Biomedical speech signal processing: concepts, algorithms, and contemporary challenges

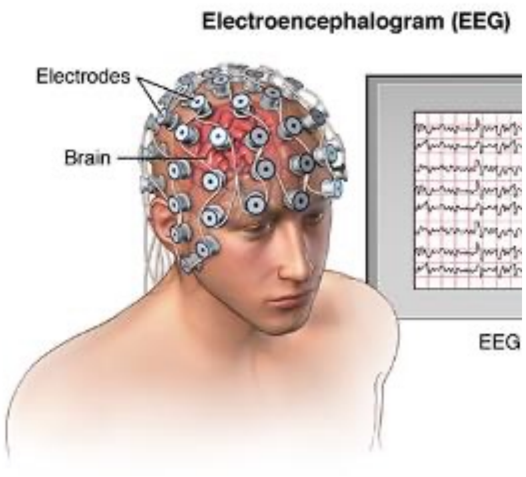
Dr Athanasios Tsanas ('Thanasis')
Associate Professor in Data Science
Usher Institute, Medical School
University of Edinburgh

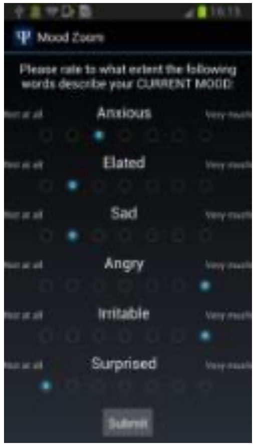
atsanas@ed.ac.uk

Research vision

DARTH

Address unmet clinical needs

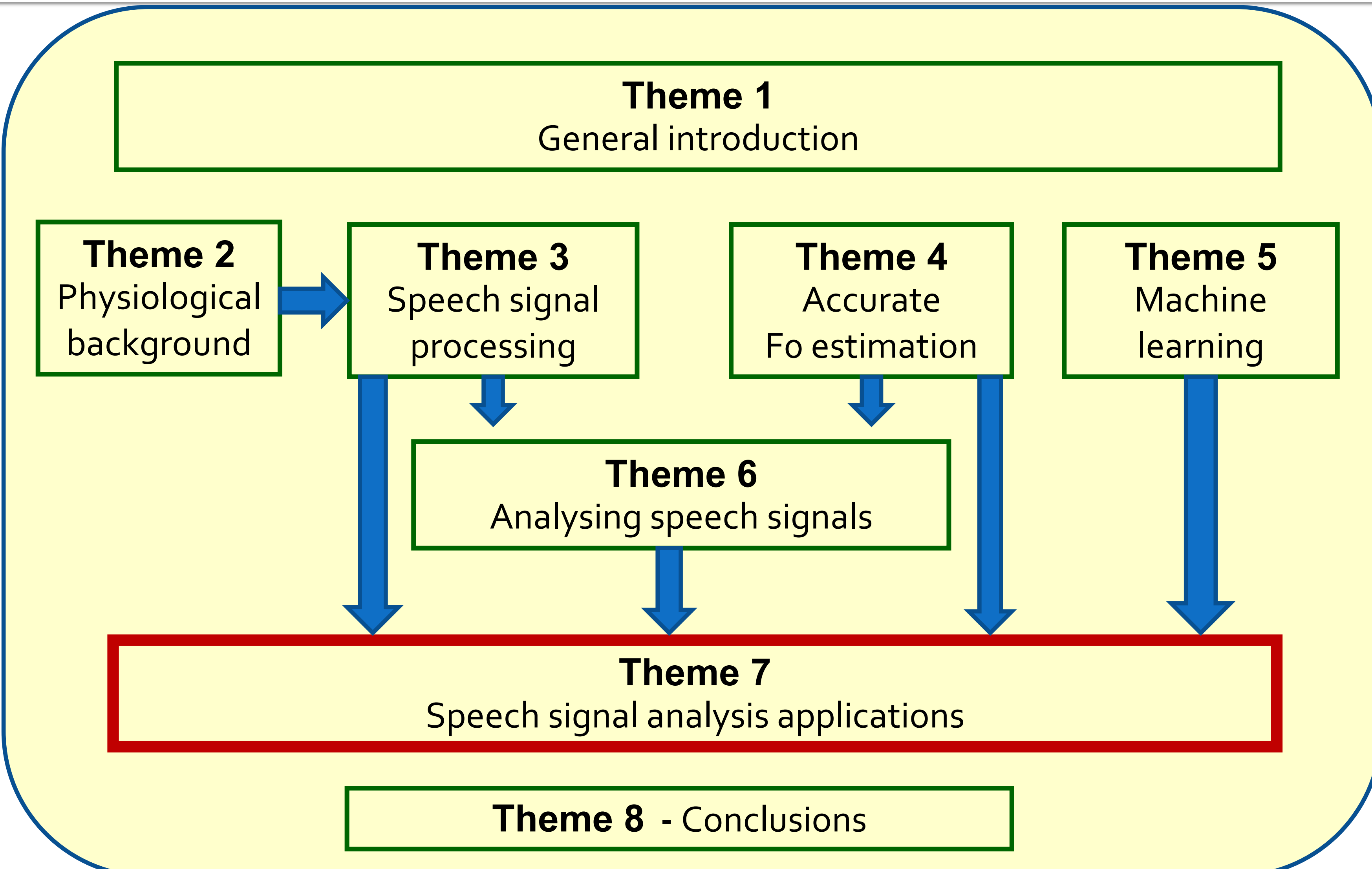




Assessment & Monitoring

- Frequent
- Unobtrusive
- Personalised
- Affordable
- Massive scale

Presentation structure



Speech types

Spontaneous speech

- Casual speech
- Describing e.g. image and eliciting emotions

Reading prescribed text

- Linguistically rich text, eliciting specific combinations
- Grandfather passage

Specific speech tasks

- Diadochokinetic test pa/ta/ka
- Sustained vowels, e.g. corner vowels /a/, /i/, /u/

Data models: two schools of thought

Differential equations

$$(+E)+(-Ri)+(-L\frac{di}{dt})+(-\frac{1}{C}q)=0$$

$$E-Ri-L\frac{di}{dt}-\frac{1}{C}q=0$$

$$E=Ri+L\frac{di}{dt}+\frac{1}{C}q$$

Differentiate both sides 1

2 $i=\frac{dq}{dt}$

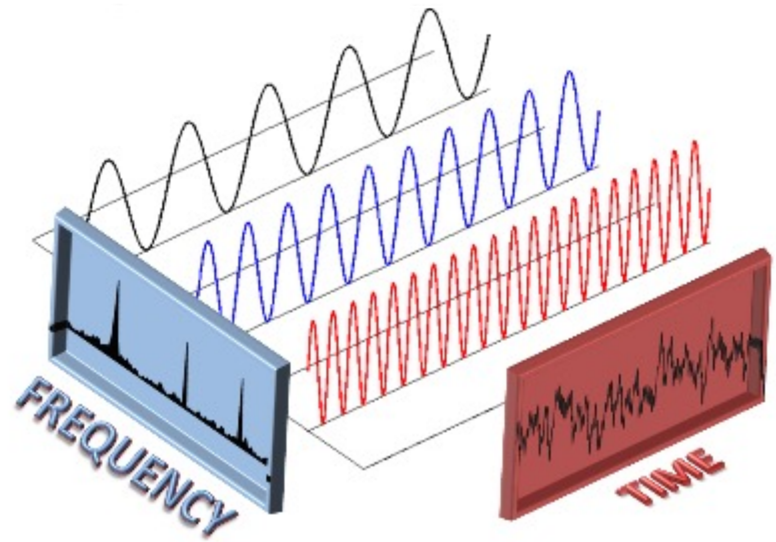
$$\frac{dE}{dt}=R\frac{di}{dt}+L\frac{d}{dt}\frac{di}{dt}+\frac{1}{C}\frac{dq}{dt}$$

$$E=R\frac{dq}{dt}+L\frac{d}{dt}(\frac{dq}{dt})+\frac{1}{C}q$$

Simplify the equation $i=\frac{dq}{dt}$

Simplify the equation

Statistics



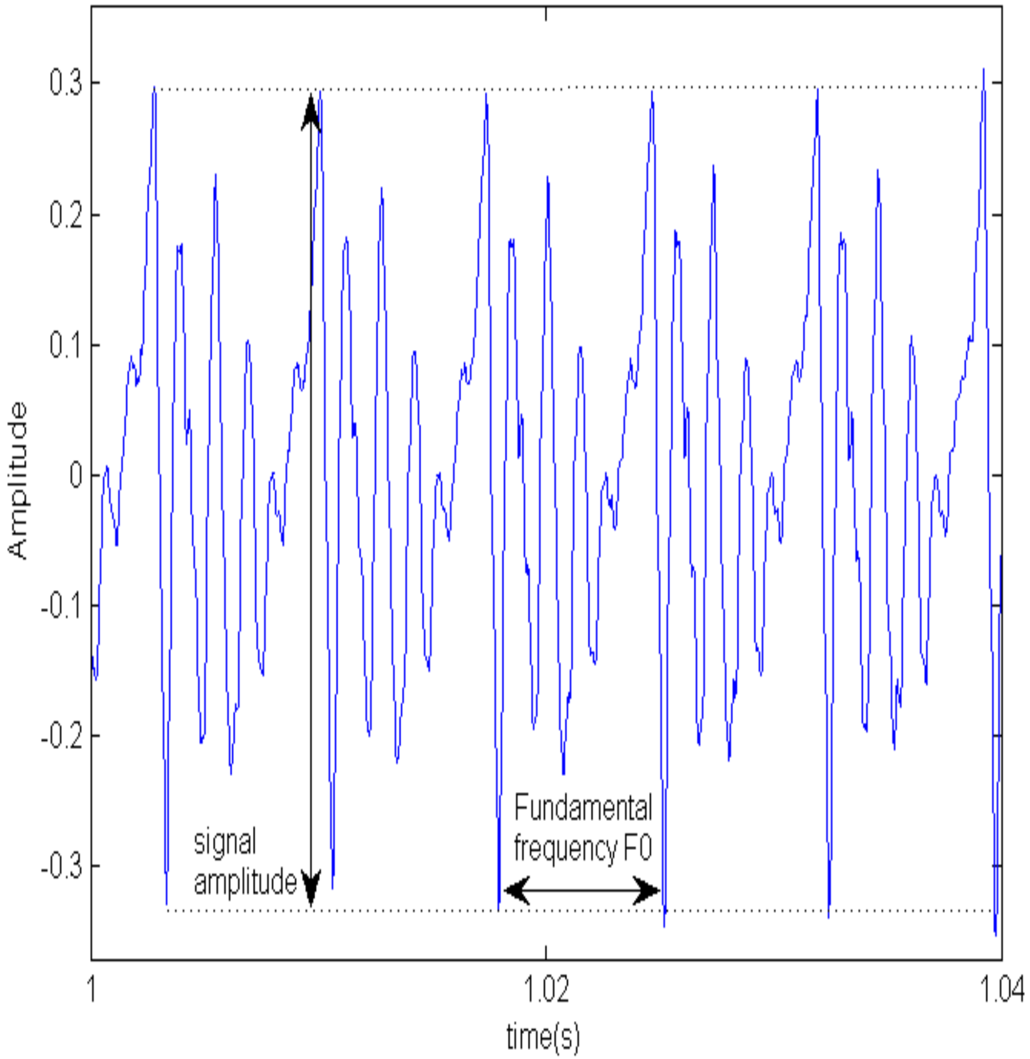
First principle models

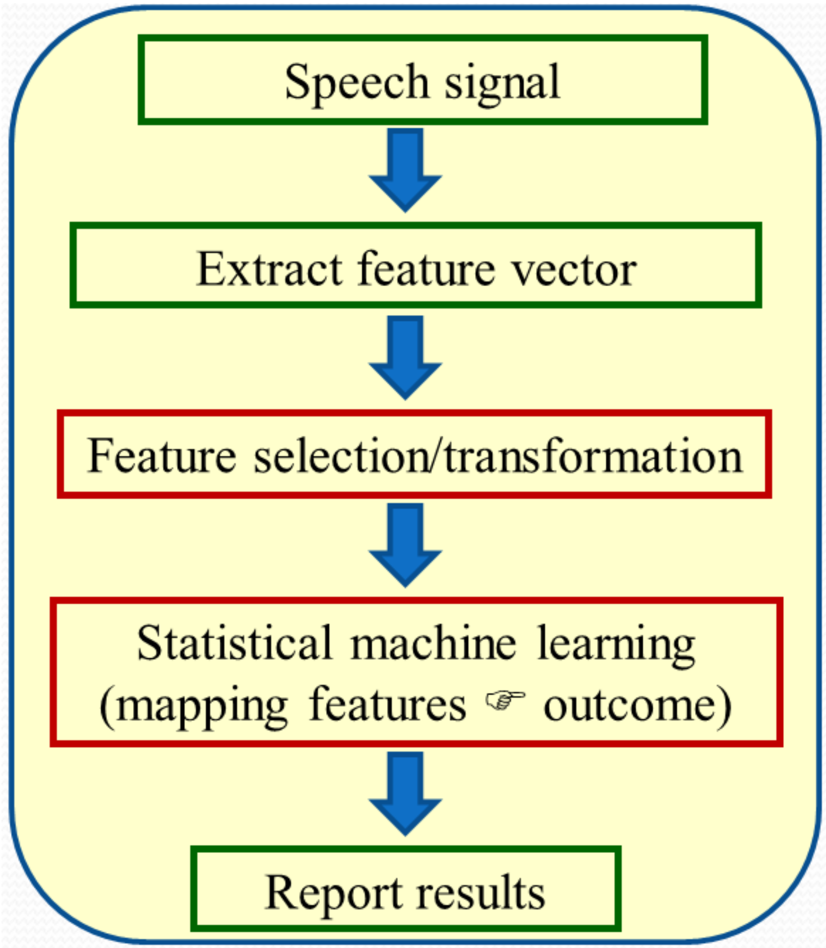
- Mechanistic insight ☺
- Difficult to match data ☹

Data driven models

- Less interpretable ☹
- Better predictions ☺

Time-series & pattern recognition





Overview of key steps

Data visualization (density estimation, scatter plots)

Exploratory analysis: hypothesis testing and statistical associations

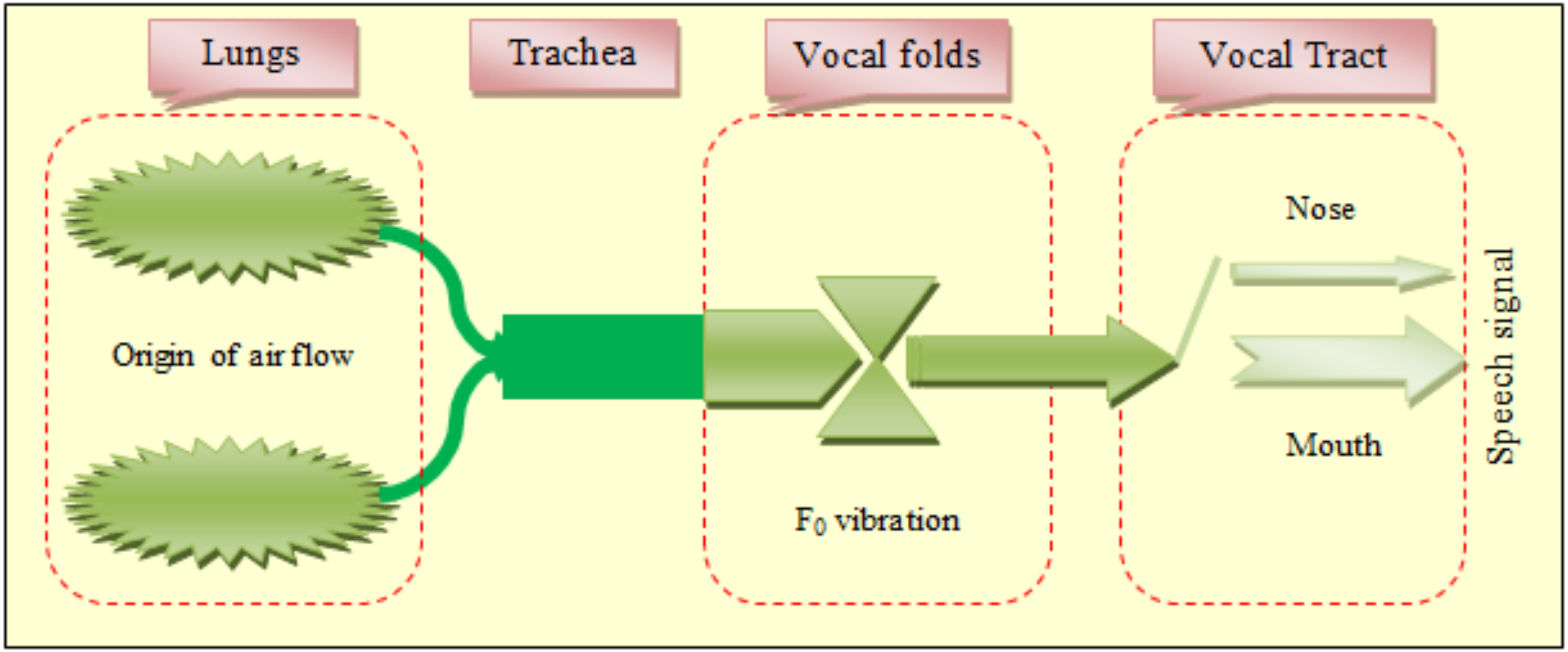
Feature selection or transformation (e.g. PCA)

Statistical mapping (regression/classification)

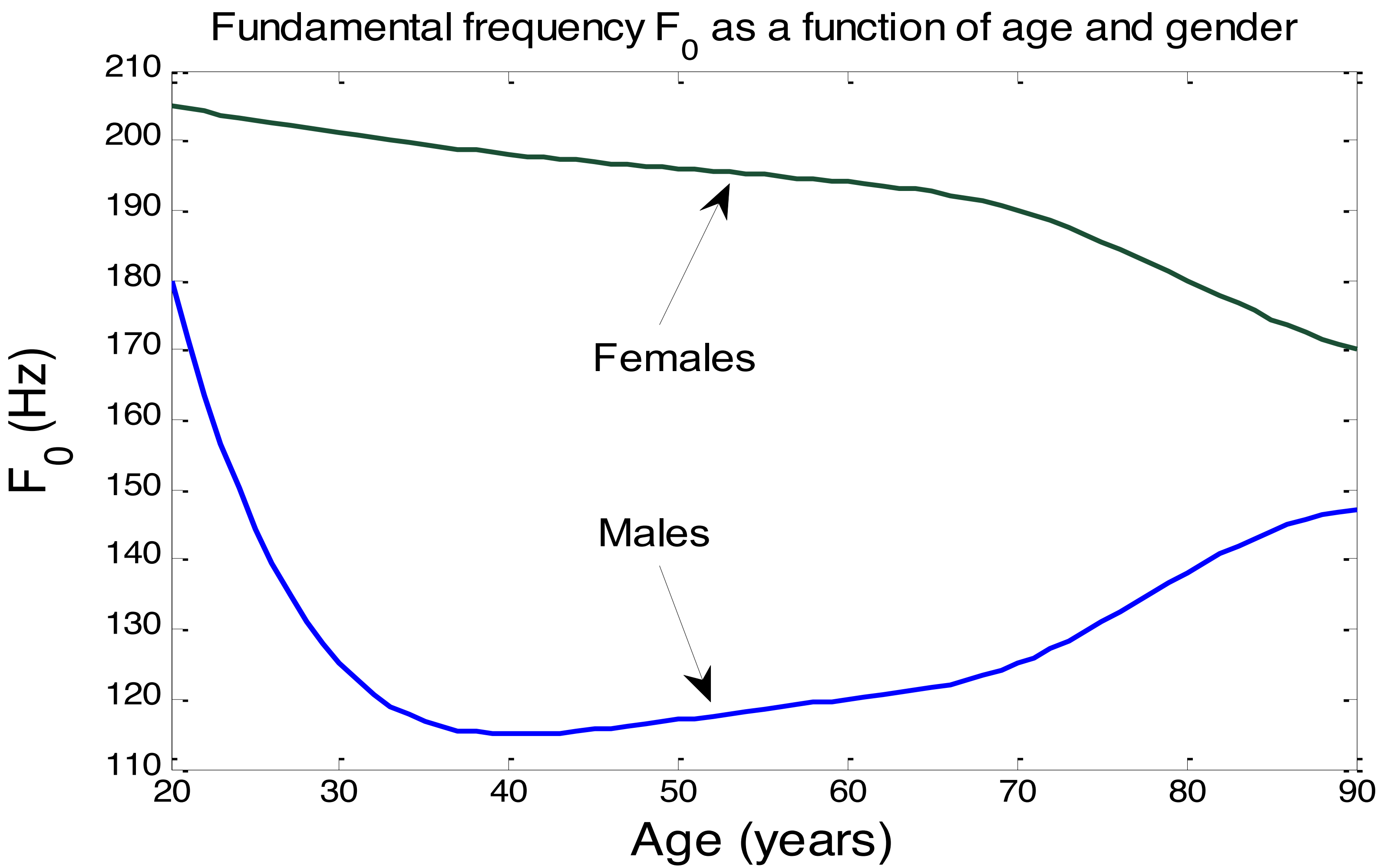
Physiological background

Theme 2

Voice production mechanism

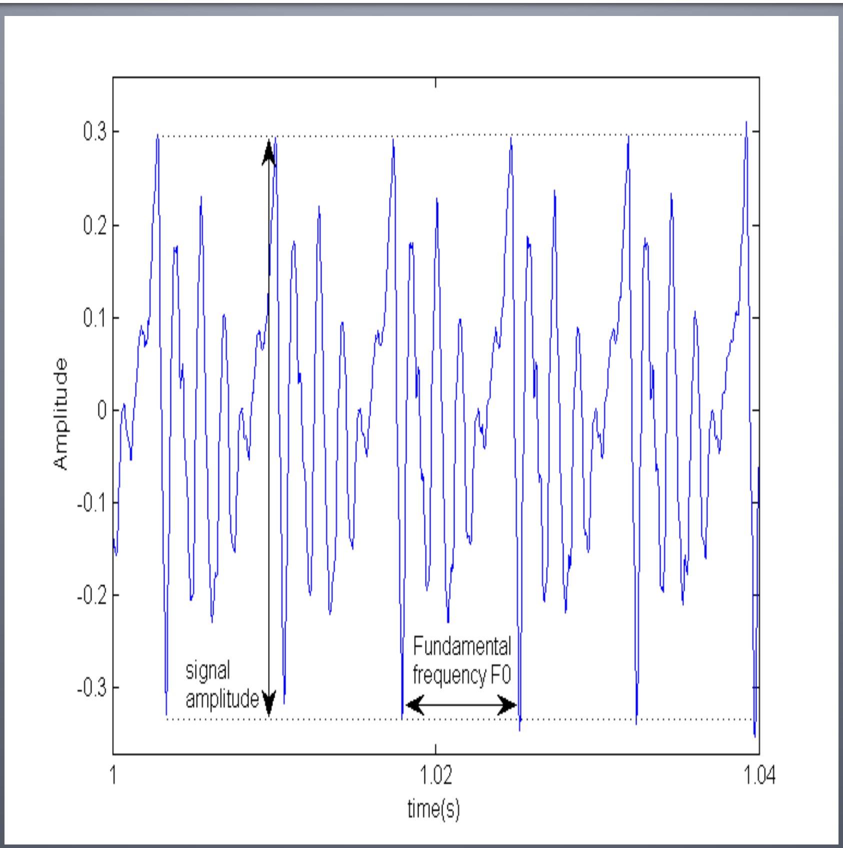


Gender-related differences

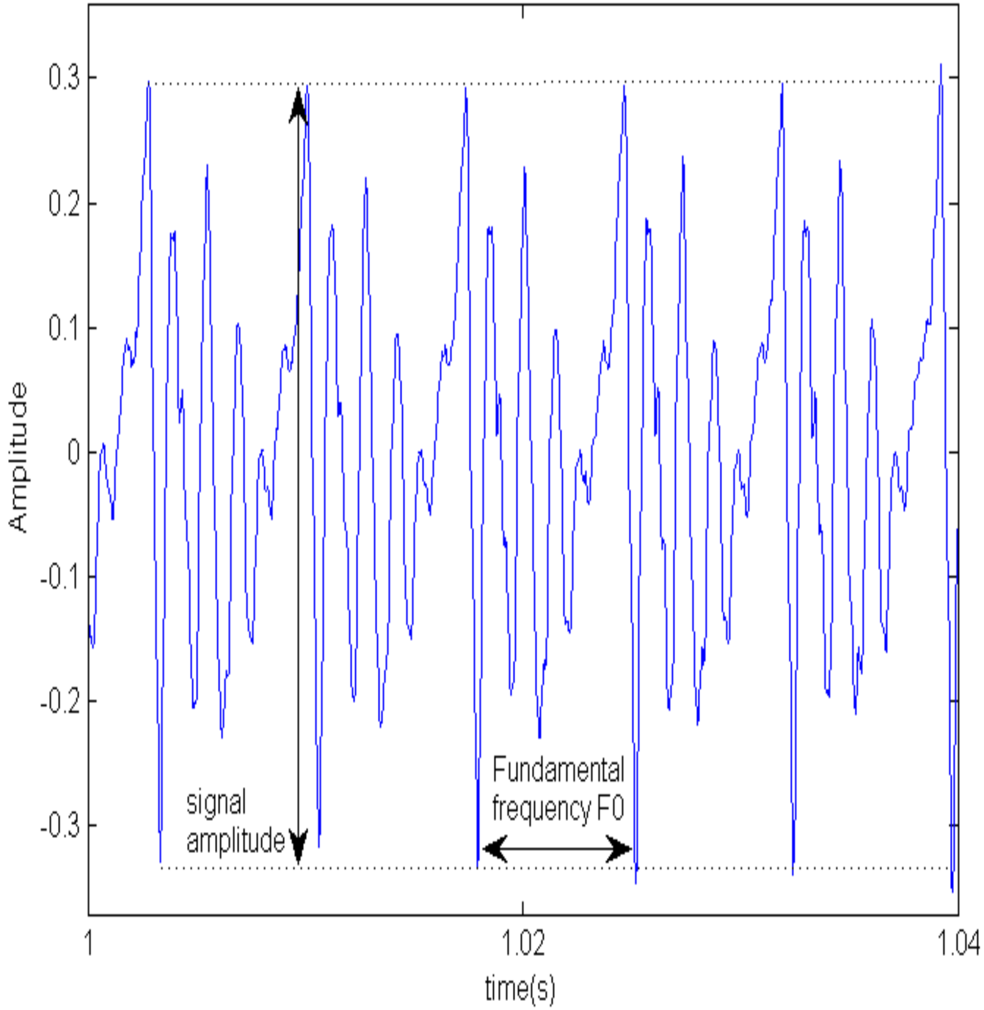


Speech signal processing algorithms

Theme 3

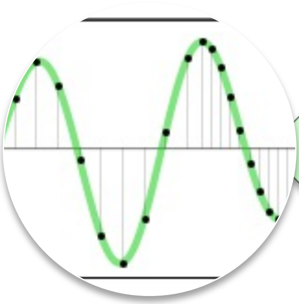


Characterising the time-series



- Can we find some patterns to describe the time-series?
- **Pattern recognition**
- As much of an art as it is science

Feature extraction

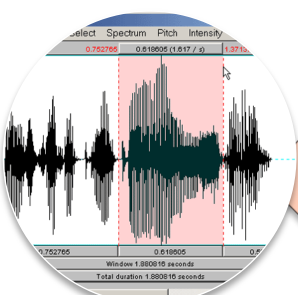


Perturbation algorithms

- Amplitude changes
- Frequency changes

Repeatability (entropy)

- Pattern consistency
- Variability



Energy

- Duration
- Signal-to-noise ratio concepts

Linear predictive coding models

- Describe a subsequent sample as a function of the previous sample:

$\hat{x}_n = a \cdot x_{n-1}$ [Strictly speaking: $\hat{x}_n = a \cdot x_{n-1} + e_n$]

- Let us generalize this idea:

$$\hat{x}_n = a_1 \cdot x_{n-1} + a_2 \cdot x_{n-2} + \cdots a_l \cdot x_{n-l} = \sum_{l=1}^{L} a_l \cdot x_{n-l}$$

- Linear Predictive Coding Coefficients (LPCCs): $\{a_l\}_{l=1}^{L}$

- We can use the LPCCs to characterise the time-series!

Perturbation algorithms

Jitter

- Quantify frequency changes

$$Jitter_{F_{0,\mathrm{abs}}} = \frac{1}{N}\sum_{i=1}^{N-1}\left|F_{0,\mathrm{i}} - F_{0,\mathrm{i+1}}\right| \qquad Jitter_{F_{0,\%}} = 100 \cdot \frac{\frac{1}{N}\sum_{i=1}^{N-1}\left|F_{0,\mathrm{i}} - F_{0,\mathrm{i+1}}\right|}{\frac{1}{N}\sum_{i=1}^{N} F_{0,\mathrm{i}}}$$

Shimmer

- Quantify amplitude changes

Energy operators

Classical squared energy operator (SEO)

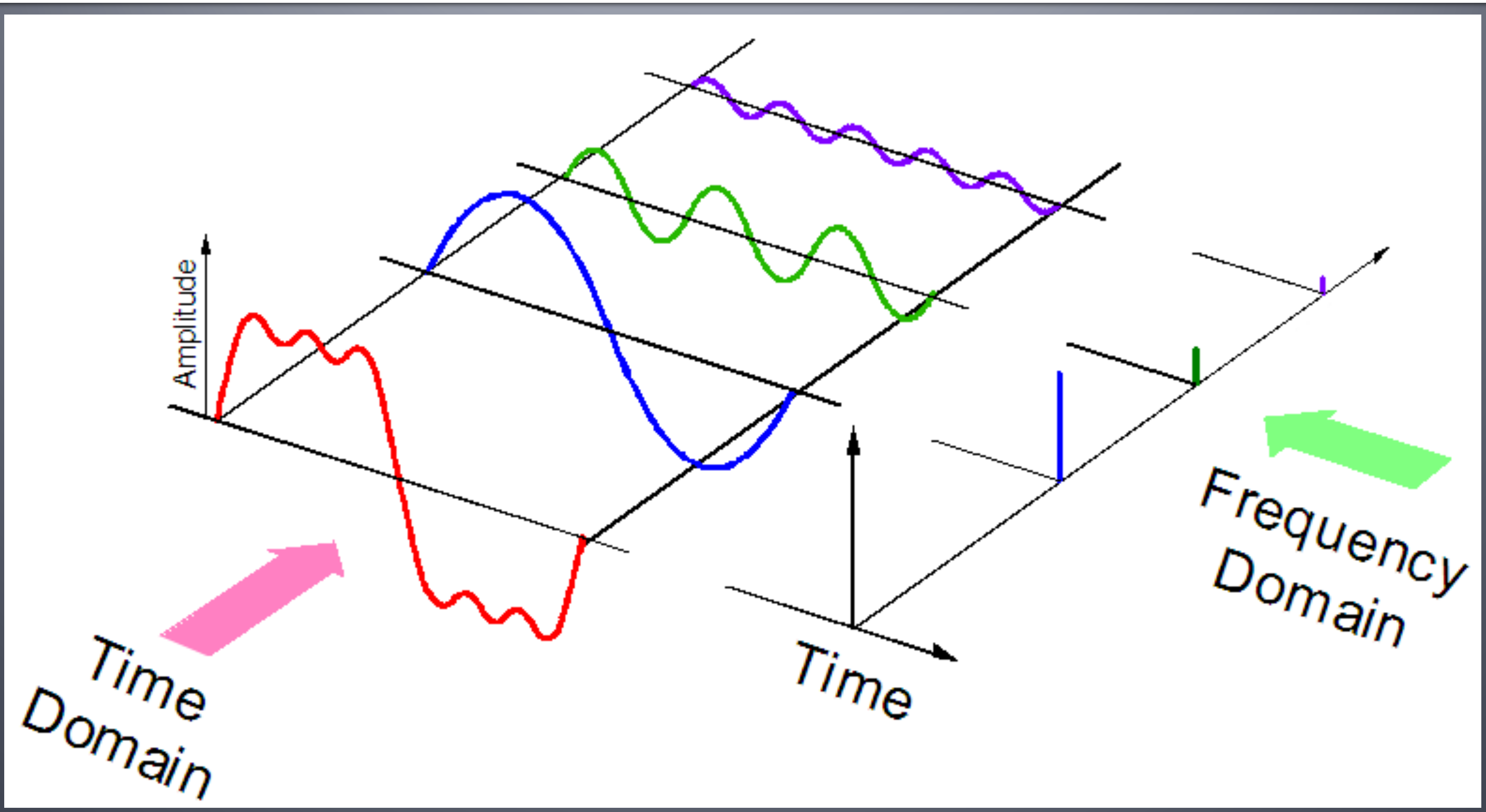
- Standard energy definition
- $SEO(x_n) = x_n^2$

Teager-Kaiser energy operator (TKEO)

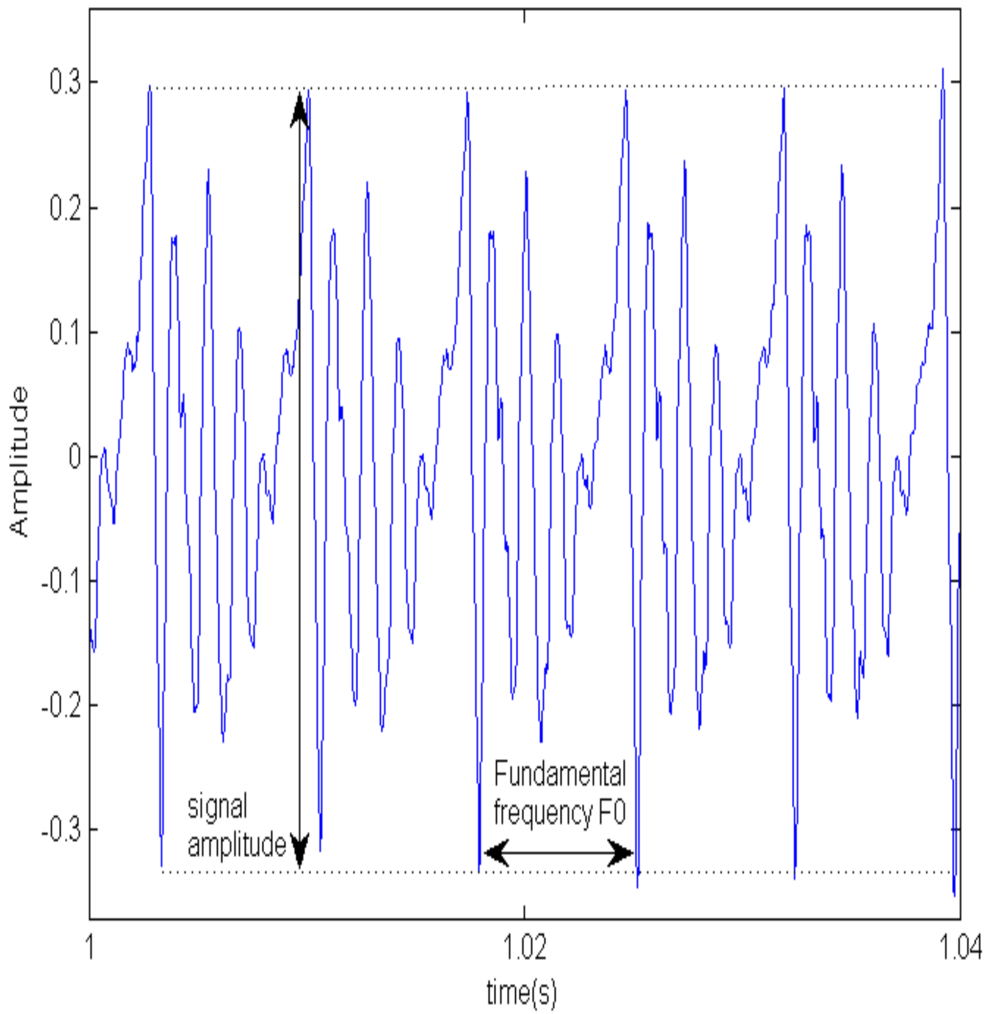
- Amplitude + frequency use in computing energy
- $TKEO(x_n) = x_n^2 - x_{n+1} \cdot x_{n-1}$

Changing domains (transformation)

Yet another approach towards extracting features…



Concept

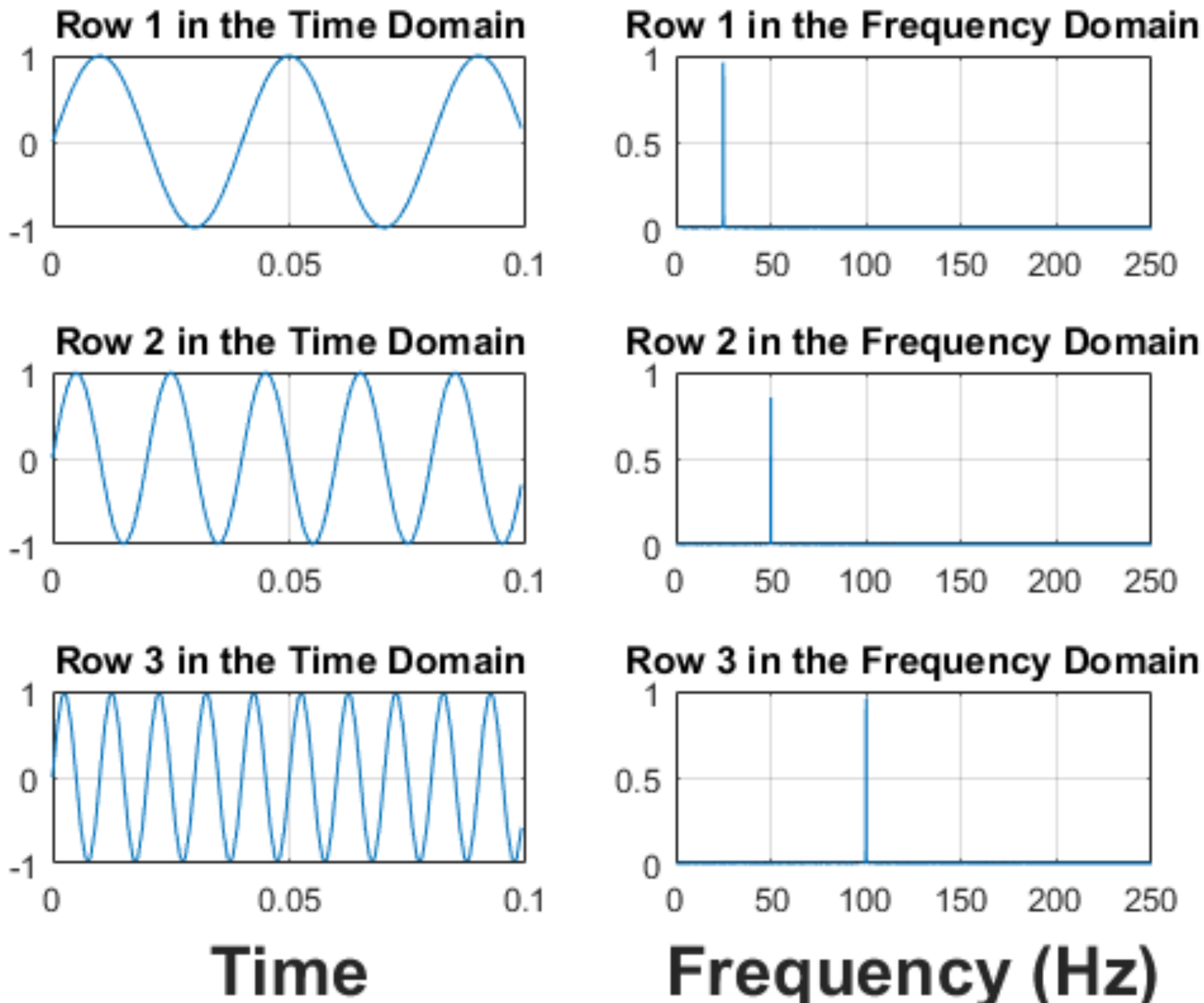


- Can we somehow describe the time series referring to different properties than using the time domain?

Fourier transform

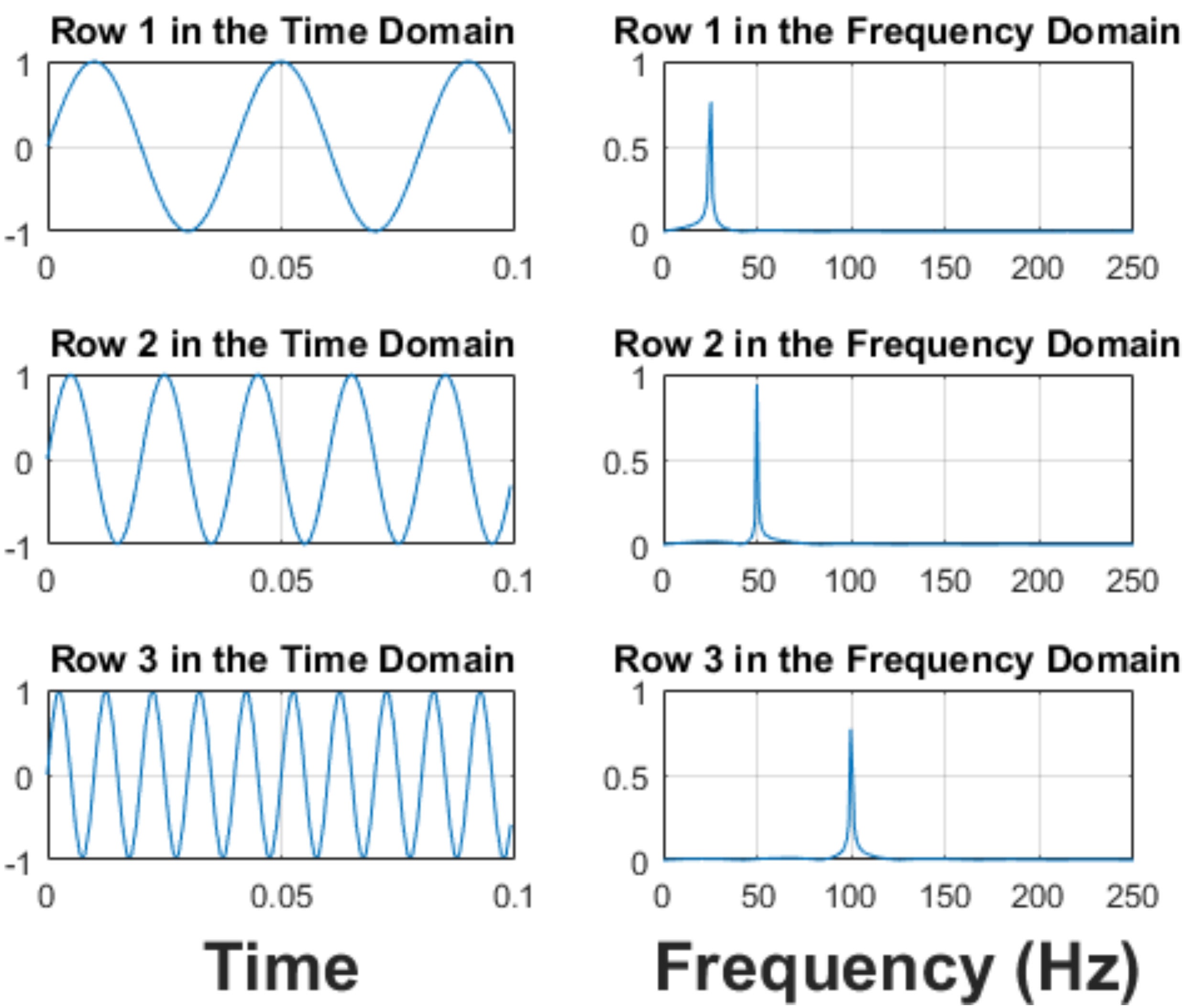
- **Express a signal as an infinite sum of sinusoids**
- The information is in the power of the coefficients
- Computed as: $X\left(e^{j\omega}\right) = \sum_{n=-\infty}^{\infty} x_n \cdot e^{-j\omega n}$
- $\omega = 2\pi f$ is the normalized digital frequency
- <u>Requires uniformly sampled data</u>

Example of Fourier transform (1)



- Sampling frequency = 1/1000
- 100000 samples
- ...all perfect, as expected!

Example of Fourier transform (2)



- Sampling frequency = 1/1000
- 1000 samples
- …not so perfect anymore!

Shortcomings of the Fourier transform

Presumes linear signal

- In practice, no infinite time-series
- Break from periodicity

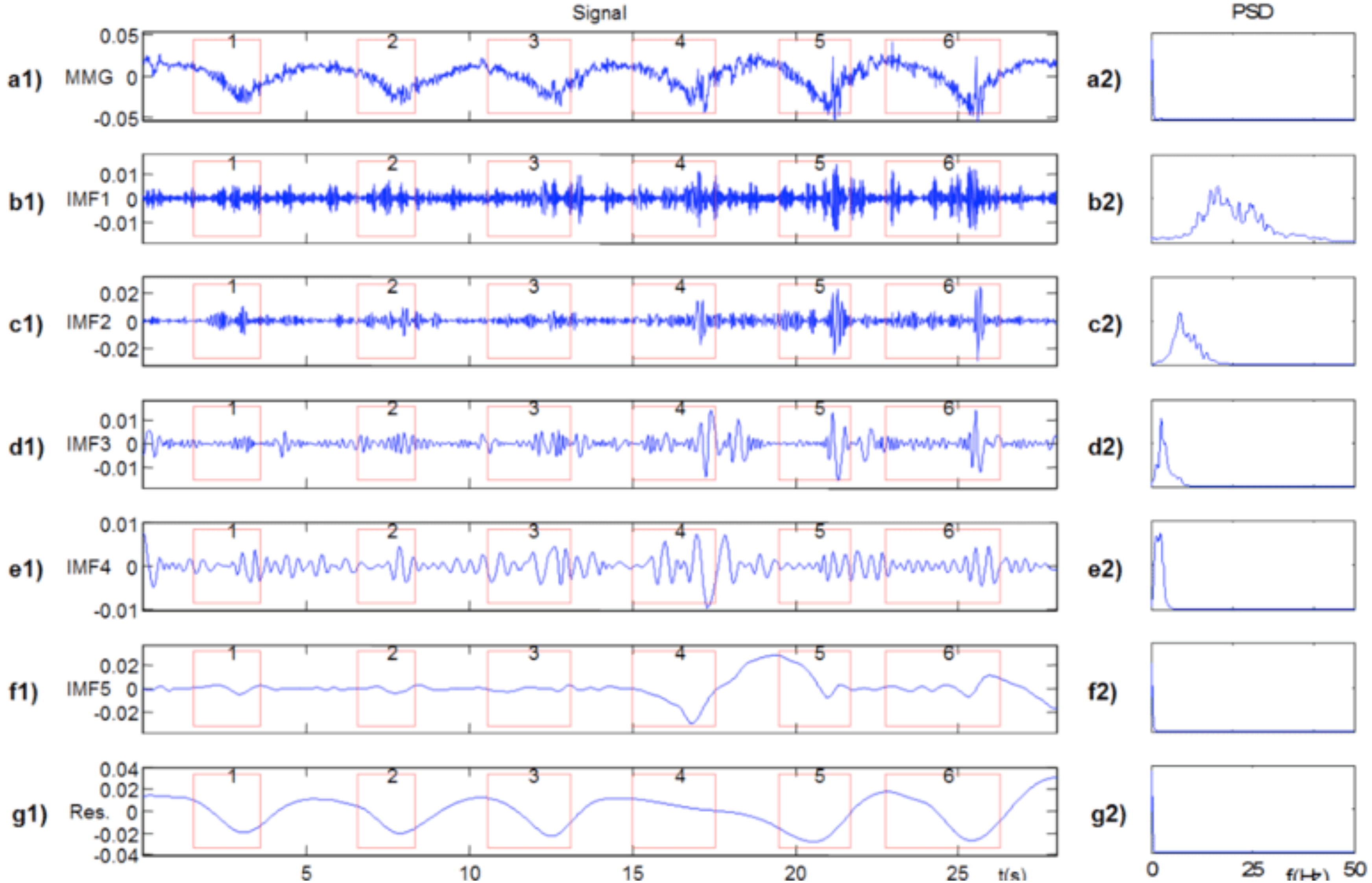
Lack of time and frequency representation

- Frequency patterns often change over time
- Fails to express frequency changes

EMD-ER

- Hilbert-Huang transform
  - decomposes the signal in its components (intrinsic mode functions)
  - high-frequency component in the signal = noise in the signal
  - lower frequencies components = signal
- Experiment to find appropriate ratios of components that denote **signal to noise ratio**

Intrinsic mode functions



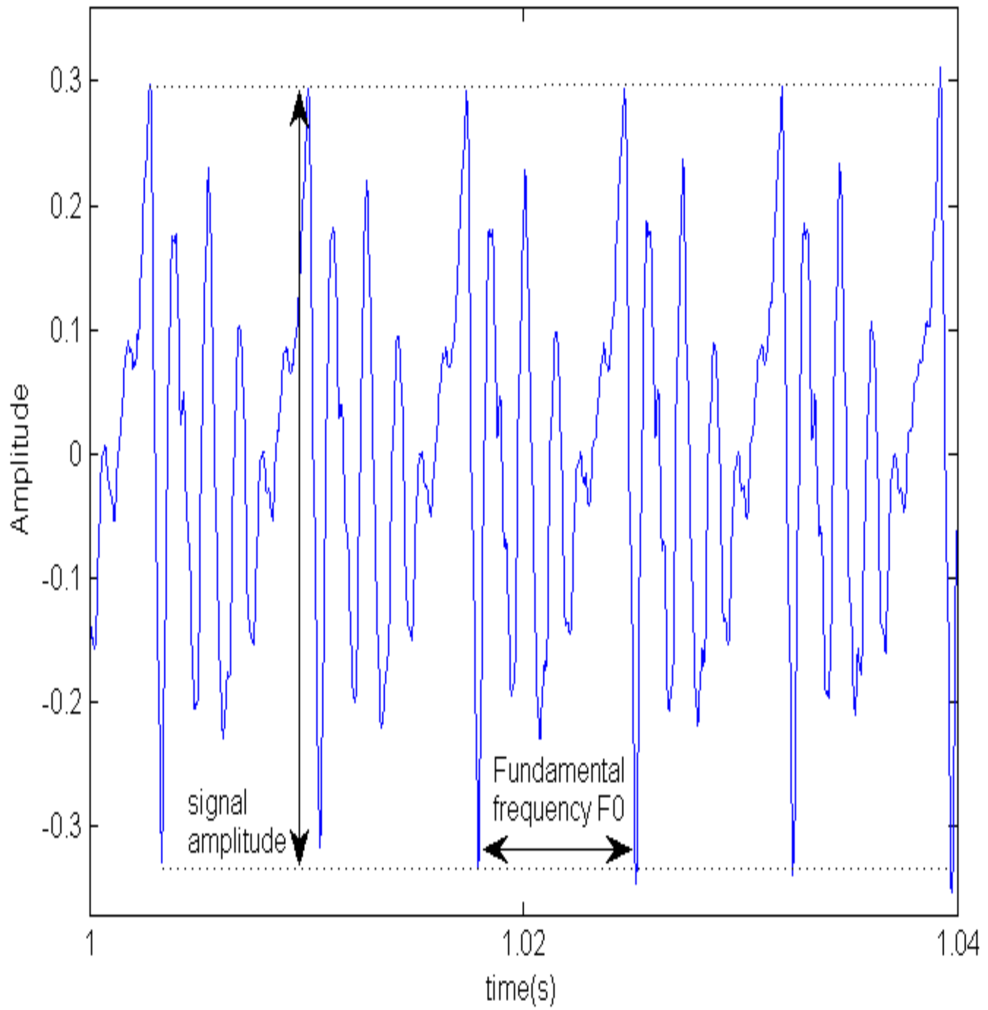
Vocal fold excitation ratio (VFER)

- Detect glottal pulses (e.g. using DYPSA)
- Scan the frequency range using 500 Hz bands
- Optimize the frequency range of ‘signal’ and ‘noise’ by experimentation
- Compute ratios, energy, and entropy values of the **signal to noise ratio** signals

F_0 estimation

Theme 4

F_0 estimation



- More than 100+ F_0 algorithms
- All with shortcomings…

Information fusion: F_0 estimation

Multiple experts/sensors/sources

No ground truth

Combine experts/sensors/sources

Best estimate ground truth

Information fusion framework

Samples	source 1	source 2	…	source M
S_1	z_{11}	z_{12}		z_{1M}
S_2	z_{21}	z_{22}		z_{2M}
S_3	z_{31}	z_{32}		z_{3M}
…	…	…		…
S_N	z_{N1}	z_{N2}		z_{NM}

y
y_1
y_2
y_3
…
y_N

Estimate

x
x_1
x_2
x_3
…
x_N

Truth

N (measurements)

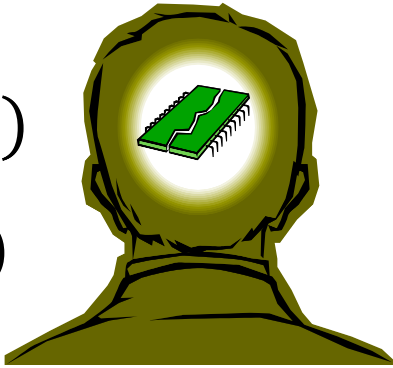
M (sensors/sources)

$$f(z_{i1}, z_{i2}, \dots z_{iM}) = y_i, : \min\{\mathcal{L}(y_i, x_i)\}$$

Kalman filter (introduction)

$$\mathbf{x}_k = \mathbf{A}_k \cdot \mathbf{x}_{k-1} + \mathbf{B}_k \cdot \mathbf{u}_k + \mathbf{w}_{k-1}, \quad \mathbf{w}_k = \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$$

$$\mathbf{z}_k = \mathbf{C}_k \cdot \mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{v}_k = \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$$



- $\mathbf{x}_k$ is the true state & $\mathbf{z}_\mathrm{k}$ is the measurement
- $\mathbf{A}_k$, $\mathbf{B}_k$, $\mathbf{C}_k$, are transition matrices, $\mathbf{u}_\mathrm{k}$ is additional input
- $\mathbf{w}_k$, $\mathbf{v}_k$, represent noise
- **How to set the covariance matrices?**

Adaptive Kalman filter (Li et al., 2008)

- Adaptive KF for each source *independently*
  - Set Q = constant, application dependent: set *a priori* to 0.1 in Li et al. (2008); 5 in Nemati et al. (2010)
  - Adapt R_k based on **signal quality indices** (algorithmic confidence)
- Combine estimates using an empirical equation *external* to KF
- Application: HR detection (Li et al., 2008), and PPG (Nemati et al., 2010)

Adaptive Kalman filter (Li et al., 2008)

- Problems with that approach
  - State noise is almost certainly not constant in practice
  - No interactions directly in the KF design

Adaptive Kalman filter (Tsanas et al., 2014)

- Apply adaptive KF to all measurements *collectively* at instant *i*
  - Adapt $\{Q_k, \mathbf{R}_k\} = f(\text{confidence in algorithm/sensor})$
- No need for additional step *external* to the KF to combine sensors

- Work with matrices instead of vectors

Updating KF covariances

$$\mathrm{Q}_k = \begin{cases} 1\,, & \text{if } \left(\mathrm{SQI}_{k,\mathrm{NDF}}\right) < 0.8 \text{ and } \left(\mathrm{SQI}_{k,\mathrm{SWIPE}}\right) < 0.8 \\ 3 + \left| \dfrac{1}{L} \displaystyle\sum_{j=1\ldots L:\, \mathrm{SQI}_{k,j} > 0.8} \left[\left(\mathrm{z}_{k,j} - \tilde{\mathrm{x}}_k \right) \cdot \mathrm{SQI}_{k,j} \right] \right|, & \text{otherwise. } (L = \left\| \mathrm{SQI}_{k,j} > 0.8 \right\|, L \le M) \end{cases}$$

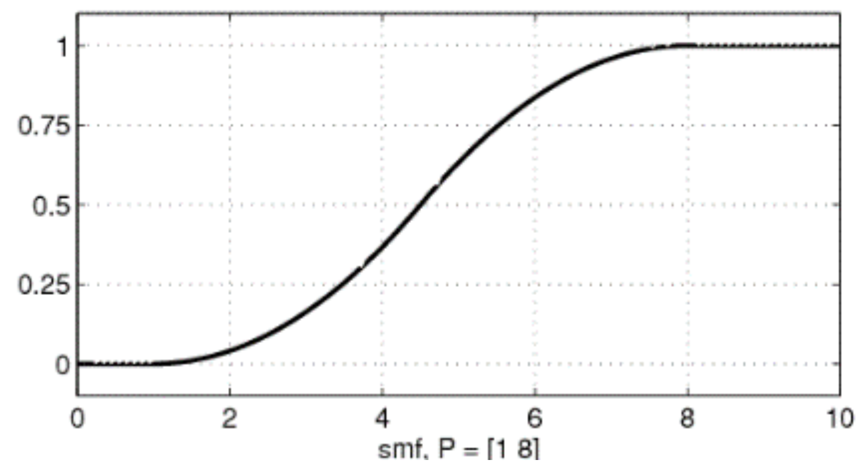
$$\mathbf{R}_k = \mathbf{R}_{k_0} \odot \exp\left(1/\mathbf{SQI}_k{}^2 - 1\right)$$

$$\mathbf{x}_k = \mathbf{A}_k \cdot \mathbf{x}_{k-1} + \mathbf{B}_k \cdot \mathbf{u}_k + \mathbf{w}_{k-1}, \quad \mathbf{w}_k = \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$$

$$\mathbf{z}_k = \mathbf{C}_k \cdot \mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{v}_k = \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$$

$$\mathbf{SQI}_{\boldsymbol{k}} = \mathbf{1} + \mathbf{b}_k - \mathbf{p1}_k - \mathbf{p2}_k - \mathbf{p3}_k - \mathbf{p4}_k$$

- S-membership curve



Bonuses and penalties

$\mathbf{p1}_k = 0.25 \cdot S_S(|\mathbf{z}_k - \mathbf{z}_{k-1}|, 0, 100)$ absolute differences in successive F0 estimates

$\mathbf{p2}_k = 0.25 \cdot S_S(|\mathbf{z}_k - \mathbf{z}_{robust}|, 0, 100)$, difference from robust estimate

$\mathbf{p3}_k = 0.75 \cdot S_S(|\mathbf{z}_k - \tilde{\mathbf{x}}_k|, 0, 100)$, difference with a priori estimate

$\mathbf{p4}_k = 0.75 \cdot S_S(|\mathbf{z}_k - \mathbf{1} \cdot \check{\mathrm{x}}_{kbest}|, 0, 50)$, difference from best expert

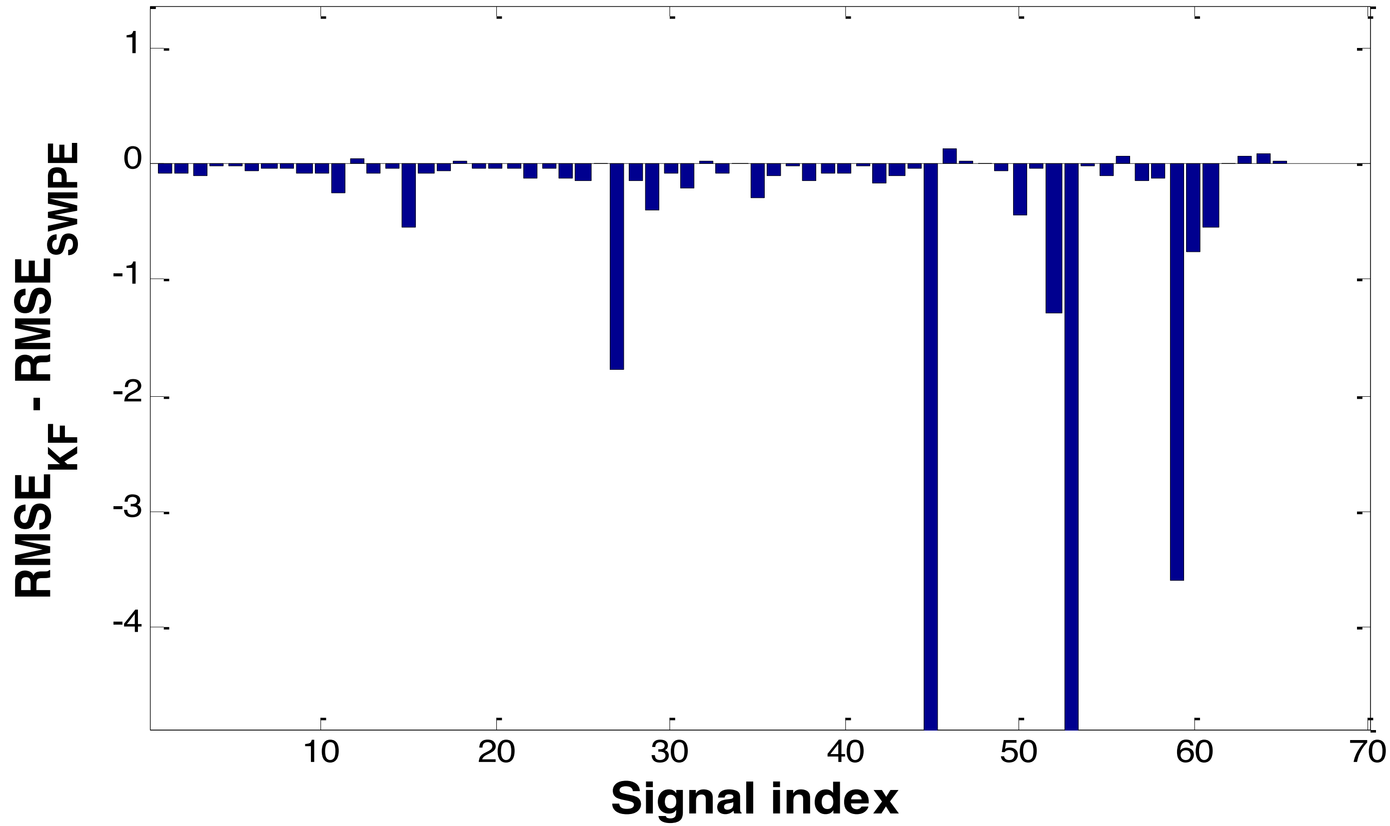
$\mathbf{b}_k = \mathbf{1} \cdot \left(1 - \mathrm{p4}_{k,(best)}\right) - \left[1 - S_S(|\mathbf{z}_k - \mathbf{1} \cdot \check{\mathrm{x}}_{kbest}|, 0, 100) \odot \left(\mathbf{p}_{k,1} + \mathbf{p}_{k,2}\right)\right]$,
bonus for being close to the best expert if heavily penalised

if $\mathrm{p4}_{k,(best)} < 0.2$ we give additional bonus to the best F_0 estimation algorithm: $\mathrm{b}_{k,(best)} = 3$

Error metrics for F_0 algorithms

Algorithm	MAE (Hz)	MRE (%)	RMSE (Hz)
DYPSA	14.42 ± 26.32	5.54 ± 8.44	25.86 ± 32.89
PRAAT1	29.22 ± 57.23	13.28 ± 24.08	31.67 ± 57.10
PRAAT2	29.05 ± 56.86	13.21 ± 24.00	31.47 ± 56.71
RAPT	28.30 ± 63.47	8.63 ± 17.98	34.21 ± 65.89
SHRP	18.78 ± 47.77	6.85 ± 16.86	26.91 ± 55.21
SWIPE	**3.06 ± 7.01**	**1.18 ± 2.48**	**6.22 ± 13.46**
YIN	16.36 ± 47.34	6.16 ± 16.32	23.35 ± 51.77
NDF	15.12 ± 60.66	4.16 ± 15.24	17.66 ± 60.87
TEMPO	50.67 ± 99.23	17.69 ± 31.08	53.21 ± 100.92
XSX	33.43 ± 52.11	16.85 ± 25.90	39.57 ± 56.81
OLS	4.08 ± 7.76	1.55 ± 2.62	7.58 ± 13.82
IRLS	3.17 ± 7.03	1.23 ± 2.49	6.53 ± 13.57
KF	**2.49 ± 5.04**	**0.97 ± 1.82**	**4.95 ± 9.19**

Information fusion results

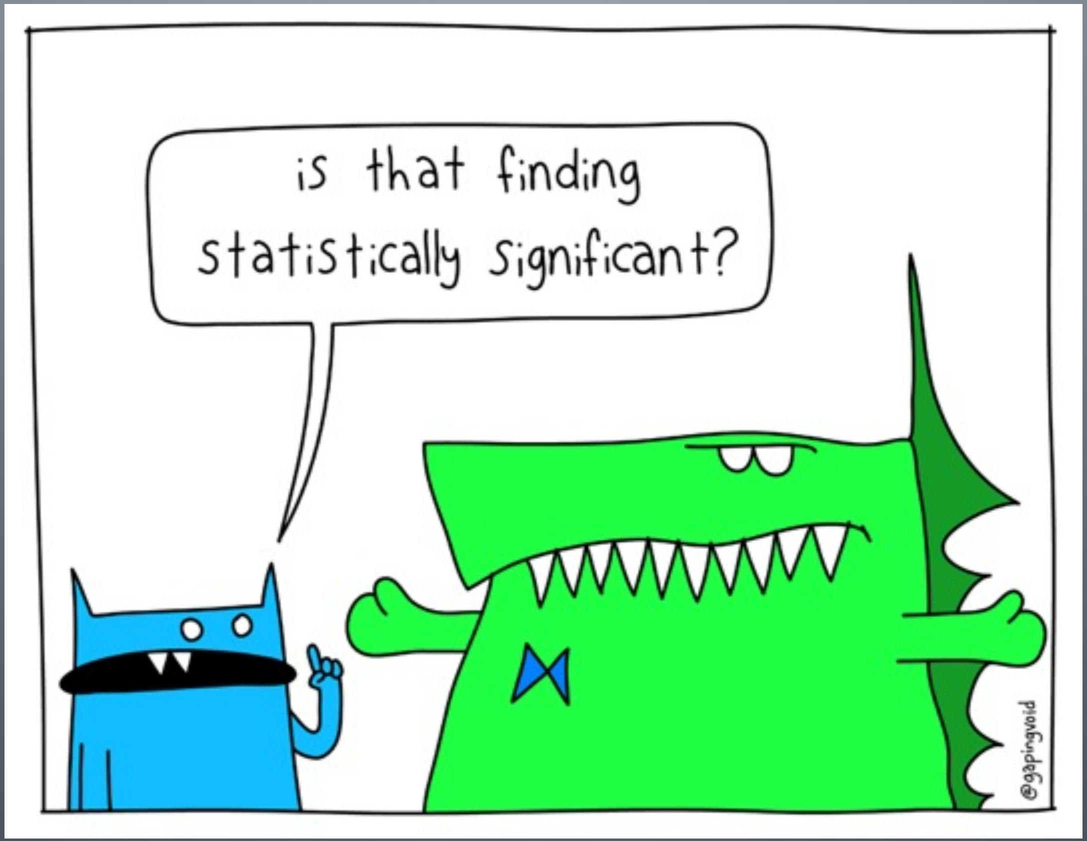


- F0 fusion better than the single best F0 expert

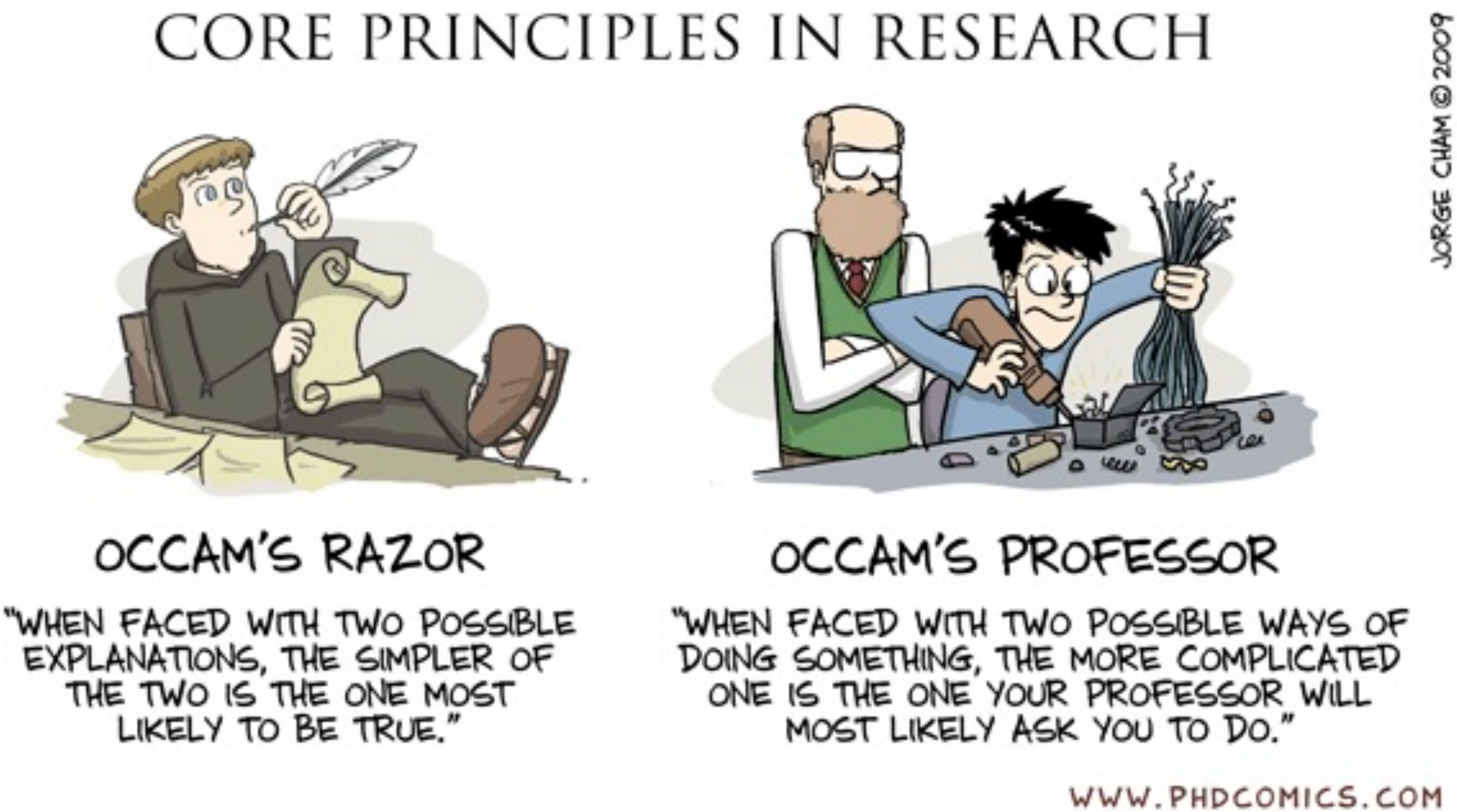
A. Tsanas, M. Zañartu, M.A. Little, C. Fox, L.O. Ramig, G.D. Clifford: *Robust fundamental frequency estimation in sustained vowels: detailed algorithmic comparisons and information fusion with adaptive Kalman filtering*, **Journal of the Acoustical Society of America**, Vol. 135, pp. 2885-2901, 2014

Statistical machine learning

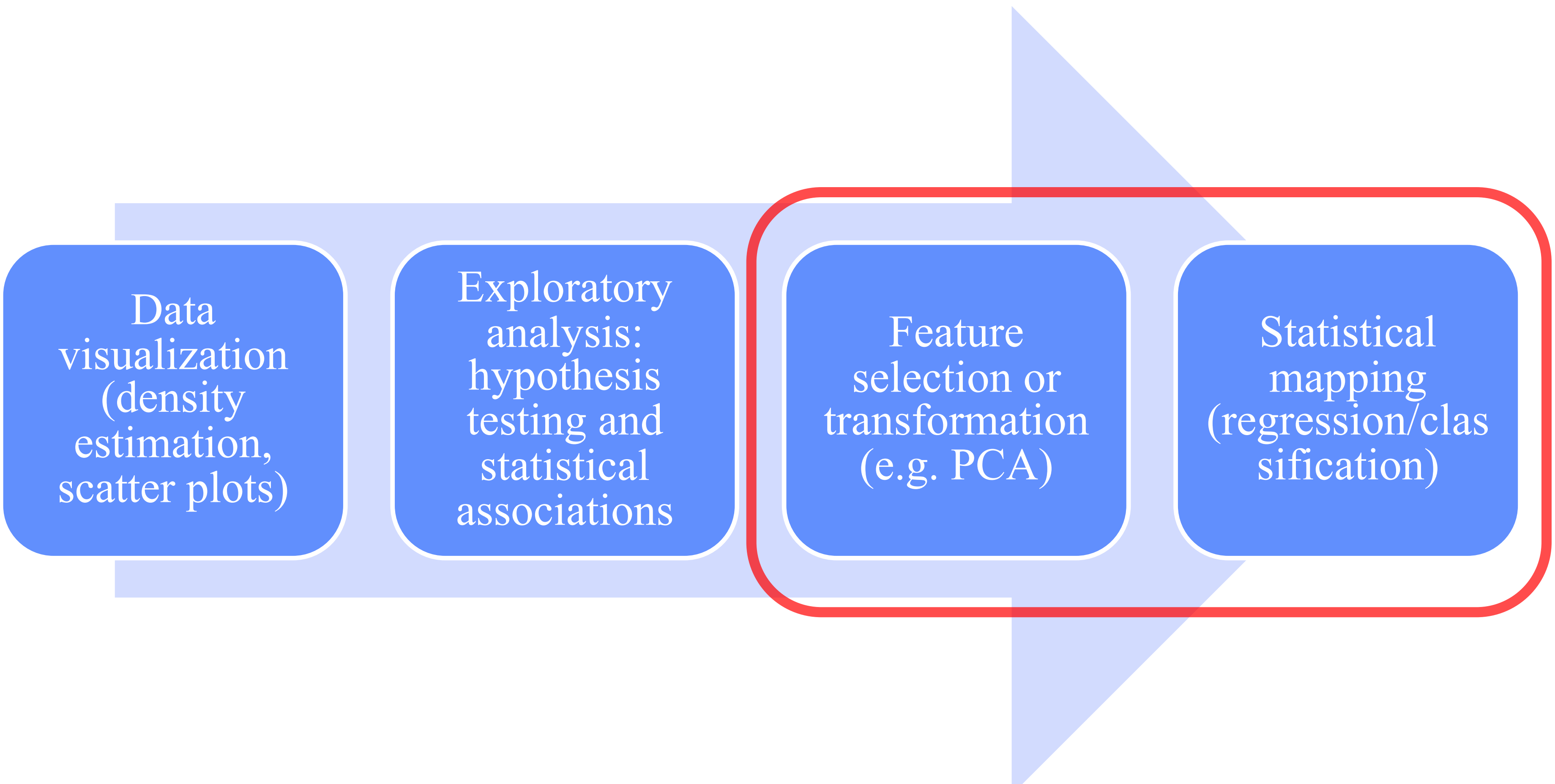
Theme 5



Principle of parsimony



Summary of key analysis processes



Overview of data analysis

Feature generation from raw data → Feature selection or transformation → Statistical mapping

X **y**

N

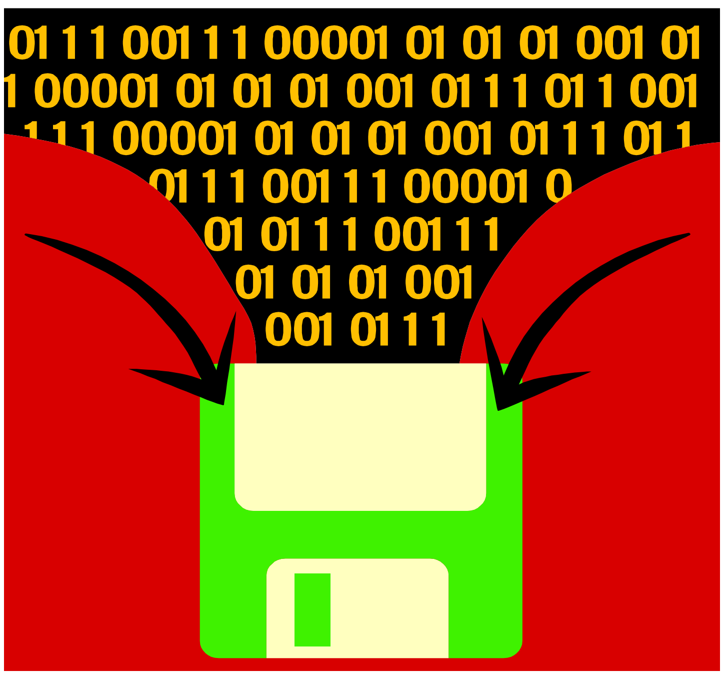
Subjects	feature1	feature2	...	feature M	result
P1	3.1	1.3		0.9	1
P2	3.7	1.0		1.3	2
P3	2.9	2.6		0.6	1
…					…
P_N	1.7	2.0		0.7	3

M (features or characteristics) outcome

- Depending on the problem, "features" can be demographics, genes, …
- $\mathbf{y} = f(\mathbf{X})$, f: mechanism **X**: feature set **y**: outcome

Introduction to the problem

- Many features M ☞ **Curse of dimensionality**
- Obstruct interpretability and detrimental to learning process



Solution to the problem

- Reduce the initial feature space M into m (where $m < M$)

- **Feature selection**
- **Feature transformation**

Feature selection advantages

- **Interpretable**
- Retain domain expertise

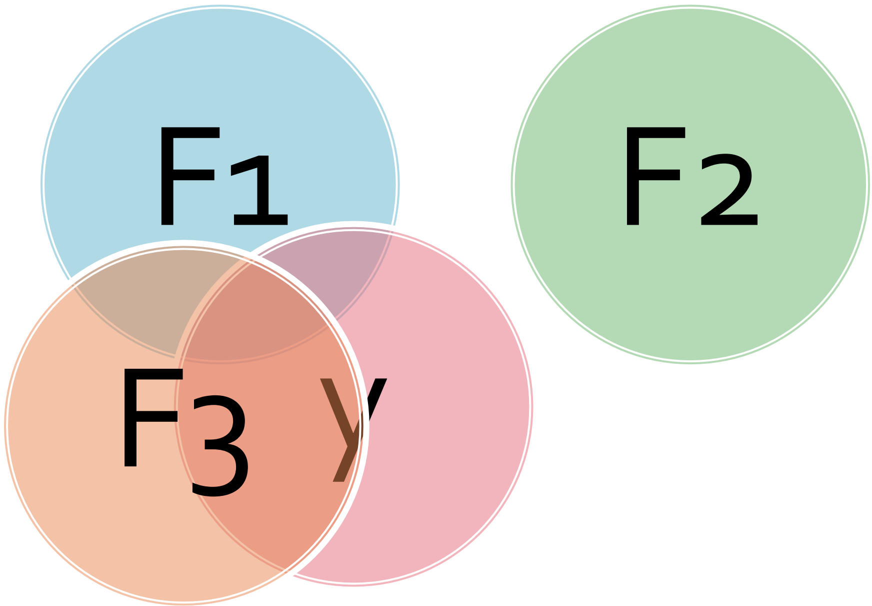
- Often is the only useful approach in practice (e.g. in micro-array data)
- Saves on resources on data collection or data processing

Feature selection (introduction)

Discard non-contributing features towards predicting the outcome

Concept of relevance

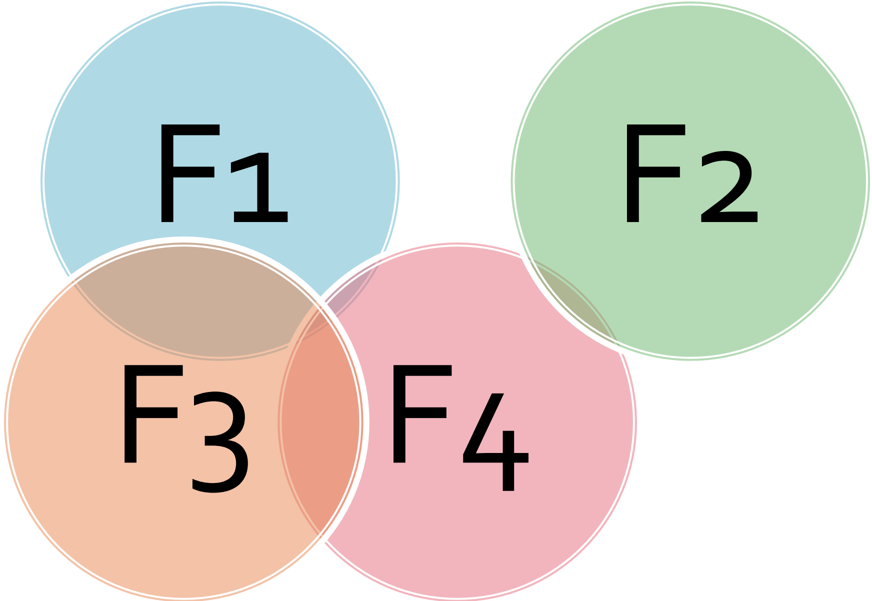
- **Maximum relevance**: features (F) and response (y)



- **Which features would you choose? In which order?**

Concept of redundancy

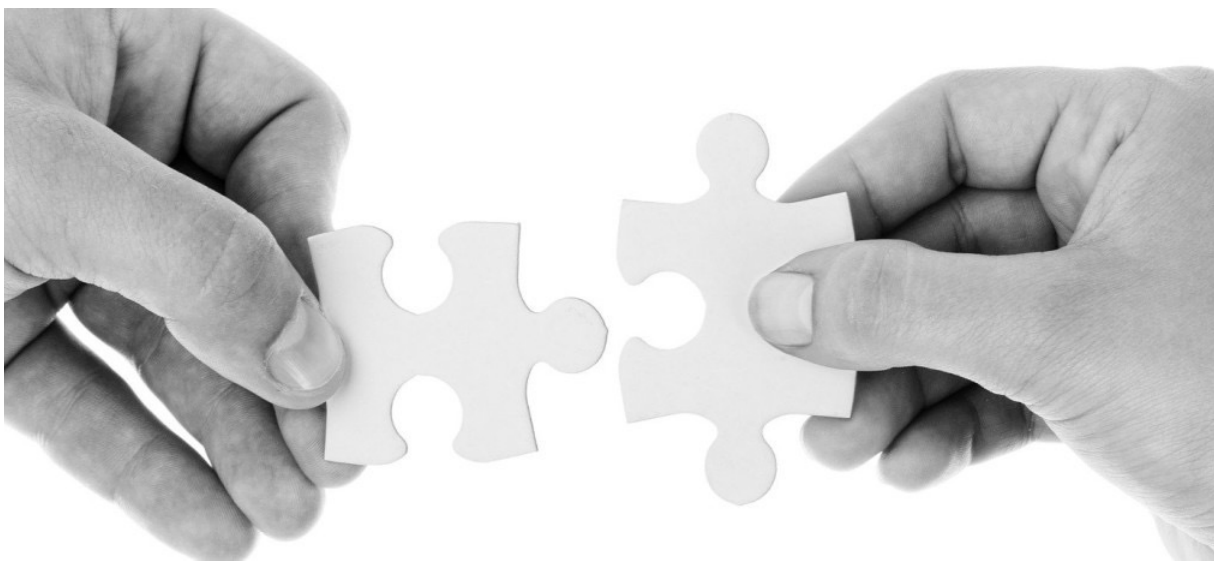
- **Minimum redundancy** amongst features in the subset



- **Which features would you choose? In which order?**

Concept of complementarity

- **Conditional relevance** (feature **interaction)**



A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

- Features are **jointly** highly predictive of outcome

minimum Redundancy Maximum Relevance (mRMR)

- Compromise: relevance and redundancy 
- Does not account for interactions and non-pairwise redundancy 
- **Generally works well**
- $\text{mRMR} \stackrel{\text{def}}{=} \max\limits_{j \in Q-S} \left[I(\mathbf{f}_j, \mathbf{y}) - \frac{1}{|S|} \sum_{s \in S} I(\mathbf{f}_j, \mathbf{f}_s) \right]$

❑ $|S|$ is the cardinality of the selected subset
❑ Q contains the indices of all possible features

Relevance, redundancy and complementarity trade-off (RRCT)

"The best result will come when everyone in the group doing what is best for himself… and the group."

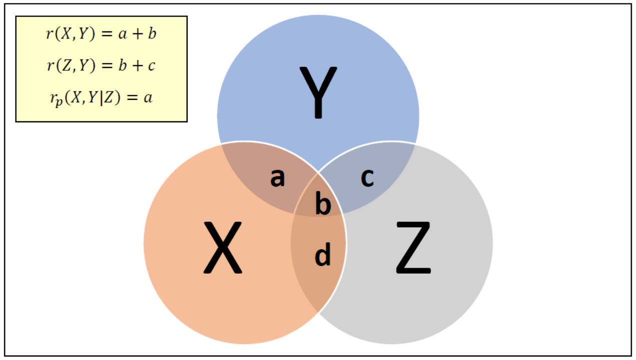
John Nash, Nobel prize winner

- RRCT is an effective algorithmic formulation of Nash's concept

$$\mathrm{RRCT} \stackrel{\text{def}}{=} \max_{j \in Q-S}\left[r_{\mathrm{IT}}(\mathbf{f}_j,\mathbf{y}) - \frac{1}{|S|}\sum_{s \in S} r_{\mathrm{IT}}(\mathbf{f}_j,\mathbf{f}_s) + \mathrm{sign}\left(r_{\mathrm{p}}(\mathbf{f}_j,\mathbf{y}|S)\right)\cdot \mathrm{sign}\left(r_{\mathrm{p}}(\mathbf{f}_j,\mathbf{y}|S) - r(\mathbf{f}_j,\mathbf{y})\right)\cdot r_{\mathrm{p,IT}}\right]$$

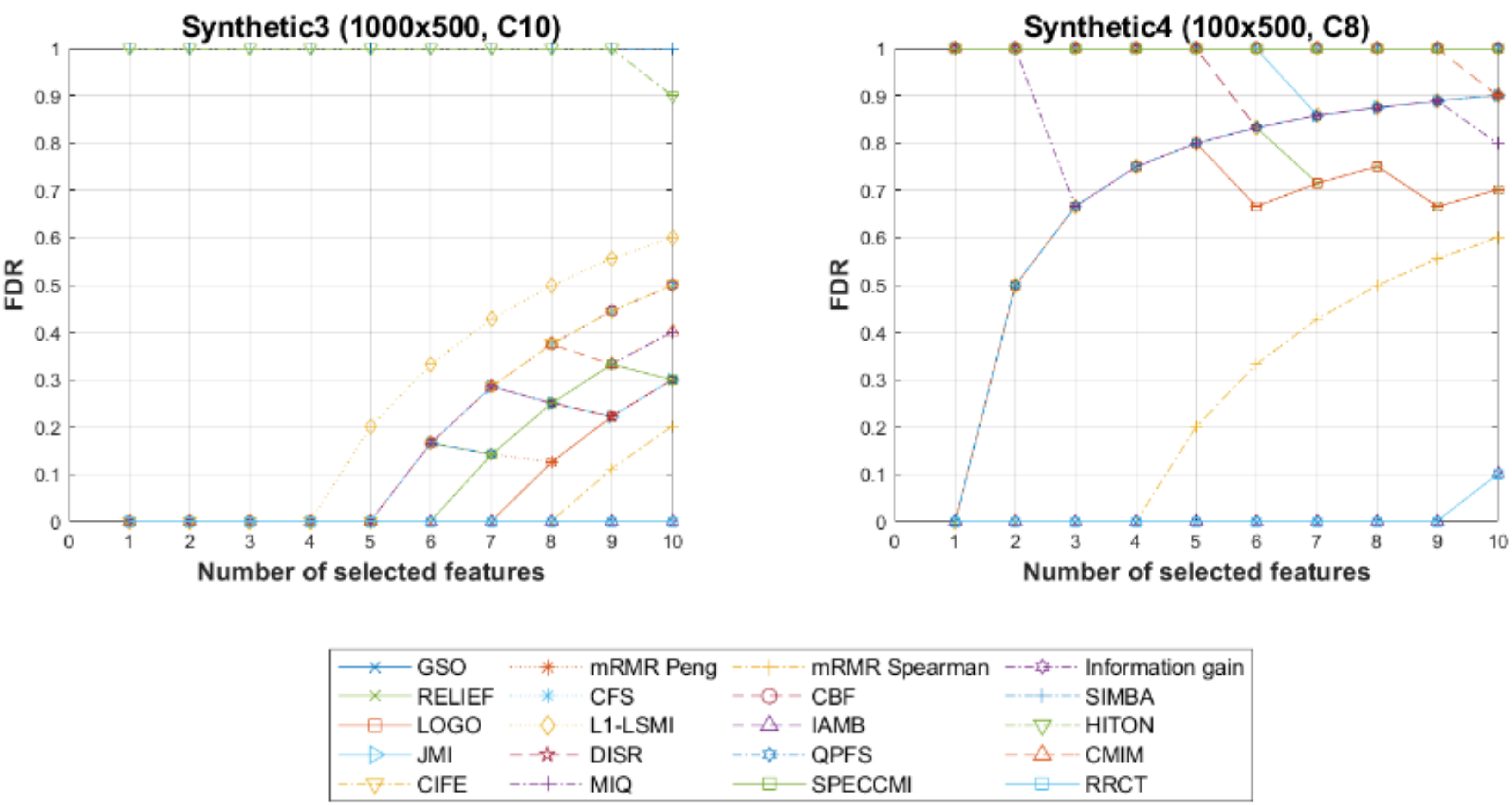
A. Tsanas: *Relevance, redundancy and complementarity trade-off (RRCT): a generic, efficient, robust feature selection tool*, **Patterns**, Vol. 3, pp. 100471, 2022 [doi: https://doi.org/10.1016/j.patter.2022.100471]

RRCT properties



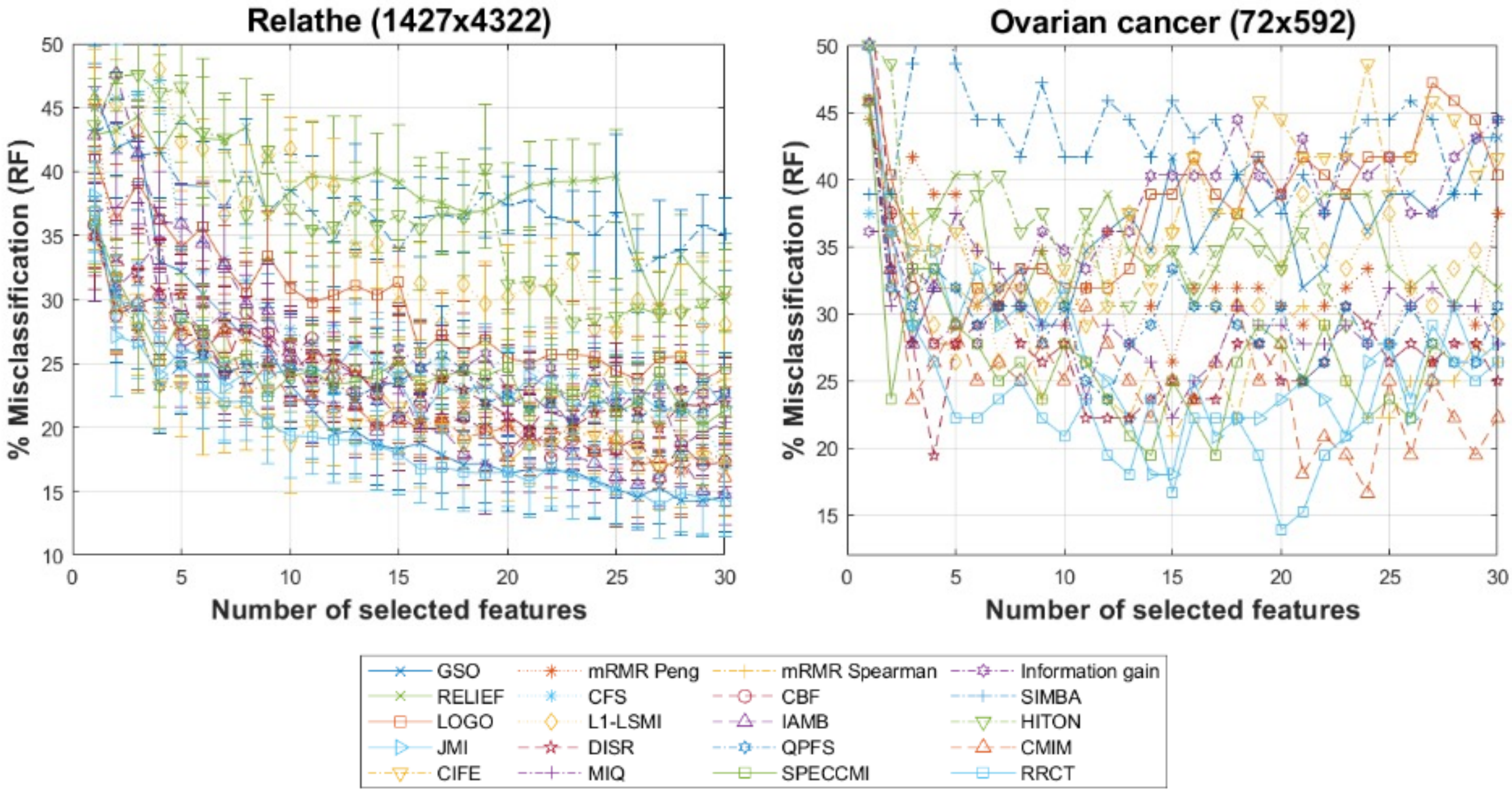
- Information theoretic feature selection algorithm
- Solid theoretical underpinning
- Applicable in regression and classification problems
- Inherently tackles mixed-type variables

Indicative FS comparisons (synthetic datasets)



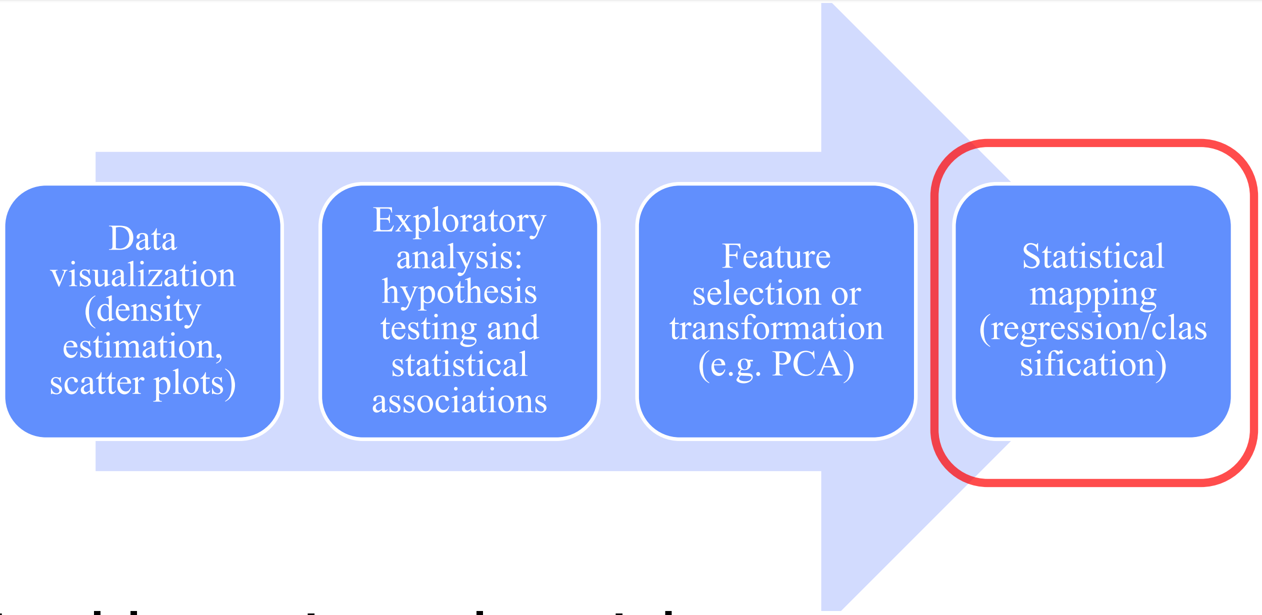
- False Discovery Rate (FDR, **lower score in the y-axis is better**)

Indicative FS comparisons (real-world datasets)



- Misclassification using a RF classifier (**lower score in the y-axis is better**)

Statistical mapping



- Statistical learning algorithms
  - Support Vector Machines
  - Random Forests
  - Gradient Boosted Machines
  - ....

Biomedical speech signal processing applications

Theme 6

Case study I

Speech and neurodegenerative disorders

Problem with current assessment

Physical presence in the clinic

- Cumbersome for those living in remote areas
- People need to take days off work

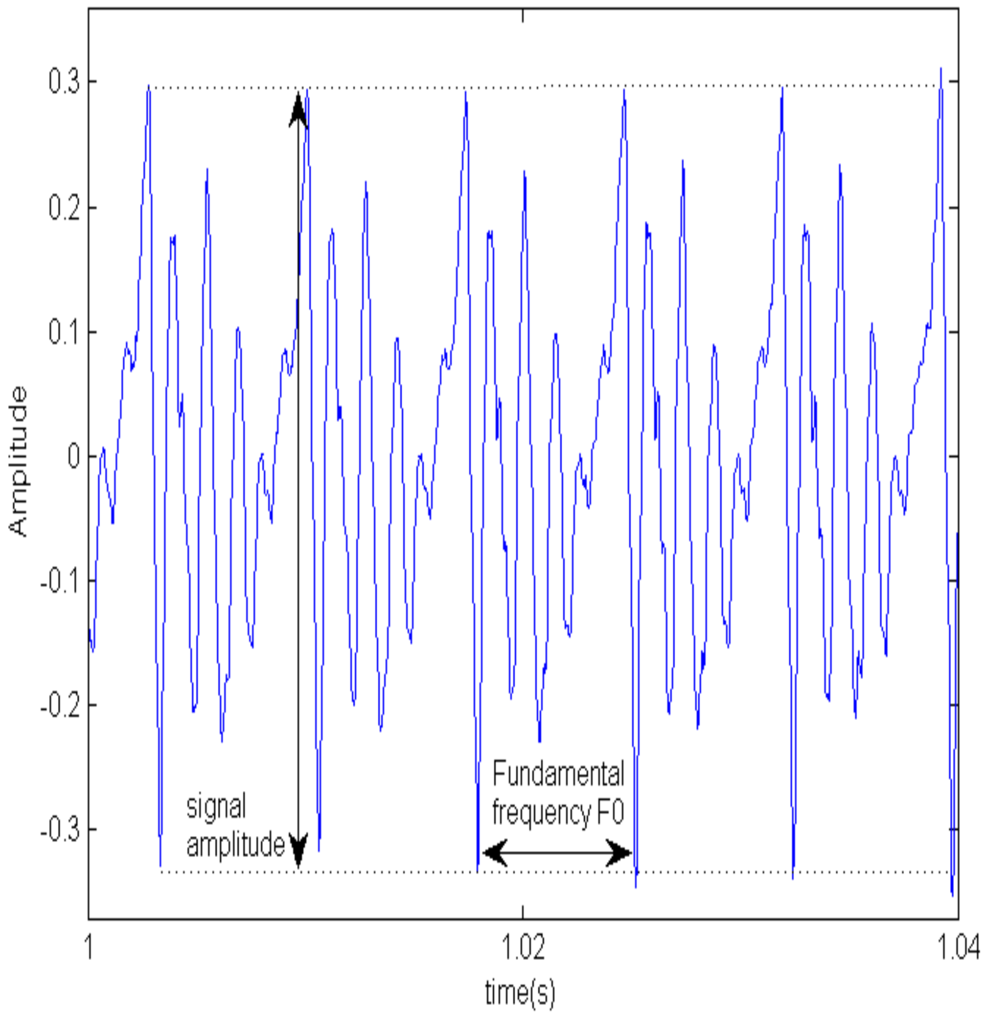
Limited time window

- Time consuming
- Snapshot: does not capture daily variability

No ground truth

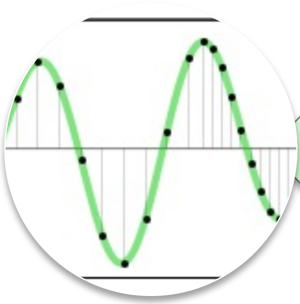
- Subjective, depends on rater’s experience
- Inter-rater variability

Characterizing the time-series



- Can we find some patterns to describe the time-series?
- **Pattern recognition**
- As much of an art as it is science

Feature extraction

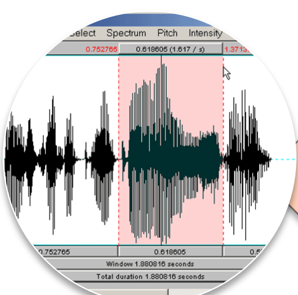


Perturbation algorithms

- Amplitude changes
- Frequency changes

Repeatability (entropy)

- Pattern consistency
- Variability



Energy

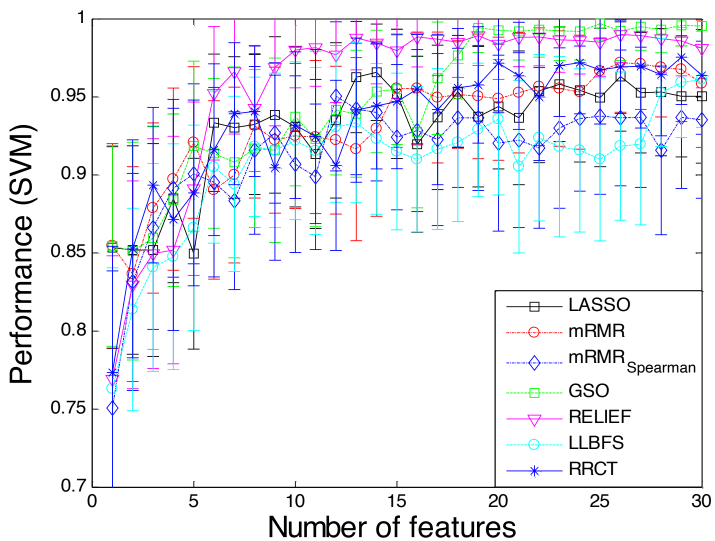
- Duration
- Signal-to-noise ratio concepts

Project 1: Parkinson's vs Control

- **NCVS dataset (PD subjects versus healthy controls)**
  - 43 subjects (33 PD), 263 sustained vowel /a/ phonations
  - Age- and gender- matched healthy controls
  - Controlled acoustic environment

The National Center for Voice & Speech

- **Binary differentiation problem**
  - About 93% accuracy in the literature
  - 98.6 % in my study



A. Tsanas et al.: Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease, **IEEE Transactions on Biomedical Engineering**, Vol. 59, pp. 1264-1271, 2012

Project 2: Parkinson's monitoring

- **AHTD dataset** (PD diagnosis up to 5 years at trial onset)
  - 42 PD subjects (28 males), 5875 sustained vowel phonations
  - Patient follow-up: assessment at baseline, 3- & 6-months
  - 6 phonations weekly

Massachusetts Institute of Technology

  - **Aim: replicate a clinical scale ranging 0 – 176 points**

Quantifying symptom severity

Unified Parkinson's Disease Rating Scale (UPDRS)

comprises three components and 44 sections in total, each section spans the range 0-4

Component 1
Mentation, behavior and mood
4 sections (1-4)

Includes mentation, thought disorder, depression, and motivation/initiative

Component 2
Activities of daily living
13 sections (5-17)

Ability to complete daily tasks unassisted, e.g. dressing, walking, writing

Component 3
Motor (motor-UPDRS)
27 sections (18-44)

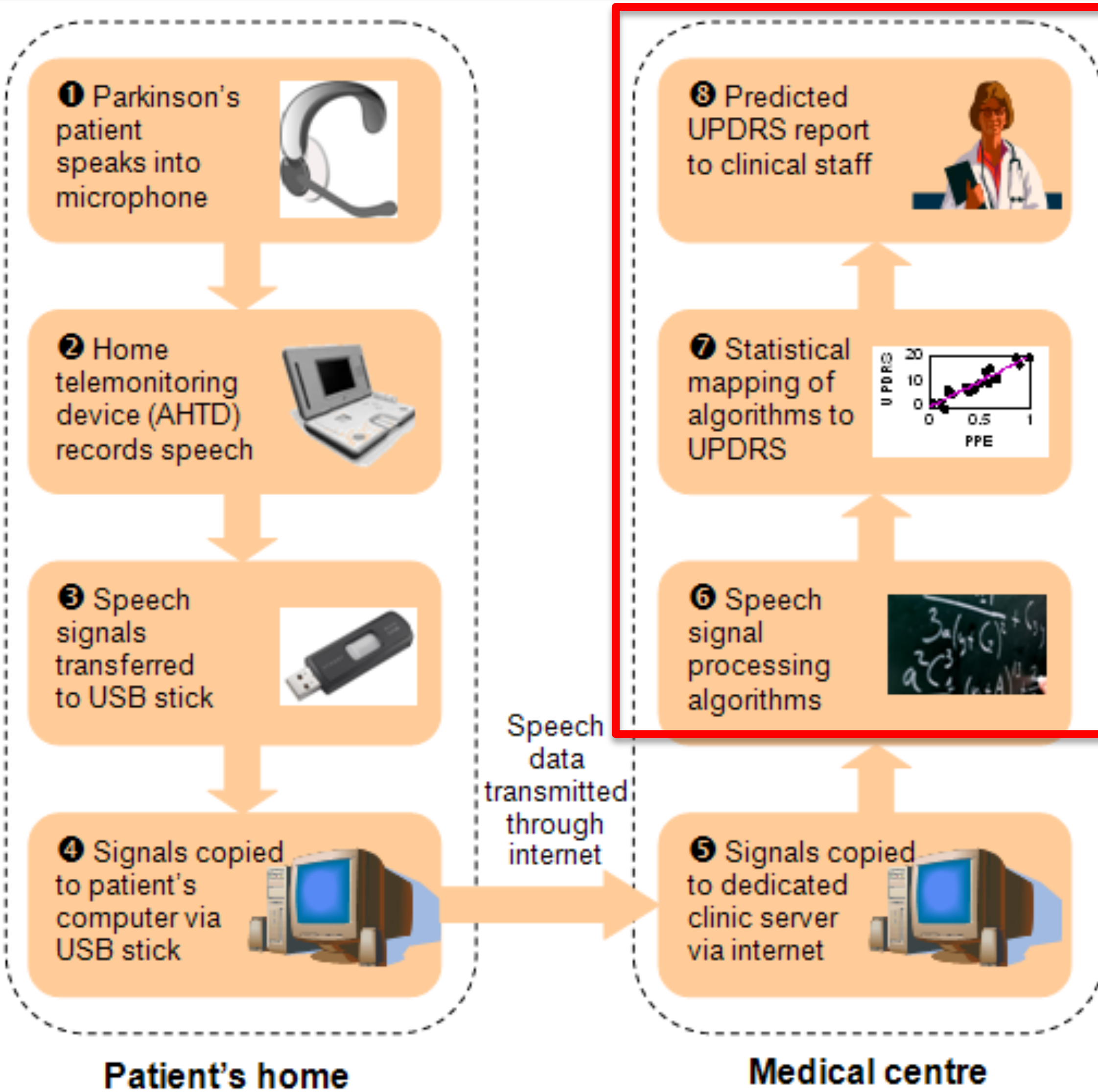
Muscle problems e.g. tremor, rigidity, posture, stability, bradykinesia

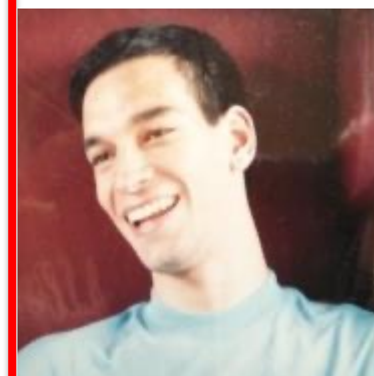
Section 5: Speech – the clinician assesses whether the subject's vocal output is *understandable* during casual discussion.

Section 18: Speech – the clinician assesses whether the subject's vocal output is *expressive* during casual discussion.

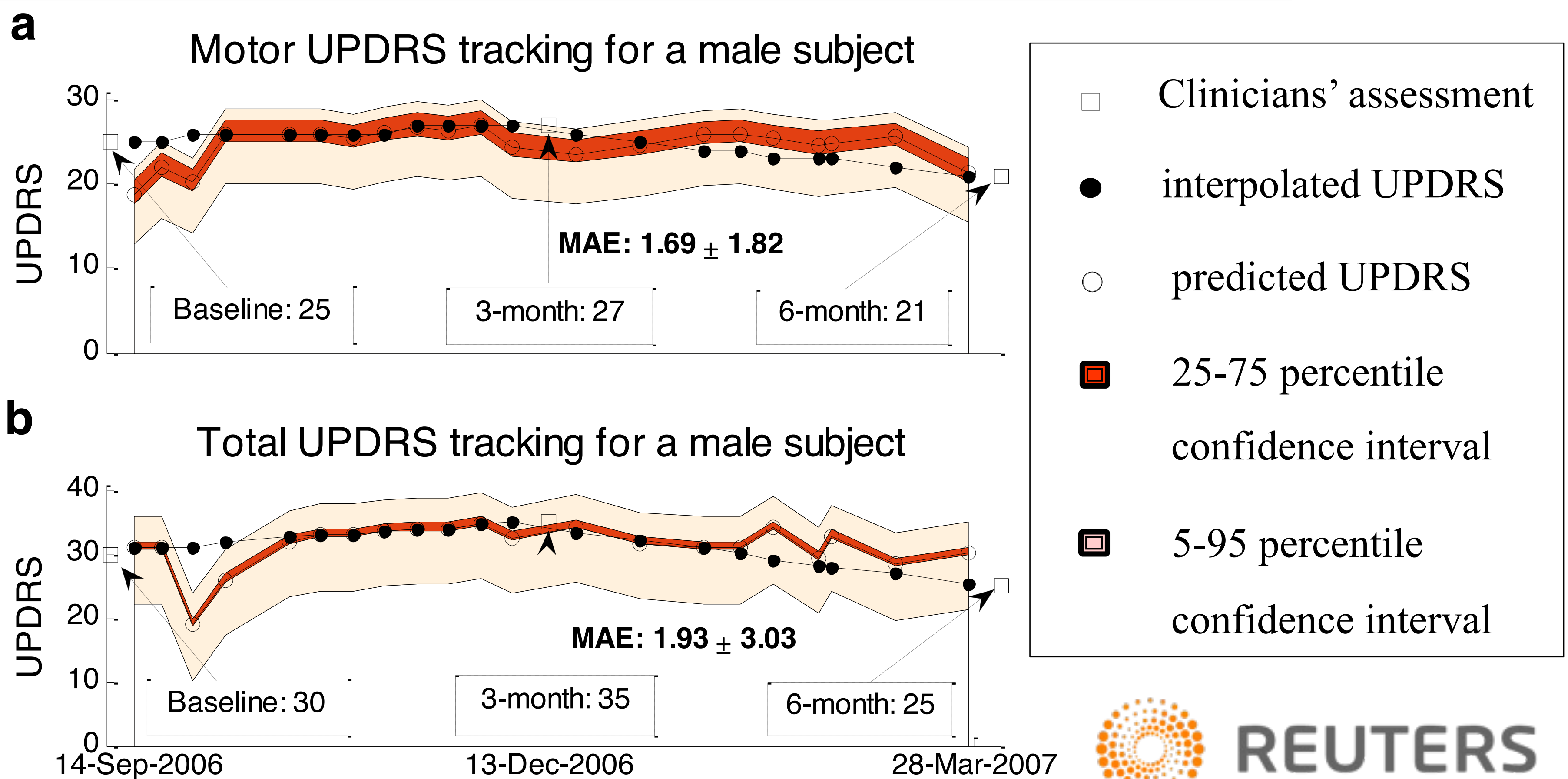
Proposed solution

Telemedicine:
the dawn of a new era





Remote assessment

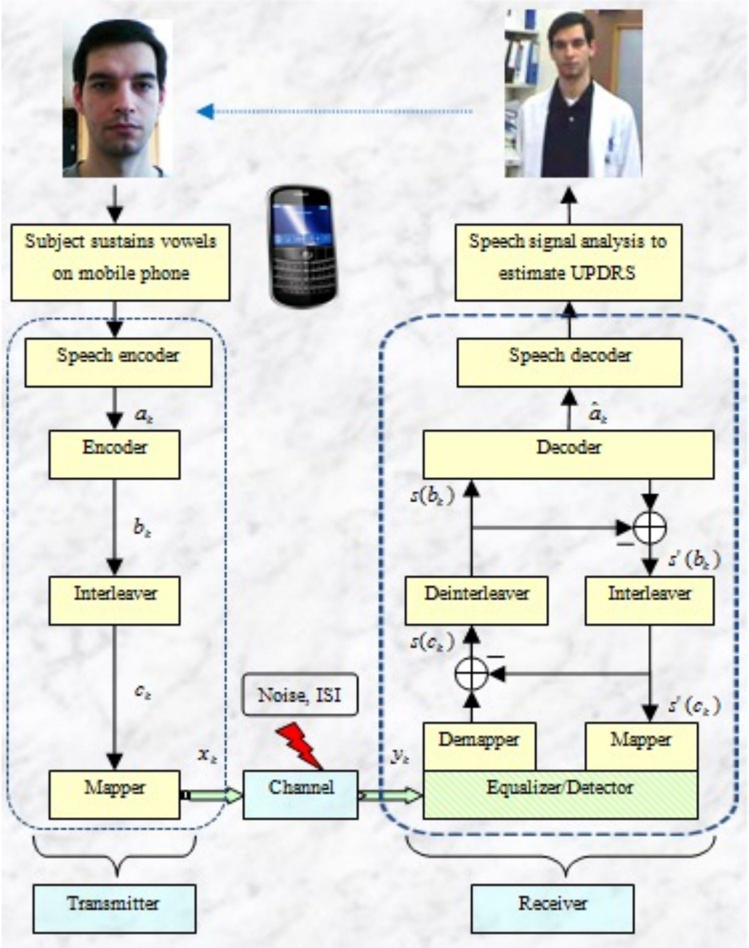


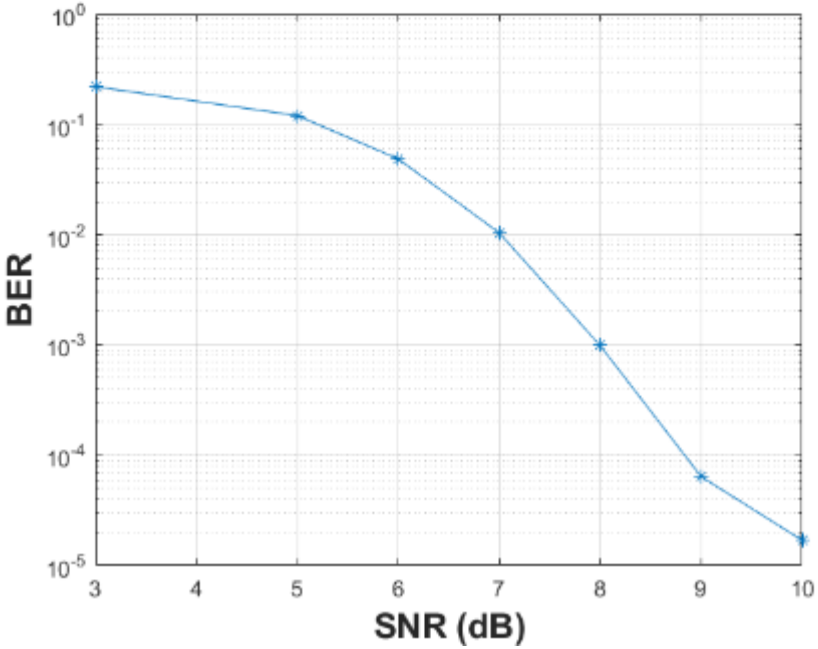
A. Tsanas, M.A. Little, P.E. McSharry, L.O. Ramig: Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity, **Journal of the Royal Society Interface**, Vol. 8, pp. 842-855, 2011

Top selected features

MALES (33 dysphonia measures)					FEMALES (33 dysphonia measures)				
Dysphonia measure	Motor UPDRS		Total UPDRS		Dysphonia measure	Motor UPDRS		Total UPDRS	
	MI	R	MI	R		MI	R	MI	R
6^{th} MFCC coef	0.106	-0.277	0.102	-0.294	Log energy	0.179	-0.458	0.170	-0.487
8^{th} MFCC coef	0.106	0.276	0.095	0.259	Std F0, Rapt	0.205	0.475	0.216	0.470
$VFER_{SNR,TKEO}$	0.077	-0.076	0.077	-0.108	10^{th} MFCC coef	0.112	0.239	0.107	0.250
$VFER_{mean}$	0.076	0.154	0.089	0.13	PPE	0.118	0.436	0.105	0.396
8^{th} delta MFCC	0.073	0.181	0.093	0.205	12^{th} MFCC coef	0.094	0.204	0.088	0.261
12^{th} delta MFCC	0.048	0.172	0.054	0.167	$IMF_{SNR,TKEO}$	0.075	-0.127	0.067	-0.067
0^{th} MFCC coef	0.079	0.171	0.097	0.197	8^{th} MFCC coef	0.114	-0.341	0.092	-0.255
2^{nd} MFCC coef	0.082	-0.149	0.084	-0.182	11^{th} MFCC coef	0.078	0.127	0.100	0.187
3^{rd} MFCC coef	0.071	0.091	0.077	0.067	$IMF_{NSR,SEO}$	0.099	-0.117	0.065	-0.058
2^{nd} delta MFCC	0.047	0.130	0.050	0.125	GNE_{mean}	0.090	0.035	0.086	-0.062
3^{rd} delta MFCC	0.046	0.169	0.054	0.161	3^{rd} delta MFCC	0.070	0.149	0.064	0.119
Std F0, Sun	0.046	0.144	0.050	0.129	HNR_{std}	0.072	0.224	0.066	0.195
9^{th} MFCC coef	0.075	-0.194	0.073	-0.153	5^{th} MFCC coef	0.113	0.173	0.115	0.188
7^{th} MFCC coef	0.079	-0.066	0.108	0.007	2^{nd} delta MFCC	0.055	0.172	0.056	0.206
4^{th} delta MFCC	0.041	0.001	0.044	0.007	$GNE_{SNR,TKEO}$	0.036	0.038	0.042	0.033
$GNE_{SNR,TKEO}$	0.023	0.074	0.024	0.089	10^{th} delta MFCC	0.071	-0.064	0.066	-0.079
$Shimmer_{A0,abs}$	0.042	-0.079	0.058	-0.135	GQ_{open}	0.061	0.256	0.057	0.248

Remote assessment using phones



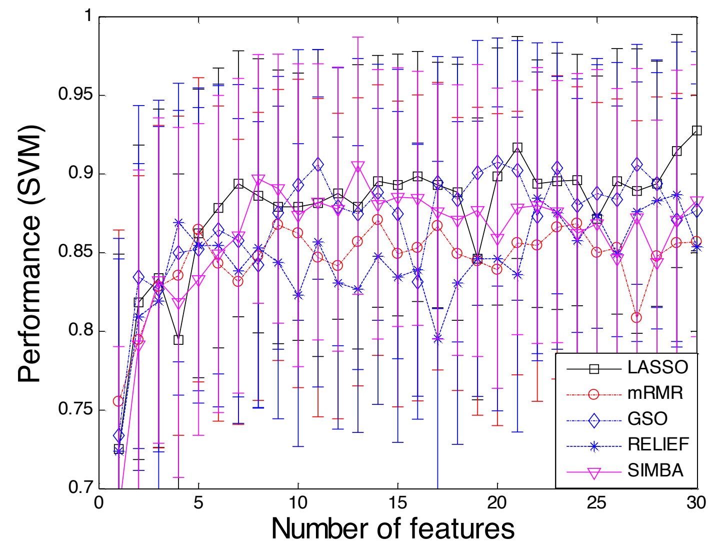


- Replicate UPDRS within ~3.5 points from ground truth

A. Tsanas, M.A. Little, L.O. Ramig: *Remote assessment of Parkinson's disease symptom severity using the simulated cellular mobile telephone network*, **IEEE Access**, Vol. 9, pp. 11024-11036, 2021

Project 3: Rehabilitation

- **LSVT dataset** (voice rehabilitation)
  - 14 PD subjects (8 male)
  - 156 phonations
  - Home rehabilitation protocol
  - **Aim: assess if voice is improving**



A. Tsanas, M.A. Little, C. Fox, L.O. Ramig: *Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease*, **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, Vol. 22, pp. 181-190, 2014 (My algorithms were used to validate a commercial product)

Project 4: Parkinson's Voice Initiative

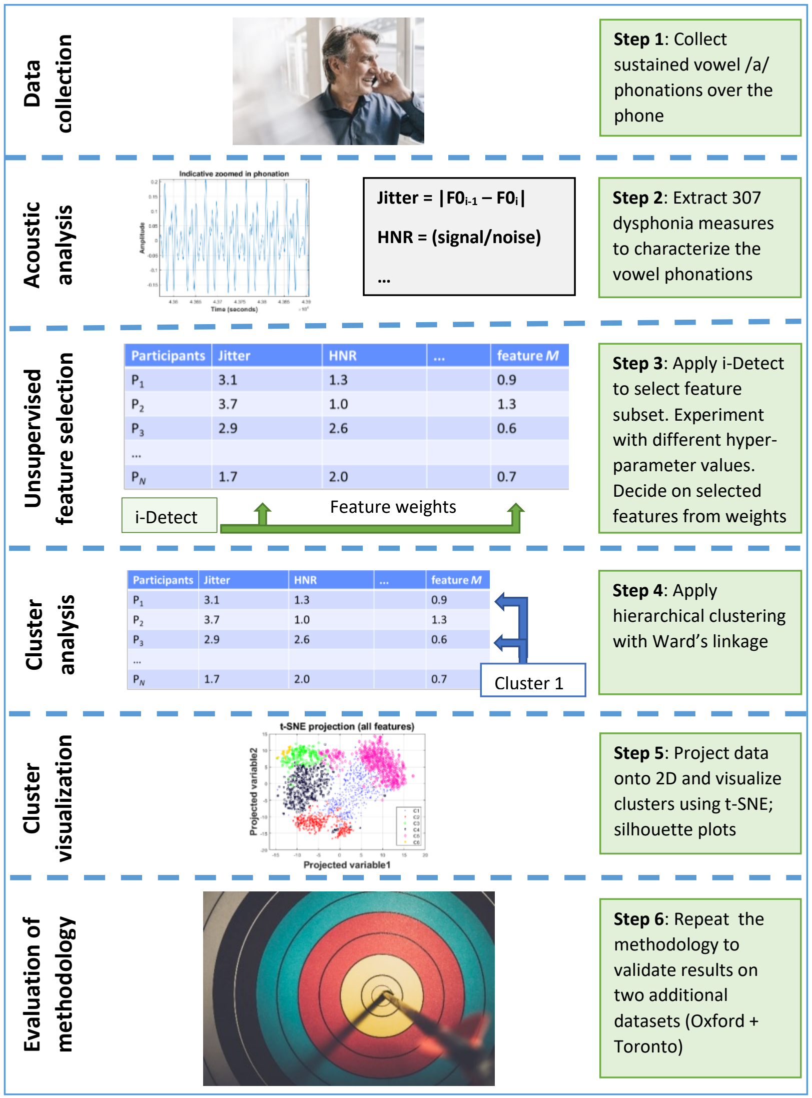
- **PVI dataset** (PD diagnosis self-reported)
  - 19,000+ phonations
  - 7 geographical locations (US, UK, Spain, Canada…)
  - Highly imbalanced: ~ 10% PD, 90% HC

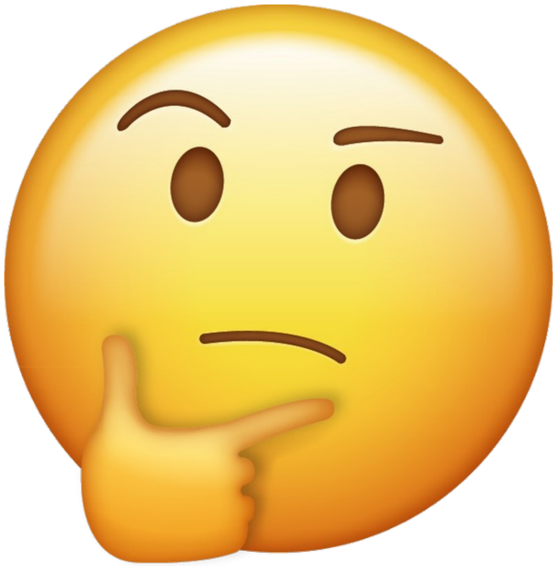
PVI
Parkinson's Voice Initiative

aculabcloud
A true cloud telephony platform

S. Arora, L. Baghai-Ravary, A. Tsanas: *Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality speech*, **Journal of Acoustical Society of America**, Vol. 145(5), 2871-2884, 2019

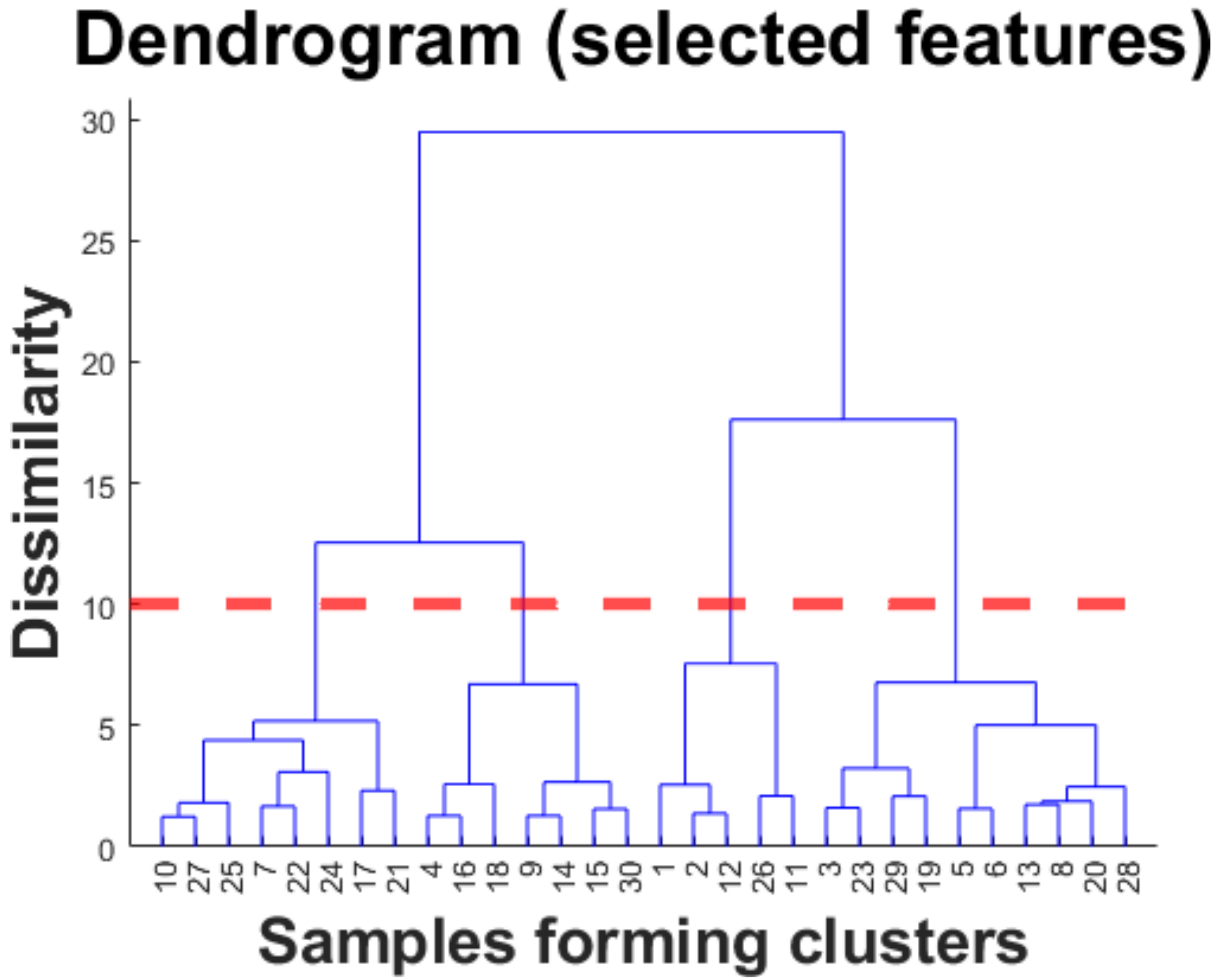
Project 5: Establishing PD subtypes



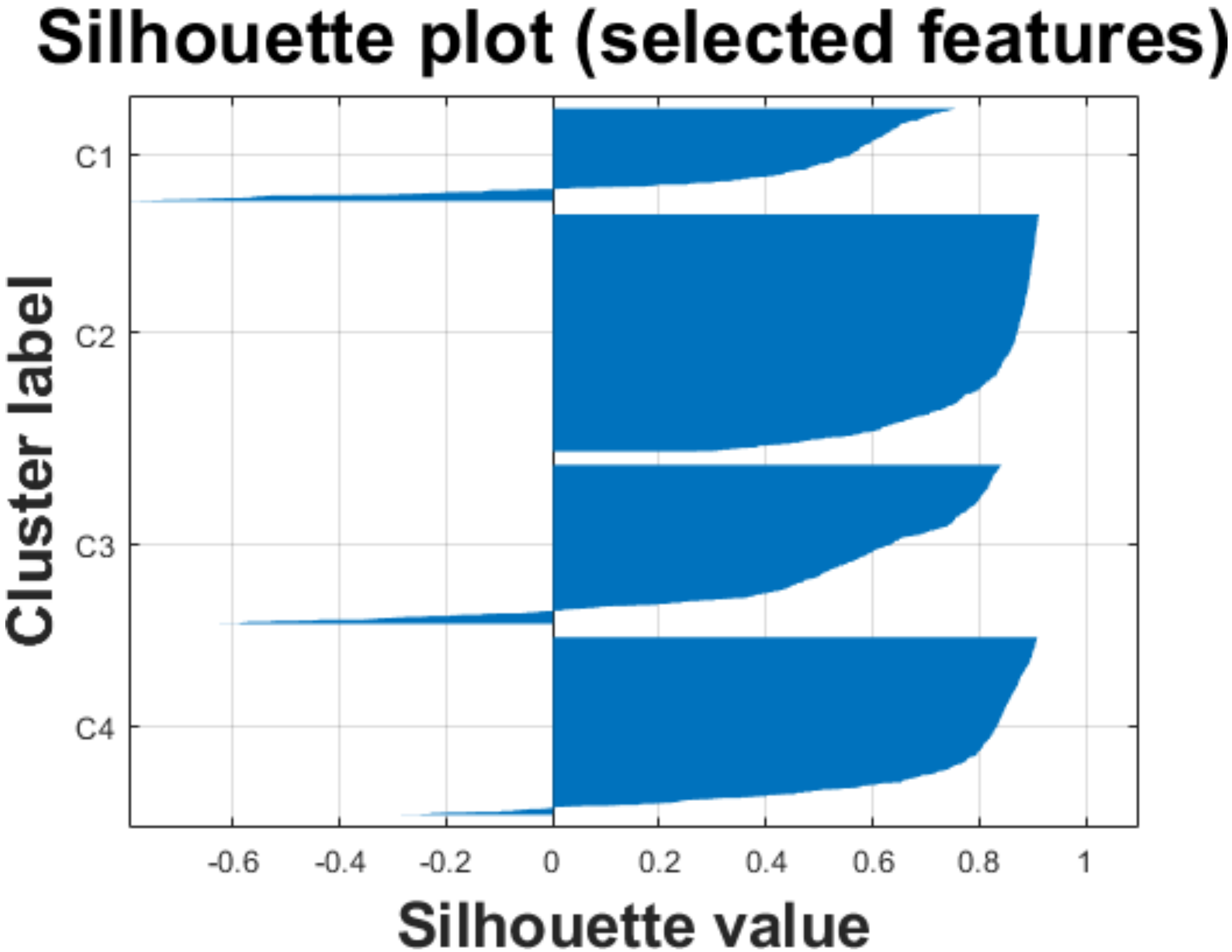


- Parkinson's as an umbrella term…
- **Can we define useful subtypes?**

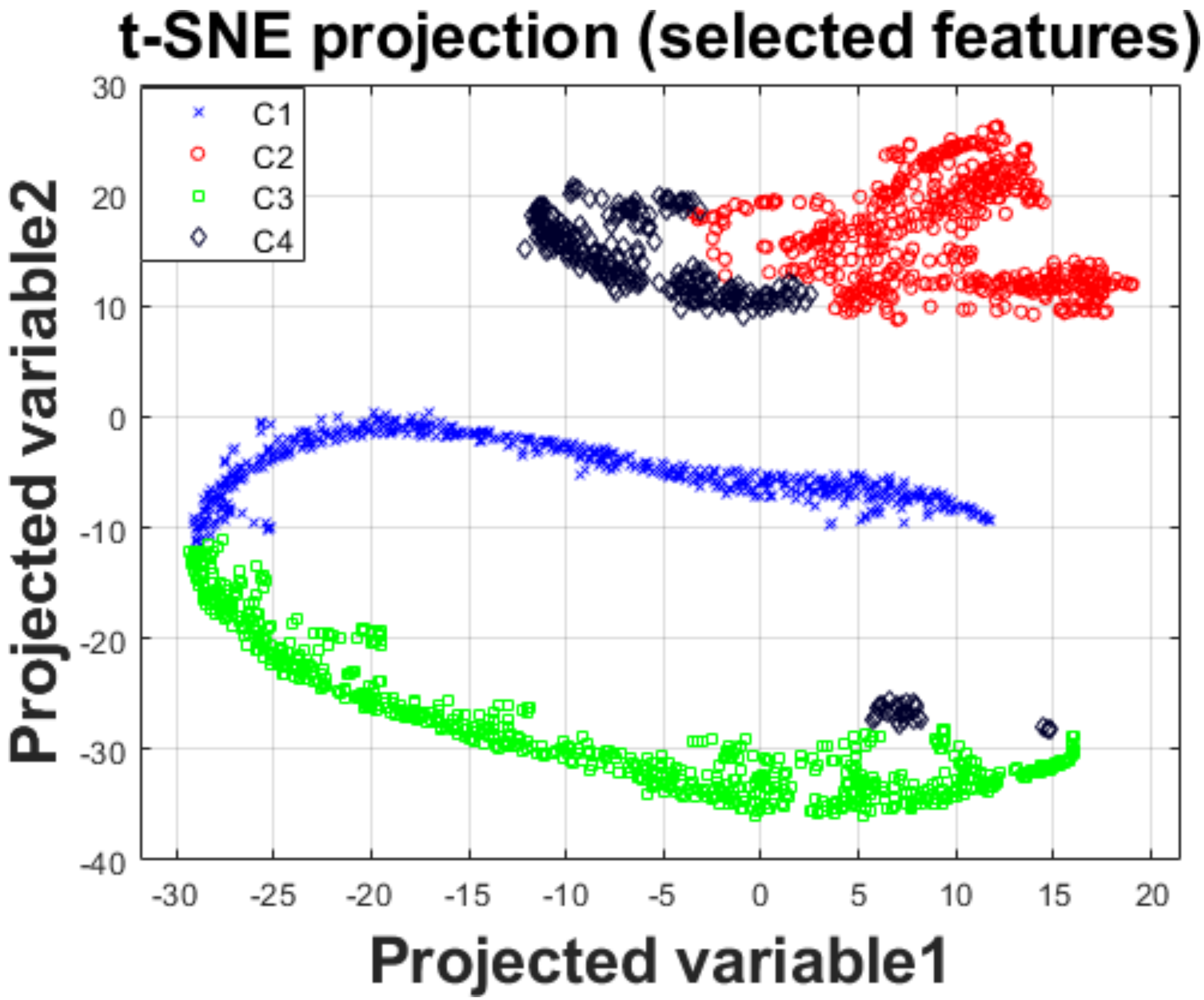
Hierarchical clustering



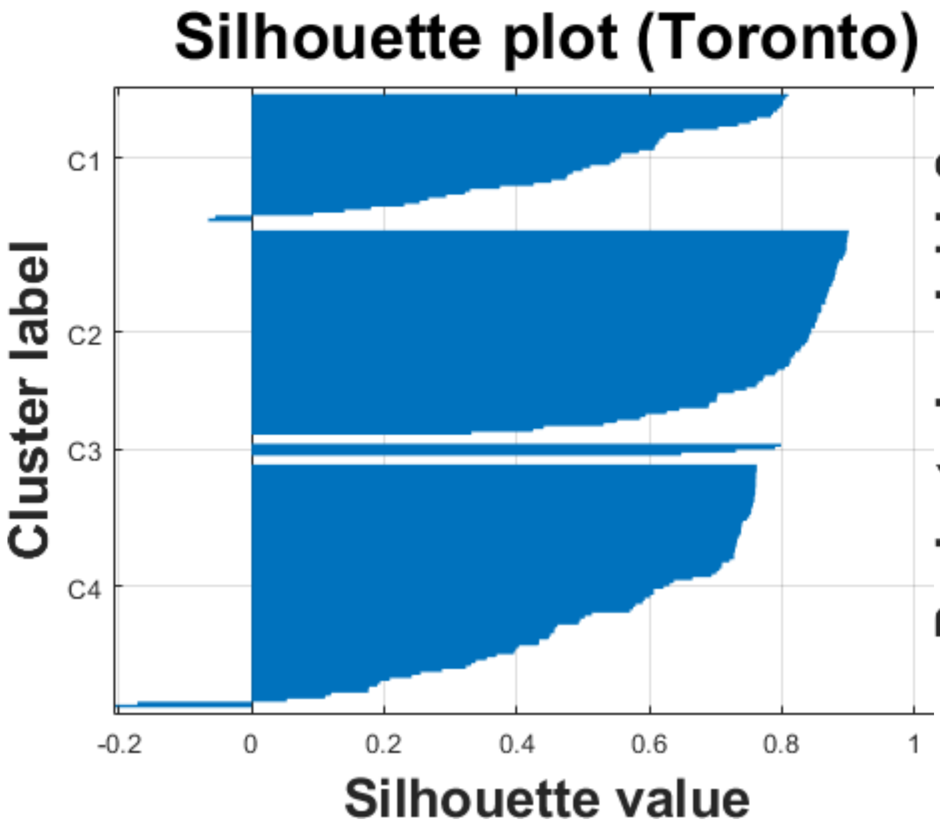
Establishing PD subtypes

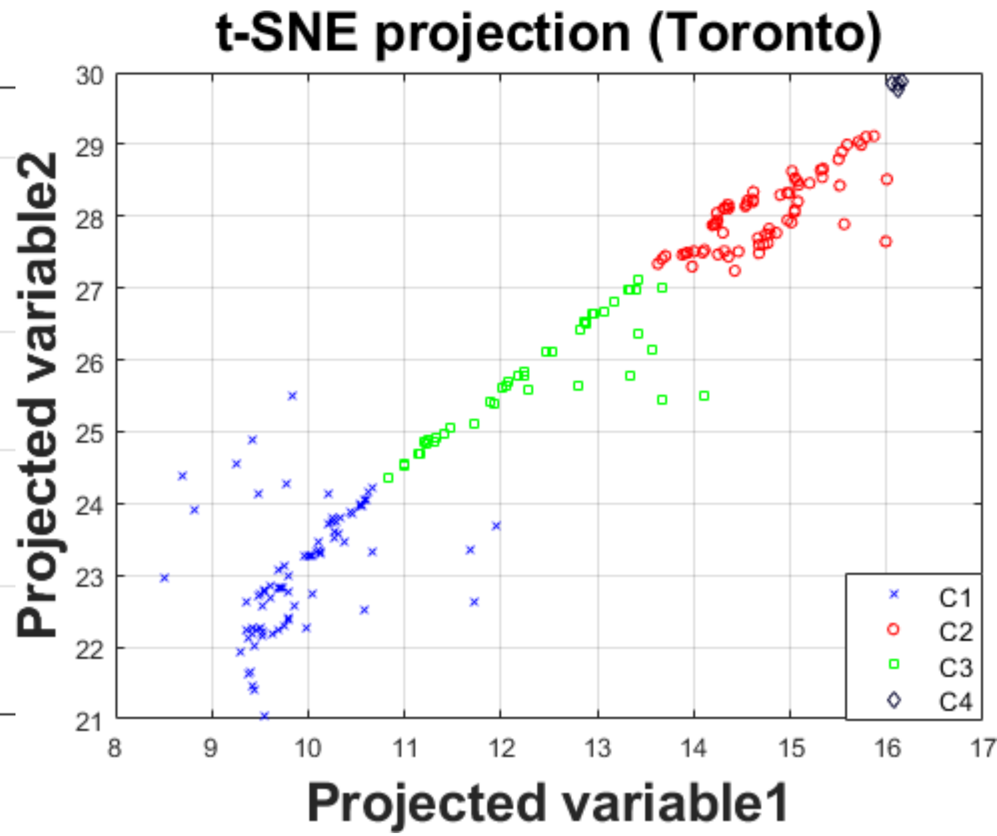


Establishing PD subtypes (Boston)



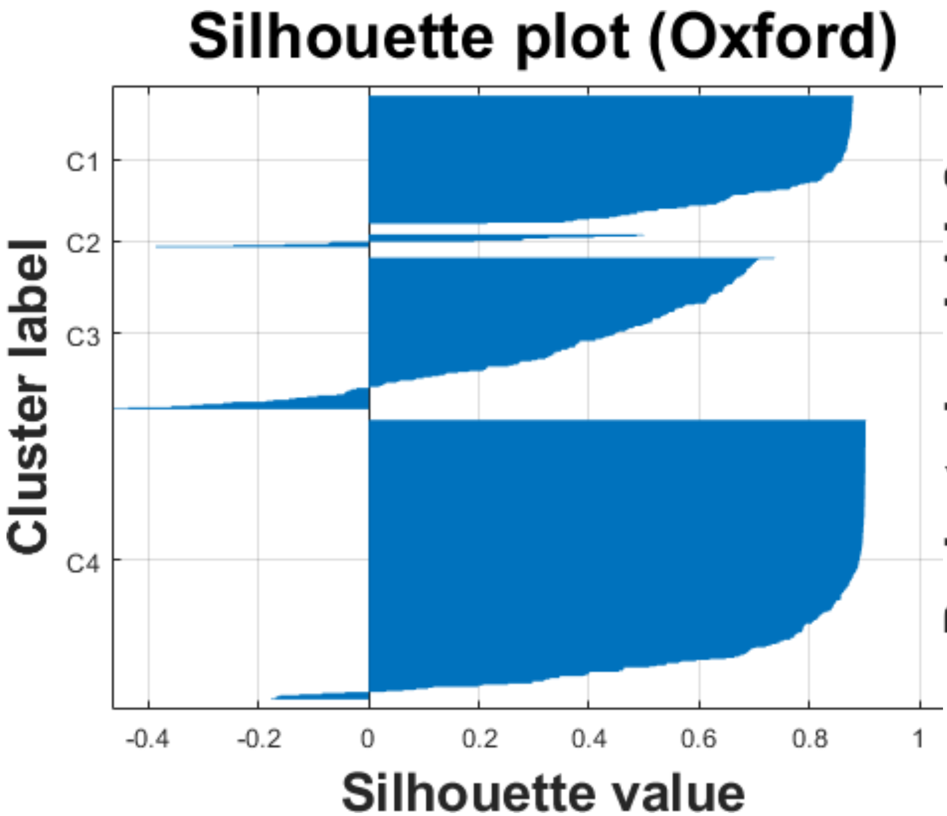
Generalizing PD subtypes (Toronto)

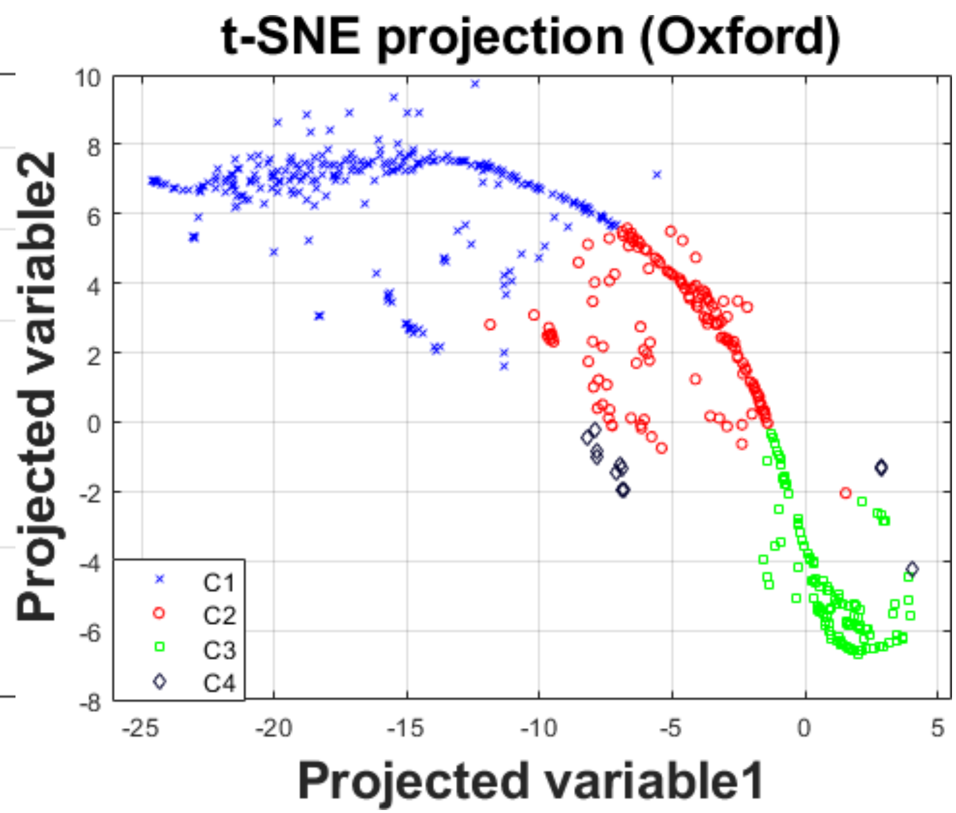




- Use the hyper-parameters from Boston
- Out of sample true generalization

Generalizing PD subtypes (Oxford)

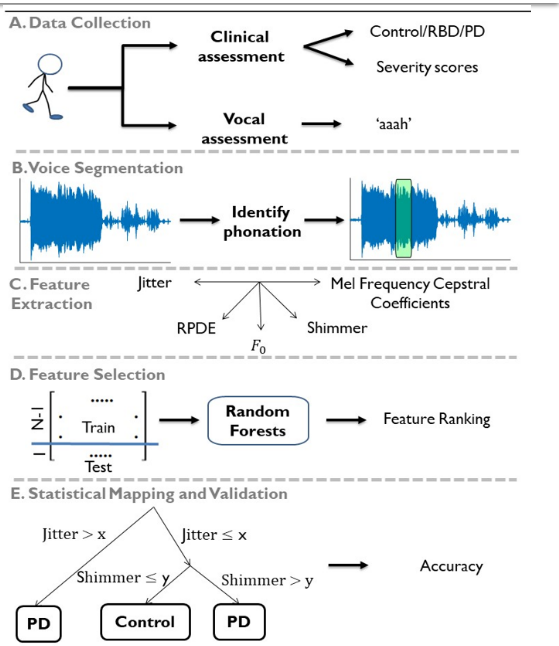


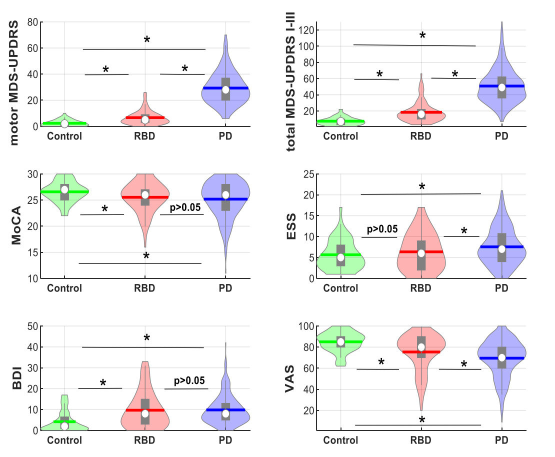


A. Tsanas, S. Arora: *Data-driven subtyping of Parkinson's using acoustic analysis of sustained vowels and cluster analysis: findings in the Parkinson's voice initiative study*, **Nature Computer Science**, (in press) 2022

Project 6: Early PD diagnosis?

- **PD-Toronto data**
  - Associate LRRK2 genetic mutations and PD voice
- **OPDC data**
  - 92 HC, 112 RBD, 335 PD
    (largest database in the world)





S. Arora, N.P. Visanji, T.A. Mestre, A. Tsanas, […], C. Marras: *Investigating voice as a biomarker for leucine-rich repeat kinase 2-associated Parkinson's disease: a pilot study*, **Journal of Parkinson's Disease**, Vol. 8(4), pp. 503-510, 2018

S. Arora, C. Lo, M. Hu, A. Tsanas: *Smartphone speech testing for symptom assessment in rapid eye movement sleep behavior disorder and Parkinson's disease*, **IEEE Access**, 2021

Conclusions

Voice signals

- Information-rich signals
- Can be used in diverse biomedical applications

High quality data

- Absolutely critical
- Standardize equipment

Advanced signal processing

- Key: knowledge of physiology
- Need to develop new, targeted algorithmic tools

Data and code available:

DARTH

https://www.darth-group.com

Data Analytics Research and Technology in Healthcare (DARTH group)

HOME | TEAM | PROJECTS | PUBLICATIONS | SOFTWARE | DATA | TEACHING | PHOTOS | CONTACT

DARTH group mission

Ever-increasing healthcare provision demands strain national health systems globally, which struggle to meet patients' needs. We passionately believe we can design and provide effective solutions which will revolutionize contemporary healthcare delivery

- Data from different projects
- Toolboxes (Matlab code)
  - Time-series analysis
  - Actigraphy
  - ML methods

Collaborators

Acknowledgements

THE UNIVERSITY OF EDINBURGH

UNIVERSITY OF OXFORD

US, Spain, Greece, Australia, UK

intel

LSVT GLOBAL Revolutionizing Rehabilitation

NCVS The National Center for Voice & Speech

aculabcloud

PVI Parkinson's Voice Initiative