PLS post-processing by similarity transformation (PLS+ST): a simple alternative to OPLS

Rolf Ergon
rolf.ergon@hit.no
Telemark University College, Porsgrunn, Norway

Published in Journal of Chemometrics 2005; **19**: 1-4

Abstract

Several methods for orthogonal signal correction (OSC) based on pre-processing of the modeling data have been developed in recent years, and OPLS (orthogonal projections to latent structures) is a well known algorithm. The main result from these methods is a reduction in the number of final components in partial least squares (PLS) regression, while the predictions are virtually unchanged (identical for OPLS). This raises the question whether the same or similar results can be obtained in a more direct way using an ordinary PLS model as starting point, and as shown in the present paper this can indeed be done by use of a simple similarity transformation. This post-processing PLS+ST method is compared with OPLS, assuming a single response variable. The PLS+ST factorization of the data matrix $\mathbf{X}$ is just a similarity transformation of the non-orthogonalized PLS factorization, while OPLS is a similarity transformation of the orthogonalized PLS factorization. The predictions are therefore identical, but the residuals are somewhat different. A theoretically founded modification of the orthogonalized PLS factorization, and a corresponding modification of OPLS, leads to identical factorizations for all these methods, within similarity transformations. The PLS+ST vs. OPLS comparison also leads to an alternative post-processing method, using the ordinary PLS algorithm twice, with predetermined and permuted loading weights vectors in the second step. A limited comparison with post-processing using principal components of predictions (PCP) or canonical correlation analysis (CCA) is included.

KEYWORDS: Orthogonal signal correction; OPLS; post-processing; similarity transformation

1 Introduction

Several methods for orthogonal signal correction (OSC) have been developed in recent years, see e.g. Svensson et al. [1] for an overview and investigation. The basic OSC idea is to use a pre-processing procedure for identification and removal of variation in the regressor matrix $\mathbf{X}$ that is orthogonal to the response vector $\mathbf{y}$ (assuming the single response case), before the corrected $\mathbf{X}$ matrix is used in e.g. partial least squares (PLS) regression. As found in [1], the main result from these methods is a reduction in the number of final PLS components, while the prediction capability is virtually unchanged. This raises the question whether the same or similar results can be obtained in a more direct way using an ordinary PLS model as starting point. As shown in the paper, this can indeed be done by use of a simple similarity transformation.

A comparison with all the more or less different OSC algorithms is beyond the scope of the present paper. Instead, the OPLS (orthogonal projections to latent structures) algorithm of Trygg and Wold [2] is used as an OSC example, and the study is also limited to the single response case. Assuming an optimal number A of ordinary PLS components, we are thus concerned with the identity between

- the PLS+ST post-processing method, extracting one $\mathbf{y}$-relevant component from A original PLS components, and
- the OPLS pre-processing method, removing $A-1$ $\mathbf{y}$-orthogonal components and leaving one $\mathbf{y}$-relevant component only.

It is shown that the PLS+ST factorization of $\mathbf{X}$ is just a similarity transformation of the non-orthogonalized PLS factorization [3], while OPLS is a similarity transformation of the orthogonalized PLS factorization [3]. The predictions are therefore identical, but the residuals are somewhat different. A theoretically founded modification of the orthogonalized PLS factorization, and a corresponding modification of OPLS, leads to identical factorizations for all these methods, within similarity transformations. The PLS+ST vs. OPLS comparison also leads to an alternative post-processing method, using the ordinary PLS algorithm twice, with predetermined and permuted loading weights vectors in the second step.

The fact that OPLS and ordinary PLS predictions are identical has also been shown by Verron *et al.* [4]. A post-processing method called principal components of prediction (PCP) has earlier been presented by Langsrud and Næs [5]. Another post-processing method based on canonical correlation analysis (CCA) was recently presented by Yu and MacGregor [6], where general advantages of post-processing methods are also discussed. These PLS+PCP and PLS+CCA methods give similar although not quite the same results as the proposed PLS+ST and modified OPLS methods.

The PLS+ST method is developed in Section 2, and a comparison with OPLS is given in Section 3. A limited comparison with PLS+PCP and PLS+CCA is presented in Section 4. Details are given in a Supplementary Appendix (http://www.....).

2 The PLS+ST method

Non-orthogonalized PLS regression model

In the following we will make use of the so-called non-orthogonalized PLS factorization of Martens [3], based on modeling data in $\mathbf{X} \in \mathbb{R}^{n\times p}$ and $\mathbf{y} \in \mathbb{R}^{n\times 1}$. Assuming A components and using the orthonormal loading weights matrix $\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{W}_{2:A} \end{bmatrix} \in \mathbb{R}^{p\times A}$, the non-orthogonal score matrix $\mathbf{T} = \begin{bmatrix} \mathbf{t}_1 & \mathbf{T}_{2:A} \end{bmatrix} = \mathbf{XW} \in \mathbb{R}^{n\times A}$, and $\mathbf{q} = \begin{bmatrix} q_1 & \mathbf{q}_{2:A}^T \end{bmatrix}^T \in \mathbb{R}^{A\times 1}$, the underlying latent variables (LV) model is

$$\begin{aligned} \mathbf{y} &= \mathbf{Tq} + \mathbf{f} = \mathbf{t}_1 q_1 + \mathbf{T}_{2:A}\mathbf{q}_{2:A} + \mathbf{f} && (1) \\ \mathbf{X} &= \mathbf{t}_1\mathbf{w}_1^T + \mathbf{t}_2\mathbf{w}_2^T + \cdots + \mathbf{t}_A\mathbf{w}_A^T + \mathbf{E} = \mathbf{t}_1\mathbf{w}_1^T + \mathbf{T}_{2:A}\mathbf{W}_{2:A}^T + \mathbf{E}, && (2) \end{aligned}$$

where $\mathbf{f}$ and $\mathbf{E}$ are unmodeled residuals. A simple least squares solution results in the prediction formula [7]

$$\hat{y}_{\text{new}} = \mathbf{x}_{\text{new}}^T \hat{\mathbf{b}} = \mathbf{x}_{\text{new}}^T \mathbf{W} \left(\mathbf{W}^T\mathbf{X}^T\mathbf{X}\mathbf{W}\right)^{-1} \mathbf{W}^T\mathbf{X}^T\mathbf{y}, \tag{3}$$

and since $\mathbf{T} = \mathbf{XW}$ and $\hat{\mathbf{y}} = \mathbf{Tq}$ also in

$$\mathbf{q} = \left(\mathbf{W}^T\mathbf{X}^T\mathbf{X}\mathbf{W}\right)^{-1} \mathbf{W}^T\mathbf{X}^T\mathbf{y}. \tag{4}$$

Model transformation

The Martens factorization (2) has the special property that all score vectors except the first one are orthogonal to both $\mathbf{y}$ and $\hat{\mathbf{y}}$ (see Supplementary Appendix for proofs). This is desirable in the present context, in that the first component only contains information regarding $\mathbf{y}$. However, as shown in Fig. 1 the first score vector itself has a component orthogonal to $\hat{\mathbf{y}}$, and that has to be subtracted in order to find a first score vector in the direction of $\hat{\mathbf{y}}$. A simple way to do this is described below.

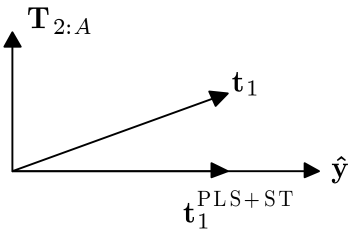

Figure 1 Score vectors in relation to $\hat{\mathbf{y}}$ for non-orthogonalized PLS factorization.

Introducing $\mathbf{M} = \begin{bmatrix} 1 & \mathbf{0} \\ q_1^{-1}\mathbf{q}_{2:A} & \mathbf{I} \end{bmatrix}$, a similarity transformation applied to the LV model (1,2) gives

$$\mathbf{y} = \begin{bmatrix} \mathbf{t}_1 & \mathbf{T}_{2:A} \end{bmatrix} \mathbf{M}\mathbf{M}^{-1} \begin{bmatrix} q_1 \\ \mathbf{q}_{2:A} \end{bmatrix} + \mathbf{f} = \mathbf{t}_1^{\mathrm{PLS+ST}} q_1 + \mathbf{f} \tag{5}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{t}_1 & \mathbf{T}_{2:A} \end{bmatrix} \mathbf{M}\mathbf{M}^{-1} \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{W}_{2:A}^T \end{bmatrix} + \mathbf{E} = \mathbf{t}_1^{\mathrm{PLS+ST}} \mathbf{w}_1^T + \mathbf{T}_{2:A} \left(\mathbf{P}_{2:A}^{\mathrm{PLS+ST}}\right)^T + \mathbf{E}, \tag{6}$$

where

$$\mathbf{t}_1^{\mathrm{PLS+ST}} = \mathbf{t}_1 + q_1^{-1}\mathbf{T}_{2:A}\mathbf{q}_{2:A}, \tag{7}$$

and where

$$\mathbf{P}_{2:A}^{\mathrm{PLS+ST}} = \mathbf{W}_{2:A} - q_1^{-1}\mathbf{w}_1\mathbf{q}_{2:A}^T \tag{8}$$

is non-orthogonal. This representation of the LV model also implies that the fitted prediction vector in accordance with Eq. (5) is $\hat{\mathbf{y}} = \mathbf{t}_1^{\mathrm{PLS+ST}} q_1$, i.e. $\mathbf{t}_1^{\mathrm{PLS+ST}}$ has the same direction as $\hat{\mathbf{y}}$, as also indicated in Fig. 1.

Prediction

Predictions may as in ordinary PLS regression be found from the formula (3). Analogously to OPLS predictions and based on $\mathbf{T}_{2:A} = \mathbf{X}\mathbf{W}_{2:A}$, we may also remove the $\mathbf{y}$-orthogonal parts of $\mathbf{X}$ and $\mathbf{x}_{\mathrm{new}}$ and compute

$$\mathbf{X}_{\mathrm{PLS+ST}} = \mathbf{X} - \mathbf{X}\mathbf{W}_{2:A} \left(\mathbf{P}_{2:A}^{\mathrm{PLS+ST}}\right)^T \tag{9}$$

and

$$\left(\mathbf{x}_{\mathrm{new}}^{\mathrm{PLS+ST}}\right)^T = \mathbf{x}_{\mathrm{new}}^T - \mathbf{x}_{\mathrm{new}}^T \mathbf{W}_{2:A} \left(\mathbf{P}_{2:A}^{\mathrm{PLS+ST}}\right)^T, \tag{10}$$

and then use the formula (3) to find

$$\hat{y}_{\mathrm{new}} = \left(\mathbf{x}_{\mathrm{new}}^{\mathrm{PLS+ST}}\right)^T \mathbf{w}_1 \left(\mathbf{w}_1^T \mathbf{X}_{\mathrm{PLS+ST}}^T \mathbf{X}_{\mathrm{PLS+ST}} \mathbf{w}_1\right)^{-1} \mathbf{w}_1^T \mathbf{X}_{\mathrm{PLS+ST}}^T \mathbf{y}. \tag{11}$$

3 Comparison with OPLS

Since the proofs of properties and results discussed below are rather technical, they are given in a Supplementary Appendix (http://www.....). In the discussion we refer to different similarity transformations of the type used in Eqs. (5) and (6).

Orthogonalized PLS regression models

The OPLS algorithm is based on the orthogonalized PLS factorization according to Wold [3]

$$\mathbf{X} = \mathbf{t}_1^{\mathrm{W}} \mathbf{p}_1^T + \mathbf{t}_2^{\mathrm{W}} \mathbf{p}_2^T + ... + \mathbf{t}_A^{\mathrm{W}} \mathbf{p}_A^T + \mathbf{E}_{\mathrm{W}} = \mathbf{T}_{\mathrm{W}} \mathbf{P}^T + \mathbf{E}_{\mathrm{W}}, \tag{12}$$

where $\mathbf{T}_{\mathrm{W}} \in \mathbb{R}^{n \times A}$ is orthogonal, while $\mathbf{P} \in \mathbb{R}^{p \times A}$ is not. The residual $\mathbf{E}_{\mathrm{W}}$ is here slightly different from the residual $\mathbf{E}$ in the Martens factorization (2). This may be corrected by using

$$\mathbf{X} = \mathbf{t}_1^{\mathrm{W}} \mathbf{p}_1^T + \mathbf{t}_2^{\mathrm{W}} \mathbf{p}_2^T + ... + \mathbf{t}_{A-1}^{\mathrm{W}} \mathbf{p}_{A-1}^T + \mathbf{t}_A^{\mathrm{W}} \mathbf{w}_A^T + \mathbf{E} = \mathbf{T}_{\mathrm{W}} \mathbf{P}^T \mathbf{W} \mathbf{W}^T + \mathbf{E}, \tag{13}$$

where $\mathbf{T}_{\mathrm{W}} \mathbf{P}^T \mathbf{W} = \mathbf{T}$. This factorization may be considered more correct, since use of $A = p$ components (the number of variables) results in $\mathbf{E} = \mathbf{0}$, while Eq. (12) then gives $\mathbf{E}_{\mathrm{W}} \neq \mathbf{0}$. From Eq. (13) also the well known estimator expression $\hat{\mathbf{b}} = \mathbf{W} \left(\mathbf{P}^T \mathbf{W} \right)^{-1} \mathbf{q}_W$ [7] follows as a simple least squares solution. Note that also Eq. (13) follows directly from the step-wise NIPALS algorithm [3], it is just a matter of where to end the last step. Also note that the only difference between $\mathbf{T}_{\mathrm{W}} \mathbf{P}^T$ and $\mathbf{T}_{\mathrm{W}} \mathbf{P}^T \mathbf{W} \mathbf{W}^T$ is that $\mathbf{p}_A^T$ is replaced by $\mathbf{w}_A^T$.

Permuted deflations

As discussed below, the OPLS algorithm makes use of a permuted loading weights matrix $\tilde{\mathbf{W}}$. With a predetermined loading weights matrix $\mathbf{W}$ it is possible to alter the deflation order in the algorithm resulting in Eq. (2), and it is rather obvious that this will not affect the final residual and predictions. A matrix $\tilde{\mathbf{W}}$ with permuted column vectors will thus give $\mathbf{X} = \tilde{\mathbf{T}} \tilde{\mathbf{W}}^T + \mathbf{E}$, i.e. $\tilde{\mathbf{T}} \tilde{\mathbf{W}}^T = \mathbf{T} \mathbf{W}^T$. The corresponding is true also for Eqs. (12) and (13), i.e. a permuted loading weights matrix $\tilde{\mathbf{W}}$ results in $\mathbf{X} = \tilde{\mathbf{T}}_{\mathrm{W}} \tilde{\mathbf{P}}^T + \mathbf{E}_{\mathrm{W}}$ and $\mathbf{X} = \tilde{\mathbf{T}}_{\mathrm{W}} \tilde{\mathbf{P}}^T \tilde{\mathbf{W}} \tilde{\mathbf{W}}^T + \mathbf{E}$ respectively, where $\tilde{\mathbf{T}}_{\mathrm{W}} \tilde{\mathbf{P}}^T = \mathbf{T}_{\mathrm{W}} \mathbf{P}^T$ and $\tilde{\mathbf{T}}_{\mathrm{W}} \tilde{\mathbf{P}}^T \tilde{\mathbf{W}} \tilde{\mathbf{W}}^T = \mathbf{T}_{\mathrm{W}} \mathbf{P}^T \mathbf{W} \mathbf{W}^T$. The differences between $\tilde{\mathbf{T}}$ and $\mathbf{T}$ etc. are thus similarity transformations only.

Original OPLS method

The OPLS method of Trygg and Wold [2] uses the factorization (assuming $A-1$ $\mathbf{y}$-orthogonal components)

$$\mathbf{X} = \mathbf{T}_{\mathrm{ortho}} \mathbf{P}_{\mathrm{ortho}}^T + \mathbf{t}_A^{\mathrm{OPLS}} \left(\mathbf{p}_A^{\mathrm{OPLS}} \right)^T + \mathbf{E}_{\mathrm{OPLS}}, \tag{14}$$

where the score vectors in $\mathbf{T}_{\mathrm{ortho}}$ are orthogonal and also orthogonal to $\mathbf{y}$, $\hat{\mathbf{y}}$ and $\mathbf{t}_A^{\mathrm{OPLS}}$. The pre-processing algorithm for finding $\mathbf{T}_{\mathrm{ortho}}$ and $\mathbf{P}_{\mathrm{ortho}}$ also finds $\mathbf{w}_1$ and an OPLS loading weights matrix $\mathbf{W}_{\mathrm{ortho}}$, and the key for understanding is that $\mathbf{W}_{\mathrm{ortho}} = -\mathbf{W}_{2:A}$. The OPLS algorithm thus gives the same results as ordinary orthogonalized PLS with permuted loading weights vectors in the order $\mathbf{w}_2$, $\mathbf{w}_3$, ... $\mathbf{w}_A$, $\mathbf{w}_1$. The difference from Eq. (12) is thus also here a similarity transformation only, i.e. $\mathbf{E}_{\mathrm{OPLS}} = \mathbf{E}_{\mathrm{W}}$.

Modified OPLS method

The PLS+ST method is based on the non-orthogonalized factorization (2), and before a comparison the OPLS method must be modified accordingly into

$$\mathbf{X} = \mathbf{T}_{\mathrm{ortho}} \mathbf{P}_{\mathrm{ortho}}^T \mathbf{W} \mathbf{W}^T + \mathbf{t}_A^{\mathrm{OPLS}} \mathbf{w}_1^T + \mathbf{E}, \tag{15}$$

where $\mathbf{T}_{\mathrm{ortho}} \mathbf{P}_{\mathrm{ortho}}^T$ still may be found using the original OPLS algorithm. This is the same result as from ordinary orthogonalized PLS using permuted loading weights vectors in the order $\mathbf{w}_2$, $\mathbf{w}_3$, ... $\mathbf{w}_A$, $\mathbf{w}_1$, followed by multiplication with $\mathbf{W} \mathbf{W}^T$ as in Eq. (13). Note that $\mathbf{E}_{\mathrm{OPLS}}$ is replaced by $\mathbf{E}$, and as argued above this modified OPLS may thus be considered more correct than original OPLS. Also note that the OPLS algorithm gives also $\mathbf{w}_1$ and $\mathbf{W} = \left[\begin{array}{cc} \mathbf{w}_1 & -\mathbf{W}_{\mathrm{ortho}} \end{array} \right]$, and the modification (15) is thus possible without use of the ordinary PLS algorithm.

Comparison with PLS+ST

The factorizations (2), (13) and (15) are all identical within similarity transformations, and so is the PLS+ST factorization (6). A detailed comparison leads to the following results:

Result 1 The second similarity transformation $\mathbf{T}_{\text{ortho}}\mathbf{P}_{\text{ortho}}^T = \mathbf{T}_{\text{ortho}}\mathbf{P}_{\text{ortho}}^T\mathbf{W}_{2:A}\left(\mathbf{P}_{\text{ortho}}^T\mathbf{W}_{2:A}\right)^{-1}\mathbf{P}_{\text{ortho}}^T$ results in the transformed OPLS score matrix $\mathbf{T}_{\text{ortho}}\mathbf{P}_{\text{ortho}}^T\mathbf{W}_{2:A} = \mathbf{T}_{2:A}$.

Result 2 The final modified OPLS component is identical with the first PLS+ST component, i.e. $\mathbf{t}_A^{\text{OPLS}}\mathbf{w}_1^T = \mathbf{t}_1^{\text{PLS+ST}}\mathbf{w}_1^T$.

Result 3 The first modified and then transformed OPLS loading matrix is identical with the PLS+ST loading matrix, i.e. $\mathbf{W}\mathbf{W}^T\mathbf{P}_{\text{ortho}}\left(\mathbf{W}_{2:A}^T\mathbf{P}_{\text{ortho}}\right)^{-1} = \mathbf{P}_{2:A}^{\text{PLS+ST}}$.

Conclusion The difference between PLS+ST and modified OPLS is thus a similarity transformation only, i.e. $\mathbf{T}_{\text{ortho}}\mathbf{P}_{\text{ortho}}^T\mathbf{W}\mathbf{W}^T = \mathbf{T}_{2:A}\left(\mathbf{P}_{2:A}^{\text{PLS+ST}}\right)^T$. Note that the score identities in Results 1 and 2 are valid also for the original OPLS. Also note that the OPLS modification involves the loading vectors only, changing $\begin{bmatrix} \mathbf{p}_1^{\text{ortho}} & \cdots & \mathbf{p}_{A-2}^{\text{ortho}} & \mathbf{p}_{A-1}^{\text{ortho}} & \mathbf{p}_A^{\text{OPLS}} \end{bmatrix}$ into $\begin{bmatrix} \mathbf{p}_1^{\text{ortho}} & \cdots & \mathbf{p}_{A-2}^{\text{ortho}} & \left(\mathbf{p}_{A-1}^{\text{ortho}}\right)^T\mathbf{w}_1\mathbf{w}_1 - \mathbf{w}_A & \mathbf{w}_1 \end{bmatrix}$. Since the corresponding difference between Eqs. (12) and (13) is only that $\mathbf{p}_A^T$ is replaced by $\mathbf{w}_A^T$, this is a minor change that may be of little practical interest.

Alternative OPLS algorithm

From the development above follows an alternative two-step OPLS algorithm:

1. Determine the loading weights matrix $\mathbf{W}$ using e.g. the orthogonalized PLS algorithm.
2. Use the orthogonalized PLS algorithm once more, but now with permuted loading weights vectors in $\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{W}_{2:A} & \mathbf{w}_1 \end{bmatrix} = \begin{bmatrix} -\mathbf{W}_{\text{ortho}} & \mathbf{w}_1 \end{bmatrix}$.

The second step will directly give $\mathbf{T}_{\text{ortho}}$, $\mathbf{P}_{\text{ortho}}$, $\mathbf{t}_A^{\text{OPLS}}$ and $\mathbf{p}_A^{\text{OPLS}}$, and thus the OPLS results, either the original factorization (14) or the modified factorization (15). Another natural choice would be to remove the least influential components first, i.e. to use the permutation $\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{w}_A & \mathbf{w}_{A-1} & \cdots & \mathbf{w}_2 & \mathbf{w}_1 \end{bmatrix}$ in the second step.

4 Comparison with PLS+PCP and PLS+CCA

In the single response case, the PLS+PCP and PLS+CCA methods give identical scores [6], and a comparison with PLS+PCP is thus to some extent relevant also for PLS+CCA. The PLS+PCP method [5] uses a factorization (with normalized loadings)

$$\mathbf{X} = \mathbf{t}_1^{\text{PCP}}\left(\mathbf{w}_1^{\text{PCP}}\right)^T + \mathbf{E}_{\text{PCP}}. \tag{16}$$

Result 4 For a single $\mathbf{y}$-relevant component the relations between the PLS+PCP and PLS+ST methods are that $\mathbf{t}_1^{\text{PCP}} \to \mathbf{t}_1^{\text{PLS+ST}} = \mathbf{t}_A^{\text{OPLS}}$ and $\mathbf{w}_1^{\text{PCP}} \to \mathbf{w}_1$ when $\hat{\mathbf{y}} \to \mathbf{y}$, i.e. with good predictions.

5 Conclusions

A simple post-processing method for separation of **y**-relevant and **y**-orthogonal variation in the **X** matrix is developed, using a non-orthogonalized PLS regression model and a similarity transformation. The method is at present restricted for use in single response cases. Within similarity transformations, the PLS+ST factorization of **X** is shown to be identical both with a modified version of the orthogonalized PLS factorization and with a correspondingly modified version of the OPLS factorization. The single extracted **y**-relevant PLS+ST component based on a PLS model with A components is thus identical with the single remaining modified OPLS component after removal of $A-1$ **y**-orthogonal components. The minor PLS and OPLS modifications involved are based on a theoretically founded interpretation of the results from the ordinary NIPALS algorithm. The original OPLS factorization, on the other hand, is within a similarity transformation shown to be identical with the ordinary orthogonalized PLS factorization. All these factorizations have common score and loading weights spaces, and they thus result in identical predictions. Both the original and modified OPLS factorizations can also be obtained by using the ordinary NIPALS algorithm twice, the second time with the predetermined but permuted loading weights vectors.

As pointed out in [2], the obvious advantages with OPLS are more parsimonious PLS representations and easier interpretation. Analysis of the **y**-orthogonal part of **X** may also be valuable. As shown in the present paper, however, an identical **y**-relevant score vector and just as informative **y**-orthogonal components can be obtained in a more direct and thus more transparent way by use of PLS+ST. Another simple alternative is to use the ordinary PLS algorithm twice, the second time with predetermined and permuted loading weights vectors.

Supplementary information with proofs is available at
(http://www.interscience.wiley.com/jpages/0886-9883/supmat/)

References

[1] Svensson O, Kourti T, MacGregor JF. An investigation of orthogonal signal correction algorithms and there characteristics. *J. Chemometrics* 2002; **16**: 176-188.

[2] Trygg J, Wold S. Orthogonal projections to latent structures, O-PLS. *J. Chemometrics* 2002; **16**: 119-128.

[3] Martens H, Næs T. *Multivariate Calibration.* Wiley: New York, 1989.

[4] Verron T, Sabatier R, Joffre R. Some theoretical properties of the O-PLS method. *J. Chemometrics* 2004; **18**: 62-68.

[5] Langsrud Ø, Næs T. Optimised score plot by principal components of prediction. *Chemometrics Intell. Lab. Syst.* 2003; **68**: 61-74.

[6] Yu H, MacGregor JF. Post processing methods (PLS-CCA): simple alternatives to preprocessing methods (OSC-PLS). *Chemometrics Intell. Lab. Syst.* 2004; **73**: 199-205.

[7] Helland IS. On the structure of partial least squares regression. *Communications in statistics* 1988; **17**: 581-607.