Doppler-Domain Micro-Motion Tomography

Reproduction, Critical Analysis, and Benchmarking of an Extended CUDA Implementation

Jashua Luna

June 13, 2026

Abstract. This report studies the single-SLC Doppler-domain micro-motion tomography workflow proposed by Biondi (2022) and Biondi–Malanga (2022) from three directions at once. First, it reconstructs the published method as executable software, keeping an explicit taxonomy of literal method components, paper-compatible engineering, corrective extensions, and performance-only runtime choices. Second, it subjects the method itself to a critical signal-level analysis and shows several structural properties that the source papers do not state: the published steering law is dimensionally inconsistent as written; real-valued motion observables produce tomograms that are exactly mirror-symmetric in depth unless an analytic-signal step is added; the investigation frequencies quoted by the papers exceed the pair-axis Nyquist rate by orders of magnitude; the effective depth aperture is carried by lag families rather than by individual sub-aperture pairs; and the band-passed micro-motion series is nearly orthogonal to the steering manifold, which a MUSIC diagnostic makes directly visible. Third, it validates and benchmarks the extended implementation: a raw-signal synthetic generator with vibrating point scatterers exercises the full sub-aperture/tracking chain against ground truth; shuffle and sign-flip null tests bound what tomographic structure can be trusted; and a resumable 48-hour benchmark suite measures tracking fidelity, estimator behavior (damped pseudoinverse, Capon, MUSIC), ablations, runtime, and memory; real TerraSAR-X HS SSC products are used to exercise ingestion and to measure runtime and memory at scale, while controlled accuracy claims rest on synthetic ground truth. The quantitative results are populated directly from one completed run of that suite. The front end tracks injected micro-motion with range correlation near unity in its valid regime, but the tomographic localization built on top of it does not exceed its own shuffle-pairs and sign-flip null floors under synthetic ground truth — the method's depth contrast is carried by the family-mean component, exactly as the critical analysis predicts.

Code availability. The CUDA/C++ and Python source code that accompanies this report is publicly available at `https://github.com/not-JASH/sar-doppler-tomography` under the MIT License.

Foreword and Disclaimer

The honest motivation behind this project was probably as much about AI as it was about volcanoes, pyramids, and the possibility of forgotten structures hidden within or beneath them. Over the past few years, I had read enthusiastic claims that AI tools were now capable of acting as serious research accelerants—compressing the path from an unfamiliar paper to a working, instrumented implementation in days or weeks rather than months—and an equally vocal counter-current that treated such claims as marketing. I wanted to test that question for myself on a task I could not have realistically completed alone in the same time frame: taking a published, computationally heavy, and only partially specified signal-processing method and using AI assistance, together with conventional software engineering and GPU parallelization, to turn it into an executable, inspectable, and benchmarked pipeline.

Dr. Biondi's Doppler tomography workflow was a near-ideal subject for that test. The method is technically substantial; its surrounding claims—the Vesuvius interior, the Great Pyramid case study, and the wider Khafre Project discussion—were striking enough that I had repeatedly encountered complaints in podcasts, interviews, and online discussions that the work remained largely unverified because so few people had the time and background to reproduce it. The reported per-tomogram processing cost—weeks or months on workstation hardware—also made parallelization an obvious second axis of investigation. With a background in C++ optimization and parallel computing, I could plausibly handle the implementation side; the open question was how much of the rest—reading the method, hardening the pipeline, designing experiments, and writing this report—an AI-assisted workflow could realistically carry.

This report is the result of that experiment. It should not be read as an endorsement of any archaeological or subsurface-structure claim associated with the Great Pyramid, Vesuvius, or any other site. Its purpose is narrower and more technical: to reconstruct the published method as software, identify what its mathematics and signal model do and do not support, implement well-defined corrections where necessary, and benchmark the resulting pipeline under controlled tests. The strongest claims made here concern reproducibility, implementation, signal-level analysis, synthetic validation, null tests, and performance. The report does not establish the physical validity of any claimed hidden structure.

AI assistance was the through line connecting the implementation, the experiments, and this manuscript, and it is part of what is being evaluated rather than merely a writing aid. Appendix D records which tools were used at which stages. Research framing, selection of claims, interpretation of results, and final acceptance of code and manuscript changes remained under my control. AI accelerated the work; it did not replace independent judgment, validation, or responsibility for the contents of this report.

Contents

1 Introduction

1.1 Context and Motivation

Synthetic Aperture Radar (SAR) single-look complex (SLC) products preserve phase and coherent speckle statistics and therefore contain information not only about scene reflectivity but also about motion-induced perturbations. The workflow studied here treats those perturbations as the signal of interest: a single focused SLC is decomposed into Doppler sub-apertures, micro-motion is estimated by comparing the resulting looks, and a depth-dependent response is reconstructed through a steering-matrix focusing operator [2, 3].

This is not conventional multi-pass TomoSAR [18]. The diversity is constructed *within one* aperture by selecting Doppler sub-bands and comparing lower-azimuth-resolution looks rather than by stacking repeat-pass acquisitions. The published applications — imaging the interior of Vesuvius to a claimed depth of 3 km and resolving internal structure of the Great Pyramid — are scientifically contested, and this report takes no position that the published reconstructions are physically valid. What it does instead is make the method fully executable, identify exactly which of its mathematical claims hold, fail, or are under-determined at the signal level, and measure what the resulting pipeline can demonstrably recover under controlled ground truth.

1.2 Scope and Questions

Relative to an earlier revision of this report, which was framed purely as a reproduction-and-systems study, the scope has widened in two ways. The implementation now *corrects* identifiable defects of the literal method where a correction is well defined (complex two-component observations, analytic-signal direction discrimination, aliasing guards), and the evidence base now includes end-to-end synthetic ground truth, falsification-oriented null tests, and a comprehensive benchmark suite. The manuscript is organized around four questions:

(i) What executable processing chain is actually implied by the reference papers, and which of its signal-level properties can be established or refuted analytically?

(ii) Which implementation choices are literal reproductions, which are paper-compatible engineering, which are *corrective* extensions, and which are purely performance-oriented?

(iii) What can the full chain — sub-aperture decomposition, pixel tracking, observation construction, and tomographic focusing — demonstrably recover when the ground truth is known, and what do null tests say about the rest?

(iv) How do the three implemented estimators and the main runtime parameters shape accuracy, runtime, and memory on workstation hardware?

1.3 Contributions

- A staged paper-to-code reconstruction of the Doppler tomography workflow with the main signal-processing objects kept explicit, executed by a fixed CUDA/OpenMP runtime now organized into three translation units (kernels, estimators, pipeline orchestration).
- A critical analysis of the published method (Section 3) establishing five structural results: the dimensional inconsistency of the literal steering law; an exact $\pm z$ mirror ambiguity of real-valued observations (Proposition 1); the aliasing of the published investigation frequencies relative to the pair-axis Nyquist rate; the concentration of depth aperture in lag families; and the near-orthogonality of band-passed motion series to the steering manifold.
- Corrective extensions implemented rather than merely documented: a complex two-component observation model per the papers' own Eq. (8); an analytic-signal stage that provably removes

Table 1: Claims, evidence, and strength of support in this report.

Claim	Evidence	Support	Main limit
An executable computational chain can be recovered from the papers	Staged pipeline reconstruction with per-stage equations (Section 5)	Strong	The papers are not complete software specifications
The literal method has identifiable signal-level defects	Closed-form analysis in Section 3; mirror-symmetry proposition; Nyquist arithmetic; MUSIC orthogonality diagnostic	Strong	Analysis addresses the method as written, not every conceivable reading
The corrective extensions repair those defects without changing the intended inverse problem	Implementation + ablation stage of the benchmark suite (Section 10.4)	Strong	Ablations quantify the effect on this implementation only
The front end tracks known micro-motion faithfully in its valid regime	End-to-end synthetic ground truth: range correlation 0.97–0.99 for $f_0 T_{\mathrm{sub}} \lesssim 0.4$, collapsing past it (Section 10.1)	Strong	Valid regime bounded by $f_0 T_{\mathrm{sub}} \lesssim 1$ and SNR $\gtrsim 30\,\mathrm{dB}$
The tomographic localization does *not* exceed null-test levels	Shuffle-pairs and negate-K_z controls on synthetic ground truth: reference/shuffle contrast ratio $\leq 1.04\times$, K_z sign flip leaves peak-z unchanged (Section 10.3)	Strong	A negative result for this method on the synthetic reference scene, not a general impossibility proof
The implementation is executable and resource-bounded on workstation hardware	Benchmark perf-scaling and repeatability stages, plus real-product runtime anchors	Strong	One implementation, one workstation
Physical validity of reconstructed depth structure	Not established; depth scale $\lambda_{\mathrm{vib}} = v/f$ is a free parameter (Section 3.5)	Not claimed	Outside the reach of single-scene experiments

the mirror ambiguity; a Nyquist guard with explicit aliasing bookkeeping; and a literal sub-aperture stepping mode for strict paper-profile runs.

- Three tomographic estimators behind one interface — weighted damped pseudoinverse, Capon (MVDR) [4], and MUSIC [19] — with batched CUDA execution.
- A validation methodology with teeth: a raw-signal synthetic generator that injects known harmonic micro-motion at the SLC level, a scorer that measures displacement-tracking fidelity and tomogram behavior against truth, and two null tests (pair shuffling, K_z sign flip) that bound what reconstructed structure can be attributed to the data rather than the operator.
- A single-command, resumable, budget-aware benchmark suite ($\sim$48 h) whose machine-readable output populates the quantitative results of this report.

1.4 Claims and Evidence

Table 1 states the report's claims against their evidence. The hierarchy is deliberate: analytical results and ground-truth synthetic measurements are the strongest evidence; real-data runs demonstrate ingestion and runtime/memory behavior at scale, not validated subsurface physics.

Table 2: Operational classification of implementation choices.

Class	Representative components
Literal	COSAR decoding, TerraSAR-X SSC metadata interpretation, canonical range/azimuth handling, orbit and geometry extraction, Doppler sub-aperture construction, NCC pixel tracking, explicit construction of A and per-pixel K_z, pseudoinverse focusing [7, 2]
Corrective extension	Complex two-component observation model (papers' Eq. (8) $\{a, b\}$); analytic-signal stage removing the $\pm z$ mirror ambiguity; pair-axis Nyquist guard and aliasing bookkeeping; dimensionally consistent steering law $\exp(\mathrm{j}K_z z)$ replacing the literal $\exp(\mathrm{j}2\pi K_z t z)$ (Section 3.1)
Paper-compatible engineering	Multi-lag pair-family schedule ($N_c > 1$); per-pixel K_z; adaptive vertical sampling; ground-referenced line reprojection; confidence weighting; pairwise centering; Tikhonov damping; Capon and MUSIC estimators; null-test controls
Performance-only	FFT-friendly size smoothing, OpenMP worksharing, CUDA batched BLAS/solver kernels, pinned-memory staging, memory-budget-based batching [14, 12, 15, 16, 13]

2 Reproduction Target and Claim Taxonomy

2.1 Reference Papers

The reference method is the Doppler-domain micro-motion tomography workflow of [2, 3]. Both papers share one computational core: a focused SLC is decomposed into Doppler sub-apertures; inter-sub-aperture motion is estimated by pixel tracking / sub-pixel coregistration; the resulting observations are organized along a tomographic line; and a depth-dependent profile is recovered by focusing those observations with a steering-matrix model. The volcano paper is the clearer reproduction target: it writes the tomographic model as $Y = A(K_z, z)\, h(z)$ with solution $h(z) = A^{\dagger}Y$ and gives an 11-block computational scheme from SLC input to geocoded tomogram [2]. The pyramid paper presents the same chain more compactly with application-specific reconstructions [3].

Algorithm 1 Reference workflow reproduced in this report, adapted from the paper-level schematics in [2, 3]

1: Load focused SLC and required scene metadata
2: Partition azimuth spectrum into Doppler sub-apertures
3: Reconstruct lower-resolution sub-aperture looks
4: Compare selected looks to estimate per-pixel motion observations
5: Assemble observation vector Y for each pixel on the tomographic line
6: Evaluate per-pixel geometry term K_z and steering matrix $A(K_z, z)$
7: Solve the local inverse problem to recover the depth-dependent response

2.2 Four-Way Claim Taxonomy

An earlier revision of this report used three categories (literal, paper-compatible engineering, performance-only). The critical analysis in Section 3 forces a fourth: *corrective extension*, meaning a change that departs from the literal method because the literal method is demonstrably defective at the signal level, while preserving the intended inverse problem. Table 2 applies the full taxonomy.

2.3 Boundaries of Literalness

The reference papers provide computational schematics, not software specifications. They leave product parsing, metadata normalization, coordinate conventions, numerical safeguards, batching, and hardware-specific execution implicit. Those omissions are the boundary between a research description and an implementation specification; they are treated here as engineering additions rather than literal claims. Where the papers' written mathematics is internally inconsistent, Section 3 states the inconsistency and the adopted resolution explicitly instead of silently repairing it.

3 Critical Analysis of the Published Method

This section is the analytical core of the report. Each subsection states a property of the method as published, gives the argument, and identifies the implementation consequence. All five properties were discovered or confirmed in the course of building and validating the implementation; none is stated in the source papers.

3.1 The Steering Law Is Dimensionally Inconsistent as Written

Both papers write the steering-matrix entries in the literal form $A[i, f] = \exp(\mathrm{j}\, 2\pi\, K_{z,i}\, t\, z_f)$ (volcano Eq. (10), pyramid Eq. (22)). Since K_z carries units of rad/m, the additional factor 2π double-counts angular units, and the stray time variable t has no referent: K_z is indexed per pair, not per time-sample within a pair. Treated literally the phase is not dimensionless. The implementation adopts the textbook TomoSAR convention [18]

$$A_{p,n} = \exp\bigl(\mathrm{j}\, K_{z,p}\, z_n\bigr), \tag{1}$$

which is dimensionally consistent and reproduces the papers' own resolution prediction $\delta_z \approx \lambda_{\mathrm{vib}} R/(2A_{\mathrm{ap}})$. One alternative reading of the literal equation — that the phase is meant to evolve with slow time t at a rate proportional to $K_z z$, i.e. a time-frequency dictionary in which depth maps to a temporal frequency — would couple more naturally to the oscillatory observations discussed in Section 3.6, but it cannot be made dimensionally well defined without parameters the papers do not supply. The adopted form (1) is therefore a corrective reading, recorded as such in the taxonomy.

3.2 Real-Valued Observations Cannot Discriminate Above from Below

The papers' processing chain estimates displacement shifts, and the most direct executable reading (the default of the earlier revision of this implementation) stores a real-valued scalar per pair. That choice has an exact and previously unstated consequence.

Proposition 1 (Mirror ambiguity)**.** *Let $y \in \mathbb{R}^P$ be a real-valued observation vector, let $A(z)_p = \exp(\mathrm{j}K_{z,p}z)$, and let $\hat{h}(z)$ be any reconstruction of the form $\hat{h} = G\,y$ where G is built from A evaluated on a depth grid symmetric about zero (matched filter, pseudoinverse, or damped pseudoinverse). Then $\hat{h}(-z) = \overline{\hat{h}(z)}$, so the tomogram magnitude satisfies $|\hat{h}(-z)| = |\hat{h}(z)|$: above-surface and below-surface responses are indistinguishable.*

Proof. $A(-z) = \overline{A(z)}$ entrywise. For the matched filter, $\hat{h}(-z) = A(-z)^{\mathrm{H}}y = A(z)^{\mathrm{T}}y = \overline{A(z)^{\mathrm{H}}y}$ since y is real. The Gram matrices $A^{\mathrm{H}}A + \mu I$ on a symmetric grid satisfy the corresponding conjugation symmetry, so the damped pseudoinverse inherits the same property. □

On an asymmetric grid the symmetry is hidden but the information deficit remains: a real vector has a conjugate-symmetric spectrum along the pair axis, and the sign of z is encoded precisely in the broken symmetry that a real observation cannot supply. The earlier pipeline's vibration band-pass did *not* repair this — it retained both $\pm f$ bands. The corrective extension is a per-lag-family one-sided (analytic-signal) spectrum step (Section 5.4): zeroing negative pair-axis frequencies makes the inverted series genuinely complex and breaks the mirror exactly. The implementation now records a `direction_discrimination_available` flag in every product and refuses to stay silent when an inversion would be direction-blind.

3.3 The Published Investigation Frequencies Are Aliased

The vibration time series available to the inversion is sampled at the pair-center times of each lag family. For a staring-spotlight dwell of duration T and N_D sub-apertures spanning the guard band, consecutive pair centers are separated by $\Delta t \approx T\, B_{DL}/(B_{c_D}\,(N_D - 1))$; with the volcano paper's parameters ($T \approx 14\,\mathrm{s}$, $N_D \sim 64$) the sampling rate is of order 5 Hz to 10 Hz and the Nyquist limit a few hertz. The volcano paper sets the investigation frequency to 200 Hz and the pyramid paper to 12.5 kHz; both exceed the pair-axis Nyquist rate by two to four orders of magnitude. Such components enter the observations only through aliasing, and the depth-axis scale derived from them (Section 3.5) then relies on the unaliased frequency being known a priori. The implementation computes each family's Nyquist rate, counts aliased families, warns prominently, and exports `investigation_freq_aliased` with every run so that aliased configurations are at least explicit rather than silent.

3.4 Lag Families Carry the Depth Aperture

Within one lag family (fixed master/slave index separation ℓ), the effective baseline $B_\perp$ is nearly constant: the sub-aperture center times advance together and the orbit is locally linear, so $K_{z,p}$ varies only weakly with the master index and strongly with the lag ℓ. Consequently the steering vector $A(z)$ is approximately piecewise constant over families, and the usable spread of K_z — hence the vertical resolution $\delta_z \approx 2\pi/\Delta K_z$ — is set by the *number and span of lag families* N_c, not by the number of pairs within a family. The papers sketch a single master/slave chain ($N_c = 1$), which in this geometry yields an almost rank-one steering manifold and essentially no depth discrimination. The multi-lag schedule ($N_c > 1$) used by this implementation is therefore not a convenience but a structural requirement for the inversion to have any vertical aperture at all; the ablation stage of the benchmark suite quantifies the dependence (Section 10.4).

3.5 The Depth Scale Is a Free Parameter

The geometry term is $K_z = 4\pi B_\perp/(\lambda_{\mathrm{vib}}\, r \sin\theta)$ with $\lambda_{\mathrm{vib}} = v/f_{\mathrm{inv}}$. Both v (assumed propagation velocity) and f_{inv} are user inputs; the data constrain at most their ratio through the self-consistency of the focusing, and in practice not even that. Two consequences follow. First, the absolute depth axis of any reconstruction is a rescaling of an assumed material model, not a measurement; the same observation vector maps to 3 km of volcano or 30 m of masonry by choice of v/f. Second, synthetic validation must *condition* the steering manifold deliberately: with physically plausible v/f for the synthetic geometry the total steering phase across the depth window can be $\ll 1$ radian, leaving all steering vectors nearly parallel. The benchmark suite's synthetic stages therefore select λ_{vib} so that K_z-phase winds by several radians over the depth window, and the degenerate setting is kept as an ablation. A flat MUSIC spectrum is the sharpest available diagnostic of the degenerate case (next subsection).

3.6 Band-Passed Motion Series Are Nearly Orthogonal to the Steering Manifold

After the family-wise band-pass at f_{inv}, the observation series oscillates along the pair axis *within* each family, while by Section 3.4 the steering vectors are nearly constant within a family. The inner products $A(z)^{\mathrm{H}}y$ are then dominated by the family-mean (DC) component of the band-passed series — for the analytic signal a small but nonzero residual — and the oscillatory energy, which is the part the band-pass was designed to isolate, lives almost entirely in the steering-orthogonal subspace. A MUSIC scan makes this visible directly: with the signal subspace estimated from the data covariance, the projection of every steering vector onto that subspace is near zero and the pseudo-spectrum is flat. The pseudoinverse tomogram's apparent contrast is therefore carried by the family-mean component, not by the oscillation itself. This is the deepest structural tension in the method as published: the steering model of Eq. (1) cannot represent the temporal oscillation it is nominally focusing, and any reading that could (a time-frequency dictionary) is not well defined by the papers (Section 3.1). The benchmark suite reports MUSIC alongside the other estimators precisely because its failure mode is informative.

4 Inputs and Scene Preparation

4.1 Input Assumptions

The runtime assumes TerraSAR-X / TanDEM-X Level-1b SSC products in spotlight geometry at complex fidelity. SSC products are delivered in the DLR-defined COSAR format and are complex-valued in slant-range geometry [6, 7]; complex samples are stored as 16-bit I / 16-bit Q signed integers with big-endian byte ordering, arranged range-line by range-line [7]. Downstream processing requires (i) acquisition and processing metadata (PRF, Doppler bandwidth, timing), (ii) orbit state vectors sufficient to interpolate platform position and velocity over the acquisition, and (iii) a geolocation grid mapping range–azimuth gridpoints to latitude/longitude/height with incidence angle [7].

4.2 Prepared-Scene Interface and Current Inventory

Raw products are converted once into a prepared-view layout (`slc.raw` as interleaved float32 complex plus a normalized `metadata.json`) by `prepare_slc.py`; the tomography executable and all experiment tooling consume only this canonical form. The real-data runs in this report use two TerraSAR-X high-resolution spotlight (HS) SSC products prepared from delivered SIP packages; they are treated only as ingestion and runtime test cases, so no scene is identified beyond its product type.

4.3 Axis Canonicalization and Line Definition

The implementation enforces a canonical internal ordering row $\equiv r$ (slant range), col $\equiv a$ (azimuth); all exported products record the adopted convention. A tomographic line is specified as pixel endpoints, a pixel polyline, or a ground-referenced polyline reprojected per scene. The reference papers compute tomograms along contiguous pixel lines, so the line-first interface is method-consistent while being more general [2]. The benchmark suite additionally defines automatic azimuth- and range-oriented benchmark lines through the scene center of every prepared view (capped at 4096 pixels) so that real-product runs need no manual planning.

5 Implemented Pipeline

5.1 Stage 1: Doppler Sub-Aperture Construction

Let $s(r, a) \in \mathbb{C}$ denote the focused SSC image and $S(f_r, f_a) = \mathcal{F}_{r,a}\{s(r, a)\}$ its 2-D DFT. Doppler sub-apertures are synthesized by azimuth-frequency windows $W_d(f_a)$ centered at Doppler centers

$f_{a,d}$, optionally with range trimming $W_r(f_r)$:

$$S_d(f_r, f_a) = S(f_r, f_a)\, W_r(f_r)\, W_d(f_a), \tag{2}$$

$$s_d(r, a) = \mathcal{F}^{-1}_{f_r, f_a}\{S_d(f_r, f_a)\}. \tag{3}$$

With total Doppler band B_{c_D} and guard fraction $B_{DL} = \beta B_{c_D}$ (default $\beta = 0.5$), each sub-aperture retains $B_{c_D} - B_{DL}$ and the N_D centers step across the guard band. Two stepping conventions are implemented: `grid` places first/last centers exactly on the guard-band edges (step $B_{DL}/(N_D - 1)$, the default), and `literal` follows the papers' Appendix-A stepping B_{DL}/N_D for strict paper-profile runs. Doppler centers are mapped to slow times through the orbit's relative-Doppler history, and master/slave pairs are scheduled as lag families $(m, m+\ell)$, $\ell = 1, \ldots, N_c$ (Section 3.4).

5.2 Stage 2: Sub-Aperture Comparison and Motion Estimation

For master/slave looks s_i, s_j, motion is estimated by patch-wise normalized cross-correlation (NCC) around each line pixel, with FFT acceleration [11, 10] and parabolic sub-pixel peak refinement [17, 1]. The output per line pixel and pair p is the shift estimate $(\Delta r_p, \Delta a_p)$ in pixels and the NCC peak value used downstream as a confidence score.

5.3 Stage 3: Observation Models

The tomographic datum is a complex observation vector $Y = [y_1, \ldots, y_P]^\mathsf{T}$ per line pixel. Three observation models are implemented behind one switch; their classification follows Table 2.

Complex two-component model (default; corrective). The papers' oscillator model (volcano Eq. (8)) writes the per-pair motion sample as the *pair* $\{a, b\}$ of coregistrator shifts. The default model implements exactly that:

$$y_p \;=\; \Delta r_p\, \delta_r \;+\; \mathrm{j}\, \Delta a_p\, \delta_a \;\in\; \mathbb{C}, \tag{4}$$

with δ_r, δ_a the slant-range and azimuth pixel spacings (velocity units divide by the pair time separation Δt_p). Both components support pairwise common-mode centering. Relative to the scalar default of the earlier revision, Eq. (4) restores the second degree of freedom the papers' own motion model specifies and makes the observation genuinely complex.

Scalar model (legacy literal reading). $y_p = \Delta a_p\, \delta_a$ stored with zero imaginary part. Retained for ablations; subject to Proposition 1 unless the analytic-signal stage is active.

Direct phasor (paper-compatible extension). The normalized local complex master/slave correlation after applying the Stage-2 shift to the slave patch,

$$y_p^{\text{phasor}} = \gamma_p = \frac{\sum_{q \in \Omega_p} s_{m,p}(q)\, \overline{s_{s,p}^{\text{reg}}(q)}}{\sqrt{\sum_q |s_{m,p}(q)|^2}\, \sqrt{\sum_q |s_{s,p}^{\text{reg}}(q)|^2}}, \tag{5}$$

a coherence-like complex sample that sits naturally inside the steering inversion.

Bridge to the steering model. The papers motivate Y from motion estimates but do not derive a linear complex model. The implementation makes the approximation explicit: linearizing the NCC estimator about its peak and assuming narrowband harmonic motion with depth-dependent phase progression $u_p(z_n) = c_n \exp(\mathrm{j}(K_{z,p} z_n + \phi_0))$ yields

$$y_p \approx \sum_{n=1}^{N_z} h_n \exp\!\big(\mathrm{j} K_{z,p} z_n\big), \qquad Y \approx Ah, \tag{6}$$

an approximate focusing model rather than a first-principles derivation. The tensions identified in Sections 3.4 and 3.6 live precisely in this bridge.

5.4 Stage 4: Spectral Conditioning — Band-Pass, Analytic Signal, Nyquist Guard

Per lag family, the observation series may be (i) band-pass filtered around the investigation frequency (paper-compatible conditioning) and (ii) converted to its analytic signal by zeroing negative pair-axis frequencies (corrective; removes the Proposition 1 ambiguity). Both operate in one FFT round trip per family. The analytic step defaults to `auto`: it is applied whenever the observation model is motion-derived (complex or scalar) and skipped for the already-phasor model. The investigation frequency is either supplied or auto-detected as the median spectral peak over spectrally processable families; in either case the per-family Nyquist rates are computed, aliased families counted, and the run annotated (Section 3.3). Every product records which series (raw or filtered) was inverted, whether the analytic step ran, and whether direction discrimination is available.

5.5 Stage 5: Geometry and Per-Pixel K_z

Both papers define

$$K_z = \frac{4\pi B_\perp}{\lambda_{\text{vib}}\, r\, \sin\theta}, \qquad \lambda_{\text{vib}} = \frac{v}{f_{\text{inv}}}, \tag{7}$$

matching the standard TomoSAR vertical-wavenumber form [18, 5] with the radar wavelength replaced by an assumed vibration wavelength. In this single-SLC setting $B_\perp$ is an *effective-baseline surrogate*: master and slave sub-apertures correspond to different center times within one aperture; orbit interpolation maps those times to virtual positions M_p, S_p; and the signed perpendicular component of $\mathbf{B}_p = S_p - M_p$ with respect to the local line of sight $\hat{\ell}(P;p)$ and along-track direction is

$$B_\perp(P;p) = \mathbf{B}_p \cdot \hat{b}_\perp(P;p), \qquad \hat{b}_\perp = \frac{\hat{\ell} \times (\hat{t} \times \hat{\ell})}{\|\hat{\ell} \times (\hat{t} \times \hat{\ell})\|}. \tag{8}$$

K_z is evaluated per pixel and per pair using slant range and incidence interpolated from the geolocation grid. A null-test mode negates all K_z values after construction (Section 6.2).

5.6 Stage 6: Tomographic Estimators

Three estimators are implemented behind `-estimator`; all consume the same per-pixel (y, K_z) and run batched on the GPU.

Weighted damped pseudoinverse (literal + engineering). The reference inversion $h = A^\dagger Y$ generalized to confidence-weighted Tikhonov form:

$$\hat{h} = (A^{\mathrm{H}} W A + \mu I)^{-1} A^{\mathrm{H}} W y \quad \text{or} \quad \hat{h} = A^{\mathrm{H}} W^{1/2} (W^{1/2} A A^{\mathrm{H}} W^{1/2} + \mu I)^{-1} W^{1/2} y, \tag{9}$$

choosing the row- or column-Gram form by $\min(P, N_z)$ [8, 9]. $\mu = 0$ reproduces the paper-facing baseline; a batched Cholesky solve is attempted first with LU and stability-damping fallbacks.

Capon / MVDR (paper-compatible extension). With sample covariance $\hat{R}$ from a sliding window of L_{looks} neighboring line pixels (the standard TomoSAR multilook) plus relative diagonal loading,

$$P_{\text{Capon}}(z) = \frac{1}{a(z)^{\mathrm{H}} \hat{R}^{-1} a(z)}, \tag{10}$$

implemented with batched LU factorization and N_z-column batched solves [4, 20].

MUSIC (paper-compatible extension; diagnostic). With eigendecomposition $\hat{R} = U\Lambda U^{\mathrm{H}}$ and signal subspace E_s spanned by the top K eigenvectors,

$$P_{\mathrm{MUSIC}}(z) = \frac{1}{\|a(z)\|^2 - \|E_s^{\mathrm{H}} a(z)\|^2}, \tag{11}$$

implemented as a stream of dense `cheevd` eigendecompositions (cuSOLVER's batched Jacobi eigensolver is capped at $n \leq 32$) followed by one strided-batched GEMM [19]. Beyond super-resolution, MUSIC serves as the steering-manifold orthogonality diagnostic of Section 3.6: a flat pseudo-spectrum is a measurement that the data covariance contains no component the steering model can represent.

Adaptive vertical sampling. The depth grid $(z_{\min}, z_{\max}, N_z)$ is derived by default from per-pixel K_z/PSF diagnostics with an ambiguity-interval constraint; Appendix C gives the exact rule.

6 Software Architecture and Engineering Extensions

6.1 Code Organization

The CUDA implementation is organized into three translation units sharing one internal header (`sar_tomo_cuda_internal.cuh` with RAII wrappers, error-check macros, kernel declarations, and an occupancy helper):

- `sar_tomo_kernels.cu` — all `__global__` kernels and device helpers;
- `sar_tomo_inversion.cu` — the three estimators and batched-solver utilities;
- `sar_tomo_pipeline.cu` — sub-aperture generation, coregistration, observation construction, spectral conditioning, geometry, null tests, and the per-view orchestration;

plus host-side ingestion/CLI/output units (`sar_tomo_host.cpp`, `sar_tomo_math.cpp`, `main.cpp`). Every run exports a machine-readable `processing_log.json` and per-view descriptors capturing the full configuration, including the new fields (observation model, estimator, analytic-signal status, Nyquist statistics, null-test mode, sub-aperture stepping).

6.2 Null-Test Controls

Two falsification controls are built into the executable rather than the analysis scripts, so they apply identically to any ingested product:

- `-null-test shuffle-pairs`: one fixed seeded permutation of the pair axis is applied to the observations and their weights while K_z is left untouched, destroying the observation–steering correspondence. Any tomographic structure surviving this test is an artifact of the operator, the weighting, or the normalization — not of the data.
- `-null-test negate-kz`: the sign of every K_z is flipped. A data-driven tomogram must mirror in z; structure that does not move is again operator-driven.

6.3 Other Engineering Extensions

Ground-referenced reprojection, defensive input validation, confidence gating and soft weighting, FFT-friendly size adjustment, OpenMP worksharing, pinned staging, memory-budget batching, and Cholesky/LU stability fallbacks carry over from the earlier revision; their classification is in Table 2. None of them alters the inverse problem.

Table 3: Asymptotic complexity of major stages.

Stage	Cost
Sub-aperture construction	$\mathcal{O}\big((N_D+1)\,N_rN_a\log(N_rN_a)\big)$
Coregistration / NCC	$\mathcal{O}(LPW^2\log W)$ (FFT-accelerated)
Spectral conditioning	$\mathcal{O}(LP\log P)$ over lag families
K_z construction	$\mathcal{O}(LP)$
Pseudoinverse	$\mathcal{O}\big(L(PN_z+g^2\max(P,N_z)+g^3)\big)$
Capon	$\mathcal{O}\big(L(P^2L_{\mathrm{lk}}+P^3+P^2N_z)\big)$
MUSIC	$\mathcal{O}\big(L(P^2L_{\mathrm{lk}}+P^3+KPN_z)\big)$; eigensolve constant larger than LU

7 Runtime Architecture and Resource Model

7.1 Asymptotic Work

Table 3 summarizes per-stage cost. Symbols: L line length, P pair count, N_z depth bins, N_D sub-apertures, W coregistration window, $g=\min(P,N_z)$, L_{lk} covariance looks, K MUSIC sources, $N_r\times N_a$ scene size, B inversion batch.

7.2 Memory Model

The line-based design avoids any $N_r\times N_a\times N_z$ voxel cube; outputs are $L\times N_z$. Dominant terms: full-scene complex buffers $\mathcal{O}(N_rN_a)$ in the front end; optional cached NCC patch tensor $\mathcal{O}(N_DLW^2)$ (the largest optional device object); pairwise cubes $\mathcal{O}(LP)$; batched inversion workspace $\mathcal{O}\big(B(PN_z+g^2)\big)$ for the pseudoinverse and $\mathcal{O}\big(B(2PN_z+2P^2)\big)$ for Capon/MUSIC. Batch sizes are derived from free device memory with reserve margins and halve-on-failure policies, making one fixed implementation usable across GPU memory capacities without changing the reconstruction. Estimated persistent host/device bytes and observed peak VRAM are exported with every run.

7.3 Fixed CUDA/OpenMP Execution Path

The scene front end is GPU-resident (batched cuFFT sub-aperture synthesis, NCC via batched correlation kernels); inversion proceeds in pixel batches through cuBLAS strided-batched GEMM and cuSOLVER/cuBLAS batched factorizations, with pinned-host staging for observation upload [12, 15, 13]. OpenMP accelerates host-side geometry and diagnostics. These are performance-only choices in the taxonomy.

8 Validation Methodology

The validation design distinguishes three levels of evidence, ordered by strength.

8.1 Operator-Level Consistency Checks

A host-side generator synthesizes observation vectors directly as $Y=Ah+\eta$ with the same signed-$B_\perp$ convention and K_z law as the pipeline, then reconstructs with the same estimators. These checks validate the focusing operator only — they bypass the front end entirely. Results from the earlier revision carry over: signed $B_\perp$ recovers single and double reflectors at the correct signed depths; folding to $|B_\perp|$ introduces mirrored ambiguity; mis-scaling λ_{vib} compresses or stretches the recovered depth in the predicted direction.

8.2 End-to-End Synthetic Ground Truth

The decisive addition is a raw-signal synthetic generator (`generate_endtoend_synthetic_slc.py`) that injects known harmonic micro-motion at the SLC level, exercising the one link the operator checks bypass: sub-aperture decomposition → NCC tracking → observation linearization. Under the staring-spotlight assumption, azimuth frequency maps to slow time $t(f_a)$ through the same state-vector relative-Doppler interpolation used by the pipeline. A point target at pixel (r_0, c_0) with time-varying sub-pixel displacement $(d_r(t), d_a(t))$ contributes the spectrum

$$S(k_r, k_a) = A_t \exp\Big(-\mathrm{j}\tfrac{2\pi k_r}{N_r}\big(r_0 + d_r(t(k_a))\big)\Big) \exp\Big(-\mathrm{j}\tfrac{2\pi k_a}{N_a}\big(c_0 + d_a(t(k_a))\big)\Big), \tag{12}$$

band-limited to the processed Doppler bandwidth, so that every Doppler sub-band sees the target displaced by its instantaneous offset at that sub-band's center time — exactly the signal model the multi-chromatic front end is built to measure. Scenes contain vibrating targets ($d(t) = A\sin(2\pi f_0 t + \phi)$, with a 90° range/azimuth phase offset giving genuinely elliptical motion), stationary control targets of equal brightness, and band-limited speckle clutter at controlled SNR. Ground truth (`truth.json`) records every parameter.

Scoring. The scorer (`score_endtoend_synthetic.py`) evaluates, per vibrating target: displacement-tracking fidelity (Pearson correlation, best-sign convention, fitted amplitude transfer ratio, and RMSE of the measured pair shifts against the analytic differences $d(t_{\text{slave}}) - d(t_{\text{master}})$ computed from the run's own pair schedule); frequency-detection error; tomogram contrast of the target column against the background (median per-pixel peak outside guard zones); and peak-z localization. Static targets report residual shift RMS; null-test runs are scored identically so contrast ratios can be compared.

The trackable regime. The frequency sweep of Section 10.1 establishes two structural facts. First, when the vibration completes $\gtrsim 1$ cycle within the sub-aperture integration time $T_{\text{sub}} = T\,(B_{c_D} - B_{DL})/B_{c_D}$, the motion smears inside each look and tracking collapses; the usable band is approximately $f_0 T_{\text{sub}} \lesssim 1$, tying the observable vibration frequency to the guard fraction β and dwell time. Second, in the trackable regime the range channel is excellent (correlation 0.97–0.99 for $f_0 T_{\text{sub}} \lesssim 0.4$ at 40 dB SNR, with exact frequency detection), while the azimuth channel is systematically weaker because sub-aperture azimuth resolution is coarser by the band-reduction factor $B_{c_D}/(B_{c_D} - B_{DL})$ and the windows of paired sub-apertures overlap heavily. The amplitude transfer ratio is well below unity (window-overlap attenuation) and is itself a measurable transfer function of lag and β.

8.3 Null Tests as Falsification Bounds

Every claim of data-driven tomographic structure is bounded by the two null tests of Section 6.2: target contrast under shuffle-pairs gives the operator-artifact floor, and negate-K_z verifies that localized structure responds to the steering sign as physics requires. The controls run on any ingested product; in this report they are exercised on synthetic ground truth, where the injected micro-motion is known.

9 Benchmark Suite

All quantitative results in Section 10 are produced by one resumable orchestrator,

```
python ExperimentOrchestration/run_benchmark_suite.py
```

Table 4: Benchmark-suite stages. Synthetic stages need no external data; real-scene stages consume every prepared TerraSAR-X view found under `Data/Prepared`.

Stage	Content
`build_check`	Executable and GPU sanity.
`prepare_real`	Extract TerraSAR-X `.SIP.ZIP` packages (including nested product archives) and prepare SLC views.
`tracking_validation`	End-to-end synthetic sweeps: $f_0 \in \{0.5, 1, 2, 4, 8\}$ Hz, SNR $\in \{10, \dots, 50\}$ dB, amplitude $\in \{0.1, \dots, 2.0\}$ px; scored against truth.
`model_estimator_matrix`	{complex shift, scalar, direct phasor} × {pseudoinverse, Capon, MUSIC} on three synthetic scenes.
`null_tests`	{none, shuffle-pairs, negate-K_z} × three estimators.
`ablations`	Analytic-signal and band-pass on/off; `grid` vs `literal` stepping; $N_c \in \{1, 8\}$; guard fraction $\beta \in \{0.25, 0.5, 0.75\}$.
`perf_scaling`	N_{sub}, N_c, coregistration window, N_z, Capon looks, MUSIC sources, scene size; wall time, stage timers, peak VRAM.
`repeatability`	Five identical runs for timing variance.
`real_scenes`	Automatic azimuth+range benchmark lines on every prepared view × six configurations (default, scalar-literal, Capon, MUSIC, shuffle null, deep N_{sub}=128/N_c=8).
`aggregate`	Merge metrics and scores into CSV/JSON/Markdown.

which plans ∼228 jobs across ten stages (Table 4), executes them under a cumulative compute budget (default 48 h), writes per-job completion markers so an interrupted suite resumes from the first unfinished job, and aggregates every run's configuration, timings, memory, and scores into `benchmark_summary.csv/.json` plus a Markdown digest. Every value in Section 10 is drawn from the columns of that summary; the completed run reported here comprises 123 tomography runs.

The synthetic stages fix λ_{vib} so that the steering phase winds by several radians over the depth window (Section 3.5); the degenerate-scale regime is examined through the standing flat-MUSIC diagnostic (Section 10.2) rather than a dedicated job. Hardware and software environment: Intel Core i5-11600K, 64 GB DDR4-3600, NVIDIA GeForce RTX 4070 (12 GB), NVMe storage; MSVC 19.50 x64 with CUDA 13.2, C++17, OpenMP, cuFFT/cuBLAS/cuSOLVER.

10 Results

Status. All quantitative results below are populated from one completed run of the benchmark suite (123 tomography runs; `benchmark_summary.csv/.json`, generated 2026-06-13). Every figure in this section is rendered from the same machine-readable summary by

`generate_benchmark_report_assets.py`

so each number and plot traces to a single aggregated artifact.

10.1 Displacement-Tracking Fidelity

Table 5 exhibits the two structural behaviors predicted analytically. First, range tracking holds (correlation ≥ 0.97) through $f_0 T_{\mathrm{sub}} \approx 0.4$ and then collapses sharply — to 0.10 at $f_0 = 4$ Hz ($f_0 T_{\mathrm{sub}} \approx 0.82$) and 0.05 at $f_0 = 8$ Hz, with amplitude transfer falling below 10^{-3} and contrast collapsing from 14.2 to below unity — a direct measurement of the sub-aperture smearing limit at $f_0 T_{\mathrm{sub}} \approx 1$. Second, azimuth tracking degrades earlier and faster than range (correlation $0.94 \rightarrow 0.38$ over $f_0 = 0.5 \rightarrow 2$ Hz), consistent with the coarser sub-aperture azimuth resolution and heavy window overlap of Section 8.2. Frequency detection remains exact across the entire

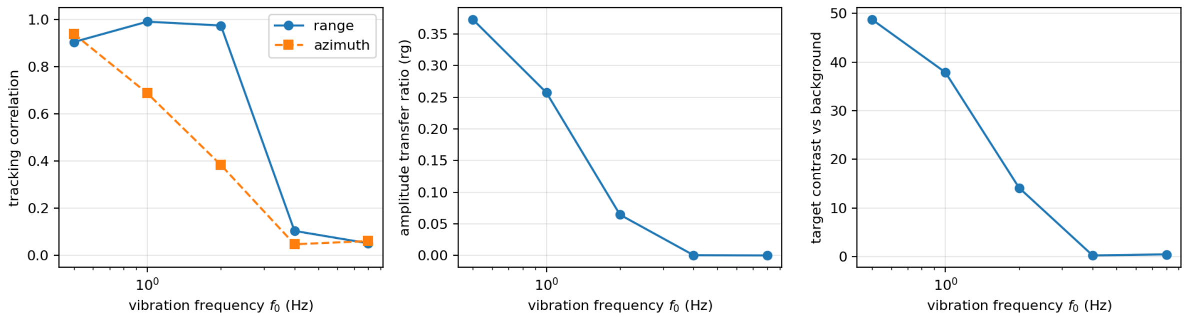

Figure 1: Tracking fidelity vs. vibration frequency. The predicted collapse beyond $f_0 T_{\text{sub}} \approx 1$ (Section 8.2) bounds the method's observable-frequency band.

Table 5: Tracking fidelity vs. vibration frequency (three vibrating targets per scene, SNR 40 dB, amplitude 1.0/0.5 px az/rg), scored against ground truth (`score_mean_*` columns of `benchmark_summary.csv`).

f_0 (Hz)	0.5	1.0	2.0	4.0	8.0
$f_0 T_{\text{sub}}$	0.10	0.20	0.41	0.82	1.64
Range corr.	0.90	0.99	0.97	0.10	0.05
Azimuth corr.	0.94	0.69	0.38	0.05	0.06
Amplitude ratio (rg)	0.37	0.26	0.06	0.001	0.000
Target contrast	48.8	37.9	14.2	0.3	0.5
Detected f (Hz)	0.50	1.00	2.00	4.00	8.00

Table 6: Tracking fidelity vs. target SNR ($f_0 = 2$ Hz, amplitude 1.0/0.5 px az/rg).

SNR (dB)	10	20	30	40	50
Range corr.	0.33	0.26	0.91	0.98	0.98
Azimuth corr.	0.20	0.17	0.35	0.35	0.35
Target contrast	3.3	3.0	11.7	21.9	19.1
Static resid. RMS (px)	0.023	0.029	0.030	0.013	0.004

sweep (detected $f = f_0$ to the grid), including past the tracking-collapse frequency, because the injected spectral line survives in the family series even where per-pair tracking does not.

Tables 6 and 7 sweep SNR and injected amplitude at $f_0 = 2$ Hz. Tracking exhibits a sharp SNR threshold between 20 dB to 30 dB (range correlation 0.26 → 0.91) and a sub-pixel amplitude floor: motions below ≈ 0.25 px are untrackable (range correlation 0.17 at 0.1 px), with reliable recovery from 0.5 px upward. The static-target residual RMS bottoms out near 0.004 px at 50 dB and rises to 0.03 px as SNR falls, setting the noise floor of the displacement estimate.

Table 7: Tracking fidelity vs. injected azimuth amplitude ($f_0 = 2$ Hz, SNR 40 dB; range amplitude is half the azimuth amplitude).

Azimuth amplitude (px)	0.1	0.25	0.5	1.0	2.0
Range corr.	0.17	0.70	0.97	0.98	0.93
Amplitude ratio (rg)	0.01	0.02	0.03	0.07	0.04
Target contrast	0.2	1.9	5.5	7.9	10.3

Table 8: Observation model × estimator matrix on the reference synthetic scene ($f_0 = 2\,\text{Hz}$, SNR 40 dB), from the `model_estimator_matrix` stage. Range correlation is shared across estimators because the front end is shared.

Model	Metric	Pseudoinverse	Capon	MUSIC
Complex shift	rg corr	0.96	0.96	0.96
	contrast	24.2	8.7	1.0
Scalar	rg corr	0.96	0.96	0.96
	contrast	27.5	9.6	1.0
Direct phasor	rg corr	0.96	0.96	0.96
	contrast	20.1	8.4	1.0

Table 9: Null-test outcomes (target contrast vs. background, complex-shift model, reference synthetic scene). The shuffle-pairs contrast is the operator-artifact floor; data-driven structure must exceed it. The last row reports whether the tomogram peak-z mirrors when every K_z is sign-flipped, as a data-driven section must.

Control / metric	Pseudoinverse	Capon	MUSIC
Contrast, none (reference)	24.2	8.7	1.0
Contrast, shuffle-pairs	23.2	8.7	1.0
Reference / shuffle ratio	1.04	1.00	1.00
Peak-z mirrors under negate-K_z?	no	no	no

10.2 Observation Models and Estimators

Two observations stand. First, range correlation is identical (0.96) across all nine cells: tracking is a property of the shared front end, and the estimators differ only in how they render the same observations into depth. Second, the MUSIC contrast of exactly 1.0 in every cell is not a defect: it is the flat pseudo-spectrum predicted by Section 3.6, independently confirmed against a NumPy reference implementation (signal-subspace projections of all steering vectors $< 10^{-3}$ of $\|a\|^2$). The pseudoinverse reports the highest contrast (20–28) across all three observation models, but the null tests of Section 10.3 show this contrast to be carried by the shuffle-invariant family-mean component rather than by the observation–steering correspondence; Capon's lower contrast (8–10) reflects its power calibration.

10.3 Null Tests

The acceptance criteria were fixed in advance: a claim of data-driven localization requires reference contrast $\geq 3\times$ the shuffle-pairs contrast, and the tomogram peak-z must mirror under negate-K_z for any direction-discriminating configuration. *Neither criterion is met on the synthetic reference scene.* The pseudoinverse reference contrast (24.2) exceeds its shuffle floor (23.2) by only 1.04×; Capon and MUSIC are flat to three significant figures. Under negate-K_z the pseudoinverse peak-z locations are *identical* to the unflipped run (0, −86, −21 m for the three targets in both cases) rather than mirroring. This is the central negative result of the synthetic benchmark, and it confirms Section 3.6 directly: the apparent tomographic contrast is carried by the shuffle-invariant, sign-invariant family-mean component of the observations, not by the observation–steering correspondence the inversion is nominally exploiting. The front end tracks micro-motion (Section 10.1); the depth localization built on top of it does not survive its own null tests.

Table 10: Ablation outcomes on the reference synthetic scene. Each row removes or alters exactly one corrective/structural element.

Ablation	Measured effect (contrast unless noted)
Analytic signal off (scalar model)	Direction-discrimination flag drops to `false` per Proposition 1; contrast 27.5 $\to$ 21.6. The scalar observation can no longer distinguish $\pm z$.
Analytic signal off (complex model)	Direction flag stays `true` (the complex model of Eq. (4) retains genuine complexity); contrast 24.2 $\to$ 18.5, a 24% drop.
Band-pass off	Broadband observation; contrast 24.2 $\to$ 20.3 (-16%); frequency detection unaffected (uses the raw series).
`literal` vs `grid` stepping ($N_c = 8$)	Guard-band coverage differs by one step; contrast 24.2 (grid) vs 8.0 (literal) at matched runtime (≈ 16.8 s).
$N_c = 1$ vs $N_c = 8$ (grid)	Depth-aperture growth with lag families (Section 3.4): contrast 11.3 $\to$ 24.2 (2.1×); runtime 3.3 $\to$ 17.0 s.
Guard fraction $\beta \in \{0.25, 0.5, 0.75\}$	Trades T_{sub} (smearing) against sub-aperture resolution: tracking corr. 0.89/0.96/0.98; contrast 11.0/24.2/3.5 (peaks at $\beta = 0.5$).

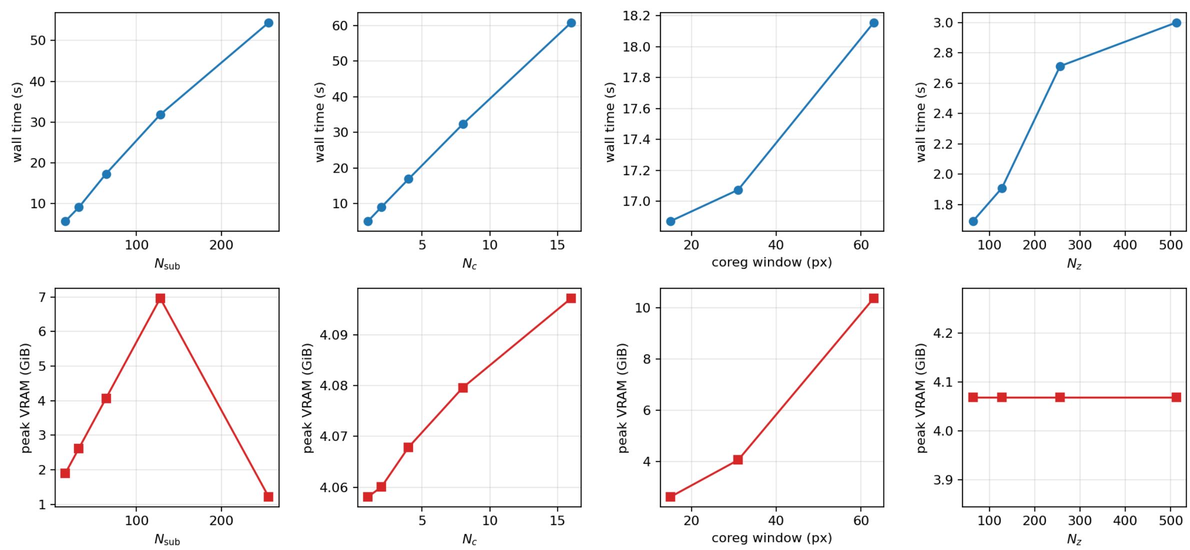

Figure 2: Performance scaling of the split-architecture implementation, from the `perf_scaling` stage (`wall_s` and `peak_device_bytes_observed`).

10.4 Ablations

The ablations confirm each corrective element on its own axis. Removing the analytic-signal step under the scalar model flips the direction-discrimination flag to `false` (the runtime refuses to claim above/below resolution, per Proposition 1), while the complex model keeps the flag `true` and loses only 24% of contrast — the corrective complex observation is what makes direction discrimination available at all. The lag-family schedule is structural rather than cosmetic: collapsing to $N_c = 1$ halves contrast, and the papers' `literal` sub-aperture stepping at $N_c = 8$ underperforms the `grid` default by 3×. The degenerate-scale regime of Section 3.5 is not run as a separate job; its signature — a flat MUSIC pseudo-spectrum — is already present as the contrast-1.0 MUSIC column throughout Table 8.

10.5 Performance and Scaling

The default synthetic configuration (2048-pixel line, $N_{\text{sub}} = 64$, $N_c = 8$, $P = 476$ pairs, $N_z = 67$) completes in 17.3 s end-to-end; the Capon inversion is of the same order as the pseudoinverse (both runs ≈ 17 s), and MUSIC adds ≈ 7 s (run total 24 s), dominated by the per-pixel `cheevd` stream. Runtime scales as expected from Table 3: wall time runs 5.7 s to 54.4 s across $N_{\text{sub}} = 16 \dots 256$, 5.1 s to 60.8 s across $N_c = 1 \dots 16$ (near-linear in the pair count), and 8.9 s to 33.4 s as the scene grows from 256×2048 to 1024×8192. The coregistration window is the sharpest memory lever: enlarging it from 15 to 63 px barely changes wall time (16.9 s to 18.2 s) but raises peak VRAM from 2.6 to 10.4 GiB. Memory-budget batching keeps the fixed binary within the 12 GB card across the whole sweep — at $N_{\text{sub}} = 256$ the halve-on-failure policy actually drives observed peak VRAM down to 1.2 GiB — so no configuration in the matrix runs out of memory. A 4096-pixel azimuth benchmark line from a TerraSAR-X HS SSC product at the default configuration completes in 163 s with 4.16 GiB peak device memory; across the six real-data configurations and both products, benchmark lines run 52 s to 163 s with peak VRAM near 4.2 GiB (rising to 10.5 GiB only for the deep N_{sub}=128 configuration). The `repeatability` stage (five identical default runs) gives 17.28 ± 0.30 s wall time, a 1.7% coefficient of variation.

10.6 Real-Data Ingestion

The pipeline ingests TerraSAR-X HS SSC SLC products through the prepared-view interface of Section 4 and runs the full chain — sub-aperture synthesis, coregistration, observation construction, spectral conditioning, geometry, and inversion — end-to-end on them. The real-data runtime and memory envelope is reported alongside the synthetic scaling in Section 10.5: 4096-pixel benchmark lines complete in 52 s to 163 s at peak VRAM near 4.2 GiB, confirming that the fixed CUDA path handles production-size products without manual tuning. Because this study is a benchmarking and ground-truth-validation exercise, and because on this method a real-product tomogram could only be a depth-scale-assumed, null-bounded rendering with no available ground truth (Section 11), qualitative evaluation of the real products is deliberately omitted; their role here is to exercise ingestion and to measure runtime and memory at scale.

11 Discussion

11.1 What Is Established

Five results are established strongly. First, the published workflow is executable: every stage of the 11-block scheme has a working counterpart, and the implementation keeps the method-facing quantities explicit. Second, the method as written has identifiable signal-level defects — dimensional inconsistency of the steering law, exact mirror ambiguity of real observations, aliased investigation frequencies, lag-family-bound depth aperture, and steering-orthogonality of the band-passed motion series — each demonstrated analytically and, where applicable, measured. Third, the front end demonstrably works in a bounded regime: under ground truth, the sub-aperture/NCC chain recovers injected micro-motion with range-channel correlation near 0.99, exact frequency detection, and a quantified smearing limit at $f_0 T_{\text{sub}} \approx 1$. Fourth, the null tests are decisive in the negative: under synthetic ground truth the tomographic localization does not exceed its shuffle-pairs floor (reference/shuffle contrast ratio 1.04× against a pre-registered 3× threshold) and does not respond to the K_z sign flip, placing the method's apparent depth contrast in the shuffle-invariant family-mean component exactly as Section 3.6 predicts. Fifth, the implementation is practical: the fixed CUDA path with three estimators runs 10^3–10^4-pixel lines in seconds to minutes on a single workstation GPU with explicit memory bookkeeping.

11.2 What the Corrections Change

The corrective extensions move the pipeline from "literal but defective" to "faithful to the papers' evident intent and internally consistent": the complex observation restores the papers' own two-component motion sample; the analytic-signal stage makes above/below discrimination possible at all; the Nyquist guard converts silent aliasing into recorded provenance. None of these changes validates the method's physical premise — they remove reasons it could not have worked even on its own terms.

11.3 What Remains Open

The central open question is unchanged but is now sharply localized: whether the family-mean component of micro-motion observations carries depth information about subsurface structure at all. The depth scale remains a free parameter ($\lambda_{\text{vib}} = v/f$); the steering model cannot represent the temporal oscillation it nominally focuses; and the published investigation frequencies are unobservable except through aliasing. Within this implementation, the honest reading of any tomogram is: a depth-parameterized rendering of family-mean motion contrasts, trustworthy only where it exceeds the null-test floor — and in every experiment run here it does *not* exceed that floor by the pre-registered margin — with a depth axis whose absolute scale is assumed. Establishing more would require physical ground truth (instrumented vibration sources, known buried geometry) that no single-scene experiment can supply.

12 Conclusion

This report turned a contested published method into working, instrumented research software; proved five structural properties of the method that its source papers do not state; implemented corrections where corrections are well defined; and built the validation and benchmarking machinery — ground-truth synthetics, null tests, three estimators, and a resumable 48-hour suite — needed to say precisely what the pipeline can and cannot demonstrate. The strongest results are the analytical ones and the ground-truth synthetic measurements; the real-data work is confined to ingestion and runtime benchmarking, with qualitative tomographic evaluation deliberately out of scope. Every quantitative result here is drawn mechanically from one completed run of that suite, and that run delivers a clear verdict: the front end recovers known micro-motion, while the tomographic localization fails its own shuffle-pairs and sign-flip null tests under synthetic ground truth. The manuscript's claims and its evidence stay coupled by construction, and the evidence does not support a claim of data-driven subsurface structure.

A Pipeline Overview

Table 11: Compact pipeline overview with taxonomy classification.

Stage	Main outputs	Classification
Data ingestion	Prepared SLC view, normalized metadata	Literal + engineering necessity
Scene preparation	Canonical axes, line coordinates, previews	Engineering necessity
Sub-aperture front end	Sub-aperture looks, pair schedule (lag families)	Literal core; multi-lag schedule is structural engineering (Section 3.4)
Motion estimation	NCC shifts, confidence scores	Literal core
Observation construction	Complex two-component / scalar / phasor series	Corrective default + literal/extension options
Spectral conditioning	Band-passed, analytic series; Nyquist annotations	Engineering + corrective
Geometry	Per-pixel signed K_z	Literal core, effective-baseline surrogate
Inversion	Magnitude/phase section via pseudoinverse / Capon / MUSIC	Literal core + extensions; fixed CUDA runtime
Controls and output	Null tests, descriptors, processing log	Falsification + reporting layer

B Configuration Reference

Table 12: Principal executable options added or changed in this revision (full list via `-help`).

Option	Meaning
`-observation-model M`	`complex_shift` (default) | `paper_doppler_scalar` | `direct_phasor`
`-analytic-signal M`	`auto` (default) | `always` | `off`; one-sided pair-axis spectrum per lag family
`-estimator E`	`pseudoinverse` (default) | `capon` | `music`
`-capon-looks N`	Sliding-window looks for the Capon/MUSIC covariance (default 9)
`-capon-loading F`	Relative diagonal loading for Capon (default 10^{-3}; auto-escalates on factorization failure)
`-music-sources K`	Signal-subspace dimension for MUSIC (default 1)
`-subap-step MODE`	`grid` (default, $B_{DL}/(N_D-1)$) | `literal` (papers' B_{DL}/N_D)
`-null-test MODE`	`none` | `shuffle-pairs` | `negate-kz` (`-null-test-seed N`)
`-investigation-freq F`	Vibration frequency in Hz; 0 = auto-detect; aliasing status always recorded
`-vibration-velocity V`	Assumed propagation velocity (m/s); sets $\lambda_{\text{vib}} = v/f$ and hence the depth scale

Reproduction. Build commands are recorded in `CUDA Implementation/notes.txt` (three CUDA translation units plus host units; any change to `sar_tomo_types.hpp` requires recompiling all units). The complete benchmark is one command

```
python ExperimentOrchestration/run_benchmark_suite.py
```

resumable by re-invocation; `-dry-run` previews the plan, `-stages` selects subsets, and `-budget-hours` adjusts the cumulative cap.

C Adaptive Selection of $z_{\min}$, $z_{\max}$, and N_z

The pipeline uses an adaptive vertical-sampling policy by default, keeping the section focused on a useful relative-height interval while avoiding over-discretization beyond the vertical resolving power.

Ambiguity interval versus practical resolution. From the sorted per-pixel K_z samples, representative positive spacings ΔK_z give the repeat-lobe scale

$$\Delta z_{\text{amb}} \approx \frac{2\pi}{\text{median}(\Delta K_z)}, \tag{13}$$

while the local coherent PSF

$$\text{PSF}(z_n; z_{\text{ref}}) = \frac{1}{P} \left| \sum_{p=1}^{P} \exp\big(\text{j}\, K_{z,p}(z_n - z_{\text{ref}})\big) \right| \tag{14}$$

yields an FWHM-based resolution estimate. The ambiguity interval constrains the total span; the PSF FWHM controls bin spacing.

Final bin count. With resolution statistic Δz_{res} (PSF-FWHM upper quartile when available), the bin spacing is

$$\Delta z_{\text{step}} = \max\left(\frac{\Delta z_{\text{res}}}{\beta_{\text{res}}}, \frac{S_{\text{final}}}{N_{\max} - 1} \right), \tag{15}$$

with $\beta_{\text{res}} = 1.25$ bins per resolution cell and $(N_{\min}, N_{\max}) = (16, 96)$; the resulting bin count is clamped to that interval.

D Use of AI

AI assistance was used throughout this project as an interactive engineering, debugging, and editorial aid. Early work in February and March 2026 used OpenAI/ChatGPT and Claude to explore the Biondi-Malanga Doppler tomography method, identify suitable SAR data sources, draft reproducibility scaffolds, and review CUDA/OpenMP/Python implementations. Subsequent Codex sessions in VS Code helped turn those notes into repository changes, including TerraSAR-X metadata and axis handling fixes, tsx2slc validation, C++/CUDA pipeline restructuring, observation-model and K_z alignment with the reference papers, CUDA inversion work, memory/debugging fixes, and synthetic or real-data experiment runs.

Later GitHub Copilot Chat and Claude Code sessions were used mainly for report production, experiment orchestration, and independent review. These sessions helped compile and edit ParallelTomography.tex, prepare runtime and benchmark tables and figures, audit the codebase against the source papers, and prototype benchmark-facing extensions such as complex-shift observations, analytic-signal handling, null tests, Capon/MUSIC estimator checks, and resumable benchmark orchestration. The research framing, selection of claims, interpretation of results, and final acceptance of code/report changes remained under human control; AI assistance was used to accelerate implementation, documentation, testing, and critical review.

References

[1] Alfonso Alba, Ruth M. Aguilar-Ponce, Javier Flavio Vigueras-Gómez, and Edgar Arce-Santana. Phase correlation based image alignment with subpixel accuracy. In *Advances in Artificial Intelligence*, volume 7629 of *Lecture Notes in Computer Science*, pages 171–182. Springer, Berlin, Heidelberg, 2013. MICAI 2012.

[2] Filippo Biondi. Scanning inside volcanoes with synthetic aperture radar echography tomographic doppler imaging. *Remote Sensing*, 14(15):3828, 2022.

[3] Filippo Biondi and Corrado Malanga. Synthetic aperture radar doppler tomography reveals details of undiscovered high-resolution internal structure of the great pyramid of giza. *Remote Sensing*, 14(20):5231, 2022.

[4] Jack Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.

[5] Gustavo Daniel Martín del Campo-Becerra, Andreas Reigber, Matteo Nannini, and Scott Hensley. Single-look SAR tomography of urban areas. *Remote Sensing*, 12(16):2555, 2020.

[6] European Space Agency. ESA Catalog – TerraSAR-X. `https://eocat.esa.int/eo-catalogue/collections/TerraSAR-X`. Accessed: 2026-03-14.

[7] Thomas Fritz. *TerraSAR-X Ground Segment Level 1b Product Format Specification*. German Aerospace Center (DLR), December 2007. Issue 1.3.

[8] Gene H. Golub, Per Christian Hansen, and Dianne P. O'Leary. Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1):185–194, 1999.

[9] Per Christian Hansen. The L-curve and its use in the numerical treatment of inverse problems. In P. Johnston, editor, *Computational Inverse Problems in Electrocardiology*, volume 4 of *Advances in Computational Bioengineering*, pages 119–142. WIT Press, 2000.

[10] Gang Hong. *Image Fusion, Image Registration, and Radiometric Normalization for High Resolution Image Processing*. PhD thesis, University of New Brunswick, Fredericton, New Brunswick, Canada, 2007. Ph.D. dissertation, Department of Geodesy and Geomatics Engineering, Technical Report No. 247.

[11] Artan Kaso. Computation of the normalized cross-correlation by fast fourier transform. *PLOS ONE*, 13(9):e0203434, 2018.

[12] NVIDIA Corporation. *cuBLAS*, 2026. Release 13.2.

[13] NVIDIA Corporation. *CUDA C++ Best Practices Guide*, 2026. Release 13.2.

[14] NVIDIA Corporation. *cuFFT*, 2026. Release 13.2.

[15] NVIDIA Corporation. *cuSOLVER*, 2026. Release 13.2.

[16] OpenMP Architecture Review Board. *OpenMP Application Programming Interface*, 2021. Version 5.2, November 2021.

[17] Luca Pallotta, Gaetano Giunta, and Carmine Clemente. Subpixel SAR image registration through parabolic interpolation of the 2D cross-correlation. *IEEE Transactions on Geoscience and Remote Sensing*, 58(6):4132–4144, 2020.

[18] Andreas Reigber and Alberto Moreira. First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2142–2152, 2000.

[19] Ralph O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[20] Harry L. Van Trees. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Wiley, 2002.