

Two-Component Covariance-Aware Uncertainty Calibration for Demosaiced One-Shot-Color Aperture Photometry

Jinxuan Liu

2511611209@hbut.edu.cn

Hubei University of Technology

Xurui Zhang

2511611208@hbut.edu.cn

Hubei University of Technology

Research Article

Keywords: Aperture photometry, color filter array, coverage probability, covariance propagation, demosaicing, image-based measurement, measurement uncertainty, one-shot-color imaging, uncertainty calibration

Posted Date: October 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11312795/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

TWO-COMPONENT COVARIANCE-AWARE UNCERTAINTY CALIBRATION FOR DEMOSAICED ONE-SHOT-COLOR APERTURE PHOTOMETRY

A PREPRINT

Jinxuan Liu and Xurui Zhang

Detroit Green Technology Institute,
Hubei University of Technology, Wuhan, China
2511611209@hbut.edu.cn, 2511611208@hbut.edu.cn
Jinxuan Liu: <https://orcid.org/0009-0003-6737-4648>

ABSTRACT

Demosaicing introduces spatial covariance by reusing noisy raw samples, causing diagonal aperture budgets to underestimate measurement uncertainty. We present a two-component budget for one-shot-color aperture photometry. The first component propagates covariance through fixed linear demosaicing using the adjoint weights $q = H^T w$, without forming an image-sized covariance matrix. The second estimates a residual field factor from native measurements at radius 8 and applies it across demosaicing routes and radii within the same field. We evaluated nominal 95% coverage and standardized-residual dispersion using adjacent-frame flux differences. At the predeclared M67 scalar-G endpoint, 38 retained pairs and 52 stars provided 1,976 star-pair observations. Full propagation increased bilinear coverage from 0.6417 to 0.9534 and reduced dispersion from 2.5310 to 0.9570. For Malvar–He–Cutler (MHC) demosaicing, coverage increased from 0.8790 to 0.9545 and dispersion decreased from 1.4812 to 0.9682. A 200-pair conditional M20 simulation reproduced this correction under known raw-noise inputs. M45 and Canon 90D measurements showed similar residual dispersions across native and full budgets, motivating the field component. With that component, Canon met both calibration criteria at radii 6 and 10. These results provide an efficient covariance-aware uncertainty budget and distinguish the demosaicing correction from residual field-level variation.

Keywords Aperture photometry, color filter array, coverage probability, covariance propagation, demosaicing, image-based measurement, measurement uncertainty, one-shot-color imaging, uncertainty calibration.

I Introduction

Aperture-flux uncertainty determines how reliably repeated measurements can be compared. One-shot-color cameras support this task in astronomy and other forms of quantitative imaging [1]. Conventional photometric budgets combine source, background, detector, and digitization variances [2]. The documented `photutils` aperture interface accepts per-pixel errors [3], while SIRIL uses source and sky-noise terms [4]. Demosaiced images additionally require the covariance created when several output pixels share the same raw sample.

A Bayer array records one color at each sensor site [5]. Photometry can use the native color sublattices or reconstruct full-resolution color planes. The AAVSO DSLR manual recommends native color separation for its default workflow and also describes tools that interpolate to full resolution [6, Sec. 5.4]. Both approaches use the same raw frame, but interpolation changes the uncertainty of an aperture sum by correlating neighboring output pixels.

Noise correlation in reconstructed images has an established theoretical basis. Fruchter and Hook analyzed its effect in drizzled images [7]. Demosaicing noise has also been used to identify reconstruction algorithms [8], and cross-channel models have supported image denoising [9, 10]. Yang and colleagues examined demosaicing in color-based

ratio measurements [11]. For aperture photometry, the practical question is whether the propagated budget gives nominal 95% intervals the expected coverage in repeated measurements.

The required pixel-noise models and linear propagation laws are available. EMVA 1288 Release 4.0 Linear specifies a photon-transfer relation under a white-noise model [12]; Release 4.0 General accommodates spatially correlated temporal noise [13]. Withagen and colleagues characterized pixel variance through the linear form $aS + b$ [14]. Noise propagation through known linear processing follows the standard covariance law [15, 17]. AstroWISP provides direct RAW/Bayer photometry with PSF/PRF modeling and pixel-error propagation [16]. We apply these principles to the specified aperture measurement on a reconstructed image, represented in raw space by $q = H^T w$.

We combine exact adjoint propagation with a native-derived field factor and evaluate the two components using paired flux differences. Background-subtracted G flux is the primary measurement; an ensemble R/G ratio tests the additional cross-channel term. Each diagonal/full comparison uses identical measured fluxes, so the change directly assesses the variance budget. Phase-specific noise laws and coordinate-only registration preserve the connection to the raw samples.

The M67 analysis separates development and evaluation at a predeclared 8-pixel radius for bilinear and Malvar–He–Cutler (MHC) demosaicing. A conditional M20 simulation tests propagation with known noise inputs. M45 provides a second field on the same instrument family, and Canon 90D provides a second camera family. Neighboring radii assess the application of a single field factor across aperture sizes.

II Proposed Method

Fig. 2 shows the adjoint budget and its cross-channel extension. The field factor s_8 scales the adjoint variance as defined below. Table I contrasts the uncertainty treatment in existing tools with ours, and Procedure IV lists the steps.

II.A Aperture measurement

Let x contain raw CFA samples in a local patch, and let H_c map those samples to output channel c . The weights w combine a fractional circular aperture with a local background annulus. Aperture weights are positive fractional areas. Annulus weights are negative and scaled to cancel a spatially constant background. The channel flux is

$$F_c = w^T H_c x. \quad (1)$$

The aperture and annulus were sampled at eight subpixels per raw pixel along each axis; registration moved the aperture coordinates without touching the images. With fractional areas a_p and b_p , the implemented weights are $w_p = a_p - (\sum a_p / \sum b_p) b_p$, with annulus radii 13–19 raw pixels. Supplementary Note S4 specifies the Bayer phases, kernels, noise maps, native normalization, and boundary handling. Throughout, “bilinear” denotes the two-green-plane averaging implementation used here, which also mixes already sampled green pixels; MHC preserves sampled green pixels.

For scalar G flux F_G the two budgets keep different terms. The diagonal budget keeps only per-pixel variance,

$$v_{G,\text{diag}} = \sum_p w_p^2 \text{Var}[(H_G x)_p], \quad (2)$$

and drops covariance between output pixels. When raw samples are independent with variance matrix Σ_x , the exact variance can be written without forming an image-sized covariance,

$$v_{G,\text{full}} = (H_G^T w)^T \Sigma_x (H_G^T w). \quad (3)$$

Proposition (adjoint budget). Assume (i) the demosaicing operator H_c is fixed and linear within each Bayer phase, (ii) raw samples are independent with diagonal variance Σ_x , and (iii) the readout is the linear form (1). Then (3) is exact conditional on these quantities. For a sparse operator the cost is $O(\text{nnz}(H) + N_{\text{raw}})$, where nnz counts nonzero entries. Fixed finite-support bilinear and MHC kernels give $O(N_{\text{raw}})$ cost. No image-sized covariance Σ_y is formed.

II.B Cross-channel ratio

The secondary quantity was an ensemble ratio $q = F_R/F_G$ after summing kept stars in a frame. Its first-order variance uses $j = (1/F_G, -F_R/F_G^2)^T$. Three budgets were compared

$$v_{q,\text{diag}} = j_R^2 v_{R,\text{diag}} + j_G^2 v_{G,\text{diag}}, \quad (4)$$

$$v_{q,\text{spatial}} = j_R^2 v_{R,\text{full}} + j_G^2 v_{G,\text{full}}, \quad (5)$$

$$v_{q,\text{full}} = v_{q,\text{spatial}} + 2j_R j_G \text{Cov}(F_R, F_G). \quad (6)$$

The spatial budget keeps covariance within each channel; the full budget adds the R–G term on top. Bilinear reconstructs R and G from disjoint raw planes, so its induced R–G covariance is zero by construction. MHC reuses green samples in gradient corrections to other channels [18], so its term can be nonzero. Section IV.E measures that increment.

II.C Two-component extension with within-field factor

We estimate a residual field factor from native measurements at the primary radius and apply it to the covariance-aware budgets,

$$s_8 = \max\left(0, \frac{\sum d_{\text{native},r8}^2}{\sum v_{\text{native},r8}} - 1\right), \quad v_{\text{total}} = v_{\text{full}}(1 + s_8). \quad (7)$$

Here $d_i = F_{i,2} - F_{i,1}$ is a normalized adjacent-frame difference, and $v_i = v_{i,1} + v_{i,2}$ is its predicted variance under zero interframe covariance. The variance is conditional on the fitted noise laws, coordinates, and frame normalization. Writing $R_8 = \sum_i d_i^2 / \sum_i v_i$ gives $1 + s_8 = \max(1, R_8)$. For $R_8 > 1$, the scaled native fitting records have unit aggregate empirical-to-predicted variance ratio; for $R_8 \leq 1$, the budget is unchanged. Coverage and residual dispersion are assessed separately from this fit.

Within each field, the retained pairs provide both factor estimation and assessment. We apply the same s_8 to bilinear-full and MHC-full at radii 6, 8, and 10. Radius 8 assesses application across routes, while radii 6 and 10 assess transfer across aperture sizes. The field-factor analysis uses the measurements from the predeclared adjoint-only endpoints. Table II and Fig. 1 report the pooled results; Table III gives native factors estimated separately at each radius. Supplementary Tables S1–S3 provide the detailed values.

Table I: Uncertainty treatment in selected interfaces and methods, as described in the cited documentation. The experimental diagonal budget is defined by (2).

Method/interface	Uncertainty treatment relevant to this study
Photutils aperture	Per-pixel error input; no interpixel covariance argument.
Siril aperture	Source and sky-noise error model in the cited interface.
EMVA 1288 Linear	White temporal-noise model; excludes correlation-inducing preprocessing.
EMVA 1288 General	Sensor characterization including spatially correlated temporal noise and preprocessing beyond the Linear restrictions.
AstroWISP	Direct RAW/Bayer photometry with PSF/PRF treatment and pixel uncertainty propagation.
Present study	Fixed-operator adjoint variance; paired difference coverage; empirical same-field multiplier and cross-radius checks.

Table II: Two-component transfer (ledger in Supplementary Table S1). The field factor s_8 is calibrated only at radius 8 on native and applied to radii 6, 8, and 10 on bilinear-full and MHC-full pooled. Intervals are 95% cluster bootstrap over frame pairs, $B = 2000$. Gates: coverage 0.92–0.98 and s_z 0.90–1.10. The factor is refitted within each bootstrap replicate.

Set	r	s_8	Coverage [95% CI]	s_z [95% CI]	Gate
M67	6	0	0.9064 [0.8601, 0.9431]	1.2002 [1.0273, 1.3733]	Fail
M67	8	0	0.9539 [0.9423, 0.9643]	0.9625 [0.9008, 1.0255]	Pass
M67	10	0	0.9162 [0.8866, 0.9398]	1.1976 [1.0925, 1.3137]	Fail
M45	6	3.39	0.9583 [0.9347, 0.9692]	0.9005 [0.7700, 1.1571]	Pass
M45	8	3.39	0.9582 [0.9458, 0.9692]	0.8868 [0.8270, 0.9741]	Fail
M45	10	3.39	0.9502 [0.9335, 0.9630]	1.0253 [0.8865, 1.1804]	Pass
Canon	6	3.61	0.9381 [0.9235, 0.9529]	1.0204 [1.0027, 1.0398]	Pass
Canon	8	3.61	0.9410 [0.9304, 0.9549]	1.0039 [0.9869, 1.0214]	Pass
Canon	10	3.61	0.9505 [0.9385, 0.9629]	0.9735 [0.9457, 0.9992]	Pass
M20	6	0	0.9510 [0.9462, 0.9564]	0.9947 [0.9788, 1.0064]	Pass
M20	8	0	0.9527 [0.9488, 0.9570]	0.9904 [0.9756, 0.9999]	Pass
M20	10	0	0.9534 [0.9488, 0.9580]	0.9876 [0.9714, 0.9993]	Pass

Table III: Native empirical-to-predicted variance ratio minus one at each radius (Supplementary Table S3). These values precede truncation. Each field’s radius-8 value is truncated at zero to obtain the field factor used for transfer.

	$r = 6$	$r = 8$	$r = 10$
M67	0.6382	−0.1612	0.3692
M45	2.9807	3.3911	5.5938
Canon	3.6414	3.6055	3.2885
M20	−0.0081	−0.0171	−0.0228

II.D Variance inflation and the autocorrelation special case

For fixed H and independent raw noise, expansion of (2) and (3) gives

$$\Phi = \frac{v_{\text{full}}}{v_{\text{diag}}} = \frac{\sum_i \sigma_i^2 (\sum_p w_p H_{pi})^2}{\sum_i \sigma_i^2 \sum_p w_p^2 H_{pi}^2}. \quad (8)$$

The denominator is positive for the measurements considered here. The operator entries encode Bayer phase and boundary handling, while w includes the negative background-annulus weights. Each raw variance σ_i^2 weights the contribution of its own sample, allowing spatially varying noise to enter the ratio explicitly.

For a single shift-invariant convolution $H_{pi} = h(p-i)$, equal raw variances, and consistent boundary treatment, define $A_h(d) = \sum_u h(u)h(u+d)$ and $A_w(d) = \sum_p w_p w_{p+d}$. Then the ratio reduces to

$$\Phi = \frac{\sum_d A_h(d)A_w(d)}{A_h(0)A_w(0)}.$$

For Bayer reconstruction, the corresponding uniform-noise expression uses phase-resolved kernel and aperture correlations, as derived in Supplementary Section S3. At the specified centered test geometry, the ratios are 6.40–6.60 for our bilinear implementation and 2.43–2.46 for MHC. Actual heteroscedastic pair-budget ratios are summarized in Supplementary Table S4. All experimental full budgets use the exact adjoint in (3).

Table IV: Calculation of the two-component budget. The field factor is estimated from native measurements at radius 8 and reused across radii within the field.

Input: fixed raw variance maps, operator H , weights w ; native radius-8 pair differences and their predicted difference variances.
1. Adjoint: $q = H^T w$, $v_{\text{full}} = q^T \Sigma_x q$ for each frame; sum both frame variances for each difference.
2. Field: s_8 from (7) on native at $r = 8$ only.
3. Total difference variance: $v_{d,\text{total}} = (v_{1,\text{full}} + v_{2,\text{full}})(1 + s_8)$ at $r = 6, 8, 10$.
Output: nominal 95% difference interval $d \pm 1.95996\sqrt{v_{d,\text{total}}}$ and its zero-coverage indicator.

II.E Calibration diagnostics

For each predefined adjacent-frame pair the target difference is zero. With observed difference d and predicted variance $v_1 + v_2$ the standardized residual is

$$z = \frac{d}{\sqrt{v_1 + v_2}}. \quad (9)$$

We report the fraction of residuals satisfying $|z| \leq 1.95996$ and their sample standard deviation s_z . A budget meets the calibration criterion when coverage is 0.92–0.98 and s_z is 0.90–1.10. The additional ratio $\sum d^2 / \sum (v_1 + v_2)$ is an uncentered empirical second-moment to predicted-variance ratio, reported as “Emp./pred.” in the tables. The gates specify tolerances around nominal coverage and dispersion, following the uncertainty terminology in JCGM 100:2008 [17]. We examined the operating characteristics with a clustered-normal simulation, $z_{ij} = \lambda(\sqrt{\rho}u_i + \sqrt{1-\rho}e_{ij})$, using M67/M20 sample sizes, 10,000 replicates, and seed 20260902. At $\lambda = 1$, joint pass probabilities were 1.0000 for $\rho = 0$ and 0.9654 for $\rho = 0.4$. They were at most 0.0048 at $\lambda = 0.8$ and 0.0381 at $\lambda = 1.2$.

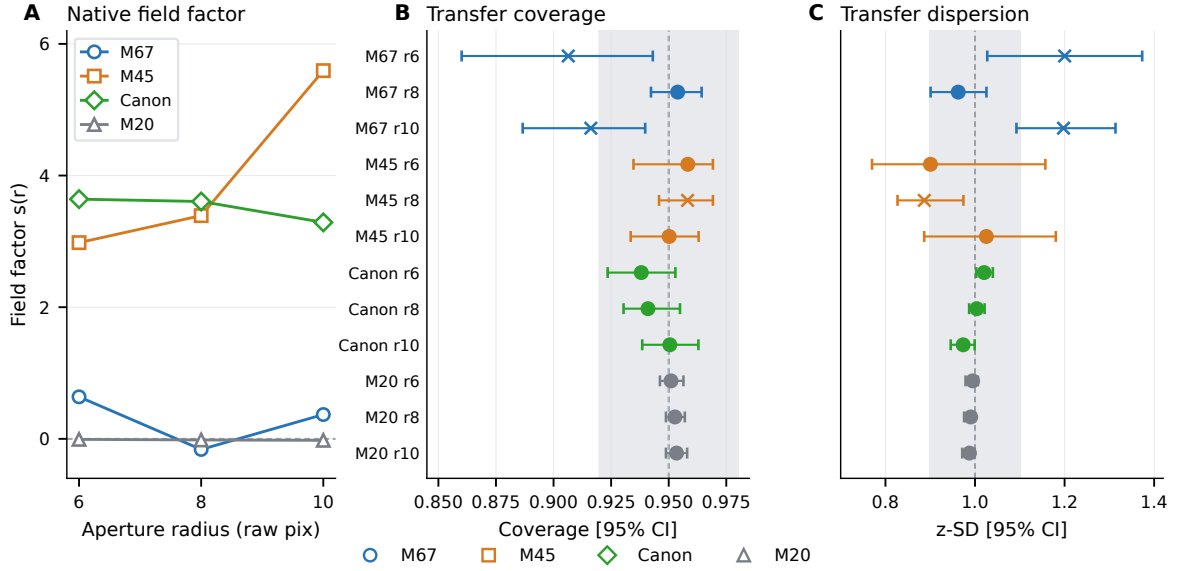


Figure 1: Two-component transfer. (A) Native field factor $s(r)$ per radius. (B) Pooled transfer coverage with 95% cluster-bootstrap intervals. (C) Pooled transfer dispersion with 95% intervals. Gray bands are the joint gates; o is Pass and x is Fail. M45 radius 8 misses dispersion by 0.013 although its interval overlaps the gate. Route-specific results are reported in Supplementary Table S2.

III Datasets and Experimental Protocol

III.A M67 Dataset

We used the public MILAN Sky Survey M67 sequence [19] archived on Zenodo [20] and verified its archive checksum. The analysis includes all 114 available raw BGGR FITS exposures, numbered 0001–0114. Each exposure lasted 10 s and was acquired with a Stellina station using a Sony IMX178 sensor. Removing a one-pixel border produced an even RRGB crop without altering interior samples.

Frames 1 to 35 were for development. Frames 36 to 114 were held for evaluation and grouped into 39 nonoverlapping adjacent pairs. Frame 114 had no partner. Development locked detection, isolation, aperture and annulus geometry, the noise law, registration thresholds, and the pass bands. The primary radius was 8 raw pixels, with radii 6 and 10 specified as sensitivity checks.

The field drifts and rotates through the sequence. Stars were found in a consensus image built from both native green phases. An affine map carried the development catalog to each frame, and only aperture coordinates were moved. Registration used the original image samples. A fit needed at least 30 matched detections and RMS at most 0.8 half-resolution pixel. Aperture centers were quantized to 0.125 raw pixel to match the fractional-area code.

The catalog holds 52 isolated unsaturated stars. Centers stay at least 24 half-resolution pixels apart, so aperture plus annulus supports, including the MHC halo, use disjoint raw samples. Positive G measurements were retained regardless of R flux. Many single-star R fluxes are not positive in 10 s, so the cross-channel endpoint is the ratio of summed R to summed G over the kept ensemble.

III.B Noise Model Calibration

For each CFA phase, pixel differences in development frames provided the variance law $\sigma_i^2 = a_i s_i + b_i$. Native-CFA aperture residuals on frames 1–35 supplied an additional phase multiplier to match development aperture repeatability. This multiplier was the squared standard deviation of the uncalibrated z and scaled slope and intercept together. The resulting phase laws were shared across all routes and held fixed during evaluation. Each frame’s multiplicative scale was the median ratio of native-G star fluxes to the development reference.

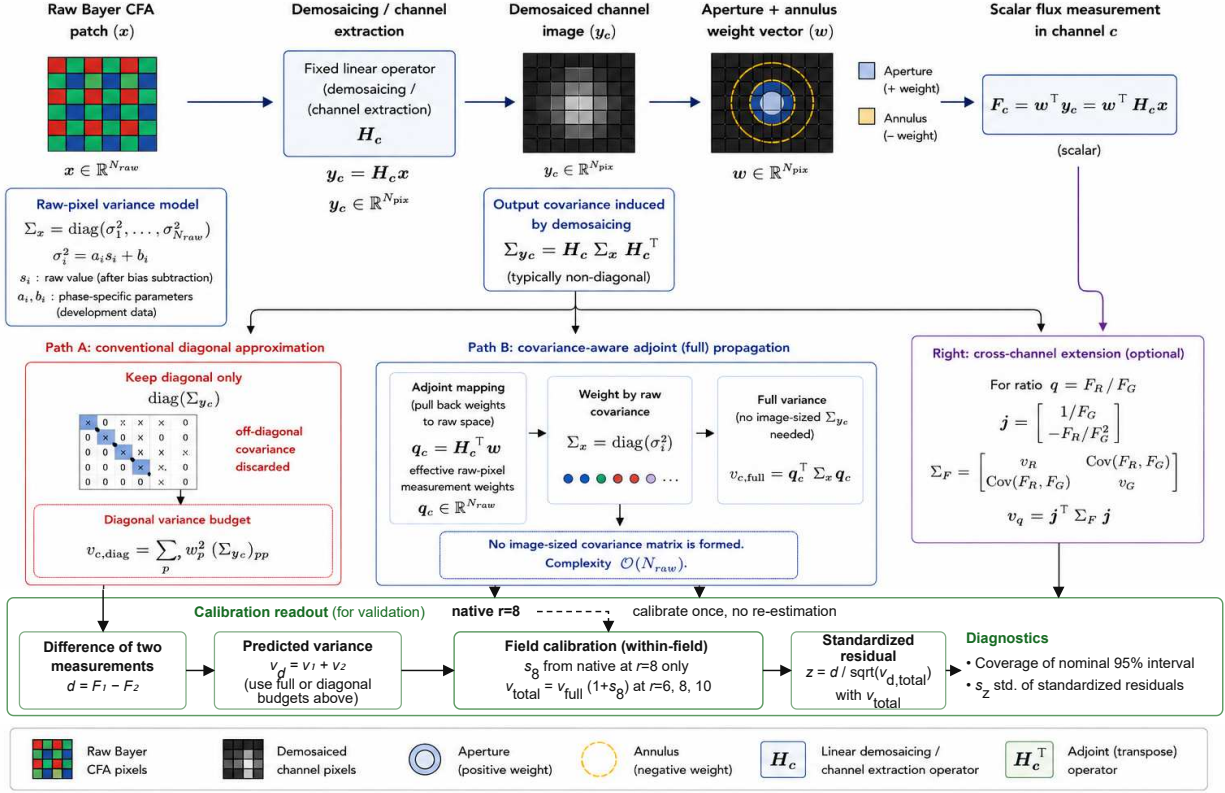


Figure 2: Two-component uncertainty framework for demosaiced Bayer imaging. The adjoint maps aperture weights to raw space through $q_c = H_c^T w$ and evaluates $v_{c,\text{full}} = q_c^T \Sigma_x q_c$. The multichannel branch includes covariance between channel measurements. The lower branch applies the native radius-8 field factor to full budgets at radii 6, 8, and 10 for paired-difference coverage and dispersion. The factor is reused across routes and radii for point estimates and refitted within bootstrap replicates. Fixed finite-support kernels give linear computational cost.

The native branch characterizes the noise model before demosaicing and provides the reference for the paired comparison of diagonal and full budgets.

III.C Evaluation Protocol and Exclusion Rule

The initial registration pass stopped at frame 99, whose RMS was 1.1325 against the fixed 0.8 limit. Before any photometric endpoint was produced, the handling rule was amended to exclude an entire predefined pair if either frame failed screening. Pair 32 (frames 98 and 99) was therefore excluded without replacement. The retained endpoint contains 38 of 39 frame pairs, 76 frames, 52 stars, and 1,976 star-pair observations (Supplementary Note S4).

III.D Controlled Simulation (M20)

The conditional simulation used a local ASI2600MC Pro M20 dataset with 15 lights, 15 matching 180-s darks, 15 flats, and 15 matching 3-s dark-flats. One calibrated light defined the fixed scene, from which 32 isolated sources were retained. Using the measured gain of $0.2548 \text{ e}^-/\text{ADU}$ and read noise of 4.193 ADU, we generated 400 frames with independent Poisson and Gaussian draws. These frames formed 200 nonoverlapping pairs. Master-dark and master-flat estimation variance was held fixed. The simulation isolates the propagation calculation from atmospheric variation and tracking errors.

III.E External Validation Protocol (M45)

The M45 sequence was acquired on 2022 November 13 and is available through MILAN [21]. We verified checksums for the archive and all 147 extracted BGGR FITS files. Each 2080×3072 frame has a 10 s exposure. M45 uses the same Stellina and IMX178 family as M67, providing a second field and date on the same instrument family.

Frames 1 to 35 were for development, frames 36 to 147 for a single evaluation arm of 56 predefined adjacent pairs. Development fixed a 34-star catalog, phase-specific scaling, registration, and geometry. The primary endpoint stayed scalar G at radius 8. Coverage and dispersion gates, the 30-pair minimum, and whole-pair exclusion after a registration failure were predeclared before the evaluation run. Radii 6 and 10 and the R/G quantities were secondary endpoints. Supplementary Note S4 gives additional protocol details.

III.F Statistical Analysis

We calculated 95% cluster-bootstrap intervals with 2,000 replicates, resampling frame pairs and retaining all stars within each sampled pair. M67 has 38 pair clusters with 1,976 star-pair observations at the primary endpoint, and M45 has 43 retained clusters. Diagonal and full budgets were compared within the same draws. Seeds were 20260901 for M67/M45 and transfer, and 20260905 for Canon adjoint-only intervals.

For Table II and Supplementary Table S1, each replicate re-estimated s_8 from sampled native radius-8 records. The same pair-ID draw selected both demosaicing routes at the target radius, preserving route-specific missingness. Separate draws at each radius produce marginal intervals. Pair counts remain the cluster counts when routes are pooled. The bootstrap conditions on the catalog, development calibration, registration, frame scales, and retained sample, and assumes exchangeable pairs. Supplementary Note S4 gives implementation details.

IV Experimental Results

IV.A Validation on Simulated Data (M20)

At the M20 8-pixel endpoint the native and both full budgets sat near 0.95 coverage and unit dispersion (Table V and Fig. 4A), while diagonal-budget dispersion reached 2.6062 for bilinear and 1.6108 for MHC. With the scene and raw-noise model fixed, the comparison isolates the effect of omitting covariance.

Table V: M20 scalar-G result at the primary 8-pixel radius. $N = 200$ frame pairs and 6,400 star-pair observations.

Route	Budget	Coverage	s_z	Emp./pred.
Native CFA	Native	0.9539	0.9905	0.9829
Bilinear	Diagonal	0.5520	2.6062	6.8108
Bilinear	Full	0.9530	0.9902	0.9825
MHC	Diagonal	0.7723	1.6108	2.6012
MHC	Full	0.9525	0.9907	0.9837

The full budgets also met both criteria at radii 6 and 10. Across the three radii, bilinear and MHC coverage was 0.9508–0.9536, with $s_z = 0.9876$ –0.9954; native measurements gave comparable results.

IV.B Primary Field Results (M67)

IV.B.1 Primary Endpoint

Native measurements met both predeclared criteria at radius 8, with coverage 0.9519 and $s_z = 0.9786$ (Table VI and Fig. 3A–B). Bilinear and MHC used the same development phase scales and per-frame normalization.

For bilinear demosaicing the diagonal interval covered 64.17% of residuals at $s_z = 2.5310$; full propagation left the measured fluxes untouched and moved coverage to 95.34% with $s_z = 0.9570$. For MHC, coverage increased from 87.90% to 95.45%, while s_z decreased from 1.4812 to 0.9682. Both operators met the gates once their spatial covariance entered the budget.

Paired against the same fluxes, full propagation cut absolute coverage error by 0.3049 [0.2699, 0.3320] for bilinear and 0.0665 [0.0392, 0.0865] for MHC. These intervals use frame pairs as the resampling unit (Fig. 3C).

Table VI: M67 scalar-G result at the predeclared 8-pixel radius. CIs are 95% cluster-bootstrap intervals over 38 frame pairs; each row contains 1,976 star-pair observations. A pass needs coverage 0.92–0.98 and $s_z = 0.90$ –1.10.

Route	Budget	Coverage [95% CI]	s_z [95% CI]	Emp./pred.	Gate
Native CFA	Native	0.9519 [0.9403, 0.9626]	0.9786 [0.9162, 1.0407]	0.8388	Pass
Bilinear	Diagonal	0.6417 [0.6134, 0.6690]	2.5310 [2.3695, 2.6923]	5.7011	Fail
Bilinear	Full	0.9534 [0.9413, 0.9646]	0.9570 [0.8957, 1.0183]	0.8106	Pass
MHC	Diagonal	0.8790 [0.8598, 0.8978]	1.4812 [1.3836, 1.5771]	1.9234	Fail
MHC	Full	0.9545 [0.9433, 0.9651]	0.9682 [0.9046, 1.0309]	0.8211	Pass

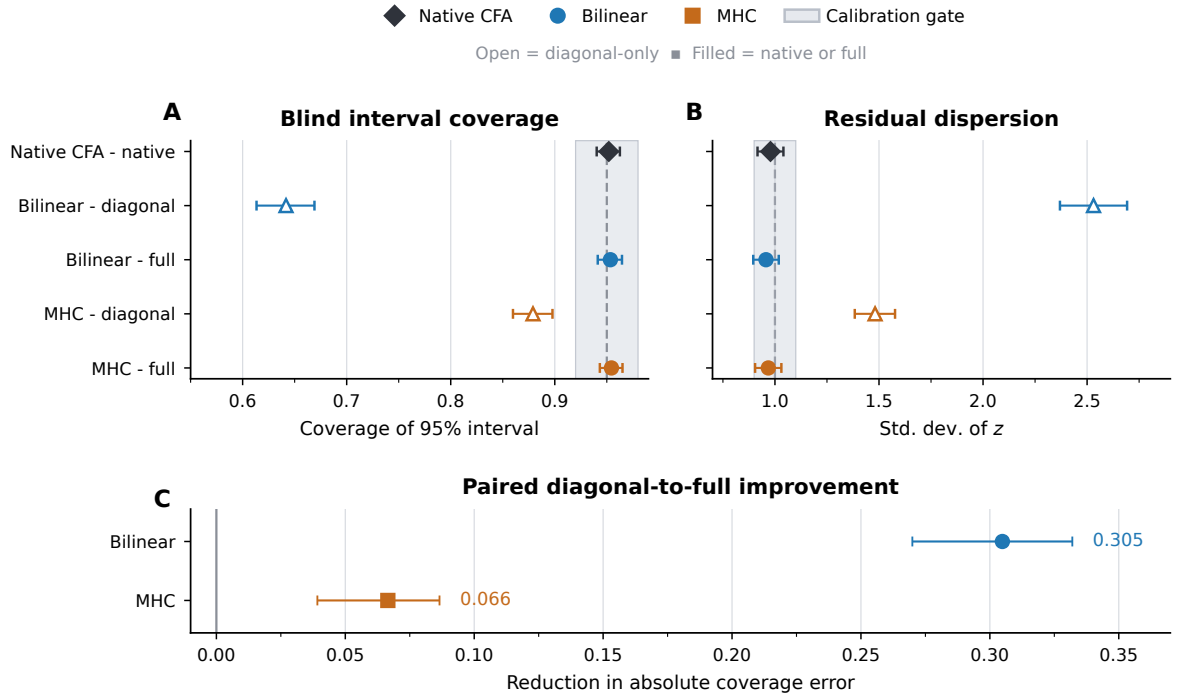


Figure 3: M67 scalar-G calibration at the predeclared 8-pixel radius. (A) Coverage of the nominal 95% interval. (B) Sample standard deviation of standardized residuals. Gray regions are the predeclared calibration gates; vertical dashed lines are the nominal values. Points and bars in A and B are estimates and 95% cluster-bootstrap intervals from 38 frame-pair clusters containing 1,976 star-pair observations. Open triangles denote diagonal-only budgets; filled symbols denote native or full budgets. (C) Paired reduction in absolute coverage error when the full budget replaces the diagonal budget. Bars are paired 95% cluster-bootstrap intervals from 2,000 replicates with seed 20260901; the frame pair is the resampling unit.

IV.B.2 Sensitivity to Aperture Radius

Radii 6 and 10 extend the comparison beyond the primary endpoint (Fig. 4B and Fig. 5B). At radius 6, s_z was 1.1874 for native, 1.2193 for full bilinear, and 1.1810 for full MHC. The corresponding values at radius 10 were 1.2067, 1.1863, and 1.2091. Coverage ranged from 0.9008 to 0.9200. Full propagation improved the demosaiced budgets, but both neighboring radii remained outside the joint criteria. Radius 8 was the predeclared primary endpoint.

Exploratory analysis used stored measurements from the 38 retained pairs without reopening the FITS files. Relative to radius 8, mean pair-level RMS increased by 0.1521–0.1984 at radius 6 and 0.2107–0.2208 at radius 10 across native and full routes. All 95% cluster-bootstrap intervals for these increases excluded zero. At radius 6, Spearman correlations were 0.806–0.827 with the aperture-growth proxy and 0.530–0.559 with registration RMS. Correlations with source-brightness quartile were near zero. The associations link aperture sensitivity to image-growth and registration diagnostics.

IV.C External Field Validation (M45)

M45 provides a second field on the same instrument family. The retained endpoint contains 43 of 56 frame pairs and the 34-star development catalog (Supplementary Note S4). The adjoint budget substantially closes the diagonal gap in Table VII: the native row sat at 0.8669 coverage with $s_z = 1.8774$. For bilinear, full propagation lifted coverage from 0.5284 to 0.8663 and cut s_z from 4.8440 to 1.8606; for MHC the move was 0.7545 to 0.8656 with s_z 2.6768 to 1.8567. The M45 full budgets end up near the native result, which itself remained outside the gates.

Applying the field factor improved coverage across all three radii (Table II). At radius 8, pooled coverage was 0.9582, while dispersion of 0.8868 fell below the 0.90 lower bound. The radius-6 pooled result met both criteria ($s_z = 0.9005$), with bilinear failing the dispersion criterion ($s_z = 0.8915$) and MHC passing ($s_z = 0.9096$). Both routes passed at radius 10. Supplementary Table S2 reports the individual route results.

Table VII: M45 scalar-G adjoint-only baseline at the predeclared 8-pixel radius. CIs are 95% cluster-bootstrap intervals over 43 retained frame-pair clusters; each row has 1,458 or 1,459 star-pair observations. The two-component transfer is reported separately in Table II.

Route	Budget	Coverage [95% CI]	s_z [95% CI]	Emp./pred.
Native CFA	Native	0.8669 [0.8076, 0.9179]	1.8774 [1.3296, 2.4087]	4.3911
Bilinear	Diagonal	0.5284 [0.4784, 0.5763]	4.8440 [3.4458, 6.1895]	29.2030
Bilinear	Full	0.8663 [0.8076, 0.9173]	1.8606 [1.3185, 2.3804]	4.2480
MHC	Diagonal	0.7545 [0.6921, 0.8112]	2.6768 [1.9091, 3.4268]	8.8690
MHC	Full	0.8656 [0.8048, 0.9184]	1.8567 [1.3200, 2.3760]	4.2384

IV.D Cross-Camera Evaluation (Canon 90D)

The Canon EOS 90D NGC 7000 sequence extends the evaluation to a different APS-C Bayer camera family (Supplementary Note S4). The endpoint used 58 common-field stars and 46 frames in 23 retained nonoverlapping pairs from 29 original pairs, yielding 525 star-pair observations at radius 8. Radii 6 and 10 and a source-note sensitivity set with 22 pairs were analyzed separately.

Pairs 7 and 10–14 were excluded before analysis because at least one frame per pair contained fewer than three whole-patch unsaturated stars. No pairs were replaced. After these exclusions, the primary set contained 23 pairs, and the source-note sensitivity set contained 22 pairs.

The Canon results confirm the diagonal-budget error in a second camera family. In Table VIII the adjoint rows move most of the way toward the native row (0.6610 coverage with $s_z = 2.1572$). Bilinear moved coverage from 0.2895 to 0.6514 with s_z 5.7047 down to 2.1562; MHC moved 0.4476 to 0.6552 with s_z 3.5522 down to 2.1548. Median full-to-diagonal variance ratios were 7.009 for bilinear and 2.720 for MHC. With the radius-8 field factor, both Canon routes met the criteria at radii 6 and 10 (Table II).

Table VIII: Canon EOS 90D NGC 7000 scalar-G adjoint-only baseline at the 8-pixel primary radius. CIs are 95% cluster-bootstrap intervals over 23 frame pairs; each row has 525 star-pair observations. The two-component transfer is reported separately in Table II.

Route	Budget	Coverage [95% CI]	s_z [95% CI]	Emp./pred.	Gate
Native CFA	Native	0.6610 [0.6163, 0.6922]	2.1572 [2.0130, 2.3562]	4.606	Fail
Bilinear	Diagonal	0.2895 [0.2520, 0.3210]	5.7047 [5.3109, 6.2424]	32.107	Fail
Bilinear	Full	0.6514 [0.6099, 0.6807]	2.1562 [2.0076, 2.3593]	4.583	Fail
MHC	Diagonal	0.4476 [0.4083, 0.4773]	3.5522 [3.3057, 3.8805]	12.458	Fail
MHC	Full	0.6552 [0.6154, 0.6822]	2.1548 [2.0043, 2.3552]	4.581	Fail

After full propagation, native, bilinear, and MHC dispersions were 2.1572, 2.1562, and 2.1548, respectively, with empirical-to-predicted variance ratios of 4.581–4.606. Dark-frame variance was 1.71–1.81 times the DNG NoiseProfile prediction. The route dispersion spread was 0.0024 for Canon, compared with 0.0207 for M45.

The 22-pair source-note sensitivity set gave coverage/ s_z of 0.667/2.140 for native, 0.657/2.139 for bilinear-full, and 0.661/2.138 for MHC-full. Dispersion decreased by about 0.017–0.018 relative to the 23-pair main set. Leave-one-pair-out s_z remained within 2.121–2.207 across routes. At radii 6–10, the full budgets had empirical-to-predicted variance ratios of approximately 4.3–4.7. The outer detector-radius quartile and a descriptive Spearman correlation near -0.10 showed little radial variation. Supplementary Note S4 gives the additional sensitivity results.

IV.E Secondary Endpoint: R/G Ratio

For the ensemble R/G endpoint, native coverage was 0.9474 with $s_z = 0.8384$ (Table IX and Fig. 4C). Both full budgets reached coverage 0.9737, with $s_z = 0.8209$ for bilinear and 0.8147 for MHC. Coverage met the criterion, while dispersion remained below its lower bound.

Table IX: M67 ensemble R/G endpoint at 8 pixels. Each row has 38 frame-pair observations.

Route	Budget	Coverage [95% CI]	s_z [95% CI]	Emp./pred.
Native CFA	Native	0.9474 [0.8684, 1.0000]	0.8384 [0.6201, 1.0307]	0.6859
Bilinear	Diagonal	0.6842 [0.5263, 0.8158]	2.1752 [1.6074, 2.6672]	4.6115
Bilinear	Full	0.9737 [0.9211, 1.0000]	0.8209 [0.6067, 1.0066]	0.6572
MHC	Diagonal	0.7632 [0.6316, 0.8947]	1.5364 [1.1340, 1.8834]	2.3011
MHC	Spatial only	0.9737 [0.9211, 1.0000]	0.8115 [0.5992, 0.9948]	0.6425
MHC	Full	0.9737 [0.9211, 1.0000]	0.8147 [0.6017, 0.9986]	0.6477

Adding the MHC R–G term on top of spatial covariance shifted predicted ratio variance by -0.804% on M67 and -1.194% in M20 (Fig. 4D): nonzero, but second order next to the spatial part at this endpoint. Bilinear’s R–G covariance is zero under the raw-independence model.

On the ratio the coverage-error cut was 0.2421 [0.0842, 0.3947] for bilinear and 0.1632 [0.0316, 0.3158] for MHC.

V Discussion

V.A Covariance correction in aperture photometry

The paired comparison shows that an accurate pixel-variance map alone is insufficient after demosaicing. Output pixels share raw samples, and the aperture sum accumulates their covariance. The adjoint weights collect each sample’s complete contribution before evaluating variance. This recovers the uncertainty of the specified measurement at linear cost for the fixed finite-support kernels used here.

M67 and conditional M20 results demonstrate the practical effect: full propagation restored near-nominal coverage and dispersion at the primary scalar-G endpoint while preserving the measured fluxes. The larger correction for our bilinear route follows its two-green-plane averaging operator. Kernel definitions therefore matter for interpreting variance inflation, even when reconstruction methods share the same name.

Spatial covariance dominated the demosaicing correction in these tests. Bilinear G had the larger diagonal-budget error despite having zero R–G covariance under the model. For MHC, adding the cross-channel term after spatial

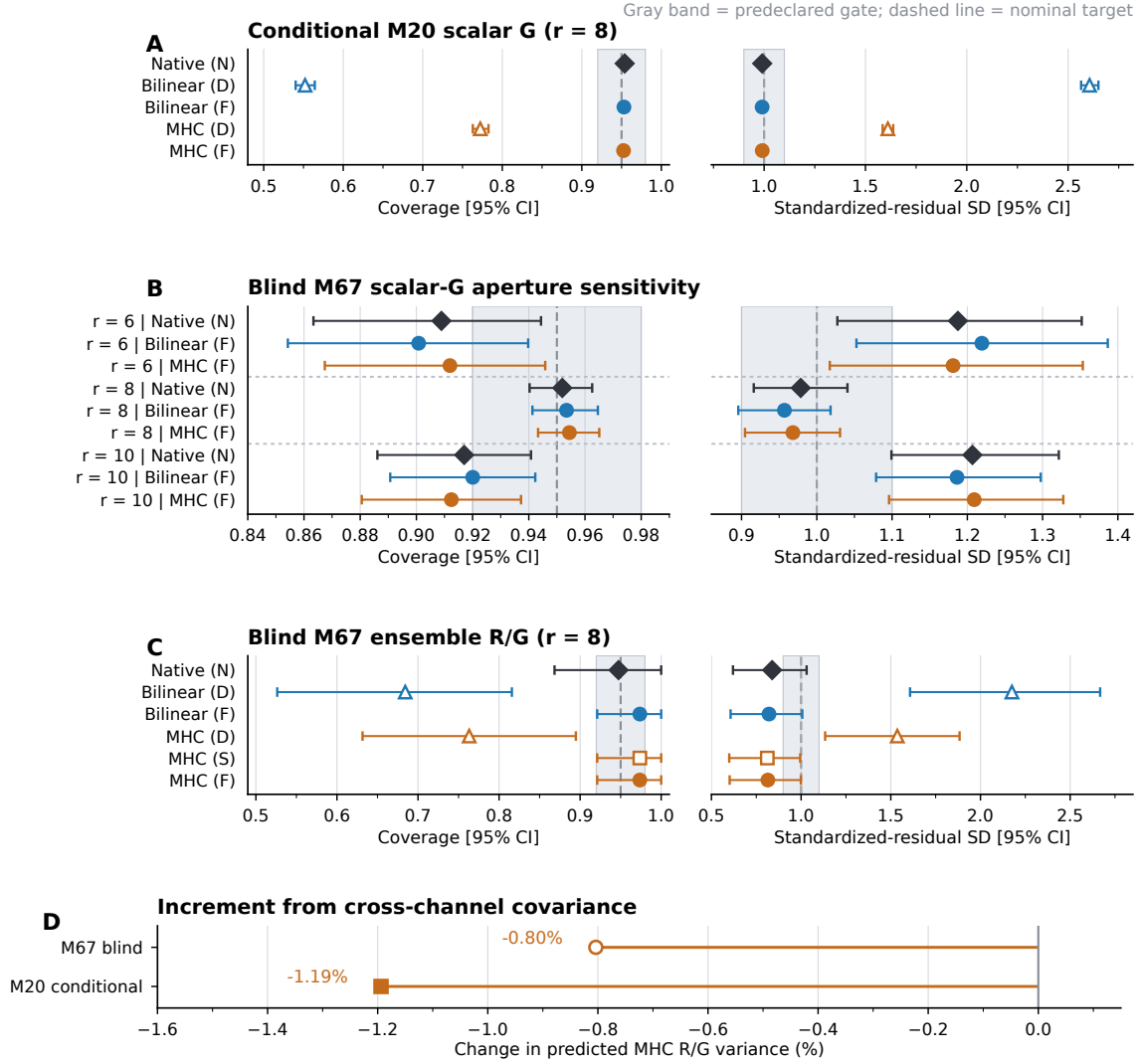


Figure 4: Calibration results outside the primary M67 scalar-G endpoint. The aligned forest plots show coverage (left) and standardized-residual standard deviation (right) on separate axes. Gray bands mark the predeclared coverage (0.92–0.98) and dispersion (0.90–1.10) gates; dashed lines mark the nominal values 0.95 and 1.00. (A) M20 scalar-G results at radius 8 from 200 simulated frame-pair clusters and 6,400 star-pair observations. (B) M67 native and full scalar-G budgets at radii 6, 8, and 10 from 38 frame-pair clusters and 1,976 star-pair observations per condition; only radius 8 enters both gates. (C) M67 ensemble R/G at radius 8 with one observation per retained frame pair ($n = 38$); native and full budgets do not meet the dispersion gate. N, D, S, and F in the row labels denote native, diagonal-only, spatial-only, and full budgets. Horizontal bars in A–C are 95% frame-pair bootstrap intervals. (D) Aggregate change in predicted MHC R/G variance after adding cross-channel covariance for M67 ($n = 38$ pairs) and conditional M20 ($n = 200$ pairs). Panel D reports the aggregate variance contrast.

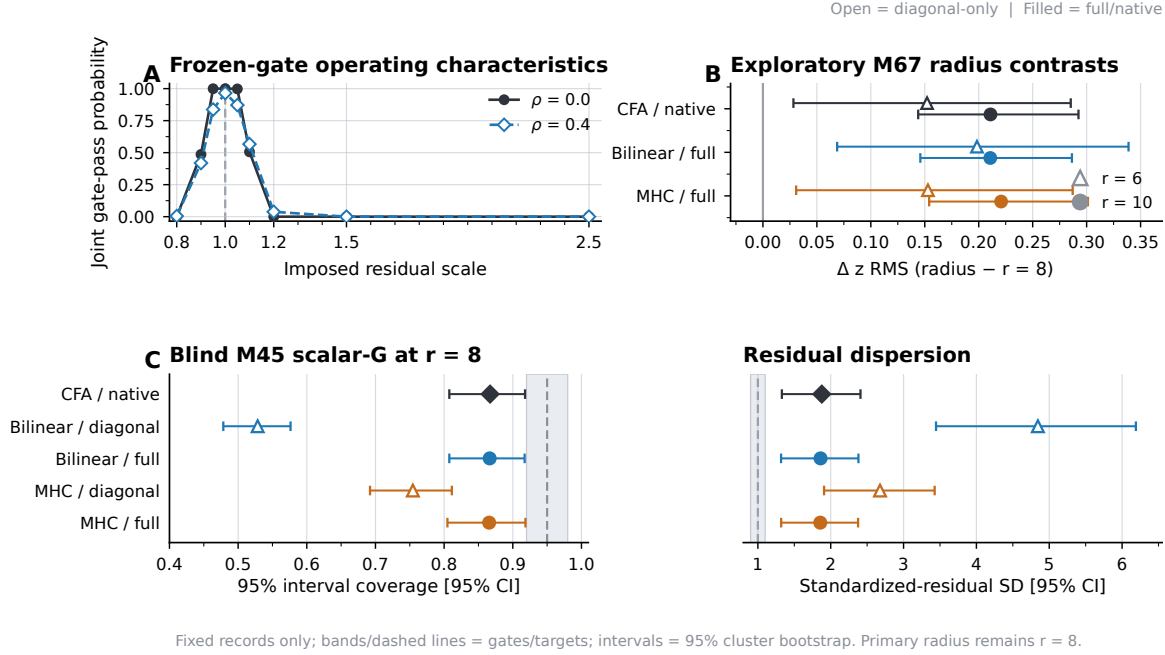


Figure 5: Operating-characteristic simulation, aperture sensitivity, and external-field results. (A) Joint gate-pass probability under the clustered-normal operating-characteristic sensitivity model for the M67 sampling dimensions. Each scale–correlation cell has 10,000 replicates with seed 20260902. (B) Exploratory paired changes in M67 pair-level standardized-residual RMS at radii 6 and 10 relative to radius 8. Points are means and bars are 95% cluster-bootstrap intervals over 38 retained frame pairs. (C) M45 scalar-G coverage and standardized-residual standard deviation at radius 8. Bars are 95% cluster-bootstrap intervals over 43 retained frame pairs; gray bands show the predeclared gates. Thirteen of 56 predefined pairs were excluded by the registration rule. Open symbols denote diagonal-only budgets and filled symbols denote native or full budgets.

propagation changed predicted R/G variance by less than 1.2% in M67 and M20. The scalar-G comparison and ratio analysis thus separate the main spatial contribution from the smaller cross-channel increment.

V.B Residual field variation and aperture transfer

M45 and Canon retained elevated dispersion after exact propagation, with closely aligned native and full results. This pattern is consistent with residual field effects shared across routes and motivates a second component estimated from native measurements. The field factor separates an empirical repeatability scale from the known reconstruction covariance.

Using one radius-8 factor across routes and radii brought both Canon routes within the criteria at radii 6 and 10. M45 passed at radius 10, while its radius-6 outcome depended on the route. M67 needed no positive factor at radius 8 but retained excess dispersion at neighboring radii. These results show where a single field factor transfers successfully and where aperture dependence remains.

V.C Limitations

The field factor was fitted and assessed on the same retained frame pairs. The reported transfer therefore describes within-field behavior; independent data are needed to assess its predictive calibration. Its remaining aperture dependence is visible in M67 and M45. The factor summarizes residual variation rather than identifying individual atmospheric, registration, or calibration contributions.

The evidence concerns relative-flux repeatability for isolated, unsaturated sources under fixed linear demosaicing. Adjacent-frame zero differences do not assess absolute photometric accuracy or shared flux bias. Propagation is

conditional on the raw-noise laws and calibration inputs, and pair-bootstrap intervals assume exchangeable pairs. These conditions define the present tests and the scope of further validation.

VI Conclusion

We developed a two-component uncertainty budget combining exact adjoint covariance propagation with a native-derived field factor. At the predeclared M67 scalar-G endpoint, full propagation increased bilinear coverage from 0.6417 to 0.9534 and MHC coverage from 0.8790 to 0.9545, with standardized-residual dispersion near one. Conditional M20 simulations reproduced the covariance correction.

The native-derived component supplied an empirical scale for residual field variation, and both Canon routes met the calibration criteria at radii 6 and 10 using the radius-8 factor. Together, these results provide an efficient way to propagate demosaicing covariance and evaluate field-level uncertainty in OSC aperture photometry.

Data and Code Availability

M67 and M45 raw sequences are publicly available from the MILAN Zenodo records [20, 21]. This release contains the manuscript and supplementary material. Analysis code and the remaining research data are not publicly released.

Acknowledgment

We thank the MILAN Sky Survey team for publishing the M67 and M45 raw sequences and the operators of the calibration archives that supported method development. We thank the observers who made the Canon 90D NGC 7000 frames available for this evaluation.

References

- [1] S. Shirmohammadi and A. Ferrero, “Camera as the instrument: The rising trend of vision based measurement,” *IEEE Instrum. Meas. Mag.*, vol. 17, no. 3, pp. 41–47, Jun. 2014, doi: 10.1109/MIM.2014.6825388.
- [2] American Association of Variable Star Observers, *AAVSO Guide to CCD Photometry*, Cambridge, MA, USA: AAVSO. [Online]. Available: <https://www.aavso.org/observing-manuals>
- [3] L. Bradley *et al.*, “photutils: Photometry tools,” version 3.0.0, Astropy Project, Apr. 17, 2026. [Online]. Available: https://photutils.readthedocs.io/en/stable/api/photutils.aperture.aperture_photometry.html. Accessed Oct. 8, 2026.
- [4] C. Richard *et al.*, “Siril: Astronomical image processing tool,” version 1.4.4, Free-Astro, Jun. 17, 2026. [Online]. Available: <https://siril.readthedocs.io/en/stable/photometry/quickphotometry.html>. Accessed Oct. 8, 2026, doi: 10.21105/joss.07242.
- [5] B. E. Bayer, “Color imaging array,” U.S. Patent 3,971,065, 1976.
- [6] American Association of Variable Star Observers, *AAVSO DSLR Observing Manual*, version 1.4, Cambridge, MA, USA: AAVSO, 2018. [Online]. Available: <https://www.aavso.org/dslr-camera-photometry-guide>
- [7] A. S. Fruchter and R. N. Hook, “Drizzle: A method for the linear reconstruction of undersampled images,” *Publ. Astron. Soc. Pacific*, vol. 114, no. 792, pp. 144–152, 2002, doi: 10.1086/338393.
- [8] J. Takamatsu, Y. Matsushita, T. Ogasawara, and K. Ikeuchi, “Estimating demosaicing algorithms using image noise variance,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2010, pp. 279–286, doi: 10.1109/CVPR.2010.5540200.
- [9] M. Lee, S. Park, and M. Kang, “Denoising algorithm for CFA image sensors considering inter-channel correlation,” *Sensors*, vol. 17, no. 6, p. 1236, 2017, doi: 10.3390/s17061236.
- [10] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim, “A holistic approach to cross-channel image noise modeling and its application to image denoising,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 1683–1691.
- [11] H. Yang, Y. Lai, X. Liu, M. A. Davies, J. R. Willmott, Y. Zhang, and J. Yang, “Experimental and simulated study of the relationship between color camera imaging and color-modeled equivalence ratio measurement,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023, doi: 10.1109/TIM.2023.3267359.

- [12] European Machine Vision Association, “EMVA Standard 1288: Standard for characterization of image sensors and cameras,” Release 4.0 Linear, 2021. [Online]. Available: https://www.emva.org/wp-content/uploads/EMVA1288Linear_4.0Release.pdf
- [13] European Machine Vision Association, “EMVA Standard 1288: Standard for characterization of image sensors and cameras,” Release 4.0 General, 2021. [Online]. Available: https://www.emva.org/wp-content/uploads/EMVA1288General_4.0Release.pdf
- [14] P. J. Withagen, F. C. A. Groen, and K. Schutte, “CCD color camera characterization for image measurements,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 1, pp. 199–203, 2007, doi: 10.1109/TIM.2006.887667.
- [15] F. A. Baqai, A. Matsui, and K. Nishio, “Method of and apparatus for analyzing noise in a signal processing system,” U.S. Patent 8,108,211, Jan. 31, 2012 (application US20080240203A1, Oct. 2, 2008).
- [16] K. Penev *et al.*, “Tools for high precision photometry from wide-field color images,” *Astron. J.*, vol. 170, no. 3, p. 191, 2025, doi: 10.3847/1538-3881/adf640, arXiv:2507.14289v2.
- [17] Joint Committee for Guides in Metrology, “Evaluation of measurement data—Guide to the expression of uncertainty in measurement,” JCGM 100:2008, 2008.
- [18] H. S. Malvar, L.-W. He, and R. Cutler, “High-quality linear interpolation for demosaicing of Bayer-patterned color images,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2004, pp. 485–488.
- [19] O. Parisot, P. Hitzelberger, P. Bruneau, G. Krebs, C. Destruel, and B. Vandame, “MILAN Sky Survey, a dataset of raw deep sky images captured during one year with a Stellina automated telescope,” *Data in Brief*, vol. 48, Art. no. 109133, 2023, doi: 10.1016/j.dib.2023.109133.
- [20] O. Parisot, “MILAN project: Raw images of deep sky objects captured with a Stellina observation station—[2022-03],” Zenodo, Jul. 2022, doi: 10.5281/zenodo.6865830.
- [21] O. Parisot, “MILAN project: Raw images of deep sky objects captured with a Stellina observation station—[2022-11],” Zenodo, Dec. 2022, doi: 10.5281/zenodo.7399412.

Supplementary Material

Two-Component Covariance-Aware Uncertainty Calibration for Demosaiced One-Shot-Color Aperture Photometry

This supplement provides pooled and per-route field-factor results (Tables S1–S2), native factors by radius (Table S3), and observed full-to-diagonal variance ratios (Table S4). Section S3 derives the phase-resolved uniform-noise identity. Section S4 gives the evaluation protocols, implemented operators, variance maps, and bootstrap details.

S1 Pooled transfer ledger

Table S1 reports the pooled transfer result at each field and radius: bilinear-full and MHC-full residuals pooled, with the field factor s_8 calibrated on native separation at radius 8 only. Intervals are 95% cluster bootstrap over frame pairs with $B = 2000$ replicates. The pooled observation count is the sum over both demosaicing routes. Per-route point values (no intervals) and gates follow in Table S2. Each bootstrap replicate refits s_8 on its sampled native radius-8 records and applies it to the sampled target records using the same pair IDs and existing missingness. Draws are separate between radii. Catalog, development calibration, registration, and normalization remain fixed.

Table S1: Pooled two-component transfer ledger (supports main Table II). s_8 is calibrated once on native at radius 8 and applied to radii 6, 8, and 10 on bilinear-full and MHC-full pooled. Intervals are 95% cluster bootstrap over frame pairs. Gates: coverage 0.92–0.98 and dispersion 0.90–1.10. The factor is refitted in each bootstrap draw.

Set	r	s_8	Coverage [95% CI]	s_z [95% CI]	Gate
M67	6	0	0.9064 [0.8601, 0.9431]	1.2002 [1.0273, 1.3733]	Fail
M67	8	0	0.9539 [0.9423, 0.9643]	0.9625 [0.9008, 1.0255]	Pass
M67	10	0	0.9162 [0.8866, 0.9398]	1.1976 [1.0925, 1.3137]	Fail
M45	6	3.39	0.9583 [0.9347, 0.9692]	0.9005 [0.7700, 1.1571]	Pass
M45	8	3.39	0.9582 [0.9458, 0.9692]	0.8868 [0.8270, 0.9741]	Fail
M45	10	3.39	0.9502 [0.9335, 0.9630]	1.0253 [0.8865, 1.1804]	Pass
Canon	6	3.61	0.9381 [0.9235, 0.9529]	1.0204 [1.0027, 1.0398]	Pass
Canon	8	3.61	0.9410 [0.9304, 0.9549]	1.0039 [0.9869, 1.0214]	Pass
Canon	10	3.61	0.9505 [0.9385, 0.9629]	0.9735 [0.9457, 0.9992]	Pass
M20	6	0	0.9510 [0.9462, 0.9564]	0.9947 [0.9788, 1.0064]	Pass
M20	8	0	0.9527 [0.9488, 0.9570]	0.9904 [0.9756, 0.9999]	Pass
M20	10	0	0.9534 [0.9488, 0.9580]	0.9876 [0.9714, 0.9993]	Pass

Table S2: Per-route transfer values (point estimates, no intervals). Each row is one demosaicing route at the same s_8 as the pooled row in Table S1. Observation counts are per route. Gate uses the original coverage 0.92–0.98 and dispersion 0.90–1.10 criteria.

Set	r	Route	n	Coverage	s_z	Gate
M67	6	bilinear	1976	0.9008	1.2193	Fail
M67	6	MHC	1976	0.9119	1.1810	Fail
M67	8	bilinear	1976	0.9534	0.9570	Pass
M67	8	MHC	1976	0.9545	0.9682	Pass
M67	10	bilinear	1976	0.9200	1.1863	Fail
M67	10	MHC	1976	0.9124	1.2091	Fail
M45	6	bilinear	1462	0.9583	0.8915	Fail
M45	6	MHC	1461	0.9582	0.9096	Pass
M45	8	bilinear	1459	0.9589	0.8879	Fail
M45	8	MHC	1458	0.9575	0.8861	Fail
M45	10	bilinear	1455	0.9498	1.0324	Pass
M45	10	MHC	1454	0.9505	1.0184	Pass
Canon	6	bilinear	525	0.9371	1.0224	Pass
Canon	6	MHC	525	0.9390	1.0193	Pass
Canon	8	bilinear	525	0.9410	1.0047	Pass
Canon	8	MHC	525	0.9410	1.0041	Pass
Canon	10	bilinear	525	0.9505	0.9751	Pass
Canon	10	MHC	525	0.9505	0.9728	Pass
M20	6	bilinear	6400	0.9508	0.9942	Pass
M20	6	MHC	6400	0.9513	0.9954	Pass
M20	8	bilinear	6400	0.9530	0.9902	Pass
M20	8	MHC	6400	0.9525	0.9907	Pass
M20	10	bilinear	6400	0.9536	0.9877	Pass
M20	10	MHC	6400	0.9531	0.9876	Pass

Pooled observation counts are 3952 (M67), 2923/2917/2909 (M45 at radii 6/8/10), 1050 (Canon), and 12800 (M20). Frame-pair counts are 38 (M67), 43 (M45), 23 (Canon), and 200 (M20). The per-route split is even except at M45, where one star-pair observation is unpaired across routes at each radius (1462/1461, 1459/1458, 1455/1454). At M45 radius 6, the pooled dispersion 0.9005 passes, whereas bilinear 0.8915 fails and MHC 0.9096 passes.

S2 Per-radius native factors

Table S3 lists the native empirical-to-predicted variance ratio minus one at each radius, before truncation. The transfer field factor s_8 is the radius-8 value of each field truncated at zero. M67 truncates a negative radius-8 value (−0.1612) to zero; M20 truncates small negative values at all three radii.

Table S3: Native empirical-to-predicted variance ratio minus one at each radius (supports main Table III). Values precede truncation; the radius-8 value of each field truncated at zero gives the transfer factor s_8 .

	$r = 6$	$r = 8$	$r = 10$
M67	0.6382	−0.1612	0.3692
M45	2.9807	3.3911	5.5938
Canon	3.6414	3.6055	3.2885
M20	−0.0081	−0.0171	−0.0228

S3 Variance ratios and phase-resolved propagation

Table S4 summarizes observed pair-budget ratios $R = v_{\text{full}}/v_{\text{diagonal}}$. Diagonal and full rows are matched by field, pair ID, star ID, route, radius, and scalar-G quantity. Each variance sums the predictions for the same two frames. The median and 2.5th/97.5th percentiles describe variation across retained star-pair observations. At radius 8, bilinear medians range from 6.854 to 7.025 across fields, while MHC medians range from 2.062 to 2.720.

For a reproducible uniform-noise identity, write a pixel as $p = 2C + \alpha$ with phase $\alpha \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$, and let $t_{\alpha\ell}(a)$ be its tap at coarse offset a in raw phase ℓ . Away from boundaries, define

$$K_{\alpha\beta}(s) = \sum_{\ell, a, b: a-b=s} t_{\alpha\ell}(a)t_{\beta\ell}(b), \quad A_{\alpha\beta}(s) = \sum_C w_{2C+\alpha}w_{2(C+s)+\beta}.$$

For independent raw samples of equal variance across all phases,

$$\Phi_{\text{uniform}} = \frac{\sum_{\alpha, \beta, s} K_{\alpha\beta}(s)A_{\alpha\beta}(s)}{\sum_{\alpha, C} w_{2C+\alpha}^2 K_{\alpha\alpha}(0)}.$$

For spatially varying variances, the general expression in the main text applies. We checked the phase sum with $\sum_i q_i^2 / \sum_i D_i$, where $D_i = \sum_p w_p^2 H_{pi}^2$, using the actual kernel tables, a 48×48 patch centered at $(23.5, 23.5)$ (zero-based raw coordinates), annulus 13–19, eightfold sampling, and equal unit raw variances. The supports stay away from the patch edges. The maximum relative error $|\Phi_{\text{phase}} - \Phi_{\text{adjoint}}|/\Phi_{\text{adjoint}}$ over the six route/radius combinations was 1.43×10^{-13} . The ratios at radii 6/8/10 were 6.39780/6.54464/6.60153 for bilinear and 2.42639/2.45453/2.46415 for MHC.

Table S4: Observed full-to-diagonal pair-variance ratios: median and 2.5th/97.5th percentiles across matched scalar-G star-pair records.

Set	r	Route	n	Median	p2.5–p97.5
M67	6	bilinear	1976	6.949	6.848–7.029
M67	6	MHC	1976	2.353	2.275–2.488
M67	8	bilinear	1976	7.025	6.940–7.071
M67	8	MHC	1976	2.324	2.227–2.463
M67	10	bilinear	1976	7.020	6.915–7.080
M67	10	MHC	1976	2.280	2.142–2.419
M45	6	bilinear	1462	6.867	6.672–6.979
M45	6	MHC	1461	2.110	1.970–2.241
M45	8	bilinear	1459	6.854	6.692–6.988
M45	8	MHC	1458	2.062	1.954–2.201
M45	10	bilinear	1455	6.820	6.690–6.959
M45	10	MHC	1454	2.023	1.936–2.154
Canon	6	bilinear	525	7.013	6.959–7.036
Canon	6	MHC	525	2.724	2.695–2.742
Canon	8	bilinear	525	7.009	6.967–7.028
Canon	8	MHC	525	2.720	2.690–2.737
Canon	10	bilinear	525	6.976	6.930–7.003
Canon	10	MHC	525	2.710	2.684–2.726
M20	6	bilinear	6400	6.971	6.871–7.005
M20	6	MHC	6400	2.656	2.529–2.712
M20	8	bilinear	6400	6.934	6.843–6.980
M20	8	MHC	6400	2.643	2.539–2.694
M20	10	bilinear	6400	6.887	6.809–6.937
M20	10	MHC	6400	2.634	2.550–2.677

S4 Evaluation protocol and implementation details (Supplementary Note S4)

This section provides additional experimental and implementation details referenced as Supplementary Note S4 in the main text.

S4.A M67 endpoint

The M67 analysis uses 114 raw BGGR FITS exposures (frames 1–114) from the public MILAN sequence. Frames 1–35 are the development subset; frames 36–114 are the evaluation subset, grouped into 39 predefined adjacent pairs. The registration check (at least 30 matched detections, RMS at most 0.8 half-resolution pixel) failed for frame 99 (RMS 1.1325). Whole-pair exclusion was adopted after this registration failure and before photometric endpoint output, retaining the original pairing and RMS threshold. Pair 32 (frames 98 and 99) was excluded without replacement. The retained endpoint contains 38 of 39 predefined pairs, 76 frames, 52 stars, and 1976 star-pair observations at radius 8. The catalog holds 52 isolated unsaturated stars with centers at least 24 half-resolution pixels apart. A positive G measurement was kept regardless of its R flux.

S4.B M45 endpoint

The M45 analysis uses 147 extracted BGGR FITS files (2080×3072 , 10 s) from the public MILAN sequence of 2022 November 13. Frames 1–35 are the development subset; frames 36–147 form a single evaluation arm of 56 predefined adjacent pairs. Development fixed a 34-star catalog with the same registration gate (30 matches, RMS at most 0.8 half-resolution pixel). The whole-pair rule excluded 13 of 56 predefined pairs, leaving 43 pairs. The retained radius-8 endpoint contains 1458–1459 star-pair observations per route (2917 pooled). M45 provides a second field and date on the same Stellina/IMX178 instrument family as M67.

S4.C M20 conditional simulation

The M20 arm uses a local ASI2600MC Pro dataset (15 lights, 15 matching 180-s darks, 15 flats, 15 matching 3-s dark-flats). One calibrated light defines the fixed scene with 32 isolated sources retained. With measured gain $0.2548 \text{ e}^-/\text{ADU}$ and read noise 4.193 ADU, 400 frames were generated by independent Poisson and Gaussian draws and grouped into 200 nonoverlapping pairs (6400 star-pair observations per route at radius 8). Master-dark and master-flat estimation variance was held fixed.

S4.D Canon 90D evaluation

The Canon EOS 90D NGC 7000 sequence uses 58 common-field stars and 46 frames in 23 retained nonoverlapping pairs from 29 original pairs (525 star-pair observations per route at radius 8). Pairs 7 and 10–14 were excluded before analysis because at least one frame per pair contained fewer than three whole-patch unsaturated stars; no pairs were replaced. Radii 6 and 10 and the 22-pair source-note sensitivity set were analyzed with the same procedure. Robustness checks (source-note set, leave-one-pair-out with s_z ranges 2.124–2.205 native, 2.122–2.207 bilinear-full, and 2.121–2.206 MHC-full, outer detector-radius quartile, Spearman correlation near -0.10 with detector radius) gave consistent sensitivity results.

S4.E Pair selection

M67 excluded 1 of 39 predefined pairs (pair 32: frames 98–99; frame 99 registration RMS 1.132537 above the 0.8 limit). M45 excluded 13 of 56 predefined pairs after registration screening under the whole-pair rule. Canon excluded pairs 7 and 10–14 for insufficient whole-patch unsaturated stars before analysis. No excluded pair was replaced or repaired in any field.

S4.F Bootstrap implementation

Intervals use 2,000 frame-pair bootstrap replicates, with seed 20260901 for M67/M45 and transfer and 20260905 for Canon adjoint-only results. Transfer draws re-estimate s_8 from native radius-8 measurements and use common pair IDs across routes at the target radius, retaining missingness. The 51st and 1950th ordered values define the transfer percentile endpoints. Draws are separate between radii; the catalog, development calibration, registration, normalization, and exclusions remain fixed. Pair resampling assumes exchangeability, including after ensemble normalization. The operating-characteristic simulation uses 10,000 replicates and seed 20260902.

S4.G Implemented Bayer operators and variance maps

M67/M45 BGGR arrays are cropped starting at raw offset (1, 1) to the largest even RGGB interior. The four raw planes are R, G_r, G_b, B at phases (0, 0), (0, 1), (1, 0), (1, 1); patch origins are even. On the full-resolution lattice, the production bilinear G kernel at red or blue sites uses the four axial green neighbors with weight $1/4$ each. At either green site it uses the sampled green with weight $1/2$ and the four diagonal samples of the other green phase with weight $1/8$ each. Thus it is the average of two separately bilinearly interpolated green planes, with unit DC gain; already sampled G values are not preserved. R and B use ordinary sampled-value, two-neighbor axial, or four-neighbor diagonal bilinear interpolation, as appropriate to phase.

MHC uses fixed 5×5 kernels [18]; its G kernel at a red/blue site has center weight $1/2$, four axial distance-one weights $1/4$, and four axial distance-two weights $-1/8$. At green sites it is the identity. The G kernels differ in both missing-green interpolation and sampled-green treatment. Each route is compared using its own diagonal and full budgets. For the uniform radius-8 test above, sample-preserving textbook bilinear gives $\Phi = 3.04969$, compared with 6.54464 for our implementation; these values are specific to the stated center and geometry.

Native G uses only the two green raw planes, with weights $2(a_i - A_G b_i / B_G)$, where A_G and B_G sum fractional aperture and annulus areas over green samples. Native R/B uses the corresponding single plane with factor 4 and that plane's area ratio. The normalization expresses native fluxes on the full-resolution area scale. Each route's diagonal/full comparison uses identical fluxes.

For M67/M45 the implemented normalized raw variance map in phase ℓ is $V_{\ell u} = a_\ell \max[(x_{\ell u} - m_\ell) / c_f, 0] + b_\ell$. Here m_ℓ is the frame's phase median, c_f is its native-G ensemble scale, and a_ℓ, b_ℓ already include the development-only phase multipliers. Fluxes are divided by c_f ; these maps are used in normalized-flux units. M20 uses the conditional Poisson/read variance $(\max(S_i, 0) / g + \sigma_{\text{read}}^2) / f_i^2$ with fixed dark and normalized-flat masters. For Canon, $V_i = [a_\ell \max(x_i - \beta_\ell, 0) + b_\ell] / (f_i^2 c_f^2)$, where β_ℓ is the stored black level, a_ℓ, b_ℓ are the DNG NoiseProfile phase laws, f_i is the normalized flat, and c_f is the native-G ensemble scale. The fixed master dark is subtracted from flux but its estimation variance is excluded. Canon's native-G reference is the per-star median over retained evaluation frames, whereas M67/M45 use their development reference. These variance maps condition on the fitted laws and calibration masters.

Out-of-range coarse-grid tap coordinates are clipped to the closest sample of the same CFA plane (edge replication, not wraparound). Repeated taps that reach one raw sample are combined before squaring for diagonal variance. Thus $D_i = \sum_p w_p^2 H_{pi}^2$ and $v_{\text{diag}} = \sum_i D_i V_i$, whereas $v_{\text{full}} = \sum_i (\sum_p w_p H_{pi})^2 V_i$. The measured aperture/annulus supports fit within 48×48 patches with the kernel halo, so clipping does not affect their nonzero weights. Registration transforms coordinates only; it does not resample the raw arrays.