

Random Matrix Theory for Modeling the Eigenvalue Distribution of Images under Upscaling

**David Vázquez-Padín, Fernando Pérez-González, and
Pedro Comesaña-Alfaro**

Signal Theory and Communications Department, University of Vigo, Vigo, Spain.

Palermo, Italy. September 18, 2017.



UniversidadeVigo

Motivation

Nowadays, **digital images** cannot be treated as an **undeniable proof of occurrence** of an event.

Motivation

Nowadays, **digital images** cannot be treated as an **undeniable proof of occurrence** of an event.

Two weeks ago: Eduardo Martins' identity was based on fake images¹



Eduardo Martins' images:



The story behind:

Stolen pictures from real war photographers:

- + slight alterations
- + reversed images

¹<http://www.bbc.com/news/world-latin-america-41174069>

Motivation

Nowadays, **digital images** cannot be treated as an **undeniable proof of occurrence** of an event.

Two weeks ago: Eduardo Martins' identity was based on fake images¹



Eduardo Martins' images:



The story behind:

Stolen pictures from real war photographers:

- + slight alterations
- + reversed images



Fraud:

- o Cheat journalists
- o Fool picture editors
- o Elude plagiarism checkers

¹<http://www.bbc.com/news/world-latin-america-41174069>

Motivation

Nowadays, **digital images** cannot be treated as an **undeniable proof of occurrence** of an event.

Two weeks ago: Eduardo Martins' identity was based on fake images¹



Eduardo Martins' images:



Daniel C. Britt' images:



¹<http://www.bbc.com/news/world-latin-america-41174069>

Problem Description

When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

Example: Eduardo Martins' fabricated identity



Original image (Max Hepworth-Povey)



Forged image

Problem Description

When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

Example: Eduardo Martins' fabricated identity



Original image (Max Hepworth-Povey)



Forged image

Problem Description

When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

Example: Eduardo Martins' fabricated identity



Original image (Max Hepworth-Povey)



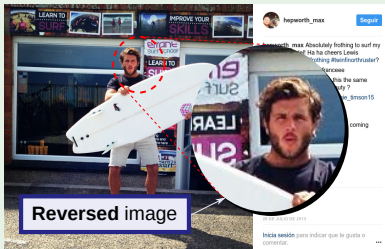
Forged image

Shutter button on the left side of the camera?

Problem Description

When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

Example: Eduardo Martins' fabricated identity



Original image (Max Hepworth-Povey)

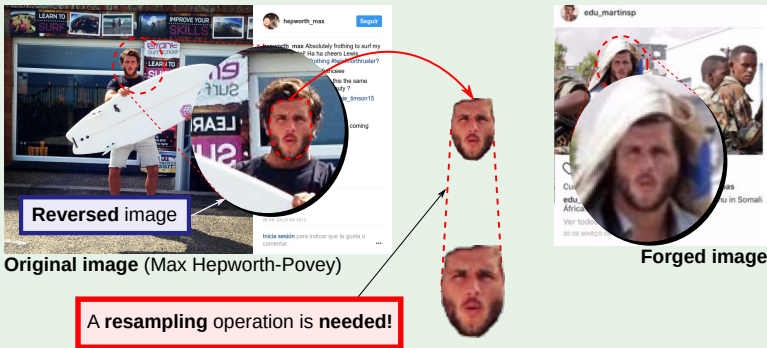


Forged image

Problem Description

When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

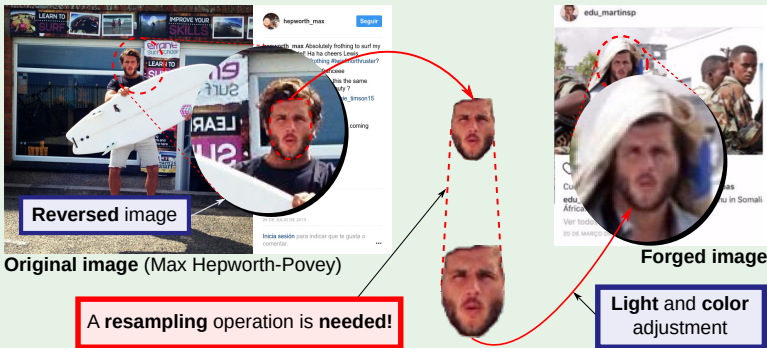
Example: Eduardo Martins' fabricated identity



Problem Description

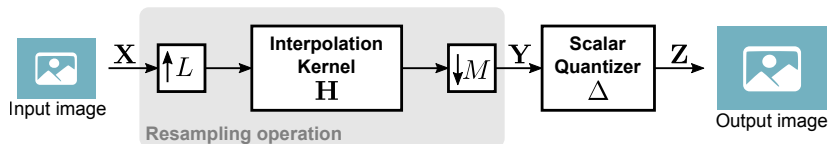
When a **credible forgery** is carried out, most of the time it is **necessary** to **adapt** added pieces to the original content.

Example: Eduardo Martins' fabricated identity



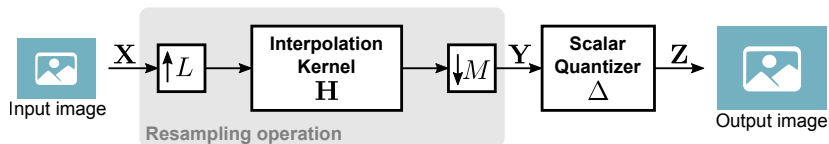
Resampling Process Description and Prior Work

Resampling operation by $\xi = \frac{L}{M} > 1$:

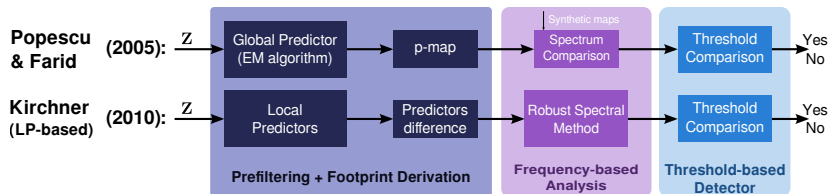


Resampling Process Description and Prior Work

Resampling operation by $\xi = \frac{L}{M} > 1$:

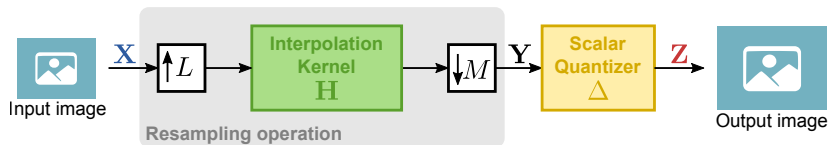


Prior Work: residue signal + spectral analysis + threshold comparison

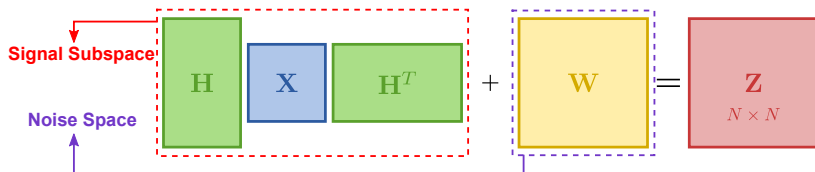


Resampling Process Description and Prior Work

Resampling operation by $\xi = \frac{L}{M} > 1$:

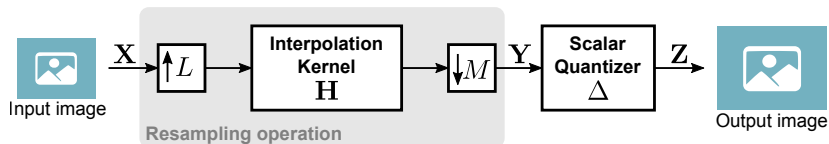


Our approach (2015): Subspace Decomposition $\rightarrow$ SVD-based analysis

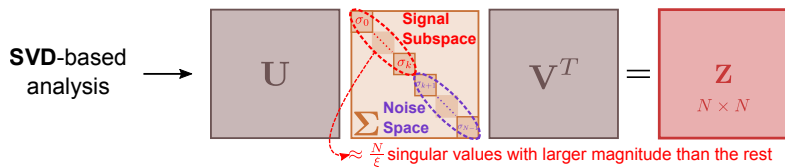


Resampling Process Description and Prior Work

Resampling operation by $\xi = \frac{L}{M} > 1$:

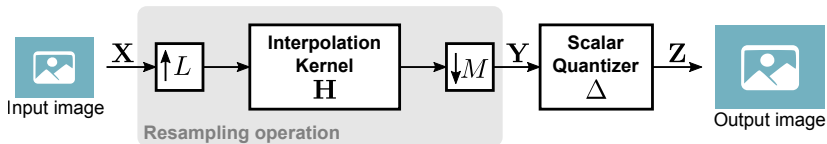


Our approach (2015): Subspace Decomposition $\rightarrow$ SVD-based analysis

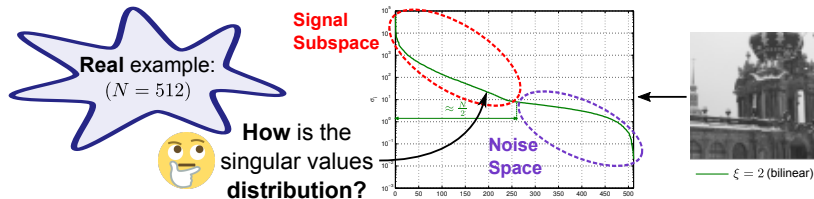


Resampling Process Description and Prior Work

Resampling operation by $\xi = \frac{L}{M} > 1$:

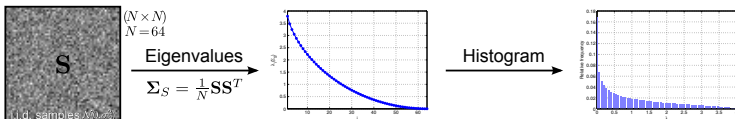


Our approach (2015): Subspace Decomposition $\rightarrow$ SVD-based analysis



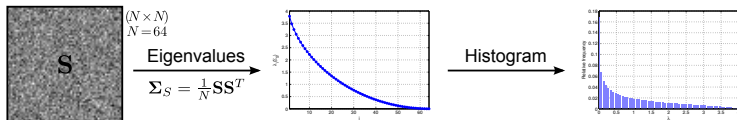
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



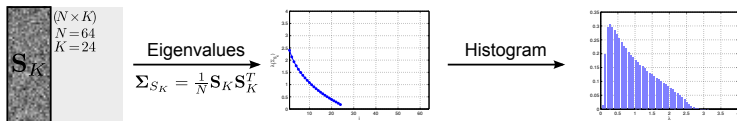
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



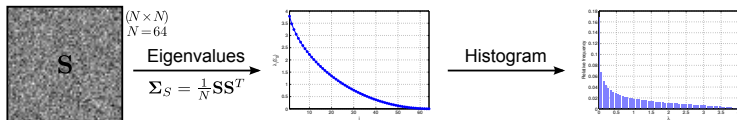
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



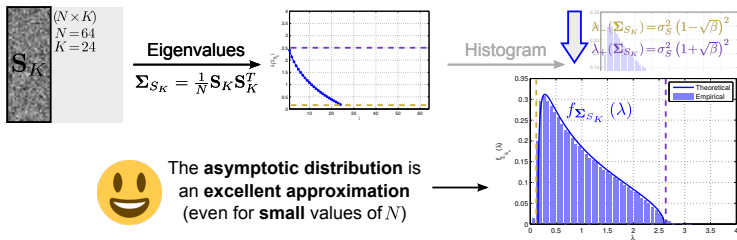
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

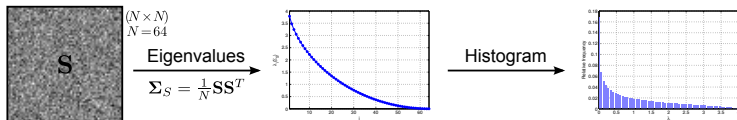
$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



The **asymptotic distribution** is an **excellent approximation** (even for **small values of N**)

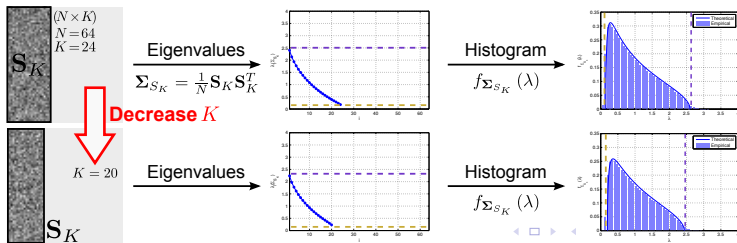
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



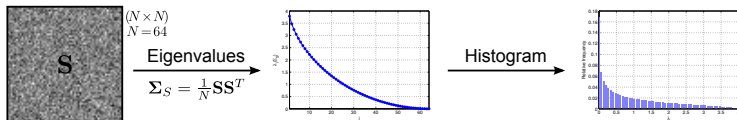
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



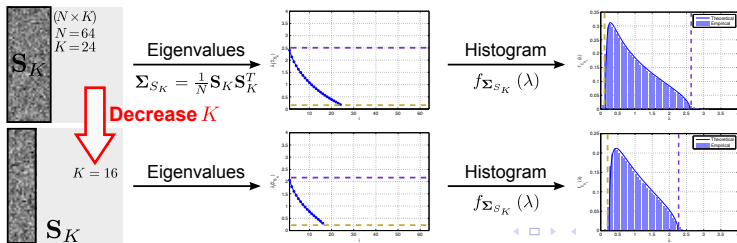
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



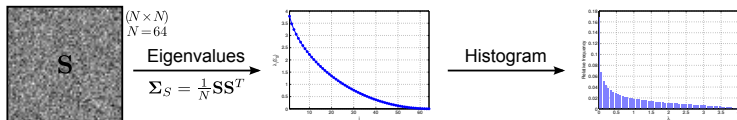
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



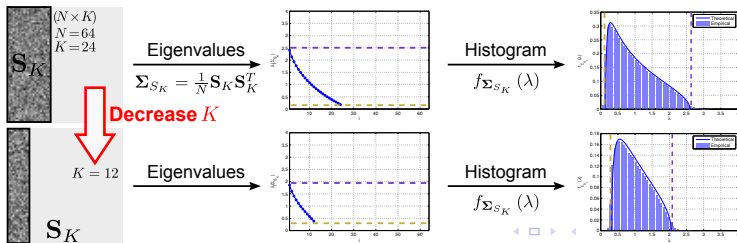
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



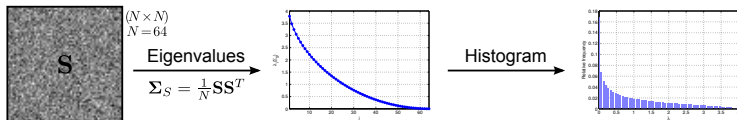
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



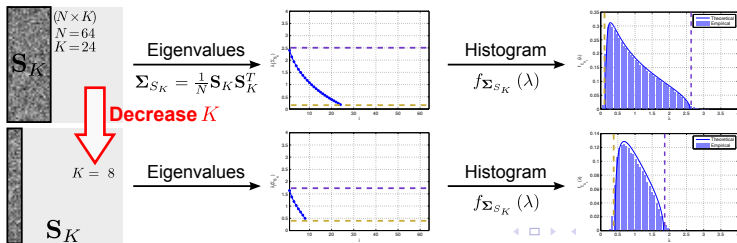
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



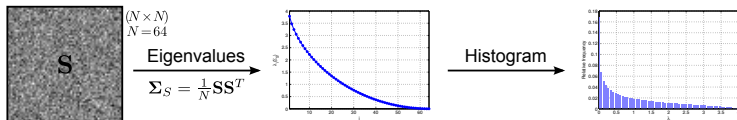
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



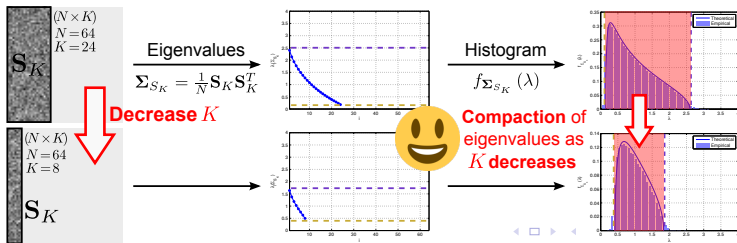
Random Matrix Theory for Eigenvalue Characterization

RMT: provides tools to model the **asymptotic eigenvalue distribution**



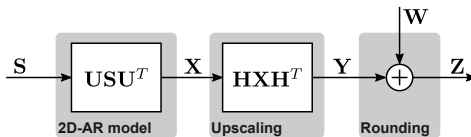
Marčenko-Pastur Law: asymptotic distribution for $N \times K$ matrices ($\beta = K/N$)

$$f_{\Sigma_{S_K}}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda_+(\Sigma_{S_K}) - \lambda)(\lambda - \lambda_-(\Sigma_{S_K}))}}{2\pi\sigma_S^2\lambda\beta}, & \text{if } \lambda_-(\Sigma_{S_K}) \leq \lambda \leq \lambda_+(\Sigma_{S_K}) \\ 0, & \text{otherwise} \end{cases}$$



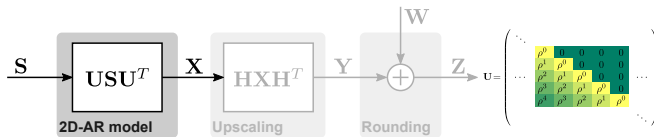
2D-AR Model and Eigenvalue Characterization

Our model: 2D-AR model $\rightarrow$ Upscaling $\rightarrow$ Rounding

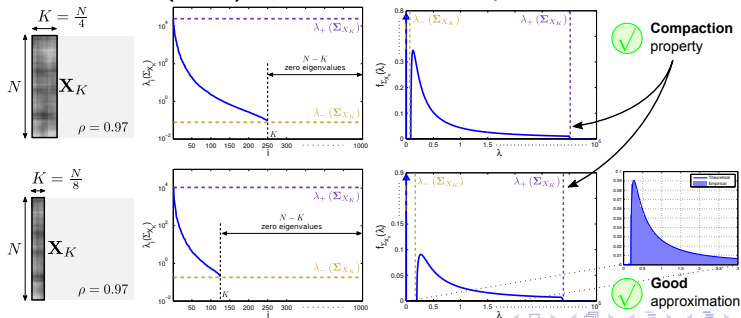


2D-AR Model and Eigenvalue Characterization

Our model: 2D-AR model $\rightarrow$ Upscaling $\rightarrow$ Rounding

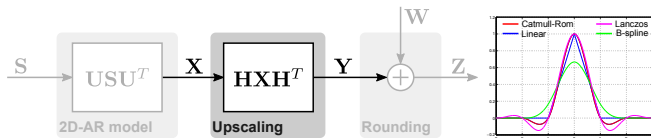


Tulino & Verdú (2004): η -transform $\rightarrow$ Stieltjes $\rightarrow$ inversion formula

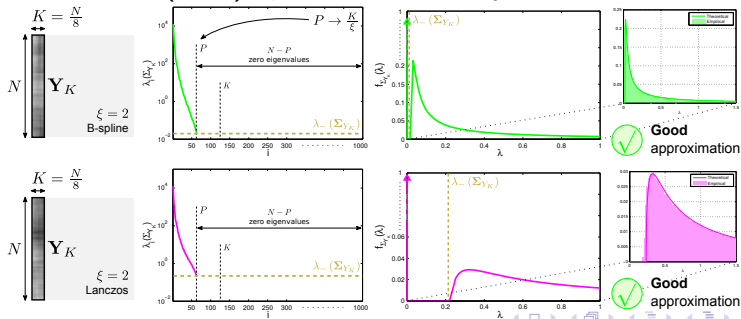


2D-AR Model and Eigenvalue Characterization

Our model: 2D-AR model $\rightarrow$ Upscaling $\rightarrow$ Rounding

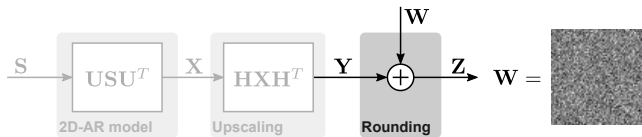


Tulino & Verdú (2004): η -transform $\rightarrow$ Stieltjes $\rightarrow$ inversion formula

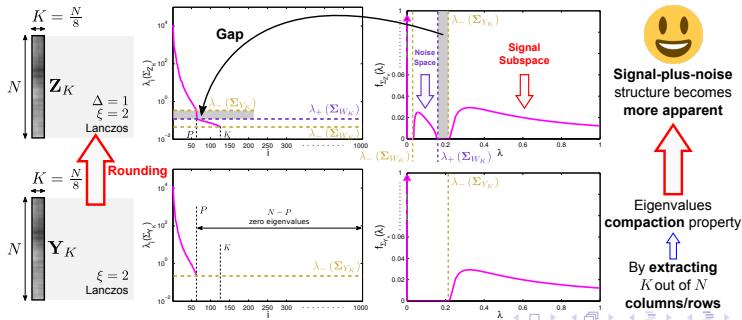


2D-AR Model and Eigenvalue Characterization

Our model: 2D-AR model $\rightarrow$ Upscaling $\rightarrow$ Rounding

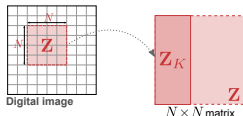


Asymptotic Gap: Tulino & Verdú for Y + Marčenko-Pastur for W



Description of Upscaling Detector

Observed Data:



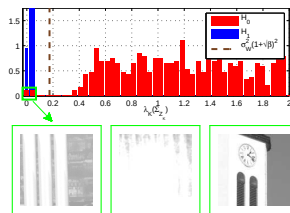
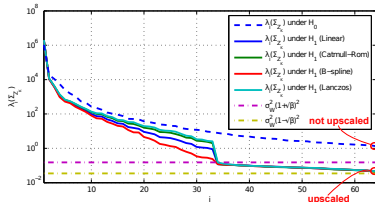
Hypothesis Test:

- $\mathcal{H}_0$: $\mathbf{Z}$ has never been upsampled
- $\mathcal{H}_1$: $\mathbf{Z}$ has been upsampled by ξ

Test statistic:

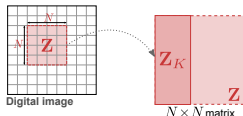
For a given K , measure the magnitude of the smallest eigenvalue of $\mathbf{Z}_K$

Example: $N = 512$, $K = 64$, $\xi = 2$ **But,** linear dependencies show up:



Description of Upscaling Detector

Observed Data:



Hypothesis Test:

- $\mathcal{H}_0$: $\mathbf{Z}$ has never been upsampled
- $\mathcal{H}_1$: $\mathbf{Z}$ has been upsampled by ξ

Test statistic: κ

$$\kappa \triangleq \begin{cases} \min_{v \in \{0, \dots, V-1\}} \Lambda_v, & \text{if } |\mathcal{S}| = 0 \\ \text{median}_{v \in \{0, \dots, V-1\} \setminus \mathcal{S}} \Lambda_v, & \text{if } 1 \leq |\mathcal{S}| < V \\ \lambda_0, & \text{if } |\mathcal{S}| = V \end{cases}$$

A set of V different submatrices are studied:

- $\Lambda \triangleq \left(\lambda_K \left(\Sigma_{Z_K}^{(0)} \right), \dots, \lambda_K \left(\Sigma_{Z_K}^{(V-1)} \right) \right)^T$
- $\mathcal{S} \triangleq \left\{ v : \lambda_K \left(\Sigma_{Z_K}^{(v)} \right) < \sigma_W^2 (1 - \sqrt{\beta})^2 \right\}$

Decision:

$$\mathcal{H}_0: \kappa \geq \sigma_W^2 (1 + \sqrt{\beta})^2$$

$$\mathcal{H}_1: \kappa < \sigma_W^2 (1 + \sqrt{\beta})^2$$

Experimental Setup

Image Dataset:

- 1317 images from the **Dresden Image Database** (Nikon cameras)

Full-frame resampling through `convert` from ImageMagick:

- **Interpolation kernels:** *Linear, Catmull-Rom, B-spline, Lanczos*
- **Set of resampling factors:** $\xi \in \{1.05, 1.1, \dots, 1.95, 2\}$

Benchmark detectors:

- **LP-based:**
 - Kirchner (2010): based on row/column predictors ($K = 3$)
- **SVD-based:**
 - Our approach (2015): subspace decomposition through SVD

Experimental Setup

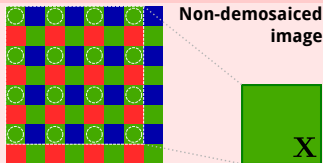
Image Dataset:

- 1317 images from the **Dresden Image Database** (Nikon cameras)

Full-frame resampling through `convert` from ImageMagick:

- **Interpolation kernels:** *Linear, Catmull-Rom, B-spline, Lanczos*
- **Set of resampling factors:** $\xi \in \{1.05, 1.1, \dots, 1.95, 2\}$

Images without demosaicing traces:

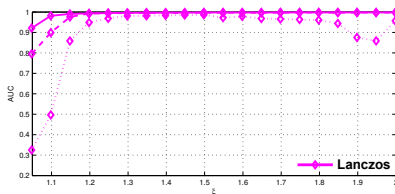
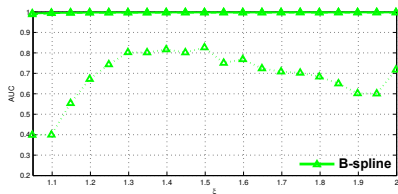
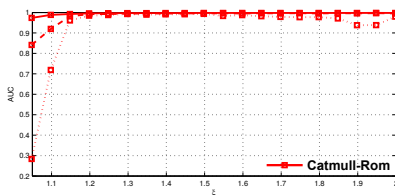
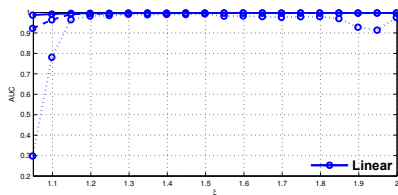


Resampling Analysis:

- **Central** block of the image:
 - green channel
- **Small** size block:
 - 32×32 pixels

Experimental Results: AUC

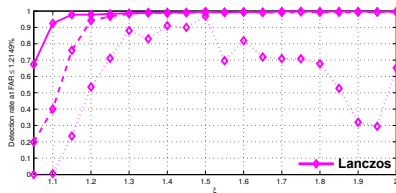
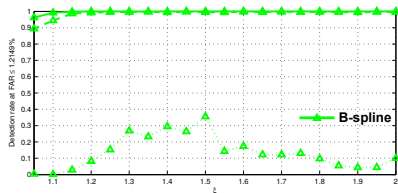
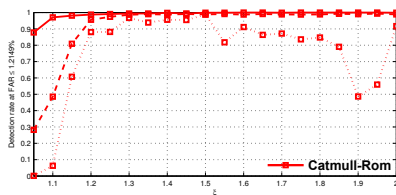
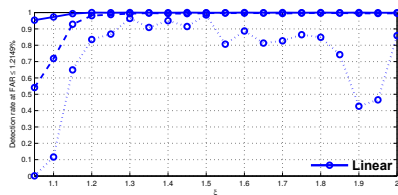
Area Under the ROC curve (AUC):



Detectors: LP-based - - - SVD-based — RMT-based (ours)

Experimental Results: Detection Rate

Detection rate at 1.2149% False Alarm Rate (FAR):



Detectors: LP-based - - - SVD-based — RMT-based (ours)

Conclusions

■ Key aspects of this work:

- 1 Processing $K < N$ consecutive columns/rows of an $N \times N$ block:
 - **eigenvalue compaction** property (Marčenko-Pastur Law)
- 2 Taking advantage of the **eigenvalue compaction** property:
 - **signal-plus-noise structure** becomes evident (asymptotic gap)
- 3 Using RMT tools for **small values** of N :
 - asymptotic limits turn out to be **excellent approximations**

■ Main contributions:

- Generalization of the **intuition behind** Marčenko-Pastur Law to **2D-AR random fields** (upscaled & quantized)
- Development of a detector that works with a **theoretical threshold**
- Promising performance with small size blocks (32×32 pixels)

■ Future research:

- Introduce a model for the demosaicing process
- Apply RMT tools for downscaling
- Address more realistic scenarios: JPEG, nonlinear filtering, etc.

Thanks for your attention!

Fernando Pérez-González
fperez@gts.uvigo.es