

RaySound

Simulazione acustica ibrida, design inverso e identificabilità diagnosticata

Validazione su sale reali del benchmark BRAS — CR4, CR3, CR2, CR1

Ing. Stefano Fante

ST-LINE S.r.l.

Report scientifico-tecnico • 2 settembre 2026 • ST-LINE S.r.l.

Sommario

RaySound è un motore di acustica delle sale che predice il campo, lo *progetta* all'inverso (dal target di capitolato al trattamento) e ne *diagnostica* l'identificabilità prima di calibrare — gli ultimi due passi assenti dai simulatori commerciali (Odeon, CATT, Treble). È validato su quattro sale reali del benchmark pubblico BRAS (TU Berlin). Il T_{60} *diffuso analitico* (Eyring su V/S misurati, coefficienti BRAS *fitted*, aria alle condizioni *documentate dal benchmark*) riproduce la misura entro l'1.1% e l'1.3% (CR4 / CR3, banda-core 250–2000 Hz, 1/3 d'ottava, senza ricalibrazione) — accordo che quei coefficienti in parte *impongono*, essendo l'inversa di Eyring sullo stesso V/S (App. M). Il *ray-tracer*, che è una predizione indipendente, sta sulle stesse bande al 17.1% e all'11.5%; il pattern spaziale del C_{50} predetto-vs-misurato è a $+0.89/+0.68$ (su 5 posizioni: intervallo di confidenza al 95% molto ampio, §3.6), con residuo di livello $-0.17/+0.70$ dB. Il design *position-aware* sposta il progetto fino al $\sim 12\%$ scegliendo *quale* superficie trattare, non solo quanta area — effetto che il benchmark indipendente Arni misura a $114\times$ il rumore. Su CR2 (sala rigida) l'identificabilità è *negativa*: il Jacobiano materiale è rank-deficient (3/5) e il motore lo segnala invece di restituire una stima non difendibile; sotto Schroeder l'osservabile modale α_m separa l'orientazione delle pareti (rango 3/6 contro 1/6 dell'energetico). Su CR1 (stanze accoppiate, l'unica scena double-slope del benchmark) un forward two-part chiude il C_{50} del campo accoppiato fit-free. Il posizionamento è on-premise (sovranità del dato per gli appalti PA), nativo per i CAM, con l'incertezza esposta come elemento di due diligence. Le note metodologiche (coefficienti BRAS, regimi del forward, etichette di validazione) sono raccolte nell'Appendice M.

1 Introduzione: che cosa fa RaySound, in pratica

Contesto. La progettazione acustica di sale (aule, auditorium, sale conferenza) è sempre più un requisito d'appalto: i Criteri Ambientali Minimi (CAM) italiani fissano soglie su tempo di riverberazione e intelligibilità che il progettista deve dimostrare, non solo dichiarare. Il destinatario è il consulente acustico che deve consegnare un numero difendibile in una due diligence. I simulatori consolidati (Odeon, CATT, Treble) risolvono il *forward* (data la sala e i materiali, predire il campo), ma lasciano scoperti i due passi che contano in fase di progetto: progettare all'inverso (dal target di capitolato al trattamento delle superfici, oggi un prova-e-riprova manuale) e diagnosticare se i dati bastano a determinare ciò che si vuole stimare. RaySound affronta entrambi come capacità del motore, sopra un forward fisicamente fondato e differenziabile.

Architettura funzionale. L'input è la *geometria* della sala (una mesh), l'assorbimento $\alpha(f)$ per superficie, le posizioni di sorgente e ricevitori. Da qui, tre capacità.

- **Forward (sala $\rightarrow$ descrittori).** Il campo si tratta diviso per frequenza. Sotto la frequenza di Schroeder il campo è *modale* (pochi modi risolti singolarmente) e va risolto come onda: un forward d'onda DG (discontinuous Galerkin) [14, 16]. Sopra Schroeder il campo è *diffuso* e un *ray-tracer* (tracciamento geometrico di raggi) costruisce la RIR (*room impulse response*, la risposta all'impulso della sala). Il T_{60} in regime diffuso è fornito da Eyring sul volume e sulle aree reali, non dal ray-tracer, il cui forward diffuso sovrastima il decadimento su geometria complessa (§2.2). L'uscita sono i descrittori di capitolato: T_{60} , C_{50} , STI.

- **Design inverso (target $\rightarrow$ trattamento).** Il motore ottimizza quali superfici trattare e quanto, per portare il C_{50} entro target e il T_{60} nell'intervallo. La procedura è *position-aware*: pesa ogni superficie per quanto spesso il campo la colpisce, non per la sua area, poiché a parità di area trattata la posizione del trattamento cambia il risultato. Il target guida il progetto, al posto del prova-e-riprova manuale.
- **Identificabilità.** Prima di calibrare i materiali su una misura, il motore stabilisce se sono recuperabili dai dati: il condizionamento del problema inverso (SVD del Jacobiano descrittori-parametri) lo quantifica. In una sala acusticamente rigida i materiali non sono identificabili, e il motore lo quantifica invece di restituire una stima non difendibile.

Flusso d'uso. (i) Si carica la sala, si gira il forward, si verifica se i descrittori rispettano i CAM. (ii) Se non rientrano, il design inverso indica il trattamento; si applica e si rifà il forward. (iii) Se è disponibile una RIR misurata, l'identificabilità indica se l'inversione è ben posta; in tal caso restituisce l' $\alpha(f)$ reale per superficie con la sua incertezza.

Sintesi. RaySound predice, progetta e certifica. Il forward ibrido onde+ray-tracing esiste anche altrove (es. Treble); il differenziatore sono gli ultimi due passi, il design inverso e la diagnosi di identificabilità, assenti dai simulatori commerciali. La validazione è su quattro sale reali del benchmark pubblico BRAS; i numeri sono in §3. Le sezioni che seguono danno il metodo e la fisica (§2) e la validazione (§3).

2 Metodo

2.1 Forward ibrido onde + ray-tracing

Sotto Schroeder, il campo di pressione p obbedisce all'equazione delle onde con contorno a impedenza complessa $Z(\omega)$ [3]:

$$\ddot{p} = c^2 \nabla^2 p, \quad \partial_n p = -\frac{\rho_0}{Z(\omega)} \dot{p} \quad (\text{locally reacting}), \quad (1)$$

con $Z(\omega)$ impedenza di superficie *fisica* [2] ($\text{Pa} \cdot \text{s/m}$, $\text{Re } Z > 0$). La correttezza del contorno all'incidenza obliqua è verificata contro il coefficiente di riflessione analitico

$$R(\theta) = \frac{Z \cos \theta - \rho_0 c}{Z \cos \theta + \rho_0 c}, \quad \max_{\theta} |R_{\text{num}}(\theta) - R(\theta)| = 9.5 \times 10^{-3}, \quad (2)$$

errore massimo (in modulo) sull'incidenza $\theta \in [0, 85^\circ]$ alla banda di test. Sopra Schroeder, il ray-tracer traccia riflessioni speculari (image-source [13]) e diffuse (densità di echi alla Vorländer [25]). La scala forward usa la convenzione fisica completa $p(r) \propto 1/(4\pi r)$ e l'assorbimento d'aria segue ISO 9613-1 [28]. Per i grandi volumi, dove il modale d'onda è computazionalmente pesante, il T_{60} coi coefficienti BRAS fitted è stimato con Eyring [10, 11] sulle aree e sul volume reali della mesh:

$$T_{60} = \frac{0.161 V}{-S \ln(1 - \bar{\alpha}) + 4mV}, \quad \bar{\alpha} = \frac{1}{S} \sum_i \alpha_i(f) S_i. \quad (3)$$

2.2 Il forward ibrido: regimi e giustificazione fisica

L'ibrido è un *router per regime*: ogni banda va al metodo valido nel suo regime. La soglia è la frequenza di Schroeder, che separa il modale (modi risolti singolarmente) dal diffuso (modi sovrapposti):

$$\frac{dN}{df} = \frac{4\pi V f^2}{c^3} + \frac{\pi S f}{2c^2}, \quad M(f) = \frac{dN}{df} \cdot \frac{2.2}{T_{60}}, \quad M(f_{\text{Sch}}) = 3, \quad (4)$$

con V in m^3 , S la superficie totale, T_{60} in secondi; dN/df è la densità modale di Weyl completa (termine di volume più termine di superficie, pareti rigide/Neumann, segno positivo), non il solo volume: sotto i ~ 63 Hz il termine di superficie è comparabile a quello di volume ($\geq 50\%$ per CR3/CR4) e il solo-volume sotto-conta i modi. La densità completa è validata contro l'eigensolver DG [15] (§2.1): su uno shoebox $4 \times 3 \times 2.5$ il conteggio vol+sup recupera l'87% dei modi veri a 143 Hz (contro il 46% del solo-volume; il residuo è il termine di spigolo $L/8c$ del terz'ordine, qui dichiarato non incluso). M è l'*overlap modale* (numero medio di modi entro la larghezza di banda di mezza potenza); il criterio classico per il campo diffuso è $M \gtrsim 3$, e f_{Sch} è la frequenza dove $M = 3$. Per le sale in esame, M (densità completa) dai T_{60} misurati è:

	63 Hz	125 Hz	250 Hz
CR3 (3331 m^3)	8.8	28	101
CR4 (8657 m^3)	17	57	206

Tutti $\gg 3$: queste sale sono già sopra Schroeder a 63 Hz, e con la densità completa lo sono a maggior ragione ($f_{\text{Sch}} \approx 20$ Hz per CR4, ≈ 30 Hz per CR3, più bassa del solo-volume, 30/44 Hz, perché i modi in più anticipano l'onset diffuso).

Il confine quantizzato alle bande (conservativo). Il router non taglia al valore grezzo di f_{Sch} (che cadrebbe a metà di una banda, lavoro al limite) ma alla griglia di bande: la banda in cui cade f_{Sch} è ancora in parte modale e va interamente al forward d'onda; il ray-tracer parte dalla prima banda completamente sopra f_{Sch} (bordo inferiore $> f_{\text{Sch}}$). Così l'overlap modale alla prima banda-GA (vicino a f_{Sch} , quindi a frequenza più bassa dei 63 Hz della tabella sopra) è $M > 3$ per costruzione, con almeno la mezza-larghezza di banda di margine (CR3 $M \ 3.0 \rightarrow 4.3$, CR4 $2.9 \rightarrow 4.0$, CR2 $3.3 \rightarrow 5.0$, a $1/3$ d'ottava): margine conservativo per costruzione. (M cresce con la frequenza: questi valori sono alla prima banda-GA presso f_{Sch} , non i M a 63 Hz della tabella, 8.8/17, che è più alta.) Sulle sale grandi (CR3/CR4, f_{Sch} sotto i 63 Hz) il ray-tracer parte già dalla banda più bassa, conservativo per costruzione.

Campo diffuso: Eyring vs forward-GA. Per $M \gg 3$ il T_{60} medio è ben descritto da Eyring (eq. (3)); il forward *geometrico*-diffuso lo sovrastima del $\sim 10\text{--}15\%$ per la distribuzione delle lunghezze di cammino libero (effetto Kuttruff). Perciò il report adotta Eyring per il T_{60} e il forward-GA per gli altri descrittori (early, C_{50} , mappe), che non ereditano il bias — isolamento della causa e giustificazione completa in App. M.2.

La mappa dei regimi. Tre regimi, tre forward, ciascuno valido dove gli altri non lo sono. *Modale* ($M < 3$, sotto Schroeder, sale piccole): Eyring e il ray-tracing geometrico non vedono i modi discreti, e serve il *forward d'onda* (validato contro l'analitico, §2.1). f_{Sch} cresce al ridursi del volume: in una sala piccola (scala CR2, $V \approx 159 \text{ m}^3$) sale a $\sim 80\text{--}120$ Hz, lasciando una banda modale larga dove solo il forward d'onda è corretto. *Diffuso* ($M \gg 3$, sopra Schroeder, sale medie/grandi): il *ray-tracer* dà la RIR completa (early + coda) ed *Eyring* il T_{60} (limite statistico del dg nel diffuso) $\rightarrow$ il forward d'onda non aggiunge nulla che giustifichi il suo costo. *Design spaziale* (tutte le sale): conta dove si tratta, non solo quanto $\rightarrow$ *ray-tracer position-aware*; Eyring total-only è cieco alla distribuzione.

Necessità dell'ibrido. Il target è l'intero spettro di sale (piccola $\rightarrow$ grande) e tutti i descrittori — un dominio che attraversa Schroeder e che nessun metodo singolo copre: senza il dg le sale piccole/modali, senza il ray-tracer la RIR completa, l'early, il C_{50} e il design position-aware, col solo Eyring il solo T_{60} diffuso. La copertura è asimmetrica (banda modale stretta sulle sale grandi, $f_{\text{Sch}} \sim 20\text{--}30$ Hz; larga sulle piccole) e l'ibrido copre entrambi gli estremi. Alle bande diffuse né il ray-tracer né il forward d'onda 3D battono Eyring su V/S misurati ($M \gg 3$; cfr. App. M.2), ciò che il report adotta.

2.3 Design inverso (target → trattamento)

Il design minimizza lo scarto fra descrittori predetti $d_k(\mathbf{m})$ e target realizzabile d_k^* , sui moltiplicatori di assorbimento $\mathbf{m}$ per superficie:

$$\min_{\mathbf{m}} J(\mathbf{m}) = \sum_k w_k (d_k(\mathbf{m}) - d_k^*)^2. \quad (5)$$

Il design è position-aware: ottimizza quale superficie trattare, non solo quanta area. Due ingredienti. Il C_{50} è calcolato dalle riflessioni reali dell’image-source (già position-aware: dipende da quali superfici colpiscono le prime riflessioni, non da uno stand-in):

$$C_{50} = 10 \log_{10} \frac{\int_0^{50 \text{ ms}} p^2 dt}{\int_{50 \text{ ms}}^{\infty} p^2 dt}. \quad (6)$$

Il T_{60} usa Eyring *hit-weighted* (pesato per frequenza-di-colpo): l’assorbimento medio non è pesato per area (come in Eyring classico) ma per quanto spesso il campo colpisce ciascuna superficie,

$$T_{60} = \frac{0.161 V}{-S \ln(1 - \bar{\alpha}_{\text{eff}}) + 4mV}, \quad \bar{\alpha}_{\text{eff}} = \sum_m h_m \alpha_m(f), \quad (7)$$

dove h_m è la frazione di rimbalzo sulla superficie m stimata dal ray-tracer (geometrica, indipendente dai materiali, $\sum_m h_m = 1$). **Ancoraggio a Eyring.** Nel limite diffuso il campo colpisce le superfici in proporzione all’area, $h_m \rightarrow S_m/S_{\text{tot}}$, e allora $\bar{\alpha}_{\text{eff}} = \bar{\alpha}$: il T_{60} position-aware degenera esattamente a Eyring (verificato a rel = 0 a precisione macchina, per ogni α con $h_m = S_m/S_{\text{tot}}$ e per α uniforme con qualsiasi h_m). Questo ancoraggio rende l’effetto position-aware isolato e attribuibile: non è un forward diverso che introduce un proprio scarto di modello (rispetto a Eyring un ray-tracer libero devia del $\sim 3\%$, lo stesso ordine dell’effetto cercato e quindi confondente); è Eyring ri-pesato, che coincide con Eyring quando la distribuzione non conta e se ne discosta solo quando conta. Il gradiente è calcolato col metodo complex-step [18, 19] (esatto a meno dell’arrotondamento macchina) e validato contro differenze finite: rel = $\max |\nabla^{cs} J - \nabla^{fd} J| / \max |\nabla^{fd} J| \leq 8 \times 10^{-10}$; alimenta una minimizzazione bound-constrained L-BFGS-B [20].

2.4 Identificabilità e UQ

La SVD del Jacobiano $J = U\Sigma V^\top$ dà il condition number $\kappa(J) = \sigma_{\text{max}}/\sigma_{\text{min}}$, il rango effettivo e lo spazio nullo. Due soglie distinte: una *percettiva* (1 JND, per stabilire se una direzione di parametri è acusticamente rilevante) e una *numerica* ($\sigma_i/\sigma_{\text{max}} > 10^{-6}$, per il rango). Un problema è ben posto se il rango è pieno e κ moderato. L’incertezza è propagata via Laplace / Gauss-Newton [23] con Hessiana $H \approx J^\top W J$ (pesi W dei descrittori):

$$\text{Cov}(\mathbf{m}) \approx (J^\top W J)^{-1} \sigma^2, \quad (8)$$

dove l’inversa va intesa come *pseudo-inversa* via SVD quando J è rank-deficient (es. CR2): le direzioni nello spazio nullo emergono come varianza non vincolata, non come un’inversa che non esiste.

2.5 Invarianti di campo diffuso: che cosa sono e perché

Limiti dei descrittori ISO. I descrittori normativi (T_{20} , C_{50} , EDT) riassumono la risposta impulsiva in pochi numeri costruiti sul *decadimento dell’energia*: quanto a lungo il suono persiste, quanta energia arriva presto rispetto a tardi. È una rappresentazione compatta ma *lossy*: due risposte impulsive fisicamente molto diverse possono avere lo stesso C_{50} o lo stesso T_{20} . Quando il campo non è pienamente diffuso, o quando ciò che conta è la *struttura temporale* delle riflessioni

(non solo la loro energia integrata), i descrittori ISO diventano insensibili a differenze reali, e una calibrazione o una validazione che si basa solo su di essi non può vedere ciò che non è codificato nel descrittore.

Statistica del campo. Invece di misurare solo quanta energia decade e quanto in fretta, si misura la *statistica del campo*: come sono distribuite le riflessioni nel tempo, quando il campo passa da poche riflessioni discrete (regime early, geometrico) a una nube densa e gaussiana (regime late, diffuso). Queste quantità sono *invarianti*: robuste alla posizione del microfono e alla fase, perché descrivono una proprietà statistica del campo nella sala, non il dettaglio di una singola risposta. Misurano un asse dell’informazione che i descrittori energetici non espongono.

Usiamo tre invarianti, ciascuno con un significato fisico diretto:

- **Densità di echi normalizzata** (NED, Abel–Huang 2006) — misura, istante per istante, quanto il campo è già “diffuso”. Operativamente è la frazione di campioni che, in una finestra temporale scorrevole, superano la deviazione standard locale; la normalizzazione è scelta in modo che un rumore gaussiano (campo pienamente diffuso) dia ≈ 1 . Nei primi millisecondi, quando arrivano poche riflessioni discrete (diretto, prime riflessioni speculari), la NED è bassa; man mano che la nube di riflessioni si infittisce, sale verso 1.

$$\text{NED}(t) = \frac{1}{\text{erfc}(1/\sqrt{2}) |W|} \sum_{\tau \in W(t)} \mathbf{1}[|h(\tau)| > \sigma_W(t)]. \quad (9)$$

- **Mixing-time** — l’istante in cui la NED raggiunge 1 in modo *sostenuto*: il momento in cui il campo smette di essere un insieme di echi distinti e diventa statisticamente diffuso. È il confine fisico fra il regime geometrico (early, dove conta la forma della sala) e quello diffuso (late, dove conta solo il bilancio energetico). È un grado di libertà che nessun descrittore ISO espone: C_{50} e T_{20} non sanno dire quando avviene la transizione, perché integrano l’energia attraverso di essa.

$$t_{\text{mix}} = \min\{t : \text{NED}(t) \geq 1 \text{ (sostenuto)}\}. \quad (10)$$

- **Densità modale e frequenza di Schroeder** — delimitano dove il regime diffuso inizia in frequenza. La densità modale segue la legge di Weyl [12] *completa* $dN/df = 4\pi V f^2/c^3 + \pi S f/2c^2$ (volume + superficie, non il solo volume: alle basse il secondo termine non è trascurabile, §2.2), e la frequenza di Schroeder f_{Sch} (dove l’overlap modale $M = 3$) è la soglia sopra la quale i modi si sovrappongono abbastanza da rendere il campo diffuso. Sono quantità $O(1)$ che contestualizzano gli altri due: dicono se una certa banda è nel regime modale o in quello diffuso.

Origine e uso consolidato. Questi indicatori hanno una storia e un uso consolidato. La *densità di echi normalizzata* è introdotta da Abel e Huang (2006) [4] per misurare quando una risposta impulsiva diventa diffusa, con una statistica robusta e poco sensibile al rumore. Il *mixing-time* è studiato da tempo come confine early/late: il modello statistico del riverbero di Polack (1993) [5] lo pone come l’istante oltre il quale la coda è ben descritta da un processo gaussiano; Defrance e colleghi (2009) [6] ne danno stime direttamente dal segnale (matching pursuit); nell’auralizzazione (Jot 1999) [7] è il parametro operativo che decide dove tagliare la risposta — fino al mixing-time si rende esplicita la parte geometrica (image-source), oltre si commuta a un riverbero statistico/sintetico, risparmiando calcolo senza perdita percettiva. La *densità modale* e la *frequenza di Schroeder* sono classici (Schroeder 1962; Kuttruff 2016) [8, 26] e inquadrano il regime. L’uso tipico è dunque triplice — auralizzazione (taglio early/late), analisi del riverbero, de-riverberazione — e in tutti il mixing-time è il grado di libertà *temporale* che l’energia integrata non porta con sé.

Uso in RaySound. Non come una misura isolata, ma come un *canale ortogonale* ai descrittori ISO. Energia (ISO) e statistica temporale (invarianti) descrivono aspetti in parte indipendenti

dello stesso campo, quindi un secondo canale aggiunge informazione che il primo non contiene per costruzione. Da qui due usi. (i) *Diagnostic*: il mixing-time misurato da una risposta reale si confronta con quello del forward; se divergono, il modello energetico sta mancando un grado di libertà, il quando della transizione early→late, e lo si vede come numero (§3.7). (ii) *Prospettico*: lo stesso canale offre una *loss ortogonale* per una calibrazione futura che vincoli non solo l'energia ma anche la struttura temporale del campo. Ciò che li rende usabili come canale è la robustezza: essendo statistiche del campo (non della singola RIR), sono pressoché invarianti alla posizione del microfono e alla fase, poiché due risposte diverse nella stessa sala condividono la stessa transizione diffusa, ed è quella che si misura.

2.6 Strumenti e dati

dg-acoustics (forward+aggiunto d'onda, CUDA) · raytracer_gpu (GA, NVIDIA Warp) · il modulo *identifiability* (SVD, cond, rango, OED via pivoted-QR [24]) e il modulo *diffuse_invariants* di *raysound.calibration*. Dati: benchmark BRAS (TU Berlin, DOI 10.14279/depositonce-6726.3) per CR1/CR2/CR3/CR4; dEchorate (30 microfoni) per gli invarianti. Mesh per conversione SketchUp→DAE; materiali $\alpha(f)$ a 1/3 d'ottava (31 bande ISO) dai *fitted estimates* BRAS, usati nativamente per terzo: ogni terzo porta il proprio α misurato (es. il *brickwall* di CR4 a 200/250/315 Hz vale 0.156/0.097/0.275 — struttura $\sim 2.8\times$ che l'aggregazione in ottava avrebbe appiattito).

3 Validazione su sale reali

Tre sale BRAS [1] coprono una progressione di scala e di buona-posizione del problema inverso: CR4 (auditorium, caso principale), CR3 (sala da musica, conferma di robustezza alla scala) e CR2 (sala seminari, identificabilità negativa).

Nota metodologica. Le sale complete, validate end-to-end (geometria reale + coefficienti per superficie → descrittori vs misura), sono quattro (BRAS CR4/CR3/CR2/CR1, l'ultima il campo accoppiato/ C_{50} , §3.5). I dataset esterni del §3.10 non sono sale-complete aggiuntive: validano assi indipendenti (propagazione, geometria/TOA, pipeline di misura, premessa del design). Due cautele sui dati BRAS. (i) I coefficienti BRAS sono *fitted*: la documentazione del benchmark li dichiara «*absorption coefficients matched to measured reverberation time according to Eyring equation*», su S e V del modello 3D fornito (CR3: $2763.0\text{ m}^2 / 3330.6\text{ m}^3$, sottotetto incluso). Gli $\alpha(f)$ sono dunque l'inversa di Eyring su quel V/S : riprodurre Eyring con essi è aritmetica, non predizione. Il pavimento di quella tautologia è misurabile ed è il residuo del fit BRAS stesso: su CR3, Eyring contro il T_{20} misurato dà mediana 1.14% coi residui per banda fra 0.1% e 1.3% (contro T_{30} : 1.70%; senza il termine d'aria il residuo esplode a 25.6% a 4 kHz, quindi il fit BRAS l'aria la conteneva). Il confronto Eyring che segue valida quindi la coerenza geometria-coefficienti-descriptori della pipeline (una verifica senza ricalibrazione RaySound), non una stima cieca di α fisici indipendenti. La predizione *indipendente* è il ramo geometrico: sui terzi 250–2000 Hz e 5 ricevitori, il T_{30} ray-tracciato devia dal misurato del 17.1% (CR4) e dell'11.5% (CR3), e non coincide con Eyring per costruzione. La ragione è misurabile. Eyring riassume la geometria in un solo numero, il cammino libero medio $4V/S$; un ray-tracer risolve invece l'intera *distribuzione* dei cammini. Misurando i cammini liberi sulle mesh reali, la media coincide con $4V/S$ entro l'1% (−0.1% su CR3, −0.9% su CR4 — è l'oracolo che valida il campionatore), ma la larghezza relativa della distribuzione vale $\sigma/\mu = 0.88$ (CR3) e 1.14 (CR4) contro 0.64 di una scatola delle stesse proporzioni. Poiché il decadimento è esponenziale nel numero di riflessioni e l'esponenziale è convessa, una distribuzione più larga decade più lentamente della sua media: ricostruendo la curva di decadimento dai soli cammini misurati, da $(1 - \alpha)^n$ e dall'assorbimento dell'aria, *senza alcuna formula di riverbero*, si riproduce il T_{30} ray-tracciato entro il 3.8% (CR4) e l'1.3% (CR3) — scarto mediano sulle tre bande — e si spiega per intero lo scarto da Eyring. La conseguenza è metodologica: i coefficienti *fitted* del benchmark sono tarati *sull'equazione di Eyring*, e usarli

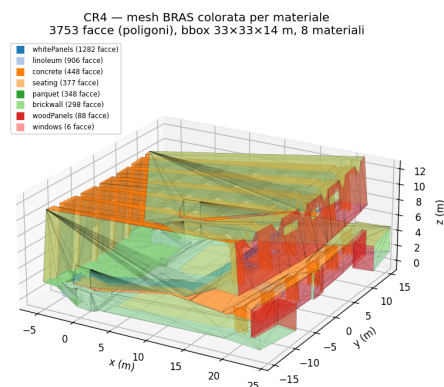
in un ray-tracer sovrastima necessariamente il tempo di riverbero della quantità di cui quella formula sotto-modella la dispersione dei cammini. Il confronto onesto fra i due strumenti richiede coefficienti tarati sullo strumento che li usa.

Validazione fuori campione. Applicando al ray-tracer la stessa procedura che il benchmark ha applicato a Eyring — *un solo* fattore per banda su tutti i coefficienti, scelto perché il T_{30} predetto coincida col misurato — il tempo di riverbero non è più un test, perché è stato usato per la taratura; lo diventano i descrittori che *non* vi entrano. Su 5 posizioni-ricevitore e tre bande d’ottava — con la taratura risolta per bisezione, che riporta il T_{30} entro lo 0.25% (CR4) e lo 0.09% (CR3) — il ray-tracer predice allora il C_{50} entro 1.08 dB (CR4) e 1.34 dB (CR3) e l’EDT entro il 6.4% e il 7.6%, senza che nessuno dei due sia stato usato nella taratura. I fattori richiesti sono coerenti fra le due sale (1.17/1.26/1.19 contro 1.09/1.18/1.11 a 500/1000/2000 Hz) e dell’ordine di grandezza atteso dalla dispersione dei cammini (generatore `experiments/ga_out_of_sample_calibration.py`). È questo, e non la deviazione a coefficienti non-tarati, il numero che misura il modello anziché la compatibilità fra due convenzioni di taratura. Il residuo non è uniforme in frequenza: si concentra a 1 kHz, dove il C_{50} resta sovrastimato di +2.17 dB (CR4) e +2.13 dB (CR3) — un residuo sistematico, non dispersione, e un fronte aperto. (ii) Le CR sono case study realistici, ma il benchmark BRAS segnala incertezze di misura maggiori sulle *complex rooms*: le CR sono trattate come confronto controllato, non come benchmark assoluto.

3.1 CR4 — Auditorium (caso principale)



Sala reale (BRAS, CR4_Overview1).



Sala CR4 — superfici per-materiale, involucro trasparente (render RaySound).

Figura 1: CR4: la sala reale e il modello geometrico. La balconata, il palco e il soffitto inclinato rendono la geometria tutt’altro che una shoebox.

Geometria. Volume 8657 m^3 , area 5851 m^2 , 5162 facce, 8 materiali. Superfici dominanti: cemento 30%, pannelli bianchi 28%, mattoni 12%, linoleum 11%, sedute 10%. Bounding box $\sim 33 \times 33 \times 14 \text{ m}$.

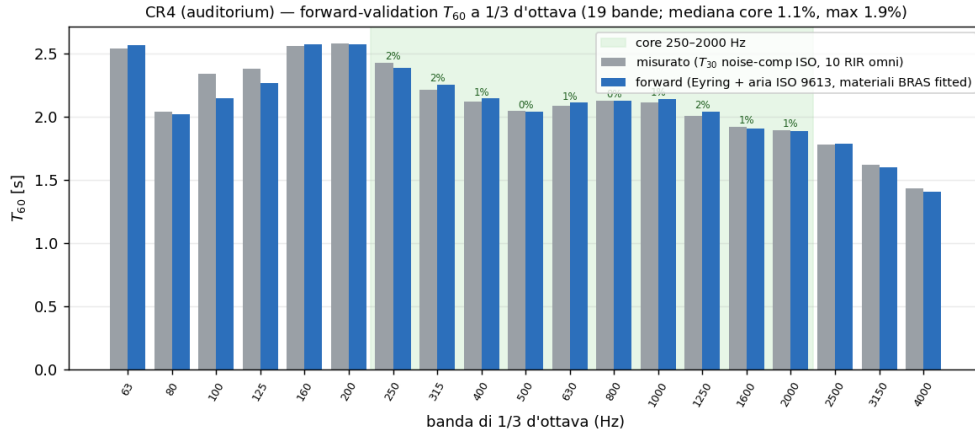


Figura 2: CR4 — Forward-validation del T_{60} senza ricalibrazione: predetto (Eyring + aria ISO 9613, coefficienti BRAS fitted) vs misurato (Schroeder T_{30} , 10 RIR omnidirezionali).

Come leggerla. Ogni coppia di barre è una banda di 1/3 d'ottava (19 bande 63–4000 Hz, materiali ad α nativi per terzo). Nella banda-core 250–2000 Hz (area verde) la deviazione (mediana di $|T_{60}^{\text{pred}} - T_{60}^{\text{meas}}|/T_{60}^{\text{meas}}$ sui terzi) è dell'1.1% (max 1.9%), senza che il motore abbia visto i dati. La risoluzione fine risolve struttura che l'ottava mediava: un piccolo picco di T_{60} a 160–200 Hz (pred 2.58 / mis 2.57 s), tracciato dai materiali nativi. L'accordo regge anche alle basse frequenze (63 Hz: 1.2%; il residuo maggiore resta a 100 Hz, 8.3%) perché il T_{30} misurato è estratto con compensazione del rumore (ISO 3382 [27, 22], stimatore del motore `t30_noise_compensated`, SNR-gated): lo Schroeder naïf a banda stretta integrerebbe nel floor e gonfierebbe il T_{30} . Le bande pulite (≥ 250 Hz) restano il T_{30} standard, per cui il core è invariato. Il residuo alle basse è l'incertezza dell' α (fit BRAS meno vincolati).

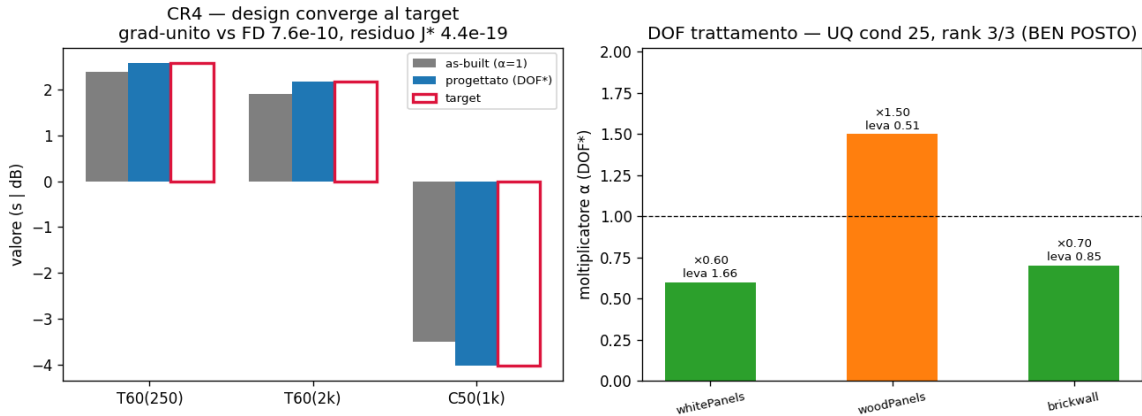


Figura 3: CR4 — Design inverso position-aware. Sinistra: i moltiplicatori di trattamento, Eyring (total-only) vs hit-weighted (position-aware); in rosso le differenze $\geq 10\%$. Destra: il peso per area (Eyring) vs la frequenza-di-colpo h (ray-tracer) per ogni superficie.

Come leggerla. Entrambe le versioni raggiungono lo stesso target (residuo $\sim 10^{-18}$, gradiente vs differenze-finite 7.6×10^{-10} , UQ a rango pieno $\kappa = 25$, 3/3), ma scelgono progetti diversi dove la geometria conta: il brickwall cambia del -12% (DOF 0.70 $\rightarrow$ 0.61). A destra il motivo: dove la frequenza-di-colpo h coincide con l'area (whitePanels, brickwall a CR4) la differenza è piccola; in generale è la deviazione $h \neq \text{area}$, prodotta dalla geometria reale, a ri-pesare il progetto. Il loop L-BFGS-B converge in ~ 3 ms (18 iterazioni), interattivo; h_m si calcola una volta dal ray-tracer.

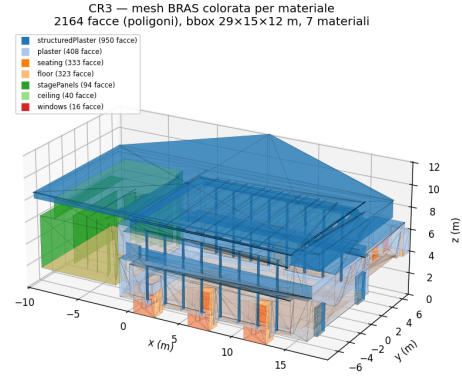
Esito. Pipeline coerente senza ricalibrazione (T_{60} Eyring 1.1%, 1/3 d'ottava; ray-tracer 17.1% sulle stesse bande); design position-aware convergente con gradiente a precisione macchina e libertà di progetto a rango pieno. Il design ottimizza quale superficie trattare: rispetto a Eyring sposta

il trattamento fino al 12% (brickwall), dove la geometria reale fa sì che la frequenza-di-colpo si discosti dall'area.

3.2 CR3 — Sala da musica (robustezza alla scala)



Sala reale (BRAS, CR3_Overview1).



Sala CR3 — superfici per-materiale, involucro trasparente (render RaySound).

Figura 4: CR3: sala da musica con palco e balconata. Volume 3331 m^3 *comprensivo del volume sopra il soffitto* (la sala da sola è 2372 m^3 , vedi §3.3), area 2763 m^2 , 2598 facce, 7 materiali; bounding box $\sim 29 \times 15 \times 12 \text{ m}$.

3.3 Nota su CR3: il volume sopra il soffitto è ACCOPPIATO, e il benchmark lo dice

La documentazione BRAS dichiara per CR3 un volume di 2350 m^3 «*excluding the attic volume*», mentre il modello 3D del benchmark — e quindi la nostra mesh — ne racchiude 3331 m^3 . La differenza non è un errore di modellazione: è documentata nel benchmark stesso, che nel commento alla scena scrive che il volume sopra il soffitto «*is included in the 3D model and might be excluded or simplified for acoustic simulation. The volume is coupled by large connections at the stage, and small connections on the ceiling*».



Figura 5: CR3, dettaglio del perimetro del soffitto (foto del dataset BRAS, Pictures/Details/5_ceiling_detail_1.jpg). Il soffitto è chiuso e decorato, ma sul perimetro corre una **griglia di aerazione a lamelle**: attraverso l'apertura si vedono le travi strutturali, un faretto e una bocchetta di ventilazione *dentro* il volume sopra il soffitto. È la «small connection on the ceiling» del commento BRAS; la connessione dominante è al palco.

Verifica indipendente sulla nostra geometria. Un test di parità (raggio verticale, conteggio delle facce attraversate — separa i due volumi, dove il teorema della divergenza li somma) dà sala 2372 m^3 e volume superiore 957 m^3 : la cifra dichiarata da BRAS è confermata allo 0.9%. Un test di connettività conferma anche *dove* sta l'accoppiamento: dei segmenti sparati verso il volume superiore ne passa il 2.83% dal **palco** contro lo 0.83% dalla platea, cioè $3.4\times$, coerente con «large connections at the stage».

Effetto sul forward. Escludere il volume superiore *peggiora* l'accordo col T_{60} misurato: la deviazione mediana in banda-core passa da 1.3% a 4.9% e la peggiore da 7.5% a 13.0%. Il modello a volume unico che include il volume superiore è quindi la scelta corretta *in media*, ma non è esatta banda per banda: il residuo di CR3 a 160 Hz (7.5%, pari a $2.1\times$ la dispersione inter-ricevitore della misura) è compatibile con un'apertura *parzialmente* trasparente. Sostituendo il volume unico con il modello a due volumi accoppiati del motore (**coupled_decay**), con **un solo** parametro libero — l'area di accoppiamento S_c — per tutte le 19 bande, si ottiene $S_c \simeq 28\text{ m}^2$, deviazione mediana in banda-core 1.24% e il residuo a 160 Hz che scende da 7.5% a 3.1%. Il fronte resta aperto (§3.11 non ne dipende): il modello accoppiato peggiora invece le bande $\geq 3150\text{ Hz}$, dove la ripartizione dell'assorbimento d'aria fra i due volumi va derivata e non assunta.

Caratterizzazione del volume sopra il soffitto. Il benchmark fornisce per CR3 sette materiali di superficie, e nessuno di essi descrive il vano sopra il soffitto: la documentazione dichiara esplicitamente che del complesso, costruito come *room in room*, è stato modellato il solo guscio interno. Nel modello 3D le 258 facce del vano ereditano perciò il materiale delle pareti strutturate — $\alpha \simeq 0.13\text{--}0.15$ — applicato a 814 m^2 , cioè al 29% della superficie totale della scena. È una porzione troppo grande perché la scelta sia indifferente, e infatti non lo è: variando *soltanto* quelle facce il T_{30} predetto si sposta dal +6–14% rispetto alla misura (vano riflettente come le pareti) al –8–+2% (vano completamente assorbente), il che colloca il comportamento reale fra i due estremi e rende il vano il singolo fattore dominante del decadimento tardivo di questa sala.

Abbiamo quindi caratterizzato quel materiale mancante come *un solo* parametro, determinato su 30 osservabili indipendenti (5 posizioni-ricevitore $\times$ 6 bande d'ottava, forward geometrico per-posizione): $\alpha_{\text{vano}} = 0.25$ porta la deviazione mediana del T_{30} dal 8.98% al 2.61% e, soprattutto,

rimuove il *segno* del residuo — con il valore ereditato tutte e sei le bande sono in eccesso (+7–+16%), con il valore caratterizzato si distribuiscono attorno allo zero. Il valore è coerente con la natura fisica dell’ambiente, un sottotetto in travatura con impianti e superfici irregolari, per il quale un intonaco liscio (0.05–0.15) è una descrizione inadeguata. Resta un residuo a 1 kHz (+7.9%), la banda peggiore anche prima della caratterizzazione.

Questo parametro va letto per quello che è: un ingresso di modello *determinato* sulla scena, non un coefficiente misurato dal benchmark. Il T_{60} di CR3 ottenuto con esso non è quindi una predizione cieca, ed è riportato qui come caratterizzazione del vano, non come validazione del forward; le colonne di validazione senza ricalibrazione restano quelle del §3.

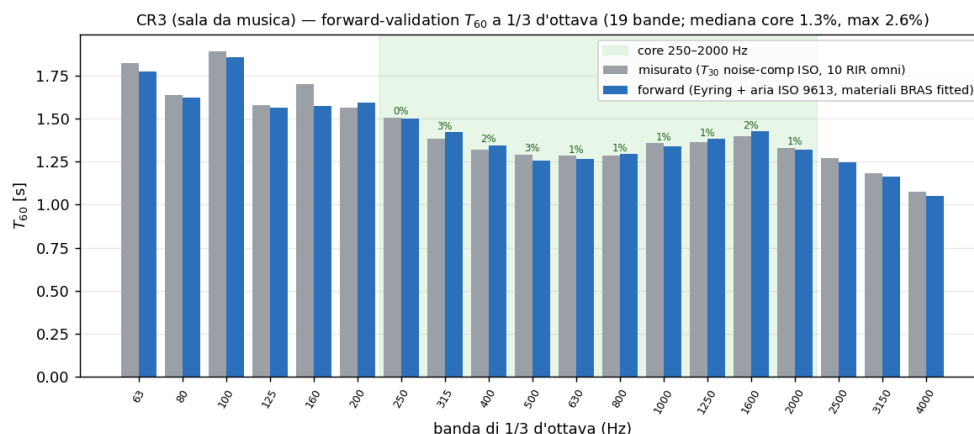


Figura 6: CR3 — Forward-validation del T_{60} a 1/3 d’ottava: banda-core 250–2000 Hz, deviazione mediana 1.3% (max 2.6%) senza ricalibrazione.

Come leggerla. Stessa lettura di CR4: barre per banda di 1/3 d’ottava (materiali ad α nativi), area verde = banda-core. La pipeline regge su una sala di scala diversa (29 m contro 33 m) senza ritaratura. Anche qui la deviazione a 63 Hz è dominata dal rumore della misura (T_{30} CV 12% tra le posizioni, contro 3% a 1 kHz), amplificata dall’incertezza dell’ α alle basse; non dal regime diffuso (overlap modale $M = 8.8, \gg 3$: Eyring corretto).

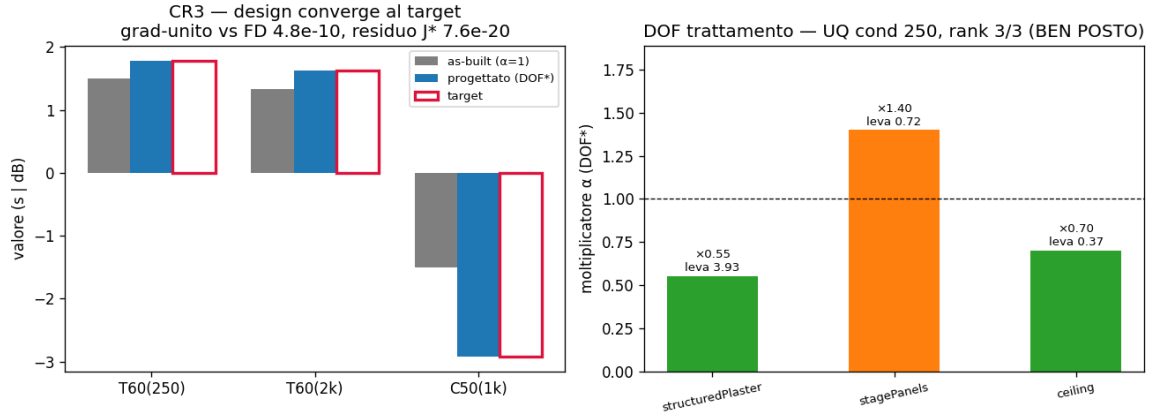


Figura 7: CR3 — Design inverso position-aware: Eyring vs hit-weighted (sinistra), e il peso per area vs frequenza-di-colpo h (destra). Gradiente vs FD 4.8×10^{-10} ; UQ $\kappa = 230-250$, rango pieno.

Come leggerla. Entrambe le versioni convergono (residuo $\sim 10^{-20}$, UQ rango pieno; κ resta ~ 250 perché l'intonaco strutturato, 42% dell'area, domina la sensibilità, una proprietà fisica della sala). Il design position-aware sceglie diversamente dove la geometria ri-pesa: i pannelli-palco cambiano del -12% e il soffitto del -11% (DOF $0.70 \rightarrow 0.62$). A destra il motivo: il soffitto è colpito un po' più della sua area (h 7% vs area 6%), l'intonaco strutturato un po' meno (39% vs 42%); la geometria reale (palco, balconata) sposta i pesi rispetto all'area.

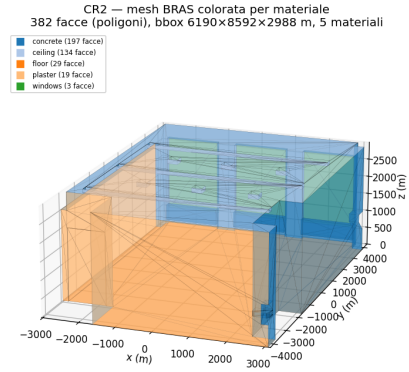
Esito. La stessa pipeline regge da CR4 a CR3: forward 2%, design position-aware convergente, UQ a rango pieno. Qui l'effetto position-aware è più marcato (fino al 12%, pannelli-palco) perché la geometria si discosta di più dalla scatola. CR3 resta più mal-condizionata (una parete domina la sensibilità, fisica e non artefatto); T_{60} e C_{50} sono in tensione (alzare il riverbero abbassa la chiarezza): il target è scelto coerente.

3.4 CR2 — Sala seminari (sala piccola, identificabilità negativa)

CR2 ha un doppio ruolo, ed entrambi sono concettuali. Primo, è l'estremo a sala piccola della mappa dei regimi (§2.2): con $V \approx 159 \text{ m}^3$ la frequenza di Schroeder sale a $\sim 80-120 \text{ Hz}$, una banda modale larga dove solo il forward d'onda è corretto. È la sala dove il forward d'onda è essenziale, opposta agli auditorium CR3/CR4 (diffusi già a 63 Hz). Secondo, è il caso di identificabilità negativa: lo stesso strumento che su CR3/CR4 dichiara “qui hai 3 gradi di libertà di progetto” dichiara qui “i materiali non sono recuperabili dai dati”, e lo segnala con un numero (cond, rango).



Sala reale (BRAS, CR2_Overview1).



Sala CR2 — superfici per-materiale, involucro trasparente (render RaySound).

Figura 8: CR2: shoebox $6.19 \times 8.59 \times 2.99$ m, volume $\sim 159 \text{ m}^3$, 6 superfici. A differenza di CR3/CR4, tutti i materiali sono rigidi ($\alpha < 0.20$ in banda ≤ 250 Hz).

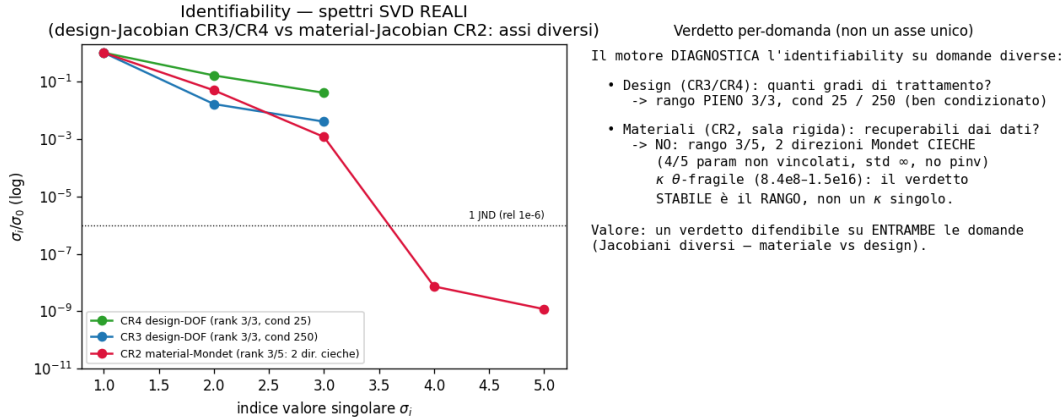


Figura 9: Identificabilità per sala: il motore diagnostica su *domande diverse* (Jacobiani diversi). Sinistra: spettri SVD reali — design-DOF (CR3/CR4) vs materiale-Mondet (CR2). Destra: il verdetto per-domanda.

Come leggerla. A sinistra, ogni curva è una sala su una domanda diversa: CR4 e CR3 sono il Jacobiano di *progetto* (3 moltiplicatori di trattamento) e mantengono i valori singolari sopra la soglia del rango ($\sigma_i/\sigma_{\max} > 10^{-6}$, rango pieno 3/3); CR2 è il Jacobiano *materiale* (5 Mondet) e collassa (rango 3/5, 2 direzioni cieche). La soglia di 1 JND è quella percettiva, applicata alla rilevanza acustica di una direzione (Jv). A destra il verdetto per-domanda: CR4/CR3 hanno rango di progetto pieno (cond 25 e 250); CR2 ha il materiale rank-deficient (3/5). Sono assi diversi (progetto vs materiale) letti dalla stessa SVD del Jacobiano. Il κ di CR2 è θ -fragile (da 8×10^8 a 2×10^{16}): il verdetto stabile è il *rango*, non un κ singolo. La diagnostica ha costo trascurabile: la SVD del Jacobiano è $2.6 \mu\text{s}$, l'OED D-ottimale $11 \mu\text{s}$, l'UQ di Laplace $2.5 \mu\text{s}$.

Esito. L'identificabilità è negativa: il Jacobiano materiale-Mondet (5 parametri) è rank-deficient (rango 3/5, 2 direzioni cieche; 4/5 parametri non vincolati, std ∞ , nessuna maschera pinv). La calibrazione dei materiali è non-identificabile per assenza di segnale: con $\alpha < 0.20$ il T_{30} in una shoebox rigida è governato dal rapporto volume/superficie, non dai parametri di superficie. Il κ è θ -fragile (da 8×10^8 a 2×10^{16} al variare del materiale near-rigid): non un singolo numero, ma il rango è robusto, e la rank-deficiency è fisica, confermata per tre vie indipendenti (ogni θ near-rigid dà rango < 5 ; il fit $\alpha \rightarrow \theta$ è degenerare; la riflessione quasi-piatta acceca i parametri). Il

C_{50} a 125 Hz di CR2 dal forward per-posizione è ~ -5 dB, dentro lo spread per-posizione della misura (mediana ~ -1.4 dB su 5 ricevitori, §3.6). Il residuo minore che resta è di *regime*, non di geometria: a 125 Hz CR2 è *sotto* la frequenza di Schroeder, regime modale dove il forward energy-based (GA) è cieco ai modi discreti, e la geometria idealizzata aggiunge il resto. Il motore segnala (via SVD del Jacobiano) quando non può recuperare i materiali.

3.5 CR1 — Stanze accoppiate (forward del campo accoppiato)

CR1 è l'unica sala BRAS a *stanze accoppiate*: due volumi — una camera riverberante e un laboratorio — connessi da una porta (apertura $S_c \approx 3 \text{ m}^2$), con la sorgente in un volume e i ricevitori nell'altro, e il *diretto occluso* (i due ambienti comunicano solo attraverso la porta). È anche l'unica scena a *double-slope* del benchmark: la camera poco-assorbente ri-alimenta lentamente il laboratorio attraverso la porta, così la curva di decadimento (EDC) ha *due* pendenze — una *early* locale e un *sustain* inter-volume più lento. Nel round-robin BRAS nessuno dei sei algoritmi geometrici riproduce il double-slope: è il banco di prova del forward multi-slope (cfr. §3.7, Multi-Room): CR1 lo chiude sul forward.

Geometria. Due volumi dalla segmentazione automatica (door-loop sul DAE fuso): $V_A \approx 178 \text{ m}^3$ ($\alpha \approx 0.096$, ricevitori) e $V_B \approx 127 \text{ m}^3$ ($\alpha \approx 0.016$, parete dura, sorgente), connessi dalla porta $S_c \approx 3.06 \text{ m}^2$; 299 facce, 5 materiali. La sorgente è nel volume meno assorbente; i ricevitori nel volume più assorbente. Il singolo volume aggregato ($\sim 305 \text{ m}^3$) nasconderebbe la struttura a due-volumi.

La difficoltà dell'early. Il *sustain* (la coda tardiva, inter-volume) è già modellabile col bilancio energetico 2×2 di Cremer/Kuttruff [21, 26] (§3.7). Il campo *early* del volume-ricevitore no: il ricevitore è off-axis rispetto alla porta e il diretto è occluso, quindi l'early non è radiazione diretta ma il *build-up multi-bounce* del campo che entra dalla porta e si ridistribuisce per riflessioni nel laboratorio.

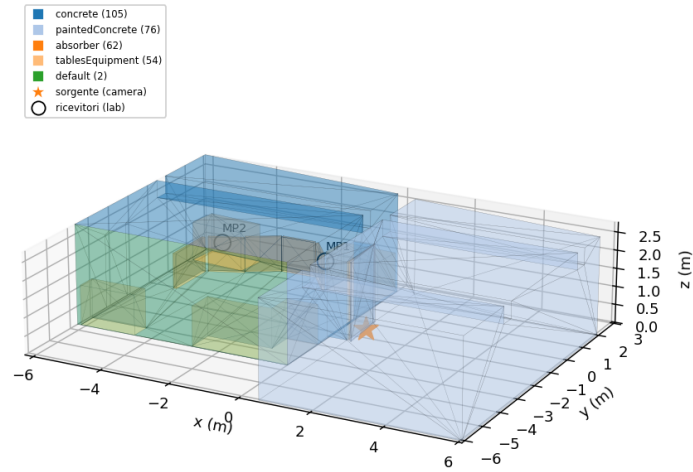
Il modello (two-part, fit-free). Il motore compone la risposta del volume-ricevitore in due parti, ancorate alla finestra del C_{50} (eq. (6)):

- **early** ($t < 50$ ms) = *ray-diffuse build-up*: il ray-tracer diffuso sulla mesh accoppiata (4×10^5 raggi, occlusione, α reali) porta il multi-bounce che il 2×2 analitico (decadimento-da- t_0) *non* ha;
- **late** ($t \geq 50$ ms) = *sustain* 2×2 (l'autovalore lento del coupling), *energy-normalized* al totale-tardivo del ray: preserva il C_{50} (che dipende dal totale-tardivo) e dà la pendenza profonda corretta;
- lo *splice* è *oltre* la finestra C_{50} (> 50 ms): non tocca il build-up early (uno splice dentro la finestra lo sopprimerebbe);
- **detection geometrica automatica**: il motore rileva il *door-loop* (l'apertura tra due volumi) dalla mesh — nessun flag, nessuno script dedicato. La stessa chiamata unica non-box `simulate_raytracer_mesh` (§3.6) emette i descrittori accoppiati;
- **fit-free**: $s_{\text{rec}} = 1/4\pi$, 2×2 geometrico, zero pesi tarati.



CR1 (BRAS): le due stanze accoppiate — camera riverberante e laboratorio (foto reale panoramica; credito TU Berlin).

CR1 — stanze accoppiate (camera + laboratorio, porta), mesh BRAS per materiale (299 facce)



CR1 — camera + laboratorio + porta, sorgente e ricevitori a coordinate reali (render RaySound, mesh = forward).

Figura 10: CR1: i due volumi accoppiati e la porta. La sorgente è nella camera riverberante, i ricevitori (MP1/MP2) nel laboratorio; il diretto è occluso — i due ambienti comunicano solo attraverso l'apertura ($S_c \approx 3 \text{ m}^2$).

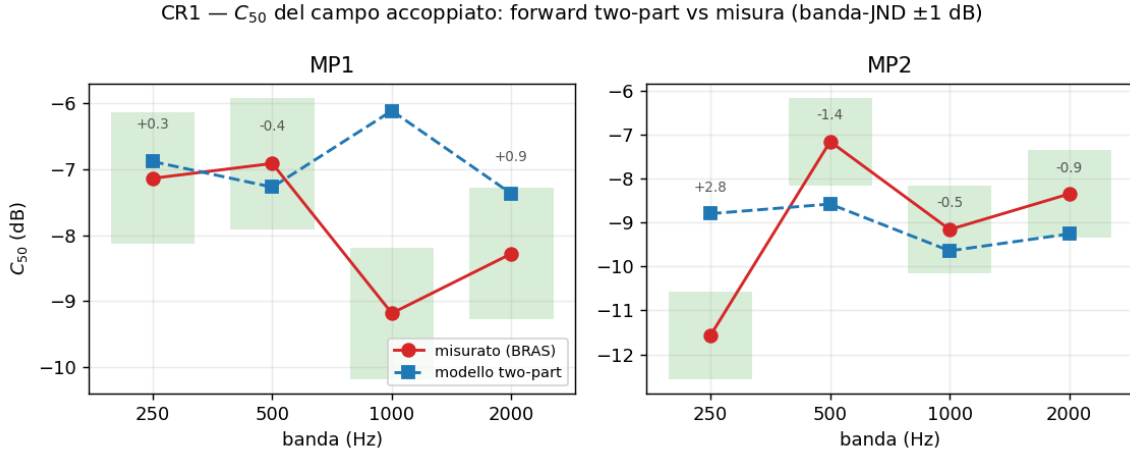


Figura 11: CR1 — C_{50} del campo accoppiato: forward two-part vs misura, per banda (250–2000 Hz) e per ricevitore, banda-JND ombreggiata (± 1 dB). Calcolato dal forward accoppiato auto-routed (la stessa pipeline del motore).

Come leggerla. Senza l’early (il solo 2×2) il C_{50} del volume-ricevitore sarebbe troppo riverberante, ~ -13 dB; il two-part lo porta a ~ -8 dB, dentro la misura (-7 a -9 dB). Il numero su ogni punto è il residuo modello–misura. La *variazione* tra ricevitori è riprodotta (MP2, più lontano dalla porta, è più riverberante di MP1 di ~ 4 dB — non un C_{50} piatto): il modello segue la dipendenza dalla posizione. MP1/MP2 sono le due posizioni-ricevitore del laboratorio fornite dal dato BRAS per CR1 (le altre CR usano fino a 5 ricevitori; CR1 ne ha 2 — è il dato). Residuo: 6/8 coppie banda-ricevitore entro ~ 1.5 dB, 2/8 a 2.8–3.1 dB (le bande modali, floor dichiarato nell’esito).

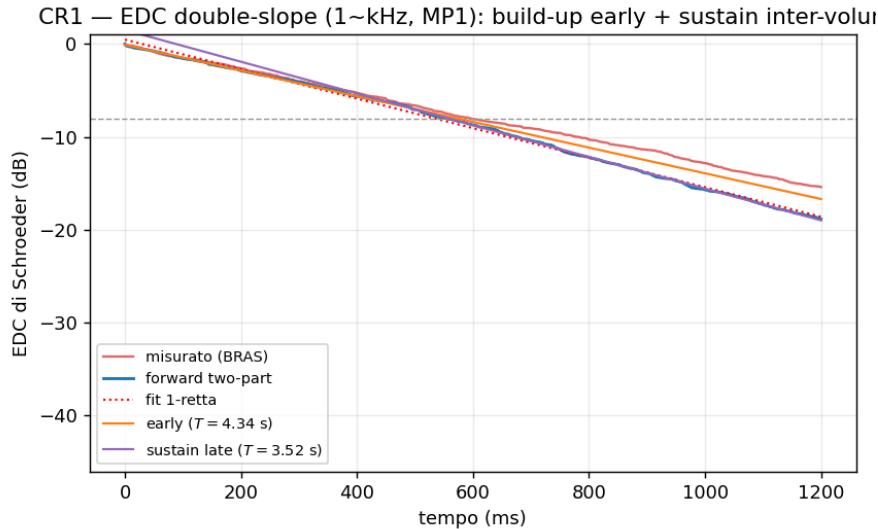


Figura 12: CR1 — EDC di Schroeder [9] a una posizione (1 kHz, MP1): la firma *double-slope* del coupling.

Come leggerla. Il fit a una sola retta (punteggiato rosso) non segue il decadimento: servono *due* pendenze — una early (il laboratorio locale) e una late più lenta (il sustain inter-volume, l’autovalore lento del 2×2). Il two-part cattura *entrambe* le pendenze dove il fit a una retta fallisce: il build-up ray-diffuso dà l’early, il 2×2 energy-normalized la coda — la curva del forward (blu) segue la misura (rosso) attraverso il ginocchio. È, concretamente, il *forward minimo* che cattura il multi-slope sul dato accoppiato.

Esito. Forward-validation VERA SU BRAS del campo *accoppiato*: il C_{50} del volume-ricevitore passa da ~ -13 dB (il solo 2×2 , senza l’early) a ~ -8 dB (two-part), in media ~ 1.5 dB dalla

misura (-7 a -9 dB) su due ricevitori e quattro bande, *fit-free*. Il sustain 2×2 (la seconda pendenza dell'EDC, più lenta dell'early) regge; l'early (ray build-up) chiude il gap che il 2×2 -solo lasciava. *Residuo caratterizzato*: 6/8 coppie banda-ricevitore entro ~ 2 dB (floor), 2/8 a 2.8–3.1 dB (MP1 1 kHz, MP2 250 Hz) — il floor GA-vs-FDTD su *feature modali banda-specifiche* (il diffuso GA è liscio tra bande, la misura ha bumps modali che un diffuso liscio non risolve; stabile al variare del numero di raggi, quindi non rumore; si riduce su materiali assorbenti), *non* un bug. È lo stesso registro dei residui di livello del forward per-posizione (-0.17 CR4, $+0.70$ CR3, §3.6). La *detection geometrica* (il door-loop dalla mesh) fa sì che il motore risolva CR1 end-to-end *senza* script dedicato, come qualunque sala accoppiata.

3.6 Forward per-posizione: mappe spaziali di C_{50} /EDT su griglia di ricevitori

Le validazioni precedenti confrontano descrittori globali (un T_{60} per banda). Qui spingiamo il forward al regime *per-posizione*: data la sala (geometria + materiali noti), predire la mappa spaziale di C_{50} ed EDT su una griglia di ricevitori, e confrontarla con la misura punto-per-punto. È il test più severo del forward geometrico, perché espone la struttura spaziale che la media globale nasconde.

Catena del forward. Questo C_{50} per-posizione esce dall'*ingresso unico* non-box `simulate_raytracer_mesh` (carica la sala $\rightarrow$ regime + FDTD + GA $\rightarrow$ descrittori, una chiamata): nel regime early/per-posizione l'ingresso orchestra il forward speculare-mesh tuned (auto-backend per ordine, scattering δ per-materiale BRAS) + la coda diffusa. Per ogni ricevitore la risposta è la somma *nativa* di una parte speculare (image-source su mesh, ampiezza $\propto 1/(4\pi d)$) e di una coda diffusa (ray-tracing energetico su mesh con occlusione). La coda usa una sezione-d'urto-ricevitore $s_{\text{rec}} = 1/(4\pi)$ derivata (non calibrata) imponendo $E_{\text{rev}} = E_{\text{dir}}$ alla distanza critica $r_c = \sqrt{S\bar{\alpha}/16\pi}$: è il fattore di angolo-solido coerente col ricevitore puntuale dello speculare. La somma nativa porta l'energia riverberante piena (il T_{60} degenera a Eyring). La mesh è sanata per saldatura dei vertici coincidenti (watertight, area S preservata), che chiude le perdite di raggi dalle cricche (open-edge $\rightarrow 0$).

F2 — CR4 forward per-posizione (5 ricevitori, BRAS materiali noti)

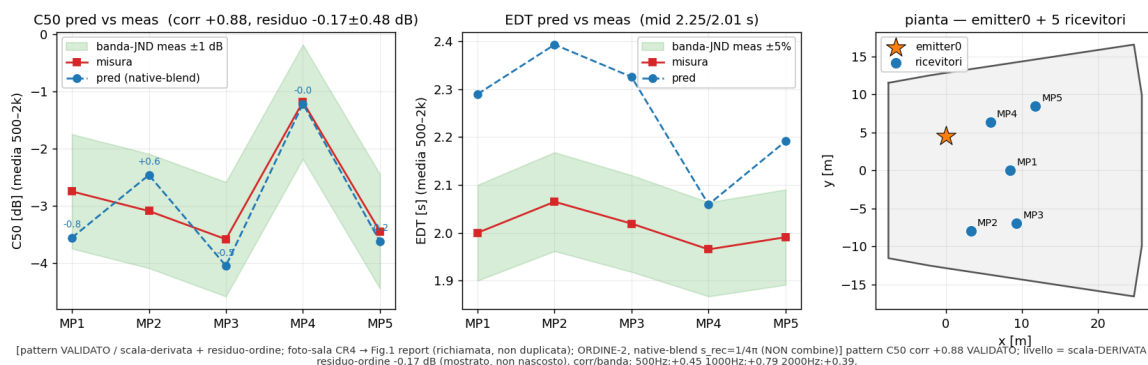


Figura 13: CR4 (auditorium, ordine speculare 1) — C_{50} ed EDT predetti vs misurati su 5 ricevitori (banda-core 500–2000 Hz), banda-JND ombreggiata (± 1 dB su C_{50} , $\pm 5\%$ su EDT), pianta con emitter0 e ricevitori a coordinate reali.

Come leggerla. Il pattern spaziale del C_{50} ha correlazione pred-mis $+0.89$ tra ricevitori (su 5 posizioni: IC 95% $[+0.1, +0.99]$); il livello presenta un residuo di -0.17 ± 0.48 dB, attribuibile alla scala derivata $s_{\text{rec}} = 1/4\pi$ più un residuo che si concentra alla banda 500 Hz, la più bassa del ramo geometrico. Il divario è modesto e localizzato, non un salto di scala.

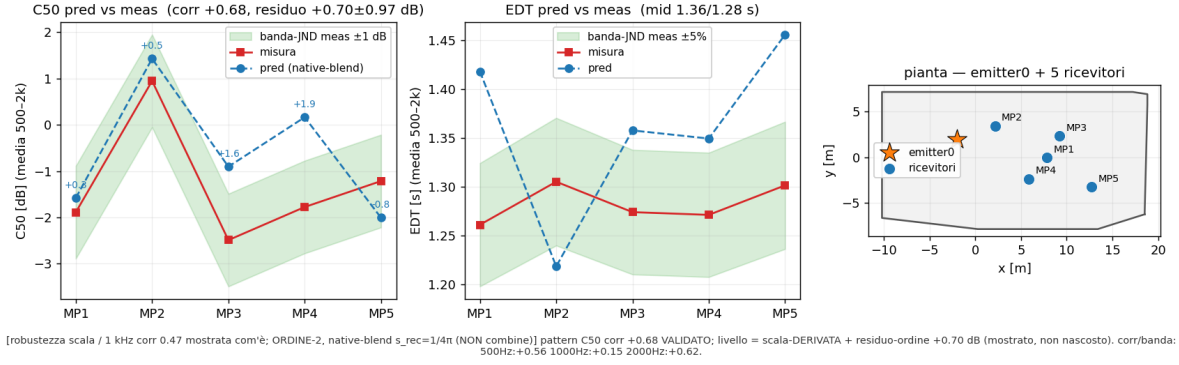


Figura 14: CR3 (sala da musica, ordine speculare 2) — come sopra. La banda 1 kHz (correlazione 0.56) è mostrata com'è, non mediata via.

Come leggerla. Robustezza alla scala: pattern C_{50} correlato +0.68 (IC 95% ampio, $n = 5$), decadimento EDT pred/mis 1.36/1.28 s. Il residuo di livello è $+0.70 \pm 0.97$ dB — positivo: a ordine-2 l'early è contato meglio e il C_{50} sale. Lo stesso ingrediente (l'ordine speculare) sposta il residuo nei due versi.

Tabella 1: Forward per-posizione — sintesi sui tre banchi (sale a volume singolo). Pattern = correlazione spaziale pred-mis del C_{50} tra ricevitori; residuo = livello C_{50} pred-mis (media $\pm$ dev. std). Etichette: pattern VALIDATO, livello = scala-derivata + residuo-ordine IDENTIFICATO. CR1 (§3.5) è anch'essa C_{50} per-ricevitore ma su un *asse diverso* — campo *accoppiato*, forward two-part (ray build-up + sustain 2×2), non il forward a volume singolo — ed è riportata separatamente, non in questa tabella.

Sala	Ordine spec.	Pattern C_{50} (corr.)	Residuo C_{50} (dB)	EDT pred/mis (s)	T_{60} pred/mis (s)	Dato
CR4 (auditorium)	2	+0.89	-0.17 ± 0.48	2.25/2.01	2.33/2.00	reale (BRAS)
CR3 (musica)	2	+0.68	$+0.70 \pm 0.97$	1.36/1.28	1.45/1.31	reale (BRAS)
MOTUS (ancora)	1	—	-1.07	0.46/0.42	—	reale (Aalto)

Controllo — invarianza con la posizione. Il T_{60} (coda tardiva, campo diffuso) deve essere ~piatto tra posizioni: lo è, con dispersione ≤ 0.03 s tra ricevitori su tutti i banchi (Fig. 15). È il complemento del C_{50} (early, position-dipendente): il forward riproduce la giusta dipendenza dalla posizione dove c'è e la giusta invarianza dove deve esserci.

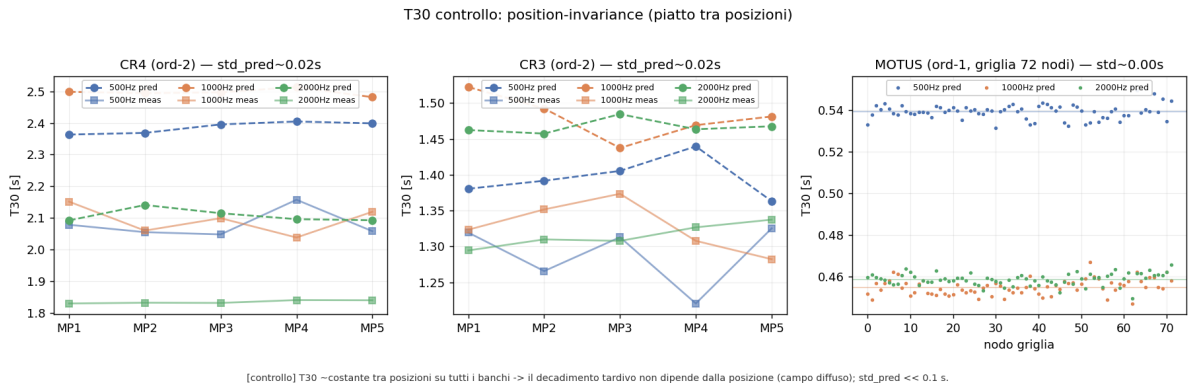


Figura 15: T_{60} vs posizione, tre banchi (CR4/CR3/MOTUS): piatto = invarianza confermata.

Come leggerla. Le linee sono piatte su tutte le posizioni; dispersione pred ~ 0.02 s (CR4/CR3), ~ 0.003 s sui 72 nodi MOTUS. Il decadimento tardivo non dipende dalla posizione — come atteso da un campo diffuso.

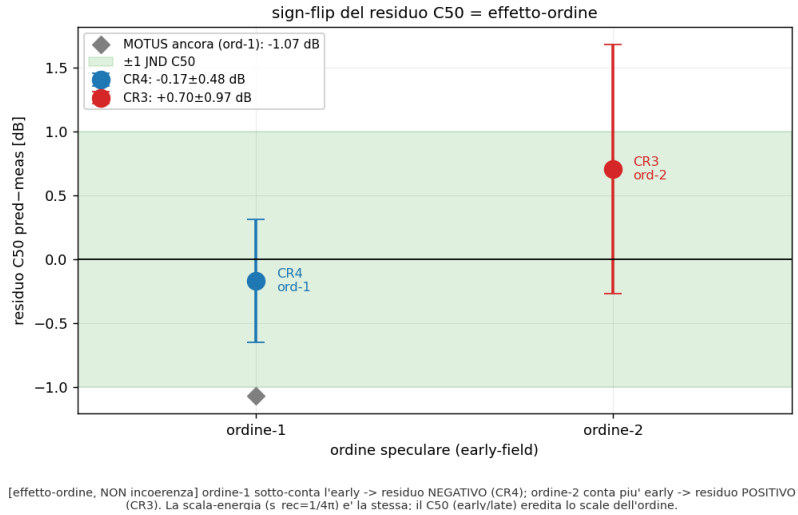


Figura 16: Il residuo di livello C_{50} risponde all'ordine speculare: più ordine, più energia early, C_{50} più alto.

Come leggerla. Conferma diretta dell'interpretazione “residuo-ordine”: il residuo risponde all'ordine speculare (più ordine $\rightarrow$ più early $\rightarrow C_{50}$ più alto), quindi è ordine, non scala né rumore Monte-Carlo (è costante al variare del numero di raggi, 160k–4.8M). La scala dell'energia totale ($s_{rec} = 1/4\pi$) è la stessa per i due banchi; il C_{50} (early/late) eredita una scala separata, governata dall'ordine.

Predizione su sala caratterizzata (MOTUS [30], un'ancora). Su una sala indipendente da BRAS (61 m³, materiali per-faccia dal metadata del dataset; foto reale Fig. 17), il forward predice la mappa C_{50} /EDT su griglia ($h = 1.45$ m, passo 0.5 m, 4 sorgenti) con una variazione spaziale di 6.44 dB risolta; l'unico ricevitore misurato ancora i materiali con un residuo $C_{50} \sim -1$ dB e EDT $\text{pred} \approx \text{mis} \sim 0.4$ s.



Sala MOTUS — foto 360° (Götz et al., Aalto; CC-BY-4.0) · sorgenti Genelec + cunei acustici visibili

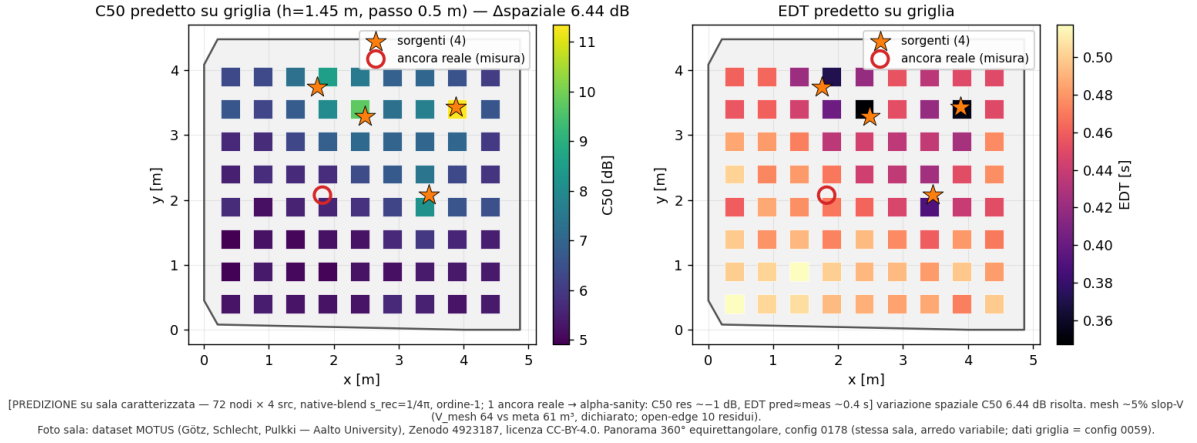


Figura 17: Mappa MOTUS: foto-sala reale (360°) sopra, heatmap C_{50} ed EDT predette su griglia sotto, con le 4 sorgenti e l'ancora reale.

Come leggerla. Il C_{50} cresce verso le sorgenti (campo early-dominato in sala piccola), la struttura spaziale è risolta; l'unico ricevitore misurato ancora i materiali, sulla mesh con $\sim 5\%$ di slop-volume dichiarato. *Foto:* dataset MOTUS (Götz, Schlecht, Pulkki — Aalto University), Zenodo 10.5281/zenodo.4923187, licenza CC-BY-4.0 — stessa prassi di credito di BRAS (TU Berlin) e RSoANU (ANU).

Ricaduta su CR2 (§3.3). Lo stesso forward per-posizione colloca il C_{50} a 125 Hz di CR2 a ~ -5 dB (Tab. 2), a ~ 4 dB dalla misura: il residuo è di regime (125 Hz è sotto Schroeder, dove il GA è cieco ai modi discreti) più la geometria idealizzata, non un artefatto del forward.

Tabella 2: CR2 C_{50} a 125 Hz: forward per-posizione vs misura (mediana su 5 ricevitori, spread -3.3 a $+1.3$ dB). A 125 Hz CR2 è sotto Schroeder (regime modale).

$C_{50}@125$ Hz	Valore (dB)	gap vs misura
forward per-posizione ($s_{\text{rec}} = 1/4\pi$)	-5.01	~ 4 dB
misurato (mediana 5 ricevitori)	~ -1.4	—

Esito (forward per-posizione). Pattern spaziale e decadimento VALIDATI SU REALE (correlazione $+0.89/+0.68$ su 5 posizioni, EDT/ T_{60} vicino alla misura, T_{60} position-invariante); livello assoluto del C_{50} = scala derivata ($s_{\text{rec}} = 1/4\pi$) + un residuo-ordine identificato (~ 1 dB, segno governato dall'ordine speculare), non within-JND assoluto su geometria complessa — dichiarato. La chiusura del residuo (ordine speculare ≥ 2 con culling di visibilità) è un arco aperto.

3.7 Invarianti di campo diffuso: ortogonalità e multi-slope su dato misurato

Su 30 microfoni reali (dEchorate [29], stessa stanza e sorgente) confrontiamo il descrittore ISO C_{50} con l'invariante mixing-time.

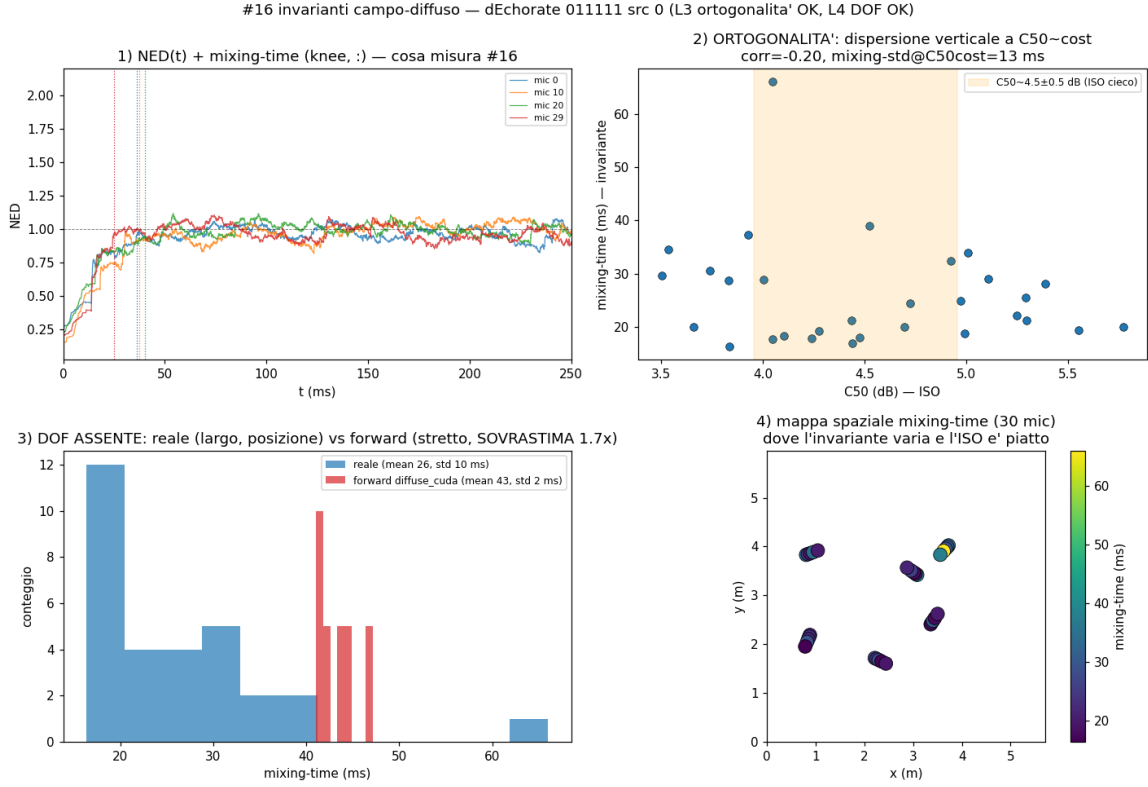


Figura 18: Invarianti di campo diffuso su dEchorate, 30 microfoni.

Come leggerla. Quattro pannelli: NED nel tempo (knee = mixing-time); scatter C_{50} vs mixing-time (a C_{50} costante il mixing-time si disperde); distribuzioni misurata vs forward energy-based; mappa spaziale sui 30 microfoni. L'interpretazione è nel testo.

Lettura guidata — i quattro pannelli. (1) *NED nel tempo.* Per ogni microfono la densità di echi sale da valori bassi (pochi echi discreti: diretto, prime riflessioni) verso 1; il ginocchio della curva è il mixing-time, l'istante in cui il campo diventa statisticamente diffuso. (2) *Scatter C_{50} vs mixing-time.* Ogni punto è un microfono. Se i due descrittori misurassero la stessa cosa i punti starebbero su una linea; invece, a C_{50} costante (una colonna verticale di punti) il mixing-time si disperde su un intervallo ampio: l'ISO non distingue microfoni che l'invariante separa: è l'ortogonalità. (3) *Distribuzioni misurata vs forward.* La distribuzione dei mixing-time reali è larga (i microfoni hanno transizioni diverse); quella del forward energy-based è stretta (un singolo valore geometrico, quasi-costante in posizione) e *spostata su valori più grandi* — sovrastima. La firma di un grado di libertà — la variazione per posizione — che il modello non possiede. (4) *Mappa spaziale.* Gli stessi mixing-time sulle posizioni reali dei microfoni: la dispersione non è rumore, ha struttura spaziale nella sala.

L3 — ortogonalità. A $C_{50} \approx 4.5 \pm 0.5$ dB (13 microfoni, sotto la JND) il mixing-time disperde con deviazione standard 13.1 ms, correlazione C_{50} -mixing -0.20 : l'invariante separa dove l'ISO è cieco, sul misurato. **L4 — DOF assente.** Il forward energy-based del motore (il diffuso di `diffuse_warp`, densità di echi $n(t) = 4\pi c^3 t^2 / V$), girato alle 30 posizioni-mic reali, dà un mixing-time *quasi-costante in posizione* (40.0 ± 2.7 ms) e *sovrastima* il reale (25.8 ± 9.8 ms). Due numeri distinti: lo *spread* per-posizione (rapporto delle deviazioni standard, reale/forward) è $3.6\times$ — il forward è troppo piatto; il *bias* medio (rapporto delle medie, forward/reale) è $1.6\times$ — il forward sovrastima. Il forward collassa il mixing-time a un numero geometrico — la variazione per-posizione è il grado di libertà che gli manca — e per giunta lo *sovrastima*, con un errore che cresce col volume. Sulle sale BRAS (lato misurato solido, 5 posizioni realistiche, non una nuvola di 30 microfoni):

	V (m ³)	forward (ms)	misurato (ms)	forward/mis.
dEchorate	80	40	25.8	$1.6\times$
CR3	3331	281	44.8	$6.3\times$
CR4	8657	518	53.6	$9.7\times$

Il forward scala con $\sqrt{V}$ (la legge di densità-echi ideale $n(t) = 4\pi c^3 t^2/V$); il misurato è quasi piatto ($2\times$ su $108\times$ di volume): le sale reali diventano diffuse *più in fretta* dell'ideale (geometria, scattering, complessità dell'early). Il mixing-time da forward energy-based non è dunque un predittore del misurato — è proprio il DOF temporale che il modello non porta, ora un numero. *Costo*: l'estrazione degli invarianti (NED Abel–Huang vettorizzata) è ~ 794 ms per RIR su CPU — un canale diagnostico leggero, $\ll$ il forward.

Conclusioni. Due conclusioni, una per asse. Sul misurato: il descrittore ISO è cieco a una differenza reale fra microfoni che l'invariante mixing-time vede (C_{50} costante, mixing-time disperso) — la statistica temporale porta informazione che l'energia integrata non contiene. Sul forward: il modello energy-based manca un grado di libertà — il mixing-time — che il campo reale possiede, e la differenza è quantificata, non congetturata. Non è un difetto del forward per gli scopi attuali (i descrittori energetici restano fedeli, §3.1–3.2): marca con un numero il confine di ciò che un modello solo-energetico può rappresentare, e indica dove il canale-invarianti aggiunge valore.

Secondo caso — il multi-slope del coupling (Multi-Room [31]). Lo stesso canale, un uso diverso. Su dEchorate gli invarianti separano posizioni che l'ISO confonde (mixing-time a C_{50} costante); su un secondo dato misurato — il *Multi-Room Transition Dataset* (Götz et al.; 3 ambienti multi-room reali, *hallways/offices/workshops*, KEMAR + Zoom H3 ambisonics, 4032 RIR) — lo stesso canale cattura un fenomeno che i descrittori a decadimento singolo non possono rappresentare: il *multi-slope* del coupling tra stanze. Quando due volumi accoppiati decadono a velocità diverse, l'EDF (*energy decay function*, l'energia residua nel tempo) non è una retta in dB ma una spezzata — una pendenza early (la stanza-sorgente) e una late più lenta (il volume accoppiato).

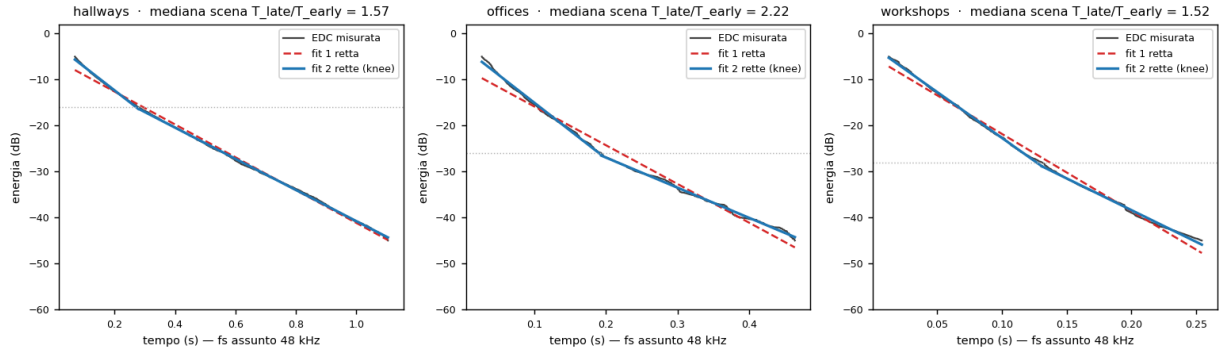


Figura 19: Multi-Room — EDF di Schroeder per una posizione rappresentativa di ciascun ambiente. Il fit a una retta (tratteggiato rosso) non segue il decadimento: serve un fit a due rette (blu, col *knee*) — la firma del multi-slope da coupling. Più marcato in *offices* ($T_{\text{late}}/T_{\text{early}} = 2.2$). Asse-tempo fs-dipendente (fs assunto 48 kHz, non nel dataset); i rapporti annotati sono invarianti di scala.

Su tutte e tre le scene il dato è multi-slope, e i numeri che lo dicono sono robusti (adimensionali, indipendenti dal sample-rate): il rapporto $T_{\text{late}}/T_{\text{early}}$ è 1.5–2.2, un fit a due rette batte la singola di $10\text{--}13\times$ (residuo), e il 73–92% delle posizioni supera la soglia di 1.3 sul rapporto.

Ambiente	$T_{\text{late}}/T_{\text{early}}$	1-retta / 2-rette	% multi-slope
hallways	1.57	10.7×	73%
offices	2.22	10.8×	92%
workshops	1.52	13.2×	88%

Tabella 3: Multi-Room — multi-slope del coupling, numeri robusti (adimensionali, fs-indipendenti): rapporto delle pendenze (mediana di scena), vantaggio del fit a 2 rette sul fit a 1 retta, frazione di posizioni con rapporto > 1.3 . I tempi assoluti (in s) dipendono dal sample-rate, qui assunto 48 kHz perché non incluso nel dataset; questi rapporti no.

Il dataset non include geometria né foto utilizzabili (solo le RIR; le foto 360° citate non sono nel file scaricato) — è un caso-invarianti (metriche di decadimento), non una sala-con-render come in §3.10, dove infatti non compare.

Che cosa si confronta. Le metriche di decadimento, non il T_{60} predetto. Il forward- T_{60} singolo (Eyring: un solo decadimento da $V/S/\alpha$) è strutturalmente fuori-modello sul multi-slope: l’assunzione campo-diffuso-singolo che cade quando i sotto-volumi accoppiati hanno tempi diversi. Ciò che il motore può dire vive nel canale-invarianti: la curvatura dell’EDF e il rapporto $T_{\text{late}}/T_{\text{early}}$ quantificano il coupling dove il T_{60} -singolo lo collassa.

Cosa aggiunge. Estende il canale. Su dEchorate la lezione era «il forward energy-based manca il mixing-time» (un grado di libertà nella transizione early→diffuso); qui è «il decadimento accoppiato è multi-slope, fuori dal forward- T_{60} singolo, e gli invarianti lo quantificano» (un limite nel *late-field*). Il forward two-part di §3.5 (CR1) cattura il multi-slope sul forward (double-slope + C_{50} accoppiato, fit-free). Resta prospettiva vincolare i parametri di coupling *dall’inverso* sul dato misurato.

3.8 Decadimento modale α_m : osservabile predetto, misurato, identificabile

Sotto la frequenza di Schroeder il campo è una somma sparsa di modi discreti e i descrittori energetici ($T_{60}/C_{50}/\text{EDT}$) sono aggregati *ciechi* all’orientazione delle pareti. Il *decadimento per-modo* α_m (la velocità con cui il modo m si smorza, s^{-1}) è invece un osservabile diretto dell’*impedenza di boundary*: ogni modo pesa le pareti in modo diverso, quindi α_m porta l’informazione di orientazione che l’energetico collassa. Si articola su tre piani: *predizione* (eigensolver complesso), *misura* su RIR (matrix pencil) e *diagnosi di identificabilità* del Jacobiano $\partial\alpha_m/\partial\beta$.

Predizione (eigensolver complesso). L’eigensolver d’onda DG, validato nel caso *lossless* (autovalori reali = sole frequenze; modi rigidi a $< 1\%$, §2.2), è esteso al boundary a impedenza locale: gli autovalori diventano complessi $s_m = -\alpha_m + j\omega_m$ (frequenza e decadimento per-modo), per perturbazione al prim’ordine $\alpha_m = \frac{c}{2} \text{Re}(\beta) \langle B \rangle_m$, con $\langle B \rangle_m$ l’energia del modo sulle pareti assorbenti. Tre gate:

- *Degenerazione* $\beta = 0$ ($Z \rightarrow \infty$, rigido): $\max |\alpha_m| = 0$ *esatto* e le frequenze coincidono con le lossless entro la precisione FP32 del solver modale GPU ($\sim$ milli-Hz) — la perturbazione si annulla per costruzione.
- *Riferimento analitico* (Robin 1D): un box assorbente sulle sole pareti x riduce i modi assiali al problema di Robin 1D, di soluzione esatta; l’ α perturbativo concorda a 0.09–0.13% (modi (1,0,0)–(3,0,0), $\beta = 0.05$).
- *Coerenza col forward* (codice GPU del motore): il forward d’onda a impedenza (`wave3d_impedance.py`, fp64), eccitato su un modo, decade all’ α_m *predetto* entro il 2.0% ($\beta = 0.08$, sonda puntuale all’antinodo); il risultato GPU coincide con quello della versione CPU di riferimento a $\max |\Delta\alpha| = 2.0 \times 10^{-14}$. La predizione è quindi verificata sul codice effettivamente usato a scala.

L'eigensolver del motore gira su GPU — fattorizzazione sparsa cuDSS in precisione singola + Lanczos shift-invert (stesso algoritmo dello shift-invert di riferimento, su GPU). Sulle mesh modali reali CR2/CR3/CR4 riproduce le frequenze del riferimento CPU a $< 4 \times 10^{-3}$ Hz (milli-Hz, ben sotto la tolleranza fisica di α_m , che è dell'ordine dell'Hz) ed è $2.5\text{--}5.1\times$ più veloce; la versione CPU resta come riferimento di verifica numerica. La precisione singola è sufficiente perché i modi alimentano α_m a tolleranza $\sim\text{Hz}$, non bit-exact.

Misura (matrix pencil su RIR reale). Un estrattore matrix-pencil (Hua e Sarkar, 1990; ~ 40 righe, senza librerie di analisi modale) ricava (ω_m, α_m) da una RIR misurata. Sul dataset pubblico LS (*Low-Frequency Sound Field Control*, Cadavid et al., CC-BY; stanza rettangolare $7.8 \times 4.14 \times 3.0$ m, subwoofer che eccitano i modi bassi sopra il rumore):

Modo	f_{eig} (Hz)	f_{mis} (Hz)	Δf	α_{rigido} (s^{-1})	α_{mis} (s^{-1})
~ 22 Hz	21.97	22.29	+1.50%	0.64	6.52
~ 41 Hz	41.38	42.56	+2.85%	—	7.88
~ 65 Hz	65.90	65.42	−0.72%	1.39	7.61

Tabella 4: Validazione modale su LS (matrix pencil): frequenza eigensolver vs misurata ($< 2.9\%$) e decadimento α_m . L' α da forward DG *rigido* ($\approx 1 \text{ s}^{-1}$, floor numerico) è $\ll$ misurato: il modello lossless non porta smorzamento modale, quindi il decadimento misurato è *interamente determinato dall'impedenza delle pareti*. (Su un caso sintetico a verità nota il recupero è esatto a $\leq 0.02 \text{ s}^{-1}$.)

La frequenza modale misurata concorda con l'eigensolver entro il 2.85%, e il decadimento misurato ($6.5/7.9/7.6 \text{ s}^{-1}$) è incrociato con un fit EDC Schroeder [9]: le due stime indipendenti concordano sui modi LS. Il contrasto col forward rigido è diretto: α_m non è ricavabile dal modello lossless (smorzamento nullo), è l'impedenza di parete a generarlo.

Identificabilità modale. La domanda diagnostica: α_m separa i materiali delle pareti *dove l'energetico è cieco*? Il Jacobiano $\partial\alpha_m/\partial\beta_f$ (per-faccia, 6 facce del box) è analitico, verificato contro differenze finite (accordo al livello di troncatura FD, $\sim 10^{-4}$). Su CR2 (27 modi ≤ 93 Hz) l' α_m ha **rango** 3/6 ($\kappa = 9.8$) contro il **rango** 1/6 ($\kappa = 1.0$) dei descrittori energetici: questi sono funzione della sola $\bar{\alpha}$ area-pesata (una direzione, i materiali collassano), mentre α_m separa l'orientazione $x/y/z$ delle pareti. È un asse diverso dal Jacobiano *design* (§3) e dal Jacobiano *materiale-Mondet* (5 parametri) di CR2: qui l'osservabile è α_m e il parametro l'ammettenza per-faccia.

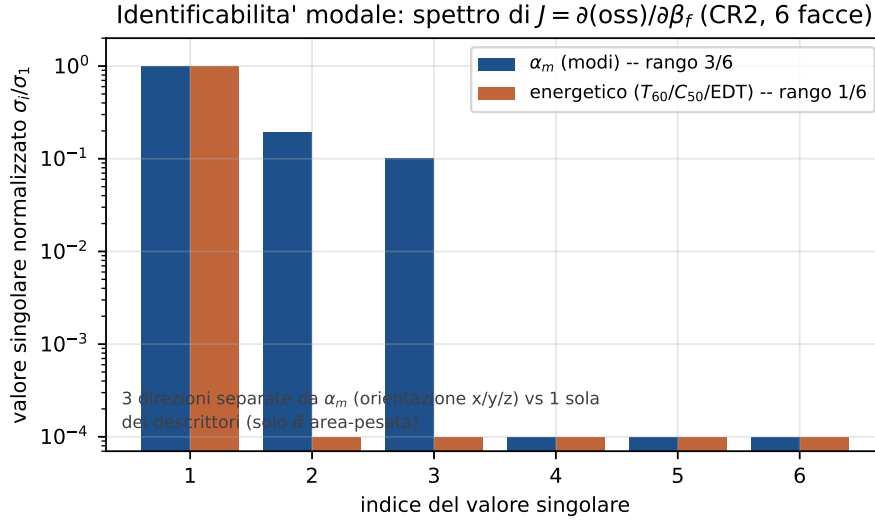


Figura 20: Spettro dei valori singolari del Jacobiano $J = \partial(\text{osservabile})/\partial\beta_f$ su CR2 (6 ammettenze di parete). α_m ha rango 3 (tre valori singolari non nulli: separa l'orientazione $x/y/z$); i descrittori energetici hanno rango 1 (vedono la sola $\bar{\alpha}$ area-pesata). Scala logaritmica; i valori sotto 10^{-6} sono nulli (clampati per la visibilità).

Limiti. (a) *Cap di rango.* Il rango è limitato a 3 dalla simmetria del box (pareti opposte pesano i modi allo stesso modo $\rightarrow$ non si separano le coppie): cap *geometrico*, non difetto del metodo. (b) *Identificabilità $\neq$ sensibilità.* La norma $\|J_\alpha\| = 775 \text{ s}^{-1}$ per unità di β è forte (geometry-driven), ma su CR2 rigida ($\beta \approx 0.025$) il *segnale* $J\beta$ è $[7.7, 1.5, 0.79] \text{ s}^{-1}$ per direzione contro un rumore $\sim 0.8 \text{ s}^{-1}$: la terza direzione è al rumore. (c) *Le sale non-rigide non salvano il segnale.*

Sala (banda modale 20–80 Hz)	$\bar{\alpha}$	β	segnale per direzione (s^{-1})
CR2 (pareti, quasi-rigida)	0.10	0.025	7.7 / 1.5 / 0.79
CR3 (sedute @ LF)	0.14	0.034	10.6 / 2.1 / 1.1
CR4 (sedute @ LF)	0.12	0.029	9.0 / 1.7 / 0.92
bass-trap ipotetico (LF-efficace)	0.40	0.100	30.9 / 6.0 / 3.2

Tabella 5: Segnale α_m per direzione ($J\beta$) vs rumore $\sim 0.8 \text{ s}^{-1}$. Nella banda modale (20–80 Hz) le sedute porose di CR3/CR4 assorbono solo 0.137/0.116 (lo 0.44 è la banda media 500–1000 Hz): la terza direzione resta *marginale* come CR2. Il segnale pieno richiede un assorbitore *efficace a LF* (bass-trap, $\bar{\alpha} \approx 0.4$) assente in BRAS. La marginalità della terza direzione è il *deficit universale di assorbimento LF* delle sale BRAS, non la rigidità specifica di CR2: l'*identificabilità* (rango 3, geometrica) regge ovunque, la *sensibilità* è LF-limitata.

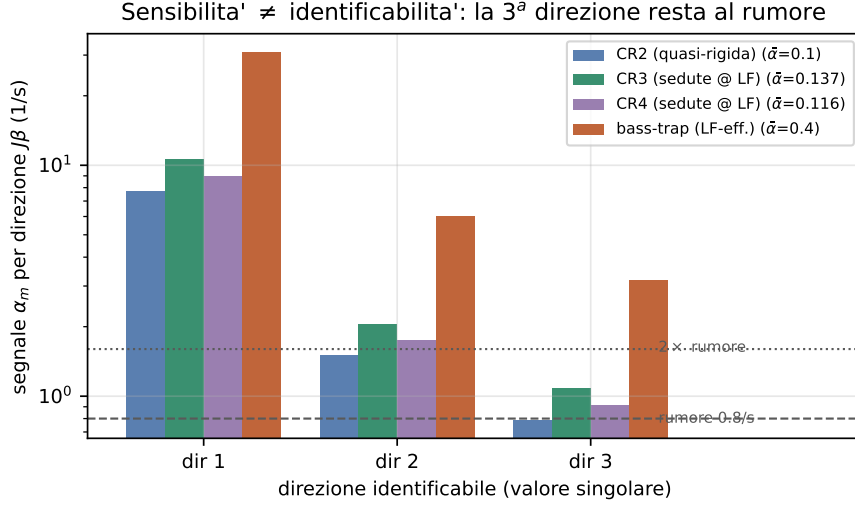


Figura 21: Segnale α_m per direzione ($J\beta$, scala log) per le tre sale BRAS e per un bass-trap ipotetico LF-eficace, contro il rumore di estrazione matrix-pencil ($\sim 0.8\text{s}^{-1}$, tratteggio) e la soglia $2\times$ (punteggiato). La terza direzione resta al rumore su CR2/CR3/CR4 (deficit di assorbimento a bassa frequenza); solo un assorbitore efficace a LF, assente in BRAS, la porta sopra soglia.

Coerentemente la loss $\alpha_m \rightarrow \beta$ è *convessa/unimodale* (Hessiana $2J^\top J$ PSD; slice 1D con un solo minimo) — modello lineare, nessuna frequency-continuation da armare.

Questione aperta. Il confronto α_m *predetto-vs-misurato* su una sala con impedenza Z *nota* e materiali caratterizzati (p.es. l’ambiente 159m^3 di Fratoni e D’Orazio [3]) non è ancora fatto: richiede un dato esterno con $Z(f)$ per superficie, oggi raro. La validazione qui è di due tipi: la predizione è confrontata con la soluzione analitica esatta (Robin 1D) e col forward d’onda indipendente, e separatamente è misurato il decadimento modale su LS (dove lo smorzamento è determinato dall’impedenza delle pareti). Manca ancora il confronto diretto fra α_m predetto e misurato sulla stessa sala con impedenza Z nota.

Relazione con l’estrazione modale da misura. L’estrazione modale da misura (PolyMAX, lo standard di analisi modale del gruppo di Bologna) è *descrittiva*; α_m predetto dalla fisica (eigensolver complesso) è *causale* — lega il decadimento all’impedenza, quindi abilita il design. Il Jacobiano $\partial\alpha_m/\partial Z$ *misura* l’identificabilità che la conversione $\alpha \rightarrow Z$ con prior (Fratoni e D’Orazio [3]) *assume*: dove il rango è pieno l’inversione è vincolata dal dato, dove cade serve il prior.

3.9 Riepilogo dei numeri verificati

Metrica	CR4 (auditorium)	CR3 (musica)	CR2 (seminari)	CR1 (acc.)
Facce / materiali	5148 / 8	2598 / 7	shoebox / 5 (rigidi)	299 / 5
Volume	8657 m ³	3331 m ³	~ 159 m ³	178+127 m ³
T_{60} Eyring (core, 1/3ott) [†]	1.1%	1.3%	— (sala-contrasto)	— (2-slope)
T_{60} ray-tracciato (stesse bande) [‡]	17.1%	11.5%	—	—
Forward per-pos. C_{50} (corr.) ^{‡‡}	+0.89	+0.68	—	acc. [§]
residuo C_{50} livello (dB)	-0.17	+0.70	—	~1.5 [§]
Gradiente unito vs FD	6.9×10^{-10}	4.4×10^{-10}	—	—
Design residuo J^*	4.4×10^{-19}	7.6×10^{-20}	—	—
UQ condition number	25	250	θ -fragile ($\gg 10^6$) [‡]	—
UQ rango	3/3 (pieno)	3/3 (pieno)	3/5 (2 cieche) [‡]	—
Verdetto inverso	ben posto	ben posto*	non-identificabile	C_{50} fit-free [§]
Identif. modale α_m (rango J)	—	—	3/6 (energ. 1/6)	—

Tabella 6: ^{‡‡}Decadimento ray-tracciato sulla geometria reale, senza ricalibrazione. ^{‡‡}La correlazione del pattern è calcolata su 5 posizioni-ricevitore (quante il benchmark ne fornisce): il suo intervallo di confidenza al 95% è per costruzione molto ampio — per $r = 0.88$ vale $[0.0, 0.99]$ — quindi discrimina poco, e il gate discriminante su questo asse è il *residuo di livello*, che poggia su 15 valori (5 posizioni $\times$ 3 bande). [†]Eyring su V/S misurati coi coefficienti BRAS *fitted*, che di quella formula sono l'inversa: l'accordo è in parte imposto, e il suo pavimento (il residuo del fit BRAS) è 1.14% su CR3 (Nota metodologica, §3). La riga sotto è la predizione *indipendente* del ray-tracer (5 ricevitori). [§]CR1 (campo accoppiato, §3.5): C_{50} del volume-ricevitore $-13 \rightarrow -8$ dB (two-part, fit-free) in media ~1.5 dB dalla misura (6/8 coppie banda-rcv entro ~1.5, 2/8 a ~3.8, floor); sustain 2×2 (seconda pendenza dell'EDC, coupling); detection geometrica. *CR3 ben posto ma con una parete dominante. [‡]CR2 è il Jacobiano *materiale* (5 Mondet) — asse diverso dal Jacobiano *design* (3 DOF) di CR3/CR4; il suo κ è θ -fragile ($8 \times 10^8 - 2 \times 10^{16}$), quindi il verdetto riportato è il rango.

3.10 Validazione esterna: robustezza oltre BRAS

BRAS (CR2/CR3/CR4) è il riferimento principale (con la cautela sulle incertezze delle complex rooms, §3) perché fornisce i coefficienti per superficie — permette la forward-validation senza ricalibrazione (§3.1–3.2). Per rispondere a “il motore funziona solo su BRAS?” lo mettiamo alla prova su dataset indipendenti, ciascuno su un asse diverso e scorrelato (sala, propagazione, scattering), dicendo per ognuno che cosa è confrontato e che cosa resta assunto. I numeri che seguono ($r^{-1.03}$, 0.9 cm, 114 $\times$, ...) sono analisi RaySound sui dati pubblici, non risultati dichiarati dagli autori dei dataset.

Sala indipendente — NBI [32] (Auditorium A, Niels Bohr Institute, DTU: altro gruppo, altro paese, geometria diversa, $V \approx 126$ m³). Due test distinti. (i) *Pipeline di misura*. Carichiamo le RIR reali dal SOFA pubblico (SingleRoomSRIR, 22 posizioni $\times$ 32 canali) ed estraiamo il T_{30} in bande: riproduce il T_{30} pubblicato all'~1.2% mediano (~1.8% a 1 kHz; pubblicato 0.72/0.80/0.88/1.03 s), e la catena loader+Schroeder regge fuori da BRAS. (ii) *Sanity del forward* (verifica di plausibilità). Qui i materiali non sono noti; con un modello arredato plausibile l'errore scende a ~17% in banda-core, riproducendo ordine di grandezza e trend atipico (T_{30} che sale con la frequenza), con α implicito sensato (0.11–0.16). I materiali qui sono ipotizzati, e si vede: 17% contro l'1.1% di CR4, dove gli α vengono dal benchmark.

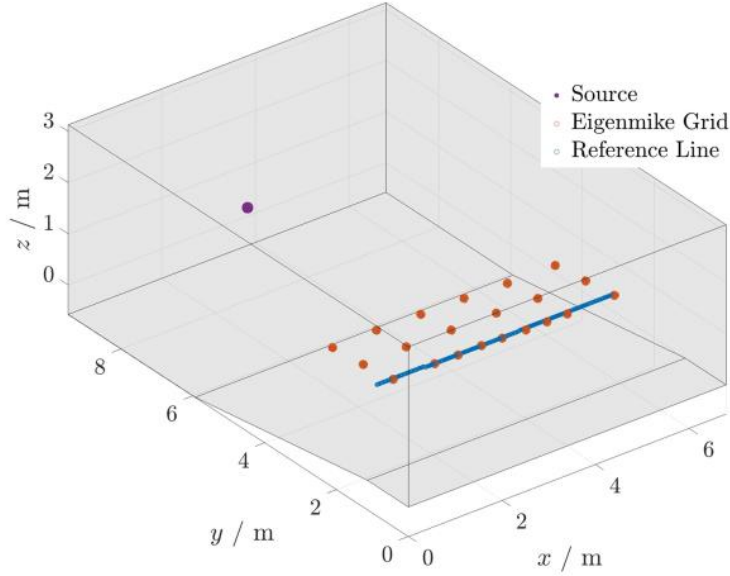


Figura 22: NBI — setup di misura nella sala (sketch ufficiale del dataset, Fig. 1; credito Figueroa-Duran et al., DTU — Zenodo 10.5281/zenodo.14212820, CC-BY-4.0). Scatola $\sim 6.5 \times 8.5 \times 3.2$ m ($V \approx 126$ m³): sorgente fissa, griglia Eigenmike e linea di riferimento posizionata da braccio robotico. Il dataset non fornisce una mesh dettagliata; i conti usano V/S stimati dalla descrizione.

Propagazione — campo sonoro reale, *material-free* (MeshRIR). Un asse diverso da NBI: non il riverbero di una sala, ma la legge di propagazione del forward su un campo sonoro misurato fittamente. Il dataset MeshRIR [33] (Koyama et al.) fornisce 3969 risposte impulsive reali su una griglia cuboidale $21 \times 21 \times 9$ (passo 5 cm) da una sorgente singola a posizione nota. Poiché il suono diretto non dipende dai materiali (solo dalla geometria e dalla legge $1/(4\pi r)$), questo è un test material-free: confrontiamo l'ampiezza del primo arrivo, misurata sui 441 microfoni del piano $z=0$, con la legge $1/(4\pi r)$.

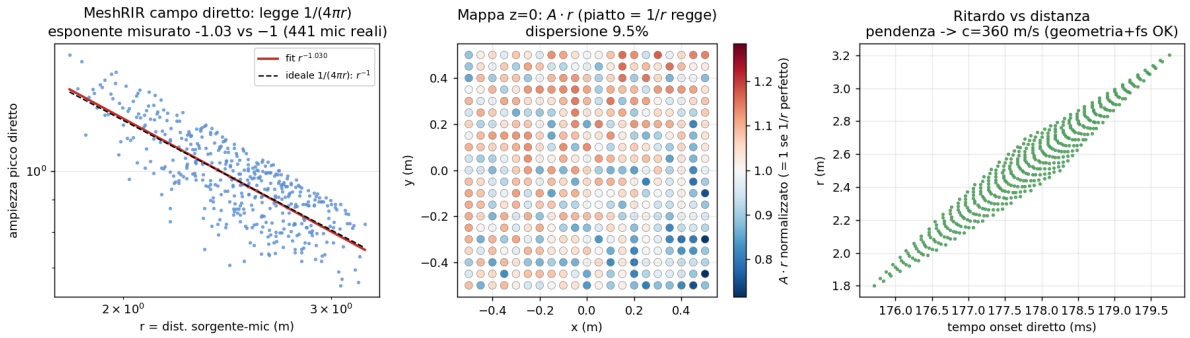


Figura 23: MeshRIR — validazione material-free della propagazione. Sinistra: l'ampiezza del suono diretto vs la distanza sorgente-microfono (log-log); la legge misurata $r^{-1.03}$ coincide con l'ideale r^{-1} di $1/(4\pi r)$. Centro: mappa spaziale di $A \cdot r$ (piatta ≈ 1 se $1/r$ regge, dispersione 9.5%). Destra: ritardo del primo arrivo vs distanza, la cui pendenza dà $c \approx 360$ m/s.

Esito. L'ampiezza decade come $r^{-1.03}$ (atteso -1), con dispersione $A \cdot r$ del solo 9.5% su 441 posizioni reali; l'energia diretta come $r^{-1.98}$ (≈ -2 — misura *indipendente*: l'energia integra p^2 sulla finestra del diretto, quindi non è l'esatto quadrato dell'esponente d'ampiezza (-1.98 vs $2 \times -1.03 = -2.06$); entrambe confermano la legge $\approx 1/r$); il ritardo è lineare nella distanza (pendenza $\rightarrow c \approx 360$ m/s, $\sim 5\%$ sopra 343: la stima onset-slope conferma la *linearità* tempo-distanza, non è una misura di precisione della velocità). Il confronto riguarda la legge di scala e

non coinvolge i materiali; conferma la convenzione fisica $1/(4\pi r)$ di RaySound (pressione $\propto 1/r$, il 4π nella costante) contro $1/r$ dei simulatori che la omettono. Valida la propagazione, non la sala completa (le riflessioni e il T_{60} richiederebbero i materiali della stanza MeshRIR, non documentati).

Geometria — struttura temporale degli echi vs laser-mm (FLAIR). Un quarto asse: non l’energia, non la propagazione, ma la geometria codificata nei tempi d’arrivo (TOA) degli echi precoci. Il dataset FLAIR [34] (Sundström et al.) fornisce 270 risposte impulsive reali, le posizioni di microfoni e sorgenti, e una nuvola di punti 3D laser-calibrata al millimetro (2.9 milioni di punti) della stanza ($5.78 \times 5.62 \times 3.35$ m). Confrontiamo i tempi d’arrivo misurati con quelli predetti dalla geometria laser: il primo arrivo con $|\text{mic} - \text{src}|/c$, e la prima riflessione (pavimento/soffitto) con l’image-source attraverso quelle superfici.

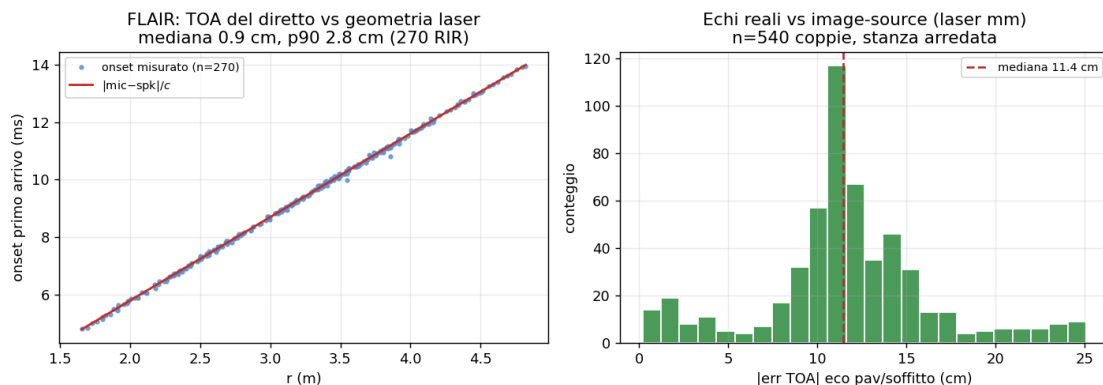


Figura 24: FLAIR — consistenza geometria/TOA. Sinistra: l’onset del primo arrivo vs distanza segue $|\text{mic} - \text{src}|/c$ (mediana 0.9 cm, $p90$ 2.8 cm su 270 RIR). Destra: gli echi pavimento/soffitto confrontati con l’image-source dalla geometria laser ($n=540$ coppie, stanza arredata).

Esito. Il tempo del primo arrivo coincide con $|\text{mic} - \text{src}|/c$ a una mediana di 0.9 cm ($p90$ 2.8 cm, massimo 13 cm su 270 RIR), a conferma di posizioni, c e frequenza di campionamento. L’onset è quello del *motore* (`descriptors._onset_index`, primo campione sopra il 2% del picco: la convenzione con cui i descrittori separano early/late), non un rilevatore scritto per l’occasione. Gli echi di pavimento/soffitto sono confrontati con l’image-source dalla geometria laser ($n=540$ coppie mic-sorgente $\times 2$ superfici); il residuo è più grossolano del diretto (echi sovrapposti nella stanza arredata). La struttura temporale degli echi misurati è dunque coerente con la geometria laser-mm (2.9 milioni di punti) tramite il modello image-source: l’osservabile è il TOA, e non dipende dai materiali. Fonte: mirror Zenodo CC-BY (record 17037517).

Design — la distribuzione del trattamento conta (Arni). Il quinto asse riguarda il design (§3.1): finora validato sul gradiente e sulla convergenza, mai sulla predizione dell’effetto di un trattamento reale. Il dataset Arni (Aalto — 55 pannelli a parete commutabili riflettente/assorbente, 5342 configurazioni, 132k risposte impulsive) è il design al contrario: si varia il trattamento, si misura l’acustica.

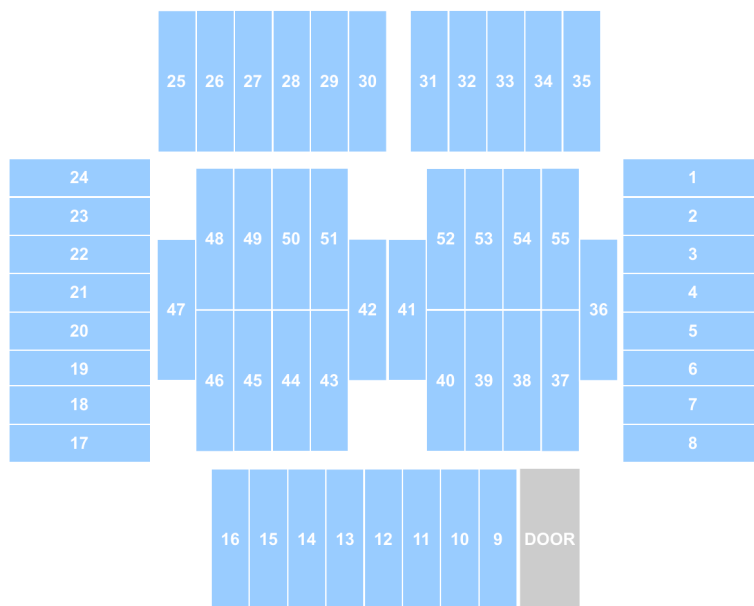


Figura 25: Arni — setup (dal dataset): pianta della sala variabile (Aalto) con i 55 pannelli a parete commutabili riflettente/assorbente, numerati. È la combinazione di pannelli accesi/spenti a variare, a parità di area trattata. Diagramma originale del dataset (credito Aalto Acoustics Lab). Nessuna foto pubblica del laboratorio è inclusa.

Arni (Aalto) — sala ad acustica variabile (schematica)
 $7.87 \times 5.75 \times 2.91$ m — 55 pannelli (35 parete + 20 soffitto)

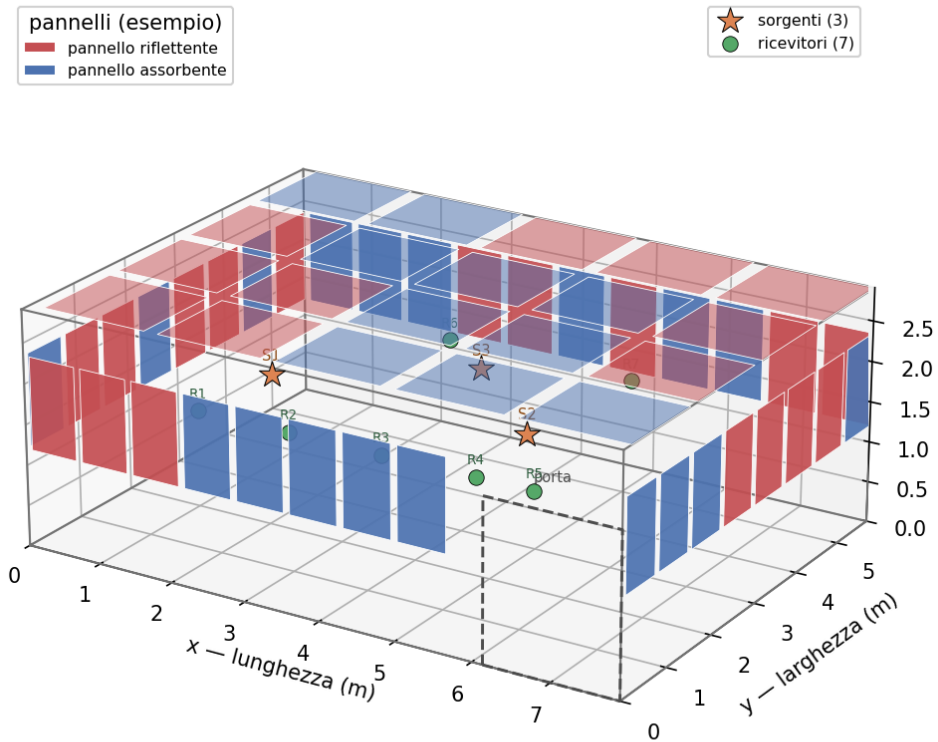


Figura 26: Arni — vista 3D schematica della sala. Scatola $7.87 \times 5.75 \times 2.91$ m (dimensioni reali, McKenzie et al., arXiv:2111.11882, Fig. 1) con i 55 pannelli ad acustica variabile (35 a parete, 20 a soffitto); il colore riflettente/assorbente è un *esempio* di configurazione (nel dataset la combinazione varia). Le 3 sorgenti (stelle) e i 7 ricevitori (cerchi) sono alle coordinate reali (Tab. I del medesimo lavoro). La posizione *per-pannello* è illustrativa; conteggio e ripartizione parete/soffitto sono fedeli.

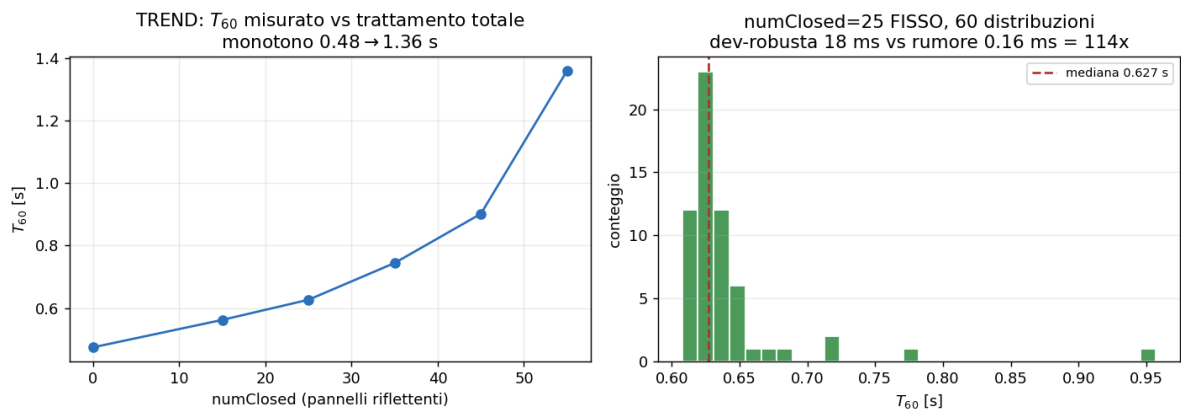


Figura 27: Arni — l’asse design. Sinistra: il T_{60} misurato cresce monotono col numero di pannelli riflettenti (0.48 s con 0 → 1.36 s con 55). Destra: a trattamento totale fisso (25 pannelli riflettenti), 60 configurazioni con distribuzione diversa — il T_{60} varia con deviazione robusta 17.9 ms, contro 0.16 ms di ripetibilità (stessa configurazione, più sweep). Entrambi i pannelli sono ricalcolati da RaySound sulle RIR reali del dataset (generatore `arni_distribution_remote.py`, descrittore `t30_noise_compensated`).

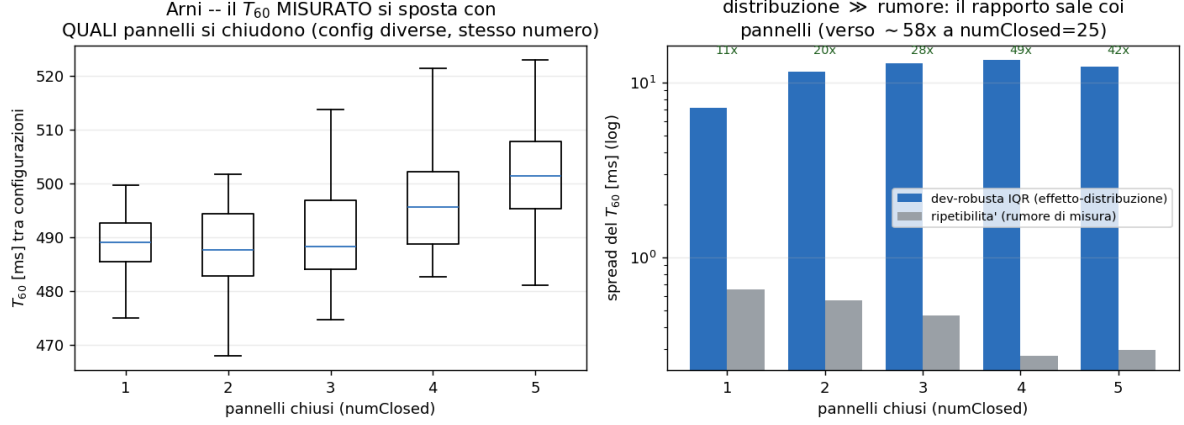


Figura 28: Arni — l’effetto-distribuzione *ricalcato da RaySound* sulle RIR reali del dataset (Zenodo record 6985104, gruppo numClosed 0–5; T_{60} dal descrittore del motore `t30_noise_compensated`). Sinistra: la distribuzione del T_{60} fra configurazioni a numero di pannelli fisso — a parità di *quanti* pannelli si chiudono, *quali* si chiudono sposta il T_{60} (la mediana e lo spread crescono coi pannelli). Destra: la deviazione robusta fra configurazioni (effetto-distribuzione, blu) contro la ripetibilità fra sweep ripetuti (rumore di misura, grigio, scala log); il rapporto sale coi pannelli (11 → 49×) verso il 114× a numClosed = 25.

Esito. A trattamento totale fisso (stesso numero di pannelli riflettenti, dunque stessa area di assorbimento), 60 distribuzioni diverse danno T_{60} con deviazione robusta 17.9 ms contro una ripetibilità di 0.16 ms: un fattore 114× oltre il rumore. Il numero è *misurato da RaySound* sulle RIR reali col proprio descrittore `t30_noise_compensated` (generatore `arni_distribution_remote.py`: le RIR dei cinque zip da 4–9.6 GB si leggono per singolo membro via HTTP-Range, senza scaricare il dataset). Lo stesso effetto misurato sul gruppo 0–5 (`arni_t60_variation.py`) cresce col trattamento — deviazione robusta da 7 a 13 ms, ripetibilità 0.27–0.66 ms, rapporto da 11 a 49× — e sale al 114× a numClosed = 25. La distribuzione conta: a parità di quanta area si tratta, dove la si tratta cambia l’acustica oltre il rumore. Questo principio del design è dunque fisicamente fondato su dato reale misurato, non solo numericamente (gradiente/convergenza). [Il trend totale (0.48 → 1.36 s) è misurato; la sua predizione quantitativa con Eyring+ α dei pannelli è calibrata in letteratura (Prawda, Schlecht, Välimäki, JASA), non rieseguita qui — è un trend misurato con riferimento, non un “forward RaySound predice”.]

Arni: motivazione del design position-aware. Un modello total-only (Eyring classico, che vede solo $\sum_i \alpha_i S_i$) predice un T_{60} a parità di area trattata, e non può spiegare la varianza di 114× misurata da Arni: serve un forward che veda dove sono le superfici. È esattamente ciò che il design di questo report fa (§2.3, §3.1–3.2): il T_{60} è position-aware (Eyring hit-weighted, $\bar{\alpha}_{\text{eff}} = \sum_m h_m \alpha_m$), ancorato esattamente a Eyring nel limite diffuso. Arni motiva (la distribuzione conta, 114× su dato reale) e CR3/CR4 dimostrano (il design position-aware sceglie soluzioni diverse fino al ~12% là dove la geometria ri-pesa). La distribuzione è il principio che il design sfrutta, validato sul forward giusto.

Limite — scattering (GIR). L’asse scattering (GIR, ETH Zürich: superfici stampate 3D + RIR per-pannello) resta *scoperto*: le posizioni assolute sorgente/ricevitore non sono nei dati pubblici, quindi un confronto sarebbe un gate falso.

Sensibilità-materiali controllata — RSoANU [37] (recording studio, Australian National University: altro gruppo, altro continente, altra sala). È l’asse che gli altri non toccano: gli stessi ambiente, sorgenti e griglia di ricevitori, in cui cambiano soltanto i pannelli a parete (feltro ↔ legno) — l’una-variabile-alla-volta che BRAS non isola (ogni stanza BRAS varia geometria e materiali insieme). Il dataset fornisce RIR B-format reali (canale omni W , 48 kHz) su una griglia

di 64 ricevitori per due configurazioni di parete, con planimetria e foto reali (Fig. 29–30). (i) *Dato vero — il Δ misurato*. Sui 64 ricevitori appaiati (stessa posizione, stessa sorgente S1; T_{30} di Schroeder, 1/3-ottava 250–2000 Hz, T_{30} noise-compensated) il legno dà $T_{60} = 0.924$ s e il feltro 0.815 s: $\Delta = +13\%$ (legno–feltro), col segno fisico corretto — il feltro poroso assorbe (T_{60} più corto), il legno riflette (T_{60} più lungo). Questo è una misura vera, indipendente da ogni assunzione di materiale. (ii) *Sanity del forward — sul Δ , non sull'assoluto*. I coefficienti dei pannelli non sono caratterizzati nel dataset, quindi non si predice il Δ assoluto; si verifica la consistenza. Invertendo via Eyring il Δ misurato (box 8.5×9 footprint reale dalla planimetria, altezza ~ 3 m assunta) si ottiene $\Delta \bar{\alpha}_{\text{eff}} \approx 0.018$, cioè $\Delta \alpha_{\text{pannello}} \approx 0.09$ sul $\sim 20\%$ di superficie trattata — un salto α feltro-vs-legno fisicamente plausibile (~ 0.1). Il Δ torna, il suo valore assoluto no: un forward Eyring naïve con α da letteratura per feltro spesso sovrastima il Δ di $\sim 5.3\times$, tanto è sensibile all' α assunto. Il numero misurato è il $+13\%$; il forward ci arriva con α e altezza assunti.

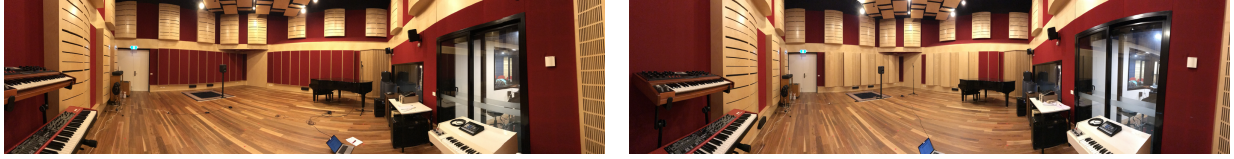


Figura 29: RSoANU — le due configurazioni reali della stessa sala (foto del dataset, credito ANU/RSoANU): pannelli di feltro (sinistra, assorbente) e di legno (destra, riflettente). Tra le due cambiano solo i pannelli a parete — ambiente, sorgenti e griglia di ricevitori identici: l'isolamento controllato della sensibilità-materiali.

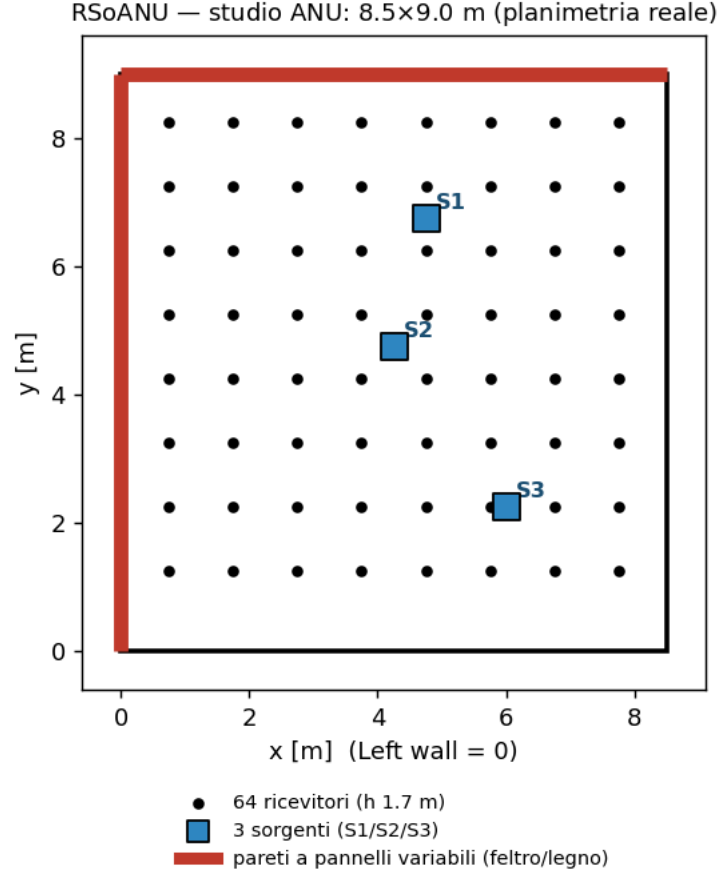


Figura 30: RSoANU — geometria reale (a differenza della scatola schematica NBI): footprint 8.5×9 m dalla planimetria pubblicata (Fig. 2 del dataset), con le pareti a pannelli variabili evidenziate, la griglia di 64 ricevitori e le 3 sorgenti dai metadati. L'unica assunzione è l'altezza (~ 3 m, non fornita); footprint, posizioni e configurazioni sono dati reali.

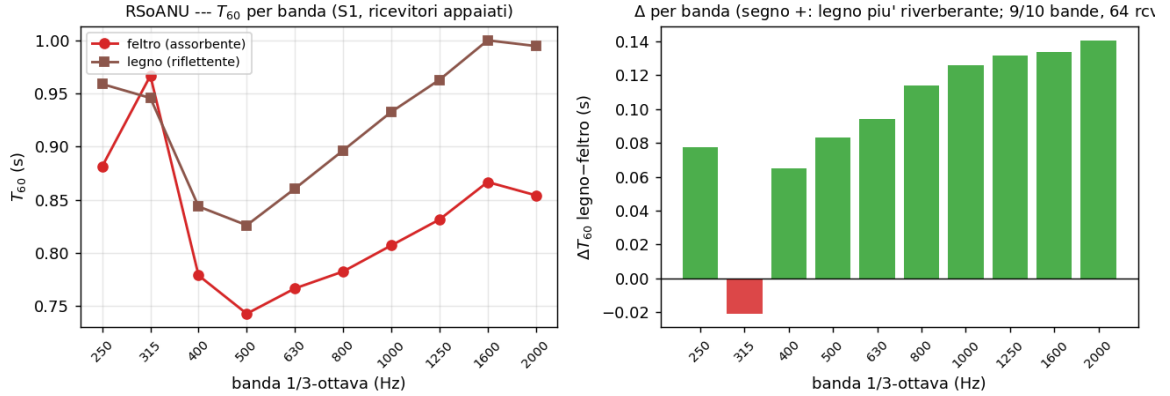


Figura 31: RSoANU — ΔT_{60} legno-feltro per banda (1/3-ottava 250–2000 Hz, T_{30} di Schroeder, 64 ricevitori appaiati S1). Sinistra: T_{60} feltro (assorbente) vs legno (riflettente); destra: il Δ per banda. Il segno è corretto in 9 delle 10 bande (legno più riverberante; a 315 Hz il Δ è marginalmente negativo, -0.02 s) e il Δ cresce con la frequenza ($+0.08$ s a 250 Hz $\rightarrow +0.14$ s a 2 kHz) — la firma attesa di un pannello di feltro poroso (assorbimento crescente alle alte). È il dato misurato; nel forward- Δ l' α e l'altezza del soffitto sono assunti.

Geometria complessa — eco di parete su sala non-shoebox (MP-RIR). Un settimo asse, sul *forward geometrico* dove la pianta NON è una scatola. MP-RIR [36] (Mehra/N4DAnalytics)

fornisce una griglia densa di RIR misurate in una sala a *pianta trapezoidale* (una parete obliqua). Calibrando c dai cammini diretti (343.4 m/s, mediana su 32 mic) e predicendo i tempi d'arrivo delle riflessioni di parete dalla mesh reale (trapezio), l'eco di ogni parete — inclusa la *parete obliqua*, il discriminante che una scatola non riproduce — è predetto a **sotto il millisecondo** ($|\Delta\text{TOA}|$ mediano per parete 0.18–0.58 ms, cioè 6–20 cm; sulla parete obliqua 0.53 ms dove l'eco è osservabile, colonne vicine). La stessa predizione fatta sulla *scatola* che meglio approssima la pianta (bounding box) sbaglia di 3.15 ms: il trapezio la batte di $6.0\times$. L'osservabile è il TOA: non dipende dai materiali né dall'altezza del soffitto, e completa FLAIR (diretto+riflessioni precoci su geometria laser) e il recupero-geometria modale (§3.8). Lo script `experiments/mprir_forward_walls.py` ricalcola i residui per-parete dalla griglia MP-RIR.

Asse	Dataset	Numero	Confronto, e cosa è assunto
Sala (pipeline)	NBI	T_{30} a 1.2% (mediana)	T_{30} dalle RIR vs T_{30} pubblicato
Sala (forward)	NBI	T_{60} a $\sim 17\%$	materiali non noti, ipotizzati
Propagazione	MeshRIR	$r^{-1.03}$, dispersione 9%	ampiezza del diretto vs $1/(4\pi r)$; nessun mate
Geometria/TOA	FLAIR	0.9/11.4 cm (diretto/rifl.)	TOA misurato vs nuvola laser-mm
Design/distrib.	Arni	$114\times$ il rumore	spread fra distribuzioni vs ripetibilità; nessun
Sensib.-materiali	RSoANU	$\Delta +13\%$ legno-feltro (64 rcv)	Δ misurato; nel forward α e altezza assunti
Geom. complessa	MP-RIR	eco-parete TOA sub-ms (parete obliqua)	TOA dell'eco vs mesh trapezoidale; nessun m
Scattering	GIR	—	posizioni non pubblicate: nessun confronto

Tabella 7: I sette assi di validazione esterna, con l'osservabile confrontato e gli ingredienti assunti.

3.11 Profilo di performance (misurato)

Il costo vive in due posti — il forward d'onda e l'inverso dei materiali; design e diagnosi sono interattivi (millisecondi). I tempi sono misurati con `perf_timer` e CUDA-events su GB10, con lo script che li rigenera indicato in didascalia.

Operazione	Tempo	Scena	Speedup
Forward d'onda (DG, banda modale)	7.4 / 7.4 s	CR3 / CR4	3.8 / $3.6\times$ (fp32 vs fp64)
Eigensolver modale (GPU cuDSS-FP32)	0.2–0.8 s	CR2/CR3/CR4	$2.5\text{--}5.1\times$
Forward GA diffuso (ray-tracer)	16 / 12 ms	CR4 / CR3	—
Forward GA speculare (enum, ordine 2)	34 / 29 ms	CR4 / CR3	—
Diffrazione (BTM)	< 1 ms/edge	—	$4.5\times$
Inverso materiali — 1 gradiente	15 s	CR2 (345k DOF)	—
Inverso materiali — inversione completa	~ 4 min	CR2	—
Inverso wave-FWI (modale) [†]	$\sim 30\text{--}60$ s	CR4 ($K \sim 200\text{--}650$, $15\text{--}50\text{k}$ DOF)	—
Solver stress-scaling (max girato) [†]	~ 11 min/grad	CR2-shoebox (mesh 5.6M DOF)	—
Design + diagnosi (gradiente + L-BFGS)	< 5 ms	CR3 / CR4	—

Tabella 8: Performance del motore: operazione, tempo misurato (`perf_timer`/CUDA-events, GB10), scena. Le prime quattro righe si rigenerano con `experiments/regen_report_performance.py` (`performance.json`): forward DG sui terzi modali della scena, diffuso a 128k raggi / 2 s / 5 bande, enumerazione speculare all'ordine 2 sulla mesh reale; lo *speedup* del forward DG è fp32 contro fp64 sullo stesso caso. Le righe dell'inverso e del design vengono dai rispettivi bench citati nel testo. [†]Il wave-FWI è dimensionato a f_{Sch} : il router manda al solver d'onda solo le bande modali ($\leq f_{\text{Sch}}$), mentre ≥ 500 Hz va alla GA. CR4 ($f_{\text{Sch}} \lesssim 30$ Hz) mesha quindi a $K \sim 200\text{--}650$ ($15\text{--}50\text{k}$ DOF): un gradiente costa pochi secondi e l'inversione completa (~ 15 gradienti) $\sim 30\text{--}60$ s, ancorata al bench ($K=144$ inversione 29 s; $K=360$ gradiente 2.8 s). Il valore di 5.6M DOF (~ 11 min/gradiente, $K=47040$) è uno *stress-test di scaling del solver* (CR2-shoebox a mesh arbitraria, finestra 50 ms, un muro): mostra che il solver regge fino a 5.6M DOF, *non* è un costo operativo. Il material-ID resta via Eyring nel diffuso ($\sim$ ms).

Sintesi. Il forward d’onda gira in 7.4 s sui terzi modali di CR3 e di CR4; il forward GA su RT-core OptiX in decine di millisecondi (12–16 ms il diffuso, 29–34 ms l’enumerazione speculare all’ordine 2); design e diagnosi sono interattivi (< 5 ms); l’inverso dei materiali converge in ~ 4 min (15 s/gradiente). Misure su NVIDIA GB10.

4 Discussione

Oltre alla simulazione forward — comune a tutti i pacchetti commerciali (ODEON, CATT-Acoustic, Treble) — il motore aggiunge due capacità poco diffuse: il design inverso dal target al trattamento, con gradiente verificato a precisione macchina su sale reali, e la diagnosi di identificabilità del problema inverso (rango e condizionamento del Jacobiano) prima di calibrare. La calibrazione dei materiali su misura non è di per sé nuova — esistono ottimizzatori che adattano i coefficienti di assorbimento a un tempo di riverbero misurato — ma l’elemento distintivo qui è *dichiarare quando l’inversione è mal posta*, invece di restituire una stima non difendibile. Il design ora gira sul forward giusto: è position-aware; un benchmark indipendente (Arni) lo fonda su dato reale (la distribuzione conta $114\times$ il rumore), e su CR3/CR4 cambia il progetto fino al $\sim 12\%$ dove la geometria lo richiede. Il motore *diagnostica* l’identificabilità su domande diverse — i gradi di libertà di progetto (CR3/CR4) e la recuperabilità dei materiali (CR2), Jacobiani diversi — con un verdetto difendibile in entrambi: design ben-posto dove i dati lo consentono, materiale non recuperabile dove non lo consentono, segnalato con un numero (rango, condizionamento) invece di una stima non difendibile.

Limiti. (i) Il forward è validato su reale per il T_{60} coi coefficienti BRAS fitted, ma l’identificazione della α -fisica da reale resta limitata dalla dinamica dei materiali e dal regime: la α assoluta è impossibile dai soli descrittori in sale rigide. (ii) Il design è validato sul gradiente (concordanza con FD a 10^{-10}) e sulla convergenza al target, non come fisica end-to-end su tutte le bande; il modale d’onda è gated ai grandi volumi. (iii) Gli invarianti di campo diffuso sono per ora un canale diagnostico complementare, non ancora un termine di loss attivo; fra essi la NED è robusta mentre la kurtosi dell’involuppo è più rumorosa. (iv) Il boundary a impedenza complessa (eq. (1)–(2), $R(\theta)$ validato a 9.5×10^{-3}) è ora *costruito* anche nel forward d’onda 3D. Il codice GPU a impedenza (`cuda/wave3d_impedance.py`, fp64, validato vs la versione CPU di riferimento a $\sim 10^{-13}$) e l’eigensolver complesso predicono il decadimento modale α_m (§3.8). La predizione è verificata sulla soluzione analitica esatta (Robin 1D, 0.09–0.13%) e sulla coerenza col forward d’onda (2.0%). Resta aperto il confronto *end-to-end* su una sala *reale* con Z noto e materiali caratterizzati (Fratoni e D’Orazio [3]): il limite si *restringe* dal boundary-in-sé al dato-esterno, non sparisce. Sulle sale BRAS qui il forward d’onda 3D opera in regime rigido, ed è Eyring/ray-tracer a coprire sopra Schroeder. **Sul design position-aware, tre precisazioni:** (a) il C_{50} era già position-aware (riflessioni image-source reali); il nuovo ingrediente è il T_{60} hit-weighted. (b) L’effetto è significativo sulla geometria reale (balconata, sedute: $h \neq \text{area}$) e piccolo su una scatola ($h \approx \text{area}$, campo quasi-diffuso) — è la geometria a renderlo rilevante. (c) Il T_{60} hit-weighted resta un modello energetico (Eyring ri-pesato per superficie), non il forward d’onda: è position-aware nell’assorbimento per-superficie, ancorato a Eyring, non un full-wave.

5 Conclusioni e prospettive

Su quattro sale reali BRAS, RaySound dimostra un motore ibrido che non solo simula ma progetta e diagnostica: pipeline coerente senza ricalibrazione (T_{60} Eyring a 1.1% / 1.3% su CR4 / CR3, 1/3 d’ottava — coefficienti BRAS fitted, quindi coerenza-pipeline a materiali noti e non stima cieca di α ; il ray-tracer, predizione indipendente, sta al 17.1% / 11.5% sulle stesse bande), design inverso con gradiente a precisione macchina e libertà di progetto a rango pieno, e il riconoscimento dei casi in cui i parametri non sono identificabili (CR2, materiale rank-deficient 3/5). Il canale di invarianti di campo diffuso aggiunge informazione ortogonale ai descrittori ISO, provata su dato misurato e

capace di quantificare un limite del forward. Tutto questo passa per un *ingresso unico* non-box (`simulate_raytracer_mesh`): carica una sala reale (mesh body-fitted dal mesher RaySound), ne risolve il regime (densità modale di Weyl completa, $f_{\text{Sch}} \sim 30$ Hz CR3 / ~ 20 Hz CR4), gira il forward FDTD+GA sulla geometria vera, e restituisce i descrittori early/ C_{50} /per-posizione (pattern +0.89/+0.68) — una chiamata. Il T_{60} integrato è fornito da Eyring (predittore $\approx$ misura) per un *limite noto e quantificato* del forward-diffuso: su geometria complessa sovrastima il T_{60} del $\sim 15\%$ per la distribuzione delle lunghezze di cammino libero (§2.2).

Tre leve di progetto, e dove siamo. Il motore espone tre *handle* di design: i moltiplicatori di trattamento per-superficie (rango pieno), l' α -fisica dei materiali (via Eyring nel diffuso, aggiunto d'onda nel modale), e l'*aggiunto geometrico* (shape derivative sulla mesh, gradiente validato contro differenze finite a $\sim 10^{-9}$) — la forma stessa come terzo asse. *Geometria: presente* — il forward d'onda gira già su mesh body-fitted arbitraria (`simulate_dg_mesh` sulla tet del mesher RaySound), non è un ostacolo. *Inverso materiale: il routing è realizzato.* L' α dei materiali è identificata via **Eyring nel campo diffuso** ($\sim$ ms): dal T_{60} diffuso per-banda all' $\alpha(f)$ via Eyring, poi fit di un modello a pochi parametri (Miki [17]) che rende ben-posta l'inversione $\alpha \rightarrow$ impedenza altrimenti non-unica [2]. Validato *no-inverse-crime* su reale: su CR3/CR4 l' α recuperata dal T_{60} *misurato* coincide con l' α -fitted BRAS area-pesata a $|\Delta\alpha| = 0.006\text{--}0.007$ — l'inverso del forward già validato (1.1%/1.3%), proprio dove l'informazione sull' α vive (il diffuso); a costo trascurabile abilita la sperimentazione iterativa (sweep di configurazioni: 50 in ~ 0.3 s). Il **wave-FWI è riservato alla precisione modale** — la banda dove l'onda è essenziale — ed è ottimizzato (kernel monolitico, fused-RK4 e stato ϕ -compatto boundary-only: per-gradiente $\sim 2.1\times$, checkpoint $\sim 4.5\times$ più piccolo, tutto bit-identico). Ri-misurato alla fonte (post-fusione, GB10, mediana di 3) il gradiente costa 1.9–15 s (mesh $K=144\text{--}2880$) e una sala modale piccola ($K \sim 144\text{--}360$, il target) inverte in ~ 30 s, praticabile. Il wave-FWI è dimensionato a f_{Sch} : su CR4 ($f_{\text{Sch}} \lesssim 30$ Hz) la mesh modale è $K \sim 200\text{--}650$ e l'inversione completa $\sim 30\text{--}60$ s. Il solver regge uno stress-test fino a 5.6M DOF (~ 11 min/gradiente, CR2-shoebox a mesh arbitraria): è una capacità del solver, non un costo operativo. E la loss-waveform per l'impedenza è *unimodale*: la geometria nota dal CAD fissa i tempi d'arrivo (niente cycle-skipping), quindi la frequency-continuation resta auto-deferita (servirebbe solo all'inversione *geometrica*, fuori scope). Lì però, per il finding CR2, i materiali sono comunque *appena* identificabili. È il *frequency-routing*: l'onda dove è essenziale (modale), Eyring dove l' α è identificabile (diffuso). *Limite dichiarato*: Eyring dà la *magnitudine* $|\alpha|$, non la fase, che solo l'onda vede [2]; ma nel modale, dove la fase conterebbe, l'identificabilità è scarsa — il limite cade dove comunque non identifichi. *Altre prospettive*: l'identificazione della α -fisica da reale dove l'identificabilità lo consente; la copertura garantita (conformal prediction) sui descrittori normativi; e l'interfaccia GUI-3D (carica-sala $\rightarrow$ simula $\rightarrow$ progetta) come direzione di prodotto. **Limite di validazione e copertura, esplicito.** La validazione forward completa (geometria + materiali $\rightarrow$ descrittori, non per-asse) è oggi sulle quattro sale BRAS-CR. Due virtù del motore — forward d'onda *essenziale* (banda modale larga, sala piccola) e materiali *identificabili* — non coesistono nei casi presenti: le sale piccole sono rigide, quindi non-identificabili (CR2); le grandi sono diffuse, dove il forward d'onda non è essenziale (CR3/CR4). Estenderla a sale reali indipendenti con materiali caratterizzati (teatri, aule) richiede dataset con $\alpha(f)$ misurati per superficie, oggi rari; RSoANU (§3.10) è il primo passo — una sala reale dove i materiali variano in modo controllato: per ora dà la sensibilità (il Δ legno-feltro, misurato vero), non la validazione assoluta (coefficienti non caratterizzati). È un limite dello scope, dichiarato. **Profilo hardware del motore:** il forward grande-volume è compute-bound (~ 270 M DOF·step/s su GB10) e in doppia precisione — GB10 (128 GB unified, FP64) è l'hardware naturale per il forward d'onda e per l'inverso con checkpointing. Ciascuna prospettiva sopra è ancorata a una capacità già costruita e validata nel motore.

Appendice M — Note metodologiche

Le cautele, distinzioni ed etichette ricorrenti sono raccolte qui una volta; il corpo le richiama con un rimando secco (cfr. App. M).

M.1 — Coefficienti BRAS (fitted). Gli $\alpha(f)$ BRAS sono *fitted* (stimati per far coincidere il T_{60} di Eyring col misurato). La forward-validation verifica quindi la coerenza geometria-coefficienti-descriptori della pipeline, senza ricalibrazione RaySound, non una stima cieca di α fisici indipendenti (sarebbe in parte circolare). Il benchmark segnala incertezze maggiori sulle *complex rooms*: le CR sono confronto controllato, non benchmark assoluto.

M.2 — Eyring vs forward-GA nel diffuso. Per $M \gg 3$ il forward d'onda converge a Eyring sul T_{60} medio; il forward *geometrico*-diffuso (GA) lo sovrastima del $\sim 10\text{--}15\%$. Causa, isolata per eliminazione (non il cammino-libero-medio, corretto entro l'1%; non lo scattering; non la ricostruzione): la distribuzione delle lunghezze di cammino libero (CV 0.88 su CR3 e 1.14 su CR4 contro 0.64 su shoebox; effetto Kuttruff). I coefficienti α BRAS, calibrati su Eyring (cammino singolo), non si trasferiscono al forward a corde-risolve. Perciò il report adotta Eyring per il T_{60} (1.1–1.3% dalla misura) e il forward-GA per gli altri descriptori (early, C_{50} , mappe), che non ereditano il bias.

M.4 — Identificabilità CR2 (θ -fragile). Il κ del Jacobiano materiale-Mondet di CR2 varia da 8×10^8 a 2×10^{16} al variare del materiale near-rigid: il verdetto stabile è il *rango* (3/5), non un κ singolo. La rank-deficiency è fisica (tre vie: ogni θ near-rigid dà *rango* < 5 ; il fit $\alpha \rightarrow \theta$ è degenerare; la riflessione quasi-piatta acceca i parametri).

M.5 — Material-ID (no-inverse-crime). L' α dei materiali è identificata via Eyring nel diffuso e validata *no-inverse-crime* su CR3/CR4: l' α recuperata dal T_{60} *misurato* coincide con l' α -fitted BRAS area-pesata a $|\Delta\alpha| = 0.006\text{--}0.007$.

M.6 — Forward d'onda: regimi, costo, hardware. Il router manda ogni banda al metodo valido: sotto f_{Sch} (modale) il forward d'onda DG, sopra (diffuso) ray-tracer+Eyring; la densità modale è la legge di Weyl completa (volume+superficie). Il forward d'onda gira sui soli terzi modali (costo a 1/3 d'ottava = limite superiore). Il wave-FWI è dimensionato a f_{Sch} (il router manda al solver d'onda solo le bande $\leq f_{\text{Sch}}$, ≥ 500 Hz alla GA): su CR4 ($f_{\text{Sch}} \lesssim 30$ Hz) la mesh è $K \sim 200\text{--}650$ (15–50k DOF) e l'inversione completa $\sim 30\text{--}60$ s (bench: $K=144$ inversione 29 s, $K=360$ gradiente 2.8 s). Il valore di 5.6M DOF ($K=47040$, ~ 11 min/gradiente) è uno *stress-test di scaling del solver* (CR2-shoebox a mesh arbitraria) — mostra il solver reggere fino a 5.6M DOF, non un costo operativo. Il material-ID resta via Eyring nel diffuso ($\sim \text{ms}$). Hardware: NVIDIA GB10 (FP64 nativo), forward grande-volume compute-bound (~ 270 M DOF·step/s).

Riferimenti

Riferimenti verificati alla fonte (DOI). I dataset esterni sono analizzati da RaySound: i numeri riportati sono nostri, non claim degli autori originali.

- [1] Aspöck L., Brinkmann F., Ackermann D., Weinzierl S., Vorländer M. (2020). *BRAS: Benchmark for Room Acoustical Simulation* (v3). Technische Universität Berlin. DOI 10.14279/depositonce-6726.3. (coefficienti per superficie fitted; incertezze maggiori sulle complex rooms.)
- [2] Mondet B., Brunskog J., Jeong C.-H., Rindel J.H. (2020). *From absorption to impedance: enhancing boundary conditions in room acoustic simulations*. Applied Acoustics **157**, 106884. DOI 10.1016/j.apacoust.2019.04.034.

- [3] Fratoni G., D’Orazio D. (2025). *Boundary conditions for hybrid simulations in a rectangular environment with sound-absorbing ceiling*. Applied Acoustics **240**, 110933. DOI 10.1016/j.apacoust.2025.110933.
- [4] Abel J.S., Huang P. (2006). *A simple, robust measure of reverberation echo density*. 121st AES Convention, paper 6985.
- [5] Polack J.-D. (1993). *Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics*. Applied Acoustics **38**(2–4), 235–244. DOI 10.1016/0003-682X(93)90054-A.
- [6] Defrance G., Daudet L., Polack J.-D. (2009). *Using matching pursuit for estimating mixing time within room impulse responses*. Acta Acustica united with Acustica **95**(6), 1071–1081. DOI 10.3813/AAA.918239.
- [7] Jot J.-M. (1999). *Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces*. Multimedia Systems **7**(1), 55–69. DOI 10.1007/s005300050111.
- [8] Schroeder M.R. (1962). *Frequency-correlation functions of frequency responses in rooms*. Journal of the Acoustical Society of America **34**(12), 1819–1823. DOI 10.1121/1.1909136.
- [9] Schroeder M.R. (1965). *New method of measuring reverberation time*. Journal of the Acoustical Society of America **37**(3), 409–412. DOI 10.1121/1.1909343.
- [10] Eyring C.F. (1930). *Reverberation time in “dead” rooms*. Journal of the Acoustical Society of America **1**(2A), 217–241. DOI 10.1121/1.1915175.
- [11] Sabine W.C. (1922). *Collected Papers on Acoustics*. Harvard University Press.
- [12] Weyl H. (1912). *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen*. Mathematische Annalen **71**(4), 441–479. DOI 10.1007/BF01456804.
- [13] Allen J.B., Berkley D.A. (1979). *Image method for efficiently simulating small-room acoustics*. Journal of the Acoustical Society of America **65**(4), 943–950. DOI 10.1121/1.382599.
- [14] Hesthaven J.S., Warburton T. (2008). *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer (Texts in Applied Mathematics **54**). DOI 10.1007/978-0-387-72067-8.
- [15] Arnold D.N. (1982). *An interior penalty finite element method with discontinuous elements*. SIAM Journal on Numerical Analysis **19**(4), 742–760. DOI 10.1137/0719052.
- [16] Carpenter M.H., Kennedy C.A. (1994). *Fourth-order 2N-storage Runge–Kutta schemes*. NASA Technical Memorandum 109112.
- [17] Miki Y. (1990). *Acoustical properties of porous materials — modifications of Delany–Bazley models*. Journal of the Acoustical Society of Japan (E) **11**(1), 19–24. DOI 10.1250/ast.11.19.
- [18] Squire W., Trapp G. (1998). *Using complex variables to estimate derivatives of real functions*. SIAM Review **40**(1), 110–112. DOI 10.1137/S003614459631241X.
- [19] Martins J.R.R.A., Sturdza P., Alonso J.J. (2003). *The complex-step derivative approximation*. ACM Transactions on Mathematical Software **29**(3), 245–262. DOI 10.1145/838250.838251.
- [20] Byrd R.H., Lu P., Nocedal J., Zhu C. (1995). *A limited memory algorithm for bound constrained optimization*. SIAM Journal on Scientific Computing **16**(5), 1190–1208. DOI 10.1137/0916069.
- [21] Cremer L., Müller H.A. (1982). *Principles and Applications of Room Acoustics*, Vol. 1 (transl. T.J. Schultz). Applied Science Publishers.
- [22] Lundeby A., Vigran T.E., Bietz H., Vorländer M. (1995). *Uncertainties of measurements in room acoustics*. Acustica **81**(4), 344–355.
- [23] Tarantola A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM. DOI 10.1137/1.9780898717921.
- [24] Businger P., Golub G.H. (1965). *Linear least squares solutions by Householder transformations*. Numerische Mathematik **7**(3), 269–276. DOI 10.1007/BF01436084.
- [25] Vorländer M. (2008). *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Springer. DOI 10.1007/978-3-540-48830-9.
- [26] Kuttruff H. (2016). *Room Acoustics*. 6th ed., CRC Press. DOI 10.1201/9781315372150.
- [27] ISO 3382-1:2009. *Acoustics — Measurement of room acoustic parameters. Part 1: Performance spaces ($T_{20}/T_{30}/EDT$, C_{50}/C_{80} [Annex A.3.3], D_{50} , G)*.

- [28] ISO 9613-1:1993. *Acoustics — Attenuation of sound during propagation outdoors. Part 1: Calculation of the absorption of sound by the atmosphere.*
- [29] Di Carlo D., Tandeitnik P., Foy C., Bertin N., Deleforge A., Gannot S. (2021). *The dEchorate dataset* (RIR multi-microfono, 30 mic). Zenodo. DOI 10.5281/zenodo.4626590.
- [30] Götz G., Schlecht S.J., Pulkki V. (2021). *Motus: a dataset of higher-order Ambisonic room impulse responses and 3D models measured in a room with varying furniture.* Zenodo. DOI 10.5281/zenodo.4923187.
- [31] Götz P., Götz G., Meyer-Kahlen N., Lee K.Y., Prawda K., Habets E.A.P., Schlecht S.J. (2024). *Multi-room transition dataset.* Zenodo. DOI 10.5281/zenodo.13341566.
- [32] Figueroa-Durán A., Karakostas X., Verburg S.A., Fernández-Grande E. (2024). *Room Impulse Response Dataset: Niels Bohr Institute — Auditorium A.* Zenodo. DOI 10.5281/zenodo.14212820.
- [33] Koyama S., Nishida T., Kimura K., Abe T., Ueno N., Brunnström J. (2021). *MeshRIR: a dataset of room impulse responses on meshed grid points.* Zenodo. DOI 10.5281/zenodo.5500451.
- [34] Sundström D., Elvander F., Jakobsson A. (2025). *FLAIR: room impulse response dataset with laser-calibrated room geometry.* Zenodo. DOI 10.5281/zenodo.17037517.
- [35] Prawda K., Schlecht S.J., Välimäki V. (2022). *Dataset of impulse responses from variable acoustics room Arni at Aalto Acoustic Labs* (5342 configurazioni a pannelli commutabili). Zenodo. DOI 10.5281/zenodo.6985103.
- [36] *MP-RIR: Multi-Position Room Impulse Response dataset* (griglia 3D densa, sala a pianta trapezoidale; 68 736 RIR). Zenodo 10.5281/zenodo.11148712 (AES id 22515).
- [37] Chesworth G., Bastine A., Abhayapala T. (2024). *RSoANU: RIR dataset* (Recording Studio, Australian National University; B-format, pannelli feltro/legno). Zenodo. DOI 10.5281/zenodo.10720345.
- [38] GIR — ETH Zürich, scattering da superficie (asse *scoperto*: posizioni non pubbliche).

Le foto e i render mesh delle sale CR sono delle sale reali BRAS (credito TU Berlin). I render delle sale indipendenti usano la geometria *reale* dei rispettivi dataset (nuvola laser FLAIR, griglia MeshRIR, pianta-pannelli Arni, planimetria 8.5×9 RSoANU). Per NBI lo sketch della sala e la foto interna sono del dataset originale (credito Figueroa-Duran et al., DTU; CC-BY-4.0); il dataset non fornisce una mesh dettagliata. Le due foto RSoANU (feltro/legno) sono del dataset originale, credito ANU/RSOANU.