PSYCHOACOUSTICALLY MOTIVATED AND BIOINSPIRED METHODS OF DIGITAL SIGNAL PROCESSING

PhD. **Vashkevich Maxim**

(vashkevich@bsuir.by)

Belarussian state university of informatics and radioelectronics (BSUIR)

Department of computer engineering

DSPA Conference

Digital Signal Processing and Its Applications

Introduction (a bit of history)

Thanks to the organizers for the invitation.

My scientific advisor: Petrovsky Alexander (1953-2019)

Википедия
Свободная энциклопедия

Заглавная страница
Рубрикация
Указатель А — Я
Избранные статьи
Случайная страница
Текущие события

Участие

Сообщить об ошибке
Сообщество
Форум
Свежие правки
Новые страницы
Справка
Пожертвовать

Инструменты

Ссылки сюда
Связанные правки
Служебные страницы
Постоянная ссылка
Сведения о странице
Цитировать страницу

Печать/экспорт

Создать книгу
Скачать как PDF
Версия для печати

В других проектах

Викисклад
Элемент Викиданных

Статья Обсуждение Читать Текущая версия Править Править код История Искать в Википедии

Петровский, Александр Александрович

Материал из Википедии — свободной энциклопедии [править | править код]

Текущая версия страницы пока не проверялась опытными участниками и может значительно отличаться от версии, проверенной 25 апреля 2019; проверки требуют 6 правок.

Петровский Александр Александрович (1 января 1953 д. Снов, БССР — 14 марта, 2019, Минск) — советский и белорусский ученый в области вычислительной техники и цифровой обработки сигналов, профессор, доктор технических наук.

Содержание [скрыть]

Петровский Александр Александрович

Дата рождения	1 января 1953
Место рождения	д. Снов, БССР
Дата смерти	14 марта 2019 (66 лет)

Биография [править | править код]

А. А. Петровский окончил Сновскую среднюю школу в 1970 году[1] и поступил учиться в Минский радиотехнический институт (МРТИ), который окончил с отличием в 1975 г. по специальности «Электронные вычислительные машины» (ЭВМ)[2]. Трудовую деятельность начал в МРТИ в должности младшего научного сотрудника[3]. Затем работал ассистентом на кафедре ЭВМ. В 1977 г. начал обучение в аспирантуре МРТИ и спустя три года защитил кандидатскую диссертацию под руководством Александра Ефимовича Леусенко. В течение следующих 9 лет (1981—1990 гг.) занимает должность доцента кафедры ЭВМ[4]. В 1985 г. принимает руководство исследовательской лабораторией микропроцессорных систем реального времени (МРТИ). В 1989 г. защитил докторскую диссертацию на тему «Построение микропроцессорных систем обработки виброакустических процессов в реальном времени» (защита проходила в Институте проблем моделирования в энергетике им. Г. Е. Пухова, г. Киев). В 1990 г. избран заведующим кафедрой электронных вычислительных средств БГУИР, которой руководил до 2017 года[5]

Aim of this talk

To give an idea of the existing **bioinspired** and **psychoacoustically motivated** methods of audio signal analysis, and to show the areas of its application.

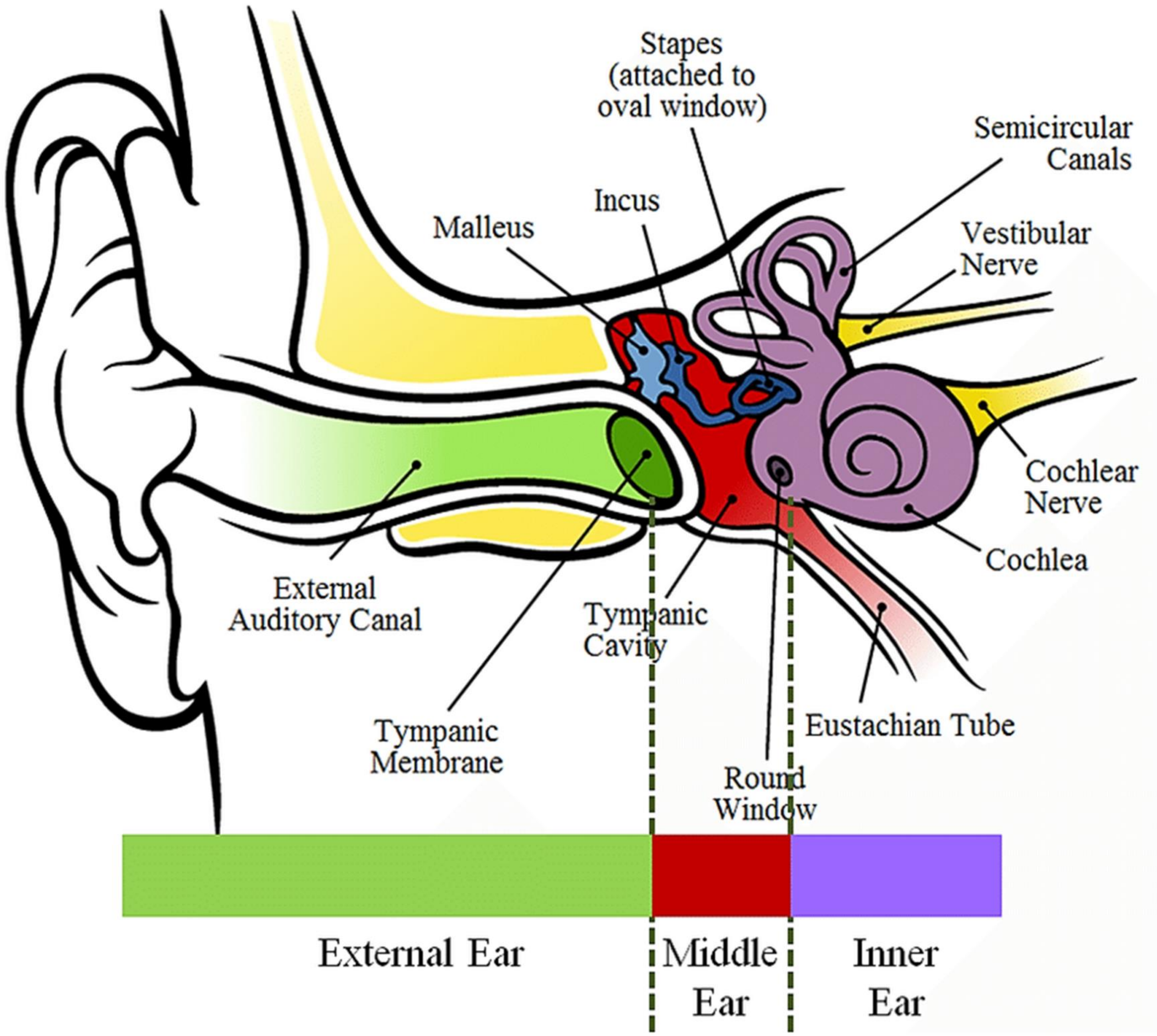
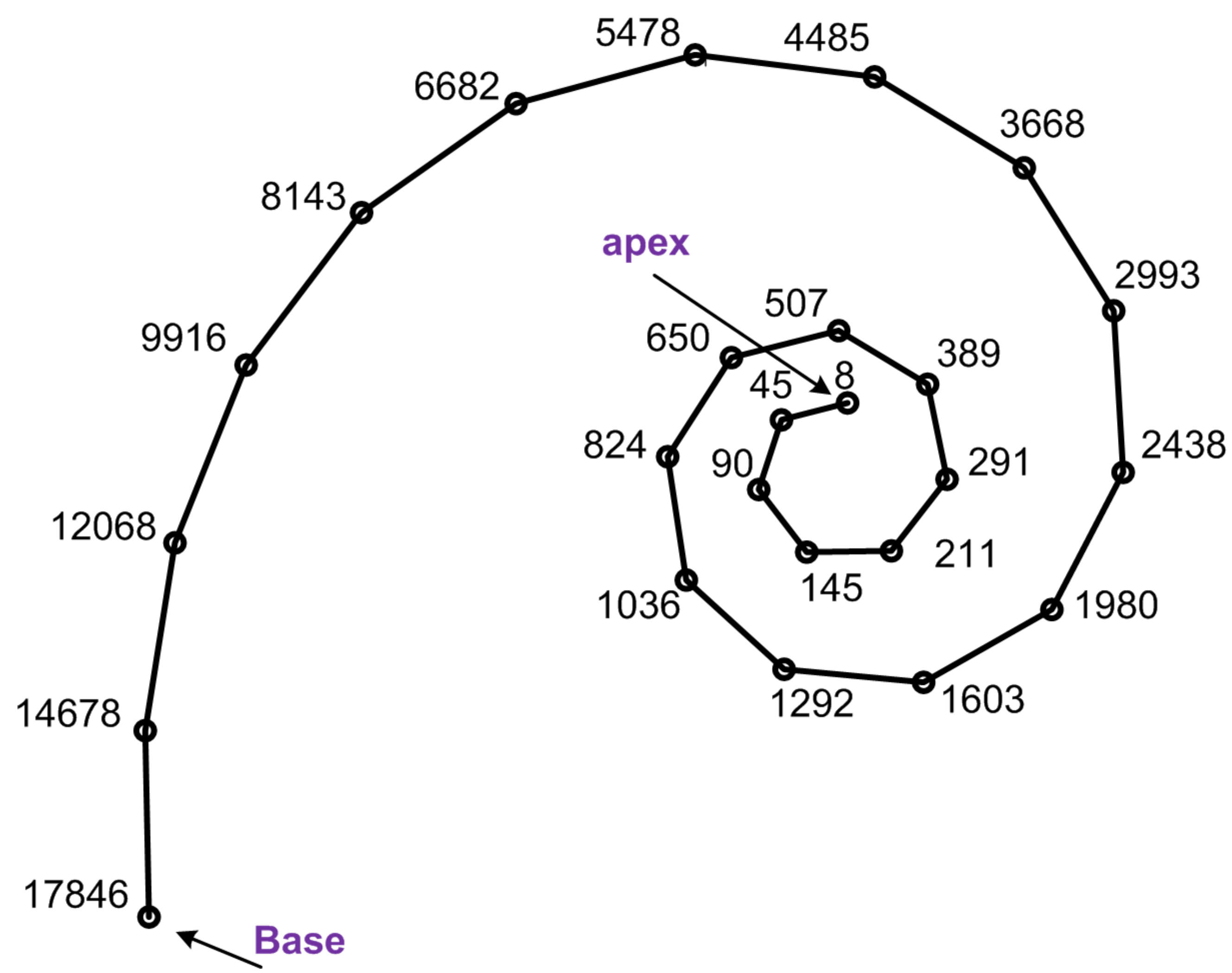


Diagram of basilar membrane. The positions of the maximum displacement are shown when stimulated with sinusoids with different frequencies (in Hz).



Two groups of methods

1) ***Bioinspired methods*** – aimed to modeling the work of the parts of the auditory system. Methods are based on parametric expressions fitted to the physiological data.

Examples of the **physiological data**:

– characteristics of the *excitation of the auditory nerve;*

– measured *displacement of the basilar membrane*;

2) ***Psychoacoustically motivated methods*** are based on patterns obtained in the study of auditory perceptions of sounds.

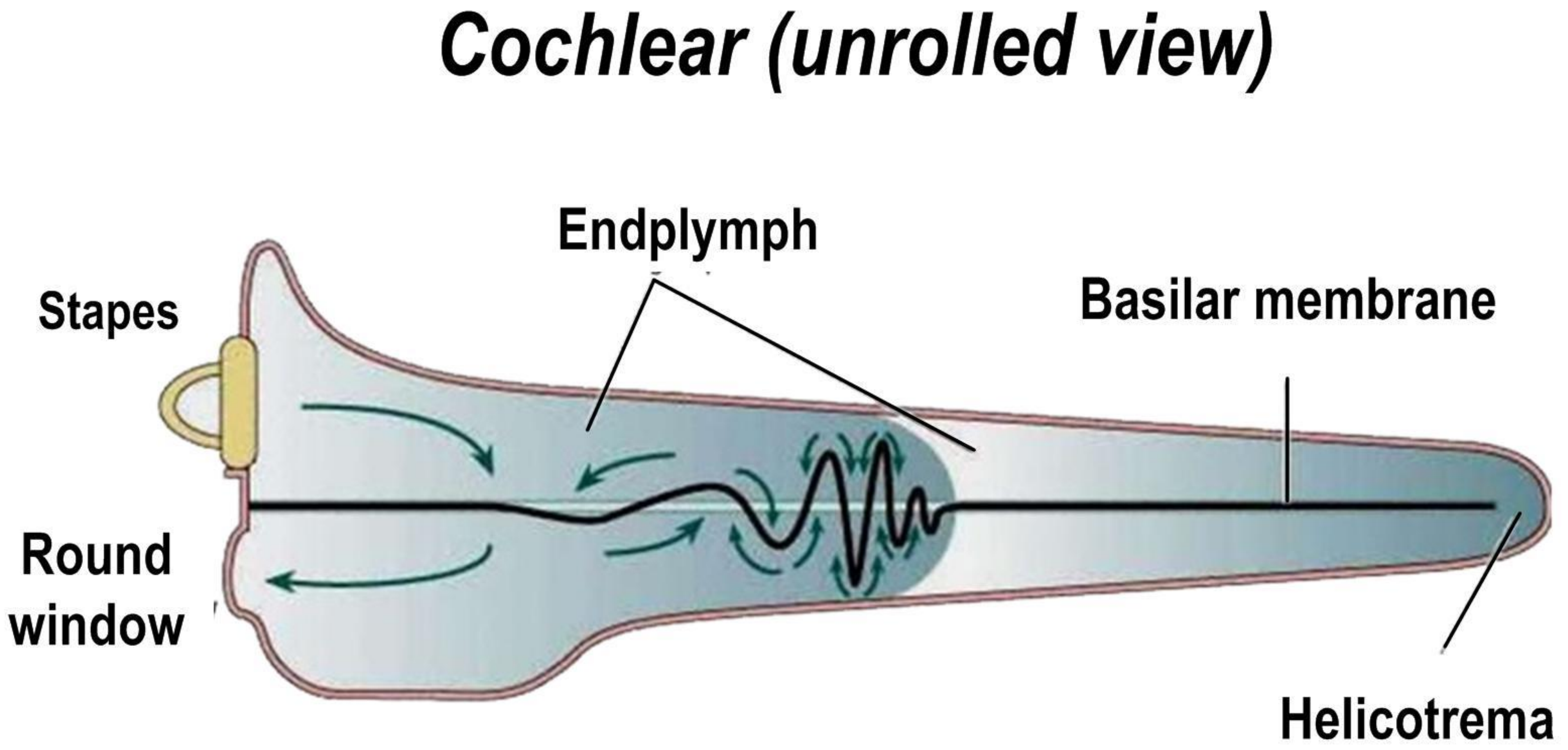
Examples of psychoacoustic patterns:

– *critical bands*;

– *pitch* and *loudness* of sound.

Auditory filters

The basilar membrane contains ~ 3000 internal hair cells. Each cell is responsible for a specific frequency range.



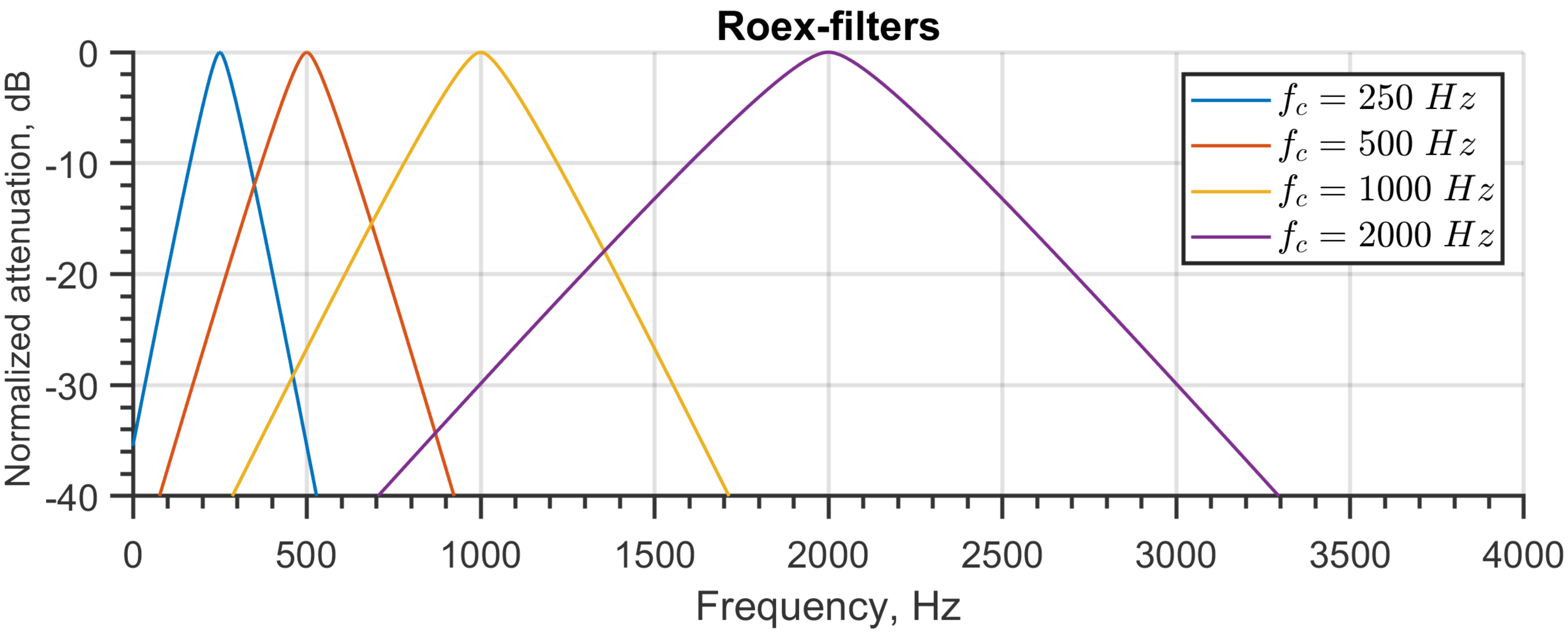
The mechanical movement of the cochlea basilar membrane is modeled using a ***bank of digital filters***. These filters are called *cochlear* or *auditory* filters.

Bioinspired auditory filters: roex-filters

Roex-filters[1] (**ro**und **ex**ponential) are describes the auditory filters in the frequency domain using round exponential functions:

$$|H(f)|^2 = (1-r)(1+pg)e^{-pg} + r, \qquad g = |f - f_c|/f_c.$$

where g – normalized deviation of frequency f from central frequency f_c.



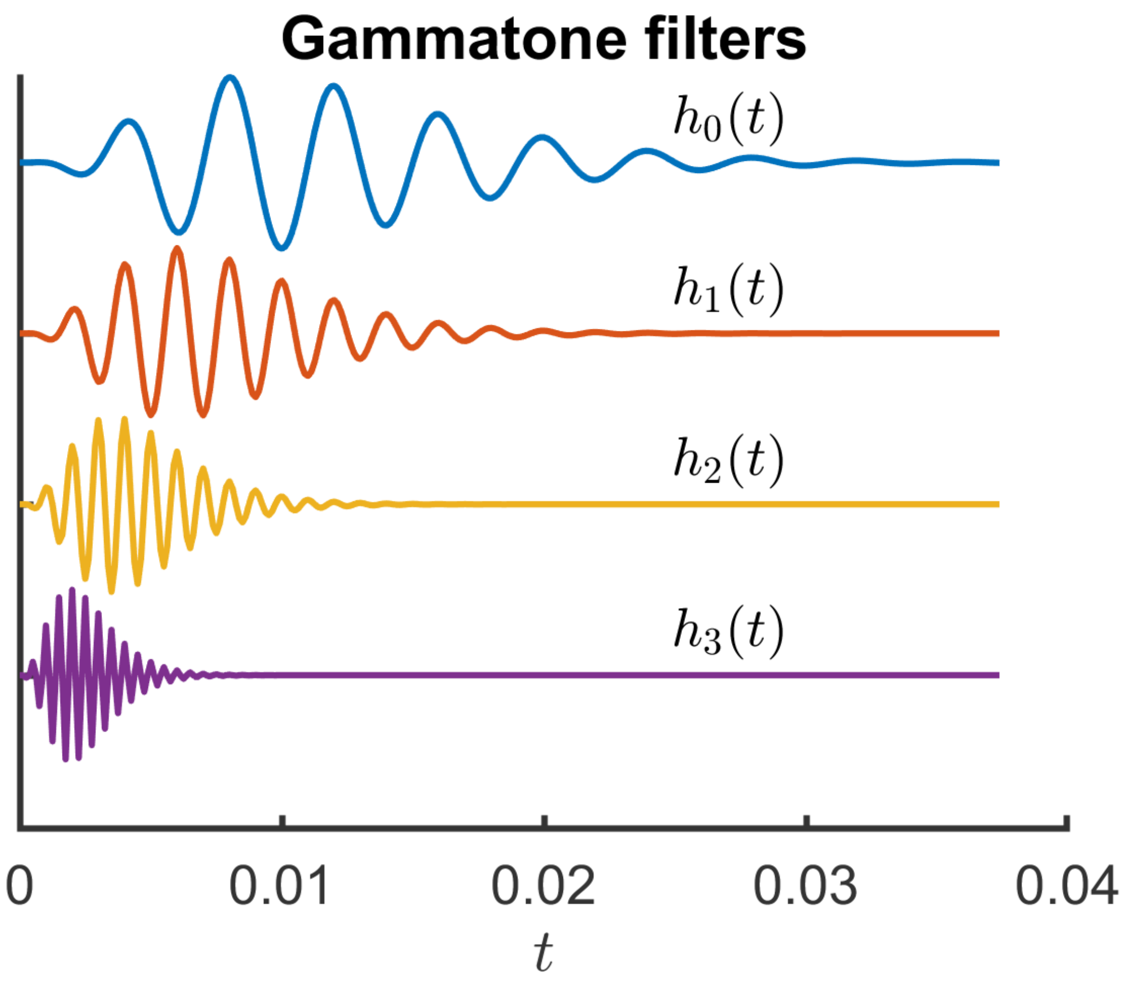
[1] The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold / R. D. Patterson, et al. // The Journal of the Acoustical Society of America. – 1982. – vol. 72, №. 6. – P. 1788–1803.

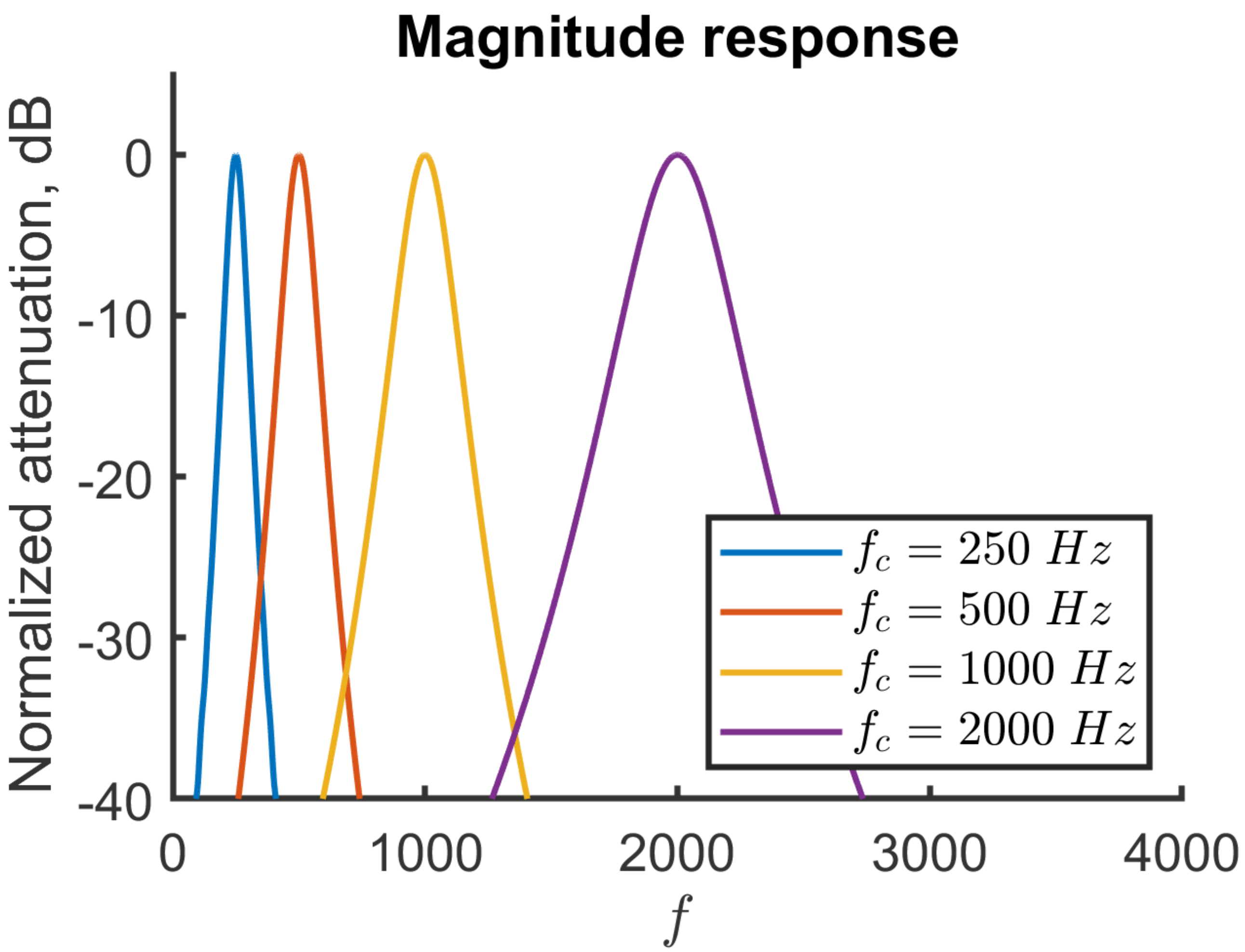
Bioinspired auditory filters: gammatone-filters

Gammatone filters[2] – description of the auditory filter in time domain:

$$h(t) = t^{(l-1)} e^{-2\pi b \mathrm{ERB}(f_c)} \cos 2\pi f_c t, \qquad t > 0.$$

where b – bandwidth parameter, l – the order.





---

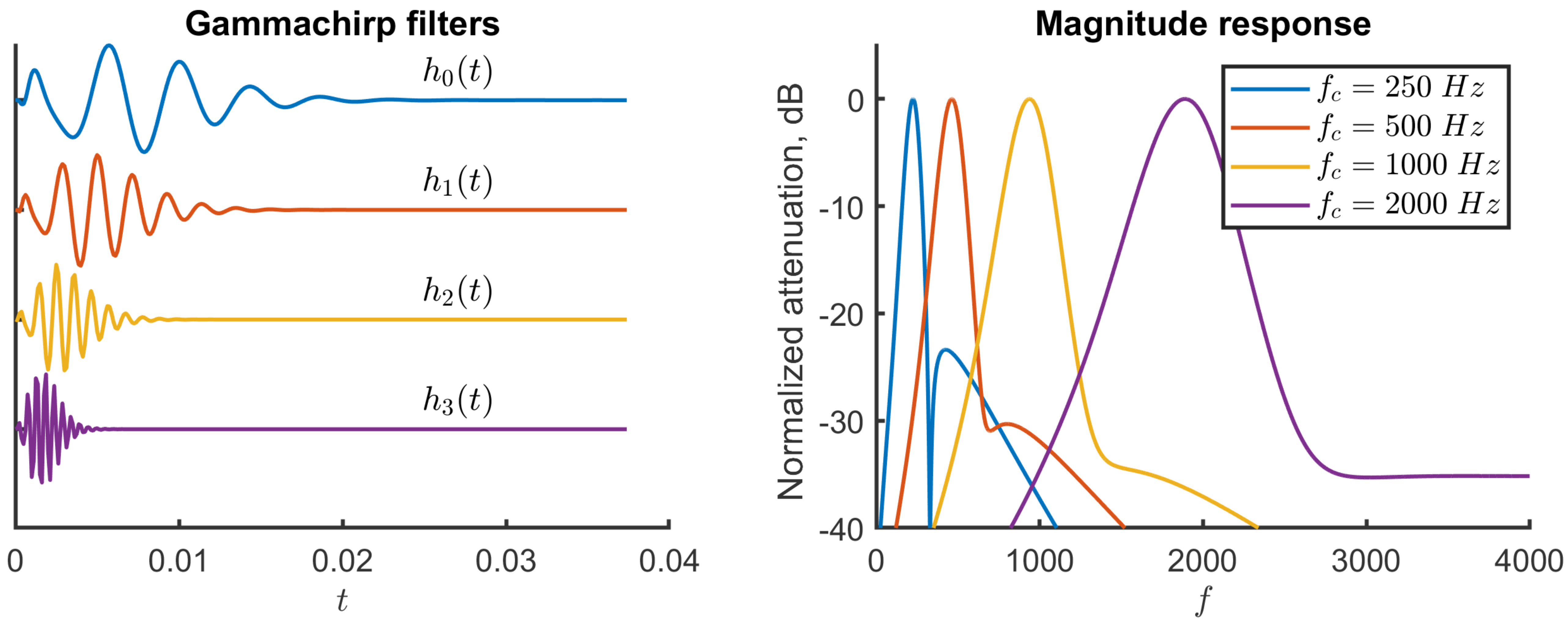
[2] An efficient auditory filterbank based on the gammatonefunction / R. D. Patterson, *et al.* // Proceeding of the IOC Speech Group on Auditory Modelling at RSRE. – 1987. – vol. 2, №. 7. – P. 1-34.

Bioinspired auditory filters: gammachirp filters

Gammachirp filters[3] – refined description of the auditory filter in the time domain:

$$h(t) = t^{(l-1)} e^{-2\pi b \mathrm{ERB}(f_r)} \cos(2\pi f_r t + c \ln t + \phi), \qquad t > 0.$$

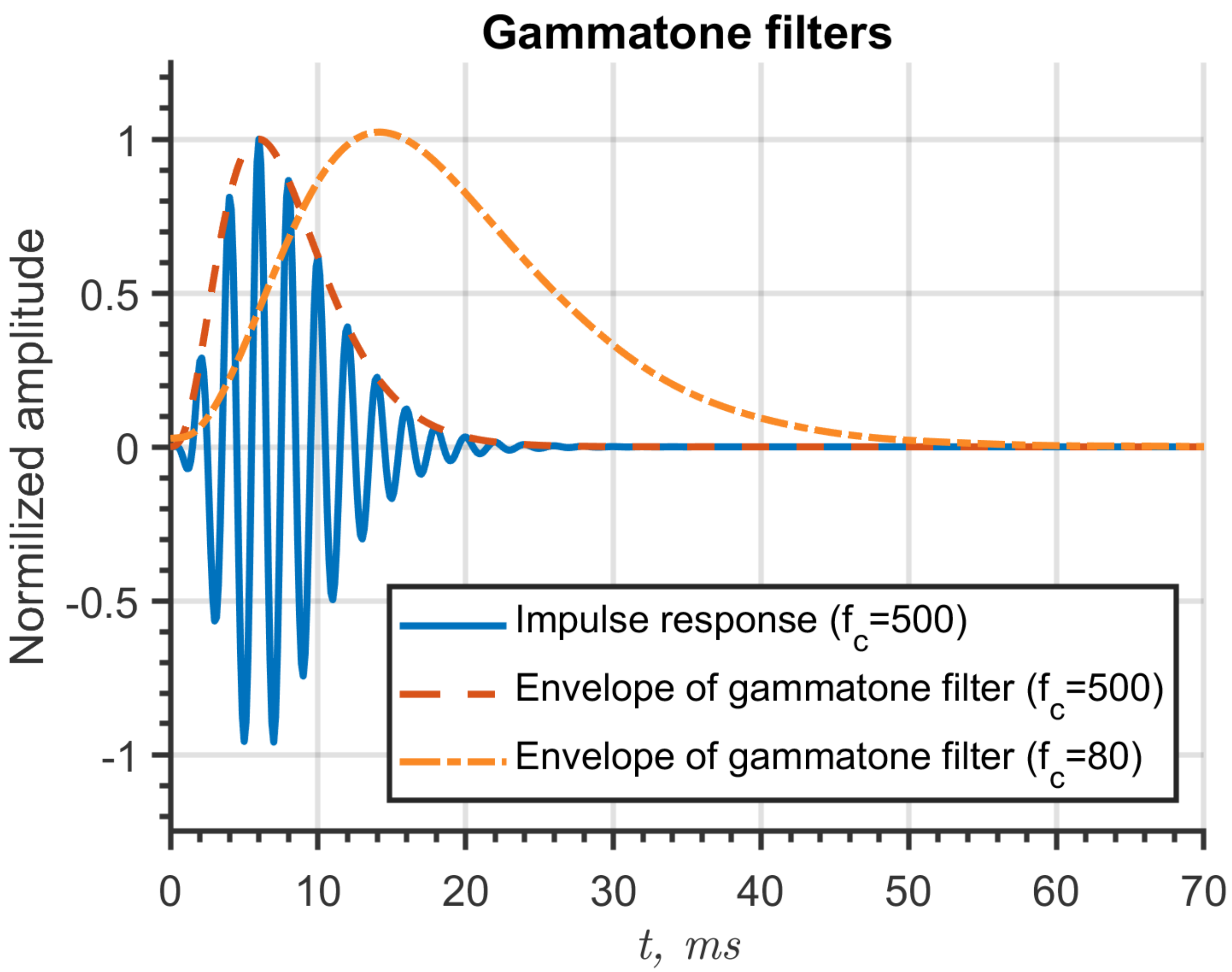
where f_r – pick frequency, c – additional parameter.



[3] Irino, T. A time-domain, level-dependent auditory filter: The gammachirp / T. Irino, R. D. Patterson // The Journal of the Acoustical Society of America. – 1997. – vol. 101, №. 1. – P. 412-419.

Practical implementation of auditory filters

The *discrete-time impulse responses* of the auditory filters are designed by sampling and windowing the continuous-time infinite-length impulse responses



Main problem: *impulse responses are very long*.

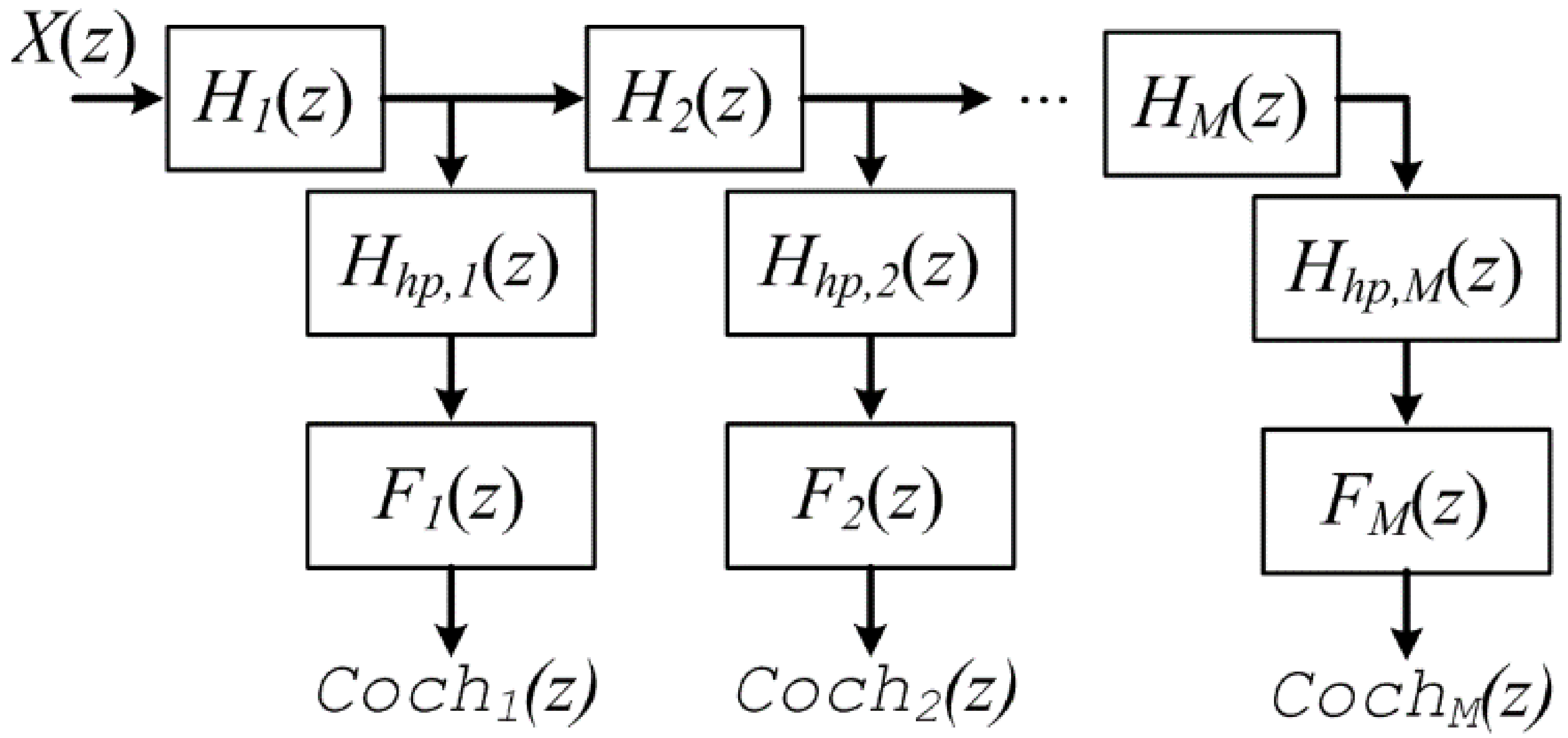
Example ($f_s = 8000$ Hz):

✓ for $\boldsymbol{f_c}$=**500 Hz** impulse response about **240 samples.**

✓ for $\boldsymbol{f_c}$=**80 Hz** impulse response about **600 samples.**

Digital time-domain model of the human cochlear (1)

J. Kates proposed[4] to model the propagation of a traveling wave on the cochlear partition using a cascade of second-order IIR digital filters.



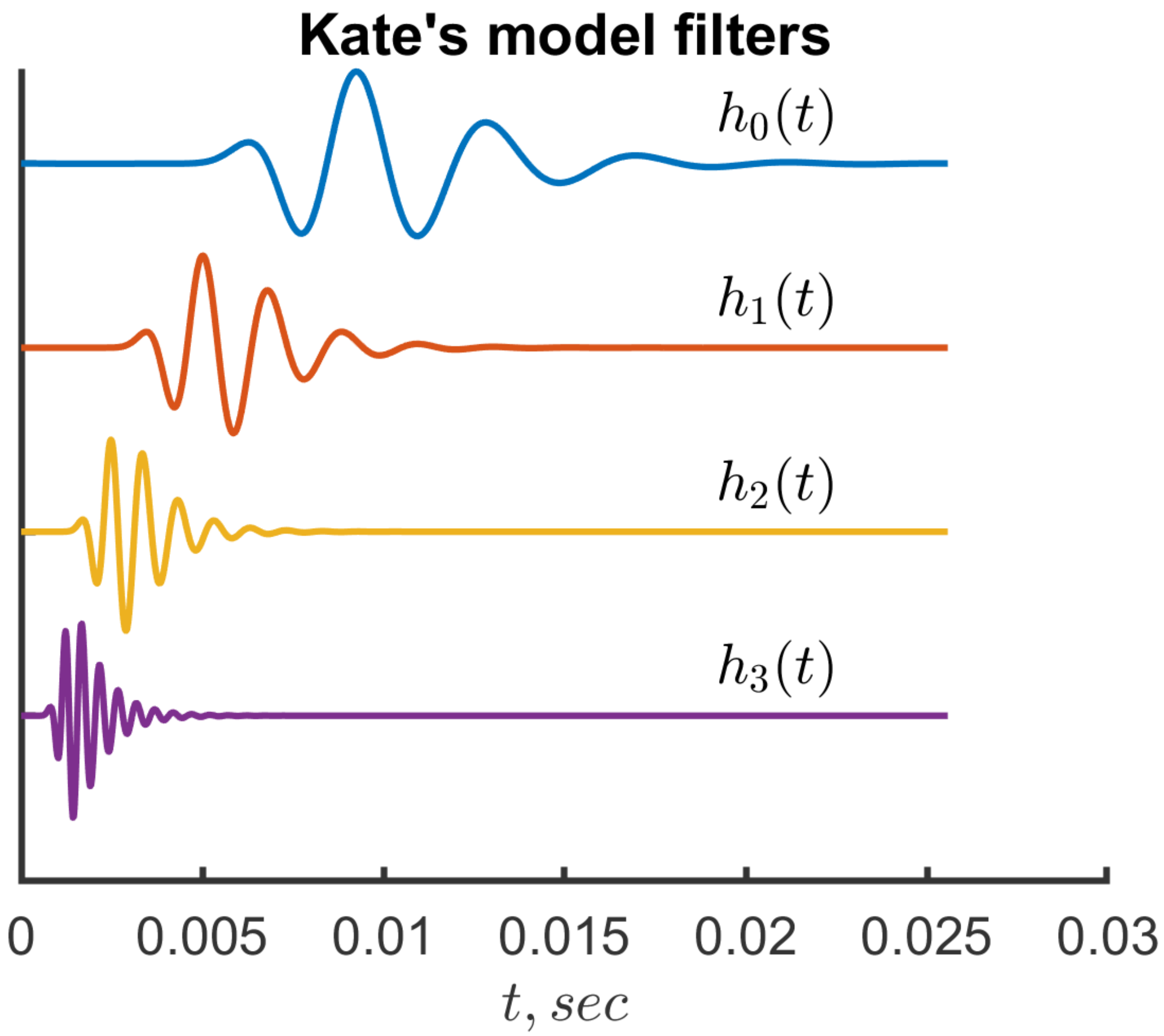
$H_{hp,k}(z)$ – one-pole highpass filter that models the pressure-to-velocity transformation; $F_k(z)$ – is a notch filter by which the total response shows two resonance frequencies, which coincides with biological observations; $H_k(z)$ – is a single section of the traveling wave filter, which provides a gain for frequencies near the resonance frequency of the filter.

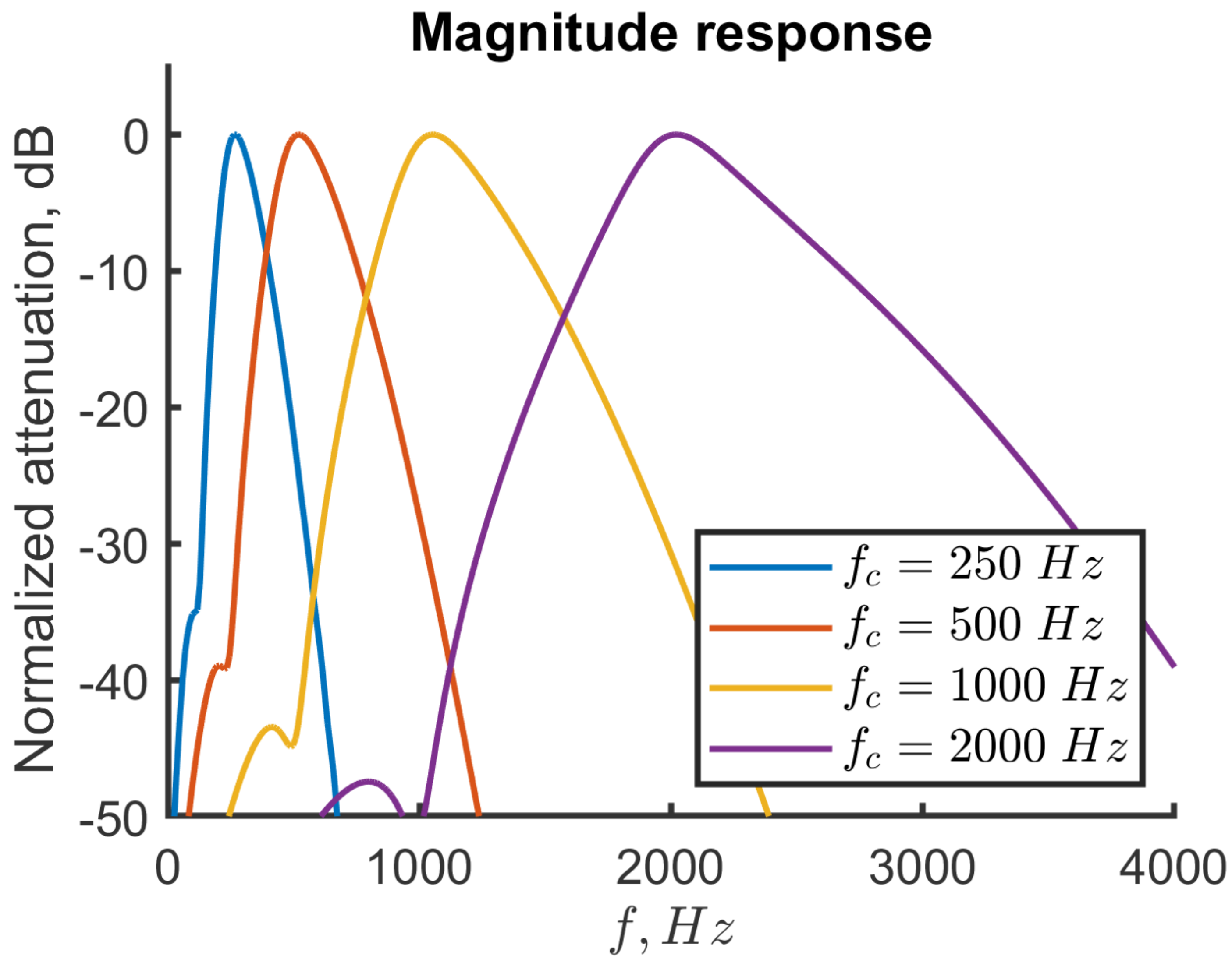
[4] Kates, J. M. A time-domain digital cochlear model // IEEE Transactions on Signal Processing. – 1991. – vol. 39, №. 12. – P. 2573–2592.

Digital time-domain model of the human cochlear (2)

Transfer function of auditory filters of Kate's model:

$$\mathrm{Coch}_k(z) = H_{hp,k}(z)F_k(z)\prod_{i=1}^{k} H_k(z).$$





Psychoacoustically motivated analysis

Drawbacks of ***bioinspired methods***:

✓ large computational burden;

✓ as a rule, resynthesize procedure is not defined;

✓ do not based on fast and effective DSP transforms;

Psychoacoustically motivated method:

✓ are based on the results of psychoacoustic experiments;

✓ are simulate frequency (sometimes temporal) resolution of hearing;

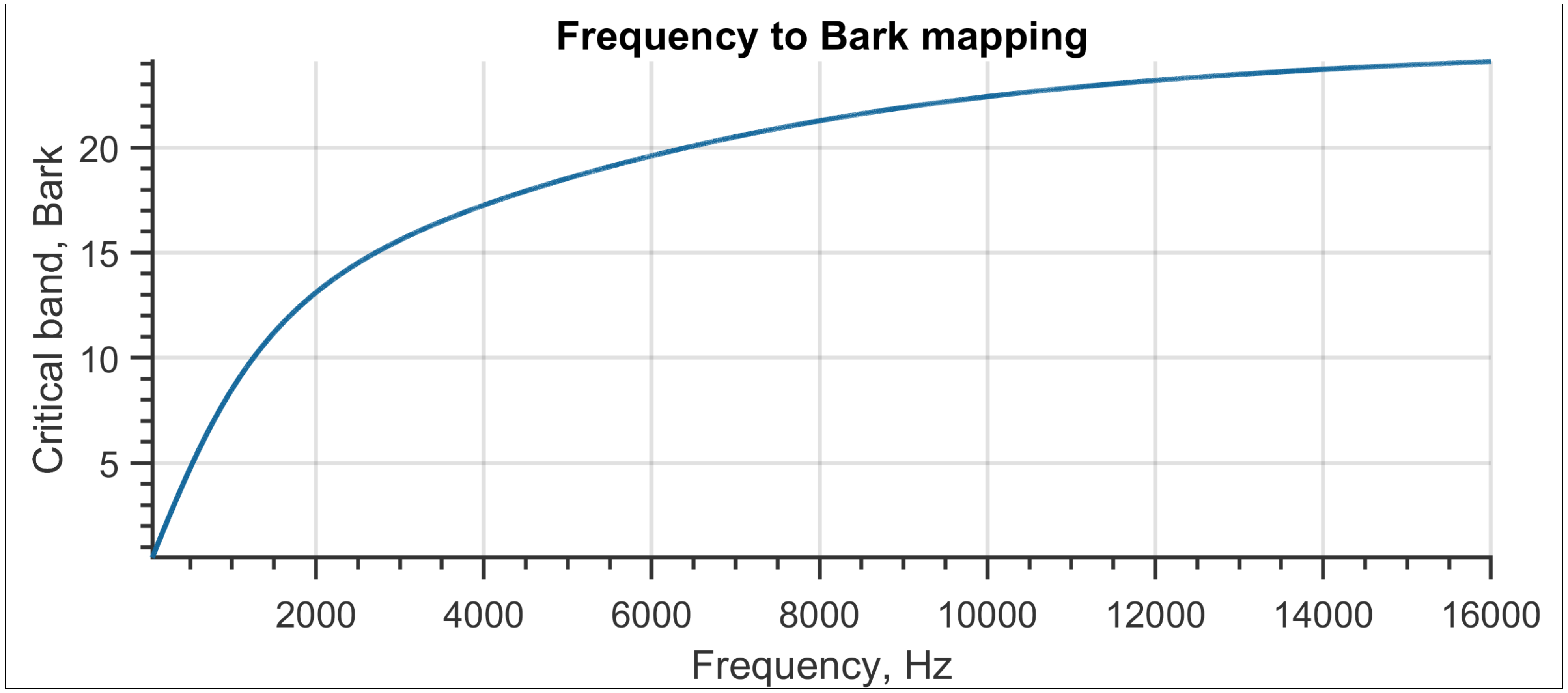
✓ are based on well-known and effective DSP transforms.

Critical bands

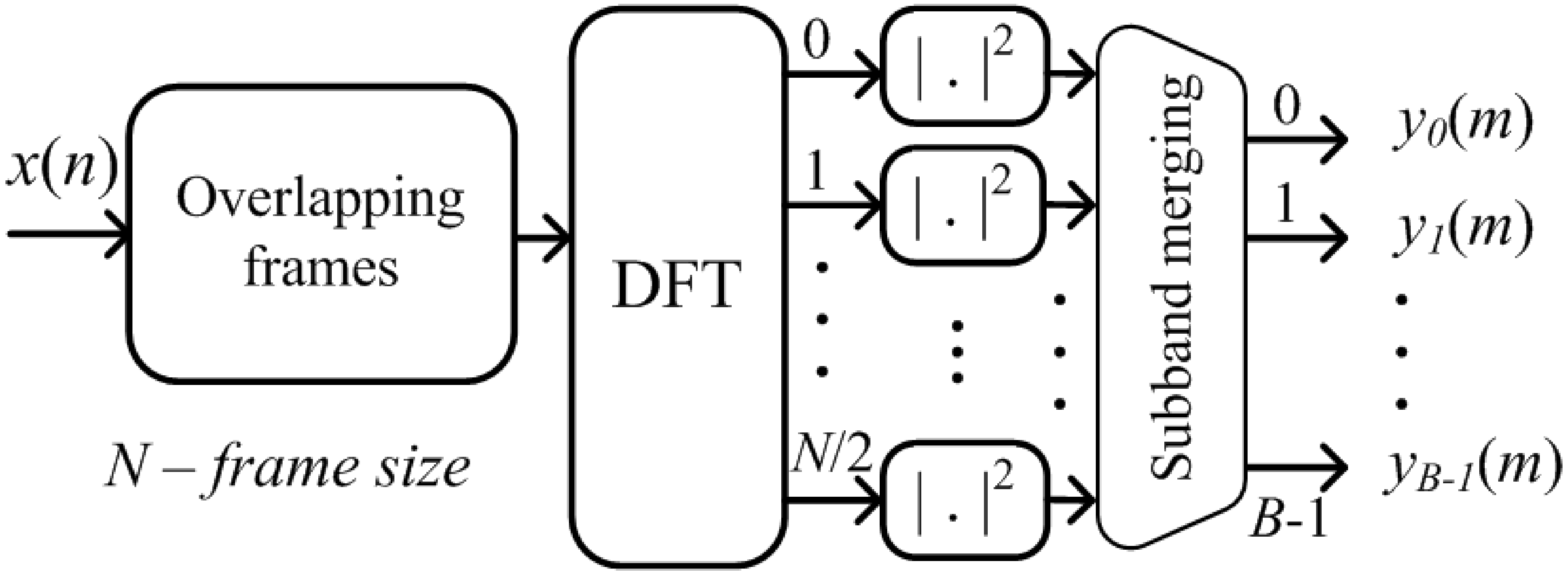
Critical bands define the frequency ranges within which the interaction of sound energies is observed.

The ***critical band*** is a frequency range (“bark”) with a width of 20% of the center frequency. The auditory range is covered by 24 barks.

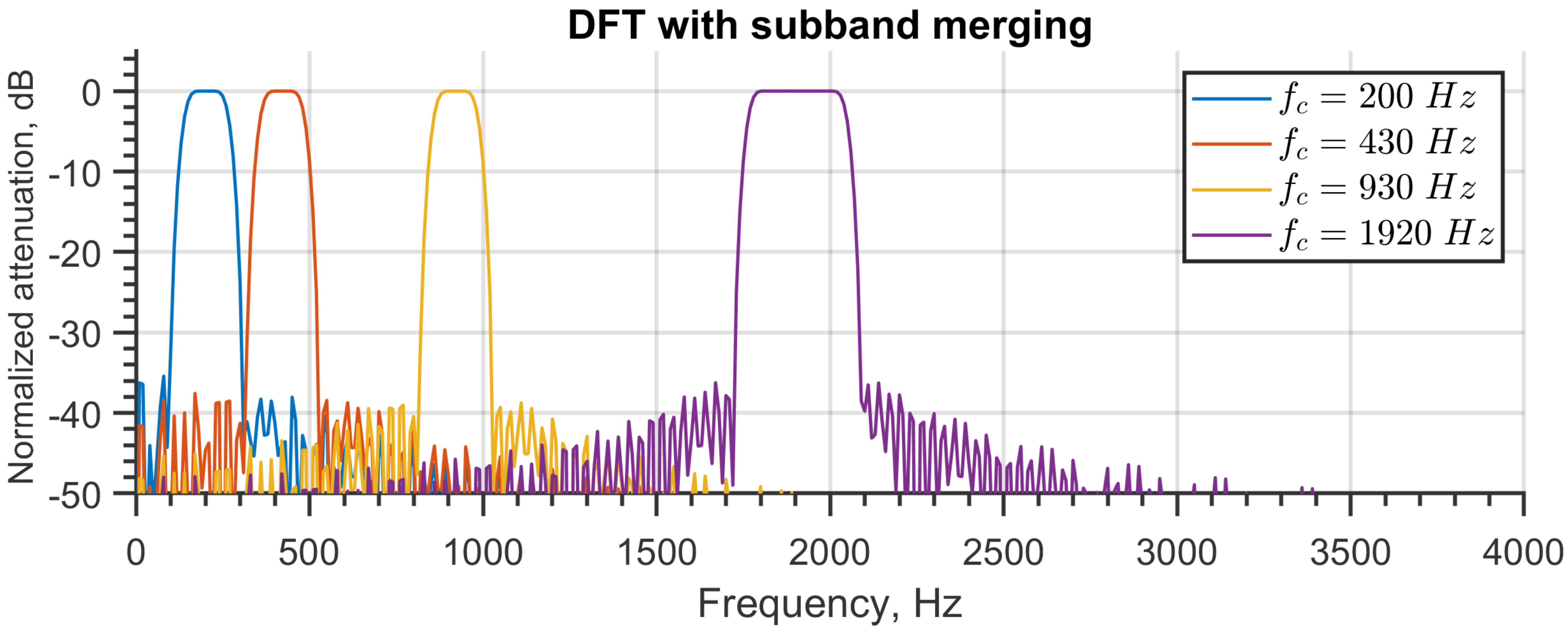
$$z(f) = 13\,\mathrm{arctg}\left(\frac{7{,}6}{10^4 f}\right) + 3{,}5\,\mathrm{arctg}\left(\left(\frac{f}{7500}\right)^2\right)\ \text{[Bark]}$$



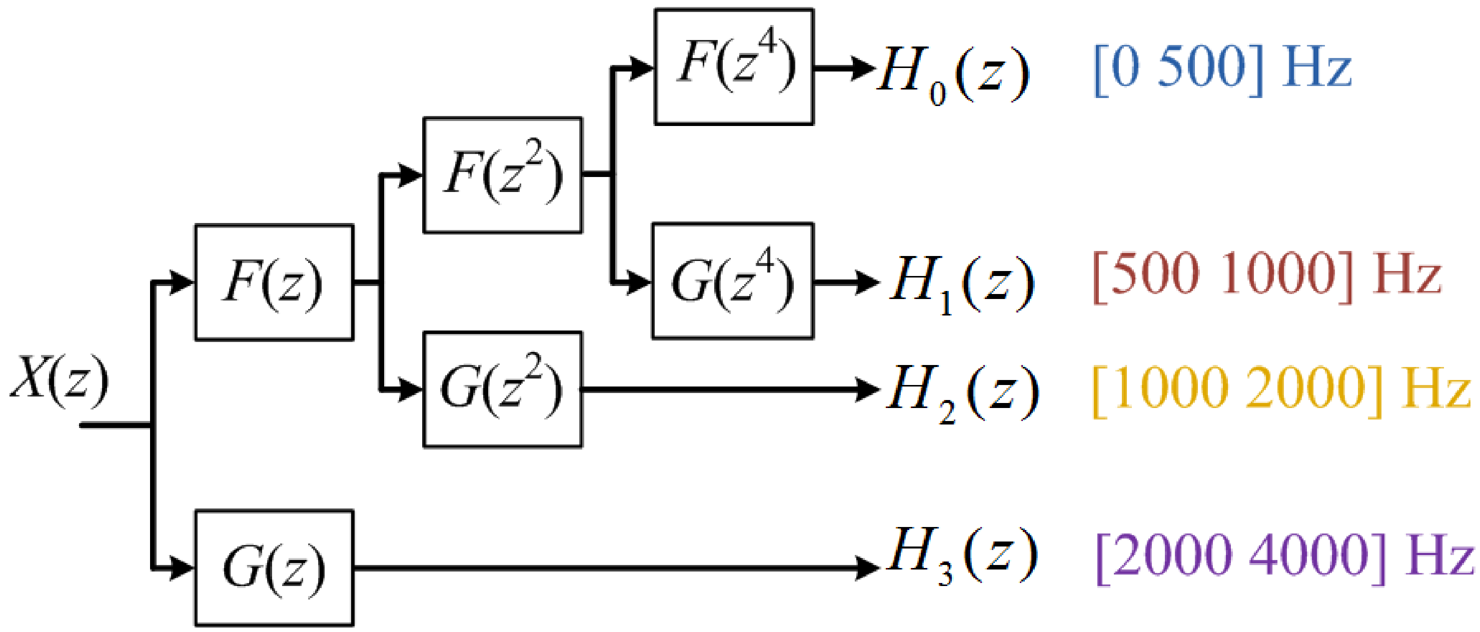
DFT with subband merging



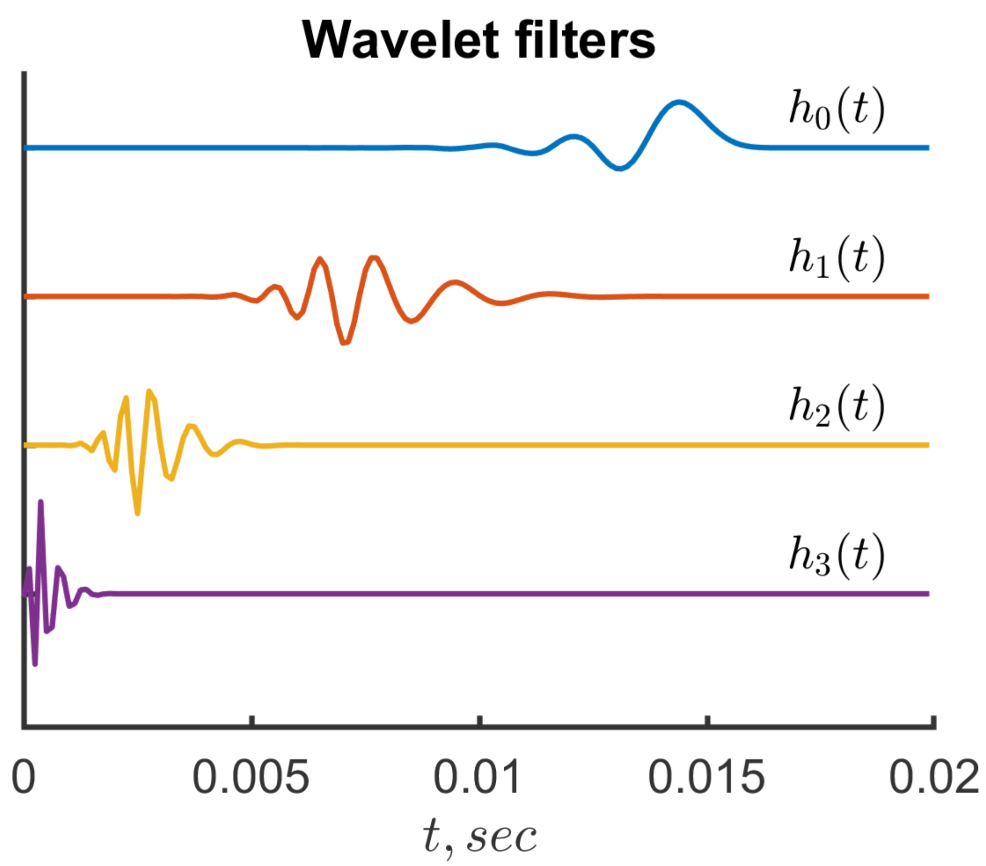
- ✓ Filters are defined in ***frequency domain***.
- ✓ Low computational complexity.
- ✓ Temporal resolution of hearing is not modeled.

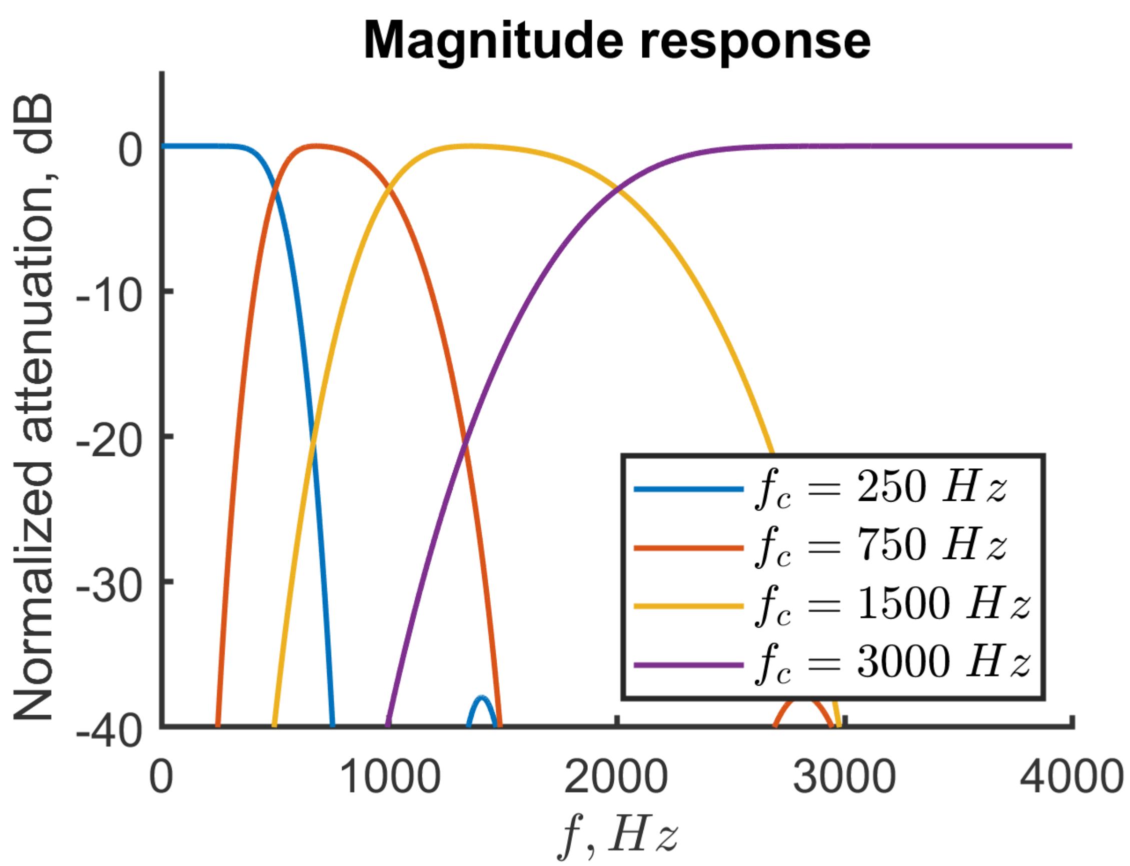


Auditory filters based on wavelets



- ✓ Temporal resolution of hearing is modeled.
- ✓ Relatively small computational complexity.
- ✓ Frequency resolution depends on filter bank structure.

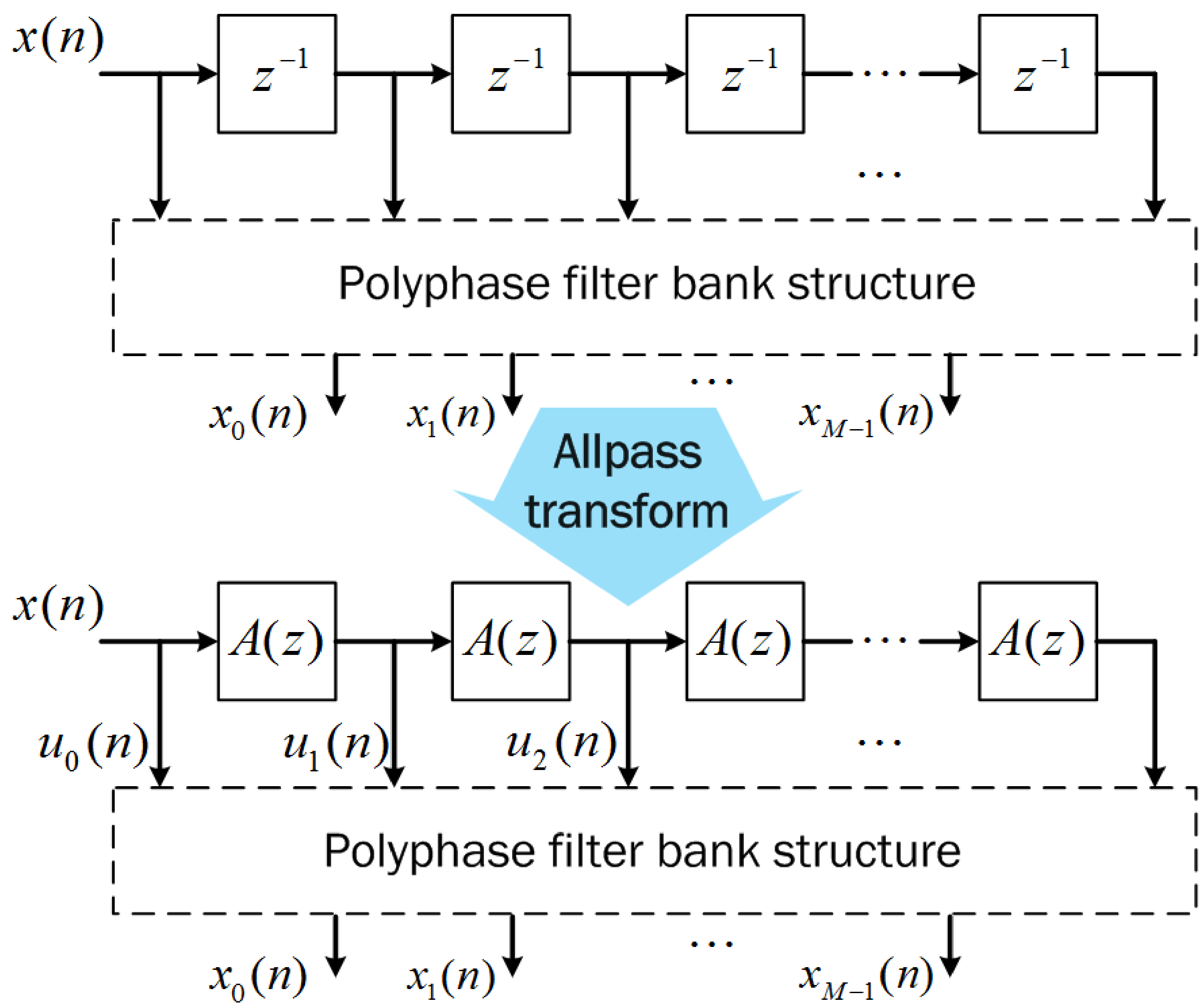




Filter bank based on allpass transform

Allpass transform allows to obtain ***nonuniform*** filter bank from ***uniform*** filter bank[5].

$$z^{-1} \rightarrow A(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}, \quad A\left(e^{j\omega}\right) = e^{j\theta(\omega)}.$$

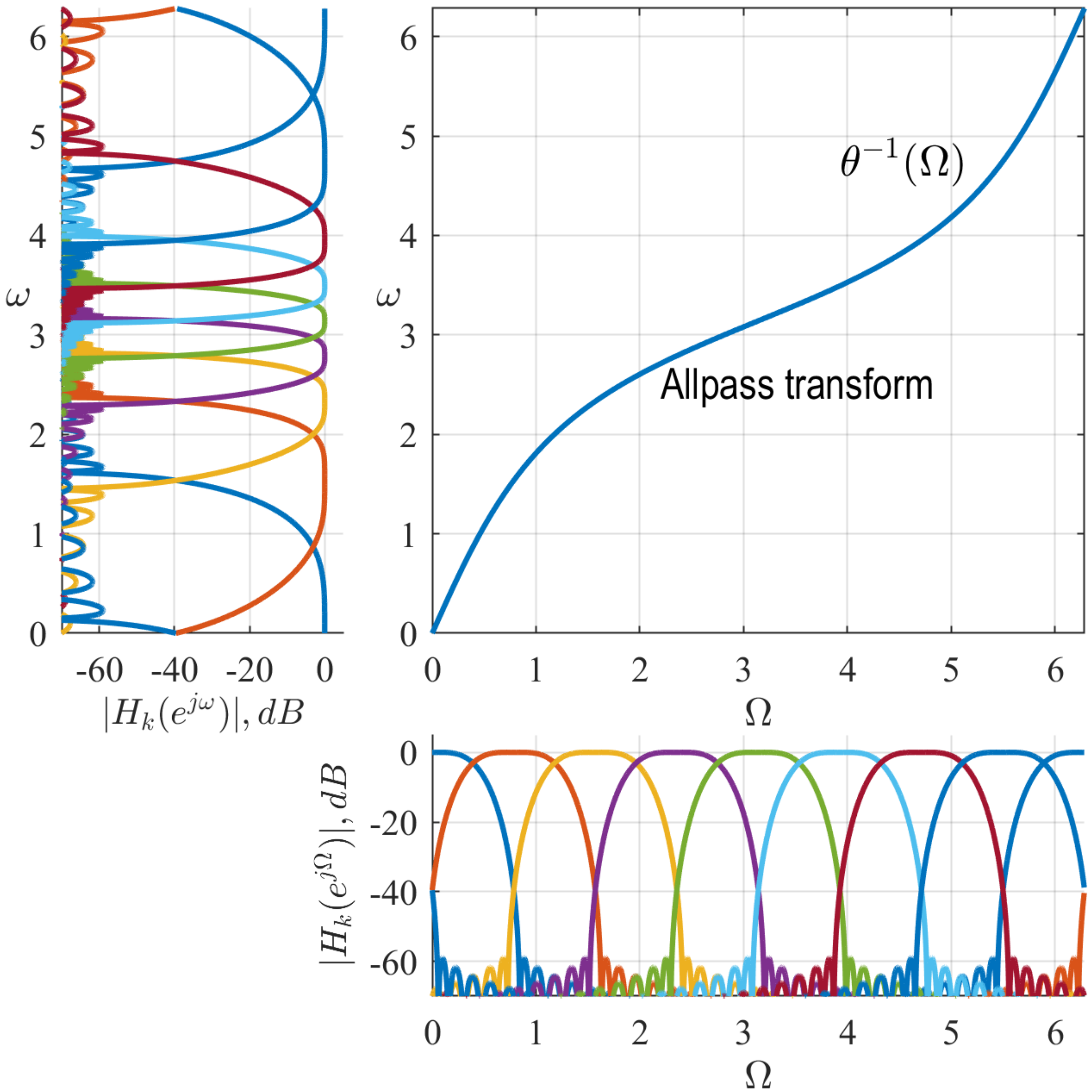


[5] Vary, P. Digital filter banks with unequal resolution // Short Communication Digest of European Signal Processing Conference (EUSIPCO), 1980. – pp. 41–42.

Filter bank based on allpass transform

Application of allpass transform to discrete-time system leads to frequency warping:

$$\theta^{-1}\colon \Omega \rightarrow \omega,$$

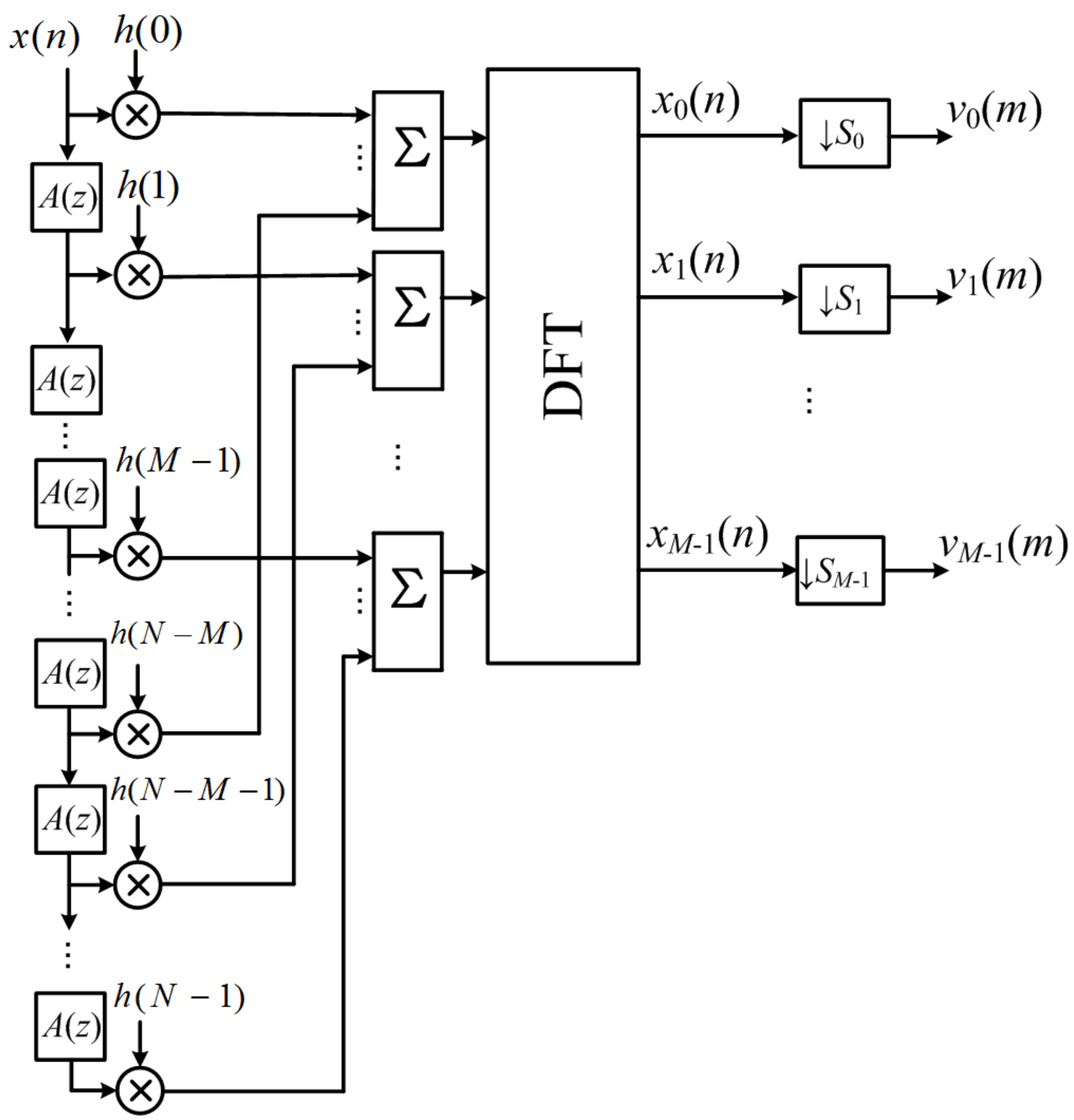


где

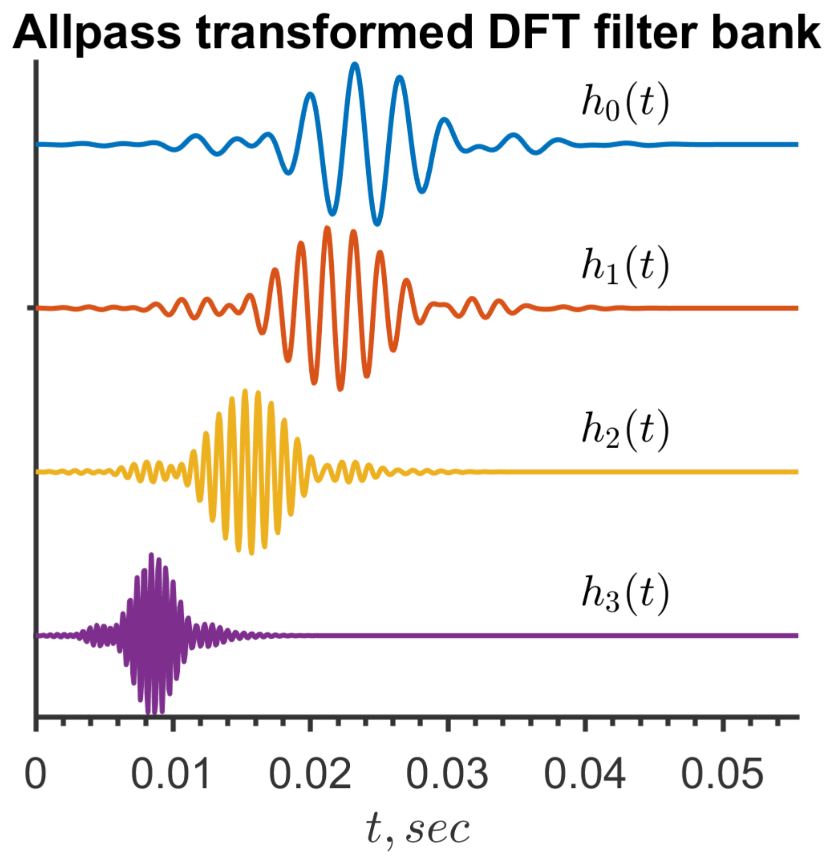
$$\theta(\omega) = \omega + \mathrm{arctg}\left(\frac{\alpha \sin \omega}{1 - \alpha \cos \omega}\right).$$

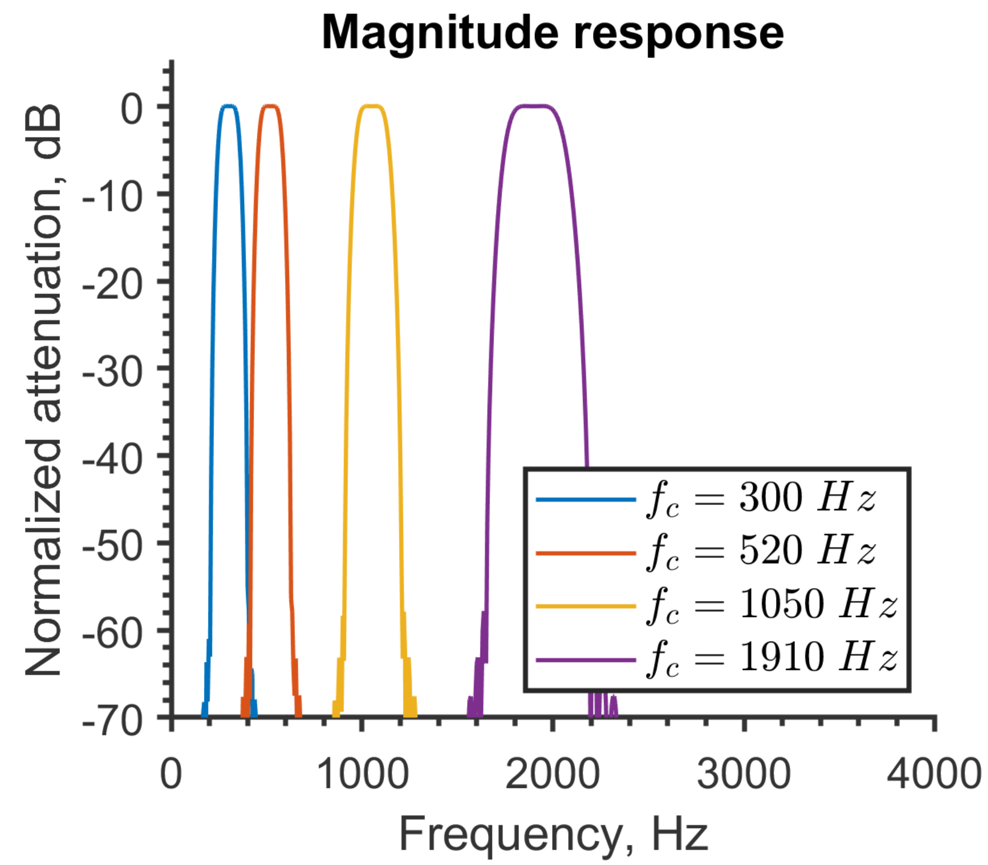
α – warping parameter.

Filter bank based on allpass transform



- ✓ Models frequency and temporal resolution of hearing.
- ✓ Relatively small computational complexity.
- ✓ Frequency resolution **does not depend** on filter bank structure.



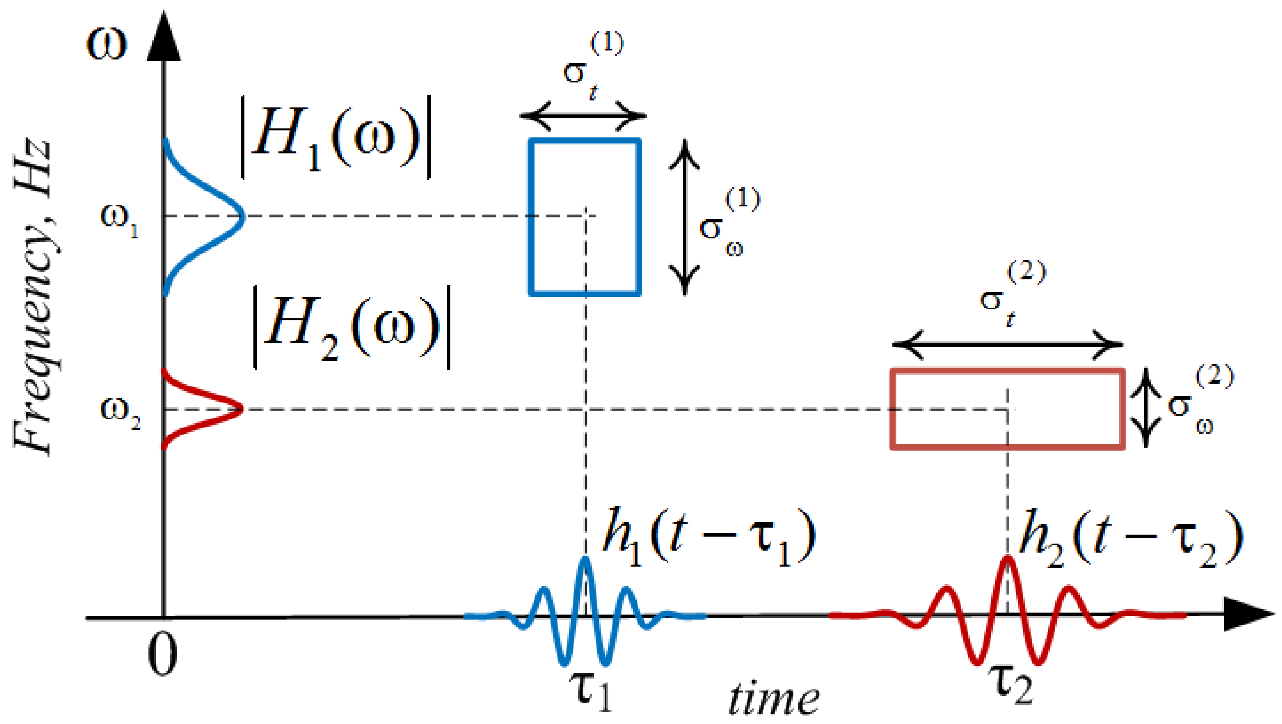


Comparison of time-frequency localization

We compared 3 different approaches to auditory filters modeling

- ✓ Gammatone filters
- ✓ Wavelet filters
- ✓ Allpass based DFT filter bank

Heisenberg rectangular[6]



Time localization

$$\sigma_t^{(k)} = \int_0^{\infty}(t-\tau_k)^2|h_k(t)|^2 dt$$

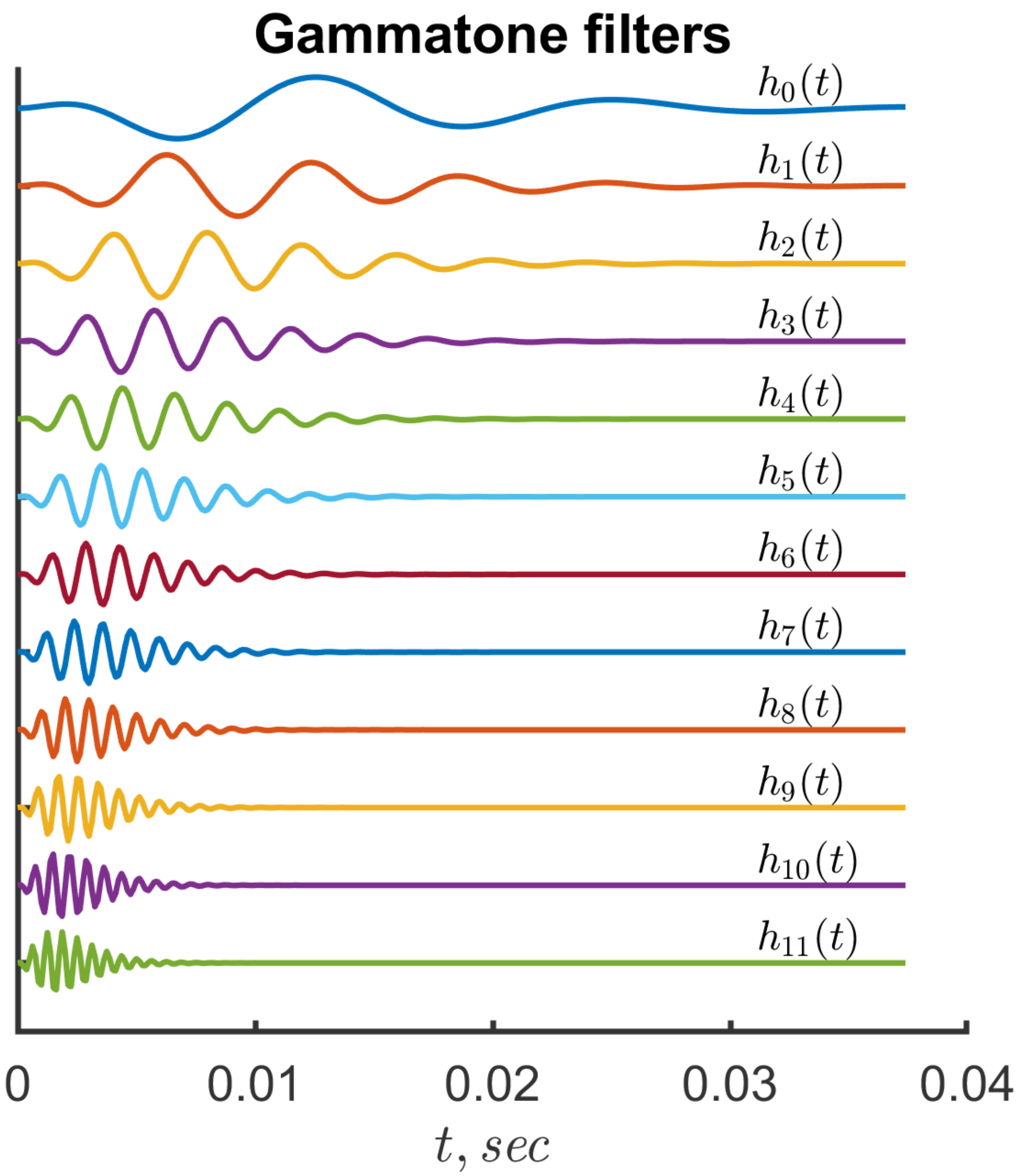
Frequency localization

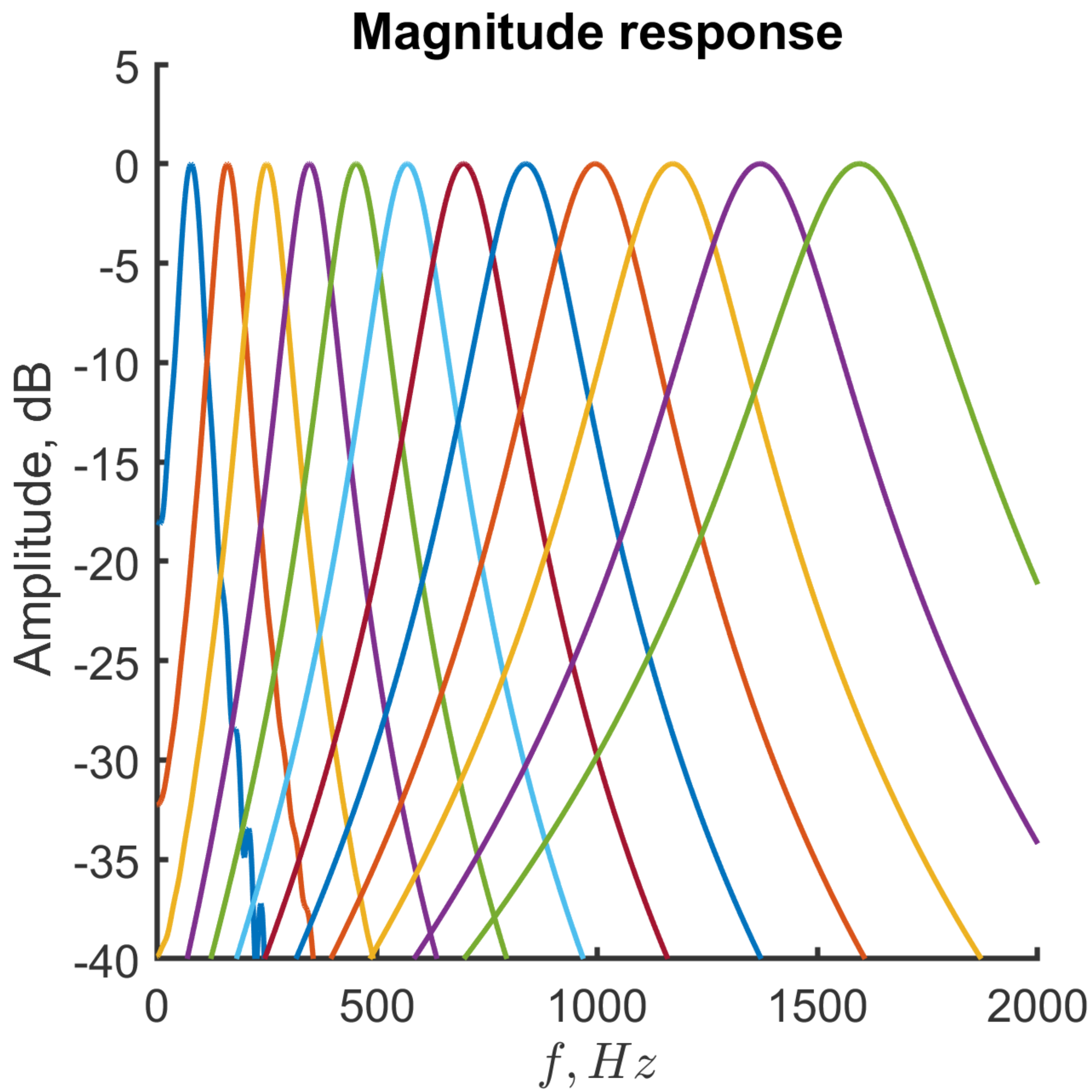
$$\sigma_\omega^{(k)} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(\omega-\omega_k)^2\left|H_k\left(e^{j\omega}\right)\right|^2 d\omega$$

[6] Вашкевич, М.И. Сравнение частотно-временных преобразований: Фурье-анализ, вейвлеты и банки фильтров на основе фазового преобразования / М.И. Вашкевич, И. С. Азаров // Цифровая обработка сигналов – 2020. – №2. – С. 13–26.

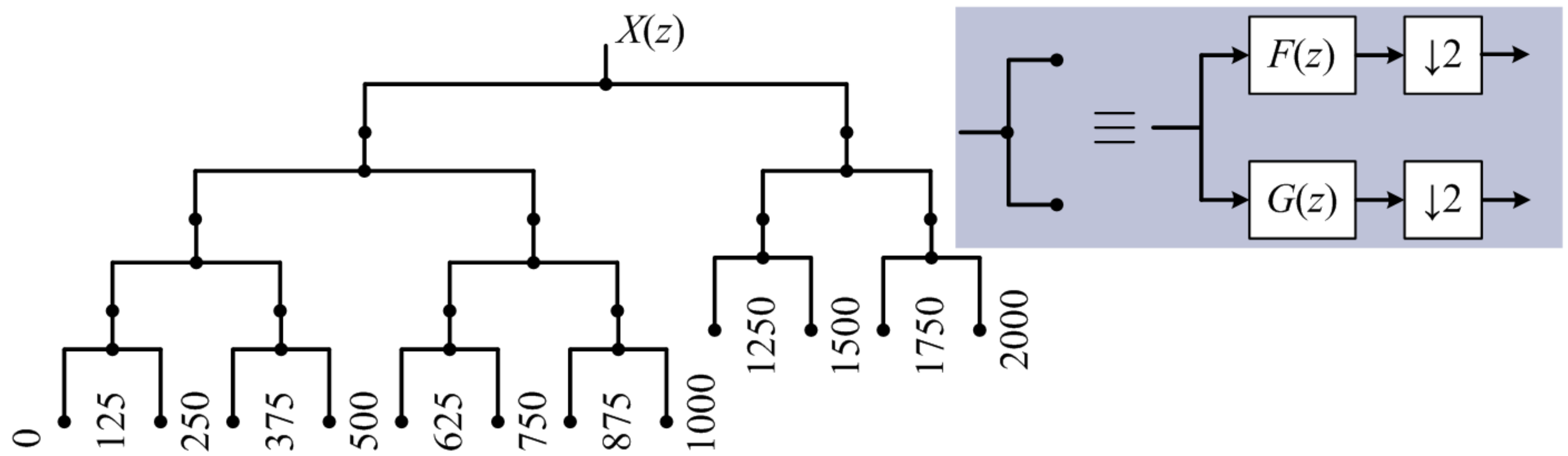
Comparison: gammatone filters

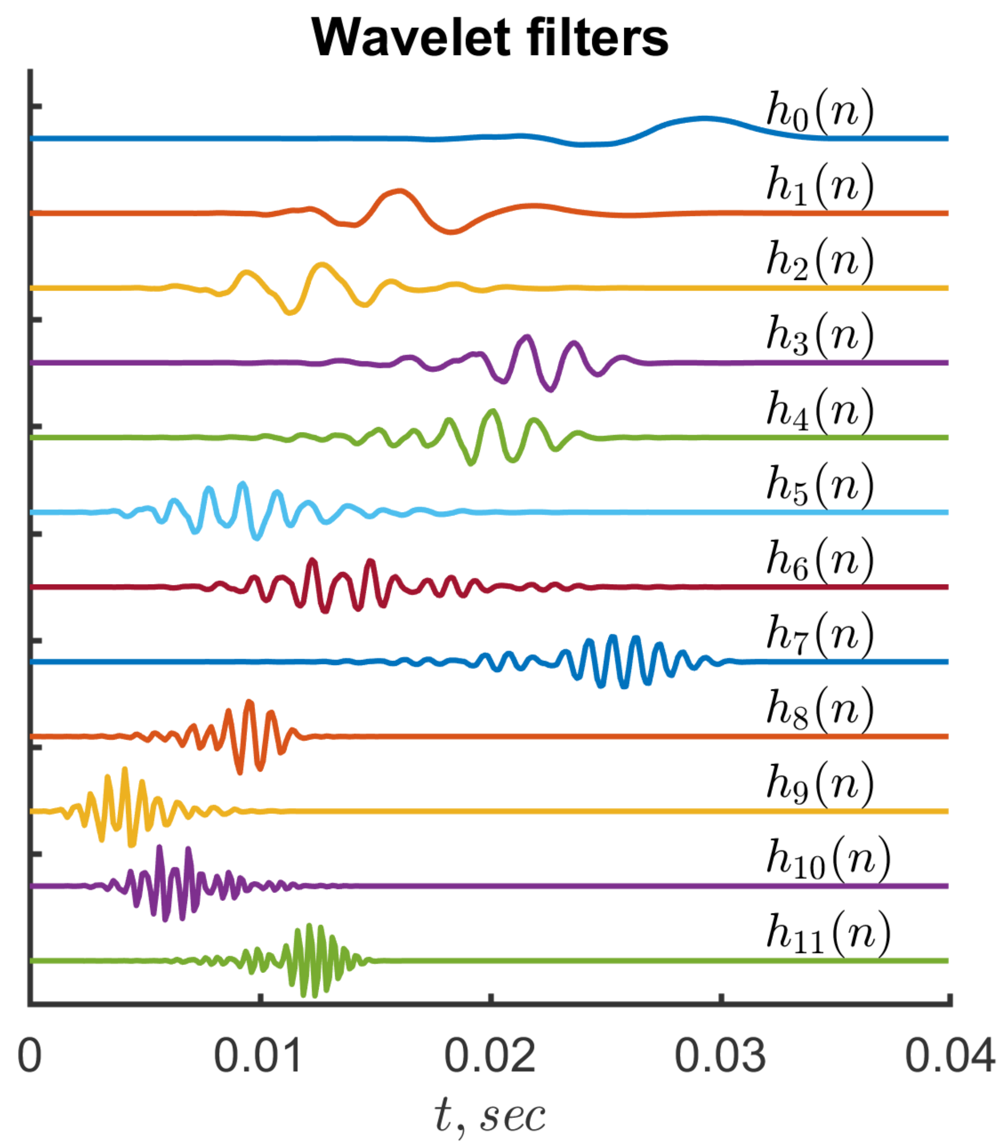
Auditory filters correspondent to first 12 critical bands.

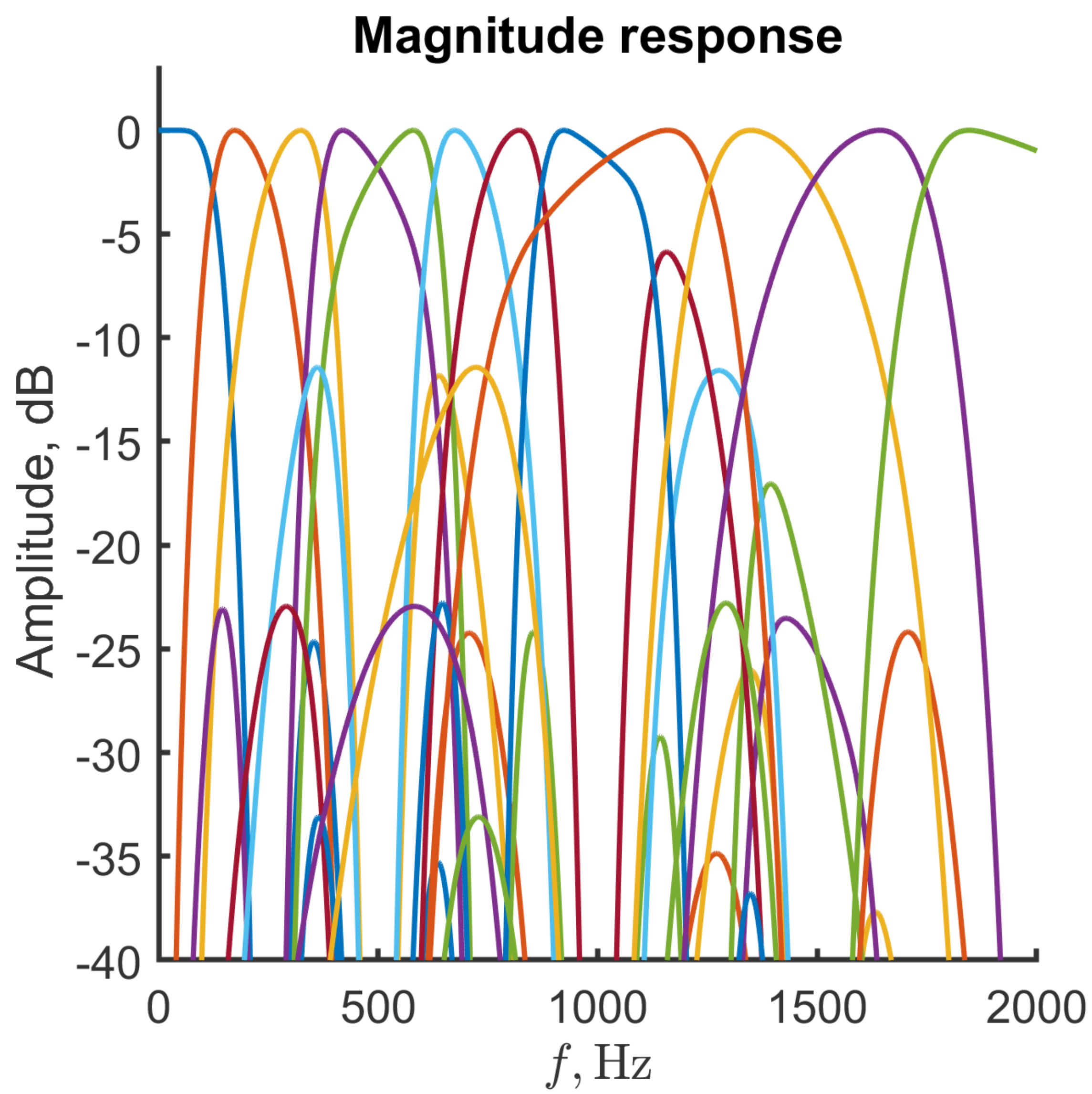




Comparison: wavelet filters

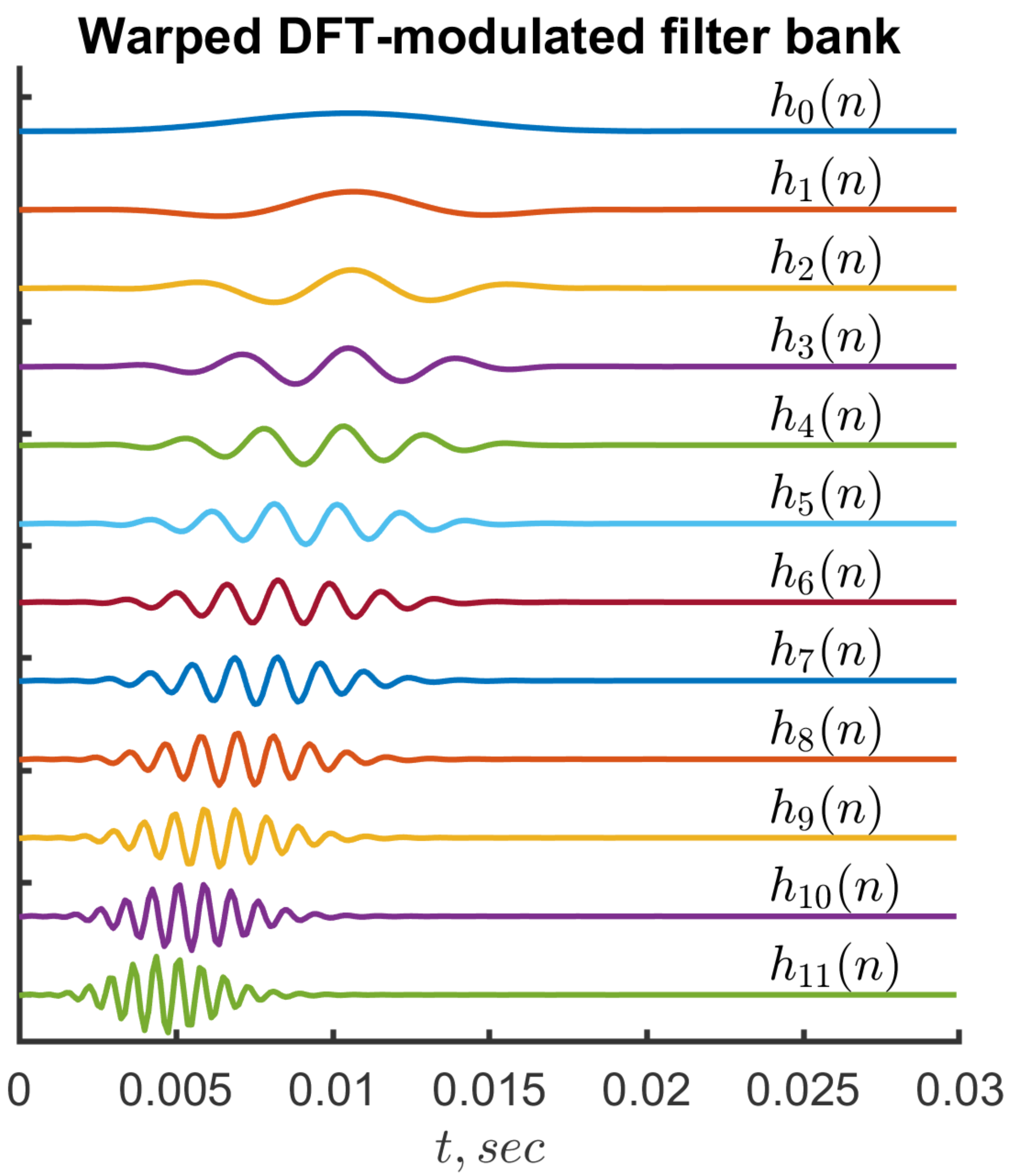


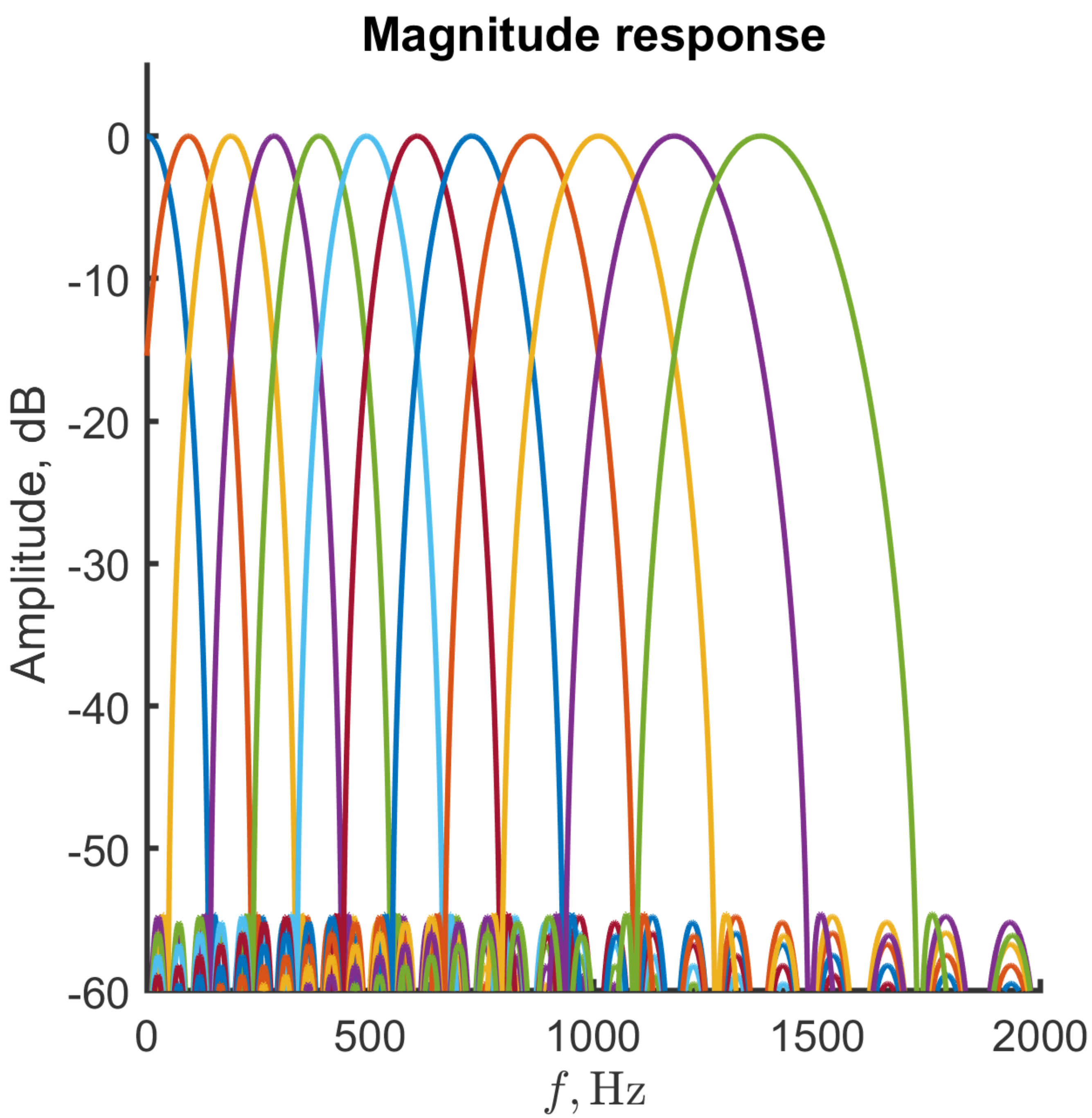




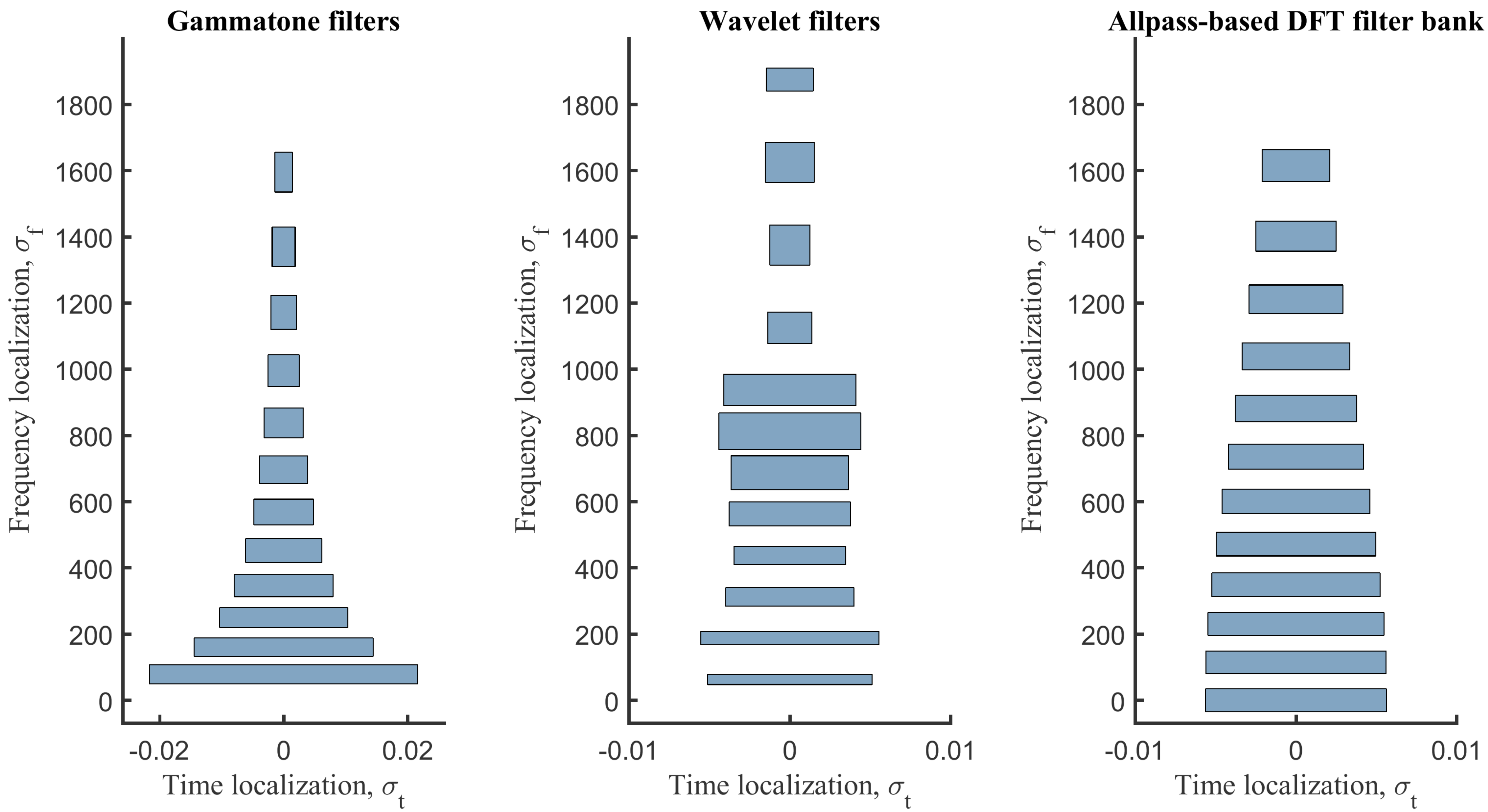
Comparison: allpass based DFT filter bank

Auditory filters correspondent to first 12 critical bands.





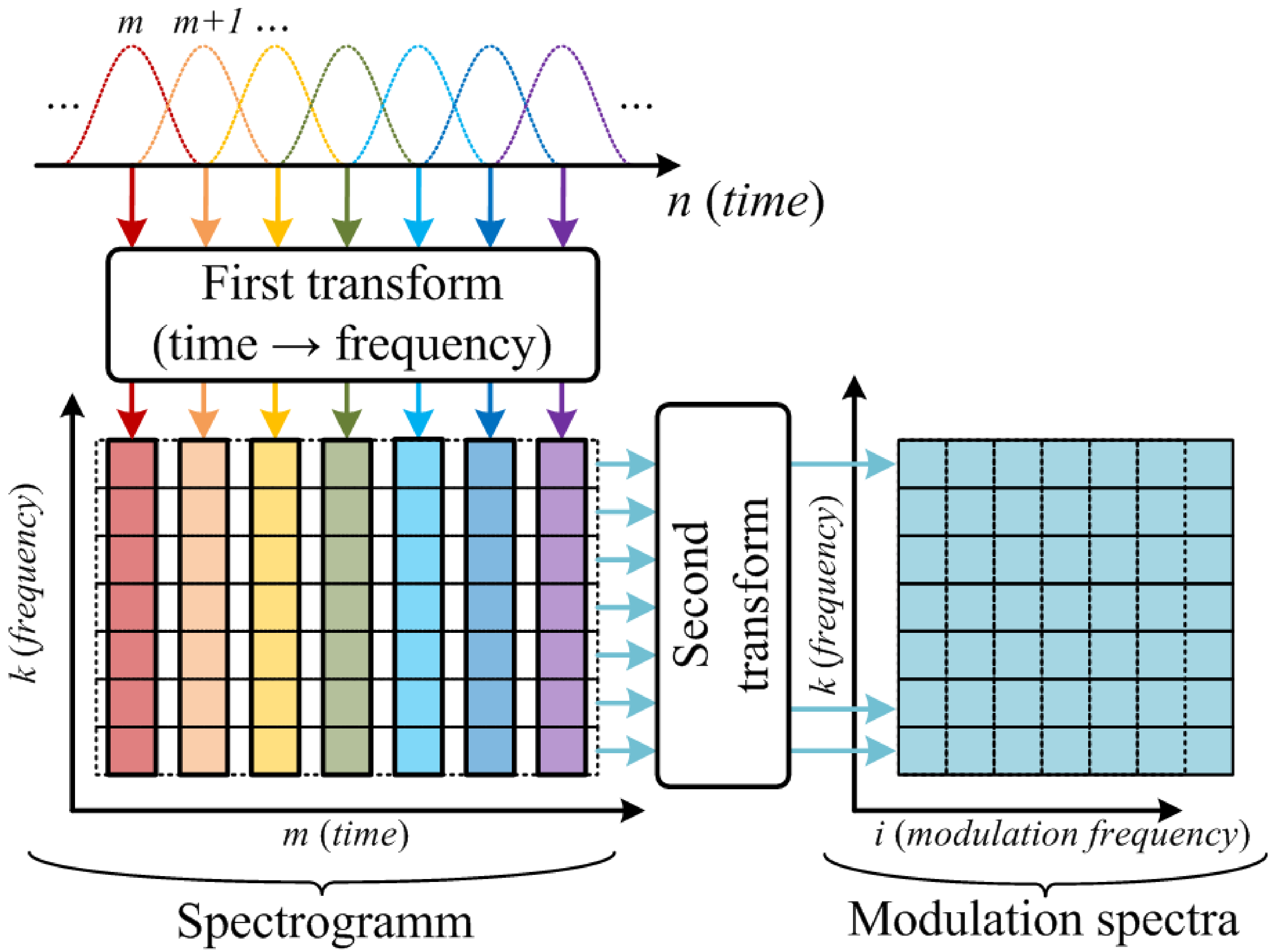
Comparison of time-frequency localization



Voice pathology detection

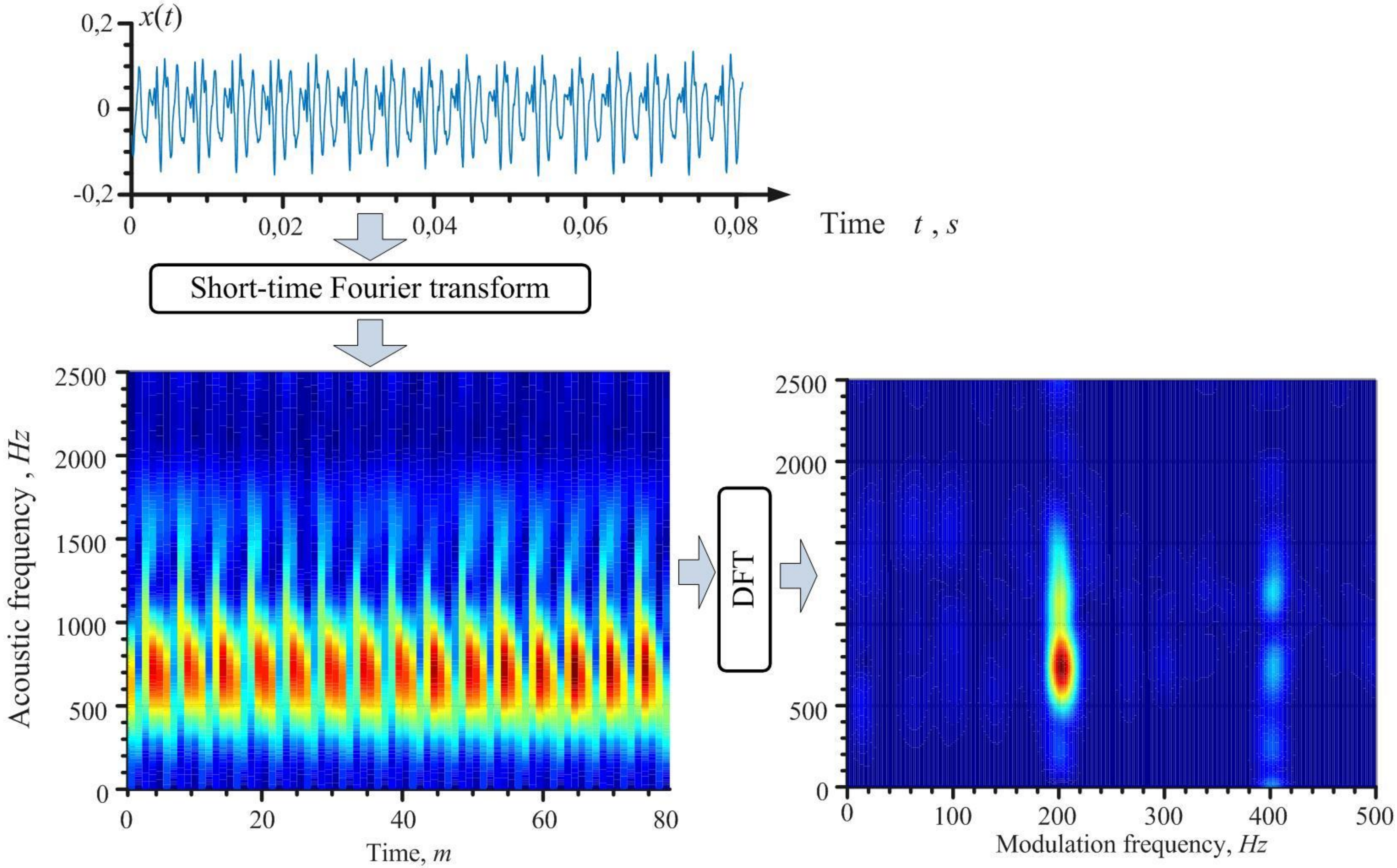
Application of psychoacoustically motivated analysis:

Voice pathology detection based on ***critical band modulation spectra analysis***[7].



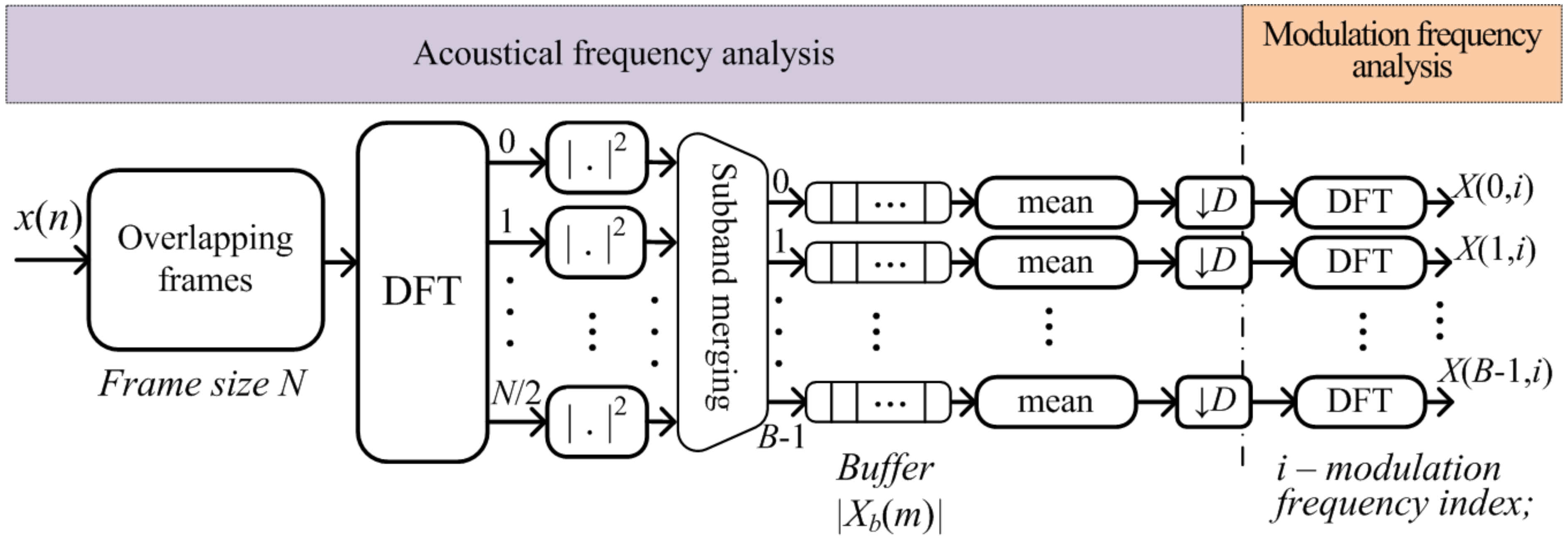
[7] Вашкевич М. И., Азаров И. С. Определение патологии голосового аппарата на основе анализа модуляционного спектра речи в критических полосах / М. И. Вашкевич, И. С. Азаров // Труды СПИИРАН. – 2020. – Т. 19. – №. 2. – С. 249–276.

Modulation spectra: example

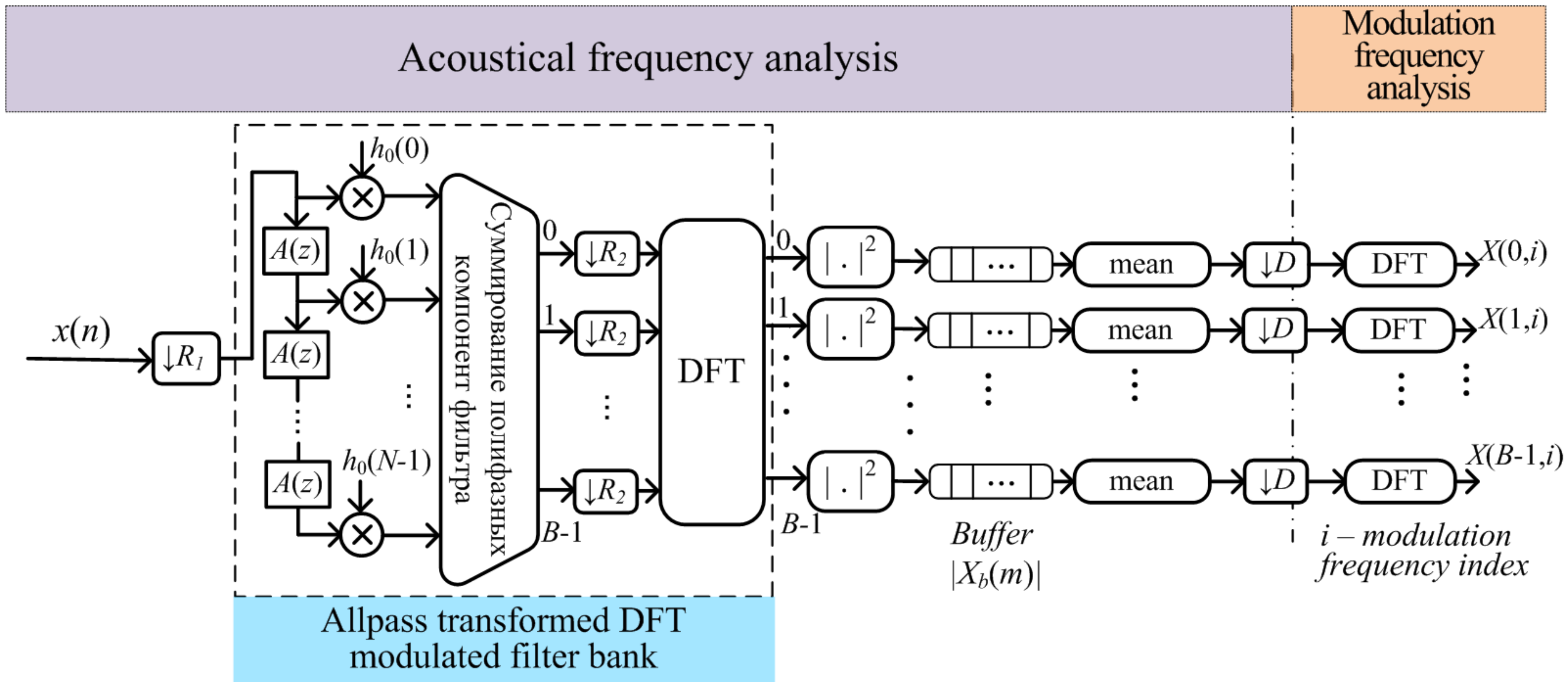


Two analysis scheams (feature extraction)

Acoustic frequency anslysis using **DFT with subband merging**

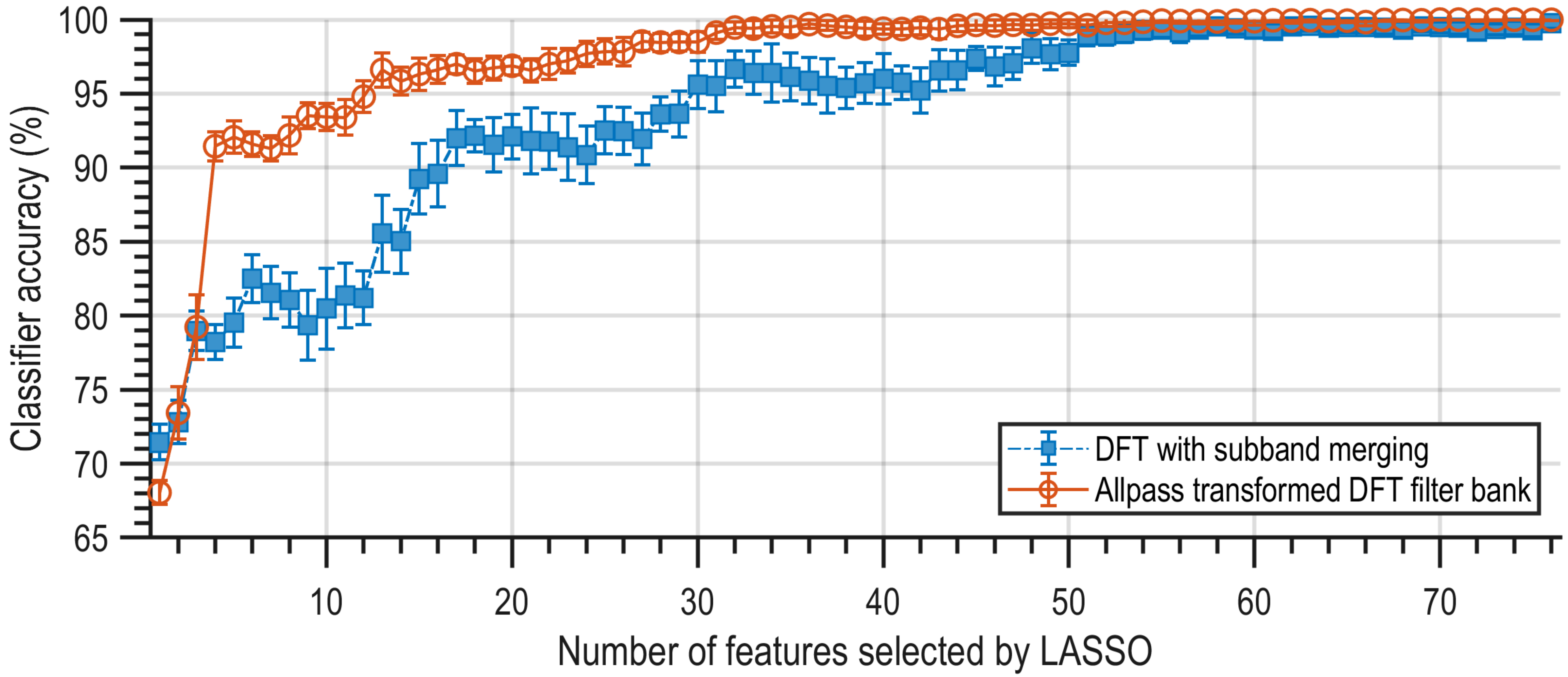


Acoustic frequency anslysis using **Allpass transformed DFT filter bank**



Classification results

Voice base: 30 recordings of patients with larynx pathology + 30 healthy controls.



Conclusion: *parameters extracted using filter bank based on allpass transform is more effective*.

Conclusion

1) *Bioinspired methods of analysis*:

– roex-filters;

– gammatone filters;

– gammachirp filters;

– Kates model.

2) *Psychoacoustically motivated analysis*:

– DFT with subband merging;

– Filter bank based on wavelets;

– Filter bank based on allpass transform;

3) *Application*: voice pathology detection based on psychoacoustically motivated analysis.

Thank you for attention!