Novel Bayesian Multiscale Methods for Speckle Removal in SAR Images

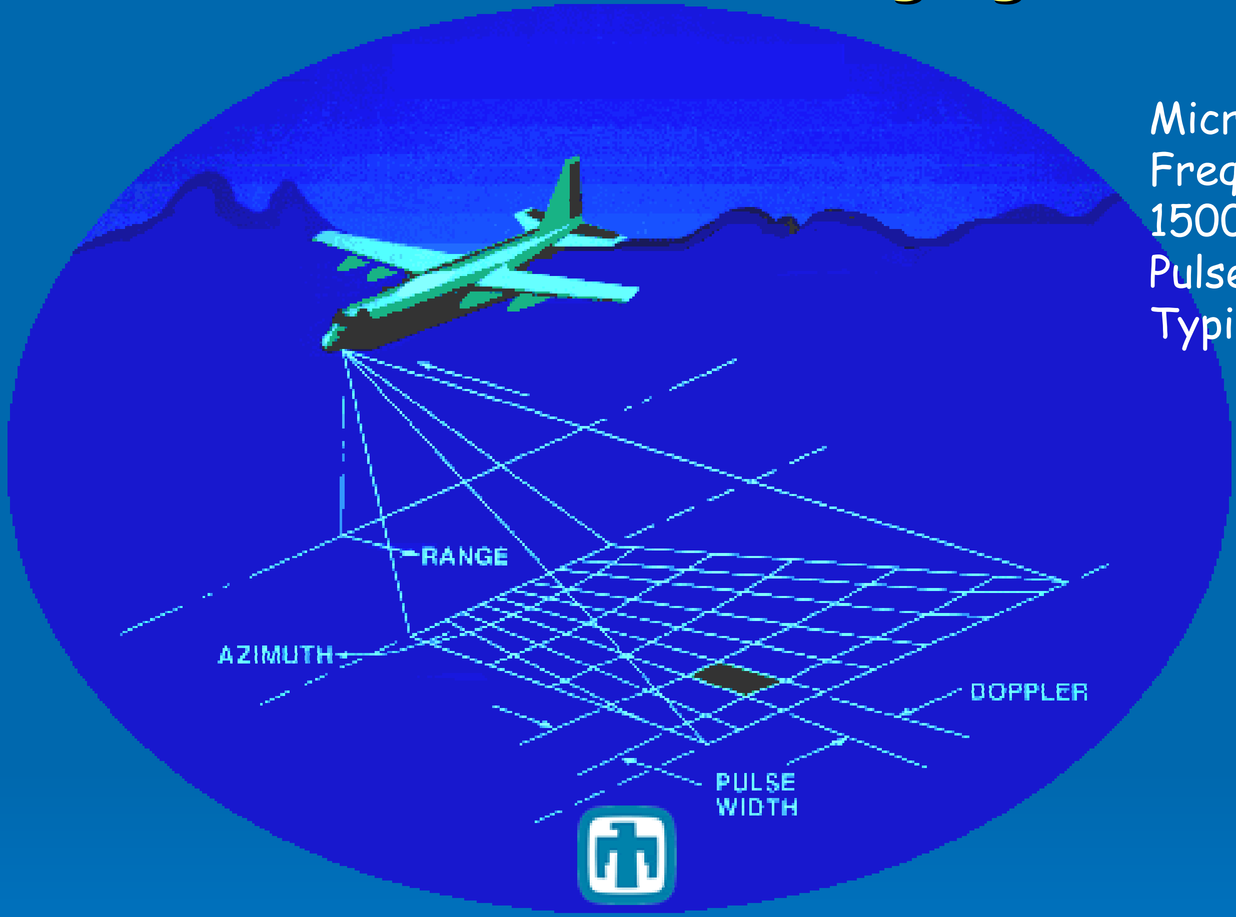
Panagiotis Tsakalides
Computer Science Department
University of Crete
Heraklion, Greece

tsakalid@ics.forth.gr

Presentation Outline

- Synthetic Aperture Radar (SAR) concept
- Heavy-tailed signals and non-Gaussian modeling
- Multiscale methods for SAR image processing
- A novel Bayesian processor for image denoising: The WIN-SAR algorithm
- SAR image denoising results

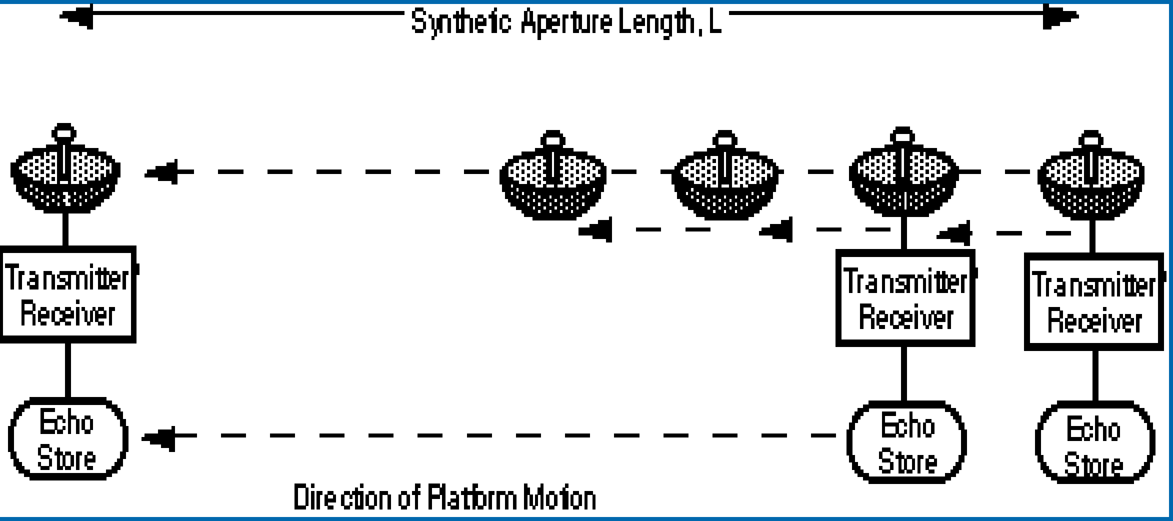
SAR Imaging Concept



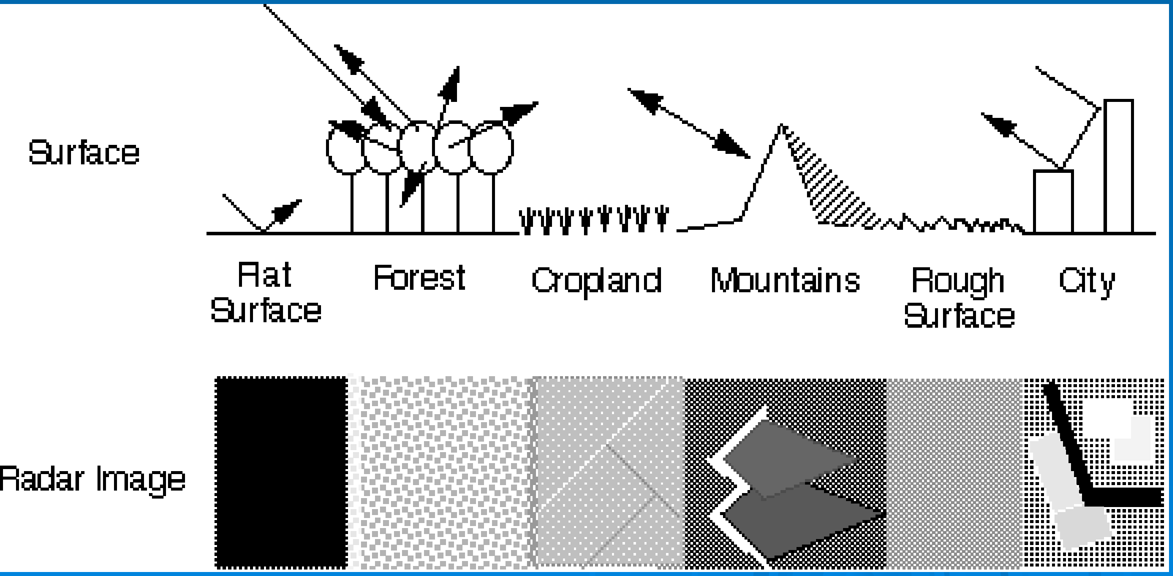
Microwave wavelenghts: 1 cm - 1 m
Frequency ranges: 300 MHz-30 GHz
1500 pulses/sec
Pulse duration: 10-50 μsecs
Typical Bandwidth: 10-200 MHz

SAR produces a two-dimensional (2-D) image. The *cross-track* dimension in the image is called *range* and is a measure of the "line-of-sight" distance from the radar to the target. The *along-track* dimension is called *azimuth* and is perpendicular to range.

SAR Imaging Concept



The length of the radar antenna determines the *resolution in the azimuth (along-track) direction* of the image: the longer the antenna, the finer the resolution in this dimension.



Each pixel in the image represents the radar *backscatter* for that area on the ground: objects approximately the size of the wavelength (or larger) appearing bright (i.e. rough) and objects smaller than the wavelength appearing dark (i.e. smooth)

SAR Qualities

- Synthetic aperture radar (SAR) systems take advantage of the long-range propagation characteristics of radar signals and the complex information processing capability of modern digital electronics to provide high resolution imagery.

- SAR complements photographic and other optical imaging capabilities because of the minimum constraints on time-of-day and atmospheric conditions and because of the unique responses of terrain and cultural targets to radar frequencies.

SIR-C/X-SAR Image of LA (space shuttle Endeavour, Oct. 1994)

- Area: 62x32 sq. miles
- L-band (24 cm) radar channel
- Horizontal polarizarion

- Very dark grey: Pacific ocean, LAX, freeway system.
- Dark grey: mountain slops
- Lighter grey: suburban areas, low-density housing
- Bright white: high-rise buildings and housing alligned parallel to radar flight track

- Can be used to map fire scars in areas prone to brush fires, such as Los Angeles

SAR Imaging Applications

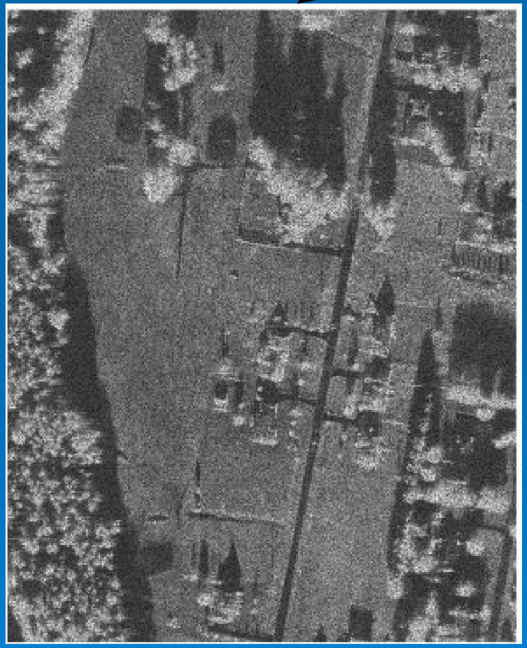
- **Civilial Applications**
  - High resolution remote sensing for mapping
  - Surface surveillance for search and rescue
  - Terrain structural information to geologists for mineral exploration
  - Sea state and ice hazard maps to navigators
  - Environmental monitoring:
    - Oil spill boundaries on water
    - Changes in delicate ecosystems
    - Air pollution monitoring in urban areas
    - Administration of natural resources

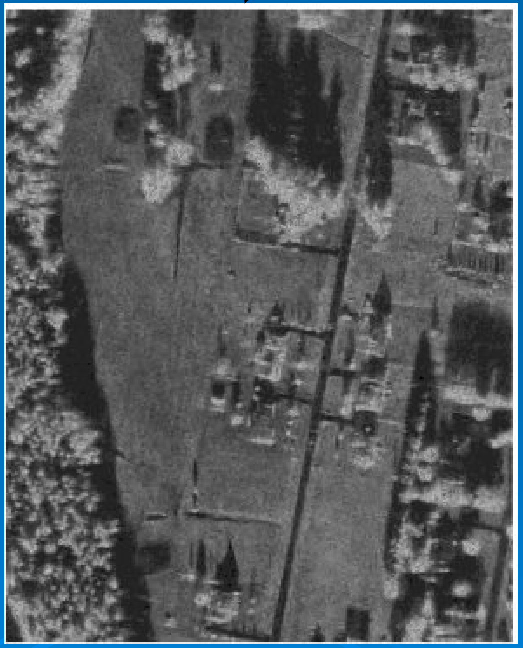
- **Military Applications**
  - Battlefield intelligence: detection and identification of potential targets to infer enemy capabilities, tactics and strategies
  - Mine detection
  - Automatic target recognition (ATR)

Problem

SAR images are inherently affected by multiplicative speckle noise, due to the coherent nature of the scattering phenomenon:

$$I(x,y) = S(x,y) \cdot \eta_m(x,y)$$





Speckle Noise (multiplicative): unit-mean, log-normal distributed.

Need to balance between speckle suppression and signal detail preservation!!!

Symmetric Alpha-Stable (SαS) Processes:

A (fairly) New Statistical Signal Processing Framework

Quotation

"The tyranny of the normal distribution is that we run the world … by attributing average levels of competence to the whole population.

What really matters is what we do with the tails of the distribution rather that the middle."

R. X. Cringely
Accidental Empires, 1992

It can also be said about least-squares in signal processing.

The Symmetric Alpha-Stable (SαS) Model

SαS Characteristic Function:

$$\phi(\omega) = e^{j\delta\omega - \gamma|\omega|^{\alpha}}$$

α: characteristic exponent, $0<\alpha<=2$ *(determines thickness of the distribution tails, $\alpha=2$: Gaussian, $\alpha=1$: Cauchy)*

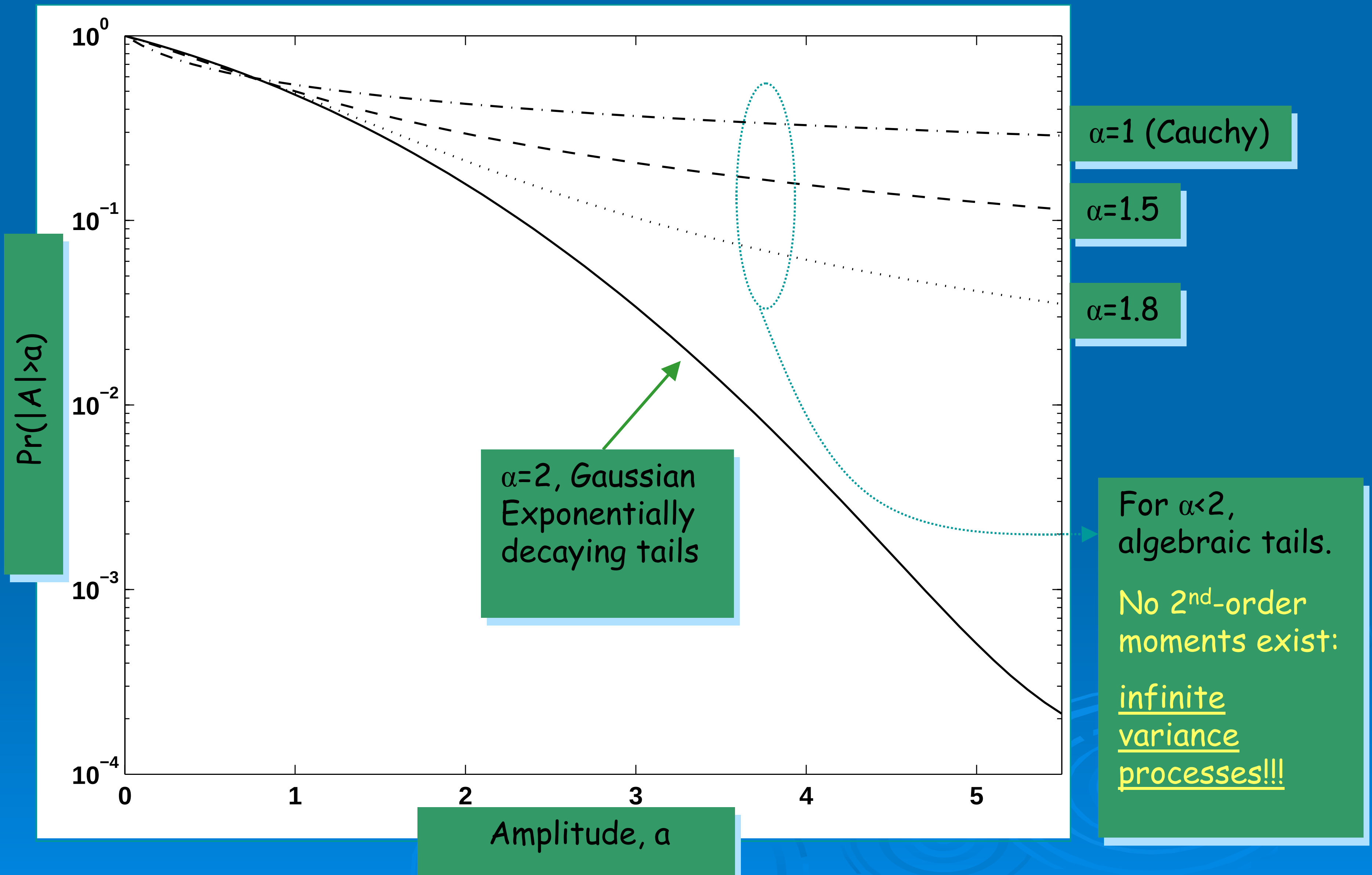
δ: location parameter *(determines the pdf's point of symmetry)*

γ: dispersion parameter, $\gamma>0$ *(determines the spread of the distribution around its location parameter)*

for Gaussian → γ = 2 x variance

for Cauchy → γ behaves like variance

SaS Probability Functions



Properties of SαS Laws

- Naturally arise as limiting processes via the Generalized Central Limit Theorem.
- Possess the stability property: The shape of a SαS r.v. is preserved up to a scale and shift under addition.
- Contain Gaussian (α=2) and Cauchy (α=1) distributions as members.
- Have heavier tails than the Gaussian: Their tail probabilities are asymptotically power laws → More likely to take values far away from the median ("Noah effect"):

$$P(X > x) \sim c_{\alpha} x^{-\alpha} \quad \text{as} \quad x \to \infty$$

Properties of SαS Laws

- Have finite *p*-order moments only for p<α:

$$Ex^p \prec \infty \quad \text{for} \quad p \prec \alpha$$

- Do not have finite second-order moments or variances:

$$Ex^2 = \infty$$

- Are self-similar processes: Exhibit long-range dependence or long memory ("Joseph effect").

Key Question!

- Since the variance if associated with the concept of power, are infinite variance distributions inappropriate for signal modeling and processing??

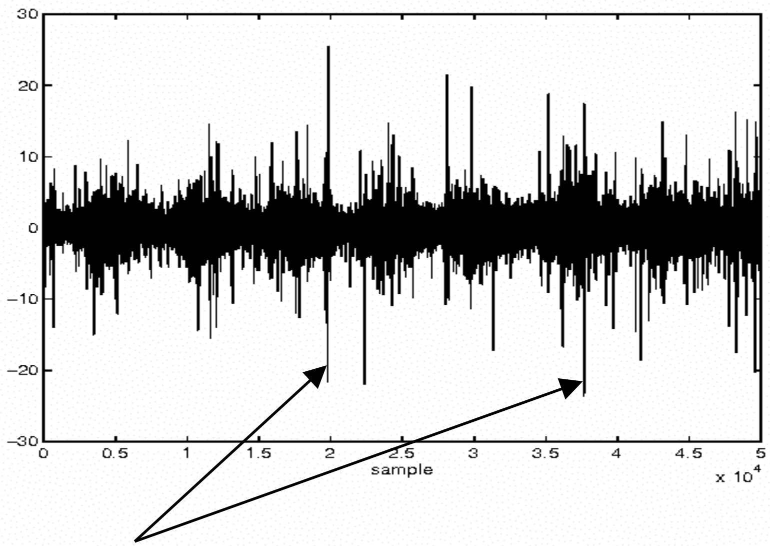
- No!! Variance is only one measure of spread! What really matters is an accurate description of the shape of the distribution. Particularly true when outliers appear in the data.

- Note that bounded data are routinely modeled by the Gaussian distribution, which has infinite support.

Real Data Modeling

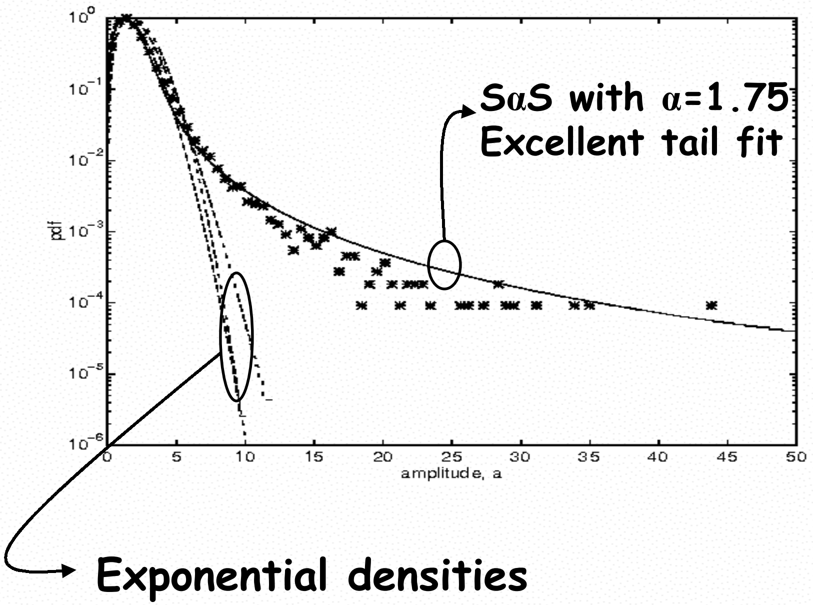
<u>Real sea clutter @ nominal sea condition:</u>

- sea state 3
- X-band radar
- 8o look-down angle
- spatial resolution of 1.52 m (5 ft)
- sampled at 40 Hz

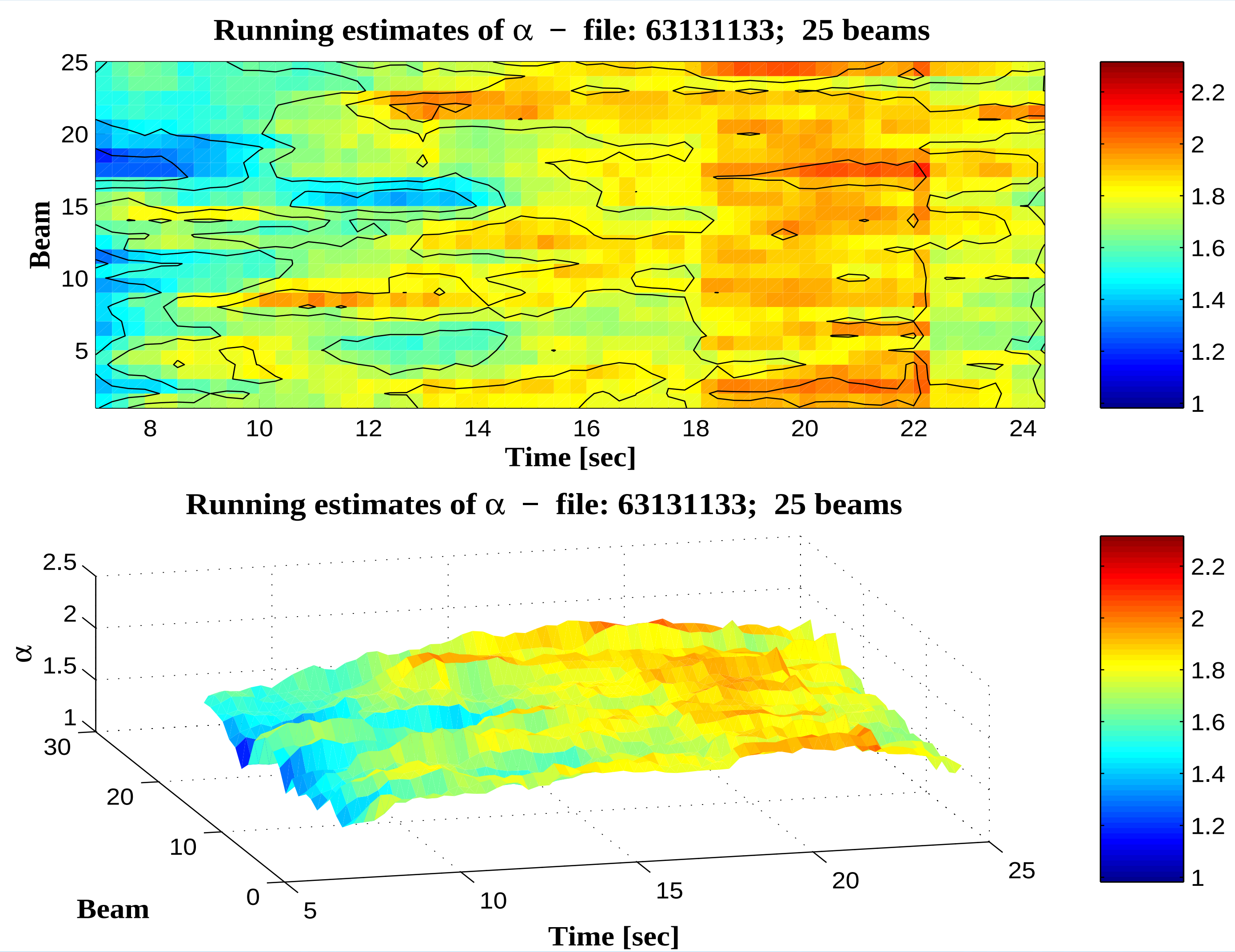


The impulsive nature of the clutter data is obvious.

<u>Clutter probability density modeling</u>



Real Data Modeling



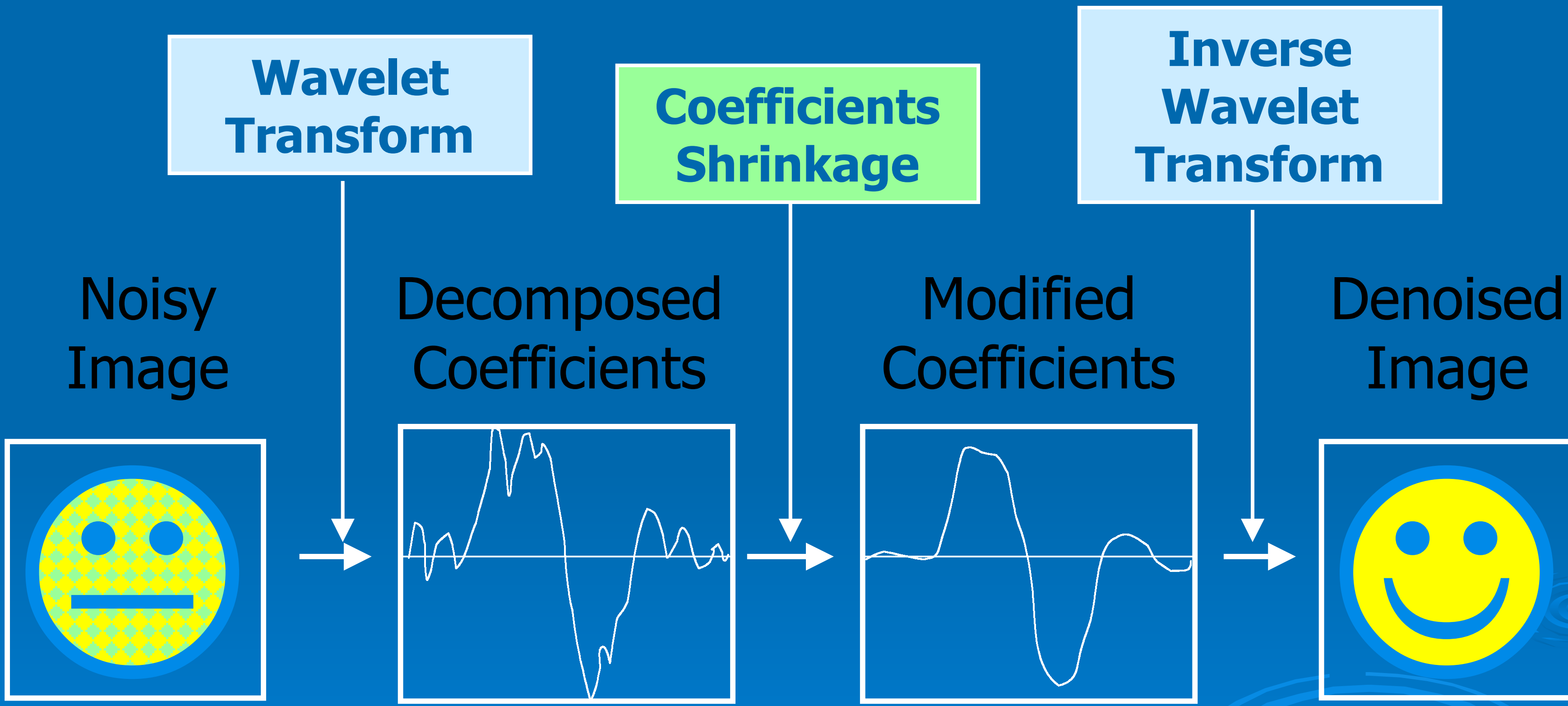
SαS Applications

- Astronomy (Holtzmark, 1919)
- Economic Time Series (Mandelbrot 60's, McCulloch 90's)
- Statistics (Zolotarev, Cambanis, Taqqu, Koutrouvelis, 70's-90's)
- Modeling of Signals and Noise:
  - Radar clutter (Tsakalides and Nikias, 1995)
  - Underwater Noise (Tsakalides and Pierce, 1997)
  - Communications Applications:
    - Telephone line noise (Stuck and Kleiner, 1974)
    - Fading in mobile systems (Hatzinakos and Llow, 1997)
    - Traffic modeling over comm. nets (Taqqu, 1996 – Petropulu, 2002)
  - Multimedia Applications:
    - Modeling, compression, watermarking, and image restoration in the DCT and Wavelet transform domains (Tsakalides et al., 1999-2002)

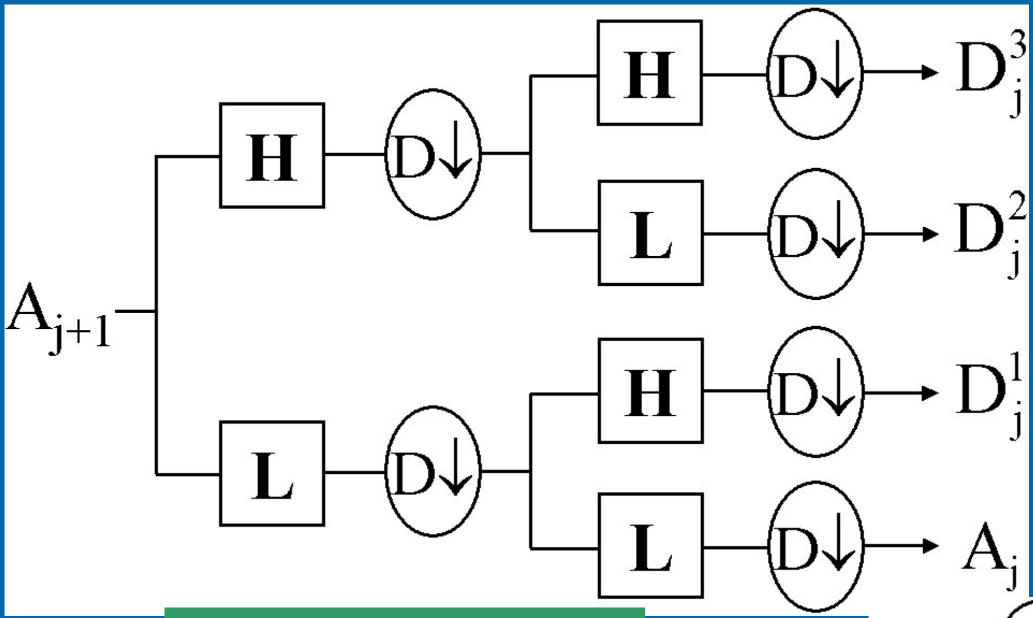
Multiscale methods for SAR image processing:

The Wavelet-based Image-Denoising Nonlinear SAR (WIN-SAR) Processor

Wavelets for Image Denoising

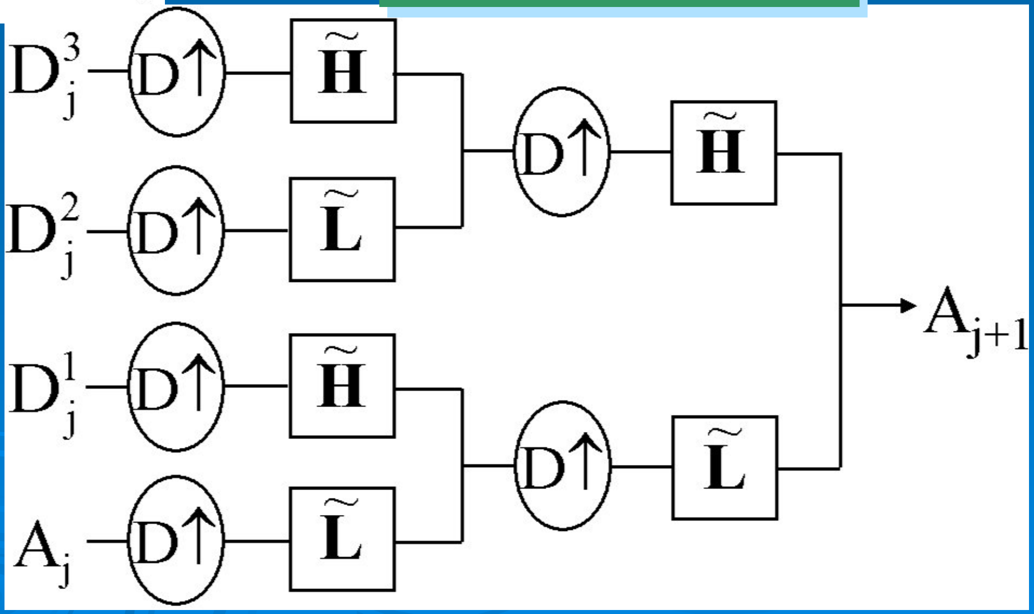


2-D Dyadic Wavelet Transform



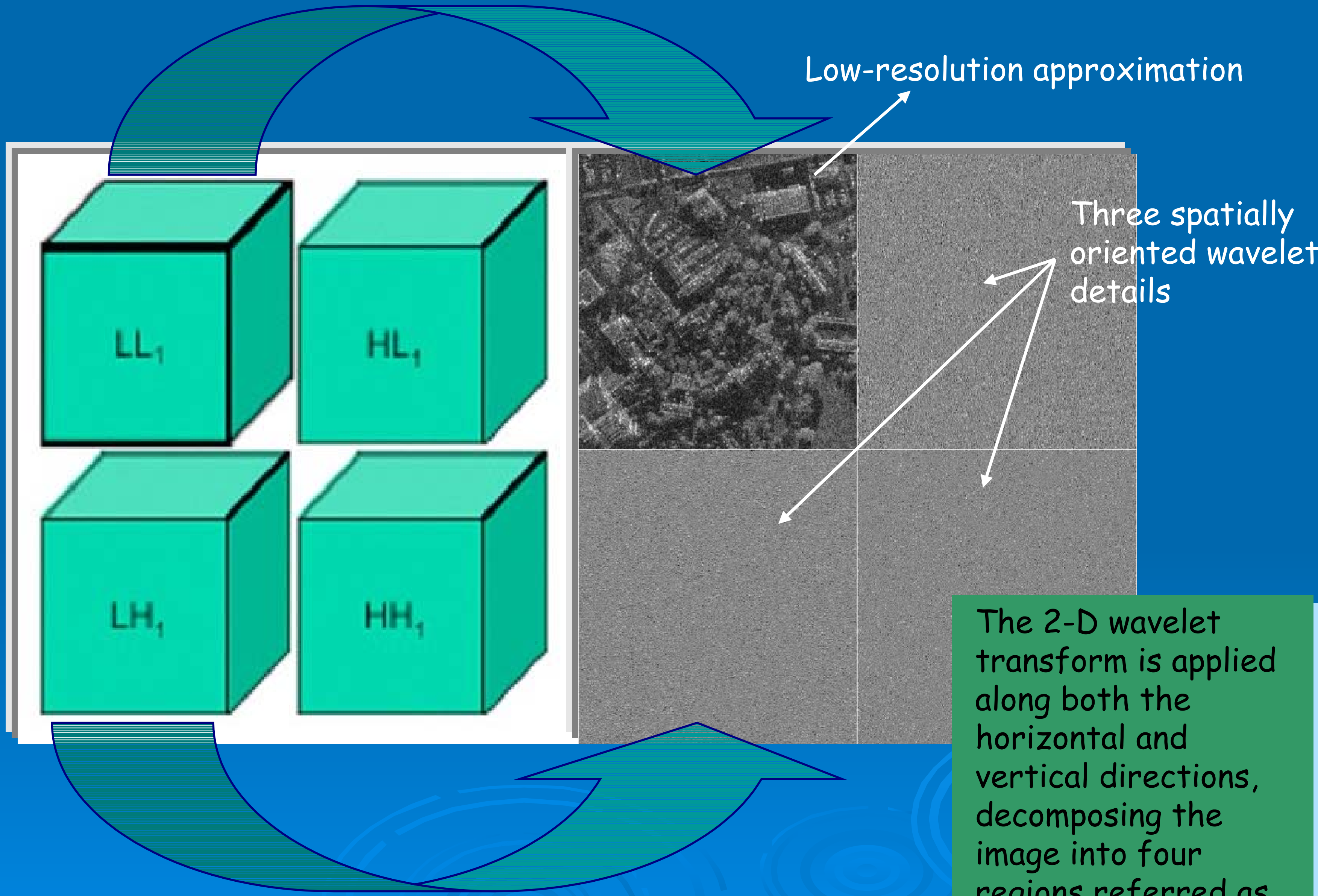
Decomposition of an image into an approximation and 3 detail subbands

Reconstruction of an image from its approximation and details



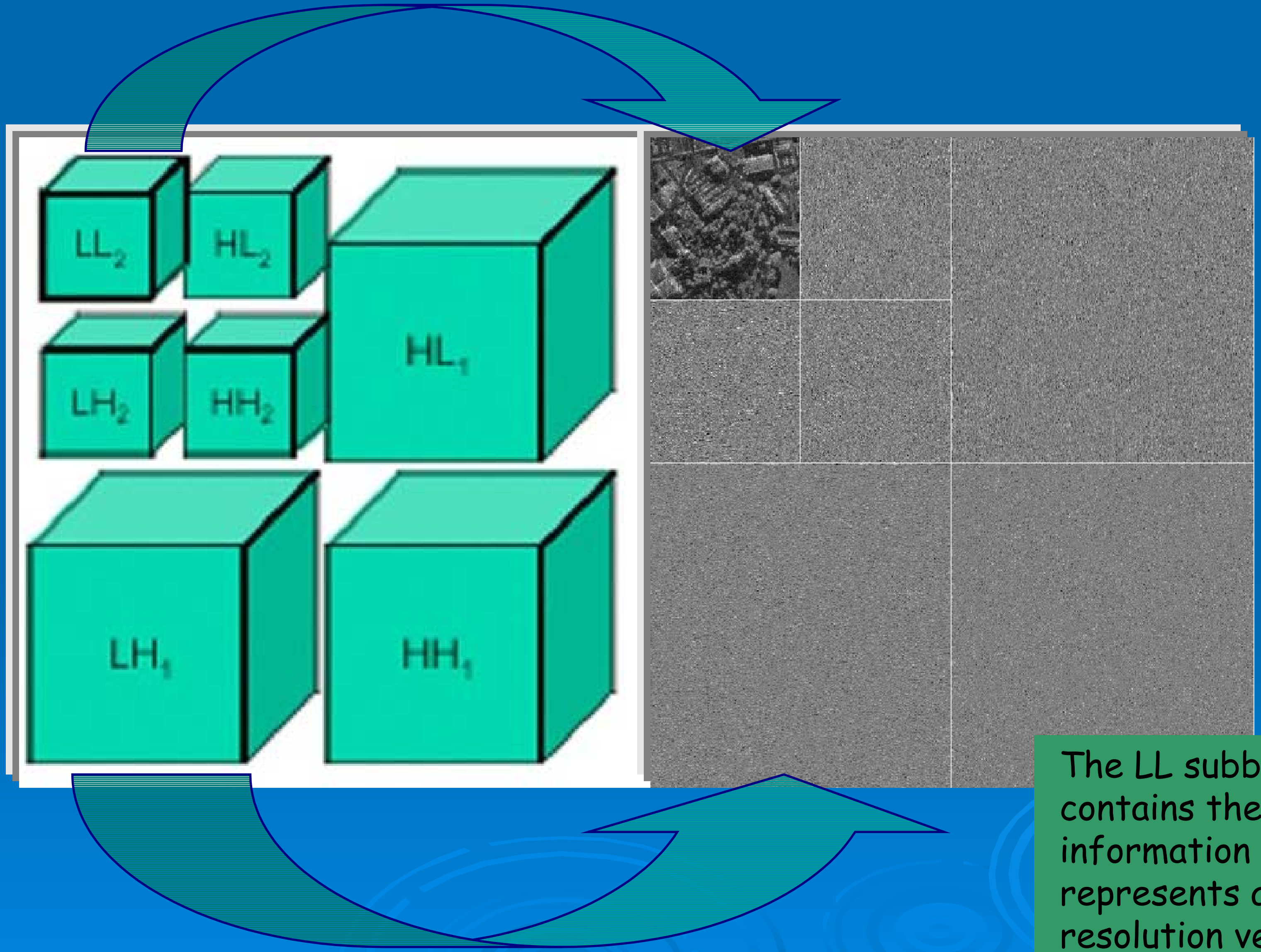
Expand a signal using a set of basis functions obtained from a single prototype: the "mother wavelet."
Result: A sequence of signal approximations a successively coarser resolutions.

Multiresolution decomposition – 1st level



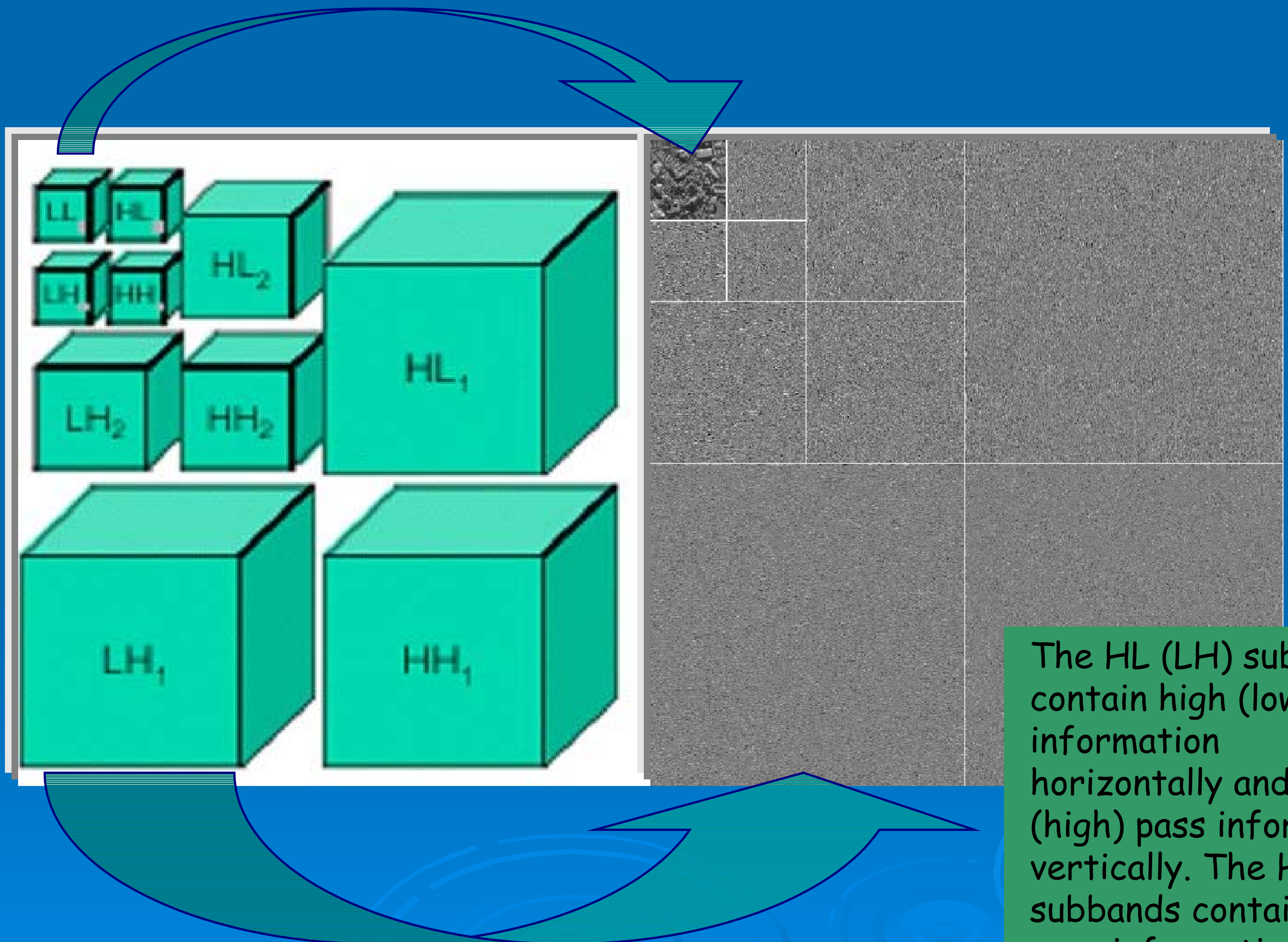
The 2-D wavelet transform is applied along both the horizontal and vertical directions, decomposing the image into four regions referred as image subbands.

Multiresolution decomposition – 2nd level



The LL subband contains the low-pass information and it represents a low resolution version of the original image.

Multiresolution decomposition – 3rd level



The HL (LH) subbands contain high (low) pass information horizontally and low (high) pass information vertically. The HH subbands contain high-pass information in both directions.

Previous Work in Wavelet-based Image Denoising

- Donoho's pioneering work: "Denoising by soft-thresholding" IEEE Trans. Inf. Theory 1995
- Simoncelli's "Noise removal via Bayesian wavelet coding," 1996
- Gagnon & Jouan's wavelet coefficient shrinkage (WCS) filter, 1997
- Simoncelli's work on texture synthesis, 1999
- Sadler's multiscale point-wise product technique, 1999
- Achim's work on heavy-tailed modeling, 2001
- Pizurica's work on inter & intra-scale statistical modeling, 2002

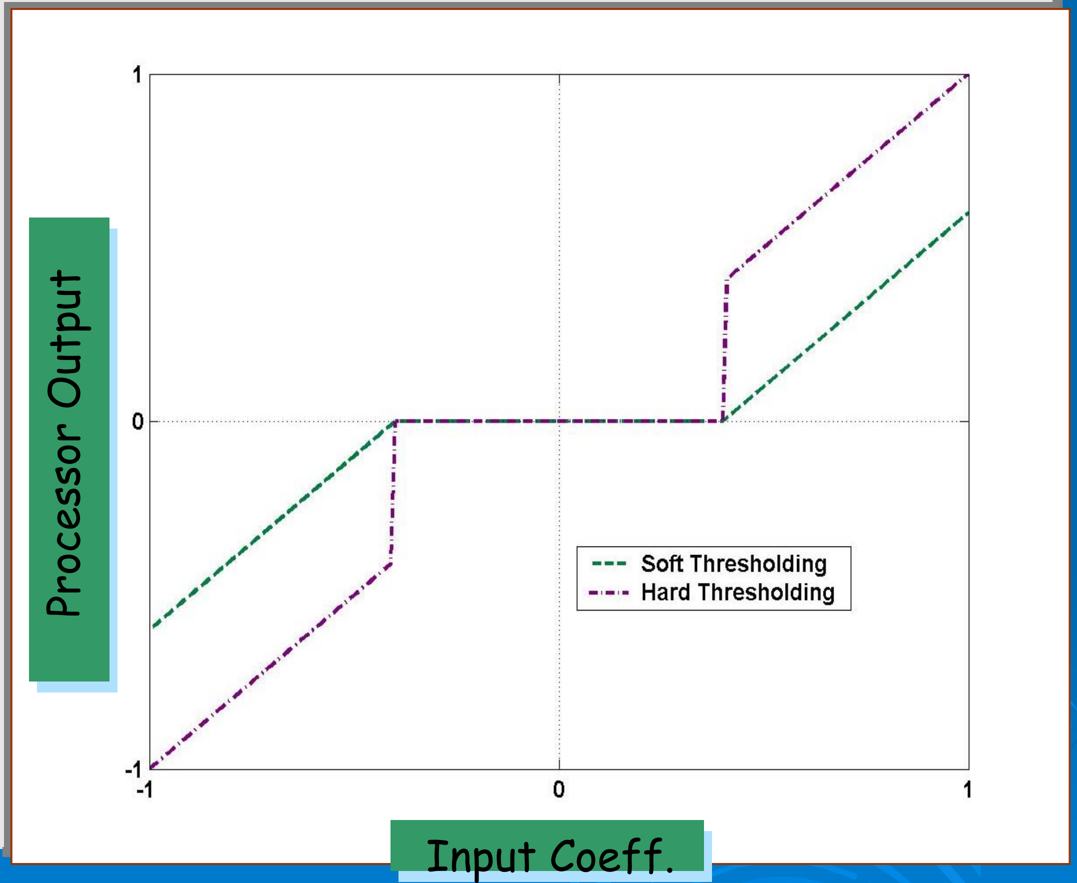
Wavelet Shrinkage Methods

- Soft Thresholding

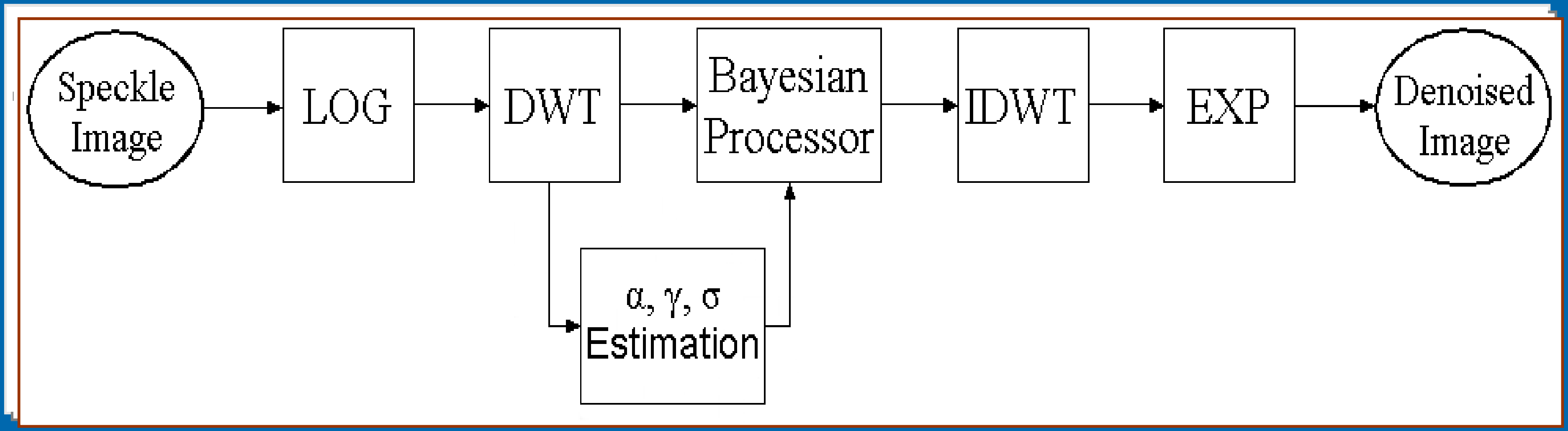
$$T_s^{soft}(s) = \begin{cases} \operatorname{sgn}(s)(|s| - t), & |s| > t \\ 0, & |s| \le t \end{cases}$$

- Hard Thresholding

$$T_s^{hard}(s) = \begin{cases} s, & |s| > t \\ 0, & |s| \le t \end{cases}$$



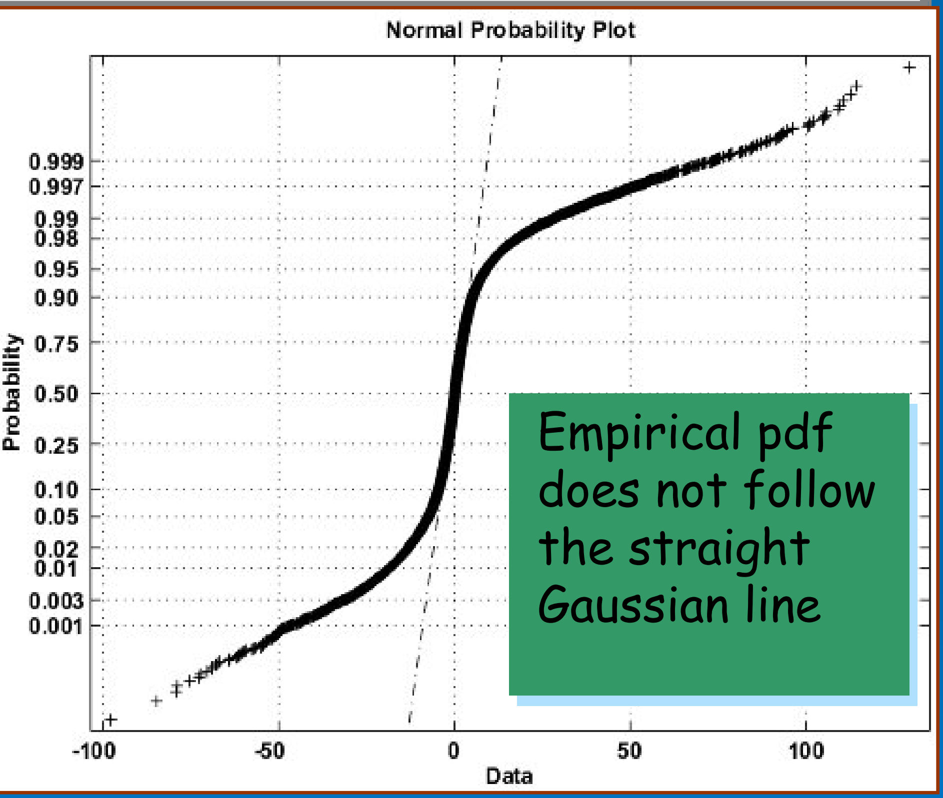
The WIN-SAR Processor

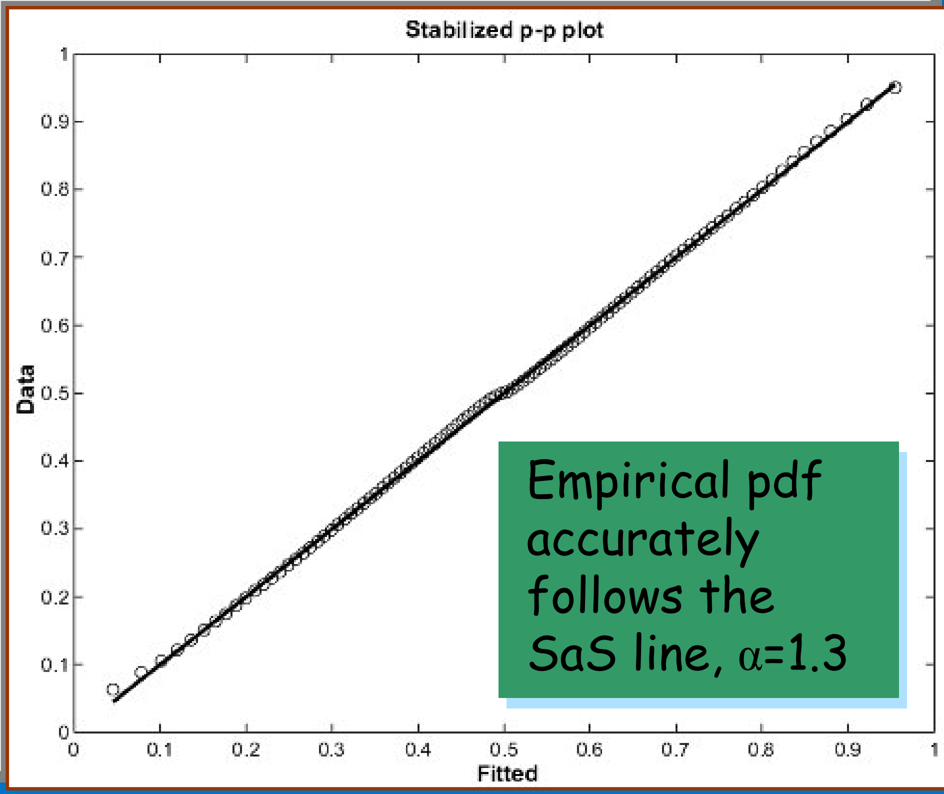


WIN-SAR fundamentals:

1. Wavelet transform the speckle SAR image.
2. SaS modeling of signal wavelet coefficients.
3. Bayesian processing of the coefficients in every level of decomposition.

Wavelet Coefficients Modeling (1)

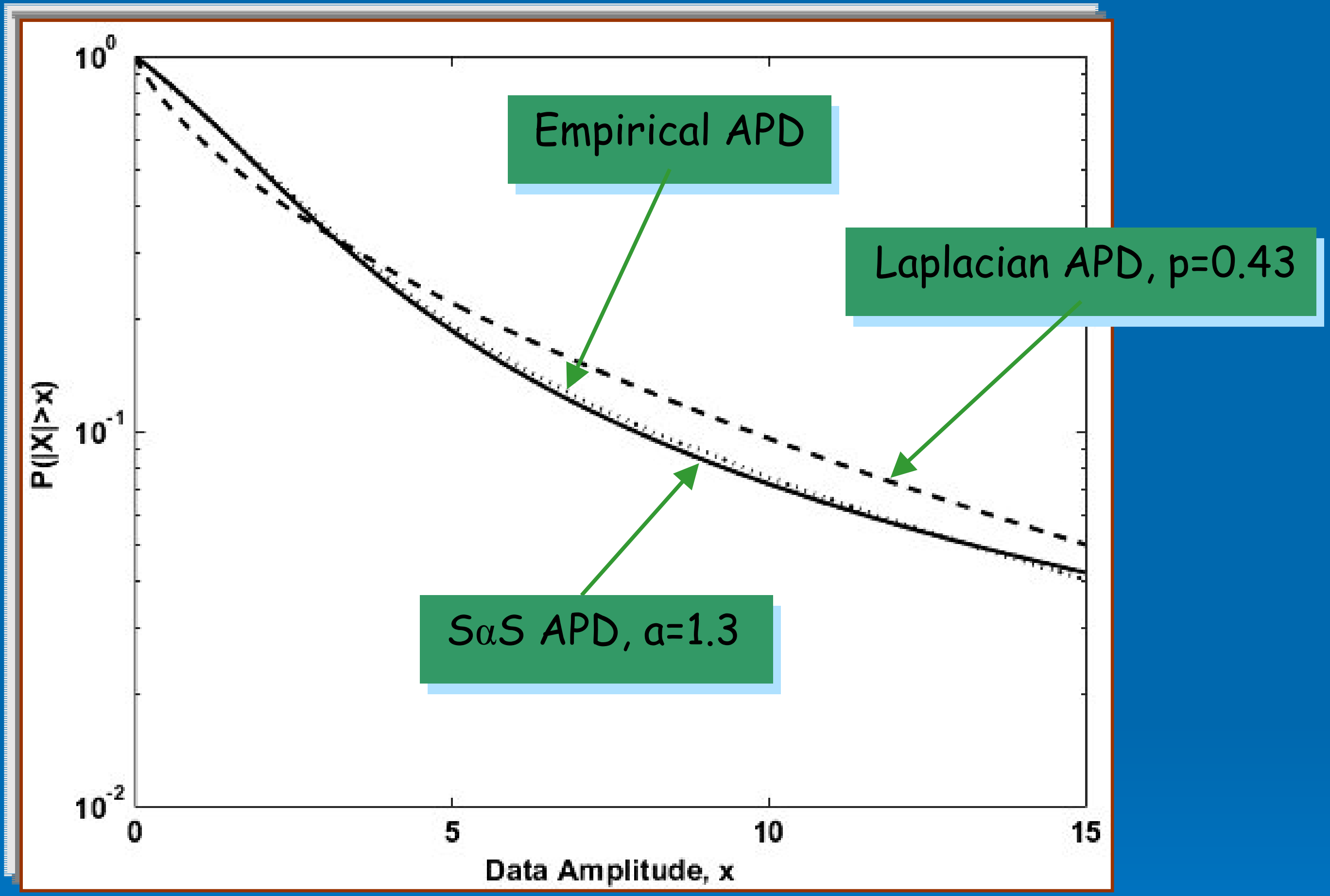




Normal and SαS probability plots of the vertical subband at the first level of decomposition of the image HB06158 from the MSTAR* collection.

* http://www.mvlab.wpafb.af.mil/public/sdms/

Wavelet Coefficients Modeling (2)



Amplitude Probability Density (APD) plot for the data of the previous slide: The SaS provides an excellent fit to both the **mode** and the **tails** of the empirical distribution.

SaS Modeling of Wavelet Subband Coefficients

Level	Image Subbands		
	Horizontal	Vertical	Diagonal
1	1.239	1.283	1.302
2	1.418	1.125	1.295
3	1.286	1.019	1.380

The tabulated key parameter α defines the degree of non-Gaussianity as deviations from the value $\alpha = 2$.

The WIN-SAR MAE Bayesian Estimator

- After applying the DWT: $\boldsymbol{d}^{i}_{j,k} = \boldsymbol{s}^{i}_{j,k} + \xi^{i}_{j,k}$

- The Bayes risk estimator of *s* minimizes the conditional risk, i.e., the loss function averaged over the conditional distribution of *s* given the measured wavelet coeffs:

$$\widehat{s}(d) = \underset{\widehat{s}(d)}{\operatorname{argmin}} \int \left| s - \widehat{s}(d) \right| \cdot P_{s|d}(s \mid d) \cdot ds$$

- The <u>mean absolute error</u> (MAE) estimator is the conditional median of s, given d, which coincides with the conditional mean (due to the symmetry of the distributions):

$$\widehat{s}(d) = \int s \cdot P_{s|d}(s \mid d) \cdot ds = \frac{\int P_{d/s}(d/s) P(s) s \cdot ds}{\int P_{d/s}(d/s) P(s) \cdot ds}$$

The WIN-SAR MAE Bayesian Estimator

$$\widehat{s}(d) = \frac{\int P_\xi(d-s)P(s)s \cdot ds}{\int P_\xi(d-s)P(s) \cdot ds} = \frac{\int P_\xi(\xi)P(s)s \cdot ds}{\int P_\xi(\xi)P(s) \cdot ds}$$

- Signal Parameter Estimation - by means of a LS fitting in the characteristic function domain:

$$\{\hat{a}_s, \hat{\gamma}_s, \hat{\sigma}\} = \underbrace{\arg\min}_{\hat{a}_s, \hat{\gamma}_s, \hat{\sigma}} \sum_i^n \left[\Phi_d(\omega_i) - \Phi_{d_e}(\omega_i)\right]^2$$

where:

$$\Phi_d(\omega) = \exp(-\gamma_s |\omega|^{\alpha_s}) \cdot \exp(-\frac{\sigma^2}{2}|\omega|^2)$$

WIN-SAR MAE Processor I/O Curves

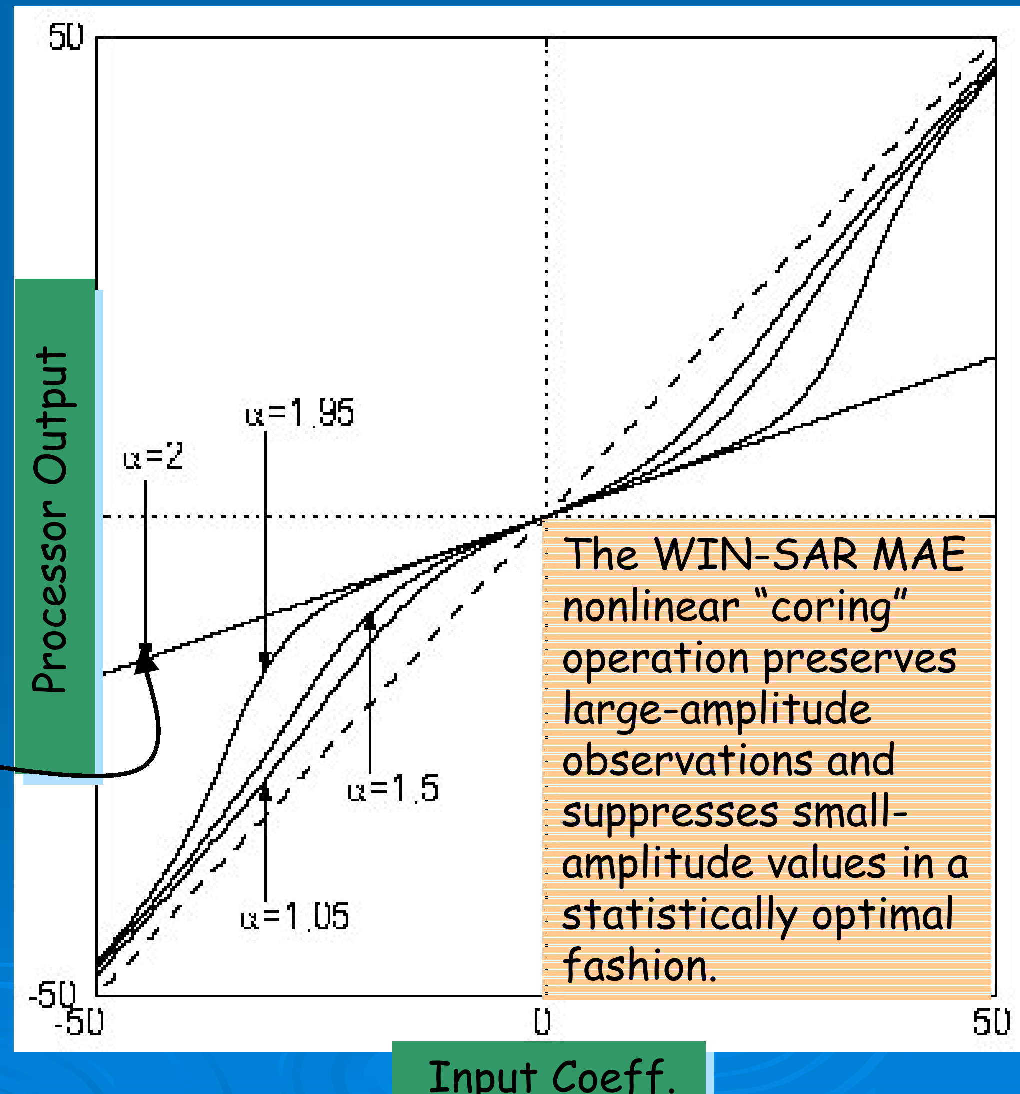
Bayesian Processing:

$$\hat{s}(d) = \frac{\int P_{\xi}(\xi)P(s)s \cdot ds}{\int P_{\xi}(\xi)P(s) \cdot ds}$$

Only for α=2 (Gaussian signal), the processing is a simple linear rescaling of the measurement:

$$\hat{s}(d) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2} d$$

For a given ratio γ/σ, the amount of shrinkage decreases as α decreases: The smaller the value of α, the heavier the tails of the signal PDF and the greater the probability that the measured value is due to the signal.



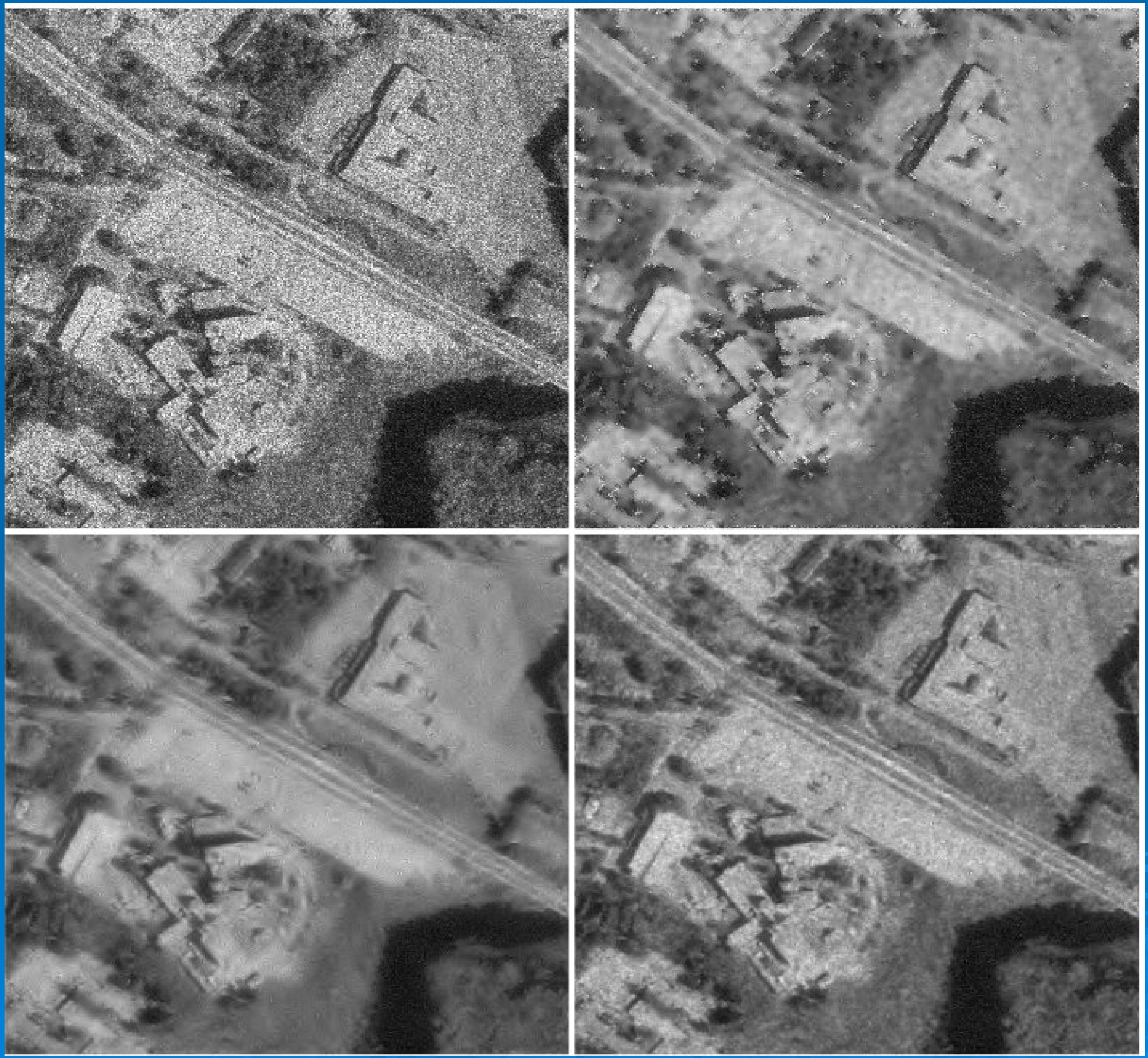
Real SAR Imagery Results (1)

Noisy Image

MSE	43.8
β	0.25
s/m	0.49

GMAP

MSE	17.1
β	0.43
s/m	0.37



Soft Thresholding

MSE	17.3
β	0.36
s/m	0.32

WIN-SAR

MSE	16.2
β	0.54
s/m	0.35

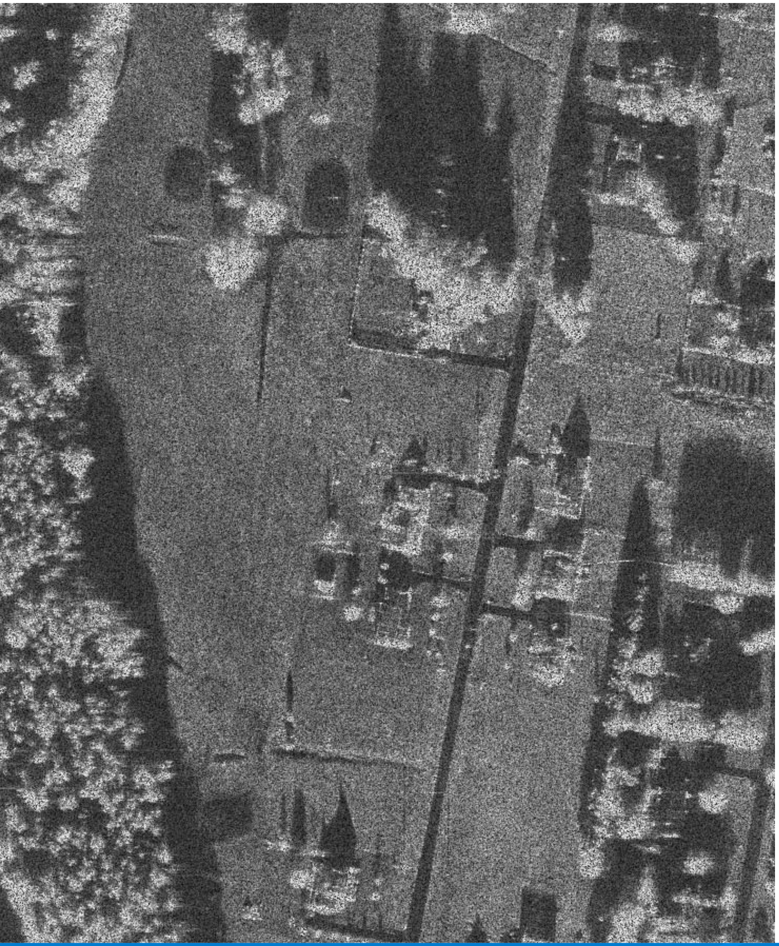
Real SAR Imagery Results (2)

Urban scene (dense set of large cross-section targets w. intermingled tree shadows

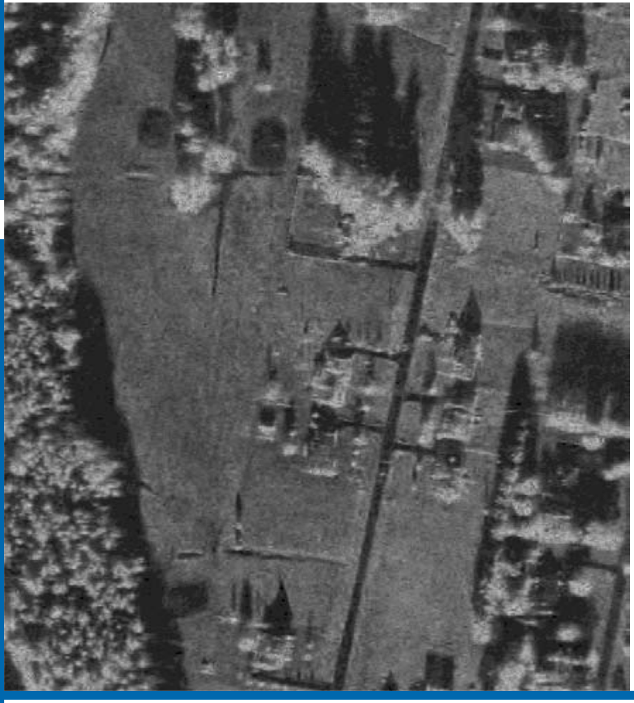
WIN-SAR

Shoft-Thresholding

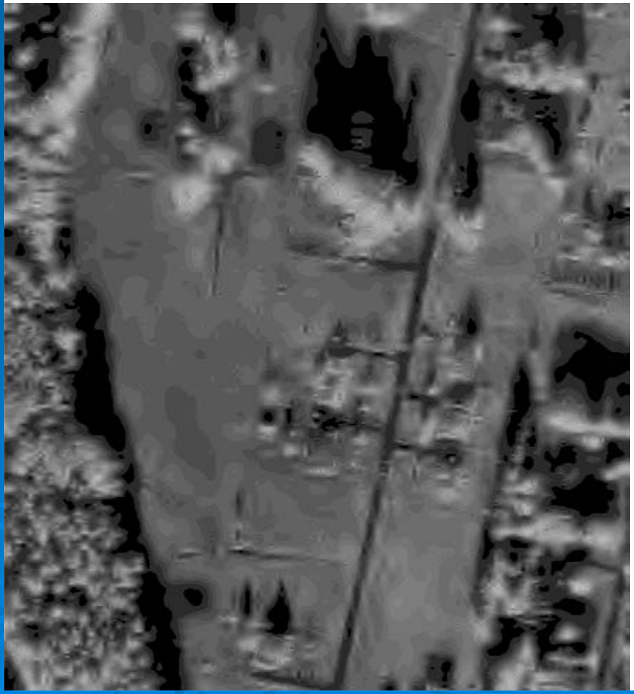
Real SAR Imagery Results (3)



Rural scene



WIN-SAR



Shoft-Thresholding

Conclusions

1. Introduced a new statistical representation for wavelet coefficients of SAR images.
2. Designed and tested Bayesian processors and found them more effective than traditional wavelet shrinkage methods, both in terms of speckle reduction and signal detail preservation.
3. Proposed processors based on solid statistical theory: do not depend on the use of any *ad hoc* thresholding parameter.
4. Future work: Analyze multiscale products for step detection and estimation.

Related Publications

1. A. Achim, A. Bezerianos, and P. Tsakalides, "SAR Image Denoising via Bayesian Wavelet Shrinkage based on Heavy-Tailed Modeling," *IEEE Transactions on Geoscience and Remote Sensing*, submitted for publication consideration, July 2002.
2. P. Tsakalides and C. L. Nikias, "High Resolution Autofocus Techniques for SAR Imaging based on Fractional Lower-Order Statistics," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 148, no. 5, pp. 267-276, October 2001.
3. A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian Multiscale Methods for Speckle Removal in Medical Ultrasound Images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 772-783, August 2001.
4. P. Tsakalides, R. Raspanti, and C. L. Nikias, "Angle/Doppler Estimation in Heavy-Tailed Clutter Backgrounds," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 419-436, April 1999.
5. P. Tsakalides and C. L. Nikias, "Robust Space-Time Adaptive Processing (STAP) in Non-Gaussian Clutter Environments," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 146, no. 2, pp. 84-94, April 1999.

The WIN-SAR MAP Bayesian Estimator

- The *MAP* estimator is the Bayes risk estimator under an uniform cost function:

$$\hat{s}\left(d\right)=\arg\max_{\hat{s}} P_{s|d}(s|d)=\arg\max_{\hat{s}} P_{d|s}(d|s)P(s)=$$

$$=\arg\max_{\hat{s}} P_{\xi}(d-s)P_{s}(s)=\arg\max_{\hat{s}} P_{\xi}(\xi)P_{s}(s)$$

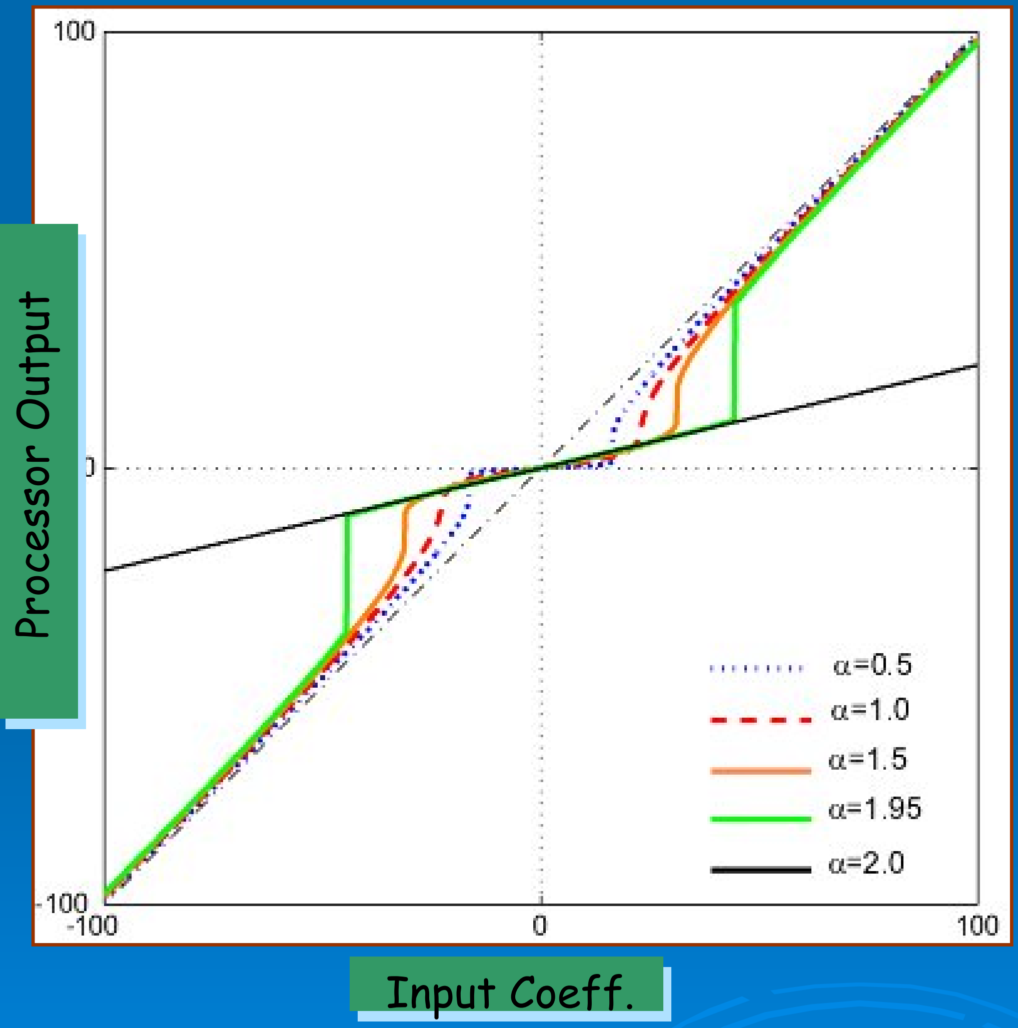
- Parameter estimation method: After estimating the level of noise σ we find the parameters α_s and γ_s by regressing

$$y=\log\left[-\left(\log\left|\Phi_{d}(\omega)\right|^{2}+\sigma^{2}\omega^{2}\right)\right]$$

on $w=\log|\omega|$ in the model: $y_k=\mu+\alpha\cdot w_k+\varepsilon_k$

where: $\mu=\log(2\gamma)$, ε_k - error term, and $(\omega_\kappa, \kappa=1,\ldots,\mathrm{K})\in \boldsymbol{R}$

WIN-SAR MAP processor I/O curves



The plots illustrate the processor dependency on the parameter α of the signal *prior* PDF. For a given ratio γ/σ, the amount of shrinkage decreases as α decreases. The intuitive explanation for this behavior is that the smaller the value of α, the heavier the tails of the signal PDF and the greater the probability that the measured value is due to the signal.