ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
"Antonio Ruberti"
CONSIGLIO NAZIONALE DELLE RICERCHE

F. Carravetta

2D-RECURSIVE MODELLING OF HOMOGENEOUS DISCRETE GAUSSIAN MARKOV FIELDS

R. 10-08, 2010

F. Carravetta – Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Roma, Viale Manzoni 30, 00185 Italy,
e-mail: `francesco.carravetta@iasi.cnr.it`.

ISSN: 1128–3378

Collana dei Rapporti dell'Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", CNR

viale Manzoni 30, 00185 ROMA, Italy

tel. ++39-06-77161
fax ++39-06-7716461
email: iasi@iasi.cnr.it
URL: http://www.iasi.cnr.it

Abstract

A 2D-stochastic realization is obtained for an homogeneous discrete Gaussian Markov-field in two dimensions. In specific, it is shown that such Markov fields can be represented as the (unique) pathwise solution of a *2D-recursive* system of equations, with a locally-correlated “forcing-noise”. The idea of stochastic “dynamic” system in two dimensions gets realized this way. Two types of domain are considered: spheric and rectangular.

Key words: Realization theory, modeling, stochastic systems, stochastic fields.

1. Introduction

The problem of getting dynamical models from certain nice classes of non-Markov stochastic processes has received attention in last decades. Both one-dimensional and two-dimensional (2D) models have been dealt with, the first class being also known as *reciprocal models* [1]-[8], the second as *Markov fields* [9]-[14] . What such kinds of stochastic processes share is the reliability in modelling physical/engineering systems where some intrinsic *non causal* phenomena is prominent. As a typical example in 2D, it should be mentioned Image Processing, where the concept of *Markov field* has been indeed a widely used tool. The reader is referred to [9]-[14] as a brief list of works more pertaining to the present article, and in particular to the book of Guyon [14] for a thorough description of Markov-fields theory.

Such "double sweep" property (it has been named this way in [4]) of reciprocal processes, is a result of paramount importance, in that it yields a *recursive* model of a non-causal process, a clever way indeed of managing the acausal feature. The problem of getting such decomposition properties has been also considered in two dimensions. In the paper [11] a discrete 2D-Markov field, defined in a bounded rectangle, has been considered, and a *nearest-neighbor* model has been built up which is converted into an equivalent 1D two-point boundary-value descriptor system, indeed the same kind of system formerly considered in [4]. The paper [11] is being here considered subsequent to [4] from a logical (not chronological) point of view. Indeed, although the results issued first (by Krener: [2], [3]) were concerning one dimensional processes – we also point out in passing [16], [17] as to the more general topic of acausal realization – Krener and Levy have been faced with representing Markov-fields during about the same time period, so some overlapping appears in the literature between the results issued in one and two dimensions. The way used in [11] as to reducing the 2D-model into one dimension consists in gathering all values of a column in a single vector. Then, such column vectors are shown satisfying a two-point descriptor system evolving in the discrete "time" represented by the column index. However, reducing this way the model to one-dimension seems to be quite expensive from a computational point of view when the number of rows is large. Generally speaking, the question as to whether the nice "double sweep" property, outlined above, of the one-dimensional model of [4] can be extended or not to a 2D-Markov field has remained unsolved.

In the present paper I propose a solution to this issue for a Gaussian, discrete, homogeneous 2D-Markov field. I will show that, under some technical assumption, the field gets always represented as a *2D-recursive* system, without introducing a state-space of augmented dimension. Also, the 2D-recursive model here presented accounts of a more general correlation scheme than in [11], i.e. the whole set of eight sites surrounding each site of the field is here assumed as the nearest-neighbor set. An explaination of the advantage of such $2D$-recursive model with respect to the model presented in [11], will be given the end of §IV, at the stage of model equations having been written down. The model will be worked out accounting at the same time of two different domains for the field: spheric and rectangular.

The paper is organized as follows, in §II and III the set up is presented, as well as some essential structural properties of the model. Section IV include the main result, above outlined. Finally, §V is a conclusive section, where also further possible developments are pointed out.

2. Model Setting and Local Properties

We will need the following notation. Given two real random variables θ, ρ, $\mathbf{E}\{\theta\}$, and $\mathbf{E}\{\theta|\rho\}$ denote expected value, and conditional expectation respectively. Given a finite set of ma-

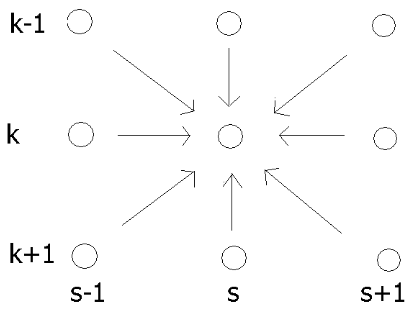

Figure 1: Correlation structure

trices $M_1, ..., M_l$, all having the same number of columns, we will write: $\text{col}(M_1, ..., M_l) = [M_1^T, ..., M_l^T]^T$, where $\cdot^T$ denotes transpose. Throughout the paper, the following convention will be widely used (for vectors *only*, not for matrices): if $\xi(i,j)$ is a two-index vector process, $(i,j) \in \mathbf{I} \times \mathbf{J}$; $\mathbf{I}, \mathbf{J}$ discrete intervals with N_I, N_J number of elements respectively, then

$$\begin{aligned}
&\xi(s) = \text{col}(\xi(0,s), ..., \xi(N_I - 1, s)) \\
&\xi = \text{col}(\xi(0), ..., \xi(N_J - 1)) \\
&\xi^-(s) = \text{col}(\xi(0,s), ..., \xi(N_I - 2, s)) \\
&\xi'(s) = \text{col}(\xi(1,s), ..., \xi(N_I - 2, s)) \\
&\xi' = \text{col}(\xi(1), ..., \xi(N_J - 2)) \\
&\xi^- = \text{col}(\xi(0), ..., \xi(N_J - 2)).
\end{aligned} \tag{1}$$

Moreover, for $\mathbf{S} \subset \mathbf{I} \times \mathbf{J}$, $\xi[\mathbf{S}] = \{\xi(i,j), (i,j) \in \mathbf{S}\}$. We are dealt with a Gaussian discrete *Markov field* in two dimensions, i.e. a Gaussian stochastic process $\{x(k,s), (k,s) \in \mathbb{N}^2\} \subset \mathcal{V}(\mathbb{R}^n)$, taking values at each $\mathbb{N}^2$-element, which also are said *sites* – or *pixels*, while thinking of an image – where $\mathcal{V}(\mathbb{R}^n)$ denotes a space of square-integrable random vectors in $\mathbb{R}^n$, such that (without loss of generality) $\mathbf{E}\{x(k,s)\} = 0$, $\forall (k,s) \in \mathbb{N}^2$ and satisfying the *Markov-field property*, which, in a sufficient to our purposes generality level, can be stated as follows. It is defined with respect to the eight *nearest neighbour sites* surrounding each site in $\mathbb{N}^2$. In Fig. 1 the underlying correlation structure between pixels we should have in mind is sketched.

Markov field property : For any $(k,s) \in \mathbb{N}^2$, let $\mathcal{I}(k,s)$ be the nearest-neighbour set:

$$\mathcal{I}(k,s) = \{(i,j) \neq (k,s), \;\; (i,j) = (k \pm p, s \pm p'), \; p, p' = 0, 1\}.$$

Then, the Markov-field property is characterized by the following identity:

$$\mathbf{E}\{x(k,s) | x(i,j), (i,j) \neq (k,s)\} = \mathbf{E}\{x(k,s) | x[\mathcal{I}(k,s)]\}.$$

If in addition the Markov field is Gaussian then

$$\mathbf{E}\{x(k,s) | x[\mathcal{I}(k,s)]\} = \sum_{(\alpha,\beta) \in \mathcal{I}'} F_{\alpha,\beta}(k,s) x(k+\alpha, s+\beta), \tag{2}$$

where $\mathcal{I}' = \{(q,p), \quad q,p = \pm 1, 0\} \setminus \{(0,0)\}$, for a suitable set of *projection matrices*: $\{F_{\alpha,\beta}(k,s), (\alpha,\beta) \in \mathcal{I}'\}$. For the Markov field $\{x(k,s), (k,s) \in \mathbb{N}^2\}$ we have the following basic result.

(*Representation theorem)*

Theorem 2.1. *For any $(k,s) \in \mathbb{N}^2$, the Markov field $x(k,s)$ satisfies the following representation*

$$x(k,s) = \sum_{(\alpha,\beta)\in\mathcal{I}'} F_{\alpha,\beta}(k,s)x(k+\alpha,s+\beta) + d(k,s), \tag{3}$$

where $\{d(k,s),(k,s) \in \mathbb{N}^2\}$ is characterized as follows

$$\mathbf{E}\{d(k,s)d^T(l,r)\} = 0, \text{ for } (l,r) \notin \mathcal{I}(k,s) \cup \{(k,s)\}, \tag{4}$$

$$\begin{aligned} D(k,s) &= \mathbf{E}\{d(k,s)d^T(k,s)\} \\ &= \mathbf{E}\{d(k,s)x^T(k,s)\}. \end{aligned} \tag{5}$$

Also, for any $\mathbb{R}^n$-valued Borel function ϕ (orthogonality property):

$$\mathbf{E}\{d(k,s)\phi^T(x(i,j))\} = 0, \quad \text{for} \quad (i,j) \neq (k,s). \tag{6}$$

Finally, for $(\alpha,\beta) \in \mathcal{I}'$, the autocorrelation function of d results in

$$\begin{aligned} \mathbf{E}\{d(k,s)d^T(k+\alpha,s+\beta)\} &= -F_{\alpha,\beta}(k,s)D(k+\alpha,s+\beta), \\ &= -D(k,s)F^T_{-\alpha,-\beta}(k+\alpha,s+\beta). \end{aligned} \tag{7}$$

Proof. Let us *define* $d(k,s)$ by using eq. (3), so, by the Markov field property and (2), the Theorem reduces in proving just eqs. (4), (5), (6), and (7). For, first of all notice that $\mathbf{E}\{d(k,s)x^T(i,j)\} = 0$ for $(i,j) \neq (k,s)$ by construction of d, therefore, since the field is Gaussian orthogonality implies independence, and (6) follows. Now, on account of (2), by eq. (3) the following relation is readily derived

$$\begin{aligned} \mathbf{E}\{d(k,s)d^T(i,j)\} = \mathbf{E}\{d(k,s)x^T(i,j)\} \\ -\mathbf{E}\{d(k,s)\mathbf{E}^T\{x(i,j)|x[\mathcal{I}(i,j)]\} \end{aligned} \tag{8}$$

which left hand side of vanishes for $(i,j) \notin \mathcal{I}(k,s) \cup \{(k,s)\}$ by the only just proven orthogonality property (6). Thus (4) follows, as well as identity (5) by setting $(i,j) = (k,s)$. As to the case $(i,j) \in \mathcal{I}(k,s)$, notice that even now $\mathbf{E}\{d(k,s)x^T(i,j)\} = 0$, (by (6)), then by using (2) eq. (8) rewrites

$$\begin{aligned} \mathbf{E}\{d(k,s)d^T(i,j)\} &= -\sum_{(\alpha,\beta)\in\mathcal{I}'} F_{\alpha,\beta}(k,s)\mathbf{E}\{x(k+\alpha,s+\beta)d^T(i,j)\} \\ &= -\sum_{(\alpha,\beta)\in\mathcal{I}'} \mathbf{E}\{d(k,s)x^T(i+\alpha,j+\beta)\}F^T_{\alpha,\beta}(i,j). \end{aligned}$$

Now, by the orthogonality property, for any $(i,j) \in \mathcal{I}(k,s)$ seven terms are zero in each of the above summations, which results in eq. (7). ∎

From eq. (3), one readily derives ($\delta(\cdot,\cdot)$: discrete Dirac function in $\mathbb{N}^2$) :

$$\begin{aligned} \mathbf{E}\{x(k,s)\, x^T(l,r)\} &= D(k,s)\delta(k-l,s-r) \\ &+ \sum_{(\alpha,\beta)\in\mathcal{I}'} F_{\alpha,\beta}(k,s)\mathbf{E}\{x(k+\alpha,s+\beta)x^T(l,r)\}. \end{aligned} \tag{9}$$

Let us fix an (arbitrary) ordering in the set of eight elements $\mathcal{I}'$, say $(\alpha_1,\beta_1)...(\alpha_8,\beta_8)$ and define the block matrices:

$$\begin{aligned} F(k,s) &= [F_{\alpha_1,\beta_1}(k,s)\cdots F_{\alpha_8,\beta_8}(k,s)];\\ R(k,s) &= [R_1(k,s)\cdots R_8(k,s)];\\ P(k,s) &= \begin{bmatrix} P_{1,1}(k,s) & \dots & P_{1,8}(k,s)\\ & \ddots & \\ P_{8,1}(k,s) & \dots & P_{8,8}(k,s)\end{bmatrix}; \end{aligned} \tag{10}$$

where, for $i,j=1,...,8$:

$$\begin{aligned} R_i(k,s) &= \mathbf{E}\{x(k,s)x^T(k+\alpha_i,s+\beta_i)\},\\ P_{i,j}(k,s) &= \mathbf{E}\{x(k+\alpha_i,s+\beta_i)x^T(k+\alpha_j,s+\beta_j)\}. \end{aligned} \tag{11}$$

The projection matrices are univocally defined, provided a *local non-singularity* property holds.

Theorem 2.2. *Suppose the Markov field $x[\mathbb{N}^2]$ is locally non-singular at site (k,s), that is: $P(k,s)>0$*[1]*. Then the matrix coefficients in the representation (3) can be uniquely calculated at (k,s) as the solution of the linear matrix equation*

$$F(k,s)P(k,s)=R(k,s). \tag{12}$$

Moreover, the matrix $D(k,s)$ is non-singular and is given by

$$D(k,s)=\mathbf{E}\{x(k,s)x^T(k,s)\}-F(k,s)R^T(k,s). \tag{13}$$

Proof. Assembling, by using the block matrices (10), the eight equations got from (9) for $(l,r)\in\mathcal{I}(k,s)$, directly entails eq. (12), which has an unique solution by assumption. Equation (13) is readily derived from (9) by setting $(l,r)=(k,s)$. Non-singularity of $D(k,s)$ follows by the local non-singularity assumption by a standard argument that can be found in [4] (p. 1015). ∎

The following *normalized* form of representation (3) is to be used. Let $\mathcal{I}''=\mathcal{I}'\backslash\{(-1,0),(1,0)\}$, and define:

$$\begin{aligned} M_0(k,s)=D^{-\frac{1}{2}}(k,s);\quad M_-(k,s)&=D^{-\frac{1}{2}}(k,s)F_{-1,0}(k,s);\\ M_+(k,s)&=D^{-\frac{1}{2}}(k,s)F_{1,0}(k,s); \end{aligned} \tag{14}$$

$$S_{\alpha,\beta}(k,s)=D^{-\frac{1}{2}}(k,s)F_{\alpha,\beta}(k,s);\quad \text{for}\ \ (\alpha,\beta)\in\mathcal{I}'', \tag{15}$$

where $D^{\frac{1}{2}}(\cdot,\cdot)$ is the matrix square root of the symmetric matrix $D(\cdot,\cdot)$, which inverse of is well defined by Theorem 2. Thus, the representation equation (3) can be rewritten

$$\begin{aligned} M_0(k,s)x(k,s)=&M_-(k,s)x(k-1,s)+M_+(k,s)x(k+1,s)\\ &+\textstyle\sum_{(\alpha,\beta)\in\mathcal{I}''}S_{\alpha,\beta}(k,s)x(k+\alpha,s+\beta)+e(k,s), \end{aligned} \tag{16}$$

with $e(\cdot,\cdot)=D^{-\frac{1}{2}}(\cdot,\cdot)d(\cdot,\cdot)$ unitary-covariance Gaussian locally correlated 'noise' field.

[1]Notice that the non-singularity of $P(k,s)$ does not depend of the particular ordering choosen for the 8-tuple of the nearest-neighbour pixels.

3. Well-posedness for Spheric or Rectangular Domains

In this section I focus my attention on *finite* domains, which cases of are by far the more usual in practice for a Markov field to occur in conjunction with, and look over conditions for the Markov field being determined in such domains by eq. (16). Two cases of finite domain will be considered: the *rectangle* and *the sphere*, both relevant from an application point of view. As to the *rectangular case*, in specific, it will be shown that the restriction of the Markov field $\{x(k,s),(k,s)\in\mathbb{N}^2)\}$ to a finite *rectangle*, namely $\mathcal{R}=(0,N_r)\times(0,N_c)\subset\mathbb{N}^2$, is uniquely determined, samplewise, by the 'forcing noise' of the local representation (16), $e(k,s)$ for $(k,s)\in\mathcal{R}$, and by the Markov field values at the *boundary*, i.e. $\{x(k,s),(k,s)\in\partial\mathcal{R}\}$, where $\partial\mathcal{R}=\overline{\mathcal{R}}\backslash\mathcal{R}$, and $\overline{\mathcal{R}}=[0,N_r]\times[0,N_c]$. As to the *spheric case*, the finite domain is the *discrete sphere* (hereinafter simply *the sphere*), namely $\mathcal{S}$, which, as a set, is equal to $\overline{\mathcal{R}}$, but algebrically is actually different in that, performing the sum of two elements, say $(k,s),(l,r)\in\mathcal{S}$, the result $(k+l,s+r)$ has to be intended modulo N_r and modulo N_c for the first and second component respectively. In other words, we should imagine the sites $(k,r)\in\mathcal{S}$ being likewise placed upon a sphere's surface as the intersections of parallels and meridians on the earth surface. It will be shown that the 'forcing noise' $\{e(i,j),(i,j)\in\mathcal{S}\}$ is sufficient (i.e. partial knowledge is not required for $x(i,j)$ at any point) for determining samplewise $x(k,s)$ at each pixel (k,s) of the sphere. These Markov field's properties are important as they will be used in the next section in the proof of the main result. Both cases, spheric and rectangular, will be treated at the same time by using a specific notation, as follows. The symbol $\mathcal{T}$ ($\overline{\mathcal{T}}$) will denote any of the sets $\mathcal{R},\mathcal{S}$ ($\overline{\mathcal{R}},\mathcal{S}$), defined above, and hereinafter, consistently with convention (1), the Markov-field at issue will be denoted by x (as well as the other fields involved, for instance: e). We say that the representation (16) is well posed on $\overline{\mathcal{T}}$ (i.e. on both rectangular and spheric domains) if there exists the (samplewise) map $\underline{e}\rightarrow x$, where $\underline{e}$ is defined as

$$\underline{e}=\begin{cases} e=\mathrm{col}(e(0),...,e(N_c)), & (\mathcal{S}) \\ \mathrm{col}(x(0),\overline{e}(1),...,\overline{e}(N_r-1),x(N_c)). & (\mathcal{R}) \end{cases} \tag{17}$$

and, defining $\bar{x}(j,s)=\sum_{i=\pm1,0}(-1)^i x(j,s+i)$:

$$\overline{e}(s)=\begin{cases} e(s)=\mathrm{col}(e(0,s),...,e(N_r,s)), & (\mathcal{S}) \\ \mathrm{col}(\bar{x}(0,s),e(1,s),...,e(N_r-1,s),\bar{x}(N_r,s)), & (\mathcal{R}) \end{cases}$$

Thus, [2] note that, as to the field $\underline{e}$, denoting $\delta(\cdot)$ the Dirac function in $\mathbb{N}$, it is

$$\underline{e}(k,s)=e(k,s),\quad (k,s)\in\mathcal{T} \tag{18}$$

$$\underline{e}(k,s)=x(k,s)-\sum_{i=\pm1}x(k,s+i)(1-\delta(s))(1-\delta(s-N_c)),\qquad (k,s)\in\partial(\overline{\mathcal{R}}) \tag{19}$$

Before giving the result, some new facts needs to be introduced.

Given three finite sequences, A,B,C, of $\nu+1$ matrices in $\mathbb{R}^{n\times n}$ each: $\{L(0),...,L(\nu)\}$, $L=$

[2] It might seem redundant defining $\overline{e}(s)$ for $\mathcal{T}=\mathcal{S}$, as in (17) it is used just for $\mathcal{T}=\mathcal{R}$. Nonetheless, the notation $\overline{e}(s)$ is to be used in eq. (22) for $\mathcal{T}=\mathcal{S}$.

A, B, C, we define the following $(\nu+1)\times(\nu+1)$-blocks *circulant* matrix:

$$\mathcal{F}_\nu^{\mathcal{S}}(A,B,C)=\begin{bmatrix} B(0) & -C(0) & 0 & \dots & -A(0)\\ -A(1) & B(1) & -C(1) & 0 & \dots\\ 0 & -A(2) & B(2) & -C(2) & \dots\\ & & \ddots & & \\ -C(\nu) & 0 & \dots & -A(\nu) & B(\nu)\end{bmatrix} \tag{20}$$

and, the $(\nu+1)\times(\nu+1)$-blocks matrix

$$\mathcal{F}_\nu^{\mathcal{R}}(A,B,C)=\begin{bmatrix} I & 0 & & \dots & 0\\ -A(1) & B(1) & -C(1) & 0 & \dots\\ 0 & -A(2) & B(2) & -C(2) & \dots\\ & & \ddots & & \\ 0 & & \dots & 0 & I\end{bmatrix}. \tag{21}$$

Equation (16) can be assembled for $k=0,...,N_r$ on the sphere as follows (recall that, on the unit sphere, it is $x(-1,s)=x(N_r,s)$, and $x(N_r+1,s)=x(0,s)$), or for $k=1,...,N_r-1$ in the *interior* of the rectangle, i.e. $\mathcal{R}$: it turns out the following compact equation including both $\mathcal{S}$ and $\mathcal{R}$ cases:

$$M(s)x(s)=S^-(s)x(s-1)+S^+(s)x(s+1)+\overline{e}(s), \tag{22}$$
$$M(s)=\mathcal{F}_{N_r}^{\mathcal{T}}(M_-(\cdot,s),M_0(\cdot,s),M_+(\cdot,s)), \tag{23}$$
$$S^\pm(s)=\mathcal{F}_{N_r}^{\mathcal{T}}(-S_{-1,\pm1}(\cdot,s),S_{0,\pm1}(\cdot,s),-S_{1,\pm1}(\cdot,s)). \tag{24}$$

Equation (22) can in turn be assembled either for $s=0,...,N_c$ (when $\mathcal{T}=\mathcal{S}$, as before, accounting of the modulo-N_c arithmetics on the sphere), or for $s=1,...,N_c-1$ (when $\mathcal{T}=\mathcal{R}$):

$$\mathbf{\Psi}x=\underline{e}, \qquad \mathbf{\Psi}=\mathcal{F}_{N_c}^{\mathcal{T}}(S^-,M,S^+). \tag{25}$$

Also, the following technical assumption will be done.

Non-singularity assumption: The Markov field x is said to be non-singular if $\mathbf{E}\{xx^T\}>0$.

Notice that, if x is non-singular, then it is locally non-singular, as the covariance matrix $P(k,s)$ defined in (10), collecting all mutual covariances between pixels in the nearest-neighbour, is non-singular as well: $P(k,s)>0$, $\forall(k,s)\in\mathcal{T}$. [3]

Theorem 3.1. *Suppose that the Markov field x is non-singular. Then representation (3) is well-posed on $\overline{\mathcal{T}}$. In particular, there exist kernel functions:* $\Gamma(\cdot,\cdot;\cdot,\cdot):\overline{\mathcal{T}}\times\overline{\mathcal{T}}\times\overline{\mathcal{T}}\times\overline{\mathcal{T}}\times\to\mathbb{R}^{n\times n}$, *and* $\Gamma(\cdot;\cdot):\overline{\mathcal{T}}\times\overline{\mathcal{T}}\to\mathbb{R}^{n(N_r+1)\times n(N_r+1)}$, *and a square $n^2(N_r+1)(N_c+1)$-dimensioned matrix Γ, such that*

$$x(k,s)=\sum_{i=0}^{N_r}\sum_{j=0}^{N_c}\Gamma(k,s;i,j)\underline{e}(i,j), \tag{26}$$

$$x(s)=\sum_{i=0}^{N_c}\Gamma(s;i)\underline{e}(i), \tag{27}$$

$$x=\Gamma\underline{e}. \qquad \Gamma=\mathbf{\Psi}^{-1}=\mathbf{E}\{xx^T\}(\mathbf{E}\{\underline{e}x^T\})^{-1} \tag{28}$$

[3] Indeed it is easily recognized that $P(k,s)$ agrees with some diagonal sub-matrix of $\mathbf{E}\{xx^T\}$.

Moreover, $\forall(k,s) \in \mathcal{S}$ *(for spheric domain), it is*

$$
\begin{array}{ll}
M_0(k,s) = M_0^T(k,s), & M_+(k,s) = M_-^T(k+1,s); \\
S_{1,1}(k,s) = S_{-1,-1}^T(k+1,s+1); & S_{0,1}(k,s) = S_{0,-1}^T(k,s+1);
\end{array}
\tag{29}
$$

which hold for a rectangular domain as well, as far as the involved matrices are calculated in $\mathcal{R}$ *(e.g. inside the rectangle).*

Proof. By eq. (25) one has $\mathbf{\Psi}\mathbf{E}\{xx^T\} = \mathbf{E}\{\underline{e}x^T\}$. On account of the orthogonality property (6), it is $\mathbf{E}\{e(s)x^T(l)\} = I_{n(N_r+1)\times n(N_r+1)}\delta(s-l)$, thus, for $\mathcal{T} = \mathcal{S}$, $\mathbf{E}\{\underline{e}x^T\} = I$, where I is the $n^2(N_r+1)(N_c+1)$-dimensioned identity, and, hence, $\mathbf{\Psi} = \mathbf{\Psi}^T > 0$. Since $\mathbf{\Psi} = \mathcal{F}_{N_c}^{\mathcal{S}}(S^-, M, S^+) > 0$ it follows $\mathcal{F}_{N_c}^{\mathcal{R}}(S^-, M, S^+)$ is non-singular as well. Indeed, $\mathcal{F}_{N_c}^{\mathcal{R}}(S^-, M, S^+)$ has the same structure of $\mathcal{F}_{N_c}^{\mathcal{S}}(S^-, M, S^+)$, and can be transformed into the latter by suitably substituting some block-rows (those indeed with I and 0 blocks). One comes such way to the matrix $\mathbf{\Psi}$ of a field x defined on $\mathcal{S}$, but equal to the former in every $(k,s) \in \mathcal{R}$. So, since $\mathcal{F}_{N_c}^{\mathcal{S}}(S^-, M, S^+)$, and $\mathcal{F}_{N_c}^{\mathcal{R}}(S^-, M, S^+)$ have the same rank (one gets the latter by replacing linear independent rows in the former) it follows that it is an invertible matrix as well. Thus, eq. (25) admits an unique solution given by (28) in both cases $\mathcal{S}$ and $\mathcal{R}$. The kernel functions in (26) and (27) are accordingly identified as suitable block sub-matrices of Γ. Moreover, for $\mathcal{T} = \mathcal{S}$ as shown only just before it is $\mathbf{\Psi} = \mathbf{\Psi}^T$, thus $S^+(s) = S^{-T}(s+1)$, and $M(s) = M^T(s)$, which imply (29). For rectangular domains, as we have seen we can extend the field defined on $\mathcal{R}$ up to the sphere, equalities (29) hold as well, provided they involve projection matrices in $\mathcal{R}$. ∎

Notice that, from (28), for spheric domain the kernel functions are simply given by the covariances $\Gamma(k,s;i,j) = \mathbf{E}\{x(k,s)x^T(i,j)\}$, and $\Gamma(s;i) = \mathbf{E}\{x(s)x^T(i)\}$.

Remark: As to the *non-singularity assumption* (as well as its *local* version considered in §II) under which Theorem 3 (and Theorem 2) have been proven, such a technical assumption is somewhat classical (indeed we also find it in [4]) and allows a lot of simplifications while treating the subject. Nonetheless, it is not a so very crucial assumption in that even if got no more verified, much of the basic results here presented would keep to be true as well. For instance, eq. (3) keeps to hold, although eq. (12) has no more an unique solution. Also, the normalized form (16) no more can be written (at least, with a nonsingular M_0), but the unnormalized equation (3) could still be used, though actually *multiple* representations could in principle occur, as many as the (multiple) solutions of eq. (12) are.

4. 2D-Recursive Representation

The aim is now to show that representation (16) can be 'solved' (in the stochastic sense, i.e. samplewise), in a similar way as a 1-D stochastic dynamic system does. To this purpose we need the following preliminary result, which proof of can be found in ([8], Lemma 5.3), or (in a somewhat different form) in [4]: if A, B, C are matrix sequences such that $C(k) = A^T(k+1)$, $B(k) = B^T(k) > 0$, for $k = 0 \ldots \bar{\nu}$, $\bar{\nu} = \nu - 1$, then the following $(\nu-1)\times(\nu+1)$-blocks three-banded matrix

$$
\Phi_\nu(A,B,C) = \begin{bmatrix} -A(1) & B(1) & -C(1) & 0 & \ldots \\ 0 & -A(2) & B(2) & -C(2) & \ldots \\ & & \ddots & & \\ & & -A(\bar{\nu}) & B(\bar{\nu}) & -C(\bar{\nu}) \end{bmatrix}
\tag{30}
$$

admits the following decomposition[4]:

$$\Phi_\nu(A,B,C) = L(C)\mathbf{T}^{-1}H(C), \quad \mathbf{T} = \text{diag}\{T(i)\}_{i=0,\dots,\bar{\nu}} \tag{31}$$

$$L(C) = \begin{bmatrix} I & -(CT)(1) & 0 & \dots & \\ 0 & I & -(CT)(2) & & \\ & & \ddots & & \\ 0 & \dots & & I & -(CT)(\bar{\nu}) \end{bmatrix}$$
$$H(C) = \begin{bmatrix} -(TC^T)(0) & I & 0 & \dots \\ 0 & -(TC^T)(1) & I & \\ & & \ddots & \\ 0 & \dots & -(TC^T)(\bar{\nu}) & I \end{bmatrix} \tag{32}$$

where for short $(CT)(i)$ stands for $C(i)T(i)$ (and so TC^T as well), and the matrix $T(i) > 0$ for $i = 0, ..., \bar{\nu}$, satisfies the backward recursive equation:

$$T^{-1}(i-1) = B(i) - C(i)T(i)C^T(i), \quad T(\nu) = 0. \tag{33}$$

Let us turn back to the field equation in compact form $\mathbf{\Psi}x = \underline{e}$ (eq. (25)), which implies, on account of convention (1) $\mathbf{\Psi'}x = \underline{e}'$, where $\mathbf{\Psi'} = \Phi_{N_c}(S^-, M, S^+)$ (notice $\mathbf{\Psi'}$ is the same in both cases $\mathcal{S}$, $\overline{\mathcal{R}}$). By applying decomposition (31) to $\mathbf{\Psi'}$ results in $L(S^+)\mathbf{Q}^{-1}H(S^+)x = \underline{e}'$, which is in turn equivalent to the following pair of equations, where a new auxiliary variable $\gamma = \text{col}(\gamma(0), ..., \gamma(N_c))$ is introduced, and convention (1) applied:

$$L(S^+)\gamma^- = \underline{e}', \qquad \gamma^- = \mathbf{Q}^{-1}H(S^+)x. \tag{34}$$

Note that $\mathbf{Q} = \text{diag}\{Q(s)\}_{s=0,\dots,N_c-1}$, $Q(s) > 0$ for $s = 0, ..., N_c-1$ satisfying

$$Q^{-1}(s-1) = M(s) - S^+(s)Q(s){S^+}^T(s), \quad Q(N_c) = 0. \tag{35}$$

On account of $L(S^+)$'s, and $H(S^+)$'s structures, from eqs. (34) we get the following *backward* and *forward* equations for the processes γ and x respectively[5]:

$$\gamma(s) = S^+(s+1)Q(s+1)\gamma(s+1) + \overline{e}(s+1), \tag{36}$$
$$x(s+1) = {S^+}^T(s)x(s) + \gamma(s). \tag{37}$$

In the following Theorem 4, it will be assumed that x is an *homogeneous* field, i.e. a *stationary* process in both the orthogonal directions k, s, as well as along diagonals. Thus all projection matrices are constant, as well as the correlation function, and hereinafter the symbol without arguments is used to denote them. The correlation gets fully described by three matrices, namely D, Δ_d, Δ_h: $D \equiv D(k,s)$, $\Delta_\mu \equiv \mathbf{E}\{d(k,s)d^T(k+\alpha, s+\beta)\}$, $\forall(\alpha,\beta) \in \mathcal{I}'_\mu$, where $\mu = d, h$ (merely, a lexicographic substitution), and

$$\mathcal{I}'_d = \{(-1,1), (1,1), (1,-1), (-1,-1)\},$$
$$\mathcal{I}'_h = \mathcal{I}' \setminus \mathcal{I}'_d.$$

[4]It can be verified by direct calculation.
[5]By (17) it is $\underline{e}'(s) = \overline{e}(s)$.

The matrices Δ_d and Δ_h are the cross-covariances between (any) (k, s) and the nearest-neighbour sites in the *diagonal* and *horizontal/vertical* directions, respectively. Also, by eqs. (7) it is: $\Delta_\mu = -F_{\alpha,\beta}D = -DF^T_{\alpha,\beta}$, $\forall(\alpha,\beta) \in \mathcal{I}'_\mu$ $(\mu = d, h)$ Thus,

$$
\begin{aligned}
&\Delta_d = \Delta_d^T, \quad \Delta_h = \Delta_h^T,\\
&F_{-1,0}=F_{1,0}=F_{0,-1}=F_{0,1}=-\Delta_h D^{-1},\\
&F_{-1,1}=F_{1,1}=F_{1,-1}=F_{-1,-1}=-\Delta_d D^{-1},
\end{aligned}
$$

and on account of eqs. (14), (15), we can define the matrices σ_d, σ_h:

$$
\begin{aligned}
&\sigma_h = \sigma_h^T = S_{0,\beta} = M_- = M_+, \quad \forall\beta = \pm 1,\\
&\sigma_d = \sigma_d^T = S_{\alpha,\beta}, \quad \forall(\alpha,\beta) \in \mathcal{I}'_d.
\end{aligned}
$$

The matrix parameters in eq. (22) are constant, i.e $M(s) \equiv M$, $S^+(s) \equiv S^+$, $S^-(s) \equiv S^-$, and in particular, by definition of $M, S^\pm$ given in (23), (24) one has

$$
\begin{aligned}
&S^+ = S^{+T} = S^- = S = \mathcal{F}^{\mathcal{T}}_{N_r}(-\sigma_d, \sigma_h, -\sigma_d),\\
&M(s) = M = M^T = \mathcal{F}^{\mathcal{T}}_{N_r}(\sigma_h, M_0, \sigma_h).
\end{aligned}
$$

Theorem 4.1. *Let the Markov field x be homogeneous. Then, for any $(k, s) \in \mathcal{T}$, x satisfies (samplewise) the following recursive scheme:*

$$
\begin{aligned}
x(k,s) = x^0(k,s) &+ \sum_{i=0}^{N_r} \Gamma(k,s;i,0)(\underline{e}(i,0) - x^0(i,0))\\
&+\sum_{i=0}^{N_r} \Gamma(k,s;i,N_c)\underline{e}(i,N_c) + \sum_{j=1}^{N_c-1} \Gamma(k,s;0,j)\underline{e}(0,j)\\
&+\sum_{j=1}^{N_c-1} \Gamma(k,s;N_r,j)\Big(\underline{e}(N_r,j) - \sum_{l=\pm1,0} (-1)^l x^0(N_r,j+l)\Big)
\end{aligned} \tag{38}
$$

where Γ is the proper kernel function (defined by Theorem 3) for the case at issue ($\mathcal{S}$ or $\mathcal{R}$), and x^0 is the solution of the following 2D-recursive equations:

$$
\begin{aligned}
\gamma_b(k,s) = -\sigma_d\mathbf{q}(k+1)\gamma_b(k+1,s) + \gamma(k+1,s)&\\
-e(k+1,s+1)&
\end{aligned} \tag{39}
$$

$$
\tilde{\gamma}(k+1,s+1) = \mathbf{q}(k)(\gamma_b(k,s) - \sigma_d\tilde{\gamma}(k,s+1)), \tag{40}
$$

$$
\gamma(s) = \mathbf{P}(s)\tilde{\gamma}(s), \quad \tilde{\gamma}(s) = \mathrm{col}(\tilde{\gamma}(0,s), ..., \tilde{\gamma}(N_r,s)), \tag{41}
$$

$$
\tilde{\gamma}(k,0) = 0, \quad k = 0,...,N_r, \tag{42}
$$

$$
\tilde{\gamma}(0,s) = 0, \quad \gamma_b(N_r-1,s) = 0, \quad s = 0,...,N_c, \tag{43}
$$

$$
\gamma_f(k,s) = -\sigma_d\mathbf{q}(k+1)\gamma_f(k+1,s) + x^0(k,s+1) - \gamma(k,s) \tag{44}
$$

$$
x^0(k+1,s) = \mathbf{q}(k)(\gamma_f(k,s) - \sigma_d x^0(k,s)), \tag{45}
$$

$$
x^0(k,N_c) = 0, \quad k = 0,...,N_r, \tag{46}
$$

$$
\gamma_f(N_r-1,s) = 0, x^0(0,s) = 0, \quad s = 0,...,N_c. \tag{47}
$$

The matrices $\mathbf{P}(s) = \mathbf{P}^T(s) > 0$ are such that $\mathbf{P}^{-1}(s) = Q(s)$ and the $Q(s)$'s satisfy, for $s = 0,...,N_c-1$, the backward recursive equation $Q^{-1}(s-1) = M - SQ(s)S$, $Q(N_c) = 0$, whereas $\mathbf{q} = \mathrm{diag}\{\mathbf{q}(0),...,\mathbf{q}(N_r-1)\} > 0$, the diagonal blocks satisfying the backward equation $\mathbf{q}^{-1}(k-1) = \sigma_h - \sigma_d\mathbf{q}(k)\sigma_d$, with $\mathbf{q}(N_c) = 0$.

Proof. Let us consider the homogeneous version of eqs (36), (37), which, noting that it is $\overline{e}'(s) = e'(s)$, can be rewritten as follows:

$$\gamma'(s) = S'\tilde{\gamma}(s+1) + e'(s+1), \tag{48}$$
$$x'(s+1) = S'x(s) + \gamma'(s). \tag{49}$$

where $\tilde{\gamma}(s) = Q(s)\gamma(s)$, $S' = \Phi_{N_r}(-\sigma_d, \sigma_h, -\sigma_d)$, and convention (1) has been used. Also, note that, by the same convention (in a reversed sense), the $2D$-process $\gamma(\cdot,\cdot)$ is well defined. Notice that $Q(s)$ actually depends of s even though the matrix coefficients M, S^+, S^- do not. The triple of constant sequences $-\sigma_d, \sigma_h, -\sigma_d$, as such, do satisfy the condition for decomposition (31), so we can apply it to S', with $\nu = N_r$, and rewrite eqs. (48), (49) as

$$L(\sigma_d)\mathbf{q}^{-1}H(\sigma_d)\tilde{\gamma}(s+1) = \gamma'(s) - e'(s+1), \tag{50}$$
$$L(\sigma_d)\mathbf{q}^{-1}H(\sigma_d)x(s) = x'(s+1) - \gamma'(s). \tag{51}$$

where $\mathbf{q}$ is accordingly defined. Equation (50) is equivalent to the following pair of equations, where a new auxiliary variable $\gamma_b(s) = \mathrm{col}(\gamma_b(0,s), ..., \gamma_b(N_r, s))$ is introduced: $L(\sigma_d)\gamma_b^-(s) = \gamma'(s) - e'(s+1)$, $\gamma_b^-(s) = \mathbf{q}^{-1}H(\sigma_d)\tilde{\gamma}(s+1)$, which the backward and forward equations (39), (40) are got from, on account of $L(\sigma_d)$'s, and $H(\sigma_d)$'s structures. Similarly, with $\gamma_f(s) = \mathrm{col}(\gamma_f(0,s), ..., \gamma_f(N_r, s))$, eq. (51) is equivalent to $L(\sigma_d)\gamma_f^-(s) = x'(s+1) - \gamma'(s)$, and $\gamma_f^-(s) = \mathbf{q}^{-1}H(\sigma_d)x(s)$, which result in the backward/forward equations (44), (45). Note that such 2D-recursive equations holds only in $(0, N_r) \times (0, N_c)$, and the field x has to be *one* of its solutions. In order to uniquely determine x, let us consider the *particular* solution, namely x^0, obtained by setting to the null function all the inizializing functions, i.e. the boundary conditions (42),(43), and (46), (47). Such x^0 is a particular case of the $\mathcal{R}$ case, with boundary conditions at the border, and as such it has to satisfy (26) as well, but x^0 replaces x in (19) (i.e. at the border). Thus, using eq. (26) once for the 'true' x, then for x^0, and subtracting, results in eq. (38) ∎

The $2D$-recursive scheme (39)-(47) realizes x^0, the *particular solution*, from which the field x is next recovered, in both cases $\mathcal{S}$ or $\mathcal{R}$, by using eq. (38). The $2D$-recursive scheme is composed by two macro-steps:

1) Forward 'brushing' of columns, ($s = 0, 1, ...N_c$), and $\gamma(\cdot,\cdot)$ computation. $\gamma_b(\cdot, s)$ is calculated first, by eq. (39) backward-iterated from $\gamma_b(N_r - 1, 0)$ – which is given in (43). Then, $\tilde{\gamma}(\cdot, s)$ (hence γ by (41)), is calculated by eq. (40) forward-iterated from $\tilde{\gamma}(0, s) = 0$ – i.e. the first identity in (43) – and using the previously calculated values $\gamma_b(k, s)$.

2) Backward 'brushing' of columns, ($s = N_c - 1, ..., 0$), and $x^0(\cdot,\cdot)$ computation. Using γ, previously calculated, and backward-iterating eq. (44) from final conditions given in (47), one gets $\gamma_f(\cdot, s)$. Next, all values of $x^0(k, s)$, stem from forward-iterating eq. (45) starting with $x^0(0, s) = 0$.

Thus, for a spheric domain $\mathcal{S}$, the map $e \rightarrow x$ gets realized recursively. For rectangular domain $\overline{\mathcal{R}}$, e is defined only in the interior $\mathcal{R}$, and what is realized is the map $(\overline{x}, e) \rightarrow x$, where $\overline{x}(i,j)$, $(i,j) \in \partial\mathcal{R}$, are the field's values along the domain's border.

Finally, it should be stressed that the advantage of the here proposed model with respect to [11] is mainly computational, rather than conceptual. Both approaches make use of representation (22), but while there this is the *final* representation, here this is further worked out so that eqs. (39)-(47) are got from it. Of course, a dynamic system can be associated as well to (22) in the form of eqs. (36), (37), but this leads to assuming the 'columns' $\gamma(s)$, and $x(s)$ as the 'state' of the system, which in turn implies that each system's iteration involves multiplying twice for a matrix. In order to perform a comparison with the $2D$-recursive scheme, let's view such double multiplication as a double 'brushing' of the column: then the total number of scannings is $2\cdot(N_r+1)$ times (twice the rows number) for each column. On the contrary, in the $2D$-recursive scheme one has to scan each column just four times, no matter of which the rows number is.

5. Conclusion

The main result of the paper is Theorem 4 where, for homogeneous Markov fields, the set of 2D recursive equations (39)–(45) has been shown to be equivalent to the original nearest-neighbour representation, realizing such way the idea of stochastic 'dynamic' system in two dimensions. The 2D-recursive scheme developed in Theorem 4, allows to scan all the domain by passing through each column four times: twice, while iterating eqs. (39), (40), and twice again for eqs. (44), (45). It should be stressed that – while thinking of Image-Processing with x a scalar gray-level – it allows of course a speedy scanning of the whole image. The 2D-recursive scheme here presented is the first, necessary, step towards the definition of recursive algorithms to be applied to a Gaussian Markov-field in any specific application. As an example, we mention the problem of finding the optimal smoothing estimate of field's values observed under Gaussian noise. In future works it will be shown that such smoothing algorithms behave similarly to the model here described, in the sense they are 2D-recursive and preserve the nice scanning properties, mentioned above, of such model. Also, an important topic is the possible definition of a 2D-recursive scheme for a $2D$-Markov field taking values in a finite set of symbols. As argued in §II, some weakening of the nonsingularity assumption is possible, in that such assumption does affect essentially the unicity, and not the existence, of the representation at issue. Indeed, studying a singular version of representation (16) – i.e. with a non-invertible M_0 – as well as the case of a non homogeneous field, are among further developments of the present work.

References

[1] B. Jamison, Reciprocal processes, *Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 30, 1974, pp. 65-86.

[2] A.J. Krener, Reciprocal processes and the stochastic realization problem for acausal systems. In: *Modelling, Identification and Control*, C.I. Byrnes and A. Lindquist Eds. North-Holland, Amsterdam, 1986, pp. 197-211.

[3] A.J. Krener, Reciprocal diffusions and stochastic differential equation of second order, *Stochastics*, vol. 24, 1988, pp. 393-422.

[4] B. C. Levy, R. Frezza, and A.J. Krener, Modelling and estimation of discrete-time Gaussian reciprocal processes, *IEEE Trans. on Aut. Cont.*, vol. 35, 1990, pp. 1013-1022.

[5] B. C. Levy, R. Frezza, and A. J. Krener, Gaussian reciprocal processes and self-adjoint stochastic differential equations of second order, *Stochastics and stochastic reports.* vol. 34, 1991, pp. 29-56.

[6] B. Levy, and A.J. Krener, Dynamics and kinematics of reciprocal diffusions, *J. Math. Phys.* vol. 34, 1993, pp. 1846-1875.

[7] A.J. Krener, Reciprocal diffusions in flat space, *Probab. Theory and Related Fields,* vol. 107, 1997, 243-281.

[8] F. Carravetta, Nearest-Neighbour Modelling of Reciprocal Chains, *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 80, 2008, pp. 525-584.

[9] P. Levy, *A special problem of Brownian motion, and a general theory of Gaussian random functions*, *Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability*, vol. 2, 1956, pp. 133-175.

[10] B. C. Levy, 1990, Noncausal estimation for discrete Gauss-Markov random fields. *Proc. Int. Symp. Math. Theory Networks Syst. (MTNS-89)*, Birkhauser, Boston, 1990.

[11] B.C. Levy, M.B. Adams, and A.S. Willsky, Solution and linear estimation of 2-D nearest-neighbor models, *Proc. IEEE,* vol. 78, 1990, pp. 627-641.

[12] R. Kindermann, and J.L. Snell, Markov Random Fields and Their Applications, Contemporary Mathematics, vol. 1. Providence, R.I. Am. Math. Soc., 1980

[13] H. Derin, and H. Elliott, Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields, *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 9, 1987, pp. 39-55.

[14] X. Guyon, *Random Fields on a Network: Modeling, Statistics, and Applications*, Springer Verlag, NY; 1991.

[15] D.A. Castanon, B.C. Levy, and A.S. Willsky, Algorithms for the incorporation of predictive information in surveillance theory. *Int. J. Systems Sci.*, vol. 16, 1985, pp. 367-382.

[16] A.J. Krener, Boundary value linear systems, *Asterisque*, vol. 75-76, 1980, pp. 149-165.

[17] A.J.Krener, Acausal realization theory, Part 1: Linear deterministic systems, *SIAM J. Control and Opt.*, vol 25, 1987, pp. 499-525.