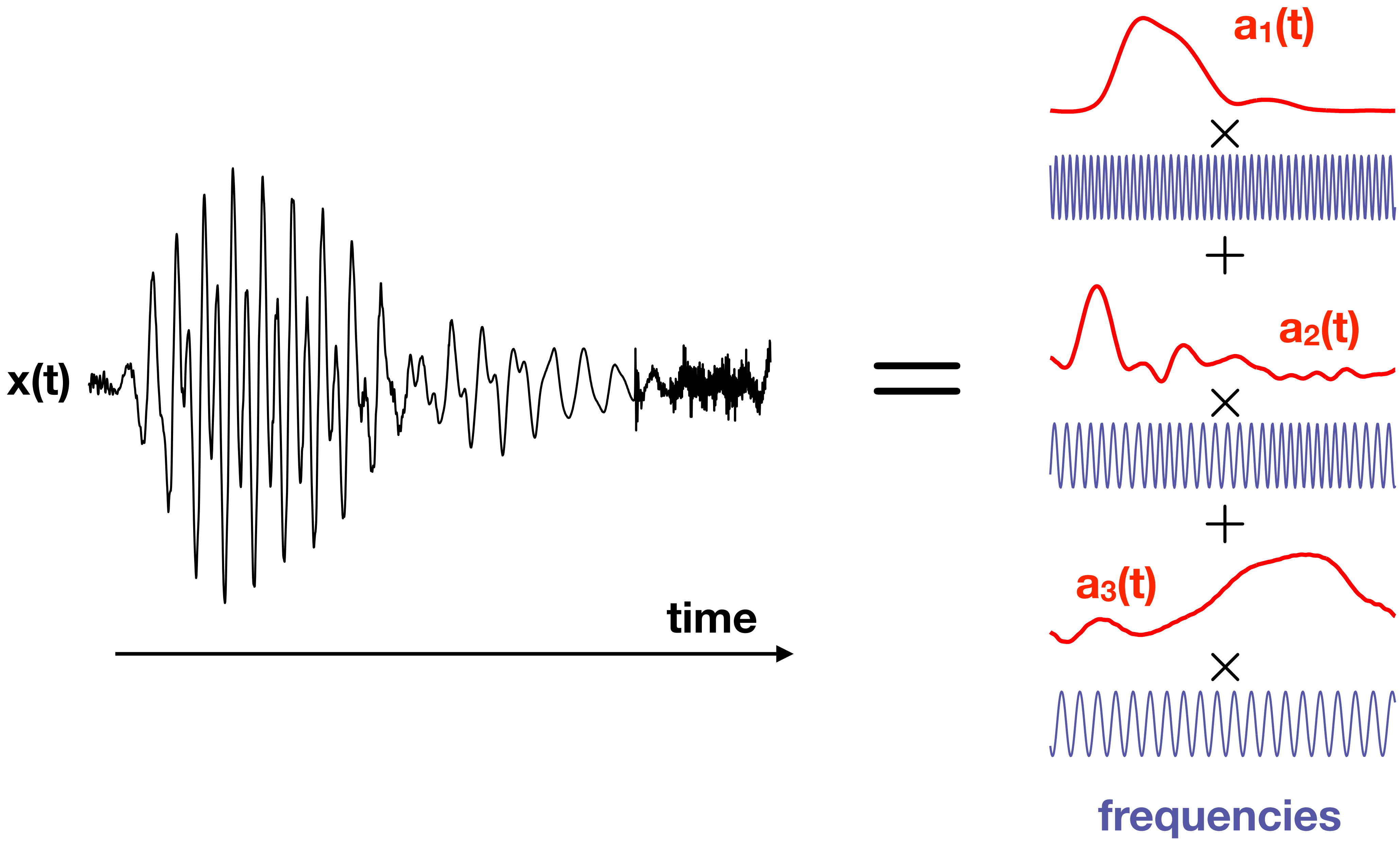
DS-GA 3001.008 Modeling time series data

L12. Spectral methods

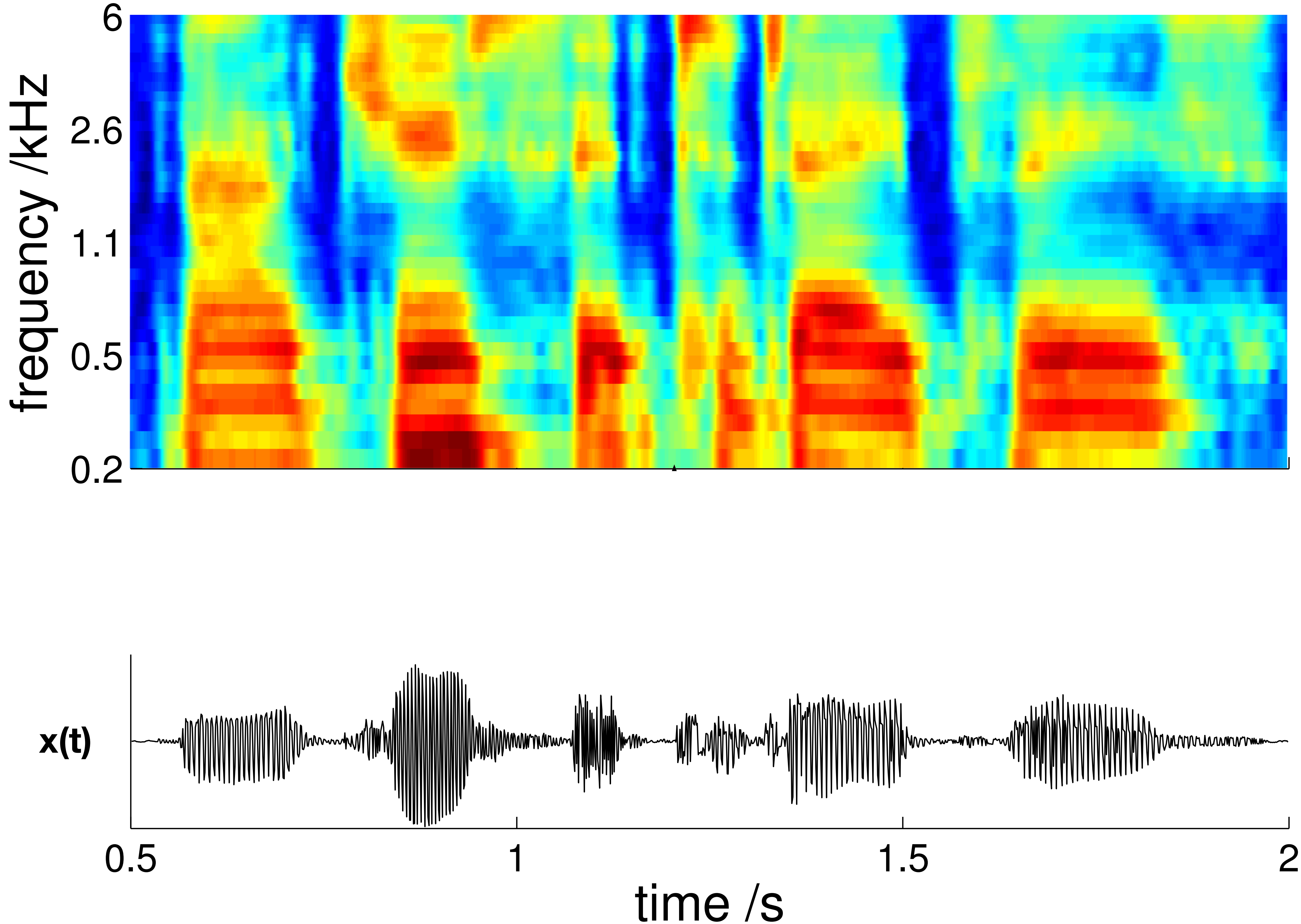
Instructor: Cristina Savin

NYU, CNS & CDS

Change of representation: from time to frequencies



Motivation:
periodic structure is by def. predictable
if it’s there in the data then we should take advantage of it



frequency /kHz
6
2.6
1.1
0.5
0.2
x(t)
0.5
1
1.5
2
time /s

Overview:

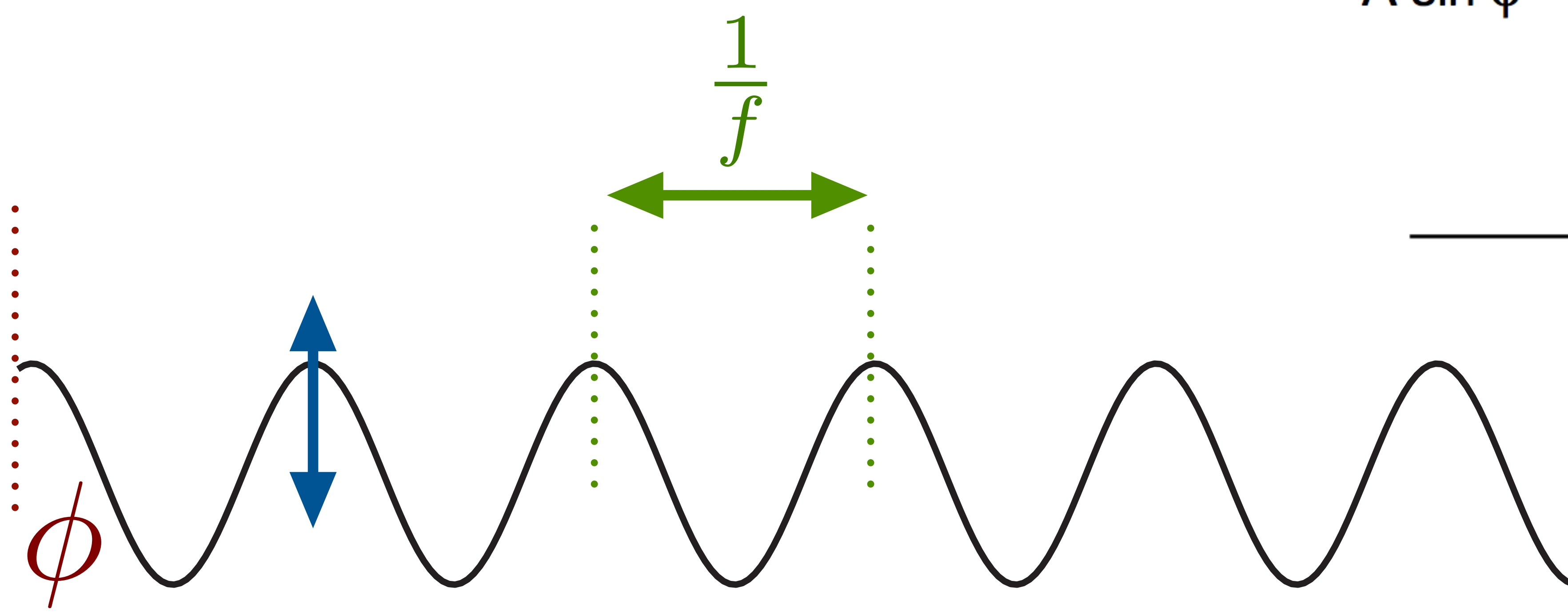
The traditional signal processing view: Fourier transform

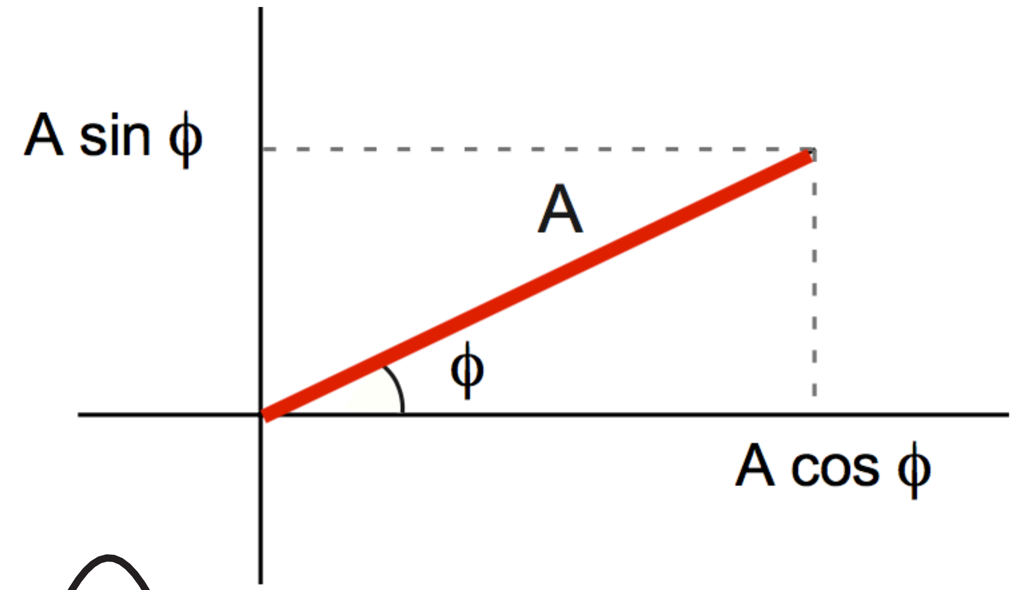
Stats perspective (chp. 4 in tsa4.pdf)

Probabilistic spectral analysis

Modeling real-world data

Reminder basics frequency analysis





$$x_t = \underbrace{A}_{\text{amplitude}}\cos(\underbrace{2\pi\omega}_{\text{frequency}} t + \underbrace{\phi}_{\text{phase}})$$

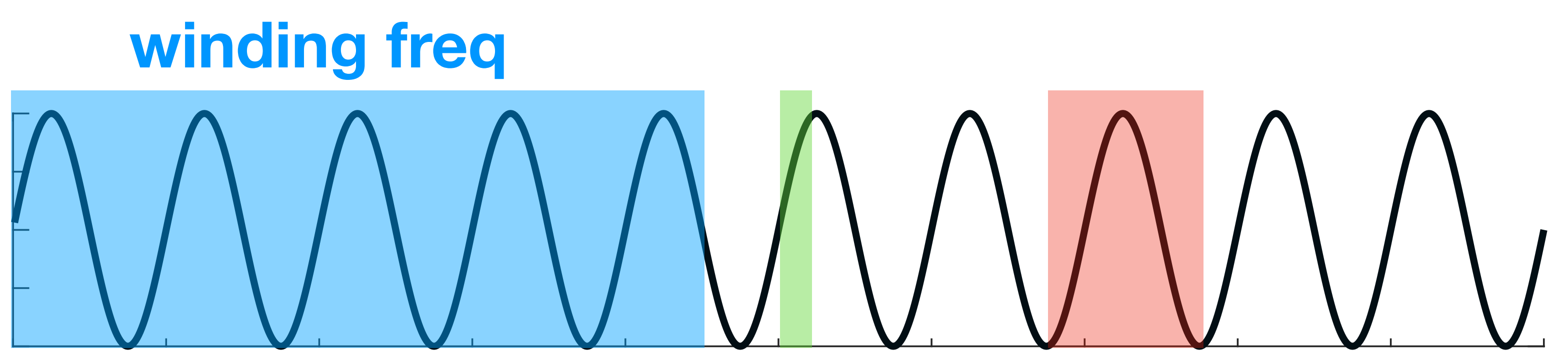
The Fourier transform

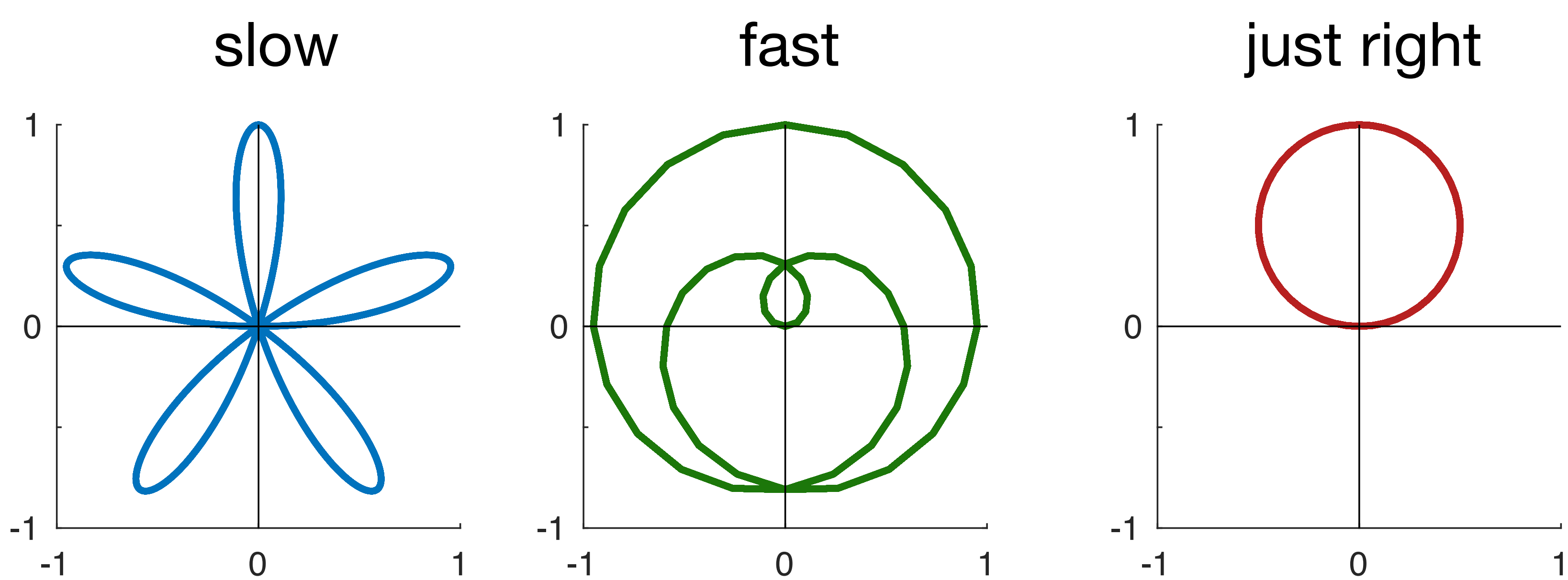
$$F(\omega) = \int_{-\infty}^{\infty} x(t) e^{-2\pi i \omega t} dt$$

***inverse**

**mathematically convenient

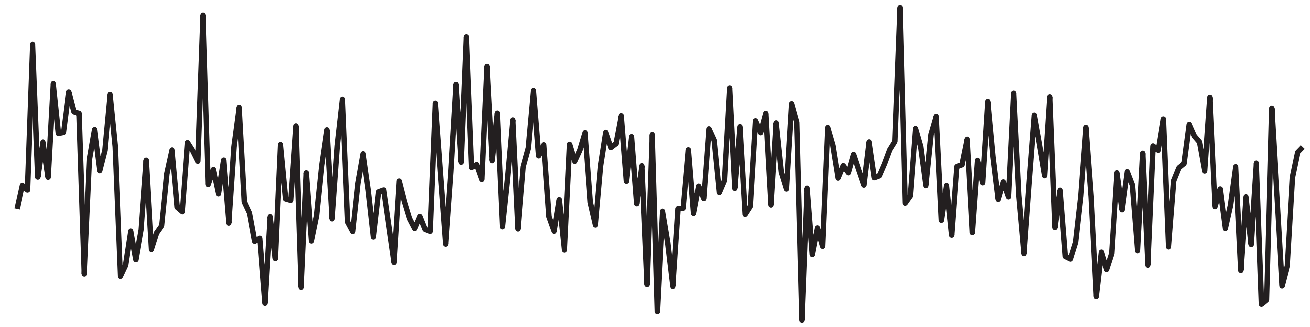
***discrete vs continuous





Gentle introduction: "3 blue 1 brown Fourier"

Stationary time series

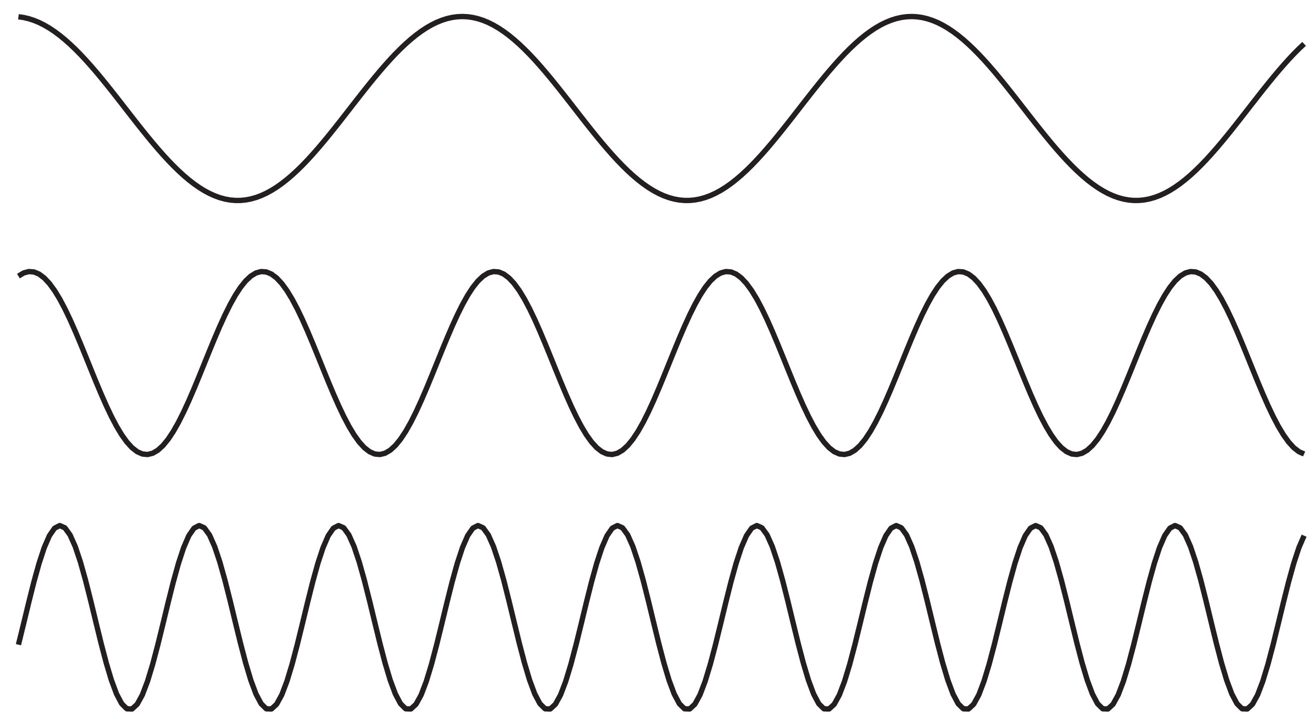


frequency domain

$a_1 \times$

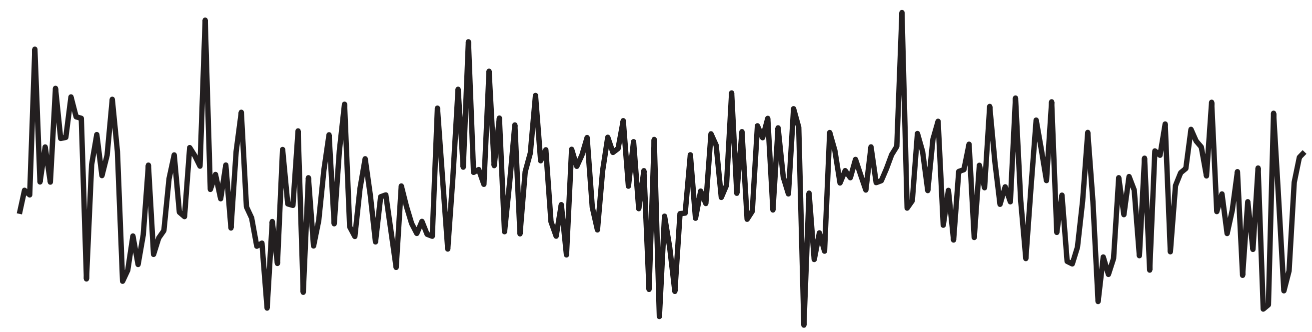
$a_2 \times$

$a_3 \times$

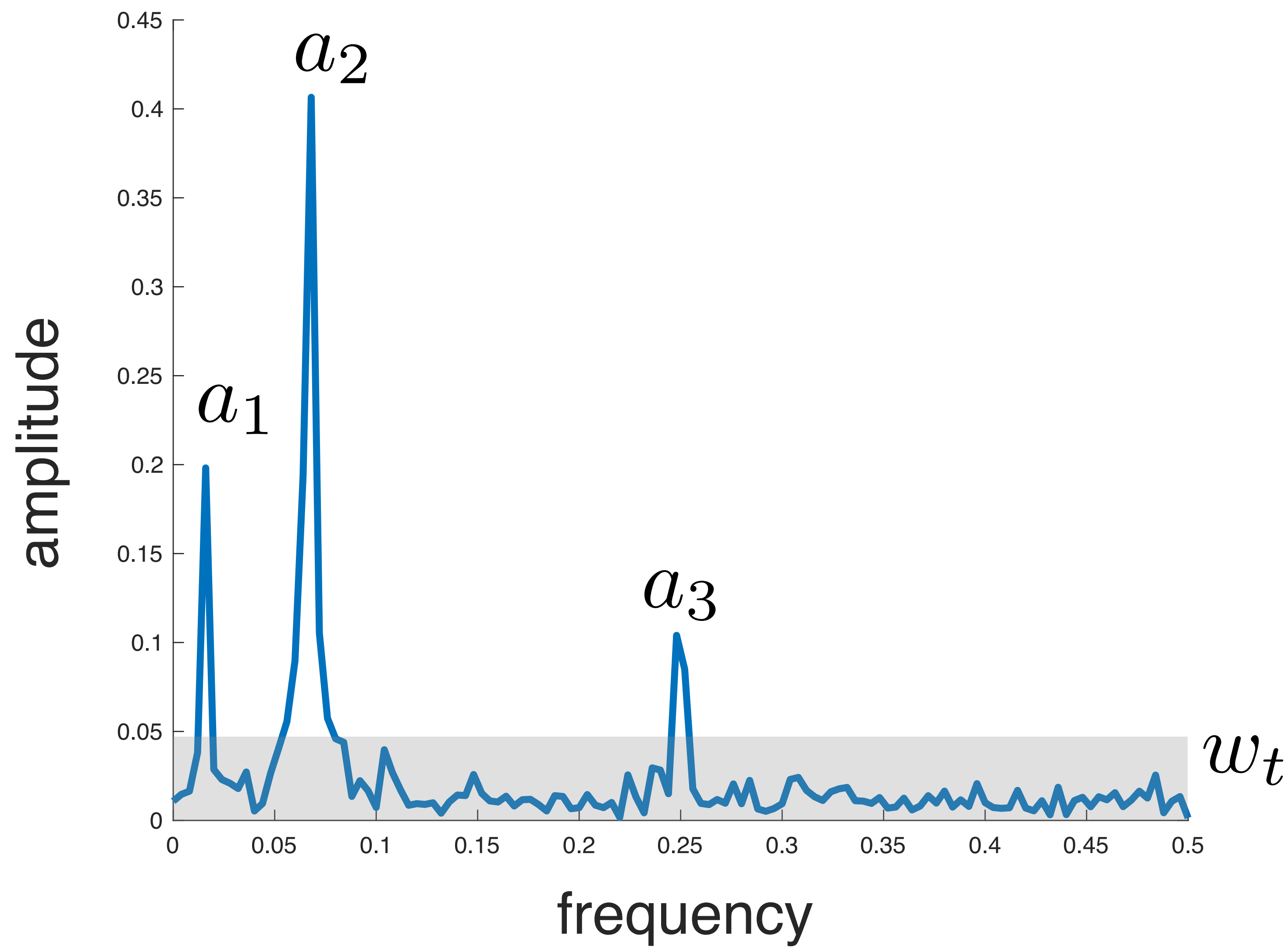


+ white noise

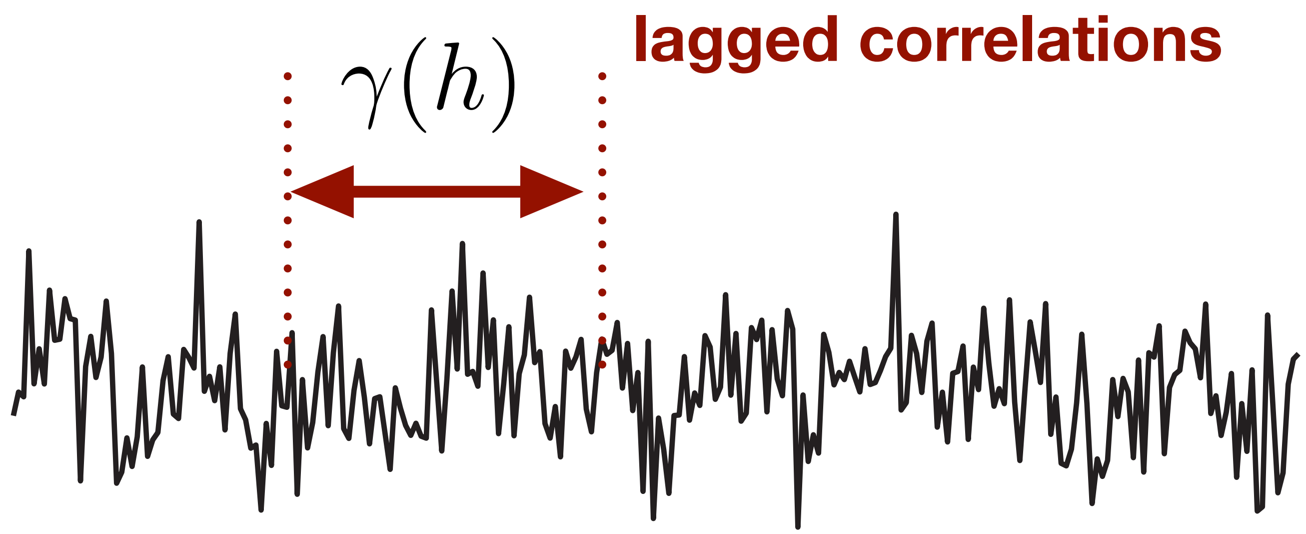
Stationary time series



frequency domain



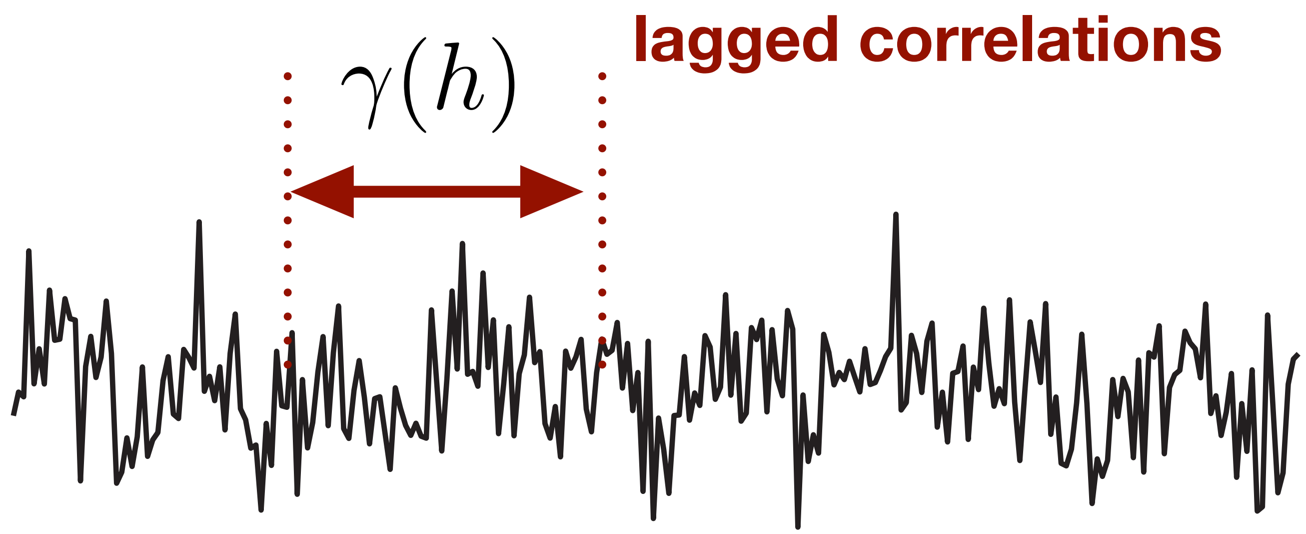
Stationary time series



ARIMA-style

$$
\begin{aligned}
\gamma_x(h) = \operatorname{cov}(x_{t+h}, x_t) &= \operatorname{cov}(U_1 c_{t+h} + U_2 s_{t+h}, U_1 c_t + U_2 s_t) \\
&= \operatorname{cov}(U_1 c_{t+h}, U_1 c_t) + \operatorname{cov}(U_1 c_{t+h}, U_2 s_t) \\
&\quad + \operatorname{cov}(U_2 s_{t+h}, U_1 c_t) + \operatorname{cov}(U_2 s_{t+h}, U_2 s_t) \\
&= \sigma^2 c_{t+h} c_t + 0 + 0 + \sigma^2 s_{t+h} s_t = \sigma^2 \cos(2\pi\omega h),
\end{aligned}
$$

Stationary time series



ARIMA-style

$$x_t = \sum_{k=1}^{q} \left[U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t) \right],$$

$$\gamma_x(h) = \sum_{k=1}^{q} \sigma_k^2 \cos(2\pi\omega_k h),$$

Discrete time-discrete frequencies

$$x_t = a_0 + \sum_{j=1}^{(n-1)/2} \left[a_j \cos(2\pi t\, j/n) + b_j \sin(2\pi t\, j/n)\right],$$

Nyquist freq/ folding freq

$$a_j = \frac{2}{n}\sum_{t=1}^{n} x_t \cos(2\pi t j/n) \quad \text{and} \quad b_j = \frac{2}{n}\sum_{t=1}^{n} x_t \sin(2\pi t j/n).$$

$$\mathcal{O}\left(n^2\right)$$

Scaled periodogram

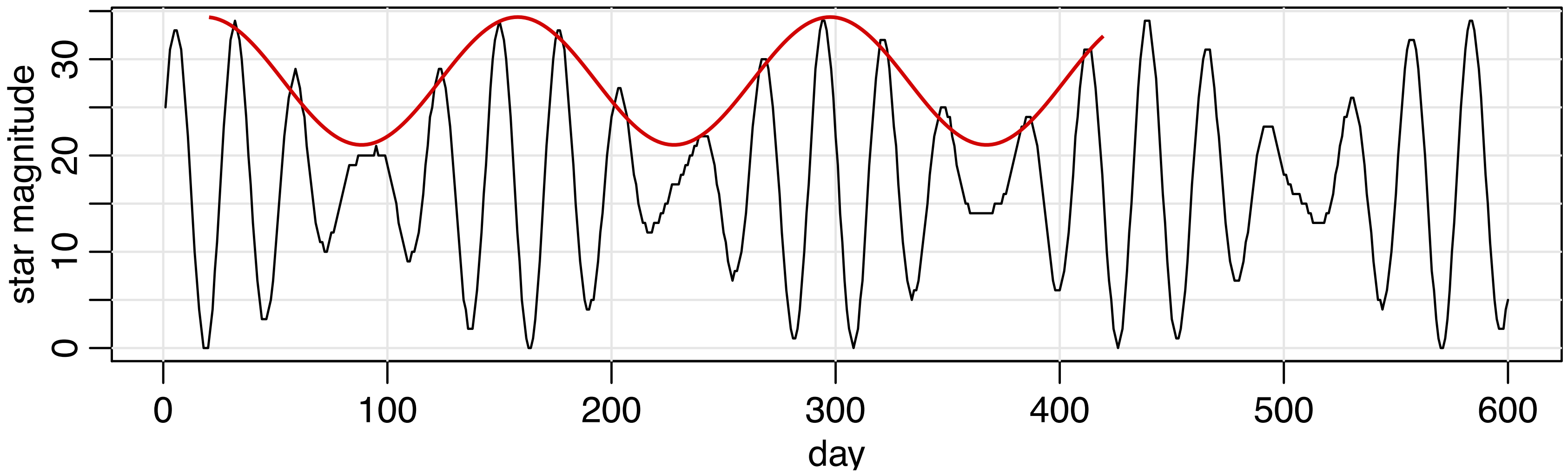
$$P(j/n) = a_j^2 + b_j^2 = \frac{4}{n}\,|d(j/n)|^2.$$

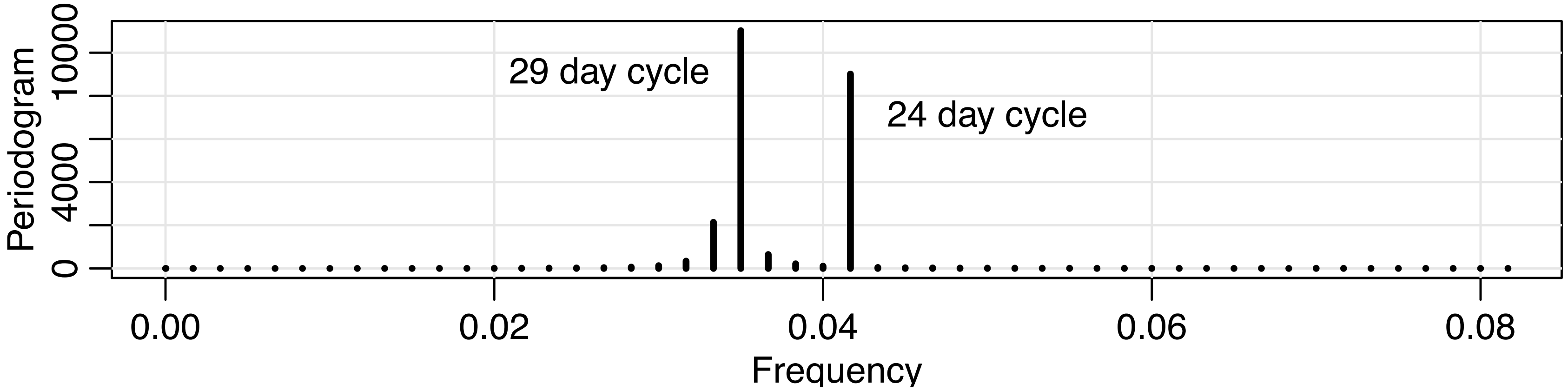
computed efficiently using **discrete Fourier transform DFT**

$$\begin{aligned} d(j/n) &= n^{-1/2}\sum_{t=1}^{n} x_t \exp(-2\pi i t j/n) \\ &= n^{-1/2}\left(\sum_{t=1}^{n} x_t \cos(2\pi t j/n) - i\sum_{t=1}^{n} x_t \sin(2\pi t j/n)\right), \end{aligned}$$

$$\mathcal{O}(n\log(n))$$

amplitude modulated signal





Spectral density

$$x_t = U_1 \cos(2\pi\omega_0 t) + U_2 \sin(2\pi\omega_0 t)$$

U_1 and U_2 are uncorrelated zero-mean random variables with equal variance σ^2.

$$\gamma(h) = \sigma^2 \cos(2\pi\omega_0 h) = \frac{\sigma^2}{2} e^{-2\pi i\omega_0 h} + \frac{\sigma^2}{2} e^{2\pi i\omega_0 h} = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega h} dF(\omega)$$

spectral distribution function

$$F(\omega) = \begin{cases} 0 & \omega < -\omega_0, \\ \sigma^2/2 & -\omega_0 \leq \omega < \omega_0, \\ \sigma^2 & \omega \geq \omega_0. \end{cases}$$

~cumulative variance

always exists, unique

$$dF(\omega) = f(\omega)\, d\omega,$$

If the autocovariance function, $\gamma(h)$, of a stationary process satisfies

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \tag{4.15}$$

then it has the representation

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega h} f(\omega)\, d\omega \quad h = 0, \pm 1, \pm 2, \ldots \tag{4.16}$$

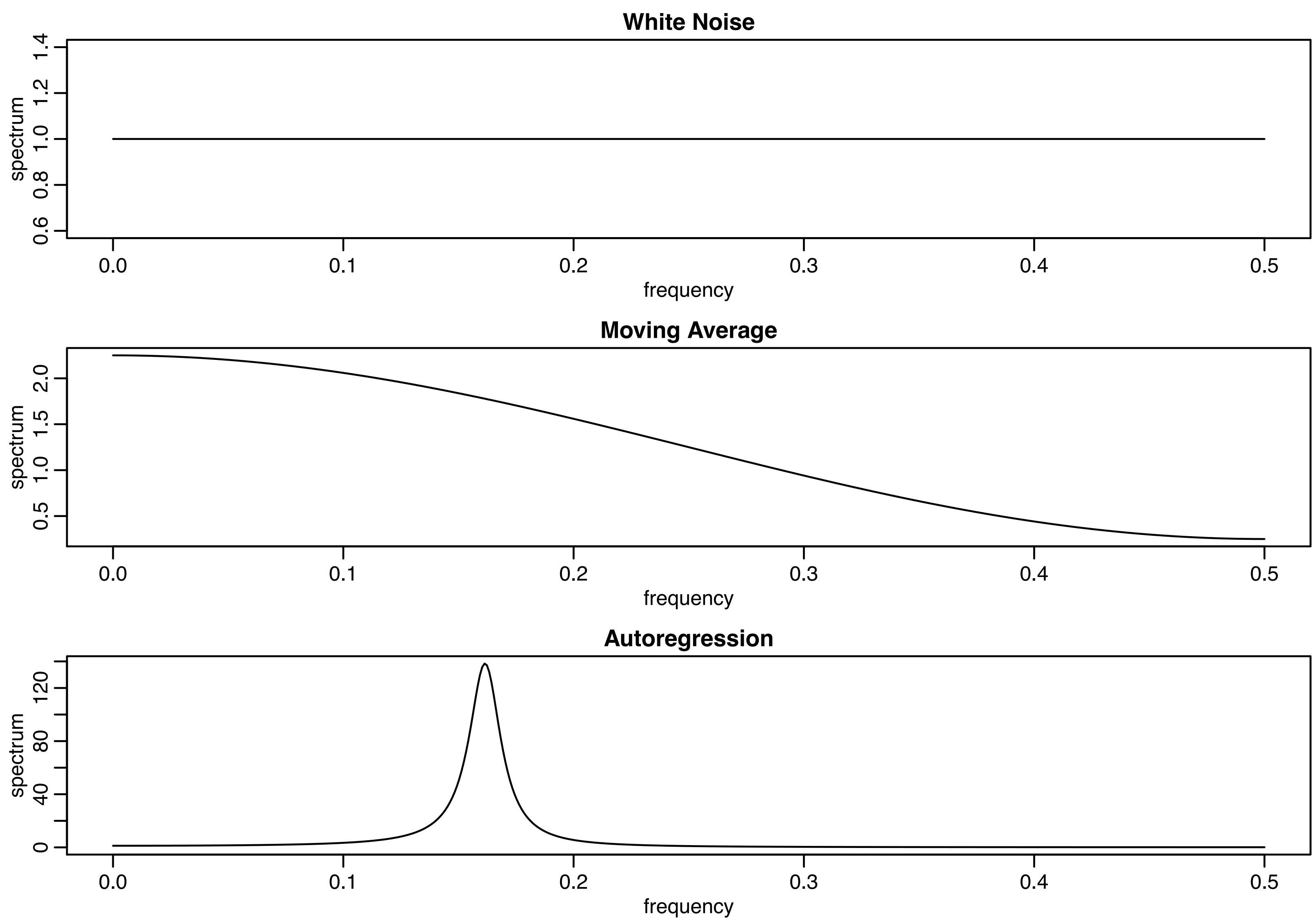
as the inverse transform of the spectral density,

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h} \quad -1/2 \leq \omega \leq 1/2. \tag{4.17}$$

One-to-one map covariance function and f, invertible

Spectral analogue of a probability density

Example: white noise $f_w(\omega) = \sigma_w^2$ **flat spectrum**



check at home for AR(1), MA(1)

Linear combinations, convolutions, filters

$$y_t = \sum_{j=-\infty}^{\infty} \boxed{a_j} x_{t-j}, \qquad \sum_{j=-\infty}^{\infty} |a_j| < \infty.$$

Impulse response

Fourier Transform

frequency response

$$A(\omega) = \sum_{j=-\infty}^{\infty} a_j \, \mathrm{e}^{-2\pi i \omega j},$$

spectral density

$$f_y(\omega) = |A(\omega)|^2 \, f_x(\omega),$$

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j},$$

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{e}^{2\pi i\omega h} \, f(\omega) \, d\omega$$

Spectral analysis as PCA

$$\mathrm{cov}(X) = \mathit{\Gamma}_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix}.$$

For n sufficiently large, the eigenvalues of $\mathit{\Gamma}_n$ are

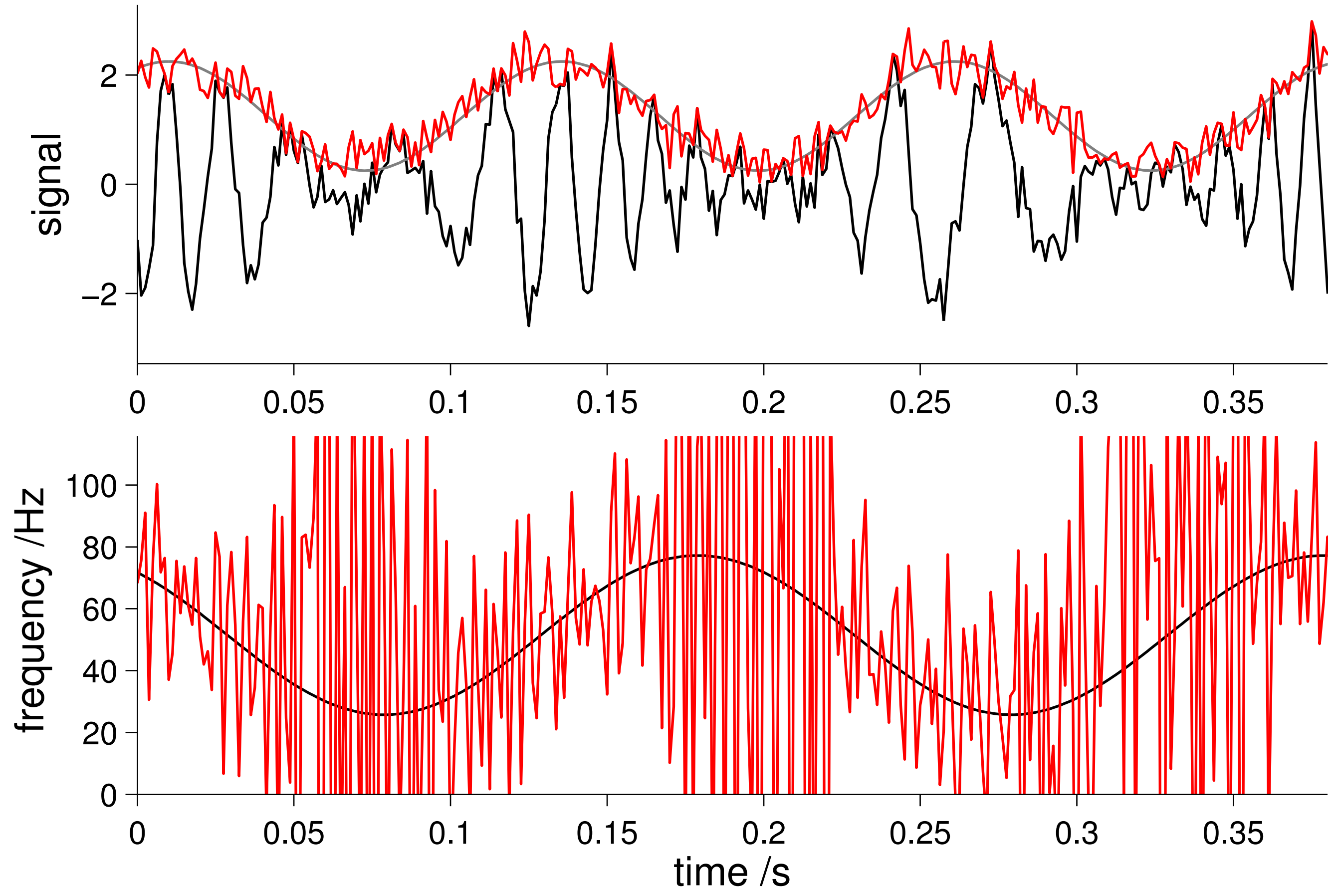
$$\lambda_j \approx f(\omega_j) = \sum_{h=-\infty}^{\infty} \gamma(h)\, \mathrm{e}^{-2\pi i h j/n},$$

A Bayesian perspective

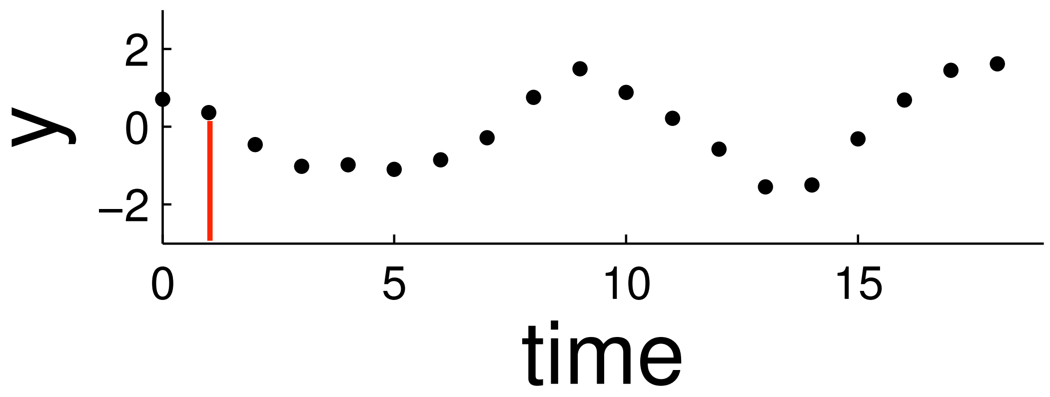
Why bayesian:

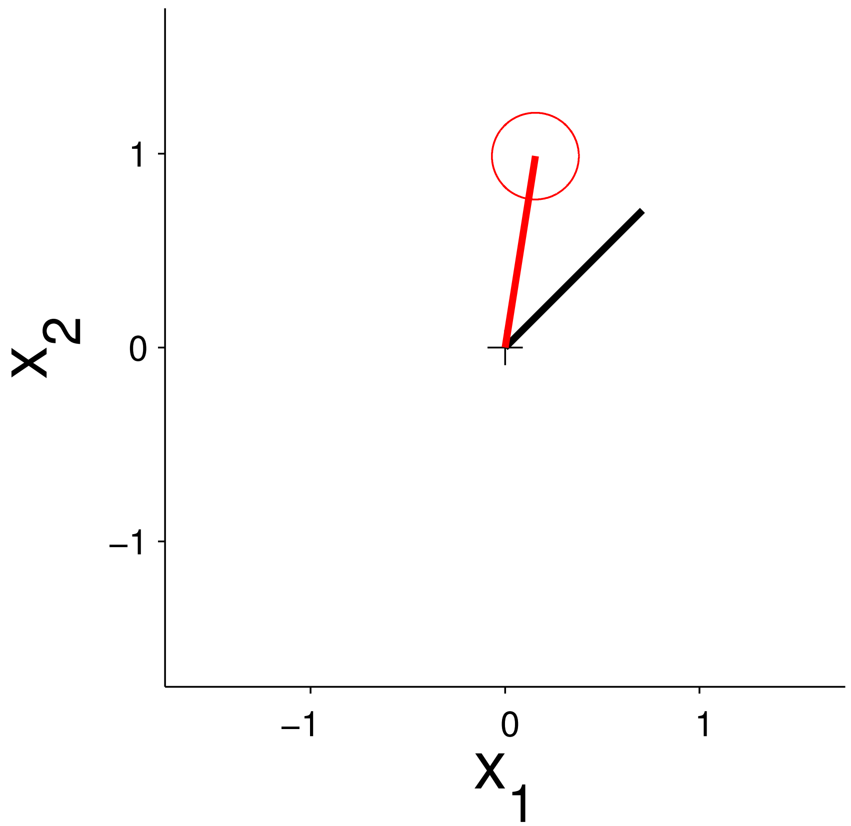
No explicit notion of measurement noise

No way to deal with missing data



A Bayesian perspective





$$\mathrm{y}(t) = \Re\left(\mathrm{a}(t)\exp(i\phi(t))\right)$$

$$\mathrm{y}_t = [1,0]\mathrm{x}_t + \sigma_{\mathrm{y}}\eta_t$$

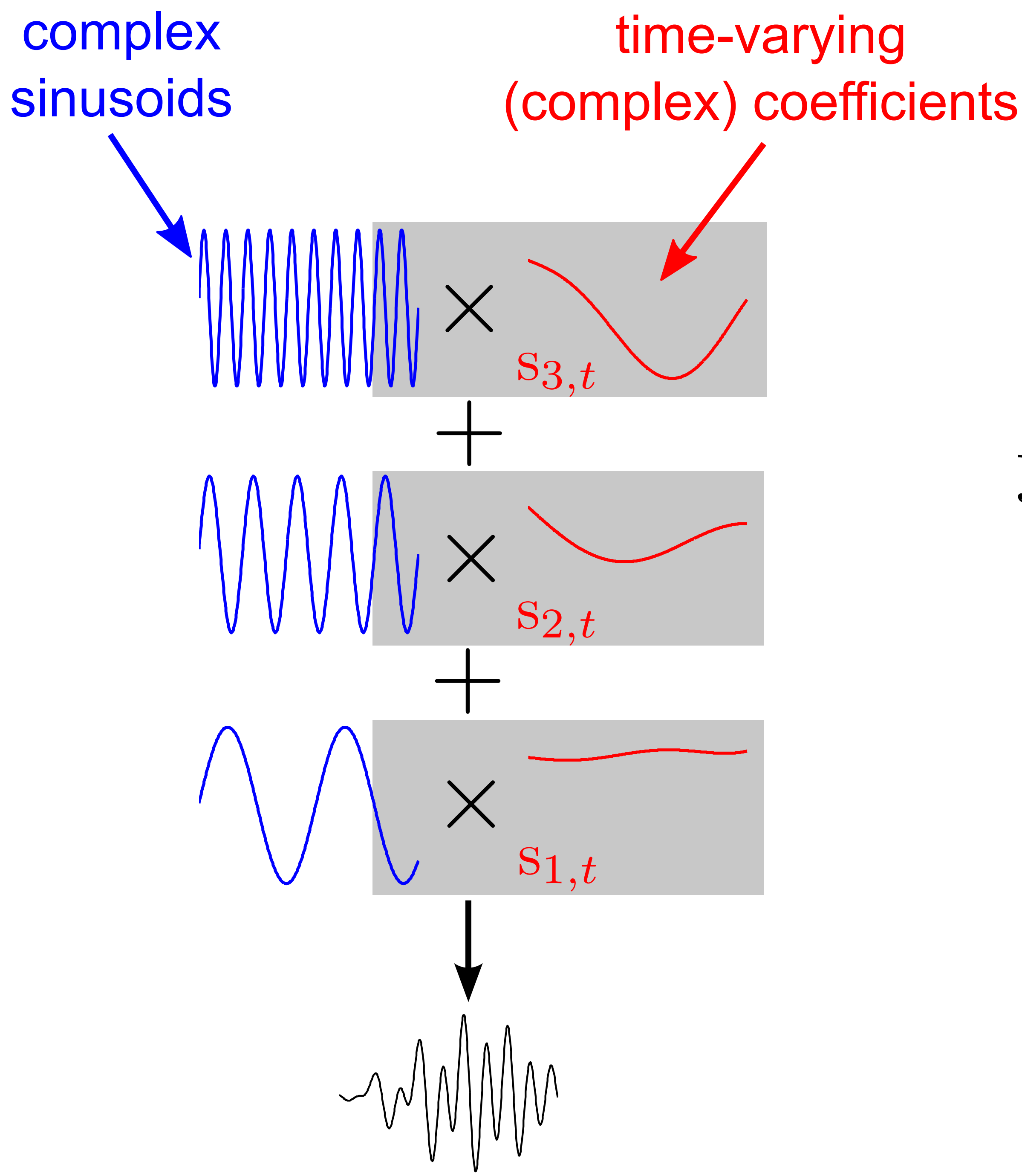
$$\mathrm{x}_t = \lambda \mathrm{R}(\bar{\omega})\mathrm{x}_{t-1} + \sigma_{\mathrm{x}}^2 \boldsymbol{\epsilon}_t$$

Just an extended Kalman filter

Rotation matrix: $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Qi and Minka 2002, Cemgil and Godshill 2005.

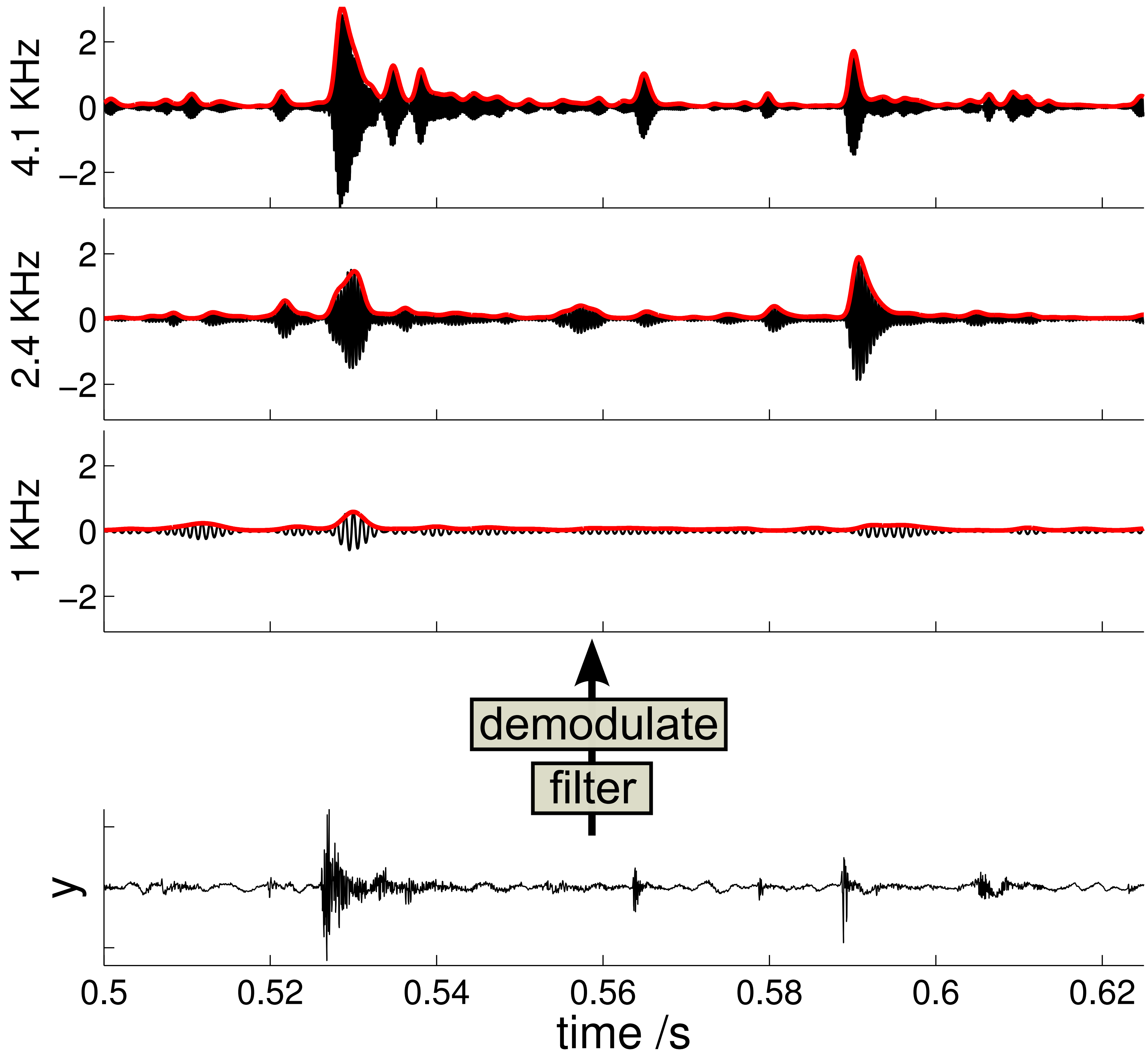
Generative model in complex domain



$$y_t = \sum_d \Re\left(e^{i\omega_d t} s_{d,t}\right) + \sigma_y \eta_t$$

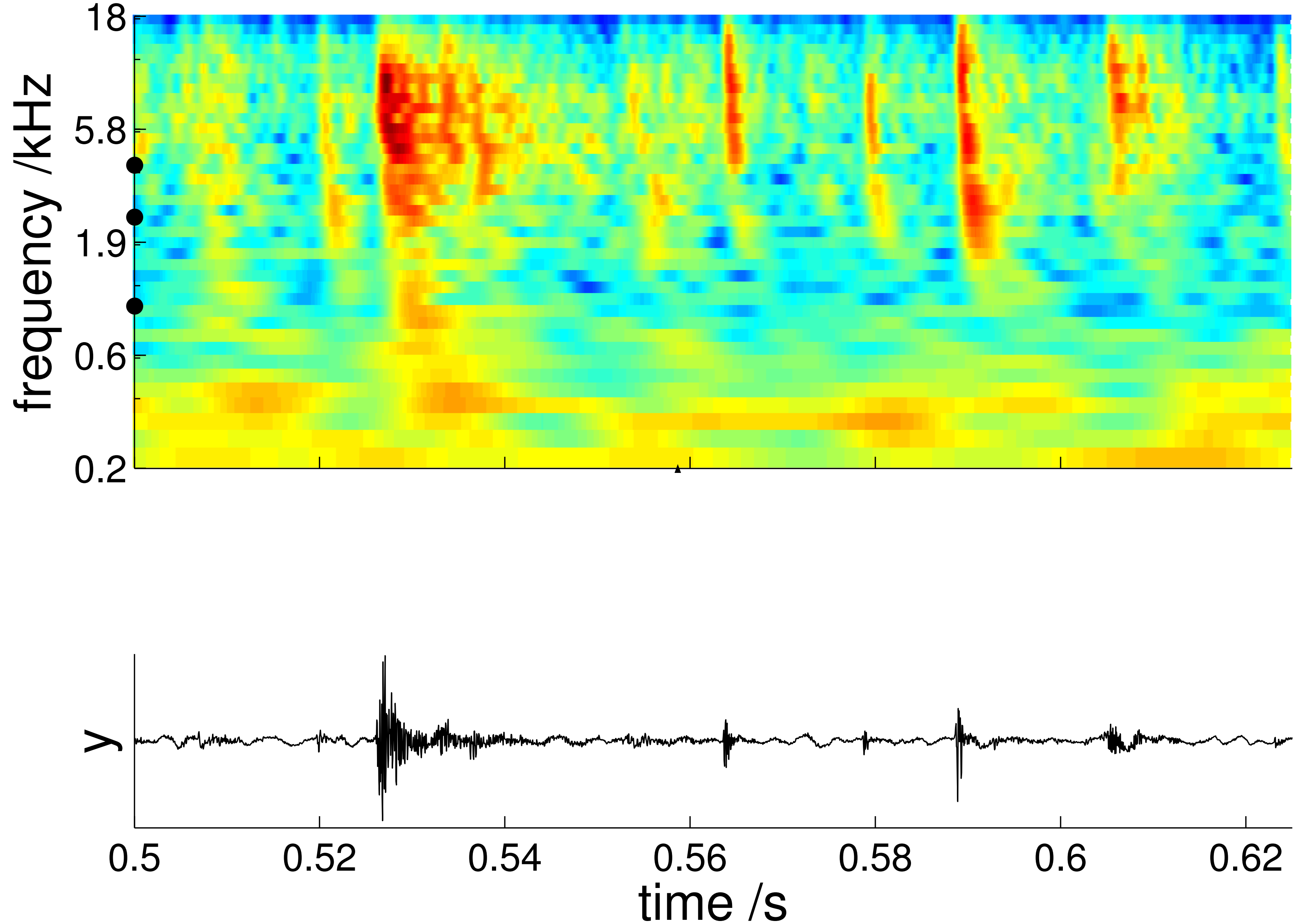
$$\Re\left(s_{d,t}\right) \sim \mathcal{GP}(0, \Gamma)$$

$$\Im\left(s_{d,t}\right) \sim \mathcal{GP}(0, \Gamma)$$

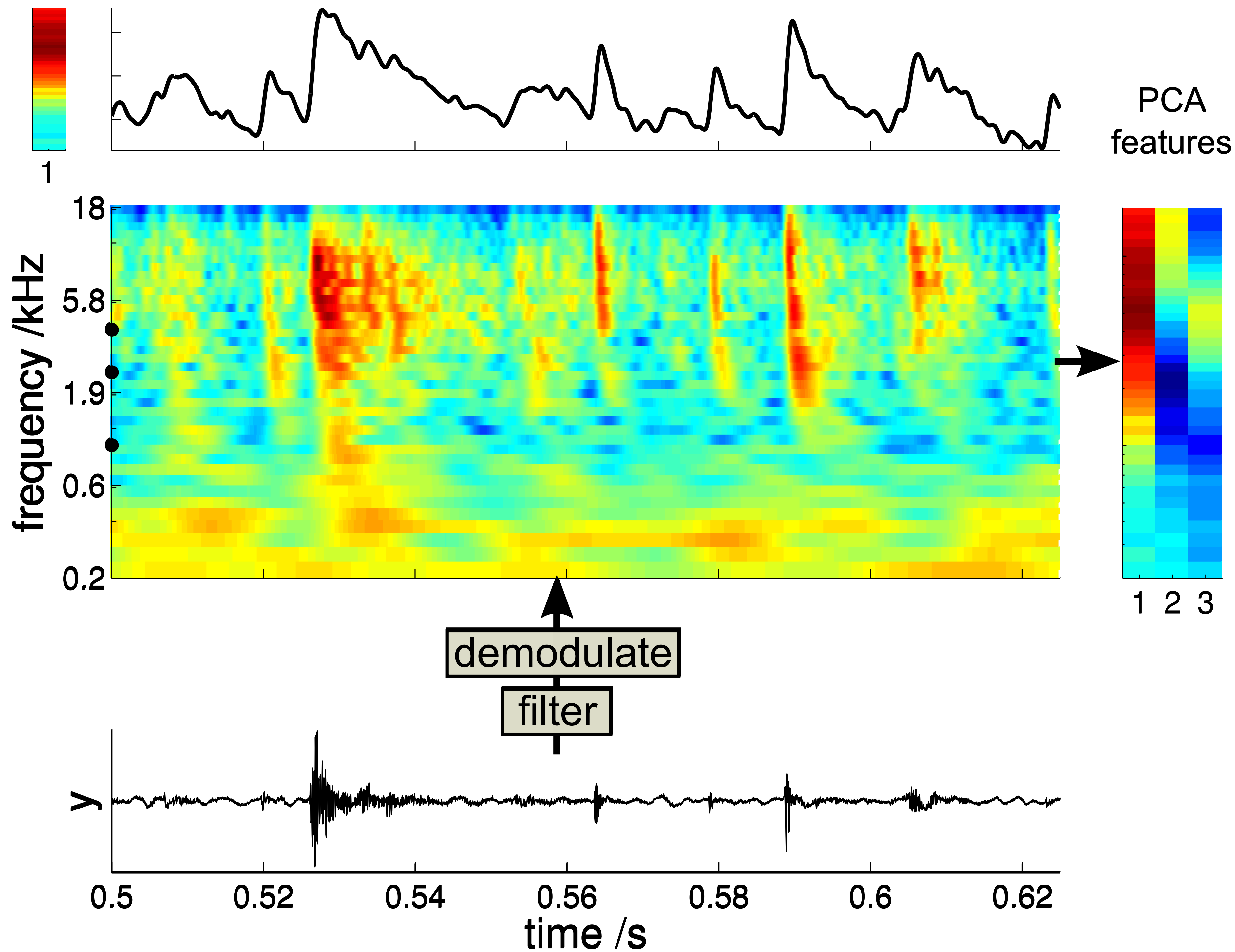


4.1 KHz
2.4 KHz
1 KHz
2
0
−2
demodulate
filter
y
0.5
0.52
0.54
0.56
0.58
0.6
0.62
time /s

The interesting cases are not stationary

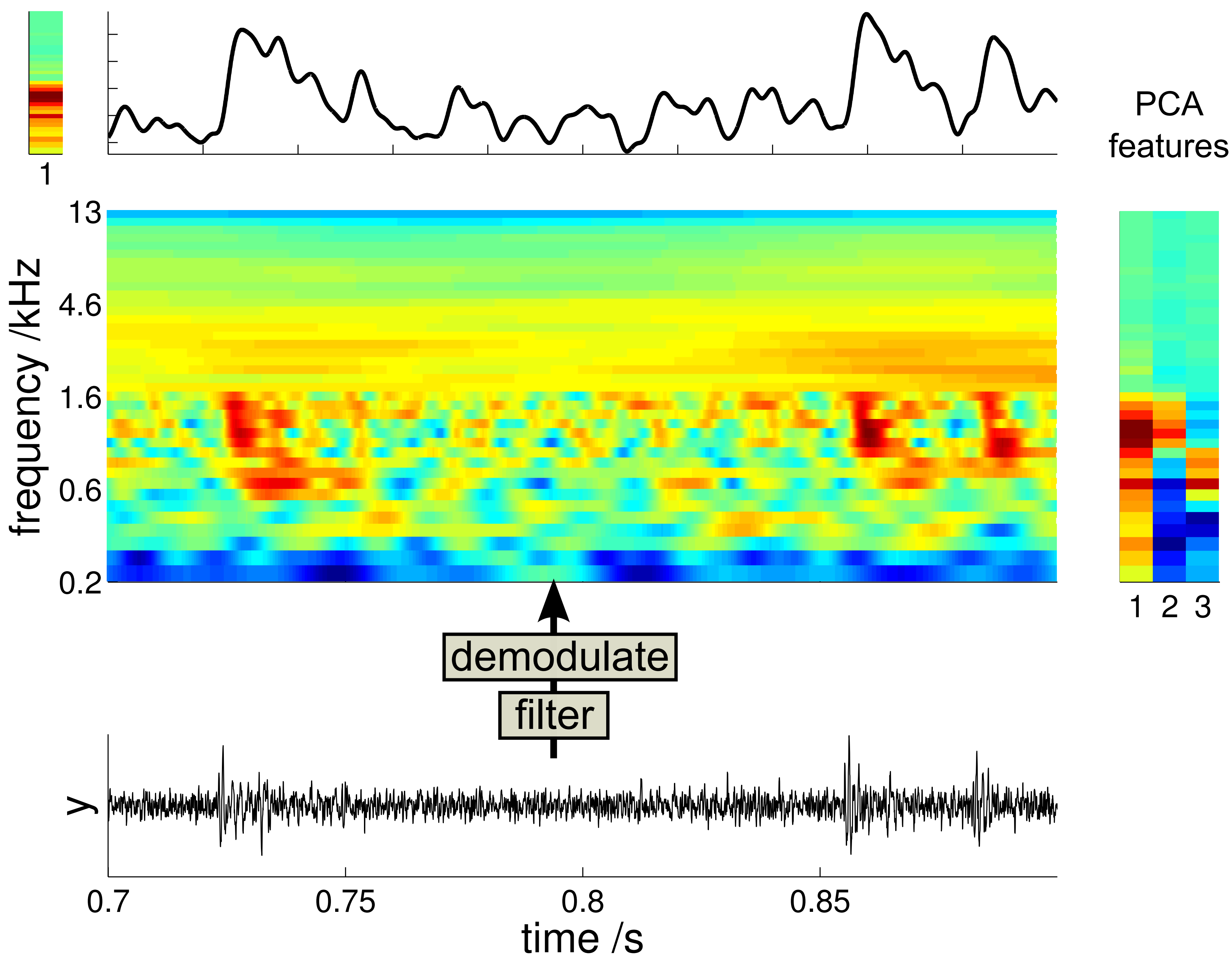


Fire



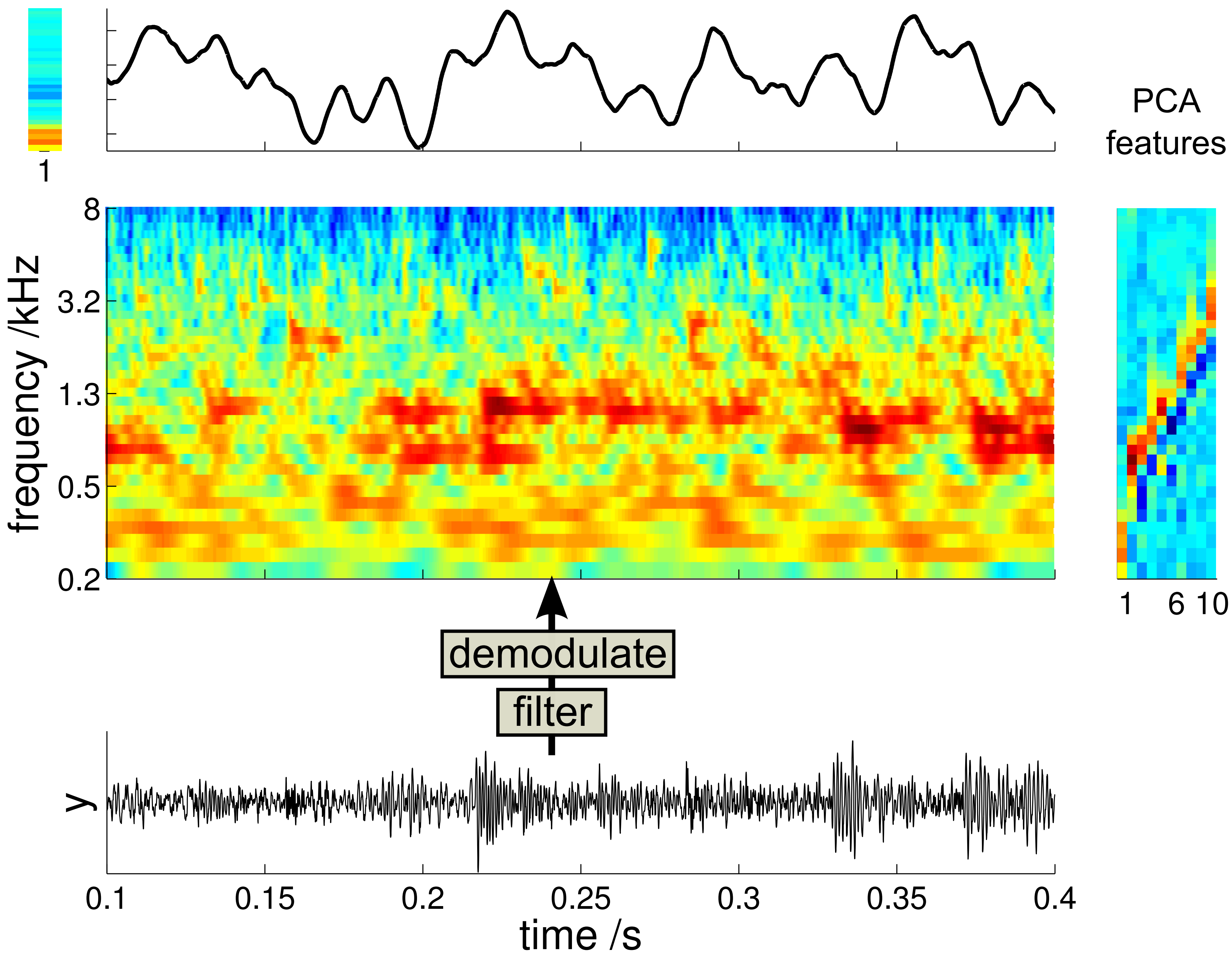
PCA
features
1
18
5.8
1.9
0.6
0.2
frequency /kHz
1
2
3
demodulate
filter
y
0.5
0.52
0.54
0.56
0.58
0.6
0.62
time /s

Wind



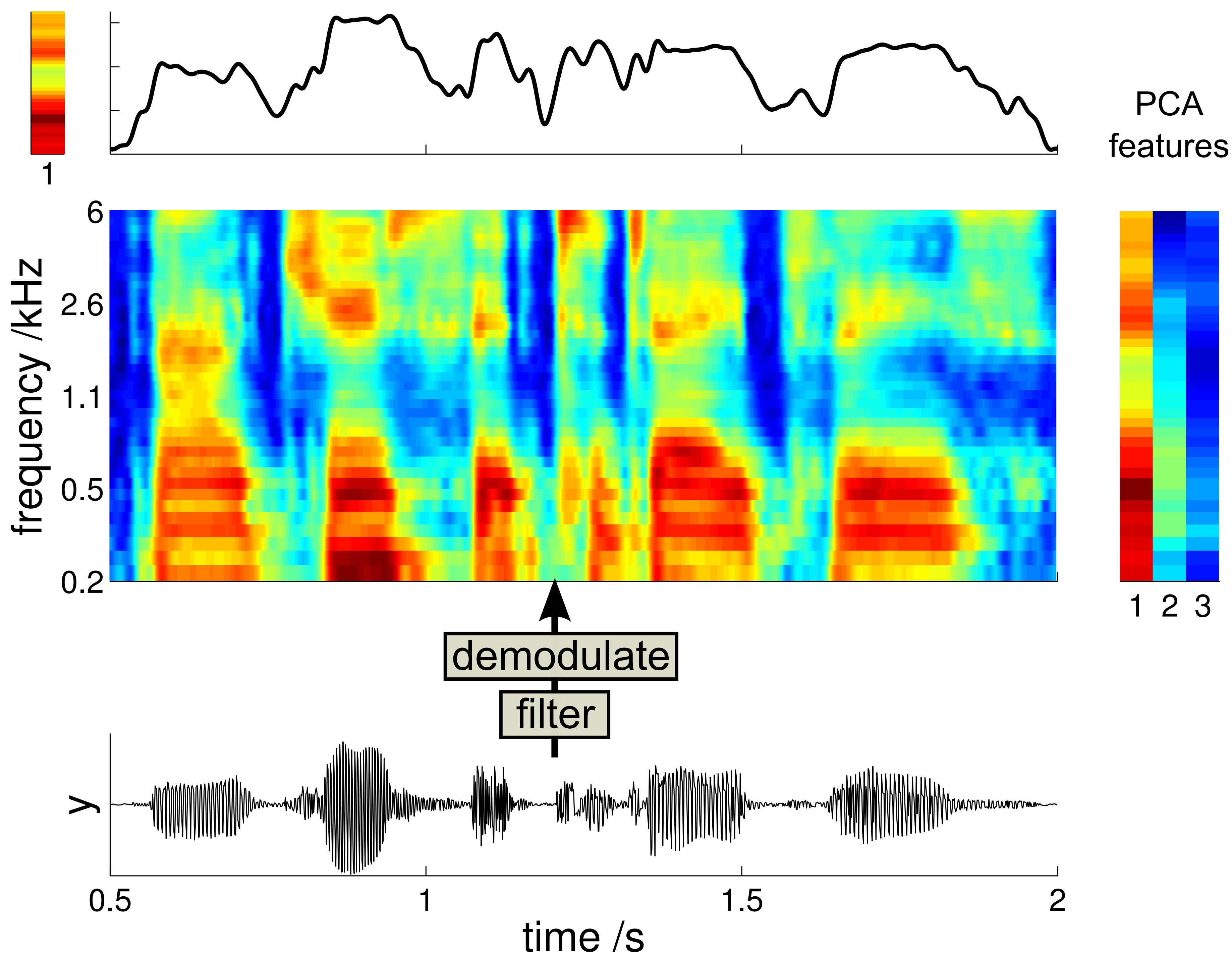
1
PCA
features
13
4.6
1.6
0.6
0.2
frequency /kHz
1 2 3
demodulate
filter
y
0.7
0.75
0.8
0.85
time /s

Water



1
PCA
features
8
3.2
1.3
0.5
0.2
frequency /kHz
1
6
10
demodulate
filter
y
0.1
0.15
0.2
0.25
0.3
0.35
0.4
time /s

Speech



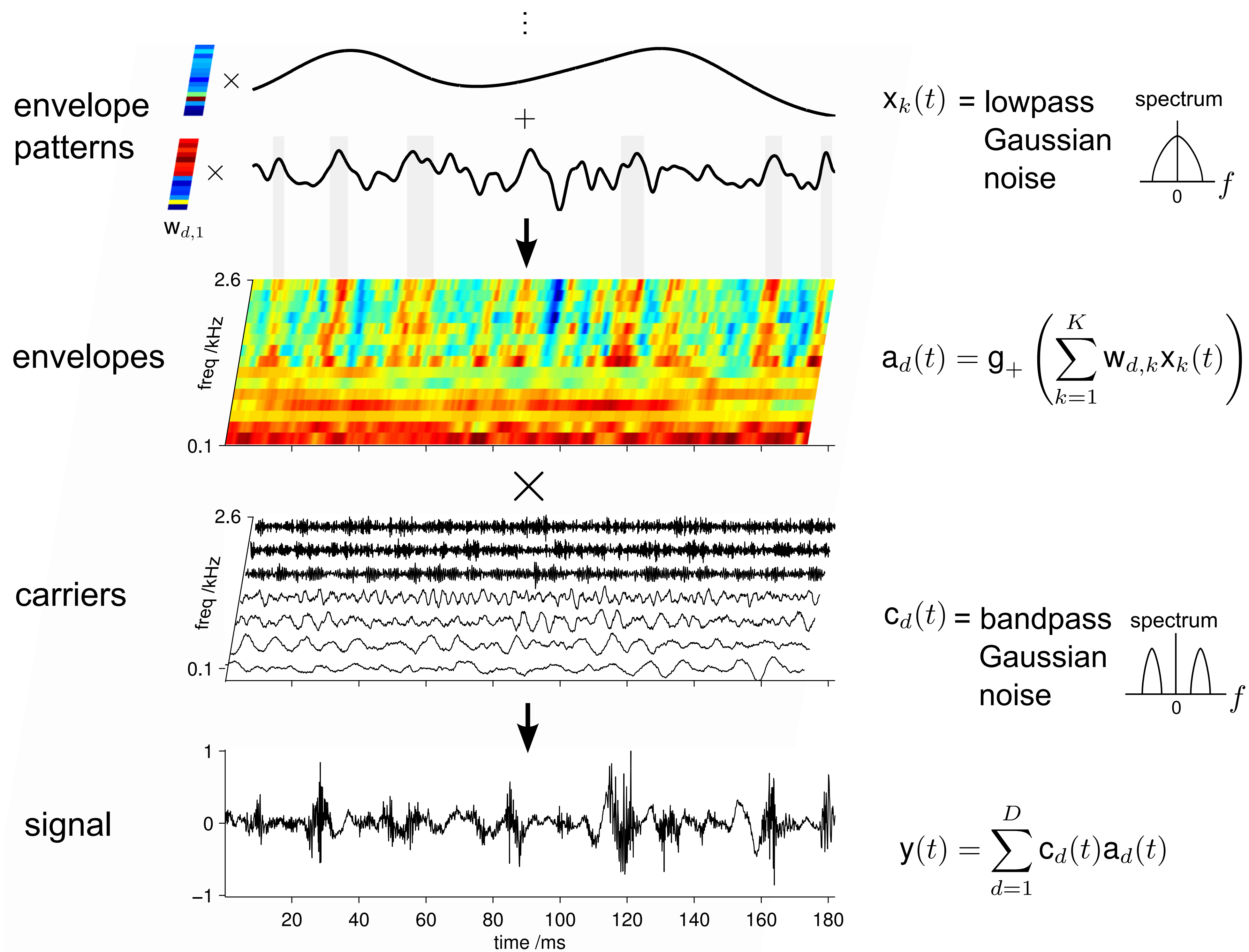
1
PCA
features
6
2.6
1.1
0.5
0.2
frequency /kHz
1 2 3
demodulate
filter
y
0.5
1
1.5
2
time /s

Spectral analysis reveals a lot of potentially useful statistical regularities in real-world data

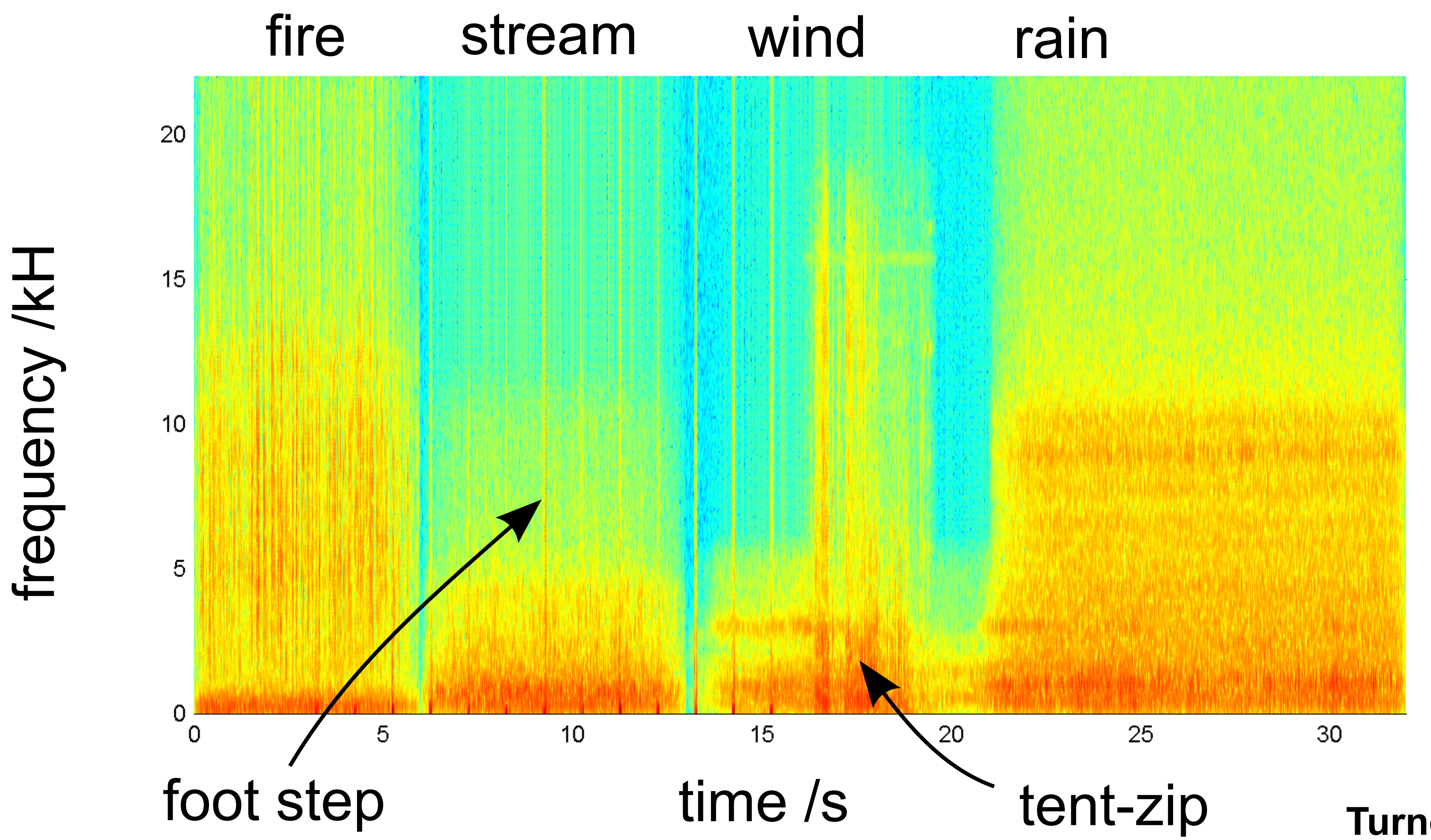
energy in sub-bands (power-spectrum)
patterns of co-modulation
time-scale of the modulation
depth of the modulation (sparsity)

We can build probabilistic models that capture such phenomena and fit them to data

Statistical Model

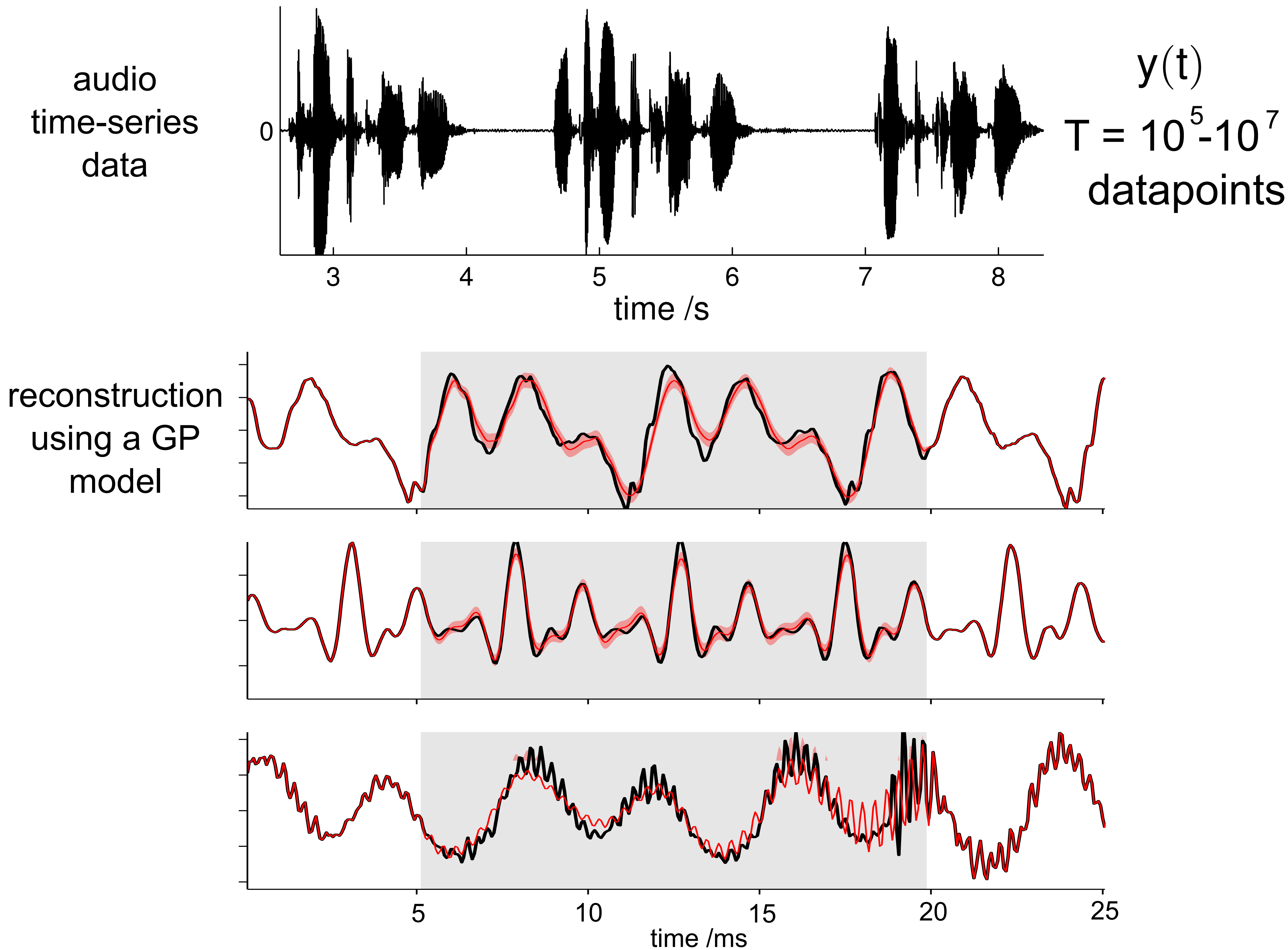


Sound Generation Demo



Turner, 2010

Application: speech imputation



robustness
adaptation

Probabilistic signal processing → **Classical signal processing**

Classical signal processing → **Probabilistic signal processing**

fast methods
important variables

extra reading:

Turner, R. E. and Sahani, M.
"Time-frequency analysis as probabilistic inference”
IEEE Transactions on Signal Processing. 2014

Turner, R. and Sahani M. "Probabilistic amplitude and frequency demodulation.”
Advances in Neural Information Processing Systems. 2011.

Turner, R. E., and Sahani, M.
"Demodulation as probabilistic inference."
IEEE Transactions on Audio, Speech, and Language Processing 19.8 (2011): 2398-2411.