FINDING SIGNAL IN RANDOM NOISE: SMOOTH HARD THRESHOLDING FOR SINGULAR VALUES WITH SURE

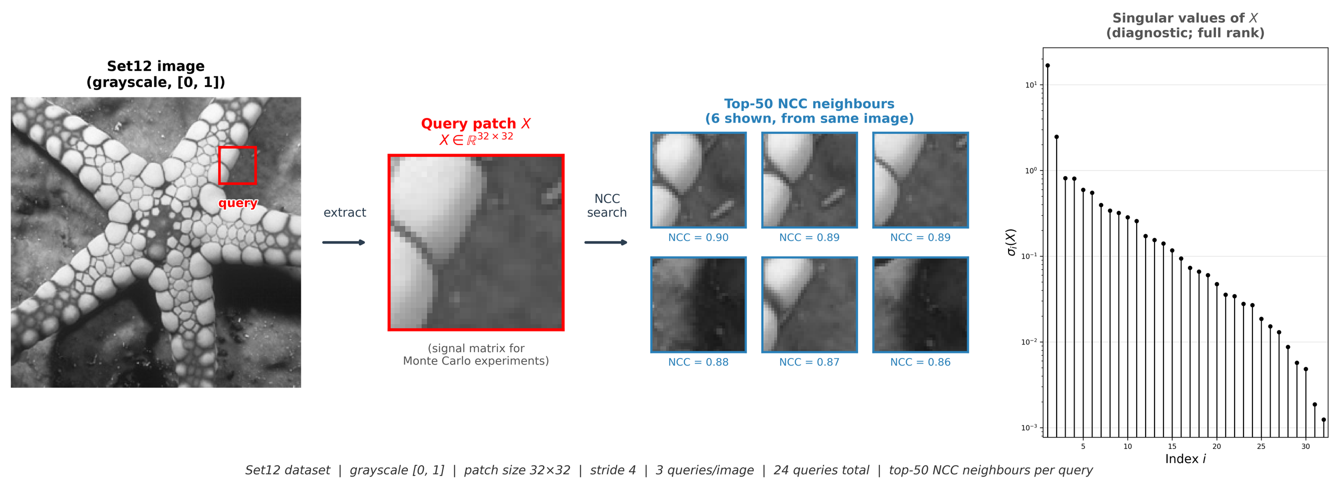
Guanzhong Yang
Department of Mathematics, Imperial College London

1. Images Become Low-Rank Patch Matrices

A noisy patch group is modeled as

$$Y = X + W, \qquad W_{ij} \sim N(0, \tau^2).$$

Natural images are locally self-similar: grouping similar patches concentrates the main signal into a few large singular values.



Self-similarity turns repeated local structure into an approximately low-rank matrix.

Thus, for self-similarity-based image denoising, a standard local workflow is:

patch group → SVD → choose rank h → denoised group

Grouping similar patches concentrates signal into a few large singular values, so we can denoise image by discarding small signal values.

2. Classical Energy Matching

Let a patch-group matrix have SVD

$$P = U \operatorname{diag}(\sigma_1, \ldots, \sigma_k) V^T,$$

and let $P_h = U \operatorname{diag}(\sigma_1, \ldots, \sigma_h, 0, \ldots, 0) V^T$ be its rank-h truncation. Classical energy matching chooses

$$\widehat{h}_{\mathrm{EM}} = \min \left\{ h : \sum_{i=h+1}^{k} \sigma_i^2 \le mn\tau^2 \right\}.$$

Its implicit surrogate is that fitted residual energy behaves like noise energy:

$$\mathbb{E} \left\| Y - \widehat{X} \right\|_F^2 \approx mn\tau^2.$$

But for a data-fitted estimator,

$$\mathbb{E} \left\| Y - \widehat{X} \right\|_F^2 = \mathbb{E} \left\| \widehat{X} - X \right\|_F^2 + mn\tau^2 - 2\mathbb{E} \left\langle \widehat{X} - X, W \right\rangle.$$

The covariance term is generally nonzero because $\widehat{X}$ adapts to noise in Y.

Traditional energy matching method chooses the rank based on a biased risk estimate.

3. SURE Estimates Risk Without Knowing X

For Gaussian noise, known variance, and a fixed estimator with sufficient regularity,

$$\mathrm{SURE}(Y) = \left\| \widehat{X}(Y) - Y \right\|_F^2 + 2\tau^2 \operatorname{div}(\widehat{X})(Y) - mn\tau^2$$

satisfies

$$\mathbb{E}\, \mathrm{SURE}(Y) = \mathbb{E} \left\| \widehat{X}(Y) - X \right\|_F^2.$$

Thus SURE lets us rank denoisers using Y and $\widehat{X}(Y)$, even though the clean matrix X is unknown. The divergence term is the degrees-of-freedom correction missing from residual energy.

SURE provides an unbiased risk estimate without knowing X under certain circumstances.

4. Smooth Thresholding Gives A Computable Rank Score

Hard singular-value thresholding is natural for rank selection, but it is discontinuous, so SURE is not directly justified. We smooth it by

$$f_{\omega,\lambda}(x) = \frac{x}{1 + \exp[-\omega(x - \lambda)]}, \qquad \omega > 0.$$

For fixed (ω, λ), the spectral estimator satisfies the SURE regularity conditions. As $\omega \to \infty$, it approaches hard thresholding away from $x = \lambda$. Candidate thresholds are placed between adjacent singular values, so each candidate corresponds to retaining a rank h.

For a fixed observed matrix and finitely many candidates, strict inequalities between limiting scores are preserved for all sufficiently large ω. Thus the hard-threshold limit gives a computable score from the observed singular values:

$$S_0(Y) = \|Y\|_F^2 - mn\tau^2,$$

$$S_h(Y) = -mn\tau^2 + \sum_{i=h+1}^{k} \sigma_i^2 + 2\tau^2 \left[|m - n| h + h + 2 \sum_{i=1}^{h} \sum_{\substack{\ell=1 \\ \ell \ne i}}^{k} \frac{\sigma_i^2}{\sigma_i^2 - \sigma_\ell^2} \right], \qquad h \ge 1.$$

We select

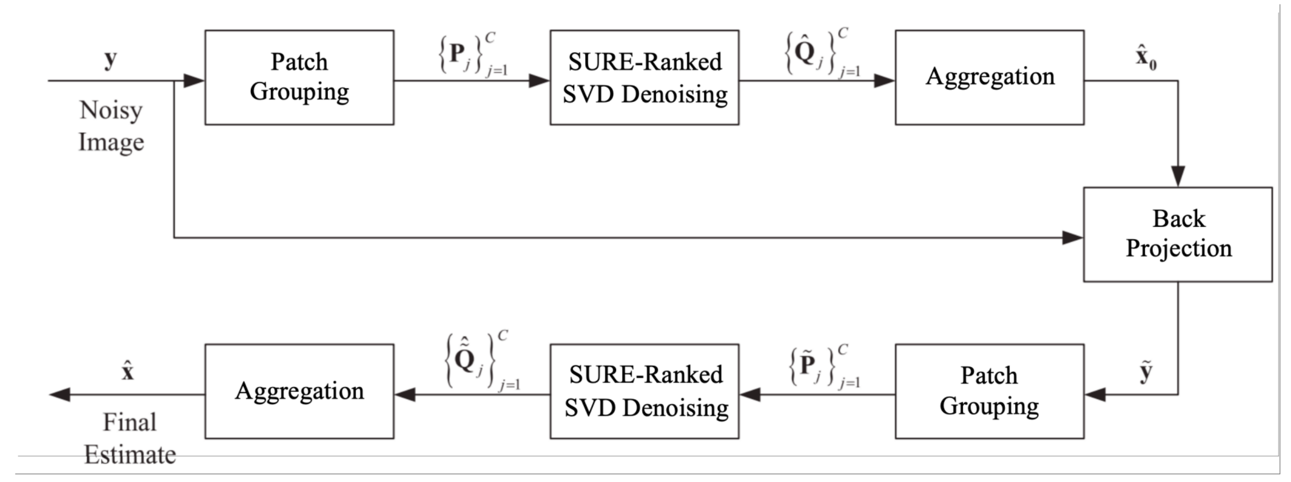
$$\widehat{h} \in \arg \min_{0 \le h \le k} S_h(Y).$$

Caveat: fixed-threshold SURE is unbiased before data-adaptive minimization; after choosing h from Y, the selected value is a criterion, not an unbiased post-selection risk estimate.

The limiting SURE score gives a rank rule; after selection, it is a criterion, not unbiased risk.

5. Same Pipeline, New Rank Rule

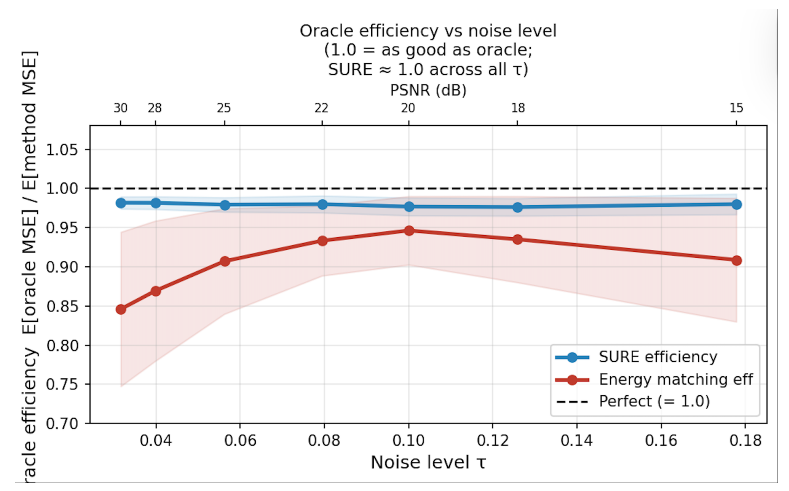
The image algorithm keeps patch grouping, SVD truncation, aggregation, back projection, and refinement. Only the local rank-selection rule changes.



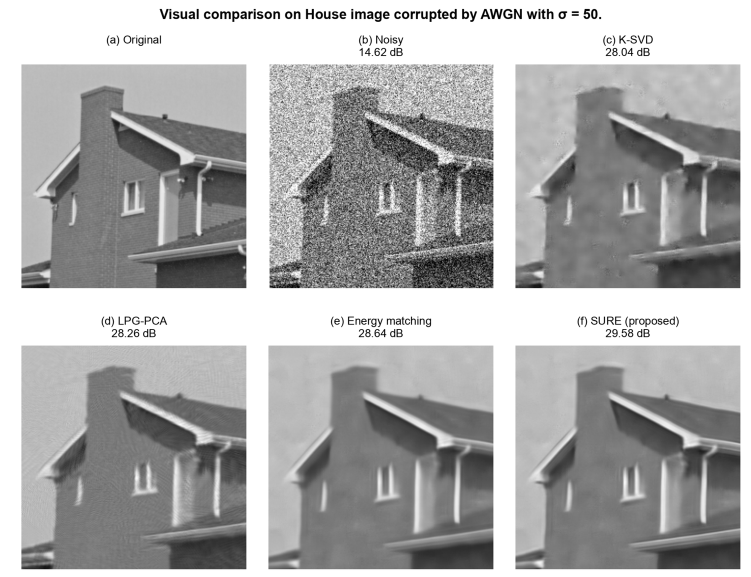
Patch groups → SVD → rank by S_h → aggregate denoised patches.

6. Numerical Evidence

The first check is matrix-level: does the selected rank behave like the oracle risk-minimizing rank? The second is image-level: does changing only the rank rule improve denoising in high noise?



SURE remains close to oracle efficiency; EM drops as noise increases.



At $\sigma = 50$, this realization is visibly cleaner under SURE rank selection.

For BSD68, each image–noise pair is tested with both rank rules on the same noisy realization; Δ = SURE minus energy matching.

σ	ΔPSNR	win	p	ΔSSIM	win	p
10	−0.517	4%	1.000	−0.0068	10%	1.000
30	−0.050	44%	0.978	−0.0016	44%	0.992
50	**+0.118**	**65%**	**0.041**	**+0.0129**	**68%**	**0.042**

One-sided Wilcoxon signed-rank tests show significant positive shifts only in the high-noise regime.

The gain is not uniform; at high noise, SURE gives statistically supported PSNR/SSIM improvement.

References, Code and Paper

[1] C. M. Stein. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, 9(6):1135–1151, 1981.
[2] E. J. Candès, C. A. Sing-Long, and J. D. Trzasko. Unbiased risk estimates for singular value thresholding and spectral estimators. *IEEE Transactions on Signal Processing*, 61(19):4643–4657, 2013.
[3] Q. Guo, C. Zhang, Y. Zhang, and H. Liu. An efficient SVD-based method for image denoising. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(5):868–880, 2016.
[4] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006.
[5] L. Zhang, W. Dong, D. Zhang, and G. Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43(4):1531–1549, 2010.
[6] A. Mukherjee, K. Chen, N. Wang, and J. Zhu. On the degrees of freedom of reduced-rank estimators in multivariate regression. *Biometrika*, 102(2):457–477, 2015.
[7] B. Efron. How biased is the apparent error rate of a prediction rule? *Journal of the American Statistical Association*, 81(394):461–470, 1986.
[8] B. Efron. The estimation of prediction error: covariance penalties and cross-validation. *Journal of the American Statistical Association*, 99(467):619–632, 2004.
[9] N. R. Hansen. A comment on Stein's unbiased risk estimate for reduced rank estimators. arXiv:1708.09657, 2017.

Generative AI acknowledgement. Beyond literature search, LaTeX troubleshooting and feedback, I used Prism to polish the poster text and Kiro to support coding and debugging.

GitHub repository   arXiv preprint