Observability Considerations for Cislunar Angles-Only Relative Navigation

Andrea López
University of Colorado Boulder
Jackson Kulik
Utah State University

Jay McMahon
University of Colorado Boulder

ABSTRACT

Angles-only relative navigation using passive optical sensors is attractive for cislunar proximity operations due to its low size, weight, and power requirements, but it has a fundamental limitation: bearing measurements alone cannot instantaneously recover range. In the Earth-Moon Circular Restricted Three-Body Problem (CR3BP), the nonlinear multi-body dynamics can provide weak range observability over time through state coupling and geometric evolution, but this observability is local, directional, and strongly dependent on both the orbital geometry and the relative configuration between spacecraft. This paper develops and applies several complementary nonlinear observability tools, two established (the Hermann-Krener observability matrix, the empirical observability Gramian) and two proposed extensions (a sigma-point observability Gramian conditioned on achievable estimation uncertainty, and higher-order measures capturing second-order curvature), to characterize the directional and time-varying observability of cislunar angles-only relative navigation across two representative orbit families.

1. INTRODUCTION

Cislunar space is transitioning from a sparsely populated, mission-specific environment to one supporting sustained scientific, commercial, and exploration activity [1, 2]. Lunar landers, communication and navigation relays, logistics vehicles, and the Gateway outpost in a near rectilinear halo orbit will operate concurrently [3, 4], and rendezvous, inspection, servicing, and formation-keeping among them all require relative navigation. Many of these vehicles face size, weight, and power constraints that preclude active sensors, and the ground-based tracking cadence available in cislunar space cannot meet the update rates that proximity operations demand [1]. Passive optical sensors are the natural alternative, and angles-only relative navigation using line-of-sight measurements has accordingly received sustained attention [5, 6].

However, angles-only comes with a fundamental limitation: a bearing measurement gives the direction to a target but does not constrain its distance. Woffinden and Geller [7] showed that under linear relative dynamics the range to a target is unobservable from bearing measurements at any order. The obstruction is exact, not numerical. Subsequent work established that nonlinear dynamics restore some range observability: Kaufman et al. [8] derived nonlinear observability criteria for the two-body problem using differential-geometric methods, Butcher et al. [9] showed that higher-fidelity relative models improve observability and shorten convergence times relative to Hill-Clohessy-Wiltshire dynamics, and Kulik and Savransky [10] used higher-order state transition tensors to exploit nonlinear coupling for angles-only relative orbit determination. Gong et al. [11] survey the strategies developed to strengthen observability in Earth orbit, classifying them by whether they exploit dynamics, multiple lines of sight, maneuvers, or sensor placement. In the cislunar regime, Greaves and Scheeres [5, 6] extended the LiAISON framework to line-of-sight measurements, and LeGrand et al. [12, 13] developed bearings-only initial relative orbit determination and angles-only tracking of cislunar objects.

Assessing nonlinear observability quantitatively, as opposed to establishing that it exists, rests on a small set of established tools. The Hermann-Krener rank condition [14] tests local observability through the gradients of successive Lie derivatives of the measurement function, and its singular value decomposition yields directional sensitivity scores

[15]. The empirical observability Gramian of Krener and Ide [16] answers the complementary finite-window question by perturbing and re-simulating, requiring no analytical derivatives. Both have been applied in the cislunar context: Fowler and Paley [17] used the Gramian condition number to compare orbit families for space domain awareness, and Givens et al. [18] computed the Hermann-Krener matrix by automatic differentiation for cislunar constellation design. Powel and Morgansen [19] related the empirical Gramian to the Fisher information matrix and the Cramér-Rao bound, and Alaeddini and Morgansen [20] used observability measures to design trajectories that improve estimation. In each of these applications the observability assessment informs a decision made before the fact, at the design or planning stage, and the quantity it returns is treated as a fixed property of the geometry. The present work asks the same tools a question they are not usually asked: whether what they return is sufficient to inform decisions taken inside an estimator, while it runs.

Both tools share two properties that have received little scrutiny. The first concerns what "nonlinear" refers to: H-K folds the dynamics into the assessment through successive Lie derivatives and EOG propagates perturbations through the full nonlinear flow, so neither replaces the dynamics with a linear model, but both are nevertheless linear in the state deviation, H-K by evaluating gradients at a point and EOG by a symmetric difference that cancels second-order content. The second is that neither is conditioned on what an estimator believes about the state. Whether the resulting first-order ordering of directions is sufficient depends on what it is used for. For comparing candidate geometries at the design stage it plausibly is; for decisions internal to an estimator, which act on directions and are taken at the uncertainty the filter actually carries, it is less clear, and the question does not appear to have been examined.

Two tools are introduced alongside the two established ones. A sigma-point observability Gramian conditions the assessment on an estimator's assumed covariance instead of an arbitrary perturbation scale, and inherits an interpretation through the Fisher information matrix. A higher-order Gramian retains the second-order term of the same measurement map, reporting the curvature that the first-order tools cancel or absorb. All four are applied to lead-follower geometries in two CR3BP orbit families chosen because they differ in first-order observability structure: a northern butterfly orbit, fully three-dimensional with a relative geometry that evolves substantially over an orbital period, and a distant retrograde orbit, near-planar and close to circular, whose relative geometry poses a challenge for first-order observability tools.

Directional observability information is relevant wherever a decision is made about a direction in state space: sensor tasking, maneuver design for observability enhancement, and the internal representation an estimator maintains. The last is the motivating case here. Gaussian mixture filters are effective in this problem precisely because they represent uncertainty a single Gaussian cannot [21, 22], but they incur a cost proportional to the number of components, so component management decides whether that capability is affordable. Those decisions are made, at present, without reference to what the measurements can resolve.

A second result follows from the analysis, and concerns the tools themselves. The established tools require more care in near-planar geometries than is generally recognised: the Gramian must not be assembled explicitly, a minimum measurement redundancy is needed before the weakest direction is determined at all, and the admissible perturbation scale narrows to about one order of magnitude. Configurations outside these bounds return a plausible spectrum that is wrong along the direction of interest. The higher-order quantities proposed here predict the scale at which the first-order tools are valid. Measured costs add a second correction: they are set by the depth of differentiation a tool requires rather than by the number of trajectories it propagates, which makes the instantaneous tool the most expensive of the four in exactly the geometry that needs it most.

Section 2 reviews linear observability, the angles-only limitation, and what changes for nonlinear systems. Section 3 formulates the four tools. Section 4 describes the scenarios. Section 5 establishes the conditions under which the four tools agree, and section 6 presents the observability structure of the two geometries. Section 7 compares the tools on practical grounds and section 8 discusses implications for mixture filtering.

2. BACKGROUND

2.1 Problem setup

The scenario considered throughout this work is an observer spacecraft estimating the relative state of a target using bearing measurements alone, with both spacecraft on periodic orbits in the Earth-Moon system. The relative dynamics are those of the Circular Restricted Three-Body Problem (CR3BP). Writing the absolute position of a spacecraft in the

Earth-Moon rotating frame as $\boldsymbol{r}$, the equations of motion are

$$\begin{bmatrix} \dot{\boldsymbol{r}} \\ \ddot{\boldsymbol{r}} \end{bmatrix} = \boldsymbol{f}_{\text{CR3BP}}(\boldsymbol{r}, \dot{\boldsymbol{r}}) = \begin{bmatrix} \dot{\boldsymbol{r}} \\ -\mu \dfrac{\boldsymbol{r} - \boldsymbol{r}_1}{\|\boldsymbol{r} - \boldsymbol{r}_1\|^3} - (1-\mu) \dfrac{\boldsymbol{r} - \boldsymbol{r}_2}{\|\boldsymbol{r} - \boldsymbol{r}_2\|^3} + 2\dot{\boldsymbol{r}} \times \hat{\boldsymbol{i}}_k \end{bmatrix}, \tag{1}$$

with μ the Earth-Moon mass parameter, $\boldsymbol{r}_1, \boldsymbol{r}_2$ the positions of the primary and secondary bodies, and $\hat{\boldsymbol{i}}_k$ the unit vector along the rotation axis. The six-dimensional relative state to be estimated is $\boldsymbol{x} = [\boldsymbol{\rho}^\top, \dot{\boldsymbol{\rho}}^\top]^\top$, with $\boldsymbol{\rho} = \boldsymbol{r}_{\text{target}} - \boldsymbol{r}_{\text{observer}}$, propagated by differencing eq. (1) applied to each spacecraft.

The observer carries a passive optical sensor that measures the line-of-sight direction to the target as a pair of bearing angles, azimuth θ and elevation ϕ,

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{\rho}) + \boldsymbol{v} = \begin{bmatrix} \theta \\ \phi \end{bmatrix} + \boldsymbol{v} = \begin{bmatrix} \arctan(\rho_T / \rho_R) \\ \arcsin(\rho_N / \|\boldsymbol{\rho}\|) \end{bmatrix} + \boldsymbol{v}, \tag{2}$$

where ρ_R, ρ_T, ρ_N are the components of $\boldsymbol{\rho}$ in the observer's radial/transverse/normal (RTN) frame, with R along the Moon-observer direction, N normal to the observer's orbital plane about the Moon, and T completing the right-handed triad, and $\boldsymbol{v} \sim \mathcal{N}(0, R)$ is zero-mean measurement noise.

2.2 From linear to nonlinear observability

Classical observability theory asks whether the state of a linear time-invariant system $\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{u}$, $\boldsymbol{z} = C\boldsymbol{x} + D\boldsymbol{u}$ can be recovered from a sequence of measurements. The question is settled by the rank of the observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}, \tag{3}$$

where n is the state dimension [23]. If $\mathcal{O}$ is full rank any initial state is recoverable from measurements over an arbitrarily short interval; if not, directions in its null space never affect the measurements. The test is binary and depends only on A and C, not on the state or the trajectory, and neither property carries over to nonlinear systems.

For angles-only navigation the test fails, for a geometric reason rather than a numerical one. A bearing gives the direction to the target and nothing else: two targets on the same line of sight produce the same pair of angles, and the measurement does not depend on relative velocity. At any single instant the measurement Jacobian $H = \partial \boldsymbol{h} / \partial \boldsymbol{x} \in \mathbb{R}^{2\times 6}$ therefore has rank two, and its four-dimensional null space contains the line-of-sight direction together with all three velocity directions,

$$\mathcal{N}(H) = \operatorname{span}\left\{ \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{0} \end{bmatrix} \right\} \oplus \left\{ \begin{bmatrix} \boldsymbol{0} \\ \delta\dot{\boldsymbol{\rho}} \end{bmatrix} : \delta\dot{\boldsymbol{\rho}} \in \mathbb{R}^3 \right\}. \tag{4}$$

Neither range nor velocity appears in the instantaneous measurement information.

Stacking measurements over time reduces this four-dimensional null space to one: the direction that scales position and velocity together by a common positive factor. A bearing sequence constrains everything transverse to it, including the orientation of the relative velocity and the ratio of range-rate to range, which is what the rotation rate of the line of sight measures, but recovers neither absolute range nor absolute range-rate, since both lie along it. Under linear relative dynamics that direction survives permanently: scaling the initial relative state scales the trajectory by the same factor at every future time, so the relative geometry is identical in shape and schedule and differs only in size. A bearing sensor measures shape and not size, so the two states generate the same measurement sequence for all time however many blocks are stacked. This is the result of Woffinden and Geller [7] under Clohessy-Wiltshire dynamics, and no observation duration, cadence, or sensor accuracy removes it. Range becomes recoverable only when something breaks that replica argument: maneuvers, offsetting the sensor from the observer's center of mass, additional lines of sight, or dynamics under which a scaled initial state does not produce a scaled trajectory [11].

The CR3BP falls in the last of these categories. Its gravitational terms do not scale linearly with distance from the primaries, so a scaled initial relative state does not produce a scaled relative trajectory, and the replica argument holds only to first order and only instantaneously. All range information available in cislunar angles-only navigation originates in that discrepancy, which restricts how observability can be assessed. Linearizing along a reference trajectory and applying the same rank test makes the answer time-varying and trajectory-dependent, replacing a yes or no with a statement about a particular state over a particular window. More importantly it removes the effect of interest: the range information is nonlinear, so a linearized test can report a direction as unobservable while the full nonlinear system does distinguish nearby states along it.

Observability for nonlinear systems is instead defined locally, in terms of distinguishability. A state is locally weakly observable if the measurements it generates differ from those generated by every other state in some neighborhood of it [14]. The statement is local by construction, since a nonlinear system can be locally well-behaved and still globally ambiguous. Hermann and Krener give a sufficient rank condition for this property based on the gradients of successive Lie derivatives of the measurement function along the dynamics, which retains the nonlinear coupling that linearization discards, and whose singular value decomposition gives directional information instead of a bare rank [14, 18]. The construction is instantaneous, however, and says nothing about how quickly information accumulates. The empirical observability Gramian of Krener and Ide [16] answers that finite-time question instead, by perturbing the initial state, re-simulating, and accumulating the resulting deviations in measurement space, requiring no analytical derivatives.

2.3 Degree of observability

For navigation and orbit determination a rank test is rarely the useful question. Full rank means a state is observable in principle, not that a filter recovers it within a mission-relevant window at a realistic cadence and noise level; cislunar angles-only geometries exist that are full rank at every point along an orbit while the smallest singular value sits orders of magnitude below the largest [17]. What matters is how much information the measurements provide and in which directions, relative to the measurement noise and over the available window. A spectral decomposition supplies both, and for a Gaussian mixture filter it is that directional structure rather than a single scalar that is actionable, since splitting, merging, and pruning are all decisions about directions.

3. METHODOLOGY

Four tools are used to characterize observability along a relative trajectory. Two are established and were introduced in section 2.2; two are proposed here.

With the exception of H-K, all take the same form. Over a window of W measurement steps, each accumulates a 6×6 Gramian

$$\mathcal{W} = \sum_{j=1}^{W} D_j^\top R^{-1} D_j, \tag{5}$$

where $D_j \in \mathbb{R}^{2\times 6}$ is the sensitivity of the step-j bearing prediction to a deviation in the initial relative state, weighted by the measurement noise covariance R. The tools differ only in how D_j is obtained: by symmetric finite differences about the reference state (EOG), by linear regression through a sigma-point ensemble drawn from an assumed covariance (SP-EOG), or by analytic differentiation (HOG). The eigendecomposition of $\mathcal{W}$ supplies directional observability information in the same form in every case, so the tools can be compared directly. Section 5.4 checks that they agree.

The window length is a parameter of the analysis rather than a property of the system, and the choice here reflects an estimation viewpoint. The window is taken *forward* from the current epoch because the decisions it is meant to inform are made now and act on measurements not yet received. No measurement values enter any of the quantities below: every finite-time tool is built from predicted sensitivities along a reference trajectory, so what is reported is a prediction of the information the next W measurements will supply, and the assessment can be made before the window is flown. This differs from the question usually asked in trajectory or constellation design, where the window spans a completed arc. Section 5.2 shows the window length also sets a floor, below which there is not enough information to determine the smallest singular value.

Neither H-K nor EOG references what an estimator actually believes about the state, and both isolate an exactly linear sensitivity by construction, H-K by differentiation and EOG by a symmetric difference that cancels second-order content. The two proposed tools each relax one of these. SP-EOG conditions the assessment on an assumed covariance P, asking whether the measurements can distinguish between states that an estimator considers equally plausible, instead of between the reference state and an arbitrary small deviation. HOG retains the second-order term explicitly, quantifying the curvature that the other tools cancel or absorb, and with it the deviation size over which their linear characterization remains valid. Table 1 summarizes the four.

Table 1: Observability tools used in this paper.

Tool	Time support	Sensitivity from	Order retained	Conditioned on P
H-K	instantaneous, order k	Lie derivatives (AD)	first	no
EOG	window W	symmetric finite differences	first	no
SP-EOG	window W	sigma-point regression	blended	yes
HOG	window W	Taylor expansion (AD)	first, second†	optional

†Truncation at second order is a choice, not a limitation of the construction.

State metric. The position and velocity blocks of the relative state carry different units, so any eigendecomposition of a matrix acting on the full state implicitly fixes a timescale relating the two. Leaving that choice implicit amounts to adopting the nondimensionalising time unit, which is a property of the Earth-Moon system and not of the relative motion under study. All tools are therefore evaluated in coordinates scaled by

$$S = \operatorname{diag}(1, 1, 1, v_{\text{char}}, v_{\text{char}}, v_{\text{char}}), \qquad v_{\text{char}} = \langle \|\dot{\boldsymbol{\rho}}\| / \|\boldsymbol{\rho}\| \rangle, \tag{6}$$

averaged along the reference trajectory. Since v_{char} is the mean rate at which the relative geometry itself evolves, this equates a velocity error with the position error it accumulates over the time in which that geometry changes appreciably. Rank and null space are invariant under this transformation, but singular value magnitudes and the ordering of directions are not.

Numerical form. The eigenvalues of $\mathcal{W}$ are the squared singular values of the underlying whitened sensitivity matrix, so assembling $\mathcal{W}$ explicitly squares the dynamic range of the problem and halves the available precision. Throughout this paper the spectrum of every window Gramian is obtained from the singular value decomposition of the whitened sensitivity matrix

$$A = \begin{bmatrix} R^{-1/2} D_1 \\ \vdots \\ R^{-1/2} D_W \end{bmatrix} S \in \mathbb{R}^{2W \times 6}, \qquad A = U \Sigma V^\top, \tag{7}$$

so that the directional observability magnitudes are the singular values σ_i directly and the directions are the columns of V, without $\mathcal{W}$ being formed. Section 5.1 shows that this is necessary rather than merely preferable in the near-planar geometry considered here. The state metric S is included in the definition of A, so every singular value and direction reported in this paper is that of the scaled map. Singular values are indexed in increasing order, $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_6$, so that σ_1 corresponds to the smallest singular value and $\boldsymbol{u}_1$ to its right singular vector, corresponding to the least-observable direction. Results are reported in terms of σ_i throughout, since that is the quantity the square-root form returns and it shares units with the first-order response of section 3.4. The eigenvalues of the assembled Gramian are their squares, $\lambda_i = \sigma_i^2$, and are used only in section 5.1, where the comparison between the assembled and square-root forms is itself the subject.

3.1 Hermann-Krener (H-K) observability

The H-K observability matrix stacks the gradients of successive Lie derivatives of the measurement function along the dynamics, evaluated at a single instant on the reference trajectory. Up to order k,

$$O_k = \begin{bmatrix} \nabla L_{\boldsymbol{f}}^0 \boldsymbol{h}(\boldsymbol{x}) \\ \nabla L_{\boldsymbol{f}}^1 \boldsymbol{h}(\boldsymbol{x}) \\ \vdots \\ \nabla L_{\boldsymbol{f}}^k \boldsymbol{h}(\boldsymbol{x}) \end{bmatrix}, \tag{8}$$

where $L_{\boldsymbol{f}}^0 \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})$, $L_{\boldsymbol{f}}^1 \boldsymbol{h}(\boldsymbol{x}) = \nabla \boldsymbol{h}(\boldsymbol{x}) \boldsymbol{f}(\boldsymbol{x})$, and higher orders follow recursively as $L_{\boldsymbol{f}}^{k+1} \boldsymbol{h}(\boldsymbol{x}) = \nabla L_{\boldsymbol{f}}^k \boldsymbol{h}(\boldsymbol{x}) \boldsymbol{f}(\boldsymbol{x})$. The derivatives are computed by automatic differentiation, which avoids manual derivation of the Lie derivatives and gives exact values rather than finite-difference approximations [18]. The system is locally observable at $\boldsymbol{x}$ if O_k is full rank for some k, and the singular value decomposition of O_k gives directional sensitivity scores. Both are instantaneous properties: the order k controls how much of the dynamics is folded into the assessment, not how long the measurements are collected for.

The number of Lie derivatives required for full rank is not known in advance and depends on the geometry, so k must be increased until the rank test passes, and each additional order requires differentiating through the previous one. Section 5.2 shows that the order at which full rank is first attained differs between the two geometries considered here.

3.2 Empirical observability Gramian (EOG)

The EOG obtains D_j in eq. (5) by finite differences, requiring no analytical derivative beyond the ability to integrate the dynamics. The reference state is perturbed by $\pm\delta_{\text{pos}}$ along each position axis and $\pm\delta_{\text{vel}}$ along each velocity axis, each perturbed state is propagated through the nonlinear dynamics, and the resulting noiseless bearing predictions are differenced at every step in the window. Perturbation magnitudes follow

$$\delta_{\text{pos}} = \epsilon, \qquad \delta_{\text{vel}} = v_{\text{char}}\,\epsilon, \tag{9}$$

so that the perturbation is isotropic in the scaled coordinates of eq. (6). The common alternative $\delta_{\text{vel}} = 2\pi\epsilon/T_{\text{orbit}}$ expresses the same intent through the orbital timescale, which is a fair choice for near circular orbits but not a representative value for more complex geometries, such as the butterfly. Then D_j is obtained as

$$D_j = \left[\cdots \quad \frac{\boldsymbol{h}(\boldsymbol{\phi}(t_j;\boldsymbol{x}_0+\delta_i\boldsymbol{e}_i)) - \boldsymbol{h}(\boldsymbol{\phi}(t_j;\boldsymbol{x}_0-\delta_i\boldsymbol{e}_i))}{2\delta_i} \quad \cdots\right] \in \mathbb{R}^{2\times 6}, \tag{10}$$

the i-th column being the symmetric difference of the step-j measurement prediction under a perturbation along state axis i, divided by the perturbation applied along that axis. Because the difference is symmetric, the resulting sensitivity is exactly first order, and second-order content cancels by construction.

3.3 Sigma-point observability Gramian (SP-EOG)

SP-EOG replaces EOG's axis-aligned, symmetric, arbitrarily sized perturbation with an ensemble of $2n+1$ Van der Merwe sigma points ($n = 6$) drawn from an assumed mean and covariance $(\hat{\boldsymbol{x}}_k, P_k)$ at the start of the window,

$$\boldsymbol{\chi}_0 = \hat{\boldsymbol{x}}_k, \qquad \boldsymbol{\chi}_i = \hat{\boldsymbol{x}}_k + \left[\sqrt{(n+\lambda)P_k}\right]_i, \qquad \boldsymbol{\chi}_{n+i} = \hat{\boldsymbol{x}}_k - \left[\sqrt{(n+\lambda)P_k}\right]_i, \quad i = 1,\ldots,n, \tag{11}$$

$$w_m^{(0)} = \frac{\lambda}{n+\lambda}, \qquad w_c^{(0)} = w_m^{(0)} + 1 - \alpha^2 + \beta, \qquad w_m^{(i)} = w_c^{(i)} = \frac{1}{2(n+\lambda)}, \ \ i \geq 1, \tag{12}$$

with $\lambda = \alpha^2(n+\kappa) - n$ and $[\sqrt{\cdot}]_i$ the i-th column of a Cholesky factor. Sigma points are indexed by $i = 0,\ldots,2n$ and window steps by $j = 1,\ldots,W$, matching eq. (5). Each sigma point is propagated through the nonlinear relative dynamics to every step in the window, producing a noiseless bearing prediction $\boldsymbol{z}_j^{(i)}$. The weighted mean prediction and the cross-covariance between the initial-time deviations and the step-j measurement deviations are

$$\bar{\boldsymbol{z}}_j = \sum_i w_m^{(i)}\,\boldsymbol{z}_j^{(i)}, \qquad P_{xz,j} = \sum_i w_c^{(i)}\left(\boldsymbol{\chi}_i - \hat{\boldsymbol{x}}_k\right)\left(\boldsymbol{z}_j^{(i)} - \bar{\boldsymbol{z}}_j\right)^\top, \tag{13}$$

from which an effective sensitivity follows by linear regression through the ensemble,

$$D_j^\top = P_k^{-1} P_{xz,j}, \tag{14}$$

accumulating as in eq. (5), and evaluated in the square-root form of eq. (7).

D_j is fit through the nonlinear propagation of a finite-spread ensemble rather than differentiated at a point, so curvature over the sampled ellipsoid is folded into the regression. Whether $\mathcal{W}_{\text{SP}}$ departs from $\mathcal{W}_{\text{EOG}}$ therefore depends on the ensemble spread. Within the radius over which the measurement map is effectively linear, the regression recovers the same first-order sensitivity as EOG, which section 5.3 shows is the case for the covariances used here. A larger ensemble captures some of the curvature EOG cancels, but one so large that the map is no longer injective over it recovers no meaningful sensitivity at all. Locating that upper limit is left to future work.

Weighting the accumulation by the measurement noise covariance, as in eq. (5), makes $\mathcal{W}_{\text{SP}}$ the Fisher information matrix of the initial state given the measurement sequence [19], and its eigenvalues carry units of achievable estimation accuracy instead of an arbitrary perturbation scale.

The source of P used in this work is referred to as PCRB-P. This is the achievable estimation covariance along the reference trajectory, obtained from a truth-pinned square-root UKF. At every step the filter state is held to the truth trajectory and updated against the truth's own noiseless measurement, so the innovation is zero and only the covariance evolves,

$$P_k^- = \mathrm{predict}(P_{k-1}^+, Q{=}0), \qquad P_k^+ = \mathrm{update}(P_k^-, R), \tag{15}$$

giving the best-case achievable covariance under the true nonlinear dynamics and measurement model.

3.4 Higher-order Gramian (HOG)

The three tools above return a first-order picture, whether by construction or by regression, and none of them reports how large the neglected second-order content actually is. HOG answers that by expanding the same measurement map to second order and retaining the quadratic term rather than cancelling it.

Let $G : \mathbb{R}^6 \to \mathbb{R}^{2W}$ map a relative state $\boldsymbol{x}$ to the stacked noiseless bearing predictions over a window of W steps, and consider a deviation $\delta\boldsymbol{x}$ to that state. The Taylor expansion of G about the reference trajectory is

$$G(\boldsymbol{x}+\delta\boldsymbol{x}) = G(\boldsymbol{x}) + G_1(\boldsymbol{x})\,\delta\boldsymbol{x} + \tfrac{1}{2}\,\delta\boldsymbol{x}^\top G_2(\boldsymbol{x})\,\delta\boldsymbol{x} + O(\|\delta\boldsymbol{x}\|^3), \tag{16}$$

with $G_1 \in \mathbb{R}^{2W\times 6}$ the Jacobian and $G_2 \in \mathbb{R}^{2W\times6\times6}$ the second-derivative tensor, both computed exactly by automatic differentiation. The expansion is truncated at second order here, but nothing in the construction requires it: an order-m truncation needs the m-th derivative tensor, whose evaluation cost and storage grow as n^m in the state dimension, and the quantities defined below extend accordingly. Writing a deviation of size ε along a unit direction $\boldsymbol{u}$ as $\delta\boldsymbol{x} = \varepsilon\boldsymbol{u}$, the linear and quadratic terms become $\varepsilon(G_1\boldsymbol{u})$ and $\frac{\varepsilon^2}{2}\boldsymbol{c}(\boldsymbol{u})$, with $\boldsymbol{c}(\boldsymbol{u})_i := \boldsymbol{u}^\top G_2[i]\,\boldsymbol{u}$. Weighting both by the stacked measurement noise $\mathcal{R} = I_W \otimes R$ gives the first- and second-order response magnitudes

$$\mathrm{fo}(\boldsymbol{u}) := \|\mathcal{R}^{-1/2} G_1 \boldsymbol{u}\|, \qquad \mathrm{so}(\boldsymbol{u}) := \|\mathcal{R}^{-1/2}\boldsymbol{c}(\boldsymbol{u})\|. \tag{17}$$

Evaluated along the right singular vectors of $A = \mathcal{R}^{-1/2}G_1$, $\mathrm{fo}(\boldsymbol{u}_i) = \sigma_i$, so the first-order part of HOG is the exact-derivative limit of the EOG Gramian and the two agree numerically at an appropriate perturbation scale (section 5.3). HOG's contribution is what can be extracted from G_2. The two terms are equal in size at

$$\varepsilon^\star(\boldsymbol{u}) = \frac{2\,\mathrm{fo}(\boldsymbol{u})}{\mathrm{so}(\boldsymbol{u})}, \tag{18}$$

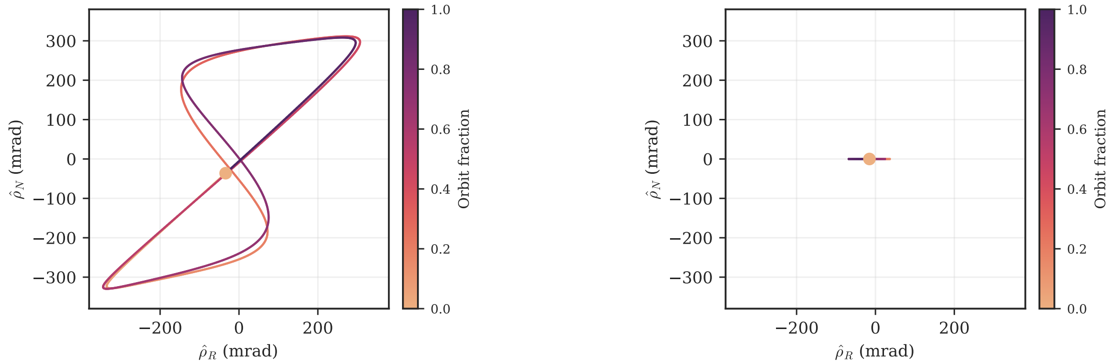
so $\varepsilon^\star(\boldsymbol{u})$ is a direction-dependent radius beyond which the second-order term is no longer a small correction, and the linear characterization shared by the other three tools, and by any filter relying on the same linearization, ceases to hold along $\boldsymbol{u}$.[1]

This study evaluates $\mathrm{so}(\boldsymbol{u})$ and $\varepsilon^\star(\boldsymbol{u})$ only along the six right singular vectors of A, which are the directions the first-order tools already single out. That is a deliberately narrow slice: G_2's directional structure is in general unrelated to G_1's, and nothing requires the direction of greatest curvature to lie near any first-order singular vector, so restricting to the first-order basis samples the second-order structure rather than characterizing it. The sample is enough for a targeted question: does curvature exist specifically where the first-order picture reports little information? A large $\mathrm{so}(\boldsymbol{u}_1)$ means information is present along the least-observable direction but inaccessible to any linear sensitivity, and by extension to a single Gaussian; a small one means that direction is uninformative at second order too. Locating the true extrema of $\mathrm{so}(\boldsymbol{u})$ over the unit sphere is left to future work.

4. SIMULATION SETUP

Two CR3BP orbit families are considered, chosen because they differ in relative geometry rather than for mission-specific reasons. The northern butterfly orbit is fully three-dimensional and its relative geometry evolves substantially over an orbital period; the distant retrograde orbit is near-planar and close to circular, so its relative geometry is nearly confined to a plane and varies little with phase. Figure 1 shows both projected onto the observer's image plane: the butterfly traces a well-spread two-dimensional figure, the DRO a narrow nearly straight line. Planar relative geometries are known to be problematic for angles-only observability [7, 17], and section 5.2 quantifies how that difficulty appears in the tools used here. The butterfly serves as the case in which the first-order tools are expected to be adequate and the DRO as the case in which they are tested.

[1] Summarizing this second-order vs. first-order behavior by a single ratio $\rho(\boldsymbol{u}) := \mathrm{so}(\boldsymbol{u})/\mathrm{fo}(\boldsymbol{u})$ maximized over $\boldsymbol{u}$ is degenerate. The ratio scales linearly with $\|\boldsymbol{u}\|$ rather than being scale-free, so its maximizer depends on an arbitrary normalization, and the natural choice collapses the optimizer onto the least-observable direction regardless of the structure of G_2.



(a) Northern butterfly orbit

(b) Distant retrograde orbit (DRO)

Fig. 1: Projection of the relative trajectories onto the observer's image plane. Color indicates progression along the orbit.

In both families the observer and target are placed in a lead-follower configuration on the same periodic orbit, with the target's initial state obtained by propagating the observer's state forward along the orbit by an offset corresponding to a relative separation of 0.5% of the orbit period. The resulting separation distances can be seen in figs. 2 and 3. Observability quantities for the finite-time tools are evaluated on a sliding forward window of W measurement steps, advanced along the reference trajectory with a fixed stride, over one full orbital period.

A measurement is assumed available at every epoch, with no field-of-view mask, exclusion angle, illumination condition, or range limit, and the sensor is taken to point at the target throughout. The quantities reported are therefore an upper bound on the information the next W measurements could supply, since any real sensor constraint removes measurements from the window. This matters most in the butterfly, where the elevation angle passes through zero and the separation varies by a factor of two over an orbital period.

Table 2 collects the scenario and numerical parameters. The two orbits are taken at closely matched Jacobi constants so that the comparison between them reflects the difference in relative geometry rather than in energy, and the measurement cadence is set to $T_{\text{orbit}}/1000$ in both cases so that a window of W steps spans the same fraction of an orbit in each.

Automatic differentiation through an adaptive-step integrator would differentiate through the step-size controller, whose step selections are discontinuous in the state, so all propagations use a fixed-step integrator. Section 5.1 shows that results are insensitive to the step size over the range tested.

Table 3 summarizes the initial conditions for the two relative trajectories, in nondimensional units. Both the observer and the relative state are expressed in the rotating frame of the CR3BP, with the relative state defined as the target minus the observer. The observer's initial state is taken from a JPL periodic-orbit database, and the relative state is obtained by propagating the observer's state forward along the orbit by an offset corresponding to a relative separation of 0.5% of the orbit period.

Figures 2 and 3 show the evolution of the relative trajectory in range, azimuth, and elevation coordinates, and the corresponding rates. The butterfly exhibits significant variation in all three coordinates and their rates, while the DRO shows limited variation, particularly in elevation, reinforcing its near-planar nature.

5. TOOL CONFIGURATION AND CROSS-TOOL AGREEMENT

The four tools differ in construction, so before their results can be read as statements about the system, the conditions under which they agree have to be established. Three choices matter: how the Gramian spectrum is computed, how much measurement information each tool is given, and how far from the reference trajectory the sampling-based tools probe.

Table 2: Scenario and numerical parameters.

Parameter	Symbol	Value
Scenario		
Mass parameter	μ	$1.215058560962404e-2$
Length unit	LU	389,703 km
Time unit	TU	382,981 s
Relative separation	–	0.5% of orbit period
Measurement cadence	–	$T_{\text{orbit}}/1000$
Measurement noise (1σ)	σ_{meas}	10^{-4} rad
Northern butterfly orbit		
Orbit period	T_{orbit}	5.173114 TU
Jacobi constant	C_J	3.0501307061
Distant retrograde orbit		
Orbit period	T_{orbit}	1.054876 TU
Jacobi constant	C_J	3.0501085628
Observability tools		
Window length	W	5 (see section 5.2)
Window stride	–	10
EOG perturbation	ϵ	10^{-6} (see section 5.3)
Sigma-point parameters	α, β, κ	1, 2, 0
Integrator step divisor	–	20 (see section 5.1)

Table 3: Initial conditions for the two relative trajectories, in nondimensional units.

	Northern butterfly		Distant retrograde	
	Observer	Relative	Observer	Relative
x	0.9375794205582276	-0.00011226420899801415	0.9115899420466983	3.9216244273165834e-05
y	0.0	-0.004720003492731605	0.0	2.5745680462596920e-03
z	0.15321507174956706	-0.00020539462721458346	0.0	0.0
$\dot{x}$	0.0	-0.008676418138965032	0.0	1.4868879684516350e-02
$\dot{y}$	-0.18261821484292484	0.0004083469239579973	0.4882158943627189	-2.6617376674670234e-04
$\dot{z}$	0.0	-0.01588169195074546	0.0	0.0

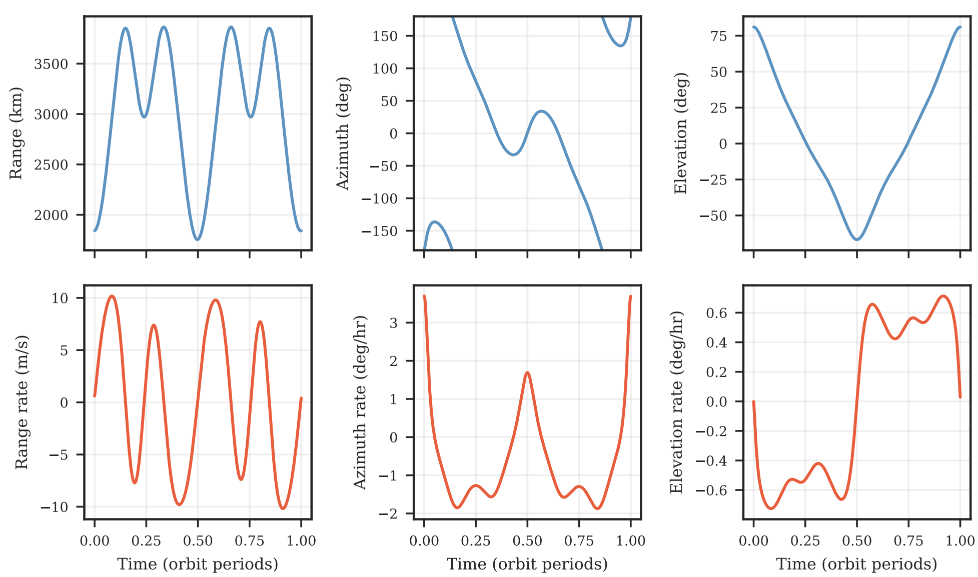


Fig. 2: Evolution of the relative trajectory in range, azimuth, and elevation coordinates, and the corresponding rates, for the northern butterfly orbit.

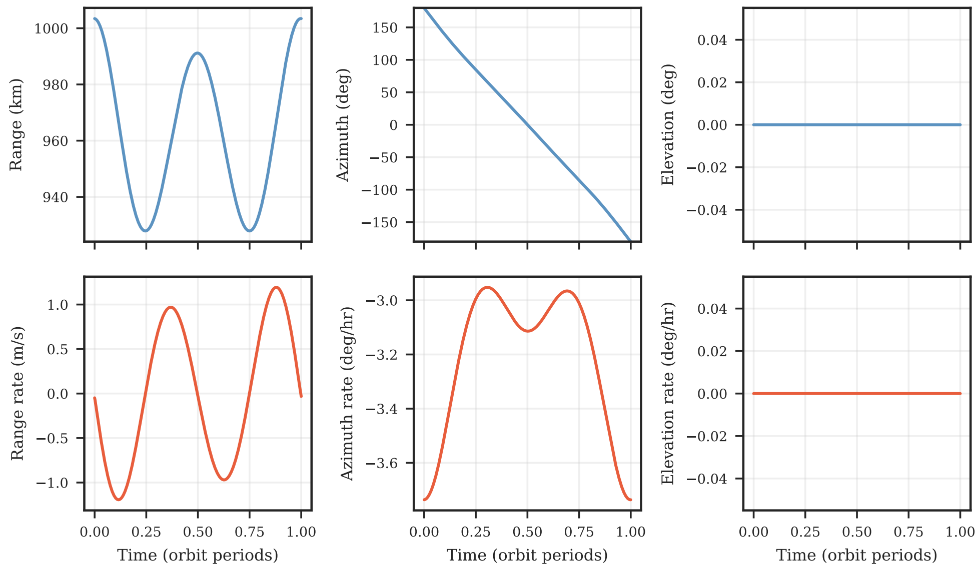


Fig. 3: Evolution of the relative trajectory in range, azimuth, and elevation coordinates, and the corresponding rates, for the distant retrograde orbit (DRO).

Table 4: Conditioning of the Gramian computation.

Orbit	Form	$\lambda_{\min}$ (range)	$\lambda_{\max}$ (range)	cond (max)	n_{neg}
Butterfly	full	$[3.141\times10^{-2}, 7.949\times10^{1}]$	$[5.551\times10^{12}, 2.463\times10^{13}]$	1.767×10^{14}	0/100
	square-root	$[3.112\times10^{-2}, 7.949\times10^{1}]$	$[5.551\times10^{12}, 2.463\times10^{13}]$	1.335×10^{7}	0/100
DRO	full	$[-1.468\times10^{-2}, 2.597\times10^{-2}]$	$[7.534\times10^{13}, 8.809\times10^{13}]$	6.671×10^{18}	47/100
	square-root	$[3.192\times10^{-11}, 6.306\times10^{-5}]$	$[7.534\times10^{13}, 8.809\times10^{13}]$	1.599×10^{12}	0/100

5.1 Numerical form

Every spectrum reported here is taken from the singular value decomposition of the whitened sensitivity matrix A of eq. (7) rather than from the assembled Gramian $\mathcal{W} = A^\top A$, and in the near-planar geometry that is a requirement rather than a preference. In the DRO, A has a maximum condition number of $\mathrm{cond}(A) \approx 1.6\times10^{12}$, which double precision resolves comfortably; $\mathcal{W}$ has one of $\sim 10^{18}$, which it does not. Computed from $\mathcal{W}$ directly at $W = 5$, the smallest eigenvalue comes back negative in about half the windows, at a magnitude matching the roundoff floor $\varepsilon_{\mathrm{mach}}\lambda_{\max}$ and with a sign that changes between runs. A negative eigenvalue is not a small value but an impossible one, since $\mathcal{W}$ is positive semidefinite by construction, so it points to the computation and not the system.

Taking the spectrum from the singular value decomposition of A removes the effect. Table 4 reports the number of windows returning a negative smallest eigenvalue, together with the range of $\lambda_{\min}$ and $\lambda_{\max}$, and the maximum condition number, for both geometries. The condition number reported for the assembled form is that of $\mathcal{W}$ and for the square-root form that of A, and the two should differ by a square. In the butterfly they do: $(1.335\times10^{7})^2 = 1.78\times10^{14}$, which is what the assembled form returns, and the geometry is insensitive to the choice of form because 10^{14} is still manageable in double precision. In the DRO they do not. The assembled 6.671×10^{18} is four orders below the square of 1.599×10^{12}, because the assembled $\lambda_{\min}$ has been replaced by the roundoff floor rather than resolved. The reported condition number is not a measurement of the problem but an artifact of the same precision loss that produces the negative eigenvalues. The square-root form returns no negative values in any window of either geometry, and is necessary rather than merely preferable in the DRO.

Automatic differentiation propagates gradients through every operation in the computational graph, so differentiating G_1 and G_2 through an adaptive-step integrator would differentiate through the step-size controller itself, whose accept/reject decisions are discontinuous in the state. All propagations feeding G_1 and G_2 therefore use a fixed step size. Over a tenfold range of the number of integration steps per measurement step, the butterfly spectrum is unchanged to four significant figures, and in the DRO $\sigma_{\max}$ is unchanged while $\sigma_{\min}$ varies by under 20% with no trend in the divisor. The results are not sensitive to that choice.

5.2 Information threshold

Each tool has a parameter controlling how much measurement information enters the assessment: the Lie derivative order k for H-K, and the window length W for the three window-based tools. Both admit a minimum value at which the stacked sensitivity is square, with as many rows as the six-dimensional state and no redundancy: $k = 2$ gives three sets of two angle rows, and $W = 3$ gives three measurement pairs.

Table 5: Rank, singular value range and conditioning at and just above the minimum information setting: Lie derivative order k for H-K, window length W for the window tools.

Orbit	Setting	% deficient	$\sigma_{\min}$ (range)	$\sigma_{\max}$ (range)	cond (max)
Butterfly	H-K, $k = 2$	0.0%	$[1.88\times10^{3}, 1.82\times10^{4}]$	$[2.09\times10^{7}, 6.39\times10^{7}]$	1.36×10^{4}
	H-K, $k = 3$	0.0%	$[4.94\times10^{3}, 1.41\times10^{5}]$	$[9.20\times10^{7}, 6.84\times10^{8}]$	1.36×10^{5}
	$W = 3$	0.0%	$[1.65\times10^{-2}, 1.95\times10^{0}]$	$[1.82\times10^{6}, 3.84\times10^{6}]$	1.11×10^{8}
	$W = 4$	0.0%	$[4.26\times10^{-2}, 4.77\times10^{0}]$	$[2.11\times10^{6}, 4.44\times10^{6}]$	4.94×10^{7}
	$W = 5$	0.0%	$[1.76\times10^{-1}, 8.92\times10^{0}]$	$[2.36\times10^{6}, 4.96\times10^{6}]$	1.34×10^{7}
DRO	H-K, $k = 2$	100.0%	$[3.86\times10^{-11}, 1.55\times10^{-8}]$	$[2.45\times10^{8}, 2.59\times10^{8}]$	6.51×10^{18}
	H-K, $k = 3$	0.0%	$[2.15\times10^{4}, 1.41\times10^{5}]$	$[7.44\times10^{8}, 1.99\times10^{9}]$	6.18×10^{4}
	$W = 3$	100.0%	$[9.68\times10^{-13}, 1.36\times10^{-10}]$	$[6.72\times10^{6}, 7.27\times10^{6}]$	6.96×10^{18}
	$W = 4$	0.0%	$[2.52\times10^{-6}, 2.80\times10^{-3}]$	$[7.76\times10^{6}, 8.39\times10^{6}]$	3.21×10^{12}
	$W = 5$	0.0%	$[5.65\times10^{-6}, 7.94\times10^{-3}]$	$[8.68\times10^{6}, 9.39\times10^{6}]$	1.60×10^{12}

At those minimum settings the two geometries behave differently. In the DRO both the H-K matrix at $k = 2$ and the window Gramian at $W = 3$ are exactly rank-deficient by one at every sampled point along the orbit, with a maximum condition number above 10^{18}. The butterfly is full rank and well conditioned throughout, at 1.11×10^{8} in the worst case. Adding one Lie derivative order, or one further measurement pair to the window, resolves the deficiency in both cases.

The same threshold appears through two independent parameters, which indicates a property of the relative geometry and not of either tool. The near-planar configuration requires redundant measurements before its weakest direction is determined at all, and the butterfly does not. This is a more precise statement than the usual one that planar angles-only geometries are rank-deficient: they are rank-deficient at minimum information, and observable but ill-conditioned at low information.

The window quantities in table 5 are computed from HOG's exact first-order part. EOG and SP-EOG, configured as in section 5.3, return the same rank and similar singular value ranges and are omitted for brevity.

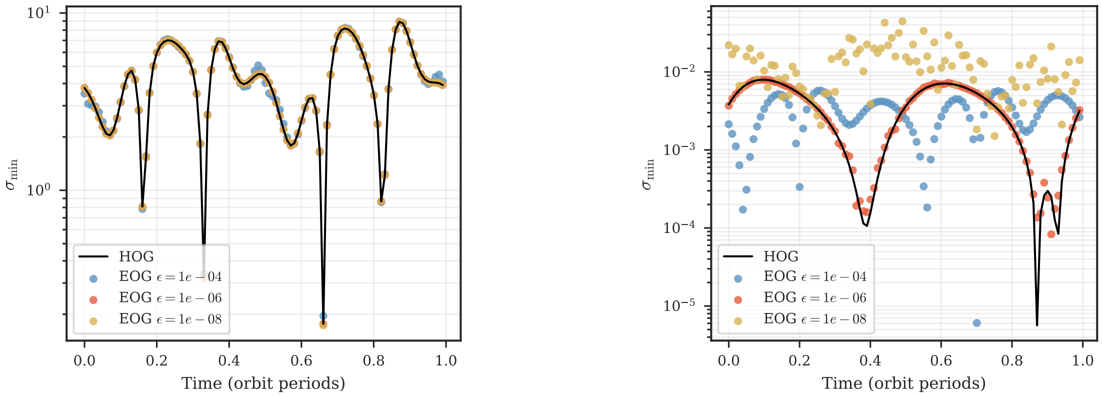
5.3 Sampling scale

EOG and SP-EOG both sample the nonlinear flow at a finite distance from the reference trajectory, EOG through the perturbation ϵ and SP-EOG through the assumed covariance P. HOG's first-order part requires neither and provides the reference against which both are calibrated. Figure 4 shows the smallest singular value from HOG's G_1 alongside EOG at several ϵ. In the butterfly every value between 10^{-4} and 10^{-8} reproduces the exact result throughout the orbit, including the depth and phase of the dips.

In the DRO the tolerance is much narrower. At $\epsilon = 10^{-6}$ the agreement is close in both magnitude and phase dependence, with the residual departures confined to the two phases at which $\sigma_{\min}$ collapses and where a first-order description of the measurement response is least accurate. At $\epsilon = 10^{-4}$ the smallest singular value is high by tens of percent and the phase structure is lost. At $\epsilon = 10^{-8}$ differencing error dominates and the result is high by more than an order of magnitude across the whole orbit. The admissible range is therefore about one order of magnitude wide, and neither of its edges announces itself: outside it EOG returns a smooth, plausible curve of the wrong value.

The reason the two geometries differ in where that range lies, and in how wide it is, is a property of the measurement map along the weakest direction rather than of the tool, and is taken up in section 6.2.

SP-EOG's covariance plays the same role as ϵ. Figure 5 compares the full spectra of EOG and SP-EOG at PCRB-P for both geometries: they agree across all six directions and throughout the orbit, with visible departures confined to

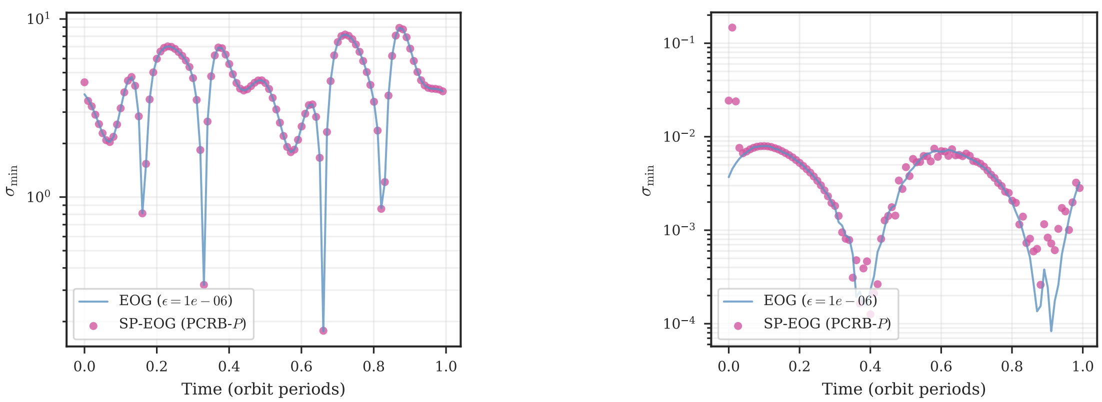


(a) Northern butterfly

(b) Distant retrograde

Fig. 4: Smallest eigenvalue of the window Gramian along both trajectories, computed from HOG's exact first-order part and from EOG at several perturbation scales ϵ.

the smallest singular value, in the few steps before the covariance converges and, in the DRO, at the phases where the first-order response collapses. Those are the same phases at which the EOG perturbation scale becomes critical, and the two tools are limited there by the same conditioning rather than by their different sampling schemes.



(a) Northern butterfly

(b) Distant retrograde

Fig. 5: Smallest eigenvalue of the window Gramian along both trajectories, computed from EOG and SP-EOG at PCRB-P.

That agreement is expected: PCRB-P is a converged estimation covariance, so the ensemble it generates is narrow, and a regression through a narrow ensemble recovers the same linear sensitivity a symmetric difference does. The statement is conditional. At small covariance the linear characterization is adequate and either tool serves, so conditioning on P buys interpretability rather than a different value; the two should differ only once the covariance is large enough for the sampled ellipsoid to extend past the linear-dominated neighbourhood.

5.4 Agreement

Figures 6 and 7 show the six singular values returned by all four tools along both trajectories, at the configurations established above: $W = 5$ for the window-based tools, and for H-K the lowest order at which the geometry reaches full rank, $k = 2$ in the butterfly and $k = 3$ in the DRO (section 5.2).

The three window-based tools agree throughout, in every direction and at every phase, which is the claim the rest of the paper relies on. H-K sits one to two orders above them and is not expected to agree: its rows are gradients of successive Lie derivatives and carry successive inverse powers of time, whereas the window tools stack W copies of the

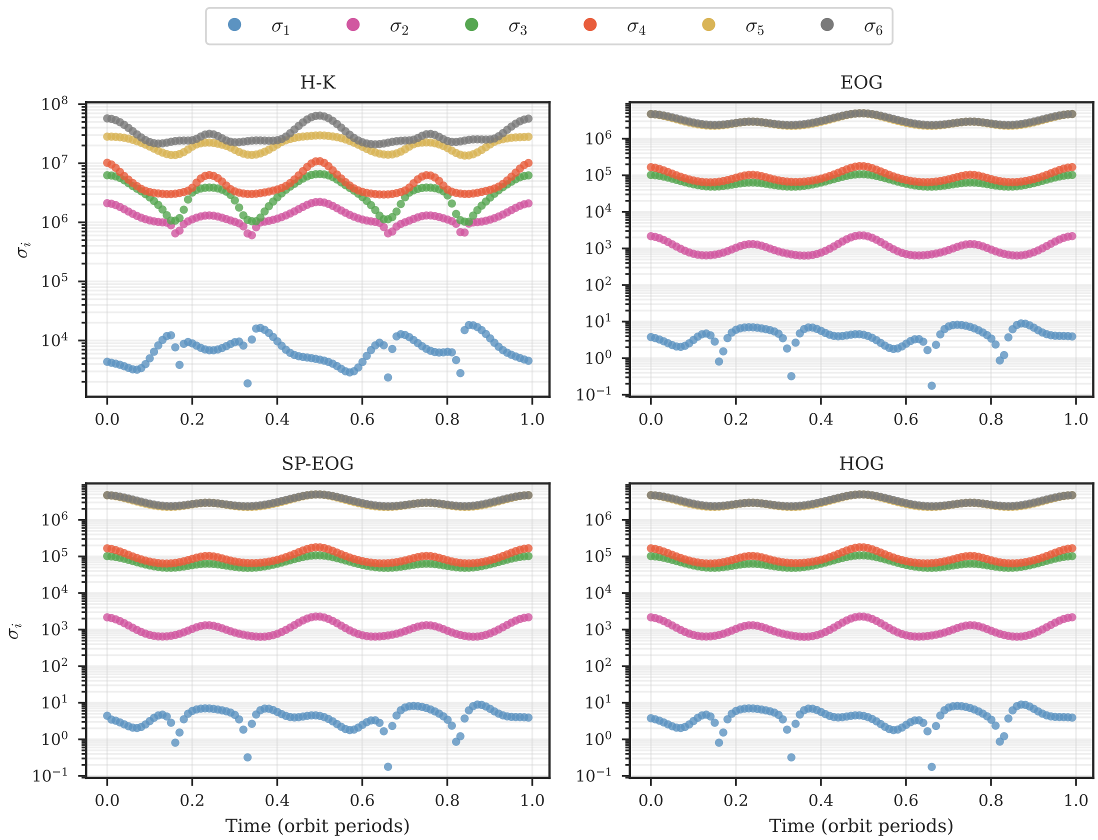


Fig. 6: Six eigenvalues of the window Gramian for the butterfly geometry, computed from each of the four tools.

same noise-weighted sensitivity, so only the ordering and the directions are comparable. That ordering is preserved, with the same four-order separation between σ_1 and the remaining five.

Because the splitting and pruning arguments of section 8 act on directions, the directions themselves also have to agree. Figure 8 shows the principal angle between the least-observable direction returned by each tool and the one returned by HOG's exact first-order part, which is the only parameter-free reference available. EOG and SP-EOG are indistinguishable from the HOG reference at plotting resolution, below 0.01° throughout both orbits: the DRO's weakest direction is recovered identically by all three window-based tools despite its first-order response lying four orders below every other. H-K departs by at most 2.2° in the DRO and 1.6° in the butterfly. Its instantaneous construction therefore recovers essentially the same weakest direction as the window-based tools, a point returned to in section 7.

The DRO $k = 2$ null vector is also shown. Since the H-K matrix is exactly rank-deficient there (section 5.2), that null vector rather than a smallest singular vector is the object to compare, and it agrees with the reference to below 0.01° over the second half of the orbit, more closely than the $k = 3$ singular vector does anywhere. The weak direction is therefore fixed by the structure of the rank deficiency, and the 2.2° residual at $k = 3$ is the cost of resolving it as a small singular value rather than an exact null vector at finite precision. The 1.6° departure in the butterfly reads the same way, with no exact deficiency at any order to anchor it.

6. OBSERVABILITY STRUCTURE

6.1 First-order structure

Figures 9 and 10 show the first-order response $\text{fo}(\boldsymbol{u}_i)$ along the six directions of eq. (17), ordered from least to most observable, over one orbital period at $W = 5$.

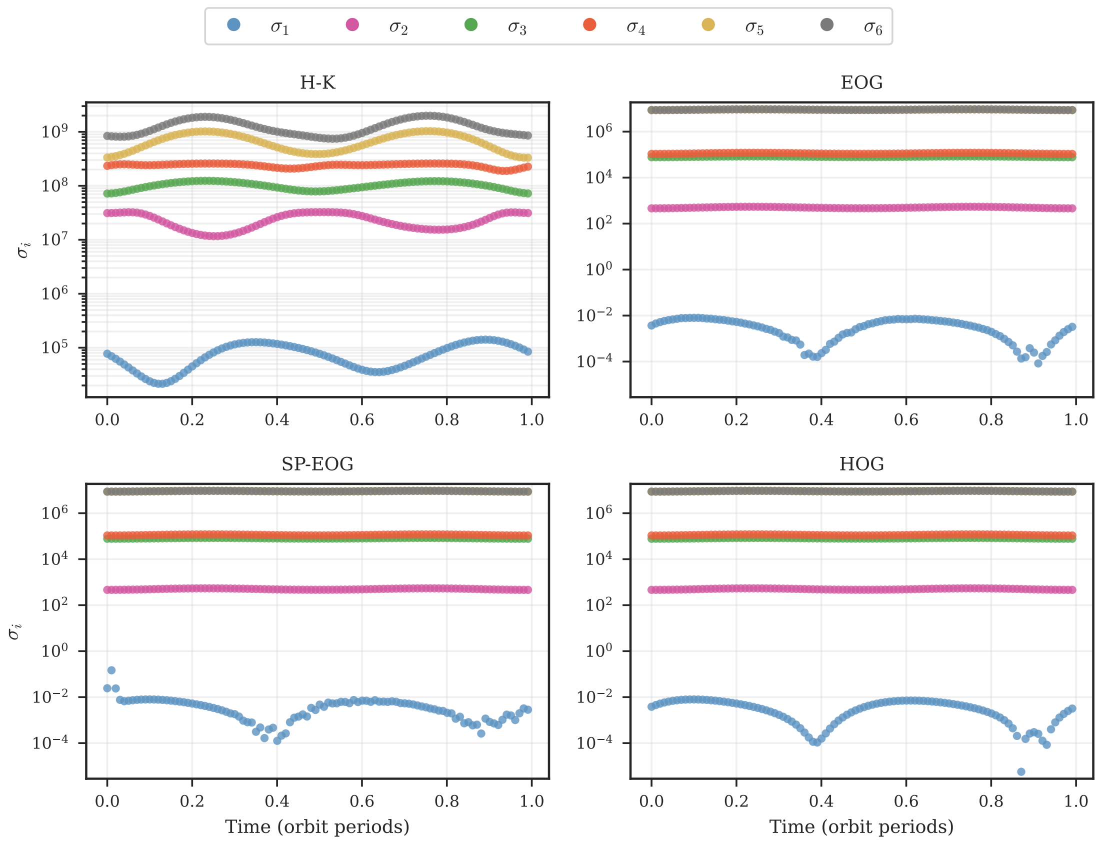
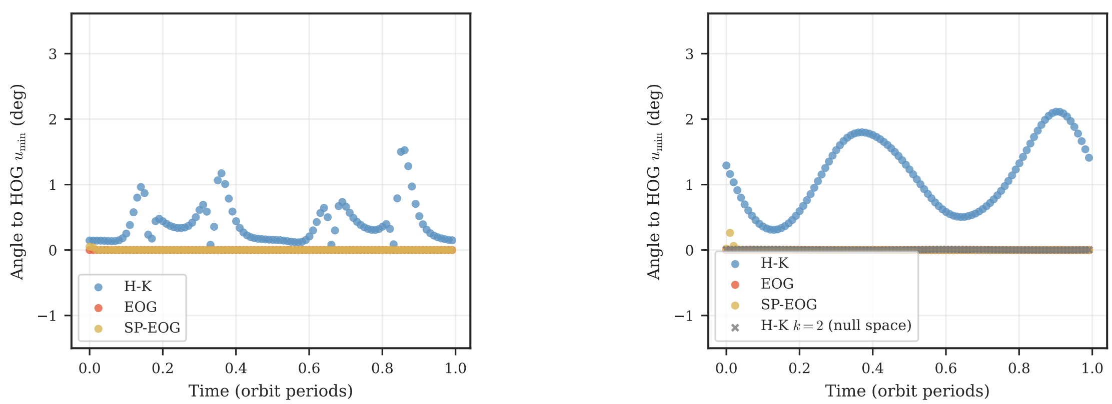


Fig. 7: Six eigenvalues of the window Gramian for the DRO geometry, computed from each of the four tools.



(a) Northern butterfly

(b) Distant retrograde

Fig. 8: Principal angle between the least-observable direction returned by each tool and the one returned by HOG's exact first-order part, along both trajectories.

The two geometries differ in the shape of the spectrum. In the butterfly the six values are graded: they span roughly six orders of magnitude, from $\mathrm{fo}(\boldsymbol{u}_1) \approx 10^0$ along the weakest to $\mathrm{fo}(\boldsymbol{u}_6) \approx 10^7$ along the strongest, with no direction separated from the rest by a large gap. All directions carry information and the weakest carries six orders less than the strongest. In the DRO the spectrum is more separated: five directions lie between $\mathrm{fo}(\boldsymbol{u}_2) \approx 10^3$ and $\mathrm{fo}(\boldsymbol{u}_5) \approx 10^7$, and the smallest singular value sits near $\mathrm{fo}(\boldsymbol{u}_1) \approx 5 \times 10^{-2}$, four to five orders below the next-weakest and nine below the

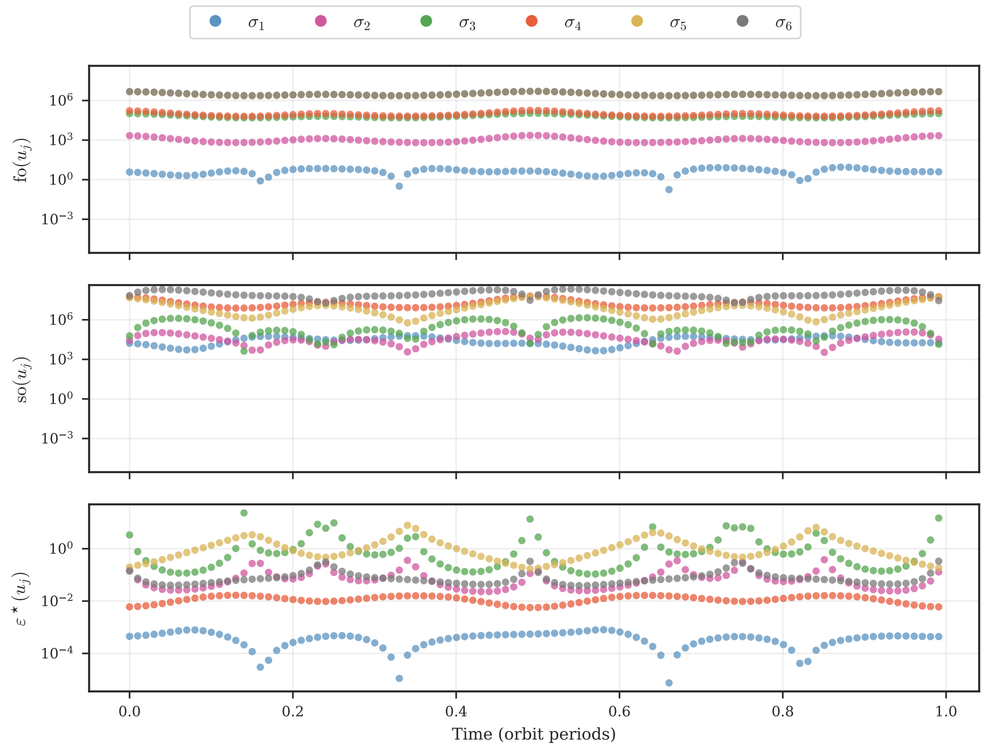


Fig. 9: First- and second-order response along the six right singular vectors of the window Gramian, and their critical radius $\varepsilon^\star$ for the butterfly geometry.

strongest, with no intermediate values at any orbital phase. The gap is a property of the geometry: it is invariant to integration step size (section 5.1) and returned identically by all three window-based tools at an appropriate sampling scale (section 5.3).

The two geometries differ in conditioning as well as in shape: at $W = 5$ the maximum condition number is 1.3×10^7 in the butterfly against 1.6×10^{12} in the DRO. The shape of the spectrum is the more informative distinction, since it identifies a single direction as the source of the ill-conditioning in the DRO case rather than an unevenness spread across all six.

6.2 Second-order structure

Figures 9 and 10 also show the second-order response $\mathrm{so}(\boldsymbol{u}_i)$ and the breakdown radius $\varepsilon^\star(\boldsymbol{u}_i)$ along the same six directions. Both are evaluated on the first-order singular vectors, so what follows describes the second-order content of those particular directions and not the second-order structure as a whole.

In the butterfly the second-order response is distributed more evenly than the first (four orders across the six directions against six), but the least-observable direction still carries the least curvature, $\mathrm{so}(\boldsymbol{u}_1) \approx 10^4$, separated from $\mathrm{so}(\boldsymbol{u}_6)$ by about three orders rather than the six that separate $\mathrm{fo}(\boldsymbol{u}_1)$ from $\mathrm{fo}(\boldsymbol{u}_6)$. The breakdown radius $\varepsilon^\star = 2\,\mathrm{fo}/\mathrm{so}$ is therefore still smallest along $\boldsymbol{u}_1$, collapsing below 10^{-5} in v_{char}-scaled units at four points per period: the direction the measurements say least about is also the one over which a linear description fails soonest, so a single-Gaussian filter is least accurate exactly where it is least appropriate.

The DRO separates the two properties instead of compounding them: along its weakest direction, whose first-order response is four orders below every other, the second-order response is $\mathrm{so}(\boldsymbol{u}_1) \sim 10^5$, third smallest of the six. This second-order response is three orders above the minimum, and flat across the orbit. Measurement information is available there that is inaccessible to any first-order analysis.

The reverse holds along the third-ranked direction: strongly observed at first order, $\mathrm{fo}(\boldsymbol{u}_3) \approx 10^5$, but $\mathrm{so}(\boldsymbol{u}_3) \approx 10^2$ is three orders below every other direction and flat across the period, so a single Gaussian remains adequate there over

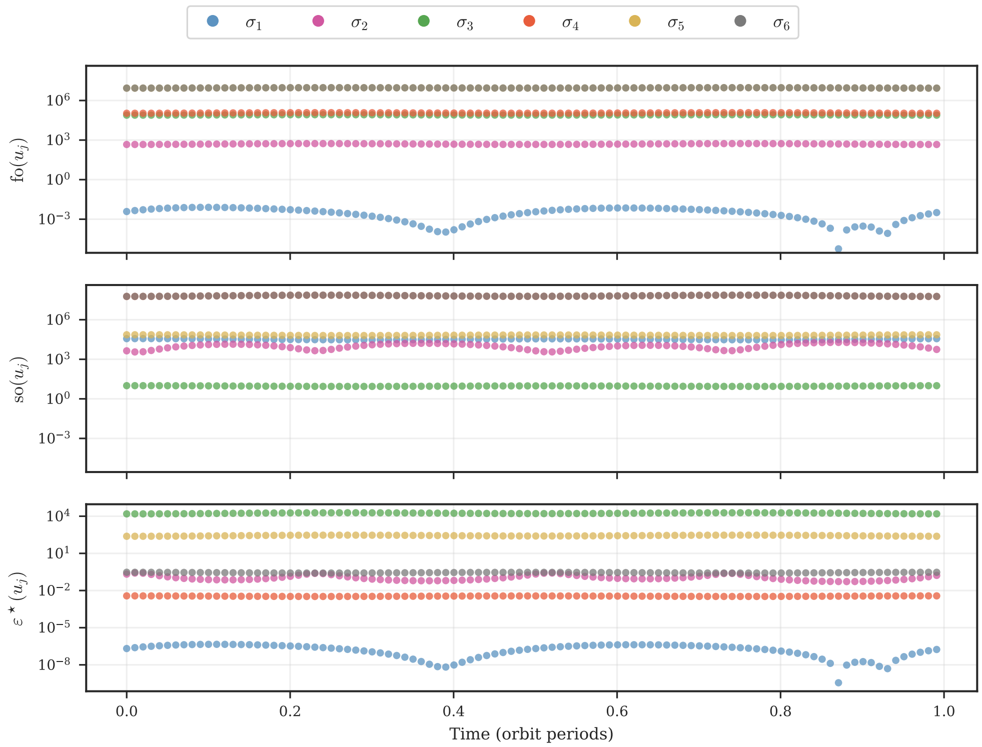


Fig. 10: First- and second-order response along the six right singular vectors of the window Gramian, and their critical radius $\varepsilon^\star$ for the DRO geometry.

Table 6: Comparison of the perturbation scale ϵ to the breakdown radius $\varepsilon^\star$ along the least-observable direction, for position and velocity blocks.

Orbit	ϵ	$\delta_{\text{pos}}/\varepsilon^*_{\text{pos}}$	$\delta_{\text{vel}}/\varepsilon^*_{\text{vel}}$
Butterfly	1×10^{-4}	$[1.649\times10^{-1}, 1.733\times10^{1}]$	$[5.524\times10^{-1}, 6.227\times10^{1}]$
	1×10^{-6}	$[1.649\times10^{-3}, 1.733\times10^{-1}]$	$[5.524\times10^{-3}, 6.227\times10^{-1}]$
DRO	1×10^{-4}	$[3.117\times10^{2}, 3.965\times10^{5}]$	$[1.860\times10^{3}, 2.345\times10^{6}]$
	1×10^{-6}	$[3.117\times10^{0}, 3.965\times10^{3}]$	$[1.860\times10^{1}, 2.345\times10^{4}]$

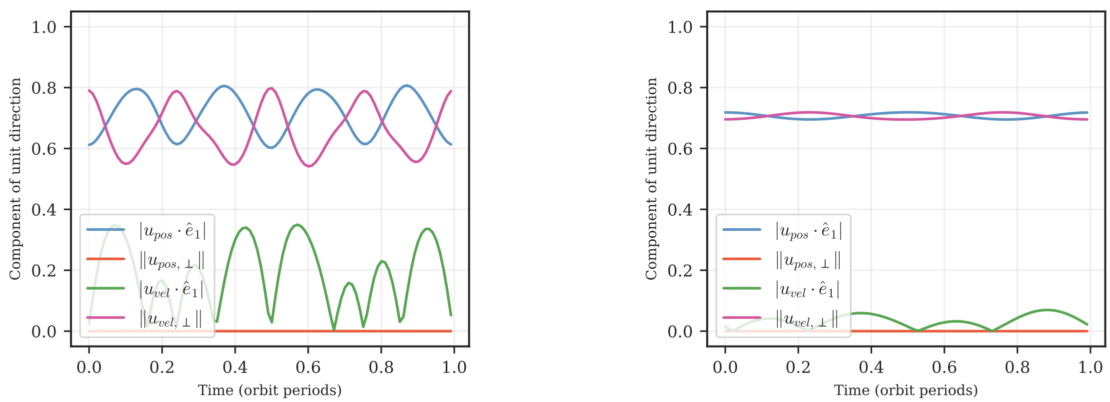
any deviation of practical interest. $\varepsilon^\star$ separates the two most sharply, spanning more than ten orders across the six DRO directions from $\boldsymbol{u}_1$ at the bottom to $\boldsymbol{u}_3$ at the top: a first-order ranking of $\boldsymbol{u}_3$ above $\boldsymbol{u}_1$ is correct but says nothing about linear validity, which fails at deviations far smaller than any realistic uncertainty along $\boldsymbol{u}_1$ and holds far beyond it along $\boldsymbol{u}_3$.

The breakdown radius also accounts for the perturbation-scale sensitivity of section 5.3. Table 6 reports the ratio of the applied perturbation to $\varepsilon^\star(\boldsymbol{u}_1)$: in the butterfly $\epsilon = 10^{-4}$ is comparable to $\varepsilon^\star$ and produced the small errors seen there, while in the DRO the same perturbation exceeds it by two to six orders and the result was unusable. The ratio orders the two geometries correctly but is not a strict threshold: the DRO EOG remains usable at $\epsilon = 10^{-6}$ despite a ratio above unity throughout. $\varepsilon^\star$ supplies an indication of scale rather than a bound. The contrast is sharpest at the information threshold of section 5.2: at $W = 3$ the DRO's first-order Gramian is exactly rank-deficient, so its weakest direction is a true null vector carrying no first-order information at all, while its second-order response along that same direction is unchanged.

The first- and second-order responses carry little information about one another: the DRO ranks the same direction first in one and third in the other, and the butterfly's agreement (both orderings place the same direction last) is a coincidence of that geometry rather than a reassurance, since reading it as license to infer curvature from a first-order spectrum gives the wrong answer in the DRO at the same Jacobi constant. Wherever a decision depends on curvature, G_2 has to be evaluated. Both quantities are evaluated only on the singular vectors of the first-order map, without locating where the second-order response is largest or smallest in an absolute sense.

6.3 Identification and phase dependence of the weakest direction

The preceding results order the state-space directions by information content without identifying them physically. Figure 11 decomposes the least-observable direction in the line-of-sight frame, separating the position and velocity blocks and resolving each into a component along the line of sight and a component in the plane transverse to it. The line-of-sight unit vector is $\hat{\boldsymbol{e}}_1 = \boldsymbol{\rho}/\|\boldsymbol{\rho}\|$, and the transverse plane is its orthogonal complement, so that for either block $\boldsymbol{b}$ the two reported quantities are $|\boldsymbol{b}\cdot\hat{\boldsymbol{e}}_1|$ and $\|\boldsymbol{b}_\perp\|$.



(a) Northern butterfly

(b) Distant retrograde

Fig. 11: Decomposition of the least-observable direction into position and velocity blocks, and into line-of-sight and transverse components, along both trajectories.

The five curves are components of a unit vector in the scaled coordinates of eq. (6), so their squares sum to one. Two features hold in both geometries at every phase, and both follow from the weakest direction being the scale direction. The cross-line-of-sight position component is identically zero, since the position block is parallel to $\boldsymbol{\rho}$ and therefore to $\hat{\boldsymbol{e}}_1$ by construction: the weakest direction carries no displacement in the plane the bearing angles measure directly. The velocity block is parallel to $\dot{\boldsymbol{\rho}}$ and lies almost entirely in the transverse plane, with a radial fraction of 0.03–0.06 in the DRO and up to 0.35 in the butterfly, consistent with the measured range rates. The two blocks carry equal norm because $\|\dot{\boldsymbol{\rho}}\|/v_{\text{char}} \approx \|\boldsymbol{\rho}\|$ by definition of v_{char}. Together these identify the weakest direction as

$$\boldsymbol{u}_{\min} \propto \begin{bmatrix} \boldsymbol{\rho} \\ \dot{\boldsymbol{\rho}} \end{bmatrix}, \tag{19}$$

the direction along which the entire relative state is scaled by a common factor. This is the direction that Woffinden and Geller [7] showed to be exactly unobservable under linear relative dynamics. The nonlinearity of the CR3BP renders it weakly observable rather than unobservable, but does not displace it: over the windows considered here it remains the least informed direction in both geometries, at every orbital phase.

The transverse velocity content of this direction should not be read as transverse velocity being weakly observed: a bearing rate measures $\dot{\rho}_\perp/\rho$, and scaling $\boldsymbol{\rho}$ and $\dot{\boldsymbol{\rho}}$ together leaves that ratio unchanged, so transverse velocity appears in $\boldsymbol{u}_{\min}$ only as the component needed for that cancellation. The components of $\boldsymbol{u}_{\min}$ are not separately interpretable, only the combination is.

The state metric is what makes this visible: in raw nondimensional coordinates the velocity components dominate numerically and $\|\boldsymbol{u}_{\text{pos}}\|^2 \sim 10^{-2}$, inviting the incorrect reading that the weakest direction is velocity-dominated. Under the scaling of eq. (6), $\|\boldsymbol{u}_{\text{pos}}\|^2 = (1 + r^2/v_{\text{char}}^2)^{-1}$ with $r = \|\dot{\boldsymbol{\rho}}\|/\|\boldsymbol{\rho}\|$, so the split is one half when r equals its orbit average, and the same expression accounts for the difference between geometries: the DRO's r is nearly constant, returning 0.48–0.52 across the period, while the butterfly's position and velocity fractions exchange weight twice per period between roughly 0.61 and 0.80, tracking the factor-of-two variation in separation. The weakest direction is thus a fixed physical object in the DRO and a slowly rotating one in the butterfly.

Its *strength* varies more than its orientation, set by the relative geometry of figs. 2 and 3. In the DRO, $\sigma_{\min}$ falls by two orders of magnitude near 0.39 and 0.87 of the period, coinciding with maxima of the *positive* range rate, while the negative extrema fall at plateaux rather than minima, consistent with a window that accumulates forward over five steps. In the butterfly, the four minima per period fall at the four maxima of inter-spacecraft range, consistent with angular sensitivity scaling inversely with separation. The correspondence locates the minima but not their depth: the butterfly range varies by a factor of 2.2, which under inverse-range scaling would predict a comparable variation in $\sigma_{\min}$, whereas the observed dips exceed an order of magnitude; in the DRO a 1.2 m/s range rate displaces the separation by under 0.04% over the window, far too little to produce a two-order reduction on its own. The geometry sets the phase at which the weakest direction becomes weakest; some further mechanism sets the magnitude.

Since the second-order response along this direction is nearly constant over the orbit, $\varepsilon^\star$ inherits the same phase dependence, falling to 10^{-9} in scaled units at the DRO minima. The tools disagree, however, on where that minimum falls: the window-based tools place it near 0.39 and 0.87 of the period, H-K near 0.12 and 0.62, approximately in antiphase. The instantaneous construction of H-K therefore identifies the correct direction but not the correct phase at which it is weakest, and is insufficient for scheduling or tasking decisions that depend on that phase.

7. PRACTICAL COMPARISON OF THE FOUR TOOLS

The four tools answer overlapping questions at different costs and with different configuration burdens. Table 7 summarizes them on grounds that are countable.

Table 7: Practical comparison of the four observability tools. Wall-clock times are per window, measured over a sweep of $K \approx 100$ windows spanning one orbital period of the DRO, with each tool at the setting that geometry requires.

	H-K	EOG	SP-EOG	HOG
Propagations per evaluation	1	12	13	1
Levels of nested differentiation	k	none	none	1 (G_1), 2 (G_2)
Wall clock per window	~10 s (k=3)	20–30 ms	20–40 ms	7.5 s
Free parameters	k	ϵ, W	P, α,β,κ, W	W
Minimum information, DRO	$k = 3$	$W = 4$	$W = 4$	$W = 4$
Admissible sampling range, DRO	n/a	~1 order	~1 order	n/a
Angle to reference $\boldsymbol{u}_{\min}$	$\leq 2.2°$	$< 0.01°$	$< 0.01°$	reference
Misconfiguration visible in output	yes (rank drop)	no	no	n/a

The measured costs separate the four into two classes about two orders of magnitude apart, not along the line the propagation count suggests: EOG and SP-EOG evaluate in tens of milliseconds despite propagating twelve and thirteen trajectories against one, while H-K and HOG take seconds. What costs is differentiating through the propagation, not the propagation itself, and each additional level of nested differentiation costs close to an order of magnitude more than the last. H-K at $k = 2$ and HOG's first-order part each carry one level and cost the same, 1.5–3.3 s. HOG's second-derivative tensor adds a second level, at 7.5 s; H-K at $k = 3$ adds one more, at roughly 10 s.

This contradicts the natural reading of the table's first row: at the order the DRO requires for full rank, H-K is the most expensive of the four, more so than HOG to second order. Its instantaneous construction makes it cheap in trajectories and expensive in derivatives, and the derivatives dominate. These are wall-clock figures from one implementation rather than operation counts, so a hand-derived or sparsity-exploiting formulation would move them; what is structural is the separation between sampling-based and differentiation-based tools, and the cost of each additional level of nesting.

How a misconfiguration presents itself cuts the other way: EOG and SP-EOG each require a sampling scale whose admissible range narrows to about one order of magnitude in the near-planar geometry (section 5.3), and a choice outside it returns a smooth, plausible spectrum that is wrong along the direction of interest, with nothing in the output to indicate it. H-K reports its own misconfiguration instead: too low a Lie derivative order returns a rank deficiency, stating that the tool is outside its regime rather than giving a wrong answer. HOG, needing no sampling scale, identifies where the admissible range lies for the other two.

What they answer differs as well. Figure 8 shows H-K recovers the same least-observable direction as the window-based tools, to within 2.2° in the DRO and 1.6° in the butterfly, despite being evaluated instantaneously, while section 6.3 shows it places the phases at which that direction is weakest roughly half a period away from where the window-based tools place them. H-K is therefore a reliable, inexpensive indicator of *which* direction is weak but not *when* it is weakest. This might be sufficient for screening a candidate geometry, insufficient for scheduling or tasking decisions that depend on the phase at which observability degrades.

Cost and failure behaviour together argue for a division of labour that is not the order the tools are usually introduced in. EOG and SP-EOG are the only two cheap enough to run repeatedly, SP-EOG the only one also conditioned on the filter's belief, but neither can tell it has been configured wrongly. HOG returns the most for its cost: it needs no sampling scale, sets the one the cheaper tools should use, and alone separates a direction carrying no information from one carrying information a linear sensitivity cannot reach. H-K is instantaneous and self-reporting, but in a near-planar geometry it requires $k = 3$ and becomes the most expensive evaluation in the set.

8. IMPLICATIONS FOR GAUSSIAN MIXTURE FILTERING

The motivation for observability-informed component management was given in section 1. What the preceding results add is a basis for two decisions that existing criteria make on the dynamics alone.

Which direction to split along

Splitting is worthwhile in a direction where a single Gaussian misrepresents the uncertainty and the measurements can eventually separate the resulting components. The first condition is about curvature and the second about information, and section 6.2 shows the two are not ordered together in the same way in the two geometries. In the butterfly the least-informed direction also carries the least curvature, so splitting along it buys little the first-order ranking did not already predict. In the DRO the extremes lie in different places. Splitting along the weakest direction there might return more than its first-order score suggests, and the candidate the second-order analysis identifies would not be selected by any first-order criterion. What matters is the comparison of orderings within a geometry, not of magnitudes across geometries.

When to split

The breakdown radius supplies a potential trigger as well as a direction. A mixture component with covariance P extends a distance $\sigma_P(\boldsymbol{u}) = \sqrt{\boldsymbol{u}^\top P \boldsymbol{u}}$ along $\boldsymbol{u}$, while $\varepsilon^\star(\boldsymbol{u})$ is the distance over which the measurement map remains effectively linear along the same direction. The ratio

$$\eta(\boldsymbol{u}) = \gamma \frac{\sigma_P(\boldsymbol{u})}{\varepsilon^\star(\boldsymbol{u})} \tag{20}$$

states whether the component's own uncertainty extends past the region in which its Gaussian representation is valid, and $\eta > 1$ is a candidate splitting condition. The factor γ sets the confidence level at which that comparison is made: $\gamma = 1, 2, 3$ compares the region's 1σ, 2σ, 3σ extent against $\varepsilon^\star$, so a user can choose how conservative the trigger is.

Which components to protect from pruning

Standard pruning removes components whose weights fall below a threshold, assuming a low weight indicates a low-value hypothesis. Components separated along a direction with a small singular value produce nearly identical predicted measurements, so their likelihoods carry little information about which is more accurate and the assumption fails [21]. SP-EOG evaluates that separation at the spread the filter actually maintains, which makes it usable online as a protection criterion rather than only as an offline diagnostic.

Cost and extension to a filter

The two criteria are not equally affordable, and section 7 says which is which. SP-EOG evaluates in a few tens of milliseconds per window, the same order as a single square-root UKF prediction step, so the pruning-protection criterion can be applied online as stated. HOG is more than two orders of magnitude more expensive, since $\varepsilon^\star$ requires the second-derivative tensor, so the trigger of eq. (20) cannot be evaluated per component at filter cadence in this form.

That cost asymmetry implies a split treatment: $\sigma_P(\boldsymbol{u})$ is per-component and cheap, so tabulating $\varepsilon^\star$ offline against orbital phase would make the trigger affordable, at the cost of assuming a component that has drifted from the reference still sees the reference curvature, an assumption the present results only partly support, since the second-order response along the weakest direction is flat across the DRO orbit (section 6.2) but not the butterfly. In a filter both quantities would instead be evaluated against each component's own belief rather than a reference trajectory, and the two tools are linked operationally through α: the sigma-point spread that sets SP-EOG's evaluation is the same spread at which $\varepsilon^{\star}$ marks the onset of curvature, so α can be chosen to keep the regression inside or deliberately outside the linear regime. Integrating these criteria into an adaptive mixture filter and evaluating them against observability-agnostic baselines is left to future work.

9. CONCLUSIONS

Four observability tools were applied to cislunar angles-only relative navigation in two CR3BP orbit families, two of them established and two proposed here: a sigma-point Gramian conditioned on an estimator's assumed covariance, and a higher-order Gramian retaining the second-order term of the composed measurement map.

Three conditions have to be met before the four tools can be compared, and each is more consequential in the near-planar geometry. The spectrum must be taken from the singular value decomposition of the whitened sensitivity matrix and not from the assembled Gramian, which squares a dynamic range already at the limit of double precision. Each tool requires a minimum amount of measurement information before the weakest direction is determined at all: at the minimum Lie derivative order and the minimum window length, both of which give as many rows as states, the DRO is exactly rank-deficient at every epoch while the butterfly is not. The sampling scale of EOG and SP-EOG must lie in an admissible range that narrows to about one order of magnitude in the DRO, and the breakdown radius from the second-order analysis predicts where that range lies. A configuration outside any of these bounds returns a plausible spectrum that is wrong along the direction of interest, with nothing in the output to indicate it.

The least-observable direction was identified physically in both geometries. Its position block lies along the line of sight, its cross-line-of-sight position component is identically zero, and its velocity block is distributed in the same proportion as the relative velocity itself, which identifies it as the direction along which the entire relative state is scaled by a common factor: the direction shown to be exactly unobservable under linear relative dynamics. The nonlinearity of the CR3BP renders it weakly observable rather than unobservable but does not displace it, at any orbital phase in either geometry.

Cost is set by depth of differentiation rather than by number of propagations, which reverses the apparent ordering of the tools: the sampling-based pair evaluate in tens of milliseconds per window, the differentiation-based pair in seconds, and H-K becomes the most expensive of the four at the Lie derivative order the near-planar geometry needs for full rank. Only the sampling-based tools are viable at filter cadence, and only the second-order tool reports where their admissible sampling scale lies.

The first- and second-order analyses order the directions differently, and by different amounts in the two geometries: in the DRO the weakest direction ranks only third in curvature, while in the butterfly the two orderings agree on which direction is worst but disagree in degree, a property of that geometry rather than a general relation. The rank deficiency of planar angles-only geometries reflects the first-order description at minimum information, not an absence of information in that direction. Because component management in a Gaussian mixture filter consists of directional decisions, and first-order observability cannot distinguish an uninformative direction from a nonlinearly informative one, first-order metrics alone may not suffice to make them.

These results are limited to two orbit families, lead-follower configurations at a single separation, a single cadence of $T_{\text{orbit}}/1000$, and windows of three to five steps; because the window tools accumulate over steps rather than elapsed time, cadence and window length are not separable here, and the phase dependence of section 6.3 describes a window spanning half a percent of an orbit. The second-order response was evaluated only along the singular vectors of the first-order map, establishing that the two orderings differ but not locating the directions of greatest curvature, and a sensor measurement was assumed available at every epoch. No closed-loop filter results are presented; the proposed component management criteria are a motivating rather than a demonstrated outcome, and integrating them into an adaptive mixture filter against observability-agnostic baselines is left to future work.

REFERENCES

[1] Carolin Frueh, Kathleen Howell, Kyle J. DeMars, and Surabhi Bhadauria. Cislunar space situational awareness. In *31st AAS/AIAA Space Flight Mechanics Meeting*, 2021.

[2] Sahil Bhatia, Khushi Shah, and Jay Kamdar. Review of strategies for cis-lunar space traffic management. In *73rd International Astronautical Congress (IAC)*, September 2022.

[3] Steve Creech, John Guidi, and Dan Elburn. Artemis: An overview of NASA's activities to return humans to the moon. In *2022 IEEE Aerospace Conference (AERO)*, pages 1–7. IEEE, March 2022.

[4] David J. Israel, Kendall D. Mauldin, Christopher J. Roberts, Jason W. Mitchell, Antti A. Pulkkinen, La Vida D. Cooper, Michael A. Johnson, Steven D. Christe, and Cheryl J. Gramling. LunaNet: a flexible and extensible lunar exploration communications and navigation infrastructure. In *2020 IEEE Aerospace Conference*, pages 1–14. IEEE, March 2020.

[5] Jesse A. Greaves and Daniel J. Scheeres. Autonomous optical-only spacecraft-to-spacecraft absolute tracking and maneuver classification in Cislunar space. *Journal of Guidance, Control, and Dynamics*, 46(11):2092–2109, 2023.

[6] Jesse Greaves and Daniel Scheeres. Spacecraft to spacecraft absolute tracking for autonomous navigation of a distributed space system from relative sensors. *J Astronaut Sci*, 71(5):46, October 2024.

[7] David C. Woffinden and David K. Geller. Observability criteria for angles-only navigation. *IEEE Trans. Aerosp. Electron. Syst.*, 45(3):1194–1208, July 2009.

[8] Evan Kaufman, T. Alan Lovell, and Taeyoung Lee. Nonlinear observability for relative orbit determination with angles-only measurements. *J of Astronaut Sci*, 63(1):60–80, March 2016.

[9] Eric A. Butcher, Jingwei Wang, and T. Alan Lovell. On Kalman filtering and observability in nonlinear sequential relative orbit estimation. *Journal of Guidance, Control, and Dynamics*, 40(9):2167–2182, 2017.

[10] Jackson Kulik and Dmitry Savransky. State transition tensors for passive angles-only relative orbit determination. In *33rd AAS/AIAA Space Flight Mechanics Meeting*, 2023.

[11] Baichun Gong, Sha Wang, Shuang Li, and Xianqiang Li. Review of space relative navigation based on angles-only measurements. *Astrodyn*, 7(2):131–152, June 2023.

[12] Keith A. LeGrand, Kyle J. DeMars, and Henry J. Pernicka. Bearings-only initial relative orbit determination. *Journal of Guidance, Control, and Dynamics*, 38(9):1699–1713, September 2015.

[13] Keith A. LeGrand, Aneesh V. Khilnani, and John L. Iannamorelli. Bayesian angles-only Cislunar space object tracking. In *33rd AAS/AIAA Space Flight Mechanics Meeting*, 2023.

[14] Reinhard Hermann and Arthur J. Krener. Nonlinear controllability and observability. *IEEE Trans. Automat. Contr.*, 22(5):728–740, October 1977.

[15] Jo Ryeong Yim. *Autonomous spacecraft orbit navigation*. PhD thesis, Texas A&M University, 2002.

[16] Arthur J. Krener and Kayo Ide. Measures of unobservability. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 6401–6406, December 2009.

[17] Erin E. Fowler and Derek A. Paley. Observability metrics for space-based Cislunar domain awareness. *The Journal of the Astronautical Sciences*, 70:10, March 2023.

[18] Mathew Givens, Ben Tatman, Anthony Zara, and Galen Savidge. Nonlinear observability of Cislunar relative navigation using automatic differentiation. In *2024 AAS/AIAA Astrodynamics Specialist Conference*, Denver, Colorado, August 2024.

[19] Nathan D. Powel and Kristi A. Morgansen. Empirical observability Gramian rank condition for weak observability of nonlinear systems with control. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 6342–6348, December 2015.

[20] Atiye Alaeddini and Kristi A. Morgansen. Trajectory design for a nonlinear system to insure observability. In *2014 European Control Conference (ECC)*, pages 2520–2525, June 2014.

[21] Andrea López, Jay W McMahon, and Hanspeter Schaub. Square root filters for cislunar angles-only relative orbit estimation. *The Journal of the Astronautical Sciences*, 73(4):67, 2026.

[22] Kyle J. DeMars, Robert H. Bishop, and Moriba K. Jah. Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. *Journal of Guidance, Control, and Dynamics*, 36(4):1047–1057, July 2013.

[23] Chi-Tsong Chen and Chi-Tsong Chen. *Linear system theory and design*, volume 301. Holt, Rinehart and Winston New York, 1984.