Scan to know paper details and
author's profile

On Some Spatial Considerations of the Tabulated (Categorical) Stationary Series (Natural Modelling; Probability Restrictions; Markovian Dependence)

Chrysoula Dimitriou-Fakalou

ABSTRACT

A spatial model for the strictly stationary series that have been discretized into a specified number of categories, is presented: special emphasis is concentrated on the finite two-sided Markovian structure. The new suggestion puts forward an all-random model, relying on a collection of unobserved series, with variables that are defined on different sample spaces. Imitating the linear ARMA series (that employ the spectral densities though), symmetry restrictions (via Bernoulli variables) and time reversal are explored and succeed to a certain extent. Subject to an, applicable to any distribution, assignment of variables' values into (k +1) ranges, and to the selection of the serial order p and q, the general Table Auto-Linear Moving-Average (k, p, q) equation provides for the spatial, all-moments stationary as well as infinite homogeneous Markovian, dependence.

Keywords: auto-linear; bernoulli; categorization; latent equalizer predictor; strictly stationary.

Classification: LCC Code: QA1-QA939

Language: English

Great Britain Journals Press

LJP Copyright ID: 925681
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 22 | Issue 8 | Compilation 1.0

On Some Spatial Considerations of the Tabulated (Categorical) Stationary Series (Natural Modelling; Probability Restrictions; Markovian Dependence)

Chrysoula Dimitriou-Fakalou

ABSTRACT

A spatial model for the strictly stationary series that have been discretized into a specified number of categories, is presented: special emphasis is concentrated on the finite two-sided Markovian structure. The new suggestion puts forward an all-random model, relying on a collection of unobserved series, with variables that are defined on different sample spaces. Imitating the linear ARMA series (that employ the spectral densities though), symmetry restrictions (via Bernoulli variables) and time reversal are explored and succeed to a certain extent. Subject to an, applicable to any distribution, assignment of variables' values into (k +1) ranges, and to the selection of the serial order p and q, the general Table Auto-Linear Moving-Average (k, p, q) equation provides for the spatial, all-moments stationary as well as infinite homogeneous Markovian, dependence.

Keywords: auto-linear; bernoulli; categorization; latent equalizer predictor; strictly stationary.

Author: Chrysoula Dimitriou-Fakalou, self-employed scholar, 1st floor, 6 Kastelorizou Street, Ano Voula Attikis, Athens 16673, Greece. email: my.submission@mail.com

I. INTRODUCTION

Up to this day, the theory of the other than ARMA-modelled stationary time series has been enriched with some decent attempts: the direct aim of Lomnicki and Zaremba (1955) or Keenan (1982), was on the strictly stationary processes by reducing to categorical $0-1$ dependencies; the Discrete ARMA models of Jacobs and Lewis (1978) have placed the focus on general mixtures; the bilinear model with deterministic coefficients offers a net solution, by writing the variable as a sum of products of its lagged and error lagged values raised to natural powers, but suffers from a lack of interpretation and demands some immediate action regarding measurability; both the INARMA model (for instance, find a primary form in McKenzie (1985) and then Du and Li (1991)) and the suggestion of Cui and Lund (2009) work for count data, with the first one crossing the variables of interest with errors, and the second one only conveniently combining information from discrete time renewal processes.

London Journal of Research in Science: Natural and Formal

Evidently, blending the different equations in order to escape the linear series modelling, has been a common mentality: for example, one could catch up on switching between different ARMA equations in Francq and Zakoïan (2001). The threshold models, an excellent account of which resides in Tong (1990), may be considered as a world popular illustration of non-linearity that caused a lot of this. Finally, the latest suggestion of a TARMA model by Dimitriou-Fakalou (2019) is to fulfil the promise of encompassing the strictly stationary dependence, via a 'non-parametric' assignment of the variables' values into a fixed number of ranges: under causality, Markov chain conditional probability approaches become its special case. Since this paper will be translating the TARMA merits for the spatial series in the line transect, a reminder of existing transitions from time to space must follow.

For the linear series, two main streams developed in the past: the simultaneous bilateral auto-regressive models of Whittle (1954) and the conditional auto-normal processes of Besag (1974). The first exhibited what happens when the standard ARMA model uses its variables from both past and future: the high risks for the inference were exposed and the causal solution has been treasured as the right way to go, subject to the unilateral (more than one) index ordering of Whittle (1954). Besag (1974), on the other hand, worked with conditional arguments and proposed a Gaussian model that equally treats both ends of the axis; that second approach would often relate to a parameterization with intrinsic restrictions. Remaining in the 'best linear predictor' series department, those authentic definitions triggered Dimitriou-Fakalou (2010) to introduce the term 'auto-linear', combining the simultaneous (rather than conditional) element from Whittle (1954) together with the spatial (rather than unilateral) element from Besag (1974).

Trying to record some attempts for the spatial modelling outside the Gaussian distribution, Besag (1974) also introduced the auto-logistic schemes (for a fixated number of sites in space rather than a process in the transect) applied on Bernoulli variables that do, in general, exhibit a higher than second-order dependence: nevertheless, those definitions were accompanied by the assumption that the quantities relating more than two variables were set as null (leaving the analysis only to *cliques consisting of single sites and pairs of sites*) and shattered any hope of going further than the relevant pairwise elements. The reader is encouraged to look at Cressie (1993) for a beautiful description of the different perspectives in spatial statistics, which is rich in diverse material: nevertheless when working in space, the dilemma of either sticking to linearity for the infinite lattice processes, or resorting to conditional probability arguments and under the knowledge of a basic distribution for the finite lattice variables, remains.

Hence a discretized strictly stationary series indexed regularly in the straight line is to be modelled: a two-sided, random coefficient equation is the proposed solution satisfying the spatial needs. Due to the severe difficulties that arise when scrutinizing the spatial dependencies, the assumption of causality that frees the 'future' errors from any dependence with 'present' and 'past' variables will still be wrapped inside the new definitions. Section 3 introduces a special case of the model, which results in the spatial pth order conditional probabilities, and Section 3.1 explains the thinking that supports the new models: the arguments relate to writing the *latent equalizer predictor*, as this has been explained already in Section 2. Section 3.2 is looking for ways to get to the restrictions that tie Bernoulli variables in space, and Section 3.3 examines whether time reversal can be achieved when one goes higher than the Gaussianity of random variables and there is no moment generating function to ease the derivations; this is versus the spectral density of an ARMA(p, q) series that reveals how 2^{p+q} different ARMA equations share the same auto-covariance function (Brockwell and Davis (1991)). Finally, *Definition 1* in Section 4 introduces the Table Auto-Linear Moving-Average model and perfects the previous suggestions, by approximating the infinite spatial Markovian dependence.

II. REMINDERS: THE 'LATENT EQUALIZER PREDICTOR' APPROACH

Suppose that $\{X_t,\ t \in \mathbb{Z}\}$ is a strictly stationary time series which, either naturally or by discretization, takes the values 0, v_1,...v_k with probability one. For a fixed positive integer p, this could be modelled as

$$X_t := I_t^{(X_{t-1},\ldots,X_{t-p})}, \tag{1}$$

where $\{I_t^{(i_1,\ldots,i_p)},\ t \in \mathbb{Z}\}$, $i_1,\ldots,i_p = 0, v_1,\ldots,v_k$ are sequences of independent in time and identically distributed random vectors with marginal probability masses

$$\pi_w^{(i_1,\ldots,i_p)} := \mathbb{P}(I_t^{(i_1,\ldots,i_p)} = w),\ w = 0, v_1,\ldots,v_k.$$

If the index functions

$$f_x(y) = \frac{\prod_{x^*=0,v_1,\ldots,v_k,\ x^*\neq x}(y - x^*)}{\prod_{x^*=0,v_1,\ldots,v_k,\ x^*\neq x}(x - x^*)} = \begin{cases} 1, & \text{if } y = x \\ 0, & \text{if } y = 0, v_1,\ldots,v_k,\ y \neq x \end{cases}$$

are considered, then (1) is equivalent to

$$f_x(X_t) = \sum_{i_1,\ldots,i_p=0,v_1,\ldots,v_k} f_x(I_t^{(i_1,\ldots,i_p)}) \cdot f_{i_1}(X_{t-1}) \ldots f_{i_p}(X_{t-p}),\ x = 0, v_1,\ldots,v_k. \tag{2}$$

Under the condition of (*table*) causality (see Dimitriou-Fakalou (2019)), it is secured that $I_t^{(\ldots)}$ is independent of X_{t-l}, $l > 0$, so that (1) implies that

$$\mathbb{P}(X_t = x \mid X_{t-n} = x_n,\ n \in \mathbb{N}) \equiv \mathbb{P}(X_t = x \mid X_{t-n} = x_n,\ n = 1,\ldots,p) = \pi_x^{(x_1,\ldots,x_n)} \tag{3}$$

for $x, x_n = 0, v_1,\ldots,v_k$, $n \in \mathbb{N}$, and hence $\{X_t\}$ exhibits the pth order (homogeneous) Markovian structure.

Model (1) uses the *latent equalizer predictor* to define $\{X_t,\ t \in \mathbb{Z}\}$ and this is explained here further. For the linear series, we write the best (linear) predictor of X_t based on, say, X_{t-i} for some decided values of the $i \neq 0$: this is a linear function of the X_{t-i} with deterministic coefficients, i.e. a random variable; knowledge of the X_{t-i} leads to a realization of the predictor. The difference of X_t minus the predictor is defined as the *prediction error*, so that what singles out the best linear predictor as unique amongst other linear functions, is that *the best linear prediction error is uncorrelated with the relevant* X_{t-i}.

For the discretized series, we write the latent equalizer predictor of X_t based on X_{t-i}, $i \neq 0$, which is a random variable that is set *equal* to X_t, using the information according to X_{t-i} as it is their function; for each one of their combined value an X_t variable is assigned and, hence, knowledge of the X_{t-i} leads to a random variable, say, the conditional predictor with known distribution. What has characterized the latent equalizer predictor of X_t based on X_{t-i}, $i \neq 0$, is that the distribution of the conditional predictor remains unchanged under the presence of the relevant condition or not, implying some form of *independence* with it: for instance, in this section $\mathbb{P}(I_t^{(i_1,\ldots,i_p)} = w \mid X_{t-1} = i_1, \ldots, X_{t-p} = i_p) = \mathbb{P}(I_t^{(i_1,\ldots,i_p)} = w)$ verifies that $I_t^{(i_1,\ldots,i_p)}$ is independent of $f_{i_1}(X_{t-1}) \ldots f_{i_p}(X_{t-p})$.

Prediction examples: The best linear prediction of X_t based on some X_{t-i}, $i \neq 0$, is a linear function of the X_{t-i}; the intercept and coefficients are determined according to the first and second moments of the variables (parameters which are known or to be estimated). The latent equalizer prediction (of X_t based on some X_{t-i}, $i \neq 0$) is the realization of a random variable; its (discrete) distribution are the parameters which are known or to be estimated. In this section's notation, when $X_{t-1} = i_1,\ldots X_{t-p} = i_p$ the latent equalizer prediction of X_t can be the realization of a random variable with identical distribution as that of $I_t^{(i_1,\ldots,i_p)}$.

III. SPATIAL MODELING

Next suppose that instead of the time axis, the strictly stationary process of interest, say $\{X_s,\ s \in \mathbb{Z}\}$ takes place in the spatial line transect (see Whittle, 1954); this can still be expressed as in (1), yielding a unilateral representation based on the variables of one side of the transect X_{s-i}, $i > 0$, but it is explored here what is the spatial analogue. Subject to the condition of causality that links (1) to an adequate strictly stationary process $\{X_s\}$, it is now written that

$$X_s := Y_s^{[(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]}, \tag{4}$$

where for $\mathbf{i}_P = (i_{P,p}, \ldots, i_{P,1})$ (and $\mathbf{i}_P^\tau = (i_{P,1}, \ldots, i_{P,p})$ still a row vector), $\mathbf{i}_F = (i_{F,1}, \ldots, i_{F,p})$, it holds that (set any $i_{P,1}, \ldots, i_{P,p}$, $i_{F,1}, \ldots, i_{F,p} = 0, v_1, \ldots, v_k$)

$$m_s^{[\mathbf{i}_P;\mathbf{i}_F]} \cdot Y_s^{[\mathbf{i}_P;\mathbf{i}_F]} := M_s^{[\mathbf{i}_P;\mathbf{i}_F]}, \tag{5}$$

which must be accompanied by

$$m_s^{[\mathbf{i}_P;\mathbf{i}_F]} := \sum_{w=0,v_1,\ldots,v_k} f_w(I_s^{(\mathbf{i}_P^\top)}) \cdot f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\ldots,i_{P,p-1})}) \ldots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w)})$$

and

$$M_s^{[\mathbf{i}_P;\mathbf{i}_F]} := I_s^{(\mathbf{i}_P^\top)} \cdot f_{i_{F,1}}(I_{s+1}^{(I_s^{(\mathbf{i}_P^\top)},i_{P,1},\ldots,i_{P,p-1})}) \ldots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},I_s^{(\mathbf{i}_P^\top)})});$$

the last equation results in

$$\begin{aligned} f_w(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}) &= f_w(I_s^{(\mathbf{i}_P^\top)}) \cdot f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\ldots,i_{P,p-1})}) \ldots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w)}),\ w = v_1,\ldots,v_k, \\ f_0(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}) &= 1 - \sum_{w=v_1,\ldots,v_k} f_w(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}). \end{aligned}$$

Note that if $m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = 0$ then $M_s^{[\mathbf{i}_P;\mathbf{i}_F]} \equiv I_s^{(\mathbf{i}_P^\top)} \cdot m_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ is equal to zero as well. Hence, it is clarified that if the system (5) is null, then the variable Y cannot be considered at all, i.e. $Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ is defined on the sample space $\{m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = 1\}$. Provided that $m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = 1$, (5) can be replaced by $f_y(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}) = f_y(M_s^{[\mathbf{i}_P;\mathbf{i}_F]})$.

In addition, (4) is equivalent to

$$\begin{aligned} f_y(X_s) = \sum_{i_{P,p},\ldots,i_{P,1},i_{F,1},\ldots i_{F,p}=0,v_1,\ldots,v_k} & f_y(Y_s^{([\mathbf{i}_P;\mathbf{i}_F]}) \cdot f_{i_{P,p}}(X_{s-p}) \ldots f_{i_{P,1}}(X_{s-1}) \\ & f_{i_{F,1}}(X_{s+1}) \ldots f_{i_{F,p}}(X_{s+p}) \qquad (6) \end{aligned}$$

for all $y = 0, v_1, \ldots, v_k$. Directly from (6), it can be deduced that for any $y = 0, v_1, \ldots, v_k$, it holds that

$$\begin{aligned} & \mathbb{P}(X_s = y \mid X_{s-p} = i_{P,p}, \ldots, X_{s-1} = i_{P,1}, X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}) = \\ & \mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]} = y \mid X_{s-p} = i_{P,p}, \ldots, X_{s-1} = i_{P,1}, X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}) \ (7) \end{aligned}$$

$(\{X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p}, X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}\} \subseteq \{m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = 1\})$.

3.1 Interpretations

The variable $m_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ is Bernoulli and it holds that

$$\mathbb{P}(m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)=\sum_{w=0,v_1,\ldots,v_k}\mathbb{E}\{f_w(I_s^{(\mathbf{i}_P^\tau)})\}\,\mathbb{E}\{f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\ldots,i_{P,p-1})})\}\ldots\mathbb{E}\{f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w)})\} \tag{8}$$

(remember that I are assumed to be 'serially' independent on different spots of the line transect). The variable $M_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ is 0 or $1\cdot w$, $w=0,v_1,\ldots,v_k$, and it holds that

$$\begin{aligned}&\mathbb{P}(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w,m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)=\\&\mathbb{E}\{f_w(I_s^{(\mathbf{i}_P^\tau)})\}\,\mathbb{E}\{f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\ldots,i_{P,p-1})})\}\ldots\mathbb{E}\{f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w)})\}\end{aligned} \tag{9}$$

(for $w=v_1,\ldots,v_k$, it holds additionally that $\mathbb{P}(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w)=\mathbb{P}(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w,m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)$). It can be written, in general, then

$$\begin{aligned}\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w)\;&\equiv\;\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w\mid m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)=\mathbb{P}(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w\mid m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)=\\&\frac{\mathbb{P}(M_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w,m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)}{\mathbb{P}(m_s^{[\mathbf{i}_P;\mathbf{i}_F]}=1)},\quad w=0,v_1,\ldots,v_k.\end{aligned} \tag{10}$$

In the other hand, it holds that

$$\begin{aligned}&\mathbb{P}(X_s=w\mid(X_{s-1},\ldots,X_{s-p})=\mathbf{i}_P^\tau,(X_{s+1},\ldots,X_{s+p})=\mathbf{i}_F)=\\&\frac{\mathbb{P}(X_s=w,X_{s+1}=i_{F,1},\ldots,X_{s+p}=i_{F,p}\mid X_{s-1}=i_{P,1},\ldots,X_{s-p}=i_{P,p})}{\sum_{w^*=0,v_1,\ldots,v_k}\mathbb{P}(X_s=w^*,X_{s+1}=i_{F,1},\ldots,X_{s+p}=i_{F,p}\mid X_{s-1}=i_{P,1},\ldots,X_{s-p}=i_{P,p})}=\\&\frac{\mathbb{E}\{f_w(I_s^{(i_{P,1},\ldots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\ldots,i_{P,p-1})})\ldots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w)})\}}{\sum_{w^*=0,v_1,\ldots,v_k}\mathbb{E}\{f_{w^*}(I_s^{(i_{P,1},\ldots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(w^*,i_{P,1},\ldots,i_{P,p-1})})\ldots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\ldots,i_{F,1},w^*)})\}}\end{aligned}$$

based on the representation $X_s=I_s^{(X_{s-1},\ldots,X_{s-p})}$. By combining the derivation above together with (7) as well as (8), (9) and (10), it can be concluded that

$$\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w)=\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w\mid(X_{s-1},\ldots,X_{s-p})=\mathbf{i}_P^\tau,(X_{s+1},\ldots,X_{s+p})=\mathbf{i}_F). \tag{11}$$

In fact using similar arguments, it can be shown (for any $x_l=0,v_1,\ldots,v_k$, $|l|=p+1,p+2,\ldots$), that

$$\begin{aligned}&\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w)=\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w\mid\ (X_{s-1},\ldots,X_{s-p})=\mathbf{i}_P^\tau,\\&\qquad(X_{s+1},\ldots,X_{s+p})=\mathbf{i}_F,X_{s-l}=x_l,\ |l|=p+1,p+2,\ldots).\end{aligned} \tag{12}$$

It has been obvious already that $Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ (properly defined) is independent of X_{s-i}, $i>0$, as the former is a function of I_{s+j}, $j\in\mathbb{N}_0$ and the latter is a function of I_{s-i-j}, $j\in\mathbb{N}_0$: this is no improvement over what was there before defining Y_s, i.e. the random variable I_s being independent of X_{s-i}, $i>0$. The additional element that the equation (11) (together with (12)) brings in, is that it allows the probability of Y to remain unchanged under the condition of interest or not (implying a form of independence of $Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}$ with each Bernoulli variable on the set $\{(X_{s-1},\dots,X_{s-p})=\mathbf{i}_P^{\tau},\ (X_{s+1},\dots,X_{s+p})=\mathbf{i}_F, X_{s-l}=x_l,\ |l|=(p+1),\dots,(p+n)\},\ n\in\mathbb{N}_0$), subject to shrinking the sample space for the variable; this seems to justify $Y_s^{[(X_{s-p},\dots,X_{s-1});(X_{s+1},\dots,X_{s+p})]}$ being a latent equalizer predictor of X_s based on X_{s-i}, $|i|\in\mathbb{N}$.

3.2 Symmetrical Probabilities

For variables with two values only ($k=1$), say v_0, v_1 ($v_0\neq v_1$), there is interest whether it holds that

$$\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]}=w)\equiv\mathbb{P}(Y_s^{[\mathbf{i}_F^{\tau};\mathbf{i}_P^{\tau}]}=w),\ w=v_0,v_1 \tag{13}$$

(it is reminded that the two random variables are defined on different spaces and the index 's' for both is used 'loosely'). If (13) takes place, then

$$\frac{\mathbb{E}\{f_w(I_s^{(i_{P,1},\dots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(w,i_{P,1},\dots,i_{P,p-1})})\dots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\dots,i_{F,1},w)})\}}{\sum_{w^*=v_0,v_1}\mathbb{E}\{f_{w^*}(I_s^{(i_{P,1},\dots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(w^*,i_{P,1},\dots,i_{P,p-1})})\dots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\dots,i_{F,1},w^*)})\}}=$$
$$\frac{\mathbb{E}\{f_w(I_s^{(i_{F,1},\dots,i_{F,p})})f_{i_{P,1}}(I_{s+1}^{(w,i_{F,1},\dots,i_{F,p-1})})\dots f_{i_{P,p}}(I_{s+p}^{(i_{P,p-1},\dots,i_{P,1},w)})\}}{\sum_{w^*=v_0,v_1}\mathbb{E}\{f_{w^*}(I_s^{(i_{F,1},\dots,i_{F,p})})f_{i_{P,1}}(I_{s+1}^{(w^*,i_{F,1},\dots,i_{F,p-1})})\dots f_{i_{P,p}}(I_{s+p}^{(i_{P,p-1},\dots,i_{P,1},w^*)})\}}$$

and, consequently, (13) results in the key equality

$$\begin{aligned}&\mathbb{E}\{f_{v_0}(I_s^{(i_{P,1},\dots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(v_0,i_{P,1},\dots,i_{P,p-1})})\dots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\dots,i_{F,1},v_0)})\}\\&\mathbb{E}\{f_{v_1}(I_s^{(i_{F,1},\dots,i_{F,p})})f_{i_{P,1}}(I_{s+1}^{(v_1,i_{F,1},\dots,i_{F,p-1})})\dots f_{i_{P,p}}(I_{s+p}^{(i_{P,p-1},\dots,i_{P,1},v_1)})\}=\\&\mathbb{E}\{f_{v_1}(I_s^{(i_{P,1},\dots,i_{P,p})})f_{i_{F,1}}(I_{s+1}^{(v_1,i_{P,1},\dots,i_{P,p-1})})\dots f_{i_{F,p}}(I_{s+p}^{(i_{F,p-1},\dots,i_{F,1},v_1)})\}\\&\mathbb{E}\{f_{v_0}(I_s^{(i_{F,1},\dots,i_{F,p})})f_{i_{P,1}}(I_{s+1}^{(v_0,i_{F,1},\dots,i_{F,p-1})})\dots f_{i_{P,p}}(I_{s+p}^{(i_{P,p-1},\dots,i_{P,1},v_0)})\}.\end{aligned} \tag{14}$$

Lemma 1: If the order $p = 1$ or $p = 2$ then (13) definitely holds.

Proof: For $p = 1$, (14) becomes

$$\mathbb{E}\{f_{v_0}(I_s^{(i_{P,1})})f_{i_{F,1}}(I_{s+1}^{(v_0)})\}\mathbb{E}\{f_{v_1}(I_s^{(i_{F,1})})f_{i_{P,1}}(I_{s+1}^{(v_1)})\} =$$
$$\mathbb{E}\{f_{v_1}(I_s^{(i_{P,1})})f_{i_{F,1}}(I_{s+1}^{(v_1)})\}\mathbb{E}\{f_{v_0}(I_s^{(i_{F,1})})f_{i_{P,1}}(I_{s+1}^{(v_0)})\}$$

or $\pi_{v_0}^{(i_{P,1})}\pi_{i_{F,1}}^{(v_0)}\pi_{v_1}^{(i_{F,1})}\pi_{i_{P,1}}^{(v_1)} = \pi_{v_1}^{(i_{P,1})}\pi_{i_{F,1}}^{(v_1)}\pi_{v_0}^{(i_{F,1})}\pi_{i_{P,1}}^{(v_0)}$, which holds when $i_{P,1} = i_{F,1}$ as well as when $v_1 = i_{P,1} \neq i_{F,1} = v_0$ or $v_0 = i_{P,1} \neq i_{F,1} = v_1$.

For $p = 2$, the equality under question reduces to

$$\pi_{v_0}^{(i_{P,1},i_{P,2})}\,\pi_{i_{F,1}}^{(v_0,i_{P,1})}\pi_{i_{F,2}}^{(i_{F,1},v_0)} \cdot \pi_{v_1}^{(i_{F,1},i_{F,2})}\pi_{i_{P,1}}^{(v_1,i_{F,1})}\pi_{i_{P,2}}^{(i_{P,1},v_1)} =$$
$$\pi_{v_1}^{(i_{P,1},i_{P,2})}\pi_{i_{F,1}}^{(v_1,i_{P,1})}\pi_{i_{F,2}}^{(i_{F,1},v_1)} \cdot \pi_{v_0}^{(i_{F,1},i_{F,2})}\pi_{i_{P,1}}^{(v_0,i_{F,1})}\pi_{i_{P,2}}^{(i_{P,1},v_0)}$$

and the reader may try all different values to verify its validity. □

Next it is explained why (13) is investigated whether it holds. If it is true, it may be derived as well that

$$\mathbb{P}(X_s = w \mid (X_{s-1}, \ldots, X_{s-p}) = \mathbf{i}_P^\tau, (X_{s+1}, \ldots, X_{s+p}) = \mathbf{i}_F) =$$
$$\mathbb{P}(X_s = w \mid (X_{s-1}, \ldots, X_{s-p}) = \mathbf{i}_F, (X_{s+1}, \ldots, X_{s+p}) = \mathbf{i}_P^\tau). \tag{15}$$

In fact, it was Besag (1974) crystallizing that symmetry conditions often take place regarding the deterministic coefficients for spatial models of interest: for the auto-normal schemes (i.e. Gaussian variables that are characterized by an up to second-order dependence), those coefficients would be found in $\mathbb{E}\{X_s|X_{s-i},\ i = \pm 1, \ldots, \pm p\}$. Later, in the simultaneous Auto-Linear model of Dimitriou-Fakalou (2010), the symmetrical coefficients were demonstrated to be (proportional to) the auto-covariances of the Y latent process, which was a moving-average series with pairs of variables uncorrelated when more than p (i.e. the order of the model) steps away. Hence the short investigation that took place here regarding Y symmetrical probabilities is natural and just.

For Bernoulli variables and $p = 1$, the dependence is still linear (the conditional expectation of X_s given X_{s-i}, $i \geq 1$ is a linear function of X_{s-1}, the conditional variance is no constant though), so that an even auto-covariance function might describe it. The material point is that

$$\mathbb{E}\{X_{s-n}^{(v_P)}X_s^{(w)}X_{s+n}^{(v_F)}\} = \mathbb{E}\{X_{s-n}^{(v_F)}X_s^{(w)}X_{s+n}^{(v_P)}\}, \quad n \in \mathbb{N}, \tag{16}$$

for any $w, v_P, v_F = 0, 1$, where the operator $X^{(i)}$, $i = 0, 1$ generates either X $(i = 1)$ or $1 - X$ $(i = 0)$, when X is in $\{0, 1\}$. The general Bernoulli property (16) results in

$$\mathbb{P}(X_s = w \mid X_{s-n} = v_P, X_{s+n} = v_F) = \mathbb{P}(X_s = w \mid X_{s-n} = v_F,\ X_{s+n} = v_P),$$

which easily justifies all symmetrical probabilities (15) for $p = 1$.

Lemma 2: For any fixed order $p \in \mathbb{N}$ ($k = 1$ and values v_0, v_1), it holds that

$$\begin{aligned}&\mathbb{P}(X_s = v_1 \mid X_{s-n} = v_1,\ X_{s+n} = v_0,\ n = 1, \ldots, p) = \\ &\mathbb{P}(X_s = v_1 \mid X_{s-n} = v_0, X_{s+n} = v_1,\ n = 1, \ldots, p) \equiv \frac{\pi_{v_1}^{(v_1,\ldots,v_1)}}{\pi_{v_1}^{(v_1,\ldots,v_1)} + \pi_{v_0}^{(v_0,\ldots,v_0)}},\end{aligned} \tag{17}$$

which can be linked to a natural interpretation.

Proof: Obvious. □

Lemma 3: For any fixed order $p \in \mathbb{N}$ ($k = 1$ and values v_0, v_1), it is true that the 'oscillating' conditional probabilities are symmetrical, i.e.

$$\begin{aligned}&\mathbb{P}(X_s = w \mid X_{s-1} = v_1, X_{s+1} = v_0, X_{s-2} = v_0, X_{s+2} = v_1, \ldots \\ &\qquad (X_{s+p-1} =)X_{s-p} \neq X_{s+p}(= X_{s-p+1})) = \\ &\mathbb{P}(X_s = w \mid X_{s-1} = v_0, X_{s+1} = v_1, X_{s-2} = v_1, X_{s+2} = v_0, \ldots \\ &\qquad (X_{s+p-1} =)X_{s-p} \neq X_{s+p}(= X_{s-p+1}))\end{aligned}$$

holds for both $w = v_0, v_1$.

Proof: If p is even, it holds (for either $w = v_0, v_1$) that

$$\begin{aligned}&\mathbb{P}(X_s = w, X_{s+1} = v_0, X_{s+2} = v_1, \ldots X_{s+p} = v_1 \mid X_{s-1} = v_1, \ldots, X_{s-p} = v_0) = \\ &\mathbb{P}(X_s = w, X_{s+1} = v_1, X_{s+2} = v_0, \ldots, X_{s+p} = v_0 \mid X_{s-1} = v_0, \ldots, X_{s-p} = v_1);\end{aligned}$$

if p is odd, it can be shown that

$$\frac{\mathbb{P}(X_s = w, X_{s+1} = v_0, X_{s+2} = v_1, \ldots, X_{s+p} = v_0 \mid X_{s-1} = v_1, \ldots, X_{s-p} = v_1)}{\mathbb{P}(X_s = w, X_{s+1} = v_1, X_{s+2} = v_0, \ldots, X_{s+p} = v_1 \mid X_{s-1} = v_0, \ldots, X_{s-p} = v_0)}$$
$$= \frac{\pi_{v_0}^{(v_1\ v_0 \ldots v_1)}}{\pi_{v_1}^{(v_0\ v_1 \ldots v_0)}}, \ w = v_0, v_1. \quad \square$$

In addition to the Y latent series of the auto-linear model (Dimitriou-Fakalou, 2010), for any $p, k \in \mathbb{N}$, the variables Y as defined in Section 2 also follow a moving-average like pattern, in the sense that Y_s is *independent* of Y_{s+i}, $|i| \geq p+1$ (p-dependence). Due to the discretization within a finite number of categories and the extensive use of index variables, it is possible to compute joint probabilities of the Y variables. It is reminded that symmetry conditions for the original X modelled naturally in space, such as (15) when it holds, translate for the spatial latent equalizer predictors Y, such as (13), respectively. Nevertheless, the varying conditions under which the different Y series are considered could be making an essential difference to the various equalities, so the probabilities of Y or M should be computed with care. Below, there is a couple of examples of unobvious restrictions on the parameter probabilities of the spatial Y or M.

<u>Probability restriction examples</u>: We consider Bernoulli variables and order $p = 1$. We start with

$$\mathbb{P}(M_s^{[0;1]} = 1) = \pi_1^{(0)}\ \pi_1^{(1)} \neq \pi_1^{(1)}(1 - \pi_1^{(1)}) \equiv \mathbb{P}(M_s^{[1;0]} = 1), \tag{18}$$

as opposed to the special case of (13) when we know that

$$\mathbb{P}(Y_s^{[0;1]} = 1) \ = \ \frac{\pi_1^{(0)}\pi_1^{(1)}}{\pi_1^{(0)}\pi_1^{(1)} + (1-\pi_1^{(0)})\pi_1^{(0)}} = \frac{\pi_1^{(1)}(1-\pi_1^{(1)})}{\pi_1^{(1)}(1-\pi_1^{(1)}) + (1-\pi_1^{(1)})(1-\pi_1^{(0)})} \equiv \mathbb{P}(Y_s^{[1;0]} = 1) : \tag{19}$$

though the left-hand or right-hand side of (18) seems to be 'proportional' to the left-hand or right-hand, respectively, side of (19), an equality does not become apparent from (18); $Y_s^{[0;1]}$ and $Y_s^{[1;0]}$ are defined on different spaces resulting, here, in $\mathbb{P}(m_s^{[0;1]} = 1) \neq \mathbb{P}(m_s^{[1;0]} = 1)$.

We continue with interest in moments (probabilities) of two random variables, and we try to discover analogues to the *even* auto-covariance function of the auto-linear Y latent series by Dimitriou-Fakalou (2010). Hence we consider the case $\{M_s^{[0;1]} = 1, M_{s+1}^{[1;0]} = 1\}$ and compare it to $\{M_s^{[1;0]} = 0, M_{s+1}^{[0;1]} = 0\}$, in terms of probability: this seems to be the closest possible to reversing the order in the transect of the two M-variables, i.e. $M^{[0;1]}$ and $M^{[1;0]}$.

An adequate space must be taking place for either case. In the first case, we write $\{mm_s^{[0;1][1;0]} = 1\} := \left\{ I_s^{(0)} = 1, I_{s+2}^{(I_{s+1}^{(1)})} = 0 \right\}$ ('mm' is one symbol), with probability of occurrence

$$\begin{aligned} \mathbb{P}(mm_s^{[0;1][1;0]} = 1) &= \pi_1^{(0)} \, [\mathbb{P}(I_{s+1}^{(1)} = 1) \, \mathbb{P}(I_{s+2}^{(1)} = 0) + \mathbb{P}(I_{s+1}^{(1)} = 0) \, \mathbb{P}(I_{s+2}^{(0)} = 0)] \\ &= \pi_1^{(0)} \, (1 - \pi_1^{(1)}) \, (\pi_1^{(1)} + 1 - \pi_1^{(0)}), \end{aligned}$$

and followed by

$$\begin{aligned} \mathbb{P}(M_s^{[0;1]} = 1, M_{s+1}^{[1;0]} = 1 \mid mm_s^{[0;1][1;0]} = 1) &\equiv \frac{\mathbb{P}(I_s^{(0)} = 1, \; I_{s+1}^{(1)} = 1, \; I_{s+2}^{(1)} = 0)}{\mathbb{P}(mm_s^{[0;1][1;0]} = 1)} \\ &= \frac{\pi_1^{(1)}}{1 - \pi_1^{(0)} + \pi_1^{(1)}}. \end{aligned}$$

In the other hand, we write $\{mm_s^{[1;0][0;1]} = 1\} := \left\{ I_s^{(1)} = 0, I_{s+2}^{(I_{s+1}^{(0)})} = 1 \right\}$, and consider the events $\{M_s^{[1;0]} = 0\} = \{I_s^{(1)} = 0\} \cup \{I_{s+1}^{(I_s^{(1)})} = 1\}$ and $\{M_{s+1}^{[0;1]} = 0\} = \{I_{s+1}^{(0)} = 0\} \cup \{I_{s+2}^{(I_{s+1}^{(0)})} = 0\}$. Observe that

$$\mathbb{P}(M_s^{[1;0]} = 0, M_{s+1}^{[0;1]} = 0 \mid mm_s^{[1;0][0;1]} = 1) = \frac{\mathbb{P}(I_s^{(1)} = 0, I_{s+1}^{(0)} = 0, I_{s+2}^{(0)} = 1)}{\mathbb{P}(mm_s^{[1;0][0;1]} = 1)} = \frac{(1 - \pi_1^{(1)}) \, (1 - \pi_1^{(0)}) \, \pi_1^{(0)}}{(1 - \pi_1^{(1)}) \, \pi_1^{(0)} \, [\pi_1^{(1)} + (1 - \pi_1^{(0)})]} \equiv \frac{(1 - \pi_1^{(0)})}{\pi_1^{(1)} + 1 - \pi_1^{(0)}}.$$

Not only is the sum of the two conditional probabilities equal to unity (this might be attributed to the fact that the realizations 0 and 1 have to switch as well), but even more so it is true that

$$\begin{aligned}
\mathbb{P}(M_s^{[0;1]} = 1, M_{s+1}^{[1;0]} = 1 \mid mm_s^{[0;1][1;0]} = 1) &= \mathbb{P}(X_s = 1 \mid X_{s-1} = 1, X_{s+1} = 0) \\
&\equiv \mathbb{P}(X_s = 1 \mid X_{s-1} = 0, X_{s+1} = 1), \\
\mathbb{P}(M_s^{[1;0]} = 0, M_{s+1}^{[0;1]} = 0 \mid mm_s^{[1;0][0;1]} = 1) &= \mathbb{P}(X_s = 0 \mid X_{s-1} = 1, X_{s+1} = 0) \\
&\equiv \mathbb{P}(X_s = 0 \mid X_{s-1} = 0, X_{s+1} = 1),
\end{aligned}$$

((17) might be referred to) and the relation between the symmetrical conditional (from past and future) probabilities of X, translates to the joint probabilities of M-adjusted (Y-like variables). This basic result can be considered as one of a plethora of unobvious restrictions that do take place regarding the Y series dependence: those might concern joint probabilities of not necessarily just two variables (especially for $p \geq 2$). Thanks to the model (4), with Y defined from (5), following M and m defined from the I series, there should be no danger that the restrictions might be overlooked by the spatial parameterization. The building block of the I series (hiding behind the m, M and Y) should make sure that any probability is computable and any relationship in the probabilities may reveal itself.

3.3 Time Reversal

It is easily demonstrated that once the pth order Markovian dependence resulting from the 'causal' representation in the transect takes place, the same holds for the unilateral representation from the other side: hence, for any $n \in \mathbb{N}_0$, it holds that

$$\begin{gathered}
\mathbb{P}(X_s = x_0 \mid X_{s+1} = x_1, \ldots, X_{s+p} = x_p, \ldots, X_{s+p+n} = x_{p+n}) = \\
\frac{\mathbb{P}(X_s = x_0, X_{s+1} = x_1, \ldots, X_{s+p} = x_p, \ldots, X_{s+p+n} = x_{p+n})}{\mathbb{P}(X_{s+1} = x_1, \ldots, X_{s+p} = x_p, \ldots, X_{s+p+n} = x_{p+n})} = \\
\frac{\mathbb{P}(X_s = x_0, X_{s+1} = x_1, \ldots, X_{s+p} = x_p)}{\mathbb{P}(X_{s+1} = x_1, \ldots, X_{s+p} = x_p)} \cdot \\
\frac{\mathbb{P}(X_{s+p+n} = x_{p+n}, \ldots, X_{s+p+1} = x_{p+1} \mid X_{s+p} = x_p, \ldots, X_{s+1} = x_1, X_s = x_0)}{\mathbb{P}(X_{s+p+n} = x_{p+n}, \ldots, X_{s+p+1} = x_{p+1} \mid X_{s+p} = x_p, \ldots, X_{s+1} = x_1)},
\end{gathered}$$

where straight from the p-variate memoryless property (3), the last ratio is equal to unity, so that one can continue by writing

$$\frac{\mathbb{P}(X_s = x_0, X_{s+1} = x_1, \ldots, X_{s+p} = x_p)}{\mathbb{P}(X_{s+1} = x_1, \ldots, X_{s+p} = x_p)} = \mathbb{P}(X_s = x_0 \mid X_{s+1} = x_1, \ldots, X_{s+p} = x_p),$$

i.e. a causal $\{X_s,\ s \in \mathbb{Z}\}$ satisfying (3) also satisfies

$$\begin{aligned}&\mathbb{P}(X_s = x_0 \mid X_{s+1} = x_1, \ldots, X_{s+p} = x_p, \ldots, X_{s+p+n} = x_{p+n}) = \\ &\mathbb{P}(X_s = x_0 \mid X_{s+1} = x_1, \ldots, X_{s+p} = x_p), \text{ for any } n \in \mathbb{N}_0.\end{aligned} \tag{20}$$

Additionally, a more general statement concerning a series of variables that occur holds as well, i.e.

$$\begin{aligned}&\mathbb{P}(X_{s-n} = x_{-n}, \ldots, X_{s-1} = x_{-1},\ X_s = x_0 \mid X_{s+1} = x_1, \ldots, X_{s+p} = x_p) = \\ &\prod_{m=0}^{n} \mathbb{P}(X_{s-m} = x_{-m} \mid X_{s-m+1} = x_{-m+1}, \ldots, X_{s-m+p} = x_{-m+p}\ ), \text{ for any } n \in \mathbb{N}_0.\end{aligned} \tag{21}$$

Thanks to (20) and (21), we may *continue under the following assumption*: there can also be $(k+1)^p$ sequences, say $\{J_s^{(j_1,\ldots,j_p)},\ s \in \mathbb{Z}\}$, $j_1, \ldots, j_p = 0, v_1, \ldots, v_k$ of independent (on different points of the transect and from either sequence) and identically distributed variables, with marginal distributions

$$\rho_w^{(j_1,\ldots,j_p)} := \mathbb{P}(J_s^{(j_1,\ldots,j_p)} = w) \equiv \mathbb{P}(X_s = w \mid X_{s+1} = j_1, \ldots, X_{s+p} = j_p)$$

for $w = 0, v_1, \ldots, v_k$, *such that it can be written for any* $s \in \mathbb{Z}$, that

$$X_s \equiv J_s^{(X_{s+1},\ldots,X_{s+p})}. \tag{22}$$

Then it *must* hold that

$$\rho_w^{(j_1,\ldots,j_p)} \equiv \pi_{j_p}^{(j_{p-1},\ldots,j_1,w)} \cdot \frac{\psi_{(j_{p-1},\ldots,j_1,w)}}{\psi_{(j_p,\ldots,j_1)}}, \tag{23}$$

where $\psi_{(i_1,\ldots,i_p)} := \mathbb{P}(X_{s-1} = i_1, \ldots, X_{s-p} = i_p)$ are the 'stationary' probabilities, which have been obtained as a unique solution to the linear system (i.e. satisfying)

$$\psi_{(i_1,\ldots,i_p)} = \sum_w \pi_{i_1}^{(i_2,\ldots,i_p,w)} \cdot \psi_{(i_2,\ldots,i_p,w)}.$$

The linear system

$$\psi^*_{(i_1,\dots,i_p)} = \sum_w \rho_{i_p}^{(i_{p-1},\dots,i_1,w)} \cdot \psi^*_{(w,i_1,\dots,i_{p-1})}$$

must yield the same stationary distribution $\psi^* = \psi$ to be the unique solution: by substitution in the system above of $\rho_{i_p}^{(i_{p-1},\dots,i_1,w)}$ according to (23), it is easy to verify that ψ is a solution indeed. Additionally, $\{X_s\}$ must be a 'causal' (from the 'future') process based on $\{J_s^{(\dots)}\}$ (see Dimitriou-Fakalou (2019)).

Following an identical argument as from the I to the Y, it can be written

$$n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} := \sum_{w=0,v_1,\dots,v_k} f_w(J_s^{(\mathbf{i}_F)}) \cdot f_{i_{P,1}}(J_{s-1}^{(w,i_{F,1},\dots,i_{F,p-1})}) \dots f_{i_{P,p}}(J_{s-p}^{(i_{P,p-1},\dots,i_{P,1},w)})$$

and

$$N_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} := J_s^{(\mathbf{i}_F)} \cdot f_{i_{P,1}}(J_{s-1}^{(J_s^{(\mathbf{i}_F)},i_{F,1},\dots,i_{F,p-1})}) \dots f_{i_{P,p}}(J_{s-p}^{(i_{P,p-1},\dots,i_{P,1},J_s^{(\mathbf{i}_F)})}),$$

so that it is defined

$$n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} \cdot Y_s^{*[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} := N_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]}, \tag{24}$$

i.e. $Y_s^{*[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]}$ is defined on the sample space $\{n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} = 1\}$. Then it *must* hold that

$$X_s = Y_s^{*[(X_{s+p},\dots,X_{s+1});(X_{s-1},\dots,X_{s-p})]} \equiv Y_s^{[(X_{s-p},\dots,X_{s-1});(X_{s+1},\dots,X_{s+p})]}. \tag{25}$$

Equation (25) sets $J_s^{(\mathbf{i}_F)} \equiv I_s^{(\mathbf{i}_P^\tau)}$ but only *provided that* $X_{s-n} = i_{P,n}$, $X_{s+n} = i_{F,n}$, $n = 1,\dots,p$.

In fact, regardless of the values of X_{s-n}, $n = \pm 1,\dots,\pm p$, it could be considered *for every* $\mathbf{i}_P$ and $\mathbf{i}_F$, whether it holds that $m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} \equiv 1$; then, provided that

$$\begin{aligned}
&\mathbb{P}(Y_s^{[\mathbf{i}_P;\mathbf{i}_F]} = Y_s^{*[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} = w \mid m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} \equiv 1) = \\
&\mathbb{P}(I_s^{(\mathbf{i}_P^\tau)} = J_s^{(\mathbf{i}_F)} = w \mid m_s^{[\mathbf{i}_P;\mathbf{i}_F]} = n_s^{[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]} \equiv 1) = \\
&\mathbb{P}(X_s = w \mid (X_{s-1},\dots,X_{s-p}) = \mathbf{i}_P^\tau, (X_{s+1},\dots,X_{s+p}) = \mathbf{i}_F)
\end{aligned} \tag{26}$$

holds for all $w = 0, v_1, \ldots, v_k$, it would be fair to conclude that $Y_s^{[\mathbf{i}_P;\mathbf{i}_F]} = Y_s^{*[\mathbf{i}_F^\tau;\mathbf{i}_P^\tau]}$ (defined on the intersection of sample spaces) has all the elements, to take the role of the (conditional) latent equalizer predictor of X_s given $X_{s-n} = i_{P,n}$, $X_{s+n} = i_{F,n}$, $n = 1, \ldots, p$ and X_{s-n}, $|n| \geq p+1$: as it has been explained already, $Y_s^{[\ldots;\ldots]}$ (properly defined) is independent of X_{s-n}, $n \in \mathbb{N}$ and, for the same reasons, $Y_s^{*[\ldots;\ldots]}$ (properly defined) would be independent of X_{s+n}, $n \in \mathbb{N}$. It is mandatory for every $s \in \mathbb{Z}$, that for *at least one* out of the $(k+1)^{2p}$ possible $i_{P,1}, \ldots, i_{P,p}$, $i_{F,1}, \ldots, i_{F,p} = 0, v_1, \ldots, v_k$, it holds that $m_s = 1$ and $n_s = 1$ and $Y_s = Y_s^*$, together with

$$\begin{aligned}
&\mathbb{P}(Y_s^{[(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]} = Y_s^{*[(X_{s+p},\ldots,X_{s+1});(X_{s-1},\ldots,X_{s-p})]} = w \mid X_{s-i},\ i \neq 0) = \\
&\quad \mathbb{P}(Y_s^{[(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]} = w \mid X_{s-i},\ i \neq 0) \equiv \\
&\quad \mathbb{P}(Y_s^{*[(X_{s+p},\ldots,X_{s+1});(X_{s-1},\ldots,X_{s-p})]} = w \mid X_{s-i},\ i \neq 0);
\end{aligned}$$

it remains a question to be answered though, whether (26) can or should be attempted to hold. Formally, one would be able to consider for every $s \in \mathbb{Z}$, that the latent equalizer predictor of X_s based on X_{s-n}, $|n| \in \mathbb{N}$, is $Y_s^{[(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]} = Y_s^{*[(X_{s+p},\ldots,X_{s+1});(X_{s-1},\ldots,X_{s-p})]}$ and should proceed conditionally via Y *or* Y^*.

IV. AN INTRODUCTION TO THE TALMA

Write $\mathbf{j}_P = (j_{P,q}, \ldots, j_{P,1})$ (and $\mathbf{j}_P^\tau = (j_{P,1}, \ldots, j_{P,q})$), $\mathbf{j}_F = (j_{F,1}, \ldots, j_{F,q})$ and consider $\{I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)},\ s \in \mathbb{Z}\}$ to be a series of independent and identically distributed $(k+1)^{p+q}$-vectors $(i_{P,1}, \ldots, i_{P,p},\ j_{P,1}, \ldots, j_{P,q} = 0, v_1, \ldots, v_k)$, which will be used as a building block for the series of interest. Write

$$\pi_x^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} := \mathbb{P}(I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = x),\ x = 0, v_1, \ldots, v_k, \quad \sum_{x=0,v_1,\ldots,v_k} \pi_x^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = 1,$$

as well as $I_s^{(\mathbf{0}_p|\mathbf{0}_q)} \equiv I_s$ and $\pi_x^{(\mathbf{0}_p|\mathbf{0}_q)} \equiv \pi_x$.

Definition 1: $\{X_s,\ s \in \mathbb{Z}\}$ will be called a Table Auto-Linear Moving-Average process of order (k, p, q) and it will be written $\{X_s\} \sim \text{TALMA}(k, p, q)$, if it holds that

$$X_s := Y_s^{([(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]|[(Y_{s-q},\ldots,Y_{s-1});(Y_{s+1},\ldots,Y_{s+q})])}, \ s \in \mathbb{Z}, \qquad (27)$$

paired with

$$Y_s \ := \ y_s^{([(X_{s-p},\ldots,X_{s-1});(X_{s+1},\ldots,X_{s+p})]|[(Y_{s-q},\ldots,Y_{s-1});(Y_{s+1},\ldots,Y_{s+q})])}, \ s \in \mathbb{Z},$$

where, for every $\mathbf{i}_P$, $\mathbf{i}_F$, $\mathbf{j}_P$ and $\mathbf{j}_F$, it is set

$$\begin{aligned} m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} \ := \ & f_{j_{F,1}}(I_{s+1}) \cdot \ldots \cdot f_{j_{F,q}}(I_{s+q}) \cdot \\ & f_{i_{F,1}}(I_{s+1}^{((I_s^{((\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}}))},i_{P,1},\ldots,i_{P,p-1})|(I_s,j_{P,1},\ldots,j_{P,q-1}))}) \cdot \ldots \cdot \\ & f_{i_{F,p}}(I_{s+p}^{((i_{F,p-1},\ldots,i_{F,1},I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})})|(I_{s+p-1}(?=j_F),\ldots,j_{F,1},I_s,j_{P,1},\ldots,j_P))}), \quad (28) \end{aligned}$$

followed by defining on the sample space $\{m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1\}$ only, the variables

$$y_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} := I_s, \quad \text{and} \quad Y_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} := I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})}.$$

1. *Definition 1* presents the spatial analogue to the Table ARMA (k, p, q) equation (see Dimitriou-Fakalou (2019)), which models the infinite Markovian dependence (see Dimitriou-Fakalou (2022)), in the same way that the standard time series ARMA is an AR(∞). Indeed it can be verified for the process of interest $\{X_s, \ s \in \mathbb{Z}\}$ that $X_s = I_s^{((X_{s-1},\ldots,X_{s-p})|(I_{s-1},\ldots,I_{s-q}))}$, $s \in \mathbb{Z}$: it is a prerequisite that this is causal, implying strictly stationary, (as in Dimitriou-Fakalou (2019)) and invertible (as in Dimitriou-Fakalou (2022)). Additionally, the main latent process $\{I_s\}$ converts into $\{Y_s\}$, which shares the same realizations but, each time, the y-value is picked from a different probability rule, depending on the fixed values of X and I from both sides. Note that, if (28) is re-written as

$$\begin{aligned} & f_{j_{F,1}}(Y_{s+1}) \cdot \ldots \cdot f_{j_{F,q}}(Y_{s+q}) \cdot \\ & f_{i_{F,1}}(I_{s+1}^{((I_s^{((\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}}))},i_{P,1},\ldots,i_{P,p-1})|(I_s,j_{P,1},\ldots,j_{P,q-1}))}) \cdot \ldots \cdot \\ & f_{i_{F,p}}(I_{s+p}^{((i_{F,p-1},\ldots,i_{F,1},I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})})|(I_{s+p-1}(?=j_F),\ldots,j_{F,1},I_s,j_{P,1},\ldots,j_P))}), \end{aligned}$$

then the definition of $y_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])}$ resembles that of a TARMA (k, q, p) caused *from the future*, though the comparisons should be done with care (this is due to merging the different y processes into one Y for the 'AR' part and considering a p-dependence for the 'MA' part). Still, it should be clear that X_s and $y_{s+l}^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])}$, $Y_{s+l}^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])}$, for *positive* (and valid) l, are independent (in the same way that X_s and I_{s+l}, $l > 0$, are independent).

To exhibit the independence of X_s and $y_{s-l}^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])}$, $Y_{s-l}^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])}$, $l > 0$, some form of time reversal would be required, i.e. writing X_s as a 'causal' (from the future) function of another sequence of, say J, iid variables, with the same variables from the past only determining the values of y, Y. Nevertheless, not only would time reversal be expected to be an arduous (if not impossible) task, but also it is *not* really needed, in order to solidify the vital property of the TALMA model (as in (29) and the statement following it, which conclude this point).

So without resorting to any other than the I building blocks, first see that

$$\begin{gathered}\{X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}, \\ I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q},\ I_{s+1} = j_{F,1}, \ldots, I_{s+q} = j_{F,q}\} \subseteq \\ \{m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1\}.\end{gathered}$$

Secondly, it is easy to spot

$$\begin{aligned}\mathbb{P}(X_s = x, Y_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}, \\ & I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q},\ I_{s+1} = j_{F,1}, \ldots, I_{s+q} = j_{F,q}) = \\ \mathbb{P}(I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = x,\ I_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q}, \\ & m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1) \equiv \\ \mathbb{P}(I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = x, I_s = y \quad | \quad & m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1),\end{aligned}$$

where the last equality is due to the fact that X_{s-l}, I_{s-n}, $l, n > 0$ are jointly independent of $I_s^{(\ldots)}$, $m_s^{(\ldots)}$. Identical arguments write, for any $l_1, m_1 \in \mathbb{N}$, that

$$\begin{aligned}\mathbb{P}(X_s = x, Y_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ X_{s+1} = i_{F,1}, \ldots, X_{s+p} = i_{F,p}, \\ & I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q},\ I_{s+1} = j_{F,1}, \ldots, I_{s+q} = j_{F,q}, \\ & X_{s-p-l},\ l = 1, \ldots, l_1,\ I_{s-q-m},\ m = 1, \ldots, m_1) = \\ \mathbb{P}(I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})} = x, I_s = y \quad | \quad & m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1).\end{aligned}$$

Thirdly, set for convenience $o := \max\{p, q\}$, $\mathbf{i}_F^* = (\mathbf{i}_F, i_{F,p+1}, \ldots, i_{F,o})$, $\mathbf{j}_F^* = (\mathbf{j}_F, j_{F,q+1}, \ldots, j_{F,o})$, and slightly shrink the sample space to $\{m_s^{*([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} = 1\} \subseteq \{m_s^{([\mathbf{i}_P;\mathbf{i}_F]|[\mathbf{j}_P;\mathbf{j}_F])} = 1\}$, where

$$\begin{aligned}m_s^{*([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} \quad := \quad & f_{j_{F,1}}(I_{s+1}) \cdot \ldots \cdot f_{j_{F,o}}(I_{s+o}) \cdot \\ & f_{i_{F,1}}(I_{s+1}^{((I_s^{((\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}}))}, i_{P,1}, \ldots, i_{P,p-1})|(I_s, j_{P,1}, \ldots, j_{P,q-1}))}) \cdot \ldots \cdot \\ & f_{i_{F,p}}(I_{s+p}^{((i_{F,p-1}, \ldots, i_{F,1}, I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})})|(j_{F,p-1}, \ldots, j_{F,1}, I_s, j_{P,1}, \ldots, j_P))}) \cdot \ldots \cdot \\ & f_{i_{F,o}}(I_{s+o}^{((i_{F,o-1}, \ldots)|(j_{F,o-1}, \ldots, j_{F,1}, I_s))})\end{aligned}$$

(the first line fixes the case $p - 1 > q$, so that $j_{F,p-1}$ exists, and the $\cdot \ldots \cdot$ last line takes care of when $q > p$, so that 'I_s' is the last entry in the MA part), so that it becomes clear why it holds that

$$\begin{aligned}\mathbb{P}(I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})} = x, I_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q}, \\ & m_s^{*([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} = 1) \equiv \\ \mathbb{P}(I_s^{(\mathbf{i}_P^{\mathcal{T}}|\mathbf{j}_P^{\mathcal{T}})} = x, I_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ I_{s-1} = j_{P,1}, \ldots, I_{s-q} = j_{P,q}, \\ & m_s^{*([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} = 1, \\ & X_{s+o+n},\ n = 1, \ldots, n_1,\ I_{s+o+n^*},\ n^* = 1, \ldots, n_2),\end{aligned}$$

for any $n_1, n_2 \in \mathbb{N}$: a similar relation can be found in Section 3.1.

To sum it up, the TALMA(k, p, q) model for $\{X_s\}$, is such that the probability

$$\begin{aligned}\mathbb{P}(X_s = x, Y_s = y \quad | \quad & X_{s-1} = i_{P,1}, \ldots, X_{s-p} = i_{P,p},\ X_{s-p-l},\ l = 1, \ldots, l_1, \\ & Y_{s-1} = j_{P,1}, \ldots, Y_{s-q} = j_{P,q},\ Y_{s-q-m},\ m = 1, \ldots, m_1, \\ & X_{s+1} = i_{F,1}, \ldots, X_{s+o} = i_{F,o},\ X_{s+o+n},\ n = 1, \ldots, n_1, \\ & Y_{s+1} = j_{F,1}, \ldots, Y_{s+o} = j_{F,o},\ Y_{s+o+n^*},\ n^* = 1, \ldots, n_2),\end{aligned} \tag{29}$$

remains *unchanged* for any l_1, m_1, n_1, $n_2 \in \mathbb{N}_0$.

The adjustment from m_s to m_s^* (which is not the case for the linear series) might be attributed to the fact that the model is multiplicative on the AL and MA parts, i.e. those two are tangled together.

2. It is clarified next how to employ the special merit of the TALMA equation, which has been attached to the use of the partitioned Y latent series. When $q = 0$, the TAL(k,p) process is none other than what was described in Section 3. It is accustomed that the inclusion of the moving-average part with order $q \in \mathbb{N}$, offers a parsimonious way to model the *infinite Markovian dependence*: its *spatial version* is what the TALMA is all about.

   To demonstrate how to approximate it, first for any $n \in \mathbb{N}_0$, write the 'conditions'

$$\mathcal{P}_{s,(n)}(x_{-\mu},\ \mu = (n+1),\ldots,(n+p),\ y_{-\nu},\ \nu = (n+1),\ldots,(n+q)) :=$$
$$\{X_{s-\mu} = x_{-\mu},\ \mu = (n+1),\ldots,(n+p),\ Y_{s-\nu} = y_{-\nu},\ \nu = (n+1),\ldots,(n+q)\},$$

and

$$\mathcal{F}_{s,(n)}(x_m,\ y_m,\ m = (n+1),\ldots,(n+o)) :=$$
$$\{X_{s+m} = x_m,\ Y_{s+m} = y_m,\ m = (n+1),\ldots,(n+o)\}.$$

Then define the probabilities

$$\omega_{(x_0,y_0)}^{([(x_{-p},\ldots,x_{-1});(x_1,\ldots,x_o)]|[(y_{-q},\ldots,y_{-1});(y_1,\ldots,y_o)])} := \mathbb{P}(X_s = x_0, Y_s = y_0 \mid \mathcal{P}_{s,(0)},\ \mathcal{F}_{s,(0)}).$$

More specifically, for any $n \in \mathbb{N}_0$, the quantity

$$\mathbb{P}(X_s = x_0 \mid X_{s+m} = x_m,\ m = \pm 1,\ldots,\pm n,\ X_{s+\mu} = x_\mu,\ \mu = -(n+p),\ldots$$
$$-(n+1),(n+1),\ldots,(n+o),\ Y_{s+\nu} = y_\nu,\ \nu = -(n+q),\ldots,-(n+1),$$
$$(n+1),\ldots,(n+o)) \tag{30}$$

will be of interest. To compute (30), first re-write it as

$$
\begin{aligned}
&\{\sum_{y_m,\ m=0,\pm1,\ldots,\pm n} \mathbb{P}(X_s = x_0,\ Y_s = y_0,\ X_{s+m} = x_m,\ Y_{s+m} = y_m,\ m = \pm1,\ldots,\pm n \mid \\
&\qquad X_{s+\mu} = x_\mu,\ \mu = -(n+p),\ldots,-(n+1),(n+1),\ldots,(n+o), \\
&\qquad Y_{s+\nu} = y_\nu,\ \nu = -(n+q),\ldots,-(n+1),(n+1),\ldots,(n+o)\ )\}\ / \\
&\{\sum_{x_0^*,y_m,\ m=0,\pm1,\ldots,\pm n} \mathbb{P}(X_s = x_0^*,\ Y_s = y_0,\ X_{s+m} = x_m,\ Y_{s+m} = y_m,\ m = \pm1,\ldots,\pm n \mid \\
&\qquad X_{s+\mu} = x_\mu,\ \mu = -(n+p),\ldots,-(n+1),(n+1),\ldots,(n+o), \\
&\qquad Y_{s+\nu} = y_\nu,\ \nu = -(n+q),\ldots,-(n+1),(n+1),\ldots,(n+o)\ )\};
\end{aligned}
$$

then observe the important relation

$$
\begin{aligned}
&\frac{\mathbb{P}(X_{s+m} = xx_m,\ Y_{s+m} = yy_m,\ m = 0,\pm1,\ldots,\pm n \mid \mathcal{P}_{s,(n)},\mathcal{F}_{s,(n)})}{\mathbb{P}(X_{s+m} = xx_m^*,\ Y_{s+m} = yy_m^*,\ m = 0,\pm1,\ldots,\pm n \mid \mathcal{P}_{s,(n)},\mathcal{F}_{s,(n)})} = \prod_{i=-n}^{n} \\
&\{\mathbb{P}(X_{s+i} = xx_i, Y_{s+i} = yy_i \mid \mathcal{P}_{s,(n)},\ X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n, \\
&\ \ldots, i-1,\ X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*,\ j^* = i+1,\ldots,n,\ \mathcal{F}_{s,(n)})\}\ / \\
&\{\mathbb{P}(X_{s+i} = xx_i^*, Y_{s+i} = yy_i^* \mid \mathcal{P}_{s,(n)},\ X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n, \\
&\ \ldots, i-1,\ X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*,\ j^* = i+1,\ldots,n,\ \mathcal{F}_{s,(n)})\}, \quad (31)
\end{aligned}
$$

which finds its roots in Besag (1974) together with a sketch of proof. Indeed, by taking its right-hand side, it can be re-written

$$
\begin{aligned}
&\prod_{i=-n}^{n} \{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,i, \\
&\qquad X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*,\ j^* = i+1,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}\ / \\
&\qquad \{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,i-1, \\
&\qquad X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*,\ j^* = i,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\};
\end{aligned}
$$

for $i = -n$, the factor becomes

$$
\begin{aligned}
&\{\mathbb{P}(X_{s-n} = xx_{-n},\ Y_{s-n} = yy_{-n},\ X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*, \\
&\qquad j^* = -n+1,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}\ / \\
&\{\mathbb{P}(X_{s+j^*} = xx_{j^*}^*,\ Y_{s+j^*} = yy_{j^*}^*,\ j^* = -n,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}
\end{aligned}
$$

with its denominator being exactly what is wanted and the numerator will cancel with the denominator of the next factor (for $i = -n+1$), which is altogether

$$\{\mathbb{P}(X_{s-n} = xx_{-n},\ Y_{s-n} = yy_{-n},\ X_{s-n+1} = xx_{-n+1},\ Y_{s-n+1} = yy_{-n+1},$$
$$X_{s+j^*} = xx^*_{j^*},\ Y_{s+j^*} = yy^*_{j^*},\ j^* = -n+2,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}\ /$$
$$\{\mathbb{P}(X_{s-n} = xx_{-n},\ Y_{s-n} = yy_{-n},\ X_{s+j^*} = xx^*_{j^*},\ Y_{s+j^*} = yy^*_{j^*},$$
$$j^* = -n+1,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}$$

and so on. Similarly, the one before the end factor ($i = n-1$)

$$\{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,n-1,$$
$$X_{s+n} = xx^*_n,\ Y_{s+n} = yy^*_n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}\ /$$
$$\{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,n-2,\ X_{s+n-1} = xx^*_{n-1},$$
$$Y_{s+n-1} = yy^*_{n-1},\ X_{s+n} = xx^*_n,\ Y_{s+n} = yy^*_n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\},$$

will have its denominator cancelled with the numerator of the previous factor ($i = n-2$) and its numerator cancelled with the denominator of the last one for $i = n$, which is

$$\{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}\ /$$
$$\{\mathbb{P}(X_{s+j} = xx_j,\ Y_{s+j} = yy_j,\ j = -n,\ldots,n-1,$$
$$X_{s+n} = xx^*_n,\ Y_{s+n} = yy^*_n \mid \mathcal{P}_{s,(n)},\ \mathcal{F}_{s,(n)})\}:$$

as for the numerator of the last part above, together with the denominator of the first factor, those are combined to verify exactly the statement under question.

Once (31) is valid, given the same condition $\mathcal{P}_{s,(n)}$, $\mathcal{F}_{s,(n)}$, the distribution

$$\mathbb{P}(X_{s+m} = xx_m,\ Y_{s+m} = yy_m,\ m = 0, \pm 1, \ldots, \pm n \mid \mathcal{P}_{s,(n)}, \mathcal{F}_{s,(n)}),$$

i.e. for any xx_m, yy_m, $m = 0, \pm 1, \ldots, \pm n$, can be determined: this should be obvious how. As for the right-hand side of (31), thanks to the conclusion regarding the probability (29) remaining unchanged, it suffices to know the probabilities ω to have at hand all that is required for it, then for (30). Straight from (30), the infinite spatial Markovian dependence is implied as $n \to \infty$, and the result is always contained within the prespecified TALMA parameters ω.

3. The parameters relating to *Definition 1* are none other than the probabilities ω, and those should be expressed in terms of the original π, i.e.

$$
\begin{aligned}
\omega_{(x,y)}^{([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} &\equiv \mathbb{P}(I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = x, I_s = y \mid m_s^{*([\mathbf{i}_P;\mathbf{i}_F^*]|[\mathbf{j}_P;\mathbf{j}_F^*])} = 1) \\
&= \mathbb{P}(I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)} = x, I_s = y \mid I_{s+1} = j_{F,1}, \ldots, I_{s+o} = j_{F,o}, \\
I_{s+1}^{((I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)}, i_{P,1}, \ldots, i_{P,p-1})|(I_s, j_{P,1}, \ldots, j_{P,q-1}))} &= i_{F,1}, \ldots, I_{s+o}^{((i_{F,o-1},\ldots)|(j_{F,o-1},\ldots,j_{F,1},I_s))} = i_{F,o}) :
\end{aligned}
$$

particular emphasis must be paid on making sure that the condition considered each time has a strictly *positive* probability of occuring (this can be referred to as *positivity* condition). For example, when it is set $(\mathbf{i}_P, \mathbf{j}_P) = \mathbf{0}_{p+q}$ (i.e. $I_s^{(\mathbf{i}_P|\mathbf{j}_P)} \equiv I_s$) this has to be accounted for in the condition first. Other cases $(\mathbf{i}_P, \mathbf{j}_P) \neq \mathbf{0}_{p+q}$ rely on the *interdependence* of $I_s^{(\mathbf{i}_P^\tau|\mathbf{j}_P^\tau)}$ and I_s, referring to variables at the same point $s \in \mathbb{Z}$ (rather than the independence of I when considered at different points of the transect): nothing has been said about that here, but the fixes for the causality (or invertibility) of the process usually demand that the distributions of the different $I_s^{(\ldots)}$ must be really near (see Dimitriou-Fakalou (2019) and Dimitriou-Fakalou (2022)).

V. FINAL POINTS: WHY TALMA?

Given the correspondence from the causal TARMA(k, p, q) to the TALMA(k, p, q) equation, as well as the fact that it is the usual tack, from a sample of observed values to derive optimal inference results via step-by-step projections on the *past* available, one might wonder, why bother with the TALMA model?

The reasons can be counted on three fingers, the first one being that it might be *desirable* to *understand* how the (stationary) spatial dependencies extend. Secondly, it might be *preferable* to *naturally clothe* the spatial relations using the TALMA not TARMA equation, when the variables of interest are indexed on more than one coordinates, mainly, in order to avoid the conventional (half-plane) unilateral order of Whittle (1954): instead of writing the conditional probability of one value based on its 'previous' ones, all (other than itself) values can contribute to the information. Finally, there is the *practical reason of prediction in spatial settings*: for example, the TAL($q = 0$) equation provides directly with the whole distribution of the variable of interest based on values from a finite number of 'neighbours' from all around, when additional information can be discarded.

In the cases of multi-indexed discretized stationary processes, it is still advised that simultaneous spatial models would better not be attempted directly: the previous care of causal arrangements in the lattice that will, nevertheless, remain hidden, is essential for the validity of the spatial considerations.

REFERENCES

1. Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion), Journal of the Royal Statistical Society, Series B, 36, 192-236.
2. Brockwell, P.J. and Davis, R.A. (1991). Time Series: Theory and Methods, second edition, Springer-Verlag, New-York.
3. Cressie, N.A.C. (1993). Statistics for Spatial Data, revised edition, Wiley, New York.
4. Cui, Y. and Lund, R.B. (2009). A new look at time series of counts, Biometrika, 96, 781-792.
5. Dimitriou-Fakalou, C. (2010). Statistical Inference for Spatial Auto-Linear Processes, Journal of Statistical Theory and Practice, 4, 345-365.
6. Dimitriou-Fakalou, C. (2019). The Table Auto-Regressive Moving-Average model for (categorical) stationary series: statistical properties (causality; from the all random to the conditional random), Journal of Nonparametric Statistics, 31, 31-63.
7. Dimitriou-Fakalou, C. (2022). The Table Auto-Regressive Moving-Average Model for (Categorical) Stationary Series: Mathematical Perspectives (Invertibility; Maximum Likelihood Estimation), Open Journal of Statistics, 12, 385-407.
8. Du, J.G. and Li, Y. (1991). The integer-valued auto-regressive (INAR(p)) model, Journal of Time Series Analysis, 12, 129-142.
9. Francq, C. and Zakoïan, J.-M. (2001). Stationarity of multivariate Markov-switching ARMA models, Journal of Econometrics, 102, 339-364.

10. Jacobs, P.A. and Lewis, P.A.W. (1978). Discrete time series generated by mixtures I: correlational and runs properties, Journal of the Royal Statistical Society, Series B, 40, 94-105.
11. Keenan, D.M. (1982). A time series analysis of binary data, Journal of the American Statistical Association, 77, 816-821.
12. Lomnicki, Z.A. and Zaremba, S.K. (1955). Some applications of zero-one processes, Journal of the Royal Statistical Society, Series B, 17, 243-255.
13. McKenzie, E. (1985). Some simple models for discrete variate time series, Water Resources Bulletin, 21, 645-650.
14. Tong, H. (1990). Non-linear Time Series: A Dynamical Systems Approach, Oxford University Press, Oxford.
15. Whittle, P. (1954). On Stationary Processes in the Plane, Biometrika, 41, 434-449.

This page is intentionally left blank