2nd NOSE II Workshop

Linköping, 18 - 21 May 2003

# Signal processing methods for drift compensation


Ricardo Gutierrez-Osuna

PRISM

Department of Computer Science

Texas A&M University

# Outline

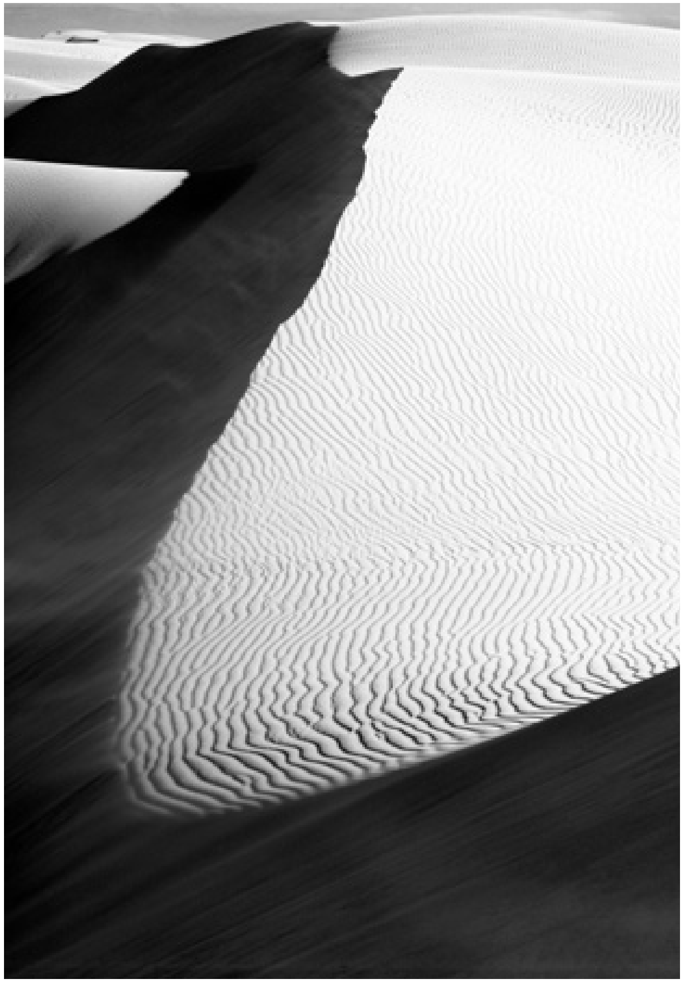
- Sources of "drift"
- Compensation approaches
- Univariate methods
- Multivariate methods
- Discussion

*This material is based upon work supported by the* ***National Science Foundation*** *under CAREER Grant No. 9984426 / 0229598*

# What is drift?

"A gradual change in any quantitative characteristic that is supposed to remain constant"

*[Holmberg and Artursson, 2003], and references therein*



# Sources of "drift"

- True drift
  - Aging (reorganization of sensing layer)
  - Poisoning (irreversible binding)
- Experimental noise
  - Short-term drift (warm-up, thermal trends)
  - Memory effects (hysteresis, sampling sequence)
  - Environmental (pressure, temp., seasonal)
  - Odor delivery (flow rate, outgass, condensation)
  - Matrix effects (background, humidity)
  - Sample degradation (oxidation, decarbonation)

# Approaches for drift compensation

- Drift-free sensors
  - Duh!
- Reference sensors
  - Differential or ratiometric measurements [Choi et al., 1985]
- Excitation
  - Temperature modulation [Roth et al., 1996]
- Frequent re-calibration
  - Unavoidable
- Careful experimental design
  - Avoid systematic errors
- Feature extraction
  - Transient response analysis [Wilson and DeWeerth, 1995]
- **Signal processing**
  - **The focus of this tutorial**

# Signal processing for drift compensation

- Univariate
  - Compensation applied to each sensor independently
    - Frequency analysis
    - Baseline manipulation
    - Differential measurements (w/ calibrant)
    - Multiplicative correction (w/ calibrant)
- Multivariate
  - Compensation applied to the response across sensors
    - Adaptive clustering
    - System identification
    - Calibration transfer (w/ calibrant)
    - Orthogonal signal correction and deflation (w/ calibrant)

# Univariate Techniques

# Frequency analysis

- Drift, noise and odor information occur at different time scales [Artursson et al., 2000]
    - Noise = high frequencies
    - Drift = low frequencies
  - Perform separation in the frequency domain
    - Filter banks [Davide et al., 1996]
    - Discrete Wavelet Transform
  - Drawbacks
    - Requires long-term time series to be collected
    - Time series has "gaps‟, variable sampling rates

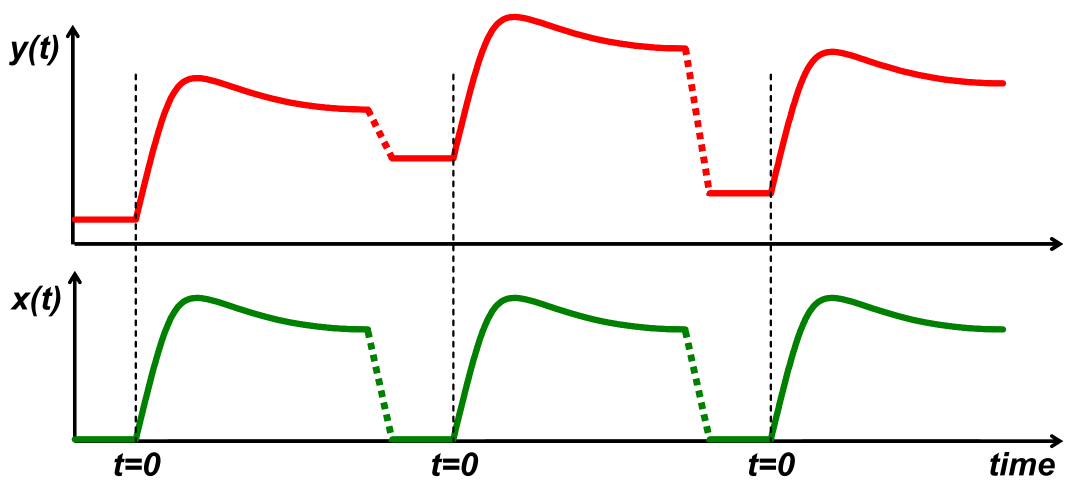
# Baseline manipulation (1)

- Basics
  - The simplest form of drift compensation
    - “Remove" the sensor response in the recovery cycle prior to sample delivery
      - Also used for pre-processing [Gardner and Bartlett, 1999]
    - A local technique, processes one "sniff" at a time

# Baseline manipulation (2)

- Differential
  - Corrects additive $\delta_A$ or baseline drift

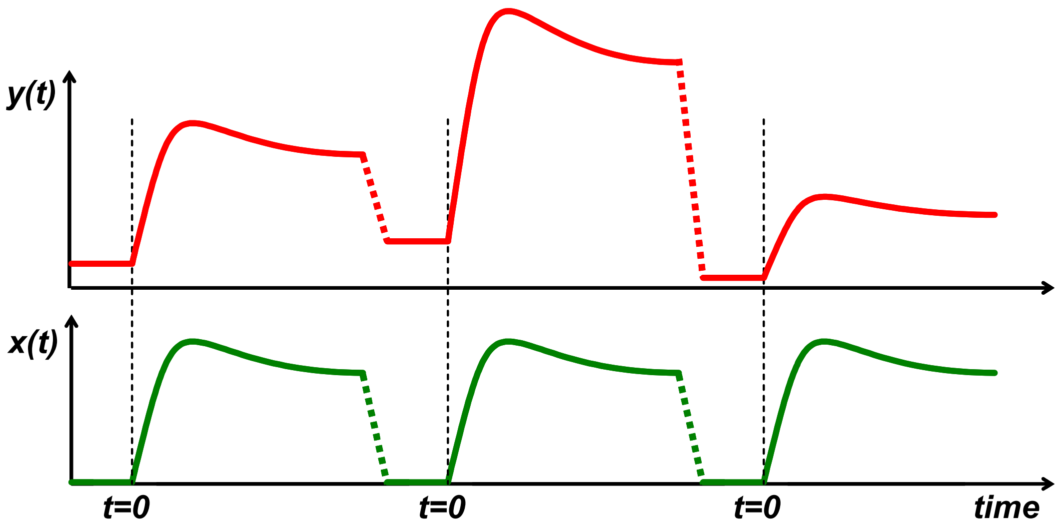
$$\hat{x}_s(t) = \underbrace{y_s(t)}_{\substack{\text{measured}\\ \text{response}}} - y_s(0) = \left(x_s(t) + \delta_A\right) - \left(x_s(0) + \delta_A\right) = \underbrace{x_s(t)}_{\substack{\text{ideal}\\ \text{response}\\ \text{(w/o drift)}}} - x_s(0)$$



# Baseline manipulation (3)

- Relative
  - Corrects multiplicative $(1+\delta_M)$ or sensitivity drift

$$\hat{x}_s(t) = \frac{y_s(t)}{y_s(0)} = \frac{x_s(t)\left(1+\delta_M\right)}{x_s(0)\left(1+\delta_M\right)} = \frac{x_s(t)}{x_s(0)}$$



# Baseline manipulation (4)

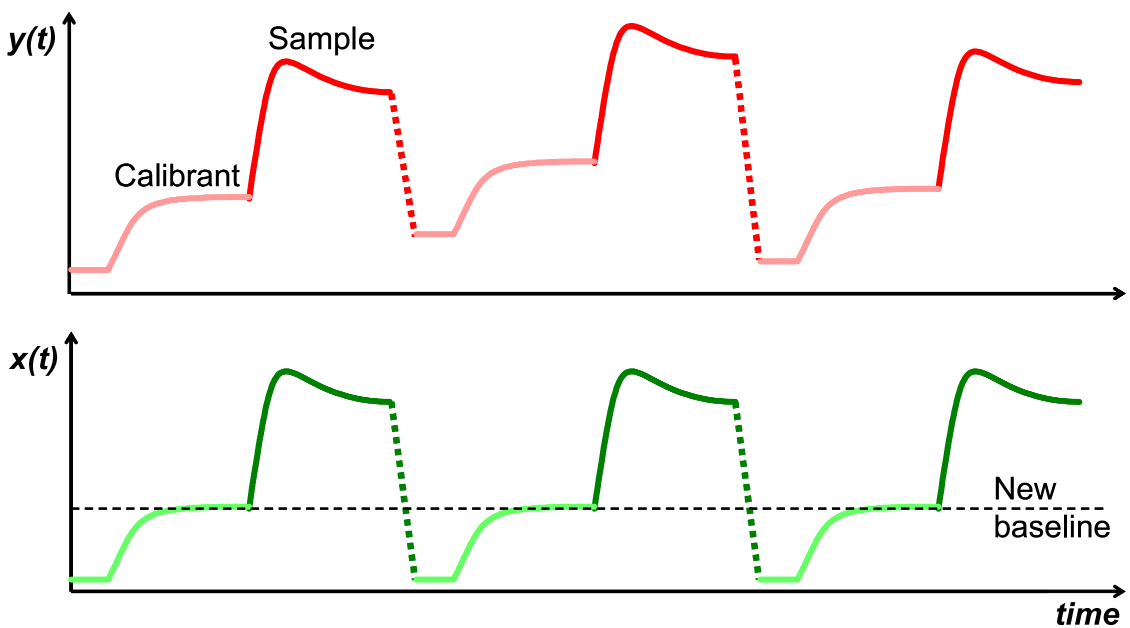
- Fractional
  - Percentual change in the sensor response

$$y_s(t) = \frac{x_s(t) - x_s(0)}{x_s(0)}$$

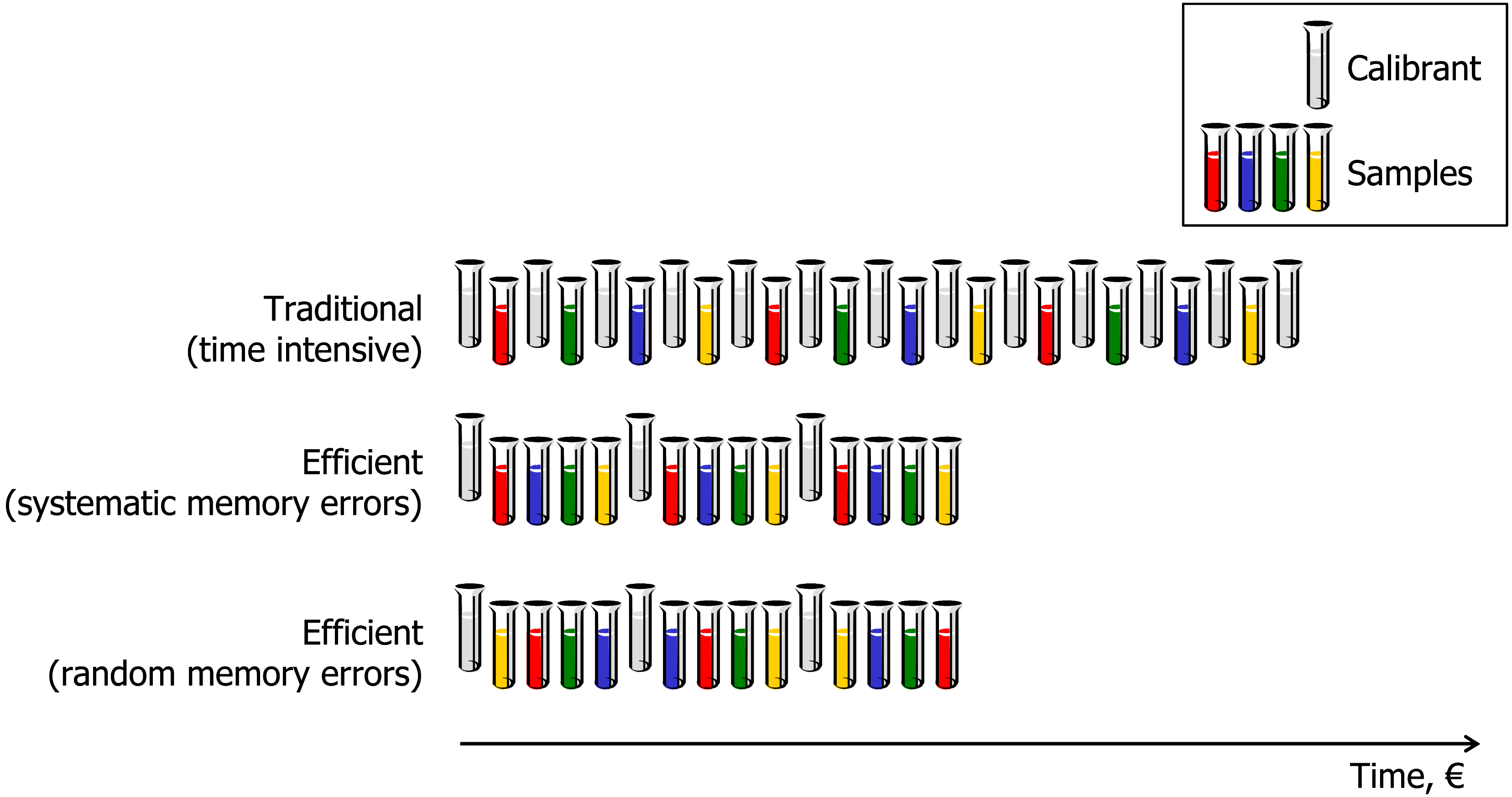
  - Properties
    - Yields a dimensionless measurement
    - Normalizes sensor responses, but can amplify noisy channels
    - Fractional conductance shown to be the most suitable method for MOS sensors [Gardner, 1991]

# Reference gas [Fryder et al., 1995]

- Equivalent to diff. BM, except a calibrant is used
  - Calibrant must be chemically stable over time AND highly correlated with samples [Haugen et al., 2000]



# Calibration schedule

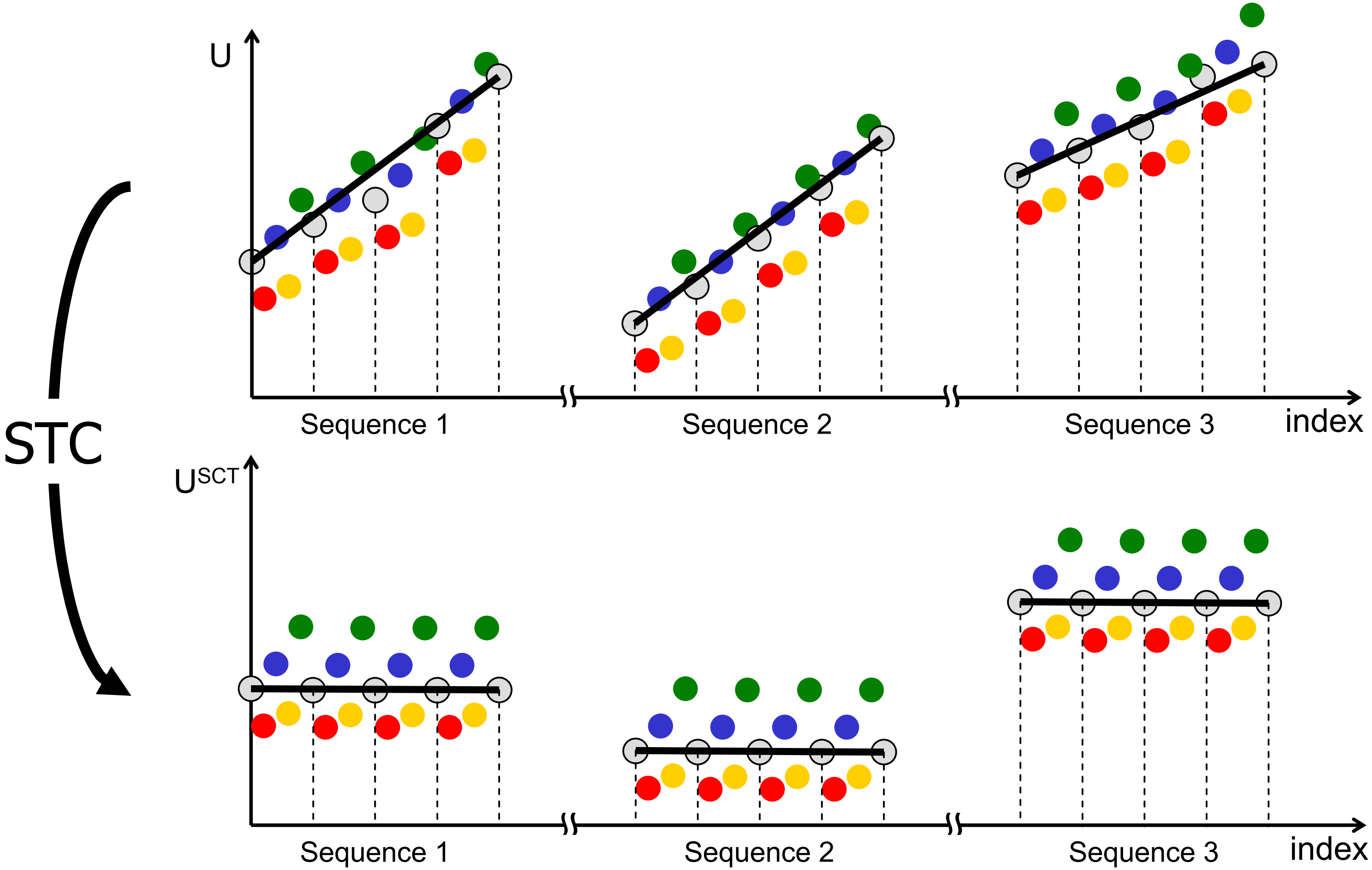



*adapted from [Salit and Turk, 1998]*

# Multiplicative correction [Haugen et al., 2000]

- Basic idea
  - Model temporal variations in a calibration gas with a multiplicative correction factor
  - Apply the same correction to the samples
  - Perform this process first on short-term trends, then on long-term fluctuations
- Properties
  - Heuristic, global technique
  - Practical for industrial applications
  - Compensates for short- and long-term drift

# Short-term correction (1)



# Short-term correction (2)

- For each sequence
  - Compute a correction factor for each calibration sample

$$q_{n,seq} = \frac{U_{1,seq}^{cal}}{U_{n,seq}^{cal}}$$

  - Build regression model for series $\{q_{1,seq}, q_{6,seq}, q_{11,seq}, \ldots\}$

$$\hat{q}_{n,seq} = a x_n + b$$
$$\text{with } x_n = \mathrm{mod}(n, N_{sam}+1)$$

  - Correct all samples

$$U_{n,seq}^{STC} = U_{n,seq}\hat{q}_{n,seq}$$

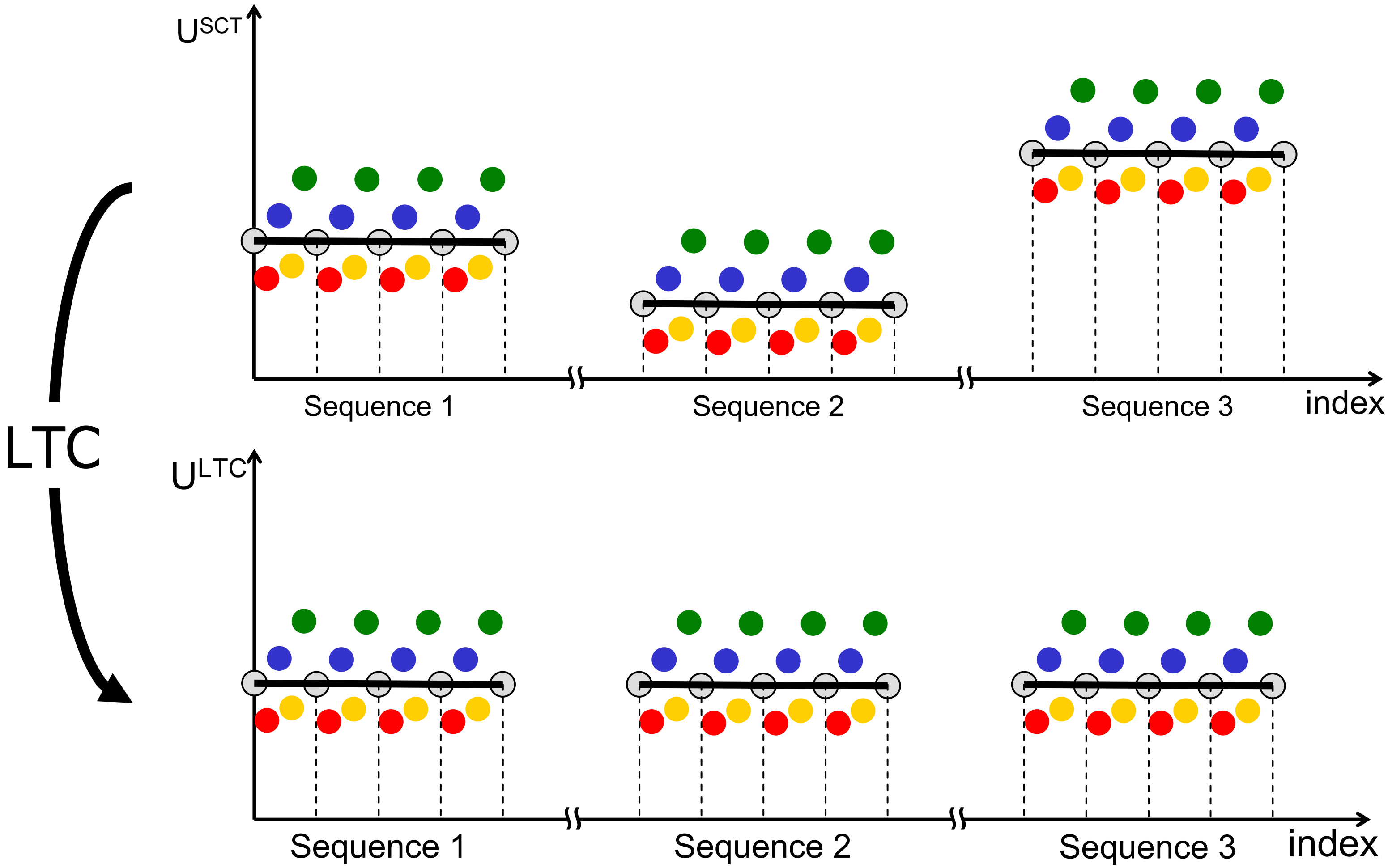
| Name | Class | Index | $x_n$ |
|---|---|---|---|
| $U_1^{cal}$ | Calibration | 1 | 0 |
| $U_2$ | Sample | 2 | 0 |
| $U_3$ | Sample | 3 | 0 |
| $U_4$ | Sample | 4 | 0 |
| $U_5$ | Sample | 5 | 0 |
| $U_6^{cal}$ | Calibration | 6 | 1 |
| $U_7$ | Sample | 7 | 1 |
| $U_8$ | Sample | 8 | 1 |
| $U_9$ | Sample | 9 | 1 |
| $U_{10}$ | Sample | 10 | 1 |
| $U_{11}^{cal}$ | Calibration | 11 | 2 |
| $U_{12}$ | Sample | 12 | 2 |
| $U_{13}$ | Sample | 13 | 2 |
| $U_{14}$ | Sample | 14 | 2 |
| $U_{15}$ | Sample | 15 | 2 |
| … | … | … | … |

$q_{6,seq}$

$\hat{q}_{9,seq}$

*(seq subindex omitted for clarity)*

# Long-term correction (1)



# Long-term correction (2)

- For each sequence
  - Compute a correction factor for the first calibration sample

$$f_{seq} = \frac{U_{1,1}^{cal}}{U_{1,seq}^{cal}}$$

  - Correct all samples

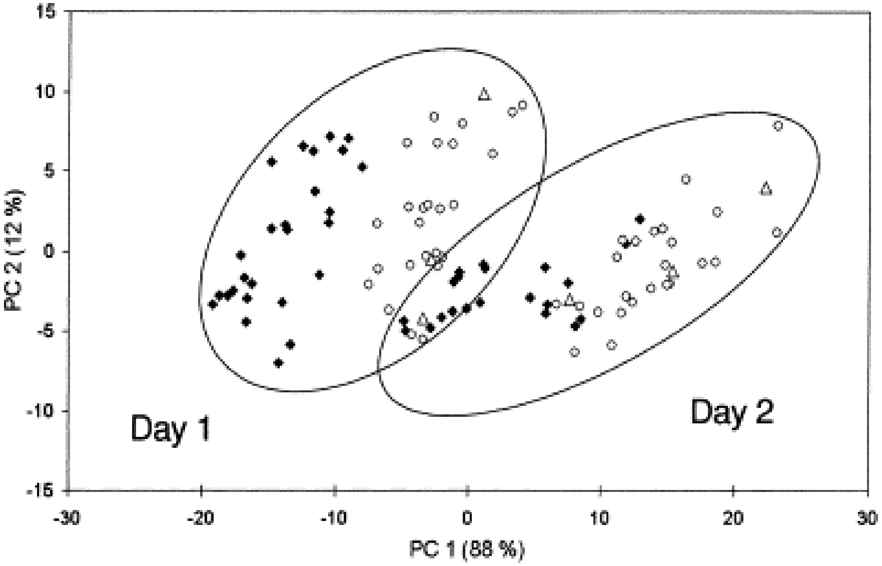
$$U_{n,seq}^{LTC} = \underbrace{U_{n,seq}\,\hat{q}_{n,seq}}_{U_{n,seq}^{STC}}\, f_{seq}$$

$f_2$

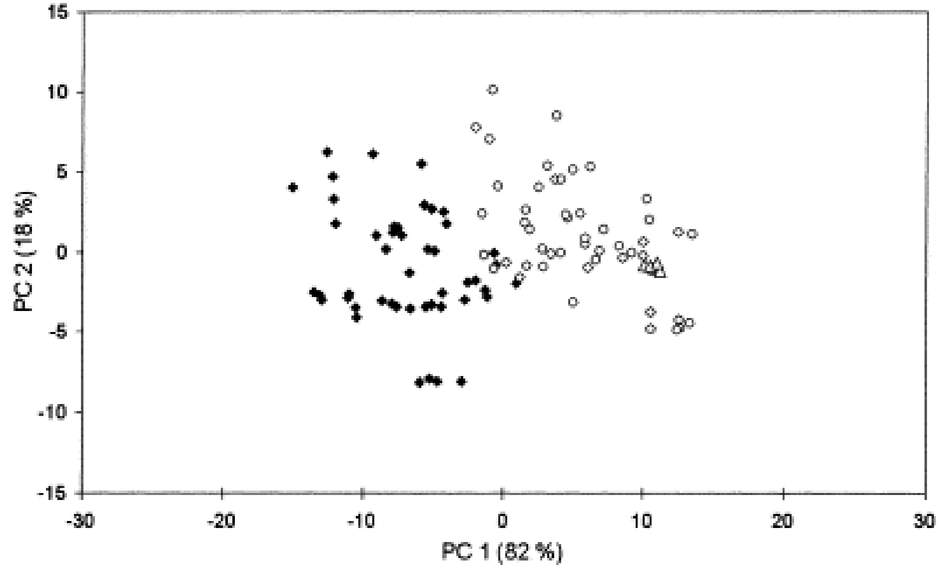
$f_3$

| Name | Class | Index | $x_n$ | Seq |
|---|---|---|---|---|
| $U_{1,1}^{cal}$ | Calibration | 1 | 0 | 1 |
| $U_{2,1}$ | Sample | 2 | 0 | 1 |
| $U_{3,1}$ | Sample | 3 | 0 | 1 |
| … | | | | 1 |
| $U_{6,1}^{cal}$ | Calibration | 6 | 1 | 1 |
| $U_{7,1}$ | Sample | 7 | 1 | 1 |
| $U_{8,1}$ | Sample | 8 | 1 | 1 |
| … | | | | |
| $U_{1,2}^{cal}$ | Calibration | 1 | 0 | 2 |
| $U_{2,2}$ | Sample | 2 | 0 | 2 |
| $U_{3,2}$ | Sample | 3 | 0 | 2 |
| … | | | | 2 |
| $U_{6,2}^{cal}$ | Calibration | 6 | 1 | 2 |
| $U_{7,2}$ | Sample | 7 | 1 | 2 |
| $U_{8,2}$ | Sample | 8 | 1 | 2 |
| … | … | … | … | |
| $U_{1,3}^{cal}$ | Calibration | 1 | 0 | 3 |
| $U_{2,3}$ | Sample | 2 | 0 | 3 |
| $U_{3,3}$ | Sample | 3 | 0 | 3 |
| … | | | | 3 |
| $U_{6,3}^{cal}$ | Calibration | 6 | 1 | 3 |
| $U_{7,3}$ | Sample | 7 | 1 | 3 |
| $U_{8,3}$ | Sample | 8 | 1 | 3 |
| … | … | … | … | |

# Performance of multiplicative correction




*PCA plot of uncorrected milk samples measured over 2 days: pasteurized milk (○) , oxidized pasteurized milk (◆), calibration samples (△).*




*PCA plot of drift corrected milk samples measured over 2 days: reference milk (○), oxidized milk (◆), calibration samples (△).*

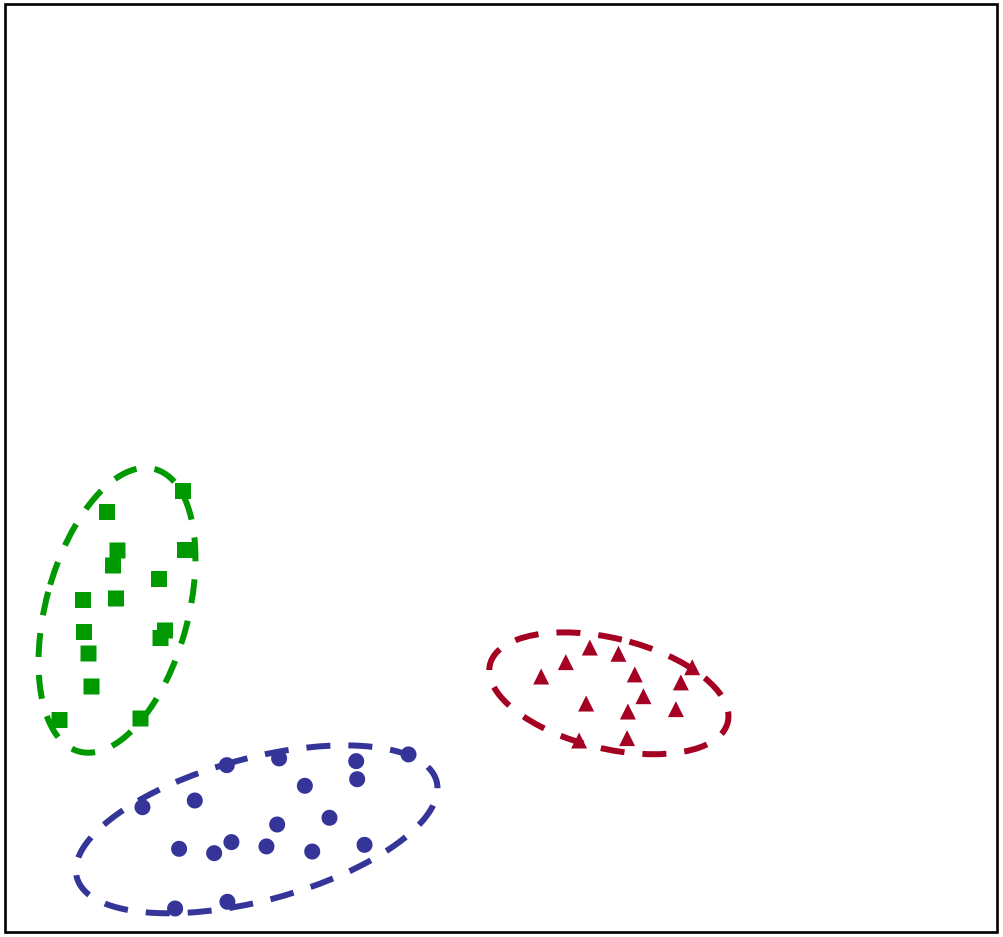
*from [Haugen et al., 2000]*

# Multivariate Techniques

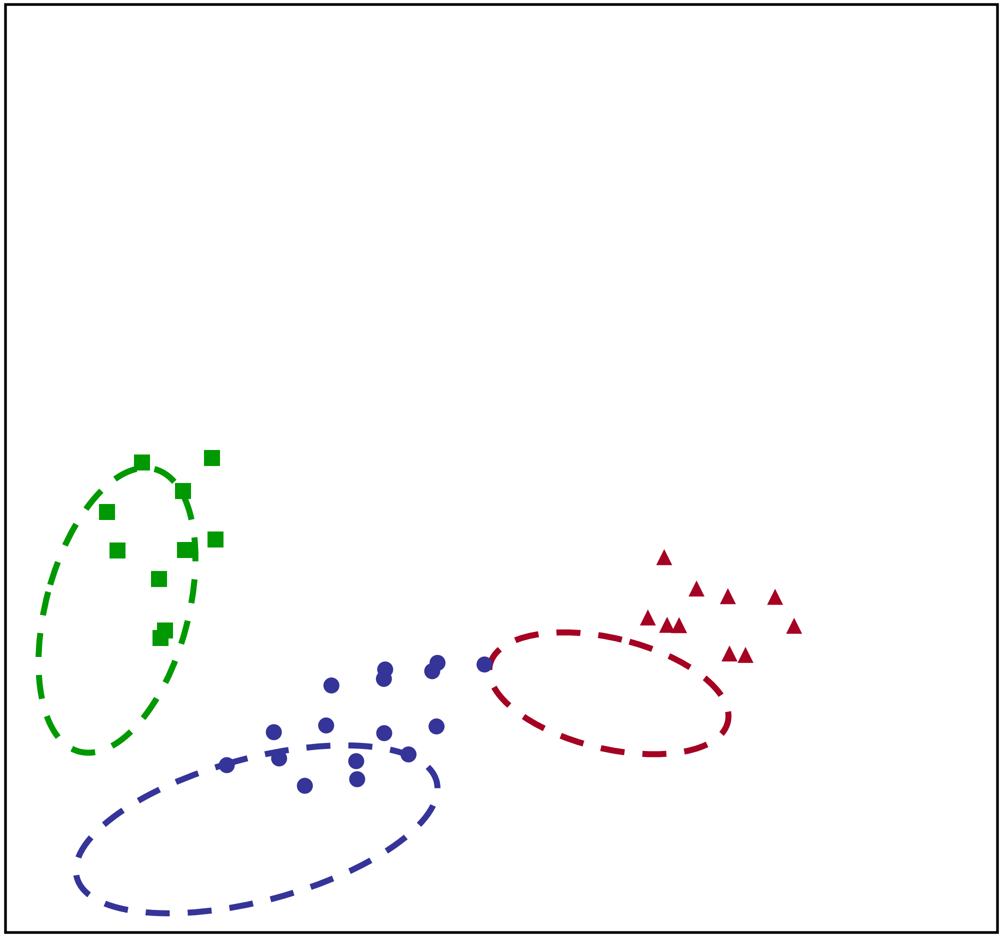
# Adaptive clustering

- Basic idea
  - Model the distribution of examples with a codebook
  - Assign an incoming (unknown) sample to the "closest" class
  - Adapt class parameters to incorporate information from the newly classified example

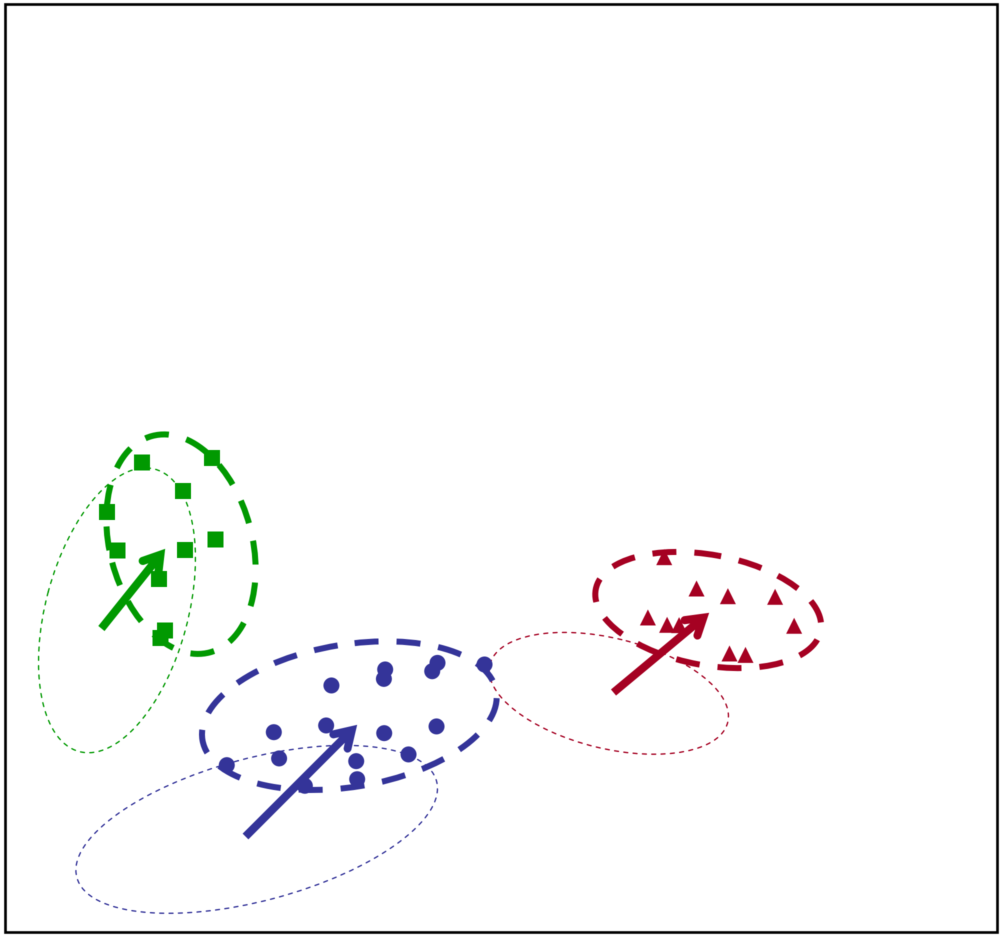
# Adaptive clustering (1)



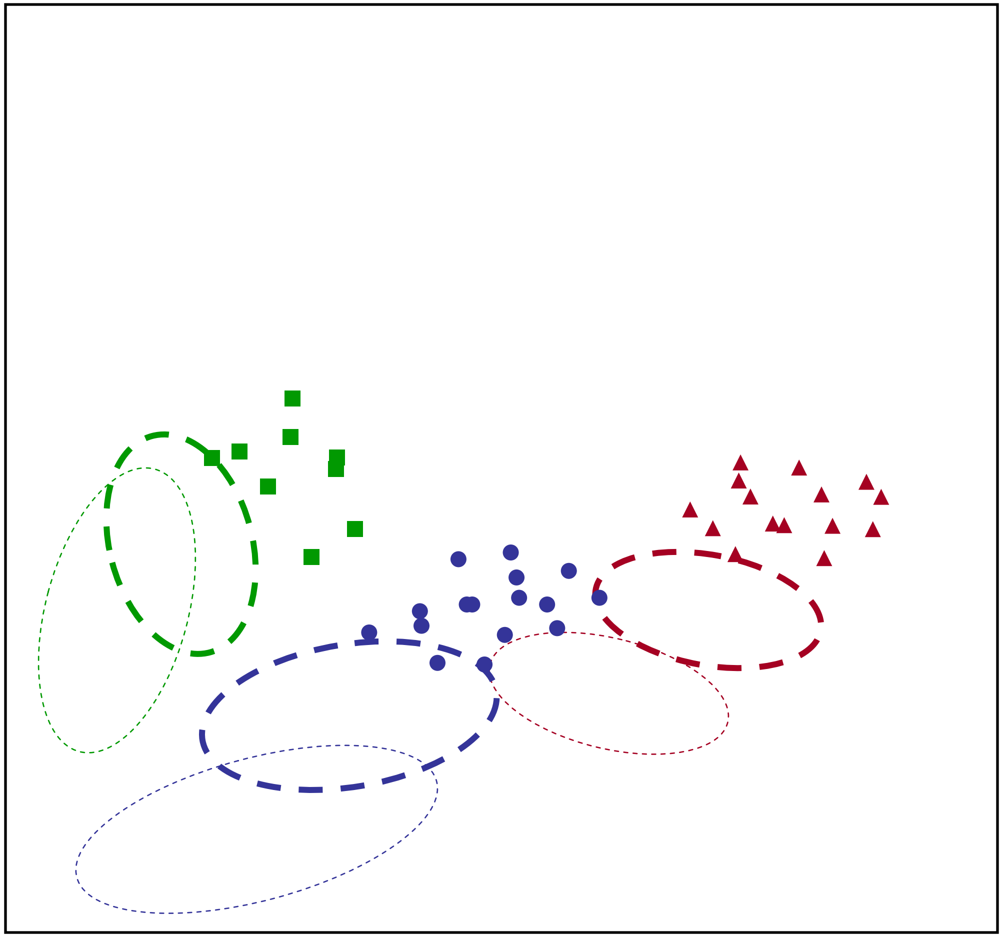
# Adaptive clustering (2)



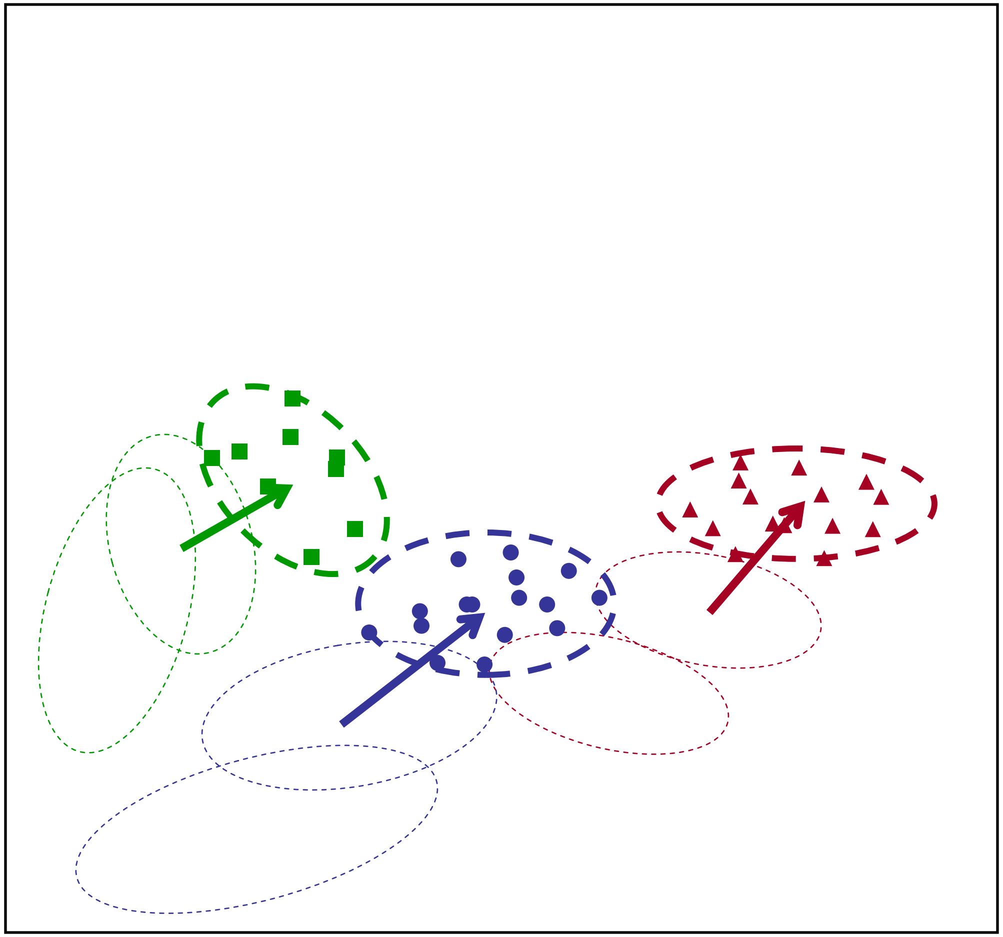
# Adaptive clustering (3)



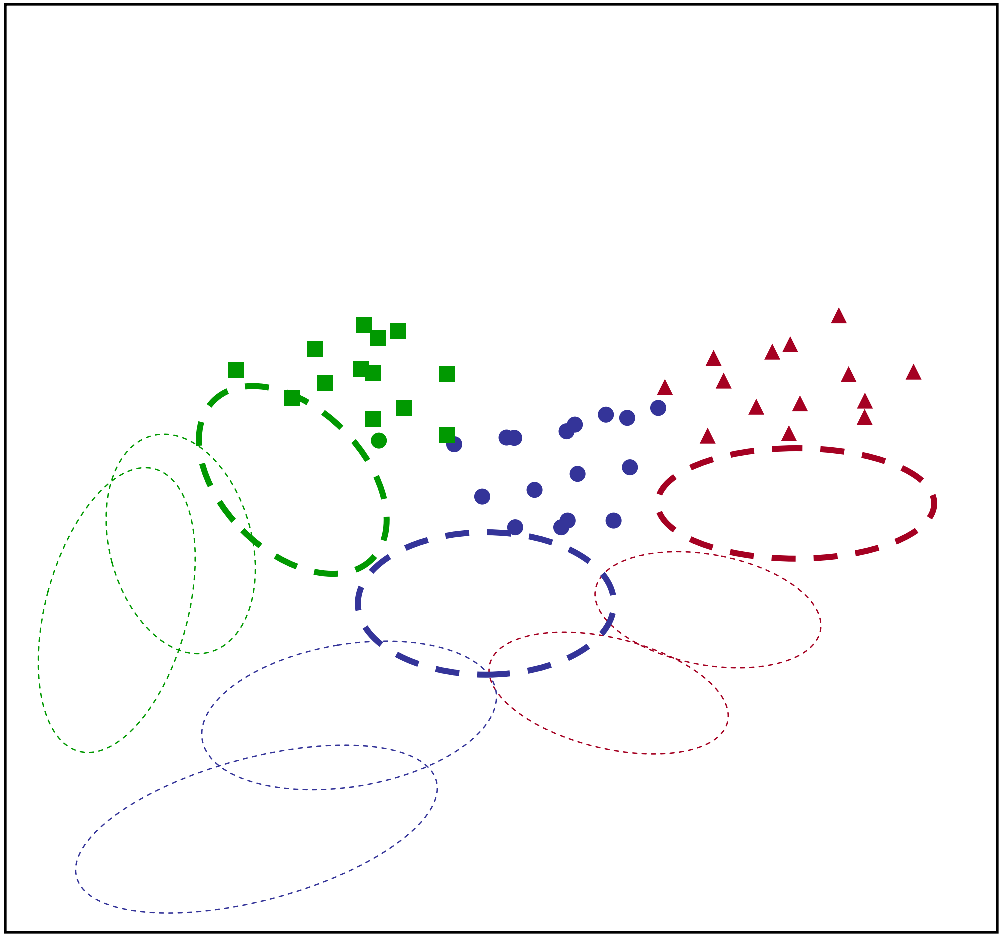
# Adaptive clustering (4)



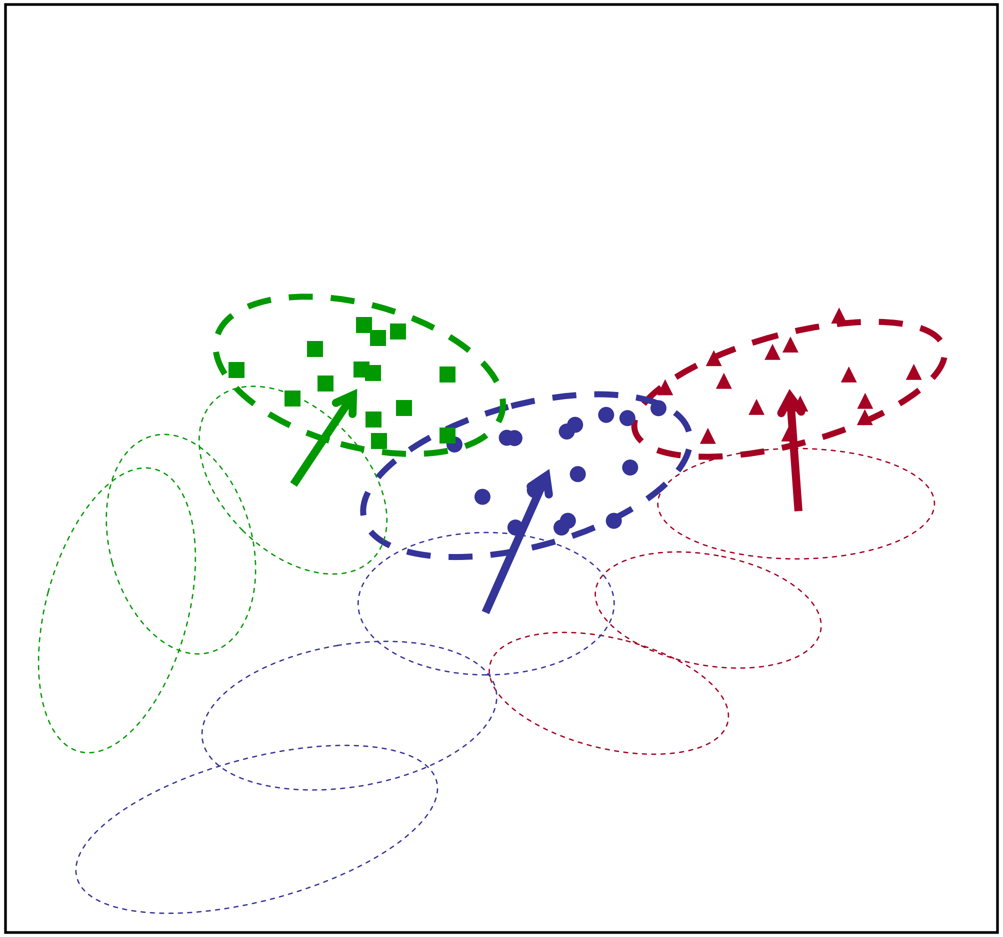
# Adaptive clustering (5)



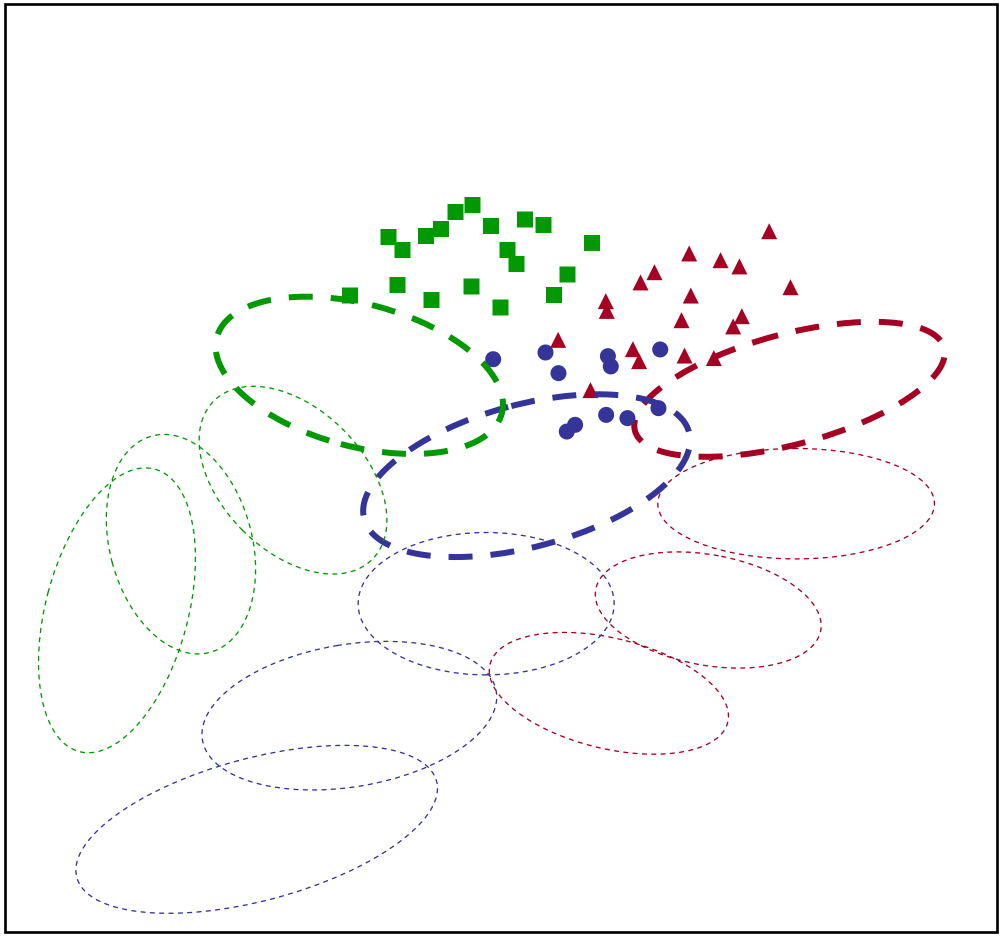
# Adaptive clustering (6)



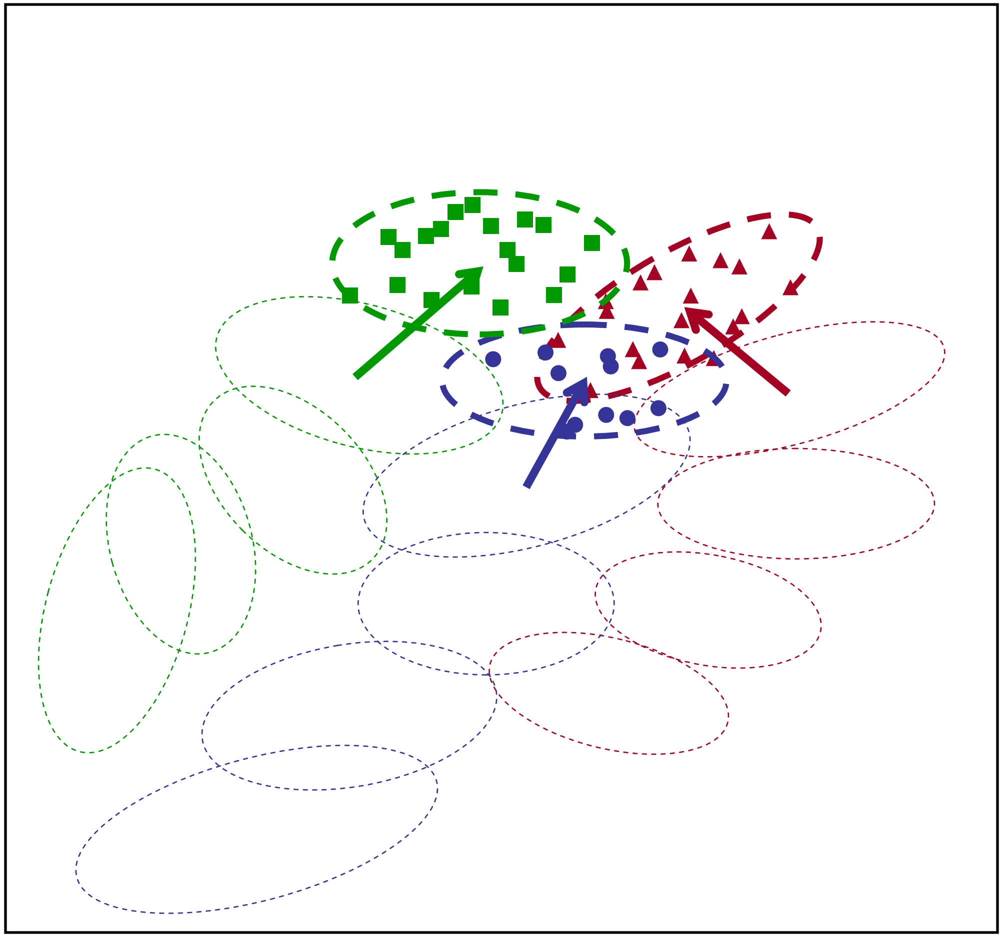
# Adaptive clustering (7)



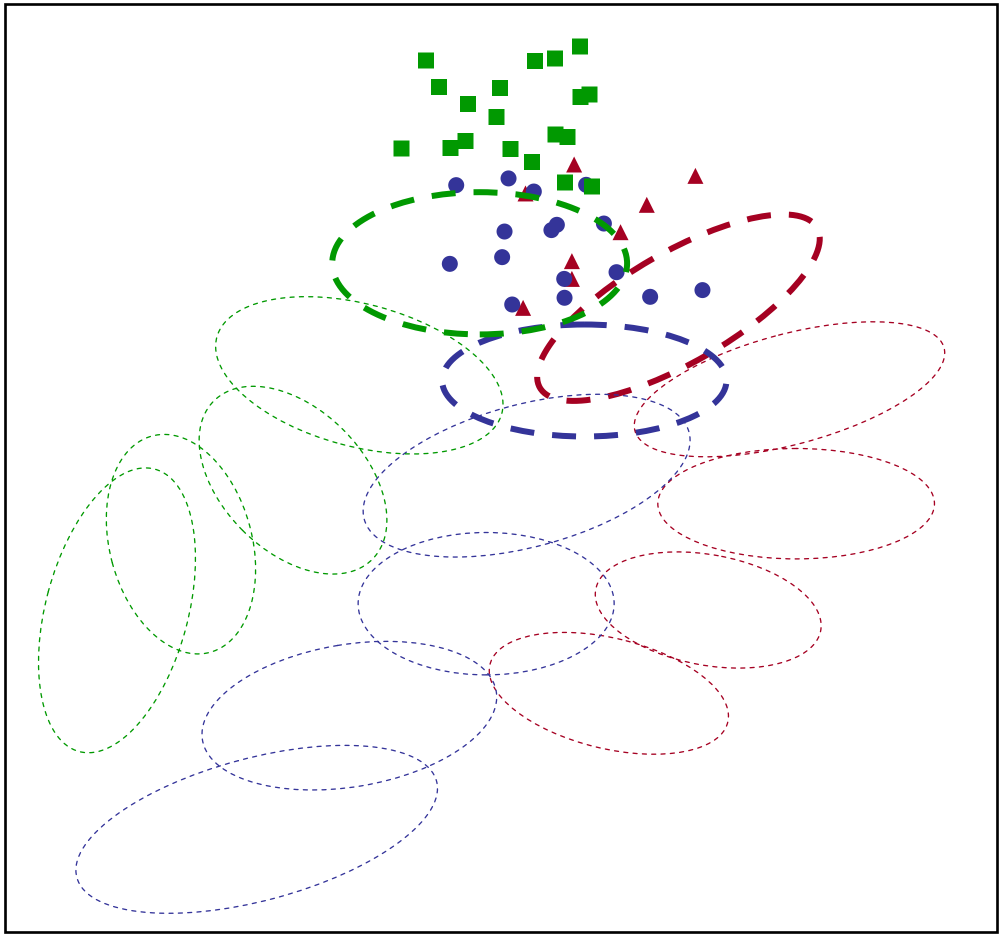
# Adaptive clustering (8)



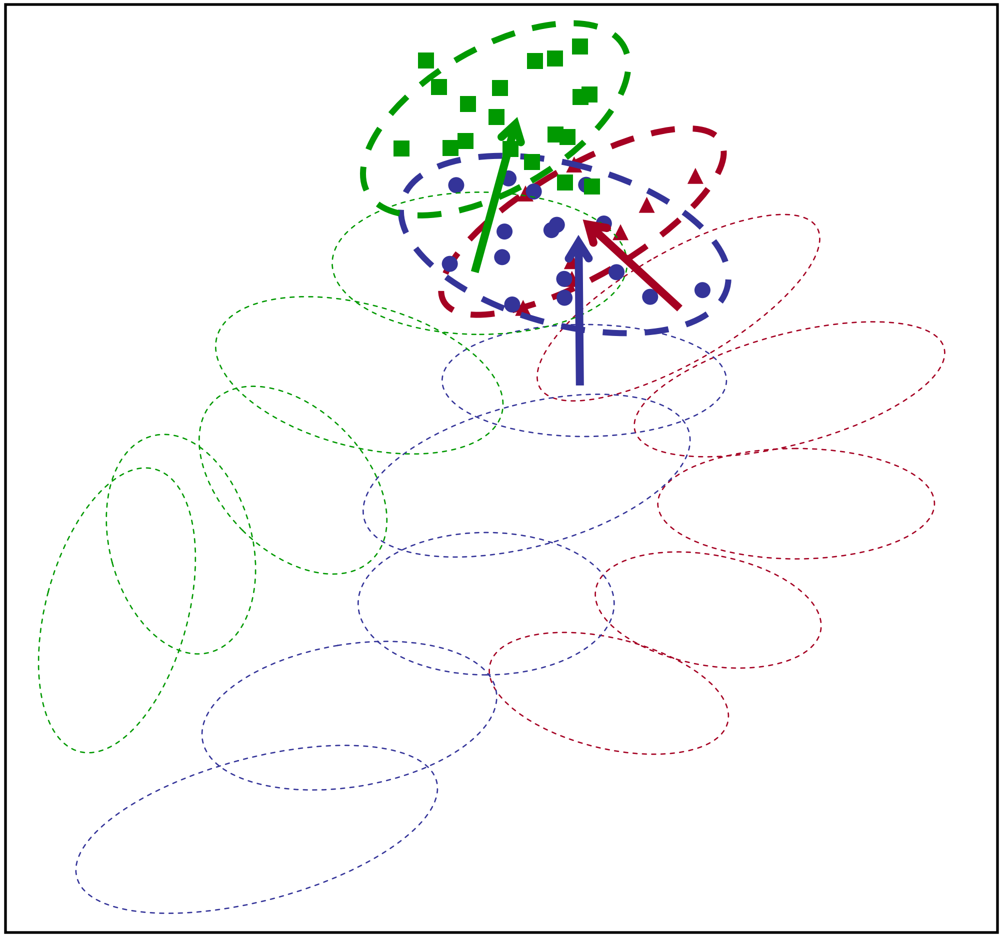
# Adaptive clustering (9)



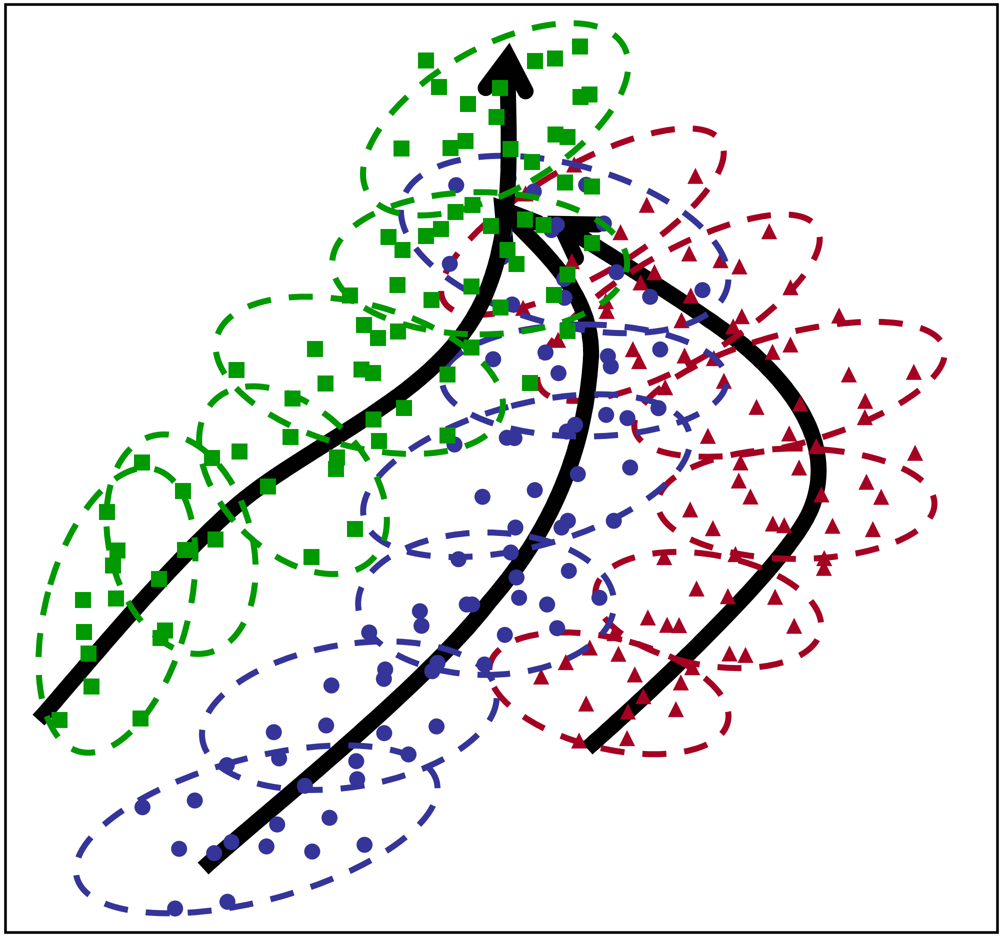
# Adaptive clustering (10)



# Adaptive clustering (11)



# Adaptive clustering (12)



# Adaptive clustering methods

- Mean updating
  - One cluster center per class [Holmberg et al., 1996]
- Kohonen self-organizing maps
  - One SOM common to all classes [Davide et al., 1994; Marco et al., 1997]
  - A separate SOM for each class [Distante et al., 2002]
- Adaptive Resonance Theory
  - ART is slightly different; new clusters can be created [Gardner et al., 1996]

# Adaptive clustering: discussion

- Algorithm relies on correct classification
  - Miss-classifications will eventually cause the model to lose track of the class patterns
- All odors need to be sampled frequently to prevent their patterns to drift too far

- This problem has also been addressed in the machine learning literature
  - see [Freund and Mansour, 1997] and refs. therein

# System identification [Holmberg et al., 1996]

- Basic idea
  - Chemical sensor responses co-vary over time
    - This "**common-mode**" behavior can be modeled with a dynamic model (e.g., ARMAX)

$$\sum_{n=0}^{|A|} a_n \underbrace{y_s(k-n)}_{\text{sensor } s} = \sum_{\substack{i=1 \\ i \neq s}}^{|S|} \sum_{n=0}^{|B|} b_{in} \underbrace{y_i(k-n)}_{\substack{\text{all other} \\ \text{sensors}}} + \sum_{n=0}^{|C|} c_n \underbrace{e(k-n)}_{\text{white noise}}$$

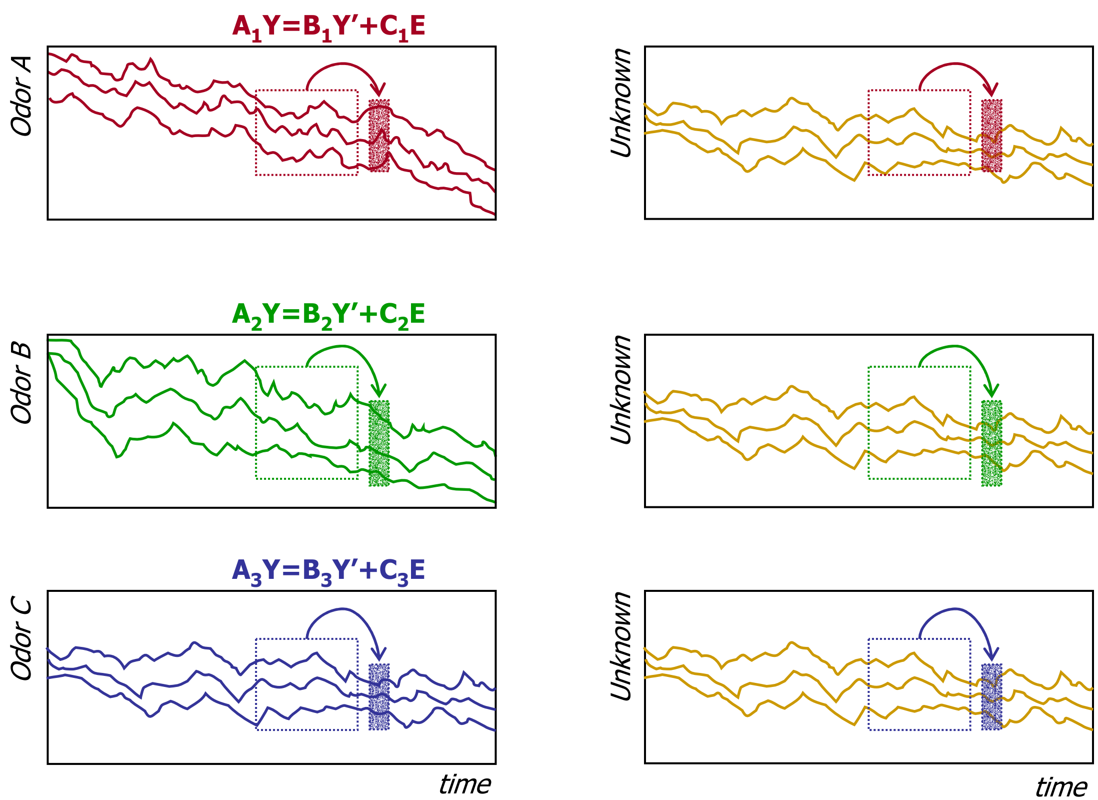
      - where $y_s(k)$ is the response of sensor 's' at time k, and $y_i(k)$ is the response from *all other* sensors
      - Model parameters {A,B,C} may be adapted over time with a recursive least-squares procedure [Holmberg et al., 1997]

# System identification: classification

- Each odor/sensor pair has a unique dynamic behavior that can be used as a fingerprint
  - Build a dynamic model for each sensor/odor
  - When an unknown odor is presented, predict its behavior of each sensor with each of the models
    - The method requires that multiple (consecutive) samples of the unknown odor be collected
  - The model with lowest prediction error corresponds to the true odor

# System identification: illustration

**MODEL DYNAMIC BEHAVIOR FOR EACH ODOR/SENSOR** → **APPLY EACH O/S MODEL TO UNKNOWN SAMPLE** → **CHOOSE ODOR MODEL WITH LOWEST ERROR**




ERROR 1

ERROR 2

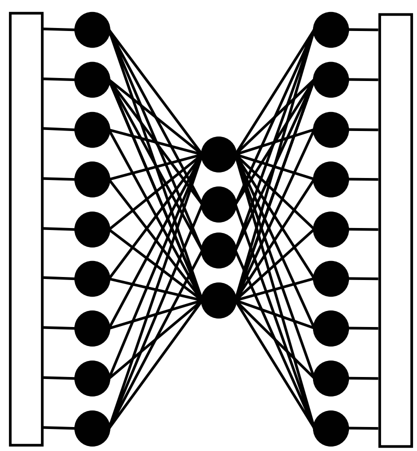
**ERROR 3**

# Calibration transfer

- Learn a regression mapping (e.g., MLPs, PLS) from drifting calibrant samples onto a baseline $t_0$
  - MLPs for PyMS [Goodacre and Kell, 1996]
  - PLS for e-noses [Tomic et al., 2002]
  - MLPs for e-noses [Balaban et al., 2000]

***Training phase:***
*Drifting calibrant samples at times $t_1$, $t_2$, …$t_N$*

***Recall phase:***
*Drifting odor samples at times $t_1$, $t_2$, …$t_N$*



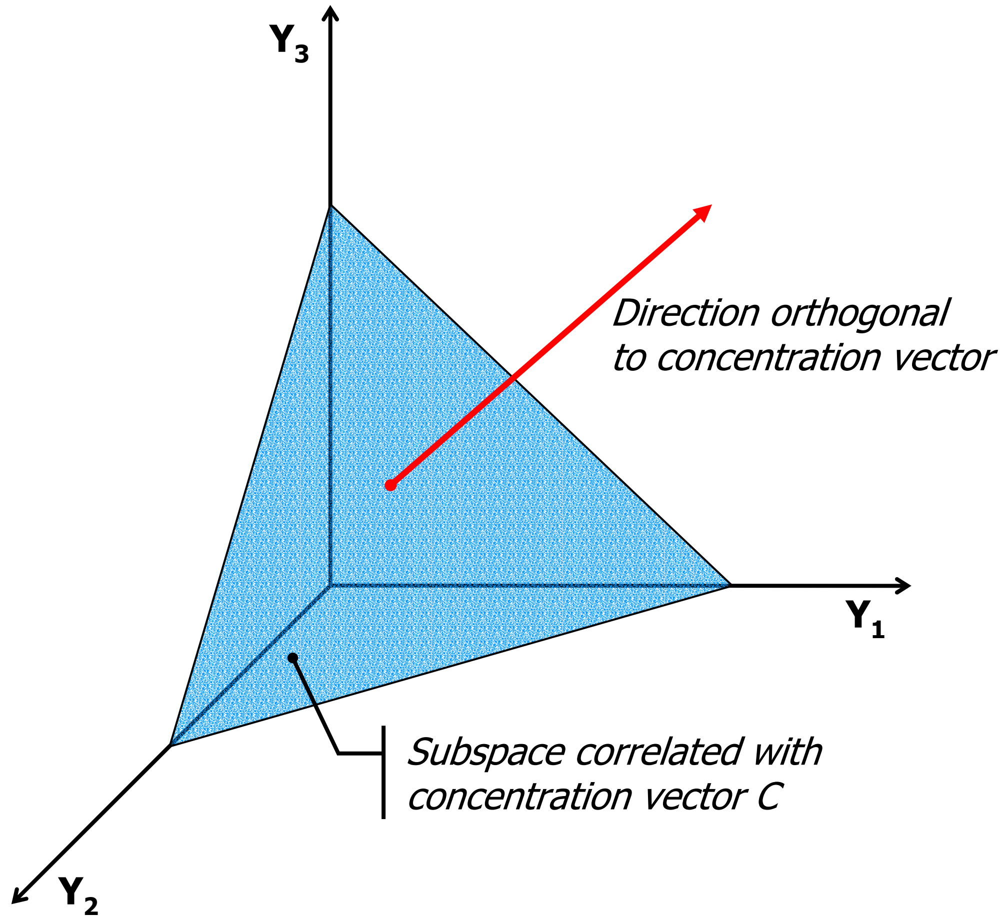
***Training phase:***
*Baseline calibrant sample at time $t_0$*

***Recall phase:***
*Corrected odor samples at times $t_1$, $t_2$, …$t_N$*

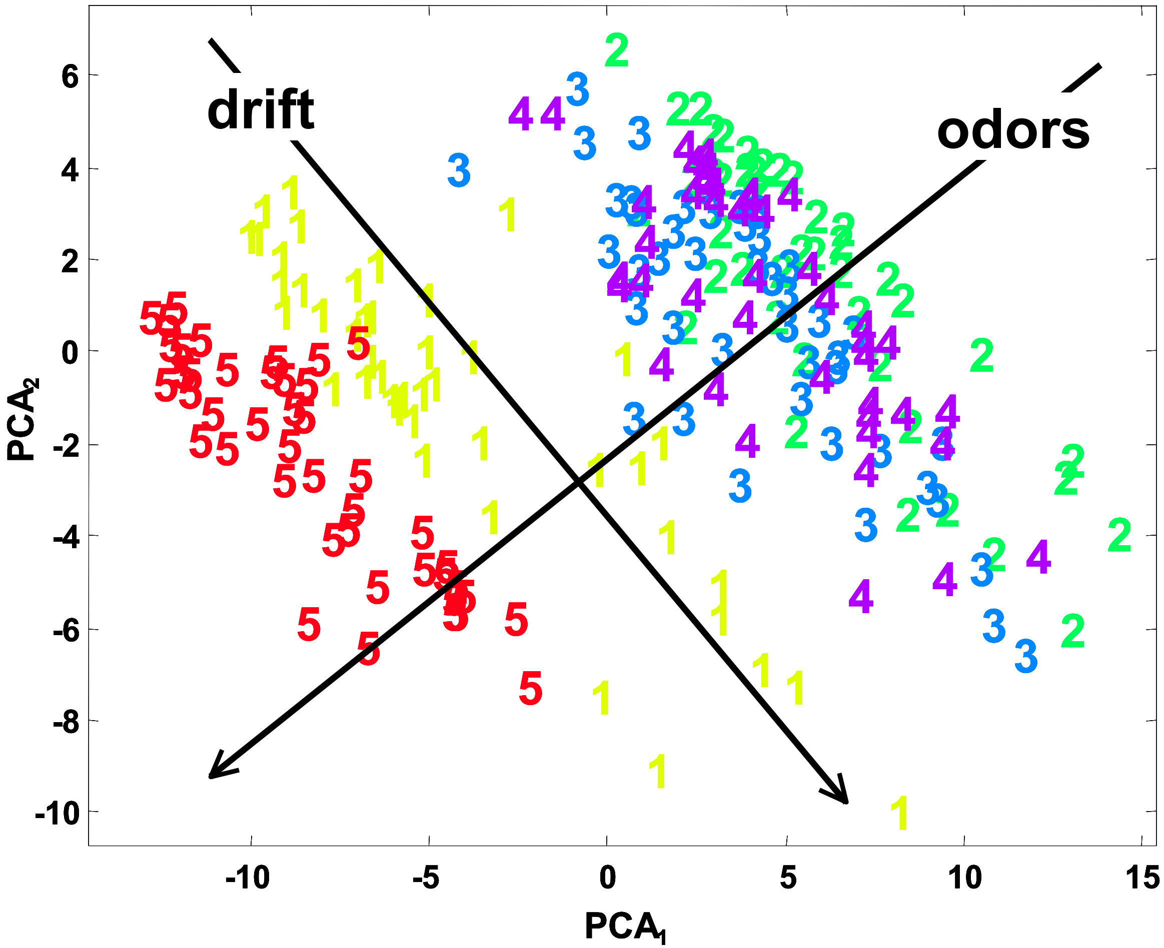
# Orthogonal signal correction [Wold et al., 1998]

- Basic idea
  - Assume dataset matrices
    - Y: sensor-array data (independent variables)
    - C: concentration vector or class label (dependent)
  - Subtract from Y factors that account for as much of the variance in Y as possible <u>AND</u> are orthogonal to C

# Orthogonal signal correction: intuition



# Orthogonal signal correction: an example



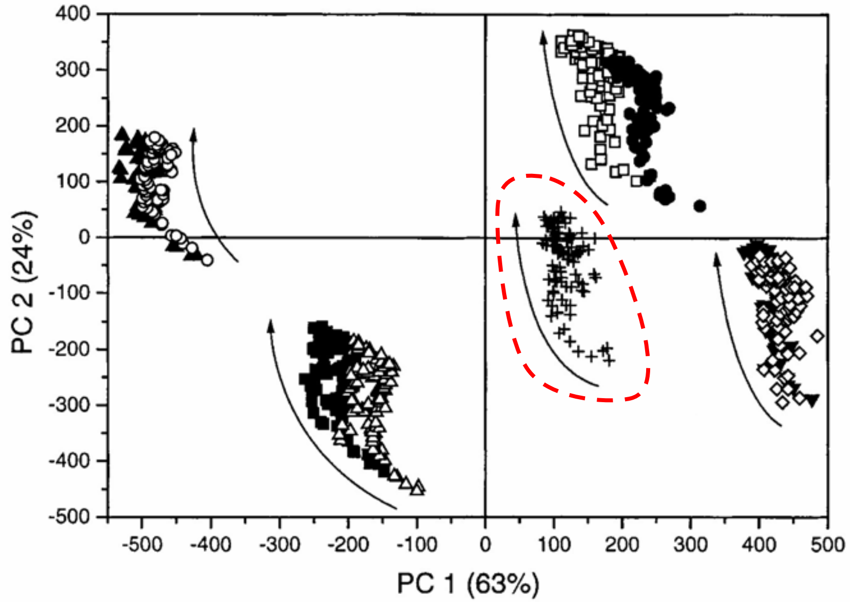
# Component correction [Artursson et al., 2000]

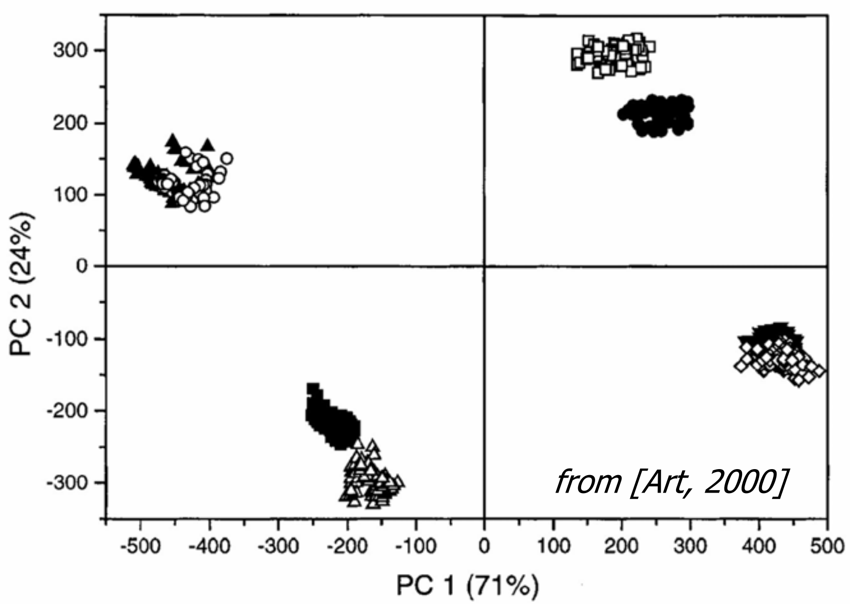
- Basic idea
  - Drift is associated with the principal components of variance in a calibration gas
  - These directions are removed from the multivariate sensor response by means of a bilinear transformation
- Algorithm

$$x_{corrected} = x - (x \cdot v_{cal})v_{cal}^T$$

  - where $v_{cal}$ is the first eigenvector of the calibration data $x_{cal}$

# Component correction results

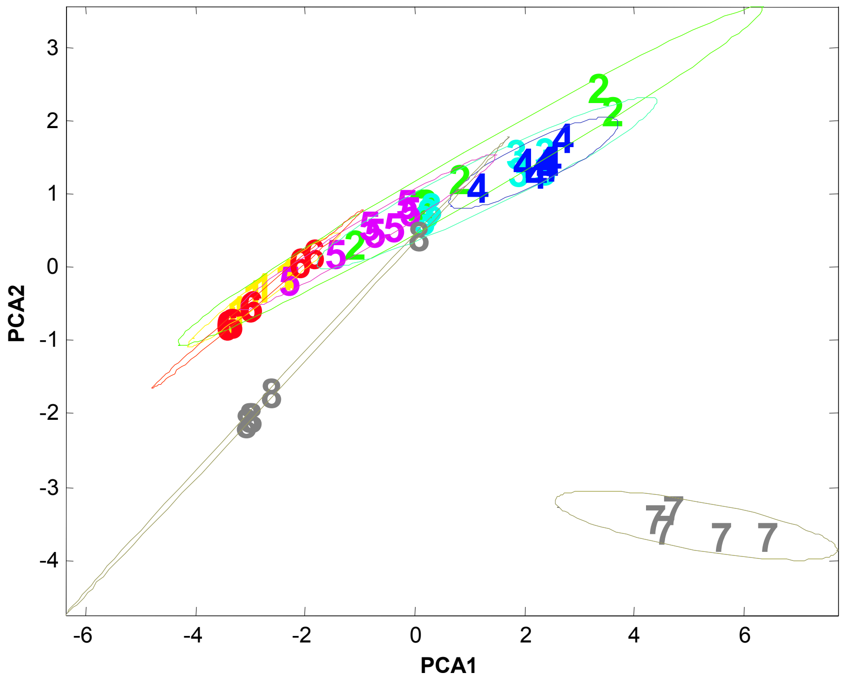


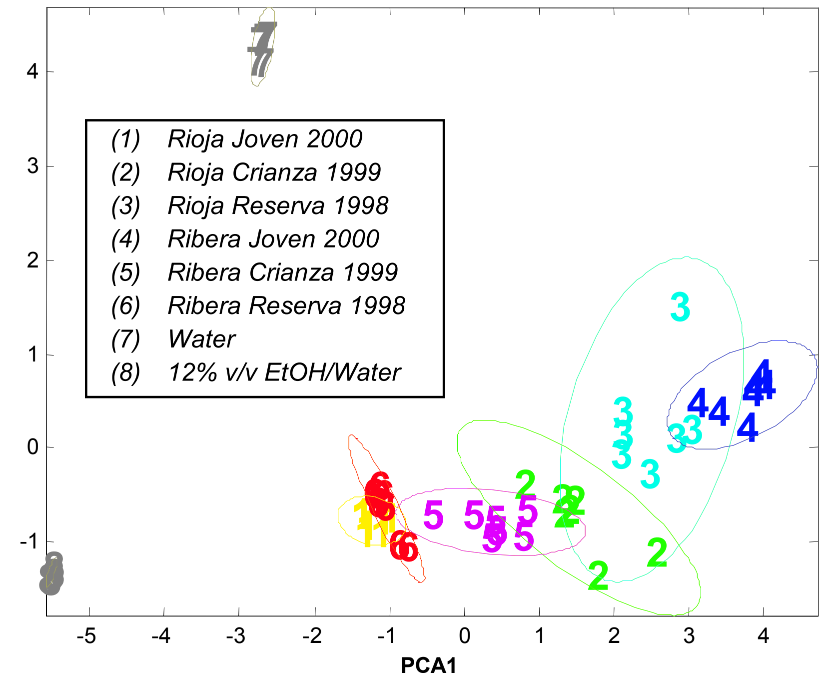




*PCA scatter plot of uncorrected samples for eight gas mixtures. Arrows indicate the direction of drift. The center cluster (+) is the calibration gas.*

*PCA scatter plot after component correction. The calibration gas (+), no longer shown in the figure, has been used to estimate and remove the principal direction of drift ($v_{cal}$).*

# Component correction results (2)



# Component deflation [Gutierrez, 2000]

- Basic idea
  - Identify variables 'x' whose variance can be attributed to drift or interferents
    - E.g., response to a wash/reference gas, time stamps, temperature, pressure, humidity, etc.
  - Measure 'y', the sensor-array response to an odor
  - Remove variance in 'y' that can be explained by 'x' (by means of regression/deflation)

- Related to target rotation [Esbensen et al., 1987]
  - "…removal of undesired information provided that there are variables uniquely connected to that information" [Christie, 1996]

# Component deflation algorithm

- Find linear projections x′=Ax and y′=By that are maximally correlated

$$\{A,B\} = \arg\max[\rho(Ax,By)]$$

  - How? Canonical Correlation Analysis (CCA) or PLS
  - Interpretation: x′ and y′ are low-dimensional projections that summarize the linear dependencies between x and y

- Find regression model $y_{pred}=Wy'$

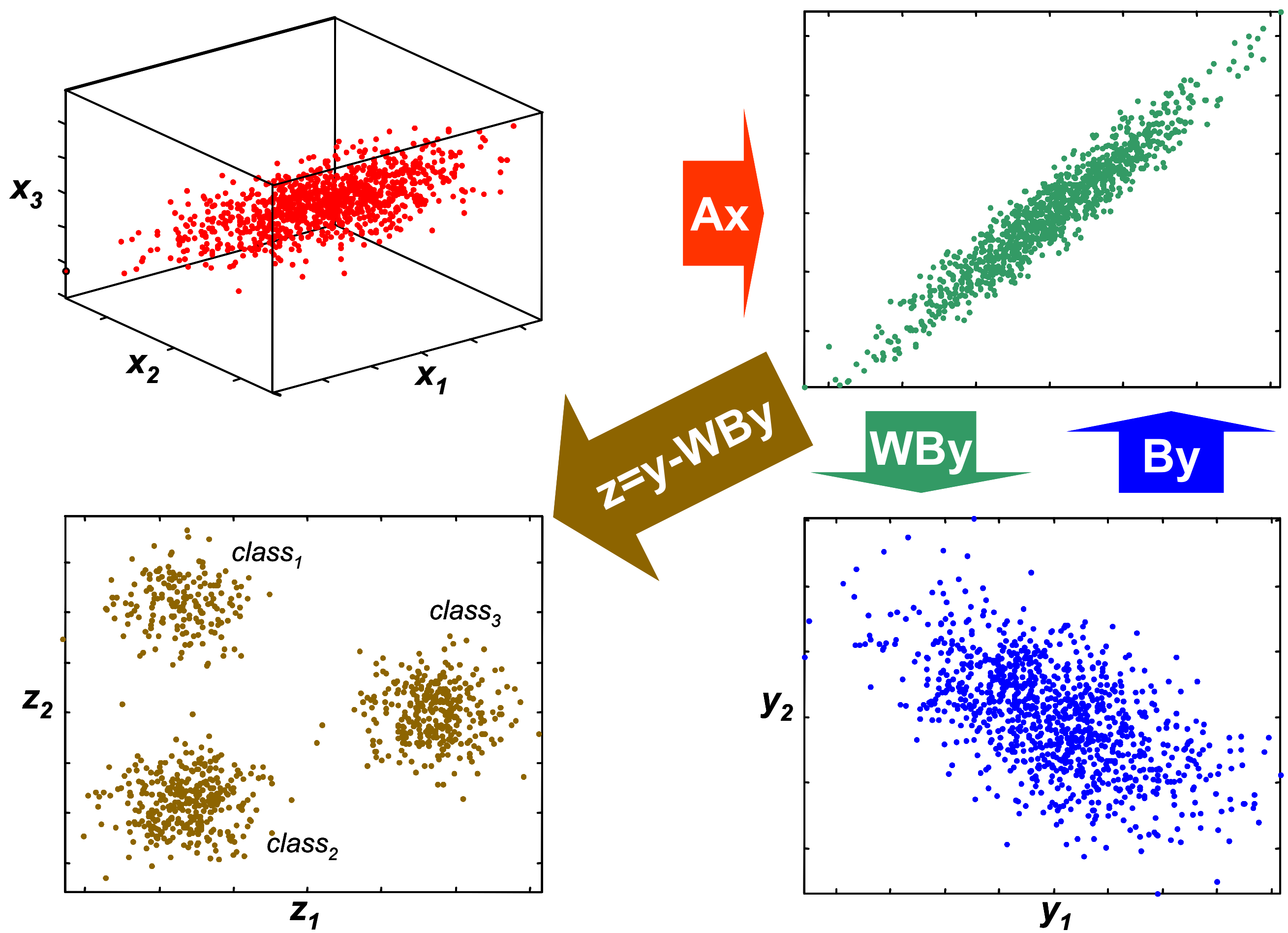
$$W = \arg\min|y - Wy'|^2$$

  - Interpretation: $y_{pred}$ contains the variance in the odor vector y that can be explained by y′ and, as a result of the CCA stage, by x

- Deflate y and use the residual z as a corrected sensor response

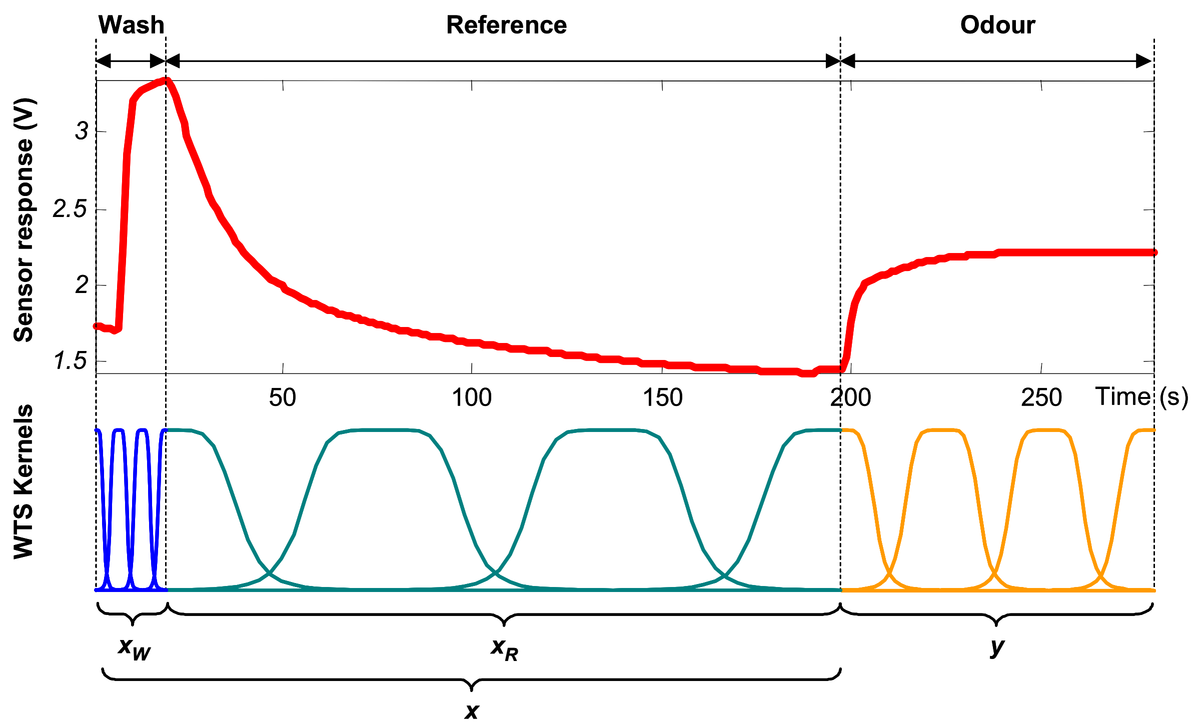
$$z = y - y_{pred} = y - WBy$$

# Component deflation: interpretation



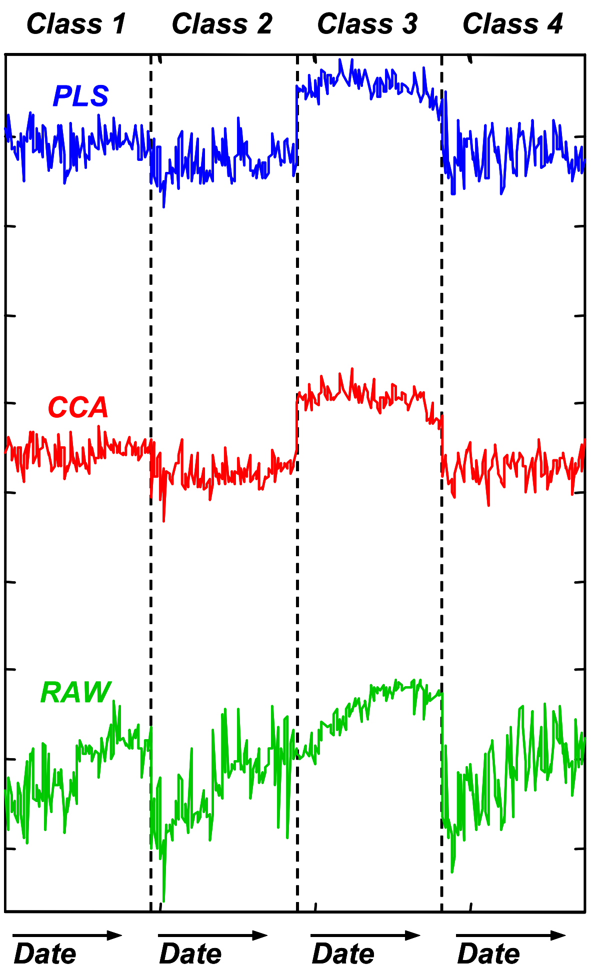
# Component deflation: motivation

- Exploit transient information in wash/reference cycle
  - To capture temporal trends, augment vector x with the time stamp of each sample (also in [Artursson et al., 2000])

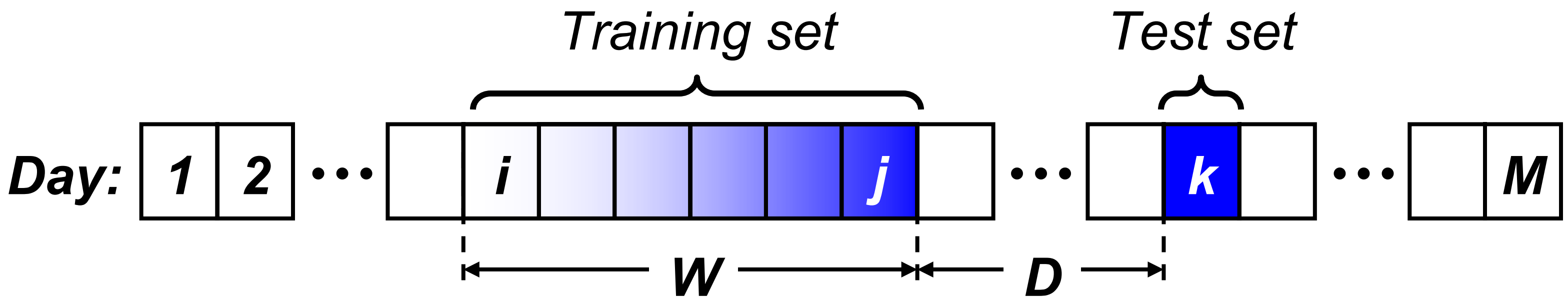


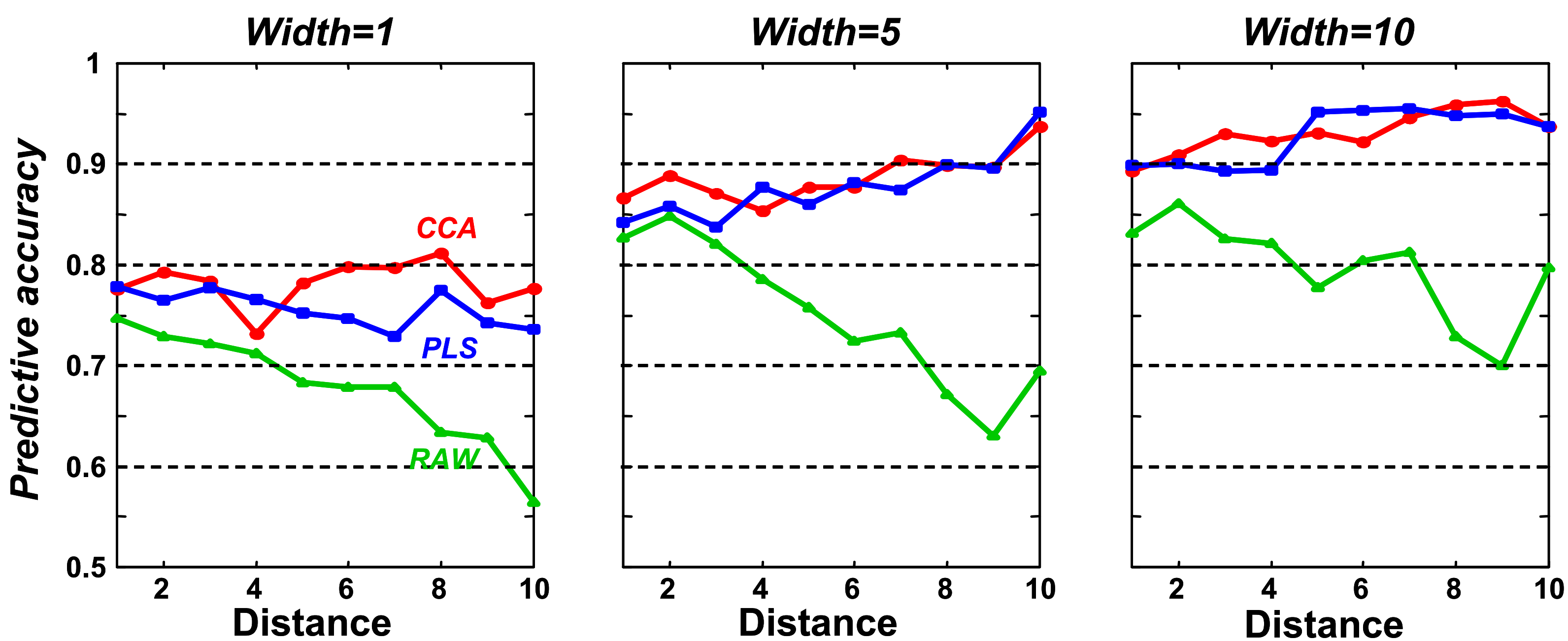
# Component deflation performance

- Database
  - Three months of data collection
    - Four cooking spices
    - Twenty four days
    - Four samples per day and spice
  - Ten metal-oxide sensors
- Plot shows one particular transient feature
  - Examples are sorted
    - By odor class
    - Within a class, by date



# Influence of aging and training set size



# Discussion of univariate methods

- Frequency decomposition
  - Mostly useful for diagnostics and analysis
- Baseline manipulation
  - Attractive for its simplicity, but not an effective drift-correction approach for chemical sensors
- Differential measurements wrt calibrant
  - First-order approximation, but does not exploit cross-correlations
- Multiplicative correction wrt calibrant
  - Empirically shown to work, heuristic, does not exploit cross-correlations

# Discussion of multivariate methods (1)

- Adaptive clustering
  - Requires equiprobable and frequent sampling
  - Correct identification is essential to ensure that the adapting distributions track the odors
  - Does not use information from a calibrant
- System identification
  - Builds a separate drift model for each odor, but requires multiple consecutive samples
  - Can be used as a template-matching classifier
  - Does not use information from a calibrant

# Discussion of multivariate methods (2)

- Calibration transfer
  - Shown to work in MS and e-nose data
  - Black-box approach
- Orthogonal Signal Correction and Deflation
  - <u>In my experience, the best approach</u>
  - Why? Time tested in chemometrics, uses multivariate and calibrant information, and IT IS SIMPLE

# References (1)


- M. Holmberg and T. Artursson, "Drift Compensation, Standards and Calibration Methods," in T. C. Pearce, S. S. Schiffman, H. T. Nagle and J. W. Gardner (Eds.), Handbook of Machine Olfaction: Electronic Nose Technology, Weinheim, Germany: Wiley-VCH, 2002, pp.325-346.
- S.-Y. Choi, K. Takahashi, M. Esani and T. Matsuo, "Stability and sensitivity of MISFET hydrogen sensors," in *Proc. Int. Conf. Solid State Sensors and Actuators*, Philadelphia, PA, 1985.
- M. Roth, R. Hartinger, R. Faul and H.-E. Endres, "Drift reduction of organic coated gas-sensors by temperature modulation," *Sensors and Actuators B 36(1-3)*, pp. 358-362, 1996.
- D.M. Wilson and S.P. DeWeerth, "Odor discrimination using steady-state and transient characteristics of tin-oxide sensors," *Sensors and Actuators B 28(2)*, pp. 123-128. 1995.
- J.W. Gardner and P.N. Bartlett, <u>Electronic Noses. Principles and Applications</u>, New York: Oxford University Press, 1999.
- J.W. Gadner, "Detection of Vapours and Odours from a Multisensor Array using Pattern recognition. Part 1. Principal Components and Cluster Analysis," in *Sensors and Actuators B 4*, pp. 109-115, 1991.
- M. Fryder, M. Holmberg, F. Winquist and I. Lündstrom, "A calibration technique for an electronic nose," in *Eurosensors IX*, 1995, pp. 683-686.
- M.L. Salit and G.C Turk, "A Drift Correction Procedure," Analytical Chemistry 70(15), pp. 3184-3190, 1998.
- J.E. Haugen, O. Tomic and K. Kvaal, "A calibration method for handling the temporal drift of solid-state gas sensors," *Analytica Chimica Acta 407*, pp. 23-39, 2000.
- T. Artursson, T. Eklöv, I. Lundström, P. Mårtensson, M. Sjöström and M. Holmberg, "Drift correction for gas sensors using multivariate methods," *Journal of Chemometrics 14(5-6)*, pp. 711-723, 2000.

# References (2)


- F.A.M. Davide, C. Di Natale, M. Holmberg and F. Winquist, "Frequency Analysis of Drift in Chemical Sensors," in Proc. 1st Italian Conf. on Sensors and Microsystems, Rome, Italy, 1996, pp. 150-154.
- R. Goodacre and D.B. Kell, "Correction of mass spectral drift using artificial neural networks," *Analytical Chemistry 68*, pp. 271-280, 1996.
- O. Tomic, H. Ulmer and J.E. Haugen, "Standardization methods for handling instrument related signal shifts in gas-sensor array measurement data," *Analytica Chimica Acta 472*, 99-111, 2002.
- M.O. Balaban, F. Korel, A.Z. Odabasi and G. Folkes, "Transportability of data between electronic noses: mathematical methods," *Sensors and Actuators B 71*, pp. 203-211, 2000.
- S. Wold, H. Antti, F. Lindgren and J. Öhman, "Orthogonal signal correction of near-infrared spectra," *Chemometrics and Intelligent Laboratory Systems 44*, pp. 175-185, 1998.
- K. Esbensen, L. Lindqvist, I. Lundholm, D. Nisca and S. Wold, "Multivariate Modeling of Geochemical and Geophysical Data," *Chemometrics and Intelligent Laboratory Systems 2*, pp. 161-175, 1987.
- O.H.J. Christie, "Data Laundering by Target Rotation in Chemistry-Based Oil Exploration," Journal of Chemometrics 10, pp. 453-461, 1996.
- M. Holmberg, F. Winquist, I. Lundström, F. Davide, C. DiNatale, and A. D'Amico, "Drift counteraction for an electronic nose," *Sensors and Actuators B 36(1-3)*, pp. 528-535, 1996.
- F.M. Davide, C. Di Natale, and A. D'Amico, "Self-organizing Multisensor Systems for Odour Classification: Internal Categorization, Adaptation, and Drift Rejection," *Sensors and Actuators B 18-19*, pp. 244-250, 1994.
- S. Marco, A. Pardo, A. Ortega and J. Samitier, "Gas Identification with Tin Oxide Sensor Array and Self-Organizing Maps: Adaptive Correction of Sensor Drifts," in Proc. IEEE Instrumentation and Measurement Technology Conference, Orrawa, Canada, May 19-21, 1997, pp. 904-907.

# References (3)


- C. Distante, P. Siciliano and K.C. Persaud, "Dynamic Cluster Recognition with Multiple Self-Organising Maps," Pattern Analysis and Applications 5, pp. 306-315, 2002.
- J. W. Gardner, E. L. Hines and C. Pang, "Detection of vapours and odours from a multisensor array using pattern recognition: self-organizing Adaptive Resonance Techniques," *Measurement + Control 29*, pp. 172-178, 1996.
- Y. Freund and Y. Mansour, "Learning under persistent drift," in *Proc. European Conference on Computational Learning Theory*, 1997, pp. 109-118.

# Questions

# Thank you