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

LẬP TRÌNH ÂM THANH

CHƯƠNG 14

Tích hợp AI vào sản phẩm

Chương 14 · 5 tiết · Mã học phần 531833 · 75 tiết · Học kỳ 5

Hà Nội – 2026

Sau chương này người học phải làm được

- Đo sáu chỉ số đầu tiên của tệp vừa nhận từ mô hình; xác định các khối xử lý cần dùng.
- Phân biệt hai cách bỏ lệch một chiều; dò vách phổ để suy ra tần số sinh thật.
- Rút gọn tỉ số L/M; cài và đo phép đổi tần số bằng bộ lọc đa pha; chọn số cạnh có căn cứ.
- Giải thích vì sao bỏ mẫu không lọc là lỗi không sửa được; phân biệt đỉnh mẫu với đỉnh thật.
- Chọn mức giao hàng và trần bộ giới hạn từ số đo; chọn luật mờ chéo theo hệ số tương quan.
- Dựng phép đo có kiểm soát cho mối nối; xếp năm khối của đường ống đúng thứ tự.
- Lập ngân sách độ trễ và thời gian máy; viết bộ kiểm bàn giao và đặt dung sai hồi quy.

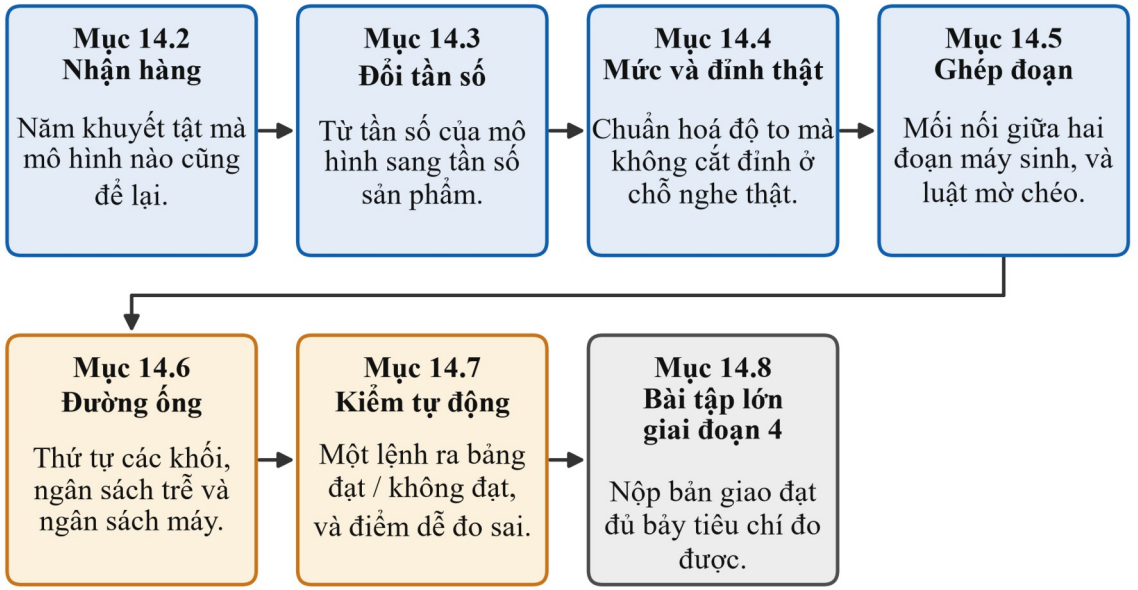
Một chương không có khái niệm mới, và đó là chủ ý

Mọi thứ đã học: lấy mẫu, FIR (Finite Impulse Response – đáp ứng xung hữu hạn), LUFS (Loudness Units Full Scale – đơn vị độ to so với toàn thang), đỉnh thật, bộ giới hạn, mờ chéo. Cái mới là thứ tự các khối: ba trong bốn thứ tự hợp lý vẫn cho ra tệp không đạt. Mục 14.6 đo cả bốn.

MỤC 14.1

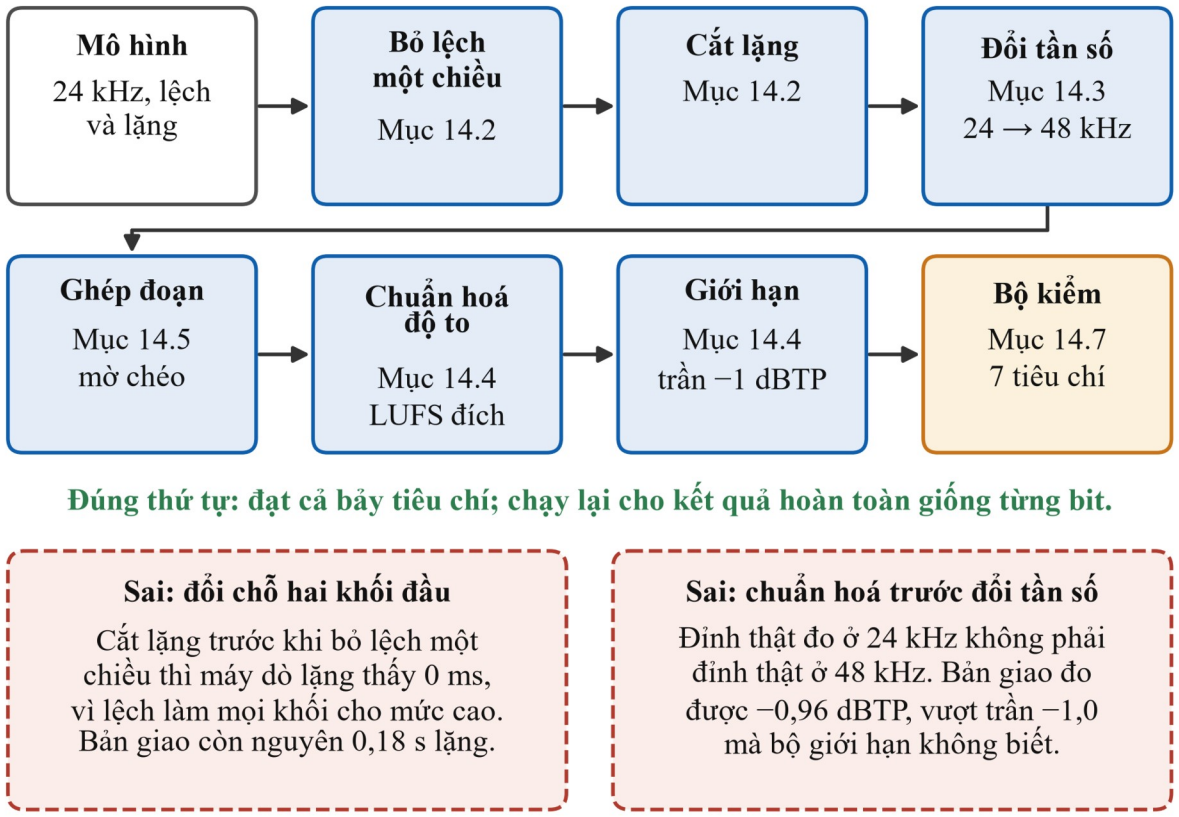
Vị trí chương học trong học phần

- Chương 13 dừng ở mảng số do mô hình trả về; chương này biến nó thành tệp giao được.
- "Có thể giao" là bảy con số đo được, mỗi số có một ngưỡng.
- Gần như không có khái niệm mới: chương đặt các kỹ thuật đã học vào đúng thứ tự.



Từ nhận hàng đến nộp bản giao: bảy mục kỹ thuật, mỗi mục kết thúc bằng một con số đo được.

Bảy khối của đường ống hậu kỳ và hai cách xếp sai



Hình 14.2 — Bảy khối của đường ống hậu kỳ, và hai cách xếp sai hay gặp

Bảy khối và câu hỏi riêng của mỗi khối

Khối	Việc của nó	Câu hỏi quyết định
Bỏ lệch một chiều	Đưa trung bình về 0	Trừ trung bình hay thông cao
Cắt lặng	Bỏ phần im hai đầu	Ngưỡng và phần chừa lại
Đổi tần số	Sang tần số sản phẩm	Bộ lọc dài bao nhiêu
Ghép đoạn	Nối nhiều đoạn	Luật mờ chéo, độ dài
Chuẩn hoá độ to	Về mức giao hàng	LUFS đích và trần
Giới hạn	Không vượt trần	Trần đặt ở đâu
Bộ kiểm	Nói đạt hay không	Số tiêu chí, nguồn ngưỡng

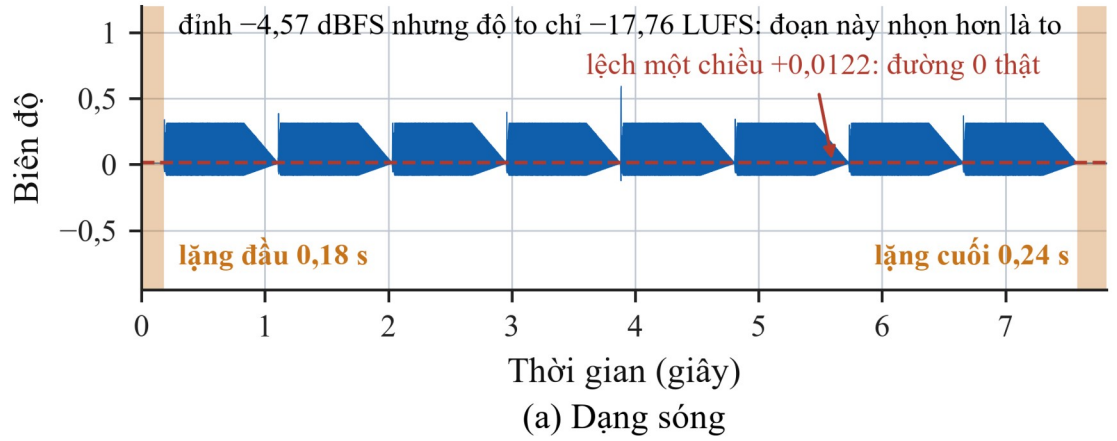
Mỗi khối thay đổi thứ mà khối sau đo được

Lệch một chiều làm máy dò lặng mù; chưa đổi tần số thì đỉnh thật đo ra chưa phải của bản giao. Thứ tự không phải chuyện thẩm mỹ: hai trong bốn cách xếp cho tệp không đạt, dù nghe không ai phân biệt.

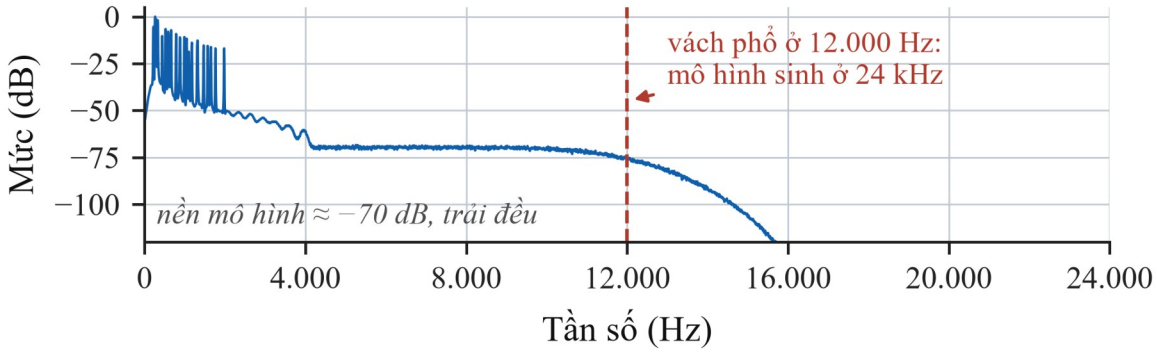
MỤC 14.2

Kết quả AI đến tay người dùng ở trạng thái nào

- Vật liệu giả lập mang năm khuyết tật cố ý; mọi con số chạy lại được.
- Đo sáu chỉ số trước khi làm bất cứ việc gì khác.
- Lệch một chiều, lặng hai đầu và vách phổ: ba khuyết tật đo được ngay.



(a) Dạng sóng



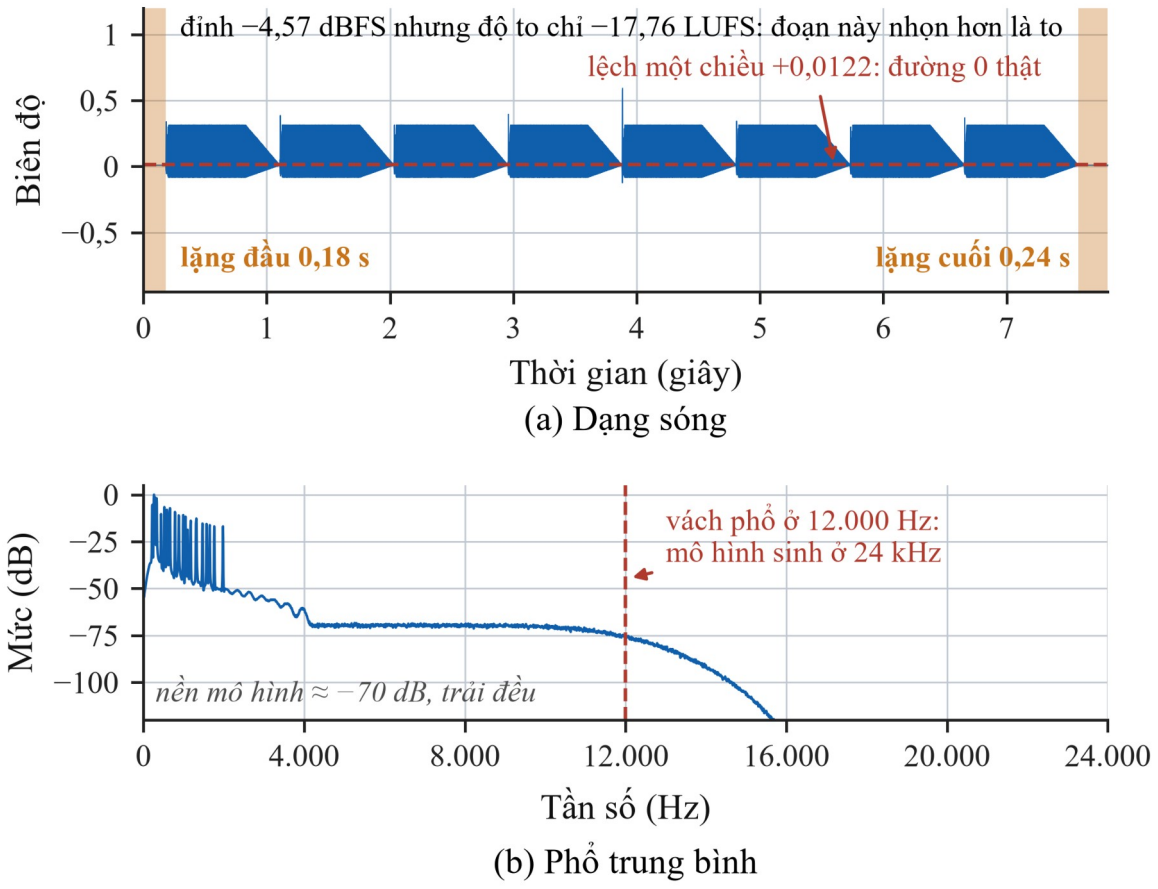
(b) Phổ trung bình

Năm khuyết tật cố ý đặt vào vật liệu

Khuyết tật	Vì sao mô hình thật cũng có	Mục xử lý
Lấy mẫu 24 kHz	Huấn luyện ở tần số thấp cho rẻ	14.3
Lệch một chiều 0,004–0,012	Không có ràng buộc trung bình bằng 0	14.2.3
Mức trôi giữa các đoạn	Mỗi lần gọi mô hình độc lập	14.4
Lặng 0,18–0,32 giây	Phần thừa được lấp bằng im lặng	14.2.4
Nền nhiễu khoảng −70 dB	Sai số lượng tử của tầng dựng sóng	14.7

Đo trước, nghe sau

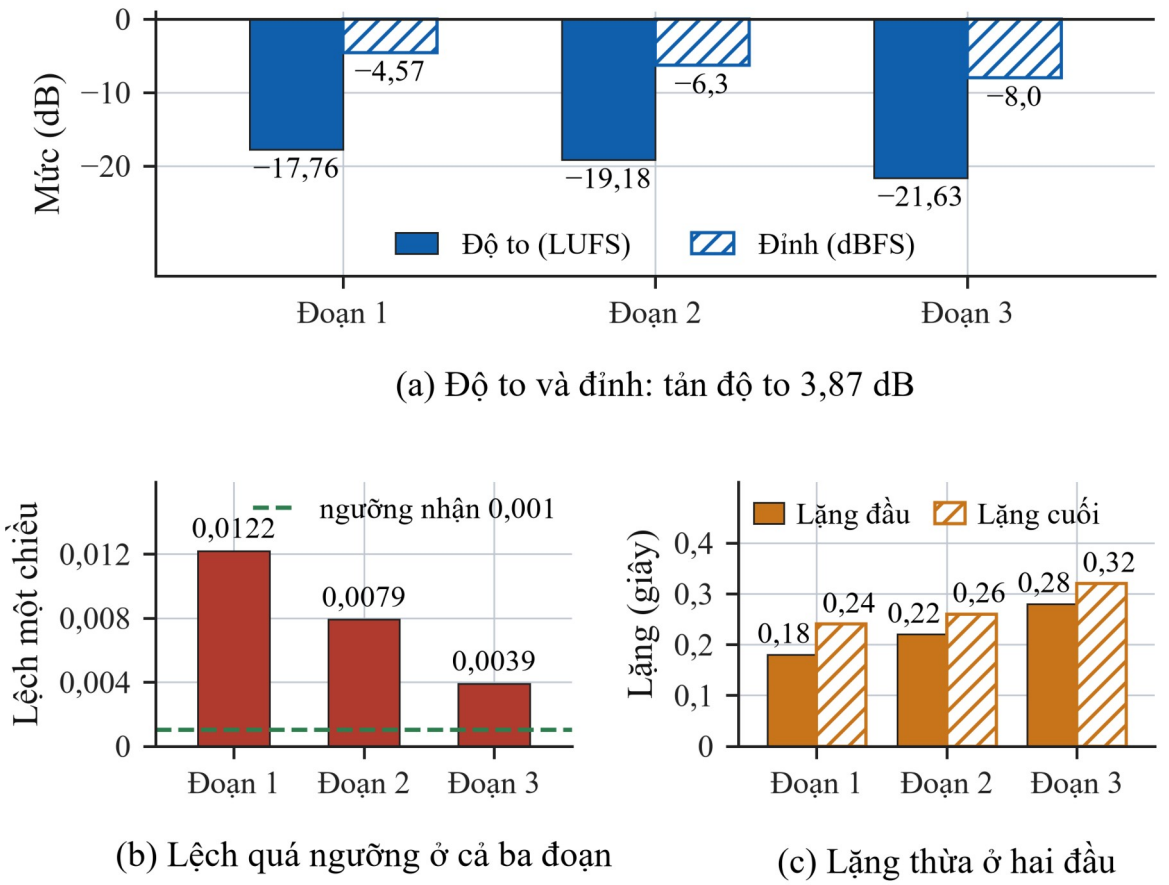
- Việc đầu tiên khi nhận tệp là đo sáu chỉ số, không phải nghe.
- Sáu con số cho biết ngay tệp cần những khối xử lý nào.
- Hình: dạng sóng ở trên, phổ trung bình ở dưới.



Hình 14.3 — Một đoạn vừa nhận: dạng sóng ở trên, phổ trung bình ở dưới

Ba đoạn chênh nhau 3,87 dB độ to

- Độ to lần lượt −17,76 · −19,18 · −21,63 LUFS.
- Ai cũng nghe ra: như ba lần thu khác nhau, không phải một bản nhạc.
- Đoạn 1: đỉnh −4,57 dBFS (decibel so với mức toàn thang, Full Scale), độ to −17,76 LUFS, chênh hơn mười ba decibel.
- Vật liệu do máy sinh thường nhọn hơn bản thu thật.



Hình 14.4 — Ba đoạn vừa nhận, đo trên ba nhóm chỉ số

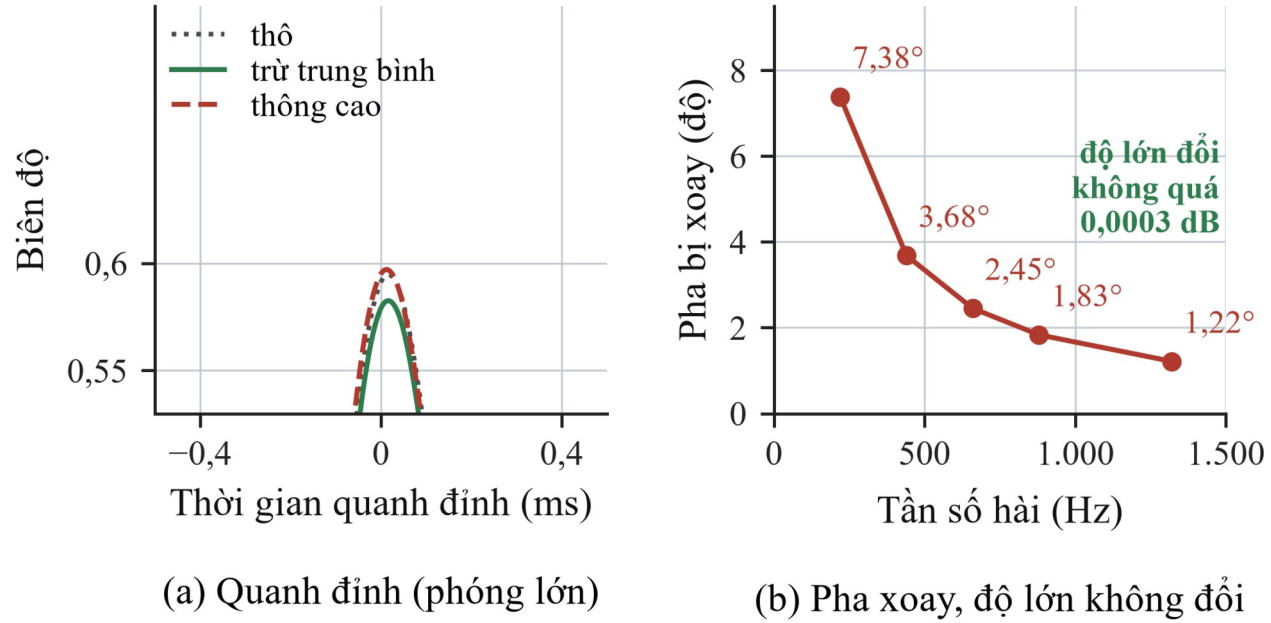
Lệch một chiều và cách trừ

$$d = \frac{1}{N}\sum_{n=0}^{N-1} x[n], \qquad x_{\text{sạch}}[n] = x[n] - d$$

(14.1)

- d là giá trị trung bình, N là số mẫu; x sạch là tín hiệu sau khi trừ phần lệch.
- Tai không nghe thấy 0 Hz, nhưng mọi phép đo và khối sau đều bị sai lệch.
- Đoạn 1: d = 0,0122, tức dạng sóng bị đẩy lên 1,2 phần trăm toàn thang.

Hai cách bỏ lệch một chiều và vì sao chúng khác nhau



Trừ trung bình

đỉnh −4,75 dBFS: lấy lại 0,181 dB
độ to giữ nguyên −17,76 LUFS

Thông cao 20 Hz

đỉnh −4,51 dBFS: mất 0,057 dB
độ to vẫn −17,76 LUFS

Hình 14.5 — Hai cách bỏ lệch một chiều, và vì sao chúng khác nhau

Hai cách bỏ lệch một chiều, đo trên đoạn 1

	Trừ trung bình	Thông cao 20 Hz
Lệch còn lại	0,0000	0,0000
Đỉnh sau khi dọn	−4,75 dBFS	−4,51 dBFS
So với trước (−4,57)	Lấy lại 0,181 dB	Mất 0,057 dB
Độ to sau khi dọn	−17,76 LUFS	−17,76 LUFS
Xử lý lệch trôi theo thời gian	Không	Có
Có xoay pha không	Không	Có

Đỉnh và độ to khác nhau về bản chất

Bộ lọc xoay pha các hài nên chúng cộng lại khác đi: đỉnh cao hơn 0,057 dB. Độ to đo năng lượng, không phụ thuộc pha, nên không đổi. Vì vậy phải đo đỉnh ở cuối đường ống, không đo ở giữa.

Đoạn thô cho 0,00 giây lặng; sau khi bỏ lệch cho 0,18 và 0,24 giây

Khối im của đoạn thô có mức −38,3 dB, cao hơn ngưỡng −64,6 dB tới 26 dB. Bỏ lệch một chiều phải là khối đầu tiên. Khi cắt, chừa lại 20 ms để không chặt phần lên của nốt đầu.

Vách phổ: tần số sinh thật của mô hình

$$f_{\text{vách}} = \min f \ \text{saocho} \ \bar{S}(g) \le \bar{S}_{\max} - \Delta \ \text{vớimọi} \ g \ge f, \qquad \Delta = 25 \ \text{dB}$$

(14.2)

- Tệp ghi 48.000 Hz chưa chắc có thông tin tới 24.000 Hz.
- Vách là tần số nhỏ nhất mà từ đó trở lên mọi tần số thấp hơn đỉnh phổ ít nhất 25 dB.
- Dò theo vách chứ không theo ngưỡng tuyệt đối, vì mức nền mỗi tệp một khác.

MỤC 14.3

Đổi tần số lấy mẫu

- Tỉ số hữu tỉ L/M quyết định chi phí và độ khó của bộ lọc.
- Bộ lọc đa pha cho cùng kết quả với ít phép nhân hơn L lần.
- Ba cách đổi cách nhau 67 dB; bỏ mẫu không lọc là lỗi không sửa được.

Rút gọn tỉ số thành L và M

$$g = \gcd(f_{\text{vào}}, f_{\text{ra}}), \qquad L = \frac{f_{\text{ra}}}{g}, \qquad M = \frac{f_{\text{vào}}}{g} \tag{14.3}$$

- Rút gọn tỉ số tần số ra trên tần số vào thành hai số nguyên L và M.
- Chèn L−1 số 0 giữa mỗi cặp mẫu, lọc thông thấp, giữ lại mỗi M mẫu một mẫu.
- Hai số này quyết định chi phí và độ khó của bộ lọc.

Tỉ số tối giản của sáu cặp tần số hay gặp

Từ	Sang	L	M	Nhận xét
44.100	48.000	160	147	Đắt, hai số đều lớn
48.000	44.100	147	160	Đắt, ngược chiều
24.000	48.000	2	1	Rẻ nhất, chỉ gấp đôi
22.050	48.000	320	147	Đắt nhất trong bảng
16.000	44.100	441	160	Rất đắt
32.000	48.000	3	2	Rẻ

Chọn tần số mô hình có tỉ số đơn giản với tần số sản phẩm

Mô hình 24 kHz lên 48 kHz tốn L = 2; mô hình 22,05 kHz lên cùng chỗ tốn L = 320. Chất lượng như nhau, nhưng bộ nhớ và độ dài bộ lọc khác hẳn.

Khối thứ nhất: chèn số 0

$$x_L[m] = x[m/L] \text{ khi } L \mid m ; \qquad x_L[m] = 0 \text{ cácchỉsốcònlại}$$

(14.4)

- Tín hiệu trung gian có L lần số mẫu; cứ L−1 mẫu lại là một số 0.
- Phổ của nó chứa L bản sao của phổ gốc, gọi là các ảnh.

Khối thứ hai: lọc thông thấp cắt ảnh

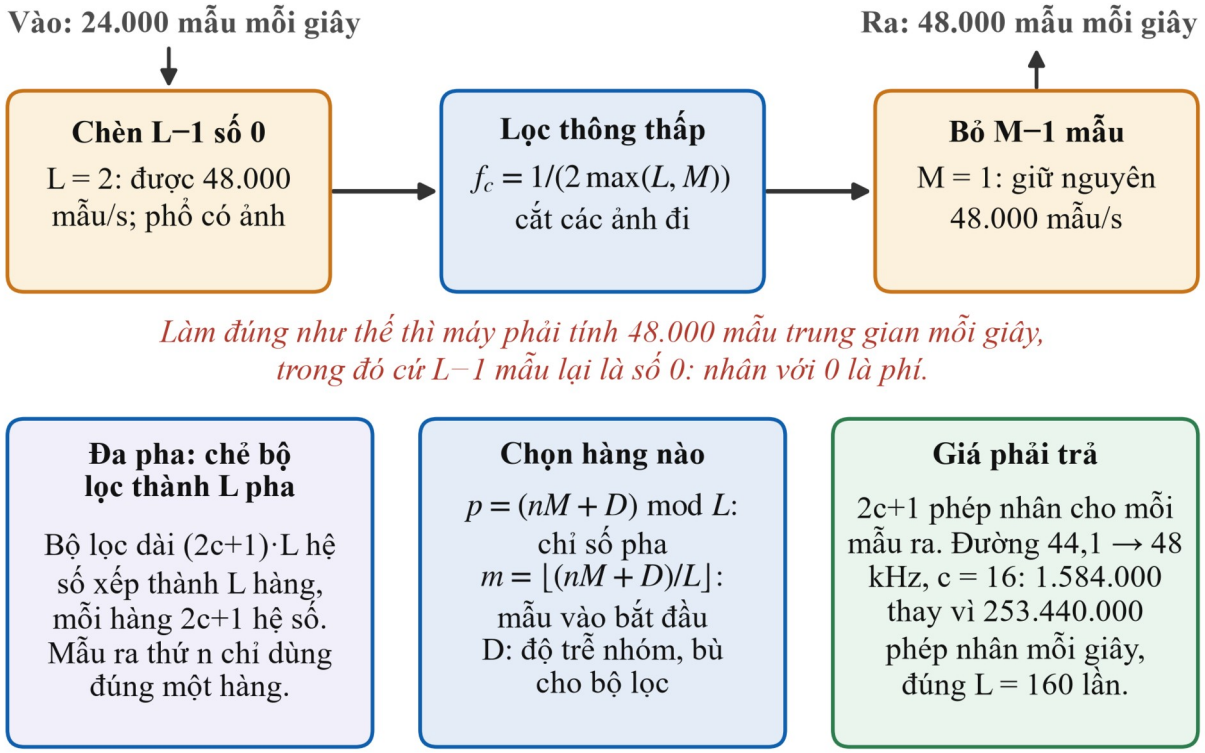
$$f_c = \frac{1}{2\max(L, M)} \quad (\text{theo tần số trung gian } Lf_{\text{vào}})$$

(14.5)

- Tần số cắt phải nhỏ hơn nửa Nyquist của cả đầu vào lẫn đầu ra.
- Vì vậy tần số cắt lấy theo số lớn hơn trong L và M.
- Khối thứ ba giữ lại mỗi M mẫu một mẫu.

Ba khối của cách hiểu, một vòng lặp của cách làm

- Làm đúng ba khối thì phải tính 48.000 mẫu trung gian mỗi giây, một nửa là số 0.
- Nhân hệ số với số 0 là phí.
- Khi L = 160, có 159 trên 160 phép nhân là phí.



Hình 14.6 — Ba khối của cách hiểu, và một vòng lặp của cách làm

Chẻ bộ lọc thành L hàng

$$y[n] = \sum_{k=0}^{K-1} g_p[k]\, x[m-k], \qquad p = (nM + D) \bmod L, \qquad m = \left\lfloor \frac{nM + D}{L} \right\rfloor$$

(14.6)

- Bộ lọc dài (2c+1)·L hệ số chẻ thành L hàng, mỗi hàng 2c+1 hệ số.
- p là chỉ số hàng (pha); m là mẫu vào nơi tích chập bắt đầu.
- D là độ trễ nhóm, cộng vào để tín hiệu ra căn đúng với tín hiệu vào.

Toàn bộ phép đổi tần số bằng đa pha

```
def doi_sr_da_pha(x, L, M, so_canh=16, beta=8.6):
    h   = bo_loc_doi_sr(L, M, so_canh, beta)
    K   = h.size // L
    G   = h.reshape(K, L).T              # (L, K) — L pha
    tre = (h.size - 1) // 2
    n   = np.arange(int(np.floor(x.size * L / M)))
    t   = n * M + tre
    p, m = t % L, t // L                 # pha nao, bat dau o dau
    idx = (m[:, None] - np.arange(K)[None, :]) + dem
    return (xp[idx] * G[p]).sum(axis=1)
```

Mỗi mẫu ra chỉ dùng đúng một hàng của bộ lọc.

Chi phí không phụ thuộc L

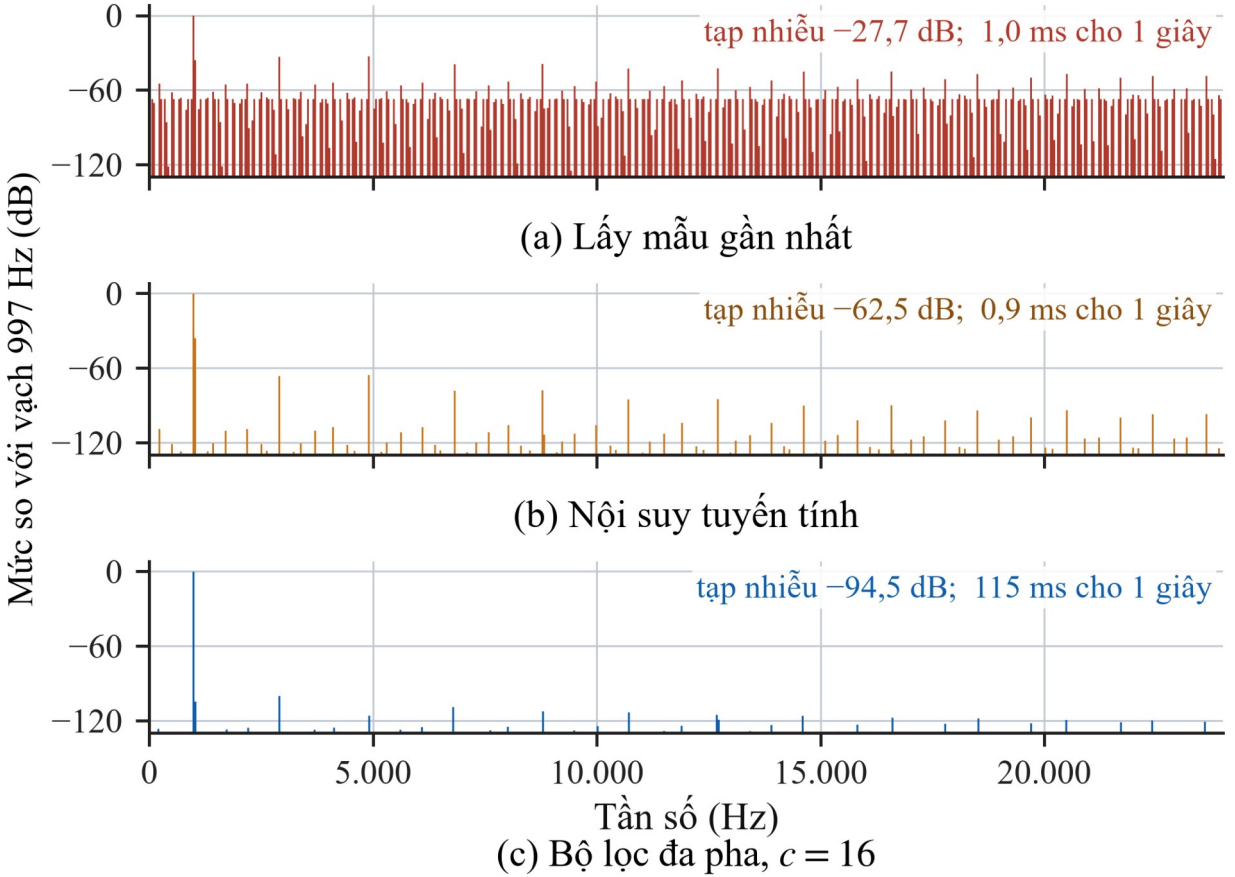
$$C_{\text{nhân}} = (2c+1)f_{\text{ra}}, \qquad \tau_{\text{lọc}} = \frac{(2c+1)L-1}{2Lf_{\text{vào}}}$$

(14.7)

- Mỗi mẫu ra tốn 2c+1 phép nhân.
- Đường 44,1 lên 48 kHz, c = 16: 1.584.000 phép nhân mỗi giây.
- Làm đúng ba khối: 253.440.000 phép nhân mỗi giây, nhiều hơn đúng 160 lần, bằng L.

Phổ hình sin 997 Hz sau ba cách đổi tần số

- Cách 1: lấy mẫu vào gần nhất. Cách 2: nội suy tuyến tính.
- Cách 3: bộ lọc đa pha.
- Vạch phổ nào khác 997 Hz đều do phép đổi sinh ra; tổng của chúng đo mức tạp nhiễu.



Hình 14.7 — Phổ của một hình sin 997 Hz sau ba cách đổi tần số

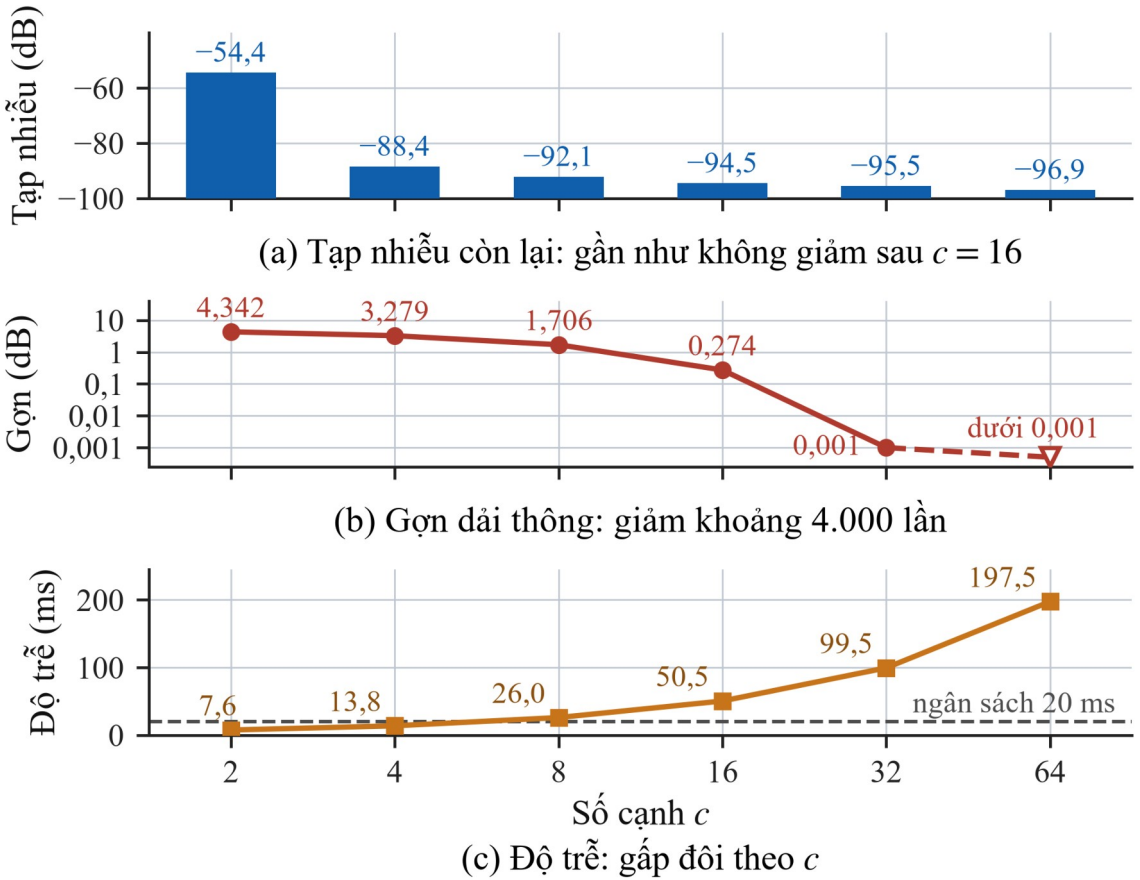
Ba cách đổi tần số, đo trên cùng một hình sin

Cách	Rác còn lại	Thời gian / 1 giây	Dùng khi nào
Lấy mẫu gần nhất	−27,7 dB	1,0 ms	Không bao giờ
Nội suy tuyến tính	−62,5 dB	0,9 ms	Đổi tỉ số liên tục
Đa pha c = 16	−94,5 dB	115 ms	Mặc định cho hậu kỳ

Chênh 67 dB; dựng bản ba phút mất khoảng hai mươi giây, cái giá không đáng bàn cãi.

Ba hệ quả của số cạnh c

- Đo trên đường 44,1 lên 48 kHz.
- Ba hệ quả: độ triệt ảnh, gợn dải thông, độ trễ.
- Ba hệ quả này không cùng chiều.



Hình 14.8 — Ba hệ quả của số cạnh, đo trên đường 44,1 lên 48 kHz

Số cạnh và ba hệ quả của nó

c	Tap / mẫu ra	Rác còn lại	Gợn dải thông	Độ trễ
2	5	−54,4 dB	4,342 dB	7,6 ms
4	9	−88,4 dB	3,279 dB	13,8 ms
8	17	−92,1 dB	1,706 dB	26,0 ms
16	33	−94,5 dB	0,274 dB	50,5 ms
32	65	−95,5 dB	0,001 dB	99,5 ms
64	129	−96,9 dB	< 0,001 dB	197,5 ms

Độ triệt gần đứng yên sau c = 16; gợn giảm bốn nghìn lần; độ trễ gấp đôi theo c.

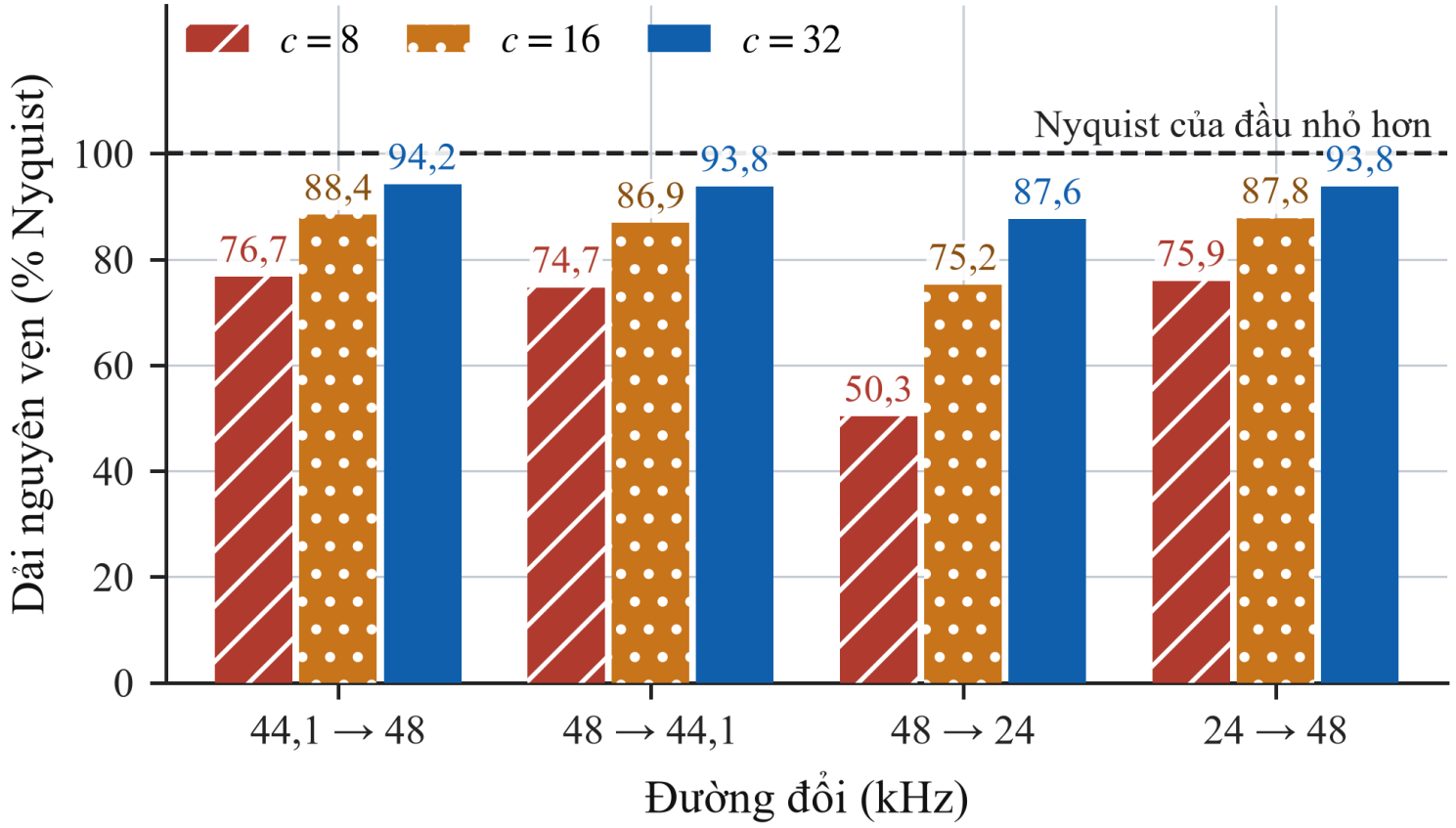
Dải giữ được không bằng Nyquist

$$B_{\text{giữ}} = \frac{\min(f_{\text{vào}}, f_{\text{ra}})}{2}(1-\rho)$$

(14.8)

- Hiểu lầm: đi xuống 24 kHz thì giữ mọi thứ dưới 12.000 Hz.
- Điều đó chỉ đúng nếu bộ lọc là bức tường thẳng đứng.
- ρ là phần dải bị dải chuyển tiếp ăn mất; phụ thuộc c và L/M; chỉ đo mới biết.

Phần dải giữ nguyên vẹn trong phạm vi 0,1 dB



Hình 14.9 — Phần dải giữ nguyên vẹn trong phạm vi 0,1 decibel

Dải giữ nguyên vẹn trong 0,1 dB, theo số cạnh

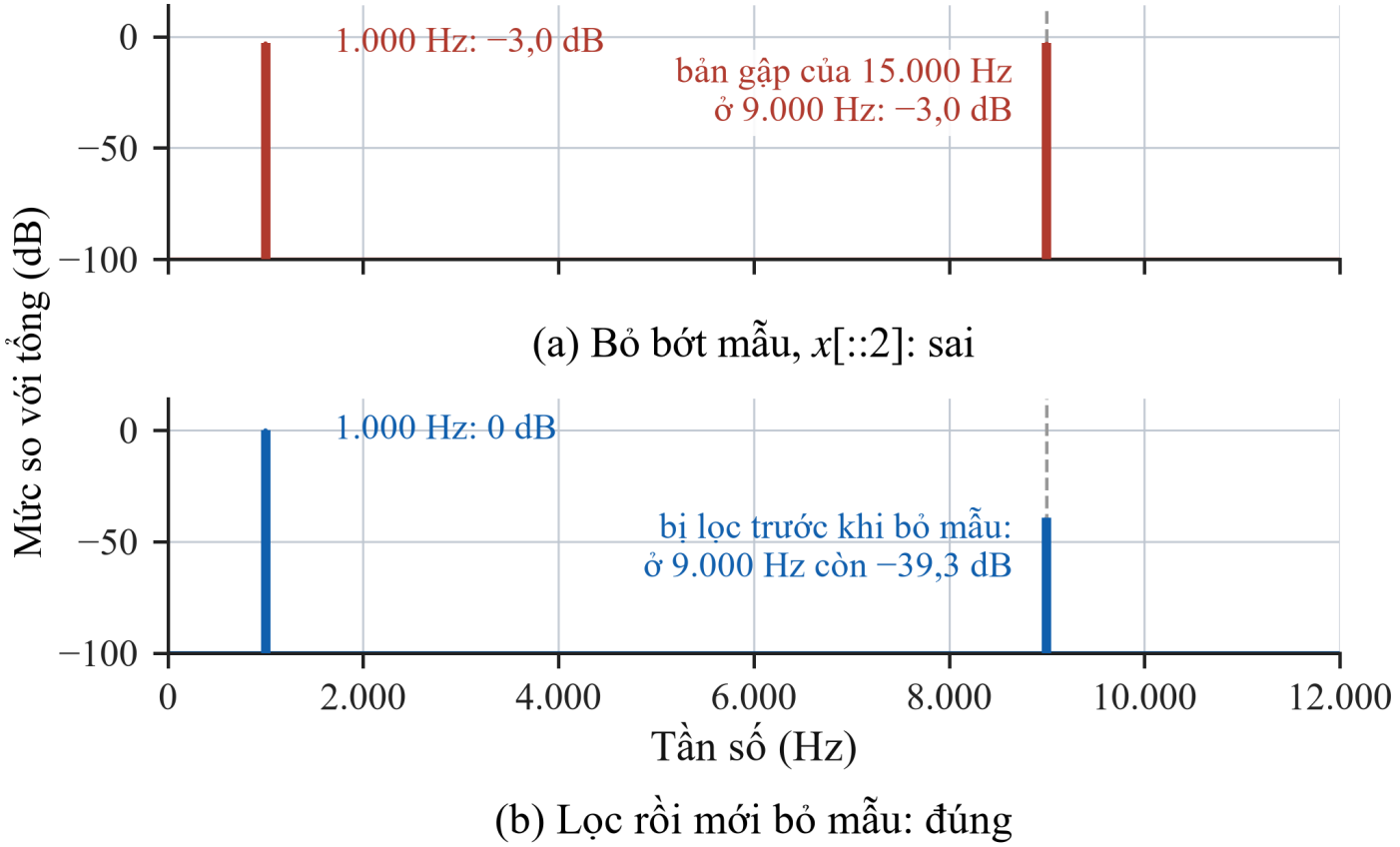
Đường đổi	Nyquist	c = 8	c = 16	c = 32
44,1 → 48 kHz	22.050 Hz	76,7 %	88,4 %	94,2 %
48 → 44,1 kHz	22.050 Hz	74,7 %	86,9 %	93,8 %
48 → 24 kHz	12.000 Hz	50,3 %	75,2 %	87,6 %
24 → 48 kHz	12.000 Hz	75,9 %	87,8 %	93,8 %

Đi xuống 24 kHz với c = 8 chỉ giữ một nửa dải: trên 6 kHz đã suy giảm.

Khi đi xuống tần số thấp hơn, dùng c lớn hơn: c = 32 thay vì 16

Cái giá là gấp đôi số phép nhân; phần thưởng là mười hai phần trăm dải. Phép đi xuống không đảo ngược được: phần dải mất không lấy lại bằng cách nâng lên trở lại.

Cùng một tín hiệu đưa xuống 24 kHz theo hai cách



Hình 14.10 — Cùng một tín hiệu đưa xuống 24 kHz theo hai cách

Bỏ bớt mẫu so với lọc rồi mới bỏ

	Bỏ bớt mẫu (x[::2])	Lọc rồi mới bỏ
Thành phần ngoài 1.000 Hz	−3,0 dB	−39,3 dB
Hình sin 15.000 Hz đi đâu	Gập xuống 9.000 Hz, gần nguyên mức	Bị cắt trước khi bỏ mẫu
Có nghe ra không	Có, một nốt lạ	Không
Sửa được sau khi lưu không	Không	—

Phần gập xuống lẫn trong dải nghe, không phân biệt được với nội dung thật

Mọi lỗi khác của chương sửa được bằng cách chạy lại đường ống từ bản gốc; lỗi này thì không, trừ khi còn giữ bản 48 kHz. Quy tắc: giữ bản gốc ở tần số cao nhất, chỉ đổi xuống ở bước cuối cùng.

MỤC 14.4

Chuẩn hoá mức và đỉnh thật

- Đưa ba đoạn về cùng độ to nhưng đỉnh thật có thể vượt trần.
- Đỉnh thật nằm giữa hai mẫu và hiện ra sau khi đổi tần số.
- Mức đích và trần là hai tham số; trần có quyền phủ quyết.

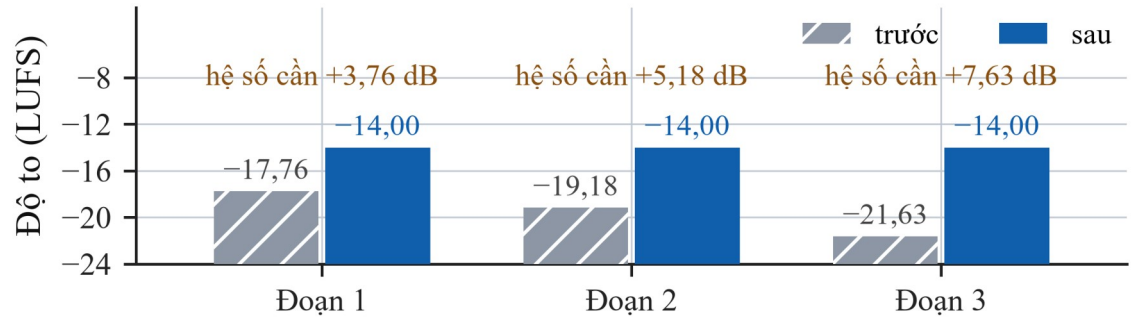
Hệ số chuẩn hoá có kèm trần

$$G = L_{\text{đích}} - L_{\text{đo}}, \qquad G_{\text{thật}} = \min\left(G,\ T_{\text{trần}} - P_{\text{tp}}\right)$$

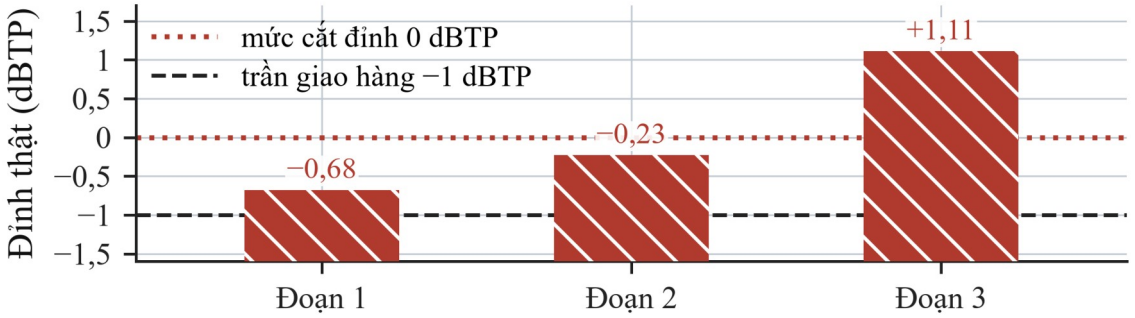
(14.9)

- Hệ số G tính từ hiệu giữa mức đích và mức đo được.
- Không nhân đúng G nếu đỉnh thật vượt trần: phải hạ thêm.
- Khi ấy độ to không đạt đích nữa; vì vậy công thức có phần lấy giá trị nhỏ hơn.

Chuẩn hoá ba đoạn về −14 LUFS và hậu quả với đỉnh thật



(a) Độ to trước và sau khi chuẩn hoá về −14 LUFS



(b) Đỉnh thật sau chuẩn hoá: cả ba đoạn vượt trần

Hình 14.11 — Chuẩn hoá ba đoạn về −14 LUFS, và hậu quả với đỉnh thật

Chuẩn hoá ba đoạn về −14 LUFS

Đoạn	Trước	Hệ số cần	Sau	Đỉnh thật sau
1	−17,76 LUFS	+3,76 dB	−14,00 LUFS	−0,68 dBTP
2	−19,18 LUFS	+5,18 dB	−14,00 LUFS	−0,23 dBTP
3	−21,63 LUFS	+7,63 dB	−14,00 LUFS	+1,11 dBTP

Tản độ to về 0, nhưng cả ba vượt trần −1 dBTP (decibel đỉnh thật, True Peak); đoạn 3 vượt cả 0 dBTP.

Đỉnh thật theo BS.1770

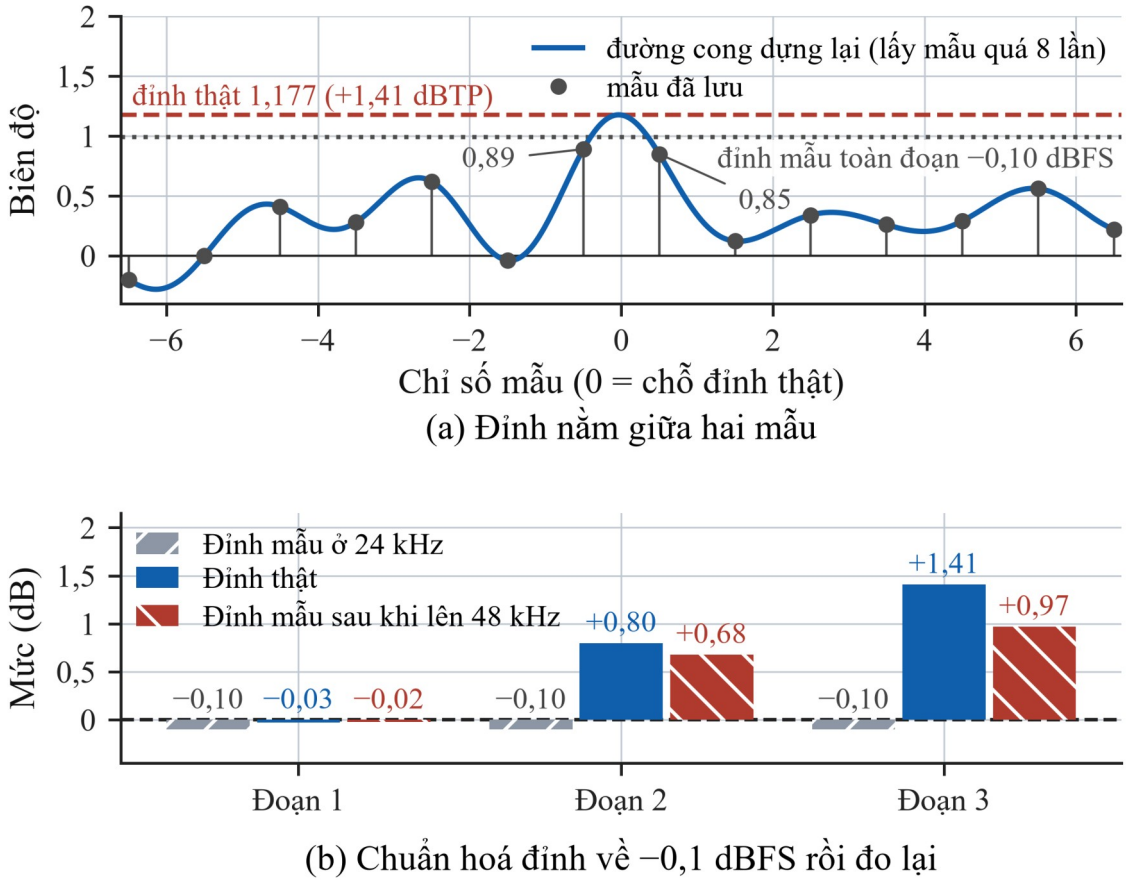
$$P_{\text{tp}} = 20 \log_{10} \max_{n} \left| x_{\uparrow 4}[n] \right|, \qquad \text{PLR} = P_{\text{tp}} - L_{\text{đo}}$$

(14.10)

- Đỉnh mẫu chỉ nhìn các mẫu đã có.
- Bộ chuyển số sang tương tự dựng đường cong qua các mẫu; đường cong có thể vọt cao hơn giữa hai mẫu.
- BS.1770: đỉnh thật là giá trị lớn nhất của đường cong, đo bằng lấy mẫu quá ít nhất bốn lần.

Đỉnh thật nằm giữa hai mẫu

- Hình (a): hai mẫu liền nhau ở 0,89 và 0,85.
- Sóng thật giữa chúng vọt tới 1,177, tức +1,41 dBTP.
- Đỉnh mẫu chỉ là −0,10 dBFS.



Hình 14.12 — Đỉnh thật nằm giữa hai mẫu, và hậu quả sau khi đổi tần số

Ba đoạn chuẩn hoá ĐỈNH về −0,1 dBFS rồi đo lại

Đoạn	Đỉnh mẫu 24 kHz	Đỉnh thật	Chênh	Đỉnh mẫu sau khi lên 48 kHz
1	−0,10 dBFS	−0,03 dBTP	+0,07 dB	−0,02 dBFS
2	−0,10 dBFS	+0,80 dBTP	+0,90 dB	+0,68 dBFS
3	−0,10 dBFS	+1,41 dBTP	+1,51 dB	+0,97 dBFS

Tệp "an toàn" ở 24 kHz bị cắt đỉnh ở 48 kHz

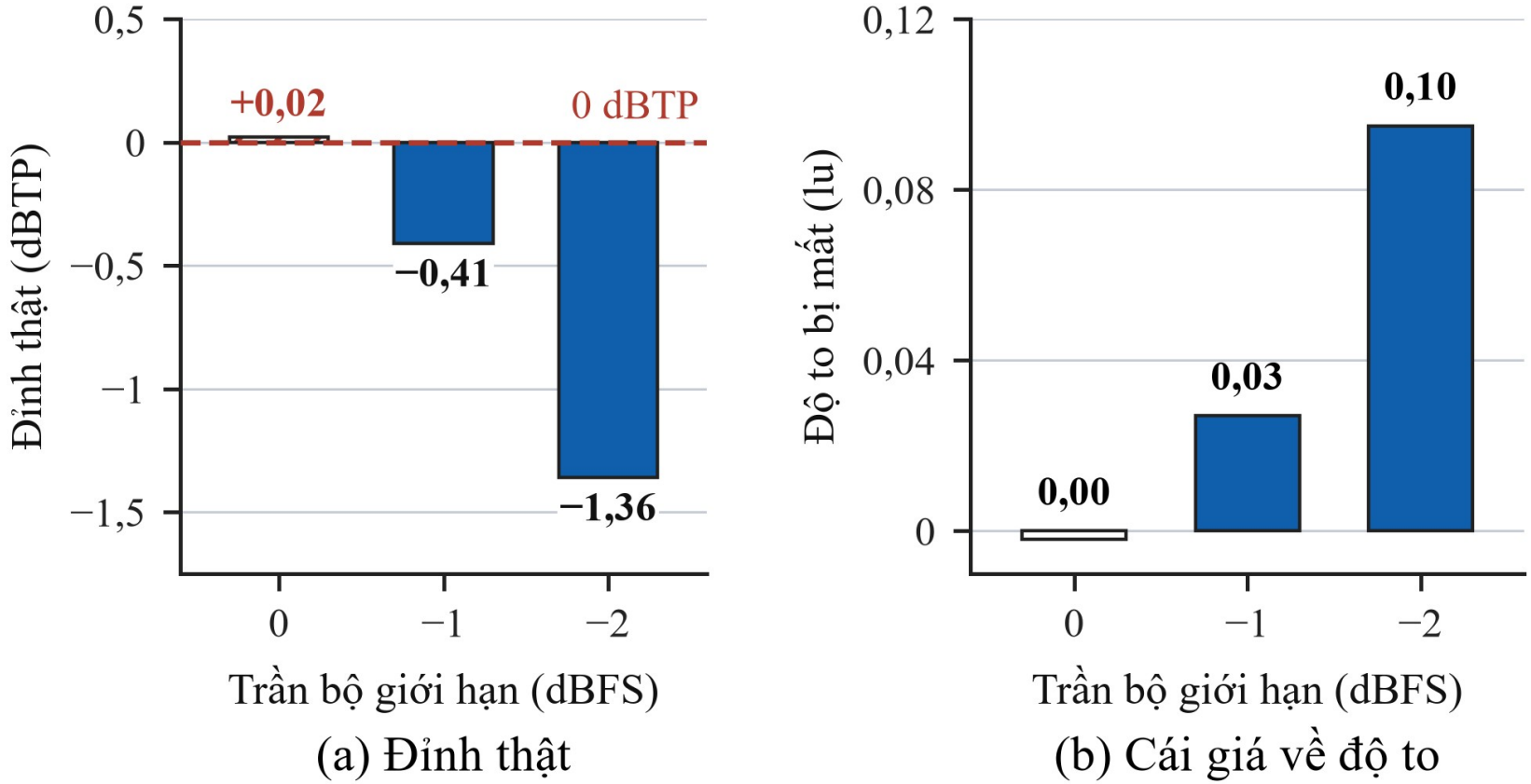
Trước khi đổi tần số, không mẫu nào chạm trần. Sau khi lên 48 kHz, đỉnh mẫu của đoạn 2 và đoạn 3 vượt 0 dBFS. Đỉnh thật là đỉnh mẫu của tệp sau khi đổi tần số, đổi định dạng hoặc qua codec.

Ba cách đặt mức cho bản giao, đo trên đoạn 1

Cách	Độ to	Đỉnh thật	Sau khi lên 48 kHz	Kết luận
Chuẩn hoá đỉnh −0,1 dBFS	−13,35 LUFS	−0,03 dBTP	0,000 % mẫu cắt	Đạt
Chuẩn hoá LUFS −14	−14,00 LUFS	−0,68 dBTP	0,000 % mẫu cắt	Đạt
Chuẩn hoá LUFS −9	−9,00 LUFS	+4,32 dBTP	0,201 % mẫu cắt	Vỡ

Cách thứ ba làm khoảng 377 mẫu mỗi giây chạm trần; mỗi mẫu là một chỗ méo.

Ba mức trần, đo sau khi đã đổi lên 48 kHz



Hình 14.13 — Ba mức trần, đo sau khi đã đổi lên 48 kHz

Mọi phép đo và đặt mức phải nằm ở cuối đường ống

Điều hứa với người nhận là con số của tệp giao, không phải của một bản trung gian. Mục 14.6 đo thêm hai cách xếp sai nữa, và cả hai sai vì cùng lý do này.

MỤC 14.5

Ghép đoạn và mối nối

- Hai đoạn do mô hình sinh không tương quan, nên luật đẳng công suất đúng.
- Phép đo mối nối phải có kiểm soát; hai phép đo đầu tiên đã sai.
- Một mili giây mờ chéo đã đủ xoá bậc thang; độ dài còn lại là cân bằng với chồng nội dung.

Hai luật mờ chéo

$$f_{\text{ra}}(t) = \cos\frac{\pi t}{2}, \qquad f_{\text{vào}}(t) = \sin\frac{\pi t}{2}, \qquad t \in [0,\ 1] \tag{14.11}$$

- Luật tuyến tính dùng hai đường thẳng; luật đẳng công suất dùng cosin và sin.
- Ở giữa chỗ chuyển: tuyến tính cho hai hệ số 0,500; đẳng công suất cho 0,707.

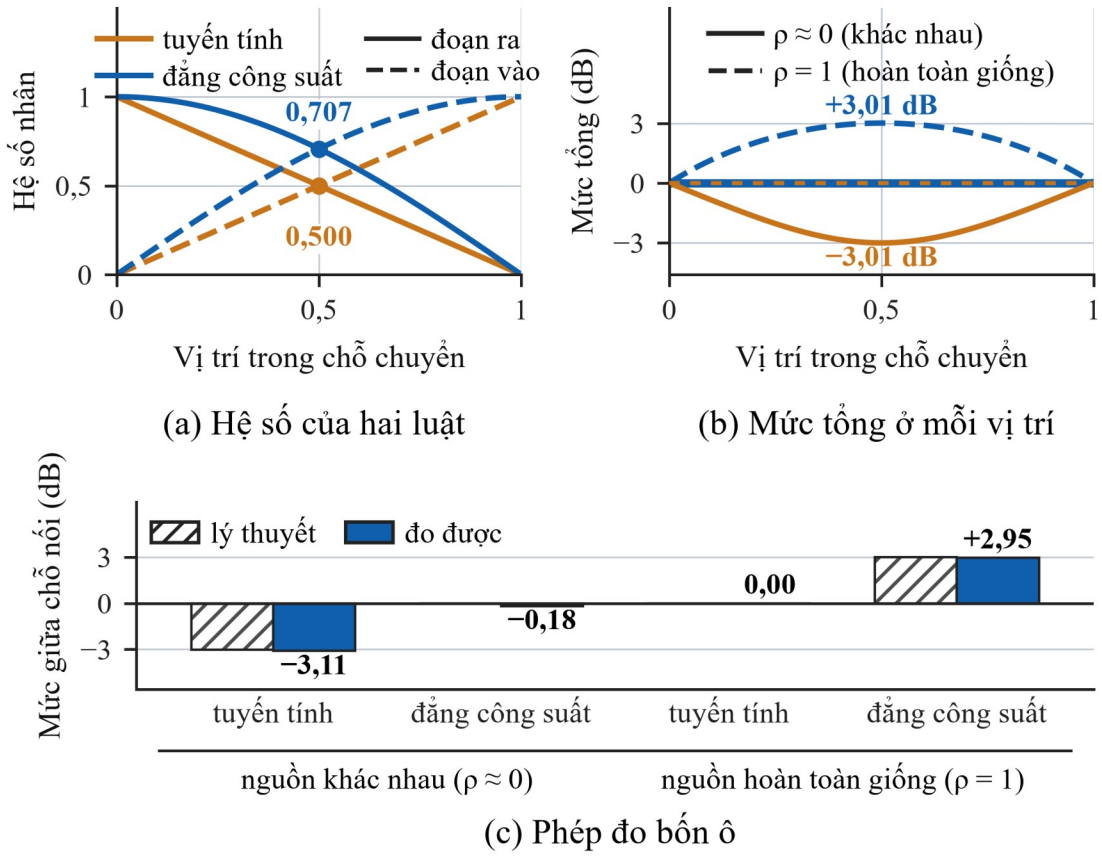
Mức tổng ở mỗi vị trí trong chỗ chuyển

$$P(t) = f_{\text{ra}}^2 + f_{\text{vào}}^2 + 2\rho f_{\text{ra}} f_{\text{vào}}$$

(14.12)

- Số hạng thứ ba chứa hệ số tương quan ρ và quyết định luật nào đúng.
- ρ = 0: luật đẳng công suất giữ mức không đổi.
- ρ = 1: luật tuyến tính mới giữ mức.

Hai luật mờ chéo, mức ở giữa chỗ nối, và phép đo bốn ô



Hình 14.14 — Hai luật mờ chéo, mức ở giữa chỗ nối, và phép đo bốn ô

Bốn ô: hai loại nguồn nhân hai luật

Nguồn	Luật	Lý thuyết	Đo được	Sai
Khác nhau (ρ ≈ 0)	Tuyến tính	−3,01 dB	−3,11 dB	0,10 dB
Khác nhau (ρ ≈ 0)	Đẳng công suất	0,00 dB	−0,18 dB	0,18 dB
Giống hệt (ρ = 1)	Tuyến tính	0,00 dB	+0,00 dB	0,00 dB
Giống hệt (ρ = 1)	Đẳng công suất	+3,01 dB	+2,95 dB	0,06 dB

Đo hệ số tương quan trước, chọn luật sau

Tuyến tính với hai đoạn không tương quan cho chỗ hõm 3 dB; đẳng công suất với hai đoạn giống hệt cho chỗ vồng 3 dB. Với kết quả do mô hình sinh, câu trả lời gần như luôn là đẳng công suất.

Ba nguyên tắc của phép đo có kiểm soát

- **Phép thử có kiểm soát:** cắt một đoạn tại vị trí k rồi bỏ 1.000 mẫu; hai nửa chỉ khác nhau chỗ nối.
- **Giữ chỗ nối đứng yên:** ghép sao cho tâm chỗ chuyển luôn rơi vào mẫu k, dù mờ chéo dài bao nhiêu.
- **Lặp ở nhiều vị trí, lấy trung vị:** bài chạy 372 vị trí k khác nhau.

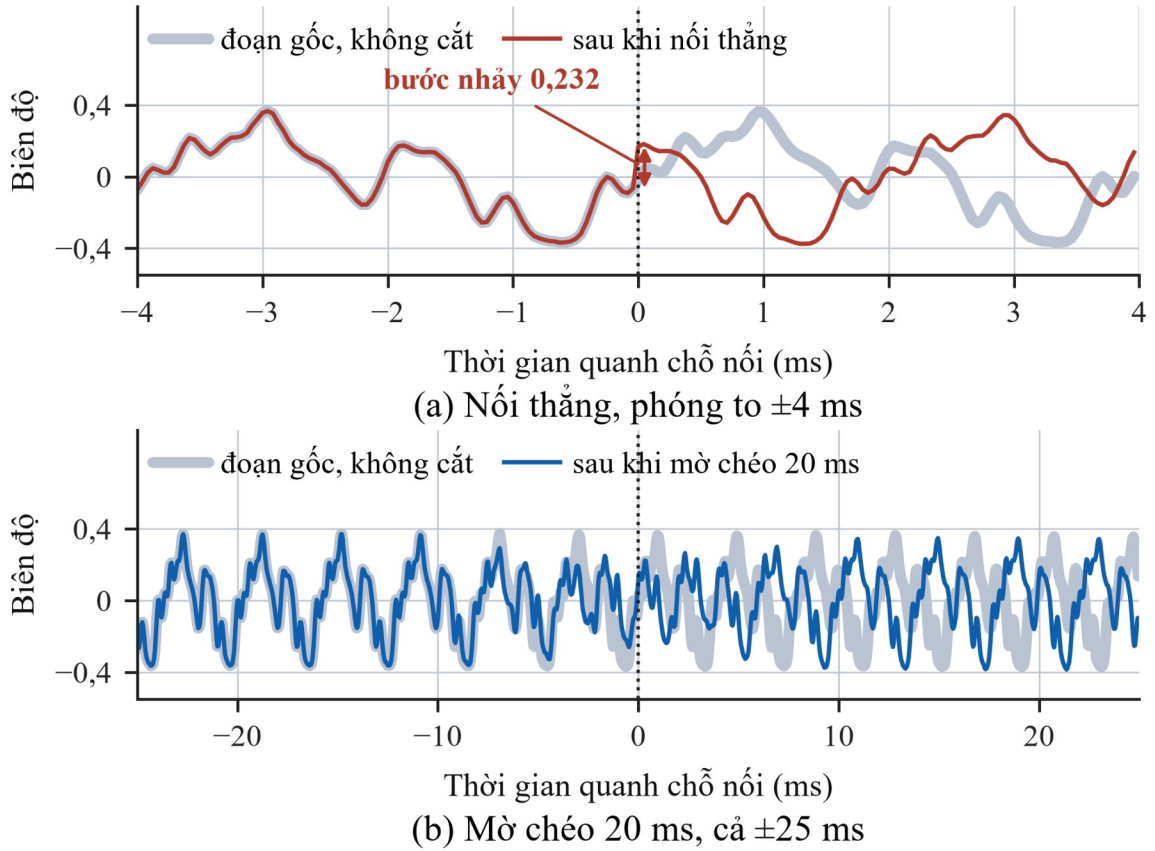
Tỉ số bước nhảy tại mối nối

$$\delta = \max_{n \in W} |y[n+1] - y[n]|, \qquad r = \frac{\delta}{\delta_{\text{nén}}}$$

(14.13)

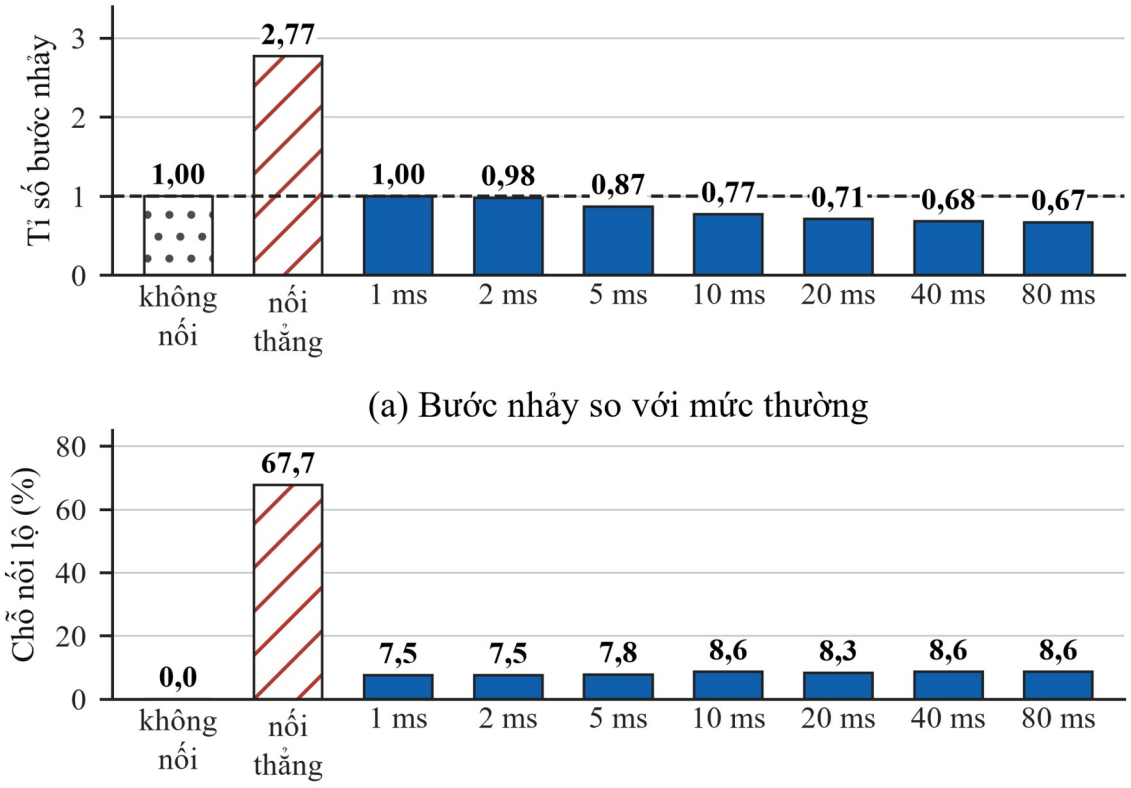
- Bước nhảy lớn nhất giữa hai mẫu liền nhau trong cửa sổ hẹp 2 ms quanh chỗ nối.
- Chia cho bước nhảy đo ở đúng chỗ ấy trên đoạn chưa cắt.
- Tỉ số bằng 1: chỗ nối không khác gì chỗ bình thường.

Cùng một chỗ cắt, nối thẳng và nối bằng mờ chéo 20 ms



Hình 14.15 — Cùng một chỗ cắt, nối thẳng và nối bằng mờ chéo 20 mili giây

Bước nhảy ở mối nối, trung vị trên 372 chỗ cắt



Hình 14.16 — Bước nhảy ở mối nối, trung vị trên 372 chỗ cắt

Nối thẳng gấp 2,77 lần mức thường; mờ chéo 1 ms đưa về 1,00

Mờ chéo dài hơn không xoá bậc thang tốt hơn. Vẫn dùng 20 đến 40 ms vì cột cuối cho thấy chồng nội dung tăng theo độ dài: chọn độ dài là cân hai thứ, không chỉ giảm bước nhảy.

MỤC 14.6

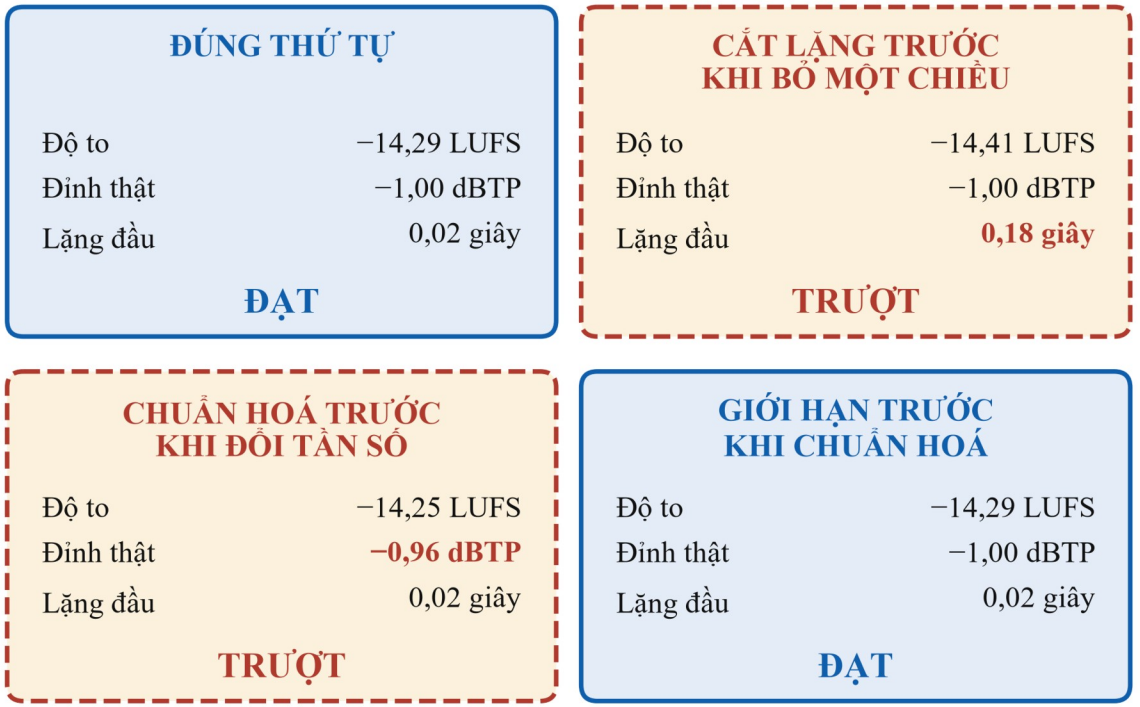
Đường ống hậu kỳ: thứ tự và ngân sách

- Bốn thứ tự của năm khối; hai trong số đó cho tệp không đạt.
- Đổi tần số nhiều chặng: một phép đo suýt cho kết luận sai.
- Hai ngân sách độc lập: độ trễ và thời gian máy.

Đường ống hậu kỳ, đúng thứ tự

```
def duong_ong_hau_ky(x, sr_vao, sr_ra=48000, lufs_dich=-14.0,
                     tran_db=-1.0, so_canh=16, cat_lang=True):
    y = bo_lech_mot_chieu(x, sr_vao)          # 1 — phai dau tien
    if cat_lang:
        y = cat_lang_dau_cuoi(y, sr_vao)      # 2 — sau khi het lech
    if int(sr_ra) != int(sr_vao):
        y = doi_sr(y, sr_vao, sr_ra, so_canh) # 3 — truoc moi phep do
    y, _, _ = chuan_hoa_lufs(y, lufs_dich, tran_db, sr_ra)
    y, _    = gioi_han(y, tran_db, 5.0, 50.0, sr_ra)
    return y
```

Bốn cách xếp cùng năm khối, và kết quả



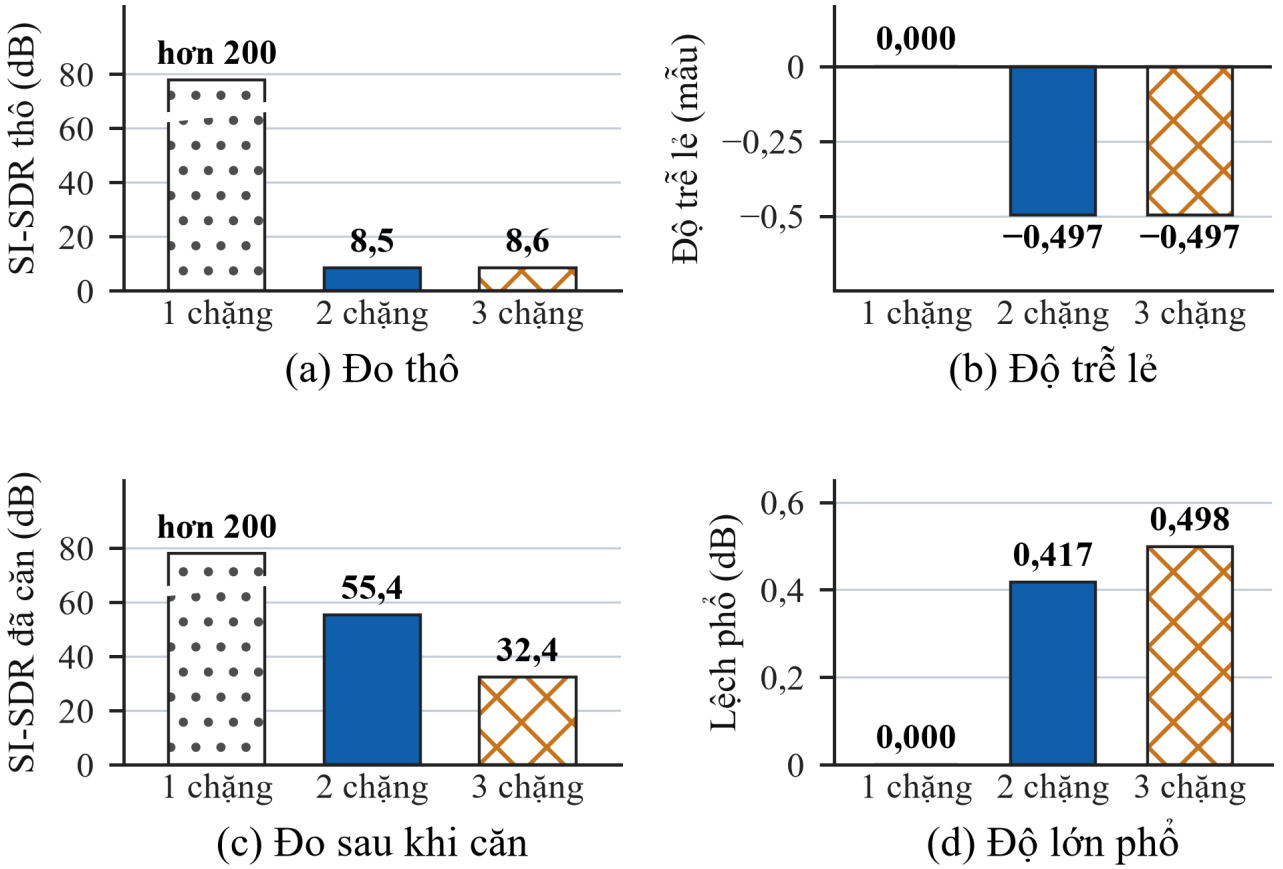
Hai cách xếp sai đều cho ra tệp nghe "vẫn được"; chỉ bộ kiểm mới chỉ ra chỗ sai.

Trượt vì: lặng đầu 0,18 giây còn nguyên · đỉnh thật −0,96 dBTP vượt trần −1,0.

Hình 14.17 — Bốn cách xếp cùng năm khối, và kết quả

Ba đường đổi, đo thô và đo sau khi căn

- Câu hỏi: đi vòng 24 → 44,1 → 48 kHz sai lệch thêm bao nhiêu so với đổi thẳng?
- SI-SDR (Scale-Invariant Signal-to-Distortion Ratio – tỉ số tín hiệu trên méo bất biến tỉ lệ), thước đo dùng suốt chương 13, cho một con số gây bất ngờ.



Hình 14.18 — Ba đường đổi tần số, đo thô và đo sau khi căn

8,5 dB là con số sai: nó phạt nửa mẫu trễ

SI-SDR so từng mẫu một nên phạt rất nặng độ trễ lẻ −0,497 mẫu. Dịch lại đúng nửa mẫu thì SI-SDR nhảy từ 8,5 lên 55,4 dB; lệch độ lớn phổ trung vị chỉ 0,0001 dB. Đi vòng qua 44,1 kHz gần như không làm sai lệch gì.

Sai số do trễ lẻ tăng theo tần số

$$|E(f)| \approx 2\pi f \Delta t \, |X(f)|$$

(14.14)

- Dịch một khoảng Δt làm pha xoay tỉ lệ với tần số.
- Vì vậy biên độ sai khác cũng tỉ lệ với tần số.
- Đường ba chặng sai thật: 32,4 dB sau căn, vì mất hẳn phần dải trên 11 kHz.

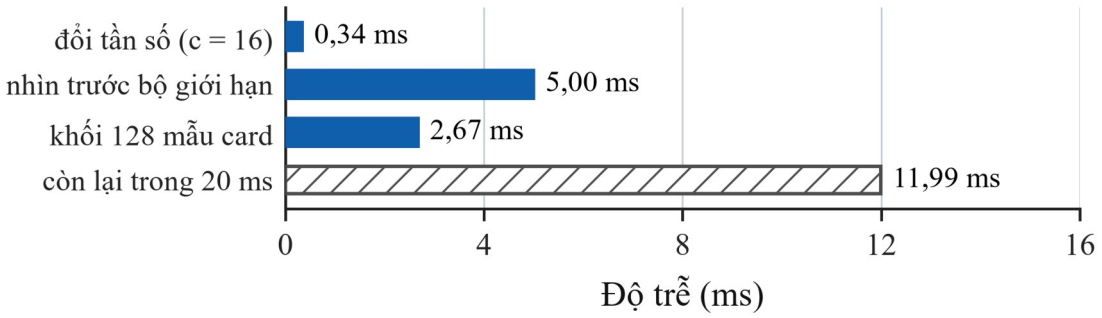
Độ trễ và thời gian máy là hai chi phí độc lập

$$\tau_{\text{tổng}} = \sum_i \tau_i , \qquad \tau_{\text{còn}} = \tau_{\text{hạn}} - \tau_{\text{tổng}}$$

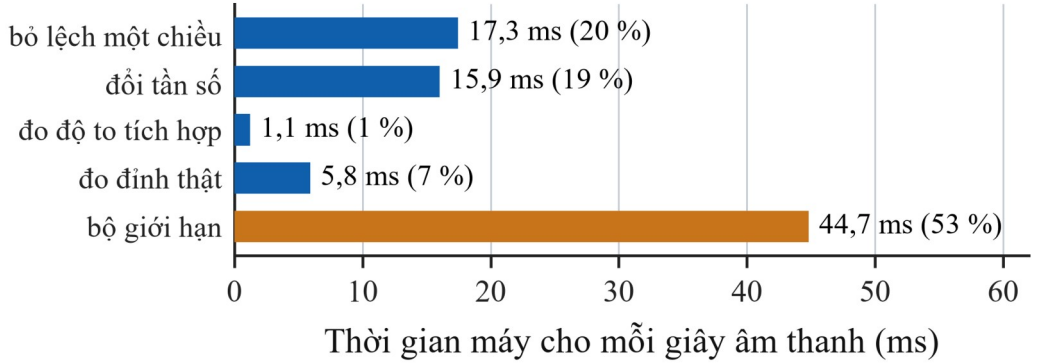
(14.15)

- Độ trễ quyết định đường ống có dùng được khi chơi trực tiếp hay không.
- Thời gian máy quyết định dựng một bản dài mất bao lâu.
- Nhầm lẫn giữa hai loại chi phí là lỗi thường gặp khi thiết kế.

Hai ngân sách của cùng một đường ống



(a) Độ trễ: ba khối dùng 8,01 ms, còn dư 11,99 ms



(b) Thời gian máy: tổng 84,8 ms, nhanh gấp 11,8 lần thời gian thực

Hình 14.19 — Hai ngân sách của cùng một đường ống

MỤC 14.7

Đo và kiểm tự động

- Bộ kiểm bàn giao: một lệnh, bảy dòng đạt hay không đạt.
- Đọc bảng kiểm theo thứ tự phụ thuộc, không đọc rời từng dòng.
- Dung sai hồi quy phải đo ra, không đoán ra.

Bộ kiểm bàn giao chạy trên đoạn 1

Tiêu chí	Ngưỡng	Bản thô	Đạt?	Sau ống	Đạt?
độ to tích hợp (LUFS)	−14,0 ± 0,5	**−17,76**	**trượt**	**−14,29**	đạt
đỉnh thật (dBTP)	≤ −1,0	−4,47	đạt	**−1,00**	đạt
lệch một chiều	≤ 0,0010	**0,0122**	**trượt**	**0,0000**	đạt
lặng đầu (ms)	≤ 250	0	đạt	**20**	đạt
lặng cuối (ms)	≤ 250	0	đạt	**0**	đạt
mẫu chạm trần (%)	≤ 0,000	0,000	đạt	**0,000**	đạt
vách phổ (Hz)	≥ 12.000	12.000	đạt	**13.078**	đạt

Đoạn 1: bản thô đạt 5/7, sau đường ống đạt 7/7.

Dòng lặng đầu của bản thô đọc 0 ms và được tính là đạt, nhưng kết quả này sai: lệch một chiều làm mọi khối đều có tín hiệu nên máy dò không thấy lặng. Một tiêu chí trượt có thể che một tiêu chí khác.

Hình 14.20 — Bộ kiểm bàn giao chạy trên đoạn 1

Dòng "lặng đầu" đạt không có nghĩa là không có lặng

Bản thô trượt "lệch một chiều", mà lệch một chiều làm máy dò lặng mù. Khi tiêu chí nền tảng trượt, mọi tiêu chí phụ thuộc vào nó coi là chưa đo được. Thứ tự phụ thuộc: lệch một chiều trước lặng hai đầu; tần số lấy mẫu trước đỉnh thật; cả hai trước độ to.

Ba phép thử để đặt dung sai hồi quy

Chạy lại nguyên tham số

hoàn toàn giống từng bit
sai lệch lớn nhất: 0
kiểm bằng so sánh bằng nhau

Đổi c từ 16 xuống 8

SI-SDR 40,4 dB
độ to lệch −0,026 lu
đỉnh thật lệch 0,00 dB

Đổi c từ 16 lên 32

SI-SDR 43,5 dB
độ to lệch −0,088 lu
đỉnh thật lệch 0,00 dB

Dung sai đặt ra từ các phép thử này

Đổi một tham số hợp lệ làm độ to xê dịch tới 0,088 lu và đỉnh thật không xê dịch. Vậy đặt dung sai hồi quy ở 0,2 lu và 0,05 dBTP: lớn hơn dao động hợp lệ, nhỏ hơn ngưỡng nghe ra. Vượt dung sai là lỗi thật, không phải nhiễu đo.

Hình 14.21 — Ba phép thử để đặt ra dung sai hồi quy

MỤC 14.8

Mốc bài tập lớn: giai đoạn 4

- Giai đoạn 4 không đòi kỹ thuật mới: ghép thứ đã có thành quy trình một lệnh.
- Bốn thứ phải nộp; bộ kiểm phải ra 7 trên 7.
- Bảy mươi trên một trăm điểm chấm bằng số đo.

Bốn thứ phải nộp ở giai đoạn 4

1. Tệp WAV bản giao

48.000 Hz, 16 hoặc 24 bit, dài 60 đến 180 giây, ghép từ ít nhất ba đoạn nguồn khác nhau

2. Mã dựng nên tệp

chạy được bằng một lệnh, không sửa tay bước nào, không có tham số hằng số trong hàm

3. Bản ghi đo

kết xuất của bộ kiểm bàn giao: bảy dòng, ghi rõ số đo và ngưỡng, kèm bản ghi của bản thô

4. Một trang giải trình

chọn tần số nào và vì sao; c bằng bao nhiêu; luật mờ chéo nào và hệ số tương quan đo được

Hai điều kiện cần, không thương lượng

Bộ kiểm phải ra 7/7: một dòng trượt là bản nộp chưa đạt, dù nghe hay. Chạy lại mã phải cho ra tệp hoàn toàn giống; nếu dùng số ngẫu nhiên thì phải cố định hạt giống.

Hình 14.22 — Bốn thứ phải nộp ở giai đoạn 4

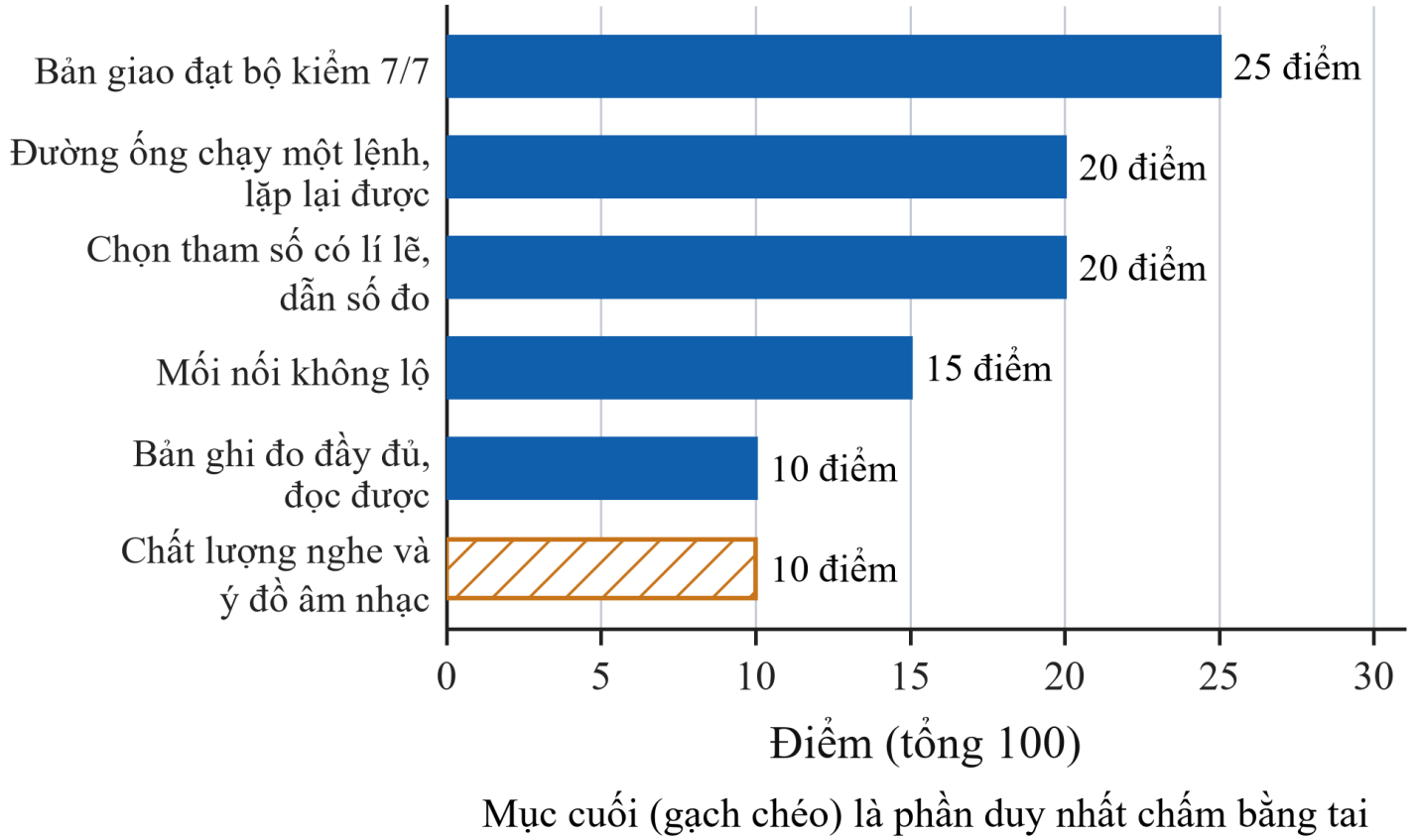
Bốn thứ phải nộp và tiêu chuẩn của từng thứ

Nộp gì	Tiêu chuẩn
Tệp WAV bản giao	48.000 Hz, 16 hoặc 24 bit, 60 đến 180 giây, ghép từ ít nhất ba đoạn
Mã dựng nên nó	Chạy một lệnh từ nguồn tới tệp giao; không tham số hằng số lẫn trong hàm
Bản ghi đo	Kết xuất bộ kiểm bảy dòng, kèm bản ghi đo của bản thô
Một trang giải trình	Tần số, c và luật mờ chéo: mỗi lựa chọn dẫn số đo

Bộ kiểm phải ra 7 trên 7, và chạy lại phải cho tệp giống hệt

Một dòng trượt là bản nộp chưa đạt, dù nghe hay đến đâu. Tiêu chí không áp dụng được thì nêu lý do trong trang giải trình và đổi ngưỡng có căn cứ, không bỏ dòng. Dùng số ngẫu nhiên thì phải cố định hạt giống.

Thang chấm giai đoạn 4, tổng 100 điểm



Hình 14.23 — Thang chấm giai đoạn 4, tổng 100 điểm

Thang chấm chi tiết

Mục	Điểm	Chấm bằng gì
Bản giao đạt bộ kiểm 7/7	25	Chạy bộ kiểm trên tệp nộp
Đường ống một lệnh, lặp lại được	20	Chạy hai lần, so hai tệp
Chọn tham số có lí lẽ, dẫn số đo	20	Đọc trang giải trình
Mối nối không lộ	15	Đo bước nhảy tại từng mối nối
Bản ghi đo đầy đủ, đọc được	10	Đủ trước và sau; có ngưỡng
Chất lượng nghe và ý đồ âm nhạc	10	Nghe; phần duy nhất chấm bằng tai

Ba lỗi khiến bản nộp trượt nhiều nhất

- **Đo không ở cuối đường ống:** đo đỉnh thật ngay sau bộ giới hạn rồi mới đổi tần số (mục 14.6.1).
- **Ghép mà không đo tương quan:** dùng luật tuyến tính vì "tự nhiên hơn", gây chỗ hõm ba decibel tại mỗi mối nối.
- **Sót tham số hằng số trong thân hàm:** thường là 48000 viết thẳng trong phép tính; đổi tần số đích sẽ sai âm thầm.

Mọi tham số phải nằm ở một chỗ duy nhất.

MỤC 14.9

Thực hành, bài tập và tổng kết

- Ba bài thực hành, mỗi bài khoảng mười lăm phút, không cần tệp ngoài.
- Bảy bài tập, trong đó hai bài chạy lại phép đo trên vật liệu của người học.
- Bốn điều đáng mang ra khỏi chương.

160 / 147

tỉ số tối giản của 44,1 và 48 kHz; lí do phải dùng đa pha

−94,5 dB

tạp của bộ lọc đa pha (c = 16), so với −27,7 dB của lấy mẫu gần nhất

+1,41 dBTP

đỉnh thật của đoạn đã chuẩn hoá đỉnh về −0,1 dBFS

2,77 lần

bước nhảy ở mối nối thẳng, so với mức bình thường của đoạn

0,497 mẫu

độ trễ lẻ do đi vòng qua 44,1 kHz, đủ làm SI-SDR đọc sai 47 dB

7 / 7

số dòng bộ kiểm phải đạt trước khi gọi là một bản giao

Chương 15: bản quyền, đạo đức, phát hiện vi phạm

Chương này lo chất lượng kỹ thuật của bản giao; chương sau lo câu hỏi bản ấy có được phép tồn tại hay không: quyền đối với giọng, deepfake âm thanh, dấu vết codec, cách bảo vệ sản phẩm của chính người làm.

Thực hành trên vật liệu do mô-đun sinh ra

- **Thực hành 1, sáu chỉ số đầu tiên:** đo sáu chỉ số của ba đoạn, lặp lại sau khi bỏ lệch một chiều; chỉ ra chỉ số nào đổi.
- **Thực hành 2, ba cách đổi tần số:** đổi hình sin 997 Hz từ 44,1 lên 48 kHz bằng ba cách, vẽ phổ, đối chiếu mục 14.3.4.
- **Thực hành 3, hai luật mờ chéo:** ghép nhiễu trắng độc lập rồi hai đoạn giống hệt; bốn con số khớp bảng 14.19 trong 0,2 dB.

Bài tập 1 đến 4

- **Bài 1, dò vách phổ:** dò ba tệp nhạc bất kỳ, đoán các khâu đã qua; giữ tệp có vách ở khoảng 15,5 kHz cho chương 15.
- **Bài 2, đường cong số cạnh:** dựng lại bảng 14.10 cho đường 48 xuống 24 kHz; chỉ ra khác biệt lớn nhất.
- **Bài 3, đi xuống không đảo ngược:** 48 → 24 → 48 kHz, đo SI-SDR trên tín hiệu đầy dải và tín hiệu dưới 9,6 kHz.
- **Bài 4, đỉnh thật của chính mình:** đo đỉnh mẫu và đỉnh thật của năm tệp; đếm số tệp vượt 0 dBTP.

Bài tập 5 đến 7

- **Bài 5, mờ chéo cho giọng nói:** lặp lại phép đo mục 14.5.3 trên giọng nói; đưa khuyến nghị độ dài riêng.
- **Bài 6, thêm tiêu chí thứ tám:** có số đo, ngưỡng và lập luận theo cách mục 14.7.3; gợi ý: dải độ to, cân bằng trái phải, tương quan hai kênh.
- **Bài 7, điểm dễ sai của thước đo:** dựng lại phép đo mục 14.6.2 trên vật liệu riêng; nêu kết luận sai nếu chỉ nhìn con số đầu.

Cùng những khối ấy, xếp sai thứ tự thì cho tệp không đạt

Chỉ có phép đo mới phân biệt được hai trường hợp, tai thì không. Chương 15 chuyển sang câu hỏi khác: bản giao có được phép tồn tại không (quyền đối với giọng, deepfake âm thanh, dấu vết codec).

Những điều cần nhớ

1. Chương không có khái niệm mới: giá trị nằm ở việc xếp các khối đã học theo đúng thứ tự.
2. Mọi phép đo và đặt mức phải nằm ở cuối đường ống; con số hứa với người nhận là của tệp giao.
3. Đổi tần số bằng bộ lọc đa pha: −94,5 dB rác, chi phí không phụ thuộc L; bỏ mẫu không lọc là lỗi không sửa được.
4. Đỉnh thật khác đỉnh mẫu; trần −1 dBTP cùng mức −14 LUFS là cặp tham số giao hàng của chương.
5. Chọn luật mờ chéo theo hệ số tương quan đo được; một mili giây đã xoá bậc thang, 20 đến 40 ms cân với chồng nội dung.
6. Một tiêu chí trượt có thể che tiêu chí khác; thước đo có thể kết luận sai (nửa mẫu trễ làm SI-SDR đọc 8,5 dB).
7. Dung sai hồi quy phải đo ra từ dao động hợp lệ, không đoán.