Programma 'Short Mobility' 2010
Relazione scientifica sull'attività di ricerca svolta.

I contenuti dell'attività di ricerca svolta dal sottoscritto (proponente e fruitore) presso la School of Electrical Engineering dell'Università di Adelaide-SA (Australia), nel periodo 23 - Aprile 2010, 24 Maggio 2010, nell'ambito del programma 'Short Mobility' 2010, sono contenuti in un lavoro in corso di preparazione e che sarà sottoposto per la pubblicazione alla rivista 'IEEE Transactions on Automatic Control':

Autori: Francesco Carravetta; Langford B. White.

Titolo: Recursive Optimal Fixed-Interval Smoothing for Hidden Reciprocal Processes. (Part I and II).

Si tratta di un lavoro diviso in due articoli (Part I, II), che saranno sottoposti alla suddetta rivista in due tempi diversi. La parte I è quasi ultimata e sarà sottoposta al più in un mese. L'attività di ricerca in oggetto ha riguardato aspetti del problema e ottenuto risultati che saranno collocati sia nella parte I che nella parte II.

Si allega alla presente un draft-paper, che include molti dei risultati ottenuti durante il soggiorno e che verranno inseriti nella parte I. Si allega inoltre una nota, riguardante la derivazione dello smoother ottimo per un processo reciproco, che è stata elaborata dallo scrivente interamente durante il soggiorno, e che verrà inserita nella parte II.

Il problema specifico affrontato dal lavoro è quello della realizzazione stocastica e del filtraggio/smoothing ottimo (a minima varianza) per processi reciproci. La parte I del lavoro riguarda la realizzazione stocastica, la parte II il filtraggio/smoothing. Si premette una breve descrizione del problema che non compare nel draft o nella nota allegate (compariranno nell'abstract e nell'Introduzione, ancora non scritte).
Un campo di Markov è un modello di segnale (ci limitiamo al caso di segnali a stati finiti su un insieme finito di siti, questi segnali sono detti anche 'catene') descritto da una famiglia di transizioni di probabilità di un sito dal suo 'nearest neighbour' ('intorno contiguo', cioè l'insieme dei siti connessi al sito. Un HRM (Hidden Reciprocal Model) è un campo di Markov a una dimensione (in cui cioè l'insieme dei siti è assimilabile ad un intervallo discreto di numeri naturali), detto anche processo reciproco, e che in aggiunta è dotato di una famiglia di probabilità, dato il segnale ad un sito, di misurazione di una qualche quantità osservabile in quel sito.
Un HRM non è un modello 'dinamico' usuale perché un campo di Markov, anche ridotto ad una sola dimensione, non è in generale un processo di Markov, e ciò a sua volta implica che non può esistere un modello generale causale di generazione del segnale stesso. Si sottolinea qui che un modello causale invece, è la base per la costruzione di un qualsiasi algoritmo ricorsivo inerente al processo (e cioè praticamente realizzabile con strumenti di calcolo ordinari).

Il problema dell'esistenza di processi reciproci fu inizialmente posto da Schroedinger, nel 1932, [1] nel tentativo di descrivere matematicamente un elettrone: nella concezione della meccanica quantistica infatti un elettrone è assimilabile ad una distribuzione di probabilità nello spazio (una segmento reale limitato, nel caso prototipale), tuttavia problemi matematici insorgono nel momento in cui si vuole determinare tale distribuzione a partire da 1) una famiglia di transizioni di probabilità 2) due condizioni al contorno assegnate alle estremità del segmento. Infatti una condizione al contorno determina il processo come processo di Markov. Due condizioni non sono invece sempre soddisfacibili da un modello di Markov. Questo portò successivamente Jamison [2] a definire la classe dei processi reciproci, che è caratterizzata dal fatto di essere una generalizzazione del concetto di processo di Markov. Hammersey e Clifford, [3] Grimmett [4] generalizzarono poi ulteriormente questa nozione al caso di processi definiti su grafi. Il problema della realizzazione stocastica (cioè di descrizione del processo attraverso una equazione, di cui il processo stesso sia soluzione) fu posto, e risolto nel caso Gaussiano, successivamente in altri lavori di Levi e Krener [5]. In particolare in [5], nel caso di processo reciproco Gaussiano a tempo discreto è introdotto il concetto di modello reciproco. Un modello reciproco ha la notevole proprietà di essere, pur non-causale, decomponibile in una coppia di equazioni ricorsive (cioè causali) che 'corrono' in direzioni opposte lungo la linea. Ciò rappresenta un'ottima generalizzazione, dal punto di vista operativo, del concetto di 'sistema dinamico' in situazioni non-causali. In un recente lavoro dello scrivente [6] l'esistenza di modelli reciproci è stata dimostrata nel caso di catene reciproche generali. Nello stesso lavoro algoritmi di stima ricorsiva sono stati definiti sulla base di questi modelli reciproci, e la loro efficacia nel caso sub ottimo polinomiale è stata dimostrata.

Nel presente lavoro, l'attenzione è rivolta a catene reciproche canoniche (a valori cioè nella base canonica di uno spazio vettoriale a dimensione finita). Ogni catena reciproca è equivalente ad una catena canonica (a scapito eventualmente di una crescita delle dimensioni del sistema). La rappresentazione canonica tuttavia implica una serie di proprietà semplificatrici. Ad esempio, mentre in [6] il modello reciproco è polinomiale, con un grado massimo proporzionale al numero di stati della catena, nel presente lavoro si dimostra che una catena reciproca canonica ha sempre un modello quadratico. Nello stesso lavoro inoltre, la soluzione di altri problemi inerenti alla realizzazione stocastica, come la simulazione del modello, o la determinazione dei coefficienti del sistema a partire da statistiche standard, è affrontata e risolta. Per quanto riguarda il problema della stima, diverse soluzioni, bayesiane e non, sono proposte per la definizione di stimatori ricorsivi ottimi (in [6] gli stimatori proposti sono tutti sub-ottimi).

Analiticamente, per quanto riguarda il draft-paper allegato: il Lemma 3 e i Teoremi 1,2 contengono i principali risultati teorici sulla cui base la realizzazione stocastica è costruita. Specificamente la realizzazione stocastica è l'equazione (40).
La variabile di 'stato' è la quantità Z, funzione della coordinata t (che non è necessariamente tempo, più comunemente è una coordinata spaziale) mentre la funzione 'e' , è un 'rumore' (correlato ad un passo) che svolge la funzione sistemica di 'input', e possiede la notevole proprietà di essere incorrelato con Z a differenti t. Si noti che l'equazione non è ricorsiva, tuttavia le matrici M, coefficienti dell'equazione sono legate fra loro dalle relazioni (41), questo implica che il modello è reciproco, secondo la definizione data da Levi [5] nel caso Gaussiano (e ripresa dallo scrivente in [6] nel caso non canonico a stati finiti). Questo implica, come illustrato in un'altra sezione del draft (si veda il capitolo a pag 14 'the Levi's double swapping procedure') la sua equivalenza con una coppia di equazioni ricorsive che 'corrono' in direzioni opposte (a titolo illustrativo si vedano le eq. (103), (104): queste sono relative allo stimatore dello stato, quelle proprie dello stato sono del tutto simili, ma pilotate dal 'rumore': 'e', anziché dalle osservazioni). Alla fine del draft è anche indicato un modo per costruire uno smoother lineare-ottimo per il processo reciproco. Questo risultato, pur subottimo, è tuttavia di interesse in quanto costituisce un algoritmo ricorsivo lineare, con complessità computazionale di tipo n^2 (con n dimensione del sistema). Questo risultato sarà comunque probabilmente spostato nella successiva parte II del lavoro, riguardante specificamente la stima.

Sempre in relazione all'attività svolta presso l'Università di Adelaide nel mese di short mobility, si allega la nota dal titolo
'Note-May2010', contenente la derivazione dello smoother ottimo per il processo reciproco. Il punto di partenza è l'equazione (1) della nota, già descritta prima, che assieme alle condizioni (2) rappresenta un modello reciproco. Le eq.ni (6), (7) rappresentano il modello equivalente backward/forward. (9) è l'equazione equivalente delle misure.
La tecnica si basa sul cambio di misura descritto dalla derivata di Radon Nykodim (13). L'eq. (17) mostra che nella nuova misura le osservazioni Y costituiscono una sequenza iid (independent identically distributed, in particolare la distribuzione è uniforme) indipendente da Z. L'identità (20) assicura che Z continua ad essere reciproco anche nella nuova misura. L'algoritmo di stima ottima consiste nel calcolo ricorsivo della distribuzione condizionata non-normalizzata definita da (23) per la stima di filtraggio, e da (24) per quella di interpolazione (lo smoother). Nel presente caso canonico a stati finiti queste distribuzioni coincidono con le stime non-normalizzate (25), (26). Sfruttando il modello reciproco (1) e la sua rappresentazione equivalente (6), (7) si arriva alle equazioni ricorsive (37), (38) (per il filtro e lo smoother rispettivamente), dove H è la matrice definita da (36), e c(Y) è la funzione (lineare) definita da (28) (l'eq- (28) è in forma esponenziale, ma è facile mostrare che vale anche una forma lineare, sempre a causa della canonicità del processo).
Si noti che le equazioni ricorsive della stima ottima sono uguali per il filtro e per lo smoother (la differenza essendo data da differenti condizioni al contorno). Lo stimatore ottimo è quindi di tipo bilineare (la qual cosa è conforme con la ben nota equazione, bilineare, di Mortensen-Zakai, per il caso tempo-continuo, valori-continui, che fornisce la densità di probabilità condizionata e non normalizzata per un generico sistema tempo continuo stocastico non lineare).
La sua complessità computazionale è anch'essa n^2, con n : dimensione di Z.

Bibliografia

[1] E. Schroedinger. 'Theorie Relativiste de l'electron et l'interpretation de la mecanique quantique'. Ann. Inst. H. Poincare 2, 269—310 (1932)

[2] B. Jamison. 'Reciprocal Processes' *Z. Wahrscheinlichkeitstheorie verw. Gebiete 30, 65—86 (1974)*

[3] J.M. Hammersley, P. Clifford. 'Markov Fields on Finite Graphs and Lattices'. *Unpublished Paper (1971)*

[4] G.R. Grimmet. 'A Theorem about Random Fields'. *Bull. London Math. Soc. 5, 81—84 (1973)*

[5] B. C. Levy, R. Frezza, A. J. Krener. 'Modeling and Estimation of Discrete-Time Gaussian Reciprocal Processes', *Trans. On Aut. Contr. 35, 9, 1013—1023 (1990)*

[6] F. Carravetta. 'Nearest-neighbour modelling of reciprocal chains'. *Stochastics: An International Journal of Probability and Stochastic Processes. 80, 6, 525--584 (2008).*

IL PROPONENTE

Francesco Carravetta

Note-MAY2010

We have (I rename e_t of the draft as v_t to distiguish of e canonical base vector of $\mathbb{R}^n$):

$$M_t^0 Z_t + M_t^- Z_{t-1} + M_t^+ Z_{t+1} = v_t, \tag{1}$$

$$M_t^+ = {M_{t+1}^-}^T, \quad M_t^0 = {M_t^0}^T \tag{2}$$

and let's consider the above equation in $0 < t < N$ (in the 'interior' of the interval). Let's apply the Levy's decomposition, so the above are equivalent to

$$\xi_t = M_{t+1}^+ Q_{t+1} \xi_{t+1} + v_t \tag{3}$$

$$Z_{t+1} = Q_t M_t^- Z_t + Q_t \xi_t \tag{4}$$

with $M_t^+ = {M_{t+1}^-}^T$, so let

$$A_t = M_t^+ Q_t; \tag{5}$$

and (3), (4), rewrites

$$\xi_t = A_{t+1} \xi_{t+1} + v_t \tag{6}$$

$$Z_{t+1} = A_t^T Z_t + Q_t \xi_t \tag{7}$$

where $Q_t = Q_t^T > 0$ comes from solving a backward matrix difference equation (like eq. (48) in the draft). Recall that (from the draft):

$$Z_t = Z_t^+ - \mathbf{E}\{Z_t^+\};$$
$$Z_t^+ = X_{t-1} \otimes X_{t+1},$$

The *uncentered* observation equation is

$$Y_t^+ = \widetilde{C}_t X_t + W_t, \tag{8}$$

which, in terms of Z^+, rewrites

$$Y_{t+1}^+ = C_t Z_t^+ + W_{t+1}, \tag{9}$$

and $Y_t = Y_t^+ - \mathbf{E}\{Y_t^+\}$, which obviously generates the same σ-algebra. Let ζ_t a process defined on $\{[0, \nu]\}$, ν a positive integer. We use the following notation for the related σ-algebras:

$$\mathcal{F}_t^\zeta = \sigma(\zeta_0, ..., \zeta_t);$$
$$\mathcal{F}_{-t}^\zeta = \sigma(\zeta_0, ..., \zeta_{t-1}, \zeta_{t+1}, ..., \zeta_\nu). \tag{10}$$

We assume Dirichelet boundary condition, so X_t is defined on $[0, N]$ (and X_0, X_N are completely observable), and Y_t, Z_t are defined on $(0, N)$. For ζ a process defined on $[0, N]$ we denote $\mathcal{F}^\zeta = \mathcal{F}_N^\zeta$. Let $f_1, ..., f_m$, $e_1, ..., e_n$, the two bases in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Also, let's use Elliott's notation $\langle \cdot \cdot \cdot \rangle$ for the scalar product. Note that $\mathcal{F}^X = \mathcal{F}^Z$. Let's view the processes X, Z, Y as living on the probability space $(\Omega, \mathcal{F}^X \vee \mathcal{F}^Y, \mathbf{P})$ Accordingly with Elliott, eq. (1.6) pag. 210, we assume

$$\mathbf{P}\{\langle Y_t^+, f_i \rangle = 1 | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} = \mathbf{P}\{\langle Y_t^+, f_i \rangle = 1 | X_t\}.$$

Let's perform a change of measure $P \to \overline{P}$, governed by the RN-derivative $\Lambda = \Lambda_N = d\mathbf{P}/d\overline{\mathbf{P}}$:

$$\frac{d\overline{\mathbf{P}}}{d\mathbf{P}} = \Lambda_N$$

$$\Lambda_t = \prod_{l=0}^{t} \lambda_l \tag{11}$$

$$\lambda_{l+1} = \prod_{i=1}^{m} \left(\frac{1}{m \cdot \mathbf{P}\big\{\langle Y_l^+, f_i\rangle = 1 | \mathcal{F}^X \vee \mathcal{F}_{-l}^Y\big\}} \right)^{\langle Y_l^+, f_i\rangle} \tag{12}$$

As in Elliot, it's possible to show that:

$$\left.\frac{d\overline{\mathbf{P}}}{d\mathbf{P}}\right|_{\mathcal{F}_t^X \vee \mathcal{F}_{t-1}^Y} = \Lambda_t, \tag{13}$$

i.e. the restricion of the Radon-Nykodim derivative $d\mathbf{P}/d\overline{\mathbf{P}}$ over the sub-σ-algebra $\mathcal{F}_t^X \vee \mathcal{F}_{t-1}^Y$ – namely, the Radon-Nykodim derivative of the corresponding measures restrictions – agrees with Λ_t. From the above definition, and accounting of W_t being independent of $\mathcal{F}^X \vee \mathcal{F}_{-t}^Y$, it follows that

$$\begin{aligned}
\mathbf{P}\{\langle Y_t^+, f_i\rangle = 1 | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \quad &= \mathbf{E}\{\langle Y_t^+, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \\
&= \mathbf{E}\{\langle C_{t-1} Z_{t-1}^+ + W_t, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \\
&= \mathbf{E}\{\langle C_{t-1} Z_{t-1}^+, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \\
&= \langle C_{t-1} Z_{t-1}^+, f_i\rangle = \langle \widetilde{C}_t X_t, f_i\rangle
\end{aligned}$$

as Z_t^+ is $\mathcal{F}^X$-measurable. Note that, Λ_t results in a $\mathcal{F}_{t-1}^X \vee \mathcal{F}_{t-1}^Y$-measurable function (i.e. it is measurable with respect to a sub-σ-algebra of the σ-algebra where the two measures which define it got restricted to). Thus λ_t can be rewritten as

$$\begin{aligned}
\lambda_{t+1} \quad &= \prod_{i=1}^{m} \left(\frac{1}{m\langle C_{t-1} Z_{t-1}^+, f_i\rangle} \right)^{\langle Y_t^+, f_i\rangle} = \prod_{i=1}^{m} \left(\frac{1}{m\langle \widetilde{C}_t X_t, f_i\rangle} \right)^{\langle Y_t^+, f_i\rangle} \\
&= \sum_{i=1}^{m} \frac{\langle Y_t^+, f_i\rangle}{m\langle C_{t-1} Z_{t-1}^+, f_i\rangle} = \sum_{i=1}^{m} \frac{\langle Y_t^+, f_i\rangle}{m\langle \widetilde{C}_t X_t, f_i\rangle}
\end{aligned}$$

where the expression involving the summation are easily verified by exploiting Y^+ being canonical-base valued. Thus Λ_t is $\mathcal{F}_{t-2}^Z \vee \mathcal{F}_{t-1}^Y$-measurable, and $\mathcal{F}_{t-1}^X \vee \mathcal{F}_{t-1}^Y$-measurable.

In the following it will be used the Kallianpur-Striebel formula (sometimes also referred to as Bayes formula), which expresses a link between the two measures, as to the conditional expectations. Using the above setting, let $\mathcal{G}$, $\mathcal{H}$ two sub-σ-algebras of $\mathcal{F}^X \vee \mathcal{F}^Y$, and φ any $\mathcal{H}$-measurable random vector, it is

$$\overline{\mathbf{E}}\{\varphi | \mathcal{G}\} = \frac{\mathbf{E}\{\Lambda\varphi | \mathcal{G}\}}{\mathbf{E}\{\Lambda | \mathcal{G}\}}. \tag{14}$$

Note that, in formula (14) the RN-derivative Λ can be replaced by any restriction Λ_t, provided that the sub-σ-algebra where such restriction applies ($\mathcal{F}_t^X \vee \mathcal{F}_{t-1}^Y$ as for Λ_t) includes $\mathcal{H}$.

By definition, it's easy to verify that

$$\mathbf{E}\{\lambda_{t+1}/\mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} = 1. \tag{15}$$

As a matter of fact

$$\begin{aligned}\mathbf{E}\{\lambda_{t+1}/\mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} &= \mathbf{E}\left\{\sum_{i=1}^m \frac{\langle Y_t^+, f_i\rangle}{m\cdot \mathbf{E}\big\{\langle Y_t^+, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\big\}}\middle| \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\right\} \\ &= \sum_{i=1}^m \frac{1}{m} = 1. \end{aligned}\tag{16}$$

One has

$$\begin{aligned}\overline{\mathbf{P}}\{\langle Y_t^+, f_r\rangle = 1 | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} &= \overline{\mathbf{E}}\{\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \\ &= \frac{\mathbf{E}\{\Lambda_{t+1}\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}}{\mathbf{E}\{\Lambda_{t+1} | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}} \\ &= \frac{\Lambda_t}{\Lambda_t}\cdot\frac{\mathbf{E}\{\lambda_{t+1}\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}}{\mathbf{E}\{\lambda_{t+1} | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}} \\ &= \mathbf{E}\{\lambda_{t+1}\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} \\ &= \mathbf{E}\left\{\prod_{i=1}^m \left(\frac{1}{m\cdot \mathbf{P}\big\{\langle Y_t^+, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\big\}}\right)^{\langle Y_t^+, f_i\rangle}\langle Y_t^+, f_r\rangle \middle| \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\right\} \\ &= \mathbf{E}\left\{\sum_{i=1}^m \frac{\langle Y_t^+, f_i\rangle\langle Y_t^+, f_r\rangle}{m\cdot \mathbf{E}\big\{\langle Y_t^+, f_i\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\big\}}\middle| \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\right\} \\ &= \frac{1}{m}\frac{\mathbf{E}\{\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}}{\mathbf{E}\{\langle Y_t^+, f_r\rangle | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\}} = \frac{1}{m} \end{aligned}\tag{17}$$

thus $\overline{\mathbf{P}}\{\langle Y_t^+, f_r\rangle = 1 | \mathcal{F}^X \vee \mathcal{F}_{-t}^Y\} = \overline{\mathbf{P}}\{\langle Y_t^+, f_r\rangle = 1\}$, and hence

$$Y_t \perp \mathcal{F}^X; \quad Y_t \perp \mathcal{F}_{-t}^Y. \tag{18}$$

and in particular

$$Y_t \perp \mathcal{F}^Z, \tag{19}$$

under the measure $\overline{\mathbf{P}}$. Moreover,

$$\begin{aligned}\overline{\mathbf{E}}\{X_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\} &= \frac{\mathbf{E}\{\Lambda_t X_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\}}{\mathbf{E}\{\Lambda_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\}} \\ &= \frac{\Lambda_{t-1}}{\Lambda_{t-1}}\mathbf{E}\{\lambda_t X_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\} \\ &= \mathbf{E}\left\{\sum_{i=1}^m \frac{\langle \widetilde{C}_{t-1}X_{t-1} + W_{t-1}, f_i\rangle X_t}{m\langle \widetilde{C}_{t-1}X_{t-1}, f_i\rangle}\middle| \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\right\}\end{aligned}$$

the term depending of W_{t-1} vanishes, as W_{t-1} is independent of $\mathcal{F}^X \vee \mathcal{F}_{t-2}^Y$, thus:

$$\overline{\mathbf{E}}\{X_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\} = \mathbf{E}\{X_t | \mathcal{F}_{-t}^X \vee \mathcal{F}_{t-2}^Y\},$$

and taking the conditional expectation with respect to $\mathcal{F}^X_{-t}$:

$$\overline{\mathbf{E}}\{X_t|\mathcal{F}^X_{-t}\} = \mathbf{E}\{X_t|\mathcal{F}^X_{-t}\}, \tag{20}$$

which directly entails being X a reciprocal process under the measure $\overline{\mathbf{P}}$ as well.

As a result, we have that under $\overline{\mathbf{P}}$ we have that Y is a white process independent of X (and, hence, of Z), and Z keeps on satisfying the reciprocal representation (1), as well as the (equivalent) backward/forward representation (6), (7).

The optimal filter and smoother of the reciprocal process.

First of all note that $v_t = Z_t - \mathbf{E}\{Z_t|\mathcal{F}^X_{-t}\}$, and Z_t is $\mathcal{F}^X_{-t}$-measurable. Hence v_t is $\mathcal{F}^X_{-t}$-measurable, thus, from the above, under measure $\overline{\mathbf{P}}$, v_t is orthogonal to all of Y:

$$v_t \perp \mathcal{F}^Y. \tag{21}$$

Let

$$\overline{\Lambda}_t = \left.\frac{d\mathbf{P}}{d\overline{\mathbf{P}}}\right|_{\mathcal{F}^X_t \vee \mathcal{F}^Y_{t-1}} = \Lambda_t^{-1} \qquad (t = 0, ..., N). \tag{22}$$

Let's define the unnormalized and uncentered conditional (filtering) distribution of $Z_t : q_t(e_r), \quad r = 1, ..., n$:

$$q_t(e_r) = \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1}\langle Z_t^+, e_r\rangle|\mathcal{F}^Y_t\}. \tag{23}$$

and the unnormalized, uncentered, conditional smoothing distribution of Z_t : $q^{\mathrm{s}}_t(e_r), \quad r = 1, ..., n$:

$$q^{\mathrm{s}}_t(e_r) = \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1}\langle Z_t^+, e_r\rangle|\mathcal{F}^Y\}. \tag{24}$$

The corresponding unnormalized, uncentered, filtering (smoothing) estimate of Z_t will be denoted by ζ_t (ζ^{s}_t):

$$\zeta_t = \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+|\mathcal{F}^Y_t\}. \tag{25}$$

$$\zeta^{\mathrm{s}}_t = \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+|\mathcal{F}^Y\}. \tag{26}$$

Let $\mathcal{G}$ denote any of σ-algebras $\mathcal{F}^Y$, $\mathcal{F}^Y_t$. One has

$$\begin{aligned}
&\overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+|\mathcal{G}\} \\
&= \overline{\mathbf{E}}\{\overline{\Lambda}_t \prod_{i=1}^m m \cdot \langle C_{t-1} Z_{t-1}^+, f_i\rangle^{\langle Y_t^+, f_i\rangle} Z_t^+|\mathcal{G}\} \\
&= m \cdot \overline{\mathbf{E}}\{\overline{\Lambda}_t \sum_{i=1}^m \langle C_{t-1} Z_{t-1}^+, f_i\rangle\langle Y_t^+, f_i\rangle Z_t|\mathcal{G}\} \\
&= m \cdot \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^n Z^{+j}_{t-1}\left(\sum_{i=1}^m C^{i,j}_{t-1}\langle Y_t^+, f_i\rangle\right) Z_t^+|\mathcal{G}\right\} \\
&= \sum_{j=1}^n \overline{\mathbf{E}}\{\overline{\Lambda}_t\langle Z_{t-1}^+, e_j\rangle Z_t^+|\mathcal{G}\} \prod_{i=1}^m m C^{i,j}_{t-1}{}^{\langle Y_t^+, f_i\rangle},
\end{aligned} \tag{27}$$

Denote

$$c_s(Y_t^+) = \prod_{i=1}^{m} mC_s^{i,j\,\langle Y_t^+, f_i\rangle} \tag{28}$$

so, denoting:

$$h_t = \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_t^+ \middle| \mathcal{F}_{t-1}^Y\right\}$$

$$h_t^{\mathrm{s}} = \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_t^+ \middle| \mathcal{F}^Y\right\}$$

on account of Y_t being, under $\overline{\mathbf{P}}$, independent of $\mathcal{F}^X \vee \mathcal{F}_{t-1}^Y$, it is

$$\overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+ | \mathcal{F}_t^Y\} = h_t c_{t-1}(Y_t) \tag{29}$$

$$\overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+ | \mathcal{F}^Y\} = h_t^{\mathrm{s}} c_{t-1}(Y_t). \tag{30}$$

Now, as we have seen just before, under $\overline{\mathbf{P}}$ the reciprocal equation (1) holds. Also, the uncentered process Z^+ satisfies the same equation (see the draft), so it is

$$Z_t^+ = M_t^{0^{-1}} v_t - M_t^{0^{-1}} M_t^- Z_{t-1}^+ - M_t^{0^{-1}} M_t^+ Z_{t+1}^+. \tag{31}$$

In particular v_t under $\overline{\mathbf{P}}$ keeps on satisfying the orthogonality property $v_t \perp Z_s$, $s \neq t$. Thus, $\mathcal{G}$ now denoting $\mathcal{F}^Y$ or $\mathcal{F}_{t-1}^Y$:

$$\begin{aligned}\overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_t^+ \middle| \mathcal{G}\right\} &= -M_t^{0^{-1}} M_t^- \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_{t-1}^+ \middle| \mathcal{G}\right\} \\ &\quad - M_t^{0^{-1}} M_t^+ \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_{t+1}^+ \middle| \mathcal{G}\right\} \\ = -M_t^{0^{-1}} M_t^- \overline{\mathbf{E}}\left\{\overline{\Lambda}_t Z_{t-1}^+ | \mathcal{G}\right\} &- M_t^{0^{-1}} M_t^+ \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_{t+1}^+ \middle| \mathcal{G}\right\}. \end{aligned} \tag{32}$$

Now, let's consider eq. (7). As the process ξ_t, by (6), is some (linear) function of $v_t, v_{t+1}, ..., v_{N-1}$, it has the same independence properties, i.e. it is independent of $Z_0, ..., Z_{t-1}$, and – under $\overline{\mathbf{P}}$ – independent of $\mathcal{F}^Y$ as well. Thus, replacing the right side of eq. (7) for Z_{t+1}^+, and erasing the v-depending term, results in

$$\overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_{t+1}^+ \middle| \mathcal{G}\right\} = A^T \overline{\mathbf{E}}\left\{\overline{\Lambda}_t \sum_{j=1}^{n} \langle Z_{t-1}^+, e_j\rangle Z_t^+ \middle| \mathcal{G}\right\}. \tag{33}$$

By using (33) in (32), substituting for $\mathcal{G}$ the σ-algebras $\mathcal{F}^Y$, $\mathcal{F}_{t-1}^Y$ and then recognizing the terms h_t, h_t^{s}, $\zeta_t, \zeta_t^{\mathrm{s}}$ we get the following equations

$$h_t = -M_t^{0^{-1}} M_t^- \zeta_{t-1} - M_t^{0^{-1}} M_t^+ A_t^T h_t,$$

$$h_t^{\mathrm{s}} = -M_t^{0^{-1}} M_t^- \zeta_{t-1}^{\mathrm{s}} - M_t^{0^{-1}} M_t^+ A_t^T h_t^{\mathrm{s}}$$

which, on account of (5), leads to

$$h_t = H_t \zeta_{t-1}, \tag{34}$$

$$h_t^{\mathrm{s}} = H_t \zeta_{t-1}^{\mathrm{s}} \tag{35}$$

with

$$H_t = -\left(I + M_t^{0^{-1}} M_t^+ Q_t M_t^{+^T}\right)^{-1} M_t^{0^{-1}} M_t^- .^1 \tag{36}$$

Finally, substituting (34), (35) in (29) and (30), recognizing $\zeta, \zeta^{\mathrm{s}}$ in the right hand side, results in

$$\zeta_t = H_t \zeta_{t-1} c_{t-1}(Y_t), \tag{37}$$

$$\zeta_t^{\mathrm{s}} = H_t \zeta_{t-1}^{\mathrm{s}} c_{t-1}(Y_t), \tag{38}$$

which give the optimal filter, and smoother respectively, equations, for the unnormalized and *unconstrained* estimates. This means that they holds just in the 'interior' of the interval $[0, N]$ (i.e. in $(0, N)$).

In order to recover the normalized estimate, notice that:

$$\begin{aligned}
\sum_{i=1}^{n} \langle \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} Z_t^+ | \mathcal{G}\}, e_i \rangle &= \sum_{i=0}^{n} \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} \langle Z_t^+, e_i \rangle | \mathcal{G}\} \\
&= \overline{\mathbf{E}}\left\{ \overline{\Lambda}_{t+1} \sum_{i=0}^{n} \langle Z_t^+, e_i \rangle | \mathcal{G} \right\} \\
&= \overline{\mathbf{E}}\{\overline{\Lambda}_{t+1} | \mathcal{G}\},
\end{aligned}$$

which shows that the normalizing factor is the sum of the components of the unnormalized estimate:

$$\mathbf{E}\{Z_t^+ | \mathcal{F}_t^Y\} = \frac{\zeta_t}{\sum_{i=1}^n \langle \zeta_t, e_i \rangle}; \qquad \mathbf{E}\{Z_t^+ | \mathcal{F}^Y\} = \frac{\zeta_t^{\mathrm{s}}}{\sum_{i=1}^n \langle \zeta_t^{\mathrm{s}}, e_i \rangle}; \tag{39}$$

Note that, both filter and smoother satisfy *the same* unconstrained equation. Indeed, what makes the difference is taking into account of boundary conditions, which will also provides the actual estimates.

[1]Notice that the inverse of $I + Q_t M_t^{+^T} M_t^{0^{-1}} M_t^+$ is always well defined

Recursive Optimal Fixed-Interval Smoothing for Hidden Reciprocal Processes

Francesco Carravetta
Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti"
Consiglio Nazionale delle Ricerche
Viale Manzoni 30, 00185 Roma, Italy
Email: carravetta@iasi.cnr.it, Telephone: (039) 06–7716410
Fax: (039) 06–7716461
Langford B. White
University of Adelaide
Consiglio Nazionale delle Ricerche
Viale Manzoni 30, 00185 Roma, Italy
Email: carravetta@iasi.cnr.it, Telephone: (039) 06–7716410
Fax: (039) 06–7716461

Abstract

....

I. INTRODUCTION

.... [8] [12], [6], [7] [14], [15] [3],

II. XXXXXXXXX

....

....

III. XXXXXXXXXXXXX

Lemma (your basic lemma got independent of the choice of base elements $e_1, ..., e_n$). *Let $X \in \mathcal{S}$, where $\mathcal{S}$ is any base of $\mathbb{R}^n$, and let V be a vector random variable jointy distributed with X, and suppose $\mathbf{Pr}\{X = e_j\} \neq 0$ for all $j = 0, ..., n-1$, then*

$$\mathbf{E}\{VX^T\} = 0 \quad <=> \quad \mathbf{E}\{V|X\} = 0. \tag{1}$$

Proof. First of all notice that, by definition of conditional expectation:

$$0 = \mathbf{E}\{(V - \mathbf{E}\{V|X\})X^T\} = \mathbf{E}\{VX^T\} + \mathbf{E}\{\mathbf{E}\{V|X\}X^T\}, \tag{2}$$

thus, part $<=$ of the thesis is straightforward. In order to proof part $=>$, let V_i denote the i-th entry of V, so

$$\mathbf{E}\{VX^T\} = 0 \quad => \mathbf{E}\{V_iX\} = 0, \quad \forall i = 0, ..., n-1.$$

Consider

$$0 = \mathbf{E}\{V_iX\} = \sum_{j=0}^{n-1} \mathbf{Pr}\{X = e_j\}\mathbf{E}\{V_i|X = e_j\}e_j, \quad i = 0, ..., n-1 \tag{3}$$

as $e_1, ..., e_n$ are linearly independent, the above identity entails (on account of hypotheses)

$$\mathbf{E}\{V_i|X = e_j\} = 0, \quad i, j = 0, ..., n-1 \tag{4}$$

which in turn implies $\mathbf{E}\{V|X\} = 0$. •

The following lemma establishes a technical result on certain conditional expectations of $Z_t^{\pm[i]}$, for any i positive integer.

Lemma 3

For each $t \in Z_n$ and i positive integer

$$\mathbf{E}\{{Z_{t-1}^{+}}^{[i]}|X_s, s \neq t\} = \mathbf{E}\{{Z_{t-1}^{+}}^{[i]}|X_{t-1}, X_{t+1}\}, \tag{5}$$

$$\mathbf{E}\{{Z_{t+1}^{-}}^{[i]}|X_s, s \neq t\} = \mathbf{E}\{{Z_{t+1}^{-}}^{[i]}|X_{t-1}, X_{t+1}\}, \tag{6}$$

Proof.

Let's perform the calculations for $Z_t^{+[i]}$ (the calculations for $Z_t^{-[i]}$ are similar). One has

$$\begin{aligned}\mathbf{E}\{{Z_{t-1}^{+}}^{[i]}|X_s, s\neq t\} &= X_{t-2}^{[i]}\otimes \mathbf{E}\{X_t^{[i]}|X_s, s\neq t\}\\ &= X_{t-2}^{[i]}\otimes \mathbf{E}\{X_t^{[i]}|X_{t-1}, X_{t+1}\}. \end{aligned} \tag{7}$$

Let ϕ_t an X_{t-1}, X_{t+1}-measurable process:

$$\begin{aligned}&\mathbf{E}\Big\{\Big(X_{t-2}^{[i]}\otimes X_t^{[i]} - X_{t-2}^{[i]}\otimes \mathbf{E}\{X_t^{[i]}|X_{t-1}, X_{t+1}\}\Big)\phi_t^T\Big\}\\ &= \mathbf{E}\Big\{\Big(X_{t-2}^{[i]}\otimes (X_t^{[i]} - \mathbf{E}\{X_t^{[i]}|X_{t-1}, X_{t+1}\})\Big)\phi_t^T\Big\}\\ &= \mathbf{E}\Big\{(X_t^{[i]} - \mathbf{E}\{X_t^{[i]}|X_s, s\neq t\})\gamma(X_{t-2}^{[i]}, \phi_t^T)\Big\} = 0\end{aligned} \tag{8}$$

where $\gamma(\cdot,\cdot)$ is some bilinear map. The above equality being verifyied for any ϕ_t, which is X_{t-1}, X_{t+1}-measurable, and the previously proven property (7), entail

$$\mathbf{E}\{X_{t-2}^{[i]}\otimes X_t^{[i]}|X_s, s\neq t\} = \mathbf{E}\{X_{t-2}^{[i]}\otimes X_t^{[i]}|X_{t-1}, X_{t+1}\}. \tag{9}$$

With similar calculations for $Z_t^{-[i]}$ we obtain

$$\mathbf{E}\{X_t^{[i]}\otimes X_{t+2}^{[i]}|X_s, s\neq t\} = \mathbf{E}\{X_t^{[i]}\otimes X_{t+2}^{[i]}|X_{t-1}, X_{t+1}\}. \tag{10}$$

•

In the following we use the notation $p_t^+(u,j;k,l)$, $p_t^-(u,j;k,l)$ with $0\leq u,j,k,l\leq n$ and $t\in\mathbb{N}$ integers, as for the *four points* transition probability functions of the reciprocal process X, which are defined as follows. Let $e_u, e_j, e_k, e_l\subset\{e_1,...,e_n\}$ then

$$p_t^+(u,j;k,l) = \Pr\{X_{t-2}=e_u, X_t=e_j|X_{t-1}=e_k, X_{t+1}=e_l\}, \tag{11}$$

$$p_t^-(u,j;k,l) = \Pr\{X_{t+2}=e_u, X_t=e_j|X_{t-1}=e_k, X_{t+1}=e_l\}. \tag{12}$$

Theorem 1

For each $t\in Z_n$, and i positive integer, there are $n^{2i}\times n^{2i}$ matrices $G_{i,t}^+$ and $G_{i,t}^-$ such that

$${Z_{t+1}^{-}}^{[i]} = G_{i,t}^- Z_t^{-[i]} + d_{i,t}^- \tag{13}$$

$${Z_{t-1}^{+}}^{[i]} = G_{i,t}^+ Z_t^{+[i]} + d_{i,t}^+ \tag{14}$$

where the residual processes are MA(1), and satisfy $\mathbf{E}\{d_{i,t}^{\pm}|X_s, s \neq t\} = 0$ *for all* $t \in Z_N$. *With* $h_i, s_i = 1, ..., n^{2i}$, *denoting* $\{G_{i,t}^{\pm}\}_{h_i,s_i}$ *the* (h_i, s_i)*-th entry of matrix* $G_{i,t}^{\pm}$ *it is:*

$$\{G_{i,t}^{+}\}_{h_i,s_i} = p_t^{+}(\alpha_1(h), \alpha_2(h); \alpha_1(s), \alpha_2(s)), \tag{15}$$

$$\{G_{i,t}^{-}\}_{h_i,s_i} = p_t^{-}(\alpha_1(h), \alpha_2(h); \alpha_1(s), \alpha_2(s)), \tag{16}$$

where the functions $\alpha_1(\cdot), \alpha_2(\cdot)$ *are defined as*

$$\alpha_1(v) = \left[\frac{v-1}{n^i}\right] + 1; \qquad \alpha_2(v) = |v-1|_{n^i} + 1, \tag{17}$$

v *being an integer index, and* $h, s : 1, ..., n^2$ *defined by*

$$h = \frac{h_i + n^{2(i-1)} - 1}{n^{2(i-1)}}; \qquad s = \frac{s_i + n^{2(i-1)} - 1}{n^{2(i-1)}}. \tag{18}$$

Proof

Let

$$d_{i,t}^{+} = X_{t-2}^{[i]} \otimes X_t^{[i]} - \mathbf{E}\{X_{t-2}^{[i]} \otimes X_t^{[i]}|X_s, s \neq t\}, \tag{19}$$

$$d_{i,t}^{-} = X_{t+2}^{[i]} \otimes X_t^{[i]} - \mathbf{E}\{X_{t+2}^{[i]} \otimes X_t^{[i]}|X_s, s \neq t\}, \tag{20}$$

so that, $\mathbf{E}\{d_{i,t}^{\pm}|X_s, s \neq t\} = 0$ for all $t \in Z_n$. From () and (), noting that $U_{i,t} = X_t^{[i]} - \mathbf{E}\{X_t^{[i]}|X_s, s \neq t\}$ and (7) we have

$$\mathbf{E}\{d_{i,t}^{\pm} {d_{i,s}^{\pm}}^T\} = \mathbf{E}\{(X_{t\mp2}^{[i]} \otimes U_{i,t})({X_{s\mp2}^{[i]}}^T \otimes U_{i,s}^T)\} = \mathbf{E}\{(X_{t\mp2}^{[i]} {X_{s\mp2}^{[i]}}^T) \otimes (U_{i,t}U_{i,s}^T)\}. \tag{21}$$

Let us denote $D_{i,h,l}^U = \mathbf{E}\{U_{i,h}{U_{i,l}}^T\}$. $U_{i,t}$ being some Borel function of X_{t-1}, X_t, X_{t+1}, and also $U_{i,t} \perp X_s\ \forall s \neq t$, entails, for $s = t$ (using (21)):

$$U_{i,t} \perp X_{t\mp2} \Rightarrow \mathbf{E}\{d_{i,t}^{\pm} {d_{i,t}^{\pm}}^T\} = \mathbf{E}\{X_{t\mp2}^{[i]} {X_{t\mp2}^{[i]}}^T\} \otimes D_{i,t}^U. \tag{22}$$

For $s = t-1$:

$$U_{i,t}, U_{i,t-1} \perp X_{t\mp2}, X_{t-1\mp2} \Rightarrow \mathbf{E}\{d_{i,t}^{\pm} {d_{i,t-1}^{\pm}}^T\} = \mathbf{E}\{X_{t-2}^{[i]} {X_{t-1\mp2}^{[i]}}^T\} \otimes D_{i,t,t-1}^U, \tag{23}$$

whilst $s < t-1$:

$$U_{i,t} \perp X_{t\mp2}, X_{s\mp2}, U_{i,s} \Rightarrow \mathbf{E}\{d_{i,t}^{\pm} {d_{i,s}^{\pm}}^T\} = \mathbf{E}\{(X_{t\mp2}^{[i]} {X_{s\mp2}^{[i]}}^T) \otimes (\mathbf{E}\{U_{i,t}\}U_{i,s}^T)\} = 0. \tag{24}$$

For $s = t+1$

$$U_{i,t}, U_{i,t+1} \perp X_{t\mp2}, X_{t+1\mp2} \Rightarrow \mathbf{E}\{d_{i,t}^{\pm} {d_{i,t+1}^{\pm}}^T\} = \mathbf{E}\{X_{t-2}^{[i]} X_{t+1\mp2}^T\} \otimes D_{i,t,t+1}^U. \tag{25}$$

and finally for $s > t+1$:

$$U_{i,s} \perp X_{t\mp 2}, X_{s\mp 2}, U_{i,t} \Rightarrow \mathbf{E}\{d^{\pm}_{i,t} {d^{\pm}_{i,s}}^T\} = \mathbf{E}\{(X^{[i]}_{t\mp 2} {X^{[i]}_{s\mp 2}}^T) \otimes (U_{i,t}\mathbf{E}\{U^T_{i,s}\})\} = 0, \tag{26}$$

Thus $d^{+}_{i,\cdot}$ and $d^{-}_{i,\cdot}$ are a one-step correlated (MA(1)) noises, with covariances which can be computed by the covariances of $X^{[i]}_t$ and of $U_{i,t}$. Now consider the term

$$\mathbf{E}\{{Z^{+}_{t-1}}^{[i]} | X_{t-1}, X_{t+1}\} = \mathbf{E}\{X^{[i]}_{t-2} \otimes X^{[i]}_t | X_{t-1}, X_{t+1}\}, \tag{27}$$

which, for $i = 1$, are the n^2 values of the joint conditional probabilities of X_{t-2} and X_t given the n^2 values taken by X_{t-1} and X_{t+1}, or equivalently, the n^2 values taken by Z^{+}_t, i.e. $e'_1, ..., e'_{n^2}$ (e'_h being the h-th canonical vector of $\mathrm{I\!R}^{n^2}$). These conditional probabilities can be appropriately arranged in an $n^2 \times n^2$ matrix $G^{+}_{1,t}$ so that

$$\mathbf{E}\{Z^{+}_{t-1} | Z^{+}_t\} = G^{+}_{1,t} Z^{+}_t. \tag{28}$$

Matrix $G^{+}_{1,t}$ can be readily calculated by observing that, for $h = 1, ..., n^2$:

$$\mathbf{E}\{\langle X_{t-2} \otimes X_t, e'_h\rangle | X_{t-1} = e_k, X_{t+1} = e_r\} = \Pr\{\langle X_{t-2} \otimes X_t, e'_h\rangle = 1 | X_{t-1} \otimes X_{t+1} = e_k \otimes e_r\}. \tag{29}$$

Let s denote the position of the 1-entry in the vector $e_k \otimes e_r$, whose expression is:

$$s = (k-1)n + r, \tag{30}$$

and note that the function $(k, r) \to s$ is invertible, with inverse given by:

$$k = \left[\frac{s-1}{n}\right] + 1; \qquad r = |s-1|_n + 1, \tag{31}$$

where $[\cdot]$ and $|\cdot|_n$ denote integer part, and modulo (n), respectively. Thus, α_1, α_2 denoting the inverse maps $s \to k$, and $s \to r$, in (31) respectively, eqs. (28), (29) are ensued by

$$\{G^{+}_{1,t}\}_{h,s} = p^{+}(u, j; k, r), \tag{32}$$

with $u = \alpha_1(h), j = \alpha_2(h)$, $k = \alpha_1(s), r = \alpha_2(s)$, that is eq (15) for $i = 1$.

For $i > 1$, one has ${Z^{+}_{t-1}}^{[i]} \in \{{e'_1}^{[i]}, ..., {e'_{n^2}}^{[i]}\}$, a subset of n^2 elements in the set of n^{2i} canonical base vectors of $\mathrm{I\!R}^{2i}$, and as to the probabilities it is

$$\Pr\{{Z^{+}_{t-1}}^{[i]} = {e'_h}^{[i]} | X_{t-1}, X_{t+1}\} = \Pr\{|X_{t-2} \otimes X_t = {e'_h}^{[i]} | X_{t-1}, X_{t+1}\}, \quad j = 1, ..., n \tag{33}$$

that is: the *same* $n^2 \times n^2$ values of the joint conditional probabilities only just considered for $i = 1$ are to be considered for $i > 1$ as well, just they have to be arranged appropriately in

$n^{2i} \times n^{2i}$ matrices $G^{+}_{i,t}$ [1]. In order to do this, first recall that X_{t-1}, X_{t+1} generate the same σ-algebra as $X^{[i]}_{t-1} \otimes X^{[i]}_{t+1}$ for any integer i, thus in place of eq. (29) we can write

$$\begin{aligned}
&\mathbf{E}\{\langle X^{[i]}_{t-2} \otimes X^{[i]}_{t}, {e'_h}^{[i]}\rangle | X_{t-1} = e_k, X_{t+1} = e_r\} \\
&= \Pr\{\langle X^{[i]}_{t-2} \otimes X^{[i]}_{t}, {e'_h}^{[i]}\rangle = 1 | X^{[i]}_{t-1} \otimes X^{[i]}_{t+1} = e^{[i]}_k \otimes e^{[i]}_r\} \\
&= \Pr\{\langle {Z^{+}_{t-1}}^{[i]}, {e'_h}^{[i]}\rangle = 1 | {Z^{+}_{t}}^{[i]} = {e'_s}^{[i]}\} \\
&= \{G^{+}_{i,t}\}_{h_i, s_i} \\
&= \Pr\{\langle X_{t-2} \otimes X_{t}, e'_h\rangle = 1 | X_{t-1} \otimes X_{t+1} = e_k \otimes e_r\} \\
&= \Pr\{\langle Z^{+}_{t-1}, e'_h\rangle = 1 | Z^{+}_{t} = e'_s\} \\
&= \{G^{+}_{1,t}\}_{h,s},
\end{aligned} \tag{34}$$

where $h_i, s_i : 1, ..., n^{2i}$, and $h, s : 1, ..., n^2$ denote the position of the 1-entry in ${e'_h}^{[i]}, {e'_s}^{[i]}$, and e'_h, e'_s respectively. It's easy to verify that

$$h_i = n^{2(i-1)}(h-1) + 1; \quad s_i = n^{2(i-1)}(s-1) + 1, \tag{35}$$

thus, (34) entails formula (15) for $i > 0$.

As to Z^{-}, following the same lines as for Z^{+} we get

$$\mathbf{E}\{{Z^{-}_{t+1}}^{[i]} | Z^{-}_{t}\} = G^{-}_{i,t} {Z^{-}_{t}}^{[i]}, \tag{36}$$

which holds for any $i > 0$, after which (13) follows from (19). Thus representation (14) is derived, and matrix $G^{-}_{i,t}$ defined as well, by means of formula (16). •

We now have our main result for this section which is a self-adjoint reciprocal model for any of the *centred* processes $Z_{i,t} = {Z^{+}_{t}}^{[i]} - \mathbf{E}\{{Z^{+}_{t}}^{[i]}\}$. We require the use of the centred process when dealing with processes indexed on the circle, because such processes have zero mean, and any canonical random variable cannot have zero mean, as its mean is a probability mass function.

The following matrices will be used: given two vectors $a, b \in \mathbb{R}^n$, for any positive integer i it's well known (see [references]) the existence of a matrix $K_{n^{2i}} \in \mathbb{R}^{n^{2i} \times n^{2i}}$ such that $a^{[i]} \otimes b^{[i]} = K_{n^{2i}}(b^{[i]} \otimes a^{[i]})$. Such matrix $K_{n^{2i}}$ is named *commutation matrix*, and is an invertible $0,1$-matrix with the property: ${K_{n^{2i}}}^{-1} = K_{n^{2i}}$ (which follows immediately by definition).

[1] Notice that, the $G^{-}_{i,t}$ have many repeated entries for $i > 1$, unlike $G^{-}_{1,t}$

Theorem 2

For any positive integer i, let $Z_{i,t} = Z_t^{+[i]} - \mathbf{E}\{Z_t^{+[i]}\}$, then for each $t \in \mathbb{Z}_n$, defining the $n^{2i} \times n^{2i}$ matrices $M_{i,t}^0, M_{i,t}^+, M_{i,t}^-$: [2]

$$M_{i,t}^0 = (D_{i,t+1}^{1/2})^\dagger, \tag{37}$$

$$M_{i,t}^+ = -\frac{1}{2}(D_{i,t+1}^{1/2})^\dagger G_{i,t+1}^+, \tag{38}$$

$$M_{i,t}^- = -\frac{1}{2}(D_{i,t+1}^{1/2})^\dagger K_{n^{2i}} G_{i,t-1}^- K_{n^{2i}}, \tag{39}$$

with $D_{i,t} = \mathbf{E}\{d_{i,t}^+ {d_{i,t}^+}^T\}$, it is

$$M_{i,t}^0 Z_{i,t} + M_{i,t}^+ Z_{i,t+1} + M_{i,t}^- Z_{i,t-1} = e_{i,t}, \tag{40}$$

$${M_{i,t}^0}^T = M_{i,t}^0 \geq 0; \quad M_{i,t}^+ = {M_{i,t+1}^-}^T \tag{41}$$

where $e_{i,t}$ is a MA(1) process satisfying

$$\mathbf{E}\{e_{i,t} Z_{i,s}^T\} = \delta_{s,t} I_{n^{2i}}. \tag{42}$$

whose one-step-correlation is given by

$$\mathbf{E}\{e_{i,t} {e_{i,t+1}}^T\} = (D_{i,t+1}^{1/2})^\dagger \left(\mathbf{E}\{X_{t-1}^{[i]} {X_t^{[i]}}^T\} \otimes D_{i,t+1,t+2}^U \right) (D_{i,t+2}^{1/2})^\dagger \tag{43}$$

Proof. Consider

$$\begin{aligned} d_{i,t+1}^+ &= X_{t-1}^{[i]} \otimes X_{t+1}^{[i]} - \mathbf{E}\{X_{t-1}^{[i]} \otimes X_{t+1}^{[i]} | X_s, s \neq t\} \\ &= K_{n^{2i}}(X_{t+1}^{[i]} \otimes X_{t-1}^{[i]} - \mathbf{E}\{X_{t+1}^{[i]} \otimes X_{t-1}^{[i]} | X_s, s \neq t\}) \\ &= K_{n^{2i}} d_{i,t-1}^-. \end{aligned} \tag{44}$$

Taking equation (13), substituting $t-1$ for t, multiplying by the commutation matrix $K_{n^{2i}}$ gives

$$\begin{aligned} K_{n^{2i}} {Z_t^-}^{[i]} &= K_{n^{2i}} G_{t-1}^- {Z_{t-1}^-}^{[i]} + K_{n^{2i}} d_{i,t-1}^-, \quad \Leftrightarrow \\ {Z_t^+}^{[i]} &= K_{n^{2i}} G_{t-1}^- K_{n^{2i}} {Z_{t-1}^+}^{[i]} + d_{t+1}^+, \end{aligned} \tag{45}$$

[2] $A^{1/2}$ denotes a matrix square-root of a symmetric matrix A, $A^\dagger$ a left pseudo-inverse of it.

noting that

$$K_{n^{2i}} {Z_{t-1}^{-}}^{[i]} = X_t^{[i]} \otimes X_{t-2}^{[i]} = {Z_{t-1}^{-}}^{[i]}. \tag{46}$$

Now take equation (14), replace t by $t+1$, add to (45) and divide by two to give

$$Z_t^{+[i]} = \left(G_{i,t+1}^{+} {Z_{t+1}^{+}}^{[i]} + K_{n^{2i}} G_{t-1}^{-} K_{n^{2i}} {Z_{t-1}^{+}}^{[i]} \right)/2 + d_{i,t+1}^{+}, \tag{47}$$

which, on account of (37)-(39), turns into the *normalized reciprocal representation*:

$$M_{i,t}^{0} Z_t^{+[i]} + M_{i,t}^{+} {Z_{t+1}^{+}}^{[i]} + M_{i,t}^{-} {Z_{t-1}^{+}}^{[i]} = e_{i,t}, \tag{48}$$

with

$$e_{i,t} = (D_{i,t+1}^{1/2})^{\dagger} d_{i,t+1}^{+}. \tag{49}$$

Also, expression (43) is straightforward from (25). Taking expectations in (48) gives

$$M_{i,t}^{0} \mathbf{E}\{Z_t^{+[i]}\} + M_{i,t}^{+} \mathbf{E}\{{Z_{t+1}^{+}}^{[i]}\} + M_{i,t}^{-} \mathbf{E}\{{Z_{i,t-1}^{+}}^{[i]}\} = 0, \tag{50}$$

so, eq. (40) comes from subtracting (50) to (48). Only rests to prove the second identity in (41). For, write eq (13) and apply (45), (46) to give

$$\begin{aligned} {Z_{t+1}^{-}}^{[i]} &= G_{i,t}^{-} Z_t^{-[i]} + d_{i,t}^{-} = G_{i,t}^{-} K_{n^{2i}} Z_t^{+[i]} + K_{n^{2i}} d_{i,t+2}^{+}, \quad \Rightarrow \\ 0 = \mathbf{E}\{{Z_{t+1}^{+}}^{[i]} e_{i,t}^{T}\} &= K_{n^{2i}} G_{i,t}^{-} K_{n^{2i}} \mathbf{E}\{Z_t^{+[i]} e_{i,t}^{T}\} + \mathbf{E}\{d_{i,t+2}^{+} e_{i,t}^{T}\}. \end{aligned} \tag{51}$$

Also, take eq. (14), replace t by $t+1$, and take expectations upon multiplying by $e_{i,t+1}^{T}$ to give

$$0 = \mathbf{E}\{Z_t^{+[i]} e_{i,t+1}^{T}\} = G_{i,t+1}^{+} \mathbf{E}\{{Z_{t+1}^{+}}^{[i]} e_{i,t+1}^{T}\} + \mathbf{E}\{d_{i,t+1}^{+} e_{i,t+1}^{T}\}. \tag{52}$$

Using property (42), premultiplying (52) by $(D_{i,t+2}^{1/2})^{\dagger}$, and (51) by $(D_{i,t+1}^{1/2})^{\dagger}$:

$$-(D_{i,t+1}^{1/2})^{\dagger} \mathbf{E}\{d_{i,t+2}^{+} e_{i,t}^{T}\} = (D_{i,t+1}^{1/2})^{\dagger} K_{n^{2i}} G_{i,t}^{-} K_{n^{2i}}, \tag{53}$$

$$-(D_{i,t+2}^{1/2})^{\dagger} \mathbf{E}\{d_{i,t+1}^{+} e_{i,t+1}^{T}\} = (D_{i,t+2}^{1/2})^{\dagger} G_{i,t+1}^{+}. \tag{54}$$

By definition, given in (49), one has

$$\begin{aligned} \mathbf{E}\{d_{i,t+1}^{+} e_{i,t+1}^{T}\} &= \mathbf{E}\{d_{i,t+1}^{+} {d_{i,t+2}^{+}}^{T}\} (D_{i,t+2}^{1/2})^{\dagger}, \\ \mathbf{E}\{d_{i,t+2}^{+} e_{i,t}^{T}\} &= \mathbf{E}\{d_{i,t+2}^{+} {d_{i,t+1}^{+}}^{T}\} (D_{i,t+1}^{1/2})^{\dagger}, \end{aligned}$$

which used in (53) and (54) yield immediately the desired result. •

Note, however, that $\{Z_t\}$ is base-valued in $\mathbb{R}^{n^2}$ but no more canonical-base-valued, as it is for $\{Z_t^+\}$.

Now, suppose the reciprocal chain X_t living on a *discrete circle*, i.e. $t \in [0, N] \subset \mathbb{N}$, where the aritmetics has to be intended modulo N. On the discrete circle the advantage is that we are able to get well-posed a reciprocal representation as (**??**) even without boundary condition. This means that, first, the following *global* representation can be obtained by assembling (**??**) for $t = 0, ..., N$:

$$MZ = e, \tag{55}$$

$$M = \begin{bmatrix} M_0^0 & M_0^+ & 0 & \dots & M_0^- \\ M_1^- & M_1^0 & M_1^+ & 0 & \dots \\ 0 & M_2^- & M_2^0 & M_2^+ & \dots \\ & & \ddots & & \\ M_N^+ & 0 & \dots & M_N^- & M_N^0 \end{bmatrix}, \tag{56}$$

$$Z = \mathrm{col}(Z_0, ..., Z_N); \qquad e = \mathrm{col}(e_0, ..., e_N), \tag{57}$$

second, by supposing $\mathbf{E}\{ZZ^T\} > 0$ (non-singularity assumption), since by (42) it is $\mathbf{E}\{eZ^T\} = I$, one has [3]

$$M = (\mathbf{E}\{ZZ^T\})^{-1}, \tag{58}$$

and in particular, on account of (**??**), $M = M^T > 0$. Thus $Z = M^{-1}e$, that is: the reciprocal model (**??**) on the discrete circle admits, samplewise, an unique solution[4] given the 'input noise' e (well-posedness on the circle). Also, it follows that on the discrete circle it is $\mathbf{E}\{Z_t\} = 0$, $t = 0, ..., N$. Thus, it should be noticed that Z^+ the *un-centered* process, (which is canonical-base-valued, hence it has a non-zero expectation) *cannot be* a solution of the reciprocal model on the circle.

[3]Notice that $\mathbf{E}\{ZZ^T\}$ can be calculated from the two points joint probabilities of Z (and, from four-points joint probabilities of X), all joint-probabilities being in turn computable by some given statistics by following the lines depicted in previous chapters. Then, eq. (58) yields a way to compute all the needed coefficients of the reciprocal representation.

[4]Also notice that the unicity is guaranteed by the non-singularity assumption. If this is not the case, the reciprocal representation holds as well, but it is not unique.

Now, suppose the following observation equation is given:

$$Y_t = C_t X_t + W_t. \tag{59}$$

with W_t white noise $\mathbf{E}\{W_t\} = 0$, $\mathbf{E}\{W_t W_t^T\} = R_t > 0$. One has

$$Y_{t-2} \otimes Y_t = (C_{t-2} \otimes C_t) Z_t + W_{t-2} \otimes W_t + (C_{t-2} X_{t-1}) \otimes W_t + W_{t-2} \otimes (C_t X_t). \tag{60}$$

so

$$\mathcal{Y}_t = \mathcal{C}_t Z_t + \mathcal{W}_t, \tag{61}$$

with

$$\mathcal{Y}_t = Y_{t-2} \otimes Y_t, \qquad \mathcal{C}_t = C_{t-2} \otimes C_t, \tag{62}$$

$$\mathcal{W}_t = W_{t-2} \otimes W_t + (C_{t-2} X_{t-2}) \otimes W_t + W_{t-2} \otimes (C_t X_t). \tag{63}$$

Note that the process $\mathcal{W}_t$ is two-step correlated, $\mathbf{E}\{\mathcal{W}_t\} = 0$ (hence, also $\mathbf{E}\{\mathcal{Y}_t\} = 0$), and $\mathbf{E}\{\mathcal{W}_t \mathcal{W}_t\} = \mathcal{Q}_t$, with $\mathcal{Q}_t$ given by

....

Now, define the *global observation* $\mathcal{Y}$:

$$\mathcal{Y} = \mathrm{col}(\mathcal{Y}_0, \mathcal{Y}_1, ..., \mathcal{Y}_N), \tag{64}$$

and the two sub-vectors $\mathcal{Y}^o$, and $\mathcal{Y}^e$, collecting the measurement $\mathcal{Y}_t$ for t odd, and even respectively (suppose wlog N odd):

$$\mathcal{Y}^o = \mathrm{col}(\mathcal{Y}_2, \mathcal{Y}_3, ..., \mathcal{Y}_{N-1}), \tag{65}$$

$$\mathcal{Y}^e = \mathrm{col}(\mathcal{Y}_0, \mathcal{Y}_1, ..., \mathcal{Y}_{N-3}). \tag{66}$$

Thus

$$\mathcal{Y}^o = \mathcal{C}^o Z^o + \mathcal{W}^o, \tag{67}$$

$$\mathcal{Y}^e = \mathcal{C}^e Z^e + \mathcal{W}^e, \tag{68}$$

where

$$\mathcal{W}^o = \mathrm{col}(\mathcal{W}_2, \mathcal{W}_3, ..., \mathcal{W}_{N-1}), \tag{69}$$

$$\mathcal{W}^d = \mathrm{col}(\mathcal{W}_0, \mathcal{W}_1, ..., \mathcal{W}_{N-3}), \tag{70}$$

$$Z^o = \mathrm{col}(Z_2, Z_3, ..., Z_{N-1}), \tag{71}$$

$$Z^e = \mathrm{col}(Z_0, Z_1, ..., Z_{N-3}), \tag{72}$$

$$\mathcal{C}^o = \mathrm{diag}(\mathcal{C}_2, \mathcal{C}_3, ..., \mathcal{C}_{N-1}), \tag{73}$$

$$\mathcal{C}^e = \mathrm{diag}(\mathcal{C}_0, \mathcal{C}_1, ..., \mathcal{C}_{N-3}). \tag{74}$$

We have $\mathbf{E}\{\mathcal{W}_t\} = 0$, $\forall t$, and from eqs. (67), (68) we deduce $\mathbf{E}\{\mathcal{Y}_t\} = 0$, $\forall t$. Moreover, since $\mathcal{W}_t$ is a one-step correlated process, we have for the covariances $R_{\mathcal{W}^\alpha \mathcal{W}^\alpha}$, $\alpha = o, e$, (merely a lexicographic substitution)

$$R_{\mathcal{W}^o \mathcal{W}^o} = \mathrm{diag}(\mathcal{Q}_2, \mathcal{Q}_3, ..., \mathcal{Q}_{N-1}), \tag{75}$$

$$R_{\mathcal{W}^e \mathcal{W}^e} = \mathrm{diag}(\mathcal{Q}_0, \mathcal{Q}_1, ..., \mathcal{Q}_{N-3}). \tag{76}$$

Thus, we have by the linear estimation formulas:

$$\widehat{Z^o} = R_{Z^o Z^o} \mathcal{C}^{oT} R^{-1}_{\mathcal{W}^o \mathcal{W}^o} \left(\mathcal{Y}^o - \mathcal{C}^o \widehat{Z^o} \right) \tag{77}$$

$$\widehat{Z^e} = R_{Z^e Z^e} \mathcal{C}^{eT} R^{-1}_{\mathcal{W}^e \mathcal{W}^e} \left(\mathcal{Y}^e - \mathcal{C}^e \widehat{Z^e} \right) \tag{78}$$

and aggregating

$$\widehat{\widetilde{Z}} = R_{\widetilde{Z}\widetilde{Z}^T} \widetilde{\mathcal{C}}^T R^{\dagger}_{\widetilde{\mathcal{W}}\widetilde{\mathcal{W}}} \left(\widetilde{\mathcal{Y}} - \widetilde{\mathcal{C}} \widehat{\widetilde{Z}} \right) \tag{79}$$

where $\widetilde{Z} = \mathrm{col}(Z^o, Z^e)$, $\widetilde{\mathcal{Y}} = \mathrm{col}(\mathcal{Y}^o, \mathcal{Y}^e)$, $R^{\dagger}_{\mathcal{W}\mathcal{W}} = \mathrm{diag}(R^{-1}_{\mathcal{W}^o \mathcal{W}^o}, R^{-1}_{\mathcal{W}^e \mathcal{W}^e})$, $\widetilde{\mathcal{C}} = \mathrm{diag}(\mathcal{C}^o, \mathcal{C}^e)$. Define $Z = \mathrm{col}(Z_0, Z_1, ..., Z_N)$, and let Γ be the swapping-rows matrix such that $Z = \Gamma \widetilde{Z}$. We have that Γ is an orthogonal matrix $\Gamma^{-1} = \Gamma^T$. Also, by post-multiplying it, a corresponding swapping of columns gets performed, so $R_{ZZ} = \Gamma R_{\widetilde{Z}\widetilde{Z}} \Gamma^T$. By using Γ eq. (79) turns in

$$M\widehat{Z} = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}} \left(\mathcal{Y} - \mathcal{C}\widehat{Z} \right) \tag{80}$$

where $\mathcal{C} = \mathrm{diag}(\mathcal{C}_0, \mathcal{C}_1, ..., \mathcal{C}_N)$, $R^{\dagger}_{\mathcal{W}\mathcal{W}} = \mathrm{diag}(\mathcal{Q}_0^{-1}, \mathcal{Q}_1^{-1}, ..., \mathcal{Q}_N^{-1})$, $\mathcal{Y} = \mathrm{col}(\mathcal{Y}_0, \mathcal{Y}_1, ..., \mathcal{Y}_N)$, and $M = R^{-1}_{\mathcal{Z},\mathcal{Z}}$, given by (56).

Now, denoting $\mathcal{V}(\mathrm{I\!R})$ the space or scalar-valued random variables, for η a random vector denote

$$\mathcal{L}(\eta) = \{\xi \in \mathcal{V}(\mathrm{I\!R}) : \xi = V^T\eta, \text{for some vector } V\},$$

the space linearly spanned by η. We have the following Lemmas:

Lemma 2. Let $\mathcal{Y} = Y_0 \otimes \cdots \otimes Y_N$, $Y = \mathrm{col}(Y_0, .., Y_N)$, $Y_i \in \{e_1, ..., e_m\}$ the canonical base of $\mathrm{I\!R}^m$. Then, $\mathcal{Y}^{[2]}$ is a quadratic function of Y, e.g. there exists a matrix $\mathcal{Q} \in \mathrm{I\!R}^{m^{2(N+1)} \times m^2(N+1)^2}$ such that $\mathcal{Y}^{[2]} = \mathcal{Q}Y^{[2]}$.

Proof. Let $\{\phi_1, ..., \phi_{m^{N+1}}\} \subset \mathrm{I\!R}^{m(N+1)}$ be the set of values of Y, and $\{e'_1, ..., e'_{m^{N+1}}\} \subset \mathrm{I\!R}^{m^{N+1}}$ be the set of values of $\mathcal{Y}$. Define the vectors $\{\widetilde{\phi}_1, ..., \widetilde{\phi}_{m^{N+1}}\} \subset \mathrm{I\!R}^{m^{N+1} \cdot m(N+1)}$ as

$$\widetilde{\phi}_i = \mathrm{col}(0, ..., 0, \phi_i, 0, ..., 0), \quad i = 1, ..., m^{N+1} \tag{81}$$

where each '0' is a vector of m^{N+1} zeroes, and ϕ_i lies at the i-th macro-position. Let us prove that, for any $i = 1, ..., m^{N+1}$ a matrix $\Psi_i \in \mathrm{I\!R}^{m(N+1) \times (m^{N+1} \cdot m(N+1))}$ there exists, such that

$$\Psi_i \widetilde{\phi}_j = e'_i \delta_{i,j}, \quad j = 1, ..., m^{N+1}. \tag{82}$$

For, let us solve the system of matrix equations (82). Stacking it for each i yields

$$\Psi_i \widetilde{\Phi} = \begin{bmatrix} 0 & \ldots & 0 & e'_i & 0 & \ldots & 0 \end{bmatrix},$$

with $\widetilde{\Phi} = [\widetilde{\phi}_1, ..., \widetilde{\phi}_{m^{N+1}}]$, which is equivalent to

$$(\widetilde{\Phi}^T \otimes I_{m^{N+1}})\mathrm{Vec}(\Psi_i) = e'_i \otimes e'_i.$$

Let's further stack each of the equations above to give

$$(\widetilde{\Phi}^T \otimes I_{m^{N+1}})Q = \underline{e}^T, \tag{83}$$

with $Q = [\mathrm{Vec}(\Psi_1), ..., \mathrm{Vec}(\Psi_{m^{N+1}})]$, and $\underline{e}^T = [{e'_1}^{[2]}, ..., {e'_{m^{N+1}}}^{[2]}]$. Finally from (83) one derives:

$$(I_{m^{N+1}} \otimes \widetilde{\Phi}^T \otimes I_{m^{N+1}})\mathrm{Vec}(Q) = \underline{e}, \tag{84}$$

which can be solved indeed giving $\mathrm{Vec}(Q)$ as a solution, for $\widetilde{\Phi}$ is full-rank by construction, so $I_{m^{N+1}} \otimes \widetilde{\Phi}^T \otimes I_{m^{N+1}}$ is full-rank as well.

Now, define $\widetilde{Y}$ as the random variable which takes on the values $\{\widetilde{\phi}_1, ..., \widetilde{\phi}_{m^{N+1}}\}$[5]. Also, let's define $\widetilde{\mathcal{Y}} = \mathrm{col}(\widetilde{Y}, ..., \widetilde{Y})$ (m^{N+1} times). Using eqs. (82) gives

$$\mathcal{Y}^{[2]} = \mathbf{K}\widetilde{\mathcal{Y}}, \tag{85}$$

where $\mathbf{K} = \mathrm{diag}(\Psi_1, ..., \Psi_{m^{N+1}})$. Moreover it is obviously defined a matrix $\mathbf{K}_2$ such that

$$\widetilde{\mathcal{Y}} = \mathbf{K}_2\widetilde{Y}. \tag{86}$$

By (85) and (86), the Lemma is proven as soon as one shows that there exists a matrix $\mathbf{K}_3$ such that

$$\widetilde{Y} = \mathbf{K}_3 Y^{[2]}. \tag{87}$$

For, note that the set of Y-values $\{\phi_1, ..., \phi_{m^{N+1}}\}$ have the following property ${\phi_i}^T\phi_j = (N+1)$, if and only if $i = j$, otherwise the scalar product is less than $N+1$ (and ≥ 0). Thus

$$\widetilde{Y} = \frac{1}{N+1}\varphi(Y)Y, \tag{88}$$

where $\varphi(Y)$ is the matrix-valued function defined as

$$\varphi(Y) = \left[\, I_{m(N+1)} \cdot \gamma({\phi_1}^T Y) \quad \ldots \quad I_{m(N+1)} \cdot \gamma({\phi_{m^{N+1}}}^T Y) \,\right]^T, \tag{89}$$

with $\gamma : \mathrm{I\!R} \to \mathrm{I\!R}$ defined as

$$\gamma(x) = \begin{cases} 1 & \text{if} \quad x = N+1; \\ 0 & \text{otherwise}; \end{cases} \tag{90}$$

Identity (88) can be directly verifyied by substituting the Y-values ϕ_i, $i = 1, ..., m^{N+1}$. It shows that the map $Y \to \widetilde{Y}$ is quadratic, which entails the existence of a suitable $\mathbf{K_3}$ in (87). •

Lemma 3 *Let $\widehat{Z}$ the optimal linear estimate of Z with respect to $\mathcal{Y}^{[2]}$. It is $\widehat{Z} = \mathbf{E}\{Z|Y\}$.*

Proof. The random vector $\mathcal{Y}^{[2]}$ takes values on the set of m^{N+1} linearly independent vectors $\{{e'_1}^{[2]}, ..., {e'_{m^{N+1}}}^{[2]}\}$, so by Lemma 1: $\Pi(Z|\mathcal{Y}^{[2]}) = \mathbf{E}\{Z|Y\}$ [6]. Moreover by Lemma 2, such optimal estimate is a quadratic function of Y. Let $\mathcal{Q}(Y)$ the space of all quadratic functions of Y, the projection $\Pi(Z|\mathcal{Q}(Y))$ cannot give, obviously, an error-variance lower than the optimal.

[5] Note that is this a linearly independent set of vectors in $\mathrm{I\!R}^{m^{N+1}}$, unlike $\{\phi_1, ..., \phi_{m^{N+1}}\} \subset \mathrm{I\!R}^{m(N+1)}$ which are not linearly independent in the space they belong to.

[6] Indeed, $\mathcal{Y}^{[2]}$ and Y generate the same σ-algebra

Moreover, it cannot give it greater as well, as we know the optimal estimate is a quadratic function. Thus $\Pi(Z|\mathcal{Q}(Y)) = \mathbf{E}(Z|Y)$. •

Thus, on account of Lemma 2., eq. (80) defines (locally) a reciprocal equation for the optimal smoothing estimate driven by the innovation process. This can be solved with the usual Levy's double swapping method giving two processes running in opposite directions.

The Levy's double-swapping procedure.

Given three finite sequences, A, B, C, of $\nu+1$ matrices in $\mathbb{R}^{n\times n}$ each: $\{L(0), ..., L(\nu)\}$, $L = A, B, C$, we define the following $(\nu+1)\times(\nu+1)$-blocks *circulant* matrix:

$$\mathcal{F}_\nu(A,B,C) = \begin{bmatrix} B(0) & C(0) & 0 & \dots & A(0) \\ A(1) & B(1) & C(1) & 0 & \dots \\ 0 & A(2) & B(2) & C(2) & \dots \\ & & \ddots & & \\ C(\nu) & 0 & \dots & A(\nu) & B(\nu) \end{bmatrix}. \tag{91}$$

If A, B, C are matrix sequences such that $C(k) = A^T(k+1)$, $B(k) = B^T(k)$, for $k = 0...\bar{\nu}$, $\bar{\nu} = \nu - 1$, then the following $(\nu-1)\times(\nu+1)$-blocks three-banded matrix

$$\Phi_\nu(A,B,C) = \begin{bmatrix} A(1) & B(1) & C(1) & 0 & \dots \\ 0 & A(2) & B(2) & C(2) & \dots \\ & & \ddots & & \\ & & A(\bar{\nu}) & B(\bar{\nu}) & C(\bar{\nu}) \end{bmatrix}. \tag{92}$$

admits the following decomposition:

$$\Phi_\nu(A,B,C) = L(C)\mathbf{T}^{-1}H(C), \quad \mathbf{T} = \mathrm{diag}\{T(i)\}_{i=0,\dots,\bar{\nu}} \tag{93}$$

$$L(C) = \begin{bmatrix} I & -(CT)(1) & 0 & \dots & \\ 0 & I & -(CT)(2) & & \\ & & \ddots & & \\ 0 & \dots & & I & -(CT)(\bar{\nu}) \end{bmatrix} \quad H(C) = \begin{bmatrix} -(TC^T)(0) & I & 0 & \dots \\ 0 & -(TC^T)(1) & I & \\ & & \ddots & \\ 0 & \dots & -(TC^T)(\bar{\nu}) & I \end{bmatrix} \tag{94}$$

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:

$$T^{-1}(i-1) = B(i) - C(i)T(i)C^T(i), \quad T(\nu) = 0. \tag{95}$$

Now, eq. (80) can be rewritten:

$$\left(M + \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{C}\right) \widehat{Z} = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \tag{96}$$

and as $\mathcal{C}, R^{\dagger}_{\mathcal{W}\mathcal{W}}$ are the diagonal matrices defined above, by using notation (91) it is

$$M + \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{C} = \mathcal{F}_N(M^-, M^0 + \mathcal{C}^T \mathcal{Q}^{-1}\mathcal{C}, M^+), \tag{97}$$

where $\mathcal{C}^T \mathcal{Q}^{-1}\mathcal{C}$ stands for the sequence $\{\mathcal{C}_t^T \mathcal{Q}_t^{-1}\mathcal{C}_t\}$, and $M^{\pm}, M^0$ stand for the sequences $\{M_t^{\pm}\}, \{M_t^0\}$, respectively. Also, the following decomposition holds

$$\Phi_N(M^-, M^0 + \mathcal{C}^T \mathcal{Q}^{-1}\mathcal{C}, M^+) = L(M^+)\mathbf{\Lambda^{-1}} H(M^+), \tag{98}$$

where $\mathbf{\Lambda} = \mathrm{diag}\{\Lambda_0, ..., \Lambda_{N-2}\}$ and the matrix $\Lambda_t > 0$ for $t = 0, ..., N-1$, satisfies the backward recursive equation:

$$\Lambda_{t-1}^{-1} = M_t^0 + \mathcal{C}_t^T \mathcal{Q}_t^{-1}\mathcal{C}_t - M_t^+ \Lambda_t M_t^{+T}, \quad \Lambda_{N-1} = 0. \tag{99}$$

Now, eq. (96), on account of (98), implies the following

$$L(M^+)\mathbf{\Lambda}^{-1} H(M^+)\widehat{Z} = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \tag{100}$$

which is equivalent to the following pair

$$L(M^+)\zeta = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \tag{101}$$

$$\mathbf{\Lambda}^{-1} H(M^+)\widehat{Z} = \zeta. \tag{102}$$

By the structure of matrices $L(), H()$ we realize that the global eqs. (101), (102), have as a local counterpart the following couple of *backward* and *forward* equations

$$\zeta_t = M_{t+1}^+ \Lambda_{t+1}\zeta_{t+1} + \mathcal{C}_t^T \mathcal{Q}_t^{-1}\mathcal{Y}_t, \tag{103}$$

$$\widehat{Z}_{t+1} = \mathcal{C}_t^T \widehat{Z}_t + \zeta_t. \tag{104}$$

Now, by defining $\Gamma = \left(M + \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{C}\right)^{-1}$ [7], eq. (96) rewrites $\widehat{Z} = \Gamma \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y}$, from which:

$$\widehat{Z}_t = \sum_{s=0}^{N} \Gamma(t,s)\mathcal{C}_s^T \mathcal{Q}_s^{-1}\mathcal{Y}_s, \tag{105}$$

[7] The inverse is well defined, as $M > 0$

where the Kernel function $\Gamma(\cdot,\cdot)$ gets readily defined by a suitable partitioning of the matrix Γ. By setting $\zeta_{N-1}=0$ we recursively get a solution, say $\widehat{Z^0}$, of eqs. (103), (104), which is a solution of the global equation:

$$\Phi_N(M^-, M^0+\mathcal{C}^T\mathcal{Q}^{-1}\mathcal{C}, M^+)\widehat{Z}^0 = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \tag{106}$$

so, on account of the relation between $\Phi(\cdot,\cdot,\cdot)$ and $\mathcal{F}(\cdot,\cdot,\cdot)$, it is

$$\mathcal{F}(M^-, M^0+\mathcal{C}^T\mathcal{Q}^{-1}\mathcal{C}, M^+)\widehat{Z}^0 = \begin{bmatrix} M_0^0\widehat{Z}_0^0 + M_0^+\widehat{Z}_1^0 + M_N^0\widehat{Z}_N^0 \\ \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}}\mathcal{Y} \\ M_N^+\widehat{Z}_0^0 + M_N^-\widehat{Z}_{N-1}^0 + M_N^0\widehat{Z}_N^0 \end{bmatrix}, \tag{107}$$

where the first and last blocks are just identities. Noting that

$$\mathcal{F}(M^-, M^0+\mathcal{C}^T\mathcal{Q}^{-1}\mathcal{C}, M^+) = \Gamma^{-1},$$

equation (107) entails

$$\begin{aligned}\widehat{Z}_t^0 = \sum_{s=1}^{N-1}\Gamma(t,s)\mathcal{C}_s^T\mathcal{Q}^{-1}\mathcal{Y}_s + \Gamma(t,0)\left(M_0^0\widehat{Z}_0^0 + M_0^+\widehat{Z}_1^0 + M_N^0\widehat{Z}_N^0\right)\\ +\Gamma(t,N)\left(M_N^+\widehat{Z}_0^0 + M_N^-\widehat{Z}_{N-1}^0 + M_N^0\widehat{Z}_N^0\right),\end{aligned} \tag{108}$$

and subtracting eq. (105) results finally in

$$\begin{aligned}\widehat{Z}_t = \widehat{Z}_t^0 + \Gamma(t,0)\left(\mathcal{C}_0^T\mathcal{Q}_0^{-1}\mathcal{Y}_0 - M_0^0\widehat{Z}_0^0 + M_0^-\widehat{Z}_1^0 - M_N^0\widehat{Z}_N^0\right)\\ +\Gamma(t,N)\left(\mathcal{C}_N^T\mathcal{Q}_N^{-1}\mathcal{Y}_N - M_N^+\widehat{Z}_0^0 + M_N^-\widehat{Z}_{N-1}^0 - M_N^0\widehat{Z}_N^0\right).\end{aligned} \tag{109}$$

In conclusion, the optimal smoothing estimate of Z_t, on the discrete circle $[0,N]$ is calculated by first iterating, using a measurement path, the backward/forward equations (103), (104), next adjusting the solution $\widehat{Z}_t^0$ so obtained, by means of eq. (109).

The optimal smoothing estimates of the hidden reciprocal chain, i.e $\widehat{X}_t$, for $t=0,...,N$, can be recovered by $\widehat{Z}_t$ by first recovering $\widehat{Z}_t^+$, using the definition of Z_t, i.e. $Z_t = Z_t^+ - \mathbf{E}\{Z_t^+\}$, (hence $\widehat{Z}_t = \widehat{Z}_t^+ - \mathbf{E}\{X_{t-1}\otimes X_t\}$) then noting that

$$\widehat{Z}_t^+ = \mathbf{E}\{X_{t-2}\otimes X_t|Y\} = \sum_{i,j=1}^{n}(e_i\otimes e_j)\mathbf{Pr}\{X_{t-2}=e_i, X_t=e_j|Y\}, \tag{110}$$

so $\widehat{Z}_t^+$ is actually the collection of all joint probabilities. Thus it is

$$\widehat{X}_t = \sum_{i,j=1}^{n} e_j\mathbf{Pr}\{X_{t-2}=e_i, X_t=e_j|Y\}\mathbf{Pr}\{X_{t-2}=e_i\}. \tag{111}$$

The case of Dirichelet Boundary conditions.

We consider now the case of an HRM defined over the whole set of integers $\{0, \pm 1, \pm 2, ...\}$, aiming to solve the optimal smoothing problem on a finite discrete interval. We show that this problem can be solved, provided that the underlying reciprocal process $\{X_t\}$ can be fully observed at *two points* of the integer domain of definition, and the observation model

$$Y_t = C_t X_t + W_t, \tag{112}$$

(the same just considered in the cyclic case), is assigned *between* these two points. In other words, the smoothing problem can be solved under the classical assumptions that assures the existence of the solution in the continuous-valued, Gaussian, case. Accordingly, we call the case at issue: the Dirichelet boundary condition case. Without loss of generality we assume the two-points boundary is $t = -1$, $t = N$, that is

$$X_{-1} = \overline{X}_{-1}; \qquad X_N = \overline{X}_N, \tag{113}$$

where $\overline{X}_{-1}, \overline{X}_N$ are two random vectors assumed directly *observable* without noise, whereas the noisy observation Y_t, satisfying (112), is defined for $t \in (-1, N)$.

As to the stochastic realization issue for the reciprocal chain $\{X_t\}$, it makes no difference with respect to the cyclic case, so the reciprocal model (**??**) holds, with the same meaning of symbols, but t being now any integer. The global state Z is the same aggregate of Z_t-values as before, $Z = \mathrm{col}(Z_0, Z_1, ..., Z_N)$, but the global measurements vector $\mathcal{Y}$ is defined only in the *interior* of the interval $[0, N]$, hence $\mathcal{Y} = \mathrm{col}(\mathcal{Y}_1, \mathcal{Y}_2, ..., \mathcal{Y}_{N-1})$, with $\mathcal{Y}_t = Y_{t-1} \otimes Y_t$ defined for $t = 1, ..., N-1$, so only measurements Y_t in $(-1, N)$ are involved. As a conseqence, we get a slightly different equation, than (80), from the linear estimation formula:

$$\Phi_N(M^-, M^0, M^+)\widehat{Z} = \mathcal{C}^T R^{\dagger}_{\mathcal{W}\mathcal{W}} \left(\mathcal{Y} - \mathcal{C}\widehat{Z}\right) \tag{114}$$

where $\Phi_N(\cdot, \cdot, \cdot)$ is the matrix-builder defined in (92), and – besides $\mathcal{Y}$ – $R^{\dagger}_{\mathcal{W}\mathcal{W}}$, and $\mathcal{C}$, have a different definition as well:

$$\mathcal{C} = \mathrm{diag}(\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_{N-1}),$$
$$R^{\dagger}_{\mathcal{W}\mathcal{W}} = \mathrm{diag}(\mathcal{Q}_1^{-1}, \mathcal{Q}_2^{-1}, ..., \mathcal{Q}_{N-1}^{-1}),$$

as $\mathcal{W} = \mathrm{col}(\mathcal{W}_1, ..., \mathcal{W}_{N-1})$ ($\mathcal{W}_t$ defined as in the cyclic case, but $t = 1, ..., N-1$).

Equation (114) can be rewritten

$$\Phi_N(M^-, M^0 + \mathcal{C}^T \mathcal{Q}^{-1}\mathcal{C}, M^+)\widehat{Z} = \mathcal{C}^T R^\dagger_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \tag{115}$$

and defining the matrix M_d:

$$M_d = \begin{bmatrix} I & 0 & \dots & & 0 \\ M_1^- & M_1^0 + \mathcal{C}_1^T \mathcal{Q}_1^{-1}\mathcal{C}_1 & M_1^+ & 0 & \vdots \\ 0 & M_2^- & M_2^0 + \mathcal{C}_2^T \mathcal{Q}_2^{-1}\mathcal{C}_2 & M_2^+ & \\ & & \ddots & & \\ & & M_{N-1}^- & M_{N-1}^0 + \mathcal{C}_{N-1}^T \mathcal{Q}_{N-1}^{-1}\mathcal{C}_{N-1} & M_{N-1}^+ \\ 0 & \dots & & 0 & I \end{bmatrix}, \tag{116}$$

from (115) we comes to

$$M_d\widehat{Z} = \mathrm{col}\left(\widehat{Z}_0, \mathcal{C}^T R^\dagger_{\mathcal{W}\mathcal{W}}\mathcal{Y}, \widehat{Z}_N\right). \tag{117}$$

Under the non-singularity assumption we have seen $M > 0$, where M is the circulant matrix (56) built up with the projection matrices $M^\pm, M^0$ of the process $\{X_t\}$, thus, under the same hypothesis, the matrix (116) is invertible as well[8], so we can extract, similarly as for eq. (105), a kernel-function $\Gamma_d(\cdot,\cdot)$ from a matrix $\Gamma_d = M_d^{-1}$, such a way eq. (117) leads to

$$\widehat{Z}_t = \sum_{s=1}^{N-1} \Gamma_d(t,s)\mathcal{C}_s^T \mathcal{Q}_s^{-1}\mathcal{Y}_s + \Gamma_d(t,0)\widehat{Z}_0 + \Gamma_d(0,N)\widehat{Z}_N. \tag{118}$$

Note that eq. (96), corresponding to the cyclic case, agrees with (115) *inside* the discrete circle, and in particular the matrix $\Phi_N(M^-, M^0 + \mathcal{C}^T \mathcal{Q}^{-1}\mathcal{C}, M^+)$ involved in eq. (115) is the same as for the cyclic case, so it undergoes to the same decomposition as in (98). Thus we come to the same recursive algorithm, given by the couple of backward/forward eqs. (103), (104), yet worked out before. These equations gives a *particular* solution, namely $\widehat{Z}_t^0$, for $t = 1, ..., N-1$, which could be even the same as for the cyclic case. The *actual* solution $\widehat{Z}_t$, $t = 0, ..., N$, has to be calculated by taking into account of the boundary conditions (113).

To this purpose, first of all, using the relation $Z_t = Z_t^+ - \mathbf{E}\{Z_t^+\}$, where

$$\mathbf{E}\{Z_t^+\} = \mathbf{E}\{\widehat{Z_t^0}^+\} = \mathbf{E}\{X_{t-1} \otimes X_t\}$$

[8]By deleting the first and last rows, and the first and last columns, of (116) we get a matrix with the same determinant. This matrix agrees with one of the minors of the matrix $M + \mathcal{C}^T R^\dagger_{\mathcal{W},\mathcal{W}}\mathcal{C}$ we have yet recognized, while considering the cyclic case, being non-singular.

so the mean values are all known for any t by hypothesis, one derives $\widehat{Z_t^0}^+$ for $t = 1, ..., N-1$, and hence $\widehat{X_0^0}$ and $\widehat{X_N^0}$ from formula (111). Thus we have $\widehat{Z_0^0}^+ = \overline{X}_{-1} \otimes \widehat{X_0^0}$, and $\widehat{Z_N^0}^+ = \widehat{X_{N-1}^0} \otimes \overline{X}_N$, (and hence $\widehat{Z_0^0}, \widehat{Z_N^0}$ as well, by subtracting the means). As $\widehat{Z}_t^0$, $t = 0, ..., N$, is a particular solution, it has to satisfy eq. (118) as well, so taking the difference eventually we get

$$\widehat{Z}_t = \widehat{Z}_t^0 + \Gamma_d(t,0)(\widehat{Z}_0^0 - \widehat{Z}_0) + \Gamma_d(0,N)(\widehat{Z}_N^0 - \widehat{Z}_N). \tag{119}$$

The above equation for $t = 0$, and $t = N$, yields

$$\widehat{Z}_0 = \widehat{Z}_0^0 + \Gamma_d(0,0)(\widehat{Z}_0^0 - \widehat{Z}_0) + \Gamma_d(0,N)(\widehat{Z}_N^0 - \widehat{Z}_N), \tag{120}$$

$$\widehat{Z}_N = \widehat{Z}_N^0 + \Gamma_d(N,0)(\widehat{Z}_0^0 - \widehat{Z}_0) + \Gamma_d(N,N)(\widehat{Z}_N^0 - \widehat{Z}_N), \tag{121}$$

which is a system of equations, such that $\widehat{Z}_0, \widehat{Z}_N$ can be calculated from (as the *unique* solutions, to within stochastic equivalence i.e. for any sample of $\widehat{Z}_0^0, \widehat{Z}_N^0$ we get the corresponding sample of $\widehat{Z}_0, \widehat{Z}_N$). Then, using eq. (119), one has the process $\widehat{Z}_t$ got calculated for any $t = 0, ..., N$.

Finally, using again the relation $\widehat{Z}_t = \widehat{Z}_t^+ - \mathbf{E}\{X_{t-1} \otimes X_t\}$, one get $\widehat{Z}_t^+$, for $t = 0, ..., N$, so $\widehat{X_t}$ can be recovered using the same equation as for the cyclic case, i.e. eq. (111).

IV. Conclusions

References

[1] 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.

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

[3] 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.

[4] 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.

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

[6] 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.

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

[8] 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.

[9] 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.

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

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

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

[13] 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.

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

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