIMS numbers in Tab. 11 and Tab. 13 correspond to different training scales: HPatches matched-budget versus full-scale Setting-A. These complement the headline margins (Tab. 1) by giving the absolute scores behind them at full scale: PaSteer-free improves over exact mainly on the large-rotation axes and splits while staying within a fraction of a point of exact near upright. Protocol caveats specific to AIMS are deferred to Sec. C.

IMC-PT [13]	upright	C4	SO2
exact	66.4/80.7/89.5	41.5/55.6/67.2	22.2/31.3/41.1
PaSteer-free	66.0/80.6/89.6	**47.3/62.7/75.2**	**28.7/40.1/50.6**

Table 12. **IMC-PT full-scale Setting-A controlled comparison.** IMC-PT 2020 [13] validation (1500 pairs/axis), frozen DeDoDe-B [9], pose AUC@{5,10,20}°. PaSteer-free improves over exact on the large-rotation C4/SO2 axes (up to +8.0/+9.5 AUC) while staying within a fraction of a point of exact on upright.

steerer	AP All	AP A.O.
exact expm	48.63	43.40
PaSteer-free (ours)	**50.11**	**45.24**
margin	+1.48	+1.85

Table 13. **AIMS [31] full-FMAP controlled comparison.** Exact versus PaSteer-free on the full FMAP evaluation set (323 frames, 21 777 pairs), frozen DeDoDe-B [9], full-scale Setting-A, C_8 maximum-similarity matching with identical candidate set and positives. AP All is over all pairs and AP A.O. is the large-rotation (any-orientation) split; PaSteer-free improves on both. Protocol caveats are in Sec. C.

Table 13 reports the controlled AIMS full-FMAP comparison: the exact and relaxed steerers share the frozen DeDoDe-B descriptor, candidate set, positives, C_8 maximum-similarity matcher, and training budget, with only the steerer mechanism differing. PaSteer-free improves over exact on both the overall and the large-rotation (A.O.) split. The absolute level reflects the AIMS protocol of Sec. C.

model	Roto	MD upr.	MD C4	MD SO2
exact steerer (joint)	**91.9**	**49.7**	**46.1**	**42.3**
PaSteer-free (ours, joint)	90.7	45.8	41.5	38.2
PaSteer-free (pinned, ours, joint)	–	47.7	42.2	39.2

Table 14. **Joint training on DeDoDe-G.** Full-scale joint descriptor–steerer training on DeDoDe-G [9]. We report Roto-360 [17] correct matches @5px and MegaDepth-1500 [19] pose AUC@5° for upright, C4, and SO2 rotations. All rows use the same scenes, schedule, generator core, and initialization; only the steerer mechanism differs. Unlike the frozen-descriptor setting, the exact steerer keeps a small edge after the descriptor co-adapts, while pinning $\tilde{\rho}(0) = I$ recovers part of the PaSteer-free upright gap.

C. AIMS evaluation protocol

This section explains why our AIMS average precision differs from the published Steerers number and why the within-protocol exact-vs-PaSteer comparison is the relevant evidence. The AIMS [31] average precision we report on a frozen descriptor (Setting-A) is in the 46–50 AP range, whereas Steerers [5] report $\approx$ 60 AP. The two numbers are produced by different evaluation protocols and are not directly comparable: they differ in the matcher, the candidate set, and the positive-label protocol. The reported $\approx$ 60 AP uses a Procrustes matcher with the released, MegaDepth-trained [19] SO2-Freq1-B model and a candidate retrieval set and positive-label margin that are not public. Our controlled study instead fixes the public FMAP candidate and positive set and uses maximum-similarity matching over C_8 probe angles for every method, the protocol under which all steerers in this paper are compared on equal footing. Under this protocol the released exact Setting-A steerer scores $\approx$ 46 AP (44.9 on the 162-frame subset, 46.3 on the full set), which is the within-protocol anchor; under the Procrustes and full-FMAP protocol the same pipeline reproduces 57.7 AP, matching the published setting. The relevant quantity is therefore the within-protocol exact-vs-PaSteer-free comparison (Tab. 13).

D. Joint training on DeDoDe-G

Table 14 shows full-scale joint descriptor–steerer training on the larger DeDoDe-G [9] descriptor (DINOv2 ViT-L backbone), with identical scenes, schedule, generator core, and initialization so that only the steerer mechanism differs; the exact generator is held skew-symmetric to remain a valid representation. It complements the DeDoDe-B joint result of the main paper (Sec. 4.5), where the two arms reach parity once the descriptor co-adapts.

Under joint training on the larger G descriptor the exact steerer marginally *outperforms* PaSteer-free on the rotation axes (C4, SO2) and on upright; pinning the upright identity $\tilde{\rho}(0) = I$ of the relaxed steerer (the PaSteer-free (pinned) row, two seeds) recovers about half of the gap but does not reverse it. This is the regime opposite to the frozen DeDoDe-G Setting-A result, where the relaxed steerer dominates the exact one by a wide margin on every rotation axis (Tab. 3). This result delimits the scope of PaSteer: relaxation helps in the frozen-descriptor retrofit regime, whereas a jointly trained descriptor can adapt to, or even benefit from, the exact representation constraint.

E. HPatches training, evaluation, and test splits

Roto-360 [17] applies in-plane rotations to the 116 HPatches [1] sequences (prefix `i_` for illumination sequences, `v_` for viewpoint sequences). For the matched-budget steerers we do not introduce a new evaluation boundary: the 10 sequences of the REKD/RELF Roto-360 `debug` split [17, 18] are used verbatim as the Roto-360 matching test set and never enter training or evaluation. The remaining non-`debug` sequences are placed in a fixed order and shuffled once with a fixed random seed; the first 60 form the steerer training set and the next 12 the steering-error evaluation set, leaving the remaining 34 unused. We fix the seed so that this partition is deterministic and identical for every steerer variant: the exact and relaxed arms are then trained and evaluated on exactly the same sequences, so any difference between them reflects the steering form alone rather than a different split. This $60/12$ partition is newly defined here.

These HPatches splits govern only the matched-budget, HPatches-trained steerers and the steering-error and theory results: Tabs. 5 and 6 and the matched blocks of Tabs. 4, 10 and 11 (the "ours, HPatches" rows). The full-scale Setting-A results (Tabs. 2, 3 and 12) and the joint training of Tabs. 9 and 14 are trained on MegaDepth and do not use this partition.

Train (60). `i_books, i_boutique, i_bridger, i_brooklyn, i_castle, i_chestnuts, i_crownday, i_crownnight, i_dc, i_dome, i_fenis, i_fruits, i_gonnenberg, i_greenhouse, i_greentea, i_indiana, i_kions, i_kurhaus, i_leuven, i_lionday, i_lionnight, i_melon, i_miniature, i_nijmegen, i_partyfood, i_pencils, i_pinard, i_pool, i_salon, i_ski, i_smurf, i_steps, i_yellowtent, v_abstract, v_adam, v_artisans, v_boat, v_bricks, v_calder, v_cartooncity, v_churchill, v_circus, v_coffeehouse, v_dirtywall, v_dogman, v_graffiti, v_home, v_laptop, v_machines, v_maskedman, v_pomegranate, v_posters, v_soldiers, v_strand, v_sunseason, v_underground, v_vitro, v_wapping, v_war, v_wounded.`

Evaluation set for steering error (12). `i_autannes, i_ktirio, i_londonbridge, i_tools, i_veggies, i_zion, v_azzola, v_beyus, v_bird, v_busstop, v_man, v_samples.`

Roto-360 test (the REKD/RELF `debug` split, 10). `i_ajuntament, i_bologna, i_resort, i_table, i_troulos, v_birdwoman, v_courses, v_gardens, v_tabletop, v_wormhole.`

F. Qualitative comparison

Figures 6 to 9 show qualitative comparisons on MegaDepth-1500 [19], IMC-PT [13], Roto-360 [17], and AIMS [31], illustrating dataset by dataset what the relaxed steerer recovers that an exact representation cannot. Each figure compares three steerers on one image pair, all on the same frozen DeDoDe-B [9] descriptor, the same SO2 detector, and the same keypoints, so any difference is attributable to the steerer alone: **(a) no steer** (the raw upright descriptor), **(b) the exact Steerers [5] representation** $\rho(\theta) = \exp(\theta D)$, and **(c) PaSteer-free**. The color convention is fixed throughout: green lines are ground-truth-verified correct matches, blue lines (panel (c)) are the correct matches PaSteer-free recovers but the exact steerer misses, and faint orange lines are rejected matches (subsampled); the number of inlier lines drawn is capped for legibility while the panel titles report the true counts. We rotate the synthetic-rotation pairs (MegaDepth-1500, IMC-PT) to $112°$ and the Roto-360 pair to $70°$, angles off the C_8 probe grid $\{0°, 45°, \dots\}$ (the "C_8-dip" angles of Fig. 5) where the exact representation is weakest; AIMS pairs carry their own intrinsic rotation and are not rotated synthetically.

Across all four datasets the pattern is the same: without steering the upright descriptor collapses under large rotation, the exact steerer recovers most of the structure, and PaSteer-free recovers a further set of correct matches (blue) that the representation constraint leaves unmatched. The effect is largest on AIMS, where the pair-relative rotation distribution $q(\Delta\theta)$ is closest to uniform (Sec. 4.4) and the cost of forcing a single exact representation over all of $\mathrm{SO}(2)$ is highest. The qualitative pattern is consistent with the steering-error and non-homomorphism analyses in Sec. A and Sec. 4.2.

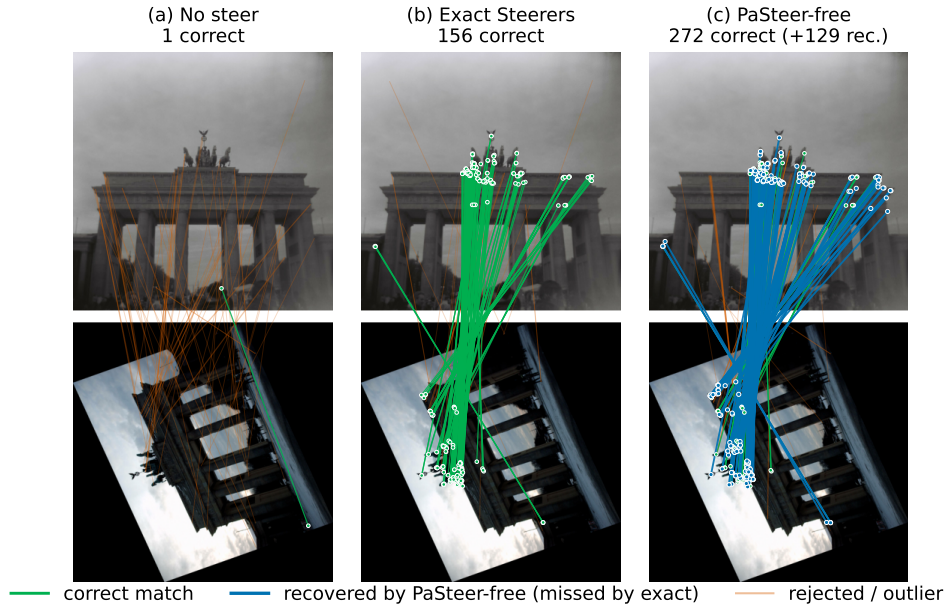


Figure 6. Qualitative results on MegaDepth-1500 [19] (scene `0022_0.5_0.7`, pair 2), target rotated 112° in-plane. No steer: 1 correct match; exact Steerers [5]: 156; PaSteer-free: 272 ($+129$ recovered).

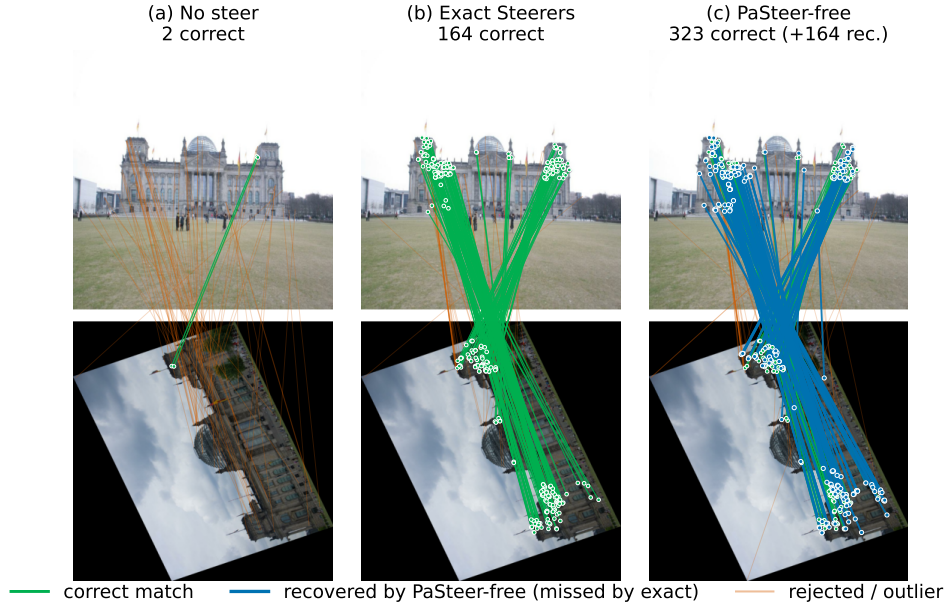


Figure 7. Qualitative results on IMC-PT 2020 [13] (`reichstag`), target rotated 112°. No steer: 2 correct; exact: 164; PaSteer-free: 323 ($+164$ recovered).

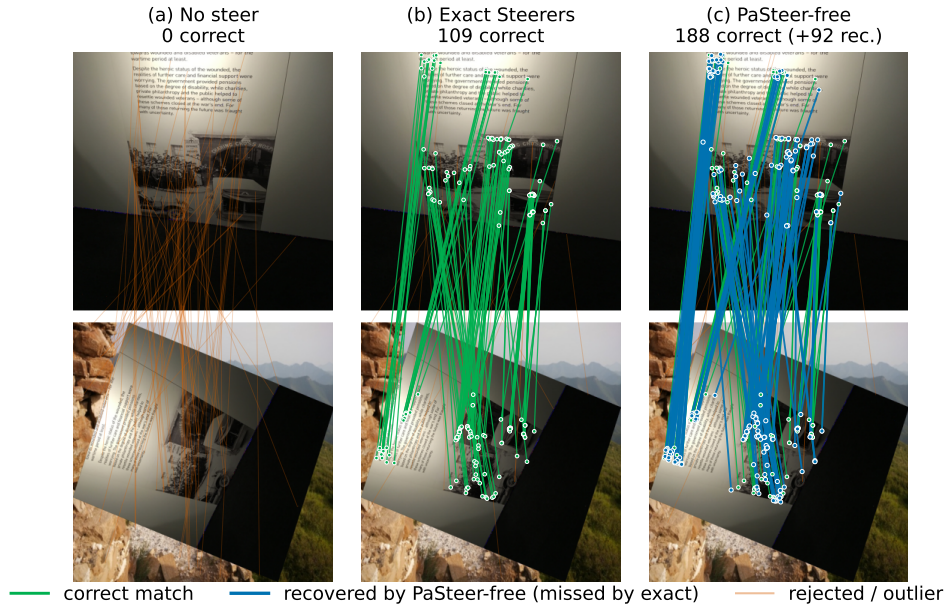


Figure 8. Qualitative results on Roto-360 [17] (sequence `v_charing`), pure in-plane rotation of 70° with a ground-truth homography. No steer: 0 correct; exact: 109; PaSteer-free: 188 ($+92$ recovered). Inliers are GT-homography reprojection ≤ 5 px.

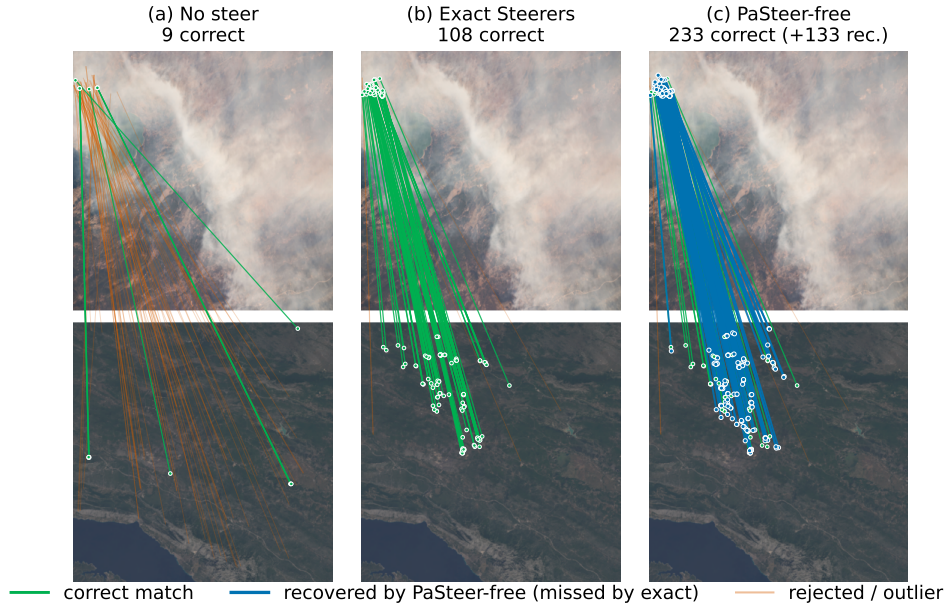


Figure 9. Qualitative results on AIMS [31] (astronaut frame `ISS052-E-27335` vs. a satellite tile), relative rotation $\approx 51^\circ$. No steer: 9 correct; exact: 108; PaSteer-free: 233 ($+133$ recovered). Inliers are MAGSAC homography inliers.