PERMANENT-TRANSITORY DECOMPOSITION IN VAR MODELS WITH COINTEGRATION AND COMMON CYCLES†

Alain Hecq, Franz C. Palm and Jean-Pierre Urbain

I. INTRODUCTION

The purpose of this paper is to derive permanent-transitory decompositions of nonstationary multiple times series generated by a finite order Gaussian VAR(\(p\)) model with both cointegration and serial correlation common features. For cointegrated processes (without common cyclical features restrictions), several permanent-transitory decompositions have been extensively used in empirical and theoretical analyses: these include the multivariate extension of the Beveridge-Nelson (B-N) decomposition proposed by Stock and Watson (1988), the observable permanent-transitory decomposition of Gonzalo and Granger (1995) where the components are identified as being combinations of the observable series. Engle and Kozicki (1993) introduced the concepts of serial correlation common features in cointegrated VAR models and Vahid and Engle (1993) gave conditions under which the cyclical component of multiple time series (≡ transitory component) can be represented by a reduced number of common stochastic cycles. The decompositions into permanent and transitory components in the latter framework were initially obtained under the strong assumption that the number of common cycles arising from common cyclical features and the number of common trends arising from cointegration add up to the number of variables in the system. In this case, the Stock-Watson and the Granger-Gonzalo decompositions coincide (Proietti, 1997). An interesting extension was recently obtained by Proietti (1997) who showed how feasible decompositions could be obtained even in the case where the sum of the number of common feature vectors and the number of cointegrating vectors is less than the dimension of the system.

In this paper we extend the aforementioned analyses to the finite order

†A previous version of this paper was presented at NASM98, ESEM98 and at the 1998 TI Workshop on Time Series Econometrics held in Rotterdam. The authors gratefully acknowledge the conference participants, Tommaso Proietti and three anonymous referees for helpful comments and suggestion on an earlier version of this paper. The usual disclaimer applies.
Gaussian VAR models with two cointegration and common cyclical features structures recently analyzed by Hecq, Palm and Urbain (1998). In the standard case both long-run and short-run dynamics matrices are nested in such a way that linear combinations of the variables in first differences are innovation processes w.r.t. to past information set. The second class of structures they considered arises when linear combinations of the variables in first differences corrected for the long run (cointegration) relations are innovation processes w.r.t. to past information. The framework is similar to that of Vahid and Engle (1993), but less restrictive as they explicitly consider linear combinations of the first differenced I(1) variables that are allowed to be predictable at low frequencies. The implications of both cases for the number of common cycles are quite different. In the standard case, the upper bound on the number of common feature vectors is given by the number of common trends. In the latter case the upper bound is simply the number of time series minus one. The former and latter models are labelled strong form (SF) and weak form reduced rank structures (WF hereafter) respectively. We believe that in applications such as business cycle analysis it is important to dispose of the strong assumption that long-run relationships between variables and short-run dynamics are interrelated. In particular, an interesting issue is concerned with the question whether the observed variables in first differences corrected for the long-run relationships exhibit common features. A macroeconomic application will illustrate the usefulness of our theoretical results for time series modeling.

We propose a permanent-transitory decomposition which satisfies three criteria. Firstly, the decomposition we study is expressed in terms of observable variables and only involves quantities already available from the vector error correction model (VECM) and the estimation of common features and cointegrating vectors. The approach we follow is similar to that of Proietti (1997). Secondly, the decomposition takes into account cointegrating and common cyclical feature restrictions. An important case is when the number of common feature vectors is equal to or smaller than the number of common trends in the SF framework. We extend this analysis to the case where the number of common feature vectors has as upper bound the number of time series minus one in the WF structure. Thirdly, this decomposition should not only be a permanent-transitory decomposition but also a common trend-common cycle decomposition. Intuitively this means that common trends disappear when premultiplying the vectors of time series by the cointegrating matrix and the common feature matrix annuls the common cycles.

The paper is organized as follows. In Section II we introduce the basic assumptions, the models discussed in Hecq, Palm and Urbain (1998) and the state space representation of the cointegrated process. Permanent-transitory decompositions are derived in Section III. A small empirical illustration concludes.
II. VAR MODELS WITH COINTEGRATION AND COMMON FEATURES

Model assumptions

Consider a Gaussian Vector Autoregression of finite order $p$ (VAR($p$)) model for an $n$-dimensional I(1) vector time series $\{x_t, t = 1, \ldots, N\}$:

$$x_t = m + \sum_{i=1}^{p} \Pi_i x_{t-i} + \varepsilon_t, \quad t = 1, \ldots, N,$$  

(1)

with fixed initial values of $x_{-p+1}, \ldots, x_0$, $m$ is a vector of constants and $\varepsilon_t$ is a $n$-dimensional homoskedastic Gaussian mean innovation process relative to $\mathcal{F}_t = \{x_{t-1}, x_{t-2}, \ldots, x_1\}$ with nonsingular covariance matrix $\Omega$. Let $L$ denote the lag operator and define $\Pi(L) = I_n - \sum_{i=1}^{p} \Pi_i L^i$. We make the following assumption:

**Assumption 1 (Cointegration):** In the VAR model (1), we assume that

1. $\text{rank}(\Pi(1)) = r$, $0 < r < n$, so that $\Pi(1)$ can be expressed as $\Pi(1) = -\alpha \beta'$, with $\alpha$ and $\beta$ both $(n \times r)$ matrices of full column rank $r$;
2. the characteristic equation $|\Pi(\xi)| = 0$ has $n - r$ roots equal to 1 and all other roots outside the unit circle.

Assumption 1 implies (see Johansen, 1995) that the process $x_t$ is cointegrated of order $(1, 1)$. The columns of $\beta$ span the space of cointegrating vectors, and the elements of $\alpha$ are the corresponding adjustment coefficients or factor loadings. Decomposing the matrix lag polynomial $\Pi(L) = \Pi(1) L + \Gamma(L)(1 - L)$, with $\Gamma(L) = I_n - \sum_{j=1}^{p} \Gamma_j L^j$, $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$ ($j = 1, \ldots, p - 1$) and defining $\Delta = (1 - L)$, we obtain the vector error correction model:

$$\Delta x_t = m + \alpha \beta' x_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} + \varepsilon_t, \quad t = 1, \ldots, N,$$  

(2)

where $\alpha \beta' = -\Pi(1) = -(I_n - \sum_{j=1}^{p} \Pi_j)$. Throughout this paper we will also assume that $p$ is known. Serial correlation common features (SCCF, see Engle and Kozicki, 1993) hold for the VECM (2), if there exists a $(n \times s)$ matrix $\tilde{\beta}$, whose columns span the cofeature space, such that $\tilde{\beta}'(\Delta x_t - m) = \tilde{\beta}' \varepsilon_t$ is a $s$-dimensional vector mean innovation process with respect to the information available at time $t$. Consequently, serial correlation common features hold if the cofeature matrix $\tilde{\beta}'$ satisfies the following two conditions:

**Assumption 2:** $\tilde{\beta}' \Gamma_j = 0_{(s \times n)}$, $j = 1 \ldots p - 1$  

(3)

**Assumption 3:** $\tilde{\beta}' \Pi(1) = -\tilde{\beta}' \alpha \beta' = 0_{(s \times n)}$  

(4)

Assumption 2 implies that $\tilde{\beta}'$ must lie in the intersection of the left null
spaces of the matrices describing the short run dynamics $\Gamma_j$, $j = 1 \ldots p - 1$. Given that $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$, $j = 1, \ldots, p - 1$ and $\Pi(1) = I_n - \sum_{j=1}^{p} \Pi_j$, Assumption 3 implies that $\beta'(I_n - \Pi_1) = 0(s \times m)$, e.g. $\Pi_1$ must have eigenvalues equal to one with multiplicity equal to $s$ and the corresponding eigenvectors must lie in the intersection of the left null spaces of the $\Gamma_j$ matrices. Cointegrated VAR models satisfying both Assumptions 2 and 3 are considered in detail in Vahid and Engle (1993). To distinguish between models that satisfy either both Assumptions 2 and 3 or Assumption 2 only, Hecq, Palm and Urbain (1998) introduce the two following definitions:

**Definition 1. (Strong Form Reduced Rank Structure):** If in addition to Assumption 1 (cointegration) both Assumptions 2 and 3 hold, the implied reduced rank structure of the VECM (2), will be labelled a strong form reduced rank structure (SF). Under SF, there exists a $(n \times s)$ matrix $\beta'$, whose columns span the cofeature space, such that $\beta'(\Delta x_t - m) = \beta' \epsilon_t$ is a $s-$dimensional vector mean innovation process with respect to $\mathcal{T}_t$.

**Definition 2. (Weak Form Reduced Rank Structure):** If in addition to Assumption 1 (cointegration) only Assumption 2 holds, the implied reduced rank structure of the VECM (2), will be labelled a weak form reduced rank structure (WF). Under WF, there exists a $(n \times s)$ matrix $\beta'$, whose columns span the cofeature space, such that $\beta'(\Delta x_t - m - \alpha \beta' x_{t-1}) = \beta' \epsilon_t$ is a $s$-dimensional vector mean innovation process with respect to $\mathcal{T}_t$.

The implications of these two classes of models are discussed in Hecq, Palm and Urbain (1998) where inferential issues are investigated and a mixed form (MF) is also proposed. The contribution of this paper is to consider permanent-transitory decompositions for these two classes of models, extending herewith the existing work considered by Proietti (1997). At this stage it is already important to note an important difference between SF and WF. In the latter case both the possible number $s$ may be greater than $n - r$ but has to remain $\leq n - 1$ and the corresponding $n - s$ common dynamic factors consist of linear combinations of the lagged first differences only. Notice that we could easily extend these definitions to the case where only part of the short-run components of the models disappears. This type of reduced rank structures has been previously mentioned by Ahn and Reinsel (1988) for stationary processes, Tiao and Tsay (1989) for VARMA models and by Reinsel and Ahn (1992) and Ahn (1997) for cointegrated VAR processes. Notice that the WF restrictions are generally not invariant to alternative vector error correction representations such as those where $y_{t-p}$ appears in levels instead of $y_{t-1}$. The implications of the lack of invariance are discussed in more details in Hecq, Palm and Urbain (1998).
III. PERMANENT-TRANSITORY DECOMPOSITIONS

We consider the following multivariate version of the Beveridge-Nelson (B-N) permanent-transitory (P-T) decomposition:

**Definition 3.** *(B-N Permanent-Transitory Decomposition)* Let \( x_t \) be a \( n \)-dimensional integrated process of order one generated by (2). A permanent-transitory decomposition of \( x_t \) is a pair of processes \( (\mu_t, \psi_t) \), such that:

1. \( \mu_t \) is a random walk process while \( \psi_t \) is a covariance stationary process,
2. \( \text{Var}(\Delta \mu_t) \) and \( \text{Var}(\Delta \psi_t) \) are strictly positive,
3. \( x_t = \mu_t + \psi_t \),
4. both \( \mu_t \) and \( \psi_t \) should be functions of the observable variables,
5. if there exists a cointegrating matrix \( \beta \) satisfying Assumption 1 and a cofeature matrix \( \tilde{\beta} \) satisfying either both Assumption 2 and 3 or Assumption 3 only, then \( \beta \psi_t = 0 \) and \( \tilde{\beta}' \mu_t = 0 \).

The definition is in line with those used by Quah (1991) and Gonzalo and Granger (1995) with the exception that we require \( \mu_t \) to be a random walk while Quah (1991) and Gonzalo and Granger (1995) assume \( \mu_t \) to be difference-stationary. \( \mu_t \) is then called the (common) permanent component of \( x_t \) while \( \psi_t \) represents the (common) transitory covariance stationary component of \( x_t \). In line with the existing related literature the latter component will often be labelled the common cycle although one should recognize that situations may occur where this terminology is not appropriate.\(^1\) The decomposition may also be labelled orthogonal if we add the requirement that \( \Delta \mu_t \) is uncorrelated with \( \psi_t \) at all leads and lags. Contrary to Quah (1991) we will not require the last property to hold in general. It is interesting however to contrast this definition with that underlying the Gonzalo-Granger (1995) definition of a P-T decomposition. The major difference comes from the random walk nature of the permanent component and from \((v)\) which requires that the cointegrating matrix annihilates the common stochastic trends of the multivariate process \( x_t \), while the cofeature matrix should annihilate the transitory (cyclical) component. The Gonzalo and Granger (1995) definition requires the common factors to be linear in the observable variables and also requires the change in the permanent component not to be Granger-caused by the transitory component. This last requirement is satisfied under \((i)-(v)\). Notice also under \((i)-(v)\) the decomposition is not only a permanent-transitory decomposition but also a common permanent-transitory decomposition.

\(^1\)To avoid confusion, it should be noticed from the outset that the term common cyclical features refers to a particular type of commonality leading to specific reduced rank structures. This concept should not be confused with the concept of cycle used in business cycle analyses (see the discussion in Cubadda, 1999). On the other hand, the concept of common cycles (in contrast to common cyclical features) refers to the common transitory component in particular permanent-transitory decompositions.
Cointegrated VAR Models

Let us first rewrite the VECM (2) in a state space form such that

\[ \Delta x_t = Z f_t \]  

\[ f_t = c + T f_{t-1} + Z' \epsilon_t \quad t = 1, \ldots, N \]  

with

\[
\begin{bmatrix}
\Delta x_t \\
\Delta x_{t-1} \\
\vdots \\
\Delta x_{t-p+2} \\
\beta' x_{t-1} 
\end{bmatrix} =
\begin{bmatrix}
\Gamma_1 + \alpha \beta' & \Gamma_2 & \cdots & \Gamma_{p-1} & \alpha \\
I_n & 0_{n \times n} & \cdots & 0_{n \times n} & 0_{n \times r} \\
0_{n \times n} & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\beta' & 0_{r \times n} & \cdots & \cdots & I_r 
\end{bmatrix}
\]

where \( f_t \) a \((n(p-1)+r) \times 1\) vector, \( Z = [I_n, 0_{n \times n}, \ldots, 0_{n \times r}] \) a \((n \times (n(p-1)+r))\) matrix, \( T \) a \((n(p-1)+r) \times (n(p-1)+r)\) matrix, and \( c' = [m', 0_{(1\times n)}, \ldots, 0_{(1\times n)}, 0_{(1\times r)}] \) a vector of dimension \( 1 \times (n(p-1)+r) \). (5) is the measurement equation while (6) is the transition equation.\(^2\) Stability conditions for the state-space representation (5)–(6) are satisfied under Assumption 1 and the condition that \( \Gamma(1) - \alpha \beta' \) and \( \beta' (\Gamma(1) - \alpha \beta')^{-1} \alpha \) have rank \( n \) and \( r \) respectively. The nonsingularity of \( \Gamma(1) - \alpha \beta' \) is satisfied under the assumption that the variables are at most integrated of order one and hence these rank conditions play the same role as the condition that \( \alpha \Gamma(1) \beta \) is of rank \( n-r \) in Johansen (1995), see also Proietti (1997).

Consider the non-zero drift case\(^3\) (i.e. when the unconditional mean \( E(\Delta x_t) \neq 0 \)) of the multivariate Beveridge-Nelson (B-N) decomposition where the trend (\( \mu_t \equiv \) permanent component) is represented by a random walk and the cycle (\( \psi_t \equiv \) transitory component) by a weakly stationary process:

\[ x_t = \mu_t + \psi_t, \]  

where

\[
\mu_t = x_t + \lim_{l \to \infty} \sum_{i=1}^l \Delta \tilde{x}_{t+i|t} - E(\Delta x_t) 
\]

is the trend component, that is the value the series would take if it were on

\(^2\)Note that Proietti (1997) writes the VAR in an interim multiplier representation system while we keep the form with the error correction term lagged of one period. It turns out that the results are identical but keeping the error correction at lag one is more in line with Definitions 1 and 2.

\(^3\)The most interesting cases arise when the drift is non-zero, as it allows for an autonomous upward or downward movement in the data which is often encountered in empirical work. The driftless model is thus just a particular case of the more general one.
its long-run path. $\Delta \tilde{x}_{t+1|t}$ denotes the $i$th-step ahead best linear unbiased forecast of $\Delta x_t$ based on the information set $\mathcal{J}_t$ that is the optimal prediction in the minimum mean square sense. Following Proietti (1997) and Harvey (1989) we assume that the mean rate of the drift is constant over the sample. By successive substitution of (6) we get the expected value of the drift $c_t^* = (\sum_{i=0}^{N-1} T^i)c$ which under the stability conditions is equal to $(I_{(n(p-1)+r)} - T)^{-1}c = c^*$ when $N \to \infty$. We then transfer both $c$ and $c^*$ to the measurement equation to get:

$$\Delta x_t = Z f_t^* + Z c^*$$

(9)

$$f_t^* = T f_{t-1}^* + Z^T \epsilon_t, \quad t = 1, \ldots, N,$$

(10)

where $f_t^* = f_t - c^*$. Since $\Delta \tilde{x}_{t+1|t} = Z T f_{t|t}^*$, where $f_{t|t}^*$ is the updated estimate of the state vector yielded by the Kalman filter, we know that under the stability conditions the expression $\sum_{i=1}^{\infty} T^i$ converges to $(I_{(n(p-1)+r)} - T)^{-1} T$ and $\psi_t = -Z(I_{(n(p-1)+r)} - T)^{-1} T f_{t|t}^*$. Since the components of $f_t^*$ have already been observed at time $t$, we have $f_{t|t}^* = f_t^*$. After some straightforward but tedious matrix manipulations we may state the following proposition for the VECM representation of a VAR. The proposition restates Proietti’s (1997) result for the interim multiplier representation.

**Proposition 1.** Under Assumption 1 and for the VECM (2) case, the decomposition: $x_t = \mu_t + \psi_t$, provides a B-N permanent ($\mu_t$) – transitory ($\psi_t$) decomposition of $x_t$ where

$$\psi_t = -(I - P)(\Gamma(1) - \alpha \beta')^{-1} \Gamma^*(L) \Delta x_t + P x_t$$

$$= \psi_{2t} + \psi_{1t}$$

(11)

$$\mu_t = (I - P)(\Gamma(1) - \alpha \beta')^{-1} \Gamma(L) x_t,$$

(12)

where $\Gamma(L) = I_n - \Gamma_1 L - \cdots - \Gamma_{p-1} L^{p-1} = \Gamma(1) + \Delta \Gamma^*(L)$, with $\Gamma(1) = (I_n - \Gamma_1 - \cdots - \Gamma_{p-1})$, $\Gamma^*(L) = \Gamma_0^* + \Gamma_1^* L + \cdots + \Gamma_{p-2}^* L^{p-2}$, $\Gamma_j = \sum_{i=j+1}^{p-1} \Gamma_i$ and where

$$P = (\Gamma(1) - \alpha \beta')^{-1} \alpha [\beta' (\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta'$$

is a $(n \times n)$ matrix satisfying the following set of properties

\[\text{Notice that the only difference between the drift and the driftless case is that the updated estimate of the state at time } t \text{ is equal to } f_{t|t}^* = f_t - c^*. \text{ That means that the unconditional mean, which is constant over the sample, is substracted from the observations.}
\]

\[\text{A detailed proof is available upon request. In the sequel, we will always denote the orthogonal complement of any } n \times s \text{-dimensional matrix } B, \text{ with } n > s \text{ and rank}(B) = s, \text{ by the } n \times (n-s) \text{ matrix } B_\perp \text{ such that } B' B_\perp = 0 \text{ with rank}(B_\perp) = n-s \text{ and rank}(B' B_\perp) = n. \text{ We then say that } B_\perp \text{ spans the left null space of } B \text{ and } B' \text{ spans the left null space of } B_\perp.\]
1. \( \beta' P = \beta' \)
2. \( P(\Gamma(1) - \alpha\beta')^{-1}\alpha = (\Gamma(1) - \alpha\beta')^{-1}\alpha \)
3. \( P^i = P \)
4. \( \alpha'_\perp \Gamma(1)P = 0 \)
5. \( P\beta_\perp = 0 \)
6. \( \text{rank}(P) = r \) and \( \text{rank}(I - P) = n - r \)

The proof is similar to the one proposed in Proietti (1997) and omitted here to save space. Using the expressions for \( c^* = (I_{(n(p-1)+r)} - T)^{-1}c \) we obtain:

\[
E(\Delta x_t) = E(\Delta x_{t-1}) = \cdots E(\Delta x_{t-p+1}) = Zc^* = (I - P)(\Gamma(1) - \alpha\beta')^{-1}m
\]

so that the trend/cycle decomposition for the non-zero drift case is obtained by replacing \( \Delta x_t \) by \( \Delta x_t - Zc^* \) in equation (11). We can either subtract the empirical mean of \( P^i x_t \) in equation (11) or use the expression \((\Gamma(1) - \alpha\beta')^{-1}a\beta'(\Gamma(1) - \alpha\beta')^{-1}a^t\) in (14). In order to make the common trends appearing more explicitly, we use the fact that \( \alpha(a'\alpha)^{-1}a' + \alpha'_\perp(a'_\perp a'_\perp)^{-1}a'_\perp = I_n \), insert this expression in front of \( \Gamma(L)x_t \) in (12) and by property 2 of the \( P \) matrix given in Proposition 1 we obtain:

\[
\mu_t = (I - P)(\Gamma(1) - \alpha\beta')^{-1}[\alpha(a'\alpha)^{-1}a' + \alpha'_\perp(a'_\perp a'_\perp)^{-1}a'_\perp] \Gamma(L)x_t
\]

\[
= (I - P)(\Gamma(1) - \alpha\beta')^{-1}\alpha'_\perp(\alpha'\alpha)^{-1}\alpha'_\perp \Gamma(L)x_t
\]

Hence the common trends are given by \( \alpha'_\perp \Gamma(L)x_t \) (see inter alia Johansen, 1995) and the corresponding loadings are \((I - P)(\Gamma(1) - \alpha\beta')^{-1} \times \alpha'_\perp(\alpha'\alpha)^{-1}\alpha'_\perp \Gamma(L)x_t \). The former part is specific to cointegrated processes and has the cointegrating relationships as cycle generators while the latter are the short-run fluctuations in a stricter sense. If we take the first difference of \( \mu_t \) and substitute (2) for \( \Gamma(L)\Delta x \) we obtain:

\[
\Delta \mu_t = (I - P)(\Gamma(1) - \alpha\beta')^{-1}\alpha'_\perp(\alpha'\alpha)^{-1}\alpha'_\perp \Gamma(L)\Delta x_t
\]

\[
= (I - P)(\Gamma(1) - \alpha\beta')^{-1}\alpha'_\perp(\alpha'\alpha)^{-1}\alpha'_\perp(\alpha\beta'x_{t-1} + m + \epsilon_t)
\]

\[
= (I - P)(\Gamma(1) - \alpha\beta')^{-1}\alpha'_\perp(\alpha'\alpha)^{-1}\alpha'_\perp(m + \epsilon_t)
\]

which shows that the trend component of the decomposition is a random walk with \( m \neq 0 \) or without \( m = 0 \) drift. This last expression also shows that any perturbation of the transitory component \( \psi_t \) does not have any long-run effects on the permanent component.

In the next subsection we take the common cyclical feature restrictions
into account and distinguish between the two models introduced in the previous section.

**WF versus SF**

The previous section restated the results of Proietti (1997) which allow to get a B-N type decomposition where the permanent and the transitory parts are linear combinations of observable variables. Now we show how to take into account common feature restrictions both in SF and WF cases.

In order to impose common feature restrictions, consider the orthogonal projection of the transitory component $\psi_t$ into the space spanned by $\beta$ and its orthogonal complement. We therefore postmultiply the orthogonal matrix $\beta' \beta^{-1} \beta' + \beta' \beta_{\perp}^{-1} \beta_{\perp} = I_n$ by the transitory part $\psi_t = \psi_{2t} + \psi_{1t}$. This transitory part is now formed by four elements, i.e. $\psi_t = \psi_{2At} + \psi_{1At} + \psi_{2Bt} + \psi_{1Bt}$ with $\beta$ and $\beta_{\perp}$ being involved in the $A$ and $B$ parts respectively. Propositions 2 and 3 show which part may be simplified in the SF and the WF cases respectively.

**Proposition 2.** If Assumptions 1, 2 and 3 hold (SF), the B-N decomposition of $x_t$ is given by

$$
\psi_{S,t} = \psi_{2Bt} + \psi_{1Bt}
$$

(17)

$$
= -\tilde{\beta}_{\perp} (\tilde{\beta}_{\perp}' \beta_{\perp})^{-1} \tilde{\beta}_{\perp}' (I - P)(\Gamma(1) - \alpha \beta')^{-1} \Gamma^*(L) \Delta x_t
$$

$$
+ \tilde{\beta}_{\perp} (\tilde{\beta}_{\perp}' \beta_{\perp})^{-1} \tilde{\beta}_{\perp}' P x_t
$$

$$
\mu_{S,t} = x_t - \psi_{S,t} = (I - P)(\Gamma(1) - \alpha \beta')^{-1} \alpha_{\perp} (\alpha_{\perp}' \alpha_{\perp})^{-1} \alpha_{\perp}' \Gamma(L) x_t
$$

(18)

**Proof.** To show that the first two parts disappear, notice that by Assumption 2, $\beta' (\Gamma(1) - \alpha \beta') = \beta'$, $\beta' (\Gamma(1) - \alpha \beta')^{-1} = \beta' (\Gamma(1) - \alpha \beta') \times (\Gamma(1) - \alpha \beta')^{-1} = \beta'$ and $\beta' P = 0$, so that $\beta' \psi_{2At} = 0$ and $\beta' \psi_{1At} = 0$.

The decomposition presented in Proposition 2 satisfies all of the conditions in Definition 3, since the common cyclical feature restrictions are explicitly taken into account and $\beta' \psi_{S,t} = 0$ by definition of orthogonal complements.

**Proposition 3.** If Assumptions 1 and 2 hold (WF), the B-N decomposition of $x_t$ is given by

$$
\psi_{W,t} = \psi_{2Bt} + \psi_{1Bt} + \psi_{2At} + \psi_{1At}
$$

(19)

$$
\mu_{W,t} = x_t - \psi_{W,t} = (I - P)(\Gamma(1) - \alpha \beta')^{-1} \alpha_{\perp} (\alpha_{\perp}' \alpha_{\perp})^{-1} \alpha_{\perp}' \Gamma(L) x_t
$$

(20)
where

\[
\psi_{2Bt} = -\tilde{\beta}_1(\tilde{\beta}_1')^{-1} \tilde{\beta}_1'(I - P)(\Gamma(1) - \alpha \beta'\Gamma^*)(L)\Delta x_t
\]

\[
\psi_{1Bt} = \tilde{\beta}_1(\tilde{\beta}_1')^{-1} \tilde{\beta}_1' P x_t
\]

\[
\psi_{2At}^* = -\tilde{\beta}(\tilde{\beta}')^{-1} \beta' \alpha[\beta'(\Gamma(1) - \alpha \beta'\Gamma^*)^{-1} \beta'(\Gamma(1) - \alpha \beta'\Gamma^*)(L)\Delta x_t
\]

\[
\psi_{1At}^* = +\tilde{\beta}(\tilde{\beta}')^{-1} \beta' [\alpha[\beta'(\Gamma(1) - \alpha \beta'\Gamma^*)^{-1} \beta' + \alpha \beta'] x_t
\]

The proof follows directly from the properties of the \( P \) matrix and the results reported in the Appendix. This decomposition is in agreement with our conditions (\( i \))–(\( iv \)) from Definition 3 and does moreover explicitly take WF common feature restrictions into account. However, from (19), we see that the decomposition of \( x_t \) into \( \psi_{W,t} \) and \( \mu_{W,t} \) does not decompose \( x_t \) into common trends-common cycles due to the presence of two additional stationary terms in \( \psi_{W,t} \), that is \( \psi_{2At}^* \) and \( \psi_{1At}^* \). These two terms do not disappear when premultiplied by \( \tilde{\beta} \) so that condition (\( v \)) which is required to obtain a common permanent-transitory decomposition of \( x_t \) is not satisfied. In this case, \( x_t \) is the sum of three different components: a common stochastic trend component (\( \mu_t \)), a (weak form) common stochastic cycles component (\( \psi_{2Bt} + \psi_{1Bt} \)) and an additional ‘uncommon’ transitory component that is not annihilated by the cofeature combinations due to the long-run predictability of the linear combinations under the WF.

In order to obtain a B-N type decomposition similar to those presented in Proposition 2 satisfying condition (\( v \)) of Definition 3, we have to consider a new variable defined as

\[
x_t^* = x_t - (\psi_{2At}^* + \psi_{1At}^*). \quad (21)
\]

We then state the following result which follows directly from the definition of \( x_t^* \):

**Proposition 4.** If Assumptions 1 and 2 hold (WF), the B-N decomposition of

\[
x_t^* = x_t - (\psi_{2At}^* + \psi_{1At}^*)
\]

is given by \( x_t^* = \mu_{W,t}^* + \psi_{W,t}^* \) where:
\[
\psi_{W,t}^* = \psi_{2Bt} + \psi_{1Bt}
\]
\[
= -\tilde{\beta}_\perp (\tilde{\beta}_\perp')^{-1} (I - P)(1 - \alpha \beta')^{-1} \Gamma^* (L) \Delta x_t
\]
\[
+ \tilde{\beta}_\perp (\tilde{\beta}_\perp')^{-1} \Gamma^* P x_t
\]
\[
\mu_{W,t}^* = x_t^* - \psi_{W,t}^* = (I - P)(1 - \alpha \beta')^{-1} \alpha_\perp' \Gamma (L) x_t
\]
\[
= \mu_{W,t}
\]

It is easily seen that the common cyclical component \( \psi_{W,t}^* \) disappears if we consider \( \tilde{\beta}' x_t^* \). Remark that the form of the common trends is identical in Propositions 3 and 4 and is given by \( \alpha_\perp' \Gamma (L) x_t \). In order to interpret \( x_t^* \), let us consider \( \tilde{\beta}' \Delta x_t^* \). According to Definition 2, we have that \( \tilde{\beta}' (\Delta x_t - \alpha \beta' x_{t-1} - m) = \tilde{\beta}' \epsilon_t \). From (21), (19) and using results from Appendix A, we have
\[
\tilde{\beta}' \Delta x_t^* = \tilde{\beta}' \Delta x_t - \tilde{\beta}' (\Gamma (1) - \alpha \beta') P (\Gamma (1) - \alpha \beta')^{-1} \Gamma (L) \Delta x_t
\]
\[
= \tilde{\beta}' \Delta x_t - \tilde{\beta}' (\Gamma (1) - \alpha \beta') P (\Gamma (1) - \alpha \beta')^{-1} (\alpha \beta' (x_{t-1} + m + \epsilon_t)
\]
\[
= \tilde{\beta}' \Delta x_t - \tilde{\beta}' (\alpha \beta' x_{t-1} + m^* + \epsilon_t^*),
\]
where the last equality follows from the property of the \( P \) matrix and where
\[
m^* = (\Gamma (1) - \alpha \beta') P (\Gamma (1) - \alpha \beta')^{-1} m
\]
\[
= -\alpha [\beta' (\Gamma (1) - \alpha \beta')^{-1} \beta' (\Gamma (1) - \alpha \beta')^{-1} m
\]
\[
= \alpha E (\beta' x_{t-1})
\]
\[
\epsilon_t^* = (\Gamma (1) - \alpha \beta') P (\Gamma (1) - \alpha \beta')^{-1} \epsilon_t.
\]

Consequently, it appears that subtracting \( \psi_{2At}^* \) and \( \psi_{1At}^* \) from the level is equivalent to adjusting the first differences \( \Delta x_t \) for the long-run relationships as was done in Definition 2, e.g. adjusting for the predictable long-run component. Because the common cycles are for \( x_t^* \) and not for \( x_t \) we call them weak form common cycles. Further, it is worth mentioning that these different decompositions are expressed in terms of observable variables and only involve quantities already available from the VECM form and the estimation of common features and cointegrating vectors.

**Relation to Other Decompositions**

It may be of some interest to briefly relate or contrast the decompositions discussed until now with some of those existing in the literature. One may first notice that a factor decomposition in the sense of Gonzalo-Granger...
(1995) is easily obtained, see also Proietti (1997), by adding the first part of \( \psi_t \) in (11) to \( \mu_t \) in (12), which yields:

\[
\mu_t^{G-G} = (I - P)(\Gamma(1) - \alpha \beta')^{-1}(\Gamma(L) - \Delta \Gamma^*(L))x_t
\]

which gives the decomposition \( x_t = Px_t + (I - P)x_t \), the first and the second r.h.s. parts being respectively the Gonzalo-Granger’s Transitory and Permanent components. Since \( \Gamma(L) - \Delta \Gamma^*(L) = \Gamma(1) \), it appears that the common trends are given by \( \alpha'_1 \Gamma(1)x_t \) and not by \( \alpha'_1 x_t \) as in Gonzalo-Granger (1995). It is also worth mentioning that this decomposition only exists under the strict restrictions that the matrix \( (\beta : \alpha_\perp) \) is full rank. While this always holds in a VAR(1) cointegrated models, it generally does not carry over to higher order cointegrated VAR systems (see for example Exercise 4.3 in Johansen, 1995). Under the SF with \( s + r = n \), the Gonzalo-Granger permanent component is given by \( \mu_t^{G-G} = (I - P)(\Gamma(1) - \alpha \beta')^{-1}\Gamma(1)x_t \) and the Gonzalo-Granger transitory component is then given by \( \psi_t^{G-G} = Px_t \), with \( x_t = \mu_t^{G-G} + \psi_t^{G-G} \). It is also worth mentioning that this decomposition only exists under the strict restrictions that the matrix \( (\beta : \alpha_\perp) \) is full rank. While this always holds in a VAR(1) cointegrated models, it generally does not carry over to higher order cointegrated VAR systems (see for example Exercise 4.3 in Johansen, 1995). Under the SF with \( s + r = n \), the Gonzalo-Granger permanent component is given by \( \mu_t^{G-G} = (I - P)(\Gamma(1) - \alpha \beta')^{-1}\Gamma(1)x_t \) and the Gonzalo-Granger transitory component is then given by \( \psi_t^{G-G} = Px_t \), with \( x_t = \mu_t^{G-G} + \psi_t^{G-G} \). Consequently the B-N and Gonzalo-Granger (1995) decompositions coincide under SF and when \( s + r = n \). In this case the space that generates the common trends is such that \( sp(\beta) = sp(\alpha_\perp) \) and so we obtain the Gonzalo-Granger common trends \( \alpha'_1 x_t \). This decomposition always trivially meets condition (v) of Definition 3 since any matrix belonging to the space generated by the columns of \( \alpha'_1 \Gamma(1) \), (or \( \alpha'_1 \) if \( s + r = n \)) annihilates the transitory component \( \psi_t^{G-G} \) without any further rank restrictions that are required to satisfy the definition of common feature used in Definitions 1 and 2. It is indeed easily seen that the ‘feature’ underlying the Gonzalo-Granger decomposition is nothing but the presence of the error correction term of the VECM which corresponds to the case where only Assumptions 1 and 3 hold.

The common factor decompositions of Escribano and Peña (1994) or Kasa (1992) such as \( x_t = \beta(\beta' \beta)^{-1}\beta'x_t + \beta_\perp(\beta'_1 \beta_\perp)^{-1}\beta'_1 x_t \) have a similar property since again any matrix belonging to the column space of \( \beta_\perp \) is a common feature matrix with respect to those decompositions without generally satisfying the standard definition of serial correlation common feature.6

IV. APPLICATION

Background

A vast amount of empirical macroeconomic literature has studied the long-run implications of the real business cycle models, see e.g. Neusser (1991),

6Notice that these decomposition are moreover not P-T decompositions in the sense of Gonzalo and Granger (1995).
King et al. (1991), Kunst and Neusser (1991), Mellander et al. (1992). With the exception of the work of Issler and Vahid (1996), little work has however been done on short-run co-movements in the neoclassical growth model. As in Issler and Vahid (1996) we analyze a small real business cycle (RBC) model with common trends and common cycles present between U.S. per capita real consumption, investment and output. We relax the hypotheses about the number of common feature and cointegrating vectors. More formally, consider the following trivariate system for the logarithms of income $y_t$, consumption $c_t$ and investment $i_t$ put forward by King, Plosser and Rebelo (1988) and analyzed by Issler and Vahid (1996):

\[ c_t = x_t^p + \bar{c} + \pi_c \hat{k}_t \] (25)

\[ i_t = x_t^p + \bar{i} + \pi_i \hat{k}_t \] (26)

\[ y_t = x_t^p + \bar{y} + \pi_y \hat{k}_t, \] (27)

where $x_t^p = x_{t-1}^p + \varepsilon_t^p$ is the common trend, that is a random walk measuring among other the impact of technology process, $\bar{y}$, $\bar{c}$ and $\bar{i}$ are the constant steady state values, $\hat{k}_t$ is the common cycle, that is the stationary transitory deviation of the capital stock from its steady state value and $\pi_c$, $\pi_i$ and $\pi_y$ are constant parameters. $\varepsilon_t^p$ and $\hat{k}_t$ may be correlated. Issler and Vahid (1996) impose the two cointegrating relationships $c_t - y_t$ and $i_t - y_t$ implied by the model (25)–(27) and test for the number of common cycle.

We analyze the period 1954:1–1996:4, that is 172 quarterly observations. Data prior to 1954:1 were used as initial observations in regressions that contain lags. Notice that we had the observations from 1948:1 but we preferred, as King et al. (1991) suggested, to exclude turbulent periods during Korean War, price control and Treasury-Fed agreement. The data used are the revised and newly published in the Survey of Current Business (May 1997) national account for the United States. The variables are $c_t$: personal consumption expenditures, $i_t$: gross private domestic investment and the output $y_t$ is the GDP less the government expenditures. The three variables have been divided by the size of the civilian population above sixteen years of age. The series are seasonally adjusted and transformed into natural log.

**Cointegration and Common Feature Analysis**

The model that best characterizes the covariance structure of the data is a VAR of order 5 (using LR statistics) with an unrestricted intercept in the short run.

Table 1 presents the test statistics (corrected for small samples) for the Johansen (1995) rank test of the number of cointegrating relationships and the 5 percent critical values. The results in Table 1. indicate that we cannot
reject the presence of two cointegrating vectors whose coefficients are not far from the theoretical ones i.e. \[c_t - y_t\] and \[i_t - y_t\] are both \(I(0)\).\(^7\) The estimated cointegrated relationships are respectively \[c_t = 0.958 (0.016) y_t\] and \[i_t = (0.045) y_t\] where asymptotic standard errors are reported in parentheses. The likelihood ratio test for unit long-run elasticities in both vectors (e.g. stationarity of the great ratios), \(\chi^2(2)\) distributed under the null, yields a value of 3.07 so that these restrictions are not rejected at the 5 percent level. A sub-sample analysis however reveals that these restrictions are not supported for the whole sample period and caused a slight shift in the transitory component. These restrictions are henceforth not imposed in the sequel.\(^8\)

For the determination of the number of common feature vectors, we use the sequential likelihood ratio test approach proposed in Hecq, Palm and Urbain (1998) based on a two-step reduced rank regression approach in which \(r\) is first determined, while ignoring restrictions on the short-run dynamics of the model. Once \(r\) is determined and \(\beta\) (the cointegrating matrix) known, \(s\) may be determined by standard reduced rank regression. The rationale behind this simple two-step analysis is that the determination of \(r\) and the estimation of \(\beta\) are not affected in terms of asymptotic efficiency by the presence of the nested reduced rank structure (see for example Ahn, 1997). Define the \(N \times n\) matrices \(W_1 = \Delta X = (\Delta x_1, \ldots, \Delta x_N)', X_1 = (x_0, \ldots, x_{N-1})', Z_1 = \Delta X^*\) with \(\Delta X^*\) being the LS residuals from the multivariate regression of \(\Delta X\) on \(X_1\).\(^9\) Define the \(N \times (n(p-1)+r)\) matrix \(W_2 = [Z_2, X_1\beta]\) with \(Z_2\) being the \(N \times n(p-1)\) matrix \((\Delta x_{N-1}, \ldots, \Delta x_{N-p+1})\). Under the maintained hypothesis of a SF reduced rank structure, the sequence of common feature Gaussian likelihood ratio test statistics for \(H_0: \text{rank}(\beta) \geq s\) against \(H_a: \text{rank}(\beta) < s\) can be shown [see Velu et al. (1986)] to be \(\xi_s = -N \sum_{i=1}^{s} \log(1 - \lambda_i), s = 1, \ldots, n-r\), where \(0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-r} < 1\) are the ordered eigenvalues of the symmetric matrix \((W_1 W_1')^{-1/2} W_1' W_2 (W_2' W_2)^{-1} W_2' W_1 (W_1 W_1')^{-1/2}\). For known \(r\) and \(\beta\), \(\xi_s\) is asymptotically \(\chi^2\)-distributed with \(s(n(p-1)+r) - s(n-s)\) degrees of freedom.

\(^7\)The traditional Augmented Dickey Fuller unit root tests statistics also strongly reject the null hypothesis. For instance for an ADF(4) we get the values \(\tau_{\mu} = -3.31, \tau_{\nu} = -3.97\) for the variable \(c_t - y_t\) and \(\tau_{\mu} = -4.14, \tau_{\nu} = -4.23\) for the variable \(i_t - y_t\).

\(^8\)We thank an anonymous referee for raising this issue.

---

### Table 1: Cointegration Tests

| \(r = 0\) | Max.Eig.Test 28.08* | 95% cv 21.0 | Trace Test 45.96* | 95% cv 29.7 |
| \(r \leq 1\) | 14.39* | 14.1 | 17.88* | 15.4 |
| \(r \leq 2\) | 3.48 | 3.8 | 3.48 | 3.8 |

---

7 The traditional Augmented Dickey Fuller unit root tests statistics also strongly reject the null hypothesis. For instance for an ADF(4) we get the values \(\tau_{\mu} = -3.31, \tau_{\nu} = -3.97\) for the variable \(c_t - y_t\) and \(\tau_{\mu} = -4.14, \tau_{\nu} = -4.23\) for the variable \(i_t - y_t\).

8 We thank an anonymous referee for raising this issue.
freedom under the null (Vahid and Engle, 1993). For the weak form a
similar approach is followed by considering the statistics
\[
W \hat{\gamma}_N P_s \hat{\lambda}_1 \log(1 - \hat{\lambda}_i), \ s = 1, \ldots, n - 1, \text{ where } 0 \leq \hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \ldots \leq \hat{\lambda}_{n-1} < 1
\]
are the ordered eigenvalues of the symmetric matrix
\[
(Z_1 Z_1)^{-1/2} Z_1 Z_2 (Z_2 Z_2)^{-1} Z_2 Z_1 (Z_1 Z_1)^{-1/2}.
\]
\(W\) has an asymptotic \(\chi^2\)-distribution with \(s(n(p - 1)) - s(n - s)\) degrees of freedom under the null. We compute results obtained both for the statistics presented above and for small sample corrected versions thereof where \(W\) and \(S\) are respectively premultiplied by the factors \((N - n(p - 1))/N\) and \((N - n(p - 1) - r)/N\), see Hecq (1999) for further details. These statistics are denoted by \(W_{cor}\) and \(S_{cor}\). For given \(r\), the likelihood ratio test statistic for the null hypothesis of a SF against the alternative of a WF, for each possible common feature rank \(s = \max(1, s - r + 1) \ldots n - r\), is given by
\[
\hat{\xi}_SW = -N \sum_{i=1}^s \log(1 - \hat{\lambda}_i)/(1 - \hat{\lambda}_i).
\]
Conditional on known \(r\) and \(\beta\), \(SW\) has an asymptotic \(\chi^2\)-distribution with degrees of freedom equal to the number of restrictions \(rs\) imposed under the \(H_0\). If the null hypothesis is rejected, one can proceed further in determining \(s\) by testing the number of zero squared canonical correlations between \(Z_1\) and \(Z_2\).

For \(p = 5\) and \(r = 2\) fixed, we obtain the tests statistics and the \(p\)-values reported in Table 2.

<table>
<thead>
<tr>
<th>(r = 2)</th>
<th>(-N \sum_{i=1}^s \ln(1 - \lambda_i))</th>
<th>df</th>
<th>(\text{Prob} &gt; \chi^2_{df})</th>
<th>(\text{Prob} &gt; \chi^2_{df})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s \geq 1)</td>
<td>23.36</td>
<td>12.06</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>(s \geq 2)</td>
<td>(54.04)</td>
<td>33.10</td>
<td>(26)</td>
<td>22</td>
</tr>
<tr>
<td>(s = 3)</td>
<td>(145.02)</td>
<td>108.53</td>
<td>(42)</td>
<td>36</td>
</tr>
</tbody>
</table>

\(^9\)A dummy was introduced in 1975:01 to capture the effect of a sharp decrease in the investment growth rate.
order to find the vectors with the meaningful economic interpretation, results in the following two cofeature relationships $\Delta c_t^* - 0.498 \Delta y_t^*$ and $\Delta i_t^* - 4.764 (0.483) \Delta y_t^*$ where asymptotic standard errors are reported in parentheses and where a * indicates that the corresponding variables are expressed in deviation from long-run effects. It is seen that these two vectors match pretty well business cycle stylized facts, e.g. that consumption is smoother than output, investment is more volatile than output and there is a single synchronous cycle.

In Figure 1 the log of the variables and their permanent components are given. The cycles in these graphs are computed using Proposition 3, that is to say, all WF common feature restrictions are imposed. However this decomposition is not a Common trend-Common cycle decomposition stricto sensus as condition (v) of the B-N decomposition is not satisfied.

Figure 2 compares for each time series the cyclical/transitory part of $x_t = (c_t, i_t, y_t)'$. Figure 3 presents the BN cycles of the original output ($y_t$) series and of the deviation of output from the long run effects ($y_t^*$). In Figure 4 the latter is compared with the transitory component obtained from the GG decomposition. The shaded areas indicate the NBER peak to trough periods. The unique (weak form) common cyclical component is given in Figure 5 with the shaded areas again indicating the NBER peak to trough periods. These graphs show the business cycles where fluctuations of investment, consumption and output are in agreement with the real business conditions.
Figure 2. Cyclical components

Figure 3. Output cycles with WF for $y_t$ and $y_t^*$
Figure 4. Output BN cycle with WF for $y_t^*$ & GG cycle

Figure 5. Single weak form common cycle
cycle model introduced above. These graphs could be used for example to date expansion and contractions periods of business cycles. Notice that the trends extracted for $c_t$ and $y_t$ in Figure 1 do not appear smoother than the observed series. Engle and Issler (1995) report a similar finding for quarterly per capita disposable income and per capita total consumption for 1953.2–1984.4. This result is due to the strong negative correlation between the permanent and transitory parts of these series.

V. CONCLUSION

In this paper we extend the Proietti (1997) Beveridge-Nelson decomposition to the case in which combinations of the first differenced variables corrected for long-run effects are white noise. In this framework, labeled weak form reduced rank structure, it is possible to have a larger number of common feature vectors than in the standard serial correlation common feature case studied initially by Vahid and Engle (1993). Similarly to the Gonzalo-Granger decomposition, the different decompositions considered in this paper are expressed in terms of observable variables and only involve quantities already available from the VECM form and the estimation of common features and cointegrating vectors. The decompositions are hence computationally easy to obtain.

An application shows the feasibility of the approach and it potential value since it leads to selecting a WF when it is more in line with economic stylized facts than a SF. The simple neo-classical real business cycle model used in this illustration is one example.

Department of Quantitative Economics
University of Maastricht

Date of receipt of Final Manuscript: May 2000.

REFERENCES


Vahid and Issler (1996) have shown that using common feature restrictions, the percentage of variance attributed to the permanent innovation is less than from the previous work by King et al. (1991) for instance. As already mentioned, these authors only allow for one common feature vector. Detailed results on the percentage of variance attributed to the permanent innovation both under SF and WF are available upon request.


APPENDIX A

Proof of Proposition 3.

Under WF, from (11)–(12) we have

\[
\psi_{w,t} = -\tilde{\beta}_\perp(\tilde{\beta}_\perp \tilde{\beta}_\perp)^{-1}\tilde{\beta}_\perp(I - P)(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
+ \tilde{\beta}_\perp(\tilde{\beta}_\perp \tilde{\beta}_\perp)^{-1}\tilde{\beta}_\perp P x_t - \tilde{\beta}(\tilde{\beta}' \tilde{\beta})^{-1}\tilde{\beta}'(I - P)(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
+ \tilde{\beta}(\tilde{\beta}' \tilde{\beta})^{-1}\tilde{\beta}' P x_t.
\]

We shall now derive the expression (19) for \(\psi^*_{2At}\) and \(\psi^*_{1At}\). First note that

\[
\tilde{\beta}'(\Gamma(1) - \alpha \beta') = \beta' = \beta' \alpha \beta' \quad \text{so that} \quad \beta' = \beta'(\Gamma(1) - \alpha \beta') + \beta' \alpha \beta',
\]

hence, \(\beta'(\Gamma(1) - \alpha \beta')^{-1} = \beta' + \beta' \alpha \beta'(\Gamma(1) - \alpha \beta')^{-1}\). From the definition of \(P\), we also have

\[
\tilde{\beta}' P = \tilde{\beta}' \alpha [\beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta'
\]

\[
+ \tilde{\beta}' \alpha \beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha [\beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta'
\]

\[
= \tilde{\beta}' \alpha [\beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta' + \tilde{\beta}' \alpha \beta'. \tag{28}
\]

Since the columns of \(\tilde{\beta}'\) span the cofeature space it holds that

\[
\tilde{\beta}'(P - I)(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
= \tilde{\beta}' \alpha [\beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta'(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
+ \tilde{\beta}' \alpha \beta'(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
- \tilde{\beta}' \alpha \beta'(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t
\]

\[
= \tilde{\beta}' \alpha [\beta'(\Gamma(1) - \alpha \beta')^{-1} \alpha]^{-1} \beta'(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t \tag{29}
\]

Multiplying respectively (28) and (29) by \(\tilde{\beta}(\tilde{\beta}' \tilde{\beta})^{-1}\), \(\psi^*_{2At}\) and \(\psi^*_{1At}\) as given in (19) can be directly obtained.

Using (28)–(29) we obtain a compact expression that can be used to compute the sum of these two components.
\[ \tilde{\beta}'(P - I)(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t + \tilde{\beta}'P x_t \]

\[ = \tilde{\beta}'\alpha[\tilde{\beta}'(\Gamma(1) - \alpha \beta')^{-1}\alpha]^{-1}\beta'(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t \]

\[ + \tilde{\beta}'\alpha[\beta'(\Gamma(1) - \alpha \beta')^{-1}\alpha]^{-1}\beta' x_t + \tilde{\beta}'\alpha \beta' x_t \]

\[ = \tilde{\beta}'(\Gamma(1) - \alpha \beta')P(\Gamma(1) - \alpha \beta')^{-1}[\Gamma^*(L)\Delta x_t + \Gamma(1)x_t] \]

\[ - \tilde{\beta}'(\Gamma(1) - \alpha \beta')P(\Gamma(1) - \alpha \beta')^{-1}\alpha \beta' x_t + \tilde{\beta}'\alpha \beta' x_t \]

(30)

Since \( P(\Gamma(1) - \alpha \beta')^{-1}\alpha = (\Gamma(1) - \alpha \beta')^{-1}\alpha \), (see property 2 of Proposition 1), (30) finally becomes

\[ \tilde{\beta}'(P - I)(\Gamma(1) - \alpha \beta')^{-1}\Gamma^*(L)\Delta x_t + \tilde{\beta}'P x_t \]

\[ = \tilde{\beta}'(\Gamma(1) - \alpha \beta')P(\Gamma(1) - \alpha \beta')^{-1}\Gamma(L)x_t \]

\[ = (\tilde{\beta}' - \tilde{\beta}'\alpha \beta')P(\Gamma(1) - \alpha \beta')^{-1}\Gamma(L)x_t \]

\[ \blacksquare \]