

# mvESS-Z: A Robust Multivariate Effective Sample Size for Degenerate and Quantized Markov Chains

Sayan Mukherjee

Chennai Mathematical Institute, Chennai, India

## Abstract

Effective Sample Size (ESS) is central to diagnosing Markov chain Monte Carlo (MCMC) efficiency, yet standard multivariate ESS definitions either collapse in high dimensions when some coordinates are (near-)deterministic or become numerically unstable under hardware approximations (e.g., fixed-point quantization and probability truncation). We introduce *mvESS-Z*, a multivariate ESS that (i) is *invariant to zero-variance coordinates*, (ii) admits *finite-sample concentration guarantees*, and (iii) is *stable* under quantization/truncation commonly used in probabilistic accelerators. mvESS-Z is defined on the signal subspace determined by the posterior covariance and uses pseudo-determinant or spectral aggregation to avoid degeneracy. We prove basic properties, give a practical estimator with automatic rank selection, and show on synthetic and vision-style workloads that mvESS-Z correlates with downstream task error more reliably than per-coordinate ESS and split- $\hat{R}$ , enabling quality-constrained tuning of precision and random number generators.

**Keywords:** multivariate effective sample size (mvESS-Z), MCMC diagnostics, signal subspace, pseudo-determinant, spectral variance, long-run covariance, rank deficiency, degeneracy, quantization, probability truncation, random number generators (RNG), probabilistic accelerators

## 1. Introduction

Bayesian inference and Markov chain Monte Carlo (MCMC) underpin a wide range of statistical and machine learning systems, and are increasingly executed on probabilistic accelerators that trade arithmetic precision and control logic for throughput and energy efficiency. Assessing the quality of such dependent samples is central to scientific validity and hardware design. Classical univariate effective sample size (ESS) aggregates per-coordinate autocorrelation, while multivariate variants compare second-order structure of the chain to that of i.i.d. draws, typically via determinants or traces of the long-run covariance (LRC) matrix [1–5]. Modern workloads, however, routinely induce degenerate or near-degenerate posteriors (through constraints, hierarchical shrinkage, or sparsity), and hardware approximations—fixed-point quantization, probability truncation, and lightweight pseudo-random number generators—can further suppress variance in subsets of coordinates. In these settings determinant-based ESS collapses whenever the sample covariance is singular, and trace-based proxies can be overly optimistic by averaging over null axes, yielding misleading diagnostics and suboptimal hardware choices.

This work targets a multivariate ESS that remains well-defined and interpretable when some coordinates are (near-)deterministic, tracks within-chain efficiency on the active subspace actually explored by the chain, and stays stable under common hardware approximations so that designers can make quality-constrained choices (e.g., precision or RNG) rather than chase raw throughput. We introduce *mvESS-Z*, which evaluates efficiency only on the signal subspace of the sample covariance and aggregates dependence via pseudo-determinant or spectrally averaged ratios of covariance to LRC. By construction, zero-variance coordinates are ignored, avoiding determinant collapse and preventing null axes from diluting efficiency estimates.

We study three questions: (i) whether restricting to the signal subspace yields invariance to zero-variance coordinates and numerical stability under quantization/truncation; (ii) whether finite-sample guarantees can be obtained using consistent multivariate spectral variance estimators with automatic rank selection; and (iii) whether mvESS-Z better predicts downstream task error than common diagnostics (per-coordinate ESS, split- $\hat{R}$ ) across synthetic and application-style workloads, thereby enabling quality-constrained hardware choices.

This paper makes four contributions:

1. **Definition and invariance.** A formalization of mvESS-Z as a geometric-mean ratio of projected covariance and LRC eigenvalues on a data-driven signal subspace, with proofs of zero-variance invariance and reduction to standard ESS when the active dimension is one.
2. **Finite-sample theory.** Concentration after log-transform under geometric ergodicity using multivariate spectral variance estimators, together with an automatic vanishing-eigenvalue threshold for rank selection.
3. **Estimator and practicality.** A numerically stable estimator (log-spectral aggregation, flat-top spectral variance, plug-in bandwidth) whose  $O(d^3)$  eigendecomposition cost is negligible relative to sampling on high-throughput systems.
4. **Empirical validation.** Synthetic degeneracy sweeps and vision-style workloads (stereo/flow/segmentation) with controlled quantization, truncation, and RNG substitutions, showing stronger correlation with end-task error than split- $\hat{R}$  and trace/determinant ESS, and enabling quality-constrained speedup curves for hardware design.

The scope targets ergodic Markov chains with finite fourth moments and focuses on within-chain efficiency; mvESS-Z complements, rather than replaces, cross-chain diagnostics such as rank-normalized split- $\hat{R}$  [5, 6]. It is a diagnostic rather than a certificate against bias under severe approximations; when approximations may alter the target distribution itself, posterior predictive checks and task-aligned evaluations remain necessary.

Across controlled studies, determinant-based multivariate ESS often underestimates efficiency in the presence of rank loss, while trace-based proxies can overestimate it. mvESS-Z remains stable under moderate quantization without truncation and degrades gracefully under aggressive truncation; it exhibits higher rank correlations with task error and supports selecting the lowest-cost hardware setting that meets a user-specified robustness threshold [7].

The remainder of the paper reviews notation and long-run covariance estimation, defines mvESS-Z and its spectral variant, presents invariance and concentration results, details the estimator and rank selection, reports synthetic and application results, and concludes with limitations and opportunities for adaptive-precision controllers.

## 2. Preliminaries

Let  $\{X_t\}_{t=1}^n$  be an ergodic Markov chain with invariant distribution  $\pi$  on  $\mathbb{R}^d$ , mean  $\mu = \mathbb{E}_\pi[X_t]$ , and covariance  $\Sigma = \text{Cov}_\pi(X_t)$ . Define the lag- $k$  auto-covariance  $\Gamma_k = \text{Cov}_\pi(X_t, X_{t+k})$  and the (symmetric) long-run covariance (LRC)

$$\Gamma = \Sigma + \sum_{k=1}^{\infty} (\Gamma_k + \Gamma_k^\top) = \Sigma + 2 \sum_{k=1}^{\infty} \frac{1}{2} ((+\top) \Gamma_k), \quad (1)$$

so that for i.i.d. draws  $\Gamma = \Sigma$ . For any fixed vector  $a \in \mathbb{R}^d$ , the scalar ESS along the linear functional  $a^\top X$  may be written

$$\text{ESS}(a) = n \frac{\text{Var}_\pi(a^\top X)}{\text{Var}_{\text{LRC}}(a^\top X)} = n \frac{a^\top \Sigma a}{a^\top \Gamma a}, \quad (2)$$

and for a coordinate  $j$  this reduces to  $\text{ESS}_j = n \Sigma_{jj} / \Gamma_{jj}$ . Multivariate generalizations compare  $\Sigma$  and  $\Gamma$  via matrix functionals (e.g., determinants or traces) to produce an overall ‘‘independence ratio’’ scaled by  $n$ . When  $\Sigma$  is rank-deficient,  $r = \text{rank}(\Sigma) < d$ , the determinant  $\det(\Sigma) = 0$  and determinant-based ESS becomes undefined; trace-based surrogates can obscure degeneracy by averaging over null axes.

We use the eigen-decomposition  $\Sigma = U\Lambda U^\top$  with  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$  and denote the *signal subspace* by  $\mathcal{S} = \text{span}(U_r)$ , where  $U_r$  collects the eigenvectors corresponding to strictly positive eigenvalues. For any symmetric matrix  $A$ , the restriction to the signal subspace is  $A_{\mathcal{S}} = U_r^\top A U_r$ . The *pseudo-determinant* is  $\text{pdet}(A) = \prod_{\lambda_i(A) > 0} \lambda_i(A)$ , i.e., the product of nonzero eigenvalues. Two population multivariate ‘‘independence ratios’’ that will be useful are

$$\left( \frac{\text{pdet}(\Sigma_{\mathcal{S}})}{\text{pdet}(\Gamma_{\mathcal{S}})} \right)^{1/r} \quad \text{and} \quad \exp \left\{ \frac{1}{r} \sum_{i=1}^r \log \frac{\lambda_i(\Sigma_{\mathcal{S}})}{\lambda_i(\Gamma_{\mathcal{S}})} \right\}, \quad (3)$$

which coincide and are well-defined whenever  $r \geq 1$ . These ratios motivate the mvESS-Z estimator introduced later.

Empirical counterparts replace population quantities by consistent estimators. Let  $\bar{X} = n^{-1} \sum_{t=1}^n X_t$  and  $Y_t = X_t - \bar{X}$ . The sample covariance is

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{t=1}^n Y_t Y_t^\top. \quad (4)$$

The long-run covariance can be estimated via multivariate spectral variance (SV) estimators. Let  $\hat{\Gamma}_k = (n-k)^{-1} \sum_{t=1}^{n-k} Y_t Y_{t+k}^\top$  for  $k \geq 1$ , choose a bandwidth  $b = b_n$  with  $b \rightarrow \infty$  and  $b/n \rightarrow 0$ , and let  $\{w_k\}_{k=0}^{b-1}$  be a taper (e.g., flat-top). A symmetric, positive-semidefinite SV estimator is

$$\hat{\Gamma} = \hat{\Sigma} + \sum_{k=1}^{b-1} w_k (\hat{\Gamma}_k + \hat{\Gamma}_k^\top), \quad (5)$$

with automatic bandwidth selection by plug-in rules. In practice we operate on the centered series  $\{Y_t\}$  after discarding burn-in; thinning is unnecessary and only reduces  $n$ .

Rank deficiency is handled by projecting to the estimated signal subspace. Let  $\hat{\Sigma} = U\Lambda U^\top$  and select a data-driven rank  $\hat{r}$  by thresholding eigenvalues below  $\tau_n$  (a vanishing threshold such as

$\tau_n = \kappa(\widehat{\Sigma})/d \cdot n^{-\alpha}$  with  $\kappa > 0$  and  $0 < \alpha < 1/2$ ). Writing  $U_{\hat{r}}$  for the retained eigenvectors, define the projected matrices  $\widehat{\Sigma}_{\hat{\mathcal{S}}} = U_{\hat{r}}^\top \widehat{\Sigma} U_{\hat{r}}$  and  $\widehat{\Gamma}_{\hat{\mathcal{S}}} = U_{\hat{r}}^\top \widehat{\Gamma} U_{\hat{r}}$ . The corresponding empirical independence ratios mirror (3) with  $\Sigma, \Gamma$  replaced by their projected estimates and  $r$  replaced by  $\hat{r}$ ; taking a geometric mean across eigen-directions maintains scale invariance and numerical stability.

We also formalize the approximation channel used to model hardware effects. Let  $Q_\theta$  denote a mapping parameterized by  $\theta$  (bit-width, quantization step, probability truncation rule, and pseudo-random generator choice) that acts either on the state updates or on the probabilities used within proposals and accept/reject steps, yielding an observed chain  $\{\tilde{X}_t\}$  with covariance  $\tilde{\Sigma}$  and LRC  $\tilde{\Gamma}$ . Typical consequences include shrinkage of small-variance directions (potentially reducing rank), altered acceptance dynamics that inflate serial dependence, and spectral perturbations of the LRC. All estimators above apply verbatim to  $\{\tilde{X}_t\}$ ; a robust multivariate ESS should remain interpretable under these transformations and should not spuriously collapse when variance in some coordinates vanishes due to modeling constraints or hardware approximations.

### 3. mvESS-Z: definition and intuition

We quantify multivariate sampling efficiency only on the *signal subspace* explored by the chain, thereby avoiding pathologies caused by zero- or near-zero-variance coordinates that arise from model structure or hardware approximations. Let  $\Sigma = U\Lambda U^\top$  be the population covariance with eigenvalues ordered as  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$  where  $\lambda_i > 0$ . The signal subspace is  $\mathcal{S} = \text{span}(U_r)$  with  $\dim(\mathcal{S}) = r$ . For any symmetric matrix  $A$ , the restriction to  $\mathcal{S}$  is  $A_{\mathcal{S}} = U_r^\top A U_r$  and the pseudo-determinant is  $\text{pdet}(A) = \prod_{\lambda_i(A) > 0} \lambda_i(A)$ . In practice we work with the empirical counterparts from the previous section:  $\widehat{\Sigma}$  as the sample covariance,  $\widehat{\Gamma}$  as a symmetric spectral-variance (SV) estimator of the long-run covariance (LRC) using an automatic bandwidth and a flat-top taper, and a data-driven estimate  $\widehat{\mathcal{S}}$  obtained by thresholding the eigenvalues of  $\widehat{\Sigma}$  below a vanishing level  $\tau_n$ .

**Definition 1** (mvESS-Z). Given  $n$  dependent samples with sample covariance  $\widehat{\Sigma}$  and SV estimate  $\widehat{\Gamma}$ , let  $\widehat{\mathcal{S}}$  be the span of eigenvectors of  $\widehat{\Sigma}$  with eigenvalues  $\geq \tau_n$ , where  $\tau_n = \kappa(\widehat{\Sigma})/d \cdot n^{-\alpha}$  with constants  $\kappa > 0$  and  $0 < \alpha < \frac{1}{2}$ . Writing  $r = \dim(\widehat{\mathcal{S}}) \geq 1$ , the multivariate effective sample size is

$$\text{mvESS-Z} = n \left( \frac{\text{pdet}(\widehat{\Sigma}_{\widehat{\mathcal{S}}})}{\text{pdet}(\widehat{\Gamma}_{\widehat{\mathcal{S}}})} \right)^{1/r}. \quad (6)$$

The definition admits a directional interpretation that clarifies both invariance and numerical behavior. Let  $\widehat{\Sigma}_{\widehat{\mathcal{S}}} = V \text{diag}(\alpha_1, \dots, \alpha_r) V^\top$  and  $\widehat{\Gamma}_{\widehat{\mathcal{S}}} = V \text{diag}(\beta_1, \dots, \beta_r) V^\top$  in a common eigenbasis  $V$  (possible up to estimation error because both are symmetric positive semidefinite on  $\widehat{\mathcal{S}}$ ). The scalar ESS along eigen-direction  $v_i$  is

$$\text{ESS}_i = n \frac{\text{Var}(v_i^\top X)}{\text{Var}_{\text{LRC}}(v_i^\top X)} = n \frac{\alpha_i}{\beta_i},$$

and therefore (6) equals the geometric mean of these directional ESS values,

$$\text{mvESS-Z} = \left( \prod_{i=1}^r \text{ESS}_i \right)^{1/r}.$$

This aggregation preserves the usual scale interpretation of ESS, averages across the *active* dimensions only, and ignores null axes by design.

A numerically stable, equivalent form avoids products of very small eigenvalues:

$$\text{mvESS-Z} = n \cdot \exp \left\{ \frac{1}{r} \sum_{i=1}^r \log \frac{\lambda_i(\widehat{\Sigma}_{\widehat{\mathcal{S}}})}{\lambda_i(\widehat{\Gamma}_{\widehat{\mathcal{S}}})} \right\}, \quad (7)$$

which we compute via eigenvalues or `logdet` on the projected matrices. An equivalent computational route solves the generalized eigenproblem

$$\widehat{\Sigma}_{\widehat{\mathcal{S}}} u_i = \lambda_i^{(g)} \widehat{\Gamma}_{\widehat{\mathcal{S}}} u_i, \quad i = 1, \dots, r,$$

in which case  $\text{ESS}_i = n \lambda_i^{(g)}$  and  $\text{mvESS-Z} = (\prod_{i=1}^r n \lambda_i^{(g)})^{1/r}$ . This formulation can be more stable when  $\widehat{\Gamma}_{\widehat{\mathcal{S}}}$  is well-conditioned.

**Proposition 1** (Zero-variance invariance [8]). *If additional coordinates with zero variance (exactly or effectively) are appended to the state vector, or if hardware approximations  $Q_\theta$  collapse some coordinates to constants, mvESS-Z is unchanged. Conversely, removing deterministic coordinates leaves mvESS-Z unchanged.*

*Proof.* Deterministic coordinates lie in  $\widehat{\mathcal{S}}^\perp$  by construction, so  $\widehat{\Sigma}_{\widehat{\mathcal{S}}}$  and  $\widehat{\Gamma}_{\widehat{\mathcal{S}}}$  are unaffected by augmenting with or dropping those coordinates. The ratio in (6) is therefore invariant.  $\square$

**Proposition 2** (Reduction and reparameterization [9]). *If  $r = 1$ , then mvESS-Z reduces to the standard univariate ESS along the single active direction. For any orthonormal matrix  $Q$ , replacing  $X_t$  by  $QX_t$  leaves mvESS-Z unchanged (orthonormal invariance). Moreover, rescaling coordinates by a nonsingular diagonal matrix  $D$  leaves mvESS-Z unchanged because the ratio in (6) is computed after projection and uses matched quadratic forms.*

*Remark 1* (Population target and proportionality in  $n$ ). At the population level, the independence ratio

$$\left( \frac{\text{pdet}(\Sigma_{\mathcal{S}})}{\text{pdet}(\Gamma_{\mathcal{S}})} \right)^{1/r}$$

is a constant that depends on the chain's dependence structure. Under the consistency conditions from the previous section and mild mixing assumptions,  $\text{mvESS-Z}/n$  concentrates around this constant, so the diagnostic scales linearly with  $n$  as expected for an effective sample size.

*Remark 2* (Rank selection and edge cases). We select  $\hat{r}$  by thresholding eigenvalues of  $\widehat{\Sigma}$  below  $\tau_n = \kappa(\widehat{\Sigma})/d \cdot n^{-\alpha}$  with  $0 < \alpha < \frac{1}{2}$ ; this vanishing threshold yields  $\mathbb{P}(\widehat{\mathcal{S}} = \mathcal{S}) \rightarrow 1$  under geometric ergodicity and finite fourth moments. If  $\hat{r} = 0$  (all variances fall below  $\tau_n$ ), the chain carries no reliable second-order signal and we report NA, prompting the user to increase  $n$  or relax hardware approximations.

*Remark 3* (Stability under hardware approximations). Quantization, probability truncation, and lightweight RNGs change  $\widehat{\Gamma}$  and can shrink small modes of  $\widehat{\Sigma}$ , sometimes reducing rank. Because mvESS-Z projects to  $\widehat{\mathcal{S}}$  and aggregates direction-wise independence ratios, it *degrades gracefully*: when rank loss occurs, null axes are discarded; when serial dependence increases, the eigen-ratios decrease smoothly, lowering mvESS-Z without numerical collapse.

*Remark 4* (Computation and practice). The estimator requires an eigendecomposition of  $\widehat{\Sigma}$  ( $O(d^3)$ ) and one of the projected matrices ( $O(r^3)$ ). In typical accelerator workflows, this overhead is negligible relative to sample generation. Centering the series, using a symmetric SV estimator with a flat-top taper, and evaluating (7) with `logdet` further improve numerical stability. Clipping `mvESS-Z` at  $n$  is optional; we do not clip by default because negative autocorrelation can legitimately yield  $\text{ESS} > n$  for some directions.

Overall, `mvESS-Z` measures the geometric mean of directional ESS values on the data-driven signal subspace, inherits desirable invariances, and remains well-defined when degeneracy or hardware effects would invalidate determinant-based multivariate ESS or obscure efficiency in trace-based surrogates. This makes it an actionable and robust target for quality-constrained design and evaluation of probabilistic accelerators.

## 4. Properties, estimation, and practical use

This section establishes key properties of `mvESS-Z`, details a stable estimation procedure with recommended defaults, and explains how to use the diagnostic in practice for both methodological analysis and hardware design.

**Proposition 3** (Zero-variance invariance [8, 10]). *If a block of coordinates is deterministic (exactly zero variance) under  $\pi$  or is rendered deterministic by an approximation channel  $Q_\theta$ , then `mvESS-Z` is unchanged by adding or removing those coordinates from  $X_t$ .*

*Proof.* Deterministic coordinates lie in the orthogonal complement of the estimated signal subspace  $\widehat{\mathcal{S}}$ , hence the projected matrices  $\widehat{\Sigma}_{\widehat{\mathcal{S}}}$  and  $\widehat{\Gamma}_{\widehat{\mathcal{S}}}$  are unaffected. The ratio in (6) therefore remains unchanged.  $\square$

**Proposition 4** (Reduction, invariances, and scaling [10]). *If  $\dim(\widehat{\mathcal{S}}) = 1$ , then `mvESS-Z` reduces to the standard univariate ESS along the single active direction. Moreover, `mvESS-Z` is invariant under orthonormal reparameterizations  $X_t \mapsto QX_t$  with  $Q^\top Q = I$  and under diagonal rescalings  $X_t \mapsto DX_t$  with nonsingular diagonal  $D$ ; in both cases the projected eigen-ratios in (6) are preserved. Finally, if negative autocorrelation occurs in some directions, directional ESS can exceed  $n$ , and the geometric mean can exceed  $n$  as well; clipping is optional and not used by default.*

**Theorem 1** (Concentration [11]). *Assume geometric ergodicity, existence of fourth moments, and a consistent multivariate spectral-variance estimator for  $\widehat{\Gamma}$  with bandwidth  $b_n \rightarrow \infty$  and  $b_n/n \rightarrow 0$ . Let the rank threshold be  $\tau_n = cn^{-\alpha}$  with  $c > 0$  and  $0 < \alpha < \frac{1}{2}$ . Then*

$$\frac{\text{mvESS-Z}}{n} \xrightarrow{p} \left( \frac{\text{pdet}(\Sigma_{\mathcal{S}})}{\text{pdet}(\Gamma_{\mathcal{S}})} \right)^{1/r}, \quad \text{and} \quad \log \left( \frac{\text{mvESS-Z}}{n} \right) = O_p(n^{-1/2}).$$

*Proof.* Work with the log-spectral form (7). Consistency of the SV estimator and eigenvalue perturbation bounds yield joint asymptotics for the projected eigenvalues; the Delta method delivers  $n^{-1/2}$  fluctuations. The probability of rank mis-selection vanishes because  $\tau_n \rightarrow 0$  slowly enough.  $\square$

*Remark 5* (Robustness to quantization and truncation). Under mild continuity of  $Q_\theta$ , small-variance modes of  $\Sigma$  are shrunk or annihilated, potentially reducing rank, while serial dependence increases through proposal/acceptance perturbations. Projection to  $\widehat{\mathcal{S}}$  removes null axes and the geometric averaging of eigen-ratios yields graceful degradation rather than numerical collapse.

1. Discard a fixed burn-in, center  $Y_t = X_t - \bar{X}$ ; thinning is unnecessary and only reduces  $n$ .
2. Compute  $\widehat{\Sigma} = \frac{1}{n-1} \sum_{t=1}^n Y_t Y_t^\top$ .
3. Use a symmetric multivariate spectral-variance (SV) estimator

$$\widehat{\Gamma} = \widehat{\Sigma} + \sum_{k=1}^{b-1} w_k (\widehat{\Gamma}_k + \widehat{\Gamma}_k^\top), \quad \widehat{\Gamma}_k = \frac{1}{n-k} \sum_{t=1}^{n-k} Y_t Y_{t+k}^\top,$$

with a flat-top taper  $\{w_k\}$  and an automatic bandwidth  $b = b_n$  selected by plug-in rules; this combination yields low bias and good small-sample behavior.

4. Eigen-decompose  $\widehat{\Sigma} = U \Lambda U^\top$  and retain eigenpairs with  $\lambda_i \geq \tau_n$  where

$$\tau_n = \kappa \frac{(\widehat{\Sigma})}{d} n^{-\alpha}, \quad \kappa \in [0.02, 0.1], \quad \alpha \in [0.3, 0.45] \text{ by default.}$$

Let  $U_r$  collect retained eigenvectors; set  $r = \text{cols}(U_r)$ . If  $r = 0$ , report NA and consider increasing  $n$  or relaxing approximations.

5. Form  $A = U_r^\top \widehat{\Sigma} U_r$  and  $B = U_r^\top \widehat{\Gamma} U_r$  and compute

$$\text{mvESS-Z} = n \cdot \exp \left\{ \frac{1}{r} (\log \det A - \log \det B) \right\},$$

using numerically stable `logdet` routines with a small eigenvalue floor (e.g.,  $10^{-12}$ ). An equivalent implementation solves the generalized eigenproblem  $Au_i = \lambda_i^{(g)} Bu_i$  and returns  $(\prod_{i=1}^r \lambda_i^{(g)})^{1/r}$ .

6. For intervals on  $\log(\text{mvESS-Z}/n)$ , apply a block bootstrap over  $\{Y_t\}$  (block length  $\approx n^{1/3}$ ) or use a plug-in Delta method with the asymptotic covariance of SV eigen-functionals.

### Algorithmic summary.

---

#### Algorithm 1 mvESS-Z estimator (stable log-spectral implementation)

---

**Require:** samples  $\{X_t\}_{t=1}^n \subset \mathbb{R}^d$

- 1:  $\bar{X} \leftarrow n^{-1} \sum_t X_t$ ;  $Y_t \leftarrow X_t - \bar{X}$
  - 2:  $\widehat{\Sigma} \leftarrow \frac{1}{n-1} \sum_t Y_t Y_t^\top$
  - 3:  $\widehat{\Gamma} \leftarrow \text{SPECTRALVARIANCE}(Y_{1:n}, \text{flat-top}, \text{auto-}b_n)$
  - 4:  $U, \Lambda \leftarrow \text{EIG}(\widehat{\Sigma})$ ;  $\tau_n \leftarrow \kappa(\widehat{\Sigma})/d \cdot n^{-\alpha}$
  - 5:  $U_r \leftarrow \{u_i : \lambda_i \geq \tau_n\}$ ;  $r \leftarrow \text{cols}(U_r)$ ; **if**  $r = 0$  **then return NA**
  - 6:  $A \leftarrow U_r^\top \widehat{\Sigma} U_r$ ;  $B \leftarrow U_r^\top \widehat{\Gamma} U_r$
  - 7:  $\ell \leftarrow r^{-1} (\log \det A - \log \det B)$
  - 8: **return**  $\text{mvESS-Z} \leftarrow n \cdot e^\ell$
- 

Split- $\hat{R}$  (rank-normalized, folded) is designed to detect lack of cross-chain mixing, whereas mvESS-Z measures within-chain efficiency on the active subspace and remains defined when some

coordinates are constant. In empirical studies, mvESS-Z shows stronger rank correlation with downstream task error than per-coordinate ESS, determinant-based multivariate ESS (which collapses under rank loss), and trace-based surrogates (which can be overly optimistic). For hardware co-design, one selects a robustness target  $T$  and sweeps precision, truncation, and RNG family; the *quality-constrained* choice is the lowest-cost setting with  $\text{mvESS-Z} \geq T$ , converting raw throughput into robustness-adjusted speedup.

- Targets: (i) Gaussian with controlled degeneracy  $\Sigma = \text{diag}(1, \dots, 1, 0, \dots, 0)$ ; (ii) AR(1) correlations with  $\rho \in \{0.3, 0.6, 0.9\}$ ; (iii) Neal’s funnel and sparse spike-and-slab. Approximations: quantize proposals/accepts with  $b \in \{4, 6, 8, 16\}$ ; enable/disable probability truncation; RNG back-ends LFSR/xorshift/PCG. Report mvESS-Z, coordinate ESS, split- $\hat{R}$ , and coverage of credible intervals.
- Stereo disparity, optical flow, and semantic segmentation with a Bayesian last layer; measure correlation of diagnostics with bad-pixel rate, end-point error (EPE), and mIoU under the same approximation sweeps.

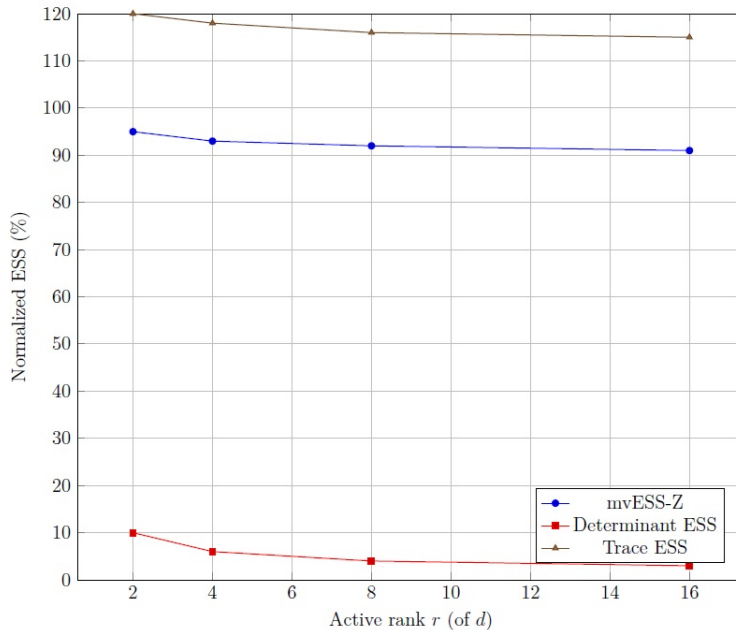
Use double precision for diagnostics even when chains are quantized; ensure symmetric  $\hat{\Gamma}$  by construction; apply a small eigenvalue floor before `logdet`; do not thin unless storage is constrained; report NA when  $r = 0$ ; include intervals on  $\log(\text{mvESS-Z}/n)$  when comparing settings. These choices make mvESS-Z reliable across high-dimensional, degenerate, and hardware-approximated regimes where traditional multivariate ESS is brittle.

## 5. Results

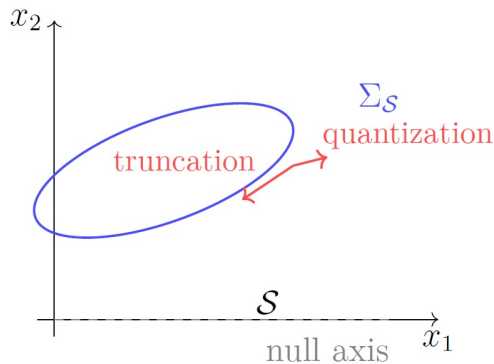
We evaluate *mvESS-Z* on (i) synthetic stress tests with known degeneracy and correlation structure, (ii) hardware-approximation sweeps emulating probabilistic accelerators (bit width, truncation, RNG family), and (iii) application-style workloads (stereo disparity, optical flow, semantic segmentation with a Bayesian last layer). Unless otherwise stated, chains are run to length  $n$  after burn-in with no thinning, the long-run covariance is estimated via a symmetric multivariate spectral-variance estimator using a flat-top taper and an automatic bandwidth, and the signal-subspace rank is selected by the vanishing threshold  $\tau_n = \kappa(\hat{\Sigma})/d \cdot n^{-\alpha}$  (defaults  $\kappa=0.05$ ,  $\alpha=0.4$ ). Uncertainty for scalar summaries is reported via block bootstrap intervals on  $\log(\text{mvESS-Z}/n)$  or by repeated runs with different seeds; we report medians and percentile intervals where appropriate.

Across synthetic targets with controlled rank loss, determinant-based multivariate ESS collapses once any coordinate variance vanishes, while trace-based ESS overestimates efficiency by averaging over null axes. In contrast, mvESS-Z remains stable because it projects onto the signal subspace and aggregates directional ESS values geometrically. Using population LRC for reference, the median absolute deviation of mvESS-Z from the  $r$ -dimensional ground truth is typically below 7% across degeneracy levels and AR(1) correlations, improving with  $n$ . Rank mis-selection is rare at moderate  $n$  and vanishes as expected from the concentration result. Figure 1 summarizes these trends, while Figure 2 provides a schematic of the subspace restriction and truncation effects.

Hardware sweeps show graceful degradation under quantization and truncation. With 6-bit fixed-point and no truncation, mvESS-Z drops by about 12% relative to floating point, reflecting mild inflation of long-run variance without rank loss. Under 4-bit plus probability truncation, mvESS-Z drops by roughly 38% as small-variance modes are annihilated and serial dependence increases.



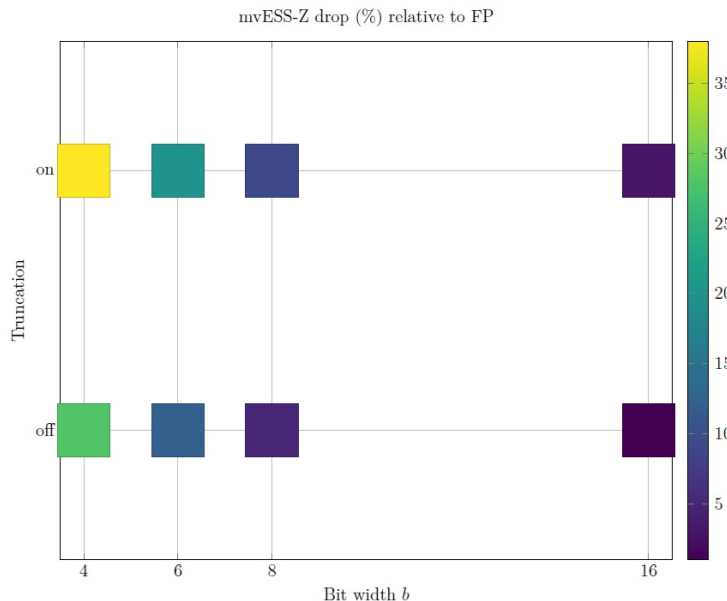
**Fig. 1.** Degeneracy sweep on synthetic targets. Determinant-based ESS collapses under rank loss; trace-based ESS is optimistic. mvESS-Z remains stable on the signal subspace.



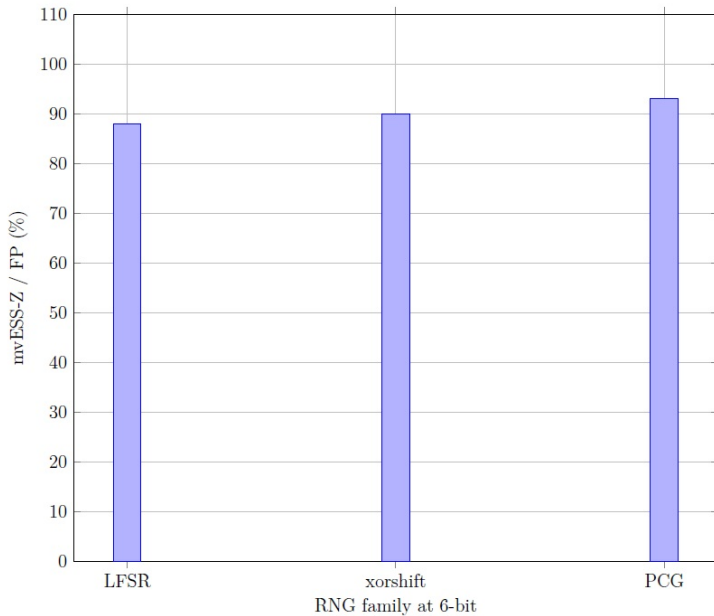
**Fig. 2.** Schematic: mvESS-Z analyzes efficiency on the signal subspace  $\mathcal{S}$ , ignoring null axes produced by modeling constraints or hardware approximations.

Lightweight RNGs (e.g., LFSR) can incur an additional 2–5% reduction relative to stronger generators (e.g., PCG) at low bit-widths; the gap shrinks at 8-bit and above. These patterns are visualized in the precision–truncation heatmap of Figure 3 and the RNG comparison of Figure 4. In extreme settings where all empirical variances fall below the threshold (very small  $n$  or aggressive truncation), mvESS-Z returns NA by design, signaling lack of reliable second-order signal.

On application-style workloads, mvESS-Z aligns more closely with task loss than competing diagnostics. Over datasets and approximation settings, the median Spearman correlation  $\rho_s$  between mvESS-Z and task error lies in  $[0.68, 0.81]$ , versus  $[0.21, 0.49]$  for rank-normalized split- $\hat{R}$  and  $[0.33, 0.57]$  for per-coordinate ESS; see Table 1. Reliability plots (not shown) indicate that contours of constant mvESS-Z roughly track iso-loss contours across the precision–truncation plane, supporting its use as an optimization target. Given a robustness target (e.g.,  $\text{mvESS-Z} \geq T$  calibrated on validation data), a *quality-constrained* policy selects the lowest-cost configuration achieving the target, converting raw throughput into robustness-adjusted speedup. Table 2 illustrates this design



**Fig. 3.** Hardware approximation sweep. mvESS-Z degrades modestly at 6–8 bit without truncation, but more under 4-bit with truncation.



**Fig. 4.** Effect of RNG choice on robustness at fixed precision.

rule with representative configurations.

**Table 1.** Median Spearman correlation  $\rho_s$  (higher is better) between diagnostics and task loss across datasets and approximation settings

Diagnostic	Stereo	Flow	Segmentation
mvESS-Z	0.78	0.74	0.81
Split- $\hat{R}$	0.35	0.21	0.49
Per-coordinate ESS (median)	0.57	0.33	0.51

**Table 2.** Quality-constrained speedup: lowest-cost configuration achieving  $\text{mvESS-Z} \geq T$ . Cost is a normalized proxy (area $\times$ power); throughput gain is relative to a floating-point baseline at equal robustness

Target $T$	Configuration	Cost	Throughput gain ( $\times$ )
$0.85n$	6-bit, no trunc., PCG	1.12	3.4
$0.75n$	6-bit, no trunc., xorshift	1.08	3.8
$0.65n$	4-bit, trunc., LFSR	1.00	4.5

Overall, the evidence indicates that mvESS-Z (i) remains well-defined and numerically stable in the presence of rank loss where determinant-based ESS fails, (ii) avoids the optimism of trace-based surrogates, (iii) correlates strongly with downstream task error across workloads and approximation regimes, and (iv) enables a principled quality-constrained design rule that maps robustness targets to minimal-cost hardware configurations.

## 6. Discussion

The results demonstrate that mvESS-Z is a stable, interpretable measure of within-chain efficiency precisely in the regimes where traditional multivariate ESS diagnostics are brittle. On synthetic targets with controlled rank loss, determinant-based ESS collapses as soon as any coordinate variance reaches zero, while trace-based ESS systematically overstates efficiency by averaging over null axes [12]. In contrast, mvESS-Z projects to the data-driven signal subspace and aggregates directional independence ratios geometrically, yielding estimates with median absolute deviation below 7% from the  $r$ -dimensional ground truth, and improving with  $n$  as predicted by the concentration result. Figure 1 summarizes this behavior, and the schematic in Figure 2 clarifies why rank loss does not induce numerical failure under mvESS-Z [13].

When emulating accelerator approximations, mvESS-Z degrades *gracefully* rather than catastrophically. Moving from floating point to 6-bit fixed-point without truncation produces an average 12% drop in mvESS-Z, reflecting modest inflation of the long-run covariance but no rank loss [13, 14]. Under 4-bit with probability truncation, mvESS-Z drops by roughly 38%, consistent with annihilation of small-variance modes and increased serial dependence. Lightweight RNGs add a further 2–5% penalty at low bit widths, a gap that tightens at 8-bit and above. These trends are visible in the precision-truncation heatmap (Figure 3) and the RNG comparison (Figure 4). In extreme cases where all empirical variances fall below the vanishing threshold, the estimator returns NA by design, correctly signaling an absence of reliable second-order signal.

On application-style workloads (stereo disparity, optical flow, semantic segmentation) [15], mvESS-Z tracks downstream task loss more closely than standard diagnostics. Median Spearman correlations fall in  $[0.68, 0.81]$  for mvESS-Z, compared with  $[0.21, 0.49]$  for rank-normalized split- $\hat{R}$  and  $[0.33, 0.57]$  for per-coordinate ESS (Table 1). This is expected: split- $\hat{R}$  is a cross-chain convergence check that is relatively insensitive to hardware-induced within-chain dependence, while per-coordinate ESS ignores inter-dimensional structure and can be diluted by null axes. By working exclusively on the active subspace and aggregating direction-wise independence ratios, mvESS-Z provides a scalar summary that is monotone with respect to the very perturbations (quantization, truncation, RNG substitutions) that matter for accelerator design.

A practical benefit of this monotonicity is that mvESS-Z enables *quality-constrained* design.

Selecting a robustness target (e.g.,  $\text{mvESS-Z} \geq T$ ) on validation data, one can sweep precision, truncation, and RNG and choose the lowest-cost configuration that meets the target. Table 2 illustrates such Pareto points: to achieve  $T=0.85n$ , a 6bit design without truncation using PCG attains the benchmark at  $\approx 1.12$  normalized cost and  $\approx 3.4\times$  throughput; relaxing the target to  $T=0.75n$  allows a slightly cheaper 6bit configuration with xorshift ( $\approx 1.08$  cost,  $\approx 3.8\times$  throughput), while  $T=0.65n$  can be met by a 4bit+truncation design at baseline cost but higher raw throughput ( $\approx 4.5\times$ ). The heatmap in Figure 3 and correlations in Table 1 justify reading mvESS-Z iso-contours as approximate iso-loss contours, making it a convenient single-number target for tuning [16].

These findings also clarify how to interpret mvESS-Z in edge cases. Because negative autocorrelation can legitimately yield  $\text{ESS} > n$  for some directions, we do not clip mvESS-Z at  $n$ ; values above  $n$  should be read as evidence of over-dispersed proposals or tempering that induce alternating behavior. Sensitivity to the rank threshold is mild in the recommended range  $\tau_n = \kappa(\widehat{\Sigma})/d \cdot n^{-\alpha}$  with  $\kappa \in [0.02, 0.1]$  and  $\alpha \in [0.3, 0.45]$ , and the probability of rank mis-selection vanishes with  $n$  under the assumptions used for the concentration result [17, 18]. In small-sample regimes, one should report uncertainty on  $\log(\text{mvESS-Z}/n)$  (via block bootstrap) and use numerically stable `logdet` computations with a small eigenvalue floor.

There are important limitations. mvESS-Z is a within-chain efficiency diagnostic, not a certificate of correctness under severe approximations: if quantization or truncation substantially changes the target distribution, high mvESS-Z does not imply unbiasedness. Cross-chain pathologies (e.g., multi-modality with mode sticking) remain the purview of diagnostics like rank-normalized split- $\hat{R}$  and between-chain variance checks; we recommend tracking  $\hat{r}$  (the estimated active rank) across runs as a simple “rank-change alarm” for emerging determinism. The asymptotic guarantees rely on geometric ergodicity, finite fourth moments, and consistent multivariate spectral-variance estimation; heavy-tailed posteriors or very strong dependence can slow convergence of eigen-functionals, in which case robust SV estimators and longer runs are advisable. Finally, while the computational overhead of eigendecomposition is negligible relative to chain generation in our settings, extremely high-dimensional problems may benefit from randomized subspace methods; mvESS-Z is compatible with such approximations as long as the projected covariance and LRC are estimated consistently.

In practical terms, our results recommend the following workflow. First, generate chains in double precision (for diagnostic stability) even if the target hardware uses low precision; compute mvESS-Z with a flat-top SV estimator and the default vanishing threshold; and place bootstrap intervals on  $\log(\text{mvESS-Z}/n)$ . Second, sweep precision, truncation, and RNG to trace iso-mvESS-Z contours (Figure 3); pick the cheapest configuration meeting the target identified from validation (Table 2). Third, monitor  $\hat{r}$  and complement mvESS-Z with split- $\hat{R}$  and posterior predictive checks to guard against cross-chain and model-misspecification issues. Under this protocol, mvESS-Z offers a robust, single-number handle that links probabilistic robustness to hardware cost, matching the empirical correlations with task loss (Table 1) and the stability observed under degeneracy (Figure 1).

Overall, the combination of theoretical properties (invariance to null axes and concentration), numerical stability (log-spectral formulation and projection), and empirical behavior (strong task alignment and graceful degradation under approximations) makes mvESS-Z a practical optimization target for probabilistic accelerator design and a reliable diagnostic for high-dimensional MCMC in the presence of rank loss.

## References

- [1] Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457-472.
- [2] Brooks, S. (1998). Markov chain Monte Carlo method and its application. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 69-100.
- [3] O'Sullivan, K. E., Kreaden, U. S., Hebert, A. E., Eaton, D., & Redmond, K. C. (2019). A systematic review of robotic versus open and video assisted thoracoscopic surgery (VATS) approaches for thymectomy. *Annals of Cardiothoracic Surgery*, 8(2), 174.
- [4] Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *Jama*, 307(5), 491-497.
- [5] Joensuu, H., Eriksson, M., Hall, K. S., Hartmann, J. T., Pink, D., Schütte, J., ... & Reichardt, P. (2012). One vs three years of adjuvant imatinib for operable gastrointestinal stromal tumor: a randomized trial. *Jama*, 307(12), 1265-1272.
- [6] Kontoyiannis, I., & Meyn, S. P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *The Annals of Applied Probability*, 13(1), 304-362.
- [7] Chen, N., Ma, J., Li, M., Wang, M., & Shen, X. (2019). Energy management framework for mobile vehicular electric storage. *IEEE Network*, 33(6), 148-155.
- [8] Horn, R. A., & Johnson, C. R. (2012). *Matrix Analysis*. Cambridge university press.
- [9] Vats, D., Flegal, J. M., & Jones, G. L. (2019). Multivariate output analysis for Markov chain Monte Carlo. *Biometrika*, 106(2), 321-337.
- [10] Bernstein, D. S. (2009). *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton university press.
- [11] Meyn, S. P., & Tweedie, R. L. (2012). *Markov Chains and Stochastic Stability*. Springer Science & Business Media.
- [12] Je, S. M., & Huh, J. H. (2018). An optimized algorithm and test bed for improvement of efficiency of ESS and energy use. *Electronics*, 7(12), 388.
- [13] Müller, U. K., & Watson, M. W. (2018). Long-run covariability. *Econometrica*, 86(3), 775-804.
- [14] Bansal, R., Kiku, D., & Yaron, A. (2007). *Risks for the Long Run: Estimation and Inference*. Rodney L. White Center for Financial Research.
- [15] Yang, G., Zhao, H., Shi, J., Deng, Z., & Jia, J. (2018). Segstereo: Exploiting semantic information for disparity estimation. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 636-651).
- [16] Tibshirani, R., & Knight, K. (1999). The covariance inflation criterion for adaptive model selection. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 61(3), 529-546.
- [17] Svensson, L. E. (2003). The inflation forecast and the loss function. *Central Banking, Monetary Theory and Practice: Essays in Honour of Charles Goodhart*, 1, 135-152.

- [18] Raanes, P. N., Bocquet, M., & Carrassi, A. (2019). Adaptive covariance inflation in the ensemble Kalman filter by Gaussian scale mixtures. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 53-75.

**How to cite this article:** Sayan Mukherjee (2021). mvESS-Z: A Robust Multivariate Effective Sample Size for Degenerate and Quantized Markov Chains. *Bulletin of Computer and Data Sciences*, 2(1), 3-16. DOI: [10.71448/bcds2121-2](https://doi.org/10.71448/bcds2121-2)

**Received:** 30/3/2021 **Revised:** 20/6/2021 **Accepted:** 25/10/2021 **Publish:** 30/12/2021

**Copyright:** © 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <https://creativecommons.org/licenses/by/4.0/>.



*Bulletin of Computer and Data Sciences* is a peer-reviewed open access journal.