Robust and adaptive representations for audio, video and other signals

Dr. Richard E. Turner (ret26@cam.ac.uk)

Computational and Biological Learning Lab
Department of Engineering
University of Cambridge
computer perception

vision and audition

new problems

machine learning

adaptive robustness

model constraints

tools and theories

neuroscience
Computer perception

Vision and audition

New problems

Adaptive robustness

Machine learning

Signal processing

Neuroscience

Model constraints

Tools and theories
signal

audio

video

brain recordings

speech recognition

source separation

hearing devices

voice manipulation

object recognition

EEG analysis
signal

audio
video
brain recordings

application

speech recognition
source separation
hearing devices
voice manipulation
object recognition
EEG analysis
signal

audio

video

brain recordings

application

speech recognition

source separation

hearing devices

voice manipulation

object recognition

EEG analysis

audio

video

brain recordings
signal

audio
video
brain recordings

application

- speech recognition
- source separation
- hearing devices
- voice manipulation
- object recognition
- EEG analysis
signal representation application

- audio
- video
- brain recordings

filter 1
filter 2
filter 3

speech recognition
source separation
hearing devices

voice manipulation
object recognition
EEG analysis
signal application representation application

signal representation application

audio video brain recordings

filter 1 filter 2 filter 3

speech recognition source separation hearing devices

voice manipulation object recognition EEG analysis
signal application representation application

filter 1

filter 2

filter 3

speech recognition
source separation
hearing devices

voice manipulation
object recognition
EEG analysis

audio
video
brain recordings

signal representation application
current methods:
- not robust
- not adaptive
signal representation application

audio
video
brain recordings

filter 1
filter 2
filter 3

speech recognition
source separation
hearing devices
voice manipulation
object recognition
EEG analysis

current methods:
not robust
not adaptive
Core problem

\[ y(t) \]
Core problem

\[ y(t) = \cos(\phi(t)) \]
Ill posed

\[ y(t) = \cos(\phi(t)) \]

\[ a(t) \]
Constraints

\[ a(t) \]

\[ y(t) = c(t) \cos(\phi(t)) \]

\[ \text{frequency } \frac{d}{dt} \phi(t) \]
Problems with current approaches
Problems with current approaches

Problems with current approaches

Problems with current approaches

Problems with current approaches

Principled approach

- **Ill posed** problem - explicitly state assumptions
- **Soft constraints** - interpolate through noise
- **Adaptability** - learn time-scales
- Solution: **probabilistic approach**
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \mathcal{R} [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t)) ] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \mathbb{R} \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i \phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re[a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re[a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re [a(t) \exp(i\phi(t))] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach

\[ y(t) = \Re \left[ a(t) \exp(i\phi(t)) \right] \]
New twist on standard approach
New twist on standard approach

\[ y_t = [1, 0]x_t + \sigma_y \eta_t \]

\[ x_t = R x_{t-1} + \sigma \epsilon_t \]

soft constraints

match signal

rotate slowly
New twist on standard approach

![Diagram](image)

soft constraints
match signal
\[ y_t = [1, 0]x_t + \sigma_y \eta_t \]
rotate slowly
\[ x_t = Rx_{t-1} + \sigma \epsilon_t \]

Estimation: Kalman Smoother
\[ p(x_t | y_{1:T}) \]

Turner, 2010
Running example
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude

![Graph showing phasors with large initial amplitude.](image)
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude
Phasor - large initial amplitude

![Graph showing phasors with large initial amplitude]
Phasor - large initial amplitude

\[ \begin{align*}
&\text{time} \\
&x_1 \\
&x_2 \\
&y
\end{align*} \]
Phasor - large initial amplitude

Large amplitude:
fixed frequency
Phasor dynamics - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude

\[ x_1(t) = \cos(\omega t) \]
\[ x_2(t) = \sin(\omega t) \]

where \( \omega \) is the angular frequency and \( t \) is time.
Phasor - small initial amplitude
Phasor - small initial amplitude

![Diagram showing phasor with small initial amplitude]
Phasor - small initial amplitude
Phasor - small initial amplitude
Phasor - small initial amplitude

Small amplitude: unconstrained frequency
Summary

Observed

• Large amplitude $\implies$ strong constraint on frequency

• Small amplitude $\implies$ no constraint on frequency

Desired

• Large amplitude $\implies$ no constraint on frequency

• Small amplitude $\implies$ strong constraint on frequency

Solution

• Place independent constraints on amplitude and frequency
Bivariate Isotropic Gaussian: $p(x_1, x_2) = \text{Norm}(x; \mu, \sigma^2 I)$
Restrict to unit circle: $p(\theta | x_1^2 + x_2^2 = 1)$
von Mises:  \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: $p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu))$
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: $p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu))$
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: 

\[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises:  

\[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \( p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \)
von Mises: \[ p(\theta) = \frac{1}{Z(k)} \exp(k \cos(\theta - \mu)) \]
Bivariate Anisotropic Gaussian: \( p(x_1, x_2) = \text{Norm}(\mathbf{x}; \mu, \Sigma) \)
$$p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2)))$$
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
\[ p(\theta) = \frac{1}{Z(k)} \exp(k_1 \cos(\theta - \mu_1) + k_2 \cos(2(\theta - \mu_2))) \]
Time-series version: Dynamics (Prior)
Time-series version: Estimation (Posterior)
Summary

- **Powerful method** for time-series of circular variables
- **Connects** simple Gaussians to difficult circular variables
- **Estimation**: Kalman smoothing with a novel moment-matching step
RESULTS
Synthetic signals
Synthetic signals: Hilbert
Synthetic signals: New approach
Synthetic signals - Missing Data
Synthetic signals - Missing Data
Synthetic signals - Missing Data
Natural signals: running example
Natural signals: running example
Natural signals: running example
Natural signals: Speech
Summary

- state of the art demodulation
- higher quality
- higher computational cost

• applications: voice manipulation
  brain recording (EEG analysis)
  cochlear implants
  auditory science
Audio modelling

![Spectrogram showing various audio events like fire, stream, wind, rain, footstep, tent-zip, and time in seconds. Turner, 2010]
Audio modelling

Turner, 2010
Audio modelling

Turner, 2010
Statistical texture synthesis

• Old approach: build **detailed physical models** (e.g. rain drops)

• New approach
  – **train model** on your favourite texture
  – **sample** from the prior, and then from the likelihood.

• Waveform unique, but statistically matched to original

• Often perceptually indistinguishable
Audio denoising

Turner, 2010
Probabilistic signal processing

Classical signal processing

?
Probabilistic signal processing

Classical signal processing

fixed

Gaussian

?
Probabilistic signal processing ? Classical signal processing

fixed Gaussian

STFT filter bank & Hilbert spectrogram

Turner et al. in prep.
Probabilistic signal processing

Classical signal processing

- robustness adaptation
- fast methods
- important variables

STFT
filter bank & Hilbert spectrogram

Turner et al. in prep.
video
signal processing

learning

Slow Feature Analysis
(classical method: object recognition, gesture recognition)

Turner et al. 2007
state of the art video denoising
Turner, submitted.
Probabilistic approach is powerful and general

- Probabilistic language **unifying** frame-work
- **Robust to noise**
- **Adapt to the signal**
Current and future projects

**machine learning**
- approximate inference

**vision**
- complex transformations
- occlusion

**neuroscience**
- analysis brain recordings

**time-series modelling**

---

**Dr. Maneesh Sahani, UCL**
**Prof. Zoubin Ghahramani, Cambridge**

---

**Prof. Eero Simoncelli, NYU**
**Dr. Jörg Lücke, Frankfurt**

---

**Prof. John O'Keefe, UCL**
**Dr. Bob Carlyon, Cambridge**
**Dr. David Baguley, Cambridge**
Additional slides
Estimation: Soften constraints
Estimation: Soften constraints
Estimation: Soften constraints
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP

\[ t - 2 \quad t - 1 \quad t \quad t + 1 \quad t + 2 \]
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP

t−2  t−1  t  t+1  t+2
Estimation via EP

\[ t+2 \quad t+1 \quad t \quad t-1 \quad t-2 \]
Estimation via EP

t−2 t−1 t t+1 t+2
Estimation via EP
Estimation via EP

\[ t, t-1, t-2, t+1, t+2 \]
Estimation via EP
Estimation via EP
Estimation via EP

t−2  t−1  t  t+1  t+2
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP
Estimation via EP

t−2  t−1  t  t+1  t+2
Estimation via EP
Estimation via EP
Estimation via EP
Synthetic signals: Hilbert
Synthetic signals: Filter + Hilbert

![Graph showing synthetic signals with time and frequency axes. The signal is plotted against time and frequency, with shaded areas indicating specific time intervals.]
Synthetic signals: Probabilistic

![Graph showing synthetic signals with time and frequency axes.]
EEG Data

raw

filtered

frequency /Hz

time /s

0 0.5 1 1.5
0 5 10
0 0.5 1 1.5
0 5 10
EEG Data
Probabilistic signal processing

Cemgil & Godsill

Classical signal processing

Filter Bank & Hilbert

X

estimation

Turner et al. in prep.
Probabilistic signal processing

Cemgil & Godsill

Freq shift

Qi & Minka

Classical signal processing

Filter Bank & Hilbert

Freq shift

STFT

X ←→ Z

Estimation

X ←→ Z

Turner et al. in prep.
Probabilistic signal processing

Cemgil & Godsill

Qi & Minka

freq shift

abs

abs

a

Amplitudes

Classical signal processing

Filter Bank & Hilbert

STFT

freq shift

abs

abs

a

Spectrogram

estimation

Turner et al. in prep.
Probabilistic signal processing \[\leftrightarrow\] Classical signal processing

Cemgil & Godsill

freq shift

Amplitudes

a

abs

Qi & Minka

freq shift

STFT

Filter Bank & Hilbert

abs

estimation

Spectrogram

a

turner et al. in prep.