Application of the theory of KM20-Langevin equations to the non-linear prediction problem for the one-dimensional strictly stationary time series

Dedicated to Professor Kiyoshi Ito on his seventy-seven birthday

By Yasunori OKABE and Takashi OOTSUKA

(Received Feb. 5, 1993)
(Revised Aug. 16, 1993)

§ 1. Introduction.

We are inspired by Masani-Wiener's work ([4]) of the non-linear prediction problem of a one-dimensional discrete time strictly stationary process. The purpose of the present paper is to give computable algorithms for the non-linear predictor by applying the theory of KM20-Langevin equations.

We have already applied in [7] the theory of KM20-Langevin equations to the linear prediction problem for the multi-dimensional weakly stationary time series and given a refinement of Wiener-Masani's work in [13], [14] and [3] by obtaining computable algorithms for the linear predictor. The results in [7] play supplementary but useful roles in the present approach to the non-linear problem, as will be explained.

Let \( X = (X(n); n \in \mathbb{Z}) \) be a real-valued strictly stationary time series on a probability space \((\Omega, \mathcal{B}, P)\) with mean zero. We shall impose the following two hypotheses which are the same as in [4]:

(H.1) \( X \) is essentially bounded, i.e., there exists a positive constant \( C > 0 \) such that \( |X(n)(\omega)| \leq C \) for any \( n \in \mathbb{Z} \) and almost all \( \omega \in \Omega \);

(H.2) For any distinct integers \( n_1, n_2, \ldots, n_k (k \in \mathbb{N}) \) the spectrum of the distribution function of the \( k \)-dimensional random variable \((X(n_1), X(n_2), \ldots, X(n_k))\) has positive Lebesgue measure.

The non-linear predictor \( \hat{X}(\nu) \) of the future \( X(\nu), \nu > 0 \), on the basis of the present and past \( X(l), l \leq 0 \), is defined by

\[
\hat{X}(\nu) = E(X(\nu) | \sigma(X(l); l \leq 0)).
\]

Masani and Wiener ([4]) have obtained a representation for the non-linear problem.

This research was partially supported by Grant-in-Aid for Science Research No. 03452011, the Ministry of Education, Science and Culture, Japan.
predictor as follows:

\[
E\{X(v) \mid \sigma(X(l) ; l \leq 0)\} = \lim_{n \to \infty} Q_n(X(0), X(-1), \ldots, X(-m_n)),
\]

where, for each \(n \in \mathbb{N}\), \(m_n\) is a nonnegative integer depending on \(n\), and \(Q_n\) is a real polynomial in \(m_n+1\) variables whose coefficients can be theoretically calculated in terms of the moments of the time series \(X\).

However, as Kallianpur has given some comments in [12], the representation (1.1) of the non-linear predictor lacks for computable algorithm which is fit for the application to applied science, because the determination of the coefficients of the polynomials \(Q_n\) involves the calculation of the determinants of matrices of different sizes, coming from their method of Schmidt’s orthogonalization. On the other hand, Masani and Wiener have suggested in [4] that certain computable algorithm for the non-linear predictor may be obtained by means of the linear predictor for a suitably defined, infinite-dimensional, weakly stationary time series.

Following their suggestion, we shall derive an \(\mathbb{R}^\infty\)-valued weakly stationary time series \(\mathbf{x} = (x(n) ; n \in \mathbb{Z})\) and consider the \(d_q+1\)-dimensional subprocesses \(X^{(q)}(n) ; n \in \mathbb{Z}\) generated by the first \(d_q+1\)-components of \(\mathbf{x}\). We remark that \(d_1 = 0\), \(d_q\) is increasing to \(\infty\) as \(q \to \infty\) and \(X^{(q)} = X\). According to the theory of KM\(_d\)-Langevin equations ([5], [6], [9]), for each \(q \in \mathbb{N}\), the linear predictor for the \(d_q+1\)-dimensional subprocess \(X^{(q)}\) can be calculated from the KM\(_d\)-Langevin data \(\mathcal{L}\mathcal{D}(X^{(q)})\) which, corresponding to the fluctuation-dissipation theorem, is obtained from the computable algorithm in terms of the correlation function of \(X^{(q)}\). By obtaining a new algorithm computing the KM\(_d\)-Langevin data \(\mathcal{L}\mathcal{D}(X^{(q)})\) from the KM\(_d\)-Langevin data \(\mathcal{L}\mathcal{D}(X^{(q-1)})\) \((q=2, 3, \ldots)\), we can practically solve the non-linear prediction problem for the original time series \(X\), because the non-linear predictor for \(X\) can be obtained as the limit as \(q \to \infty\) of the first component of the linear predictors for \(X^{(q)}\).

Now we shall explain the contents of this paper. In § 2, according to [5] and [9], we shall recall and rearrange the theory of KM\(_d\)-Langevin equations for a \(d\)-dimensional weakly stationary time series \(Z = (Z(n) ; |n| \leq N)\), where \(d, N\) are fixed natural numbers. In particular, we shall introduce the KM\(_d\)-Langevin data \(\mathcal{L}\mathcal{D}(Z)\) associated with the time series \(Z\) which consists of the triplet of the forward and backward KM\(_d\)-Langevin delay functions, the forward and backward KM\(_d\)-Langevin partial correlation functions, and the forward and backward KM\(_d\)-Langevin fluctuation functions. The KM\(_d\)-Langevin data \(\mathcal{L}\mathcal{D}(Z)\), together with the forward and backward KM\(_d\)-Langevin forces, will determine the forward and backward KM\(_d\)-Langevin equations describing the time evolution of the time series \(Z\). We can obtain a concrete expression for the linear predictor for the time series \(Z\) in terms of the KM\(_d\)-Langevin data.
Application of the theory of \( \mathcal{KM}_2 \)-Langevin equations

Furthermore, associated with a \( d \)-dimensional weakly stationary time series \( \mathbf{Z} = (Z(n); n \in \mathbb{Z}) \), we can construct the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{Z}) \).

Section 3 will develop the theory of the \( \mathcal{KM}_2 \)-Langevin equations and obtain a new formula between the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{Z}) \) and the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{Y}) \), where the time series \( \mathbf{Y} \) is a \( d' \)-dimensional local and weakly stationary time series generated by the first \( d' \)-components of the series \( \mathbf{Z} \) (\( 1 \leq d' < d \)).

In the last section, we shall return to the real-valued strictly stationary time series \( \mathbf{X} = (X(n); n \in \mathbb{Z}) \) with mean zero satisfying conditions (H.1) and (H.2). By modifying the idea in Masani and Wiener ([4]), we shall derive an \( \mathbb{R}^m \)-valued weakly stationary time series \( \mathbf{X} = (X(n); n \in \mathbb{Z}) \) and consider the \( d_q + 1 \)-dimensional subprocesses \( \mathbf{X}^{(q)} = (X^{(q)}(n); n \in \mathbb{Z}) \) generated by the first \( d_q + 1 \)-components of \( \mathbf{X} \). We remark that the first components of \( X^{(q)}(n) \) are equal to \( X(n) \) (\( q \in \mathbb{N}, n \in \mathbb{Z} \)) and the construction of the time series \( \mathbf{X}^{(q)} \) with dimension \( d_q + 1 \) is fit for the application to data analysis. Applying the results in Section 3 to these time series \( \mathbf{X}^{(q)} \), we shall obtain an algorithm computing the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{X}^{(q)}) \) from the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{X}^{(q-1)}) \) (\( q = 2, 3, \ldots \)). Thus the non-linear prediction problem for the original real valued strictly stationary time series \( \mathbf{X} \) can be practically solved as follows:

\[
E(X(v)|\sigma(X(l); l \leq 0)) = \text{the first component of } \lim_{N \to \infty} \sum_{N, q = 0}^{N} Q_q(X^{(q)})(N+v, N; N-k)X^{(q)}(-k),
\]

where, for each \( q \in \mathbb{N} \), the \( M(d_q+1; \mathbb{R}) \)-valued function \( Q_q(X^{(q)})(\cdot, \cdot; \cdot) \) is called the forward prediction function associated with the time series \( X^{(q)} \) in the theory of the \( \mathcal{KM}_2 \)-Langevin equations, which can be recursively calculated from the \( \mathcal{KM}_2 \)-Langevin data \( \mathcal{L}_{\mathcal{D}}(\mathbf{X}^{(q)}) \). By using the results in [7], furthermore, we can theoretically obtain an algorithm for the limit as \( N \to \infty \) of the forward prediction functions \( Q_q(X^{(q)})(N+v, N; N-k) \) for any fixed \( q, v \in \mathbb{N}, k \in \mathbb{N}^* (\equiv \mathbb{N} \cup \{0\}) \).

As the application of the theory of \( \mathcal{KM}_2 \)-Langevin equations to data analysis, we are going to develop a new project of the stationary, causal and prediction analysis ([9], [8], [10]).

The authors would like to thank the referee for valuable advices.

Section 2. The theory of \( \mathcal{KM}_2 \)-Langevin equations.

We shall recall the theory of \( \mathcal{KM}_2 \)-Langevin equations from [5], [9].

Let \( d \) and \( N \) be any natural numbers. Let \( \mathbf{Z} = (Z(n); |n| \leq N) \) be any \( d \)-dimensional real-valued local and weakly stationary time series on a
probability space \((\Omega, \mathcal{B}, P)\) with covariance matrix function \(R^Z\):

\[
R^Z(n) = E(Z(n)^t Z(0)) \quad (|n| \leq N).
\]

Then we define, for each \(n \in \mathbb{N}, 1 \leq n \leq N\), two block Toeplitz matrices \(T^+_n(Z), T^-_n(Z) \in M(nd; \mathbb{R})\) by

\[
T^+_n(Z) = \begin{pmatrix}
R^Z(0) & R^Z(\pm 1) & \cdots & R^Z(\pm (n-1)) \\
R^Z(\pm 1) & R^Z(0) & \cdots & R^Z(\pm (n-2)) \\
\vdots & \vdots & \ddots & \vdots \\
R^Z(\pm (n-1)) & R^Z(\pm (n-2)) & \cdots & R^Z(0)
\end{pmatrix},
\]

It is to be noted that

\[
T^+_n(Z) = T^-_n(Z) = R^Z(0).
\]

In this subsection, we treat the case where the following condition holds:

\[
T^+_n(Z), T^-_n(Z) \in GL(nd; \mathbb{R}) \quad (1 \leq n \leq N).
\]

We remark that condition (2.5) is equivalent to

\[
\{Z_j(n); 1 \leq j \leq d, |n| \leq N\} \text{ is linearly independent in } L^2(\Omega, \mathcal{B}, P),
\]

where \(Z(n)=^t(Z_1(n), \ldots, Z_d(n))\).

For any \(d\)-dimensional square-integrable stochastic process \(Y=(Y(n); l \leq n \leq r)\) with a discrete time parameter space defined on the probability space \((\Omega, \mathcal{B}, P)\) \((l, r \in \mathbb{Z}, 1 < r)\), we define, for any \(m, n \in \mathbb{Z}, l \leq m \leq n \leq r\), a real closed subspace \(L^m_n(Y)\) of \(L^2(\Omega, \mathcal{B}, P)\) by

\[
L^m_n(Y) = \text{the closed linear hull of } \{Y_j(k); 1 \leq j \leq d, m \leq k \leq n\}.
\]

According to the method of innovation, we introduce the \(d\)-dimensional forward (resp. backward) KM\(_2\)O-Langevin force \(v^+_n(Z)\) (resp. \(v^-_n(Z)\)) as follows:

\[
\begin{align*}
v^+_n(Z)(n) &= Z(n) - P_{X_{n+1}^0,Z}(n) \quad (0 \leq n \leq N) \\
v^-_n(Z)(m) &= Z(m) - P_{X_{m+1}^0,Z}(m) \quad (-N \leq m \leq 0),
\end{align*}
\]

where \(L^0_n(Z) = L^0_n(Z) = \{0\}\).

For each \(n \in \mathbb{N}^*, 0 \leq n \leq N\), let \(V_+(Z)(n)\) (resp. \(V_-(Z)(n)\)) be the covariance matrix of \(v^+_n(Z)(n)\) (resp. \(v^-_n(Z)(-n)\)). We call the function \(V_+(Z)(\cdot)\) (resp. \(V_-(Z)(\cdot)\)) the forward (resp. backward) KM\(_2\)O-Langevin fluctuation function. The following causal relation holds among \(Z, v_+(Z)\) and \(v_-(Z)\):
Causal relation ([5], [9]).

\[(2.9)\] \(\nu_+(Z)(0) = \nu_-(Z)(0) = Z(0).\)

\[(2.10)\] \(E(\nu_+(Z)(\pm m)\nu_-(Z)(\pm n)) = \delta_{mn}V_\pm(Z)(n) \quad (0 \leq m, n \leq N).\)

\[(2.11)\]
\[\mathcal{L}_n^\pm(Z) = \mathcal{L}_n^\pm(\nu_+(Z)) \quad (0 \leq n \leq N).\]

\[(2.12)\]
\[\mathcal{L}_n^\pm(Z) = \mathcal{L}_n^\pm(\nu_-(Z)) \quad (0 \leq n \leq N).\]

Let the system \(\mathcal{S}(Z)\) of elements in \(M(d, \mathbb{R})\) be the \(K_M\)-O-Langevin data associated with the process \(Z:\)

\[\mathcal{S}(Z) = \{r_+(Z)(n, k), r_-(Z)(n, k), \delta_+(Z)(m), \delta_-(Z)(m), V_+(Z)(l), V_-(Z)(l);\]

\[k, m, n \in \mathbb{N}, 1 \leq k < n \leq N, 1 \leq m \leq N, l \in \mathbb{N}^*, 0 \leq t \leq N\].

We know that \(Z\) satisfies the forward (resp. backward) \(K_M\)-O-Langevin equation (2.12+) (resp. (2.12-)):

\[K_M\text{-O-LANGEVIN EQUATIONS ([5], [9]).}\]

\[(2.12+)\] \(Z(\pm n) = - \sum_{k=1}^{n-1} \gamma_+(Z)(n, k)Z(\pm k) - \delta_+(Z)(n)Z(0) + \nu_+(Z)(\pm n) \quad (1 \leq n \leq N).\)

In the sequel we adopt a convention to make the summation running the empty set 0. We call the function \(\gamma_+(Z)(\cdot, \cdot)\) (resp. \(\gamma_-(Z)(\cdot, \cdot)\)) the forward (resp. backward) \(K_M\)-O-Langevin delay function associated with the process \(Z\).

The function \(\delta_+(Z)(\cdot)\) (resp. \(\delta_-(Z)(\cdot)\)) is said to be the forward (resp. backward) \(K_M\)-O-Langevin partial correlation function associated with the process \(Z\).

Concerning the relation between the Toeplitz matrices and the \(K_M\)-O-Langevin fluctuation functions, we can use the \(K_M\)-O-Langevin equations to see that

\[(2.13)\] \(\det T_n^\pm(Z) = \prod_{k=0}^{n-1} \det V_\pm(Z)(k) \quad (1 \leq n \leq N).\)

If follows from (2.5) and (2.13+) that

\[(2.14)\] \(V_+(Z)(n), V_-(Z)(n) \in GL(d, \mathbb{R}) \quad (0 \leq n \leq N).\)

The fluctuation-dissipation theorem (FDT) stated in §1 is the following:

\[FDT ([2], [1], [11], [15], [5], [9]).\] For any \(n, k \in \mathbb{N}, 1 \leq k < n \leq N,\)

\[(2.15)\] \(\gamma_+(Z)(n, k) = \gamma_+(Z)(n-1, k-1) + \delta_+(Z)(n)\gamma_+(Z)(n-1, n-k-1);\)

\[(2.16)\] \(V_+(Z)(n) = (I - \delta_+(Z)(n)\delta_+(Z)(n))V_+(Z)(n-1);\)

\[(2.17)\] \(\delta_-(Z)(n)V_+(Z)(n-1) = V_-(Z)(n-1)\delta_+(Z)(n);\)

\[(2.18)\] \(\delta_-(Z)(n)V_+(Z)(n) = V_-(Z)(n)\delta_+(Z)(n),\)

where we put
\[ (2.19) \quad \gamma_+(Z)(m, 0) = \delta_+(Z)(m) \quad \text{and} \quad \gamma_-(Z)(m, 0) = \delta_-(Z)(m) \quad (1 \leq m \leq N). \]

The relations \((2.16)\) and \((2.17)\) in FDT come from the following relation:

**BURG’S RELATION** ([11], [15], [5], [9]). For any \( n \in \mathbb{N} \), \( 1 < n \leq N \),

\[ (2.20) \quad \sum_{k=0}^{n-1} \gamma_+(Z)(n, k) R^2(k+1) = \sum_{k=0}^{n-1} R^2(k+1) \gamma_-(Z)(n, k). \]

FDT implies that both the KM20-Langevin delay and fluctuation functions can be recursively calculated from the KM20-Langevin partial correlation functions. On the other hand, the latter can be obtained from the correlation function \( R^2 \) by the following formulæ:

**KM20-LANGEVIN PARTIAL CORRELATION FUNCTIONS** ([2], [1], [11], [15], [5], [9]). For any \( n \in \mathbb{N} \), \( 1 \leq n \leq N \),

\[ (2.21) \quad \delta_+(Z)(n) = - (R^2(\pm n) + \sum_{k=0}^{n-2} \gamma_\pm(Z)(n-1, k) R^2(\pm(k+1))) V_\pm(Z)(n-1)^{-1}. \]

For any \( m, n \in \mathbb{N}^* \), \( 0 \leq n \leq m \leq N \), we define \( P_+(Z)(m, n) \), \( P_-(Z)(m, n) \) and \( e_+(Z)(m, n) \), \( e_-(Z)(m, n) \) by

\[ (2.22) \quad P_+(Z)(m, n) = E((Z(\pm m))^t \nu_+(Z)(\pm n)) V_+(Z)(n)^{-1/2} \]

and

\[ (2.23) \quad e_+(Z)(m, n) = E((Z(m) - P_+(Z)(m, n)) (Z(m) - P_+(Z)(m, n))), \]

\[ (2.24) \quad e_-(Z)(m, n) = E((Z(\pm m) - P_-(Z)(m, n)) (Z(\pm m) - P_-(Z)(m, n)) ). \]

We call the function \( P_+(Z)(\cdot, \cdot) \) (resp. \( P_-(Z)(\cdot, \cdot) \)) the forward (resp. backward) prediction function and the function \( e_+(Z)(\cdot, \cdot) \) (resp. \( e_-(Z)(\cdot, \cdot) \)) the forward (resp. backward) prediction error function. Then we know

**PREDICTION FORMULÆ** ([5], [9]). (i) For any \( m, n \in \mathbb{N}^* \), \( 0 \leq n \leq m \leq N \),

\[ (2.24) \quad P_+(Z)(m, n) = \sum_{k=0}^{n} P_+(Z)(m, k) V_+(Z)(k)^{-1/2} \nu_+(Z)(k); \]

\[ (2.25) \quad P_-(Z)(m, n) = \sum_{k=0}^{n} P_-(Z)(m, k) V_-(Z)(k)^{-1/2} \nu_-(Z)(-k). \]

(ii) For any \( m, n \in \mathbb{N}^* \), \( 0 \leq n < m \leq N \),

\[ (2.26) \quad P_+(Z)(m, n) = \sum_{k=0}^{n} Q_+(Z)(m, n ; k) Z(k); \]

\[ (2.27) \quad P_-(Z)(m, n) = \sum_{k=0}^{n} Q_-(Z)(m, n ; k) Z(-k). \]

Here the \( M(d ; \mathbb{R}) \)-valued prediction functions \( P_+(Z)(\cdot, \cdot) \) and \( Q_+(Z)(\cdot, \cdot; \cdot) \)
can be determined by the following algorithms:

**Prediction Algorithms ([5], [9]).**

(i) For any \( m, k \in \mathbb{N}^* \), \( 0 \leq k \leq m \leq N \),

\[
P_s(Z)(m, k) = \begin{cases} 
V_s(Z)(k)^{1/2} & \text{if } m = k \\
-\sum_{l=0}^{m-k-1} \gamma_s(Z)(m, l)P_s(Z)(l, k) & \text{if } m > k + 1.
\end{cases}
\]

(ii) For any \( m, n, k \in \mathbb{N}^* \), \( 0 \leq k \leq n < m \leq N \),

\[
Q_s(Z)(m, n ; k) = -\sum_{i=n+1}^{m-1} \gamma_s(Z)(m, i)Q_s(Z)(i, n ; k) - \gamma_s(Z)(m, k).
\]

Finally, the prediction error functions can be calculated by the following formulae:

**Prediction Error Formulae ([5], [9]).**

(i) For any \( m, n \in \mathbb{N}^* \), \( 0 < n < m \leq N \),

\[
e_s(Z)(m, n) = \sum_{k=n+1}^{m} P_s(Z)(m, k) P_s(Z)(m, k).
\]

(ii) In particular, for any \( n \in \mathbb{N} \), \( 0 < n < N \),

\[
e_s(Z)(n, n-1) = (I - \delta_s(Z)(n)\delta_s(Z)(n)) \cdots (I - \delta_s(Z)(1)\delta_s(Z)(1))R_s(0).
\]

[2.2] Let \( Z=(Z(n) ; n \in \mathbb{Z}) \) be any \( d \)-dimensional real-valued weakly stationary time series on a probability space \((\Omega, \mathcal{F}, P)\) with covariance function \( R^2 \). In this subsection, we treat the case where the following condition holds:

\[
(Z_j(n) ; 1 \leq j \leq d, n \in \mathbb{Z}) \text{ is linearly independent in } L^2(\Omega, \mathcal{F}, P),
\]

where \( Z(n)=\langle Z_1(n), \ldots, Z_d(n) \rangle \).

By restricting the time parameter space, we have a \( d \)-dimensional real-valued local and weakly stationary time series \( Z_N=(Z(n) ; |n| \leq N) \) \((N \in \mathbb{N})\). It then can be seen that the system \( \{\mathcal{L} \mathcal{D}(Z_N) ; N \in \mathbb{N}\} \) of the KM\( 2O \)-Langevin data \( \mathcal{L} \mathcal{D}(Z_N) \) \((N \in \mathbb{N})\) satisfies the following consistency condition:

\[
\gamma_s(Z_{N+1})(n, k) = \gamma_s(Z_N)(n, k) \quad (1 \leq k < n \leq N);
\]

\[
\delta_s(Z_{N+1})(n) = \delta_s(Z_N)(n) \quad (1 \leq n \leq N);
\]

\[
V_s(Z_{N+1})(n) = V_s(Z_N)(n) \quad (0 \leq n \leq N).
\]

Therefore, we can construct a KM\( 2O \)-Langevin data \( \mathcal{L} \mathcal{D}(Z) \) associated with the process \( Z \):

\[
\mathcal{L} \mathcal{D}(Z) = \{\gamma_s(Z)(n, k), \delta_s(Z)(m), V_s(Z)(l) ; k, m, n \in \mathbb{N}, k < n, l \in \mathbb{N}^*\}.
\]
§ 3. A new formula for the KM$_4$O-Langevin data.

Let $d, d^{(1)}, d^{(2)}, N$ be any natural numbers such that $d=d^{(1)}+d^{(2)}$ and let $Z=(Z(n); |n| \leq N)$ be any $d$-dimensional local and weakly stationary time series satisfying condition (2.6). We divide the components of $Z(n)$ into two blocks $Y(n)$ and $W(n)$, i.e.,

$$Z(n) = \begin{pmatrix} Y(n) \\ W(n) \end{pmatrix} \quad (|n| \leq N),$$

where $Y(n)=({Z_1(n), \ldots , Z_{d^{(1)}}(n)})$ and $W(n)=({Z_{d^{(1)}+1}(n), \ldots , Z_{d^{(1)}+d^{(2)}}(n)})$. It is to be noted that $Y=(Y(n); |n| \leq N)$ (resp. $W=(W(n); |n| \leq N)$) is a $d^{(1)}$-dimensional (resp. $d^{(2)}$-dimensional) weakly stationary time series satisfying condition (2.6).

In this section, we discuss how the KM$_4$O-Langevin data associated with $Z$ is calculated by those associated with $Y$ and $W$. We define the mutual correlation function $R^{YW}$ of $Y$ and $W$:

$$R^{YW}(n) = E(Y(n)W(0)) \quad (|n| \leq N).$$

Let $\mathcal{L}_d(Z)$ (resp. $\mathcal{L}_d(Y)$ and $\mathcal{L}_d(W)$) be the KM$_4$O-Langevin data associated with $Z$ (resp. $Y$ and $W$). We divide the components of matrices $\gamma_{s}(Z)(n, k)$ and $\delta_{s}(Z)(n)$ into four blocks $\gamma_{s}^{pq}(Z)(n, k)$ and $\delta_{s}^{pq}(Z)(n)$, for $p, q \in \mathbb{N}, 1 \leq p, q \leq 2$, i.e.,

$$\gamma_{s}(Z)(n, k) = \begin{pmatrix} \gamma_{s}^{11}(Z)(n, k) & \gamma_{s}^{12}(Z)(n, k) \\ \gamma_{s}^{21}(Z)(n, k) & \gamma_{s}^{22}(Z)(n, k) \end{pmatrix}$$

and

$$\delta_{s}(Z)(n) = \begin{pmatrix} \delta_{s}^{11}(Z)(n) & \delta_{s}^{12}(Z)(n) \\ \delta_{s}^{21}(Z)(n) & \delta_{s}^{22}(Z)(n) \end{pmatrix},$$

where $\gamma_{s}^{pq}(Z)(n, k)=((\gamma_{s}(Z)(n, k)))_{d^{(p-1)}+1 \leq \mu \leq d^{(p-1)}+d^{(p)}, d^{(q-1)}+1 \leq \nu \leq d^{(q-1)}+d^{(q)}}$ with $d^{(0)}=0$ and $\delta_{s}^{pq}(Z)(n)=\gamma_{s}^{pq}(Z)(n, 0)$.

Furthermore, we divide the components of $\nu_{s}(Z)(n)$ into two blocks $\nu_{s}^{1}(Z)(n)$ and $\nu_{s}^{2}(Z)(n)$, i.e.,

$$\nu_{s}(Z)(n) = \begin{pmatrix} \nu_{s}^{1}(Z)(n) \\ \nu_{s}^{2}(Z)(n) \end{pmatrix},$$

where $\nu_{s}^{1}(Z)(n)=({\nu_{s}^{11}(Z)(n), \ldots , \nu_{s,d^{(1)}}(Z)(n)})$ and $\nu_{s}^{2}(Z)(n)=({\nu_{s,d^{(1)}+1}(Z)(n), \ldots , \nu_{s,d^{(1)}+d^{(2)}}(Z)(n)})$. Then, for any $n \in \mathbb{N}, 1 \leq n \leq N$, the KM$_4$O-Langevin equations (2.12) for $Z$ are represented as follows:
Application of the theory of KM20-Langevin equations

\[ Z(\pm n) = - \sum_{k=1}^{n-1} \left( \delta^1_\pm(Z)(n, k) \gamma^1_\pm(Z)(n, k) + \delta^2_\pm(Z)(n, k) \gamma^2_\pm(Z)(n, k) \right) \left( Y(\pm k) \right) \]
\[ - \left( \delta^1_\pm(Z)(n) \gamma^1_\pm(Z)(n) + \delta^2_\pm(Z)(n) \gamma^2_\pm(Z)(n) \right) \left( Y(0) \right) + \left( \nu^1_\pm(Z)(\pm n) \right). \]

By noting (3.1), we have

\[ Y(\pm n) = - \sum_{k=1}^{n-1} \gamma^1_\pm(Z)(n, k) Y(\pm k) - \sum_{k=1}^{n-1} \gamma^2_\pm(Z)(n, k) W(\pm k) \]
\[ - \delta^1_\pm(Z)(n) Y(0) - \delta^2_\pm(Z)(n) W(0) + \nu^1_\pm(Z)(\pm n); \]
\[ W(\pm n) = - \sum_{k=1}^{n-1} \gamma^1_\pm(Z)(n, k) Y(\pm k) - \sum_{k=1}^{n-1} \gamma^2_\pm(Z)(n, k) W(\pm k) \]
\[ - \delta^1_\pm(Z)(n) Y(0) - \delta^2_\pm(Z)(n) W(0) + \nu^2_\pm(Z)(\pm n). \]

We shall obtain other formulae, different from (2.21), by which the KM20-Langevin partial correlation functions \( \delta_\pm(Z)(\cdot) \) and \( \sigma_\pm(Z)(\cdot) \) are recursively calculated from \( \mathcal{L}(Y) \), \( \mathcal{L}(W) \) and \( R^{YW} \) together with (2.15). For this purpose, we define \( B_+(Y|W)(l, k) \), \( B_-(Y|W)(l, k) \), \( B_+(W|Y)(l, k) \) and \( B_-(W|Y)(l, k) \) by

\[ B_\pm(Y|W)(l, k) = R^{YW}(\pm l) + \sum_{j=0}^{k-2} R^{YW}(\pm(l-k+j+1)) \gamma_\pm(W)(k-1, j) \]
\[ B_\pm(W|Y)(l, k) = R^{YW}(\pm l) + \sum_{j=0}^{k-2} R^{YW}(\pm(l-k+j+1)) \gamma_\pm(Y)(k-1, j) \]
for any \( k, l \in \mathbb{N}, 1 \leq k \leq N, 0 \leq l \leq N. \)

**Theorem 3.1.** For any \( n \in \mathbb{N}, 1 \leq n \leq N, \)

\[ \delta_\pm(Z)(n) = \begin{pmatrix} \delta_\pm(Y)(n)V_\pm(Y)(n-1) & 0 \\ 0 & \delta_\pm(W)(n)V_\pm(W)(n-1) \end{pmatrix} \]
\[ - \sum_{k=0}^{n-1} \gamma_\pm(Z)(n-1, k) \begin{pmatrix} 0 & B_\pm(Y|W)(k+1, n) \\ B_\pm(W|Y)(k+1, n) & 0 \end{pmatrix} \left( V_\pm(Z)(n-1)^{-1} \right), \]
where

\[ \gamma_\pm(Z)(j, j) = I \text{ and } \gamma_\pm(Z)(j, j) = I \quad (0 \leq j \leq N). \]

**Proof.** We prove the plus part. We shall rewrite the first term \( F \) of the right-hand side of the plus part of (2.21) for any fixed \( n \in \mathbb{N}, 1 \leq n \leq N: \)

\[ F = - \left( R^Z(\pm n) + \sum_{k=0}^{n-1} \gamma_\pm(Z)(n-1, k)R^Z(\pm(k+1)) \right). \]
We divide the components of matrix $F$ into four blocks $F_{pq}$ for $p, q \in \mathbb{N}$, $1 \leq p, q \leq 2$, i.e.,

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix},$$

where $F_{pq} = (F_{pq})_{d(q-1)+d(q-1)+d(p-1)+d(p-1)}$, $d(q-1)+d(q-1)+d(p-1)+d(p-1)$.

At first we rewrite the $(1, 1)$-block $F_{11}$ of $F$ as follows:

$$F_{11} = -\left( R^F(n) + \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^F(k+1) + \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)RW^Y(k+1) \right).$$

We shall rewrite the second term of the equation above; by using equation (2.12), we see from (2.10) and (2.11) that

$$\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)E(Y(k-n+2)'Y(-n+1))$$

$$= \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)E(Y(k-n+2)'(-\sum_{j=0}^{n-2} \gamma_{-1}(Y)(n-1, j)Y(-j)))$$

$$= -\sum_{k=0}^{n-2} \sum_{j=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^F(k-n+j+2)\gamma_{-1}(Y)(n-1, j)$$

$$= -\sum_{k=0}^{n-2} \sum_{j=0}^{n-2} \gamma_{11}(Z)(n-1, k)E(Y(k)'Y(n-j-2))\gamma_{-1}(Y)(n-1, j).$$

On the other hand, by using equation (3.4), we see from (2.10) and (2.11) that

$$E\left( -\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)Y(k) \right)Y(n-j-2)$$

$$= E(Y(n-1)'Y(n-j-2)) - E(\left( \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)W(k) \right)Y(n-j-2))$$

$$= R^F(j+1) + \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^W(k-n+j+2).$$

Further, by virtue of Burg's relation (2.20), we see
Application of the theory of KM\(_2\)-Langevin equations

\[ n-2 \sum_{k=0}^{n-1} \gamma_{11}(Z)(n-1, k)R_Y(k+1) \]

\[ = \sum_{k=0}^{n-1} \gamma_{11}(Y)(n-1, k)R_Y(k+1) \]

\[ + \sum_{j=0}^{n-2} \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)R_{XY}(k-n+j+2)\gamma_{-1}(Y)(n-1, j). \]

According to the definition of \( B_+(W|Y)(\cdot, *) \), we see from (2.20) that

\[ F^{11} = -\left( R^{Y}(n)+\sum_{k=0}^{n-2} \gamma_{11}(Y)(n-1, j)R^{Y}(k+1) \right) \]

\[ - \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k) \left( R^{XY}(k+1)+\sum_{j=0}^{n-2} R^{XY}(k-n+j+2)\gamma_{-1}(Y)(n-1, j) \right) \]

\[ = \delta_{1}(Y)(n)V_{-1}(Y)(n-1)-\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)B_+(W|Y)(k+1, n). \]

Therefore, according to (3.8), we get

(a) \[ F^{11} = \delta_1(Y)(n)V_{-1}(Y)(n-1)-\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)B_+(W|Y)(k+1, n). \]

Secondly, we rewrite the (2, 1)-block \( F^{21} \) of \( F \) as follows:

\[ F^{21} = -\left( R^{XY}(n)+\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^{Y}(k+1)+\sum_{j=0}^{n-2} \gamma_{11}(Z)(n-1, k)R_{WY}(k+1) \right). \]

We shall rewrite the second term of the equation above; by using equation (2.12), we see from (2.10) and (2.11) that

\[ \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^{Y}(k+1) \]

\[ = \sum_{k=0}^{n-2} E\left( -\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)Y(k) \right) \gamma_{-1}(Y)(n-1, j). \]

On the other hand, by using equation (3.5), we have from (2.10) and (2.11) that

\[ E\left( -\sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k)Y(k) \right) \gamma_{-1}(Y)(n-1, j) \]

\[ = R^{XY}(j+1)+\sum_{j=0}^{n-2} \gamma_{11}(Z)(n-1, k)R^{XY}(k-n+j+2). \]

Therefore, we obtain

\[ F^{21} = -\left( R^{XY}(n)+\sum_{k=0}^{n-2} R^{XY}(k+1)\gamma_{-1}(Y)(n-1, j) \right) \]

\[ - \sum_{k=0}^{n-2} \gamma_{11}(Z)(n-1, k) \left( R^{XY}(k+1)+\sum_{j=0}^{n-2} R^{XY}(k-n+j+2)\gamma_{-1}(Y)(n-1, j) \right). \]
According to the definition of $B_+(W|Y)(\cdot, \cdot)$ in (3.7+) and (3.8), we get

(b) \begin{align*}
F^{21} &= -\sum_{k=0}^{n-1} \gamma_{21}^{Z}(Z)(n-1, k)B_+(W|Y)(k+1, n).
\end{align*}

Similarly, we can show

(c) \begin{align*}
F^{12} &= -\sum_{k=0}^{n-1} \gamma_{12}^{Z}(Z)(n-1, k)B_+(Y|W)(k+1, n)
\end{align*}
and

(d) \begin{align*}
F^{22} &= \delta_+(W)(n)V_-(W)(n-1) - \sum_{k=0}^{n-1} \gamma_{22}^{Z}(Z)(n-1, k)B_+(Y|W)(k+1, n).
\end{align*}

Thus we can conclude from (a), (b), (c) and (d) that the plus part holds. In the same way, the minus part is proved. (Q.E.D.)

As stated in §2, $V_+(Z)(\cdot)$ and $V_-(Z)(\cdot)$ are recursively calculated from $\delta_+(Z)(\cdot)$ and $\delta_-(Z)(\cdot)$ by (2.16±). However, we can obtain other formulae for the KM20-Langevin fluctuation functions $V_\pm(Z)(\cdot)$, similar to Theorem 3.1.

**THEOREM 3.2.** For any $n \in \mathbb{N}$, $0 \leq n \leq N$,

$$V_+(Z)(n) = \begin{pmatrix} V_+(Y)(n) & 0 \\ 0 & V_+(W)(n) \end{pmatrix} + \sum_{k=0}^{n} \gamma_{(n,n-k)}(Z)(n) \begin{pmatrix} 0 & B_+(Y|W)(k+1, n+1) \\ B_+(W|Y)(k, n+1) & 0 \end{pmatrix}.$$

**PROOF.** We divide the components of matrices $V_\pm(Z)(n)$ into four blocks

$$V_{\pm}^{pq}(Z)(n)$$
for $p, q \in \mathbb{N}$, $1 \leq p, q \leq 2$, i.e.,

$$V_+(Z)(n) = \begin{pmatrix} V_+^{11}(Z)(n) & V_+^{12}(Z)(n) \\ V_+^{21}(Z)(n) & V_+^{22}(Z)(n) \end{pmatrix},$$
where $V_{\pm}^{pq}(Z)(n) = ((V_\pm(Z)(n))_{k}d^{(p-1-\lambda)}_{0} + d^{(p-1)_{0} + d^{(p-1)}}_{0} + d^{(p-1)}_{0} + d^{(p-1)}_{0} + d^{(p-1)}_{0} + d^{(p-1)}_{0})$.

We prove only the plus part, because the minus part is proved in the same way. By using equation (3.4+) for $Z$, it follows from (2.10+) and (2.11+) that

$$V_+^{11}(Z)(n) = E(\nu_{1}(Z)(n)Y(n)) + E(\nu_{1}(Z)(n)(\sum_{k=0}^{n-1} \gamma_{11}^{Z}(Z)(n, k)Y(k)))$$

$$+ E(\nu_{1}(Z)(n)(\sum_{k=0}^{n-1} \gamma_{11}^{Z}(Z)(n, k)W(k))) = E(\nu_{1}(Z)(n)Y(n)).$$

Further, by using equation (2.12+) for $Y$ and noting (2.10+) and (2.11+) that
Application of the theory of KMqO-Langevin equations

\[ V^\dagger (Z)(n) = E(\nu^\dagger (Z)(n)^t(-\sum_{k=0}^{n-1} \gamma_+(Y)(n, k)Y(k)) + E(\nu^\dagger (Z)(n)\nu_+(Y)(n)) \]

\[ = E(\nu^\dagger (Z)(n)^t\nu_+(Y)(n)). \]

By using equation (3.4+) for \( Z \), we see that

\[ V^\dagger (Z)(n) = E(Y(n)^t\nu_+(Y)(n)) + E\left( \left( \sum_{k=0}^{n-1} \gamma^\dagger (Z)(n, k)Y(k) \right)^t\nu_+(Y)(n) \right) \]

\[ + E\left( \sum_{k=0}^{n-1} \gamma^\dagger (Z)(n, k)W(k)^t\nu_+(Y)(n) \right) \]

\[ = V_+(Y)(n) + \sum_{k=0}^{n-1} \gamma^\dagger (Z)(n, k)E(W(k)^t\nu_+(Y)(n)). \]

On the other hand, by using equation (2.12+) for \( Y \),

\[ V^\dagger (Z)(n) = V_+(Y)(n) + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)E(W(n-l)^t\nu_+(Y)(n)) \]

\[ = V_+(Y)(n) + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)E(W(n-l)^t\nu_+(Y)(n)) \]

\[ + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)E(W(n-l)^t\nu_+(Y)(n)) \]

\[ = V_+(Y)(n) + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)R^{WY}(-l) \]

\[ + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)\sum_{j=0}^{n-1} R^{WY}(-(l-n+j))\gamma_+(Y)(n, j). \]

Therefore, according to the definition of \( B_-(W|Y)(\cdot, \cdot) \) in (3.7+) and (3.8),

\[ V^\dagger (Z)(n) = V_+(Y)(n) + \sum_{k=0}^{n-1} \gamma^\dagger (Z)(n, n-l)B_-(W|Y)(k, n+1). \]

In the same way as in \( V^\dagger (Z)(n) \), it follows from (3.4+), (3.5+), (2.10+), (2.11+) and (2.12+) that

\[ V^\dagger (Z)(n) = E(\nu^\dagger (Z)(n)^tY(n)) \]

\[ = E(\nu^\dagger (Z)(n)^t\nu_+(Y)(n)) \]

\[ = E(W(n)^t\nu_+(Y)(n)) + \sum_{k=0}^{n-1} \gamma^\dagger (Z)(n, k)E(W(k)^t\nu_+(Y)(n)) \]

\[ = R^{WY}(0) + \sum_{i=1}^{n-1} R^{WY}(n-l)^t\gamma_+(Y)(n, l) + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)R^{WY}(-l) \]

\[ + \sum_{i=1}^{n-1} \gamma^\dagger (Z)(n, n-l)\sum_{j=0}^{n-1} R^{WY}(-(l-n+j))\gamma_+(Y)(n, j). \]
Therefore, according to the definition of $B_-(W|Y)(\cdot, \cdot)$ in (3.7) and (3.8),
\[ V^+(Z)(n) = \sum_{k=0}^{n} \gamma^+\eta(Z)(n, n-k)B_-(W|Y)(k, n+1). \]

Similarly, we obtain
\[ V^+(Z)(n) = \sum_{k=0}^{n} \gamma^\eta(Z)(n, n-k)B_-(W|Y)(k, n+1) \]
and
\[ V^+(Z)(n) = V^+(Z)(n) + \sum_{k=0}^{n} \gamma^\eta(Z)(n, n-k)B_-(Y|W)(k, n+1). \]

Thus we can conclude from (e), (f), (g) and (h) that the plus part holds.
\[ \text{(Q. E. D.)} \]

§ 4. The non-linear prediction problem.

Let $X=(X(n) ; n \in \mathbb{Z})$ be a one-dimensional strictly stationary time series on a probability space $(\Omega, \mathcal{F}, P)$ with mean zero. Moreover we impose the same hypotheses as in Masani-Wiener [4]:

(H.1) $X$ is essentially bounded;

(H.2) for any distinct integers $(n_1, \cdots, n_k)$ the spectrum of the distribution function of the $k$-dimensional random variable $^t(X(n_1), \cdots, X(n_k))$ has positive Lebesgue measure.

For any subset $\mathcal{A}$ of $L^2(\Omega, \mathcal{F}, P)$, we denote by $[\mathcal{A}]$ the closed subspace of $L^2(\Omega, \mathcal{F}, P)$, generated by all elements of $\mathcal{A}$.

To obtain the non-linear predictor $\hat{X}(\nu)=E(X(\nu)|\sigma(X(l) ; l \leq 0))$ is reduced to getting a projection of $X(\nu)$ ($\nu \in \mathbb{N}$) as follows:

LEMMA 4.1 (Masami-Wiener [4]).

(i) $E(X(\nu)|\sigma(X(l) ; l \leq 0)) = P_{\mathcal{A}^\nu_{\omega_2}}X(\nu)$ ($\nu \in \mathbb{N}$),

where $\mathcal{A}^\nu_{\omega_2} = \left\{ 1, \Pi \left( X(n_k)^{n_k} ; m \in \mathbb{N}^*, p_k \in \mathbb{N}, n_k \in \mathbb{Z} (0 \leq k \leq m), n_0 < \cdots < n_m \leq 0 \right) \right\}$. 

(ii) $\left\{ 1, \Pi \left( X(n_k)^{n_k} ; m \in \mathbb{N}^*, p_k \in \mathbb{N}, n_k \in \mathbb{Z} (0 \leq k \leq m), n_0 < \cdots < n_m \leq 0 \right) \right\}$ is linearly independent in $L^2(\Omega, \mathcal{F}, P)$.

We shall obtain certain computable algorithm for $\hat{X}(\nu)$. For that purpose, we shall show the following lemma.
LEMMA 4.2.

\[ E(X(\nu) | \sigma(X(t) ; t \leq 0)) = P_{X^{\infty}} X(\nu) \quad (\nu \in \mathbb{N}), \]

where

\[ K^{\infty} = \left[ \prod_{k=0}^{m} X(n-k)^p_k - E \left( \prod_{k=0}^{m} X(n-k)^p_k \right) ; m \in \mathbb{N}^*, n \leq 0, \right. \]
\[ p_0 \in \mathbb{N}, p_k \in \mathbb{N}^* \quad (1 \leq k \leq m) \].

PROOF. By Lemma 4.1(i), what we need to prove is that \( P_{X^{\infty}} X(\nu) = P_{X^{\infty}} X(\nu) \) for any \( \nu \in \mathbb{N} \). For any \( m \in \mathbb{N}^*, n \leq 0, p_0 \in \mathbb{N}, p_k \in \mathbb{N}^* \quad (1 \leq k \leq m) \), there exist \( M \in \mathbb{N}^*, q_i \in \mathbb{N}, n_i \in \mathbb{Z} \quad (0 \leq i \leq M), n_0 < \cdots < n_M \leq 0 \) such that

\[ \prod_{k=0}^{m} X(n-k)^p_k = \prod_{i=0}^{M} X(n_i)^{q_i}, \]

it can be seen that

\[ \mathcal{M}^{\infty} \subseteq K^{\infty} = [1]. \]

Therefore, we see that \( P_{X^{\infty}} X(\nu) = P_{X^{\infty}} X(\nu) = E(X(\nu)) = 0 \). Thus, it follows that Lemma 4.2 holds. (Q. E. D.)

For the purpose of parametrizing the infinite-dimensional subspace \( K^{\infty} \), we define a subset \( A \) of \( \{0, 1, 2, \ldots\}^\mathbb{N}^* \) by

\[ A = \{ \mathbf{p} = (p_0, p_1, p_2, \ldots) \in \{0, 1, 2, \ldots\}^\mathbb{N}^* ; p_0 \geq 1 \text{ and there exists } m \in \mathbb{N}^* \text{ such that } p_m \neq 0, p_k = 0 \quad (k \geq m+1) \}. \]

For any \( \mathbf{p} \in A \), a one-dimensional strictly stationary time series \( \varphi_{\mathbf{p}} = (\varphi_{\mathbf{p}}(n) ; n \in \mathbb{Z}) \) is introduced by

\[ \varphi_{\mathbf{p}}(n) = \prod_{k=0}^{\infty} X(n-k)^{p_k} \]

and we set

\[ G = \{ \varphi_{\mathbf{p}} ; \mathbf{p} \in A \}. \]

We shall order the elements of \( G \) to arrange them in a sequence \( \{ \varphi_j ; j \in \mathbb{N}^* \} \). For each \( q \in \mathbb{N} \), we define a subset \( A_q \) of \( A \) and a subset \( G^{(q)} \) of \( G \) by

\[ A_q = \{ \mathbf{p} = (p_0, p_1, \ldots) \in A ; q = \sum_{k=0}^{\infty} (k+1) \cdot p_k \} \quad \text{and} \quad G^{(q)} = \{ \varphi_{\mathbf{p}} ; \mathbf{p} \in A_q \}. \]

Then we have the disjoint union

\[ G = \bigcup_{q \in \mathbb{N}} G^{(q)}. \]

Now we shall order the elements of \( G \). For any \( \varphi_{\mathbf{p}} \in G^{(q)} \) and \( \varphi_{\mathbf{p}'} \in G^{(q')} \), we say that \( \varphi_{\mathbf{p}} \) precedes \( \varphi_{\mathbf{p}'} \) if and only if \( q < q' \) or \( q = q' \) and in addition, there
exists \( k_0 \in \mathbb{N}^* \) such that \( p_k = p_{k_0}^* (0 \leq k \leq k_0 - 1) \) and \( p_{k_0} > p_{k_0}^* \). Then we have

\[
G = \{ \varphi_j ; j \in \mathbb{N}^* \}
\]

and

\[
G^{(q)} = \{ \varphi_{d_0 + 1}, \varphi_{d_0 + 2}, \ldots, \varphi_{d_q} \},
\]

where

\[
d_q = \text{the number of } \left\{ \bigcup_{r=1}^{q} G^{(r)} \right\} - 1
\]

and

\[
(\varphi_{d_0 + 1}(n), \varphi_{d_0 + 2}(n), \ldots, \varphi_{d_q}(n)) = (X(n)^0, X(n)^1 X(n-1), \ldots, X(n)X(n-q+2)).
\]

For example,

\[ (d_1, d_2, d_3, d_4) = (0, 1, 3, 6) \]

and

\[
(\varphi_0(n), \varphi_1(n), \varphi_2(n), \varphi_3(n), \varphi_4(n), \varphi_5(n), \varphi_6(n)) = (X(n), X(n)^{2} X(n-1), X(n)^{3} X(n-1), X(n)^{4} X(n-1), X(n)X(n-2)).
\]

By using the system \( G = \{ \varphi_j ; j \in \mathbb{N}^* \} \), we define \( X^{(q)}(n); n \in \mathbb{Z} \) and \( Y^{(q)}(n); n \in \mathbb{Z} \) by

\[
X^{(q)}(n) = \begin{pmatrix}
\varphi_0(n) - E(\varphi_0(n)) \\
\varphi_1(n) - E(\varphi_1(n)) \\
\vdots \\
\varphi_{d_q}(n) - E(\varphi_{d_q}(n))
\end{pmatrix}
\]

and

\[
Y^{(q)}(n) = \begin{pmatrix}
\varphi_{d_0 + 1}(n) - E(\varphi_{d_0 + 1}(n)) \\
\varphi_{d_0 + 2}(n) - E(\varphi_{d_0 + 2}(n)) \\
\vdots \\
\varphi_{d_q}(n) - E(\varphi_{d_q}(n))
\end{pmatrix}.
\]

Then, by virtue of Lemma 4.1(ii), we have the following lemma.

**Lemma 4.3.**

(i) For any \( q \in \mathbb{N} \), \( X^{(q)}(n) \) is a \( d_0 + 1 \)-dimensional weakly stationary time series satisfying condition (2.30).

(ii) \( X^{(1)} = X \).

(iii) \( X^{(q)}(n) = \begin{pmatrix} X^{(q-1)}(n) \\ Y^{(q)}(n) \end{pmatrix} \) \( (q = 2, 3, \ldots) \).

(iv) \[ \left[ \bigcup_{N=0}^{\infty} \bigcup_{q=1}^{\infty} L_0^0(X^{(q)}) \right] = \mathcal{K}_0^{\infty} \).

We shall show how the non-linear predictor of \( X \) is expressed by using the
Application of the theory of $\text{KM}_2\text{O}$-Langevin equations

THEOREM 4.1. For any $\nu > 0$,
\[ E(X(\nu) | \sigma(X(l); l \leq 0)) = \text{the first component of } \lim_{N \to \infty} \left( \sum_{k=0}^{N} Q_{\nu}(X^{(q)})(N+\nu, N; N-k)X^{(q)}(-k) \right). \]

PROOF. By Lemmas 4.2 and 4.3(iv), we have
\[ E(X(\nu)) | \sigma(X(l); l < 0)) = \lim_{N \to \infty} P_{X^{(q)}(N)} X^{(q)} \]
\[ = \text{the first component of } \lim_{N \to \infty} P_{X^{(q)}(N)} X^{(q)}(\nu). \]
By applying the prediction formula (2.25+) to the time series $X^{(q)}$, we have
\[ P_{X^{(q)}(N)} X^{(q)}(\nu) = U(-N)P_{X^{(q)}(N)}^{0} X^{(q)}(N+\nu) \]
\[ = U(-N) \left( \sum_{k=0}^{N} Q_{\nu}(X^{(q)})(N+\nu, N; k)X^{(q)}(-k) \right) \]
\[ = \sum_{k=0}^{N} Q_{\nu}(X^{(q)})(N+\nu, N; k)X^{(q)}(-k) \]
\[ = \sum_{k=0}^{N} Q_{\nu}(X^{(q)})(N+\nu, N; N-k)X^{(q)}(-k), \]
where $U(-N)$ is a unitary operator from $\mathcal{L}_{X^{(q)}}(X^{(q)})$ to $\mathcal{L}_{N}(X^{(q)})$ such that $U(-N)X^{(q)}(n)=X^{(q)}(n-N) \ (0 \leq n \leq N)$. Therefore, we get Theorem 4.1. (Q.E.D.)

We shall explain the structure of algorithm computing the coefficients $Q_{\nu}(X^{(q)})(\cdot, *, *) \ (q \in \mathbb{N})$ in Theorem 4.1. Let $\mathcal{D}(X^{(q)})$ (resp. $\mathcal{D}(X^{(q-1)})$ and $\mathcal{D}(Y^{(q)})$) be the $\text{KM}_2\text{O}$-Langevin data associated with $X^{(q)}$ (resp. $X^{(q-1)}$ and $Y^{(q)}$). By (2.27+),
\[ Q_{\alpha}(X^{(q)})(m, n; k) = \sum_{l=n+1}^{m-1} \gamma_{\alpha}(X^{(q)})(m, l)Q_{\alpha}(X^{(q)})(l, n; k) - \gamma_{\alpha}(X^{(q)})(m, k), \]
which implies that, for each fixed $q \in \mathbb{N}$, $Q_{\alpha}(X^{(q)})(\cdot, *, *)$ can be calculated from $\mathcal{D}(X^{(q)})$. By virtue of FDT, $\mathcal{D}(X^{(q)})$ can be recursively calculated from the $\text{KM}_2\text{O}$-Langevin partial correlation functions $\delta_\alpha(X^{(q)})(\cdot)$. By applying Theorem 3.1 to the time series $X^{(q)}$, we obtain an algorithm computing $\delta_\alpha(X^{(q)})(\cdot)$ in Theorem 4.2. The crux is that the $\delta_\alpha(X^{(q)})(\cdot)$ can be calculated from $\mathcal{D}(X^{(q-1)})$, $\mathcal{D}(Y^{(q)})$ and $R^{X^{(q-1)}Y^{(q)}} (q=2, 3, \ldots)$. 

THEOREM 4.2. For any $n, q \in \mathbb{N}$, $2 \leq q$, 

\[ \delta_s(X^{(q)}) (n) = \begin{pmatrix} \delta_s(X^{(q-1)}) (n) V_s(X^{(q-1)}) (n-1) & 0 \\ 0 & \delta_s(Y^{(q)}) (n) V_s(Y^{(q)}) (n-1) \end{pmatrix} \]

\[ - \sum_{k=0}^{n-1} \gamma_s(X^{(q)}) (n-1, k). \]

\[ B_s(X^{(q-1)} | Y^{(q)}) (k+1, n) \]

\[ B_s(Y^{(q)} | X^{(q-1)}) (k+1, n) \]

\[ V_s(X^{(q)}) (n-1)^{-1}, \]

where

\[ \gamma_s(X^{(q)}) (j, j) = 1 \quad \text{and} \quad \gamma_s(X^{(q)}) (j, j) = 1 \quad (j \in \mathbb{N}^*). \]

Finally we shall make a comment concerning the global behavior of the prediction functions \( Q_s(X^{(q)}) (N + \nu, N; N-k) \) as \( N \to \infty \) in order to complete the representation for the non-linear predictor in Theorem 4.1. For that purpose, we need the following stronger condition (H.3) than (H.2), besides (H.1):

(H.3) For each \( q \in \mathbb{N} \), the weakly stationary process \( X^{(q)} \) has the spectral density matrix function \( \Delta(X^{(q)}) (\theta) \) defined on \( [-\pi, \pi) \) such that

\[ \log(\det(\Delta(X^{(q)}))) \in L^1([-\pi, \pi]). \]

By Theorems 4.2, 5.1 and 5.2 in [7], we find that, for each \( q \in \mathbb{N} \), the following limits exist:

\[ \gamma_s(X^{(q)}) (k) \equiv \lim_{n \to \infty} \gamma_s(X^{(q)}) (n, n-k) \quad (k \in \mathbb{N}^*); \]

\[ P_s(X^{(q)}) (k) \equiv \lim_{n \to \infty} P_s(X^{(q)}) (n, n-k) \quad (k \in \mathbb{N}^*). \]

Moreover they satisfy the following recursive relations: for any \( k \in \mathbb{N} \),

\[ \begin{cases} P_s(X^{(q)}) (0) = V_s(X^{(q)}) / \sqrt{3} \\ P_s(X^{(q)}) (k) = - \sum_{l=0}^{k-1} \gamma_s(X^{(q)}) (k-l) P_s(X^{(q)}) (l). \end{cases} \]

By virtue of Theorem 6.5 in [7], we can theoretically obtain the algorithms for the limits as \( N \to \infty \) of the prediction functions \( Q_s(X^{(q)}) (N + \nu, N; N-k) \) for any \( q, \nu \in \mathbb{N}, \ k \in \mathbb{N}^* \): the limits

\[ Q_s(X^{(q)}) (\nu, k) \equiv \lim_{N \to \infty} Q_s(X^{(q)}) (N + \nu, N; N-k) \]

exist and they satisfy the following recursive relations:

\[ Q_s(X^{(q)}) (\nu, k) = - \sum_{l=1}^{\nu-1} \gamma_s(X^{(q)}) (\nu-l) Q_s(X^{(q)}) (l, k) - \gamma_s(X^{(q)}) (\nu+k). \]
Application of the theory of $K\text{M}_2\text{O}$-Langevin equations

References


Yasunori Okabe
Department of Mathematics
Faculty of Science
Hokkaido University
Sapporo 060
Japan
Present Address
Department of Mathematical Engineering
and Information Physics
Faculty of Engineering
University of Tokyo
Tokyo 113
Japan

Takashi Ootsuka
Department of Mathematics
Faculty of Science
Hokkaido University
Sapporo 060
Japan
Present Address
High School of
Abashiri Minamigaoka
Abashiri 093
Japan