TRƯỜNG ĐẠI HỌC XÂY DỰNG HÀ NỘI
KHOA CÔNG NGHỆ THÔNG TIN

LẬP TRÌNH ÂM THANH

CHƯƠNG 7

Tăng cường chất lượng

Chương 7 · 5 tiết · Mã học phần 531833 · 75 tiết · Học kỳ 5

Hà Nội – 2026

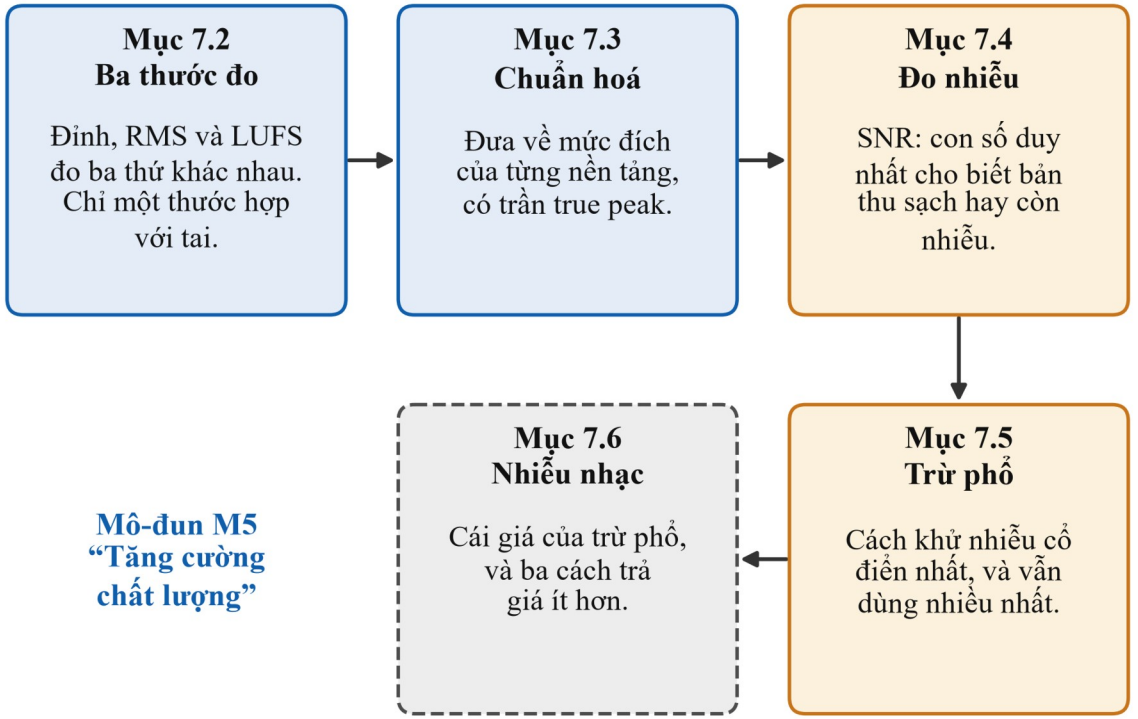
Sau chương này người học phải làm được

- Giải thích vì sao ba tệp cùng mức đỉnh vẫn không to bằng nhau; chọn thước đo phù hợp cho từng tình huống.
- Đo độ to theo LUFS (Loudness Units Full Scale – đơn vị độ to so với toàn thang) ở ba thang thời gian, đọc LRA (Loudness Range – dải độ to), giải thích vai trò của hai cổng trong BS.1770.
- Đo true peak và nêu vì sao trần luôn nằm dưới 0 dBFS (decibel so với mức toàn thang, Full Scale).
- Chuẩn hoá một tệp về mức đích của một nền tảng, đủ bốn bước, có kiểm trần.
- Đo SNR (Signal-to-Noise Ratio – tỉ số tín hiệu trên nhiễu) của bản thu chỉ có bản đã trộn; quyết định có nên khử nhiễu hay không.
- Cài đặt phép trừ phổ, chọn hai tham số có căn cứ, giải thích vì sao pha được giữ nguyên.
- Nhận ra và đo nhiễu nhạc; nêu ba cách giảm nó cùng cái giá của mỗi cách.
- Sắp đúng thứ tự các bước; nêu cái giá bằng số của việc chuẩn hoá trước khi khử nhiễu.

MỤC 7.1

Vị trí chương học trong học phần

- Chương 5 và 6 dạy làm ra một bản dựng; chương này hỏi bản dựng ấy đã nghe được chưa.
- Hai việc phải làm: chỉnh mức cho đúng đích và hạ nhiễu nền xuống.
- Cả hai việc đều thay một cuộc tranh cãi bằng một con số đo được.



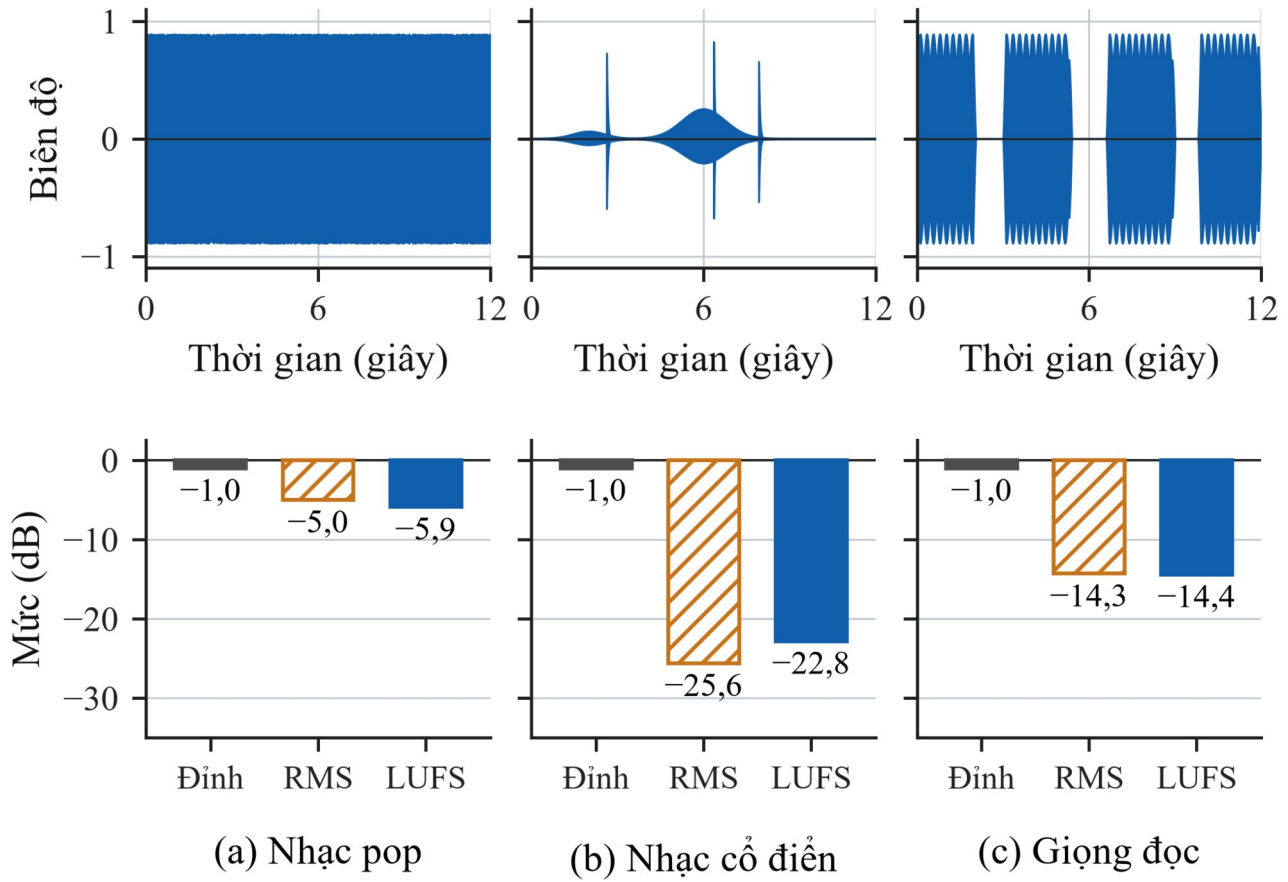
Hai việc: chỉnh mức cho đúng đích, và hạ nhiễu nền xuống. Cả hai đều có một con số đo được thay cho một cuộc tranh cãi, và đều có một cái giá phải trả.

Hai việc, hai bức tường

Chuẩn hoá độ to đụng trần true peak; khử nhiễu đụng nhiễu nhạc. Không việc nào có lời giải đúng tuyệt đối. Vì thế chương này không dạy "làm thế nào cho đúng" mà dạy "đo cái gì để chọn có ý thức".

Ba tệp cùng đỉnh, ba độ to

- Chuẩn hoá ba tệp (pop, cổ điển, giọng đọc) về đỉnh −1 dBFS rồi nghe ở cùng mức loa.
- Cột đỉnh của cả ba bằng nhau chính xác.
- RMS (Root Mean Square – giá trị hiệu dụng) chênh nhau 20,6 decibel; LUFS chênh nhau 17,0 decibel.
- Tai xác nhận: ba tệp không to bằng nhau.



Hình 7.2 — Ba tệp cùng đỉnh −1 dBFS, đo bằng ba thước khác nhau

Đỉnh không trả lời câu hỏi "to hay nhỏ"

Đỉnh trả lời câu hỏi "có vượt thang không" — một câu hỏi kỹ thuật, không liên quan tới việc tiếng nghe to hay nhỏ. Muốn biết to hay nhỏ phải dùng một thước khác: thước mà cả ngành đang dùng.

Ba nguyên tắc học

- **Kiểm lại từng con số:** mọi giá trị sinh ra từ mô-đun sigtools và kiểm được bằng vài dòng mã.
- **Hai chỗ không có lời giải tuyệt đối:** trần true peak và nhiễu nhạc; chỉ có lựa chọn có ý thức.
- **Nghe là bắt buộc:** nhiễu nhạc nhìn phổ đồ thì thấy, nhưng chỉ nghe mới hiểu vì sao nó khó chịu.

Ba thứ gần giống mà thuộc chương khác

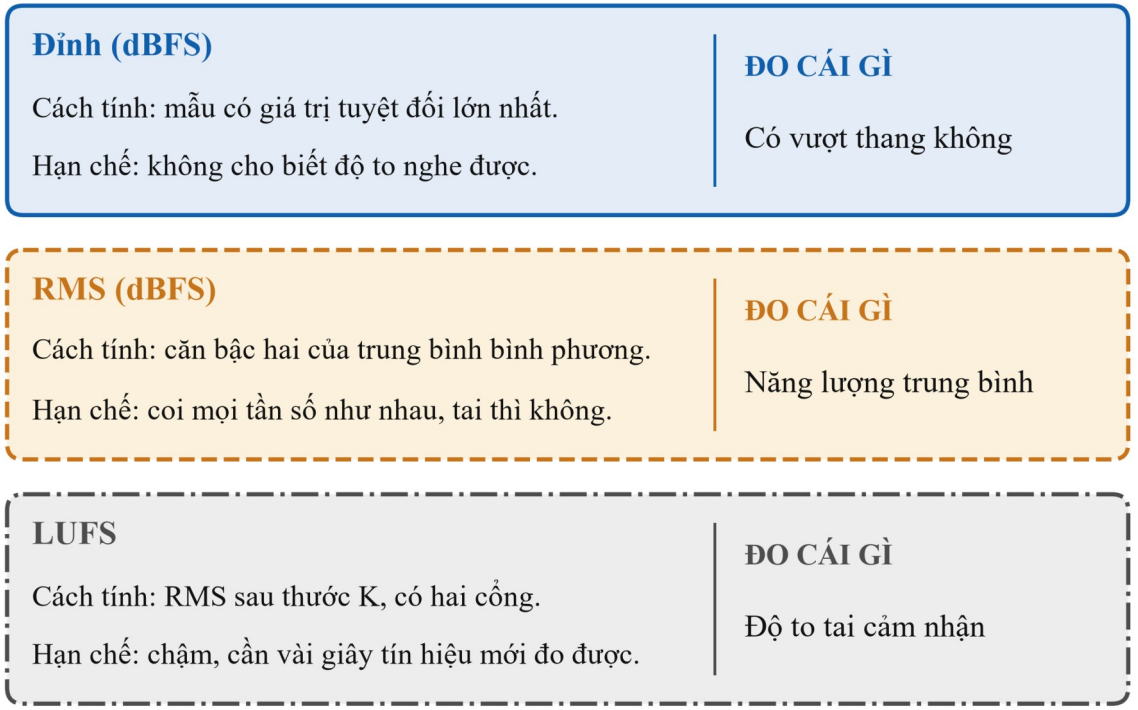
Thứ dễ gộp nhầm	Vì sao không thuộc chương 7	Thuộc chương
Bộ nén động, bộ giới hạn	Đổi dải động; cần ngưỡng, tỉ số, thời gian tấn công	Chương 8
Cân bằng tần số (EQ)	Cần thiết kế bộ lọc	Chương 8
Khử nhiễu bằng học máy	Không dựa trên giả định nhiễu dừng; cần dữ liệu huấn luyện ngoài phạm vi học phần	Ngoài chương trình

Khung học trừ phổ — đo trước, chọn tham số có ý thức, đo lại cả hai mặt — dùng được với mọi công cụ, kể cả hộp đen.

MỤC 7.2

Ba thước đo mức

- Đỉnh, RMS và LUFS đo ba thứ khác nhau và không thay thế nhau.
- LUFS có trọng số theo tần số, bám sát cảm nhận của tai.
- Một tệp có ba con số độ to theo ba thang thời gian, cộng thêm LRA.



Ba thước không thay thế nhau.

Một bản giao nộp đúng quy cách phải kiểm cả ba.

Mức đỉnh và mức hiệu dụng

$$L_{\text{đỉnh}} = 20\log_{10}\max_{n}|x[n]|\,; \qquad L_{\text{rms}} = 20\log_{10}\sqrt{\frac{1}{N}\sum_{n}x^{2}[n]}$$

(7.1)

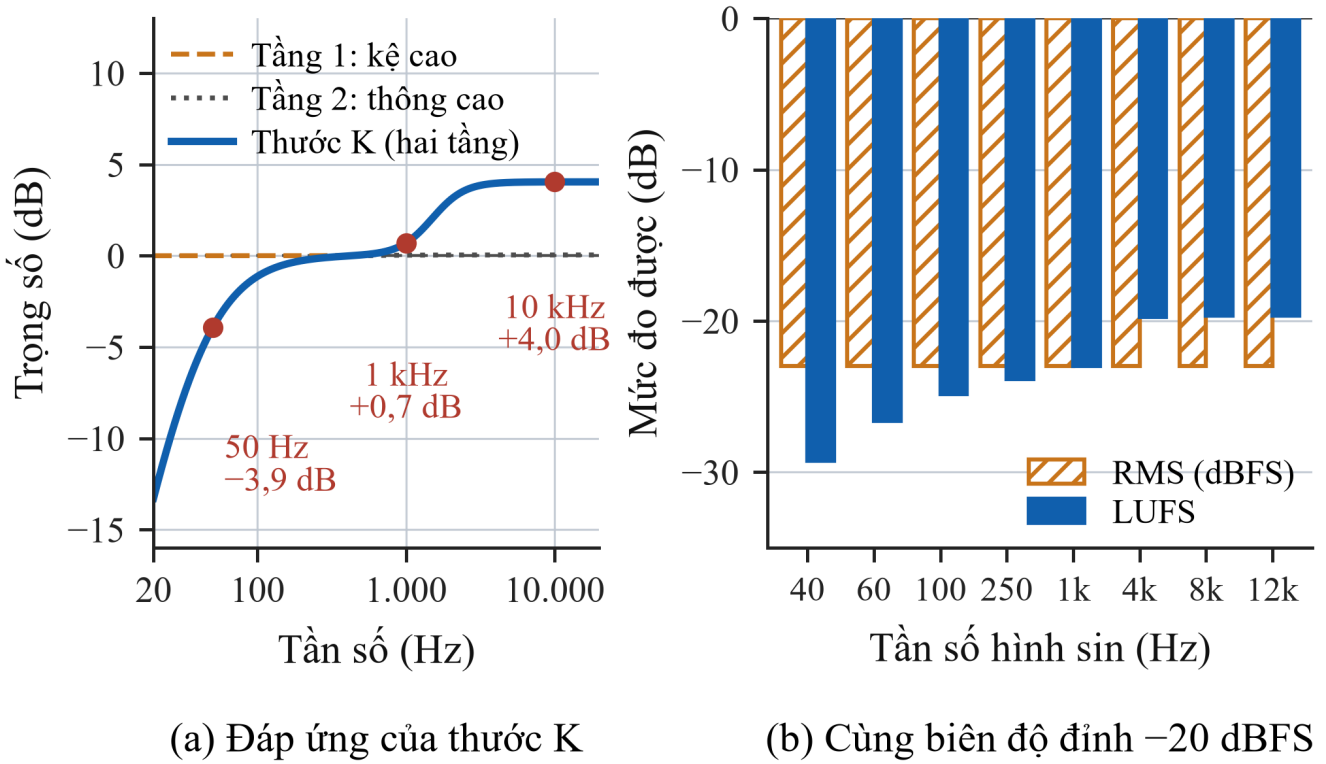
- Ldinh là mức đỉnh: mẫu có giá trị tuyệt đối lớn nhất.
- Lrms là căn bậc hai của trung bình bình phương trên N mẫu.
- Cả hai tính bằng dBFS, tức decibel so với toàn thang.
- Khoảng cách giữa hai mức là hệ số đỉnh (mục 5.7).

RMS coi mọi tần số như nhau

- Tai người không như vậy: kém nhạy với dải trầm, hơi nhạy hơn với dải 2–5 kHz.
- Đó là nội dung đường đẳng âm ở mục 1.5.
- Hai bản nhạc có RMS bằng nhau chính xác vẫn có thể nghe khác nhau vài decibel.
- Cần một thước có trọng số theo tần số.

Thước K: hai bộ lọc mắc nối tiếp

- **Tầng một — kệ cao:** nâng dải trên 2 kHz lên gần 4 dB, mô phỏng ảnh hưởng của đầu người.
- **Tầng hai — thông cao:** −3,9 dB ở 50 Hz, −13,3 dB ở 20 Hz.
- **Sau hai tầng:** bình phương trung bình như RMS, rồi cộng các kênh có trọng số.



(a) Đáp ứng của thước K

(b) Cùng biên độ đỉnh −20 dBFS

Hình 7.4 — Đáp ứng của thước K, và hệ quả khi đo

Độ to theo LUFS

$$z_i = \frac{1}{T}\sum_n y_i^2[n]\,; \qquad L_K = -0{,}691 + 10\log_{10}\sum_i G_i z_i$$

(7.2)

- zi là bình phương trung bình của kênh i sau thước K, trên khoảng thời gian T.
- Gi là trọng số kênh: trái, phải, giữa lấy 1; hai kênh vòng lấy 1,41.
- Hằng số −0,691 hiệu chỉnh để hình sin 997 Hz đọc đúng theo quy ước của chuẩn.
- 1 lu bằng đúng 1 dB; lu dùng cho hiệu hai mức, LUFS dùng cho mức tuyệt đối.

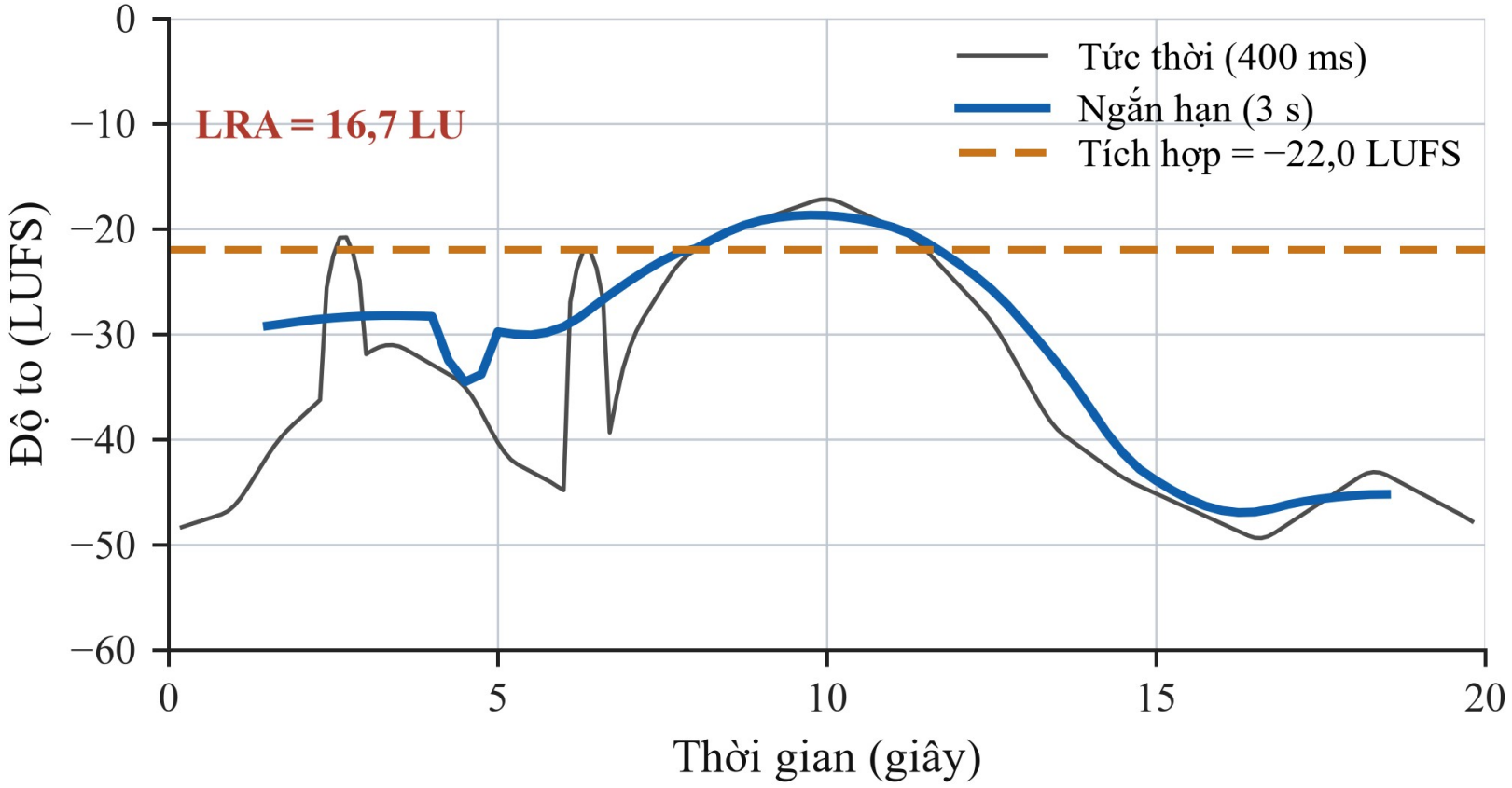
Tám hình sin cùng biên độ, LUFS khác nhau

RMS của tám hình sin bằng nhau chính xác: −23,0 dBFS. LUFS thì không: hình sin 40 Hz đọc −29,3 còn hình sin 8 kHz đọc −19,7. Chênh 9,6 decibel, đúng chiều tai cảm nhận. Đó là toàn bộ khác biệt giữa RMS và LUFS.

Ba thang thời gian của phép đo độ to

Tên	Cửa sổ	Dùng để làm gì
Tức thời (momentary)	400 ms	Dò chỗ đột ngột quá to; nhảy theo từng nốt nên khó đọc
Ngắn hạn (short-term)	3 s	Mượt hơn; người dựng nhìn khi cân các đoạn với nhau
Tích hợp (integrated)	Cả tệp, có cổng	Một con số cho cả tệp; nền tảng phát nhìn vào số này

Ba đường độ to trên cùng một đoạn nhạc cổ điển



Hình 7.5 — Ba đường độ to trên cùng một đoạn nhạc cổ điển

Đọc con số LRA

LRA	Nghĩa là	Thường gặp ở
Dưới 3 lu	Gần như phẳng, đã bị nén rất mạnh	Nhạc pop hiện đại, quảng cáo phát thanh
4 – 8 lu	Biến thiên vừa phải	Podcast, phim truyền hình, nhạc nhẹ
9 – 20 lu	Dải động rộng	Nhạc cổ điển, phim rạp
Trên 20 lu	Rất rộng; cần kiểm xem có phải lỗi dựng không	Bản thu hoà nhạc chưa xử lý

LRA là khoảng giữa phân vị 10 % và 95 % của độ to ngắn hạn, sau khi bỏ các khối quá nhỏ.

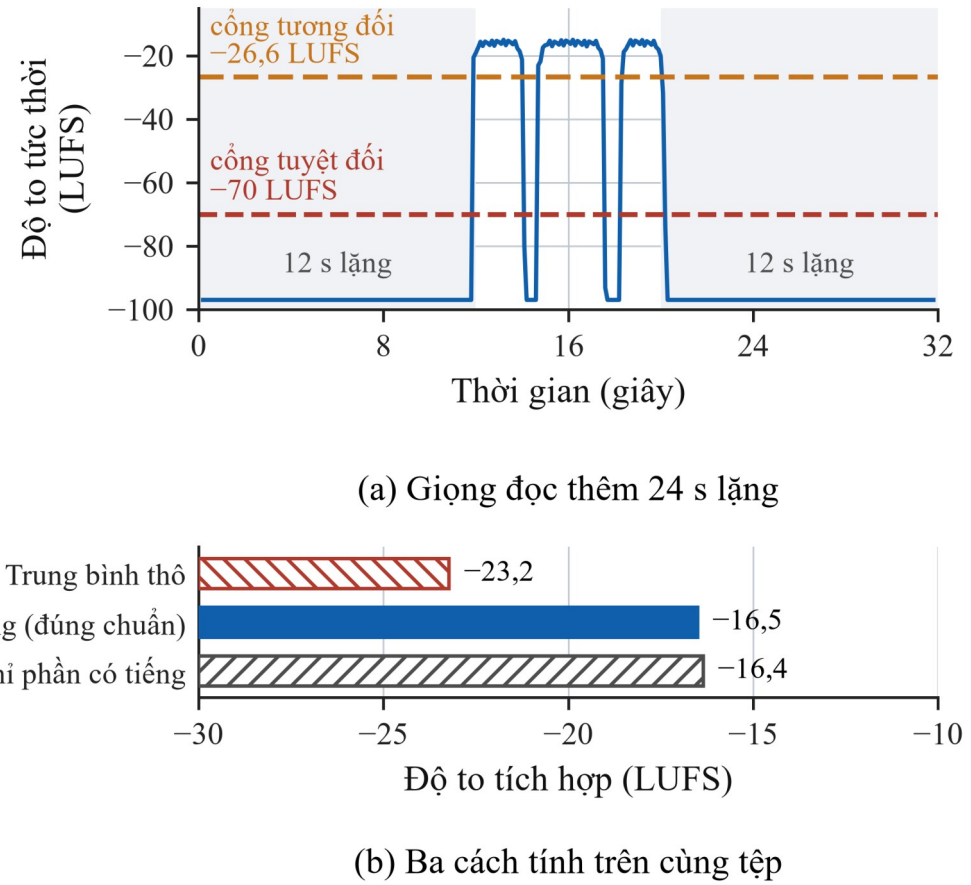
Vì sao cần cổng khi đo tích hợp

$$\Gamma_a = -70\ \text{LUFS}; \qquad \Gamma_r = L_K(\{j : l_j > \Gamma_a\}) - 10 \tag{7.3}$$

- Đoạn im lặng dài kéo con số tụt xuống; podcast là thể loại chịu ảnh hưởng nhiều nhất.
- **Cổng tuyệt đối:** bỏ mọi khối dưới −70 LUFS.
- **Cổng tương đối:** tính trung bình các khối còn lại, rồi bỏ khối thấp hơn mức ấy 10 lu.

Giọng đọc thêm 24 giây lặng, tính ba cách

- Trung bình thô: −23,2 LUFS.
- Có hai cổng: −16,5 LUFS.
- Chỉ đo phần có tiếng: −16,4 LUFS.
- Hai cổng chỉ lệch 0,1 decibel so với khi tự tay cắt phần lặng.



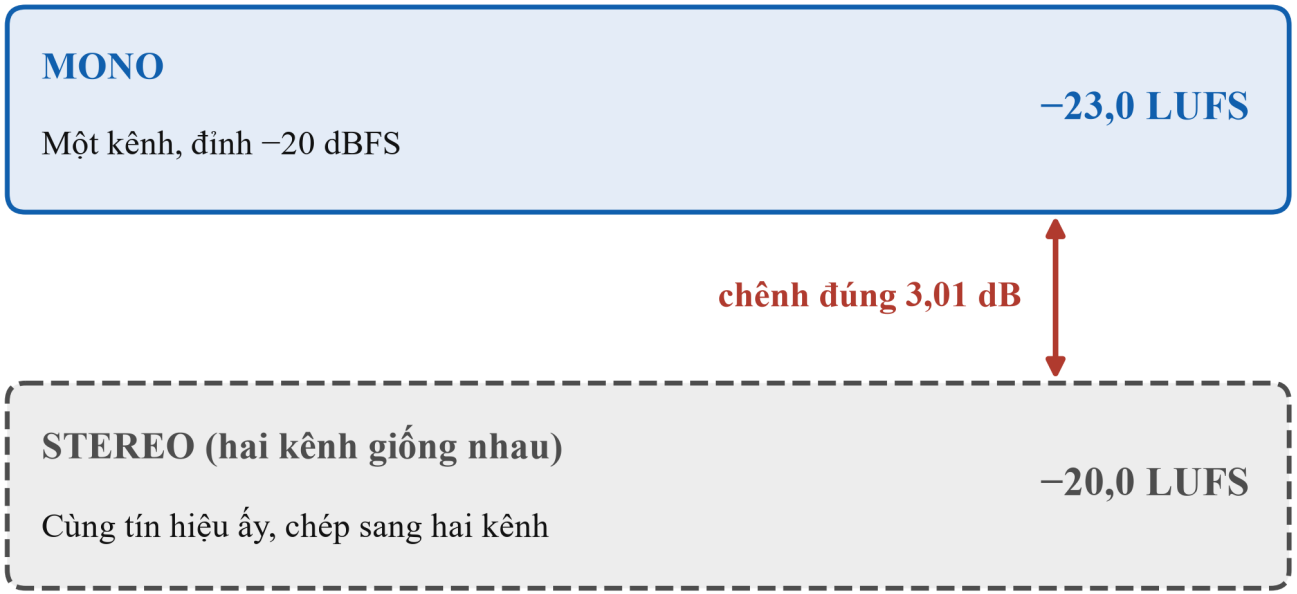
Hình 7.6 — Cùng một giọng đọc, thêm 24 giây lặng, tính ba cách

Cách kiểm cổng của một phần mềm đo

Đo một tệp có 20 giây im lặng ở đầu, rồi đo lại sau khi cắt phần lặng. Nếu hai kết quả lệch quá 0,3 dB thì phần mềm không dùng cổng đúng chuẩn. Hàm do_to_tich_hop trong sigtools thực hiện đúng hai cổng của BS.1770-4.

Mono và stereo cho hai con số LUFS

- Các kênh được cộng vào nhau trước khi lấy lô-ga-rít.
- Chép mono sang hai kênh làm tổng tăng gấp đôi: cộng thêm đúng 3,01 decibel.
- Luôn ghi rõ mono hay stereo khi báo LUFS.
- Không so LUFS của tệp mono với tệp stereo; chuyển về cùng số kênh trước.



Cùng một tiếng, cùng biên độ, cùng mức đỉnh, mà con số LUFS khác nhau 3 dB chỉ vì số kênh.

Khi giao nộp phải ghi rõ tệp mono hay stereo, và không so LUFS của tệp mono với LUFS của tệp stereo.

Hình 7.7 — Cùng một tín hiệu, mono và stereo, hai con số LUFS

Đo ba thước trên một tệp

```
import numpy as np, soundfile as sf
import am

x, sr = sf.read("ban_dung.wav")

print("Đỉnh      :", am.peak_dbfs(x),        "dBFS")
print("True peak :", am.true_peak_dbtp(x),   "dBTP")
print("RMS       :", am.rms_dbfs(x),         "dBFS")
print("LUFS      :", am.do_to_tich_hop(x, sr))
print("LRA       :", am.dai_do_to(x, sr),    "LU")

# bốn dòng này nên có trong MỌI kịch bản giao nộp
```

Năm con số ấy là hồ sơ đầy đủ về mức của một tệp: in ra trước và sau mọi xử lý.

MỤC 7.3

Chuẩn hoá theo độ to

- True peak là ràng buộc bắt buộc, vì sóng thật có thể vượt mẫu cao nhất.
- Mức đích phụ thuộc nơi giao nộp; không có một mức chung.
- Chuẩn hoá đủ bốn bước, và khi trần chặn lại thì phải chọn.

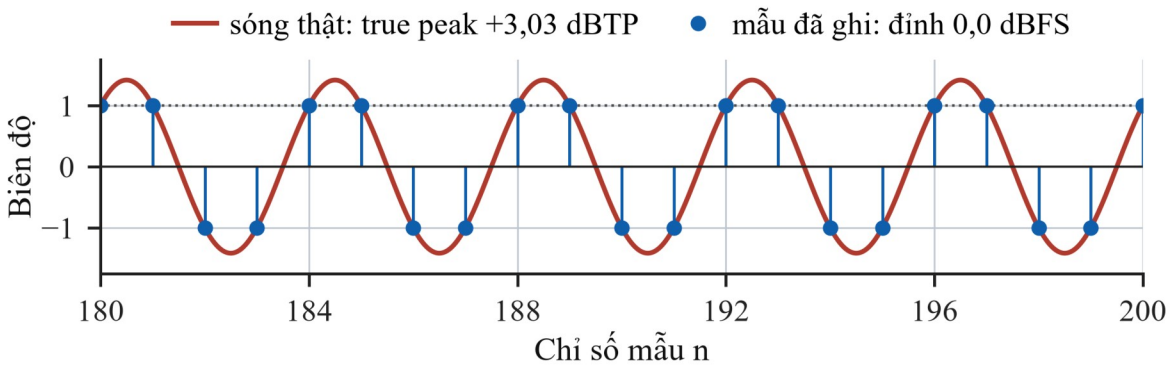
Vì sao trần nằm dưới 0

$$\mathrm{TP} = 20 \log_{10} \max_{m} \left| x_{\uparrow M}[m] \right|, \qquad M \geq 4$$

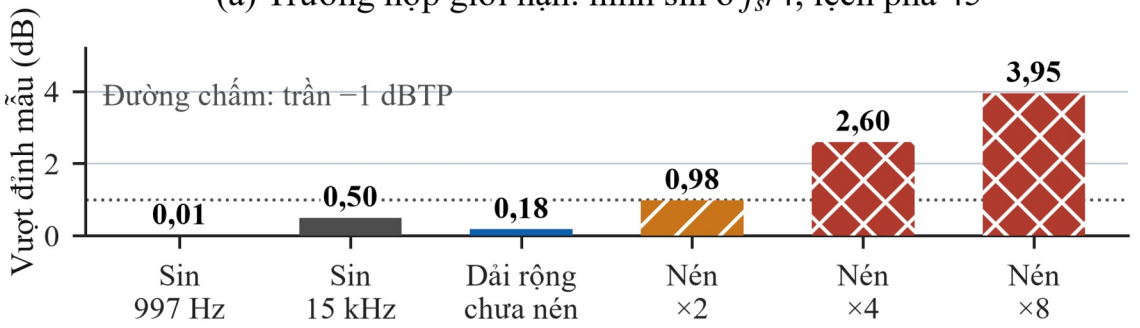
(7.4)

- Đỉnh mẫu chỉ nhìn các mẫu đã ghi.
- Bộ chuyển số–tương tự dựng sóng liên tục đi qua các mẫu; sóng có thể vượt mẫu cao nhất.
- Đo theo BS.1770: lấy mẫu quá lên ít nhất bốn lần rồi tìm giá trị lớn nhất.
- Kết quả ghi bằng dBTP (decibel đỉnh thật, True Peak).

Đỉnh mẫu và true peak trên sáu loại tư liệu



(a) Trường hợp giới hạn: hình sin ở $f_s/4$, lệch pha 45°



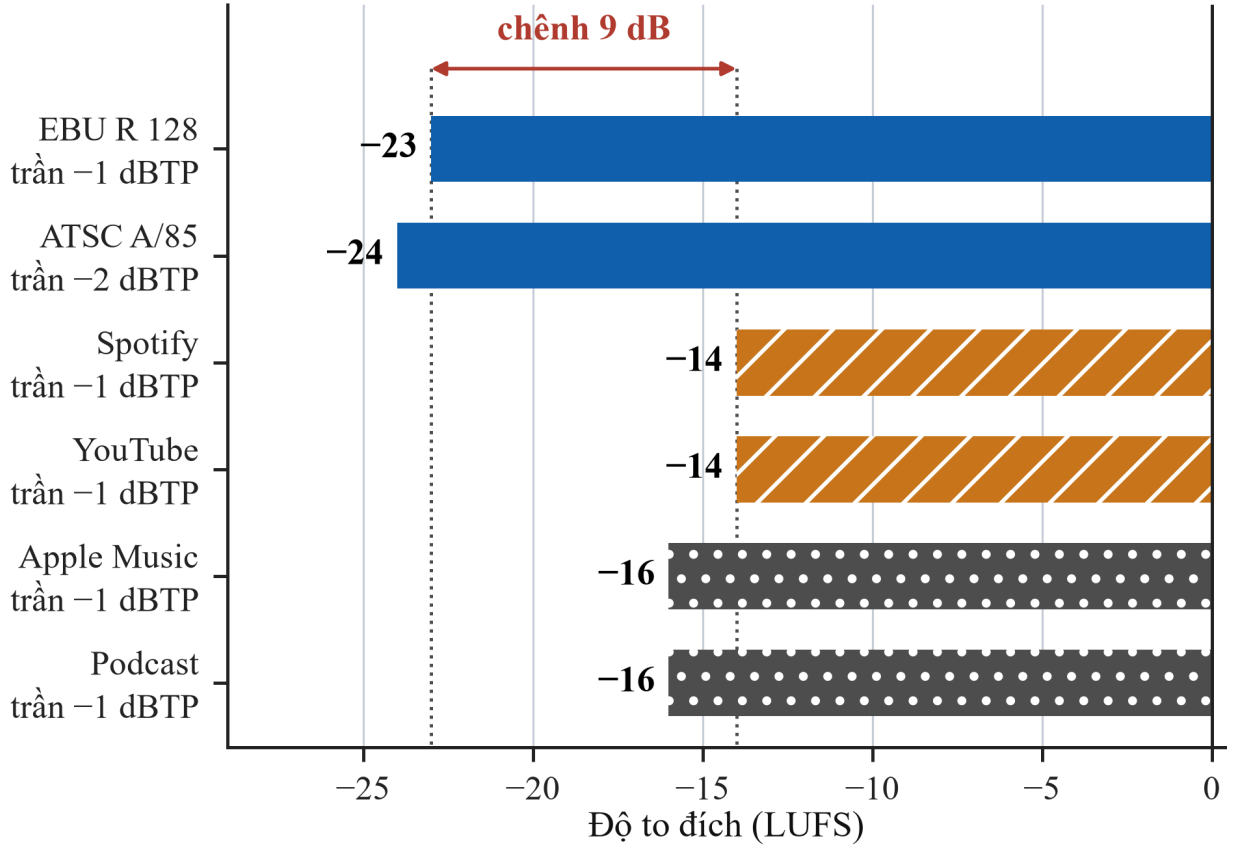
(b) Sáu loại tư liệu cùng chuẩn hoá về 0,0 dBFS

Hình 7.8 — Đỉnh mẫu và true peak trên sáu loại tư liệu

Vì sao trần là −1 chứ không phải 0

Mọi chuẩn phát sóng và nền tảng trực tuyến đặt trần dưới 0 dBFS, thường là −1 dBTP; phát thanh Bắc Mỹ là −2 dBTP. Lý do một: chừa chỗ cho phần tăng giữa hai mẫu. Lý do hai: bộ mã hoá MP3, AAC, Opus làm đỉnh dịch đi, nên tệp đúng 0,0 dBFS có thể vượt và méo.

Mức đích và trần của các nơi giao nộp



Hình 7.9 — Mức đích và trần của sáu nơi giao nộp

Các chuẩn thường gặp

Nơi giao nộp	LUFS đích	Trần	Vì sao là con số ấy
EBU R 128 (phát thanh, TH châu Âu)	−23	−1 dBTP	Chừa dải động cho tiếng động lớn
ATSC A/85 (truyền hình Bắc Mỹ)	−24	−2 dBTP	Đường truyền cũ kém chịu đỉnh
Spotify, YouTube	−14	−1 dBTP	Người nghe ở nơi ồn, tai nghe điện thoại
Apple Music	−16	−1 dBTP	Cân bằng giữa độ to và dải động
Podcast (khuyến nghị chung)	−16	−1 dBTP	Chủ yếu lời nói, nghe khi di chuyển

Chênh giữa −23 và −14 là chín decibel: nộp sai chuẩn thì bản dựng bị nghe là nhỏ hẳn so với các bài khác.

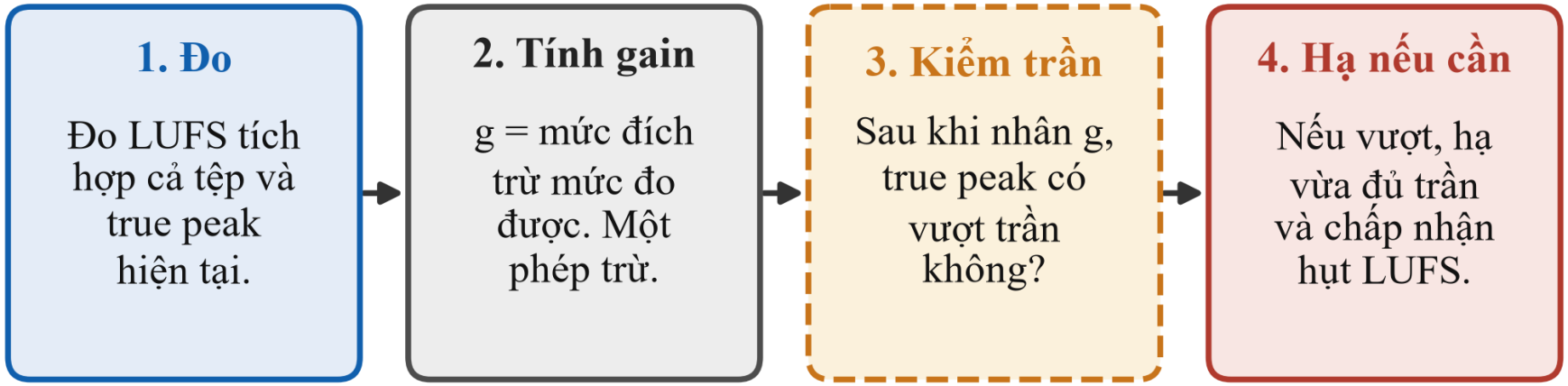
Lượng gain cần cộng

$$g_{\mathrm{dB}} = L_{\text{đích}} - L_{\text{đo}}\,; \qquad g_{\text{cuối}} = \min\left(g_{\mathrm{dB}},\ \mathrm{TP}_{\text{trần}} - \mathrm{TP}_{\text{đo}}\right)$$

(7.5)

- gdB là hiệu giữa mức đích và mức đo được.
- Nếu cộng chừng ấy mà true peak vượt trần thì lấy giá trị nhỏ hơn trong hai lựa chọn.
- Chấp nhận độ to hụt so với đích.

Bốn bước chuẩn hoá theo độ to



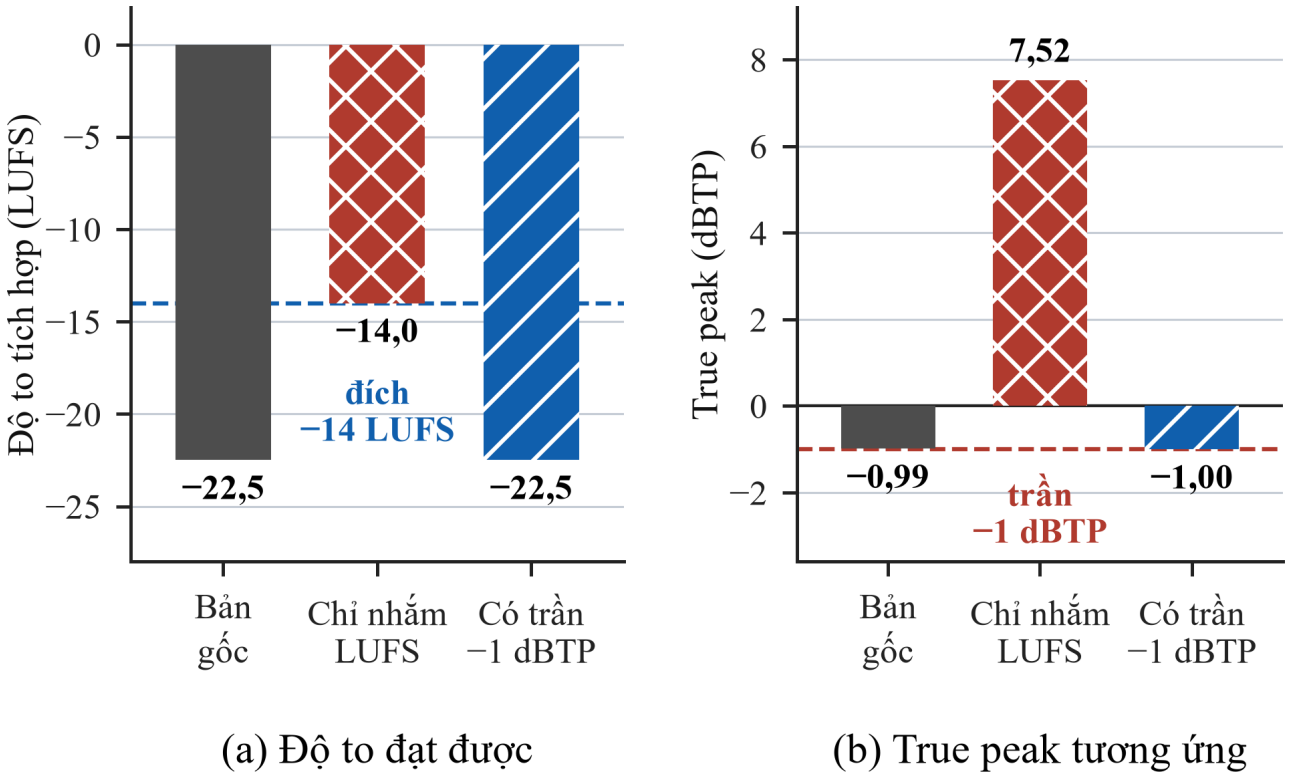
Bước 4 hay bị bỏ qua; bỏ qua thì tệp giao nộp bị nền tảng từ chối hoặc bị hạ mức tự động ở khâu phát.

Nếu phải hạ nhiều hơn 1–2 dB ở bước 4 thì vấn đề nằm ở chỗ khác: bản dựng đã bị nén quá tay từ trước. Luôn chuẩn hoá ở bước cuối.

Hình 7.10 — Bốn bước chuẩn hoá theo độ to

Nhạc cổ điển đưa về −14 LUFS

- Đỉnh cao hơn độ to trung bình 21,8 decibel; đại lượng này là PLR (Peak-to-Loudness Ratio – tỉ số đỉnh trên độ to).
- Không kiểm trần: true peak tăng lên +7,52 dBTP, méo mạnh.
- Đặt trần −1 dBTP: độ to dừng ở −22,5 LUFS, hụt 8,5 decibel.

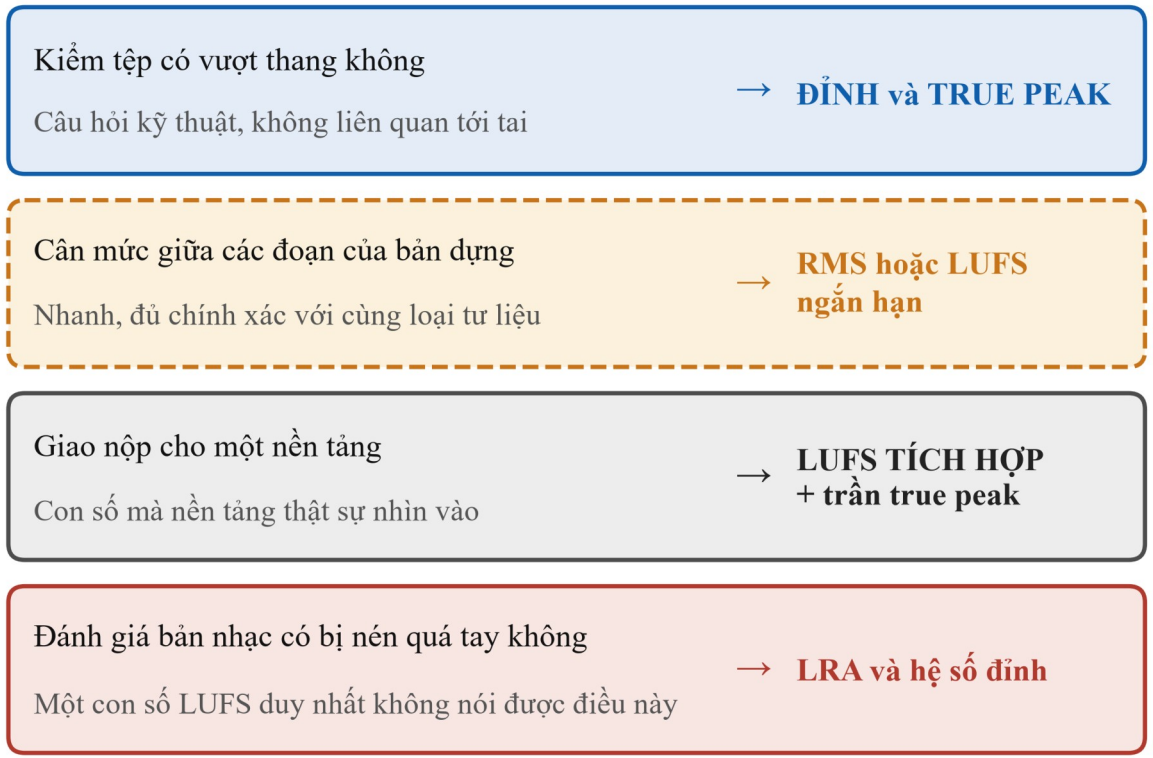


Hình 7.11 — Một bản nhạc cổ điển đưa về −14 LUFS, có và không có trần

Phải chọn, và mọi chuẩn đều ưu tiên trần

Hụt 1–2 dB thì chấp nhận được, vì nền tảng tự chỉnh lại ở khâu phát. Hụt tới 8 dB thì vấn đề không nằm ở phép chuẩn hoá: hoặc bản dựng có PLR quá cao và cần nén nhẹ trước (chương 8), hoặc tư liệu không hợp nền tảng nhắm −14 LUFS.

Bốn tình huống và thước đo tương ứng



Dùng nhầm thước là nguyên nhân của phần lớn các bản giao nộp bị trả lại.

Hình 7.12 — Bốn tình huống và thước đo tương ứng

Chuẩn hoá đủ bốn bước

```
def chuan_hoa_lufs(x, dich_lufs=-23.0, tran_tp=-1.0, sr=44100):
    """Chuẩn hoá theo độ to, CÓ trần true peak."""
    # 1. đo
    hien = do_to_tich_hop(x, sr)
    # 2. tính gain
    g = dich_lufs - hien
    y = x * db_to_lin(g)
    # 3. kiểm trần
    tp = true_peak_dbtp(y)
    bi_chan = tp > tran_tp
    # 4. hạ thêm nếu cần
    if bi_chan:
        g += (tran_tp - tp)
        y = x * db_to_lin(g)
    return y, g, bi_chan


y, g, chan = chuan_hoa_lufs(x, -16.0, -1.0, sr)
if chan:
    print(f"Tran chan lai. Do to that: "
          f"{do_to_tich_hop(y, sr):.1f} LUFS")
```

Giá trị trả về thứ ba cho người gọi biết trần có chặn hay không.

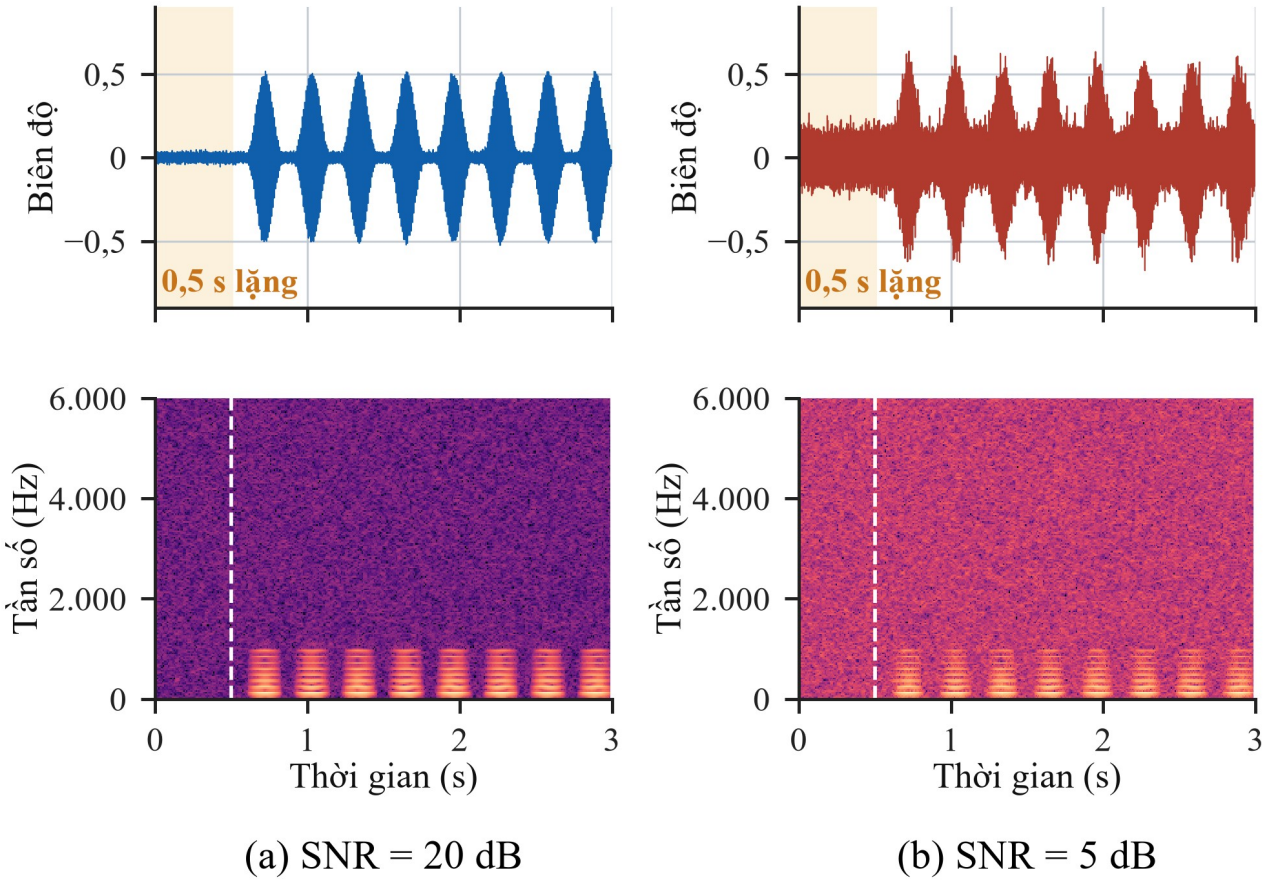
MỤC 7.4

Nhiễu và cách đo nó

- Chương này chỉ bàn nhiễu nền dừng: quạt, điều hoà, ù nguồn, tiếng rít tiền khuếch đại.
- SNR cho một con số cho cả bản thu; phổ nhiễu cho phân bố theo tần số.
- Cả hai đo được từ một đoạn lặng ngắn đầu tệp.

Cùng một giọng đọc ở hai mức nhiễu

- Nhiễu nền dừng có phổ gần như không đổi theo thời gian.
- Giả định này làm cho toàn bộ mục 7.5 hoạt động được.
- Tiếng nói chuyện bên cạnh, xe chạy qua, cửa đóng nằm ngoài phạm vi.
- Những thứ ấy thường phải cắt bỏ bằng tay.



Hình 7.13 — Cùng một giọng đọc ở hai mức nhiễu, nhìn hai cách

Tỉ số tín hiệu trên nhiễu

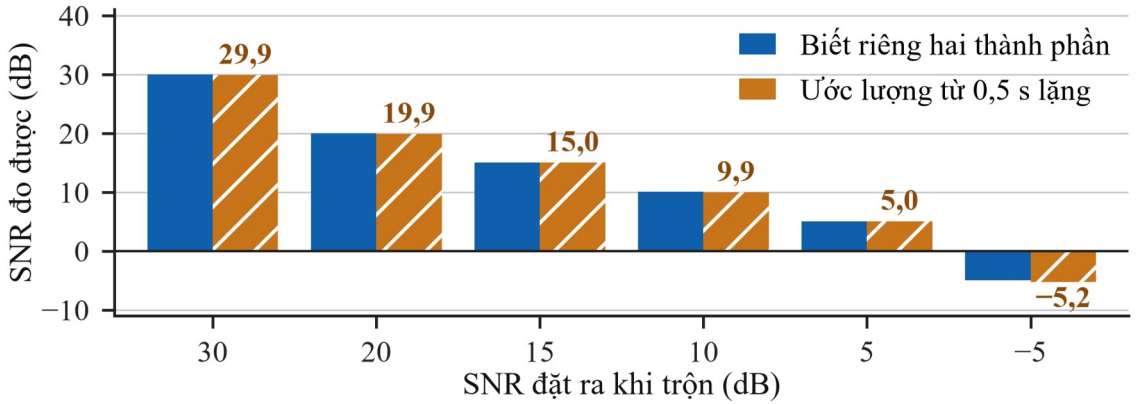
$$\mathrm{SNR}_{\mathrm{dB}} = 10\log_{10}\frac{P_{\text{tín}}}{P_{\text{nhiễu}}} = 10\log_{10}\frac{P_{\text{tổng}} - P_{\text{nhiễu}}}{P_{\text{nhiễu}}}$$

(7.6)

- Vế đầu là định nghĩa: tỉ số công suất tín hiệu trên công suất nhiễu, tính bằng decibel.
- Vế sau là cách tính khi chỉ có bản đã trộn.
- Đo công suất nhiễu trên một đoạn lặng, rồi trừ khỏi công suất tổng để ra công suất tín hiệu.

Hai cách đo SNR

- Ước lượng từ nửa giây im lặng chính xác tới một phần mười decibel.
- So sánh với khi biết riêng hai thành phần.
- Chỉ cần nửa giây yên tĩnh là đo được chất lượng cả bản thu.



(a) Hai cách đo cho cùng một kết quả

Trên 40 dB	Phòng thu đã xử lý âm
25 – 40 dB	Bản thu tốt trong phòng yên
15 – 25 dB	Thu tại nhà, có quạt hoặc điều hoà
5 – 15 dB	Thu ngoài trời, quán, hội trường
Dưới 5 dB	Lời nói đã khó nghe rõ

(b) Đọc con số SNR: đo trước khi khử nhiễu

Hình 7.14 — Hai cách đo SNR, và cách đọc con số ấy

Thói quen đáng tập từ hôm nay

Mỗi lần bấm máy, thu dư ba tới năm giây yên tĩnh trước khi bắt đầu nói hoặc chơi nhạc. Ba giây ấy dùng để đo SNR và ước lượng phổ nhiễu. Không có chúng, cả mục 7.5 không dùng được.

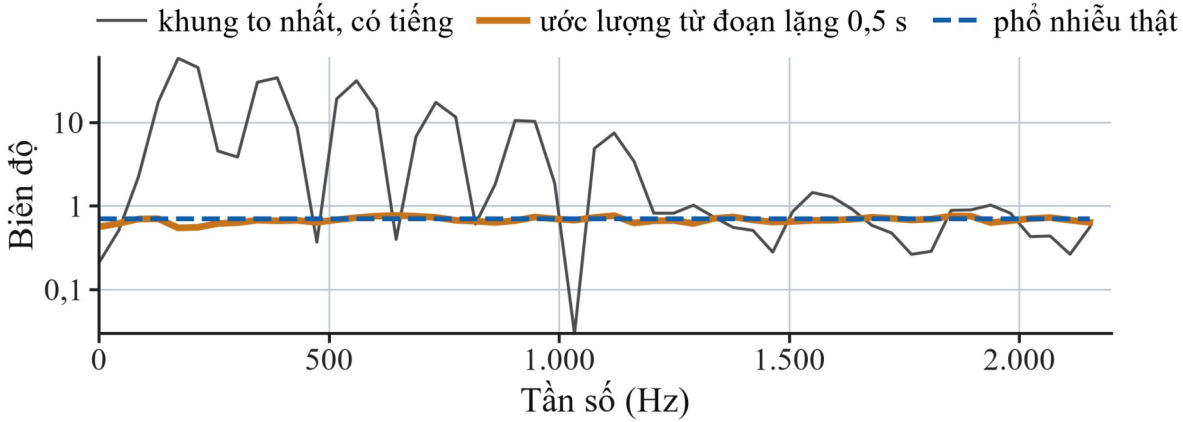
Ngưỡng SNR và ý nghĩa thực tế

SNR	Tình huống điển hình	Nên làm gì
Trên 40 dB	Phòng thu đã xử lý âm học	Không cần khử
25 – 40 dB	Phòng yên tĩnh, bản thu tốt	Không cần khử; khử chỉ thêm nhiễu nhạc
15 – 25 dB	Thu tại nhà, có quạt, điều hoà	Khử nhẹ, hệ số trừ vượt 1 – 1,5
5 – 15 dB	Thu ngoài trời, quán, hội trường	Khử vừa, hệ số khoảng 2, san theo khung
Dưới 5 dB	Lời nói đã khó nghe rõ	Cân nhắc thu lại

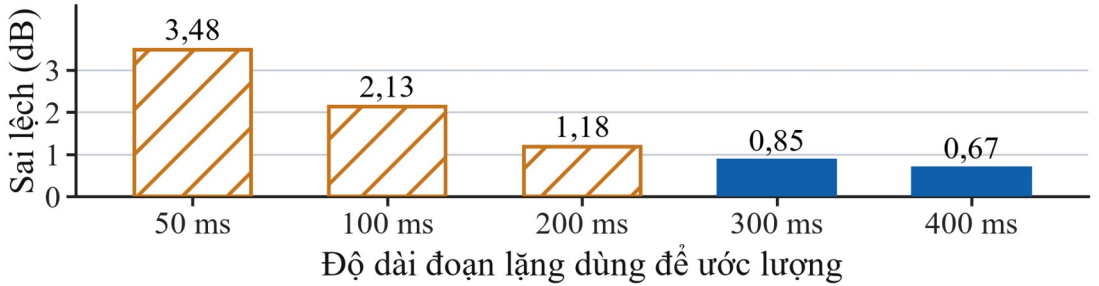
Câu hỏi đầu tiên của mọi việc khử nhiễu là có nên khử hay không.

Phổ nhiễu ước lượng từ đoạn lặng

- Khử nhiễu cần cả một phổ, không chỉ một con số.
- Phép ước lượng bám sát phổ thật dù chỉ nhìn nửa giây đầu tệp.
- Đoạn lặng 50 ms: sai lệch trung bình 3,48 dB.
- Đoạn lặng 400 ms: sai lệch trung bình 0,67 dB.



(a) Phổ nhiễu ước lượng bám sát phổ thật



(b) Đoạn lặng càng dài, ước lượng càng chuẩn

Hình 7.15 — Phổ nhiễu ước lượng từ đoạn lặng, so với phổ nhiễu thật

Ước lượng phổ nhiễu và đo SNR

```
import am

x, sr = am.doc_wav("thu_tho.wav")

# doan lang: 0,0 den 0,45 giay
snr = am.snr_tu_doan_lang(x, 0.0, 0.45, sr)
print(f"SNR = {snr:.1f} dB")

if snr < 25:
    N = am.pho_nhieu_uoc_luong(x, 0.0, 0.45, sr)
    # N la mot mang bien do, dai n_fft // 2 + 1
else:
    print("Da du sach — khong can khu")
```

Giả định duy nhất, và chỗ nó sụp

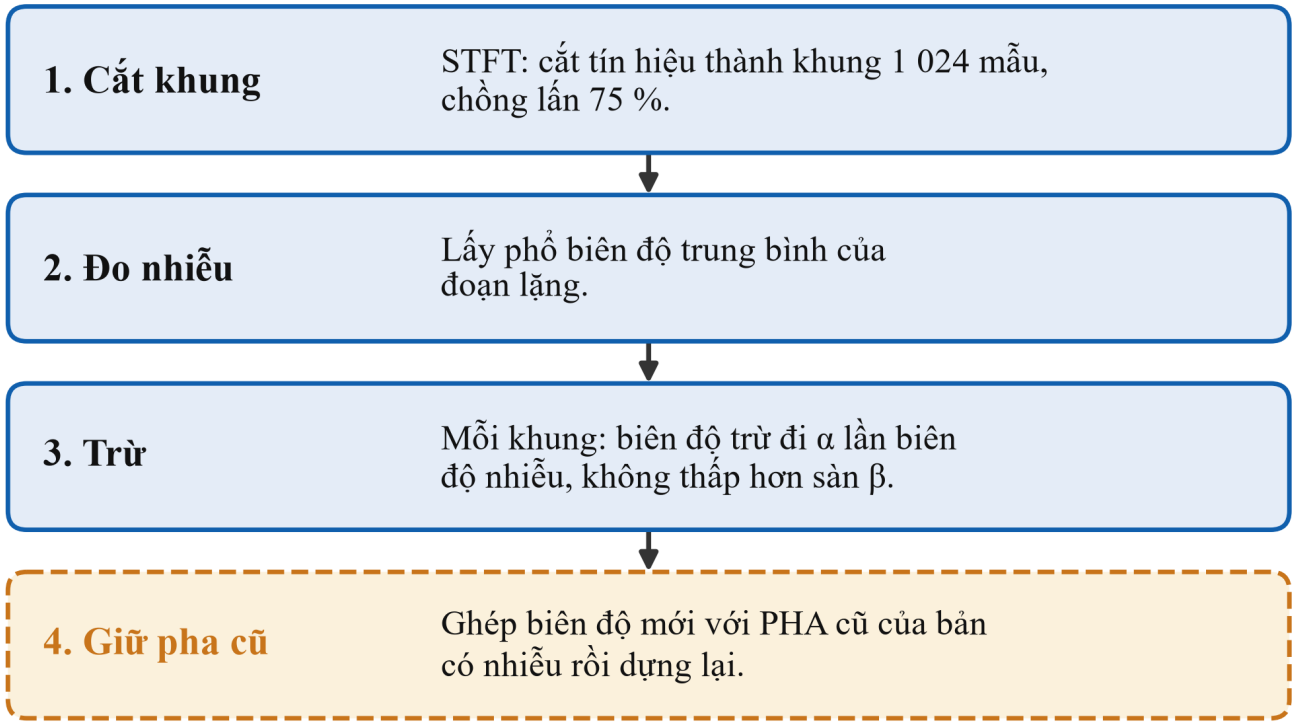
Phổ nhiễu không đổi theo thời gian: nhờ vậy phổ đo trên nửa giây đầu dùng được cho cả tệp dài mười phút. Giả định sai với tiếng nói chuyện, xe chạy qua, mưa lúc to lúc nhỏ. Cách kiểm: so phổ ở đầu và cuối tệp; lệch quá 3 dB thì nhiễu không dừng.

MỤC 7.5

Khử nhiễu bằng trừ phổ

- Ở mỗi khung, lấy biên độ đo được trừ biên độ nhiễu; phần còn lại là tín hiệu.
- Hai tham số phải chọn: alpha (trừ vượt) và beta (sàn phổ).
- Pha của bản có nhiễu được giữ nguyên khi dựng lại.

Bốn bước của phép trừ phổ



Bước 4 là chỗ đáng ngạc nhiên nhất: pha không được xử lý gì cả. Tai ít nhạy với sai pha trong dải bị nhiễu lấn (xem mục 7.5.3).

Hình 7.16 — Bốn bước của phép trừ phổ

Phép trừ phổ

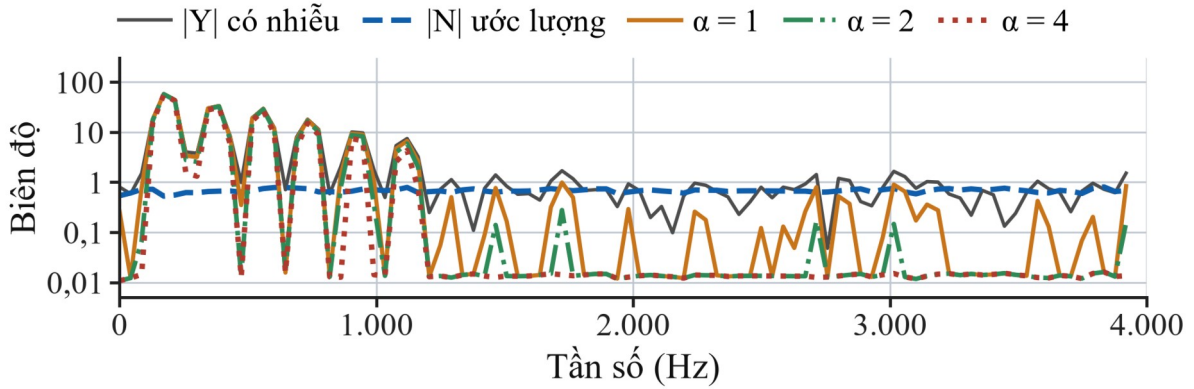
$$|\hat{S}[k]| = \max\left(|Y[k]| - \alpha|\hat{N}[k]|,\ \beta|\hat{N}[k]|\right)$$

(7.7)

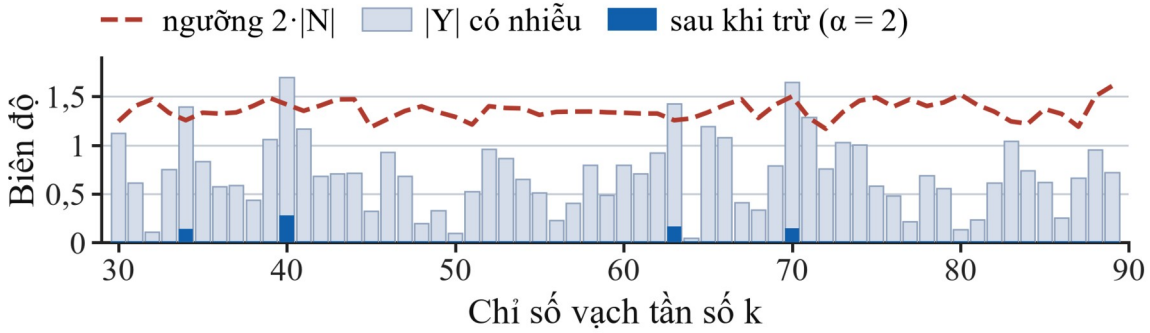
- Y[k] là phổ khung đang xét, N̂[k] là phổ nhiễu ước lượng, Ŝ[k] là phổ tín hiệu ước lượng.
- **Alpha — hệ số trừ vượt:** trừ nhiều hơn ước lượng để bù sai số; alpha bằng 2 là trừ gấp đôi.
- **Beta — sàn phổ:** không cho biên độ tụt dưới beta lần mức nhiễu, tránh những vạch về đúng không.

Chỗ nhiễu nhạc sinh ra

- Trong một khung, chỉ vạch vượt ngưỡng alpha·|N̂| sống sót: bốn trên sáu mươi vạch.
- Năm mươi sáu vạch còn lại bị đưa xuống sàn.
- Khung sau, một tập bốn vạch khác sống sót, vì nhiễu ngẫu nhiên.
- Mỗi khung một tập vạch nhấp nháy: hậu quả ở mục 7.6.



(a) Một khung phổ trừ ở ba mức α



(b) Chỉ 4 trong 60 vạch sống sót; 56 vạch bị nén xuống sàn

Hình 7.17 — Một khung phổ, ba mức trừ, và chỗ nhiễu nhạc sinh ra

Biên độ mới, pha cũ

$$\hat{s}[n] = \mathrm{ISTFT}\left(\left|\hat{S}[k]\right| \cdot e^{j\varphi_Y[k]} \right) \tag{7.8}$$

- **Lý do một:** pha của nhiễu ngẫu nhiên hoàn toàn, không có gì để lấy trung bình nên ước lượng rất khó.
- **Lý do hai:** ở vạch nhiễu lấn át, tai không phân biệt được pha đúng sai; ở vạch tín hiệu mạnh, pha đo được vốn gần đúng.
- Pha quan trọng, nhưng không quan trọng như nhau ở mọi chỗ (mục 4.5, 6.5).

Toàn bộ thuật toán, và cách dùng nó

```
def tru_pho(x, pho_nhieu, alpha=2.0, beta=0.02,
            n_fft=1024, hop=256):
    X, i, w, tt = _stft(x, n_fft, hop)
    bien = np.abs(X)
    pha  = np.angle(X)        # giữ nguyên
    N = pho_nhieu[None, :]
    con = np.maximum(bien - alpha * N, beta * N)
    return _istft(con * np.exp(1j * pha),
                  i, w, tt, n_fft, hop)
```

(a) Toàn bộ thuật toán (chín dòng)

```
# 1. Đo SNR trước; trên 25 dB thì dừng
snr = snr_tu_doan_lang(x, 0.0, 0.45, sr)
if snr > 25: raise SystemExit("đủ sạch")
# 2. Ước lượng phổ nhiễu từ đoạn lặng
N = pho_nhieu_uoc_luong(x, 0.0, 0.45, sr)
# 3. Trừ phổ, rồi đo lại cả hai mặt
y = tru_pho(x, N, alpha=2.0)
print(san_nhieu_nen(y, sr))      # nền im đi
print(do_nhieu_nhac(y, 0.0, 0.45))  # cái giá
```

(b) Cách dùng và hai phép đo bắt buộc

Hình 7.18 — Toàn bộ thuật toán, và cách dùng nó

Trừ phổ — bản đầy đủ

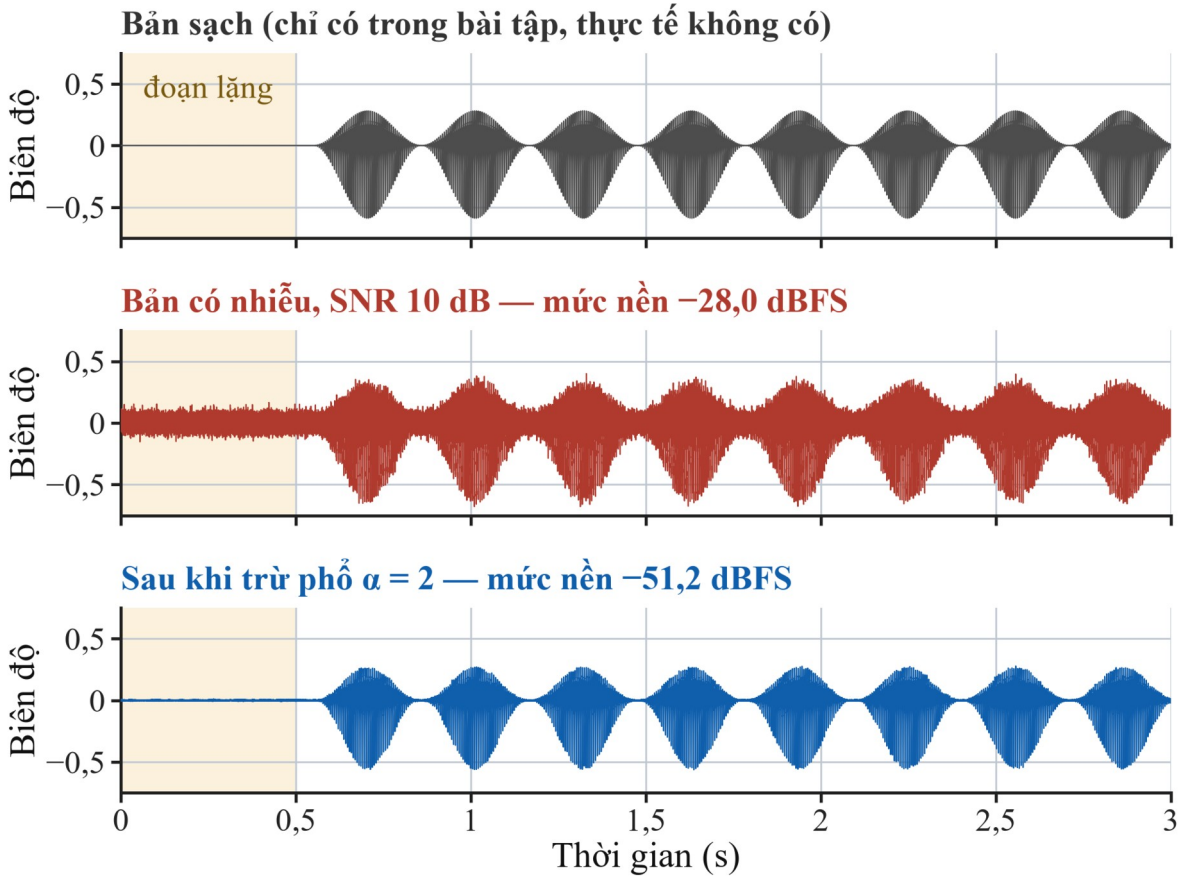
```
def tru_pho(x, pho_nhieu, alpha=2.0, beta=0.02,
            n_fft=1024, hop=256):
    """Tru pho: bien do tru di alpha lan bien do nhieu."""
    X, i, w, tt = _stft(x, n_fft, hop)
    bien = np.abs(X)
    pha  = np.angle(X)              # GIU NGUYEN
    N = np.asarray(pho_nhieu, dtype=float)[None, :]
    con = np.maximum(bien - alpha * N, beta * N)
    return _istft(con * np.exp(1j * pha), i, w, tt, n_fft, hop)


# dung: bao gio cung do TRUOC va do LAI SAU
print("SNR truoc :", snr_tu_doan_lang(on, 0.0, 0.45, sr))
N = pho_nhieu_uoc_luong(on, 0.0, 0.45, sr)
y = tru_pho(on, N, alpha=2.0)
print("Nen sau   :", san_nhieu_nen(y, sr), "dBFS")
print("Nhieu nhac:", do_nhieu_nhac(y, 0.0, 0.45, sr))
```

Ba dòng cuối quan trọng hơn thuật toán: đo cả hai mặt, nền im đi bao nhiêu và trả giá bao nhiêu.

Bản sạch, bản có nhiễu, bản đã trừ phổ

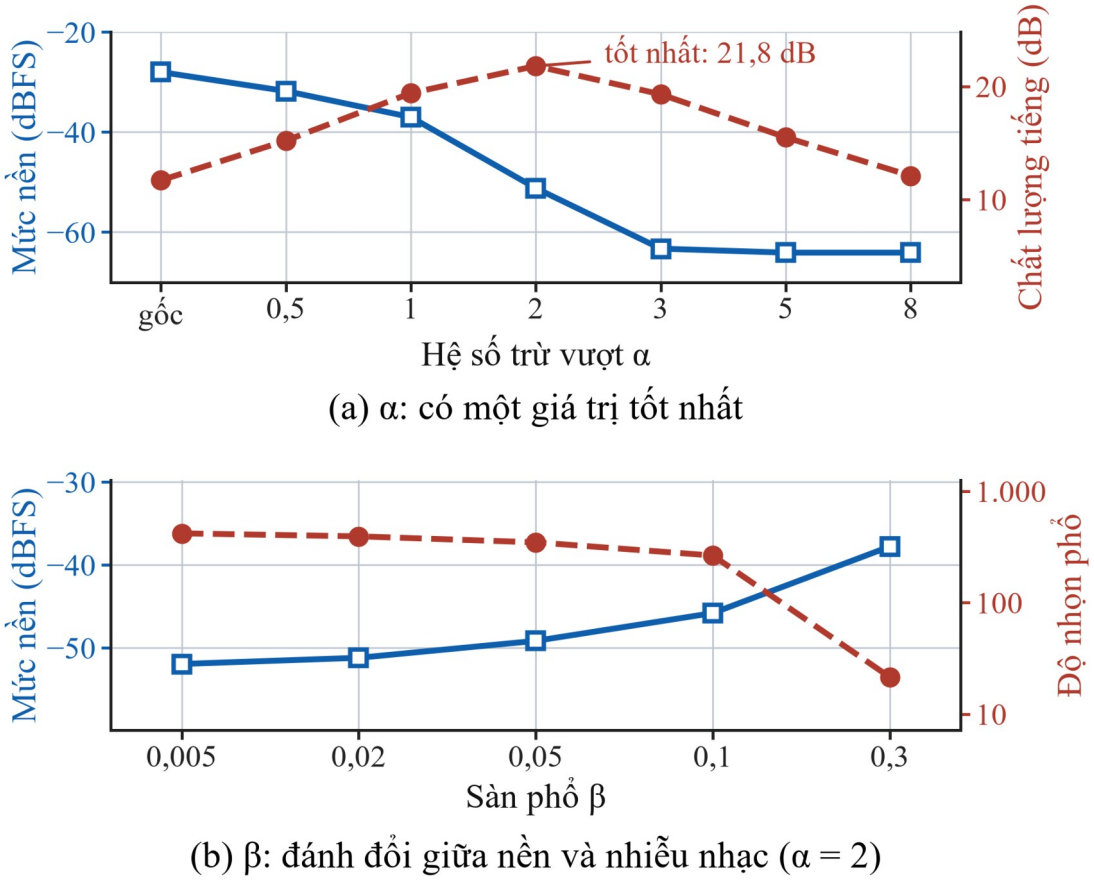
- Dạng sóng có vẻ hoàn hảo: nửa giây đầu phẳng lì.
- Mức nền hạ từ −28,0 xuống −51,2 dBFS, tức 23 decibel.
- Dạng sóng không cho biết cái còn lại nghe như thế nào.
- Đó là nội dung mục 7.6.



Hình 7.19 — Bản sạch, bản có nhiễu, và bản đã trừ phổ

Vai trò của alpha và của beta

- **Alpha có giá trị tốt nhất:** chất lượng tiếng cực đại 21,8 dB ở alpha bằng 2, tụt xuống 12,1 dB ở alpha bằng 8.
- **Beta là sự đánh đổi thuần tuý:** beta nhỏ thì nền im hơn nhưng nhiễu nhạc nhiều hơn; beta lớn thì ngược lại.



Hình 7.20 — Vai trò của alpha và của beta

Chọn hai tham số theo tình huống

Tình huống	alpha	beta	Vì sao
SNR trên 25 dB	—	—	Đừng khử
Lời nói, SNR 15 – 25 dB	1,0 – 1,5	0,05	Bớt chú ý tới nền, không xoá nền
Lời nói, SNR 5 – 15 dB	2,0	0,02	Vùng khử nhiễu có ích; nên san khung
Nhạc, mọi mức SNR	1,0 – 1,5	0,1	Giữ đuôi vang và nhạc cụ nhỏ
Chỉ cần nghe rõ lời	3,0	0,01	Chấp nhận nhiễu nhạc; chỉ cho tư liệu ghi chép

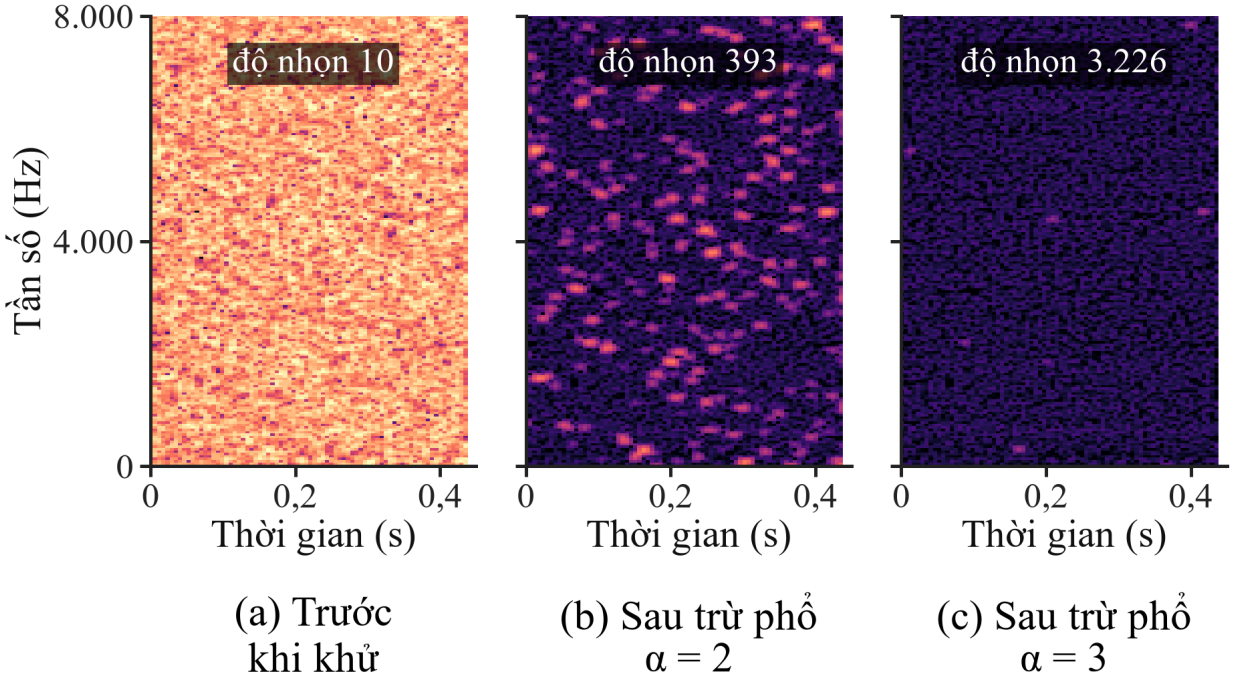
MỤC 7.6

Nhiễu nhạc — cái giá của trừ phổ

- Sau khi trừ phổ, nền không im lặng mà thành chuỗi tiếng lích rích như chuông gió.
- Độ nhọn của phổ công suất biến cảm nhận ấy thành một con số.
- Không cách khử nào thắng ở cả ba đại lượng: mức nền, chất lượng tiếng, nhiễu nhạc.

Nửa giây lặng nhìn bằng phổ đồ

- Nhiễu nền tự nhiên là một tấm màn đều.
- Sau trừ phổ, tấm màn tách thành những chấm sáng rời rạc.
- Mỗi vạch nhấp nháy là một tiếng lích rích ở đúng tần số của nó.
- Nguyên nhân: mỗi khung một tập vạch khác nhau vượt ngưỡng.



Hình 7.21 — Nửa giây lặng, nhìn bằng phổ đồ, trước và sau khi khử

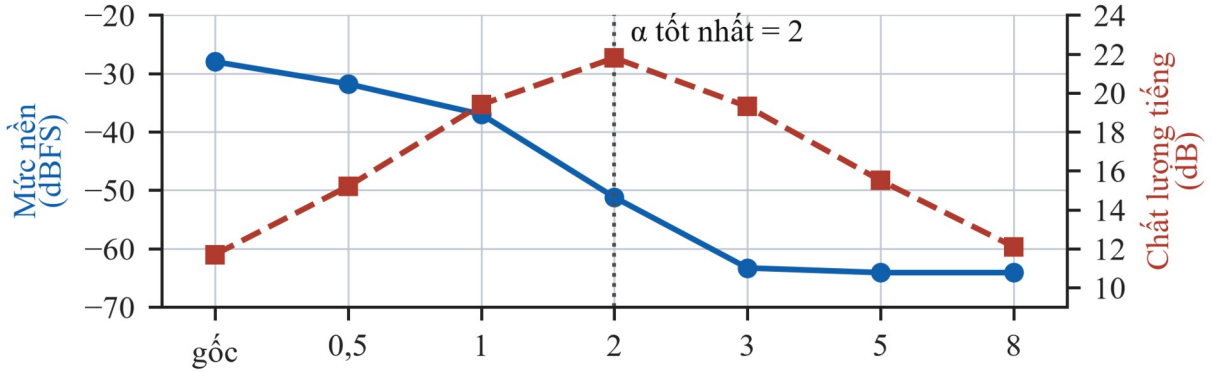
Độ nhọn của phổ công suất

$$\kappa = \frac{1}{K\sigma_P^4}\sum_k (P[k]-\mu_P)^4$$

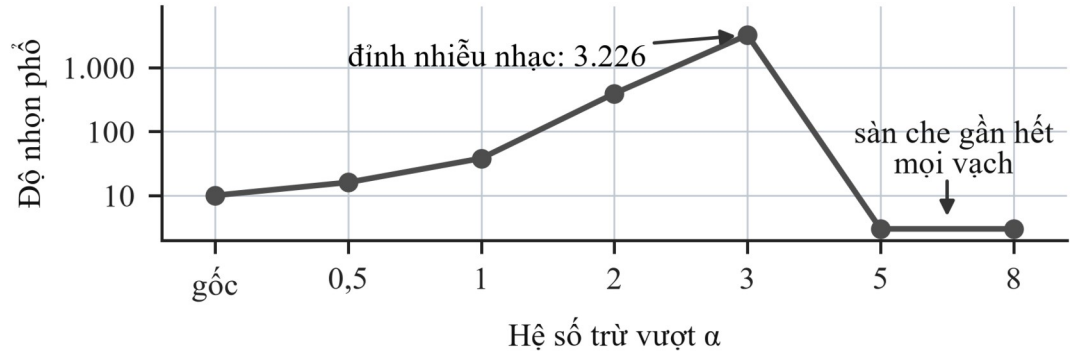
(7.9)

- Độ nhọn (kurtosis) của phổ công suất trong đoạn lặng; mu và sigma là trung bình và độ lệch chuẩn của P[k].
- Nhiễu nền tự nhiên trải đều nên độ nhọn thấp; nhiễu nhạc là những đỉnh cô lập nên độ nhọn cao.
- Ví dụ của bài: nhiễu tự nhiên có độ nhọn 10; sau trừ phổ với alpha bằng 3 lên 3 226.

Mức nền, chất lượng tiếng và nhiễu nhạc theo alpha



(a) Mức nền (đường liền) và chất lượng tiếng (đường đứt)



(b) Độ nhọn phổ ở đoạn lặng

Hình 7.22 — Mức nền, chất lượng tiếng và nhiễu nhạc theo alpha

Ba đại lượng theo hệ số trừ vượt alpha

alpha	Mức nền (dBFS)	Chất lượng tiếng (dB)	Độ nhọn phổ
Gốc (chưa khử)	−28,0	11,7	10
0,5	−31,8	15,2	16
1	−37,0	19,4	38
2	−51,2	21,8 (tốt nhất)	393
3	−63,3	19,3	3 226 (đỉnh)
5	−64,1	15,5	3
8	−64,1	12,1	3

Đọc ba cột cùng lúc

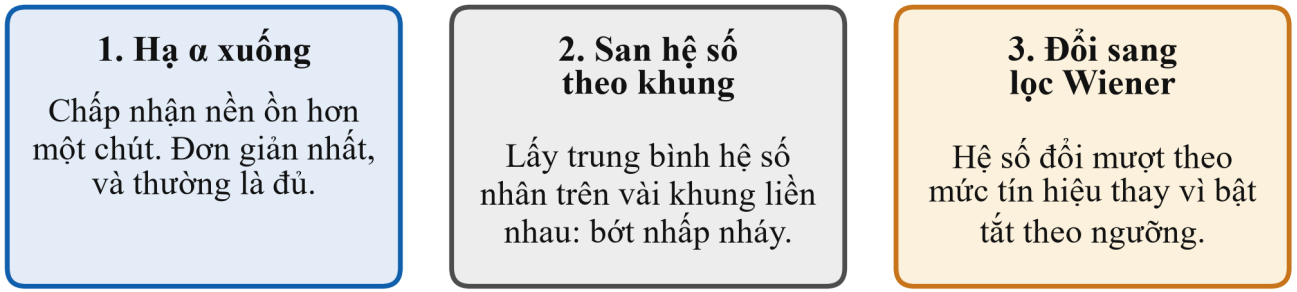
- **Mức nền:** giảm đều rồi bão hoà; từ alpha bằng 3 sàn beta chặn lại.
- **Chất lượng tiếng:** cực đại ở alpha bằng 2; ở alpha bằng 8 còn kém hơn khi chưa khử.
- **Nhiễu nhạc:** đỉnh ở alpha bằng 3; từ alpha bằng 5 độ nhọn về gần 3 vì sàn che cả vạch của tiếng.
- **Hệ quả:** phải nhìn cột chất lượng tiếng bên cạnh, không chỉ nhìn độ nhọn.

Mỗi phép xử lý có ít nhất hai mặt

Chỉ nhìn mức nền thì alpha bằng 8 là lựa chọn tuyệt vời. Mặt thứ hai thường khó đo hơn mặt thứ nhất, và việc của người kỹ thuật là tìm ra cách đo nó. Điều này lặp lại ở chương 8 với bộ nén, chương 10 với hiệu ứng, chương 12 với nén dữ liệu.

Ba cách giảm nhiễu nhạc, và cách thứ tư

- Cách thứ hai rẻ nhất: san hệ số nhân của mỗi vạch trên vài khung liền nhau.
- Hệ số bớt nhảy loạn; đổi lại tiếng nhoè nhẹ theo thời gian.
- Cách thứ ba: đổi công thức hệ số nhân sang liên tục (bộ lọc Wiener).



Cả ba đều đổi "nền im hơn" lấy "ít nhiễu nhạc hơn". Không cách nào cho không cái gì.

Cách thứ tư, cách tốt nhất, không nằm trong xử lý: thu lại cho sạch hơn. Ba mươi giây tắt điều hoà đáng giá hơn mọi thuật toán ở đây.

Hình 7.23 — Ba cách giảm nhiễu nhạc, và cách thứ tư

Trừ phổ có san hệ số theo khung

```
def tru_pho_san_khung(x, pho_nhieu, alpha=2.0, beta=0.02,
                      san=3, n_fft=1024, hop=256):
    X, i, w, tt = _stft(x, n_fft, hop)
    bien = np.abs(X)
    pha  = np.angle(X)
    N = np.asarray(pho_nhieu, dtype=float)[None, :]
    # he so nhan cua tung vach, tung khung
    he_so = (np.maximum(bien - alpha * N, beta * N)
             / np.maximum(bien, 1e-12))
    # SAN he so tren "san" khung lien nhau
    nhan = np.ones(int(san)) / float(san)
    he_so = sg.lfilter(nhan, 1.0, he_so, axis=0)
    return _istft(bien * he_so * np.exp(1j * pha),
                  i, w, tt, n_fft, hop)
```

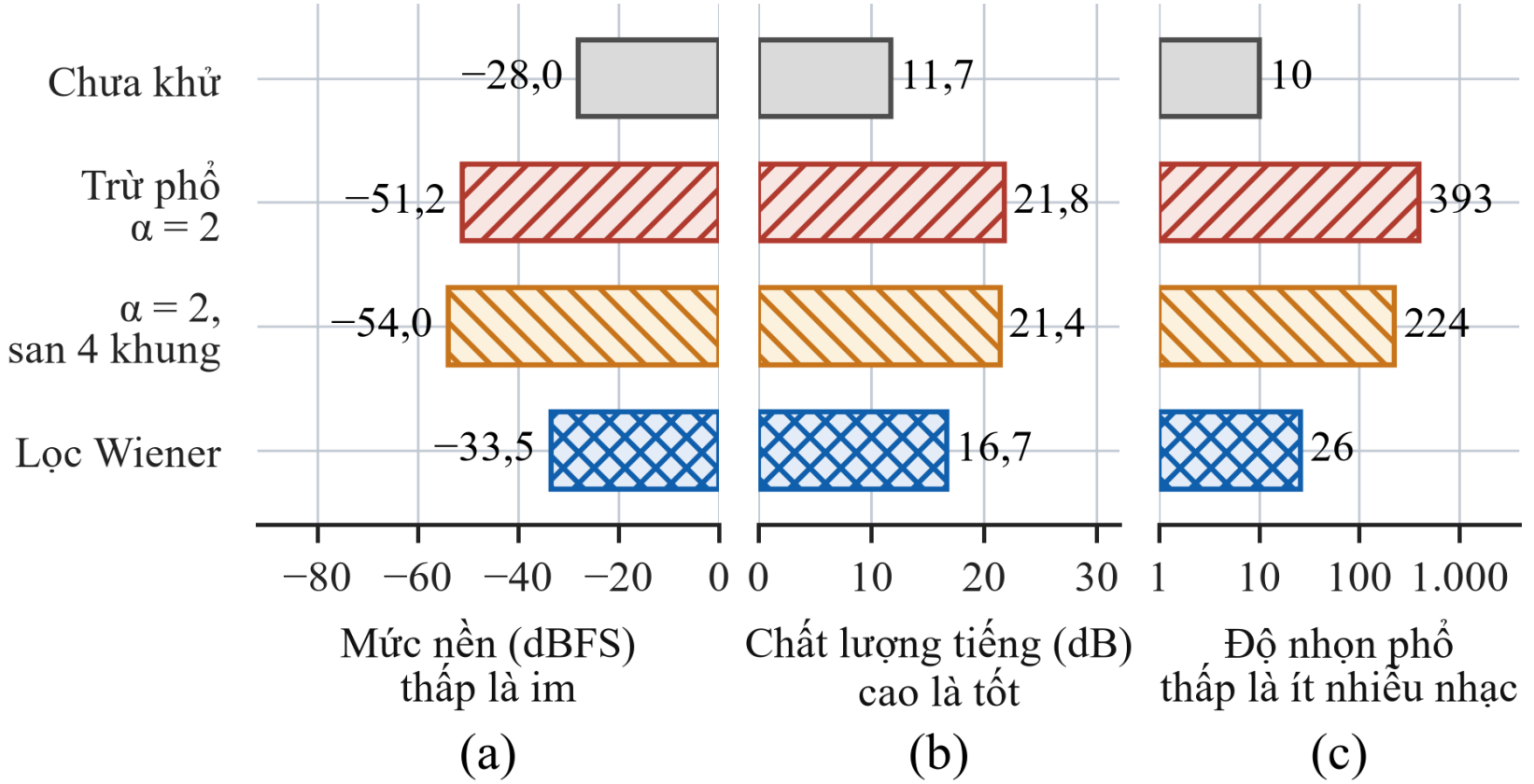
Hệ số nhân đổi liên tục

$$H_w[k] = \frac{\left|\hat{S}[k]\right|^2}{\left|\hat{S}[k]\right|^2 + \left|\hat{N}[k]\right|^2}$$

(7.10)

- Bộ lọc Wiener ở dạng đơn giản nhất.
- Tín hiệu mạnh hơn nhiễu nhiều: hệ số gần 1; nhiễu lấn át: hệ số gần 0.
- Ở giữa, hệ số chuyển dần chứ không nhảy.
- Chính chỗ chuyển dần làm bộ lọc ít sinh nhiễu nhạc.

Ba cách khử nhiễu, đo bằng ba đại lượng



Hình 7.24 — Ba cách khử nhiễu, đo bằng ba đại lượng

7.6.5 · BẢNG 7.8

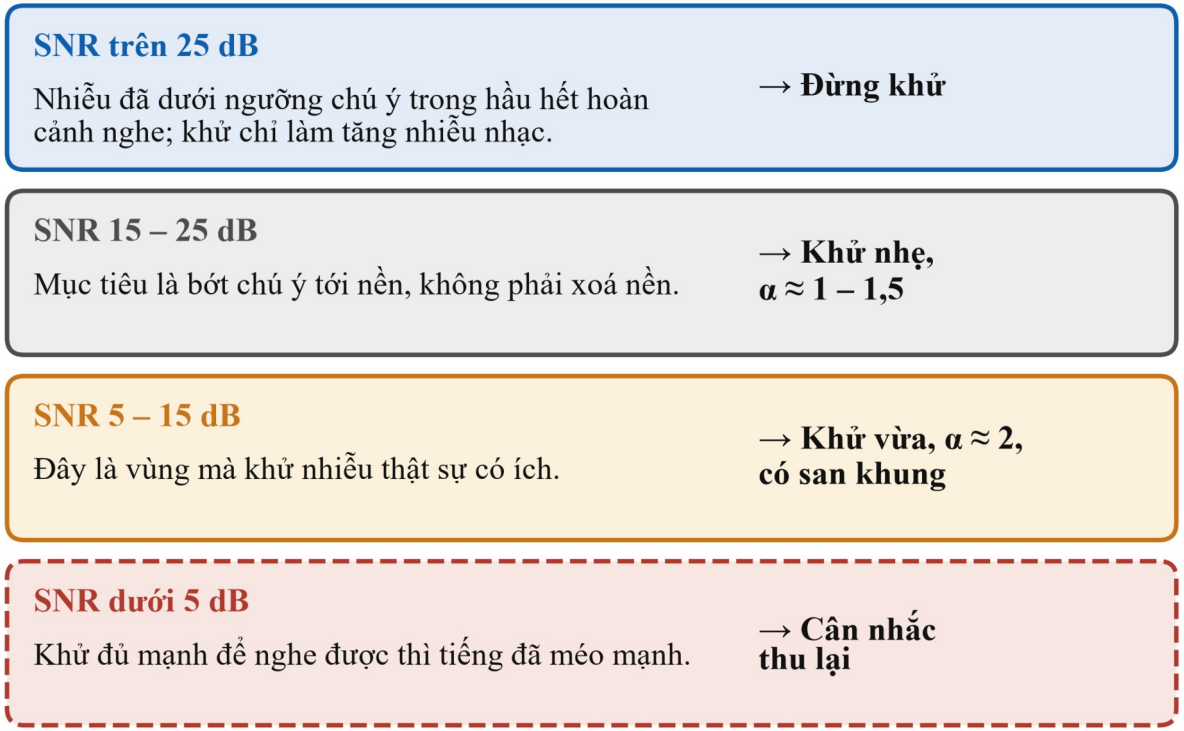
Ba cách khử trên cùng một bản thu SNR 10 dB

Cách	Mức nền (dBFS)	Chất lượng tiếng (dB)	Độ nhọn phổ
Chưa khử	−28,0	11,7	10
Trừ phổ alpha = 2	−51,2	21,8	393
alpha = 2, san 4 khung	−54,0	21,4	224
Lọc Wiener	−33,5	16,7	26

San bốn khung rẻ nhất và là lựa chọn mặc định; Wiener ít nhiễu nhạc nhất nhưng chỉ hạ nền 5,5 decibel.

Bốn vùng SNR và quyết định tương ứng

- Quyết định đầu tiên là có khử hay không, dựa trên một con số đo được.
- Trên 25 decibel, nhiễu đã dưới ngưỡng chú ý trong hầu hết hoàn cảnh nghe.
- Khi ấy khử chỉ đổi thứ không ai để ý lấy thứ ai cũng để ý.



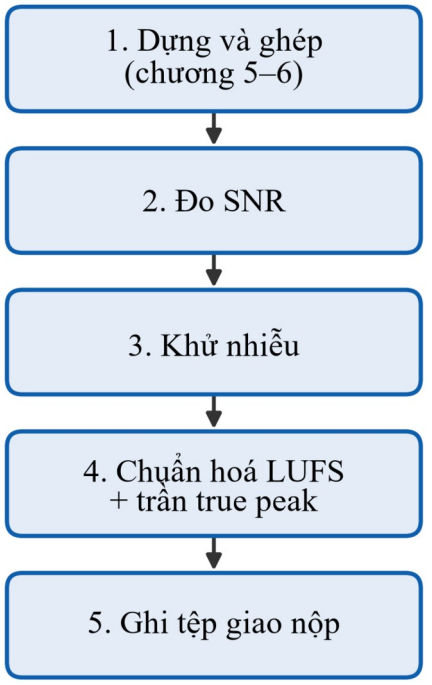
Hình 7.25 — Bốn vùng SNR và quyết định tương ứng

Cách tốt nhất: thu lại cho sạch

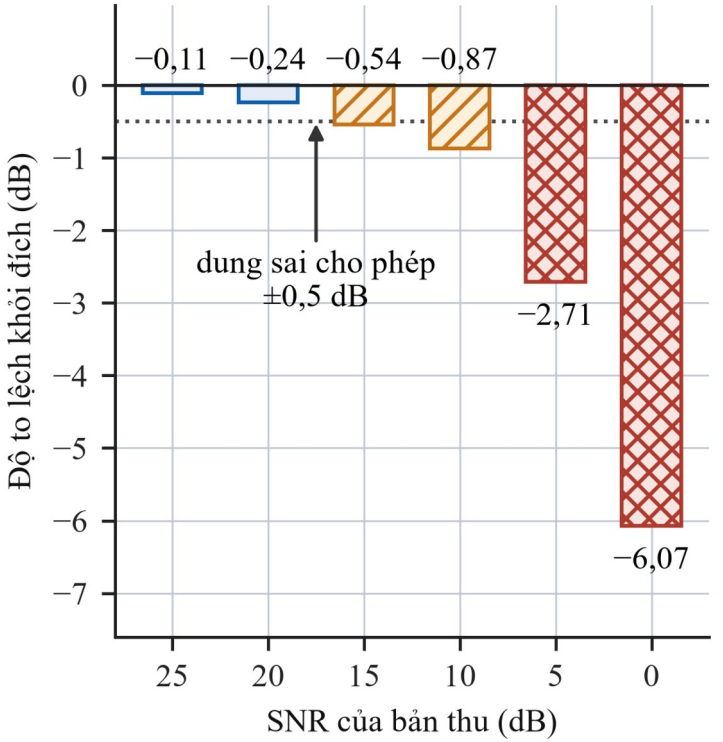
Ba mươi giây tắt điều hoà, dời micro gần miệng thêm hai mươi xăng-ti-mét, hoặc treo một tấm chăn sau lưng người nói đáng giá hơn mọi tham số alpha và beta cộng lại. Khử nhiễu dùng khi không thu lại được.

Thứ tự đúng, và cái giá của làm sai thứ tự

- Chuẩn hoá luôn là bước cuối trước khi ghi tệp.
- Chuẩn hoá trước khi khử nhiễu: phần nhiễu bị lấy đi mang theo một phần độ to.
- Hụt 0,11 decibel ở SNR 25 dB: không đáng kể.
- Hụt 6,07 decibel ở SNR 0 dB: sai hẳn.



(a) Thứ tự đúng



(b) Độ hụt độ to khi chuẩn hoá trước

Hình 7.26 — Thứ tự đúng, và cái giá của làm sai thứ tự

MỤC 7.7

Mốc bài tập lớn — giai đoạn 2

- Nộp trước chương 10, làm tiếp trên bản dựng đã nộp ở giai đoạn 1.
- Nếu bản giai đoạn 1 quá sạch (SNR trên 25 dB), thu thêm ba mươi giây ở nơi có nhiễu nền rõ.
- Tiêu chí có trọng số lớn nhất là giải thích được lựa chọn alpha.

Năm bước của giai đoạn 2

1. Đo trước	Năm con số (đỉnh, true peak, RMS, LUFS tích hợp, LRA) và SNR của bản giai đoạn 1.
2. Chọn nền tảng	Một nền tảng cụ thể trong bảng ở mục 7.3.2; ghi rõ LUFS đích và trần true peak.
3. Khử nhiễu	Nếu SNR dưới 25 dB. Ghi rõ α, β, có san khung hay không và lý do chọn.
4. Chuẩn hoá	Đủ bốn bước ở mục 7.3.3, làm sau khi khử nhiễu.
5. Đo lại	Bảng năm con số trước và sau, cộng mức nền và độ nhọn phổ, kèm nhận xét bằng tai.

Bốn tiêu chí chấm: LUFS đúng đích ±0,5 dB · true peak không vượt trần · có bảng đo trước và sau · giải thích được lựa chọn α.

Trọng số lớn nhất nằm ở tiêu chí cuối: chọn α bao nhiêu không quan trọng bằng việc giải thích được vì sao chọn con số ấy.

Hình 7.27 — Năm bước của giai đoạn 2

Nội dung phải nộp

Hạng mục	Yêu cầu cụ thể
Bảng đo trước	Năm con số của bản giai đoạn 1, cộng SNR đo từ đoạn lặng
Chọn nền tảng	Một nền tảng ở bảng 7.4, ghi rõ LUFS đích và trần
Khử nhiễu	Nếu SNR dưới 25 dB: ghi alpha, beta, san khung và lý do chọn
Chuẩn hoá	Đủ bốn bước, làm sau khử nhiễu; ghi rõ trần có chặn hay không
Bảng đo sau	Cùng năm con số, cộng mức nền và độ nhọn phổ nếu có khử
Nhận xét	Nửa trang: nghe khác gì, đánh đổi nào chưa thoả đáng

Bốn tiêu chí chấm giai đoạn 2

Tiêu chí	Điểm	Cách kiểm
LUFS đúng đích ±0,5 dB	2	Đo bằng bộ đo độc lập; trần chặn lại phải nói rõ, vẫn đủ điểm
True peak không vượt trần	2	Đo bằng lấy mẫu quá 4 lần; vượt trần mất trọn hai điểm
Đủ bảng đo trước và sau	2	Đủ năm con số ở cả hai bảng; khớp khi chạy lại
Giải thích được lựa chọn alpha	4	Nêu được các giá trị đã thử và vì sao dừng ở giá trị chọn

Vì sao tiêu chí cuối có trọng số lớn nhất

Ba tiêu chí đầu máy chấm được và có đáp án duy nhất; chúng kiểm quy trình. Tiêu chí cuối kiểm xem người học hiểu mình đánh đổi cái gì lấy cái gì. Bài chọn alpha bằng 1,5 và giải thích được sẽ điểm cao hơn bài chọn alpha bằng 2 mà không nói được lý do.

Ba lỗi làm mất điểm nhiều nhất ở các khoá trước

- **Chuẩn hoá trước khi khử nhiễu:** bản nộp hụt so với đích vì người nộp không đo lại trên tệp cuối.
- **Khử nhiễu trên tệp đã sạch:** không đo SNR trước; độ nhọn phổ ở bảng đo sau tố cáo điều đó.
- **Báo LUFS mono, nộp stereo (hoặc ngược lại):** chênh đúng 3,01 decibel.

MỤC 7.8

Thực hành, bài tập và tổng kết

- Ba bài thực hành, mỗi bài khoảng mười phút, trên tư liệu của chính người học.
- Bảy bài tập có đáp số bằng số.
- Sáu câu tự kiểm; chưa trả lời được thì quay lại đúng mục.

Làm ngay sau khi đọc chương

- **Bài 1 — ba tệp, ba độ to:** chuẩn hoá về đỉnh −1 dBFS, xếp hạng bằng tai rồi đo LUFS để đối chiếu.
- **Bài 2 — trần chặn lại:** đưa tệp có PLR lớn nhất về −14 LUFS; ghi độ to thật đạt được và nhận xét độ hụt.
- **Bài 3 — nghe nhiễu nhạc:** thử alpha bằng 1, 2, 3, 6; mô tả bằng lời rồi so với độ nhọn phổ.

Bài tập của chương 7

Bài	Nội dung	Đáp số
1	Tệp −31 LUFS: cộng bao nhiêu dB để đạt EBU R 128, Spotify?	+8 dB; +17 dB
2	Tệp mono −18 LUFS chép sang stereo thì đo được bao nhiêu?	−14,99 LUFS
3	LUFS −20, true peak −0,4 dBTP; về −16 LUFS, trần −1 dBTP?	−20,6 LUFS, hụt 4,6 dB
4	Viết hàm bao_cao_muc(x, sr) in năm con số, cảnh báo nếu vượt −1 dBTP	Hàm bắt được tệp vượt trần
5	Trộn nhiễu trắng để đạt SNR 12 dB; kiểm lại bằng đoạn lặng	Lệch dưới 0,2 dB
6	Quét alpha 0,5 tới 6, bước 0,5; vẽ ba đường	Alpha tốt nhất 1,5 – 2,5
7	Chứng minh bằng số: chuẩn hoá trước khử nhiễu cho kết quả sai	Sai lệch lớn hơn ở SNR thấp

Sáu câu tổng kết chương 7

Ba tệp cùng đỉnh −1 dBFS có to bằng nhau không?	Không. Trên ba tư liệu của bài, LUFS chênh nhau 17,0 dB; đỉnh không nói gì về độ to.
Thước K làm gì?	Hạ dải trầm (−3,9 dB ở 50 Hz) và nâng dải cao (+4,1 dB ở 10 kHz) cho khớp với tai người.
Vì sao LUFS cần hai cổng?	Không có cổng, 24 s im lặng kéo con số tụt 6,7 dB; tệp nhiều khoảng lặng bị chuẩn hoá quá mức.
True peak khác đỉnh mẫu thế nào?	Đo sau khi lấy mẫu quá 4 lần. Tư liệu bị cắt đỉnh có thể vượt tới +3,95 dB so với đỉnh mẫu.
Đo SNR bằng cách nào khi chỉ có bản đã trộn?	Lấy công suất nhiễu từ một đoạn lặng, rồi trừ khỏi công suất tổng. Sai lệch dưới 0,1 dB.
Nhiễu nhạc từ đâu ra, và trả giá thế nào?	Từ những vạch nhấp nháy qua ngưỡng. Đỉnh ở $\alpha = 3$ với độ nhọn gấp hơn ba trăm lần nhiễu tự nhiên.

Chương 8 mở mô-đun M6 (hiệu ứng) từ chính vấn đề đã gặp ở mục 7.3.4: bộ nén động.

Hình 7.28 — Sáu câu tổng kết chương 7

Từ chương 7 sang chương 8

- Chương 8 bắt đầu từ chỗ bản nhạc cổ điển bị trần chặn: PLR quá cao thì không chuẩn hoá về mức to được.

- Công cụ chữa là bộ nén động.
- Giữ lại bản nhạc cổ điển; chuẩn bị thêm một đoạn thu có mức lên xuống thất thường.
- Đọc lại mục 4.6 (phương trình sai phân) và mục 6.4 (phép trễ).

Bốn phép kiểm cho các hàm mới

```
# 1. GIA TRI DA BIET — tinh duoc bang tay
#    sin 1 kHz -20 dBFS mono phai doc -23,0 LUFS
x = sine(1000, 4.0, amp=0.1)
assert abs(do_to_tich_hop(x, sr) + 23.0) < 0.05

# 2. TINH CHAT BAT BIEN — dung voi moi dau vao
#    chuyen mono sang stereo thi LUFS tang dung 3,01
st = np.stack([x, x], axis=1)
assert abs(do_to_tich_hop(st, sr)
           - do_to_tich_hop(x, sr) - 3.01) < 0.05

# 3. TRUONG HOP BIEN — dau vao o ria mien xac dinh
#    doan ngan hon mot khoi 400 ms thi khong do duoc
assert not np.isfinite(do_to_tich_hop(sine(1000, 0.1), sr))

# 4. SO VOI CACH CHAM — hai cach, mot ket qua
#    true peak phai >= dinh mau, luon luon
assert true_peak_dbtp(x) >= peak_dbfs(x) - 1e-9
```

Mô-đun am tăng từ 18 lên 34 hàm; mỗi hàm mới kèm ít nhất một phép kiểm trong test_am.py.

Trừ phổ với alpha bằng 0

Hàm khử nhiễu khó kiểm bằng một dấu bằng, nhưng có một tính chất luôn đúng: trừ phổ với alpha bằng 0 phải trả về gần đúng tín hiệu vào. Phép kiểm này bắt mọi lỗi ở khâu cắt khung và dựng lại — chỗ hay sai nhất khi tự cài đặt STFT (Short-Time Fourier Transform – biến đổi Fourier thời gian ngắn).

Những điều cần nhớ

1. Đỉnh, RMS và LUFS đo ba thứ khác nhau: ba tệp cùng đỉnh −1 dBFS có LUFS chênh nhau 17 decibel.
2. LUFS dùng thước K và hai cổng; trần true peak luôn dưới 0 dBFS, thường −1 dBTP.
3. Chuẩn hoá đủ bốn bước: đo, tính gain, kiểm trần, hạ thêm nếu cần; khi trần chặn lại thì giữ trần.
4. SNR đo từ đoạn lặng quyết định có khử nhiễu hay không: trên 25 dB thì đừng khử.
5. Trừ phổ giữ nguyên pha, chọn alpha và beta có căn cứ; cái giá là nhiễu nhạc, đo bằng độ nhọn phổ.
6. Mọi phép xử lý có hai mặt; chuẩn hoá luôn là bước cuối, sau khử nhiễu.