ROBUST STATISTICS FOR TESTING THE HOMOGENEITY OF COVARIANCE SCALE AND SHAPE

AN APPLICATION TO SAR CHANGE DETECTION

Seminar at NUS 10th April 2019

Ammar Mian

ABOUT THIS TALK

We consider the problem of SAR change detection through covariance matrix equality testing.

Approach based on mutivariate statistical analysis. Pre-requisites:

- Basic algebra
- Basic statistics

Slides are available at: https://ammarmian.github.io/

For questions: ammar.mian@centralesupelec.fr

PLAN OF THE PRESENTATION

1. The Change detection problem
2. Covariance-based approaches
3. Statistics of decision under Gaussian assumption
4. Extension to Elliptical distributions
5. Results on SAR images

THE CHANGE DETECTION PROBLEM

WHAT WE WANT TO DO ?

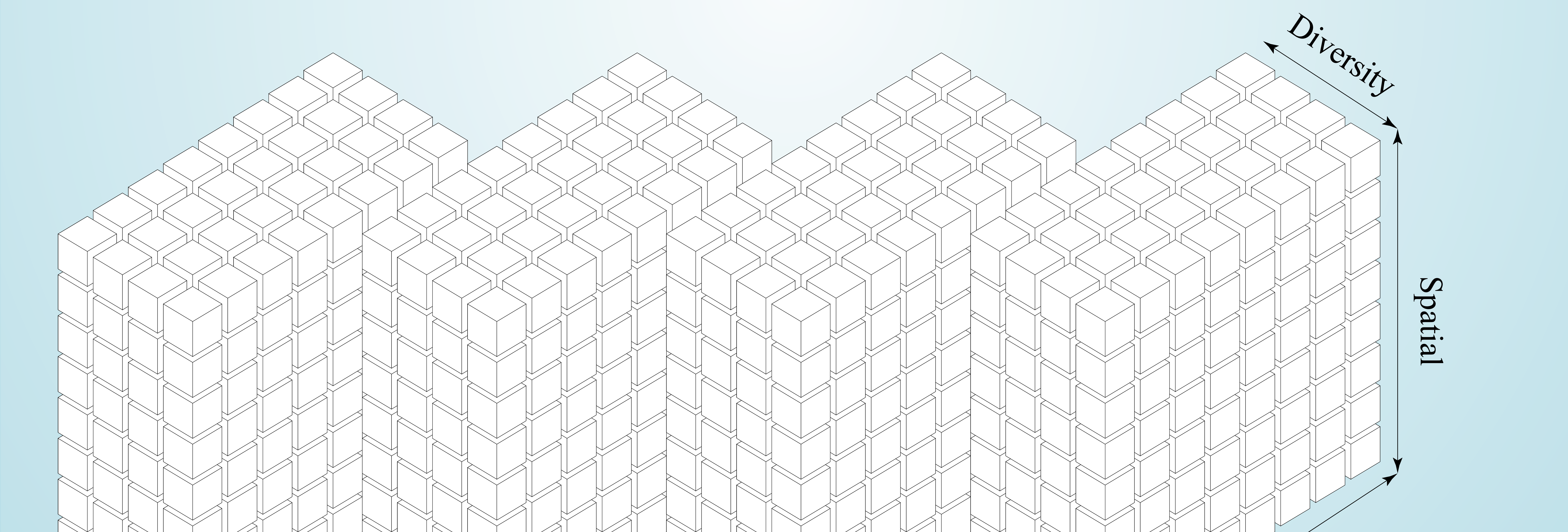
Considering a **time series** of SAR images, we want to **detect** spatial areas where notceable changes have occured.

Black RockCity before the Burning Man Festival, Nevada, USA - TerraSAR-X SpotLight Acquisition © 2011 Astrium Services / Infoterra GmbH

Black Rock City after the Burning Man Festival, Nevada, USA - TerraSAR-X SpotLight Acquisition © 2011 Astrium Services / Infoterra GmbH

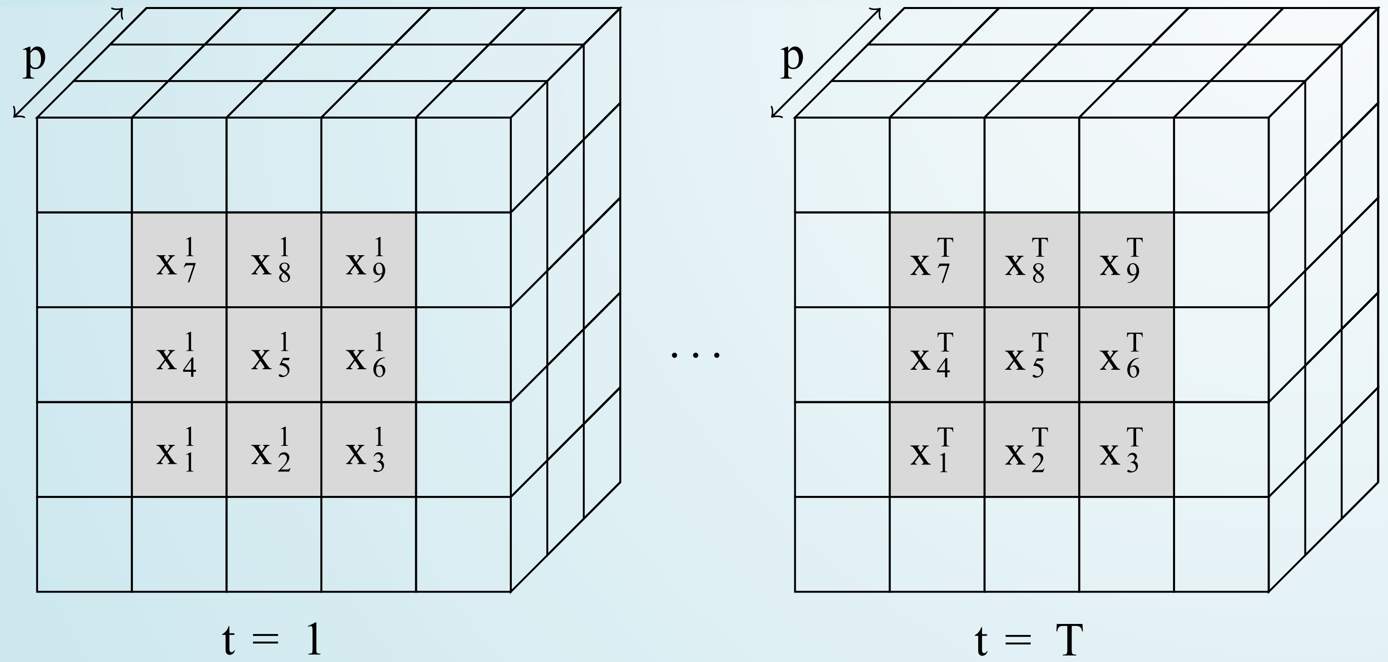
ASSUMPTIONS

- The images are co-registred
- Unsupervised methodology
- Data is complex and multivariate



COVARIANCE-BASED APPROACHES

DESCRIPTION OF THE APPROACH



We consider a sliding windows approach to consider spatially local data.

On this window, we assume i.i.d observations:

$$\forall (k,t)\, \mathbf{x}_k^t \sim \mathbb{C} \quad (\mathbf{0}_p, \boldsymbol{\Sigma}_t)$$

Idea: Compare the covariances $\boldsymbol{\Sigma}_t$ over time.

INTEREST OF THE APPROACH

- The SAR data is typically **noisy** (speckle noise) and can be **multivariate** (polarimetric for example).
- Covariance accounts for the local correlations between the observed noisy data. A change in the scene is likely to impact this matrix in a way.
- Using statistical detection theory allow to have theoretical results on false alarms or detection performance.

THE DETECTION PROBLEM

Denote by $\{ \quad _1,\ldots, \quad _T\}$ a collection of T mutually independent samples of i.i.d p-dimensional complex vectors: $\quad _t = [\mathbf{x}_1^t,\ldots,\mathbf{x}^t \quad]$.

We assume $\forall(k,t)$, $\mathbb{E}\{\mathbf{x}_k^t\} = \mathbf{0}_p$ and we denote $\boldsymbol{\Sigma}_t = \tau_t \boldsymbol{\xi}_t$ the shared covariance matrices among the elements of $\quad _t$. $\boldsymbol{\xi}_t$ is the shape matrix ($Tr(\boldsymbol{\xi}_t) = p$) and τ_t is the scale.

We want to choose between the following alternatives:

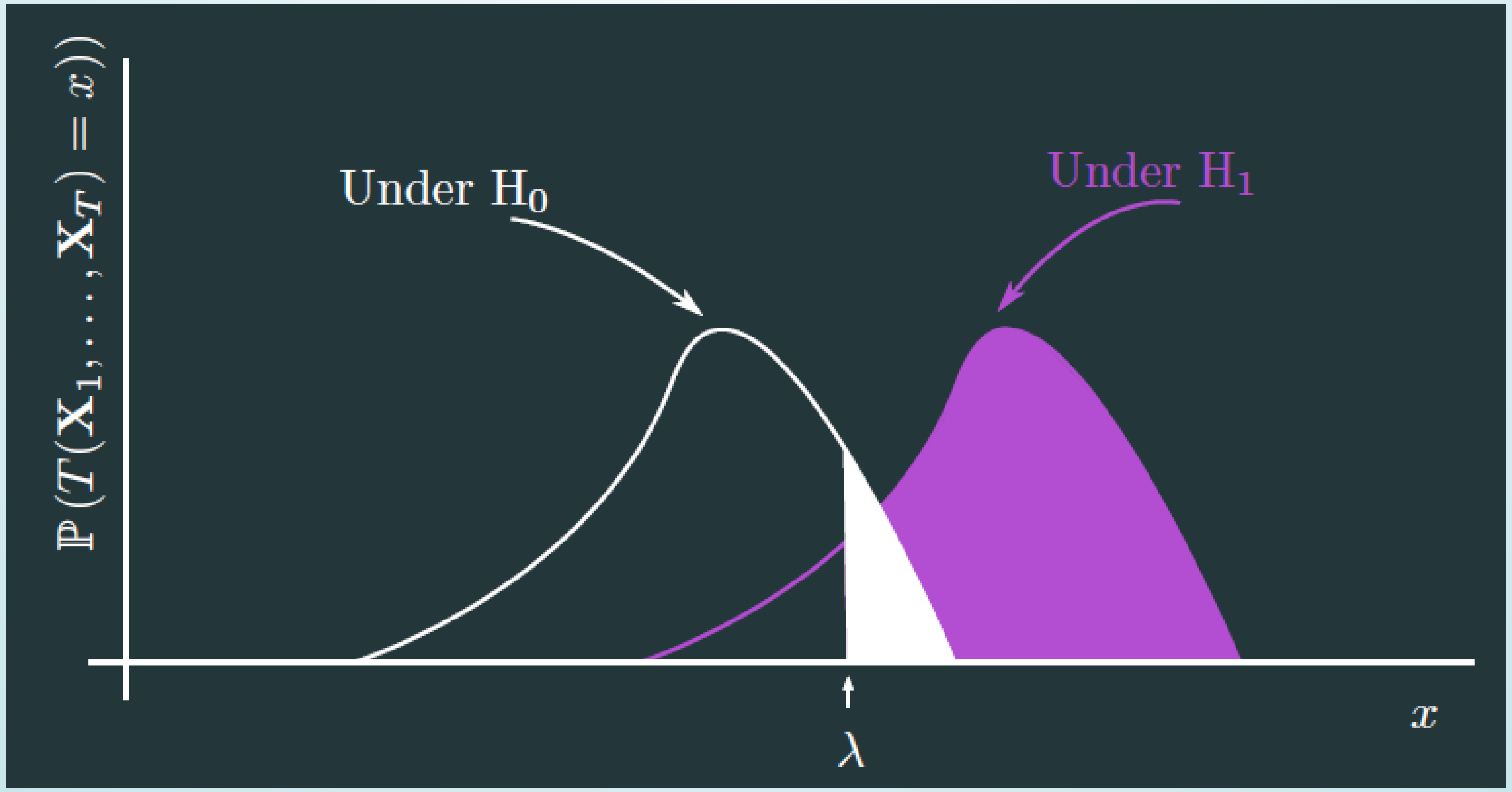
$$
\begin{aligned}
\mathrm{H}_0 &: \quad \boldsymbol{\Sigma}_1 = \ldots = \boldsymbol{\Sigma}_T = \boldsymbol{\Sigma}_0 \,, \\
\mathrm{H}_1 &: \quad \exists(t,t'),\ \boldsymbol{\Sigma}_t \quad \boldsymbol{\Sigma}_{t'}
\end{aligned}
$$

STATISTIC OF DECISION

We want to obtain:

- a statistic of decision T: $\begin{array}{rl} \mathbb{C}^{p\times} \times \cdots \times \mathbb{C}^{p\times} & \to \mathbb{R}^+ \\ {}_1, \ldots, {}_T & \to T({}_1, \ldots, {}_T) \end{array}$
- a threshold λ

So that $\mathbb{P}\,(T({}_1, \ldots, {}_T) > \lambda/\mathrm{H}_1)$ is high while $\mathbb{P}\,(T({}_1, \ldots, {}_T) > \lambda/\mathrm{H}_0)$ is low.



STATISTICS OF DECISION UNDER GAUSSIAN ASSUMPTION

STATISTICS OF DECISION UNDER GAUSSIAN ASSUMPTION (1/2)

Suppose $\forall t, \forall k, \mathbf{x}_k^t \sim \mathbb{C}\quad(\mathbf{0}_p, \mathbf{\Sigma}_t)$ so that $p_{\mathbf{x}_k^t;\mathbf{\Sigma}_t}(\mathbf{x}_k^t;\mathbf{\Sigma}_t) = \frac{1}{\pi^p|\mathbf{\Sigma}_t|}\mathrm{exp}\mathrm{tr}\left\{\mathbf{S}_k^t\mathbf{\Sigma}_t^{-1}\right\}$, where $\mathbf{S}_k^t = \mathbf{x}_k^t{\mathbf{x}_k^t}^{\mathrm{H}}$.

Many statistic exists but the options can be reduced to[1] :

- The Generalized Likelihood Ratio Test (GLRT) statistic:

$$\hat{\Lambda}_{\mathrm{G}} = \frac{\left|\hat{\mathbf{\Sigma}}_0^{\mathrm{SC}}\right|^T}{\quad_{t=1}^{T}\left|\hat{\mathbf{\Sigma}}_t^{\mathrm{SC}}\right|} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda, \text{ where :}$$

$$\forall t, \hat{\mathbf{\Sigma}}_t^{\mathrm{SC}} = \frac{1}{\quad}\quad_{k=1}\mathbf{S}_k^t \text{ and } \hat{\mathbf{\Sigma}}_0^{\mathrm{SC}} = \frac{1}{T}\quad_{t=1}^{T}\hat{\mathbf{\Sigma}}_t^{\mathrm{SC}}\quad.$$

[1] D. Ciuonzo, V. Carotenuto and A. De Maio, "On Multiple Covariance Equality Testing with Application

STATISTICS OF DECISION UNDER GAUSSIAN ASSUMPTION (2/2)

- The t_1 statistic obtained from either Terell or Rao tests:

$$\hat{\Lambda}_{\mathrm{t}_1} = \sum_{t=1}^{T} \mathrm{Tr}\left[\left(\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}}\left(\hat{\boldsymbol{\Sigma}}_0^{\mathrm{SC}}\right)^{-1} - \mathbf{I}_p\right)^2\right] \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda.$$

- The Wald statistic:

$$\hat{\Lambda}_{\ \mathrm{ald}} = \sum_{t=2}^{T} \mathrm{Tr}\left[\left(\mathbf{I}_p - \hat{\boldsymbol{\Sigma}}_1^{\mathrm{SC}}\,(\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}})^{-1}\right)^2\right]$$

$$- q \sum_{t=1}^{T} (\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}})^{-T} \otimes (\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}})^{-1}, \mathrm{vec} \sum_{t=2}^{T} \quad t \quad \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda,$$

SOME PROPERTIES OF THE STATISTICS

CFARness:

The GLRT, t_1 and Wald statistic have the CFAR property with regards to the covariance parameter.

*Proof:*The statistics are invariant for the group of transformation $\mathcal{G} = \left\{ \quad \mathbf{x}_k^t|, \quad \in \mathbb{S}_p^{\mathbb{H}} \right\}$.

Distribution under $\mathbb{H}_0$ (F-Approximation[1]):

Under null hypothesis, we have: $2(1-c)\ln(\hat{\Lambda}_{\mathrm{B}}) \sim \chi^2\left((T-1)p(p+1)\right)$ where

$c = \dfrac{T^2-1}{T(\quad -1)} \times \dfrac{2p^2+3p-1}{6(T-1)(p+1)}$ and $\hat{\Lambda}_{\mathrm{B}}$ is a modified version of $\hat{\Lambda}_{\mathrm{G}}$

There are similar results for others statistics.

[1] Box, G. E. P. “A General Distribution Theory for a Class of Likelihood Criteria.” Biometrika, vol. 36, no. 3/4, 1949, pp. 317–346.

FALSE ALARM/THRESHOLD RELATIONSHIP

Under H_0 we have [1]:

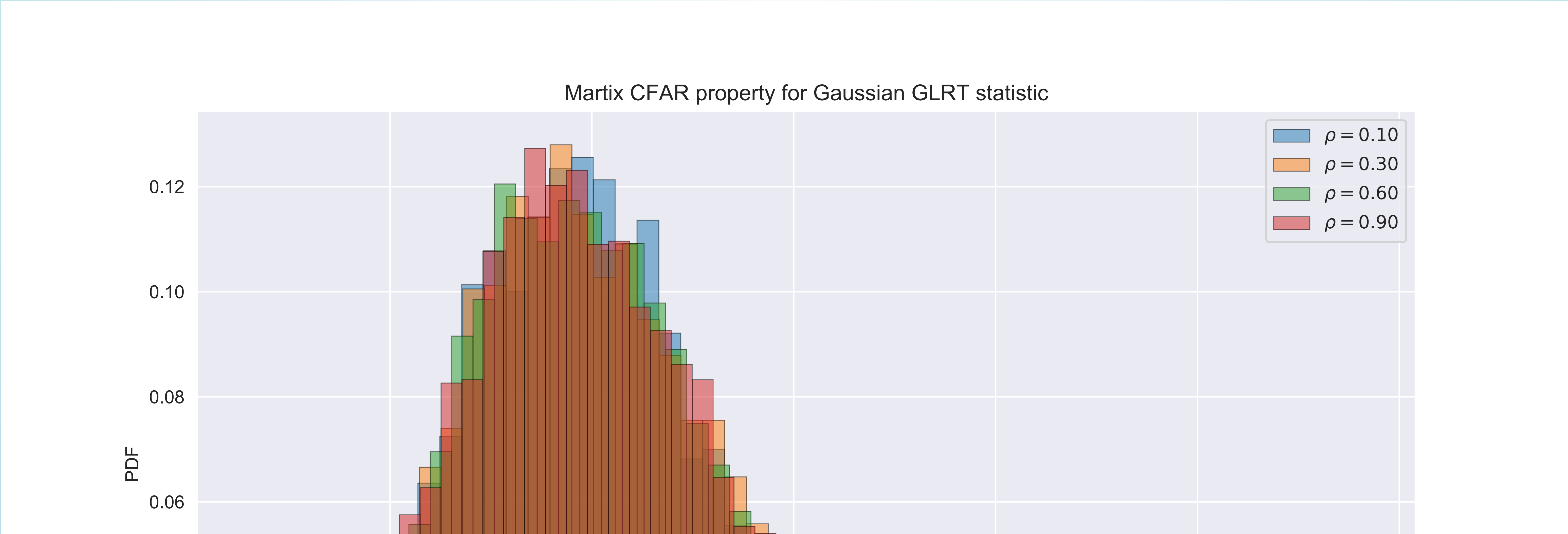
$$P \quad 2\rho\log(\hat{\Lambda}_{\mathrm{G}}) \leq z \quad \approx P\left\{\chi^2(f^2) \leq z\right\} + \omega_2 \left[P\left\{\chi^2(f^2+4) \leq z\right\} - P\left\{\chi^2(f^2) \leq z\right\}\right]$$

$$f = (T-1)p^2,\ \rho = 1 - \frac{(2p^2-1)}{6(T-1)p} \quad \frac{T}{\quad} - \frac{1}{\quad T} \quad ,$$

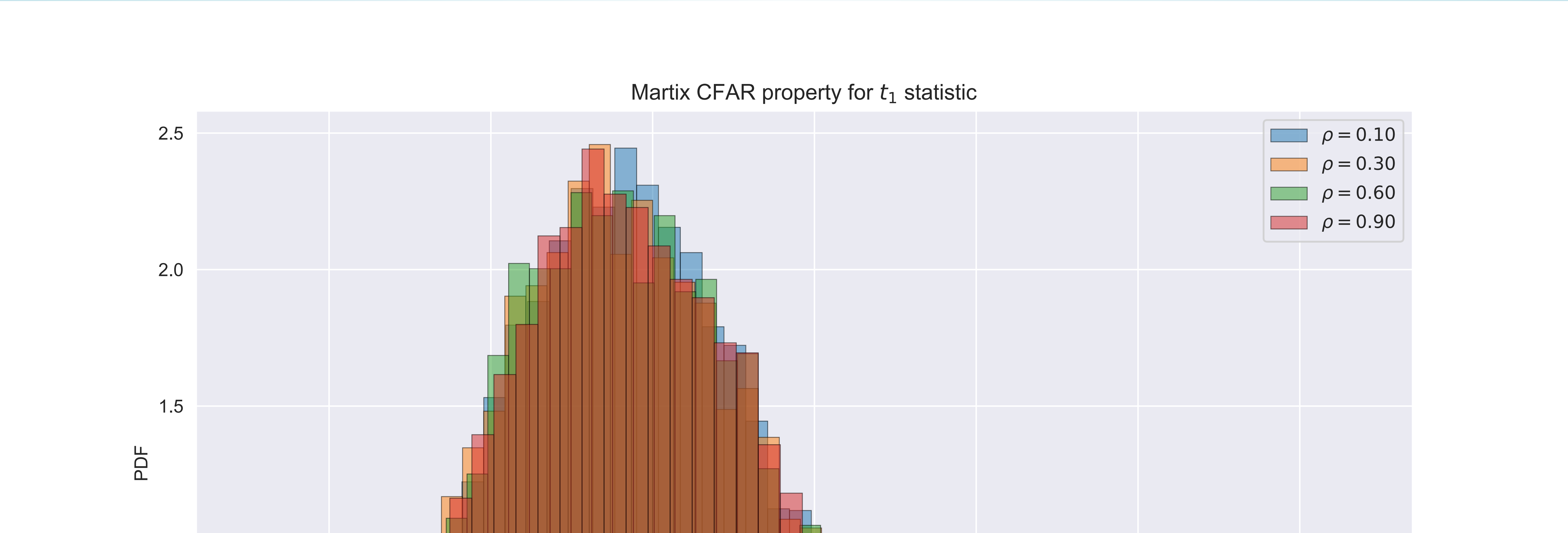
$$\omega_2 = \frac{p^2(p^2-1)}{24\rho^2} \quad \frac{T}{\quad 2} - \frac{1}{(\quad T)^2} \quad - \frac{p^2(T-1)}{4} \quad 1 - \frac{1}{\rho} \quad {}^2$$

[1] Anderson, T. W. (1962). An introduction to multivariate statistical analysis (No. 519.9 A53). New York: Wiley.

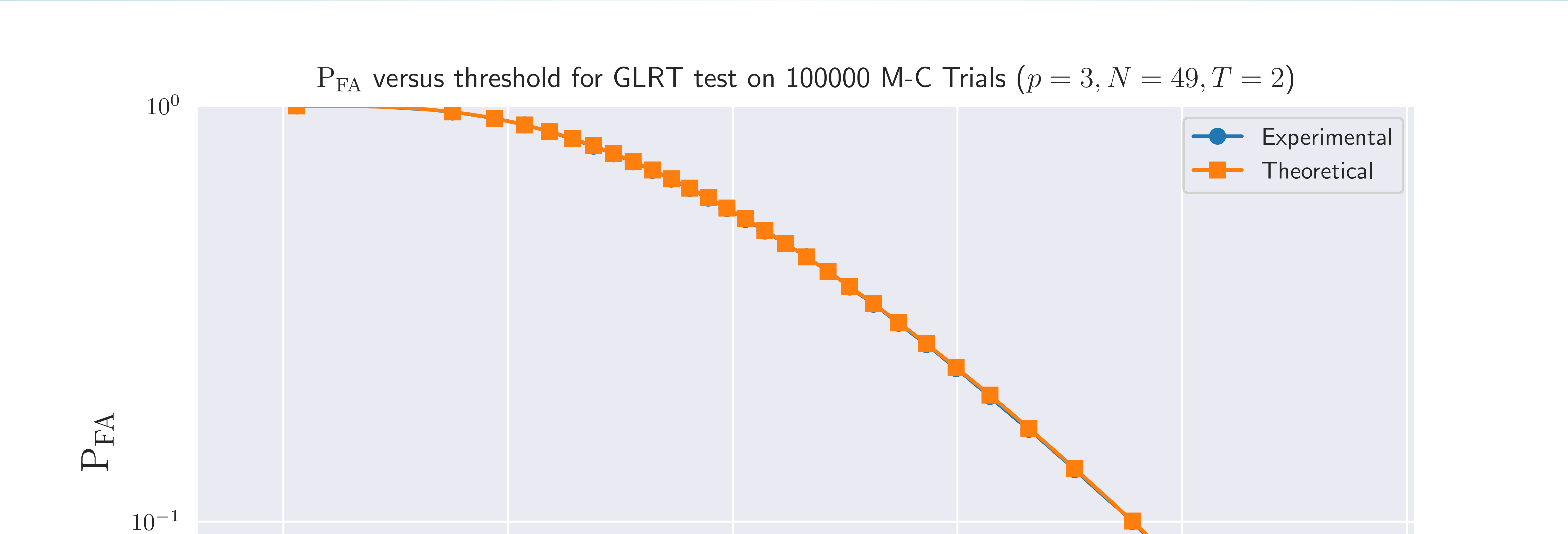
EXPERIMENTAL VALIDATION OF MATRIX CFAR PROPERTY (1/2)



EXPERIMENTAL VALIDATION OF MATRIX CFAR PROPERTY (2/2)



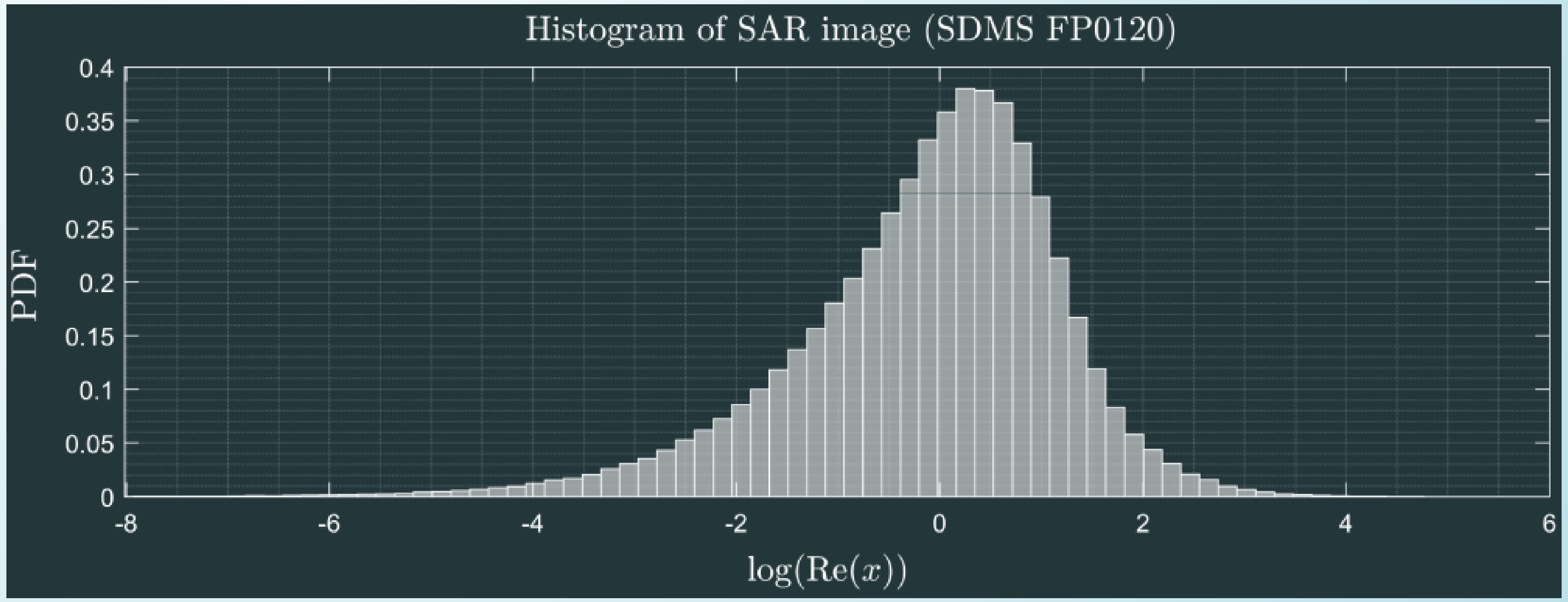
EXPERIMENTAL VALIDATION OF DISTRIBUTINO UNDER $\mathbf{H}_0$



EXTENSION TO ELLIPTICAL DISTRIBUTIONS

NON-GAUSSIANITY HIGH-RESOLUTION IMAGES

Sometimes the Gaussian hypothesis is not accurate !



To model this kind of distribution the family of elliptical distributions have been introduced[1]

$\mathbb{C}ES$ DISTRIBUTIONS

Probability distribution:

$$p_{\mathbf{x};g;\mathbf{\Sigma}}(\mathbf{x};g;\mathbf{\Sigma}) = \mathfrak{C}_{p,g}|\mathbf{\Sigma}|^{-1}g(\mathbf{x}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\mathbf{x})$$

where g is a density generator function with some rgularity conditions, $\mathbf{\Sigma} \in \mathbb{S}_p^{\mathbb{H}}$ is the scatter matrix and $\mathfrak{C}_{p,g}$ is a normalisation constant.

Example:

Multivariate Student-t: $\dfrac{\Gamma \quad \frac{\nu+g}{2} \quad |\Sigma|^{-1/2}}{(\pi\nu)^{\frac{1}{2}}\Gamma \quad \frac{\nu}{2} \quad 1+\frac{\delta(\mathbf{y},\mu,\Sigma)}{\nu} \quad ^{\frac{1}{2}(\nu+g)}}$

Problem: How do we test the equality of scatter matrix in this model ? In peculiar, can we test scale and shape separately ?

APPROACH 1: BOOTSRAPPING THE GAUSSIAN GLRT

Since the distribution of Gaussian-derived detector is known[1] under $\mathbb{C}ES$ model it is possible to correct it in order to obtain one keeping the CFAR property in $\mathbb{C}ES$ context.

Hallin proposed to use the following test[2]:

$$= 2 \quad {}_{1\leq t\leq t'\leq T} \quad {}_{;t,t'} \geq \chi^2 \quad \frac{(T-1)p(p+1)}{2} \quad {}_{1-\mathrm{P_{FA}}}$$, where:

$$_{;t,t'} = \frac{1}{4(1+\hat{\kappa}_p)} \mathrm{Tr}\left[(\hat{\boldsymbol{\Sigma}}_0^{\mathrm{SC}} \quad)^{-1} (\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}} \quad - \hat{\boldsymbol{\Sigma}}_{t'}^{\mathrm{SC}} \quad)^2 \right] - \frac{\hat{\kappa}_p}{(p+2)\hat{\kappa}_p + 2} \mathrm{Tr}^2 \left[(\hat{\boldsymbol{\Sigma}}_0^{\mathrm{SC}} \quad)^{-1} (\hat{\boldsymbol{\Sigma}}_t^{\mathrm{SC}} \quad - \hat{\boldsymbol{\Sigma}}_{t'}^{\mathrm{SC}} \quad) \right] \quad ,$$

and $\hat{\kappa}_p = p(p+1)/2 \quad {}_{t=1}^{T} \quad {}_{k=1} d^4(\mathbf{x}_k^t, \hat{\boldsymbol{\Sigma}}_0^{\mathrm{SC}} \quad) - 1$and $d(\mathbf{x}, \boldsymbol{\Sigma}) = \|\boldsymbol{\Sigma}^{-1/2}\mathbf{x}\|$

[1] Yanagihara, H., Tonda, T., and Matsumoto, C. (2005).The effects of nonnormality on asymptotic distributions of some likelihood ratio criteria for testing covariance structures under normal assumption. Journal of Multivariate Analysis, 96(2):237{264.

[2] Marc Hallin, Davy Paindaveine, Optimal tests for homogeneity of covariance, scale, and shape,

APPROACH 2: GLRT UNDER $\mathbb{C}\mathcal{ES}$ MODEL

We compute the following statistic:

$$\hat{\Lambda} = \frac{\max_{\mathbf{\Sigma}_1,\dots,\mathbf{\Sigma}_T} \; p_{\ 1,\dots,\ T;\mathbf{\Sigma}_1,\dots,\mathbf{\Sigma}_T}(\ 1,\dots,\ T;\mathbf{\Sigma}_1,\dots,\mathbf{\Sigma}_T/\mathrm{H}_1)}{\max_{\mathbf{\Sigma}_0} \; p_{\ 1,\dots,\ T;\mathbf{\Sigma}_0}(\ 1,\dots,\ T;\mathbf{\Sigma}_0/\mathrm{H}_0)}$$

which yields:

$$\hat{\Lambda}^g = \frac{|\hat{\mathbf{\Sigma}}_0\ |^T}{\underset{t=1}{\overset{T}{}}\ |\hat{\mathbf{\Sigma}}_t\ |} \ \underset{t=1}{\overset{T}{}}\ \underset{k=1}{} \frac{g(\mathbf{x}_k^{t\ \mathrm{H}}\{\hat{\mathbf{\Sigma}}_0\ \}^{-1}\mathbf{x}_k^t)}{g(\mathbf{x}_k^{t\ \mathrm{H}}\{\hat{\mathbf{\Sigma}}_t\ \}^{-1}\mathbf{x}_k^t)} \ \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{}}\ \lambda,$$

where: $\hat{\mathbf{\Sigma}}_t \ = f_t(\hat{\mathbf{\Sigma}}_t\ )$, $\hat{\mathbf{\Sigma}}_0\ = \frac{1}{T}\underset{t=1}{\overset{T}{}} f_t(\hat{\mathbf{\Sigma}}_0\ )$ and $f_t(\mathbf{\Sigma}) = \frac{1}{\ }\underset{k=1}{} \frac{-g'(\mathbf{x}_k^{t\ \mathrm{H}}\{\hat{\mathbf{\Sigma}}\}^{-1}\mathbf{x}_k^t)}{g(\mathbf{x}_k^{t\ \mathrm{H}}\{\hat{\mathbf{\Sigma}}\}^{-1}\mathbf{x}_k^t)}\mathbf{x}_k^t\mathbf{x}_k^{t\ \mathrm{H}}$

Problem: We need to know g !!!

APPROACH 2BIS: SELECT $g = \frac{1}{x^p}$

The statistic reads:

$$\hat{\Lambda}^{1/x^p} = \frac{\left|\hat{\mathbf{\Sigma}}_0\right|^T}{\prod_{t=1}^{T} \left|\hat{\mathbf{\Sigma}}_t\right|} \prod_{t=1}^{T} \prod_{k=1} \frac{\left(\mathbf{x}_k^{t\,\mathrm{H}} \{\hat{\mathbf{\Sigma}}_t\}^{-1} \mathbf{x}_k^t\right)^p}{\left(\mathbf{x}_k^{t\,\mathrm{H}} \{\hat{\mathbf{\Sigma}}_0\}^{-1} \mathbf{x}_k^t\right)^p} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda,$$

Properties[1]: - This statistic is valid for testing the **Shape matrix** for any elliptical distribution.

- The distribution of $2\log(\hat{\Lambda}^g)$ under H_0 is assyptotically that of a $\chi^2\left((T-1)p(p+1)\right)$

Problem: We can't test a change in the scale !

[1] A. Mian, J. Ovarlez, G. Ginolhac and A. M. Atto, "A Robust Change Detector for Highly Heterogeneous Multivariate Images," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, 2018, pp. 3429-3433.

APPROACH 3: COMPOUND-GAUSSIAN MODEL (1/2)

We consider the Compound-Gaussian model: $\mathbf{x}_k^t \sim \overline{\tau_k^t}\,\mathbb{C}\quad(\mathbf{0}_p, \boldsymbol{\xi}_t)$, where τ_k^t are assumed to be **deterministic** and $Tr(\boldsymbol{\xi}) = p$. $\rightarrow$ We can integrate these unknown parameters in the GLRT[1]:

$$\hat{\Lambda}_{\mathrm{T}} = \frac{|\hat{\mathbf{\Sigma}}_0^{\mathrm{T}}|^T}{\mathop{}_{t=1}^{T} |\hat{\mathbf{\Sigma}}_t|} \mathop{}_{k=1} \left(\frac{\mathop{}_{t=1}^{T} {\mathbf{x}_k^t}^{\mathrm{H}} \{\hat{\mathbf{\Sigma}}_t^{\mathrm{T}}\}^{-1} \mathbf{x}_k^t}{\mathop{}_{t=1}^{T} {\mathbf{x}_k^t}^{\mathrm{H}} \{\hat{\mathbf{\Sigma}}_0\}^{-1} \mathbf{x}_k^t} \right)^{Tp}_{p} \mathop{}^{\mathrm{H}_1}_{\mathrm{H}_0} \lambda, \text{ where :}$$

$$\hat{\mathbf{\Sigma}}_t = \frac{p}{} \mathop{}_{k=1} \frac{\mathbf{x}_k^t {\mathbf{x}_k^t}^{\mathrm{H}}}{{\mathbf{x}_k^t}^{\mathrm{H}} \{\hat{\mathbf{\Sigma}}\}^{-1} \mathbf{x}_k^t} \text{ and } \hat{\mathbf{\Sigma}}_0^{\mathrm{T}} = \frac{p}{} \mathop{}_{k=1} \frac{\mathop{}_{t=1}^{T} \mathbf{x}_k^t {\mathbf{x}_k^t}^{\mathrm{H}}}{\mathop{}_{t=1}^{T} {\mathbf{x}_k^t}^{\mathrm{H}} \{\hat{\mathbf{\Sigma}}^{\mathrm{T}}\}^{-1} \mathbf{x}_k^t}$$

APPROACH 3: COMPOUND-GAUSSIAN MODEL (2/2)

We can also test the scale parameters only and ignoring a change in the shape matrix.
$\rightarrow$ The GLRT[1] yields:

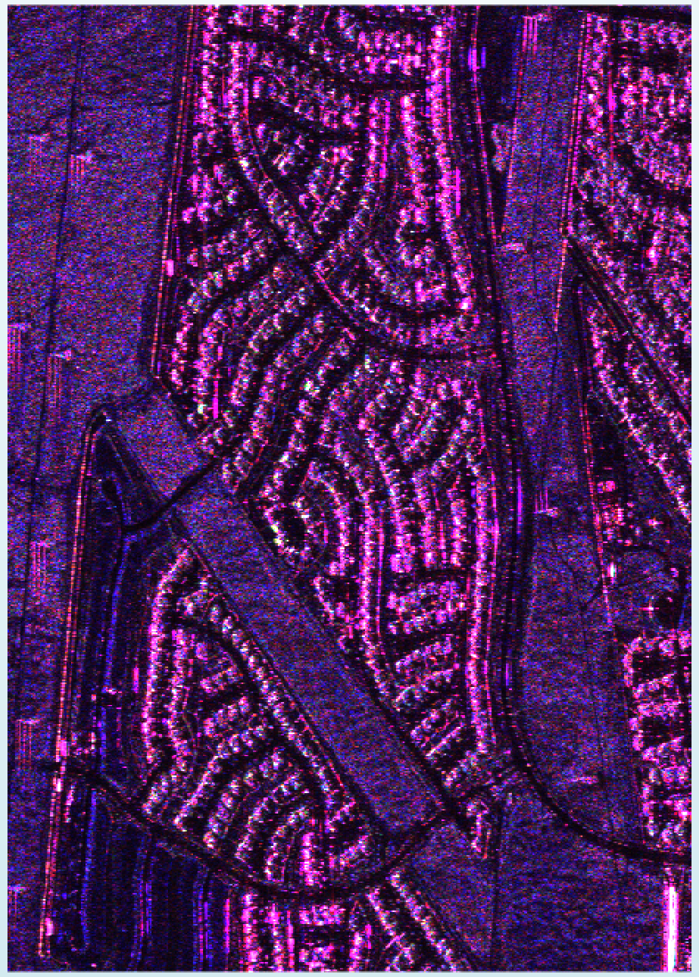
$$\hat{\Lambda}_{\text{Tex}} = \prod_{t=1}^{T} \frac{\left|\hat{\boldsymbol{\Sigma}}_t^{\text{Tex}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_t\right|} \prod_{k=1} \frac{\left(\sum_{t=1}^{T} {\mathbf{x}_k^t}^{\text{H}} \{\hat{\boldsymbol{\Sigma}}_t^{\text{Tex}}\}^{-1} \mathbf{x}_k^t\right)^{Tp}}{\prod_{t=1}^{T} \left({\mathbf{x}_k^t}^{\text{H}} \{\hat{\boldsymbol{\Sigma}}_0\}^{-1} \mathbf{x}_k^t\right)^{p}} \underset{\text{H}_0}{\overset{\text{H}_1}{\gtrless}} \lambda, \text{ where :}$$

$$\hat{\boldsymbol{\Sigma}}_t = \frac{p}{} \sum_{k=1} \frac{\mathbf{x}_k^t {\mathbf{x}_k^t}^{\text{H}}}{{\mathbf{x}_k^t}^{\text{H}} \{\hat{\boldsymbol{\Sigma}}\}^{-1} \mathbf{x}_k^t} \text{ and } \hat{\boldsymbol{\Sigma}}_t^{\text{Tex}} = \frac{p}{} \sum_{k=1} \frac{\sum_{t=1}^{T} \mathbf{x}_k^t {\mathbf{x}_k^t}^{\text{H}}}{\sum_{t=1}^{T} {\mathbf{x}_k^t}^{\text{H}} \{\hat{\boldsymbol{\Sigma}}^{\text{Tex}}\}^{-1} \mathbf{x}_k^t}$$

RESULTS ON SAR IMAGES

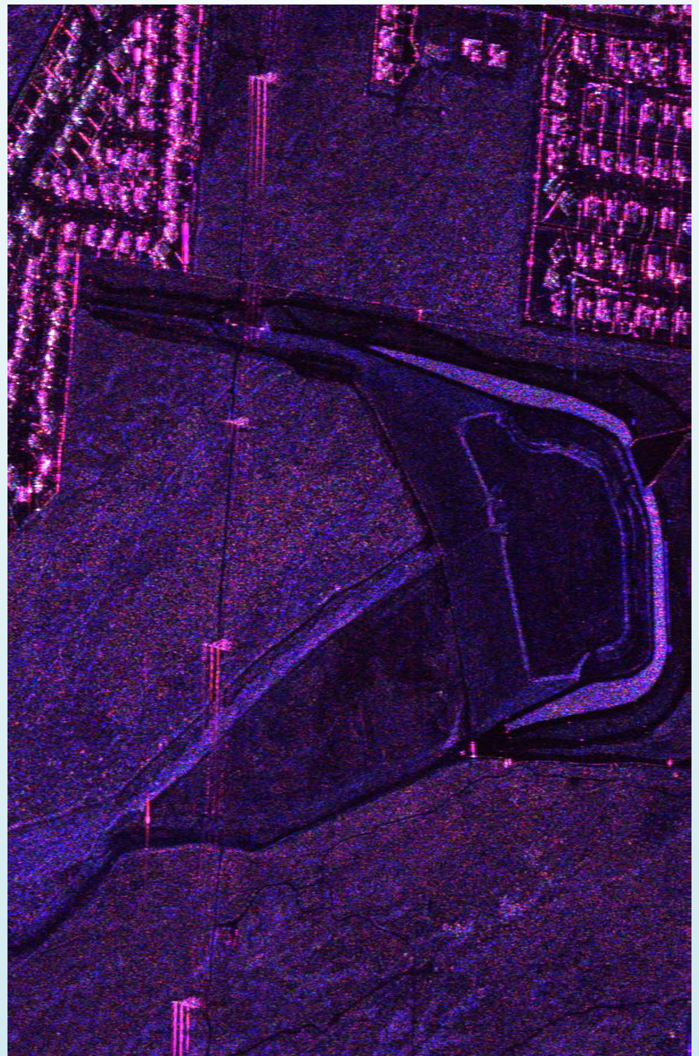
THE UAVSAR DATASET

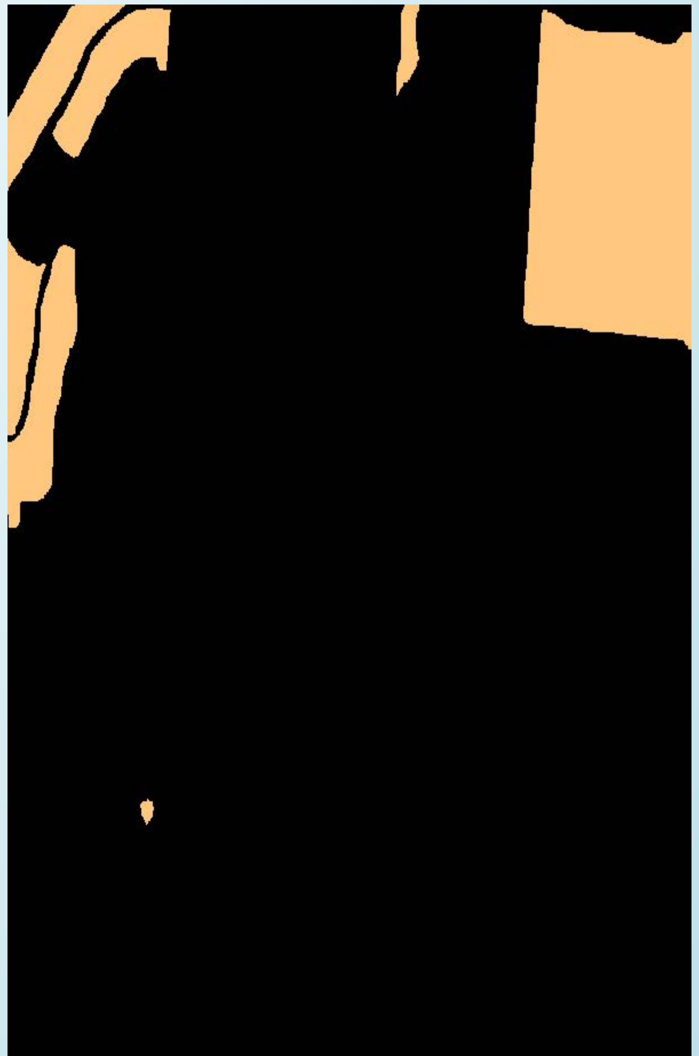
- Polarimetric data $(p = 3)$
- Dimensions: 2360px by 600 px
- Resolution: 1m67 (Range) and 1m (Azimuth)



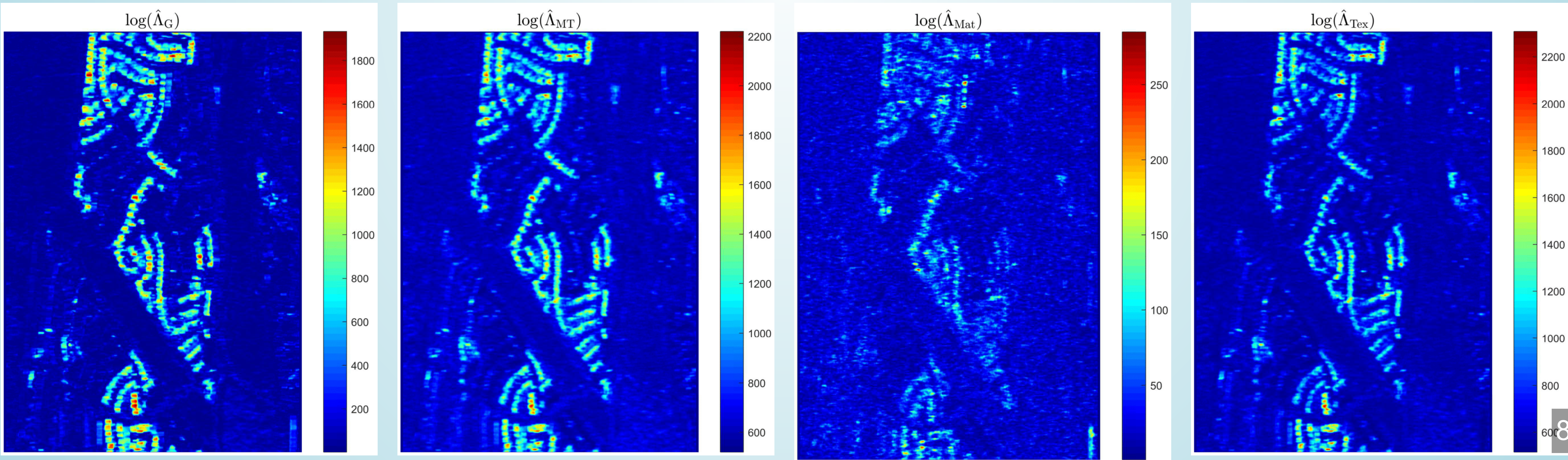
THE UAVSAR DATASET (2/2)

- Polarimetric data ($p = 3$)
- Dimensions: 2360px by 600 px
- Resolution: 1m67 (Range) and 1m (Azimuth)

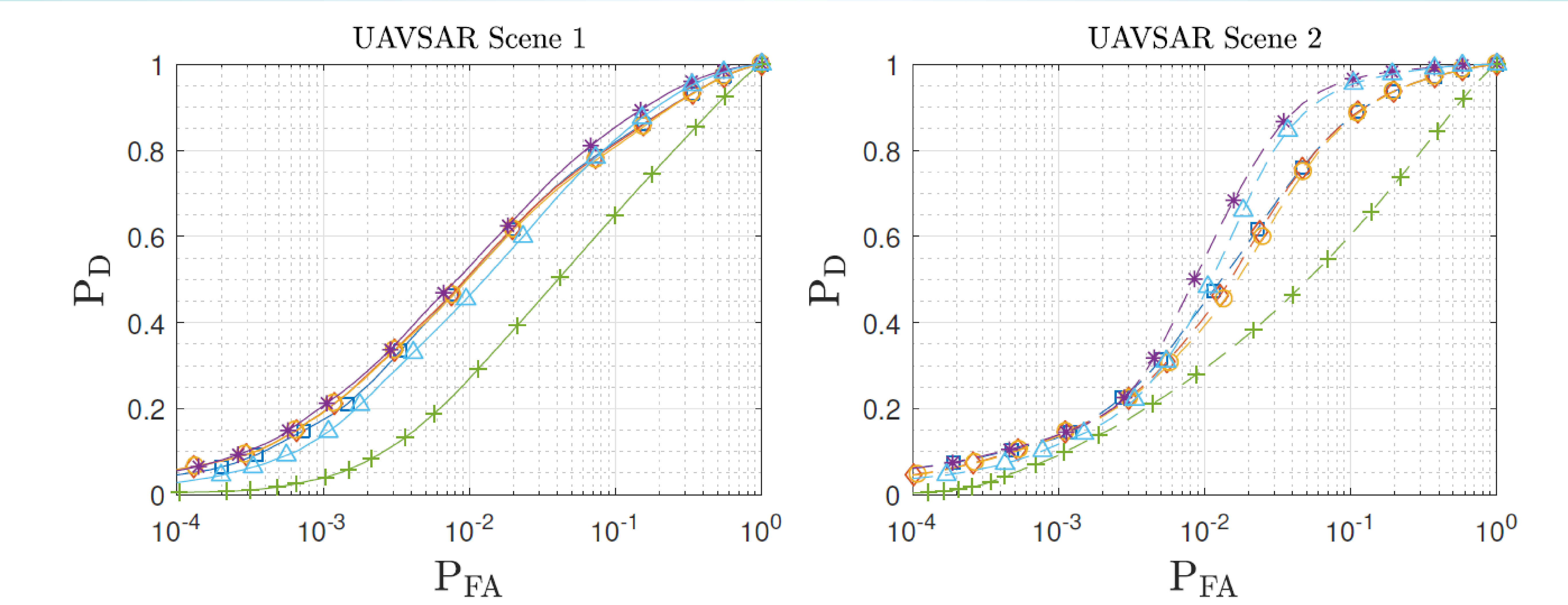




RESULTS (1/2)

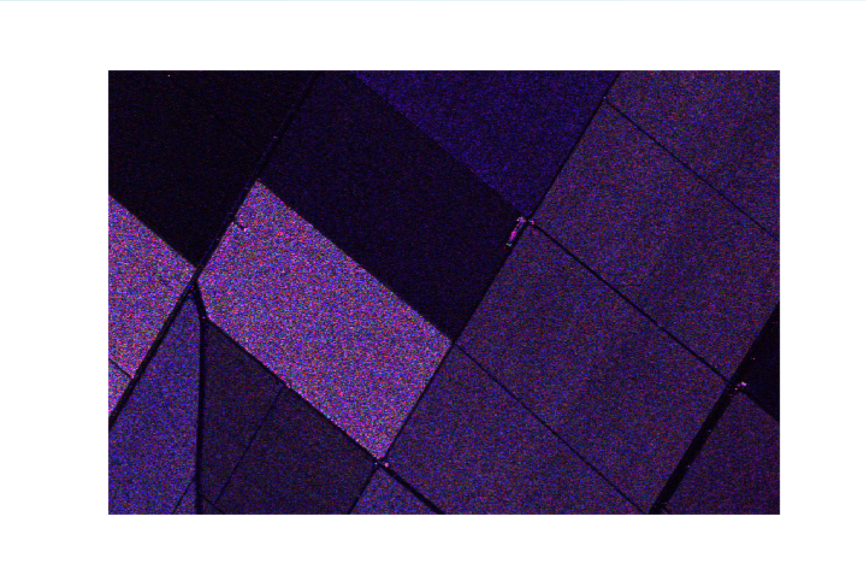

$\log(\hat{\Lambda}_{\mathrm{G}})$
1800
1600
1400
1200
1000
800
600
400
200
$\log(\hat{\Lambda}_{\mathrm{MT}})$
2200
2000
1800
1600
1400
1200
1000
800
600
$\log(\hat{\Lambda}_{\mathrm{Mat}})$
250
200
150
100
50
$\log(\hat{\Lambda}_{\mathrm{Tex}})$
2200
2000
1800
1600
1400
1200
1000
800

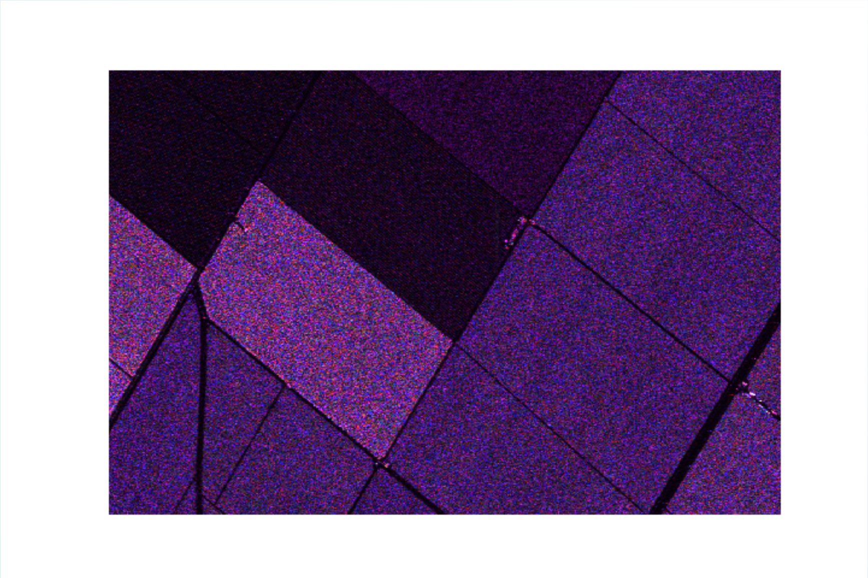
RESULTS (2/2)

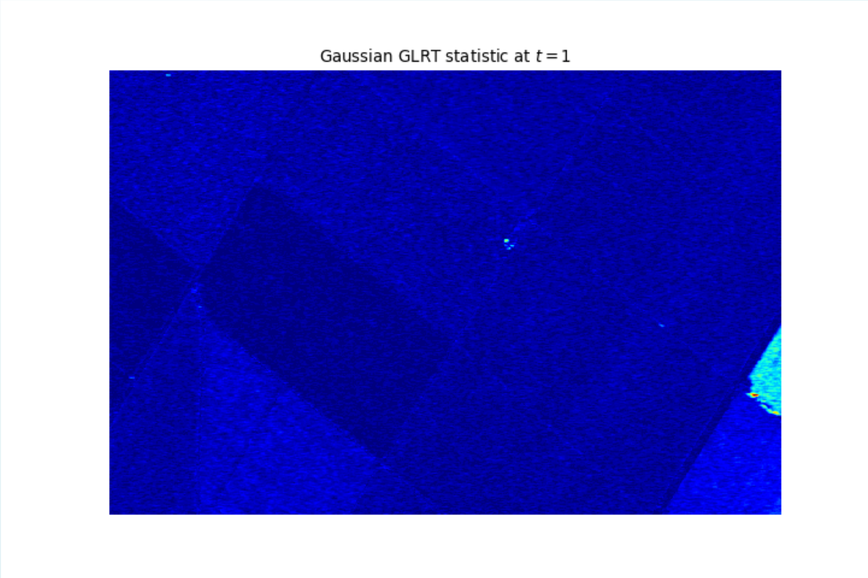


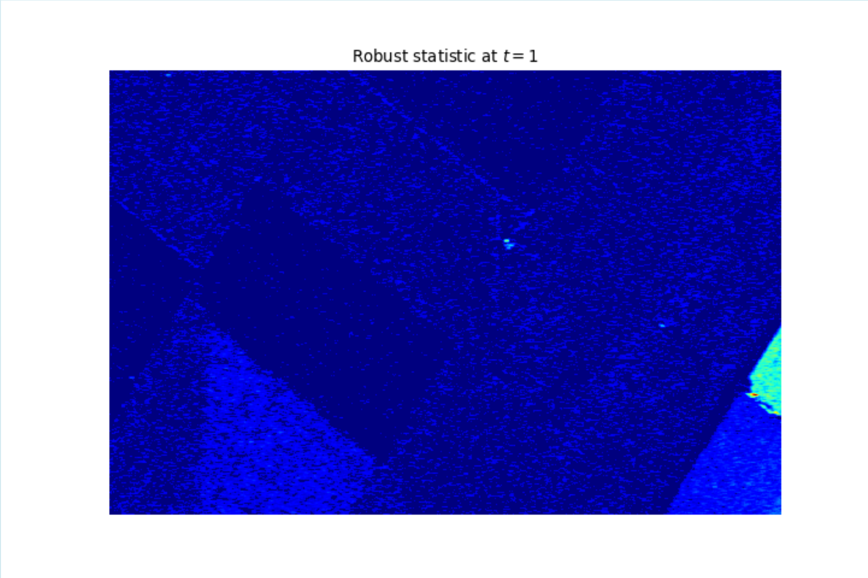
UAVSAR Scene 1
UAVSAR Scene 2
P_D
P_{FA}
1
0.8
0.6
0.4
0.2
0
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}

RESULT ON ANOTHER DATASET

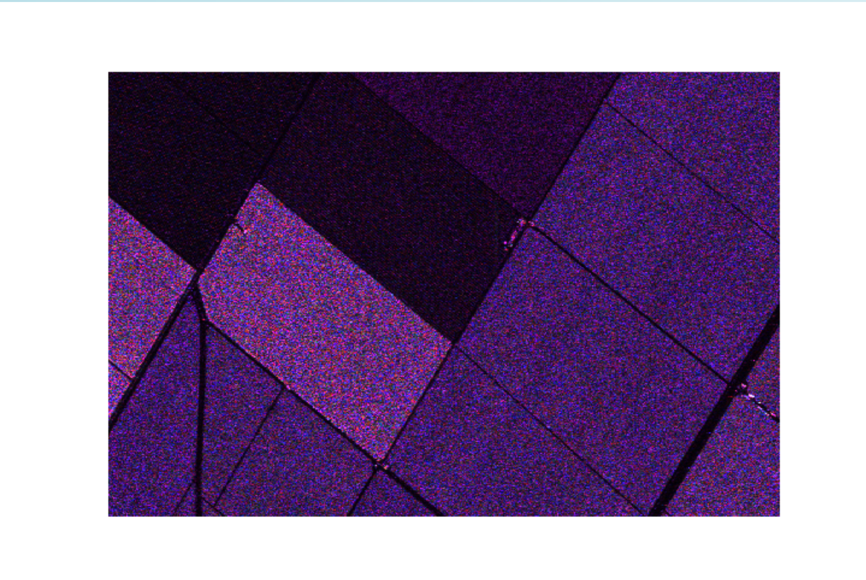


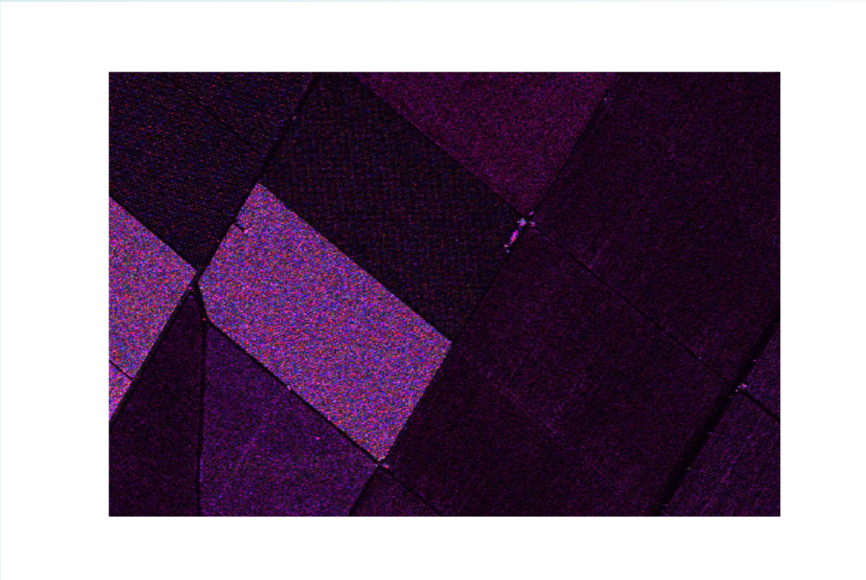


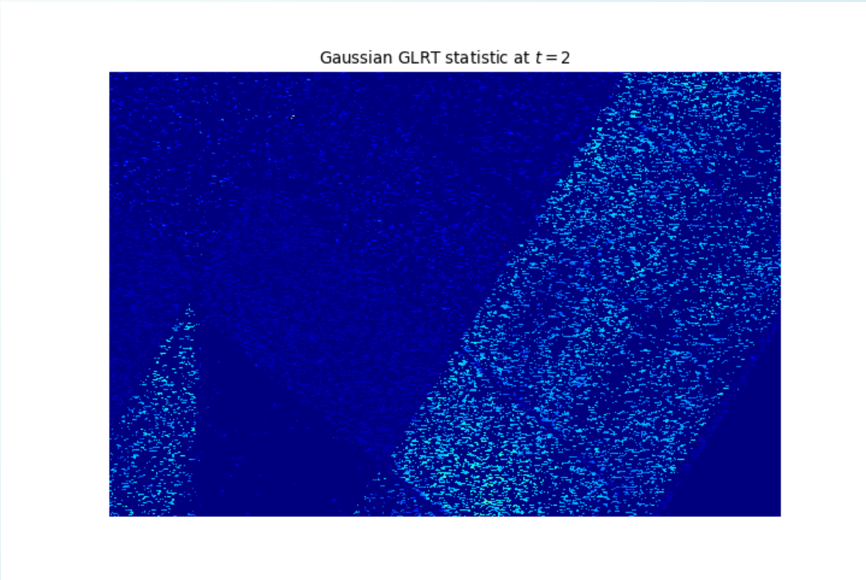


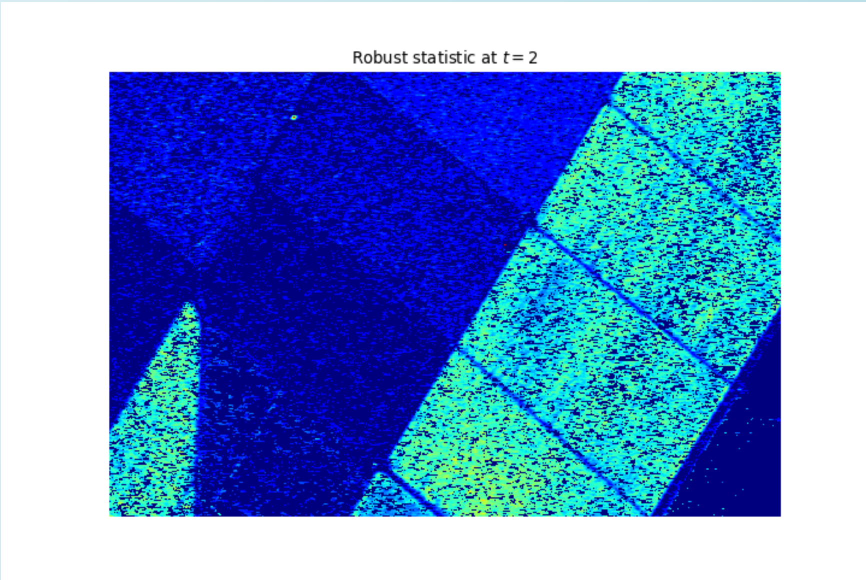


RESULT ON ANOTHER DATASET

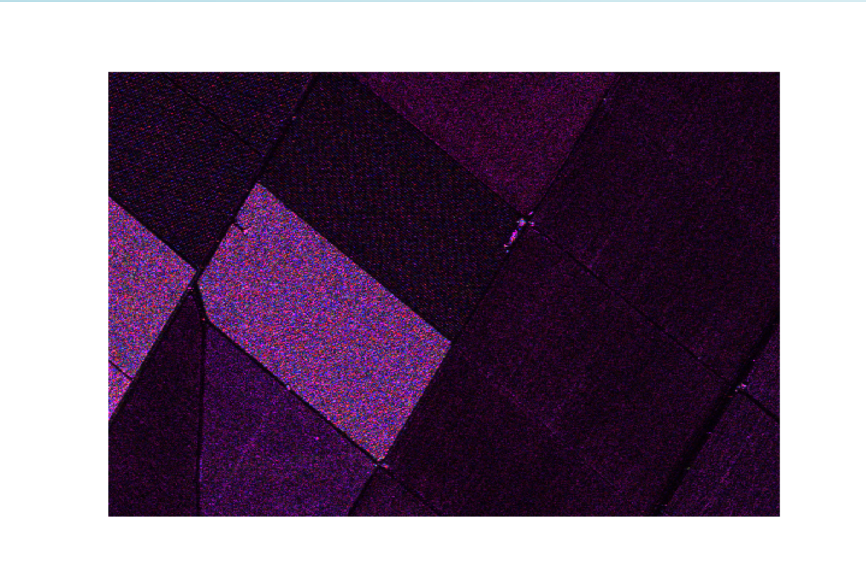


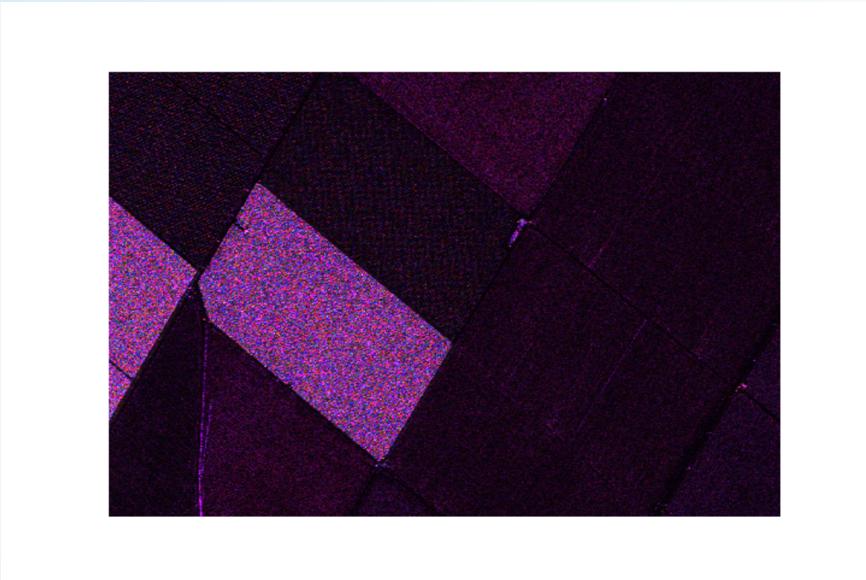


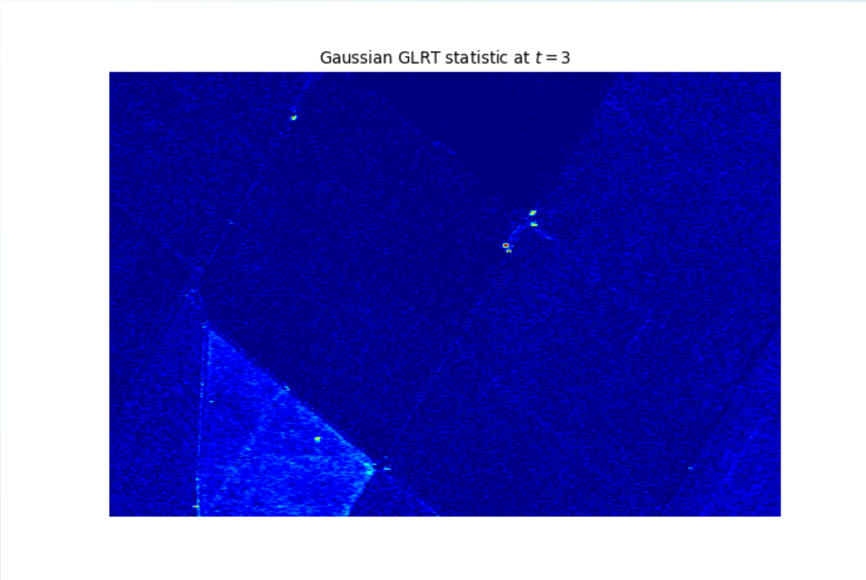


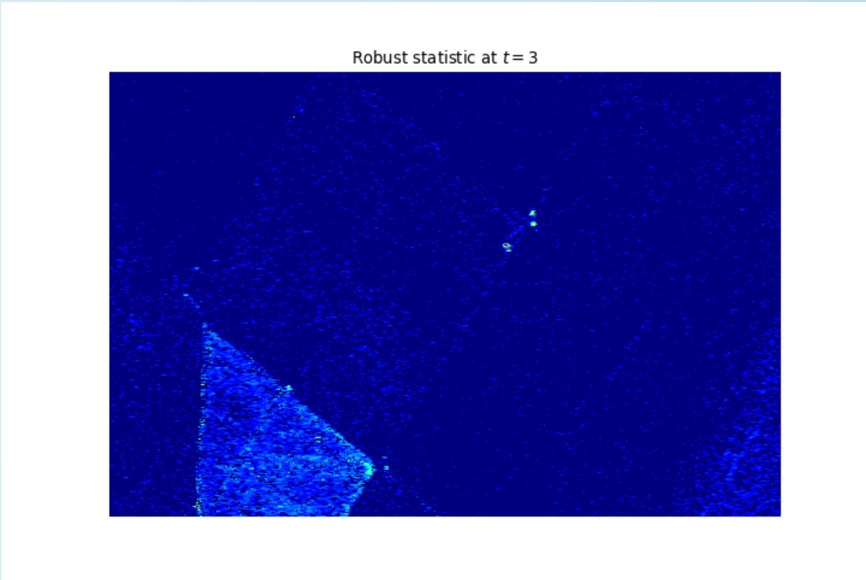


RESULT ON ANOTHER DATASET

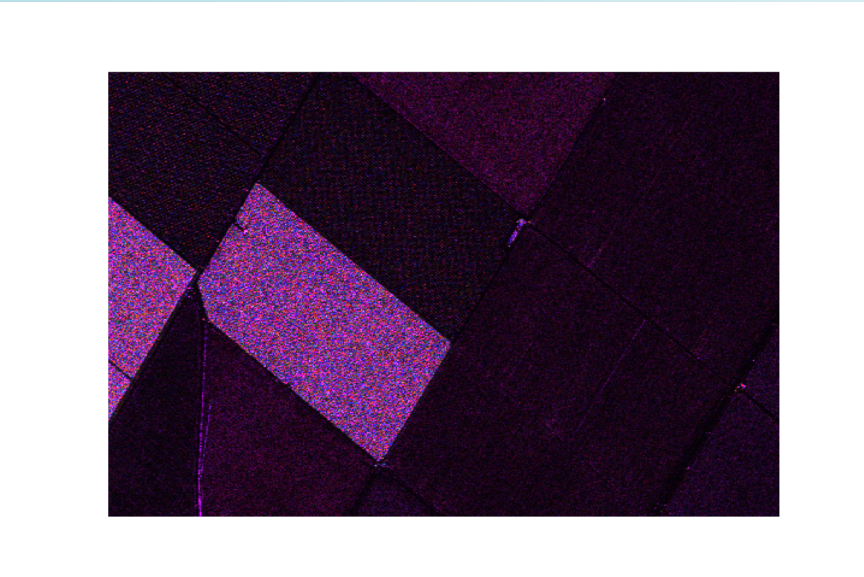


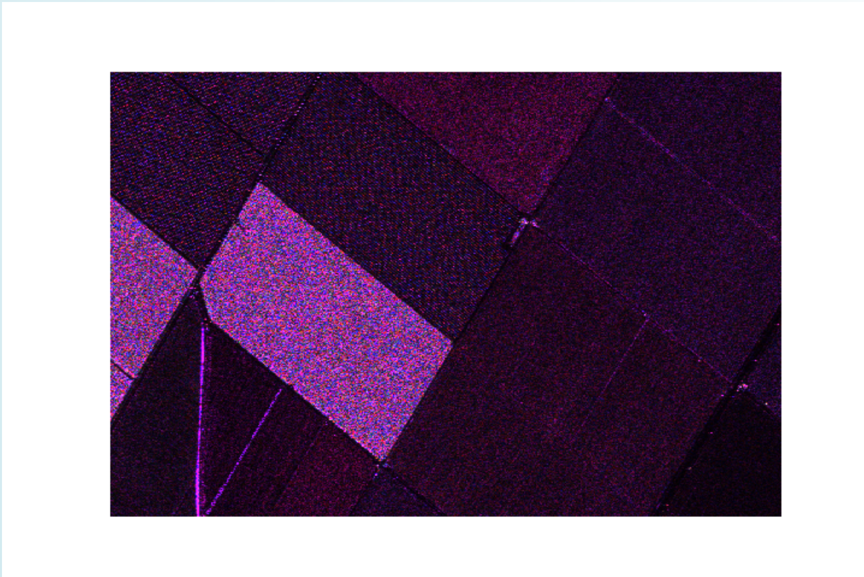


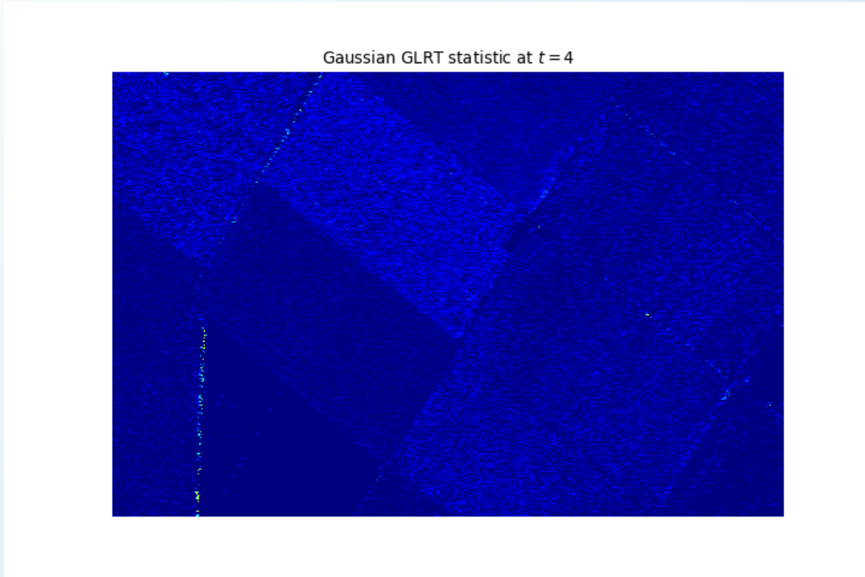


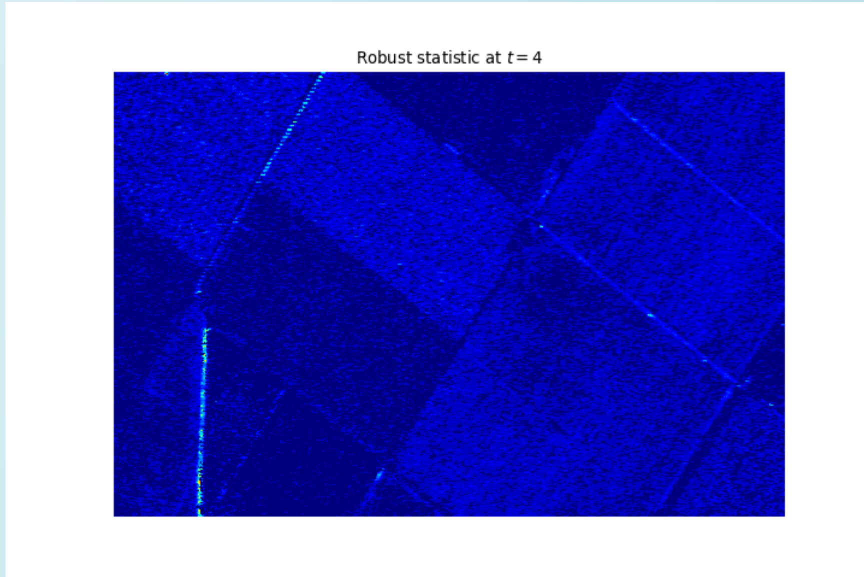


RESULT ON ANOTHER DATASET

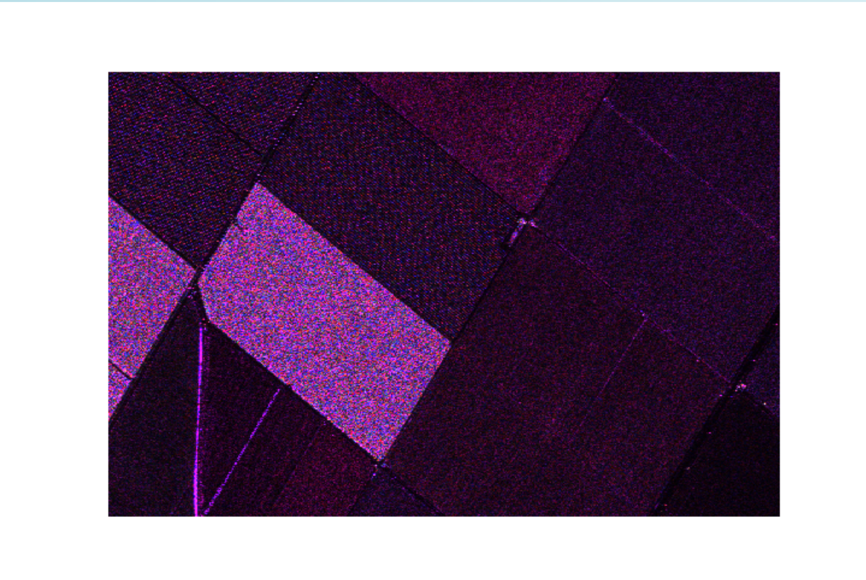


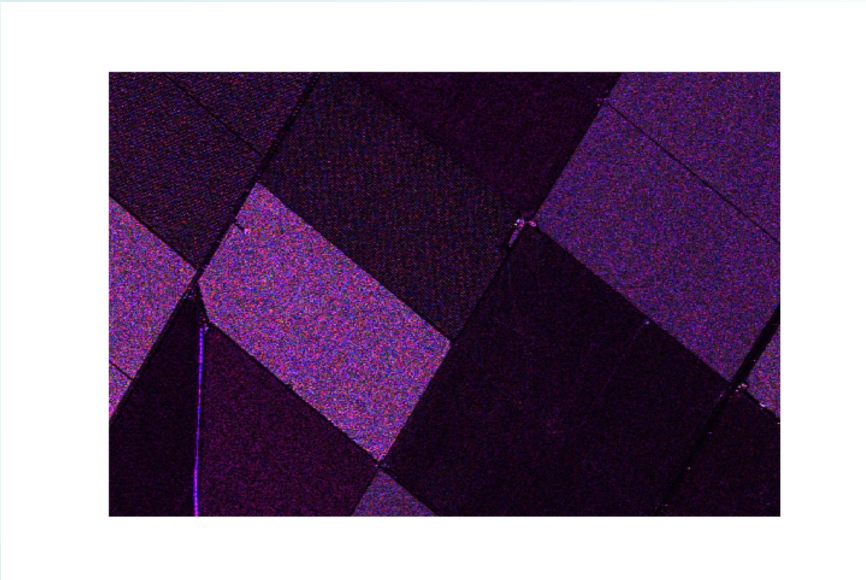


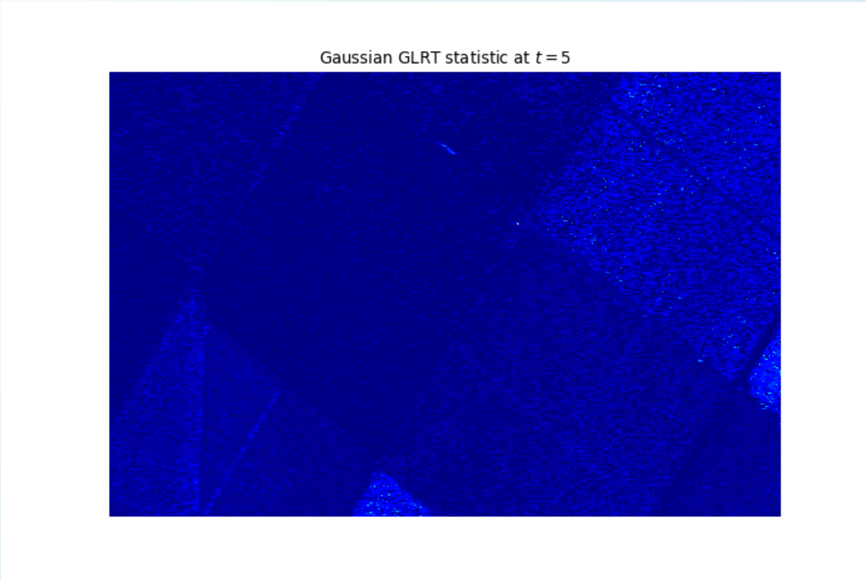


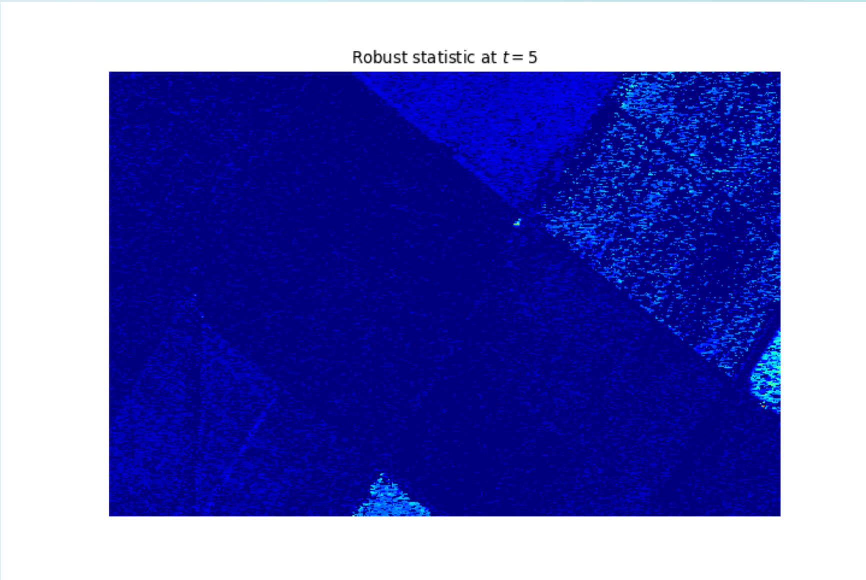


RESULT ON ANOTHER DATASET

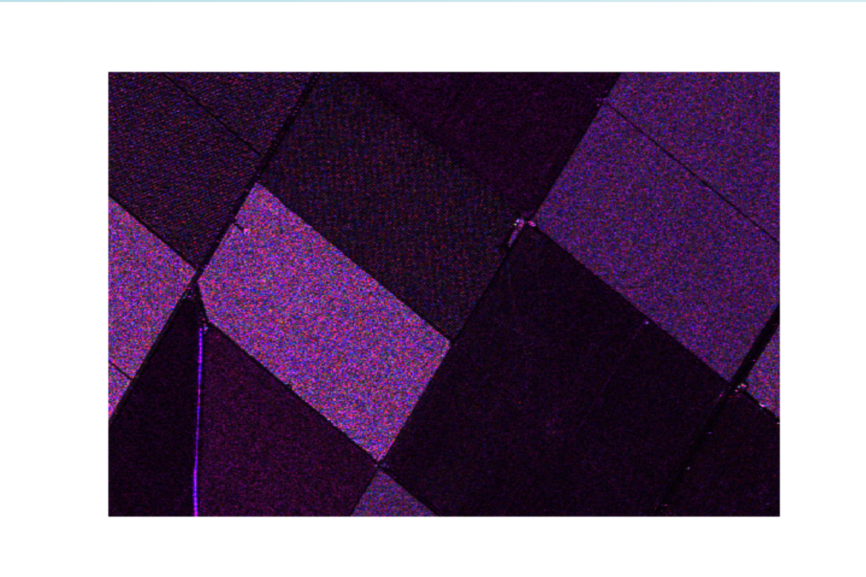


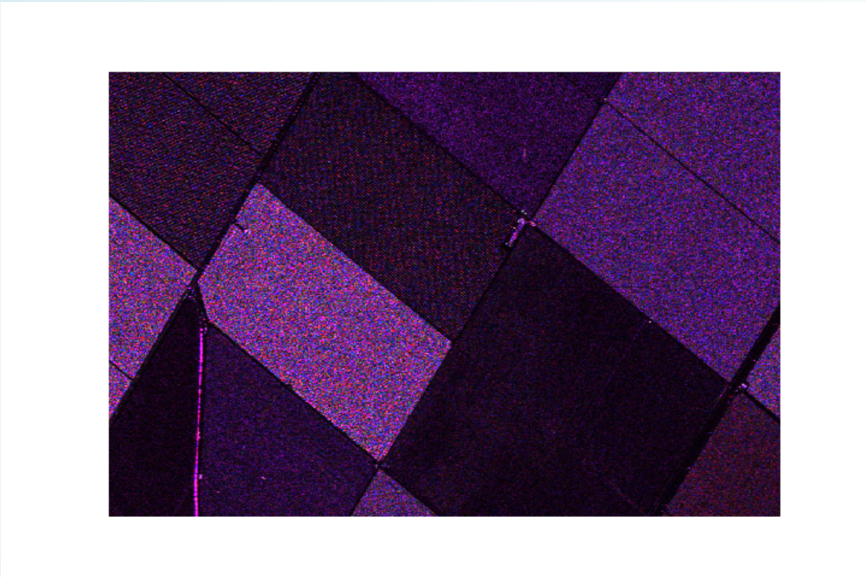


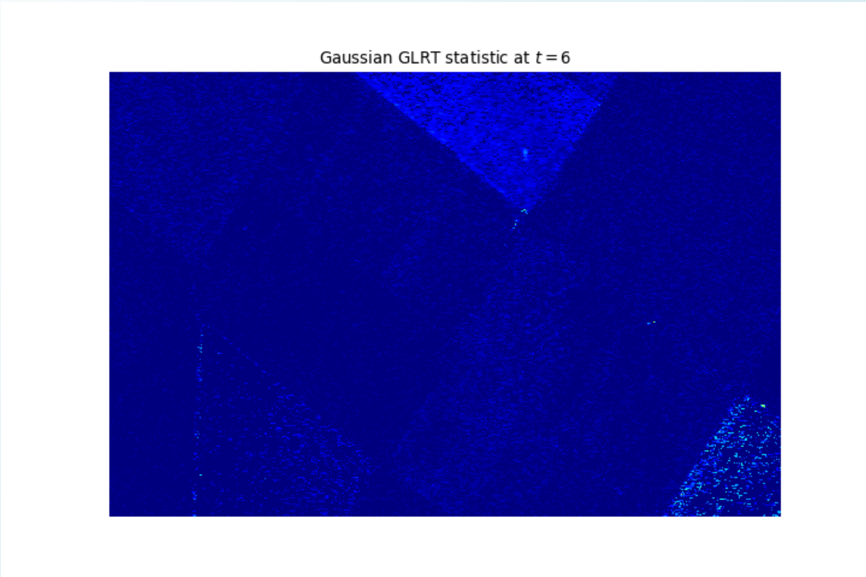


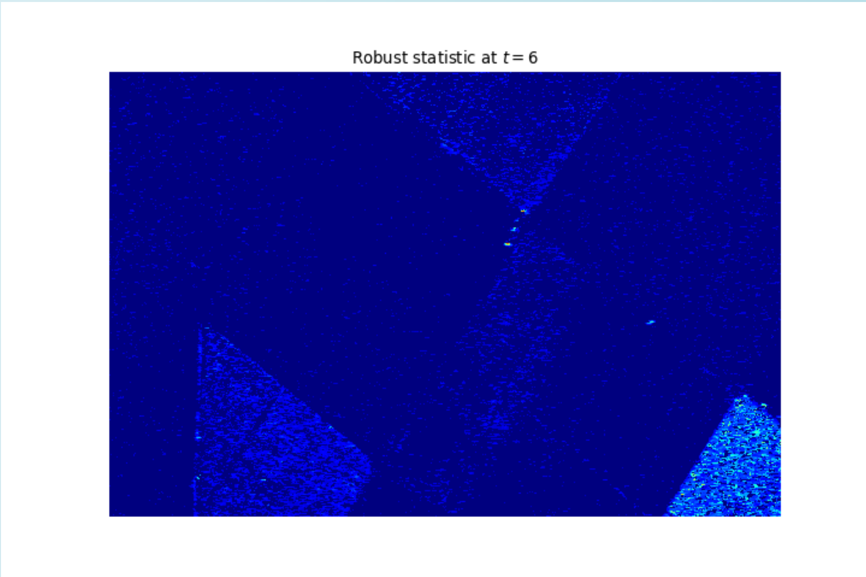


RESULT ON ANOTHER DATASET

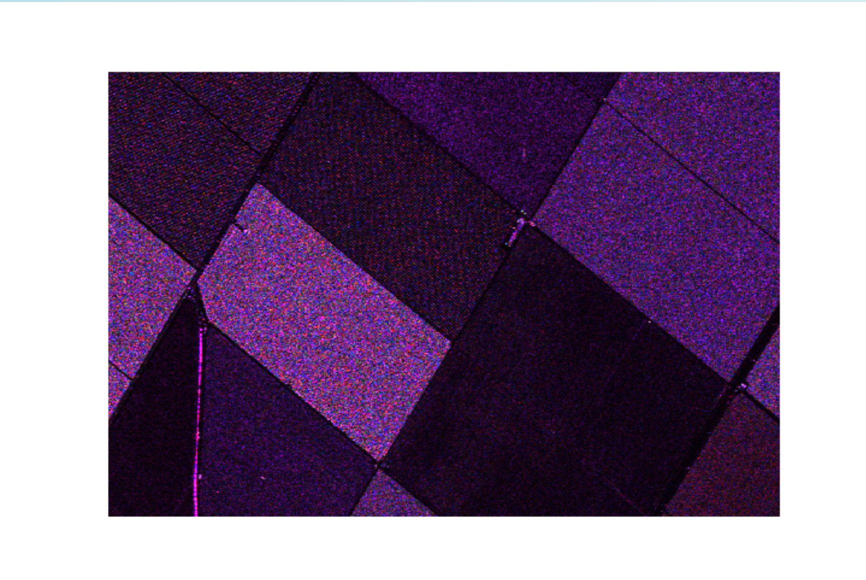


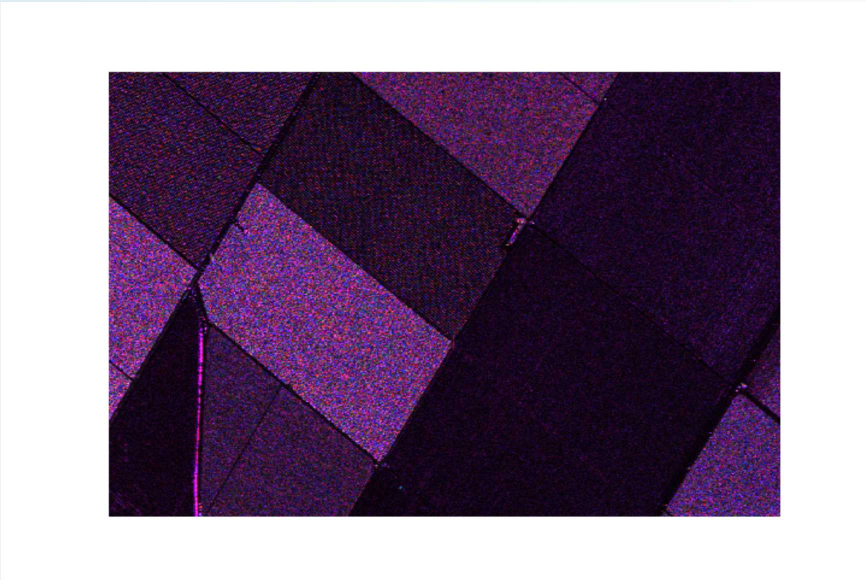


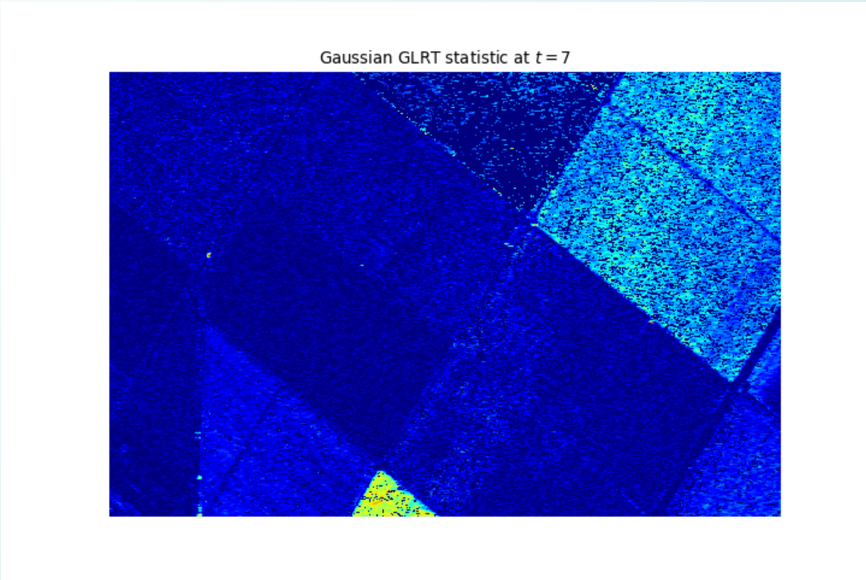


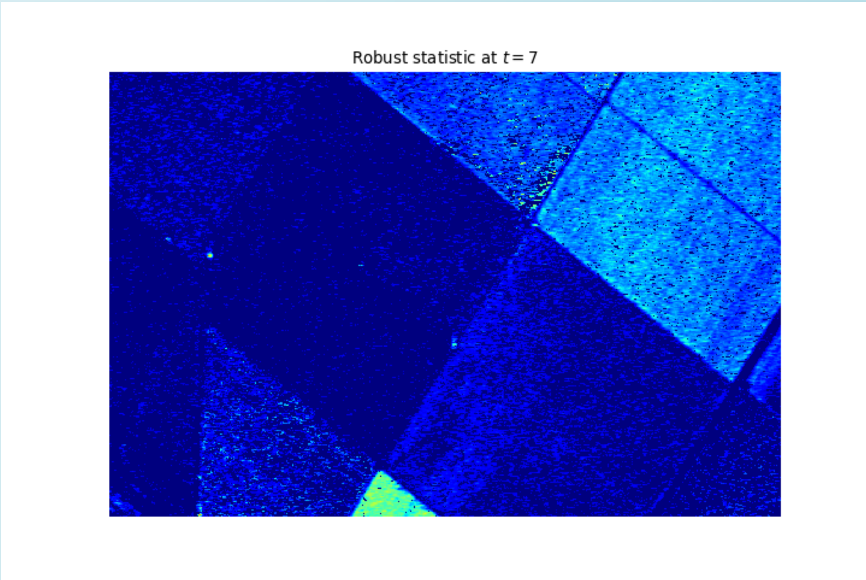


RESULT ON ANOTHER DATASET

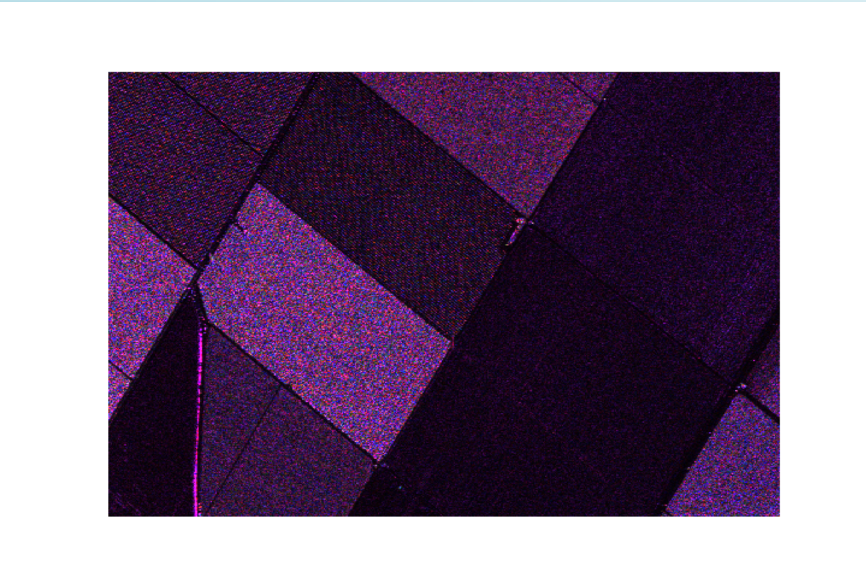


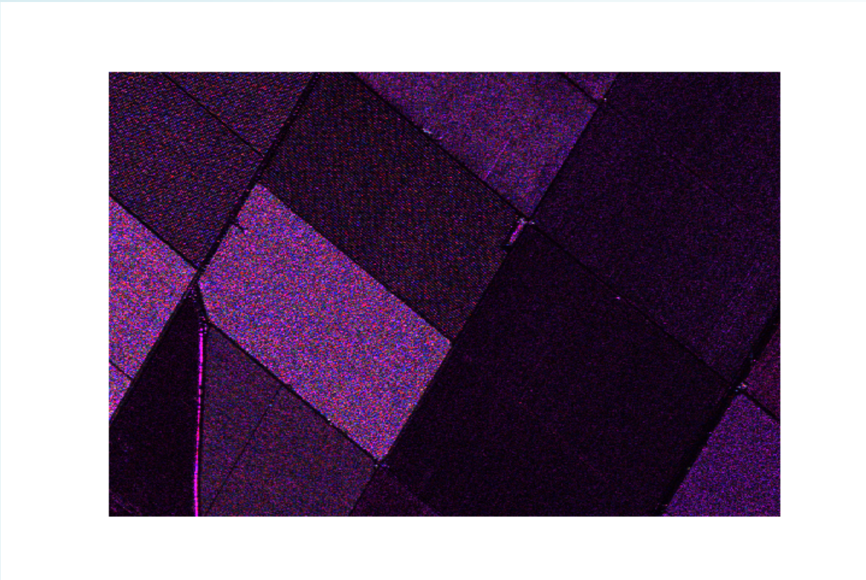


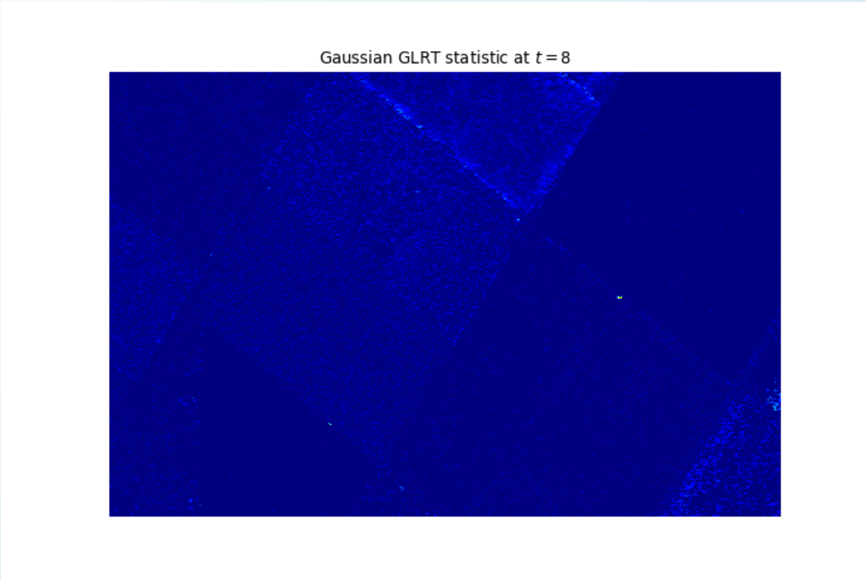


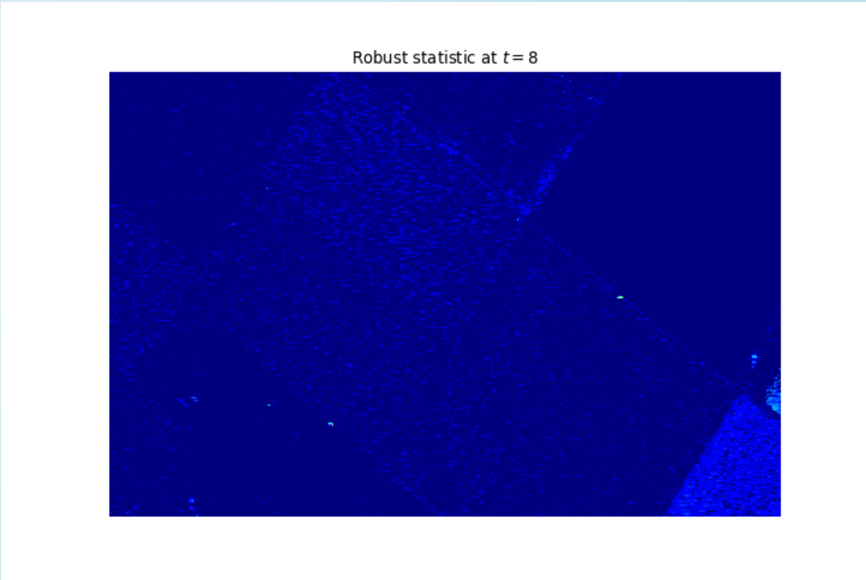


RESULT ON ANOTHER DATASET

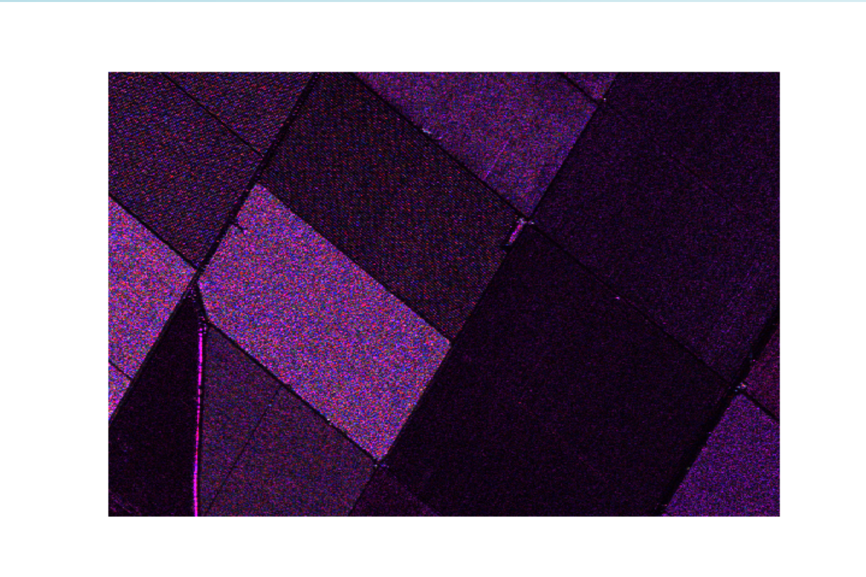


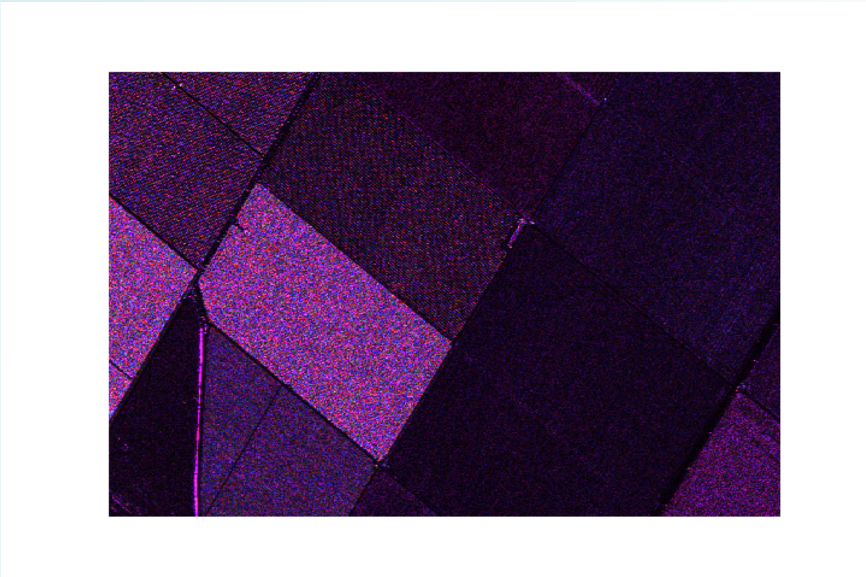


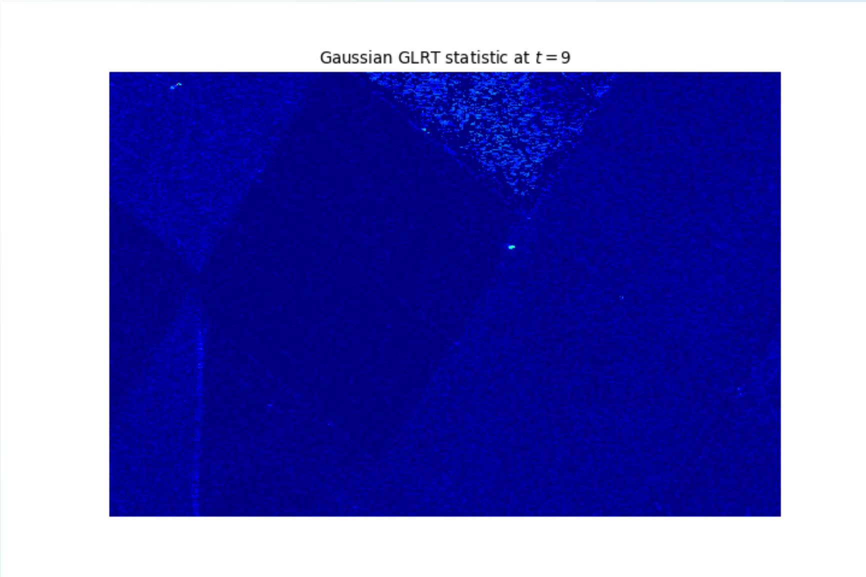


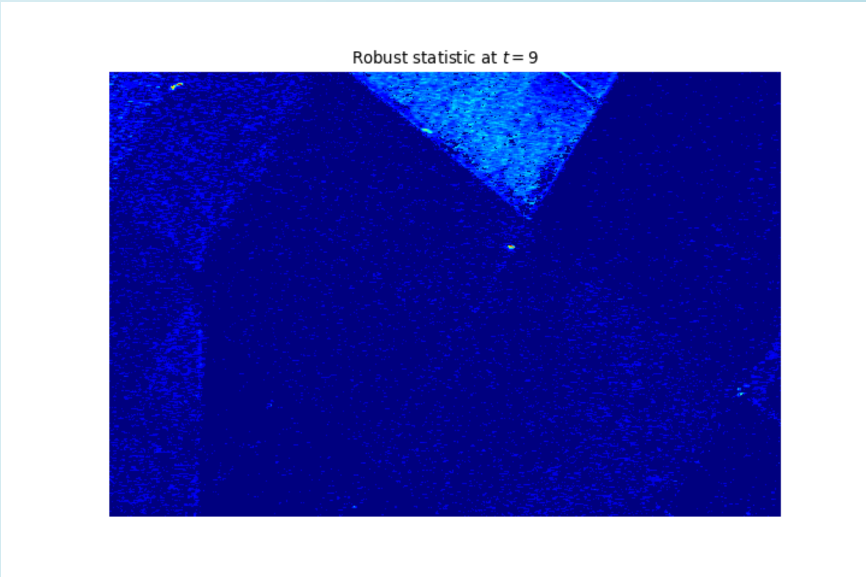


RESULT ON ANOTHER DATASET

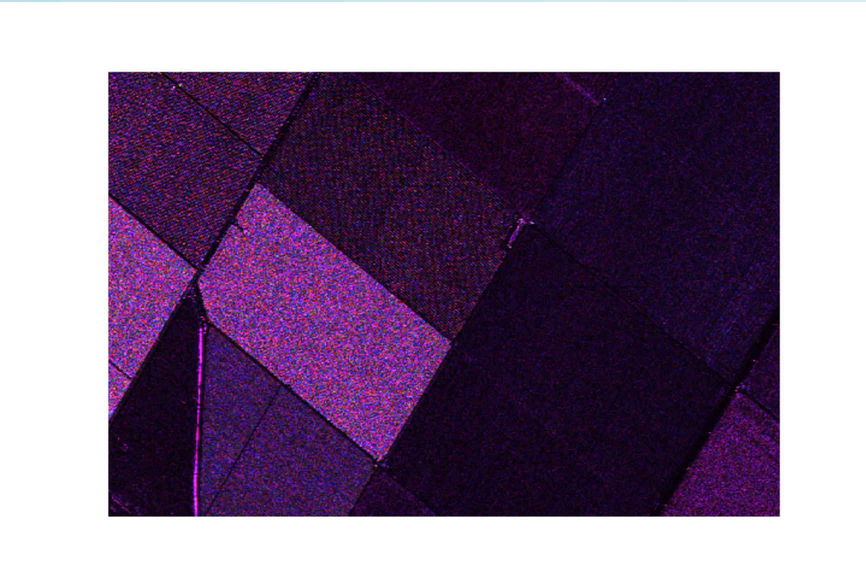


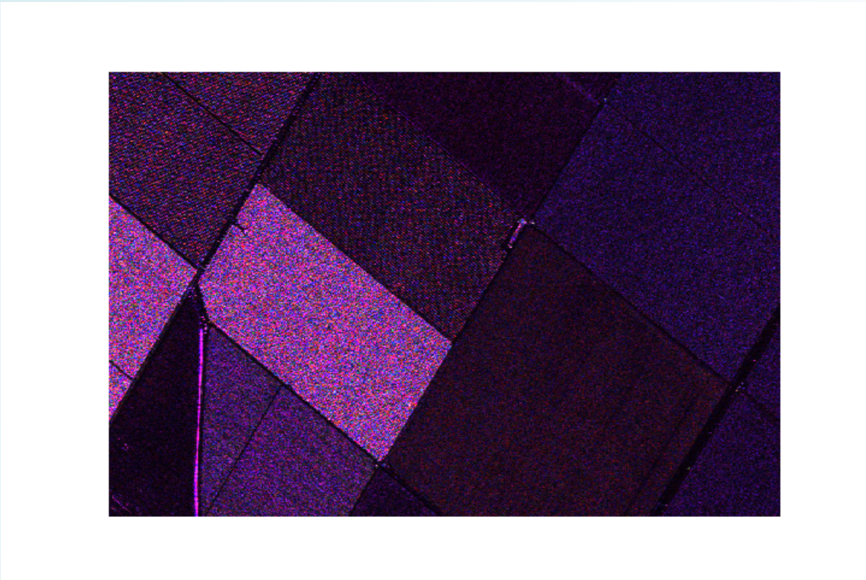


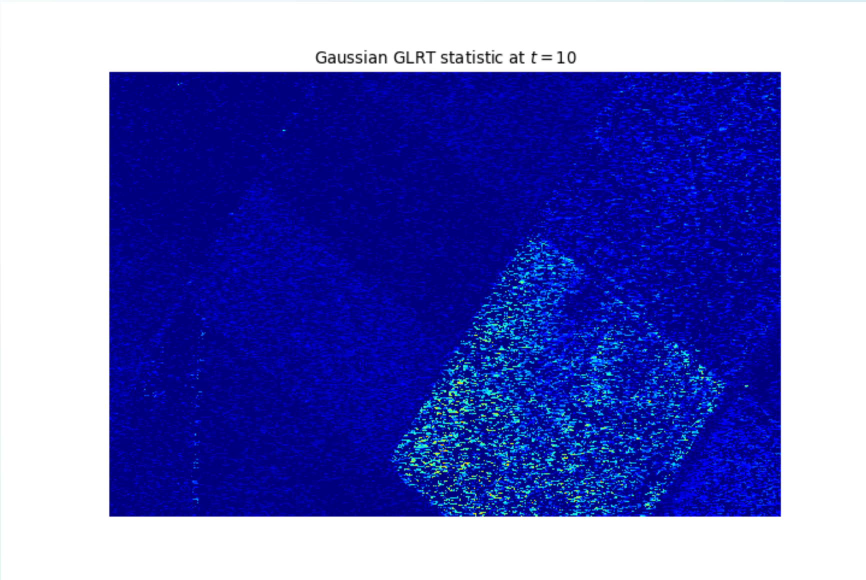


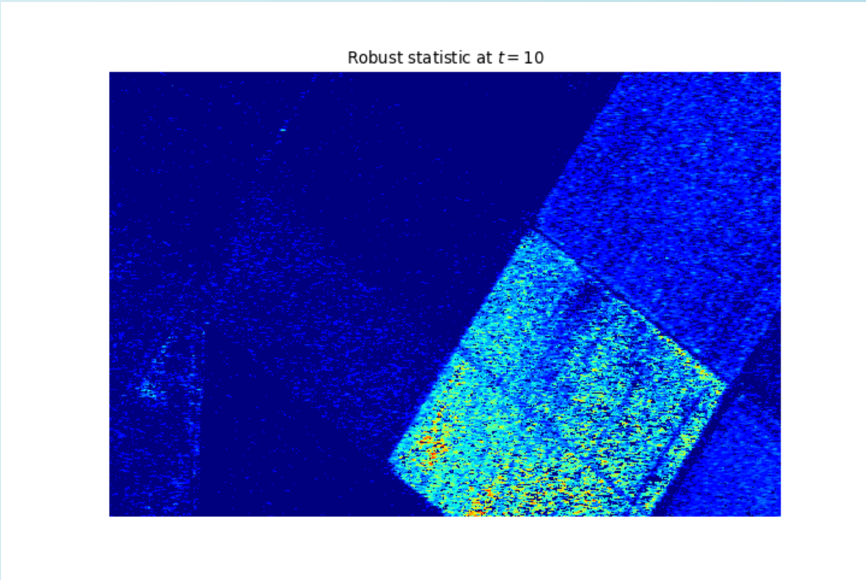


RESULT ON ANOTHER DATASET

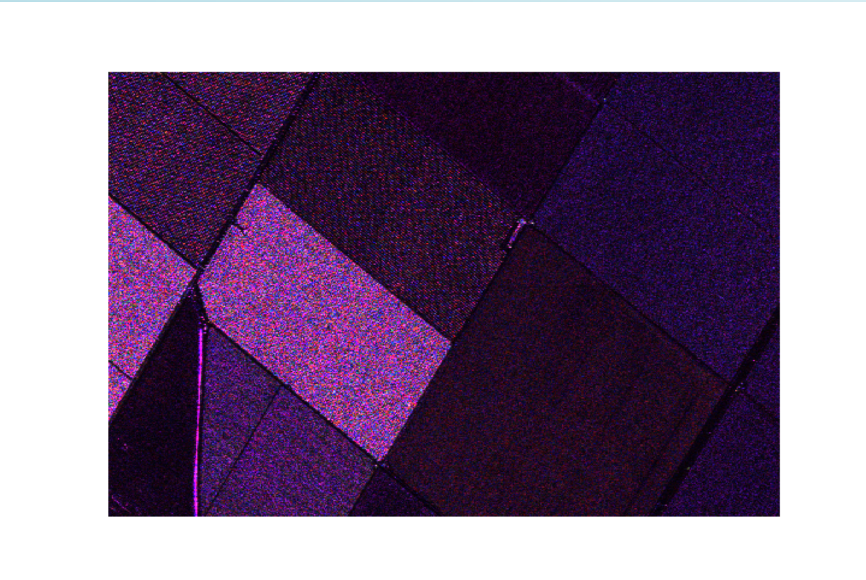


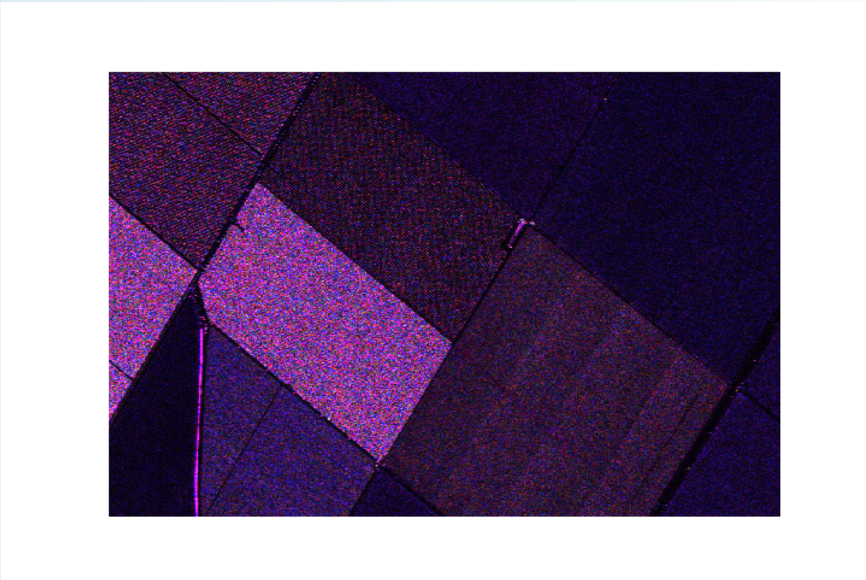


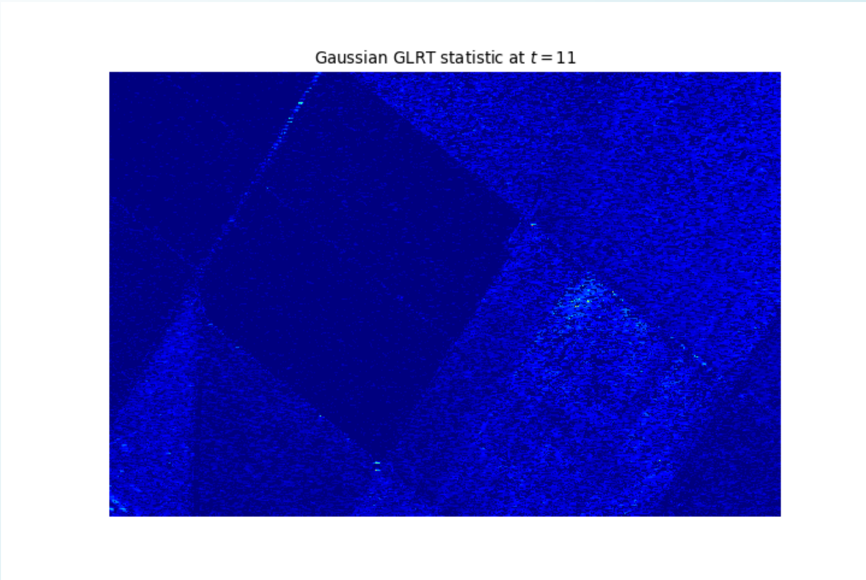


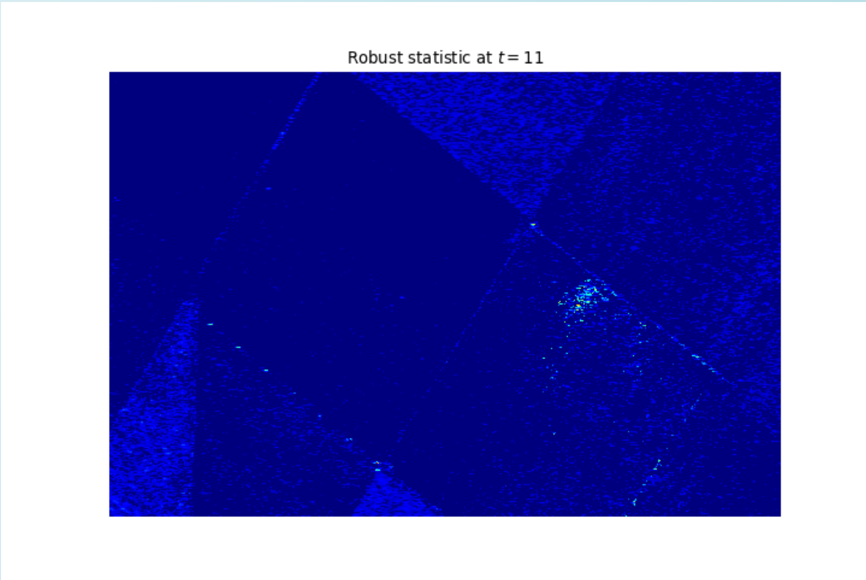


RESULT ON ANOTHER DATASET

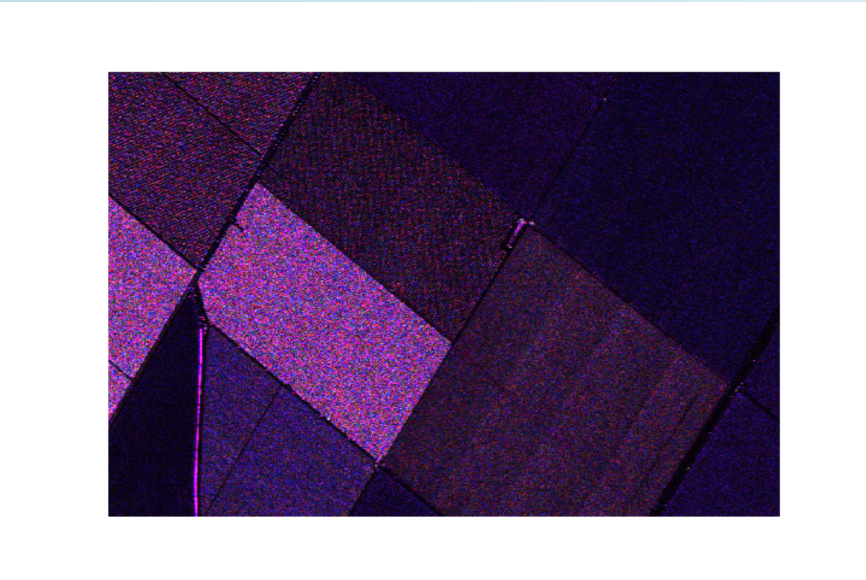


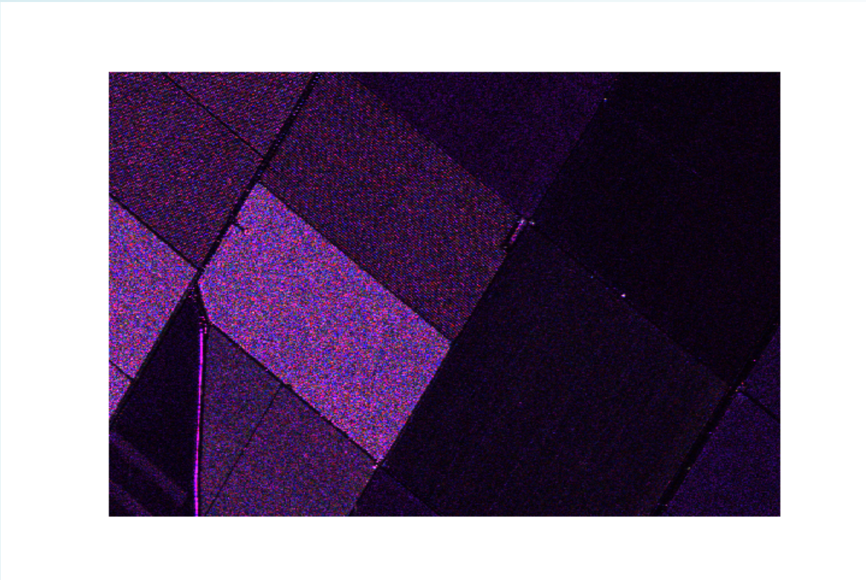


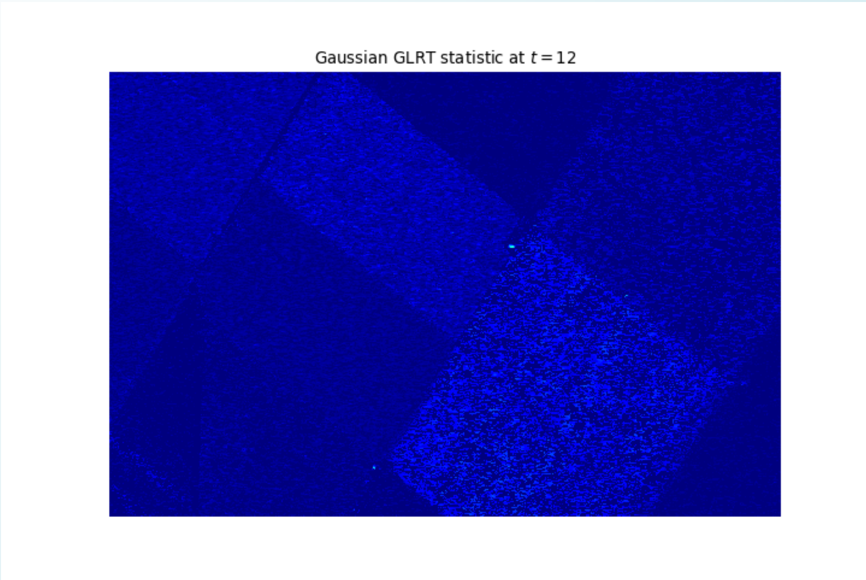


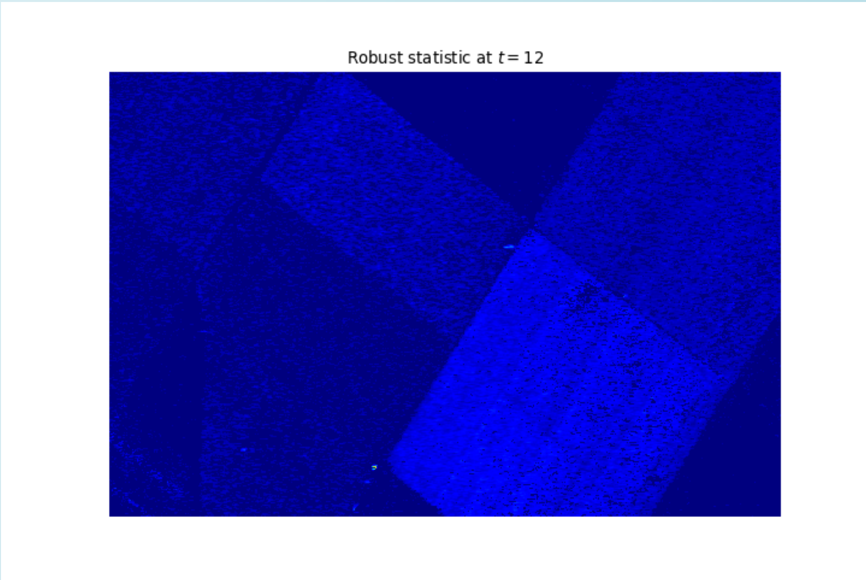


RESULT ON ANOTHER DATASET

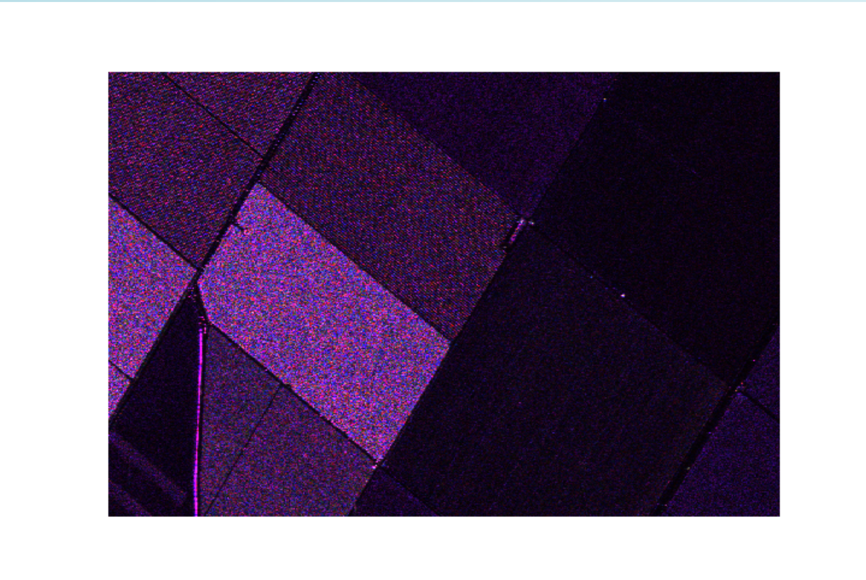


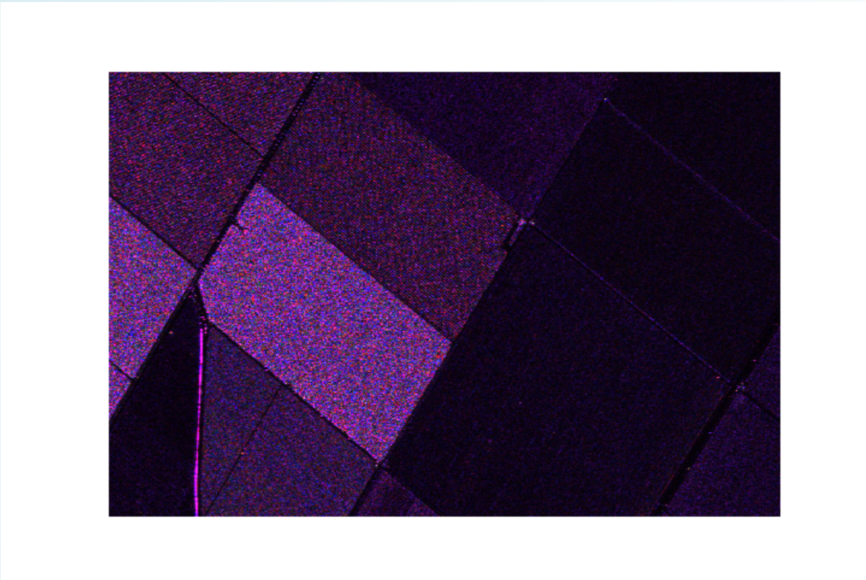


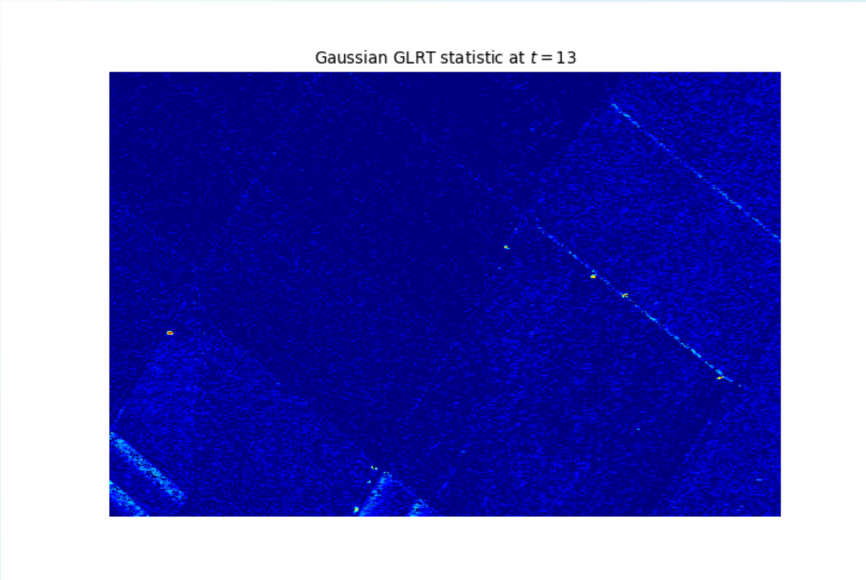


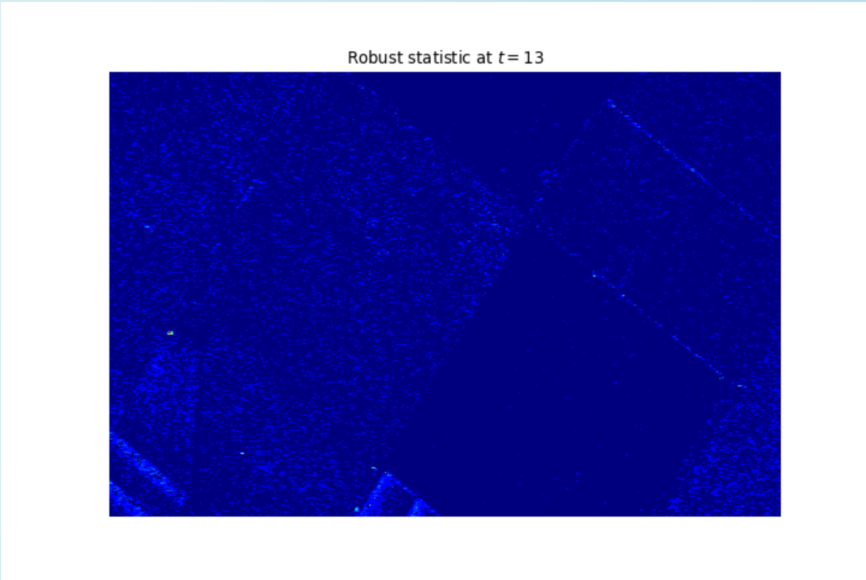


RESULT ON ANOTHER DATASET

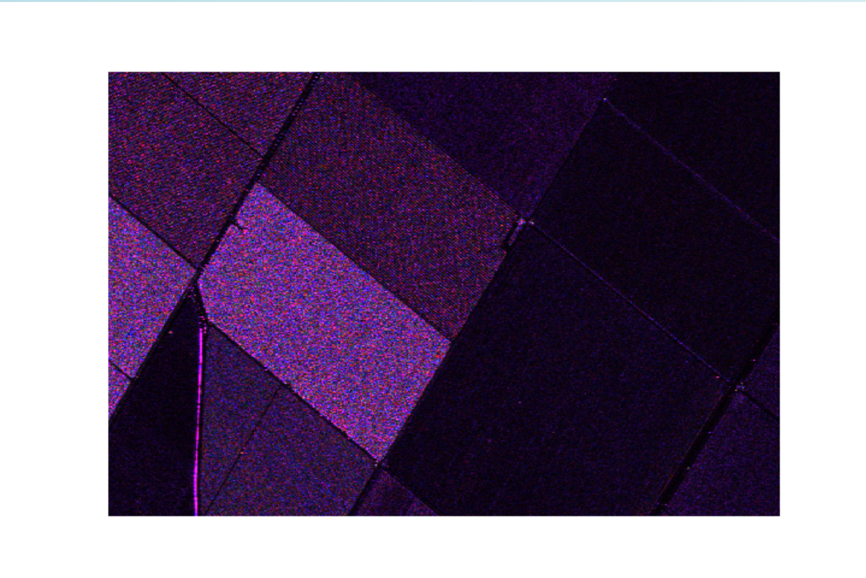


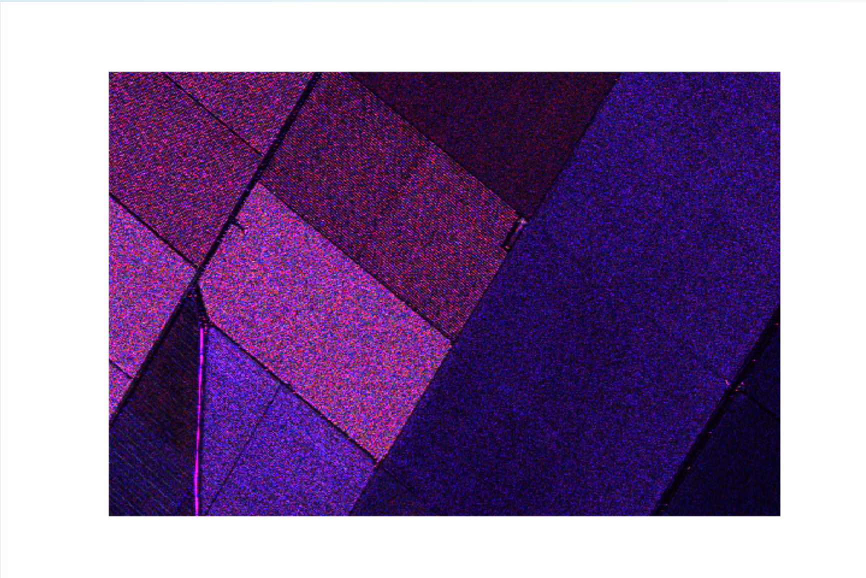


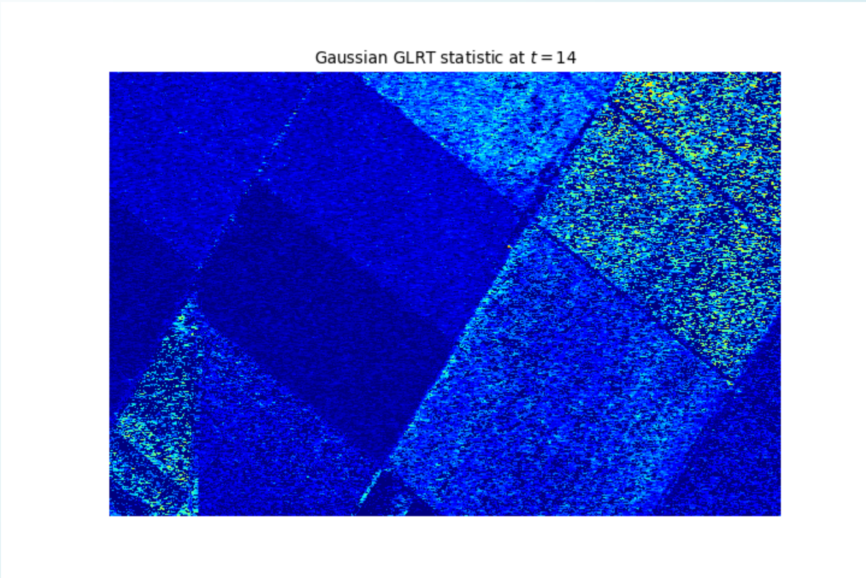


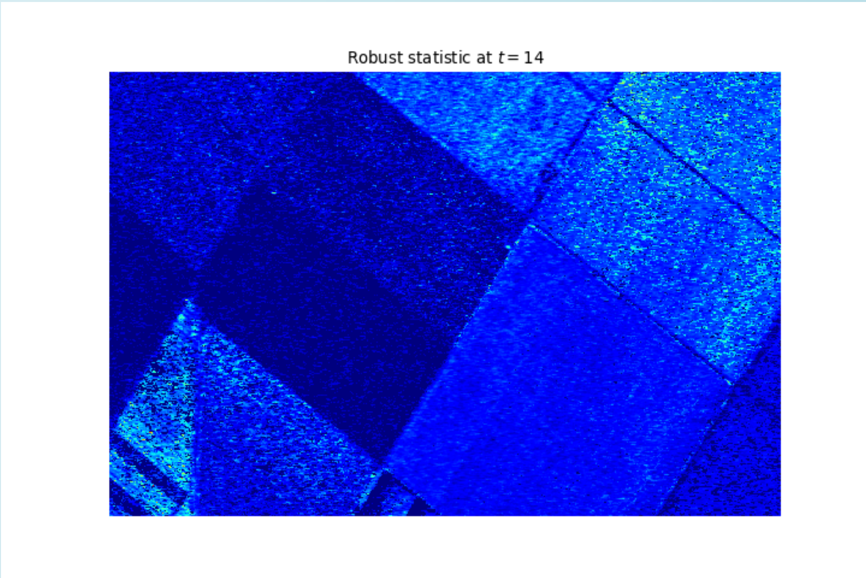


RESULT ON ANOTHER DATASET

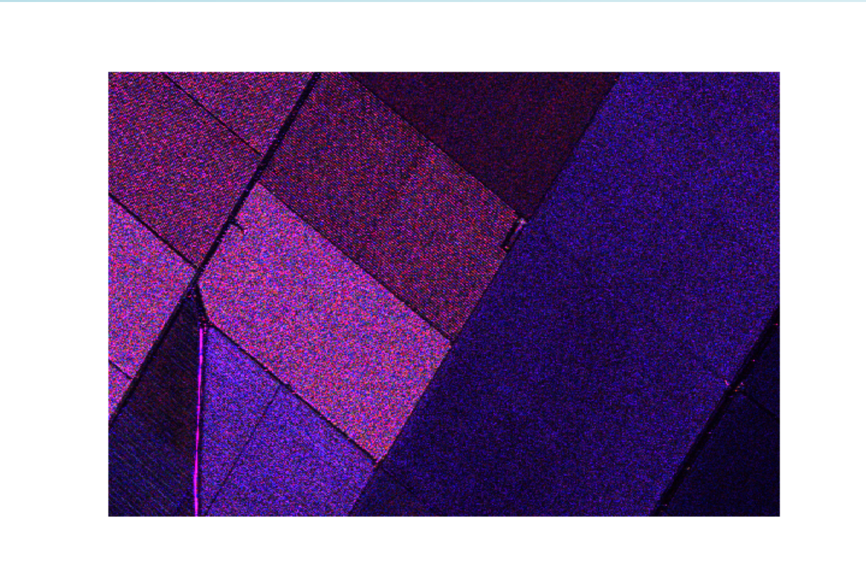


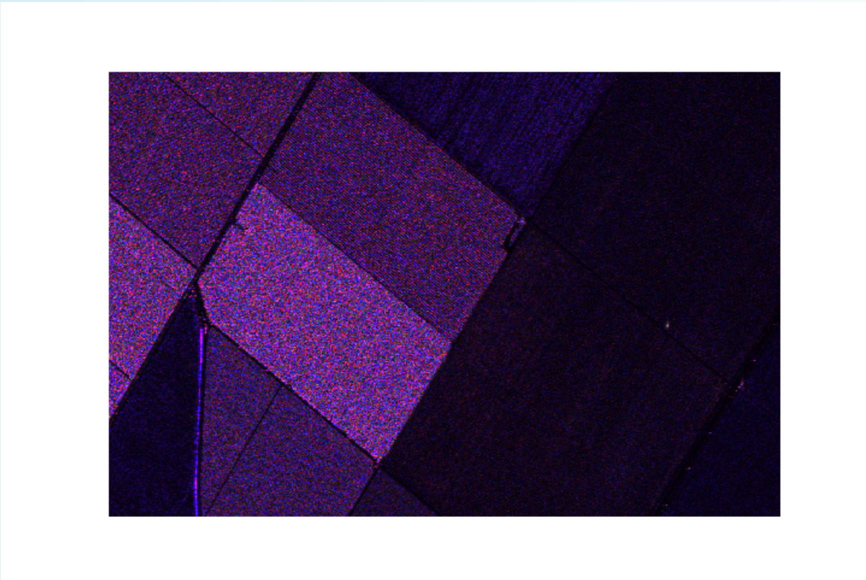


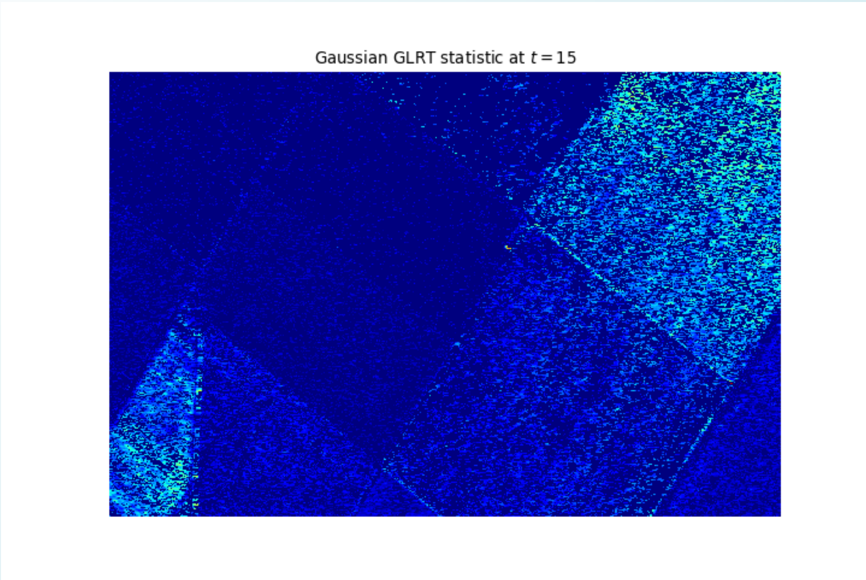


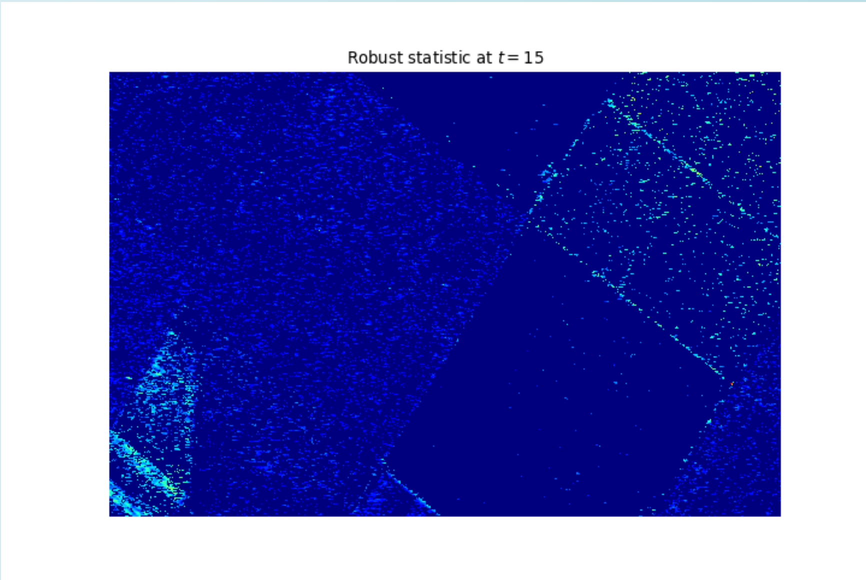


RESULT ON ANOTHER DATASET

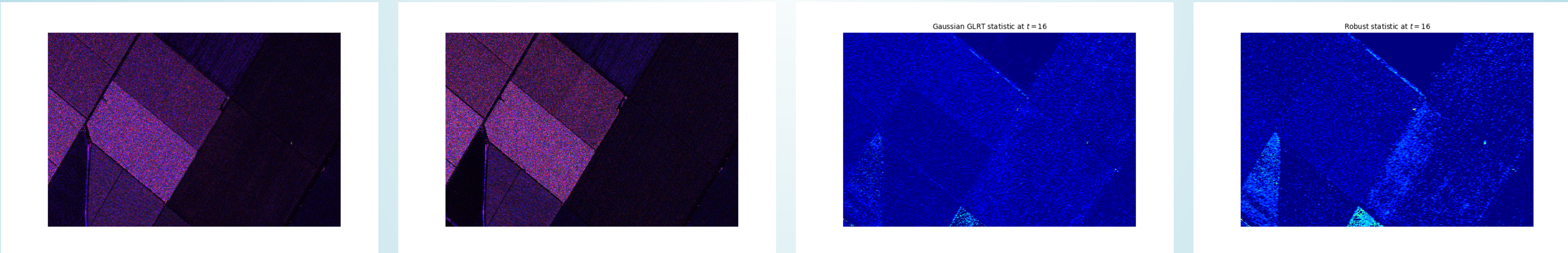








RESULT ON ANOTHER DATASET



THANKS FOR YOUR ATTENTION !