

Multidimensional rational covariance extension with applications to Wiener system identification

Axel Ringh^{*}, Johan Karlsson^{*}, and Anders Lindquist^{†*}

^{*}Department of Mathematics
KTH Royal Institute of Technology
Stockholm, Sweden

[†]Department of Automation
Shanghai Jiao Tong University, China
Shanghai, China

14th of December 2017, CDC2017



- Introduction
 - Motivation - texture generation
 - Wiener system identification
- Multidimensional rational covariance extension problem
 - Problem derivation
 - Solution via convex optimization
- Examples in texture generation

- Motivated by the use of thresholded Gaussian random fields to model porous materials [1], we are interested in generating **binary textures**.

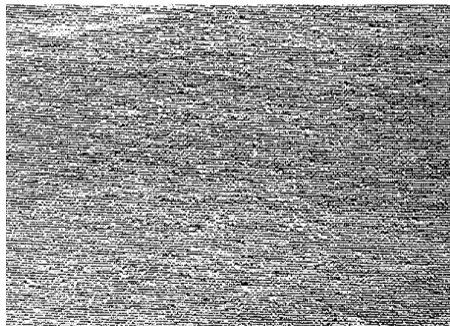


Figure: Example of a texture

- Motivated by the use of thresholded Gaussian random fields to model porous materials [1], we are interested in generating **binary textures**.
- We want to estimate a system that can generate similar textures
 \rightsquigarrow **multidimensional Wiener systems identification**.

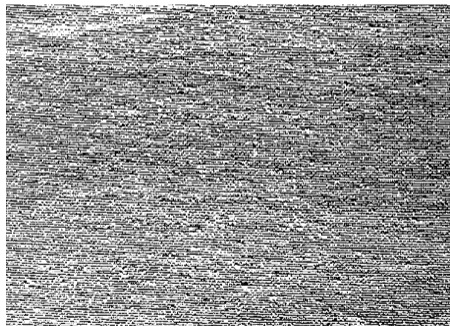


Figure: Example of a texture

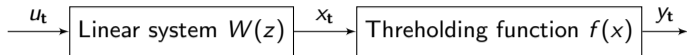


Figure: A Wiener system with thresholding as static nonlinearity.

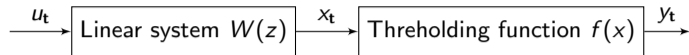


Figure: A Wiener system with thresholding as static nonlinearity.

- $\{u_t; \mathbf{t} \in \mathbb{Z}^d\}$ be a zero-mean Gaussian white noise input.

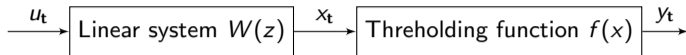


Figure: A Wiener system with thresholding as static nonlinearity.

- $\{u_t; \mathbf{t} \in \mathbb{Z}^d\}$ be a zero-mean Gaussian white noise input.
- The linear dynamical system is a strictly causal *autoregressive-moving-average* (ARMA) filter

$$x_t + \sum_{\mathbf{k} \in \Lambda_+ \setminus \{0\}} a_{\mathbf{k}} x_{t-\mathbf{k}} = \sum_{\mathbf{k} \in \Lambda_+} b_{\mathbf{k}} u_{t-\mathbf{k}} \Leftrightarrow W(\mathbf{z}) = \frac{\sum_{\mathbf{k} \in \Lambda_+} b_{\mathbf{k}} \mathbf{z}^{\mathbf{k}}}{\sum_{\mathbf{k} \in \Lambda_+} a_{\mathbf{k}} \mathbf{z}^{\mathbf{k}}} = \frac{b(\mathbf{z})}{a(\mathbf{z})}$$

where $\Lambda_+ \subset \mathbb{Z}^2$ is the support of the filter.

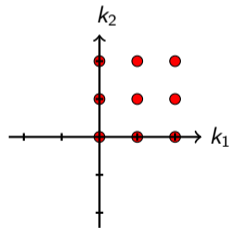


Figure: Example of Λ_+ .

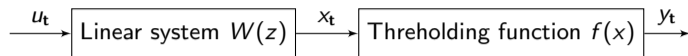


Figure: A Wiener system with thresholding as static nonlinearity.

- $\{u_t; t \in \mathbb{Z}^d\}$ be a zero-mean Gaussian white noise input.
- The linear dynamical system is a strictly causal *autoregressive-moving-average* (ARMA) filter

$$x_t + \sum_{k \in \Lambda_+ \setminus \{0\}} a_k x_{t-k} = \sum_{k \in \Lambda_+} b_k u_{t-k} \Leftrightarrow W(z) = \frac{\sum_{k \in \Lambda_+} b_k z^k}{\sum_{k \in \Lambda_+} a_k z^k} = \frac{b(z)}{a(z)}$$

where $\Lambda_+ \subset \mathbb{Z}^2$ is the support of the filter.

- Thresholding function

$$f(x) = \begin{cases} 1 & x > \tau \\ 0 & \text{otherwise} \end{cases}$$

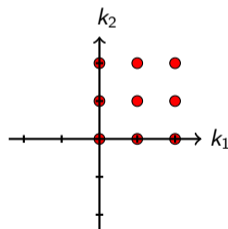


Figure: Example of Λ_+ .

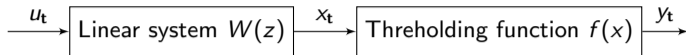


Figure: A Wiener system with thresholding as static nonlinearity.

- $\{u_t; \mathbf{t} \in \mathbb{Z}^d\}$ be a zero-mean Gaussian white noise input.
- The linear dynamical system is a strictly causal *autoregressive-moving-average* (ARMA) filter

$$x_t + \sum_{k \in \Lambda_+ \setminus \{0\}} a_k x_{t-k} = \sum_{k \in \Lambda_+} b_k u_{t-k} \Leftrightarrow W(z) = \frac{\sum_{k \in \Lambda_+} b_k z^k}{\sum_{k \in \Lambda_+} a_k z^k} = \frac{b(z)}{a(z)}$$

where $\Lambda_+ \subset \mathbb{Z}^2$ is the support of the filter.

- Thresholding function

$$f(x) = \begin{cases} 1 & x > \tau \\ 0 & \text{otherwise} \end{cases}$$

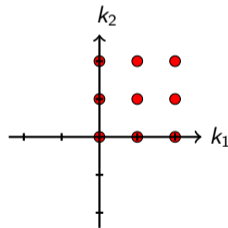


Figure: Example of Λ_+ .

Goal: From samples (y_t) we want to identify τ and W .

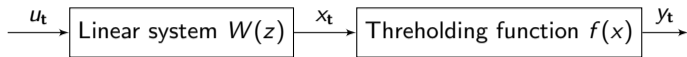


Figure: A Wiener system with thresholding as static nonlinearity.

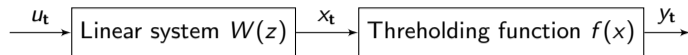


Figure: A Wiener system with thresholding as static nonlinearity.

- Identifying the threshold parameter:
 - Since u_t is zero-mean and Gaussian and $W(z)$ is linear, x_t is zero-mean and Gaussian.
 - $\mathbb{E}[y_t] = P(y_t = 1) = P(x_t > \tau) = 1 - P(x_t \leq \tau) = 1 - \phi(\tau)$, where ϕ is the Gaussian CDF
 \rightsquigarrow we can estimate τ as $\tau_{\text{est}} = \phi^{-1}(1 - \mathbb{E}[y_t])$.

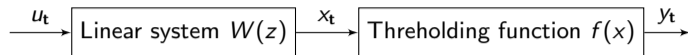


Figure: A Wiener system with thresholding as static nonlinearity.

- Identifying the threshold parameter:
 - Since u_t is zero-mean and Gaussian and $W(z)$ is linear, x_t is zero-mean and Gaussian.
 - $\mathbb{E}[y_t] = P(y_t = 1) = P(x_t > \tau) = 1 - P(x_t \leq \tau) = 1 - \phi(\tau)$, where ϕ is the Gaussian CDF
 \rightsquigarrow we can estimate τ as $\tau_{\text{est}} = \phi^{-1}(1 - \mathbb{E}[y_t])$.
- In steady-state x_t is second order stationary process, i.e., the covariances are independent of the absolute time \mathbf{t} : $c_k := \mathbb{E}[x_t x_{t-k}]$. Assume that $c_0 = 1$ (normalization). Estimate covariances c_k :

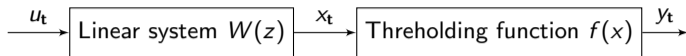


Figure: A Wiener system with thresholding as static nonlinearity.

- Identifying the threshold parameter:
 - Since u_t is zero-mean and Gaussian and $W(z)$ is linear, x_t is zero-mean and Gaussian.
 - $\mathbb{E}[y_t] = P(y_t = 1) = P(x_t > \tau) = 1 - P(x_t \leq \tau) = 1 - \phi(\tau)$, where ϕ is the Gaussian CDF
 \rightsquigarrow we can estimate τ as $\tau_{\text{est}} = \phi^{-1}(1 - \mathbb{E}[y_t])$.
- In steady-state x_t is second order stationary process, i.e., the covariances are independent of the absolute time \mathbf{t} : $c_k := \mathbb{E}[x_t x_{t-k}]$. Assume that $c_0 = 1$ (normalization). Estimate covariances c_k :
 - Let $r_k := \mathbb{E}[y_{t-k} y_t] - \mathbb{E}[y_{t-k}] \mathbb{E}[y_t]$ be the covariances of the process y_t .
 - Since x_t is Gaussian a theorem by Price [1] gives the following relationship between the covariances

$$r_k = \int_0^{c_k} \frac{1}{2\pi\sqrt{1-s^2}} \exp\left(-\frac{\tau^2}{1+s}\right) ds.$$

Since the integrand is positive, the mapping can be inverted (numerically).

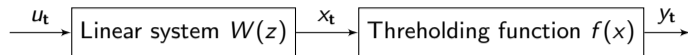


Figure: A Wiener system with thresholding as static nonlinearity.

- Identifying the threshold parameter:
 - Since u_t is zero-mean and Gaussian and $W(z)$ is linear, x_t is zero-mean and Gaussian.
 - $\mathbb{E}[y_t] = P(y_t = 1) = P(x_t > \tau) = 1 - P(x_t \leq \tau) = 1 - \phi(\tau)$, where ϕ is the Gaussian CDF
 \rightsquigarrow we can estimate τ as $\tau_{\text{est}} = \phi^{-1}(1 - \mathbb{E}[y_t])$.
- In steady-state x_t is second order stationary process, i.e., the covariances are independent of the absolute time \mathbf{t} : $c_k := \mathbb{E}[x_t x_{t-k}]$. Assume that $c_0 = 1$ (normalization). Estimate covariances c_k :
 - Let $r_k := \mathbb{E}[y_{t-k} y_t] - \mathbb{E}[y_{t-k}] \mathbb{E}[y_t]$ be the covariances of the process y_t .
 - Since x_t is Gaussian a theorem by Price [1] gives the following relationship between the covariances

$$r_k = \int_0^{c_k} \frac{1}{2\pi\sqrt{1-s^2}} \exp\left(-\frac{\tau^2}{1+s}\right) ds.$$

Since the integrand is positive, the mapping can be inverted (numerically).

- From the covariances c_k , estimate the linear system $W(z)$.

[1] R. Price. A useful theorem for nonlinear devices having Gaussian inputs. *IRE Transactions on Information Theory*, 4(2), 69-72, 1958.

Multidimensional rational covariance extension

Derivation of the problem

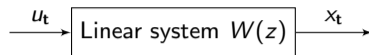


Figure: A linear system.

From covariance data $\{c_k\}_{k \in \Lambda}$ we want to estimate a linear system $W(z)$.

- The **power spectral density** $\Phi(e^{i\theta})$ of a stochastic process $\{x_t; t \in \mathbb{Z}^d\}$ is defined as the nonnegative function such that

$$c_k := \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i(k,\theta)} \Phi(e^{i\theta}) d\theta, \quad k \in \mathbb{Z}^2 \quad \iff \quad \Phi(e^{i\theta}) = \sum_{k \in \mathbb{Z}^2} c_k e^{-i(k,\theta)}.$$

Multidimensional rational covariance extension

Derivation of the problem

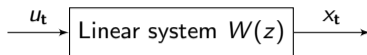


Figure: A linear system.

From covariance data $\{c_k\}_{k \in \Lambda}$ we want to estimate a linear system $W(z)$.

- The **power spectral density** $\Phi(e^{i\theta})$ of a stochastic process $\{x_t; t \in \mathbb{Z}^d\}$ is defined as the nonnegative function such that

$$c_k := \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i(k,\theta)} \Phi(e^{i\theta}) d\theta, \quad k \in \mathbb{Z}^2 \quad \iff \quad \Phi(e^{i\theta}) = \sum_{k \in \mathbb{Z}^2} c_k e^{-i(k,\theta)}.$$

- Moreover,

$$\Phi(e^{i\theta}) = |W(e^{i\theta})|^2 \Phi_u(e^{i\theta}) = |W(e^{i\theta})|^2 = \frac{|b(e^{i\theta})|^2}{|a(e^{i\theta})|^2} = \frac{\sum_{k \in \Lambda} p_k e^{-i(k,\theta)}}{\sum_{k \in \Lambda} q_k e^{-i(k,\theta)}} = \frac{P(e^{i\theta})}{Q(e^{i\theta})}.$$

where P and Q are **trigonometric polynomials**, and $\Lambda = \Lambda_+ - \Lambda_+$ (Minkowski set difference).

Multidimensional rational covariance extension

Derivation of the problem

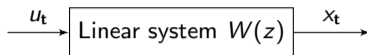


Figure: A linear system.

From covariance data $\{c_k\}_{k \in \Lambda}$ we want to estimate a linear system $W(z)$.

- Summarizing, this gives us the following problem:

Problem formulation – Approximate multidimensional rational covariance extension problem

Given a sequence of covariances $c = (c_k)_{k \in \Lambda}$ find a positive function $\Phi(e^{i\theta})$ so that

$$\begin{cases} c_k \approx \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i(k,\theta)} \Phi(e^{i\theta}) d\theta, & k \in \Lambda \\ \Phi(e^{i\theta}) = \frac{P(e^{i\theta})}{Q(e^{i\theta})}, & P \text{ and } Q \in \bar{\mathfrak{P}}_+. \end{cases}$$

- Notation for nonnegative trigonometric polynomials:

$$\bar{\mathfrak{P}}_+ = \{p := (p_k)_{k \in \Lambda} \in \mathbb{C}^{|\Lambda|} \mid p_{-k} = p_k^*, P(e^{i\theta}) := \sum_{k \in \Lambda} p_k e^{-i(k,\theta)}, P(e^{i\theta}) \geq 0 \text{ for all } \theta \in \mathbb{T}^2\}$$

- **Rational covariance extension:**

- R. Kalman, 1981.
- T.T. Georgiou, 1983.
- T.T. Georgiou, 1987.
- C.I. Byrnes, A. Lindquist, S.V. Gusev, and A.S. Matveev, 1995.
- C.I. Byrnes, S.V. Gusev, and A. Lindquist, 1998.
- C.I. Byrnes, T.T. Georgiou, and A. Lindquist, 2000.
- C.I. Byrnes, P. Enqvist, and A. Lindquist, 2001.
- P. Enqvist, 2004.
- H.I. Nurdin, and A. Bagchi, 2006.
- T.T. Georgiou, and A. Lindquist, 2008.
- A. Ferrante, M. Pavon, and M. Zorzi, 2012.
- M. Zorzi, 2014.

- **Periodic/Circulant problem:**

- A. Lindquist and G. Picci, 2013.
- A. Lindquist, C. Masiero, and G. Picci, 2013.
- A. Ringh, and A. Lindquist, 2014.
- G. Picci, and B. Zhu, 2017.

- **Matrix valued**

- A. Blomqvist, A. Lindquist, R. Nagamune, 2003.
- F. Ramponi, A. Ferrante, and M. Pavon, 2009.
- M. Pavon, and A. Ferrante, 2013.
- M. Zorzi, 2014.
- B. Zhu, 2017.
- B. Zhu, and G. Baggio, 2017.

- **Multidimensional problem:**

- T.T. Georgiou, 2005.
- T.T. Georgiou, 2006.
- A. Ringh, J. Karlsson, and A. Lindquist, 2015.
- A. Ringh, J. Karlsson, and A. Lindquist, 2016.
- J. Karlsson, A. Lindquist, and A. Ringh, 2016.
- A. Ringh, J. Karlsson, and A. Lindquist, 2017.

Notation: let \mathfrak{C}_+ be the interior of dual cone of $\bar{\mathfrak{P}}_+$:

$$\mathfrak{C}_+ := \{c \in \mathbb{C}^{|\Lambda|} \mid c_{-k} = c_k^*, \langle c, p \rangle := \sum_{k \in \Lambda} c_k p_k^* > 0 \text{ for all } p \in \bar{\mathfrak{P}}_+ \setminus \{0\}\}$$

Multidimensional rational covariance extension

Solution with exact matching

Notation: let \mathfrak{C}_+ be the interior of dual cone of $\bar{\mathfrak{P}}_+$:

$$\mathfrak{C}_+ := \{c \in \mathbb{C}^{|\Lambda|} \mid c_{-k} = c_k^*, \langle c, p \rangle := \sum_{k \in \Lambda} c_k p_k^* > 0 \text{ for all } p \in \bar{\mathfrak{P}}_+ \setminus \{0\}\}$$

Theorem

Given a sequence c , the primal problem

$$(P) \quad \min_{\Phi > 0} \int_{\mathbb{T}^2} \left(P \log \frac{P}{\Phi} + \Phi - P \right) \frac{d\theta}{(2\pi)^2}$$

subject to $c_k = \int_{\mathbb{T}^2} e^{i(k, \theta)} \Phi(e^{i\theta}) \frac{d\theta}{(2\pi)^2}, \quad k \in \Lambda,$

has a solution if and only if $c \in \mathfrak{C}_+$. Moreover, for every $c \in \mathfrak{C}_+$ and trigonometric polynomial $P > 0$ the solution to (P) is given by

$$\Phi(e^{i\theta}) = \frac{P(e^{i\theta})}{\hat{Q}(e^{i\theta})}$$

where \hat{Q} is the unique solution to the dual problem

$$(D) \quad \min_{q \in \bar{\mathfrak{P}}_+} \langle c, q \rangle - \int_{\mathbb{T}^2} P \log(Q) \frac{d\theta}{(2\pi)^2}.$$

Theorem

Given any sequence c , for ε large enough the primal problem

$$(P) \quad \min_{\Phi > 0, \tilde{c}} \int_{\mathbb{T}^2} \left(P \log \frac{P}{\Phi} + \Phi - P \right) \frac{d\theta}{(2\pi)^2}$$

$$\text{subject to } \tilde{c}_k = \int_{\mathbb{T}^2} e^{i(k, \theta)} \Phi(e^{i\theta}) \frac{d\theta}{(2\pi)^2}, \quad \mathbf{k} \in \Lambda,$$

$$\|\tilde{c} - c\|^2 \leq \varepsilon^2,$$

has an optimal solution given by

$$\Phi(e^{i\theta}) = \frac{P(e^{i\theta})}{\hat{Q}(e^{i\theta})},$$

where \hat{Q} is the unique solution to the dual problem

$$(D) \quad \min_{q \in \tilde{\mathfrak{P}}_+} \langle c, q \rangle - \int_{\mathbb{T}^2} P \log(Q) \frac{d\theta}{(2\pi)^2} + \varepsilon \|q - e\|,$$

where $e \in \mathbb{C}^{|\Lambda|}$, $e_0 = 1$ and $e_k = 0$ for $\mathbf{k} \in \Lambda \setminus \{0\}$.

- How do we obtain the system from the spectrum?
- In one dimension, **spectral factorization** as a **sum-of-one-square** is always possible:

$$\frac{P(e^{i\theta})}{Q(e^{i\theta})} \xrightarrow[\text{Trivial}]{\text{Spectral factorization}} \frac{|b(e^{i\theta})|^2}{|a(e^{i\theta})|^2}$$

- How do we obtain the system from the spectrum?
- In one dimension, **spectral factorization** as a **sum-of-one-square** is always possible:

$$\frac{P(e^{i\theta})}{Q(e^{i\theta})} \xrightleftharpoons[\text{Trivial}]{\text{Spectral factorization}} \frac{|b(e^{i\theta})|^2}{|a(e^{i\theta})|^2}$$

- Not true in the two dimensions - only factorization as **sum-of-several-squares** can be guaranteed [1, 2]:

$$\frac{P(e^{i\theta})}{Q(e^{i\theta})} = \frac{\sum_{k=1}^{\ell} |b_k(e^{i\theta})|^2}{\sum_{k=1}^m |a_k(e^{i\theta})|^2}.$$

- **Open questions:**

- How to construct a realization from such a spectrum?
- Is it possible to characterize P for which Q is a sum-of-one-square?

[1] M.A. Dritschel. On factorization of trigonometric polynomials. *Integral Equations and Operator Theory*, 49(1), 11-42, 2004.

[2] J.S. Geronimo, and M.J. Lai. Factorization of multivariate positive Laurent polynomials. *Journal of Approximation Theory*, 139(1-2), 327-345, 2006.

- How do we obtain the system from the spectrum?
- In one dimension, **spectral factorization** as a **sum-of-one-square** is always possible:

$$\frac{P(e^{i\theta})}{Q(e^{i\theta})} \xrightleftharpoons[\text{Trivial}]{\text{Spectral factorization}} \frac{|b(e^{i\theta})|^2}{|a(e^{i\theta})|^2}$$

- Not true in the two dimensions - only factorization as **sum-of-several-squares** can be guaranteed [1, 2]:

$$\frac{P(e^{i\theta})}{Q(e^{i\theta})} = \frac{\sum_{k=1}^{\ell} |b_k(e^{i\theta})|^2}{\sum_{k=1}^m |a_k(e^{i\theta})|^2}.$$

- **Open questions:**

- How to construct a realization from such a spectrum?
- Is it possible to characterize P for which Q is a sum-of-one-square?
- We resort to a heuristic, obtained by “abusing” results in [3].

[1] M.A. Dritschel. On factorization of trigonometric polynomials. *Integral Equations and Operator Theory*, 49(1), 11-42, 2004.

[2] J.S. Geronimo, and M.J. Lai. Factorization of multivariate positive Laurent polynomials. *Journal of Approximation Theory*, 139(1-2), 327-345, 2006.

[3] J.S. Geronimo, and H.J. Woerdeman. Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables. *Annals of Mathematics*, 839-906, 2004.



Figure: A Wiener system with thresholding as static nonlinearity.

Algorithm for Wiener system identification with thresholding

Input: $\{y_t\}$

1: Estimate threshold parameter: $\tau_{\text{est}} = \phi^{-1}(1 - E[y_t])$ from the data.

2: Estimate covariances: $r_k := E[y_{t-k}y_t] - E[y_{t-k}]E[y_t]$ from the data.

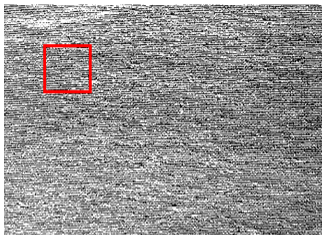
3: Compute covariances $c_k := E[x_{t-k}x_t]$ by using the relation $r_k = \int_0^{c_k} \frac{1}{2\pi\sqrt{1-s^2}} \exp\left(-\frac{\tau^2}{1+s}\right) ds$

4: Estimate a rational spectrum using the theory developed here.

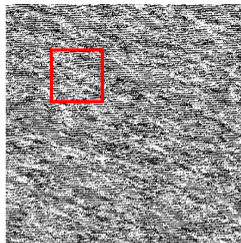
5: Apply a *heuristic*, approximate factorization procedure.

Output: τ_{est} , coefficients for the linear dynamical system

Example in texture generation



(a) Texture.



(b) Reconstruction.

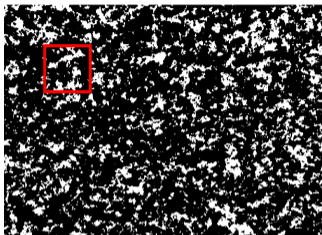


(c) Close-up of the texture.

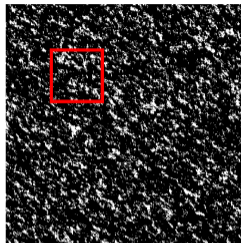


(d) Close-up of the reconstruction.

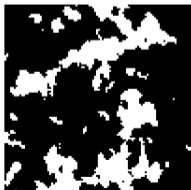
Example in texture generation



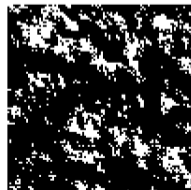
(e) Texture.



(f) Reconstruction.

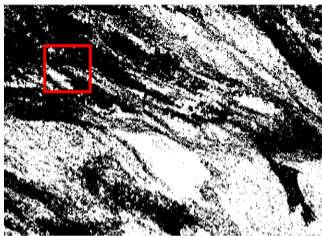


(g) Close-up of the texture.

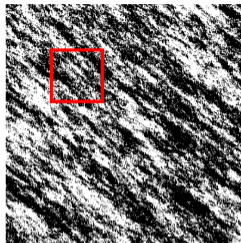


(h) Close-up of the reconstruction.

Example in texture generation



(i) Texture.



(j) Reconstruction.



(k) Close-up of the texture.



(l) Close-up of the reconstruction.

We can do estimation of rational multidimensional spectra.
But we need to better understand what they mean in terms of dynamical systems.

Thank you for your attention!

Questions?