

# Quality-Relevant Process Monitoring

**S. Joe Qin**

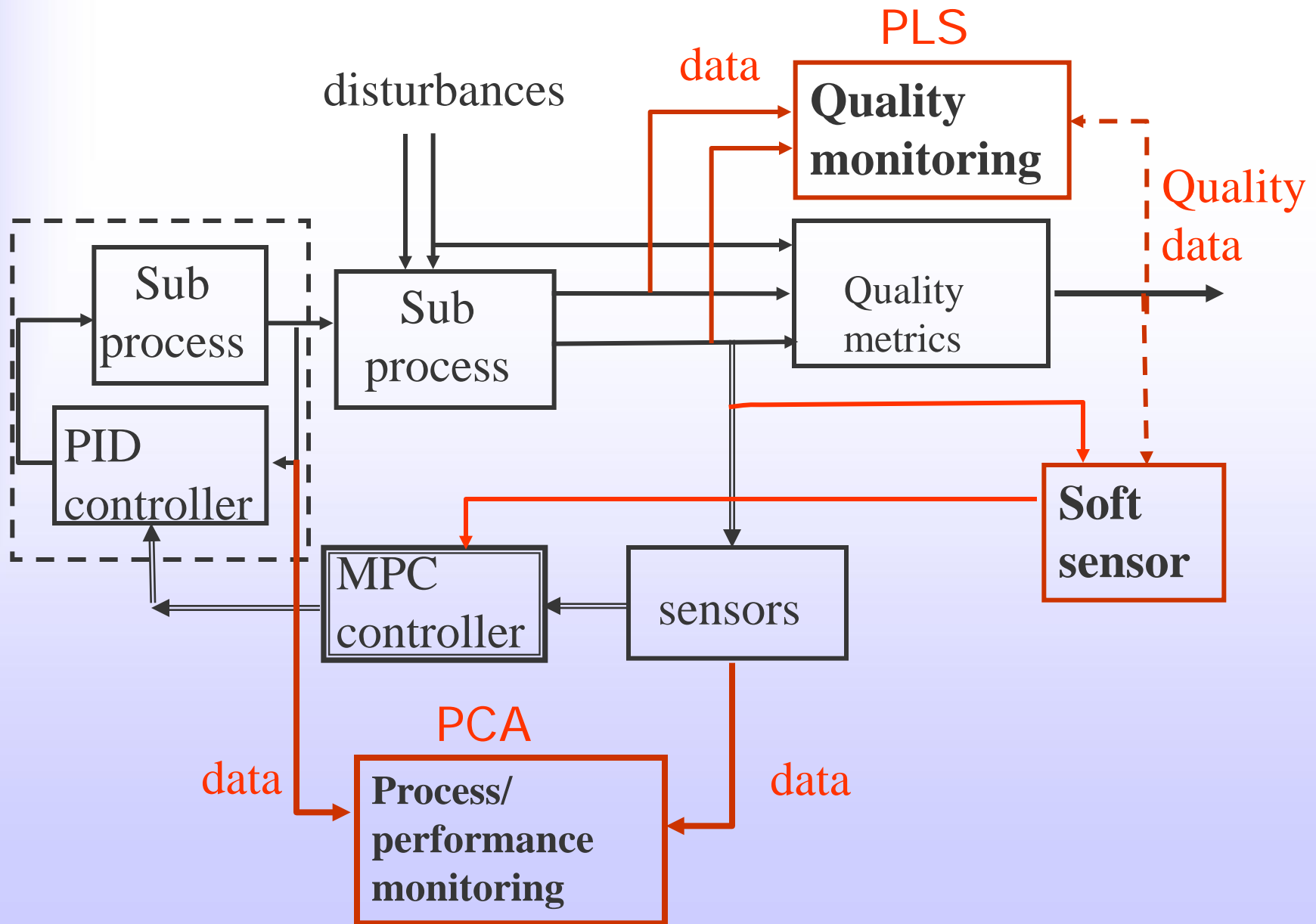
**Department of Chemical Engineering and Material  
Sciences**

**Department of Electrical Engineering  
University Southern California  
Los Angeles, CA, U.S.A.**

[sjin@usc.edu](mailto:sjin@usc.edu)

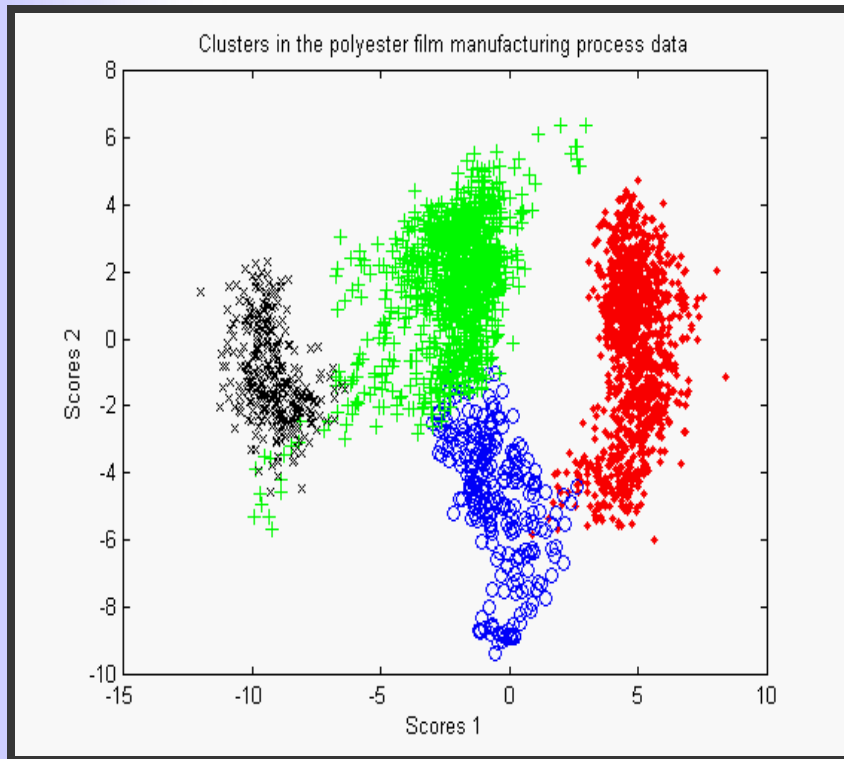
**Nov. 7-12, 2010, Salt Lake City, UT**

# Data-driven process monitoring



# Process and Quality variables

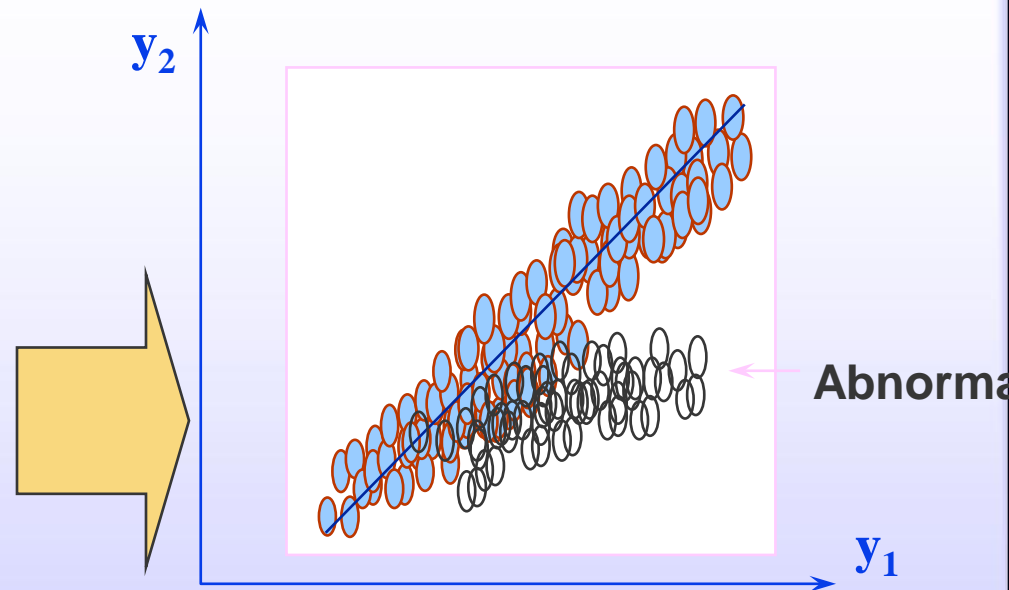
## Process data



PCA monitoring: X alone

PLS monitoring: X guided by Y

## Quality data



- **PCA based monitoring focuses on process data variation only**
  - 'unsupervised' analysis of process data
  - Nuisance alarms
- **PLS based monitoring monitors process data by its co-variation with quality data**
  - 'supervised' analysis of process data
  - removes nuisance alarms caused by PCA

# PCA diagnosis methods

- **Currently applied in many areas**
  - **Fault detection: SPE, T2, or combined**
  - **Fault diagnosis: contribution plots; fault identification via reconstruction**
- **Recent progress**
  1. **Understand the weakness of contribution plots**
  2. **Unifying many diagnosis methods and suggest relative contributions to be used**
  3. **Kernel methods for nonlinear data monitoring**

# Unifying many diagnosis methods

- Carlos Alcalá and S. Joe Qin (2010). Analysis and Generalization of Fault Diagnosis Methods for Process Monitoring. Revised for *J. of Process Control*. (Special Issue in honor of T.J. McAvoy)
  - Most contribution plot methods do not have statistically equal contributions when no fault is present
  - Suggest relative contributions which possess this property

# Nonlinear, Kernel PCA methods

- C. Alcalá and S.J. Qin (2010). Reconstruction-based Contribution for Process Monitoring with Kernel Principal Component Analysis, to appear in *I&EC Research*. (Special Issue in honor of T.F. Edgar)

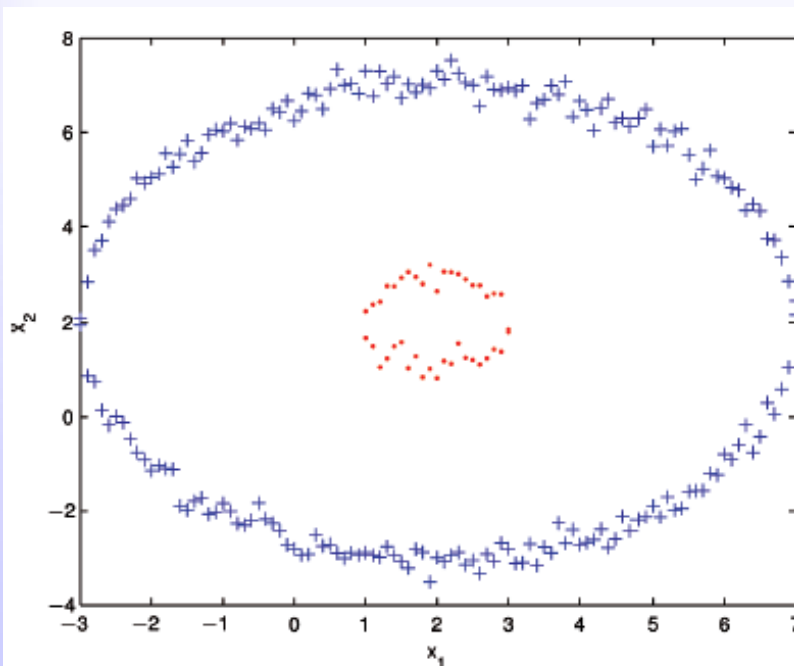


Figure 1. Normal and faulty data in a bidimensional space.

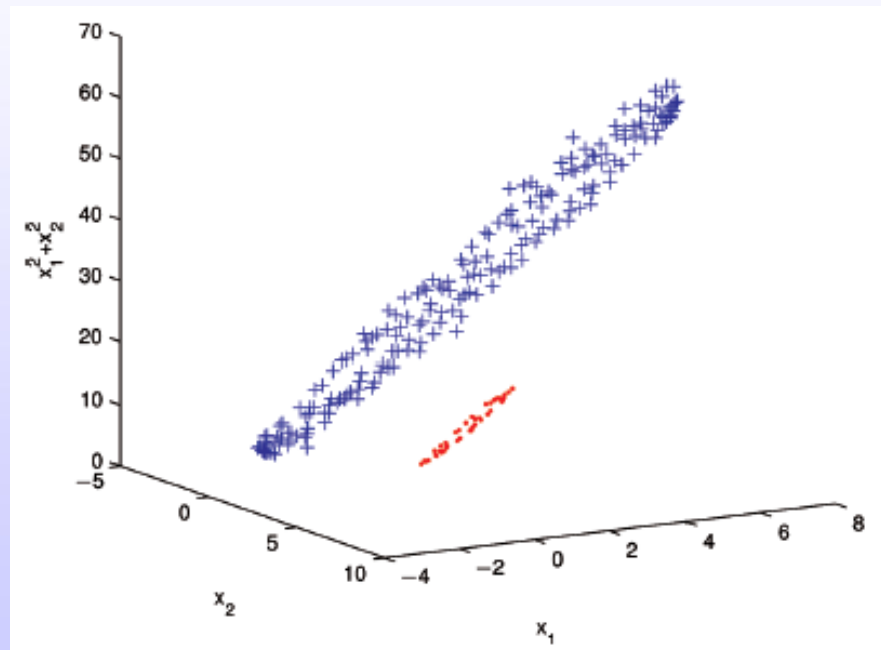
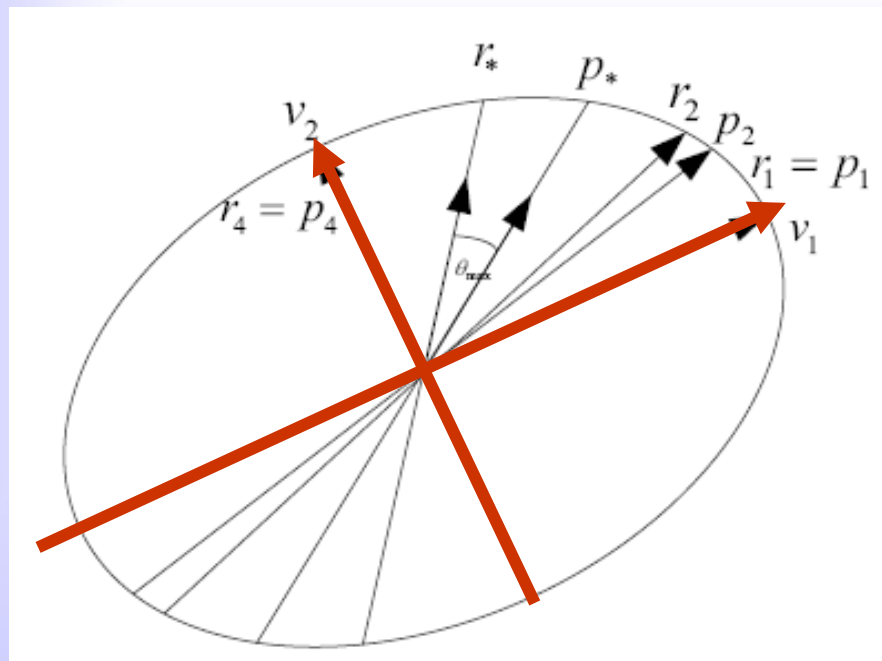


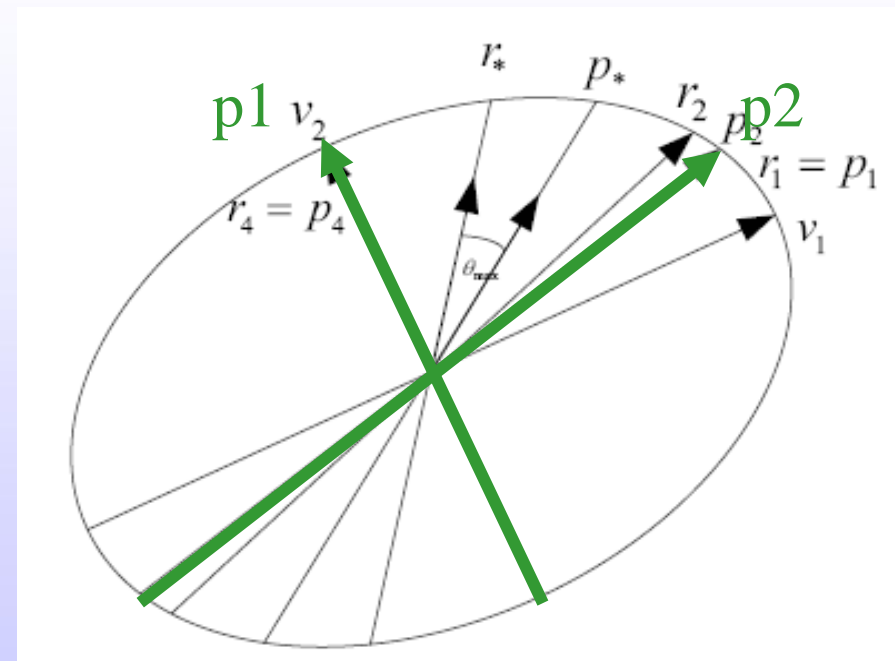
Figure 2. Normal and faulty data in a tridimensional space.

# PLS: Impact of Y on X-space Decomposition

- PLS partition of X, depending on whether Y lines up with the major X directions or not



PCA directions

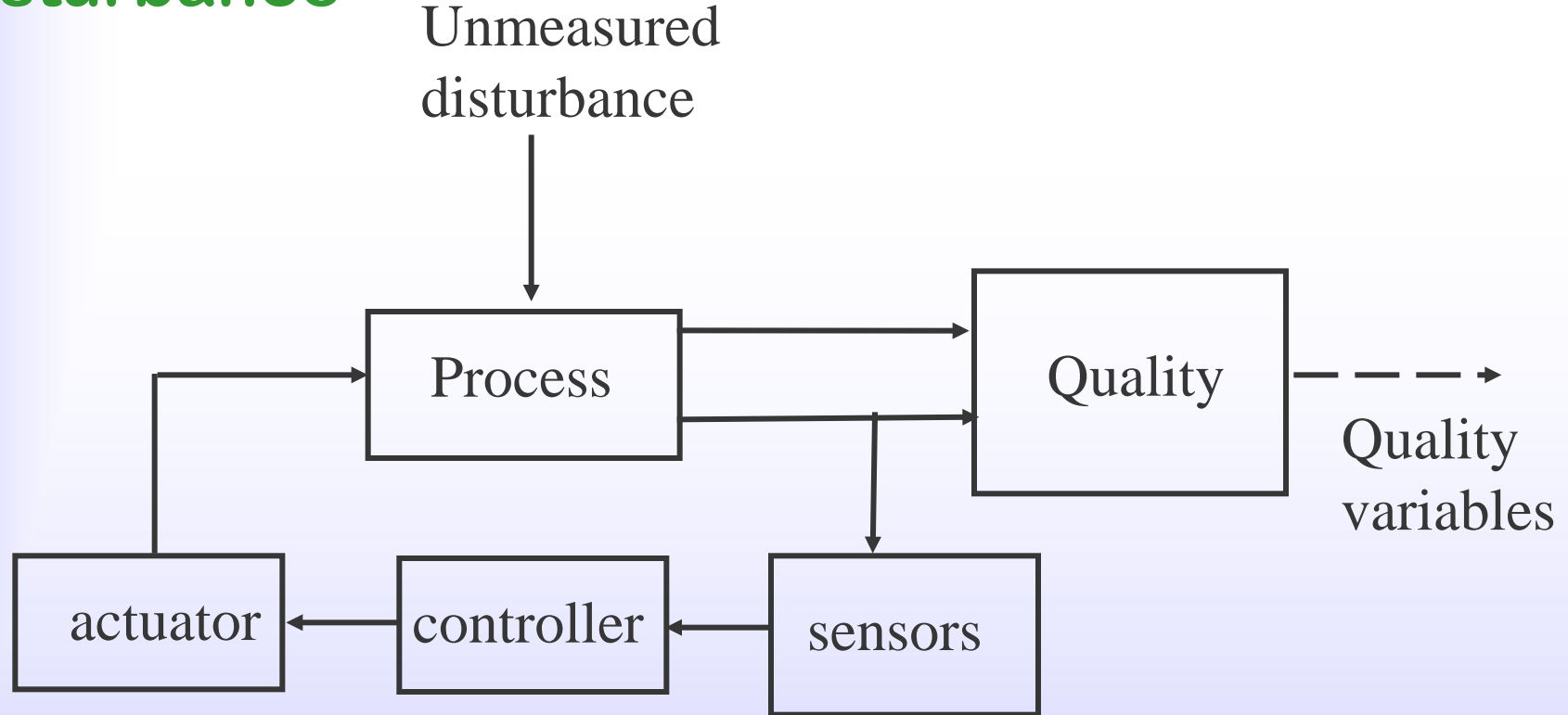


PLS directions

## PCA-like, Unsupervised Monitoring

- ❑ Pros: plenty of process data, easy to use given normal data
- ❑ Cons: need normal data to define 'normal'
- ❑ Cons: out-of-control in process data does not always point to a 'quality' problem
- ❑ Cons: Measured variables being normal does not guarantee the quality is normal because of unmeasured contributors to quality

# Example: Controller responds to unmeasured disturbance



- Tracking measured variables alone would signal an alarm even though the control does its job to reject the unmeasured disturbance

# PLS-Based Process Monitoring

- PLS-based monitoring uses quality data  $Y$  to guide the partition of process data  $X$ , which is different from PCA partition of  $X$ -space
- Impact of  $Y$  on the structural modeling of  $X$ -space
  - PLS is the de facto method for modeling  $X$  and  $Y$
  - PLS factors and residuals are interpreted in the same way as PCA factors and residuals
  - Lack of understanding of the impact of  $Y$  on the decomposition of  $X$ -space

## Recent work

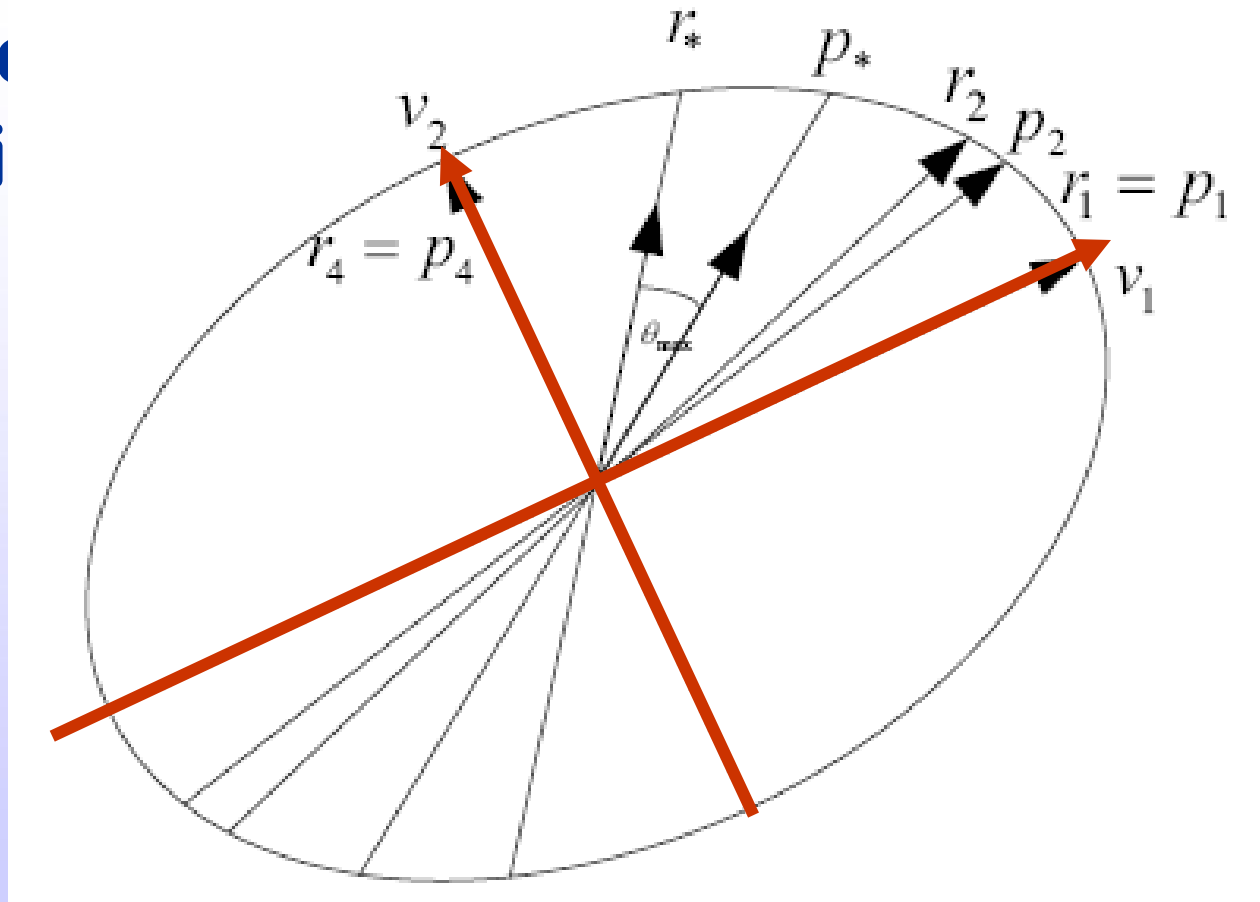
- Gang Li, S. Joe Qin, and Donghua Zhou (2008). Geometric properties of partial least squares for process monitoring, submitted to *Automatica*.
  - Gives a fairly thorough understanding of X-space decomposition guided by Y
- Donghua Zhou, Gang Li, and S. Joe Qin (2008). Total projection to latent structures for process monitoring, accepted by *AIChE Journal*.
  - PCA-like interpretation of PLS partition is not adequate. Additional projections (i.e.,

# PCA vs. PLS for monitoring

<p>PCA model</p> $\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad \mathbf{T} = \mathbf{XP}$	<p>PLS model</p> $\mathbf{X} = \mathbf{TP}^T + \mathbf{E}$ $\mathbf{Y} = \mathbf{TQ}^T + \mathbf{F} \quad \mathbf{T} = \mathbf{XR}$
<p>PCA projection</p> $\mathbf{P} = \mathbf{R}$	<p>PLS projection</p> $\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}}$ $\hat{\mathbf{x}} = \mathbf{PR}^T \mathbf{x} \in S_p \equiv \text{Span}\{\mathbf{P}\}$ $\tilde{\mathbf{x}} = (\mathbf{I} - \mathbf{PR}^T) \mathbf{x} \in S_r \equiv \text{Span}\{\mathbf{R}\}^\perp$
<p>PCA monitoring</p> $T^2 = \mathbf{t}^T \mathbf{\Lambda}^{-1} \mathbf{t} \sim \frac{A(n^2 - 1)}{n(n - A)} F_{A, n-A, \alpha}$ $Q = \ \tilde{\mathbf{x}}\ ^2 \sim g\chi_{h, \alpha}^2$	<p>PLS monitoring</p> $T^2 = \mathbf{t}^T \mathbf{\Lambda}^{-1} \mathbf{t} \sim \frac{A(n^2 - 1)}{n(n - A)} F_{A, n-A, \alpha}$ $Q = \ \tilde{\mathbf{x}}\ ^2 \sim g\chi_{h, \alpha}^2$
	<p><b>PLS residual not always ‘small’</b></p>
	<p><b>PLS scores not all related to Y</b>  <b>PLS residual ‘faults’ can affect Y</b></p>

# Impact of Y on X-space Decomposition

- PLS partition of X departs from PCA decomposition of X depending on whether the direction of Y is aligned with the direction of X



PCA  
directions

# Total PLS for Monitoring

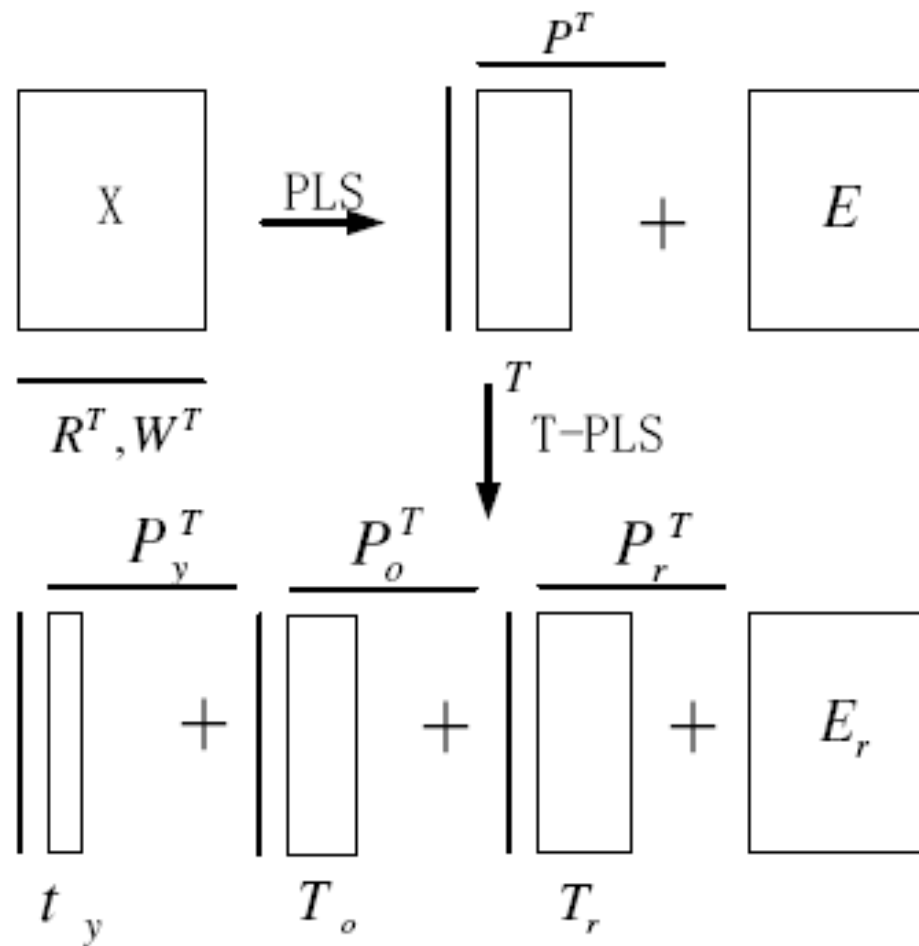


Fig. 1. T-PLS structure

# Four subspaces – Total PLS

## 1. Two related to scores

1. subspace of X-space that is solely responsible in predicting Y
2. subspace of X-space that is explored by the PLS objective but does not predict Y

## 2. Two related to residuals

1. subspace of X-space that is not 'useful' for the PLS objective, but has significant variation or excitation in X-space
2. subspace of X-space that is not excited in the X-space of the data

## 3. Detail will be presented by Carlos Alcalá

## PLS-based monitoring papers

- Gang Li, S. Joe Qin, and Donghua Zhou (2010). Geometric properties of partial least squares for process monitoring, *Automatica*, 46, 204-210.
- Donghua Zhou, Gang Li, and S. Joe Qin (2010). Total projection to latent structures for process monitoring, *AIChE Journal*, 56, 168-178.
- Gang Li, S. Joe Qin, and Donghua Zhou (2010). Output relevant fault reconstruction and fault subspace extraction in Total PLS models. Accepted by *I&EC Research*.